PATTERNS OF RUBISCO CONCENTRATIONS IN GLACIERIZED AND NON-GLACIERIZED ENVIRONMENTS ACROSS JONES SOUND, NU

by

Megan Roberts

Submitted in partial fulfilment of the requirements
for the degree of Masters of Science

at

Dalhousie University
Halifax, Nova Scotia
September 2022

Table of Contents

List of Tables

List of Figures

Abstract

Tidewater glaciers can enhance delivery of nutrients to the surface ocean. The Canadian Arctic Archipelago (CAA) has a high density of tidewater glaciers, but their influence on the marine environment, particularly phytoplankton, is understudied. Here we develop and apply a targeted metaproteomic approach to measure Rubisco concentrations in phytoplankton communities across Jones Sound, NU. Because Rubisco is the protein used for carbon fixation, we can estimate potential rates of primary production from these measurements. We found these estimates to be within expected ranges. We found that Rubisco produced by *Chaetoceros* (diatom) is higher at glacierized stations while Rubisco from *Micromonas* (Chlorophyta), is enhanced at non-glacierized sites. This suggests that future climate scenarios may favour smaller phytoplankton groups, like *Micromonas*, with downstream consequences for food webs and carbon cycling. This study broadens our understanding of the impact tidewater glaciers in the CAA, now and in a warmer future.

List of Abbreviations used

ACN: acetonitrile

ASV: amplicon sequence variant

BCA: bicinchoninic acid

CAA: Canadian Arctic Archipelago

CCGS: Canadian Coast Guard Ship

Chl a: Chlorophyll a

CTD: conductivity, temperature, depth

DNA: deoxyribonucleic acid

DTT: dithiothreitol

FA: formic acid

HDPE: High Density Polyethylene

HPLC: High-performance liquid chromatography

MS: mass spectrometer

NCBI: National Centre for Biotechnology Information

PAR: Photosynthetically active radiation

PCA: Principal component analysis

PCR: Polymerase chain reaction

POC: particulate organic carbon

PP: primary productivity

QIIME: Quantitative Insights into Microbial Ecology

RbcL: Rubisco large subunit

RNA: Ribonucleic acid

S/Y: Sailing yacht

SCM: subsurface chlorophyll maximum

SDS: sodium dodecyl sulfate

SRM: selected reaction monitoring

TEAB: Triethylammonium bicarbonate

Acknowledgements

This thesis would not have been possible without immense support and guidance from my two supervisors Dr. Erin Bertrand and Dr. Maya Bhatia. I am truly grateful for being part of such a collaborative and diverse project led by such strong women. Erin, I have eternal gratitude for your thoughtful and patient insight and leadership in science and otherwise along the way. Maya, your unwavering dedication and drive is admirable and I am thankful for every necessary pep talk. I am so appreciative to every member of the Bertrand lab who have all made this degree more enjoyable in so many ways, from input on data to excursions outside the lab – Elden, Scott McCain, Loay, Insa, Maria, Cat, Catalina, Lena and Scott Pollara. To everyone of the 'Ice to Ocean' team – Stephanie Waterman, Dave Burgess, Andrew Hamilton, Patrick Williams, Patrick White and Jenifer Spence and my committee – Dr. Zoe Finkel and Dr. Emmanuel Devred for input and discussion along the way. I am also thankful to Amundsen Science, the crew of the *CCGS Amundsen*, and chief scientist Jean Carlos for support and logistics in the field. I consider myself lucky to have met and worked with Patrick Williams during fieldwork aboard the *Amundsen*. I couldn't have imagined a better sampling partner to filter water with at any and all hours. I need to thank the many people who have contributed to this dataset including: Jean-Eric Tremblay for nutrient data collected on the *CCGS Amundsen*, Maria Cavaco and Patrick White from the Bhatia lab for processing and analyzing DNA samples and Gianpaolo Cardellini for assistance with the RbcL database. I owe a particular thanks to Dr. Elden Rowland for his training and knowledge of protein and mass spectrometry work in the lab. I can't thank every close friend and family member enough for never-ending support through this degree, and in particular Cait for being an unwavering source of belief and encouragement.

Chapter 1: Introduction

1.1 Increased ice mass loss due to global warming

Ice masses around the world are melting at unprecedented rates due to anthropogenic climate change (Bliss et al., 2014; King et al., 2020). This warming is amplified in the Arctic region (Rantanen et al., 2022) and melt rates of ice masses have increased markedly over the past two decades (Gardner et al., 2011; Noël et al., 2018). Recent glacier mass balance modelling efforts (inclusive of all ice caps and glaciers outside of Antarctica) predicts these trends will continue to gradually increase over the twenty-first century (Bliss et al., 2014). The Canadian Arctic Archipelago (CAA), where approximately 14% of the Earth's glaciers and ice caps are located, is no exception. In contrast to Antarctica or Greenland, the CAA is comprised of smaller but numerous (>300) tidewater glaciers, and therefore ice masses in this region are more susceptible to summer air temperature extremes and are predicted to continue to face amplified glacier mass loss (Box et al., 2019; Cook et al., 2019). In the future, warming air temperatures are likely to continue to result in accelerated melting of CAA tidewater glaciers with further retreat (increased runoff over the next century) and eventually disappearance (subsequent centuries) (Cook et al., 2019).

1.2 Glacier melt – ocean interactions

In high latitude regions, increased runoff from tidewater glaciers (ice masses where the terminus of the glacier is in contact with the ocean) and the interaction of this runoff with the coastal marine environment has the potential to significantly impact marine primary production (Hopwood et al., 2018). The delivery of dissolved freshwater-sourced and

marine-sourced chemical species via the buoyant glacial meltwater plume to the surface ocean has been found to enhance primary productivity in some regions (Arrigo et al., 2017; Halbach et al., 2019; Kanna et al., 2018; Meire, Mortensen, et al., 2016). This is in part due to the timing of the melt, which largely comes at the height of summer when surface nutrient concentrations are low due to drawdown during the initial spring bloom (Tremblay et al., 2015), but when light availability is still high. Simultaneously, the sediment load delivered with glacial melt can also have negative impacts on primary productivity. This is because turbid plumes surrounding glacier termini can increase light attenuation into the water column, causing phytoplankton growth to be light limited in these areas (Halbach et al., 2019; Murray et al., 2015). Runoff from tidewater glaciers enters the marine water column at a depth determined by several factors, including the thickness of the ice at the terminus, the grounding line depth (maximum depth of the ice front at the ocean) and the location of the submarine discharge channels (Slater et al., 2022). Once in the ocean, the resulting buoyant freshwater has the potential to entrain and deliver deeper seawater to the surface. This is notably in contrast to land-terminating glaciers where freshwater runoff flows onto the surface of the ocean, stratifying the water column and restricting mixing (Hopwood et al., 2018). Recent studies conducted around Greenland, where the average grounding line depth of tidewater glaciers is relatively deep, at 280 m, have identified two general mechanisms by which glaciers can facilitate the delivery of nutrients to the surface ocean (Morlighem et al., 2017). These are entrainment-driven upwelling of macronutrients (nitrate, silicate, phosphate) at the base of tidewater glaciers and direct delivery of macro- and micro-nutrients such as silicate, phosphorous and trace metals via the glacial melt itself (Bhatia et al., 2013; J. Hawkings

et al., 2016; J. R. Hawkings et al., 2017; Meire, Meire, et al., 2016). Recently in the northern CAA, where tidewater glaciers have a much shallower average grounding line depth of 150 m (Van Wychen et al., 2014), these same mechanisms of nutrient addition have also been observed (Bhatia et al., 2021; Williams et al., 2021). Both mechanisms deliver vital nutrients to phytoplankton growth, however in the CAA, nitrate is often a limiting nutrient during the summer (Randelhoff et al., 2020; Tremblay & Gagnon, 2009). Since nitrate concentrations in glacial meltwater from this region are low (< 2 μM, Williams et al. 2021), upwelling of nutrient rich deep seawater near tidewater glaciers appears to be the mechanism for delivering this key nutrient (Bhatia et al., 2021; Williams et al., 2021). Thus these tidewater glacier influenced regions are of particular interest for this study. Broadening our understanding of phytoplankton communities in close proximity to tidewater glaciers and how they respond to glacial runoff is important because over the next century runoff from marine-terminating glaciers will likely continue to increase (Cook et al., 2019) with important possible consequences for future marine primary productivity.

Despite the high density of tidewater glaciers in the CAA, until recently, there have been no comprehensive studies investigating interactions between phytoplankton and glaciers in this region. Jones Sound is a marine waterway located south of Ellesmere Island and north of Devon Island in Nunavut, and is home to the Inuit hamlet of Ausuittuq (Grise Fiord), the northernmost community in Canada (Fig. 1). The residents of this community rely on this marine region for subsistence hunting and have observed increased productivity at the termini of the glaciers (pers. comm. J. Qaapik, Grise Fiord Rangers). Some initial, pioneering work in this region in the 1960s found higher spring-

time (under the sea ice) macronutrient concentrations in a fiord with a tidewater glacier relative to one without, concluding that these elevated concentrations are due to activity of the tidewater glaciers and that this may positively impact phytoplankton (Apollonio, 1973). Nearly 60 years later, a new study, albeit one conducted in the summer (open water) season, drew similar conclusions based on a survey of a broader range of tidewater glaciers in the region (Bhatia et al., 2021). The surface waters surrounding tidewater glaciers had significantly higher concentrations of macronutrients and micronutrients compared to surface waters adjacent to coastal areas without tidewater glaciers. However, chlorophyll *a* (Chl *a*) data collected at the same time has not yet clarified whether these glacier-driven nutrient additions promote increased phytoplankton growth (Bhatia et al., 2021). To further investigate the impact of glacial melt on the phytoplankton community in the CAA, we aim to quantify the response of these marine primary producers in regions influenced by tidewater glaciers, and those regions lacking this influence, using a novel protein-based approach that gives additional insight beyond that of Chl *a*.

1.3 Phytoplankton in the CAA

Previous studies of phytoplankton in the CAA have identified patterns and trends in primary productivity in this region as well the dominant groups facilitating this process. The North Water polynya, located just northeast of Jones Sound between Ellesmere Island and Greenland at the northern edge of Baffin Bay, is the closest region within the CAA in proximity to Jones Sound with extensive past studies of phytoplankton dynamics. Primary productivity (PP) and Chl *a* measurements have been used to observe both long term trends and seasonal shifts in phytoplankton biomass in this region (Blais et al., 2017; Tremblay et al., 2012; Varela et al., 2013). Studies examining the

phytoplankton community assemblages across the CAA have helped identify when and where prevalent phytoplankton groups are productive in the region. For example, results obtained using 18S rRNA gene and transcript sequencing, reveal that *Micromonas spp.* appear to be ubiquitously present and important across most of the CAA under diverse environmental conditions (Joli et al., 2018; Kalenitchenko et al., 2019; Monier et al., 2015). In comparison, results obtained via microscopy and 18S sequencing, show that diatoms such as *Chaetoceros, Thalassiosira* and *Fragilariopsis* sp. tend to dominate during bloom periods (Booth et al., 2002; Kalenitchenko et al., 2019), with other groups such as dinoflagellates and haptophytes, specifically *Phaeocystis*, also commonly present during or after bloom periods (Kalenitchenko et al., 2019). These studies examining phytoplankton community composition are often paired with primary production (^{14}C, ^{13}C-based), Chl a, and particulate organic carbon (POC) measurements to assess the contribution of different groups to primary production. To date, no studies have systematically explored the influence of glacial melt on marine phytoplankton communities in the CAA. In this study we measure Rubisco concentrations produced by major phytoplankton groups in the CAA. Rubisco is the enzyme that catalyzes fixation of atmospheric CO_2 and is therefore directly involved in primary productivity. By measuring Rubisco concentrations associated with different phytoplankton groups, we aim to fill a knowledge gap on how primary producers are responding to glacial input in the marine environment using a new method.

1.4 Rubisco

In this study, we apply a novel mass spectrometry-based technique to measure the concentration of the protein Rubisco in phytoplankton. Rubisco enzymes found in

eukaryotic phytoplankton are composed of eight large subunits and eight small subunits (Hartman & Harpel, 1994). The large subunits of Rubisco are directly involved in the catalytic chemistry of the carbon fixation reaction (Andrews 1988) while the small subunits seem to help with specificity of the reaction (Spreitzer, 2003) but are not essential for catalysis (Andersson & Backlund, 2008). For this reason, in this study we measure the concentration of Rubisco large subunits (RbcL). Rubisco is an inefficient enzyme, largely due to the fact that when oxygen is present it will compete with CO_2 at the active site and initiate a photorespiratory pathway (Lorimer & Andrews, 1973). To account for this inefficiency, Rubisco is present in relatively large concentrations in marine phytoplankton (Losh et al., 2013; Young, Goldman, et al., 2015), thereby making it an accessible target to quantify in mixed microbial communities using mass spectrometry. Surveys of Rubisco-encoding genes and their expression have been made in the ocean: qPCR assays have been used to measure RbcL gene abundance (John et al., 2007) and gene expression (Pujari et al., 2019) from different phytoplankton taxa and quantitative western blots have been used to obtain measurements of Rubisco as a percentage of the total protein in communities (Losh et al., 2013; Young, Goldman, et al., 2015; Young, Kranz, et al., 2015). Measurements of Rubisco enzyme activity and Chl a patterns have been made in Jones Sound, revealing that Rubisco activity was highly correlated with Chl a concentration and distributed in similar patterns throughout the mixed layer of the water column (Harrison et al., 1987). However, these past analyses of Rubisco activity in marine phytoplankton have not provided absolute quantification of the protein, and have not been able to resolve which phytoplankton groups are responsible for Rubisco production.

Here we develop a targeted metaproteomic method that measures the concentration of RbcL produced by different phytoplankton groups thought to be important in the CAA. This gives us quantitative insight into what major phytoplankton groups are producing Rubisco and allows us to compare patterns across different environments throughout Jones Sound. We then us these measurements to estimate maximum potential carbon fixation rates in these environments and apportion that carbon fixation to key phytoplankton groups quantitatively. It is important to understand how different groups of phytoplankton are contributing to carbon fixation and responding to glaciers because different phytoplankton groups have distinct impacts on marine biogeochemical cycles and arctic food webs (Miller and Wheeler, 2012). Gaining understanding of how tidewater glaciers are influencing phytoplankton in the marine environment presently is vital to helping us predict what might happen to marine productivity in coastal regions influenced by glaciers in the future as these ice masses continue to melt, retreat from the ocean, and eventually disappear.

Chapter 2: Methods

2.1 Study Site

Jones Sound is a waterway north of Devon and south of Ellesmere Islands in the Canadian Arctic Archipelago (CAA). Water flows into Jones Sound from Nares Strait to the east and Hell Gate to the west and currents generally flow west along the north side of the sound and east along the south, largely exiting into Baffin Bay (Figure 1, inset). Jones Sound is also home to the northern most Inuit hamlet of Ausuittuq (Grise Fiord), whose residents use the area as hunting grounds for 'country food' (seal, narwhal, beluga). There are numerous tidewater glaciers that drain into Jones Sound from ice fields and caps on Ellesmere and Devon Islands; in this study we conducted near-terminus sampling of the ocean at four of these glaciers. Glaciers targeted here include Sydkap and Jakeman glaciers on Ellesmere Island draining the Sydkap and Manson ice fields respectively, as well as Sverdrup and Belcher glaciers on Devon Island draining the Devon ice cap (Figure 1). During summer months (July-August) surface nutrient concentrations have been found to be elevated in the marine region proximal to these glaciers as a result of upwelling of deep seawater (Bhatia et al 2021). These four glaciers also fall along a continuum of how strong their influence is on the marine environment based on the proportion of the ice terminus in the water, grounding line depth and subglacial meltwater flux (Bhatia et al 2021). Based on these parameters, these four glaciers order from Belcher glacier as a strongly-influenced-tidewater-glacier-system to Jakeman glacier as a weakly-influenced-tidewater-glacier-system with Sydkap and Sverdrup glaciers falling in between (Bhatia et al. 2021). Two near-shore non-glacierized sites were also sampled including Grise Fiord on Ellesmere Island and a site near Truelove Inlet on Devon Island

(Figure 1). In addition to these nearshore sites, offshore sampling occurred across open Jones Sound, the eastern gateway and just outside on the western edges of Baffin Bay.

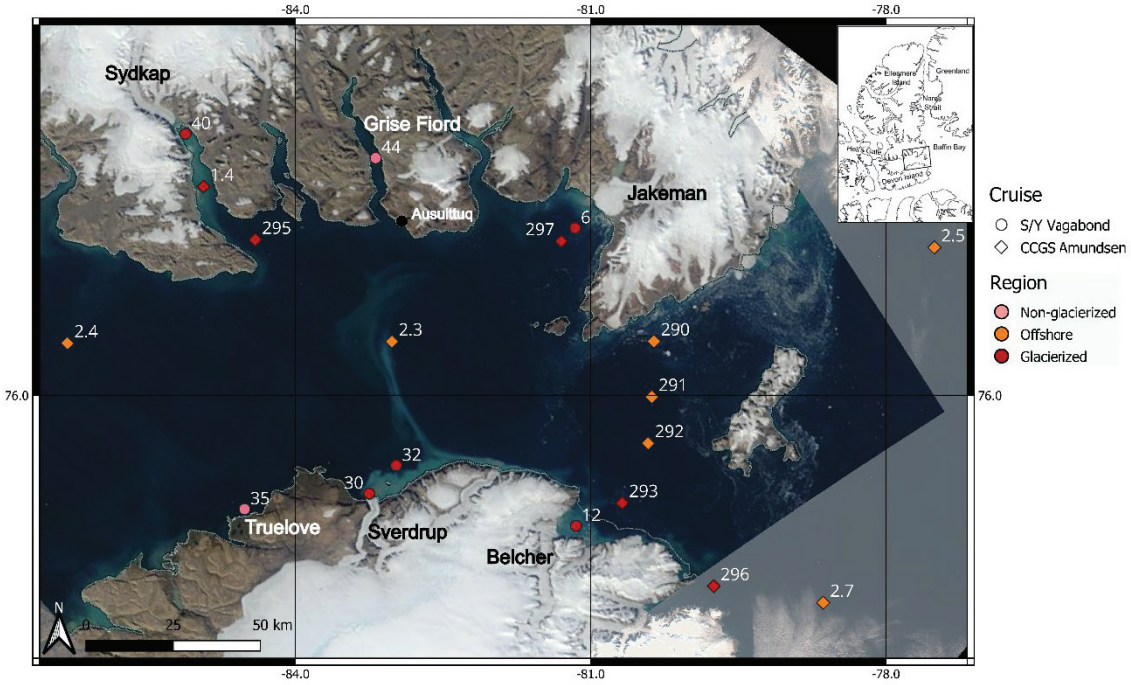


Figure 1. Map of study area in Jones Sound, NU in the Canadian Arctic Archipelago. Six distinct near-coast study sites were sampled. Two non-glacierized sites: Grise Fiord and Truelove and four glacierized sites: Jakeman, Belcher, Sverdrup and Sydkap. Diamonds indicated locations of stations sampled during Leg 2b of the *CCGS Amundsen* 2019 summer cruise, and circles indicate the stations sampled from the *S/Y Vagabond,* also in summer 2019. Stations coloured red are considered glacierized sites, pink sites are nearshore non-glacierized sites while orange stations are >20km offshore. *CCGS* Amundsen stations were sampled for Rubisco protein at 4 depths, three in the euphotic zone and one below, while *S/Y Vagabond* stations were sampled only at the subsurface chlorophyll a maximum depth. At all stations, CTD, Chlorophyll a, dissolved O_2, PAR and nutrient samples were also collected. Samples for chloroplast 16S rRNA gene sequencing were collected at Vagabond stations only, also at the chlorophyll maximum depth.

2.2 Sample Collection

Samples from the Jones Sound region were collected in late summer of 2019 during two concurrent expeditions, August 6-10th on the *CCGS Amundsen* and July 22-August 16 on the *S/Y Vagabond*. Coordinates and sampling details for all stations are outlined in Table S1. Drs Maya Bhatia, Erin Bertrand and Dave Burgess conducted the sampling aboard

S/Y Vagabond, while Patrick Williams and I conducted sampling aboard *CCGS Amundsen.*

On the *CCGS Amundsen*, in situ measurements of the water column profiles were collected using a conductivity, temperature and depth profiler (CTD; Sea-Bird SBE-011, Sea-Bird Electronics Inc.) with additional sensors for fluorescence (Seapoint chlorophyll fluorometer), light intensity (PAR; LI-COR Biosciences) and dissolved oxygen (Sea-bird electronic / SBE 43). Seawater to filter for protein samples was collected in 12L Niskin bottles attached to a rosette system. Water was transported from the Niskin bottle to the filtration apparatus using clean 10L HDPE carboys. Water was filtered sequentially through 3 and 0.2 µm pore-size polycarbonate filter membranes using a peristaltic pump, tygon and silicon tubing. Filters were stored in 2 mL cryovials and frozen in a -80°C freezer. We collected water from 4 depths at each station. These depths were generally surface water at 5m, the subsurface chlorophyll maximum (SCM), a second deep chlorophyll max (if present), and a deep sample at least 100m below any detectable fluorescence signal from the sensors. Nutrients were also sampled directly from the niskin bottles through a glass fiber filter (GF/F, Whatman) into acid cleaned polyethylene tubes and stored in the dark at 4°C, until measured on-board the *CCGS Amundsen* within a few hours of collection. Nutrient concentrations were determined by the Tremblay Lab as in Kalentichenko et al. 2019.

The *S/Y Vagabond* collected profiles of the water column using a RBR maestro^3 multichannel logger (RBR Ltd., Ottawa) equipped with sensors for conductivity, temperature, pressure, dissolved oxygen, photosynthetic active radiation, Chl *a*, and turbidity. Seawater for nutrient, protein and DNA samples was collected using 10L Go-

Flo bottles at the SCM depth. For DNA and protein, water was taken from the Go-Flo into 4 L polycarbonate bottles. It was then filtered through a 0.22 µm Sterivex filter unit via peristaltic pumping. 1L was filtered for DNA samples and between 2-5 L were filtered for protein samples. Residual water was pushed through the Sterivex with a clean syringe. The filter was then capped and flash frozen and stored in a dry shipper charged with liquid nitrogen for transport back to the laboratory, at which point samples were transferred to a -80C freezer. Nutrients samples were filtered through a 0.22 µm polyethersulfone (PES) syringe filter into clean (3.7% HCl soaked, MilliQ water rinsed), 3x sample-rinsed 20 mL HDPE plastic scintillation vial and frozen (-20C) until analysis. Nutrient concentrations were analyzed within 1 month of collection on a Skalar SAN++ Continuous Flow Nutrient Analyzer at the CERC.OCEAN Laboratory (McGrath et al., 2019).

2.3 RbcL sequence database

To select appropriate RbcL peptide sequences for this study, first a database of RbcL protein sequences was compiled. Protein sequences (FASTA) similar to Type 1 Ribulose-1,5-bisphosphate carboxylase/oxygenase RbcL from *E. huxleyi* (Tabita et al. 2007/2008) were retrieved using the "protein BLAST" software from the National Center for Biotechnology Information (NCBI) website (https://blast.ncbi.nlm.nih.gov/Blast.cgi) using the default settings, with one exception. This was that searches were restricted to photosynthetic microalgae Rubisco by excluding several organism taxonomic ids: liverworts (taxid: 3195), mosses (taxid: 3208), ferns (taxid: 241806), gymnosperms (taxid: 71049), flowering plants (taxid: 3398), seed plants (taxid: 58024), hornworts (taxid: 13809), vascular plants (taxid: 58023), bryophyta (taxid: 3208), angiospermae

(taxid: 3398), bacteria (taxid: 2), bacterium species (taxid: 77133), green plants (taxid: 33090), red algae (taxid: 2763), euryarchaeota (taxid: 28890), brown algae (taxid: 2870), crenarchaeotes (taxid: 28889), and viruses (taxid: 10239). All other search algorithm parameters were left as the default. The following databases were used to source sequences: refseq_protein, swissprot, pdb, and MMETSP collection of transcriptomes (Keeling et al 2014). A multiple alignment with the collected sequences was conducted in Geneious Prime (version 2.4). Sequences less than 76% identical at the amino acid level to any other sequence in the alignment similar were deleted. This cut-off value was chosen as it has been previously used in molecular phylogenetic studied for Rubisco allowing for the identification of *bona fide* Rubisco sequences (Pujari et al., 2019; Tabita et al., 2008; Young et al., 2016). Dr. Elden Rowland and Gianpaolo Cardellini assisted with Rubisco database construction.

2.4 Peptide Selection

To be able to measure the Rubisco concentration of important phytoplankton groups, peptides that were conserved across relevant taxonomic groups for our study area were chosen *a priori*. At the time of peptide selection, DNA and RNA sequencing data from the sites sampled for protein data was not yet available. Therefore, we based Rubisco peptide decisions on the available literature of phytoplankton community composition across the CAA at the time (Joli et al., 2018; Kalenitchenko et al., 2019; Lovejoy, C, 2014). I leveraged our Rubisco database to attempt to select sequences from the following phytoplankton groups pertinent to the CAA: diatoms, haptophytes, green algae (specifically *Micromonas sp.*) and dinoflagellates. MUSCLE alignments of sequences from each of these groups were prepared and inspected for tryptic peptides conserved

within each group. The consensus sequence from each alignment was put into the tool PeptideCutter (Gasteiger et al. in Walker 2005) and any potentially useful peptides were checked for taxonomic specificity using the program Unipept (Gurdeep Singh et al., 2019). I was able to identify tryptic peptides to measure all groups except for dinoflagellates. These peptides are summarized in Figure 2.

2.5 Protein Extraction

Protein was extracted from 0.2 and 3 µm polycarbonate filters and 0.22 µm Sterivex filters. To remove the filter from the Sterivex units, the plastic cartridge was pried opened with plyers while kept frozen on a bed of dry ice and a sterile blade was used to cut around perimeter of the filter area and the filter placed into a 2 mL cryovial. While on ice, 750 µL of 2% SDS extraction buffer (0.1 M Tris/HCl pH 7.5, 5% glycerol, 5 mM EDTA, 2% SDS) was added to each sample. After 10 minutes the samples were heated at 95 °C at 350 RPM for 15 minutes. Next, the filters were sonicated on ice for 1 minute each using a Qsonica Sonicator (Newtown, CT). Settings for the sonicator were 50% amplitude, 125 W; pulse 15s ON, 15s OFF. After sonication the filters were incubated for 30 min. at room temperature with a gentle vortex every 10-15 min. At this point the supernatant was separated from the filter and transferred into a clean 2 mL centrifuge tube. The sample was centrifuged at 15 000 g for 30 min. This was sometimes increased to 45 min. if the cell pellet was still dispersed or not well formed which happened most often for the 3 µm filters. The supernatant was transferred to a new 2 mL centrifuge tube (safe-lock, Eppendorf). The tube was weighed before and after the transfer of the supernatant to determine the volume of extract obtained, assuming a density of 1 mg/mL.

Protein concentration was determined using a Micro BCA Protein Assay Kit (Thermo Scientific™).

2.6 Protein Digestion

Between 20 – 50 µg of protein was digested with trypsin using S-trap mini columns (Protifi, USA). Protein extracts were reduced by adding 5 mM dithiothreitol (DTT, Sigma-Aldrich 97%) dissolved in 50 mM ammonium bicarbonate (Ambic) and incubated at 37 °C for 1 hour, 350 RPM. After cooling down to room temperature, the extract was alkylated by adding 15 mM iodoacetamide (Sigma-Aldrich, BioUltra) dissolved in 50 mM Ambic. Sample was vortexed briefly, spun down, and left to incubate for 30 minutes at room temperature in the dark. 5mM DTT was added again at room temperature and vortexed. Next, 12% phosphoric acid was added to obtain 1.2% concentration in the sample and S-Trap buffer (90% aqueous methanol in 100 mM TEAB, pH7.1; TEAB acidified to 7.1 using 85% phosphoric acid) was added in volumes 7 times that of the volume of protein extract being digested. For samples in which the addition of S-trap buffer would result in a total volume greater than 2 mL, samples were transferred to a 5 mL centrifuge tube (Protein LoBind, Eppendorf) with a 'rinse' of the 2 mL tube with S-trap buffer to minimize sample loss. After addition of the S-trap buffer, sample was loaded onto the S-traps in 600 µL aliquots. A vacuum extraction manifold (Waters) was used for this step and the subsequent washes of the S-trap. The S-traps were then washed with ten 600 µL volumes of S-trap buffer. The first wash aliquot was used to rinse the sample tube to reduce sample loss. S-trap columns were removed, spun at 2000 x g and moved to a clean 2 mL centrifuge tube. Protein captured on the S-trap columns was digested with trypsin (Thermo Scientific™ Pierce™ Trypsin Protease, MS Grade) for 12-

16 hours at 37 °C (1:25 ratio trypsin to total protein). The trypsin was loaded onto columns in a volume of 125 µl 50 mM TEAB at pH 8. Digested samples were then eluted off the S-trap column through a series of washes: 80 µL of 50 mM Ambic, 80 µL 0.2% aqueous formic acid and 80 µL 50% acetonitrile containing 0.2% formic acid. Columns were spun down at 4 000 g for 1 minute after each addition. The S-trap column was removed and peptide solutions were transferred to a 1.7 mL centrifuge tube (protein lobind, Costar) and dried down in a vacufuge (Eppendorf) for 2-4 hours (V-AQ setting). The dry peptides were then stored at -80 until reconstitution. Dried peptides were reconstituted in a volume of 1% FA, 3% ACN to a concentration of 0.5 µg/µL. The peptide concentration was confirmed using a Micro BCA Protein Assay Kit (Thermo Scientific^{TM}) with a tryptic digested BSA as the protein standard. I performed all lab work associated with protein extraction, digestion and preparation for mass spectrometer analysis with guidance from Dr. Rowland.

2.7 Targeted Mass Spectrometry

Targeted metaproteomic analysis was performed using a Dionex Ult 3000 UPLC coupled with a TSQ Quantiva mass spectrometer (MS). The MS was equipped with a heated low flow capillary ESI probe (HESI-II) with the following settings: spray voltage of 3500 V, sheathgas 5, auxillary gas 2, ion transfer tube temperature of 325 °C, vapor gas 70 °C and Chrom filter setting of 10 s. Samples were diluted with 1% FA, 3% ACN to a final peptide concentration of 0.167 µg/µL. Each sample was spiked with either 3.3 or 6.6 fmol/µl of each heavy isotope-labelled internal standard for each peptide (Thermo Scientific, Supplemental Table 1). 6 µl injections were performed in triplicate. Selected reaction monitoring (SRM) transitions were optimized on our instrument using the

Quantiva transition optimization tool. The method contained 86 transitions (Table S2), 15 ms dwell time, Q1 and Q3 resolution was set to 0.7 (FWHM), automatically calibrated RF lens setting, and a collision gas pressure of 2.5 mTorr. All raw targeted metaproteomic data obtained from the mass spectrometer was processed using Skyline-daily software (Pino 2017). RbcL peptide concentrations were calculated from the mass spectrometry data by multiplying the peak area of each peptide of interest by the ratio of moles of the heavy isotope-labeled version of that peptide added to the peak area corresponding to that heavy isotope labeled peptide. For peptides containing the amino acid methionine, which can be oxidized, peak areas of oxidized and non-oxidized forms of the peptide were summed together. Concentrations of each peptide from 3 and 0.2 µm size fraction filters, if available, were summed. Dr. Rowland assisted with liquid chromatography and mass spectrometry. I analysed the targeted mass spectrometry data with guidance from Dr. Rowland.

2.8 DNA Extraction

DNA was extracted from the 0.2 µm sterivex filters using a phenol chloroform protocol. Sterivex filters were thawed on ice. 500 mL of TE buffer, 20 µL RNAseA (10mg/mL) and 150 µL of fresh lysozyme were added to each sterivex. Sterivex then incubated for 1 hour at 37°C while rotating. 100 µL of proteinase K(10mg/mL) and 100 µL 10% SDS were added to the sterivex and incubated for another 20 minutes at 55°C. Lysate from moved from sterivex cartridges using a syringe into a clean 15 mL centrifuge tube. Equal volume of Phenol:chloroform:IAA (25:24:1) was added to each sample. Tubes were inverted for 10s and spun at 2500 g for 5 minutes, or until an aqueous layer with no debris on top of the milky layer formed. This upper aqueous layer was then transferred to a clean 15 mL centrifuge tube. An equal volume of Chloroform:IAA (24:1) was added to

the new tubes. Tubes were again inverted for 10s and spun at 2500 g for 5 minutes, or until an aqueous layer with no debris on top of the milky layer formed. Aqueous layer was transferred to new Amicon filter with 650 µL of TE buffer. Amicon filter was spun at 3500g for 10 minutes. Wash filter two more times with ~1.5 mL of TE buffer at 3500 g for 6 minutes. Remaining volume transferred from Amicon filter (~250 µL) into 1.5 mL epi tube and stored at -80°C. DNA was quantified using a Qubit. Maria Cavaco from the Bhatia Lab at the University of Alberta conducted these DNA extractions.

2.9 DNA sequencing and computational analyses

16S rRNA were amplified genes using universal prokaryotic primers 515F (Parada et al., 2016) and 926R (Quince et al., 2011). Each primer also contained a six-base index sequence for sample multiplexing (Bartram et al., 2011). The PCR mix (25µL total volume) contained 1 × Q5 reaction buffer, 0.5 µM forward primer, 0.5 µM reverse primer, 200 µM dNTPs, 0.500 U Q5 polymerase (New England Biolabs, Ipswich,M.A, U.S.A) and 2.5 µL of genomic template. Genomic extracts with DNA concentrations of greater than 5 ng μL^{-1} were diluted 1:100 in nuclease-free water. The PCR was performed as follows: 95°C for 3 minutes, 35 cycles of 95°C for 30 seconds, 60°C for 30 seconds, 70°C for 1 minute and a final extension of 70°C for 10 minutes. Pooled 16S rRNA gene amplicons were purified using Nucleomag beads and a 4.5 pM library containing 50% PhiX Control v3 (Illumina, Canada Inc., NB, Canada) was sequenced on a MiSeq instrument (Illumina Inc., CA, USA) using a 2 × 250 cycle MiSeq Reagent Kit v3 (Illumina Canada Inc). The MiSeq reads were demultiplexed using MiSeq Reporter software version 2.5.0.5. Each read pair was assembled using the paired-end assembler for Illumina sequences (PANDAseq; Masella, Bartram and Truszkowski, 2012) with a

quality threshold of 0.9, dictating that 90% of overlapping reverse and forward reads must match in order to assemble reads into read pairs. Bioinformatics were performed with using the Quantitative Insights Into Microbial Ecology II pipeline QIIME 2 v.2022.2.0 (Bolyen et al., 2019). Adapters and primers were trimmed from the raw sequence data using cutadapt. Forward and reverse reads were then truncated, denoised, merged through their overlaps, and filtered for chimeras using DADA2 denoise-paired commands implemented in QIIME2 (Callahan et al., 2016). Sequences were clustered into amplicon sequence variants (ASVs) using the q2-feature-classifier (Bokulich et al., 2018) classify-sklearn naïve Bayes taxonomy classifier against the SILVA v.138 reference database (Klindworth et al., 2013). ASVs identified as Mitochondria were removed. ASVs assigned to Chloroplast were extracted from the entire 16S dataset, and further classified against the PhytoRef database (Decelle et al., 2015). Patrick White and Maria Cavaco analysed the sequencing data.

2.10 Filtering data for 16S Chloroplast rRNA amplicon sequencing comparison

In order to compare the RbcL concentration data with 16S chloroplast rRNA gene amplicon sequencing data, we classified the ASVs based on how they relate to the groups that our RbcL peptides target. We conducted two separate analyses, one examining how the 16S rRNA gene amplicon data compared to the groups we expect to target with our broad RbcL peptide, and a second examining how they compared to the taxon-specific RbcL peptides. For the second analysis, the 3 diatom peptides were grouped together for simplicity. Any ASVs in the 16S chloroplast data that were not measured by any of the RbcL peptides were classified as ‘other’. The remaining ASVs were grouped into one of 3 groups: diatom, haptophytes or *Micromonas sp.* Within each of these 3 groups there

were two distinctions based on how much taxonomic resolution the ASV was assigned: 1) ASVs from phytoplankton groups that contain enough taxonomic specificity to suggest that their originating organisms would encode the corresponding RBCL peptide and 2) ASVs that are associated with the assigned group, but lack enough taxonomic specificity to be certain that their originating organism would encode the associated RBCL peptide. Patrick White and I collaboratively conducted this sequence filtering.

2.11 Calculations of Rubisco as a percentage of total protein

Form 1 Rubisco contains 8 of each large and small subunits, therefore equimolar stoichiometry can be used in conjunction with the molecular weights of 55kDa and 15kDa for RbcL and RbcS respectively (Baker et al. 1975) to make calculations of Rubisco as a % of total protein. The following equation is used to calculate the %Rubisco of total protein for each peptide in our samples:

$$RbcL\ peptide\ in\ situ\ concentration\ \frac{fmol}{L} \times \text{vol. seawater filtered } (L)$$

$$\times \frac{1\ mol}{1X10^{15} fmol} = total\ mol\ RbcL\ in\ sample$$

$$mol\ RbcL = mol\ RbcS$$

$$Mass\ of\ RbcL = mol\ of\ RbcL\ \times 55\ 000 \frac{g\ RbcL}{mol}$$

$$Mass\ of\ RbcS = mol\ of\ RbcS\ \times 15\ 000 \frac{g\ RbcL}{mol}$$

$$(RbcL\ mass + RbcS\ mass) \div total\ sample\ protein\ (g) \times 100$$

$$= \%Rubisco\ of\ total\ protein$$

2.12 Estimating maximum potential carbon fixation rates

RbcL concentrations were converted to maximum potential carbon fixation rates. Since Rubisco contains 8 large subunits and has 8 active sites, 1 mol RbcL = 1 mol active site. Assuming CO_2 saturation and all active sites are active, we can use a maximum turnover rate of 0.4 mol C/mol RbcL (0°C; Young et al. 2015) to estimate the carbon fixation rates using the following equation:

$$Carbon\ Fixation\ Rate\ (mg\ C\ m^{-3}d^{-1}) = RbcL\left(\frac{fmol}{L}\right) \times K_{cat}\left(\frac{molC}{active\ site\ \cdot s}\right) \times \frac{1\ mol\ active\ site}{1\ mol\ RbcL} \times \frac{12.011\ g\ C}{mol} \times \frac{1000\ L}{m^3} \times \frac{86400\ s}{day}$$

Chapter 3: Results

3.1 Rubisco large subunit peptides:

In this study we measured six Rubisco large subunit (RbcL) peptides that were chosen to target what we believed to be the most relevant phytoplankton groups for our study area. Figure 2 summarizes the six RbcL peptides identified and measured in this study. The YESGVIPYAK peptide sequence is found in a large portion of eukaryotic phytoplankton RbcL sequences, excluding green algae and dinoflagellates. The remaining five peptides were selected to measure RbcL from specific groups of phytoplankton that are likely to be important in our study region (Freyria et al. 2021, Kalenitchenko et al. 2019, Joli et al. 2018). DYVAEGPQILR measures RbcL from a specific portion of the group Haptophyta which includes *Phaeocystis* species. FLYCMEGINR, FLNCMEGINR and FLNCLEGINR are found in the same region of diatom RbcL sequences and are specific to different groups within; FLYCMEGINR targets *Chaetoceros* species, FLNCMEGINR targets *Fragilariopsis* and *Pseudo-Nitzschia* species, and FLNCLEGINR targets *Thalassiosira* species. NVTLGFVDLMR measures a subset of green algae, mainly *Micromonas sp.* A comprehensive multiple alignment of RbcL sequences used for this analysis can be found in Appendix A (Figure S1).

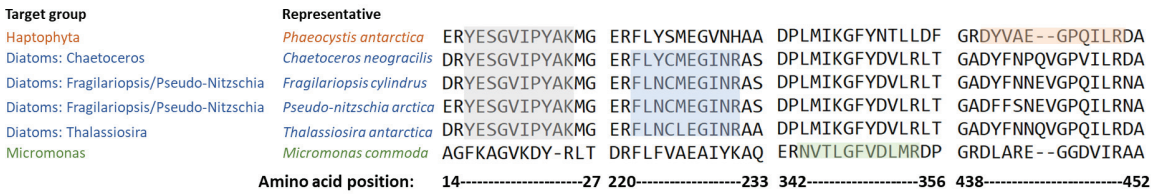


Figure 2. Partial multiple alignment of Rubisco large subunit (RbcL) sequences from select phytoplankton strains that represent the target groups for each of the RbcL peptides being measured in this study. 'YESGVIPYKAK' is a RbcL peptide that is found across a broad range of eukaryotic phytoplankton but is absent from green algae. The gap present in this alignment at the region of the highlighted Haptophyta peptide is not missing information; the diatom sequences have additional amino acids in this region that the Haptophyta does not. Only regions of the alignment containing the peptides of interest are displayed here.

A linear regression, shown in Figure 3, can be used to examine the relationship between these RbcL peptides and help us gain a sense of the quality of these novel measurements. The figure displays the correlation between the broad eukaryotic RbcL peptide and the sum of the specific RbcL peptide concentrations measured in this study. In a single-organism proteomic study, we would generally expect two peptides from the same protein to be present in cells at a relatively constant 1:1 proportion. Similarly, we expect a strong linear relationship between our broad eukaryotic RbcL peptide and the sum of the specific RbcL peptides. With an R^2 value of 0.86, this regression indicates there is such a relationship between the broad peptide and the sum of individual peptide concentrations, bringing a measure of confidence about our method. Because the broad eukaryotic RbcL peptide sequence is found across a broader range of phytoplankton than the individual group RbcL peptides measured here, we do not except a 1:1 ratio.

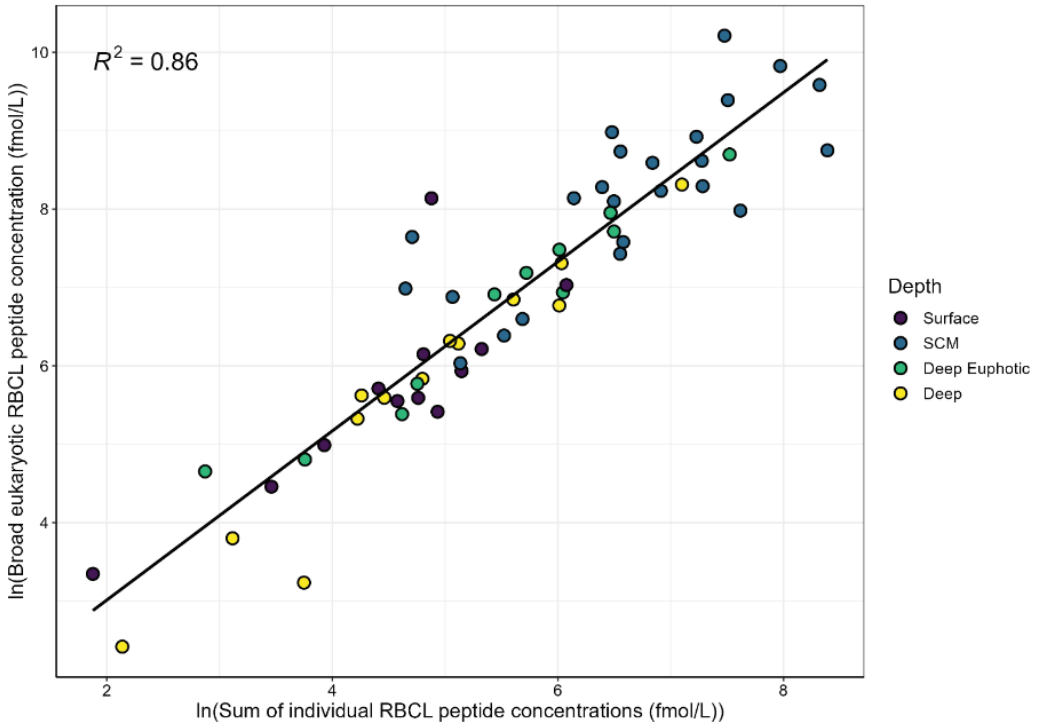


Figure 3. Linear regression (R^2 displayed) of broad eukaryotic Rubisco large subunit (RbcL) peptide (YESGVIPYAK) concentration versus the sum of the specific taxonomic RbcL peptide concentrations (DYVAEGPQILR, FLNCMEGINR, FLNCLENGINR, FLYCMEGINR). *Micromonas* RbcL peptide (NVTLGFVDLMR) concentration is excluded from the sum of individual taxonomic peptides since the broad RbcL peptide is not found within green algae. Samples are coloured by depth. SCM = subsurface chlorophyll maximum.

Proportion of phytoplankton community captured by RbcL peptides:

We compared results from our targeted protein-based method to a method more commonly used to explore phytoplankton community composition: chloroplast 16S rRNA gene amplicon sequencing (16S chloroplast), in order to 1) help us understand what portion of the phytoplankton community these RbcL peptides are likely to capture, and 2) assess how it compares to a commonly used method for phytoplankton community composition analysis and. Stacked bars in Figure 4 display relative proportions of the 16S chloroplast data from 6 different stations sampled on the *S/Y Vagabond* displayed as their expected coverage by both the broad eukaryotic RbcL peptide in combination with the *Micromonas* RbcL peptide (Figure 4A) as well as the taxon-specific RbcL peptides

(Figure 4B). All stations were sampled at the SCM except station 30 at Sverdrup glacier which also had a second, deep sample (z = 110m). The combination of specific RbcL peptides as well as the broad RbcL peptide appear to detect a large proportion of the phytoplankton community's Rubisco sequences in most cases. 'Other' in the 16S chloroplast rRNA gene amplicon sequencing data (Figure 4B; sequences that are not likely to be associated with organisms with Rubisco sequences that would contain the specific RbcL peptides) made up <15% of the reads with the exception of station 40 in South Cape Fiord. At this site, we estimate that 51% of the sequences are associated with organisms with Rubisco sequences that do not include our specific RbcL peptides. However, the majority of the sequences at station 40 that are not captured by the taxon-specific RbcL peptides are covered by the broad eukaryotic RbcL peptide (Figure 4A). The breakdown of the composition of the 'other' section for each of these samples in Figure 4B can be found in Figure S1. When we compare the RbcL concentration data in relative proportion and compare to the 16S chloroplast grouped into coverage by taxon-specific peptides (Figure 4C) we find the average proportion of *Micromonas* is higher in the RbcL data (29.5% ±16.9) relative to the 16S chloroplast data (5.10% ±7.62). The average proportion of Haptophyta and diatoms are similar across the two data types; Haptophyta make up 22.2% ±18.6 of the RbcL data and 17.1% ±18.4 of the 16S chloroplast data, while diatoms make up 48.2% ±29.4 of the RbcL data and 62.5% ±35.2 of the 16S chloroplast reads data.

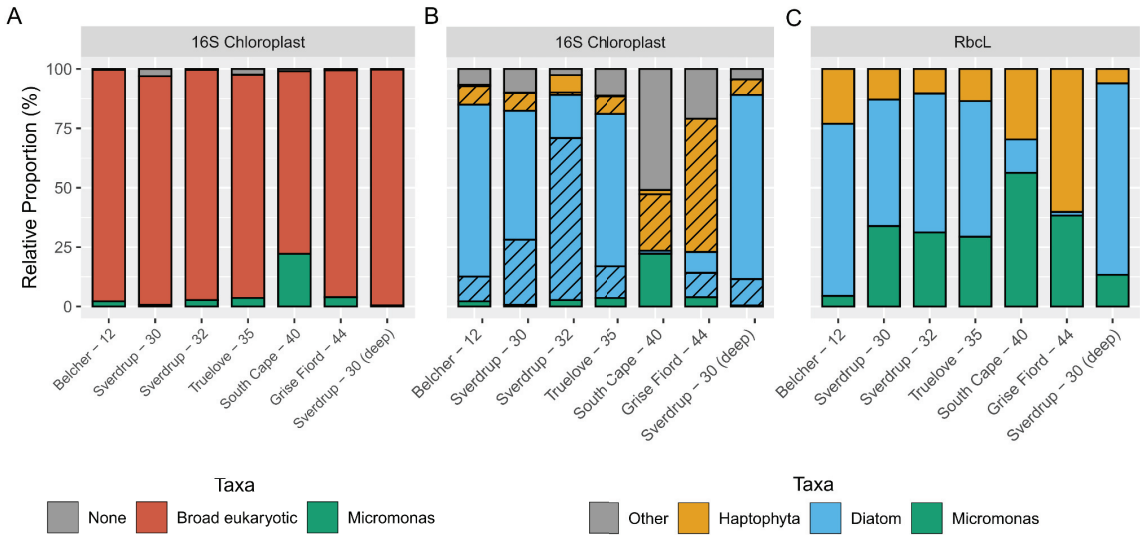


Figure 4. Comparison of phytoplankton community composition assessed via 16S chloroplast rRNA gene amplicon sequencing (A, B) and our taxon-specific RbcL quantification method. Unless otherwise noted, these samples were collected aboard the S/Y Vagabond and originated at the subsurface chlorophyll maximum. The 'deep' sample was collected at 110 m. Stations 12,30,32, and 40 are glacierized while stations 35 and 44 are non-glacierized. Panel A describes the 16S chloroplast sequencing data according to whether we expect the broad eukaryotic RbcL peptide to be present in Rubisco sequences associated with the 16S chloroplast amplicon sequence variants (ASVs). Panel B describes the same 16S sequencing data according to whether we expect these ASVs to be associated with Rubisco sequences encoding the taxon-specific RbcL peptides, combining the three diatom peptides. 'Other' in panels A and B represents amplicon sequence variants (ASVs) from phytoplankton groups that are not measured by any of the specific taxonomic RbcL peptides. Hashed bars in panel B indicate ASVs that are associated with the assigned group but lack enough taxonomic specificity to be certain that their originating organism would encode the associated RbcL peptide. The solid bars in A and B represent ASVs from phytoplankton groups that contain enough taxonomic specificity to suggest that their originating organisms would encode the corresponding RbcL peptide. Panel C describes the proportion of RbcL ascribed to each phytoplankton group based on peptide measurements.

3.3 Relationships between nutrients, light, Chlorophyll a, total protein and RbcL through the water column

We observed a diverse set of patterns and relationships between nutrients, light, Chl *a*, total protein and RbcL throughout the water column from the stations sampled by the *CCGS Amundsen* (Figure 5&6). Across these 12 stations, we measure a large range of RbcL concentrations from 0.011 to 14.5 pmol/L. RbcL concentrations of the diatom and

haptophyte groups reach maximum values most often at the SCM depth where nutrient concentrations have started to increase, while RbcL concentrations from the smaller green algae *Micromonas sp.* are commonly higher in surface waters where there are lower nutrient concentrations. Total protein concentration is sometimes higher at the surface depth (first depth, z = 5m) where nitrate concentration is low and the broad eukaryotic RbcL remains close to 0 pmol/L. However the maximum concentrations of both total protein and the broad eukaryotic RbcL peptide usually both occur at the subsurface chlorophyll maximum (SCM). The depth of the euphotic zone ranges from 41 to 90 meters (calculated as the depth where PAR = 0.1% of the surface value, Banse 2004), and across our stations, the SCM peak occurs as shallow as 7m and as deep as 40m, while the depth of the nutricline ([nitrate] > 0.05 μM (Cermeno et al. 2008)) varies from 2m to 33m. Examining patterns in these depth profile sets across our study region helps us 1) understand relationships between these variables and 2) explore how they are impacted by the presence of tidewater glaciers and other features across Jones Sound.

Stations in open Jones Sound, 2.3 and 2.4 (Figure 5a) and b) >20 km from shore, have two of the shallowest euphotic zone depths at 41 and 46 m respectively. Station 2.3 has a small Chl a peak (9 μg/L) at 15 m and a bigger, more distinct peak at 33 m (25 μg/L) while station 2.4 has one distinct Chl a peak at 25 m (22 μg/L). Although the SCM peak is quite distinct at both these stations, patterns in the protein and RbcL depth profiles are not so defined. Total protein concentration and *Chaetoceros* RbcL are quite high in the surface sample at station 2.3 around 30 μg/L and 0.125 pmol/L respectively. We then see RbcL concentration levels from both the broad eukaryotic peptide and all specific groups measured at both the small Chla peak and SCM peak at station 2.3, but

similar to other stations, *Chaetoceros* is still producing the highest RbcL concentration (0.125 pmol/L). At station 2.4, RbcL concentrations stay high just below the SCM as the chlorophyll peak drops off. At 2.4 we also see RbcL measurements below the euphotic zone, where RbcL concentrations typically drop close to 0. At the deepest depth sampled (140 m), RbcL concentrations remain high for both the broad peptide and all specific phytoplankton groups targeted (from 0.045 pmol/L (*Haptophyta)* to 0.170 pmol/L (*Thalassiosira*).

Across the eastern gateway to Jones Sound, stations 290, 291 and 292 (Figure 5c), d) and e)), have mid-range SCM depths between 15 to 30m. The euphotic zone depth at station 291 (77 m) is deeper than that at stations 290 and 292 (55 and 49 m) but all are relatively deep compared to stations closer to glaciers. The nutricline is relatively shallow at these stations 291 and 292, starting at 9.5 and 10.5 m respectively while the nutricline of station 290 is deeper around 30 m which is more similar to some of the glacierized stations such as 1.4 and 295 in South Cape Fiord and station 296 on the outskirt of Jones Sound. At station 290, Chl a, total protein and all RbcL peptides peak at the SCM depth. We see a clear SCM at 29m, with the largest RbcL concentrations coming from *Chaetoceros* (0.474 pmol/L) and *Micromonas* (0.375 pmol/L) similar to many other stations, and smaller contributions from Haptophyta (0.151 pmol/L) and *Thalassiosira* (0.078 pmol/L). The small Chl a SCM peak at 15m (5 µg/L), just below the nutricline, at station 291 corresponds well with where we see the highest broad eukaryotic RbcL concentration with the highest specific RbcL concentration coming from the *Chaetoceros* group (~0.250 pmol/L). At station 292 the total protein concentration (23 µg/L) is high in the surface water while the RbcL concentrations stays close to 0 pmol/L. Station 292 has

a much higher Chl a concentration at its SCM (21 µg/L) which also corresponds well with the peak in the broad eukaryotic RbcL concentration. *Chaetoceros* has the highest specific RbcL concentration at the SCM (0.421 pmol/L) while the RbcL concentration from *Micromonas* (0.232 pmol/L) increases at a deeper depth (31m) within the euphotic zone.

East of Jones Sound, stations 2.5 and 2.7 (Figure 5a and b) on the northwest edge of Baffin Bay, have relatively deep euphotic zone depths (90m) and the depth of the SCM is 35 and 33 meters respectively. The highest RbcL concentration from the broad eukaryotic peptide appears at the SCM depth for station 2.7, but at station 2.5 we see the broad eukaryotic RbcL value peak at a smaller, deeper Chla peak at 55m. *Chaetoceros* produced the highest concentration of RbcL (0.319 and 0.627 pmol/L) at both these sites followed by *Micromonas* (0.136 and 0.345 pmol/L). At station 2.5, the highest *Micromonas* RbcL concentration occurs at the surface depth, not the SCM depth.

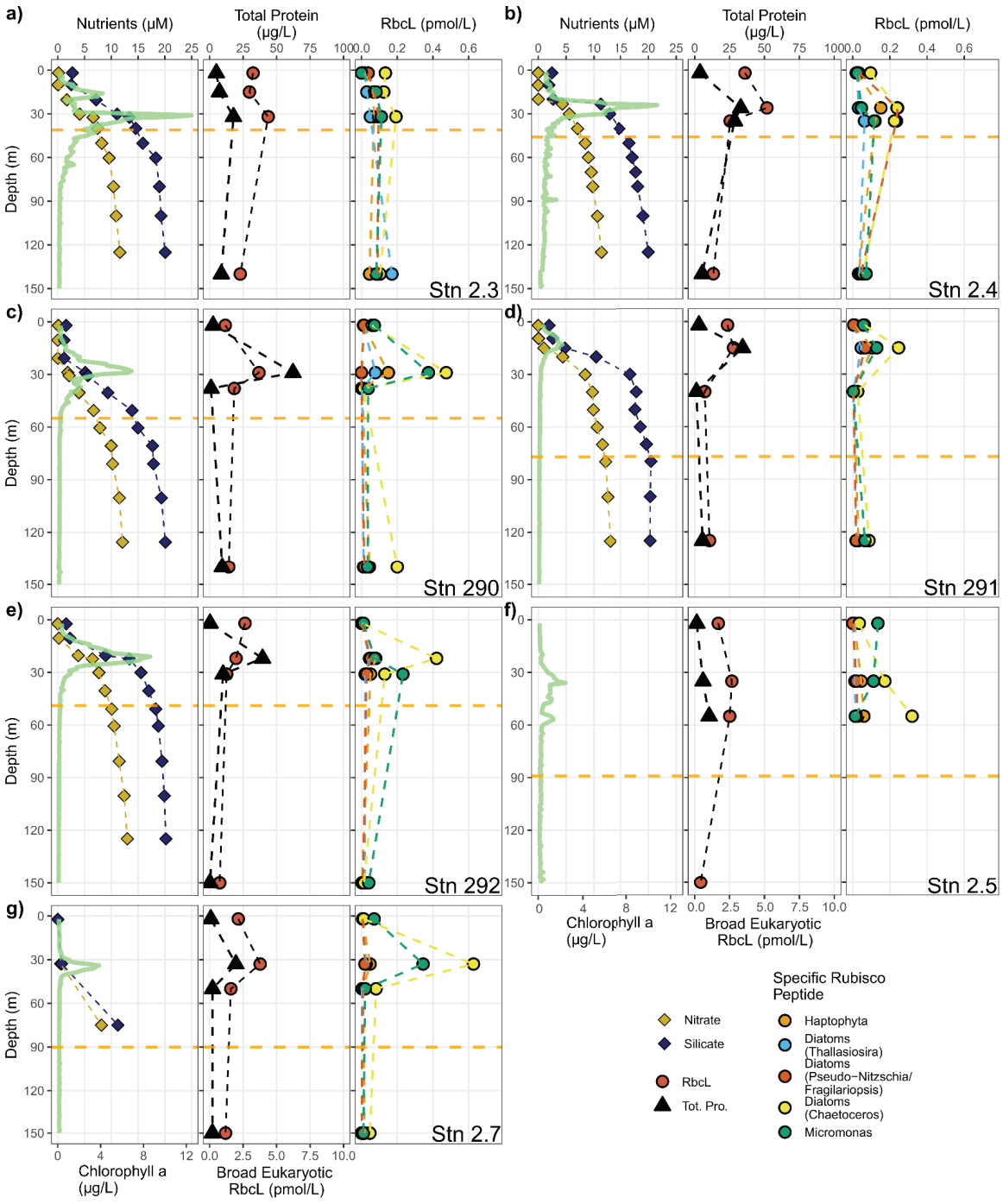


Figure 5. Depth profiles of sensor Chl a and bottle measurements of nitrate and silicate concentrations (left), total protein and broad eukaryotic Rubisco large subunit (RbcL) concentration (middle) and RcbL concentrations for 5 specific phytoplankton groups (right) for non-glacierized stations sampled from *CCGS Amundsen*: 2.3(a), 2.4(b), 290(c), 291(d), 292(e), 2.5(f), 2.7(g). Yellow dotted line across all panels indicates the estimated euphotic zone depth calculated as 0.1% of surface PAR for that station. Note differing x-axis scales between Figure 5 and Figure 6 for RbcL and total protein panels.

At the stations within glacierized South Cape Fiord, 1.4 and 295, the euphotic zone depth is 43 and 46m respectively (Figure 6a and b)), similar to euphotic zone depth observed in the open Jones Sound stations. At station 1.4, 18.5km from Sydkap Glacier and 9km from a smaller tidewater glacier also flowing into South Cape Fiord, two small Chl a peaks occur at 6 and 35m with the deeper peak being the SCM. Total protein and RbcL concentrations are highest in the surface water, with the highest RbcL concentration derived from *Micromonas*. Apart from *Chaetoceros*, whose RbcL concentration increases to 0.540 pmol/L at the SCM depth, RbcL concentrations for all targeted phytoplankton groups stay relatively low at the SCM (< 0.050 pmol/L). At station 295, located approximately 39 km away from Sydkap Glacier, a broad SCM occurs at 35m, below the nutricline depth of 25m. At this station, total protein concentration is highest at the surface, whereas RbcL concentrations are highest at the SCM depth. Of note, among our targeted groups, all have low (i.e. close to 0 pmol/L) RbcL concentrations in the surface water, apart from *Micromonas*, while *Chaetoceros* again has the highest RbcL concentration at the SCM.

Station 297 (Figure 6c) is located in front of Jakeman Glacier. The euphotic zone depth of this station is 45 m. The Chl a depth profile at station 297 shows a broad peak spanning from approximately 15 to 35 m. The nutricline starts very shallow at 2 m, but we see a much sharper increase in both nitrate and silicate (above 5 μM) towards the base of the broad Chl a peak. The total protein concentration is highest in the surface water and at 20 m, staying consistent around 65 μg/L at both of these depths. The maximum broad eukaryotic RbcL concentration appears the SCM, with the biggest RbcL concentration coming from *Chaetoceros* (0.935 pmol/L), but the *Psuedo-*

Nitzschia/Fragilariopsis sp. and *Micromonas sp.* groups also producing high concentrations of RbcL (0.403 and 0.347 pmol/L respectively). RbcL concentrations otherwise stay relatively low at all other depths. Depth profiles from two tidewater glaciers on the northeast edge of Devon Island, Belcher Glacier (station 293) and an unnamed glacier (station 296) just outside Jones Sound (Figure 6d and e) show defined peaks of the broad RbcL peptide at the SCM depth but more complicated patterns in RbcL concentrations among our targeted phytoplankton groups. Station 293 is located 20 km away from Belcher Glacier, and the euphotic zone depth here is 49 m, with a shallow SCM depth of 7 m right below the relatively shallow nutricline (7.7m). At this station, as at several other stations, the total protein and the broad eukaryotic RbcL concentrations peak at the same depth as the Chl a, but the patterns in RbcL concentration from individual phytoplankton groups are more unpredictable. High RbcL concentrations from the green algae *Micromonas* (0.249 pmol/L) and diatom *Chaetoceros* (0.336 pmol/L) are evident in the surface water, while the diatom groups *Pseudo-Nitschia/Fragilariopsis* instead peak slightly lower in the water column at the SCM depth.

Station 296 sits 21 km away from an unnamed tidewater glacier on Devon Island. The depth of the euphotic zone here is 44 m and there is a distinct SCM at 26 m, peaking at the same depth as we see an increase in nitrate (nutricline depth = 26 m). At the surface, the total protein concentration is elevated (110 µg/L) while RbcL concentrations remain close to zero. Below the surface, the broad eukaryotic RbcL concentration is highest at the SCM, with *Chaetoceros, Pseudo-Nitschia/Fragilariopsis*, and Haptophyta contributing the highest RbcL concentrations (2.365, 0.937 and 0.578 pmol/L respectively), with *Micromonas* remaining low at the SCM (0.012 pmol/L).

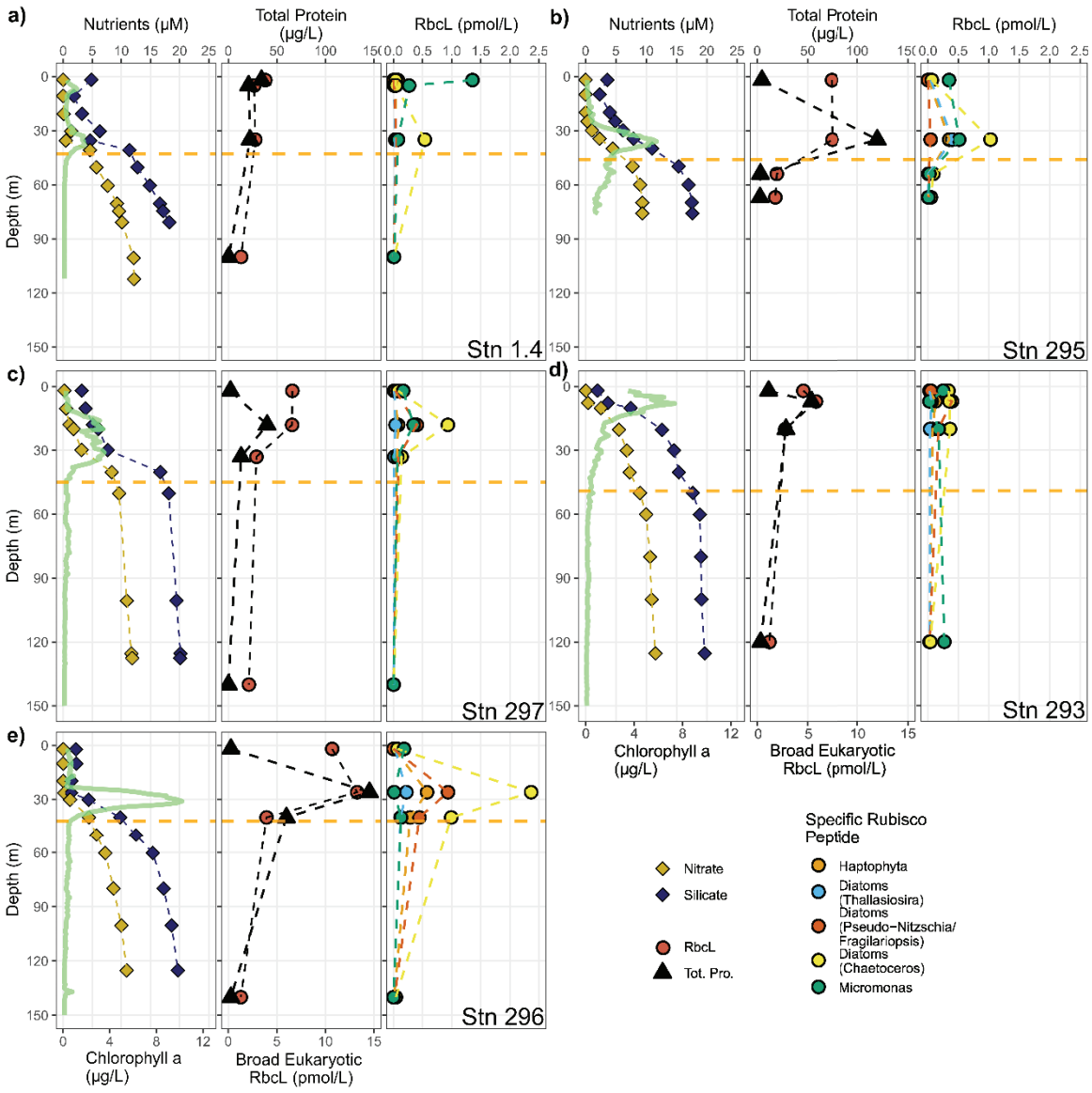


Figure 6. Depth profiles of sensor Chl a and bottle measurements of nitrate and silicate concentrations (left), total protein and broad eukaryotic Rubisco large subunit (RbcL) concentration (middle) and RcbL concentrations for 5 specific phytoplankton groups (right) for glacierized stations sampled from *CCGS Amundsen*: 1.4(a), 295(b), 297(c), 293(d), 296(e). Yellow dotted line across all panels indicates the estimated euphotic zone depth calculated as 0.1% of surface PAR for that station. Note differing x-axis scales between Figure 5 and Figure 6 for RbcL and total protein panels.

3.4 Regional patterns in Rubisco:

Grouping our stations based on proximity to glacier systems, we observe differences in the magnitude of Rubisco being produced, and type of phytoplankton

producing it, in these different marine environments. We can divide our stations into three regional categories: offshore (>20 km from shore), glacierized (<20 km from tidewater glacier) and non-glacierized (<20 km from shore, but > 20 km from any tidewater glacier). RbcL concentrations from each of the *Micromonas*, Haptophyta and the *Pseudo-nitszchia/Fragilariopsis* and *Chaetoceros* diatom groups appear across all three regions while RbcL from the *Thalassiosira* group is mostly present across glacierized stations only. The average RbcL concentration from *Thalassiosira* in the glacierized category is 0.175 pmol/L ± 0.149 while the average concentration across offshore stations is lower (0.0411 pmol/L ± 0.210) and even lower for non-glacierized stations (0.025 pmol/L ±1.91). However, this higher average RbcL concentration of *Thalassiosira sp.* at glacierized stations is not statistically significant compared to the offshore sites ($p = 0.55$, Wilcoxon Rank Sum test). The low number of sites classified as non-glacierized disqualifies this category from significance tests. Out of the other target group RbcL concentrations, we see significantly higher *Chaetoceros* RbcL concentrations at glacierized stations (0.820 pmol/L ± 0.630) compared to off-shore (0.339 pmol/L ± 0.171) with $p<0.05$ (Wilcoxon Rank Sum test). Although we see a larger number of phytoplankton groups contributing RbcL in the glacierized sites, our highest RbcL and Chla values come from the non-glacierized regions. Grise Fiord specifically has high concentrations of Haptophyta (4.291 pmol/L) and *Micromonas* (2.729 pmol/L) with very small concentrations from the three diatom groups (<0.050 pmol/L each).

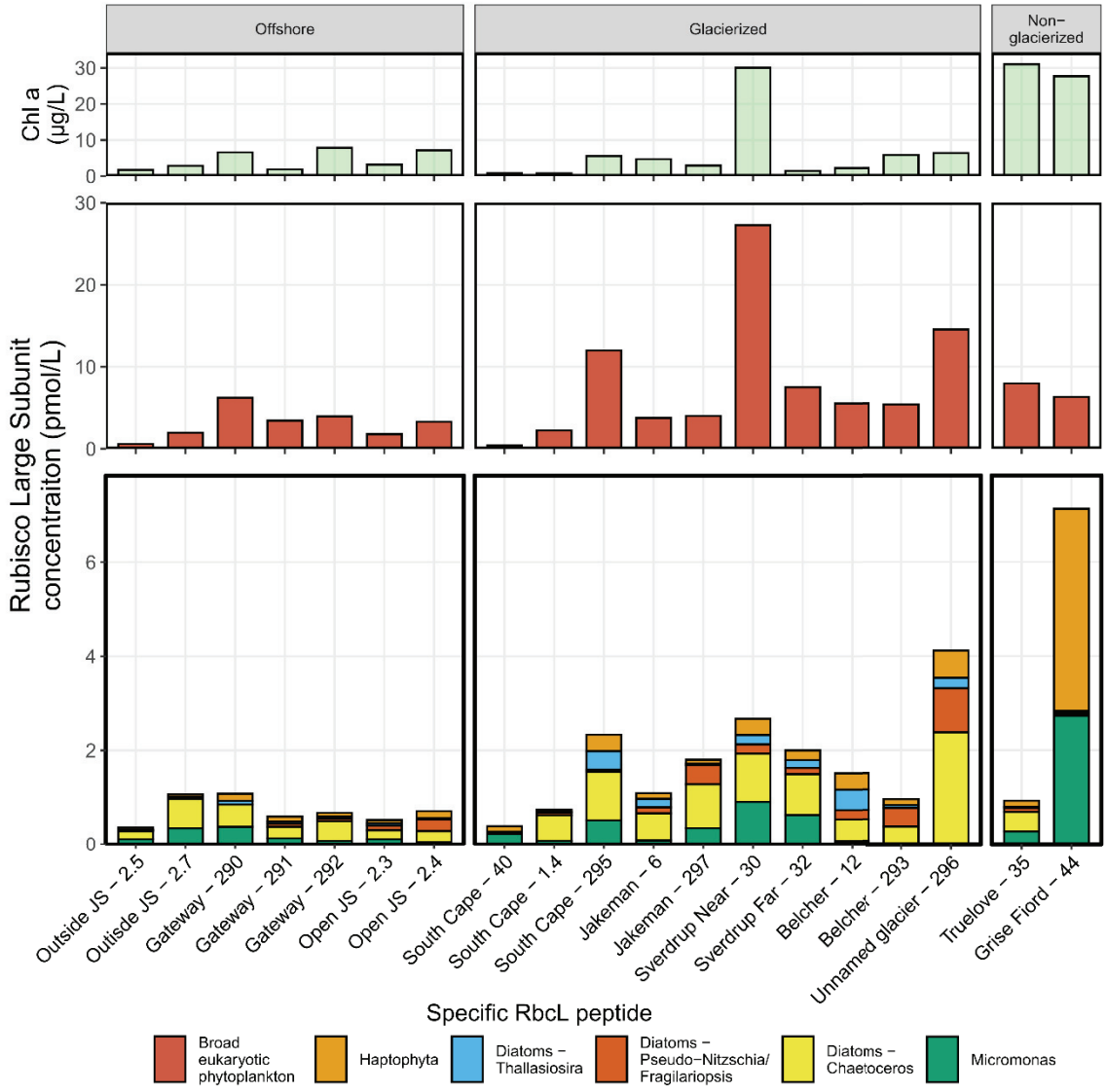


Figure 7. Bar plot of Chlorophyll a sensor values (top), broad eukaryotic RbcL peptide (middle), and specific RbcL peptide concentrations for five phytoplankton groups (bottom) at the sub-surface chlorophyll maximum (scm) depth for all stations sampled on the *S/Y Vagabond* and *CCGS Amundsen* in 2019.

3.5 Correlation of RbcL trends with environmental conditions

A Principapal Component Analysis (PCA) of the RbcL peptides (Figure 7) and a Spearman's rank correlation matrix of the RbcL peptides compared with environmental factors (Figure 8) reveal environmental drivers of patterns in our RbcL concentrations. The PCA of RbcL concentrations from all stations at the SCM depth reveals that the

Micromonas and haptophyta peptides group away from the three diatom peptides as well the broad eukaryotic peptide. RbcL concentrations from diatom groups and the broad eukaryotic RbcL peptide are driven by glacierized sites while haptophyte and *Micromonas* RbcL concentrations are driven by non-glacierized sites (Figure 7).

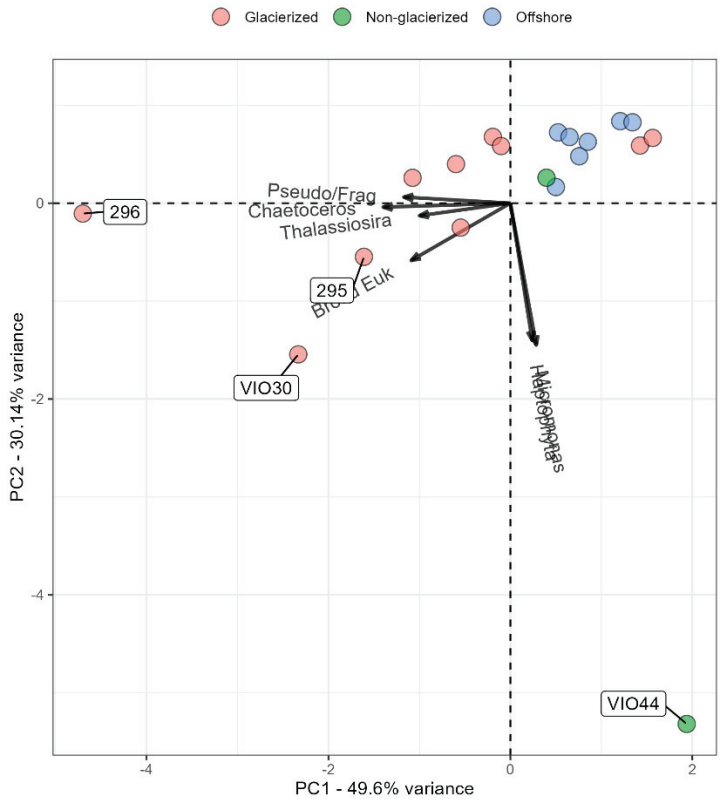


Figure 8. Principal component analysis (PCA) of Rubisco large subunit (RbcL) peptide concentrations from all stations and depths sampled. Data are coloured by location (red = glacierized, green = non-glacierized and blue = offshore).

The correlogram highlights that *Micromonas* RbcL patterns are distinct from the other target RbcL groups, displaying relationships that indicate trends driven by properties throughout the water column (Figure 8). This analysis showed a significant ($p<0.01$) negative correlation between *Micromonas* and salinity, density, nitrate and silicate. In contrast, the Haptophyta, diatom, and broad eukarotyic peptides show no significant relationship with these factors. Additionally, *Micromonas* also had a significant ($p<0.05$)

positive correlation with temperature and PAR, which was not present in any of the other peptides examined in this study.

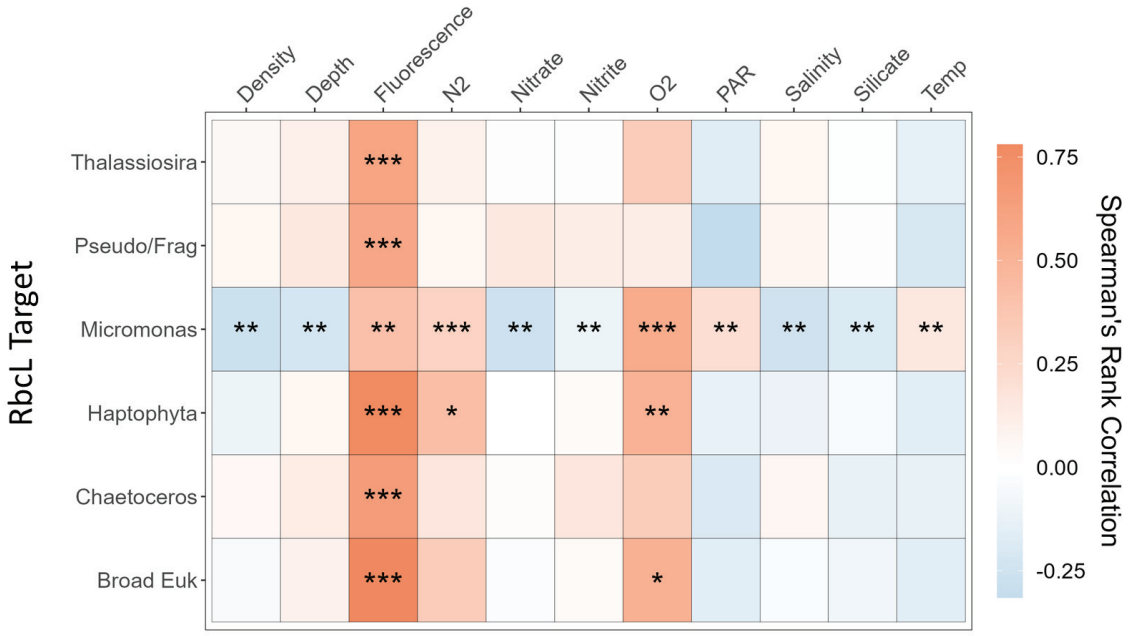


Figure 9. Correlogram displaying pairwise comparisons of environmental factors and RbcL peptide concentrations. Colour represents strength/direction of correlation using the Spearman's rank coefficient. Red = positive correlation, blue = negative correlation. Asterisks within squares represent the significance level *p*-value from Spearman's *Rho* ($p < 0.05$ *, $p < 0.01$ **, $p < 0.001$ ***). Blank squares indicate the correlation is insignificant. Environmental variables: water temperature (Temp), photosynthetically available radiation (Par), dissolved oxygen (O2), Brunt-Vaisala Frequency (N2), Chl *a* fluorescence (Fluo), salinity (Sal).

Chapter 4: Discussion

4.1 16S chloroplast comparison

We expect that the broad eukaryotic RbcL peptide as well as the selection of specific RbcL peptides measured in this study, targeting three groups of diatoms, a green algae group (*Micromonas* sp.), and haptophytes, covered a large proportion of the phytoplankton community that was identified by chloroplast 16s rRNA gene sequencing (Figure 4). This implies that we are capturing the major players in these communities with our RbcL peptides, and that we can make meaningful conclusions about the Jones Sound phytoplankton community as a whole using these peptide data. However, there are some caveats and exceptions. Specifically, at station 40 located in South Cape Fiord where the RbcL peptides appear to only capture ~50% of the photosynthetic community (Figure 4), approximately half of the 'other' section here is composed of organisms of the family *Chroomonadaceae* (Figure S1). This family falls within the class *Cryptophyta* which are believed to be capable of mixotrophy (Jansson et al., 1996). Cryptophyes have previously been identified in northern Baffin Bay in early fall, exhibiting a significant negative correlation with nitrogen (Blais 2017), and also in the western CAA in early spring, before the diatom-dominated spring bloom (Terrado et al. 2011). Collectively, these reports indicate Cryptophytes are likely to be present pre- or post-bloom conditions in this region and additional peptides to quantify their contributions to RbcL in the CAA should be developed. Initially we hypothesized that photosynthetic dinoflagellates would make significant contributions to the phytoplankton community in our samples based on previous surveys of the eukaryotic microbial communities in the CAA (Freyria et al., 2021; Kalenitchenko et al., 2019; Lovejoy et al., 2002). However, these studies utilized

18S rRNA gene amplicon sequencing which captures both the heterotrophic and photosynthetic microbial community, and thus, these previous works were only able to speculate on the photosynthetic capabilities of dinoflagellates present in the CAA. Chloroplast 16S rRNA sequencing data presented here (Figure S1) does not include dinoflagellates, but paired 18S rRNA sequencing data (J. Spence, P. White, unpublished) has resolved the presence of this group in Jones Sound. Taken together, these results suggest that dinoflagellate species present in this region are largely heterotrophic and not contributing to carbon fixation.

Our 16S chloroplast rRNA gene sequencing and RbcL peptides resolved relatively similar proportions of our targeted groups (Figure 4). This is in part surprising, as the abundance of these biomolecules per cell or per unit biomass is expected to vary differently between taxonomic groups (Needham & Fuhrman, 2016). We also expect that RbcL concentrations per cell or per unit biomass would change depending on the physiological status of the plankton (eg. growth phase, nutritional status) (Losh et al., 2013) while chloroplast 16S copies would be driven largely by number of chloroplasts, which can also vary with physiological changes but to a lesser degree than expected with RbcL concentrations (Needham & Fuhrman, 2016). The target group with the most notable proportional difference between these two datasets is *Micromonas,* suggesting that they may produce a higher Rubisco to chloroplast ratio compared to the other plankton groups assessed here. However, while the amount of Rubisco produced by *Micromonas* has not yet been assessed in culture, measurements from other green algae suggest that Rubisco makes up a smaller percentage of total protein in this group relative to diatoms (Losh et al 2013).

4.2 Ecological and biogeochemical significance of target phytoplankton groups

Diatoms are classically considered to be important for carbon export into the deep ocean as well as quality food for higher trophic levels in the marine food ecosystem (Smetacek, 1999). For this reason, in this study, we are particularly interested in whether nutrient additions at the termini of tidewater glaciers promote the presence of diatom-dominated communities during summer months in Jones Sound. Our results show that glacierized locations have a significantly ($p<0.05$, Wilcoxon Rank Sum test) higher RbcL contribution from the diatom *Chaetoceros* sp. compared to offshore sites (Figure 7). *Chaetoceros* sp. have been reported to be the dominant diatom group across the CAA during bloom periods (Booth et al. 2002, Comeau et al. 2011, Kalenitchenko et al. 2019, Freyria et al. 2021). However, because *Chaetoceros* sp. are a smaller centric diatom, they have also been reported to persist in-between bloom periods, when nutrient concentrations are depleted (Booth et al. 2002, Tremblay et al. 2006). Thus, the ubiquity of this species during both bloom and in-between periods could have important implications for carbon export throughout the year. Although we see higher concentrations of *Chaetoceros* RbcL at stations classified as glacierized, RbcL concentrations from this group are present across virtually all stations sampled in this study with the exception of the non-glacierized site, Grise Fiord (Figure 7). Additionally, although not significant ($p > 0.05$, Wilcoxon Rank Sum test), average RbcL concentrations from the diatom group *Thalassiosira* are higher at our glacierized stations compared to offshore (Figure 7). *Thalassiosira* include larger diatoms that are usually present at the onset of spring blooms in the CAA, but whose populations decline soon after once nutrient concentrations being to fall (Lovejoy 2002, Tremblay et al 2006,

Freyria et al. 2021). This pattern of greater *Thalassiosira* sp. abundance under replete nutrient conditions is consistent with our finding of higher RbcL concentrations from *Thalassiosira* sp. near tidewater glaciers, which can supply nutrients to the surface waters throughout the summer (Bhatia et al., 2021).

Micromonas sp. have been identified as an important and ubiquitous picophytoplankton group across the Arctic (Balzano et al., 2012; Joli et al., 2017). These small cells can often dominate the phytoplankton community in the CAA (Lovejoy et al. 2007) and the dominant strain *M. polaris* has adapted to low temperatures and light in order to thrive in polar conditions throughout the year (Lovejoy et al 2006, Liefer et al. 2018). RbcL concentrations from *Micromonas* are elevated in the surface water more than any other targeted group from this study (Figures 5 and 6) and exhibit a significant negative correlation with nutrients and salinity (Figure 8). Generally, smaller phytoplankton, like *Micromonas*, are able to thrive in low-nutrient conditions, and play important roles in nutrient cycling and the microbial loop (Fenchel, 2008). In this study, the presence of *Micromonas* in more stratified, nutrient-deplete environments within Jones Sound is consistent with this role. The highest *Micromonas* RbcL concentrations we observe appear in Grise Fiord (Figure 6), a fiord with no influence from glaciers, low nitrate concentrations (Bhatia et al. 2021) and a distinct lack of diatom RbcL (Figure 6). Though our dataset is limited, it suggests that *Micromonas* could become increasingly important in CAA coastal regions as tidewater submarine discharge and its associated nutrient delivery to surface water declines in the future. Finally, though smaller phytoplankton species, like *Micromonas,* have not traditionally been considered major players in carbon export from the surface to the deep ocean, new evidence shows this

could change when *Micromonas* is the majority of the phytoplankton community (Bachy et al., 2022). Data from Fram Strait reveals *Micromonas* comprising up to ~10% of all chloroplast amplicon sequences in sediment traps measured below a *Micromonas* dominated phytoplankton community (Bachy et al., 2022).

RbcL concentrations from Haptophyta, mainly *Phaeocystis* in this region, are present in low concentrations throughout the SCM depths across most stations, but, similar to *Micromonas,* are markedly elevated in Grise Fiord (Figure 6). *Phaeocystis* sp. can be small free-living cells or form colonies when in bloom conditions (Lancelot et al., 1994). When present in smaller numbers this phytoplankton group is most likely contributing to nutrient cycling and the microbial loop, but when conditions are right for *Phaeocystis* to form a larger colony it also has the potential to contribute to carbon export (Verity et al., 2007)

At several stations in this study (Fig 5, panels a,b,c,d,e), we see elevated RbcL concentrations at the deepest depth sampled, below the euphotic zone. We do not expect that this RbcL was produced *in-situ* at that depth, but rather that it is likely present at depth as a result of sinking events. At station 290, we see an elevated RbcL concentration from just one specific target group at depth, *Chaetoceros* (Figure 5c), but at other stations (2.3, 2.4, 291 - Figure 5 a, b and d) we see elevated deep RbcL concentrations from several target groups, suggesting that multiple species are contributing to this export. Previous work has reported the utility of Rubisco measurements to examine carbon export to the deep ocean (Orellana & Hansell, 2012). Orellana & Hansell measure Rubisco at depth in the form of a polymer microgel which falls along the continuum of dissolved and particulate organic material and so these data are not fully analogous to our

RbcL measurements made in particulate form in this study. Nonetheless, Rubisco in microgel form was found in concentrations up to ~20 ng mL^{-1} at depths of 2000 m and deeper in the North Pacific. This polymer microgel form seems to protect proteins from biodegradation and enables detection in the very deep ocean, allowing the tracking of carbon export from phytoplankton. Future efforts should examine the relationship between particulate and microgel Rubisco concentrations during export events to explore the utility of deep RbcL measurements to track carbon export.

4.3Rubisco as a fraction of total protein

Using our dataset, we can calculate the percentage of total protein that is made up of Rubisco. This is of interest because it allows us to compare these data to previous studies, and because it has the potential to provide some insight into the trophic status as well as physiological status of the plankton community. Since our broad eukaryotic RbcL peptide targets such a large portion of the phytoplankton community (excluding green algae, Figure 2), we can use the concentrations of this peptide in conjunction with the *Micromonas* RbcL peptide to estimate what percentage of the total protein is made up of Rubisco. In our SCM depth samples this calculated percentage ranges from 0.19 - 2.0%. Previous studies have shown that temperature appears to be the primary driver of what percentage of total protein is made up of Rubisco in pure cultures and in phytoplankton-dominated field samples (Losh et al., 2013; Young et al., 2016; Young, Kranz, et al., 2015). Another factor that also appears to influence the percentage of protein that is comprised of Rubisco is nutrient starvation, but the magnitude of this impact appear small compared to temperature (Losh et al., 2013). Field observations of the Rubisco fraction of total protein in Southern Ocean plankton samples and cultures of polar diatom

F. cylindrus are high, between 10-20%, which is attributed to particularly poor temperature adaptation (Young, Goldman, et al., 2015; Young, Kranz, et al., 2015) . In contrast, cultures grown at 20°C under nutrient-replete conditions revealed relatively low Rubisco as a percent of total protein, from 1.4% to 3.7% across 8 species of phytoplankton including diatoms, haptophytes and green algae. An environmental community off the Southern California coast exhibited even lower Rubisco fractions (0.3% to 2.5%) when incubated at the *in situ* temperature of 16°C. Perhaps surprisingly, the calculated percentage range of Rubisco in total protein from our data, collected in an Arctic region with water temperatures hovering around 0°C, align better with low values from mesophilic cultures and environmental data than the psychrophilic diatom cultures and Southern Ocean environmental samples. It is important to note that all previous data used to calculate % Rubisco were generated using western blots rather than quantitative mass spectrometry. Our laboratory has performed peptide-based determinations of % Rubisco in the polar diatom *F. cylindrus*, with results ranging from 3-8% (C. Albury et al. unpublished). This range is lower than that reported by Young et al. (2015) and is more consistent with our field data. Future work should continue to examine any systematic differences in these approaches to measuring Rubisco content. Additionally, the percentages we measure in Jones Sound could be driven lower by higher contributions of heterotrophic microbial biomass compared to the intense phytoplankton bloom conditions examined in the Southern Ocean by Young et al. (2015).

4.4 Estimating maximum carbon fixation potential

Rubisco is a very inefficient enzyme with low turnover rates, and, as such, is usually considered to be the rate limiting step in the Calvin cycle if light is plentiful and not

limiting ATP production (Badger et al., 1998; MacIntyre & Geider, 1996). Thus, we expect that the Rubisco concentrations we measure are likely a reflection of carbon fixation rates in the absence of light limitation. In light limited situations we expect that measured Rubisco concentrations would not reflect carbon fixation rates. Since light is plentiful in the Arctic summer months, we assume that Rubisco, and its slow carboxylation rate, are likely the rate-limiting step in carbon-fixation at many of our sampling locations. However, at some of our stations, light penetration in the water column is restricted by sediment delivery from glaciers, suggesting that light limitation is very possible. Given this, as well as the fact that we have not yet demonstrated a direct relationship between carbon fixation rates and our Rubisco measurements, we conservatively consider our measurements to be a reflection of phytoplankton group-specific carbon fixation potential.

When we convert our Rubisco concentrations into maximum potential carbon fixation rate estimates, we see our data fits within the published literature of primary productivity measurements made *in situ* in nearby regions (Table 1). Our estimated maximum potential carbon fixation rates are reasonable, given previous primary productivity measurements made in Jones Sound during a similar time of year ranged from 0.001-2.016 with an average of 0.43 mgC $m^{-3}d^{-1}$ (Irwin et al. 1980). The range and average values of primary productivity from the broader context of Baffin Bay are higher than those of our calculated rates but still within a similar magnitude (Table 1). From our Rubisco concentration data we can also calculate taxon-specific rates of carbon fixation. The average rates of carbon fixation across diatoms, haptophytes and *Micromonas sp*. is similar (0.12, 0.16 and 0.15 mg C $m^{-3}d^{-1}$). Although these different phytoplankton groups

have similar average estimated rates of carbon fixation, diatoms are contributing more to carbon fixation at stations closer to glaciers while haptophytes and *Micromonas* contribute more at non-glacierized sites (Figure 7).

Table 1. Ranges and mean values of primary productivity rates across Jones Sound and Baffin Bay taken at the subsurface chlorophyll maximum (SCM) depth.

Location	Date (month-year)	Taxonomic Coverage	P.P. (mgC m^{-3} day^{-1})		Reference
			Range	Average *(n)*	
Baffin Bay	Sep-78	Total community	5.7-25.1	13.2 (12)	Harrison et al. 1982[a]
Jones Sound	Aug/Sep-83	Total community	0.001 - 2.02	0.43 (8)	Irwin et al. 1980[a]
Jones Sound	Aug-19	Total community	0.01-11.7	1.25 (19)	Present study[b]
Jones Sound	Aug-19	Diatoms	0-1.55	0.12 (19)	Present study[b]
Jones Sound	Aug-19	Haptophytes	0.02-1.78	0.16 (19)	Present study[b]
Jones Sound	Aug-19	*Micromonas sp.*	0.01-1.13	0.15 (19)	Present study[b]

[a]Carbon fixation rates measured via ^{14}C stable isotope tracer experiments
[b]Carbon fixation rates estimated from RbcL concentration data

Chapter 5. Conclusions

5.1 Overview

Through the development and application of a targeted metaproteomic approach to measure the concentration Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunits (RbcL) from key phytoplankton groups in glacierized and non-glacierized regions in Jones Sound, NU, we assessed: 1) the utility of a suite of RbcL peptide measurements in the CAA and 2) the amount of Rubisco different phytoplankton groups are contributing and 3) the maximum carbon fixation potential. In this work, we established that we were able to capture the majority of the phytoplankton community with the RbcL peptides measured in our study, albeit with additional improvements that could be made to optimize the proportion of the community resolved. Measurements of RbcL concentrations reveal that the diatom group *Chaetoceros* appears in higher abundance close to tidewater glaciers and that the small picoeukaryote *Micromonas* is ubiquitously present across diverse conditions. This latter finding indicates that in the future, *Micromonas* will most likely be an important part of the carbon cycle and coastal Arctic marine food webs.

5.2 Limitations and future works

It is evident there is room for improvement in the selection of RbcL peptides in order to fully resolve all the phytoplankton groups present at our stations. Published methodology for measuring RbcL concentrations with Western Blots uses RbcL global antibodies designed against peptide tags which are conserved across all photosynthesizers (Campbell et al., 2003). In contrast, our tryptic peptide-based approach, via targeted mass spectrometry, requires a suite of measurements, rather than a single one, to be inclusive

of the phytoplankton community as a whole. This has its advantages and disadvantages in that it allows us to obtain phytoplankton group-specific resolution with our measurements, but it also means that these peptides need to be continually re-evaluated as we learn more about the composition of phytoplankton communities and as these communities change.

An important additional parameter that can help us understand the effect of glaciers on phytoplankton growth is turbidity, as glaciers can deliver large loads of suspended sediment to the ocean. This particulate matter in the glacial melt has the potential to limit light penetration in surface waters, so may impact phytoplankton growth (Hopwood et al., 2020). Future work requires more careful consideration of turbidity and suspended sediment measurements in relation to carbon fixation potential. In particular, it would be useful to have higher resolution sampling near glacier termini to assess the spatial extent of the impact the turbid glacial plume on phytoplankton composition and function (e.g., protein production, nutrient uptake). Future efforts should also consider measuring gross and net primary productivity rates in addition to calculations of the proportion of total protein that can be attributed to each phytoplankton group of interest via global metaproteomics (McCain et al., 2022). The addition of these measurements will allow us to best leverage our Rubisco concentrations to learn about phytoplankton physiology and biochemistry *in-situ*.

5.3 Implications and significance

The work produced by this thesis has both scientific and societal significance. Scientifically, it fills a knowledge gap about the influence of tidewater glaciers on phytoplankton communities in the CAA. This is done using a novel method which yields

information about not only which phytoplankton groups are facilitating primary production but also quantifies each groups contribution. Societally, this work was conducted in Jones Sound, home to the Inuit hamlet of Ausuittuq (Grise Fiord). Hunting of marine mammals in Jones Sound is vital to the food security of the people living in this community. As such, they are deeply interested in projections of ecosystem health for this region. Measuring carbon fixation potential across a spectrum of glacierized and non-glacierized environments within Jones Sound helped us gain insight into how phytoplankton communities at the base of marine food webs may respond to ongoing environmental change in the Arctic. Since 2019, our team has continued to work collaboratively with the community of Ausuittuq, making more extensive measurements across Jones Sound with the view to integrate these data into conceptual and numerical models of future ocean change in this region. As our collaborations with the community grow and more knowledge and skills are shared between Southern scientists and community members, data collection and monitoring has evolved into a year-round program, led by local Inuit, spanning the ice-covered to open water seasons. Collaboration with local communities is key for studies in remote high Arctic regions like Jones Sound, where data has been most commonly gathered only in the summer. Knowledge sharing is imperative to attaining a full appreciation of the historical context of the sites we study and also to designing research which is mutually beneficial to both southern scientists and northern community members.

Jones Sound is a marine high Arctic region that is surrounded by ice caps and fields, and thus, one that is inundated with meltwater input from tidewater and land-terminating glaciers. Developing metaproteomic-based approaches to quantitatively

assess phytoplankton group-specific carbon fixation potential will be important for continued monitoring of this dynamic seascape as glaciers in the region react to climate change. Moreover, if adapted and applied to other marine regions, the tool presented in this thesis for assessing carbon fixation potential across phytoplankton groups presents exciting opportunities to gain additional understanding about which groups are contributing significantly to marine primary productivity and the physiological basis of those contributions.

References

Andersson, I., & Backlund, A. (2008). Structure and function of Rubisco. *Plant Physiology and Biochemistry*, *46*(3), 275–291. https://doi.org/10.1016/j.plaphy.2008.01.001

Apollonio, S. (1973). Glaciers and Nutrients in Arctic Seas. *Science*, *180*(4085), 491–493. https://doi.org/10.1126/science.180.4085.491

Arrigo, K. R., van Dijken, G. L., Castelao, R. M., Luo, H., Rennermalm, Å. K., Tedesco, M., Mote, T. L., Oliver, H., & Yager, P. L. (2017). Melting glaciers stimulate large summer phytoplankton blooms in southwest Greenland waters: Glaciers Stimulate Phytoplankton Blooms. *Geophysical Research Letters*, *44*(12), 6278–6285. https://doi.org/10.1002/2017GL073583

Bachy, C., Sudek, L., Choi, C. J., Eckmann, C. A., Nöthig, E.-M., Metfies, K., & Worden, A. Z. (2022). Phytoplankton Surveys in the Arctic Fram Strait Demonstrate the Tiny Eukaryotic Alga Micromonas and Other Picoprasinophytes Contribute to Deep Sea Export. *Microorganisms*, *10*(5), Article 5. https://doi.org/10.3390/microorganisms10050961

Badger, M. R., Andrews, T. J., Whitney, S. M., Ludwig, M., Yellowlees, D. C., Leggat, W., & Price, G. D. (1998). The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO2-concentrating mechanisms in algae. *Canadian Journal of Botany*, *76*(6), 1052–1071. https://doi.org/10.1139/b98-074

Balzano, S., Gourvil, P., Siano, R., Chanoine, M., Marie, D., Lessard, S., Sarno, D., & Vaulot, D. (2012). Diversity of cultured photosynthetic flagellates in the northeast Pacific and Arctic Oceans in summer. *Biogeosciences*, *9*(11), 4553–4571. https://doi.org/10.5194/bg-9-4553-2012

Bartram, A. K., Lynch, M. D. J., Stearns, J. C., Moreno-Hagelsieb, G., & Neufeld, J. D. (2011). Generation of Multimillion-Sequence 16S rRNA Gene Libraries from Complex Microbial Communities by Assembling Paired-End Illumina Reads. *Applied and Environmental Microbiology*, *77*(11), 3846–3852. https://doi.org/10.1128/AEM.02772-10

Bhatia, M. P., Kujawinski, E. B., Das, S. B., Breier, C. F., Henderson, P. B., & Charette, M. A. (2013). Greenland meltwater as a significant and potentially bioavailable source of iron to the ocean. *Nature Geoscience*, *6*(4), 274–278. https://doi.org/10.1038/ngeo1746

Bhatia, M. P., Waterman, S., Burgess, D. O., Williams, P. L., Bundy, R. M., Mellett, T., Roberts, M., & Bertrand, E. M. (2021). Glaciers and Nutrients in the Canadian Arctic Archipelago Marine System. *Global Biogeochemical Cycles*, *35*(8), e2021GB006976. https://doi.org/10.1029/2021GB006976

Blais, M., Ardyna, M., Gosselin, M., Dumont, D., Bélanger, S., Tremblay, J.-É., Gratton, Y., Marchese, C., & Poulin, M. (2017). Contrasting interannual changes in phytoplankton productivity and community structure in the coastal Canadian Arctic Ocean. *Limnology and Oceanography*, *62*(6), 2480–2497. https://doi.org/10.1002/lno.10581

Bliss, A., Hock, R., & Radić, V. (2014). Global response of glacier runoff to twenty-first century climate change. *Journal of Geophysical Research: Earth Surface*, *119*(4), 717–730. https://doi.org/10.1002/2013JF002931

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., Huttley, G. A., & Gregory Caporaso, J. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, *6*(1), 90. https://doi.org/10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), Article 8. https://doi.org/10.1038/s41587-019-0209-9

Booth, B. C., Larouche, P., Bélanger, S., Klein, B., Amiel, D., & Mei, Z.-P. (2002). Dynamics of Chaetoceros socialis blooms in the North Water. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(22), 5003–5025. https://doi.org/10.1016/S0967-0645(02)00175-3

Box, J. E., Colgan, W. T., Christensen, T. R., Schmidt, N. M., Lund, M., Parmentier, F.-J. W., Brown, R., Bhatt, U. S., Euskirchen, E. S., Romanovsky, V. E., Walsh, J. E., Overland, J. E., Wang, M., Corell, R. W., Meier, W. N., Wouters, B., Mernild, S., M\aard, J., Pawlak, J., & Olsen, M. S. (2019). Key indicators of Arctic climate change: 1971–2017. *Environmental Research Letters*, *14*(4), 045010. https://doi.org/10.1088/1748-9326/aafc1b

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), Article 7. https://doi.org/10.1038/nmeth.3869

Campbell, D. A., Cockshutt, A. M., & Porankiewicz-Asplund, J. (2003). Analysing photosynthetic complexes in uncharacterized species or mixed microalgal communities using global antibodies. *Physiologia Plantarum*, *119*(3), 322–327. https://doi.org/10.1034/j.1399-3054.2003.00175.x

Cook, A. J., Copland, L., Noël, B. P. Y., Stokes, C. R., Bentley, M. J., Sharp, M. J., Bingham, R. G., & Broeke, M. R. van den. (2019). Atmospheric forcing of rapid marine-terminating glacier retreat in the Canadian Arctic Archipelago. *Science Advances*, *5*(3), eaau8507. https://doi.org/10.1126/sciadv.aau8507

Decelle, J., Romac, S., Stern, R. F., Bendif, E. M., Zingone, A., Audic, S., Guiry, M. D., Guillou, L., Tessier, D., Gall, F. L., Gourvil, P., Santos, A. L. D., Probert, I., Vaulot, D., Vargas, C. de, & Christen, R. (2015). PhytoREF: A reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. *Molecular Ecology Resources*, *15*(6), 1435–1445. https://doi.org/10.1111/1755-0998.12401

Fenchel, T. (2008). The microbial loop – 25 years later. *Journal of Experimental Marine Biology and Ecology*, *366*(1), 99–103. https://doi.org/10.1016/j.jembe.2008.07.013

Freyria, N. J., Joli, N., & Lovejoy, C. (2021). A decadal perspective on north water microbial eukaryotes as Arctic Ocean sentinels. *Scientific Reports*, *11*(1), 8413. https://doi.org/10.1038/s41598-021-87906-4

Gardner, A. S., Moholdt, G., Wouters, B., Wolken, G. J., Burgess, D. O., Sharp, M. J., Cogley, J. G., Braun, C., & Labine, C. (2011). Sharply increased mass loss from glaciers and ice caps in the Canadian Arctic Archipelago. *Nature; London*, *473*(7347), 357–360. http://dx.doi.org.ezproxy.library.dal.ca/10.1038/nature10089

Gurdeep Singh, R., Tanca, A., Palomba, A., Van der Jeugt, F., Verschaffelt, P., Uzzau, S., Martens, L., Dawyndt, P., & Mesuere, B. (2019). Unipept 4.0: Functional Analysis of Metaproteome Data. *Journal of Proteome Research*, *18*(2), 606–615. https://doi.org/10.1021/acs.jproteome.8b00716

Halbach, L., Vihtakari, M., Duarte, P., Everett, A., Granskog, M. A., Hop, H., Kauko, H. M., Kristiansen, S., Myhre, P. I., Pavlov, A. K., Pramanik, A., Tatarek, A., Torsvik, T., Wiktor, J. M., Wold, A., Wulff, A., Steen, H., & Assmy, P. (2019). Tidewater Glaciers and Bedrock Characteristics Control the Phytoplankton Growth Environment in a Fjord in the Arctic. *Frontiers in Marine Science*, *6*. https://doi.org/10.3389/fmars.2019.00254

Harrison, W. G., Li, W. K. W., Smith, J. C., Head, E. J. H., & Longhurst, A. R. (1987). Depth profiles of plankton, particulate organic matter and microbial activity in the eastern Canadian Arctic during summer. *Polar Biology*, *7*(4), 207–224. https://doi.org/10.1007/BF00287417

Hartman, F. C., & Harpel, M. R. (1994). Structure, Function, Regulation, and Assembly of D-Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase. *Annual Review of Biochemistry*, *63*(1), 197–232. https://doi.org/10.1146/annurev.bi.63.070194.001213

Hawkings, J. R., Wadham, J. L., Benning, L. G., Hendry, K. R., Tranter, M., Tedstone, A., Nienow, P., & Raiswell, R. (2017). Ice sheets as a missing source of silica to the polar oceans. *Nature Communications*, *8*(1), Article 1. https://doi.org/10.1038/ncomms14198

Hawkings, J., Wadham, J., Tranter, M., Telling, J., Bagshaw, E., Beaton, A., Simmons, S.-L., Chandler, D., Tedstone, A., & Nienow, P. (2016). The Greenland Ice Sheet as a hot spot of phosphorus weathering and export in the Arctic. *Global Biogeochemical Cycles*, *30*(2), 191–210. https://doi.org/10.1002/2015GB005237

Hopwood, M. J., Carroll, D., Browning, T. J., Meire, L., Mortensen, J., Krisch, S., & Achterberg, E. P. (2018). Non-linear response of summertime marine productivity to increased meltwater discharge around Greenland. *Nature Communications*, *9*(1), 3256. https://doi.org/10.1038/s41467-018-05488-8

Jansson, M., Blomqvist, P., Jonsson, A., & BergstrÖm, A.-K. (1996). Nutrient limitation of bacterioplankton, autotrophic and mixotrophic phytoplankton, and heterotrophic nanoflagellates in Lake Örträsket. *Limnology and Oceanography*, *41*(7), 1552–1559. https://doi.org/10.4319/lo.1996.41.7.1552

John, D. E., Patterson, S. S., & Paul, J. H. (2007). Phytoplankton-Group Specific Quantitative Polymerase Chain Reaction Assays for RuBisCO mRNA Transcripts in Seawater. *Marine Biotechnology*, *9*(6), 747–759. https://doi.org/10.1007/s10126-007-9027-z

Joli, N., Gosselin, M., Ardyna, M., Babin, M., Onda, D. F., Tremblay, J.-É., & Lovejoy, C. (2018). Need for focus on microbial species following ice melt and changing freshwater regimes in a Janus Arctic Gateway. *Scientific Reports*, *8*(1), Article 1. https://doi.org/10.1038/s41598-018-27705-6

Joli, N., Monier, A., Logares, R., & Lovejoy, C. (2017). Seasonal patterns in Arctic prasinophytes and inferred ecology of Bathycoccus unveiled in an Arctic winter metagenome. *The ISME Journal*, *11*(6), 1372–1385. https://doi.org/10.1038/ismej.2017.7

Kalenitchenko, D., Joli, N., Potvin, M., Tremblay, J.-É., & Lovejoy, C. (2019). Biodiversity and Species Change in the Arctic Ocean: A View Through the Lens of Nares Strait. *Frontiers in Marine Science; Lausanne*. http://dx.doi.org/10.3389/fmars.2019.00479

Kanna, N., Sugiyama, S., Ohashi, Y., Sakakibara, D., Fukamachi, Y., & Nomura, D. (2018). Upwelling of Macronutrients and Dissolved Inorganic Carbon by a Subglacial Freshwater Driven Plume in Bowdoin Fjord, Northwestern Greenland. *Journal of Geophysical Research: Biogeosciences*, *123*(5), 1666–1682. https://doi.org/10.1029/2017JG004248

King, M. D., Howat, I. M., Candela, S. G., Noh, M. J., Jeong, S., Noël, B. P. Y., van den Broeke, M. R., Wouters, B., & Negrete, A. (2020). Dynamic ice loss from the Greenland Ice Sheet driven by sustained glacier retreat. *Communications Earth & Environment*, *1*(1), Article 1. https://doi.org/10.1038/s43247-020-0001-2

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, *41*(1), e1. https://doi.org/10.1093/nar/gks808

Lancelot, C., Wassmann, P., & Barth, H. (1994). Ecology of Phaeocystis-dominated ecosystems. *Journal of Marine Systems*, *5*(1), 1–4. https://doi.org/10.1016/0924-7963(94)90012-4

Lorimer, G. H., & Andrews, T. J. (1973). Plant Photorespiration—An Inevitable Consequence of the Existence of Atmospheric Oxygen. *Nature*, *243*(5406), Article 5406. https://doi.org/10.1038/243359a0

Losh, J. L., Young, J. N., & Morel, F. M. M. (2013). Rubisco is a small fraction of total protein in marine phytoplankton. *New Phytologist*, *198*(1), 52–58. https://doi.org/10.1111/nph.12143

Lovejoy, C. (2014). *Changing Views of Arctic Protists (Marine Microbial Eukaryotes) in a Changing Arctic*. 10.

Lovejoy, C., Legendre, L., Martineau, M.-J., Bâcle, J., & von Quillfeldt, C. H. (2002). Distribution of phytoplankton and other protists in the North Water. *Deep Sea Research Part II: Topical Studies in Oceanography*, *49*(22), 5027–5047. https://doi.org/10.1016/S0967-0645(02)00176-5

MacIntyre, H. L., & Geider, R. J. (1996). Regulation of Rubisco activity and its potential effect on photosynthesis during mixing in a turbid estuary. *Marine Ecology Progress Series*, *144*(1/3), 247–264.

McCain, J. S. P., Allen, A. E., & Bertrand, E. M. (2022). Proteomic traits vary across taxa in a coastal Antarctic phytoplankton bloom. *The ISME Journal*, *16*(2), Article 2. https://doi.org/10.1038/s41396-021-01084-9

Meire, L., Meire, P., Struyf, E., Krawczyk, D. W., Arendt, K. E., Yde, J. C., Pedersen, T. J., Hopwood, M. J., Rysgaard, S., & Meysman, F. J. R. (2016). High export of dissolved silica from the Greenland Ice Sheet. *Geophysical Research Letters*, *43*(17), 9173–9182. https://doi.org/10.1002/2016GL070191

Meire, L., Mortensen, J., Rysgaard, S., Bendtsen, J., Boone, W., Meire, P., & Meysman, F. J. R. (2016). Spring bloom dynamics in a subarctic fjord influenced by tidewater outlet glaciers (Godthåbsfjord, SW Greenland). *Journal of Geophysical Research: Biogeosciences*, *121*(6), 1581–1592. https://doi.org/10.1002/2015JG003240

Monier, A., Comte, J., Babin, M., Forest, A., Matsuoka, A., & Lovejoy, C. (2015). Oceanographic structure drives the assembly processes of microbial eukaryotic communities. *The ISME Journal*, *9*(4), Article 4. https://doi.org/10.1038/ismej.2014.197

Morlighem, M., Williams, C. N., Rignot, E., An, L., Arndt, J. E., Bamber, J. L., Catania, G., Chauché, N., Dowdeswell, J. A., Dorschel, B., Fenty, I., Hogan, K., Howat, I., Hubbard, A., Jakobsson, M., Jordan, T. M., Kjeldsen, K. K., Millan, R., Mayer, L., … Zinglersen, K. B. (2017). BedMachine v3: Complete Bed Topography and Ocean Bathymetry Mapping of Greenland From Multibeam Echo Sounding Combined With Mass Conservation. *Geophysical Research Letters*, *44*(21), 11,051-11,061. https://doi.org/10.1002/2017GL074954

Murray, C., Markager, S., Stedmon, C. A., Juul-Pedersen, T., Sejr, M. K., & Bruhn, A. (2015). The influence of glacial melt water on bio-optical properties in two contrasting Greenlandic fjords. *Estuarine, Coastal and Shelf Science*, *163*, 72–83. https://doi.org/10.1016/j.ecss.2015.05.041

Needham, D. M., & Fuhrman, J. A. (2016). Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. *Nature Microbiology*, *1*(4), Article 4. https://doi.org/10.1038/nmicrobiol.2016.5

Noël, B., van de Berg, W. J., Lhermitte, S., Wouters, B., Schaffer, N., & van den Broeke, M. R. (2018). Six Decades of Glacial Mass Loss in the Canadian Arctic Archipelago. *Journal of Geophysical Research: Earth Surface*, *123*(6), 1430–1449. https://doi.org/10.1029/2017JF004304

Orellana, M. nica V., & Hansell, D. A. (2012). Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO): A long-lived protein in the deep ocean. *Limnology and Oceanography*, *57*(3), 826–834. https://doi.org/10.4319/lo.2012.57.3.0826

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Pujari, L., Wu, C., Kan, J., Li, N., Wang, X., Zhang, G., Shang, X., Wang, M., Zhou, C., & Sun, J. (2019). Diversity and Spatial Distribution of Chromophytic Phytoplankton in the Bay of Bengal Revealed by RuBisCO Genes (rbcL). *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.01501

Quince, C., Lanzen, A., Davenport, R. J., & Turnbaugh, P. J. (2011). Removing Noise From Pyrosequenced Amplicons. *BMC Bioinformatics*, *12*(1), 38. https://doi.org/10.1186/1471-2105-12-38

Randelhoff, A., Holding, J., Janout, M., Sejr, M. K., Babin, M., Tremblay, J.-É., & Alkire, M. B. (2020). Pan-Arctic Ocean Primary Production Constrained by Turbulent Nitrate Fluxes. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00150

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, *3*(1), Article 1. https://doi.org/10.1038/s43247-022-00498-3

Slater, D. A., Carroll, D., Oliver, H., Hopwood, M. J., Straneo, F., Wood, M., Willis, J. K., & Morlighem, M. (2022). Characteristic Depths, Fluxes, and Timescales for Greenland's Tidewater Glacier Fjords From Subglacial Discharge-Driven Upwelling During Summer. *Geophysical Research Letters*, *49*(10), e2021GL097081. https://doi.org/10.1029/2021GL097081

Smetacek, V. (1999). Diatoms and the Ocean Carbon Cycle. *Protist*, *150*(1), 25–32. https://doi.org/10.1016/S1434-4610(99)70006-4

Spreitzer, R. J. (2003). Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Archives of Biochemistry and Biophysics*, *414*(2), 141–149. https://doi.org/10.1016/S0003-9861(03)00171-1

Tabita, F. R., Hanson, T. E., Satagopan, S., Witte, B. H., & Kreel, N. E. (2008). Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1504), 2629–2640. https://doi.org/10.1098/rstb.2008.0023

Tremblay, J.-É., Anderson, L. G., Matrai, P., Coupel, P., Bélanger, S., Michel, C., & Reigstad, M. (2015). Global and regional drivers of nutrient supply, primary production and CO2 drawdown in the changing Arctic Ocean. *Progress in Oceanography*, *139*, 171–196. https://doi.org/10.1016/j.pocean.2015.08.009

Tremblay, J.-É., & Gagnon, J. (2009). The effects of irradiance and nutrient supply on the productivity of Arctic waters: A perspective on climate change. In J. C. J. Nihoul & A. G. Kostianoy (Eds.), *Influence of Climate Change on the Changing Arctic and Sub-Arctic Conditions* (pp. 73–93). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9460-6_7

Tremblay, J.-É., Robert, D., Varela, D. E., Lovejoy, C., Darnis, G., Nelson, R. J., & Sastri, A. R. (2012). Current state and trends in Canadian Arctic marine ecosystems: I. Primary production. *Climatic Change*, *115*(1), 161–178. https://doi.org/10.1007/s10584-012-0496-3

Van Wychen, W., Burgess, D. O., Gray, L., Copland, L., Sharp, M., Dowdeswell, J. A., & Benham, T. J. (2014). Glacier velocities and dynamic ice discharge from the Queen Elizabeth Islands, Nunavut, Canada. *Geophysical Research Letters*, *41*(2), 484–490. https://doi.org/10.1002/2013GL058558

Varela, D. E., Crawford, D. W., Wrohan, I. A., Wyatt, S. N., & Carmack, E. C. (2013). Pelagic primary productivity and upper ocean nutrient dynamics across Subarctic and Arctic Seas. *Journal of Geophysical Research: Oceans*, *118*(12), 7132–7152. https://doi.org/10.1002/2013JC009211

Verity, P. G., Brussaard, C. P., Nejstgaard, J. C., van Leeuwe, M. A., Lancelot, C., & Medlin, L. K. (2007). Current understanding of Phaeocystis ecology and biogeochemistry, and perspectives for future research. *Biogeochemistry*, *83*(1), 311–330. https://doi.org/10.1007/s10533-007-9090-6

Walker, J. M. (Ed.). (2005). *The Proteomics Protocols Handbook*. Humana Press. https://doi.org/10.1385/1592598900

Williams, P. L., Burgess, D. O., Waterman, S., Roberts, M., Bertrand, E. M., & Bhatia, M. P. (2021). Nutrient and Carbon Export From a Tidewater Glacier to the Coastal Ocean in the Canadian Arctic Archipelago. *Journal of Geophysical Research: Biogeosciences*, *126*(9), e2021JG006289. https://doi.org/10.1029/2021JG006289

Young, J. N., Goldman, J. A. L., Kranz, S. A., Tortell, P. D., & Morel, F. M. M. (2015). Slow carboxylation of Rubisco constrains the rate of carbon fixation during Antarctic phytoplankton blooms. *New Phytologist*, *205*(1), 172–181. https://doi.org/10.1111/nph.13021

Young, J. N., Heureux, A. M. C., Sharwood, R. E., Rickaby, R. E. M., Morel, F. M. M., & Whitney, S. M. (2016). Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *Journal of Experimental Botany*, *67*(11), 3445–3456. https://doi.org/10.1093/jxb/erw163

Young, J. N., Kranz, S. A., Goldman, J. A. L., Tortell, P. D., & Morel, F. M. M. (2015). Antarctic phytoplankton down-regulate their carbon-concentrating mechanisms under high CO2 with no change in growth rates. *Marine Ecology Progress Series*, *532*, 13–28. https://doi.org/10.3354/meps11336

Appendix A: Supporting Information

Table S1. Summary of stations and depths sampled for protein on *S/Y Vagabond* and *CCGS Amundsen* cruises during July and August 2019.

Cruise	Station #	Sampling date	Latitude	Longitude	Depths (m)
Amundsen	2.5	2019-08-06	76.358	-77.506	2, 35, 55, 150
Amundsen	290	2019-08-06	76.132	-80.357	2, 29, 38, 140
Amundsen	291	2019-08-06	75.996	-80.378	2, 15, 40, 125
Amundsen	292	2019-08-06	75.882	-80.416	2, 22, 31, 150
Amundsen	293	2019-08-07	75.733	-80.678	2, 7, 20, 120
Amundsen	2.3	2019-08-07	76.131	-83.020	2, 15, 32, 140
Amundsen	2.4	2019-08-08	76.127	-86.317	2, 26, 35, 140
Amundsen	1.4	2019-08-08	76.503	-84.934	2, 5, 35, 100
Amundsen	295	2019-08-08	76.377	-84.408	2, 35, 54, 87
Amundsen	297	2019-08-09	76.373	-81.299	2, 18, 33, 140
Amundsen	296	2019-08-10	75.524	-79.749	2, 26, 40, 140
Amundsen	2.7	2019-08-10	75.482	-78.635	2, 33, 50, 150
Vagabond	6	2019-07-31	76.40443	-81.1576	40
Vagabond	12	2019-08-02	75.67577	-81.1484	20
Vagabond	30	2019-08-06	75.75667	-83.24993	10,110
Vagabond	32	2019-08-07	75.8268	-82.97588	20
Vagabond	35	2019-08-09	75.71782	-84.51685	33
Vagabond	40	2019-08-14	76.62732	-85.11707	13
Vagabond	44	2019-08-15	76.57053	-83.18493	31

Table S2. Selected reaction monitoring mass spectrometry parameters for peptides measured in this study. *1 has the greatest signal and 5 has the lowest.

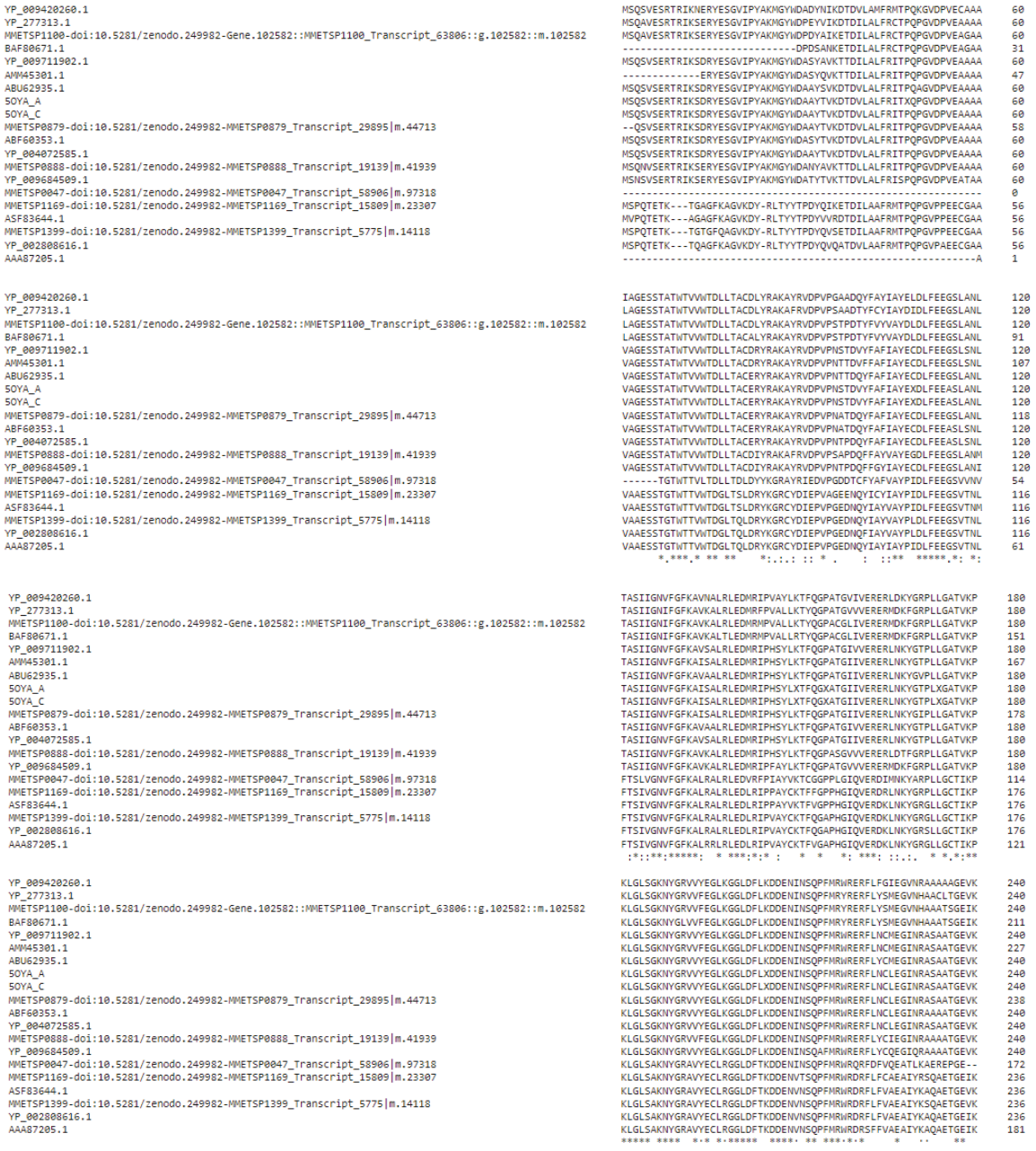
Peptide name	Peptide (charge state)	Precursor (m/z)	Product (m/z)	Fragment ion	Collision Energy (V)	SRM rank *	RT (min)
Broad Euk Rubisco Large subunit	YESGVIPYAK(+2)	563.79	293.11	b2	16.9	3	18.3
			478.27	y4	24.8	2	
			591.35	y5	16.5	4	
			747.44	y7	19.3	5	
			834.47	y8	18.2	1	
	YESGVI(13C15N)PYAK(+2)	567.30	293.11	b2	16.9	3	
			478.27	y4	24.8	2	
			598.37	y5	16.5	4	
			754.46	y7	19.3	5	
			841.49	y8	18.2	1	
Frag/Pseudo Rubisco, Large subunit	FLNC[+57.0]MEGINR(+2)	627.29	993.42	y4	21.9	2	20.6
			879.38	y5	22.4	3	
			719.35	y6	23.2	4	
			588.31	y7	23	5	
			459.27	y8	24.5	1	
	FLNC[+57.0]MEGI(*13C15N*)NR(+2)	630.80	1000.44	y4	21.9	2	
			886.40	y5	22.4	3	
			726.37	y6	23.2	4	
			595.33	y7	23	5	
			466.28	y8	24.5	1	
Frag/Pseudo Rubisco, Large subunit, Oxidized	FLNC[+57.0]M[+16.0]EGINR(+2)	635.29	1009.42	y4	21.9	1	16.6
			895.38	y5	22.4	3	
			735.35	y6	23.2	5	
			588.31	y7	23	4	
			459.27	y8	24.5	2	
	FLNC[+57.0]M[+16.0]EGI(*13C15N*)NR(+2)	630.80	1016.44	y4	21.9	1	
			902.39	y5	22.4	3	
			742.36	y6	23.2	5	
			595.33	y7	23	4	
			466.28	y8	24.5	2	
Chaetoceros Rubisco, Large subunit	FLYC[+57.0]MEGINR(+2)	651.80	1042.44	y8	21.6	1	22.6
			879.381	y7	23	2	
			719.35	y6	23.8	3	
			588.31	y5	22.9	5	
			459.27	y4	25.5	4	
	FLYC[+57.0]MEGI*(13C15N)*NR (heavy)(+2)	655.31	1049.46	y8	21.6	1	
			886.40	y7	23	2	
			726.36	y6	23.8	3	
			595.3272	y5	22.9	5	
			466.2846	y4	25.5	4	
Chaetoceros Rubisco, Large subunit, oxidized	FLYC[+57.0]M[+16.0]EGINR(+2)	659.80	1058.439	y8	21.6	1	18.2
			895.376	y7	23	3	
			735.3454	y6	23.8	4	
			588.31	y5	22.9	5	
			459.2674	y4	25.5	2	

	Peptide (charge state)	Precursor (m/z)	Product (m/z)	Fragment ion	Collision Energy (V)	SRM rank *	
	FLYC[+57.0]M[+16.0]EGI*(13C15N)*NR (heavy)(+2)	663.31	1065.457	y8	21.6	1	
			902.3932	y7	23	3	
			742.3626	y6	23.8	4	
			595.3272	y5	22.9	5	
			466.2846	y4	25.5	2	
Thalassiosira Rubisco, Large subunit	FLNC[+57.0]LEGINR(+2)	618.31	975.4676	y8	21.4	3	21.7
			861.4247	y7	22.9	5	
			701.3941	y6	22.7	4	
			588.31	y5	22.4	2	
			459.2674	y4	23.8	1	
	FLNC[+57.0]LEGI*(13C15N)*NR (heavy)(+2)	621.82	982.4848	y4	21.4	3	
			868.4419	y5	22.9	5	
			708.4112	y6	22.7	4	
			595.3272	y7	22.4	2	
			466.2846	y8	23.8	1	
Micromonas Rubisco, Large subunit	NVTLGFVDLMR(+2)	632.84	1051.56	y9	19.7	2	29.6
			950.5128	y8	20.9	3	
			837.4287	y7	20	1	
			780.4073	y6	19.6	4	
	NVTLGFVDL*(13C15N)*MR (heavy)(+2)	636.35	1058.578	y9	19.7	2	
			957.53	y8	20.9	3	
			844.4459	y7	20	1	
			787.4244	y6	19.6	4	
Micromonas Rubisco, Large subunit, oxidized	NVTLGFVDLM[+16.0]R(+2)	640.84	1067.555	y9	19.7	2	25.9
			966.5077	y8	20.9	4	
			853.4237	y7	20	1	
			796.4022	y6	19.6	3	
	NVTLGFVDL*(13C15N)*M[+16.0]R (heavy)(+2)	644.35	1074.573	y9	19.7	2	
			973.5249	y8	20.9	4	
			860.4408	y7	20	1	
			803.4194	y6	19.6	3	
Phaeocystis Rubisco, Large subunit	DYVAEGPQILR(+2)	630.83	982.568	y9	21.4	5	21.4
			883.4996	y8	20.5	2	
			812.4625	y7	19.8	4	
			683.4199	y6	20.3	1	
			378.166	b3	18	3	
	DYVAEGPQI*(13C15N)*LR (heavy)(+2)	634.34	989.5852	y9	21.4	5	
			890.5168	y8	20.5	2	
			819.4796	y7	19.8	4	
			690.437	y6	20.3	1	
			378.166	b3	18	3	

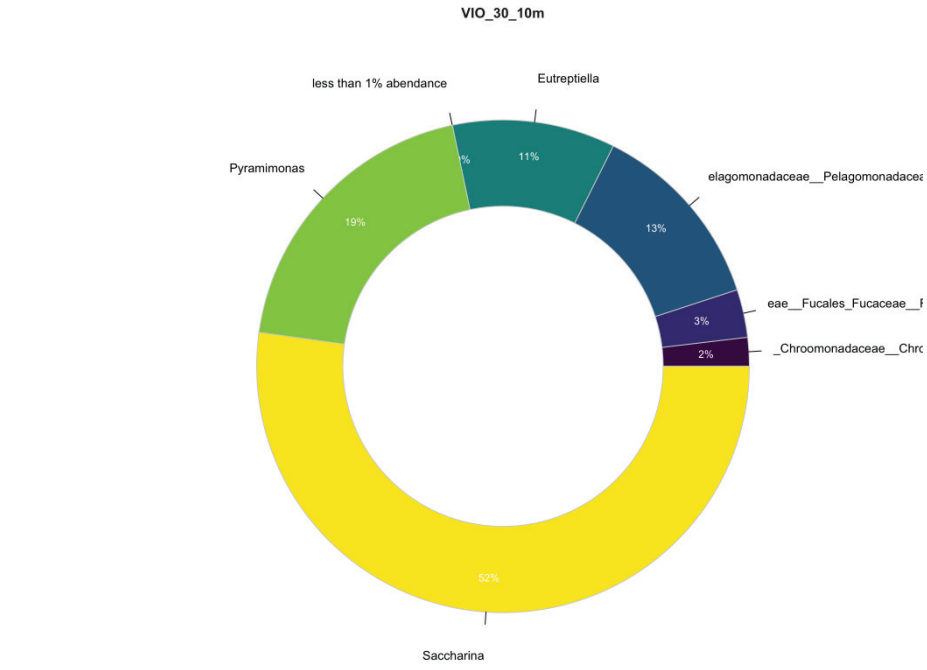
Table S3. RbcL sequence accession numbers and associated organisms for reference to multiple alignment in Figure S2.

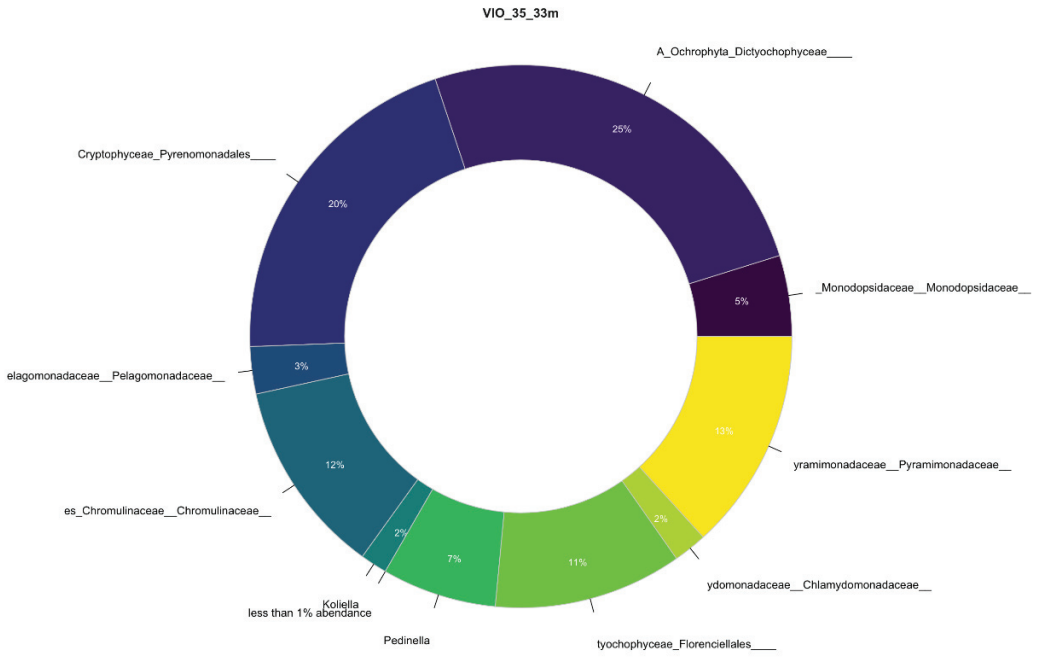
Accession #	Organism
MMETSP1100	Phaeocystis antarctica
ABU62935.1	Chaetoceros neogracilis
YP_009711902.1	Fragilariopsis cylindrus
AMM45301.1	Pseudo-nitzschia arctica
ABF60353.1	Thalassiosira antarctica
>YP_002808616.1	Micromonas commoda
>YP_009684509.1	Florenciella parvula
>YP_277313.1	Emiliania huxleyi
>ASF83644.1	Chlamydomonas reinhardtii
>AAA87205.1	Micromonas pusilla
>MMETSP0888	Pelagomonas calceolata
>MMETSP1399	Bathycoccus prasinos
>YP_009684509.1	Florenciella parvula
>BAF80671.1	Phaeocystis pouchetii
>MMETSP0047	Chroomonas cf. mesostigmatica
>5OYA_A	Chaetoceros socialis: chain A
>5OYA_C	Chaetoceros socialis: chain C
>YP_004072585.1	Thalassiosira oceanica CCMP1005
>MMETSP0879	Thalassiosira weissflogii
>MMETSP1169	Pyramimonas obovata
>MMETSP0929	Ostreococcus mediterraneus
>MMETSP1471	Pycnococcus provasolii
>YP_009420260.1	Cryptomonas curvata

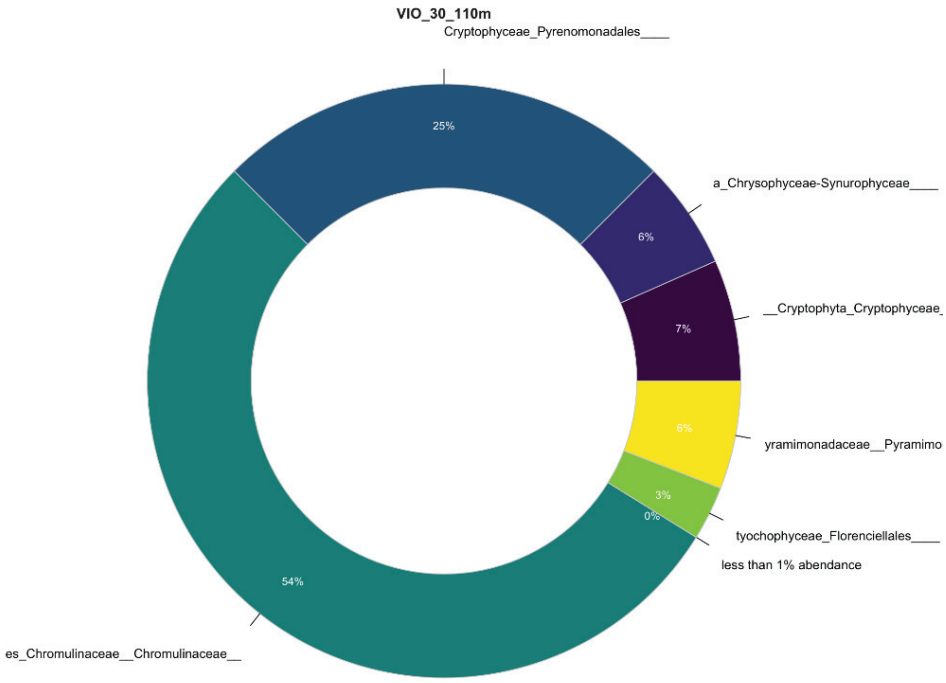
Figure S1. CLUSTAL multiple sequence alignment of RbcL sequences. See Table S3 for taxon key

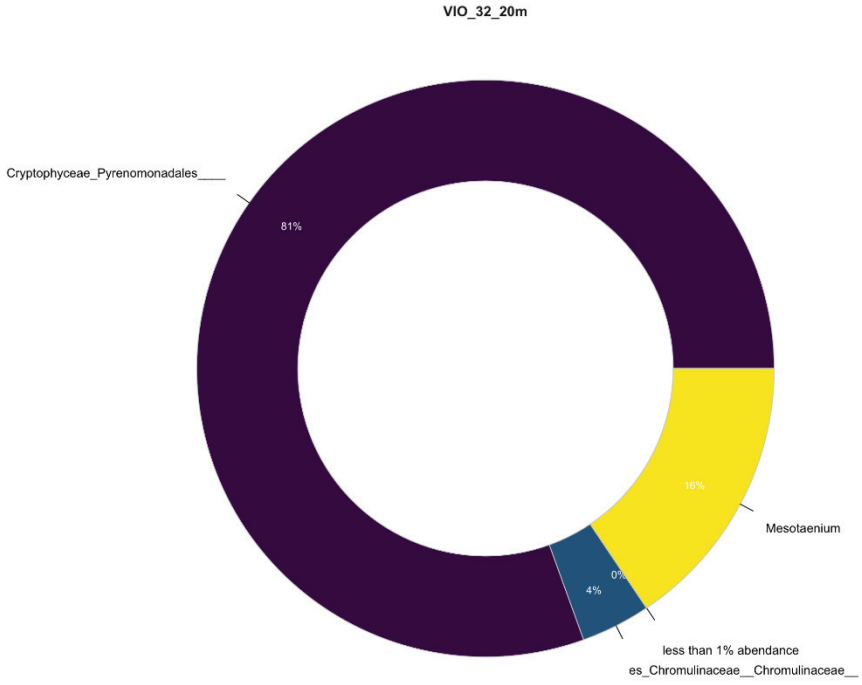


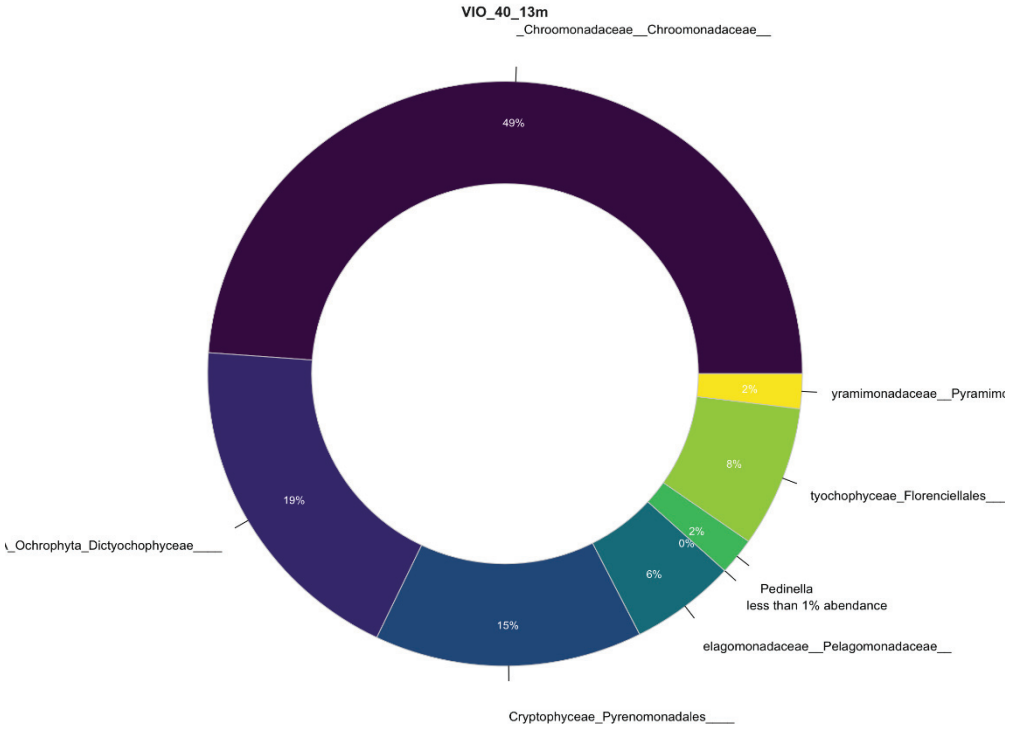
```
YP_009420260.1                                                                GHYFNVTAGTMEDMYERAEFCREIGSVICMIDLV-IGYTAIQSMGIWARQNSMILHLHRA    299
YP_277313.1                                                                   GHYLNTTAATMEDMYERANFARDLGSVIVMHIDLV-IGYTAIQSMGKWNSRDNDVILHLHRA    299
MMETSP1100-doi:10.5281/zenodo.249982-Gene.102582::MMETSP1100_Transcript_63806::g.102582::m.102582  GHYLNSTAATMEDMYERAEFAVELGSVIVMHIDLV-IGYTAIQSMAKWNCRKVDCILHLHRA    299
BAF80671.1                                                                    GHYLNSTAATMEDMYERAEFAVELGSVIVMHIDLV-IGYTAIQYMAKWNCRKVDCILHLHRA    270
YP_009711902.1                                                                GSYLNVTAATWEEVLKRCEYAKQVGSIVVMHIDLV-MGYTAIQSAAIWARENDMLLHLHRA    299
AMM45301.1                                                                    GSYLNVTAATMEEVLKRCEYAKEVGSIVVMHIDLV-MGYTAIQTTALWARENDMLLHLHRA    286
ABU62935.1                                                                    GSYLNVTAATMEEVYKRAEYAKQVGSVIIMIDLV-MGYTAIQSIALWSRDNDMLLHLHRA    299
5OYA_A                                                                        GSYLNVTAATMEEVYKRAEYAKAIGSIVVMIDLV-MGYTAIQSIAYWARENDMLLHLHRA    299
5OYA_C                                                                        GSYLNVTAATMEEVYKRAEYAKAIGSIVVMIDLV-MGYTAIQSIAYWARENDMLLHLHRA    299
MMETSP0879-doi:10.5281/zenodo.249982-MMETSP0879_Transcript_29895|m.44713      GSYLNVTAATMEEVYKRAEYAKLIGSVIIMIDLV-MGYTAIQSIAYWARENDMLLHLHRA    297
ABF60353.1                                                                    GSYLNITAATMEEVYKRAEYAKAIGSVVVMIDLV-MGYTAIQSIAYWARENDMLLHLHRA    299
YP_004072585.1                                                                GSYLNVTAATMEEVYKRAEYAKQIGSIVIMIDLV-MGYTAIQSIAYWARENDMLLHLHRA    299
MMETSP0888-doi:10.5281/zenodo.249982-MMETSP0888_Transcript_19139|m.41939      GSYLNMTAGNNDEMITRGEFAKELGSVIVMIDLV-IGYTAIQTAAIWARANMHLLHLHRA    299
YP_009684509.1                                                                GSYLNVTAGDMEAVYERCNYAKELGTIVVMIDLV-MGYTAIQSAAKWARKNDMILHLHRA    299
MMETSP0047-doi:10.5281/zenodo.249982-MMETSP0047_Transcript_58906|m.97318      ----------------------------------------------------------    172
MMETSP1169-doi:10.5281/zenodo.249982-MMETSP1169_Transcript_15809|m.23307      GHYLNVTAATTEEVIKRAEFARDLGVPIVMHDYITGGFTTNTSLAAYCRDNGLLLHIHRA    296
ASF83644.1                                                                    GHYLNATAGTCEEMMKRAVCAKELGVPIIMHDYLTGGFTANTSLAIYCRDNGLLLHIHRA    296
MMETSP1399-doi:10.5281/zenodo.249982-MMETSP1399_Transcript_5775|m.14118       GHYLNATAGNVDEMIKRAVCAKELGMPIVMHDYLTAGFTANTTLAHYCRDHGLLLHIHRA    296
YP_002808616.1                                                                GHYLNVTAATSEEMIKRAACAKELGLPIIMHDYITGGFTSNTSLAIYCRDHGLLLHIHRA    296
AAA87205.1                                                                    GHYLNATADTLEEMLKRAVCAKELGLPIVMHDYITGGFTANTSLAIYCRDHGLLLHIHRA    241

YP_009420260.1                                                                GNSTYSRQKTHGMNFRVICKWMRMAGVDHIHAGTVVGKLEGDPLMVKGFYNVLLQAKLDV    359
YP_277313.1                                                                   GNSTYSRQKNHGMNFRVICKWMRMSGCDHIHAGTVVGKLEGDPLMIKGFYNTLLDTKTEV    359
MMETSP1100-doi:10.5281/zenodo.249982-Gene.102582::MMETSP1100_Transcript_63806::g.102582::m.102582  GNSTYSRQKNHGMNFRVICKWMRMAGVDHIHAGTVVGKLEGDPLMIKGFYNTLLDFKSDI    359
BAF80671.1                                                                    GNSTYSRQKNHGMNFRVICKWMRMAGVDHIHAGTVVGKLEGDPLMIKGFYNTLLDFKSDI    330
YP_009711902.1                                                                GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIKGFYDVLRLTTLDV    359
AMM45301.1                                                                    GNSTYARQKNHGINFRVICKWMRMCGVDHIHAGTVVGKLEGDPLMIKGFYDVLRLTHLDV    346
ABU62935.1                                                                    GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIKGFYDVLRLTNLEV    359
5OYA_A                                                                        GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIXGFYDILRLTELEV    359
5OYA_C                                                                        GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIXGFYDILRLTELEV    359
MMETSP0879-doi:10.5281/zenodo.249982-MMETSP0879_Transcript_29895|m.44713      GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIXGFYDILRLTELEV    357
ABF60353.1                                                                    GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIKGFYDVLRLTELEV    359
YP_004072585.1                                                                GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTVVGKLEGDPLMIKGFYDILRLTELEV    359
MMETSP0888-doi:10.5281/zenodo.249982-MMETSP0888_Transcript_19139|m.41939      GNSTYARQKNHGINFRVICKWMRMAGVDHIHAGTVVGKLEGDPLMIQGFYDTLLKTKLSI    359
YP_009684509.1                                                                GNSTYARQKNHGINFRVICKWMRMSGVDHIHAGTIVGKLEGDPLMVQGFYETLLCVKSEV    359
MMETSP0047-doi:10.5281/zenodo.249982-MMETSP0047_Transcript_58906|m.97318      ----------------------------------------------------------    172
MMETSP1169-doi:10.5281/zenodo.249982-MMETSP1169_Transcript_15809|m.23307      MHAVIDRQRNHGIHFRVLAKALRMSGGDHLHSGTVVGKLEGEREVTLGFVDLMRDAYIEK    356
ASF83644.1                                                                    MHAVIDRQRNHGIHFRVLAKALRMSGGDHLHSGTVVGKLEGEREVTLGFVDLMRDDYVEK    356
MMETSP1399-doi:10.5281/zenodo.249982-MMETSP1399_Transcript_5775|m.14118       MHAVIDRQRNHGIHFRVLAKALRMSGGDHLHSGTVVGKLEGERNVTLGFVDLMRDAYVEK    356
YP_002808616.1                                                                MHAVIDRQRNHGIHFRVLAKALRMSGGDHLHSGTVVGKLEGERNVTLGFVDLMRDPIVEK    356
AAA87205.1                                                                    MHAVIDRQRNHGIHFRVLAKALRMSGGDHLHSGTVVGKLEGERNVTLGFVDLMRDPIVEK    301

YP_009420260.1                                                                NLPQGLFFAQDWGSLAKCLPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHPDGIQAG    419
YP_277313.1                                                                   NLPQGLFFAQDWASLRKCVPVASGGIHCGQMHQLINYLGDDVVLQFGGGTIGHPDGIQAG    419
MMETSP1100-doi:10.5281/zenodo.249982-Gene.102582::MMETSP1100_Transcript_63806::g.102582::m.102582  NLPCGMFFAQDWASLRKCVPVASGGIHCGQMHQLIDYLGDDVILQFGGGTIGHPDGIQAG    419
BAF80671.1                                                                    NLPCGMFFAQDWASLRKCVPVASGGIHCGQMHQLIDYLGDDVVLQFGGGTIGHPDGIQAG    390
YP_009711902.1                                                                NLPYGIFFEMSWASLRRCMPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHPDGIQAG    419
AMM45301.1                                                                    NLPFGVFFEMPWASLRRCMPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHPDGIQAG    406
ABU62935.1                                                                    NLPFGIFFEMGWASLRKCMPVASGGIHCGQMHQLIYYLGDDVVLQFGGGTIGHPDGIQAG    419
5OYA_A                                                                        NLPFGIFFEMDWASLRRCMPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHPDGIQAG    419
5OYA_C                                                                        NLPFGIFFEMDWASLRRCMPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHPDGIQAG    419
MMETSP0879-doi:10.5281/zenodo.249982-MMETSP0879_Transcript_29895|m.44713      NLPYGIFFEMDWASLRRCMPVASGGIHCGQMHQLVHYLGDDVVLQFGGGTIGHPDGIQAG    417
ABF60353.1                                                                    NLPFGIFFEMDWASLRRCMPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHPDGIQAG    419
YP_004072585.1                                                                NLPFGIFFEMDWASLRRCMPVASGGIHCGQMHQLIHYLGDDVVLQFGGGTIGHSDGIQAG    419
MMETSP0888-doi:10.5281/zenodo.249982-MMETSP0888_Transcript_19139|m.41939      DLPKGIFFDMDWASLRKCLPVASGGIHCGQMHQLLNYLGDDVVLQFGGGTIGHPDGIQAG    419
YP_009684509.1                                                                NLPKGLFFAQDWASLRKCLPVASGGIHCGQMHQLLNYLGDDVVLQFGGGTIGHPDGIQAG    419
MMETSP0047-doi:10.5281/zenodo.249982-MMETSP0047_Transcript_58906|m.97318      ----------------------------------------------------------    172
MMETSP1169-doi:10.5281/zenodo.249982-MMETSP1169_Transcript_15809|m.23307      DRSRGVYFTQDWCGLAGVMPVASGGIHVWHMPALVEIFGDDSCLQFGGGTLGHPWGNAPG    416
ASF83644.1                                                                    DRSRGIYFTQDWCSMPGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGTLGHPWGNAPG    416
MMETSP1399-doi:10.5281/zenodo.249982-MMETSP1399_Transcript_5775|m.14118       DRDRGVYFTQDWASLPGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGTLGHPWGNAPG    416
YP_002808616.1                                                                DRDRGVYFTQDWGSLPGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGTLGHPWGNAPG    416
AAA87205.1                                                                    DRDRGVYFTQNWGSLPGVMP----------------------------------------    321

YP_009420260.1                                                                ATANRVALECMVVARNEGRDYLA--EGPQILKTAAKSCGPLQTALDLWKDITFNYASTDT    477
YP_277313.1                                                                   ATANRVALECMVLARNEGRDYIA--EGPQILRDAAKTCGPLQTALDLWKDITFNYASTDT    477
MMETSP1100-doi:10.5281/zenodo.249982-Gene.102582::MMETSP1100_Transcript_63806::g.102582::m.102582  ATANRVALECMVLARNEGRDYVA--EGPQILRDAAKTCGP--------------------    457
BAF80671.1                                                                    ATANRVALECMVLARNEGRDYVA--EGPQILRDAAKTCGPLQTALDLWKDITFNYASTDT    448
YP_009711902.1                                                                ATANRVALEAMVMARNEGADYFNNEVGPQILRNAAKTCGPLQTALDLWKDISFNYTSTDT    479
AMM45301.1                                                                    ATANRVALEAMVLARNEGADFFSNEVGPQILRNAAKTCGPLQTALDLWKDISFNYTSTDT    466
ABU62935.1                                                                    ATANRVALEAMVMARNEGADYFNPQVGPVILRDAAKSCGPLQTALDLWKD-----------    469
5OYA_A                                                                        ATANRVALEAMVLARNEGADYFNNQVGPQILRDAAKTXGPLQTALDLWKDISFNYTSTDT    479
5OYA_C                                                                        ATANRVALEAMVLARNEGADYFNNQVGPQILRDAAKTXGPLQTALDLWKDISFNYTSTDT    479
MMETSP0879-doi:10.5281/zenodo.249982-MMETSP0879_Transcript_29895|m.44713      ATANRVALEAMVLARNEGQDYFNNQVGPQILRDAAKTCGPLQTALDLWKDISFNYTSTDT    477
ABF60353.1                                                                    ATANRVALEAMVLARNEGADYFNNQVGPQILRDAAKTCGPLQTALDLWKDISFNYTSTDT    479
YP_004072585.1                                                                ATANRVALEAMVLARNEGADYFNTQVGPQILRDAAKTCGPLQTALDLWKDISFNYTSTDT    479
MMETSP0888-doi:10.5281/zenodo.249982-MMETSP0888_Transcript_19139|m.41939      ATANRVALEAMVLARNEGRNYIE--EGPQILRDTAKMCGPLQTALDLWKDISFNYASTDT    477
YP_009684509.1                                                                ATANRVALEAMVMARNEGRDYVA--EGPEILQDAAKMCGPLQTALDLWKGISFNYTSTDT    477
MMETSP0047-doi:10.5281/zenodo.249982-MMETSP0047_Transcript_58906|m.97318      ----------------------------------------------------------    172
MMETSP1169-doi:10.5281/zenodo.249982-MMETSP1169_Transcript_15809|m.23307      AAANRVALEACVQARNEGRDLAR--EGGDVLRAACKWSPELAAACEVWKEIKFEFEAMDT    474
ASF83644.1                                                                    AAANRVALEACTQARNEGRDLAR--EGGDVIRSACKWSPELAAACEVWKEIKFEFDTIDK    474
MMETSP1399-doi:10.5281/zenodo.249982-MMETSP1399_Transcript_5775|m.14118       ASANRVALEACTQARNEGRDLAR--EGGDVIRAACKWSPELAAACEVWKEIKFEFDTVDT    474
YP_002808616.1                                                                ASANRVALEACTQARNEGRDLAR--EGGDVIRAACKWSPELAAACEVWKEIKFEFDTVDT    474
AAA87205.1                                                                    ----------------------------------------------------------    321

YP_009420260.1                                                                ADFIETATSNR------------------------------------------------    488
YP_277313.1                                                                   ADFVETATANV------------------------------------------------    488
MMETSP1100-doi:10.5281/zenodo.249982-Gene.102582::MMETSP1100_Transcript_63806::g.102582::m.102582  ----------------------------------------------------------    457
BAF80671.1                                                                    ADFVETATANR------------------------------------------------    459
YP_009711902.1                                                                ADFAVTPTANV------------------------------------------------    490
AMM45301.1                                                                    ADFAVT-----------------------------------------------------    472
ABU62935.1                                                                    ----------------------------------------------------------    469
5OYA_A                                                                        ADFSETATANR------------------------------------------------    490
5OYA_C                                                                        ADFSETATANR------------------------------------------------    490
MMETSP0879-doi:10.5281/zenodo.249982-MMETSP0879_Transcript_29895|m.44713      ADFAETATANR------------------------------------------------    488
ABF60353.1                                                                    ADFAETATANR------------------------------------------------    490
YP_004072585.1                                                                ADFAETATANR------------------------------------------------    490
MMETSP0888-doi:10.5281/zenodo.249982-MMETSP0888_Transcript_19139|m.41939      ADFADTPTANV------------------------------------------------    488
YP_009684509.1                                                                ADFAETPTANV------------------------------------------------    488
MMETSP0047-doi:10.5281/zenodo.249982-MMETSP0047_Transcript_58906|m.97318      ----------------------------------------------------------    172
MMETSP1169-doi:10.5281/zenodo.249982-MMETSP1169_Transcript_15809|m.23307      LMMETSP-doi-zenodo--MMETSPTranscriptm-MMETSPTranscriptg-ORFM    528
ASF83644.1                                                                    L----------------------------------------------------------    475
MMETSP1399-doi:10.5281/zenodo.249982-MMETSP1399_Transcript_5775|m.14118       L----------------------------------------------------------    475
YP_002808616.1                                                                L----------------------------------------------------------    475
AAA87205.1                                                                    ----------------------------------------------------------    321
```


VIO_30_10m
less than 1% abendance
Eutreptiella
Pyramimonas
elagomonadaceae__Pelagomonadacea
eae__Fucales_Fucaceae__
_Chroomonadaceae__Chro
11%
13%
3%
2%
19%
52%
Saccharina

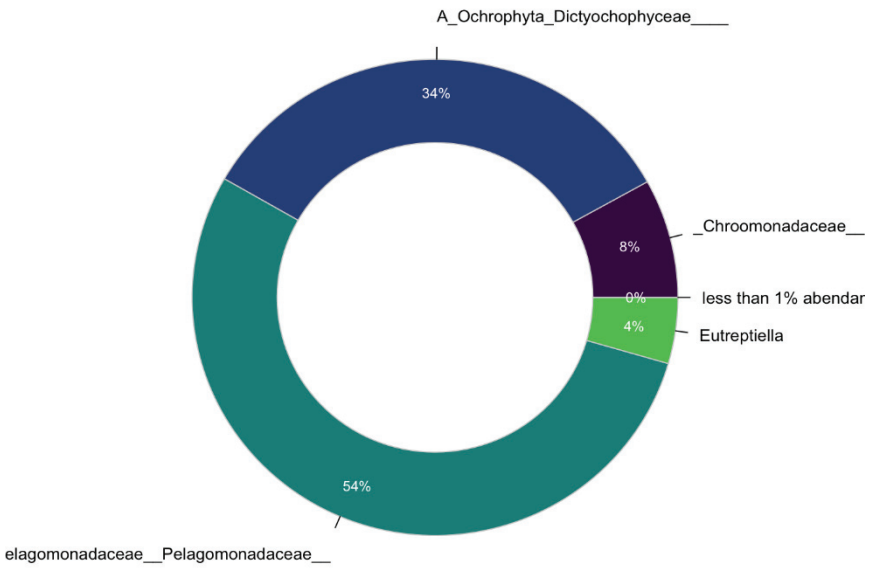

VIO_35_33m
A_Ochrophyta_Dictyochophyceae____
Cryptophyceae_Pyrenomonadales____
elagomonadaceae__Pelagomonadaceae__
es_Chromulinaceae__Chromulinaceae__
Koliella
less than 1% abendance
Pedinella
tyochophyceae_Florenciellales____
ydomonadaceae__Chlamydomonadaceae__
yramimonadaceae__Pyramimonadaceae__
_Monodopsidaceae__Monodopsidaceae__
25%
20%
3%
12%
2%
7%
11%
2%
13%
5%


VIO_30_110m
Cryptophyceae_Pyrenomonadales____
25%
a_Chrysophyceae-Synurophyceae____
6%
__Cryptophyta_Cryptophyceae_
7%
6%
yramimonadaceae__Pyramimo
3%
0%
tyochophyceae_Florenciellales____
less than 1% abendance
54%
es_Chromulinaceae__Chromulinaceae__


VIO_32_20m
Cryptophyceae_Pyrenomonadales____
81%
16%
Mesotaenium
0%
4%
less than 1% abendance
es_Chromulinaceae__Chromulinaceae__


VIO_40_13m
_Chroomonadaceae__Chroomonadaceae__
49%
2%
yramimonadaceae__Pyramim
8%
tyochophyceae_Florenciellales__
2%
0%
Pedinella
less than 1% abendance
6%
elagomonadaceae__Pelagomonadaceae__
15%
Cryptophyceae_Pyrenomonadales____
19%
_Ochrophyta_Dictyochophyceae____

VIO_12_20m


A_Ochrophyta_Dictyochophyceae____
34%
_Chroomonadaceae__
8%
0%
less than 1% abendar
4%
Eutreptiella
54%
elagomonadaceae__Pelagomonadaceae__

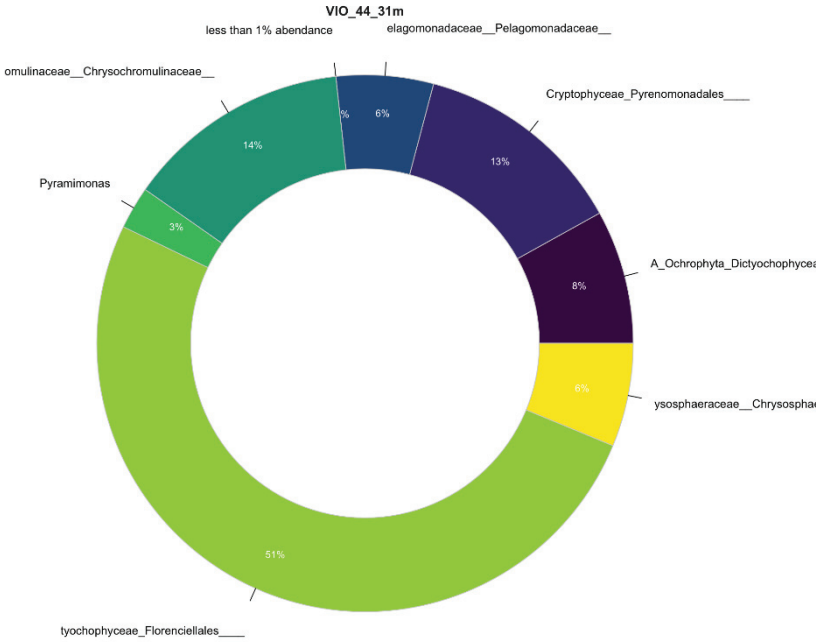


Figure S2. Proportions of phytoplankton groups in 'other' for stations measured by chloroplast 16S rRNA amplicon gene sequencing presented in Figure 4.