A TRANSCRIPTOMIC COMPARISON OF PHYSIOLOGICAL RESPONSES TO IRON AND LIGHT IN SOUTHERN OCEAN DIATOMS

Carly Moreno

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 Marine Sciences Department in the College of Arts and Sciences.

Chapel Hill
2015

Approved by:

Adrian Marchetti

Nicolas Cassar

Barbara MacGregor

ABSTRACT

Carly Moreno: A transcriptomic comparison of physiological responses to iron and light in Southern Ocean diatoms
(Under the direction of Adrian Marchetti)

Iron and light are two important abiotic factors that influence diatom growth and distribution in the Southern Ocean (SO). Through a combination of physiological and transcriptomic approaches, I have explored the molecular underpinnings of nine SO diatoms that allow for adaptation and/or acclimation to low iron and light conditions. SO diatoms used in this study ranged across five orders of magnitude in size and displayed various degrees of resistance to iron and light limitation. Specifically, we investigated the presence or absence of 22 key genes involved in iron acquisition and homeostasis, photosynthesis, and nitrogen assimilation. SO diatoms have a variety of unique resource utilization strategies coupled with gene repertoires that allow them to take advantage of ecological niches or play important roles in phytoplankton blooms. Certain diatom genes, such as B_{12}-independent methionine synthase (MetE) and flavodoxin, were found to exhibit biogeographical patterns in distribution that favor high-nutrient, low-chlorophyll regions.

To my family – for their love and support
To my stepfather – for being one of the best people in my life

ACKNOWLEDGMENTS

This work would not have been possible without the many individuals who guided and supported me along the way. First, I thank those who supplied me with seawater from the Western Antarctic Peninsula on which I performed diatom isolations: Rachel Eveleth, Yajuan Lin, and Naomi Shelton. A very special thanks to Sarah Davies (UNC) for guidance and instruction for polar diatom transcriptome analysis. I thank Jihyuk Kim and Elaine Monbureau (UNC) for help in creating Matlab figures and gene distribution maps. I thank Matt Kanke, Natalie Cohen, and Weida Gong for assistance with transcriptome and python scripting. I am grateful to Spencer Nelson and Jamal Benjamin (UNC) for culturing of diatom isolates used in this study. Last, but not least, I would like to thank my beautiful family – my grandmother, my sisters, and my two fathers – for their amazing support these past three years. With their unwavering love, I will always have the courage and motivation to get through difficult times.

This work was partially funded by the NSF Antarctic Organisms and Ecosystems Program and the Gordon and Betty Moore Foundation. Partial funding was also made possible by the Gates Millennium Scholarship Program for which I remain continuously grateful for their past and future support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBRREVIATIONS

AOX	mitochondrial alternative oxidase
ASSY	argininosuccinate synthase
Chl *a*	chlorophyll *a*
Cu	copper
CREG	cellular repressor of E1A-stimulated genes
CYTC6	cytochrome c_6
Fe	iron
FLDA	flavodoxin
FRE	ferric reductase
FTN	ferritin
FTR	iron (III) permease
HNLC	high-nutrient, low chlorophyll
ISIP	iron-starvation induced protein
LTER	Long-Term Ecological Research project
MCU	multi-copper oxidase
MetE	B_{12}-independent methionine synthase
MetH	B_{12}-dependent methionine synthase
MMETSP	Marine Microbial Eukaryote Transcriptome Sequencing Project
Mn/Fe SOD	manganese/Fe superoxide dismutase
NCP	net community productivity
NiR	nitrate reductase
NR	nitrite reductase

NRAMP	natural resistance associated macrophage protein
PCYN	plastocyanin
PET	photosynthetic electron transport
PetF	ferredoxin
POC	particulate organic carbon
RHO	rhodopsin
ROS	reactive oxygen species
SO	Southern Ocean
WAP	Western Antarctic Peninsula

THESIS OBJECTIVES

Southern Ocean ecosystems are being altered due to anthropogenic climate change, potentially resulting in shifts in the composition, diversity, and growth of primary producers, including diatoms. Diatoms are a key group of phytoplankton in the SO and serve as the base of most marine food webs in Antarctic polar waters (Saba et al. 2014). Since the Palmer long term ecological research (LTER) project was established 20 years ago along the western Antarctic Peninsula (WAP), researchers have recorded dramatic changes in the ecology and oceanography or the region, especially at lower latitudes north of Palmer station (Saba et al. 2014). The WAP is a particularly important area to study because in the past 50 years it has experienced an increase in the mean annual air temperature of 7°C and a decline in sea ice extent and duration (Ducklow et al. 2013), phytoplankton biomass (Montes-Hugo et al. 2009), krill (Saba et al. 2014), and Adelie penguin populations (Ducklow et al 2007).

Within the Southern Ocean, the two main environmental variables that have been identified as influencing phytoplankton growth are iron and light availability. Iron is supplied naturally to the SO in by aeolian deposition, melting of sea ice, resuspension of sediments, and deep winter mixing; however, rapid biological uptake and particle scavenging create a high-nutrient, low chlorophyll (HNLC) condition in which primary productivity is limited by iron availability (Cassar et al. 2007; Edwards and Sedgewick 2001; Death et al. 2014). The mixed-layer depth, which influences mean light levels, is an important driver of phytoplankton growth and the occurrence of blooms. Deep winter mixed layers (>100m) can result in phytoplankters, like diatoms, experiencing considerable periods of light limitation (Sallée et al. 2010). As warming climate changes the dynamics of light and iron availability in the SO, there could be both positive and negative effects on diatom growth and primary productivity. For example, the

warming and freshening of the surface layer would enhance vertical stratification, resulting in elevated mean irradiances; however, increased stratification might lead to more seasonal nutrient limitation. The interactive effects of these environmental factors are still missing from models of primary productivity.

Although there have been recent advances in our understanding of how iron and light influence the physiology of SO diatoms, such as the possible lack of an interactive effect between iron and light (Strzepek et al., 2012), few studies have investigated the molecular underpinnings of the distinct physiological responses of polar diatoms to iron and light limitation. The main objective of this study is to identify the genetic and cellular mechanisms that may mediate a specific adaptation or acclimation to low iron and light in SO diatoms using transcriptomics. Transcriptomics is the study of an organism's mRNA transcripts, or expressed genes, under specific circumstances using high-throughput sequencing. The benefits of analyzing transcriptomes as opposed to genomes is they relate directly to the physiological status of the cell and they are much easier to sequence than genomes. We also utilized publically available transcriptomes from the Marine Microbial Eukaryote Transcriptome Project (Keeling et al. 2013), in conjunction with transcriptomes of SO diatoms I recently isolated, to understand the molecular bases of physiological adaptations to low iron and light conditions.

From Western Antarctic waters, I have isolated three species of polar pennate diatoms (*Navicula lanceolata*, *Fragilariopsis cylindrus* and *Pseudo-nitzschia subcurvata*) and six centric diatoms (*Thalassiosira antarctica, Actinochilus actinocyclus, Proboscia alata., Chaetoceros debilis, Chaetoceros rostratus*, and *Eucampia antarctica*) and performed a detailed examination of their growth under varying iron and light conditions, along with transcriptomic analysis. We are particularly interested in protein-encoding genes and metabolic pathways that have

previously been identified as being affected by iron and or light. Ultimately, by investigating gene expression repertoires in conjunction with the effects of varying iron and light conditions on growth, we can gain a better understanding of how differences in cellular mechanisms can have far-reaching effects on diatom distribution and bloom occurrence, as well as processes that influence net community productivity and global biogeochemistry.

CHAPTER 1: THE PHYSIOLOGICAL EFFECTS OF IRON AND LIGHT LIMITATION ON SOUTHERN OCEAN DIATOMS

Introduction

Diatoms are key organisms in determining net community production, and the overall health of the Antarctic ecosystem (Ducklow, et al., 2007). Understanding the physiological adaptation to changing light and iron regimes of polar diatoms along the western Antarctic Peninsula is crucial to defining the role they have in sustaining Antarctic food web structure and to what extent they can potentially affect biogeochemical cycles in the Southern Ocean. This chapter will focus on the physiological adaptations of diatoms to limiting iron and light conditions in terms of biophysical properties such as growth rate, photosynthetic efficiency and biovolume, while providing essential background information.

Oceanography of the Southern Ocean: primary production, iron and light, and the rapidly changing Western Antarctic Peninsula

The Southern Ocean encircling Antarctica is an important driver of ocean circulation and climate, biogeochemical cycling, and polar ecosystem structure and productivity (Sabine et al., 2004, Gruber et al., 2009, Smith and Comiso, 2007). Although the Southern Ocean (SO) is a relatively small ocean basin, it serves as the main intersection between the Atlantic, Pacific, and Indian Oceans as well as between surface and deep waters. The principle current, the Antarctic Circumpolar Currrent (ACC), is comprised of complex oceanographic zones and fronts with varying physical and chemical signatures (Talley et al., 2011). Coastal regions along the

Antarctic continent are particularly important as sites of upwelled waters rich in nutrients and CO_2 (Arrigo et al., 2008). When SO waters subsequently subduct to form subantarctic intermediate and mode waters, they can ultimately fuel productivity in low latitudes when they return to the surface (Sarmiento et al., 2004).

Primary production in the SO is highly variable spatially and temporally. Intense phytoplankton blooms can occur when environmental conditions are favorable (Arrigo et al., 2008). Production "hot spots' include polynyas (regions of open water surrounded by sea ice) (Tremblay and Smith, 2007), the seasonal ice zone (Smith and Nelson, 1986), and narrow continental shelves (Sweeney, 2003). The phytoplankton communities in these regions, mainly comprised of diatoms and prymnesiophytes such as *Phaeocystis*, are responsible for the large quantity of dissolved inorganic carbon (DIC) fixed into organic matter by photosynthesis (Arrigo et al., 2010). This organic matter can be recycled in surface waters, remineralized at depth, or transported to deep ocean sediments via the biological pump, the efficiency of which determines how much CO_2 the ocean can draw down (Ducklow et al., 2001). Phytoplankton, especially diatoms, also serve as the base of the polar marine food web, and support a wide variety of Antarctic fauna like krill, fish, penguins, and whales (Saba et al., 2014). However, despite plentiful nitrate and phosphate in the surface waters throughout the Southern Ocean, primary productivity is patchy at best (Moore and Abbott, 2000).

In the Southern Ocean, abundant nutrients remain unutilized in the euphotic zone (Boyd, 2002; Falkowski 1998). It is a high-nutrient low-chlorophyll (HNLC) region in which phytoplankton growth and biomass are limited by the availability of the micronutrient iron and low light (Boyd, 2002). The high nutrient concentrations correspond to lower than expected values of chlorophyll *a*, and consequently lower values of primary production despite various

sources of iron (Sokolov and Rintoul, 2007). Iron is naturally added to the SO by aeolian deposition, sea ice melt water, upwelling, and wind driven mixing over shallower coastal regions (Cassar et al., 2007; Edwards and Sedgewick 2001; Death et al., 2014). Iron limitation is supported by findings from artificial iron fertilization experiments, where large amounts of iron are added to surface waters, leading to massive phytoplankton blooms (Boyd et al., 2000). Thus, if more iron were available to phytoplankton, it is predicted that primary production would increase and blooms would be more common (de Baar et al., 2005).

All organisms have a requirement for iron because it is an essential component of enzymes required for photosynthesis, respiration, reduction of oxidized nitrogen and sulfur compounds, and nitrogen fixation (Raven, 1999). Iron limitation in diatoms results in decreased growth rates, chlorophyll-*a* contents, and photosynthetic efficiency (Timmermans et al., 2010; Arrigo et al., 2010; Strzepek et al., 2012, Alderkamp et al., 2012). Iron limitation can also cause diatoms to change their silicon-to-nitrogen consumption ratios. During periods of low iron input, diatoms take up more silicon relative to nitrogen from seawater to incorporate in their frustules, leaving surface waters depleted in silicic acid (Takeda, 1998). One proposed mechanism for increased silica precipitation on frustules of iron-limited diatoms is the upregulation of spermine synthase, a key enzyme in the polyamine synthesis pathway, which is associated with precipitating silica on the cell wall (Nunn et al., 2013). This regulation, along with decreasing cell size, could have evolved to allow faster sinking to the deep ocean possibly resulting in reduced bacterial colonization and degradation (Marchetti & Cassar, 2009; Nunn et al., 2013). In this way diatoms essentially control the silicon cycle and indirectly the global carbon cycle (Assmy et al., 2013).

Light is another important factor in regulating productivity in the surface mixed layer. Because phytoplankton drift freely in the oceans, they are sensitive to dynamic physical ocean processes such as stratification, changes in mixed layer depth, and sea ice extent and duration (Deutsch, et al., 2009). In the SO, light availability varies daily and seasonally. Because the mixed layer can also be very deep, phytoplankton may spend considerable periods of time below the euphotic zone. This means diatoms must be able to exploit highly variable light environments in order to survive, or lie dormant when light is absent for long periods of time. It has been proposed that light is the major limiting factor during the austral spring and autumn (Smith and Gordon, 1997), whereas Fe limits primary production during the summer (van Oijen, et al., 2004). However, co-limitation of iron and light is also possible at different times and regions (Tremblay and Smith, 2007; Boyd et al., 2002; Sedwick et al., 2007).

The WAP has been studied intensively for the past two decades by the Palmer Long-Term Ecological Research Project (LTER; Fig. 1). This is an area of rapidly changing climate, declining sea ice coverage, and deglaciation (Schofield et al., 2010). It is also experiencing the fastest increase in mean atmospheric temperature in the world, with an increase of 7°C since 1951 (Meredith, 2005). Because biological systems in the WAP are sensitive to ice seasonality, the recent warming and decrease in sea ice has resulted in significant decreases in phytoplankton biomass, including shifts from large diatoms to small flagellated cryptophytes (Montes-Hugo,

2009), as well as decreases in primary production, krill abundance, and the Adelie penguin populations that are found north of Palmer Station (Saba et al., 2014).

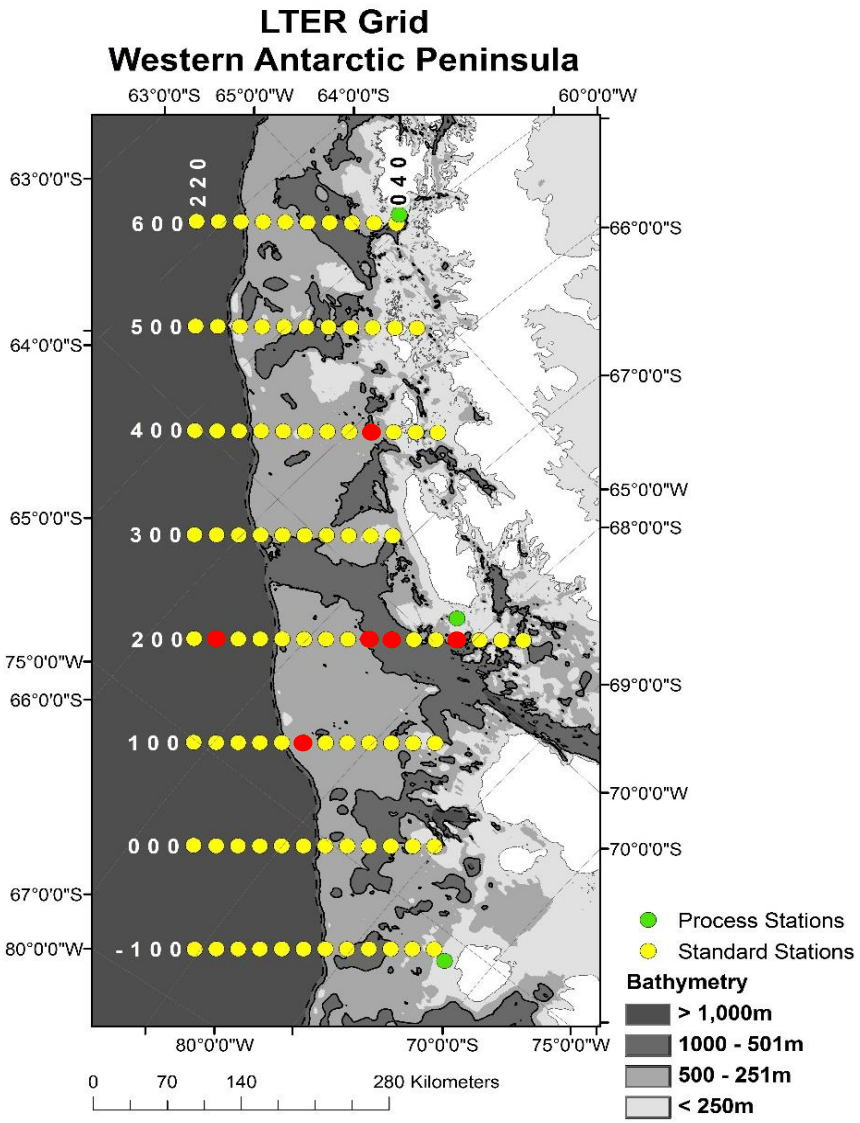


Figure 1.1. Map of the PalmerLTER sampling grid along the WAP. Locations of isolated diatoms are in red. Yellow stations are 20 km apart and increase in number right to left. Each line is 100 km apart from 63S to 80S. *C. cf* sociales and *C. rostratus* were isolated from 200.-020 in Marguerite Bay; *T. antarctica* was isolated at 200.180 an off-shelf region. In the mid-shelf regions, *E. antarctic* was isolated from 200.020; *A. actinochilus* was isolated from 400.040; *P. alata* and *P. subcurvata* were isolated from 200.040; *N. lanceolata* was isolated from 100.040.

Effects of iron and light on polar diatom physiology and metabolism

It is well known that iron and light are the primary environmental bottom-up controls that influence primary productivity and diatom abundances in the Southern Ocean (Boyd, 2002; Strzepek et al., 2012). Recent investigations have revealed multiple physiological mechanisms for increasing iron acquisition, efficiently allocating intracellular iron, and decreasing iron requirements in polar diatoms in order to adapt or acclimate to low iron environments. These include reducing cellular iron-to-carbon ratios (Strzepek et al., 2012), substituting iron-containing redox catalysts with proteins that can use other metal cofactors (Peers and Price, 2006), activating a high-affinity iron uptake system (Raven, 1990), performing luxury uptake of iron (Marchetti et al., 2009), and increasing surface area-to-volume (SA:V) ratios (Sunda & Huntsman, 1995). Increasing SA:V ratios are useful under nutrient deficiency because it decreases the diffusive boundary layer and increases iron uptake rates.

Intracellular iron requirements, or quotas, are measured as Fe:C ratios and are a measure of the iron demands of the cell. Fe requirements of a diatom can be inferred by either measuring the concentration of Fe in the medium that supports growth and calculating a half-saturation constant in relation to growth rates (K_μ), or by measuring their intracellular Fe content. Diatoms isolated from the SO have particularly low iron contents and higher iron use efficiencies, or the amount of C fixed per mole of Fe per day (Marchetti and Maldonado, 2016). Strzepek et al. (2011) found that intracellular iron concentrations normalized to either carbon or cell volume were extremely low for SO diatoms. For example, *Proboscia inermis* had the lowest cellular Fe:C ratio recorded (0.4 µmol:mol with a relative growth rate of 0.6 d^{-1} in relation to the high iron treatment), while the model temperate diatom, *T. oceanica*, had an Fe:C of 0.58 µmol:mol and relative growth rate of 0.35 d^{-1}. Both the large pennate and centric diatoms in their study had

iron requirements that were at least twofold lower than other examined oceanic species. Although the diatoms used in their study were large (>20 um in some cases), and thus, had lower cell SA:V ratios, they appeared to compensate for this by greatly reducing their iron requirements.

Raven et al. (1999) theoretically calculated that most of the iron required by a cell is allocated to catalytic proteins used in photosynthetic electron transport, specifically the photosynthetic units (PSUs). Iron is used in the major photosynthetic complexes including: photosystem I (PS1), photosystem II (PSII), and the cytochrome b_6f complex. The iron requirements of the cell can be reduced by changing the ratio of each complex in response to growth, irradiance, and iron availability (Raven et al., 1999; Sunda and Huntsman 1997). For example, because PSII requires less iron than PSI and cytochrome b_6f, oceanic diatoms can become more enriched in PSII relative to the other complexes (Strzepek and Harrison, 2004; Green et al., 1991).

Because the photosynthetic architecture of diatoms requires 50-80% of the available intracellular iron, photoacclimation and iron interactions are intimately linked (Raven 2013). Laboratory studies have documented distinct physiological responses to low irradiances when cells are grown in iron-limiting conditions such as increasing the number or size of photosynthetic subunits, altering light harvesting elements, and decreasing Chl *a* contents. (Timmermans et al., 2010; Arrigo et al., 2010; Strzepek et al., 2012, Alderkamp et al., 2012). Although some of these strategies have higher iron requirements, polar diatoms appear to have the ability to maintain low cellular iron demands when grown under low irradiances. Strzepek et al. (2012) hypothesized that polar diatoms may increase the size rather than the number of PSU when grown under low irradiances. PSU's are composed of the iron-rich PSI, PSII and the

cytochrome b_6f complex and they perform photosynthetic electron transport (PET). By increasing the size of the PSU's, they could lower their iron requirements and grow at a higher relative growth rate. This strategy would contradict previous research regarding the antagonistic relationship between light and iron, in which photoacclimation strategies to low light cause cellular iron requirements to increase (Raven et al. 1990).

The interaction of iron and light in the SO is particularly important for diatoms because they can experience environmental conditions both low in iron availability, and with variable light levels due to deep mixed layers. For example, *F. cylindrus*, a common bloom former in the Southern Ocean, has been shown to thrive in shallow mixed layers where light levels can be high, and under sea ice where light levels just below the ice can be as low as 0.1% of surface irradiance. This is because at high light levels, *F. cylindrus* can maintain low concentrations of photosynthetic pigments and high concentrations of photoprotective pigments, while in low light regimes, this species can absorb green light between 500-575nm, which is found underneath sea ice (Arrigo et al., 2010). This could be due to *F. cylindrus* containing the chromophore retinal, a pigment used in rhodopsin (Marchetti et al., 2015). Rhodopsin (*RHO*) is a light-driven proton pump that could potentially act as an alternative to photosynthesis in low iron conditions. While the chromophore retinal requires a small amount of iron, the rhodopsin protein uses a much smaller amount than the photosynthetic apparatus. This could indicate that diatoms may lower their iron requirements through activating the rhodopsin gene. *Pseudo-nitzschia granii* has been shown to increase transcripts of *RHO* under low iron conditions (Marchetti et al. 2015). However, it is still unclear if the rhodopsin in diatoms is producing ATP as an alternative to photosynthesis.

Understanding the physiological adaptation to changing light and iron availability of polar diatoms along the WAP is crucial to defining their role in the Antarctic food web and to what extent they can potentially affect biogeochemical cycles in the SO. The goal of this chapter was to understand how iron and light limitation affects the growth rate, photosynthetic efficiency, and biovolume of polar diatoms. We sought to determine if polar diatoms grown under iron and light limitation were indeed co-limited by iron and light, or if they exhibited other types of resource limitation. Ultimately, this research will expand our existing knowledge of how iron and light interactions affect polar and temperate species (Strzepek et al. 2012; Sunda & Huntsman et al. 1997)

Materials and methods

Diatom isolation and identification - Eight species of diatoms were isolated from the Western Antarctic Peninsula along the PalmerLTER sampling grid in March 2014 (see Fig. 1.1 for locations). Isolations were performed using an Olympus CKX41 inverted microscope by single cell isolation with micropipette (Anderson 2005). The identification of diatom 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). Amplification of the nuclear 18S rDNA region was achieved with standard PCR protocols using eukaryotic-specific, universal 18S forward and reverse primers. Primers were obtained from Medlin et al. (1982) and are as follows: 18AF 5'-AACCTGGTTGATCCTGCCAGT-3' and 18BR 5'- TGATCCTTCTGCAGGTTCACCTAC -3'. The length of the region amplified is approximately 1600 base pairs (bp). *Pseudo-nitzschia* species are difficult to identify by 18S rRNA gene sequence, therefore, additional support of the taxonomic identification of *P.*

subcurvata was provided through sequencing of the 18S-ITS1-5.8S regions. Amplification of this region was performed with 18SF-euk and 5.8SR_euk primers according to Hubbard et al. (2008), which are as follows: 18SF-euk 5'-CTTATCATTTAGAGGAAGGTGAAGTCG-3' and 5.8SR-euk 5'-CTGCGTTCTTCATCGTTGTGG-3'. PCR products were purified using either QIAquick PCR Purification Kit (Qiagen) or ExoSAP-IT (Affymetrix) and sequenced by Sanger DNA sequencing (Genewiz). Sequences were edited using Geneious Pro software (http://www.geneious.com, Kearse et al., 2012) and BLASTn sequence homology searches were performed against the NCBI nucleotide non-redundant (nr) database to determine species with a cutoff identity of 98%. In the case of *P. subcurvata*, a specific polymorphic fragment of the ITS1 region was further investigated by aligning PnAll primers (Hubbard et al., 2008) to the 18S-ITS1-5.8S region, resulting in a fragment of 171 bp. Hubbard et al. (2008) found the length of this fragment to be unique to a given *Pseudo-nitzschia* species despite very similar sequences. My analysis of this fragment resulted in the highest sequence homology and length similarity, an ITS1 fragment of 171 bp, to *P. subcurvata* (Strain: 1-F, Accession number: DQ329205).

Diatom phylogenetic analysis was performed with Geneious Pro and included 71 additional diatom 18S rDNA sequences from publically available genomes and transcriptomes, including all those in the Marine MicroEukaryote Transcriptome Sequencing Project (MMETSP) database. Diatom sequences were trimmed to the same length and aligned with MUSCLE (Edgar 2004). A phylogenetic tree was created in Mega with the Maximum-likelihood method of tree reconstruction, the Jukes-Cantor genetic distance model (Jukes and Cantor 1969), and 100 bootstrap replicates.

Growth conditions, physiological characteristics and biovolumes - Isolates were maintained at 4°C in constant irradiance at either 10 µmol photons m^{-2} s^{-1} (low light) or 90 µmol

photons m^{-2} s^{-1} (standard light) and with media containing two iron concentrations. Cultures were grown in synthetic seawater medium, AQUIL, enriched with filter sterilized vitamin and trace metal (buffered with 100 μmol L^{-1} EDTA) solutions as well as 300 μmol L^{-1} nitrate, 200 μmol L^{-1} silicate and 20 μmol L^{-1} phosphate. Premixed Fe-EDTA (1:1) was added separately for total Fe concentrations of either 1370 nmol L^{-1} (pFe 19) or 3.1 nmol L^{-1} (pFe 21.7) to achieve high iron and low iron media, respectively. All media preparation and subsampling were performed under a positive-pressure, trace metal clean laminar flow hood. Cultures were grown in acid-washed 28 mL polycarbonate centrifuge tubes (Nalgene) and maintained in exponential phase by dilution. Specific growth rates were calculated from the linear regression of the natural log of *in vivo* chlorophyll *a* fluorescence using a Turner 10-AU fluorometer (Brand et al. 1981).

Photophysiological parameters were measured with a Fluorescence Induction Relaxation System (FIRe) (Satatlantic). Samples were dark acclimated for at least 10 minutes and measurements were taken of each culture for photosynthetic efficiency (F_v:F_m), and functional absorption cross-section of PSII, (σ_{PSII}). FIRe parameters were set to measure single turnover flash of PSII reaction centers (single closure event) with a sample delay of 100, and a total of 50 samples (Gorbunov and Falkowski, 2004).

Cell dimensions and biovolume measurements - To estimate biovolumes (V) of each diatom species, frustules were viewed using an Olympus BX61 Upright Wide Field Microscope with the differential interference contrast (DIC) imaging mode and a 60X/1.42 Oil PlanApo N objective lens. Valve apical length (AL), transapical width (TW), and pervalvar height (PH) dimensions were estimated with Scion Image software (Informer Technologies, Inc.). Diatom biovolumes were calculated according to Hillebrand et al. (1999) and were estimated by

modeling the cells after geometric shapes. Pennate diatoms were modeled after elliptic prisms and centric diatoms were modeled after cylinders. The equations used are:

$$\text{Cylinder:} \quad V = \frac{\pi}{4} \cdot \mathrm{h} \cdot \mathrm{d}^2$$

$$\text{Prism on elliptic base:} \quad V = \frac{\pi}{4} \cdot AL \cdot TW \cdot PH$$

For all species except *T. antarctica*, at least four cells were imaged and measured. *T. antarctica* had a sample size n = 1.

Data analysis - Statistical analyses of growth rates and photophysiological data were performed with SigmaPlot 12.5 (SysStat Software Inc.). To test for significant differences between treatments, Two-Way Analysis of Variance (ANOVA) was performed with a significance level set to $p<0.05$. ANOVA also tests for normality using Shapiro-Wilks and Equal Variance tests. Because ANOVA does not test all interactions, an unpaired t-test was performed between –FeLL and +FeSL for μ, F_v:F_m, or σ_{PSII}. All tests passed the Shapiro-Wilks Normality tests unless otherwise stated, in which case p-values are representative of the Mann-Whitney Rank Sum test. Post-hoc Tukey tests were also performed in order to determine which treatments differed significantly ($p < 0.05$). Plots were created using Matlab software.

Results

Diatom isolates – Isolates collected for this study include two raphid pennate diatoms: *Navicula lanceolata* and *Pseudo-nitzschia subcurvata*; four radial centric diatoms: *Actinocyclus actinochilus, Eucampia antarctica, Proboscia alata*, and *Thalassiosira antarctica*; and two bi-multipolar centrics: *Chaetoceros cf sociales* and *Chaetoceros rostratus*. Our diatom collection does not include an araphid pennate. The diatoms were isolated from the following stations: 200.-020 in Marguerite Bay; 200.180, an off-shelf region; and 400.040, 200.040, and 100.040 in

mid-shelf regions (Table 1). Despite originating from a relatively small geographic region off the Western Antarctic Peninsula, these diatoms represent diverse lineages. Full-length 18S rDNA sequences were successfully obtained for eight of the isolates which were aligned with 76 18S rDNA sequences obtained from the MMETSP diatom database (Fig. 1.2). Of the diatoms sequenced as part of the MMETSP, four of our polar isolate transcriptomes are unique, as *A. actinochilus*, *N. lanceolata*, *C. rostratus*, and *C. sociales* have not been previously sequenced. Transcriptome sequencing of *E. antarctica* was unsuccessful within our first round of transcriptome sequencing; however, the MMETSP database also contains a sequenced strain of this species.

Table 1.1. Isolates with strain designation, sampling location along the PalmerLTER, and accession numbers of best BLAST hit in NCBI Genbank.

Polar diatoms	Strain ID	Date Isolated	PalmerLTER Station	Latitude	Longitude	GenBank best hit	Isolator
Thalassiosira antarctica	UNC1401	Mar-14	200.180	-66.5795	-72.7391	DQ514874.1	C. Moreno
Eucampia antarctica	UNC1402	Mar-14	200.020	-67.6418	-70.2688	X85389.1	C. Moreno
Actinocyclus actinochilus	UNC1403	Mar-14	400.040	-66.2540	-67.3366	AY485506.1	C. Moreno
Navicula lanceolata	UNC1404	Mar-14	100.040	-68.1121	-72.3461	KC771158.1	C. Moreno
Proboscia alata	UNC1405	Mar-14	200.040	-67.5111	-70.5890	AY485525.1	C. Moreno
Pseudo-nitzschia subcurvata	UNC1406	Mar-14	200.040	-67.5111	-70.5890	AY485490.1	C. Moreno
Chaetoceros cf *sociales*	UNC1407	Mar-14	200.-020	-67.2956	-69.6654	HM581778.1	C. Moreno
Chaetoceros rostratus	UNC1408	Mar-14	200.-020	-67.2956	-69.6654	X85391.1	C. Moreno
Fragilariopsis cylindrus	UNC1301	Mar-13					C. Moreno

Figure 1.2. Phylogenetic tree generated through Maximum-likelihood analysis of 18S rDNA sequences from MMETSP diatoms and new isolates from the WAP denoted with a yellow star. Centric diatoms are highlighted in green (bi-polar) and blue (radial) and pennate diatoms are highlighted in purple (raphid) and red (araphid). Node labels show percent consensus support for tree arrangement, and branch lengths represent divergence between two nodes in the tree (substitutions per site of sequence alignment). The outgroup of the rooted tree is *Bolidomonas pacifica* RCC208, a eukaryotic, picoplanktonic heterokont.

Growth parameters – Growth rates are expressed as specific growth rate (μ; d^{-1}) or as relative growth rate in relation to the maximum specific growth rate (μ_{max}) determined in iron-replete and standard light conditions (μ:μ_{max}). Other physiological parameters such as F_v:F_m and σ_{PSII} are expressed in specific terms. Nutrient limitation of iron and light, or an additive interaction of both these resources, was diagnosed in an isolate if the mean relative growth rate and F_v:F_m were significantly reduced ($p < 0.05$).

Among the isolates grown in iron-replete, standard light conditions (+FeSL), μ_{max} ranged from 0.25 d^{-1} (*A. actinochilus*) to 0.72 d^{-1} (*P. subcurvata*; Fig. 1.3A); F_v:F_m values were between 0.473 (*E. antarctica*) and 0.545 (*C. sociales*; Fig. 1.3B); and σ_{PSII} ranged from 222 for *T. antarctica* to 511 in *F. cylindrus* (Fig. 1.3C).

The relative growth rates of all isolates in low iron cultures (-FeSL) were reduced by 12 – 64%. Except for *P. alata*, all diatoms experienced some degree of iron limitation, as their μ:μ_{max} were significantly reduced (ANOVA, $p<0.05$) (Fig. 1.3D). F_v:F_m is known to decrease in iron-limited cells; however, in our study this measure was less responsive to iron stress than was growth rate. However, for *N. lanceolata*, *T. antarctica*, and *C. cf sociales*, F_v:F_m was reduced 19 – 50% (ANOVA, $p<0.05$), indicating these diatoms were experiencing moderate to severe nutrient stress (Fig. 1.3B). In addition, σ_{PSII} was observed to significantly increase (ANOVA, $p<0.05$) under low iron conditions, which is also characteristic of Fe limitation in diatoms (Fig. 1.3C). While a decrease in F_v:F_m was not observed in *F. cylindrus* and *E. antarctica*, an increase in σ_{PSII} was observed.

Low irradiance (+FeLL) physiologically stressed all diatoms, except *F. cylindrus*, resulting in reductions in growth rates (Fig. 1.4A). For the remaining species, μ:μ_{max} was observed to decrease from 12 – 60% ($p < 0.05$). F_v:F_m did not decrease, and in the case of *A.*

actinochilus, P. subcurvata, C. rostratus, and *E. antarctica*, F_v:F_m significantly increased from 7 – 25% ($p < 0.05$; Fig. 1.4B). σ_{PSII} did not significantly change from +FeSL treatments (Fig. 1.4C).

When cells were grown under combined low iron and low light growth conditions (-FeLL), all isolates experienced significant decreases in μ:μ_{max} of at least 35% ($p < 0.05$; Fig. 1.5A) relative to +FeSL. However, decreases in F_v:F_m were only evident in *F. cylindrus* (Mann-Whitney, $p = 0.01$), *P. alata*, and *N. lanceolata* (Fig. 1.5B). An increase in σ_{PSII} of ~30% ($p<0.05$) was observed for *F. cylindrus, P. alata* (Mann-Whitney; $p = 0.024$), and *C. sociales* (Fig. 1.5C). Steady-state measurements of F_v:F_m and σ_{PSII} in *T. antarctica* and *N. lanceolata* under -FeLL conditions were not obtainable due to cessation of growth in these isolates.

We also determined interactive effects of low iron and light in our isolates (Fig. 1.6 A-I). Diatoms exhibiting no interactive effects include *C. cf sociales, P. subcurvata, A. actinochilus, C. rostratus, N. lanceolata,* and *T. antarctica*. Growth rates in these –FeLL cultures were not significantly different from either –FeSL or +FeLL treatments. Diatoms that only appeared to be limited by light were *E. antarctica* and *C. rostratus*. *C. rostratus* also showed a small degree of iron limitation, but only in light saturating conditions. *P. alata* and *F. cylindrus* both exhibited additive interaction between iron and light in our treatments.

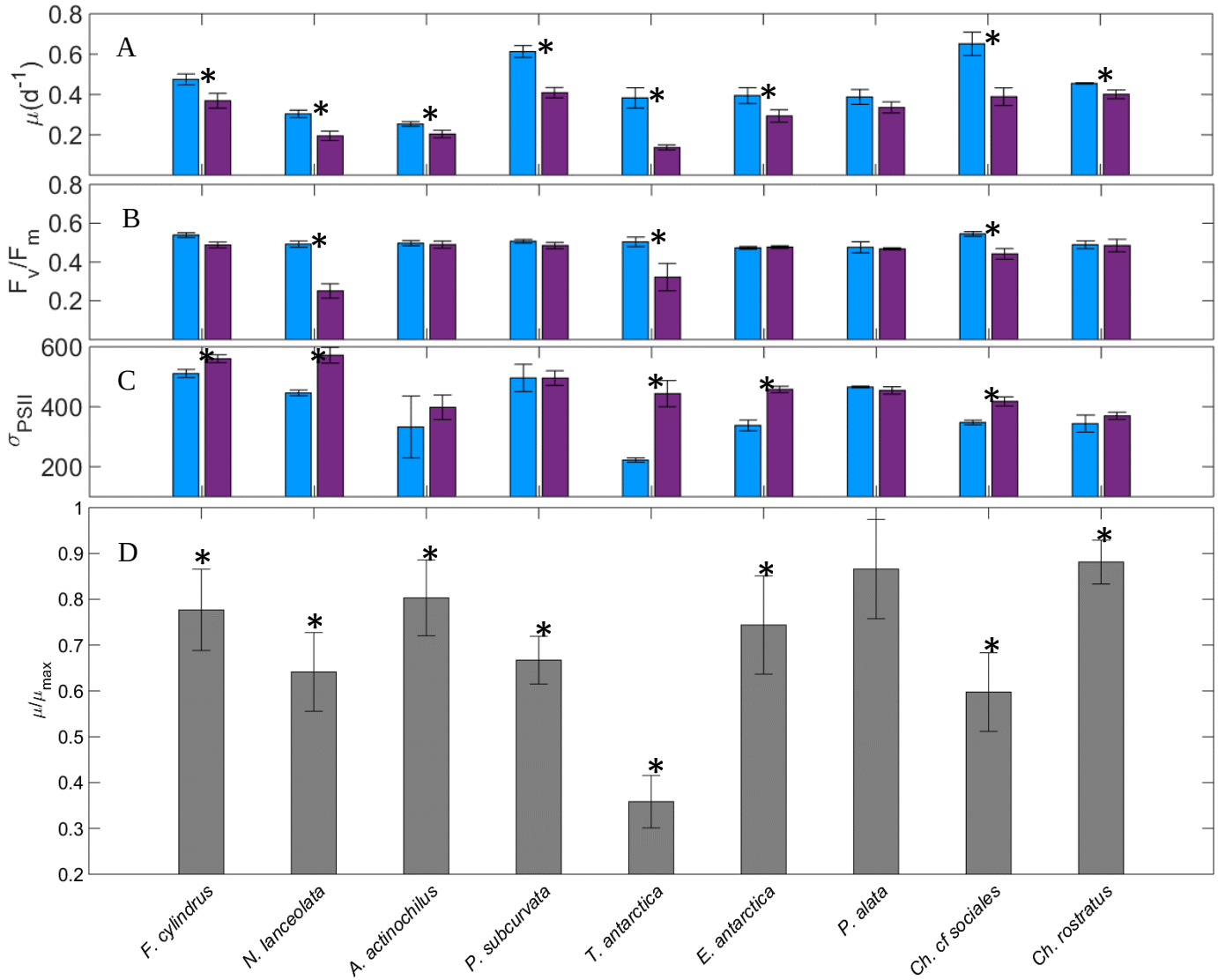


Figure 1.3. A-D. Iron effect. Comparison of specific growth rate (A), F_v:F_m (B), and σ_{PSII} (C) in +FeSL and –FeSL conditions. Relative values of growth rates for each isolate are shown in D. Error bars represent the standard error of the mean (n > 3 for all +FeSL and –FeSL treatments). Isolates labeled with an asterisk indicate the –FeSL treatment is significantly different from the +FeSL treatment (Two-Way ANOVA, $p < 0.05$).

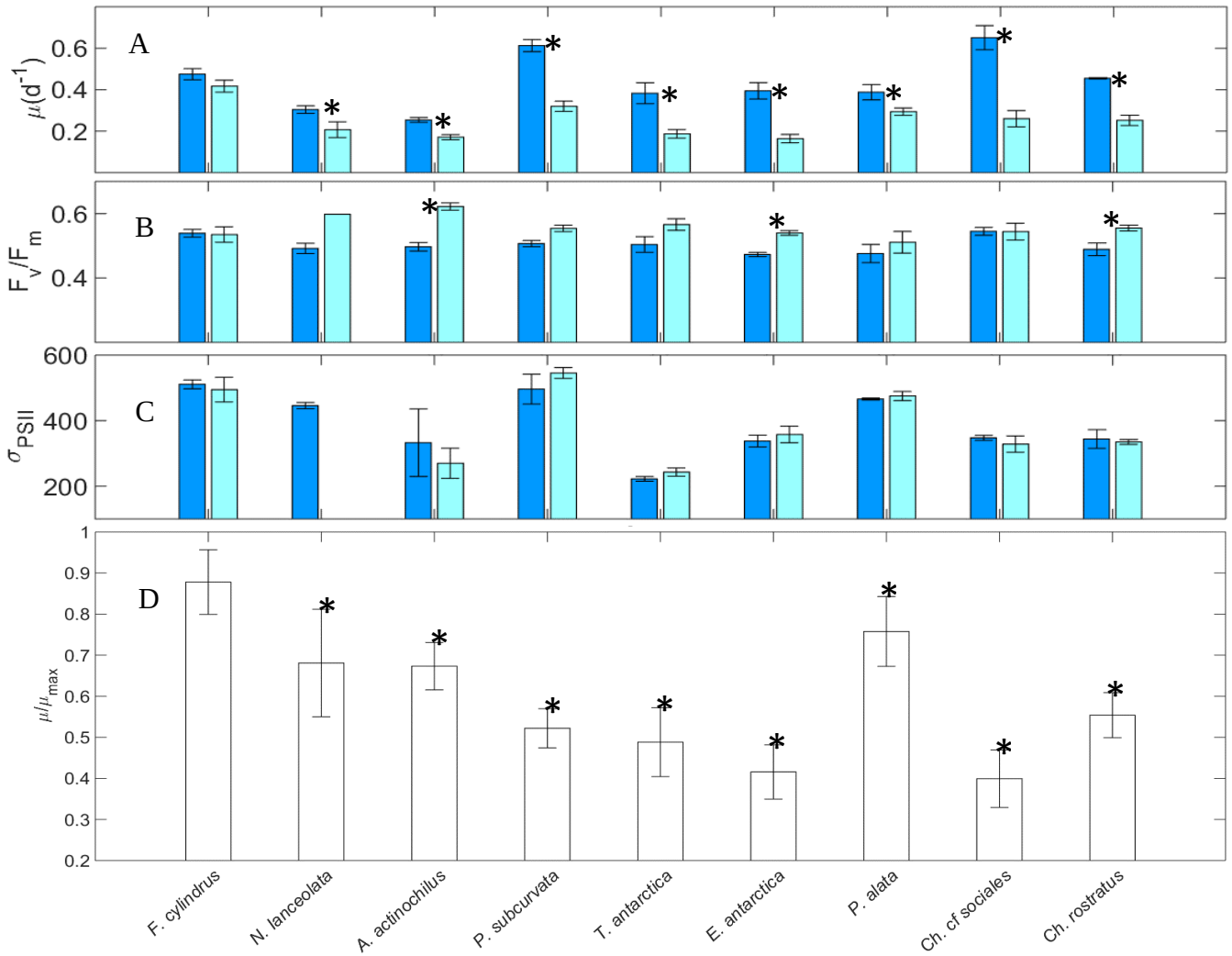


Figure 1.4. A-D. Light effect. Comparison of specific growth rate (A), F_v:F_m (B), and σ_{PSII} (C) in +FeSL and +FeLL conditions. Relative values of growth rates for each isolate are shown in D. Error bars represent the standard error of the mean (n > 3 for +FeSL; n > 3 +FeLL treatments, except for *N. lanceolata* F_v:F_m and σ_{PSII} n = 1). Treatments labeled with an asterisk are significantly different from the +FeSL treatment (Two-Way ANOVA, $p < 0.05$).

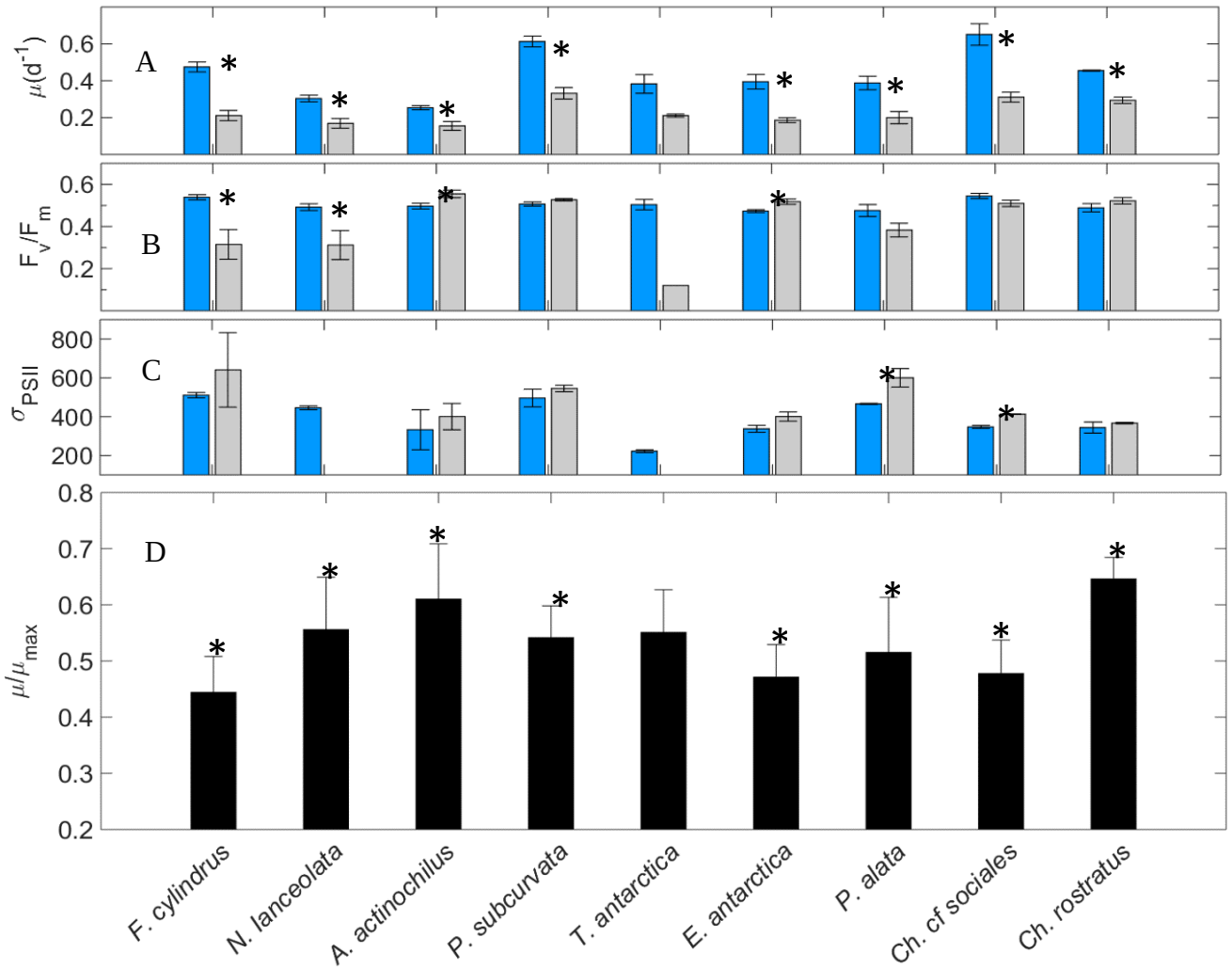


Figure 1.5. A-D. Combined Fe and light effect. Comparison of specific growth rate (A), F_v:F_m (B), and σ_{PSII} (C) in +FeSL and –FeLL conditions. Relative values of growth rate for each isolate are shown in D. Error bars represent the standard error of the mean (n > 3 for +FeSL; n > 3 -FeLL treatments, except for *T. antarctica* F_v:F_m and σ_{PSII} n = 1; *N. lanceolata* F_v:F_m and σ_{PSII} n = 2). Treatments labeled with an asterisk are significantly different from the +FeSL treatment (Student's t-test, $p<0.05$).

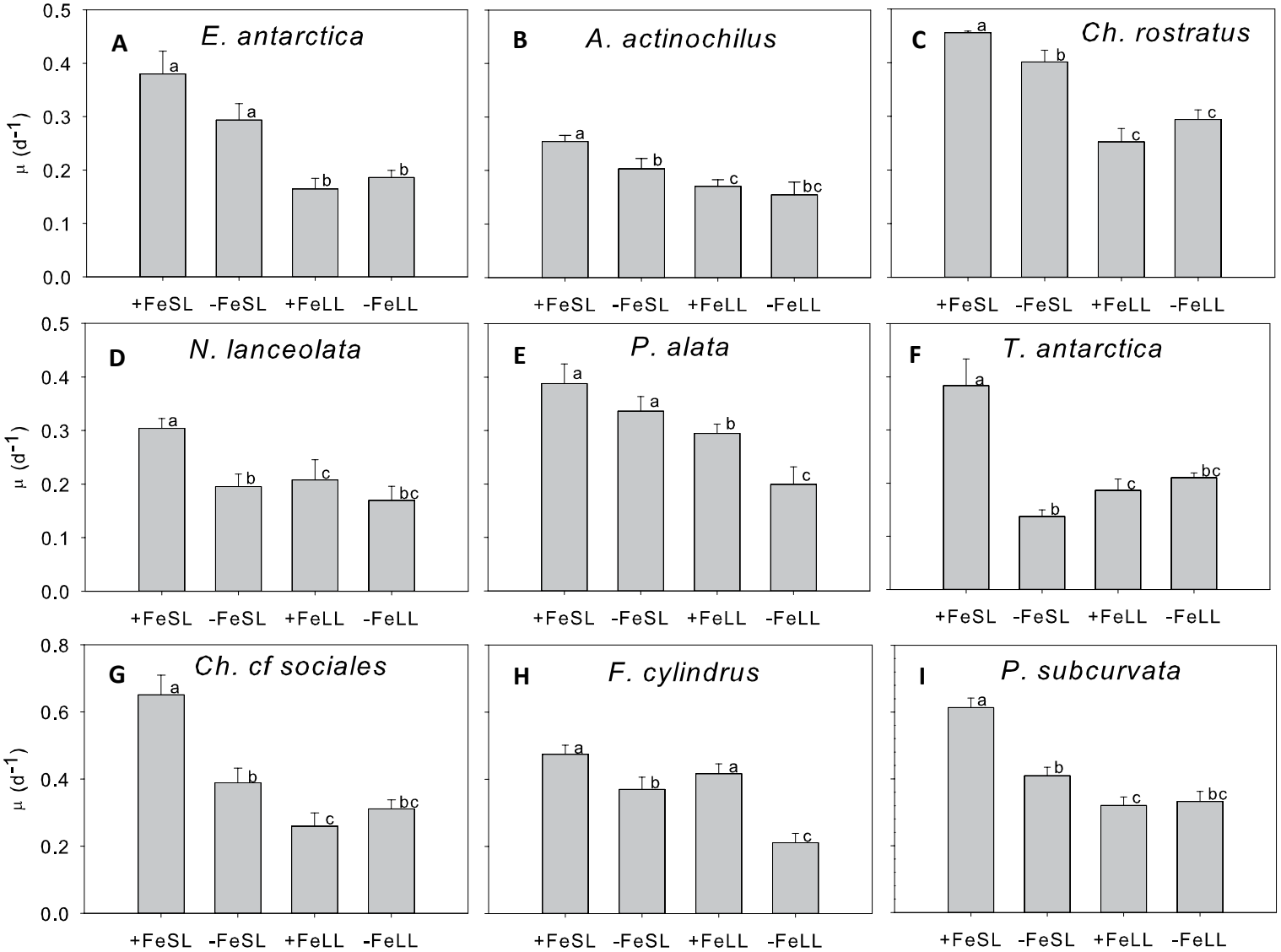


Figure 1.6. A-I. Comparison of growth rates of each isolate in each treatment. Letters denote significant differences in growth rates (ANOVA, $p < 0.05$.; Student's t-test or Mann-Whitney, $p < 0.05$). Error bars represent one standard error of the mean.

Cell dimensions and biovolume – The diatoms isolated for this study represent a wide range of cell sizes and resulting biovolumes. Biovolumes for the isolates ranged by five orders of magnitude from 50 µm^3 for *F. cylindrus* to >10^5 µm^3 for *T. antarctica* (Fig. 1.7). To examine how cell size affects a diatom's ability to adapt to low iron and/or light conditions we measured diatom biovolume and compared it µ:µ$_{max}$ under low iron (Fe effect) and low light (light

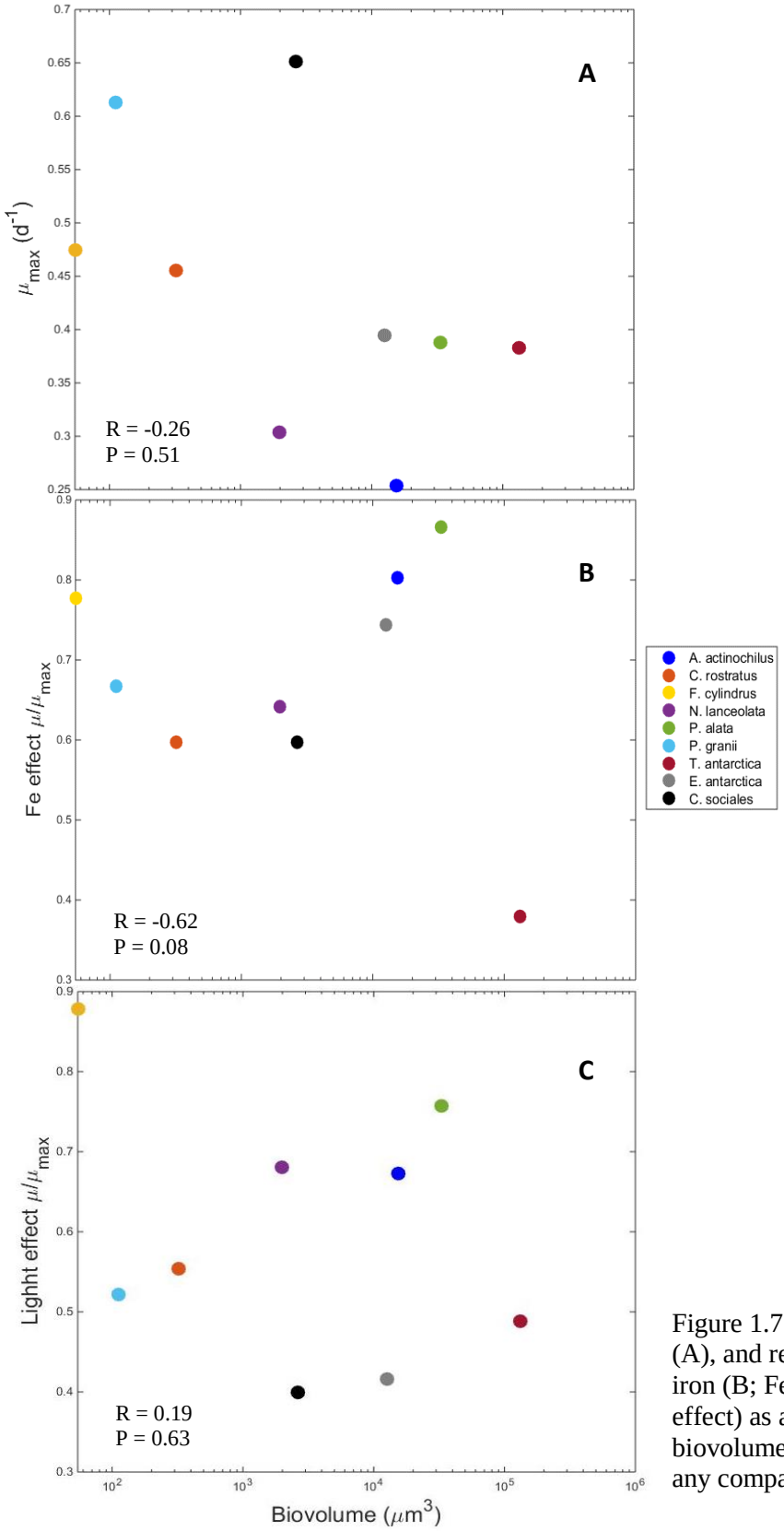


Figure 1.7. A-C. Maximum growth rates (A), and relative growth rates under low iron (B; Fe effect), and low light (C; light effect) as a function of log transformed biovolume. No significant correlation for any comparison.

effect) conditions. In replete conditions, μ_{max} was not significantly correlated with biovolume (r = -0.26, *p* = 0.51). Similarly, under low iron and low light conditions there was not a significant correlation between μ:μ_{max} and biovolume (iron; r = -0.06, *p* = 0.08 and light; r = 0.19, *p* = 0.63).

Discussion

This is the first study in which the physiology of nine Southern Ocean diatoms have been directly compared. The number and diversity of diatoms, including both centrics and pennates, isolated from this region is valuable as a way to understand how diatoms are able to survive under low iron and light conditions. Diatoms are a diverse lineage among the stramenopiles and those that reside in HNLC waters employ unique adaptations and acclimations to variable environmental conditions, such as low Fe:C ratios, and high relative growth rates under low light and iron availability (Strzepek et al. 2012) .

Maximum specific growth rates determined for *F. cylindrus, E. antarctica, P. alata, T. antarctica,* and *A. actinochilus* are comparable to previously published growth rates of the same or closely related species (Strzepek et al. 2012; Agusti & Duarte 2000; Timmermans et al. 2004). Strzepek et al. (2012) used DFB, a stronger siderophore than EDTA, resulting in slightly lower growth rates in the two diatoms used in their study (*E. antarctica, P. inermis*). To my knowledge, growth rates under varying iron and light conditions of three of our diatoms have not previously been published, including *C. rostratus, C. sociales,* and *N. lanceolata*.

As a way to measure degree of iron limitation, the photophysiological measurements F_v:F_m and σ_{PSII} were also collected. In iron-limited cells, F_v:F_m is expected to decline as the light-harvesting antenna systems transfer electrons to PSII reaction centers less efficiently. This decrease in efficiency is supported by previous research in culture and field-based experiments

(Green et al 1991; Petrou et al. 2014; Boyd et al. 2000; Trimborn et al. 2014). Conversely, σ_{PSII} characteristically increases under iron-limited conditions as a result of an increase in the ratio of PSII antenna complexes relative to reaction center complexes (Green et al. 1991). The increase in σ_{PSII} has also been proposed to be due to an increase in the size of the absorption cross section area (Strzepek et al. 2012).

In our low iron treatment, this pattern was evident in only *C. cf sociales*, *T. antartica*, and *N. lanceolata*, indicating these diatoms were experiencing moderate to severe iron limitation. *P. alata* did not exhibit any of the characteristics associated with iron limitation. While a decrease in μ:μ_{max} and F_v:F_m was not observed in *F. cylindrus* and *E. antarctica*, an increase in σ_{PSII} was seen, perhaps implying these species were just experiencing the onset of iron-limited growth when grown in the low iron medium used in this study.

Light limitation resulted in reduced μ:μ_{max} in all diatoms except for *F. cylindrus*, which did not exhibit significant reduction in μ, F_v:F_m, or σ_{PSII}, indicating this species was not experiencing light stress under the low light condition used in this study. For most other diatoms growing under low light, neither F_v:F_m nor σ_{PSII} changed relative to the replete treatment, except for *A. actinochilus*, *E. antarctica*, and *C. rostartus*, in which F_v:F_m actually increased. These trends in F_v:F_m and σ_{PSII} might suggest Southern Ocean diatoms have different mechanisms to maintain a healthy photosynthetic apparatus along with high photosynthetic efficiency, despite a reduced growth. Interestingly, diatoms that overwinter in the SO and experience months of darkness appear to retain a functional photosynthetic electron transport chain (Peters and Thomas 1996). This would help them rapidly acclimate to more favorable light conditions when the mixed layer depth shoals in spring. These observations suggest F_v:F_m and σ_{PSII} are not good indicators of light stress in Southern Ocean diatoms.

In many of these treatments, μ, F_v:F_m, and σ_{PSII} were appreciably decoupled. In terms of iron limitation, this may indicate several of our diatoms were only moderately iron-limited under our culture conditions. In order to produce growth and physiological parameters with characteristics of iron limitation, a further reduction in iron concentrations and/or a stronger iron chelator, such as desferrioxamine B mesylate (DFB), would need to be added to the culture medium. Under light-limiting conditions, F_v:F_m did not decrease, indicating SO diatoms may not have been experiencing much stress and maintained efficient electron transfer between light harvesting antennae and reaction centers. It may also indicate that SO diatoms have unknown novel mechanisms for maintaining high levels of photosynthetic efficiency. In terms of light limitation in SO diatoms, F_v:F_m may not be a good indicator of stress. μ was the most responsive physiological parameter for iron and/or light stress in our study. Overall, SO diatoms have larger σ_{PSII} and lower maximum F_v:F_m than temperate diatoms, which allows them to take advantage of low iron and light conditions (Strzepek et al. 2012 & Sugget et al. 2009)

Three environmental stress responses were observed: no interaction between iron and light, an additive interaction between iron and light, and a response in which light was the main limiting resource (Fig. 1.6). A stress response with no interaction is defined as a response in which resource limitation of two potential limiting nutrients is no greater than one of the two limiting nutrients by itself. In this example, reducing both iron and light produces the same reduction in μ as though either iron or light were limiting. *C. cf sociales, P. subcurvata, A. actinochilus, C. rostratus, N. lanceolata,* and *T. antarctica* growth can be classified as having no interaction between iron and light as –FeLL cultures were not significantly different from either –FeSL or +FeLL treatments (Fig. 1.6). This would indicate a synergistic relationship between iron and light does not exist in these polar diatoms. In contrast, Sunda and Huntsman (1995)

found that in temperate diatoms, low-light, iron-limited cells had reduced growth rates compared to high-light, iron-limited cells due to increased iron demands in order to efficiently use the photosynthetic electron transport chain under low light. It is interesting the diatoms that exhibit this growth response include among the smallest and largest diatoms used in this study. Small diatoms may be able to cope with low iron and low light without a decrease in growth because of increased SA:V ratios, decreased diffusive boundary layers, and increased iron uptake rates. Similar findings were observed in *C. brevis*, a small diatom grown in natural SO water (without EDTA) which thrived under low iron and low light conditions (Timmermans et al. 2001; Oijen et al. 2004). In contrast, large diatoms are able to exhibit this growth pattern between iron and light because they minimize their Fe:C ratios (Strzepek et al 2012). SO diatoms also have the ability to reduce iron strongly bound to organic complexes such as EDTA, DFB, and other siderophores, possibly making the pool of bioavailable iron larger in this region (Strzepek et al. 2011).

E. antarctica demonstrated only light limitation in this study, with no significant iron effect on growth. It is possible that *E. antarctica* was not iron-limited in our study as we used EDTA and not a stronger siderophore such as DFB. Previous researchers, who added DFB to induce iron limitation in this species (Strzepek et al. 2011), also observed that *E. antarctica* had high relative growth rates and near-constant Fe:C ratios even in low irradiances. In the case of *C. rostratus*, the patterns in growth are slightly more complicated, but are mainly driven by irradiance as well. Under light saturating conditions, low iron availability resulted in slightly reduced growth rate (12%; $p < 0.001$), however, a decrease in light resulted in a much larger decrease (45%; $p < 0.001$). In other words, the effect of iron limitation on growth rate was much

less pronounced in low light cultures than light saturated cultures, indicating *C. rostratus* is more sensitive to change in light than iron.

P. alata and *F. cylindrus* demonstrated an additive interaction between iron and light. In both of these isolates, growth rates were most reduced in the combined low iron and low light treatment compared to either variable alone. These two diatoms are also at the extremes of the size distribution among our diatoms. *P. alata* and *F. cylindrus* are the second to largest and the smallest cells, respectively, yet have similar functional characteristics in relation to iron and light limitation. The lack of synergistic effects observed for most of the isolates support the proposal by Strzepek et al. (2012), in which iron and light limitation in SO diatoms result in a distinctive photoacclimatory response that increases the number but not the size of PSU, which minimizes iron demands.

Phytoplankton cell size is known to correlate well with resource utilization and growth (Epply et al 1969). Nutrient uptake rates decrease due to diffusion limitation with increasing cell size (Marchetti and Maldonado, 2016), and growth typically decreases with increasing size (Banse 1976). Thus, small cells should dominate HNLC regions; however, iron kinetics are not the only factor responsible for community composition, allowing many large diatoms to exist in the SO for several biophysical reasons. First, they have large vacuoles with high nutrient storage capabilities, which could be useful under fluctuating nutrient regimes (Raven 1987). Second, their large size can serve as protection from zooplankton grazing and allows them to control their depth in the euphotic zone by varying their ballast (Armstrong et al. 2002). At the biochemical level, SO diatoms also have adaptive strategies to acquire iron and efficiently use light so they can survive in the SO. When cells are iron limited they can substitute iron-containing redox catalysts with proteins that can use other metal-cofactors (Peers and Price, 2006), activate a high-

affinity iron uptake system (Raven, 1990), perform luxury uptake of iron (Marchetti et al., 2009), and reduce Fe:C ratios (Strzepek et al., 2012). However, the physiological trade-off of reducing iron quotas in response to iron limitation is that open ocean diatoms compromise their ability to quickly acclimate to rapidly changing irradiance (Strzepek and Harrison, 2004).

Diatoms in this study displayed a weak, insignificant negative correlation between decreasing growth rate and increasing biovolume. This poor correlation between maximum growth rates and biovolume was also found among other SO diatoms (Strzepek et al. 2011). Maximum growth rate appears to be species specific and in the SO is likely determined by temperature and nutrient resource (nutrient) utilization strategies of each diatom. In our survey, it appears there are several types of diatoms exhibiting different resource utilization strategies, including diatoms that grow quickly and utilize nutrient pulses like *C. cf sociales* and *P. subcurvata* (high u_{max} but low $u:u_{max}$ in –FeLL); diatoms that grow slowly but can also take advantage of nutrient pulses perhaps by having large storage capacities such as *E. antarctica* and *P. alata* (large vacuoles); and diatoms that have low growth rates but can survive well under severe resource limitation, such as *N. lanceolata* and *C. rostratus* (high $u:u_{max}$ in –FeLL).

Interestingly, the ability of SO diatoms to cope with low iron and low light also does not scale with biovolume. This could be a result of not being iron limited in our medium, but it also suggests that large, SO diatoms have efficient mechanisms and gene repertoires that can maintain higher $\mu:\mu_{max}$ under these stressful conditions. On the other hand, smaller cells, such as *F. cylindrus*, *P. subcurvata*, and *C. rostratus* are able to maintain relatively high $\mu:\mu_{max}$, perhaps due to their increased SA:V ratios.

The oceanography of the sites from which these diatoms were isolated could partially explain the growth patterns observed. *C. cf sociales* and *N. lanceolata* were isolated from coastal

stations (200.-020 and 100.040) and likely represent coastal and/or sea ice dependent diatoms adapted to higher iron concentrations and shallower mixed layers associated with upwelling along the shelf. *T. antarctica* was isolated from an off-shelf region (200.180). *Thalassiosira* species have been shown to be associated with coastal areas and sea ice blooms (Lin et al 2015; Garibotti et al. 2005). *P. alata* and *E. antarctica* were isolated from a more southerly station associated with the sea ice edge bloom (200.040) where iron and light can interact synergistically. These large diatoms would benefit from their storage-adapted lifestyle in this region as the sea ice edge can be a source of nutrients and iron pulses and is associated with a shallower mixed layer.

Because of the importance of light in the SO, further research should investigate the effects of high light or diel light cycles on these diatoms. SO diatoms overwinter in near darkness or within sea ice where light levels routinely reach 0.1% of surface irradiance for months at a time. Further research with SO diatoms should investigate the elemental ratios (C:N:Si) including iron quotas and iron-use-efficiencies (amount of C per Fe used per day) to understand how these diatoms change their cellular composition in times of iron stress and how these physiological adaptations could relate to biogeochemical cycles. In particular, during the PalmerLTER 2014 field season, the same year in which these diatoms were isolated, it was observed that four diatom genera, including *Probosica*, *Thalassiosira*, *Pseudo-nitzschia*, and *Stellarima*, contributed to 88% of the net community productivity (NCP) variance (Lin et al. submitted). Thus, understanding cellular and molecular mechanisms of iron and light limitation within specific diatoms could help in our understanding of how these particularly important diatoms influence biogeochemical cycles.

Phenotypic variability in growth rates and photophysiology as a function of iron and/or light status observed among the diatoms examined in this study could likely be explained by differences in gene repertoires in relation to iron-requiring processes such as photosynthesis and nitrogen assimilation. In the next chapter, I will investigate if SO diatoms show a correlation between presence (or inferred absence) of particular genes and the physiological parameters discussed in this chapter. Elucidating the genotypic variability among SO diatoms could help in understanding phytoplankton community composition and succession, nutrient cycling, and carbon and silica sequestration in HNLC regions, which is especially important in light of current rates of climate change in the WAP region.

CHAPTER 2: INVESTIGATING THE TRANSCRIPTOMES OF SEVEN DIATOMS IN RELATION TO IRON AND LIGHT STATUS

Introduction

The first chapter addressed phenotypic acclimation to iron and light limitation. Here, I will investigate genotypic adaptations that have possibly resulted from selective pressure from low iron and light conditions. Despite numerous physiological investigations (Strzepek et al., 2012; Arrigo et al., 2010; Alderkamp et al., 2012) on the effects of iron and light limitation on Southern Ocean diatoms, there are no studies that use modern molecular techniques such as next-generation sequencing in conjunction with physiological studies to elucidate how SO diatoms cope with light and iron stress (Park et al. 2010; Strauss, 2012).

Diatoms are a diverse lineage comprised of two major clades. Centrics are older and divided further into radial or bi-multipolar varieties; pennates diverged later and are comprised of araphid and raphid pennates (Kooistra et al. 2007). The first two diatoms to have their genomes sequenced, the centric *T. pseudonana* and the pennate *P. tricornutum*, have been shown to be more divergent than fish and humans despite their relatively recent evolution (Bowler et al. 2010). In ancient oceans diatoms became successful 190 Mya, when the atmosphere contained almost eight times higher CO_2 concentrations and higher levels of dissolved inorganic iron (Armbrust 2009). Because iron has a high capacity as an electron donor and acceptor, diatoms likely evolved to use iron as a cofactor in many catalytic proteins involved in photosynthesis and

other biochemical reactions. Unfortunately, in modern oceanic environments iron is rapidly oxidized and precipitated. The selective pressure in low iron environments has resulted in diatoms evolving strategies including luxury storage, protein substitution, photoacclimation, and reduced cell size. In this study we investigated 22 genes (isogroups) from nine different metabolic groups including photosynthetic iron-dependent and independent proteins, iron storage, high-affinity iron-uptake, transport proteins, vitamin synthesis, superoxide dismutase and alternative oxidase, nitrogen assimilation, and proteins involved in the urea cycle.

We investigated two protein substitution pairs that are used in photosynthesis and that can be exchanged for each other in low iron conditions. The first protein substitution pair we investigated was cytochrome c_6 (CYTC6) and plastocyanin (PCYN). All complete diatom genome sequences have a gene encoding the iron-containing electron transporter CYTC6 (Raven et al., 2013), but only examined diatoms that come from iron-limited regions, such as *T. oceanica* and *F. cylindrus*, appear to have genes that encode for PCYN, an alternative, copper-containing electron transporter (Peers and Price, 2006). In addition, several diatoms, including *P. inermis*, transcribe more than one copy of PCYN (Groussman et al 2015). *T. oceanica* constitutively expresses plastocyanin even under iron-replete conditions, suggesting that plastocyanin expression is a permanent adaptation to open ocean, low iron environments (Peers and Price, 2006).

The second major protein substitution that can occur under low iron conditions is the substitution of ferredoxin (PetF) for flavodoxin (FLDA). PetF is a non-heme iron-sulfur protein that serves as the terminal electron acceptor and reduces NADP+ to NADPH which, together with ATP, provides energy to drive the light-independent Calvin Cycle. PetF has a high redox potential, but its 2Fe-2S protein cluster results in increased iron cell quotas. In low-iron

environments, diatoms have been found to replace PetF with FLDA, a protein that contains flavin-mononucleotide as a prosthetic group, as the redox cofactor instead of iron (La Roche et al. 1996). The ratio of PetF to FLDA has been used as a molecular indicator of iron stress (La Roche et al. 1996); however, multiple copies of FLDA (I and II) exist in certain diatom species and may not all be regulated by iron availability (Whitney et al. 2011, Groussman et al. 2015).

Table 2.1. Summary of protein substitutions in the photosynthesis electron transport chain

PROTEINS THAT REQUIRE IRON	**NON-IRON CONTAINING PROTEIN EQUIVALENTS**
Ferredoxin (PetF)	Flavodoxin (FLDA)
Cytochrome c_6 (CYTC6)	Plastocyanin (PCYN)

Two other proteins involved in photosynthesis were investigated, one that requires iron and one that does not. The first protein, plastid terminal oxidase (PTOX), is used in an alternative electron photosynthetic pathway from PSII to PTOX. PTOX requires two iron atoms, but this short pathway provides an electron shunt after PSII allowing photosynthesis to bypass the PSI complex, which requires 12 Fe atoms (three 4Fe-4S centers; Behrenfeld & Milligan 2013). In some diatoms, rhodopsins (RHO) can also be used as an iron/light management strategy. Diatom RHO is a putative light-driven proton pump that can be used for ATP synthesis and requires a small amount of iron for retinal synthesis. Recently, Marchetti et al. (2015) found that RHO in a temperate diatom, *P. granii*, was highly expressed at both the gene and protein level under iron-limited growth conditions. This suggests that RHO might be used when photosynthesis is compromised at low iron concentrations.

Physiological and molecular studies of *T. oceanica* and *P. tricornutum* have shown that diatoms employ a high affinity iron uptake system under low iron conditions. A ferric reductase (FRE) first reduces Fe(III) to Fe(II), which is then re-oxidized by a multi-copper oxidase (MCU) which is paired with or in close proximity to an iron permease (FTR) that transports Fe(III) into the cell (Marchetti & Maldonado, 2016). Interestingly, SO diatoms can acquire iron complexed to strong organic ligands; however, the mechanisms are unknown (Strzepek et al. 2011).

There are a number of other controlled iron uptake systems that diatoms can use. An iron-starvation induced protein 2a (ISIP2a) has been shown to be widely expressed among phytoplankton (Morrissey et al. 2015). It is highly expressed under low iron conditions and concentrates Fe(III) at the cell surface without the use of MCU (Morrissey et al. 2015). In *T. oceanica* and *P. tricornutum*, other low-iron responsive genes include *ISIP1*, *ISIP3*, natural resistance-associated macrophage proteins (*NRAMP*), and cellular repressor of E1A-genes (*CREG*; Allen et al. 2008). Transcripts for these genes have been identified as putative iron receptors because they are up-regulated under iron limitation and targeted to the secretory pathway (Lommer et al. 2012; Allen et al. 2008).

Once iron has reached the interior of the cell, it can be used immediately or stored for later use. Luxury uptake is the ability to acquire and safely store iron in excess of what the cell needs to grow, and often employs storage proteins called ferritin. Almost all raphid pennate diatoms for which whole-genome sequences or transcriptomes are available, such as those in the MMETSP, contain a ferritin gene, whereas ferritin genes in centric diatoms appear to be more randomly distributed (Groussman et al. 2015).

Iron is used in many other aspects of diatom metabolism besides photosynthesis. For example, the proteins involved in nitrogen uptake, including nitrate (NR) and nitrite reductase

(NiR), require iron cofactors. In a proteomic study of *T. pseudonana*, iron-limited cells were found to significantly up-regulate their nitrogen recycling pathways, resulting in conservation of iron normally allocated to NR, NiR, and other reduced nitrogen assimilation steps (Nunn et al., 2013; Lommer et al., 2012). In addition, Nunn et al. (2013) suggest further adaptations under iron limitation, including recycling of N in the urea cycle to create polyamines for nitrogen storage and silica precipitation. As a representative of the urea cycle, we chose argninosuccinate synthase (*ASSY*). It is one of the first enzymes in the urea cycle and was found to be highly expressed in an iron-enriched metatranscriptome study in an HNLC region (Marchetti et al. 2012).

In addition to iron, diatoms require vitamin B_{12}, which has been shown to be limiting in certain regions of the SO (Bertrand and Saito 2007). To our current knowledge B_{12} is only synthesized by bacteria. Diatom requirements for this nutrient depend on the type of methionine synthase they possess. Some diatoms only contain MetH, which requires B_{12}, while other diatoms also contain the B_{12}-independent methionine synthase MetE. Studies have shown that iron and B_{12} can act synergistically to enhance growth in iron-limited waters (Koch et al. 2011). In the SO, the ability to retain MetE would be advantageous as these areas can be co-limited by both nutrients. For example, *F. cylindrus*, a major bloom former in the SO, contains MetE allowing it to flourish in conditions that might be limiting to other phytoplankton that are auxotrophic for B_{12} such as following a large iron-induced bloom (Ellis et al. 2015)

Diatoms minimize reactive oxygen species (ROS) resulting from metabolic byproducts by using superoxide dismutases (SOD), which convert superoxide to a molecular oxygen and hydrogen peroxide (Peers & Price 2004). There are four isoforms of SODs, containing either manganese, copper/zinc, iron, or nickel. *T. pseudonana* uses Mn-SOD as its dominant SOD in

times of iron limitation (Wolfe-Simon et al. 2006). Many algae have both Fe-SOD and Mn-SOD; by expressing Mn-SOD, they can decrease their iron demands, especially when iron availability is low. Mitochondrial alternative oxidases (AOX) are similar to SODs in their structure and activity, and help to reduce ROS in the mitochondria (Maxwell et al 1999). They have a lower requirement for iron than other terminal oxidases (Bowler et al. 2010). Electrons from photosynthesis can also be transported via the malate shunt to the mitochondria, where AOX again can help with reducing ROS (Yoshida et al 2007).

These are just a few of the many genes affected by iron (and possibly light) limitation. Some have been shown to have a direct role in iron homeostasis, photosynthesis, and nitrogen assimilation, while others have putative functions in other pathways that are affected by iron and light. The focus here was to investigate the presence or absence of their transcripts within the transcriptomes of the polar diatom isolates described in Chapter 1, the genome of *F. cylindrus*, and other diatom transcriptomes made available by the MMETSP (Keeling, et al., 2014). Over 46 different diatoms were sequenced as part of this program, which has facilitated a better understanding of the various strategies diatoms employ to cope with iron and light limitation, as well as their ecology and influence on marine biogeochemistry.

Materials and Methods

RNA extractions, transcriptome library preparation and sequencing - Cultures grown for high throughput sequencing of mRNA were grown in acid-washed 2L polycarbonate bottles in iron-replete (pFe 19) conditions and standard light (90 µmol photons m^{-2} s^{-1}). After reaching late exponential/early stationary phase, cultures were harvested onto polycarbonate filters (3.0 µm pore size, 25 mm) and stored at -80°C. Total RNA was extracted using the RNAqueous 4PCR

Kit (Ambion) according to the manufacturer's protocols. Initial bead beating was also performed to remove all cells from the filters and eluted in water. Residual genomic DNA was eliminated by DNAse 1 digestion at 37 ºC for 45 min. Sample concentrations were quantified using a Qubit 2000 Fluorometer, while an Agilent Bioanalyzer 2100 was used to determine RNA integrity. mRNA libraries were generated with ~2 µg of total RNA and prepared with the Illumina TruSeq Stranded mRNA Library Preparation Kit. Eight samples were individually barcoded pooled prior to sequencing on the Illumina MiSeq platform at the High Throughput Sequencing Facility (HTSF) at UNC-Chapel Hill. Sequencing resulted in ~2 million paired-end reads of 2x300bp per sample

Sequence assembly and annotation - Illumina TruSeq adapters and poly-A tails were trimmed from raw reads using the Fastx_toolkit clipper function. Fastq_quality_filter was used to remove poor quality sequences, such that remaining sequences had a minimum quality score of 20 with a minimum of 80% of bases within a read meeting this quality score requirement. Any remaining raw sequences less than 50 base pairs in length were also removed. A custom Perl script repaired reads to maintain pairing information, and redundant data were removed using fastx_collapser. Merged files were de novo assembled using Trinity (Grabherr et al., 2011) with 750 GB of memory and 24 CPUs at the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. The resulting assembly was filtered to remove contigs less than 200 bp in length. Trinity-assembled contigs which exhibited sequence overlap were grouped into isogroups which were then used for sequence homology searches (BLASTx E-value $\leq 10^{-4}$) against the EuKaryotic Orthologous Groups (KOG; Altschul S., et al. 1990) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (Kanehisa, 2006).

To assess the quality of our transcriptome assemblies, the protein-coding sequences were used to determine completeness and contiguity. Completeness was first assessed by using the computational method CEGMA (core eukaryotic gene mapping approach), which characterizes the full and partial representation of 248 core eukaryotic genes represented in diverse taxa (Parra et al 2009). Partial representation of the 248 core eukaryotic genes is defined as an alignment of a protein to its protein family in the euKaryotic clusters of Orthologous Groups (KOG) that exceeds a minimum alignment score. A second metric of completeness was performed by evaluating KEGG mapping results using the KAAS server and the single-directional best hit method (http://www.genome.jp/kegg/kaas/). Conserved complexes and pathways such as the ribosome, spliceosome, and proteasome were evaluated to determine transcriptome assembly completeness. Finally contiguity, or the distribution of longest-assembled contigs mapped to reference genes, was determined as according to Martin and Wang (2011) with custom scripts.

For each transcriptome, unassembled sequence reads were aligned to the final Trinity assembly using Bowtie 2 (Langmead 2012). Mapped reads were normalized by the Reads per Kilobase per Million reads method (RPKM; Mortazavi et al. 2008). RPKM is a measure of the relative molar concentrations of mRNA and is calculated from the number of reads that map to a gene region (C), the exon length of the base-pairs in a gene (L), and the total number of mapped reads in a transcriptome (N).

$$RPKM = \frac{C \cdot 10^9}{L \cdot N}$$

RPKM takes into consideration two biases from Illumina sequencing runs: library size and gene length. However, it has been shown in differential expression analyses that normalizing by gene length can introduce a bias toward less expressed genes (Oshlack and Wakefield 2009). In addition, the mean expressed transcript length can vary between samples, so by removing this

normalization factor, expression is more comparable across samples and species (Li et al 2009). While RPKM has been shown to be inconsistent between samples, it is still a useful unit as a semi-quantitative measure for identifying presence and absence of key genes of interest in this study.

Presence/absence and biogeographical distribution of key genes of interest - 22 key protein-encoding genes of interest were used in the study to investigate the molecular basis of iron and light limitation in polar diatoms. These genes include: ferritin (*FTN*), iron-starvation-induced protein -1, -2a, and -3 (*ISIP1, ISIP2a* and *ISIP3*), flavodoxin (*FLDA*), ferredoxin (*petF*), plastocyanin (*PCYN*), cytochrome c_6 (*CYTC6*), cobalamin-independent methionine synthase (*MetE*), Fe- and Mn- superoxide dismutases (*SOD*), mitochondrial alternative oxidase (*AOX*), multicopper oxidase (*MCO*), iron (III) permease (*FTR),* ferric reductase (*FRE*), rhodopsin (*RHO*), plastid terminal oxidase (*PTOX*), cellular repressor of EA1-stimulated genes (*CREG*), nitrite and nitrate reductase (*NiR, NR*), and argininosuccinate synthase (*ASSY*). Reference sequences for each of these genes were obtained from the *F. cylindrus* and *P. tricornutum* JGI genome portals and *T. psuedonana* and *T. oceanica* NCBI and GenBank repositories (Table 3). Reference sequences were identified in the transcriptomes by translated nucleotide homology searches (tBLASTn) with an e-value cutoff of $<10^{-5}$. The RPKM read count unit was determined for isogroups that were homologous with the reference query.

Subsequently, reference sequences were identified in the MMETSP protein database by BLASTp (e-value $<10^{-5}$) homology searches among the diatom transcriptomes. The 76 diatoms sequenced as part of the MMETSP include raphid and araphid pennates, and radial and bi-multipolar centrics isolated from all ocean basins including coastal, open-ocean, polar and HNLC regions. The transcriptomes and their associated latitude and longitude were obtained

from iMicrobe Data Commons (Project Code CAM_P_0001000) and the National Center for Marine Algae and Microbiota (NCMA). Quantitative read counts, through RPKM or other differential analysis, could not be used for comparing genes among samples because of inherent differences in culturing and sequencing conditions. Therefore, a semi-quantitative comparison was made by investigating the presence and absence of genes within the MMETSP database. Importantly, because gene expression as well as transcriptome depth varies among culture conditions and the absence of a gene could perhaps be due to transcriptional regulation, this data provides evidence of presence, but not absolute absence. Custom Matlab scripts allowed global biogeographical distribution of key genes of interest to be mapped (courtesy N. Cassar).

Results and Discussion

Sequencing assembly and annotation – Illumina MiSeq sequencing of ten pooled samples resulted in raw sample libraries averaging one million paired-end reads. Library size ranged from 681,141 (*N. lanceolata*) to 1,932,817 (*C. rostratus*) paired-end reads. Minimum length of a read after trimming was 200 bp in all transcriptomes and the maximum read lengths was 687 bp (Table 2.1). Trinity assemblies yielded transcriptomes with 10,659 (*C. cf sociales*, N50 =315) to 33,912 (*P. subcurvata*, N50=935) contigs. Mapping reads back to the assembled transcriptome resulted in mapping efficiencies of 24% (*A. actinochilus*) to 54% (*T. antarctica*). Functional annotation was provided with KAAS BLAST comparison against the KEGG database. The single-directional best-hit method was used, yielding transcriptomes with 13 to 27% functionally annotated contigs.

Table 2.2. Statistics of sequencing, assembly, and quality metrics for completeness and contiguity.

	C. sociales	***P. alata***	***A. actinochilus***	***N. lanceolata***	***T. antarctica***	***C. rostratus***	***P. subcurvata***
Transcriptome size (Mb)	3.6	11.4	18.4	22.1	22.3	21.8	22.9
Raw sequence reads (bp)	1188055	996739	1158920	681141	1363451	1932817	1031210
Average read length (bp)	338	411	453	493	567	590	687
Number of contigs	10659	27792	40647	44909	39240	36912	33912
Number of isogroups	9944	25796	35132	42346	32250	32392	28884
Maximum length (bp)	5810	7209	5975	7283	6551	6817	8191
N50	315	440	511	568	747	815	935
Contiguity*	0.07	0.11	0.2	0.16	0.23	0.19	0.25
CEGMA**	19%	38%	56%	65%	72%	69%	75%
Mapping efficiency	47%	51%	24%	50%	54%	51%	46%
Functional annotation - KEGG	2,604 [27%]	4,481 [17%]	5461 [19%]	5,781 [13%]	5625 [15%]	5968 [18%]	5,607 [19%]

* Threshold 0.75

** Partial completeness of genome based on 248 core eukaryotic genes (CEGs)

Table 2.3. Presence/absence of key genes of interest in 7 diatom transcriptomes. Presence of a gene is highlighted in green, with semi-qualitative RPKM values shown. Absence of gene is denoted in red. *E. antarctica* is not included as sequencing of this transcriptome was not successful during the first sequencing run.

		C. sociales	**P. alata**	**T. antarctica**	**N. lanceolata**	**A. actinochilus**	**C. rostratus**	**P. subcurvata**
Photosynthetic Fe-dependent proteins	Ferredoxin	32	-	-	646	104	-	5
	Cytochrome c_6	443	40	3092	612	701	99	33
Photosynthetic Fe-independent proteins	Plastocyanin	86	532	161	223	2891	2478	528
	Rhodopsin	-	14417	-	45	217	2962	117
	Flavodoxin	39	19857	732	32	770	45293	759
	Plastid terminal oxidase	186	13	28	11	81	173	155
Storage	Ferritin	-	-	-	53	-	86	46
High-affinity Fe uptake	Ferric reductase	-	-	912	96	-	-	30
	Iron permease	30	-	-	34	863	-	-
	Multi-copper oxidase	-	-	224	92	-	-	-
Transporter/Receptor	Iron starvation induced protein 2a	8	80	448	252	489	1274	1454
	Iron starvation induced protein 3	15	87	150	4	215	517	133
	Iron starvation induced protein 1	-	482	120	6	1531	579	22
	Natural resistance against macrophage	28	-	329	37	43	283	15
	Cellular repressor of EA-1	-	203	123	4	32	133	44
Vitamin synthesis	B_{12}-methionine independent synthase	-	104	138	146	39	31	-
	B_{12}-methionine dependent synthase	24	245	150	83	79	337	257
Reducing ROS	Mn/Fe superoxide dismutase	-	295	478	24	60	695	50
	Mitochondrial alternative oxidase	-	-	51	12	17	-	212
Nitrate assimilation	Nitrite reductase	13	592	1187	47	587	481	233
	Nitrate reductase	65	466	609	187	5263	1182	177
Urea cycle	Argininosuccinate synthase	132	2688	2703	111	1276	263	82

.

Transcriptome quality – In addition to reporting N50 values, each de novo assembly was assessed for contiguity and completeness. Contiguity was calculated at the 0.75 level and yielded transcriptomes with 0.07 (*C. sociales*) to 0.25 (*P. subcurvata*) coverage values, indicating there are many small transcripts within our transcriptomes. Ideally, the distribution of contigs that cover a reference gene should be close to 1. However, in the case of *N. lanceolata*, a mid-range quality transcriptome, contiguity was 0.16 (Fig. 2.1).

The second quality metric, completeness, was calculated with CEGMA and KEGG mapping tools. CEGMA calculated the partial representation of 248 highly conserved core eukaryotic genes. The partial representation of a protein is an alignment of a protein and the KOG for that protein family that exceeds a minimum alignment score (Parra et al. 2009). Completeness for our transcriptomes ranged from 18.5% (*C. sociales*) to 75% (*P. subcurvata*). By using the KEGG mapping tool we were also able to determine that highly conserved pathways such as the ribosome and spliceosome pathways were between 65% and 80% complete for *N. lanceolata* (KAAS maps not shown for all transcriptomes; Fig. 2.2). Overall, we have small transcriptomes characterized by small contigs and isogroups. In some transcriptome libraries, our sequencing depth may not have deep enough to fully assess the presence of key genes of interest used in this study, especially in the smaller transcriptomes of *C. sociales* and *P. alata* and therefore should be interpreted with caution (Table 2.1).

Figure 2.1. Contiguity of *N. lanceolata* transcriptome. Contiguity was measured at the 0.75 level resulting in 16% of the transcripts covering half of the transcriptome. *N. lanceolata* represents a mid-size transcriptome.

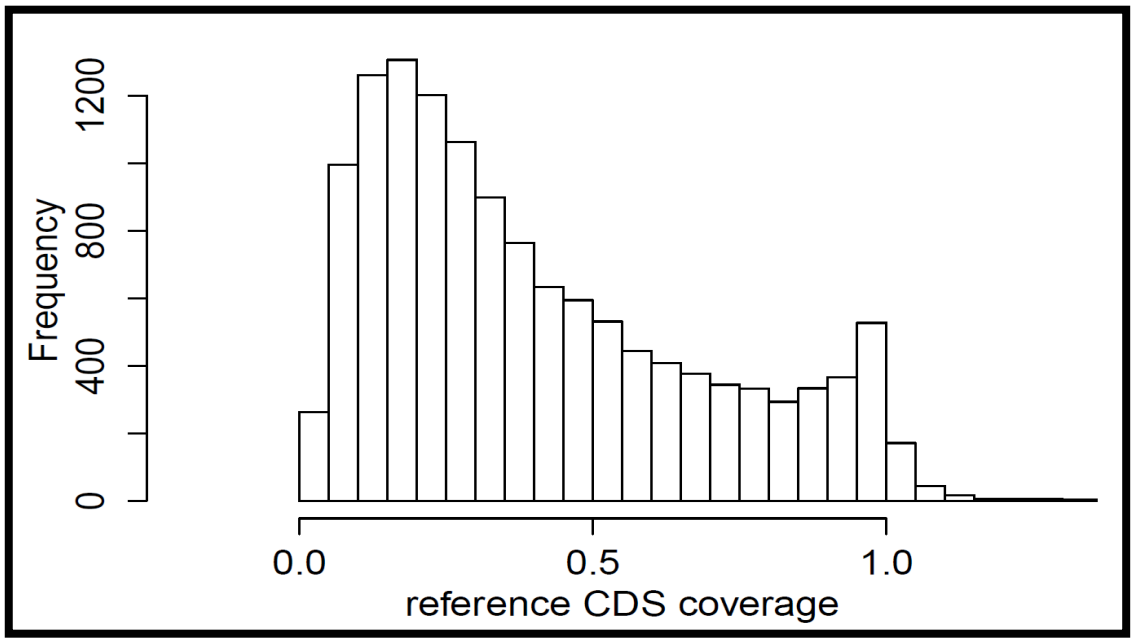
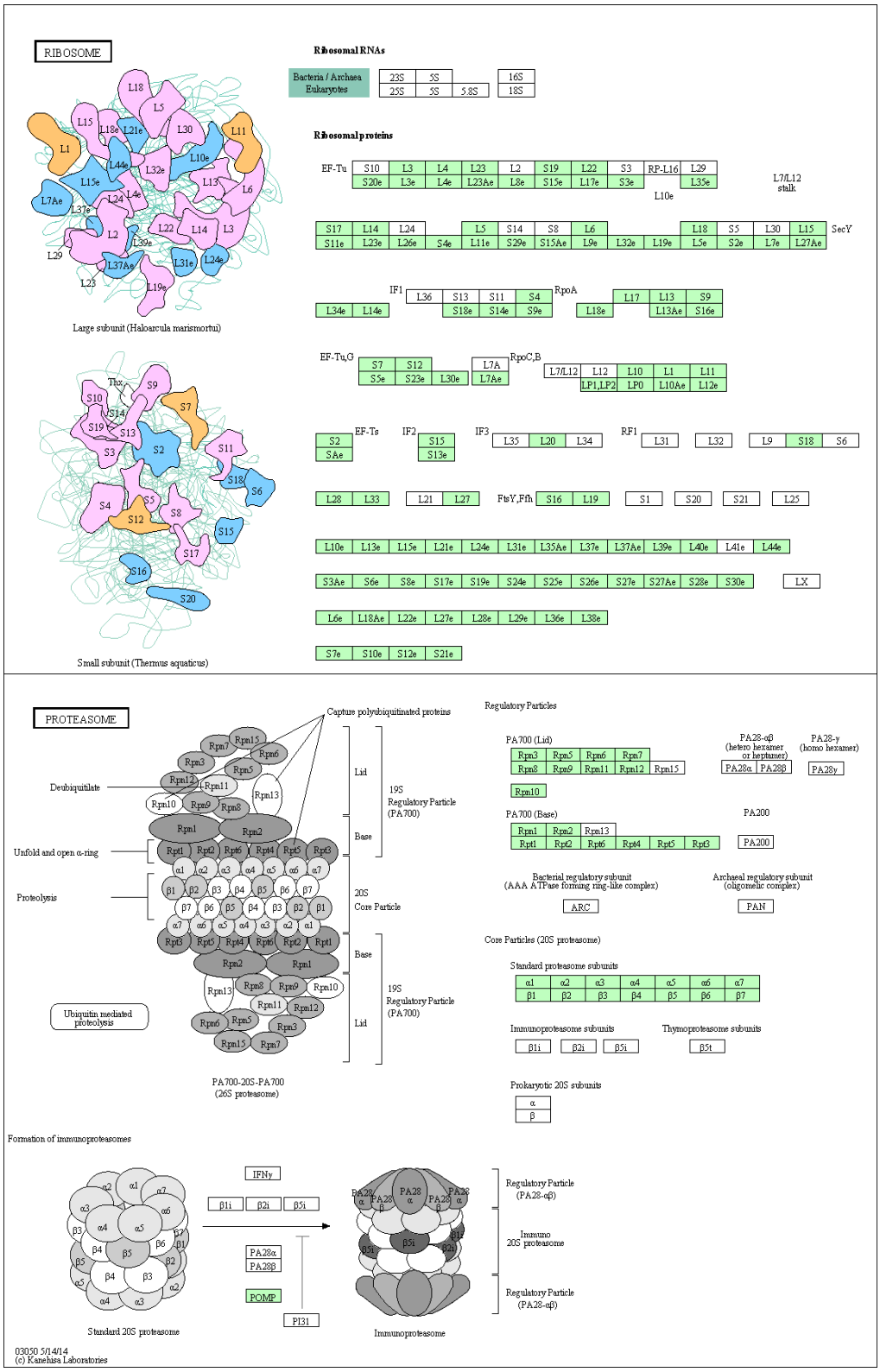


Figure 2.2. *N. lanceolata transcriptome* quality by analyzing highly conserved ribosome and proteasome pathways. Ribosome pathway = 80% complete, proteasome pathway = 65% complete;



Presence/absence of genes in seven diatom transcriptomes – We investigated 22 protein-encoding genes potentially involved in an iron and/or light stress physiological response. These genes fall into nine functional categories, including photosynthetic iron-dependent and independent proteins, iron storage proteins, high-affinity iron-uptake proteins, transport proteins, vitamin synthesis proteins, superoxide dismutase and alternative oxidase proteins, nitrogen assimilation proteins, and proteins involved in the urea cycle (Table 2.3). We also calculated RPKM as a way to semi-qualitatively confirm likelihood of presence of a gene.

Transcripts for the redox substitution pairs *CYTC6* and *PCYN*, well known photosynthetic mobile electron carriers, were found in all of our diatoms, suggesting that while *PCYN* is likely beneficial for growth in HNLC environments, it does not preclude the expression of *CYTC6*. These two proteins are functionally very similar in regard to their molecular mass as well as their ability to transfer electrons, therefore, their expression in the SO might largely be based on whether copper or iron is limiting (Raven et al. 1999). Because PCYN synthesis requires tenfold more copper compared to CYTC6, the availability of copper to phytoplankton is important in the SO. In field experiments from another HNLC region in the North Pacific ocean, copper has not been shown to be limiting and is likely sufficient for growth throughout the world's ocean (Peers & Price 2006). However, if copper did became limiting, the principle benefit of transcribing *CYTC6* is that it would be useful in times of copper limitation (Peers & Price, 2006). No studies of copper concentrations in the WAP surface waters have been performed, and it would be interesting to see if there was a relationship between copper limitation and decreased expression of *PCYN*.

There were overall higher RPKM values for *PCYN* than *CYTC6* in each transcriptome, except in the case of *T. antarctica* where RPKM for *CYTC6* was ~19-fold higher, possibly suggesting this species still relies primarly on CYTC6. In previous research, it has been shown that *PCYN* is widely distributed among algal lineages, perhaps originating early in diatom history, indicating that diatoms that do not contain *PCYN* have lost the gene (Groussman et al. 2015). In addition, several diatoms, including *Fragilariopsis kerguelensis* and *Pseudo-nitzschia heimii*, were found to transcribe more than one copy of *PCYN* (Groussman et al 2015). Having multiple paralogs of *PCYN* might result in one paralog with a new function or could increase the capacity of PCYN to transfer electrons, so while *T. antarctica* contains *PCYN*, it could potentially have another function. In Chapter 1, *T. antarctica* also had the lowest relative growth rate when grown under low iron conditions, indicating transcription of *CYTC6* might impact the fitness of this species in relation to low iron, and possibly light.

The second redox pair we investigated was *PetF* and *FLDA*, which are the terminal electron acceptors in photosynthetic electron transfer (PET) chain. While *FLDA* was present in all transcriptomes, PetF was only present in *C. cf sociales, N. lanceolata, A. actinochilus*, and *P. subcurvata*. In *N. lanceolata*, there were ~20 times more transcripts of *PetF*, as measured by RPKM. While the other diatoms have minimal transcription of *PetF* (RPKM < 100), *N. lanceolata* seems to rely primarly on this protein. While *N. lanceolata* had a low relative growth rate in low iron treatments, this species had all genes that we queried for, indicating it possibly makes up for its reliance on iron-containing proteins such as PetF and CYTC6 with other proteins involve in iron storage and uptake, such as FTN, the high-affinity iron uptake system. *N. lanceolata*,

which is anticipated to have high iron requirements, also possessed MetE which likely helps it to avoid co-limitation of iron and B_{12}.

Five diatoms, including *T. antarctica, P. alata, F. cylindrus, C. rostratus,* and *E. antarctica* (inferred from MMETSP strain, but included in the physiological studies of Chapter 1) lacked detectable *PetF* transcripts. Although our transcriptomes were not deeply sequenced and our growth conditions limit what transcripts might be detected, the possible lack of PetF suggest that these diatoms are using FLDA to alleviate low iron stress and that this might be a permanent adaptation. Other polar diatoms, such as *Fragilariopsis curta* and a *Pseudo-nitzschia* sp., constitutively express *FLDA* and do not transcribe *PetF* (Pankowski & McMinn, 2009). While PCYN has a lower redox potential than PetF, it would be useful in alleviating iron limitation by lowering iron requirements in PET (Marchetti & Maldonado 2016). RPKM of *FLDA* are elevated in *P. alata* and *C. rostratus* (>19,000), perhaps reflecting a need to transcribe more copies to counteract the lower redox potential. It also implies that these species may rely heavily on this strategy, and their high relative growth rates (>0.87 d^{-1}) could be attributed to the presence (and high expression) of FLDA. Multiple copies of *FLDA* (I and II) exist in certain diatom species and may contribute to increasing redox potential for PET in these diatoms, or it is possible they are not regulated by iron availability, but rather perform another function (Groussman et al. 2015; Whitney et al. 2011).

In some diatoms, the production of rhodopsin (RHO) can be used as another iron/light management strategy in addition to changing the photosynthetic architecture. Of our sequenced diatoms, only two species lacked a detectably expressed rhodopsin gene, *T. antarctica* and *C. cf sociales*. The *C. sociales* transcriptome is small (1 million

bp) and is characterized by low quality isogroups (N50 =315), therefore, it is possible this species has *RHO*, but our sequencing was not deep enough to capture its transcripts. In addition, *C. cf sociales* did not exhibit an interaction between iron and light; RHO possession might help account for the relatively small effect of iron limitation that we observed. This evidence, along with its transcriptome quality score, suggests *C. cf sociales* could contain *RHO* within its gene repertoire; however, deeper transcriptomes are needed. Lack of *RHO* transcripts in *T. antarctica* might contribute to its observed low relative growth rates. On the other hand, *P. alata* had high RPKM values for *RHO* and high relative growth rates, indicating that *P. alata* might lower its iron requirements through activating RHO.

High-affinity iron uptake systems are used by diatoms in transporting iron across the cell surface and are comprised of a ferric reductase (FRE), multi-copper oxidase (MCO), and an iron (III) permease (FTR). Five diatoms had at least one component of a high-affinity iron uptake system. Interestingly, *P. alata*, *E. antarctica*, and *C. rostratus* did not contain detectable transcripts for the high-affinity iron uptake system. Previous research has also shown that the same (or closely related) species lack detectable gene expression of the high-affinity Fe uptake system (*P. alata, E. antarctica, C. neogracile*; Groussman et al. 2015). Despite lacking detectable transcripts, these diatoms had high relative growth rates under low iron conditions in this study, indicating either they were not severely iron limited or they utilized alternative iron transporters or cell surface receptors.

These three diatoms expressed genes for all other transporters queried in this study, including the ISIPs, NRAMP, and CREG. The exception is *P. alata* lacked a

detectable homolog of NRAMP. ISIPs and CREG have been shown to have multiple paralogs resulting from either duplication events or transposition (Lommer et al. 2012). Having duplicated genes might allow the cell to increase the number of transporters, creating a higher density, and increase the overall affinity of the cell for iron. Another hypothetical mechanism for iron uptake is the use of short-term iron storage molecules on the cell surface, similar to transferrins identified in *D. salina*; however, there is no similar mechanism known in diatoms presently (Kustka et al. 2007; Allen et al. 2008). Because *P. alata, E. antarctica*, and *C. rostratus* are speculated not to have a high-affinity iron uptake system similar to the one found in yeast (FRE, MCU, and FTR) (Van Ho et al. 2012), they might employ any of these other iron uptake strategies as well as increase their storage capabilities while decreasing their iron requirements.

F. cylindrus and *N. lanceolata* are the only diatoms that possessed identifiable transcripts for the entire high-affinity iron uptake system in addition to all the transporters for which we queried. While *F. cylindrus* and *P. alata* seemed functionally similar, as they both displayed an additive interaction response between iron and light, their acquisition strategies appear to be fundamentally different. This may allow them to take advantage of different ecological niches, such as open ocean ecosystems or ice edge environments, with respect to iron and light limitation in the WAP. It would be interesting to investigate the role of ISIP1 and ISIP3 as well as other mechanisms of iron acquisition in diatoms.

Interestingly, the only diatoms shown to transcribe an *FTN* homolog are *N. lanceolata* and *P. subcurvata*, both raphid pennate diatoms, and *C. rostratus*, a bi-multipolar centric. *FTN* has been shown to be present in all four clades of diatoms, with

paralogs present in many species, and may have been introduced by lateral gene transfer from a bacterium (Groussman et al. 2015). The ability of ferritin to quickly store iron for later use is advantageous for diatom genera like *Pseudo-nitzschia* and *Fragilariopsis* as these diatoms often form large blooms after iron fertilization and continue dividing after iron has returned to background levels (Marchetti et al., 2009). However, *FTN* is likely not required by all HNLC diatoms, such as in centrics which may possess a vacuolar storage system for iron (Nuester et al. 2012).

Because vitamin B_{12} may at times be variable or limiting to diatom growth in the WAP, diatoms that only contain *MetH* would be affected to a greater extent than diatoms that possess both *MetH* and *MetE* (Bertrand et al 2007). Only three of our diatoms lacked a detectable *MetE* gene, including *E. antarctica, P. subcurvata*, and *C. cf sociales*. While the lack of *MetE* transcripts in *C. cf sociales* may be due to a low quality transcriptome, a number of other Pseudo-nitzschia species have been shown to lack a functional *MetE* gene (Ellis et al. 2015). Phytoplankton bloom succession in the WAP could be affected by concentrations of B_{12}, such that B_{12} auxotrophs would bloom first and draw down B_{12}, while diatoms that do not require B_{12} would bloom once B_{12} concentrations were depleted (Sañudo-Wilhelmy et al. 2014; Hoppe et al. 2013). Bertrand et al. (2007) showed that B_{12} availability in the Ross Sea influenced the phytoplankton community, favoring the auxotrophic *P. subcurvata* to bloom when B_{12} was also added to iron enrichments. *E. antarctica* and *P. subcurvata* could be early bloom formers, while diatoms like *P. alata, N. lanceolata*, and *T. antarctica* are late bloom formers. Indeed, *P. inermis* has been shown to bloom in late austral summer, accounting for greater than 90%

of the diatom biomass observed in February 2006 in Ryder Bay (Adelaide Island, Antarctica; Annett et al., 2010).

The last protein-encoding gene which showed variation in the presence and absence of transcripts within our transcriptomes was *AOX*. The AOX protein serves as an electron sink for excess reductants diverted from photosynthesis to the mitochondria in times of photoinhibition (Prihoda et al. 2011). *P. alata, C. cf sociales*, and *C. rostratus* did not have any detectable transcripts for an *AOX* homolog in this study. They may use another terminal oxidase for an electron sink, but a recent study has verified the importance of AOX in balancing the ATP/NADPH ratio in PET and optimizing carbon fixation/growth (Bailleul et al. 2015). It found in *P. tricornutum*, *AOX* facilitates the interaction between the plastid and mitochondria with decreases in *AOX* expression in high light, linked to decreased PET capacity (Bailleul et al. 2015).

Genes that encode putative transporters (*CREG, ISIP1, ISIP3*), and components of nitrogen assimilation and the urea cycle (*NR, NiR*, and *ASSY*), and *MetH* were found in all of our transcriptomes and throughout the majority of diatoms in the MMETSP. Although these genes are shared throughout diatom phylogeny it would be interesting to understand how they evolved and under what conditions they are most highly expressed.

Our transcript data suggest a variety of gene repertoires in our diatom isolates, likely resulting in different iron requirements and physiological responses to iron and light limitation. Differences in resource utilization strategies as well as biovolume identified in Chapter 1 could be a direct result of differences in gene repertoires or acclimation strategies used by diatoms. *P. subcurvata, F. cylindrus*, and *C. cf sociales* were grouped physiologically based on having relatively higher μ_{max} and smaller size.

Despite the lower quality transcriptome in *C. sociales*, they appear to have similar gene repertoires with the exception that *PetF* was detected in *P. subcurvata*, and *MetE* in *F. cylindrus*. The storage-adapted species include *T. antarctica, P. alata, A. actinochilus* and *E. antarctica*. Given their large size and low μ_{max}, it is surprising that these diatoms, overall, appear to lack *FTN* and a complete high-affinity uptake system. *T. antarctica* and *A. actinochilus* do have a component of the uptake system; however, this does not seem to benefit them, as they have the lowest μ_{max} of the large diatoms. Because of their low μ:μ_{max}, *T. antarctica* and *A. actinochilus* might not be competitive in environments with very low iron or light availability as compared to other large diatoms present. The survivalists, *N. lanceolata* and *C. rostratus*, have low μ_{max}, high μ:μ_{max} in extreme conditions (-FeLL) and mid-range size. They do not detectably transcribe many of the same genes within their genome; for example, *C. rostratus* putatively lacks the high-affinity uptake system, *PetF*, and *AOX*. Despite differences in their gene repertoires, they share functional characteristics.

In some species, the range in growth rates observed is likely a result of a combination of a diatom's ability to phenotypically acclimate to iron and light limitation, such as changing the size of PSUs or lowering iron requirements, and their genetic adaptations to low iron and light as represented by their gene repertoires. For example, centric diatoms that acclimate to low iron conditions by altering their photosynthetic architecture likely have permanent adaptations in their gene repertoires that include containing *PCYN, FLDA*, and *RHO* and not *PetF*. In other cases, the molecular basis of low iron and light tolerance have yet to be resolved, as in the case of *FTN* and high-affinity iron acquisition. While *FTN* may be the molecular basis for iron storage in some

diatoms, we still cannot identify genes involved in vacuolar storage of Fe in other diatoms, such as the large centrics in this study. In addition, most diatoms in this study lacked identifiable transcripts of some or all of the high affinity iron uptake system (except *N. lanceolata*), indicating these diatoms likely have unique ways of acquiring iron. In this area, more research is needed to understand the molecular basis of iron acquisition and storage in diatoms.

Overall, there appear to be diverse ways in which diatoms are able to cope with low iron and light conditions, with some genes that may be directly involved in phenotypic acclimation, while others leave questions remaining as to how diatoms acclimate/adapt to low iron. In any case having a variety of mechanisms among diatoms may allow them to take advantage of different ecological niches or play a specific role in phytoplankton bloom succession. Ultimately, the biophysical and genetic makeup of these diatoms contribute to the overall phytoplankton composition and their effects on biogeochemical cycles.

To understand if the ecophysiology of these diatoms has a biogeographical distribution throughout surface waters of the ocean, we performed sequence homology searches of our key genes of interest in the MMETSP database, a publically available transcriptome of *Pseudo-nitzschia granii*, an in-house transcriptome of *Gramanema islandica*, and the JGI genome portal of *F. cylindrus* (BLASTp results for these three genes and diatom location information is provided in the appendices section). Although a much larger set of transcriptomes and diatoms would be needed for a thorough understanding of the biogeographical distribution of our key genes of interest, we can still gain insight into the extent of geographical influence on distribution of these genes. In the

case of most genes, we did not observe a trend of presence or absence (data not shown) with location. However, *MetE*, *FLDA*, and *PetF* transcripts did display biogeographical patterns associated with location of isolation (Fig. 2.3). In the case of *MetE*, diatoms from HNLC regions, in particular the SO, were more likely to possess transcripts to possibly reduce co-limitation of B_{12} and iron (Fig.2.3A). Bacterial production may be driving this pattern as well, because it has been observed that the SO has low abundances of bacteria and cyanobacteria, which are producers of B_{12} (Boyd et al. 2000). There also seems to be a geographical influence on diatoms that contain *FLDA*, with diatoms that possess this gene more often isolated from HNLC regions (Fig. 2.3B). Since iron is limiting in these regions, it makes sense that diatoms would retain *FLDA*, and possibly eliminate *PetF* in their gene repertoires. In agreement with this, we observed a general absence of *PetF* transcript recovery in diatoms isolated from the SO and other HNLC regions (Fig. 2.3C). Thus, the presence or absence of a single gene might have ecological consequences, especially in the case of *MetE*, *PetF*, and *FLDA*; however, the complete distribution of species with and without these genes would have to be investigated.

An application of this combined physiological and transcriptomic study is that we now have a collection of genes that may possibly serve as good molecular indicators for iron and or light limitation. It has been proposed that the expression ratio of flavodoxin to ferrodoxin (*FLDA*:*PetF*) may be an indicator for iron stress; however, multiple copies of *FLDA* may have different functions apart from accepting electrons in photosynthesis, and *FLDA* expression could be affected by factors besides iron limitation (Whitney et al. 2011). In addition, not all diatoms appear to have *PetF*, or even if they do, it is not preferentially used even under high iron conditions. Recently, a molecular indicator of

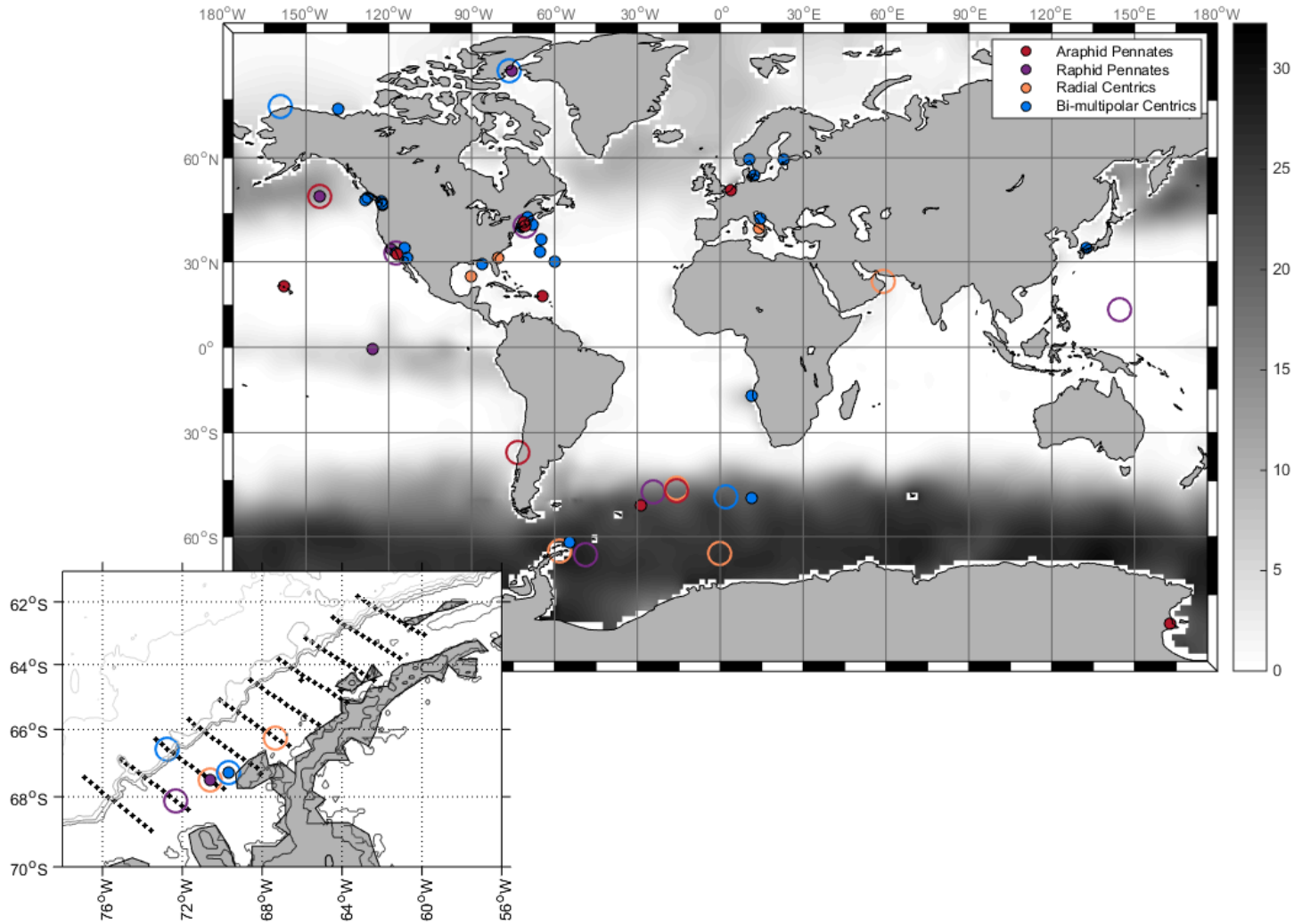
FLDA and *ISIP3* has been proposed in *T. oceanica*, in which the expression of these genes is inversely related to iron concentrations (Chappell et al. 2015). While having the expression of two genes investigated in one diatom provides valuable information about the nutritional status of that species, and possibly genus, it might not provide information about the entire community. Although this study does not provide a comparative analysis of gene expression, the use of transcriptomes is proposed as molecular indicators in future studies of iron and light nutritional status in SO diatoms. However the genes employed must be specific to each particular diatom.From our biogeographical study, it is clear that *MetE*, *PetF*, *FLDA*, *FTN*, *RHO*, *AOX*, and the *ISIP*s should be among the genes analyzed in a transcriptome, because they appear to be present in most (but not all) diatom clades and likely reflect iron and light limitation.

By combining transcriptomic and biophysical approaches, a better understanding of how diatoms survive under iron and light limitation can be gained. Overall, there appear to be diverse ways in which diatoms are able to cope with low iron and light conditions, with the presence (and inferred absence) of some genes helping to resolve the molecular mechanisms used to phenotypically acclimate or adapt to low iron and light. However, some acclimation/adaptation mechanisms remain unresolved and more research is needed to fully elucidate how diatoms survive in the harsh conditions of the SO. In any case, having a variety of mechanisms among diatoms may allow them to take advantage of different ecological niches or play a specific role in phytoplankton bloom succession. Future work with SO diatoms should include obtaining intracellular elemental and iron quotas, comparative gene expression, and photochemistry analyses to better understand the adaptation and acclimation strategies that diatoms possess. This

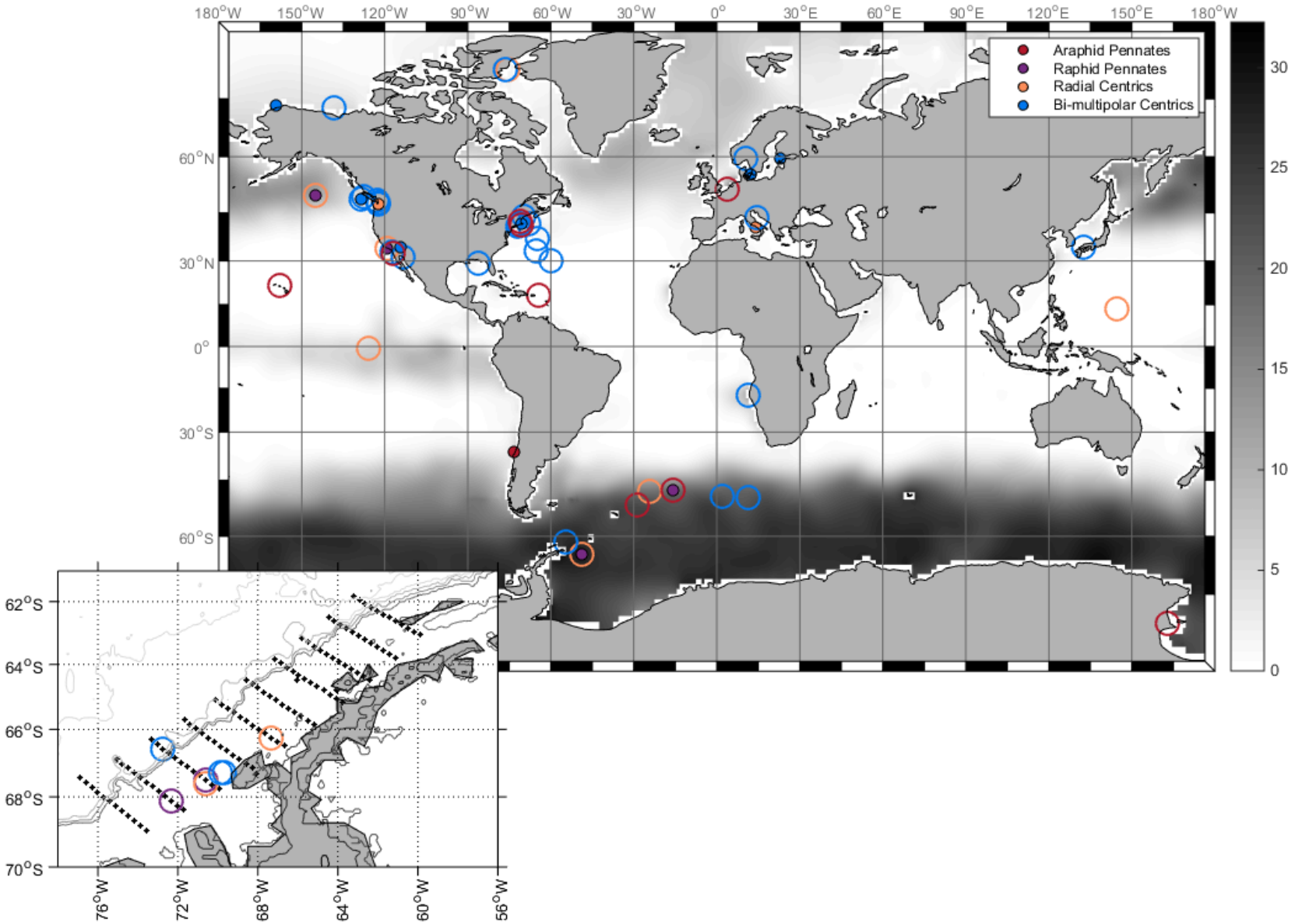
data will be helpful for modelers who rely on linking iron to carbon dynamics in order to understand the role diatoms play in global biogeochemical cycles.

Figure 2.3 A-C. Distribution of key genes of interest from isolated diatoms sequenced as part of this study and the MMETSP. An open circle represents a diatom isolate that has the gene and a closed circle represents an isolated diatom where the gene is absent in the transcriptome. Sea surface nitrate concentrations are shown in grey scale with darker colors corresponding to higher nitrate (μmol L^{-1}). Inset graph is a close-up of the WAP region.

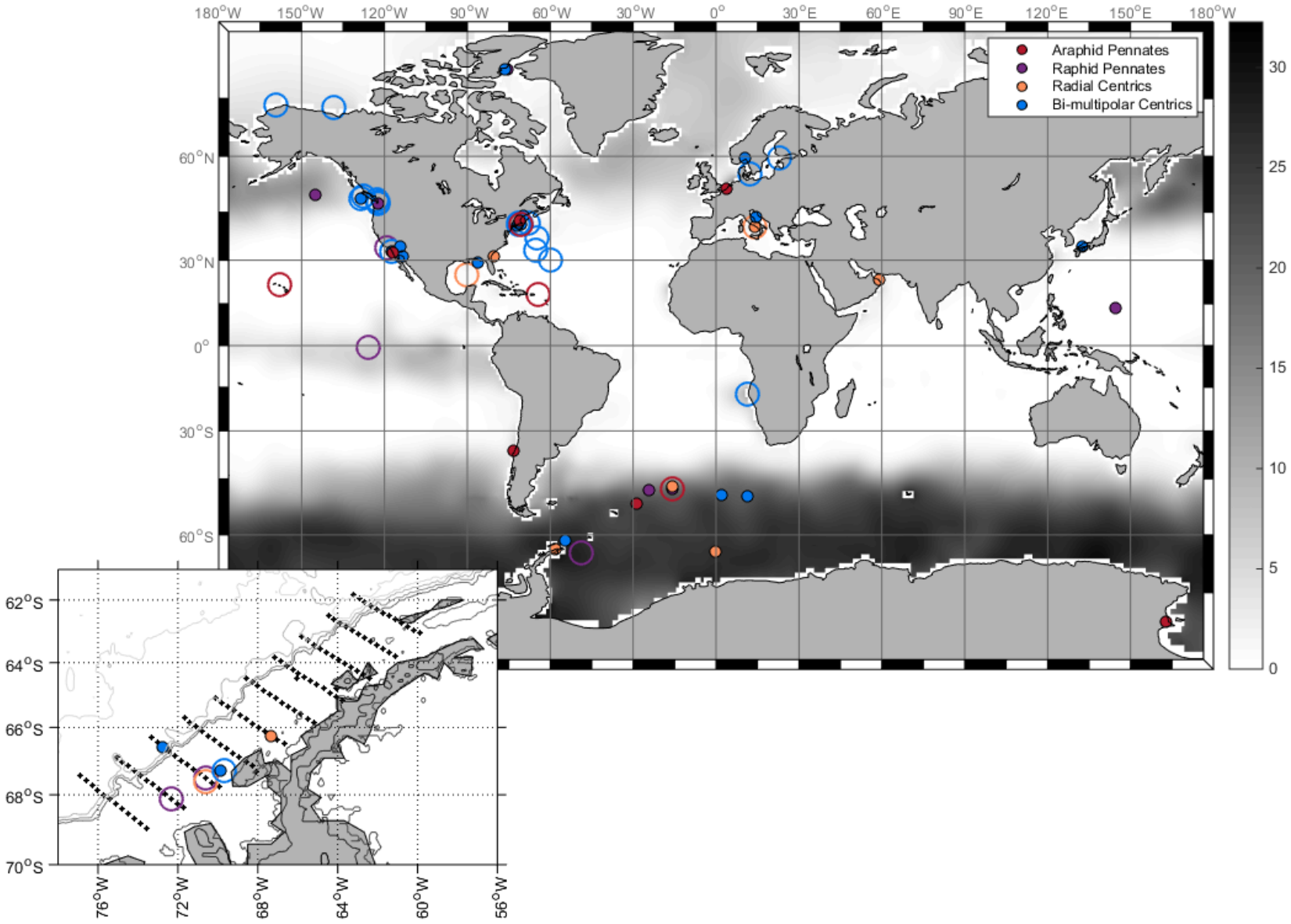
A – B_{12}-independent methionine synthase (*MetE*)



B – Flavodoxin (*FLDA*)



C – Ferredoxin (*PefF*)



APPENDIX 1: REFERENCE GENES

Appendix 1: Reference genes used in this study along with associated accession number, source repository, and diatom

Gene	Diatom	Source	Accession #
Ammonium transporter (*AMT*)	*P. tricornutum*	GenBank	ACI65096
Mitochondrial alternative oxidase (*AOX*)	*F. cylindrus*	JGI	212777
Argininosuccinate synthase (*ASSY*)	*P. tricornutum*	GenBank	ACI65191.1
Cellular repressor of EA-1 genes (*CREG*)	*P. tricornutum*	JGI	51183
Cytochrome C6 (*CYTC6*)	*P. tricornutum*	GenBank	ACI65608
Flavodoxin (*FLDA*)	*T. oceanica*	GenBank	EJK73369
Ferric reductase (*FRE*)	*P. tricornutum*	GenBank	XP_002179812
Iron permease (*FTR*)	*P. tricornutum*	GenBank	EED95805
Iron starvation induced protein 1 (*ISIP1*)	*P. tricornutum*	JGI	55031
Iron starvation induced protein 2a (*ISIP2a*)	*P. tricornutum*	GenBank	EEC48748
Iron starvation induced protein 3 (*ISIP3*)	*P. tricornutum*	JGI	47674
B_{12}-independent methionine synthase (*MetE*)	*F. cylindrus*	GenBank	AIL25366
B_{12}-dependent methionine synthase (*MetH*)	*F. kerguelensis*	GenBank	AIM62184
Manganese superoxide dismutase (*Mn-SOD*)	*F. cylindrus*	JGI	239458
Multicopper oxidase (*MUCOX*)	*F. cylindrus*	JGI	181250
Nitrite reductase (*NiR*)	*T. oceanica*	GenBank	EJK78098
Natural resistance against macrophage progein (*NRAMP*)	*F. cylindrus*	JGI	197170
Plastocyanin (*PCYN*)	*T. oceanica*	GenBank	EJK71623
Ferredoxin (*PetF)*	*T. pseudonana*	NCBI	XP_002297522
Plastid terminal oxidase (*PTOX*)	*P. tricornutum*	JGI	37377
Rhodopsin (*RHO*)	*P. granii*	GenBank	AJA37445
Ferritin (*FTN*)	*P. granii*	GenBank	ACI30660

APPENDIX 2: DIATOM LOCATION INFORMATION

Appendix 2. Location (latitude and longitude) of isolate sequenced as part of the MMETSP, and used in gene biogeography maps

Genus	Species	Strain	MMETSP ID	Latitude	Longitude
Asterionellopsis	glacialis	CCMP134	MMETSP0705	32.655	-117.14
Asterionellopsis	glacialis	CCMP1581	MMETSP1394	51.7333	3.7667
Asterionellopsis	glacialis	Unknown	MMETSP0713	41.325	-70.5667
Cyclophora	tenuis	ECT3854	MMETSP0397	21.514	-157.8362
Licmophora	paradoxa	CCMP2313	MMETSP1360	42.41915	-70.91349
Staurosira	complex sp.	CCMP2646	MMETSP1361	-36.5062	-73.1225
Synedropsis	recta cf	CCMP1620	MMETSP1176	-77.8333	163
Thalassionema	frauenfeldii	CCMP 1798	MMETSP0786	18.47	-64.575
Thalassionema	nitzschioides		MMETSP0693	41.325	-70.5667
Thalassiothrix	antarctica	L6-D1	MMETSP0152	-51.83333	-28.66667
Thalassionema	nitzschioides	L26-B	MMETSP0156	-47.86667	-15.8
Chaetoceros	affinis	CCMP159	MMETSP0088	29.42	-86.105
Chaetoceros	brevis	CCMP164	MMETSP1435	-61.2833	-54.7361
Chaetoceros	curvisetus	Unknown	MMETSP0716	32.9	-117.255
Chaetoceros	debilis	MM31A-1	MMETSP0150	-49.6	2.1
Chaetoceros	dichaeta	CCMP1751	MMETSP1447	-49.9	11.55
Chaetoceros	neogracile	CCMP1317	MMETSP0751	32.85	-117.35
Chaetoceros	sp.	UNC1202	MMETSP1429	48.816	-128.666
Cyclotella	meneghiniana	CCMP 338	MMETSP1057	43.8442	-69.641
Dactyliosolen	fragilissimus	Unknown	MMETSP0580	41.325	-70.5667
Ditylum	brightwellii	GSO103	MMETSP1002	49.654	-127.436
Ditylum	brightwellii	GSO104	MMETSP1010	41.325	-70.57
Ditylum	brightwellii	GSO105	MMETSP0998	47.74	-122.417
Eucampia	antarctica	CCMP1452	MMETSP1437	-77.8333	163
Extubocellulus	spinifer	CCMP396	MMETSP0696	31.3172	-113.56
Helicotheca	tamensis	CCMP826	MMETSP1171	30	-60
Minutocellus	polymorphus	CCMP3303	MMETSP1434	-17.5	11.25
Skeletonema	dohrnii	SkelB	MMETSP0562	41.57361	-71.40861
Skeletonema	grethea	CCMP 1804	MMETSP0578	41.4331	-71.455
Skeletonema	japonicum	CCMP2506	MMETSP0593	34.3	132.4
Skeletonema	marinoi	FE60	MMETSP1040	43.5	14.5
Skeletonema	marinoi	FE7	MMETSP1039	43.5	14.5
Skeletonema	marinoi	SM1012Hels-07	MMETSP0319	59.51	23.15
Skeletonema	marinoi	UNC1201	MMETSP1428	48.816	-128.666
Skeletonema	marinoi	skelA	MMETSP0918	41.5736105	-71.4086105
Skeletonema	menzelii	CCMP793	MMETSP0603	41.566	-70.5842
Attheya	septentrionalis	CCMP2084	MMETSP1449	77.8136	-76.3697
Chaetoceros	cf. neogracile	RCC1993	MMETSP1336	70.68333333	-138.3666667
Detonula	confervacea	CCMP 353	MMETSP1058	41.6	-71.18333333

Minutocellus	polymorphus	RCC2270	MMETSP1322	71.18333333	-159.4166667
Triceratium	dubium	CCMP147	MMETSP1175	30	-60
Corethron	pennatum	L29A3	MMETSP0171	-47.33333	-15.65
Coscinodiscus	wailesii	CCMP2513	MMETSP1066	31.4322	-80.358
Leptocylindrus	danicus var. apora	B651	MMETSP0322	40.8083	14.25
Leptocylindrus	danicus var. danicus	B650	MMETSP0321	40.8083	14.25
Leptocylindrus	danicus	CCMP1856	MMETSP1362	25	-90
Nitzschia	punctata	CCMP561	MMETSP0744	32.99	-117.255
Proboscia	alata	PI-D3	MMETSP0174	-64.0038	-0.0025
Proboscia	inermis	CCAP1064/1	MMETSP0816	-63.25	-58.3333
Pseudo-nitzschia	arenysensis	B593	MMETSP0329	40.8083	14.25
Pseudo-nitzschia	delicatissima	B596	MMETSP0327	40.8083	14.25
Rhizosolenia	setigera	CCMP 1694	MMETSP0789	23.5643	58.853
Skeletonema	marinoi	SM1012Den-03	MMETSP0320	55.5837	12.412
Thalassiosira	antarctica	CCMP982	MMETSP0902	59.5	10.6
Thalassiosira	gravida	GMp14c1	MMETSP0492	41.92732	-67.85788
Thalassiosira	miniscula	CCMP1093	MMETSP0737	32.9	-117.255
Thalassiosira	oceanica	CCMP1005	MMETSP0970	33.1833	-65.25
Thalassiosira	punctigera	Tpunct2005C2	MMETSP1067	47.6901	-122.40041
Thalassiosira	rotula	CCMP3096	MMETSP0403	49.65	-127.4338
Thalassiosira	rotula	GSO102	MMETSP0910	48.629353	-122.957542
Thalassiosira	sp.	FW	MMETSP1059	34.27944	-114.22222
Thalassiosira	weissflogii	CCMP1010	MMETSP1405	37	-65
Thalassiosira	weissflogii	CCMP1336	MMETSP0878	41.11	-72.1
Amphiprora	paludosa	CCMP125	MMETSP1065	41.525	-70.6736
Amphiprora	sp.	CCMP467	MMETSP0724	32.9	-117.255
Amphora	coffeaeformis	CCMP127	MMETSP0316	41.525	-70.6736
Entomoneis	sp.	CCMP2396	MMETSP1443	77.8333	-75.55
Fragilariopsis	cylindrus	CCMP1102		-64.08	-48.7033
Fragilariopsis	kerguelensis	L2-C3	MMETSP0906	-48.1	-24.25
Fragilariopsis	kerguelensis	L26-C5	MMETSP0733	-48	-16
Fragilariopsis	radiata	13vi08-1A	MMETSP0418	13.4427	144.6428
Pseudo-nitzschia	delicatissima	UNC1205	MMETSP1432	48.816	-128.666
Pseudo-nitzschia	fraudulenta	WWA7	MMETSP0850	34.0847	-119.05
Pseudo-nitzschia	granii	UNC1204		50	-145
Pseudo-nitzschia	heimii	UNC1101	MMETSP1423	50	-145
Pseudo-nitzschia	pungens	cf. cingulata	MMETSP1060	47.6901	-122.40041
Pseudo-nitzschia	pungens	cf. pungens	MMETSP1061	47.6901	-122.40041
Stauroneis	constricta	CCMP1120	MMETSP1352	-0.7475	-126.03

APPENDIX 3: PetF SUPPLEMENTARY INFORMATION

Appendix 3: MMETSP diatom BLASTp results of PetF with e-value, latitude and longitude. Diatom identifiers include strain ID, MMETSP ID, and subject ID. (-) indicates no BLAST result.

Genus	Species	Strain	MMETSP_ID	Subject ID	e-value	Latitude	Longitude	Clade
Thalassionema	nitzschioides	L26-B	MMETSP0156	CAMPEP_0194203890	4.00E-40	-47.867	-15.800	Araphid Pennate
Thalassionema	frauenfeldii	CCMP 1798	MMETSP0786	CAMPEP_0178916612	2.00E-38	18.470	-64.575	Araphid Pennate
Asterionellopsis	glacialis	Unknown	MMETSP0713	CAMPEP_0184886684	2.00E-37	41.325	-70.567	Araphid Pennate
Thalassionema	nitzschioides		MMETSP0693	CAMPEP_0205839574	3.00E-37	41.325	-70.567	Araphid Pennate
Cyclophora	tenuis	ECT3854	MMETSP0397	CAMPEP_0116568138	1.00E-27	21.514	-157.836	Araphid Pennate
Asterionellopsis	glacialis	CCMP134	MMETSP0705	-	-	32.655	-117.140	Araphid Pennate
Asterionellopsis	glacialis	CCMP1581	MMETSP1394	-	-	51.733	3.767	Araphid Pennate
Licmophora	paradoxa	CCMP2313	MMETSP1360	-	-	42.419	-70.913	Araphid Pennate
Staurosira	complex sp.	CCMP2646	MMETSP1361	-	-	-36.506	-73.123	Araphid Pennate
Synedropsis	recta cf	CCMP1620	MMETSP1176	-	-	-77.833	163.000	Araphid Pennate
Thalassiothrix	antarctica	L6-D1	MMETSP0152	-	-	-51.833	-28.667	Araphid Pennate
Thalassiosira	weissflogii	CCMP1010	MMETSP1408	CAMPEP_0203187078	2.00E-64	37.000	-65.000	Bi(multi) polar centric
Minutocellus	polymorphus	CCMP3303	MMETSP1434	CAMPEP_0197725580	2.00E-48	-17.500	11.250	Bi(multi) polar centric
Minutocellus	polymorphus	RCC2270	MMETSP1322	CAMPEP_0185799760	3.00E-47	71.183	-159.417	Bi(multi) polar centric
Skeletonema	menzelii	CCMP793	MMETSP0603	CAMPEP_0183648216	8.00E-46	41.566	-70.584	Bi(multi) polar centric
Thalassiosira	gravida	GMp14c1	MMETSP0492	CAMPEP_0201598230	9.00E-46	41.927	-67.858	Bi(multi) polar centric
Thalassiosira	rotula	GSO102	MMETSP0910	CAMPEP_0196129530	3.00E-45	48.629	-122.958	Bi(multi) polar centric
Thalassiosira	miniscula	CCMP1093	MMETSP0737	CAMPEP_0183720344	3.00E-44	32.900	-117.255	Bi(multi) polar centric
Skeletonema	dohrnii	SkelB	MMETSP0562	CAMPEP_0206646526	1.00E-43	41.574	-71.409	Bi(multi) polar centric
Skeletonema	marinoi	UNC1201	MMETSP1428	CAMPEP_0197216810	1.00E-43	48.816	-128.666	Bi(multi) polar centric
Thalassiosira	oceanica	CCMP1005	MMETSP0970	CAMPEP_0206708282	1.00E-43	33.183	-65.250	Bi(multi) polar centric
Thalassiosira	rotula	CCMP3096	MMETSP0403	CAMPEP_0184774218	1.00E-43	49.650	-127.434	Bi(multi) polar centric
Skeletonema	marinoi	SM1012Den-03	MMETSP0320	CAMPEP_0115946752	5.00E-43	55.584	12.412	Bi(multi) polar centric
Skeletonema	marinoi	SM1012Hels-07	MMETSP0319	CAMPEP_0115919778	1.00E-42	59.510	23.150	Bi(multi) polar centric

Skeletonema	marinoi	skelA	MMETSP0918	CAMPEP_0184911864	4.00E-42	41.574	-71.409	Bi(multi) polar centric
Ditylum	brightwellii	GSO104	MMETSP1013	CAMPEP_0204810610	9.00E-41	41.325	-70.570	Bi(multi) polar centric
Thalassiosira	punctigera	Tpunct2005C2	MMETSP1067	CAMPEP_0172528362	3.00E-39	47.690	-122.400	Bi(multi) polar centric
Triceratium	dubium	CCMP147	MMETSP1175	CAMPEP_0197432434	8.00E-38	30.000	-60.000	Bi(multi) polar centric
Chaetoceros	cf. neogracile	RCC1993	MMETSP1336	CAMPEP_0119449874	2.00E-37	70.683	-138.367	Bi(multi) polar centric
Skeletonema	grethea	CCMP 1804	MMETSP0578	CAMPEP_0201696492	2.00E-37	41.433	-71.455	Bi(multi) polar centric
Ditylum	brightwellii	GSO103	MMETSP1002	CAMPEP_0204662216	1.00E-36	49.654	-127.436	Bi(multi) polar centric
Ditylum	brightwellii	GSO105	MMETSP1001	CAMPEP_0185671684	7.00E-33	47.740	-122.417	Bi(multi) polar centric
Helicotheca	tamensis	CCMP826	MMETSP1171	CAMPEP_0185732772	1.00E-28	30.000	-60.000	Bi(multi) polar centric
Chaetoceros	neogracile	CCMP1317	MMETSP0754	CAMPEP_0195496542	6.00E-11	32.850	-117.350	Bi(multi) polar centric
Attheya	septentrionalis	CCMP2084	MMETSP1449	-	-	77.814	-76.370	Bi(multi) polar centric
Chaetoceros	affinis	CCMP159	MMETSP0088	-	-	29.420	-86.105	Bi(multi) polar centric
Chaetoceros	brevis	CCMP164	MMETSP1435	-	-	-61.283	-54.736	Bi(multi) polar centric
Chaetoceros	curvisetus	Unknown	MMETSP0716	-	-	32.900	-117.255	Bi(multi) polar centric
Chaetoceros	debilis	MM31A-1	MMETSP0149	-	-	-49.600	2.100	Bi(multi) polar centric
Chaetoceros	dichaeta	CCMP1751	MMETSP1447	-	-	-49.900	11.550	Bi(multi) polar centric
Chaetoceros	sp.	UNC1202	MMETSP1429	-	-	48.816	-128.666	Bi(multi) polar centric
Cyclotella	meneghiniana	CCMP 338	MMETSP1057	-	-	43.844	-69.641	Bi(multi) polar centric
Dactyliosolen	fragilissimus	Unknown	MMETSP0580	-	-	41.325	-70.567	Bi(multi) polar centric
Eucampia	antarctica	CCMP1452	MMETSP1437	-	-	-77.833	163.000	Bi(multi) polar centric
Extubocellulus	spinifer	CCMP396	MMETSP0696	-	-	31.317	-113.560	Bi(multi) polar centric
Skeletonema	japonicum	CCMP2506	MMETSP0593	-	-	34.300	132.400	Bi(multi) polar centric
Skeletonema	marinoi	FE60	MMETSP1040	-	-	43.500	14.500	Bi(multi) polar centric
Skeletonema	marinoi	FE7	MMETSP1039	-	-	43.500	14.500	Bi(multi) polar centric
Thalassiosira	antarctica	CCMP982	MMETSP0902	-	-	59.500	10.600	Bi(multi) polar centric
Thalassiosira	sp.	FW	MMETSP1059	-	-	34.279	-114.222	Bi(multi) polar centric
Thalassiosira	weissflogii	CCMP1336	MMETSP0878	-	-	41.110	-72.100	Bi(multi) polar centric
Detonula	confervacea	CCMP 353	MMETSP1058	-	-	41.600	-71.400	Radial Centric
Leptocylindrus	danicus var. danicus	B650	MMETSP0321	CAMPEP_0116040274	2.00E-27	40.808	14.250	Radial Centric
Leptocylindrus	danicus	CCMP1856	MMETSP1362	CAMPEP_0196810402	9.00E-07	25.000	-90.000	Radial Centric

Coscinodiscus	wailesii	CCMP2513	MMETSP1066	-	-	31.432	-80.358	Radial Centric
Leptocylindrus	danicus var. apora	B651	MMETSP0322	-	-	40.808	14.250	Radial Centric
Nitzschia	punctata	CCMP561	MMETSP0744	-	-	32.990	-117.255	Radial Centric
Proboscia	alata	PI-D3	MMETSP0174	-	-	-64.004	-0.003	Radial Centric
Proboscia	inermis	CCAP1064/1	MMETSP0816	-	-	-63.250	-58.333	Radial Centric
Rhizosolenia	setigera	CCMP 1694	MMETSP0789	-	-	23.564	58.853	Radial Centric
Corethron	pennatum	L29A3	MMETSP0169	-	-	-47.333	-15.650	Radial Centric
Fragilariopsis	cylindrus	CCMP1102		-	0	-64.080	-48.703	Raphid Pennate
Amphiprora	sp.	CCMP467	MMETSP0724	CAMPEP_0168749370	2.00E-50	32.900	-117.255	Raphid Pennate
Pseudo-nitzschia	fraudulenta	WWA7	MMETSP0851	CAMPEP_0201200450	5.00E-39	34.085	-119.050	Raphid Pennate
Pseudo-nitzschia	delicatissima	UNC1205	MMETSP1432	CAMPEP_0197277910	3.00E-22	48.816	-128.666	Raphid Pennate
Stauroneis	constricta	CCMP1120	MMETSP1352	CAMPEP_0119563612	3.00E-18	-0.748	-126.030	Raphid Pennate
Amphora	coffeaeformis	CCMP127	MMETSP0316	-	-	41.525	-70.674	Raphid Pennate
Entomoneis	sp.	CCMP2396	MMETSP1443	-	-	77.833	-75.550	Raphid Pennate
Fragilariopsis	kerguelensis	L2-C3	MMETSP0906	-	-	-48.100	-24.250	Raphid Pennate
Fragilariopsis	kerguelensis	L26-C5	MMETSP0733	-	-	-48.000	-16.000	Raphid Pennate
Fragilariopsis	radiata	13vi08-1A	MMETSP0418	-	-	13.443	144.643	Raphid Pennate
Pseudo-nitzschia	arenysensis	B593	MMETSP0329	-	-	40.808	14.250	Raphid Pennate
Pseudo-nitzschia	delicatissima	B596	MMETSP0327	-	-	40.808	14.250	Raphid Pennate
Pseudo-nitzschia	granii	UNC1204		-	-	50.000	-145.000	Raphid Pennate
Pseudo-nitzschia	heimii	UNC1101	MMETSP1423	-	-	50.000	-145.000	Raphid Pennate
Pseudo-nitzschia	pungens	cf. cingulata	MMETSP1060	-	-	47.690	-122.400	Raphid Pennate
Pseudo-nitzschia	pungens	cf. pungens	MMETSP1061	-	-	47.690	-122.400	Raphid Pennate
Amphiprora	paludosa	CCMP125	MMETSP1065	-	-	41.525	-70.674	Raphid Pennate

APPENDIX 4: FLDA SUPPLEMENTARY INFORMATION

Appendix 4: MMETSP diatom BLASTp results of FLDA with e-value, latitude and longitude. Diatom identifiers include strain ID, MMETSP ID, and subject ID. (-) indicates no BLAST result.

Genus	Species	Strain	MMETSP_ID	Subject ID	e-value	Latitude	Longitude	Clade
Asterionellopsis	glacialis	Unknown	MMETSP0713	CAMPEP_0184878458	3.00E-94	41.325	-70.5667	Araphid Pennate
Cyclophora	tenuis	ECT3854	MMETSP0397	CAMPEP_0116542030	7.00E-92	21.514	-157.8362	Araphid Pennate
Asterionellopsis	glacialis	CCMP1581	MMETSP1394	CAMPEP_0197151076	2.00E-86	51.7333	3.7667	Araphid Pennate
Asterionellopsis	glacialis	CCMP134	MMETSP0705	CAMPEP_0205131002	2.00E-82	32.655	-117.14	Araphid Pennate
Synedropsis	recta cf	CCMP1620	MMETSP1176	CAMPEP_0119014672	2.00E-75	-77.8333	163	Araphid Pennate
Thalassionema	nitzschioides	L26-B	MMETSP0156	CAMPEP_0194213668	3.00E-71	-47.86667	-15.8	Araphid Pennates
Thalassionema	frauenfeldii	CCMP 1798	MMETSP0786	CAMPEP_0178927282	4.00E-62	18.47	-64.575	Araphid Pennate
Thalassiothrix	antarctica	L6-D1	MMETSP0152	CAMPEP_0194133880	9.00E-50	-51.83333	-28.66667	Araphid Pennate
Licmophora	paradoxa	CCMP2313	MMETSP1360	CAMPEP_0202453368	3.00E-48	42.41915	-70.91349	Araphid Pennate
Thalassionema	nitzschioides		MMETSP0693	CAMPEP_0205835806	8.00E-46	41.325	-70.5667	Araphid Pennate
Staurosira	complex sp.	CCMP2646	MMETSP1361	-	-	-36.5062	-73.1225	Araphid Pennate
Attheya	septentrionalis	CCMP2084	MMETSP1449	CAMPEP_0198294680	5.00E-93	77.8136	-76.3697	Bi(multi) polar centric
Thalassiosira	oceanica	CCMP1005	MMETSP0970	CAMPEP_0206693386	0	33.1833	-65.25	Bi(multi) polar centric
Thalassiosira	rotula	GSO102	MMETSP0910	CAMPEP_0196143042	2.00E-116	48.629353	-122.957542	Bi(multi) polar centric
Thalassiosira	rotula	CCMP3096	MMETSP0403	CAMPEP_0184777058	2.00E-115	49.65	-127.4338	Bi(multi) polar centric
Thalassiosira	gravida	GMp14c1	MMETSP0492	CAMPEP_0201603546	1.00E-111	41.92732	-67.85788	Bi(multi) polar centric
Thalassiosira	antarctica	CCMP982	MMETSP0902	CAMPEP_0201885636	2.00E-108	59.5	10.6	Bi(multi) polar centric
Detonula	confervacea	CCMP 353	MMETSP1058	CAMPEP_0172315214	9.00E-105	41.6	-71.4	Polar Centric
Thalassiosira	punctigera	Tpunct2005C2	MMETSP1067	CAMPEP_0172527408	8.00E-104	47.6901	-122.40041	Bi(multi) polar centric
Cyclotella	meneghiniana	CCMP 338	MMETSP1057	CAMPEP_0172291170	2.00E-102	43.8442	-69.641	Bi(multi) polar centric
Skeletonema	marinoi	FE7	MMETSP1039	CAMPEP_0180848120	3.00E-101	43.5	14.5	Bi(multi) polar centric
Dactyliosolen	fragilissimus	Unknown	MMETSP0580	CAMPEP_0184854792	6.00E-101	41.325	-70.5667	Bi(multi) polar centric
Skeletonema	marinoi	FE60	MMETSP1040	CAMPEP_0180892888	6.00E-100	43.5	14.5	Bi(multi) polar centric
Skeletonema	dohrnii	SkelB	MMETSP0562	CAMPEP_0206660158	3.00E-96	41.57361	-71.40861	Bi(multi) polar centric
Eucampia	antarctica	CCMP1452	MMETSP1437	CAMPEP_0197836854	1.00E-95	-77.8333	163	Bi(multi) polar centric

Minutocellus	polymorphus	CCMP3303	MMETSP1434	CAMPEP_0197726062	2.00E-95	-17.5	11.25	Bi(multi) polar centric
Chaetoceros	curvisetus	Unknown	MMETSP0717	CAMPEP_0204622136	3.00E-93	32.9	-117.255	Bi(multi) polar centric
Chaetoceros	sp.	UNC1202	MMETSP1429	CAMPEP_0197237438	2.00E-91	48.816	-128.666	Bi(multi) polar centric
Chaetoceros	cf. neogracile	RCC1993	MMETSP1336	CAMPEP_0119447330	3.00E-89	70.68333333	-138.3666667	Bi(multi) polar centric
Chaetoceros	neogracile	CCMP1317	MMETSP0751	CAMPEP_0195359084	5.00E-89	32.85	-117.35	Bi(multi) polar centric
Chaetoceros	affinis	CCMP159	MMETSP0088	CAMPEP_0203646006	9.00E-87	29.42	-86.105	Bi(multi) polar centric
Chaetoceros	dichaeta	CCMP1751	MMETSP1447	CAMPEP_0198264464	7.00E-83	-49.9	11.55	Bi(multi) polar centric
Chaetoceros	brevis	CCMP164	MMETSP1435	CAMPEP_0197738932	1.00E-46	-61.2833	-54.7361	Bi(multi) polar centric
Extubocellulus	spinifer	CCMP396	MMETSP0696	CAMPEP_0178508480	1.00E-46	31.3172	-113.56	Bi(multi) polar centric
Thalassiosira	miniscula	CCMP1093	MMETSP0737	CAMPEP_0183702364	1.00E-46	32.9	-117.255	Bi(multi) polar centric
Chaetoceros	debilis	MM31A-1	MMETSP0149	CAMPEP_0194083366	2.00E-46	-49.6	2.1	Bi(multi) polar centric
Ditylum	brightwellii	GSO105	MMETSP0998	CAMPEP_0185650814	3.00E-46	47.74	-122.417	Bi(multi) polar centric
Ditylum	brightwellii	GSO104	MMETSP1010	CAMPEP_0204738864	1.00E-43	41.325	-70.57	Bi(multi) polar centric
Skeletonema	japonicum	CCMP2506	MMETSP0593	CAMPEP_0201723242	1.00E-43	34.3	132.4	Bi(multi) polar centric
Ditylum	brightwellii	GSO103	MMETSP1002	CAMPEP_0204687074	3.00E-43	49.654	-127.436	Bi(multi) polar centric
Helicotheca	tamensis	CCMP826	MMETSP1171	CAMPEP_0185726538	3.00E-42	30	-60	Bi(multi) polar centric
Skeletonema	marinoi	skelA	MMETSP0918	CAMPEP_0184919536	1.00E-41	41.5736105	-71.4086105	Bi(multi) polar centric
Thalassiosira	weissflogii	CCMP1010	MMETSP1405	CAMPEP_0203098210	1.00E-41	37	-65	Bi(multi) polar centric
Skeletonema	grethea	CCMP 1804	MMETSP0578	CAMPEP_0201702732	4.00E-41	41.4331	-71.455	Bi(multi) polar centric
Triceratium	dubium	CCMP147	MMETSP1175	CAMPEP_0197444866	4.00E-40	30	-60	Bi(multi) polar centric
Thalassiosira	weissflogii	CCMP1336	MMETSP0878	CAMPEP_0171354524	3.00E-38	41.11	-72.1	Bi(multi) polar centric
Minutocellus	polymorphus	RCC2270	MMETSP1322	-	-	71.18333333	-159.4166667	Bi(multi) polar centric
Skeletonema	marinoi	SM1012Den-03	MMETSP0320	-	-	55.5837	12.412	Bi(multi) polar centric
Skeletonema	marinoi	SM1012Hels-07	MMETSP0319	-	-	59.51	23.15	Bi(multi) polar centric
Skeletonema	marinoi	UNC1201	MMETSP1428	-	-	48.816	-128.666	Bi(multi) polar centric
Skeletonema	menzelii	CCMP793	MMETSP0603	-	-	41.566	-70.5842	Bi(multi) polar centric
Thalassiosira	sp.	FW	MMETSP1059	-	-	34.27944	-114.22222	Bi(multi) polar centric
Corethron	pennatum	L29A3	MMETSP0169	CAMPEP_0194266102	5.00E-43	-47.33333	-15.65	Radial Centric
Coscinodiscus	wailesii	CCMP2513	MMETSP1066	CAMPEP_0172504384	5.00E-106	31.4322	-80.358	Radial Centric

Leptocylindrus	danicus var. apora	B651	MMETSP0322	CAMPEP_0116050568	1.00E-90	40.8083	14.25	Radial Centric
Leptocylindrus	danicus var. danicus	B650	MMETSP0321	CAMPEP_0116026706	2.00E-79	40.8083	14.25	Radial Centric
Leptocylindrus	danicus	CCMP1856	MMETSP1362	CAMPEP_0196811168	2.00E-79	25	-90	Radial Centric
Nitzschia	punctata	CCMP561	MMETSP0744	CAMPEP_0178742078	5.00E-74	32.99	-117.255	Radial Centric
Proboscia	alata	PI-D3	MMETSP0174	CAMPEP_0194371862	2.00E-42	-64.0038	-0.0025	Radial Centric
Proboscia	inermis	CCAP1064/1	MMETSP0816	CAMPEP_0171310236	9.00E-73	-63.25	-58.3333	Radial Centric
Rhizosolenia	setigera	CCMP 1694	MMETSP0789	CAMPEP_0178950596	4.00E-81	23.5643	58.853	Radial Centric
Amphiprora	paludosa	CCMP125	MMETSP1065	CAMPEP_0172470380	9.00E-43	41.525	-70.6736	Raphid Pennate
Pseudo-nitzschia	heimii	UNC1101	MMETSP1423	CAMPEP_0197189390	4.00E-84	50	-145	Raphid Pennate
Pseudo-nitzschia	fraudulenta	WWA7	MMETSP0850	CAMPEP_0201117996	7.00E-84	34.0847	-119.05	Raphid Pennate
Entomoneis	sp.	CCMP2396	MMETSP1443	CAMPEP_0198146148	9.00E-81	77.8333	-75.55	Raphid Pennate
Amphiprora	sp.	CCMP467	MMETSP0724	CAMPEP_0168749830	1.00E-79	32.9	-117.255	Raphid Pennate
Stauroneis	constricta	CCMP1120	MMETSP1352	CAMPEP_0119562926	2.00E-75	-0.7475	-126.03	Raphid Pennate
Fragilariopsis	radiata	13vi08-1A	MMETSP0418	CAMPEP_0116846204	1.00E-52	13.4427	144.6428	Raphid Pennate
Fragilariopsis	kerguelensis	L2-C3	MMETSP0906	CAMPEP_0196001262	4.00E-44	-48.1	-24.25	Raphid Pennate
Amphora	coffeaeformis	CCMP127	MMETSP0316	CAMPEP_0170673568	2.00E-37	41.525	-70.6736	Raphid Pennate
Fragilariopsis	kerguelensis	L26-C5	MMETSP0733	CAMPEP_0170785742	2.00E-37	-48	-16	Raphid Pennate
Fragilariopsis	cylindrus	CCMP1102		-	0	-64.08	-48.7033	Raphid Pennate
Pseudo-nitzschia	granii	UNC1204		-	0	50	-145	Raphid Pennate
Pseudo-nitzschia	arenysensis	B593	MMETSP0329	-	-	40.8083	14.25	Raphid Pennate
Pseudo-nitzschia	delicatissima	B596	MMETSP0327	-	-	40.8083	14.25	Raphid Pennate
Pseudo-nitzschia	delicatissima	UNC1205	MMETSP1432	-	-	48.816	-128.666	Raphid Pennate
Pseudo-nitzschia	pungens	cf. cingulata	MMETSP1060	-	-	47.6901	-122.40041	Raphid Pennate
Pseudo-nitzschia	pungens	cf. pungens	MMETSP1061	-	-	47.6901	-122.40041	Raphid Pennate

APPENDIX 5: MetE SUPPLEMENTARY INFORMATION

Appendix 5: MMETSP diatom BLASTp results of MetE with e-value, latitude and longitude. Diatom identifiers include strain ID, MMETSP ID, and subject ID. (-) indicates no BLAST result.

Genus	Species	Strain	MMETSP_ID	Subject ID	e-value	Latitude	Longitude	Clade
Asterionellopsis	glacialis	CCMP134	MMETSP0705	-	-	32.655	-117.140	Araphid Pennate
Asterionellopsis	glacialis	CCMP1581	MMETSP1394	-	-	51.733	3.767	Araphid Pennate
Asterionellopsis	glacialis	Unknown	MMETSP0713	-	-	41.325	-70.567	Araphid Pennate
Cyclophora	tenuis	ECT3854	MMETSP0397	-	-	21.514	-157.836	Araphid Pennate
Licmophora	paradoxa	CCMP2313	MMETSP1360	-	-	42.419	-70.913	Araphid Pennate
Staurosira	complex sp.	CCMP2646	MMETSP1361	CAMPEP_0202487798	0	-36.506	-73.123	Araphid Pennate
Synedropsis	recta cf	CCMP1620	MMETSP1176	-	-	-77.833	163.000	Araphid Pennate
Thalassionema	frauenfeldii	CCMP 1798	MMETSP0786	-	-	18.470	-64.575	Araphid Pennate
Thalassionema	nitzschioides		MMETSP0693	-	-	41.325	-70.567	Araphid Pennate
Thalassiothrix	antarctica	L6-D1	MMETSP0152	-	-	-51.833	-28.667	Araphid Pennate
Thalassionema	nitzschioides	L26-B	MMETSP0156	CAMPEP_0194228602	0	-47.867	-15.800	Araphid Pennates
Chaetoceros	affinis	CCMP159	MMETSP0088	-	-	29.420	-86.105	Bi(multi) polar centric
Chaetoceros	brevis	CCMP164	MMETSP1435	-	-	-61.283	-54.736	Bi(multi) polar centric
Chaetoceros	curvisetus	Unknown	MMETSP0716	-	-	32.900	-117.255	Bi(multi) polar centric
Chaetoceros	debilis	MM31A-1	MMETSP0150	CAMPEP_0194126000	0	-49.600	2.100	Bi(multi) polar centric
Chaetoceros	dichaeta	CCMP1751	MMETSP1447	-	-	-49.900	11.550	Bi(multi) polar centric
Chaetoceros	neogracile	CCMP1317	MMETSP0751	-	-	32.850	-117.350	Bi(multi) polar centric
Chaetoceros	sp.	UNC1202	MMETSP1429	-	-	48.816	-128.666	Bi(multi) polar centric
Cyclotella	meneghiniana	CCMP 338	MMETSP1057	-	-	43.844	-69.641	Bi(multi) polar centric
Dactyliosolen	fragilissimus	Unknown	MMETSP0580	-	-	41.325	-70.567	Bi(multi) polar centric
Ditylum	brightwellii	GSO103	MMETSP1002	-	-	49.654	-127.436	Bi(multi) polar centric
Ditylum	brightwellii	GSO104	MMETSP1010	-	-	41.325	-70.570	Bi(multi) polar centric
Ditylum	brightwellii	GSO105	MMETSP0998	-	-	47.740	-122.417	Bi(multi) polar centric
Eucampia	antarctica	CCMP1452	MMETSP1437	-	-	-77.833	163.000	Bi(multi) polar centric
Extubocellulus	spinifer	CCMP396	MMETSP0696	-	-	31.317	-113.560	Bi(multi) polar centric

Helicotheca	tamensis	CCMP826	MMETSP1171	-	-	30.000	-60.000	Bi(multi) polar centric
Minutocellus	polymorphus	CCMP3303	MMETSP1434	-	-	-17.500	11.250	Bi(multi) polar centric
Skeletonema	dohrnii	SkelB	MMETSP0562	-	-	41.574	-71.409	Bi(multi) polar centric
Skeletonema	grethea	CCMP 1804	MMETSP0578	-	-	41.433	-71.455	Bi(multi) polar centric
Skeletonema	japonicum	CCMP2506	MMETSP0593	-	-	34.300	132.400	Bi(multi) polar centric
Skeletonema	marinoi	FE60	MMETSP1040	-	-	43.500	14.500	Bi(multi) polar centric
Skeletonema	marinoi	FE7	MMETSP1039	-	-	43.500	14.500	Bi(multi) polar centric
Skeletonema	marinoi	SM1012Hels-07	MMETSP0319	-	-	59.510	23.150	Bi(multi) polar centric
Skeletonema	marinoi	UNC1201	MMETSP1428	-	-	48.816	-128.666	Bi(multi) polar centric
Skeletonema	marinoi	skelA	MMETSP0918	-	-	41.574	-71.409	Bi(multi) polar centric
Skeletonema	menzelii	CCMP793	MMETSP0603	-	-	41.566	-70.584	Bi(multi) polar centric
Attheya	septentrionalis	CCMP2084	MMETSP1449	CAMPEP_0198287942	0	77.814	-76.370	Bi(multi) polar centric
Chaetoceros	cf. neogracile	RCC1993	MMETSP1336	-	-	70.683	-138.367	Bi(multi) polar centric
Detonula	confervacea	CCMP 353	MMETSP1058	-	-	41.600	-71.183	Bi(multi) polar centric
Minutocellus	polymorphus	RCC2270	MMETSP1322	CAMPEP_0185820162	3.00E-53	71.183	-159.417	Bi(multi) polar centric
Triceratium	dubium	CCMP147	MMETSP1175	-	-	30.000	-60.000	Bi(multi) polar centric
Corethron	pennatum	L29A3	MMETSP0171	CAMPEP_0194315660	0	-47.333	-15.650	Radial Centric
Coscinodiscus	wailesii	CCMP2513	MMETSP1066	-	-	31.432	-80.358	Radial Centric
Leptocylindrus	danicus var. apora	B651	MMETSP0322	-	-	40.808	14.250	Radial Centric
Leptocylindrus	danicus var. danicus	B650	MMETSP0321	-	-	40.808	14.250	Radial Centric
Leptocylindrus	danicus	CCMP1856	MMETSP1362	-	-	25.000	-90.000	Radial Centric
Nitzschia	punctata	CCMP561	MMETSP0744	-	-	32.990	-117.255	Radial Centric
Proboscia	alata	PI-D3	MMETSP0174	CAMPEP_0194374316	0	-64.004	-0.003	Radial Centric
Proboscia	inermis	CCAP1064/1	MMETSP0816	CAMPEP_0171311326	0	-63.250	-58.333	Radial Centric
Pseudo-nitzschia	arenysensis	B593	MMETSP0329	-	-	40.808	14.250	Radial Centric
Pseudo-nitzschia	delicatissima	B596	MMETSP0327	-	-	40.808	14.250	Radial Centric
Rhizosolenia	setigera	CCMP 1694	MMETSP0789	CAMPEP_0178977680	1.00E-31	23.564	58.853	Radial Centric
Skeletonema	marinoi	SM1012Den-03	MMETSP0320	-	-	55.584	12.412	Radial Centric

Thalassiosira	antarctica	CCMP982	MMETSP0902	-	-	59.500	10.600	Radial Centric
Thalassiosira	gravida	GMp14c1	MMETSP0492	-	-	41.927	-67.858	Radial Centric
Thalassiosira	miniscula	CCMP1093	MMETSP0737	-	-	32.900	-117.255	Radial Centric
Thalassiosira	oceanica	CCMP1005	MMETSP0970	-	-	33.183	-65.250	Radial Centric
Thalassiosira	punctigera	Tpunct2005C2	MMETSP1067	-	-	47.690	-122.400	Radial Centric
Thalassiosira	rotula	CCMP3096	MMETSP0403	-	-	49.650	-127.434	Radial Centric
Thalassiosira	rotula	GSO102	MMETSP0910	-	-	48.629	-122.958	Radial Centric
Thalassiosira	sp.	FW	MMETSP1059	-	-	34.279	-114.222	Radial Centric
Thalassiosira	weissflogii	CCMP1010	MMETSP1405	-	-	37.000	-65.000	Radial Centric
Thalassiosira	weissflogii	CCMP1336	MMETSP0878	-	-	41.110	-72.100	Radial Centric
Amphiprora	paludosa	CCMP125	MMETSP1065	CAMPEP_0172449172	7.00E-41	41.525	-70.674	Raphid Pennate
Amphiprora	sp.	CCMP467	MMETSP0724	CAMPEP_0168748186	0	32.900	-117.255	Raphid Pennate
Amphora	coffeaeformis	CCMP127	MMETSP0316	-	-	41.525	-70.674	Raphid Pennate
Entomoneis	sp.	CCMP2396	MMETSP1443	-	-	77.833	-75.550	Raphid Pennate
Fragilariopsis	cylindrus	CCMP1102		-	0	-64.080	-48.703	Raphid Pennate
Fragilariopsis	kerguelensis	L2-C3	MMETSP0906	CAMPEP_0196006278	0	-48.100	-24.250	Raphid Pennate
Fragilariopsis	kerguelensis	L26-C5	MMETSP0733	CAMPEP_0170764404	0	-48.000	-16.000	Raphid Pennate
Fragilariopsis	radiata	13vi08-1A	MMETSP0418	CAMPEP_0116868728	0	13.443	144.643	Raphid Pennate
Pseudo-nitzschia	delicatissima	UNC1205	MMETSP1432	-	-	48.816	-128.666	Raphid Pennate
Pseudo-nitzschia	fraudulenta	WWA7	MMETSP0850	-	-	34.085	-119.050	Raphid Pennate
Pseudo-nitzschia	granii	UNC1204		-	-	50.000	-145.000	Raphid Pennate
Pseudo-nitzschia	heimii	UNC1101	MMETSP1423	-	-	50.000	-145.000	Raphid Pennate
Pseudo-nitzschia	pungens	cf. cingulata	MMETSP1060	-	-	47.690	-122.400	Raphid Pennate
Pseudo-nitzschia	pungens	cf. pungens	MMETSP1061	-	-	47.690	-122.400	Raphid Pennate
Stauroneis	constricta	CCMP1120	MMETSP1352	-	-	-0.748	-126.030	Raphid Pennate

REFERENCES

Agusti S. and Duarte C.M. 2000. Experimental induction of a large phytoplankton bloom in Antarctic coastal waters. Mar. Ecol. Prog. Ser. 206: 73–85.

Altschul S., et al. (1990) Basic Local Alignment Search Tool Journal of Molecular Biology 215(3):403-410.

Andersen RA. (Ed. 2005). *Algal Culturing Techniques: A Book for All Phycologists*. Elsevier.

Annett AL., Damien SC., Crosta X., Clarke A., Raja S. Ganeshram RS. (2010). Seasonal progression of diatom assemblages in surface waters of Ryder Bay, Antarctica. Polar Biol. 33:13–29

Armstrong, R. A., C. Lee, J. I. Hedges, S. Honjo, and S. G. Wakeham (2002), A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative Andersen, Robert A., (Ed.) 2005. Algal Culturing Techniques: A Book for All Phycologists. Elsevier. ISBN: 0-12-088426-7

Arrigo, K. R., van Dijken, G., & Long, M. (2008). Coastal Southern Ocean: A strong anthropogenic CO_2 sink. *Geophysical Research Letters*, *35*(21), L21602. doi:10.1029/2008GL035624

Arrigo, K. R., Mills, M. M., Kropuenske, L. R., van Dijken, G. L., Alderkamp, A.-C., & Robinson, D. H. (2010). Photophysiology in two major southern ocean phytoplankton taxa: photosynthesis and growth of Phaeocystis antarctica and Fragilariopsis cylindrus under different irradiance levels. *Integrative and Comparative Biology*, *50*(6), 950–66. doi:10.1093/icb/icq021

Assmy, P., Smetacek, V., Montresor, M., Klaas, C., Henjes, J., Strass, V. H., Wolf-Gladrow, D. (2013). Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic Circumpolar Current. *Proceedings of the National Academy of Sciences of the United States of America*. doi:10.1073/pnas.1309345110

Banse K. (1976). Rates of growth, respirations and photosynthesis of unicellular algae as relation to cell-size – review. J. Phycol. 12:135-140.

Behrenfeld MJ, Milligan AJ (2013) Photophysiological Expressions of Iron Stress in Phytoplankton. Annual Review of Marine Science 5:217-246.

Blaby-Haas CE, Merchant SS. 2012. The ins and outs of algal metal transport. Biochimica et Biophysica Acta Bioenergetics 1823, 1531–1552.

Boyd, P. W., Watson, a J., Law, C. S., Abraham, E. R., Trull, T., Murdoch, R., … Zeldis, J. (2000). A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature*, *407*(6805), 695–702. doi:10.1038/35037500

Boyd, P. W. (2002). Review: Environmental factors controlling phytoplankton processes in the Southern Ocean, *861*(October 2001), 844–861.

Brand R., Guillard R., and Murphy L. (1981). A method for the rapid and precise determination of acclimated phytoplankton reproductive rates. J. Plankton Res. 3: 193–201.

Cassar, N., Bender, M. L., Barnett, B. A., Fan, S., Moxim, W. J., Levy, H., & Tilbrook, B. (2007). The Southern Ocean biological response to aeolian iron deposition. *Science (New York, N.Y.), 317*(5841), 1067–70. doi:10.1126/science.1144602.

Chappell PD., Whitney LP., Joselynn WR. et al. (2015). The genetic indicators of iron limitation in wild populations of *Thalassiosira oceanica* from the northeast Pacific Ocean. The ISME Journal (2015) 9, 592–602; doi:10.1038/ismej.2014.171.

Death R. et al. Antarctic ice sheet fertilizes the Southern Ocean. Biogeosciences, 11, 2635–2644, (2014).

de Baar, H. J. W et al. (2005). Synthesis of iron fertilization experiments: From the Iron Age in the age of enlightenment. J. Geophys. Res. Oceans 110: C09S16, doi:10.1029/2004JC002601.

Deutsch, C., Hofmann, E., Ito, T., Lovenduski, N., Russell, J., Sarmiento, J. (2009) A report of the Southern Ocean Scoping Workshop sponsored by the Ocean Carbon and Biogeochemistry Program and the NSF Office of Polar Programs. Princeton, June, 2009.

Ducklow, H.W., D.K. Steinberg, and K.O. Buesseler. 2001. Upper ocean carbon export and the biological pump. *Oceanography* 14(4):50–58.

Ducklow, H. W., Baker, K., Martinson, D. G., Quetin, L. B., Ross, R. M., Smith, R. C., … Fraser, W. (2007). Marine pelagic ecosystems: the West Antarctic Peninsula. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1477), 67–94. doi:10.1098/rstb.2006.1955

Ducklow, H.W., W.R. Fraser, M.P. Meredith, S.E. Stammerjohn, S.C. Doney, D.G. Martinson, S.F. Sailley, O.M. Schofield, D.K. Steinberg, H.J. Venables, and C.D. Amsler. 2013. West Antarctic Peninsula: An ice-dependent coastal marine ecosystem in transition. Oceanography 26(3):190–203

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792-1797.

Edwards R. and Sedwick P. Iron in East Antarctic snow: Implications for atmospheric iron deposition and algal production in Antarctic waters. Geophys Res Lett 28(20): 3907–3910 (2001).

Ellis, K.A., Cohen N.R., Moreno C., Marchetti A. (Submitted). Cobalamin-independent methionine syntghase distribution and influence on vitamin B_{12} growth requirements in marine diatoms. Phycology.

Epply R.W., Rogers J.N., and McCarthy J.J. (1969). Half-saturation constants for uptake of nitrate and ammonium by marine phytoplankton. *Limno. Oceanogr*. 14: 912-920.

Falkowski, P. G. (1998). Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science, 281*(5374), 200–206. doi:10.1126/science.281.5374.200

Garibotti IA, Vernet MA, Smith RC, Ferrario ME (2005) Interannual variability in the distribution of the phytoplankton standing stock across the seasonal sea-ice zone west of the Antarctic Peninsula. J Plankton Res 27: 825–843.

Gorbunov, M.Y. & Falkowski, P.G. (2005). Fluorescence induction and relaxation (FIRe) technique and instrumentation for monitoring photosynthetic processes and primary production in aquatic ecosystems, in: A. van der Est, D. Bruce (Eds.), Photosynthesis: Fundamental Aspects to Global Perspectives, Proc. 13th International Congress of Photosynthesis, Montreal, Allen Press, Springer, pp 1029-1031

Greene, R. M., Geider, R. J. & Falkowski, P. G. Effect of iron limitation on photosynthesis in a marine diatom. Limnol. Oceanogr. 36, 1772–1782 (1991).

Gruber, N., Gloor, M., Mikaloff, Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., Takahashi, T. (2009). Oceanic sources, sinks, and transport of atmospheric CO 2. *Global Biogeochemical Cycles*, *23*(1),. doi:10.1029/2008GB003349

Hubbard, K., Rocap, G. & Armbrust, E. (2008). Inter- and intraspecific community structure within the diatom genus Pseudo-nitzschia (Bacillariophyceae). Journal of Phycology 44:637-649.

Jukes TH & Cantor CR (1969) Evolution of protein molecules. In Munro HN, editor, Mammalian Protein Metabolism, pp. 21-132, Academic Press, New York.

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., Worden, A. Z. 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 Biology*, *12*(6), e1001889. doi:10.1371/journal.pbio.1001889

Kroger N, Deutzmann R, Bergsdorf C, Sumper M (2000) Species-specific polyamines from diatoms control silica morphology. PNAS 97: 14133–14138.

Kustka AB, Allen AE, Morel FM. (2007). Sequence Analysis and Transcriptional Regulation of Iron Acquisition Genes in Two Marine Diatoms 1. J Phycol. 43:715–729.

La Roche J, Boyd PW, McKay RML, Geider RJ. (1996). Flavodoxin as an in situ marker for iron stress in phytoplankton. Nature, 382:802–805.

Lin Y, Cassar N, Marchetti A, Ducklow H and Li Z. (submitted). Specific eukaryotic plankton are good predictors of carbon export in the Southern Ocean. Nature Geoscience.

Litchman E. (2007) Resource competition and the Ecological success of phytoplankton. In *Evolution of Primary Producers in the Sea*. Falkowki PG., Knoll AH. (Eds.). Academic Press.

Maldonado and Price. Influence of N substrate on Fe requirements of marine centric diatoms. 1996. Mar Ecol Prog Ser 141: 161-172.

Marchetti, A. and Maldonado, M.T. (2016). Iron. In *The Physiology of Microalgae.* Borowitzka M, Beardall J and Raven J (Eds.). Springer Publishing Company.

Marchetti A, Catlett D, Hopkinson B, Ellis K and Cassar N (2015). Marine eukaryotic proteorhodopsins and their potential role in coping with low iron availability. The ISME Journal advance online publication; doi: 10.1038/ismej.2015.74

Marchetti, A., Schruth, D. M., Durkin, C. a, Parker, M. S., Kodner, R. B., Berthiaume, C. T., Armbrust, E. V. (2012). Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(6), E317–25. doi:10.1073/pnas.1118408109

Marchetti, A., Parker, M. S., Moccia, L. P., Lin, E. O., Arrieta, A. L., Ribalet, F., Armbrust, E. V. (2009). Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature*, *457*(7228), 467–70. doi:10.1038/nature07539.

Maxwell D.P., Want Y., and McIntoch L. (1999). The alternative oxidase lowers mitochondrial reactive oxygen production in plant cells. Proc Natl Acad Sci: 96(14): 8271–8276

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71: 491-499

Meredith, M. P. (2005). Rapid climate change in the ocean west of the Antarctic Peninsula during the second half of the 20th century. *Geophysical Research Letters*, *32*(19), L19604.

Montes-Hugo, M. et al. (2009). Recent changes in phytoplankton communities associated with rapid regional climate change along the Western Antarctic Peninsula. Science 323, 1470–1473.

Moore, J. K., & Abbott, M. R. (2000). Phytoplankton chlorophyll distributions and primary production in the Southern Ocean. *Journal of Geophysical Research*, *105*(C12), 28709.

Nuester J, Vogt S, Twining BS. Localization of Iron Within Centric Diatoms of the Genus Thalassiosira.J Phycol 2012, 48:626–634.

Pankowski A, McMinn A (2009) Iron availability regulates growth, photosynthesis, and production of ferredoxin and flavodoxin in Antarctic sea ice diatoms. Aquatic Biology 4:273-288.

Park S, Jung G, Hwang YS, Jin E. (2010) Dynamic response of the transcriptome of a psychrophilic diatom, Chaetoceros neogracile, to high irradiance. Planta 231:349–360.

Parra G., Bradnam K., Ning Z., Keane T., and Korf I. (2009). Assessing the gene space in draft genomes" Nucleic Acids Research, 37(1): 298-297.

Peers G, Price NM (2006) Copper-containing plastocyanin used for electron transport by an oceanic diatom. Nature 441:341–344.

Peters E and Thomas DN. (1996). Prolonged darkness and diatom mortality – a comparison of polar and temperate Thalssiosira species. Journal of Plankton Research 18, 953-968.

Petrou K., Trimborn S., Rost, Ralph PJ, and Hassler CS. (2014). The impact of iron limitation on the physiology of the Antarctic diatom Chaetoceros simplex. Mar Biol (2014) 161:925–937

Prihoda, J., Tanaka, A., de Paula, W.B.M., Allen, J.F., Tirichine, L., and Bowler, C. (2012). Chloroplast-mitochondria cross-talk in diatoms. Journal of Experimental Botany, 63(4): 1543-1557.

Raven JA. (1990). Predictions of Mn and Fe use efficiencies of phototrophic growth as a function of light availability for growth and of C assimilation pathway. New Phytol. 116:1–18.

Raven JA, Evans MCW, Korb RE. (1999). The role of trace metals in photosynthetic electron transport in O2-evolving organisms. *Photosynthesis Research* 60, 11–149.

Raven, JA. (2013). Iron acquisition and allocation in stramenopile algae. *Journal of Experimental Botany*, *64*(8), 2119–27.

Saba, G. K., Fraser, W. R., Saba, V. S., Iannuzzi, R. a, Coleman, K. E., Doney, S. C., Ducklow, H. W., Martinson D. G., Miles T. N., Paterson-Fraser, D. L., Stammerjohn, S. E., Steinberg D. K., Schofield, O. M. (2014). Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula. *Nature Communications*, *5*, 4318. doi:10.1038/ncomms5318

Sabine, C. L., Feely, R. A, Gruber, N., Key, R. M., Lee, K., Bullister, J. L.,ios, A. F. (2004). The oceanic sink for anthropogenic CO2. *Science (New York, N.Y.)*, *305*(5682), 367–71.

Sallée J. B. et al Zonally asymmetric response of the Southern Ocean mixed-layer depth to the Southern Annular Mode. Nature Geoscience 3, 273 - 279 (2010).

Sañudo-Wilhelmy, S. a, Gómez-Consarnau, L., Suffridge, C. & Webb, E. a 2014. The role of B vitamins in marine biogeochemistry. *Ann. Rev. Mar. Sci.* 6:339–67.

Sarmiento, J. L., Gruber, N., Brzezinski, M. A, & Dunne, J. P. (2004). High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature*, *427*(6969), 56–60.

Schofield, O., Ducklow, H. W., Martinson, D. G., Meredith, M. P., Moline, M. a, & Fraser, W. R. (2010). How do polar marine ecosystems respond to rapid climate change? *Science (New York, N.Y.)*, *328*(5985), 1520–3. doi:10.1126/science.1185779.

Scion Image Software. 2015. Informer Technologies, Inc. http://scion-image.software.informer.com/.

Smith, W. O., & Comiso, J. C. (2008). Influence of sea ice on primary production in the Southern Ocean: A satellite perspective. *Journal of Geophysical Research*, *113*(C5), C05S93.

Sokolov, S., & Rintoul, S. R. (2007). On the relationship between fronts of the Antarctic Circumpolar Current and surface chlorophyll concentrations in the Southern Ocean. *Journal of Geophysical Research*, *112*(C7), C07030. doi:10.1029/2006JC004072

Sunda WG, Swift DG, Huntsman SA. 1991. Low iron requirement for growth in oceanic phytoplankton. Nature 351:55–57.

Sunda, W.G. and Huntsman, S.A. (1995). Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Mar Chem* 50: 189-206.

Sunda, W. G., & Huntsman, S. A. (1997). Interrelated influence of iron, light and cell size on marine phytoplankton growth, Nature. *390*(6658), 389–392. doi:10.1038/37093

Sweeney, C. (2003), The annual cycle of surface water CO2 and O2 in the Ross Sea: A model for gas exchange on the continental shelves of Antarctica, in Biogeochemistry of the Ross Sea, Ant. Res. Ser., vol. 78, pp. 295–312, AGU, Washington, D. C.

Strauss, Jan (2012) A genomic analysis using RNA-Seq to investigate the adaptation of the psychrophilic diatom Fragilariopsis cylindrus to the polar environment. Doctoral thesis, University of East Anglia.

Strzepek RF, and Harrison PJ (2004). Photosynthetic architecture differs in coastal and oceanic diatoms. Nature 431: 689–692, doi:10.1038/nature02954.

Strzepek RF., Maldonado MT., Hunter KA., Frew RD., Boyd PW. (2011). Adaptive strategies by Southern Ocean phytoplankton to lessen iron limitation: Uptake of organically complexed iron and reduced cellular iron requirements. Limnol. Oceanogr. 56: 1983–2002,

Strzepek R.F., Hunter K.A., Frew R.D., Harrison P.J., & Boyd P.W. (2012). Iron-light interactions differ in Southern Ocean phytoplankton. Limnology and Oceanography, 57(4), 1182–1200.

Takeda, S. (1998). Influence of iron availability on nutrient consumption ratio of diatoms in oceanic waters, *393*(June), 1–4.

Talley L.D., Pickard G.L., Emery W.J., Swift J.H., 2011. *Descriptive Physical Oceanography: An Introduction* (Sixth Edition), Elsevier, Boston, 560 pp.

Timmermans KR. Gerringa LJA., de Baar HJW, van der Wagt B., Veldhuis MJW., do Jon JTM, and Croot PL. 2001. Growth rates of large and small Southern Ocean diatoms in relation to availability of iron in natural seawater. Limnol. Oceanogr. 46: 260–266.

.Timmermans, K. R., Wagt, B. Van Der, & de Baar, H. J. W. (2004). Growth rates, half saturation constants, and silicate, nitrate, and phosphate depletion in relation to iron availability of four large open-ocean diatoms from the Southern Ocean. *Limnology and Oceanography*, *49*(6), 2141–2151.

Tremblay, J.-E., and W. O. Smith Jr. (2007). Phytoplankton processes in polynyas, in *Polynyas: Windows to the World's Oceans*, edited by W. O. Smith Jr. and D. G. Barber, pp. 239– 270, Elsevier, Amsterdam.

Trimborn S., Thoms S., Petrou K, Kranz SA., Rost B. (2014). Photophysiological responses of Southern Ocean phytoplankton to changes in CO2 concentrations: Short-term versus acclimation effects. Journal of Experimental Marine Biology and Ecology 451 44–54.

Van Ho A, Ward DM, Kaplan J (2002) Transition metal transport in yeast. Annual Reviews in Microbiology 56:237-261.

van Oijen T., van Leeuwe MA., Gieskes WWC and de Baar HJW. (2004) Effects of iron limitation on photosynthesis and carbohydrate metabolism in the Antarctic diatom Chaetoceros brevis (Bacillariophyceae), European Journal of Phycology, 39:2, 161-171.

Whitney LP, Lins JJ, Hughes MP, Wells ML, Chappell PD, Jenkins BD. 2011. Characterization of putative iron responsive genes as species-specific indicators of iron stress in thalassiosiroid diatoms. Front Microbiol 2:234. doi: 10.3389/fmicb.2011.00234.

Wood PM. Interchangeable copper and iron proteins in algal photosynthesis. Studies on plastocyanin and cytochrome c-552 in *Chlamydomonas*. Eur J Biochem 1978 Jun 1;87(1):9–19. pmid:208838