UNIVERSITY OF CALIFORNIA SAN DIEGO

Adaptive physiology and structuring of microbial communities in ephemeral Antarctic environments.

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

in

Marine Biology

by

Angela Zoumplis

Committee in charge:

Professor Andrew Allen, Chair
Professor Jeff Bowman
Professor Theresa Gaasterland
Professor Rob Knight
Professor Maria Vernet

2022

The Dissertation of Angela Zoumplis is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego
2022

DEDICATION

For Nowell and Anna Zoumplis

EPIGRAPH

There's been some whiskey drinking. The blue-tinged ice cubes in our glasses—older, we are told, than the very idea of whiskey. It's warm tonight by local standards, which can see temperatures drop to 50 below and beyond. So, as one does in the Dry Valleys of Antarctica, at the bottom of the world, I go to the beach and play Frisbee.

I pick my way across the ice-covered lake, unsteady on my crampons, and flop gratefully down on soft sand, staring up at a midnight sun that never sets. Behind me a few yards away, looming overhead, is the massive, 200-foot-high wall of a glacier. In the other direction, what looks very much like Mars.

Antarctica is the last un-fucked-up place on Earth.

- Anthony Bourdain

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

Many thanks to everyone who made this dissertation possible. It's truly been an amazing experience and an absolute dream to get to work in some incredible places with even more incredible people.

Andy, getting a "yes" from you for an internship at the J. Craig Venter Institute almost 10 years ago was one of the luckiest points of my life. From San Diego to the Antarctic Dry Valleys, thanks for always being an email (or two emails) away. While I still have a lot to learn here, I've gained so much professionally and personally from our chats. You're an extraordinary scientist and always one that I've looked up to. I've always been proud to be a part of the Allen lab and though it's been a steep learning curve- thanks very much for not giving up on me. Your patience, time and resource investments and optimism have been greatly appreciated. I know a lot of hard work went on behind the scenes by you to make things happen for this thesis and will forever be thankful to have had these opportunities. Very grateful to have had the chance to work with you. Thank you for everything.

To my committee, thank you all for the encouragement, support and for sticking with me over the many years of this PhD. Terry Gaasterland, thank you so very much for your guidance in bioinformatics, brainstorming and designing projects with me, and for your motivational emails through the rough parts of research and writing that made a big difference in my outlook for this thesis. To Maria Vernet, thank you for consistently checking up on my progress (which lit a fire to be a bit more efficient with my time) and your creative, thoughtful insights into these projects. The innovative, polar work that comes out of your lab is absolutely inspiring. Jeff Bowman, thank you for the interesting Antarctic biology discussions, troubleshooting and perspectives in navigating through work on these sometimes difficult ecosystems. I am very appreciative that you

shared your time, in between all your ridiculously cool expeditions and projects, to help us out with this dissertation. Rob Knight, I can't tell you how nerve wracked I was to ask you to be on my committee (I'm still a little nervous). Your past work has led the way in microbial ecology and the questions we ask here, and for sure into the future, are a direct result of your inventive and impactful studies. Thanks very much for the feedback and conversations and for your approachability. Best. Committee. Ever.

A special thanks to the people who I share this PhD with (seriously, if you'd like a piece of the degree I will mail it to you); Hong Zheng, Erin Garza, Drishti Kaul and Bethany Kolody, none of this would be possible without you all. Your tireless efforts, endless lab, phone and zoom conversations, and tenacity in working with me is saint-like. In addition, I am so happy to you all friends/buddies and you've made my life in grad school much better than everyone said it would be. To everyone in the Allen lab, the J. Craig Venter Institute and Scripps Institution of Oceanography that accepted a continuous barrage of questions and who have shared their knowledge with me, many, many thanks. To Rachel Diner, Aubrie O'Rourke, Vince Bielinski, Flip McCarthy, Mark Moosburner, Tyler Coale, Ariel Rabines, Daniel Yee, Chris Dupont, Greg Rouse, Meredith Wright and Bogumil Karas, thank you all for your help with this research and my development as a scientist. A special thanks to Erin Bertrand and Jeff McQuaid for taking on intern Angela. You are both brilliant mentors, excellent scientists and I greatly appreciate the many hours spent leading me to a set of foundational skills in molecular microbiology research that got me through this PhD.

To the PIs and student teams in the McMurdo Long Term Ecological Research program, Mike Gooseff and Tina Takacs-Vesbach, thank you for giving me a dream opportunity and an extra science family to share ideas and experiences. Diane McKnight, I am so thankful for you and

am so stoked to have collaborated with you on these fantastic stream diatoms. Your studies are the reason why I became interested in Antarctic research and am so happy to have you as a mentor. Nick Schulte and Josh Darling thank you for the many times you pulled me out of the ice (figuratively, but also literally). I will cherish the time spent on the Stream/Algae Oops teams and with our F6 Family band. Art DeVries, thank you so much for opening up your lab and home to us and dedicating your time for work on the (really cool!) antifreeze assays. Jill Mikucki, thank you for going into your lab and freezers, in the middle of a pandemic, to get us more shots at putting together this research. To Deneb Karentz, the coordinators and participants of the NSF Antarctic Biology Training course, thank you for giving me a really great head start in this field and for insisting on waterproof pants.

To all my SIO friends and cohort; Jennifer Listen! Le, Jack Pan, Nati Delherbe & Alvaro Plominsky, Kiefer Forsch, Beverly French, Kaitlyn "I win cars" Lowder, Win Hongjamrassilp, Daniel Conley, Garfield Kwan, Canon Purdy, Sarah Maher, Sarah Schwenck (Schwenks a lot!) I have so much enjoyed our laughs, trips and fun times together. James Giammona, Abhishek Sareen, and Cecilia Chow, you all have been honorarily inducted into the above group. To Jeane Wong (Yoda) and the LXS crew, I am in awe of your unyielding determination to create equity in education. Thank you for letting me be a small part of this massive undertaking. I have nothing but admiration for all you do and continue to do.

A very special thanks to the people who got me here; My undergraduate professors; Wenda Ribeiro, Edward Weiss, Kathleen Brunke, Michael Meyer and Rick Sherwin. Your dedication to teaching and enthusiasm for all things science changed the course of my life.

Mom, Dad, Dorian, Shannon, Gavin and Lilah (and Doodley), I love you all so very much and thank you for always being little lights in the dark. This thesis is for you.

Chapter one has been prepared as a submission to the ISME Journal. A. Zoumplis, B. C. Kolody, D. Kaul, H. Zheng, P. Venepally, D. M. McKnight, A. DeVries, A. E. Allen. "Rapid destabilization of nitrogen-fixing microbial mats in a high-flow Dry Valleys meltwater stream." The dissertation author was the primary investigator and author of this paper.

Chapter two has been prepared as a submission to the PNAS Journal. A. Zoumplis, D. Kaul, N. Schulte, H. Zheng, P. Venepally, D. M. McKnight, A. E. Allen. "An intrusive, marine eukaryotic assemblage in the Antarctic Dry Valleys." The dissertation author was the primary investigator and author of this paper.

Chapter three, in part is currently being prepared for submission for publication of the material. A. Zoumplis, E. Garza, D. Yee, A. DeVries, V. Bielinski, A. E. Allen. "Heterologous expression and targeted delivery of sea ice diatom ice binding proteins." The dissertation author was the primary investigator and author of this paper.

VITA

2011 B. S. Biology, Christopher Newport University
2021 M. S. Oceanography, Scripps Institution of Oceanography, UCSD
2022 Ph. D. Marine Biology, Scripps Institution of Oceanography, UCSD

PUBLICATIONS

Zoumplis, A., Kolody, B. C., Kual, D., Zheng, H., McKnight, D. M., DeVries, A., Allen, A. E., Rapid destabilization of nitrogen-fixing microbial mats in a high-flow Dry Valleys meltwater stream. (in prep for ISME)

Zoumplis, A., Kaul, D., Schulte, N., Zheng, H., Venepally, P., McKnight, D. M., Allen, A. E. An intrusive, marine eukaryotic assemblage in the Antarctic Dry Valleys. (in prep for PNAS)

Schulte, N. O., Khan, A. L., Smith, E. W., **Zoumplis, A**., Kaul, D., Allen, A. E., McKnight, D. M. (2021). Blowin' in the wind: Dispersal, structure, and metacommunity dynamics of aeolian diatoms in the McMurdo Sound region, Antarctica. *Journal of Phycology*. https://doi.org/10.1111/jpy.13223

O'Rourke, A.*, **Zoumplis, A.***, Wilburn, P.*, Lee, M. D., Lee, Z., Vecina, M., Mercader, K. (2020). Following the Astrobiology Roadmap: Origins, Habitability and Future Exploration. *Current issues in molecular biology*, 38(1), 1-32. https://doi.org/10.21775/cimb.038.001 (*denotes equal contribution)

Nguyen, M. T., Lapointe, S., Jennings, B., **Zoumplis, A**. (2011). Gulf of Mexico Air Quality: CALIPSO Support for Gulf of Mexico Air Quality Relating to the Deepwater Horizon Oil Spill. *Earthzine* (No. NF1676L-11726). https://ntrs.nasa.gov/citations/20130012931

FIELDS OF STUDY

Molecular Microbial Ecology

ABSTRACT OF THE DISSERTATION

Adaptive physiology and structuring of microbial communities in ephemeral Antarctic environments.

by

Angela Zoumplis

Doctor of Philosophy in Marine Biology

University of California San Diego, 2022

Professor Andrew Allen, Chair

Antarctica hosts an expanse of ecological niches that approach the biophysical limits of life. Many landscape features, including sea ice, transient meltwater streams and ponds are defined by their impermanence. Instability within these systems is further complicated by the impingement of global climate change. Microbial communities flourishing within short-lived polar oases display a consortium of cold-thriving bacteria, eukaryotic algae, protists and micro-invertebrates. This thesis explores the diversity of these microbiota, the environmental drivers shaping biogeographical distribution patterns and the molecular underpinnings of responses to rapid physical and chemical fluctuations.

In Chapter 1, we utilize meta-omics tools to characterize the diversity and functionality of microbial mats in the Antarctic Dry Valleys trainset streams over spatiotemporal gradients. Findings shed light on the rapid turnovers in community structure following high flow and desiccation periods. Results provide a new view of active intra-stream diversity, biotic interactions and alterations in ecosystem function over a natural hydrological regime. Additionally, we document novel antifreeze activity in several Dry Valleys endemic taxa. In Chapter 2, we conduct a broad survey of molecular diversity on aquatic habitats of Victoria land, Antarctica. This research identifies salinity as a major driver of microbial diversity and identifies a marine eukaryotic community assemblage at the base of the Taylor Glacier. In Chapter 3 we focus on understanding the adaptations and acclimation responses of polar and temperate diatoms to cold stress while exploring the functionality of ice-binding protein (IBP) coding sequences. Heterologous expression and targeted delivery of sea ice diatom antifreeze proteins into a temperate diatom resulted in a significant reduction of mortality upon exposure to freeze/thaw cycling. Data in this chapter emphasizes the efficacy and predicts the essentiality of this adaptation to cold-thriving organisms.

INTRODUCTION

High latitude regions feature a variety of episodic and temporary habitats, many of which, at times, constitute the most biodiverse and productive communities relative to the regional landscape (Fernández-Méndez et al. 2015, Van Horn et al. 2016, Arrigo 2017). These short-lived environments can be more sensitive to climate change as they are tightly linked with seasonal fluctuations in temperature and hydrological regimes with little buffer compared to more permanent systems . In recent years, studies have documented unprecedented shifts in the advancements and retreats of sea ice and the flow cycles of intermittent meltwater streams (Doran et al. 2008, Stammerjohn et al. 2012, Fountain et al. 2014, Gooseff et al. 2017). Significant climate driven variations in these ephemeral communities lead to alterations to food web structures and whole ecosystem function (Gooseff et al. 2017, Michel et al. 2019). Despite their direct influence on polar ecosystems, little is known about the molecular diversity and *in situ* function of ephemeral biota.

This dissertation utilizes a combination of biochemical assays, environmental genomics, and diatom engineering and bioinformatics tools to explore temporal and spatial patterns of diversity and ecophysiology of polar microbes. In Chapter 1 we collected samples from spatially distance and phenotypically distinct microbial mats across a high-flow Dry Valleys meltwater stream along with flow measurements over the austral summer season to observe the hydrological drivers of stream community structure and function. We observed a significant destabilization of the nitrogen-fixing community of the marginal mats during peak flow times. Functionally, the mats were distinct though synchronies in gene expression were identified at the high and low flow intervals throughout the season. Chapter 2 uses an extensive Victorialand dataset to investigate intrusive marine eukaryotes in the Antarctic valley network. A molecular survey of polar

terrestrial, aeolian and freshwater aquatic and nearby marine environments led to the identification of an active, relic microbial community within the Taylor Valley. Further study into the functional attributes of this community showed activity related to primary metabolisms, stress coping mechanisms and cell signaling. Chapter 3 applies innovations in diatom genome editing tools to an investigation of polar adaptive strategies and whether ice binding proteins (IBPs) confer freeze resistance to the model temperate diatom, Phaeodactylum tricornutum. Results show a significant reduction of mortality after a freeze thaw event in transformants with intracellularly localized IBPs over the wildtype. Overall, this dissertation describes the drivers of diversity and explores the functional basis for survival of polar microbial communities in ephemeral habitats.

REFERENCES FOR THE INTRODUCTION

1. Arrigo, K. R. (2017) Sea ice as a habitat for primary producers. *Sea ice***,** 352-369.

2. Doran, P. T., C. P. McKay, A. G. Fountain, T. Nylen, D. M. McKnight, C. Jaros & J. E. Barrett (2008) Hydrologic response to extreme warm and cold summers in the McMurdo Dry Valleys, East Antarctica. *Antarctic Science,* 20**,** 499-509.

3. Fernández-Méndez, M., C. Katlein, B. Rabe, M. Nicolaus, I. Peeken, K. Bakker, H. Flores & A. Boetius (2015) Photosynthetic production in the central Arctic Ocean during the record sea-ice minimum in 2012. *Biogeosciences,* 12**,** 3525-3549.

4. Fountain, A. G., J. S. Levy, M. N. Gooseff & D. Van Horn (2014) The McMurdo Dry Valleys: a landscape on the threshold of change. *Geomorphology,* 225**,** 25-35.

5. Gooseff, M. N., J. E. Barrett, B. J. Adams, P. T. Doran, A. G. Fountain, W. B. Lyons, D. M. McKnight, J. C. Priscu, E. R. Sokol & C. Takacs-Vesbach (2017) Decadal ecosystem response to an anomalous melt season in a polar desert in Antarctica. *Nature ecology & evolution,* 1**,** 1334-1338.

6. Michel, L. N., B. Danis, P. Dubois, M. Eleaume, J. Fournier, C. Gallut, P. Jane & G. Lepoint (2019) Increased sea ice cover alters food web structure in East Antarctica. *Scientific reports,* 9**,** 1-11.

7. Stammerjohn, S., R. Massom, D. Rind & D. Martinson (2012) Regions of rapid sea ice change: An inter-hemispheric seasonal comparison. *Geophysical Research Letters,* 39.

8. Van Horn, D. J., C. R. Wolf, D. R. Colman, X. Jiang, T. J. Kohler, D. M. McKnight, L. F. Stanish, T. Yazzie & C. D. Takacs-Vesbach (2016) Patterns of bacterial biodiversity in the glacial meltwater streams of the McMurdo Dry Valleys, Antarctica. *FEMS Microbiology Ecology,* 92**,** fiw148.

Rapid destabilization of nitrogen-fixing microbial mats in a high-flow Dry Valleys meltwater stream

A. Zoumplis[1,2], B. Kolody[3], D. Kaul[2], H. Zheng[2], P. Venepally[2], D. M. McKnight[4], C. Takacs-Vesbach[5], A. DeVries[6], A. E. Allen[1,2*]

1. Scripps Institution of Oceanography, University of California, San Diego, CA 92093
2. Microbial and Environmental Genomics Group, J. Craig Venter Institute, La Jolla, CA 92037
3. Earth and Planetary Science, University of California, Berkeley, Berkeley, CA, 94720
4. Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, 80309
5. Department of Biology, University of New Mexico, Albuquerque, NM, 87131
6. Evolution, Ecology and Behavior, University of Illinois at Urbana-Champaign, Urbana, IL 61801

*Correspondence

Abstract

The meltwater streams of the McMurdo Dry Valleys are hot spots of biological diversity in the climate-sensitive polar desert landscape. Microbial mats, largely comprised of cyanobacteria, dominate the streams which flow for a brief window of time (~10 weeks) over the austral summer. These communities, critical to nutrient and carbon cycling, display previously uncharacterized patterns of rapid destabilization and recovery upon exposure to variable and physiologically detrimental conditions. Here, we characterize changes in biodiversity, transcriptional responses and activity of microbial mats in response to hydrological disturbance over spatiotemporal gradients. While diverse metabolic strategies persist between marginal mats and main channel mats, data collected at mid-season peak meltwater flow revealed a homogenization of the mat communities, directly influencing the biogeochemical roles of this stream ecosystem. Gene expression pattern analyses identified strong functional sensitivities of nitrogen-fixing marginal mats to changes in hydrological activities. Stress response markers detailed the environmental challenges of each microhabitat and the molecular mechanisms underpinning survival in a polar desert ecosystem at the forefront of climate change. At mid and end points in the flow cycle, mobile genetic elements were upregulated across all mat types indicating high degrees of genome evolvability and transcriptional synchronies. Additionally, we identified novel antifreeze activity in the stream microbial mats indicating the presence of ice-binding proteins (IBPs). Cumulatively, these data provide a new view of active intra-stream diversity, biotic interactions and alterations in ecosystem function over a high flow hydrological regime.

Introduction

The McMurdo Dry Valleys (MDVs) are considered one of the harshest environments on Earth (Doran, Lyons and McKnight 2010). Low temperatures, rapid freeze/thaw cycles, aridity, variations in light regimes, steep chemical and salt gradients, and nutrient bioavailability all pose challenges to life in this evolving landscape (Barrett et al. 2007, Doran et al. 2002a, Fountain et al. 2010, Hawes and Schwarz 2001, McKnight et al. 1999, Toner, Sletten and Prentice 2013). In the early 1990's the MDVs experienced a decadal cooling trend (Doran et al. 2002b). In 2006 this cooling trend terminated and the MDVs have since experienced higher and more variable temperatures with high-flow and flood events occurring more frequently (Gooseff et al. 2017, Obryk et al. 2020). Warming events have resulted in rapid and sustained effects on microbial communities in the polar desert landscape (Nielsen et al. 2012, Fountain et al. 2016, Andriuzzi et al. 2018). Data presented in this study on the stability of keystone species and processes in response to natural, hydrological variation sheds light on the future predictions of ecosystem stability in response to climate-driven variation in the Antarctic Dry Valleys.

Microbes dominate several ecosystems within the polar desert including the glaciers, soils, ponds, meltwater streams, and perennial ice covered lakes (Adams et al. 2014, Cary et al. 2010, Jungblut et al. 2005, Kohler et al. 2015, Sommers et al. 2019, Wharton Jr, Parker and Simmons Jr 1983). The glacial meltwater streams constitute the most biodiverse habitat of the MDVs (Esposito et al. 2008, Van Horn et al. 2016, Wharton Jr et al. 1983). Throughout the austral summer, rising temperatures and increasing solar radiation cause glacial melting that flows downwards, saturating stream beds and underlying hyporheic zones (Wlostowski et al. 2016). Streams flow is characterized by daily pulses driven by changing sun angles and punctuated periods of high flow

for two to three months before returning to a desiccated, frozen state over the austral winter (McKnight and Tate 1997).

Most of the biomass in the MDVs is in the form of benthic microbial mats (Davey and Clarke 1992). Cyanobacteria form the basis of these cohesive mats which include bacteria, eukaryotic algae, protists and micro-invertebrates (Davey and Clarke 1992, Vincent 2000). Organisms within the mats are thought to be highly adapted to life in a polar desert. Adaptations such as rapid metabolomic and reproductive reactivation, efficient photosystems, and nutrient scavenging protect against intermittent water activity, low light levels and oligotrophy, respectively (Hawes and Schwarz 2001, McKnight et al. 2007, Varin et al. 2012). Stream microbial mats typically form along stream margins and within main channels (Alger 1997). Flow dynamics of the streams create hydration gradients between these locations establishing niches of microbial communities (Alger 1997, Kohler et al. 2015). Within an individual stream, the pigment diversity associated with stream mats of differing habitats is a potential indication of selective pressure related to solar irradiance (Marizcurrena et al. 2019). Orange mats are predominately found in the main stream channel forming cohesive benthic mats (Vincent et al. 1993). Green mats, typically, are found in the main stream channel attached to rocks, with black mats forming near the stream margin (Alger 1997). The observed red mats are variably located intermediately between margin and channel habitats. The presence and biomass of these mats have been historically monitored based on photosynthetic pigments (Vincent et al. 1993, Howard-Williams et al. 1986, Kohler et al. 2015, Van Horn et al. 2016).

Placement of a mat within a stream may affect the degree of light refraction, availability of nutrients, exposure to low temperatures and periodic desiccation and other factors that may drive community structure (Howard-Williams et al. 1986). Furthermore, these hydrological and

environmental conditions are stream dependent and change throughout the austral summer with peak flow occurring during the month of January (Howard-Williams et al. 1986). Past studies have emphasized flow dynamics and hydrological regimes as important drivers of stream bacteria and diatom community structures (Esposito et al. 2006, Kohler et al. 2015, McKnight et al. 1999, Stanish, Nemergut and McKnight 2011, Van Horn et al. 2016). Limitations of biodiversity studies conducted in the meltwater streams include classifications based on morphological characterizations which may underestimate the number of species within communities. Likewise, DNA methods may be placing heavy emphasis on nonfunctioning organisms and potentially inactive genetic material. Characterizing the active communities of the meltwater streams to *in situ* flow cycles is critical to understanding the stability of this ecosystem in a climate sensitive region.

To date, there has not been an in depth look at the active biodiversity and transcriptional response of *in situ* meltwater stream communities to hydrological regimes over ecologically meaningful spatial and temporal scales. Here, we characterize within stream mat functional variation over an austral season and establish the molecular underpinnings of survival in this unique and changing environment. Our results provide a new view into the critical roles of microbial mats in the biological uptake and biogeochemical cycling of nutrients in a rapidly changing, climate-sensitive environment. The patterns identified in this study can provide insights into microbial community responses to future climate-driven hydrological changes and flood events in a warming polar ecosystem. Furthermore, results identify genes and pathways with potential novel biotechnological application.

Materials and Methods

Sample Collection

Algal mat samples were collected during the 2016-2017 field season from a pre-existing microbial sampling transect established by the McMurdo Long Term Ecological Research (MCM-LTER) program (Alger 1997). Sampling was conducted at the Canada Stream transect (-77.6132414, 163.0526875) within the Taylor Valley of the McMurdo Dry Valleys (Fig. 1A). The hydrological record from the Canada Stream has been maintained consistently with annual flow data measurements available since 1990. Canada Stream is the most consistently flowing stream in the Fryxell basin receiving flow from the East facing tongue of the Canada Glacier that drains into two ponds prior to entering a well-defined channel (Wlostowski et al. 2016). Flow in Canada Stream begins earlier and ends later in the season than other Fryxell basin streams and has a greater probability of high flows (100 L/s) which can result in the scouring of microbial mats (Wlostowski et al. 2016, Cullis, Stanish and McKnight 2014).

Microbial mats were collected from the Canada Stream margin, channel and intermediate sites at four time points throughout the summer season to capture flow dynamics that are hypothesized to drive microhabitat diversity. Orange and green mats were collected from the main channel. Red and black mats were collected near and on the stream margin respectively (Fig. 1B). Mat samples for transcriptomics and rRNA analyses were collected in triplicates using an EtOH-sterilized #13 brass cork borer (227 mm^2) and forceps, and transferred into sterile 2 ml cryovials and covered with ~1ml of RNALater. Cryovials were then flash frozen in liquid nitrogen and stored at -80C. Mat samples collected for the purpose of antifreeze activity assays were collected in sterile whirlpack bags and stored at -20C.

Methods

RNA and DNA were extracted from frozen, RNAlater™ (Thermo Fisher Scientific) preserved microbial mat samples in accordance with NucleoMag RNA (Macherey-Nagel) and NucleoMag Plant (Macherey-Nagel) kit protocols. The V4-V5 region and V4 region of the 16S and 18S ribosomal RNA (rRNA) was amplified, independently from RNA and DNA using universal primers (Amaral-Zettler et al. 2009, Parada, Needham and Fuhrman 2016, Stoeck et al. 2010). Full methods for sample processing for sequencing are provided in Supplementary File 1. Libraries were sequenced using the Illumina MiSeq platform (2 x 300bp) generating a total of ~2M reads with an average number of 150,000 reads per sample for both 16S and 18S datasets. Sequences were quality filtered and amplicon sequence variants (ASVs) were generated using the DADA2 (Callahan et al. 2016) module in QIIME2 (version 2019.4) (Caporaso et al. 2010) with default thresholds for expected error and the following length truncation parameters: 250 bp for forward truncation (--p-trunc-len-f) and 200 bp for reverse truncation (--p-trunc-len-r).

Taxonomic annotations for 16S and 18S rRNA ASVs were assigned through rRNA reference databases including SILVA (version 132) and NCBI-NT/NR, respectively. Alpha and beta diversity metrics were calculated using the "core-metrics-phylogenetic" method using default parameters and visualized using 'Emperor' tool in QIIME 2 (version 2019.4).

Metatranscriptomes were generated from isolated mRNA. cDNA synthesis and simultaneous amplification of polyA as well as total RNA results in substantial coverage of prokaryotic and eukaryotic mRNA (Bertrand et al. 2015, Dupont et al. 2015). Metatranscriptomic libraries were sequenced on the Illumina HiSeq 4000 platform generating approximately 100 Gb/ 400 million reads per lane, ~5 Gb/10 million paired reads per sample. Post-sequencing, reads were filtered and trimmed for removal of primers, adapters, and low quality sequences. Ribopicker

(version 0.4.3) (Schmieder, Lim and Edwards 2012) was used to identify and remove rRNA. Transcripts were assembled into contigs with CLC Genomics (version) and open reading frames (ORFs) were predicted with FragGeneScan (version 1.31) (Rho, Tang and Ye 2010). ORFs were annotated de novo for function against the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2012), Pfam (Finn et al. 2014) and hidden Markov model (HMM) searches (Finn, Clements and Eddy 2011), and the in-house comprehensive reference database, PhyloDB (version 1.076) for taxonomic annotation of these ORF sequences.

For differential expression analysis, between mat type, triplicate samples collected on 12/18/16, 1/9/17 and 1/28/17 were pooled based on mat type. Pairwise differential expression analyses of transcripts between mat types were performed using the R package edgeR (version 3.10.5) (Robinson, McCarthy and Smyth 2010). Counts were normalized using the "calcNormFactors" function and an extract test with tagwise dispersion was used to test for differential expression across mat types. Resulting p-values were False Discovery Rate (FDR) corrected for multiple testing (Benjamini-Hochberg) and FDR < 0.05 was used as a significance threshold.

In order to more comprehensively annotate ORF function, ORFs were grouped into clusters of similar amino acid sequences using the Markov Cluster Algorithm (MCL). Directional edge weights were set as the ratio of pairwise- to self- BLASTP scores, and default parameters were used to assign ORFs to clusters. A consensus cluster annotation was called by clusters being statistically enriched in annotations with a Fisher's exact test ($p < 0.05$). Clusters that were annotated as having the same function were grouped into "functional clusters." A weighted gene co-expression network analysis (WGCNA) (version 1.70.3) (Langfelder and Horvath 2008) was

performed as previously described (Kolody et al. 2019) to identify groups, or modules, of functional clusters with synchronous expression patterns across the four mat types. Only functional clusters with at least 500 reads in 80% of samples were included in the WGCNA analysis. WGCNA was run on library-normalized counts. Modules were detected using the blockwiseModules function using a power function exponent, b, of 14 to optimize scale-free topology, a minimum module size of 30 clusters, and parameters detectCutHeight= 0.995, reassignThreshold = 0, mergeCutHeight = 0.5. The top 5 most abundant clusters in each module were plotted using igraph and tcltk (Fig. 7).

Orange, green, red and black mat samples were collected from the Canada stream and stored and shipped at -20 °C for antifreeze activity assays. One gram of each mat type was added to a 2 ml microcentrifuge tube filled with 200 µl) distilled water. The mats were then put through 5 freeze thaw cycles (-20 °C for 30 mins followed by a 10 min thaw at room temperature). Samples were spun down for 10 mins at 5000 rpm at 4 °C and the supernatant was collected.

A Clifton Nanoliter Cryoscope (Clifton, New Jersey) was used to determine the freezing and melting points of the lysed mat samples. The cryoscope is fitted with a cold stage utilizing a temperature-controlled Peltier sample holder. Temperatures are accurate to 0.01 °C. Each of the 600 um diameter sample holder wells were loaded with heavy microscope immersion oil (Type B). Ten nanoliters of the collected microbial mat supernatant was inserted into the center of each oil-filled well with a micropipette. Controls were distilled water (Control 1) and 1000 mOsmol standard (Control 2).

Once loaded, samples were cooled to -40 °C. The frozen wells were warmed until only a single isolated 10um ice crystal was left in the well. Melting points of the isolated crystal were determined by lowing and raising the temperatures until the melting velocity was undetectable.

The melting points of solutions with no antifreeze protein activity were at equilibrium with the freezing point. Single ice crystals were slowly cooled at a rate of 0.074 °C/min. The non-equilibrium freezing points were recorded and the difference from the equilibrium melting point is reported as the thermal hysteresis. In additional experiments, the single, small ice crystal underwent a 30 min period of annealing at a temperature slightly below the melting point before cooling after which the thermal hysteresis was recorded.

Results

Sensitivity of community structuring to meltwater influx

Ash free dry mass (AFDM) and chlorophyll *a* analyses were conducted on mat samples collected early January. AFDM values were significantly differentiated among black/red mats vs. green/orange mats (Table 1). Black mats reported the highest ratio of AFDM to chl *a*, followed by red mats (Table 1). This result is likely due to elevated detritus deposition to the stream marginal mats as well as higher proportions of non-chlorophyll possessing organisms (Fig. 2b). Eukaryotic species richness was highest in orange and early season black mats (Fig. S1). Sharp drop-offs in observed species for eukaryotes occurred in black and red mat as the season progressed (Fig. S2). Prokaryotic species richness remained relatively consistent throughout the season across mat types (Fig. S3, S4).

Table 1.1 Ash Free Dry Mass (AFDM), Chlorophyll a (Chl a) and AFDM to Chl a ratio values for green, orange, red and black mats collected from 1/9/17.

Color	AFDM	Chl a	AFDM:Chl a ratio
Green	7.84	2.91	2.69
Orange	14.3	3.06	4.69
Red	22.8	2.52	9.06
Black	27.8	0.229	121

Across 14 samples, 15,477 ASVs were generated with 8,611 prokaryotic, 16S (V4-V5 region) and 6,866 eukaryotic, 18S (V4 region) sequences. Prokaryotic 16S rRNA communities across all mat types were dominated by Cyanobacteria (80%). Other major groups of bacteria included Proteobacteria (8%) Bacteroidetes (6%), Chloroflexi (2%) and Planctomycetes (0.80%) (Fig. 2a). The dominance of Cyanobacteria over the non-cyanobacterial sequences in the prokaryotic community is in contrast to the relatively low cyanobacteria abundances (~35%) observed in other mat community studies in similar polar, coastal and freshwater regions (Fig. 3) (Bolhuis and Stal 2011, Sorokovikova et al. 2013, Van Horn et al. 2016, Varin et al. 2012). We attribute these differences to DNA vs. RNA sequencing where more of the active community is revealed (Fig. 3). RNA sequencing is preferred for determining the viable community over DNA and morphological methods given its rapid degradation in non-viable cells (Blazewicz et al. 2013).

Of the eukaryotes, Opisthokonta averaged a high proportion (66%) of the 18S rRNA community across all mats types, followed by Alveolata (14%), Stramenopiles (4%), Rhizaria (4%), Viridiplantae (4%) and Amoebozoa (1%) (Fig. 2b). A less dominant group, Apusozoa (0.05%), displayed a limited presence in black and orange mats only. A more detailed look into the dominant Opisthokonta reveals taxa group variation among mat types. Rotifera averaged over half of the metazoan community for orange and red mats, whereas green mats are dominated by Tardigrada. Black mats have the highest proportion of Nematoda of the mat types (Fig. S5).

Additional sampling was conducted throughout the season in order to detect community changes over the austral summer. The eukaryotic communities of orange, green and red mats remained relatively stable over time, while black mats on the stream margin experienced strong shifts in taxa groups in both the prokaryotic and eukaryotic communities (Fig. 2). Early season black mats show peaks in Alveolata and Stramenopile, while Opisthokonta dominate the mid-

season sample. Late season black mats show a decline in Opisthokonta and an increase in Amoebozoa and Rhizaria. Mid-season black mats, where stream flow is at its maximum, resemble the taxa structure of the green and orange stream bed mats. This shift is evident across multiple taxa groups, particularly in the dominant cyanobacteria communities where black mats experience a mid-season shift from the nitrogen-fixing *Nostoc* and *Wilmottia* to *Tychonema* and *Scytonema* before a reversion in late season (Fig. 2).

Principle coordinate analysis (PCoA) was performed on prokaryotic (16S) and eukaryotic (18S) rRNA communities using weighted unifrac distance metric. Significant clustering was found in both 16S and 18S rRNA data (p-value < 0.01, ANOSIM R-statistic 0.307 and 0.480 respectively) (Fig. 4). Beta diversity clustering in prokaryotes reflected a similar community structure as observed in the taxonomic analyses. November, early season, red and black mat types cluster together. Though displaying distinct community signatures, mid-season mats of all types generally cluster together. Red mats appear to have the most dissimilar prokaryotic communities among the four mat types, largely dominated by the cyanobacteria, *Wilmottia* (Fig. 2C, 4). Eukaryotic clustering patterns show clustering between black, red and orange mats with an isolated green mat cluster (Fig. 4). Clustering was strongest among the Rhizaria and Metazoa groups indicating higher similarities within samples of a particular mat type and higher dissimilarities between the different mat types (p-value < 0.01, ANOSIM R-statistic 0.525 and 0.456 respectively). Regardless of temporal distance, eukaryotes were grouped by mat type indicating a lesser impact of hydrological drivers on eukaryotic community structuring (Fig. 4).

The nitrogen-fixing, black marginal mats and red, intermediate mats were the most susceptible to shifts in prokaryotic taxa groups over time with channel mat taxa remaining stable. These shifts were rapid (time period of weeks) indicating rapid community turnover with changes

in flow patterns. Marginal mats are subject to scouring during high flow periods. Particulate organic matter (POM) sampled over diel flow cycles showed an increase in POM concentration with increased water levels (Cullis et al. 2014). During these daily flow pulses, shear forces act on the microbial mats (Cullis et al. 2014). The most predominate form of POM in transport through the streams and hyporheic zones is *Nostoc* derived, as determined by the isotopic signature of N-fixation (Kohler et al. 2018). These results support our findings of unstable nitrogen-fixing marginal mat communities. A reduction in flow near the end of the austral summer season resulted in the recovery of *Nostoc* and nitrogen fixers that dominated the marginal mats at the beginning of the sampling period. The rapid recovery of this community to its pre-flow state indicates resilience of these critical, keystone microbes.

Past studies have sampled mat communities across several streams in the Miers, Taylor and Wright Valleys to detect patterns in prokaryotic diversity (Van Horn et al. 2016). Though phenotypically and spatially these mats appear similar among streams, there is significant variation in seasonal flow conditions (Stanish et al. 2011). Ultimately, these intra-stream processes and physical controls on diversity could explain the lack of homogeneity among meltwater stream mats and potentially why a black, marginal mat from a stream at peak flow may appear similar to orange, channel mats at other locations, among other observed discrepancies.

Nutrient cycling and stress responses are significantly differentiated by mat type

Across all mat types, orange and black mats had the highest percentage of significantly differentially expressed genes (34%) followed by green and black mats (23%) (Fig. 5A). Furthermore, black and red mats had no significantly differentially expressed genes (Fig. 5A). The transcriptional profiles of the orange and black mats showed uniquely enriched genes

corresponding to geobiological processes and stress responses (Fig. 5B). Identification of these functional disparities broadens our understanding of the structuring of these communities and the degree of physiological plasticity required for survival in this transient environment.

Nitrogen cycling

Diazotrophs and their role in the Dry Valleys microbial communities have long been documented through numerous diversity and nitrogen (N) cycling activity studies (Coyne et al. 2020, Kohler et al. 2018, McKnight et al. 2004). N-fixing cyanobacteria are known to be a primary source of N to meltwater steams along with glacier derived input (Howard-Williams, Priscu and Vincent 1989). Black marginal mats in particular are thought to be an important niche for generating bioavailable N for the biosphere. Differential gene expression analysis between black and orange mat types provides new insights into the N cycling strategies between these communities. Black mats show upregulation of nitrogenase genes in several species of cyanobacteria including *Anabena*, *Calothrix*, and *Nostoc* (Fig. 5B). Orange mats displayed an abundance of urease genes indicating catabolism of urea into NH_4 and CO_2 (Fig. 5B). The ability to utilize and repurpose organic N is likely an important adaptation in this N limited environment. In the meltwater streams, urea is produced by micro-invertebrates. Our findings suggest that *Anabena*, *Leptolyngbya*, *Microcoleus*, *Nostoc* and *Synechococcus* are key taxa for urea utilization (Fig. 5B).

Our functional analysis did not detect any genes related to nitrification in any of the four mats types. Potential for nitrification and the accumulation of N has primarily been found to occur in the stream sediment and hyporheic zone (Hopkins et al. 2006, Singley et al. 2021). As such, nitrification genes were not expected to have been dominantly expressed in our mat samples. In

regard to N-fixing capabilities of *Nostoc*-dominated black mats, our data show influence of these mats on N cycling varies over the season as nitrogen fixation genes are upregulated during early and late low-flow time points while expression declines during mid-season peak flow (Fig 5C). Orange mat expression of the urease gene indicative of organic N (urea) utilization increased from early season to late season (Fig. 5C).

Carbon cycling

The identification of carbon fixation related enzymes revealed an abundance of autotrophic organisms in the mats. A search for alternate carbon source utilization yielded support for the catabolic degradation of starches, cellulose, and hemicellulose among cyanobacteria. The use of exogenous carbon sources is not surprising in cyanobacteria evolved in ecosystems undergoing natural, extended periods of darkness. In this dataset, we identify several genera of cyanobacteria potentially capable of heterotrophic metabolism including *Leptolyngbya*, *Microcoleus*, *Nostoc, Moorea, and Calothrix* (Fig. 5B).

Stress Responses

Low temperatures and freeze events have been shown to upregulate stress responses such as lipid biosynthesis, fatty acid desaturases and heat shock proteins (HSP), in cold dwelling microorganisms (Raymond-Bouchard and Whyte 2017, Králová 2017). The biosynthesis and transport of compatible solutes including glycine betaine, ectoine and trehalose prevent water loss from cells as ice forms (Chua et al. 2018) . Here, we find marginal mats primarily undergoing temperature stress as indicated by upregulation of HSP family genes (Fig. 5B). At the beginning and end of the austral summer season, black mats are periodically exposed to air temperatures

during periods of low flow. Air temperatures in the Fryxell basin averaged -2.92 °C (σ 3.16) and -1.45 °C (σ 2.22) throughout the December and January months, respectively. These temperatures are significantly lower than the average water temperatures recorded by the Canada stream gauges for December 4.08 °C (σ 2.14) and January 3.48 °C (σ 2.51). Freeze/thaw cycling can result in the upregulation of genes encoding heat shock proteins which is seen in our data (Fig. 5B). Additionally, desiccation and oxidative stress responses have also been identified in marginal mats (Fig. 5B). In contrast, channel mat organisms expressed elevated osmoprotectant transcripts (Fig. 5B). Cyanobacteria species from both mat types appear to be nitrogen and phosphorous limited though black mats exhibit a more taxonomically diverse response to these nutrient deficiencies (Fig 5B).

This differential expression analysis enabled the discernment of specific contributions of mat types to biogeochemical processes, including nitrogen and carbon cycling, and identified stress responses providing insight into environmental challenges. Observing gene expression over the austral summer season provided insights into the abruptness and destabilization of critical functions (i.e. nitrogen fixation) in lotic periods.

Weighted correlation network analysis reveals hydrological drivers of functional synchronies

The weighted correlation network analysis (WGCNA) identified eight modules (Fig. 6) of co-expressed functional clusters. A majority of these clusters fall into modules that have expression patterns that appear to be responding to mat type and flow dynamics (Fig. 6, 7). The number of clusters in each module ranges from 1610 (module 1) to 44 (module 8) (Fig. 6). Module

1 displays a clear difference in activity established by the upregulation in green and orange mats compared to the black and red mats (Fig. 7). Top gene activities of modules 1, 3, 5 appear to be driven by high stream flow (Fig. 7). Functions exhibiting peaks in hydration periods include photosystem II, eukaryote related cytochrome c oxidase and histone ORFs (Fig. 6, 7). Modules 2, 4, 6 and 7 are correlated to low flow regimes where top functions include photosystem I, binding proteins including fasciclin and PF12849 (periplasmic binding protein domain), and PF12680 (antibiotic formation) (Fig. 6, 7). Likely, photoinhibition of the PSII is caused by exposure to higher irradiance levels and accumulation of reactive oxygen species (ROS) in the low flow modules whereas we see higher expression of the protected PSI (Gururani, Venkatesh and Tran 2015, Murata et al. 2007). The ancient fasciclin domain is present across many taxa groups and is important in cell interaction with the external environment (Seifert 2018). As a significant component in the extracellular matrix, fasciclin plays a variety of roles in cell adhesion and stress management (Meng et al. 2020, Rai et al. 2020, Seifert 2018). Here, we see elevated transcripts of cyanobacterial fasciclin proteins in periods of desiccation and freeze/thaw (Fig. 6, 7). Module 7 is comprised largely of transposases that are upregulated across all mat types at the end of flow indicating a potential stress response to the start of the desiccation period (Fig. 6, 7). Orange and black mats both exhibit an upregulation of genes encoded by viruses during periods of high flow apparent in module 8 (Fig. 6, 7).

Functionally, black marginal mats behaved more similarly to red intermediate mats than to distal channel mats. The black, marginal mats appear to be the most transcriptionally sensitive to hydrological activity within the stream. External environment triggering occurs throughout the austral summer season in these mats as visualized by mat-type specific upregulation of gene clusters in response to changes in flow over time. Modules 2 and 6 display an upregulation in the

function of marginal mats during low flow intervals (Fig. 7). Conversely, gene expression patterns in the black, marginal mats appears to synchronize with other mat types over peak flow in modules 5 and 8 (Fig. 7). The cluster expression synchronies in module 5 are functionally annotated as metazoan which can in part, be due to a shift in the marginal mat community to metazoan taxa with increased flow (Fig. 2B, 6).

Mobile genetic elements including transposases, homing endonucleases and intron maturases are found to be upregulated in the microbial mats at mid and end points in the flow cycle (Fig. 6, 7). These genes may be part of an adaptive toolkit for the mats resulting in potential trait plasticity in response to the stress of meltwater stream variability. Upregulation of RNA-directed DNA polymerase functional clusters in black and orange mats during peak flow potentially suggest virus mediated alterations of cyanobacterial host physiology or host defenses against viral activity (Fig. 6, 7). Infection-related host phenotypes have not been determined in this study though viral genetic exchange can be advantageous for host evolution.

Novel thermal hysteresis and antifreeze activity in Dry Valleys microbial mats

All mat types across Canada Stream displayed some degree of thermal hysteresis (TH) in the antifreeze assays (Fig. 8). The strongest hysteresis in the 'no anneal' time trials was the black mat with a TH of 0.17°C followed by green mats (0.056 °C), red mats (0.037 °C) and orange mats (0.015 °C). Ice crystal seeds grew as a disc for orange mats which is consistent with controls. Orange mat ice crystal visualization along with thermal hysteresis results indicate little to no antifreeze activity. Black and green mat supernatants displayed faceting on a-axis (prism planes). Red mats displayed growth in patches on basal plane as well as some faceting on a-axis.

An additional antifreeze assay was performed using the same methods previously described with an extra step of a 30 minute anneal before cooling the ice crystal. Annealing time along with concentration and slower cooling rates have had significant impacts in increasing the TH activity (Knight and DeVries 2009, Kubota 2011, Takamichi et al. 2007). In this investigation the additional annealing time resulted in a significant increase in TH activity and binding in red (1.4 °C), black (0.72 °C) and green mats (0.11 °C). Black and green mats had relatively high TH variability among samples. Orange mats (0.065 °C) remained at little to no antifreeze activity. Dendritic bursts were evident in black and red mats upon reaching the freezing point threshold (Fig. S7). Green mats displayed evidence of prism plane faceting. These results support the idea that annealing time has a strong influence on the thermal hysteresis activity.

A Hidden Markov Model (HMM) search identified ORFs of the conserved ice-binding-like protein Pfam (PF11999). This Pfam includes the domain of unknown function (DUF3494) which contains ice binding proteins (IBPs) from a diverse array of organism with varying magnitudes of function (Vance et al. 2019). The search yielded a large percentage of prokaryotic derived proteins with several stramenopile and fungi variations. The counts were highest in early season black and red mats in comparison to the lower count orange and green mats (Fig. 5C). The taxonomic representation in the highly expressing black and red mats consists of *Singulisphaera acidiphila* and *Streptomyces sviceus* belonging to the Planctomycetes and Actinobacteria groups, respectively, with the top Eukaryotic hit being the fungi *Zymoseptoria tritici* (Fig. S8). Alternatively, the Planctomycetes species *Pirellula staleyi* and the stramenopile, *Ochromonas* were the most dominant IBP encoding taxa in the orange and green mats (Fig. S8). No taxa described here are exclusive to any one mat.

There is potential for novel structural IBP diversity in the microbial mats as up to 5% of the IBP expressing community in early season black and red mats were from unclassified taxa groups (Fig. S8). Applications of such antifreeze activity described in this analysis expands to food, medical and biotechnology industries. It appears red mats, in particular have moderate antifreeze activity, with a TH reaching ~ 1°C (Bar Dolev, Braslavsky and Davies 2016), under certain freezing conditions. While we have reported a higher TH activity with longer annealing times, additional factors known to increase TH activity include higher IBP concentration and slower freezing rates (Bar Dolev et al. 2016).

Conclusion

The findings in this study detail changes in the active community structures and gene expression patterns of microbial mats in a high-flow meltwater stream. The variability and functional potential of stream communities may play important roles in landscape processes over time. Black, marginal mats contribute heavily to the N_2 fixation and are subsequently responsible for N inputs after depletion of glacier-derived N (Kohler et al. 2018). These mats, though critical to the sustainability of the meltwater ecosystem, undergo destabilization of community structure and function. While we have established some of the mechanistic means of black mat resilience to the harsh, rapidly fluctuating environment, the instability of keystone community members with hydrological variation is striking, representing a distinctive feature of ecosystem function for these streams. The polar desert is a region prone to exacerbated effects of the warming climate (Levy et al. 2018, Fountain et al. 2014). Past studies have shown the dramatic and persistent effects on diversity of microbial mat taxa during flood pulse events (Barrett et al. 2008, Nielsen et al. 2012,

Fountain et al. 2016). With future projections of climate-driven variation (Gooseff et al. 2011), it is unknown if these mats will continue to adapt and persist in their polar desert niche.

There is a lack of molecular biodiversity and functional expression data for hydrological and solar transition periods in the Dry Valleys meltwater streams. Data on responses of lake bacterioplankton, phytoplankton and protists taxa groups to bimodal solar cycling indicate degrees of metabolic plasticity that should be further explored with modern molecular approaches (Vick-Majors, Priscu and Amaral-Zettler 2014, Bielewicz et al. 2011, Vick and Priscu 2012, Morgan-Kiss et al. 2016). Likewise, additional stream mat sampling over spatio-temporal gradients and continuous periods can provide a better picture of the biogeochemical contributions and stability of stream biota within the rapidly changing polar desert system. Moreover, this study provides relevant insight to other stream habitats as functionally important species, such as Nostoc, are prevalent across a wide range of climate-sensitive ecosystems.

Finally, we have presented information on the ice shaping abilities and thermal hysteresis activities on the meltwater stream microbial mats. In particular, black marginal mats and red intermediate mats hosts microbes that have potential uses in food, biotechnology and medical fields. The autecology and cultivability of microbes from these mats is a challenge to the further characterization of taxa specific antifreeze activity. Additional studies on focused on recombinant expression characterization, purification and stability of these potentially novel IBPs could lead to a better understanding of survival through freeze/thaw cycling but also industrial application.

Conflict of Interest

The authors declare no conflicts of interest

Acknowledgements

Fieldwork was supported by the National Science Foundation Grant OPP-1637708 (to the MCM LTER). This study was funded by NSF Antarctic Sciences Awards NSF-OPP 0732822 and NSF-OPP-1043671 (to A.E.A.), Gordon and Betty Moore Foundation Grant GBMF3828 (to A.E.A.); NSF Ocean Sciences Award NSF-OCE-1136477 and NSF-OCE-1756884 (to A.E.A.) We thank the members of the MCM LTER algae and streams teams that contributed to the field collections of samples and analyses of the hydrological record. In addition, we thank Ariel Rabines for help with sample processing and Josh Darling for his work on the Canada Stream hydrograph depicted in Figure 1.

Chapter one has been prepared as a submission to the ISME Journal. A. Zoumplis, B. C. Kolody, D. Kaul, H. Zheng, P. Venepally, D. M. McKnight, A. DeVries, A. E. Allen. "Rapid destabilization of nitrogen-fixing microbial mats in a high-flow Dry Valleys meltwater stream." The dissertation author was the primary investigator and author of this paper.

References

1. Adams, B. J., D. H. Wall, R. A. Virginia, E. Broos & M. A. Knox (2014) Ecological biogeography of the terrestrial nematodes of Victoria Land, Antarctica. *ZooKeys*, 29.

2. Alger, A. (1997) Ecological processes in a cold desert ecosystem: the abundance and species distribution of algal mats in glacial meltwater streams in Taylor Valley, Antarctica. *Occasional paper/University of Colorado*.

3. Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow & S. M. Huse (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PloS one,* 4**,** e6372.

4. Andriuzzi, W., B. Adams, J. Barrett, R. Virginia & D. Wall (2018) Observed trends of soil fauna in the Antarctic Dry Valleys: early signs of shifts predicted under climate change. *Ecology,* 99**,** 312-321.

5. Bar Dolev, M., I. Braslavsky & P. L. Davies (2016) Ice-binding proteins and their function. *Annual review of biochemistry,* 85**,** 515-542.

6. Barrett, J., R. Virginia, D. Wall, P. Doran, A. Fountain, K. Welch & W. Lyons (2008) Persistent effects of a discrete warming event on a polar desert ecosystem. *Global Change Biology,* 14**,** 2249-2261.

7. Barrett, J. E., R. A. Virginia, W. B. Lyons, D. M. McKnight, J. C. Priscu, P. T. Doran, A. G. Fountain, D. H. Wall & D. Moorhead (2007) Biogeochemical stoichiometry of Antarctic dry valley ecosystems. *Journal of Geophysical Research: Biogeosciences,* 112.

8. Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen & K. Xu (2015) Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proceedings of the National Academy of Sciences,* 112**,** 9938-9943.

9. Bielewicz, S., E. Bell, W. Kong, I. Friedberg, J. C. Priscu & R. M. Morgan-Kiss (2011) Protist diversity in a permanently ice-covered Antarctic lake during the polar night transition. *The ISME journal,* 5**,** 1559-1564.

10. Blazewicz, S. J., R. L. Barnard, R. A. Daly & M. K. Firestone (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *The ISME journal,* 7**,** 2061-2068.

11. Bolhuis, H. & L. J. Stal (2011) Analysis of bacterial and archaeal diversity in coastal microbial mats using massive parallel 16S rRNA gene tag sequencing. *The ISME journal,* 5**,** 1701-1712.

12. Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson & S. P. Holmes (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods,* 13**,** 581-583.

13. Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich & J. I. Gordon (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods,* 7**,** 335-336.

14. Cary, S. C., I. R. McDonald, J. E. Barrett & D. A. Cowan (2010) On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology,* 8**,** 129-138.

15. Chan, Y., J. D. Van Nostrand, J. Zhou, S. B. Pointing & R. L. Farrell (2013) Functional ecology of an Antarctic dry valley. *Proceedings of the National Academy of Sciences,* 110**,** 8990-8995.

16. Chua, M. J., R. L. Campen, L. Wahl, J. J. Grzymski & J. A. Mikucki (2018) Genomic and physiological characterization and description of Marinobacter gelidimuriae sp. nov., a psychrophilic, moderate halophile from Blood Falls, an Antarctic subglacial brine. *FEMS microbiology ecology,* 94**,** fiy021.

17. Coyne, K. J., A. E. Parker, C. K. Lee, J. A. Sohm, A. Kalmbach, T. Gunderson, R. León-Zayas, D. G. Capone, E. J. Carpenter & S. C. Cary (2020) The distribution and relative ecological roles of autotrophic and heterotrophic diazotrophs in the McMurdo Dry Valleys, Antarctica. *FEMS microbiology ecology,* 96**,** fiaa010.

18. Cullis, J. D., L. F. Stanish & D. M. McKnight (2014) Diel flow pulses drive particulate organic matter transport from microbial mats in a glacial meltwater stream in the McMurdo Dry Valleys. *Water Resources Research,* 50**,** 86-97.
19. Davey, M. C. & K. J. Clarke (1992) FINE STRUCTURE OF A TERRESTRIAL CYANOBACTERIAL MAT FROM ANTARCTICA 1. *Journal of phycology,* 28**,** 199-202.

20. Doran, P. T., W. B. Lyons & D. M. McKnight. 2010. *Life in Antarctic deserts and other cold dry environments: astrobiological analogs*. Cambridge University Press.

21. Doran, P. T., C. P. McKay, G. D. Clow, G. L. Dana, A. G. Fountain, T. Nylen & W. B. Lyons (2002a) Valley floor climate observations from the McMurdo Dry Valleys, Antarctica, 1986–2000. *Journal of Geophysical Research: Atmospheres,* 107**,** ACL 13-1-ACL 13-12.

22. Doran, P. T., J. C. Priscu, W. B. Lyons, J. E. Walsh, A. G. Fountain, D. M. McKnight, D. L. Moorhead, R. A. Virginia, D. H. Wall & G. D. Clow (2002b) Antarctic climate cooling and terrestrial ecosystem response. *Nature,* 415**,** 517-520.

23. Dupont, C. L., J. P. McCrow, R. Valas, A. Moustafa, N. Walworth, U. Goodenough, R. Roth, S. L. Hogle, J. Bai & Z. I. Johnson (2015) Genomes and gene expression across light

and productivity gradients in eastern subtropical Pacific microbial communities. *The ISME journal,* 9**,** 1076-1092.

24. Esposito, R., S. Horn, D. M. McKnight, M. Cox, M. Grant, S. Spaulding, P. Doran & K. Cozzetto (2006) Antarctic climate cooling and response of diatoms in glacial meltwater streams. *Geophysical Research Letters,* 33.

25. Esposito, R., S. Spaulding, D. M. McKnight, B. Van de Vijver, K. Kopalová, D. Lubinski, B. Hall & T. Whittaker (2008) Inland diatoms from the McMurdo dry valleys and James Ross Island, Antarctica. *Botany,* 86**,** 1378-1392.

26. Finn, R., J. Mistry, J. Tate, P. Coggill & A. Heger (2014) Pfam: the protein families database. Nuclei. Acids Re.

27. Finn, R. D., J. Clements & S. R. Eddy (2011) HMMER web server: interactive sequence similarity searching. *Nucleic acids research,* 39**,** W29-W37.

28. Fountain, A. G., J. S. Levy, M. N. Gooseff & D. Van Horn (2014) The McMurdo Dry Valleys: a landscape on the threshold of change. *Geomorphology,* 225**,** 25-35.

29. Fountain, A. G., T. H. Nylen, A. Monaghan, H. J. Basagic & D. Bromwich (2010) Snow in the McMurdo dry valleys, Antarctica. *International Journal of Climatology: A Journal of the Royal Meteorological Society,* 30**,** 633-642.
30. Fountain, A. G., G. Saba, B. Adams, P. Doran, W. Fraser, M. Gooseff, M. Obryk, J. C. Priscu, S. Stammerjohn & R. A. Virginia (2016) The impact of a large-scale climate event on Antarctic ecosystem processes. *Bioscience,* 66**,** 848-863.

31. Gooseff, M. N., J. E. Barrett, B. J. Adams, P. T. Doran, A. G. Fountain, W. B. Lyons, D. M. McKnight, J. C. Priscu, E. R. Sokol & C. Takacs-Vesbach (2017) Decadal ecosystem response to an anomalous melt season in a polar desert in Antarctica. *Nature ecology & evolution,* 1**,** 1334-1338.

32. Gooseff, M. N., D. M. McKnight, P. Doran, A. G. Fountain & W. B. Lyons (2011) Hydrological connectivity of the landscape of the McMurdo Dry Valleys, Antarctica. *Geography Compass,* 5**,** 666-681.

33. Gururani, M. A., J. Venkatesh & L. S. P. Tran (2015) Regulation of photosynthesis during abiotic stress-induced photoinhibition. *Molecular plant,* 8**,** 1304-1320.

34. Hawes, I. & A. M. Schwarz (2001) Absorption and utilization of irradiance by cyanobacterial mats in two ice-covered antarctic lakes with contrasting light climates. *Journal of Phycology,* 37**,** 5-15.

35. Hopkins, D., A. Sparrow, B. Elberling, E. Gregorich, P. Novis, L. Greenfield & E. Tilston (2006) Carbon, nitrogen and temperature controls on microbial activity in soils from an Antarctic dry valley. *Soil Biology and Biochemistry,* 38**,** 3130-3140.

36. Howard-Williams, C., J. C. Priscu & W. F. Vincent (1989) Nitrogen dynamics in two Antarctic streams. *Hydrobiologia,* 172**,** 51-61.

37. Howard-Williams, C., C. L. Vincent, P. A. Broady & W. F. Vincent (1986) Antarctic stream ecosystems: variability in environmental properties and algal community structure. *Internationale Revue der Gesamten Hydrobiologie und Hydrographie,* 71**,** 511-544.

38. Jungblut, A. D., I. Hawes, D. Mountfort, B. Hitzfeld, D. R. Dietrich, B. P. Burns & B. A. Neilan (2005) Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environmental microbiology,* 7**,** 519-529.

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

40. Knight, C. & A. DeVries (2009) Ice growth in supercooled solutions of a biological "antifreeze", AFGP 1–5: An explanation in terms of adsorption rate for the concentration dependence of the freezing point. *Physical Chemistry Chemical Physics,* 11**,** 5749-5761.

41. Kohler, T. J., L. F. Stanish, S. W. Crisp, J. C. Koch, D. Liptzin, J. L. Baeseman & D. M. McKnight (2015) Life in the main channel: long-term hydrologic control of microbial mat abundance in McMurdo Dry Valley streams, Antarctica. *Ecosystems,* 18**,** 310-327.

42. Kohler, T. J., L. F. Stanish, D. Liptzin, J. E. Barrett & D. M. McKnight (2018) Catch and release: Hyporheic retention and mineralization of N-fixing Nostoc sustains downstream microbial mat biomass in two polar desert streams. *Limnology and Oceanography Letters,* 3**,** 357-364.

43. Kolody, B., J. McCrow, L. Z. Allen, F. Aylward, K. Fontanez, A. Moustafa, M. Moniruzzaman, F. Chavez, C. Scholin & E. Allen (2019) Diel transcriptional response of a California Current plankton microbiome to light, low iron, and enduring viral infection. *The ISME journal,* 13**,** 2817-2833.

44. Králová, S. (2017) Role of fatty acids in cold adaptation of Antarctic psychrophilic Flavobacterium spp. *Systematic and applied microbiology,* 40**,** 329-333.

45. Kubota, N. (2011) Effects of cooling rate, annealing time and biological antifreeze concentration on thermal hysteresis reading. *Cryobiology,* 63**,** 198-209.

46. Langfelder, P. & S. Horvath (2008) WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics,* 9**,** 1-13.

47. Levy, J. S., A. G. Fountain, M. Obryk, J. Telling, C. Glennie, R. Pettersson, M. Gooseff & D. Van Horn (2018) Decadal topographic change in the McMurdo Dry Valleys of Antarctica: Thermokarst subsidence, glacier thinning, and transfer of water storage from the cryosphere to the hydrosphere. *Geomorphology,* 323**,** 80-97.

48. Marizcurrena, J. J., M. F. Cerdá, D. Alem & S. Castro-Sowinski. 2019. Living with pigments: the colour palette of Antarctic life. In *The ecological role of micro-organisms in the antarctic environment*, 65-82. Springer.

49. McKnight, D. M., D. K. Niyogi, A. S. Alger, A. Bomblies, P. A. Conovitz & C. M. Tate (1999) Dry valley streams in Antarctica: ecosystems waiting for water. *Bioscience,* 49**,** 985-995.

50. McKnight, D. M., R. L. Runkel, C. M. Tate, J. H. Duff & D. L. Moorhead (2004) Inorganic N and P dynamics of Antarctic glacial meltwater streams as controlled by hyporheic exchange and benthic autotrophic communities. *Journal of the North American Benthological Society,* 23**,** 171-188.

51. McKnight, D. M. & C. Tate (1997) Canada stream: a glacial meltwater stream in Taylor Valley, south Victoria Land, Antarctica. *Journal of the North American Benthological Society,* 16**,** 14-17.

52. McKnight, D. M., C. Tate, E. Andrews, D. Niyogi, K. Cozzetto, K. Welch, W. Lyons & D. Capone (2007) Reactivation of a cryptobiotic stream ecosystem in the McMurdo Dry Valleys, Antarctica: a long-term geomorphological experiment. *Geomorphology,* 89**,** 186-204.

53. Meng, J., B. Hu, G. Yi, X. Li, H. Chen, Y. Wang, W. Yuan, Y. Xing, Q. Sheng & Z. Su (2020) Genome-wide analyses of banana fasciclin-like AGP genes and their differential expression under low-temperature stress in chilling sensitive and tolerant cultivars. *Plant cell reports,* 39**,** 693-708.

54. Morgan-Kiss, R., M. Lizotte, W. Kong & J. Priscu (2016) Photoadaptation to the polar night by phytoplankton in a permanently ice-covered Antarctic lake. *Limnology and Oceanography,* 61**,** 3-13.

55. Murata, N., S. Takahashi, Y. Nishiyama & S. I. Allakhverdiev (2007) Photoinhibition of photosystem II under environmental stress. *Biochimica et Biophysica Acta (BBA)-Bioenergetics,* 1767**,** 414-421.

56. Nielsen, U. N., D. H. Wall, B. J. Adams, R. A. Virginia, B. A. Ball, M. N. Gooseff & D. M. McKnight (2012) The ecology of pulse events: insights from an extreme climatic event in a polar desert ecosystem. *Ecosphere,* 3**,** 1-15.

57. Obryk, M. K., P. T. Doran, A. G. Fountain, M. Myers & C. P. McKay (2020) Climate from the McMurdo dry valleys, Antarctica, 1986–2017: Surface air temperature trends and redefined summer season. *Journal of Geophysical Research: Atmospheres,* 125**,** e2019JD032180.

58. Parada, A. E., D. M. Needham & J. A. Fuhrman (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental microbiology,* 18**,** 1403-1414.

59. Rai, R., S. Singh, A. Chatterjee, K. K. Rai, S. Rai & L. Rai (2020) All4894 encoding a novel fasciclin (FAS-1 domain) protein of Anabaena sp. PCC7120 revealed the presence of a thermostable β-glucosidase. *Algal Research,* 51**,** 102036.

60. Raymond-Bouchard, I. & L. G. Whyte. 2017. From transcriptomes to metatranscriptomes: cold adaptation and active metabolisms of psychrophiles from cold environments. In *Psychrophiles: from biodiversity to biotechnology*, 437-457. Springer.

61. Rho, M., H. Tang & Y. Ye (2010) FragGeneScan: predicting genes in short and error-prone reads. *Nucleic acids research,* 38**,** e191-e191.

62. Robinson, M. D., D. J. McCarthy & G. K. Smyth (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-140.

63. Schmieder, R., Y. W. Lim & R. Edwards (2012) Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics,* 28**,** 433-435.

64. Seifert, G. J. (2018) Fascinating fasciclins: a surprisingly widespread family of proteins that mediate interactions between the cell exterior and the cell surface. *International journal of molecular sciences,* 19**,** 1628.

65. Singley, J. G., M. N. Gooseff, D. M. McKnight & E. Hinckley (2021) The Role of Hyporheic Connectivity in Determining Nitrogen Availability: Insights from an Intermittent Antarctic Stream. *Journal of Geophysical Research: Biogeosciences,* 126**,** e2021JG006309.

66. Sommers, P., J. L. Darcy, D. L. Porazinska, E. Gendron, A. G. Fountain, F. Zamora, K. Vincent, K. M. Cawley, A. J. Solon & L. Vimercati (2019) Comparison of microbial communities in the sediments and water columns of frozen cryoconite holes in the McMurdo Dry Valleys, Antarctica. *Frontiers in microbiology,* 10**,** 65.

67. Sorokovikova, E. G., O. I. Belykh, A. S. Gladkikh, O. V. Kotsar, I. V. Tikhonova, O. A. Timoshkin & V. V. Parfenova (2013) Diversity of cyanobacterial species and phylotypes in biofilms from the littoral zone of Lake Baikal. *Journal of Microbiology,* 51**,** 757-765.

68. Stanish, L. F., D. R. Nemergut & D. M. McKnight (2011) Hydrologic processes influence diatom community composition in Dry Valley streams. *Journal of the North American Benthological Society,* 30**,** 1057-1073.

69. Stoeck, T., D. Bass, M. Nebel, R. Christen, M. D. Jones, H. W. BREINER & T. A. Richards (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular ecology,* 19**,** 21-31.

70. Takamichi, M., Y. Nishimiya, A. Miura & S. Tsuda (2007) Effect of annealing time of an ice crystal on the activity of type III antifreeze protein. *The FEBS journal,* 274**,** 6469-6476.

71. Toner, J. D., R. S. Sletten & M. L. Prentice (2013) Soluble salt accumulations in Taylor Valley, Antarctica: implications for paleolakes and Ross Sea Ice Sheet dynamics. *Journal of Geophysical Research: Earth Surface,* 118**,** 198-215.

72. Van Horn, D. J., C. R. Wolf, D. R. Colman, X. Jiang, T. J. Kohler, D. M. McKnight, L. F. Stanish, T. Yazzie & C. D. Takacs-Vesbach (2016) Patterns of bacterial biodiversity in the glacial meltwater streams of the McMurdo Dry Valleys, Antarctica. *FEMS Microbiology Ecology,* 92**,** fiw148.

73. Vance, T. D., M. Bayer-Giraldi, P. L. Davies & M. Mangiagalli (2019) Ice-binding proteins and the 'domain of unknown function'3494 family. *The FEBS journal,* 286**,** 855-873.

74. Varin, T., C. Lovejoy, A. D. Jungblut, W. F. Vincent & J. Corbeil (2012) Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. *Applied and environmental microbiology,* 78**,** 549-559.

75. Vick, T. J. & J. C. Priscu (2012) Bacterioplankton productivity in lakes of the Taylor Valley, Antarctica, during the polar night transition. *Aquatic microbial ecology,* 68**,** 77-90.

76. Vick-Majors, T. J., J. C. Priscu & L. A. Amaral-Zettler (2014) Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. *The ISME journal,* 8**,** 778-789.

77. Vincent, W., M. Downes, R. Castenholz & C. Howard-Williams (1993) Community structure and pigment organisation of cyanobacteria-dominated microbial mats in Antarctica. *European Journal of Phycology,* 28**,** 213-221.

78. Vincent, W. F. 2000. Cyanobacterial dominance in the polar regions. In *The ecology of cyanobacteria*, 321-340. Springer.

79. Wharton Jr, R. A., B. C. Parker & G. M. Simmons Jr (1983) Distribution, species composition and morphology of algal mats in Antarctic dry valley lakes. *Phycologia,* 22**,** 355-365.

80. Wlostowski, A. N., M. N. Gooseff, D. M. McKnight, C. Jaros & W. B. Lyons (2016) Patterns of hydrologic connectivity in the McMurdo Dry Valleys, Antarctica: a synthesis of 20 years of hydrologic data. *Hydrological Processes,* 30**,** 2958-2975.

Figures

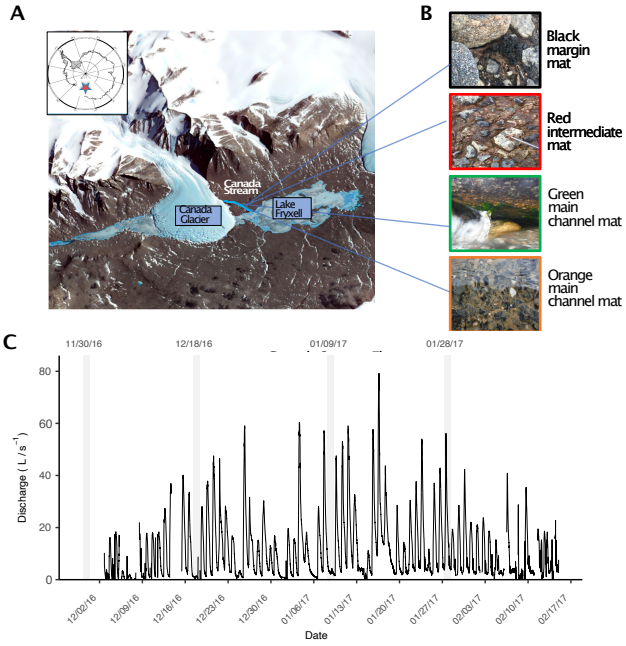

Figure 1.1 Sampling site and stream features. **(A)** Map of Lake Fryxell Basin, Taylor Valley, Antarctica showing placement of the Canada meltwater stream. (**B**) Visuals of the four mat types found along the Canada stream margins and within the main channel. (**C**) Seasonal flow rate (L/s) of Canada stream averaged weekly over 10 weeks.

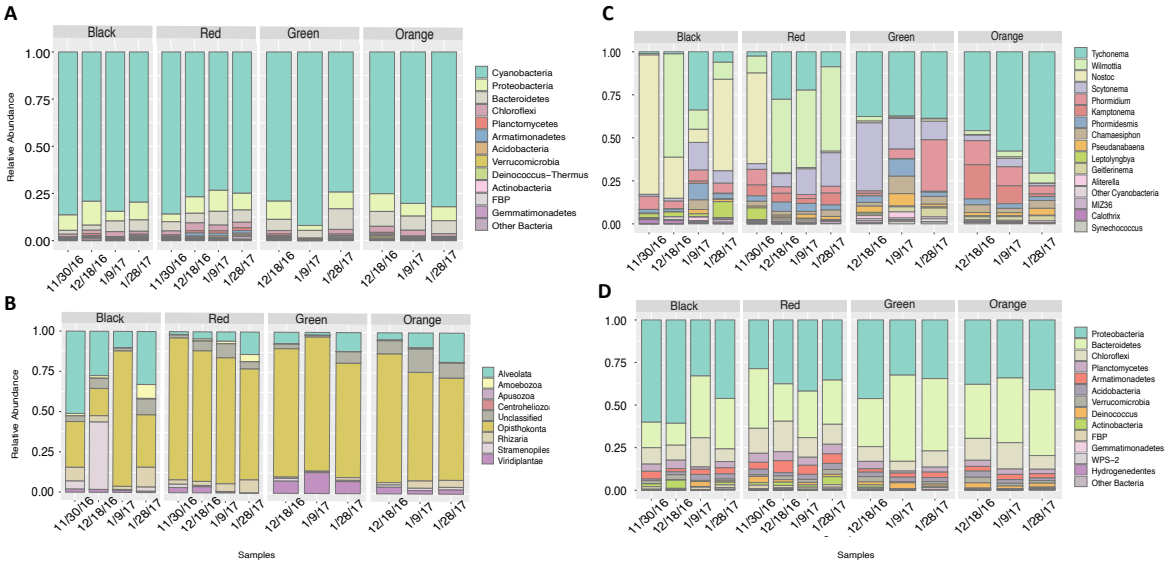

Figure 1.2 Relative abundance of (**A**) prokaryotic, (**B**) eukaryotic, (**C**) cyanobacterial, (**D**) and non-cyanobacterial community members from non-plastid 16S and 18S rRNA ASVs.

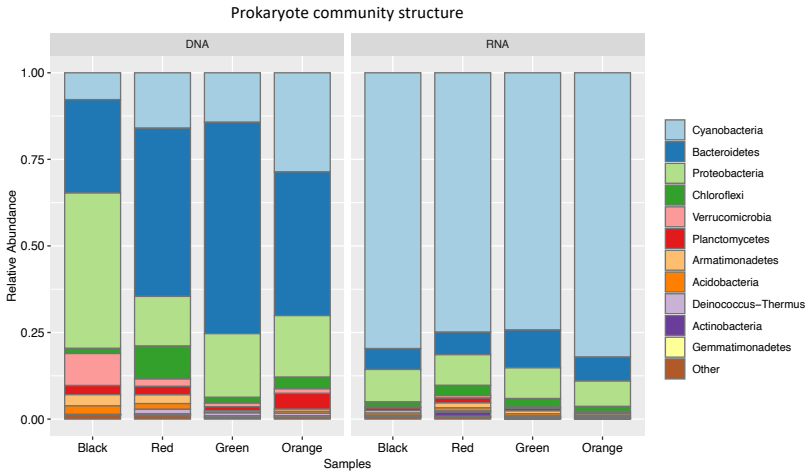

Figure 1.3 Comparison of prokaryote relative abundances from 01/28/17 black, red, green and orange mats showing differences between 16S rDNA and 16S rRNA derived communities.

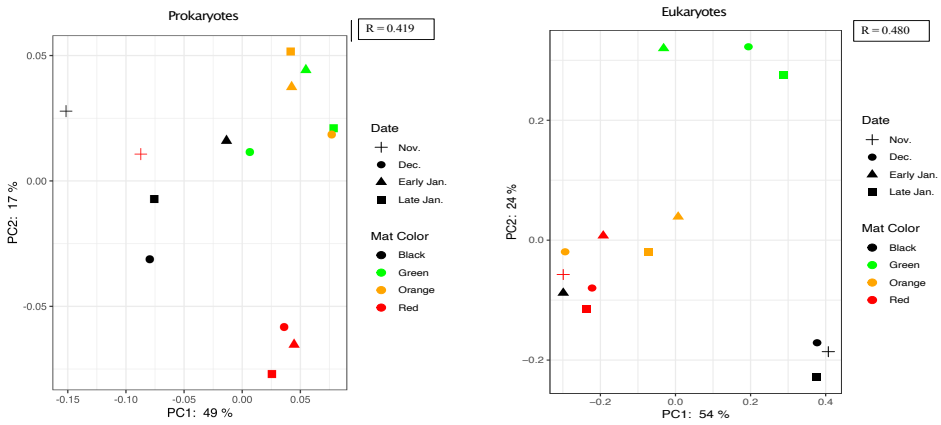

Figure 1.4 Principal coordinate analysis (PCoA) of UniFrac weighted distance matrices with ANOSIM values generated from (left)16S rRNA and (right) 18S rRNA ASVs. Shapes of plotted points denote sample collection date and colors denote mat type of origin.

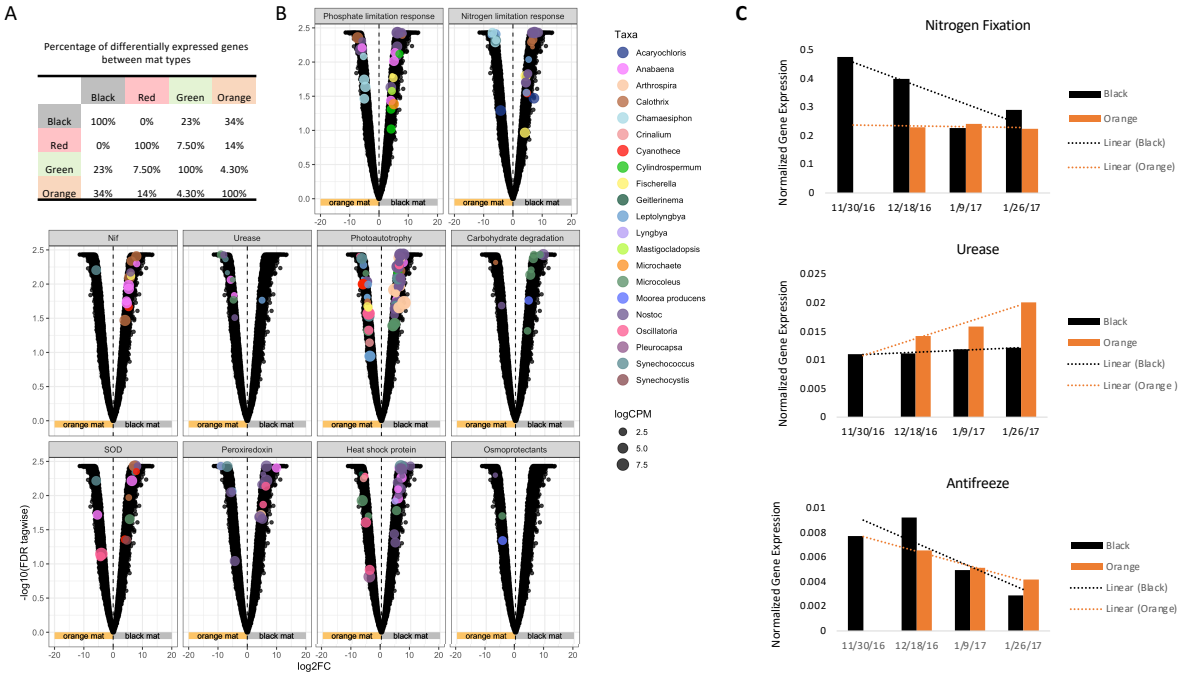

Figure 1.5 (A) Table showing the percentage of significantly differentially expressed ORFs between mat types. (**B**) Differential expression of ORFs between orange and black mat triplicates (12/18/16, 1/9/17 and 1/28/17). The following facet labels represent the collection of pathway-critical genes, referenced in Chan et al. 2013 (Chan et al. 2013), to depict functional processes: Nutrient limitation: "Phosphate limitation response", "Nitrogen limitation response". Nitrogen cycling: Nitrogen fixation "Nif", "Urease". Carbon cycling "Photoautotrophy", "Carbohydrate degradation". Stress responses: Superoxide Dismutase "SOD", "Peroxiredoxin", "Heat shock protein" and "Osmoprotectants". Significantly differentially expressed ORFs are colored by taxonomy and scaled by abundance. (**C**) Library normalized expression of (top) nitrogen fixation, (middle) urease, and (bottom) antifreeze ORFs from black and orange mats over time.

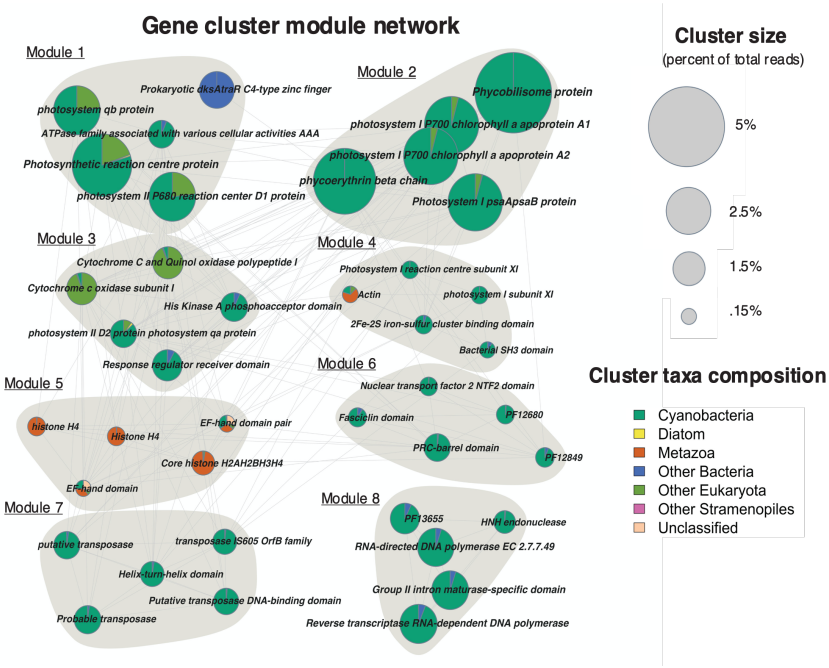

Figure 1.6 Visualization of networks of co-expression modules as determined by WGCNA. Pie charts represent highly abundant functional clusters of ORFS colored by taxonomic contribution and sized by percent of total reads.

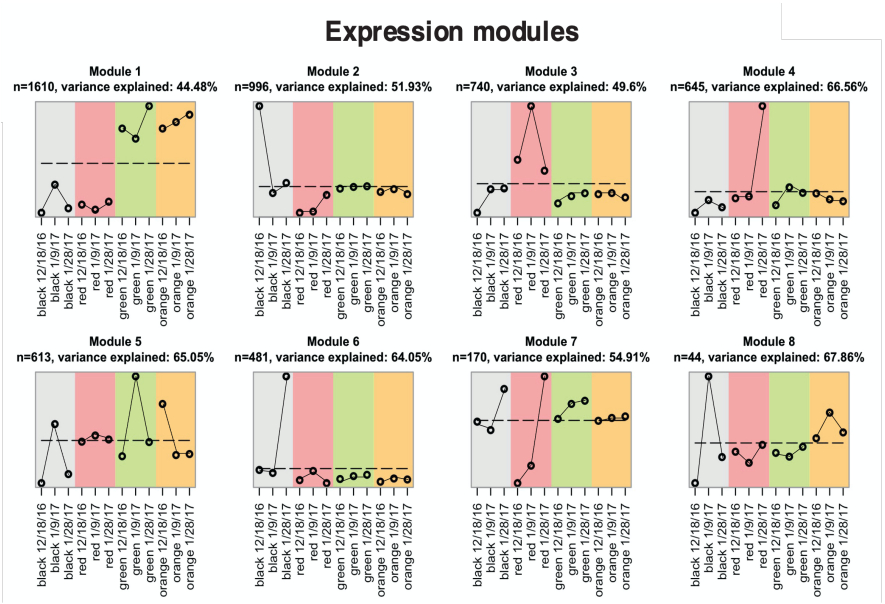

Figure 1.7 Co-expression patterns of modules (depicted in Fig. 7) colored by mat type. Subtitle describes number (n) of ORFs in each module as well as percentage of variance explained by each module's expression profile.

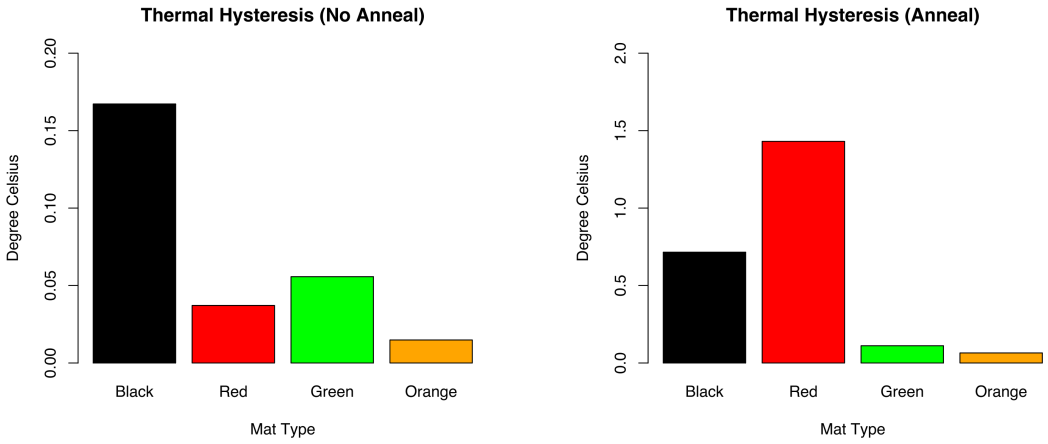

Figure 1.8 Thermal hysteresis antifreeze activities of black, red, green and orange mats with no annealing period (left) and after a thirty minute annealing period (right).

Supporting Figures

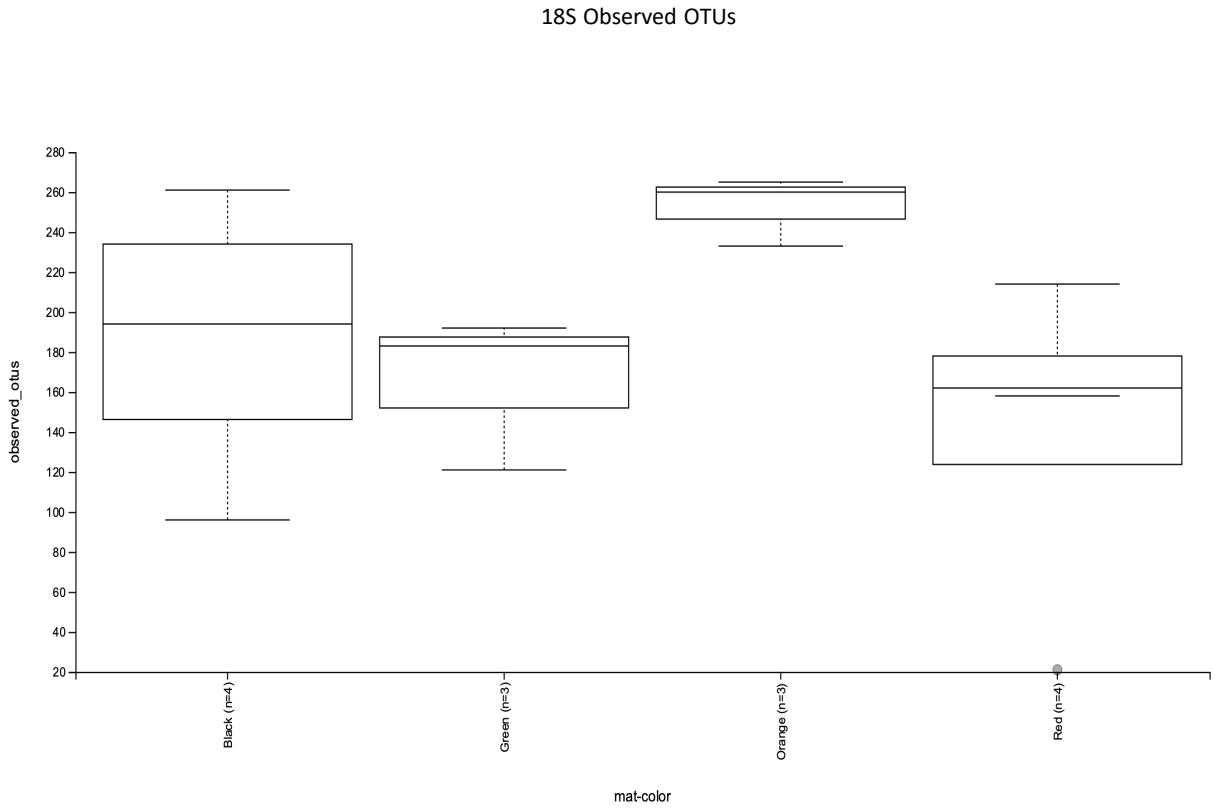

Figure S1 Boxplot of observed richness of 18S rRNA OTUs for mat samples grouped by black, green, red and orange mat types.

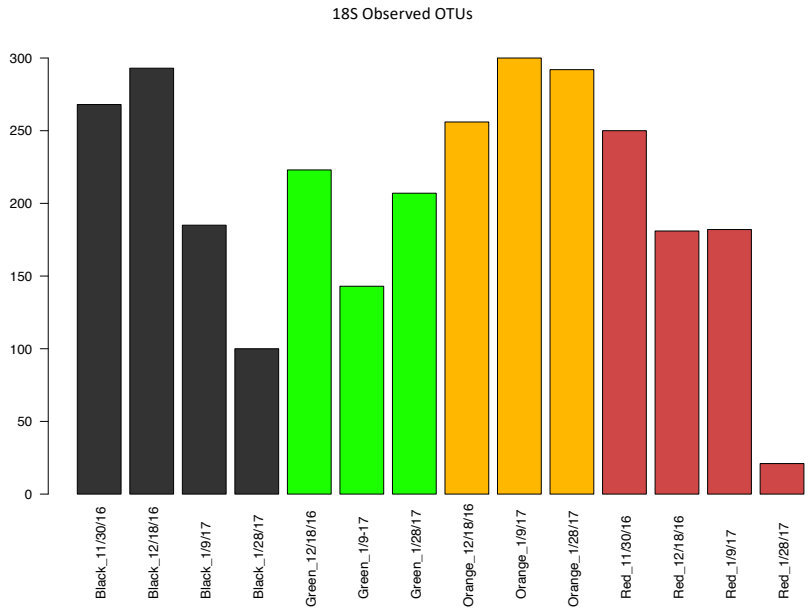

Figure S2 Bar graph of observed richness of 18S rRNA OTUs for black, green, red andorange mat individual samples.

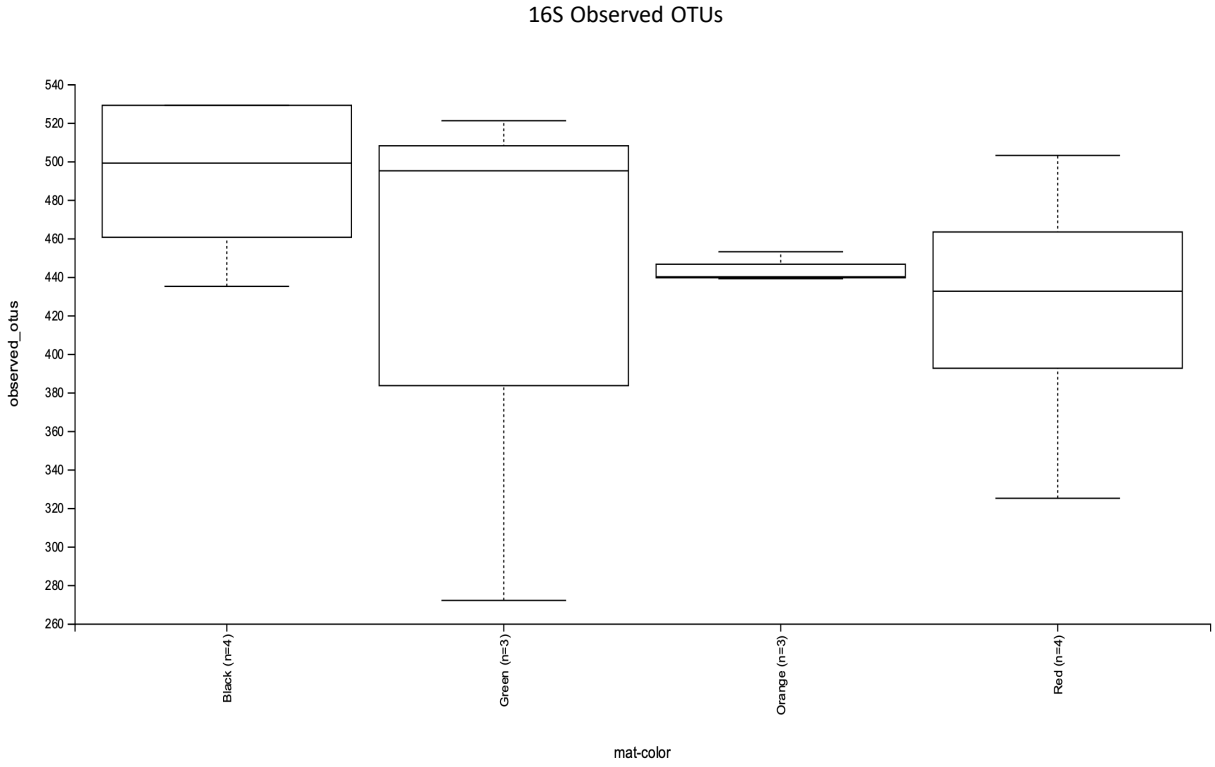

Figure S3 Boxplot of observed richness of 16S rRNA OTUs for mat samples grouped by black, green, red and orange mat types.

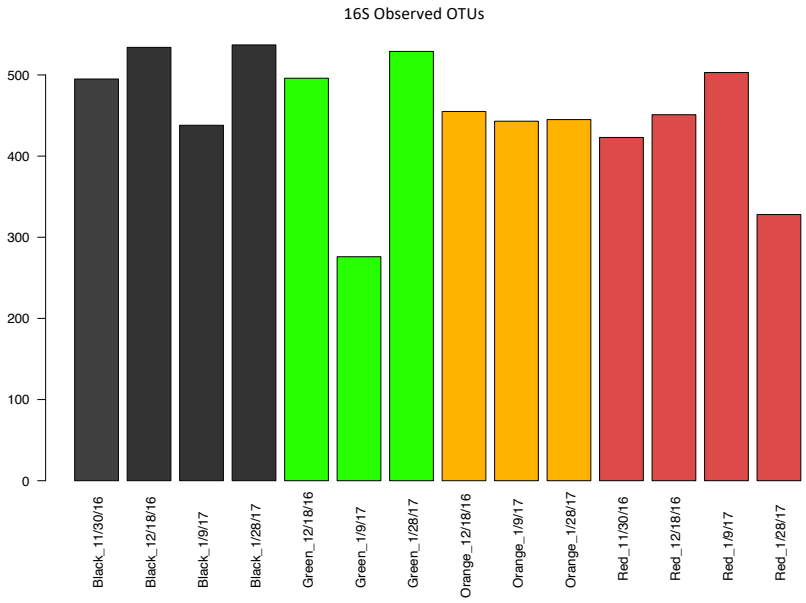

Figure S4 Bar graph of observed richness of 16S rRNA OTUs for black, green, red andorange mat individual samples.

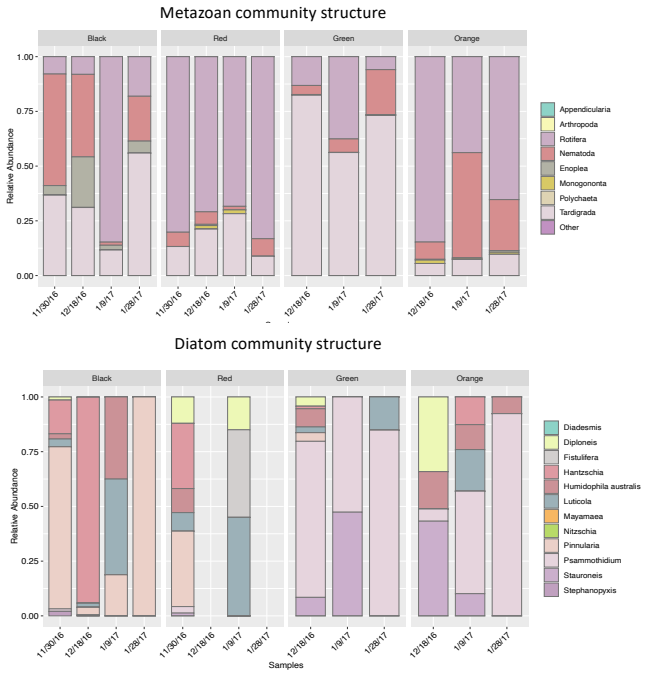

Figure S5 Relative abundance of (**top**) metazoan and (**bottom**) diatom community members derived from 18S rRNA ASVs. Diatoms were not detected in 12/18/16 or 1/28/17 red mat samples.

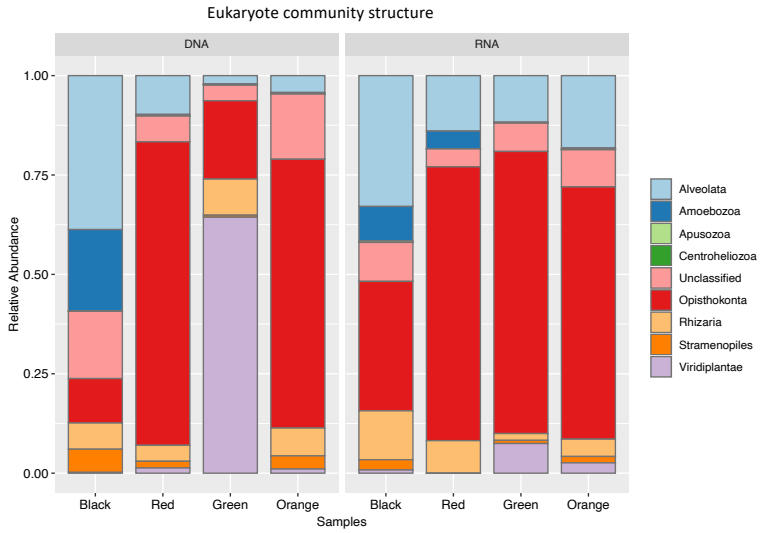

Figure S6 Comparison of eukaryote relative abundances from 01/28/17 black, red, green and orange mats showing differences between 16S rDNA and 16S rRNA derived communities.

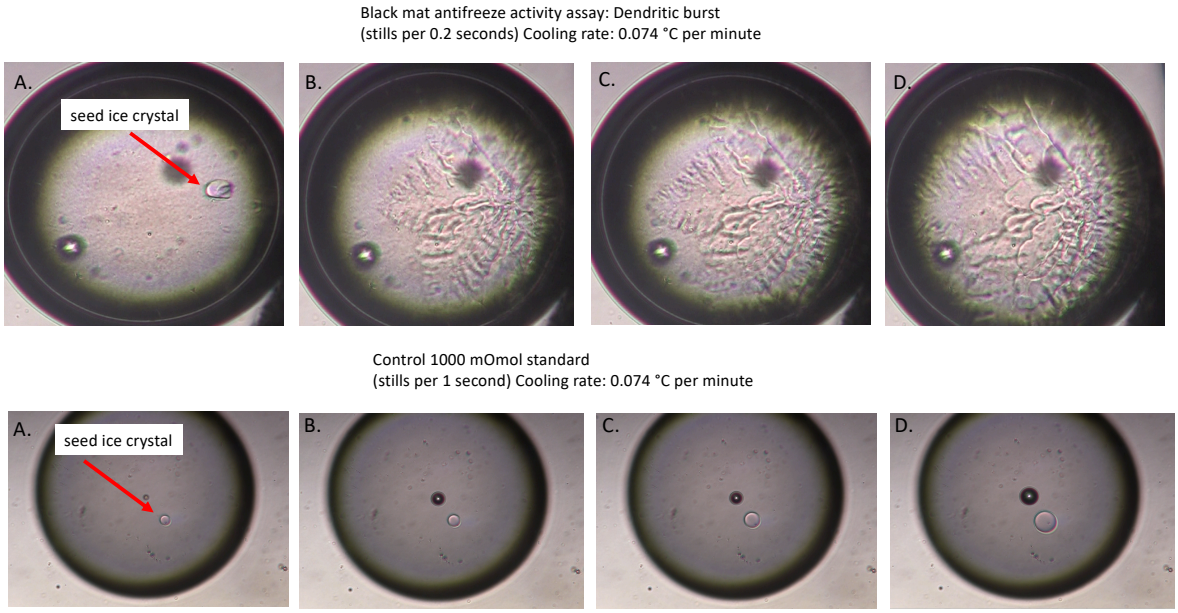

Figure S7 (top) Visualization of a dendritic burst of a seed ice crystal cooled at a rate of 0.074°C/min in a well with black mat sample lysate. Images are taken sequentially (**A**) 0.2 sec (**B**) 0.4 sec (**C**) 0.6 sec (**D**) 0.8 sec. (**bottom**) A seed ice crystal cooled at a rate of 0.074°C/min in a well with 1000 mOsmol standard. Images are taken sequentially (**A**) 1 sec (**B**) 2 sec (**C**) 3 sec (**D**) 4 sec

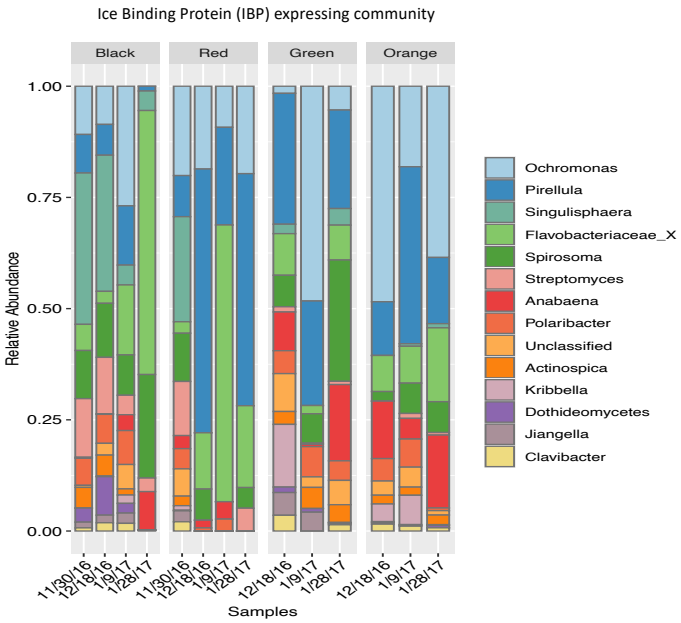

Figure S8 Relative abundance of ice-binding protein (IBP) expressing community members derived from taxonomically annotated ORFs from Pfam (PF11999).

Supporting Information File 1

16S and 18S amplification, cDNA synthesis and sequencing protocol

A.) RNA and DNA were extracted from samples in accordance with NucleoMag RNA (Macherey-Nagel) and NucleoMag Plant (Macherey-Nagel) kit protocols

B.) 1-10 ng of total RNA was used to generate cDNA using the Life Technologies. SuperScript III First Strand Synthesis system with random hexamer primers.

C.) 4 µl of each sample template was used in a 25 µl PCR reaction, which contained, and a final primer concentration of 200 nM.

D.) Primers:

a. 16S V4-V5 region
515F-Y (5′-GTGYCAGCMGCCGCGGTAA) and
926R (5′-CCGYCAATTYMTTTRAGTTT) (Parada et al. 2015)
b. 18S V4 region
F) CCAGCASCYGCGGTAATTCC
R) ACTTTCGTTCTTGATYRA modified from (Stoeck et al. 2010)

E.) PCR Reaction:

a. 5x buffer: 5ul
b. TruFi Taq DNA Polymerase: 0.25ul
c. Primer Mix: 1ul
d. Templet: 4ul
e. H20: 9.75

F.) A no template negative control for cDNA synthesis was used as a negative control for subsequent PCRs reactions.

G.) Thermocycling conditions included an initial denaturation at 95°C for 1 minutes, 30 cycles of 95°C for 15 seconds, 56°C for 15 seconds, 72°C for 30 minutes

H.) Reactions were cleaned using Ampure XP beads (Beckman Coulter, Brea CA). The final products were resuspended in 45 µL of elution buffer

I.) 1 µL was used to quantify the final product using' PicoGreen Quant-IT assay (thermoFihser) 20 ng and 30 ng of each 16S or 18S amplicon respectively were pooled separately for sequencing.

J.) 1-2 ul were used for final quantification on an Agilent TapeStation.

Metatranscriptomics amplification, cDNA synthesis and sequencing protocol

A.) 150ng of total RNA as input, ribosomal RNA was removed using Ribo-Zero Magnetic kits (Illumina).

B.)The rRNA-deplete total RNA was used for cDNA synthesis by Ovation RNA-Seq System V2 (TECAN, Redwood City, USA).

C.) Double stranded cDNA was fragmented using Covaries E210 system with the target size of 400bp.

D.) 100ng of fragmented cDNA as input into the Ovation Ultralow System V2 (TECAN, Redwood City, USA), following the manufactures protocol.

E.) Ampure XP beads (Beckman Coulter) were used for final library purification.

F.) Library quality was analyzed on a 2200 TapeStation System with Agilent High Sensitivity DNA 1000 ScreenTape System (Agilent Technologies, Santa Clara, CA, USA).

An intrusive marine eukaryotic assemblage in the Antarctic Dry Valleys

A. Zoumplis[1,2], D. Kaul[2], N. Schulte[3], H. Zheng[2], P. Venepally[2], D. M. McKnight[3], A. E. Allen[1,2]

1. Scripps Institution of Oceanography, University of California, San Diego, CA 92093
2. Microbial and Environmental Genomics Group, J. Craig Venter Institute, La Jolla, CA 92037
3. Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, 80309

Abstract

The McMurdo Dry Valleys (MDVs) region is highly sensitive to climate warming and undergoes rapid ecological response to shifts in temperature and hydrological disturbance which in turn can produce lasting changes in microbial community structure (Freckman and Virginia 1997, Stanish et al. 2012, Monteiro et al. 2020). Geochemical evidence suggests past warm intervals produced a rise in global average temperature and sea level that contributed to substantial flooding within the Antarctic valley network (Denton et al. 1993, Lyons et al. 2005, Bao et al. 2008). These floods would have persisted for millions of years before retreating (Denton et al. 1993, Berry Lyons, Frape and Welch 1999). The goal of this study was to assess the MDVs for evidence of microbial remnants of marine incursions. A survey of molecular diversity was

conducted with over 160 aquatic, sediment and aeolian samples collected across freshwater, marine and terrestrial habitats around Victoria Land, Antarctica. Results of this survey reveal a previously unknown, dominant, marine eukaryotic community surrounding the Taylor Glacier terminus (~30km inland of the McMurdo Sound). Marine indicator species were identified across several taxa groups including diatoms, haptophytes, dinoflagellates and ciliates in red hued ice, mud and sediment samples from the Taylor glacier site. Metatranscriptomics analyses on the intrusive community depicts an active, metabolizing community with upregulated functions in cell motility, signaling, stress resistance and morphological transformation. These findings shed light on the mechanisms of intrusive microbiota survivability and their suitability in osmotically variable habitats.

Introduction

Diatoms are unicellular microalgae with intricate, silica based cell walls. They account for ~ 40% of carbon export in ocean ecosystems; a number that is thought to be underestimated as a result of species variability (Tréguer et al. 2018, Pierella Karlusich, Bowler and Biswas 2021). Life cycle, morphological features, elemental composition and metabolic activity of these phytoplankton are all critical factors in understanding CO_2 sequestration which can inform climate models (Wilson et al. 2018, Tréguer et al. 2018, Irion et al. 2021). Beyond their known importance in primary productivity and biogeochemical flux, diatoms have long been recognized as indicators of global environmental change (Dixit et al. 1992). There are more than 100,000 species of diatoms making them among the most diverse algae in the world (Kooistra et al. 2003). The presence or absence of certain diatoms can be suggestive of terrain type as most genera are confined to either marine or freshwater habitats, though transitions across salinity barriers due occur (Sims, Mann

and Medlin 2006, Alverson, Jansen and Theriot 2007). Due to their sensitivity in response to environmental pressures, strong habitat preferences, and relatively short generation times, diatom communities and fossil assemblages have been used reanimate past climate events and biological responses to these events (Taylor, Whitehead and Domack 2001, Matul et al. 2018, Kato 2020, Pinseel et al. 2021). The remote MDVs hosts a simplified food web with limited top-down and grazing controls (James, Hall and Laybourn-Parry 1998, Virginia and Wall 1999). Thus, polar desert ecosystems are an ideal environment to observe physical, chemical and climatic drivers of community change.

Metabolically active diatoms constitute a major component of the microbial communities found in multiple landscape features within the MDVs including transient meltwater streams (Wharton Jr et al. 1983, Esposito et al. 2008, Kohler et al. 2015, Sommers et al. 2019). These meltwater streams constitute the most biodiverse habitat of the MDVs (Van Horn et al. 2016). Since 1993, the McMurdo Dry Valleys Long-Term Ecological Research program (MCM-LTER) has collected biodiversity, chemistry and flow measurements from over 35 ephemeral streams. Parallel environmental metadata is also available from meteorological stations set up around the valley floor and atop glaciers. These maintained, publicly available datasets include records from a site referred to as the “red river” which is more commonly known in literature as Blood Falls.

Unlike most streams in the MDVs, Blood Falls is primarily sourced through a conduit leading to a subglacial brine beneath the Taylor Glacier (Badgeley et al. 2017). Recent studies have indicated the brine system is more extensive than previously thought, extending ~5.7 km eastward of the Taylor Glacier (Mikucki et al. 2015). The entrapped iron-rich brine periodically flows out of the Taylor Glacier displaying a characteristic red hue (Mikucki et al. 2004). Major ion chemistry and isotopic composition analysis have provided evidence for an ancient, seawater

origin of the englacial brine (Lyons et al. 2019). Further support for the marine source theory has come from the discovery of bacteria in the extensive brine system that share close 16S rRNA gene identity to marine bacteria (Mikucki et al. 2004, Mikucki and Priscu 2007, Campen et al. 2019). While no eukaryotes have been detected in the brine itself, the surrounding area of the Taylor Glacier terminus (TGT) is of key site of interest for relic, marine life as the high salt and nutrient content and periodic stream flow creates a harsh, but feasible habitat (Mikucki et al. 2009, Campen et al. 2019). Despite the implications for such a community, little is known about the eukaryotic diversity surrounding this area.

Diatom frustule depositions from Antarctic marine taxa have been discovered in several non-marine locations across the continent leading to debates on the stability and dynamics of major glaciers and ice sheets over past warming periods (Webb et al. 1984, Kellogg and Kellogg 1996, Scherer et al. 2016). The focal point of these debates has been the origin of emplacement of marine fossils and whether wind or oceanic/subglacial processes are responsible for transport (Stroeven and Prentice 1997, Scherer et al. 2004). The taxa constituting these diatomaceous sediments have largely been identified by morphology, none of which have shown an active presence in the deposit sites.

This investigation on the molecular characterization of active eukaryotic diversity of the TGT site and the surrounding MDVs aquatic and aeolian and the McMurdo Sound (MCMS) systems has led to the discovery of an active and likely relic eukaryotic marine community. The marine eukaryotic assemblage is inclusive of diatoms, haptophytes, dinoflagellates, ciliates and copepods. This site, which is host to previously hidden biodiversity and endemism in the polar desert and elucidates patterns of biological dispersal, divergence, and connectivity between polar desert aquatic ecosystems and the adjacent marine environment. Additionally, we identify a suite

of primary metabolic and environmental response pathways characterizing the functional activities of the highly adaptive, intrusive community.

Methods

Sample collection

Samples collected from the TGT site (77.7167° S, 162.2667° E) included ice, red-hued mud, red- hued sediment, and non-red hued sediment (Figure 1.) and from MDVs lake and ephemeral stream microbial mats (109 samples). Following sterile, DNA/RNA free collection, samples from all sites were preserved for genomic analysis in RNAlater and stored in the liquid nitrogen. These samples were shipped on dry ice to maintain a temperature of -80°C. Temperature, salinity, and conductivity were measured onsite. This study additionally utilizes amplicon library datasets from samples collected previously in 2015 from the MCMS sea ice and water column (27 samples), and in 2017 from several Dry Valleys aeolian collectors (15 samples). A description of the Dry Valleys aeolian material and MCMS collected data and metadata used in the experiments below has been published (Schulte et al. 2021) and (Bertrand et al. 2015) respectively. The MDVs samples extend across four valleys containing a distance gradient from the MCMS (Fig 1.).

Environmental sequencing and bioinformatics

RNA was extracted from frozen, RNAlater preserved samples using the protocol from the NucleoMag RNA (Macherey-Nagel) kits. 1-10 ng of total RNA was used to generate cDNA using the Life Technologies. SuperScript III First Strand Synthesis system with random hexamer primers. The 16S V4-V5 region F (5′-GTGYCAGCMGCCGCGGTAA) R (5′-CCGYCAATTYMTTTRAGTTT) (Parada et al. 2016) and 18S V4 region F (5′-

CCAGCASCYGCGGTAATTCC) R (5′-ACTTTCGTTCTTGATYRA (Stoeck et al. 2010) were targeted for amplicon sequencing. 4 µl of each sample was used in a 25 µl PCR reaction which additionally included 5ul of 5X buffer, 0.25ul of TruFi Taq DNA Polymerase, 1ul of 200nM Primer Mix and 9.75ul of H2O. Thermocycling conditions were as follows: 95°C for 1 minutes, 30 cycles of 95°C for 15 seconds, 56°C for 15 seconds, 72°C for 30 minutes. Ampure XP beads (Beckman Coulter, Brea CA) were used to clean reaction products. After quantification, 20 ng and 30 ng of each 16S or 18S amplicon, respectively, were pooled separately for sequencing. Amplicon libraries underwent sequencing using the Illumina MiSeq platform (2 x 300bp) generating a total of ~22M reads with an average number of 90,000 reads per sample for both 16S and 18S datasets. Sequences were filtered for quality, followed by the generation of amplicon sequence variants (ASVs) from the DADA2 (Callahan et al. 2016) module in QIIME2 (version 2019.4) (Caporaso et al. 2010). Default thresholds were used for expected error and length truncation parameters: 250 bp for forward truncation (--p-trunc-len-f) and 200 bp for reverse truncation (--p-trunc-len-r).

The 16S rRNA amplicon sequence variants (ASVs) were taxonomically classified against the SILVA database (version 132), while the 18S rRNA ASVs were annotated using NCBI-NT/NR. Relative abundance values were computed for all ASVs and taxonomy bar plots were created using ggplot2 (version 3.3.5). Alpha and beta diversity analyses were performed for metrics including “shannon” and “observed_otus” and “weighted_unifrac” and “unweighted_unifrac”, respectively. Principle coordinates were calculated and subsequently imported into R (version 4.1.2) with qiime2R (version 0.99.6) and visualized using ggplot2.

Metatranscriptomics

Extracted total RNA was depleted of rRNA for mRNA enrichment. Following cDNA synthesis, metatranscriptomic libraries were sequenced on the Illumina HiSeq 4000 platform. Sequencing data was filtered and trimmed for removal of primers, adapters, and poor quality sequences. Ribopicker (version 0.4.3) (Schmieder et al. 2012) was used to identify and remove remaining rRNA. Transcripts were assembled into contigs with CLC Genomics (version 10.1.1) and open reading frames (ORFs) were predicted with FragGeneScan (version 1.31) (Rho et al. 2010). The following databases were used for the de novo functional annotation and taxonomic annotations of ORFs; the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al. 2012), Pfam (Finn et al. 2014) and hidden Markov model (HMM) searches (Finn et al. 2011), and the in-house comprehensive reference database, PhyloDB (version 1.076).

Species indicator analysis

We identified eukaryotic ASVs that were associated with each habitat or combination of two habitats using indicator species analysis of ASV relative abundance (Dufrêne and Legendre 1997, Cáceres and Legendre 2009). Indicator species analysis generates an indicator value for each ASV, which is the product of ASV specificity (probability that the sample belongs to the sample group given the relative abundance of the ASV) and sensitivity (probability of detecting the ASV in samples belonging to the sample group). Indicators for habitat combinations were calculated for hierarchically organized groups and reflect ASVs with similar relative abundances between two habitats and not others (De Cáceres, Legendre and Moretti 2010). A minimum indicator value of 0.5 was accepted, ensuring that the probability of a least one of specificity or sensitivity was > 0.7 for a given indicaor . Indicator values near 1 represent ASVs with the highest association with the habitat or habitat combination. Statistical significance of indicator values was determined from

999 permutations of samples across habitat types, and only ASVs with $p < 0.05$ were retained. Indicator species analysis was conducted using the function 'multipatt' in the R package indicspecies v.1.7.9. Furthermore, a four way Venn diagram was created (http://bioinformatics.psb.ugent.be/webtools/Venn/) to depict overlap in the number of unique and shared ASVs between the following sample sites: MCMS; TGT; MDV aquatic; MDV aeolian. A chi squared test was performed to test the significance of ASVs distribution differences between environments.

Results

A total of 7077 eukaryotic, 18S V4 region ASVs were generated across 168 samples (18 TGT, 14 Dry Valleys aeolian, 109 Dry Valleys aquatic, 27 MCMS). The dominant eukaryotic community member differed across the environmental types with Stramenopiles (41.5%), Opisthokonta (49.8%) and Viridiplantae (30.1%) dominating the MCMS, Dry Valleys aquatic, and Dry Valleys aeolian habitats, respectively (Fig 2.). Across samples, Viridiplantae was most abundant in the TGT samples with the exception of the red mud samples which had the highest proportion of Unclassified (31.1%) ASVs followed by Alveolata (23.9%) (Fig 2.). The red mud samples, along with the Dry Valleys aeolian samples, were unique in major eukaryotic community structure displaying a significant decrease in Opisthokonta in comparison to the Dry Valleys aquatic samples (p-value .007) (Fig 2.). A visible but less significant increase in Alveolata (p-value 0.092) and Haptophyceae (p-value 0.099) is apparent in the red mud samples over the Dry Valleys aquatic samples (Fig 2.). In addition, a slight increase in Cryptophyta is detected within the MCMS and TGT red mud samples (Fig 2.).

A further breakdown of Alveolata and Bacillariophyta groups detects unique taxa in the TGT environment and additional similarities to the MCMS community (Fig 3.). The dinoflagellate, *Gymnodiniaceae* is highly abundant in the aforementioned environments comprising ~40% and ~30% of the TGT and marine communities, respectively (Fig 3.). Its presence is minimal in the Dry Valleys aeolian samples (~0.03%) and undetectable in the Dry Valleys aquatic environment (Fig 3.). The marine dinoflagellate, *Protoperidiniaceae* as well as other uncultured marine dinoflagellates and marine ciliates follow this pattern of biogeographical distribution (Fig 3.).

The Bacillariophyta community presents the most striking difference in taxa between the polar samples. The MCMS diatom community consists of marine species from the *Nitzschia, Amphiprora, Fragilariopsis, Pseudo-nitzschia, Navicula*, and *Chaetoceros* genera (Fig 3.). Likewise, the dominant Dry Valleys aquatic community is comprised of Dry Valleys endemic diatom species within the following genera: *Hantzschia, Luticola, Humidophile, Diploneis, Stauroneis* and *Psammothidium* (Fig 3.). The dominant taxa at the TGT features a combination of marine and Dry Valleys endemic diatoms including *Hantzschia* (~45%), *Pseudo-nitzschia* (~18%), *Fragilariopsis* (11%), and *Guinardia* (8%) (Fig 3.). The top 80% of classified species in red mud samples collected at the Taylor Valley terminus is exclusively composed of marine diatoms with *Hantzschia* (3.2%) and *Fistulifera* (.70%) in the minor community (Fig 3.). Similarly, the red sediment samples host 60% of marine diatom taxa groups *Guinardia* and *Pseudo-nitzschia* in the dominant community (Fig 3.). The ice, and surrounding sediments of this location contain the largest proportion of Dry Valleys endemic taxa, *Hantzschia* (~80%), while the marine taxa are in the minor fraction (Fig 3.). The Dry Valleys aeolian samples, comparably, contain a majority of the Dry Valleys endemic taxa (~90%), with minor marine components (Fig 3.).

Both weighted and unweighted Unifrac Principle coordinate analysis (PCoA) plots were generated for the Bacillariophyta group (Fig 4.). Results depict a strong clustering of samples, colored by environment, between the TGT and MCMS communities, and the Dry Valleys aquatic and Dry Valleys aeolian communities (Fig 4.). The weighted Unifrac PCoA plot identifies minimally overlapping 95% confidence ellipses between the Dry Valleys aquatic and marine environments with no ellipse overlap between the aeolian and marine spaces (Fig 4.). Several samples of Dry Valleys aquatic and aeolian environments cluster in the marine environment space (Fig 4.). The unweighted Unifrac PCoA plot shows no overlap of Dry Valleys aquatic or aeolian ellipses and fewer samples breaching into the marine cluster (Fig 4.).

A Venn diagram was created to show the relationships of ASVs between environments (Fig 5.). This analysis revealed the highest proportion of shared MCMS ASVs to be with the TGT environment (Fig 5.). There is a significant difference (pvalue < 0.03) of shared marine and TGT ASVs (9.34% of total TGT ASVs) over the shared marine and Dry Valleys aquatic ASVs (1.15% of total Dry Valleys aquatic ASVs) (Fig 5.). Unique diatom ASVs from the TGT sample include the following marine species: *Nitzschia stellata, Pseudo-nitzschia australis, Actinocyclus actinochilus, Guinardia solstherfothii, Thalassionema nitzschioides,* and *Cerataulina pelagica*. The Dry Valleys aquatic samples contained two unique marine ASVs annotated as *Pseudo-nitzschia australis* and *Chaetoceros didymus* originating from the Lawson and Miers streams. No unique marine diatom ASVs were found in the Dry Valleys aeolian samples. Several ASVs from the Bacillariophyta, Ciliophora, Dinozoa, Haptophyta and Metazoa divisions are significant co-indicators for both MCMS and TGT environments based on the relative abundances of these ASVs (Fig 6.). In the Metazoan group, significant ASV indicators in the TGT samples are primarily shared with the Dry Valleys aquatic samples. *Calanus pacificus*, *Pseudo-nitzschia australis,*

Acutuncus antarticus and several uncultured species have ASV indicators solely for the TGT habitat (Fig 6.).

Metatranscriptomics analysis showed a high percentage of unclassified open reading frames (ORFs) among the TGT samples (>80% of total Eukaryotic ORFs). Alveolata were the most dominant taxa in the classified mRNA community structure for the red mud, red sediment, and ice samples (Fig 7., S3., S4., S5.). In the sediment samples we see an increased proportion of Hacrobia relative to Alveolata (Fig S5.). Within the classified ORFs, genes derived from the Ferritin-like domain were the most highly expressed in the red mud samples (Fig 7.). UTP--glucose-1-phosphate uridylyltransferase, Myosin head, and GMC-oxidoreductase were the most abundant of annotated ORFs in the red sediment, ice, and sediment samples, respectively (Fig S3., S4., S5.) All samples displayed evidence for active eukaryotic metabolisms as well as cell motility, signaling and structuring processes.

Discussion

Findings in this study show dominance of marine eukaryotic lineages in the TGT region relative to the endemic Dry Valleys community, most strikingly in the diatom community. This pattern is visible, though less prominent in the prokaryotic structuring (Fig 2.). Bacteria with marine origins have been noted in Blood Falls brine samples and include the following genera *Psychrobacter*, *Marinobacter*, *Pseudomonas*, and *Acinetobacter* (Mikucki and Priscu 2007). These genera are also present in the minor community of our TGT environment samples. While the prokaryote community in our TGT samples appears to have more overlap with Dry Valleys aquatic and aeolian samples than the MCMS environment compared to the eukaryotic community

(Fig S1.), this may be due to shorter generation times, increased likelihood for aeolian dispersal and greater tolerance for variable conditions (less habitat filtering).

The TGT site hosts a unique physical environment where sediment is sporadically hydrated by iron-rich cryopreserved seawater in addition to glacial meltwater (Lyons et al. 2019). A growing collection of studies emphasizes the importance of allopatric speciation and distinct biogeographies in diversification processes of microorganisms (Hanson et al. 2012, Singer et al. 2019, Pinseel et al. 2020). The remote, isolated nature of the Dry Valleys separated by the Transantarctic mountains and limited aeolian dispersal (Schulte et al. 2021) creates the likelihood for opportunities of allopatric speciation. In a comparison of ASV diversity, the TGT site were enriched in shared MCMS ASVs (Fig 5.). Unique marine ASVs were identified from the relic site across several taxa groups indicating a potentially endemic nature to the eukaryotes recovered on a sequence-level.

The Taylor Glacier, situated at the easternmost part of the East Antarctic Ice Sheet (EAIS), has undergone numerous advances and retreats during past warming and cooling events (Scherer et al. 2008, Naish et al. 2009, Levy et al. 2012). While there is evidence of a fjord-like Taylor Valley, the reach and timing of the marine waters remains a question (Hendy et al. 1977, Porter and Beget 1981, Higgins, Hendy and Denton 2000, Toner et al. 2013). The subglacial brine beneath the Taylor Glacier is the only surface reaching of the widespread aquifers (Mikucki et al. 2015). The findings presented in this study supports the idea of an ancient marine incursion into the Taylor Glacier through the detection of a persistent marine eukaryotic community that is potentially seeded and maintained by a supply of iron-rich, nutrient replete brine.

Much of the refutation marine source hypotheses as evidenced by marine depositions is the likelihood of aeolian emplacement of these communities (Scherer et al. 2004, Scherer et al. 2016).

The McMurdo Dry Valleys are regarded as one of the windiest regions in the world; However summer-time, coastal easterly winds are low-intensity relative to the westerly (down-valley), katabatic winds prominent during the winter months (Katurji et al. 2019). Several biodiversity assessments on aerosolized bacterial diversity of varying distance to Ross Sea and marine environments found little to no significant marine influence in the composition of aerosol communities (Hughes et al. 2004, Pearce et al. 2010, Bottos et al. 2014, Schulte et al. 2021). In addition, past studies on MDVs aeolian dispersal and connectivity of biodiversity have emphasized the unlikelihood of interbasin and intervalley transport of airborne microbiota with intrabasin dispersal more likely (Schulte et al. 2021, Katurji et al. 2019). This is primarily a result of the separation of valleys and basins by obstructive landscape features including mountain ranges (Katurji et al. 2019).

A limited amount of marine species were detected in our Explorer Cove and Lake Bonney wind collector, and Lawson stream site samples. A small presence of marine species is expected in these areas due to localized scale aeolian dispersal. These sites are the closest in proximity of sampled sites to the MCMS and TGT, respectively (Fig 1.); Therefore, this finding is consistent with previous work on MDV airborne community patterns. Additionally, no unique marine ASVs were discovered in the wind samples. A unique marine diatom ASV was detected in the Lawson stream site most likely originating from the TGT community.

The probability of colonization success is largely affected by viability of the colonizing organism (or propagule) (Wynn-Williams 1991). Aeolian dispersal selects for small size class cells (Schulte et al. 2021); however both short and long distance wind travel create physiological detrimental for these cells owing to particle collision and fragmentation during suspension . The heavy silicification and armored properties of some diatoms and dinoflagellates can be beneficial

in this mode of transport. Marine diatoms are often significantly larger in size and less silicified than freshwater diatoms (Conley, Kilham and Theriot 1989, Litchman, Klausmeier and Yoshiyama 2009). The dinoflagellate, *Gymnodiniaceae*, a species indicator taxon for MCMS and TGT red mud environments, lacks an armored structure making it a particularly difficult to study due to its susceptibility for cell damage (Escarcega-Bata et al. 2021). Subsequently, the likelihood of viability for the extensive wind transport of these taxa is minimal.

The settlement habitat (environmental filter) is another dynamic for establishing a dominant intrusive community. The findings in this study suggest salinity as the major environmental driver behind the unique distribution and composition of MDVs microbial communities. Viable marine eukaryotes deposited at the Taylor Glacier terminus site are matched in a semi-favorable environment compared to other landscape features and locations. The chemistry of the brine that sporadically saturates the stream channel is comprised of high concentrations of chloride, sulfate, and iron (Mikucki et al. 2009). Relative to the meltwater streams, the hypersaline brine is higher in dissolved nitrogen, phosphorus and organic carbon (Mikucki et al. 2009). In a less favorable turn, meltwater periodically flows from the Taylor Glacier, episodically mixing with brine, supplying oligotrophic water and resulting high osmotic variability (Lyons et al. 2019). The terminus site undergoes frequent desiccation over the austral summer though occasionally experiences short brine flow occurrences during the winter months (Carr, Carmichael and Pettit 2021). This event is distinct from the other ephemeral aquatic communities in the Dry Valleys as high salt content sustains the supraglacial lake and hydrological flow is independent of the warm summer temperatures. Nonetheless, the community is often exposed to sustained desiccation.

Among the taxa groups described in the TGT community analysis, several have cyst, spore or resting cell life cycle stages (Belmonte and Rubino 2019, Tang et al. 2021, Rembauville et al. 2018). Notably, diatoms such as the marine species *Chaetoceros socialis*, can transition from an active cell into a spore that remains viable for at least 9 months (Pelusi et al. 2020). The duration of spore viability of planktonic diatoms is lengthened significantly by low temperatures and anoxic conditions (Ellegaard and Ribeiro 2018). Moreover, diatoms have the ability to transform rapidly from resting and active stages upon improved environmental conditions (Pelusi et al. 2020). Comparably, encystment is commonly seen in the life cycles of dinoflagellate and ciliate species (Belmonte and Rubino 2019). The formation of resting stages is a likely coping mechanism of protists to avoid the adverse, variable conditions present at the sample site (Belmonte and Rubino 2019). The upregulation of genes within the ferritin-like domain in the red mud samples (Fig 7.) provides support to the potential for transitions between active and resting protist stages as overexpressed ferritin is known to play a role in the morphological transformation and dormancy in microorganisms (Garcia-Morales et al. 2017, Liu et al. 2022).

The marine eukaryotic community discovered at the Taylor Valley terminus site is dominant, active and diverse. Taken together, with evidence for past marine incursions in the polar desert, observations of limited, localized aeolian dispersal and sequence divergences we provide findings indicative of marine relic microbial communities in the McMurdo Dry Valleys. In testing this hypothesis, we conducted thorough molecular samplings on the communities of the TGT site (Blood Falls), aeolian collectors and combined these with extensive genomic datasets from the MCMS and Dry Valleys. The sequencing of environmental 18S rRNA from these sites revealed insights into the unique community diversification and activity at the TGT site. Further comparisons of ASVs identified a divergence of the TGT endemics from their oceanic

counterparts. These results show the potential for intrusive, marine microbiota to colonize areas of the polar desert as well as existing degrees of connectivity between these habitats. Additionally we report functional mechanisms for the ability of intrusive marine taxa to persist in the polar desert, overcome the stresses of a desert environment.

Conclusion

This study follows up on critical questions left by prior studies on the possible effects of ancient sea level rise and marine incursions into the MDVs. With atmospheric CO_2 (350-450ppm) increasing, the projected temperature rise in the near future could result in higher sea levels over the next centuries/millenniums (Pagani et al. 2010, Gasson, DeConto and Pollard 2016, Westerhold et al. 2020). The discovery of this relic community at the TGT provides us with a better understanding of the extent to which past climatic drivers and warming events have shaped ecosystems. This research expands polar desert eukaryotic taxa diversity and has generated novel amplicon sequence variants (ASVs) of particular importance in polar environments, where the application of molecular tools to biogeochemical cycling has been hindered by a lack of sequence data and suitable reference databases. In addition, the eukaryotic community structures of the surrounding oceanic and MDVs-based aeolian collectors sheds light on the limited but existing marine intrusions into the present day ecosystem.

Conflict of Interest

The authors declare no conflict of interest

Acknowledgements

This study was supported by the National Science Foundation Grant OPP-1637708 (to the MCM LTER). Additionally, this research was funded by NSF Antarctic Sciences Awards, NSF-OPP 0732822 and NSF-OPP-1043671 (to A.E.A.), Gordon and Betty Moore Foundation Grant GBMF3828 (to A.E.A.); NSF Ocean Sciences Award NSF-OCE-1136477 and NSF-OCE-1756884 (to A.E.A.). Thank you to the MCM-LTER PIs, student and logistics field teams. Thank you to Rob Lampe and Ariel Rabines for support in the sample and bioinformatic processing of this data.

Chapter two has been prepared as a submission to the PNAS Journal. A. Zoumplis, D. Kaul, N. Schulte, H. Zheng, P. Venepally, D. M. McKnight, A. E. Allen. "An intrusive, marine eukaryotic assemblage in the Antarctic Dry Valleys." The dissertation author was the primary investigator and author of this paper.

References

1. Adams, B. J., D. H. Wall, R. A. Virginia, E. Broos & M. A. Knox (2014) Ecological biogeography of the terrestrial nematodes of Victoria Land, Antarctica. *ZooKeys***,** 29.

2. Alger, A. (1997) Ecological processes in a cold desert ecosystem: the abundance and species distribution of algal mats in glacial meltwater streams in Taylor Valley, Antarctica. *Occasional paper/University of Colorado*.

3. Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow & S. M. Huse (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PloS one,* 4**,** e6372.

4. Andriuzzi, W., B. Adams, J. Barrett, R. Virginia & D. Wall (2018) Observed trends of soil fauna in the Antarctic Dry Valleys: early signs of shifts predicted under climate change. *Ecology,* 99**,** 312-321.

5. Bar Dolev, M., I. Braslavsky & P. L. Davies (2016) Ice-binding proteins and their function. *Annual review of biochemistry,* 85**,** 515-542.

6. Barrett, J., R. Virginia, D. Wall, P. Doran, A. Fountain, K. Welch & W. Lyons (2008) Persistent effects of a discrete warming event on a polar desert ecosystem. *Global Change Biology,* 14**,** 2249-2261.

7. Barrett, J. E., R. A. Virginia, W. B. Lyons, D. M. McKnight, J. C. Priscu, P. T. Doran, A. G. Fountain, D. H. Wall & D. Moorhead (2007) Biogeochemical stoichiometry of Antarctic dry valley ecosystems. *Journal of Geophysical Research: Biogeosciences,* 112.

8. Bertrand, E. M., J. P. McCrow, A. Moustafa, H. Zheng, J. B. McQuaid, T. O. Delmont, A. F. Post, R. E. Sipler, J. L. Spackeen & K. Xu (2015) Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proceedings of the National Academy of Sciences,* 112**,** 9938-9943.

9. Bielewicz, S., E. Bell, W. Kong, I. Friedberg, J. C. Priscu & R. M. Morgan-Kiss (2011) Protist diversity in a permanently ice-covered Antarctic lake during the polar night transition. *The ISME journal,* 5**,** 1559-1564.

10. Blazewicz, S. J., R. L. Barnard, R. A. Daly & M. K. Firestone (2013) Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *The ISME journal,* 7**,** 2061-2068.

11. Bolhuis, H. & L. J. Stal (2011) Analysis of bacterial and archaeal diversity in coastal microbial mats using massive parallel 16S rRNA gene tag sequencing. *The ISME journal,* 5**,** 1701-1712.
12. Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson & S. P. Holmes (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods,* 13**,** 581-583.

13. Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich & J. I. Gordon (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods,* 7**,** 335-336.

14. Cary, S. C., I. R. McDonald, J. E. Barrett & D. A. Cowan (2010) On the rocks: the microbiology of Antarctic Dry Valley soils. *Nature Reviews Microbiology,* 8**,** 129-138.

15. Chan, Y., J. D. Van Nostrand, J. Zhou, S. B. Pointing & R. L. Farrell (2013) Functional ecology of an Antarctic dry valley. *Proceedings of the National Academy of Sciences,* 110**,** 8990-8995.

16. Chua, M. J., R. L. Campen, L. Wahl, J. J. Grzymski & J. A. Mikucki (2018) Genomic and physiological characterization and description of Marinobacter gelidimuriae sp. nov., a psychrophilic, moderate halophile from Blood Falls, an Antarctic subglacial brine. *FEMS microbiology ecology,* 94**,** fiy021.

17. Coyne, K. J., A. E. Parker, C. K. Lee, J. A. Sohm, A. Kalmbach, T. Gunderson, R. León-Zayas, D. G. Capone, E. J. Carpenter & S. C. Cary (2020) The distribution and relative ecological roles of autotrophic and heterotrophic diazotrophs in the McMurdo Dry Valleys, Antarctica. *FEMS microbiology ecology,* 96**,** fiaa010.

18. Cullis, J. D., L. F. Stanish & D. M. McKnight (2014) Diel flow pulses drive particulate organic matter transport from microbial mats in a glacial meltwater stream in the McMurdo Dry Valleys. *Water Resources Research,* 50**,** 86-97.

19. Davey, M. C. & K. J. Clarke (1992) FINE STRUCTURE OF A TERRESTRIAL CYANOBACTERIAL MAT FROM ANTARCTICA 1. *Journal of phycology,* 28**,** 199-202.

20. Doran, P. T., W. B. Lyons & D. M. McKnight. 2010. *Life in Antarctic deserts and other cold dry environments: astrobiological analogs*. Cambridge University Press.

21. Doran, P. T., C. P. McKay, G. D. Clow, G. L. Dana, A. G. Fountain, T. Nylen & W. B. Lyons (2002a) Valley floor climate observations from the McMurdo Dry Valleys, Antarctica, 1986–2000. *Journal of Geophysical Research: Atmospheres,* 107**,** ACL 13-1-ACL 13-12.

22. Doran, P. T., J. C. Priscu, W. B. Lyons, J. E. Walsh, A. G. Fountain, D. M. McKnight, D. L. Moorhead, R. A. Virginia, D. H. Wall & G. D. Clow (2002b) Antarctic climate cooling and terrestrial ecosystem response. *Nature,* 415**,** 517-520.

23. Dupont, C. L., J. P. McCrow, R. Valas, A. Moustafa, N. Walworth, U. Goodenough, R. Roth, S. L. Hogle, J. Bai & Z. I. Johnson (2015) Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. *The ISME journal,* 9**,** 1076-1092.

24. Esposito, R., S. Horn, D. M. McKnight, M. Cox, M. Grant, S. Spaulding, P. Doran & K. Cozzetto (2006) Antarctic climate cooling and response of diatoms in glacial meltwater streams. *Geophysical Research Letters,* 33.

25. Esposito, R., S. Spaulding, D. M. McKnight, B. Van de Vijver, K. Kopalová, D. Lubinski, B. Hall & T. Whittaker (2008) Inland diatoms from the McMurdo dry valleys and James Ross Island, Antarctica. *Botany,* 86**,** 1378-1392.

26. Finn, R., J. Mistry, J. Tate, P. Coggill & A. Heger (2014) Pfam: the protein families database. Nuclei. Acids Re.

27. Finn, R. D., J. Clements & S. R. Eddy (2011) HMMER web server: interactive sequence similarity searching. *Nucleic acids research,* 39**,** W29-W37.

28. Fountain, A. G., J. S. Levy, M. N. Gooseff & D. Van Horn (2014) The McMurdo Dry Valleys: a landscape on the threshold of change. *Geomorphology,* 225**,** 25-35.

29. Fountain, A. G., T. H. Nylen, A. Monaghan, H. J. Basagic & D. Bromwich (2010) Snow in the McMurdo dry valleys, Antarctica. *International Journal of Climatology: A Journal of the Royal Meteorological Society,* 30**,** 633-642.

30. Fountain, A. G., G. Saba, B. Adams, P. Doran, W. Fraser, M. Gooseff, M. Obryk, J. C. Priscu, S. Stammerjohn & R. A. Virginia (2016) The impact of a large-scale climate event on Antarctic ecosystem processes. *Bioscience,* 66**,** 848-863.

31. Gooseff, M. N., J. E. Barrett, B. J. Adams, P. T. Doran, A. G. Fountain, W. B. Lyons, D. M. McKnight, J. C. Priscu, E. R. Sokol & C. Takacs-Vesbach (2017) Decadal ecosystem response to an anomalous melt season in a polar desert in Antarctica. *Nature ecology & evolution,* 1**,** 1334-1338.

32. Gooseff, M. N., D. M. McKnight, P. Doran, A. G. Fountain & W. B. Lyons (2011) Hydrological connectivity of the landscape of the McMurdo Dry Valleys, Antarctica. *Geography Compass,* 5**,** 666-681.

33. Gururani, M. A., J. Venkatesh & L. S. P. Tran (2015) Regulation of photosynthesis during abiotic stress-induced photoinhibition. *Molecular plant,* 8**,** 1304-1320.

34. Hawes, I. & A. M. Schwarz (2001) Absorption and utilization of irradiance by cyanobacterial mats in two ice-covered antarctic lakes with contrasting light climates. *Journal of Phycology,* 37**,** 5-15.

35. Hopkins, D., A. Sparrow, B. Elberling, E. Gregorich, P. Novis, L. Greenfield & E. Tilston (2006) Carbon, nitrogen and temperature controls on microbial activity in soils from an Antarctic dry valley. *Soil Biology and Biochemistry,* 38**,** 3130-3140.

36. Howard-Williams, C., J. C. Priscu & W. F. Vincent (1989) Nitrogen dynamics in two Antarctic streams. *Hydrobiologia,* 172**,** 51-61.

37. Howard-Williams, C., C. L. Vincent, P. A. Broady & W. F. Vincent (1986) Antarctic stream ecosystems: variability in environmental properties and algal community structure. *Internationale Revue der Gesamten Hydrobiologie und Hydrographie,* 71**,** 511-544.

38. Jungblut, A. D., I. Hawes, D. Mountfort, B. Hitzfeld, D. R. Dietrich, B. P. Burns & B. A. Neilan (2005) Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environmental microbiology,* 7**,** 519-529.

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

40. Knight, C. & A. DeVries (2009) Ice growth in supercooled solutions of a biological "antifreeze", AFGP 1–5: An explanation in terms of adsorption rate for the concentration dependence of the freezing point. *Physical Chemistry Chemical Physics,* 11**,** 5749-5761.

41. Kohler, T. J., L. F. Stanish, S. W. Crisp, J. C. Koch, D. Liptzin, J. L. Baeseman & D. M. McKnight (2015) Life in the main channel: long-term hydrologic control of microbial mat abundance in McMurdo Dry Valley streams, Antarctica. *Ecosystems,* 18**,** 310-327.

42. Kohler, T. J., L. F. Stanish, D. Liptzin, J. E. Barrett & D. M. McKnight (2018) Catch and release: Hyporheic retention and mineralization of N-fixing Nostoc sustains downstream microbial mat biomass in two polar desert streams. *Limnology and Oceanography Letters,* 3**,** 357-364.

43. Kolody, B., J. McCrow, L. Z. Allen, F. Aylward, K. Fontanez, A. Moustafa, M. Moniruzzaman, F. Chavez, C. Scholin & E. Allen (2019) Diel transcriptional response of a California Current plankton microbiome to light, low iron, and enduring viral infection. *The ISME journal,* 13**,** 2817-2833.

44. Králová, S. (2017) Role of fatty acids in cold adaptation of Antarctic psychrophilic Flavobacterium spp. *Systematic and applied microbiology,* 40**,** 329-333.

45. Kubota, N. (2011) Effects of cooling rate, annealing time and biological antifreeze concentration on thermal hysteresis reading. *Cryobiology,* 63**,** 198-209.

46. Langfelder, P. & S. Horvath (2008) WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics,* 9**,** 1-13.

47. Levy, J. S., A. G. Fountain, M. Obryk, J. Telling, C. Glennie, R. Pettersson, M. Gooseff & D. Van Horn (2018) Decadal topographic change in the McMurdo Dry Valleys of Antarctica: Thermokarst subsidence, glacier thinning, and transfer of water storage from the cryosphere to the hydrosphere. *Geomorphology,* 323**,** 80-97.

48. Marizcurrena, J. J., M. F. Cerdá, D. Alem & S. Castro-Sowinski. 2019. Living with pigments: the colour palette of Antarctic life. In *The ecological role of micro-organisms in the antarctic environment*, 65-82. Springer.

49. McKnight, D. M., D. K. Niyogi, A. S. Alger, A. Bomblies, P. A. Conovitz & C. M. Tate (1999) Dry valley streams in Antarctica: ecosystems waiting for water. *Bioscience,* 49**,** 985-995.

50. McKnight, D. M., R. L. Runkel, C. M. Tate, J. H. Duff & D. L. Moorhead (2004) Inorganic N and P dynamics of Antarctic glacial meltwater streams as controlled by hyporheic exchange and benthic autotrophic communities. *Journal of the North American Benthological Society,* 23**,** 171-188.

51. McKnight, D. M. & C. Tate (1997) Canada stream: a glacial meltwater stream in Taylor Valley, south Victoria Land, Antarctica. *Journal of the North American Benthological Society,* 16**,** 14-17.

52. McKnight, D. M., C. Tate, E. Andrews, D. Niyogi, K. Cozzetto, K. Welch, W. Lyons & D. Capone (2007) Reactivation of a cryptobiotic stream ecosystem in the McMurdo Dry Valleys, Antarctica: a long-term geomorphological experiment. *Geomorphology,* 89**,** 186-204.

53. Meng, J., B. Hu, G. Yi, X. Li, H. Chen, Y. Wang, W. Yuan, Y. Xing, Q. Sheng & Z. Su (2020) Genome-wide analyses of banana fasciclin-like AGP genes and their differential expression under low-temperature stress in chilling sensitive and tolerant cultivars. *Plant cell reports,* 39**,** 693-708.

54. Morgan-Kiss, R., M. Lizotte, W. Kong & J. Priscu (2016) Photoadaptation to the polar night by phytoplankton in a permanently ice-covered Antarctic lake. *Limnology and Oceanography,* 61**,** 3-13.

55. Murata, N., S. Takahashi, Y. Nishiyama & S. I. Allakhverdiev (2007) Photoinhibition of photosystem II under environmental stress. *Biochimica et Biophysica Acta (BBA)-Bioenergetics,* 1767**,** 414-421.

56. Nielsen, U. N., D. H. Wall, B. J. Adams, R. A. Virginia, B. A. Ball, M. N. Gooseff & D. M. McKnight (2012) The ecology of pulse events: insights from an extreme climatic event in a polar desert ecosystem. *Ecosphere,* 3**,** 1-15.

57. Obryk, M. K., P. T. Doran, A. G. Fountain, M. Myers & C. P. McKay (2020) Climate from the McMurdo dry valleys, Antarctica, 1986–2017: Surface air temperature trends and redefined summer season. *Journal of Geophysical Research: Atmospheres,* 125**,** e2019JD032180.

58. Parada, A. E., D. M. Needham & J. A. Fuhrman (2016) Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental microbiology,* 18**,** 1403-1414.

59. Rai, R., S. Singh, A. Chatterjee, K. K. Rai, S. Rai & L. Rai (2020) All4894 encoding a novel fasciclin (FAS-1 domain) protein of Anabaena sp. PCC7120 revealed the presence of a thermostable β-glucosidase. *Algal Research,* 51**,** 102036.

60. Raymond-Bouchard, I. & L. G. Whyte. 2017. From transcriptomes to metatranscriptomes: cold adaptation and active metabolisms of psychrophiles from cold environments. In *Psychrophiles: from biodiversity to biotechnology*, 437-457. Springer.

61. Rho, M., H. Tang & Y. Ye (2010) FragGeneScan: predicting genes in short and error-prone reads. *Nucleic acids research,* 38**,** e191-e191.

62. Robinson, M. D., D. J. McCarthy & G. K. Smyth (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-140.

63. Schmieder, R., Y. W. Lim & R. Edwards (2012) Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics,* 28**,** 433-435.

64. Seifert, G. J. (2018) Fascinating fasciclins: a surprisingly widespread family of proteins that mediate interactions between the cell exterior and the cell surface. *International journal of molecular sciences,* 19**,** 1628.

65. Singley, J. G., M. N. Gooseff, D. M. McKnight & E. Hinckley (2021) The Role of Hyporheic Connectivity in Determining Nitrogen Availability: Insights from an Intermittent Antarctic Stream. *Journal of Geophysical Research: Biogeosciences,* 126**,** e2021JG006309.

66. Sommers, P., J. L. Darcy, D. L. Porazinska, E. Gendron, A. G. Fountain, F. Zamora, K. Vincent, K. M. Cawley, A. J. Solon & L. Vimercati (2019) Comparison of microbial communities in the sediments and water columns of frozen cryoconite holes in the McMurdo Dry Valleys, Antarctica. *Frontiers in microbiology,* 10**,** 65.

67. Sorokovikova, E. G., O. I. Belykh, A. S. Gladkikh, O. V. Kotsar, I. V. Tikhonova, O. A. Timoshkin & V. V. Parfenova (2013) Diversity of cyanobacterial species and phylotypes in biofilms from the littoral zone of Lake Baikal. *Journal of Microbiology,* 51**,** 757-765.

68. Stanish, L. F., D. R. Nemergut & D. M. McKnight (2011) Hydrologic processes influence diatom community composition in Dry Valley streams. *Journal of the North American Benthological Society,* 30**,** 1057-1073.

69. Stoeck, T., D. Bass, M. Nebel, R. Christen, M. D. Jones, H. W. BREINER & T. A. Richards (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular ecology,* 19**,** 21-31.

70. Takamichi, M., Y. Nishimiya, A. Miura & S. Tsuda (2007) Effect of annealing time of an ice crystal on the activity of type III antifreeze protein. *The FEBS journal,* 274**,** 6469-6476.

71. Toner, J. D., R. S. Sletten & M. L. Prentice (2013) Soluble salt accumulations in Taylor Valley, Antarctica: implications for paleolakes and Ross Sea Ice Sheet dynamics. *Journal of Geophysical Research: Earth Surface,* 118**,** 198-215.

72. Van Horn, D. J., C. R. Wolf, D. R. Colman, X. Jiang, T. J. Kohler, D. M. McKnight, L. F. Stanish, T. Yazzie & C. D. Takacs-Vesbach (2016) Patterns of bacterial biodiversity in the glacial meltwater streams of the McMurdo Dry Valleys, Antarctica. *FEMS Microbiology Ecology,* 92**,** fiw148.

73. Vance, T. D., M. Bayer-Giraldi, P. L. Davies & M. Mangiagalli (2019) Ice-binding proteins and the 'domain of unknown function'3494 family. *The FEBS journal,* 286**,** 855-873.

74. Varin, T., C. Lovejoy, A. D. Jungblut, W. F. Vincent & J. Corbeil (2012) Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. *Applied and environmental microbiology,* 78**,** 549-559.

75. Vick, T. J. & J. C. Priscu (2012) Bacterioplankton productivity in lakes of the Taylor Valley, Antarctica, during the polar night transition. *Aquatic microbial ecology,* 68**,** 77-90.

76. Vick-Majors, T. J., J. C. Priscu & L. A. Amaral-Zettler (2014) Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. *The ISME journal,* 8**,** 778-789.

77. Vincent, W., M. Downes, R. Castenholz & C. Howard-Williams (1993) Community structure and pigment organisation of cyanobacteria-dominated microbial mats in Antarctica. *European Journal of Phycology,* 28**,** 213-221.

78. Vincent, W. F. 2000. Cyanobacterial dominance in the polar regions. In *The ecology of cyanobacteria*, 321-340. Springer.

79. Wharton Jr, R. A., B. C. Parker & G. M. Simmons Jr (1983) Distribution, species composition and morphology of algal mats in Antarctic dry valley lakes. *Phycologia,* 22**,** 355-365.

80. Wlostowski, A. N., M. N. Gooseff, D. M. McKnight, C. Jaros & W. B. Lyons (2016) Patterns of hydrologic connectivity in the McMurdo Dry Valleys, Antarctica: a synthesis of 20 years of hydrologic data. *Hydrological Processes,* 30**,** 2958-2975.

Figures

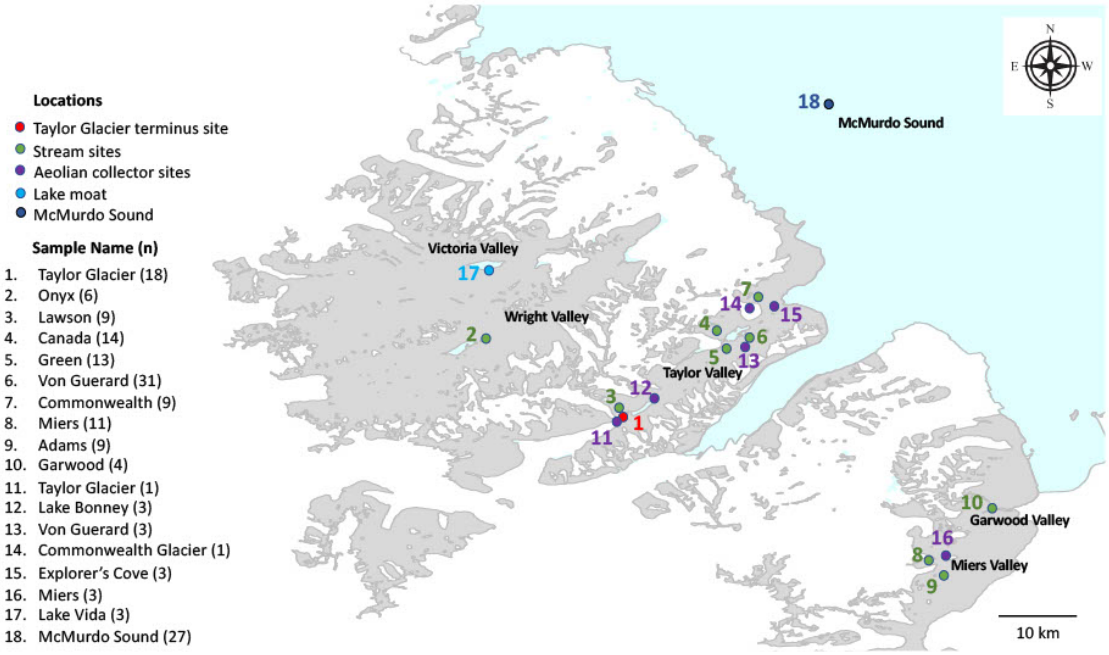

Figure 2.1 Map of Victorialand sampling locations

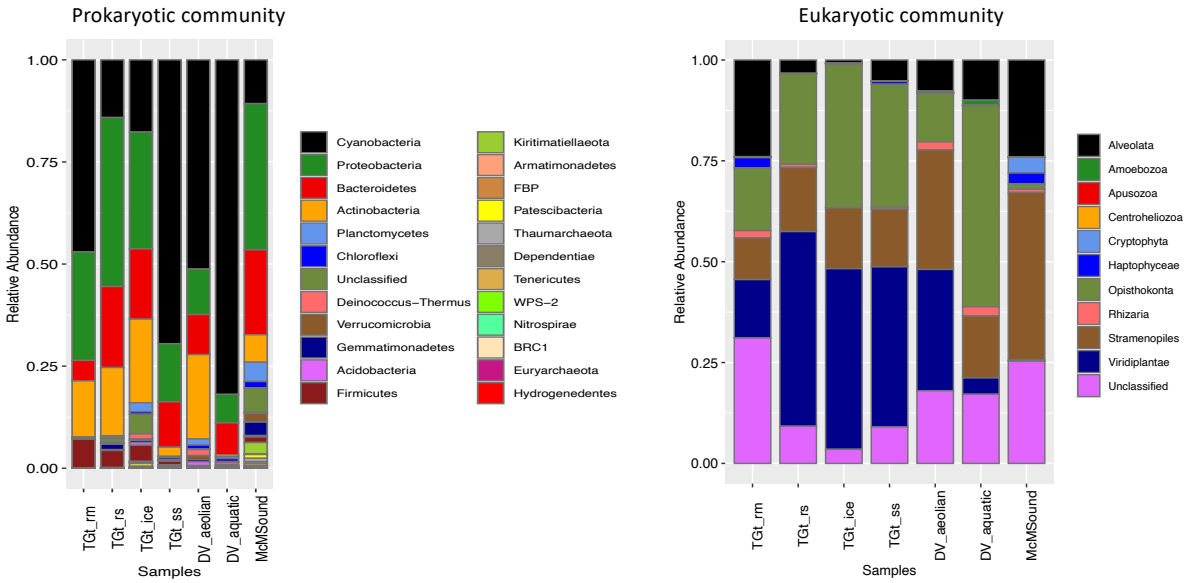

Figure 2.2 Relative abundance of (**left**) prokaryotic and (**right**) eukaryotic community members from non-plastid 16S and 18S rRNA ASVs.

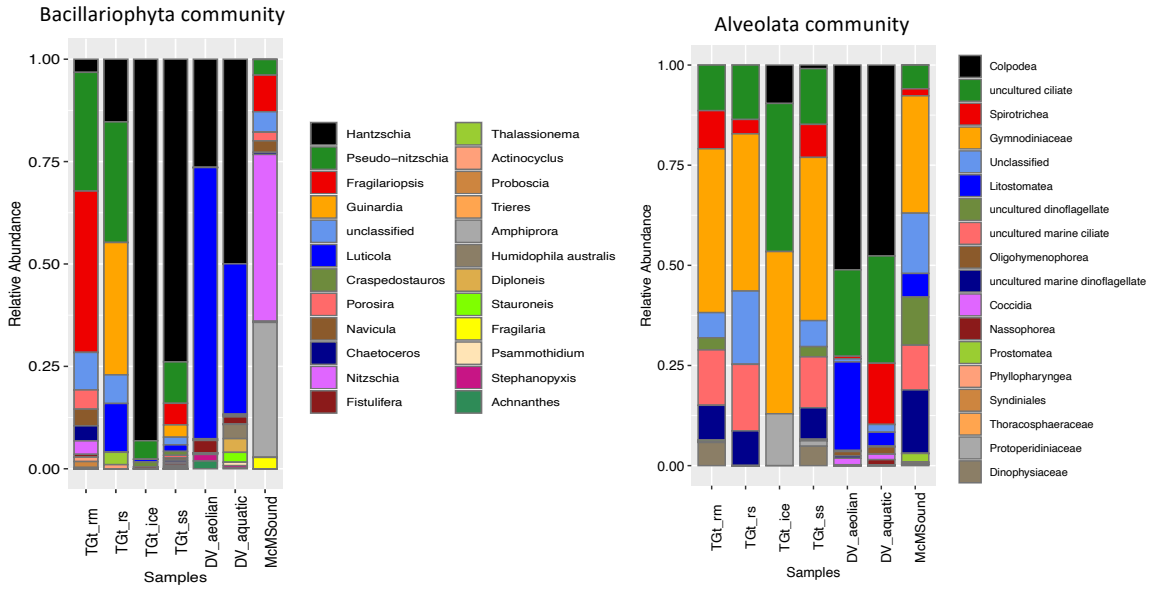

Figure 2.3 Relative abundance of (**left**) Bacillariophyta and (**right**) Alveolata community members from non-plastid 16S and 18S rRNA ASVs.

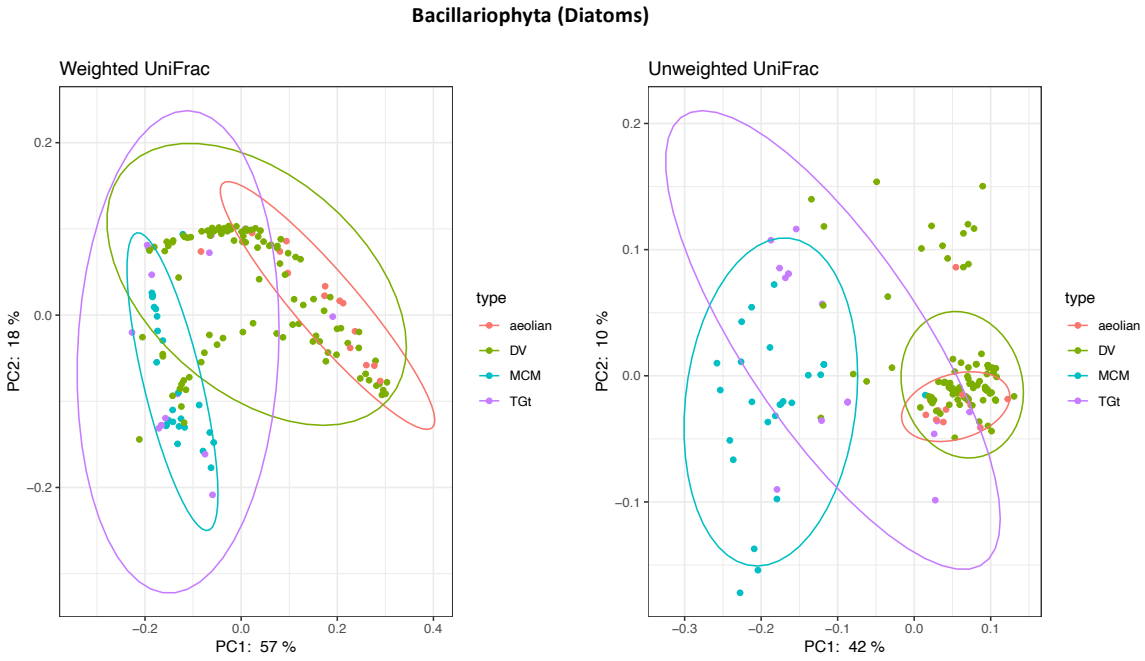

Figure 2.4 Principal coordinate analysis (PCoA) of Bacillariophyta **(left)** weighted and **(right)** unweighted UniFrac distance matrices with 95% confidence ellipses. Color of plotted points denote sample environment where "aeolian", "DV", "MCM" and "TGt" represent Dry Valleys aeolian, Dry Valleys aquatic, McMurdo Sound and Taylor Glacier terminus collection sites, respectively.

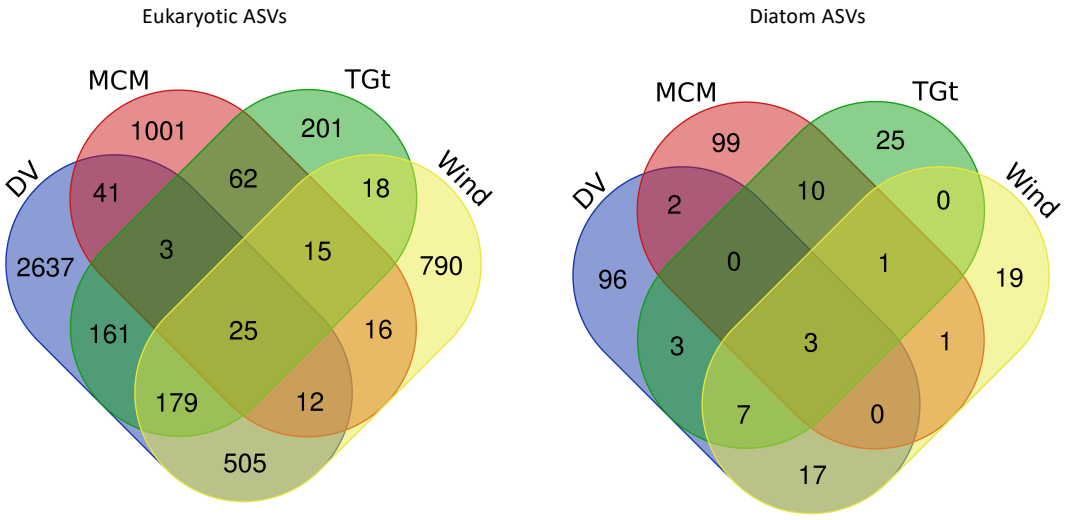

Figure 2.5 A Venn diagram illustrating the relationships between **(left)** eukaryotic and **(right)** diatom ASVs from each sample environment. Colored circles are labeled to denote sample environments where "Wind", "DV", "MCM" and "TGt" represent Dry Valleys aeolian, Dry Valleys aquatic, McMurdo Sound and Taylor Glacier terminus collection sites, respectively.

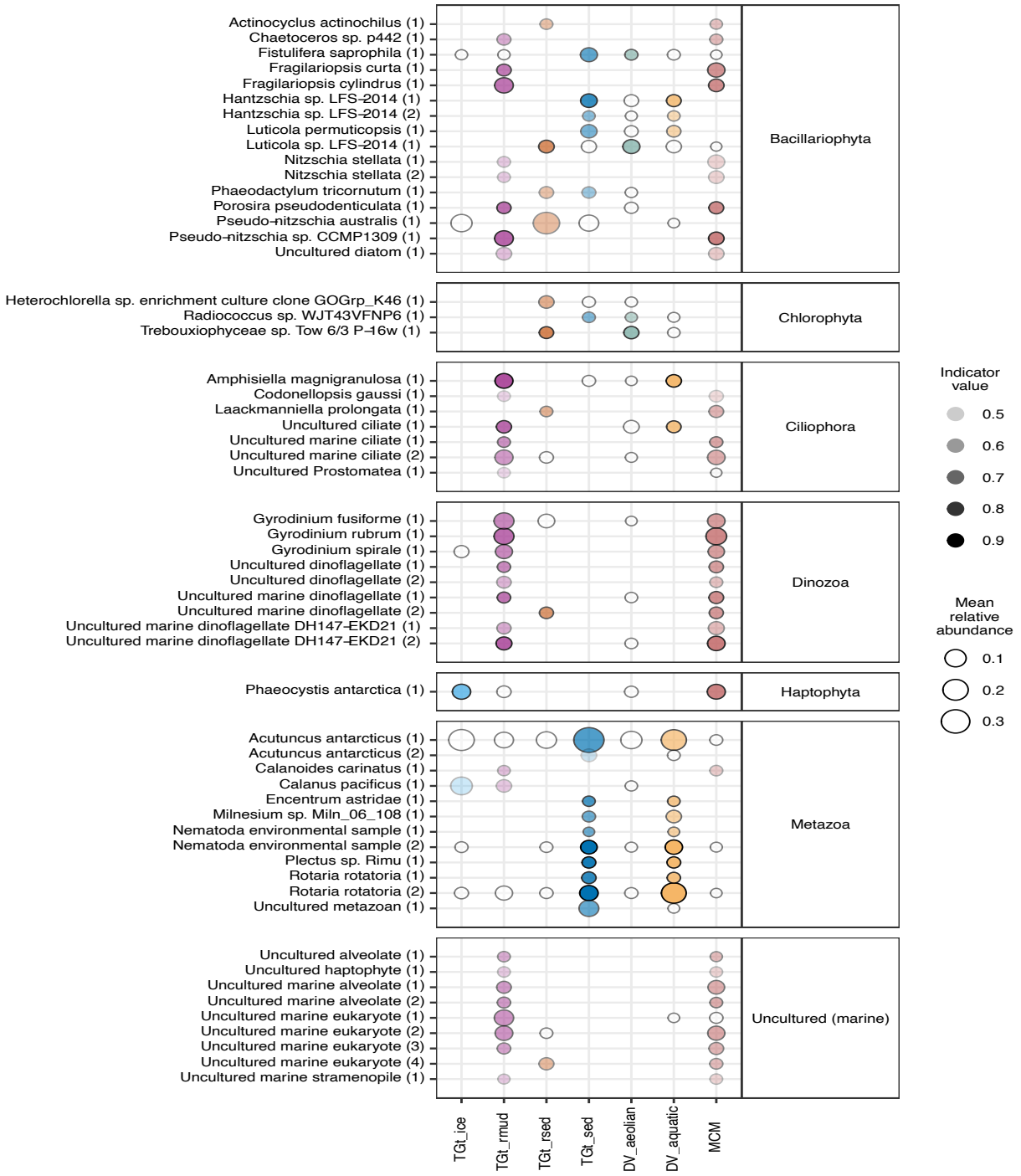

Figure 2.6 ASVs from select eukaryotic phyla that are indicators of each Blood Falls habitat or a combination of a Blood Falls habitat and a second habitat. Each white-filled dot reflects the relative abundance of an ASV within a habitat for which it was not an indicator.

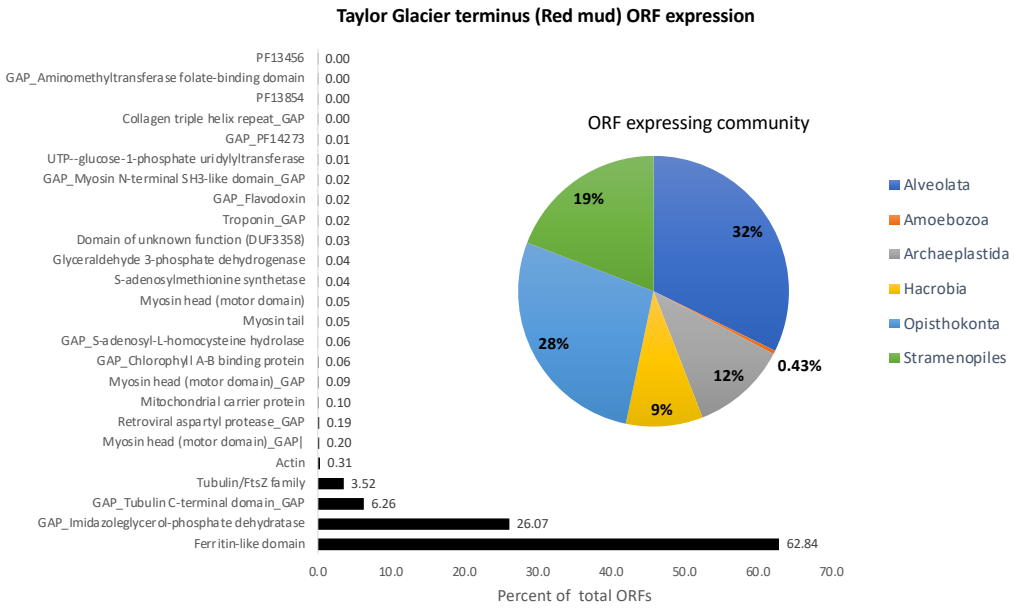

Figure 2.7 The top 25 most highly expressed ORFs in the Taylor Glacier terminus red mud samples. The pie chart depicts the relative abundance of taxonomically classified ORF expressing community members.

Supporting Figures

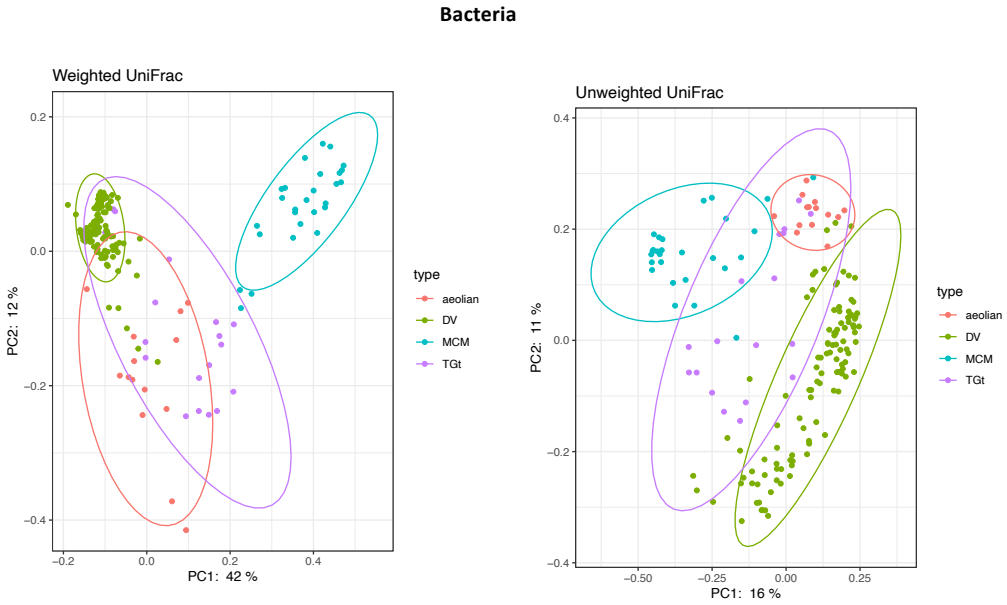

Figure S1 Principal coordinate analysis (PCoA) of bacteria **(left)** weighted and **(right)** unweighted UniFrac distance matrices with 95% confidence ellipses.

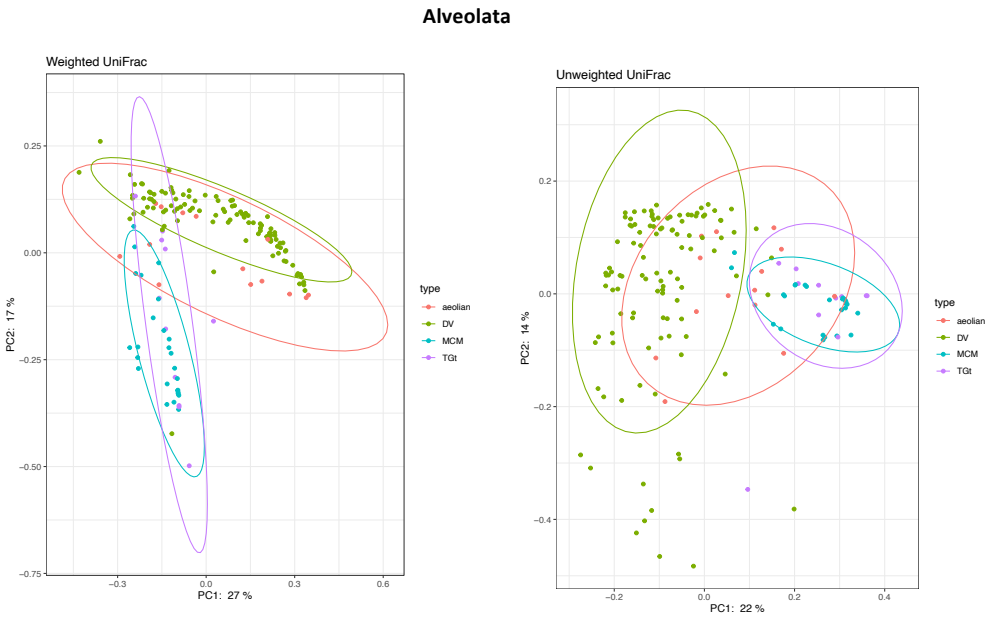

Figure S2 Principal coordinate analysis (PCoA) of bacteria **(left)** weighted and **(right)** unweighted UniFrac distance matrices with 95% confidence ellipses.

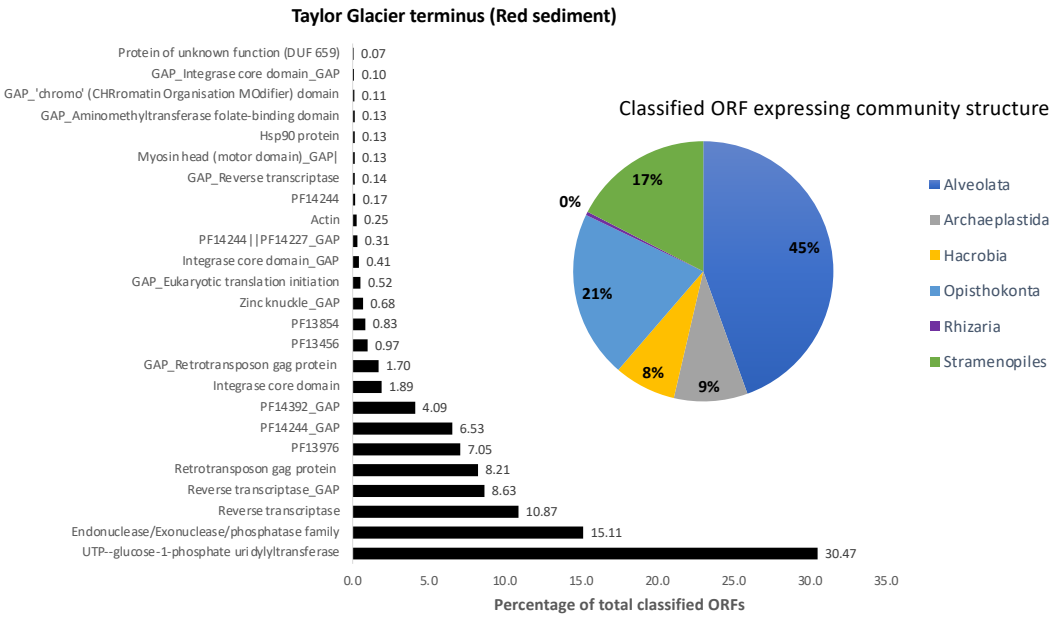

Figure S3 The top 25 most highly expressed ORFs in the Taylor Glacier terminus red sediment samples. The pie chart depicts the relative abundance of taxonomically classified ORF expressing community members.

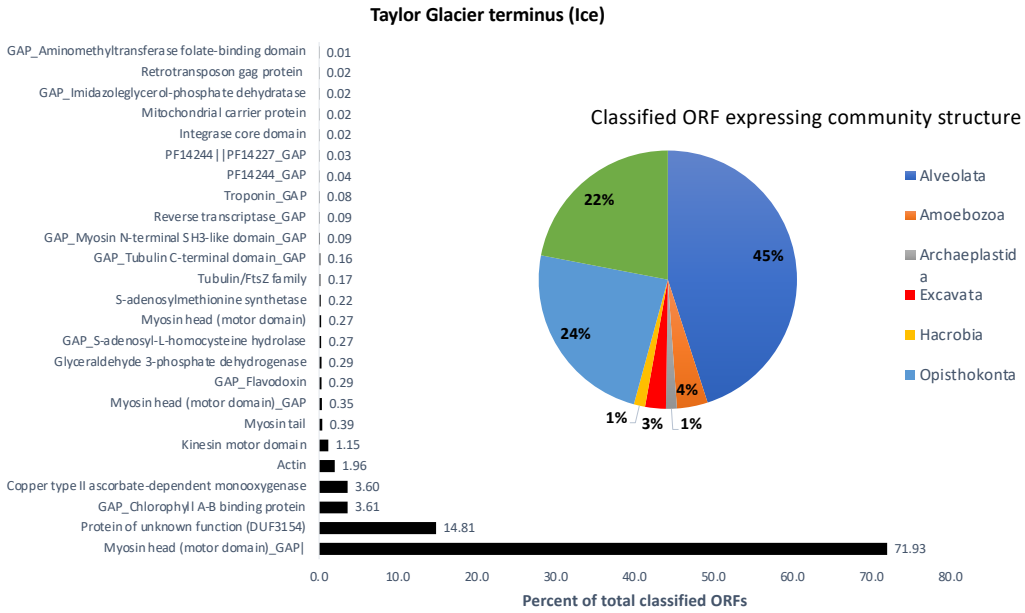

Figure S4 The top 25 most highly expressed ORFs in the Taylor Glacier terminus ice samples. The pie chart depicts the relative abundance of taxonomically classified ORF expressing community members.

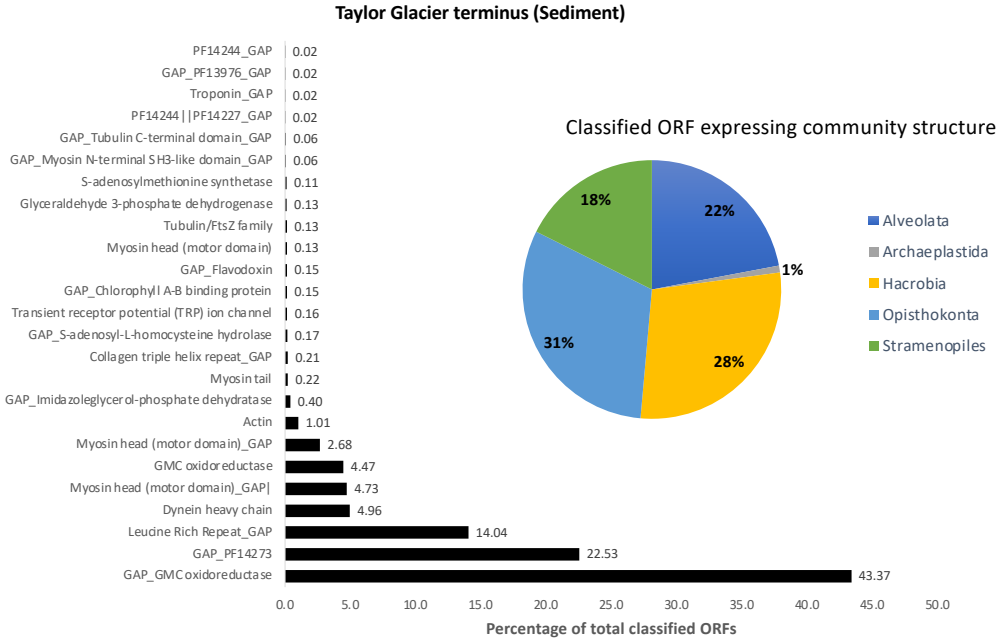

Figure S5 The top 25 most highly expressed ORFs in the Taylor Glacier terminus sediment samples. The pie chart depicts the relative abundance of taxonomically classified ORF expressing community members.

Heterologous expression and targeted delivery of sea ice diatom ice binding proteins.

A. Zoumplis, E. Garza, D. Yee, A. DeVries, V. Bielinski, A. E. Allen

Abstract

The production of ice binding proteins (IBPs) is an effective coping strategy in cold environments and one that is found across many taxa groups. Significant applications for these proteins have been utilized in the biotechnology, medicine and food industries. The impact of global change and concurrent freeze thaw cycling has recently placed agricultural production and the development of freeze protection tools for plants and soil microbes at the forefront of IBP research. Here, we investigate the efficacy and use of a moderate IBP from the polar diatom, *Fragilariopsis cylindrus*, in protecting a temperate diatom, *Phaeodactylum tricornutum* from freeze thaw cycling. Additionally, we investigated several individual and combined targeted localizations of the IBP on cell health and mortality rates. The findings in this study indicate a significant improvement in cell recovery after a freeze thaw event of the individual chloroplast, nucleus, cytoplasm localized IBP transformant lines over the wild type *P. tricornutum*. A triple stacked trait transformant was developed with the aforementioned localizations and displayed a significant increase in maximum quantum yield of photosystem II and reductions in mortality after freeze thaw cycling over the wildtype.

Introduction

Global climate change has led to an increase in the frequency and severity of regional freeze thaw cycling events (Wipf et al. 2015, Gavazov et al. 2017). Erratic changes in temperature cause changes in soil communities (Liang 2021), lysis of cells (Mohr and Stein 1970, McLellan 1989, Liu et al. 2014), subsequent ecosystem losses of carbon and nutrients in the form of greenhouse gases (Pelster et al. 2019), leaching (Liu et al. 2014) and other detrimental impacts on the agricultural industry. Freeze thaw events have had severe impacts on a variety of internationally produced crops such as citrus and wheat, among others (Ito et al. 2018, Willick, Tanino and Gusta 2021). The demand for staple foods is expected to increase while crop abandonment occurrences and yield losses are expected to worsen (Ferrarezi, Rodriguez and Sharp 2020). Polar phytoplankton offer exploitable defenses against freezing conditions including cryoprotectants (Bayer-Giraldi et al. 2010). Genomic engineering tool development in diatoms has led to many promising developments utilizing specialized products in biotechnological, biofuel and agroecological advancements (Gatenby et al. 2003, Levitan et al. 2014, Daboussi et al. 2014).

Diatoms are significant contributors to the primary productivity of oceanic food webs, particularly the Southern Ocean, providing food to keystone species such as Krill (Lizotte 2001). They also play a vital role in global CO2 uptake and are responsible for ~40% of carbon export. (Tréguer et al. 2018) Antarctica provides dynamic ecological niches for these eukaryotic phytoplankton ranging from coastal sea ice to inland freshwater glacial streams. Ice associated diatoms possess adaptations not found in temperate species that allow for function in steep gradients of light, salinity, nutrients and temperature as well as periodic freeze-thaw events and desiccation-hydration cycles (Fiala and Oriol 1990, Teufel and Morgan-Kiss 2018, Garcia-Lopez, Alcazar and Cid 2021).

These environmental factors, combined and individually, pose obstacles for life. Low temperatures cause a variety of detrimental effects including the loss of membrane fluidity, reduction of chemical reaction rates, and lysing of cells by ice crystal formation (Morgan-Kiss et al. 2006). One way organisms combat low temperatures is by producing ice-binding proteins (IBPs) (Janech et al. 2006). Species of bacteria, fish, insects and phytoplankton all retain versions of these proteins, which are responsible for protecting cells from freezing damage (Raymond and Kim 2012). IBPs are secreted to the external environment where they may bind to the prism or basal planes of an ice crystal allowing for pitting (Raymond and Kim 2012). Pitting in sea ice results in small brine pocket formation within the ice that protects the cell from freezing injuries (Raymond and Kim 2012).

The flood of data generated through metagenomic and transcriptomic studies of Antarctic marine ecosystems contains a large distribution and diversity of IBPs, which predicts the essentiality of this adaptation to cold-thriving organisms (Uhlig et al. 2015). Diversity and abundance of eukaryotic IBP transcripts measured in Antarctic and Arctic ice cores (lower 10 cm) revealed abundances of IBP transcripts matching those of core cellular function reference genes including actin, light-harvesting proteins and protochlorophyllide reductase (Uhlig et al. 2015).

Sea ice brine pocket communities are dominated by the polar diatom *Fragilariopsis cylindrus* (Mock and Thomas 2005). Ice binding proteins possessed by this species (fcIBPs) are believed to have bacterial origins due to a proposed horizontal gene transfer (Raymond and Kim 2012). Isoforms of these genes have been characterized in terms of protein-ice interaction on a microstructural level (Raymond and Kim 2012, Kondo, Mochizuki and Bayer-Giraldi 2018). IBPs are known to increase the porosity of sea ice as visualized by differences in ice morphology (Janech et al. 2006). More recently, the thermal hysteresis activity of fcIBPs has been determined to be

moderate ~0.1 ˚C -0.3 ˚C (Kondo et al. 2018). Differential gene expression of *F. cylindrus* ice binding proteins in response to salt and cold shock have resulted in the up regulation and down regulation of several identified isoforms (Mock and Valentin 2004, Krell et al. 2008, Bayer-Giraldi et al. 2010)

The fully sequenced temperate diatom, *Phaeodactylum tricornutum* optimally grows around temperatures extending from 18° C to 22 ° C (William and Morris 1982). When exposed to low temperatures ~5° C, the growth rate and cell function of *P. tricornutum* greatly decrease (William and Morris 1982). Both *P. tricornutum* and *F. cylindrus* are in the same class of diatoms, Bacillariophyceae, which contain the raphid pennate diatoms. The raphe, a parallel-plate capillary, is functionally significant as it is responsible for the secretion of mucilage that aids in movement and surface adhesion (Gordon and Drum 1970). Studies on *P. tricornutum* has shown the ability to transfer heterologous genes via conjugation to the temperate species of diatom producing novel functions (Karas et al. 2015).

The goal of this investigation was to determine the efficacy of a polar diatom ice binding protein in the reduction of mortality of a non-polar diatom after exposure to freeze thaw cycling. Novel approaches utilized in this study integrated diatom ecophysiology and synthetic genomics in order to: 1.) Examine a known freeze thaw coping strategy. 2.) Observe the effects of targeted and native signal IBP transformation to a temperate diatom in conferring an improved tolerance to this condition. 3.) Compare wildtype and IBP transformed temperate diatom growth in response to repeated freeze thaw cycles.

Methods

Heterologous expression of fcIBP ID: 268004 in E.coli

Total RNA extracted from *F. cylindrus* cells was used in synthesizing cDNA for the PCR amplification of IBP coding sequence (ID: 268004). This coding sequence was selected based on alignment with other known ice binding proteins and the possession of a signal peptide. A blast of this protein to the Phyre2 Protein Fold Recognition Server revealed a similarity to a Flavobacterium antifreeze protein (Figure 1.). The presence of beta sheets indicates this protein is likely to hydrogen bond to ice crystal surfaces (Fig 1.). Once amplified, the coding sequence was cloned into a pBAD expression vector. This specific vector contains an arabinose inducible promoter. Full-length and modified versions of ice binding proteins with no signal peptide were cloned into this vector. Proteins were fused with His-FLAG and thioredoxin tags. Descriptions of expression vectors are in Table 1. Expression vectors were transformed into chemically competent E.coli. The His-FLAG tags allowed for a Cobalt resin pull down of the IBP. The full protocol for small scale expression testing of E.coli clones adapted from (Savitsky et al. 2010) can be found in Supplementary File 1.

Targeted delivery of fcIBP into *P. tricornutum*

The *P. tricornutum* species is an ideal candidate for transformation and observing the efficiency of individual IBP genes in regards to the following; 1.) This diatom does not naturally possess any IBP homologous genes as revealed by the sequencing of the *P. tricornutum* genome (Bowler et al 2008) 2.) The similarities of *P. tricornutum* in morphological features and relatedness to *F. cylindrus* (Lundholm, Daugbjerg and Moestrup 2002) 3.) The development of molecular and synthetic tools for delivery of plasmids into *P. tricornutum* via conjugation (Karas et al. 2015). Signal peptides, otherwise referred to as localization signals, are responsible for the transport of proteins to various organelles in a cell or the extracellular matrix. Conserved pre-

sequences around 20 amino acid in length have been identified in *P. tricornutum* to transport proteins into the cytoplasm and across chloroplast, nucleus, ER and mitochondria membranes. A full list of signal peptide localizations and references can be found in (Table 1.). These signal peptides were amplified from *P. tricornutum* and were added to the N-terminus of the fcIBP. A mRuby3 tag was attached to the end of C-terminus for visualization.

Table 3.1 References for modified *P. tricornutum* signal peptides used in the targeting IBP to various extracellular and intracellular locations.

Signal peptide target location	**References**
Extracellular matrix (secreted)	(Erdene-Ochir et al. 2019)
Chloroplast	(Apt et al. 2002)
Endoplasmic reticulum (ER)	(Kilian and Kroth 2005)
Mitochondria	(Kroth et al. 2008)
Chloroplast (pyrenoid compartment)	(Tachibana et al. 2011)
Nucleus	(Kalderon et al. 1984)

Strains and Growth Conditions

Escherichia coli (NEB5α) cells used for vector construction came from New England Biolabs and Epi300 cells used for diatom conjugations were from Lucigen. XJ autolysis cells (Zymo Research) were used for large-scale IBP production. For cloning, all strains were grown on Luria Bertani (LB) agar or in broth supplemented with the following antibiotics as needed: spectinomycin (50 µg ml-1), kanamycin (50 µg ml-1), and gentamycin (20 µg ml-1). *Phaeodactylum tricornutum* CCAP1055/1 cultures were grown in L1 medium containing 8.8 mM nitrate or 880 µM ammonium. The diatoms were grown in atmospheric levels of CO2 with 50 µE m-2s-1 light following a 14:10 diel cycle at 18˚C. For conjugations, diatoms were grown in 12-well plates containing a 50% mixture of 2% agar, 45% L1 media, and 5% LB broth, respectively. Spectinomycin (50 µg ml -1) was used to supplement the media to reduce unwanted bacterial growth from the diatom cultures. Following conjugations, diatom cultures were replated on agar

plates consisting of one-part L1 ammonium medium and one-part 2% agar supplemented with zeocin (Invitrogen) at a concentration of 40 µg ml-1 for selection of exconjugants. Transgenic cultures were maintained in L1 ammonium broth with zeocin until they were transferred to L1 nitrate media for induction experiments.

Cloning Methods

All of the vectors made in this study were constructed using the uLoop Assembly system (Pollak et al. 2020). Briefly, the DNA components of the IBP expression cassettes (promoter-signal peptide-IBP-mRuby3 or 3x stop-terminator) were made into individual level 0 (L0) parts and entered into the uLoop library. This allowed for the quick and easy assembly of each of the expression cassettes into multiple level 1 expression vectors using Golden Gate assembly (Engler and Marillonnet 2014). Each IBP expression vector was then used in a level 2 assembly with a diatom conjugation vector (centromere-oriT-Shble) and two spacer vectors (noncoding sequence used to fill a space in an assembly) to create a single conjugatable vector expressing a signal peptide-IBP fusion construct that can be used in *P. tricornutum*. DNA parts were amplified using PrimeStar Polymerase and purified using a DNA Clean & Concentrator kit from Zymo Research. Each L0 part was sequenced before further assemblies were performed to ensure that no mutations were introduced during PCR or cloning. Colony PCR, which was performed to double check expression cassette size, used SapphireAmp Fast polymerase.

Conjugation Method

Conjugations between *E. coli* harboring a pTA-Mob mobilization vector, to transfer the IBP expression plasmids, and *P. tricornutum* were performed as previously described in (Diner et

al. 2016). Three positive isolates for each vector containing mRuby3 were used in microscopy experiments to ensure that each of the signal peptides was localizing to the predicted cellular location. Once each signal peptide was confirmed to be correct, the signal peptide-IBP vectors lacking the mRuby3 tag were conjugated into *P. tricornutum* using the same method as above. Effectiveness of the transformation was determined by: A.) Visualization of the fluorescent protein tag in localized target of transformant cultures. B.) Increased ability of transformant to withstand freeze thaw cycles over wildtype *P. tricornutum*.

Freeze-thaw recovery assays

Transformants and control lines were grown up in 25ml of F/2 medium with 880 µM ammonium with zeocin added to transformant and mRuby3 control lines and placed in an incubator with 50 µE m-2s-1 light (14:10 diel cycle) at 18˚C. At late exponential phase, 100,000 cells were collected from each culture and transferred into 1 ml of F/2 medium with 880 µM nitrate in a 2ml microcentrifuge tube and placed in the incubator for 24 hours. Experiment cultures were then placed in a -20 ˚C freezer for 2 hours. After freezing, they were thawed at room temperature for 30 minutes. For growth recovery experiments, 5000 cells (~50ul) from each 1ml thawed culture, were transferred into 25mls of F/2 medium with 880 µM Nitrate (no zeocin). Cultures were placed in an incubator in the same conditions described above. Cell counts and fluorescence were measured daily for indication of growth and recovery.

For assays including the triple stacked trait transformants, the above freeze thaw experiment was followed with the following adjustments: Transformants and control lines were grown up in 25ml of F/2 medium with 880 µM nitrate with zeocin added to transformant cultures. During exponential phase 1.5 million cells were collected, spun down at 3000rpm for 10 mins, and

diluted into 3mls of fresh media and placed in 5 ml culture tubes for the duration of the freeze-thaw assay. After which 20,000 cells were plated on F/2 medium plates with 8.8 mM of Nitrate (no zeocin). The remaining cultures were analyzed for the maximum quantum efficiency of photosystem II photochemistry (Fv/Fm).

Results

The C-Terminal His-FLAG tag, no signal peptide IBP, yielded the greatest amount of protein in the soluble fraction (Fig 2.). Clones containing C-Terminal His-Flag tags with full-length IBP as well as N-Terminal His FLAG and N-Terminal His-Flag + thioredoxin with no signal peptide IBP also produced smaller yields of soluble protein. Proteins were around 24kDA in size with the thioredoxin tag added vectors slightly larger. All clones with expressed protein and NHT control were grown up into 5ml cultures overnight, spun down, frozen for antifreeze activity and thermal hysteresis assays. Thermal hysteresis (TH) activity of clones averaged around 0.30 C with no significant differences in attached tags. This TH effect is found to be comparable to other moderate fcIBP TH activities reported in (Kondo et al. 2018).

For the *P. tricornutum* transformants, we visualized IBPs in the targeted nucleus, cytoplasm, chloroplast, chloroplast (pyrenoid compartment), and mitochondria localizations (Fig 3.). Imaging of the mRuby3 (green) revealed a strong overlap with the plastid autofluorescence (red), nucleus (blue, Hoechst 33342) and mitochondria (orange, MitoTracker orange) for the chloroplast, nucleus and mitochondrial *P. tricornutum* lines, respectively (Fig 3.). The secreted and ER transgenic lines were ambiguous in IBP presence as expected in the former and with the latter forming a mRuby3 fluorescent pattern resembling targeted ER compartmentalization from other studies.

The nucleus, cytoplasm, chloroplast, chloroplast (pyrenoid compartment) IBP transformants showed significant growth recovery during exponential phase over the wtPt, mRuby3 and NH^4 controls after the freeze-thaw period (Fig 4A.). The secretion, ER, mitochondria IBP transformants showed improved, but not significant, growth recovery over controls (Fig 4B.). The controls had significantly higher standard deviation values over the transformants (Fig 4.). The afore-mentioned results were used in the development of a triple stacked trait *P. tricornutum* transformant that included the three chloroplast, cytoplasm and nucleus IBP localizations. Cytoplasm-chlorophyll-nucleus IBP stacked lines displayed a significant increase in Fv/Fm values (p-value 0.0001) (Fig 5A.) and also in colony forming unit (CFU) counts (p-value 0.03) over the wtPt controls (Fig 5B.).

Discussion

The native signal peptide on the fcIBP 268004 is responsible for the secretion the IBP into the extracellular matrix where it protects the cell from lysis by an external ice crystal. Ice crystals can form inside the cell particularly in rapid cooling conditions (McLellan 1989). Many organisms have IBPs localized intracellularly inhibiting ice crystal formation and preventing internal cell damage (Ahlgren et al. 1988, Kristiansen et al. 1999, Kawahara et al. 2007). Our strategy in reducing freeze thaw mortality was to employ the use of extracellular and intracellular native *P. tricornutum* signaling peptides to protect cell critical structures from ice crystal growth. A growth recovery experiment showed a significant recovery in the chloroplast, cytoplasm and nucleus localized IBP transformants over the wild type *P. tricornutum* after freeze thaw exposure. Additionally, these transformants outperformed ER, secretory and mitochondrial localized IBP transformants.

While the secreted IBP provides only extracellular protection for the cell, the internalized structures are still at risk. One reason why the intracellular localized IBP lines may have an advantage is the possibility of external and internal IBPs protection of the cell, with the external IBPs being the result of any mortality events and lysis of transformants. The advantage of the cytoplasm, chlorophyll and nucleus lines over might be temperate diatom specific over other targeted organelles as the extent of freeze thaw damage to internalized structures varies from organism to organism (Mohr and Stein 1970, Anjum 2003, Yang et al. 2020) and whether other mechanisms of cryoprotection are present (Schilling et al. 2002). Freezing temperature and duration heavily impact viability of cells and structural damage (McLellan 1989).

In the case of diatoms, limited ultrastructure studies have revealed several structural changes as a result of exposure to freezing temperatures and subsequent thawing (McLellan 1989). After freezing at 0.5 ˚C per minute to -40 ˚C, diatoms from the genus *Stephanodiscus* revealed a less dense and granular cytoplasm, disruption of lamellar organization and membrane disorganization in chloroplasts, vacuolar leakage, indiscernible mitochondria and other unrecognizable, highly disorganized organelles (McLellan 1989). These structural damages, most likely the result of mechanical shearing due to ice crystal formation, are possibly alleviated in the transgenic lines by the IBP-mediated ice crystal growth inhibition.

Conclusion

The data in this study provides a more detailed glimpse into the adaptive strategies of polar sea ice diatoms by examining the effectiveness of a known coping strategy, ice binding proteins, in conferring freeze thaw resistance. The findings of significant reductions in mortality and improved recovery of IBP transformant lines over the wildtype *P.tricornutum* display the

essentiality of this ice active molecule in preventing freezing damage to cells. In addition, we show an improved freeze resistance using targeted compartmentalization signal peptides of the IBP over native extracellular localizations. The aforementioned insights into the generation of artificial extremophiles may one day be useful in future biotechnological and agricultural tool development in preventing low temperature cell lysis and eventual transfer into industry relevant species.

Acknowledgements

Chapter three, in part is currently being prepared for submission for publication of the material. A. Zoumplis, E. Garza, D. Yee, A. DeVries, V. Bielinski, A. E. Allen. "Heterologous expression and targeted delivery of sea ice diatom ice binding proteins." The dissertation author was the primary investigator and author of this paper.

References

1. Ahlgren, J. A., C. Cheng, J. Schrag & A. DeVries (1988) Freezing avoidance and the distribution of antifreeze glycopeptides in body fluids and tissues of Antarctic fish. *Journal of Experimental Biology,* 137**,** 549-563.

2. Anjum, M. (2003) Cellular damage to the callus cells of potato subjected to freezing. *Biologia plantarum,* 46**,** 1-6.

3. Apt, K. E., L. Zaslavkaia, J. C. Lippmeier, M. Lang, O. Kilian, R. Wetherbee, A. R. Grossman & P. G. Kroth (2002) In vivo characterization of diatom multipartite plastid targeting signals. *Journal of Cell Science,* 115**,** 4061-4069.

4. Bayer-Giraldi, M., C. Uhlig, U. John, T. Mock & K. Valentin (2010) Antifreeze proteins in polar sea ice diatoms: diversity and gene expression in the genus Fragilariopsis. *Environmental Microbiology,* 12**,** 1041-1052.

5. Daboussi, F., S. Leduc, A. Marechal, G. Dubois, V. Guyot, C. Perez-Michaut, A. Amato, A. Falciatore, A. Juillerat & M. Beurdeley (2014) Genome engineering empowers the diatom Phaeodactylum tricornutum for biotechnology. *Nature communications,* 5**,** 1-7.

6. Diner, R. E., V. A. Bielinski, C. L. Dupont, A. E. Allen & P. D. Weyman (2016) Refinement of the diatom episome maintenance sequence and improvement of conjugation-based DNA delivery methods. *Frontiers in bioengineering and biotechnology,* 4**,** 65.

7. Engler, C. & S. Marillonnet. 2014. Golden gate cloning. In *DNA cloning and assembly methods*, 119-131. Springer.

8. Erdene-Ochir, E., B.-K. Shin, B. Kwon, C. Jung & C.-H. Pan (2019) Identification and characterisation of the novel endogenous promoter HASP1 and its signal peptide from Phaeodactylum tricornutum. *Scientific reports,* 9**,** 1-10.

9. Ferrarezi, R. S., K. Rodriguez & D. Sharp (2020) How historical trends in Florida all-citrus production correlate with devastating hurricane and freeze events. *Weather,* 75**,** 77-83.

10. Fiala, M. & L. Oriol (1990) Light-temperature interactions on the growth of Antarctic diatoms. *Polar Biology,* 10**,** 629-636.

11. Garcia-Lopez, E., P. Alcazar & C. Cid (2021) Identification of Biomolecules Involved in the Adaptation to the Environment of Cold-Loving Microorganisms and Metabolic Pathways for Their Production. *Biomolecules,* 11**,** 1155.

12. Gatenby, C. M., D. M. Orcutt, D. A. Kreeger, B. C. Parker, V. A. Jones & R. J. Neves (2003) Biochemical composition of three algal species proposed as food for captive freshwater mussels. *Journal of Applied Phycology,* 15**,** 1-11.

13. Gavazov, K., J. Ingrisch, R. Hasibeder, R. T. Mills, A. Buttler, G. Gleixner, J. Pumpanen & M. Bahn (2017) Winter ecology of a subalpine grassland: effects of snow removal on soil respiration, microbial structure and function. *Science of the Total Environment,* 590**,** 316-324.

14. Gordon, R. & R. W. Drum (1970) A capillarity mechanism for diatom gliding locomotion. *Proceedings of the National Academy of Sciences,* 67**,** 338-344.

15. Ito, A., T. Sakaue, O. Fujimaru, A. Iwatani, T. Ikeda, D. Sakamoto, T. Sugiura & T. Moriguchi (2018) Comparative phenology of dormant Japanese pear (Pyrus pyrifolia) flower buds: a possible cause of 'flowering disorder'. *Tree physiology,* 38**,** 825-839.

16. Janech, M. G., A. Krell, T. Mock, J. S. Kang & J. A. Raymond (2006) Ice-binding proteins from sea ice diatoms (bacillariophyceae) 1. *Journal of Phycology,* 42**,** 410-416.

17. Kalderon, D., B. L. Roberts, W. D. Richardson & A. E. Smith (1984) A short amino acid sequence able to specify nuclear location. *Cell,* 39**,** 499-509.

18. Karas, B. J., R. E. Diner, S. C. Lefebvre, J. McQuaid, A. P. Phillips, C. M. Noddings, J. K. Brunson, R. E. Valas, T. J. Deerinck & J. Jablanovic (2015) Designer diatom episomes delivered by bacterial conjugation. *Nature communications,* 6**,** 1-10.

19. Kawahara, H., Y. Iwanaka, S. Higa, N. Muryoi, M. Sato, M. Honda, H. Omura & H. Obata (2007) A novel, intracellular antifreeze protein in an antarctic bacterium, Flavobacterium xanthum. *Cryoletters,* 28**,** 39-49.

20. Kilian, O. & P. G. Kroth (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *The Plant Journal,* 41**,** 175-183.

21. Kondo, H., K. Mochizuki & M. Bayer-Giraldi (2018) Multiple binding modes of a moderate ice-binding protein from a polar microalga. *Physical Chemistry Chemical Physics,* 20**,** 25295-25303.

22. Krell, A., B. Beszteri, G. Dieckmann, G. Glöckner, K. Valentin & T. Mock (2008) A new class of ice-binding proteins discovered in a salt-stress-induced cDNA library of the psychrophilic diatom Fragilariopsis cylindrus (Bacillariophyceae). *European Journal of Phycology,* 43**,** 423-433.

23. Kristiansen, E., S. Pedersen, H. Ramløv & K. Zachariassen (1999) Antifreeze activity in the cerambycid beetle Rhagium inquisitor. *Journal of Comparative Physiology B,* 169**,** 55-60.

24. Kroth, P. G., A. Chiovitti, A. Gruber, V. Martin-Jezequel, T. Mock, M. S. Parker, M. S. Stanley, A. Kaplan, L. Caron & T. Weber (2008) A model for carbohydrate metabolism in the diatom Phaeodactylum tricornutum deduced from comparative whole genome analysis. *PloS one,* 3**,** e1426.

25. Levitan, O., J. Dinamarca, G. Hochman & P. G. Falkowski (2014) Diatoms: a fossil fuel of the future. *Trends in biotechnology,* 32**,** 117-124.

26. Liang, B. W. 2021. *The Effect of Freeze-Thaw Cycles on Microbial Resilience Along a Crop Biodiversity Gradient*. Michigan State University.

27. Liu, J., B. Ulén, G. Bergkvist & H. Aronsson (2014) Freezing–thawing effects on phosphorus leaching from catch crops. *Nutrient Cycling in Agroecosystems,* 99**,** 17-30.

28. Lizotte, M. P. (2001) The contributions of sea ice algae to Antarctic marine primary production. *American zoologist,* 41**,** 57-73.

29. Lundholm, N., N. Daugbjerg & Ø. Moestrup (2002) Phylogeny of the Bacillariaceae with emphasis on the genus Pseudo-nitzschia (Bacillariophyceae) based on partial LSU rDNA. *European Journal of Phycology,* 37**,** 115-134.

30. McLellan, M. (1989) Cryopreservation of diatoms. *Diatom Research,* 4**,** 301-318.

31. Mock, T. & D. N. Thomas (2005) Recent advances in sea-ice microbiology. *Environmental Microbiology,* 7**,** 605-619.

32. Mock, T. & K. Valentin (2004) PHOTOSYNTHESIS AND COLD ACCLIMATION: MOLECULAR EVIDENCE FROM A POLAR DIATOM 1. *Journal of Phycology,* 40**,** 732-741.

33. Mohr, W. & M. Stein (1970) Ultrastructural damage in "protectively" treated, frozen-thawed, tomato fruit parenchyma. *Cryobiology,* 7**,** 113-121.

34. Morgan-Kiss, R. M., J. C. Priscu, T. Pocock, L. Gudynaite-Savitch & N. P. Huner (2006) Adaptation and acclimation of photosynthetic microorganisms to permanently cold environments. *Microbiology and molecular biology reviews,* 70**,** 222-252.

35. Pelster, D. E., M. H. Chantigny, P. Rochette, N. Bertrand, D. A. Angers, B. J. Zebarth & C. Goyer (2019) Rates and intensity of freeze–thaw cycles affect nitrous oxide and carbon dioxide emissions from agricultural soils. *Canadian Journal of Soil Science,* 99**,** 472-484.

36. Pollak, B., T. Matute, I. Nuñez, A. Cerda, C. Lopez, V. Vargas, A. Kan, V. Bielinski, P. von Dassow & C. L. Dupont (2020) Universal loop assembly: open, efficient and cross-kingdom DNA fabrication. *Synthetic Biology,* 5**,** ysaa001.

37. Raymond, J. A. & H. J. Kim (2012) Possible role of horizontal gene transfer in the colonization of sea ice by algae. *PloS one,* 7**,** e35968.

38. Savitsky, P., J. Bray, C. D. Cooper, B. D. Marsden, P. Mahajan, N. A. Burgess-Brown & O. Gileadi (2010) High-throughput production of human proteins for crystallization: the SGC experience. *Journal of structural biology,* 172**,** 3-13.

39. Schilling, S. M., H. A. Sror, D. K. Hincha, J. M. Schmitt & C. A. Köhn. 2002. Cryoprotectin, A Cabbage Protein Protecting Thylakoids from Freeze-Thaw Damage. In *Plant Cold Hardiness*, 195-210. Springer.

40. Tachibana, M., A. E. Allen, S. Kikutani, Y. Endo, C. Bowler & Y. Matsuda (2011) Localization of putative carbonic anhydrases in two marine diatoms, Phaeodactylum tricornutum and Thalassiosira pseudonana. *Photosynthesis Research,* 109**,** 205-221.

41. Teufel, A. G. & R. M. Morgan-Kiss. 2018. Physiological and Biochemical Adaptations of Psychrophiles. In *Extremophiles*, 185-208. CRC Press.

42. Tréguer, P., C. Bowler, B. Moriceau, S. Dutkiewicz, M. Gehlen, O. Aumont, L. Bittner, R. Dugdale, Z. Finkel & D. Iudicone (2018) Influence of diatom diversity on the ocean biological carbon pump. *Nature Geoscience,* 11**,** 27-37.

43. Uhlig, C., F. Kilpert, S. Frickenhaus, J. U. Kegel, A. Krell, T. Mock, K. Valentin & B. Beszteri (2015) In situ expression of eukaryotic ice-binding proteins in microbial communities of Arctic and Antarctic sea ice. *The ISME journal,* 9**,** 2537-2540.

44. William, K. L. & I. Morris (1982) Temperature adaptation in Phaeodactylum tricornutum Bohlin: photosynthetic rate compensation and capacity. *Journal of Experimental Marine Biology and Ecology,* 58**,** 135-150.

45. Willick, I. R., K. K. Tanino & L. V. Gusta (2021) The impact of global climate change on the freezing tolerance of winter cereals in Western Canada. *Journal of Agronomy and Crop Science,* 207**,** 88-99.

46. Wipf, S., M. Sommerkorn, M. I. Stutter, E. J. Wubs & R. Van Der Wal (2015) Snow cover, freeze-thaw, and the retention of nutrients in an oceanic mountain ecosystem. *Ecosphere,* 6**,** 1-16.

47. Yang, C., J. Peng, X. Li, D. Liang, Z. Yang & Y. Zhang (2020) The mechanism underlying overwintering death in poplar: the cumulative effect of effective freeze–thaw damage. *Journal of Forestry Research,* 31**,** 219-229.

Figures

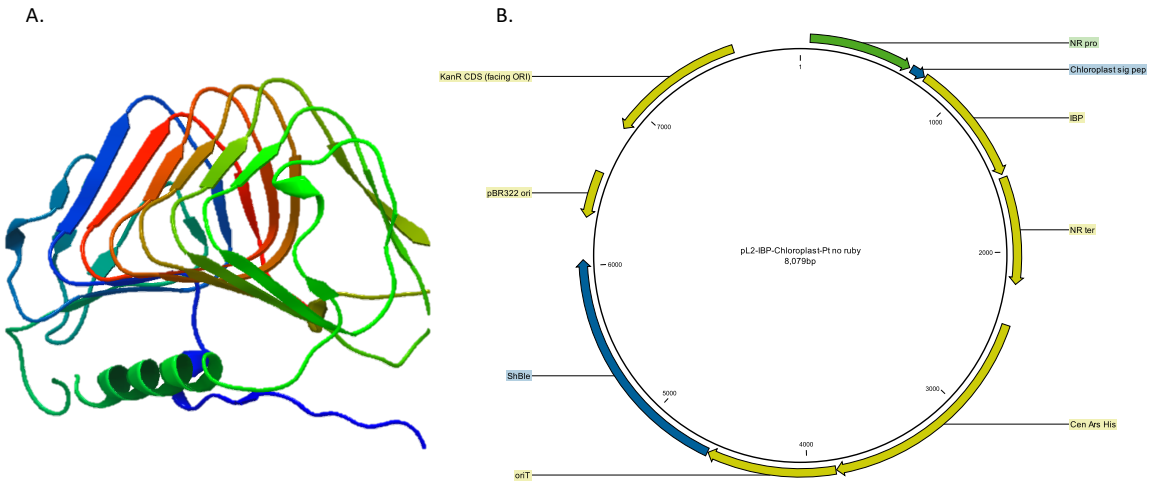

Figure 3.1 (**A**) A Phyre2 protein structure blast result for *Fragilariopsis cylindrus* ice binding protein (fcIBP ID 268004) revealed similarities to a *Flavobacterium* antifreeze protein. The presence of Beta-sheets indicate strong hydrogen bonding potential of this protein to ice crystal planes. **(B)** The plasmid construct used in transforming the fcIBP into *Phaeodactylum tricornutum*.

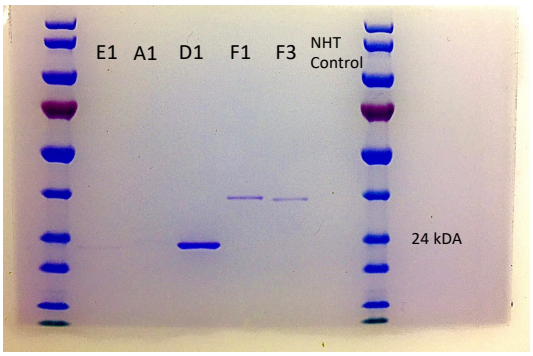

Figure 3.2 (**A**) SDS-PAGE gel visualizing soluble fraction of protein (fcIBP ID 268004) from *E.coli* transformants. Expression vectors coincide with labels as follows: **E1** IBP with N terminal His Flag, **A1/D1** IBP with C terminal His Flag, **F1/F3** IBP + thioredoxin, **NHT control** N-Terminal His Flag no IBP.

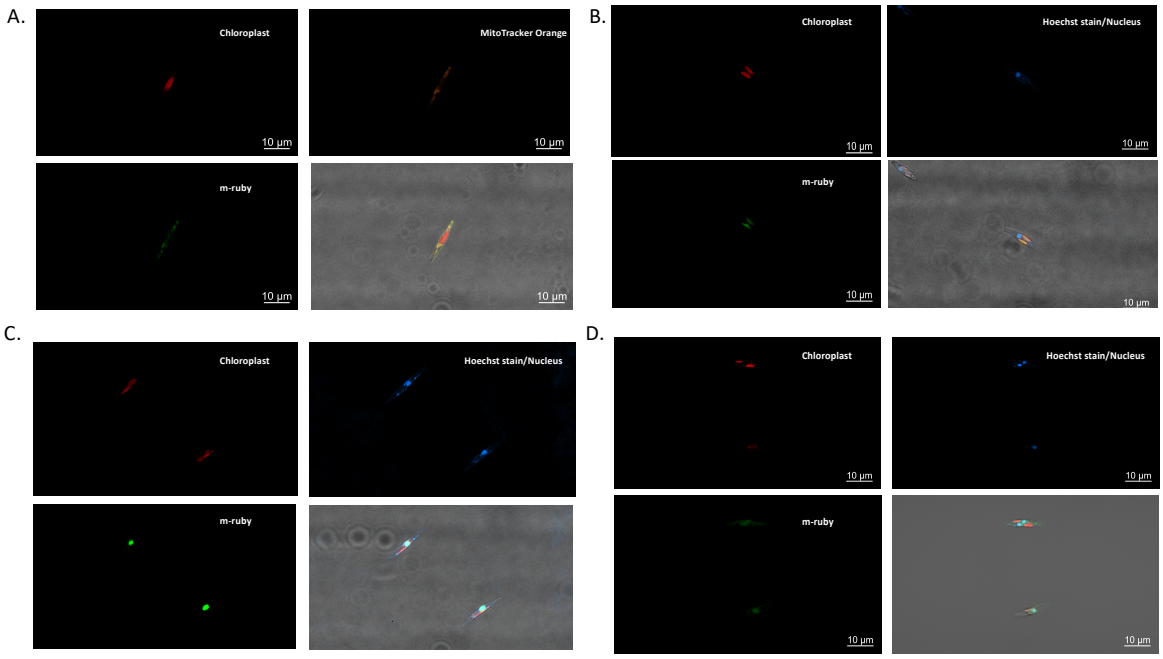

Figure 3.3 Imaging of fcIBP targeted localizations depicted by the green mRuby3 fluorescence for the **(A)** mitochondria **(B)** chloroplast **(C)** nucleus and **(D)** cytoplasm transformants.

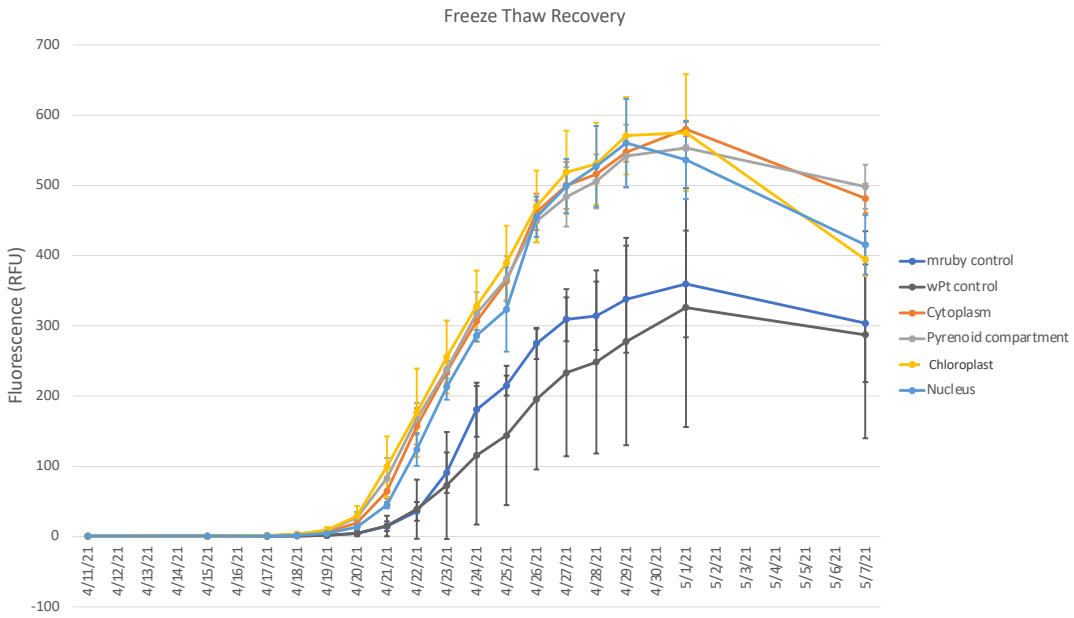

Figure 3.4 Growth curves measured by fluorescence with standard deviation error bars showing recovery of cytoplasm, chloroplast, chloroplast (pyrenoid compartment) transformant lines over wildtype (wtPt) and mRuby3 control lines after a freeze thaw period.

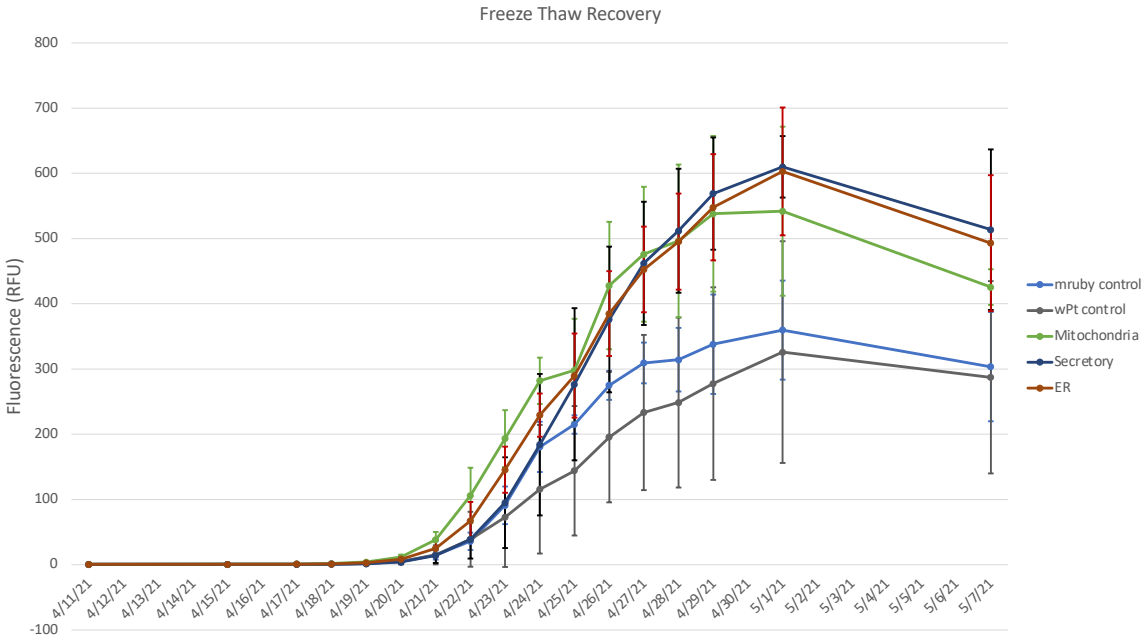

Figure 3.5 Growth curves measured by fluorescence with standard deviation error bars showing recovery of mitochondria, secretory and ER transformant lines over wildtype (wtPt) and mRuby3 control lines after a freeze thaw period.

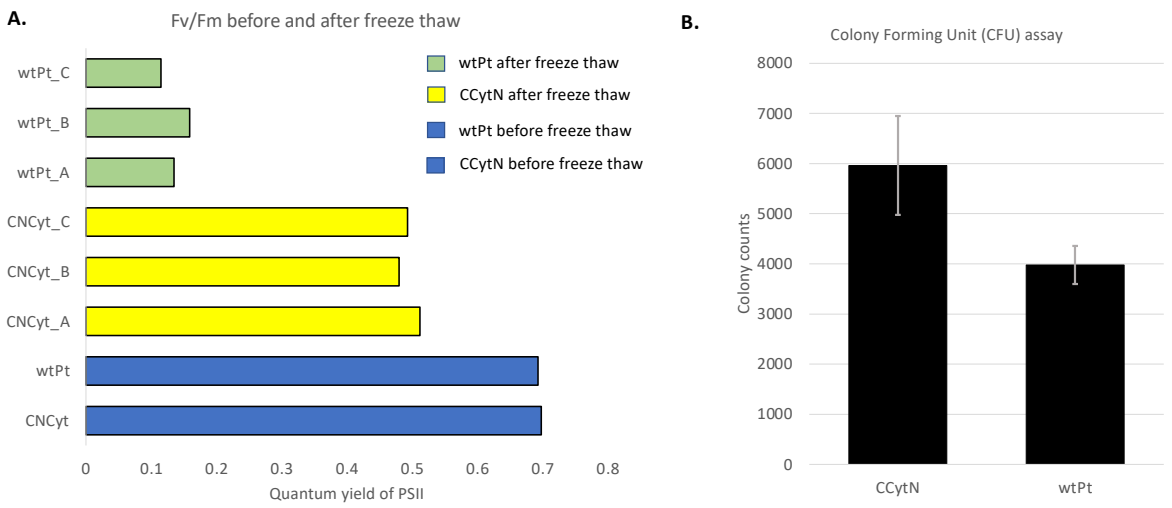

Figure 3.6 (A) Comparison of quantum yields of PSII and **(B)** colony forming units (CFUs) in wildtype (wtPt) and chloroplast-nucleus-cytoplasm (CNCyt) localized fcIBP transformants before and after freeze thaw event.

Supporting Information File 1

Small-scale Expression Testing of E coli Clones

Start overnight culture in 4 mL LB with appropriate antibiotic for selection from glycerol stock or culture plate (for pBAD, use 10 ug/ml Tet)
Incubate overnight at 30oC, shaking at 225 rpm
On day of expression, dilute overnight culture 1:20 into Terrific Broth with antibiotic (10mL total volume, so 0.5 mL cells + 9.5 mL TB, can be done in 50 mL conical vial)
Grow for 4 hrs at appropriate temp
Transfer culture to 18oC incubator and lower shaking to 200 rpm for additional 1 hr
Add L-arabinose to a final volume of 0.5% (250uL of 20% L-arabinose in 10mL culture)
Allow cultures to grow with arabinose overnight at 18oC
Pellet cells with 10 min spin at 3000 x g
Either freeze away pellets at -80oC or continue with processing (below)

Cobalt resin pulldown protocol for E coli

Sonication Lysis Buffer for E coli and Pt
50mM NaPO4 (pH 7.5-8.0)
500mM NaCl
0.1% Trition X-100
10mM imidazole
1 mg/mL lysozyme
10uM B-ME
1 EDTA-free Protease inhibitor tablet/10mL buffer

Wash Buffer
50mM NaPO4 (pH 7.5-8.0)
500mM NaCl
30mM imidazole

Elution Buffer
50mM NaPO4 (pH 7.5-8.0)
500mM NaCl
250mM imidazole

For chemical lysis of E coli, use B-PER Bacterial Protein Extraction Reagent (Thermo cat# 72248) supplemented with:
1mg/mL lysozyme
2U/ml DNAse or Benzonase

Also needed: Talon resin (Clonetech cat #635501)

Note: Talon resin binds best at pH 8.0, is sensitive to EDTA and high amounts of reducing reagents. Use TCEP if you want to get into the 1mM range, but lower amounts work just fine for us at this scale. 5% glycerol can also be added to all 3 buffers for sonication if necessary (depends on your protein).

Protocol:

1. Harvest 10mL overnight expression culture (E coli) or late log phase Pt by centrifugation for 10 min at 3000 x g
2. Remove supernatant by decanting and resuspend pellet in 800ul lysis buffer and transfer to a new 1.5mL Eppendorf tube
3. For B-PER chemical lysis, incubate tubes at RT for 15min with mixing
4. For sonication, hit with 6x 15 sec pulses on ice (use sonicator in Syn Bio hood room, can't change the settings so use what it is, then do 15sec on/15sec off pulses - your Pt should look like iced tea once fully lysed, nice & clear)
5. After lysis, clarify lysates by centrifuging at 4oC for 10 min at 12000 x g
6. While lysates are spinning, equilibrate the Talon resin by washing 3x with 1-5 volumes of resin bedding
7. For each pulldown you do, you will need 25ul of Talon resin. This comes supplied as a 50% slurry in ethanol, so calculate how many pulldowns you will be doing, then add 2. So, if you have 8, estimate you will need for 10 rxns.
   - 7a. For 10 rxns, shake the slurry well and remove 500ul resin slurry and transfer to new Eppendorf tube
   - 7b. Pellet resin by spinning in microfuge for 30sec at 3000 x g, then remove supernatant and add 1 mL of lysis buffer (B-PER or Phosphate buffer, whichever you used for lysis)
   - 7c. Vortex resin briefly, then pellet again and wash beads 2x with 1mL lysis buffer
   - 7d. After 3rd spin, remove supernatant and add 250uL of lysis buffer to produce 500uL of equilibrated 50% resin
8. Transfer 700ul of supernatant to new Eppendorf tube
9. Add 50uL equilibrated resin to tube and incubate for 1hr at RT with end over end mixing
10. After 1hr, pellet resin with 30sec 3000 x g spin at RT, then remove supernatant.
11. Wash resin 3x with 1mL of wash buffer, pelleting resin in between each washing and removing supernatant
12. After 3rd wash, remove supernatant and resuspend resin in 50uL of Elution buffer
13. Incubate elutions for 5 min at RT, mixing a few times by flicking the tube
14. Pellet resin with 30sec spin at 3000 x g, then carefully remove supernatant and transfer to a new Eppendorf tube
15. 10uL of bacterially expressed eluate run on SDS-PAGE gel should show you expressed protein via Coomassie stain

CONCLUSION

The microbial communities in polar environments perform vital functions that are critical for ecosystem stability. Climate driven changes in community structure, and keystone species, can disrupt ecosystem function as a whole leading to a potential collapse in a warming climate. Studies on the molecular diversity and adaptive strategies of remote, snow and ice free regions of Antarctica are limited even though these regions are hotspots of biodiversity. This dissertation set out to establish a baseline of molecular diversity and explore the functional roles of the microbiota inhabiting ephemeral polar ecosystems.

Polar microbes are incredibly resilient to rapid chemical and physical fluctuations; However, future warming projections and climate variation puts these ecosystems at risk for dramatic and persistent changes. In Chapters 1 and 2, we emphasize the distinct community structures and specialized functional roles of microbial communities across the Dry Valleys landscape. Hydrological drivers have heavily drive the community structuring of the Dry Valleys meltwater stream microbial mats. During periods of high flow, mats on the stream margin undergo a drastic overturn of taxa resulting a major decline of nitrogen-fixing cyanobacteria, *Nostoc.* At lower flow times, near the end of the austral summer season, the *Nostoc* recovered displaying a striking resilience of this keystone taxa.

A broadened survey across multiple aquatic and aeolian habitats in the Dry Valleys and nearby, adjacent marine systems revealed the influence of past climate events on the polar desert ecosystem. A dominant, intrusive marine community located at the terminus of the Taylor Glacier showed the long standing effects of climate warming and increased connectivity on the landscape biota. Thought to be a relic community, likely brought into the Dry Valleys through ancient flooding, the taxa at this location are sporadically supplied with nutrient and iron rich hypersaline

brine that creates a suitable habitat for non-endemic marine taxa. A metatransciptomics analysis revealed an upregulation of Ferritin-like genes that may be responsible for the formation of spores and cysts that are likely stress coping mechanisms of intrusive species. Both of these chapters describe resilience strategies and stress responses of polar organisms at a time of unprecedented landscape change as well as impacts of climate on ecosystem functioning.

Well-known coping strategies utilized by highly adaptive polar microbes in cold, ephemeral environments can be exploited for their use in anticipated climate change related events. Freeze-thaw cycling is one such damaging effect of climate variation that has the potential to be mediated through the use of bioproducts. Ice binding proteins are prevalent among sea ice microbes including polar diatoms. In Chapter 3 we explored the use of IBPs in protecting non-polar diatoms from freeze thaw events. Results showed significant reductions in mortality after a freeze thaw cycle among IBP transformed diatoms over wildtype diatoms. These findings reiterate the incredible efficacy of polar adaptive mechanisms and provides evidence for the ability to confer freeze resistance with the use of IBPs.

This dissertation provides insights into the climate driven patterns of ephemeral microbial community strucure and function across spatiotemporal scales. Data from this dissertation serves as a baseline of molecular biodiversity in climate sensitive regions. In addition, findings provide the framework for a better understanding of the complex ecophysiology and adaptive strategies of microbes supporting rapidly changing polar landscape ecosystems.