UNIVERSITY OF CALIFORNIA, SAN DIEGO

Iron, ice and advection: the influence of the Antarctic marine environment on the physiology and distribution of phytoplankton

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Marine Biology

by

Jeffrey B. McQuaid

Committee in Charge:

Andrew E. Allen, Chair
Eric E. Allen
Katherine A. Barbeau
Mark H. Ellisman
Mark Hildebrand

2017

The dissertation of Jeffrey B. McQuaid is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego
2017

DEDICATION

To Katherine and Alexander

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Chapter one has been prepared as a Letter to *Nature*. "Carbonate sensitive phytotransferrin controls high affinity iron uptake in diatoms". McQuaid, Jeffrey B.; Kustka, Adam B.; Oborník, Miroslav; Horák, Aleš; McCrow, John P.; Karas Bogumil J.; Zheng, Hong; Kindeberg, Theodor; Andersson, Andreas J.; Barbeau, Katherine A.; Allen, Andrew E. We thank J. Badger for early contributions to phylogenetic analyses, A. Dickson for pH analysis, K. Forsch for CSV measurements, and E. Bertrand for Trace-metal clean techniques. This study was supported by the National Science Foundation (NSF-MCB-1024913, NSF-ANT-1043671, NSF-OCE--0727997, United States Department of Energy Genomics Science program (DE-SC00006719 and DE-SC0008593) and Gordon and Betty Moore Foundation grant GBMF3828 (AEA), and the Czech Science Foundation, project 15-17643S (MO and AH).

Chapter two has been prepared as a short communication to the *International Society for Microbial Ecology (ISME) Journal*. McQuaid, Jeffrey B.; McCrow, John P.; Hoffman, Jeffrey M.; Zheng, Hong; Espinoza, Joshua; Allen, Andrew E. "Metagenomic survey of East Antarctica reveals microbial advection into the deep ocean".

Chapter three has been prepared as a submission to the *International Society for Microbial Ecology (ISME) Journal*. McQuaid, Jeffrey B.; McCrow, John P.; Hoffman, Jeffrey M.; Zheng, Hong; Cziko,

Paul A.; Valas, Ruben E.; Allen, Andrew E. "Sea Ice Transcriptome reveals unique stresses of conducting photosynthesis in a semisolid matrix."

VITA

1992: Bachelor of Science, St. Lawrence University

2000: Master of Science, The University of Arizona

2017: Doctor of Philosophy, Scripps Institution of Oceanography, University of California, San Diego

PUBLICATIONS

Diner, Rachel E., Chari M. Noddings, Nathan C. Lian, Anthony K. Kang, **Jeffrey B. McQuaid**, Jelena Jablanovic, Josh L. Espinoza, Ngocquynh A. Nguyen, Miguel A. Anzelmatti, Jakob Jansson, Vincent A. Bielinski, Bogumil J. Karas, Christopher L. Dupont, Andrew E. Allen, Philip D. Weyman "Diatom Centromeres Suggest a Novel Mechanism for Nuclear Gene Acquisition." *bioRxiv* (2016): 096016.

Bertrand, Erin M., John P. McCrow, Ahmed Moustafa, Hong Zheng, **Jeffrey B. McQuaid**, Tom O. Delmont, Anton F. Post, Rachel E. Sipler, Jenna L. Spackeen, Kai Xu, Deborah A. Bronk, David A. Hutchins, and Andrew E. Allen. "Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge." *Proceedings of the National Academy of Sciences* 112, no. 32 (2015): 9938-9943.

Karas, Bogumil J., Rachel E. Diner, Stephane C. Lefebvre, **Jeffrey B McQuaid**, Alex P.R. Phillips, Chari M. Noddings, John K. Brunson, Ruben E. Valas, Thomas J. Deerinck, Jelena Jablanovic, Jeroen T.F. Gillard, Karen Beeri, Mark H. Ellisman, John I. Glass, Clyde A. Hutchison III, Hamilton O. Smith, J. Craig Venter, Andrew E. Allen, Christopher L. Dupont & Philip D. Weyman. "Designer diatom episomes delivered by bacterial conjugation." *Nature communications* 6 (2015).

Morrissey, Joe, Robert Sutak, Javier Paz-Yepes, Atsuko Tanaka, Ahmed Moustafa, Alaguraj Veluchamy, Yann Thomas Hugo Botebol, François-Yves Bouget, **Jeffrey B. McQuaid**, Leila Tirichine, Andrew E. Allen, Emmanuel Lesuisse, Chris Bowler, "A novel protein, ubiquitous in marine phytoplankton, concentrates iron at the cell surface and facilitates uptake." *Current Biology* 25, no. 3 (2015): 364-371.

Wilkins, D., Frederico M. Lauro, T. J. Williams, Matthew Z DeMaere, Mark V. Brown, Jeffrey M. Hoffman, Cynthia Andrews-Pfannkoch, **Jeffrey B. McQuaid,** Martin J. Riddle, Steven R. Rintoul, and Ricardo Cavicchioli. "Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics." *Environmental microbiology* 15, no. 5 (2013): 1318-1333.

Yooseph, Shibu, Cynthia Andrews-Pfannkoch, Aaron Tenney, **Jeffrey B. McQuaid**, Shannon Williamson, Mathangi Thiagarajan, Daniel Brami, Lisa Zeigler-Allen, Jeff Hoffman, Johannes B. Goll, Douglas Fadrosh, John Glass, Mark D. Adams, Robert Friedman, J. Craig Venter. A Metagenomic Framework for the Study of Airborne Microbial Communities. *PLOS ONE* (2013) 8(12).

Brown, Mark V., Federico M Lauro, Matthew Z DeMaere, Les Muir, David Wilkins, Torsten Thomas, Martin J Riddle, Jed A Fuhrman, Cynthia Andrews-Pfannkoch, Jeffrey M Hoffman, **Jeffrey B McQuaid**, Andrew Allen, Stephen R Rintoul, Ricardo Cavicchioli. "Global biogeography of SAR11 marine bacteria." *Molecular systems biology* 8, no. 1 (2012): 595.

ABSTRACT OF THE DISSERTATION

Iron, ice and advection: how the physiology and distribution of marine organisms is influenced by the extremes of the Antarctic environment

by

Jeffrey B. McQuaid

Doctor of Philosophy in Marine Biology

University of California, San Diego 2017

Andrew E. Allen, Chair

The Southern Ocean is a key driver of biogeochemical processes and the earth's climate, influencing the global distribution of water masses, nutrients, and atmospheric CO_2 drawdown rates. The physical and geochemical environment of the Southern Ocean influences and is in turn influenced by the plankton of the euphotic zone. One of the key controlling factors which limits the growth and productivity of phytoplankton is the availability of the micronutrient iron. In the first

chapter, I characterize a novel high affinity iron binding phytotransferrin in the diatom *Phaeodactyulum tricornutum,* which reveals that carbonate-coordinated ferric iron binding evolved independently and convergently from metazoan transferrin. Biochemical manipulations of the seawater carbonate system demonstrate that the concentration of the carbonate ion co-limits uptake rates of unchelated ferric iron. As phytotransferrin sequences have broad taxonomic distribution and are abundant in marine environmental genomic datasets, this suggests that ocean acidification will negatively impact a globally important eukaryotic iron acquisition mechanism. In the second chapter, I analyze a metagenomic survey of the distribution and genomic capacity of planktonic organisms in the continental shelf regions of East Antarctica and demonstrate an advective connection between surface and deep populations of ammonia oxidizing Archaea. In the third chapter, I use environmental transcriptomics (metatranscriptomics) to evaluate cellular- and ecosystem-level responses of an ice algae community living and fixing carbon on the underside of Antarctic sea ice. Taken together, these three chapters reveal some of the remarkable adaptations which marine plankton have used to adapt to life in one of the Earth's most extreme environments, the Southern Ocean.

INTRODUCTION

The physical, chemical and biological processes which occur in the Southern Ocean have a critical and outsized effect on global biogeochemistry, with direct ramifications to the Earth's overall climate. The Southern Ocean can be thought of as a window of exchange between the atmosphere and the deep ocean: unimpeded circulation of the Antarctic Circumpolar Current drives the upwelling of nutrient-rich deep waters to the surface of the ocean where they fuel productive polar ecosystems. At the same time, the nearshore regions around Antarctica are one of the few places on Earth where equilibrated surface waters sink to depth, driving overturning circulation which is a critical component of the earth's climate. This large-scale upwelling and sinking explains why the Southern Ocean (defined as the ocean south of the subtropical convergence) occupies 20% of the ocean surface area yet absorbs 50% of the absorbed heat (Barnett et al., 2005) and 40% of carbon dioxide (Sabine et al., 2004). At the surface of the ocean and occupying the interface of between the atmosphere and the deep ocean, are phytoplankton. Phytoplankton moderate CO_2 uptake via the fixation of carbon dioxide as well as affect the chemical and nutrient signature of surface waters, and as these water masses are subducted and redistributed throughout the global ocean basins, they bear the chemical signatures of Antarctic phytoplankton from their time at the surface.

Phytoplankton in the Southern Ocean are exposed to and survive in one of the more extreme environments on the planet. In addition to cold

temperatures, two environmental extremes directly impact the growth and productivity of phytoplankton: the lack of light through much of the polar winter, and the low concentrations of dissolved iron throughout the summer growing season. Concerning the latter, phytoplankton growing in the Southern Ocean contend with some of the lowest dissolved iron concentrations on the planet (Kluder et al., 2011). In the Southern Ocean, dissolved iron is lower than any other ocean basin and far lower than most terrestrial environments making the Southern Ocean an extreme environment for organisms to obtain the micronutrient iron (Boyd et al., 2007). The concentration of dissolved iron in the Southern Ocean surface waters is frequently sub-nanomolar, resulting in concentration of unchelated 'free' ferric iron in the picomolar range (Gledhill and Buck, 2012; Barbeau, 2001). This concentration is far below the affinity constants of described microbial iron transporters (Blaby-Haas and Merchant, 2012), and creating a degree of uncertainty as to how planktonic organisms obtain iron at such low concentrations of iron.

Another environmental extreme which planktonic organisms encounter in the Southern Ocean is coverage of the water column by sea ice. At the peak, sea ice covers twenty million square kilometers of the Southern Ocean, nearly doubling the size of the Antarctic continent (Arrigo et al., 1997). Algae, particularly diatoms, colonize this environment, growing in the cracks and brine channels and frequently obtaining high densities. The brine channels constitute one of the more extreme

environments where eukaryotic plants grow: carbon fixation and the subsequent drawdown of inorganic carbon creates a highly alkaline environment where the pH can rise above 11 (Gleitz, 1995). Brine rejection during ice formation creates salinities in excess of 160 ppt (Mock, 2003) oxygen can builds up to 200% of saturation (Kuhl, 2001) and the sheer density of organisms can lead to diffusion limitation of nutrients, metals and carbon (Arrigo et al., 2004). Though low iron environments and sea ice affect millions of square kilometers of Southern Ocean habitats, the adaptations which organisms use for these environments are not fully understood.

In chapter one, a phylogenomic analysis is combined with biochemical characterizations and reverse genetics to identify the high affinity uptake mechanism of the model diatom *Phaeodactylum tricornutum.* Transcripts encoding this protein are among the most abundant in the Ross Sea (Bertrand, 2015). We show that *P. tricornutum* uses a phytotransferrin, a functional analog of transferrin which convergently and independently evolved carbonate-coordinated ferric iron binding. Knockouts strains deficient in the phytotransferrin protein are unable to access picomolar concentrations of unchelated ferric iron, and we found that function could be restored in the knockouts by complementing with human transferrin, strongly suggesting a shared mechanism between the two proteins. Like transferrin, phytotransferrin is taken into the cell via endocytosis, and manipulations of the seawater

carbonate concentration showed that uptake was directly dependent on the concentration of carbonate ion (CO_3^{2-}).

Further analysis demonstrated that uptake of iron has a second order dependence on the concentrations of iron and carbonate ion, and that the interaction is synergistic rather than additive. We found that the interaction occurs across a range of environmentally relevant concentrations of seawater carbonate ion, revealing that the concentration of carbonate ion limits iron uptake, an interaction previously unknown. As phytotransferrin sequences have broad taxonomic distribution (Morrissey et al. 2015) and are abundant in marine environmental genomic datasets (Marchetti et al. 2012), this suggests that ocean acidification and the decline in the concentration of seawater carbonate ion will negatively impact a globally important eukaryotic iron acquisition mechanism. In the Southern Ocean, carbonate ion concentrations are already among the lowest in the world, and by the year 2100, carbonate ion concentrations will likely decrease by 40% (McNeil, 2008) potentially exacerbating iron limitation and negatively impacting productivity of the Southern Ocean.

In chapter two, we examine data from a metagenomic survey undertaken of the Mertz polynya, a region of open water where Antarctic bottom water formation takes place. In latent heat polynyas like the Mertz, ice production is high, generating ultra-cold and ultra-dense high-salinity shelf waters which sink and fuel meridional overturning circulation.

The ice-free nature of polynyas also creates higher phytoplankton productivity due to direct insolation of the water surface (Arrigo, 2003). An examination of microeukaryotic community DNA in continental shelf and slope regions confirmed the influence of silica in structuring planktonic communities. An analysis of ultra-cold and ultra-dense ice shelf water emerging from the Mertz Glacier Tongue revealed archaeal populations living under the shelf, confirming observations from the Ross and Wedell ice shelf regions (Vick-Majors, 2016). We found that ammonia oxidizing archaea are advected down the Antarctic continental slope to form Antarctic Bottom Water. We hypothesize that ice shelves in regions of active bottom water formation serve as reservoirs of ammonia oxidizing archaea, seeding the deep ocean with continuous populations of surface-derived archaeal communities.

In chapter three, we use genomics to characterize the biological stresses and community dynamics of sea ice algal communities. The sea ice surrounding Antarctica is one of the largest seasonal habitats on earth, reaching a maximum of 20 million square kilometers in the early spring. Algae, particularly diatoms, adhere to and grow on the underside of the ice, and represent an early and critical component of the Antarctic food web. Diatoms grow on the undersides of sea ice for light-field stability and to escape predation, but our understanding of how photosynthesis and carbon fixation are conducted while embedded in a semisolid ice matrix is limited by subsampling of a heterogenous environment. We employ

environmental transcriptomics (meta-transcriptomics) to evaluate cellular- and community-level responses to the sea ice environment. We find that transcripts for a biophysical carbon concentrating mechanism (CCM) are strongly upregulated, as are transcripts encoding freeze tolerance, nutrient and metal acquisition, and protection from reactive oxygen species (ROS). In aggregate, the transcriptional data describe an environment where gasses and nutrients are diffusion-limited, creating unique stresses, and confirming and building upon physical and chemical measurements of the sea ice environment.

Taken together, these three chapters paint a portrait of the extremes of Antarctica and the Southern Ocean, as well as the amazing biological adaptations which organisms use to thrive in that environment. Polar regions including the Southern Ocean are among most sensitive regions to climate change, and are also among the fastest changing environments on the planet (McNeil, 2008). While the effects of warming and acidification on the phytoplankton community are both unknown and unpredictable, there will likely be winners and losers, adding another element of uncertainty to our understanding of the future effects of climate change.

REFERENCES FOR THE INTRODUCTION

Arrigo, Kevin R., Denise L. Worthen, Michael P. Lizotte, Paul Dixon, Gerhard Dieckmann. Primary Production in Antarctic Sea Ice. *Science* **276** (1997): 394–97.

Arrigo, Kevin R, and David N Thomas. Large Scale Importance of Sea Ice Biology in the Southern Ocean. *Antarctic Science* **16** (2004): 471–86

Arrigo, K. R. Phytoplankton dynamics within 37 Antarctic coastal polynya systems. J. Geophys. Res. 108, 3271 (2003).

Barbeau, K., Rue, E. L., Bruland, K. W., & Butler, A. Photochemical cycling of iron in the surface ocean mediated by microbial iron (III)-binding ligands. *Nature*, **413**, 409-413 (2001).

Barnett, T.P., Pierce, D.W., AchutaRao, K.M., Gleckler, P.J., Santer, B.D., Gregory, J.M. and Washington, W.M., 2005. Penetration of human-induced warming into the world's oceans. *Science*, **309** (2005): 284-287

Bertrand, Erin M, John P McCrow, Ahmed Moustafa, Hong Zheng, Jeffrey B McQuaid, Tom O Delmont, Anton F Post, et al. Phytoplankton-Bacterial Interactions Mediate Micronutrient Colimitation at the Coastal Antarctic Sea Ice Edge. *Proceedings of the National Academy of Sciences of the United States of America* **112** (2015): 9938–43.

Blaby-Haas, C. E., & Merchant, S.S. The ins and outs of algal metal transport. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1823, no. 9 (2012): 1531-1552

Boyd, P.W., Jickells, T., Law, C.S., Blain, S., Boyle, E.A., Buesseler, K.O., Coale, K.H., Cullen, J.J., De Baar, H.J., Follows, M. and Harvey, M. Mesoscale iron enrichment experiments 1993-2005: Synthesis and future directions. *Science* **315** (2007): 612–617

Gledhill, M., & Buck, K. N. The organic complexation of iron in the marine environment: a review. *Front. Microbiol*. **3**, (2012): 1–17

Gleitz, Markus, Michiel Rutgers v.d. Loeff, David N Thomas, Gerhard S Dieckmann, and Frank J Millero. Comparison of Summer and Winter Inorganic Carbon, Oxygen and Nutrient Concentrations in Antarctic Sea Ice Brine. *Marine Chemistry* 51 (1995): 81–91.

Klunder, M. B., P. Laan, R. Middag, H. J. W. De Baar, and J. C. Van Ooijen. Dissolved iron in the Southern Ocean (Atlantic sector). *Deep Sea Research Part II: Topical Studies in Oceanography* **58**, no. 25 (2011): 2678-2694.

Kühl, Michael, Ronnie N. Glud, Jens Borum, Rodney Roberts, and Søren Rysgaard. Photosynthetic Performance of Surface-Associated Algae below Sea Ice as Measured with a Pulse-Amplitude-Modulated (PAM) Fluorometer and O2 Microsensors. *Marine Ecology Progress Series* **223** (2001): 1–14.

Marchetti, A. et al. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 317-325 (2012).

McNeil, B.I. and Matear, R.J. Southern Ocean acidification: A tipping point at 450-ppm atmospheric CO2. *Proceedings of the National Academy of Sciences*, **105** (2008.) pp.18860-18864.

Morrissey, J. et al. A novel protein, ubiquitous in marine phytoplankton, concentrates iron at the cell surface and facilitates uptake. *Curr. Biol.* **25**, 364–371 (2015).

Mock, Thomas, and David N. Thomas. Recent Advances in Sea-Ice Microbiology. *Environmental Microbiology* **7** (2005): 605–19

Sabine, C.L., Feely, R.A., Gruber, N., Key, R.M., Lee, K., Bullister, J.L., Wanninkhof, R., Wong, C., Wallace, D.W., Tilbrook, B. and Millero, F.J. The oceanic sink for anthropogenic CO2. *Science*, **305** (2004): 367-371.

CHAPTER 1: CARBONATE SENSITIVE PHYTROTRANSFERRIN CONTROLS HIGH AFFINITY IRON UPTAKE IN DIATOM

1.1 ABSTRACT

In vast areas of the ocean the scarcity of iron controls the growth and productivity of phytoplankton (Moore et al. 2013; Boyd et al. 2007). While most dissolved iron in the marine environment is complexed to organic molecules (Gledhill and Buck, 2012), picomolar amounts of unchelated ferric iron (Fe') are maintained within the euphotic zone and are an important source of iron for eukaryotic phytoplankton, particularly diatoms (Morel et al., 2008). Genome-enabled studies of Fe' utilization in diatoms have revealed novel iron-responsive transcripts (Allen et al. 2008; Lommer et al. 2012), including ferric iron concentrating ISIP2a (Morrissey et al. 2015), but the mechanism behind picomolar Fe' acquisition remains unknown. Here we show that ISIP2a is a phytotransferrin which independently and convergently evolved carbonate-coordinated ferric iron binding. Deletion of ISIP2a disrupts high affinity iron uptake in the diatom *Phaeodactylum tricornutum*, and uptake is restored by complementation with human transferrin. ISIP2a is internalized via endocytosis, and manipulations of the seawater carbonic acid system reveal a second order dependence on the concentrations of Fe' and carbonate (CO_3^{2-}). In *P. tricornutum*, the synergistic interaction of [Fe'] and [CO_3^{2-}] occurs across a range of environmentally relevant concentrations of carbonate ion, revealing that [CO_3^{2-}] co-limits Fe' uptake. Phytotransferrin sequences

have broad taxonomic distribution (Morrissey et al. 2015) and are abundant in marine environmental genomic datasets (Marchetti et al. 2012; Bertrand et al. 2015), suggesting that ocean acidification and the decline in seawater [CO_3^{2-}] will negatively impact a globally important eukaryotic iron acquisition mechanism.

1.2 METHODS

1.2.1 Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

1.2.2 Phylogenetic analysis

To maximize the dimensions of our phylogenetic analysis, we built a series of Hidden Markov Models (HMM) from a manually curated multiple peptide alignments. HMMs were assembled using transferrin and ISIP2a sequences obtained from NCBI via BLAST. Transferrin and ISIP2a HMMs were used to iteratively obtain sequences from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP; Caron et al. 2014). For inclusion into the transferrin and phytotransferrin clades, sequences were manually screened to contain the conserved metal binding residues and the carbonate binding region (Extended data Fig. 3a). Sequences were aligned using genafpair exhaustive algorithm included in MAFFT (Katoh and Standley, 2013). The alignment was then manually

inspected and ambiguous regions were removed in SEAVIEW 4 . Phylogenetic analysis was conducted using maximum likelihood under the gamma-corrected LG matrix in RAxML 8.2.8[33]. The best topology, as well as non-parametric branching support was inferred using fast algorithm (-f a option) from 1000 replicates. Due to the high divergence of sequences, we have alternatively used Bayesian inference under the probabilistic CAT algorithm with numbers of site-categories limited to 40 (C40 model), combined with empirical exchange rates as defined by the LG model in Phylobayes 4.1 (Lartillot et al., 2009) Two independent MCMC chains were run until they converged (i.e. maximum observed discrepancy was below 0.2) and minimum effective size of model parameters exceeded 100.

1.2.3 Divergence time estimation

Fossil calibration points for multicellular taxa (metazoans and green plants) were compiled from http://fossilcalibrations.org (Extended Data Table 4). Calibration of diatoms and haptophytes (mineralizing protozoans) was adopted according to Berney (Berney and Pawlowski, 2006) and Parfrey (Parfey et al., 2011). Respective values are summarized in Table 4. The best-scoring topology from the Bayesian inference (see above) was then used for an estimation of divergence using UGAM and CIR relaxed clock models in Phylobayes 4.1. Alternatively, we have employed an uncorrelated log-normal clock model as implemented in BEAST 2 (Drummond, 2012). Here the MCMC were run for 30 million generations. Convergence as well as effective sample size was evaluated

using Tracer and the dating visualized on the tree using TreeAnnotator in the BEAST package.

1.2.4 Predictions of signal peptides, transmembrane domains and GPI anchors

All the sequences used for phylogenetics reconstructions were searched for the presence of ER signal peptide (SP), transmembrane domains (TD) and glycosylphosphatidylinositol anchor (GPI). ER signal peptides were predicted using SignalP (Nielsen et al., 1997) and TargetP (Emanuelsson et al., 2000). Transmembrane domains were searched using TMHMM (Moller et al., 2001) and GPI anchors were predicted by PredGPI (Pierleoni et al., 2008). The results are shown in a tree in supplementary figure 1.

1.2.5 Cell lines and cultivation

All cultivation and manipulation of *Phaeodactylum tricornutum* strain CCMP 632 was conducted using sterile trace-metal clean techniques in Aquil media (Sunda et al., 2005) unless otherwise noted. To remove contaminating iron, prepared media was first passed through a Chelex 100 resin (Bio-Rad Laboratories) and microwave-sterilized in acid cleaned polycarbonate bottles prior to amendment with filter steriziled, Chelexed nutrient stocks and f/2 vitamin amendments. For cultures grown for use in short term uptake rates: Cultures were grown in media amended with macronutrients 880 µM NO_3, 36 µM PO_4 106 µM SiO_4 and trace metals

(100 nM Zn, 48 nM Mn, 40 nM Co, 40 nM Cu, 10 nM Se and 100 nM Ni) with 100 μM EDTA. Vitamin and metal-EDTA solutions were added to artificial seawater and equilibrated overnight. Iron was pre-equilibrated overnight in a 1:1.25 ratio of iron to EDTA prior to addition. Cultures were maintained at 17°C at 300 μmol quanta m^{-2} s^{-1} in a 10 h:14 h dark: light cycle, and all culture manipulations were done under a class-100 HEPA filter. Background contaminating iron was measured by flow injection analysis at multiple times and varied from 0.2 – 0.7 nM total dissolved iron over the course of the experimentation. Fe', the sum of all unchelated iron species, was calculated as in Sunda (Sunda et al., 2005). Cultures grown for growth rate analysis: Cultures of *P. tricornutum* (CCMP 632) were grown in a modified Aquil medium (Sunda et al., 2005) containing EDTA, Co, Cu, Mo, Mn, Zn at final concentrations of 3×10^{-4} M, 1.51×10^{-7} M, 1×10^{-7} M, 3.36×10^{-7} M, 1×10^{-8} M, and 2.39×10^{-7} M, with variable Fe concentrations. Cells were grown at 320 μmol photons m^{-2} s^{-1} under a 12/12 light/dark cycle using cool-white fluorescent bulbs. Under these conditions, Fe'= 0.00068 of total Fe. A maximum of 750 pmol L^{-1} Fe' was assumed because of iron hydroxide precipitation, but for simplicity, we report the results as Fe' even at concentrations exceeding this maximum. Cell growth was monitored daily using a Turner AU-10 fluorometer for *in vivo* chlorophyll fluorescence, and steady state growth rates were computed from linear regressions of ln (relative fluorescence) versus time.

1.2.6 TALEN knockouts, transferrin complements, and site directed mutagenesis

Transcription activator-like effector nucleases (TALENs) were used to disrupt the ISIP2a gene. TALEN construction protocols followed the methods as described for *P. tricornutum* (Weyman et al., 2015). Briefly, a unique pair of 20 nucleotide length sequences were designed to span the first intron junction of ISIP2a (GenBank ID: XP_002179762.1). TALEN constructs were assembled into the pTH vector (Weyman et al., 2015) and sequence verified. A homologous recombination (HR) plasmid containing resistance to Nourseothricin and driven by the promoter/terminator for the light harvesting promoter FcpA (Zaslavskaia et al., 2000) was assembled into pUC-19 using Gibson cloning (Gibson, et al., 2009). Both HR and pTH plasmids were introduced into *P. tricornutum* via particle bombardment (Falicatore et al., 1999). Transformants were pre-screened by PCR for correct insertion of the Nourseothricin resistance cassette and ΔISIP2a gene knockouts were subsequently confirmed by western blot using custom-made Genomic Antibodies (OriGene) targeting the ISIP2a protein.

Complement sequences for the N- and C-domains of human serum transferrin (HST) were codon optimized for the diatom *Thalassiosira pseudonana* and fused *in silico* to the signal peptide and C-terminal transmembrane anchor of ISIP2a. These fusion products were then synthesized on a BioXp 3200 DNA synthesizer (SGI DNA). The YFP complement was fused via Gibson cloning to the signal and

transmembrane sequences of ISIP2a, while the ISIP2a complement was obtained by amplifying *P. tricornutum* cDNA (GeneScript). PCR based site directed mutagenesis strains were made with base pair mismatch primers and amplified from the ISIP2a complement strain. All constructs were placed under control of the *P. tricornutum* Nitrate Reductase (NR) promoter and terminator (Poulsen et al., 2005), sequence verified, and maintained on pUC-19 vectors (n = 3 for each construct). All complement sequences were Gibson assembled into the cargo plasmid p0521s (containing a selection marker for bleomycin; Karas et al., 2015) and introduced into the ΔISIP2a knockout strain via conjugation with *Escherichia coli* Epi300 cells as described (Karas et al., 2015). Positive transformants were selected on dual Nourseothricin/Zeocin antibiotic plates (200 and 100 μg ml^{-1}, respectively) and verified by PCR and sequencing. Cell lines were maintained in 7.5 nM FeEDTA Aquil under positive antibiotic selection until testing.

1.2.7 Short term ^{59}Fe uptake assays

Cell strains were grown without antibiotic selection and acclimated in Aquil at 15 pM Fe' for a minimum of 5 transfers prior to assay. 15 pM Fe' was found to reduce growth rate in WT *P. tricornutum* by ~15%, with an approximate Fv/Fm of 0.55, indicative of iron-limited (but not iron-starved) growth (Allen et al., 2008) Cell counts were conducted on a Beckman Quanta cell cytometer, cross-calibrated by direct microscopic counts using a hemocytometer. To avoid changes to physiology and

media pH, cultures were grown and assayed at low cell densities (between 1 – 3 x 10^5 cells ml^{-1}). ^{59}FeEDTA was made by equilibrating ^{59}Fe with EDTA at pH 7 for 24 hours prior to addition to growth media, and desferrioxamine B was loaded with ^{59}Fe by equilibrating at pH 3 overnight before adjusting to pH 8 with high purity NaOH. The ratio of ^{59}Fe to unloaded DFB was 4:5. Uptake media was made by pre-equilibrating 7.5 nM ^{59}FeEDTA media and gravimetrically adding a 1:10 volume to the cell cultures. During uptake assays, cells were shielded from actinic light and periodically filtered onto 2.0 µm PTFE filters, washed with Ti-citrate-EDTA solution (Hudson and Morel, 1989) and preserved in Ecolite (MP Biomedicals) prior to liquid scintillation counting. Short term uptake assays did not exceed 3 hrs.

1.2.8 DIC manipulation experiments

Media for the DIC manipulation experiments was made by acidifying Chelex-treated bicarbonate-free Aquil media to pH 3 with ultra-high purity HCl and bubbling with filtered N_2 for 60 minutes to strip out dissolved CO_2. To buffer against large (±0.1 pH) changes in pH, Chelex-treated Tris-HCl was added at a final concentration of 2 mM, and the pH was adjusted back to 8.0 using ultra-high purity ammonium hydroxide. Ammonium hydroxide was chosen to reduce contaminating concentrations of carbonate ion. Trace-metal clean F/2 nutrients were added and all media was stored in an N_2/O_2 purged, CO_2 impermeable glove box prior to use. DIC was added to sub-aliquoted bottles of media at concentrations of 0,

0.5, 1, 1.5, 2, 4 and 6 mM DIC L^{-1}, and media were pre-equilibrated with ^{59}FeEDTA or ^{59}FeFOB 24 hours prior to assay. *P. tricornutum* cells were grown in 2.7 L square polycarbonate bottles in Aquil media containing 15 pM Fe' (7.5 nM Fe EDTA), and were grown on a 10/14 light/dark schedule. Cultures were grown to early exponential phase (3 x 10^5 cells ml^{-1}) and harvested by centrifugation at 1000 x g in acid-cleaned centrifuge bottles. Cells were washed and centrifuged 2x in DIC-free Aquil synthetic seawater salts (without nutrients), and re-suspended in the DIC-adjusted ^{59}Fe media. Uptake reaction volume was 150 ml, and media was removed from the media bottles by acid cleaned pipette, to avoid unnecessary aeration. Uptake rates were assayed as outlined above, and samples for pH and total carbon analysis removed from parallel (non-radioactive) cultures, poisoned with $HgCl_2$, sealed with Apiezon L vacuum grease (M&I Materials), and stored in borosilicate stoppered bottles for analysis (Dickson et al., 2007).

1.2.9 Approximation of Fe′ and carbonate rate constants (DIC manipulation experiments)

The short-term uptake rates determined from an iron limited culture incubated across a range of DIC and Fe' levels (Extended Data Table 1) were used to determine a second order rate constant for Fe' uptake and to solve for pseudo-first order rate constants with respect to Fe' and carbonate ion. These results are expressed on a per cell basis, and pertain to cultures of *P. tricornutum* grown under this condition[51], as the starting

ISIP2a protein content per cell was identical (but unknown) in all short-term uptake experiments. Uptake rates determined from the two treatments where no DIC (as $NaHCO_3$) was added were excluded. The slope of uptake versus the product of [Fe'] and [CO_3^{2-}] (1.22 x 10^{-3} mol Fe $cell^{-1}$ h^{-1} (M Fe')$^{-1}$) (M CO_3^{2-})$^{-1}$; Extended Data Figure 6a), was used as the starting condition from which an initial CO_3^{2-} rate constant was determined. Assuming Fe uptake rate is first order with respect to Fe' at low Fe (Kustka et al., 2007) (and supported by the ratios of rates observed at varied Fe' at identical [CO_3^{2-}]) the uptake rates measured at Fe' ranging from 2-45 pM (Extended Data Figure 6b) and at variable [CO_3^{2-}] were collapsed by dividing these rates by the Fe' levels (Extended Data Figure 6c). The slope of this quantity versus [CO_3^{2-}] revealed a rate constant for Fe'-normalized Fe uptake of 1.27 x 10^{-3} mol Fe $cell^{-1}$ h^{-1} (M CO_3^{2-}) Then, a pseudo-first order rate constant of Fe' for Fe uptake, k = 0.962 mol Fe $cell^{-1}$ h^{-1} (M Fe')$^{-1}$, was estimated as the quotient of the second order rate constant and the estimated CO_3^{2-} rate constant. The constituent rate constants for Fe uptake were solved by iteration using these initial conditions. Each round was performed with 16000 pairs of trial rate constants, allowed to vary randomly from 0.33 to 3-fold, from the initial or previously iterated values. Residual sum of square (RSS) values were calculated for each pair of trial rate constants. The median constants from the 20 pairs of trial constants with the lowest RSS values (which had coefficients of variation of 0.01% or less) were used for the next trial.

Trials were ceased when constants did not appreciably differ and no trend with trial number was apparent. For each constant, the average of the last four trials are reported; the iterated constants were 0.782 mol (± 0.05) mol Fe $cell^{-1}$ h^{-1} (M Fe')$^{-1}$, and 0.00156 (± 0.00087) mol Fe $cell^{-1}$ h^{-1} (M CO_3^{2-}) for Fe' and carbonate, respectively (about 20% different from initial estimates). The fit of the data using the derived second order rate constants versus empirically measured rates is shown in Extended Data Figure 6d.

1.2.10 pH/pCO_2 manipulation experiments

2 liters of chelexed treated Aquil media in 2.7 L polycarbonate bottles was bubbled with filtered N_2/O_2 mixtures containing 50, 200, 1000 and 5000 ppm CO_2 plus an ambient air (~400 ppm) control. Tubing and lines were acid cleaned PTFE, with acid soaked and cleaned plastic aerators (Lee's Discard-a-Stone) used to maximize gas transfer. In the first experiment (Fe' unconstrained), both high (100 nM FeEDTA) and low (10 nM FeEDTA) media was pre-equilibrated with bubbled gas for 48 hrs. prior to inoculation with a low-iron acclimated culture of wild-type *P. tricornutum* cells. Cells were monitored twice daily for cell counts, chlorophyll and Fv/Fm. Cultures were grown to a low cell density (< 3 x 10^5 cells ml^{-1}) prior to the uptake assay. ^{59}FeEDTA was pre-equilibrated overnight in filter sterilized media from each culture to assure that media chemistry would be similar. Assays were a mix of 135 ml culture and 15 ml ^{59}Fe uptake media, blended without aeration, periodically filtered onto

2.0 µm PTFE filters, washed with Ti-citrate-EDTA solution[50] and preserved in Ecolite (MP Biomedicals) prior to liquid scintillation counting. Short term uptake assays did not exceed 2 hrs. 500 ml of culture was preserved for DIC and Total Alkalinity, with the remaining culture centrifuged and stored for protein expression analysis. Both ISIP2a and β-carbonic anhydrase were analyzed on protein per cell normalized samples (Invitrogen).

For the Fe' constrained experiments, the Fe-to-EDTA ratio was varied to constrain pH-induced changes in [Fe']. For the low iron cultures, we attempted to constrain Fe to ~20 pM Fe' across all five pH/pCO_2 manipulations, and in the high iron cultures, Fe' was constrained to ~200 pM Fe'. In order to keep the maximum concentration of EDTA to 1mM, the total concentration of FeEDTA was reduced to 50 nM in the Fe-replete cultures, and 10 nM in the Fe-limited cultures. Cultures were analyzed as above.

1.2.11 Characterization of endocytosis

To generate fluorescently-labeled cell lines, full length ISIP2a and ISIP1 genes (including 600 base pair native promoter and 300 bp native terminator regions) were amplified with Phusion high fidelity polymerase (New England Biolabs). RFP and YFP protein tags were fused onto constructs using Gibson cloning and assembled into a pUC-19 vector. Full-length constructs were re-amplified, sequence verified, and assembled into p0521s using yeast assembly (Karas, et al., 2015). Clones were PCR screened, with positive clones re-sequenced. For microscopy, cells were

grown to low density (2 x 10^5 cells ml^{-1}) in 7.5 nM FeEDTA Aquil media, centrifuged briefly and resuspended in Aquil containing 300 nM FeEDTA. Cells were immediately visualized on a Leica SP5 confocal microscope equipped with a 100x oil immersion objective. A 514 nm laser was used to visualize YFP, RFP and plastid autofluorescence, with emission monitored at 520–550 nm, 610 – 635 nm, and 700 – 740 nm, respectively. The membrane impermeable dye FM 1-43 was used to stain the outer membrane: wild-type *P. tricornutum* was grown at 7.5 nM FeEDTA, centrifuged briefly and resuspended in Aquil containing 300 nM FeEDTA and 5 µg ml^{-1} of FM 1-43 (Thermo-Fisher). Cells were visualized within 10 minutes of dye addition (excitation 488 nm/emission 610 nm) by confocal microscopy. To evaluate the effect of endocytosis inhibitors, we added 5 µM of the clathrin-mediated endocytosis inhibitor Pitstop2 (N-[5-(4-Bromobenzylidene)-4-oxo- 4,5-dihydro- 1,3-thiazol- 2-naphthalene- 1-sulfonamide, Abcam Inc.) to short term uptake media containing 7.5 nM Fe:EDTA, mixed and assayed immediately, as outlined above.

1.2.12 Analysis of seawater carbonate chemistry

Media samples for pH, TA and DIC were collected following protocols as described in Dickson (Dickson et al., 2007). All samples were collected in 500 ml Pyrex borosilicate glass bottles and poisoned with 200 µl of a saturated $HgCl_2$ solution to arrest metabolic activity, and bottles were sealed using Apiezon-L sealant. For the DIC manipulation experiments, pH and DIC were measured: pH was measured on the total pH scale using

meta-cresol purple (Carter et al., 2013) and DIC analyses were carried out using an AIRICA analyzer (Marianda, Inc.). For the pH/pCO_2 experiments, total alkalinity (TA) was determined using an open-cell potentiometric acid titration system equipped with a Metrohm 876 Dosimat Plus automated titrator and a Metrohm Ecotrode Plus pH electrode, as developed by the Dickson Lab at Scripps Institution of Oceanography. Performance and accuracy of the analyses was evaluated through analysis of Certified Reference Material (CRM) provided by the Dickson lab. Accuracy and Reference Material (CRM) provided by the Dickson lab. Accuracy and precision ± 2 μmol L^{-1} was both DIC and TA analyses. The complete media carbonic-acid system including carbonate ion concentration were calculated based on DIC and TA data based on in situ temperature and salinity using CO2SYS (Lewis et al., 1998)

1.2.13 Statistical analyses

No statistical methods were used to predetermine sample size. All *n* numbers represent biological replicates, and the experiments were not randomized or blinded. All the data in bar graphs are expressed as mean ± s.e.m, with paired Student's *t*-test (two-tailed). Statistical analysis was performed in R (R core team, 2014) , using linear regression (lm), two-way ANOVA (aov), and correlation tests (cor.test), where appropriate, to compare the relationship of Fe' uptake to [Fe'], [CO_3^{2-}] and the [Fe'][CO_3^{2-}] product. Test statistics and p-values are summarized in Extended Data Table 3. $P < 0.05$ was considered significant.

1.3 RESULTS AND DISCUSSION

The iron cargo protein transferrin is thought to have originated in metazoa as an adaptation to multicellularity (Lambert et al., 2005). In vertebrates, serum transferrin circulates in the blood, binding iron with high affinity ($K_d^{Fe'} \approx$ 1-10 pM) and distributing iron to cells via receptor-mediated endocytosis (Baker et al., 2003). Transferrin binds iron tightly by coordinating ferric iron with an exogenous carbonate (CO_3^{2-}). Following endocytosis, the carbonate anion is protonated and ferric iron is reduced, disrupting coordination and initiating the release of iron (Aisen et al., 1978). Recent discoveries of ferric iron concentrating ISIP2a proteins in diatoms (Lommer et al., 2012; Morrissey et al., 2105) and transferrin in the green alga *Dunaliella salina* (Fisher et al., 1997) have suggested that transferrin-like iron acquisition mechanisms are not limited to multicellular organisms. We extracted sequences from a novel marine microeukaryote database (Caron et al., 2016) to reconstruct the phylogenetic history of ISIP2a and transferrin, extending both protein families to a common origin in bacterial periplasmic binding proteins (PBP; Figures 1 and 4). While transferrins and the ISIP2a clade of algal transferrin-like proteins, or phytotransferrins (Anderson and Morel, 1982), are distantly related through anion-binding bacterial PBP, transferrin and phytotransferrin are functional analogs which convergently evolved iron binding. This independent evolution of iron coordination marks the third occasion (Bruns

et al., 1997) where ferric binding proteins have arisen from bacterial PBP and highlights the critical role of the coordinating anion in the transferrin superfamily.

Estimates of the timing of the divergence of transferrin and phytotransferrin from bacterial PBP range from 675 – 913 million years ago, consistent with the inferred changes to Neoproterozoic marine redox states (Och et al., 2012) and the biological requirements to develop novel mechanisms for accessing dilute ferric iron (Figure 5). An archaeal origin of endocytosis-mediated transferrin is intriguing and supports recent discoveries of membrane and vesicular trafficking capacity in marine archaea (Zaremba-Niedzwiedzka et al., 2017). Transferrin was likely vertically transmitted from Archaea to Eukarya, while the roots of phytotransferrin in cyanobacteria and α-proteobacteria suggest Eukaryotic acquisition via endosymbiotic gene transfer. Importantly, transferrin and phytotransferrin are simultaneously present in the genomes of deeply branching prasinophytes, indicating early acquisition by chlorophyte algae (Figure 4). However, the absence of transferrin from rhodophytes and secondary endosymbiotic algae suggests an equally early loss of transferrin from most marine algae. Phytotransferrins are frequently transmembrane anchored (Extended Date Figure 1), a selective advantage for unicellular phytoplankton which would facilitate cell surface iron binding and receptor-less internalization.

To examine the contribution of phytotransferrin to high affinity iron uptake in diatoms, we used TALE nucleases (Weyman et al., 2015) to disrupt *ISIP2a* in the diatom *P. tricornutum* (Figure 6b). In artificial seawater where [Fe'] is precisely controlled with excess EDTA, deletion of *ISIP2a* reduces growth rates of *P. tricornutum* at low pM [Fe'] (Fig. 2a). Deletion of *ISIP2a* reduces short term ^{59}Fe uptake rates at 15 pM [Fe'] by >90%, confirming *ISIP2a* to be critical to high affinity iron uptake (Figure 2b). *ISIP2a* is one of four homologs of phytotransferrin in the *P. tricornutum* genome, however at low [Fe'], ISIP2a accounts for 95% of summed phytotransferrin transcript abundance (Smith et al., 2016). Disruption of *ISIP2a* does not affect uptake rates of iron bound to the siderophore ferrioxamine B, suggesting a separate mechanism for uptake of complexed Fe (Figure 6c).

To determine whether *ISIP2a* is a functional analog of transferrin, we codon-optimized both domains of soluble human serum transferrin (HTF) and fused each lobe to the signal peptide and transmembrane anchor of *ISIP2a* (Figure 2e). These gene products, along with native *ISIP2a* and a YFP-fusion control, were reintroduced into *ΔISIP2a* by conjugation with *E. coli* (Karas et al., 2015). The N-lobe of HTF fully restored high affinity iron uptake equal to that of reintroduced *ISIP2a,* while the C-lobe partially restored uptake (Figure 2b), suggesting a shared binding and internalization mechanism. In human serum transferrin, mutation of the conserved tyrosine iron-binding ligands disrupts iron

binding (Lambert, 2005). In ISIP2a, site directed mutagenesis of the corresponding tyrosines to phenylalanines (Figure 2d) reduced short term uptake rates by half, suggesting that phytotransferrin uses a similar iron coordination mechanism (Figure 2c).

Endocytosis is a requirement for the internalization of transferrin-bound iron (Harding, 1983). In *P. tricornutum*, Fe' uptake rates are sharply reduced when low [Fe'] acclimated cultures are incubated with the clathrin-mediated endocytosis inhibitor Pitstop2 (Figure 2d). Red fluorescent protein labeled ISIP2a localizes to the outer membrane, small internal vesicles and larger endosome-like structures, a pattern consistent with protein internalization via endocytosis (Figure 2e). A putative iron-regulated receptor ISIP1 (Allen et al., 2008; Lommer et al., 2012) was co-localized but did not show any cell-surface interaction with ISIP2a. When stained with membrane impermeable dyes, vesicles contain portions of the outer membrane (Figure 6d). An endocytosis-mediated iron acquisition pathway enables the direct binding and internalization of ferric iron, highlighting a eukaryotic-specific mechanism which likely contributes to the ecological dominance of eukaryotic phytoplankton in iron limited high-nutrient oceans (Boyd et al., 2007).

In both human (Aisen et al., 1978) and algal (Fisher et al., 1997) transferrin, the coordinating carbonate ion is required for the efficient binding and internalization of iron. To examine the role of carbonate in phytotransferrin-mediated Fe' uptake, we measured short-term Fe' uptake

rates of *P. tricornutum* in pH-controlled media amended with dissolved inorganic carbon (DIC; Table 1). Uptake rates of a low-iron acclimated culture of *P. tricornutum* were linearly correlated to [CO_3^{2-}] across a range of [Fe'] (Figure 3a), consistent with a transferrin-like uptake mechanism (Aisen et al., 1978). Uptake rates of iron complexed to Ferrioxamine-B were unaffected by changes in [CO_3^{2-}], further evidence that the uptake of iron complexes proceeds through an alternate mechanism. Fe' uptake rates in *ΔISIP2a* were reduced 85 - 90% but were still influenced by [CO_3^{2-}], likely reflecting the activity of less abundant ISIP2a homologs.

In serum transferrin, iron binds through the synergistic interaction of Fe' and CO_3^{2-}, with neither substrate binding in the absence of the other (Baker et al., 2003). In the DIC manipulation experiments this synergistic interaction (Slinker, 1998) was significant ($p = 2x10^{-6}$). Furthermore, uptake rates plotted as a function of the [Fe'][CO_3^{2-}] product fit a linear model, exhibiting the second order rate dependence characteristic of transferrin (Fig. 3b and Figure 9). As [ISIP2a] was uniform across all uptake measurements, we used the linear regression from the DIC-amended measurements to derive first order rate constants approximations for Fe' and CO_3^{2-} of 0.782 ± 0.05 and 0.00156 ± 0.00087, respectively (Figure 9). This dependence of Fe' uptake rates on [CO_3^{2-}] is linear between 50 and 400 μmol L^{-1} CO_3^{2-} and within the range of the current and year 2100 projected concentrations of seawater carbonate ion (Doney et al., 2009). These results imply that [CO_3^{2-}] currently co-limits

Fe' uptake rates, and that reductions in seawater [CO_3^{2-}] have a negative impact on the ability of *P. tricornutum* to access Fe'.

To further evaluate the interaction of [Fe'] and [CO_3^{2-}] on Fe' uptake, we manipulated the pH/pCO_2 of iron-replete and iron-limited *P. tricornutum* cultures. As pCO_2 was raised from 50 to 5000 ppm CO_2, pH dropped from 8.6 to 7.5, reducing both [Fe'] and [CO_3^{2-}] (Table 2). As pCO_2 increases and pH decreases, [Fe'] decreases at a greater rate than [CO_3^{2-}], therefore changes to [Fe'] appear to drive changes iron uptake rates (Figure 8 a, b), consistent with similar observations on seawater acidification and Fe' bioavailability (Shi et al., 2010). However, our analysis reveals that the [Fe'] and [CO_3^{2-}] interaction is significant (p = 0.00375), and the resulting [Fe'][CO_3^{2-}] product has a more significant fit to a linear model than Fe' alone (Figure 3c and Table 3). These results yield a more accurate understanding of Fe' uptake rates and highlight the importance of the synergistic carbonate ion to high-affinity iron uptake.

Elevated levels of ISIP2a expression is indicative of iron-limited growth (Allen et al., 2008), and in the iron-replete cultures, ISIP2a concentration increases with pCO_2 (Fig. 3d), though the increased expression of ISIP2a may be in response to [Fe'], [CO_3^{2-}] or both. To determine whether changes in [CO_3^{2-}] alone can induce iron limitation and ISIP2a expression, we repeated the above pCO_2 manipulation experiments, altering the Fe:EDTA ratio to constrain changes in [Fe'] while allowing [CO_3^{2-}] to freely vary with pH (Table 3). In these Fe'-constrained cultures,

^{59}Fe uptake rates were correlated with [CO_3^{2-}] and did not correlate with changes in [Fe'] (Extended Data 5 d,e). As before, the [Fe'][CO_3^{2-}] interaction term best explained Fe' uptake rates (Fig. 3c dashed line and Fig.7 d-f). Despite sufficient [Fe'], ISIP2a expression was elevated in the high pCO_2/low [CO_3^{2-}] culture, suggesting that insufficient carbonate ion can induce iron limitation in an otherwise iron-replete culture (Fig. 3e). The notion that high pCO_2 can induce an iron-limited response has also been described in the iron-binding protein Fea1/H43 from *Chlamydomonas reinhardtii,* the only other characterized phytotransferrin (see Figure 1). Fea1/H43 was initially identified as a high pCO_2-inducible biomarker (Hanawa et al., 2007), only later shown to be an iron binding protein (Allen et al., 2007) whose dual regulation by iron and CO_2 was not understood (Baba et al., 2011).

1.4 CONCLUSION

We show that high-affinity iron uptake in *P. tricornutum* is mediated by a phytotransferrin which is exquisitely sensitive to the concentration of carbonate ion. For *P. tricornutum* growing at 5 pM Fe', a reduction in pH from 8.1 to 7.8 and a rise in pCO_2 from 400 to 800 ppm results in a 44.5% decrease in Fe' uptake rates. We view these results as a natural starting point to analyze the influence of carbonate ion on iron uptake rates and acquisition strategies in natural assemblages of marine phytoplankton.

1.5 ACKNOWLEDGEMENTS

Chapter one has been prepared as a Letter to *Nature*. "Carbonate sensitive phytotransferrin controls high affinity iron uptake in diatoms". McQuaid, Jeffrey B.; Kustka, Adam B.; Oborník, Miroslav; Horák, Aleš; McCrow, John P.; Karas Bogumil J.; Zheng, Hong; Kindeberg, Theodor; Andersson, Andreas J.; Barbeau, Katherine A.; Allen, Andrew E. We thank J. Badger for early contributions to phylogenetic analyses, A. Dickson for pH analysis, K. Forsch for CSV measurements, and E. Bertrand for Trace-metal clean techniques. This study was supported by the National Science Foundation (NSF-MCB-1024913, NSF-ANT-1043671, NSF-OCE--0727997, United States Department of Energy Genomics Science program (DE-SC00006719 and DE-SC0008593) and Gordon and Betty Moore Foundation grant GBMF3828 (AEA), and the Czech Science Foundation, project 15-17643S (MO and AH).

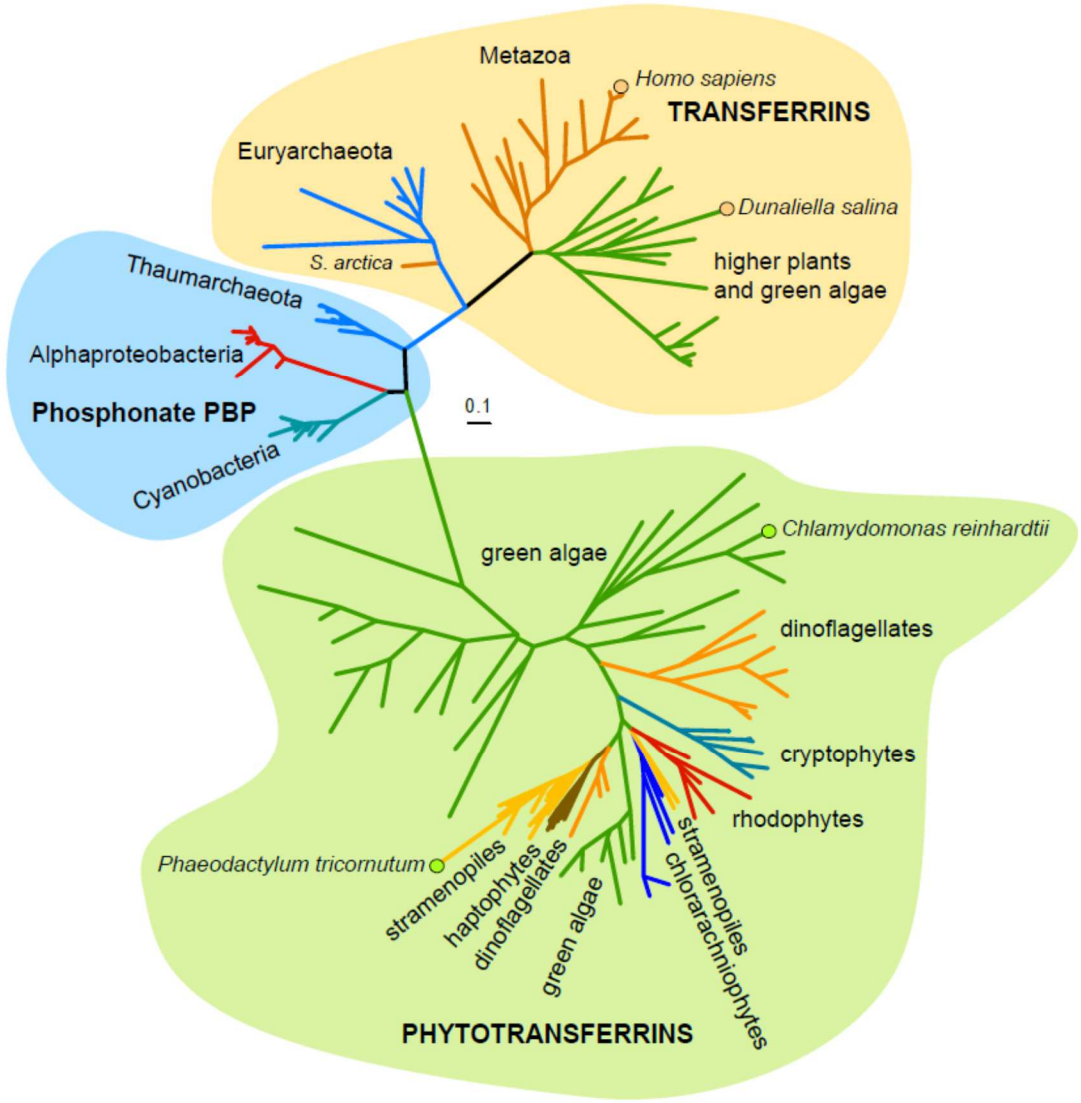

Figure 1: Divergence and phylogenetic distribution of transferrin and phytotransferrin**.** Maximum likelihood phylogenetic tree inferred from amino acid sequences of the divergence of transferrin and phytotransferrin from ancestral phosphonate periplasmic binding proteins (PBP). Circles at branch ends denote iron-binding proteins noted in text: human serum transferrin (*H. sapiens*), triplicated transferrin (Ttf1, *D. salina*), Fe-assimilation protein (Fea1, *C. reinhardtii*) and ISIP2a (*P. tricornutum*). Scale bar indicates 0.1 amino acid substitutions per position.

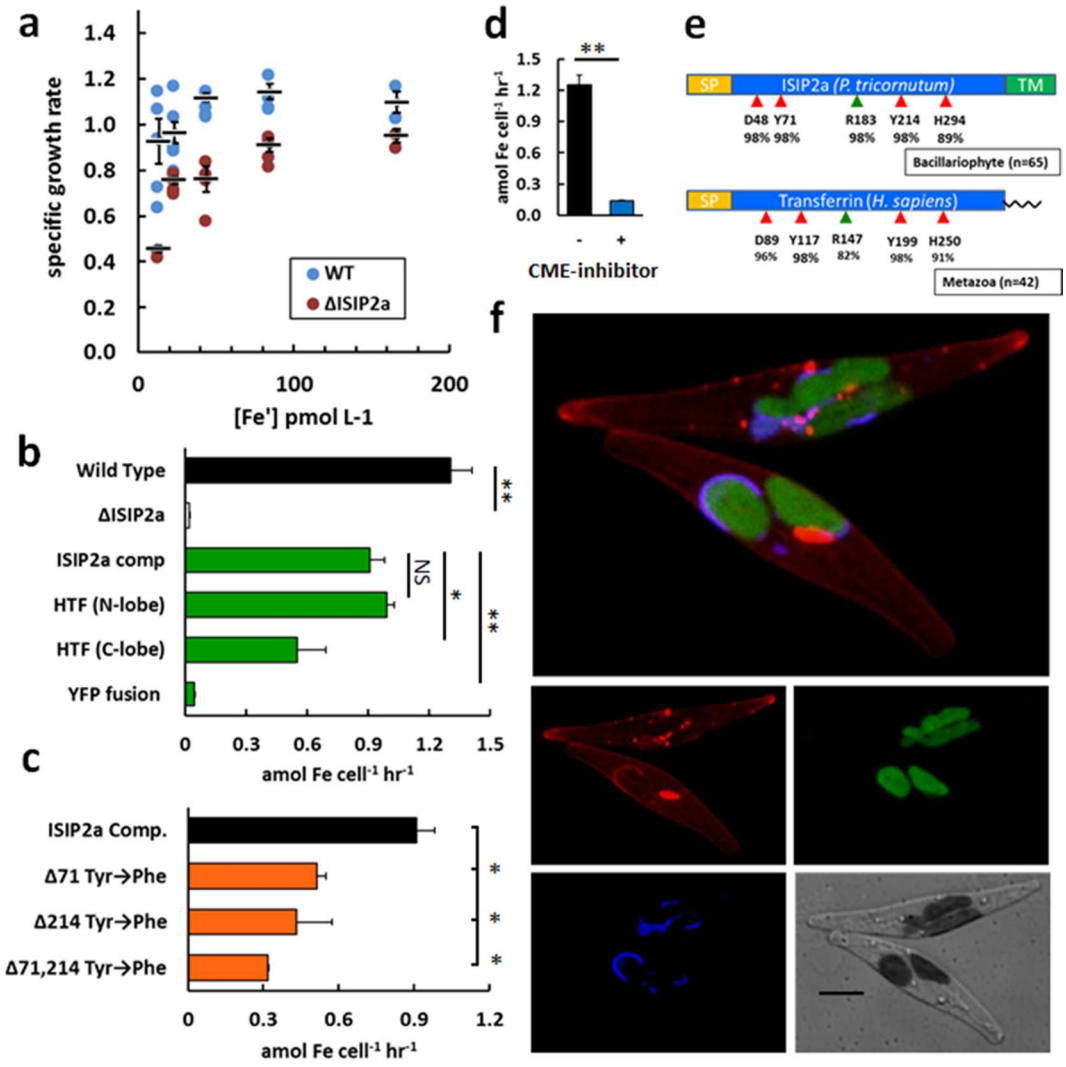

Figure 2: Biochemical characterization of ISIP2a **a)** Specific growth rates, wild-type vs. ΔISIP2a. Bars are mean ± s.e.m. **b)** Short-term ^{59}Fe uptake assays, 15 pM Fe', wild-type vs. knockout and complements. **c)** Short-term ^{59}Fe uptake assays, 15 pM Fe', site-directed mutagenesis strains **d)** Effect of clathrin mediated endocytosis (CME) inhibitor on short-term 15 pM ^{59}Fe' uptake rates **e)** Locations of conserved iron (red) and carbonate (green) coordinating amino acids in *P. tricornutum* phytotransferrin (ISIP2a) and human serum transferrin (N-lobe), with percent conservation in bacillariophyte and metazoan sequences **f)** Outer membrane and vesicular localization of ISIP2a (red), co-localized with an uncharacterized iron-regulated receptor ISIP1 (blue) and plastid (green) in iron-limited *P. tricornutum*. Scale bar is 5 μm. **b-d)** minimum 3 biological replicates (cultures or mutants), error bars are s.e.m, paired t-test, NS not significant, *P < 0.05, ** P < 0.01

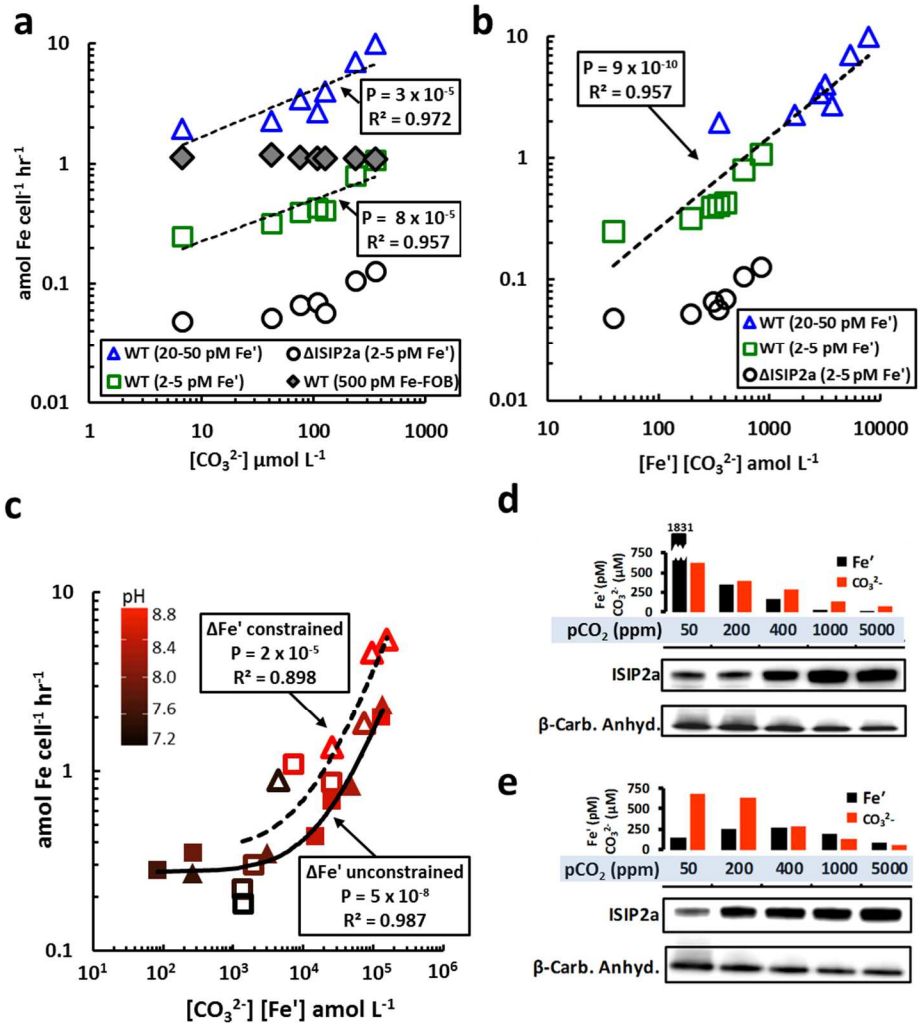

Figure 3: Carbonate and iron interact synergistically to control the rate of Fe′ uptake and expression of ISIP2a **a)** DIC manipulation experiment, uptake vs. $[CO_3^{2-}]$ for a low-iron adapted WT culture assayed at low (2-5 pM) and high (20-50 pM) Fe′, and at 500 pM ^{59}Fe-FerrioxamineB. Uptake rates of ΔISIP2a are shown for contrast. **b)** data from a) plotted as Fe′ uptake rates vs. interaction product $[Fe'][CO_3^{2-}]$ **c)** pH/pCO_2 manipulation experiments. Solid line, filed shapes: Both [Fe′] and $[CO_3^{2-}]$ allowed to covary with pH. Dashed line, hollow shapes: [Fe′] constrained while $[CO_3^{2-}]$ allowed to vary with pH. Squares, Triangles = low, high iron, respectively. Linearized graphs see Extended Data 5. **d, e)** Immunoblots of ISIP2a and B-carbonic anhydrase from pH/pCO_2 manipulations, Fe-replete cultures. [Fe′], $[CO_3^{2-}]$ concentrations (bar chart) and CO_2 levels (50 - 5000 ppm) for d) Fe′ unconstrained and e) Fe′ constrained cultures **a-c)** R^s and P-values refer to linear models (See Extended Data Table 3)

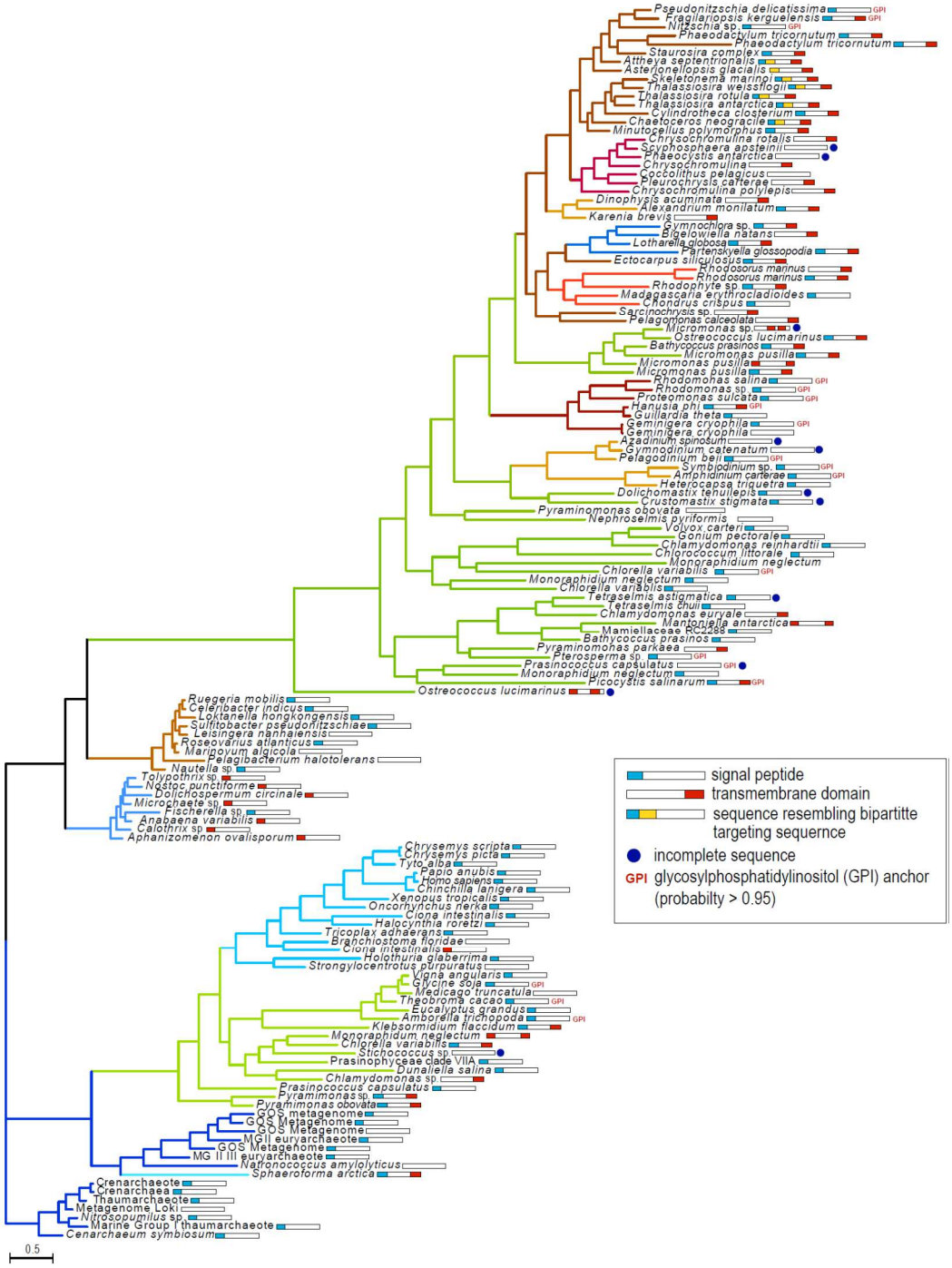

Figure 4: Bayesian (Phylobayes) phylogenetic tree inferred from amino acid sequences of the divergence of transferrin and phytotransferrin from phosphonate-binding Periplasmic Binding Proteins (PBP). Signal (blue) and Transmembrane (red) peptides are shown to right of sequence. Scale bar is 0.5 substitutions/position.

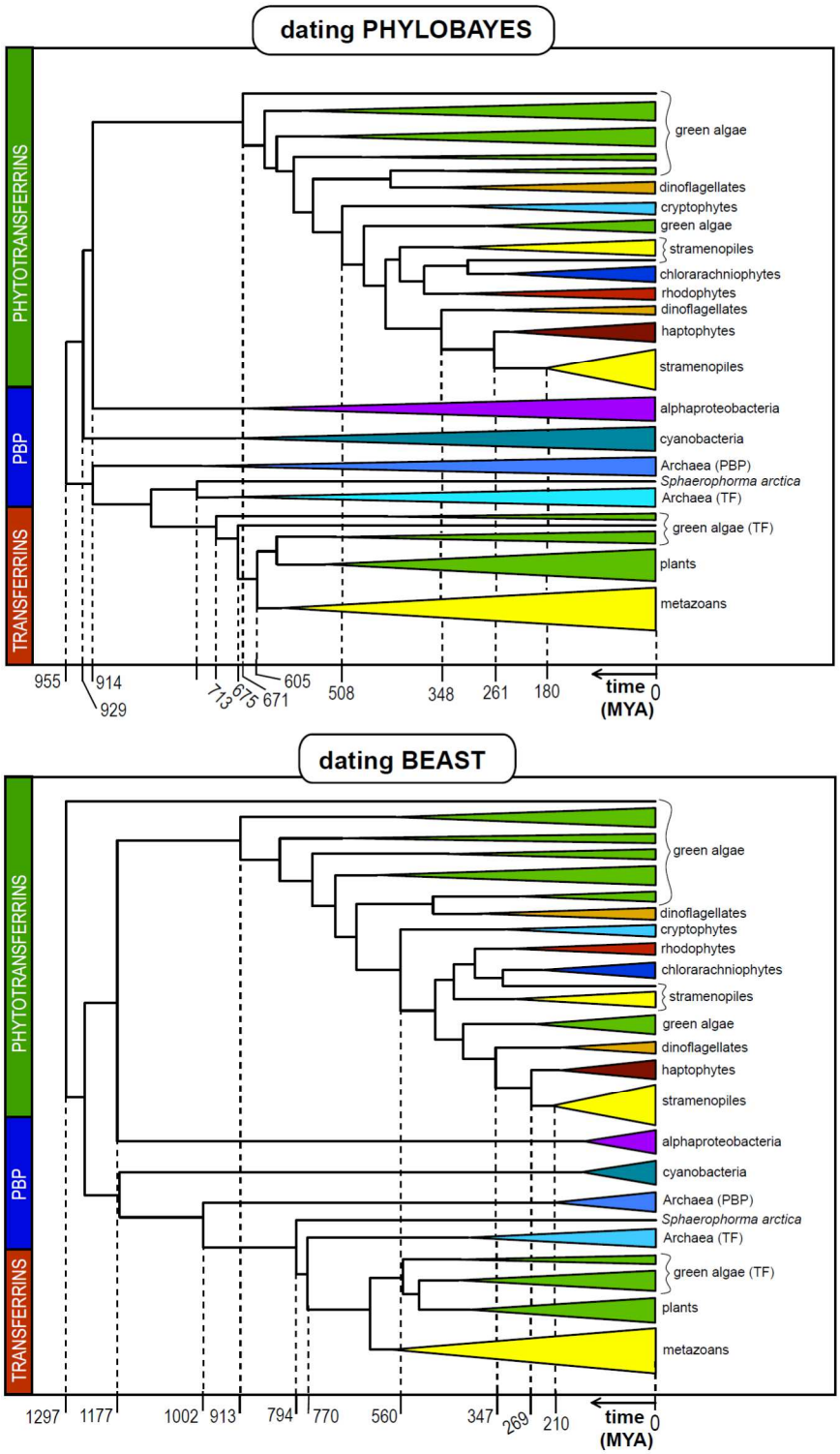

Figure 5: Estimate divergence times between transferrin, phytotransferrin and phosphonate PBP using Phylobayes (top) and Beast (bottom).

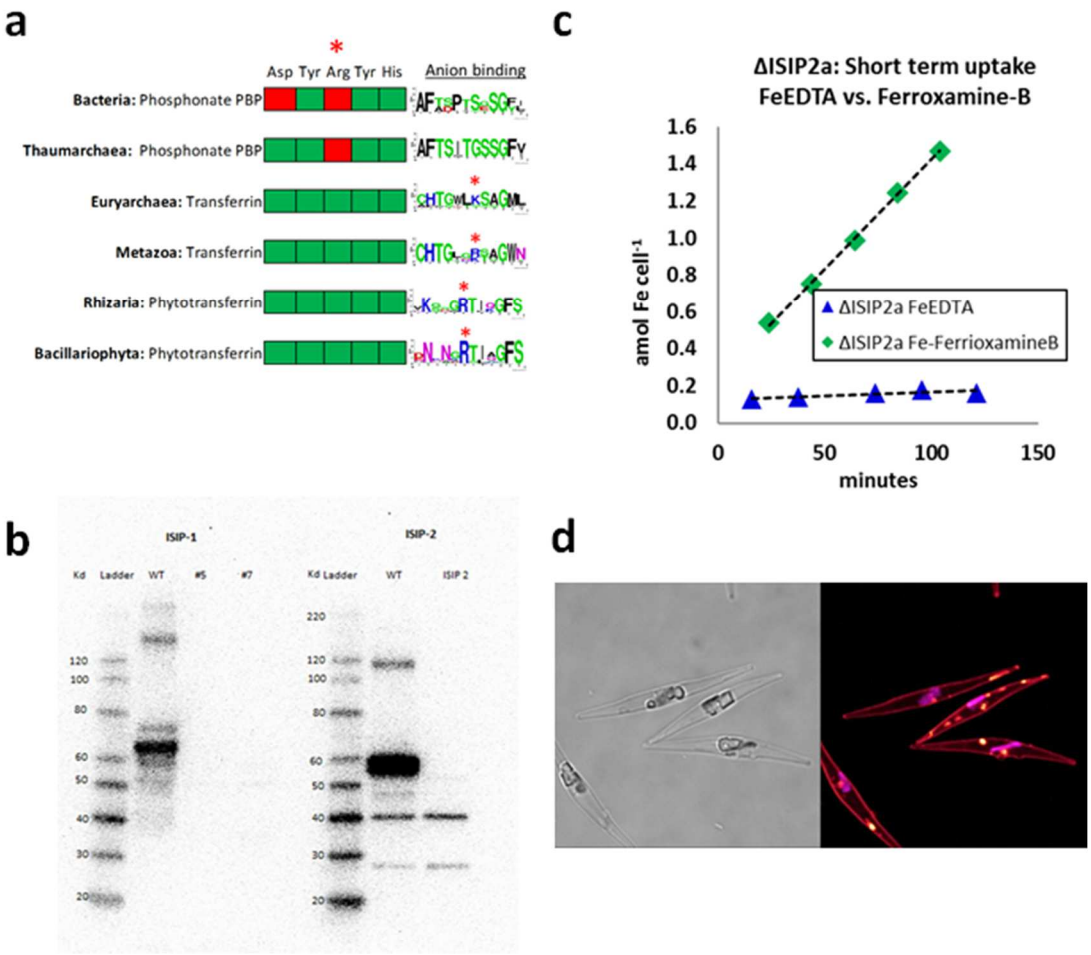

Figure 6: Biochemical characterization of ISIP2a. **a)** Conservation among iron and carbonate-coordinating (*) amino acids between PBP, phytotransferrin and transferrin in select phyla. presence (green) and absence (red) of conserved putative iron (Asp, Tyr, Tyr and His) and carbonate (Arg/Lys). Logo tag of conservation plot of the anion-binding region. **b)** Western blot of WT and ΔISIP1, ΔISIP2a TALEN knockout strains. **c)** Fe-ferrioxamine B vs. Fe-EDTA uptake in ΔISIP2a. **d)** FM 1-43 impermeable outer membrane stain of iron-limited cells pulsed with Fe.

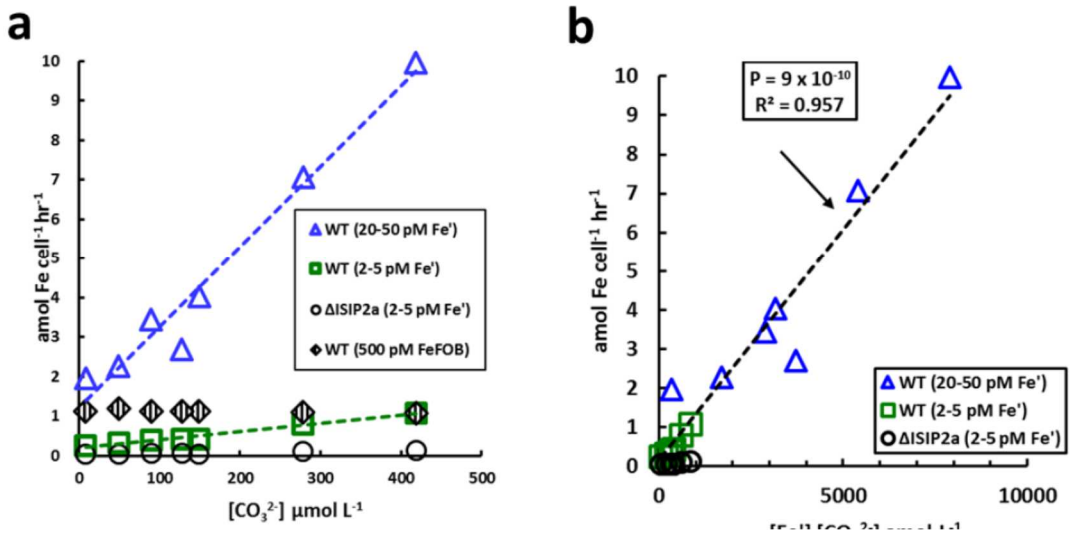

Figure 7: a) linearized version of Figure 3a. b) linearized version of Figure 3b.

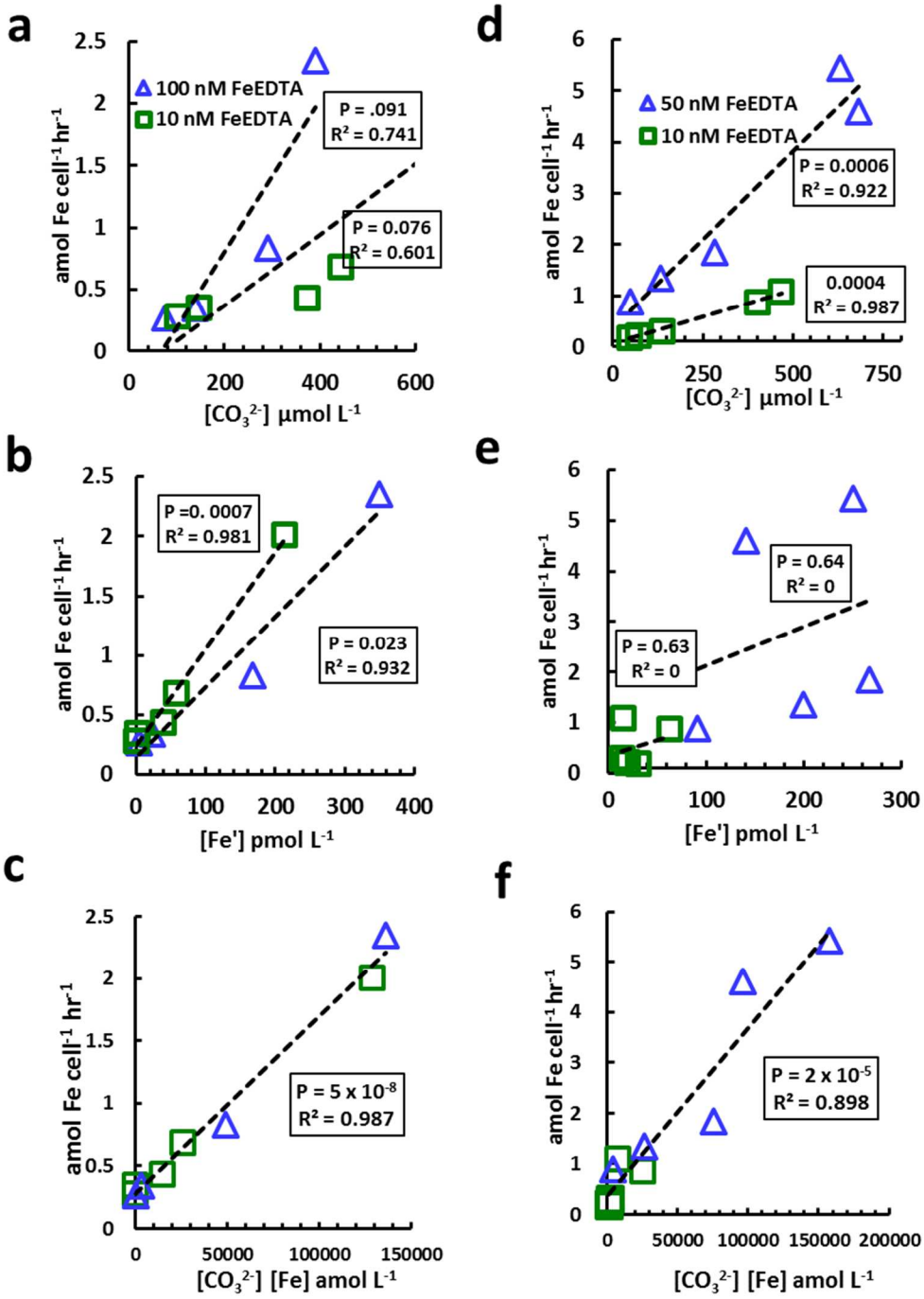

Figure 8: Linearized plots of iron uptake rates, pH/pCO_2 manipulation experiments **a-c)** Unconstrained Fe' experiments a) Carbonate vs. uptake b) iron vs. uptake, and c) [Fe'][CO_3^{2-}] product vs uptake rates. P-vaules, R^2 refer to fit to linear models. **d-f)** Unconstrained Fe' d) Carbonate vs. uptake e) iron vs. uptake, and f) [Fe'], [CO_3^{2-}] product vs uptake. P-values, R^2 refer to fit to linear models. For statistical data see Table 3.

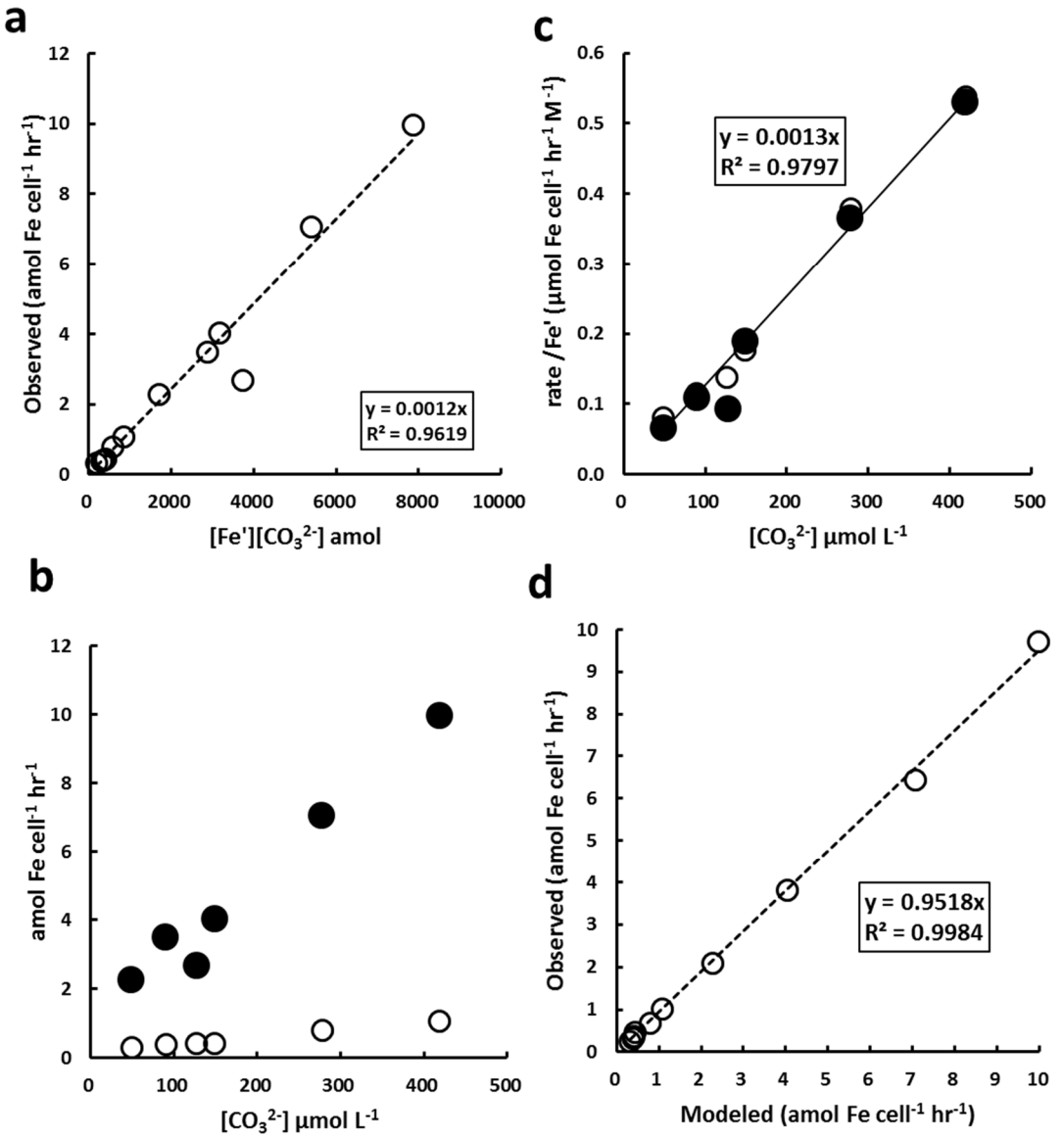

Figure 9: Derivation of second-order and constitutive rate constants for iron uptake in *P. tricornutum* grown with varied DIC and Fe′. **a)** Iron uptake rate versus the product of Fe′ and carbonate concentrations. Regression excludes two observations where DIC was omitted from the media. **b)** Iron uptake rates for treatments incubated with low (2-5 pM; open symbols) and high (20-50 pM; closed symbols) Fe′, versus carbonate concentration. **c)** Uptake rates normalized to the concentrations of Fe′, plotted against $[CO_3^{2-}]$. The slope has units of mol Fe cell^{-1} h^{-1} (M Fe′)$^{-1}$ (M CO_3^{2-})$^{-1}$, equivalent to the pseudo-first order uptake rate with respect to carbonate. **d)** Rates of uptake calculated as a function of Fe′ and $[CO_3^{2-}]$ versus measured rates.

Table 1: Measured (bold) and derived values for DIC manipulation experiments. From Top: Wild Type at 10 nM Fe-EDTA , Wild Type at 100 nM Fe-EDTA , ΔISIP2a at 10 nM Fe-EDTA, Wild Type at 500 pM Fe-FOB. Note TA not measured due to presence of 2mM Tris·HCl in media.

	DIC Treatment	Fe' uptake amol Fe'cell^{-1}·hr^{-1}	[DIC] µmol kg^{-1}	[TA] µmol kg^{-1}	pH	[CO_3] µmol kg^{-1}	[Fe'] pmol L^{-1}	[Fe'][CO_3] fmol L^{-1}	[EDTA] µmol L^{-1}	[Fe_{TOTAL}] nM	log K'$_{dark}$, 20°C
10 nM FeEDTA WildType PT	6 mM DIC	**1.076**	**5557**	-	**7.87**	419	2.0	0.84	100	10.8	-7.731
	4 mM DIC	**0.800**	**3647**	-	**7.88**	278	2.1	0.59	100	10.9	-7.716
	2 mM DIC	**0.412**	**1884**	-	**7.90**	149	2.3	0.35	100	11.0	-7.674
	1.5 mM DIC	**0.433**	**1437**	-	**7.95**	127	3.1	0.40	100	11.0	-7.545
	1 mM DIC	**0.396**	**970**	-	**7.97**	90	3.4	0.31	100	10.8	-7.499
	0.5 mM DIC	**0.318**	**505**	-	**8.00**	49	4.0	0.20	100	10.9	-7.436
	no added DIC	**0.250**	**75**	-	**8.03**	8	4.9	0.04	100	10.9	-7.347

	DIC Treatment	Fe' uptake amol Fe'cell^{-1}·hr^{-1}	[DIC] µmol kg^{-1}	[TA] µmol kg^{-1}	pH	[CO_3] µmol kg^{-1}	[Fe'] pmol L^{-1}	[Fe'][CO_3] fmol L^{-1}	[EDTA] µmol L^{-1}	[Fe_{TOTAL}] nM	log K'$_{dark}$, 20°C
100 nM FeEDTA WildType PT	6 mM DIC	**9.967**	**5557**	-	**7.87**	419	18.8	7.89	100	101.4	-7.731
	4 mM DIC	**7.071**	**3647**	-	**7.88**	278	19.4	5.42	100	101.1	-7.716
	2 mM DIC	**4.032**	**1884**	-	**7.90**	149	21.3	3.17	100	100.5	-7.674
	1.5 mM DIC	**2.694**	**1437**	-	**7.95**	127	29.1	3.71	100	102.0	-7.545
	1 mM DIC	**3.433**	**970**	-	**7.97**	90	32.1	2.88	100	101.2	-7.499
	0.5 mM DIC	**2.264**	**505**	-	**8.00**	49	34.7	1.71	100	94.7	-7.436
	no added DIC	**1.960**	**75**	-	**8.03**	8	45.2	0.36	100	100.4	-7.347

	DIC Treatment	Fe' uptake amol Fe'cell^{-1}·hr^{-1}	[DIC] µmol kg^{-1}	[TA] µmol kg^{-1}	pH	[CO_3] µmol kg^{-1}	[Fe'] pmol L^{-1}	[Fe'][CO_3] fmol L^{-1}	[EDTA] µmol L^{-1}	[Fe_{TOTAL}] nM	log K'$_{dark}$, 20°C
10 nM FeEDTA ΔISIP2a	6 mM DIC	**0.129**	**5557**	-	**7.87**	419	2.0	0.84	100	10.8	-7.731
	4 mM DIC	**0.107**	**3647**	-	**7.88**	278	2.1	0.59	100	10.9	-7.716
	2 mM DIC	**0.057**	**1884**	-	**7.90**	149	2.3	0.34	100	10.9	-7.674
	1.5 mM DIC	**0.070**	**1437**	-	**7.95**	127	3.1	0.40	100	10.9	-7.545
	1 mM DIC	**0.066**	**970**	-	**7.97**	90	3.5	0.31	100	11.0	-7.499
	0.5 mM DIC	**0.052**	**505**	-	**8.00**	49	4.0	0.20	100	10.9	-7.436
	no added DIC	**0.049**	**75**	-	**8.03**	8	4.9	0.04	100	10.9	-7.347

	DIC Treatment	Fe' uptake amol Fe'cell^{-1}·hr^{-1}	[DIC] µmol kg^{-1}	[TA] µmol kg^{-1}	pH	[CO_3] µmol kg^{-1}	[FeFOB] pmol L^{-1}	[Fe'][CO_3] fmol L^{-1}
500 pM FeFOB WildType PT	6 mM DIC	**1.090**	**5557**	-	**7.87**	419	500	-
	4 mM DIC	**1.106**	**3647**	-	**7.88**	278	500	-
	2 mM DIC	**1.117**	**1884**	-	**7.90**	149	500	-
	1.5 mM DIC	**1.116**	**1437**	-	**7.95**	127	500	-
	1 mM DIC	**1.123**	**970**	-	**7.97**	90	500	-
	0.5 mM DIC	**1.187**	**505**	-	**8.00**	49	500	-
	no added DIC	**1.126**	**75**	-	**8.03**	8	500	-

Table 2: Measured (in bold) and derived values for pCO_2/pH manipulation experiments. Top: pCO_2 manipulations where the Fe-EDTA ratio was held constant, and Fe' concentrations were allowed to free-float with changes in pH. Bottom: pCO_2 manipulations where the pH was allowed to change while the Fe to EDTA ratio manipulated, constraining changes in [Fe'] to a narrow range (see column 7).

		pCO_2 TREATMENT	Fe' uptake amol cell^{-1} hr^{-1}	[DIC] µmol kg^{-1}	[TA] µmol kg^{-1}	pH	$[CO_3]$ µmol kg^{-1}	[Fe'] pmol L^{-1}	[Fe'][CO_3] fmol L^{-1}	[EDTA] µmol L^{-1}	[FeEDTA] nM	log K'$_{DARK}$, 21.5°C
pCO_2/pH manipulation, [Fe'] unconstrained	10 nM FeEDTA	50 ppm	**2.01**	**2052.2**	**2937.42**	8.657	605.2	214	129.3	100	14.4	-5.829
		200 ppm	**0.68**	**1902.9**	**2582.82**	8.519	443.8	58	25.7	100	8.5	-6.164
		400 ppm	**0.43**	**1802.4**	**2391.5**	8.453	372.9	40	14.9	100	8.4	-6.325
		1000 ppm	**0.35**	**2162.3**	**2391.01**	7.901	146.5	2	0.3	100	8.5	-7.664
		5000 ppm	**0.28**	**2487.9**	**2623.08**	7.665	100.2	0.8	0.1	100	14.2	-8.237
	100 nM FeEDTA	50 ppm	**7.36**	**1978.7**	**2891.33**	8.695	619.9	1831§	1135.3	100	100.0	-5.737
		200 ppm	**2.35**	**1999.7**	**2596.31**	8.397	390.6	350	136.5	100	100.9	-6.460
		400 ppm	**0.83**	**1988.9**	**2446.12**	8.272	291.8	168	49.0	100	97.5	-6.764
		1000 ppm	**0.35**	**2092.1**	**2311.23**	7.89	138.3	22	3.1	100	109.4	-7.691
		5000 ppm	**0.27**	**2287.9**	**2386.29**	7.579	75.8	3	0.3	100	97.5	-8.446

§The calculated concentration of Fe' exceeds the 750 pmol Fe' L^{-1} precipitation limit, therefore this datapoint was not used in data models

		pCO_2 TREATMENT	Fe' uptake amol cell^{-1} hr^{-1}	[DIC] µmol kg^{-1}	[TA] µmol kg^{-1}	pH	$[CO_3]$ µmol kg^{-1}	[Fe'] pmol L^{-1}	[Fe'][CO_3] fmol L^{-1}	[EDTA] µmol L^{-1}	[FeEDTA] nM	log K'$_{DARK}$, 21.5°C
pCO_2/pH manipulation, [Fe'] constrained	10 nM FeEDTA	50 ppm	**1.089**	**1586.5**	**2335.35**	8.633	468.0	16	8.4	1000	12.0	-5.888
		200 ppm	**0.865**	**1802.9**	**2433.89**	8.471	402.5	63	25.5	100	12.1	-6.281
		400 ppm	**0.301**	**1874.1**	**2103.35**	7.915	138.1	14	1.9	20	12.0	-7.630
		1000 ppm	**0.221**	**2087.8**	**2182.7**	7.557	69.5	19	1.3	2	12.2	-8.499
		5000 ppm	**0.182**	**2899.9**	**2873.93**	7.233	45.6	30	1.4	0.2	11.7	-9.286
	50 nM FeEDTA	50 ppm	**4.605**	**1991.0**	**2986.83**	8.754	682.8	141	96.0	1000	55.2	-5.594
		200 ppm	**5.432**	**2315.3**	**3210.06**	8.609	630.7	250	157.8	250	55.2	-5.946
		400 ppm	**1.847**	**1632.3**	**2103.2**	8.333	282.3	267	75.4	50	55.1	-6.616
		1000 ppm	**1.348**	**1987.4**	**2200.78**	7.868	132.1	200	26.4	5	55.4	-7.744
		5000 ppm	**0.888**	**1917.5**	**1975.22**	7.44	48.9	91	4.5	1	55.4	-8.783

Table 3: Statistical analyses of DIC and pH/pCO_2 manipulations

Linear regression analysis, Fe' uptake

	FIGURE	data	term	DF	RSE	adj. R^2	F-stat	P-value	comment
DIC manipulation (pH constrained)	**3a**	WT Low Fe'	$[CO_3^{2-}]$	5	4.97×10^{-17}	0.972	212	2.76×10^{-5}	
	3a	WT High Fe'	$[CO_3^{2-}]$	5	6.13×10^{-16}	0.957	135	8.29×10^{-5}	
	3a	ΔISIP2a Low Fe'	$[CO_3^{2-}]$	5	8.54×10^{-18}	0.921	71	3.85×10^{-4}	
	Ext. 4c	WT Low Fe'	[Fe']	5	1.86×10^{-16}	0.611	10	2.32×10^{-2}	
	Ext. 4c	WT High Fe'	[Fe']	5	1.95×10^{-15}	0.565	9	3.13×10^{-2}	
	Ext. 4c	ΔISIP2a Low Fe'	[Fe']	5	2.12×10^{-17}	0.514	7	4.23×10^{-2}	
	3b	WT (Combined Low/High)	$[Fe'][CO_3^{2-}]$	12	5.96×10^{-16}	0.957	292	8.70×10^{-10}	
pCO_2/pH manipulation (Fe' unconstrained)	**Ext. 5a**	WT Low Fe'	$[CO_3^{2-}]$	3	4.54×10^{-16}	0.601	7	7.68×10^{-2}	Not Significant
	Ext. 5a	WT High Fe'	$[CO_3^{2-}]$	2	4.92×10^{-16}	0.741	10	9.06×10^{-2}	Not Significant
	Ext. 5b	WT Low Fe'	[Fe']	3	9.81×10^{-17}	0.981	212	7.02×10^{-4}	
	Ext. 5b	WT High Fe'	[Fe']	2	2.52×10^{-16}	0.932	42	2.30×10^{-2}	
	Ext. 5c.	WT (Combined Low/High)	$[Fe'][CO_3^{2-}]$	7	9.13×10^{-17}	0.987	588	5.16×10^{-8}	
pCO2/pH manipulation (Fe' constrained)	**Ext. 5d**	WT Low Fe'	$[CO_3^{2-}]$	3	4.85×10^{-17}	0.987	292	4.36×10^{-4}	
	Ext. 5d	WT High Fe'	$[CO_3^{2-}]$	3	5.74×10^{-16}	0.922	48	6.15×10^{-4}	
	Ext. 5e	WT Low Fe'	[Fe']	3	4.60×10^{-16}	0	0	6.34×10^{-1}	Not Significant
	Ext. 5e	WT High Fe'	[Fe']	3	2.27×10^{-15}	0	0	6.47×10^{-1}	Not Significant
	Ext. 5f	WT (Combined Low/High)	$[Fe'][CO_3^{2-}]$	8	5.90×10^{-16}	0.898	80	1.93×10^{-5}	

Significant interactions: 2-way Analysis of Variation (ANOVA), Fe' uptake

FIGURE	data	term	DF	Mean Sq	F	p-value
3b	**DIC manipulation (pH constrained)**	[Fe']	1	3.71×10^{-29}	131	4.55×10^{-7}
		$[CO_3^{2-}]$	1	4.13×10^{-29}	146	2.73×10^{-7}
		$[Fe'][CO_3^{2-}]$	1	2.68×10^{-29}	95	2.02×10^{-6}
3c Solid Line	**pH/pCO2 manipulation (Fe' unconstrained)**	[Fe']	1	4.53×10^{-30}	450	4.33×10^{-6}
		$[CO_3^{2-}]$	1	1.10×10^{-31}	11	$2.14\ 10^{-2}$
		$[Fe'][CO_3^{2-}]$	1	2.63×10^{-31}	26	3.75×10^{-3}
3c Broken Line	**pH/pCO2 manipulation (Fe' constrained)**	[Fe']	1	1.38×10^{-29}	43	5.89×10^{-4}
		$[CO_3^{2-}]$	1	1.22×10^{-29}	38	8.10×10^{-4}
		$[Fe'][CO_3^{2-}]$	1	2.76×10^{-30}	9	2.58×10^{-2}

Table 4: Fossil calibration points used as minimum and maximum constraints used for the estimate of divergence of transferrins and phytotransferrins from common prokaryotic ancestor.

	maximum	minimum
node	age of appearance	
mammal	165	62
amniote	333	318
tetrapod	351	337
vertebrate	450	421
metazoan	636	550
embryophyte	133	125
diatom	185	115
haptophyte	260	203

1.6 REFERENCES

Aisen, P., Leibman, A., & Zweier, J. L. Stoichiometric and site characteristics of the binding of iron to human transferrin. *J. Biol. Chem.* **253**, 1930–1937 (1978).

Allen, A. E. et al. Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10438–43 (2008).

Allen, M. D., del Campo, J. a, Kropat, J. & Merchant, S. S. FEA1, FEA2, and FRE1, encoding two homologous secreted proteins and a candidate ferrireductase, are expressed coordinately with FOX1 and FTR1 in iron-deficient *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **6**, 1841–52 (2007).

Anderson, M. A., & Morel, F. M. The influence of aqueous iron chemistry on the uptake of iron by the coastal diatom *Thalassiosira weissflogii*. *Limnol. Oceanogr.* **27,** 789-813 (1982).

Baba, M., Hanawa, Y., Suzuki, I. & Shiraiwa, Y. Regulation of the expression of H43/Fea1 by multi-signals. *Photosynth. Res*. **109**, 169–177 (2011).

Baker, H. M., Anderson, B. F. & Baker, E. N. Dealing with iron: common structural principles in proteins that transport iron and heme. *Proc. Natl. Acad. Sci. U. S. A*. **100**, 3579–3583 (2003).

Barbeau, K., Rue, E. L., Bruland, K. W., & Butler, A. Photochemical cycling of iron in the surface ocean mediated by microbial iron (III)-binding ligands. *Nature*, **413**, 409-413 (2001).

Berney, C., & Pawlowski, J. A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. *Proc Royal Soc London B: Biol Sci*, **273** 1867-1872. (2006)

Bertrand, E. M. et al. Phytoplankton -bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 9938–9943 (2015).

Boyd, P. W., et al. Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science* **315**, 612–617 (2007).

Bruns, Christopher M., et al. Structure of *Haemophilus influenza* Fe^{3+}-binding protein reveals convergent evolution within a superfamily. *Nature structural biology* **4,** 919-924 (1997).

Caron, D.A., et al. Probing the evolution, ecology, and physiology of marine protists using transcriptomics. *Nat. Rev. Microb.* **15**, 6-20 (2016).

Carter, B. R., Radich, J. A., Doyle, H. L., & Dickson, A. G. An automated system for spectrophotometric seawater pH measurements. *Limnol. Oceanogr. Methods*, **11** (2013)

Dickson, A. G., Sabine, C. L., & Christian, J. R. Guide to best practices for ocean CO2 measurements. (2007).

Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. Ocean acidification: the other CO2 problem. *Marine Science* **1,** (2009).

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular biology and evolution*, **29** 1969-1973 (2012)

Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J.l of Molecular Biology* **300**: 1005-1016 (2000).

Falciatore, A., Casotti, R., Leblanc, C., Abrescia, C., & Bowler, C. Transformation of nonselectable reporter genes in marine diatoms. *Marine Biotechnology*, **1** 239-251. (1999).

Fisher, M., Gokhman, I., Pick, U. & Zamir, A. A structurally novel transferrin-like protein accumulates in the plasma membrane of the unicellular green alga Dunaliella salina grown in high salinities*. J. Biol. Chem.* **272**, 1565–1570 (1997).

Fisher, M., Zamir, A. & Pick, U. Iron Uptake by the Halotolerant Alga Dunaliella Is Mediated by a Plasma Membrane Transferrin. *J. Biol. Chem.* **273**, 17553-17558 (1998).

Gibson, D. G., et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods*, ***6***, 343-345 (2009).

Gledhill, M., & Buck, K. N. The organic complexation of iron in the marine environment: a review. *Front. Microbiol*. **3**, 1–17 (2012).

Gouy, M., Guindon, S. and Gascuel, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol*, ***27***, 221-224. (2010)

Hanawa, Y. et al. Induction of a high-CO_2-inducible, periplasmic protein, H43, and its application as a high-CO_2-responsive marker for study of the

high-CO_2-sensing mechanism in Chlamydomonas reinhardtii. *Plant Cell Physiol.,* **48**, 299–309 (2007).

Harding, Clifford, John Heuser, and Philip Stahl. "Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes." J. Cell Boil. **97**, 329-339 (1983).

Hudson, R. J., & Morel, F. M. Distinguishing between extra-and intracellular iron in marine phytoplankton. *Limnology and Oceanography*, **34**, 1113-1120 (1989).

Karas, B. J., et al. Designer diatom episomes delivered by bacterial conjugation. *Nature communications*, ***6*** (2015).

Katoh, K. and Standley, D.M., MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, **30,** 772-780 (2013)

Kustka, A. B., Allen, A. E., & Morel, F. M. Sequence analysis and transcriptional regulation of iron acquisition genes in two marine diatoms1. *Journal of Phycology*, **43** 715-729. (2007)

Lambert, L. A, Perri, H., Halbrooks, P. J. & Mason, A. B. Evolution of the transferrin family: conservation of residues associated with iron and anion binding. *Comp. Biochem. Physiol.* **142**, 129–41 (2005).

Lartillot, N., Lepage, T., & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics*, **25** 2286-2288. (2009).

Lewis, E., Wallace, D., & Allison, L. J. *Program developed for CO2 system calculations* (p. 38). Tennessee: Carbon Dioxide Information Analysis Center, managed by Lockheed Martin Energy Research Corporation for the US Department of Energy. (1998)

Lommer, Markus, et al. "Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation." *Genome biology* **13**, 1-20 (2012)

Marchetti, A. et al. Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 317-325 (2012).

Moore, C. M., et al. Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**, 701–710 (2013).

Moller S, Croning MDR, Apweiler R (2001) Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* **17**:646-653

Morel, F. M. M., Kustka, A. B. & Shaked, Y. The role of unchelated Fe in the iron nutrition of phytoplankton. *Limnol. Oceanogr*. **53**, 400–404 (2008).

Morrissey, J. et al. A novel protein, ubiquitous in marine phytoplankton, concentrates iron at the cell surface and facilitates uptake. *Curr. Biol.* **25**, 364–371 (2015).

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10**:1-6 (1997)

Och, L. M. & Shields-Zhou, G. A. The Neoproterozoic oxygenation event: Environmental perturbations and biogeochemical cycling. *Earth-Science Rev*. **110**, 26–57 (2012).

Parfrey, L. W., Lahr, D. J., Knoll, A. H., & Katz, L. A. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc. Natl. Acad. Sci. U. S. A*. **108** 13624-13629. (2011).

Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**:785-786 (2011)

Pierleoni A, Martelli PL, Casadio R. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics* **9**:392 (2008).

Poulsen, N., & Kröger, N. A new molecular tool for transgenic diatoms. *Febs Journal*, **272**, 3413-3423 (2005).

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Smith, S. R., J. T. Gillard, et. al. Transcriptional Orchestration of the Global Cellular Response of a Model Pennate Diatom to Diel Light Cycling under Iron Limitation. *PLoS Genetics*, **12** (12), *PLoS Genetics* (2016).

Shi, D., Xu, Y., Hopkinson, B. M. & Morel, F. M. M. Effect of Ocean Acidification on Iron Availability to Marine Phytoplankton. *Science* **327**, 676–679 (2010).

Slinker, B. K. The statistics of synergism. *J. Mol. Cell. Cardiol*. **30**, 723–731 (1998).

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312-1313. (2014).

Sunda, William G., Neil M. Price, and Francois MM Morel. "Trace metal ion buffers and their use in culture studies." Algal culturing techniques (2005): 35-63.

Weyman, P. D. et al. Inactivation of Phaeodactylum tricornutum urease gene using transcription activator-like effector nuclease-based targeted mutagenesis. *Plant Biotechnol. J.* **13**, 460–470 (2015).

Zaremba-Niedzwiedzka, K., et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017)

Zaslavskaia, L. A., Lippmeier, J. C., Kroth, P. G., Grossman, A. R., & Apt, K. E. Transformation of the diatom Phaeodactylum tricornutum (Bacillariophyceae) with a variety of selectable marker and reporter genes. *Journal of Phycology*, **36** 379-386. (2000).

CHAPTER 2: METAGENOMIC SURFEY OF MERTZ GLACIER POLYNYA REVEALS MICROBIAL ADVECTION INTO THE DEEP OCEAN.

2.1 ABSTRACT

As part of the 2007/08 International Polar Year, a metagenomic survey was undertaken in the Mertz Polynya region of East Antarctica to document the diversity and genomic potential of marine planktonic communities. An examination of microeukaryotic community DNA in continental shelf and slope regions confirmed the influence of silica in structuring planktonic communities. An analysis of ultra-cold and ultra-dense Ice Shelf Water emerging from underneath the Mertz Glacier Tongue revealed ice shelf archaeal populations becoming entrained in the formation of dense shelf water. These ammonia oxidizing archaea are ultimately advected down the continental slope to form Antarctic Bottom Water (AABW). We hypothesize that ice shelves in regions of active bottom water formation serve as reservoirs of ammonia oxidizing archaea, seeding the deep ocean with continuous populations of surface-derived archaeal communities.

2.2 INTRODUCTION

Antarctic continental shelf regions contain a variety of ecotopes and biogeochemical gradients, thus are excellent natural laboratories for studying the rapidly changing dynamics of ice cover, stratification, convective mixing and nutrient availability on phytoplankton community structure and function. Of particular interest are latent-heat polynya ecosystems, ice free regions where the intense activity of katabatic winds

drive sea ice away from the continent, creating highly productive zones of open ocean (Arrigo, 2003) and fueling the production of dense shelf waters which sink and form Antarctic Bottom Water (AABW; Rintoul, 1998). The relatively high productivity of polynya ecosystems compared to the surrounding water has been attributed to early insolation and ice-melt stratification of the water column (Sambrotto, 2003). Polynyas are particularly known for large blooms of diatoms (Beans, 2008) dinoflagellates (Dennet et al. 2001) and haptophytes (Gowing et al. 2001), and rates of primary production can exceed 1g C m^{-2} d^{-1} (Arrigo et al, 2000). The importance of polynyas in past climate changes is apparent in the thick layers of diatomaceous sediment present in polynya regions (Presti et al, 2003). The intensity and rate of climate change in polar regions has affected East Antarctic polynyas – these changes include rising sea surface temperatures (Steig, 2009), changes in the extent of sea ice coverage, freshening and reduction in AABW formation (Marsland, 2007; Broeker, 1999), wind circulation changes (Boning, 2008) and acidification due to increased CO_2 dissolution into polar waters (McNeal, 2008).

As part of the International Polar Year (IPY), the Collaborative East Antarctic Marine Census (CEAMARC) was initiated under the auspices of the Census of Antarctic Marine Life (CAML). The collaboration involved scientists from 30 countries and 14 research vessels in a coordinated survey of marine life and oceanography in the continental waters surrounding East Antarctica, and took place during the austral summer

(Dec 2007 – Jan 2008). The objective of the research cruise was to investigate biological diversity and understand the potential effects and influences of ongoing climate change. In addition to research targeting the continental shelf areas, the Climate of Antarctica and the Southern Ocean (CASO) reoccupied the WOCE SR03 transect from Tasmania to Adèlie Land in Antarctica. The goal of CASO is to obtain a synoptic circumpolar snapshot Climate of the Southern Ocean and to understand the complex interactions between the atmosphere, ocean and the biogeochemical markers which define zonal and meridional transport.

Here we describe a size-fractionated metagenomic survey of Antarctic surface and mesopelagic waters. A companion study examined prokaryotic distribution and functional genomics of the surface waters (Wilkins et al. 2013). The results presented here will focus on eukaryotic diversity in the larger size classes (0.8 – 20 μm). In addition, these results will examine five unpublished samples collected in the bathypelagic zone (180 – 3600 m) and establish a mechanism by which surface population of ammonia-oxidizing Thaumarchaea are advected into the deep ocean via the formation of Antarctic Bottom Water.

2.3 METHODS

2.3.1 Physiochemical data collection.

During the CEAMARC-AU0803 IPY 2007 cruise, the RSV Aurora Australis occupied stations over the continental shelf near Adèlie-George V

Land between December 22, 2007 and January 19, 2008. A total of 130 CTD casts were collected during the CEAMARC cruise, of which 74 were taken on the shelf between 138°E and 149°E and between the Antarctic coast and the shelf break (about 65°40'S at 140°E), providing a near-synoptic snapshot of the region (Hosie et al. 2011; G. D. Williams et al. 2010)

CEAMARC hydrographic casts were acquired using a SeaBird SBE9-plus CTD, and dissolved oxygen was measured with a SBE43 sensor. All data were processed using standard filtering techniques to eliminate high frequency noise remove outliers (Rosenberg and Rintoul, 2010). CTD casts were supplemented by water samples collected with Niskin bottles mounted on a 24 bottle rosette frame (Rosenberg and Rintoul, 2010). Measurements of water sample salinity and dissolved oxygen content were used to calibrate the CTD data.

2.3.2 Water collection, size fractionation and sample preservation

At each station, water was collected by filtration of 200L of seawater using a shipboard diaphragm pump. Water was pre-filtered through a 20 µm mesh filter to exclude large zooplankton, and the entire volume was passes sequentially across three 293 mm PES filters consisting of 3.0 µm, 0.8 µm and 0.1 µm size class filters. Filters were stored in 1X Tris-EDTA (TE) buffer (Invitrogen) containing 50mM EGTA and 50mM EDTA and frozen at -80°C. For the bathypelagic samples, water was collected by filling all 24 bottle of the ships Niskin-bottle rosette, and samples were

obtained within 10 meters of the bottom. Filtration and preservation proceeded as above. Coordinates and environmental metadata, including temperature, nutrients, and Chlorophyll a, is given for the 23 sample sites in Table 1.

2.3.3 Sample extraction and library construction

In the laboratory, filters were thawed to 20°C and extracted by placing sliced filter fragments into 50 ml conical centrifuge tubes with lysis buffer, followed by phenol-chloroform extraction and ethanol purification as described (Rusch, 2007). Isolated DNA was quality verified on agarose gels prior shearing and adaptor ligation. Agencourt AMPure beads (Beckman) were used to purify the product followed by emulsion PCR and sequencing on a 454 GS FLX 25Ti (Roche).

To remove artificial replicates in 454-Titanium generated sequences the approach reported by Gomez-Alvarez et al. (2009) was used with the following parameters: raw sequences sharing >90% nucleotide identity beginning with the same 3 nucleotides were removed. 454-sequence post processing, including chimera identification and frameshift identification, was performed as described previously (Zeigler Allen et al. 2012; Yooseph et al. 2013).

2.3.4 Metagenome Sequence Annotation

FragGeneScan was used to predict the Open Reading Frames (ORF), and these ORFs were annotated *de-novo* by mapping protein clusters against PhyloDB, an inhouse database containing public sequences and

data from the Marine Microbial Eukaryote Transcriptome Sequencing Project (Keeling et al. 2014), KEGG, Ensembl, IMG and several in-house assemblies. A Lineage Probability Index (LPI) was used to assign a taxonomic classification to that ORF (Podell et al. 2007) and prevent incorrect taxonomic assignment of conserved sequences. LPI was also used to screen for and remove bacterial/archaeal contaminants from the reference dataset.

2.4 RESULTS AND DISCUSSION

2.4.1 Eukaryotic diversity in the euphotic zone

During the international polar year CEAMARC cruise, the Mertz Polynya was largely ice free, allowing samples to be collected from the Mertz Glacier Tongue to the French Antarctic research station Dumont D'Urville, and out to the edge of the continental slope at 65°S. After the Ross and Weddell Sea Polynyas, the Mertz Polynya is the third largest source area for the formation of Antarctic Bottom Water (Rintoul 1998; Guy D. Williams et al. 2008; Shadwick, Tilbrook, and Williams 2014). AABW is a critical component of global overturning circulation (Talley 2013).

The Mertz Polynya remains ice free through the winter due to the intense katabatic winds which blow sea ice offshore (Rintoul 1998). The physical barrier of the ice shelf-like Mertz Glacial Tongue impedes the replenishment of sea ice from the east, creating 'ice factory' like conditions where increased brine exclusion and latent heat loss lead to dense water

formation (Guy D. Williams et al. 2008). In addition to fueling the production of AABW, the lack of ice cover on the polynya allows for increased phytoplankton productivity in early spring as most of the Antarctic littoral zone is covered in pack and land-fast ice. Like many polynyas, the Mertz Polynya is highly productive with a diverse assemblage of phytoplankton (Kevin R. Arrigo 2003; Wright et al. 1996; Marchant et al. 2000; Strutton et al. 2000).

Our metagenomic survey occupied seven surface water stations north of the polar front in the Southern Ocean High-Nutrient Low Chlorophyll (HNLC) region and eleven surface water stations south of the Polar Front in the Mertz Polynya region. The survey also included five bathypelagic stations on the continental shelf and abyssal plain. At all stations the large filter fractions (3 – 20 μm) were dominated by eukaryotic DNA, and Shannon Diversity indices of the taxonomic assignments to these metagenomic reads revealed higher eukaryotic diversity in the nearshore polynya regions when compared to the HNLC regions beyond the shelf. Prokaryotic diversity had an opposite distribution, with a higher Shannon index in the HNLC region and lower in the Mertz Polynya. The diversity index was highest for all fractions in the bottom water samples, and similar high diversity has been found other in meso- and bathypelagic samples from other deep ocean microbial assemblages (Shi et al. 2011; Tolar et al. 2016; Konstantinidis et al. 2009).

Taxonomic analysis of the 0.8 – 20 μm filters revealed that the eukaryotic community structure falls into four groupings correlated to seawater temperature and silica concentration (Figures 11, 12). Nearshore areas encompassing the Mertz Polynya and the continental shelf region of East Antarctica are high in silica (50 – 78 μmol kg^{-1}; Table 5), and these stations were enriched in silicified phytoplankton. At least 30% of the reads mapped to bacillariophyte, including large centric (e.g. *Chaetoceros*) and chain pennates (*Nitzschia* and *Pseudonitzschia* spp.), as well as the cosmopolitan sea ice diatom *Fragilariopsis cylindrus* (Figure 10) The silicified heterokont order Dictyocha was also enriched in the areas south of the polar front, and *Phaeocystis antarctica* was present at all stations and was particularly abundant on the 0.8 μm filter, indicating the free-living (non-colonial) form was dominant (DiTullio et al. 2000).

In contrast, the regions beyond the Antarctic Divergence retain high levels of nitrate and phosphate but are low in dissolved silica (Table 5). Metagenomic reads from beyond the Antarctic Divergance were enriched in non-silicified organisms: pelagophytes, isochrysid haptophytes, and prymnesiophytes. Among silicified phytoplankton, there is a shift away from heavily silicified diatoms and towards open ocean diatoms with lighter frustules like the Southern Ocean cosmopolitan *Fragilariopsis kerguelensis*, as well as smaller species (e.g. *Thalassiosira miniscula*).

North of the subtropical front, summer seawater temperatures rise above 10°C and eukaryotic reads from these stations show the emergence

of chlorophyte algae, with reads from diatoms and *P. Antarctica* reduced in the dataset. North of the subtropical front is also the region where cyanobacteria begin to appear in significant numbers (Wilkins et al. 2013). The bottom water samples had fewer identifiable reads and contained numerous euphotic zone organisms, particularly armored phytoplankton, suggesting transport via sinking particles.

Iron is a critical limiting micronutrient and exerts a major control on phytoplankton productivity and community structure. Recently a ferric iron binding protein, ISIP2a, has been characterized in the diatom *Phaeodactylum tricornutum* (Chapter 1, this dissertation). ISIP2a is a phytotransferrin which allows eukaryotic phytoplankton to access picomolar concentrations of unchelated iron (Fe'), concentrations typical for the southern ocean (Morel, Kustka, and Shaked 2008). A plot of ISIP2a abundance in the metagenomic dataset shows widespread geographic and taxonomic distribution among southern ocean phytoplankton groups (Figure 10), though the abundance cannot be estimated due to the lack of genome size information for many organisms.

2.4.2 Advection of surface microbes to the deep

The Mertz Polynya is an active site of Antarctic Bottom Water formation (Rintoul 1998). While most of this production takes place in the winter, AABW continues to flow over the Adèlie sill through the summer (Lacarra et al. 2011). Dense shelf waters are formed through the mixing of High Salinity Shelf Waters (HSSW), formed through wintertime brine

exclusion, and ultra-cold, dense Ice Shelf Water (ISW), which emerges from under the floating Mertz Glacier Ice Shelf (Mertz Glacier Tongue or MGT). ISW sinks to depth in the Mertz and Adèlie Basins, where it overflows the Adèlie sill and advects to the deep ocean, mixing with Circumpolar Deep Water to create AABW. This water subsequently spreads across the southern portions of the Atlantic, Indian and Pacific Ocean basins, in part propelling global overturning circulation (Talley 2013).

A surface sample of ISW emerging from the Mertz Glacier Tongue (Buchannan Bay Station 353) had a large number of reads mapping to Thaumarchaea (Figure 14). These small archaea are ubiquitous in the wintertime waters surrounding Antarctica and facilitate the chemolithotrophic oxidation of ammonia (DeLong et al. 1994) while fixing carbon via the 3-Hydroxypropionate/4-Hydroxybutyrate pathway (Swan et al. 2014). Among the most abundant reads matching Thaumarchaea from Buchannan Bay were ammonia transporters and genes for propionyl-CoA carboxylase, highlighting the importance of chemolithotrophy in these organisms (Alonso-Saez et al. 2012). In the summer, Thaumarchaea are largely absent in Antarctic surface waters (Grzymski et al. 2012), except under ice shelves, where ammonia-oxidizing archaea can make up a substantial fraction of the microbial community (Vick-Majors et al. 2016; Christner et al. 2014). Thus the discovery of Thaumarchaea in Mertz Glacier ISW is consistent with high Thaumarchaeal populations which have

been discovered underneath the Ross Ice Shelf (Vick-Majors et al. 2016) and associated shelf outflow waters (Alonso-Sáez et al. 2011).

Marine Thaumarchaea are divided into 2 clades; a deep-water clade and a shallow clade, and this division is globally consistent across all major oceanic basins (Luo et al. 2014; Swan et al. 2014). The Thaumarchaeal metagenomic reads from the Mertz Glacier ISW outflow recruit almost exclusively to surface-isolated clades like *Cenarchaeaum symbiosum* and *Nitrosopumilus* spp. (Figure 15). As cold, dense ISW enters the polynya, the water mass subducts under the warmer Antarctic Surface Water, mixes with High Salinity Shelf Waters (HSSW) and fills the depths of the Mertz and Adèlie Basins with modified dense Shelf Water (Figure 16; G. D. Williams et al. 2010; Guy D. Williams et al. 2008). Four samples of this dense shelf water were taken at 325, 815, 1170 and 1210 meters, and all had a similar community signature (Figure 16) with most of the samples recruiting to *Nitrosopumilus* (Figure 15). Dense Shelf Water crosses out of the Mertz and Adèlie Basins at the Adèlie Sill, where it cascades down the continental shelf, entraining circumpolar deep water along the way to become AABW (Guy D. Williams et al. 2008). A sample taken at 56°25’S, 140°6’E, in the retroflected AABW current on the abyssal plain, revealed that the surface Thaumarchaeota signal was still present, though a large number of the reads now recruited to SCGC AAA007-O23, a deepwater single amplified genome first obtained from 770 m deep in the N. Atlantic Gyre (Luo et al. 2014). The large distance traveled by the surface archaea

suggesting that bacterial populations, like chemical tracers, can be advected considerable distances from their source.

2.5 CONCLUSIONS

The hypothesis that microorganisms can be entrained and advected via the movement of water masses is not new. In observations of water masses north of the Antarctic polar front, Wilkins et al. (2010) observed an advection effect on community structure in Antarctic Mode and Intermediate Water Masses. However, here we go a step further, identifying a continental source and a specific population (ammonia-oxidizing archaea) which is advected and distributed into the deep ocean. This potential source of diversity may have implications for the cycling of nitrogen through the deep ocean.

2.6 ACKNOWLEDGEMENTS

Chapter two has been prepared as a short communication to the *International Society for Microbial Ecology (ISME) Journal*. McQuaid, Jeffrey B.; McCrow, John P.; Hoffman, Jeffrey M.; Zheng, Hong; Espinoza, Joshua; Allen, Andrew E. "Metagenomic survey of East Antarctica reveals microbial advection into the deep ocean".

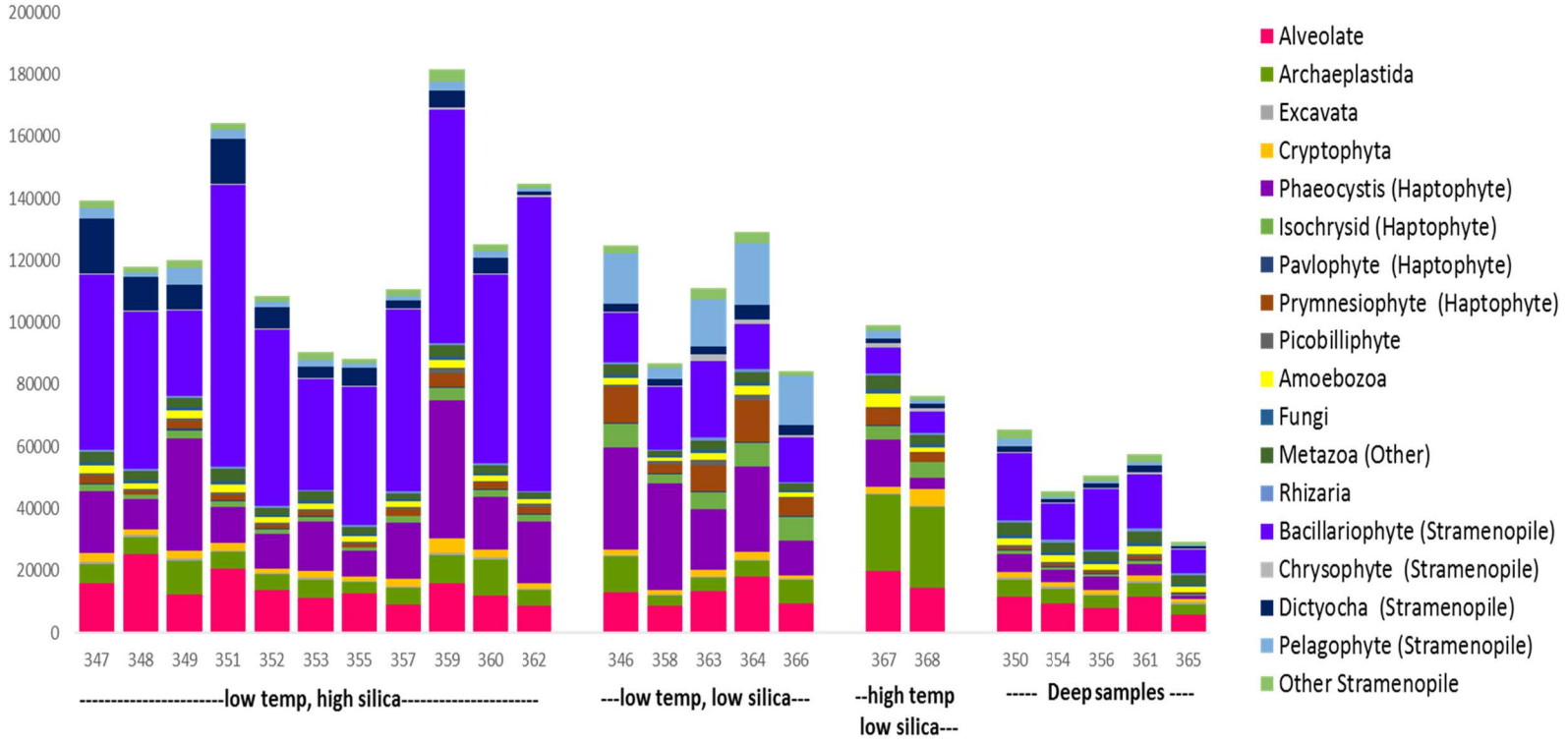

Figure 10: The taxonomid distribution of eukaryotic reads from surface and deep stations on the CEAMARC cruise. Data are clustered by silica content, water temperature, and depth.

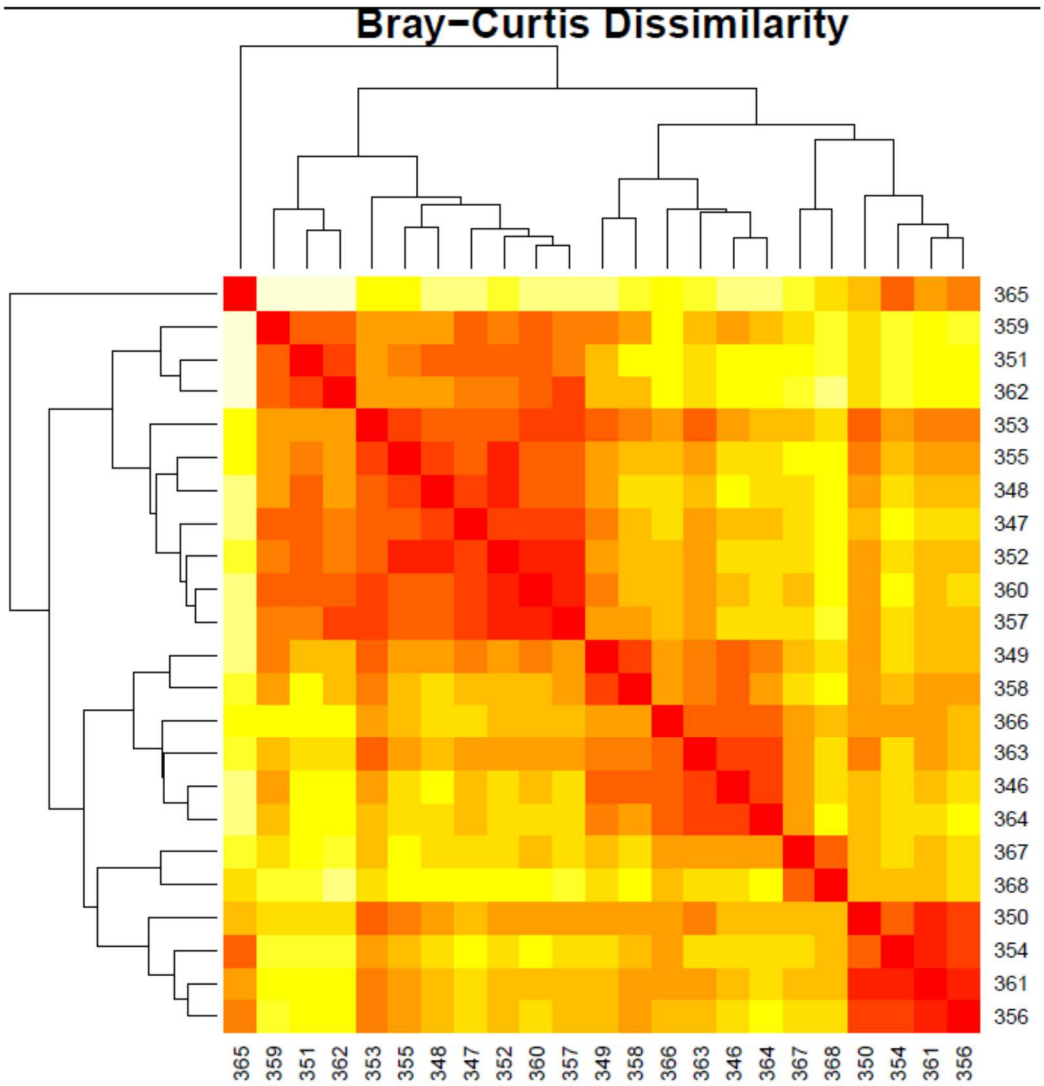

Figure 11: Bray-Curtis dissimilarity matrix of the eukaryotic fractions (3.0 and 0.8 um) for all 19 metagenomic stations

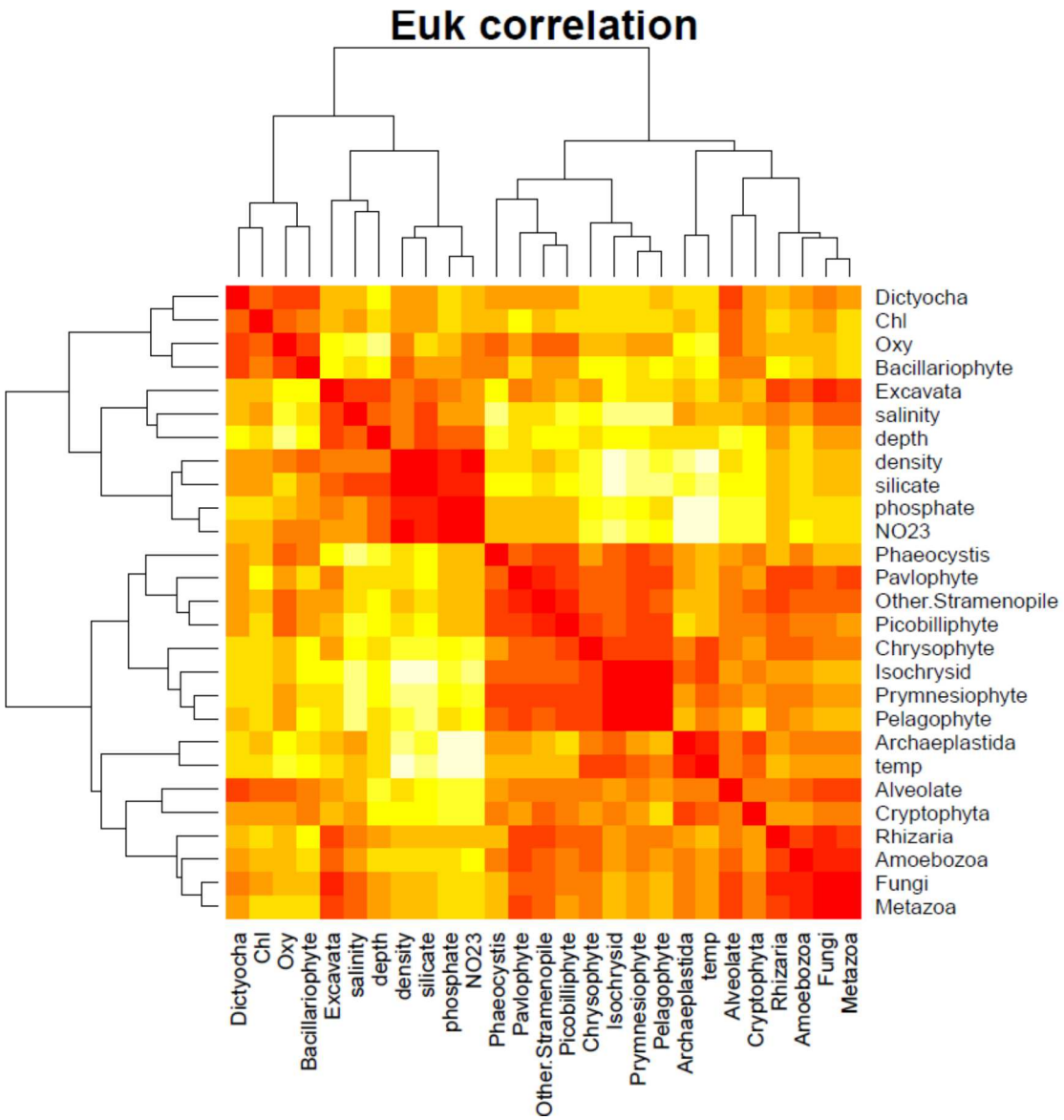

Figure 12: Bray-Curtis dissimilarity matrices of various eukaryotic phytoplankton groups (from fig. 10) and physiochemical parameters

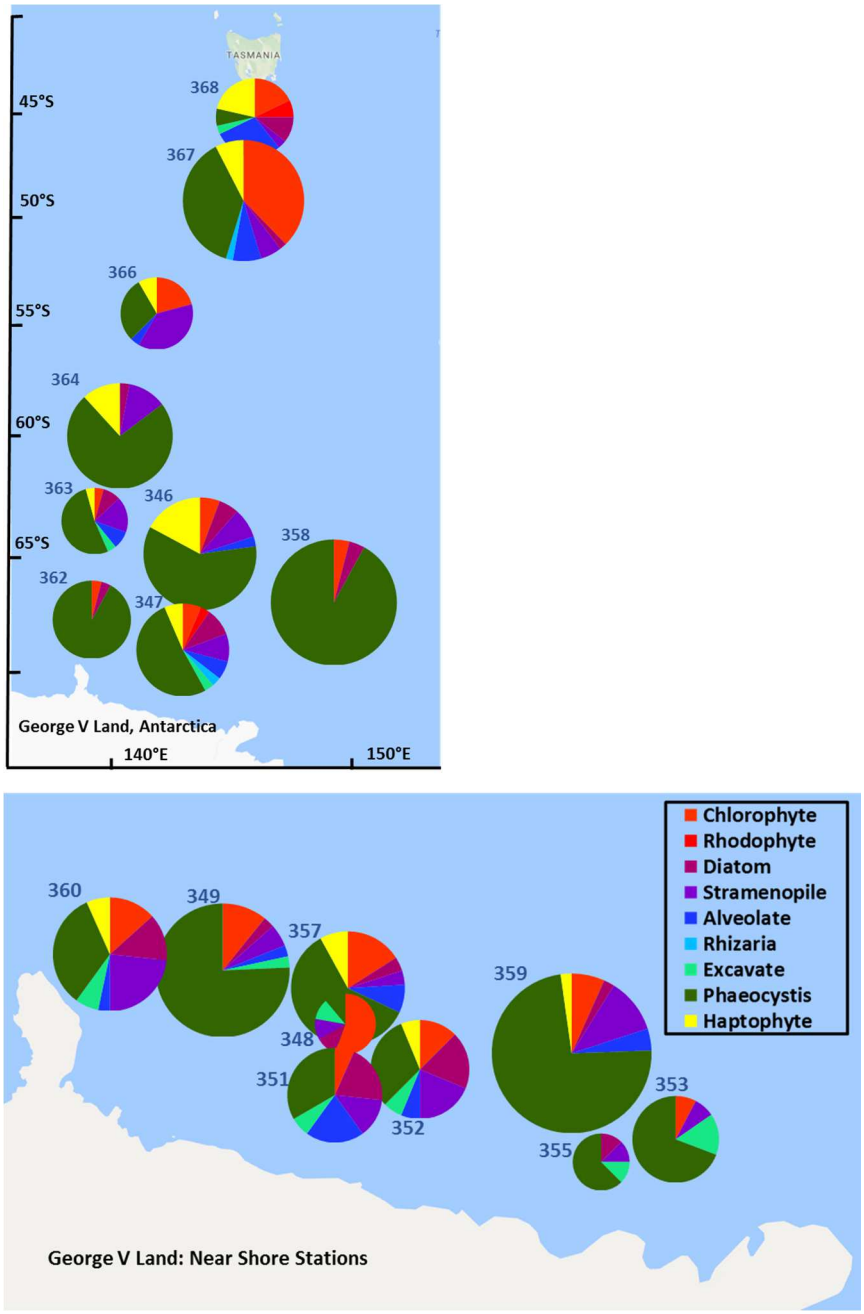

Figure 13: Taxonomic distribution of the iron binding ISIP2a across the Southern Ocean transect (top) and in the Mertz Polynya (bottom).

Figure 14: Library-normalized prokaryotic taxonomic diversity for Mertz Polynya surface samples (left) and deep samples (right), depth of sample in meters (top of column). Column in center is ice shelf water taken from Buchannan Bay adjacent to Mertz Glacier Tongue. At bottom is potential density referenced to the surface.

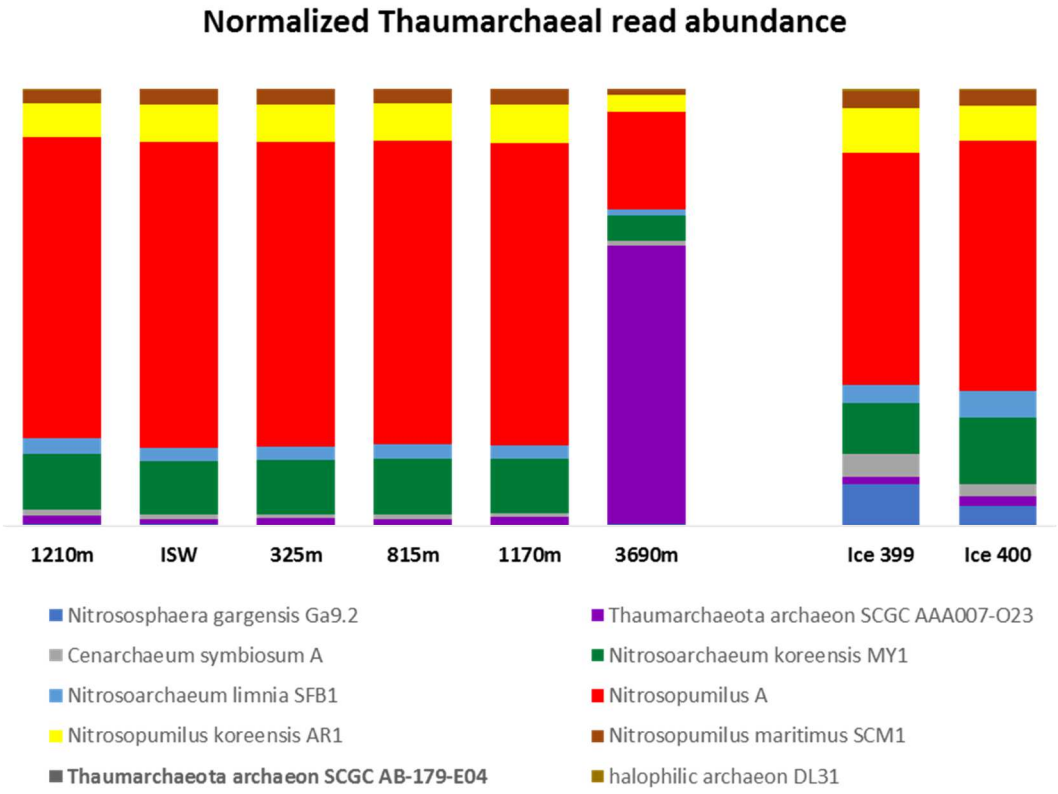

Figure 15: Recruitment of Thaumarchaeal reads from station 353 Ice Shelf Water (ISW) and the five deep-water stations: 350, 354, 356, 361 and 365. *Nitrospumilus* spp are shallow clades, while SCGC AAA007-O23 is a mesopelagic Thaumarchaea isolated from 770 m deep in the N. Atlantic gyre.

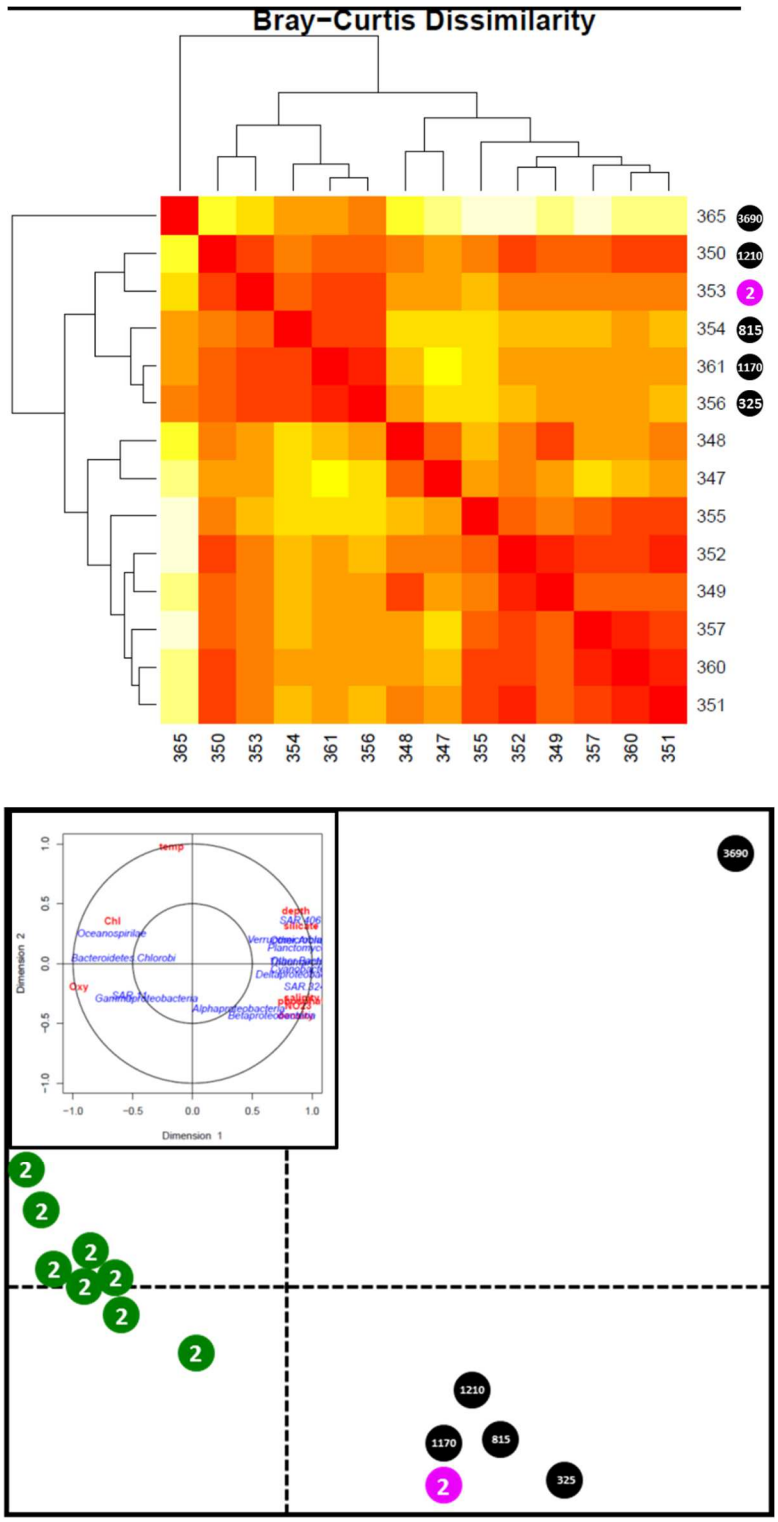

Figure 16: a) Bray-Curtis dissimilarity of prokaryotic samples from surface and bathypelagic samples b) correspondence analysis of surface samples (green), ice shelf water (pink) and bathypelagic samples (black)

Table 5: Environmental metadata for stations occupied during the CEAMARC cruise.

Station number	depth (m)	Chl a µg/l	Temp. (ITS-90)	PSS78	O_2 µmol kg^{-1}	Silicate µmol kg^{-1}	NO_3+NO_2 µmol kg^{-1}	PO_4 µmol kg^{-1}	Depth
346	2	0.25	2.924	33.762	339.1	9.6	25.2	1.5	4294
347	2	3.201	-0.6745	34.2444	380.78	54.4	25.2	1.43	443
348	2	12.611	-0.6399	34.3206	380.66	51.7	24.39	1.4	649
349	2	3.745	-1.2769	34.4381	347.69	69	27.11	1.81	365
350	330	0	-1.8558	34.5515	319.79	77.7	30.71	2.14	365
351	2	1.642	-0.6541	34.3874	357.49	61.1	25.78	1.63	597
352	2	1.037	-0.8353	34.3431	356.8	58.5	27.06	1.64	169
353	2	-0.225	-1.8154	34.4626	315.43	78.5	34.02	2.23	178
354	1320	0	-1.8808	34.7369	313.93	78.8	34.55	2.26	1209
355	2	8.431	-0.8874	34.3703	356.19	60.3	25.85	1.56	891
356	890	0	-1.8944	34.695	312.16	78.5	30.59	2.09	816
357	2	2.671	-0.4526	34.2101	351.28	52.1	27.48	1.67	533
358	2	0.267	-0.0349	33.588	343.46	38.1	27.19	1.82	3561
359	2	2.4	0.047	34.1126	351.92	52.8	26.3	1.58	364
360	2	7.503	-0.6601	34.1173	363.52	62.4	24.54	1.53	308
361	1170	0	-1.7753	34.5887	304.59	76.9	30.19	2.1	1144
362	2	0.191	0.8135	33.5919	354.81	39.5	26.8	1.8	1027
363	2	0.1	3.547	33.754	372.32	11	25.66	1.57	4473
364	2	0.5	4.16	33.713	330.05	1.89	10.47	1.29	3693
365	3690	0	0.4785	34.6932	216.03	135.3	32.91	2.3	3699
366	2	0.25	7.656	33.837	287.83	2.8	17.51	1.21	3180
367	2	0.2	10.94	34.43	280.59	0.24	6.28	0.65	3490
368	2	1.25	14.24	34.715	259.29	0.04	1.67	0.34	3201

2.7 REFERENCES

Alonso-Saez, L. et al. Role for urea in nitrification by polar marine Archaea. *Proc. Natl. Acad. Sci.* 109, 17989–17994 (2012).

Alonso-Sáez, L., Andersson, A., Heinrich, F. & Bertilsson, S. High archaeal diversity in Antarctic circumpolar deep waters. *Environ. Microbiol.* Rep. 3, 689–697 (2011).

Arrigo, K. R. Phytoplankton dynamics within 37 Antarctic coastal polynya systems*. J. Geophys. Res*. 108, 3271 (2003).

Christner, B. C. et al. A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* 512, 310–313 (2014).

DiTullio, G. R. et al. Rapid and early export of Phaeocystis antarctica blooms in the Ross Sea, Antarctica. *Nature* 404, 595–598 (2000).

Grzymski, J. J. et al. A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J.* 6, 1883–900 (2012).

Hosie, G. et al. CEAMARC, the Collaborative East Antarctic Marine Census for the Census of Antarctic Marine Life (IPY # 53): An overview*. Polar Sci*. 5, 75–87 (2011).

Keeling, P. J. et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol.* 12, e1001889 (2014).

Konstantinidis, K. T., Braff, J., Karl, D. M. & DeLong, E. F. Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific Subtropical Gyre. *Appl. Environ. Microbiol.* 75, 5345–5355 (2009).

Lacarra, M., Houssais, M. N., Sultan, E., Rintoul, S. R. & Herbaut, C. Summer hydrography on the shelf off Terre Adélie/George V Land based

on the ALBION and CEAMARC observations during the IPY. *Polar Sci*. 5, 88–103 (2011).

Luo, H. et al. Single-cell genomics shedding light on marine Thaumarchaeota diversification. *ISME J.* 8, 732–6 (2014).

Marchant, H., Davidson, a, Wright, S. & Glazebrook, J. The distribution and abundance of viruses in the Southern Ocean during spring. *Antarct. Sci.* 12, 414–417 (2000).

Morel, F. M. M., Kustka, a. B. & Shaked, Y. The role of unchelated Fe in the iron nutrition of phytoplankton. *Limnol. Oceanogr.* 53, 400–404 (2008).

Podell, S. et al. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. *Genome Biol*. 8, R16 (2007).

Rintoul, S. R. On the origin and influence of Adélie Land Bottom Water. *Ocean. Ice, Atmos. interactins Antarct. Cont. margin* 75, 151–171 (1998).

Shadwick, E. H., Tilbrook, B. & Williams, G. D. Carbonate chemistry in the Mertz Polynya (East Antarctica): Biological and physical modification of dense water outflows and the export of anthropogenic CO2. *J. Geophys. Res. Ocean.* 119, 1–14 (2014).

Shi, Y., Tyson, G. W., Eppley, J. M. & DeLong, E. F. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J.* 5, 999–1013 (2011).

Strutton, P. G., Gri, F. B., Waters, R. L., Wright, S. W. & Bindo, N. L. Primary productivity of the coast of East Antarctica (80° to 150°E): January to March 1996. 47, 2281–2298 (2000).

Swan, B. K. et al. Genomic and metabolic diversity of marine group i thaumarchaeota in the mesopelagic of two subtropical gyres. *PLoS One* 9, (2014).

Tolar, B. B., Ross, M. J., Wallsgrove, N. J. & Liu, Q. Contribution of ammonia oxidation to chemoautotrophy in Antarctic coastal waters. *ISME* 10, 1–15 (2016).

Talley, L. D. Closure of the global overturning circulation through the Indian, Pacific and Southern Oceans: schematics and transports*. J. Chem. Inf. Model.* 53, 1689–1699 (2013).

Vick-Majors, T. J. et al. Biogeochemistry and microbial diversity in the marine cavity beneath the McMurdo Ice Shelf, Antarctica. *Limnol. Oceanogr.* 61, 572–586 (2016).

Wilkins, D. et al. Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environ. Microbiol*. 15, 1318–33 (2013).

Williams, G. D., Bindoff, N. L., Marsland, S. J. & Rintoul, S. R. Formation and export of dense shelf water from the Adelie depression, East Antarctica. *J. Geophys. Res. Ocean*. 113, 1–12 (2008).

Williams, G. D. et al. Antarctic bottom water from the Adelie and George v Land Coast, East Antarctica (140-149??E). *J. Geophys. Res. Ocean*. 115, 1–29 (2010).

Wright, S. W. et al. Analysis of phytoplankton of the Australian sector of the Southern Ocean: Comparisons of microscopy and size frequency data with interpretations of pigment HPLC data using the 'CHEMTAX' matrix factorisation program. *Mar. Ecol. Prog. Ser.* 144, 285–298 (1996).

Yooseph, S. et al. A metagenomic framework for the study of airborne microbial communities. *PLoS One* 8, (2013).

Zeigler Allen, L. et al. Influence of nutrients and currents on the genomic composition of microbes across an upwelling mosaic. *ISME J*. 6, 1403–14 (2012).

CHAPTER 3: ANTARCTIC SEA ICE TRANSCRIPTOME REVEALS THE STRESSES OF PHOTOSYNTHESIS CONDUCTED IN A FROZEN MATRIX

3.1 ABSTRACT

The sea ice surrounding Antarctica is one of the largest seasonal habitats on earth, reaching a surface area maximum of 20 million square kilometers during the early spring. Algae, particularly diatoms, grow on the underside of the ice, and represent an early and critical component of the Antarctic food web. Diatoms grow on the undersides of sea ice for light-field stability and to escape predation, but the difficulties surrounding photosynthesis and carbon fixation while embedded in ice been difficult to ascertain due to microheterogeneity. Here we use environmental transcriptomics (meta-transcriptomics) to evaluate cellular- and community-level responses to the sea ice environment as compared to free-living organisms in the adjacent water column. Transcripts for a biophysical carbon concentrating mechanism (CCM) are strongly upregulated, as are transcripts encoding freeze tolerance, nutrient and metal acquisition, and protection from reactive oxygen species (ROS).

3.2 INTRODUCTION

The sea ice surrounding Antarctica is one of the largest biomes on earth, equal in size to tundra and desert habitats, and is dynamic, growing to a seasonal maximum of 20 million km^2, effectively doubling the area of the Antarctic continent. Sea ice is an important habitat for marine phytoplankton, particularly diatoms, and algae associated with sea ice contribute between 10 – 25 % of seasonal primary productivity of ice-

covered regions of the Southern Ocean (Arrigo et al. 1997). Primary productivity of algae embedded within or growing on sea ice can exceed 1g C m^{-2} day^{-1} (Arrigo et al. 1997), a number comparable to that of a productive ocean region. This high productivity is concentrated in a narrow layer at the bottom 20 cm of the ice column, and is temporally constrained as well: sea ice productivity reaches a maximum in the period between November and December, followed by a rapid decline as warming sea ice melts, releasing phytoplankton into the water column (Arrigo, 2004)

While ice algae can colonize a number sea ice microhabitats, including cracks, surface ponds and platelet ice, the algal communities which inhabit the bottom layers of the ice community grow to high densities (Thomas and Dieckmann, 2008). Colonizing and growing on the underside of sea ice has distinct advantages for algae: absent heavy snow cover, there is sufficient light for photosynthesis, providing light-field stability (Grossi, 1987; Arrigo, 2003). In the early spring, the ice structure of bottom ice is semi-porous and friable, allowing for diffusion of carbon and nutrients (Arrigo, 1999) yet still solid enough to allow protection from grazers (Thomas and Dieckmann, 2002). Tidal currents at the ice-water interface ensure enhanced hydrodynamic flow, both tangentially across the ice surface and vertically into the brine channels (Padman, Erofeeva, and Joughin 2003). In Antarctica, sea ice averages up to 2.0 m thick, and has sharp vertical gradients in temperature and

salinity (McMinn, 1999; Mock, 2003). Temperatures at the top of the ice column can drop below -25°C, consistent with air temperatures, and salinities can exceed 160 ppt in brine channels close to the ice surface. Sea ice algae growing at the bottom of the ice column experience salinities and temperatures closer to that of the adjacent water (Mock, 2003).

The advantages of light-field stability and protection from predation are often offset by the sheer density of organisms: ice algal biomass can 200 mg chlorophyll a m^{-2}, but this biomass is confined to a narrow band of a few centimeters (Figure 17). This high density of organisms in narrow channels can lead to a drawdown of dissolved inorganic carbon (DIC) and subsequent rise in pH. Obtaining representative pH measurements from concentrated brine-filled micro-channels is challenging, however a few representative pH values have been recorded ranging from 8 – 11, consistent with elevated rates of DIC drawdown (Gleitz et al. 1995; Søgaard et al. 2010). High pH values are indicative of low DIC, suggesting that sea ice algae are experiencing carbon limitation. A corollary effect of the drawdown of DIC from carbon fixation is the buildup of oxygen: microelectrodes implanted in the ice have revealed concentrations of O_2 exceeding 200% of saturation (Gleitz et al. 1995; Kühl et al. 2001; Papadimitriou et al. 2007). Sea ice biomass between November and December, these large standing stocks overwhelm the diffusion rates of carbon and nutrients, possibly impacting productivity and growth rates (Arrigo and Thomas 2004).

While physical and chemical measurements of bulk sea ice using microelectrodes and microcores have begun to reveal the extremes of the physiochemical sea ice environment, the effects of these extremes on the sea ice community are largely unknown. Here we use environmental genomics to describe the community structure of the bottom ice algal community and the cellular responses at a molecular level and compare those responses to those of free-living diatoms from the underlying water column.

3.3 METHODS

3.3.1 Sample collection

Environmental samples were collected over two seasons: in the early spring, when the bottom ice algal community is at maximal density (Jan 2009) and in the late spring (Nov 2009) when the sea ice is thinning and the bottom ice community has been released into the water column due to melting. In the early spring, three stations were occupied in a transect extending west from Cape Evans (Ross Island, McMurdo Sound): 399 (166.424°E, 77.650°S), 400 (166.168°E, 77.659°S), and 401 (165.750°E, 77.673°S). Boreholes were drilled at each site using a 9 cm Kovacs corer using a clean stainless steel drill head with a polyurethane body. The sea ice at all three sites was approximately 1.6 m thick, and snow cover was 0.3 m thick at the eastern end of the transect, decreasing to 0.05 m at the westward end. The bottom ice community was sampled by removing multiple cores, placing the cores into a lightproof container

and scraping the friable ice layer which contained the ice algae into clean sterile 50 ml falcon tubes. Tubes were quickly frozen in liquid nitrogen for later processing. To sample the under-ice planktonic environment, a line measure was used to place the tubing intake at 3m, or 1.5 m below the bottom of the ice. For these samples, 200 L of seawater was peristaltically pumped through a 200 µm mesh pre-filter and then serially passed across three 293 mm PES filters, resulting in the following size fractions: 200 – 3.0 µm, 3.0– 0.8 µm and 0.8 - 0.1 µm. Filters were rapidly transferred to storage buffer (RNALater, Thermo) and frozen at –80°C for later processing. For meta-transcriptomics, 20 L of 200 µm pre-filter seawater was peristaltically pumped through 0.22 µm Sterivex filters (Millipore), dewatered, capped and immediately frozen in liquid nitrogen for later analysis. A second season of late spring samples were collected Jan 20 – 24, 2009, in the McMurdo Sound (165.4744°E, 77.6166°S). Ice cores removed at these sites revealed there was no significant bottom algae community, thus only water column samples were taken.

3.3.2 16S/18S ribosomal RNA amplification and sequencing

Nucleic acids were extracted from all three size class filters using the Global Ocean Sampling protocol as previously described (Rusch, 2007). For the sea ice samples, 5 – 10 g subsamples of sea ice algae were slowly melted into an equal volume of RNALater, and the nucleic acids were extracted with PlantZol (Thermo) following the manufacturer's instructions (Thermo). Extracted nucleic acids were subsampled for the

construction of 16S/18S rRNA libraries: the V3 – V5 region of the 16S rRNA was amplified using eubacterial primers 341F (5'-CCTACGGGNG-GCWGCAG-3') (Lane et al. 1986) and 926R (5'-CCGTCAATTCMTTTRAGT-3') (Herlemann et al. 2011), producing amplicons approximately 500 bp long. Similarly, the V4 region of the eukaryotic 18S rRNA was amplified with primers TAReuk454FWD1 (5'-CCAGCASCYGCGGTAATTCC-3') and TAReukREV3 (5'-ACTTTCGTTCTTGATYRA-3') (Stoeck et al. 2010). Adapter sequences and 10bp multiplex identifier (MID) barcodes were appended to both sets of primers for 454 sequencing and deconvolution of the datasets. PCR conditions were 95°C for 2 minutes, 30 cycles of 95°C for 20 seconds, 56°C for 30 seconds, 72°C for 5 minutes. PCR reactions were cleaned with AMpure XP beads (Beckman) and visualized on agarose gels prior to quantification using the PicoGreen QuantIt assay (Thermo). 20 ng of each 16S and 18S amplicons were then pooled. The number of molecules needed for emulsion PCR (emPCR) was estimated by using the KAPA Biosystems Library Quantification Kit (Kapa Biosystems), a Biomek FX (using butanol) was used to automate the emulsion-disruption process, and the Robotic Enrichment Module (Roche) was used for subsequent bead enrichment steps. For each filter set (and sea ice sample) approximately $1 – 2 \times 10^5$ 16S and 18S sequences were generated for analysis.

3.3.3 Metatranscriptome preparation

Filter layers were removed from the Sterivex capsules with a sterile scalpel and placed directly into Trizol reagent (Thermo) and extracted

following the manufacturer's instructions. Approximately 300 – 1200 ng RNA were recovered from each sterivex filter. RNA was prepped with the RNAeasy MinElute kit (Qiagen). Two types of metatranscriptomes were prepared: Eukaryotic-specific Poly-A transcritpomes and TotalRNA trancriptomes. Poly-A libraries were made by amplifying 100 ng of DNAse-treated community RNA with MessageAmp II rRNA Amplification kit (Thermo) via two rounds of in-vitro transcription at 37C for 14h with T7 Oligo(dT) primers. These amplified constructs were then converted to double-stranded cDNA with SuperScript III first- and second-strand synthesis kits (Thermo). The resulting cDNA was purified on agarose gels and the 0.3 – 3.0 kb size range was extracted using a QIAquick Gel extraction kit and purified using AMpureXP magnetic beads

For TotalRNA libraries, prepped RNA (from above) was depleted of ribosomal RNA with the Ribo-Zero magnetic kit (Illumina) using a 2:1:1 mixture of plant, animal and bacterial RNAs to remove rRNA. This rRNA depleted mixture was purified using an Agencourt RNAClean XP kit, and 2 ng of this mixture was used as template for cDNA synthesis. The Ovation RNA-seq System V2 (NuGEN) was used, which applies both random hexamers and polyA primers. One microgram of this cDNA pool was then fragmented to a 200 bp in size, and Illumina libraries were then constructed with the TruSeq RNA sample prep kit (V2) with an end-repair step. These libraries were paired-end sequenced on an Illumina HiSeq, generating 5 – 10 Gb of sequence per library.

3.3.4 Meta-transcritpome analysis

Output from the HiSeq was trimmed and filtered to remove primers, adaptors and rRNA sequences using Ribopicker v. 0.4.3 (Schmieder, Lim, and Edwards 2012) and CLC Assembly Cell (CLC bio) was then used to assemble the reads into contigs. FragGeneScan was used to predict the Open Reading Frames (ORF), and these ORFs were annotated *de-novo* for function by comparing against TigrFam, Pfam, KEGG, and KO assignments. A Lineage Probability Index (LPI) was used to assign a taxonomic classification to that ORF (Podell et al. 2007) and prevent incorrect taxonomic assignment of conserved sequences as well as bacterial/archaeal contaminants in the reference dataset. For groups with a high LPI for e.g. diatoms, edgeR (Robinson, McCarthy, and Smyth 2009) was used to normalize libraries and assign fold change values to determine differential expression. A Benjamini-Hochberg ($p < 0.05$) test was used in comparisons between the diatoms living in the sea ice vs. the water column. These phylogenetic groupings were group-normalized as reads per kilobase per million diatom reads mapped (RPKM) for comparisons between the different samples (early spring ice vs. water, and early vs late spring water communities).

3.3.5 Microscopy

Light microscopy images were generated by gentle melting filaments of sea ice diatoms into filtered Ross sea water and imaging on a Zeiss Axioscop under differential interference contrast (DIC). Scanning

electron micrographs were made of the sea ice community by critical-point drying frozen samples of the ice community, sputter-coating the samples with chromium on an Emitech K575X Sputter Coater and imaging on a FEI Scios DualBeam FIB/SEM.

3.3.6 Thermal hysteresis and ice recrystallization

Ice binding proteins were isolated from five hundred milligrams of bottom ice community by thawing samples into 0.2 µm filtered Ross sea water followed by five rounds of freeze-fracturing on liquid nitrogen followed by thawing in an ice bath. Cellular debris was clarified by centrifugation for 5 minutes at 14,000 x g. Non-equilibrium freezing and melting points were determined on a Clifton nanoliter freezing-point milli-osmometer (Clifton Technical Physics) which was modified for computer control of the temperature ramp rates, resulting in ultrafine temperature resolution (0.002 °C). Freezing and melting of the ice crystals was visualized on an inverted microscope at 200x magnification, and ice crystals were formed by freezing sub-microliter droplets of the clarified lysate and seawater controls in immersion oil (at -20°C) followed by warming to melt all of the crystals. To determine freezing/melting hysteresis, rapid warming and cooling was used around the freezing point to overcome hysteresis. To determine the hysteresis freezing point, the temperature was lowered at ~0.1°C per minute until rapid growth of the ice crystal was observed. To determine the melting hysteresis, the sample was warmed at 0.03 °C per minute until the last ice crystal disappeared.

For the ice recrystallization assay, 10 nL of either lysate or seawater control was frozen at -20°C and ice recrystallization was monitored over the course of 60 minutes.

3.4 RESULTS AND DISCUSSION

3.4.1 Phylogenetic distribution of organisms in the sea ice environment

Three separate paired samples were taken of the sea ice community and the underlying water column in the early spring (November), and three water column samples were taken in the early summer (January), when the bottom ice community had detached from the sea ice underside (Figure 17). We determined the distribution of organisms in ice and water communities by extracting community DNA, amplifying the V3/V4 variable domain of the 16S/18S rRNA gene, and comparing an in-house database, PhyloDB 2.0. We mapped distribution and diversity of microplanktonic organisms (<200 μm) in sea ice and in the underlying water column (Figure 18, 19). Sea ice biomass and productivity peaks between the months of November and December, particularly for land-fast bottom ice communities (Arrigo 1999). Across all three ice cores the sea ice was dominated by pennate diatoms, with >70% of the sea ice OTUs matching 18S rRNA sequences from four pennate diatom genera: Ampiphora, Fragilariopsis, Thalassionema and Nitzschia (Figure 18). These genera are Non-pennate diatoms consisted of less than 5% of the OTUs, and included several unarmored dinoflagellates and OTUs corresponding to the centric diatom Porosira. This near total dominance of the sea ice by pennate

diatoms has also been observed in sea ice communities from the Ross (Sedwick and DiTullio 1997) Weddell (Ackley, Buck, and Taguchi 1979) and East Antarctica regions (McMinn et al. 2007). Prokaryotic rRNA was less than 5% of the rRNA recovered from the TotalRNA libraries, underscoring that early spring bottom ice communities are highly simplified, eukaryotic-dominated systems where prokaryotes play a minor role (Figure 20). Light and scanning electron microscopy confirmed the dominance of pennate diatoms and the lack attached or free-living prokaryotes (Figure 17). The OTUs present in the ice were related to Bacteriodetes marine bacteria known for attaching to surfaces and other organisms and degrading high molecular weight compounds (Fernandez-Gomez et al. 2013); however few of these bacteria were observed in scanning electron micrographs (Figure 17). A companion study on the under ice dissolved and particulate metal concentrations determined the sea ice biomass in the samples to vary between 2.4 - 9.5 mg L^{-1}, which is over a thousand times more concentrated than the underlying water column, where biomass was generally < 1 µg L^{-1}(Noble et al. 2013).

In contrast to the autotrophic diatom-dominated sea ice community, the organisms inhabiting the water column underneath the sea ice were dominated by heterotrophic alveolates, with diatoms and other autotrophic organisms comprising less than 10% of OTUs (Figures 18, 20). Alveolate OTUs were primarily associated with unarmored heterotropic Gymnodinid dinoflagellates, parasitic/symbiotic dinoflagellates (e.g. Amoebophyra and

Syndinialid parasites of free-living dinoflagellates) and ciliophores. Non alveolate plankton included radiolarian protozoa and picobiliphytes, both of which are non-obligate autotrophs. 16S amplification of the bacterial fraction revealed an abundance of bacteria involved in the oxidation of reduced nitrogen and sulfur compounds, including many bacterial lineages associated with winter Antarctica conditions. These include OTUs corresponding to δ- and β-proteobacteria (Nitrospina and Nitrosospira), both of which oxidize nitrite to nitrate, and sulfite oxidizers of the SUP05 lineage. The primers used for the construction of the 16S libraries are eubacterial-specific and do not amplify archaea; however abundant archaeal sequences encoding AmoA, nitrite reductase and ammonium transporter were found in the below-ice water column, as described below.

In early summer (late January), fast ice still covers much of McMurdo Sound, however the porous ice which supports the bottom ice community has melted, and the bottom ice community has been released into the water column. 18S analysis did not find any of these organisms to be abundant in the late spring water column. Of sea ice diatoms, only members of the cosmopolitan Fragilariopsis were present. Many of the other diatoms were centric species (Chaetoceros) which had no representation in sea ice. species were present. Ice core sampling revealed no bottom community, and 18S revealed a change in the under ice community. Aside from Fragilariopsis spp, few of the sea ice diatoms were present in the water column. Mamiellate prasinophytes, absent in

the early spring, were in abundance, as were non-diatom like members of the MAST groups 7 and 8 (de Vargas et al. 2015).

3.4.2 Functional genomics: antifreeze prevention

The photosynthetic strategies and cellular metabolism of cells embedded in ice likely differ from those of free-living planktonic orgamisms. While the advantages of light-field stability and protection from grazers are apparent, decades of research also have revealed the disadvantages related to life in ice, including ice crystal-induced cellular damage, localized depletion of CO_2 and nutrients, and elevated concentrations of brine salts, oxygen and pH (for a review see Arrigo and Thomas, 2004). We elected to look at the algal community response to these challenges by examining differential expression patterns of diatom-specific transcripts from the sea ice and comparing them to the smaller community of free-living diatoms in the water column. Both Total RNA (which examines RNA from all organisms) and eukaryote-specific PolyA-RNA transcriptional libraries were used to cross-validate each other.

The most abundant transcripts in the diatom sea ice community were for ice binding proteins (IBP), accounting for 4% of total diatom transcripts (Figures 21, 22). Ice binding proteins function differently from traditional antifreeze solutions in that IBPs adhere to and modify the surface of ice, inhibiting crystal growth and inducing thermal hysteresis between the melting and freezing points (Davies 2014). In addition, IPBs

alter the freezing point in a manner which is not directly proportional to their concentration, and this non-colligative effect allows ice binding proteins to function at concentrations far lower than antifreeze compounds. Thus IBPs do not affect cellular osmolarity (Bayer-Giraldi et al. 2010). In the sea ice diatom transcriptome, ice binding proteins have a number of architectures, the most common being a single domain protein with a signal peptide, suggesting a secretory pathway. There are also multiple domain IBPs and IPB-fusion proteins. Particularly common are 'ice adhesins', IBPs used by diatoms and other organisms to adhere to the ice layer (Bayer-Giraldi et al. 2010). Ice adhesions are large multimeric proteins composed of an ice binding domain, followed by 2 – 15 bacterial immunoglobin-like domains and a transmembrane anchor. The immunoglobin domains form a tether, allowing the organism to adhere to the ice surface without direct membrane contact. The sea ice algal communities in McMurdo Sound are mostly raphid pennates, which suggests that gliding motility and use of ice adhesins assist in the colonization of brine channels in the ice underside. Light microscopy of filaments of these diatoms growing in a brine channel reveal both dense packing of the diatoms in the channel, along with a heavy exopolysaccharide layer (EPS: Figure 17).

As thermal hysteresis is a characteristic for the activity of ice binding proteins, we examined the ability of lysate from the sea ice community to alter the freezing properties of seawater. Ice binding

proteins depress the freezing temperature by bending the plane of ice formation, which reduces the favorability of crystalline lattice formation by raising the melt temperature, the melting and reforming of ice crystals is inhibited, thus thermal hysteresis prevents cell-damaging crystal growth (Davies 2014). Using a nanoliter milliosmometer attached to a cold-stage inverted microscope, we determined clarified sea ice algal lysates exerted a differential between the melting (-0.74C) and freezing (-0.91C) point of seawater, indicative of the activity of ice binding proteins rather than osmotic solutes or antifreeze compounds (Cziko et al. 2014). This activity was not seen in seawater and brine-channel controls. In an ice recrystallization assay, ice algal extracts prevented the recrystallization and growth of ice crystals (Figure 23).

3.4.3 Carbon Concentrating mechanisms

The high density of actively photosynthesizing algae in sea results in localized drawdown of dissolved inorganic carbon (DIC) throughout the brine channels and pore space of the bottom sea ice (Thomas and Dieckmann 2002). The drawdown of DIC can markedly elevate the pH compared to the underlying seawater (Mock and Thomas 2005). Brine channel pH measurements range from pH 8.0 up to pH 11 (Gleitz et al. 1995; Delille et al. 2007; Papadimitriou et al. 2007), underscoring the extent to which high standing stocks of bottom ice algae can induce localized depletion of inorganic carbon. At the three sea ice sampling locations, the bottom ice algae community ranged from 2.1 – 9.5 mg C^{-1},

representing a 1000 fold increase in biomass as compared to the water column, which rarely exceeded 1 μg C L^{-1} (Noble et al. 2013). Compounding the high standing stocks is the apparent lack of heterotrophic bacteria to re-mineralize organic matter back to CO_2 and complete the 'microbial loop'. Only 10% of the reads in bottom ice Total RNA libraries were bacterial in origin, and this lack of bacteria was also observed in light and scanning electron micrographs (Figure 17). High biomass and carbon drawdown combined with poor diffusivity likely result in sea ice algal communities being an extreme environment for carbon limitation.

In order to saturate the carbon-fixing enzyme RuBisCO with sufficient carbon dioxide, pennate diatoms leverage their complex internal membrane system as a biophysical carbon concentrating mechanism (CCM), components of which have only recently been described (Matsuda et. al, in press). In the pennate diatom *Phaoedactylum tricornutum*, the biophysical CCM works by actively pumping bicarbonate ions through the outer membrane (Nakajima, Tanaka, and Matsuda 2013) and across all four internal plastic membranes into the plastid lumen. Inside the lumen, a θ-carbonic anhydrase (Kikutani et al. 2016) converts the bicarbonate to CO_2, saturating the thylakoid-penetrating pyrenoid (Matsuda, in press). The diffusion of CO_2 away from the pyrenoid and across the chloroplast membranes is reduced by Zn-containing β-carbonic anhydrases which

have been localized in the periplastid space and the CER lumen (Satoh et al. 2001; Tachibana et al. 2011).

In the sea ice transcriptome, transcripts for outer membrane sodium-dependent bicarbonate transporters are among the most abundant in the mRNA pool, second only in abundance to ice- and chlorophyll-binding proteins (Fig. 22). In *Phaeodactylum*, transcripts for bicarbonate transporters are rarely abundant, only becoming significant when the CO_2 (aq.) concentration falls below 50 ppm (Dupont, in prep). Therefore, we consider the high copy numbers encoding these proteins to be indicators of carbon-limiting conditions. The expression of bicarbonate transporters in sea ice is 4 to 5-fold higher than in the underlying water column (Figure 23). Similarly regulated at a transcriptional level are β-carbonic anhydrases, and transcripts for these are among the top 100 transcripts, 4 fold more abundant in sea ice diatoms than free-living diatoms (Fig. 22).

As carbon is likely at low concentration in the interstitial channels of sea ice, we examined transcripts encoding abundances of proposed C4 carbon concentrating mechanisms (Kroth et al. 2008). C4 models rely on the plastic-localized pyruvate phosphate dikinase (PPDK) to convert pyruvate to phospho*enol*pyruvate (PEP) for export and carboxylation in the CER (Kroth et al. 2008). However, transcripts for PPDK were in low abundance and not significantly differentially expressed (<1-fold) in ice vs the water column. The lack of differential expression was consistent

similar reverse genetics approaches which suggest that carbon concentration is accomplished vie aa biophysical rather than biochemical concentrating mechanism (Haimovich-Dayan et al. 2013).

3.4.4 Oxygen supersaturation

The combination of an active carbon concentrating mechanism and photosynthesis in a confined space leads not only to the drawdown of inorganic carbon, but to the buildup of high concentrations of oxygen. While some phytoplankton blooms can lead to localized supersaturation of oxygen, the highest concentrations of oxygen have been recorded in sea ice, exceeding 200% of the underlying seawater (Gleitz et al. 1995; Kühl et al. 2001). This is likely exacerbated by the continual irradiance, which does not allow for nighttime diffusion of the accumulated oxygen, as well as the lack of a heterotrophic community which can actively consume oxygen as it is produced. While elevated concentrations of oxygen are not toxic, high oxygen can lead to the accumulation of reactive oxygen species (ROS), including superoxide, hydrogen peroxide, and hydroxyl radicals (Raven 1991). Superoxide dismutases are metalloenzymes used to disproportion superoxide radicals, producing molecular oxygen or hydrogen peroxide, and transcripts encoding all protein representing all three SOD families (Fe/Mn, Cu/Zn and Ni SODs) are present in the sea ice transcriptome. However, only Fe/Mn SODs are abundant (in the top 25% of transcripts), and none of the SOD enzymes are differentially expressed when compared to seawater (Fig. 22), suggesting that metal limitation

may be impacting reliance on these enzymes (see below). In contrast, enzymes which do not require a metal cofactor such as catalases and peroxidases exhibit differential expression. Peroxiredoxin and thioredoxin (which restores activity to peroxiredoxin) are also abundant and exhibit 3-fold differential expression when compared to water column diatoms. Pennate diatoms under iron limitation are known to increase tocopherol production, but enzymes of this biosynthetic pathway did not show differential expression.

Conditions of high oxygen/low carbon dioxide are canonical conditions for increased oxygenase activity of RuBisCO and photorespiration. The complete photorespiration pathway in diatoms has not been fully elucidated, with a potential pathway routed through the peroxisome (the glyoxylate cycle) as well as a proposed mitochondrial pathway (the C2 pathway) (Kroth et al. 2008; Davis et al. 2016). In the sea ice transcriptome, many of the mitochondrial enzymes are upregulated, including glycine decarboxylase (P-protein, 2.6-fold DE), alanine glutamine amino transferase (3 fold DE) and pyruvate decarboxylase (2.5 fold DE). In addition, the peroxisomal isoform of glycolate oxidase was not abundant nor differentially expressed in the ice algal community, whereas the mitochondrial isoform (GOX2) was 2.1-fold differentially expressed, further suggesting a C2-type mitochondrial pathway in operation. However, as most of these enzymes are also part

of intersecting metabolic pathways, caution will need to interpret the dataset until a more complete pathway for diatoms has been validated.

3.4.5 Acquisition of nutrients and micronutrients

As the diffusion of carbon and oxygen appear to be affected by both the high standing stocks and physical constraints imposed by the ice channel network, we hypothesized that other nutrients and micronutrients might be similarly affected. We compared ice and water column transcriptomes to TransportDB 2.0 (Elbourne et al. 2016) to determine the types of transporters are differentially expressed in the sea ice community, manually curating those transporters where available. Among the most abundant transcripts in the ice algal community, second only to the outer membrane Na^+/HCO_3^- transporters in abundance, are transcripts for the ferric iron-binding protein ISIP2a (Figure 22). Iron Stress Induced Protein 2a, or ISIP2a, is described extensively in this dissertation in Chapter 1, and is the second most abundant transporter in the sea ice transcriptome. Transcripts for ISIP2a are 2.5-fold more abundant among diatoms in the sea ice than water column. ISIP2a is a outer membrane phytotransferrin which uses the synergistic anion carbonate to bind ferric iron, therefore the high abundance and differential expression of this phytotransferrin could be a result of iron limitation, carbonate limitation, or both. Recent studies by Sedwick et al (2011) and Bertrand et al (2011) have noted that early spring iron limitation in the Ross Sea. While the pH of the brine channels is likely alkaline, favoring the carbonate ion species, the high

rates of drawdown can deplete the carbonate ion and affect the ability of the algae to obtain iron via ISIP2a. Indeed, when *Phaeodactylum tricornutum* is grown under iron replete conditions but low carbon (50 ppm CO_2), ISIP2a is among the most abundant transcripts, indicating that carbonate chemistry directly influences iron bioavailability (Dupont, in prep).

The high abundance and differential expression of ISIP2a is indicative of iron stress, and this stress is manifest via alternative metal acquisition and metal substitution strategies. Highly abundant in the transcriptome is another iron stress protein of unknown function, ISIP-1, as well as the iron substitution protein flavodoxin, both of which are in the top 100 most expressed transcripts. In the diatom *Thalassiosira oceanica,* the copper containing electron transport protein plastocyanin can substitute for the iron containing cytochrome b6 (Peers and Price 2006), and while these transcripts are not abundant, they are highly differentially expressed (3.6-fold DE). Elevated copper requirements are evident in the differential expression of a putative copper-transporting haloacid P-type ATPase (Figure 22). Additional differentially expressed transcripts include Ferric Reductases, an FTR iron permease, a Ni/Co transporter and a cobalamin transporter, the latter thought to be a co-limiting micronutrient in the early spring Ross Sea (Bertrand et al. 2015). A putative Zip-Zinc outer membrane permease is highly abundant, and

may be transporting ferrous iron or zinc, with the latter required for Zn-containing carbonic anhydrases.

Transporters for nitrate, nitrite and urea are both abundant and differentially expressed in the sea ice (Figure 22). Of the major diatom nitrogen transporters, only the ammonia transporter was more highly expressed in the water column (Figure 22). Due to the high pH of the brine channels (from DIC drawdown), most ammonium ion is present as the uncharged ammonia species, and thus is not available for internalization via ammonium ion transporters. Phosphate and silicon transporters were also elevated in the sea ice transcriptome, with a clear preference in the sea ice for sodium-dependent phosphate transporters of the SLC34 family (Figure 22).

3.5 CONCLUSION

For planktonic algae in the early spring of the Ross Sea, sea ice provides an important refuge from predation as well as a stable light field for optimizing photosynthesis and light harvesting antennae. However, the results here reveal the downsides to life in the ice channels: reduced rates of diffusion impact the ability to obtain sufficient carbon, nutrients and metals, thus cells living in the sea ice must devote a large portion of their transcriptional and cellular energy to the concentration of carbon and the production of transporters. The buildup of oxygen creates hazards through slower dissipation of reactive oxygen species, and this also functions as metabolic drag. The impact of high oxygen and low carbon on

photorespiratory pathways is not known. Despite all these impediments to growth, in the early spring in the Ross Sea, the bottom ice algae community is highly productive, vastly exceeding the biomass of the underlying water column.

3.6 ACKNOWLEDGEMENTS

Chapter three has been prepared as a submission to the *International Society for Microbial Ecology (ISME) Journal*. McQuaid, Jeffrey B.; McCrow, John P.; Hoffman, Jeffrey M.; Zheng, Hong; Cziko, Paul A.; Valas, Ruben E.; Allen, Andrew E. "Sea Ice Transcriptome reveals unique stresses of conducting photosynthesis in a semisolid matrix."

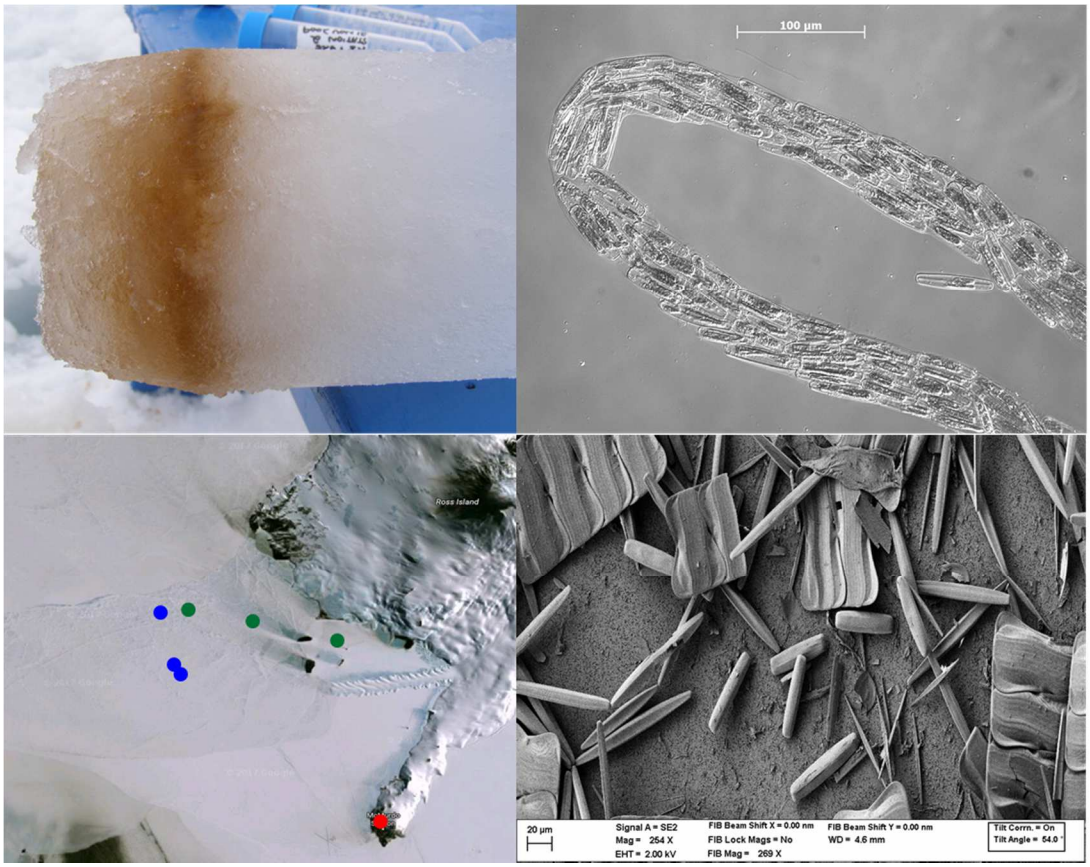

Figure 17: Bottom ice algae community and sample locations in McMurdo Sound. Clockwise from upper left: A representative core from station 400 showing the high density of the bottom ice algal community. Pennate diatoms as they grow in a brine channel. SEM of mixed Fragilariopsis, Nitzschia and Amphiprora spp., note general lack of prokaryotes attached or retained on 0.2 µm filter. Sample locations, early spring sites are in green, late spring sites are in blue, McMurdo Station is in red.

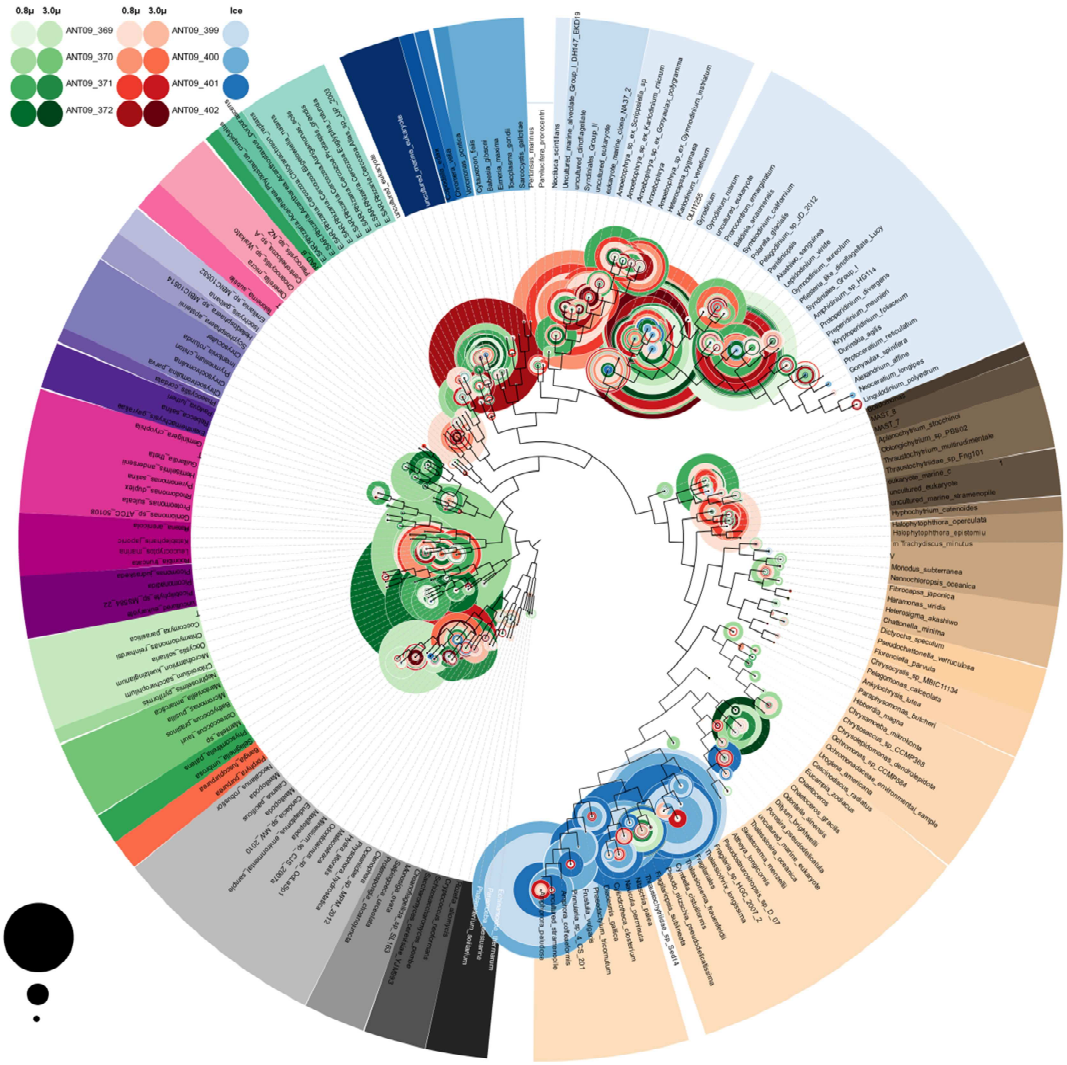

Figure 18: Phylogenetic mapping of the 18S rRNA from the sea ice (blue circles) and underlying water column (red circles), in contrast to the water column community of the last spring (green circles).

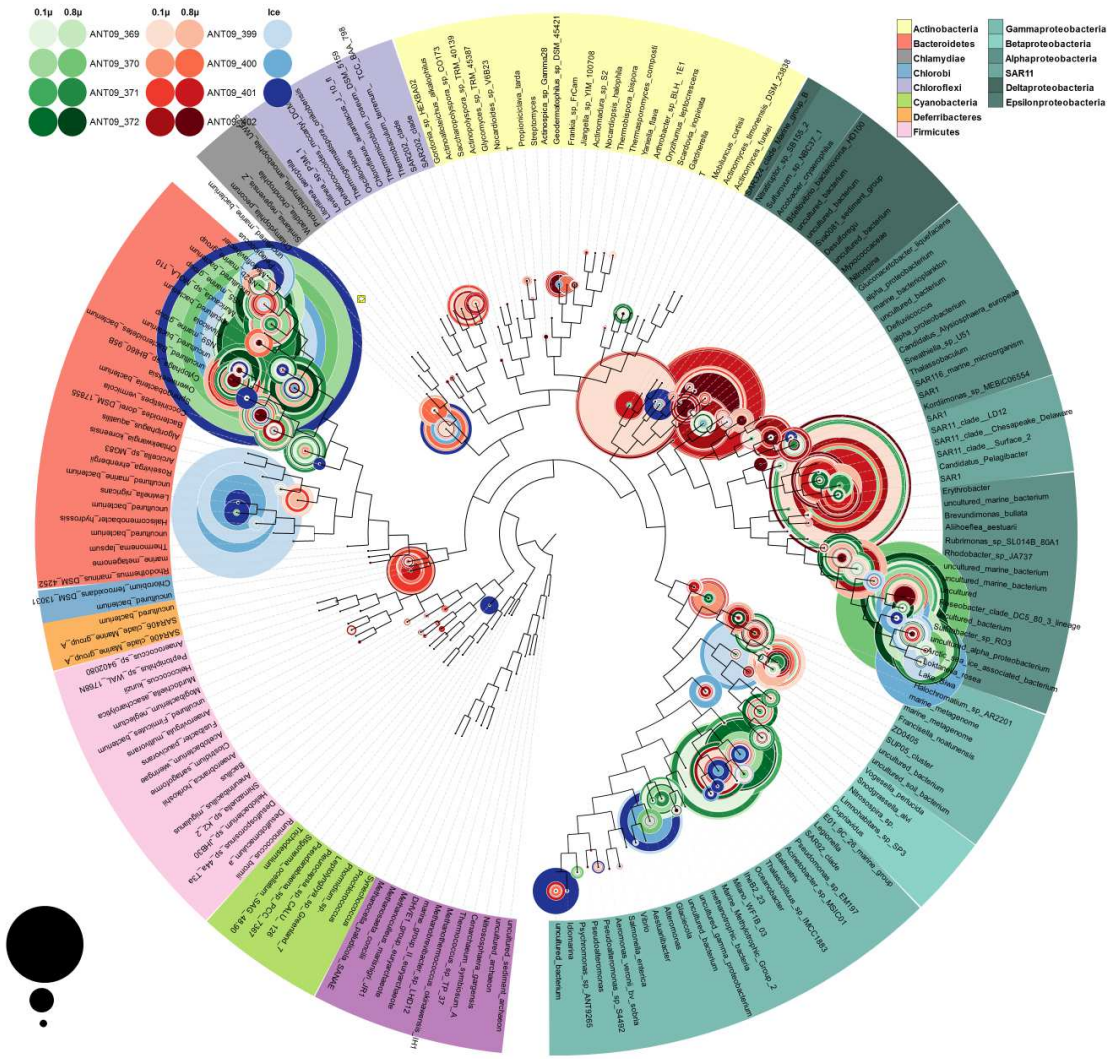

Figure 19: Distribution of prokaryotic organisms based on amplification of 16S rRNA from the sea ice (blue circles) and underlying water column (red circles), in contrast to the water column community of the last spring (green circles).

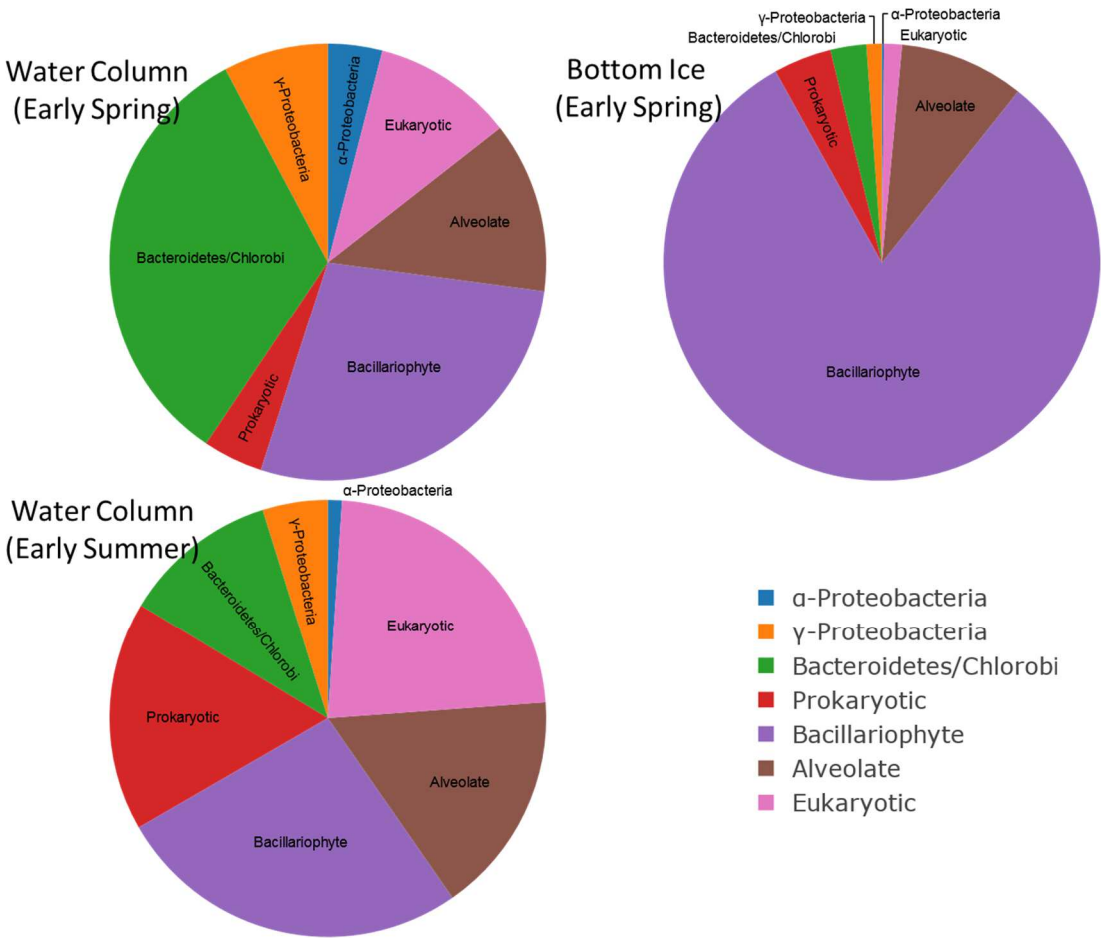

Figure 20: Phylogenetic assignments of reads from Total RNA libraries for (top) early spring ice and water column meta-transcriptomes, and late spring water column meta-transcriptome (after ice community has sloughed)

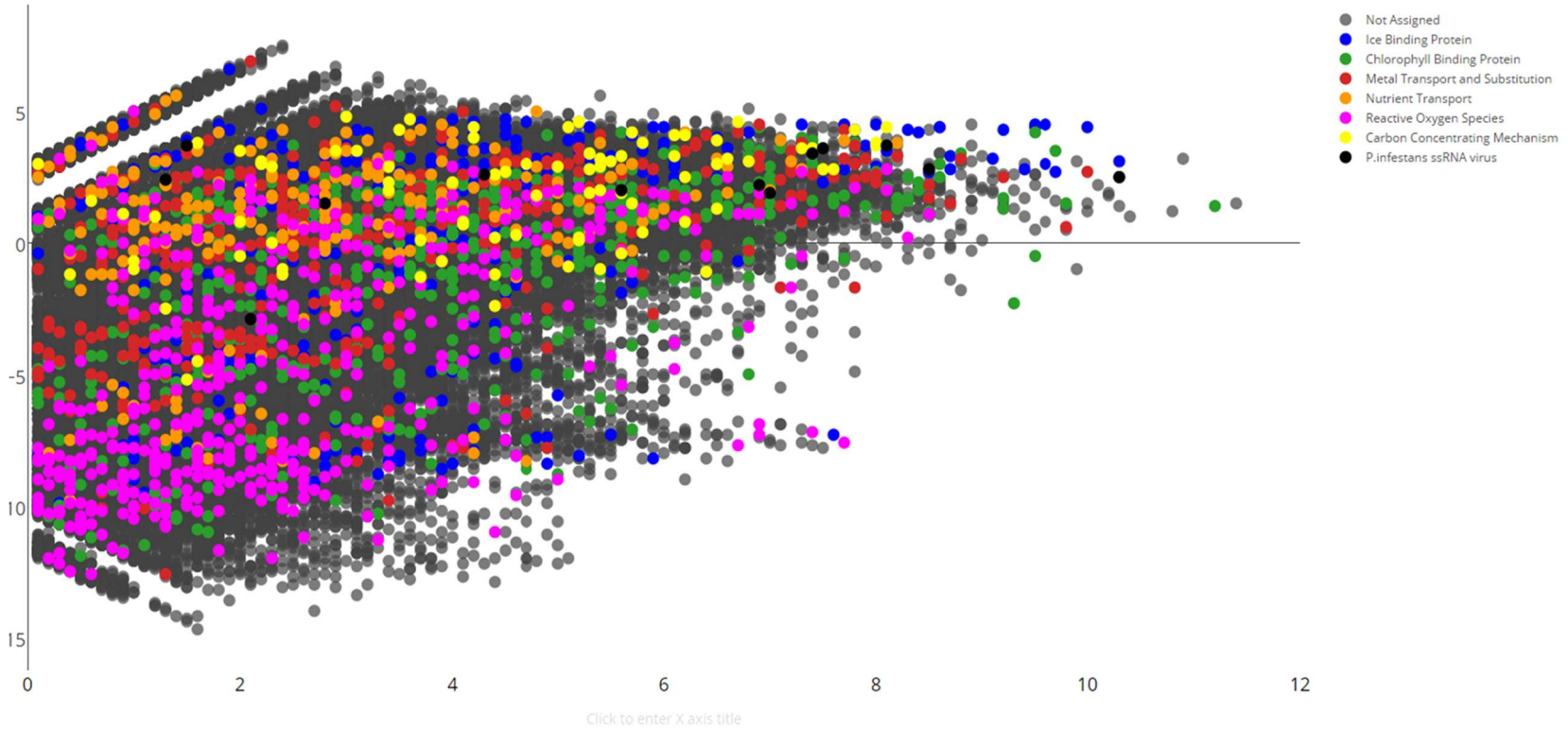

Figure 21: Log_2fold change of abundance of transcripts in sea ice (Y-axis, positive values) and in water column (Y-axis, negative values) from OligoDT libraries, with fold abundance on X-axis. Transcripts corresponding to hypothesized stresses are overlaid on top.

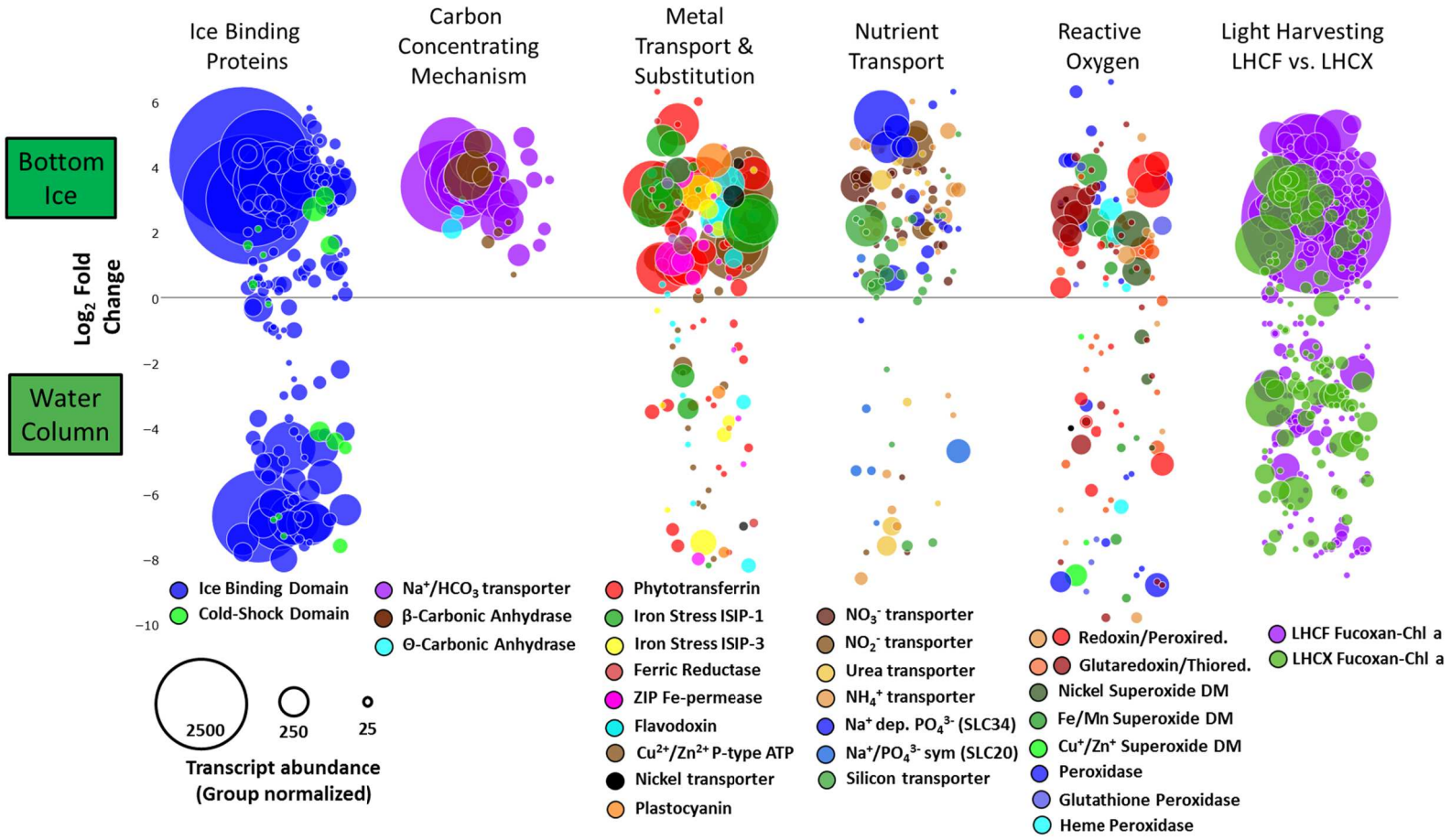

Figure 22: Bubble plot of 6 differentially expressed modules for freeze-tolerance, carbon concentration, metal- and nutrient transport, reactive oxygen and light harvesting. Bubble area is proportional to group normalized transcript abundance, differential expression is represented by the log_2fold change (y-axis).

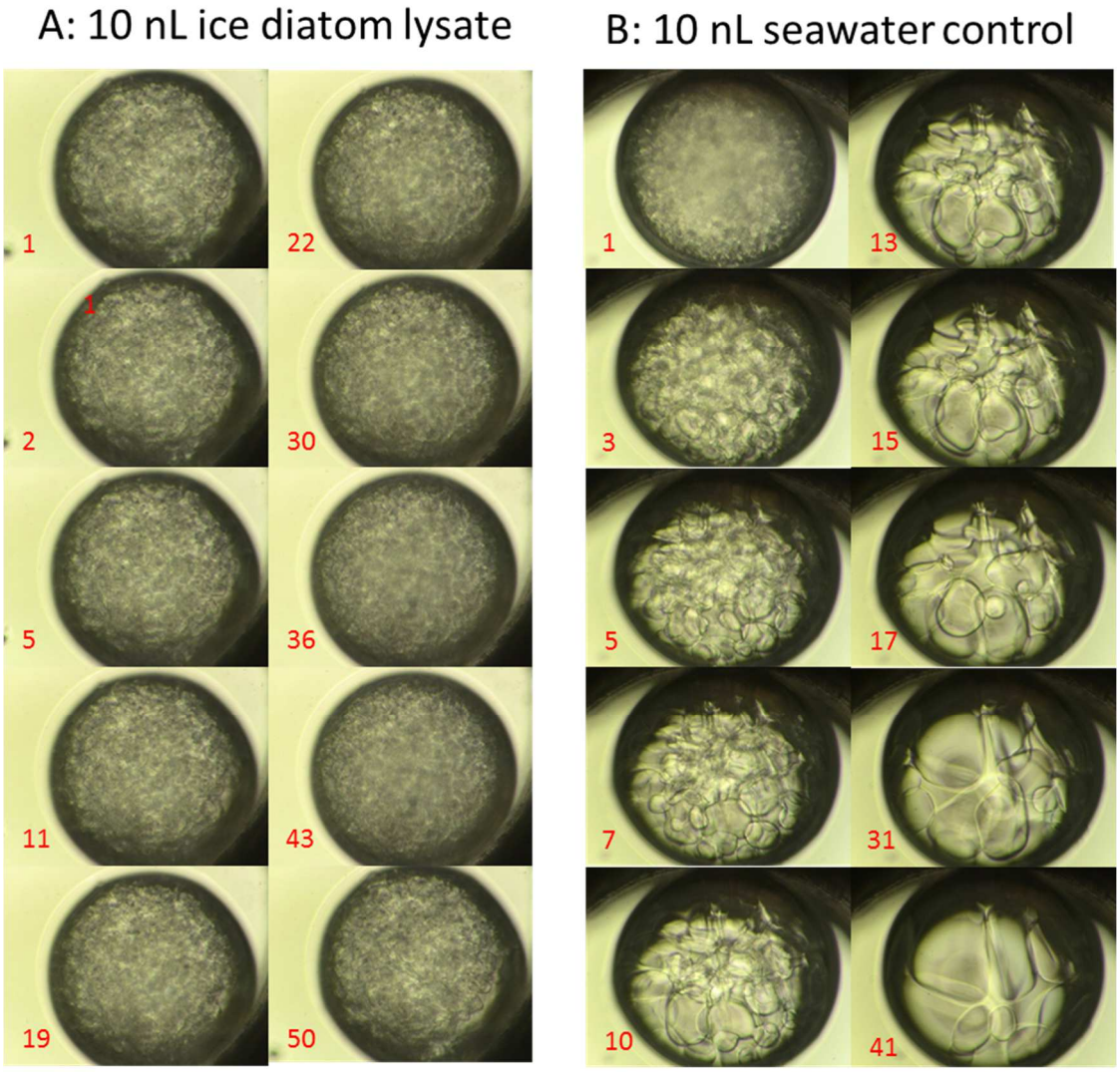

Figure 23: Ice recrystallization assay, elapsed time in minutes, lower left corner. Left: diatom lysate prevents recrystallization of and reformation of larger ice crystals. Right: a seawater control – note rapid growth of large ice crystals at expense of smaller crystals.

3.7 REFERENCES

Ackley, S. F., Buck, K. R. and Taguchi, S. Standing Crop of Algae in the Sea Ice of the Weddell Sea Region. *Deep Sea Research Part A, Oceanographic Research Papers* 26 (1979): 269–81.

Allen, A.E., Laroche J., Maheswari, U., Lommer, M., Schauer, N., Lopez P.J., Finazzi, G., Fernie, A.R. and Bowler, C. "Whole-Cell Response of the Pennate Diatom Phaeodactylum Tricornutum to Iron Starvation". *Proceedings of the National Academy of Sciences of the United States of America* 105 (2008): 10438–43.

Alonso-Saez, L., Waller, A.S., Mende, D. R., Bakker, K., Farnelid, H., Yager, P. L., Lovejoy, C. et al. "Role for Urea in Nitrification by Polar Marine Archaea". *Proceedings of the National Academy of Sciences* 109 (2012): 17989–94.

Alonso-Sáez, L., Andersson, A., Heinrich, F., Bertilsson, S. "High Archaeal Diversity in Antarctic Circumpolar Deep Waters". *Environmental Microbiology* Reports 3 (2011): 689–97.

Arrigo, K. R. et al. "Phytoplankton Community Structure and the Drawdown of Nutrients and CO2 in the Southern Ocean". *Science* 283 (1999): 365–67.

Arrigo, K. R, Thomas, D. N. "Large Scale Importance of Sea Ice Biology in the Southern Ocean". *Antarctic Science* 16 (2004): 471–86.

Arrigo, K. R. "Phytoplankton Dynamics within 37 Antarctic Coastal Polynya Systems". *Journal of Geophysical Research* 108 (2003): 3271.

Arrigo, K. R., Worthen, D.L., Lizotte, M.P., Dixon, P., Dieckmann, G. "Primary Production in Antarctic Sea Ice". *Science* 276 (1997): 394–97.

Barnett, T. P. "Penetration of Human-Induced Warming into the World's Oceans". *Science* 309 (2005): 284–87.

Maddalena, B., Uhlig, C., John, U., Mock, T., Valentin K. "Antifreeze Proteins in Polar Sea Ice Diatoms: Diversity and Gene Expression in the Genus Fragilariopsis". *Environmental Microbiology* 12 (2010): 1041–52.

Bertrand, Erin M, John P McCrow, Ahmed Moustafa, Hong Zheng, Jeffrey B McQuaid, Tom O Delmont, Anton F Post, et al. Phytoplankton-Bacterial Interactions Mediate Micronutrient Colimitation at the Coastal Antarctic

Sea Ice Edge. *Proceedings of the National Academy of Sciences of the United States of America* 112 (2015): 9938–43.

Boyd, P W, T Jickells, C S Law, S Blain, E a Boyle, K O Buesseler, K H Coale, et al. Mesoscale Iron Enrichment Experiments 1993-2005: Synthesis and Future Directions. *Science* (New York, N.Y.) 315 (2007): 612–17.

Christner, Brent C., John C. Priscu, Amanda M. Achberger, Carlo Barbante, Sasha P. Carter, Knut Christianson, Alexander B. Michaud, et al. A Microbial Ecosystem beneath the West Antarctic Ice Sheet. *Nature* 512 (2014). Nature Publishing Group: 310–13.

Cziko, Paul a., Arthur L. DeVries, Clive W. Evans, and Chi-Hing Christina Cheng. Antifreeze Protein-Induced Superheating of Ice inside Antarctic Notothenioid Fishes Inhibits Melting during Summer Warming. *Proceedings of the National Academy of Sciences* 111 (2014): 14583–88.

Davies, Peter L. 2014. Ice-Binding Proteins: A Remarkable Diversity of Structures for Stopping and Starting Ice Growth. Trends in Biochemical *Sciences* 39 (11). Elsevier Ltd: 548–55.

Davis, Aubrey, Raffaela Abbriano, Sarah R. Smith, and Mark Hildebrand. 2016. Clarification of Photorespiratory Processes and the Role of Malic Enzyme in Diatoms. *Protist* 168 (1). Elsevier GmbH.: 134–53..

Delille, Bruno, Bruno Jourdain, Alberto V Borges, and Daniel Delille. 2007. Biogas (CO2 , O2 , Dimethylsulfide) Dynamics in Spring Antarctic Fast Ice *Limnology and Oceanography* 52 (4): 1367–79.

DeLong, Edward F, Ke Y Wu, Barbara B Prézelin, and Raffael V Jovine. 1994. High Abundance of Archaea in Antarctic Marine Picoplankton. *Nature*.

DiTullio, G R, J M Grebmeier, K R Arrigo, M P Lizotte, D H Robinson, a Leventer, J P Barry, M L VanWoert, and R B Dunbar. 2000. Rapid and Early Export of Phaeocystis Antarctica Blooms in the Ross Sea, Antarctica. *Nature* 404 (6778): 595–98.

Elbourne, Liam D H, Sasha G. Tetu, Karl A. Hassan, and Ian T. Paulsen. 2016. OUP Accepted Manuscript. *Nucleic Acids Research* 45 (November 2016): 1–15.

Engelen, S, P Hingamp, M Sieracki, Cd Vargas, S Audic, N Henry, J Decelle, et al. 2015. Eukaryotic Plankton Diversity in the Sunlit Ocean. *Science* 348 (MAY): 1–12.

Fernandez-Gomez, Beatriz, Michael Richter, Margarete Schuler, Jarone Pinhassi, Silvia G Acinas, Jose M Gonzalez, and Carlos Pedros-Alio. 2013. Ecology of Marine Bacteroidetes: A Comparative Genomics Approach. The ISME Journal 7 (5): 1026–37. doi:10.1038/ismej.2012.169.
Gledhill Martha, and Kristen N. Buck. 2012. The Organic Complexation of Iron in the Marine Environment: A Review. *Frontiers in Microbiology* 3 (FEB): 1–17.

Gleitz, Markus, Michiel Rutgers v.d. Loeff, David N Thomas, Gerhard S Dieckmann, and Frank J Millero. 1995. Comparison of Summer and Winter Inorganic Carbon, Oxygen and Nutrient Concentrations in Antarctic Sea Ice Brine. *Marine Chemistry* 51 (2): 81–91.

Grzymski, Joseph J, Christian S Riesenfeld, Timothy J Williams, Emilie Long, Flavia Evans, Mathew Z Demaere, Federico M Lauro, et al. 2012. A Metaproteomic Assessment of Winter and Summer Bacterioplankton from Antarctic Peninsula Coastal Surface Waters. *The ISME Journal* 6 (10). Nature Publishing Group: 1883–1900.

Haimovich-Dayan, Maya, Nitsan Garfinkel, Daniela Ewe, Yehouda Marcus, Ansgar Gruber, Heiko Wagner, Peter G. Kroth, and Aaron Kaplan. 2013. The Role of C4 Metabolism in the Marine Diatom Phaeodactylum Tricornutum. *New Phytologist* 197 (1): 177–85.

Herlemann, Daniel Pr R, Matthias Labrenz, Klaus Jürgens, Stefan Bertilsson, Joanna J Waniek, Anders F Andersson, Klaus Ju, Stefan Bertilsson, Joanna J Waniek, and Anders F Andersson. 2011. Transitions in Bacterial Communities along the 2000 Km Salinity Gradient of the Baltic Sea. *The ISME Journal* 5: 1571–79.

Hosie, Graham, Philippe Koubbi, Martin Riddle, Catherine Ozouf-Costaz, Masato Moteki, Mitsuo Fukuchi, Nadia Ameziane, Takashi Ishimaru, and Anne Goffart. 2011. CEAMARC, the Collaborative East Antarctic Marine Census for the Census of Antarctic Marine Life (IPY # 53): An Overview. *Polar Science* 5 (2). Elsevier B.V. and NIPR: 75–87.

Keeling, Patrick J., Fabien Burki, Heather M. Wilcox, Bassem Allam, Eric E. Allen, Linda a. Amaral-Zettler, E. Virginia Armbrust, 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.

Kikutani, Sae, Kensuke Nakajima, Chikako Nagasato, Yoshinori Tsuji, Ai Miyatake, and Yusuke Matsuda. 2016. Thylakoid Luminal Θ-Carbonic Anhydrase Critical for Growth and Photosynthesis in the Marine Diatom

Phaeodactylum Tricornutum. *Proceedings of the National Academy of Sciences of the United States of America* 113 (35): 9828–33.

Konstantinidis, Konstantinos T., Jennifer Braff, David M. Karl, and Edward F. DeLong. 2009. Comparative Metagenomic Analysis of a Microbial Community Residing at a Depth of 4,000 Meters at Station ALOHA in the North Pacific Subtropical Gyre. *Applied and Environmental Microbiology* 75 (16): 5345–55.

Kroth, Peter G, Anthony Chiovitti, Ansgar Gruber, Veronique Martin-Jezequel, Thomas Mock, Micaela Schnitzler Parker, Michele S Stanley, et al. 2008. A Model for Carbohydrate Metabolism in the Diatom Phaeodactylum Tricornutum Deduced from Comparative Whole Genome Analysis. *PloS One* 3 (1): e1426.

Kühl, Michael, Ronnie N. Glud, Jens Borum, Rodney Roberts, and Søren Rysgaard. 2001. Photosynthetic Performance of Surface-Associated Algae below Sea Ice as Measured with a Pulse-Amplitude-Modulated (PAM) Fluorometer and O2 Microsensors. *Marine Ecology Progress Series* 223 (1): 1–14.

Lacarra, M., M. N. Houssais, E. Sultan, S. R. Rintoul, and C. Herbaut. 2011. Summer Hydrography on the Shelf off Terre Adélie/George V Land Based on the ALBION and CEAMARC Observations during the IPY. *Polar Science* 5 (2). Elsevier B.V. and NIPR (2011): 88–103.

Lane, J, Bernadette Pace, Gary J Olsen, David a Stahl, Mitchell L Sogin, Norman R Pace, David J Lane, David a Stahlt, and Mitchell L Sogint. 1986. Rapid Determination of 16S Ribosomal RNA Sequences for Phylogenetic Analyses. *Evolution* 83.

Lommer, Markus, Michael Specht, Alexandra-Sophie Roy, Lars Kraemer, Reidar Andreson, Magdalena a Gutowska, Juliane Wolf, et al. 2012. Genome and Low-Iron Response of an Oceanic Diatom Adapted to Chronic Iron Limitation. *Genome Biology* 13 (7): R66.

Luo, Haiwei, Bradley B Tolar, Brandon K Swan, Chuanlun L Zhang, Ramunas Stepanauskas, Mary Ann Moran, and James T Hollibaugh. 2014. Single-Cell Genomics Shedding Light on Marine Thaumarchaeota Diversification. *The ISME Journal* 8 (3). Nature Publishing Group: 732–36.

Marchant, H, a Davidson, S Wright, and J Glazebrook. 2000. The Distribution and Abundance of Viruses in the Southern Ocean during Spring. *Antarctic Science* 12 (04): 414–17.

Marchetti, Adrian, David M Schruth, Colleen a Durkin, Micaela S Parker, Robin B Kodner, Chris T Berthiaume, Rhonda Morales, Andrew E Allen, and E Virginia Armbrust. 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.

McMinn, a., K. G. Ryan, P. J. Ralph, and a. Pankowski. 2007. Spring Sea Ice Photosynthesis, Primary Productivity and Biomass Distribution in Eastern Antarctica, 2002-2004. *Marine Biology* 151 (3): 985–95.

Mock, Thomas, and David N. Thomas. 2005. Recent Advances in Sea-Ice Microbiology. *Environmental Microbiology* 7 (5): 605–19.

Moore, C M, M M Mills, K R Arrigo, I Berman-Frank, L Bopp, P W Boyd, E D Galbraith, et al. 2013. Processes and Patterns of Oceanic Nutrient Limitation. Nature Geosci 6 (9). *Nature* Publishing Group: 701–10.

Morel, François M. M., a. B. Kustka, and Y. Shaked. 2008. The Role of Unchelated Fe in the Iron Nutrition of Phytoplankton. *Limnology and Oceanography* 53 (1): 400–404.

Morrissey, Joe, Robert Sutak, Javier Paz-Yepes, Atsuko Tanaka, Ahmed Moustafa, Alaguraj Veluchamy, Yann Thomas, et al. 2015. A Novel Protein, Ubiquitous in Marine Phytoplankton, Concentrates Iron at the Cell Surface and Facilitates Uptake. *Current Biology* 25 (3): 364–71.

Nakajima, Kensuke, Atsuko Tanaka, and Yusuke Matsuda. 2013. SLC4 Family Transporters in a Marine Diatom Directly Pump Bicarbonate from Seawater. *Proceedings of the National Academy of Sciences of the United States of America* 110 (5): 1767–72.

Noble, A. E, Moran, D.M., Allen A.E., and Saito, M. A. 2013. Dissolved and Particulate Trace Metal Micronutrients under the McMurdo Sound Seasonal Sea Ice: Basal Sea Ice Communities as a Capacitor for Iron. *Frontiers in Chemistry* 1 (October): 25.

Padman, Lawrence, Svetlana Erofeeva, and Ian Joughin. 2003. Fronts and Upper Ocean Thermal Variability South of New Zealand. *Antarctic Science* 15 (1): 141–52.

.

Papadimitriou, S., D.N. Thomas, H. Kennedy, C. Haas, H. Kuosa, a. Krell, and G.S. Dieckmann. 2007. Biogeochemical Composition of Natural Sea Ice Brines from the Weddell Sea during Early Austral Summer. *Limnology and Oceanography* 52 (5): 1809–23.

Peers, Graham, and Neil M Price. 2006. Copper-Containing Plastocyanin Used for Electron Transport by an Oceanic Diatom. *Nature* 441 (7091): 341–44.

Podell, Sheila, Terry Gaasterland, Ev Koonin, Ks Makarova, L Aravind, Jg Lawrence, H Ochman, et al. 2007. DarkHorse: A Method for Genome-Wide Prediction of Horizontal Gene Transfer. *Genome Biology* 8 (2): R16.

Raven, John a. 1991. Plant Responses to High O2 Concentrations: Relevance to Previous High O2 Episodes. *Global and Planetary Change* 5 (1-2): 19–38.

Rintoul, S. R. 1998. "On the Origin and Influence of Adélie Land Bottom Water. Ocean, Ice, and Atmosphere: Interactins at the Antarctic Continental Margin" 75: 151–71.

Robinson, Mark D., Davis J. McCarthy, and Gordon K. Smyth. 2009. edgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* 26 (1): 139–40.

Sabine, Christopher L., Richard a. Feely, Nicolas Gruber, Robert M. Key, Kitack Lee, John L Bullister, Rik Wanninkhof, et al. 2004. The Oceanic Sink for Anthropogenic CO2. *Science* 305 (5682): 367–71.

Satoh, Dan, Yasutaka Hiraoka, Brian Colman, and Yusuke Matsuda. 2001. Physiological and Molecular Biological Characterization of Intracellular Carbonic Anhydrase from the Marine Diatom Phaeodactylum tricornutum1. *Plant Physiology* 126: 1459–70.

Schmieder, Robert, Yan Wei Lim, and Robert Edwards. 2012. Identification and Removal of Ribosomal RNA Sequences from Metatranscriptomes. *Bioinformatics* 28 (3): 433–35.

Sedwick, Peter N., and Giacomo R. DiTullio. 1997. Regulation of Algal Blooms in Antarctic Shelf Waters by the Release of Iron from Melting Sea Ice. *Geophysical Research Letters* 24 (20): 2515.

Shadwick, E. H., B. Tilbrook, and G. D. Williams. 2014. Carbonate Chemistry in the Mertz Polynya (East Antarctica): Biological and Physical Modification of Dense Water Outflows and the Export of Anthropogenic CO2. *Journal of Geophysical Research: Oceans* 119 (1): 1–14.

Shi, Yanmei, Gene W Tyson, John M Eppley, and Edward F DeLong. 2011. Integrated Metatranscriptomic and Metagenomic Analyses of Stratified Microbial Assemblages in the Open Ocean. *The ISME Journal* 5 (6). Nature Publishing Group: 999–1013.

Søgaard, Dorte Haubjerg, Morten Kristensen, Søren Rysgaard, Ronnie Nøhr Glud, Per Juel Hansen, and Karen Marie Hilligsøe. 2010. Autotrophic and Heterotrophic Activity in Arctic First-Year Sea Ice: Seasonal Study from Malene Bight, SW Greenland. *Marine Ecology* Progress Series 419: 31–45.

Stoeck, Thorsten, David Bass, Markus Nebel, Richard Christen, Meredith D M Jones, Hans Werner Breiner, and Thomas a. Richards. 2010. Multiple Marker Parallel Tag Environmental DNA Sequencing Reveals a Highly Complex Eukaryotic Community in Marine Anoxic Water. *Molecular Ecology* 19 (SUPPL. 1): 21–31.

Strutton, Peter G, F Brian Gri, Raechel L Waters, Simon W Wright, and Nathaniel L Bindo. 2000. Primary Productivity Of The Coast of East Antarctica (80° to 150° E): January to March 1996 47 (March 1996): 2281–98.

Swan, Brandon K., Mark D. Chaffin, Manuel Martinez-Garcia, Hilary G. Morrison, Erin K. Field, Nicole J. Poulton, E. Dashiell P Masland, et al. 2014. Genomic and Metabolic Diversity of Marine Group I Thaumarchaeota in the Mesopelagic of Two Subtropical Gyres. *PLoS ONE* 9 (4).

Tachibana, Masaaki, Andrew E. Allen, Sae Kikutani, Yuri Endo, Chris Bowler, and Yusuke Matsuda. 2011. Localization of Putative Carbonic Anhydrases in Two Marine Diatoms, Phaeodactylum Tricornutum and Thalassiosira Pseudonana. *Photosynthesis Research* 109 (1-3): 205–21.

Talley, Lynne D. 2013. Closure of the Global Overturning Circulation through the Indian, Pacific and Southern Oceans: Schematics and Transports. *Journal of Chemical Information and Modeling* 53 (9): 1689–99.

Thomas, D N, and G S Dieckmann. 2002. Antarctic Sea Ice--a Habitat for Extremophiles. *Science* 295 (5555): 641–44.

Tolar, B B, M J Ross, N J Wallsgrove, and Q Liu. 2016. Contribution of Ammonia Oxidation to Chemoautotrophy in Antarctic Coastal Waters. The ISME 10 (11). *Nature* Publishing Group: 1–15.

Vick-Majors, Trista J., Amanda Achberger, Pamela Santibáñez, John E. Dore, Timothy Hodson, Alexander B. Michaud, Brent C. Christner, et al. 2016. Biogeochemistry and Microbial Diversity in the Marine Cavity beneath the McMurdo Ice Shelf, Antarctica. *Limnology and Oceanography* 61 (2): 572–86.

Wilkins, David, Federico M Lauro, Timothy J Williams, Matthew Z Demaere, Mark V Brown, Jeffrey M Hoffman, Cynthia Andrews-Pfannkoch, et al. 2013. Biogeographic Partitioning of Southern Ocean Microorganisms Revealed by Metagenomics. *Environmental Microbiology* 15 (5): 1318–33.

Williams, G. D., S. Aoki, S. S. Jacobs, S. R. Rintoul, T. Tamura, and N. L. Bindoff. 2010. Antarctic Bottom Water from the Ad??lie and George v Land Coast, East Antarctica (140-149??E). *Journal of Geophysical Research: Oceans* 115 (4): 1–29.

Williams, Guy D., N. L. Bindoff, Simon J. Marsland, and Steve R. Rintoul. 2008. Formation and Export of Dense Shelf Water from the Ad??lie Depression, East Antarctica. *Journal of Geophysical Research*: *Oceans* 113 (4): 1–12.

Wright, S. W., D. P. Thomas, H. J. Marchant, H. W. Higgins, M. D. Mackey, and D. J. Mackey. 1996. Analysis of Phytoplankton of the Australian Sector of the Southern Ocean: Comparisons of Microscopy and Size Frequency Data with Interpretations of Pigment HPLC Data Using the 'CHEMTAX' Matrix Factorisation Program. *Marine Ecology Progress Series* 144 (1-3): 285–98.

Yooseph, Shibu, Cynthia Andrews-Pfannkoch, Aaron Tenney, Jeff McQuaid, Shannon Williamson, Mathangi Thiagarajan, Daniel Brami, et al. 2013. A Metagenomic Framework for the Study of Airborne Microbial Communities. *PLoS ONE* 8 (12).

Zeigler Allen, Lisa, Eric E Allen, Jonathan H Badger, John P McCrow, Ian T Paulsen, Liam D H Elbourne, Mathangi Thiagarajan, et al. 2012. Influence of Nutrients and Currents on the Genomic Composition of Microbes across an Upwelling Mosaic. *The ISME Journal* 6 (7). Nature Publishing Group: 1403–14.