UNIVERSITY OF CALIFORNIA
SANTA CRUZ

The Molecular Diversity and Physiology of Polar Phytoplankton

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

DOCTOR OF PHILOSOPHY

in

OCEAN SCIENCES

By

Maria Hamilton

June 2021

The dissertation of Maria Hamilton is approved:

_____

Professor Alexandra Z. Worden, Chair

_____

Associate Professor Phoebe Lam

_____

Professor Raphael M. Kudela

_____

Professor Maria Vernet

_____

Quentin Williams

Acting Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

The Molecular Diversity and Physiology of Polar Phytoplankton

Maria Hamilton

Little is known about the molecular diversity of the phytoplankton inhabiting the fjords of the West Antarctic Peninsula (WAP), despite the status of this region as being among the most vulnerable to future warming. Additionally, there are many gaps in our knowledge of how polar phytoplankton will respond to the effects of climate change or about the unique life history strategies they may have evolved. The main objective of this dissertation is to develop an improved understanding of the physiology of a model polar phytoplankton species and to characterize phytoplankton community structure in terms of molecular diversity in a climate-sensitive polar environment. The phytoplankton community in a WAP fjord was surveyed by sequencing the V9 region of the 18S rRNA gene and the V1-V2 region of the plastid derived 16S rRNA gene (Chapter 2). In addition, photosynthetic cells were enumerated by flow cytometry. A novel cryptophyte was observed and the phylogeny and global distribution of this cryptophyte was then examined. Among other phytoplankton, the picoprasinophyte, *Micromonas polaris*, was also observed in these Antarctic samples. With the genome of *M. polaris* CCMP2099 newly in hand, we aimed to improve our understanding of how polar phytoplankton may respond to the future effects of climate change by examining *M. polaris'* physiological and transcriptional responses to changing CO_2 and nitrate conditions in continuous-flow

photo-bioreactors (Chapter 3). Finally, we explored the potential survivability and transcriptional responses by this alga as it descends to the deep ocean, building on prior observations of *M. polaris* in North Atlantic Deep Water, using a series of *in situ* incubation experiments (Chapter 4). Collectively, these studies provide baseline information on polar phytoplankton community structure and physiological capability, which will help to guide future assessments of the phytoplankton in these regions under a changing climate.

Acknowledgements

First and foremost, I would like to thank my PhD advisor, Alexandra Worden, for her support and guidance throughout my time in her lab, and for shaping me into the scientist that I am today. Additional thanks to my dissertation reading committee for taking the time to offer constructive feedback and advice: Phoebe Lam, Raphe Kudela, and Maria Vernet. A tremendous amount of support, both emotional and academic, was collectively provided by members of the Worden Lab- thank you for all of the guidance and assistance at sea, in the lab, and through the computer. In particular, I would like to thank Sebastian Sudek for all of the supremely useful science and science-adjacent conversations, and for being such an incredibly patient teacher of lab and fieldwork techniques. Additionally, I'd like to thank the members of the Fjord Eco project, as well as the captain and crew of the R/V *Lawrence M. Gould* and RVIB *Nathaniel B. Palmer*, who efforts were essential for the work included in Chapter 2 of this dissertation. A particular thank you goes to Martina Mascioni for her extensive feedback and assistance with manuscript editing. Furthermore, I'd like to thank the MBARI support staff and scientists, as well as the captain and crew of the R/V *Western Flyer*, and the pilots of the ROV *Doc Ricketts*- specifically the helpful and creative, Knute Brekke. Finally, special thanks to my biological family and to my Santa Cruz family for providing their unwavering love and support for me on this journey.

This research was financially supported by grants from the Gordon and Betty Moore Foundation, the National Science Foundation, and the David and Lucille Packard Foundation.

Chapter 1: Introduction

The organisms living in polar marine environments, both in the Arctic and Antarctic, have had to adapt to conditions that are very different from the rest of the global ocean. These high latitude regions are characterized by cold temperatures, extreme seasonal changes in day/night length, and influences from the cryosphere (Kirst and Wiencke 1995). Polar phytoplankton have evolved to survive, and even thrive, at high latitudes and are important for their contributions to global carbon fixation and for their role as primary producers, forming the base of marine food webs. Understanding the ecology and physiology of polar phytoplankton in the context of future climate change is critical, as high latitude regions are believed to be among those most affected by warming (Doney et al. 2012).

Warming is only one aspect of climate change. Increased concentrations of dissolved CO_2 and reduced surface nutrients are additional factors that are predicted to alter phytoplankton communities. The reduced surface nutrients are a result of greater stratification of the water column, due to surface warming. The stronger density gradient restrains the ability for nutrient-rich deep water to replenish the surface via entrainment processes. These effects have been well documented in the subtropical gyres of the Atlantic and Pacific, but are also being investigated in other regions, including the Arctic (Polovina et al. 2008; Tremblay et al. 2008). While an increase in CO_2 availability may positively affect some taxa during bloom scenarios where dissolved CO_2 can become limiting, subsequent ocean acidification could negatively affect others, resulting in phytoplankton "winners" and "losers" under

future climate change (Dutkiewicz et al. 2015; Bach et al. 2017). In polar regions, the effect of increased atmospheric CO_2 may be more pronounced, as decreased temperatures increase the solubility of gases into the ocean (Hoegh-Guldberg and Bruno 2010; Qi et al. 2017). Increased CO_2 is thought to alter the phytoplankton community structure; in some Arctic regions warming has already been reported and/or predicted to accelerate a shift towards picophytoplankton (operational diameter of <2 or <3µm depending on the research group) dominance in Arctic waters (Li et al. 2009; Brussaard et al. 2013; Hoppe et al. 2018). As the Arctic ocean warms, the reduction of summer sea ice and subsequent increase in irradiation is predicted to result in increased annual primary production (Arrigo et al. 2008; Steinacher et al. 2010). However, others have suggested that nitrogen, not light, limits photosynthetic growth in much of the Arctic, leaving the smallest sized phytoplankton as the best competitors under warming scenarios (Li et al. 2009; Tremblay and Gagnon 2009; Grebmeier et al. 2010).

Across the Arctic, the prasinophyte *Micromonas polaris* (*Micromonas* clade E2) appears to be among the most common picophytoplankton based on non-quantitative molecular analyses, and seemingly inhabits the ecological niche typically occupied by picocyanobacteria in warmer waters (Lovejoy et al. 2006; Lovejoy et al. 2007; Lovejoy et al. 2011; Kilias et al. 2014). This recently described species has also been found in Arctic sea ice (Belevich et al. 2018). *Micromonas* cells are 1-3 µm long, with one flagellum, a single chloroplast, and no cell wall (Manton and Parke 1960). *M. polaris* was previously considered endemic to the Arctic, constrained

geographically by its psychrophilic nature (Lovejoy et al. 2007). It has now been reported in the Antarctic, based on metagenome analyses and barcode amplicon sequencing studies (Simmons et al. 2015; Trefaut et al. 2021). This suggests the alga utilizes some sort of global transport mechanism and one theory is that *M. polaris,* as a cyst, spore, or other type of resting-stage cell, is transported longitudinally via the North Atlantic Deep Water (NADW) that is formed in the Labrador and Nordic Seas and upwelled in the Southern Ocean (Simmons et al. 2015). In support of this hypothesis, molecular analyses have demonstrated that *M. polaris* sequences are present in the NADW about 27 years into its North to South flow in the Atlantic Ocean, but not in Gyre samples from the water column above the NADW (Simmons et al. 2015).

Although there have not been any direct observations of an alternate life stage of *Micromonas* than that seen in standard culture conditions, there is some evidence that a cell-walled stage may be possible. Genes encoding hydroxyproline-rich glycoproteins (HRGP), which are known to be the components of cell walls in plants and other algae, have been identified in the genomes of two *Micromonas* species, as well as genes related to sexual reproduction (Worden et al. 2009). Sexual reproduction is linked with the formation of a resistant cell-walled life stage in a number of phytoplankton species, including the green alga, *Chlamydomonas reinhardtii* (Goodenough et al. 2007; Ellegaard and Ribeiro 2018). Additionally, some Antarctic prasinophytes have been observed to form cysts, including three species of *Pyramimonas*: *P. tychotreta*, *P. gelidicola*, and *P. amylifera* (Hargraves

and Gardiner 1980; van den Hoff et al. 1989; Daugbjerg et al. 2000). While cysts and other resting-stage cell forms have not been well studied in prasinophyte green algae, this phenomenon has been observed and widely researched in diatom and dinoflagellate species (Ellegaard and Ribeiro 2018). It is also thought that the ability to produce a cyst could be a survival method for polar phytoplankton during winter periods of 24-hour darkness (Mcminn and Martin 2013).

Compared with much of the global ocean, much less is known about the molecular diversity of phytoplankton in polar environments. The combination of remote geography and obstructive sea ice conditions has hindered oceanographic research in these regions, particularly in the Antarctic. Around the continent, there are still only a small number of studies that have investigated phytoplankton community structure using a molecular approach, relative to the Arctic (López-García et al. 2002; Wolf et al. 2013; Abele et al. 2017; Trefaut et al. 2021). In combination with classical methods (pigment analyses or microscopy), molecular-based diversity studies are needed for improved identification of taxa and a more developed understanding of the photosynthetic protists in Antarctic waters. The need to establish baseline knowledge on these communities is especially urgent, as polar regions are experiencing some of the most rapid changes to climate on Earth (Doney et al. 2012). I aim to address this need in Chapter 2 of this dissertation, which specifically examines protistan diversity in an Antarctic fjord, with a focus on phytoplankton community composition from a molecular perspective in two seasons. The primary method used in this chapter was sequencing of the ribosomal RNA (rRNA) gene amplicons, a technique that has been

widely adopted in investigations of microbial communities in a variety of environments, including the ocean. Here, primers that targeted the V1-V2 region of the 16S rRNA gene were used to target the plastid-containing members of the community, while primers targeting the V9 region of the 18S rRNA gene were also employed to capture the eukaryotic microbe community as a whole.

M. polaris was among the members of the Antarctic photosynthetic community identified in Chapter 2, and in Chapters 3 and 4, I sought to further our understanding of this model polar phytoplankton's physiological and molecular capabilities with both laboratory and *in situ* incubation experiments. The RNA-Seq approach utilized to examine gene expression in these later two chapters was aided by the recent sequencing of the *M. polaris* genome, performed by the Worden Lab. While a full functional annotation remains underway, access to the genome has allowed for genome-guided assembly of the transcriptomes generated in these experiments. As part of this dissertation, a select number of relevant genes were also manually annotated.

The objectives of this dissertation were to:

1. Characterize the protistan community in a fjord in the West Antarctic Peninsula, and link phytoplankton dynamics to environmental conditions.
2. Assess the physiological and transcriptional response of the polar phytoplankton species, *Micromonas polaris*, to altered CO_2 and nitrate conditions.

3. Evaluate the viability of *M. polaris* after incubations in the deep sea, and examine the transcriptional response of this alga to extreme conditions.

This dissertation aims to enhance our understanding of the phytoplankton community structure in polar waters, as well as explore the physiology of the polar phytoplankton species, *Micromonas polaris*.

Chapter 2: Spatio-temporal variability in the molecular diversity of Antarctic protistan communities reveals dominance by a novel cryptophyte alga associated with warming influences*

*This chapter has been submitted for publication in the journal mBio

Abstract

The Andvord fjord in the West Antarctic Peninsula is known for its productivity and aggregations of megafauna. Although considered a highly climate sensitive ecosystem, the molecular diversity of protists there is virtually unknown, including photosynthetic taxa that underpin productivity. Because seasonally resolved data is completely lacking, we performed spring and fall expeditions pursing diversity, phytoplankton abundances, and how community profiles connect to sea ice and meltwater extent, parameters heavily influenced by climate change. 18S rRNA Amplicon Sequence Variant (ASV) analyses revealed diverse predatory protists, spanning multiple eukaryotic supergroups, alongside enigmatic heterotrophs like Picozoa. Among photosynthetic protists, cryptophyte contributions were marked, and analysis of plastid-derived 16S rRNA ASVs supported 18S ASV results, including a dichotomy between cryptophytes and diatoms seen in other Antarctic regions. We demonstrate that this discordance reflects different community structures, specifically diverse diatoms, including *Chaetoceros sp.* and *Fragilariopsis cylindrus*, versus a single dominant cryptophyte. Phylogenetic reconstructions and electron microscopy performed on field samples, demonstrated the dominant cryptophyte alga represents an evolutionarily distinct, uncultivated lineage (named herein Clade VII). Analysis of global data, showed that *Teleaulax amphioxeia* and *Plagioselmis prolonga* dominate cryptophytes at a pan-oceanic scale, including Arctic samples. However, they were not detected in the WAP, which is dominated by the novel cryptophyte, with trace,

consistent contributions from the cultured alga *Geminigera crypophila*. Furthermore, Clade VII cell abundances correlate with increased meltwater and temperature. Thus, our studies indicate that a newly identified dominant primary producer thrives in conditions that will be enhanced in Antarctica by climate-induced warming.

Introduction

The West Antarctic Peninsula (WAP) is considered to be among the most climate-sensitive regions on Earth (Turner et al. 2005). The marine environment in this physically complex region has high levels of primary productivity, in turn supporting a variety of benthic and pelagic life, including many marine mammals (Thiele et al. 2004; Ducklow et al. 2007; Ware et al. 2011). Unlike lower latitude oceans, where both cyanobacteria and photosynthetic unicellular eukaryotes contribute to primary production via photosynthesis, the WAP food web is underpinned solely by photosynthetic protists, but the taxa that dominate these communities and the environmental factors that shape them are not well resolved. In fact, from a molecular diversity perspective neither the photosynthetic protists, nor the heterotrophic protists that consume them have ever been studied in the context of seasonal and spatial variability.

To date, much of the research on WAP primary producers has been performed using microscopy and pigment-based analyses. These studies indicate that WAP austral summer phytoplankton blooms are typically dominated by diatoms or cryptophytes, or that blooms of smaller flagellates (including cryptophytes) follow the initial diatom bloom stage (Ducklow et al. 2007). The haptophyte alga *Phaeocystis* and other small flagellates are also present (Garibotti et al. 2005; Mendes et al. 2013), although the species comprising the latter are difficult to identify by microscopy or HPLC-pigment analysis (Wright et al. 2009). In nearshore environments, blooms by the small flagellated prasinophyte, *Pyramimonas* sp. and

unidentified unarmored dinoflagellates have also been reported (Rodriguez et al. 2002; Mascioni et al. 2019). Although the controls acting directly on these important primary producer communities, such as predators, are unknown, several studies have suggested that phytoplankton community dynamics are shaped by the timing of sea ice retreat and the extent of surface water stratification.

Longer-term trends also have been identified for WAP phytoplankton communities. The most striking being a shift towards cryptophyte dominance, which is thought to be linked to increased warming based on multi-year pigment analyses on transects across the WAP (Mendes et al. 2018). Several WAP studies have also examined the molecular diversity of microbial eukaryotes, including phytoplankton, generally using the 18S rRNA gene (Luria et al. 2014; Abele et al. 2017; Lin et al. 2017; Rozema et al. 2017; Fuentes et al. 2019; Trefaut et al. 2021). However, there is a paucity of clearly contextualized repeat sampling of molecular diversity, which, together with limited information at taxonomic levels connecting to acclimatization and evolutionary processes, poses problems for assessing future change.

Additional to the general challenges of working in Antarctica, are the glacio-marine fjords that pepper the WAP coastline — narrow inlets with steeply elevated sides containing at least one tidewater glacier at its terminus (Howe et al. 2010). The physical complexity contributed by the presence of multiple fjords is overlain by the marked seasonal changes and differences in input of melted sea ice. This results in potential differences in fjord productivity levels and in how the interface between the cryosphere and ocean manifests. Although still understudied from a biological

perspective, WAP fjords are known to be colder and have a generally weaker meltwater influence than Arctic fjords (Pan et al. 2020). Lower meltwater inputs result in reduced upper ocean stratification and reduced inner fjord turbidity, factors that generally enhance phytoplankton growth and productivity in polar fjords.

Andvord Bay is a WAP fjord with high primary productivity and an abundance of diverse megafauna (Grange and Smith 2013). Five phytoplankton groups have been reported in Andvord Bay and the coastal WAP based on HPLC pigment analyses, specifically, cryptophytes, diatoms, prasinophytes, dinoflagellates, and "unidentified small phytoflagellates" (Mendes et al. 2013; Mendes et al. 2018; Pan et al. 2020). Although the taxonomic composition of the latter, or indeed several of these phytoplankton groups, is currently unknown, the small flagellates reportedly comprise 38% of the total Chlorophyll *a* (Chl *a*) summed across austral spring stations (Pan et al. 2020). In another study, small, unidentified flagellates (≤5 μm) comprised up to 90% of the phytoplankton community in samples from along the WAP coast, including Neko Harbor in Andvord Bay, based on microscopy, and unidentified cryptophytes reached up to 75% of the total phytoplankton in a bloom (Mascioni et al. 2019).

Here, we characterize the molecular diversity of unicellular eukaryotes in Andvord Bay and adjacent waters for the first time over two sampling seasons, austral spring and fall. High-throughput amplicon (V9 18S rRNA) sequencing (Amaral-Zettler et al. 2009) was used to characterize diversity of both photosynthetic and heterotrophic protists. We then more deeply examined the photosynthetic community

using V1-V2 16S rRNA primers that recover both bacterial and plastid-derived sequences, and are more reflective of photosynthetic organismal relative abundances than 18S rRNA gene amplicons, due to more constrained 16S gene copy numbers (Needham and Fuhrman 2016; Choi et al. 2020). This was combined with flow cytometric cell enumeration, which provided a quantitative understanding of phytoplankton community gradients in the fjord, in addition to diversity assessments and their association with seasonal changes. After recognizing a seemingly novel cryptophyte Amplicon Sequence Variant (ASV), we performed full-length gene sequencing and field microcopy studies to establish its phylogenetic relationships and morphology. These studies revealed an uncultured cryptophyte that forms a distinct clade separate from that containing the canonical Antarctic cryptophyte, *Geminigera cryophila* (Taylor and Lee 1971). Not only is the novel lineage present in multiple Antarctic settings, its abundance is correlated with water temperature and the degree of ice melt, factors that are directly connected to climate change.

Materials and Methods

Environmental sampling

Cruises to Andvord Bay, Gerlache Strait, and a western Antarctic shelf station, were performed from 27 November – 20 December 2015 (LMG 15-10) and 4–26 April 2016 (NBP 16-03). Niskin bottles for water collection were deployed on a rosette system with dual sensors measuring depth, temperature, salinity, fluorescence,

and light transmission. Microscopy sampling was according to surface light level percentages (50%, 12%, and 1%), with 3% (final concentration) Lugol's fixation. Flow cytometry samples were preserved with 0.25% EM grade glutaraldehyde (final concentration) (Cuvelier et al. 2010). DNA samples were filtered onto 0.2 µm 47 mm Supor filters under low vacuum, and frozen at -80^{o}C alongside flow cytometry samples. Methods for nutrients, Chl *a*, meltwater fraction, and mixed layer depth measurements are in (Pan et al. 2020).

DNA extraction and amplicon sequencing

DNA was extracted using a modified protocol of the QIAGEN DNeasy plant kit (Demir-Hilton et al. 2011), quantified using the QuBit dsDNA high-sensitivity assay (Life Technologies, Grand Island, NY, USA). For amplicon sequencing it was diluted with TE pH 8 to 1 ng µl^{-1}, with V1-V2 16S rRNA amplicon sequencing, PCR reactions as in (Sudek et al. 2015), each including 5 ng template, 5 µl 10X buffer, 1U Hi-Fi Taq, 1.6 µl 50 mM MgSO4 (Life Technologies), and 200 nM each of forward primer 27F_ill 5′-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGagrgttygatymtggctcag □3′ and reverse primer 338RPL_ill 5′□ GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGgcwgccwcccgtaggwgt □3′; capital letters represent Illumina linker sequences on the 27F/338R primer pair (Daims et al. 1999). Reactions were cycled for 2 min at 94°C, 15 sec at 94°C (30

times), 30 sec at 55°C, 1 min at 68°C and 7 min 68°C (for elongation). Purification was done using the MinElute kit (QIAGEN, Valencia, CA, USA), and product presence and removal of primer-dimer verified by electrophoresis. V9 18S rRNA amplicons were generated as above using the 1389F/1510R primer pair (Amaral-Zettler et al. 2009). Samples were sequenced using Illumina MiSeq v3 chemistry.

18S rRNA ASV generation and analyses

18S ASVs were generated by trimming the raw amplicon sequences of primers using cutadapt v3.2 and inputting them into the DADA2 pipeline (v1.19). The identity of subsequent ASVs was determined via the classifier tool implemented in Qiime2 (Bolyen et al. 2019). For taxonomic classification, V9 sequences were phylogenetically mapped onto an 18S rRNA gene reference phylogenetic tree (Wideman et al. 2020) constructed using the PR^2 database v.4.454 (Guillou et al. 2013), with short sequences (<400 bp) and/or sequences not spanning the V9 region removed, as well as metazoans. Sequence redundancy was limited by retaining only representative sequences after clustering, such that the final alignment contained 20,939 18S sequences. The maximum-likelihood tree was reconstructed using RAxML v.8.2 (Stamatakis 2006) under the GTR model with CAT approximation. Sequences used to build the tree, along with the 885 ASV V9 sequences, were aligned using MAFFT (Katoh and Standley 2013). Aligned V9 ASV sequences were then placed onto the PR^2 reference maximum-likelihood tree using the RAxML

evolutionary placement algorithm, EPA-ng v0.3.6 (Barbera et al. 2019), under GTRCAT. Subsequent to phylogenetic mapping, and for display purposes, taxa with long branches were removed from the phylogenetic tree if the terminal branch was longer than 3 substitutions per site, using information from the Newick utilities package (Junier and Zdobnov 2010) and the tree rendered using the R package ggtree (Yu et al. 2017). Note that no V9 ASVs mapped to these long branches.

Phytoplankton cell enumeration

Flow cytometry samples were analyzed on an Influx flow cytometer (BD, San Jose, CA, USA) as described (Cuvelier et al. 2010) using Fluoresbrite YG 0.75 μm beads as standards. WinList 7.0 (Verity Software House, Topsham, ME, USA) was used to analyze listmodes. Cryptophytes were distinguished from other phytoplankton based on forward angle light scatter (FALS) and phycoerythrin-derived orange fluorescence. For cell counts and dimension measurements, samples were observed at maximum amplification (400X) using an inverted optical microscope Leica DMIL LED. For scanning electron microscopy (SEM), sample aliquots were filtered onto 0.2 μm polyamide filters and dehydrated through an ethanol dilution series (25%, 50%, 75%, 100%) with final critical point dehydration. Specimens were sputter-coated with gold–palladium and examined using a Jeol JSM-6360 LV and Zeiss NTS SUPRA 40. Cryptophyte biomass was estimated by approximating average cell volume based on models from (Sun and Liu 2003) and the average cryptophyte size

measured herein. Cellular carbon content was estimated using a carbon:volume ratio for cryptophytes (Menden-Deuer and Lessard 2000). Estimates of non-cryptophyte phytoplankton biomass was generated by approximating average cell volume, assuming a spherical cell shape and 5.5 μm average length, and 237 fg C μm^{-3} as the carbon:volume ratio (Worden et al. 2004).

QPCR and standard curves were performed with a *Micromonas* primer-probe set, MicroGen08 (MicGen08F-TGTTCAAAGCGGGCTTA, MicGen08R-ATGCCCCCAACTGTTCCTCTTAA, MicGen08P- 6-FAM CCATGCTGAAATATTCAAG MGBNFQ), according to methods in (Demir-Hilton et al. 2011) using an Applied Biosystems Real Time PCR System (7500) (Foster City, CA, USA). Cycling conditions were 10 min at 95°C, followed by 45 cycles at 95°C for 15 sec and 60°C for 1 min. Inhibition tests with dilutions between 1:4 and 1:4000, established 1:40 as being appropriate for templates. Assay detection limits were 10 template copies $well^{-1}$. rDNA copies ml^{-1} were calculated by accounting for the volume of seawater filtered, template added and dilution factor used.

Plastid ASV generation and analyses

For V1-V2 16S rRNA amplicons, low-quality merged sequences and primers (at 100% match) were removed using cutadapt v1.16. A 10% read length window with a Q25 running-quality threshold was implemented for trimming low-quality bases. This data was input to the DADA2 pipeline (v1.14), and paired-end sequences

were merged with a 20-nucleotide overlap and no mismatches. The resulting ASVs were run through PhyloAssigner to determine phylogenetic classification (Choi et al. 2020). The taxonomic identity of select ASVs was also examined using BLASTn searches against the GenBank nr database.

Cryptophyte full-length 18S rRNA gene sequencing and phylogenetic analyses

To obtain full-length sequences, clone libraries were generated from 14 April Gerlache Strait DNA using a universal eukaryotic 18S rRNA gene primer set (Moon-Van Der Staay et al. 2000) as in (Cuvelier et al. 2010). After cloning and purification, sequencing was performed using Big Dye Terminator chemistry on an AB3730xl sequencer (Applied Biosystems), with initial screening using the 502F primer for the taxa of interest, and then with plasmid targeted M13F and M13R primers. GeneStudio v2.2.0.0 was used to manually curate the assemblies.

The cryptophyte 18S rRNA gene sequence generated was used as a BLASTn (Altschul et al. 1990) query against GenBank nr, to retrieve sequences from related cultured taxa, and from environmental sequences, in independent searches, that were then combined with sequences from prior cryptophyte reconstructions (Laza-Martínez et al. 2012; Daugbjerg et al. 2018). For multiple identical sequences we performed some subsampling and, after preliminary alignment and tree building, further curated to remove sequences that had multiple unresolved bases, indicating poor sequence quality, or that were much shorter than the bulk of near full-length sequences. The

resulting 93 representative sequences were aligned using MAFFT v.7.271 (Katoh and Standley 2013), using the –auto option, and ambiguous positions were removed using trimAl v.1.2 (-gt 0.3, -st 0.001) (Capella-Gutiérrez et al. 2009). In order to retain sequence FJ032651, which was one of the closest to our full-length cryptophyte sequence, we reduced the number of 5´ positions included (for the entire alignment), after observing that tree building with 'missing positions' settings was not appropriately representing branch lengths. The final alignment consisted of 1649 positions. The tree was inferred using RAxML v.8.2.9 (-m GTRGAMMAI) (Stamatakis 2015) and branch support was assessed with 1,000 non-parametric bootstrap replicates. Additionally, a Bayesian Inference (BI) phylogenetic reconstruction analysis was performed with MrBayes (Ronquist et al. 2012) using the same model of evolution with two independent runs of 2,500,000 generations with four chains each (i.e., one cold and three heated), sampling every 250 generations. After a burn-in of the first 25% of trees, posterior probabilities for node supports were computed.

To investigate TPG cryptophyte distributions, relative abundance of ASVs from this lineage was determined for Tara Oceans surface samples filtered onto a >0.8 μm size fraction (for consistency this required omitting some stations) (Ibarbalz et al. 2019). A representative sequence database for each statistically supported TPG clade was created and used as BLASTn queries against all ASVs identified as "Cryptophyta" (Ibarbalz et al. 2019) using PR^2. ASVs with 100% identity to a representative sequence for each TPG clade were identified and their relative

abundance computed within the total "Cryptophyta" amplicons; those with <10 reads in a sample were considered "not detected" for the respective station.

Statistical analyses

The R package, Hmisc (v.4.1.1) was used to calculate Pearson correlations and was visualized via the package corrplot (v.0.84). Rarefaction curves were created and diversity metrics calculated (Shannon index and species richness) using the R package, vegan (v2.5-6), for the ASVs generated from the 18S and 16S (plastid only) datasets.

Results

Molecular diversity of microbial eukaryotes in Andvord Bay and beyond

We analyzed the molecular diversity of the microbial eukaryotic community in Andvord Bay and the adjacent Gerlache Strait, the conduit to Bransfield Strait and the continental margin surrounding the South Shetland Islands (Zhou et al. 2002) (Fig. 1). V9 18S rRNA sequencing of four austral spring (Nov-Dec) and two austral fall (April) surface samples resulted in 1,256,577 total amplicons post quality control (Table S1). Rarefaction analysis indicated the depth of sequencing was saturated for all samples (Fig. S1). Average species richness was higher in austral spring (396 ± 83) than fall (226 ± 41) samples, although differences in sample numbers could

influence these results. Shannon Diversity estimates were similar between austral spring (3.49 ± 0.28) and fall (3.27 ± 0.34). Of the 885 total V9 18S ASVs, 627 were unique to spring samples, and 58 were only detected in fall samples. Some of the spring-'unique' ASVs reached up to 4% of total amplicons within a single sample, while many other seasonally unique ASVs were low in relative abundance (<0.6% of total amplicons), including all those present in fall samples.

We next analyzed the 18S ASVs against a published maximum-likelihood phylogenetic reference tree of near full length 18S rRNA gene sequences (Wideman et al. 2020). The reference tree contained sequences from all major eukaryotic groups, including data from described, cultured taxa and environmental studies. Our Antarctic 18S ASVs mapped to 242 nodes, and demonstrated that each of the major eukaryotic groups in the phylogenetic reconstruction were represented in the Fjord (Fig. 2A). The most diverse groups were the Dinophyta and Stramenopila, comprising 272 and 278 ASVs, respectively, and spanning 55 and 44 nodes (Fig. 2A). The Cryptista and Haptista (Keeling and Burki 2019), represented solely by cryptophytes and haptophytes herein, contained the most abundant single ASV. Furthermore, all cryptophyte ASVs were placed at a single node, suggesting high ASV identity (Fig. 2A).

We next delineated protistan consumer ASVs from those of primary producers (Amaral-Zettler et al. 2009) using a compilation of literature-based information. Eight eukaryotic groups were classified as photosynthetic, and 13 as frequently heterotrophic or of mixed-nutritional modes (Fig. 2B). Lineages with members that

are considered either heterotrophic or mixotrophic, defined herein as the capability for photosynthesis and phagotrophy, were included in the heterotrophic/mixed category. Because the V9 did not allow confident delineation of photosynthetic, mixotrophic and heterotropic dinoflagellates, all were placed in the heterotrophic/mixed category. Multiple lineages of the TSAR (Telonemia, Stramenopila, Alveolata, Rhizaria) supergroup assemblage belong to this category, with high relative abundances of rhizarians, ciliates, dinoflagellates, syndiniales, and MArine STramenopiles (MASTs). The latter comprised MAST-1, 2, 3, 7, 8, which were important contributors to the heterotrophic/mixed community in both austral spring and fall, based on relative amplicon abundance. Finally, picozoans, choanoflagellates, and some fungal taxa were also detected – and generally present in both seasons.

Several different eukaryotic lineages were represented among the clearly photosynthetic portion of the community. The major groups were Viridiplantae (green algae within the supergroup, Archaeplastida), Stramenopila, Cryptista, and Haptista. Among green algae were 16 mamiellophyte ASVs and 5 pyramimonad ASVs. Within the Stramenopila, 151 different diatom ASVs, 9 bolidophytes, 12 dictyochophytes, and 9 pelagophytes, collectively contributed to overall relative abundances (Fig. 2B). Cryptophytes comprised up to 71% of total photosynthetic 18S rRNA gene amplicons. In contrast to the many diatom ASVs, cryptophytes were represented by 22 18S ASVs (Fig. 2B), 20 of which had contributions <0.8% of total cryptophyte 18S ASVs (Table S2). A strong dominant was present, $_{18S}$ASV1,

representing 93% of all cryptophytes in the samples, while the second most relatively abundant taxon ($_{18S}$ASV32) contributed just 4%.

Cryptophyte abundance and biomass

Our next aim was to determine actual abundances of cryptophytes, which can be unambiguously identified via the phycoerythrin (PE) pigment they contain (Olson et al. 1989). This was important because we observed contrasting patterns of cryptophyte and diatom ASVs, with respective relative dominance by one or the other. These patterns could arise as a function of how relative abundances are generated, if one taxon increases in relative abundance, the other may have stayed the same in terms of cellular abundance, or may have decreased, but regardless will appear to decline in amplicon analyses. Using flow cytometry, we observed highest total phytoplankton abundances in austral spring, corresponding with Chl *a* results (Fig. 3, Table S3), while fall cell abundances and Chl *a* concentrations were low, with maximum measurements of ~1,303 cells mL^{-1} (flow cytometry) and 0.80 µg Chl *a* L^{-1}. The highest concentrations in spring were observed in the upper 20 m at each station, reaching a maximum of 10,204 cells mL^{-1} at IBA (Dec 2, 1.5 m; Fig. 3). Variability in cell abundance occurred throughout the fjord and adjacent waters, with a mean of 4,465 cells mL^{-1} (± 2,294, n=42) within the fjord, and 5,511 cells mL^{-1} (± 1,482, n=26) outside the fjord. Austral spring phytoplankton cell abundances were

correlated with Chl *a* at depths above 60 m (r = 0.82, *p*-value = 2.2×10^{-16}) and with increased meltwater (r = 0.54, *p*-value = 1.6×10^{-6}).

Cryptophytes were detected by cytometry at all spring stations, except SB, and had highest abundances within the fjord. These values corresponded well with microscopy-based confirmation of cryptophyte observations (r = 0.92, *p*-value = 3.2×10^{-16}), although absolute counts were systematically lower for the latter (Fig. S2). The cryptophytes observed with microscopy all corresponded to a singular morphotype. The greatest number of cryptophytes was found at IBA (2 Dec, 1.5 m), which also featured the highest number of total phytoplankton cells, >30% of which were crypotophytes (i.e., >3,000 cryptophyte cells mL^{-1}; Fig. 3). While this station had the highest cell numbers, Chl *a* ranked as the 2nd highest, with MBA 28 Nov having higher overall Chl *a*, likely reflecting differences in pigment concentration potentially associated with taxonomic composition that do not link directly to cell numbers. The abundance of cryptophytes and other photosynthetic cells enumerated by flow cytometry was correlated (r = 0.76, *p*-value = 2.2×10^{-16}). Moreover, cryptophyte abundance was negatively correlated with distance from the glacier (r = -0.28, *p*-value = 0.021) and salinity (r = -0.39, *p*-value = 1.0×10^{-3}) in the upper 60 m, while positively correlated with meltwater fraction (r = 0.36, *p*-value = 2.0×10^{-3}) and temperature (r = 0.55, *p* value = 9.1×10^{-7}).

Plastid Diversity Across the Fjord

Notable contributions of a single cryptophyte, while the other phytoplankton groups were more diverse based on analyses of 18S ASVs led us to further investigate WAP primary producers. We therefore examined phytoplankton diversity and community composition using plastid-derived 16S rRNA ASVs in a greater number of samples to search for potential gradients across the physically complex fjord. Rarefaction analyses showed that 16S amplicon saturation was reached in the 25 surface samples sequenced (Fig. S3). The number of plastid sequences recovered ranged from 5,043 to 157,703 in spring samples and 1,025 to 19,698 in fall (Table S1), out of 182,191 ± 65,051 and 91,466 ± 59,257 average amplicon sequences, respectively. The number of plastid-derived sequences and relative percentages out of all 16S amplicons (including bacteria) was correlated with Chl *a* concentrations ($r = 0.71$, p-value = 6.3×10^{-5}; Fig. 4A). Average Shannon diversity examined at the plastid ASV level was not significantly different between spring (2.39 ± 0.54) and fall (2.42 ± 0.54). Furthermore, while no relationship was detected between phytoplankton diversity and meltwater percentages, phytoplankton diversity did exhibit a weak, but positive response to distance from the glacial terminus ($r = 0.46$, p-value = 0.03).

Dynamics of Distinct Phytoplankton Lineages

To examine phylogenetic relationships between phytoplankton community members and identify specific taxa, sequences were characterized via statistical placement on reference 16S rRNA reconstructions (Vergin et al. 2013) previously

developed for analysis of cyanobacteria and eukaryotic phytoplankton (Choi et al. 2020). The results showed representatives from the same eukaryotic 'supergroups' (Keeling and Burki 2019) as 18S rRNA ASV analyses, specifically, Viridiplantae (Archaeplastida), Stramenopila, Cryptista, and Haptista (Fig. 4A). As expected, cyanobacterial sequences, which would be recovered by the primers used, were not detected. The Viridiplantae were represented by two main lineages, pyramimonads and mamiellophytes. Four eukaryotic phytoplankton groups were identified within the Stramenopila, specifically, diatoms, pelagophytes, dictyochophytes, and bolidophytes. Finally, the Cryptista and Haptista were represented only by cryptophytes and haptophytes, respectively. Among the general trends observed was, again, contrasting dominance by diatoms or cryptophytes, particularly in the austral spring (Fig. 4A).

Among the Viridiplantae, pyramimonad ASVs had higher relative amplicon abundances, except on 15 April at Gerlache (Fig. 4A). Their relative amplicon abundances were negatively correlated with temperature, and positively correlated with increased phosphate concentrations (Fig. 4B). The four pyramimonad ASVs were assigned to a single "Last common Ancestor (LCA)" of Prasinophyte Class I node; more precise classification was not possible because 16S rRNA gene sequences from characterized taxa are not available. However, 9 Mamiellophyceae ASVs were detected, three belonging to the picoeukaryote *Micromonas polaris* and 6 to *Bathycoccus*. In both seasons, mamiellophytes contributed <10% of total plastid amplicons, reaching a maximum contribution of 6.8% in the austral spring and 7.4%

in fall, and similarly low numbers of *Micromonas* cells were detected by qPCR (Fig. S4). Spring and fall Mamiellophyceae relative abundances were positively correlated with distance from the glacier ($r = 0.53$, p value = 6.0×10^{-3}) (Fig. 4B).

Diatoms were the most highly represented stramenopiles (Fig. 4A) and also showed the greatest diversity of the phytoplankton groups detected, with a total of 66 16S ASVs. Several diatom ASVs had high nucleotide identity to known species, and some were detected only at specific stations, while others presented a more consistent background presence (Fig. 4C). $_{16S}$ASV8 was one of the most relatively abundant diatoms in austral spring in the outer regions of the fjord and the Gerlache Strait, and also in trace amounts (<0.2% of diatom ASV abundances) at the inner most stations during both seasons. It had 99.3% nt similarity to *Chaetoceros calcitrans*, a species so far unreported in Antarctic waters, moreover *Chaetoceros* cells were seen by microscopy although further delineation was not possible based on morphological characteristics. $_{16S}$ASV12 peaked in relative abundance at SB (4 April) in the austral fall (Fig. 4A). This diatom was likely *Odontella weissflogii*, which was observed by light microscopy (Round et al. 1990) (Fig. S5), but has no 16S sequence in GenBank nr. *O. weissflogii* is phylogenetically close to the odontelloid *Triceratium dubium* based on other gene markers (Theriot et al. 2010), and with which $_{16S}$ASV12 had 98.2% nt identity. Unfortunately, 16S rRNA gene sequences for many other diatoms are unavailable as well, including the Thalassiosirales. For example, within the fjord during austral spring the most abundant diatom ASV was $_{16S}$ASV41 (Fig. 4C). Based on top BLASTn results, $_{16S}$ASV41, along with $_{16S}$ASV17, $_{16S}$ASV25, $_{16S}$ASV74,

$_{16S}$ASV124, and $_{16S}$ASV126, are all likely members of the Thalassiosirales. Many Thalassiosirales were observed by microscopy, with species of *Shionodiscus* and *Thalassiosira* being notable (Fig. S5). In austral fall, $_{16S}$ASV48, which had 100% nt identity to *Fragilariopsis cylindrus*, was present at all but Stations SB (4 April) and OBB (19 April) and had among the highest relative abundances for diatoms. Diatoms as a whole showed a significant positive correlation with distance from the glacier and depth of the mixed layer, defined here as the region of the upper water column where surface wind-driven mixing results in homogenous temperature and other physical oceanographic characteristics (Fig. 4B).

The other major stramenopile taxa groups observed were bolidophytes, dictyochophytes, and pelagophytes, which at times comprised considerable percentages of plastid amplicons. For example, bolidophytes sometimes exceeded diatoms and consisted of 9 ASVs, most of which lacked high identity matches in GenBank, except the second most abundant ($_{16S}$ASV58) which had 99.6% identity to *Triparma laevis*, reported in the WAP previously based on 18S rRNA genes (Ichinomiya et al. 2016), and originally found in the North Pacific (Tajima et al. 2016). Dictyochophytes consisted of 13 ASVs, with six being phylogenetically placed with *Helicopedinella*, three as sister to the *Pseudopedinella* and CCMP2098 clades sensu Choi et al. (Choi et al. 2020), and one each with *Florenciella parvula*, *Florenciella* clade II and *Dictyocha*. Pelagophytes showed the least 16S-based diversity of all the stramenopiles, with three ASVs, placed with environmental clades PEC-I, PEC-V, and PEC-IX, respectively (Choi et al. 2020). Pelagophytes and

dictyochophytes each exhibited a significant negative correlation between relative amplicon abundances and temperature (Fig. 4B) and the former were negatively correlated with day length. Additionally, pelagophyte, dictyochophyte, and bolidophyte amplicon abundances were each positively correlated with phosphate concentrations.

Cryptophytes exhibited higher relative abundances than haptophytes, and other groups of phytoplankton discussed thus far, apart from the diatoms taken as a whole (Fig. 4A). Haptophytes were dominated by $_{16S}$ASV68 and $_{16S}$ASV151, the first having 100% nt identity to *Phaeocystis antarctica* and the second 98% identity to *Chrysochromulina parva* (Table S2). Both had relatively low contributions, with the maximum being 6% of total plastid amplicons at IBB during the fall (Fig. 4A). Moreover, haptophyte relative abundance was negatively correlated with temperature and day length (Fig. 4B). For cryptophytes, on average 85.2 ± 10.0% (n=25) of their 16S rRNA amplicons were formed by two sequence variants ($_{16S}$ASV2 and $_{16S}$ASV4) at each station that differed from each other by two base pairs, while 7 other ASVs formed the rest of the cryptophyte community. An "uncultured marine cryptophyte clone" sequence from near the South Shetland Islands (59.3792 S 55.7742 W and ~700 km north of our sampling region) had the closest (99.6%) nt identity to our dominant cryptophyte 16S ASV, however, the closest available 16S rRNA gene sequence from a cultured species was *Teleaulax amphioxeia* at 97.1% nt identity. In austral spring, cryptophytes comprised up to 93% and in fall, when Chl *a* levels and phytoplankton abundances were low, cryptophytes formed up to 45% of total plastid

amplicons. For both seasons, cryptophyte relative abundance was negatively correlated with distance from the glacier (r = -0.42, *p*-value = 0.038) (Fig. 4B).

A Novel Cryptophyte Dominates

The studies performed thus far established a covariance between diatoms, made up of many known taxa, and a dominant but seemingly unknown cryptophyte ($_{16S}$ASV2). Like cryptophyte $_{16S}$ASV2, the corresponding dominant $_{18S}$ASV1 did not appear to come from a cultured cryptophyte. $_{18S}$ASV1 did have 100% nt identity to an Antarctic sequence from Ace Lake (HQ111513), which came from a cryptophyte that was briefly in culture but then lost (van den Hoff et al. 2020). The length of amplicons generated herein precluded phylogenetic analysis – apart from statistical placement of ASVs on curated reference trees. We therefore constructed 18S rRNA gene clone libraries, using DNA from a Gerlache Strait sample, and recovered full-length sequences that matched $_{18S}$ASV1 at 100% nt identity.

The full length sequence of the dominant cryptophyte in our study branched within the morphologically variable cryptophyte Teleaulax/Plagioselmis/Geminigera (TPG) lineage (Hoef-Emden 2008) (Fig. 5A). Although the innermost nodes (backbone) of the topology were generally unsupported, we were able to identify 9 clades within the TPG lineage moderately (>70/0.7) to fully supported by bootstrap values and posterior probabilities. All Antarctic sequences belonged to two statistically supported clades. Our WAP sequence formed a clade (named herein

Clade VII) with the Ace Lake sequence, coming from the other side of Antarctica, and a Southern Ocean sequence from near the South Shetland Islands (FJ032651; Fig. 5B). Sequences from the cultured Antarctic cryptophyte, *Geminigera cryophila*, branched in Clade (IV), as did an one from Bayly Bay (HQ111512), near Ace Lake (van den Hoff et al. 2020). Five other clades contained cultured taxa (II, III, V, VIII, IX). Three were devoid of cultured representatives, Clade VII from our study, Clade I containing Sargasso Sea environmental clones, and Clade VI, containing East China Sea environmental clones. Additionally, while sequence identities were often quite high or overlapping (Fig. S6), each clade had consistent, unique synapomorphies that differentiated them clearly from other clades (Fig. S7). Newly identified Clade VII had a total of 8 unique nucleotide polymorphisms, relative to Clade IV and VI (across the entire gene sequence).

Based on the phylogenetic reconstruction, Clade VII appeared to represent a novel species within the TPG lineage. We then paired the phylogenetic results with light and scanning electron microscopy studies. All cryptophytes observed in our samples (n = 128), with the exception of 3 samples from SB (20 Dec), corresponded to a singular morphotype, which we therefore considered to be the WAP cryptophyte identified via molecular analysis and flow cytometry. This morphotype also matched descriptions of WAP cryptophytes observed by (Mascioni et al. 2019). Light microscopy measurements indicated it was 11.3 ± 2.1 μm length and 5.1 ± 0.6 μm width (mean ± standard deviation of 41 individuals from both austral spring and fall). Morphologically, the novel cryptophyte differs from *G. cryophila*, which has a

rounded or ovoid shape, a featureless periplast and a characteristic warty aspect due to the accumulation of lipid droplets in the peripheral cytoplasm (Taylor and Lee 1971), as well as from *Teleaulax* species which have a teardrop shape and featureless periplast (Laza-Martínez et al. 2012). The novel cryptophyte had a teardrop shape, warty appearance, and hexagonal to rectangular plates that comprised the periplast (Fig. 5C). Further, it had a conical tail with no plates, two flagella, slightly shorter than the body emerging sideways from a ventral furrow. These characteristics aligned with physical traits attributed to *Plagioselmis* genus, yet the specifics of its morphological characters, did not fit to any currently described taxon (Mascioni et al. 2019). These results led us to conclude that Clade VII captures members of an unknown cryptophyte species, that hereafter we refer to as Clade VII, to avoid naming prior to reevaluation of the entire TPG lineage.

Cryptophyte Clade Distributions Across the Oceans Expose Niche Differentiation

While the overall topology of the 18S rRNA gene phylogenetic reconstruction was not well supported, many clades had support as well as clade-specific synapomorphies, especially in the 18S V9 region. Like most clades, Clades III and V differentiated from all other clades within the V9, but in this case alone did not resolve from each other (Fig. S7). Each clade (Fig. 5A) featured 100% nt identity within itself for the V9 region, with the exception of environmental Clade VI (East China Sea) and the freshwater clade (IX), where there were three variants per clade

that differed by 1 nt each (Fig. S7). We exploited the clade-identifying synapomorphies to examine TPG distributions in Tara Oceans V9 18S data (Ibarbalz et al. 2019). The results show that cryptophytes are broadly distributed from pole to pole, as supported by prior studies on different ocean regions (Hoef-Emden 2008). Furthermore, our analyses showed that the TPG are the dominant marine cryptophyte lineage throughout the global ocean (Fig. 6A).

Our more specific synapomorphy-based analyses provided unprecedented insight into the molecular diversity of TPG cryptophytes and environments inhabited. Members of Clade II (*Teleaulax amphioxeia*) and Clade III/V (*Plagioselmis prolonga*/*Teleaulax gracilis*) were present in temperate and tropical environments, as well as high Arctic latitudes, but were not detected in the WAP and Southern Ocean (Fig. 6E, G). Moreover, although they were frequently detected together, this was not always the case. The newly discovered Clade VII is limited to high southern latitudes, and dominates in the region of Antarctica (WAP) sampled by Tara (Fig. 6B). Members of Clade IV (*Geminigera cryophila*) were found at just one WAP station, although present in multiple samples of our study. At the Tara Oceans WAP sites, Clade IV had much lower contributions to TPG relative abundance than Clade VII.

Finally, we estimated contributions of this cryptophyte to total phytoplankton biomass in Andvord Bay and the Gerlache Strait using the established cell size, conversion factors, and flow cytometry data and found that the average cryptophyte contribution in the top 20 m was 29 ± 13% in austral spring (n=46). We did not perform these estimates for fall because cell counts were less statistically robust (due

to low phytoplankton numbers). It should also be noted that our estimates of total biomass are only rough calculations, due to lack of precision in assigning biomass values to the non-cryptophyte cell populations, which had a range of forward angle light scatter values, indicating a range of sizes.

Discussion

Studies examining protistan diversity using high-throughput 18S rRNA or 16S rRNA (for phytoplankton only) sequencing have generally been performed using two different approaches, survey sampling, with individual samples from multiple ocean sites (Ibarbalz et al. 2019), or time-series of multi-expedition regional sampling. The latter provides resolution and facilitates connection of diversity and community dynamics to seasonal environmental variations (Needham et al. 2017; Bolaños et al. 2020; Choi et al. 2020). With respect to polar regions, there is survey coverage of the Arctic (Lovejoy et al. 2006; Bachy et al. 2011; Ibarbalz et al. 2019), and information is more limited on Antarctica (Wolf et al. 2015; Luo et al. 2016; Zoccarato et al. 2016), reflecting the challenges in accessing Southern Ocean and Antarctic coastal zones. The overall physical dynamics of Arctic and Antarctic seas differ, as do the fjords. Arctic fjords receive strong annual meltwater inputs, providing cold freshwater and terrigenous sediments originating from glaciers, which leads to a significant down-fjord gradient in both pelagic and benthic habitats (Wlodarska-Kowalczuk et al. 2005). In Antarctic fjords, meltwater influences are more mild, such that nutrients are

still injected, and some stratification occurs, but without the impacts of sediment loading and high turbidity associated with a strong meltwater inputs. Meltwater extent is highly sensitive to climate change, a factor that motivated this research and our study design for investigating extant communities in Antarctic fjords.

Protistan Communities in Polar Environments

The major protistan groups seen herein (Fig. 2) are also present in Arctic fjords (Kubiszyn et al. 2014; Zhang et al. 2019) and a number have been reported in WAP studies. Our study location was ~60 km from a region sampled near Anvers Island (Luria et al. 2014), a distinct environment from fjords that has similarities to the Gerlache Strait. Our location was intermediate to the coastal South Shetland Islands (Luo et al. 2016; Abele et al. 2017; Trefaut et al. 2021) and Marguerite Bay (Rozema et al. 2017), a true bay (not a fjord). The sequence analysis resolution for the two studies of Southern Ocean waters near the South Shetland Islands was most similar to ours, although they used the V4 18S region, not V9. Overarching broad similarities were apparent in the protists seen as one moves from Southern Ocean waters into fjord environments. The more 'mysterious' and small-sized heterotrophic microbial eukaryotes, also present in Arctic fjords and seas (Bachy et al. 2011; Zhang et al. 2019), are picozoa, telonemids, and MASTs. While the ecosystem roles of uncultured MAST lineages, and even picozoa are unclear, several are known predatory heterotrophic nanoflagellates, including MAST-1, MAST-2, MAST-7, and

MAST-8 (Lin et al. 2012; Massana et al. 2014). Community composition herein was similar to the most comparable South Shetland Island study (Luo et al. 2016), however overall diversity in the latter was less than in our study, likely due to their use of 97% nt identity OTUs, which would minimize diversity estimates.

The results suggest that dinoflagellates are important predators based on 18S ASV relative abundances (Fig. 2B). They are also highly diverse, and confirm early hypotheses based on more limited 18S rRNA gene sequencing, that there are small, still undescribed dinoflagellates in the Southern Ocean (López-García et al. 2002; Gast et al. 2006). They included ASVs affiliated with potentially kleptoplastidic gymnodiniales, that acquire and use plastids from other phytoplankton, in addition to operating as heterotrophs through seasonal periods of complete darkness (Bachy et al. 2011; Torstensson et al. 2015).

Phytoplanktonic 18S and 16S ASVs (Fig. 2 and Fig. 4) also corresponded generally, as well as with other WAP and Arctic studies. The higher taxonomic resolution of our analysis approach for 16S ASVs, and the fact that gene copy numbers are more constrained than for the 18S rRNA gene (Needham and Fuhrman 2016; Choi et al. 2020), combined with the greater number of samples sequenced, provided new insights into dynamics and diversity of these important phytoplankton communities, as well as linkages between specific taxa and environmental factors.

The Molecular Diversity of Antarctica's 'Small Flagellated' Phytoplankton

HPLC and microscopy studies have highlighted the important contributions of "small flagellates" (also termed "mixed flagellates") in coastal WAP environments, including in Andvord Bay (Pan et al. 2020). This group included multiple phytoplankton taxa (i.e., pigmented), some identifiable by microscopy and others not. Using our marker gene studies, we surmise that the "mixed flagellates" of Andvord Bay (Pan et al. 2020) include prasinophytes, bolidophytes, pelagophytes, dictyochophytes, and haptophytes. Many of these small photosynthetic flagellates may have been mis-assigned by HPLC as being diatoms or assigned to a broad category of "other phytoplankton" in some WAP studies. We surmise that *Phaeocystis antarctica* is important in these "other phytoplankton" or "mixed flagellates" groups, as it is seen in both the 18S and 16S analyses, and has been elsewhere in the WAP during austral spring (Selz et al. 2018). *P. antarctica* has a secondary role important in plankton interactions, as a prey and plastid source for some kleptoplastidic Antarctic dinoflagellates (Gast et al. 2007) observed in our data. Additionally, both bolidophytes and dictyochophytes have pigment overlaps with diatoms, such that HPLC analyses would classify them as diatoms (Choi et al. 2020). This misidentification has important biogeochemical implications, as the bolidophytes lack a diatom-like frustule, having either siliceous plates or being unsilicified altogether (Guillou et al. 1999; Ichinomiya et al. 2016), like dictyochophytes. Importantly, both haptophytes and dictyochophytes contain some predatory phototosynthetic mixotrophs, which alters the environmental constraints

and selective processes that act on them as compared to purely photosynthetic taxa (Worden et al. 2015).

The prasinophytes observed would also fall into the "mixed flagellates" or "other phytoplankton" categories of prior Antarctic studies. The dominant pyramimonad 18S ASV had 100% nt identity to *Pyramimonas gelidicola*, a species isolated from Antarctic waters (van den Hoff et al. 1989; Bell and Laybourn-Parry 2003). Both our 18S and 16S datasets contained *M. polaris* and *Bathycoccus prasinos* ASVs. These two have been reported in the South Shetland Islands, Marguerite Bay (Rozema et al. 2017; Trefaut et al. 2021), the Amundsen Sea, and circumpolar waters surrounding Antarctica (Delmont et al. 2015; Simmons et al. 2015). *M. polaris*, thought endemic to Arctic waters (Lovejoy et al. 2007), prior to high throughput sequencing, was detected in Southern Ocean metagenomes through highly conserved introns and flanking sequence (Simmons et al. 2015). Finally, a study of Fildes Bay (South Shetland Islands), where *Micromonas* comprised up to 46.7% of the total 18S amplicons, reported a possible new variant (in addition to *M. polaris*), based on V4 18S sequence data (Trefaut et al. 2021). This differs with our observation of two *M. polaris* 18S V9 ASVs (100 and 99.4% nt identity to the cultured strain), however the V4 is known to resolve some prasinophyte groups better than V9 (Monier et al. 2016). Additionally, we found *M. polaris* consistently present, but at low relative abundances in both 16S and 18S amplicon data, as confirmed by qPCR (Fig. S4), and contrasting with high relative abundances in Fildes Bay. The reasons for this disparity remain unclear, and more systematic sampling of the WAP will be needed to identify

environmental drivers behind proliferation of different taxa in specific Antarctic regions.

Phytoplankton community structure and links to environmental parameters

A factor thought to contribute to productivity and aggregations of large mammals in Andvord Bay, is that it is "dynamically quiet", lacking the strong wind forcing of the ambient ocean, and facilitating some level of stratification. We observed a significant positive correlation between the depth of the mixed layer and relative diatom amplicon abundance, an indication of requirements for higher nutrient concentrations than available when the water column is more strongly stratified (Fig. 3B). Otherwise, the macronutrient that correlated most strongly with multiple phytoplankton groups was phosphate. Bolidophytes, dictyochophytes, pelagophytes, and pyramimonads all exhibited positive correlations with increased phosphate concentrations (Fig. 4B). Results from an analysis of "mixed flagellates" in Andvord Bay (Pan et al. 2020) contrasted with this finding, with no relationship seen. This difference likely arises from merging of taxonomic groupings identified herein, including haptophytes, as "mixed flagellates", although some, but not all showed a positive relationship with phosphate in our study (i.e., haptophytes did not). These results highlight the importance of examining phytoplankton taxa at a high taxonomic resolution that connects with niche partitioning and evolutionary distance.

While macronutrient availability is an important factor shaping phytoplankton communities in many regions of the global ocean, in polar fjords, glacial dynamics are also extremely important. Distance from the glacier exhibited the strongest relationships with relative amplicon abundances of different phytoplankton groups. Specifically, distance from the glacier terminus was positively related to relative mamiellophyte and diatom amplicon abundances (Fig. 4B). In contrast, cryptophyte amplicon relative abundances were negatively correlated with distance from the glacier, and their cell abundances were also generally highest in the inner part of the fjord (Fig. 4). The significant positive correlations with cryptophyte cell abundance and increased temperature as well as glacial meltwater, and with decreased salinity all point to cryptophyte proliferation being enhanced due to meltwater inputs. This suggests that the glacier plays an important role in shaping the phytoplankton community in Andvord Bay, despite its characterization as a "cold" fjord with relatively little influence from glacial meltwater in terms of physical oceanography (Lundesgaard et al. 2020). Furthermore, the multiple correlations observed for the Clade VII cryptophyte highlight the fact that statistics were stronger because we were examining the dynamics of a single organism in relation to environmental factors, unlike the other phytoplankton groups, which were more diverse and could not be separated in flow cytometry data and only sometimes via relative abundances, likely limiting detection of discrete responses to environmental variables.

As discussed, diatoms are important stramenopiles in Andvord Bay, and elsewhere in both Arctic fjords and Antarctic environments. However, diatoms are

diverse (Malviya et al. 2016) and our data shows that patterns arise at the 16S ASV level that are obscured at the group level. For example, *F. cylindrus* ($_{16S}$ASV48), a well-recognized member of the WAP phytoplankton community (Mendes et al. 2018), was present at low relative abundance at all spring stations within the fjord and Gerlache Strait, but dominated diatom amplicons at SB (20 Dec) (Fig. 4C). In fall the opposite occurred, where this taxon occurred at higher relative amplicon abundances within the fjord and strait. One explanation of taxon-specific trends revealed by our molecular dataset could be that there is a seasonal pattern of an initial austral spring presence of *F. cylindrus* out on the shelf, with a transition towards the nearshore environment, including within the fjords, during the fall. Other diatom ASVs showed differing patterns.

A unique feature of polar fjords is that there are occasional katabatic wind events, driven by strong winds originating on the continental ice sheet (Lundesgaard et al. 2020). Our sampling included three days at Sill 3 that spanned an austral spring katabatic wind event (Lundesgaard et al. 2020). Prior to the event (29 Nov and 8 Dec), cryptophyte amplicons dominated plastid sequences (Fig. 4A) and cryptophyte cells were abundant. Post-event (17 Dec) relative cryptophyte amplicon frequency was lower, and diatoms dominated amplicons. The strong down-fjord winds induced wind-activated mixing that replenished photic zone nutrients (Lundesgaard et al. 2020), with nitrate higher towards the end of the event (13 Dec) and drawdown observed by 17 Dec (Ekern 2017). Thus, katabatic wind events appear to have a role in resetting phytoplankton communities, and the relative contributions of different

lineages in Antarctic fjords, adding a consideration beyond the oft discussed influences of ice retreat and the degree of temperature-induced surface water stratification.

Molecular insights into the novel Clade VII WAP cryptophyte and overall TPG Lineage

While cryptophytes have been observed as one of the key phytoplankton taxa in the WAP by a multitude of studies, almost nothing was known about their diversity. *Geminigera cryophila* is the only Antarctic cryptophyte in culture, and hence the only one where molecular sequence data has been connected with morphological information. Our findings show a unique lineage defined here as TPG Clade VII which is phylogenetically and morphologically distinct from *G. cryophila* (Clade IV; Fig. 5).

The phylogenetic analyses call attention to the current disconnect between evolutionary analyses and the morphologically-defined genera within the TPG lineage (Fig. 5A). This may partially stem from life history-dependent dimorphism, as observed in the non-TPG genus, *Cryptomonas* (Laza-Martínez et al. 2012). Challenges for the TPG lineage include limited representation of some clades in culture, resolution of the marker sequences used, possible erroneous naming in GenBank, and low quality marker gene sequences for some reference taxa. *Plagioselmis prolonga* and *Teleaulax amphioxeia* have been proposed to be the same

genus and species, with *P. prolonga* representing the haploid state and *T. amphioxeia* the diploid state, primarily based on morphological changes within cultures (Altenburger et al. 2020). Here, both phylogenetic analyses and conserved V9 polymorphisms distinguished *Plagioselmis* and *Teleaulax*, but this does not preclude their being from the same genus (Laza-Martínez et al. 2012).

Additionally challenges have come from environmental studies. For example, the Ace Lake (HQ111513, Clade VII) and Bayly Bay (HQ111512, Clade IV) cryptophytes are proposed cryptomorphs of *G. cryophila* (van den Hoff et al. 2020), although it is unclear if this was observed for both the Ace Lake and the Bayly Bay culture (both now lost). The suggested dimorphism, is a warty, ovoid cell type (campylomorph form) seen in aged cell cultures, for which microscopy images are lacking. Given their placement in separate bootstrap supported clades, as well as conserved polymorphisms between the V9 sequences of Clade IV and Clade VII (Fig. 5, Fig. S7), and the fact that this phenomenon not been observed in cultured *G. cryophila* strains, possible cryptomorphy seems unlikely, at least between the Ace Lake strain and *G. cryophila*. An overall reevaluation of genus level differentiation within the TPG lineage would nevertheless be exceptionally helpful, and would benefit from targeted PCR and genome sequencing efforts paired with cell imaging, if not cultures. Finally, of the 9 TPG Clades we delineated using phylogenetic methods, apart from Clade IX, containing freshwater isolate *Plagioselmis nannoplanctica* and otherwise only freshwater sequences, all were comprised of marine environmental sequences or marine isolates.

Linking High Resolution Taxonomy to Local and Global Spatial Community Structure

Our molecular and detailed microscopic analysis identified a key phytoplankter in a taxonomically meaningful way and established important environmental linkages. Identification of evolutionarily conserved polymorphisms between clades within the V9 18S region underpinned our ability to place findings in a broader context using Tara Oceans data. Globally, TPG group members appear to dominate over other cryptophytes (Fig. 6A). Some cryptophyte clades were broadly distributed in tropical and temperate waters (Clade I) (Fig. 6H) or also extended into subpolar Arctic waters as high as 69° North (Clades VI, VIII) (Fig. 6C, D). Clade VII, discovered and described herein, had much higher relative abundances than *G. cryophila* during our studies, dominating the cryptophyte community during both WAP expeditions, as well as dominating Tara Oceans cryptophyte data from the Southern Ocean (Fig. 5, 6). We found no evidence for cryptophyte Clade VII being in Arctic fjords, nor generally in Arctic waters, based on Tara Oceans amplicon data, and the cultured *G. cryophila* (Clade VI) appeared to be absent as well (Ibarbalz et al. 2019) (Fig. 6B, F). In contrast, the cryptophytes seen in the Arctic are also present in tropical and temperate waters extending to 60° South (Clades II and III/V), but not detected in Antarctic waters (Fig. 6E, G).

It has been suggested that endemism is common to protists from both poles, based on pyrosequencing (Wolf et al. 2015), but as sequencing depth has increased,

some of these conclusions have been invalidated. For example, for *Micromonas polaris*, metagenomic and genome analyses led to the proposal that it is transported from the Arctic to Antarctic, as a cyst or other form, via the North Atlantic Deep Water (NADW), where it is also detected (Simmons et al. 2015). The NADW, as a mechanism of ecosystem connectivity, may also be relevant for other bipolar species that have a dormant phase. The unique oceanographic factors that contribute to what appears to be isolation of plankton in Antarctic waters warrant further investigation. The dominance of the Clade VII cryptophyte during our two sampling periods in Andvord Bay above all other cryptophytes, suggests extreme niche differentiation, which may play a role in the endemism of Clade VII cryptophytes and in the phylogeography of cryptophytes as a whole.

Even if highly niche specialized, we also observed that when the absolute abundance of non-cryptophyte phytoplankton increased, so also did the cryptophyte in spring months ($r=0.75$, p-value$=2.2\times10^{-16}$) as did the relative proportion of cryptophytes ($r=0.93$, p-value$=3.8\times10^{-31}$). This indicates that the phytoplankton community collectively responded positively to certain environmental cues, but that the cryptophyte response (in terms of cell numbers) was more pronounced. The maximum cryptophyte abundances observed here were similar to another WAP study, being between 3,000 to 4,000 cells mL^{-1} (Biggs et al. 2019). Our FCM counts were also within range of those for blooms reported in microscopy studies (compiled in (Mascioni et al. 2019)), although the abundance observed in blooms at Anvers Island (Garibotti et al. 2005) were much higher, reaching 15.7×10^{3} cells mL^{-1}.

What can these findings tell us about the oscillation between cryptophytes and diatoms observed in our amplicon data, and reported previously in the Antarctic using other methods (Schofield et al. 2017)? Although not all studies employed statistical analyses, and none identified the cryptophyte(s) or diatoms involved, speculations regarding Antarctica's diatom-cryptophyte dichotomy include cryptophyte preference for lowered salinity (Mendes et al. 2013), and/or increased temperature (Mendes et al. 2018; Pan et al. 2020), or (conflictingly) increases in response to low salinity, colder waters (Schofield et al. 2017). Cryptophyte blooms have also been observed in connection with increased stratification (and a subsequent decrease in mixed layer depth) (Mendes et al. 2013; Mendes et al. 2018). We find that 1.) the communities contrast sharply in cryptophytes having a single taxonomic dominant, while diatoms are diverse and the relative abundances of specific diatoms shift over time; 2.) the cryptophyte present appears to be specialized to Antarctic, whereas the diatoms present seem to have much larger distributions; 3.) Clade VII cryptophyte abundance is strongly correlated with increased ice melt, temperature, stratification, and lower salinity, all factors that are predicted to be enhanced by climate change; the diatoms showed a significant relationship with seasonally reduced stratification, potentially reflecting greater nutrient requirements. These results were reflected in both seasonal and short-time scale events, the latter being exemplified by the katabatic wind event captured in our study.

Finally, why are the overall phytoplankton group patterns so different in Antarctic fjords versus the Arctic? In the Arctic, diatoms appear to be the main bloom

forming taxa, with blooms typically seen in the spring in either the inner (Calleja et al. 2017) or outer fjord (Piwosz et al. 2009). While cryptophytes have also been reported in many Arctic fjords, other algal groups, including the prymnesiophytes and photosynthetic dinoflagellates, have been noted as more important components of the phytoplankton community when diatoms are lower in abundance (Piquet et al. 2014; Calleja et al. 2017). Our results point to dominance of the Clade VII cryptophyte in the WAP and fjords within, and suggest a great importance to blooms based on HPLC and microscopy studies in the same region (Mascioni et al. 2019; Pan et al. 2020). One potential consideration is that the cryptophytes seen in the Arctic, appear to be generalists, also distributed throughout temperate and tropical waters. In contrast, Clade VII so far has not been observed elsewhere in the ocean, suggesting its proliferation is connected to a specialization strategy that differs from the generalist strategy seemingly pertinent to other TPG clades, seen in the Arctic, as well as diatoms that are bipolar or even more globally distributed.

Conclusions

Knowledge of the extant diversity in Antarctica and its 'hot-spot' fjords is a priority. We elucidated protistan diversity by combining results from two expeditions designed to examine seasonal differences in an Antarctic fjord noted for high productivity and diverse megafauna. Our studies reveal variations in how diversity of different phytoplankton lineages manifest in this important polar region, and

connections to climate-sensitive seasonal perturbations. The novel cryptophyte discovered was the dominant cryptophyte during our study and in Tara Oceans survey data from the Southern Ocean, almost to the exclusion of other cryptophytes, and appears to be endemic to Antarctica. Unfortunately, this cryptophyte has not been cultured, and its food web linkages are unclear, unlike many diatoms, which are known components of fisheries food webs. Importantly, the abundance of the Clade VII cryptophyte was positively correlated with increased temperatures and glacial meltwater inputs, along with decreased salinity. While WAP fjords are considered to be "cold", with weaker meltwater influence, relative to Arctic fjords, they are not immune to the effects of climate induced warming. The dominance of a single cryptophyte over diverse diatoms when glacial freshwater inputs were maximal, may well reflect future trends, with as yet unknown impacts on the larger food web structure.

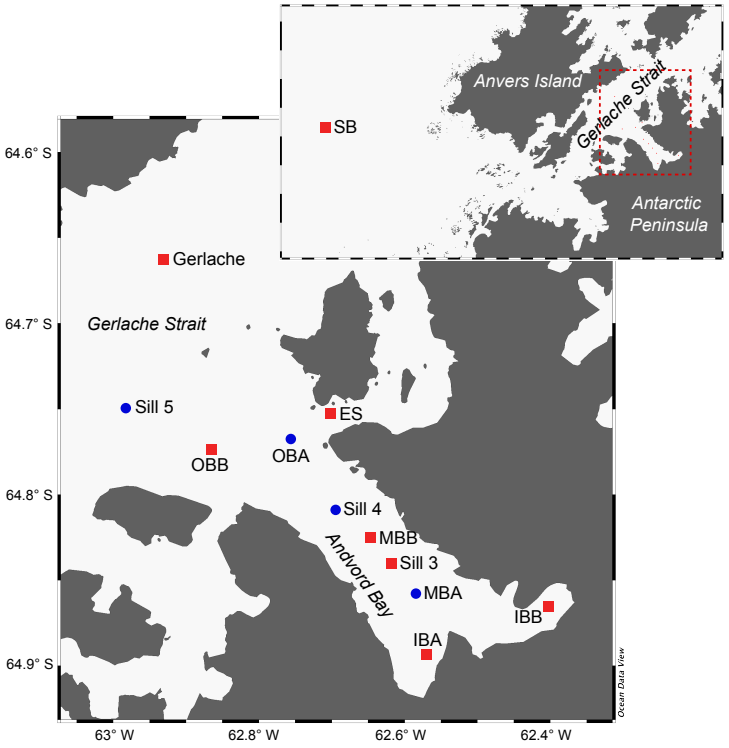


Fig 2.1 Sampling stations in the present study, including Station B on the Antarctic continental shelf (SB) and stations within and adjacent to Andvord Bay in the Gerlache Strait. The map in the foreground shows a greater view of the west coast of the Antarctic Peninsula and Anvers Island, with the red box outlining the map of Andvord Bay shown in the background. Red squares indicate stations that were sampled and analyzed during both the austral spring and fall, while blue circles represent stations that were only sampled and analyzed during the austral spring.

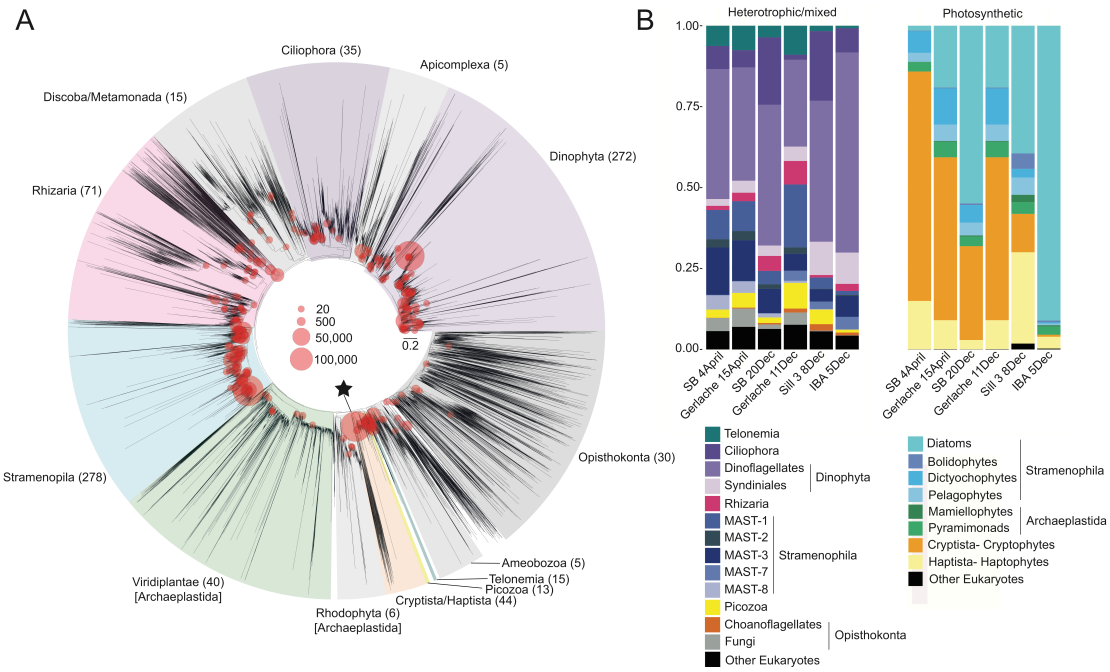


Fig 2.2 Eukaryotic diversity and distributions at Andvord Bay and Gerelache Strait Stations. **a** Maximum-likelihood phylogeny based on 18S rRNA gene reference sequences (near full-length), onto which combined 18S V9 ASVs from the austral spring and fall samples sequenced were mapped. The size of the red circle is proportional to the number of sequences mapped to each node. Major eukaryotic groups are highlighted, and the parentheses next to each group label list the number of ASVs represented. The black star indicates the node representing cryptophytes. **b** Relative 18S rRNA gene amplicon frequency for this subset of samples, based on Qiime classification, with a subsequent partitioning based on a literature review of nutritional types. The "Heterotrophic/mixed category" includes strict heterotrophs, as well as sequences classified into broad taxonomic groups that include both heterotrophic and mixotrophic (capable of phagotrophic nutrition and photosynthesis) representatives, as well as groups where the nutritional mode could not be identified

(i.e., dinoflagellates). Only groups represented at above 10% relative contributions in one or more samples are shown. Note, MAST-12 was also present but fall under the "Other Eukaryotes" classification in fig. 2b, along with other ASVs in taxonomic groups of low relative abundance (<10% in all samples) and ASVs classified as "unidentified" based on Qiime2 (Table S2). The "Other Eukaryotes" category under photosynthetic eukaryotes comprised ASVs in taxonomic groups that were either present at low relative abundance (<10% in all samples) or lacked a taxonomic classification beyond "Archaeplastida". All metazoans were excluded from the analysis. All samples represented were also sequenced for 16S rRNA amplicons (see Fig 2.4).

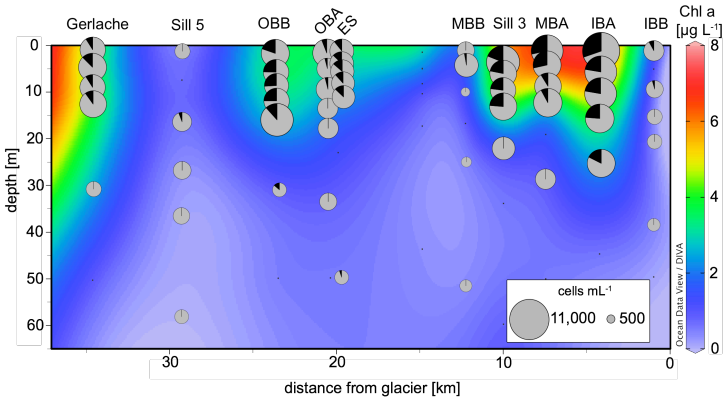


Fig 2.3 Flow cytometry-based phytoplankton abundances across austral spring depth profiles. Cross section of the upper 65 m of the spring transect showing Chl *a* concentration in the background with cell abundance overlain (pie charts). Black represents the proportion of total cells that were identified as cryptophytes, while grey represents all other photosynthetic cells. The size of each pie chart is scaled to the total number of cells enumerated.

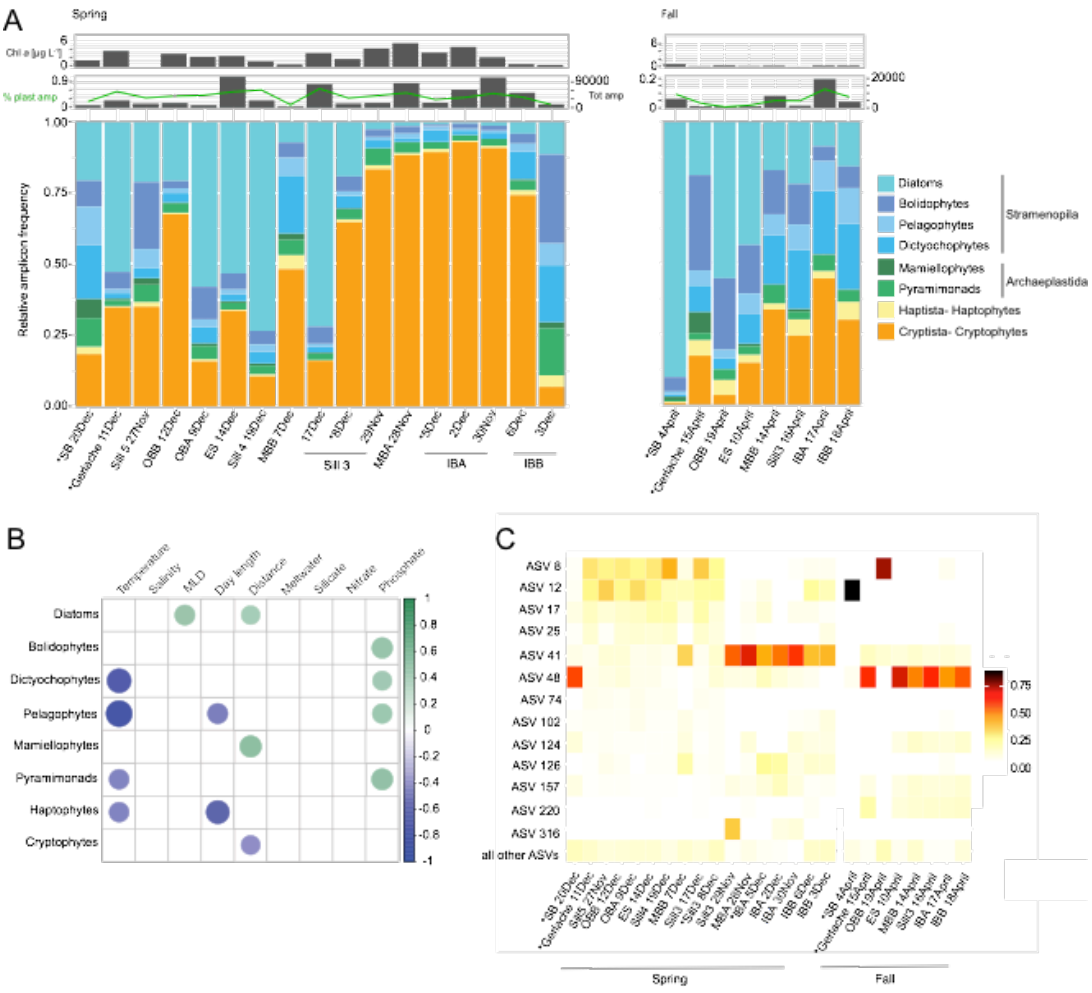


Fig 2.4 Phytoplankton diversity, distribution, and environmental relationships. **a** Relative 16S rRNA gene amplicon frequency for samples taken during austral spring and fall based on PhyloAssigner classification. The topmost grey bar plot represents Chl *a* concentrations corresponding to each DNA sample. The lower grey bar plot shows the total number of plastid amplicons sequenced with the percentage of total amplicons that were identified as plastid sequences overlaid as a light green line. *, indicates samples that were also analyzed by 18S rRNA amplicon sequencing (see Fig. 2.2). **b** Heatmap showing the significant Pearson's correlations (p value ≤ 0.05)

between environmental factors and relative amplicon frequency of phytoplankton taxa. Pairwise correlations without an assigned colored dot represent correlations that are not significant. **c** Heatmap of the distribution of diatom ASVs as the percent of total PhyloAssigner-classified diatom amplicons. Shown are the top 10 ASVs, as well as ASVs that made up greater than 10% of the total diatom amplicons in at least one sample. The less frequently observed diatom ASVs were summed and represented here as “all other ASVs”.

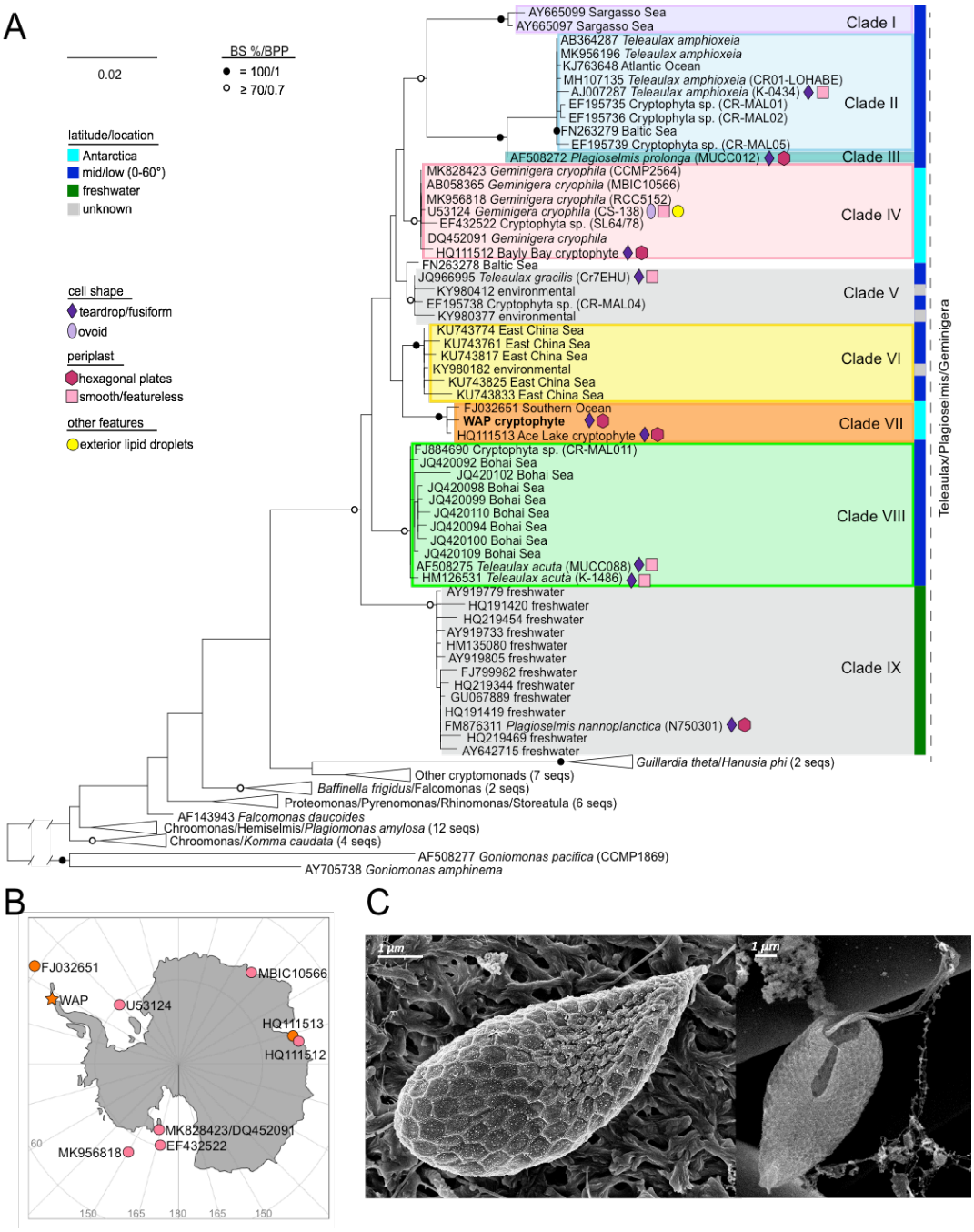


Fig 2.5 The phylogeny, morphology, and distribution of the dominant cryptophyte in this study. **a** Maximum likelihood phylogeny of 18S rRNA genes from cryptophytes within the Teleaulax/Plagioselmis/Geminigera lineage revealing the position of the

WAP cryptophyte clone sequence (in bold). Taxon names are taken directly from GenBank and include strain identifiers accession numbers. Environmental sequences are named by their location or as "environmental" when the location could not be determined. Symbols represent key morphological traits. The right-hand colored bar signifies the latitude at which the sequence containing sample was taken (or the latitude of isolation for sequences from cultured representatives). Potential clades are highlighted and identified with a roman numeral. **b** Map of the Antarctic continent with stations where full-length cryptophyte sequences originated from, either as cultured isolates or environmental samples. **c** SEM images of the WAP cryptophyte where the external morphological features can be appreciated.

.

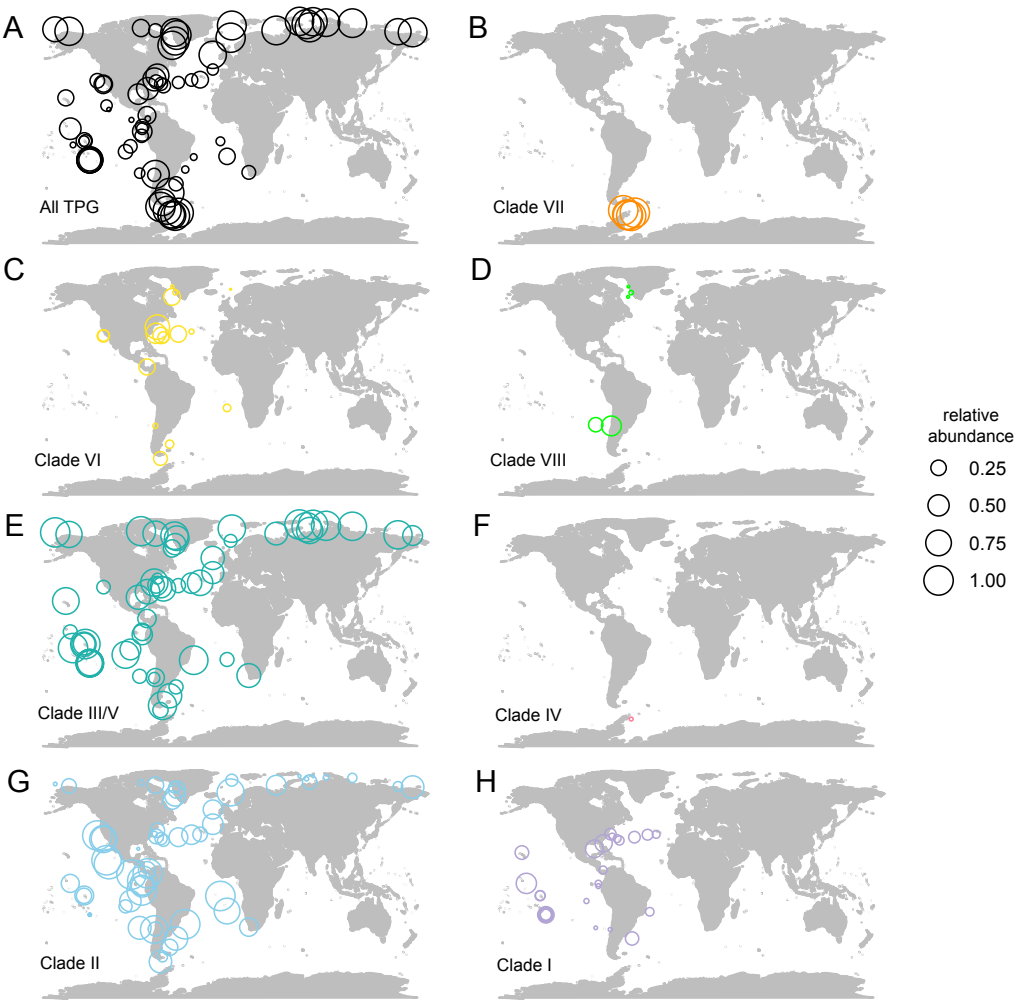


Fig 2.6 Global cryptophyte distributions in the surface ocean. **a-h** Maps of stations from the *Tara Oceans* expedition where surface seawater samples (>0.8 µm filter size fraction) were taken for V9 18S rRNA gene sequencing (data from Ibarbalz et al. 2020). Rings are sized to represent relative amplicon abundance. The first map (**a**) shows the relative abundance of all TPG clade cryptophyte amplicons out of the total cryptophyte amplicons in each sample. The following maps show the relative abundance of amplicons representing Clade VII (**b**), Clade VI (**c**), Clade VIII (*Teleaulax acuta*) (**d**), Clade III/V (*Plagioselmis prolonga/Teleaulax gracilis*) (**e**),

Clade IV (*Geminigera cryophila*) (**f**), Clade II (*Teleaulax amphioxeia*) (**g**), and Clade I (**h**) out of the total TPG clade cryptophytes. Classifications were based on the phylogenetic reconstruction of cryptophytes using the full-length 18S rRNA gene (fig. 5a) and the TPG clades designated within. Specific polymorphisms in the V9 region that connected to each clade were identified, which enabled the designation of representative sequences and their corresponding ASV within the *Tara Oceans* dataset for each clade.

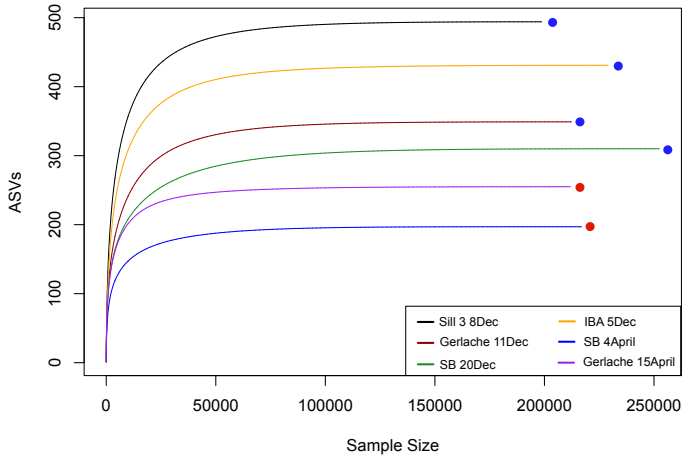


Fig S2.1 Rarefaction curve of the 18S rRNA gene amplicon ASVs. Austral spring samples are indicated with a blue dot, and fall samples with a red dot. Sequence numbers are provided in Table S1.

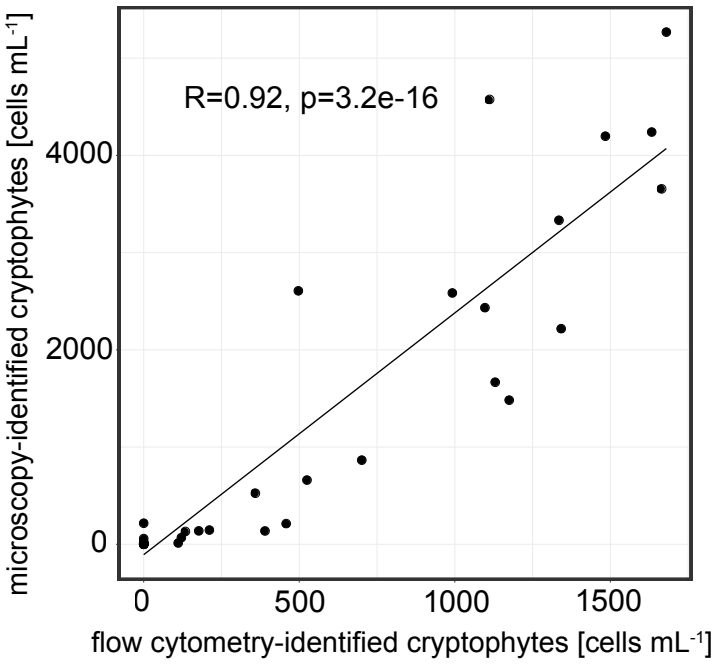


Fig S2.2 The relationship between cryptophyte cells enumerated with flow cytometry vs. with light microscopy.

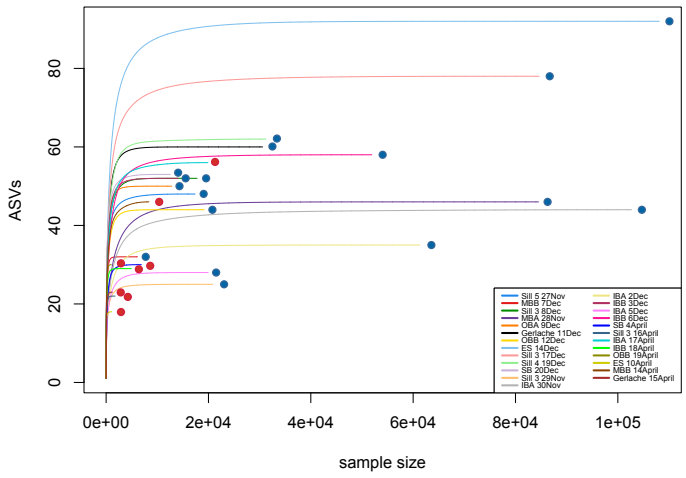


Fig S2.3 Rarefaction curves for the V1-V2 16S rRNA plastid ASVs. Austral spring samples are indicated with a blue dot, and fall samples with a red dot. Sequence numbers are provided in Table S1.

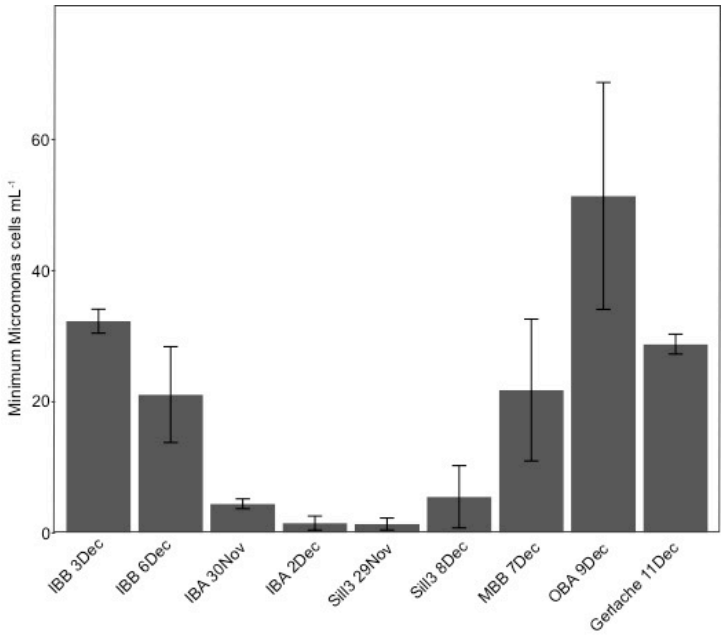


Fig. S2.4 Minimum Micromonas cells mL^{-1} as estimated using qPCR for a subset of austral spring samples. Error bars show the standard deviation of three replicates.

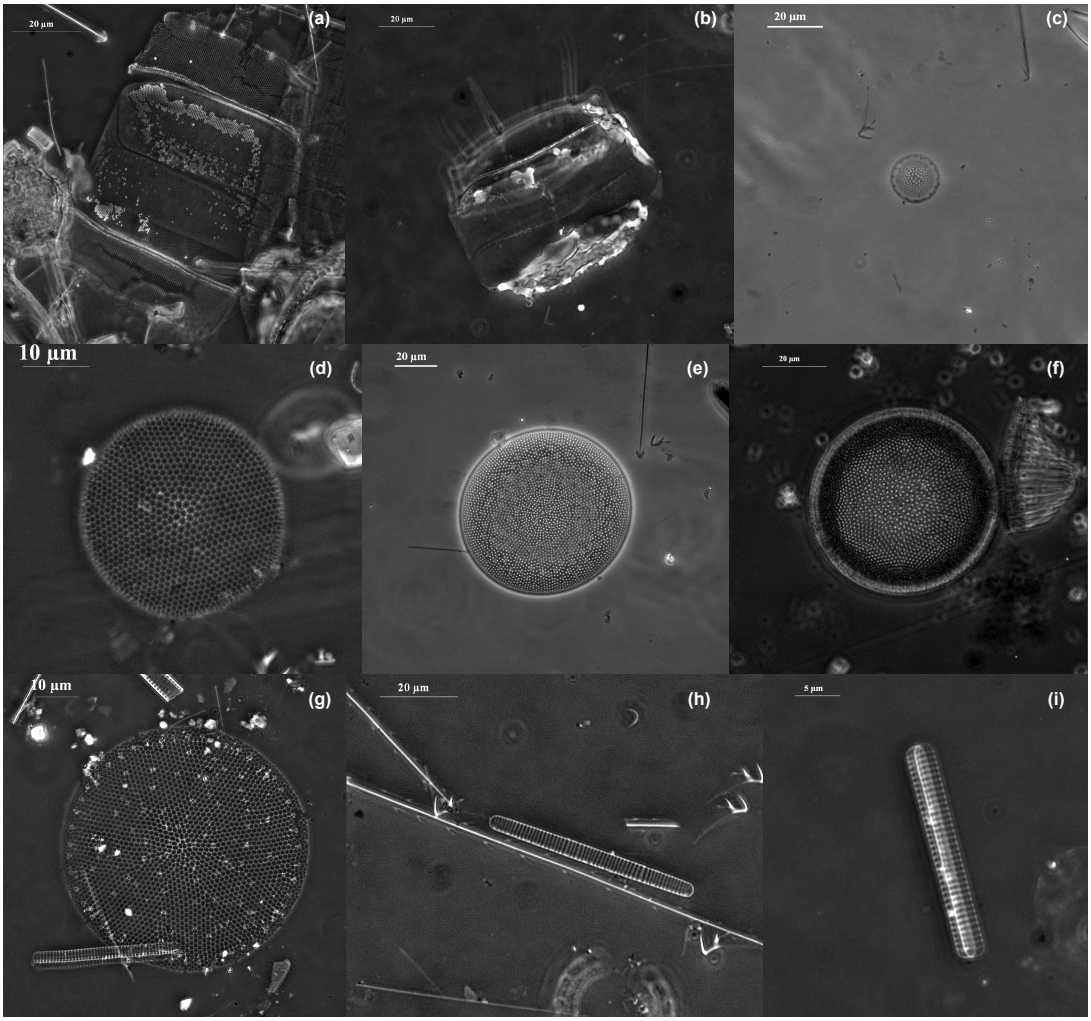


Fig S2.5 Microscopy images showing a selection of diatoms identified, including *Odontella weissflogii* (**a,b**), *Shionodiscus gracilis* (var. gracilis) (**c**), *Shionodiscus poroseriatus* (**d**), *Thalassiosira lentigiosa* (**e**), *Thalassiosira oliveriana* (**f**), *Thalassiosira* cf. *tumida* (**g**), *Fragilariopsis cylindrus* (**h, i**).

	sequence	1	2	3	4	5	6	7	8
1	EF43252 Cryptophyta sp. (SL64/78)	-	99.6	99.71	98.34	99.43	99.14	91.22	97.36
2	U53124 Geminigera cryophila (CS-138)	99.6	-	99.77	98.39	99.48	99.2	91.39	97.36
3	DQ452091 Geminigera cryophila	99.71	99.77	-	98.51	99.71	99.43	91.39	97.65
4	MK956818 Geminigera cryophila (RCC5152)	98.34	98.39	98.51	-	98.22	97.93	92.67	97.63
5	HQ111512 Bayly Bay Cryptophyte	99.43	99.48	99.71	98.22	-	99.14	91.16	97.36
6	HQ111513 Ace Lake Cryptophyte	99.14	99.2	99.43	97.93	99.14	-	91.85	98.11
7	FJ032651 environmental	91.22	91.39	91.39	92.67	91.16	91.85	-	92.97
8	Gerlache Cryptophyte	97.36	97.36	97.65	97.63	97.36	98.11	92.97	-

Fig S2.6 Sequence identity matrix of aligned Antarctic cryptophyte (Clades IV and VII) 18S rRNA gene sequences.

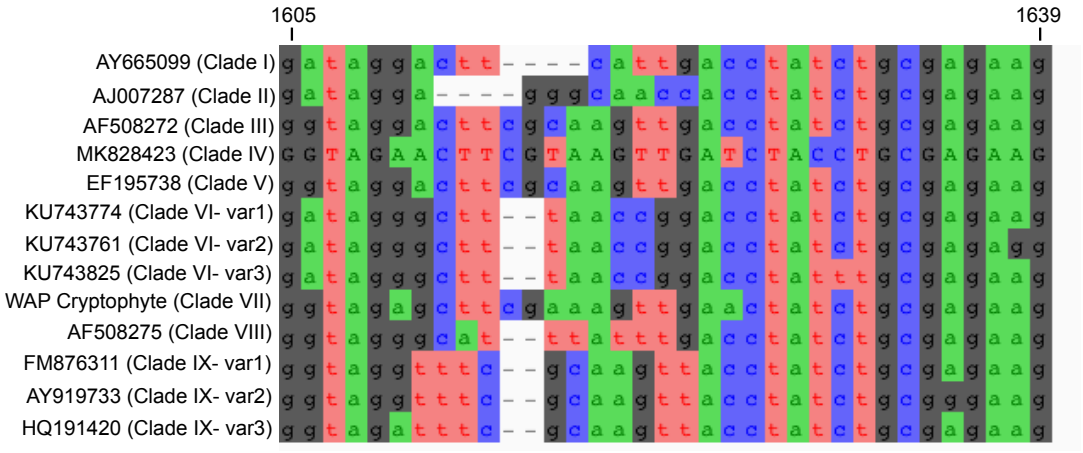


Fig S2.7 A subsection of the alignment of the 18S rRNA gene from cryptophytes representing the clades in the fig. 5a phylogenetic tree, revealing numerous clade-specific polymorphisms within the V9 region. Gaps represent positions present in other cryptophytes (in the full alignment), but absent from the subset shown.

Please see attached file "mhamilton_dissertation_TableS2.1.xlsx"

Supplementary Table S2.1 Station names and depths where samples were taken for DNA filtration, the numbers of total V9 18S rRNA gene amplicon sequences, total V1-V2 16S rRNA gene amplicon sequences, and plastid sequences.

Please see attached file "mhamilton_dissertation_TableS2.2.xlsx"

Supplementary Table S2.2 List of all 18S and 16S ASVs and their taxonomic/nutritional categorization based on Qiime (18S) or PhyloAssigner (16S) classification.

Please see attached file "mhamilton_dissertation_TableS2.3.xlsx"

Supplementary Table S2.3 Flow cytometry-based counts of total eukaryotic phytoplankton cells and those that contain phycoerythrin (attributed herein to cryptophytes; Olson et al. 1989) from spring and fall stations.

Chapter 3: The Evolution of the Marine Green Alga *Micromonas polaris* and Responses to Changes in CO_2 and Nitrate Availability

Abstract

Links between increased dissolved CO_2 and reduced surface nutrient concentrations have been documented, and are predicted to expand and intensify in multiple ocean regions. Less is known about how the interaction between these altered conditions will affect phytoplankton, particularly in polar regions. The prasinophyte, *Micromonas polaris*, is present in Arctic and Antarctic waters; it is among the most abundant picophytoplankton in the Arctic. We present an 106-day experiment performed in replicated continuous-flow photobioreactors that tested biological transitions in this species in response to variations in CO_2 and nitrate availability. Changes in physiological factors and gene expression were examined under variations of CO_2/nitrate supply, including ambient /replete (400 ppm/150 μM), high/replete (1000/150), ambient/low (400/10), and high /low (1000/10). Highest average growth rates occurred under replete conditions, with no statistically significant differences between growth rates, chlorophyll fluorescence, or photosynthetic efficiency under ambient CO_2/replete nitrate and high CO_2/replete nitrate. Growth rates were lowest under low nitrate supply periods, and physiological parameters only varied statistically when compared to the high nitrate supply periods. Thus, nitrate had a stronger influence than CO_2 with respect to gross physiology. Several genes associated with the CO_2 carbon concentrating mechanism (CCM) were identified within the newly sequenced *M. polaris* genome. Similarly to the results from physiological measurements, these genes only responded to nitrate limitation,

but not increased CO_2, which suggests a possible connection between nutrient availability and inorganic carbon assimilation.

Introduction

The physical effects of increased anthropogenic CO_2 on the global ocean have been well documented, including changes to SST, pH, sea ice cover, and surface nutrient concentrations (Naduvinamani and Marali 2008; Hansen et al. 2010; Hoegh-Guldberg and Bruno 2010). Understanding how these altered abiotic conditions may, in turn, affect marine phytoplankton, especially in rapidly changing polar regions, is critical. However, to date there are few model species for environmentally relevant polar phytoplankton in which such studies can be performed. To date, there are only two genomes that have been sequenced including the diatom, *Fragilariopsis cylindrus* (Mock et al. 2017) and the green alga *Chlamydomonas* sp. ICE-L (Zhang et al. 2020). The later is evolutionarily much closer to the prasinophytes that often dominate in some polar regions, but this sea ice associated *Chlamydomonas* sp. ICE-L is likely more representative of a highly specialized niche, distinct from the greater polar open ocean environment. Diatoms, on the other hand, are a core component of the spring phytoplankton assemblage in the Arctic, but are from a different eukaryotic supergroup than the prasinophytes. Thus, while the *Fragilariopsis cylindrus* and *Chlamydomonas* sp. ICE-L genomes are important contributions, additional sequenced genomes from widespread and ecologically relevant polar species are needed to more fully investigate algal physiology and molecular acclimation to shifting conditions.

Availability of multiple relevant models systems is important because the responses of phytoplankton to environmental change, including increased dissolved

CO_2, are by no means heterogeneous. Laboratory culture and mesocosm experiments have revealed a variety of responses to increased CO_2, both direct and indirect, among phytoplankton groups, and even between ecotypes (Schaum et al. 2013; Dutkiewicz et al. 2015; Bach et al. 2017). This diversity of responses by phytoplankton is underlain by the molecular biology of each taxon. It has been shown that changes in gene expression are unique to individual species, despite the fact that they share many of the same genes involved in carbon fixation (Hennon et al. 2017). One of the systems that has received the most attention in marine phytoplankton studies is the carbon concentrating mechanisms (CCMs) that help concentrate CO_2 around the active site of Rubisco, as this carbon-fixing enzyme actually has a poor affinity for CO_2 (Badger et al. 1998). Instead of binding with CO_2, Rubisco will sometimes bind and fix oxygen, resulting in a loss of fixed carbon. This alternate metabolic pathway is known as photorespiration (Bauwe et al. 2010). Under future climate change conditions of increased availability of dissolved CO_2, it would be expected that photorespiration rates would decrease along with the dependence on energetically expensive CCMs (Badger et al. 1998). Decreased expression of CCM genes under elevated CO_2 has been reported for a number of phytoplankton, including the well-studied diatom, *Thalassiosira pseudonana* and the coccolithophore haptophyte *Emiliania huxleyi* (Rokitta et al. 2012; Hennon et al. 2014; Hennon et al. 2017). But the relationship between CCMs, photorespiration, and CO_2 may not be so simple, as two other species of haptophytes, *Gephyrocapsa oceanica* (calcifying coccolithophore) and *Chrysochromulina polylepis* (non-calcifying), were found to

have either no significant difference or an increase in CCM and photorespiration gene expression, respectively (Hennon et al. 2017). Furthermore, the same study found that the raphidophyte *Heterosigma akashiwo* also showed an increase in expression of these genes, suggesting that they may have additional functions in other metabolic pathways (Hennon et al. 2017). This is just one example of how variable the responses of different phytoplankton taxa can be, despite the relative conservation across the genes encoding CCMs.

Another challenge of investigating the phytoplankton response to CO_2 is the consideration of interactive effects. Experiments combining increased CO_2 and nutrient-limited treatments have been able to shed light on some of these effects, particularly in regards to nitrogen limitation (King et al. 2011; Li et al. 2012). Nitrogen is an important component of Rubisco, and when this nutrient is limiting, maintaining carbon fixation with less Rubisco protein is one way of improving N-use efficiency. This could be achieved by increasing CCM activity (Young and Beardall 2005). But the exact relationship between nitrogen, CCMs, and photorespiration in phytoplankton is unclear. This is because of conflicting results wherein some studies show decreased affinity for CO_2 (considered here as a proxy for CCM activity), and others indicate an increased affinity under nitrate limitation (Beardall et al. 1982; Young and Beardall 2005; Li et al. 2012). Experiments investigating the effects of increased CO_2 under low nitrate conditions, as opposed to nitrate-replete, are also closer to real life circumstances. Phytoplankton growth is limited by low nitrate in

much of the ocean, and low nutrient regions are believed to increase with future climate change (Falkowski 1997; Polovina et al. 2008).

In the especially climate-sensitive polar regions, there has been some research on how the phytoplankton community structure may be altered by increased CO_2 using mesocosms (Meakin and Wyman 2011; Brussaard et al. 2013; Schulz et al. 2017). However, few studies have focused on the response of individual taxa, or have considered interactive effects. There has been some work performed using *Micromonas* isolated from Arctic waters, one of which reported a marginal increase in growth rate with increased pCO_2 under both continuous and dynamic light conditions, but no concordant increase in POC production (White et al. 2020). Another found a non-linear response in *Micromonas* to increased pCO_2 under higher temperatures (Hoppe et al. 2018). This clearly highlights the need for comprehensive investigations involving a molecular approach to understanding interactive effects.

In this chapter, I have analyzed an experiment on the polar alga *Micromonas polaris* (CCMP2099) designed to observe its response to shifts that include elevated CO_2 and decreased nitrate. The experimental design, including the use of continuous-flow photo-bioreactors and an extended timespan, was important in that it is allows for a acclimated state of nutrient limited growth to be achieved. While there are several additional challenges and potential biases involved with bioreactor experiments, including divergence between replicate columns and possible evolutionary processes, these experiments are important for attempting to replicate oceanic conditions of limitation compared to traditional "batch culturing" techniques.

Alongside daily physiological measurements, changes in gene expression following conditions of ambient CO_2/replete nitrate, high CO_2/replete nitrate, ambient CO_2/low nitrate, and high CO_2/low nitrate were examined at a genome-wide scale. Genes involved in the carbon concentrating mechanism were given special attention. The published manuscript for this chapter will include an analysis and annotation of the *M. polaris* genome, which has been recently sequenced in the Worden lab. The genome automatic annotation is part of a collaborative effort involving Stephane Rombauts, at Ghent University.

Materials and Methods

Photo-bioreactors

The experiment was conducted in four continuous flow photo-bioreactors custom built at the Monterey Bay Aquarium Research Institute (MBARI) (Wilson et al. 2010) and modified as in Guo et al. (2018). Each photo-bioreactor is equipped with programmable LED lights, a magnetic stir bar, and sensors for temperature, pH, oxygen, and fast repetition rate fluorometry (FRRf).

Culturing setup and sample collection

Axenic *M. polaris* CCMP2099 was grown in 2 L glass photo-bioreactor vessels under a sinusoidal 14:10h light:dark cycle with irradiance reaching a

maximum midday peak of 150 μE m^{-2} sec^{-1}. Cultures were maintained at 8 °C and the duration of the experiment lasted for 106 days in order to capture acclimated responses for each condition. Discrete samples for flow cytometry and RNA-seq were harvested 5-hrs after lights on. Live flow cytometry samples were collected daily and cells were enumerated using an Accuri C6 (BD biosciences, San Jose, CA) before and during the experiment. RNA-seq samples were harvested throughout the experiment with less frequency. For these samples, 120 mL of culture was filtered onto 0.2 μm pore size 47 mm Supor filters and flash frozen in liquid nitrogen.

Nutrient, carbonate chemistry, cell volume and pigment measurements

Nitrogen concentrations within the media carboys and residual concentrations within the photo-bioreactor vessels were measured via auto-analyzer and according to methods described in Pennington & Chavez (2000). The pH of each photo-bioreactor vessel was measured continuously via a Cole-Parmer pH electrode (part no: C-29044-00) mounted within the vessel. Additionally, culture samples were removed from the photobioreactor vessel and independently measured using a Fischer Scientific A815 Plus pH meter with NBS buffer as calibration standards. Periodic DIC samples were also removed from the photobioreactor vessels and then tested according to methods described in Huang et al. (2012). Parameters for carbonate chemistry were calculated using pH and DIC measurements and the CO2SYS program (Lewis et al., 1998)

available online. DIC values were compared against known seawater standards provided by the laboratory of Dr. A Dickson (San Diego, CA, USA).

Experimental design

The experiment started with quadruplicate photo-bioreactor vessels containing stable mid-exponential growth *M. polaris* cells in sterile medium (C1-C4). The L1 medium used was made with sterile artificial seawater as a base. The basic principal behind these bioreactors is that the vessel is receiving a continuous input of media while simultaneously removing culture to maintain a stable growth rate and nutrient concentration. When the cells are moved into conditions (i.e. altered CO_2 or nitrate) that induce a growth rate response, the outflow pump rate can be adjusted to prevent washout/overgrowth of cells. The growth rates (μ) that we report are determined via daily cell enumerations and the following formula:

$$\mu = D_1 + \ln(N_2/N_1)/(t_2-t_1)$$

where D_1 is dilution rate at time t_1, and N_1 and N_2 are cell number at time 1 (t_1) and time 2 (t_2), respectively. For the first 50 days of the experiment, cells were grown with replete nitrate (150 μM) media as the supply, before media was changed, in a gradual shift for the algae, to a low-nitrate (10 μM) supply for the next 52 days. In the final 3 days the media supply was again the replete nitrate concentration. At the start of the experiment, cells were grown under ambient (400 ppm) CO_2. The CO_2 was

increased to 1000 ppm during days 21-40 and again during days 75-94. After day 73, only two of the bioreactors continued the experiment- the other two were lost, likely due to a dramatic temperature spike of unknown origin (reaching 24 °C) discovered within these columns. Hereafter, days will be referred to as T, such that e.g. T20 reflects day 20 of the experiment.

Treatments for RNA-seq were sampled as follows: ambient CO_2/replete nitrate (control group- T20), high CO_2/replete nitrate (HCO2RN- T40), ambient CO_2/transitional nitrate (ACO2TN- T71), high CO_2/low nitrate (HCO2LN- T85/T86), and ambient CO_2/low nitrate (ACO2LN- T99/T100). Due to the loss of two bioreactors (C3 and C4) and the need for statistical significance to examine differential expression, RNA samples from two sequential days were considered here to represent "biological replicates" for the HCO2LN and ACO2LN treatments.

RNA extractions, transcriptome sequencing, and RNA-seq analysis

RNA was extracted from samples representing each treatment using the ToTALLY RNA kit (Ambion) with the addition of a mechanical step as described in Duanmu et al., 2014. The extract was treated using the TurboDNA-free kit (Life Technologies, Grand Island, NY, USA) following the manufactures instructions. The sample from T71 C2 was lost to tube breakage during the extraction protocol. RNA quantity and quality was determined using a Qubit fluorometer (Life Technologies) and a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA), respectively. Library

preparation and sequencing was performed at HudsonAlpha Genome Sequencing Center (Huntsville, AL, USA). Paired-end reads of 150 bp were generated with the Illumina HiSeq platform. Reads were aligned to the *M. polaris* genome using HISAT2 (ver 2.1.0) (Kim et al. 2015). The aligned reads were then quantified using HTSeq-count (ver 0.5.4p3) (Anders et al. 2015) and input into DESeq2 (ver 1.10.1) (Love et al. 2014) for normalization and differential expression analysis. DESeq2 was utilized for normalization and examination of differential expression due its ability to provide enhanced sensitivity and precision for datasets with a small number replicates, while controlling for false positives. The internalized normalization methodology accounts for both sequencing depth and RNA composition bias, the latter of which is important for comparisons between samples.

Identification of candidate CCM genes and Micromonas commoda *orthologues*

Protein sequences for known genes associated with the model green alga *Chlamydomonas reinhardtii* biophysical CCM were compiled based on available literature (Wang et al. 2015; Aspatwar et al. 2018). These sequences were included in a blastp search against the *M. polaris* genome and positive hits were determined based on a cut-off of <E-10. Orthologues genes in the *Micromonas commoda* genome were identified based on blastn searches against *M. polaris* genes of interest, similarly with a cut-off of <E-10.

Statistical analyses

Differences between treatments and timepoints were assessed using a one-way ANOVA and subsequent post hoc Tukey HSD tests with a 95% confidence interval via the base R package, stats (v4.2.0). A multidimensional scaling analysis (MDS) was performed and plotted, and the analysis of similarities (ANOSIM) were generated using the package, vegan (v2.5-7). The MDS utilized Euclidean distance matrices. The heatmap-associated hierarchical clustering was generated via the R package, pheatmap (v1.0.12).

Results

Internal bioreactor chemistry

The internally measured pH generally reflected the CO_2 treatment (Fig 3.1). The only anomaly seen was a spike in pH probe data observed day 39 in C3 on followed by a drop on day 41 before returning to pH levels on par with the other bioreactor columns (Fig. 3.1e). The reasons for this spike in the internal pH probe measurement are unclear and are not reflected in any additional carbon chemistry parameters measured, nor in measurements of cell physiology for these days. The mean pH across all ambient CO_2 treatment periods was 8.21 (±0.16), while the mean for the high CO_2 periods was 7.78 (±0.33). DIC measurements were taken less frequently, but were still indicative of the CO_2 treatment, with a mean of 1841.32 (±57.76) μmol kg^{-1} and 1953.46 (±17.35) for the ambient and high CO_2 treatment

periods, respectively. During the initial replete nitrate phase of the experiment, the residual nitrate concentrations from the bioreactor cultures generally remained high between 55-142 μM, with day 40 and 42 samples from C2 having 39 and 40 μM residual, respectively(Fig 3.1). Residual nitrate was also stable for the low nitrate treatments, between 0.06-0.15 μM concentrations, again with one exception. The residual nitrate concentration measured on day 100 C1 and C2 was 7.73 and 6.42, respectively.

The physiological response of M. polaris to altered CO_2 and nitrate

During the course of the experiment, growth rates, photosynthetic efficiency of photosystem II (F_v/F_m), forward scatter (FSC), and chlorophyll fluorescence (FL3) all exhibited some degree of daily variation (Fig 3.2). The overall trends in these physiological effects also corresponded well between replicates. The highest growth rates were achieved during replete nitrate conditions, but dropped with the introduction of the low nitrate media, before later increasing again. We consider the initial switch to ambient CO_2/low nitrate conditions from day 50-75 to be indicative of a transitional period (referred to as ACO2TN), due to the lower growth rates and higher FSC as compared to the high CO_2/low nitrate (HCO2LN) conditions and return to ambient CO_2/low nitrate (ACO2LN) from day 94-102. The transitional period also saw some of the most dramatic swings in daily growth rate, and displayed a pattern of initial reduction in photosynthetic efficiency followed by recovery (Fig

3.2). The chlorophyll fluorescence showed a clear visual trend across the experiment in all replicates, with a dramatic decrease under the low nitrate treatment which remained consistently low until this nutrient was spiked back in on day 102 (Fig 3.2).

Variation among the treatments and timepoints was determined via a one-way ANOVA and statistical significance was found for all physiological parameters measured (p-value < 7.67e-05), except for photosynthetic efficiency when considering only the timepoints where RNA-Seq samples were taken. Statistical differentiation between specific treatments/timepoints was examined using post hoc Tukey HSD tests (p-value < 0.05). There were no significant differences between the mean growth rates, chlorophyll fluorescence, and photosynthetic efficiency measurements (which included all replicates) during the ACO2RN and HCO2RN conditional periods as a whole, nor between the individual timepoints for any physiological parameter when samples were taken for RNA-Seq analysis (T20 vs. T40) (Fig 3.3, Supplementary Table 3.1, 3.2). The only statistically significant difference between the ACO2RN and HCO2RN treatments was with forward scatter (proxy for cell size), where size appeared to be slightly elevated in the ACO2RN treatment. Both of these replete nitrate periods differed from the two low nitrate periods (HCO2LN and ACO2LN) in all parameters when averaged across time and replicates (Fig 3.3a). Among the individual timepoints, T20 (ACO2RN) differed from the low nitrate timepoints, T85/86 (HCO2LN) and T99/100 (ACO2LN), with significantly higher growth rate, forward scatter, and chlorophyll fluorescence (Fig 3.3b). However, the other replete nitrate timepoint, T40 (HCO2RN), was only

significantly different from T85/86 in terms of growth rate, and in T85/86 and T99/100 with regard to chlorophyll fluorescence. The period of transitional nitrate (ACO2TN) was inconsistently paired with either the replete or low nitrate treatments, and had (for example) a growth rate more similar to the low nitrate treatments (HCO2LN and ACO2LN) compared to the replete (ACO2RN and HCO2RN) (Fig 3.3a). The transitional nitrate period and the specific timepoint measured within occasionally differed from all other treatments, with differences in chlorophyll fluorescence when averaged across all days (Fig 3.3a) and a significantly reduced growth rate for T71 (Fig 3.3b). There were no statistically significant differences observed between the two low nitrate treatments, or the timepoints measured for RNA-Seq, for any physiological parameters.

A closer statistical analysis of the differences between replicates within individual RNA-Seq timepoints/treatments compared to the differences across treatments based on physiological response parameters was performed to evaluate the potential biases associated with grouping two separate days as one treatments (i.e. T85 and T86 as HCO2LN). Visualization of distance via a multidimensional scaling analysis (Supplementary Fig 3.1), revealed a notably wide spread between replicates within a single day in HCO2RN (T40) and ACO2TN (T71), that was not seen in the treatments where two days were measured (HCO2LN and ACO2LN). ANOSIM analyses indicated that there were differences between the T85 and T86 (HCO2LN treatment) and there were not differences between T99 and T100 (ACO2LN

treatment), however in neither case were the ANOSIM findings statistically supported (R = 1, *p*-value = 0.33; R = 0, *p*-value = 0.67, respectively).

Global patterns of gene expression

Overall transcript read alignment rates to the *M. polaris* genome were between 67.5-82.4%. One sample (T40 C3) has a lower alignment rate of 56.6% and was associated with the previously mentioned anomalous fluctuation in pH within the bioreactor column at the time of sampling, and thus, it was removed from the downstream analysis. There were no clear signs of contamination by marine bacteria among the unmapped reads, though some of the unmapped reads could potentially be classified as of viral origin (based on a blastn search). Examination of differentially expressed genes after variance stabilizing transformation of the normalized read counts showed that there were multiple of these genes that were also highly expressed across the majority of samples (Fig 3.4). Hierarchical clustering of samples based on expression of these differentially expressed genes indicated strong similarity between global expression profiles for samples within some treatments (i.e. the T99/100 ACO2LN group), and a disconnect within others (i.e. the T40 HCO2RN group) (Fig 3.4).

An analysis of differential gene expression revealed that there were no statistically significant (padj < 0.05; Wald test, Benjamini-Hochberg adjustment) differentially expressed genes under the HCO2RN treatment (Supplementary Fig

S3.2). The other three treatments, ACO2TN, HCO2LN, and ACO2LN, all had a high number of significantly differentially expressed genes. A greater proportion (56-68%) of these genes were enriched in terms of transcript expression compared to those that were depleted in these treatments (Supplementary Fig S3.2).

The molecular response to CO2 in a low nitrate environment

Differential expression was examined and 19 genes in the HCO2LN exhibited significantly increased expression relative to the ACO2LN treatment, while 458 exhibited lower expression (Fig 3.5; padj ≤ 0.05). Many of these genes were of unknown function and had no orthologues within the *M. commoda* genome (Supplementary Table S3.1). Among the most highly up-regulated genes was an orthologue to an *M. commoda* cyclin-dependent kinase annotated as a core cell-cycle gene. An additional putative protein kinase and Aldose1-/Glucose-6-phosphate 1-epimerase were also enriched in the HCO2LN treatment. The most highly down-regulated gene was one of unknown function, but likely within the zinc-finger protein domain family. Additionally found to be highly down-regulated in HCO2LN treatment was a gene with a protein domain associated with Rubisco assembly, and a putative RNA helicase. None of the CCM candidate genes were found to be differentially expressed in the HCO2LN treatment compared to ACO2LN.

Examination of CCM genes

In the newly sequenced *M. polaris* CCMP2099 genome at least seven genes were identified as putative CCM genes (Fig S3.6). Three of these genes, Miarc19g00220, Miarc01g08680, and Miarc12g01670, were each orthologues to the CCM transporter genes, HLA3, LCIA, and CCP1/2, respectively. Miarc04g07470 was found to be orthologous to LCIB, a gene involved in leaked CO_2 recapture from the pyrenoid. Two orthologues to gamma carbonic anhydrases (γ CA) were found, Miarc14g01200 and Miarc02g04470, which were also paralogous to one another. Finally, an orthologue to a pyrenoid assembly related gene that is essential to *C. reinhardtii* CCM functioning, CIA6, was found (Miarc06g00780).

Statistically significant (padj ≤ 0.05) differential expression was observed for most of these genes in a portion of the low and transitional nitrate treatments relative to the control treatment of ACO2RN (Fig 3.7). The HL3 and LCIB candidate genes were up-regulated in all three low/transitional nitrate treatments (HCO2LN, ACO2LN, ACO2TN) (Fig 3.7 a,d), while CIA6 was down-regulated in all three (Fig 3.7 g). The candidate genes, LCIA and CCP1/2, were significantly up-regulated only in the ACO2TN treatment (Fig 3.7 b,c). One of the candidate γ CA genes (Miarc14g01200) was up-regulated in the HCO2LN treatment only, while the other (Miarc02g04470) was not differentially expressed in any treatment (Fig 3.7 e,f).

Discussion

Small picoeukaryote phytoplankton, including *Micromonas*, are considered to be among the potential "winners" of climate change due, in part, to their large surface area to volume ratio which aids in resource acquisition as nutrient poor regions of the ocean expand (Li et al. 2009). Rising CO_2 and subsequent ocean acidification is thought to potentially even contribute to the proliferation of picoeukaryotes, with increased dominance over other phytoplankton observed in mesocosm studies (Brussaard et al. 2013; Schulz et al. 2017). There has been speculation that phytoplankton taxa typically dependent on a CCM to accumulate sufficient intracellular inorganic carbon at current CO_2 levels may benefit from increased CO_2 availability by being able to down-regulate this energetically expensive mechanism (White et al. 2020). Laboratory culture experiments have shown enhanced growth rates in response to CO2 enrichment, though it should be noted that there is significant variation in the magnitude of this change not only between taxa, but within a single picoeukaryote species as well (Schaum et al. 2013). Growth in arctic *Micromonas* may be slightly boosted by increased CO_2, but many question remain regarding the underlying mechanisms within the cell, as well as the influences of various interactive effects (White et al. 2020). Our findings add to the complexity of this topic, but also augment physiological interpretations with molecular data.

Stability of M. polaris *under increased* CO_2

This study shows *M. polaris* has negligible responses when subjected to increases in CO_2 under nitrogen replete conditions. Not only was no significant alteration to its molecular expression observed under replete nitrate conditions, but comparison of growth rates and physiological parameters under ambient and high CO_2 in replete nitrate media saw neither a positive nor negative responses for most aspects measured. Only a slight potential increase in size (FSC) was observed to be statistically significant. There was no significant differential gene expression observed between these two treatments, indicating that *M. polaris* cells remain more-or-less unaffected by an increase in CO_2 and reduced seawater pH. On an abecedarian level, this lack of response differs from that which is seen for Arctic *Micromonas* in two other laboratory culture-based studies. The mean growth rate observed by White et al. in a non-constant light regime shifted from a μ of 0.37 (± 0.01) in the 400 μatm CO_2 treatment to 0.41 (± 0.02) at 1000 μatm and did represent a small statistically significant increase, suggesting that the cells may have marginally benefited from the CO_2 increase (White et al. 2020). This increase is notably lower than what has seen in similar CO_2 conditions for the picoprasinophyte *Ostreococcous tauri*, with a mean growth rate of 0.76 (± 0.02) at 380 μatm CO_2 and 1.04 (±0.05) at 1000 μatm CO_2 (Schaum et al. 2013). Additionally, White et al. saw no significant increase in POC production with higher CO_2, indicating that while cell abundances may have been somewhat positively affected, actual biomass remained unaltered (White et al. 2020). Another Arctic *Micromonas* study observed a statistically significant increase in mean growth rate after a shift from 380 μatm to 1000 μatm CO_2, but this was only

found when cultures were maintained at a temperature of 6 ºC and not at a reduced temperature of 2 ºC where growth rates were stable (Hoppe et al. 2018).

The current body of work investigating the effect of increased CO_2 on polar *Micromonas* in a controlled laboratory setting (including the present study) indicates that this organism's response is neither a uniform, nor particularly strong one. These results also at first appear to be in contrast to the results of mesocosm studies that show a community shift towards picophytoplankton, and specifically *Micromonas*, at higher CO_2 levels in both polar and temperate environments (Meakin and Wyman 2011; Brussaard et al. 2013; Schulz et al. 2017). But the nature of how these observations are made does not necessarily put the seemingly differing conclusions at odds with one another. While *Micromonas* growth rates may remain stable when CO_2 is increased in an axenic culture experiment, as we show here, this steadiness could still potentially allow them to outcompete other taxa that are negatively affected by high CO_2 in a diverse assemblage of phytoplankton, more reflective of the natural environment.

The influence of nitrogen availability and interactive effects at a molecular level

An additional reality of the global ocean environment under climate change is that increases in CO_2 and subsequent ocean acidification are likely to be coupled with changes in temperature and nutrient availability. Multiple studies have acknowledged the need for incorporating interactive effects into investigations of algal responses to

enhanced CO_2, with nitrogen limitation being among the factors examined (Li et al. 2012; Hennon et al. 2014). In the present study, we did not observe any statistically significant variation between *M. polaris* cells grown with ambient versus high CO_2 under nitrate limitation, matching the lack of response seen with replete nitrate, at least in terms of measurements of cell physiology. The true physiological shift occurred as a reaction to reduced nitrate availability, with a reduction in chlorophyll fluorescence and growth rate. These results are similar to the findings of a study on the diatom, *Phaeodactylum tricornutum*, which saw a significant negative response (using CO_2 affinity as a proxy for CCM activity) with reduced nitrate for the same indicators measured, but no response to increased CO_2 in either high or low nitrate conditions (Li et al. 2012). Based solely on our observations of cell physiology (growth rate, photosynthetic efficiency, cell size, and chlorophyll fluorescence), it would appear that increased CO_2 has only a slight significant size effect on *M. polaris* in replete nitrate, and no effect in the reduced nitrate conditions that are increasingly reflective of the polar ocean environment. Underneath this perceived stability though, is a potentially important molecular response.

While the molecular profile of *M. polaris* remained unaltered after CO_2 was increased in replete nitrate conditions, matching the physiological response, significant changes to gene expression occurred at high CO_2 levels when nitrate was low. A number of genes were highly down-regulated (Fig 3.5), indicating a potential relaxation of cellular processes with the increased availability of inorganic carbon. While many of these genes were of unknown function, one of the most highly down-

regulated genes was a putative Rubsico assembly gene. The potential reduction of Rubisco production in response to increased CO_2 in a low nitrate environment could be an attempt by the cell to alleviate N-stress. While DIC uptake efficiency was not measured, it could be presumed that the activity of the CCM, coupled with increased availability of DIC in the media, was sufficient to maintain a certain level of carbon fixation and growth with less Rubsico. Expression of genes associated with the biophysical CCM, a system adapted to increase the intracellular availability of inorganic carbon for use by Rubisco, was not significantly altered by the increase of CO_2 in the low nitrate treatment. Instead, the *M. polaris* CCM was likely activated by changes in nitrate availability.

Activity of the CCM in M. polaris

Information on the green algal CCM is based primarily on experiments with *Chlamydomonas reinhardtii*, a model species of green algae for which a genetic system has been developed. Here, we attempt to construct the CCM in *M. polaris* based on manual detection and annotation of orthologous genes within the recently sequenced genome (Fig 3.6). While several candidate genes for components of the system were found in this manner, two of the carbonic anhydrases determined to be essential to the functioning of the CCM (CAH3 and CAH6) in *C. reinhardtii* (Meyer and Griffiths 2013) were absent. Of the candidate CCM genes identified in the *M. polaris* genome, most were significantly up-regulated compared to the ambient CO_2

and replete nitrate conditions in at least one of the reduced nitrate treatments (Fig 3.7). This collectively would suggest that the CCM might be activated in response to reduced nitrate availability, potentially as a way to maintain efficient DIC assimilation with less Rubisco and that CAs that are not homologous to CAH3 and CAH6 are present and substituting for their roles.

Two of these candidate genes, the CCM-associated transporter, HLA3, and LCIB were up-regulated in all three reduced nitrate treatments. LCIB is a novel protein with multiple homologues, including LCIC, LCID, and LCIE, and is critical for CCM functioning and growth in low CO_2 conditions (Wang and Spalding 2014). Its proposed functions are to recapture CO_2 and prevent leakage from the pyrenoid in low CO_2 conditions, but many details are still unknown. The precise reason for the observed up-regulation of this specific gene, regardless of CO_2 availability, when nitrate is reduced is also unclear. Another uncertainty is why significant up-regulation of the two chloroplast-localized CCM-associated transporters, LCIA and CCP1/2, was only observed during the transitional phase of the experiment and not for other two low-nitrate treatments (Fig 3.7 b,c). Of the two γ CA genes, only one exhibited significant up-regulation, which occurred under high CO_2 and low nitrate conditions (Fig 2.7e). Very little is known about the role of the γ CAs in green algae, other than their localization within the mitochondria for *C. reinhardtii* (Mitra et al. 2005; Aspatwar et al. 2018). It is unclear if the γ CAs that we have identified here in *M. polaris* are also localized to the mitochondria or elsewhere in the cell. Experimental evidence linking these proteins to the CCM or to CA activity is lacking in *C.*

reinhardtii, but a modified γ CA in cyanobacteria has been shown to be an active CA (De Araujo et al. 2014; DiMario et al. 2018).

The final CCM-associated gene that we report here, CIA6, is a more recent discovery in *C. reinhardtii*, and is proposed to be involved in pyrenoid formation (Ma et al. 2011). The pyrenoid is a specialized micro-compartment within the chloroplast, inside which Rubisco resides and carbon fixation occurs (Borkhsenious et al. 1998). The role of this structure is likely akin to that of the carboxysomes in CCM-containing cyanobacteria, which provide a barrier to CO_2, thus increasing the concentration of inorganic carbon around Rubisco. The formation of this structure is critical for optimal functioning of the CCM, and cells with a mutation in the CIA6 gene in *C. reinhardtii* have a reduced affinity for inorganic carbon (Morita et al. 1999; Ma et al. 2011). Therefore, it is curious that this gene would be down-regulated in all of the reduced nitrate conditions, when other potential components of the CCM were enriched (Fig 3.7f). The necessity of having each aspect of the *M. polaris* CCM operate cohesively is unclear. In *C. reinhardtii,* the primary suite of CCM genes (HLA3, LCI1, LCIA, CCP1/2, CAH3, CAH6, and LCIB/LCIC) is collectively controlled by a transcriptional regulator, CIA5, or the "CCM master switch" (Xiang et al. 2001; Brueggeman et al. 2012; Meyer and Griffiths 2013). We have found no CIA5 orthologue in the *M. polaris* genome, nor has it been detected in *M. commoda* or *M. pusilla* (Worden et al. 2009). The lack of a "CCM master switch" could account for the differences in expression level for the candidate CCM genes observed here.

The reason behind the significant disparity in directionality of differential expression for CIA6 compared to the rest of the CCM remains an open question.

Conclusion

The data presented here indicate that of the two aspects of climate change investigated, nitrate limitation has a much greater influence on the growth and cell physiology of the polar green alga, *Micromonas polaris*, than increased CO_2 availability. Furthermore, increased CO_2 has no significant effect on the growth rate, chlorophyll fluorescence, or photosynthetic efficiency, either positive or negative, in replete nitrate conditions. Interactive effects were not revealed via any changes to growth rate, but there was a clear molecular response to increased CO_2 when nitrate was reduced. Importantly, it should be emphasized that throughout the entire experiment *M. polaris* exhibited positive growth rates, that unlike most experiments conducted to date the cells experience limitation not starvation. Finally,elements of the green algal CCM were identified in *M. polaris* and were found to be responsive to nitrate, but not CO_2, indicating a potential for the connection between nutrient availability and inorganic carbon assimilation that should be explored in future studies.

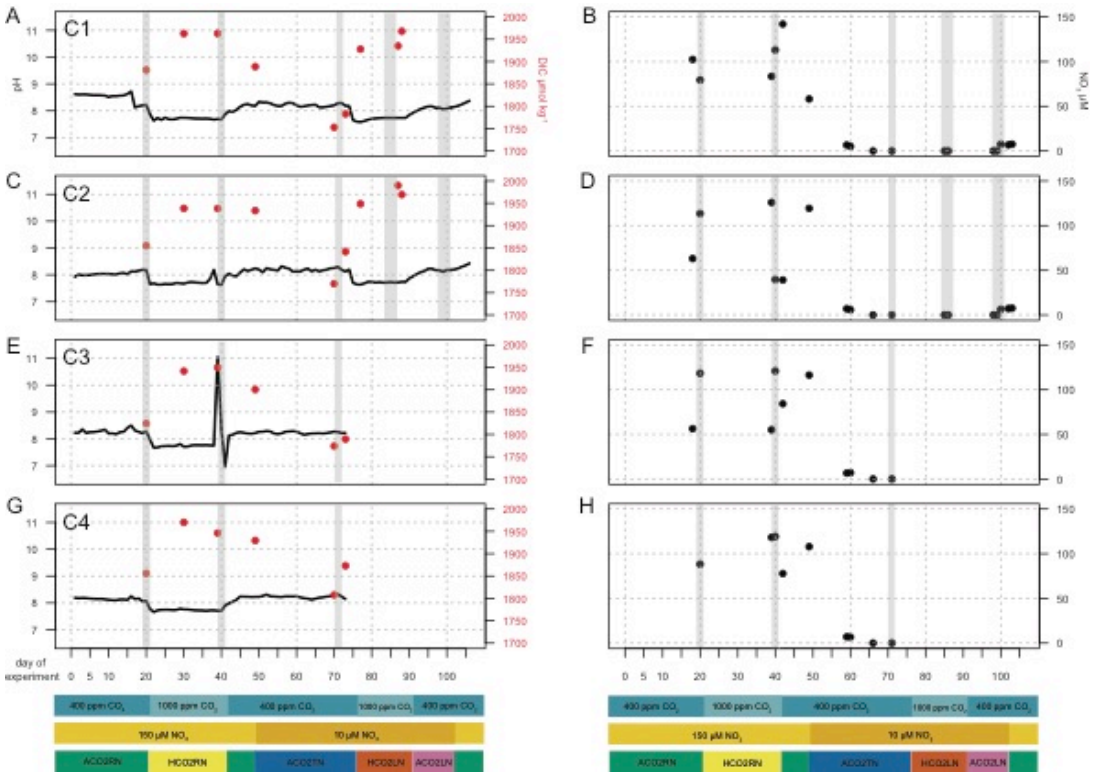


Fig 3.1 Internal conditions within the photobioreactors throughout the course of the experiment. The pH and DIC (µmol kg^{-1}) measurements are shown in the left panels (a, c, e, g), and nitrate (µM) is shown in the right panels (b, d, f, h). Panel rows are separated by bioreactor column (replicate). The yellow and blue bars underlying the panels describe the experimental treatment. Grey shading indicates days when samples were taken for RNA sequencing.

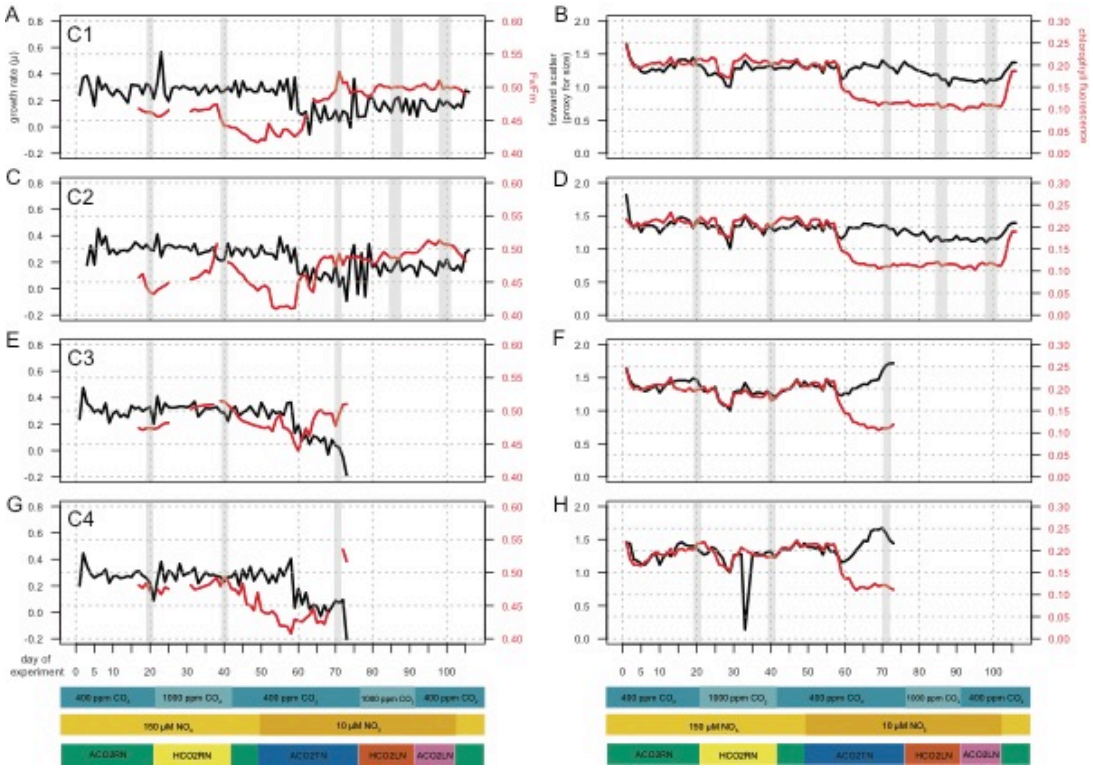
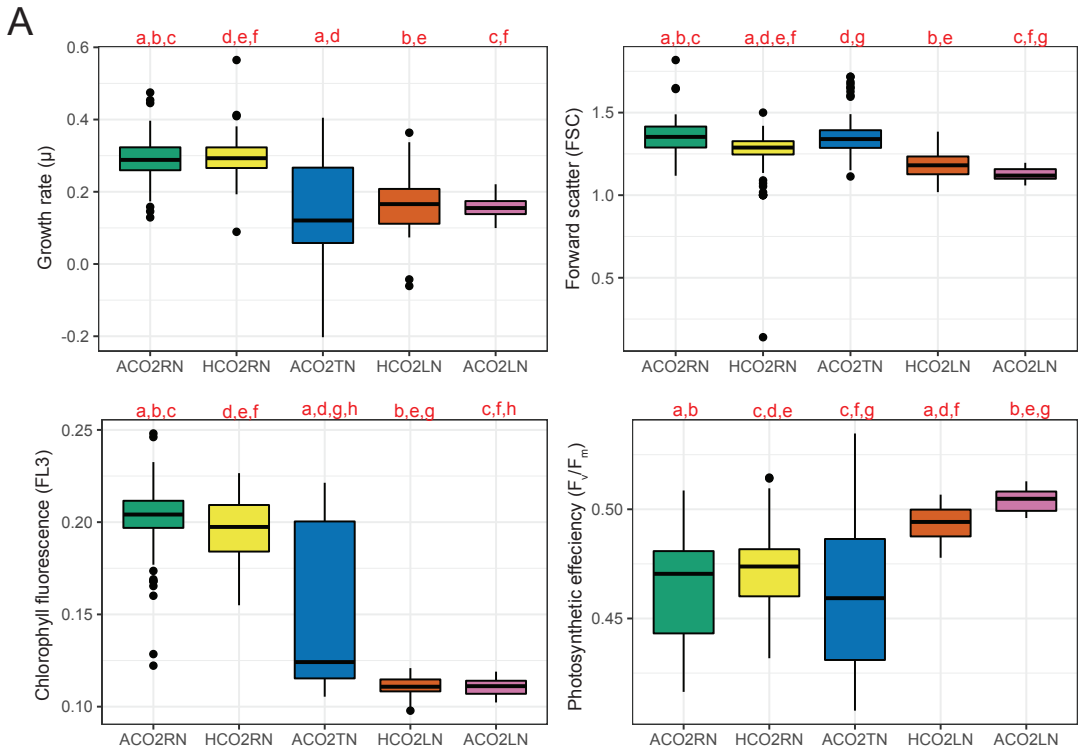


Fig 3.2 *M. polaris* physiology throughout the course of the experiment. The daily mean growth rates (μ) and Fv/Fm are shown in the left panels (a, c, e, g), and forward scatter (FSC- proxy for size) and chlorophyll fluorescence are shown in the right panels (b, d, f, h). Panel rows are separated by bioreactor column (replicate). The bars underlying the panels describe the experimental treatment. Grey shading indicates days when samples were taken for RNA sequencing.


A
a,b,c
d,e,f
a,d
b,e
c,f
Growth rate (μ)
0.6
0.4
0.2
0.0
-0.2
ACO2RN
HCO2RN
ACO2TN
HCO2LN
ACO2LN
a,b,c
a,d,e,f
d,g
b,e
c,f,g
Forward scatter (FSC)
1.5
1.0
0.5
a,b,c
d,e,f
a,d,g,h
b,e,g
c,f,h
Chlorophyll fluorescence (FL3)
0.25
0.20
0.15
0.10
a,b
c,d,e
c,f,g
a,d,f
b,e,g
Photosynthetic effeciency (Fv/Fm)
0.50
0.45

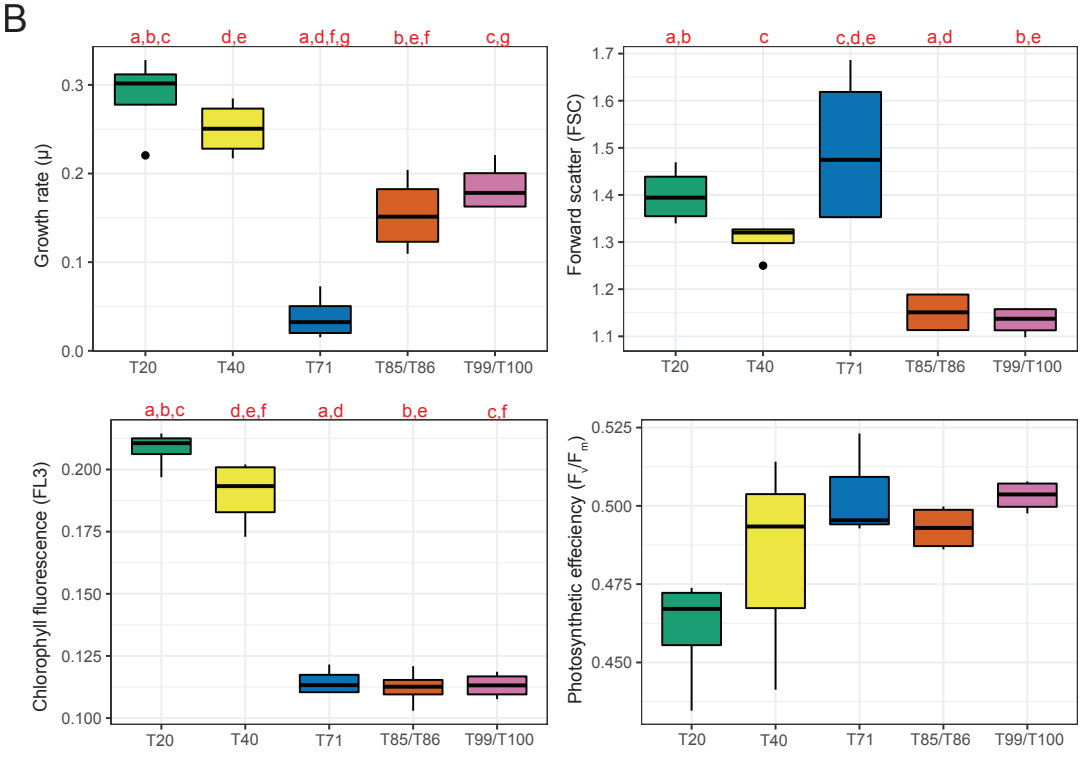

B
a,b,c
d,e
a,d,f,g
b,e,f
c,g
Growth rate (μ)
0.3
0.2
0.1
0.0
T20
T40
T71
T85/T86
T99/T100
a,b
c
c,d,e
a,d
b,e
Forward scatter (FSC)
1.7
1.6
1.5
1.4
1.3
1.2
1.1
a,b,c
d,e,f
a,d
b,e
c,f
Chlorophyll fluorescence (FL3)
0.200
0.175
0.150
0.125
0.100
Photosynthetic effeciency (Fv/Fm)
0.525
0.500
0.475
0.450

Fig 3.3 Box-and-whisker plots of daily measurements of growth rates, forward scatter, chlorophyll fluorescence, and photosynthetic efficiency for the full period representing each experimental condition (ACO2RN = ambient CO_2, replete nitrate; HCO2RN = high CO_2, replete nitrate; ACO2TN = ambient CO_2, transitional nitrate; HCO2LN = high CO_2, low nitrate; ACO2LN = ambient CO_2, low nitrate) (a) and for the individual timepoints where samples for RNA-Seq were taken (b). Red lowercase letters above plots indicate significant difference between treatments/timepoints with matching letters (Tukey's HSD test, p-value < 0.05).

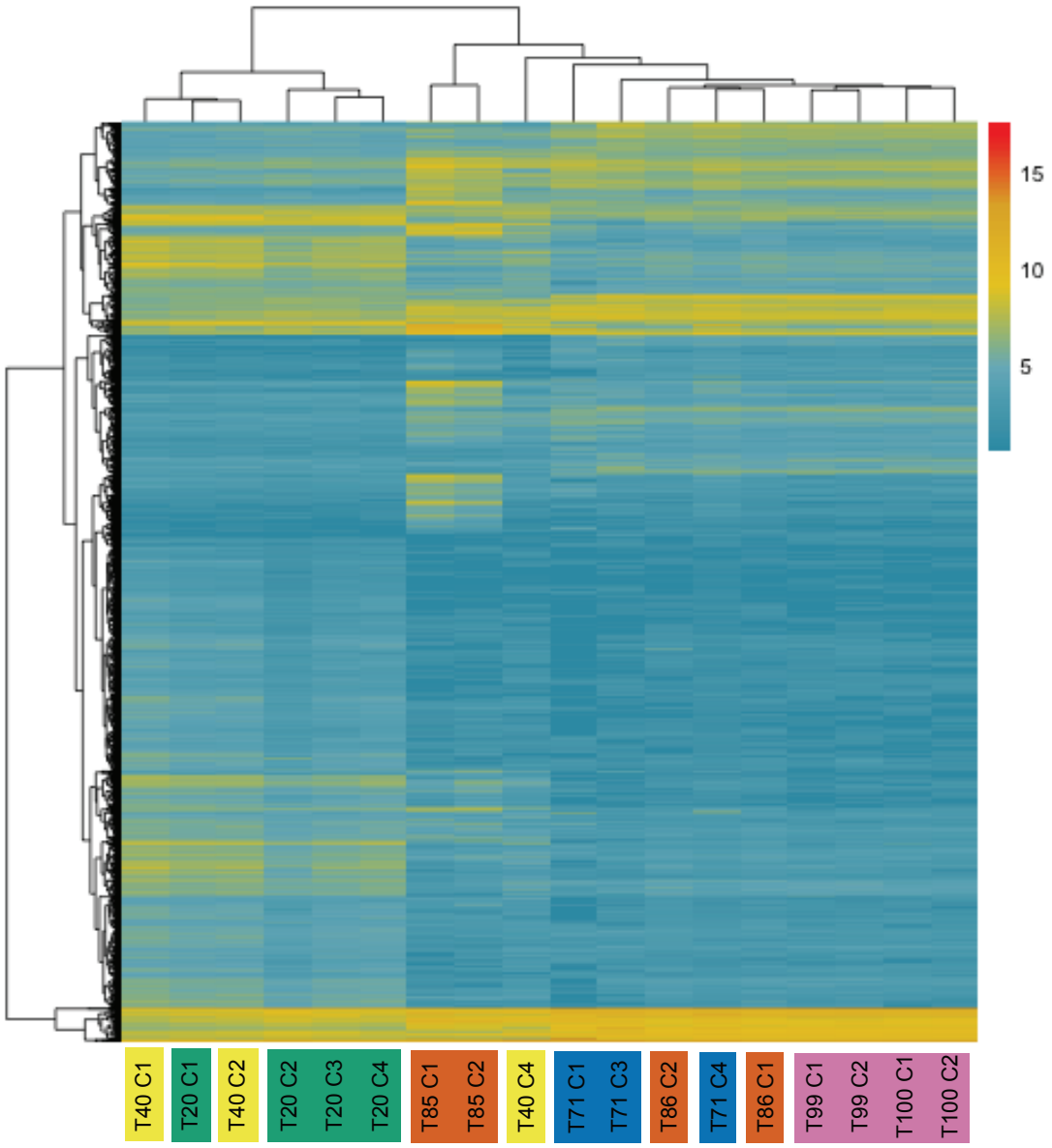


Fig 3.4 Clustered heatmap of all genes (n = 3484) that were differentially expressed in at least one treatment compared to the ambient CO_2, replete nitrate control (ACO2RN). Note that all comparison were made to the replete ambient treatment. The colorbar indicates the normalized and variance stabilizing transformed read counts.

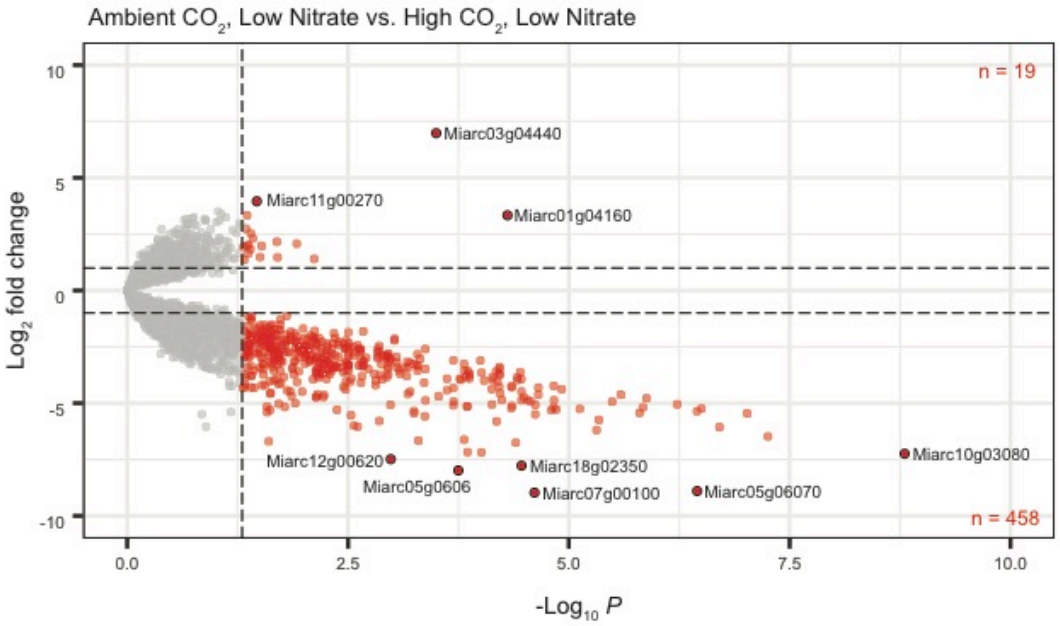


Fig 3.5 Differential gene expression in *M. polaris* during the high CO_2, low nitrate (HCO2LN) treatment, compared to the ambient CO_2, low nitrate (ACO2LN). The volcano plot shows fold changes in transcript abundances for genes that are differentially expressed. Statistically significant (padj ≤ 0.05) genes visualized with red circles, while those that do not pass the threshold of significance are grey. A selection of highly differentially expressed genes are outlined in black and labeled with the gene name. A description of these genes can be found in Supplementary Table S3.3.

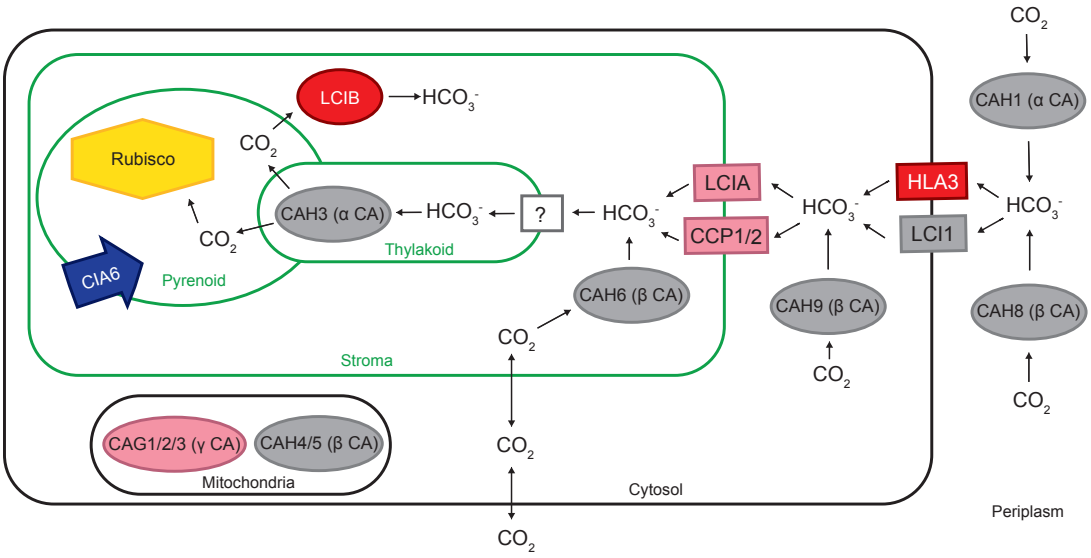


Fig 3.6 A model of the green algal CCM, based on observations in *Chlamydomonas reinhardtii*. Candidate genes present in the *M. polaris* genome are shown in color. Red indicates significant up-regulation in all low-nitrate treatments, pink indicates significant up-regulation in at least one low-nitrate treatment, and blue indicates significant down-regulation in all low-nitrate treatments. Membrane-impermeable bicarbonate (HCO_3^-) is actively transported into the cytosol, then to the stroma, then finally to the lumen of the transpyrenoid thylakoids by an as-yet unknown transporter, where it is converted to CO_2 by the carbonic anhydrase, CAH3. LCIB captures CO_2 that has leaked from the pyrenoid and converts this into HCO_3^- and is essential for CCM functioning at low levels of CO_2. Additional carbonic anhydrases localized to different regions of the cell may convert CO_2 into HCO_3^-. The CIA6 gene is required for normal formation of the pyrenoid and functioning of the CCM.

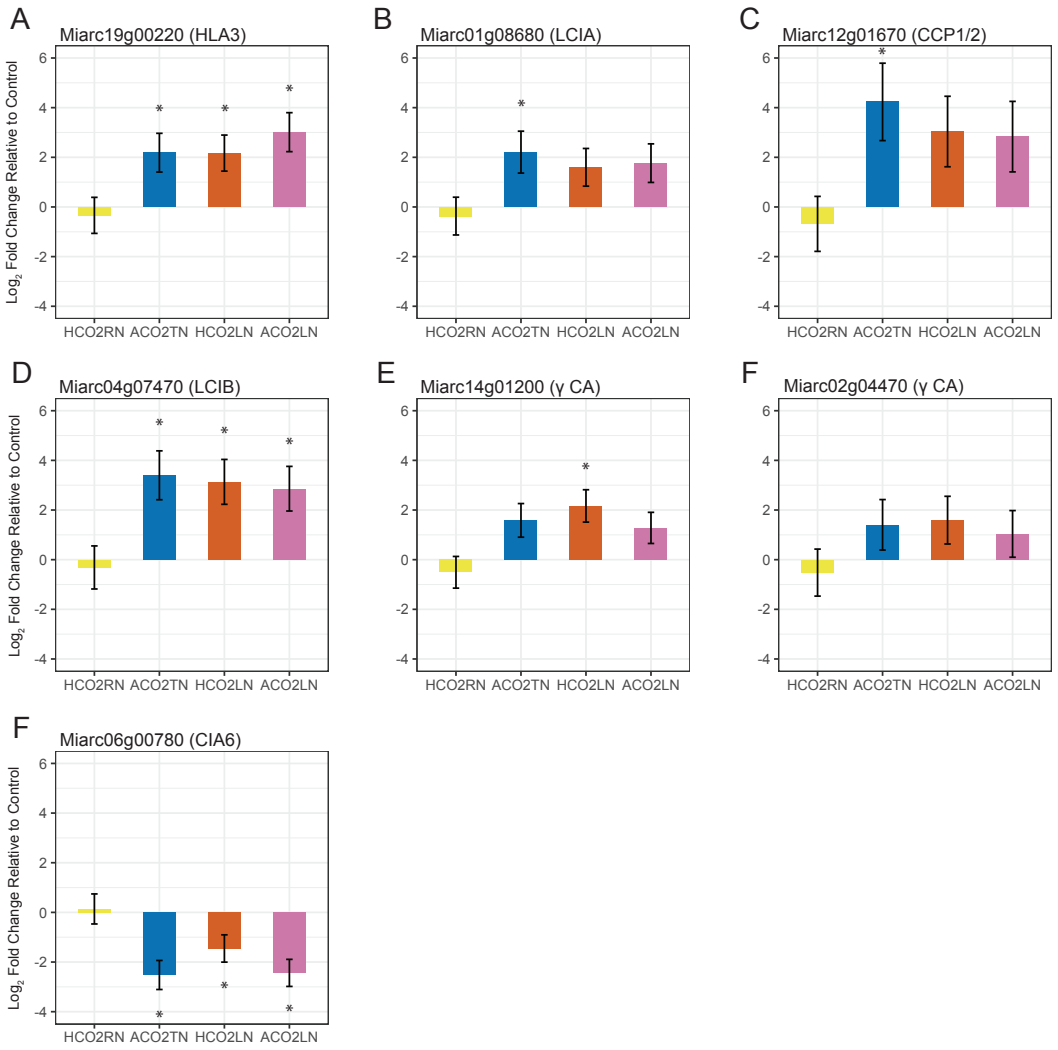
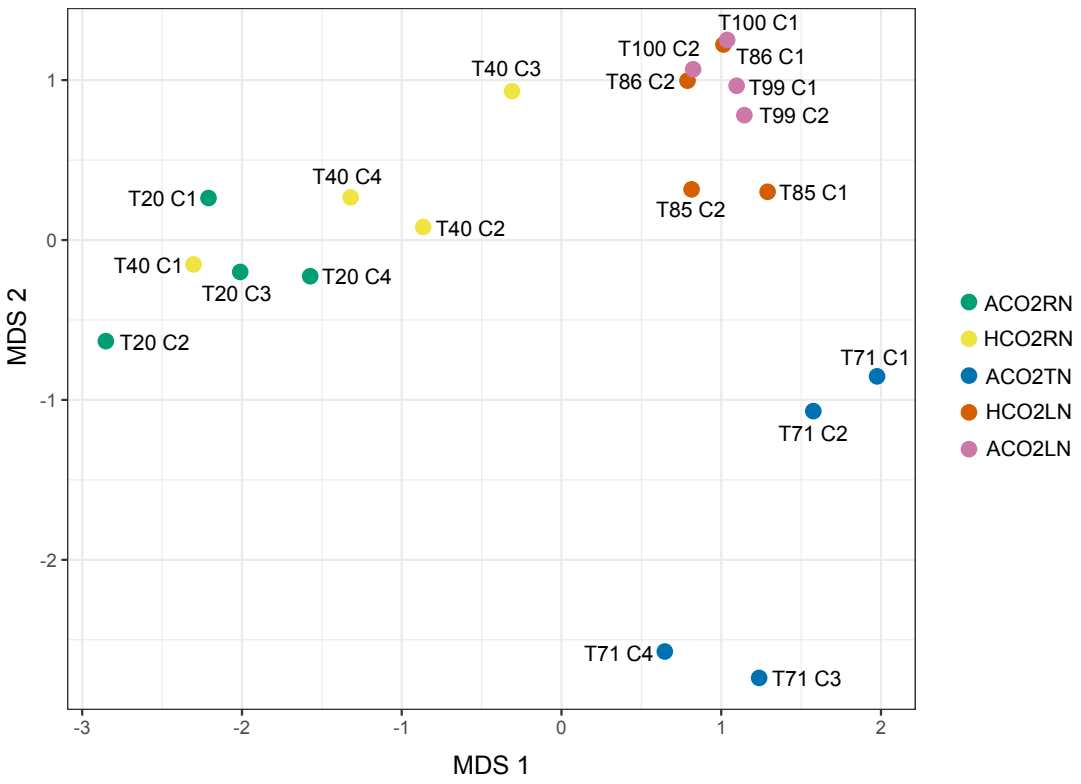
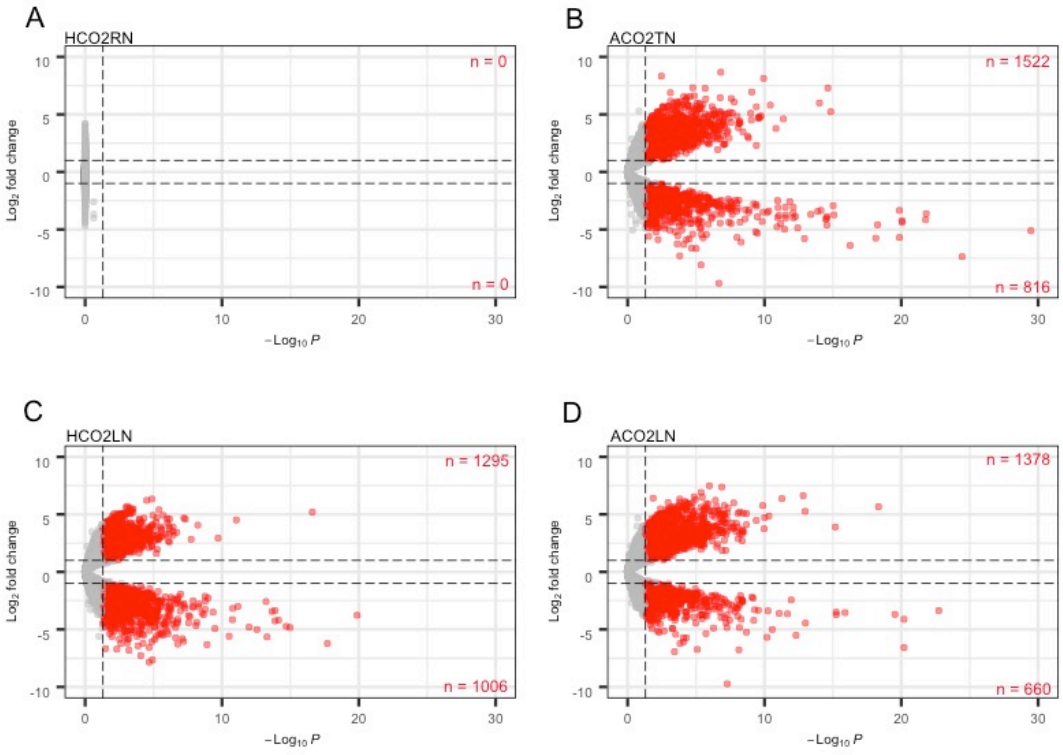


Fig 3.7 Differential expression of *M. polaris* orthologues to *C. reinhardtii* CCM genes in each treatment compared to the ambient CO_2, replete nitrate control group (ACO2RN). Candidate genes include three inorganic carbon transporters: HLA3 (a), LCIA (b), and CCP1/2 (c); a putative CO_2 recapture complex, LCIB (d); two gamma carbonic anhydrases, γ CA (e,f); and a gene required for normal formation of the pyrenoid and functioning of the CCM, CIA6 (f). Error bars represent the standard error based on the Wald test implemented by DESeq2. Asterisks indicate statistical

significance based on the adjusted p-value (padj≤0.05). Treatments include the following: HCO2RN = high CO_2, replete nitrate; ACO2TN = ambient CO_2, transitional nitrate; HCO2LN = high CO_2, low nitrate; ACO2LN = ambient CO_2, low nitrate.



Supplementary Fig S3.1 A multidimensional scaling analysis (MDS) plot based on measurements of physiological parameters (growth rate, forward scatter, chlorophyll fluorescence, and photosynthetic efficiency) for each replicate in the timepoints sampled for RNA-Seq. Samples are colored by treatment conditions: ACO2RN = ambient CO_2, replete nitrate; HCO2RN = high CO_2, replete nitrate; ACO2TN = ambient CO_2, transitional nitrate; HCO2LN = high CO_2, low nitrate; ACO2LN = ambient CO_2, low nitrate.



Supplementary Fig S3.2 Differential gene expression in *M. polaris* for the high CO_2, replete nitrate (HCO2RN) (a), the ambient CO_2, transitional nitrate (ACO2TN) (b), the high CO_2, low nitrate (HCO2LN) (c), and the ambient CO_2, low nitrate (ACO2LN) (d) treatments as compared to the ambient CO_2, replete nitrate control group (ACO2RN). The volcano plots show fold changes in transcript abundances for genes that are differentially expressed. Statistically significant (padj ≤ 0.05) genes visualized with red circles, while those that do not pass the threshold of significance are grey.

Treatment	Growth rate	FSC	FL3	F_v/F_m
ACO2RN	0.29±0.06	1.35±0.10	0.20±0.02	0.46±0.02
HCO2RN	0.29±0.06	1.25±0.17	0.20±0.02	0.47±0.02
ACO2TN	0.15±0.13	1.36±0.13	0.15±0.04	0.46±0.03
HCO2LN	0.16±0.09	1.19±0.08	0.11±0.01	0.49±0.01
ACO2LN	0.16±0.03	1.19±0.08	0.11±0.01	0.50±0.01

Supplementary Table S3.1 Mean and standard deviation of daily measurements of growth rates (a), forward scatter (FSC), chlorophyll fluorescence (FL3), and photosynthetic efficiency (F_v/F_m) for the full period representing each experimental condition: ACO2RN = ambient CO_2, replete nitrate; HCO2RN = high CO_2, replete nitrate; ACO2TN = ambient CO_2, transitional nitrate; HCO2LN = high CO_2, low nitrate; ACO2LN = ambient CO_2, low nitrate.

Treatment	Growth rate	FSC	FL3	F_v/F_m
T20	0.29±0.05	1.36±0.06	0.21±0.01	0.46±0.02
T40	0.25±0.03	1.3±0.04	0.19±0.01	0.48±0.04
T71	0.05±0.02	1.5±0.17	0.12±0.01	0.5±0.02
T85/86	0.15±0.04	1.15±0.04	0.11±0.01	0.49±0.01
T99/100	0.19±0.03	1.13±0.03	0.11±0.01	0.5±0.01

Supplementary Table S3.2 Mean and standard deviation of daily measurements of growth rates, forward scatter (FSC), chlorophyll fluorescence (FL3), and photosynthetic efficiency (F_v/F_m) for each day when samples were taken for RNA sequencing.

M. polaris Gene ID	Direction of differential expression	*M. commoda* orthologue?	Protein domains	Putative function/annotation
Miarc03g04440	+	yes	IPR004147, IPR004147	Protein kinase
Miarc11g00270	+	yes	IPR000719, IPR011009	Cyclin-dependent kinase; core cell cycle gene
Miarc01g04160	+	no	IPR008183	Aldose1-/Glucose-6-phosphate 1-epimerase
Miarc05g03210	-	yes	IPR011545, IPR014001, IPR001650, IPR014014	RNA helicase
Miarc16g01080	-	yes	none	Unknown
Miarc12g00620	-	yes	IPR038052	Rubisco assembly
Miarc18g02350	-	no	none	Unknown
Miarc05g06060	-	no	none	Unknown
Miarc05g06070	-	no	none	Unknown
Miarc07g00100	-	yes	IPR036770	Zinc finger family protein; unknown

Supplementary Table S3.3 Statistically significant (padj≤0.05) differentially expressed genes in the high CO_2, low nitrate (HCO2LN) treatment compared to ambient CO_2, low nitrate (ACO2LN).

Chapter 4: Physiological Acclimation to the Deep Sea by a Photoautotroph

Abstract

The obligate psychrophilic prasinophyte alga, *Micromonas polaris*, has been observed in both Arctic and Antarctic waters, and it has been suggested that this organism may be transported between these poles via deep sea currents in the form of a cyst or other resting-stage cell type. The survivability of *M. polaris* to deep sea conditions was tested with two *in situ* incubation experiments. Growth was measured immediately upon recovery after incubations of up to two months at 1,020 m. Extended incubation times (six days or four months) at 3,560 m resulted in a delayed growth response, but nevertheless indicated that this alga can persist at these extreme depths. Gene expression was observed for all incubations and confirmed viability and cell activity after prolonged exposure to deep sea conditions. Genes associated with meiosis were identified within the *M. polaris* genome and differential expression was examined, as sexual reproduction is linked to cyst formation in other green algal species. While there were no clear molecular indications of meiotic activity triggered by a descent to the deep sea, the delayed recovery growth along with qualitative data, including observations of cell aggregations, point towards the potential for a resting-stage cell type as a strategy for persistence.

Introduction

Phytoplankton are typically restricted to the euphotic zone in order to access the light needed for growth, but deep mixing can bring these phototrophs to depths greater than 500 m (Paluszkiewicz et al. 1994). DNA and RNA (to determine possible viability) have been detected in depths up to 1,000 m in the Arctic for *M. polaris* (Vader et al. 2014). This alga has also been seen in Arctic and North Atlantic Deep Water (up to 3,000 m), though it is unclear whether this represents viable cells or well-preserved DNA, as the RNA was not examined (Vader et al. 2014; Simmons et al. 2015). But the presence of genetically identical *M. polaris* in the Antarctic as compared to the Arctic sequence is suggestive that even after more than 100 years in the deep sea, cells may remain viable. One possible mechanism for this would be an as yet unknown cyst stage in this organism.

In the world of multi-cellular organisms, examining the different life stages of an animal, plant, or fungi is generally considered to be key to understanding that organism as a whole. But for phytoplankton, this is often overlooked, with the exception of diatoms. In diatoms, sexual reproduction, and different types of resting stages have been described and examined extensively (McQuoid and Hobson 1996; Edlund and Stoermer 1997). Other taxa have also been observed to exhibit a resting stage, including dinoflagellates, chrysophytes, and some green algae (Head 1996; Piątek et al. 2009). When discussing the resting stages of phytoplankton cells, the term "cyst" is often used as a general descriptor or umbrella term, but there are many different types (Ellegaard and Ribeiro 2018). Some varieties of resting stages involve

the formation of a thick cell wall and are part of the sexual cycle- these include hypnozygotes, zygospores, and oospores. Akinetes, resting spores/hypnospores, and diatom resting cells are not involved in sexual reproduction and feature varying degrees of morphological differences, as compared to the active cell (Ellegaard and Ribeiro 2018).

Much of what is known about phytoplankton cysts comes from marine sediment core records, as long-term resting stages typically sink out of the photic zone and become buried at the seafloor (Lundholm et al. 2011; Ellegaard and Ribeiro 2018). Germination of cysts from sediment core layers representing cells as old as 100 years has been shown for both diatoms and dinoflagellates (Lundholm et al. 2011; Ribeiro et al. 2011; Ellegaard et al. 2013), but it is unclear how these cysts are able to remain viable for such extended periods of time. Respiration rates can be on the order of 10% of the vegetative cell stages for some of these cysts based on short-term laboratory studies, which would rapidly deplete their carbohydrate storage (Binder and Anderson 1990; Kuwata et al. 1993).

The purpose of cyst formation and the factors (both exogenous and endogenous) that lead to their development and later germination are the subjects of ongoing research. The formation of cysts as part of a sexual life phase makes some sense, as a resistant cell could provide a reservoir of genetic diversity. But not all sexually reproducing phytoplankton feature a resting stage, and not all cysts are associated with sexual reproduction (Bravo and Figueroa 2014; Ellegaard and Ribeiro 2018). For many phytoplankton, a resting stage is an adaption that likely allows for

their survival during times of unfavorable environmental conditions. In some dinoflagellates, cyst formation has been linked with nutrient limitation and changes in temperature (Anderson and Lindquist 1985; Grzebyk and Berland 1996; Figueroa et al. 2005). In other phytoplankton, it is believed to be a mechanism for enduring periods of prolonged darkness, such as during the polar night (Mcminn and Martin 2013). Light has been shown to induce cyst germination in some species (Agrawal 2009), and can trigger formation in others, including the prasinophyte *Pyramimonas amylifera* (Hargraves and Gardiner 1980). Interestingly, the circadian-rhythm linked hormone, melatonin has also been revealed to induce cyst formation, suggesting that there may be endogenous controls on formation/germination that are a response to photoperiod (Balzer 1996).

Cysts are also likely to play an important role in phytoplankton species dispersal (Ellegaard and Ribeiro 2018). Resistant cells are able to survive long journeys on the feathers and feet of marine birds (Schlichting 1960) and in the ballast water of ships (Bolch and de Salas 2007). Bird and human-mediated transport is typically the explanation suggested for phytoplankton dispersal events, but the function of cysts in "natural" (i.e. via currents) dispersal has also been considered (Slapeta 2006; Ribeiro et al. 2012). A key example of this is in the prasinophyte *Micromonas polaris*, where it has been suggested that this psychrophilic alga uses the NADW to travel as a cyst from the Arctic to the Antarctic (Simmons et al. 2015). There has yet to have been any observations of cyst forms of *M. polaris*, though standard laboratory culturing methods, such as the incubation of cultures under a

temperate light:dark cycle, may be impeding possible cyst formation. Additionally, the small size of this organism would make the incidental discovery of cells with unique cyst-related morphologies highly unlikely without electron microscopy techniques.

In this chapter, the response of *M. polaris* to extreme conditions mimicking those found in the NADW, at least in terms of pressure and temperature, was examined using *in situ* incubation experiments. First, survivability of *M. polaris* after incubations was investigated by examining the recovery growth response. Secondly, changes in cell concentration over time, *in situ*, were examined. Lastly, changes in global gene expression were examined and particular attention was paid to genes linked to meiosis. Potential issues with these experiments include the unnatural rapid descent of the culture bags via the ROV and the lack of nutrient replenishment to the cultures during incubation. However, these experiments are meant to serve as an initial investigation, upon which future studies could be built.

Materials and Methods

M. polaris culture preparation

Prior to each experiment, axenic cultures of *Micromonas polaris* CCMP2099 were grown in incubators with a 14:10h light:dark cycle at 6 °C in L1 media made with autoclaved artificial seawater as a base. Cells were diluted with media to maintain mid-exponential growth and were transferred into new flasks as the volume

of culture increased. Growth rates and cell abundances were monitored using an Accuri C6 (BD biosciences, San Jose, CA) (Supplementary Fig S4.1).

Experimental design

Two incubation deployment cruises were conducted in November 2017 (IRC17), and January 2018 (DSEL18) in the Northeastern Pacific Ocean onboard the R/V *Western Flyer*. Deep sea *in situ* incubations were conducted on the seafloor at depths of 1,020 m (location = 36.7709984º, -122.08256º) and 3,560 m (36.117934º, -123.48831º), for the IRC17 and DSEL18 cruise experiments respectively (Fig 4.1a,b). The ROV *Doc Ricketts* was used for transportation of incubation bags attached to metal stands to the seafloor (Fig 4.1c). Prior to deployment, incubation bags (OriGen Biomedical, Austin, TX, USA) were filled with 200 mL of exponentially growing *M. polaris* culture and 800 mL of 0.2 μm-filtered natural seawater as a diluent.

Each set of incubations included two types of controls. One type (TC) acted as a control to test the effects of being moved from flasks into culture bags- these bags never left the ship. The second type of control (T0) tested the effects of deployment- these bags were attached to the ROV itself and were recovered immediately following deployment. IRC17 included incubation times of 1 day (T1D) and 2 months (T2M). DSEL18 featured incubations times of 6 days (T6D) and 4 months (T4M). The DSEL18 4 month incubation bags were recovered on a third cruise in May 2018, FSEL18. Each incubation time included replicates of three bags, with the exception of

the TC controls. The TC controls in both experiments each had two replicate bags. Upon recovery, 1 mL triplicate samples were taken from each bag and fixed and flash-frozen for flow cytometry. Single 200 mL and 100 mL samples were also taken for RNA and DNA, respectively, from each replicate, except for the TC controls. One replicate bag from the TC control incubation in each experiment was sampled twice for RNA and DNA to create a total of three RNA or DNA samples from this incubation timepoint. These samples were filtered onto 0.2 μm micron pore size 47 mm Supor filters and stored at -80°C. Remaining culture from each bag was then transferred to a sterile culture flask and stored in a 4 °C incubator for the remainder of the cruise (IRC17- 2 days following initial TC/T0 recovery, DSEL18- 7 days following initial TC/T0 recovery, FSEL18- 1 day following T4M recovery). The shipboard incubator was outfitted with an LED light on a 14:10h light:dark cycle for the DSEL18 cruise.

Upon returning to land, recovered cultures were moved into a 6 °C incubator. Flow cytometry samples from the recovered cultures were taken daily after return to land, and growth rates and cell abundances were monitored using an Accuri C6. For the DSEL18 TC, T0, and T6D incubations, additional daily samples from the recovered cultures were taken onboard the ship and flash-frozen for flow cytometry measurements of growth rates and cell abundances performed with an Influx flow cytometer (BD, San Jose, CA, USA) on shore as previously described in Chapter 2 of this dissertation (Cuvelier et al. 2010).

RNA extractions, transcriptome sequencing, and RNA-seq analysis

RNA was extracted, sequenced, and analyzed using the same methods and software as described in Chapter 3 of this dissertation.

Identification of candidate genes associated with meiosis and Micromonas *orthologues*

Protein sequences for candidate genes annotated as meiosis-related in *Micromonas commoda* and *Micromonas pusilla* were compiled from Worden et al. (Worden et al. 2009) and used in a blastp search against the *M. polaris* genome. Positive hits were determined based on a cut-off of <E-10. Additional *Chlamydomonas reinhardtii* sex determination genes (as described in Goodenough et al. 2007) and cell wall HRGP genes (Suzuki et al. 2000) were also used in a blastp search, and positive hits in *M. polaris* were determined based on the above described cut-off.

Results

Deep sea conditions and initial recovery growth rates

For the IRC17 experiment, the descent to the 1,020 m incubation depth took ~30 min, where cells were subjected to temperatures of 3.8 ºC at the seafloor. The

mean post-recovery growth rates (μ) for each incubation period were generally within the range of pre-cruise growth rates, with means between 0.26±0.05 d^{-1} and 0.59±0.06 d^{-1} within the first 3 days following recovery (Fig 4.2a). It should be noted that the T1D incubations showed considerable variability between replicates with recovery growth rates ranging from -0.23 to 1.09 d^{-1} for the first day of measured growth (2 days after recovery). The cultures from the T1D incubations began to universally decline in growth rate on the 4th day of recovery, while the cultures from the other incubation periods (TC, T0, and T2M) declined on the 5th day of recovery. The decline in growth can likely be attributed to a lack of nutrient replenishment, which often follows the stationary phase of cell culture growth.

For the DSEL18 experiment, the descent to 3,560 m took ~105 min, and cells were subjected to temperatures of 1.5 ºC at the seafloor during the period of their incubation. The mean post-recovery growth rates for the T0 and TC incubations were consistently low and did not exceed a rate of 0.10±0.01 d^{-1} in the 8 days following recovery (Fig 4.2b). Cell growth rates in these incubations generally did not experience a dramatic decline during the period measured, and remained steady. In contrast, the T6D incubation experienced a rapid decline in cell numbers upon recovery, which hindered the ability to measure growth. Chlorophyll-containing cell populations in all replicates became undetectable via flow cytometry after 6 days. The average concentration of cells in the T6D incubations 3 days following recovery was 108,333±62,517 cells mL^{-1}, nearly an order of magnitude less than that of the other incubation period cultures on the same day. One month after the initial recovery date,

the cultures from the T6D incubations were transferred into fresh media, where they initially did not appear to benefit from this addition. The cultures remained in the incubator for a second month, during which no observations were made. After the second month of incubation, cultures appeared to have a high cell density upon an initial visual inspection. They were then diluted with fresh media for a second time and reached an average growth rate of 0.30±0.01 d^{-1} (Fig 4.2a).

The T4M incubation cultures that were deployed during DSEL18 and recovered during FSEL18 did not initially appear to contain viable cells based on observations via flow cytometry (Supplementary Fig S4.3a). Dark green filaments were seen floating within the incubation bags before agitation upon sampling and transfer to culture flasks (Supplementary Fig S4.2b). These cultures did not immediately respond to the addition of fresh media. After one and a half months of unobserved incubation, cell growth was visually detected (Supplementary Fig S4.2c). Robust populations of fluorescent cells of a similar size to *Micromonas* were also detected via flow cytometry (Supplementary Fig S4.3b).

Patterns of gene expression in response to dramatic changes in depth

For both experiments (IRC17 and DSEL18), transcript reads from the control group samples and short-term incubation samples had standard alignment rates, ranging from 72.2-84.4%. The extended incubation samples (T2M and T4M) had poor alignment rates, between 2.0-58.0%. Examination of a selection of the

unmapped reads indicated that these samples were contaminated by a range of marine bacteria, including some obligate barophilic taxa, implying that they were able to enter the bags at depth. Despite the reduced number of mapped reads, the extended incubation samples were included in downstream analyses due to the unique nature of this treatment.

Hierarchical clustering of samples based on expression of the top 300 genes generally indicated strong similarity between global expression profiles for samples within each timepoint (including the TC control group) (Figs 4.3 and 4.4). This connection between overall molecular response and condition held true for both experiments, with the exception of the T4M timepoint from DSEL18. While these samples did not cluster with any of the other timepoints, they did not form their own group, but instead each sample represented its own outgroup. The DSEL18 T6D samples had an expression profile that differed greatly from the rest of the experiment, including the extended T4M timepoint samples. Several genes of unknown function were expressed here at a uniquely high level (Fig 4.4). When the two experiments were combined and normalized collectively, the PCA revealed a tight clustering of the replicates and of the TC, T0, and T1D incubations for both IRC17 and DSEL18 (Supplemental Fig S4.4). The replicates for the DSEL18 T6D incubation clustered together, but were very distinct from all other samples. The two extended incubations, IRC17 T2M and DSEL18 T4M, were also separate from the other samples and featured a greater spread amongst the individual replicates, particularly for T4M.

An analysis of differential gene expression for each incubation timepoint compared to the associated experimental control (i.e. TC) revealed a number of statistically significant (padj < 0.05; Wald test, Benjamini-Hochberg adjustment) differentially expressed genes, particularly within the IRC17 T2M and the DSEL18 T6D and T4M timepoints. These incubations also had a high number of uniquely differentially expressed genes (Fig 4.5). The uniquely differentially expressed genes from IRC17 T2M and DSEL18 T4M incubations were compared and 149 were shared among the two extended incubations. Of these shared genes, 5% were down-regulated, and 64% were enriched in both, the rest exhibited expression levels that conflicted in direction. For the DSEL18 T6D incubation, most (85%) of the genes that were differentially expressed were down-regulated, but increased expression of a number of genes was still observed (Fig 4.6). The majority of the most highly up-regulated genes in this incubation were of unknown function with no discernable protein domains, but one candidate gene for a P-loop NTP-ase was identified, as was a candidate gene potentially involved in cell wall/membrane/envelope biogenesis, based on the KOG classification of an orthologue from *M. commoda* (Supplementary table 4.1).

Identification and examination of meiosis-associated genes

All putative meiosis-related genes that were previously identified in either *M. commoda*, *M. pusilla*, or both, were similarly identified in the newly sequenced *M.*

polaris genome (Supplementary Table S4.2). Two additional sex-determination associated genes from *C. reinhardtii* were also found to have orthologues in *M. polaris*, GSP1 and GSM1, but expression was observed only for GSP1, where it was seen in all conditions of the two experiments (Supplementary Table S4.3). Several other key *C. reinhardtii* meiosis genes and three genes that are known HRGP cell wall components were searched for in *M. polaris*, but no orthologues were identified (Supplementary Table S4.3). Of the putative meiosis-associated genes deemed potentially essential to meiosis, only one exhibited any statistically significant differential expression in the *M. polaris* incubations, BRCA2 (Supplementary Table S4.2). This gene was uniquely up-regulated in both of the two extended incubation timepoints, IRC17 T2M and DSEL18 T4M.

Discussion

For photosynthetic microbes, the deep sea is typically thought of as a phytoplankton "graveyard", where dead and dying cells spend their final days, if they have not already been fully decomposed, their organic components remineralized and incorporated into new forms of life. Some living phytoplankton are able to survive in the deep sea, as well as in the sediment, as cysts. The formation of this resting-stage cell type in many, if not most, taxa occurs within the euphotic zone, and can be triggered by a variety of biotic and abiotic factors including nutrient limitation, anoxia, and darkness (Bravo and Figueroa 2014; Ellegaard and Ribeiro 2018). The

possibility of cyst formation taking place after or even during a cell's descent into the deep has yet to be explored. In fact, when it comes to the general physiological and molecular changes that a healthy phytoplankton cell might undergo if it were to rapidly sink, virtually nothing is known. Sinking rates of phytoplankton aggregates and cells attached to fecal pellets have been documented as being within a range of ~10-1000 m d^{-1} (Michaels 1988; Ploug et al. 2008), and sites of NADW formation in the Arctic can also bring surface water to the deep at a rapid rate (Dickson and Brown 1994). These rates are well within a reasonable timespan that would be required for a healthy phytoplankton cell to reach the deep sea intact (i.e. before the stress response leads to programmed cell death). Non-encysted diatoms with intact plasma membranes (an indication of a living cell) have been observed at bathypelagic depth ranges of 2,000-4,000 m, but recovery of cell growth was not shown (Agusti et al. 2015). By reconstructing an artificial mechanism for rapid descent, we are able to observe the molecular changes that occur when a polar green alga is transported to the deep sea, and show recovery growth rates even after prolonged persistence in the bathypelagic.

The survivability of M. polaris *after deep sea exposure*

The deep sea environment is characterized primarily by its low temperature, high pressure, and lack of light- all of which interfere with or explicitly inhibit phytoplankton growth. Polar phytoplankton may be among the taxa with a particular

potential to persist at depth and/or form cysts in response to a rapid environmental shift, as they are already well adapted to cold temperatures and some have developed strategies that allow them to survive through periods of 24-hours of darkness characteristic of high latitude winters (Berge et al. 2015; van de Poll et al. 2019). Temperature and a lack of light availability are also two factors that can be more easily tested in laboratory experiments, while the high pressure conditions present at depths >1000 m are much more difficult to replicate onshore. The unique experimental design of the present study allowed for the examination of these effects on a specific taxa of phytoplankton *in situ*. We have found that the polar prasinophyte, *Micromonas polaris*, is surprisingly resistant to extreme changes in depth, even after prolonged periods in the deep sea (Fig 4.2a). Immediately following a two month incubation at a depth of 1,020 m, *M. polaris* cells reached a mean growth rate of up to 0.59±0.06 d^{-1}, which would fall into the high end of the range of exponential growth rates reported for standard laboratory conditions (Lovejoy et al. 2007; Hoppe et al. 2018). While cells did decrease in number during this incubation period, they appear ready to instantaneously resume growth once conditions become favorable again. This result is not wholly unexpected, as *Micromonas* is among the taxa known to survive through the polar night in surface waters (Vader et al. 2014).

Incubations at a depth of 3,560 m produced very different results than those at 1,020 m. While *M. polaris* was still able to rapidly recover growth after an incubation time of zero days (immediate retrieval following descent), the intermediate and extended incubations did not (Fig 4.2b). After a period of six days at depth, there was

a significant loss of cells, and the abundance continued to decline once in the laboratory incubator with a light-dark cycle in the first week following retrieval. But eventual recovery was mediated by the addition of fresh media (Fig 4.2c). It should be noted that all incubations at the start of the experiment consisted of healthy cell populations that were diluted with filtered natural seawater, devoid of nutrient amendment, which could potentially limit recovery growth. Despite this, cells were able to resume growth immediately after two months at 1,020 m, unlike in the six day incubation at 3,560 m. This indicates that there is potentially a pressure-related effect on the cells that is only observed at a deeper depth. The four month incubation showed a similar delayed response. Upon retrieval, a viable population of *M. polaris* cells was not detected via flow cytometry (Supplemental Fig S4.3). It is unclear if the delayed recovery observed is the result of a very small number of still viable cells that would have not been initially detected as *M. polaris*, or the result of encysted cell germination. The green filaments, presumably made up of *M. polaris* cells containing chlorophyll, seen in the four month incubation bags may provide an additional clue regarding the response of the cells to depth. When faced with hostile environmental conditions, some phytoplankton, including the green alga, *Chlamydomonas*, form aggregates, as well as clusters of 4-16 cells surrounded by a cell wall called palmelloids (de Carpentier et al. 2019). Encysted *Chlamydomonas* zygotes can also form aggregates (Suzuki et al. 2000), which may confer additional stress resistance. Nitrate availability and light fluctuations (Goodenough et al. 2007; Zou et al. 2017) have been described as controlling factors for sexual reproduction (meiosis) and

subsequent cyst formation in *Chlamydomonas reinhardtii*, while additional environmental factors including temperature and salinity have been linked to cell wall formation in other green algae (de Carpentier et al. 2019). But depth-related pressure changes have yet to be considered in terms of the cellular stress response. By comparing the recovery growth rates (or lack thereof) for *M. polaris* after incubations at two different depths, we are able to discern a distinct difference between depths of 1,020 and 3,560 m. Subjugation to the most extreme depths likely results in a cellular stress response that includes the formation of cell aggregates.

The response of M. polaris *to extreme depths at a molecular level*

While the presence of *M. polaris* cysts could not be concretely observed due to the small size of the cells as well as the limitations of the techniques and equipment at hand during these experiments, the molecular response of this organism to incubations at depth was investigated. One key finding is that even after a 4 month incubation at 3,560 m, when cells appeared to be "dead", a proportion remains transcriptionally active (Fig 4.4), a clear indicator of viability. Using RNA, which degrades rapidly, as proxy for cellular activity allows for the examination of processes occurring within the cell at a specific point in time, but it is not without its limits, particularly within the context of these experiments. Outside of the response to extreme depth, act of retrieval may constitute a stress response in cells that have had time to acclimate to the deep sea environment in the moderate (six days) and

extended (two and four months) incubations. These three incubation periods each had a high number of uniquely differentially expressed genes (Fig 4.5), suggesting that a significant change to molecular processes is occurring, even if cells display instantaneous growth recovery similar to the control, as in the two month incubation at 1,020 m.

Though the two extended incubations resulted in a very different recovery growth rate response for each depth, there were a number of uniquely differentially expressed genes shared between them, indicating a universal molecular response to general deep sea conditions regardless of actual depth. Among these shared genes was an orthologue to a putative BRCA2 gene in *M. commoda* (Worden et al. 2009). Though this protein is more famously known for its association with human breast cancer, BRCA2 has been determined to be essential for meiosis in the model plant, *Arabidopsis* (Siaud et al. 2004), but it may also be involved in other cellular functions, such as transcription regulation during the plant immune response to genotoxic substances (Wang et al. 2010). We identified many potential meiosis-associated candidate genes within the *M. polaris* genome, but none of the other genes essential to this process were found to be significantly differentially expressed in any incubation (Supplemental Table 4.2). These results do not rule out the possibility of potential mating-induced zygotic cyst formation, akin to what has been described in *Chlamydomonas*, occurring in *M. polaris*. It is possible that this process did occur, just not within a timespan that would provide evidence in our post-incubation RNA samples.

While the six day incubation at 3,560 m additionally did not reveal any differential expression of candidate meiosis-related genes, it provided evidence of a strong molecular response to this condition based on the high number of genes that were uniquely differentially expressed (Fig 4.5). The majority of differentially expressed genes were down-regulated (Fig 4.6), which might indicate that a large number of non-essential and/or light-mediated functions were reduced in importance as the cell undergoes (or prepares for) encystment or deep sea acclimation. A number of genes did exhibit up-regulation in this incubation, indicating that some functions were specifically triggered by deep sea conditions after a six day period. Among the most enriched, was one candidate gene of potential interest, Miarc05g04590 (Fig 4.6). The *M. commoda* orthologue of this gene was annotated as being related to "cell wall/membrane/envelope biogenesis", based on the KOG classification. It is difficult to determine if this candidate gene is actually involved in the formation of a cyst-like cell wall in *M. polaris*, but its up-regulation under conditions that inhibit normal processes of cell growth indicate this gene's role in a stress response.

Conclusion

The data presented in this chapter provide the first evidence indicating that the polar green alga, *Micromonas polaris*, is able to survive at extreme depths of over 3,500 m for up to four months. While there was no definitive molecular evidence suggesting that zygotic cyst formation was actively occurring, a number of putative meiosis-associated genes were identified within the newly sequenced *M. polaris*

genome, indicating genetic potential for this stress response mechanism. The capability of *M. polaris* to persist in the bathypelagic for extended periods of time carries important implications regarding algal dispersal via deep ocean currents.

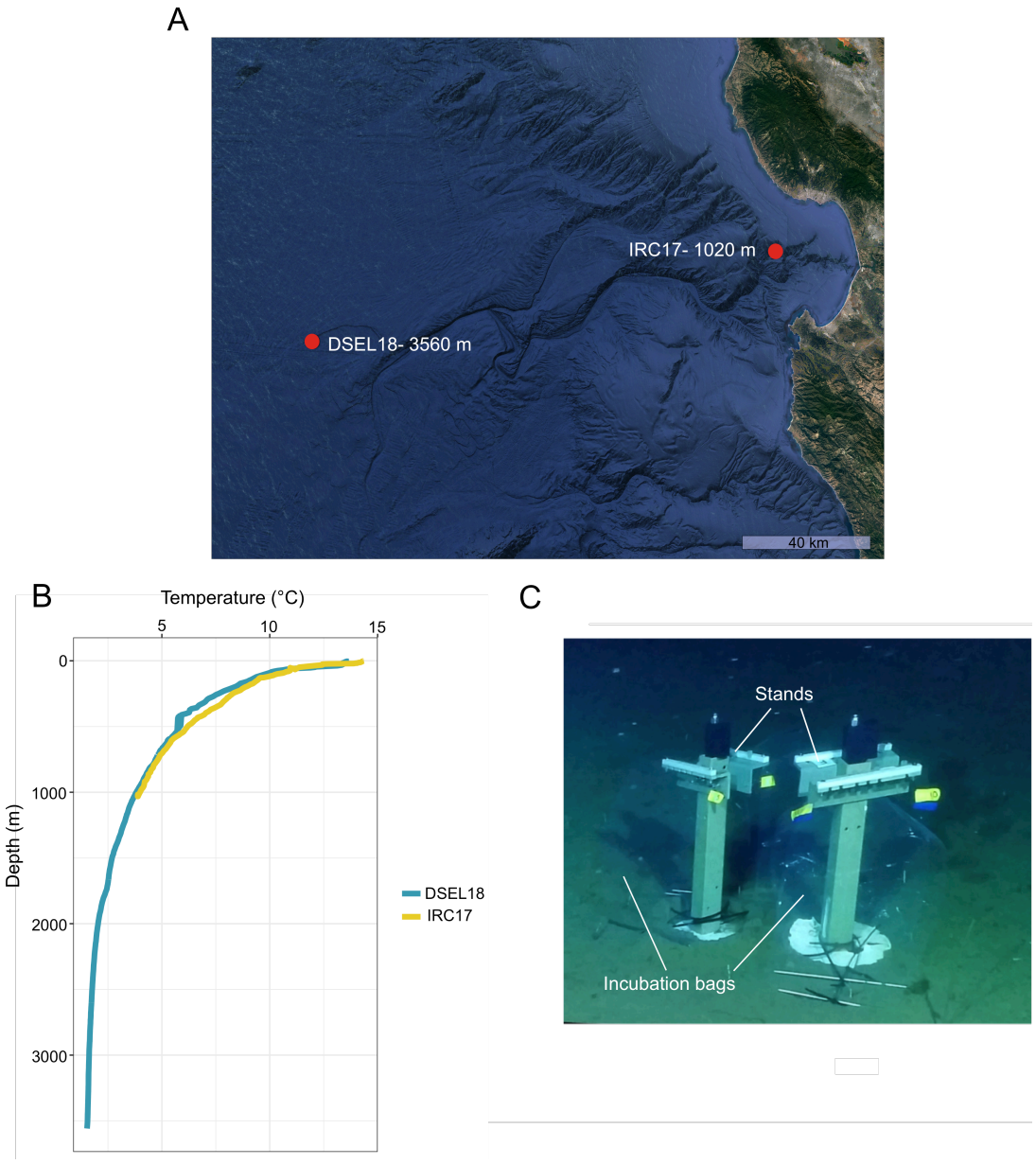


Fig 4.1 An overview of the two incubation sites and the incubation apparatus. The top panel map (a) indicates the incubation sites within and outside of the Monterey Bay, CA. The bottom panels include depth profiles showing temperature (b), and a photograph of the stands holding the incubation bags upright at the bottom of the seafloor (c).

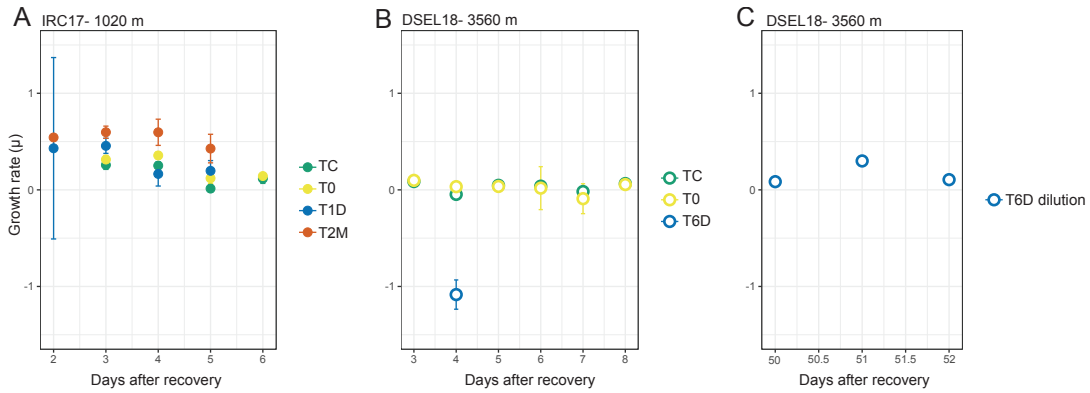


Fig 4.2 Growth of *M. polaris* cells immediately following recovery from incubation at 1,020 m depth in the IRC17 experiment (a) and 3,560 m depth in the DSEL18 experiment (b). The y axis indicates the number of days after incubations were recovered. Both the control incubations (TC) in each experiment are represented by two replicates. All other incubation timepoints include three replicates. The T6D incubation was additionally diluted with fresh media and growth was measured 50 days after recovery (c).

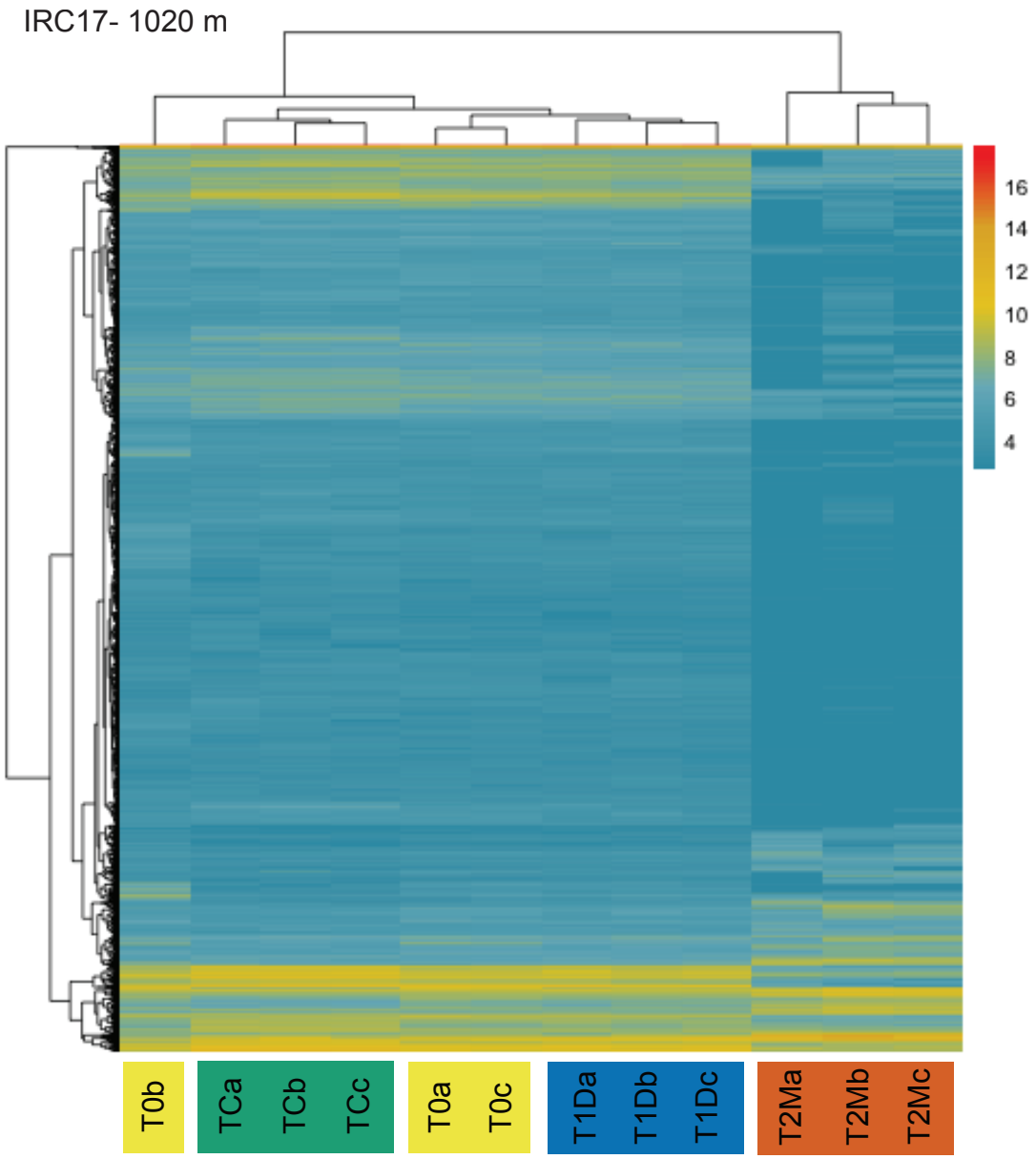


Fig 4.3 Clustered heatmap of all genes (n = 2913) that were differentially expressed in at least one treatment compared to the control (TC). The colorbar indicates the normalized and variance stabilizing transformed read counts.

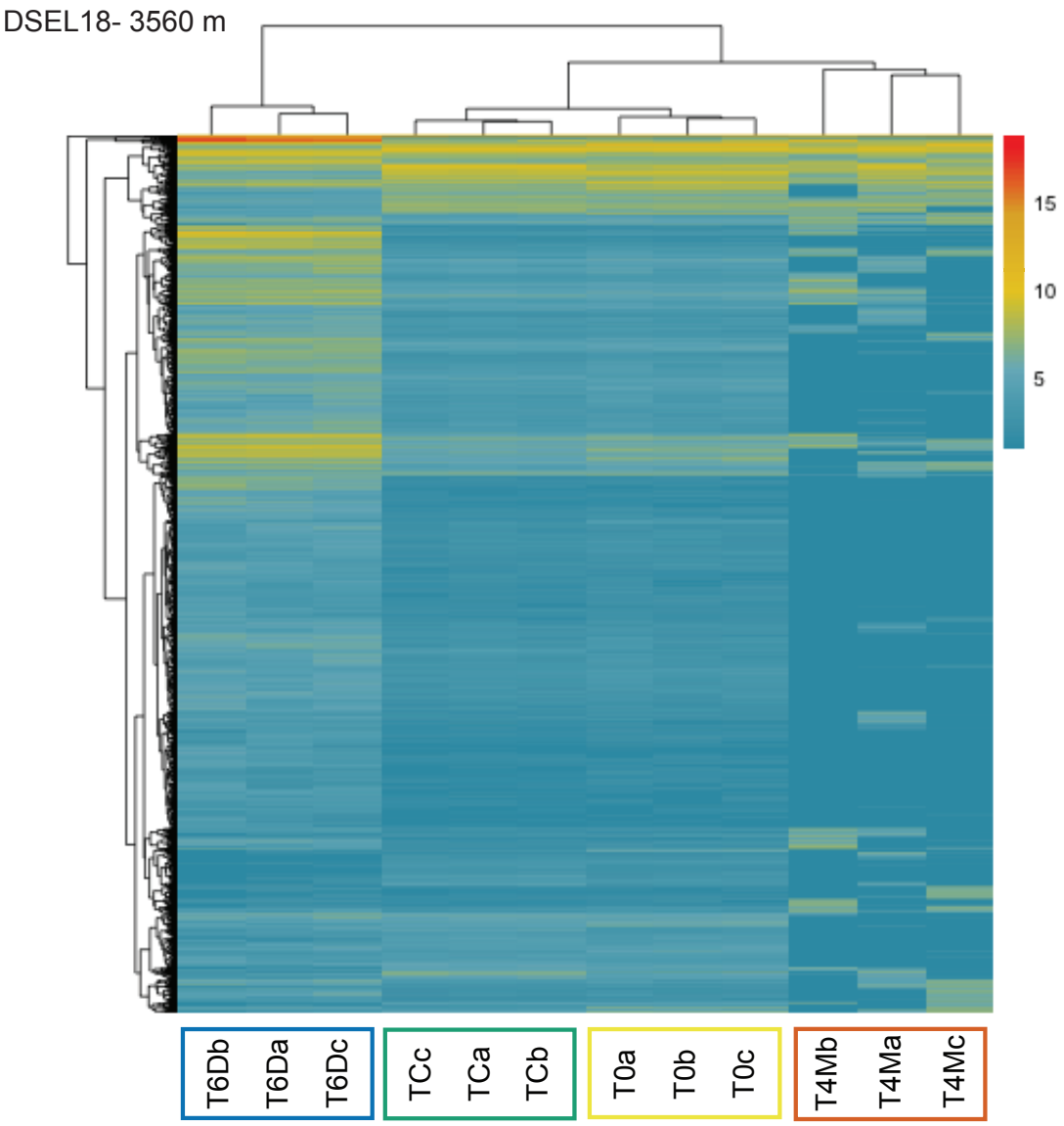


Fig 4.4 Clustered heatmap of all genes (n = 3027) that were differentially expressed in at least one treatment comparison. The colorbar indicates the normalized and variance stabilizing transformed read counts.

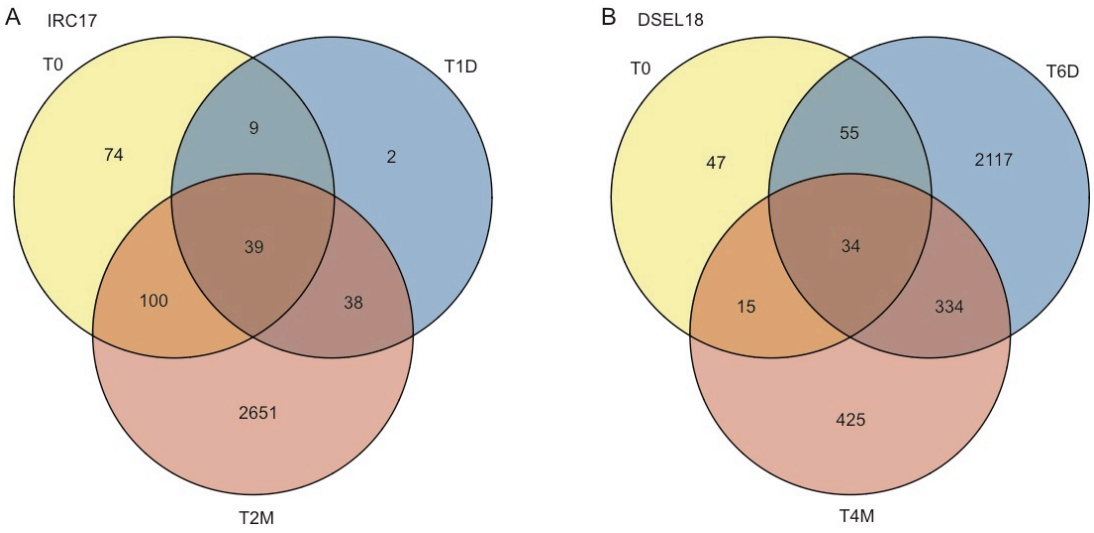


Fig 4.5 Venn diagrams showing statistically significant (padj≤0.05) differentially expressed genes (comparison to the respective TC control for each experiment) and the overlap between incubation duration for IRC17 (a) and DSEL18 (b).

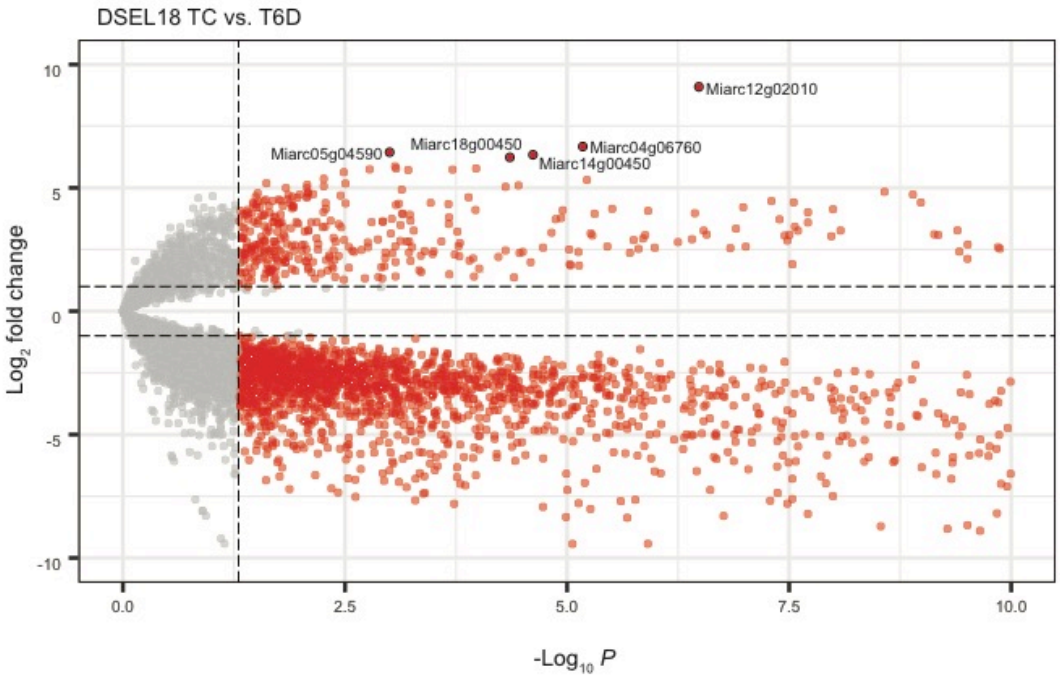
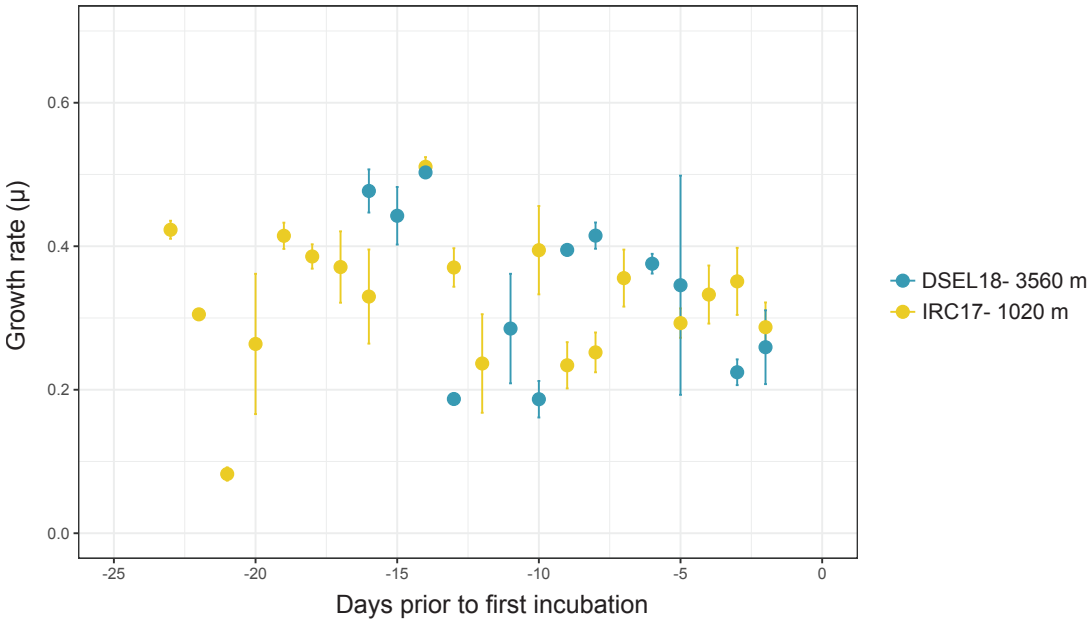
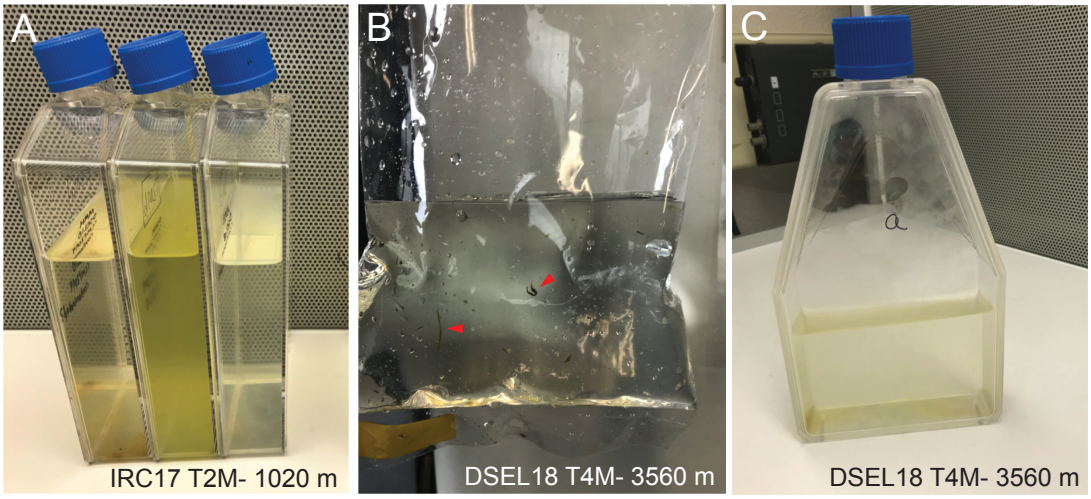


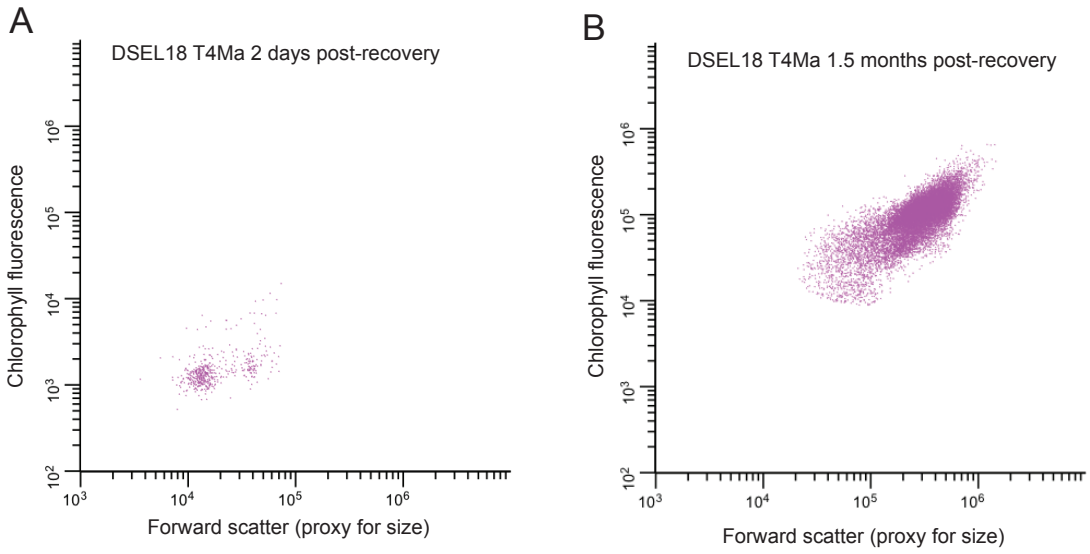
Fig 4.6 Differential gene expression in *M. polaris* during the T6D treatment, compared to the TC control in the DSEL18 experiment. The volcano plot shows fold changes in transcript abundances for genes that are differentially expressed. Statistically significant (padj≤0.05) genes visualized with red circles, while those that do not pass the threshold of significance are grey. A selection of highly up-regulated genes are outlined in black and labeled with the gene name. A description of these genes can be found in Supplementary Table S4.1.



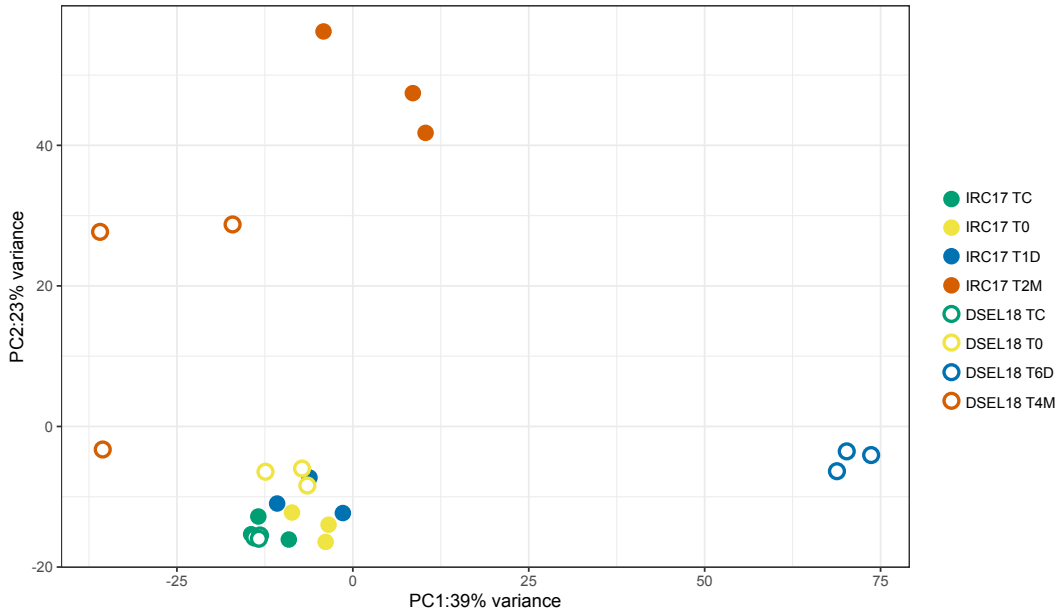
Supplementary Fig S4.1 Daily growth rates of *M. polaris* cultures prior to incubation experiments. Among the IRC17 growth measurement days, the first 8 are represented by two replicates, and the following days by ≥ 4. Among the DSEL18 growth measurement days, the first 4 are represented by two replicates, and the following days by ≥ 4.



Supplementary Fig S4.2 Photographs of *M. polaris* cultures recovered from the deep sea. The first panel shows the three replicates of the IRC17 T2M incubation approximately two months after recovery (a). The second photo (b) was taken of a DSEL18 T4M incubation bag immediately following recovery after observations of dark green filaments (red arrows) were observed inside the bag. The other two replicates of this incubation also exhibited similar filaments (not shown). Growth of recovered DSEL18 T4M cultures was detected visually after an approximately 1.5 month period in the laboratory (during which observations were not made, as it was thought that after the initial few days of observation that the cultures must have arrived back on the ship deck "dead") (c).



Supplementary Fig S4.3 Flow cytometric histograms indicating an absence of cells with a *Micromonas*-like chlorophyll signal 2 days post-recovery (a) and a clear population after a 1.5 months post-recovery (b).



Supplemental Fig S4.4 A PCA including both experiments, IRC17 and DSEL18 based on normalized gene expression. The proportion of variance explained is indicated in parentheses. Each circle represents a single sample and the color is indicative of incubation duration. Filled circles are samples from IRC17, while open circles represent DSEL18 samples.

M. polaris Gene ID	Direction of differential expression	*M. commoda* orthologue?	Protein domains	Putative function/annotation
Miarc14g00450	+	yes	IPR005046	Protein of unknown function DUF285
Miarc04g06760	+	yes	none	Unknown
Miarc05g04590	+	yes	IPR020683, IPR036770, IPR002110	cell wall/membrane/envelope biogenesis
Miarc18g00450	+	no	IPR027417	P-loop NTPase
Miarc14g00450	+	yes	none	Unknown

Supplementary Table S4.1 Statistically significant (padj≤0.05) differentially expressed genes in the T6D incubation compared to the TC control for the DSEL18 experiment.

Please see attached file "mhamilton_dissertation_TableS4.2.xlsx

Supplemental Table S4.2 Meiosis associated genes found in the *M. polaris* genome and their comparison to the *M. commoda* and *M. pusilla* genomes based on data reported in Worden et al. 2009. *The M. commoda* protein sequence for each gene was used as a blastp query against the *M. polaris* genome unless otherwise indicated. The second column indicates if the gene has been observed to be either essential to functional meiosis or specific to the process. The final column indicates any statistically significant (padj≤0.05) differential expression observed compared to the TC control for a particular experiment. NF = not found. * = *Saccharomyces cerevisiae* was used for query in blastp, **= *Clamydomonas reinhardtii* was used for query

Please see attached file "mhamilton_dissertation_TableS4.3.xlsx

Supplemental Table S4.3 *Clamydomonas reinhardtii* meiosis and cell wall associated genes. Orthologous genes found in the *M. polaris* genome are identified in the fourth column. The final column indicates any statistically significant (padj≤0.05) differential expression observed compared to the TC control for a particular experiment. NF = not found

Chapter 5: Conclusions and Perspectives

The phytoplankton that reside in the high latitude oceans of the Arctic and Antarctic are not only integral to the flow of carbon through global biogeochemical cycles, but they also comprise the base of the marine food web in these climate-sensitive environments. Without a baseline characterization of the specific taxa present in polar waters, including the environmental drivers behind their diversity and proliferation, future predictions for a rapidly changing ocean cannot be accurately made. Similarly, an understanding of the physiological capacity to withstand environmental shifts in key algal species is crucial to these predictive efforts, and may also help to answer longstanding questions on ecological and evolutionary processes.

In this dissertation, protistan community diversity was examined at a molecular level throughout the surface waters of an Antarctic fjord, Andvord Bay, in the West Antarctic Peninsula (WAP) during austral spring and fall, and a novel cryptophyte species was identified (Chapter 2). This work focused primarily on the photosynthetic community, where we observed an oscillation between relative cryptophyte and diatom contributions. Interestingly, the diatoms typically consisted of a diverse assemblage of different taxa, while the cryptophytes were dominated by a single species. High absolute abundances of cryptophyte cells, particularly within the inner fjord, were confirmed via flow cytometry and linked to increased temperature and glacial meltwater, indicating that enhanced proliferation of this taxa is likely to be expected with future climate change. Further analyses were performed for this uncultivated cryptophyte to confirm its status as a species separate from the canonical Antarctic cryptophtye, and to better characterize its evolutionary relationship to other

members of the *Teleaulax*/*Plagioselmis*/*Geminigera* (TPG) lineage. The phylogenetic reconstruction revealed a number of distinct clades within the TPG lineage and representatives from each were analyzed against a global dataset. The novel cryptophyte identified in this study was determined to be endemic to Antarctica, in contrast to other TPG cryptophytes with a global distribution. It is thus likely that the novel WAP cryptophyte is highly specialized to this particular niche environment, and future investigations of global cryptophyte biodiversity would benefit from a comparative examination of physiological adaptions across TPG cryptophyte species. A critical first step towards these endeavors would be the establishment, sustainment, and molecular and physiological characterization of WAP cryptophyte cultures, as questions remain regarding the potential for a dimorphic life history. Additionally, a reevaluation of genus level differentiation within the TPG is likely necessary, as the field moves beyond morphology-based taxonomy and hidden cryptophyte dimorphism is uncovered (Altenburger et al. 2020).

Among the taxa identified in Andvord Bay, was the picoprasinophyte, *Micromonas polaris*. Though the results from Chapter 2 of this dissertation found this phytoplankton to be low in relative abundance compared to other taxa, such as the cryptophytes and diatoms, *M. polaris* is considered to be a key component of the phytoplankton community on the other side of the globe in Arctic waters (Lovejoy et al. 2007). This alga has been predicted to be among those whose relative proliferation may be enhanced in the Arctic by climate change (Li et al. 2009), but many of the details regarding the potential cellular response remain to be elucidated. Controlled

laboratory and *in situ* experiments were performed with *M. polaris* in order to examine how this alga responds to increased CO_2 and decreased nitrate (Chapter 3), as well as investigate survivability in the deep sea, where it has been observed (Chapter 4) (Simmons et al. 2015). These two studies included an investigation of both basic physiology, as well as the molecular response to altered environmental conditions, which was aided by the recent sequencing of the *M. polaris* genome.

Results from Chapter 3 indicated that nitrate availability had a much greater effect on the physiology of *M. polaris* than CO_2; with reduced growth rates, cell size, and chlorophyll fluorescence, but increased photosynthetic efficiency under nitrate limitation. While picophytoplankton such as *M. polaris* may be predicted to outcompete larger phytoplankton in a more oligotrophic future ocean, due to the nutrient-uptake benefits of an increased surface area to volume ratio, our data shows that this organism is not immune to the negative effects of nutrient limitation. On the other hand, increases in CO_2 did not alter most aspects of *M. polaris* physiology, including growth rates, in either replete or limiting nitrate conditions. An examination of gene expression shed light on the potential underlying molecular response during these treatments. One of the key findings was that genes identified to be associated with the biophysical CO_2 concentrating mechanism (CCM) were actually only responsive to nitrate and not CO_2, despite speculation that increased CO_2 availability might reduce the need for the CCM. While the specific CCM genes examined here are unable to fully explain the potential link between inorganic carbon assimilation and nitrate availability, it is one that could be further explored within this dataset.

Specifically, the full suite of genes associated with Rubisco-assembly and functioning could be identified within the *M. polaris* genome and their expression examined. Additionally, the biophysical CCM is not the only potential CCM- genes associated with a C4 metabolism have been identified in the non-polar *Micromonas* genomes, but they have not yet been searched for in *M. polaris* or examined for CO_2-responsiveness in a green alga. Overall, the dataset produced in this chapter has the potential to provide many avenues of continued exploration of the molecular activity of *M. polaris* in relevant environmental conditions.

The *in situ* incubation experiments performed in Chapter 4 confirm the extraordinary resilience of *M. polaris* in the deep sea, which has been speculated to be the key to their bi-polar distribution (Simmons et al. 2015). Recovery growth was observed for this phytoplankton after incubations of up to four months at 3,560 m depths. Qualitative data, such as the observation of cell aggregates and highly delayed recovery growth within seemingly "dead" cultures, pointed towards the possibility of deep-sea survival via cyst formation. An examination of gene expression confirmed the viability of these cells and indicated a strong response to extended periods in the deep. Genes associated with meiosis were identified in the *M. polaris* genome and their expression examined, as cyst formation is linked to sexual reproduction in other green algae (Goodenough et al. 2007), but there were no clear indicators of increased *M. polaris* meiotic activity based on expression. Genes encoding hydroxyproline-rich glycoproteins (HRGP) known to specifically be components of the *Chlamydomonas*

reinhardtii cell wall were not identified in the *M. polaris* genome, but the full suite of HRGPs in *M. polaris* remains to be identified and their expression examined.

Future investigations into the possibility of cyst-formation in *M. polaris* could be enhanced by first fully characterizing this process in *Pyramimonas amylifera,* a polar green alga where encystment in response to light availability has been visually observed (Hargraves and Gardiner 1980). Similar to Chapter 3, the dataset generated in Chapter 4 can also be further explored from a variety of different angles. For example, all genes that were differentially expressed in all extended incubations could be classified and functionally described, in order to paint a more complete picture of the processes occurring when a phytoplankton is transported to the deep sea. Ultimately, having a fully functionally annotated *M. polaris* genome will greatly enhance our ability to draw conclusions based on global expression profiles, and automated functional annotations are currently underway.

The work collectively presented in this dissertation provides baseline information on polar phytoplankton diversity and the environmental drivers of community structure in a climate-sensitive habitat, and explores the physiological capabilities of a model polar alga. Together, the combination of oceanographic sampling and laboratory and *in situ* experimentation allows for a variety of important questions regarding the ecology and life histories of polar phytoplankton to be explored.

References

Abele D, Vazquez S, Buma AGJ, Hernandez E, Quiroga C, Held C, Frickenhaus S, Harms L, Lopez JL, Helmke E, et al. 2017. Pelagic and benthic communities of the Antarctic ecosystem of Potter Cove: Genomics and ecological implications. *Mar. Genomics* 33:1–11.

Agrawal SC. 2009. Factors Affecting Spore Germination in Algae. *Folia Microbiol. (Praha).* 54:273–302.

Agusti S, González-Gordillo JI, Vaqué D, Estrada M, Cerezo MI, Salazar G, Gasol JM, Duarte CM. 2015. Ubiquitous healthy diatoms in the deep sea confirm deep carbon injection by the biological pump. *Nat. Commun.* 6:1–8.

Altenburger A, Blossom HE, Garcia-Cuetos L, Jakobsen HH, Carstensen J, Lundholm N, Hansen PJ, Moestrup, Haraguchi L. 2020. Dimorphism in cryptophytes-The case of Teleaulax amphioxeia/Plagioselmis prolonga and its ecological implications. *Sci. Adv.* 6:1–9.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.

Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA Genes. *PLoS One* 4:1–9.

Anders S, Pyl PT, Huber W. 2015. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.

Anderson DM, Lindquist NL. 1985. Time-course measurements of phosphorus depletion and cyst formation in the dinoflagellate Gonyaulax tamarensis Lebour. *J. Exp. Mar. Bio. Ecol.* 86:1–13.

De Araujo C, Arefeen D, Tadesse Y, Long BM, Price GD, Rowlett RS, Kimber MS, Espie GS. 2014. Identification and characterization of a carboxysomal γ-carbonic anhydrase from the cyanobacterium Nostoc sp. PCC 7120. *Photosynth. Res.* 121:135–150.

Arrigo KR, van Dijken G, Pabi S. 2008. Impact of a shrinking Arctic ice cover on marine primary production. *Geophys. Res. Lett.* 35:L19603.

Aspatwar A, Haapanen S, Parkkila S. 2018. An update on the metabolic roles of carbonic anhydrases in the model alga chlamydomonas reinhardtii. *Metabolites* 8.

Bach LT, Alvarez-Fernandez S, Hornick T, Stuhr A, Riebesell U. 2017. Simulated ocean acidification reveals winners and losers in coastal phytoplankton. *PLoS*

One 12:1–22.

Bachy C, López-García P, Vereshchaka A, Moreira D. 2011. Diversity and vertical distribution of microbial eukaryotes in the snow, sea ice and seawater near the North Pole at the end of the polar night. *Front. Microbiol.* 2:106.

Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD. 1998. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO2-concentrating mechanisms in algae. *Can. J. Bot.* 76:1052–1071.

Balzer I. 1996. Encystment of Gonyaulax polyedra: Dependence on Light. *Biol. Rhythm Res.* 27:386–389.

Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, Stamatakis A. 2019. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. *Syst. Biol.* 68:365–369.

Bauwe H, Hagemann M, Fernie AR. 2010. Photorespiration: players, partners and origin. *Trends Plant Sci.* 15:330–336.

Beardall J, Griffiths H, Raven JA. 1982. Carbon isotope discrimination and the CO2 accumulating mechanism in Chlorella emersonii. *J. Exp. Bot.* 33:729–737.

Belevich TA, Ilyash L V., Milyutina IA, Logacheva MD, Goryunov D V., Troitsky A V. 2018. Photosynthetic Picoeukaryotes in the Land-Fast Ice of the White Sea, Russia. *Microb. Ecol.* 75:582–597.

Bell EM, Laybourn-Parry J. 2003. Mixotrophy in the antarctic phytoflagellate, Pyramimonas gelidicola (Chlorophyta: Prasinophyceae). *J. Phycol.* 39:644–649.

Berge J, Renaud PE, Darnis G, Cottier F, Last K, Gabrielsen TM, Johnsen G, Seuthe L, Weslawski JM, Leu E, et al. 2015. In the dark: A review of ecosystem processes during the Arctic polar night. *Prog. Oceanogr.* 139:258–271.

Biggs TEG, Alvarez-Fernandez S, Evans C, Mojica KDA, Rozema PD, Venables HJ, Pond DW, Brussaard CPD. 2019. Antarctic phytoplankton community composition and size structure: importance of ice type and temperature as regulatory factors. *Polar Biol.* 42:1997–2015.

Binder BJ, Anderson DM. 1990. Biochemical composition and metabolic activity of Scrippsiella Trochoidea (Dinophyceae) resting cysts. *J. Phycol.* 26:289–298.

Bolaños LM, Karp-Boss L, Choi CJ, Worden AZ, Graff JR, Haëntjens N, Chase AP, Della Penna A, Gaube P, Morison F, et al. 2020. Small phytoplankton dominate western North Atlantic biomass. *ISME J.* 14:1663–1674.

Bolch CJS, de Salas MF. 2007. A review of the molecular evidence for ballast water introduction of the toxic dinoflagellates Gymnodinium catenatum and the

Alexandrium “tamarensis complex” to Australasia. *Harmful Algae* 6:465–485.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37:852–857.

Borkhsenious ON, Mason CB, Moroney J V. 1998. The Intracellular Localization of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase in Chlamydomonas reinhardtii. *Plant Physiol.* 116:1585–1591.

Bravo I, Figueroa R. 2014. Towards an Ecological Understanding of Dinoflagellate Cyst Functions. *Microorganisms* 2:11–32.

Brueggeman AJ, Gangadharaiah DS, Cserhati MF, Casero D, Weeks DP, Ladungab I. 2012. Activation of the carbon concentrating mechanism by CO2 deprivation coincides with massive transcriptional restructuring in Chlamydomonas reinhardtii. *Plant Cell* 24:1860–1875.

Brussaard CPD, Noordeloos AAM, Witte H, Collenteur MCJ, Schulz K, Ludwig A, Riebesell U. 2013. Arctic microbial community dynamics influenced by elevated CO2 levels. *Biogeosciences* 10:719–731.

Calleja ML, Kerhervé P, Bourgeois S, Kędra M, Leynaert A, Devred E, Babin M, Morata N. 2017. Effects of increase glacier discharge on phytoplankton bloom dynamics and pelagic geochemistry in a high Arctic fjord. *Prog. Oceanogr.* 159:195–210.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973.

de Carpentier F, Lemaire SD, Danon A. 2019. When Unity Is Strength: The Strategies Used by Chlamydomonas to Survive Environmental Stresses. *Cells* 8:1–16.

Choi CJ, Jimenez V, Needham DM, Poirier C, Bachy C, Alexander H, Wilken S, Chavez FP, Sudek S, Giovannoni SJ, et al. 2020. Seasonal and Geographical Transitions in Eukaryotic Phytoplankton Community Structure in the Atlantic and Pacific Oceans. *Front. Microbiol.* 11:542372.

Cuvelier ML, Allen AE, Monier A, McCrow JP, Messie M, Tringe SG, Woyke T, Welsh RM, Ishoey T, Lee J-H, et al. 2010. Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc. Natl. Acad. Sci.* 107:14679–14684.

Daims H, Brühl A, Amann R, Schleifer KH, Wagner M. 1999. The domain-specific probe EUB338 is insufficient for the detection of all bacteria: Development and

evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol.* 22:434–444.

Daugbjerg N, Marchant HJ, Thomsen H a. 2000. Life history stages of *Pyramimonas tychotreta* (Prasinophyceae, Chlorophyta), a marine flagellate from the Ross Sea, Antarctica. *Phycol. Res.* 48:199–209.

Daugbjerg N, Norlin A, Lovejoy C. 2018. Baffinella frigidus gen. et sp. nov. (Baffinellaceae fam. nov., Cryptophyceae) from Baffin Bay: Morphology, pigment profile, phylogeny, and growth rate response to three abiotic factors. *J. Phycol.* 54:665–680.

Delmont TO, Murat Eren A, Vineis JH, Post AF. 2015. Genome reconstructions indicate the partitioning of ecological functions inside a phytoplankton bloom in the Amundsen Sea, Antarctica. *Front. Microbiol.* 6:1090.

Demir-Hilton E, Sudek S, Cuvelier ML, Gentemann CL, Zehr JP, Worden AZ. 2011. Global distribution patterns of distinct clades of the photosynthetic picoeukaryote Ostreococcus. *ISME J.* 5:1095–1107.

Dickson R, Brown J. 1994. The production of North Atlantic Deep Water : Sources , rates , and pathways. *J. Geophys. Res.* 99:12319–12341.

DiMario RJ, Machingura MC, Waldrop GL, Moroney J V. 2018. The many types of carbonic anhydrases in photosynthetic organisms. *Plant Sci.* 268:11–17.

Doney SC, Ruckelshaus M, Emmett Duffy J, Barry JP, Chan F, English CA, Galindo HM, Grebmeier JM, Hollowed AB, Knowlton N, et al. 2012. Climate Change Impacts on Marine Ecosystems. *Ann. Rev. Mar. Sci.* 4:11–37.

Duanmu D, Bachy C, Sudek S, Wong C, Jiménez V, Rockwell NC. 2014. Marine algae and land plants share conserved phytochrome signaling systems. 111.

Ducklow HW, Baker K, Martinson DG, Quetin LB, Ross RM, Smith RC, Stammerjohn SE, Vernet M, Fraser W. 2007. Marine pelagic ecosystems: the West Antarctic Peninsula. *Philos. Trans. R. Soc. B Biol. Sci.* 362:67–94.

Dutkiewicz S, Morris JJ, Follows MJ, Scott J, Levitan O, Dyhrman ST, Berman-Frank I. 2015. Impact of ocean acidification on the structure of future phytoplankton communities. *Nat. Clim. Chang.* 5:1002–1006.

Edlund MB, Stoermer EF. 1997. Ecological, evolutionary, and systematic significance of diatom life histories. *J. Phycol.* 918:897–918.

Ekern L. 2017. Assessing seasonal primary production in Andvord Bay, Antarctica.

Ellegaard M, Ribeiro S. 2018. The long-term persistence of phytoplankton resting stages in aquatic 'seed banks.' *Biol. Rev.* 93:166–183.

Ellegaard M, Ribeiro S, Lundholm N, Andersen TJ, Berge T, Ekelund F, Härnström

K, Godhe A. 2013. Using the sediment archive of living dinoflagellate cysts and other protist resting stages to study temporal population dynamics. *Biol. Geol. Perspect. Dinoflag.* TMS Specia:149–153.

Falkowski PG. 1997. Evolution of the nitrogen cycle and its influence on the biological sequestration of CO2 in the ocean. *Nature* 1:271–294.

Figueroa RI, Bravo I, Garcés E. 2005. Effects of nutritional factors and different parental crosses on the encystment and excystment of Alexandrium catenella (Dinophyceae) in culture. *Phycologia* 44:658–670.

Fuentes S, Arroyo JI, Rodríguez-Marconi S, Masotti I, Alarcón-Schumacher T, Polz MF, Trefault N, De la Iglesia R, Díez B. 2019. Summer phyto- and bacterioplankton communities during low and high productivity scenarios in the Western Antarctic Peninsula. *Polar Biol.* 42:159–169.

Garibotti IA, Vernet M, Ferrario ME. 2005. Annually recurrent phytoplanktonic assemblages during summer in the seasonal ice zone west of the Antarctic Peninsula (Southern Ocean). *Deep. Res. Part I Oceanogr. Res. Pap.* 52:1823–1841.

Gast RJ, Moran DM, Beaudoin DJ, Blythe JN, Dennett MR, Caron DA. 2006. Abundance of a novel dinoflagellate phylotype in the Ross Sea, Antarctica. *J. Phycol.* 42:233–242.

Gast RJ, Moran DM, Dennett MR, Caron DA. 2007. Kleptoplasty in an Antarctic dinoflagellate: Caught in evolutionary transition? *Environ. Microbiol.* 9:39–45.

Goodenough U, Lin H, Lee JH. 2007. Sex determination in Chlamydomonas. *Semin. Cell Dev. Biol.* 18:350–361.

Grange LJ, Smith CR. 2013. Megafaunal communities in rapidly warming fjords along the West Antarctic Peninsula: Hotspots of abundance and beta diversity. *PLoS One* 8.

Grebmeier JM, Moore SE, Overland JE, Frey KE, Gradinger R. 2010. Biological response to recent pacific arctic sea ice retreats. *Eos (Washington. DC)*. 91:161–162.

Grzebyk D, Berland B. 1996. Influences of temperature , salinity and irradiance on growth of Prorocentrum minimum (Dinophyceae) from the Mediterranean. 18:1837–1849.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, De Vargas C, Decelle J, et al. 2013. The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41:597–604.

Guillou L, Medlin LK, Oce O, Vaulot D. 1999. Bolidomonas : A new genus with two

species belonging to a new algal class, the Bolidophyceae (Heterokonta). *J. Phycol.* 381:368–381.

Guo J, Wilken S, Jimenez V, Choi CJ, Ansong C, Dannebaum R, Sudek L, Milner DS, Bachy C, Reistetter EN, et al. 2018. Specialized proteomic responses and an ancient photoprotection mechanism sustain marine green algal growth during phosphate limitation. *Nat. Microbiol.* 3:781–790.

Hansen J, Ruedy R, Sato M, Lo K. 2010. Global surface temperature change. *Rev. Geophys.* 48:1–29.

Hargraves PE, Gardiner WE. 1980. The life history of Pyramimonas amylifera Conrad (Prasinophyceae). *J. Plankton Res.* 2:99–108.

Head MJ. 1996. Chapter 30. Modern dinoflagellate cysts and their biological affinities. *Palynol. Princ. Appl.* 3:1197–1248.

Hennon GM, Limón MDH, Haley ST, Juhl AR, Dyhrman ST. 2017. Diverse CO2-induced responses in physiology and gene expression among eukaryotic phytoplankton. *Front. Microbiol.* 8:1–14.

Hennon GMM, Quay P, Morales RL, Swanson LM, Virginia Armbrust E. 2014. Acclimation conditions modify physiological response of the diatom Thalassiosira pseudonana to elevated CO2 concentrations in a nitrate-limited chemostat. *J. Phycol.* 50:243–253.

Hoef-Emden K. 2008. Molecular phylogeny of phycocyanin-containing cryptophytes: Evolution of biliproteins and geographical distribution. *J. Phycol.* 44:985–993.

Hoegh-Guldberg O, Bruno J. 2010. The Impact of Climate Change on the. *Ecol. Res.* 328:1523–1528.

van den Hoff J, Bell E, Whittock L. 2020. Dimorphism in the Antarctic cryptophyte Geminigera cryophila (Cryptophyceae) . *J. Phycol.*:1–11.

van den Hoff J, Burton HR, Vesk M. 1989. An encystment stage, bearing a new scale type of the antartic prasinophyte Pyramimonas gelidicola and its paleolimnological and taxonomic significance. *J. Phycol.* 25:446–454.

Hoppe CJM, Flintrop CM, Rost B. 2018. The arctic picoeukaryote micromonas pusilla benefits synergistically from warming and ocean acidification. *Biogeosciences* 15:4353–4365.

Howe JA, Austin WEN, Forwick M, Paetzel M, Harland REX, Cage AG. 2010. Fjord systems and archives : a review. *Geol. Soc. London, Spec. Publ.* 344:5–15.

Huang WJ, Wang Y, Cai WJ. 2012. Assessment of sample storage techniques for total alkalinity and dissolved inorganic carbon in seawater. *Limnol. Oceanogr. Methods* 10:711–717.

Ibarbalz FM, Henry N, Brandão MC, Martini S, Busseni G, Byrne H, Coelho LP, Endo H, Gasol JM, Gregory AC, et al. 2019. Global Trends in Marine Plankton Diversity across Kingdoms of Life. *Cell* 179:1084-1097.e21.

Ichinomiya M, Santos A., Yoshikawa S, Kamiya M, Ohki K, Audic S, de Vargas C, Vaulot D, Kuwata A. 2016. Diversity and oceanic distribution of Parmales and Bolidophytes, a group closely related to diatoms. In review at ISME journal. *ISME J.* 10:2419–2434.

Junier T, Zdobnov EM. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26:1669–1670.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* 30:772–780.

Keeling PJ, Burki F. 2019. Progress towards the Tree of Eukaryotes. *Curr. Biol.* 29:R808–R817.

Kilias E, Kattner G, Wolf C, Frickenhaus S, Metfies K. 2014. A molecular survey of protist diversity through the central Arctic Ocean. *Polar Biol.* 37:1271–1287.

Kim D, Langmead B, Salzberg SL. 2015. HHS Public Access. *Nat. Methods* 12:357–360.

King AL, Sãudo-Wilhelmy SA, Leblanc K, Hutchins DA, Fu F. 2011. CO2and vitamin B12interactions determine bioactive trace metal requirements of a subarctic Pacific diatom. *ISME J.* 5:1388–1396.

Kirst GO, Wiencke C. 1995. Ecophysiology of Polar Algae. *J. Phycol.* 31:181–199.

Kubiszyn AM, Piwosz K, Wiktor JM, Wiktor JM. 2014. The effect of inter-annual Atlantic water inflow variability on the planktonic protist community structure in the West Spitsbergen waters during the summer. *J. Plankton Res.* 36:1190–1203.

Kuwata A, Hama T, Takahashi M. 1993. Ecophysiological characterization of two life forms, resting spores and resting cells, of a marine planktonic diatom, Chaetoceros pseudocurvisetus, formed under nutrient depletion. *Mar. Ecol. Prog. Ser.* 102:245–256.

Laza-Martínez A, Arluzea J, Miguel I, Orive E. 2012. Morphological and molecular characterization of Teleaulax gracilis sp. nov. and T. minuta sp. nov. (Cryptophyceae). *Phycologia* 51:649–661.

Li W, Gao K, Beardall J. 2012. Interactive Effects of Ocean Acidification and Nitrogen-Limitation on the Diatom Phaeodactylum tricornutum. *PLoS One* 7:1–8.

Li WKW, Mclaughlin FA, Lovejoy C, Carmack EC. 2009. Smallest Algae Thrive As the Arctic. 326.

Lin Y, Cassar N, Marchetti A, Moreno C, Ducklow H, Li Z. 2017. Specific eukaryotic plankton are good predictors of net community production in the Western Antarctic Peninsula. *Sci. Rep.* 7:1–11.

Lin YC, Campbell T, Chung CC, Gong GC, Chiang KP, Worden AZ. 2012. Distribution patterns and phylogeny of marine stramenopiles in the North Pacific Ocean. *Appl. Environ. Microbiol.* 78:3387–3399.

López-García P, Rodrguez-Valera F, Pedrós-Alió C, Moreira D. 2002. Unexpected Diversity of Smal Eukaryotes in Deep-Sea Antartic Plankton. *Nature* 409:603–606.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:1–21.

Lovejoy C, Galand PE, Kirchman DL. 2011. Picoplankton diversity in the Arctic Ocean and surrounding seas. *Mar. Biodivers.* 41:5–12.

Lovejoy C, Massana R, Pedro C. 2006. Diversity and Distribution of Marine Microbial Eukaryotes in the Arctic Ocean and Adjacent Seas Diversity and Distribution of Marine Microbial Eukaryotes in the Arctic Ocean and Adjacent Seas. *Appl. Environ. Microbiol.* 72:3085–3095.

Lovejoy C, Vincent WF, Bonilla S, Roy S, Martineau MJ, Terrado R, Potvin M, Massana R, Pedrós-Alió C. 2007. Distribution, phylogeny, and growth of cold-adapted picoprasinophytes in arctic seas. *J. Phycol.* 43:78–89.

Lundesgaard Ø, Winsor P, Truffer M, Merrifield M, Powell B, Statscewich H, Eidam E, Smith CR. 2020. Hydrography and energetics of a cold subpolar fjord: Andvord Bay, western Antarctic Peninsula. *Prog. Oceanogr.* 181:102224.

Lundholm N, Ribeiro S, Andersen TJ, Koch T, Godhe A, Ekelund F, Ellegaard M. 2011. Buried alive – germination of up to a century-old marine protist resting stages. *Phycologia* [Internet] 50:629–640. Available from: http://www.phycologia.org/doi/abs/10.2216/11-16.1

Luo W, Li H, Gao S, Yu Y, Lin L, Zeng Y. 2016. Molecular diversity of microbial eukaryotes in sea water from Fildes Peninsula, King George Island, Antarctica. *Polar Biol.* 39:605–616.

Luria CM, Ducklow HW, Amaral-Zettler LA. 2014. Marine bacterial, archaeal and eukaryotic diversity and community structure on the continental shelf of the western Antarctic Peninsula. *Aquat. Microb. Ecol.* 73:107–121.

Ma Y, Pollock S V., Xiao Y, Cunnusamy K, Moroney J V. 2011. Identification of a novel gene, CIA6, required for normal pyrenoid formation in chlamydomonas reinhardtii. *Plant Physiol.* 156:884–896.

Malviya S, Scalco E, Audic S, Vincent F, Veluchamy A, Poulain J, Wincker P,

Iudicone D, De Vargas C, Bittner L, et al. 2016. Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. U. S. A.* 113:E1516–E1525.

Manton I, Parke M. 1960. Further observations on small green flagellates with special reference to possible relatives of chromulina pusilla butcher. *J. Mar. Biol. Assoc. United Kingdom* 39:275–298.

Mascioni M, Almandoz GO, Cefarelli AO, Cusick A, Ferrario ME, Vernet M. 2019. Phytoplankton composition and bloom formation in unexplored nearshore waters of the western Antarctic Peninsula. *Polar Biol.* 42:1859–1872.

Massana R, Del Campo J, Sieracki ME, Audic S, Logares R. 2014. Exploring the uncultured microeukaryote majority in the oceans: Reevaluation of ribogroups within stramenopiles. *ISME J.* 8:854–866.

Mcminn a, Martin A. 2013. Dark survival in a warming world Dark survival in a warming world. *Proc. R. Soc. B* 280:201229.

McQuoid MR, Hobson LA. 1996. Diatom Resting Stages. *J. Phycol.* 32:889–902.

Meakin NG, Wyman M. 2011. Rapid shifts in picoeukaryote community structure in response to ocean acidification. *ISME J.* 5:1397–1405.

Menden-Deuer S, Lessard EJ. 2000. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45:569–579.

Mendes CRB, Tavano VM, Dotto TS, Kerr R, de Souza MS, Eiras Garcia CA, Secchi ER. 2018. New insights on the dominance of cryptophytes in Antarctic coastal waters: a case study in Gerlache Strait. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 149:161–170.

Mendes CRB, Tavano VM, Leal MC, de Souza MS, Brotas V, Garcia CAE. 2013. Shifts in the dominance between diatoms and cryptophytes during three late summers in the Bransfield Strait (Antarctic Peninsula). *Polar Biol.* 36:537–547.

Meyer M, Griffiths H. 2013. Origins and diversity of eukaryotic CO2-concentrating mechanisms: Lessons for the future. *J. Exp. Bot.* 64:769–786.

Michaels AFAMWS. 1988. Primary production, sinking fluxes and the microbial food web. *Deep. Res.* 35:473–490.

Mitra M, Mason CB, Xiao Y, Ynalvez RA, Lato SM, Moroney J V. 2005. The carbonic anhydrase gene families of Chlamydomonas reinhardtii. *Can. J. Bot.* 83:780–795.

Mock T, Otillar RP, Strauss J, McMullan M, Paajanen P, Schmutz J, Salamov A, Sanges R, Toseland A, Ward BJ, et al. 2017. Evolutionary genomics of the cold-

Adapted diatom Fragilariopsis cylindrus. *Nature* 541:536–540.

Monier A, Worden AZ, Richards TA. 2016. Phylogenetic diversity and biogeography of the Mamiellophyceae lineage of eukaryotic phytoplankton across the oceans. *Environ. Microbiol. Rep.* 8:461–469.

Moon-Van Der Staay SY, Van Der Staay GWM, Guillou L, Vaulot D, Claustre H, Médlin LK. 2000. Abundance and diversity of prymnesiophytes in the picoplankton community from the equatorial Pacific Ocean inferred from 18S rDNA sequences. *Limnol. Oceanogr.* 45:98–109.

Morita E, Abe T, Tsuzuki M, Fujiwara S, Sato N, Hirata A, Sonoike K, Nozaki H. 1999. Role of pyrenoids in the CO2-concentrating mechanism: Comparative morphology, physiology and molecular phylogenetic analysis of closely related strains of Chlamydomonas and Chloromonas (Volvocales). *Planta* 208:365–372.

Naduvinamani NB, Marali GB. 2008. Dynamic Reynolds equation for micropolar fluid lubrication of porous slider bearings. *J. Mar. Sci. Technol.* 16:182–190.

Needham DM, Fuhrman J a. 2016. Pronounced daily succession of phytoplankton , archaea and bacteria following a spring bloom. *Nat. Microbiol.* 1:16005.

Needham DM, Sachdeva R, Fuhrman JA. 2017. Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. *ISME J.* 11:1614–1629.

Olson RJ, Zettler ER, Anderson OK. 1989. Discrimination of eukaryotic phytoplankton cell types from light scatter and autofluorescence properties measured by flow cytometry. *Cytometry* 10:636–643.

Paluszkiewicz T, Garwood RW, Denbo DW. 1994. Deep convective plumes in the ocean. *Oceanography*:37–44.

Pan B, Vernet V, Manck L, Forsch K, Ekern L, Mascioni M, Barbeau KA, Almandoz G, Orona AJ. 2020. Environmental drivers on phytoplankton taxonomic composition in an Antarctic fjord. *Prog. Oceanogr.*

Piątek J, Piątek M, Zeeb BA, El Shahed A. 2009. Chrysophyte Stomatocysts in Africa: The First Description of an Assemblage in The Recent Sediments of a Thermo-Mineral Spring in Egypt. *Phycologia* 48:13–23.

Piquet AMT, Van De Poll WH, Visser RJW, Wiencke C, Bolhuis H, Buma AGJ. 2014. Springtime phytoplankton dynamics in Arctic Krossfjorden and Kongsfjorden (Spitsbergen) as a function of glacier proximity. *Biogeosciences* 11:2263–2279.

Piwosz K, Walkusz W, Hapter R, Wieczorek P, Hop H, Wiktor J. 2009. Comparison of productivity and phytoplankton in a warm (Kongsfjorden) and a cold (Hornsund) Spitsbergen fjord in mid-summer 2002. *Polar Biol.* 32:549–559.

Ploug H, Iversen MH, Fischer G. 2008. Ballast, sinking velocity, and apparent diffusivity within marine snow and zooplankton fecal pellets: Implications for substrate turnover by attached bacteria. *Limnol. Oceanogr.* 53:1878–1886.

van de Poll WH, Abdullah E, Visser RJW, Fischer P, Buma AGJ. 2019. Taxon-specific dark survival of diatoms and flagellates affects Arctic phytoplankton composition during the polar night and early spring. 9999:1–12.

Polovina JJ, Howell E a., Abecassis M. 2008. Ocean's least productive waters are expanding. *Geophys. Res. Lett.* 35:2–6.

Qi D, Chen L, Chen B, Gao Z, Zhong W, Feely RA, Anderson LG, Sun H, Chen J, Chen M, et al. 2017. Increase in acidifying water in the western Arctic Ocean. *Nat. Clim. Chang.* 7:195–199.

Ribeiro S, Amorim A, Andersen TJ, Abrantes F, Ellegaard M. 2012. Reconstructing the history of an invasion: The toxic phytoplankton species Gymnodinium catenatum in the Northeast Atlantic. *Biol. Invasions* 14:969–985.

Ribeiro S, Berge T, Lundholm N, Andersen TJ, Abrantes F, Ellegaard M. 2011. Phytoplankton growth after a century of dormancy illuminates past resilience to catastrophic darkness. *Nat. Commun.* 2.

Rodriguez F, Varela M, Zapata M. 2002. Phytoplankton assemblages in the Gerlache and Bransfield Straits (Antarctic Peninsula) determined by light microscopy and CHEMTAX analysis of HPLC pigment data. *Deep. Res. Part II Top. Stud. Oceanogr.* 49:723–747.

Rokitta SD, John U, Rost B. 2012. Ocean Acidification Affects Redox-Balance and Ion-Homeostasis in the Life-Cycle Stages of Emiliania huxleyi. *PLoS One* 7.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539–542.

Round FE, Crawford DW, Mann DG. 1990. The Diatoms. Biology and Morphology of the Genera. Cambridge: Cambridge University Press

Rozema PD, Biggs T, Sprong PAA, Buma AGJ, Venables HJ, Evans C, Meredith MP, Bolhuis H. 2017. Summer microbial community composition governed by upper-ocean stratification and nutrient availability in northern Marguerite Bay, Antarctica. *Deep. Res. Part II Top. Stud. Oceanogr.* 139:151–166.

Schaum E, Rost B, Millar AJ, Collins S. 2013. Variation in plastic responses of a globally distributed picoplankton species to ocean acidification. *Nat. Clim. Chang.* 3:298–302.

Schlichting HE. 1960. The Role of Waterfowl in the Dispersal of Algae. *Trans. Am.*

Microsc. Soc. 79:160–166.

Schofield O, Saba G, Coleman K, Carvalho F, Couto N, Ducklow H, Finkel Z, Irwin A, Kahl A, Miles T, et al. 2017. Decadal variability in coastal phytoplankton community composition in a changing West Antarctic Peninsula. *Deep. Res. Part I Oceanogr. Res. Pap.* 124:42–54.

Schulz KG, Bach LT, Bellerby RGJ, Bermúdez R, Büdenbender J, Boxhammer T, Czerny J, Engel A, Ludwig A, Meyerhöfer M, et al. 2017. Phytoplankton blooms at increasing levels of atmospheric carbon dioxide: Experimental evidence for negative effects on prymnesiophytes and positive on small picoeukaryotes. *Front. Mar. Sci.* 4.

Selz V, Lowry KE, Lewis KM, Joy-Warren HL, Van De Poll W, Nirmel S, Tong A, Arrigo KR. 2018. Distribution of Phaeocystis antarctica-dominated sea ice algal communities and their potential to seed phytoplankton across the western Antarctic Peninsula in spring. *Mar. Ecol. Prog. Ser.* 586:91–112.

Siaud N, Dray E, Gy I, Gérard E, Takvorian N, Doutriaux MP. 2004. Brca2 is involved in meiosis in Arabidopsis thaliana as suggested by its interaction with Dmc1. *EMBO J.* 23:1392–1401.

Simmons MP, Bachy C, Sudek S, van Baren MJ, Sudek L, Ares M, Worden AZ. 2015. Intron invasions trace algal speciation and reveal nearly identical Arctic and Antarctic Micromonas populations. *Mol. Biol. Evol.* 32:2219–2235.

Slapeta. 2006. Global dispersal and ancient cryptic species in the smallest marine eukaryotes. *Mol. Biol. Evol.* 23:23–29.

Stamatakis A. 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

Stamatakis A. 2015. Using RAxML to Infer Phylogenies. *Curr. Protoc. Bioinformatics* 51:6.14.1-6.14.14.

Steinacher M, Joos F, Frölicher TL, Bopp L, Cadule P, Cocco V, Doney SC, Gehlen M, Lindsay K, Moore JK, et al. 2010. Projected 21st century decrease in marine productivity: A multi-model analysis. *Biogeosciences* 7:979–1005.

Sudek S, Everroad RC, Gehman ALM, Smith JM, Poirier CL, Chavez FP, Worden AZ. 2015. Cyanobacterial distributions along a physico-chemical gradient in the Northeastern Pacific Ocean. *Environ. Microbiol.* 17:3692–3707.

Sun J, Liu D. 2003. Geometric models for calculating cell biovolume and surface area for phytoplankton. *J. Plankton Res.* 25:1331–1346.

Suzuki L, Woessner JP, Carolina N, Uchida H, Waffenschmidt S, Goodenough UW. 2000. A zygote-specific protein with hydroxyproline-rich glycoprotein domains

and lectin-like domains involved in the assembly of the cell wall of Chlamydomonas reinhardtii (Chlrophyta). 583:571–583.

Tajima N, Saitoh K, Sato S, Maruyama F, Ichinomiya M, Yoshikawa S, Kurokawa K, Ohta H, Tabata S, Kuwata A, et al. 2016. Sequencing and analysis of the complete organellar genomes of Parmales, a closely related group to Bacillariophyta (diatoms). *Curr. Genet.* 62:887–896.

Taylor DL, Lee CC. 1971. A new cryptomonad from antarctica: Cryptomonas cryophila sp. nov. *Arch. Mikrobiol.* 75:269–280.

Theriot EC, Ashworth M, Ruck E, Nakov T, Jansen RK. 2010. A preliminary multi gene phylogeny of the diatoms (Bacillariophyta): Challenges for future research. *Plant Ecol. Evol.* 143:278–296.

Thiele D, Chester ET, Moore SE, Širovic A, Hildebrand JA, Friedlaender AS. 2004. Seasonal variability in whale encounters in the Western Antarctic Peninsula. *Deep. Res. Part II Top. Stud. Oceanogr.* 51:2311–2325.

Timothy Pennington J, Chavez FP. 2000. Seasonal fluctuations of temperature, salinity, nitrate, chlorophyll and primary production at station H3/M1 over 1989-1996 in Monterey Bay, California. *Deep. Res. Part II Top. Stud. Oceanogr.* 47:947–973.

Torstensson A, Dinasquet J, Chierici M, Fransson A, Riemann L, Wulff A. 2015. Physicochemical control of bacterial and protist community composition and diversity in Antarctic sea ice. *Environ. Microbiol.* 17:3869–3881.

Trefaut N, De la Iglesia R, Moreno-Pino M, Lopes dos Santos A, Gerikas Ribeiro C, Cristi A, Marie D, Vaulot D. 2021. Annual phytoplankton dynamics in coastal waters from Fildes Bay, Western Antarctic Peninsula. *Sci. Rep.*:2020.10.27.356600.

Tremblay J-E, Gagnon J. 2009. The effects of irradiance and nutrient supply on the productivity of Arctic waters: a perspective on climate change. In: Influence of Climate Change on the Changing Arctic and Sub-Arctic Conditions. p. 73–89.

Tremblay JÉ, Simpson K, Martin J, Miller L, Gratton Y, Barber D, Price NM. 2008. Vertical stability and the annual dynamics of nutrients and chlorophyll fluorescence in the coastal, southeast Beaufort Sea. *J. Geophys. Res. Ocean.* 113:1–14.

Turner J, Colwell SR, Marshall GJ, Lachlan-Cope TA, Carleton AM, Jones PD, Lagun V, Reid PA, Iagovkina S. 2005. Antarctic climate change during the last 50 years. *Int. J. Climatol.* 25:279–294.

Vader A, Marquardt M, Meshram AR, Gabrielsen TM. 2014. Key Arctic phototrophs are widespread in the polar night. *Polar Biol.* 38:13–21.

Vergin KL, Beszteri B, Monier A, Thrash JC, Temperton B, Treusch AH, Kilpert F, Worden AZ, Giovannoni SJ. 2013. High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J.* 7:1322–1332.

Wang S, Durrant WE, Song J, Spivey NW, Dong X. 2010. Arabidopsis BRCA2 and RAD51 proteins are specifically involved in defense gene transcription during plant immune responses. *Proc. Natl. Acad. Sci. U. S. A.* 107:22716–22721.

Wang Y, Spalding MH. 2014. LCIB in the Chlamydomonas CO2-concentrating mechanism. *Photosynth. Res.* 121:185–192.

Wang Y, Stessman DJ, Spalding MH. 2015. The CO2 concentrating mechanism and photosynthetic carbon assimilation in limiting CO2: How Chlamydomonas works against the gradient. *Plant J.* 82:429–448.

Ware C, Friedlaender AS, Nowacek DP. 2011. Shallow and deep lunge feeding of humpback whales in fjords of the West Antarctic Peninsula. *Mar. Mammal Sci.* 27:587–605.

White E, Hoppe CJM, Rost B. 2020. The Arctic picoeukaryote Micromonas pusilla benefits from ocean acidification under constant and dynamic light. *Biogeosciences* 17:635–647.

Wideman JG, Monier A, Rodríguez-Martínez R, Leonard G, Cook E, Poirier C, Maguire F, Milner DS, Irwin NAT, Moore K, et al. 2020. Unexpected mitochondrial genome diversity revealed by targeted single-cell genomics of heterotrophic flagellated protists. *Nat. Microbiol.* 5:154–165.

Wilson ST, Tozzi S, Foster RA, Ilikchyan I, Kolber ZS, Zehr JP, Karl DM. 2010. Hydrogen cycling by the unicellular marine diazotroph crocosphaera watsonii strain WH8501. *Appl. Environ. Microbiol.* 76:6797–6803.

Wlodarska-Kowalczuk M, Pearson TH, Kendall MA. 2005. Benthic response to chronic natural physical disturbance by glacial sedimentation in an Arctic fjord. 303:31–41.

Wolf C, Frickenhaus S, Kilias ES, Peeken I, Metfies K. 2013. Regional variability in eukaryotic protist communities in the Amundsen Sea. *Antarct. Sci.* 25:741–751.

Wolf C, Kilias E, Metfies K. 2015. Protists in the polar regions: Comparing occurrence in the Arctic and Southern oceans using pyrosequencing. *Polar Res.* 34:1–8.

Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ. 2015. Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science (80-.).* 347.

Worden AZ, Lee JH, Mock T, Rouzé P, Simmons MP, Aerts AL, Allen AE, Cuvelier

ML, Derelle E, Everett M V., et al. 2009. Green Evolution and Dynamic Adaptions Revealed by Genomes of the Marine Picoeukaryotes Micromonas. *Science (80-.).* 324.

Worden AZ, Nolan JK, Palenik B. 2004. Assessing the dynamic and ecology of marine picoplankton : the importance of eukaryotic component. *Limnol. Ocean.* 49:168-79.

Wright SW, Ishikawa A, Marchant HJ, Davidson AT, Van Den Enden RL, Nash G V. 2009. Composition and significance of picophytoplankton in Antarctic waters. *Polar Biol.* 32:797–808.

Xiang Y, Zhang J, Weeks DP. 2001. The Cia5 gene controls formation of the carbon concentrating mechanism in Chlamydomonas reinhardtii. *Proc. Natl. Acad. Sci. U. S. A.* 98:5341–5346.

Young EB, Beardall J. 2005. Modulation of photosynthesis and inorganic carbon acquisition in a marine microalga by nitrogen, iron, and light availability. *Can. J. Bot.* 83:917–928.

Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. 2017. Ggtree: an R Package for Visualization and Annotation of Phylogenetic Trees With Their Covariates and Other Associated Data. *Methods Ecol. Evol.* 8:28–36.

Zhang F, Cao S, Gao Y, He J. 2019. Distribution and environmental correlations of picoeukaryotes in an arctic fjord (Kongsfjorden, Svalbard) during the summer. *Polar Res.* 38:1–11.

Zhang Z, Qu C, Zhang K, He Y, Zhao X, Yang L, Zheng Z, Ma X, Wang X, Wang W, et al. 2020. Adaptation to Extreme Antarctic Environments Revealed by the Genome of a Sea Ice Green Alga. *Curr. Biol.* 30:3330-3341.e7.

Zhou M, Niiler PP, Hu JH. 2002. Surface currents in the Bransfield and Gerlache Straits, Antarctica. *Deep. Res. Part I Oceanogr. Res. Pap.* 49:267–280.

Zoccarato L, Pallavicini A, Cerino F, Fonda Umani S, Celussi M. 2016. Water mass dynamics shape Ross Sea protist communities in mesopelagic and bathypelagic layers. *Prog. Oceanogr.* 149:16–26.

Zou Y, Wenzel S, Müller N, Prager K, Jung EM, Kothe E, Kottke T, Mittag M. 2017. An animal-like cryptochrome1 controls the Chlamydomonas sexual cycle. *Plant Physiol.* 174:1334–1347.