Genomic and Functional Investigations Into Seasonally-Impacted and Morphologically-Distinct Anoxygenic Photosynthetic Cyanobacterial Mats

by

Sharon L. Grim

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Earth and Environmental Sciences)
in the University of Michigan
2019

Doctoral Committee:

Associate Professor Gregory J. Dick, Chair
Professor Joel D. Blum
Assistant Professor Jena E. Johnson
Professor Patrick D. Schloss
Professor Nathan D. Sheldon
Associate Professor Stephen A. Smith

Sharon L. Grim

sgrim@umich.edu

ORCID iD: 0000-0003-4215-7519

© Sharon L. Grim 2019

Dedication

To my family: my husband Dack Stuart, my son Wesley Stuart Grim,

my sister Karen, and my parents David and Sam Grim.

Acknowledgements

This dissertation would not have been possible without the support, engagement, and assistance of several people over the last five years.

My advisor Greg Dick has been a thoughtful, dedicated, patient, and enthusiastic mentor as I became more familiar with geomicrobiology, learned new techniques, and worked through the ups and downs of my dissertation project. My committee members Joel Blum, Jena Johnson, Pat Schloss, Nathan Sheldon, and Stephen Smith have been invaluable resources and perspectives.

Past and current members of the geomicrobiology lab have been great friends and provided feedback on my work. Thanks to Paul den Uyl, Robert Hein, Sunit Jain, Lauren Kinsman-Costello, Judith Klatt, Matthew Medina, Kevin Meyer, McKenzie Powers, Cody Sheik, Derek Smith, Hui Chien Tan, and Colleen Yancey. Katy Rico has been a wonderful field-sampling partner and conference travel buddy since we started at UM. Phoebe Aron and Naomi Levin provided water stable isotope analyses and interpretations. The labs of Tom Johengen and Allen Burton supported us in the field. The joint lab groups of Vincent Denef and Melissa Duhaime provided interdisciplinary and excellent ecological perspectives.

Collaborators provided key feedback and analyses that improved our research. Bopaiah Biddanda and his lab at Grand Valley State University have been our field partners in the Middle Island Sinkhole for many years. Jacob Waldbauer at University of Chicago provided proteomic

analyses and interpretations. Our collaborators at MPI-Bremen, Dirk de Beer and Arjun Chennu, enabled unprecedented insights into the geochemistry of MIS.

This work would not have been possible without the NOAA Thunder Bay National Marine Sanctuary dive team of John Bright, Tane Casserley, Stephanie Gandulla, Russ Green, Phil Hartmeyer, and Wayne Lusardi who sampled for us over multiple years. Captains Steve Bawks, Travis Smith, and Andrew Yagiela made sure our expeditions got there safely. I deeply appreciate their support, cheerfulness, and excitement about MIS. Thanks also to NOAA Great Lakes Environmental Research Laboratory members Steve Ruberg and Dack Stuart for field equipment, sampling, and logistic support.

Members of the Earth and Environmental Sciences department have cultivated a caring and supportive environment. Thanks to the graduate students for all the fun and engaging activities and events. Thanks also to the office staff, especially Anne Hudon, Stacy Wilkin, and Paula Frank, for keeping the graduate students sane and well nourished.

I am grateful to funding agencies for supporting my dissertation research. My time at UM has been funded by a Novak Fellowship and the Rackham Predoctoral Fellowship, as well sa NSF Grant EAR-1637066 to my advisor. The Rackham Research Grant, Rackham Conference Travel Grant, and Scott Turner Awards enabled me to expand my research and present it to my colleagues.

Last, I want to thank my family for seeing me through this adventure. My sister Karen and my parents David and Sam have cheered me on throughout the program. My son Wesley made our time here even more memorable. Finally, my husband Dack has been a pillar of support and the embodiment of patience as we established new roots in Ann Arbor. Without you, life as a PhD student would have been a little less sweet and a lot more stressful.

Table of Contents

List of Tables

List of Figures

Abstract

Cyanobacteria are key members of modern photosynthetic microbial mats, by providing organic matter and nutrients to the ecosystem. With their innovation of oxygenic photosynthesis (OP) over 2 billion years ago and the wide distribution of microbial mats in the geologic record, cyanobacteria have shaped Earth's redox history. Select modern cyanobacteria are also capable of anoxygenic photosynthesis using sulfide (AP), an older metabolism than OP and a less-understood biological mechanism that limited the rise of oxygen. This dissertation used molecular and ecological techniques to investigate modern anoxygenic cyanobacteria. In **Chapter I**, I outline the current understanding of AP cyanobacterial mats in modern and ancient ecosystems. In **Chapter II**, I described the genome of a cultured AP cyanobacterium, *Geitlerinema* sp. PCC9228. Genomic analysis of *Geitlerinema* revealed numerous adaptations to low-oxygen and sulfidic conditions, which were potentially prevalent for much of Earth's history. I applied knowledge of cyanobacterial genetic adaptation to AP from *Geitlerinema*, to the microbial mats of Middle Island Sinkhole (MIS), a submerged sinkhole impacted by low-O_2, sulfur-rich groundwater. In **Chapter III**, I characterized the impact of seasonally changing light conditions and geochemistry on the microbial community (16S rRNA genes and metagenomics) and its function (metaproteomics). The dominant AP cyanobacteria, *Phormidium* and *Planktothrix*, are abundant and active in summer when light is highest. In contrast, when light is lower in autumn, sulfide-oxidizing bacteria are more active. The shift in microbial community function has implications for oxygen and sulfur cycling in the mat. In **Chapter IV**, I applied

metagenomics and metaproteomics to distinct mat morphotypes in MIS. I observed flat purple cyanobacterially-dominated mat (‘flat’), conical purple mat (‘fingers’), white-pigmented mat (‘white’), and a mottled purple/brown pigmented mat (‘giraffe’). The cyanobacterial community shifted from *Phormidium* and *Planktothrix* in fingers and some flat mat, to *Pseudanabaena* and *Spirulina* in giraffe and white mats. Sulfide-oxidizing bacteria and sulfate-reducing bacteria were widely observed in all mat types. The genomes and proteins suggest functional niche similarity with regard to sulfide and oxygen cycling between abundant cyanobacteria in different mats. Understanding the role of cyanobacteria in shaping the function and appearance of modern microbial mats informs the interpretation of the chemical environment, metabolisms, and biogeochemical impact of microbial mats through Earth history.

Chapter I: Introduction

Through their wide geographical range, diverse metabolisms, and prevalence in geologic history, photosynthetic microbial mats have fundamentally transformed Earth's surface. Cyanobacteria are pivotal members of these mats through their hallmark metabolism, oxygenic photosynthesis (OP), which provides O_2 and organic carbon that sustains a network of interacting microbial metabolisms within mats. The innovation of OP was instrumental in Earth's oxidative shift in the Proterozoic Era, and the proliferation of free O_2 in the atmosphere shaped the nitrogen, carbon, sulfur, iron, and oxygen cycles (Lyons et al., 2014). Widely available O_2 is linked to the maturation of the sulfur cycle, and the emergence of aerobic heterotrophy and other metabolisms that continue to function in extant mats (Butterfield, 2009; Catling et al., 2005; Fike et al., 2015). In modern microbial mats we continue to observe dynamic microbially-mediated redox conditions similar to those of early Earth. Exploring contemporary microbial mats as analogs to the ancient mats that proliferated on Earth provides insight into the organisms and metabolisms responsible for the development and moderation of biogeochemical cycling in past and current systems.

1.1 The role of cyanobacteria in oxygen and sulfur cycling

Oxygenic photosynthesis was the most influential biological mechanism behind the significant accumulation of free O_2 in the atmosphere during the Great Oxidation Event (GOE) 2.4-2.1 billion years ago (Ga) (Lyons et al., 2014). The coupling of two photosystems to oxidize H_2O may have occurred as early as the Archean (Planavsky et al., 2014; Xiong et al., 2000), but geochemical conditions permitted substantial permanent atmospheric O_2 accumulation only in the Proterozoic (Holland, 2002). Though the GOE refers to Earth's first widespread shift in oxidation, the biological and geological processes responsible for the subsequent 2 Ga of protracted and low atmospheric oxygenation until the rise to near modern levels during the Neoproterozoic Oxidation Event 0.6 Ga (Canfield & Teske, 1996; Fike et al., 2015), are still under debate.

A biological mechanism that may have slowed the rise of O_2 is anoxygenic photosynthesis (AP) using sulfide as the electron donor (Johnston et al., 2009). AP relies upon H_2S and other electron donors that are easier to oxidize than H_2O, the electron donor for OP. This metabolism transfers electrons from H_2S oxidation to photosystem I, producing elemental sulfur instead of free O_2. Though euxinia was once thought to be pervasive and enduring throughout the Proterozoic (Canfield, 1998), more recent models suggest that ferruginous conditions dominated, with only episodic and/or regional euxinia (Planavsky et al., 2011; Poulton et al., 2010; Reinhard et al., 2013). With periods of euxinia throughout the Archean and Proterozoic, being equipped for sulfide exposure through AP would have conveyed a metabolic advantage to these flexible cyanobacteria in shifting redox conditions. Before the advent of OP, AP was likely an important source of organic matter to ecosystems (Hamilton et al., 2016; Johnston et al., 2009), but its contribution to modern ecosystems is not well known. The products

of AP, elemental sulfur or sulfur intermediates, may be recycled by AP organisms or other sulfur cycling microorganisms (Haas et al., 2018), whose products in turn stimulate AP. However, the role of this metabolism in the ancient sulfur cycle, the interactions between functional groups, and how they interface with environmental conditions are not well understood. Much remains unknown about the distribution of AP in ancient cyanobacteria, the impact of AP in modern ecosystems, and the influence of AP on overall O_2 production.

Modern cyanobacteria have a range of responses to sulfide. When exposed to even 1-10 uM of sulfide, enough to inhibit electron transport to photosystem II (Miller & Bebout, 2004; Oren et al., 1979), cyanobacteria may either completely shut down OP, partially perform OP, facultatively switch between OP and AP, or completely switch to AP (Cohen et al., 1986). In general, cyanobacteria that regularly experience sulfidic conditions are more tolerant of higher sulfide levels than those with infrequent or no exposure (Miller & Bebout, 2004). Sulfide also impacts functioning of other bacteria and mitochondria (Theissen et al., 2003). To detoxify sulfide, organisms can use sulfide-quinone reductase (SQR, encoded by *sqr*) to oxidize sulfide (Theissen et al., 2003). In AP cyanobacteria, SQR transfers electrons from H_2S oxidation to plastoquinone, which is the intersection point between photosystems I and II (Klatt, de Beer, et al., 2016). The different versions of *sqr* in cyanobacteria point to a complicated evolutionary history of AP. At least three variants of *sqr* have been identified in cyanobacteria (Dick et al., 2018; Grim & Dick, 2016): type I is most widely seen in AP cyanobacteria but is most closely related to green sulfur bacteria type IV (Marcia et al., 2010), type II likely originated in ancestral bacteria coping with sulfidic conditions in the late Proterozoic and is observed across different domains and phyla (Theissen et al., 2003), and type VI which is most closely related to *sqr* expressed in green sulfur bacteria at high sulfide levels (Chan et al., 2009). The diverse

phylogenetic history of cyanobacterial *sqr* suggests lateral gene transfer occurred between GSB and cyanobacteria in sulfidic environments, providing the ability to couple sulfide oxidation with phototrophic growth to cyanobacteria (Gregersen et al., 2011).

Recent research has further elucidated the physiological response of AP cyanobacteria to redox changes. Modern AP cyanobacteria need a prolonged sulfide 'priming' period to synthesize SQR (Bronstein et al., 2000; Klatt, Al-Najjar, et al., 2015), suggesting the time and concentration of H_2S exposure dictates the onset of AP. Further, within those capable of simultaneous AP and OP, O_2 production has been observed to decrease with increasing H_2S levels (Cohen et al., 1986; Klatt, Al-Najjar, et al., 2015; Klatt, de Beer, et al., 2016). Light may also influence the balance of AP and OP, due to OP needing both photosystems but AP requiring only photosystem I and thus less light to operate (Klatt, de Beer, et al., 2016; Klatt, Meyer, et al., 2016). These nuanced responses and adaptations to sulfide exposure reflect different modern habitats and redox conditions (Miller & Bebout, 2004), but draw upon adaptations that have been shaped through evolutionary history. However, our ability to infer sulfide physiologies and the distribution of AP cyanobacteria from the geologic record is extremely limited.

1.2 The geobiological role of microbial mats

Microbial mats are found in a wide range of ecosystems across Earth, from extremes of temperature in ice-covered Antarctic lakes (Sumner et al., 2015) and hydrothermal hot springs (Ward et al., 2006), to temperate coastal shelves (Noffke, 2010) and submerged sinkholes (Biddanda et al., 2006; Ruberg et al., 2008). Mats are hotspots of biogeochemistry in that they condense diverse microorganisms, metabolisms, and geochemical gradients into sub-millimeter scales (Dick et al., 2018). As well as being globally distributed, mats are found throughout Earth

history and are the oldest records of life (Bosak et al., 2013; Lalonde & Konhauser, 2015; Noffke & Awramik, 2013; Stal, 2012). Modern mats harbor physical characteristics and microbial morphotypes similar to fossil mats, and present tangible analogs to aid in understanding the geologic record (Noffke, 2010).

Photosynthetic microbial mats are a consortium of phylogenetically and functionally diverse microorganisms that depend upon each other for metabolism and biogeochemical cycling (Konhauser, 2009; Paerl et al., 2000). Cyanobacteria and other phototrophs dominate the surfaces of mats, where they receive light for photosynthesis. They are also pioneers and stabilizers of biofilms through production of extracellular polymers (Stal, 2012). Redox gradients, ultimately influencing availability of electron donors and acceptors, dictate the abundance and presence of other functional groups in the mat (Konhauser, 2009). Heterotrophs such as aerobic bacteria and sulfate-reducing bacteria depend on organic matter from phototrophs and the availability of their electron acceptor, such as O_2 or SO_4^{2-}. These as well are sourced from the environment or from other microbes such as cyanobacteria and sulfide oxidizers. As such, while laminated and structured microbial mats are used as modern day analogs to stromatolites and microbialites (Noffke & Awramik, 2013), just as often mats are an interwoven network of distinct microbes without clear physical separation between functional groups or layers (Konhauser, 2009; Stal, 2012).

Microbial mats are of interest in understanding metabolically flexible cyanobacteria because AP members are typically found in microbial mats (Dick et al., 2018). Additionally, all known AP cyanobacteria are filamentous, a morphology that has been observed throughout the fossil record (Schirrmeister et al., 2013). Filamentous cyanobacteria are a regular fixture of photosynthetic microbial mats, and include taxa such as *Planktothrix, Geitlerinema*, and

Phormidium sp. This polyphyletic morphology is also observed in sulfide-oxidizing bacteria, such as *Beggiatoa* sp. which are a staple of sulfidic microbial mats (Flood et al., 2014). *Beggiatoa* are found in both marine and freshwater systems, and use nitrate or oxygen to oxidize sulfide and other reduced forms of sulfur from either hydrothermal sources or sulfate-reducing bacteria in sediments (MacGregor et al., 2013; Mussmann et al., 2007; Sharrar et al., 2017). The sulfide-oxidizing enzyme SQR is often found in *Beggiatoa* (Sharrar et al., 2017), as well as the Sox gene pathway for the oxidation of thiosulfate and/or the reverse dissimilatory sulfite reductase (rDsr) pathway to oxidize elemental sulfur to sulfite (Mussmann et al., 2007). Both of these genes can be transferred laterally (Gregersen et al., 2011). In different settings, filamentous cyanobacteria and sulfide-oxidizing bacteria can promote the formation of mats of similar appearance—cyanobacteria in photic systems, *Beggiatoa* in deeper settings (Flood et al., 2014).

On a diel cycle, cyanobacteria and sulfide oxidizers experience redox shifts in response to the functioning or nightly shutdown of OP (and thus the delivery of O_2), and the infiltration of H_2S and methane from deeper in the mat. Both sulfide-oxidizing bacteria and cyanobacteria can migrate in response to their physicochemical environment. Movement of *Beggiatoa* members in the mat is primarily motivated by their dual requirements for sulfide and oxygen/nitrate (Flood et al., 2014). Typically, *Beggiatoa* and other filamentous sulfide-oxidizing bacteria will adjust their position with respect to redox shifts (Klatt, Meyer, et al., 2016), whereas cyanobacteria switch to fermentation when experiencing diel patterns of darkness (Stal, 2012). Cyanobacteria exhibit phototaxis, positioning themselves vertically in the mat to optimizing light harvesting (Biddanda et al., 2015). These migrations of bacteria are believed to produce layers in microbial mats, but the potential for preservation varies greatly and depends on the geochemical environment (Stal, 2012).

In these systems, microbes are both shapers of, and responders to, geochemical cycles. On a diurnal scale, mats can be oxic during the daytime due to photosynthesis, but with the depletion of light and/or increase in sulfide, anoxic or euxinic conditions can set in (Canfield & Marais, 1993). Conditions can also change seasonally, as light levels and nutrient availability change throughout the year (Paerl et al., 2000). Microbial mats experience redox challenges, such as euxinia or anoxia, on geologically short temporal scales such as diurnal or seasonal cycles, that are similar to those faced by ancestral microbial mats during Earth's large-scale geochemical shifts. Investigations into the functionality and assembly of microbial mats on these short temporal scales are thus windows into the ecology of ancient microbial mats.

1.3 Molecular methods to study anoxygenic photosynthetic cyanobacterial genes and ecology

Physiological and molecular studies on cultures built the foundation of our understanding of anoxygenic photosynthetic cyanobacteria. The canonical AP cyanobacterium, *Geitlerinema* sp. PCC 9228 (formerly *Oscillatoria limnetica*) was first cultured in the 1970s (Cohen et al., 1975), physiologically described throughout the 1980s (Cohen et al., 1986 and references therein), and used as a reference for the first *sqr* sequence in the 1990s (Arieli et al., 1994; Theissen et al., 2003). Throughout that time a handful of other cyanobacteria were tested for AP (Garlick et al., 1977) and sulfide tolerances were described for some cyanobacteria (Miller & Bebout, 2004). However, by nature these culture-based approaches are limited to studying the organisms capable of laboratory growth and manipulation, and their *ex situ* growth may not represent *in situ* conditions and ecological interactions. Molecular techniques provide genetic and genomic investigations into cultures as well as natural communities with AP cyanobacteria.

Meta-omics approaches harness modern high throughput nucleotide sequencing technology suitable to investigating model organisms and natural ecosystems (Dick & Lam, 2015). By targeting different molecular pools (DNA and proteins), meta-omics approaches can untangle the identity of microbes in a natural community, the likely metabolisms they are performing, and the expression of functions (Dick, 2019). DNA fingerprinting through marker gene (e.g. 16S rRNA gene) analysis is a survey of microbial members and answers the question: who is there? Metagenomics takes this a step further by analyzing the DNA content of whole communities, and informing us of potential function in a microbial community: what genes does a specific organism have? With modern Illumina-based sequencing, metagenomics is powerful enough to reconstruct known and novel microbial genomes from diverse environments, and broaden our scope of the tree of life (Hug et al., 2016). Metaproteomics relies upon metagenomic references and mass spectrometry to identify the pool of proteins in a whole community. The detection of a protein indicates actual expression of function: what proteins have been synthesized to satisfy biochemical requirements?

This dissertation uses meta-omics to investigate AP and OP cyanobacteria at each end of the spectrum of community complexity, from a cyanobacterial culture to a benthic microbial mat. To investigate the genetic repertoire of AP cyanobacteria, in **Chapter II** I evaluated a draft genome of *Geitlerinema* sp. PCC 9228. I identified genes related to both photosynthetic modes and nitrogen fixation that equip *Geitlerinema* for life in fluctuating H_2S, O_2, and light availability. *Geitlerinema* possesses three types of *psbA*, a gene encoding essential proteins for photosynthesis, likely cued to oxygen and/or light availability. The regulation of oxygen/light-replete, microaerobic, and anoxic versions of *psbA* is necessary for nitrogen fixation, an O_2-sensitive process. Genes controlling nitrogen fixation in *Geitlerinema* are present in nitrogen-

fixing archaea and select anaerobic bacteria, but have not been previously investigated in cyanobacteria. *Geitlerinema* encodes two types of *sqr*, which may be cued to different sulfide concentrations and trace metal (arsenic) exposure. The first thorough physiological characterization of anoxygenic cyanobacteria was conducted 40 years ago using *Geitlerinema*, but despite advances in gene sequencing and analysis, its genome was not evaluated until this dissertation project. The genome of *Geitlerinema* broadens the known suite of genes underpinning cyanobacterial photosynthetic modes and nitrogen fixation, and their potential metabolic connectedness. Such genes may have been required for cyanobacteria to cope with fluctuating redox gradients that dominated Earth's early environments, and modern extreme systems.

In **Chapters III** and **IV**, I apply our understanding of the genetic foundation for AP established in **Chapter II** and in previous studies to a natural cyanobacterial mat in Middle Island Sinkhole (MIS), a modern O_2-poor aquatic environment in Lake Huron, Michigan. Sulfate-rich groundwater bathes the sinkhole, and bacterial sulfate reduction in the mat and sediment provides sulfide to AP cyanobacteria *Phormidium* and *Planktothrix*, as well as sulfide-oxidizing bacteria (*Beggiatoa*) (Biddanda et al., 2015; Nold, Pangborn, et al., 2010; Voorhies et al., 2012; 2016). In **Chapter III**, I investigated the impact of seasonality on the function (metaproteomics) and community structure (16S rRNA gene analysis) of the sulfur-cycling microbial population in MIS, over multiple years and seasons. Light intensity and groundwater chemistry change seasonally, which can impact the growth and function of AP cyanobacteria *Phormidium* and *Planktothrix*, sulfide-oxidizing *Beggiatoa*, and sulfate-reducing deltaproteobacteria. Cyanobacteria are abundant and dominate the community functional profile in summer months. When light levels decrease in autumn, *Beggiatoa* and sulfate reducers are

more abundant and active in the microbial mat. This research is the first to describe the protein signatures of discrete populations of sulfur cycling organisms, including AP cyanobacteria, in a low-oxygen, sulfidic mat. By shaping the microbial community from the top down, these seasonal dynamics may influence the balance of oxygen production in microbial mats on longer temporal scales.

In **Chapter IV**, I leverage metagenomics and metaproteomics to explore the microbial communities in several visually distinct microbial mat morphologies in MIS, to understand the linkage between microbial function and mat appearance. In 2015-2017, field surveys documented four mat morphotypes (flat purple mat, conical purple 'finger' mats, flat white mat, and mottled purple/brown 'giraffe' mat) with different spatiotemporal distributions. Sampling each mat type multiple times, I recovered hundreds of metagenome-assembled-genomic bins (MAGs) that represented cyanobacteria, sulfide-oxidizing gamma- and epsilonproteobacteria, sulfate-reducing deltaproteobacteria, and other bacterial members of the mat. Near-complete representative MAGs belonging to *Pseudanabaena* sp. and *Spirulina* sp. were the most often observed in cyanobacteria in white and giraffe mats, whereas representative *Phormidium* and *Planktothrix* bins were more often observed in white and finger mats. The presence of different *sqr* and *psbA* versions in these MAGs suggest functional redundancy between cyanobacteria in the different morphotypes. Two distinct sulfate-reducing deltaproteobacterial MAGs with differences in their spatiotemporal range have putatively different responses to oxygen conditions. Similarly, the spatiotemporal abundances of sulfide-oxidizing MAGs belonging to Gammaproteobacteria and Epsilonproteobacteria may also be related to oxic/sulfidic conditions and the abilities to metabolize sulfide using nitrate instead of oxygen, and sulfur disproportionation. This shift in the microbial sulfur cyclers manifested in the proteomic profiles,

with proteins belonging to *Pseudanabaena, Spirulina,* sulfide oxidizers, and sulfate reducers more abundant in giraffe and white mats compared to finger mats. The connection between microbial activity and mat morphotype in early-Earth analog microbial mats informs our interpretation of the putative function of microbes and their interaction with geochemical conditions in preserved mats.

1.4 References

Arieli, B., Shahak, Y., Taglicht, D., Hauska, G., & Padan, E. (1994). Purification and characterization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in *Oscillatoria limnetica*. *Journal of Biological Chemistry*, *269*(8), 5705–5711.

Biddanda, B. A., Coleman, D. F., Johengen, T. H., Ruberg, S. A., Meadows, G. A., Van Sumeren, H. W., et al. (2006). Exploration of a Submerged Sinkhole Ecosystem in Lake Huron. *Ecosystems*, *9*(5), 828–842. https://doi.org/10.1007/s10021-005-0057-y

Biddanda, B. A., McMillan, A. C., Long, S. A., Snider, M. J., & Weinke, A. D. (2015). Seeking sunlight: rapid phototactic motility of filamentous mat-forming cyanobacteria optimize photosynthesis and enhance carbon burial in Lake Huron's submerged sinkholes. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00930

Bosak, T., Knoll, A. H., & Petroff, A. P. (2013). The meaning of stromatolites. *Annual Review of Earth and Planetary Sciences*. https://doi.org/10.1146/annurev-earth-042711-105327

Bronstein, M., Schütz, M., Hauska, G., Padan, E., & Shahak, Y. (2000). Cyanobacterial Sulfide-Quinone Reductase: Cloning and Heterologous Expression. *Journal of Bacteriology*, *182*(12), 3336–3344. https://doi.org/10.1128/JB.182.12.3336-3344.2000

Butterfield, N. J. (2009). Modes of pre-Ediacaran multicellularity. *Precambrian Research*, *173*(1-4), 201–211. https://doi.org/10.1016/j.precamres.2009.01.008

Canfield, D. E. (1998). A new model for Proterozoic ocean chemistry. *Nature*, 1–4.

Canfield, D. E., & Marais, Des, D. J. (1993). Biogeochemical cycles of carbon, sulfur, and free oxygen in a microbial mat. *Geochemica Et Cosmochimica Acta*, *57*, 3971–3984. https://doi.org/10.1016/0016-7037(93)90347-Y

Canfield, D. E., & Teske, A. (1996). Late Proterozoic rise in atmospheric oxygen concentration inferred from phylogenetic and sulphur-isotope studies. *Nature*, *382*(6587), 127–132. https://doi.org/10.1038/382127a0

Catling, D. C., Glein, C. R., Zahnle, K. J., & McKay, C. P. (2005). Why O(2) is required by complex life on habitable planets and the concept of planetary "oxygenation time." *Astrobiology*, *5*(3), 415–438. https://doi.org/10.1089/ast.2005.5.415

Chan, L.-K., Morgan-Kiss, R. M., & Hanson, T. E. (2009). Functional Analysis of Three Sulfide:Quinone Oxidoreductase Homologs in Chlorobaculum tepidum. *Journal of Bacteriology*, *191*(3), 1026–1034. https://doi.org/10.1128/JB.01154-08

Cohen, Y., Jørgensen, B. B., Padan, E., & Shilo, M. (1975). Sulphide-dependent anoxygenic photosynthesis in the cyanobacterium *Oscillatoria limnetica*. *Nature*, *257*(5526), 489–492. https://doi.org/10.1038/257489a0

Cohen, Y., Jørgensen, B. B., Revsbech, N. P., & Poplawski, R. (1986). Adaptation to Hydrogen Sulfide of Oxygenic and Anoxygenic Photosynthesis among Cyanobacteria. *Applied and Environmental Microbiology*, *51*(2), 398–407.

Dick, G. J. (2019). *Genomic approaches in Earth and Environmental Sciences* (pp. 1–165). Wiley Blackwell.

Dick, G. J., & Lam, P. (2015). Omic Approaches to Microbial Geochemistry. *Elements*, *11*(6), 403–408. https://doi.org/10.2113/gselements.11.6.403

Dick, G. J., Grim, S. L., & Klatt, J. M. (2018). Controls on O_2 Production in Cyanobacterial Mats and Implications for Earth's Oxygenation. *Annual Review of Earth and Planetary Sciences*, *46*(1), 123–147. https://doi.org/10.1146/annurev-earth-082517-010035

Fike, D. A., Bradley, A. S., & Rose, C. V. (2015). Rethinking the Ancient Sulfur Cycle. *Annual Review of Earth and Planetary Sciences*, *43*, 593–622. https://doi.org/10.1146/annurev-earth-060313-054802

Flood, B. E., Bailey, J. V., & Biddle, J. F. (2014). Horizontal gene transfer and the rock record: comparative genomics of phylogenetically distant bacteria that induce wrinkle structure formation in modern sediments. *Geobiology*, *12*(2), 119–132. https://doi.org/10.1111/gbi.12072

Garlick, S., Oren, A., & Padan, E. (1977). Occurrence of facultative anoxygenic photosynthesis among filamentous and unicellular cyanobacteria. *Journal of Bacteriology*, *129*(2), 623–629.

Gregersen, L. H., Bryant, D. A., & Frigaard, N.-U. (2011). Mechanisms and evolution of oxidative sulfur metabolism in green sulfur bacteria. *Frontiers in Microbiology*, *2*(May), 1–14. https://doi.org/10.3389/fmicb.2011.00116/abstract

Grim, S. L., & Dick, G. J. (2016). Photosynthetic Versatility in the Genome of *Geitlerinema* sp. PCC 9228 (Formerly *Oscillatoria limnetica* "Solar Lake"), a Model Anoxygenic Photosynthetic Cyanobacterium. *Frontiers in Microbiology*, *7*(590), 1144. https://doi.org/10.3389/fmicb.2016.01546

Haas, S., de Beer, D., Klatt, J. M., Fink, A., Rench, R. M., Hamilton, T. L., et al. (2018). Low-Light Anoxygenic Photosynthesis and Fe-S-Biogeochemistry in a Microbial Mat. *Frontiers in Microbiology*, *9*, 858. https://doi.org/10.3389/fmicb.2018.00858

Hamilton, T. L., Bryant, D. A., & Macalady, J. L. (2016). The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans. *Environmental Microbiology*, *18*(2), 325–340. https://doi.org/10.1111/1462-2920.13118

Holland, H. D. (2002). Volcanic gases, black smokers, and the Great Oxidation Event. *Geochimica Et Cosmochimica Acta*, *66*(21), 3811–3826. https://doi.org/10.1016/S0016-7037(02)00950-X

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., et al. (2016). A new view of the tree of life. *Nature Microbiology*, 1–6. https://doi.org/10.1038/nmicrobiol.2016.48

Johnston, D. T., Wolfe-Simon, F., Pearson, A., & Knoll, A. H. (2009). Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age. *Proceedings of the National Academy of Sciences*, *106*(40), 16925–16929. https://doi.org/10.1073/pnas.0909248106

Klatt, J. M., Al-Najjar, M. A. A., Yilmaz, P., Lavik, G., de Beer, D., & Polerecky, L. (2015). Anoxygenic photosynthesis controls oxygenic photosynthesis in a cyanobacterium from a sulfidic spring. *Applied and Environmental Microbiology*, *81*(6), 2025–2031. https://doi.org/10.1128/AEM.03579-14

Klatt, J. M., de Beer, D., Häusler, S., & Polerecky, L. (2016). Cyanobacteria in Sulfidic Spring Microbial Mats Can Perform Oxygenic and Anoxygenic Photosynthesis Simultaneously during an Entire Diurnal Period. *Frontiers in Microbiology*, *7*(116), 440. https://doi.org/10.3389/fmicb.2016.01973

Klatt, J. M., Meyer, S., Häusler, S., Macalady, J. L., de Beer, D., & Polerecky, L. (2016). Structure and function of natural sulphide-oxidizing microbial mats under dynamic input of light and chemical energy. *The ISME Journal*, *10*(4), 921–933. https://doi.org/10.1038/ismej.2015.167

Konhauser, K. O. (2009). Introduction to Geomicrobiology. John Wiley & Sons. https://doi.org/10.1371/journal.pone.0140766

Lalonde, S. V., & Konhauser, K. O. (2015). Benthic perspective on Earth's oldest evidence for oxygenic photosynthesis. *Proceedings of the National Academy of Sciences*, *112*(4), 995–1000. https://doi.org/10.1073/pnas.1415718112

Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth’s early ocean and atmosphere. *Nature*, *506*(7488), 307–315. https://doi.org/10.1038/nature13068

MacGregor, B. J., Biddle, J. F., Siebert, J. R., Staunton, E., Hegg, E. L., Matthysse, A. G., & Teske, A. (2013). Why orange Guaymas Basin Beggiatoa spp. are orange: single-

filament-genome-enabled identification of an abundant octaheme cytochrome with hydroxylamine oxidase, hydrazine oxidase, and nitrite reductase activities. *Applied and Environmental Microbiology*, *79*(4), 1183–1190. https://doi.org/10.1128/AEM.02538-12

Marcia, M., Ermler, U., Peng, G., & Michel, H. (2010). A new structure-based classification of sulfide:quinone oxidoreductases. *Proteins: Structure, Function, and Bioinformatics*, *78*(5), 1073–1083. https://doi.org/10.1002/prot.22665

Miller, S. R., & Bebout, B. M. (2004). Variation in sulfide tolerance of photosystem II in phylogenetically diverse cyanobacteria from sulfidic habitats. *Applied and Environmental Microbiology*, *70*(2), 736–744. https://doi.org/10.1128/AEM.70.2.736-744.2004

Mussmann, M., Hu, F. Z., Richter, M., de Beer, D., Preisler, A., Jørgensen, B. B., et al. (2007). Insights into the Genome of Large Sulfur Bacteria Revealed by Analysis of Single Filaments. *PLoS Biology*, *5*(9), e230. https://doi.org/10.1371/journal.pbio.0050230

Noffke, N. (2010). Geobiology: Microbial mats in sandy deposits from the Archean Era to today (pp. 1–198). Springer Science & Business Media.

Noffke, N., & Awramik, S. M. (2013). Stromatolites and MISS—differences between relatives. *GSA Today*. https://doi.org/10.1130/GSATG187A.1

Nold, S. C., Pangborn, J. B., Zajack, H. A., Kendall, S. T., Rediske, R. R., & Biddanda, B. A. (2010). Benthic bacterial diversity in submerged sinkhole ecosystems. *Applied and Environmental Microbiology*, *76*(1), 347–351. https://doi.org/10.1128/AEM.01186-09

Oren, A., Padan, E., & Malkin, S. (1979). Sulfide inhibition of photosystem II in cyanobacteria (blue-green algae) and tobacco chloroplasts. *Biochimica Et Biophysica Acta*, *546*, 270–279. https://doi.org/https://doi.org/10.1016/0005-2728(79)90045-8

Paerl, H. W., Pinckney, J. L., & Steppe, T. F. (2000). Cyanobacterial–bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. *Environmental Microbiology*, *2*(1), 11–26. https://doi.org/10.1046/j.1462-2920.2000.00071.x

Planavsky, N. J., Asael, D., Hofmann, A., Reinhard, C. T., Lalonde, S. V., Knudsen, A., et al. (2014). Evidence for oxygenic photosynthesis half a billion years before the Great Oxidation Event. *Nature Geoscience*, *7*(4), 283–286. https://doi.org/10.1038/ngeo2122

Planavsky, N. J., McGoldrick, P., Scott, C. T., Li, C., Reinhard, C. T., Kelly, A. E., et al. (2011). Widespread iron-rich conditions in the mid-Proterozoic ocean. *Nature*, *477*(7365), 448–451. https://doi.org/10.1038/nature10327

Poulton, S. W., Fralick, P. W., & Canfield, D. E. (2010). Spatial variability in oceanic redox structure 1.8 billion years ago. *Nature Geoscience*, *3*(7), 486–490. https://doi.org/10.1038/ngeo889

Reinhard, C. T., Planavsky, N. J., Robbins, L. J., Partin, C. A., Gill, B. C., Lalonde, S. V., et al.

(2013). Proterozoic ocean redox and biogeochemical stasis. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1208622110/-/DCSupplemental

Ruberg, S. A., Kendall, S. T., & Biddanda, B. A. (2008). Observations of the Middle Island Sinkhole in Lake Huron–a unique hydrogeologic and glacial creation of 400 million years. *Marine Technology Society Journal*, *42*(4), 12–21. https://doi.org/10.4031/002533208787157633

Schirrmeister, B. E., de Vos, J. M., Antonelli, A., & Bagheri, H. C. (2013). Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1209927110/-/DCSupplemental

Sharrar, A. M., Flood, B. E., Bailey, J. V., Jones, D. S., Biddanda, B. A., Ruberg, S. A., et al. (2017). Novel Large Sulfur Bacteria in the Metagenomes of Groundwater-Fed Chemosynthetic Microbial Mats in the Lake Huron Basin. *Frontiers in Microbiology*, *8*, 2104. https://doi.org/10.3389/fmicb.2017.00791

Stal, L. J. (2012). Cyanobacterial Mats and Stromatolites. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II* (pp. 65–125). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_4

Sumner, D. Y., Hawes, I., Mackey, T. J., Jungblut, A. D., & Doran, P. T. (2015). Antarctic microbial mats: A modern analog for Archean lacustrine oxygen oases. *Geology*, *43*(10), 887–890. https://doi.org/10.1130/G36966.1

Theissen, U., Hoffmeister, M., Grieshaber, M., & Martin, W. (2003). Single eubacterial origin of eukaryotic sulfide:quinone oxidoreductase, a mitochondrial enzyme conserved from the early evolution of eukaryotes during anoxic and sulfidic times. *Molecular Biology and Evolution*, *20*(9), 1564–1574. https://doi.org/10.1093/molbev/msg174

Voorhies, A. A., Biddanda, B. A., Kendall, S. T., Jain, S., Marcus, D. N., Nold, S. C., et al. (2012). Cyanobacterial life at low O_2: community genomics and function reveal metabolic versatility and extremely low diversity in a Great Lakes sinkhole mat. *Geobiology*, *10*(3), 250–267. https://doi.org/10.1111/j.1472-4669.2012.00322.x

Voorhies, A. A., Eisenlord, S. D., Marcus, D. N., Duhaime, M. B., Biddanda, B. A., Cavalcoli, J. D., & Dick, G. J. (2016). Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics. *Environmental Microbiology*, *18*(2), 358–371. https://doi.org/10.1111/1462-2920.12756

Ward, D. M., Bateson, M. M., Ferris, M. J., Kuhl, M., Wieland, A., Koeppel, A., & Cohan, F. M. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1475), 1997–2008. https://doi.org/10.1098/rstb.2006.1919

Xiong, J., Fischer, W. M., Inoue, K., Nakahara, M., & Bauer, C. E. (2000). Molecular evidence for the early evolution of photosynthesis. *Science*, *289*(5485), 1724–1730.

Chapter II: Photosynthetic Versatility in the Genome of *Geitlerinema* sp. PCC 9228 (Formerly *Oscillatoria limnetica* 'Solar Lake'), a Model Anoxygenic Photosynthetic Cyanobacterium

Sharon L. Grim and Gregory J. Dick

Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA

Originally published in 2016 in *Frontiers in Microbiology* 7: 1-14.

doi: 10.3389/fmicb.2016.01546

2.1 Abstract

Anoxygenic cyanobacteria that use sulfide as the electron donor for photosynthesis are a potentially influential but poorly constrained force on Earth's biogeochemistry. Their versatile metabolism may have boosted primary production and nitrogen cycling in euxinic coastal margins in the Proterozoic. In addition, they represent a biological mechanism for limiting the accumulation of atmospheric oxygen, especially before the Great Oxidation Event and in the low-oxygen conditions of the Proterozoic. In this study, we describe the draft genome sequence of *Geitlerinema* sp. PCC 9228, formerly *Oscillatoria limnetica* 'Solar Lake', a mat-forming diazotrophic cyanobacterium that can switch between oxygenic photosynthesis and sulfide-based anoxygenic photosynthesis (AP). *Geitlerinema* possesses three variants of *psbA,* which encodes protein D1, a core component of the photosystem II reaction center. Phylogenetic analyses indicate that one variant is closely affiliated with cyanobacterial *psbA* genes that code for a D1

protein used for oxygen-sensitive processes. Another version is phylogenetically similar to cyanobacterial *psbA* genes that encode D1 proteins used under microaerobic conditions, and the third variant may be cued to high light and/or elevated oxygen concentrations. *Geitlerinema* has the canonical gene for sulfide quinone reductase (SQR) used in cyanobacterial AP and a putative transcriptional regulatory gene in the same operon. Another operon with a second, distinct *sqr* and regulatory gene is present, and is phylogenetically related to *sqr* genes used for high sulfide concentrations. The genome has a comprehensive *nif* gene suite for nitrogen fixation, supporting previous observations of nitrogenase activity. *Geitlerinema* possesses a bidirectional hydrogenase rather than the uptake hydrogenase typically used by cyanobacteria in diazotrophy. Overall, the genome sequence of *Geitlerinema* sp. PCC 9228 highlights potential cyanobacterial strategies to cope with fluctuating redox gradients and nitrogen availability that occur in benthic mats over a diel cycle. Such dynamic geochemical conditions likely also challenged Proterozoic cyanobacteria, modulating oxygen production. The genetic repertoire that underpins flexible oxygenic/anoxygenic photosynthesis in cyanobacteria provides a foundation to explore the regulation, evolutionary context, and biogeochemical implications of these co-occurring metabolisms in Earth history.

2.2 Introduction

With a long evolutionary history and wide ecological success, cyanobacteria are pivotal mediators of Earth's geochemical cycles, most notably through oxygenic photosynthesis (OP). This metabolism emerged early in cyanobacteria (Blankenship, 2010; Farquhar et al., 2010), and oxygenic cyanobacteria that colonized newly formed continental margins and shallow seas in the Archean (Reddy and Evans, 2009; Lalonde and Konhauser, 2015) were the leading mechanism

for partial oxygenation of these shallow regions (Planavsky et al., 2014; Satkoski et al., 2015). Although, cyanobacterial OP is the major biological force behind Earth's oxygenation (Blankenship, 2010; Crowe et al., 2013; Planavsky et al., 2014), the biological and geological processes that influence cyanobacterial oxygen production and thus underpin Earth's oxygenation are still under debate (Lyons et al., 2014).

The balance of oxygenic and anoxygenic photosynthesis (AP) has been proposed as a biological mechanism to explain the delay and variability in oxygenation (Johnston et al., 2009). This includes competition between AP bacteria and OP cyanobacteria, interactions between different cyanobacterial groups with varying degrees of AP or OP specialization, as well as cellular regulation of the photosynthetic modes within metabolically flexible cyanobacterial groups. Atmospheric oxygenation is considered key evidence for OP, but the emergence and development of AP in cyanobacteria is less clear. Assuming an early evolution of photosynthetic flexibility in cyanobacteria, AP cyanobacteria may have had an important role in sustaining ancient ecosystems from the end of the Archean through the Proterozoic, especially in times of global and local variability of O_2, fixed nitrogen, and alternative electron donors for photosynthesis such as H_2S and Fe(II) (Canfield, 1998; Scott et al., 2008; Lyons et al., 2014; Sperling et al., 2015).

Molecular innovations equipped cyanobacteria for OP and nitrogen fixation in a dynamic environment. The development of OP required the linkage of two light-driven reaction centers: photosystem II, which produces oxygen from the oxidation of water; and photosystem I, which transfers electrons from plastoquinone to ferredoxin (Blankenship, 2002). This coupled system capitalized on the wide availability of H_2O compared to more limited supply of electron donors for AP such as Mn(II), Fe(II), and H_2S (Blankenship, 2002). Homologous proteins D1 and D2,

encoded by the *psbA* and *psbD* genes, form the core of PSII and anchor the water oxidizing complex (Ferreira et al., 2004; Fischer et al., 2015). Modern cyanobacteria have multiple versions of *psbA* to cope with different oxygen levels and light regimes (Mohamed et al., 1993). Cyanobacterial adaptations to an aerobic lifestyle are also reflected in the *nif* genes for nitrogen fixation, which initially emerged in methanogens in an anoxic environment (Raymond et al., 2004; Boyd et al., 2011). Reflecting increasing oxygen levels, cyanobacterial genomes lost and recruited nitrogenase (*nif*)-related genes, and shifted expression patterns and regulation that enabled nitrogen fixation in an oxic world (Boyd et al., 2015).

Studies of modern cyanobacteria have provided insights into how sulfide may have modulated the balance of OP and AP in ancient ecosystems. The influence of sulfide on cyanobacterial photosynthesis ranges from complete inhibition at even low levels of H2S to resilience or resistance to sulfide toxicity. In some cyanobacteria, sulfide exposure may induce AP (Cohen et al., 1986; Miller and Bebout, 2004), which does not include PSII and thus does not produce O_2. Instead, the sulfide quinone reductase (SQR, coded for by the *sqr* gene) oxidizes sulfide to sulfur and transfers electrons to PSI (Arieli et al., 1994; Theissen et al., 2003). Such AP cyanobacteria have been documented in hypersaline lakes (Cohen et al., 1975a), sinkholes (Voorhies et al., 2012), and sulfidic springs (Chaudhary et al., 2009; Bühring et al., 2011; Klatt et al., 2016). Studied cyanobacteria have different mechanisms for the transition between OP and AP, such as protein synthesis (Oren and Padan, 1978), a dependence on light quantity and spectrum, and kinetics and affinities between enzymes and quinones (Klatt et al., 2015a). The physiology of AP cyanobacteria and their potential importance in modern and ancient ecosystems have been previously explored, yet the genomic basis for this flexible metabolism is poorly understood.

Geitlerinema sp. PCC 9228, formerly *Oscillatoria limnetica* 'Solar Lake', is a model anoxygenic photosynthetic cyanobacterium. The organism was cultured from the low-light sulfidic hypolimnion of Solar Lake, below a layer of green and purple sulfur bacteria (Cohen et al., 1975b). Filamentous cyanobacteria such as *Geitlerinema* are rare in oxic and well-illuminated surface waters, and most numerous in the euxinic hypolimnion at which they receive a fraction of surface irradiance (Cohen et al., 1977b). In laboratory experiments at light intensities similar to or higher than *in situ* levels, *Geitlerinema* performs OP, but transitions fully to sulfide-based AP at micromolar concentrations of sulfide (Cohen et al., 1986). Under sulfidic conditions, *Geitlerinema* can also fix nitrogen and produce hydrogen (Belkin and Padan, 1978; Belkin et al., 1982). Its SQR has been isolated (Arieli et al., 1994), sequenced (Bronstein et al., 2000), and phylogenetically characterized (Pham et al., 2008; Marcia et al., 2010a; Gregersen et al., 2011). *Geitlerinema* is a model organism for studying the physiology of flexible AP/OP, diazotrophic cyanobacteria and their influence on modern and ancient systems. In this study, we analyzed a draft genome of *Geitlerinema* and characterized the genes related to nitrogen fixation, AP, and OP. These results provide a genomic foundation for metabolic flexibility in response to varying sulfide, oxygen, and light levels that was observed in previous physiology studies (Belkin and Padan, 1978; Belkin et al., 1982; Cohen et al., 1986; Shahak et al., 1987).

2.3 Materials and Methods

Culturing and Sequencing. The original strain was isolated from the sulfidic water column of Solar Lake, Israel (Cohen et al., 1975b), and was kindly provided by A. Oren for culturing. A monoalgal culture was grown in modified Chu's 11 in Turks Island Salts medium at room temperature (average 22.0◦C) and ambient light in a 125 mL Erhlenmeyer flask. We extracted

whole community DNA using the MPBio FastDNA SpinKit and Fastprep-24 Bead Beater (MP Biomedicals, Solon, OH, USA) following the default protocol, except that 0.3 g of beads were used for bead beating. DNA was quantified using Quant-IT PicoGreen (Invitrogen, Grand Island, NY, USA) and submitted to the University of Michigan DNA Sequencing Core for library preparation and Illumina HiSeq 2 × 100 bp paired-end sequencing.

Assembly. Using wrappers provided at https://github.com/Geo-omics/ scripts, reads were dereplicated with custom perl scripts, trimmed using Sickle (version 1.33) (Joshi and Fass, 2011), and assembled using IDBA-UD (version 1.1.1) (Peng et al., 2012) with the following parameters: –mink 65 –maxk 85 – step 10 –pre_correction. Tetranucleotide frequency was used to bin scaffolds by emergent self-organizing maps (Dick et al., 2009), with a minimum contig length of 500 bp and a window size of 10,000 bp (**Figure SI 2.1**). The cyanobacterial bin of scaffolds >1000 bp in length was submitted to the Integrated Microbial Genomes Expert Review (IMG- ER) automated pipeline from Joint Genomes Institute (JGI) for annotation of genes and pathways (IMG accession number: 2660238729; Supplementary Table S1). Raw reads, assembled scaffolds, and gene annotations were submitted to NCBI (project: PRJNA302164).

Phylogenetic Analysis. Phylogenetic analyses of genes of interest (**Table 2.1**) was performed with maximum likelihood and the PROTGAMMAGTR algorithm in RAxML 8.1.15 and bootstrapped 1000 times (Stamatakis, 2014). *psbA*, *sqr*, and *nifHDK* gene phylogenies included 44 cyanobacterial isolate genomes and genomic bins from Hot Lake and Middle Island Sinkhole cyanobacterial mat metagenomes (Cole et al., 2014; Voorhies et al., 2016; Supplementary Table S2). Each of these cyanobacterial genome has at least one *sqr* gene, and 30 of the 44 genomes

have a *nifHDK* gene set. *sqr* and *nifHDK* genes from *Sulfuricurvum kujiense* YK-1 DSM 16994, and other bacterial *sqr* genes (**Table SI 2.3**) were included for context. Translated genes (amino acid sequences) were aligned with clustal-omega (version 1.2.0) (Sievers et al., 2011). Alignments for *nifH*, *nifD*, and *nifK* were concatenated, and the concatenated alignment was used for analysis. *Halothece halophytica chlLNB* genes were used to root the *nifHDK* tree (Boyd et al., 2015). Flavocytochrome c:sulfide dehydrogenase (FCSD) genes formed an outgroup for the *sqr* tree (Marcia et al., 2010a). The most divergent *psbA* from *Gloeobacter kilaueensis* JS-1 was used as an outgroup for the *psbA* tree (Cardona et al., 2015). Phylogenetic trees were drawn with FigTree version 1.4.2 (Rambaut, 2012).

2.4 Results

Overview

The 4.77 Mb draft genome of *Geitlerinema* sp. PCC 9228 (henceforth "*Geitlerinema*") contains 3,969 protein coding genes on 195 scaffolds. Coverage is on average 905x across the genome. The genome has 100% of universally conserved bacterial genes expected to be present (Raes et al., 2007; Alneberg et al., 2014).

Carbon Metabolism

Geitlerinema has key genes coding for proteins involved in photosynthetic and respiratory electron flow, including photosystem II (*psbADBCEFO*), succinate dehydrogenase (*sdhABC*), type-1 NADPH dehydrogenase (*ndhA-M*), cytochrome b6f (*petADBCEJ*), photosystem I (*psaABCDEFK*), and cytochrome c oxidation (*coxABC*) (Supplementary Table S1; Mulkidjanian et al., 2006). It has genes for the Calvin- Benson-Bassham cycle, carbon

dioxide concentrating proteins (*ccmK*1 *K*2 *MN*), and Rubisco large and small subunits (*rbcLS*) for carbon fixation (**Figure 2.1**). Superoxide dismutase genes (*sodC* and *sodN*) to cope with superoxide formation during photosynthesis and aerobic respiration are also present in the genome. In ancestral cyanobacteria, these enzymes would have been critical for defense against increased production of reactive oxygen species alongside increasing O2 fluxes (Blank and Sánchez-Baracaldo, 2010; Fischer et al., 2016).

Geitlerinema has genes encoding a complete tricarboxylic acid (TCA) cycle, including 2-oxoglutarate dehydrogenase and succinyl CoA synthase to link synthesis of 2-oxoglutarate through succinyl-coA with succinate (Steinhauser et al., 2012). The genome also has the genes for acetolactate synthases and succinate-semialdehyde dehydrogenase to interconvert 2-OG and succinate through succinic semialdehyde in an alternative closure to the TCA cycle (Zhang and Bryant, 2011). Though it has *shc*, encoding for squalene-hopene cyclase, *Geitlerinema* lacks the *hpnP* gene for hopanoid methylation (Ricci et al., 2015). 2-methylhopanes, derived from 2-methylhopanoids, have been used as a bacterial biomarker in the geologic record (Summons and Lincoln, 2012). *Geitlerinema* has genes for acyl- ACP reductase and fatty aldehyde decarbonylase, key enzymes in an alkane biosynthesis pathway unique to cyanobacteria (Schirmer et al., 2010; Coates et al., 2014).

In the chlorophyll synthesis pathway, we identified both the aerobic oxidative ester cyclase *chlE* common to all cyanobacteria (Mulkidjanian et al., 2006), and the oxygen independent ester cyclase *bchE*. Functional *bchE* is common in anoxygenic phototrophic bacteria such as green sulfur bacteria and heliobacteria (Sousa et al., 2013), and is only rarely present in cyanobacteria, such as *Synechocystis* sp. PCC 6803 (Minamizaki et al., 2008). *Geitlerienema* has the light-dependent NADPH- protochlorophyllide oxidoreductase that produces chlorophyllide in

the penultimate step of the pathway. This gene originated in cyanobacteria and is limited to cyanobacteria and phototrophic eukaryotes (Blankenship, 2002; Yang and Cheng, 2004). However, the genome lacks the light-independent oxidoreductase *chlLNB*, which allows for chlorophyll synthesis in the dark and is observed in all photosynthetic phyla (Blankenship, 2002). Due to their homology, *Geitlerinema*'s genes for *nifHDK* are the closest matches to *H. halophytica's chlLNB*. Coverage estimates for *nifHDK* genes are consistent with the rest of the genome, indicating there was no mis-assembly of *chlLNB* reads into *nifHDK* genes.

Nitrogen Metabolism

The cyanobacterial genome has genes for a variety of pathways of nitrogen acquisition, including an ammonia transporter gene *amt*, cyanate lyase *cynS*, nitrate/nitrite transporters *narK* and *focA*, and nitrate assimilation-related genes nitrate and nitrite reductases *nirA*, *nirC*, and *narB* (**Figure 2.1**). The cyanobacterium also has urea transporters (*urtABCD*), urease genes (*ureABCDFG*), and genes for transport of neutral, branched, and polar amino acids.

Geitlerinema has a comprehensive operon for nitrogen fixation (*nifVXSU, nifHDKEB*; Raymond et al., 2004). Additional nitrogenase-related proteins are located on the same contig (*iscA*, $nifI_1I_2$, ferredoxin, *nifN*; **Figure 2.2**; Boyd et al., 2015). *iscA* is commonly observed in aerobic diazotrophs, and its recruitment into the genome is linked to a transition to aerobic lifestyle (Boyd et al., 2015). The cyanobacterium has *glnB*, a member of the PII signal transduction protein family that regulates nitrogen-related proteins, and *ntcA*, which controls expression of *glnB* (Forchhammer, 2004). The $nifI_1I_2$ gene is also a member of the PII protein family, but is characteristic of diazotrophic anaerobes that regulate their nitrogenase activity post-translation, such as *Desulfovibirio* and *Clostridium* (Boyd et al., 2015).

Phylogenetic analysis shows that *nifHDK* genes from *Geitlerinema* clusters with those from two cultured cyanobacterial genomes (*Pleurocapsa* sp. 7327, *Microcoleus chthonoplastes*) and a cyanobacterial genome-from-meta genomic bin (*Phormidium* OSCR; bootstrap = 100; **Figure SI 2.2**). Like *Geitlerinema*, their nitrogenase operons also hold *nifI*1 *I*2 . *Geitlerinema* has a bidirectional NiFe hydrogenase gene set (*hoxEFUYH* and *hoxW*) with its transcriptional regulator (*lexA*) and hydrogenase maturation proteins (*hypBAEDC*) (**Figure 2.1**). Unlike *Geitlerinema*, typical nitrogen-fixing cyanobacteria have an uptake hydrogenase (*hupSL*) to consume H_2 produced in nitrogen fixation. *hupSL* is under similar transcriptional regulation as nitrogen acquisition genes like dinitrogenase, making expression of *hupSL* dependent on nitrogen limitation (Tamagnini et al., 2007). In contrast, the bidirectional hydrogenase can be present in both diazotrophic and non-diazotrophic cyanobacteria and is expressed under more diverse conditions. It may, for instance, be used in fermentation or to direct electrons during photosynthesis (Tamagnini et al., 2007). *Geitlerinema* produces hydrogen during sulfide-dependent AP in the presence of bioavailable nitrogen, as well as in the absence of CO_2, suggesting the nitrogenase-independent *hox* gene set may be responsible for hydrogen evolution (Belkin and Padan, 1978; Belkin et al., 1982).

Photosystem II Assembly

The core proteins in photosystem II are encoded by *psbA* and *psbD* (Blankenship, 2002). *Geitlerinema* has one version of *psbD* and three versions of the *psbA* gene (**Figure 2.1**). The standard *psbA*, hereafter referred to as *psbA3*, is in a large clade of typical oxygenic *psbA* designated “group 4” after (Cardona et al., 2015) (**Figure 2.3**). All but one of the isolate genomes in this analysis have at least one copy of this form of *psbA*, which is used in OP in

aerobic conditions (**Table SI 2.2**). Coverage estimates and paired-end information suggest that *Geitlerinema* has two copies of *psbA3*. It is the sole gene on its contig, and pairs of reads that map to the ends of the *psbA3* match to the ends or beginnings of four other contigs. Those portions have 100% identity to the ~80 bp beginning and end of *psbA3*. The De Bruijn graph-based assembly algorithm used in this analysis frequently prematurely assembles identical copies of genes (Nagarajan and Pop, 2013). The approach accurately assembled the two copies of *psbA3*, but could not automatically bridge the copies.

One additional version of *psbA*, henceforth referred to as *psbA1*, has 85% sequence identity to *psbA3* (blastx). *psbA1* is located in a different operon but on the same scaffold near *ntcA* and photosynthetic subunits *psbN* and *psbH*. *psbA1* is in a well-supported clade with genes from *Geitlerinema* sp. PCC 7105 (2510100750) and *Geitlerinema* sp. BBD 1991 (BBD_1000995126; bootstrap = 81; **Figure 2.3**). This gene is also in a larger "group 2" (Cardona et al., 2015) that includes *psbA* genes from cyanobacterial genomic bins (MIS_1001011011, MIS_100039089) sourced from a low-oxygen cyanobacterial mat in the Middle Island Sinkhole (Voorhies et al., 2016). Group 2 genes encode D1 proteins used under growth conditions that do not favor water oxidation (Cardona et al., 2015), such as heterotrophy in the dark (Park et al., 2013), photosynthetic electron transport inhibition (Kiss et al., 2012), or oxygen- sensitive processes such as nitrogen fixation (Toepel et al., 2008; Wegener et al., 2015). Finally, the third variant, *psbA2*, is 59–88% similar to *psbA1* (discontinuously; tblastx), and 90% similar to *psbA3* (over 99% of their lengths; blastx). *psbA2* is in a clade with genes from a Middle Island Sinkhole cyanobacterium (MIS_100299244), *Geitlerinema* sp. BBD1991 (BBD_1000995127) and *Phormidium* sp. OSCR (2609132164; bootstrap = 94). Proteins from these "group 3" *psbA* genes (Cardona et al., 2015) are expressed under microaerobic conditions,

with modified electron transfer to cope with the changing redox environment (**Figure 2.1**; Sicora et al., 2004, 2009; Sugiura et al., 2012).

Sulfide Oxidation

Two copies of the gene for sulfide quinone reductase, *sqr*, are in the genome of *Geitlerinema* (**Figure SI 2.4**). This gene is involved in the oxidation of sulfide for detoxification or to harvest electrons for AP, such as in purple and green sulfur bacteria (Theissen et al., 2003; Marcia et al., 2009, 2010a; Gregersen et al., 2011). Each *sqr* is located upstream of arsenic resistance *arsR*-type genes, putatively involved in transcriptional regulation of *sqr* under conditions such as sulfide exposure (Nagy et al., 2014). One *sqr*, referred to as *sqr1* hereafter, matches the previously cloned and sequenced *sqr* gene of *Geitlerinema* (Bronstein et al., 2000). This enzyme mediates the reduction of plastoquinone and oxidation of sulfide in sulfide-dependent AP (Arieli et al., 1994). On the *sqr* phylogenetic tree, Geitlerinema's canonical *sqr* groups with other cyanobacterial versions that are considered type I (Marcia et al., 2010a; Gregersen et al., 2011; bootstrap = 100; **Figure 2.4**; **Figure SI 2.5**). Hydrogen sulfide affinities for cyanobacterial *sqr* in this cluster are high, with Km in the micromolar range (Arieli et al., 1991; Bronstein et al., 2000). *Geitlerinema*, *Coleofasciculus* (formerly *Microcoleus*), *Halothece*, and proteobacterial members in this cluster have been shown to grow with this *sqr* on sulfide-induced AP (Oren and Padan, 1978; Cohen et al., 1986; Jørgensen et al., 1986; Schütz et al., 1999).

The second *sqr* gene, referred to as *sqr2* hereafter, is 25–50% identical over discontinuous fragments to the *sqr1* gene (tblastx). It is most similar (tblastx 48–63% identity to overlapping fragments that span the length of the gene) to that of *Chloroherpeton thalassium*

ATCC35110, a green sulfur bacterium, with similar identity to the *sqr* genes of other sulfur-oxidizing bacteria. *sqr2* clusters phylogenetically with two other cyanobacterial *sqr* [*Geitlerinema* sp. PCC 7105 and BBD 1991 (bootstrap = 97)], among a group of known type VI versions (bootstrap = 100; **Figure 2.4**; **Figure SI 2.5**). They include *sqr* from thiotrophic proteobacteria (Marcia et al., 2010a) and green sulfur bacteria (Gregersen et al., 2011), including that of *Chlorobium tepidum* used in sulfide oxidation when sulfide exceeds 4 mM (Chan et al., 2009).

Trace Metal Resistance

Arsenic resistance genes *arsB* (an arsenite efflux pump) and *arsC* (arsenate reductase; Slyemi and Bonnefoy, 2011; van Lis et al., 2013), and *arsR*, the putative regulatory protein of *sqr*, are in an *arsRBC* operon downstream of *sqr2*. Upstream of *sqr2* are genes encoding glutathione S-transferase and a multidrug efflux pump (**Figure SI 2.4**). Another *arsB* is located downstream of *sqr1* but in a different operon. In combination with *arsB*, two arsenite-tranporting ATPases *arsA* genes are present in the genome, one located on the same scaffold as the nitrogenase gene suite and the other on the longest scaffold in the dataset. *Geitlerinema* does not have known arsenite oxidizing genes (*aioAB*, *arxA*) or respiratory arsenate reductase genes (*arrA*), based on gene annotation and BLAST searches against known genes (Hoeft et al., 2010; Slyemi and Bonnefoy, 2011). It has an annotated chromate transporter (*chrA*) on another scaffold that has 52% positive match (blastp) to a similarly annotated gene in *Synechocystis* sp. PCC6803 that functions as an arsenite uptake transporter for arsenite oxidation (Nagy et al., 2014; **Figure 2.1**). On another scaffold, a gene belonging to the DUF302 protein superfamily of unknown function is quite similar (75% positive, 56% identical blastp) to a *Synechocystis* gene in the *sqr1*

plasmid operon, as well as to a potential arsenic oxidase gene in *Agrobacterium* (61% positive, 44% identical, blastp; Nagy et al., 2014). *Geitlerinema* also has a methyltransferase similar (84% positive, blastx) to the *arsM* gene of *Synechocystis* used in arsenite methylation for detoxification (Yin et al., 2011).

2.5 Discussion

Previous investigations of cyanobacterial cultures capable of OP and sulfide-dependent AP targeted physiology, biochemistry, and limited genetic analyses (Cohen et al., 1986; Arieli et al., 1994; Bronstein et al., 2000; Miller and Bebout, 2004; Klatt et al., 2015a). However, the broader genetic characteristics of this cyanobacterial metabolic flexibility have not been thoroughly evaluated. In this study, we analyzed the genome of a model anoxygenic photosynthetic cyanobacterium, *Geitlerinema* sp. PCC 9228. The organism was isolated from a sulfidic, low- light environment, and numerous physiology studies have documented its ability to fix nitrogen, its high affinity for sulfide, and its oxygenic and anoxygenic photosynthetic capabilities (Cohen et al., 1975b, 1986; Belkin et al., 1982). We comprehensively evaluated the draft genome of *Geitlerinema* for key metabolic genes for oxygen, nitrogen, and carbon cycling, and provide the genetic evidence of its metabolic versatility.

Photosynthesis Is Optimized for Light and Redox.

Our results indicate that *Geitlerinema* sp. PCC 9228 and other sulfide-tolerant and/or sulfide-using cyanobacteria have multiple different types of *psbA* genes that are inferred to meet physiological and biochemical needs under changing light and redox conditions. The *psbA* gene encodes the D1 protein that directly supports the water-oxidizing cluster in PSII (Ferreira et al.,

2004; Fischer et al., 2015). Due to its role in water oxidation, this protein experiences high levels of oxidative damage and degradation. By impacting water oxidation, changes in redox such as sulfide inhibition of the water oxidizing complex in PSII (Miller and Bebout, 2004) and/or non-optimal light conditions lead to excessive energy (Klatt et al., 2015b) and influence the risk and magnitude of oxidative damage to D1 (Aro et al., 1993). Thus, cyanobacteria that experience such dynamic conditions may use another of the multiple *psbA* genes in their genomic repertoire under different light and oxygen regimes to mitigate oxidative damage (Schaefer and Golden, 1989; Campbell et al., 1999; Sicora et al., 2009; Gan et al., 2014; Cardona et al., 2015; Ho et al., 2016). *Geitlerinema* has two copies of *psbA* for OP under high oxygen and/or light levels (both *psbA3*; group 4), one gene for non- oxygen evolving PSII (*psbA1*; group 2), and one gene for microaerobic conditions (*psbA2*; group 3; **Figures 2.1** and **2.3**). We infer that these distinct photosynthetic genes reflect the adaptation of *Geitlerinema* to a variably lit, low-oxygen, and sulfidic lifestyle.

During high light conditions, when *Geitlerinema* conducts OP (Cohen et al., 1975a, 1986), the *psbA3* (group 4) genes are likely used to reduce photoinhibition and oxidative stress, as observed in *Synechococcus* sp. PCC 7942, *Synechocystis* sp. PCC6803, and *Thermosynechococcus elongatus* (Schaefer and Golden, 1989; Sicora et al., 2006; Kós et al., 2008; Sugiura et al., 2010). Other cyanobacteria synthesize identical D1 proteins from different genes at high light versus regular light conditions (Sugiura et al., 2010). Two copies of *psbA3* in its genome may equip *Geitlerinema* to continue OP in times of sufficient and/or high light, such as when mixed up into the epilimnion or in holomixis. Given that *psbA3* clusters with standard and high-light variant genes (**Figure 2.3**), it remains to be seen if one or both of *Geitlerinema*'s

two copies of *psbA3* is cued to the different light intensities the organism may experience in its natural habitat.

The *psbA1* and *psbA2* genes likely provide *Geitlerinema* with the flexibility to conduct photosynthesis and nitrogen fixation under varying oxygen and/or sulfide concentrations. With sufficient light, the cyanobacterium may rely upon *psbA2* (group 3), which is keyed for microaerobic conditions. Changing redox environment promotes synthesis of *psbA*2-like variants (group 3) in *Thermosynechococcus elongatus* (Sugiura et al., 2012) and *Synechocystis* sp. PCC6803 (Sicora et al., 2004). However, light levels are lower than required for OP in the hypolimnion of Solar Lake, the prime habitat of *Geitlerinema*. Together with continuous sulfide exposure, these conditions often favor AP in cyanobacteria (Klatt et al., 2015a, 2016) and in green sulfur bacteria (Hanson et al., 2013; Findlay et al., 2015). Groups 2 and 3 *psbA* genes in the current study are in close chromosomal proximity to *sqr* genes (**Figure 2.3**), suggesting an intriguing but currently untested linkage between modified PSII and sulfide exposure. Alternative D1 proteins have been hypothesized to be directly involved in the transfer of electrons from donors other than water (Murray, 2012), such as in sulfide oxidation during nitrogen fixation (Becraft et al., 2015; Olsen et al., 2015). However, whether alternative D1 proteins are directly participating in this unusual photochemistry or serve to inactivate PSII in AP, remains to be tested.

Very little sulfide is required to limit OP in *Geitlerinema* (Cohen et al., 1986), and in the transition between oxygenic to AP, continuous exposure induces synthesis of proteins such as SQR (Arieli et al., 1991). SQR has a higher affinity for plastoquinone than PSII (Klatt et al., 2015a), thus in this induction period and with continual sulfide exposure, *Geitlerinema* may also reformulate its photosystem II machinery to reduce oxygen production. Finally, the *psbA1* (group

2) gene is also likely used to disable oxygen production during times of nitrogen fixation. Non-oxygen producing group 2 enzymes have been implicated as structural 'placeholders' that may allow oxygen-sensitive processes such as nitrogen fixation to occur (Sicora et al., 2009; Murray, 2012; Wegener et al., 2015). Further experiments targeting gene and protein expression will confirm the physiological roles of the variant *psbA* genes.

Variant *psbA* genes such as those observed in *Geitlerinema* likely permitted ancestral and potentially AP-capable cyanobacteria to meet metabolic requirements in dynamic physicochemical conditions. The importance of AP cyanobacteria and their metabolisms on Archean and Proterozoic oxygen levels is linked to the timing of OP. Mat-forming cyanobacteria in modern hypersaline and hot spring ecosystems likely face the same environmental challenges as their stromatolite-forming ancestors (Stal, 1995; Grotzinger and Knoll, 1999). As in modern systems, the metabolisms and mat-building lifestyles of ancestral cyanobacterial populations would have promoted light and redox dynamics on spatial (depth) and temporal (diel) scales (Canfield, 2005; Blank and Sánchez-Baracaldo, 2010; Lalonde and Konhauser, 2015; Sumner et al., 2015). Genomic and physiological studies on contemporary mat-forming cyanobacteria and their responses to changing environmental parameters, such as different versions of *psbA* keyed to light and/or oxygen levels, inform potential genetic and physiological strategies in ancient cyanobacteria. Given the long geologic history of variable and low atmospheric oxygen concentrations, the first appearance of alternative *psbA* in cyanobacteria is uncertain. The functional differences and potential heterotachy in these homologous genes dictate caution when evaluating their timing and evolutionary order. However, the basal arrangements of group 2 to group 3/4, and group 3 to group 4 in the phylogenetic tree hint at an ancestral development of water oxidation when atmospheric oxygen was low (Cardona et al., 2015). The perpetuation of

diverse and evolutionarily old *psbA* genes in cyanobacterial genomes allow for efficient metabolic functioning regardless of oxygen levels/needs (Murray, 2012). These variant genes were likely retained in cyanobacterial genomes for oxygen-sensitive processes, such as nitrogen fixation, in an increasingly oxidizing environment.

Nitrogen Acquisition Strategies in Variable Redox Conditions.

Geitlerinema has a suite of genes for uptake of nitrogen (nitrate, nitrite, urea) as well as nitrogen fixation through a comprehensive nitrogenase gene suite (**Figures 2.1** and **2.2**). Sulfide-dependent AP was measured in select strains grown in nitrogen-replete media, including *Geitlerinema* sp. PCC 9228, *Coleofasciculus chthonoplastes,* and *Pseudanabaena* FS39, suggesting nitrate assimilation occurs under AP conditions (Cohen et al., 1986; Klatt et al., 2015a). *Geitlerinema* is also capable of AP-dependent nitrogen fixation (Belkin et al., 1982), potentially using sulfide to scavenge residual oxygen or donate electrons to nitrogenase (Stal, 2012).

Clues about the cyanobacterial transition from an anaerobic to aerobic lifestyle are apparent in regulatory genes for nitrogen acquisition. Due to its evolution under anoxia and strict requirement for anaerobic conditions, cyanobacterial diazotrophy in an increasingly oxidizing environment required new adaptations (Boyd et al., 2011; Stüeken et al., 2015). In the *nif* operon of *Geitlerinema* is *nif*I_1I_2, which encodes a signal transduction protein of the PII family that is present in diazotrophic archaea and select anaerobic bacteria, but has not been studied in cyanobacteria (Forchhammer, 2004; Boyd et al., 2015). In those organisms, NifI_1I_2 inhibits nitrogenase activity when the cell is no longer nitrogen limited (Kessler et al., 2001). *Geitlerinema*'s *nifHDK* gene suite is phylogenetically grouped with those from three other

cyanobacterial genomes that also have *nifI*$_1$*I*$_2$ in their nitrogenase operons (**Figure SI 2.2**), and members of this cluster are adapted to low-oxygen and/or sulfidic conditions. *C. chthonoplastes* can continue to operate OP at low sulfide concentrations and is also capable of sulfide-dependent AP (Cohen et al., 1986). Similar to *Geitlerinema* and its *psbA1*, *Pleurocapsa* sp. PCC 7327 also has a *psbA* gene that codes for a rogue D1 protein typically used in anoxic conditions (Wegener et al., 2015).

Geitlerinema Possesses Genetic Versatility for Sulfide Oxidation.

Sulfide quinone reductase (SQR, encoded by *sqr*) oxidizes sulfide to elemental sulfur, and this process can be coupled with lithotrophic or phototrophic growth in bacteria and archaea or detoxification of sulfide in eukaryotes (Theissen et al., 2003; Marcia et al., 2010a). The genome of *Geitlerinema* sp. PCC 9228 holds two *sqr* operons. The first operon has the well-studied high-affinity *sqr1* (type I), which is translated after a few hours of exposure to micromolar levels of H_2S (Oren and Padan, 1978) and enables *Geitlerinema* to grow by sulfide-based AP (Cohen et al., 1986). The transcriptional regulatory gene *arsR*, located downstream of *sqr1*, likely controls its expression (Nagy et al., 2014). Though lateral gene transfer has distributed various *sqr* types among bacteria, archaea, and eukarya (Theissen et al., 2003), type I *sqr* genes in cyanobacteria such as *Geitlerinema sqr1* form a distinct and well-supported cyanobacterial phylogenetic subclade within a bacterial clade (**Figure 2.4**). This cyanobacterial clade includes genes from AP- capable or sulfide-tolerant members such as *C. chthonoplastes*, *Synechocystis* sp. PCC6803, *H. halophytica*, and *Geitlerinema* sp. BBD (Cohen et al., 1986; Theissen et al., 2003; Marcia et al., 2010a; Nagy et al., 2014; Den Uyl et al., 2016). SQR is an evolutionarily ancient enzyme that was widespread in organisms during the Proterozoic

(Theissen et al., 2003), during which the oceans had variable redox conditions including euxinia. Although the evolutionary history of *sqr* in cyanobacteria appears to be complex and remains unresolved, the *sqr1* gene is thought to be endogenous to cyanobacteria (Theissen et al., 2003). As a critical component of sulfide-based AP, this gene would have enabled ancestral cyanobacteria to thrive during the Proterozoic, when periods of photic zone euxinia would have favored AP and tempered oxygen production (Johnston et al., 2009). Whether the endogenous type I SQR may have permitted ancestral cyanobacteria to oxidize sulfide for energy, or cyanobacteria retooled this detoxifying enzyme for AP, is an open question.

The role of the second *sqr* operon in *Geitlerinema* is unknown. In addition to sulfide consumption through *sqr1*-based AP (Oren and Padan, 1978), *Geitlerinema* has a second sulfide donation site on the immediate donor side of PSI, which is not inducible (i.e., does not require protein synthesis) and does not significantly contribute to proton translocation (Shahak et al., 1987). However, this response operates at even higher concentrations of sulfide (Km in the mM range) without saturation (Shahak et al., 1987; Arieli et al., 1991). The enzyme mediating this response to high sulfide is unknown, thus it could be encoded by this *sqr2*. Only 10 of the 44 cyanobacterial genomes and genomic bins examined in this study (including *Geitlerinema* sp. PCC 9228) have more than one version of *sqr*. Seven genomes in this subset, such as *Synechocystis* sp. PCC6803 (Nagy et al., 2014), each have two versions *sqr*, one of which is phylogenetically similar to *sqr1*, and the other a eukaryotic homolog (type II). Only three genomes, all of them *Geitlerinema* species, have genes phylogenetically similar to *sqr2* (type VI) as well as an *sqr1*-like version. The sulfide physiology of *Geitlerinema* sp. PCC7105 is unknown, but *Geitlerinema* sp. BBD is a known sulfide-resistant photosynthetic cyanobacterium (Den Uyl et al., 2016). Green sulfur bacteria, thiotrophic proteobacteria, and members of

Aquificaceae have multiple versions of *sqr* or homologs with a range of sulfide affinities. *Aquifex aeolicus* expresses its types I and VI *sqr*, similar to *Geitlerinema's sqr1* and *sqr2*, even when not growing thiotrophically (Marcia et al., 2010a,b). On the other hand, *Chlorobium* is capable of sulfide-dependent growth at mM H_2S by using its type VI *sqr*, and other versions are used at lower sulfide concentrations (Chan et al., 2009). Phylogenetic clustering of *sqr2* with green sulfur bacterial *sqrs*, including one used at high sulfide levels, supports a similar role in *Geitlerinema*. Like *sqr1*, *sqr2* also has a transcriptional regulatory gene *arsR* in the operon. These observations raise the possibility that *Geitlerinema* is capable of growing phototrophically on different sulfide levels through its *sqr1* and *sqr2*. Our results also suggest that such versatility was likely achieved through a combination of evolutionary processes including vertical descent within ancestral cyanobacteria (*sqr1*, type I clade) as well as lateral transfer from other groups (*sqr2*, type VI clade). When these genes became part of the AP-cyanobacterial repertoire is uncertain.

Chromosomal examination of cyanobacterial *sqr* and groups 2 and 3 *psbA* genes underscores a potential relationship between sulfide exposure and reformulation of the water-oxidizing complex in periods of AP and/or OP. Of the evaluated cyanobacterial genomes with *sqr*, 15 of the genomes have *sqr* types I, II, or VI in close genetic proximity to low-oxygen or anaerobic *psbA* genes. These genomes include metagenomic bins from Middle Island Sinkhole, where sulfide-based AP has been demonstrated (Voorhies et al., 2012), two other species of *Geitlerinema*, four *Cyanothece* species, and two *Leptolyngbya* species, among others. Because of variable contig lengths, it is unknown if *psbA1* and/or *psbA2* of *Geitlerinema* sp. PCC 9228 are in proximity to one or both of its *sqr* genes. However, the small intergenetic spaces between anoxygenic/micro-oxygenic *psbA* varieties and *sqr* in the other genomes hints at potential linked

transcriptional regulation of these genes. Additionally, of the cyanobacterial *sqr* genes, 27 were in close proximity to *arsR*- like transcriptional regulators. These genes and conditions of their expression are an attractive target for further studies.

Potential Trace Metal Oxidation.

Different versions of *sqr* in *Geitlerinema* sp. PCC 9228 may also be linked to trace metal metabolism and resistance. In close proximity to *sqr2* there are genes for *arsC*, *arsB*, *arsR*, and glutathione (GST) *S*-transferase, and on other contigs there are genes for arsenite transporting ATP-ases *arsA*, another *arsB*, a methyltransferase similar to arsenite methylator *arsM*, and a chromate transporter similar to a cyanobacterial arsenite importer. These genes are involved in arsenate reduction, arsenite transport, transcriptional regulation, and mediation of arsenic resistance (Oden et al., 1994; Mukhopadhyay et al., 2002; Cameron and Pakrasi, 2010; Slyemi and Bonnefoy, 2011). *Synechocystis* sp. PCC6803 uses an *arsR* gene to regulate genomic *arsBHC* expression during arsenic exposure (López-Maury et al., 2003), and is able to grow in mM concentrations of arsenite and arsenate (Sánchez-Riego et al., 2014). An *arsR*-like transcriptional regulator is also adjacent to *sqr1* in *Geitlerinema* (Bronstein et al., 2000), and the proximity of *arsR* to *sqr2* suggests a similar regulatory role. Arsenic resistance genes are widely distributed among bacteria and archaea, and may be found in environments that do not have measurable arsenite (Mukhopadhyay et al., 2002; Oremland and Stolz, 2005). The hydrology of the habitat of *Geitlerinema,* Solar Lake, is driven primarily by seawater seepage through the sand bar, with minor meteoric input (Cohen et al., 1977a). Usual sources of arsenic, such as hydrothermal hot springs in hypersaline environments or weathering of arsenic- rich clays (Oremland et al., 2009), are absent from this system. As such, these genes may have been

inherited from ancestral cyanobacteria inhabiting a metal-rich environment (Oremland et al., 2009), be used to detoxify other metals such as antimony (Nagy et al., 2014), or cope with reactive oxygen species and oxidative stress (Latifi et al., 2009; Takahashi et al., 2011).

The proximity of the *sqr2* to arsenic-related genes (*arsRBC*, glutathione *S*-transferase), taken together with experimental results from related organisms and genes, hints at an unexplored potential metabolism in *Geitlerinema*: AP using arsenite as the electron donor, as in anoxygenic bacteria (Kulp et al., 2008; Hoeft et al., 2010; Edwardson et al., 2014). *Oscillatoria*-like cyanobacterial biofilms in an arsenic-rich hot spring, and *Synechocystis* sp. PCC 6803, can perform light-dependent arsenite oxidation (Kulp et al., 2008; Nagy et al., 2014). In *Synechocystis* sp. PCC 6803, the same *sqr* and *arsR* genes from its plasmid *sqr* operon, which enable light-dependent sulfide oxidation, are expressed during arsenite (As(III)) oxidation in the light (Nagy et al., 2014). The cyanobacterium imports arsenite through a chromate transporter (*suoT*, on the same plasmid operon), stores mM arsenite intracellularly without detriment, and exports excess arsenite through *arsB* (Nagy et al., 2014). The arrangement of *Geitlerinema's* arsenic-related genes on its *sqr2* operon is similar to the plasmid *sqr* operon of *Synechocystis* sp. PCC 6803. Arsenite oxidation has been linked to electron transfer to quinones (Jiang et al., 2009) and energetics support the potential of cyanobacterial plastoquinone as an arsenite oxidant (Nagy et al., 2014), but this pathway has not been explored in cyanobacteria. The co-transcription of *sqr1* in *Synechocystis* with arsenite uptake genes during arsenite exposure (Nagy et al., 2014), and the well-characterized electron-stripping mechanism of *sqr1* on sulfide (Arieli et al., 1994), hints at a role for *sqr* in arsenite oxidation.

Arsenite-dependent cyanobacterial AP is intriguing due to the role of arsenite-based primary production on ancient Earth. Arsenic was likely more abundant on Earth's surface

during the Archean than present (Oremland et al., 2009). In anoxic marine basins that dominated the biosphere 2.7 Ga, arsenite- dependent microbial autotrophy putatively cycled nitrogen, carbon, and arsenic (Sforna et al., 2014). This metabolism continues in modern hypersaline hot springs and subsurface aquifers that have elevated arsenic levels (Oremland and Stolz, 2005; Kulp et al., 2008; Hoeft et al., 2010). The genome of *Geitlerinema* lacks proteobacterial arsenite-oxidizing genes, and instead has genes similar to those used for light-dependent arsenite oxidation in *Synechocystis* sp. PCC 6803 (Nagy et al., 2014). Verifying arsenite-dependent AP in cyanobacteria, and conclusively linking the *sqr2* gene in *Geitlerienema* to that process, would clarify the potential role of anoxygenic cyanobacteria in arsenic cycling in both modern and ancient ecosystems.

In summary, analysis of the genome of *Geitlerinema* sp. PCC9228 complements prior physiology studies by providing the genetic foundation for its metabolisms of nitrogen fixation, facultative OP, and sulfide-based AP. We find multiple versions of *psbA*, encoding a key protein for water oxidation, which may enable a sensitive response to varying conditions of light, oxygen and sulfide. Nitrogen fixation is linked to oxygen level and production via the $nifI_1I_2$ regulator in the *nif* operon and via non-oxygen producing *psbA*, respectively. Multiple versions of *sqr* likely address a range of sulfide concentrations and may also be linked to responses to metals and oxidative stress and perhaps even arsenite oxidation. Aerobic versatility encoded in the genome of *Geitlerinema*, coupled with diazotrophic regulation and concentration-specific sulfide responses, permit *Geitlerinema* to thrive in periodic sulfidic, microoxic, and poorly lit conditions of Solar Lake (Cohen et al., 1975a). Such dynamic geochemical conditions likely also challenged cyanobacteria during variable sulfide and oxygen levels of the Archean and Proterozoic (Satkoski et al., 2015; Sperling et al., 2015). This study of *Geitlerinema* and its unique gene

assemblage addresses both the ambiguous role of Archean cyanobacteria in oxygen production/mitigation prior to the development of OP and the Great Oxidation Event, as well as the contributions of AP cyanobacteria to Proterozoic biogeochemistry. The phylogeny and diversity of genes responsible for metabolic versatility in *Geitlerinema* suggest a blend of genetic strategies for the anoxic early environment—such as methanogen-like modulation of nitrogen fixation and non-water oxidizing photosynthetic proteins—with post-oxidation strategies such as specific photosynthetic proteins for micro-oxic as well as oxic conditions, and different SQRs for fluctuating sulfide concentrations. Phototrophs capable of versatile AP/OP, such as *Geitlerinema*, would have had the advantage over organisms metabolically limited to either oxic or sulfidic conditions. Their continuous photosynthesis likely supported other microorganisms with fixed nitrogen and carbon, sulfide removal, and intermittent oxygen production. Furthermore, conditional production of oxygen at variable concentrations would have had strong influences on the structure and metabolic needs of their associated microbial communities through development of oxygen refugia and/or oases. Further research into *Geitlerinema's* growth and transcriptional regulation will uncover the fine-tuned response of AP/OP cyanobacteria to changing redox conditions. In turn, we can relate the scope of these dual metabolisms and their modern physiologies to their ancestors' impacts on ecology and geochemistry as Earth slowly and discontinuously became oxygenated.

2.6 Acknowledgements

We thank Tanai Cardona and Eric Boyd for providing cyanobacterial *psbA* and *nifHDK* sequences used in this study, and for helpful discussion. Prof. Aharon Oren kindly provided the culture of *Geitlerinema* sp. PCC 9228. We thank Paul Den Uyl for culturing the cyanobacterium

and extracting DNA for sequencing, and for providing sequences from *Geitlerinema* sp. BBD for analyses. We thank Sunit Jain and Judith Klatt of the University of Michigan Geomicrobiology Lab for bioinformatics assistance and thoughtful feedback and discussion.

2.7 References

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi: 10.1038/nmeth.3103

Arieli, B., Padan, E., and Shahak, Y. (1991). Sulfide-induced sulfide-quinone reductase activity in thylakoids of *Oscillatoria limnetica*. *J. Biol. Chem.* 266, 104–111.

Arieli, B., Shahak, Y., Taglicht, D., Hauska, G., and Padan, E. (1994). Purification and characterization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in *Oscillatoria limnetica*. *J. Biol. Chem.* 269, 5705– 5711.

Aro, E.-M., Virgin, I., and Andersson, B. (1993). Photoinhibition of photosystem II. inactivation, protein damage and turnover. *Biochim. Biophys. Acta* 1143, 113–134. doi: 10.1016/0005-2728(93)90134-2

Becraft, E. D., Wood, J. M., Rusch, D., Kühl, B. M., Jensen, S. I., Bryant, D. A., et al. (2015). The molecular dimension of microbial species: 1. ecological distinctions among, and homogeneity within, putative ecotypes of *Synechococcus* inhabiting the cyanobacterial mat of mushroom spring, yellowstone national park. *Front. Microbiol.* 6:590. doi: 10.3389/fmicb.2015.00590

Belkin, S., Arieli, B., and Padan, E. (1982). Sulfide dependent electron transport in *Oscillatoria limnetica*. *Isr. J. Bot.* 31, 199–200. doi: 10.1080/0021213X.1982.10676943

Belkin, S., and Padan, E. (1978). Sulfide-dependent hydrogen evolution in the cyanobacterium *Oscillatoria limnetica*. *FEBS Lett.* 94, 291–294. doi: 10.1016/0014-5793(78)80959-4

Blank, C. E., and Sánchez-Baracaldo, P. (2010). Timing of morphological and ecological innovations in the cyanobacteria – a key to understanding the rise in atmospheric oxygen. *Geobiology* 8, 1–23. doi: 10.1111/j.1472-4669.2009.00220.x

Blankenship, R. E. (2002). *Molecular Mechanisms of Photosynthesis*, 1st Edn. Oxford: Blackwell Science Ltd.

Blankenship, R. E. (2010). Early evolution of photosynthesis. *Plant Physiol.* 154, 434–438. doi: 10.1104/pp.110.161687

Boyd, E. S., Anbar, A. D., Miller, S., Hamilton, T. L., Lavin, M., and Peters, J. W. (2011). A late methanogen origin for molybdenum-dependent nitrogenase. *Geobiology* 9, 221–232. doi: 10.1111/j.1472-4669.2011.00278.x

Boyd, E. S., Costas, A. M. G., Hamilton, T. L., Mus, F., and Peters, J. W. (2015). Evolution of molybdenum nitrogenase during the transition from anaerobic to aerobic metabolism. *J. Bacteriol.* 197, 1690–1699. doi: 10.1128/JB. 02611-14

Bronstein, M., Schütz, M., Hauska, G., Padan, E., and Shahak, Y. (2000). Cyanobacterial sulfide-quinone reductase: cloning and heterologous expression. *J. Bacteriol.* 182, 3336–3344. doi: 10.1128/JB.182.12.3336-3344. 2000

Bühring, S., Sievert, S. M., Jonkers, H. M., Ertefai, T., Elshahed, M. S., Krumholz, L. R., et al. (2011). Insights into chemotaxonomic composition and carbon cycling of phototrophic communities in an artesian sulfur- rich spring (Zodletone, Oklahoma, USA), a possible analog for ancient microbial mat systems. *Geobiology* 9, 166–179. doi: 10.1111/j.1472-4669.2010. 00268.x

Cameron, J. C., and Pakrasi, H. B. (2010). Essential role of glutathione in acclimation to environmental and redox perturbations in the cyanobacterium synechocystis sp. PCC 6803. *Plant Physiol.* 154, 1672–1685. doi: 10.1104/pp.110.162990

Campbell, D. A., Clarke, A. K., Gustafsson, P., and Öquist, G. (1999). Oxygen- dependent electron flow influences photosystem II function and *psbA* gene expression in the cyanobacterium *Synechococcus* sp. PCC 7942. *Physiol. Plant.* 105, 746–755. doi: 10.1034/j.1399-3054.1999.105420.x

Canfield, D. E. (1998). A new model for proterozoic ocean chemistry. *Nature* 396, 450–453. doi: 10.1038/24839

Canfield, D. E. (2005). The early history of atmospheric oxygen: homage to Robert M. Garrels. *Annu. Rev. Earth Planet. Sci.* 33, 1–36. doi: 10.1146/annurev.earth.33.092203.122711

Cardona, T., Murray, J. W., and Rutherford, A. W. (2015). Origin and evolution of water oxidation before the last common ancestor of the Cyanobacteria. *Mol. Biol. Evol.* 32, 1310–1328. doi: 10.1093/molbev/msv024

Chan, L.-K., Morgan-Kiss, R. M., and Hanson, T. E. (2009). Functional analysis of three sulfide:quinone oxidoreductase homologs in *Chlorobaculum tepidum*. *J. Bacteriol.* 191, 1026–1034. doi: 10.1128/JB.01154-08

Chaudhary, A., Haack, S. K., Duris, J. W., and Marsh, T. L. (2009). Bacterial and archaeal phylogenetic diversity of a cold sulfur-rich spring on the shoreline of Lake Erie, Michigan. *Appl. Environ. Microbiol.* 75, 5025–5036. doi: 10.1128/AEM.00112-09

Coates, R. C., Podell, S., Korobeynikov, A., Lapidus, A., Pevzner, P., Sherman, D. H., et al. (2014). Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. *PloS ONE* 9:e85140. doi: 10.1371/journal.pone.0085140

Cohen, Y., Jørgensen, B. B., Padan, E., and Shilo, M. (1975a). Sulphide-dependent anoxygenic photosynthesis in the cyanobacterium *Oscillatoria limnetica*. *Nature* 257, 489–492. doi: 10.1038/257489a0

Cohen, Y., Jørgensen, B. B., Revsbech, N. P., and Poplawski, R. (1986). Adaptation to hydrogen sulfide of oxygenic and anoxygenic photosynthesis among cyanobacteria. *Appl. Environ. Microbiol.* 51, 398–407. .

Cohen, Y., Krumbein, W. E., Goldberg, M., and Shilo, M. (1977a). Solar Lake (Sinai). 1. Physical and chemical limnology. *Limnol. Oceanogr.* 22, 597–608. doi: 10.4319/lo.1977.22.4.0597

Cohen, Y., Krumbein, W. E., and Shilo, M. (1977b). Solar Lake (Sinai). 2. Distribution of photosynthetic microorganisms and primary production. *Limnol. Oceanogr.* 22, 609–620. doi: 10.4319/lo.1977.22.4.0609

Cohen, Y., Padan, E., and Shilo, M. (1975b). Facultative anoxygenic photosynthesis in the cyanobacterium *Oscillatoria limnetica*. *J. Bacteriol.* 123, 855–861.

Cole, J. K., Hutchison, J. R., Renslow, R. S., Kim, Y.-M., Chrisler, W. B., Engelmann, H. E., et al. (2014). Phototrophic biofilm assembly in microbial-mat-derived unicyanobacterial consortia: model systems for the study of autotroph-heterotroph interactions. *Front. Microbiol.* 5:109. doi: 10.3389/fmicb.2014.00109

Crowe, S. A., Døssing, L. N., Beukes, N. J., Bau, M., Kruger, S. J., Frei, R., et al. (2013). Atmospheric oxygenation three billion years ago. *Nature* 501, 535–538. doi: 10.1038/nature12426

Den Uyl, P. A., Richardson, L. L., Jain, S., and Dick, G. J. (2016). Unraveling the physiological roles of the cyanobacterium *Geitlerinema* sp. BBD and other black band disease community members through genomic analysis of a mixed culture. *PloS ONE* 11:e0157953. doi: 10.1371/journal.pone. 0157953

Dick, G. J., Andersson, A. F., Baker, B. J., Simmons, S. L., Thomas, B. C., Yelton, A. P., et al. (2009). Community-wide analysis of microbial genome sequence signatures. *Genome Biol.* 10:R85. doi: 10.1186/gb-2009-10-8-r85

Edwardson, C. F., Planer-Friedrich, B., and Hollibaugh, J. T. (2014). Transformation of monothioarsenate by haloalkaliphilic, anoxygenic photosynthetic purple sulfur bacteria. *FEMS Microbiol. Ecol.* 90, 858–868. doi: 10.1111/1574-6941.12440

Farquhar, J., Zerkle, A. L., and Bekker, A. (2010). Geological constraints on the origin of oxygenic photosynthesis. *Photosynth. Res.* 107, 11–36. doi: 10.1007/s11120-010-9594-0

Ferreira, K. N., Iverson, T. M., Maghlaoui, K., Barber, J., and Iwata, S. (2004). Architecture of the photosynthetic oxygen-evolving center. *Science* 303, 1831– 1838. doi: 10.1126/science.1093087

Findlay, A. J., Bennett, A. J., Hanson, T. E., and Luther, G. W. (2015). Light- dependent sulfide oxidation in the anoxic zone of the chesapeake bay can be explained by small populations of phototrophic bacteria. *Appl. Environ. Microbiol.* 81, 7560–7569. doi: 10.1128/AEM.02062-15

Fischer, W. W., Hemp, J., and Johnson, J. E. (2015). Manganese and the evolution of photosynthesis. *Orig. Life Evol. Biosph.* 45, 351–357. doi: 10.1007/s11084-015- 9442-5

Fischer, W. W., Hemp, J., and Valentine, J. S. (2016). How did life survive Earth's great oxygenation? *Curr. Opin. Chem. Biol.* 31, 166–178. doi: 10.1016/j.cbpa.2016.03.013

Forchhammer, K. (2004). Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiol. Rev.* 28, 319–333. doi: 10.1016/j.femsre.2003.11.001

Gan, F. S. Z., Rockwell, N. C., Martin, S. S., Lagarias, J. C., and Bryant, D. A. (2014). Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* 345, 1312–1317. doi: 10.1126/science.1256963

Gregersen, L. H., Bryant, D. A., and Frigaard, N. U. (2011). Mechanisms and evolution of oxidative sulfur metabolism in green sulfur bacteria. *Front. Microbiol.* 2:116. doi: 10.3389/fmicb.2011.00116

Grotzinger, J. P., and Knoll, A. H. (1999). Stromatolites in precambrian carbonates: evolutionary mileposts or environmental dipsticks? *Annu. Rev. Earth Planet. Sci.* 27, 313–358. doi: 10.1146/annurev.earth.27.1.313

Hanson, T. E., Luther, G. W. III., and Findlay, A. J. (2013). Phototrophic sulfide oxidation: environmental insights and a method for kinetic analysis. *Front. Microbiol.* 4:382. doi: 10.3389/fmicb.2013.00382

Ho, M. Y., Shen, G., Canniffe, D. P., Zhao, C., and Bryant, D. A. (2016). Light- dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II. *Science* 353. doi: 10.1126/science.aaf9178

Hoeft, S. E., Kulp, T. R., Han, S., Lanoil, B., and Oremland, R. S. (2010). Coupled arsenotrophy in a hot spring photosynthetic biofilm at Mono Lake, California. *Appl. Environ. Microbiol.* 76, 4633–4639. doi: 10.1128/AEM.00545-10

Jiang, J., Bauer, I., Paul, A., and Kappler, A. (2009). Arsenic redox changes by microbially and chemically formed semiquinone radicals and hydroquinones in a humic substance model quinone. *Environ. Sci. Technol.* 43, 3639–3645. doi: 10.1021/es803112a

Johnston, D. T., Wolfe-Simon, F., Pearson, V., and Knoll, A. H. (2009). Anoxygenic photosynthesis modulated proterozoic oxygen and sustained Earth's middle age. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16925–16929. doi: 10.1073/pnas.0909248106

Jørgensen, B. B., Cohen, Y., and Revsbech, N. P. (1986). Transition from anoxygenic to oxygenic photosynthesis in a microcoleus chthonoplastes cyanobacterial mat. *Appl. Environ. Microbiol.* 51, 408–417.

Joshi, N. A., and Fass, J. N. (2011). *Sickle: A Sliding-Window, Adaptive, Quality- Based Trimming Tool for FastQ Files*. Available at: https://github.com/najoshi/s ickle.

Kessler, P. S., Daniel, C., and Leigh, J. A. (2001). Ammonia switch-off of nitrogen fixation in the methanogenic archaeon *Methanococcus maripaludis*: mechanistic features and requirement for the novel GlnB homologues, NifI(1) and NifI(2). *J. Bacteriol.* 183, 882–889. doi: 10.1128/JB.183.3.882-889. 2001

Kiss, É, Kós, P. B., Chen, M., and Vass, I. (2012). A unique regulation of the expression of the *psbA*, *psbD*, and *psbE* genes, encoding the D1, D2 and cytochrome b559 subunits of the photosystem II complex in the chlorophyll d containing cyanobacterium *Acaryochloris marina*. *Biochim. Biophys. Acta* 1817, 1083–1094. doi: 10.1016/j.bbabio.2012.04.010

Klatt, J. M., Al-Najjar, M. A. A., Yilmaz, P., Lavik, G., de Beer, D., and Polerecky, L. (2015a). Anoxygenic photosynthesis controls oxygenic photosynthesis in a cyanobacterium from a sulfidic spring. *Appl. Environ. Microbiol.* 81, 2025–2031. doi: 10.1128/AEM.03579-14

Klatt, J. M., Haas, S., Yilmaz, P., de Beer, D., and Polerecky, L. (2015b). Hydrogen sulfide can inhibit and enhance oxygenic photosynthesis in a cyanobacterium from sulfidic springs. *Environ. Microbiol.* 17, 3301–3313. doi: 10.1111/1462- 2920.12791

Klatt, J. M., Meyer, S., Häusler, S., Macalady, J. L., de Beer, D., and Polerecky, L. (2016). Structure and function of natural sulphide-oxidizing microbial mats under dynamic input of light and chemical energy. *ISME J.* 10, 921–933. doi: 10.1038/ismej.2015.167

Kós, P. B., Deák, Z., Cheregi, O., and Vass, I. (2008). Differential regulation of *psbA* and *psbD* gene expression, and the role of the different D1 protein copies in the Cyanobacterium *Thermosynechococcus elongatus* BP-1. *Biochim. Biophys. Acta* 1777, 74–83. doi: 10.1016/j.bbabio.2007.10.015

Kulp, T. R., Hoeft, S. E., Asao, M., Madigan, M. T., Hollibaugh, J. T., Fisher, J. C., et al. (2008). Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from mono lake, California. *Science* 321, 967–970. doi: 10.1126/science.1160799

Lalonde, S. V., and Konhauser, K. O. (2015). Benthic perspective on Earth's oldest evidence for oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 995–1000. doi: 10.1073/pnas.1415718112

Latifi, A., Ruiz, M., and Zhang, C.-C. (2009). Oxidative stress in cyanobacteria. *FEMS Microbiol. Rev.* 3, 258–278. doi: 10.1111/j.1574-6976.2008.00134.x

López-Maury, L., Florencio, F. J., and Reyes, J. C. (2003). Arsenic sensing and resistance system in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* 185, 5363–5371. doi: 10.1128/JB.185.18.5363-5371.2003

Lyons, T. W., Reinhard, C. T., and Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506, 307–315. doi: 10.1038/nature13068

Marcia, M., Ermler, U., Peng, G., and Michel, H. (2009). The structure of *Aquifex aeolicus* sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9625–9630. doi: 10.1073/pnas.0904165106

Marcia, M., Ermler, U., Peng, G., and Michel, H. (2010a). A new structure-based classification of sulfide:quinone oxidoreductases. *Proteins* 78, 1073–1083. doi: 10.1002/prot.22665

Marcia, M., Langer, J. D., Parcej, D., Vogel, V., Peng, G., and Michel, H. (2010b). Characterizing a monotopic membrane enzyme. biochemical, enzymatic and crystallization studies on *Aquifex aeolicus* sulfide:quinone oxidoreductase. *Biochim. Biophys. Acta* 1798, 2114–2123. doi: 10.1016/j.bbamem.2010.07.033

Miller, S. R., and Bebout, B. M. (2004). Variation in sulfide tolerance of photosystem II in phylogenetically diverse cyanobacteria from sulfidic habitats. *Appl. Environ. Microbiol.* 70, 736–744. doi: 10.1128/AEM.70.2.736-744.2004

Minamizaki, K., Mizoguchi, T., Goto, T., Tamiaki, H., and Fujita, Y. (2008). Identification of two homologous genes, *chlAI* and *chlAII*, that are differentially involved in isocyclic ring formation of chlorophyll a in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 283, 2684–2692. doi: 10.1074/jbc.M708954200

Mohamed, A., Eriksson, J., Osiewacz, H. D., and Jansson, C. (1993). Differential expression of the *psbA* genes in the cyanobacterium *Synechocystis* 6803. *Mol. Gen. Genet.* 238, 161–168. doi: 10.1007/BF00279543

Mukhopadhyay, R., Rosen, B. P., Phung, L. T., and Silver, S. (2002). Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiol. Rev.* 26, 311– 325. doi: 10.1111/j.1574-6976.2002.tb00617.x

Mulkidjanian, A. Y., Koonin, E. V., Makarova, K. S., Mekhedov, S. L., Sorokin, A., Wolf, Y. I., et al. (2006). The cyanobacterial genome core and the origin of photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13126–13131. doi: 10.1073/pnas.0605709103

Murray, J. W. (2012). Sequence variation at the oxygen-evolving centre of photosystem II: a new class of 'rogue' cyanobacterial D1 proteins. *Photosynth. Res.* 110, 177–184. doi: 10.1007/s11120-011-9714-5

Nagarajan, N., and Pop, M. (2013). Sequence assembly demystified. *Nat. Rev. Genet.* 14, 157–167. doi: 10.1038/nrg3367

Nagy, C. I., Vass, I., Rákhely, G., Vass, I. Z., Tóth, A., Duzs, A., et al. (2014). Coregulated genes link sulfide:quinone oxidoreductase and arsenic metabolism in *Synechocystis* sp. strain PCC6803. *J. Bacteriol.* 196, 3430–3440. doi: 10.1128/JB.01864-14

Oden, K. L., Gladysheva, T. B., and Rosen, B. P. (1994). Arsenate reduction mediated by the plasmid-encoded ArsC protein is coupled to glutathione. *Mol. Microbiol.* 12, 301–306. doi: 10.1111/j.1365-2958.1994.tb01018.x

Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., LaButti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. comparative genomics of *Synechococcus* strains with different light responses and in situ diel transcription patterns of associated putative ecotypes in the mushroom spring microbial mat. *Front. Microbiol.* 6:604. doi: 10.3389/fmicb.2015.00604

Oremland, R. S., Saltikov, C. W., Wolfe-Simon, F., and Stolz, J. F. (2009). Arsenic in the evolution of earth and extraterrestrial ecosystems. *Geomicrobiol. J.* 26, 522–536. doi: 10.1080/01490450903102525

Oremland, R. S., and Stolz, J. F. (2005). Arsenic, microbes and contaminated aquifers. *Trends Microbiol.* 13, 45–49. doi: 10.1016/j.tim.2004.12.002

Oren, A., and Padan, E. (1978). Induction of anaerobic, photoautotrophic growth in the cyanobacterium *Oscillatoria limnetica*. *J. Bacteriol.* 133, 558–563.

Park, J.-J., Lechno-Yossef, S., Wolk, C. P., and Vieille, C. (2013). Cell-specific gene expression in anabaena variabilis grown phototrophically, mixotrophically, and heterotrophically. *BMC Genomics* 14:759. doi: 10.1186/1471-2164-1 4-759

Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi: 10.1093/bioinformatics/bts174

Pham, V. H., Yong, J.-J., Park, S.-J., Yoon, D.-N., Chung, W.-H., and Rhee, S.- K. (2008). Molecular analysis of the diversity of the sulfide:quinone reductase (Sqr) gene in sediment environments. *Microbiology* 154, 3112–3121. doi: 10.1099/mic.0.2008/018580-0

Planavsky, N. J., Asael, D., Hofmann, A., Reinhard, C. T., Lalonde, S. V., Knudsen, A., et al. (2014). Evidence for oxygenic photosynthesis half a billion years before the great oxidation event. *Nat. Geosci.* 7, 283–286. doi: 10.1038/ngeo2122

Raes, J., Korbel, J. O., Lercher, M. J., von Mering, C., and Bork, P. (2007). Prediction of effective genome size in metagenomic samples. *Genome Biol.* 8:R10. doi: 10.1186/gb-2007-8-1-r10

Rambaut, A. (2012). *FigTree Version 1.4.0*. Available at: http://tree.bio.ed.ac.uk/so ftware/figtree

Raymond, J., Siefert, J. L., Staples, C. R., and Blankenship, R. E. (2004). The natural history of nitrogen fixation. *Mol. Biol. Evol.* 21, 541–554. doi: 10.1093/molbev/msh047

Reddy, S. M., and Evans, D. A. D. (2009). Palaeoproterozoic supercontinents and global evolution: correlations from core to atmosphere. *Geol. Soc. Spec. Publications* 323, 1–26. doi: 10.1144/SP323.1

Ricci, J. N., Michel, A. J., and Newman, D. K. (2015). Phylogenetic analysis of HpnP reveals the origin of 2-methylhopanoid production in alphaproteobacteria. *Geobiology* 13, 267–277. doi: 10.1111/gbi.12129

Sánchez-Riego, A. M., López-Maury, L., and Florencio, F. J. (2014). Genomic responses to arsenic in the cyanobacterium *Synechocystis* sp. PCC 6803. *PloS ONE* 9:e96826. doi: 10.1371/journal.pone.0096826

Satkoski, A. M., Beukes, N. J., Li, W., Beard, B. L., and Johnson, C. M. (2015). A redox-stratified ocean 3.2 billion years ago. *Earth Planet. Sci. Lett.* 430, 43–53. doi: 10.1016/j.epsl.2015.08.007

Schaefer, M. R., and Golden, S. S. (1989). Differential expression of members of a cyanobacterial *psbA* gene family in response to light. *J. Bacteriol.* 171, 3973–3981.

Schirmer, A., Rude, M. A., Li, X., Popova, E., and del Cardayre, S. B. (2010). Microbial biosynthesis of alkanes. *Science* 329, 559–562. doi: 10.1126/science.1187936

Schütz, M., Maldener, I., Griesbeck, C., and Hauska, G. (1999). Sulfide-quinone reductase from *Rhodobacter capsulatus*: requirement for growth, periplasmic localization, and extension of gene sequence analysis. *J. Bacteriol.* 181, 6516– 6523.

Scott, C., Lyons, T. W., Bekker, A., Shen, Y., Poulton, S. W., Chu, X., et al. (2008). Tracing the stepwise oxygenation of the proterozoic ocean. *Nature* 452, 456–459. doi: 10.1038/nature06811

Sforna, M. C., Philippot, P., Somogyi, A., van Zuilen, M. A., Medjoubi, K., Schoepp-Cothenet, B., et al. (2014). Evidence for arsenic metabolism and cycling by microorganisms 2.7 billion years ago. *Nat. Geosci.* 7, 811–815. doi: 10.1038/ngeo2276

Shahak, Y., Arieli, B., Binder, B., and Padan, E. (1987). Sulfide-dependent photosynthetic electron flow coupled to proton translocation in thylakoids of the cyanobacterium *Oscillatoria limnetica*. *Arch. Biochem. Biophys.* 259, 605–615. doi: 10.1016/0003-9861(87)90527-3

Sicora, C., Wiklund, R., Jansson, C., and Vass, I. (2004). Charge stabilization and recombination in photosystem II containing the D1′ protein product of the *psbA1* gene in *Synechocystis* 6803. *Phys. Chem. Chem. Phys.* 6, 4832–4837. doi: 10.1039/B406695B

Sicora, C. I., Appleton, S. E., Brown, C. M., Chung, J., Chandler, J., Cockshutt, A. M., et al. (2006). Cyanobacterial psbA families in *Anabaena* and *Synechocystis* encode trace, constitutive and UVB-induced D1 isoforms. *Biochim. Biophys. Acta* 1757, 47–56. doi: 10.1016/j.bbabio.2005.11.002

Sicora, C. I., Ho, F. M., Salminen, T., Styring, S., and Aro, E.-M. (2009). Transcription of a 'Silent' cyanobacterial *psbA* gene is induced by microaerobic conditions. *Biochim. Biophys. Acta* 1787, 105–112. doi: 10.1016/j.bbabio.2008.12.002

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. *Mol. Syst. Biol.* 7, 539–539. doi: 10.1038/msb.2011.75

Slyemi, D., and Bonnefoy, V. (2011). How prokaryotes deal with arsenic. *Environ. Microbiol. Rep.* 4, 571–586. doi: 10.1111/j.1758-2229.2011.00300.x

Sousa, F. L., Shavit-Grievink, L., Allen, J. F., and Martin, W. F. (2013). Chlorophyll biosynthesis gene evolution indicates photosystem gene duplication, not photosystem merger, at the origin of oxygenic photosynthesis. *Genome Biol. Evol.* 5, 200–216. doi: 10.1093/gbe/evs127

Sperling, E. A., Wolock, C. J., Morgan, A. S., Gill, B. C., Kunzmann, M., Halverson, G. P., et al. (2015). Statistical analysis of iron geochemical data suggests limited late proterozoic oxygenation. *Nature* 523, 451–454. doi: 10.1038/nature14589

Stal, L. J. (1995). Physiological ecology of cyanobacteria in microbial mats and other communities. *New Phytol.* 131, 1–32. doi: 10.1111/j.1469- 8137.1995.tb03051.x

Stal, L. J. (2012). "Cyanobacterial mats and stromatolites," in *Ecology of Cyanobacteria II*, ed. B. A. Whitton (Dordrecht: Springer), 65–125. doi: 10.1007/978-94-007-3855-3_4

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Steinhauser, D., Fernie, A. R., and Araújo, W. L. (2012). Unusual cyanobacterial TCA cycles: not broken just different. *Trends Plant Sci.* 17, 503–509. doi: 10.1016/j.tplants.2012.05.005

Stüeken, E. E., Buick, R., Guy, B. M., and Koehler, M. C. (2015). Isotopic evidence for biological nitrogen fixation by molybdenum-nitrogenase from 3.2 Gyr. *Nature* 520, 666–669. doi: 10.1038/nature14180

Sugiura, M., Kato, Y., Takahashi, R., Suzuki, H., Watanabe, T., Noguchi, T., et al. (2010). Energetics in photosystem II from *Thermosynechococcus elongatus* with a D1 protein encoded by either the *psbA1* or *psbA3* gene. *Biochim. Biophys. Acta* 1797, 1491–1499. doi: 10.1016/j.bbabio.2010.03.022

Sugiura, M., Ogami, S., Kusumi, M., and Un, S. (2012). Environment of TYRZ in photosystem II from *Thermosynechococcus elongatus* in which PsbA2 Is the D1 protein. *J. Biol. Chem.* 287, 13336–13347. doi: 10.1074/jbc.M112. 340323

Summons, R. E., and Lincoln, S. A. (2012). "Biomarkers: informative molecules for studies in geobiology," in *Fundamentals of Geobiology*, eds A. H. Knoll, D. E. Canfield, and K. O. Konhauser (Chichester: John Wiley & Sons, Ltd) doi: 10.1002/9781118280874.ch15

Sumner, D. Y., Hawes, I., Mackey, T. J., Jungblut, A. D., and Doran, P. T. (2015). Antarctic microbial mats: a modern analog for archean lacustrine oxygen oases. *Geology* 43, 887–890. doi: 10.1130/G36966.1

Takahashi, H., Kopriva, S., Giordano, M., Saito, K., and Hell, R. (2011). Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annu. Rev. Plant Biol.* 62, 157–184. doi: 10.1146/annurev-arplant-042110-103921

Tamagnini, P., Leitão, E., Oliveira, P., Ferreira, D., Pinto, F., Harris, D. J., et al. (2007). Cyanobacterial hydrogenases: diversity, regulation and applications. *FEMS Microbiol. Rev.* 31, 692–720. doi: 10.1111/j.1574-6976.2007.00085.x

Theissen, U., Hoffmeister, M., Grieshaber, M., and Martin, W. (2003). Single eubacterial origin of eukaryotic sulfide:quinone oxidoreductase, a mitochondrial enzyme conserved from the early evolution of eukaryotes during anoxic and sulfidic times. *Mol. Biol. Evol.* 20, 1564–1574. doi: 10.1093/molbev/msg174

Toepel, J., Welsh, E., Summerfield, T. C., Pakrasi, H. B., and Sherman, L. A. (2008). Differential transcriptional analysis of the cyanobacterium *Cyanothece* sp. strain *ATCC* 51142 during light-dark and continuous-light growth. *J. Bacteriol.* 190, 3904–3913. doi: 10.1128/JB.00206-08

van Lis, R., Nitschke, W., Duval, S., and Schoepp-Cothenet, B. (2013). Arsenics as bioenergetic substrates. *Biochim. Biophys. Acta* 1827, 176–188. doi: 10.1016/j.bbabio.2012.08.007

Voorhies, A. A., Biddanda, B. A., Kendall, S. T., Jain, S., Marcus, D. N., Nold, S. C., et al. (2012). Cyanobacterial life at low O2: community genomics and function reveal metabolic versatility and extremely low diversity in a great lakes sinkhole mat. *Geobiology* 10, 250–267. doi: 10.1111/j.1472-4669.2012.00322.x

Voorhies, A. A., Eisenlord, S. D., Marcus, D. N., Duhaime, M. B., Biddanda, B. A., Cavalcoli, J. D., et al. (2016). Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics. *Environ. Microbiol.* 18, 358–371. doi: 10.1111/1462-2920.12756

Wegener, K. M., Nagarajan, A., and Pakrasi, H. B. (2015). An atypical *psbA* gene encodes a sentinel D1 protein to form a physiologically relevant inactive photosystem II complex in cyanobacteria. *J. Biol. Chem.* 290, 3764–3774. doi: 10.1074/jbc.M114.604124

Yang, J., and Cheng, Q. (2004). Origin and evolution of the light-dependent protochlorophyllide oxidoreductase (LPOR) genes. *Plant Biol.* 6, 537–544. doi: 10.1055/s-2004-821270

Yin, X.-X., Chen, J., Qin, J., Sun, G.-X., Rosen, B. P., and Zhu, Y.-G. (2011). Biotransformation and volatilization of arsenic by three photosynthetic cyanobacteria. *Plant Physiol.* 156, 1631–1638. doi: 10.1104/pp.111.178947

Zhang, S., and Bryant, D. A. (2011). The tricarboxylic acid cycle in cyanobacteria. *Science* 334, 1551–1553. doi: 10.1126/science.1210858

2.8 Tables and Figures

Table 2.1 Integrated Microbial Genomes (IMG) accession numbers of genome and select genes of interest of *Geitlerinema* sp. PCC 9228.

Description	IMG ID
Genome ID	2660238729
nifH	2663545465
nifD	2663545468
nifK	2663545469
psbA1	2663545011
psbA2	2663544218
psbA3	2663548115
sqr1	2663545736
sqr2	2663547046

Figure 2.1. Metabolic schematic of *Geitlerinema* sp. PCC 9228. Genes for major processes [*psbA1-3, sqr1, ccm, rbcLS, nifHDK, hox, arsABC, ure, urt, focA, narK, amt*, glutathione *S*-transferase (GST)], and associated regulatory genes (*arsR, ntcA, nifI*$_1$*I*$_2$) are presented. Suppression of functions is indicated with red dashed lines ending in a flat line. Other interactions are indicated with black dashed lines ending with open circles. Transfer and production of key reactants and products are indicated with solid lines. Putative arsenite import is indicated with a dotted line, and anoxygenic photosynthesis-related processes have a solid blue line.

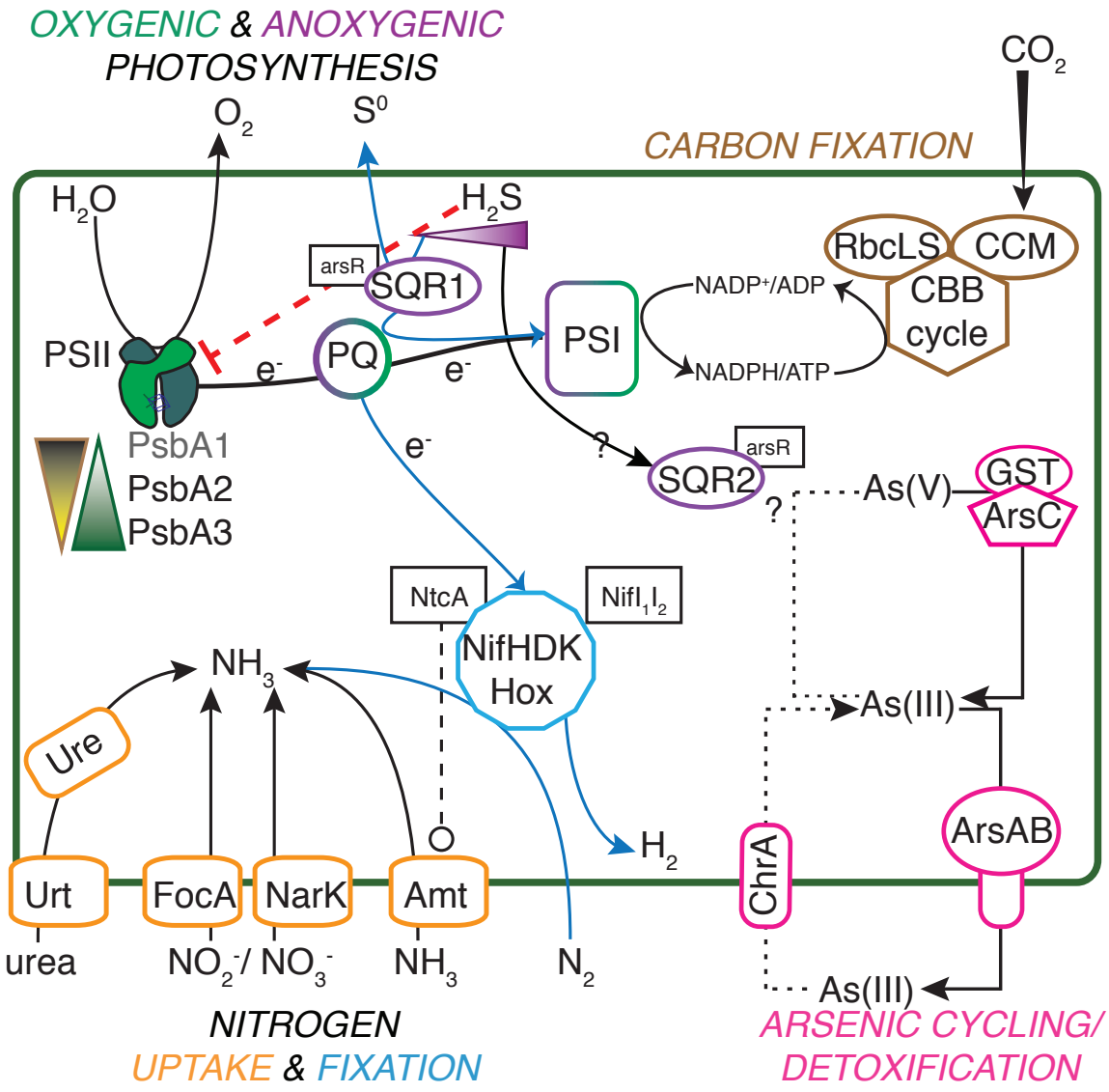


Figure 2.2. Schematic of nitrogenase genes in Geitlerinema sp. PCC 9228. *nifXSU, iscA, nifHI1I2DK, fdxN*, and *nifENB* are arranged in an apparent operon. *nifV* and a regulatory arsenic-related gene *arsA* are located upstream of the *nifHDK* operon.

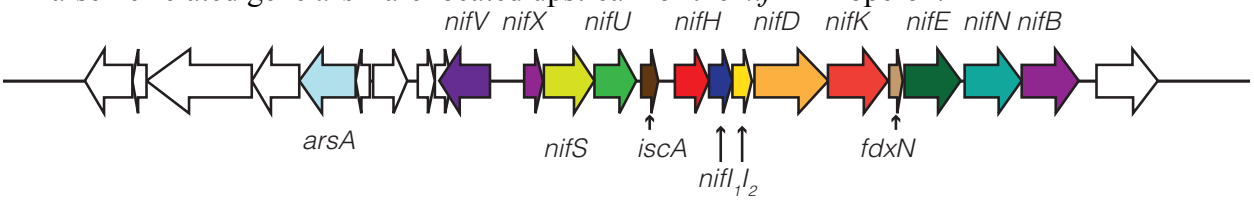


Figure 2.3. Phylogenetic tree of *psbA*. Genes are colored by groups 1–4 modeled after (Cardona et al., 2015). The outgroup is a *Gloeobacter kilaueensis JS psbA*, which is most similar to *psbD* and is unable to oxidize water (Cardona et al., 2015). Purple indicates group 1, red genes belong to group 2, orange genes are members of group 3, and blue standard *psbA*s belong to group 4. Genetic proximity to *sqr* types I (green circle), II (maroon squares), and VI (blue stars) *sqr* genes are indicated. Fourteen *psbA* genes are five or fewer genes from *sqr*, and the largest gap is 19 genes. *Geitlerinema* sp. PCC 9228 has members of groups 2–4.

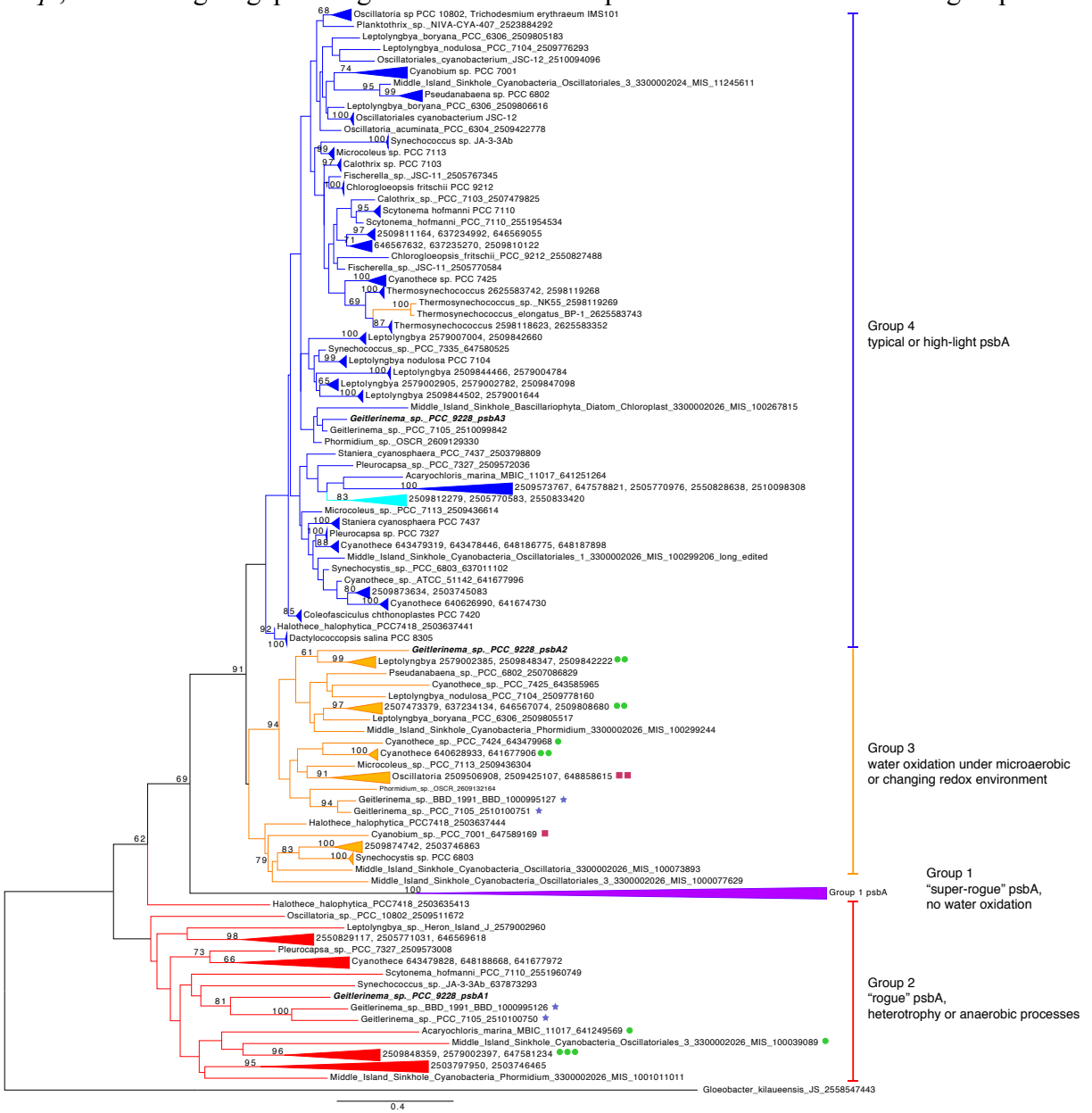
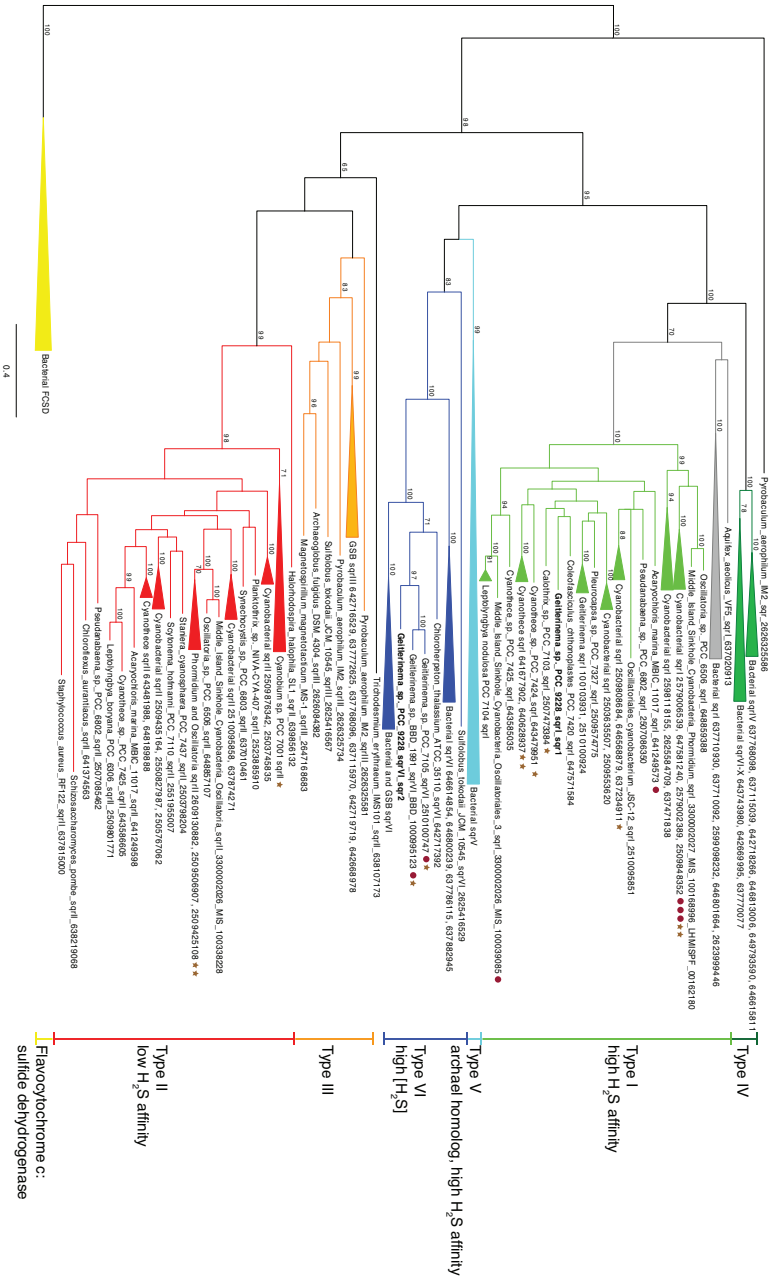


Figure 2.4. Phylogenetic tree of sulfide quinone reductase (*sqr*). Genes are colored by types I–VI modeled after (Marcia et al., 2010a). Bacterial FCSD (yellow) includes representatives of flavocytochrome c:sulfide dehydrogenase. Green sulfur bacteria have *sqr* belonging to types III (orange), V (light blue), IV (dark green), and VI (blue). Cyanobacterial *sqr* are limited to types I (light green), II (red), and VI. Proximity to *psbA* versions on the same contig (see Figure 3) are indicated with red circles (group 2) and brown stars (group 3). No *sqr* genes in this analysis were genetically proximal to groups 1 or 4 *psbA* genes. *Geitlerinema* sp. PCC 9228 has types VI and I *sqr*.



2.9 Appendix A

Supporting Information

Table SI 2.1. This and next 3 pages: IMG accession numbers of genome and genes analyzed from *Geitlerinema* sp. PCC 9228.

Description	IMG ID	gene
photosystem II P680 reaction center D1 protein-1	2663545011	psbA1
photosystem II P680 reaction center D1 protein-2	2663544218	psbA2
photosystem II P680 reaction center D1 protein-3	2663548115	psbA3
sulfide:quinone oxidoreductase 1	2663545736	sqr1
sulfide:quinone oxidoreductase 2	2663547046	sqr2
photosystem I P700 chlorophyll a apoprotein A1	2663545162	psaA
photosystem I P700 chlorophyll a apoprotein A2	2663545352	psaB large fragment 1089bp
photosystem I P700 chlorophyll a apoprotein A2	2663545589	psaB large fragment 1098bp
photosystem I P700 chlorophyll a apoprotein A2	2663545163	psaB fragment 483bp
photosystem I P700 chlorophyll a apoprotein A2	2663544697	psaB fragment 483bp
photosystem I P700 chlorophyll a apoprotein A2	2663546384	psaB fragment 84bp
photosystem I P700 chlorophyll a apoprotein A2	2663545213	psaB fragment 84bp
photosystem I subunit 7	2663546302	psaC
photosystem I subunit 2	2663546891	psaD
photosystem I subunit 4	2663544894	psaE
photosystem I subunit 3	2663544496	psaF
photosystem I subunit 8	2663544693	psaI
photosystem I subunit 10	2663546545	psaK
photosystem I subunit 11	2663544694	psaL
photosystem I subunit 12	2663545652	psaM
succinate dehydrogenase subunit A	2663545631	sdhA
succinate dehydrogenase subunit B	2663547796	sdhB
succinate dehydrogenase subunit C	2663544771	sdhC
NADH dehydrogenase subunit H	2663546838	ndhA
NADH dehydrogenase subunit N	2663546346	ndhB
NADH dehydrogenase subunit A	2663545237	ndhC
NADH dehydrogenase subunit M	2663545570	ndhD
NADH dehydrogenase subunit M	2663545034	ndhD
NADH dehydrogenase subunit M	2663545915	ndhD
NADH dehydrogenase subunit M	2663544356	ndhD
NADH dehydrogenase subunit K	2663546835	ndhE
NADH dehydrogenase subunit L	2663545571	ndhF
NAD(P)H-quinone oxidoreductase subunit 5	2663545035	ndhF
NAD(P)H-quinone oxidoreductase subunit 5	2663544357	ndhF
NADH dehydrogenase subunit J	2663546836	ndhG
NADH dehydrogenase subunit D	2663546817	ndhH
NAD(P)H-quinone oxidoreductase subunit I	2663546837	ndhI
NAD(P)H-quinone oxidoreductase subunit J	2663545235	ndhJ
NADH dehydrogenase subunit B	2663545236	ndhK
NAD(P)H-quinone oxidoreductase subunit L	2663545074	ndhL
NAD(P)H-quinone oxidoreductase subunit M	2663546659	ndhM
NAD(P)H-quinone oxidoreductase subunit N	2663547377	ndhN
apocytochrome f	2663547308	petA
cytochrome b6	2663545656	petB
cytochrome b6-f complex iron-sulfur subunit	2663544168	petC
cytochrome b6-f complex iron-sulfur subunit	2663544784	petC
cytochrome b6-f complex iron-sulfur subunit	2663547307	petC
cytochrome b6-f complex subunit 4	2663545657	petD
plastocyanin	2663548000	petE
ferredoxin	2663545357	petF

ferredoxin	2663545748	petF
ferredoxin	2663547777	petF
cytochrome b6-f complex subunit 5	2663546903	petG
ferredoxin--NADP+ reductase	2663546780	petH
cytochrome b6-f complex subunit 7	2663545911	petM
cytochrome b6-f complex subunit 8	2663545827	petN
cytochrome c6	2663546305	petJ
cytochrome c6	2663547966	petJ
photosystem II P680 reaction center D2 protein	N/A	psbD
photosystem II CP47 chlorophyll apoprotein	2663547980	psbB
photosystem II CP43 chlorophyll apoprotein	2663546355	psbC
photosystem II CP43 chlorophyll apoprotein	2663546353	psbC
photosystem II CP43 chlorophyll apoprotein	2663547553	psbC
photosystem II cytochrome b559 subunit alpha	2663545240	psbE
photosystem II cytochrome b559 subunit beta	2663545241	psbF
photosystem II PsbH protein	2663545025	psbH
photosystem II PsbI protein	2663547338	psbI
photosystem II PsbJ protein	2663545243	psbJ
photosystem II PsbK protein	2663545201	psbK
photosystem II PsbL protein	2663545242	psbL
photosystem II PsbM protein	2663547055	psbM
PsbN protein	2663545026	psbN
photosystem II oxygen-evolving enhancer protein 1	2663547292	psbO
photosystem II oxygen-evolving enhancer protein 2	2663545303	psbP
photosystem II PsbT protein	2663547981	psbT
photosystem II PsbU protein	2663547487	psbU
photosystem II cytochrome c550	2663546873	psbV
photosystem II cytochrome c550	2663546874	psbV
photosystem II PsbX protein	2663544266	psbX
photosystem II PsbY protein	2663544160	psbY
photosystem II PsbZ protein	2663547094	psbZ
photosystem II Psb27 protein	2663547854	psb27
photosystem II protein	2663544713	psb28
photosystem II Psb28-2 protein	2663547285	psb28-2
cytochrome c oxidase subunit 1	2663547835	coxA
cytochrome c oxidase subunit 2	2663547834	coxB
cytochrome c oxidase subunit 3	2663547836	coxC
carbon dioxide concentrating mechanism protein CcmK	2663545038	ccmK1
carbon dioxide concentrating mechanism protein CcmK	2663545037	ccmK2
carbon dioxide concentrating mechanism protein CcmL	2663547648	ccmL
carbon dioxide concentrating mechanism protein CcmM small-chain	2663545041	ccmM
carbon dioxide concentrating mechanism protein CcmM small-chain	2663545040	ccmM
carbon dioxide concentrating mechanism protein CcmM	2663545039	ccmM
carbon dioxide concentrating mechanism protein CcmO	2663545231	ccmO
carbon dioxide concentrating mechanism protein CcmN	2663545042	ccmN
ribulose 1,5-bisphosphate carboxylase large subunit	2663545044	rbcL
rubisco chaperonin	2663545045	rbcX
ribulose 1,5-bisphosphate carboxylase small subunit	2663545046	rbcS
nickel superoxide dismutase	2663547183	sodN
superoxide dismutase, Fe-Mn family	2663545076	sodC
2-oxoglutarate dehydrogenase	2663544142	
succinyl-CoA synthetase alpha subunit	2663546151	sucD
succinyl-CoA synthetase (ADP-forming) beta subunit	2663545720	sucC
acetolactate synthase, large subunit	2663546224	ilvB, ilvG, ilvI
acetolactate synthase, large subunit	2663545283	ilvB, ilvG, ilvI
acetolactate synthase, small subunit	2663545134	ilvH, ilvN
succinate-semialdehyde dehydrogenase / glutarate-semialdehyde dehydrogenase	2663546225	gabD
squalene-hopene cyclase	2663547443	shc

anaerobic magnesium-protoporphyrin IX monomethyl ester cyclase	2663545681	bchE
anaerobic magnesium-protoporphyrin IX monomethyl ester cyclase	2663545682	bchE
Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase	2663547315	chlE
Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase	2663545872	chlE
cyanate lyase	2663544394	cynS
nitrate/nitrite transporter	2663544392	narK
ferredoxin-nitrite reductase	2663544395	nirA
formate/nitrite transporter focA/nirC	2663544393	focA/nirC
ferredoxin-nitrate reductase	2663544391	narB
ammonium transporter	2663545067	amt
urease subunit gamma	2663547720	ureA
urease subunit beta	2663547721	ureB
urease	2663544172	ureC
urease accessory protein	2663544955	ureD
urease accessory protein	2663546629	ureF
urease accessory protein	2663546627	ureG
urea-binding protein	2663546626	urtA
urea transport system substrate-binding protein	2663546591	urtA
urea ABC transporter membrane protein	2663546625	urtB
urea ABC transporter membrane protein	2663546624	urtC
urea ABC transporter ATP-binding protein	2663546623	urtD
urea ABC transporter ATP-binding protein	2663546622	urtE
homocitrate synthase NifV	2663545460	nifV
nitrogen fixation protein NifB	2663545461	nifX
cysteine desulfurase	2663545462	nifS
Fe-S cluster assembly protein NifU	2663545463	nifU
iron-sulfur cluster assembly protein	2663545464	iscA
Mo-nitrogenase iron protein subunit NifH	2663545465	nifH
nitrogen regulatory protein P-II family	2663545466	nifI1
nitrogen regulatory protein P-II family	2663545467	nifI2
nitrogenase molybdenum-iron protein alpha chain	2663545468	nifD
Mo-nitrogenase MoFe protein subunit NifK	2663545469	nifK
(2Fe-2S) ferredoxin	2663545470	fdxN
nitrogenase molybdenum-cofactor synthesis protein NifE	2663545471	nifE
nitrogenase molybdenum-iron protein NifN	2663545472	nifN
nitrogen fixation protein NifB	2663545473	nifB
nitrogen regulatory protein P-II family	2663546645	glnB
hydrogenase nickel incorporation protein HypA/HybF	2663545443	hypA
hydrogenase nickel incorporation protein HypB	2663545442	hypB
hydrogenase expression/formation protein HypD	2663545445	hypD
hydrogenase expression/formation protein HypE	2663545444	hypE
Hydrogenase maturation protein, carbamoyltransferase HypF	2663547362	hypF
NAD(P)-dependent nickel-iron dehydrogenase diaphorase component subunit HoxE	2663544431	hoxE
NAD(P)-dependent nickel-iron dehydrogenase flavin-containing subunit	2663544432	hoxF
NAD(P)-dependent nickel-iron dehydrogenase diaphorase component subunit HoxU	2663544433	hoxU
NAD(P)-dependent nickel-iron dehydrogenase subunit HoxY	2663544434	hoxY
NAD(P)-dependent nickel-iron dehydrogenase catalytic subunit	2663544435	hoxH
hydrogenase maturation protease	2663544439	hoxW
LexA-binding, inner membrane-associated putative hydrolase	2663545182	lexA
global nitrogen regulator NtcA, cyanobacterial	2663545019	ntcA
GST	2663547047	GST
multidrug efflux pump	2663547048	multidrug efflux pump
padR	2663547050	padR
arsenite efflux ATP-binding protein ArsA	2663545455	arsA
transcriptional regulator, ArsR family	2663545310	arsR
arsenite efflux ATP-binding protein ArsA	2663546112	arsA
arsenical pump membrane protein	2663545756	arsB

transcriptional regulator, ArsR family	2663547029	arsR
arsenite transporter, ACR3 family	2663547028	arsB
arsenate reductase	2663547027	arsC
chromate transporter	2663547951	chrA
Methyltransferase domain-containing protein	2663546259	arsM
Uncharacterized conserved protein DUF302	2663544170	

Table SI 2.2. This and next page: IMG accession numbers of cyanobacterial genomes and genomic bins used in phylogenetic analysis, and associated *nifHDK*, *psbA*, and *sqr* genes. Phylogenetic clustering of *psbA* (G1-4) from (Cardona et al. 2015) and this study, and *sqr* (Type I-VI) from (Marcia, Ermler, et al. 2010; Gregersen et al. 2011) are indicated.

	IMG Accession Numbers									NCBI Accession numbers				
Organism	**Genome**	***sqr***	**type (I-VI)**	***psbA***	**type (1-4)**	**Near a *sqr*?**	***nifH***	***nifD***	***nifK***	***sqr***	***psbA***	***nifH***	***nifD***	***nifK***
Acaryochloris marina MBIC 11017	641228474	641249573	I	641249569	2	641249573				YP_001514822	WP_012161111.1			
		641249598	II	641251264	4					YP_001514847	WP_010482631.1			
Anabaena variabilis ATCC 29413	646564504	646568879	I	646569618	2		646569746	646569747	646569748	YP_323880	WP_011320804.1	WP_011320927.1	WP_011320928.1	WP_011320929.1
				646567074	3		646569419	646569420	646569433		WP_011318398.1	YP_324416.1	YP_324417.1	YP_324430.1
				646569055	4						WP_011318412.1			
				646567632	4						WP_011318925.1			
Calothrix sp. PCC 7103	2507262048	2507473384	I	2507473379	3	2507473384				WP_019496967	WP_019496962.1			
				2507479825	4						WP_019492713.1			
				2507482675	4						WP_019495472.1			
				2507478832	4						WP_019491774.1			
				2507482017	4						WP_019494835.1			
				2507483403	4						WP_019488209.1			
Chlorogloeopsis fritschii PCC 9212	2548877023	2550827987	II	2550828637	1		2550827568	2550827567	2550827556	WP_016872789	WP_016873418.1	WP_016872368.1	WP_016872367.1	WP_016872357.1
				2550829117	2						WP_016873890.1			
				2550833420	4						WP_016878157.1			
				2550827488	4						WP_016872293.1			
				2550828669	4						WP_016873445.1			
				2550834064	4						WP_016878921.1			
				2550828638	4						WP_016873419.1			
Coleofasciculus chthonoplastes PCC 7420	647533184	647571584	I	647568289	4		647565360	647565357	647565356	ZP_05030405	ZP_05031153.1	WP_006099029.1	WP_006098789.1	WP_044205237.1
				647572437	4						ZP_05027093.1			
Cyanobacterium aponium PCC 10605	2503707009	2503745835	II	2503746465	2					WP_015219410	WP_015220005.1			
				2503746863	3						WP_015220386.1			
				2503745083	4						WP_015218695.1			
Cyanobium sp. PCC 7001	647533126	647589167	II	647589169	3	647589167				ZP_05045491	WP_006910002.1			
		647589188	II	647588948	4					ZP_05045506	WP_006911636.1			
				647589899	4						WP_006909897.1			
Cyanothece sp. ATCC 51142	641522622	641677902	I	641677972	2		641675026	641675027	641675028	YP_001804821	WP_009545353.1	WP_009546639.1	WP_009546640.1	WP_009546641.1
				641677906	3	641677902					WP_009545419.1			
				641677996	4						WP_008274631.1			
				641674730	4						WP_009546761.1			
Cyanothece sp. CCY0110	640612201	640628937	I	640628933	3	640628937	640625205	640625206	640625207	ZP_01731467	WP_008277783.1	WP_008274042.1	WP_008274043.1	WP_008274045.1
				640626990	4						EAZ91041.1			
Cyanothece sp. PCC 7424	643348533	643479951	I	643479828	2		643480377	643480376	643480375	YP_002379007	WP_012598947.1	WP_015954153.1	WP_015954152.1	WP_015954151.1
		643481988	II	643479968	3	643479951				YP_002376992	WP_012599084.1			
				643479319	4						WP_012598449.1			
				643478446	4						WP_012597589.1			
Cyanothece sp. PCC 7425	643348534	643585035	I	643585965	3		643585358	643585359	643585360	YP_002482763	WP_012628391.1	WP_012627797.1	WP_012627798.1	WP_012627799.1
		643586605	II	643584498	4					YP_002484311	WP_012626959.1			
				643584070	4						WP_012626542.1			
Cyanothece sp. PCC 7822	648028021	648189888	II	648188668	2		648188591	648188592	648188593	YP_003890172	WP_013323757.1	WP_013323681.1	WP_013323682.1	WP_013323683.1
				648186775	4						WP_013321097.1			
				648187898	4						WP_013323018.1			
Dactylococcopsis salina PCC 8305	2509276056	2509553620	I	2509554132	4					WP_015228940	WP_015229407.1			
				2509552870	4						WP_015228284.1			
Fischerella sp. JSC-11	2505679024	2505767062	II	2505770977	1		2505767548	2505767547	2505767546	WP_009459992	WP_009453704.1	WP_009459216.1	WP_009459215.1	WP_009459214.1
				2505771031	2						WP_009453755.1			
				2505770584	4						WP_009455234.1			
				2505767345	4						WP_026086577.1			
				2505770976	4						WP_009453703.1			
				2505770583	4						WP_009455233.1			
Geitlerinema sp. BBD 1991	3300001677	BBD_100103931	I	BBD_1000995126	2	BBD_1000995123	BBD_100100267	BBD_100100268	BBD_100100269	N/A	CAA68778.1	N/A	N/A	N/A
		BBD_1000995123	VI	BBD_1000995127	3	BBD_1000995123	BBD_1001010119	BBD_1001010120	BBD_1001010121		WP_010874309.1	N/A	N/A	N/A
Geitlerinema sp. PCC 7105	2510065011	2510100924	I	2510100750	2	2510100747				WP_035492693	WP_010871214.1			
		2510100747	VI	2510100751	3	2510100747				WP_017659174.1	WP_017659179.1			
				2510099842	4						WP_017662363.1			
***Geitlerinema* sp. PCC 9228**	2660238729	2663545736	I	2663545011	2		2663545465	2663545468	2663545469	N/A	N/A	N/A	N/A	N/A
		2663547046	VI	2663544218	3					N/A	N/A			
				2663548115	4						N/A			
Geminocystis herdmanii PCC 6308	2509601046	2509876342	II	2509874742	3					WP_017295350	WP_017293773.1			
				2509873634	4						WP_017292699.1			
Gloeobacter kilaueensis JS	2558309063	N/A		2558547443	0						AGY58976.1			
Halothece halophytica PCC7418	2503538028	2503635507	I	2503635413	2		2503636510	2503636511	2503636512	WP_015224891	WP_015224801.1	WP_015225829.1	WP_015225830.1	WP_041596559.1
				2503637444	3						WP_015226718.1			
				2503637441	4						WP_015226715.1			
Leptolyngbya boryana PCC 6306	2509601031	2509801771	II	2509805517	3		2509804224	2509804225	2509804226	WP_017286582	WP_017290327.1	WP_017289029.1	WP_017289030.1	WP_017289031.1
				2509805183	4						WP_017289997.1			
				2509806616	4						WP_017288020.1			
Leptolyngbya nodulosa PCC 7104	2509601026	2509778855	I	2509778160	3		2509775993	2509775992	2509775991	WP_017301594	WP_017300941.1	WP_017298849.1	WP_017298848.1	WP_040696575.1
		2509773782	I	2509776293	4					WP_017296689	WP_017299141.1			
				2509775115	4						WP_017297994.1			
				2509773891	4						WP_017296794.1			
				2509777846	4						WP_017300650.1			
Leptolyngbya sp. Heron Island J	2576861623	2579006539	I	2579002960	2		2579006069	2579006068	2579006067	WP_023074994	WP_023071503.1	WP_023074533.1	WP_023074532.1	WP_036053974.1
		2579002389	I	2579002397	2	2579002389				WP_023070947	WP_023070955.1			
				2579002385	3	2579002389					WP_023070943.1			
				2579004784	4						WP_023073286.1			
				2579007004	4						WP_031293852.1			
				2579002782	4						WP_023071336.1			
				2579002905	4						WP_023071448.1			
				2579001644	4						WP_031291203.1			
Leptolyngbya sp. PCC 7375	2509601039	2509848352	I	2509848359	2	2509848352	2509842133	2509842134	2509842135	WP_006520144	WP_006520151.1	WP_006514484.1	WP_006514485.1	WP_036019883.1
				2509842222	3						WP_006514564.1			
				2509848347	3	2509848352					N/A			
				2509844466	4						WP_006516643.1			
				2509842660	4						WP_006514967.1			
				2509847098	4						WP_006516801.1			
				2509844502	4						WP_006516678.1			

	IMG Accession Numbers									NCBI Accession numbers				
Organism	**Genome**	***sqr***	**type (I-VI)**	***psbA***	**type (1-4)**	**Near a *sqr*?**	***nifH***	***nifD***	***nifK***	***sqr***	***psbA***	***nifH***	***nifD***	***nifK***
Acaryochloris marina MBIC 11017	641228474	641249573	I	641249569	2	641249573				YP_001514822	WP_012161111.1			
		641249598	II	641251264	4					YP_001514847	WP_010482631.1			
Anabaena variabilis ATCC 29413	646564504	646568879	I	646569618	2		646569746	646569747	646569748	YP_323880	WP_011320804.1	WP_011320927.1	WP_011320928.1	WP_011320929.1
				646567074	3		646569419	646569420	646569433		WP_011318398.1	YP_324416.1	YP_324417.1	YP_324430.1
				646569055	4						WP_011318412.1			
				646567632	4						WP_011318925.1			
Calothrix sp. PCC 7103	2507262048	2507473384	I	2507473379	3	2507473384				WP_019496967	WP_019496962.1			
				2507479825	4						WP_019492713.1			
				2507482675	4						WP_019495472.1			
				2507478832	4						WP_019491774.1			
				2507482017	4						WP_019494835.1			
				2507483403	4						WP_019488209.1			
Chlorogloeopsis fritschii PCC 9212	2548877023	2550827987	II	2550828637	1		2550827568	2550827567	2550827556	WP_016872789	WP_016873418.1	WP_016872368.1	WP_016872367.1	WP_016872357.1
				2550829117	2						WP_016873890.1			
				2550833420	4						WP_016878157.1			
				2550827488	4						WP_016872293.1			
				2550828669	4						WP_016873445.1			
				2550834064	4						WP_016878921.1			
				2550828638	4						WP_016873419.1			
Coleofasciculus chthonoplastes PCC 7420	647533184	647571584	I	647568289	4		647565360	647565357	647565356	ZP_05030405	ZP_05031153.1	WP_006099029.1	WP_006098789.1	WP_044205237.1
				647572437	4						ZP_05027093.1			
Cyanobacterium aponium PCC 10605	2503707009	2503745835	II	2503746465	2					WP_015219410	WP_015220005.1			
				2503746863	3						WP_015220386.1			
				2503745083	4						WP_015218695.1			
Cyanobium sp. PCC 7001	647533126	647589167	II	647589169	3	647589167				ZP_05045491	WP_006910002.1			
		647589188	II	647588948	4					ZP_05045506	WP_006911636.1			
				647589899	4						WP_006909897.1			
Cyanothece sp. ATCC 51142	641522622	641677902	I	641677972	2		641675026	641675027	641675028	YP_001804821	WP_009545353.1	WP_009546639.1	WP_009546640.1	WP_009546641.1
				641677906	3	641677902					WP_009545419.1			
				641677996	4						WP_008274631.1			
				641674730	4						WP_009546761.1			
Cyanothece sp. CCY0110	640612201	640628937	I	640628933	3	640628937	640625205	640625206	640625207	ZP_01731467	WP_008277783.1 EAZ91041.1	WP_008274042.1	WP_008274043.1	WP_008274045.1
				640626990	4									
Cyanothece sp. PCC 7424	643348533	643479951	I	643479828	2		643480377	643480376	643480375	YP_002379007	WP_012598947.1	WP_015954153.1	WP_015954152.1	WP_015954151.1
		643481988	II	643479968	3	643479951				YP_002376992	WP_012599084.1			
				643479319	4						WP_012598449.1			
				643478446	4						WP_012597589.1			
Cyanothece sp. PCC 7425	643348534	643585035	I	643585965	3		643585358	643585359	643585360	YP_002482763	WP_012628391.1	WP_012627797.1	WP_012627798.1	WP_012627799.1
		643586605	II	643584498	4					YP_002484311	WP_012626959.1			
				643584070	4						WP_012626542.1			
Cyanothece sp. PCC 7822	648028021	648189888	II	648188668	2		648188591	648188592	648188593	YP_003890172	WP_013323757.1	WP_013323681.1	WP_013323682.1	WP_013323683.1
				648186775	4						WP_013321097.1			
				648187898	4						WP_013323018.1			
Dactylococcopsis salina PCC 8305	2509276056	2509553620	I	2509554132	4					WP_015228940	WP_015229407.1			
				2509552870	4						WP_015228284.1			
Fischerella sp. JSC-11	2505679024	2505767062	II	2505770977	1		2505767548	2505767547	2505767546	WP_009459992	WP_009453704.1	WP_009459216.1	WP_009459215.1	WP_009459214.1
				2505771031	2						WP_009453755.1			
				2505770584	4						WP_009455234.1			
				2505767345	4						WP_026086577.1			
				2505770976	4						WP_009453703.1			
				2505770583	4						WP_009455233.1			
Geitlerinema sp. BBD 1991	3300001677	BBD_100103931	I	BBD_1000995126	2	BBD_1000995123	BBD_100100267	BBD_100100268	BBD_100100269	N/A	CAA68778.1	N/A	N/A	N/A
		BBD_1000995123	VI	BBD_1000995127	3	BBD_1000995123	BBD_1001010119	BBD_1001010120	BBD_1001010121		WP_010874309.1	N/A	N/A	N/A
Geitlerinema sp. PCC 7105	2510065011	2510100924	I	2510100750	2	2510100747				WP_035492693	WP_010871214.1			
		2510100747	VI	2510100751	3	2510100747				WP_017659174.1	WP_017659179.1			
				2510099842	4						WP_017662363.1			
***Geitlerinema* sp. PCC 9228**	2660238729	2663545736	I	2663545011	2		2663545465	2663545468	2663545469	N/A	N/A	N/A	N/A	N/A
		2663547046	VI	2663544218	3					N/A	N/A			
				2663548115	4						N/A			
Geminocystis herdmanii PCC 6308	2509601046	2509876342	II	2509874742	3					WP_017295350	WP_017293773.1			
				2509873634	4						WP_017292699.1			
Gloeobacter kilaueensis JS	2558309063	N/A		2558547443	0						AGY58976.1			
Halothece halophytica PCC7418	2503538028	2503635507	I	2503635413	2		2503636510	2503636511	2503636512	WP_015224891	WP_015224801.1	WP_015225829.1	WP_015225830.1	WP_041596559.1
				2503637444	3						WP_015226718.1			
				2503637441	4						WP_015226715.1			
Leptolyngbya boryana PCC 6306	2509601031	2509801771	II	2509805517	3		2509804224	2509804225	2509804226	WP_017286582	WP_017290327.1	WP_017289029.1	WP_017289030.1	WP_017289031.1
				2509805183	4						WP_017289997.1			
				2509806616	4						WP_017288020.1			
Leptolyngbya nodulosa PCC 7104	2509601026	2509778855	I	2509778160	3		2509775993	2509775992	2509775991	WP_017301594	WP_017300941.1	WP_017298849.1	WP_017298848.1	WP_040696575.1
		2509773782	I	2509776293	4					WP_017296689	WP_017299141.1			
				2509775115	4						WP_017297994.1			
				2509773891	4						WP_017296794.1			
				2509777846	4						WP_017300650.1			
Leptolyngbya sp. Heron Island J	2576861623	2579006539	I	2579002960	2		2579006069	2579006068	2579006067	WP_023074994	WP_023071503.1	WP_023074533.1	WP_023074532.1	WP_036053974.1
		2579002389	I	2579002397	2	2579002389				WP_023070947	WP_023070955.1			
				2579002385	3	2579002389					WP_023070943.1			
				2579004784	4						WP_023073286.1			
				2579007004	4						WP_031293852.1			
				2579002782	4						WP_023071336.1			
				2579002905	4						WP_023071448.1			
				2579001644	4						WP_031291203.1			
Leptolyngbya sp. PCC 7375	2509601039	2509848352	I	2509848359	2	2509848352	2509842133	2509842134	2509842135	WP_006520144	WP_006520151.1	WP_006514484.1	WP_006514485.1	WP_036019883.1
				2509842222	3						WP_006514564.1			
				2509848347	3	2509848352					N/A			
				2509844466	4						WP_006516643.1			
				2509842660	4						WP_006514967.1			
				2509847098	4						WP_006516801.1			
				2509844502	4						WP_006516678.1			

Table SI 2.3. This and next page: IMG accession numbers and/or NCBI accession numbers for bacterial and archaeal *sqr* genes used in phylogenetic analysis. Phylogenetic Type I-VI from (Marcia, Ermler, et al. 2010; Gregersen et al. 2011) are indicated.

Taxonomy	IMG	Type	NCBI
Acidithiobacillus caldus ATCC 51756	645601439	V	ZP_05293376
Allochromatium vinosum DSM 180	646614749	FCSD	ADC62031
	646615811	IV	YP_003444098
	646614854	VI	YP_003443166
Aquifex aeolicus VF5	2626082283	FCSD	WP_010880096
	637020913	I	NP_214500
	637019911	VI	NP_213539
Archaeoglobus fulgidus DSM 4304	2626084382	III	NP_069393
Arcobacter nitrofigilis DSM 7299	646813006	IV	YP_003654976
Beggiatoa sp. PS	2623997857	FCSD	N/A
	2623999446	I	N/A
Chlorobaculum parvum NCIB 8327	642719719	III	YP_001998682
	642719525	V	YP_001998488
	642719655	VI	YP_001998618
Chlorobium chlorochromatii CaD3	637772625	III	YP_378420
Chlorobium ferrooxidans DSM 13031	639205487	VI	ZP_01385814
Chlorobium limicola DSM 245	642667635	FCSD	YP_001942098
	642668978	III	YP_001943408
	642669995	IV-X	YP_001944405
Chlorobium tepidum TLS	637115970	III	NP_661917
	637115039	IV	NP_661023
	2625429993	V	WP_010932556
	637116031	VI	NP_661978
Chloroflexus aurantiacus	641374563	II	YP_001637460
Chloroherpeton thalassium ATCC 35110	642716529	III	YP_001995567
	642718266	IV	YP_001997268
	642717392	VI	YP_001996408
Halorhodospira halophila SL1	639856132	II	YP_001003231
Magnetospirillum magnetotacticum MS-1	2647168683	III	WP_009869729
Pelodictyon luteolum DSM 273	637768096	III	YP_374002
	637768098	IV	YP_374004
	637770077	IV-X	YP_375931
	637769162	VI	YP_375032
Persephonella marina EX-H1	643743980	IV-X	YP_002731229
	643743019	VI	YP_002730289
Pyrobaculum aerophilum IM2	2626325586	Unk	WP_011008794
	2626325734	III	WP_011008939
	2626325581	III	WP_011008789
Rhodobacter capsulatus DSM155	2599098232	I	CAA66112
Rhodopseudomonas palustris HaA2	637882945	VI	YP_484673
Rhodopseudomonas palustris TIE-1	642711469	FCSD	YP_001990584
Schizosaccharomyces pombe	638219068	II	NP_596067
Staphylococcus aureus RF122	637815000	II	YP_415550
Sulfolobus tokodaii JCM 10545	2625416567	III	WP_010978596
	2625414418	V	WP_010980568
	2625416529	VI	WP_010978633
Sulfuricurvum kujiense DSM 16994	649793590	IV	YP_004060860
	649793598	VI	YP_004060868
Sulfurimonas autotrophica DSM 16294	648191284	V	YP_003891564
Thermoplasma volcanium	2625869763	V	BAB60371
Thiobacillus denitrificans ATCC 25259	637710930	I	YP_315983
	637710092	I	YP_315165
Thiomicrospira crunogena XCL-2	637786115	VI	YP_391650
Thiomonas intermedia K12	646801664	I	YP_003643765

	646802624	V	YP_003644708
	646800883	V	YP_003642999
	646800239	VI	YP_003642362

Figure SI 2.1. Emergent self-organized map (ESOM) of *Geitlerinema* sp. PCC 9228 genomic contigs, clustered on the basis of tetranucleotide frequency.

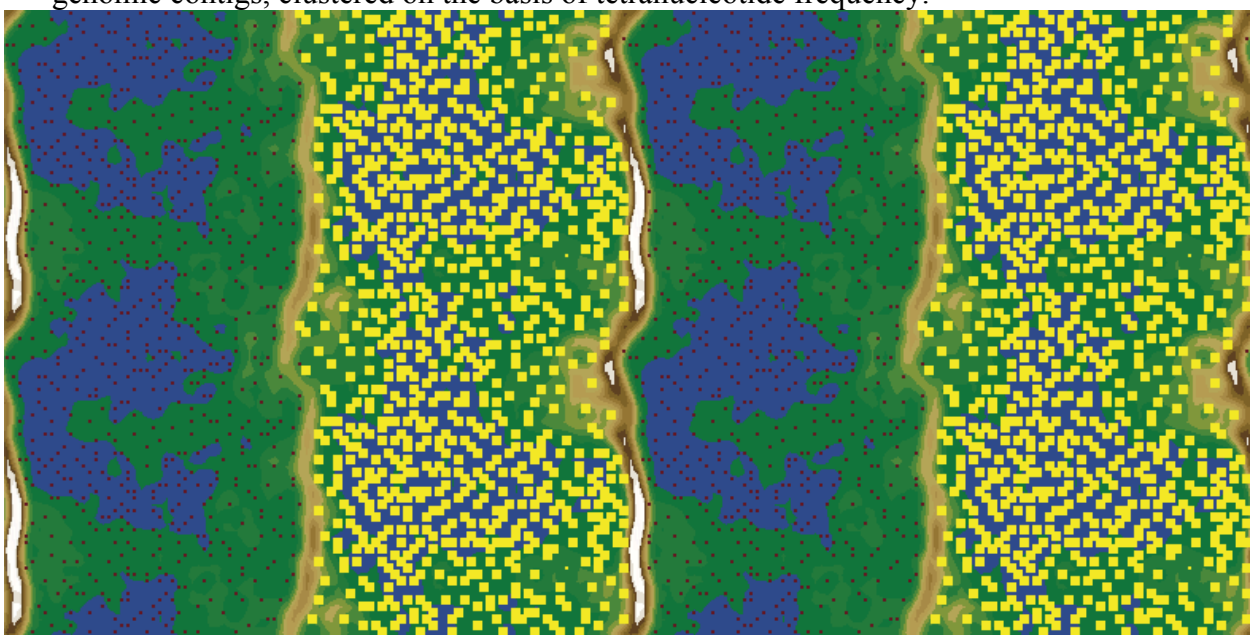


Figure SI 2.2. Phylogenetic tree of *nifHDK*. The clade with *Geitlerinema* sp. PCC 9228 includes three other cyanobacterial genomes that also have *nif*I_1I_2. The outgroup is the homologous *chlLNB* gene suite from *Halothece halophytica*.

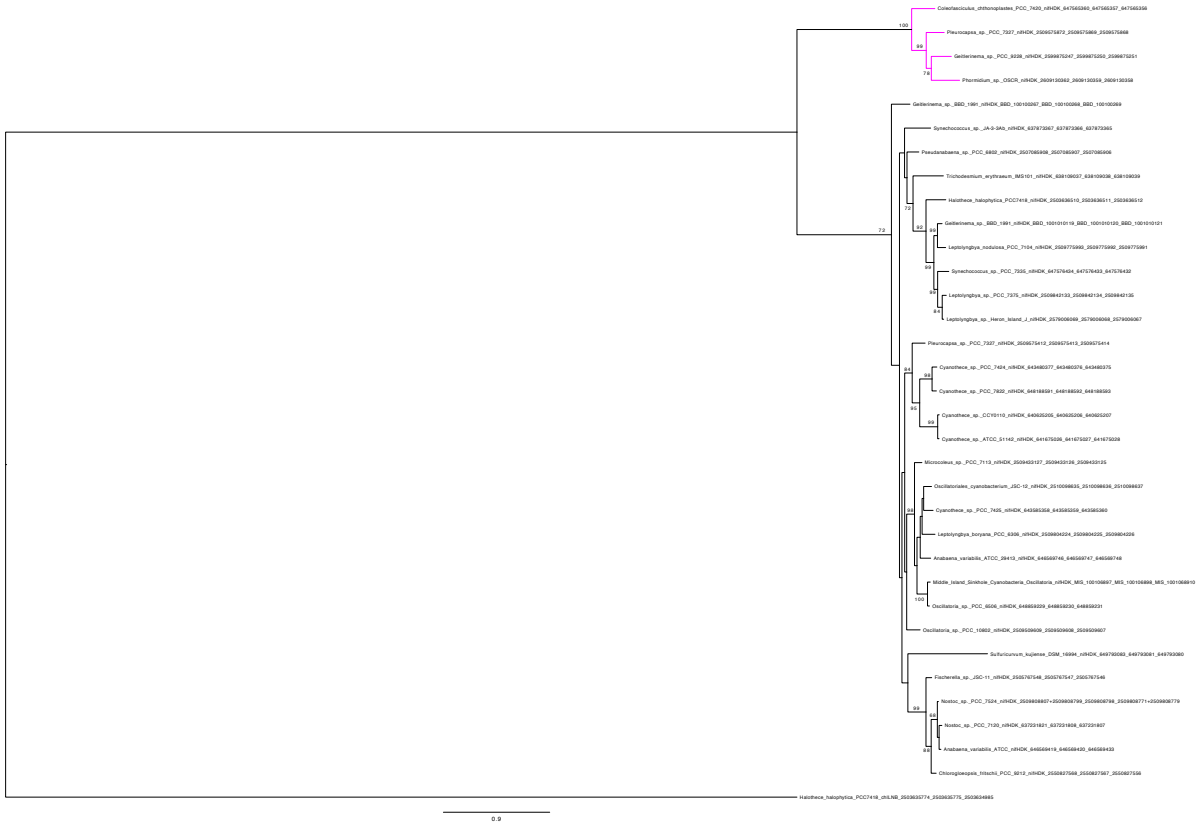


Figure SI 2.3. Full phylogenetic tree of *psbA* (with non-collapsed clades). Genes are colored by groups 1–4 modeled after (Cardona et al., 2015). The outgroup is a *Gloeobacter kilaueensis JS psbA* that is most like *psbD* and is unable to oxidize water. Purple indicates group 1, red genes belong to group 2, orange genes are members of group 3, and blue standard psbAs belong to group 4. Genetic proximity to types I (green circle), II (maroon squares), and VI (blue stars) *sqr* genes are indicated. Fourteen *psbA* genes are five or fewer genes from sqr, and the largest gap is 19 genes. *Geitlerinema* sp. PCC 9228 has members of groups 2–4.

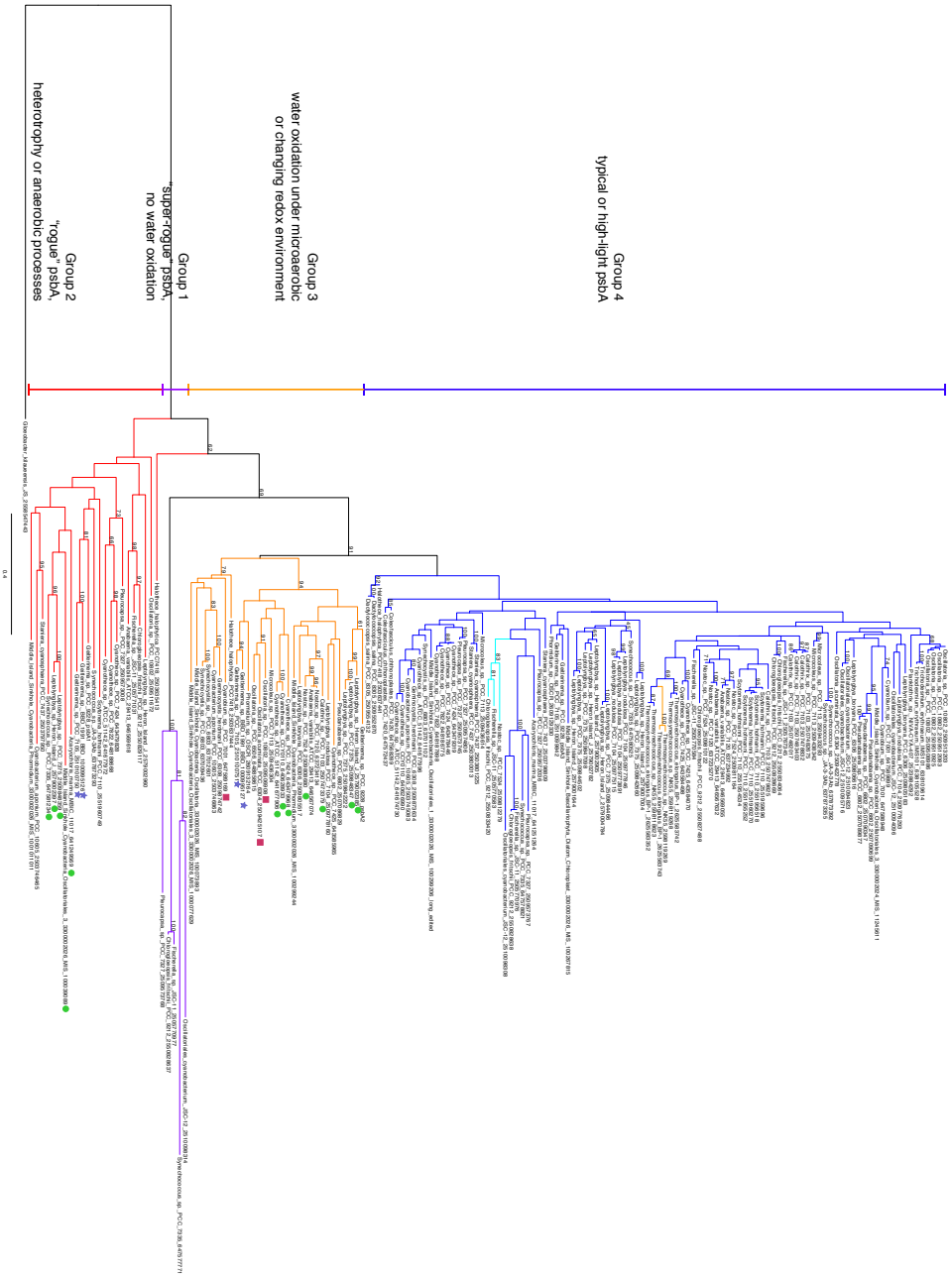


Figure SI 2.4. Schematic of *sqr* operons in *Geitlerinema* sp. PCC 9228. (A) Canonical *sqr* (type I) with regulatory gene *arsR* collocated downstream. (B) Novel *sqr* (type VI) with *arsR*, arsenite transport gene *arsB*, and arsenate reductase *arsC* all downstream, and glutathione *S*-transferase and a multidrug efflux pump are upstream.

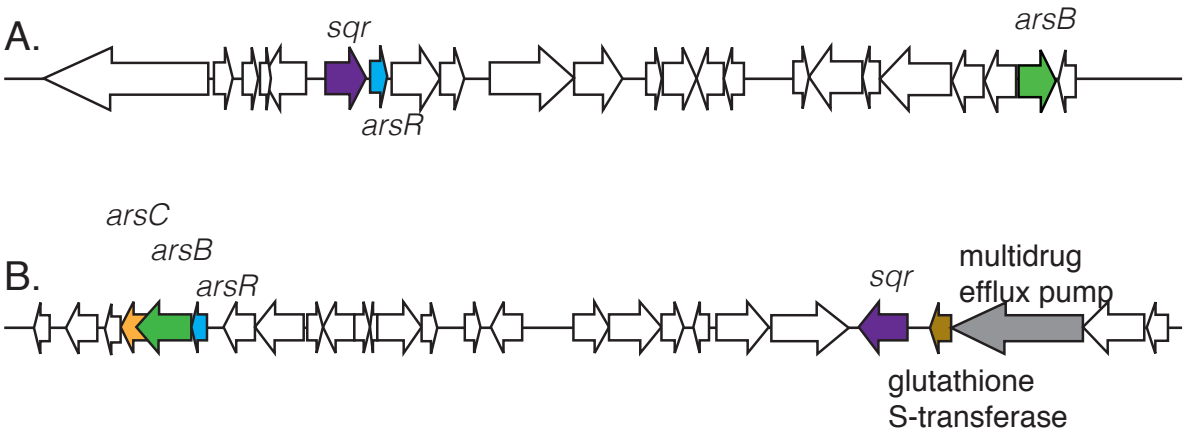
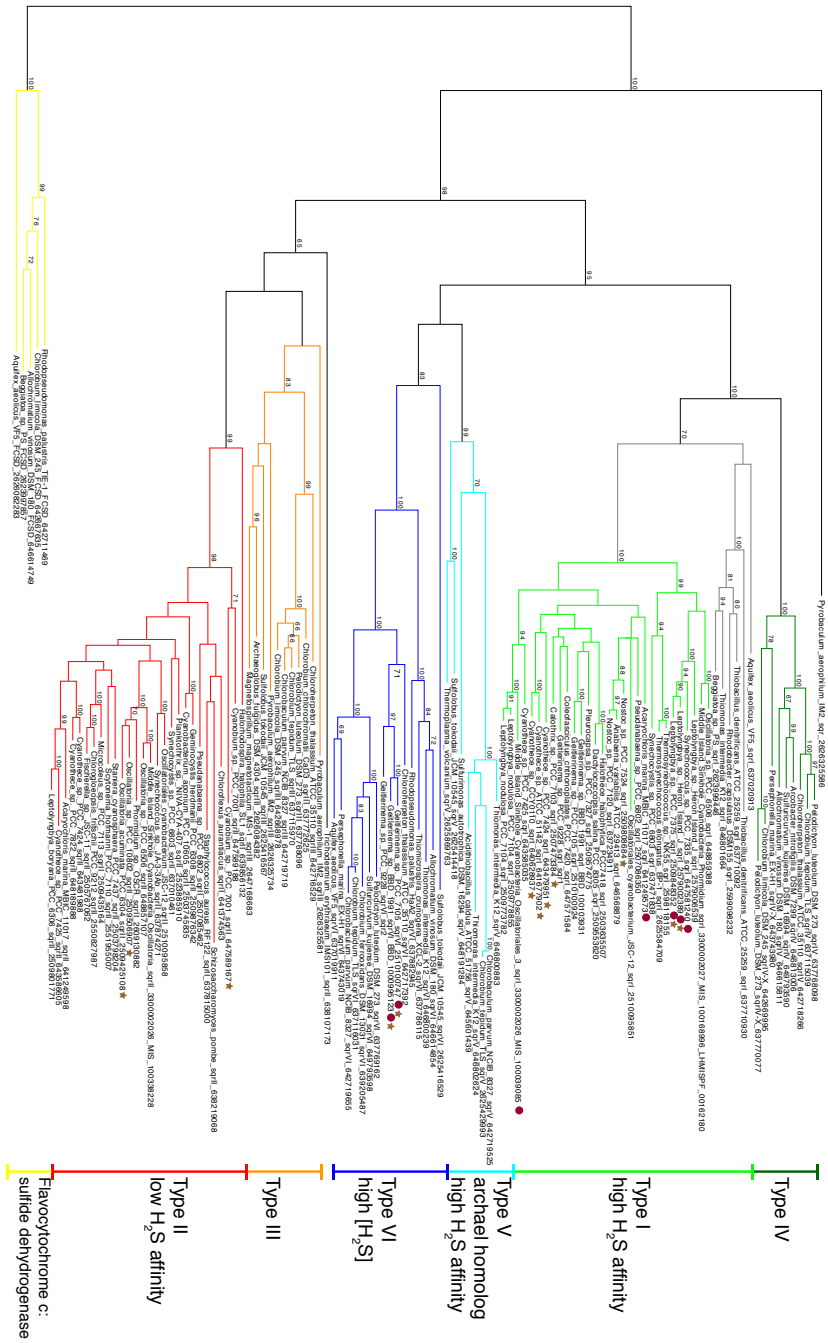


Figure SI 2.5. Full phylogenetic tree of sulfide quinone reductase (*sqr*). Genes are colored by types I–VI modeled after (Marcia et al., 2010a). Bacterial FCSD (yellow) includes representatives of flavocytochrome c:sulfide dehydrogenase. Green sulfur bacteria have *sqr* belonging to types III (orange), V (light blue), IV (dark green), and VI (blue). Cyanobacterial *sqr* are limited to types I (light green), II (red), and VI. Proximity to *psbA* versions on the same contig (see Figure 3; Supplementary Figure S3) are indicated with red circles (group 2) and brown stars (group 3). No *sqr* genes in this analysis were genetically proximal to groups 1 or 4 *psbA* genes. *Geitlerinema* sp. PCC 9228 has types VI and I *sqr*.



Chapter III: Seasonal Shifts in Community Composition and Proteome in a Sulfur-Cycling Cyanobacterial Mat

3.1 Abstract

In phototrophic microbial mats, cyanobacteria provide nutrients, fixed carbon, and oxygen through photosynthesis. Due to the reliance of cyanobacteria on light for photosynthesis, seasonality and changing irradiance alter their metabolisms, which can impact the population structure and metabolisms of other microbes in the mat. In this study, we explored temporal shifts in irradiance, water chemistry, community structure and function of a microbial mat in Middle Island Sinkhole (MIS), a low-oxygen, sulfate-rich, submerged environment. Mixing between low-oxygen groundwater and well-oxygenated lake water controls MIS water chemistry, which along with light varies with season. Through both 16S rRNA gene sequencing and quantitative proteomics profiles, we observed a relationship between the changing environment and the activity and abundance of cyanobacterial (*Phormidium*, *Planktothrix*) populations. In summer and across the seasons, the differential abundance of wavelength-specific phycobiliproteins and key metabolic proteins suggests a physiological response of the cyanobacterial community to different light levels. Sulfide-oxidizing bacteria, namely *Beggiatoa*, were abundant and metabolically active in the microbial mat, and likely interdependent on sulfur reducers and the cyanobacteria. The abundance and identity of dominant sulfate-reducing bacteria shift throughout the year, indicating changing environmental conditions and niches for sulfate-reducing bacteria. Our study reveals that in addition to altering the light conditions for

cyanobacterial photosynthesis, seasonal variation extends as well to the sulfur cycling bacteria in this microbial mat. This temporal perspective aids our understanding of diversity and metabolic functions through different seasons in this and other temperate microbial mat systems.

3.2 Introduction

Through the innovation of oxygenic photosynthesis (OP), cyanobacteria shaped the geosphere and biosphere by driving the rise of oxygen ~2.4 Ga (Lyons et al., 2014). In modern systems, cyanobacteria are a core constituent of photosynthetic microbial mats by providing organic matter, fixed nitrogen, and O_2 for other microorganisms and their metabolisms (Dick et al., 2018). These cyanobacterial products stimulate the metabolism of other microbes such as sulfate-reducing bacteria, sulfide-oxidizing bacteria, and methanogens, which contribute to geochemical fluxes such as methane and sulfide (Bolhuis et al., 2014). Alternatively, anoxygenic photosynthetic (AP) cyanobacteria that use H_2S can sustain microbial mats in sulfidic conditions that are toxic to most cyanobacteria (Biddanda et al., 2015; de Beer et al., 2017; Klatt, Meyer, et al., 2016). Though AP is an infrequently observed metabolism in cyanobacteria, nearly all AP cyanobacteria are mat formers (Dick et al., 2018). With their diverse metabolic functions and ecological strategies, cyanobacteria are geared to establish microbial mats in both ubiquitous and extreme environments not unlike those of early Earth.

Oxygenic cyanobacteria can inhabit different niches and present as different ecological species, or ecotypes, based on adaptations to light levels and wavelengths of light (Ward et al., 2006). For example, *Planktothrix rubescens*, commonly found in deeper stratified lakes and equipped with pigments that capture the blue-green light that is dominant at depth, performs optimally in low light intensities (Oberhaus et al., 2007). Oxygenic *Phormidium*-dominated

cyanobacterial mats were more common and more productive at shallower depths in an Antarctic ice-covered lake (Moorhead et al., 1997). Distinct OP cyanobacterial mats dominated by oscillatorial members similar to *Phormidium* and *Planktothrix* were located at different optimal depths in various ice-covered Antarctic lakes based on the amount of light that passed through the ice (Hawes & Schwarz, 2001).

When photosynthetic conditions shift, cyanobacteria respond by modulating their cellular content of proteins (the "proteome"). To maximize light harvesting in changing conditions, cyanobacteria regulate the amount of phycobiliproteins (Hihara et al., 2001) and adjust their pigment content through complementary chromatic acclimation (CCA) (Bryant, 1982). Adjusting their phycobiliprotein contents permit cyanobacteria to most effectively capture the available light when both the quantity and quality of irradiance changes. Previous studies have identified shifts in other cyanobacterial proteins under variable photosynthetic conditions (Cobley et al., 2002). Thioredoxins target a variety of processes and proteins involved with photosynthesis, including Rubisco in carbon fixation and phycobiliproteins in light harvesting (Florencio et al., 2006). Cyanobacteria that experience prolonged darkness tightly control ATP synthase, essential in both photosynthesis and respiration (Imashimizu et al., 2011). Such changes in protein expression and synthesis allow cyanobacteria to adapt to their changing environment, and survive under stress.

Light levels are also tightly intertwined with the balance of AP and OP and the redox geochemistry of cyanobacterial mats. For metabolically flexible cyanobacteria, sulfide is often the preferred electron donor over water, thus sulfide-driven AP controls OP such that in the presence of sulfide no OP will take place (Klatt, Al-Najjar, et al., 2015). Higher light levels stimulate higher rates of AP and faster depletion of sulfide. If rates of sulfide consumption

sufficiently outpace rates of sulfide supply (*e.g.*, from diffusion or local sulfate reduction), then sulfide is depleted and OP takes over. Hence, oxygenic cyanobacteria may produce oxygen oases in hypoxic conditions (Sumner et al., 2015). AP, OP, and light are intertwined due to the differential employment of photosystems I and II (PSI and PSII) in these photosynthetic modes. AP uses only PSI, and thus requires less light energy than OP (Klatt, de Beer, et al., 2016). To date, studies on cyanobacterial light adaptation have characterized geochemical cycling of in situ OP/AP populations on diel scales (Hamilton et al., 2018; Klatt, Meyer, et al., 2016); the molecular responses of AP and OP cyanobacteria to the seasonally changing environment are unknown.

In addition to light, physical and geochemical gradients influence the abundance and activities of microbes in mat systems. In a typical mat, diverse phototrophs including cyanobacteria, photosynthetic eukaryotes such as diatoms, and anoxygenic phototrophs, may be found at mat surfaces, but light availability, sulfide concentration, and temperature define their contribution to primary production (Camacho et al., 2005; Cohen et al., 1975). Low oxygen levels can constrain eukaryotic populations but benefit anoxygenic phototrophs, who also have a lower light limit than oxygenic counterparts. Both oxygen and sulfide influence sulfide oxidizers, who depend on both oxygen availability and abundant sulfide, and sulfate reducers, some of whom may tolerate low oxygen levels (Canfield & Marais, 1991). In such systems located at high latitudes, seasonality may be a strong influence on the environment. Weather-driven events as well as yearly cycles can shift the delivery of nutrients and light to microbial populations (Bolhuis et al., 2014). Thus sampling to capture the temporal variability in the environment provides a better ecological picture of microbial mats.

We observe flexible AP/OP cyanobacteria (*Phormidium* sp.) in the low-oxygen microbial mats of Middle Island Sinkhole, Lake Huron (MIS) (Voorhies et al., 2012). Located 23 m deep, MIS is bathed in low-oxygen, high-sulfate groundwater (Ruberg et al., 2008). Sulfate-reducing bacteria in the mat and sediment provide H_2S for the AP cyanobacteria and sulfide-oxidizing bacteria (Kinsman-Costello et al., 2017). To understand the relationship between environment and microbial community and function, from June to September in multiple years we measured select environmental parameters likely to be influenced by seasonality (light, groundwater chemistry), analyzed microbial community composition and structure via 16S rRNA genes, and evaluated proteins to untangle key microbial functions. Ultimately, these interactions between geochemical gradients, microbial members, and physicochemical parameters shape the microbial composition, dominant metabolisms, and biogeochemical functioning of the mat.

3.3 Materials and methods

Sample collection. We sampled the microbial community of Middle Island Sinkhole (located at 45° 11.914 N, 83° 19.671 W), a 23.0 m deep sinkhole of approximately 125 m length and 100 m width in Lake Huron (Baskaran et al., 2016; Kinsman-Costello et al., 2017; Ruberg et al., 2008). We collected samples once in 2009, and one to three times a year between 2011-2015. Scuba divers from the NOAA Thunder Bay National Marine Sanctuary collected intact flat purple mat and sediment cores using 20 x 7 cm clear polycarbonate tubes and rubber stoppers. Cores were kept upright, in the dark and on ice or at 4˚C until they were sampled within 48 hr of collection. Microbial mat was removed intact from cores, homogenized, stored in 2mL microcentrifuge tubes, and frozen at -80˚C until DNA extraction.

Physicochemical measurements. To measure photosynthetically available radiation (PAR) at 23.0 m depth in the sinkhole, we used a LiCor LI-192 underwater quantum sensor (LiCor Biotechnology) (sampled in 2009-2013, and 2017), a compact optical profiling system for UV light in natural waters (UV C-OPS; Biospherical Instruments Inc.) (Cory et al., 2016) (in 2014), a hyperspectral profiler (HyperPro II profiler, Sea-Bird) (2015-2016), and HOBO loggers mounted 0.25 – 0.75 m above the sediment surface (Onset Computer Corporation) (in 2014-2017). Because HOBO loggers measured light in lux or lumens, their light measurements were calibrated to PAR quantitation through hyperspectral profiling of the 23.0 m-deep light field at the same time as the loggers. *k*-attenuation coefficients were calculated from the linear relationship

$$ln(I_{23}) = ln(I_D) - k(D - 23)$$

where *I* is irradiance at the last measured depth *D* in meters in vertical profiles obtained through LiCor, C-OPs, and HyperPro II instruments. These *k* values were used to project HOBO-acquired values to uniform depth at 23.0 m.

To identify seasonal changes in MIS groundwater chemistry, divers collected 60 mL syringes of groundwater at specific locations in the sinkhole for conductivity and water stable isotope measurements. Specific conductivity in 2014-2018 was measured using a handheld probe (Yellow Springs Instruments Inc.), as well as calculated from ionic composition. We measured major element anion concentrations in water samples including chloride (Cl^-), fluoride (Fl^-), bromide (Br^{2-}), sulfate (SO_4^{2-}), and nitrate (NO_3^-) using electrolytic-suppression ion chromatography (Dionex, Thermo Scientific). We measured concentrations of major cation elements including calcium (Ca^{2+}), sodium (Na^+), potassium (K^+), and magnesium (Mg^{2+}) with inductively coupled plasma-optical emission spectroscopy (ICP-OES, Perkin Elmer). Sulfate

measurements are reported from samples collected in 2012-2013 (Kinsman-Costello et al., 2017), as well as samples newly collected in 2014-2016. We measured $\delta^{18}O$ and δD in select samples from 2015-2018 to trace the influence of water reservoirs and mixing in the groundwater. Isotope measurements were performed on a Picarro L2130-i cavity ringdown spectrometer with an A0211 high-precision vaporizer and attached autosampler (Picarro). The Picarro ChemCorrect software was used to monitor samples for organic contamination. Precision was better than 0.1‰ and 0.3‰ for $\delta^{18}O$ and δD, respectively. For that analysis, we also measured samples from a fountain outside the Alpena, Michigan library in late summer and autumn, which based on its source depth (1267 feet below surface) putatively taps the same aquifer as the source of MIS groundwater (Ruberg et al., 2008). We compared surface water measurements in this study to previous measurements (Baskaran et al., 2016; Jasechko et al., 2014). Using those measurements, we established the Alpena library fountain and surface water as endpoints of a linear mixing model to explain the observed values in the sinkhole. Values were also compared to the Great Lakes mean water line (Jasechko et al., 2014).

Light manipulation experiments. To evaluate whether light is the dominant influence on community structure, we conducted in situ (n = 3) and laboratory (n = 7) experiments that altered the intensity of light available to the microbial mat. In situ, we deployed nylon mesh tents that reduced light intensity by 50% or 90% over 1 m^2 areas of microbial mat. We collected intact microbial mat and sediment cores from these tented areas, as well as mat adjacent and outside the tented areas, one day and 6-9 days after initial deployment, and for some 36-39 days after initial deployment. These samples were processed the same as other field samples for DNA extraction. In the laboratory, we used 60mL syringes to subsample freshly collected mat and

sediment cores. These “microcores” were kept at 8-10˚C in water baths filled with sinkhole groundwater source, and subjected to 100%, 50%, or 10% of in situ light intensity. We used colored film to maintain a light field similar to *in situ* conditions underwater (Rosco, Stamford, CT). Incubations proceeded for up to 82 days, and we sampled initially every 2-3 days for 1 week to capture the most rapid change in community structure, then a final time point at 6 to 82 days. We had one field and four laboratory experiments that began in autumn, and two field and three laboratory experiments that began in late summer. At sampling time points, we collected up to 0.25g wet mat material and immediately froze them at -80˚C until DNA extraction.

DNA preparation and sequencing. We extracted up to 0.5 g of wet mat material using a modified version of the MPBio Fast DNA Spin Kit for Soil (MP Biomedical, Santa Anna, CA, USA). In summary, 0.3 g of beads (corresponding to one large bead, seven medium beads, and an equal volume of small beads), sodium phosphate buffer, and MT buffer was used to chemically and mechanically lyse cells, in either the FastPrep instrument for 45 s (samples up to 2013), or horizontal lysis on a vortex mixer for 10 min at speed 7 (2014-2015). After protein precipitation, DNA was cleaned, pelleted, and resuspended in up to 100 uL nuclease-free water. DNA was stored at 4˚C for immediate quantification and -20˚C for long term.

We used the PicoGreen assay (Invitrogen, Carlsbad, CA, USA) to quantitate double stranded DNA. Samples were diluted to between 1-25 ng/uL and submitted to the University of Michigan Host Microbiome Core for Illumina library preparation and sequencing (Kozich et al., 2013; Seekatz et al., 2015). Bacterial primers 515F/806R were used to amplify the 16S rRNA gene v4 region in a reaction mixture consisting of 5 μL of 4 μM equimolar primer set, 0.15 μL of AccuPrime Taq DNA High Fidelity Polymerase, 2 μL of 10x AccuPrime PCR Buffer II (Thermo

Fisher Scientific), 11.85 µL of PCR-grade water, and 1-10 µL of DNA template. Thermocycling was an initial denaturation at 95˚C for 2 min, 30 cycles of 95˚C for 20 s, 55˚C for 15 s, 72˚C for 5 min, and a final extension of 72˚C for 10 min. PCR products were cleaned and normalized using SequalPrep Normalization Plate Kit (Thermo Fisher Scientific), then quantified and pooled equimolarly according to Kapa Biosystems Library qPCR MasterMix (ROX Low) Quantification kit for Illumina platforms. An Agilent Bioanalyzer kit confirmed library size and purity, and the library pool was sequenced on the Illumina MiSeq using a 500 cycle V2 kit with 15% PhiX for diversity.

Quantitative proteomics. Technical replicates of samples from June (n = 3), July (n = 4), and October (n = 3) 2015 were evaluated through quantitative proteomics. From 0.25 to 5.0g of wet mat material, proteins were extracted, isobarically labeled, and analyzed on LC-MS/MS as described in (Waldbauer et al., 2017). Briefly, proteins were extracted from mat material using a denaturing and reducing extraction buffer (1% SDS, 10% glycerol, 10 mM dithiothreitol, 200 mM Tris, pH 8) at 95˚C for 20 min. Cysteines were alkylated by addition of 40 mM iodoacetamide and incubation in the dark for 30 min. Where not otherwise specified, all solid reagents were dissolved in LC/MS-grade water (Fisher Optima). Proteins were purified by a modified eFASP (enhanced filter-aided sample preparation) protocol (Erde et al., 2014), using Vivacon 500 concentrators (30 kDa nominal cutoff, Sartorius). Proteins were digested with MS-grade trypsin (Thermo Pierce) at 37 °C overnight, and peptides were eluted from the concentrator and dried by vacuum centrifugation. Peptide isobaric labeling is described at protocols.io (dx.doi.org/10.17504/protocols.io.d2i8cd). The C-terminal of peptides was labeled with either ^{18}O or ^{16}O, and the N-terminal of peptides was demethylated with either dideuterated

(D_2) formaldehyde (^{16}O-labeled) or unlabeled (H_2) formaldehyde (^{18}O-labeled). A standard composed of all samples was also labeled and run alongside samples to provide quantitative comparison between samples and between separate runs.

We used protein sequences predicted from previously generated metagenomes (Voorhies et al., 2016) to identify proteins, and metagenome-assembled genomic bins refined through anvi'o (Eren et al., 2015) to link proteins to specific organisms. Briefly, we used 'anvi-gen-contigs-database' to profile the metagenomic co-assembly of 15 samples from 2007-2012. The software used Prodigal v2.6.0 (Hyatt et al., 2010) to call genes, and identified single-copy genes belonging to Bacteria (Campbell et al., 2013) and Archaea (Rinke et al., 2013) through HMMER (Eddy, 2011). Commands 'anvi-centrifuge' and anvi-import-taxonomy' incorporated gene-level taxonomy calls into the database made using centrifuge (Kim et al., 2016). 'anvi-profile' reconciled the mapped reads against the co-assembly to generate differential coverage and tetranucleotide frequency (tnf) information in a profile database, with a minimum scaffold length of 5000 bp. A combination of automated binning using CONCOCT (Alneberg et al., 2014) and refinement based off of tnf and coverage generated 69 metgenome-assembled-genomic bins (MAGs). To recover additional scaffolds that were not long enough (likely due to the assembler not reconciling these highly conserved contigs), we retained cyanobacterially identified bins and iteratively compared tnf and coverage of scaffolds that were 2500-5000bp, 1500-2500bp, and 1000-1500bp long. Our oscillatorial cyanobacterial bins had 96.4% completion/8.63% redundancy (unknown Oscillatorial, 4.4Mbp), 94.96%/97.12% (*Phormidium*, 12.5Mbp), 82.01%/62.59% (*Planktothrix*, 6.8Mbp), as well as a 5.1Mbp collection of cyanobacterial scaffolds that were not able to be confidently binned. Additional bins of interest are listed in **Table SI 3.1**.

16S rRNA gene bioinformatic analysis. Raw pairs of sequencing reads (250 bp) were quality trimmed and merged using 'iu-merge-pairs', which is a program in illumina-utils (available from https://github.com/merenlab/illumina-utils) (Eren et al., 2013), using minimum quality score of 25, minimum overlap of 200 bp, and at points of divergence in the overlap the higher quality basecall was retained. Merged reads with five or fewer mismatches were kept for Minimum Entropy Decomposition v. 2.1 (Eren et al., 2014) using the following parameters: -d 4 -N 3 --min-substantive-abundance 5 -V 3 --relocate-outliers. We used GAST (Huse et al., 2008) to call taxonomy using the curated SILVA database, and confirmed with BLASTN against SILVA 123 (Pruesse et al., 2007) (bacteria and archaea), and PhytoRef (Decelle et al., 2015) (chloroplasts). mothur v. 1.33 (Schloss et al., 2009) was used to check for chimeras de novo, and putatively-chimeric nodes that did not have taxonomy assigned via SILVA 123 and GAST were removed. We searched for sulfate-reducing genera in the Deltaproteobacteria using a taxonomic search for "sulf" or "thio". The read analysis is outlined here: https://hackmd.io/s/r1CGeQs_G

Statistical analyses. We used the R statistical environment (R Core Team, 2015) in RStudio (RStudio Team, 2014) to analyze nodes and proteins. For environmental samples, we used Morisita-Horn metric to calculate a distance matrix on Hellinger-transformed bacterial relative abundances, as input for nonmetric multidimensional scaling (NMDS) with autotransformation = TRUE. For correlation network analyses, we retained bacterial genera that were at least 0.1% of the total community (n = 67). We generated one network using rcorr on a Bray-Curtis distance matrix calculated from the read counts of genera, and after a Benjamini-Hochberg false discovery rate correction, we retained all correlations $p < 0.001$. We used CoNetinR (Faust &

Raes, 2016) to generate three additional network matrices using Spearman's rho correlation, Pearson correlation, and Bray-Curtis, implementing the ReBoot procedure as described in the original CoNet (Faust & Raes, 2012) and Benjamini-Hochberg FDR correction, to keep correlations $p < 0.001$. The final correlation network contained nodes and edges whose direction (positive or negative) were supported by all four matrices, with the edges representing the average of the measures. The final network was visualized in Gephi and in a correllogram via corrplot, with the standard deviation of the mean in the lower triangle. We used amova (Excoffier et al., 1992) in the R package *pegas* (Paradis, 2010) for testing significant difference in bacterial community structure between months and seasons, and LEFSE in the Galaxy Project (Afgan et al., 2018) to identify taxa as biomarkers for seasons and months; those taxa are indicated with significance values. For evaluation of differential abundance of proteins between seasons, we calculated the weighted mean and weighted standard deviation of the $\log_2$normalized abundance ratios of samples taken from the same month. Paired t-tests ($p < 0.05$) corrected with Benjamini-Hochberg false discovery rate were used to retain significantly different weighted mean abundances.

3.4 Results

Light is predictably seasonal.

We measured PAR over several years with varying frequency and methods. For each month, an average and quartiles of maximum daily light intensity were calculated from the summarized daily maxima and the episodic measurements. The highest average daily maximum PAR we observed in a month was 106 μmol photons m^{-2} s^{-1} in July, and the lowest was 2.93 μmol m^{-2} s^{-1} in March (**Figure 3.1**). *k*-extinction coefficients for PAR ranged from 0.12 to 0.14,

were not related to season, and were consistent with previous measurements in oligotrophic Lake Huron (Yousef et al., 2017). Our observations of the microbial mat were limited to months May-June (hereafter referred to as "early summer"), July-August ("late summer"), and September-October ("autumn"). Within that timeframe PAR is generally highest in July, and lowest in May and October. The lowest observed light values were recorded via loggers deployed over multiple months in the sinkhole 0.25 m – 0.75 m above the sediment surface. While it is to be expected that light is generally lower in the winter in a temperate aquatic system, these low measurements may be also compounded by one or a combination of turbidity in the water column due to phytoplankton growth, microbial mat overgrowth on the loggers, and episodic shading from clouds, snow and/or ice.

The quality of light also varied between seasons in Middle Island Sinkhole. Based on 7 hyperspectral profiles of wavelengths between 348-801 nm in 2015-2016, at 23 m depth in the MIS arena the most available wavelengths of light are blue-green (**Figure SI 3.1, Table SI 3.2**). As expected, the amount of blue and green light is directly related to overall PAR availability, whereas in all casts red light is generally below 0.002 μmol m^{-2} s^{-1} nm^{-1}. In the visible spectrum, red wavelengths are the first to be absorbed in the water column, so the lack of red light at 23 m depth is unsurprising.

We observed subtle differences in availability of other wavelengths of light (**Figure SI 3.2**). The average attenuation coefficient for green light (530-570 nm) in July (average of four casts: k = 0.109 ± 0.007 m^{-1}) was generally higher than in June (average of 2 casts k = 0.099 ± 0.009 m^{-1}) and October (1 cast, k = 0.103 m^{-1}) (**Table SI 3.3**), indicating green light is more rapidly attenuated in late summer than in the other months. Hyperspectral profiles (n = 2) at an open water site outside of the sinkhole (Kinsman-Costello et al., 2017) in June 2015 were not

different from those in the sinkhole, indicating that water column dynamics are similar between the two sites. On the other hand, attenuation of red light (620-670 nm) does not change between the seasons (average *k* in June = 0.382 ± 0.041 m^{-1}, July = 0.381 ± 0.034 m^{-1}, October = 0.373 m^{-1}), corroborating largely physical influences (such as DOM absorption in the upper water column) on red light availability are similar throughout the year.

Groundwater chemistry is distinct and seasonally variable in the sinkhole.

Water venting from the sinkhole alcove is thermally and chemically distinct from overlying Lake Huron water, and as it flows through the sinkhole it slowly mixes with lake water (Baskaran et al., 2016; Ruberg et al., 2008). Generally, the specific conductivity of Alpena fountain was on average 5171 μS cm^{-1}, groundwater as it vented from the dominant source (hereafter referred to as "alcove groundwater source") was on average 2732 μS cm^{-1}, and it attenuated to an average of 1813 μS cm^{-1} as it distributed in the bowl-shaped arena of the sinkhole (hereafter referred to as "MIS arena water") (**Table 3.1**, **Figure SI 3.3**). These are higher than lake surface water (226 μS cm^{-1}), and are similar to previously reported values for MIS (Baskaran et al., 2016).

As expected for reservoirs removed from atmospheric influence, groundwater source and MIS water were depleted in both $\delta^{18}O$ and δD compared to lake surface (**Table 3.2, Figure 3.2**). The source varied between -10.1 to -11.6‰ for $\delta^{18}O$ and -71.0 to -79.1‰ for δD. The $\delta^{18}O$ and δD values for MIS water were -8.83 to -10.37‰ and -62.29 to -71.88‰ respectively. Isotope values for the Alpena fountain were tightly constrained and even more depleted compared to the sinkhole (-12.11 to -12.42‰ $\delta^{18}O$ and -82.62 to -86.91‰ δD), and the lake surface was relatively enriched in $\delta^{18}O$ (-6.99 to -7.25‰) and δD (-52.04 to -54.10‰). The metric d-excess, the

difference between measured δD and expected δD calculated from measured $\delta^{18}O$ and the Great Lakes Mean Water Line (GLWL) (Jasechko et al., 2014), was highest for Alpena fountain samples, on average 10.48‰, and predictably decreased with increasing proportion of surface water. Groundwater source samples (d-excess = 8.17-10.40‰) and MIS water (d-excess = 4.83-6.34‰) had increasingly higher amounts of surface water (average d-excess = 1.03‰) (**Figure SI 3.4**). Our measured d-excess for surface waters is within range of the previously measured d-excess for a wide distribution of samples from the surface of Lake Huron (Jasechko et al., 2014).

We also identified seasonal variability in groundwater chemistry. We constructed a linear mixing model using the $\delta^{18}O$ and δD values of the surface and the Alpena library fountain as endpoints (**Figure 3.2**), allowing us to estimate the proportion of these two endmembers in other water samples. (It is possible that the groundwater source from the aquifer mixes with lake water beneath the lake, such as in a hyporheic zone, en route to MIS, however we are not able to measure this impact.) The proportion of water in the sinkhole layer that is sourced from the groundwater aquifer increases from early to late summer, and decreases late summer to autumn. Samples of MIS water from early summer contained on average 39% aquifer source water, and late summer water samples contained an average of 55% aquifer source water. In contrast, autumn sinkhole water samples had on average 47% aquifer source water. We observed a similar trend in the alcove source waters as well, and corroborated the linear mixing model using ion concentrations (SO_4^{2-}, Fl^-, Na^+) (**Figure SI 3.5**). Increased meteoric recharge in spring and autumn from snowmelt and rain likely impacts the discharged water, leading to heavier values of $\delta^{18}O$ and δD. Additionally, increased terrestrial runoff through late spring and storm-induced mixing in autumn also likely reduces the proportion of the aquifer source and enriches the isotopic signature of sinkhole water.

Variability in values of MIS water may be partially attributed to sampling location. While there are small local sources of groundwater around the circumference of the sinkhole arena, the largest source and dominant influence on the isotopic composition is from the alcove groundwater source. The proportion of surface water mixed into the groundwater layer generally increases with greater distance from the alcove source (**Figure SI 3.6**). Samples in 2015 were collected at random within the sinkhole, in 2016 in georeferenced locations, and in 2017-2018 at one specific location. The distance between these sampling locations and the alcove was as much as 50 m, thus samples taken farther from the groundwater source were more impacted by surface geochemistry.

Microbial community structure varies seasonally.

We evaluated the bacterial community observed in the microbial mat samples from 2009-2015. Samples were categorized as "early summer" (from May and June, n = 9), "late summer" (July and August, n = 13), and "autumn" (September-October, n = 12). Diatom chloroplast 16S rRNA genes ranged in abundance from 0.14-44% (average: 12.4%) of all sequencing reads, belonged primarily to Bacillariophyta;Cymbellaceae (**Table SI 3.4**), and were removed from further analysis due to nonlinear relationships between 16S rRNA gene abundances and eukaryotes.

Bacterial communities were dissimilar by month (AMOVA, $p < 0.05$) and by season ($p < 0.08$) (**Figure 3.3, Figure SI 3.7, Table SI 3.5**). Early summer communities are dissimilar to autumn communities, and summer populations show a degree of mixing of the two seasonal groups. Select bacterial populations of interest varied in relative abundance between the seasons (**Figure 3.4**). Of the cyanobacteria, filamentous members *Phormidium* and *Planktothrix* were

most abundant in these samples. On a seasonal scale, *Phormidium* was more abundant in summer samples (on average 33 and 35% of the total bacterial community in early and late summer respectively) compared to autumn samples (4.2%) (LEFSe, $p < 0.05$), whereas *Planktothrix* is more frequently observed in autumn and early summer samples (8.4-8.0% compared to 1.9% in late summer) (LEFSe, $p < 0.05$) (**Table SI 3.4**). Other cyanobacterial groups were typically 5% or less of the community.

Based on taxonomy, putative sulfur-oxidizing bacteria (hereafter "SOB") in this community are primarily gammaproteobacterium *Beggiatoa* and a cohort of 5 epsilonprotebacteria (*Arcobacter*, *Sulfurospirillum*, *Sulfurovum*, *Sulfuricurvum*, *Sulfurimonas*). Across seasons, *Beggiatoa* is at least as abundant (4.8-11%) as the pool of epsilonproteobacteria SOB (2.4-4.8%). Other gammaproteobacteria and epsilonproteobacteria comprised 0-5% of the community on average. Deltaproteobacteria, including putative sulfur and sulfate reducers (hereafter "SRB") based on taxonomy such as *Desulfonema* and *Desulfocapsa,* ranged from 4.8% on average of the total bacterial community in early summer samples to 13% in late summer and autumn (**Table SI 3.6**). Other proteobacterial members and Bacteroidetes each constitute 12-18% on average throughout the year. Firmicutes, Spirochaetes, and Verrucomicrobia were each on average 5% or less of the bacterial community. Other bacterial groups, including Acidobacteria, Chloroflexi, Chlorobi, and Planctomycetes, were infrequently observed (<1% typically).

To evaluate whether light is the dominant influence on community structure, we conducted experiments *in situ* and in the laboratory. In laboratory incubations, we subjected microbial mat intact with sediment to either 100%, 50%, or 10% of in situ light intensity, all at temperatures representative of typical sinkhole source water (8-10˚C). In none of these

experiments did we observe a change in community structure related to light intensity (**Figures SI 3.8 and 3.9**). Unlike samples collected *in situ*, NMDS analysis on these light-manipulated samples performed poorly (stress > 0.18, data not shown), suggesting they were strongly divergent from an expected "core" composition based on season or light intensity.

Populations of sulfur cyclers are linked.

We used correlation network analysis to understand potential positive and negative relationships between the abundances of 60 bacterial genera in the microbial mat (**Figure 3.5, Figure SI 3.10**). All relationships evaluated were positive in direction, and had $r \geq 0.50$. *Beggiatoa* was solely correlated with *Desulfonema*, which was also correlated with Bacteroidetes;*Paludibacter* and *Desulfobacterium*. *Sulfuricurvum* linked together other putative sulfide-oxidizing taxa (*Sulfurospirillum*, *Sulfurovum*) as well as *Desulfobacterium*. *Sulfurovum* in turn linked *Sulfurimonas*, *Desulfocapsa*, *Desulfobulbus*, *Desulfomicrobium*, and *Desulfobacula* to the group. *Thiothrix*, another sulfide-oxidizing gammaproteobacterium, was correlated with *Desulfomicrobium*. *Planktothrix*'s sole correlation was with *Paludibacter*. *Phormidium* was not correlated with any taxa.

Increased abundance of putative sulfur reducers in autumn.

The pool of SRB was larger in late summer and autumn samples (**Table SI 3.6**). Within this group, however, we observed some seasonal peaks in specific genera. Many taxa followed the pattern of *Desulfonema*, which was more abundant in late summer and autumn (5.9-6.0% compared to 0.34%, LEFSe, $p < 0.05$). In contrast, *Desulfocapsa* is more abundant in summer samples compared to autumn (average 2.7-3.0% versus 0.93%). Another SRB, *Desulfotalea*, is

more abundant in early summer than later seasons (0.0012% vs 0.00061% in late summer and <0.0001% in autumn).

Community function differs by season.

Quantitative proteomics analyses showed that the abundances of specific proteins belonging to specific genomic bins varied across the seasons (**Table SI 3.7**). As with our 16S rRNA gene analysis, we divided the months into seasons (June considered early summer, July as late summer, September as autumn). Among proteins assigned to genome bins, *Phormidium* dominated the protein pool in early and late summer (5671 of 7795 total spectra in June and July), whereas diatom proteins (Bacillariophyta) were most abundant in autumn (348 of 657 spectra observed in September). Across the dataset, 53% of observed spectra were attributed to phycobilisome-related proteins, of which *Phormidium* was the dominant contributor. It is likely that we collated multiple *Phormidium* strains in our MAG bin, and were not able to resolve proteins to strain level (**Table SI 3.1**). We were also unable to resolve several observed proteins to specific MAG bins (referred to as "unknown organism"). When possible, proteins were assigned taxonomy based on the protein's contig in IMG.

Phormidium's phycobiliproteins varied in abundance between seasons. Allophycocyanin subunits ApcA and ApcB of an A/B/C operon in Phormidium (3300002026_MIS_100162309-11) were significantly more abundant in early and late summer compared to autumn (**Figure 3.6**). An ApcF protein (3300002027_MIS_101595453) was less abundant in early summer compared to late summer. Other phycobilisome proteins composing phycoerythrin and phycocyanin had dynamic abundances across the seasons (**Table SI 3.7**). We observed PsaB, a photosystem I protein (3300002026_MIS_100342443), in significantly lower abundance in

autumn compared to late summer. Two Thioredoxin TrxA in *Phormidium* (3300002027_MIS_100354662 and 3300002027_MIS_101999963) were both significantly less abundant in autumn compared to the summer months.

Other *Phormidium* proteins related to growth and photosynthesis also varied in abundance between seasons. Chaperonin GroEL of an operon also containing GroES (3300002027_MIS_100694094-5) was significantly less abundant in autumn compared to the summer months. GroES of a different operon (3300002024_MIS_11767773) in *Phormidium* had a similar significant differential abundance. A ribosomal protein L12 RplL (3300002024_MIS_11047972) was significantly lower in autumn than in late summer. Superoxide dismutase (3300002024_MIS_10334211) and a rhodanese involved in sulfur cycling (3300002026_MIS_1001011015) and on the same contig as a microaerobic PsbA and a type B sulfide-quinone reductase (Grim & Dick, 2016), were abundant exclusively in summer and not observed in other months. *Phormidium*'s photosystem II protein PsbA was not recovered from these samples. We also were not able to measure SQR, the sulfide quinone reductase responsible for anoxygenic photosynthesis in cyanobacteria.

Planktothrix, the other dominant cyanobacterium in the mat, also had differential abundance patterns in photosynthesis-related proteins. CpeE of a CpeD/E operon (3300002024_MIS_10075351-2) was much more abundant in late summer than in early summer. AtpD the beta subunit of F0F1-type ATP synthase on an operon also containing the epsilon subunit (3300002027_MIS_100753761-2), as well as PsaD (3300002027_MIS_100129452), were much more abundant in late summer compared to autumn. A *Pseudanabaena* GroEL (3300002026_MIS_1000534011) was less abundant in early compared to late summer. Unknown cyanobacteria also had significantly differentially abundant phycoerythrin (Cpe),

phycocyanin (Cpc), and phycoerythrocyanin (Pec) proteins in summer compared to autumn (**Figure 3.6, Table SI 3.7**).

We recovered at least two contigs of diatom chloroplast genome(s), of 65,732 bp and 42,711 bp in length. We observed more MS/MS spectra belonging to diatom PsbA (22 across the dataset) compared to PsaA (6). Of the other significantly differentially abundant proteins, we observed the beta subunit of F0F1-type ATP synthase (3300002026_MIS_1000010375) from one contig, and the large subunit of Rubisco (3300002026_MIS_1000043450) from the other contig, in much lower abundance in late summer compared to the other season units. Another putative chloroplast contig (21,320 bp) contained a PsaF (3300002026_MIS_1000265837) in significantly lower abundance in late summer compared to the other months. The lengths of these contigs are not enough to represent 2 separate chloroplasts (Brembu et al., 2014), and may belong to the same chloroplast.

Three other F0F1 ATP synthase subunits belonging to different organisms were also differentially abundant. A beta subunit belonging to *Desulfotalea* (3300002024_MIS_11625411) was more abundant in autumn compared to late summer. A diatom's alpha subunit (3300002024_MIS_10280361) was more abundant in autumn compared to early summer. Additionally, another alpha subunit belonging to Chloroflexi;*Anaerolinea* (3300002024_MIS_10189961) was most abundant in late summer, and least abundant in autumn.

Two *Beggiatoa* proteins, a glyceraldehyde 3-phosphate dehydrogenase GapA (3300002024_MIS_10898911) and an outer membrane protein (3300002027_MIS_101719261), were significantly more abundant in autumn compared to late summer, and not observed in early summer. A Betaproteobacteria;*Rhodoferax* FbaA fructose-bisphosphate aldolase

(3300002024_MIS_10200811) was more abundant in late summer compared to early summer. We observed a protein of unknown function, putatively an outer membrane protein (3300002026_MIS_100111357) that belonged to a Gammaproteobacteria;Methylococcaceae member, that was more abundant in autumn compared to late summer.

Several proteins of unknown metagenome-assembled genomic bins had intriguing abundance patterns, but were not found to be significantly different across season units. A dissimilatory sulfite reductase gamma subunit DsrC (3300002027_MIS_100764955) and an alpha/beta subunit DsrA (3300002027_MIS_101406783), both putatively belonging to *Beggiatoa* were observed in higher abundance in autumn compared to late summer. Another *Beggiatoa* DsrA (3300002024_MIS_11408701) and DsrC (3300002024_MIS_11106572) were observed only in autumn. Additional proteins belonging to putative sulfide-oxidizing bacteria (**Table SI 3.7**) were observed exclusively or more often in autumn. A sulfate-reducing DsrA (3300002024_MIS_10135141) and DsrC (3300002027_MIS_101924972) were observed in late summer only. Similarly, a Desulfobacteraceae member's beta subunit of adenosine-5'-phosphosulfate reductase (3300002027_MIS_100225507), for dissimilatory sulfate reduction, was less apparent in autumn compared to summer.

3.5 Discussion

Microbial mats are ideal natural systems to examine the principles of microbial ecology because the organisms sort by metabolism and redox gradients into functional niches (Bolhuis et al., 2014; Ward et al., 1998). In this study, we examined the relationship between the physicogeochemical environment, microbial community composition, and important metabolic functions of a natural microbial mat in a temperate aquatic environment over several months and

years. The structure of the microbial community and the activities of key microbes such as the dominant cyanobacterium *Phormidium*, the dominant sulfide-oxidizer *Beggiatoa*, and sulfate and sulfur reducers shifts between seasons (**Figure 3.7**). Predictable and seasonal changes in light and water chemistry are the likely influences on these microbial members. These shifts in metabolism in tandem with shifts in physicochemistry invite further investigation into other microbial mat systems through the lens of temporal heterogeneity.

To better characterize environmental drivers of seasonality in the microbial mat in the microbial mat, we measured physicochemical parameters that are likely to change over a year in the temperate Lake Huron sinkhole. In this system, we evaluated the changes in light (which has the potential to influence photosynthesis) and groundwater chemistry (which would determine sulfate delivery). Light peaked in summer months and was substantially lower in the other months. Within the months evaluated, water in MIS was more enriched in groundwater (and thus sulfate) in late summer. Due to different adaptations in light harvesting in cyanobacteria and sulfate reduction in the mat, these abiological parameters have the potential to shape community composition, metabolism, and oxygen and sulfide gradients.

Seasonality, through the parameters of groundwater chemistry and irradiance, impacts the abundance and functions of the dominant cyanobacteria in the microbial mat (**Figure 3.7**). In addition to contributing to the oxygen budget through oxygenic photosynthesis, key cyanobacteria *Phormidium* and *Planktothrix* are implicated in sulfur cycling through anoxygenic photosynthetic (AP) gene sulfide quinone reductase (*sqr*) (Dick et al., 2018; Voorhies et al., 2012). In addition, *Phormidium*, *Planktothrix*, and other filamentous cyanobacteria in the sinkhole have variant photosystem II (*psbA*) genes that are optimized for low oxygen conditions (Dick et al., 2018; Grim & Dick, 2016). Though we did not observe variant PsbA and SQR

proteins in our dataset, *Phormidium* and other cyanobacteria demonstrated a more nuanced response in different seasons.

For photosynthesis, cyanobacteria absorb light through phycobilisomes, composed of a core of allophycocyanin, and linker phycobiliproteins, which may include phycoerythrin, phycocyanin, and phycoerythrocyanin. The most direct relationship we observed between our protein profiles, 16S rRNA gene data, and seasonality is evident in phycobiliprotein abundance. The abundance distribution of *Phormidium* allophycocyanin corresponds with its 16S rRNA gene relative abundance in that allophycocyanin was more abundant in early summer (May and June) and late summer (July and August) compared to autumn (September and October). Similarly, in these months light was most available. Previous pigment measurements in *Phormidium*-dominated terrestrial mats collected in autumn, identified higher concentrations of phycoerythrin with lower light conditions, but did not measure allophycocyanin (Snider et al., 2017). In freshwater cyanobacteria, phycobiliprotein synthesis increased with increasing irradiance until reaching an optimum rate at 50-60 μmol m^{-2} s^{-1}, after which the abundance of pigments decreases (Wyman & Fay, 1986) . Our dataset supports a positive relationship of allophycocyanin levels and photosynthetic growth in *Phormidium* with irradiance. Such a linkage invites investigation into using allophycocyanin measurements to quantify the abundance and/or relative growth rate of *Phormidium* in this and other mat systems.

Seasonal changes in the abundance of particular phycobiliproteins, phycoerythrin and phycocyanin, were more complex. Different phycoerythrin and phycocyanin proteins are abundant at different times in the year (**Table SI 3.7**), suggesting the presence of multiple copies and isoforms of phycobilisome rod-core linker proteins in the *Phormidium* genome and/or strain heterogeneity in the *Phormidium* metagenome-assembled genomic bin (Guan et al., 2007;

Rinalducci et al., 2008). In complementary chromatic acclimation (CCA), cyanobacteria can either regulate synthesis of just phycoerythrin (group II organisms), or of both phycoerythrin and phycocyanin (group III organisms) (Tandeau de Marsac & Houmard, 1993). In terrestrial mats dominated by *Phormidium*, PE was more abundant than PC in autumn (Snider et al., 2017). Though in the absence of cultures we cannot confirm to which CCA group the MIS *Phormidium* belongs, different levels and/or isoforms of phycocyanin and phycoerythrin in the season units may be a response to changing light conditions.

The higher abundance of *Phormidium* ribosomal proteins and GroESL proteins in late summer compared to autumn is likely a reflection of more favorable growth conditions for the cyanobacterium. GroEL and GroES chaperonins are required for growth (Georgopoulos & Welch, 1993), and their proportion in the proteome increases linearly with growth rate (Jahn et al., 2018). Further, increased light intensity increases expression of *groESL* promoters (Rajaram et al., 2014). In *Synechocystis*, *groESL* is not transcribed in the dark (Glatz et al., 1997), and the abundance of chaperonins and ribosomal proteins is constrained by light and carbon limitation (Jahn et al., 2018).

Lower thioredoxin levels also corroborate suboptimal photosynthetic conditions for *Phormidium* in autumn. Light energy from photosystem I is used to reduce thioredoxin (TrxA), which interacts with enzymes in the Calvin cycle (carbon fixation) and oxidative pentose phosphate cycle (carbon catabolism) (Tandeau de Marsac & Houmard, 1993). By moderating levels of reduced TrxA, cyanobacteria control key cellular and metabolic processes, such as carbon fixation, nitrogen metabolism, oxidative stress response, and light harvesting (Blankenship, 2014). In *Synechocystis* sp. 6803, *trxA* transcripts decreased when the aerobic culture was transferred to lower light levels (Navarro et al., 2000; Pérez-Pérez et al., 2009). The

lower abundance of TrxA in autumn could reflect *Phormidium*'s lower relative abundance in autumn, but may also imply a shift in their metabolism due to changing photosynthetic conditions. Altogether, *Phormidium*'s differential abundance of phycobiliproteins, thioredoxin, and chaperonins reflects a change in growth rate and/or abundance of *Phormidium* linked to varying growing conditions with the season units.

Temporal patterns in abundance of pigments from other cyanobacteria point to potential niche partitioning of the photosynthetic bacterial community. We observed abundance patterns of allophycocyanin, phycocyanin, and phycoerythrocyanin from *Planktothrix*, *Pseudanabaena*, and unknown oscillatorial cyanobacteria that were higher in summer than in autumn. *Planktothrix*'s statistically significant proteins--a phycoerythrin linker protein CpeE, the beta subunit of an ATP synthase, and a photosystem I component PsaD--were more abundant in late summer than in the other season units, suggesting higher photosynthetic activity during the late summer months. ATP synthesis is integral to photosynthesis and respiration in cyanobacteria, and is dependent upon light and redox conditions. The catalytic sites of the thylakoid membrane-peripheral F1 component of F0F1 ATP synthase are on the 3 beta subunits (Hisabori et al., 2013), which likely require greater synthesis during periods of increased activity. Given the higher abundance of *Planktothrix* phycoerythrin rod-linker protein CpeE and photosystem I PsaD, this increased ATP synthase activity in the summer is likely due to heightened photosynthesis. The detection of non-*Phormidium* pigments in our Phormidium-dominated proteomics dataset points to active light harvesting in these low-abundance members. As with *Phormidum*, it is possible that these organisms may modulate pigment content through CCA, an intriguing topic for future physiological studies on these organisms. The potential for niche

habitation through CCA or different light optima would permit these cyanobacteria to coexist in the microbial mat.

Though abundance of eukaryotic 16S rRNA gene (*i.e.*, chloroplast sequences) is not a quantitative metric of growth or cell abundance, when diatom chloroplasts--mainly of Bacillariophyta; Cymbellaceae--were observed in early summer (on average 13% of the reads), proteins from their chloroplasts such as ATP synthase and Rubisco were also present. However, given their low 16S rRNA gene levels in autumn, diatoms were at the same time more active than expected. Two ATP synthase subunits belonging to diatoms were more abundant in autumn, when diatom 16S rRNA gene abundance is generally lower (on average 8.2% of the 16S rRNA gene reads). In fact, they contributed over half of the protein spectra in autumn. The overwhelming abundance of diatom proteins in autumn indicates these potentially less abundant members are nevertheless quite active.

The differential abundance and changing functional profiles of sulfide-oxidizing bacteria and sulfur/sulfate-reducing deltaproteobacteria across the seasons have implications for sulfur and oxygen cycling in MIS (**Figure 3.7**). *Beggiatoa* is a dominant sulfur oxidizer in the sinkhole mat, contributing up to 35% of bacterial sequences. It is especially abundant in autumn, when we also observe higher abundances of key *Beggiatoa* proteins, such as glyceraldehyde 3-phosphate dehydrogenase, involved in the Calvin-Benson-Bassham cycle, an outer membrane protein likely involved in motility (Yu & Kaiser, 2006), and reverse dissimilatory sulfite reductase proteins. As a filamentous motile bacterium, *Beggiatoa* can move to an optimal vertical position of sulfide and oxygen within the microbial mat (Jørgensen & Revsbech, 1983). In stratified mats, *Beggiatoa* can constrain cyanobacterial photosynthesis by covering the mat and obscuring light (Klatt, Meyer, et al., 2016). Low light also reduces anoxygenic photosynthetic sulfide oxidation,

thus sulfide is more available for chemolithotrophs (Klatt, Meyer, et al., 2016). Taken together with light measurements and higher autumn sulfide levels previously measured (Kinsman-Costello et al., 2017), the increased representation of these proteins, along with *Beggiatoa*'s higher relative abundance in 16S rRNA gene datasets, point to light as an indirect influence on the seasonal change in abundance of this sulfide oxidizer.

The 16S rRNA gene abundance and proteomic profiles of sulfate-reducing deltaproteobacteria suggest interactions with *Beggiatoa* in the sulfur cycle in MIS. The correlation between relative abundances of *Beggiatoa* and *Desulfonema*, the most abundant SRB we observed, may reflect a metabolic interaction in which sulfide produced through *Desulfonema's* sulfate reduction is used by Beggiatoa. The higher abundance of deltaproteobacterial SRB in late summer and autumn, corroborates previously observed higher sulfide availability in the mat and sediment later in the year (Kinsman-Costello et al., 2017). However, in this environment abundance does not always equate to increased functioning. The higher autumn abundance of an ATP synthase subunit of a putative sulfate reducer *Desulfotalea*, a psychrophile that thrives in cold systems (Rabus et al., 2004), indicates activity despite lower abundance in 16S rRNA gene datasets in autumn. Other proteins from unbinned sequences involved in sulfate reduction (DsrA, DsrC, ApsR) were observed in late summer samples either exclusively or in greater abundance. *Desulfocapsa*, another putative sulfur reducer that may grow through inorganic sulfur disproportionation as opposed to dissimilatory sulfate reduction (Finster et al., 2013; Janssen et al., 1996), is also abundant early in the year in contrast to *Desulfonema*. At up to 3.0% relative abundance, *Desulfocapsa* may play an as-of-yet unexplored role in metabolizing elemental sulfur, thiosulfate, and sulfite in the sinkhole. These glimpses into the distribution and activities of sulfate reducers and S disproportionators--key sources of sulfide to

the mat--shed light into their feedback and control on photosynthetic and chemosynthetic metabolisms.

Proteins belonging to other bacteria point towards additional functionally important community members and metabolisms. *Rhodoferax*, a purple nonsulfur bacterium tolerant of cool conditions (Kaden et al., 2014; Madigan et al., 2000), was up to 20% of the bacterial community, and had a carbon fixation protein that was highly abundant in later summer. The ability of some to conduct photosynthesis (Madigan et al., 2000) hints at a supporting role for *Rhodoferax* in providing fixed carbon to the community. We observed an abundant putative outer membrane protein belonging to a gammaproteobacterial methanotroph in high abundance in autumn, supporting active methanogenesis in the system. *Anareolinea* (Chloroflexi), which we observed more often later in the year (up to 0.9% of the bacterial community), has an ATP synthase subunit more abundant in late summer. Characterized isolates of Anaerolineae are typically mesophiles found in anaerobic digesters and groundwater (Yamada, 2006), though genes for aerobic respiration and/or O_2 detoxification were observed in cultured genomes as well as in two recent metagenome-assembled genomes recovered from sulfidic hot springs (Ward et al., 2018). Sulfide-tolerant *Anaerolinea* members capable of oxygen detoxification or flexible aerobic respiration would thrive through redox changes in this microbial mat underpinned by cyanobacterial photosynthesis. Additional genomic and proteomic examination and targeted cultivation of these and other microbial members would untangle their roles in oxygen and sulfur cycling in MIS.

The patterns of community composition and proteomes described above suggest that seasonality in physicochemical conditions influences the dominant metabolisms and organisms in MIS cyanobacterial mats. Photosynthetically active irradiance (PAR) measured in the sinkhole

increased in early summer, peaked in late summer, and declined in autumn. In addition to quantity of light changing, the spectral quality of irradiance in the MIS water column varied between the measured months, which may impact the light harvesting pigments of cyanobacteria. Hyperspectral profiles in the sinkhole from 2015-2016 indicate green light (which phycoerythrin maximally absorbs) is most available in early summer, and least available in late summer, potentially due to growth of surface colonial cyanobacteria (Fahnenstiel & Carrick, 1991). As in marine environments, at 23 m deep in the sinkhole, red light is minimally available, and is not different between the seasons measured. The paucity of red light (which phycocyanin maximally absorbs) throughout the year suggests that modulating at least phycoerythrin abundance (group II response) would be the most likely response to changing light between the seasons. These dynamics in light intensity and quality have ramifications for the cyanobacterial pigment content, and promote growth of ecotypes optimized for different light niches.

Seasonal patterns in sinkhole water chemistry may also impact delivery and availability of sulfate and other metabolic resources. $\delta^{18}O/\delta D$ isotope and specific conductivity measurements confirm that the benthic water layer in the sinkhole is composed of seasonally-changing proportions of Lake Huron surface water and venting sulfate-rich groundwater. Assuming that the composition of venting source groundwater is invariable across the year, our model suggests that early summer and autumn MIS waters have a higher proportion of surface water. The high proportion of venting source groundwater observed in late summer samples aligns with high delivery of sulfate to the system, which may impact microbial sulfur cycling. Sulfate measured in mat porewater and overlying groundwater (Kinsman-Costello et al., 2017) is slightly lower but fits our predictions from the linear mixing model. These data suggest that both the source groundwater as well as surface-driven factors, most importantly seasonal availability

of light, strongly influence sinkhole water chemistry, and their interactions have the potential to shape the metabolisms and structure of the microbial community.

3.6 Conclusion

In this study, we evaluated the relationship between seasonally-changing irradiance and water chemistry, and microbial community structure and function in the microbial mat of Middle Island Sinkhole. Our proteomics surveys highlighted the key role of cyanobacterial photosynthesis in the community, and revealed an active and dynamic community of sulfur-oxidizing and sulfate-reducing bacteria within the mat. Proteins detected from other bacteria suggest that they play functionally important roles in the mat despite their lower abundance.

The microbial community in summer experiences high light and plentiful sulfate, thanks to the elevated proportion of high-sulfate groundwater in the sinkhole. These conditions seem favorable for the growth of the dominant putative AP-cyanobacterium *Phormidium*, as well as a sulfide-tolerant cyanobacterium *Planktothrix*. Sulfate-reducing bacteria with potentially high O_2 tolerance thrive in summer, and are instrumental in transforming sulfate to sulfide for the AP cyanobacteria and sulfide-oxidizing bacteria. Physicochemical conditions change from summer to autumn, including less available light in the sinkhole due to lower solar irradiation, and lower sulfate delivery due to less groundwater and/or more surface water mixing into the sinkhole. Growth and photosynthesis of *Phormidium* is worse in autumn, but photosynthetic eukaryotes such as diatoms may have different optimal growth conditions and are thus abundant and active. The community of sulfate-reducing bacteria is also present in autumn, but the identity of its dominant members is different from summer potentially due to different ideal sulfate and O_2

conditions. Sulfide-oxidizing bacteria, especially *Beggiatoa*, are highly abundant and active in autumn, and may play a greater role in S cycling than AP cyanobacteria.

Our results headline the responses of a putatively anoxygenic photosynthetic cyanobacterial community to shifting irradiance, thereby affecting the C, S, and O budgets. Importantly, this linkage between cyanobacterial activity, sulfur cycling, and seasonally-changing environment encourages interpretations of mats in the geological record relative to seasonality. Our research adds to the understanding of how microbial mats respond to dynamic environments over seasonal time scales.

3.7 Acknowledgements

We are grateful to NOAA Thunder Bay National Marine Sanctuary and divers for field site access, facilities support, and sampling. We thank Judith Klatt, Katy Rico, Matthew Medina, Bopaiah Biddanda, Steve Ruberg, Arjun Chennu, Allen Burton, Michelle Hudson, Phoebe Aron, and Naomi Levin for field sampling, data collection, processing, and analyses.

3.8 References

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544. https://doi.org/10.1093/nar/gky379

Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*(11), 1144–1146. https://doi.org/10.1038/nmeth.3103

Baskaran, M., Novell, T., Nash, K., Ruberg, S. A., Johengen, T., Hawley, N., et al. (2016). Tracing the Seepage of Subsurface Sinkhole Vent Waters into Lake Huron Using Radium and Stable Isotopes of Oxygen and Hydrogen. *Aquatic Geochemistry*, *22*(4), 349–374. https://doi.org/10.1007/s10498-015-9286-7

Biddanda, B. A., McMillan, A. C., Long, S. A., Snider, M. J., & Weinke, A. D. (2015). Seeking sunlight: rapid phototactic motility of filamentous mat-forming cyanobacteria optimize photosynthesis and enhance carbon burial in Lake Huron's submerged sinkholes. *Frontiers in Microbiology*, *6*. https://doi.org/10.3389/fmicb.2015.00930

Blankenship, R. E. (2014). *Molecular mechanisms of photosynthesis* (1st ed.). Blackwell Science Ltd.

Bolhuis, H., Cretoiu, M. S., & Stal, L. J. (2014). Molecular ecology of microbial mats. *FEMS Microbiology Ecology*. https://doi.org/10.1111/1574-6941.12408

Bryant, D. A. (1982). Phycoerythrocyanin and Phycoerythrin: Properties and Occurrence in Cyanobacteria. *Journal of General Microbiology*, *128*, 835–844.

Camacho, A., Rochera, C., Silvestre, J. J., Vicente, E., & Hahn, M. W. (2005). Spatial Dominance and Inorganic Carbon Assimilation by Conspicuous Autotrophic Biofilms in a Physical and Chemical Gradient of a Cold Sulfurous Spring: The Role of Differential Ecological Strategies. *Microbial Ecology*, *50*(2), 172–184. https://doi.org/10.1007/s00248-004-0156-x

Campbell, J. H., O'Donoghue, P., Campbell, A. G., Schwientek, P., Sczyrba, A., Woyke, T., et al. (2013). UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(14), 5540–5545. https://doi.org/10.1073/pnas.1303090110

Canfield, D. E., & Marais, Des, D. J. (1991). Aerobic sulfate reduction in microbial mats. *Science*, *251*, 1471–1473.

Cobley, J. G., Clark, A. C., Weerasurya, S., Queseda, F. A., Xiao, J. Y., Bandrapali, N., et al. (2002). CpeR is an activator required for expression of the phycoerythrin operon (*cpeBA*) in the cyanobacterium *Fremyella diplosiphon* and is encoded in the phycoerythrin linker-polypeptide operon (*cpeCDESTR*). *Molecular Microbiology*, *44*, 1517–1531.

Cohen, Y., Jørgensen, B. B., Padan, E., & Shilo, M. (1975). Sulphide-dependent anoxygenic photosynthesis in the cyanobacterium *Oscillatoria limnetica*. *Nature*, *257*(5526), 489–492. https://doi.org/10.1038/257489a0

Cory, R. M., Davis, T. W., Dick, G. J., Johengen, T., Denef, V. J., Berry, M. A., et al. (2016). Seasonal Dynamics in Dissolved Organic Matter, Hydrogen Peroxide, and Cyanobacterial Blooms in Lake Erie. *Frontiers in Marine Science*, *3*, 3815. https://doi.org/10.3389/fmars.2016.00054

de Beer, D., Weber, M., Chennu, A., Hamilton, T. L., Lott, C., Macalady, J., & M Klatt, J. (2017). Oxygenic and anoxygenic photosynthesis in a microbial mat from an anoxic and sulfidic spring. *Environmental Microbiology*, *19*(3), 1251–1265. https://doi.org/10.1111/1462-2920.13654

Decelle, J., Romac, S., Stern, R. F., Bendif, E. M., Zingone, A., Audic, S., et al. (2015).

PhytoREF: a reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. *Molecular Ecology Resources*, *15*(6), 1435–1445. https://doi.org/10.1111/1755-0998.12401

Dick, G. J., Grim, S. L., & Klatt, J. M. (2018). Controls on O_2 Production in Cyanobacterial Mats and Implications for Earth's Oxygenation. *Annual Review of Earth and Planetary Sciences*, *46*(1), 123–147. https://doi.org/10.1146/annurev-earth-082517-010035

Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Computational Biology*, *7*(10), e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Erde, J., Loo, R. R. O., & Loo, J. A. (2014). Enhanced FASP (eFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments. *Journal of Proteome Research*, *13*(4), 1885–1895. https://doi.org/10.1021/pr4010019

Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: an advanced analysis and visualization platform for "omics data. *PeerJ*, *3*(358), e1319. https://doi.org/10.7717/peerj.1319

Eren, A. M., Maignien, L., Sul, W. J., Murphy, L. G., Grim, S. L., Morrison, H. G., & Sogin, M. L. (2013). Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods in Ecology and Evolution*, *4*(12), 1111–1119. https://doi.org/10.1111/2041-210X.12114

Eren, A. M., Morrison, H. G., Lescault, P. J., Reveillaud, J., Vineis, J. H., & Sogin, M. L. (2014). Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high- throughput marker gene sequences. *The ISME Journal*, 1–12. https://doi.org/10.1038/ismej.2014.195

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNARestriction Data. *Genetics*, *131*, 479–491.

Fahnenstiel, G. L., & Carrick, H. J. (1991). Phototrophic Picoplankton in Lakes Huron and Michigan: Abundance, Distribution, Composition, and Contribution to Biomass and Production. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*, 379–388. https://doi.org/https://doi.org/10.1139/f92-043

Faust, K., & Raes, J. (2012). Microbial interactions: from networks to models. *Nature Publishing Group*, *10*(8), 538–550. https://doi.org/10.1038/nrmicro2832

Faust, K., & Raes, J. (2016). CoNet app: inference of biological association networks using Cytoscape. *F1000Research*, *5*, 1519. https://doi.org/10.12688/f1000research.9050.2

Finster, K. W., Kjeldsen, K. U., Kube, M., Reinhardt, R., Mussmann, M., Amann, R., & Schreiber, L. (2013). Complete genome sequence of *Desulfocapsa sulfexigens*, a marine deltaproteobacterium specialized in disproportionating inorganic sulfur compounds. *Standards in Genomic Sciences*, *8*(1), 58–68. https://doi.org/10.4056/sigs.3777412

Florencio, F. J., Pérez-Pérez, M. E., López-Maury, L., Mata-Cabana, A., & Lindahl, M. (2006). The diversity and complexity of the cyanobacterial thioredoxin systems. *Photosynthesis Research*, *89*(2-3), 157–171. https://doi.org/10.1007/s11120-006-9093-5

Georgopoulos, C., & Welch, W. J. (1993). Role of the Major Heat Shock Proteins as Molecular Chaperones. *Annual Review of Cellular Biology*, *3*, 601–634.

Glatz, A., Horváth, I., Varvasovszki, V., Kovács, E., Török, Z., & Vigh, L. (1997). Chaperonin Genes of the *Synechocystis* PCC 6803 Are Differentially Regulated under Light-Dark Transition during Heat Stress. *Biochemical and Biophysical Research Communications*, *239*, 291–297.

Grim, S. L., & Dick, G. J. (2016). Photosynthetic Versatility in the Genome of *Geitlerinema* sp. PCC 9228 (Formerly *Oscillatoria limnetica* "Solar Lake"), a Model Anoxygenic Photosynthetic Cyanobacterium. *Frontiers in Microbiology*, *7*(590), 1144. https://doi.org/10.3389/fmicb.2016.01546

Guan, X., Qin, S., Zhao, F., Zhang, X., & Tang, X. (2007). Phycobilisomes linker family in cyanobacterial genomes: divergence and evolution. *International Journal of Biological Sciences*, *3*, 434–445.

Hamilton, T. L., Klatt, J. M., de Beer, D., & Macalady, J. L. (2018). Cyanobacterial photosynthesis under sulfidic conditions: insights from the isolate Leptolyngbya sp. strain hensonii. *The ISME Journal*, *12*(2), 568–584. https://doi.org/10.1038/ismej.2017.193

Hawes, I., & Schwarz, A. M. J. (2001). Absorption and utilization of irradiance by cyanobacterial mats in two ice-covered antarctic lakes with contrasting light climates. *Journal of Phycology*, *37*(1), 5–15. https://doi.org/10.1046/j.1529-8817.1999.014012005.x

Hihara, Y., Kamei, A., Kanehisa, M., Kaplan, A., & Ikeuchi, M. (2001). DNA Microarray Analysis of Cyanobacterial Gene Expression during Acclimation to High Light. *The Plant Cell*, *13*, 793–806.

Hisabori, T., Sunamura, E.-I., Kim, Y., & Konno, H. (2013). The Chloroplast ATP Synthase Features the Characteristic Redox Regulation Machinery. *Antioxidants & Redox Signaling*, *19*(15), 1846–1854. https://doi.org/10.1089/ars.2012.5044

Huse, S. M., Dethlefsen, L., Huber, J. A., Welch, D. M., Relman, D. A., & Sogin, M. L. (2008). Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. *PLoS Genetics*, *4*(11), e1000255. https://doi.org/10.1371/journal.pgen.1000255

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bioinformatics*, *11*, 1–11.

Imashimizu, M., Bernát, G., Sunamura, E.-I., Broekmans, M., Konno, H., Isato, K., et al. (2011).

Regulation of F0F1-ATPase from Synechocystis sp. PCC 6803 by gamma and epsilon subunits is significant for light/dark adaptation. *Journal of Biological Chemistry*, *286*(30), 26595–26602. https://doi.org/10.1074/jbc.M111.234138

Jahn, M., Vialas, V., Karlsen, J., Maddalo, G., Edfors, F., Forsström, B., et al. (2018). Growth of Cyanobacteria Is Constrained by the Abundance of Light and Carbon Assimilation Proteins. *CellReports*, *25*(2), 478–486.e8. https://doi.org/10.1016/j.celrep.2018.09.040

Janssen, P. H., Schuhmann, A., Bak, F., & Liesack, W. (1996). Disproportionation of inorganic sulfur compounds by the sulfate-reducing bacterium *Desulfocapsa thiozymogenes* gen. nov., sp. nov. *Archives of Microbiology*, *166*(3), 184–192. https://doi.org/10.1007/s002030050374

Jasechko, S., Gibson, J. J., & Edwards, T. W. D. (2014). Stable isotope mass balance of the Laurentian Great Lakes. *Journal of Great Lakes Research*, *40*(2), 336–346. https://doi.org/10.1016/j.jglr.2014.02.020

Jørgensen, B. B., & Revsbech, N. P. (1983). Colorless Sulfur Bacteria, *Beggiatoa* spp. and *Thiovulum* spp., in O_2 and H_2S Microgradients. *Applied and Environmental Microbiology*, *45*(4), 1261–1270.

Kaden, R., Sproer, C., Beyer, D., & Krolla-Sidenstein, P. (2014). *Rhodoferax saidenbachensis* sp. nov., a psychrotolerant, very slowly growing bacterium within the family *Comamonadaceae*, proposal of appropriate taxonomic position of *Albidiferax ferrireducens* strain T118T in the genus *Rhodoferax* and emended description of the genus *Rhodoferax*. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 4), 1186–1193. https://doi.org/10.1099/ijs.0.054031-0

Kim, D., Song, L., Breitwieser, F. P., & Salzberg, S. L. (2016). Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Research*, *26*(12), 1721–1729. https://doi.org/10.1101/gr.210641.116

Kinsman-Costello, L. E., Sheik, C. S., Sheldon, N. D., Allen Burton, G., Costello, D. M., Marcus, D., et al. (2017). Groundwater shapes sediment biogeochemistry and microbial diversity in a submerged Great Lake sinkhole. *Geobiology*, *15*(2), 225–239. https://doi.org/10.1111/gbi.12215

Klatt, J. M., Al-Najjar, M. A. A., Yilmaz, P., Lavik, G., de Beer, D., & Polerecky, L. (2015). Anoxygenic photosynthesis controls oxygenic photosynthesis in a cyanobacterium from a sulfidic spring. *Applied and Environmental Microbiology*, *81*(6), 2025–2031. https://doi.org/10.1128/AEM.03579-14

Klatt, J. M., de Beer, D., Häusler, S., & Polerecky, L. (2016). Cyanobacteria in Sulfidic Spring Microbial Mats Can Perform Oxygenic and Anoxygenic Photosynthesis Simultaneously during an Entire Diurnal Period. *Frontiers in Microbiology*, *7*(116), 440. https://doi.org/10.3389/fmicb.2016.01973

Klatt, J. M., Meyer, S., Häusler, S., Macalady, J. L., de Beer, D., & Polerecky, L. (2016).

Structure and function of natural sulphide-oxidizing microbial mats under dynamic input of light and chemical energy. *The ISME Journal*, *10*(4), 921–933. https://doi.org/10.1038/ismej.2015.167

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, *79*(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature*, *506*(7488), 307–315. https://doi.org/10.1038/nature13068

Madigan, M. T., Jung, D. O., Woese, C. R., & Achenbach, L. A. (2000). *Rhodoferax antarcticus* sp. nov., a moderately psychrophilic purple nonsulfur bacterium isolated from an Antarctic microbial mat. *Archives of Microbiology*, *173*(4), 269–277. https://doi.org/10.1007/s002030000140

Moorhead, D. L., Wolf, C. F., & Wharton, R. A. (1997). Impact of light regimes on productivity patterns of benthic microbial mats in an antarctic lake: A modeling study. *Limnology and Oceanography*, *42*(7), 1561–1569. https://doi.org/10.4319/lo.1997.42.7.1561

Navarro, F., Martín-Figueroa, E., & Florencio, F. J. (2000). Electron transport controls transcription of the thioredoxin gene (*trxA*) in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Molecular Biology*, *43*(1), 23–32.

Oberhaus, L., Briand, J. F., Leboulanger, C., Jacquet, S., & Humbert, J. F. (2007). Comparative effects of the quality and quantity of light and temperature on the growth of *Planktothrix agardhii* and *P. rubescens*. *Journal of Phycology*, *43*(6), 1191–1199. https://doi.org/10.1111/j.1529-8817.2007.00414.x

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Pérez-Pérez, M. E., Martín-Figueroa, E., & Florencio, F. J. (2009). Photosynthetic Regulation of the Cyanobacterium *Synechocystis* sp. PCC 6803 Thioredoxin System and Functional Analysis of TrxB (Trx *x*) and TrxQ (Trx *y*) Thioredoxins. *Molecular Plant*, *2*(2), 270–283. https://doi.org/10.1093/mp/ssn070

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, *35*(21), 7188–7196. https://doi.org/10.1093/nar/gkm864

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria. Retrieved from https://www.r-project.org/

Rabus, R., Ruepp, A., Frickey, T., Rattei, T., Fartmann, B., Stark, M., et al. (2004). The genome

of Desulfotalea psychrophila, a sulfate-reducing bacterium from permanently cold Arctic sediments. *Environmental Microbiology*, *6*(9), 887–902. https://doi.org/10.1111/j.1462-2920.2004.00665.x

Rajaram, H., Chaurasia, A. K., & Apte, S. K. (2014). Cyanobacterial heat-shock response: role and regulation of molecular chaperones. *Microbiology*, *160*(Pt_4), 647–658. https://doi.org/10.1099/mic.0.073478-0

Rinalducci, S., Roepstorff, P., & Zolla, L. (2008). De novosequence analysis and intact mass measurements for characterization of phycocyanin subunit isoforms from the blue-green alga *Aphanizomenon flos-aquae*. *Journal of Mass Spectrometry*, *44*(4), 503–515. https://doi.org/10.1002/jms.1526

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J.-F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, *499*(7459), 431–437. https://doi.org/10.1038/nature12352

RStudio Team. (2014). RStudio: Integrated Development for R. Retrieved from http://www.rstudio.com/

Ruberg, S. A., Kendall, S. T., & Biddanda, B. A. (2008). Observations of the Middle Island Sinkhole in Lake Huron–a unique hydrogeologic and glacial creation of 400 million years. *Marine Technology Society Journal*, *42*(4), 12–21. https://doi.org/10.4031/002533208787157633

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, *75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Seekatz, A. M., Theriot, C. M., Molloy, C. T., Wozniak, K. L., Bergin, I. L., & Young, V. B. (2015). Fecal Microbiota Transplantation Eliminates Clostridium difficile in a Murine Model of Relapsing Disease. *Infection and Immunity*, *83*(10), 3838–3846. https://doi.org/10.1128/IAI.00459-15

Snider, M. J., Biddanda, B. A., Lindback, M., Grim, S. L., & Dick, G. J. (2017). Versatile photophysiology of compositionally similar cyanobacterial mat communities inhabiting submerged sinkholes of Lake Huron. *Aquatic Microbial Ecology*, *79*(1), 63–78. https://doi.org/10.3354/ame01813

Sumner, D. Y., Hawes, I., Mackey, T. J., Jungblut, A. D., & Doran, P. T. (2015). Antarctic microbial mats: A modern analog for Archean lacustrine oxygen oases. *Geology*, *43*(10), 887–890. https://doi.org/10.1130/G36966.1

Tandeau de Marsac, N., & Houmard, J. (1993). Adaptation of cyanobacteria to environmental stimuli: new steps towards molecular mechanisms. *FEMS Microbiology Reviews*, *104*, 119–190.

Voorhies, A. A., Biddanda, B. A., Kendall, S. T., Jain, S., Marcus, D. N., Nold, S. C., et al. (2012). Cyanobacterial life at low O_2: community genomics and function reveal metabolic versatility and extremely low diversity in a Great Lakes sinkhole mat. *Geobiology*, *10*(3), 250–267. https://doi.org/10.1111/j.1472-4669.2012.00322.x

Voorhies, A. A., Eisenlord, S. D., Marcus, D. N., Duhaime, M. B., Biddanda, B. A., Cavalcoli, J. D., & Dick, G. J. (2016). Ecological and genetic interactions between cyanobacteria and viruses in a low-oxygen mat community inferred through metagenomics and metatranscriptomics. *Environmental Microbiology*, *18*(2), 358–371. https://doi.org/10.1111/1462-2920.12756

Waldbauer, J., Zhang, L., Rizzo, A., & Muratore, D. (2017). diDO-IPTL: A Peptide-Labeling Strategy for Precision Quantitative Proteomics. *Analytical Chemistry*, *89*(21), 11498–11504. https://doi.org/10.1021/acs.analchem.7b02752

Ward, D. M., Bateson, M. M., Ferris, M. J., Kuhl, M., Wieland, A., Koeppel, A., & Cohan, F. M. (2006). Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1475), 1997–2008. https://doi.org/10.1098/rstb.2006.1919

Ward, L. M., McGlynn, S. E., & Fischer, W. W. (2018). Draft Genome Sequences of Two Basal Members of the *Anaerolineae* Class of *Chloroflexi* from a Sulfidic Hot Spring. *Genome Announcements*, *6*(25), 629. https://doi.org/10.1128/genomeA.00570-18

Wyman, M., & Fay, P. (1986). Underwater Light Climate and the Growth and Pigmentation of Planktonic Blue-Green Algae (Cyanobacteria) I. The Influence of Light Quantity. *Proceedings of the Royal Society of London*, *227*(1248), 367–380.

Yamada, T. (2006). *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. *International Journal of Systematic and Evolutionary Microbiology*, *56*(6), 1331–1340. https://doi.org/10.1099/ijs.0.64169-0

Yousef, F., Shuchman, R., Sayers, M., Fahnenstiel, G., & Henareh, A. (2017). Water clarity of the Upper Great Lakes: Tracking changes between 1998-2012. *Journal of Great Lakes Research*, *43*(2), 239–247. https://doi.org/10.1016/j.jglr.2016.12.002

Yu, R., & Kaiser, D. (2006). Gliding motility and polarized slime secretion. *Molecular Microbiology*, *63*(2), 454–467. https://doi.org/10.1111/j.1365-2958.2006.05536.

3.9 Tables and Figures

Table 3.1 Specific conductivity in water parcels. Units are in μS cm^{-1}.

Date	Season	Alpena	Alcove	Arena	Surface
20120531	Early summer			1821	
20120725	Late summer			1930	
20120927	Autumn	4798		1883	
20130509	Early summer			1793	
20130718	Late summer		3331	2086	
20130928	Autumn			1825	
20140724	Late summer		4058	1702	245
20140924	Autumn	5543	3574	1790	249
20150731	Late summer		2431	1916	237
20150731	Late summer			2161	188
20151007	Autumn		2012	1948	229
20160609	Early summer		1461	1880	219
20160609	Early summer			1640	
20160722	Late summer			1170	217
20160928	Autumn			1979	255
20170809	Late summer		2258	1856	218
20170810	Late summer			1821	205
20170910	Autumn			1787	
20170926	Autumn			1839	
20170927	Autumn			1695	
20170928	Autumn			1387	
20180710	Late summer			1985	
Average		5171	2732	1813	226

Table 3.2 $\delta^{18}O$ and δD measured in water samples from Alpena fountain, MIS Arena water and specific locations in the arena (when available), the Alcove groundwater source of the sinkhole, and the surface of the lake above the sinkhole.

Date	Location	Sample Name	Year	$\delta^{18}O$	δD
20151007	Alpena	ALP.2015.008	2015	-12.19	-83.80
20151007	Alpena	ALP.2015.014	2015	-12.12	-82.62
20170912	Alpena	ALP.2017.Sept	2017	-12.32	-84.24
20170810	Alpena	ALP.2017.Aug	2017	-12.42	-86.91
20180710	Alpena	ALP.2018.July	2018	-12.11	-83.25
20150731	Arena	MIS.2015.042	2015	-9.73	-68.77
20151007	Arena	MIS.2015.077	2015	-10.37	-71.88
20160609	P4	MIS2016.009	2016	-9.73	-67.99
20160609	P10	MIS2016.016	2016	-9.18	-65.86
20160609	G	MIS2016.023	2016	-8.98	-65.08
20160928	D3-D4	MIS.2016.Arena D3-4	2016	-9.30	-65.18
20160928	A6-B6	MIS.2016.325	2016	-9.04	-64.30
20160928	lush	MIS.2016.328	2016	-9.53	-67.17
20170531	P7	MIS.2017.012	2017	-8.84	-63.00
20170531	P7	MIS.2017.013	2017	-8.83	-62.29
20170810	P7	MIS.2017.083	2017	-10.25	-71.42
20180710	P7	P7.2018.July	2018	-9.87	-70.04
20150604	Alcove	MIS.2015.008	2015	-11.43	-77.96
20150731	Alcove	MIS.2015.047	2015	-11.47	-78.34
20151007	Alcove	MIS.2015.079	2015	-10.14	-70.99
20160609	Alcove	MIS2016.001	2016	-11.36	-77.52
20160928	Alcove	MIS.2016.322	2016	-11.28	-77.69
20170531	Alcove	MIS.2017.005	2017	-11.54	-78.32
20170809	Alcove	MIS.2017.062	2017	-11.53	-78.53
20170810	Alcove	MIS.2017.080	2017	-11.60	-79.14
20150604	Surface	MIS.2015.001	2015	-7.13	-52.87
20151007	Surface	MIS.2015.073	2015	-7.14	-53.24
20160609	Surface	MIS2016.026	2016	-7.06	-52.54
20160928	Surface	MIS.2016.319	2016	-6.99	-52.04
20170531	Surface	MIS.2017.002	2017	-7.12	-52.43
20170809	Surface	MIS.2017.056	2017	-7.21	-52.39
20170810	Surface	MIS.2017.077	2017	-7.25	-53.20
20180710	Surface	SFC.2018.July	2018	-7.20	-54.10

Figure 3.1 Maximal daily light observed each month in the sinkhole. For each month, daily maximum light values from logged data (2014-2017) and recorded light from episodic measurements (2009-2016) (grey transparent circles) were summarized in box and whisker plots and averaged per month. Boxes represent the 25-75th percentiles with whiskers covering the minimum and maximum values. The average is plotted with the thick line.

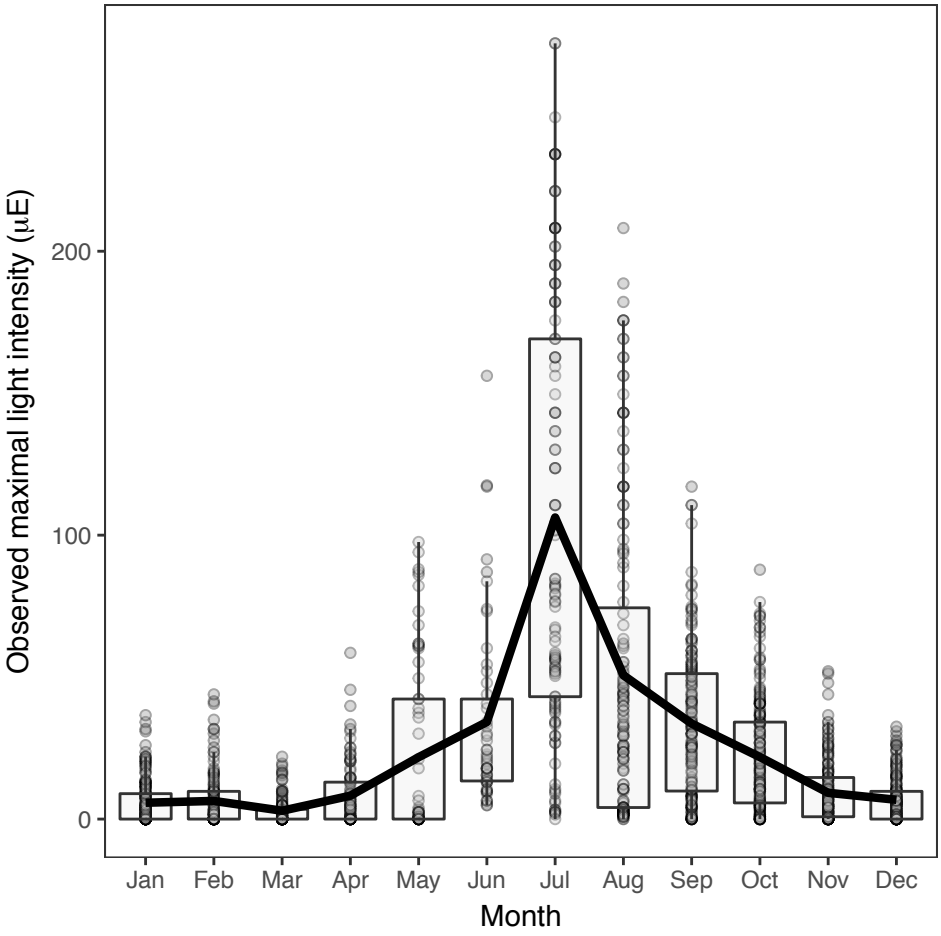


Figure 3.2 Measured $\delta^{18}O$ in water samples by location and season, and proportion of surface water in samples. Measured values are plotted in colored circles and summarized with box and whisker plots. The average $\delta^{18}O$ values for Alpena fountain and surface were used as end-members in a linear mixing model to estimate the proportion of surface water in each sample (grey triangles).

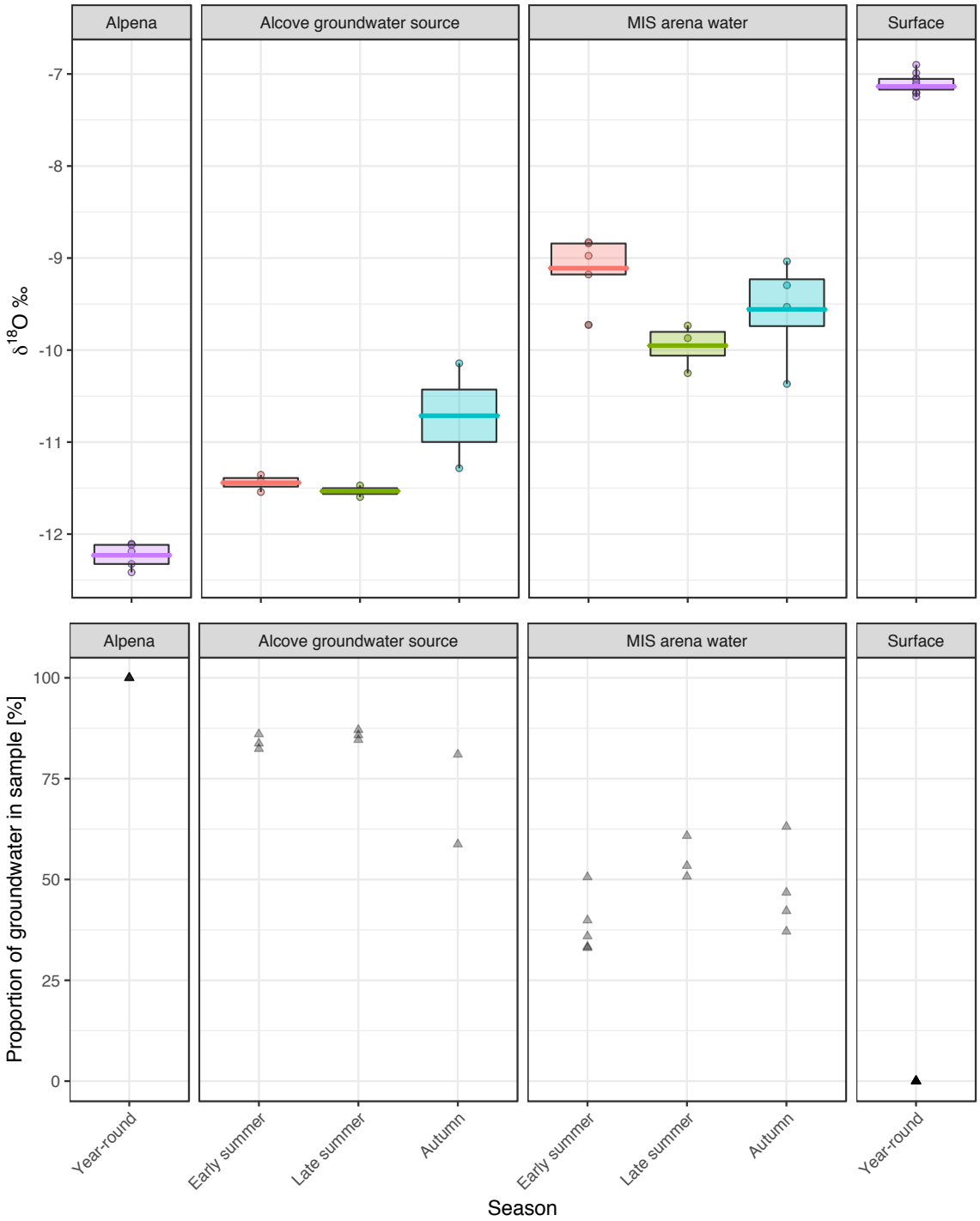


Figure 3.3 Nonmetric multidimensional scaling plot relating bacterial community structure in samples by season. A Morisita-Horn distance matrix was calculated from Hellinger-transformed bacterial relative abundances, and used to plot NMDS. Samples in May (diamonds) and June (squares) are colored green for early summer, July samples are white circles, and samples from September (upright triangles), and October (lower triangles) are colored pink.

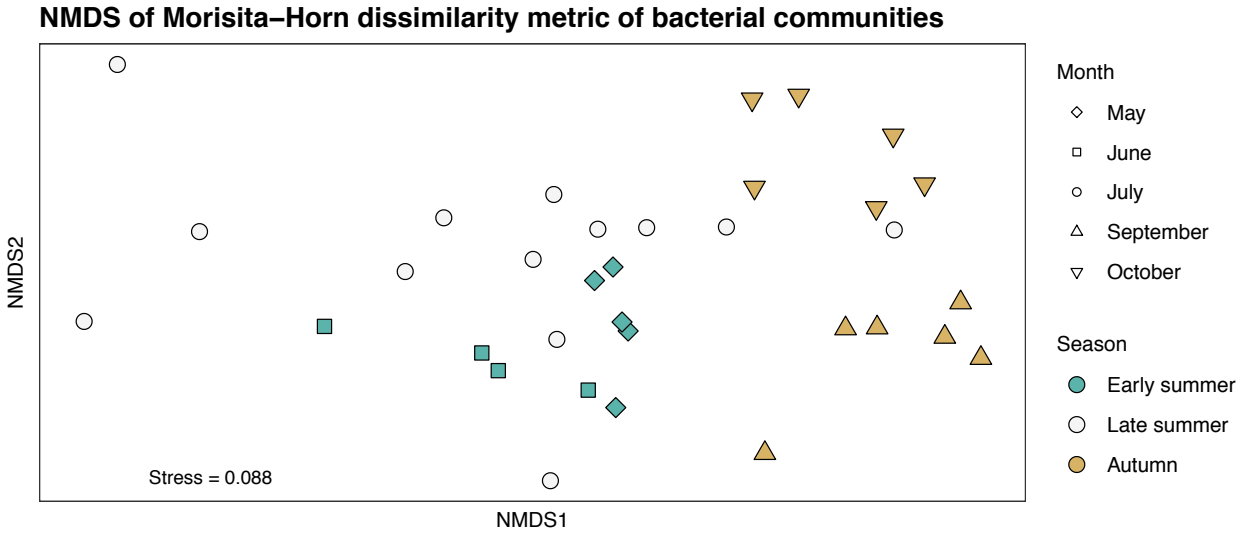


Figure 3.4 Relative abundance of relevant bacterial taxonomic groups in samples, grouped by season. Relative abundances of key genera were summed and presented, and classes and phyla without those genera are represented.

Figure 3.5 Correlation network between relative abundances of key genera. Bray-Curtis distance matrix was calculated from read counts of genera and used to generate a correlation network. Only significant ($p < 0.001$) positive ($r \geq 0.6$) correlations between specific organisms are plotted in random colors.

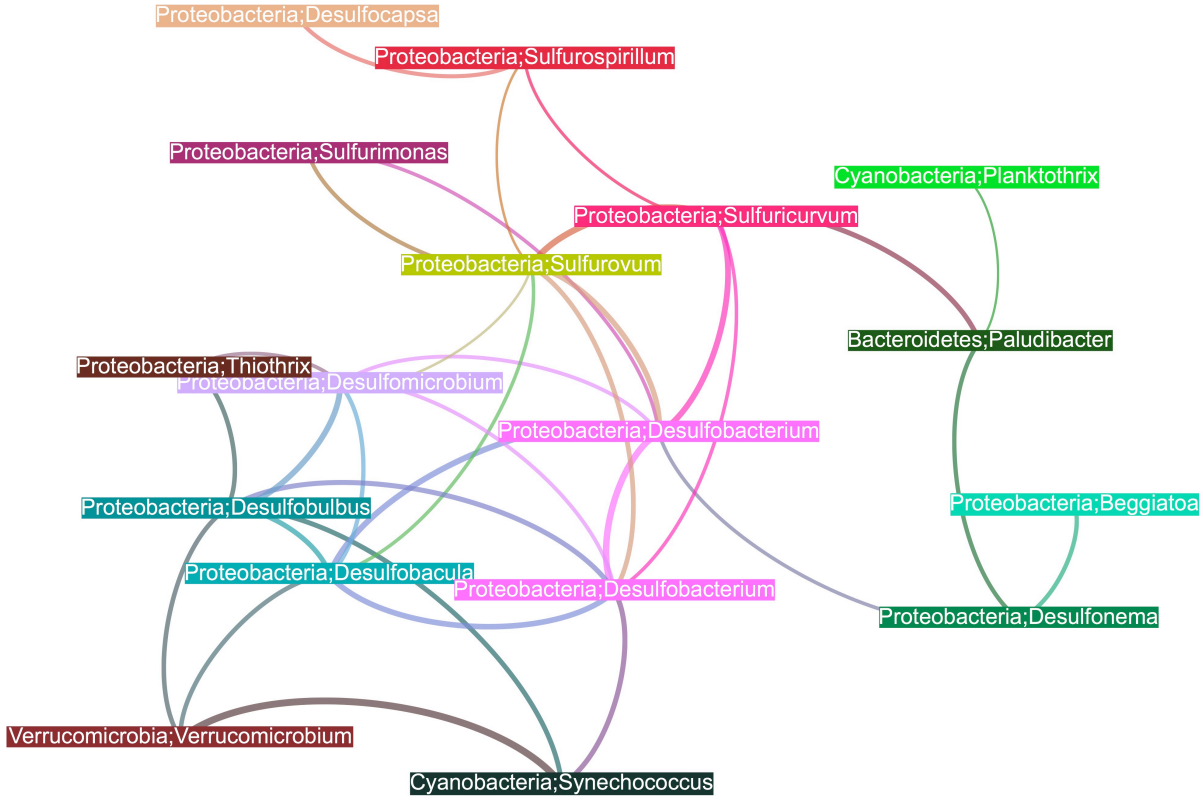


Figure 3.6 Significantly differentially abundant proteins observed in each season. Weighted-mean log_2-normalized observations of proteins were used in paired t-tests to determine significant changes in abundance. The results of paired significance testing between two season units are presented in the right hand columns: for example, proteins more abundant in early summer versus late summer are noted with "E > L". Proteins are grouped by MAG bin, and described by function.

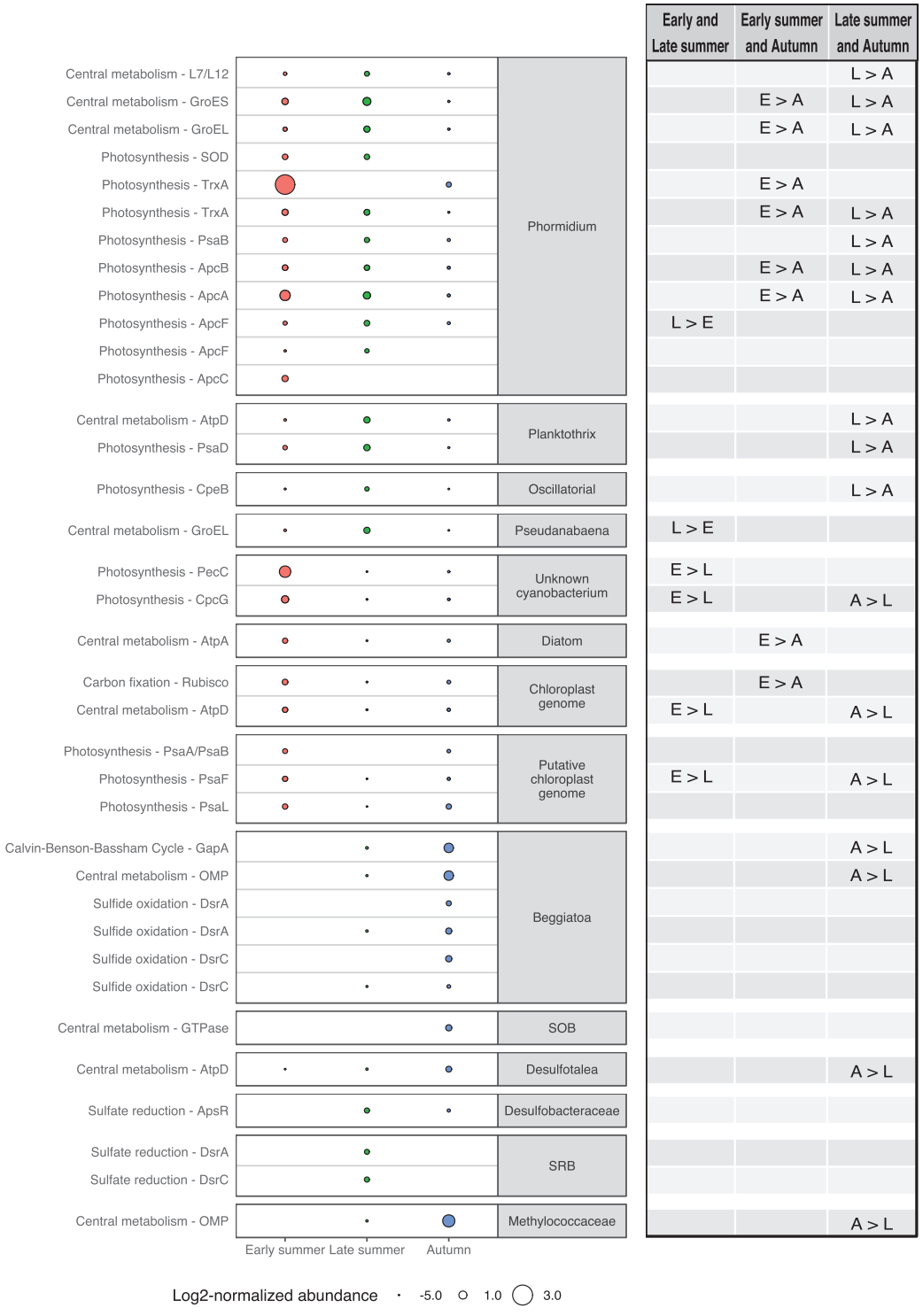
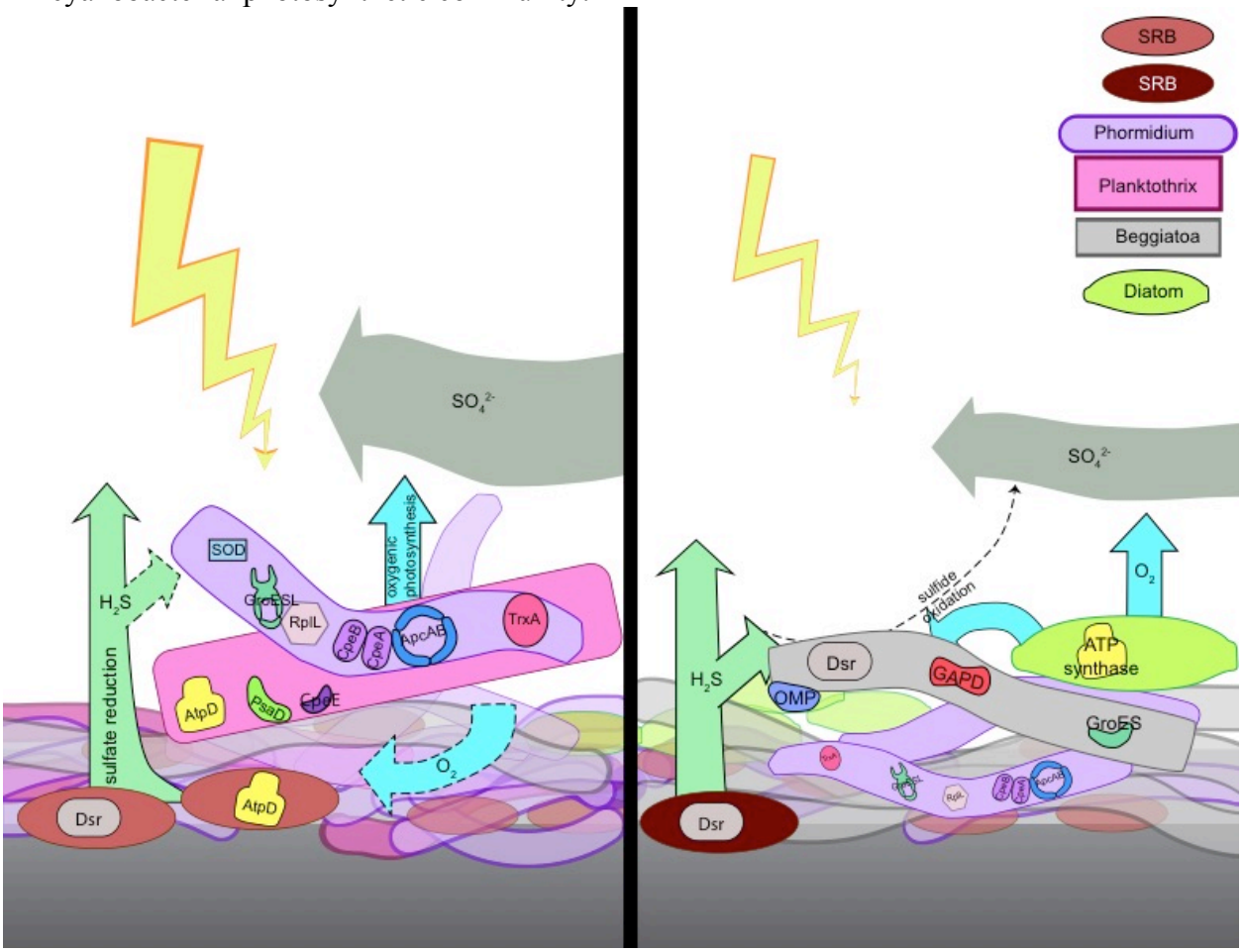


Figure 3.7 Changing physicochemical environment, functions and membership of the sulfur-cycling microbial mat in MIS. In summer (left), light intensity is high and cyanobacteria, dominated by oxygenic photosynthetic *Phormidium* and *Planktothrix*, are abundant and active. Sulfate is abundant and sulfate-reducing bacteria (SRB) provide sulfide, putatively used by *Phormidium* in anoxygenic photosynthesis. In autumn (right), light intensity and concentrations of sulfate are lower. SRB typically found lower in the sediment dominate sulfide production, and *Beggiatoa* is active and abundant. Diatoms supplant the cyanobacterial photosynthetic community.



3.10 Appendix B

Supplemental Information

Table SI 3.1. Summary table of metagenome-assembled-genomic (MAG) bins from MIS metagenomic co-assembly of 15 samples. Our oscillatorial cyanobacterial bins had 96.4% completion/8.63% redundancy (unknown Oscillatorial, 4.4Mbp), 94.96%/97.12% (*Phormidium*, 12.5Mbp), 82.01%/62.59% (*Planktothrix*, 6.8Mbp), as well as 5.1Mbp of cyanobacterial scaffolds that were not able to be confidently binned.

Bin Name	Taxonomy	Total length (bp)	Number contigs	N50	GC Content (%)	Completion (%)	Redundancy (%)
Bin_10_1	Oxalobacteraceae	1399360	102	14694	52.2	83.5	0.7
Bin_10_2	None	8793166	1065	8429	54.4	40.1	25.9
Bin_1_1_2	Unknown	3697586	235	19988	58.1	87.8	0.0
Bin_1_1_3	Unknown	2426406	236	10340	59.8	79.9	0.7
Bin_1_1_4	Unknown_ Betaproteobacteria	2692753	288	9561	56.7	81.3	2.9
Bin_1_2_1	Unknown	32047170	3420	9539	56.2	100.0	513.7
Bin_1_2_2	Unknown	3959061	339	13750	54.8	74.1	1.4
Bin_1_2_3	Desulfomicrobiaceae	1971419	184	11209	60.1	38.8	0.0
Bin_1_2_4	Verrucomicrobiaceae	2691273	293	9464	60.5	57.6	0.7
Bin_2_1	Unknown	43596255	4158	11296	39.3	100.0	982.7
Bin_2_1_1	Unknown	3435637	107	55038	48.3	95.7	1.4
Bin_2_2	Prolixibacteraceae	10282644	1303	7720	39.2	91.4	158.3
Bin_2_3	Unknown	4088197	269	19807	39.3	92.8	2.9
Bin_2_4	Bacteriovoracaceae	3220591	137	32252	41.5	87.1	2.2
Bin_2_5	Chitinivibrionaceae	4044161	360	12352	44.0	80.6	7.2
Bin_3_1	Unknown	26754205	2357	12395	39.7	93.5	363.3
Bin_3_2	Pseudanabaenaceae	330214	20	19027	43.4	4.9	0.0
Bin_3_3	Unknown	25549780	2621	9636	43.3	96.4	379.1
Bin_3_4	Unknown	3335697	298	12662	51.3	89.9	4.3
Bin_4_1	Comamonadaceae	4464091	437	11208	59.6	92.1	73.4
Bin_4_2	Comamonadaceae	4626934	192	33064	58.3	85.6	2.2
Bin_4_3_1	Spirochaetaceae	3855763	203	27591	56.7	87.8	1.4
Bin_4_3_2	Unknown	3686202	224	22146	58.1	82.0	2.2
Bin_4_3_3	Comamonadaceae	3478904	360	9789	60.3	71.2	8.6
Bin_4_3_4	Unknown	3051872	248	14317	61.2	75.5	1.4
Bin_4_3_5	Comamonadaceae	2561953	248	11544	58.7	75.5	1.4
Bin_4_3_6	Thiobacillaceae	2169430	214	11226	59.2	76.3	0.7
Bin_4_4	Comamonadaceae	3519382	181	24907	61.7	95.0	0.7
Bin_5_1	Unknown	8296834	1051	7928	45.7	88.5	95.0
Bin_5_1_1	Oscillochloridaceae	2838987	206	16267	49.6	85.6	2.2
Bin_5_2	Gallionellaceae	3964793	445	8827	48.5	83.5	53.2
Bin_5_3	Anaerolineaceae	4097665	193	30506	50.3	85.6	0.0
Bin_5_4	Desulfobacteraceae	3672135	217	21210	45.4	97.1	1.4
Bin_5_5	Desulfobacteraceae	4525442	478	9962	45.6	87.1	2.2
Bin_5_6	Unknown	1993497	236	8446	52.6	74.1	1.4
Bin_5_7	Unknown	4463472	401	12458	46.4	61.9	1.4
Bin_5_8	Methylococcaceae	2679520	199	15693	41.3	92.1	2.2
Bin_6_2_1	Unknown	2784637	345	8051	37.0	61.2	1.4
Bin_6_2_2	Unknown	3998633	434	9474	37.8	67.6	0.0
Bin_6_2_3	Unknown	1565326	222	6928	32.3	61.9	0.0

Bin_6_3	Helicobacteraceae	2205377	58	60466	37.6	98.6	0.7
Bin_6_4	Unknown	4567808	305	19067	33.9	97.1	0.7
Bin_6_5	Unknown	5045492	570	9270	37.7	82.7	65.5
Bin_6_6	Helicobacteraceae	2227546	83	39393	42.1	96.4	2.9
Bin_6_8	Unknown	3469702	357	10098	37.1	81.3	4.3
Bin_6_9	Campylobacteraceae	1969826	18	168327	30.9	97.8	0.7
Bin_6_9_1	Paludibacteraceae	3068346	236	15389	36.8	76.3	2.2
Bin_7_1	Comamonadaceae	1887820	188	10541	65.7	71.9	1.4
Bin_7_2	Unknown	3548243	264	15864	70.4	59.7	0.7
Bin_7_3	Unknown	3910175	248	19011	68.5	84.2	2.2
Bin_7_4	Unknown	5226942	268	25696	60.0	95.7	0.7
Bin_7_5	Unknown	3094174	314	10652	68.2	46.8	0.0
Bin_7_6	Unknown_Burkholderiales	2119882	272	7621	67.2	46.0	4.3
Bin_7_7	Unknown	3882231	175	28826	68.4	94.2	2.9
Bin_7_8	Unknown_Betaproteobacteria	2826391	131	33263	58.6	74.1	3.6
Bin_7_9	Paludibacteraceae	3708936	283	16092	37.2	84.9	0.0
Bin_8_1	Flavobacteriaceae	2595863	182	18366	31.6	66.9	0.7
Bin_8_2	None	27397424	2633	11361	29.2	84.2	82.7
Bin_9_3	Unknown	3787733	260	17794	42.6	84.9	2.2
Bin_9_4	Unknown	2448590	206	13927	34.3	90.6	0.0
Bin_9_5	Unknown	6620208	323	32199	37.6	91.4	2.9
Bin_9_cyano_1	Microcoleaceae	18617	3	6498	41.3	0.0	0.0
Bin_9_cyano_2	Microcoleaceae	1337148	158	8092	40.6	11.5	0.0
Bin_Oscillatoriales_3	Pseudanabaenaceae	4377283	576	15153	42.1	96.4	8.6
Bin_Phormidium	Microcoleaceae	12463200	5269	2679	43.4	95.0	97.1
Bin_Planktothrix	Unknown	6801343	3689	1761	42.9	82.0	62.6
Bin_Unknown_Cyano	Unknown	3755319	2862	1269	43.6	56.1	24.5
Ribosomal_16S	Unknown	585160	52	13169	46.2	17.3	0.0
UNBINNED_CONTIGS	Unknown	63752146	7447	8469	53.2	100.0	1028.8

Table SI 3.2. Calculated quantities of green (530-570 nm) and red (620-670 nm) light at 23m depth in the sinkhole in June 2015, October 2015, and July 2016, from hyperspectral casts.

Date	Available green (530-570 nm) light (μmol photons)	Available red (620-670 nm) light (μmol photons)
20150606	35.1	0.07
20160720	14.3	0.04
20160727	17.4	0.08
20151008	25.0	0.06

Table SI 3.3. Light quality and quantity metrics from hyperspectral profiles. Seven hyperspectral casts in MIS arena and 2 open water casts were conducted in 2015-2016, from which we determined *k*-extinction coefficients and available energy (µmol photons m^{-2} s^{-1} nm^{-1}) for each measured wavelength of light between 350-800 nm. This summarizes the *k*-extinction coefficient and energy at depth for the range of green wavelengths (530-570 nm) and the range of red wavelengths (620-670 nm), per cast, per day, and per month.

	k-extinction coefficients (m^{-1})		energy at 23 m (µmol photons nm^{-1})		average daily k-extinction coefficient (m^{-1})		standard deviation		average monthly k-extinction coefficient (m^{-1})		standard deviation	
	Green	Red	Green	Red	Green	Red	Green	Red	Green	Red	Green	Red
20150606 cast 4	0.095	0.345	0.385	0.002	0.097	0.353	0.007	0.024	NA	NA	NA	NA
20150606 cast 3	0.098	0.361	0.382	0.001	0.097	0.353	0.007	0.024	NA	NA	NA	NA
20151008 cast 1	0.103	0.373	0.175	0.000	NA	NA	NA	NA	NA	NA	NA	NA
20160727 cast 2	0.107	0.373	0.426	0.001	0.106	0.373	0.007	0.032	0.109	0.381	0.007	0.034
20160727 cast 1	0.105	0.374	0.474	0.002	0.106	0.373	0.007	0.032	0.109	0.381	0.007	0.034
20160720 cast 2	0.114	0.392	0.388	0.001	0.113	0.389	0.006	0.035	0.109	0.381	0.007	0.034
20160720 cast 1	0.112	0.385	0.413	0.001	0.113	0.389	0.006	0.035	0.109	0.381	0.007	0.034
20150606 cast 2	0.105	0.391	0.669	0.001	0.099	0.383	0.009	0.041	NA	NA	NA	NA
20150606 cast 1	0.093	0.374	0.852	0.002	0.099	0.383	0.009	0.041	NA	NA	NA	NA

Table SI 3.4. The relative abundance of diatom 16S rRNA gene reads in each sample, categorized by sample name and season unit.

Sample Name	**Sample Number**	**Relative abundance of reads (%)**
MIS.0609.01.000	Early summer.1	0.80
MIS.0611.03.000	Early summer.2	0.14
M3-3	Early summer.3	11.77
MIS.0513.01.000	Early summer.4	13.57
MIS.0513.02.000	Early summer.5	16.34
MIS.0513.03.000	Early summer.6	21.43
MIS.0513.04.000	Early summer.7	44.39
MIS.2015.027A-2	Early summer.8	5.12
MIS.2015.027A-3	Early summer.9	6.25
M1-1	Late summer.1	39.12
MIS.0713.03.000	Late summer.2	3.23
MIS.0713.07.000	Late summer.3	20.28
MIS.0713.08.000	Late summer.4	41.24
MIS.0713.09.000	Late summer.5	3.92
MIS.0713.11.000	Late summer.6	1.16
MIS.0714.013_mat	Late summer.7	12.60
MIS.0714.020_mat	Late summer.8	21.53
MIS.0714.024_mat	Late summer.9	16.36
MIS.0714.028.000	Late summer.10	0.42
MIS.2015.060A-2	Late summer.11	2.11
MIS.2015.060B-2	Late summer.12	30.52
MIS.2015.060G	Late summer.13	9.65
MIS.0912.LK1.000a	Autumn.1	1.78
MIS.0912.LK1.000b	Autumn.2	6.07
MIS.0912.LK3.000	Autumn.3	0.42
MIS.0912.LK4.000	Autumn.4	1.33
MIS.0912.LK5.000	Autumn.5	7.17
MIS.0912.LK6.000	Autumn.6	1.31
MIS.0914.003_mat	Autumn.7	16.66
MIS.0914.015.000A	Autumn.8	1.99
MIS.0914.020_mat	Autumn.9	5.18
MIS.0914.021_mat	Autumn.10	3.30
MIS.2015.100A-2	Autumn.11	31.25
MIS.2015.100B-2	Autumn.12	22.32
SeasonUnit	Mean relative abundance of reads (%)	Standard deviation
Early summer	13.31	13.65
Late summer	15.55	14.27
Autumn	8.23	9.89

Table SI 3.5. This and next page: The average and standard deviation of relative abundances of key taxonomic groups by season unit.

Taxon	SeasonUnit	Mean (%)	Standard Deviation (%)
Cyanobacteria;Phormidium	Early summer	33.55	15.94
Cyanobacteria;Phormidium	Late summer	35.55	23.92
Cyanobacteria;Phormidium	Autumn	4.19	3.96
Cyanobacteria;Planktothrix	Early summer	8.05	7.83
Cyanobacteria;Planktothrix	Late summer	1.88	4.36
Cyanobacteria;Planktothrix	Autumn	8.43	11.54
Cyanobacteria;Spirulina	Early summer	0.02	0.03
Cyanobacteria;Spirulina	Late summer	0.04	0.09
Cyanobacteria;Spirulina	Autumn	0.55	1.65
Cyanobacteria;Pseudanabaena	Early summer	0.00	0.01
Cyanobacteria;Pseudanabaena	Late summer	0.01	0.03
Cyanobacteria;Pseudanabaena	Autumn	0.15	0.30
Other Cyanobacteria	Early summer	4.36	2.48
Other Cyanobacteria	Late summer	2.47	1.96
Other Cyanobacteria	Autumn	2.43	1.42
Epsilonproteobacteria S-oxidizers	Early summer	4.81	5.07
Epsilonproteobacteria S-oxidizers	Late summer	2.41	1.93
Epsilonproteobacteria S-oxidizers	Autumn	3.11	1.66
Other Epsilonproteobacteria	Early summer	0.06	0.08
Other Epsilonproteobacteria	Late summer	0.22	0.30
Other Epsilonproteobacteria	Autumn	0.96	1.26
Gammaproteobacteria;Beggiatoa	Early summer	4.78	7.40
Gammaproteobacteria;Beggiatoa	Late summer	5.25	4.11
Gammaproteobacteria;Beggiatoa	Autumn	10.79	12.19
Other Gammaproteobacteria	Early summer	2.27	1.46
Other Gammaproteobacteria	Late summer	2.20	1.89
Other Gammaproteobacteria	Autumn	5.55	2.93
Deltaproteobacteria;Desulfonema	Early summer	0.34	0.45
Deltaproteobacteria;Desulfonema	Late summer	5.95	5.95
Deltaproteobacteria;Desulfonema	Autumn	6.06	4.97
Deltaproteobacteria;Desulfocapsa	Early summer	2.67	1.99
Deltaproteobacteria;Desulfocapsa	Late summer	3.00	2.57
Deltaproteobacteria;Desulfocapsa	Autumn	0.94	0.88
Other Deltaproteobacteria	Early summer	1.85	1.19
Other Deltaproteobacteria	Late summer	3.06	2.53
Other Deltaproteobacteria	Autumn	5.76	2.03
Other Proteobacteria	Early summer	16.44	5.23
Other Proteobacteria	Late summer	9.13	5.27
Other Proteobacteria	Autumn	12.25	6.19
Acidobacteria	Early summer	0.33	0.17
Acidobacteria	Late summer	0.52	0.55
Acidobacteria	Autumn	1.09	0.38
Bacteroidetes	Early summer	14.73	4.72
Bacteroidetes	Late summer	15.52	6.44
Bacteroidetes	Autumn	18.31	6.92
Chlorobi	Early summer	0.13	0.08
Chlorobi	Late summer	0.31	0.39
Chlorobi	Autumn	0.78	0.39
Chloroflexi	Early summer	0.23	0.11
Chloroflexi	Late summer	0.94	1.37
Chloroflexi	Autumn	1.75	0.68
Firmicutes	Early summer	1.16	0.87
Firmicutes	Late summer	3.46	5.56
Firmicutes	Autumn	2.80	1.50
Planctomycetes	Early summer	0.24	0.15

Planctomycetes	Late summer	0.93	1.00
Planctomycetes	Autumn	1.60	0.73
Spirochaetes	Early summer	1.12	1.07
Spirochaetes	Late summer	1.98	1.56
Spirochaetes	Autumn	2.87	1.46
Verrucomicrobia	Early summer	1.42	0.96
Verrucomicrobia	Late summer	2.07	2.66
Verrucomicrobia	Autumn	5.07	2.46

Table SI 3.6. This and next 3 pages: The average and standard deviation of relative abundances of Deltaproteobacterial putative sulfate reducers in each season unit.

Taxon	Season Unit	Mean (%)	Standard deviation (%)
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfarculales;Desulfarculaceae;Desulfarculus	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfarculales;Desulfarculaceae;Desulfarculus	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfarculales;Desulfarculaceae;Desulfarculus	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfarculales;Desulfarculaceae;Other	Autumn	0.06	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfarculales;Desulfarculaceae;Other	Early summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfarculales;Desulfarculaceae;Other	Late summer	0.02	0.05
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfatibacillum	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfatibacillum	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfatibacillum	Late summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfatiferula	Autumn	0.03	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfatiferula	Early summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfatiferula	Late summer	0.01	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacter	Autumn	0.03	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacter	Early summer	0.01	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacter	Late summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacterium	Autumn	0.85	0.52
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacterium	Early summer	0.19	0.15
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacterium	Late summer	0.34	0.40
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacula	Autumn	0.31	0.21
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacula	Early summer	0.03	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfobacula	Late summer	0.12	0.12
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfococcus	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfococcus	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfococcus	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfofaba	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfofaba	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfofaba	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteracea	Autumn	6.06	4.97

e;Desulfonema			
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfonema	Early summer	0.84	1.62
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfonema	Late summer	5.29	5.57
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulforegula	Autumn	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulforegula	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulforegula	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfosarcina	Autumn	0.15	0.18
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfosarcina	Early summer	0.02	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfosarcina	Late summer	0.07	0.19
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfotignum	Autumn	0.01	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfotignum	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Desulfotignum	Late summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Other	Autumn	0.24	0.16
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Other	Early summer	0.03	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobacteraceae;Other	Late summer	0.32	0.83
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfobacterium	Autumn	0.59	0.35
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfobacterium	Early summer	0.22	0.22
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfobacterium	Late summer	0.48	0.56
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfobulbus	Autumn	0.20	0.12
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfobulbus	Early summer	0.14	0.20
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfobulbus	Late summer	0.10	0.09
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfocapsa	Autumn	0.94	0.88
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfocapsa	Early summer	2.40	1.84
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfocapsa	Late summer	2.72	2.38
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfopila	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfopila	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfopila	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulforhopalus	Autumn	0.03	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulforhopalus	Early summer	0.01	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulforhopalus	Late summer	0.02	0.04
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfotalea	Autumn	0.00	0.00

Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfotalea	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfotalea	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfurivibrio	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfurivibrio	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Desulfurivibrio	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Other	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Other	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Desulfobulbaceae;Other	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Nitrospinaceae;Nitrospinaceae	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Nitrospinaceae;Nitrospinaceae	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Nitrospinaceae;Nitrospinaceae	Late summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Other	Autumn	0.06	0.06
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Other	Early summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfobacterales;Other	Late summer	0.02	0.04
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfohalobiaceae;Desulfovermiculus	Autumn	0.03	0.02
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfohalobiaceae;Desulfovermiculus	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfohalobiaceae;Desulfovermiculus	Late summer	0.01	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfomicrobiaceae;Desulfomicrobium	Autumn	0.22	0.19
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfomicrobiaceae;Desulfomicrobium	Early summer	0.21	0.20
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfomicrobiaceae;Desulfomicrobium	Late summer	0.18	0.25
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfonatronaceae;Desulfonatronum	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfonatronaceae;Desulfonatronum	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfonatronaceae;Desulfonatronum	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfovibrionaceae;Desulfovibrio	Autumn	0.10	0.05
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfovibrionaceae;Desulfovibrio	Early summer	0.04	0.06
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfovibrionaceae;Desulfovibrio	Late summer	0.04	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfurellales;Desulfurellaceae;Other	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfurellales;Desulfurellaceae;Other	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfurellales;Desulfurellaceae;Other	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Desulfuromonas	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Desulfuromonas	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Desulfuromonas	Late summer	0.00	0.00

Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Other	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Other	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Other	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Pelobacter	Autumn	0.03	0.05
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Pelobacter	Early summer	0.02	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Desulfuromonadaceae;Pelobacter	Late summer	0.02	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Geobacteraceae;Geobacter	Autumn	0.34	0.49
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Geobacteraceae;Geobacter	Early summer	0.32	0.39
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Geobacteraceae;Geobacter	Late summer	0.11	0.12
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Geobacteraceae;Geothermobacter	Autumn	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Geobacteraceae;Geothermobacter	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Geobacteraceae;Geothermobacter	Late summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Other	Autumn	0.05	0.04
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Other	Early summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Desulfuromonadales;Other	Late summer	0.02	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophaceae;Desulfobacca	Autumn	0.01	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophaceae;Desulfobacca	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophaceae;Desulfobacca	Late summer	0.00	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophaceae;Desulfomonile	Autumn	0.03	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophaceae;Desulfomonile	Early summer	0.01	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophaceae;Desulfomonile	Late summer	0.02	0.03
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophobacteraceae;Desulfovirga	Autumn	0.01	0.01
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophobacteraceae;Desulfovirga	Early summer	0.00	0.00
Bacteria;Proteobacteria;Deltaproteobacteria;Syntrophobacterales;Syntrophobacteraceae;Desulfovirga	Late summer	0.00	0.01

Table SI 3.7. This and next 2 pages: Weighted averages of log_2-normalized abundances of proteins across season unit. Log_2-normalized abundances and spectral abundance of proteins in samples from respective season units were used to generate weighted averages and standard deviations (when a protein was observed in more than 1 sample in each season unit). Proteins found to be significantly differentially abundant between early and late summer are noted with "a", between early summer and autumn with "b", and between late summer and autumn with "c".

Protein	Short name	System	Organism	Early summer	Late summer	Autumn	Significant relationships
3300002024_MIS_10189961	AtpA	Central metabolism	Anaerolinea	-0.74 ± 0.13	0.17 ± 0.2	-1.92 ± 0.25	abc
3300002024_MIS_10898911	GapA	Calvin-Benson-Bassham Cycle	Beggiatoa		-1.87 ± 1.11	1.36 ± 0.21	c
3300002024_MIS_11408701	DsrA	Sulfide oxidation	Beggiatoa			-0.12	
3300002027_MIS_101406783	DsrA	Sulfide oxidation	Beggiatoa		-2.13 ± 1.26	0.37 ± 0.98	
3300002027_MIS_101719261	OMP	Central metabolism	Beggiatoa		-2.48 ± 1.39	1.38 ± 0.39	c
3300002024_MIS_11106572	DsrC	Sulfide oxidation	Beggiatoa			0.42	
3300002027_MIS_100764955	DsrC	Sulfide oxidation	Beggiatoa		-2.67 ± 1.21	-0.97 ± 1.18	
3300002026_MIS_1000010375	AtpD	Central metabolism	Chloroplast genome	-0.01 ± 0.25	-3.37 ± 0.49	-1.37 ± 0.48	ac
3300002026_MIS_1000043450	Rubisco	Carbon fixation	Chloroplast genome	0.21 ± 0.18	-3.21 ± 0	-0.92 ± 0.11	b
3300002027_MIS_100225507	ApsR	Sulfate reduction	Desulfobacteraceae		-0.06 ± 0.95	-1.32 ± 0.18	
3300002024_MIS_11625411	AtpD	Central metabolism	Desulfotalea	-3.75	-2.25 ± 0.92	0.24 ± 0.48	c
3300002024_MIS_10280361	AtpA	Central metabolism	Diatom	-0.01 ± 0.24	-3.19	-1.43 ± 0.41	b
3300002024_MIS_10574951	Actin		Eukaryote	1.09 ± 0.42	-1.77 ± 0.3	-1.49 ± 0.23	ab
3300002027_MIS_100411962	Ubiquitin		Eukaryote	0.01	-3.64 ± 0.85	0 ± 0.3	c
3300002026_MIS_100111357	OMP	Central metabolism	Methylococcaceae		-2.29 ± 0.85	1.93 ± 0.3	c
3300002026_MIS_1001557613	CpeB	Photosynthesis	Oscillatorial	-2.91 ± 1.16	-0.49 ± 0.33	-3.74 ± 0.12	c
3300002024_MIS_11047972	L7/L12	Central metabolism	Phormidium	-1.02 ± 0.9	-0.17 ± 0.14	-2.25 ± 0.32	c
3300002024_MIS_11767773	GroES	Central metabolism	Phormidium	0.43 ± 0.29	0.92 ± 0.29	-2.87 ± 0.2	bc
3300002027_MIS_100694095	GroEL	Central metabolism	Phormidium	-0.62 ± 0.42	0.35 ± 0.31	-2.59 ± 0.19	bc
3300002024_MIS_11788891	SOD	Photosynthesis	Phormidium	0.14 ± 0.25	0.07 ± 0.29		
3300002027_MIS_100354662	TrxA	Photosynthesis	Phormidium	2.95 ± 0.47		-0.11 ± 0.18	b
3300002027_MIS_101999963	TrxA	Photosynthesis	Phormidium	0.35 ± 0.45	0.2 ± 0.3	-3.32 ± 0.48	bc
3300002026_MIS_100342443	PsaB	Photosynthesis	Phormidium	-0.22 ± 0.68	-0.09 ± 0.18	-1.44 ± 0.24	c
3300002026_MIS_1001623010	ApcB	Photosynthesis	Phormidium	0.12 ± 0.36	0.03 ± 0.28	-1.7 ± 0.15	bc

3300002026_MIS_100162309	ApcA	Photosynthesis	Phormidium	1.54 ± 0.54	0.7 ± 0.34	-1.69 ± 0.19	bc
3300002027_MIS_101595453	ApcF	Photosynthesis	Phormidium	-0.62 ± 0.01	0.17 ± 0.3	-1.53	a
3300002027_MIS_101997101	ApcF	Photosynthesis	Phormidium	-2.21	-0.52		
3300002026_MIS_1001623011	ApcC	Photosynthesis	Phormidium	0.42 ± 0.74			
3300002027_MIS_101997102	GS	Central metabolism	Phormidium	-3.22	0.76 ± 0.67		
3300002027_MIS_100694094	GroES	Central metabolism	Phormidium	2.31 ± 0.06			
3300002027_MIS_101905325	Cation transport ATPase		Phormidium		-0.1		
3300002027_MIS_100217005	CpeC	Photosynthesis	Phormidium	-0.49	-0.24 ± 0.5		
3300002027_MIS_100523011	CpcC	Photosynthesis	Phormidium	-2.43	0.48 ± 0.32		
3300002027_MIS_101945691	CpeE	Photosynthesis	Phormidium	-0.83 ± 0.04	0.41 ± 0.15	-2.34 ± 0.29	abc
3300002027_MIS_101945692	CpeD	Photosynthesis	Phormidium	-4.28	0.86 ± 0.41		
3300002027_MIS_100155491	CpcB	Photosynthesis	Phormidium	2.42 ± 0.84	-2.58 ± 1.29	-1.45 ± 0.39	ab
3300002027_MIS_100155492	CpcA	Photosynthesis	Phormidium	0.84 ± 2.03			
3300002027_MIS_100217001	CpeA	Photosynthesis	Phormidium	-2.52 ± 0.55	0.05 ± 0.31	-4.04 ± 2.25	a
3300002027_MIS_100217002	CpeB	Photosynthesis	Phormidium	-1.62 ± 0.48	-0.08 ± 0.09	-1.26 ± 1.4	
3300002027_MIS_100523012	CpcA	Photosynthesis	Phormidium	-1.01 ± 0.17	-0.47 ± 0.55	-3.65 ± 0.38	bc
3300002027_MIS_100523013	CpcB	Photosynthesis	Phormidium	-1.8 ± 0.2	0.33 ± 0.38	-0.94 ± 1.31	a
3300002027_MIS_101905321	CpcB	Photosynthesis	Phormidium	-2.89 ± 0.58	0.1 ± 0.13	-1.43 ± 0.49	a
3300002027_MIS_101906971	CpeA	Photosynthesis	Phormidium	1 ± 0.33	-2.22 ± 1.4	-0.5 ± 0.45	b
3300002027_MIS_101906972	CpeB	Photosynthesis	Phormidium	1.55 ± 0.45	-1.87 ± 1.34	-0.76 ± 0.53	ab
3300002027_MIS_101849361	Cna protein B-type domain	Central metabolism	Phormidium	3.56 ± 0.01	1.33 ± 0.26		a
3300002026_MIS_1001011015	PspE	Sulfur cycling	Phormidium		1.3		
3300002027_MIS_100753761	AtpD	Central metabolism	Planktothrix	-1.99 ± 0.89	0.41 ± 0.49	-2.26 ± 0.15	c
3300002027_MIS_100129452	PsaD	Photosynthesis	Planktothrix	-0.41 ± 1.11	0.45 ± 0.32	-2.89 ± 0.13	c
3300002024_MIS_10075351	CpeE	Photosynthesis	Planktothrix	-0.92 ± 0.27	1.29 ± 0.28	-2.77	a
3300002024_MIS_10075352	CpeD	Photosynthesis	Planktothrix	-1.92	-0.12 ± 0.39	-1.5	
3300002026_MIS_1000534011	GroEL	Central metabolism	Pseudanabaena	-1.94 ± 0.66	0.4 ± 0.34	-3.65 ± 1.02	a
3300002026_MIS_1000265839	PsaA/PsaB	Photosynthesis	Putative chloroplast genome	-0.18 ± 0.06		-1.03 ± 0.55	
3300002026_MIS_	PsaF	Photosynthesis	Putative	-0.09 ±	-4.79 ±	-1.58 ±	ac

1000265837			chloroplast genome	0.31	1.25	0.66	
3300002026_MIS_1000265832	PsaL	Photosynthesis	Putative chloroplast genome	0.03 ± 0.36	-4.31	-0.01 ± 0.2	
3300002024_MIS_10200811	Fba	Carbon fixation	Rhodoferax	-2.7 ± 0.31	0.18 ± 0.5		a
3300002026_MIS_1000340919	GTPase	Central metabolism	SOB			0.39	
3300002026_MIS_1000443213	L14	Central metabolism	SOB			0.47 ± 0.21	
3300002026_MIS_100330993	GS	Central metabolism	SOB			-0.93	
3300002026_MIS_1003449713	GroES	Central metabolism	SOB		-0.38	0.16	
3300002024_MIS_10135141	DsrA	Sulfate reduction	SRB		-0.14 ± 0		
3300002027_MIS_101924972	DsrC	Sulfate reduction	SRB		-0.08 ± 1.58		
3300002026_MIS_1000769710	PecC	Photosynthesis	Unknown cyanobacterium	1.8 ± 1.1	-3.63 ± 1.36	-2.34 ± 0.29	a
3300002026_MIS_1000769711	CpcG	Photosynthesis	Unknown cyanobacterium	0.71 ± 1.31	-4.04 ± 0.65	-2.16 ± 0.35	ac
3300002026_MIS_100076979	PecA	Photosynthesis	Unknown cyanobacterium	0.15			
3300002027_MIS_101689572	CpcA/CpcB	Photosynthesis	Unknown cyanobacterium	0.75 ± 0.5	-2.32 ± 0.95	-0.66 ± 0.15	a
3300002027_MIS_101689573	CpcA/CpcB	Photosynthesis	Unknown cyanobacterium	0.78 ± 1.52	-5.28 ± 1.1	-2.76 ± 0.54	ac
3300002027_MIS_100173913	ApcB	Photosynthesis	Unknown cyanobacterium	-0.62 ± 0.15	0.15 ± 0.46	-2.61 ± 0.64	c
3300002027_MIS_100173914	ApcA	Photosynthesis	Unknown cyanobacterium	-2.03 ± 0.55	0.21 ± 0.67	-3 ± 0.38	ac

Figure SI 3.1. Quantities of light available at 23 m in the MIS arena or in the open water outside of MIS ("open") from hyperspectral casts. The available light (y-axis) is plotted for each wavelength (x-axis) of PAR. Different sampling days are represented with colors, and separate casts from the same day are in solid or dashed lines.

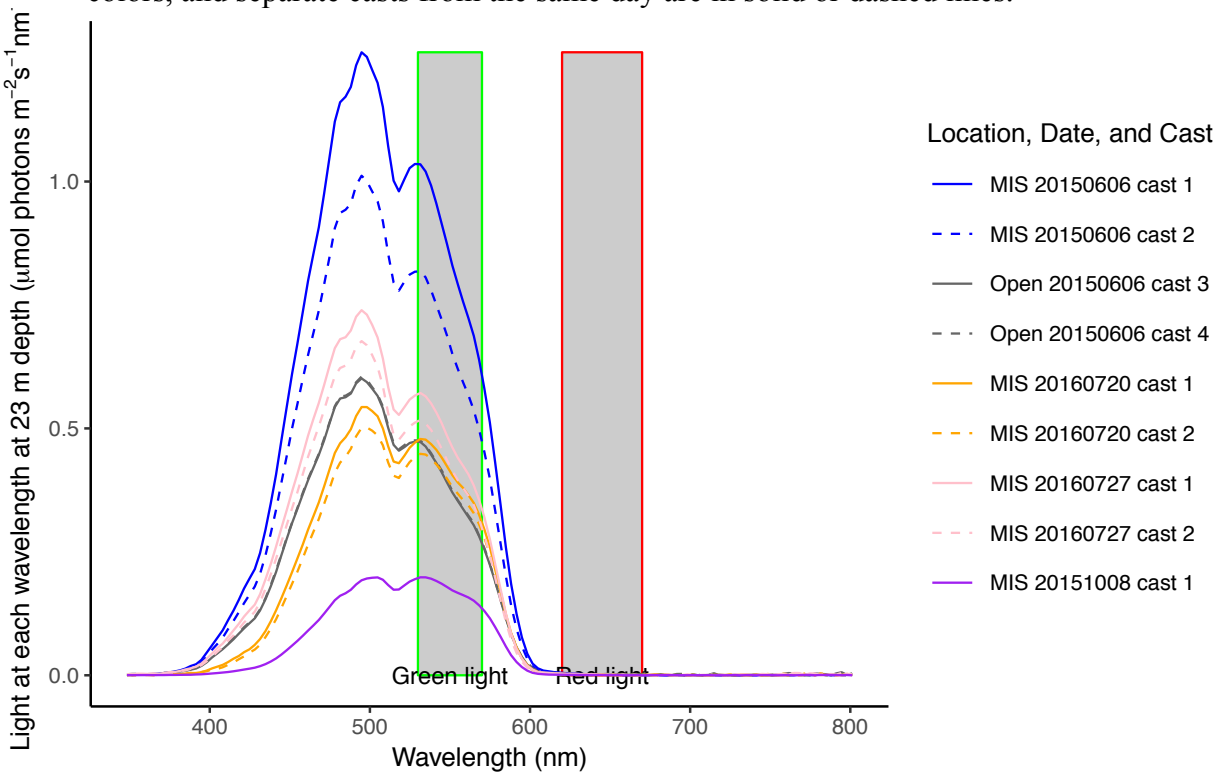


Figure SI 3.2. Calculated *k*-extinction coefficients at 23 m in the MIS arena or away from the arena ("open") from hyperspectral casts. The extinction coefficient (y-axis) is plotted for each wavelength (x-axis) of PAR. Different sampling days are represented with colors, and separate casts from the same day are in solid or dashed lines.

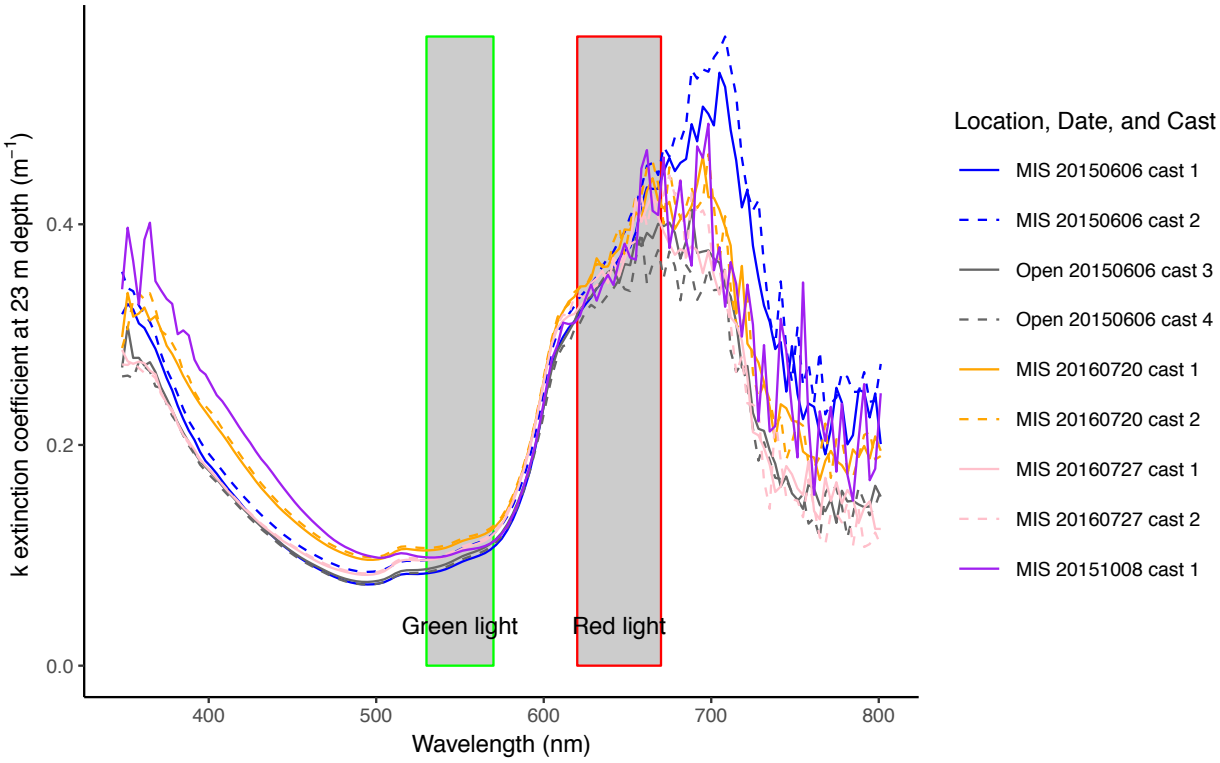


Figure SI 3.3. Specific conductivity measured with hand-held probe or calculated from ion chemistry (IC and ICPMS) in water samples. Units are µS cm^{-1}.

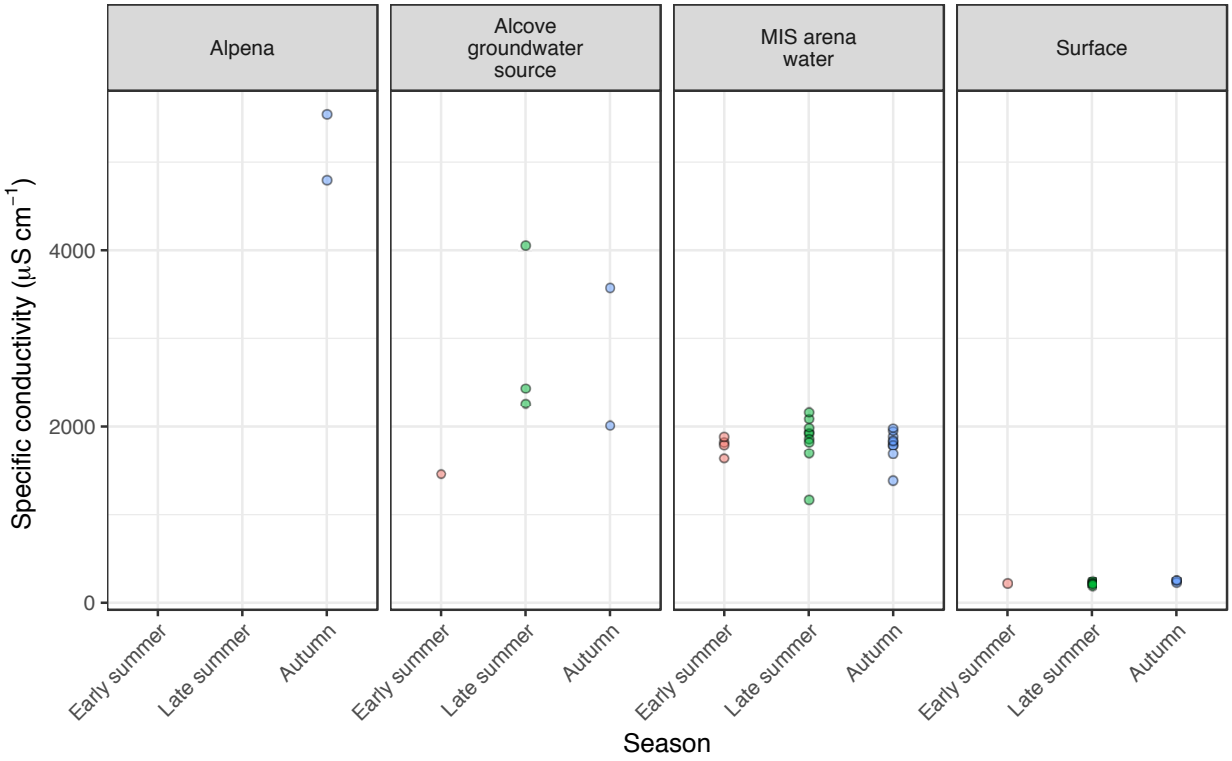


Figure SI 3.4. Deviations from the Great Lakes Mean Water Line, d-excess (difference between measured δD and modeled δD from Great Lakes Mean Water Line,) for water samples at each location per season. The mean for each season is presented as the colored horizontal line, and box and whiskers summarize 25-75th percentiles of observations (colored circles).

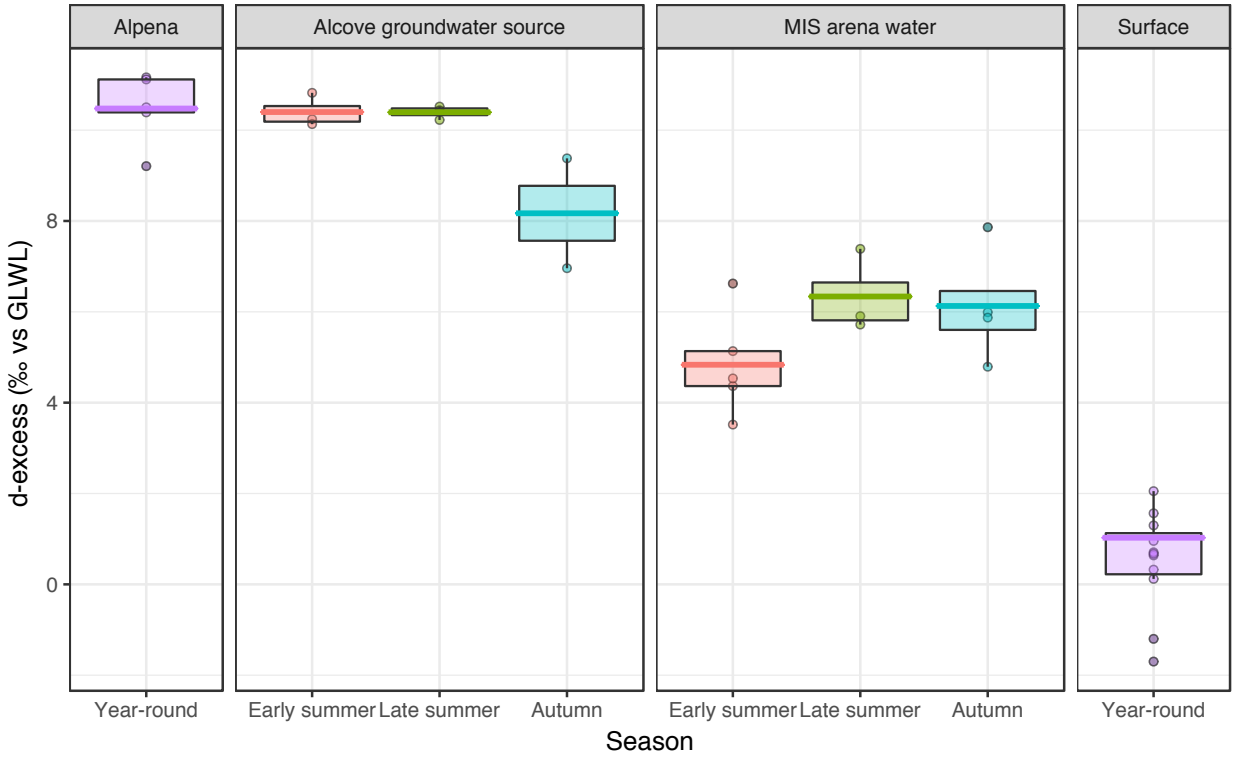


Figure SI 3.5. Measured ion concentrations in water samples compared to predicted ion concentrations based on $\delta^{18}O$-derived linear mixing model. Colored points represent measurements of ions of interest (sulfate, chloride, fluoride, sodium, potassium, magnesium, and calcium), and the horizontal colored line is the mean measured concentration per season unit. For each season, the gray box represents the mean predicted concentrations and standard deviations using the concentrations measured in Alpena fountain and the isotope-derived linear mixing model between surface water and Alpena fountain water.

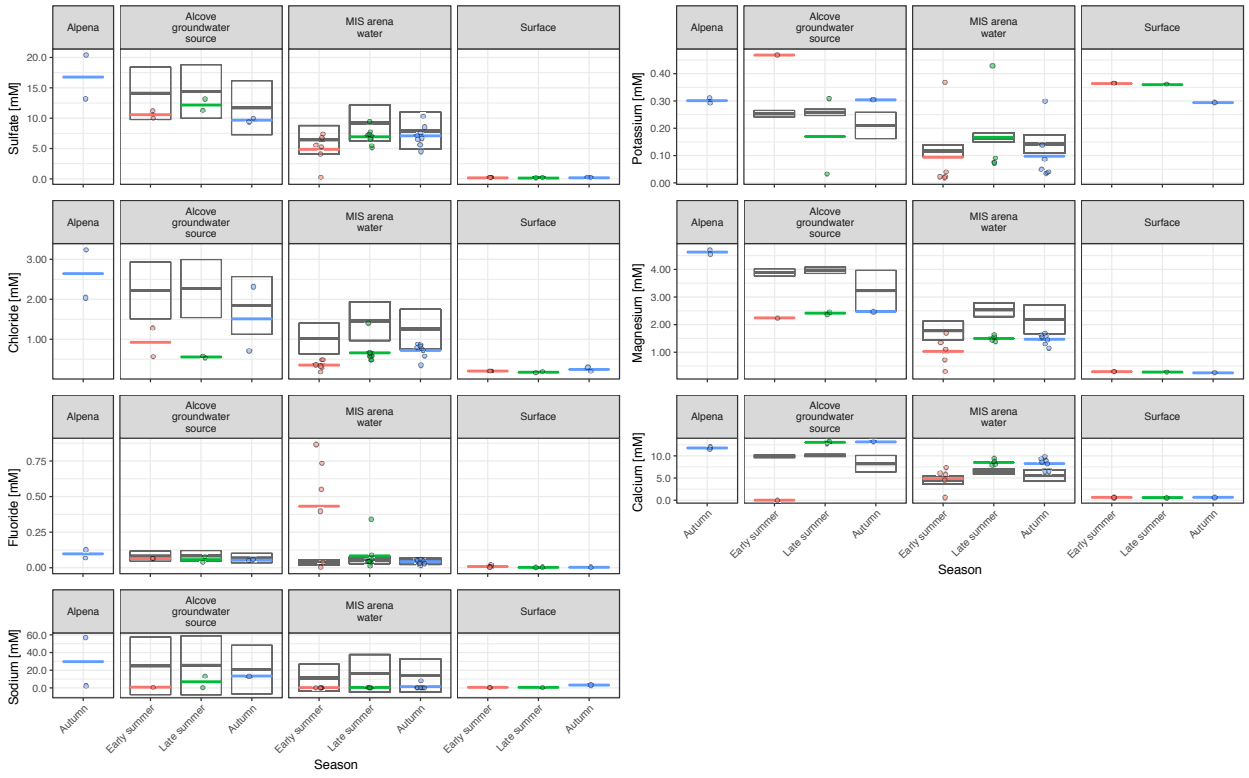


Figure SI 3.6. $\delta^{18}O$ and modeled percentage of groundwater at specific locations in MIS arena in 2016. Samples collected in late summer are marked red, and autumn samples are white.

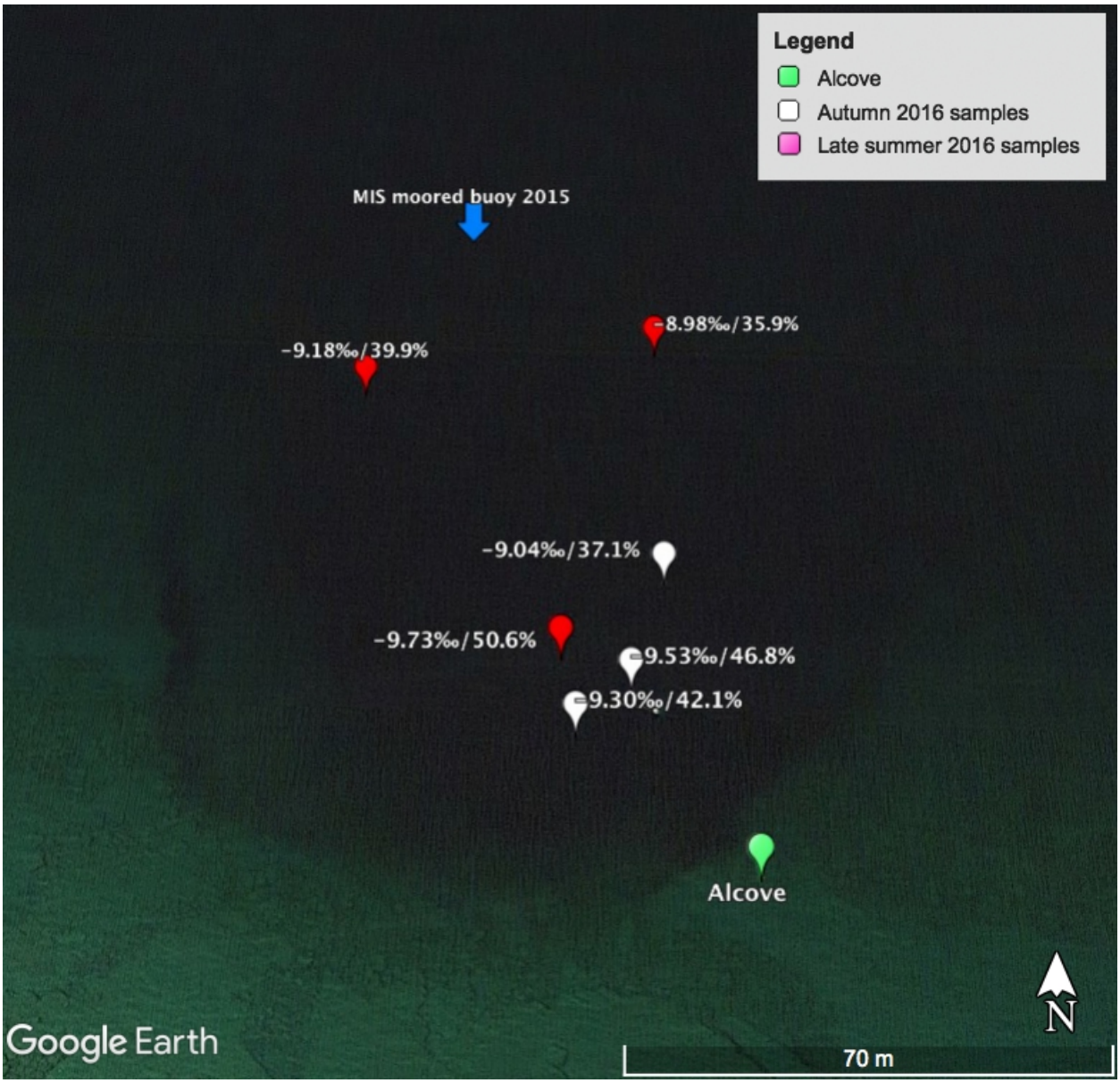


Figure SI 3.7. Relative abundance of relevant bacterial taxonomic groups in samples, grouped by month and year. Relative abundances of key genera were summed and presented, and classes and phyla without those genera are represented.

Figure SI 3.8. Change (Δ) in relative abundance of key bacterial taxa in light-manipulation experiments conducted in summer. Relative abundances at day 0 are normalized to in situ relative abundances from the collection day. The subsequent values are normalized to the relative abundance of the taxa at the prior timepoint. Ribbons show the range of Δ between replicates. Samples at 100% in situ light intensity are plotted in red, those at 50% in situ light intensity in green, and those at 10% in situ light intensity in blue.

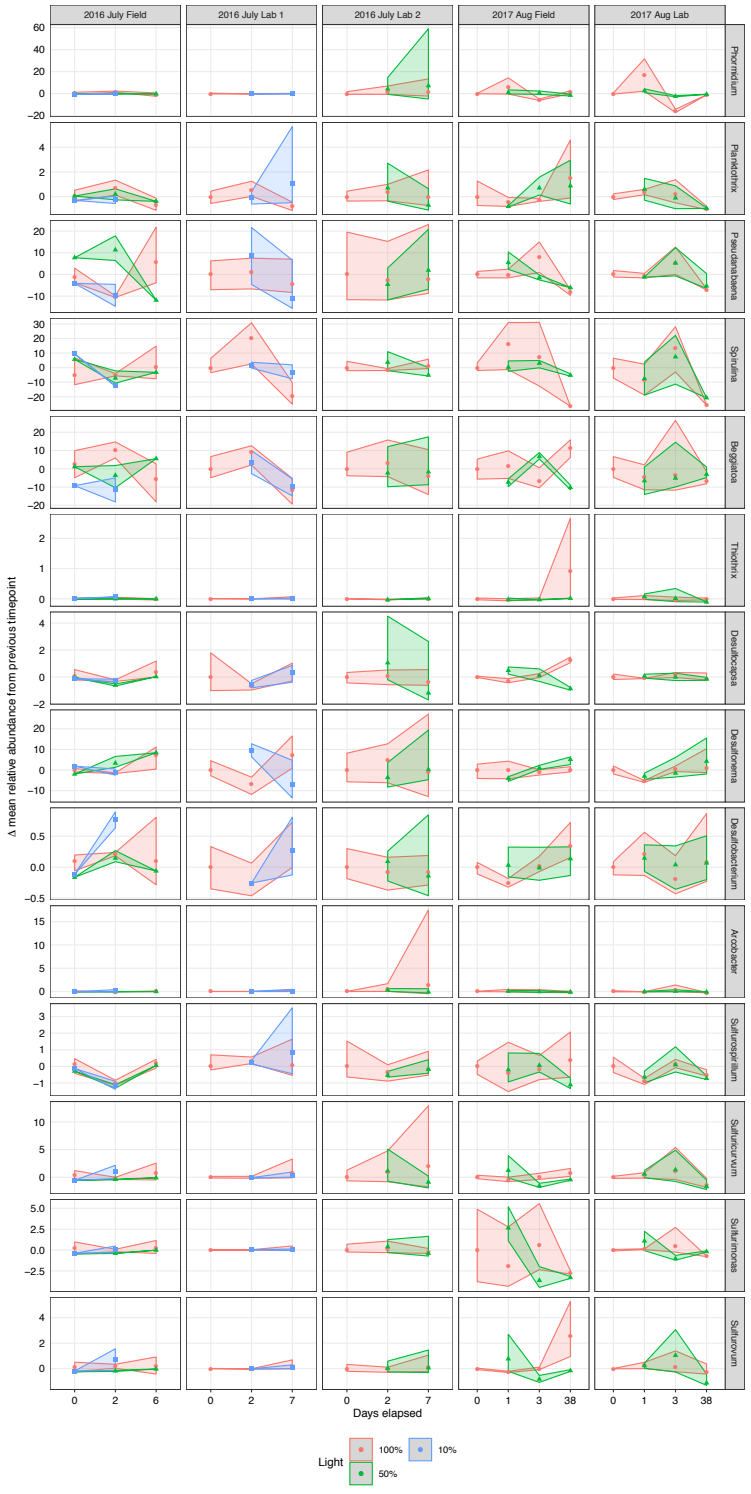


Figure SI 3.9. Change (Δ) in relative abundance of key bacterial taxa in light-manipulation experiments conducted in autumn. Relative abundances at day 0 are normalized to in situ relative abundances from the collection day. The subsequent values are normalized to the relative abundance of the taxa at the prior timepoint. Ribbons show the range of Δ between replicates. Samples at 100% in situ light intensity are plotted in red, those at 50% in situ light intensity in green, and those at 10% in situ light intensity in blue.

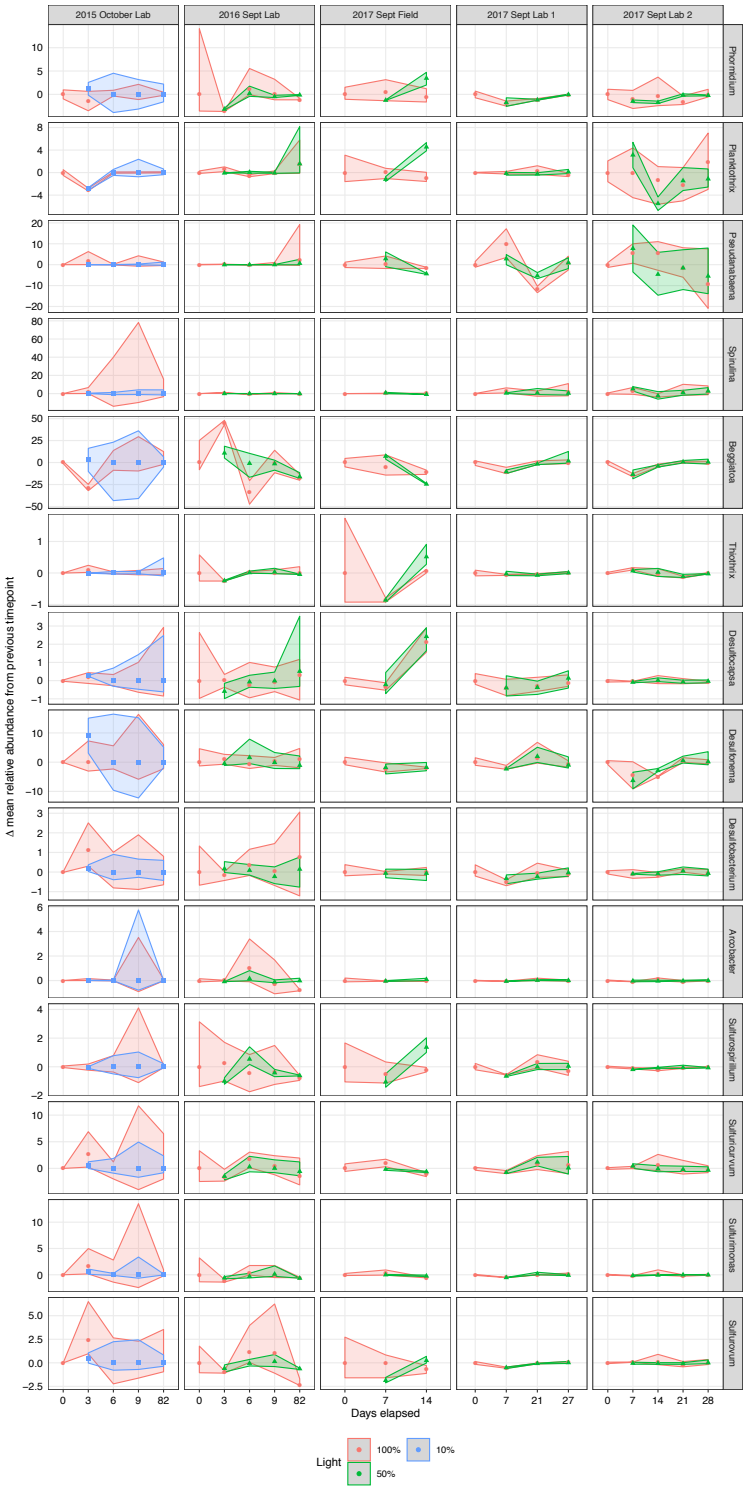


Figure SI 3.10. Correlation network showing significant ($p < 0.001$) relationships in relative abundances of 60 genera. Bray-Curtis distance matrix was calculated from read counts of genera and used to generate a correlation network.

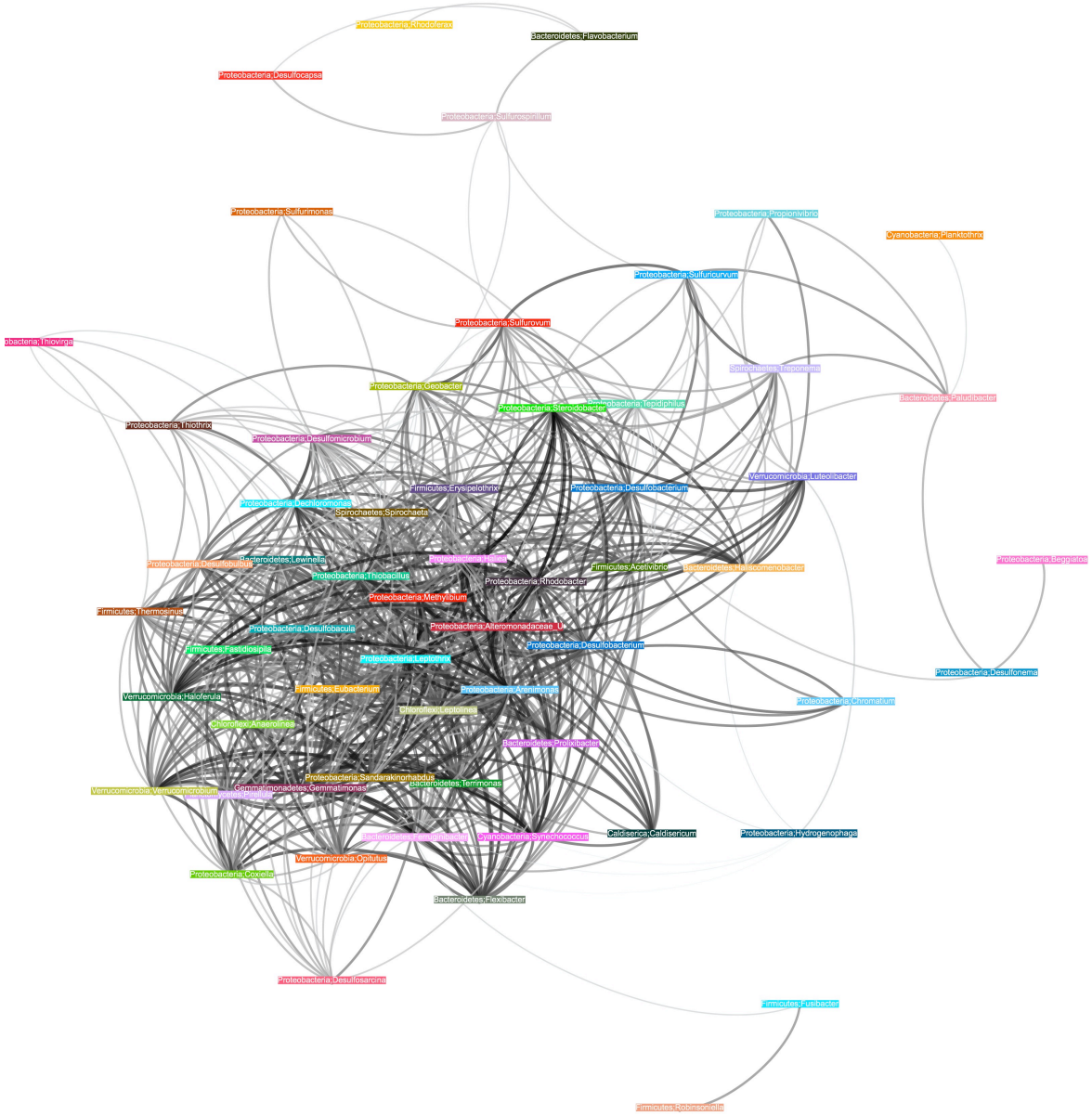
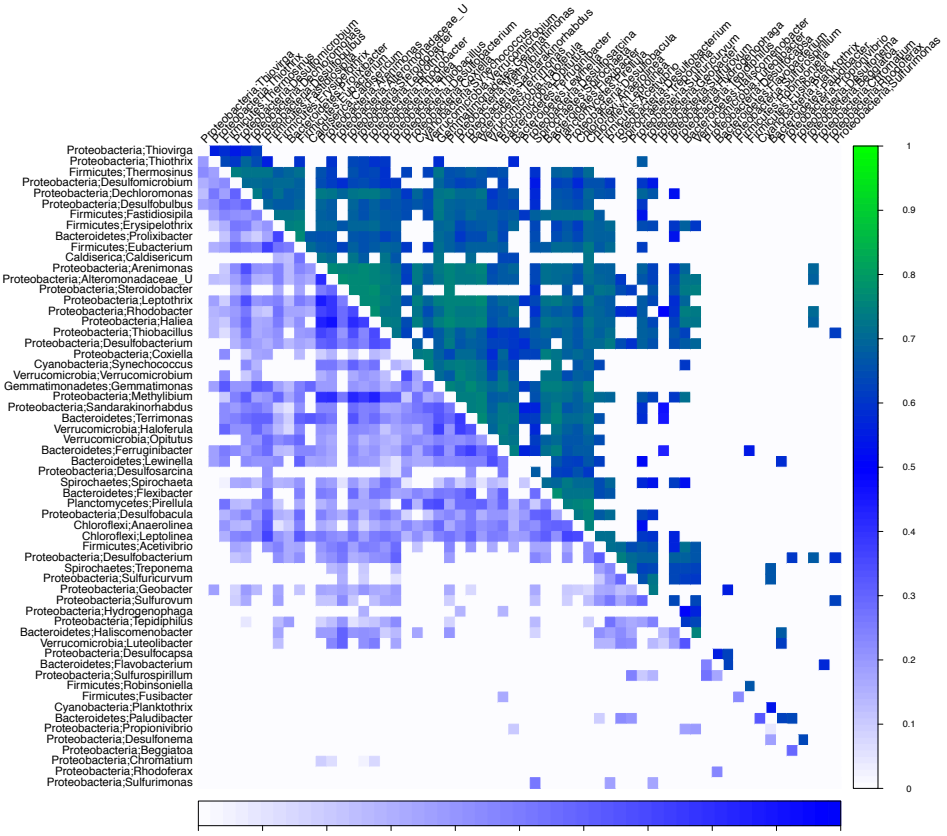


Figure SI 3.11. Heatmap of significant correlations between abundances of bacterial genera. The upper right triangle shows the strength of the average correlation (0 to 1) calculated from 4 distance matrices. The lower left triangle shows the standard deviation of the mean correlation. Genera are clustered by similarity in correlation patterns.



Chapter IV: Functional Overlap of Sulfur Cycling Bacteria Between Microbial Mat Morphotypes

4.1 Abstract

Microbial mats with diverse morphologies are widespread in the geologic record, but inferring their taxonomic composition, metabolisms, and environmental settings remains a challenge. Investigations of extant mats with settings similar to ancient Earth can reveal the links between mat geochemistry, morphology, and microbiology. In this study, we analyzed microbial community structure (16S rRNA genes and metagenomics), function (metagenomics and metaproteomics), and geochemistry (microsensors) in different mat morphotypes (flat purple, finger mat, white mat, and giraffe mat) observed in Middle Island Sinkhole, a submerged sinkhole in which the benthic environment is impacted by low-O_2, sulfur-rich groundwater. Community structure and function were substantially different between the mat types. Flat mats were classified into two groups, characterized by their dominant cyanobacteria (either *Phormidium* or *Spirulina* and *Pseudanabaena*) and spatiotemporal range. *Phormidium* was most abundant in finger mat and select flat mat, and was significantly abundant in their proteomes. In giraffe and white mats, sulfide-oxidizing gammaproteobacteria and epsilonproteobacteria, sulfate-reducing deltaproteobacteria, and cyanobacteria *Pseudanabaena* and *Spirulina* were abundant and metabolically active based on their proteomic abundance. In contrast to these differences in community composition, measurements of H_2S and O_2 vertical profiles and rates of oxygenic and H_2S-based anoxygenic photosynthesis in the cyanobacterial mats show similar

microscale geochemical environments and photosynthesis across mat morphotypes, implying that different taxonomic groups perform similar key ecosystem functions. Metagenomics revealed similar gene content encoding capacity for low-O_2 lifestyles and sulfide-based anoxygenic photosynthesis (AP) between *Phormidium* and *Pseudanabaena,* and *Planktothrix* and *Spirulina*, suggesting niche similarity between these pairs of cyanobacteria in different mats. Various groups of sulfate-reducing bacteria and sulfide-oxidizing bacteria were also differentially abundant and functionally active across different mat types, likely due to differences in O_2 tolerance (sulfate-reducing bacteria) and in N and S metabolisms (sulfide-oxidizing bacteria) inferred from the metagenomes. While the core proteome was largely conserved between mat morphotypes, the shifts in relative abundances of functional guilds of bacteria, and their significant differences in protein abundances in each mat type, point to a subtle shift in the sulfur-cycling mat community. Although the environmental drivers of these shifts remain unclear, our exploration of the connection between geochemistry, taxonomic diversity, and metabolic functions in geologically-relevant mat morphological types highlights the complexity behind ecological interpretations of microbial mats throughout Earth history.

4.2 Introduction

Microbial mats are dynamic ecosystems that showcase the linkage between organisms, metabolisms, and biogeochemical cycling. Environmental changes, such as shifts in light that impact photosynthetic activity in cyanobacteria, often induce concerted metabolic responses in different functional groups of microorganisms that manifest in changes in the mat appearance (Bosak et al., 2009; Bradley et al., 2017; Hamilton et al., 2018; Reyes et al., 2013). Geochemical variability, including changes in oxygen concentration, influence abundance and physical

positioning of sulfide-oxidizing bacteria and sulfate reducers in mats relative to O_2 due to their different ranges of aerotolerance and energetic requirements (Baumgartner et al., 2006; Klatt, Meyer, et al., 2016). Additionally, bacterial sulfate reduction (Visscher et al., 2000) and cyanobacterial photosynthesis (Kraus et al., 2018; Pepe-Ranney et al., 2012) can alter the local geochemical environment and influence lithification and mineral precipitation in mats. Such interactive processes have been implicated in the morphological appearances and preservation of microbial mats in the Archean and Proterozoic (Noffke & Awramik, 2013; Stal, 2012). Because of their consolidation of microbes and metabolisms, and their broad temporal range, microbial mats are important biogeochemical systems throughout Earth history.

The dynamics between microbes and metabolisms in modern mats have been extrapolated to understand metabolisms, microbes, and biogeochemical cycling in mats preserved in the geological record. Mats can hold both body fossils (preserved cells) as well as trace fossils (indicators of microbial activity). Common preserved mat types that are attributable to cyanobacteria include laminae, wrinkled structures, and conical or domal mat (Bosak et al., 2013; Noffke, 2010). A thorough spatiotemporal census of mat morphotype and distribution has been limited to mineralizing and laminating stromatolites and microbialites, whereas microbially-induced sedimentary structures are not as widely surveyed (Noffke & Awramik, 2013). Flat mats are observed throughout the geological record, whereas the oldest wrinkled or ripple-patterned mat is ~2.9 Ga (Noffke et al., 2003), and domal cones have been observed only since the Great Oxidation Event 2.4-2.1 Ga (Bosak et al., 2009). Phototactic growth of cyanobacteria and their relative orientation to the surface has been used to explain layering in fossil mats (Stal, 2012). Upward growth and formation of cones has been attributed to phototaxis as well, though trapped gases underlying the mat may also contribute to their structure (Bosak et

al., 2009; Noffke, 2010). Cyanobacterial mats can leave behind "wrinkle marks" in the sediment with wavy ridges and pits, but trapping of gases such as O_2 or H_2S under the EPS-rich mat, or other bacterial activities, can also generate mottled condensed mat in some areas (Flood et al., 2014; Noffke, 2010; Porada et al., 2008). Because of these diverse mechanisms of formation, mat morphotypes have not been strictly attributed to specific geochemical environments. For example, domal or cone structures in the fossil record have been interpreted to be O_2 oases (Bosak et al., 2009) but they could just as well have been from the upward ballooning of mats by sediment-derived H_2S (Noffke, 2010) or methane (Bosak et al., 2009).

Though we frequently use modern microbial mats as analogs to early Earth ecosystems, a critical gap remains in understanding the phylogenetic and ecological similarities between inhabitants of contemporary and ancient mats. Microbial mats are behind the formation of microbialites, biogenic stromatolites, and microbially induced sedimentary structures, and preserved cell morphologies are used to infer organism groups responsible for these structures. However, in modern microbiology cell shape and size are often not informative with regard to taxonomy, and thus fossilized forms lack information about community membership, metabolisms, and ecological interactions (Bosak et al., 2013; Bradley et al., 2017; Noffke & Awramik, 2013). Despite the extensive phylogenetic diversity in modern microbial taxa (Hug et al., 2016), most major biogeochemical cycles are driven by a limited set of core metabolic pathways and biochemical structures (Falkowski & Godfrey, 2008; Louca et al., 2018). Oxygenic photosynthesis (OP) arose in cyanobacteria (Xiong et al., 2000), and evidence exists for methanogenesis having an archaeal origin (Woese, 1987), but other key metabolisms, such as aerobic respiration (Brochier-Armanet et al., 2009; Soo et al., 2017), anoxygenic photosynthesis (AP) (Raymond et al., 2002; Shih et al., 2017), and sulfide oxidation (Anantharaman et al.,

2014), have been subjects of wide and recent lateral gene transfer between microorganisms across the tree of life (Gogarten & Townsend, 2005). With the widespread distribution of key metabolisms such as AP and sulfide oxidation across multiple lineages, in diverse microbial communities there can exist functional redundancy wherein phylogenetically distinct but coexisting members can perform the same metabolic function, forming functional groups or guilds (Fukami, 2015; Hubbell, 2005; Louca et al., 2018). Under similar environmental constraints, different taxa can occupy the same functional niche in different microbial ecosystems, complicating the relationship between phylogeny and ecology.

Classifying organisms by functional group or guild can be difficult when the same organism performs multiple distinct metabolisms, such as cyanobacteria capable of both OP and AP using sulfide. Metabolically flexible cyanobacteria have been observed in microbial mats in hypersaline lakes (Cohen et al., 1975), hot springs (Garcia Pichel & Castenholz, 1990), sulfidic cave systems (Klatt, Meyer, et al., 2016), and karst sinkholes (Hamilton et al., 2018; Voorhies et al., 2012). Instead of using both photosystems I (PSI) and II (PSII) and oxidizing water as they would for OP, AP cyanobacteria transfer electrons from H_2S to PSI using sulfide quinone reductase (SQR), and do not employ PSII (Arieli et al., 1994; Klatt, de Beer, et al., 2016). AP and sulfide physiologies are not ancestral traits in cyanobacteria (Dick et al., 2018; Miller & Bebout, 2004; Sanchez-Baracaldo et al., 2005), complicating the relationship between functional role and phylogeny in AP cyanobacteria. Such metabolically flexible cyanobacteria are inferred to have inhabited chemically-stratified Proterozoic seas and have been implicated in Earth's delayed oxygenation (Johnston et al., 2009). However, in the modern world, with its oxic surface and scarcity of sulfide in sunlit habitats, AP cyanobacteria are restricted to mats and biofilms in extreme settings (Cohen et al., 1975; Hamilton et al., 2018; Klatt, de Beer, et al., 2016; Klatt,

Meyer, et al., 2016). Exploring the ecology of metabolically flexible cyanobacteria in microbial mats will elucidate their role in modern biogeochemical cycling, and allow better assessments of their impact on oxygenation in Earth history.

Middle Island Sinkhole (MIS) is a submerged sinkhole in Lake Huron, Michigan with microbial mats that experiences seasonal shifts in irradiance and mixing of high-sulfate, low O_2 groundwater with fresh lake water (**Chapter III**). AP-capable cyanobacteria, sulfate-reducing bacteria, and sulfide-oxidizing bacteria form a coherent microbial mat above organic-rich sediment (Biddanda et al., 2009; Rico & Sheldon, 2019; Ruberg et al., 2008; Voorhies et al., 2012). In this study, we describe the appearance, microbial composition and function, and geochemical fluxes of four mat morphotypes observed in MIS from 2015-2017. Domal uplifted "finger" mat, white unpigmented flat mat, flat purple mat, and mottled "giraffe" patterned mat were functionally and geochemically distinct. We observed the bacterial community and functional profile shift in flat purple mat over the measured years, presumably due to bacterial interactions and responses to their changing environment. We explore the relationship between guilds of various metabolic functions (photosynthesis, sulfide oxidation, and sulfate reduction), mat appearance, and impact on geochemical cycling in these discrete mat types.

4.3 Methods and materials

Sample collection. We sampled Middle Island Sinkhole (located at 45° 11.914 N, 83° 19.671 W), a 23.0 m deep sinkhole of approximately 125 m length and 100 m width in Lake Huron (Baskaran et al., 2016; Ruberg et al., 2008). Multiple mat morphotypes were collected: flat purple mat (June, July, and October 2015; June, July and September 2016; and June, August, and September 2017); purple "finger" mats with a domal, uplifted structure (June 2015, July 2016,

and June 2017); giraffe mat, so called for cm-scale darker pigmented borders surrounding lighter pigmented areas (July 2016; at other times we did not observe that mat morphotype); and flat white mat (July 2016 and September 2017). Sampling in July 2016 occurred with respect to a transect grid (**Figure SI 4.1**), with mats collected either by the "D" line, along the "A" line, or the area from "A6" to "C4". Scuba divers from the NOAA Thunder Bay National Marine Sanctuary collected intact mat and sediment cores using 20 x 7 cm clear polycarbonate tubes and rubber stoppers. Cores were kept upright, in the dark and on ice or at 4˚C until sampling within 2 hr of collection. Microbial mat was removed intact from cores, homogenized, stored in 2mL microcentrifuge tubes, and frozen at -80˚C until DNA extraction. Samples were subjected to sequencing of the 16S rRNA gene, metaproteomics, and/or shotgun metagenomics (**Table SI 4.1**).

Micrometer scale geochemical measurements. Microsensors for O_2, H_2S, and pH were prepared and calibrated as previously described (Revsbech and Ward 1983; Jeroschewski et al., 1996; de Beer et al., 1997). Laboratory measurements (2015 and 2016, multiple profiles) focused on flat purple and finger mats, and *in situ* measurements covered flat purple mat near the D line (2016), giraffe mat in the A4-C5 area (2016), and flat white mat (2017). Sensors were positioned and visually inspected to be located at the mat surface to determine the sediment-water interface or '0' position. Measurements in the laboratory were conducted with a micromanipulator and stepper-motor for each probe. In the laboratory, rates of gross oxygenic photosynthesis were calculated using a light-dark shift approach (Revsbech and Jorgensen, 1983), and anoxygenic photosynthesis was resolved by dosing the mat with sulfide during light-dark shifts (Klatt, Haas, et al., 2015). Halogen lamps and blue-green theatre film (Rosco, Stamford, CT) were used to

produce a light regime of 58-79 μmol photons m^{-2} s^{-1} with similar light fields as *in situ*. In situ field measurements were conducted with an automated deployed micromanipulator with two probes for each parameter arranged within a circular frame, measuring at alternate steps in depth (de Beer et al., 2017). Field deployments were for 24 h cycles and took advantage of the natural light field (maximal daily intensity of 30-57 μmol photons m^{-2} s^{-1}). For flux estimates from *in situ* profiles (n = 2 each mat type), due to the probes' destruction of the substrate in their descent, measurements at approximately 2 PM (1400) local time were used.

DNA preparation and sequencing. We extracted up to 0.5 g of wet mat material using a modified version of the MPBio Fast DNA Spin Kit for Soil (MP Biomedical, Santa Anna, CA, USA). In summary, 0.3 g of beads (corresponding to one large bead, seven medium beads, and an equal volume of small beads), sodium phosphate buffer, and MT buffer was used to chemically and mechanically lyse cells, in either the FastPrep instrument for 45 s (samples up to 2013), or horizontal lysis on a vortex mixer for 10 min at speed 7 (2014-2015). After protein precipitation, DNA was cleaned, pelleted, and resuspended in up to 100 μL nuclease-free water. DNA was stored at 4˚C for immediate quantification and -20˚C for long term.

We used the PicoGreen assay (Invitrogen, Carlsbad, CA, USA) to quantitate double stranded DNA. Samples were normalized to between 1-25 ng/μL and submitted to the University of Michigan Host Microbiome Core for Illumina library preparation and sequencing (Kozich et al., 2013; Seekatz et al., 2015). Bacterial primers 515F/806R were used to amplify the 16S rRNA gene v4 region in a reaction mixture consisting of 5 μL of 4 μM equimolar primer set, 0.15 μL of AccuPrime Taq DNA High Fidelity Polymerase, 2 μL of 10x AccuPrime PCR Buffer II (Thermo Fisher Scientific), 11.85 μL of PCR-grade water, and 1-10 μL of DNA template. Thermocycling

was an initial denaturation at 95˚C for 2 min, 30 cycles of 95˚C for 20 s, 55˚C for 15 s, 72˚C for 5 min, and a final extension of 72˚C for 10 min. PCR products were cleaned and normalized using SequalPrep Normalization Plate Kit (Thermo Fisher Scientific), then quantified and pooled equimolarly according to Kapa Biosystems Library qPCR MasterMix (ROX Low) Quantification kit for Illumina platforms. An Agilent Bioanalyzer kit confirmed library size and purity, and the library pool was sequenced on the Illumina MiSeq using a 500 cycle V2 kit with 15% PhiX for diversity.

Fifteen samples representing fingers (n = 3), flat mat (5), a mixture of flat mat and finger (2), white mat (2) and giraffe mat (sampling a ridge, a pit, and homogenized material, n = 3) (**Table SI 4.1**) were shotgun sequenced for metagenomics at the UM DNA Sequencing Core. Briefly, DNA was sheared to an average size of 350bp and prepared for Illumina sequencing using the SMARTer PrepX kits (Clontech). Sequencing was done on three lanes of Illumina HiSeq 4000 2x150bp sequencing, multiplexing 6-7 samples per lane.

16S rRNA gene bioinformatic analysis. Raw pairs of sequencing reads (250 bp) were quality trimmed and merged using 'iu-merge-pairs', which is a program in illumina-utils (available from https://github.com/merenlab/illumina-utils) (Eren et al., 2013), using minimum quality score of 25, minimum overlap of 200 bp, and at points of divergence in the overlap the higher quality basecall was retained. Merged reads with five or fewer mismatches were kept for Minimum Entropy Decomposition v. 2.1 (Eren et al., 2014) using the following parameters: -d 4 -N 3 --min-substantive-abundance 5 -V 3 --relocate-outliers. We used GAST (Huse et al., 2008) to call taxonomy using the curated SILVA database, and confirmed with BLASTN against SILVA 123 (Pruesse et al., 2007) (bacteria and archaea), and PhytoRef (Decelle et al., 2015) (chloroplasts).

mothur v. 1.33 (Schloss et al., 2009) was used to check for chimeras de novo, and putatively-chimeric nodes that did not have taxonomy assigned via SILVA 123 and GAST were removed. We searched for sulfate-reducing genera in the Deltaproteobacteria using a taxonomic search for "sulf" or "thio". The read analysis is outlined here: https://hackmd.io/s/r1CGeQs_G.

Metagenomic analysis. Raw reads were quality controlled using FastQC (Simon Andrews, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), which dereplicated reads, removed sequencing adapters, and trimmed low quality ends. High quality reads were assembled using MEGAHIT v1.1.3 using --mink 21 --maxk 141 --step 12 (Li et al., 2015). Samples were individually assembled, and we used EukRep (West et al., 2018) to extract putatively-prokaryotic contigs. On the contigs not flagged as putatively eukaryotic, we used METABAT (Kang et al., 2015) to generate metagenome assembled genomes (MAGs) using tetranucleotide frequencies and differential coverage. We used EMIRGE (Miller et al., 2011) to assemble reads into 16S rRNA genes and read mapping to match assembled 16S rRNA genes to MAGs. For those genes not found in MAG bins we used bandage (Wick et al., 2015) to link them to binned contigs via their de Brujin graphs. We used Prodigal v2.6.3 (Hyatt et al., 2010) to call genes on the non-eukaryotic contigs. Bins from each sample were compared against each other, and a 'consensus' bin for each MAG was determined using dRep (Olm et al., 2017). Bins were retained that were at least 30% complete, at most 50% contaminated, and had at most 75% strain heterogeneity according to CheckM (Parks et al., 2015), and compared in dRep using ANImf algorithm of at least similarity 99.0%. Representative bins for select taxa, based off the best completion/contamination/strain heterogeneity metrics, as well as individual assemblies were uploaded to IMG/ER for analysis. Assembly statistics are provided in **Table 4.1**.

Quantitative proteomics. Samples of each mat type observed over the years (**Table SI 4.1**) were evaluated with quantitative proteomics. From 0.25 to 5.0g of wet mat material, proteins were extracted, isobarically labeled, and analyzed on LC-MS/MS as described in (Waldbauer et al., 2017). Briefly, proteins were extracted from mat material using a denaturing and reducing extraction buffer (1% SDS, 10% glycerol, 10 mM dithiothreitol, 200 mM Tris, pH 8) at 95˚C for 20 min. Cysteines were alkylated by addition of 40 mM iodoacetamide and incubation in the dark for 30 min. Where not otherwise specified, all solid reagents were dissolved in LC/MS-grade water (Fisher Optima). Proteins were purified by a modified eFASP (enhanced filter-aided sample preparation) protocol (Erde et al., 2014), using Vivacon 500 concentrators (30 kDa nominal cutoff, Sartorius). Proteins were digested with MS- grade trypsin (Thermo Pierce) at 37 °C overnight, and peptides were eluted from the concentrator and dried by vacuum centrifugation. Peptide isobaric labeling is described at protocols.io (dx.doi.org/10.17504/protocols.io.d2i8cd). The C-terminal of peptides was labeled with either ^{18}O or ^{16}O, and the N-terminal of peptides was demethylated with either dideuterated (D_2) formaldehyde (^{16}O-labeled) or unlabeled (H_2) formaldehyde (^{18}O-labeled). A standard composed of all samples was also labeled and run alongside samples to provide quantitative comparison between samples and between separate runs. We used protein sequences predicted from our *de novo* sequence metagenomes to identify proteins, and metagenome-assembled genomic bins as described above to link proteins to specific organisms.

Statistical analyses. We used the R statistical environment (R Core Team, 2015) in RStudio (RStudio Team, 2014) to analyze nodes and proteins. We used Morisita-Horn metric to calculate

a distance matrix on Hellinger-transformed bacterial relative abundances, as input for nonmetric multidimensional scaling with autotransformation = TRUE. We used 'amova' (Excoffier et al., 1992) in the R package 'pegas' (Paradis, 2010) for testing significant difference in bacterial community structure between mat morphotypes. We use the relative abundances of cyanobacterial 16S rRNA genes to sub-classify flat mat samples. To calculate the phylogenetic relatedness of the bacterial community between mat morphotypes, we used the function 'ses.mntd' in package 'picante' (Kembel et al 2010). We aligned the top 1000 most abundant bacterial sequences using clustal-omega (Sievers et al., 2011), and made a phylogenetic tree in RAxML-ng v. 0.8.0 (available at https://github.com/amkozlov/raxml-ng) (Stamatakis, 2014) using the GTR+G4 model of evolution and bootstrapped 1000 times. The phylogenetic tree and the abundance matrix of the top 1000 nodes were input to 'ses.mntd' to generate the nearest-taxon-index (NTI). Kruskal-Wallis testing ($p < 0.05$) determined if NTI varied by mat morphotype.

For evaluation of differential abundance of proteins between mat morphotype (between flat mat subtypes, and between fingers, giraffe mat, white mat, and flat mat subtypes), we used Kruskal-Wallis testing ($p < 0.05$) of groups of proteins. We calculated the weighted mean and weighted standard deviation of the $\log_2$-normalized abundance ratios of samples taken from the same mat type. Plotting of heatmaps and bubble charts was done in 'ggplot2' (Wickham, 2009). To compare clustering of samples by proteomics profiles to clustering by 16S rRNA gene abundances, we calculated a Bray-Curtis distance matrix in the 'vegan' package (Oksanen et al., 2015) on the abundance of cyanobacterial nodes of samples that were also evaluated with proteomics, and a Gower distance matrix on the normalized $\log_2$ abundances of proteins with more than one observation, as inputs for cladograms. Ordering of samples in each cladogram was

compared to evaluate the relationship of mat morphotype on clustering. The package 'dendextend' (Galili, 2015) was used to plot cladograms.

4.4 Results

Mats grouped into four broad morphotypes.

We categorized mat samples based on their morphological appearance *in situ*. Flat purple mats were the most often observed and most widely distributed morphotype throughout the sampling period, being seen in every sampling trip from 2015-2017 (**Figure 4.1**-A, B). They were characterized by their purple/brown pigmentation, and ranged in thickness from a thin < 1 mm to gelatinous 5 mm. Occasionally, this purple layer could be peeled away from a tan or grey underlying layer above the sediment. Conical finger structures (**Figure 4.1**-C) were purple mat that protruded from the sediment, putatively due to buoyant methane and sulfide gas diffusing from sediments below (Voorhies et al., 2012). They were most often observed in early and late summer months, were between 2-30 cm in height, and scattered throughout the MIS area. Giraffe mat (**Figure 4.1**-D) was characterized by cm-long and thick, dark purple pigmented borders surrounding lighter purple pigmented middles. From visual observation, there was little to no elevation change between the darker borders and lighter areas. Of all the mat morphotypes, giraffe was most limited in its spatiotemporal range, being observed only in July 2016 in the A-C area between transects 3-6 (**Figure SI 4.1**). White mat (**Figure 4.1**-E) was more often observed later in the year. It had a similar range of thickness to purple mat, but lacked the purple surface, and infrequently had a thin (< 1 mm) purple or greenish-yellow layer underneath.

Fluxes of H_2S and O_2 vary by mat type.

Microsensor measurements of both *in situ* and *ex situ* mats determined that sulfide fluxes from sediments into the mat varied only subtly across morphotypes (**Table 4.2**). Total *in situ* sulfide (S_{tot}) fluxes from the sediment into giraffe and purple mat were of similar magnitude (0.26 – 0.55 μmol S_{tot} m^{-2} s^{-1}). The range observed in white mat (0.07 to 0.91) was as large as the variability in fingers (0.06 to 0.43) and flat mat measured *ex situ* (0.41 to 0.92). Depth integrated gross rates of oxygenic photosynthesis were 0 in white mat, whereas it was of similar magnitude in other mats measured *in situ* and *ex situ* (1.26 – 2.86 μmol O_2 m^{-2} s^{-1}). A higher flux of O_2 from photosynthesis in the mat escapes to the water column in giraffe mat (1.19 to 1.66 μmol O_2 m^{-2} s^{-1}) and fingers (1.29 to 1.59, average 1.42) than in *ex situ* flat mat (1.39 to 2.86, average 1.34 μmol O_2 m^{-2} s^{-1}) and *in situ* flat mat (0.94 to 1.51). *Ex situ* measurements documented high rates of H_2S-based anoxygenic photosynthesis in flat mat (0.93 to 2.49 μmol S_{tot} m^{-2} s^{-1}) compared to fingers (0.09 to 0.21). Because the electron difference between OP and AP is 2 electrons (Klatt, de Beer, et al., 2016), the potential rate of CO_2 fixation via OP in flat mat is 2.3 to 3 times higher than via AP. The much lower measured rates of AP in fingers leads to a potential rate of CO_2 fixation via OP that is 12 to 35 times higher than via AP. We were unable to measure anoxygenic photosynthesis *in situ* due to the requirement of finely controlling light in order to measure it (Klatt, de Beer, et al., 2016).

Distinct bacterial communities characterize mat morphotypes.

The relative abundances of different cyanobacteria, epsilonproteobacterial and gammaproteobacterial sulfide oxidizers ("eSO" and "gSO", respectively), and deltaproteobacterial sulfate reducers ("SRB") varied by mat type (AMOVA, $p < 0.01$) (**Figure 4.2**). *Phormidium* was the most abundant cyanobacterial group in fingers, whereas *Spirulina* and

Pseudanabaena were the most abundant cyanobacteria in giraffe and white mat. *Planktothrix* and other cyanobacteria were typically 10% or less of the community. Of the sulfide-oxidizing bacteria, *Beggiatoa* were dominant in most samples, though in select white mat and flat mat samples, eSO were more abundant. *Desulfonema* and *Desulfocapsa* were the most abundant putative SRB in the mats, and did not vary in abundance between mat type. Members of Bacteroidetes, Chloroflexi, Spirochaetes, Verrucomicrobia, Chlorobi, and Planctomycetes were altogether generally 25% or less, but in some instances up to 50%, of the bacterial community. Giraffe mat and white mat were relatively cohesive in community structure compared to flat and finger mat (**Figure SI 4.2**).

Within flat mat (**Figure 4.3, Figure SI 4.3**), further molecular characterization revealed two subtypes with different cyanobacterial community compositions and spatiotemporal ranges. The relative abundances of the dominant cyanobacteria (*Phormidium*, *Spirulina* and *Pseudanabaena*) and eSO changed over the years. Samples collected prior to July 2016 had a higher relative abundance of *Phormidium* compared to the other groups, and the overall community composition of these samples was similar, with all grouping together in one clade (**Figure 4.3**). Samples from 2016 were more variable in their relative abundances of cyanobacterial community members. Generally, in July 2016 the samples collected near the A5-C5 transects had more *Spirulina* and *Pseudanabaena* (**Figure 4.2**, the cluster of samples from MIS.2016.60D to MIS.2016.278, all giraffe samples, and white samples MIS.2016.205 to MIS.2016.280), whereas those with more *Phormidium* were located near the D transect (sample cluster MIS.2016.062D to MIS.2016.Y.00.C, **Figure SI 4.1**). Autumn 2016 samples were more similar to the seasonal average and had higher abundances of *Phormidium* (**Figure 4.3**) (**Chapter III**). Samples from 2017 were more stable in composition compared to 2016, but

grouped primarily with *Pseudanabaena* and *Spirulina*-rich samples (**Figure 4.3**), and within the evaluated bacterial groups, are typically <65% *Phormidium*, <50% ESOB, and >50% *Spirulina* and *Pseudanabaena* (**Figure SI 4.3**). Fingers were an exception to this trend; *Phormidium* was enriched in the fingers of 2015-2016 as well as the surface of finger mat in 2017 (MIS.2017.036B). However, *Spirulina* and *Pseudanabaena* formed a large part of the community in the layers underlying the surface of the finger (MIS.2017.037B, MIS.2017.038B).

Because of the variability in cyanobacterial community structure in flat mat (**Figure 4.2**), we used the cyanobacterial 16S rRNA gene profiles to classify flat mat into either flat *Phormidium*-dominated mat ("flat-P mat") or flat *Spirulina* and *Pseudanabaena*-dominated mat ("flat-SP mat"). This categorization separated flat mat samples into two clades (**Figure 4.3**). Twenty-two samples had a higher abundance of *Phormidium* than of *Spirulina* and *Pseudanabaena*, and thus were flat-P mat, whereas 36 samples had a higher abundance of *Spirulina* and *Pseudanabaena* (>45%) than *Phormidium* and were considered flat-SP mat. Bacterial richness varied in the mat morphotypes. *Pseudanabaena* was the most abundant cyanobacterial genus in one sample, MIS.2017.124B, but its relatively low abundances of *Spirulina*, *Phormidium*, and *Planktothrix* compared to other cyanobacteria grouped it with flat-P samples (**Figure 4.3**). However, for additional analyses it is considered a member of flat-SP mats. Flat mat samples with a higher abundance of *Phormidium* have the greatest diversity of bacteria (2957 ± 960 nodes). Mat samples with more *Pseudanabaena* and *Spirulina*, which included flat mat, giraffe mat, and white mat, had intermediate richness (2279 ± 986 nodes, 2434 ± 1014 nodes, and 2516 ± 1025 nodes, respectively). Finger mats had the lowest observed richness, of 1200 ± 388 nodes.

Phylogenetic evaluation of the bacterial communities in the different mat morphotypes was determined using nearest-taxon-index (NTI) (**Figure SI 4.4**). Organisms that are more phylogenetically similar are more likely to occupy the same functional niche, so a high NTI (>2.0) indicates the observed taxa are more closely related than a randomly assembled community, and suggests environmental processes have selected for phylogenetically conserved traits ("niche-based assembly"). In contrast, a low NTI (< -2.0) occurs for communities in which the observed taxa are more distantly related than by chance, and corresponds to environmental filtering for divergent traits (Kembel, 2009). An intermediate NTI value (-2 to 2) suggests that neutral assembly processes are more responsible for the assembly of the community. A mean NTI of the communities that is < 0 indicates more overdispersion than average, or > 0 more clustering than average (Kembel, 2009). In our analyses of mat types, the mean NTI values for each group were not significantly different from each other. The average NTI for giraffe mat and white mat were 0.05 ± 0.16 and 0.12 ± 0.36 respectively, suggesting stochastic processes are behind community assembly. Flat-SP mat (mean NTI = 0.19 ± 0.32), flat-P mat (0.22 ± 0.32) and finger mat (0.37 ± 0.47) have slightly positive NTI, suggesting more clustering in the community than average.

Evidence for differential oxygen and sulfur cycling in metagenome-assembled-genomic bins.

Each metagenome sequencing dataset had between 96.6M and 167M reads. The fifteen metagenomes generated over 12Gbp of assemblies that had a total of 1.93M contigs representing bacteria and archaea. We extracted between 9 and 115 metagenome-assembled-genomic bins (MAGs) from each metagenome, and using dRep to dereplicate genomically-similar MAGs from

different samples, constructed 377 clusters of MAGs (**Table 4.1**). Based on CheckM and EMIRGE results (as well as BLAST queries of known AP cyanobacterial genes of interest (*psbA*, *sqr, nifHDK*) for cyanobacteria) we identified clusters of *Phormidium* bins (6 genomes), *Spirulina* bins (10 genomes), *Pseudanabaena* bins (10 genomes), *Planktothrix* bins (4 genomes), *Desulfonema* bins (5 genomes), *Desulfobacteraceae* bins (11 genomes), a gammaproteobacterial cluster (3 genomes, referred to as Gamma_bin_8), a separate gammaproteobacterial cluster (10 genomes, referred to as Gamma_bin_15), an epsilonproteobacterial cluster with two genomes recovered from white mats (referred to as Campbylobacterales_bin_11), and individual clusters of epsilonproteobacteria named for their deepest phylogenetic resolution (Helicobacteraceae_bin_66, *Sulfurimonas*_bin_83, Campylobacterales_bin_108, Campylobacterales_bin_21, *Sulfurospirillum*_bin_34, *Sulfurospirillum*_bin_16, *Sulfurimonas*_bin_40) (**Table 4.4, Figure 4.4**). Only the bins of *Phormidium*, *Spirulina*, and *Desulfonema* recruited EMIRGE-assembled 16S rRNA genes that confirmed their taxonomy.

The representative *Phormidium* MAG possesses two *sqr* genes that classify as type I (found in AP cyanobacteria) and II (for sulfide detoxification in bacteria) (Marcia et al., 2010) (**Figure SI 4.5**). It also has groups 2, 3, and 4 *psbA* genes for different O_2 levels from anaerobic to standard/high light conditions (Cardona et al., 2015). *Phormidium* has a nitrogenase gene suite *nifHDK* for nitrogen fixation, although it is phylogenetically different from the previously described *Phormidium nifHDK* (Grim & Dick, 2016; Voorhies et al., 2012), and a bidirectional hydrogenase *hox* gene suite to recover hydrogen from nitrogen fixation (Tamagnini et al., 2007). It is equipped with phycocyanin *cpcAB* and phycoerythryin *cpeAB* genes for harvesting light at red and green wavelengths (Tandeau de Marsac & Houmard, 1993), and 68 ATP-binding

cassette transporters for mineral and organic ions including sulfate, molybdate, phosphate, amino acids, metals, vitamin B-12, and lipopolysaccharides.

The representative *Planktothrix* MAG also possesses a *sqr* gene, of the type II variety typically for sulfide detoxification. Like *Phormidium*, *Planktothrix* also has *cpcAB* and *cpeAB*, but does not have the genes for nitrogen fixation (**Figure 4.4**). Additionally, its two *psbA* genes are suited for microaerobic and standard O_2 conditions. It also has *hox* bidirectional hydrogenase and has many of the same transporters as *Phormidium*. The *Spirulina* representative MAG has a type I *sqr* (**Figure SI 4.5**) and a group 3 *psbA*. While it encodes for *cpcAB*, it lacks *cpeAB* phycoerythrin pigment genes. *Pseudanabaena* also lacks *cpeAB* but possesses *cpcAB*. Like *Phormidium*, *Pseudanabaena* has *nifHDK*, groups 2, 3, and 4 *psbA* genes, and two *sqr* genes of types I and VI, the latter more related to green sulfur bacterial *sqr* transcribed at high sulfide levels (Chan et al., 2009). Both *Spirulina* and *Pseudanabanea* have *hox* bidirectional hydrogenases and similar transporters to *Phormidium*.

The SRB bins attributed to *Desulfonema* and *Desulfobacteraceae* are both equipped for dissimilatory sulfate reduction (**Figure 4.4**). The representative *Desulfonema* genome has two terminal oxidases, a cytochrome c oxidase *coxABCD* and a cytochrome bd quinol oxidase *cydAB*, with the former usually operating at higher oxygen levels than that latter (Rabus et al., 2004; Ramel et al., 2013). *Desulfonema* also has a membrane-bound [NiFe]-hydrogenase *hyd* for H_2 oxidation that may be coupled to O_2 reduction with its bd oxidase (Ramel et al., 2013). The representative *Desulfobacteraceae* also has *cydAB* but not the *cox* suite. In addition, *Desulfobacteraceae* has membrane-bound [NiFe]-hydrogenases *hyd* and *hya*, the latter of which may be involved in recovering hydrogenase activity after exposure to oxygen radicals (Tremblay & Lovley, 2012).

The gSO bins both possess *nifHDK* for nitrogen fixation, *narGH* for nitrate reduction, and cbb3 oxidase for O_2 reduction at low O_2 levels (Morris & Schmidt, 2013). Gamma_bin_8 has a complete reverse dissimilatory sulfate reduction pathway (reverse *dsr*, *apr*, *sat*), and *fccAB* to transform H_2S into S^0 (Mussmann et al., 2007). Gamma_bin_8 also has *hox* bidirectional hydrogenase that is analogous to bidirectional hydrogenases used by cyanobacteria to recover H_2 from diazotrophy (Tamagnini et al., 2007) and by purple sulfur bacteria for both dark fermentative and thiosulfate-driven light-dependent H_2 production (Rákhely et al., 2007), but has not been described in chemoautotrophic sulfide oxidizers. Gamma_bin_15 lacks the *apr* gene, but has *sqr* for sulfide oxidation, *psr* gene for sulfur disproportionation, and *hya* and *hyd* hydrogenases. In *Beggiatoa*, these hydrogenases may couple H_2 oxidation with sulfur reduction (Schmidt et al., 1987).

The eSO bins with high completion and low contamination metrics were primarily recovered from white mat samples, though three were recovered from different fingers (*Sulfurospirillum*_bin_34 and *Sulfurospirillum*_bin_16) and a flat mat sample from giraffe mat (*Sulfurimonas*_bin_40) (**Figure 4.4, Table SI 4.2**). Genes for flagellar assembly were observed in *Sulfurospirillum*_bin_16, *Sulfurimonas*_bin_40, *Sulfurospirillum*_bin_34, and *Sulfurimonas*_bin_83, and generally complete reverse TCA cycles were observed in all bins. Additionally, all bins had [NiFe] hydrogenases, transporters for zinc and phospholipids, and most bins with higher genome recovery had additional transporters, for compounds such as molybdate, phosphate, lipopolysaccharides, cobalt, and nickel (**Table SI 4.2**). The cluster referred to as Campylobacterales_bin_11 has two genomes with three *sqr* genes, of types II, IV, and V, for the oxidation of sulfide (**Figure SI 4.5**), and *tst* thiosulfate sulfurtransferase to reduce thiosulfate to sulfite. Neither genes for adenosine 5-phosphate sulfate reductase (APS) nor

sulfide oxidoreductase (SOR), for indirect or direct sulfite oxidation, are observed in the cluster (Takai et al., 2005). The cluster has *nifHDK* nitrogenase genes, cbb3 cytochrome c oxidase (high O_2 affinity) (Sievert et al., 2008) and a low-[O_2] *cydAB* cytochrome bd oxidase. Its group I hydrogenase is a typical membrane-bound uptake hydrogenase, whereas its group IV hydrogenase may be involved in energy and H_2 production (Vignais & Billoud, 2007).

*Sulfurimonas*_bin_83 has *sqr* genes of types II and IV, a complete denitrification pathway, a cbb3 cytochrome c oxidase, a group I [NiFe] hydrogenase and a group III bidirectional hydrogenase (Vignais & Billoud, 2007). Helicobacteraceae_bin_66 has two *sqr* genes of type II, and one each of types IV and V. Additionally, of all the eSO genomes evaluated it is only one of two with any of the Sox cluster, *soxABXY* (*soxD* was not identified) for complete oxidation of reduced sulfur compounds such as thiosulfate to sulfate (Sievert et al., 2008). Helicobacteraceae_bin_66 also has *cydAB*, group I [NiFe] hydrogenase, and transporters and genes for assimilatory nitrate reduction. Campylobacterales_bin_108 has types IV and VI *sqr* genes, genes for assimilatory nitrate reduction as well as *nifHDK*, and cytochrome c oxidase. Campylobacterales_bin_21 has two *sqr* type II genes and *sqr* type IV, cytochrome bc1 complex oxidase, and a group IV [NiFe] hydrogenase. Both of the bins found from finger metagenomes, *Sulfurospirillum*_bin_34 and *Sulfurospirillum*_bin_16, have one *sqr* type II gene, cytochrome cbb3 oxidase, cytochrome bd oxidase, group IV [NiFe] hydrogenase, and several transporters. *Sulfurospirillum*_bin_16 has genes for dissimilatory nitrate reduction, whereas *Sulfurospirillum*_bin_34 has additional groups I and II [NiFe] uptake hydrogenases, and genes involved in nitrite, nitroalkane, hydroxylamine, and nitrile metabolism, but not for transporting nitrate. The bin recovered from flat purple mat near giraffe mat, *Sulfurimonas*_bin_40, has the other *sox* cluster observed in the eSOs (*soxABCYZ*) as well as *sqr* types IV and VI, and *phs* for

thiosulfate disproportionation (Sharrar et al., 2017). It also has cytochrome cbb3 oxidase, *nifHDK*, and groups I and III [NiFe] hydrogenases.

Based off MAG recovery estimates from different samples, 16S rRNA gene levels, and proteomics information, these cyanobacterial, gSO, eSO, and SRB MAG bins are generally enriched in specific mat types (**Table 4.4**). *Pseudanabaena* and *Spirulina* are distributed throughout giraffe mat and flat mat samples that were collected near giraffe mat. Additionally, they were in finger samples from 2016 and 2017, but not from 2015. *Phormidium* and *Planktothrix* were frequently found in finger metagenomes from 2015 and 2017, and sparingly distributed throughout other mat types. Members of the *Desulfobacteraceae* cluster were found primarily in flat purple and white mat near giraffe mat as well as the giraffe metagenomes, one of the 2015 finger metagenomes, and a *Phormidium*-dominated flat mat from July 2016. *Desulfonema* cluster members were obtained from the ridge of the giraffe mat and the homogenized giraffe mat sample, flat mat near giraffe mat, and *Phormidium*-dominated flat mat from July 2016. Gamma_bin_8 members were obtained from giraffe mat and a flat mat sample near giraffe mat, whereas Gamma_bin_15 representatives were recovered from finger metagenomes from 2015 and 2017, and flat white and purple mat. All but two eSO bins were recovered from white or purple mat collected near giraffe mat, and the other two eSO bins were obtained from finger metagenomes from 2015 and 2016.

Mat proteomes do not align with community composition or morphotype.

To assess how the proteome varied as a function of community structure and mat type, we constructed a dendrogram from the Bray-Curtis distance matrix of the cyanobacterial 16S rRNA gene abundances of samples for which we had paired 16S rRNA gene profiles and

proteomic profiles (**Figure 4.5**). We used the previously mentioned abundance cutoff to categorize flat mat samples as either flat-P or flat-SP. Comparison of dendrograms revealed disconnects between community composition and proteome. In the 16S rRNA gene dendrogram (**Figure 4.5-A**), in one broad clade (group 1) are all samples from 2015 (fingers and flat mat), one finger sample from 2017 that was enriched in *Phormidium*, and select flat and white mat from 2016 that had a higher abundance of *Phormidium* than other cyanobacteria. In the other broad clade (group 2) are all giraffe mat samples and all but one sample from 2017, which covered fingers and white mat. Also in this clade are flat mat samples that are more abundant in *Spirulina* and *Pseudanabaena* than *Phormidium*. This grouping by year and mat type was not translated directly to the dendrogram from the proteomics profiles (**Figure 4.5-B**). We observed three broad clades that mixed different mat morphotypes. One clade (group A) had many finger mat samples from 2015 and 2016, and flat mat samples from 2016 that were *Phormidium*-dominant in the 16S rRNA gene profiles. The largest clade in the proteomics profile (group B) encompassed all the white mat samples (collected in 2016 and 2017), all the giraffe mat samples, fingers collected in 2017, and the remaining flat mat samples. A small clade (group C) contained only flat mat samples from July 2016 that generally had more *Spirulina* and *Pseudanabaena*, except for one sample (MIS.2016.120C) whose 16S rRNA gene profile grouped in the *Phormidium*-dominant clade. The discrepancy between cladograms drawn from cyanobacterial community structure and total community functioning is likely due to the metabolic roles of other bacteria in the microbial assemblage being more widely observed in mats, than just cyanobacterial metabolisms, as discussed below.

Functional profiles of flat, white, giraffe, and finger mats are distinct.

Across the samples we observed 2571 unique proteins that were grouped into 53 taxonomic groups. Given the distribution of flat mats across multiple clades in the proteomic dendrogram, for proteomic abundance testing we categorized the proteomics profiles of flat mats by the 16S rRNA gene dendrogram (**Figure 4.5-B**). This yielded a group of 13 flat mat samples that had more *Phormidium* than *Spirulina* and *Pseudanabaena* (flat-P), and a group of 9 flat mat samples that had more *Spirulina* and *Pseudanabaena* (flat-SP). Because of the strong differences in diversity, community composition, and dendrogram clustering, we explored the difference in protein abundance between these two mat types. After significance testing and quality control, comparison of the proteomics between these two flat mat types identified 74 proteins that significantly ($p < 0.05$) differed in abundance, of which 27 belonged to gSO and cyanobacteria (**Figure 4.6**). All *Phormidium* proteins, as well as half the *Planktothrix* proteins, that were significantly differentially abundant were more abundant in the flat-P mats, whereas the remaining significant *Planktothrix* proteins and those belonging to *Pseudanabaena* and *Spirulina* were more abundant in flat-SP mats (**Figure 4.6**). *Phormidium*'s proteins were involved in processes ranging from photosynthesis, carbon fixation, pigment production, and ribosomes; half the proteins each identified in *Planktothrix, Pseudanabaena,* and *Spirulina* were for phycoerythrin alpha subunit CpeA, phycocyanin alpha subunit CpcAB, and allophycocyanin beta subunit ApcB.

Sulfide-oxidizing gammaproteobacterial proteins were generally more abundant in flat-SP mat than in flat-P mat, and reflected processes such as signal transduction, oxidative phosphorylation, sulfur metabolism, and core processes. Interestingly, gSO proteins more abundant in flat-P mat than in flat-SP mat include a [NiFe] hydrogenase HybA and BisC, a protein that is involved in denitrification and DMSO reductase (Hoffmann et al., 1995, Nonaka

et al., 2006) (**Figure 4.6**). A sulfide-oxidizing betaproteobacterium, most closely related to *Sulfuritortus*, had a CcmK that was significantly more abundant in flat-P mat (**Figure 4.6**).

Between flat-P mat, flat-SP mat, finger mat, giraffe mat, and white mat, we identified 77 proteins that differed significantly in their weighted mean $\log_2$-normalized abundances (**Figure 4.7, Figure SI 4.6**). Forty of these proteins belonged to the dominant cyanobacteria (*Phormidium*, *Planktothrix*, *Spirulina,* and *Pseudanabaena*), SRB, or gSO. As with the prior comparison between the different types of flat purple mats, many *Phormidium* proteins related to photosynthesis, carbon fixation, and ribosomes were more abundant in fingers and flat-P mat than in other mat types. However, two *Phormidium* proteins related to ribosomes and growth were as abundant in giraffe mat as in finger and flat-P mat (**Figure 4.7**). Several phycobiliproteins belonging to other cyanobacteria were significantly differentially abundant between mat types. Phycoerythrin CpeA and phycocyanin CpcAB belonging to *Planktothrix* were more abundant in flat-SP mat and finger mat than in other mat types, though CpcAB had high variability in observed abundances (**Figure 4.7**). In addition to being abundant in flat-SP mat, *Pseudanabaena*'s CpcA was also highly abundant in flat-SP, giraffe, and white mat, compared to finger and flat-P mat.

Proteins belonging to gSO and SRB were involved in growth and metabolism, and were more abundant in flat-SP*,* giraffe, and white mats than in the other mat types. This included a gSO rDsrA protein involved in sulfide oxidation (**Figure 4.7**). gSO cytochrome proteins PetA and PetC were most abundant in giraffe and white mats. No proteins belonging to eSO were found to be significantly differentially abundant across the five mat types.

Several proteins that were not assigned to MAGs were significantly differentially abundant between mat morphotypes (**Figure 4.7**). Unidentified cyanobacterial phycobiliproteins

CpeAB from the same contig (Ga0371429_10072573-4) were highly abundant in giraffe mat. Another organism's CpeB (Ga0371430_10473701) was most abundant in flat-SP and giraffe mat, whereas a third cyanobacterium's CpeB (Ga0371429_10798201) was most abundant in flat-P and finger mats. Surprisingly, a fourth cyanobacterium CpeA (Ga0371436_11380134) was more abundant in flat-P and white mats (**Figure 4.7**). A DsrB protein was more abundant in flat mats than in other mat types. Other proteins involved in sulfur metabolism AprB (Ga0371430_10776051), translation (RpsA, FusA), and porins were more abundant in giraffe and white mats.

The core proteome of MIS mats.

Of the 2571 proteins observed in the dataset, 738 were observed in 3 or more mat types and not significantly differentially abundant between them (**Table SI 4.3**). These cosmopolitan proteins are sourced from 27 taxa representing 64 MAG bins, as well as 16 bins without taxonomy. Of the binned proteins, *Phormidium* contributed 131, followed by Gammaproteobacteria (81), Gammaproteobacterium_bin_15 in particular (22), sulfate-reducing members Desulfobacteraceae (43) and Desulfobulbaceae (2), *Spirulina* (23), *Pseudanabaena* (15), and *Planktothrix* (8). Two proteins from epsilonproteobacterial sulfide oxidizers and two from methanotrophic Methylococcaceae are also observed in the core proteome. The core proteins represent 22 different metabolic processes, such as ABC transporters; metabolism of carbon, hydrogen, sulfur, and nitrogen; photosynthesis; transcription, translation, and other genetic information processing; and metabolism of cofactors, vitamins, terpenoids, and polyketides.

Photosynthesis was the most represented process (93 proteins), primarily of allophycocyanin subunits ApcAB, phycocyanin subunits CpcAB, phycoerythrin subunits CpeAB, electron transport proteins PetABCFHJ, photosystem I proteins PsaABCDEFL, and photosystem II proteins PsbABCEHOPUV. Carbon metabolism was the next most represented process (86), and included proteins like Rubisco RbcSL, pyruvate orthophosphate dikinase PpdK, transkelotase TktAB, carbon dioxide concentrating mechanisms CcmKM, and glyceraldehyde 3-phosphate dehydrogenase Gap. Translation-involved proteins (85) included large and small subunit ribosomal proteins, and elongation factors. Genetic information processing proteins (68) were also abundant in the core proteome, such as chaperones GroEL, GroES, and DnaK, cell division protease FtsH, and ATP-dependent Clp protease. Also in the core proteome were ATPase subunits and other proteins in oxidative phosphorylation (40), proteins in signaling and cellular processes (32) like porins and peroxiredoxin, signal transduction proteins (29) such as actins and cytochromes, and sulfur metabolism proteins (23) such as adenylylsulfate reductase Apr subunits, dissimilatory sulfite reductase Dsr proteins, Fcc sulfide dehydrogenase, anaerobic dimethyl sulfoxide reductase subunits Dms, and SoxY. We observed several proteins related to mitigating oxidative stress in the core proteome. A superoxide dismutase (SOD) and a peroxiredoxin belonging to *Phormidium* were abundant in finger and flat mats. *Pseudanabaena* had two peroxiredoxin proteins that were abundant in fingers, giraffe, and flat-SP mats. SODs, peroxiredoxins, and cytochrome c peroxidases belonging to gammaproteobacteria and unknown members were also abundant in many mat types.

4.5 Discussion

Studies of the morphological appearance and metabolisms of modern microbial mats are critical to understanding the mechanisms of formation of microbial structures in the geologic record. In this research, we evaluated the community structure, functional composition, and geochemical cycling in distinct mat morphotypes observed in Middle Island Sinkhole, a low O_2 submerged sinkhole in Lake Huron, Michigan analogous to a Proterozoic microbial system (Rico & Sheldon, 2019). Bacterial community composition differed between flat microbial mat of either purple or white coloration, domal uplifted purple mat (fingers), and mottled purple/brown mat (giraffe mat). Between select mat types, variation in geochemical fluxes indicates room for a biological role in differential biogeochemical cycling. Metabolic differences in cyanobacteria, sulfur-oxidizing bacteria, and sulfate-reducing bacteria observed in metagenomes from the different mat types, and the presence of a cosmopolitan protein core across mat types, suggests functional overlap between members for H_2S, SO_4^{2-}, and O_2 cycling. The interaction of these organisms' biogeography and abundance, their potential and observed differences in H_2S and O_2 metabolism, and observed geochemical fluxes, is not clearly linked to the emergence and persistence of mat morphotypes.

Different cyanobacteria are functional counterparts in mat morphotypes.

Community structure and composition dramatically shifted between mat types observed in 2015 to 2017. Fingers and flat purple mats collected in 2015 had high proportions of *Phormidium* and *Planktothrix* compared to other cyanobacteria, as well as an abundance of *Beggiatoa* and *Desulfonema* members. Purple flat mats collected in 2016 were more heterogeneous, with mats in specific sampling locations having a higher proportion of *Pseudanabaena* and *Spirulina* (near giraffe mat, the A5-C5 transect, **Figure SI 4.1**) than of

Phormidium (near D transect). Flat mat samples collected after 2016 were more widely distributed throughout the MIS area, and reflected an overall shift in the community composition towards more *Pseudanabaena* and *Spirulina* (**Figure 4.2**). Fingers are relative havens for *Phormidium* despite the overall shift in the bacterial community structure. Sampling for white mat occurred primarily in July 2016 and September 2017, however white mat was observed but not sampled in MIS at other times in summer and autumn months.

The shift in dominant cyanobacteria from *Phormidium* and *Planktothrix* to *Pseudanabaena* and *Spirulina* in flat mats and the similarity of key S, O, and N cycling genes in their genomes, suggests functional analogy between cyanobacterial groups. Both *Phormidium* and *Pseudanabaena* have two *sqr* genes of different types, multiple versions of *psbA* genes, and *nifHDK* genes to fix nitrogen, suggesting this pair of cyanobacteria may be equipped to occupy the same functional niche of a versatile AP cyanobacterium. Different versions of PsbA are linked to varying O_2 and light levels. Group 2 anaerobic *psbA*, group 3 microaerobic *psbA*, and group 4 standard/high-light *psbA* (Cardona et al., 2015) would allow *Phormidium* and *Pseudanabaena* to adjust their capacity for O_2 evolution, thus facilitating nitrogen fixation (Aro et al., 1993; Blankenship, 2014; Sicora et al., 2009). Given that sulfide inhibits water oxidation, which occurs in PSII (Miller & Bebout, 2004), group 2 *psbA* may also be involved in sulfide metabolism, such as deactivating PSII when oxidizing sulfide in AP (Becraft et al., 2015; Murray, 2012; Olsen et al., 2015). On the other hand, *Spirulina* and *Planktothrix* do not have *nifHDK* and possess only one *sqr*, and 1-2 versions of *psbA* (either groups 3+4 or only group 4). *Pseudanabaena* and *Spirulina* are primarily abundant and more active in the proteomes of giraffe mat, and white and purple mat collected near giraffe mat in July 2016 and as well as mats sampled in 2017, whereas *Phormidium* and *Planktothrix* are most abundant and active in fingers

and flat mat collected prior to July 2016. The similar rates of gross oxygenic photosynthesis in purple, finger, and giraffe mat are also consistent with the conclusion that the metabolically dominant cyanobacterial members in those mats, namely *Phormidium* and *Pseudanabaena*, may be functionally similar.

While many similarities exist between the cyanobacteria observed between different mat types, a crucial difference between the genomes of dominant cyanobacteria of giraffe and co-located white and purple mats, and the cyanobacteria of fingers and other flat mats, may be in their light harvesting capabilities. The recovered genomes of *Pseudanabaena* and *Spirulina* both lack *cpeAB* genes for phycoerythrin, which capture green light (560 nm maximal absorption) (Bryant, 1982). They both have red-light-harvesting *cpcAB* phycocyanin (620 nm maximal absorption) genes, and *Spirulina* additionally encodes for phycoerythrocyanin, which captures wavelengths between 570-595 nm. Our genomic recoveries of each cyanobacterium were above 90%, which would suggest we effectively sequenced their genomes and thus did not miss *cpeAB* due to lack of depth. Though the two sequenced representative genomes of *Spirulina* cultures do not have *cpeAB* and only one has *pecAB,* the lack of *cpeAB* in *Pseudanabaena* is unexpected, given that four sequenced cultures' genomes code for phycoerythrin. In contrast, both the MIS representative genomes of *Planktothrix* and *Phormidium* encode for *cpcAB* as well as *cpeAB*, which other cultured members of their genera share. These phycobilisome proteins improve light harvesting in cyanobacteria: phycoerythrin or phycoerythrocyanin passes energy to phycocyanin, which transfers energy to allophycocyanin (Bryant, 1982). Only those that synthesize phycoerythrin are capable of complementary chromatic adaptation, which allows organisms to fine tune their pigment contents in response to changes in light intensity (Bryant, 1982). Phycoerythrin may be key for organisms thriving in low light conditions. Phycoerythrin-

synthesizing cyanobacteria have lower light requirements for optimal growth rate than cyanobacteria that synthesize only phycocyanin (Wyman & Fay, 1986). Further, with increasing light, rate of phycoerythrin synthesis compared to rate of phycocyanin synthesis declines (Wyman & Fay, 1986). In *Synechococcus* PCC 6301, the ratio of phycocyanin to allophycocyanin is higher at lower light intensities (Tandeau de Marsac & Houmard, 1993). Given the lack of phycocyanin in their genomes, the success and relative proliferation of *Pseudanabaena* and *Spirulina* in the mats suggests that light is not a limiting condition for their growth. As in marine systems, in the water column of MIS, red light rapidly attenuates with depth, leaving blue-green light most available at 23 m (**Chapter III**). Generally, light levels are highest in summer (**Chapter III**), but in July 2016 when *Pseudanabaena* and *Spirulina* were widely observed in the mats, the average daily maximal light intensity was 25-50% less than observed in 2015 and 2017. It is possible that the low light levels observed in 2016 may have been related to phytoplankton growth and organic matter in the water column. In Lake Huron, green light is generally least available in summer due to growth of surface colonial cyanobacteria (Fahnenstiel & Carrick, 1991), though we did not observe obvious surface phytoplankton blooms at MIS in 2016. We did not see evidence for low light levels impacting growth of *Phormidium*, since their proteins were highly abundant in mats throughout 2016. Given the lower available light in summer 2016 and the lack of genes to harvest green light, the high relative abundances of *Spirulina* and *Pseudanabaena* in the mat seems unrelated to light conditions.

The success of *Pseudanabaena* and *Spirulina* in the MIS microbial mats may be tied to improved fitness over other cyanobacteria, putatively due to more efficient energy use and/or ecological dependence on other community members. *Pseudanabaena* and *Spirulina* genomes were between 96-99% complete, and up to 4.6 Mbp in size and 4828 genes, whereas for

Phormidium and *Planktothrix* we recovered 92-93% representing a considerably larger 5668-5692 genes and 5.5-5.9 Mbp (**Table 4.4**). Even accounting for strain heterogeneity of 15-30% contributing redundancy to the MAG, the genomes of *Phormidium* and *Planktothrix* are larger than those of *Pseudanabaena* and *Spirulina.* The representative genomes of *Pseudanabaena* and *Spirulina* are smaller than available sequenced cultures (4.9-5.6 Mbp), and when using the cultures' coding densities (the number of genes per 1000 bp) to extrapolate the size of our representative MAGs, they should be between 5.0-5.2 Mbp. (A similar application to the representative *Phormidium* and *Planktothrix* genomes would yield expected genome sizes of 5.7-7.2 Mbp.) Increased variability in environmental conditions leads to microbes with larger genomes and a larger number of genes (Bentkowski et al., 2015). In contrast, more limiting environments host organisms with smaller genomes to minimize resource requirements for e.g. replication and protein synthesis (Giovannoni et al., 2014). Ecological interdependence is higher in communities with organisms of smaller genome sizes. In part due to their reduced number of genes, these microbes rely upon interactions with other organisms to acquire nutrients and resources (Anantharaman et al., 2016; Zelezniak et al., 2015). In freshwater systems, free-living organisms with smaller genomes rely upon metabolic compounds from other organisms (Garcia et al., 2018). In contrast, members of particle-associated communities tend to have larger genomes, putatively for greater metabolic flexibility to handle the dynamics in resource availability, and these communities are more diverse (Schmidt, 2018). In MIS, though their genomes were smaller than other cyanobacteria, the number of transporters in the representative *Pseudanabaena* and *Spirulina* cyanobacterial genomes did not differ from *Phormidium* and *Planktothrix* (63-72). Further, a higher abundance of *Pseudanabaena* and *Spirulina* and their smaller genomes characterize samples with less diverse bacterial communities. Due to the

extracellular polymeric substances (EPS) that bind the mat together (Voorhies et al., 2012), as well as the high density of microorganisms in mats (Dick et al., 2018), the microbial mats of MIS may buffer their inhabitants from strong environmental perturbations and nutrient limitations. This cohort of organisms likely facilitates ecological interdependencies for *Pseudanabaena* and *Spirulina* that balance their smaller genome sizes and improve their competitiveness against *Phormidium* and *Planktothrix*.

Functional niche overlap between the cyanobacteria can also influence overall community composition. To understand if environmental pressures are influencing the morphotypes' bacterial communities, we employed nearest-taxon-index (NTI) to evaluate the phylogenetic relatedness of organisms in each community. Abiotic environmental processes that impose selective pressures would filter for phylogenetically-conserved traits and lead to higher phylogenetic relatedness between coexisting taxa than by chance (Kembel, 2009; Stegen et al., 2012). On the other hand, communities that experience ecologically neutral or stochastic processes would host taxa that would not compete with each other and thus do not overlap in niche. Competition between coexisting taxa should drive ecological differentiation, eventually leading to high phylogenetic diversity in the community (Stegen et al., 2012; Violle et al., 2011). As a result, taxa that are more closely related may inhabit the same ecological niche, whereas more distantly related organisms will have more dissimilar niches, reduced competition, and better coexistence (Violle et al., 2011). The mean NTI values for each mat type group were not significantly different from each other, but they were all above 0 (**Figure SI 4.4**), suggesting a greater role for environmental filtering for conserved traits than random processes. Finger mats had the highest mean NTI of 0.37, and also had the lowest observed bacterial richness of the mats. From these characteristics, one potential conclusion is that environmental perturbations in

finger mats are relatively reduced, and there is little competition for niche space. The same cannot be said for flat-*Phormidium* mat, which had the most diverse mats and the next highest mean NTI of 0.22. Samples in which *Spirulina* and *Pseudanabaena* were more abundant than *Phormidium*, which includes many flat mats, and the giraffe and white mats, had mean NTIs that were positive but closer to 0, and intermediate in bacterial richness compared to fingers and other flat mat samples. With apparently neutral processes controlling community assembly and average richness in these mats, the abundance of *Pseudanabaena* and *Spirulina* may be related to reduced stressors in their environment, lack of niche overlap in the mats, and favorable ecological interactions with other microbes.

Differential capacity for sulfur cycling seen in phylogeny of genes and distribution of organisms.

The distribution of *sqr* genes throughout the evaluated eSO bins, and the morphotype-related abundance of these members, point to a linkage between sulfide physiologies and distinct geochemical environments in each mat type. All of the eSO members of white and purple mats collected near giraffe mat have 2 or 3 *sqr* type genes, whereas finger eSO members have only type II *sqr*. Though it is not clear if all the eSO members of white and purple mats are related to *Sulfurimonas*, all of them have type IV *sqr*, which in *Sulfurimonas* members is likely from a common ancestor and not from lateral gene transfer like other *sqr* can be acquired (Gregersen et al., 2011; Han & Perner, 2015). In *Chlorobaculum tepidum*, types IV and VI *sqr* are expressed when grown between 2-4 mM H_2S and above 4 mM respectively (Chan et al., 2009). Sulfide fluxes were generally highest in white mat compared to the other mat types (**Figure 4.5**), making different-affinity *sqr* genes a valuable asset to sulfide response in these members. Type II *sqr*

have a reportedly low affinity for H_2S (K_m ~ millimolar) (Marcia et al., 2010), and in the evaluated eSO are found in four of the white mat MAGs/clusters and the finger MAGs. These type II *sqr* genes phylogenetically group together (**Figure SI 4.5**) in a broader clade that includes cyanobacterial, yeast, and proteobacterial type II *sqr*.

Fingers experienced much lower sulfide flux from the sediment than other mat types; because they are uplifted from the sediment, essentially all sulfide is locally sourced from sulfate reducers in the mat, and is consumed by AP cyanobacteria and SO bacteria. Because of the potentially low [H_2S] observed in fingers, genes encoding for only low-affinity H_2S type II *sqr* in finger eSOs may not be solely for sulfide response. *Synechocystis* sp. PCC 6803 expresses its AP-involved *sqr* to detoxify arsenite (Nagy et al., 2014). Zn and other trace metals are micronutrients and readily incorporate into organic matter. Epsilonproteobacteria have ribosomal genes optimized for low-Zn environments (Chen et al., 2009), and though MIS groundwater is elevated in zinc compared to surrounding and surface Lake Huron water, trace metals such as Zn complex with sulfide in the presence of sulfate reduction and thus are not bioavailable (Huerta-Diaz and Morse, 1992). Zn, Ni, Cu, and Fe are enriched in microbial mats where active sulfate reduction is occurring (Huerta-Diaz et al., 2012; Rico, 2019), and when they are remobilized during sulfide-oxidizing denitrification, type II *sqr* may play a role in heavy metal tolerance (Han & Perner, 2015). All the eSO evaluated have transporters for zinc, and over half have some or all the genes for denitrification, suggesting that in addition to sulfide oxidation, these *sqr* type II genes may be involved in metal metabolism.

The distribution of the dominant cyanobacteria and their genes among the different mats also suggest potential links between geochemical conditions and mat morphotypes. All the evaluated cyanobacteria have 1-2 *sqr* genes. *Phormidium*, *Spirulina,* and *Pseudanabaena* have

the type I *sqr* identified in culture-based studies of AP, but *Phormidium* and *Pseudanabaena* have an additional *sqr*. However, the second *sqr* in *Pseudanabaena* is of type VI, previously found in a few cultures (Grim & Dick, 2016) and only recently implicated in AP in cyanobacteria (Hamilton et al., 2018), whereas *Phormidium*'s second *sqr* and *Planktothrix*'s sole *sqr* is of type II with low H_2S affinity and typically for sulfide detoxification (Marcia et al., 2010). The phylogeny of *sqr* in cyanobacteria points to horizontal gene transfer from other lineages, such as green sulfur bacteria (Gregersen et al., 2011; Sanchez-Baracaldo et al., 2005). Given that cyanobacteria and gSO share many genes through HGT (Flood et al., 2014), it is possible that these additional *sqr* genes in *Phormidium* and *Pseudanabaena* may have been acquired from bacterial neighbors such as gammaproteobacteria, to address varying sulfide and/or metal availability in their mats.

Differences in the dominant putative sulfide-oxidizing group between mat types may be linked to geochemical environments. In some mats, namely select white and flat mats from 2016, eSO members were dominant, whereas in a majority of other mats, *Beggiatoa* and other gammaproteobacteria were more abundant. Shifts between gammaproteobacteria-dominated and epsilonproteobacteria-dominated community composition in other systems has been related to concentrations of H_2S and electron donor. In H_2S-rich shallow hydrothermal vents, eSO members such as *Sulfurovum* were key pioneer organisms in the mat, and gSO members such as *Thiomicrospira* became more abundant in more established communities (Patwardhan et al., 2018). In Frassasi Spring streams, gammaproteobacteria were more abundant than epsilonproteobacteria in biofilms because the O_2 concentrations were higher than what eSO prefer (Macalady et al., 2006). Epsilonproteobacteria can couple denitrification with oxidation of reduced sulfur compounds such as sulfide and thiosulfate (Brettar et al., 2006; Bruckner et al.,

2012). Given that many of the evaluated eSO genomic bins were recovered from two white mats in which cyanobacteria were not dominant, whereas the gammaproteobacterial clusters sourced most of their members from flat purple or giraffe mat, it is likely that concentrations of H_2S and/or O_2 determine the dominant sulfur oxidizer. In the evaluated white mats, epsilonproteobacteria that were not dependent upon cyanobacterially-produced O_2 are predominant sulfur oxidizers, whereas in the cyanobacterially-dominated mats gammaproteobacterial sulfide oxidizers are dominant.

The sulfate-reducing bacterial genomes evaluated in this research highlight adaptations to different redox conditions in the microbial mat. Sulfate reducers are not limited to anoxic conditions in microbial mats (Canfield & Marais, 1991; Jørgensen & Bak, 1991), and different members have varying O_2 tolerances (Cypionka et al., 1985; Sass et al., 1997). Desulfobacteraceae;*Desulfonema* members have been found to thrive in oxic microbial mats (Minz et al., 1999), and they are the most abundant sulfate-reducing member in the MIS microbial mats (**Figure 4.2**). Further, the metagenomic bins and their distribution across mats suggest the existence of multiple ecotypes of Desulfobacteraceae with different aerotolerances and/or capacities for electron donors. The *Desulfonema* cluster grouped five MAGs sourced from only mats sampled in 2016, including giraffe mat and flat purple mat. The *Desulfobacteraceae* cluster included 11 genomes of varying completion from all mat types over the years. The *Desulfonema* representative genome possesses two oxidases for reducing O_2 at different levels, a bd-quinol oxidase and a cytochrome c oxidase for higher O_2 levels than the bd type (Rabus et al., 2004; Ramel et al., 2013). Sulfate-reducing bacteria may use cytochrome c oxidase to detoxify (Ramel et al., 2013) reduce O_2, coupled with oxidation of organic compounds, inorganic sulfur compounds, and H_2 (Dannenberg et al., 1992; Dilling & Cypionka, 1990; Hardy & Hamilton,

1981). The *Desulfobacteraceae* MAG has only a cytochrome bd-quinol oxidase, which has a high O_2 affinity and been involved more in O_2 detoxification than in energy production (Ramel et al., 2013). The distribution of the *Desulfobacteraceae* MAG in finger and flat mat samples, which typically have higher O_2 flux than other samples, is surprising if its oxidase is best suited for O_2 detoxification and not aerobic respiration.

Distinct organisms are responsible for core functions that are preserved across mat types.

Our proteomic investigation revealed subtle functional shifts between the observed mat morphotypes. Corroborating abundance patterns based on 16S rRNA genes, *Phormidium* was significantly more active in finger and flat-*Phormidium* dominated mats, and less active in other mat types (**Figure 4.6**). Significantly differentially abundant gammaproteobacterial and deltaproteobacterial proteins were observed in higher abundance in flat-*Spirulina* and *Pseudanabaena* mat, giraffe, and white mat. These proteins typically represented core processes such as ribosomes and chaperones, but also revealed the involvement of these organisms in O, S, C, and N cycling. Distinct phycobiliproteins were significantly differentially abundant in the morphotypes, supporting the role of functional niches and potentially chromatic adaptation for *Phormidium, Planktothrix, Pseudanabaena,* and other cyanobacteria. We also observed significantly higher abundance of two protein subunits of reverse dissimilatory sulfite reductase and adenylylsulfate reductase in flat *Spirulina* and *Pseudanabaena*-dominated mats, giraffe and white mats, suggesting heightened sulfide oxidation in these mat types.

Though a select group of potentially influential proteins were significantly differentially abundant between mats, nearly 1/3 of observed proteins were cosmopolitan and not different in abundance across mats (**Table SI 4.3**). The IMG-assigned taxonomies of the contigs hosting

these cosmopolitan proteins point to a core functional group of organisms. Judging from protein abundance, *Phormidium*, gammaproteobacteria, and deltaproteobacteria are substantial members of this core group. Cyanobacteria *Spirulina*, *Pseudanabaena*, and *Planktothrix* also contribute proteins to the cosmopolitan core, but are far less often observed compared to *Phormidium*. These core functional members cover many instrumental processes such as photosynthesis, carbon metabolism, translation, and sulfur metabolism. Indeed, photosynthetic proteins were the most widely observed in the core proteome, and indicate the operation of both photosystems I and II. The abundances and distribution of superoxide dismutase, peroxiredoxins, and peroxidases suggest that cyanobacteria, gammaproteobacteria, and other members are responding to oxidative stress. Sulfur metabolism, such as sulfate reduction and sulfur oxidation, is also occurring in most mat types. The abundance of chaperones and ribosomes and their distribution among many taxa, point to active growth of these organisms in the core functional group. Indeed, these processes underpinned by organisms such as *Phormidium,* gammaproteobacterial sulfide oxidizers, and deltaproteobacterial sulfate reducers, may not vary significantly in magnitude between the mat types, and it is the and/or subtle environmental differences of key proteins of select organisms that may give rise to the different morphological appearances.

Mat morphotypes as an expression of their environment filtering for cyanobacteria.

Our study determined that mat morphotypes in MIS have distinct community structures. The dominant cyanobacteria were different in abundance between the finger, giraffe mat, white mat, and flat mat morphotypes. Similarly, abundances and identities of the sulfate-reducing bacteria and sulfide-oxidizing bacteria were different between the mat types. While the

composition and structure of the bacterial communities varied, the number of proteins that differed between mat types was a fraction of the pool of proteins that were observed across many mat types. Our previous research determined that seasonality is a strong influence on the bacterial community, likely through changes in groundwater chemistry and/or light quantity. However, many of the mat morphotypes evaluated in this current study occurred in the same season, often within 50 m of each other. Further, except for the lack of photosynthesis in white mats and a much-reduced rate of AP in fingers, fluxes of total sulfide and O_2 were invariable between mat types. Thus a question emerging from this research is, what caused the appearance and distribution of the MIS mat morphotypes?

It is striking to note the proliferation of *Pseudanabaena* and *Spirulina* members in the mat starting from 2016. Prior to 2016, *Spirulina* and *Pseudanabaena* were each 1% or less of the bacterial community (**Chapter III**) and were not effectively resolved in prior metagenomes (Voorhies, 2014). In one sample from October 2014 they reached their previously observed maximum abundance of 6% and 1% for *Spirulina* and *Pseudanabaena* respectively, before we sampled in 2016. Thereafter, these cyanobacteria became staples of the bacterial community. Our proteomics dataset points to lower protein abundance of *Phormidium* (and thus lower growth) and higher abundance of *Pseudanabaena* and other cyanobacterial proteins (higher growth) in giraffe, white, and *Spirulina* and *Pseudanabaena*-dominated mat. These signatures imply that the increased relative abundance of non-*Phormidium* members in these mats is likely due to growth conditions that favor *Pseudanabaena* and *Spirulina*. After 2016 the dominance of *Phormidium* is limited to finger mats, as evident in the 16S rRNA gene and proteomic profiles.

Favorable growth conditions for *Pseudanabaena* may be related to temperature, salinity, and depth. When cultures of *Phormidium* and *Pseudanabaena* from polar environments were

subjected to different growth temperatures, *Phormidium* is more successful at cooler temperatures (4-8˚C) and *Pseudanabaena* is more successful at warmer temperature (16-23˚C) (Kleinteich et al., 2012). In Lake Huron, though the groundwater source remains at a constant 8.4˚C throughout the year, mixing with the dynamic water column changes the physicochemical regime in the sinkhole arena. While we did not target long-term shifts in the overlying lake water of the sinkhole, warming in the Great Lakes in the last decade (Zhong et al., 2019) likely altered the physicochemical properties of the water column and impacted the benthic diatom community of coastal areas in Lake Huron (Sivarajah et al., 2018), leaving room for such influences in MIS. Other environmental parameters that impact the distribution *Phormidium* and *Pseudanabaena* in microbial mats, such as salinity and depth, may be less influential in MIS. In Antarctic lakes influenced by hypersaline brine or brackish water, typically *Phormidium* was more abundant than *Pseudanabaena* members (Pessi et al., 2016). Within a depth gradient in Lake Fryxell, *Pseudanabaena* was more abundant in shallower, beige mats whereas *Phormidium* was more abundant in pink-pigmented ridge and prostrate mats that are more prevalent in deeper mats on the shelf (Jungblut et al., 2016). Light optima, sulfide tolerance, and ability to withstand physicochemical stress determines depth niches for cyanobacteria in other microbial mats (Bradley et al., 2017; Jungblut et al., 2016; Noffke, 2010; Pepe-Ranney et al., 2012; Wilmeth et al., 2018).

These experiments focused on axenic cultures, or related depth and salinity changes to differences in the growth of *Phormidium* and *Pseudanabaena.* In MIS, without clear differences in the measurements of the geochemical environment that we made (e.g. sulfide and oxygen profiles), it is likely that microbial interspecies interactions may influence community composition. The flat mats dominated by *Pseudanabaena* and *Spirulina* had lower observed

species richness than *Phormidium*-dominated mats. Generally, the sulfide-oxidizing community in *Phormidium*-dominated mats (including fingers) was more even between gammaproteobacteria and epsilonproteobacteria, and we observed a higher abundance of other cyanobacteria as well. Metagenomic evidence suggests these organisms may have more diverse potential biogeochemical influences, than those found in *Pseudanabaena* and *Spirulina* mats. However, with only a fraction of non-cyanobacterial proteins significantly differing between the two mat types, the impact of these microbes on the growth of *Pseudanabaena* and *Phormidium* is not well understood. Further work on the optimal growth conditions for MIS cyanobacteria can untangle whether *Phormidium* and *Pseudanabaena* are so functionally similar that subtle shifts in ecological conditions can influence which cyanobacterium reigns supreme, and whether they have responses and functions that shape the mat appearance.

The MIS environment is generally the same throughout and experiences the same groundwater source, and our broad geochemical measurements indicated no substantial difference in O_2 and H_2S flux between cyanobacterial mat types. We did not capture any smaller-scale sediment-derived geochemical differences, if they were present, between the areas in which the mat types are primarily found. Changes in the physicochemical environment are able to induce bottom-up shifts in the appearance of microbial mats. For example, an EPS-rich microbial mat can trap gases underneath, increasing pressure under the mat and causing it to uplift into the 'gas dome' mat type (Noffke, 2010). Modern mats in coastal environments at the right depth and experiencing the right tidal frequency will generate kinneyia wrinkle structures (Herminghaus et al., 2016). At the depth of MIS with its ~1 m layer of groundwater, tidal frequencies in the basin generated from the imbalance between the cold, high-conductivity groundwater and fresh overlying lake water, could generate ripple patterns of the right wavelength that would yield

giraffe mat (Herminghaus et al., 2016). A hydrodynamic model such as this would have to be refined to interpret the localization of giraffe mat to only one area of the sinkhole.

Our understanding of the biological and environmental mechanisms behind the formation of microbial mats in the geologic record, has been significantly deepened thanks to investigations in extant settings (Noffke, 2010). Yet without clear linkages between geochemical fluxes, microbial community structure, and metabolic functioning in the mat morphotypes observed in MIS, our study posits caution in inferring metabolisms from preserved morphotypes. Sulfide-oxidizing bacteria have rarely been implicated in the formation of preserved microbial mats because of the importance paid to phototrophs, but they are abundant and active members in the MIS mats and are key mat members in chemotrophic communities (Flood et al. 2014). AP is not a widely observed metabolism among cyanobacteria, yet in MIS this capability may be present in dominant cyanobacteria in discrete mat types. Even in fingers, the uplifted mats most similar to cones, AP is likely an important process in depleting locally produced sulfide, yet cones and domes in the geologic record have been interpreted to be O_2 oases that are observed only after the GOE (Bosak et al., 2009; Bradley et al., 2017). Though the processes behind the appearance and distribution of morphotypes in MIS are not yet well known, the presence of core functional groups and metabolisms across these mats, and the consistency in geochemical fluxes, complicates the biogeochemical role of different mat morphotypes in Earth history.

4.6 Conclusion

The Middle Island Sinkhole mat morphotypes we evaluated in this study hosted taxonomically and functionally diverse microbial communities. This spatiotemporal evaluation is in contrast to previous perspectives of MIS mats being stable, low-diversity ecosystems

(Voorhies, 2014). These mats are similar in appearance to those observed in the geologic record, and are ideal systems in which to explore the link between functional similarities, geochemistry, and phylogenetic relationship. A small pool of significantly differentially abundant proteins from key organisms, subtle variations in geochemical cycling, and local physicochemical differences may play a role in the morphological appearance of mats. The differing distributions, proteomic abundances, and genome sizes of cyanobacteria with similar S, O, and N metabolizing genes in distinct mats implicates functional similarity between microbes that characterize select mat morphotypes. Despite prominent differences in appearance, these mats also host common metabolic guilds of sulfate-reducing bacteria and sulfide-oxidizing bacteria. Finally, the metaproteomic profiles presented here provide opportunities for untangling the functional interactions of microbes in extreme ecosystems, and a framework for potential metabolisms in ancient mats.

4.7 Acknowledgements

We are grateful to NOAA Thunder Bay National Marine Sanctuary and divers for field site access, facilities support, and sampling. We thank Judith Klatt, Katy Rico, Matthew Medina, Bopaiah Biddanda, Steve Ruberg, Arjun Chennu, Dirk de Beer, Allen Burton, Michelle Hudson, Phoebe Aron, and Naomi Levin for field sampling, data collection, processing, and analyses.

4.8 References

Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., et al. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nature Communications, 7, 1–11. https://doi.org/10.1038/ncomms13219

Anantharaman, K., Duhaime, M. B., Breier, J. A., Wendt, K. A., Dick, G. J., & Toner, B. M. (2014). Sulfur oxidation genes in diverse deep-sea viruses. Science. https://doi.org/10.1126/science.1250598

Arieli, B., Shahak, Y., Taglicht, D., Hauska, G., & Padan, E. (1994). Purification and characterization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in Oscillatoria limnetica. Journal of Biological Chemistry, 269(8), 5705–5711.

Aro, E.-M., Virgin, I., & Andersson, B. (1993). Photoinhibition of photosystem II. Inactivation, protein damage and turnover. Biochimica Et Biophysica Acta, 1143(2), 113–134. https://doi.org/10.1016/0005-2728(93)90134-2

Baskaran, M., Novell, T., Nash, K., Ruberg, S. A., Johengen, T., Hawley, N., et al. (2016). Tracing the Seepage of Subsurface Sinkhole Vent Waters into Lake Huron Using Radium and Stable Isotopes of Oxygen and Hydrogen. Aquatic Geochemistry, 22(4), 349–374. https://doi.org/10.1007/s10498-015-9286-7

Baumgartner, L. K., Reid, R. P., Dupraz, C., Decho, A. W., Buckley, D. H., Spear, J. R., et al. (2006). Sulfate reducing bacteria in microbial mats: Changing paradigms, new discoveries. Sedimentary Geology, 185(3-4), 131–145. https://doi.org/10.1016/j.sedgeo.2005.12.008

Becraft, E. D., Wood, J. M., Rusch, D. B., Kühl, M., Jensen, S. I., Bryant, D. A., et al. (2015). The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of Synechococcus inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park. Frontiers in Microbiology, 6. https://doi.org/10.3389/fmicb.2015.00590

Biddanda, B. A., Nold, S. C., Ruberg, S. A., Kendall, S. T., Sanders, T. G., & Gray, J. J. (2009). Great Lakes Sinkholes: A Microbiogeochemical Frontier. Eos, Transactions American Geophysical Union, 90(8), 61–62. https://doi.org/10.1029/2009EO080001

Blankenship, R. E. (2014). Molecular mechanisms of photosynthesis (1st ed.). Blackwell Science Ltd.

Bosak, T., Knoll, A. H., & Petroff, A. P. (2013). The meaning of stromatolites. Annual Review of Earth and Planetary Sciences. https://doi.org/10.1146/annurev-earth-042711-105327

Bosak, T., Liang, B., Sim, M. S., & Petroff, A. P. (2009). Morphological record of oxygenic photosynthesis in conical stromatolites. Proceedings of the National Academy of Sciences of the United States of America, 106(27), 10939–10943. https://doi.org/10.1073/pnas.0900885106

Bradley, J. A., Daille, L. K., Trivedi, C. B., Bojanowski, C. L., Stamps, B. W., Stevenson, B. S., et al. (2017). Carbonate-rich dendrolitic cones: insights into a modern analog for incipient microbialite formation, Little Hot Creek, Long Valley Caldera, California. Npj Biofilms and Microbiomes, 1–11. https://doi.org/10.1038/s41522-017-0041-2

Brettar, I., Labrenz, M., Flavier, S., Bötel, J., Kuosa, H., Christen, R., & Höfle, M. G. (2006). Identification of a Thiomicrospira denitrificans-like epsilonproteobacterium as a catalyst for autotrophic denitrification in the central Baltic Sea. Applied and Environmental Microbiology, 72(2), 1364–1372. https://doi.org/10.1128/AEM.72.2.1364-1372.2006

Brochier-Armanet, C., Talla, E., & Gribaldo, S. (2009). The Multiple Evolutionary Histories of Dioxygen Reductases: Implications for the Origin and Evolution of Aerobic Respiration. Molecular Biology and Evolution, 26(2), 285–297. https://doi.org/10.1093/molbev/msn246

Bruckner, C. G., Mammitzsch, K., Jost, G., Wendt, J., Labrenz, M., & Jürgens, K. (2012). Chemolithoautotrophic denitrification of epsilonproteobacteria in marine pelagic redox gradients. Environmental Microbiology, 15(5), 1505–1513. https://doi.org/10.1111/j.1462-2920.2012.02880.x

Bryant, D. A. (1982). Phycoerythrocyanin and Phycoerythrin: Properties and Occurrence in Cyanobacteria. Journal of General Microbiology, 128, 835–844.

Canfield, D. E., & Marais, Des, D. J. (1991). Aerobic sulfate reduction in microbial mats. Science, 251, 1471–1473.

Cardona, T., Murray, J. W., & Rutherford, A. W. (2015). Origin and Evolution of Water Oxidation before the Last Common Ancestor of the Cyanobacteria. Molecular Biology and Evolution, 32(5), 1310–1328. https://doi.org/10.1093/molbev/msv024

Chan, L.-K., Morgan-Kiss, R. M., & Hanson, T. E. (2009). Functional Analysis of Three Sulfide:Quinone Oxidoreductase Homologs in Chlorobaculum tepidum. Journal of Bacteriology, 191(3), 1026–1034. https://doi.org/10.1128/JB.01154-08

Chen, K., Roberts, E., & Luthey-Schulten, Z. (2009). Horizontal gene transfer of zinc and non-zinc forms of bacterial ribosomal protein S4. BMC Evolutionary Biology, 9(1), 179. https://doi.org/10.1186/1471-2148-9-179

Cohen, Y., Jørgensen, B. B., Padan, E., & Shilo, M. (1975). Sulphide-dependent anoxygenic photosynthesis in the cyanobacterium Oscillatoria limnetica. Nature, 257(5526), 489–492. https://doi.org/10.1038/257489a0

Cypionka, H., Widdel, F., & Pfennig, N. (1985). Survival of sulfate-reducing bacteria after oxygen stress, and growth in sulfate-free oxygen-sulfide gradients. FEMS Microbiology Ecology, 31, 39–45.

Dannenberg, S., Kroder, M., Dilling, W., & Cypionka, H. (1992). Oxidation of H2 organic compounds and inorganic sulfur compounds coupled to reduction of O2 or nitrate by sulfate-reducing bacteria. Archives of Microbiology, 158, 93–99.

de Beer, D., Schramm, A., Santegoeds, C. M., Kühl, M. (1997). A nitrite microsensor for profiling environmental biofilms. Applied Environmental Microbiology, 63(3), 973–977.

de Beer, D., Weber, M., Chennu, A., Hamilton, T. L., Lott, C., Macalady, J., & M Klatt, J. (2017). Oxygenic and anoxygenic photosynthesis in a microbial mat from an anoxic and sulfidic spring. Environmental Microbiology, 19(3), 1251–1265. https://doi.org/10.1111/1462-2920.13654

Decelle, J., Romac, S., Stern, R. F., Bendif, E. M., Zingone, A., Audic, S., et al. (2015). PhytoREF: a reference database of the plastidial 16S rRNA gene of photosynthetic eukaryotes with curated taxonomy. Molecular Ecology Resources, 15(6), 1435–1445. https://doi.org/10.1111/1755-0998.12401

Dick, G. J., Grim, S. L., & Klatt, J. M. (2018). Controls on O2 Production in Cyanobacterial Mats and Implications for Earth's Oxygenation. Annual Review of Earth and Planetary Sciences, 46(1), 123–147. https://doi.org/10.1146/annurev-earth-082517-010035

Dilling, W., & Cypionka, H. (1990). Aerobic respiration in sulfate-reducing bacteria. FEMS Microbiology Letters, 71, 123–128.

Erde, J., Loo, R. R. O., & Loo, J. A. (2014). Enhanced FASP (eFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments. Journal of Proteome Research, 13(4), 1885–1895. https://doi.org/10.1021/pr4010019

Eren, A. M., Maignien, L., Sul, W. J., Murphy, L. G., Grim, S. L., Morrison, H. G., & Sogin, M. L. (2013). Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. Methods in Ecology and Evolution, 4(12), 1111–1119. https://doi.org/10.1111/2041-210X.12114

Eren, A. M., Morrison, H. G., Lescault, P. J., Reveillaud, J., Vineis, J. H., & Sogin, M. L. (2014). Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high- throughput marker gene sequences. The ISME Journal, 1–12. https://doi.org/10.1038/ismej.2014.195

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNARestriction Data. Genetics, 131, 479–491.

Fahnenstiel, G. L., & Carrick, H. J. (1991). Phototrophic Picoplankton in Lakes Huron and Michigan: Abundance, Distribution, Composition, and Contribution to Biomass and Production. Canadian Journal of Fisheries and Aquatic Sciences, 49, 379–388. https://doi.org/https://doi.org/10.1139/f92-043

Falkowski, P. G., & Godfrey, L. V. (2008). Electrons, life and the evolution of Earth's oxygen cycle. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1504), 2705–2716. https://doi.org/10.1098/rstb.2008.0054

Flood, B. E., Bailey, J. V., & Biddle, J. F. (2014). Horizontal gene transfer and the rock record: comparative genomics of phylogenetically distant bacteria that induce wrinkle structure formation in modern sediments. Geobiology, 12(2), 119–132. https://doi.org/10.1111/gbi.12072

Fukami, T. (2015). Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. Annual Review of Ecology, Evolution, and Systematics, 46(1), 1–23. https://doi.org/10.1146/annurev-ecolsys-110411-160340

Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. Bioinformatics, 31(22), 3718–3720. https://doi.org/10.1093/bioinformatics/btv428/-/DC1

Garcia Pichel, F., & Castenholz, R. W. (1990). Comparative anoxygenic photosynthetic capacity in 7 strains of a thermophilic cyanobacterium. Archives of Microbiology, 153(4), 344–351. https://doi.org/10.1007/BF00249003

Garcia, S. L., Buck, M., Hamilton, J. J., Wurzbacher, C., Grossart, H.-P., McMahon, K. D., & Eiler, A. (2018). Model Communities Hint at Promiscuous Metabolic Linkages between Ubiquitous Free-Living Freshwater Bacteria. mSphere, 3(3), 257. https://doi.org/10.1128/mSphere.00202-18

Giovannoni, S. J., Thrash, J. C., & Temperton, B. (2014). Implications of streamlining theory for microbial ecology. The ISME Journal, 8(8), 1553–1565. https://doi.org/10.1038/ismej.2014.60

Gogarten, J. P., & Townsend, J. P. (2005). Horizontal gene transfer, genome innovation and evolution. Nature Reviews Microbiology, 3(9), 679–687. https://doi.org/10.1038/nrmicro1204

Gregersen, L. H., Bryant, D. A., & Frigaard, N.-U. (2011). Mechanisms and evolution of oxidative sulfur metabolism in green sulfur bacteria. Front Microbiol, 2(May), 1–14. https://doi.org/10.3389/fmicb.2011.00116/abstract

Grim, S. L., & Dick, G. J. (2016). Photosynthetic Versatility in the Genome of Geitlerinema sp. PCC 9228 (Formerly Oscillatoria limnetica “Solar Lake”), a Model Anoxygenic Photosynthetic Cyanobacterium. Frontiers in Microbiology, 7(590), 1144. https://doi.org/10.3389/fmicb.2016.01546

Hamilton, T. L., Klatt, J. M., de Beer, D., & Macalady, J. L. (2018). Cyanobacterial photosynthesis under sulfidic conditions: insights from the isolate Leptolyngbya sp. strain hensonii. The ISME Journal, 12(2), 568–584. https://doi.org/10.1038/ismej.2017.193

Han, Y., & Perner, M. (2015). The globally widespread genus Sulfurimonas: versatile energy metabolisms and adaptations to redox clines. Frontiers in Microbiology, 6(185), 90. https://doi.org/10.3389/fmicb.2015.00989

Hardy, J. A., & Hamilton, W. A. (1981). The Oxygen Tolerance of Sulfate-Reducing Bacteria Isolated from North Sea Waters. Current Microbiology, 6, 259–262.

Herminghaus, S., Thomas, K. R., Aliaskarisohi, S., Porada, H., & Goehring, L. (2016). Kinneyia: A Flow-Induced Anisotropic Fossil Pattern from Ancient Microbial Mats. Frontiers in Materials, 3, 691. https://doi.org/10.3389/fmats.2016.00030

Hoffmann, T., Troup, B., Szabo, A., Hungerer, C., & Jahn, D. (1995). The anaerobic life of Bacillus subtilis: Cloning of the genes encoding the respiratory nitrate reductase system. FEMS Microbiology Letters, 131, 219–225.

Hubbell, S. P. (2005). Neutral theory in community ecology and the hypothesis of functional equivalence. Functional Ecology, 19, 166–172.

Huerta-Diaz, M. A., & Morse, J. W. (1992). Pyritization of trace metals in anoxic marine sediments. Geochimica Et Cosmochimica Acta, 56, 2681–2702.

Huerta-Diaz, M. A., Delgadillo-Hinojosa, F., Siqueiros-Valencia, A., Valdivieso-Ojeda, J., Reimer, J. J., & Segovia-Zavala, J. A. (2012). Millimeter-scale resolution of trace metal distributions in microbial mats from a hypersaline environment in Baja California, Mexico. Geobiology, 10(6), 531–547. https://doi.org/10.1111/gbi.12008

Huse, S. M., Dethlefsen, L., Huber, J. A., Welch, D. M., Relman, D. A., & Sogin, M. L. (2008). Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. PLoS Genetics, 4(11), e1000255. https://doi.org/10.1371/journal.pgen.1000255

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. Bioinformatics, 11, 1–11.

Jeroschewski, P., Steuckart, C., Kühl, M. (1996). An amperometric microsensor for the determination of H2S in aquatic environments. Analytical Chemistry, 68(24), 4351–4357.

Johnston, D. T., Wolfe-Simon, F., Pearson, A., & Knoll, A. H. (2009). Anoxygenic photosynthesis modulated Proterozoic oxygen and sustained Earth's middle age. Proceedings of the National Academy of Sciences, 106(40), 16925–16929. https://doi.org/10.1073/pnas.0909248106

Jungblut, A. D., Hawes, I., Mackey, T. J., Krusor, M., Doran, P. T., Sumner, D. Y., et al. (2016). Microbial Mat Communities along an Oxygen Gradient in a Perennially Ice-Covered Antarctic Lake. Applied and Environmental Microbiology, 82(2), 620–630. https://doi.org/10.1128/AEM.02699-15

Jørgensen, B. B., & Bak, F. (1991). Pathways and Microbiology of Thiosulfate Transformations and Sulfate Reduction in a Marine Sediment (Kattegat, Denmark). Applied and Environmental Microbiology, 57(3), 847–856.

Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ, 3(8), e1165. https://doi.org/10.7717/peerj.1165

Kembel, S. W. (2009). Disentangling niche and neutral influences on community assembly: assessing the performance of community phylogenetic structure tests. Ecology Letters,

12(9), 949–960. https://doi.org/10.1111/j.1461-0248.2009.01354.x

Klatt, J. M., de Beer, D., Häusler, S., & Polerecky, L. (2016). Cyanobacteria in Sulfidic Spring Microbial Mats Can Perform Oxygenic and Anoxygenic Photosynthesis Simultaneously during an Entire Diurnal Period. Frontiers in Microbiology, 7(116), 440. https://doi.org/10.3389/fmicb.2016.01973

Klatt, J. M., Haas, S., Yilmaz, P., de Beer, D., & Polerecky, L. (2015). Hydrogen sulfide can inhibit and enhance oxygenic photosynthesis in a cyanobacterium from sulfidic springs. Environmental Microbiology, 17(9), 3301–3313. https://doi.org/10.1111/1462-2920.12791

Klatt, J. M., Meyer, S., Häusler, S., Macalady, J. L., de Beer, D., & Polerecky, L. (2016). Structure and function of natural sulphide-oxidizing microbial mats under dynamic input of light and chemical energy. The ISME Journal, 10(4), 921–933. https://doi.org/10.1038/ismej.2015.167

Kleinteich, J., Wood, S. A., Küpper, F. C., Camacho, A., Quesada, A., Frickey, T., & Dietrich, D. R. (2012). Temperature-related changes in polar cyanobacterial mat diversity and toxin production. Nature Climate Change, 2(5), 356–360. https://doi.org/10.1038/nclimate1418

Konhauser, K. O. (2009). Introduction to Geomicrobiology. John Wiley & Sons. https://doi.org/10.1371/journal.pone.0140766

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology, 79(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Kraus, E. A., Beeler, S. R., Mors, R. A., Floyd, J. G., GeoBiology 2016, Stamps, B. W., et al. (2018). Microscale Biosignatures and Abiotic Mineral Authigenesis in Little Hot Creek, California. Frontiers in Microbiology, 9, 544. https://doi.org/10.3389/fmicb.2018.00997

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution forlarge and complex metagenomics assemblyvia succinct de Bruijn graph. Bioinformatics, 31, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033/-/DC1

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., et al. (2018). Function and functional redundancy in microbial systems. Nature Ecology & Evolution, 1–8. https://doi.org/10.1038/s41559-018-0519-1

Macalady, J. L., Lyon, E. H., Koffman, B., Albertson, L. K., Meyer, K., Galdenzi, S., & Mariani, S. (2006). Dominant microbial populations in limestone-corroding stream biofilms, Frasassi cave system, Italy. Applied and Environmental Microbiology, 72(8), 5596–5609. https://doi.org/10.1128/AEM.00715-06

Marcia, M., Ermler, U., Peng, G., & Michel, H. (2010). A new structure-based classification of sulfide:quinone oxidoreductases. Proteins: Structure, Function, and Bioinformatics, 78(5), 1073–1083. https://doi.org/10.1002/prot.22665

Miller, C. S., Baker, B. J., Thomas, B. C., Singer, S. W., Banfield, J. F., & Banfield, J. F. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. Genome Biology, 12, 1–14. https://doi.org/10.1186/gb-2011-12-5-r44

Miller, S. R., & Bebout, B. M. (2004). Variation in sulfide tolerance of photosystem II in phylogenetically diverse cyanobacteria from sulfidic habitats. Applied and Environmental Microbiology, 70(2), 736–744. https://doi.org/10.1128/AEM.70.2.736-744.2004

Minz, D., Fishbain, S., Green, S. J., Muyzer, G., Cohen, Y., Rittmann, B. E., & Stahl, D. A. (1999). Unexpected Population Distribution in a Microbial Mat Community: Sulfate-Reducing Bacteria Localized to the Highly Oxic Chemocline in Contrast to a Eukaryotic Preference for Anoxia. Applied and Environmental Microbiology, 65(10), 4659–4665.

Morris, R. L., & Schmidt, T. M. (2013). Shallow breathing: bacterial life at low O2. Nature Reviews Microbiology, 1–8. https://doi.org/10.1038/nrmicro2970

Mussmann, M., Hu, F. Z., Richter, M., de Beer, D., Preisler, A., Jørgensen, B. B., et al. (2007). Insights into the Genome of Large Sulfur Bacteria Revealed by Analysis of Single Filaments. PLoS Biology, 5(9), e230. https://doi.org/10.1371/journal.pbio.0050230

Nagy, C. I., Vass, I., Rákhely, G., Vass, I. Z., Tóth, A., Duzs, A., et al. (2014). Coregulated genes link sulfide:quinone oxidoreductase and arsenic metabolism in Synechocystis sp. strain PCC6803. Journal of Bacteriology, 196(19), 3430–3440. https://doi.org/10.1128/JB.01864-14

Noffke, N. (2010). Geobiology: Microbial mats in sandy deposits from the Archean Era to today (pp. 1–198). Springer Science & Business Media.

Noffke, N., & Awramik, S. M. (2013). Stromatolites and MISS—differences between relatives. GSA Today. https://doi.org/10.1130/GSATG187A.1

Noffke, N., Gerdes, G., & Klenke, T. (2003). Benthic cyanobacteria and their influence on the sedimentary dynamics of peritidal depositional systems (siliciclastic, evaporitic salty, and evaporitic carbonatic). Earth-Science Reviews, 62(1-2), 163–176. https://doi.org/10.1016/S0012-8252(02)00158-7

Nonaka, H., Keresztes, G., Shinoda, Y., Ikenaga, Y., Abe, M., Naito, K., et al. (2006). Complete genome sequence of the dehalorespiring bacterium Desulfitobacterium hafniense Y51 and comparison with Dehalococcoides ethenogenes 195. Journal of Bacteriology, 188(6), 2262–2274. https://doi.org/10.1128/JB.188.6.2262-2274.2006

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., OHara, R. B., et al. (2015,

January 12). Package “vegan.”

Olm, M. R., Brown, C. T., Brooks, B., & Banfield, J. F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication, 11(12), 2864–2868. https://doi.org/10.1038/ismej.2017.126

Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., LaButti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. Comparative genomics of Synechococcus strains with different light responses and in situ diel transcription patterns of associated putative ecotypes in the Mushroom Spring microbial mat. Frontiers in Microbiology, 6, 604. https://doi.org/10.3389/fmicb.2015.00604

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics, 26(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research, 25(7), 1043–1055. https://doi.org/10.1101/gr.186072.114

Patwardhan, S., Foustoukos, D. I., Giovannelli, D., Yücel, M., & Vetriani, C. (2018). Ecological Succession of Sulfur-Oxidizing Epsilon- and Gammaproteobacteria During Colonization of a Shallow-Water Gas Vent. Frontiers in Microbiology, 1–16. https://doi.org/10.3389/fmicb.2018.02970

Pepe-Ranney, C., Berelson, W. M., Corsetti, F. A., Treants, M., & Spear, J. R. (2012). Cyanobacterial construction of hot spring siliceous stromatolites in Yellowstone National Park. Environmental Microbiology, 14(5), 1182–1197. https://doi.org/10.1111/j.1462-2920.2012.02698.x

Pessi, I. S., Maalouf, P. de C., Laughinghouse, H. D., IV, Baurain, D., & Wilmotte, A. (2016). On the use of high-throughput sequencing for the study of cyanobacterial diversity in Antarctic aquatic mats. Journal of Phycology, n/a–n/a. https://doi.org/10.1111/jpy.12399

Porada, H., Ghergut, J., & Bouougri, E. H. (2008). Kinneyia-Type Wrinkle Structures--Critical Review And Model Of Formation. Palaios, 23(2), 65–77. https://doi.org/10.2110/palo.2006.p06-095r

Poulton, S. W., Fralick, P. W., & Canfield, D. E. (2010). Spatial variability in oceanic redox structure 1.8 billion years ago. Nature Geoscience, 3(7), 486–490. https://doi.org/10.1038/ngeo889

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research, 35(21), 7188–7196. https://doi.org/10.1093/nar/gkm864

R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria.

Retrieved from https://www.r-project.org/

Rabus, R., Ruepp, A., Frickey, T., Rattei, T., Fartmann, B., Stark, M., et al. (2004). The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium from permanently cold Arctic sediments. Environmental Microbiology, 6(9), 887–902. https://doi.org/10.1111/j.1462-2920.2004.00665.x

Ramel, F., Amrani, A., Pieulle, L., Lamrabet, O., Voordouw, G., Seddiki, N., et al. (2013). Membrane-bound oxygen reductases of the anaerobic sulfate-reducing Desulfovibrio vulgaris Hildenborough: roles in oxygen defence and electron link with periplasmic hydrogen oxidation. Microbiology, 159(Pt_12), 2663–2673. https://doi.org/10.1099/mic.0.071282-0

Raymond, J., Zhaxybayeva, O., Gogarten, J. P., Gerdes, S. Y., & Blankenship, R. E. (2002). Whole-Genome Analysis of Photosynthetic Prokaryotes. Science, 298, 1616–1620.

Rákhely, G., Laurinavichene, T. V., Tsygankov, A. A., & Kovács, K. L. (2007). The role of Hox hydrogenase in the H2 metabolism of Thiocapsa roseopersicina. Biochimica Et Biophysica Acta (BBA) - Bioenergetics, 1767(6), 671–676. https://doi.org/10.1016/j.bbabio.2007.02.004

Revsbech, N. P., Jørgensen, B. B. (1983). Photosynthesis of benthic microflora measured with high spatial resolution by the oxygen microprofile method: capabilities and limitations of the method. Limnology and Oceanography, 28(4), 749–759.

Reyes, K., Gonzalez, N. I., Stewart, J., Ospino, F., Nguyen, D., Cho, D. T., et al. (2013). Surface orientation affects the direction of cone growth by Leptolyngbya sp. strain C1, a likely architect of coniform structures Octopus Spring (Yellowstone National Park). Applied and Environmental Microbiology, 79(4), 1302–1308. https://doi.org/10.1128/AEM.03008-12

Rico, K. I. (2019, May 10). Controls on the sediment geochemistry of a low-oxygen Precambrian analogue. (N. D. Sheldon, Ed.).

Rico, K. I., & Sheldon, N. D. (2019). Nutrient and iron cycling in a modern analogue for the redoxcline of a Proterozoic ocean shelf. Chemical Geology, 511, 42–50. https://doi.org/10.1016/j.chemgeo.2019.02.032

RStudio Team. (2014). RStudio: Integrated Development for R. Retrieved from http://www.rstudio.com/

Ruberg, S. A., Kendall, S. T., & Biddanda, B. A. (2008). Observations of the Middle Island Sinkhole in Lake Huron–a unique hydrogeologic and glacial creation of 400 million years. Marine Technology Society Journal, 42(4), 12–21. https://doi.org/10.4031/002533208787157633

Sanchez-Baracaldo, P., Hayes, P. K., & Blank, C. E. (2005). Morphological and habitat evolution in the Cyanobacteria using a compartmentalization approach. Geobiology, 3(3),

145–165. https://doi.org/10.1111/j.1472-4669.2005.00050.x

Sass, H., Cypionka, H., & Babenzien, H.-D. (1997). Vertical distribution of sulfatereducing bacteria at the oxicanoxic interface in sediments of the oligotrophic Lake Stechlin. FEMS Microbiology Ecology, 22, 245–255.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology, 75(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Schmidt, M. L. (2018, September 21). Microhabitats Shape Bacterial Community Composition, Ecosystem Function, and Genome Traits. (V. J. Denef, Ed.).

Schmidt, T. M., Arieli, B., Cohen, Y., Padan, E., & Strohl, W. R. (1987). Sulfur Metabolism in Beggiatoa alba. Journal of Bacteriology, 169(12), 1–7.

Seekatz, A. M., Theriot, C. M., Molloy, C. T., Wozniak, K. L., Bergin, I. L., & Young, V. B. (2015). Fecal Microbiota Transplantation Eliminates Clostridium difficile in a Murine Model of Relapsing Disease. Infection and Immunity, 83(10), 3838–3846. https://doi.org/10.1128/IAI.00459-15

Sharrar, A. M., Flood, B. E., Bailey, J. V., Jones, D. S., Biddanda, B. A., Ruberg, S. A., et al. (2017). Novel Large Sulfur Bacteria in the Metagenomes of Groundwater-Fed Chemosynthetic Microbial Mats in the Lake Huron Basin. Frontiers in Microbiology, 8, 2104. https://doi.org/10.3389/fmicb.2017.00791

Shih, P. M., Hemp, J., Ward, L. M., Matzke, N. J., & Fischer, W. W. (2017). Crown group Oxyphotobacteria postdate the rise of oxygen. Geobiology, 15(1), 19–29. https://doi.org/10.1111/gbi.12200

Sicora, C. I., Ho, F. M., Salminen, T., Styring, S., & Aro, E.-M. (2009). Transcription of a "silent" cyanobacterial psbA gene is induced by microaerobic conditions. Biochimica Et Biophysica Acta, 1787(2), 105–112. https://doi.org/10.1016/j.bbabio.2008.12.002

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology, 7(1), 539–539. https://doi.org/10.1038/msb.2011.75

Sievert, S. M., Scott, K. M., Klotz, M. G., Chain, P. S. G., Hauser, L. J., Hemp, J., et al. (2008). Genome of the epsilonproteobacterial chemolithoautotroph Sulfurimonas denitrificans. Applied and Environmental Microbiology, 74(4), 1145–1156. https://doi.org/10.1128/AEM.01844-07

Sivarajah, B., Paterson, A. M., Rühland, K. M., Köster, D., Karst-Riddoch, T., & Smol, J. P. (2018). Diatom responses to 20th century shoreline development and climate warming in three embayments of Georgian Bay, Lake Huron. Journal of Great Lakes Research, 44(6),

1339–1350. https://doi.org/10.1016/j.jglr.2018.09.004

Soo, R. M., Hemp, J., Parks, D. H., Fischer, W. W., & Hugenholtz, P. (2017). On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. Science, 355(6332), 1436–1440. https://doi.org/10.1126/science.aal3794

Stal, L. J. (2012). Cyanobacterial Mats and Stromatolites. In B. A. Whitton (Ed.), Ecology of Cyanobacteria II (pp. 65–125). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_4

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities, 6(9), 1653–1664. https://doi.org/10.1038/ismej.2012.22

Takai, K., Campbell, B. J., Cary, S. C., Suzuki, M., Oida, H., Nunoura, T., et al. (2005). Enzymatic and genetic characterization of carbon and energy metabolisms by deep-sea hydrothermal chemolithoautotrophic isolates of Epsilonproteobacteria. Applied and Environmental Microbiology, 71(11), 7310–7320. https://doi.org/10.1128/AEM.71.11.7310-7320.2005

Tamagnini, P., Leitão, E., Oliveira, P., Ferreira, D., Pinto, F., Harris, D. J., et al. (2007). Cyanobacterial hydrogenases: diversity, regulation and applications. FEMS Microbiology Reviews, 31(6), 692–720. https://doi.org/10.1111/j.1574-6976.2007.00085.x

Tandeau de Marsac, N., & Houmard, J. (1993). Adaptation of cyanobacteria to environmental stimuli: new steps towards molecular mechanisms. FEMS Microbiology Reviews, 104, 119–190.

Tremblay, P.-L., & Lovley, D. R. (2012). Role of the NiFe hydrogenase Hya in oxidative stress defense in Geobacter sulfurreducens. Journal of Bacteriology, 194(9), 2248–2253. https://doi.org/10.1128/JB.00044-12

Vignais, P. M., & Billoud, B. (2007). Occurrence, Classification, and Biological Function of Hydrogenases: An Overview. Chemical Reviews, 107(10), 4206–4272. https://doi.org/10.1021/cr050196r

Violle, C., Nemergut, D. R., Pu, Z., & Jiang, L. (2011). Phylogenetic limiting similarity and competitive exclusion. Ecology Letters, 14(8), 782–787. https://doi.org/10.1111/j.1461-0248.2011.01644.x

Visscher, P. T., Reid, R. P., & Bebout, B. M. (2000). Microscale observations of sulfate reduction: Correlation of microbial activity with lithified micritic laminae in modern marine stromatolites. Geology, 28, 919–922.

Voorhies, A. A. (2014, January 24). Investigation of microbial interactions and ecosystem dynamics in a low O2 cyanobacterial mat.

Voorhies, A. A., Biddanda, B. A., Kendall, S. T., Jain, S., Marcus, D. N., Nold, S. C., et al. (2012). Cyanobacterial life at low O2: community genomics and function reveal metabolic versatility and extremely low diversity in a Great Lakes sinkhole mat. Geobiology, 10(3), 250–267. https://doi.org/10.1111/j.1472-4669.2012.00322.x

Waldbauer, J., Zhang, L., Rizzo, A., & Muratore, D. (2017). diDO-IPTL: A Peptide-Labeling Strategy for Precision Quantitative Proteomics. Analytical Chemistry, 89(21), 11498–11504. https://doi.org/10.1021/acs.analchem.7b02752

West, P. T., Probst, A. J., Grigoriev, I. V., Thomas, B. C., & Banfield, J. F. (2018). Genome-reconstruction for eukaryotes from complex natural microbial communities. Genome Research, 28(4), 569–580. https://doi.org/10.1101/gr.228429.117

Wick, R. R., Schultz, M. B., Zobel, J., & Holt, K. E. (2015). Bandage: interactive visualization of de novo genome assemblies. Bioinformatics, 31, 3350–3352. https://doi.org/10.1093/bioinformatics/btv383/-/DC1

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wilmeth, D. T., Johnson, H. A., Stamps, B. W., Berelson, W. M., Stevenson, B. S., Nunn, H. S., et al. (2018). Environmental and Biological Influences on Carbonate Precipitation Within Hot Spring Microbial Mats in Little Hot Creek, CA. Frontiers in Microbiology, 9, 1–13. https://doi.org/10.3389/fmicb.2018.01464

Woese, C. R. (1987). Bacterial Evolution. Microbiological Reviews, 51(2), 221–271.

Wyman, M., & Fay, P. (1986). Underwater Light Climate and the Growth and Pigmentation of Planktonic Blue-Green Algae (Cyanobacteria) I. The Influence of Light Quantity. Proceedings of the Royal Society of London, 227(1248), 367–380.

Xiong, J., Fischer, W. M., Inoue, K., Nakahara, M., & Bauer, C. E. (2000). Molecular evidence for the early evolution of photosynthesis. Science, 289(5485), 1724–1730.

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D. R., Bork, P., & Patil, K. R. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. Proceedings of the National Academy of Sciences of the United States of America, 112(20), 6449–6454. https://doi.org/10.1073/pnas.1421834112

Zhong, Y., Notaro, M., & Vavrus, S. J. (2019). Spatially variable warming of the Laurentian Great Lakes: an interaction of bathymetry and climate. Climate Dynamics, 52(9), 5833–5848. https://doi.org/10.1007/s00382-018-4481-z

4.9 Tables and Figures

Table 4.1 Sequencing and assembly results of metagenomic datasets.

Mat type	IMG dataset	Sample	Year	Month	Day	N50 (bp)	Total number of raw reads	Assembly length (bp)	Number of contigs	Number of prokaryote contigs > 1000 bp	Number of prokaryote MAG bins
Finger	3300033517	MIS.2015.027A3	2015	June	4	1026	167515048	692629062	1017918	61028	35
Finger	3300033507	MIS.2016.030B	2016	June	9	859	115511662	718338751	1145025	47169	23
Flat A	3300033508	MIS.2016.033B	2016	June	9	1346	121752032	909459467	1165877	158869	67
Flat A	3300033509	MIS.2016.060D	2016	July	19	1244	144984256	893189390	1252970	147732	60
Giraffe pit	3300033518	MIS.2016.213	2016	July	28	1348	111325384	701445461	933962	125804	48
Giraffe	3300033519	MIS.2016.221	2016	July	28	1309	108194844	579145956	780602	102941	38
Giraffe	3300033512	MIS.2016.G.00.C	2016	July	28	1493	108741556	644280742	811842	105112	44
Giraffe	3300033549	MIS.2016.G.06.C	2016	July	28	1003	175140662	1552361899	2352229	261951	62
Flat D	3300033514	MIS.2016.Y.00.C	2016	July	27	1158	102856012	833759304	1176210	146846	42
Flat D	3300033520	MIS.2016.Y.06.C	2016	July	27	972	102299086	960330577	1493067	158433	37
Finger	3300033515	MIS.2017.036B	2017	June	2	1378	96683920	103159288	140253	15354	12
Finger	3300033516	MIS.2017.037B	2017	June	2	1811	117807040	508884368	593262	87605	55
Flat A	3300033564	MIS.2016.255	2016	July	28	1586	130640186	505874628	628644	88597	48
Flat white	3300033562	MIS.2016.275	2016	July	28	1527	141353982	1385456446	1676959	256823	115
Flat white	3300033558	MIS.2016.280	2016	July	28	956	84363316	580554833	953497	81968	35
Flat A	3300033559	MIS.2016.135B	2016	July	25	1549	112966880	470248132	593572	76136	39
Total							2037133852	12132035487		1937751	769
Average						1309.5	119831403.1	713649146.3	990374.9	113985.4	45.2

Table 4.2 Estimates of S and O fluxes and photosynthetic rates measured *in situ* in flat purple, giraffe, and white mat, and *ex situ* in finger and flat purple mat.

	In situ 30-57 µmol photons m^{-2} s^{-1}			**Ex situ 58-79 µmol photons m^{-2} s^{-1}**	
	Purple	**Giraffe**	**White**	**Fingers**	**Purple**
Net OxyPhoto (µmol O_2 m^{-2} s^{-1})	1.44 to 2.22	1.79 to 2.36	0	1.26 to 1.59	1.39 to 2.86
AnoxyPhoto (µmol S_{tot} m^{-2} s^{-1})	ND	ND	ND	0.09 to 0.21	0.93 to 2.49
O_2 flux into water column (µmol O_2 m^{-2} s^{-1})	0.94 to 1.51	1.19 to 1.66	-0.04 to -0.26	0.67 to 1.83	0.35 to 2.12
Stot upward flux into mat (µmol S_{tot} m^{-2} s^{-1})	0.26 to 0.46	0.29 to 0.55	0.07 to 0.91	0.06 to 0.43	0.41 to 0.92

Table 4.3 Presence/absence and grouping of key genes in cyanobacterial genomes: *psbA*, *sqr, nif, cpe, cpc*

Metabolism	Gene	*Phormidium*	*Planktothrix*	*Pseudanabaena*	*Spirulina*
Photosynthesis	*psbA* group 2 (anoxic)	X		X	
	psbA group 3 (microaerobic)	X	X	X	
	psbA group 4 (oxic)	X	X	X	X
Pigments	phycoerythrin	X	X		
	phycocyanin	X	X	X	X
	phycoerythrocyanin				X
Sulfur cycling	*sqrI* (high H_2S affinity)	X		X	X
	sqrII (sulfide detoxification)	X	X		
	sqrVI (low H_2S affinity)			X	
Nitrogen fixation	*nifHDK*	X		X	

Table 4.4 This and next 2 pages: Completion and redundancy estimates for MAG bins of interest. Asterisked * bins indicate representative bin for further analyses.

Organism	Bins	Sample description	Completion (%)	Redundancy == CheckM Contamination %	Strain heterogeneity (%)	Contamination %	Genome size (bp)
Phormidium	107396_bin.15.fa	Flat A	64.65	1.31	85.71	0.19	4716872
	107397_bin.3.fa	Flat white by giraffe	46.85	0.98	80	0.20	2969681
	* new_117907_bin_8_phorm.fa	Finger	92.49	5.76	15.15	4.89	5922938
	new_117910_bin_1_46_cyano.fa	Flat A	67.71	17.21	72.28	4.77	6364074
	new_117915_bin_25_phorm.fa	Flat D	84.31	6.46	12.5	5.65	5514688
	new_117917_bin_2_4_phorm.fa	Finger	87.82	4.91	58.62	2.03	6793430
Planktothrix	new_117906_bin_15_plank.fa	Finger	75.58	2.67	43.75	1.50	4017934
	* 117910_bin.49.fa	Flat A	93.36	1.99	29.41	1.40	5514186
	117913_bin.38.fa	Giraffe	48.86	0.33	50	0.17	2704689
	117918_bin.54.fa	Finger	89.85	3.13	44.44	1.74	5236627
Pseudanabaena	new_107395_bin_27_cyano.fa	Flat A	90.5	54.66	0	54.66	7812391
	* new_107396_bin_67_cyano.fa	Flat white by giraffe	96.78	2.79	7.14	2.59	4521637
	new_107397_bin_11_cyano.fa	Flat white by giraffe	95.36	5.54	28.12	3.98	4102370
	new_107398_bin_31_cyano.fa	Flat A	91.47	8.1	6.82	7.55	4164679
	117908_bin.2.fa	Finger	43.27	0.47	0	0.47	1774606
	new_117910_bin_35_cyano.fa	Flat A	91.57	9.28	6.45	8.68	4068924
	new_117911_bin_25_cyano.fa	Giraffe pit	93.1	10.97	0	10.97	4164513
	new_117912_bin_1_cyano.fa	Giraffe ridge	92.96	5.9	10	5.31	4204160
	new_117913_bin_4_cyano.fa	Giraffe	89.97	6.96	3.03	6.75	4003132
	new_117918_bin_29_cyano.fa	Finger	84.48	19.83	0	19.83	3992695
Spirulina	107395_bin_38_plus.fa	Flat A	98.67	0.56	66.67	0.19	4537598
	107398_bin_34_plus.fa	Flat A	98.44	2.11	69.23	0.65	4547421
	117906_bin.16.fa	Finger	44.52	0.22	0	0.22	1661841
	117908_bin_4_plus.fa	Finger	94.02	1.67	25	1.25	4309314
	* 117909_bin_66_plus.fa	Flat A	98.17	0.78	25	0.59	4528925
	117910_bin_33_plus.fa	Flat A	98.22	0.56	66.67	0.19	4577207
	117911_bin_9_plus.fa	Giraffe pit	98.44	0.33	50	0.17	4738927
	117912_bin_25_plus.fa	Giraffe ridge	98.67	1.89	93.33	0.13	4899338
	117913_bin_24_plus.fa	Giraffe	98.22	0.63	25	0.47	4530499
	117918_bin_38_plus.fa	Finger	98.44	0.78	60	0.31	4593709
Desulfo	107395_bin.19.fa	Flat A	69.82	2.35	0	2.35	2557356
	* 107398_bin.22.fa	Flat A	75.59	0.32	0	0.32	3230363

	117912_bin.8.fa	Giraffe ridge	76.21	1.97	40	1.18	3023799
	117913_bin.26.fa	Giraffe	94.64	23.78	90.24	2.32	4408069
	117915_bin.18.fa	Flat D	82.09	3.91	71.43	1.12	3536614
Desulfobacteraceae	107395_bin.45.fa	Flat A	50.2	0.97	100	0.00	2473442
	107396_bin.79.fa	Flat white by giraffe	73.04	3.1	71.43	0.89	3478934
	* 107397_bin.13.fa	Flat white by giraffe	90.75	3.58	75	0.90	5329193
	107398_bin.21.fa	Flat A	54.6	1.47	80	0.29	2621878
	117907_bin.1.fa	Finger	87.85	5.94	75	1.49	4809721
	117909_bin.21.fa	Flat A	76.68	1.67	80	0.33	3467608
	117910_bin.56.fa	Flat A	66.71	2.69	100	0.00	3443140
	117911_bin.5.fa	Giraffe pit	67.22	1.61	100	0.00	3023460
	117912_bin.18.fa	Giraffe ridge	45.62	1.79	100	0.00	2375759
	117913_bin.2.fa	Giraffe	47.8	1.96	100	0.00	2193258
	117915_bin.23.fa	Flat D	72.14	4.59	90	0.46	3732848
Gamma_bin_8	117911_bin.7.fa	Giraffe pit	62.1	3.09	13.33	2.68	2117028
	117912_bin.27.fa	Giraffe ridge	55.39	0.84	16.67	0.70	1712752
	* 107398_bin.8.fa	Flat A	85.58	2.49	16.67	2.07	3161292
Gamma_bin_15	107395_bin.11.fa	Flat A	39.39	0	0	0.00	1833894
	107396_bin.57.fa	Flat white by giraffe	63.14	1.44	36.36	0.92	2211128
	107397_bin.35.fa	Flat white by giraffe	77.92	5.28	66.67	1.76	3058726
	107398_bin.37.fa	Flat A	41.66	0	0	0.00	1391670
	117907_bin.2.fa	Finger	65.17	1.06	60	0.42	2425313
	117909_bin.39.fa	Flat A	54.23	3.45	100	0.00	2377308
	117910_bin.22.fa	Flat A	30.85	0	0	0.00	1750427
	117911_bin.34.fa	Giraffe pit	35.82	0	0	0.00	1610201
	* 117915_bin.15.fa	Flat D	75.29	3.8	65.22	1.32	2835298
	117918_bin.53.fa	Finger	54.28	5.86	100	0.00	2642412
Campylobacterales_bin_11	107397_bin.16.fa	Flat white by giraffe	90.16	2.46	42.86	1.41	1984893
	* 107396_bin.11.fa	Flat white by giraffe	96.11	2.87	25	2.15	2416355
Helicobacteraceae_bin_66	107396_bin.66.fa	Flat white by giraffe	78.23	7.45	88.46	0.86	1682526
Sulfurimonas_bin_83	107396_bin.83.fa	Flat white by giraffe	86.97	2.24	14.29	1.92	1923550
Campylobacterales_bin_108	107396_bin.108.fa	Flat white by giraffe	89.55	2.87	22.22	2.23	1768385
Campylobacterales_bin_21	107397_bin.21.fa	Flat white by giraffe	68.93	6.01	82.35	1.06	1504870

Sulfurospirillum_bin_34	117906_bin.34.fa	Finger	69.93	0.69	75	0.17	1316208
Sulfurospirillum_bin_16	117908_bin.16.fa	Finger	52.99	3.18	50	1.59	1101348
Sulfurimonas_bin_40	117909_bin.40.fa	Flat A	75.04	3.48	5.26	3.30	1667688

Figure 4.1 Distinct mat morphotypes observed in MIS from 2015-2017. A: Flat purple mat dominated by *Phormidium*. B: Flat purple mat dominated by *Spirulina* and *Pseudanabaena*. C: Finger mat. D: Giraffe mat. E: White mat.

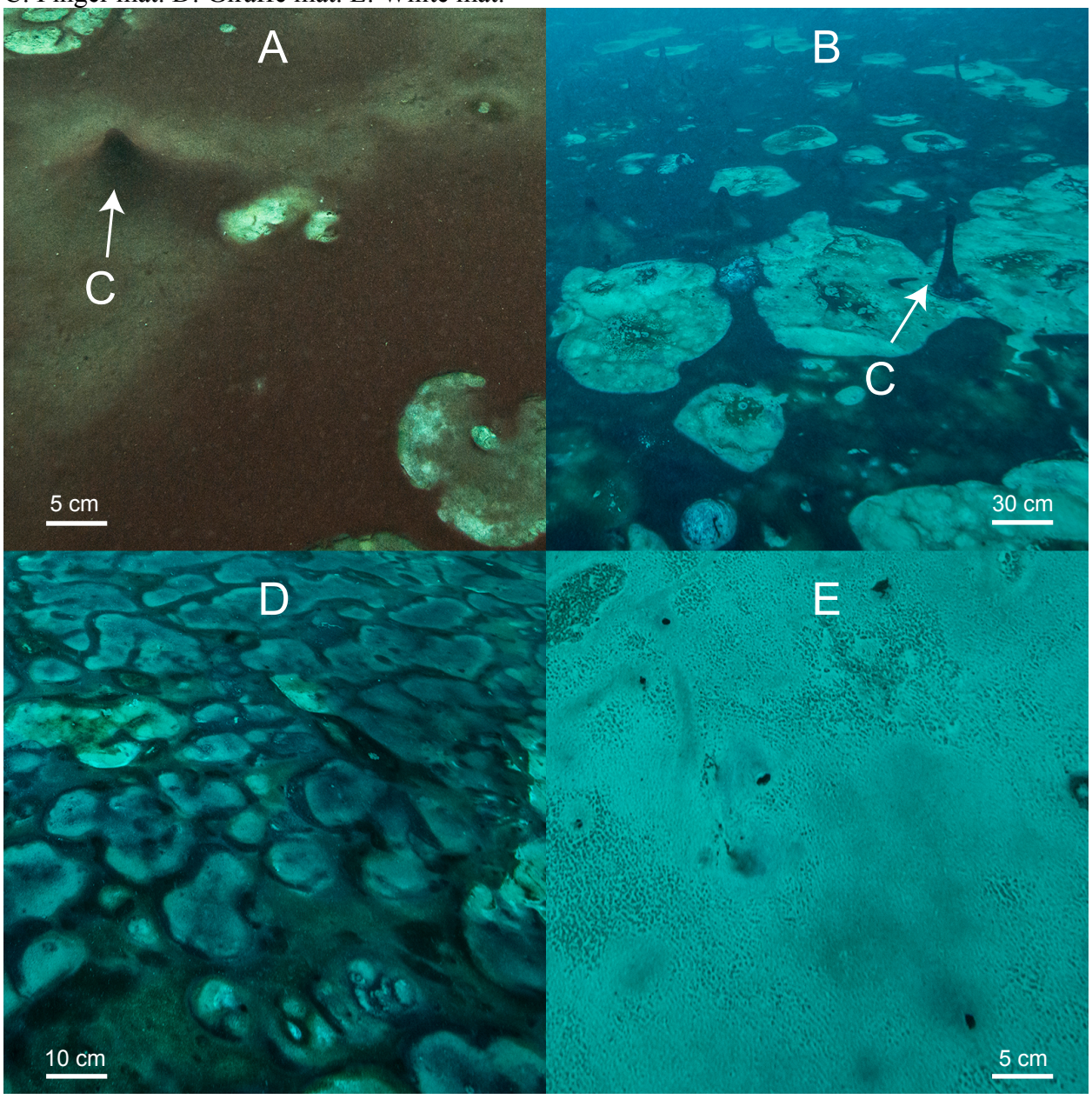


Figure 4.2 The relative abundances of 16S rRNA genes belonging to cyanobacteria, sulfur cycling bacteria (Gammaproteobacteria and Epsilonproteobacteria), sulfate-reducing bacteria (Deltaproteobacteria), and other abundant bacterial clades in mats. Samples are grouped by morphological appearance (finger, white, giraffe, flat), and labeled by collection year, month, and name.

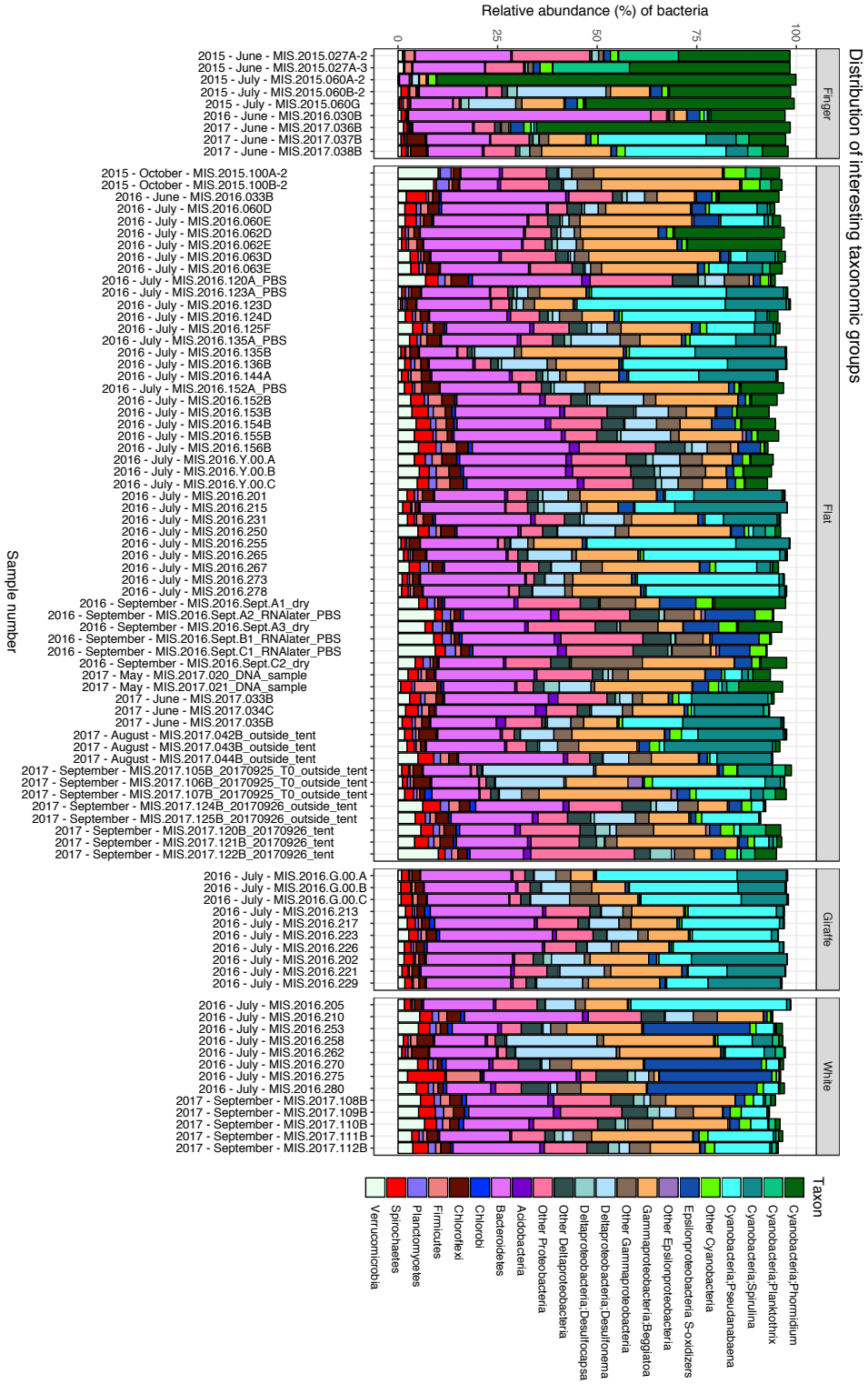


Figure 4.3 Dendrogram classifying the flat mat samples by the abundance of 16S rRNA genes from various cyanobacterial taxa. Only the abundances of cyanobacterial nodes were used in generating a Bray-Curtis distance matrix, which was the input for a dendrogram. In black are all samples in which the relative abundance of *Spirulina* and *Pseudanabaena* are more than 45% of the sequences assigned to *Spirulina, Pseudanabaena, Phormidium*, and *Planktothrix.* In red are all samples in which *Phormidium* was more abundant.

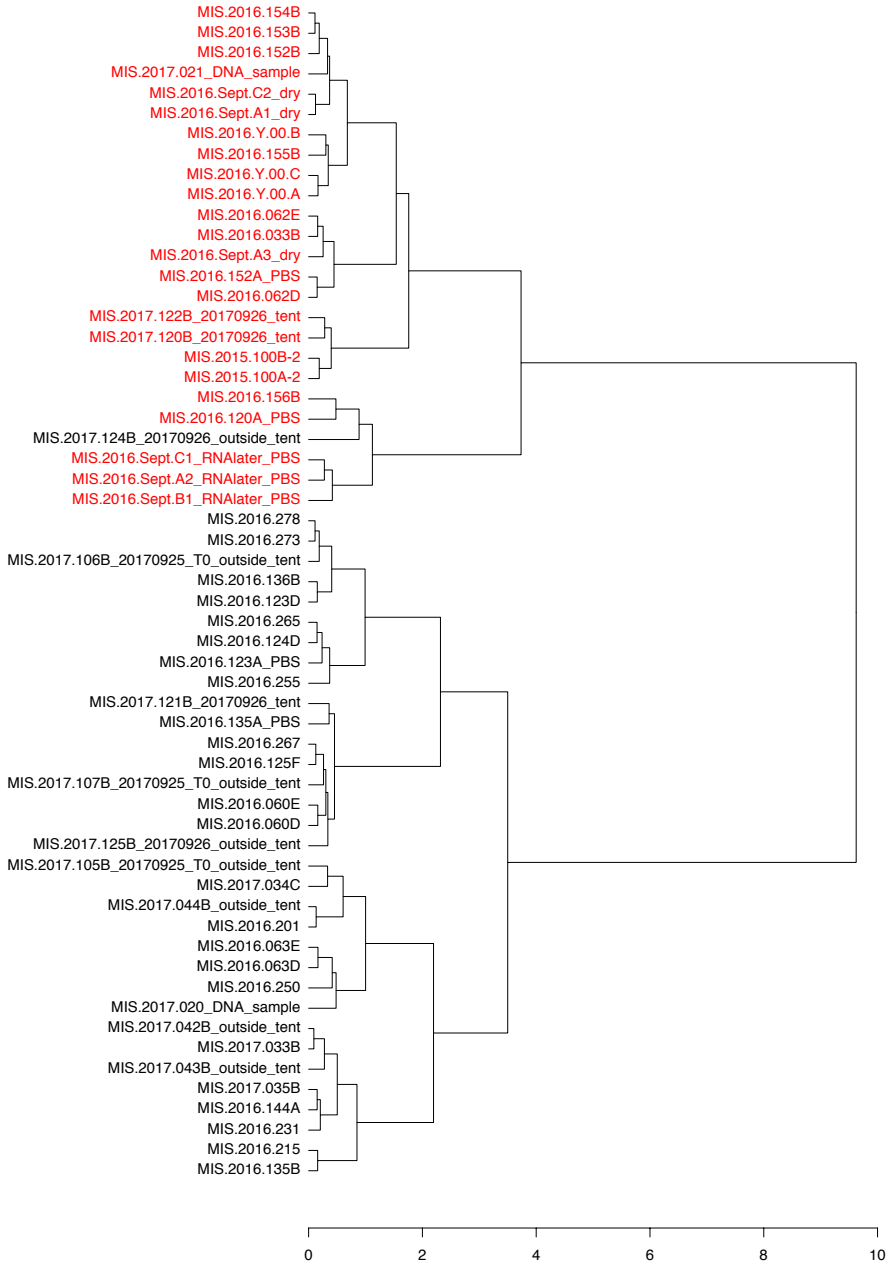


Figure 4.4 Next page: Key bacterial MAGs in each mat morphotype. A: Finger mat. B: Giraffe mat. C: White mat. Phormidium-dominated flat purple mat and Spirulina and Pseudanabaena-dominated flat purple mat are similar to A and B respectively. Important genes in biogeochemical cycling are shown. Oxygen-cycling genes are in blue and include: *psbA* photosystem II D1 protein, groups 2, 3, and 4; *cpcAB* phycocyanin and *cpeAB* phycoerythrin; cytochrome cbb3, bd-quinol, and c oxidases *cbb3, cyd*, and *cox* respectively. Sulfur-cycling genes are in green shades and include: *sqr* sulfide quinone reductase, types I, II, IV, and VI; *sat, apr, dsr* for dissimilatory sulfate reduction (SRBs) or sulfide oxidation (gSOs); *sox* for thiosulfate oxidation; and *psr* for polysulfide oxidation. Nitrogen-cycling genes are in purple and include: *nifHDK* nitrogenase; *narGH, nirA, narB* for nitrate reduction; *napAB* and *nrfAH* for nitrate reduction; denitrification with *napAB, nirS, norBC*, and *nosZ*; [NiFe] hydrogenases of groups I, II, III, and IV are NiFe *hyd, hya, hox*, and *hyf*, respectively. Transporters are in small, unlabeled rectangles and include sulfate, nitrate, molybdate, tungstate, phosphate, zinc, cobalt, nickel, manganese, iron, lipopolysaccharides, and phospholipids.

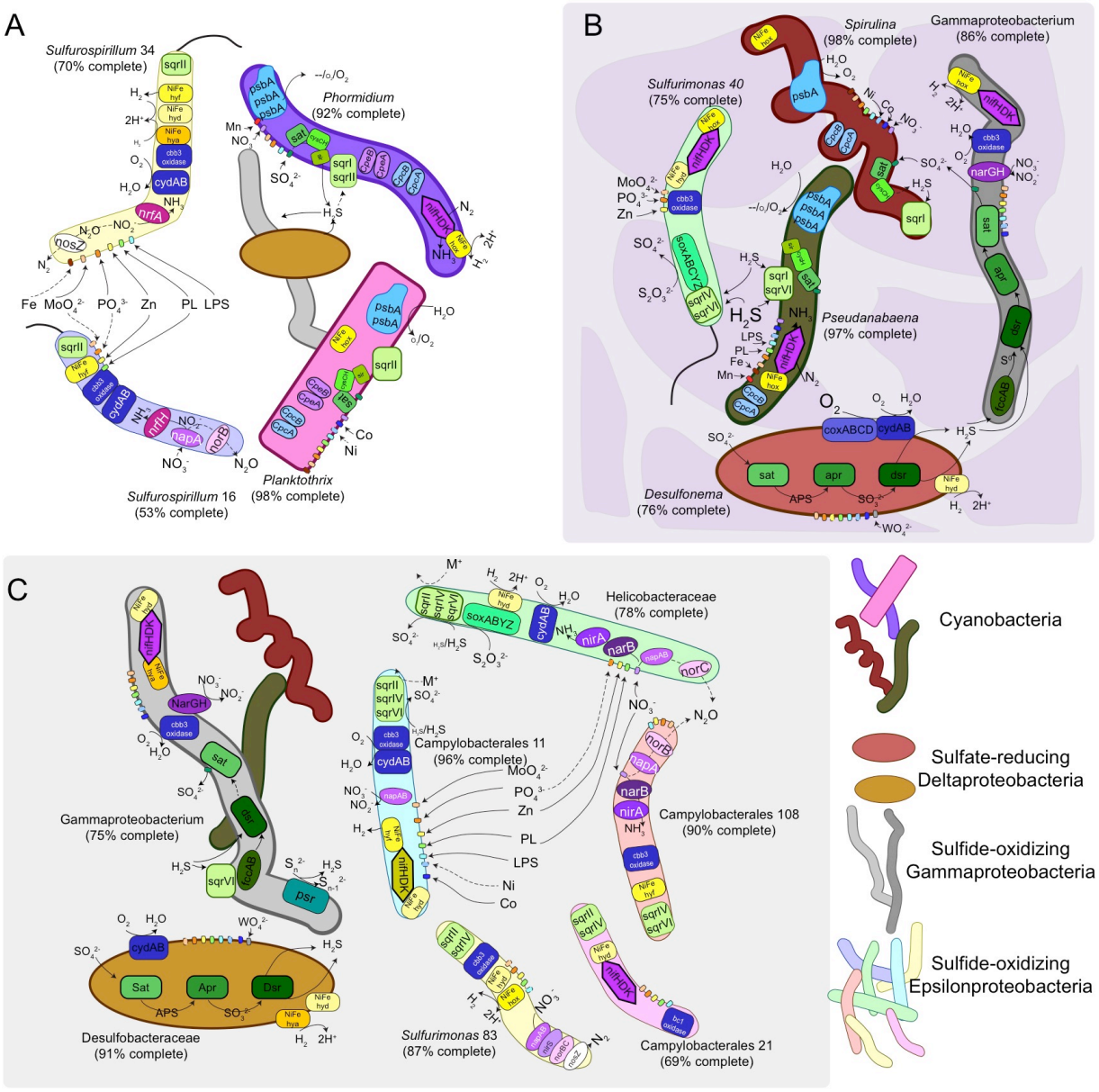

A
Sulfurospirillum 34
(70% complete)
Phormidium
(92% complete)
Sulfurospirillum 16
(53% complete)
Planktothrix
(98% complete)
B
Spirulina
(98% complete)
Gammaproteobacterium
(86% complete)
Sulfurimonas 40
(75% complete)
Pseudanabaena
(97% complete)
Desulfonema
(76% complete)
C
Helicobacteraceae
(78% complete)
Campylobacterales 11
(96% complete)
Campylobacterales 108
(90% complete)
Gammaproteobacterium
(75% complete)
Desulfobacteraceae
(91% complete)
Sulfurimonas 83
(87% complete)
Campylobacterales 21
(69% complete)
Cyanobacteria
Sulfate-reducing
Deltaproteobacteria
Sulfide-oxidizing
Gammaproteobacteria
Sulfide-oxidizing
Epsilonproteobacteria

Figure 4.5 Comparative dendrograms of samples with paired 16S rRNA gene profiles and proteomics profiles. Samples are colored by mat type: white mat is black, giraffe mat is red, finger mat is green, flat-P mat is light blue, and flat-SP mat is dark blue. The dendrogram of the Bray-Curtis dissimilarity matrix on the relative abundances of cyanobacterial 16S rRNA genes (A) clusters samples by abundances of *Phormidium, Spirulina,* and *Pseudanabaena*. The proteomic dendrogram based off a Gower dissimilarity matrix on the proteins' log_2-normalized abundances (B) does not preserve the grouping as seen in the 16S rRNA gene-based dendrogram.

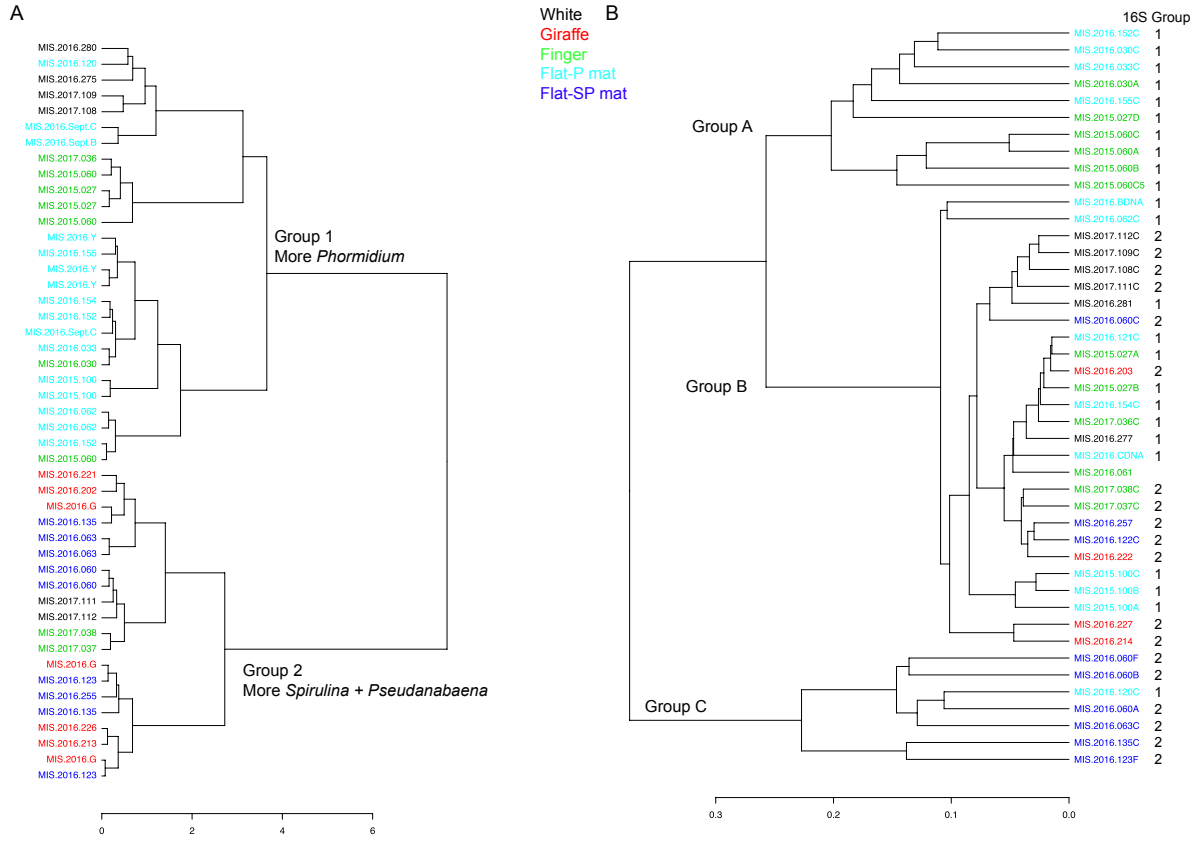


Figure 4.6 Heatmap of significantly differentially abundant proteins between flat-*Phormidium* mat and flat-*Spirulina* and *Pseudanabaena* mat, observed in specific taxa/bins. These proteins belonging to betaproteobacterial sulfide oxidizers (Betaproteobacteria SO), deltaproteobacterial sulfate reducers (SRB), gamamproteobacterial sulfide oxidizers, *Planktothrix*, *Phormidum*, *Pseudanabaena*, and *Spirulina*, are plotted by metabolic function and gene name. Proteins with higher weighted mean log_2-normalized abundances are plotted in warm grey colors, whereas proteins with lower abundances are plotted in cool grey colors. Proteins whose weighted means are larger than their standard deviations are plotted with black borders, whereas other proteins have weighted means are less than their standard deviations and are variably represented in samples of the same mat type.

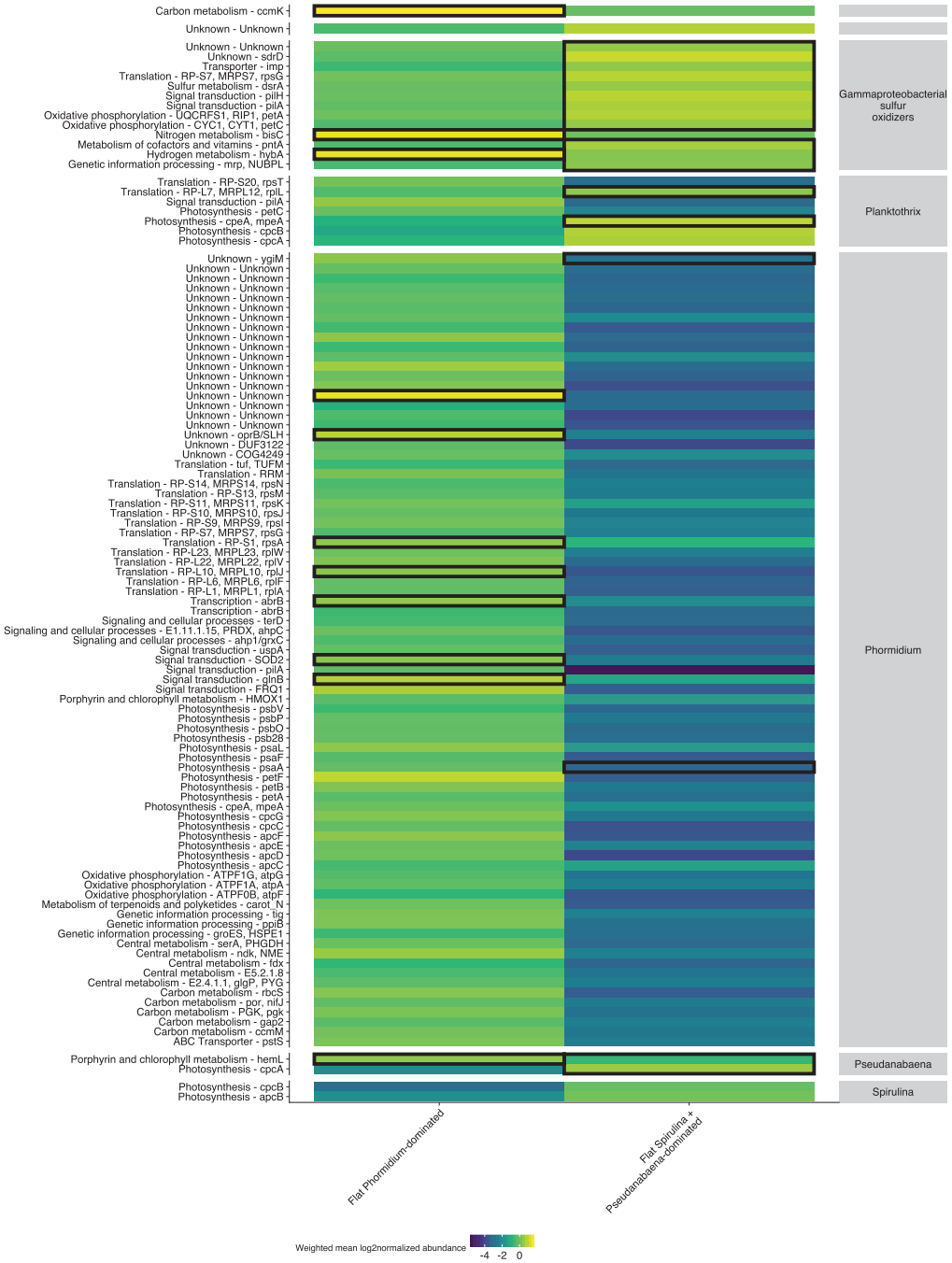
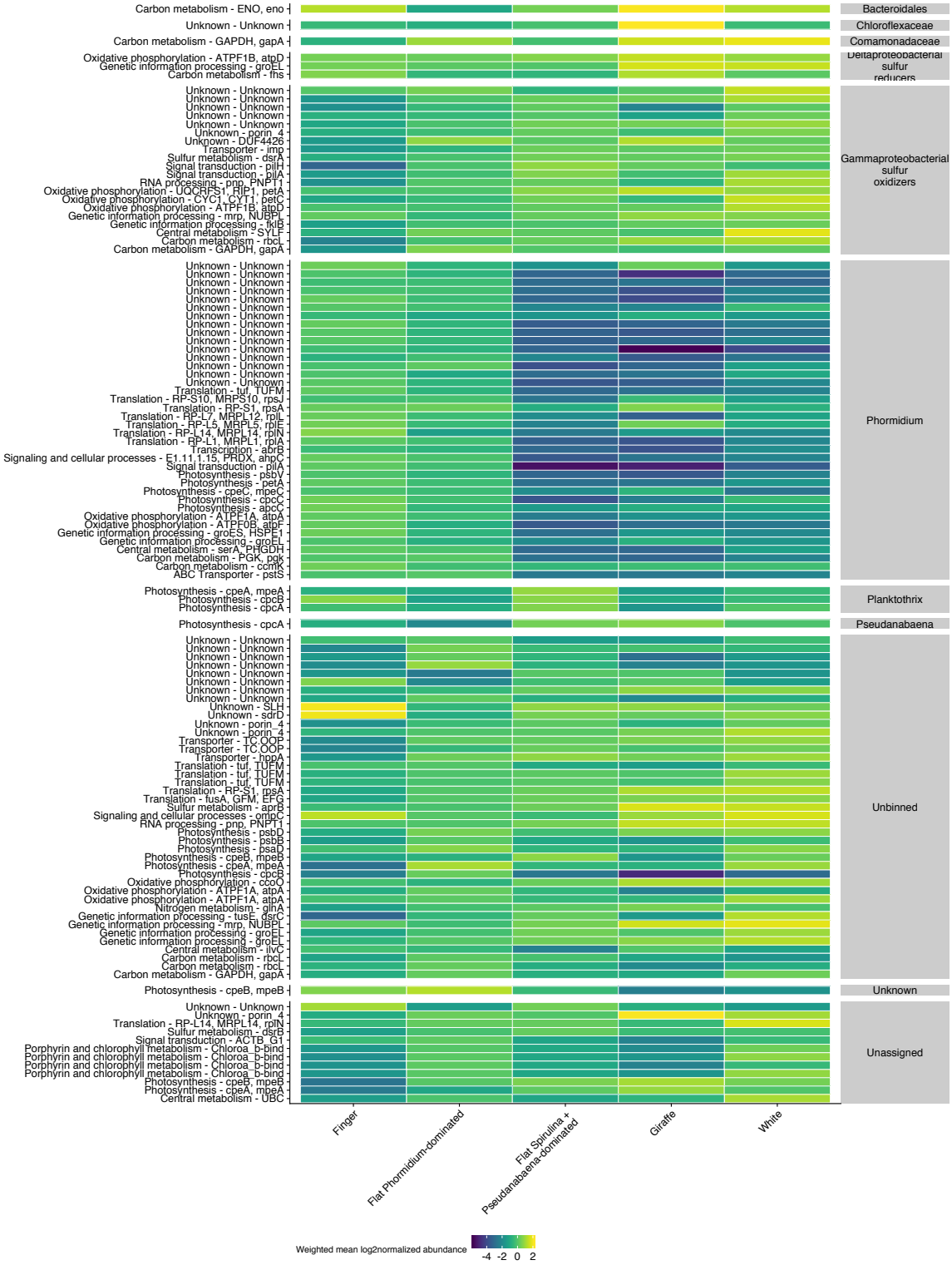


Figure 4.7 Heatmap of significantly differentially abundant proteins in mat morphotypes grouped by taxa/bin and metabolic function. Proteins with higher weighted mean log_2-normalized abundances are plotted in warm grey colors, whereas proteins with lower abundances are plotted in cool grey colors.



4.10 Appendix C

Supplemental Information

Table SI 4.1 This and next 2 pages: List of samples by mat type evaluated with 16S, metagenomics, and quantitative community proteomics.

Mat type	Sample	Year	Month	16S rRNA gene sequenced	Meta-proteomics	Meta-genomics
Finger	MIS.2015.027	2015	June	2	3	1
Finger	MIS.2015.060	2015	July	3	4	1
Finger	MIS.2016.030	2016	June	1	2	1
Finger	MIS.2016.061	2016	July		1	
Finger	MIS.2017.036	2017	June	1	1	1
Finger	MIS.2017.037	2017	June	1	1	1
Finger	MIS.2017.038	2017	June	1	1	
Flat	MIS.2015.100	2015	October	2	3	
Flat	MIS.2016.033	2016	June	1	1	1
Flat	MIS.2016.060	2016	July	2	4	1
Flat	MIS.2016.063	2016	July	2	1	
Flat	MIS.2016.120	2016	July	1	1	
Flat	MIS.2016.121	2016	July		1	
Flat	MIS.2016.122	2016	July		1	
Flat	MIS.2016.123	2016	July	2	1	
Flat	MIS.2016.124	2016	July	1		
Flat	MIS.2016.125	2016	July	1		
Flat	MIS.2016.135	2016	July	2	1	1
Flat	MIS.2016.136	2016	July	1		
Flat	MIS.2016.144	2016	July	1		
Flat	MIS.2016.156	2016	July	1		
Flat	MIS.2016.201	2016	July	1		
Flat	MIS.2016.215	2016	July	1		
Flat	MIS.2016.231	2016	July	1		
Flat	MIS.2016.250	2016	July	1		
Flat	MIS.2016.255/ MIS.2016.257	2016	July	1	1	1
Flat	MIS.2016.265	2016	July	1		
Flat	MIS.2016.267	2016	July	1		
Flat	MIS.2016.273	2016	July	1		
Flat	MIS.2016.278	2016	July	1		
Flat	MIS.2016.062	2016	July	2	1	
Flat	MIS.2016.152	2016	July	2	1	
Flat	MIS.2016.153	2016	July	1		
Flat	MIS.2016.154	2016	July	1	1	
Flat	MIS.2016.155	2016	July	1	1	
Flat	MIS.2016.Y.00	2016	July	3		1
Flat	MIS.2016.Sept.A	2016	September	3		
Flat	MIS.2016.Sept.B	2016	September	1	1	
Flat	MIS.2016.Sept.C	2016	September	2	1	
Flat	MIS.2017.020	2017	May	1		
Flat	MIS.2017.021	2017	May	1		
Flat	MIS.2017.033	2017	June	1		
Flat	MIS.2017.034	2017	June	1		
Flat	MIS.2017.035	2017	June	1		
Flat	MIS.2017.042	2017	August	1		
Flat	MIS.2017.043	2017	August	1		
Flat	MIS.2017.044	2017	August	1		
Flat	MIS.2017.105	2017	September	1		
Flat	MIS.2017.106	2017	September	1		

Flat	MIS.2017.107	2017	September	1		
Flat	MIS.2017.124	2017	September	1		
Flat	MIS.2017.125	2017	September	1		
Flat	MIS.2017.120	2017	September	1		
Flat	MIS.2017.121	2017	September	1		
Flat	MIS.2017.122	2017	September	1		
Giraffe	MIS.2016.G.00	2016	July	3		1
Giraffe	MIS.2016.213/ MIS.2016.214	2016	July	1	1	1
Giraffe	MIS.2016.217	2016	July	1		
Giraffe	MIS.2016.223	2016	July	1		
Giraffe	MIS.2016.226/ MIS.2016.227	2016	July	1	1	
Giraffe	MIS.2016.202/ MIS.2016.203	2016	July	1	1	
Giraffe	MIS.2016.221/ MIS.2016.222	2016	July	1	1	1
Giraffe	MIS.2016.229	2016	July	1		
White	MIS.2016.205	2016	July	1		
White	MIS.2016.210	2016	July	1		
White	MIS.2016.253	2016	July	1		
White	MIS.2016.258	2016	July	1		
White	MIS.2016.262	2016	July	1		
White	MIS.2016.270	2016	July	1		
White	MIS.2016.275/ MIS.2016.277	2016	July	1	1	1
White	MIS.2016.280/ MIS.2016.281	2016	July	1	1	1
White	MIS.2017.108	2017	September	1	1	
White	MIS.2017.109	2017	September	1	1	
White	MIS.2017.110	2017	September	1		
White	MIS.2017.111	2017	September	1	1	
White	MIS.2017.112	2017	September	1	1	

Table SI 4.2 Genes of interest identified in epsilonproteobacterial sulfide-oxidizing MAGs. X indicates presence of complete gene. Asterisk * indicates partial set of genes for metabolism.

	MAG	Campylobacterales_bin_11	Helicobacteraceae_bin_66	Sulfurimonas_bin_83	Campylobacterales_bin_108	Campylobacterales_bin_21	Sulfurospirillum_bin_34	Sulfurospirillum_bin_16	Sulfurimonas_bin_40
	Taxonomy	Campylobacterales	Helicobacteraceae	Sulfurimonas	Campylobacterales	Campylobacterales	Sulfurospirillum	Sulfurospirillum	Sulfurimonas
SQR type	**II**	X	X	X		X	X	X	
	IV	X	X	X	X	X			X
	VI	X	X		X				X
	Sox		*						*
Nitrogen metabolism	***nifHDK***	X			X	X			X
	Denitrification		*	X	*		*	*	
	Assimilatory	*	X		X				
	Dissimilatory						*	X	
Cytochrome	***ccb3***	X		X	X		X	*	X
	bd	X	X				X	X	
	bc1					X			
Transporters	**Nitrate**		X		X				
	Molybdate	X		*	X	*	X	*	*
	Phosphate	X	*	X	X	X	X	*	X
	Phospholipids	X	X	X		X	X	X	
	Lipopolysaccharides	X		*	X	X	X		
	Zinc	X	X	X	X	X	X	X	X
	Cobalt	X							
	Nickel	*							
	Iron						*		
Flagellar Assembly							X	*	*
Hydrogen Metabolism	**group 1**	X	X	X		X	X		X
	group 2						X		
	group 3a								
	group 3b								
	group 3d			X					X
	group 4	X			X		X	X	

Table SI 4.3 Next 25 pages: Proteins that were not significantly differentially abundant between mat morphotypes, and were observed in at least 3 mat types.

	Log2-normalized abundance in									
Protein	Finger	Flat-P	Flat-SP	Giraffe	White	Taxonomy	Category	Short name	Gene name	System
Ga0371427_10019131	1.18	NA	1.31	-0.03	0.05	Unbinned	KO:K02004	ABC.CD.P	putative ABC transport system permease protein	ABC Transporter
Ga0371428_1034062	-0.18	0.89	0.20	NA	NA	Unbinned	COG1840	afuA	ABC-type Fe3+ transport system, periplasmic component	ABC Transporter
Ga0371428_1038942	0.60	0.30	-0.58	NA	NA	Phormidium	KO:K01999	livK	branched-chain amino acid transport system substrate-binding protein	ABC Transporter
Ga0371436_10687121	NA	0.92	0.28	NA	-0.63	Unbinned	KO:K02067	mlaD, linM	phospholipid/cholesterol/gamma-HCH transport system substrate-binding protein	ABC Transporter
Ga0371436_10080863	0.57	3.54	-0.78	0.84	-3.69	Proteobacteria	KO:K02066	mlaE, linK	phospholipid/cholesterol/gamma-HCH transport system permease protein	ABC Transporter
Ga0371428_1021522	-0.74	-0.35	0.11	-0.23	-0.18	Gammaproteobacteria	KO:K02040	pstS	phosphate transport system substrate-binding protein	ABC Transporter
Ga0371429_10289202	NA	0.54	1.02	2.09	2.41	Pseudanabaena	KO:K02040	pstS	phosphate transport system substrate-binding protein	ABC Transporter
Ga0371429_10156581	1.65	1.49	NA	1.64	1.93	Unbinned	KO:K02040	pstS	phosphate transport system substrate-binding protein	ABC Transporter
Ga0371431_10574592	1.14	0.76	1.26	NA	NA	Planktothrix	KO:K02040	pstS	phosphate transport system substrate-binding protein	ABC Transporter
Ga0371428_1008485	-0.14	-0.21	-0.41	NA	NA	Chloroflexaceae	KO:K01251	ahcY	adenosylhomocysteinase	Amino acid metabolism
Ga0371428_1025892	-0.35	0.30	0.30	NA	0.31	Gammaproteobacteria	KO:K01251	ahcY	adenosylhomocysteinase	Amino acid metabolism
Ga0371428_1021442	0.02	-0.25	-2.14	NA	-1.19	Unbinned	KO:K01251	ahcY	adenosylhomocysteinase	Amino acid metabolism
Ga0371428_1025556	0.15	-0.11	-3.01	NA	NA	Phormidium	KO:K01251	ahcY	adenosylhomocysteinase	Amino acid metabolism
Ga0371428_1071264	-0.27	0.16	-1.08	NA	NA	Phormidium	KO:K00818	argD	acetylornithine aminotransferase	Amino acid metabolism
Ga0371477_1348463	0.19	1.25	-0.23	NA	NA	Geobacter	KO:K01940	argG, ASS1	argininosuccinate synthase	Amino acid metabolism
Ga0371434_10758125	0.17	0.64	NA	-1.24	-1.14	Desulfobacteraceae	KO:K03856	aroA	3-deoxy-7-phosphoheptulonate synthase	Amino acid metabolism
Ga0371437_10739572	0.66	-0.40	-3.30	NA	NA	Unbinned	KO:K01969	E6.4.1.4B	3-methylcrotonyl-CoA carboxylase beta subunit	Amino acid metabolism
Ga0371430_10296231	-2.69	-0.89	NA	NA	1.16	Bacteroidetes	KO:K00820	glmS, GFPT	glutamine---fructose-6-phosphate transaminase (isomerizing)	Amino acid metabolism
Ga0371428_1154203	1.08	0.83	0.46	0.61	1.12	Gammaproteobacterium_117915_bin_15	COG0493, COG0493	gltD	NADPH-dependent glutamate synthase beta chain or related oxidoreductase	Amino acid metabolism
Ga0371428_1162222	1.21	0.97	0.21	NA	1.28	Unbinned	COG0493	gltD	NADPH-dependent glutamate synthase beta chain or related oxidoreductase	Amino acid metabolism
Ga0371428_1013982	0.30	0.17	0.68	0.01	NA	Gammaproteobacteria	COG0493, COG1144, COG3349	gltD/porD	NADPH-dependent glutamate synthase beta subunit-like oxidoreductase/Pyruvate/2-oxoacid:ferredoxin oxidoreductase delta subunit/uncharacterized protein with NAD-binding domain and iron-sulfur cluster	Amino acid metabolism
Ga0371428_1010771	-0.85	0.18	0.29	0.85	1.33	Phormidium	COG0493, COG1144	gltD/porD	NADPH-dependent glutamate synthase beta subunit-like oxidoreductase/Pyruvate/2-oxoacid:ferredoxin oxidoreductase delta subunit	Amino acid metabolism

Ga0371429_10444373	0.65	0.04	-0.24	NA	1.48	Gammaproteobacteria	COG0493, COG1232, COG1144	gltD/porD/hemY	NADPH-dependent glutamate synthase beta subunit-like oxidoreductase/Protoporphyrinogen oxidase/Pyruvate/2-oxoacid:ferredoxin oxidoreductase delta subunit	Amino acid metabolism
Ga0371431_10224013	0.10	0.04	-1.01	NA	NA	Verrucomicrobiaceae	KO:K00789	metK	S-adenosylmethionine synthetase	Amino acid metabolism
Ga0371428_1148462	-0.02	-0.57	NA	NA	-0.56	Phormidium	KO:K00789	metK	S-adenosylmethionine synthetase	Amino acid metabolism
Ga0371428_1029051	NA	0.79	0.54	1.15	NA	Unbinned	KO:K03072	secD	preprotein translocase subunit SecD	Bacterial secretion system
Ga0371430_11686422	-0.09	0.36	-1.35	NA	0.50	Unbinned	KO:K01961	accC	acetyl-CoA carboxylase, biotin carboxylase subunit	Carbon metabolism
Ga0371428_1007462	-0.81	-0.01	0.06	NA	1.01	Gammaproteobacteria	KO:K01637	aceA	isocitrate lyase	Carbon metabolism
Ga0371428_1033043	NA	0.37	0.46	NA	0.92	Gammaproteobacteria	KO:K01682	acnB	aconitate hydratase 2 / 2-methylisocitrate dehydratase	Carbon metabolism
Ga0371429_10687486	-0.06	-0.30	-0.22	NA	-0.23	Comamonadaceae	KO:K01682	acnB	aconitate hydratase 2 / 2-methylisocitrate dehydratase	Carbon metabolism
Ga0371429_10665311	-1.04	-0.39	0.53	-0.81	NA	Unbinned	KO:K01682	acnB	aconitate hydratase 2 / 2-methylisocitrate dehydratase	Carbon metabolism
Ga0371430_11290381	-2.29	-0.21	-0.05	-0.63	0.62	Gammaproteobacterium_117915_bin_15	KO:K01682	acnB	aconitate hydratase 2 / 2-methylisocitrate dehydratase	Carbon metabolism
Ga0371428_1017936	-0.10	1.01	0.06	1.79	0.60	Desulfobacteraceae	KO:K14138	acsB	acetyl-CoA synthase	Carbon metabolism
Ga0371428_1084271	-0.62	1.97	1.14	NA	NA	Unbinned	KO:K14138	acsB	acetyl-CoA synthase	Carbon metabolism
Ga0371428_1097791	-1.29	0.54	0.28	1.40	1.39	Unbinned	KO:K14138	acsB	acetyl-CoA synthase	Carbon metabolism
Ga0371428_1058232	-0.12	0.13	-0.11	1.24	NA	Desulfobacteraceae	KO:K01895	ACSS, acs	acetyl-CoA synthetase	Carbon metabolism
Ga0371428_1161471	0.84	-0.70	0.84	0.95	NA	Desulfobacteraceae	KO:K01895	ACSS, acs	acetyl-CoA synthetase	Carbon metabolism
Ga0371428_1093253	-0.42	-0.66	0.40	0.38	0.81	Unbinned	KO:K01895	ACSS, acs	acetyl-CoA synthetase	Carbon metabolism
Ga0371428_1158462	0.90	-0.44	-0.18	1.59	0.77	Unbinned	KO:K01895	ACSS, acs	acetyl-CoA synthetase	Carbon metabolism
Ga0371428_1041777	0.67	-0.96	-3.32	NA	NA	Phormidium	KO:K01895	ACSS, acs	acetyl-CoA synthetase	Carbon metabolism
Ga0371427_10462513	-1.56	-0.09	-1.04	NA	NA	NA	KO:K01623	ALDO	fructose-bisphosphate aldolase, class I	Carbon metabolism
Ga0371428_1059043	-0.59	-0.55	0.50	NA	NA	Phormidium	KO:K03738	aor	aldehyde:ferredoxin oxidoreductase	Carbon metabolism
Ga0371429_11024881	0.01	0.18	1.33	1.29	NA	Spirulina	KO:K08696	ccmK	carbon dioxide concentrating mechanism protein CcmK	Carbon metabolism
Ga0371429_11067534	-1.85	1.51	-0.19	0.20	NA	Betaproteobacteria	COG4577	ccmK	carbon dioxide concentrating mechanism protein CcmK	Carbon metabolism
Ga0371429_10397192	NA	-0.70	1.40	2.73	1.69	Unbinned	KO:K08696	ccmK	carbon dioxide concentrating mechanism protein CcmK	Carbon metabolism
Ga0371428_1083423	0.59	0.10	-1.23	1.34	-1.39	Phormidium	KO:K08696	ccmK	carbon dioxide concentrating mechanism protein CcmK	Carbon metabolism
Ga0371428_1083425	-0.41	-0.03	-2.69	NA	-1.54	Phormidium	KO:K08698	ccmM	carbon dioxide concentrating mechanism protein CcmM	Carbon metabolism
Ga0371428_1017938	-0.80	0.07	-0.76	NA	NA	Desulfobacteraceae	KO:K00194	cdhD, acsD	acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit delta	Carbon metabolism
Ga0371428_1017935	-0.11	1.48	-0.18	NA	0.14	Desulfobacteraceae	KO:K00197	cdhE, acsC	acetyl-CoA decarbonylase/synthase, CODH/ACS complex subunit gamma	Carbon metabolism
Ga0371428_1017937	0.83	1.78	0.84	-0.21	NA	Desulfobacteraceae	KO:K00198	cooS, acsA	anaerobic carbon-monoxide dehydrogenase catalytic subunit	Carbon metabolism
Ga0371428_1065798	-0.32	1.35	-0.03	NA	NA	Desulfobacteraceae	KO:K00198	cooS, acsA	anaerobic carbon-monoxide dehydrogenase catalytic subunit	Carbon metabolism

Ga0371433_10792051	-0.86	0.83	-0.70	NA	NA	Unbinned	KO:K00382	DLD, lpd, pdhD	dihydrolipoamide dehydrogenase	Carbon metabolism
Ga0371428_1122321	-0.47	0.22	0.36	-0.05	0.73	Gammaproteobacteria	KO:K00626	E2.3.1.9, atoB	acetyl-CoA C-acetyltransferase	Carbon metabolism
Ga0371427_10935021	-0.08	0.37	-0.94	NA	NA	NA	KO:K01100	E3.1.3.37	sedoheptulose-bisphosphatase	Carbon metabolism
Ga0371427_10935022	-0.81	0.40	-0.78	NA	2.07	NA	KO:K01100	E3.1.3.37	sedoheptulose-bisphosphatase	Carbon metabolism
Ga0371428_1010121	-0.37	0.35	0.33	0.71	0.95	Gammaproteobacteria	KO:K01610	E4.1.1.49, pckA	phosphoenolpyruvate carboxykinase (ATP)	Carbon metabolism
Ga0371428_1095852	1.11	0.13	0.71	1.39	0.92	Unbinned	KO:K01689	ENO, eno	enolase	Carbon metabolism
Ga0371428_10693614	-0.05	0.55	-2.69	NA	NA	Phormidium	KO:K01689	ENO, eno	enolase	Carbon metabolism
Ga0371428_1139862	-1.71	0.32	0.13	-1.51	0.38	Gammaproteobacteria	KO:K03841	FBP, fbp	fructose-1,6-bisphosphatase I	Carbon metabolism
Ga0371428_1080931	-0.18	2.14	0.85	NA	NA	Desulfobacteraceae	KO:K00123	fdoG, fdhF, fdwA	formate dehydrogenase major subunit	Carbon metabolism
Ga0371428_1035722	0.85	1.61	0.63	0.28	0.96	Unbinned	KO:K01938	fhs	formate--tetrahydrofolate ligase	Carbon metabolism
Ga0371431_11444881	NA	2.03	0.57	-0.18	NA	Desulfobacteraceae	KO:K01938	fhs	formate--tetrahydrofolate ligase	Carbon metabolism
Ga0371429_102515718	0.59	NA	1.25	3.87	NA	Spirulina	KO:K00150	gap2	glyceraldehyde-3-phosphate dehydrogenase (NAD(P))	Carbon metabolism
Ga0371431_11310401	1.36	0.01	-0.44	NA	NA	Unbinned	KO:K00150	gap2	glyceraldehyde-3-phosphate dehydrogenase (NAD(P))	Carbon metabolism
Ga0371428_10321818	0.20	-0.35	-2.49	NA	-1.64	Phormidium	KO:K00150	gap2	glyceraldehyde-3-phosphate dehydrogenase (NAD(P))	Carbon metabolism
Ga0371476_12652433	-0.71	0.39	-0.66	NA	NA	Bacteroidales	KO:K00134	GAPDH, gapA	glyceraldehyde 3-phosphate dehydrogenase	Carbon metabolism
Ga0371427_10933292	-0.74	0.34	-0.59	NA	NA	Unbinned	KO:K00134	GAPDH, gapA	glyceraldehyde 3-phosphate dehydrogenase	Carbon metabolism
Ga0371428_1026342	-1.04	0.40	-0.12	-1.19	0.27	Unbinned	KO:K00134	GAPDH, gapA	glyceraldehyde 3-phosphate dehydrogenase	Carbon metabolism
Ga0371437_10077301	NA	0.14	0.45	-0.65	0.47	Unbinned	KO:K00134	GAPDH, gapA	glyceraldehyde 3-phosphate dehydrogenase	Carbon metabolism
Ga0371429_10545512	0.13	-0.28	-0.78	NA	-0.14	NA	KO:K00134	GAPDH, gapA	glyceraldehyde 3-phosphate dehydrogenase	Carbon metabolism
Ga0371428_1006394	0.11	-0.14	-2.74	NA	-1.77	Phormidium	KO:K00134	GAPDH, gapA	glyceraldehyde 3-phosphate dehydrogenase	Carbon metabolism
Ga0371431_102970814	NA	-1.23	0.82	4.39	1.44	Pseudanabaena	KO:K11532	glpX-SEBP	fructose-1,6-bisphosphatase II / sedoheptulose-1,7-bisphosphatase	Carbon metabolism
Ga0371428_1170091	-0.58	0.30	0.32	NA	NA	Unbinned	KO:K00600	glyA, SHMT	glycine hydroxymethyltransferase	Carbon metabolism
Ga0371428_1011905	-0.03	-0.17	-3.50	NA	NA	Phormidium	KO:K00600	glyA, SHMT	glycine hydroxymethyltransferase	Carbon metabolism
Ga0371428_1028563	2.14	0.69	0.68	0.93	0.45	Gammaproteobacteria	KO:K00031	IDH1, IDH2, icd	isocitrate dehydrogenase	Carbon metabolism
Ga0371428_1027273	1.96	1.36	0.77	0.09	1.29	Gammaproteobacteria	KO:K08691	mcl	malyl-CoA/(S)-citramalyl-CoA lyase	Carbon metabolism
Ga0371428_1011716	0.78	2.10	0.40	2.00	NA	Desulfobacteraceae	KO:K00024	mdh	malate dehydrogenase	Carbon metabolism
Ga0371428_11001619	4.07	-0.13	0.45	1.24	NA	Chloroflexaceae	KO:K00024	mdh	malate dehydrogenase	Carbon metabolism
Ga0371428_1072331	-3.93	0.23	0.41	0.64	0.70	Gammaproteobacteria	KO:K00024	mdh	malate dehydrogenase	Carbon metabolism
Ga0371428_1034841	-0.43	0.05	0.04	NA	NA	Gammaproteobacteria	KO:K00850	pfkA, PFK	6-phosphofructokinase 1	Carbon metabolism
Ga0371428_1174682	0.09	0.73	0.32	NA	NA	Unbinned	KO:K00927	PGK, pgk	phosphoglycerate kinase	Carbon metabolism
Ga0371427_11111981	-1.55	0.15	-1.21	-2.21	-1.53	NA	KO:K00927	PGK, pgk	phosphoglycerate kinase	Carbon metabolism
Ga0371429_10928611	-0.63	0.23	-0.63	-0.29	0.64	NA	KO:K00927	PGK, pgk	phosphoglycerate kinase	Carbon metabolism
Ga0371428_10035615	0.06	-0.31	-2.61	NA	-1.18	Phormidium	KO:K03737	por, nifJ	pyruvate-ferredoxin/flavodoxin oxidoreductase	Carbon metabolism

Ga0371428_10864012	1.14	-0.59	0.30	-0.08	NA	Chloroflexaceae	KO:K01006	ppdK	pyruvate, orthophosphate dikinase	Carbon metabolism
Ga0371428_1014861	-1.64	-0.17	0.16	0.05	-0.34	Gammaproteobacteria	KO:K01006	ppdK	pyruvate, orthophosphate dikinase	Carbon metabolism
Ga0371428_1070691	0.00	0.56	0.60	0.63	0.55	Unbinned	KO:K01006	ppdK	pyruvate, orthophosphate dikinase	Carbon metabolism
Ga0371429_10134122	-0.87	0.02	0.28	-1.66	-1.02	Gammaproteobacteria	KO:K01006	ppdK	pyruvate, orthophosphate dikinase	Carbon metabolism
Ga0371431_11479611	0.37	0.30	-0.06	NA	NA	Gammaproteobacteria	KO:K01006	ppdK	pyruvate, orthophosphate dikinase	Carbon metabolism
Ga0371431_10438991	0.51	-0.75	0.00	NA	NA	Unbinned	KO:K01006	ppdK	pyruvate, orthophosphate dikinase	Carbon metabolism
Ga0371428_1056121	NA	0.18	0.48	-0.27	NA	Unbinned	KO:K01720	prpD	2-methylcitrate dehydratase	Carbon metabolism
Ga0371477_1268713	1.91	2.28	-1.53	NA	0.07	Unbinned	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371428_1105941	NA	-2.02	-2.42	NA	-3.01	Unbinned	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371430_10798471	NA	-0.23	-3.17	-3.71	-1.42	Unbinned	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371430_11454571	NA	1.14	-0.18	NA	1.93	Unbinned	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371432_10469211	NA	0.88	0.91	2.27	1.45	Unbinned	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371437_10430384	NA	2.68	0.93	2.86	2.59	Porphyromonadaceae	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371428_10128315	-0.50	-0.89	-2.02	-0.68	-1.65	Phormidium	KO:K01601	rbcL	ribulose-bisphosphate carboxylase large chain	Carbon metabolism
Ga0371429_104460910	-1.99	0.23	-0.72	NA	NA	Unbinned	KO:K01602	rbcS	ribulose-bisphosphate carboxylase small chain	Carbon metabolism
Ga0371429_10997221	1.00	-0.77	-0.24	NA	NA	Unbinned	KO:K01602	rbcS	ribulose-bisphosphate carboxylase small chain	Carbon metabolism
Ga0371428_10128317	-0.68	0.19	-3.42	NA	-4.63	Phormidium	KO:K01602	rbcS	ribulose-bisphosphate carboxylase small chain	Carbon metabolism
Ga0371429_10652475	1.33	0.57	-0.16	NA	NA	Oscillatoria	KO:K01783	rpe, RPE	ribulose-phosphate 3-epimerase	Carbon metabolism
Ga0371427_10022211	-1.23	0.92	NA	NA	-0.37	NA	KO:K01783	rpe, RPE	ribulose-phosphate 3-epimerase	Carbon metabolism
Ga0371428_1023286	0.43	-0.49	-3.22	NA	NA	Phormidium	KO:K01783	rpe, RPE	ribulose-phosphate 3-epimerase	Carbon metabolism
Ga0371427_10289651	-0.01	0.55	-0.47	NA	1.48	Unbinned	KO:K00615	tktA, tktB	transketolase	Carbon metabolism
Ga0371428_1038202	-0.68	0.76	0.24	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K00615	tktA, tktB	transketolase	Carbon metabolism
Ga0371428_1046621	NA	0.35	0.33	0.30	0.47	Gammaproteobacteria	KO:K00615	tktA, tktB	transketolase	Carbon metabolism
Ga0371429_10297617	0.09	-0.31	-2.89	-2.97	NA	Unbinned	KO:K00615	tktA, tktB	transketolase	Carbon metabolism
Ga0371429_10778891	0.03	0.04	-0.96	NA	0.58	Unbinned	KO:K00615	tktA, tktB	transketolase	Carbon metabolism
Ga0371428_1089327	0.24	-0.26	-2.95	NA	-2.59	Phormidium	KO:K00615	tktA, tktB	transketolase	Carbon metabolism
Ga0371428_1057891	1.11	-0.03	0.36	0.10	NA	Gammaproteobacteria	KO:K01803	TPI, tpiA	triosephosphate isomerase (TIM)	Carbon metabolism
Ga0371435_104584211	NA	NA	-0.10	-0.13	0.67	Bacteria	KO:K01803	TPI, tpiA	triosephosphate isomerase (TIM)	Carbon metabolism
Ga0371427_10192601	0.12	0.59	-0.54	NA	NA	Unbinned	KO:K11262	ACACA	acetyl-CoA carboxylase / biotin carboxylase 1	Central metabolism
Ga0371429_10103132	0.03	0.49	-0.57	NA	NA	Unbinned	KO:K11262	ACACA	acetyl-CoA carboxylase / biotin carboxylase 1	Central metabolism
Ga0371427_11114311	-0.09	0.80	-0.57	NA	NA	NA	KO:K11262	ACACA	acetyl-CoA carboxylase / biotin carboxylase 1	Central metabolism
Ga0371428_1074023	-0.09	0.37	-2.63	NA	NA	Unbinned	KO:K03564	BCP, PRXQ, DOT5	peroxiredoxin Q/BCP	Central metabolism
Ga0371437_10094462	-1.73	0.00	-0.82	-1.69	NA	Cyanobium	KO:K03867	BGLUT	UDP-glucose:tetrahydrobiopterin glucosyltransferase	Central metabolism
Ga0371428_1006358	0.09	-0.38	-4.22	NA	NA	Phormidium	KO:K01955	carB	carbamoyl-phosphate synthase large subunit	Central metabolism
Ga0371476_10587331	-0.50	-0.96	NA	NA	-1.38	Unbinned	KO:K02337	dnaE	DNA polymerase III subunit alpha	Central metabolism
Ga0371431_11497684	NA	0.31	-1.25	-0.11	NA	Comamonadaceae	KO:K02337	dnaE	DNA polymerase III subunit alpha	Central metabolism
Ga0371428_1129145	0.29	-0.33	-2.52	NA	-0.63	Phormidium	KO:K00688	E2.4.1.1, glgP, PYG	starch phosphorylase	Central metabolism
Ga0371428_1008222	-0.03	0.14	-1.86	NA	NA	Phormidium	KO:K01792	E5.1.3.15	glucose-6-phosphate 1-epimerase	Central metabolism
Ga0371428_1074934	0.42	-0.53	-2.76	NA	NA	Phormidium	KO:K01802	E5.2.1.8	peptidylprolyl isomerase	Central metabolism
Ga0371437_11062241	-0.95	-0.29	-0.80	NA	0.45	Unbinned	COG1376	erfK	Lipoprotein-anchoring transpeptidase ErfK/SrfK [Cell wall/membrane/envelope biogenesis]	Central metabolism
Ga0371427_10653031	-2.10	0.15	0.38	NA	NA	NA	KO:K00647	fabB	3-oxoacyl-[acyl-carrier-protein] synthase I	Central metabolism
Ga0371427_11246911	0.68	0.22	-0.37	NA	NA	Unbinned	KO:K09458	fabF	3-oxoacyl-[acyl-carrier-protein] synthase II	Central metabolism
Ga0371476_10160692	-1.31	0.28	-0.83	-0.88	NA	Unbinned	KO:K00059	fabG	3-oxoacyl-[acyl-carrier protein] reductase	Central metabolism
Ga0371428_1165442	-0.99	-0.89	0.09	NA	NA	Chloroflexaceae	KO:K00059	fabG	3-oxoacyl-[acyl-carrier protein] reductase	Central metabolism

Ga0371430_10763253	1.72	0.32	0.49	NA	NA	Comamonadaceae	KO:K00059	fabG	3-oxoacyl-[acyl-carrier protein] reductase	Central metabolism
Ga0371429_11102001	-1.67	0.55	-0.57	NA	-0.57	NA	KO:K00059	fabG	3-oxoacyl-[acyl-carrier protein] reductase	Central metabolism
Ga0371428_1086202	NA	0.15	0.11	NA	0.37	Gammaproteobacterium_117915_bin_15	KO:K00208	fabI	enoyl-[acyl-carrier protein] reductase I	Central metabolism
Ga0371427_10668501	-1.08	0.45	-1.00	NA	NA	NA	KO:K00208	fabI	enoyl-[acyl-carrier protein] reductase I	Central metabolism
Ga0371428_1070952	0.55	-0.78	-3.23	NA	NA	Phormidium	COG0633	fdx	Ferredoxin	Central metabolism
Ga0371431_11166831	-1.61	-1.47	-1.95	NA	NA	Unbinned	KO:K07407	galA, rafA	alpha-galactosidase	Central metabolism
Ga0371428_1001304	NA	0.20	0.65	0.95	NA	Gammaproteobacteria	KO:K02437	gcvH	glycine cleavage system H protein	Central metabolism
Ga0371428_1067011	NA	0.59	0.06	NA	0.58	Unbinned	KO:K01652	ilvB, ilvG, ilvI	acetolactate synthase I/II/III large subunit	Central metabolism
Ga0371429_10478307	NA	0.33	0.32	0.39	NA	Comamonadaceae	KO:K01653	ilvB, ilvG, ilvI	acetolactate synthase I/II/III large subunit	Central metabolism
Ga0371428_1127552	0.26	-0.30	0.16	0.02	0.79	Gammaproteobacteria	KO:K00053	ilvC	ketol-acid reductoisomerase	Central metabolism
Ga0371428_1102213	-0.33	0.16	0.07	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K01687	ilvD	dihydroxy-acid dehydratase	Central metabolism
Ga0371428_1028161	0.14	-0.46	-2.13	NA	NA	Unbinned	KO:K01687	ilvD	dihydroxy-acid dehydratase	Central metabolism
Ga0371437_10801504	NA	0.13	-1.68	-1.69	NA	Bacteria	KO:K04487	iscS, NFS1	cysteine desulfurase	Central metabolism
Ga0371428_1046011	0.27	0.19	-2.41	NA	NA	Unbinned	KO:K16149	K16149	1,4-alpha-glucan branching enzyme	Central metabolism
Ga0371429_10037341	NA	-1.47	-0.44	NA	-1.56	Unbinned	KO:K16881	K16881	mannose-1-phosphate guanylyltransferase / phosphomannomutase	Central metabolism
Ga0371428_1050994	-0.47	0.55	-0.03	0.07	0.95	Gammaproteobacterium_117915_bin_15	KO:K00052	leuB	3-isopropylmalate dehydrogenase	Central metabolism
Ga0371427_10652301	0.14	-0.26	0.42	NA	1.60	Unbinned	COG1304	lldD	FMN-dependent dehydrogenase, includes L-lactate dehydrogenase and type II isopentenyl diphosphate isomerase	Central metabolism
Ga0371430_11136663	-2.30	1.18	0.08	NA	NA	Methylococcaceae	KO:K06078	lpp	murein lipoprotein	Central metabolism
Ga0371428_1027081	0.48	0.31	-2.14	NA	NA	Phormidium	COG1109	manB	Phosphomannomutase	Central metabolism
Ga0371429_10357402	-0.73	0.72	-0.04	-0.95	0.50	Bacteroidales	COG4799	mmdA	Acetyl-CoA carboxylase, carboxyltransferase component	Central metabolism
Ga0371428_1139223	0.18	0.43	-2.48	NA	NA	Phormidium	KO:K00940	ndk, NME	nucleoside-diphosphate kinase	Central metabolism
Ga0371437_11015641	NA	-0.50	2.63	1.30	1.79	Unbinned	KO:K00346	nqrA	Na+-transporting NADH:ubiquinone oxidoreductase subunit A	Central metabolism
Ga0371427_10825782	-0.81	0.02	-0.71	NA	0.99	NA	COG4770	pccA	Acetyl/propionyl-CoA carboxylase, alpha subunit	Central metabolism
Ga0371427_10825781	-1.26	0.16	-0.73	NA	NA	NA	KO:K01965	PCCA, pccA	propionyl-CoA carboxylase alpha chain	Central metabolism
Ga0371428_1031514	0.26	0.08	-2.20	NA	NA	Phormidium	COG0628	perM	Predicted PurR-regulated permease PerM	Central metabolism
Ga0371428_1008562	-1.64	0.61	0.41	0.47	NA	Gammaproteobacteria	KO:K00023	phbB	acetoacetyl-CoA reductase	Central metabolism
Ga0371428_1080104	-0.04	0.73	-0.04	1.06	NA	Desulfobacteraceae	KO:K01908	prpE	propionyl-CoA synthetase	Central metabolism
Ga0371430_11271411	2.78	NA	1.80	0.99	NA	Comamonadaceae	KO:K01908	prpE	propionyl-CoA synthetase	Central metabolism
Ga0371428_1010197	NA	0.37	0.44	NA	0.45	Gammaproteobacteria	KO:K00831	serC, PSAT1	phosphoserine aminotransferase	Central metabolism
Ga0371429_10559433	-0.66	0.12	0.33	NA	NA	Gammaproteobacteria	KO:K00831	serC, PSAT1	phosphoserine aminotransferase	Central metabolism
Ga0371428_1139542	-0.76	0.10	0.36	NA	0.40	Unbinned	COG2930	SYLF	Lipid-binding SYLF domain	Central metabolism
Ga0371429_10808012	1.95	-0.56	0.44	NA	0.81	Unbinned	KO:K03671	trxA	Thiol-disulfide isomerase or thioredoxin	Central metabolism
Ga0371437_10026872	-2.61	0.00	-0.72	NA	NA	Unknown	KO:K00050	ttuD	hydroxypyruvate reductase	Central metabolism
Ga0371430_10989101	0.50	0.26	-0.34	-0.53	1.39	Unbinned	COG5272	UBI4	UBI4	Central metabolism
Ga0371427_10546782	-0.06	0.23	-1.17	NA	NA	Unbinned	KO:K12447	USP	UDP-sugar pyrophosphorylase	Central metabolism
Ga0371428_1017902	0.38	-0.62	0.25	0.27	0.70	Unbinned	KO:K01447	xlyAB	N-acetylmuramoyl-L-alanine amidase	Central metabolism
Ga0371428_10321814	0.76	0.41	-3.53	NA	NA	Phormidium	KO:K09747	ybaB	Conserved DNA-binding protein YbaB	Central metabolism

Ga0371428_10929710	0.28	-0.32	-2.37	NA	NA	Phormidium	KO:K06872	ygcG	Uncharacterized membrane protein YgcG, contains a TPM-fold domain	Central metabolism
Ga0371428_1148093	-2.11	-0.12	0.00	-0.39	-1.40	Desulfobacteraceae	KO:K13954	yiaY	alcohol dehydrogenase	Central metabolism
Ga0371430_11700771	-0.26	-0.23	0.72	0.76	-0.16	Unbinned	KO:K13954	yiaY	alcohol dehydrogenase	Central metabolism
Ga0371428_1034314	-1.04	0.05	0.02	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K03694	clpA	ATP-dependent Clp protease ATP-binding subunit ClpA	Genetic information processing
Ga0371428_1015831	-0.54	-0.25	-0.95	NA	NA	Unbinned	KO:K03695	clpB	ATP-dependent Clp protease ATP-binding subunit ClpB	Genetic information processing
Ga0371428_1077077	-0.35	0.03	-0.73	NA	1.43	Unbinned	KO:K03696	clpC	ATP-dependent Clp protease ATP-binding subunit ClpC	Genetic information processing
Ga0371428_10254622	-0.21	-0.37	-1.44	NA	NA	Phormidium	KO:K03696	clpC	ATP-dependent Clp protease ATP-binding subunit ClpC	Genetic information processing
Ga0371428_1033912	-0.94	0.34	0.24	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K04043	dnaK, HSPA9	molecular chaperone DnaK	Genetic information processing
Ga0371428_10770740	-0.08	0.90	-1.03	NA	NA	Unbinned	KO:K04043	dnaK, HSPA9	molecular chaperone DnaK	Genetic information processing
Ga0371428_1124291	-0.85	0.77	0.40	NA	NA	Unbinned	KO:K04043	dnaK, HSPA9	molecular chaperone DnaK	Genetic information processing
Ga0371428_1008409	-0.69	-0.65	-0.67	0.08	-0.80	Phormidium	KO:K04043	dnaK, HSPA9	molecular chaperone DnaK	Genetic information processing
Ga0371439_10074801	0.12	0.04	-0.28	NA	NA	Unbinned	COG1651	dsbG	Protein-disulfide isomerase [Posttranslational modification, protein turnover, chaperones]	Genetic information processing
Ga0371428_1067583	-0.30	0.50	0.07	NA	NA	Gammaproteobacteria	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371428_10108185	-0.45	0.11	-1.03	NA	NA	Unbinned	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371428_1064711	0.36	0.33	-2.24	NA	NA	Unbinned	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371427_10268294	-0.94	1.13	-1.15	NA	NA	NA	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371428_10154211	0.71	-0.04	-0.11	NA	NA	Phormidium	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371428_1077515	0.09	-0.41	-2.96	NA	NA	Phormidium	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371428_1120204	0.16	-0.16	-1.38	NA	NA	Phormidium	KO:K03798	ftsH, hflB	cell division protease FtsH	Genetic information processing
Ga0371428_1001084	-0.19	-0.34	0.28	NA	NA	Gammaproteobacteria	KO:K03531	ftsZ	cell division protein FtsZ	Genetic information processing
Ga0371428_1001591	-0.23	-0.44	-1.37	NA	NA	Unbinned	KO:K03531	ftsZ	cell division protein FtsZ	Genetic information processing
Ga0371428_10225730	0.22	0.16	-2.26	NA	NA	Phormidium	KO:K03531	ftsZ	cell division protein FtsZ	Genetic information processing
Ga0371428_1050945	-1.36	1.86	-0.50	NA	NA	Desulfobacteraceae	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371428_10770741	0.34	0.38	-0.48	NA	-1.07	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371428_1143251	NA	0.20	0.23	0.39	0.51	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371429_10905038	1.51	-1.40	0.59	NA	NA	Spirulina	KO:K04077	groEL	chaperonin GroEL	Genetic information processing

Ga0371429_10293661	-0.97	-0.70	-1.41	-1.55	-0.45	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371429_10583261	-0.38	1.69	0.61	1.31	0.67	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371429_10897941	-1.52	1.07	0.26	NA	NA	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371430_10698632	NA	1.52	0.69	0.79	1.50	Comamonadaceae	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371430_10030721	NA	0.76	0.13	1.07	1.22	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371430_11007881	1.30	1.47	-0.20	NA	NA	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371430_11857721	-2.09	0.36	0.50	NA	1.02	Unbinned	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371439_10843976	-0.28	0.28	-0.91	NA	NA	Cyanobacteria	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371427_10073762	-0.17	0.54	-0.43	NA	2.07	NA	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371427_10690031	-0.51	0.30	-1.09	NA	NA	NA	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371428_1025303	-0.03	-0.02	-2.21	-0.54	-0.85	Phormidium	KO:K04077	groEL	chaperonin GroEL	Genetic information processing
Ga0371428_1050944	NA	1.53	0.55	2.15	NA	Desulfobacteraceae	KO:K04078	groES, HSPE1	chaperonin GroES	Genetic information processing
Ga0371428_1087013	4.30	NA	-0.59	3.25	NA	Chloroflexaceae	KO:K04078	groES, HSPE1	chaperonin GroES	Genetic information processing
Ga0371428_1101062	-0.46	0.27	0.42	0.81	1.26	Unbinned	KO:K04078	groES, HSPE1	chaperonin GroES	Genetic information processing
Ga0371429_10583262	1.75	-0.82	0.36	NA	NA	Unbinned	KO:K04078	groES, HSPE1	chaperonin GroES	Genetic information processing
Ga0371428_1129073	-1.84	0.96	-0.28	-0.47	-0.45	NA	KO:K11251	H2A	histone H2A	Genetic information processing
Ga0371428_1129072	-1.77	3.56	0.43	NA	NA	NA	KO:K11252	H2B	histone H2B	Genetic information processing
Ga0371428_1013262	-0.55	0.71	-0.24	NA	NA	NA	KO:K11254	H4	histone H4	Genetic information processing
Ga0371428_1037201	-1.14	0.52	0.02	NA	NA	Desulfobacteraceae	KO:K04079	HSP90A, htpG	molecular chaperone HtpG	Genetic information processing
Ga0371430_11668711	-1.87	-0.01	-0.50	0.35	-0.37	Desulfobulbaceae	KO:K04079	HSP90A, htpG	molecular chaperone HtpG	Genetic information processing
Ga0371429_10921972	-1.05	0.85	0.00	NA	NA	Unbinned	KO:K09490	HSPA5, BIP	heat shock 70kDa protein 5	Genetic information processing
Ga0371427_10261561	-0.49	0.83	-0.07	NA	NA	NA	KO:K09490	HSPA5, BIP	heat shock 70kDa protein 5	Genetic information processing
Ga0371427_10261562	1.31	0.26	-0.10	NA	NA	NA	KO:K09490	HSPA5, BIP	heat shock 70kDa protein 5	Genetic information processing
Ga0371428_1031054	0.68	0.09	-0.21	-0.44	0.69	Desulfobacteraceae	KO:K03530	hupB	DNA-binding protein HU-beta	Genetic information processing
Ga0371428_10384815	0.44	0.24	-1.52	NA	NA	Phormidium	KO:K03530	hupB	DNA-binding protein HU-beta	Genetic information processing

Ga0371428_1053391	0.21	0.30	0.32	NA	1.40	Gammaproteobacterium_117915_bin_15	KO:K05788	ihfB, himD	integration host factor subunit beta	Genetic information processing
Ga0371428_1051672	0.65	-0.25	-0.57	NA	NA	Phormidium	KO:K11996	MOCS3, UBA4	adenylyltransferase and sulfurtransferase	Genetic information processing
Ga0371428_1009433	NA	-0.13	0.40	1.56	0.62	Gammaproteobacteria	KO:K03569	mreB	rod shape-determining protein MreB and related proteins	Genetic information processing
Ga0371427_10876452	-0.89	0.17	-0.40	NA	0.66	NA	KO:K11294	NCL, NSR1	nucleolin	Genetic information processing
Ga0371428_10409916	-0.02	0.10	-2.85	NA	NA	Phormidium	KO:K03768	ppiB	peptidyl-prolyl cis-trans isomerase B	Genetic information processing
Ga0371428_1010194	NA	0.72	0.16	1.13	NA	Gammaproteobacteria	KO:K03769	ppiC	peptidyl-prolyl cis-trans isomerase C	Genetic information processing
Ga0371428_1016701	-1.13	0.53	0.09	0.41	0.61	Gammaproteobacteria	KO:K03769	ppiC	peptidyl-prolyl cis-trans isomerase C	Genetic information processing
Ga0371428_1000222	NA	0.19	0.10	NA	1.35	Gammaproteobacteria	KO:K03770	ppiD	peptidyl-prolyl cis-trans isomerase D	Genetic information processing
Ga0371428_1115969	0.34	0.42	-2.22	NA	NA	Phormidium	KO:K03969	pspA	phage shock protein A	Genetic information processing
Ga0371476_12556645	-1.41	0.33	-0.36	NA	2.11	Porphyromonadaceae	KO:K07568	queA	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	Genetic information processing
Ga0371476_11992081	-1.32	1.39	0.02	NA	NA	Unbinned	COG2003	radC	DNA repair protein RadC, contains a helix-hairpin-helix DNA-binding motif	Genetic information processing
Ga0371435_10426612	NA	NA	-0.29	-1.59	-1.90	Unknown	KO:K06179	rluC	23S rRNA pseudouridine955/2504/2580 synthase	Genetic information processing
Ga0371431_10032531	0.36	-0.01	-0.18	-1.15	0.86	NA	KO:K03500	rsmB, sun	16S rRNA (cytosine967-C5)-methyltransferase	Genetic information processing
Ga0371428_1059433	-0.32	0.44	0.54	NA	NA	Unbinned	KO:K03545	tig	trigger factor	Genetic information processing
Ga0371428_10083814	-0.01	0.10	-2.47	NA	NA	Phormidium	KO:K03545	tig	trigger factor	Genetic information processing
Ga0371428_1132216	-1.33	0.05	0.52	-0.12	0.67	Gammaproteobacteria	KO:K07235	tusD, dsrE	tRNA 2-thiouridine synthesizing protein D	Genetic information processing
Ga0371428_1140251	1.13	1.37	0.95	2.02	0.71	Desulfobacteraceae	KO:K11179	tusE, dsrC	tRNA 2-thiouridine synthesizing protein E	Genetic information processing
Ga0371428_1018575	-1.11	-0.34	0.45	-0.02	0.81	Gammaproteobacteria	KO:K11179	tusE, dsrC	tRNA 2-thiouridine synthesizing protein E	Genetic information processing
Ga0371428_1067131	-0.58	0.36	0.25	2.08	NA	NA	KO:K11179	tusE, dsrC	tRNA 2-thiouridine synthesizing protein E	Genetic information processing
Ga0371436_10658062	-0.17	-1.72	-1.55	NA	NA	Bacteroidetes	KO:K09710	ybeB	ribosome-associated protein	Genetic information processing
Ga0371428_1058832	0.00	0.98	0.35	1.58	0.92	Gammaproteobacterium_117915_bin_15	COG0437	hybA	Fe-S-cluster-containing dehydrogenase component	Hydrogen metabolism
Ga0371428_1042153	0.85	1.32	0.22	NA	0.79	Gammaproteobacteria	COG0437	hybA	Fe-S-cluster-containing dehydrogenase component	Hydrogen metabolism
Ga0371429_10376383	NA	-0.09	0.62	0.70	NA	Chloroflexi	KO:K00864	glpK, GK	glycerol kinase	Lipid metabolism
Ga0371428_1008389	-0.50	0.31	-1.22	NA	NA	Phormidium	COG1977	moaD	Molybdopterin converting factor, small subunit [Coenzyme transport and metabolism]	Metabolism of cofactors and vitamins

Ga0371476_11779023	-2.26	1.24	NA	-0.26	NA	Desulfobacteraceae	KO:K03474	pdxJ	pyridoxine 5-phosphate synthase	Metabolism of cofactors and vitamins
Ga0371428_1058401	NA	0.20	0.48	0.47	0.76	Unbinned	KO:K00325	pntB	H+-translocating NAD(P) transhydrogenase subunit beta	Metabolism of cofactors and vitamins
Ga0371428_1095737	0.51	-0.07	-3.47	NA	-2.42	Phormidium	pfam09150	carot_N	orange carotenoid protein/transport factor 2 (NTF2)-like protein	Metabolism of terpenoids and polyketides
Ga0371430_11233411	2.38	-1.60	0.92	NA	NA	Unbinned	pfam13750	big_3_3	Bacterial Ig-like domain (group 3)	NA
Ga0371428_1021431	-0.41	NA	-0.51	0.28	NA	Unbinned	pfam03160, pfam03160, pfam03160	calx_beta	Calx-beta domain	NA
Ga0371430_10421511	-0.17	1.04	0.20	NA	1.10	Unbinned	COG3291, COG3291, COG3291	COG3291	PKD repeat [Function unknown]	NA
Ga0371428_1101276	0.27	1.11	0.22	1.05	0.71	Gammaproteobacteria	COG3597	COG3597	Uncharacterized conserved protein, DUF697 family [Function unknown]	NA
Ga0371428_1017712	0.24	-0.19	-2.10	NA	NA	Phormidium	COG4249	COG4249	Uncharacterized protein, contains caspase domain [General function prediction only]	NA
Ga0371476_11475011	NA	0.40	-0.30	-1.33	-0.67	Bacteria	COG4333	COG4333	uncharacterized protein	NA
Ga0371429_11051161	2.84	-0.91	0.83	NA	NA	Unbinned	COG4803	COG4803	Uncharacterized membrane protein	NA
Ga0371428_1046923	-0.51	0.16	1.11	NA	NA	Gammaproteobacteria	COG4902	COG4902	Uncharacterized protein [Function unknown]	NA
Ga0371428_1123412	-0.84	0.00	0.10	NA	NA	Gammaproteobacteria	COG5490	COG5490	Uncharacterized protein [Function unknown]	NA
Ga0371428_10322158	0.56	-0.34	-3.95	NA	NA	Phormidium	pfam11320	DUF3122	Protein of unknown function in Cyanobacteria (DUF3122)	NA
Ga0371429_10640613	-0.11	-0.23	-2.13	NA	NA	Unbinned	pfam11866	DUF3386	Protein of unknown function (DUF3386)	NA
Ga0371428_1005742	-0.14	0.30	-3.43	NA	NA	Phormidium	pfam13319	DUF4090	Protein of unknown function (DUF4090)	NA
Ga0371428_1068282	2.11	0.64	0.41	NA	NA	NA	pfam13511	DUF4124	Domain of unknown function (DUF4124)	NA
Ga0371428_1111861	0.04	-1.09	0.08	NA	0.26	Unbinned	pfam14467	DUF4426	Domain of unknown function (DUF4426)	NA
Ga0371428_1007893	-0.21	-0.33	0.12	0.47	0.70	Gammaproteobacteria	pfam16137	DUF4845	Domain of unknown function (DUF4845)	NA
Ga0371428_10964911	-0.10	0.03	-0.12	NA	NA	Gammaproteobacteria	pfam06051	DUF928	Domain of Unknown Function (DUF928)	NA
Ga0371428_1088923	2.22	0.37	0.79	NA	1.45	Gammaproteobacteria	pfam00350	dynamin_N	Dynamin family	NA
Ga0371428_1051382	0.55	0.59	-0.34	NA	NA	Unbinned	pfam03065	Glyco_hydro_57	Glycosyl hydrolase family 57	NA
Ga0371427_10410021	-1.13	0.40	-1.08	NA	NA	NA	COG1346	lrgB	Putative effector of murein hydrolase [Cell wall/membrane/envelope biogenesis]	NA
Ga0371427_10997131	-1.49	0.40	0.25	NA	NA	Unbinned	pfam03650	MPC	Uncharacterized protein family (UPF0041)	NA
Ga0371476_10963821	NA	1.00	-1.77	NA	-2.23	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371476_12230653	0.88	0.72	-0.53	NA	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1031523	0.38	1.47	0.51	NA	NA	Desulfobacteraceae	Hypo-rule applied	NA	NA	NA
Ga0371428_1048264	0.22	-0.49	-0.76	0.10	-0.49	Desulfobacteraceae	Hypo-rule applied	NA	NA	NA
Ga0371428_1051362	-0.09	0.93	-0.54	0.11	-0.07	Desulfobacteraceae	Hypo-rule applied	NA	NA	NA
Ga0371428_1155815	0.21	1.03	-0.39	1.17	NA	Desulfobacteraceae	Hypo-rule applied	NA	NA	NA

Ga0371428_1007443	0.05	0.51	0.32	1.29	1.22	Gammaproteobacterium_117915_bin_15	Hypo-rule applied	NA	NA	NA
Ga0371428_1110232	NA	1.62	0.25	1.89	NA	Gammaproteobacterium_117915_bin_15	Hypo-rule applied	NA	NA	NA
Ga0371428_1135553	-0.06	0.46	0.78	0.43	NA	Gammaproteobacterium_117915_bin_15	Hypo-rule applied	NA	NA	NA
Ga0371428_10259610	NA	0.40	0.74	1.83	NA	Chloroflexaceae	Hypo-rule applied	NA	NA	NA
Ga0371428_1139257	1.45	0.22	-0.45	2.31	NA	Chloroflexaceae	Hypo-rule applied	NA	NA	NA
Ga0371428_1014892	-1.04	0.05	-0.09	-0.55	0.13	Gammaproteobacteria	Hypo-rule applied	NA	NA	NA
Ga0371428_1086063	-1.31	-0.24	0.36	NA	NA	Gammaproteobacteria	Hypo-rule applied	NA	NA	NA
Ga0371428_1111232	0.41	-0.21	0.35	NA	1.00	Gammaproteobacteria	Hypo-rule applied	NA	NA	NA
Ga0371428_1074672	-0.11	-0.50	0.72	1.11	0.98	Desulfobacteraceae	Hypo-rule applied	NA	NA	NA
Ga0371428_1005411	0.71	-0.85	-3.72	NA	-2.96	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1021302	-0.45	0.58	0.40	NA	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1037311	-2.56	0.84	0.45	2.71	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1048691	-0.09	-0.08	0.65	NA	0.61	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1066471	0.45	0.39	0.42	0.35	0.66	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1071765	0.67	-0.56	-4.46	NA	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1080451	-1.15	-0.41	0.27	-0.78	-0.33	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1104831	0.63	0.68	0.53	NA	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1105191	NA	-0.08	-0.14	NA	-0.59	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1117172	0.06	0.02	-0.69	NA	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1142532	0.51	-0.24	0.22	0.99	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1147141	NA	1.66	0.57	0.84	NA	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371428_1162181	-0.90	0.26	-0.07	-0.39	0.21	Unbinned	Hypo-rule applied	NA	NA	NA
Ga0371427_10176892	0.15	0.18	-0.44	NA	NA	NA	Hypo-rule applied	NA	NA	NA
Ga0371427_10402231	-1.69	0.49	-0.64	NA	NA	NA	COG0526	NA	NA	NA
Ga0371427_10573131	0.03	0.30	-1.65	NA	-1.23	NA	Hypo-rule applied	NA	NA	NA

Ga0371427_10722331	-1.59	-1.34	NA	NA	-1.15	NA	Hypo-rule applied	NA	NA	NA
Ga0371427_10947111	-0.23	0.49	-1.15	NA	1.42	NA	Hypo-rule applied	NA	NA	NA
Ga0371427_11201192	-0.64	-0.01	-0.80	NA	NA	NA	Hypo-rule applied	NA	NA	NA
Ga0371429_10193611	0.63	0.27	-0.76	NA	NA	NA	NA	NA	NA	NA
Ga0371429_10277401	-0.86	0.28	-1.40	-0.63	0.10	NA	NA	NA	NA	NA
Ga0371429_10294031	NA	0.12	0.30	0.03	0.76	NA	NA	NA	NA	NA
Ga0371429_10305275	2.44	1.80	-1.48	NA	NA	NA	NA	NA	NA	NA
Ga0371429_10419341	1.22	0.10	-1.04	NA	2.12	NA	NA	NA	NA	NA
Ga0371429_10427001	1.52	NA	0.95	2.32	0.88	NA	NA	NA	NA	NA
Ga0371429_10438631	-1.60	-0.01	-0.24	NA	NA	NA	NA	NA	NA	NA
Ga0371429_10442322	-2.09	0.56	-0.76	NA	NA	NA	NA	NA	NA	NA
Ga0371429_10698911	-0.01	0.65	-0.36	NA	NA	NA	NA	NA	NA	NA
Ga0371429_10718331	-0.17	-0.06	-0.30	NA	0.84	NA	NA	NA	NA	NA
Ga0371429_10738081	-1.29	0.14	-0.07	-0.39	-0.31	NA	NA	NA	NA	NA
Ga0371429_10798091	-0.93	-0.22	-0.72	-1.50	-0.14	NA	NA	NA	NA	NA
Ga0371429_11041357	1.39	1.20	-1.06	NA	NA	NA	NA	NA	NA	NA
Ga0371429_11090491	0.22	0.32	-0.24	NA	NA	NA	NA	NA	NA	NA
Ga0371430_10011381	1.04	0.45	-1.69	NA	NA	NA	NA	NA	NA	NA
Ga0371430_10012642	NA	0.53	0.40	NA	0.57	NA	NA	NA	NA	NA
Ga0371430_10180451	0.80	-0.72	-2.34	NA	NA	NA	NA	NA	NA	NA
Ga0371430_10256303	0.81	0.90	0.06	NA	2.12	NA	NA	NA	NA	NA
Ga0371430_10286661	-1.27	0.61	0.60	0.50	0.22	NA	NA	NA	NA	NA
Ga0371430_10300221	1.47	0.94	0.46	NA	1.09	NA	NA	NA	NA	NA
Ga0371430_10693502	-1.20	-0.31	-0.04	0.45	1.25	NA	NA	NA	NA	NA
Ga0371430_10806141	0.41	-0.54	NA	NA	-1.80	NA	NA	NA	NA	NA
Ga0371430_10877745	1.46	0.91	0.42	NA	NA	NA	NA	NA	NA	NA
Ga0371430_10922231	-0.96	-0.68	0.61	-1.04	-0.15	NA	NA	NA	NA	NA
Ga0371430_11060401	-1.56	0.46	0.38	NA	NA	NA	NA	NA	NA	NA
Ga0371430_11079521	NA	-0.19	0.70	NA	0.96	NA	NA	NA	NA	NA
Ga0371430_11331913	-2.80	-1.63	0.11	0.20	-0.12	NA	NA	NA	NA	NA
Ga0371431_10026731	NA	NA	-0.32	-0.84	-1.23	NA	NA	NA	NA	NA
Ga0371431_10064023	1.45	-0.10	0.37	NA	NA	NA	NA	NA	NA	NA
Ga0371431_10602293	-1.40	0.52	-1.86	NA	NA	NA	NA	NA	NA	NA
Ga0371431_106037325	-1.28	0.02	-1.54	NA	-0.70	NA	NA	NA	NA	NA
Ga0371431_10759852	0.14	-0.37	-0.81	NA	1.24	NA	NA	NA	NA	NA
Ga0371431_11093833	-1.61	-1.29	-1.69	NA	-1.65	NA	NA	NA	NA	NA
Ga0371431_11242061	1.24	1.32	-1.20	NA	NA	NA	NA	NA	NA	NA
Ga0371432_10132042	-0.24	NA	-0.75	-0.81	0.01	NA	NA	NA	NA	NA
Ga0371432_10198881	-0.05	0.20	-0.41	NA	NA	NA	NA	NA	NA	NA
Ga0371432_10272731	NA	0.85	0.94	NA	1.74	NA	NA	NA	NA	NA
Ga0371432_10310671	-0.47	0.33	-0.71	NA	-0.24	NA	NA	NA	NA	NA
Ga0371434_10703411	NA	1.96	-0.14	NA	2.26	NA	NA	NA	NA	NA
Ga0371434_10804241	0.80	2.44	-1.24	NA	NA	NA	NA	NA	NA	NA
Ga0371434_11097334	1.15	-2.02	-0.59	NA	-0.08	NA	NA	NA	NA	NA
Ga0371435_11398831	NA	-2.16	0.06	NA	-0.59	NA	NA	NA	NA	NA
Ga0371435_12388952	0.35	-0.02	-0.88	NA	NA	NA	NA	NA	NA	NA
Ga0371436_10117326	0.09	0.16	-2.20	NA	NA	NA	NA	NA	NA	NA

Ga0371436_114888211	0.32	1.30	-0.07	NA	NA	NA	NA	NA	NA	NA
Ga0371437_10389522	-0.97	-0.77	-1.05	NA	NA	NA	NA	NA	NA	NA
Ga0371437_10629121	-0.63	-0.44	0.60	-2.44	NA	NA	NA	NA	NA	NA
Ga0371437_11171172	0.33	-0.09	NA	-2.78	-1.14	NA	NA	NA	NA	NA
Ga0371477_1217341	0.20	-0.43	-0.19	NA	NA	NA	NA	NA	NA	NA
Ga0371428_1001543	0.80	-0.45	-3.95	NA	NA	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_1003886	0.43	1.31	-3.09	-1.71	0.51	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_1006233	0.39	-0.15	-3.26	NA	NA	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_1008504	0.05	0.50	-3.02	NA	NA	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_1025554	-1.76	1.23	NA	NA	-1.61	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_1050145	0.83	0.27	-3.06	NA	-0.19	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_1059606	0.46	-0.58	-2.44	NA	-1.20	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_10639620	0.19	0.87	-1.98	NA	NA	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_10877726	0.46	0.62	-1.76	NA	NA	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_11001728	0.83	0.16	-2.07	NA	NA	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371428_11295414	-0.05	-0.20	-2.99	NA	0.54	Phormidium	Hypo-rule applied	NA	NA	NA
Ga0371427_10318683	-0.65	-0.07	-0.63	NA	NA	NA	pfam13460	NAD_binding_10	NAD(P)H-binding	NA
Ga0371428_1039624	0.11	0.81	-3.62	NA	NA	Phormidium	pfam04966	oprB	Carbohydrate-selective porin, OprB family	NA
Ga0371429_10537741	1.21	-2.28	0.47	1.22	0.38	Spirulina	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371429_10080561	3.13	-0.28	0.70	NA	NA	Unbinned	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371429_10233701	2.12	-0.22	-2.25	-2.09	-1.12	Unbinned	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371429_10999365	0.43	0.50	-3.27	NA	0.08	Unbinned	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371431_104222215	NA	0.41	0.94	3.57	0.95	Pseudanabaena	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371428_1003484	0.59	0.30	-2.79	NA	-2.80	Phormidium	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371428_1024769	-0.55	-0.16	-0.33	-0.76	1.92	Phormidium	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371428_1026451	0.46	0.79	-2.43	NA	NA	Phormidium	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371431_11013295	2.23	NA	0.76	NA	0.73	Planktothrix	pfam04966, pfam00395	oprB/SLH	Carbohydrate-selective porin, OprB family/S-layer homology domain	NA
Ga0371476_10000431	NA	0.96	0.51	-0.04	1.49	Gammaproteobacteria	pfam07691	PA14	insert in bacterial beta-glucosidases	NA
Ga0371429_10980632	NA	-2.92	0.29	0.07	0.43	Gammaproteobacteria	pfam07691	PA14	insert in bacterial beta-glucosidases	NA

Ga0371430_11398266	-0.70	-0.11	-0.05	0.42	1.49	Comamonadaceae	pfam09361	phasin_2	Phasin protein	NA
Ga0371428_1055704	-0.60	NA	0.33	0.04	NA	Gammaproteobacteria	pfam13609	porin_4	Gram-negative porin	NA
Ga0371429_10652881	2.85	NA	0.79	1.72	NA	Unbinned	pfam04151	PPC	Bacterial pre-peptidase C-terminal domain	NA
Ga0371430_10832612	3.16	-0.85	0.52	NA	NA	NA	pfam04151	PPC	Bacterial pre-peptidase C-terminal domain	NA
Ga0371428_1040881	0.57	-0.08	-2.94	-1.73	NA	Unbinned	pfam17210	sdrD	SdrD B-like domain	NA
Ga0371428_1046022	0.16	0.51	-3.51	NA	-1.20	Unbinned	pfam17210, pfam17210	sdrD	SdrD B-like domain	NA
Ga0371429_10415953	0.53	-0.34	0.95	0.49	0.97	Gammaproteobacteria	pfam17210	sdrD	SdrD B-like domain	NA
Ga0371430_11455791	0.30	1.16	-0.38	NA	NA	Unbinned	pfam17210, pfam17210, pfam17210, pfam17210	sdrD	SdrD B-like domain	NA
Ga0371437_10083838	-2.30	0.27	-0.15	NA	NA	Campylobacterales	COG3381	torD	Cytoplasmic chaperone TorD involved in molybdoenzyme TorA maturation [Posttranslational modification, protein turnover, chaperones]	NA
Ga0371439_10101874	-0.39	NA	-1.06	-1.34	0.75	Cryomorphaceae	pfam13517, pfam13517, pfam13517	VCBS	Repeat domain in Vibrio, Colwellia, Bradyrhizobium and Shewanella	NA
Ga0371429_10383762	-0.02	0.53	-0.39	NA	2.46	NA	pfam00094	VWD	von Willebrand factor type D domain	NA
Ga0371428_10412415	0.49	-0.63	-3.03	NA	NA	Phormidium	pfam07444	Ycf66_N	Ycf66 protein N-terminus	NA
Ga0371428_1055589	0.63	0.26	-2.81	NA	NA	Phormidium	COG3103	ygiM	Uncharacterized conserved protein YgiM, contains N-terminal SH3 domain, DUF1202 family [General function prediction only]	NA
Ga0371428_1037704	0.19	-0.25	-2.77	NA	NA	Unbinned	COG5492, COG5492	yjdB	Uncharacterized conserved protein YjdB, contains Ig-like domain [General function prediction only]	NA
Ga0371428_1042154	NA	1.31	-0.09	-0.44	0.50	Gammaproteobacteria	COG0243	bisC	Anaerobic selenocysteine-containing dehydrogenase	Nitrogen metabolism
Ga0371428_1094771	NA	0.37	0.70	NA	1.11	Unbinned	COG0243	bisC	Anaerobic selenocysteine-containing dehydrogenase	Nitrogen metabolism
Ga0371429_10398651	NA	-0.08	0.15	1.25	0.21	Unbinned	COG0243	bisC	Anaerobic selenocysteine-containing dehydrogenase	Nitrogen metabolism
Ga0371430_10058331	0.13	0.49	0.16	-0.17	NA	Unbinned	COG0243	bisC	Anaerobic selenocysteine-containing dehydrogenase	Nitrogen metabolism
Ga0371429_10005272	-1.42	0.27	0.31	NA	NA	NA	KO:K00262	E1.4.1.4, gdhA	glutamate dehydrogenase (NADP+)	Nitrogen metabolism
Ga0371429_10301142	-0.80	1.01	0.57	NA	NA	NA	KO:K00262	E1.4.1.4, gdhA	glutamate dehydrogenase (NADP+)	Nitrogen metabolism
Ga0371429_10439143	-0.69	0.08	-0.49	-1.15	-0.42	NA	KO:K00262	E1.4.1.4, gdhA	glutamate dehydrogenase (NADP+)	Nitrogen metabolism
Ga0371428_1003843	-1.85	-0.43	-0.22	-1.14	-0.73	Gammaproteobacteria	KO:K01915	glnA	glutamine synthetase	Nitrogen metabolism
Ga0371429_108364118	0.66	NA	1.00	3.39	NA	Spirulina	KO:K01915	glnA	glutamine synthetase	Nitrogen metabolism
Ga0371428_1005633	0.17	-0.59	-1.97	NA	0.46	Phormidium	KO:K01915	glnA	glutamine synthetase	Nitrogen metabolism
Ga0371428_1013542	-0.07	0.60	0.22	1.02	1.32	Unbinned	KO:K04752	glnK	nitrogen regulatory protein P-II 2	Nitrogen metabolism
Ga0371428_1042152	-0.29	1.56	-0.58	NA	NA	Gammaproteobacteria	COG3301	nrfD	Formate-dependent nitrite reductase, membrane component NrfD	Nitrogen metabolism
Ga0371430_10963327	NA	3.11	0.12	NA	2.44	Methylococcaceae	KO:K10945	pmoB-amoB	methane/ammonia monooxygenase subunit B	Nitrogen metabolism
Ga0371428_1022982	-0.28	0.24	0.51	NA	NA	Gammaproteobacteria	KO:K01081	E3.1.3.5	5'-nucleotidase	Nucleotide metabolism

Ga0371428_1063502	-0.12	-0.27	-2.13	-0.03	NA	Phormidium	KO:K00088	guaB	IMP dehydrogenase	Nucleotide metabolism
Ga0371429_10224212	-0.34	-0.01	-0.66	NA	NA	Unbinned	KO:K02137	ATPeF0O, ATP5O, ATP5	F-type H+-transporting ATPase subunit O	Oxidative phosphorylation
Ga0371429_10224021	1.09	0.81	-0.36	0.14	0.76	NA	KO:K02133	ATPeF1B, ATP5B, ATP2	F-type H+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371430_10480262	-0.83	-0.34	-0.89	NA	NA	NA	KO:K02133	ATPeF1B, ATP5B, ATP2	F-type H+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371429_11093461	0.10	-0.07	-0.29	NA	NA	NA	KO:K02145	ATPeV1A, ATP6A	V-type H+-transporting ATPase subunit A	Oxidative phosphorylation
Ga0371428_10108146	-0.34	0.10	-0.51	NA	0.20	Unbinned	KO:K02109	ATPF0B, atpF	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation
Ga0371428_10108147	-1.78	0.73	-0.77	NA	NA	Unbinned	KO:K02109	ATPF0B, atpF	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation
Ga0371429_11096493	1.79	-0.17	-2.10	NA	NA	Unbinned	KO:K02109	ATPF0B, atpF	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation
Ga0371430_10349932	0.83	0.03	0.29	NA	-0.50	Unbinned	KO:K02109	ATPF0B, atpF	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation
Ga0371428_1014306	0.05	-0.10	-2.31	-2.70	-2.04	Phormidium	KO:K02109	ATPF0B, atpF	F-type H+-transporting ATPase subunit b	Oxidative phosphorylation
Ga0371428_1016895	0.07	0.59	0.02	0.11	1.08	Desulfobacteraceae	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371428_1095981	-0.52	0.59	0.58	0.62	-0.25	Desulfobacteraceae	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371428_1005231	-2.18	0.23	0.46	NA	NA	Gammaproteobacteria	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371429_11097802	0.81	NA	2.27	4.09	NA	Spirulina	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371429_10719711	-0.79	0.26	0.25	0.17	0.92	Gammaproteobacteria	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371429_10940201	-0.54	0.42	-0.21	NA	NA	Bacteroidales	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371429_10269981	-0.30	-0.22	0.63	-3.71	NA	Unbinned	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371429_10436002	-0.09	NA	1.24	2.41	1.65	Unbinned	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371429_11096491	-0.98	-0.64	-1.20	-2.20	-0.27	Unbinned	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371431_109985717	0.79	0.70	0.37	NA	NA	Planktothrix	KO:K02111	ATPF1A, atpA	F-type H+/Na+-transporting ATPase subunit alpha	Oxidative phosphorylation
Ga0371476_10045734	1.55	-0.16	0.33	NA	NA	Bacteroidales	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371428_1113412	NA	1.91	0.75	1.10	1.66	Bacteroidales	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371428_10108159	-0.65	0.72	-0.53	-0.06	0.09	Unbinned	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation

Ga0371428_1118821	0.13	0.59	0.56	-0.66	1.42	Unbinned	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371429_10405824	-0.56	1.27	0.32	NA	NA	Sulfurospirillum 16	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371429_10691745	0.47	NA	0.51	NA	0.64	Spirulina	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371430_10365353	0.78	0.17	-0.05	-0.96	1.75	Comamonadaceae	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371438_1122111	1.59	NA	1.04	2.84	NA	Unbinned	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371428_1162232	0.15	0.52	-0.03	0.06	1.92	NA	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371428_1014658	-0.02	-0.31	-1.72	-1.54	-1.43	Phormidium	KO:K02112	ATPF1B, atpD	F-type H+/Na+-transporting ATPase subunit beta	Oxidative phosphorylation
Ga0371428_10108148	-1.31	0.87	-0.35	NA	NA	Unbinned	KO:K02113	ATPF1D, atpH	F-type H+-transporting ATPase subunit delta	Oxidative phosphorylation
Ga0371429_11096492	0.90	-0.34	-2.30	NA	NA	Unbinned	KO:K02113	ATPF1D, atpH	F-type H+-transporting ATPase subunit delta	Oxidative phosphorylation
Ga0371428_10108160	0.10	0.06	-0.64	NA	2.10	Unbinned	KO:K02114	ATPF1E, atpC	F-type H+-transporting ATPase subunit epsilon	Oxidative phosphorylation
Ga0371428_1014659	-0.09	0.04	-2.49	NA	NA	Phormidium	KO:K02114	ATPF1E, atpC	F-type H+-transporting ATPase subunit epsilon	Oxidative phosphorylation
Ga0371428_1016894	0.34	1.37	0.50	NA	NA	Desulfobacteraceae	KO:K02115	ATPF1G, atpG	F-type H+-transporting ATPase subunit gamma	Oxidative phosphorylation
Ga0371428_10143010	0.18	-0.40	-2.76	NA	-0.72	Phormidium	KO:K02115	ATPF1G, atpG	F-type H+-transporting ATPase subunit gamma	Oxidative phosphorylation
Ga0371429_10255114	NA	0.64	0.97	NA	1.49	Gammaproteobacteria	KO:K00413	CYC1, CYT1, petC	ubiquinol-cytochrome c reductase cytochrome c1 subunit	Oxidative phosphorylation
Ga0371428_1158876	-0.64	0.35	0.01	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K00428	cytCP	cytochrome c peroxidase	Oxidative phosphorylation
Ga0371428_1021391	NA	-1.49	-1.48	NA	1.88	Unbinned	KO:K00428	cytCP	cytochrome c peroxidase	Oxidative phosphorylation
Ga0371427_11075022	-0.24	0.67	-0.64	NA	NA	Unbinned	KO:K01507	ppa	inorganic pyrophosphatase	Oxidative phosphorylation
Ga0371427_11115173	0.13	0.22	-0.31	NA	NA	NA	KO:K00235	SDHB, SDH2	succinate dehydrogenase (ubiquinone) iron-sulfur subunit	Oxidative phosphorylation
Ga0371429_10621063	0.20	-0.81	0.12	-1.50	0.18	Unknown	KO:K02092	apcA	allophycocyanin alpha subunit	Photosynthesis
Ga0371429_11105074	-2.01	-1.24	0.14	1.82	1.35	Pseudanabaena	KO:K02092	apcA	allophycocyanin alpha subunit	Photosynthesis
Ga0371429_10689267	-0.08	-2.41	-0.02	0.86	-1.01	Spirulina	KO:K02092	apcA	allophycocyanin alpha subunit	Photosynthesis
Ga0371434_10335453	-0.05	-0.55	-3.15	-2.23	-1.48	Unbinned	KO:K02092	apcA	allophycocyanin alpha subunit	Photosynthesis
Ga0371429_10621062	0.25	-0.86	0.18	-1.41	-0.50	Unknown	KO:K02093	apcB	allophycocyanin beta subunit	Photosynthesis
Ga0371429_11105073	-0.98	-0.49	1.28	1.60	0.77	Pseudanabaena	KO:K02093	apcB	allophycocyanin beta subunit	Photosynthesis
Ga0371429_10689268	-0.11	-2.00	-0.01	0.46	-1.46	Spirulina	KO:K02093	apcB	allophycocyanin beta subunit	Photosynthesis
Ga0371429_10347211	1.73	0.25	-2.65	2.64	NA	Unbinned	KO:K02093	apcB	allophycocyanin beta subunit	Photosynthesis
Ga0371439_10016595	0.38	-0.74	1.32	NA	NA	Cyanobacteria	KO:K02093	apcB	allophycocyanin beta subunit	Photosynthesis
Ga0371439_10663031	0.73	-2.91	NA	NA	-2.20	Unbinned	KO:K02093	apcB	allophycocyanin beta subunit	Photosynthesis
Ga0371428_1168882	0.03	-0.08	-3.95	NA	NA	Phormidium	KO:K02095	apcD	allophycocyanin-B	Photosynthesis
Ga0371428_1093111	-0.18	-0.12	-2.41	NA	NA	Phormidium	KO:K02096	apcE	phycobilisome core-membrane linker protein	Photosynthesis
Ga0371428_1005632	-0.11	0.28	-3.59	NA	-2.66	Phormidium	KO:K02097	apcF	phycobilisome core component	Photosynthesis

Ga0371434_10082331	-1.03	0.03	NA	-1.60	-1.00	Bacteroidales	COG4447	COG4447	photosystem II stability/assembly factor-like uncharacterized protein	Photosynthesis
Ga0371429_11062711	0.04	0.78	-0.88	NA	NA	NA	COG4447	COG4447	photosystem II stability/assembly factor-like uncharacterized protein	Photosynthesis
Ga0371427_11260832	NA	-1.76	0.41	-1.62	-0.94	Unbinned	KO:K02284	cpcA	phycocyanin alpha chain	Photosynthesis
Ga0371429_10538883	0.11	-1.46	-0.61	0.20	0.42	Spirulina	KO:K02284	cpcA	phycocyanin alpha chain	Photosynthesis
Ga0371436_10348192	0.92	2.01	0.45	0.80	1.57	Unbinned	KO:K02284	cpcA	phycocyanin alpha chain	Photosynthesis
Ga0371439_10478956	0.80	-1.85	-2.26	NA	-3.71	Pseudanabaena	KO:K02284	cpcA	phycocyanin alpha chain	Photosynthesis
Ga0371427_10317775	-1.52	-0.83	0.80	NA	0.17	Unbinned	KO:K02285	cpcB	phycocyanin beta chain	Photosynthesis
Ga0371427_11260831	0.40	-0.50	0.64	NA	NA	Unbinned	KO:K02285	cpcB	phycocyanin beta chain	Photosynthesis
Ga0371429_10538882	-1.16	-2.89	-0.19	0.39	-1.80	Spirulina	KO:K02285	cpcB	phycocyanin beta chain	Photosynthesis
Ga0371430_10409153	0.33	0.05	-3.00	NA	-1.36	Unbinned	KO:K02285	cpcB	phycocyanin beta chain	Photosynthesis
Ga0371431_100675410	NA	-0.91	0.94	1.05	2.15	Pseudanabaena	KO:K02285	cpcB	phycocyanin beta chain	Photosynthesis
Ga0371439_10321082	0.52	-0.39	-1.62	NA	NA	Unbinned	KO:K02285	cpcB	phycocyanin beta chain	Photosynthesis
Ga0371429_10538884	0.78	-2.96	0.58	1.40	-1.68	Spirulina	KO:K02286	cpcC	phycocyanin-associated rod linker protein	Photosynthesis
Ga0371429_10575295	-0.21	0.58	-1.39	NA	NA	Unbinned	KO:K02286	cpcC	phycocyanin-associated rod linker protein	Photosynthesis
Ga0371429_10538885	1.58	NA	1.62	NA	0.45	Spirulina	KO:K02287	cpcD	phycocyanin-associated, rod	Photosynthesis
Ga0371429_11067193	1.03	NA	2.25	3.93	NA	Spirulina	KO:K02290	cpcG	phycobilisome rod-core linker protein	Photosynthesis
Ga0371431_10128548	1.62	NA	1.89	1.01	NA	Pseudanabaena	KO:K02290	cpcG	phycobilisome rod-core linker protein	Photosynthesis
Ga0371428_1092973	-0.05	0.17	-2.66	NA	-0.82	Phormidium	KO:K02290	cpcG	phycobilisome rod-core linker protein	Photosynthesis
Ga0371427_10096134	-1.71	0.33	0.61	NA	NA	Unknown	KO:K05376	cpeA, mpeA	phycoerythrin alpha chain	Photosynthesis
Ga0371435_10476131	NA	1.59	-0.08	NA	2.25	Unbinned	KO:K05376	cpeA, mpeA	phycoerythrin alpha chain	Photosynthesis
Ga0371428_10877741	-0.63	-0.13	-2.00	-1.44	-2.11	Phormidium	KO:K05376	cpeA, mpeA	phycoerythrin alpha chain	Photosynthesis
Ga0371436_11380135	-1.95	-1.97	-1.93	-0.48	0.97	Unbinned	KO:K05377	cpeB, mpeB	phycoerythrin beta chain	Photosynthesis
Ga0371437_10994253	-3.19	-0.17	-0.47	NA	NA	Cyanobium	KO:K05377	cpeB, mpeB	phycoerythrin beta chain	Photosynthesis
Ga0371437_10438781	-1.71	2.91	0.55	-0.17	1.72	Unbinned	KO:K05377	cpeB, mpeB	phycoerythrin beta chain	Photosynthesis
Ga0371431_10732495	-1.86	-1.64	0.05	-2.17	-0.42	Planktothrix	KO:K05377	cpeB, mpeB	phycoerythrin beta chain	Photosynthesis
Ga0371429_10235543	0.39	0.57	-0.92	NA	NA	Unbinned	KO:K05378	cpeC, mpeC	phycoerythrin-associated linker protein	Photosynthesis
Ga0371431_107324913	0.70	NA	0.79	0.14	NA	Planktothrix	KO:K05379	cpeD, mpeD	phycoerythrin-associated linker protein	Photosynthesis
Ga0371429_10512061	-0.40	0.44	-1.32	NA	-0.39	NA	KO:K08907	LHCA1	light-harvesting complex I chlorophyll a/b binding protein 1	Photosynthesis
Ga0371429_10565811	-1.83	0.04	-0.78	NA	NA	NA	KO:K08907	LHCA1	light-harvesting complex I chlorophyll a/b binding protein 1	Photosynthesis
Ga0371429_11033275	0.12	0.20	-0.91	-0.20	-0.64	NA	KO:K08907	LHCA1	light-harvesting complex I chlorophyll a/b binding protein 1	Photosynthesis
Ga0371428_1148081	0.23	-3.44	-0.47	NA	-0.11	Unbinned	KO:K02630	pecC	phycoerythrocyanin-associated rod linker protein	Photosynthesis
Ga0371429_10534538	-0.05	-3.72	1.01	NA	NA	Spirulina	KO:K02630	pecC	phycoerythrocyanin-associated rod linker protein	Photosynthesis
Ga0371429_10251643	0.42	-2.56	-2.94	NA	-1.51	Unbinned	KO:K02634	petA	apocytochrome f	Photosynthesis
Ga0371428_10108179	-1.00	-0.14	-0.75	-2.06	-0.74	Unbinned	KO:K02635	petB	cytochrome b6	Photosynthesis
Ga0371428_10726411	-0.57	0.14	-2.64	NA	NA	Phormidium	KO:K02635	petB	cytochrome b6	Photosynthesis
Ga0371428_1105171	-0.43	-0.12	-2.75	NA	NA	Unbinned	KO:K02636	petC	cytochrome b6-f complex iron-sulfur subunit	Photosynthesis

Ga0371429_10691252	-0.43	-0.17	-2.30	NA	NA	Oscillatoria	KO:K02636	petC	cytochrome b6-f complex iron-sulfur subunit	Photosynthesis
Ga0371428_10322150	0.44	-0.72	-4.32	NA	-1.61	Phormidium	KO:K02636	petC	cytochrome b6-f complex iron-sulfur subunit	Photosynthesis
Ga0371428_10315114	0.77	-0.20	-3.25	NA	-1.89	Phormidium	KO:K02639	petF	ferredoxin	Photosynthesis
Ga0371429_101458613	-0.84	0.85	-3.41	NA	-3.15	Phormidium	KO:K02639	petF	ferredoxin	Photosynthesis
Ga0371427_10397601	-1.06	-0.01	-1.23	-3.04	-0.78	NA	KO:K02641	petH	ferredoxin--NADP+ reductase	Photosynthesis
Ga0371428_1136681	0.49	-0.40	-2.80	NA	NA	Phormidium	KO:K02641	petH	ferredoxin--NADP+ reductase	Photosynthesis
Ga0371429_10183096	1.24	-0.51	-1.84	NA	NA	Oscillatoria	KO:K08906	petJ	cytochrome c6	Photosynthesis
Ga0371428_1032528	0.57	-0.09	-4.15	NA	NA	Phormidium	KO:K08906	petJ	cytochrome c6	Photosynthesis
Ga0371428_1010812	-1.22	0.35	-0.83	-0.38	-0.72	Unbinned	KO:K02689	psaA	photosystem I P700 chlorophyll a apoprotein A1	Photosynthesis
Ga0371428_10070625	0.24	-0.27	-3.16	NA	-1.32	Phormidium	KO:K02689	psaA	photosystem I P700 chlorophyll a apoprotein A1	Photosynthesis
Ga0371429_10535265	-0.90	-0.17	0.52	1.07	0.61	Pseudanabaena	KO:K02690	psaB	photosystem I P700 chlorophyll a apoprotein A2	Photosynthesis
Ga0371428_1077072	NA	-0.68	-1.21	NA	2.16	Unbinned	KO:K02691	psaC	photosystem I subunit VII	Photosynthesis
Ga0371428_1010221	-0.06	0.30	-1.34	NA	-0.92	Phormidium	KO:K02691	psaC	photosystem I subunit VII	Photosynthesis
Ga0371429_10055136	1.73	NA	0.52	-0.09	-1.14	Spirulina	KO:K02692	psaD	photosystem I subunit II	Photosynthesis
Ga0371431_10887161	NA	-0.90	0.58	1.96	1.63	Pseudanabaena	KO:K02692	psaD	photosystem I subunit II	Photosynthesis
Ga0371428_10108186	-0.27	0.43	-0.82	NA	NA	Unbinned	KO:K02693	psaE	photosystem I subunit IV	Photosynthesis
Ga0371429_10585196	1.72	NA	1.58	2.16	NA	Spirulina	KO:K02693	psaE	photosystem I subunit IV	Photosynthesis
Ga0371428_1109561	1.02	0.22	-2.62	NA	NA	Phormidium	KO:K02693	psaE	photosystem I subunit IV	Photosynthesis
Ga0371428_1010814	-0.30	0.00	-0.92	-0.53	0.95	Unbinned	KO:K02694	psaF	photosystem I subunit III	Photosynthesis
Ga0371429_10878384	NA	NA	0.86	2.26	2.01	Pseudanabaena	KO:K02694	psaF	photosystem I subunit III	Photosynthesis
Ga0371429_10691084	0.38	NA	0.32	1.20	NA	Spirulina	KO:K02694	psaF	photosystem I subunit III	Photosynthesis
Ga0371429_10583822	0.96	-0.89	0.23	-0.36	-0.38	Unbinned	KO:K02694	psaF	photosystem I subunit III	Photosynthesis
Ga0371428_1043625	-0.67	-0.41	-3.50	NA	-3.58	Phormidium	KO:K02694	psaF	photosystem I subunit III	Photosynthesis
Ga0371428_1010819	-1.02	-0.20	-1.35	-2.11	-0.84	Unbinned	KO:K02699	psaL	photosystem I subunit XI	Photosynthesis
Ga0371428_1003884	-0.86	0.32	-1.64	-2.45	-1.82	Phormidium	KO:K02699	psaL	photosystem I subunit XI	Photosynthesis
Ga0371428_1009538	0.53	0.23	-2.15	NA	NA	Phormidium	KO:K08902	psb27	photosystem II Psb27 protein	Photosynthesis
Ga0371428_1089492	0.43	-0.26	-2.94	NA	-1.30	Phormidium	KO:K08903	psb28	photosystem II 13kDa protein	Photosynthesis
Ga0371428_10770751	-0.68	0.03	-0.99	-0.95	0.14	Unbinned	KO:K02703	psbA	photosystem II P680 reaction center D1 protein	Photosynthesis
Ga0371429_105730310	0.36	0.98	0.52	NA	NA	Unbinned	KO:K02703	psbA	photosystem II P680 reaction center D1 protein	Photosynthesis
Ga0371428_10020420	-0.85	0.77	-0.94	NA	NA	Phormidium	KO:K02703	psbA	photosystem II P680 reaction center D1 protein	Photosynthesis
Ga0371428_1040383	0.23	-0.12	-1.57	-0.29	-1.08	Phormidium	KO:K02704	psbB	photosystem II CP47 chlorophyll apoprotein	Photosynthesis
Ga0371428_1088461	0.32	-0.19	-1.70	NA	-2.34	Unbinned	KO:K02705	psbC	photosystem II CP43 chlorophyll apoprotein	Photosynthesis
Ga0371429_10343112	0.35	-0.64	0.35	NA	NA	Spirulina	KO:K02705	psbC	photosystem II CP43 chlorophyll apoprotein	Photosynthesis
Ga0371428_10108113	0.50	0.45	-0.50	NA	-1.50	Unbinned	KO:K02707	psbE	photosystem II cytochrome b559 subunit alpha	Photosynthesis
Ga0371428_10108137	-1.41	1.68	-0.58	NA	NA	Unbinned	KO:K02709	psbH	photosystem II PsbH protein	Photosynthesis
Ga0371431_11172544	0.76	-0.68	-2.08	NA	0.17	Unbinned	KO:K02716	psbO	photosystem II oxygen-evolving enhancer protein 1	Photosynthesis
Ga0371429_10558172	-0.68	0.36	-0.66	NA	NA	NA	KO:K02716	psbO	photosystem II oxygen-evolving enhancer protein 1	Photosynthesis
Ga0371428_1042066	0.06	-0.27	-3.05	NA	-2.09	Phormidium	KO:K02716	psbO	photosystem II oxygen-evolving enhancer protein 1	Photosynthesis
Ga0371428_1088502	0.20	-0.26	-2.64	NA	-0.97	Phormidium	KO:K02717	psbP	photosystem II oxygen-evolving enhancer protein 2	Photosynthesis
Ga0371429_10584901	3.18	NA	0.46	NA	0.56	Unknown	KO:K02719	psbU	photosystem II PsbU protein	Photosynthesis
Ga0371428_1069368	-0.30	-0.17	-3.25	NA	NA	Phormidium	KO:K02719	psbU	photosystem II PsbU protein	Photosynthesis
Ga0371428_10108167	-0.90	-0.18	-0.99	-1.85	-0.55	Unbinned	KO:K02720	psbV	photosystem II cytochrome c550	Photosynthesis
Ga0371431_11188266	2.92	-0.19	1.35	NA	NA	Planktothrix	KO:K02720	psbV	photosystem II cytochrome c550	Photosynthesis
Ga0371427_10390301	0.17	0.43	-1.04	NA	NA	NA	pfam11264	THF1	Thylakoid formation protein	Photosynthesis

Ga0371428_10108182	-0.89	0.23	-0.82	NA	0.88	Unbinned	KO:K03405	chlI, bchI	magnesium chelatase subunit I	Porphyrin and chlorophyll metabolism
Ga0371429_10983511	NA	1.12	-0.99	NA	-0.72	Unbinned	pfam00504	Chloroa_b-bind	Chlorophyll A-B binding protein	Porphyrin and chlorophyll metabolism
Ga0371428_1027461	-0.25	0.62	-3.03	NA	NA	Unbinned	KO:K10960	chlP, bchP	geranylgeranyl diphosphate/geranylgeranyl-bacteriochlorophyllide a reductase	Porphyrin and chlorophyll metabolism
Ga0371428_1066381	-0.06	-0.03	0.15	1.28	NA	Phormidium	KO:K10960	chlP, bchP	geranylgeranyl diphosphate/geranylgeranyl-bacteriochlorophyllide a reductase	Porphyrin and chlorophyll metabolism
Ga0371428_1042062	0.44	-0.06	-0.88	NA	NA	Phormidium	KO:K00228	CPOX, hemF	coproporphyrinogen III oxidase	Porphyrin and chlorophyll metabolism
Ga0371428_1071265	-0.54	-0.17	-3.21	NA	NA	Phormidium	KO:K01885	EARS, gltX	glutamyl-tRNA synthetase	Porphyrin and chlorophyll metabolism
Ga0371428_1146843	0.14	-0.99	-2.76	NA	NA	Phormidium	KO:K01749	hemC, HMBS	hydroxymethylbilane synthase	Porphyrin and chlorophyll metabolism
Ga0371429_10750353	-1.21	0.33	-0.73	NA	NA	Pseudanabaena	KO:K01845	hemL	glutamate-1-semialdehyde 2,1-aminomutase	Porphyrin and chlorophyll metabolism
Ga0371429_10220502	-0.75	0.89	0.01	NA	NA	Unbinned	KO:K01845	hemL	glutamate-1-semialdehyde 2,1-aminomutase	Porphyrin and chlorophyll metabolism
Ga0371428_1010701	-0.23	-0.24	-0.10	0.55	0.12	Phormidium	KO:K01845	hemL	glutamate-1-semialdehyde 2,1-aminomutase	Porphyrin and chlorophyll metabolism
Ga0371428_1034552	0.55	-0.36	-1.69	NA	-0.71	Phormidium	KO:K00510	HMOX1	heme oxygenase 1	Porphyrin and chlorophyll metabolism
Ga0371476_12167301	0.31	-0.47	-0.04	-0.23	-1.17	Unbinned	KO:K05692	ACTB_G1	actin beta/gamma 1	Signal transduction
Ga0371429_10013692	1.67	4.08	1.18	NA	NA	Unbinned	KO:K05692	ACTB_G1	actin beta/gamma 1	Signal transduction
Ga0371430_11561625	0.29	2.33	0.01	NA	NA	Gammaproteobacteria	KO:K05692	ACTB_G1	actin beta/gamma 1	Signal transduction
Ga0371430_10792511	0.29	0.13	-0.01	0.85	-0.04	NA	KO:K05692	ACTB_G1	actin beta/gamma 1	Signal transduction
Ga0371435_10292621	-0.06	0.68	-0.80	-0.93	NA	Unbinned	COG0784	cheY	signal transduction histidine kinase/CheY-like chemotaxis protein	Signal transduction
Ga0371428_1021734	-0.69	-0.36	-0.07	-0.32	-0.30	Gammaproteobacteria	COG4564	COG4564	Signal transduction histidine kinase	Signal transduction
Ga0371431_10923372	NA	-0.29	-0.85	1.79	NA	Desulfobacteraceae	COG4564	COG4564	Signal transduction histidine kinase	Signal transduction
Ga0371428_1028773	-0.25	0.72	NA	1.46	0.37	Phormidium	COG4564	COG4564	Signal transduction histidine kinase	Signal transduction
Ga0371437_11153703	NA	0.02	-0.36	NA	-1.03	Desulfobulbaceae	COG5001	COG5001	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	Signal transduction
Ga0371428_1121972	-0.68	0.47	-0.03	1.54	NA	Gammaproteobacterium_117915_bin_15	KO:K08738	CYC	cytochrome c	Signal transduction
Ga0371428_1016702	-0.27	0.76	0.08	1.22	0.40	Gammaproteobacteria	KO:K08738	CYC	cytochrome c	Signal transduction
Ga0371428_1165834	0.32	0.50	-0.09	NA	NA	Gammaproteobacteria	COG4654, COG4654	cytC552	Cytochrome c551/c552	Signal transduction

Ga0371428_1148763	NA	0.20	-0.17	NA	0.50	Gammaproteobacterium_117915_bin_15	COG2863, COG2863	cytc553	cytochrome c553	Signal transduction
Ga0371428_1109513	-0.22	0.29	0.38	0.26	1.72	Gammaproteobacteria	COG2863, COG2863	cytc553	cytochrome c553	Signal transduction
Ga0371428_1157634	-0.26	0.84	0.21	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K11688	dctP	C4-dicarboxylate-binding protein DctP	Signal transduction
Ga0371428_1007002	2.44	-0.02	0.79	0.76	NA	Gammaproteobacteria	KO:K04771	degP, htrA	serine protease Do	Signal transduction
Ga0371428_10456819	-0.07	0.60	-3.44	NA	-2.94	Phormidium	COG5126	FRQ1	Ca2+-binding protein, EF-hand superfamily	Signal transduction
Ga0371428_1006979	0.47	0.71	-1.36	-0.68	NA	Phormidium	KO:K04751	glnB	nitrogen regulatory protein P-II 1	Signal transduction
Ga0371428_1036062	-1.21	0.52	0.84	NA	NA	Unbinned	KO:K03406	mcp	methyl-accepting chemotaxis protein	Signal transduction
Ga0371429_11115202	2.94	0.34	-3.03	NA	-2.01	Oscillatoria	KO:K02650	pilA	type IV pilus assembly protein PilA	Signal transduction
Ga0371430_10026961	-1.95	1.92	0.56	NA	-0.41	Gammaproteobacteria	KO:K02650	pilA	type IV pilus assembly protein PilA	Signal transduction
Ga0371427_10277512	-0.93	-0.10	-0.60	NA	-0.21	Unbinned	KO:K05863	SLC25A4S, ANT	solute carrier family 25 (mitochondrial adenine nucleotide translocator), member 4/5/6/31	Signal transduction
Ga0371428_1058032	-0.23	0.09	0.41	NA	-0.48	Gammaproteobacteria	KO:K04564	SOD2	superoxide dismutase, Fe-Mn family	Signal transduction
Ga0371427_10956933	-0.54	0.48	0.00	NA	NA	NA	KO:K04564	SOD2	superoxide dismutase, Fe-Mn family	Signal transduction
Ga0371428_1042954	0.43	0.26	-2.55	NA	NA	Phormidium	KO:K04564	SOD2	superoxide dismutase, Fe-Mn family	Signal transduction
Ga0371428_1010818	0.00	0.21	-0.85	NA	NA	Unbinned	COG0464	spoVK	AAA+-type ATPase, SpoVK/Ycf46/Vps4 family	Signal transduction
Ga0371428_1040424	0.55	-0.24	-3.41	NA	NA	Phormidium	COG0589	uspA	Nucleotide-binding universal stress protein, UspA family	Signal transduction
Ga0371428_1063971	-0.14	-0.27	-3.38	NA	NA	Phormidium	COG0589, COG0589	uspA	Nucleotide-binding universal stress protein, UspA family	Signal transduction
Ga0371476_12819442	NA	NA	-0.64	-1.20	-1.46	Unbinned	KO:K08641	vanX	zinc D-Ala-D-Ala dipeptidase	Signal transduction
Ga0371429_10661834	0.83	NA	1.78	3.56	NA	Pseudanabaena	COG0678	ahp1	Peroxiredoxin [Posttranslational modification, protein turnover, chaperones]	Signaling and cellular processes
Ga0371431_10792484	2.63	NA	2.06	3.12	NA	Pseudanabaena	COG0678	ahp1	Peroxiredoxin [Posttranslational modification, protein turnover, chaperones]	Signaling and cellular processes
Ga0371428_1006397	-0.04	-0.45	-3.03	NA	-2.20	Phormidium	COG0678, COG0695	ahp1/grxC	Peroxiredoxin/Glutaredoxin [Posttranslational modification, protein turnover, chaperones]	Signaling and cellular processes
Ga0371427_10012382	-1.97	0.87	-0.92	NA	-0.39	Unbinned	KO:K17099	ANXA13	annexin A13	Signaling and cellular processes
Ga0371429_10046591	-0.39	0.41	-0.49	NA	NA	NA	KO:K06758	CHL1	L1 cell adhesion molecule like protein	Signaling and cellular processes
Ga0371430_10886383	NA	0.55	0.16	1.97	NA	NA	KO:K06758	CHL1	L1 cell adhesion molecule like protein	Signaling and cellular processes
Ga0371476_10940982	0.01	-0.16	-0.92	NA	NA	Bacteria	COG5555	COG5555	Cytolysin, a secreted calcineurin-like phosphatase	Signaling and cellular processes
Ga0371428_1036385	-0.76	0.27	0.47	NA	0.41	Gammaproteobacterium_117915_bin_15	COG1555	comEA	DNA uptake protein ComE and related DNA-binding proteins	Signaling and cellular processes
Ga0371428_1116812	-0.11	0.44	0.63	NA	1.35	Gammaproteobacteria	COG1555, COG1555	comEA	DNA uptake protein ComE and related DNA-binding proteins	Signaling and cellular processes
Ga0371429_10415483	2.52	1.03	-1.21	NA	NA	Oscillatoria	COG3678	cpxP	Periplasmic protein refolding chaperone Spy/CpxP family [Posttranslational modification, protein turnover, chaperones]	Signaling and cellular processes
Ga0371428_1136712	-0.66	-0.29	0.04	-0.15	0.59	Gammaproteobacteria	KO:K03386	E1.11.1.15, PRDX, ahpC	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	Signaling and cellular processes
Ga0371429_10317684	0.89	0.99	0.53	NA	NA	Unbinned	KO:K03386	E1.11.1.15, PRDX, ahpC	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	Signaling and cellular processes

Ga0371430_10087181	2.67	-0.95	0.67	NA	NA	Unbinned	KO:K03386	E1.11.1.15, PRDX, ahpC	peroxiredoxin (alkyl hydroperoxide reductase subunit C)	Signaling and cellular processes
Ga0371428_1011739	-1.42	-0.16	0.26	NA	-0.18	Gammaproteobacteria	COG2825	hlpA	Periplasmic chaperone for outer membrane proteins, Skp family [Cell wall/membrane/envelope biogenesis, Posttranslational modification, protein turnover, chaperones]	Signaling and cellular processes
Ga0371430_10342382	0.00	0.51	-0.52	NA	NA	Unbinned	COG1360	motB	Flagellar motor protein MotB	Signaling and cellular processes
Ga0371428_1142061	0.33	0.65	0.88	NA	NA	Desulfobacteraceae	COG2885	ompA	Outer membrane protein OmpA and related peptidoglycan-associated (lipo)proteins	Signaling and cellular processes
Ga0371430_11825691	-2.02	-0.15	-0.29	NA	NA	Comamonadaceae	COG3203	ompC	Outer membrane protein (porin)	Signaling and cellular processes
Ga0371431_10215442	0.56	0.10	-0.38	-0.13	-0.08	Unbinned	COG3203	ompC	Outer membrane protein (porin)	Signaling and cellular processes
Ga0371428_1017782	0.09	0.27	0.31	NA	NA	Gammaproteobacteria	KO:K07275	ompW	outer membrane protein	Signaling and cellular processes
Ga0371429_101937210	0.96	-1.86	1.03	0.89	0.16	Spirulina	COG3659	oprB	Carbohydrate-selective porin OprB	Signaling and cellular processes
Ga0371428_10154212	0.19	-0.16	-3.03	NA	NA	Phormidium	KO:K18640	parM	plasmid segretation protein ParM	Signaling and cellular processes
Ga0371431_10276152	NA	1.00	-1.19	NA	-1.33	Bacteroidales	COG1520, COG1520	PQQ	Outer membrane protein assembly factor BamB, contains PQQ-like beta-propeller repeat [Cell wall/membrane/envelope biogenesis]	Signaling and cellular processes
Ga0371428_1008633	0.14	-0.32	0.07	0.07	1.62	Gammaproteobacteria	COG3793	terB	Tellurite resistance protein	Signaling and cellular processes
Ga0371428_1089324	0.13	-0.59	-3.05	NA	-2.54	Phormidium	KO:K05795	terD	tellurite resistance protein TerD	Signaling and cellular processes
Ga0371428_1173461	-0.91	0.04	-0.49	NA	NA	NA	KO:K07374	TUBA	tubulin alpha	Signaling and cellular processes
Ga0371429_10274112	0.77	2.91	0.95	NA	NA	NA	KO:K07374	TUBA	tubulin alpha	Signaling and cellular processes
Ga0371428_1061033	-0.19	0.49	0.50	0.91	1.93	NA	KO:K07375	TUBB	tubulin beta	Signaling and cellular processes
Ga0371428_1059183	0.08	0.60	0.91	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K06207	typA, bipA	GTP-binding protein	Signaling and cellular processes
Ga0371428_1027441	0.00	0.29	1.00	NA	NA	Phormidium	KO:K06207	typA, bipA	GTP-binding protein	Signaling and cellular processes
Ga0371427_10089442	-3.10	0.65	-0.38	NA	NA	Flavobacterium	COG0210	uvrD	Superfamily I DNA or RNA helicase	Signaling and cellular processes
Ga0371432_10776291	-1.78	-1.16	NA	NA	-3.76	Unbinned	COG4646, COG0827	ytxK	Adenine-specific DNA methylase, N12 class	Signaling and cellular processes
Ga0371427_10872941	-2.00	0.17	-0.42	-0.87	-1.02	NA	KO:K06630	YWHAE	14-3-3 protein epsilon	Signaling and cellular processes
Ga0371428_1052838	-0.80	0.60	-0.18	-0.73	-0.01	Desulfobacteraceae	KO:K00394	aprA	adenylylsulfate reductase, subunit A	Sulfur metabolism
Ga0371428_1154663	-0.86	0.54	-0.13	NA	NA	Desulfobacteraceae	KO:K00394	aprA	adenylylsulfate reductase, subunit A	Sulfur metabolism
Ga0371430_10214921	-0.12	-0.48	-0.13	0.86	-0.08	Unbinned	KO:K00394	aprA	adenylylsulfate reductase, subunit A	Sulfur metabolism
Ga0371428_1052837	0.27	0.05	-0.24	1.39	0.71	Desulfobacteraceae	KO:K00395	aprB	adenylylsulfate reductase, subunit B	Sulfur metabolism
Ga0371428_1154662	-0.32	-0.25	0.10	NA	NA	Desulfobacteraceae	KO:K00395	aprB	adenylylsulfate reductase, subunit B	Sulfur metabolism

Ga0371429_10465142	1.47	-1.13	1.44	1.98	NA	Spirulina	COG1404	aprE	Serine protease, subtilisin family [Posttranslational modification, protein turnover, chaperones]	Sulfur metabolism
Ga0371428_1010041	-0.62	0.37	0.12	-0.03	1.25	Unbinned	KO:K07306	dmsA	anaerobic dimethyl sulfoxide reductase subunit A	Sulfur metabolism
Ga0371428_1013401	-1.30	0.53	0.02	0.88	1.11	Unbinned	KO:K07306	dmsA	anaerobic dimethyl sulfoxide reductase subunit A	Sulfur metabolism
Ga0371428_1122881	-0.84	0.37	-0.06	-0.11	-0.57	Unbinned	KO:K07306	dmsA	anaerobic dimethyl sulfoxide reductase subunit A	Sulfur metabolism
Ga0371428_1010042	-2.09	0.16	0.39	-0.14	1.10	Unbinned	KO:K07307	dmsB	anaerobic dimethyl sulfoxide reductase subunit B	Sulfur metabolism
Ga0371430_10128293	NA	NA	-0.30	0.47	-0.26	Unbinned	KO:K07307	dmsB	anaerobic dimethyl sulfoxide reductase subunit B	Sulfur metabolism
Ga0371428_10649910	0.19	1.09	-0.06	NA	NA	Desulfobacteraceae	KO:K11180	dsrA	dissimilatory sulfite reductase alpha subunit	Sulfur metabolism
Ga0371428_1163591	-0.58	1.43	0.73	NA	NA	Desulfobacteraceae	KO:K11180	dsrA	dissimilatory sulfite reductase alpha subunit	Sulfur metabolism
Ga0371428_1135681	-0.51	0.17	0.37	NA	0.02	Unbinned	KO:K11180	dsrA	dissimilatory sulfite reductase alpha subunit	Sulfur metabolism
Ga0371428_1064999	-1.16	1.48	-0.02	NA	NA	Desulfobacteraceae	KO:K11181	dsrB	dissimilatory sulfite reductase beta subunit	Sulfur metabolism
Ga0371429_10640591	NA	0.46	0.64	NA	-0.15	Gammaproteobacteria	KO:K11181	dsrB	dissimilatory sulfite reductase beta subunit	Sulfur metabolism
Ga0371430_11014352	-0.31	0.50	0.13	0.94	NA	Unbinned	KO:K11181	dsrB	dissimilatory sulfite reductase beta subunit	Sulfur metabolism
Ga0371428_1009684	NA	-0.02	0.45	0.74	0.79	Gammaproteobacterium_117915_bin_15	COG2210	dsrE2	Peroxiredoxin family protein [Energy production and conversion]	Sulfur metabolism
Ga0371428_1107443	-0.75	-0.09	0.29	1.33	0.43	Gammaproteobacterium_117915_bin_15	KO:K17229	fccB	sulfide dehydrogenase	Sulfur metabolism
Ga0371428_1036098	0.41	-1.15	0.63	NA	1.57	Unbinned	KO:K17229	fccB	sulfide dehydrogenase	Sulfur metabolism
Ga0371428_10322164	0.53	-0.52	-2.76	NA	NA	Phormidium	COG0491	gloB	Glyoxylase or a related metal-dependent hydrolase, beta-lactamase superfamily II	Sulfur metabolism
Ga0371428_1146203	-0.54	1.28	0.39	NA	NA	Desulfobacteraceae	KO:K00958	sat, met3	sulfate adenylyltransferase	Sulfur metabolism
Ga0371428_1044222	-0.27	-0.38	0.09	NA	NA	Gammaproteobacterium_117915_bin_15	KO:K17226	soxY	sulfur-oxidizing protein SoxY	Sulfur metabolism
Ga0371431_11327791	0.01	-0.30	-0.47	NA	NA	Unbinned	pfam14250	abrB	AbrB-like transcriptional regulator	Transcription
Ga0371428_1092934	-0.42	0.36	-2.12	NA	NA	Phormidium	pfam14250	abrB	AbrB-like transcriptional regulator	Transcription
Ga0371431_10471572	0.68	NA	1.06	NA	-2.55	Planktothrix	pfam14250	abrB	AbrB-like transcriptional regulator	Transcription
Ga0371435_11867512	-0.26	0.81	-3.18	NA	-3.60	Unbinned	COG1747	COG1747	Uncharacterized N-terminal domain of the transcription elongation factor GreA [Function unknown]	Transcription
Ga0371428_1038013	-0.86	0.30	0.47	0.15	1.08	Gammaproteobacteria	KO:K03704	cspA	cold shock protein	Transcription
Ga0371428_1080056	-0.01	-0.58	-0.32	NA	NA	Chloroflexaceae	COG1396	hipB	Transcriptional regulator, contains XRE-family HTH domain	Transcription
Ga0371428_1034392	-1.57	0.18	-0.05	-1.15	0.29	Gammaproteobacteria	KO:K13643	iscR	Rrf2 family transcriptional regulator, iron-sulfur cluster assembly transcription factor	Transcription
Ga0371428_10200626	0.51	0.31	0.07	NA	NA	Phormidium	KO:K03040	rpoA	DNA-directed RNA polymerase subunit alpha	Transcription
Ga0371428_1048482	-0.93	0.68	NA	0.32	NA	Gammaproteobacteria	KO:K03043	rpoB	DNA-directed RNA polymerase subunit beta	Transcription
Ga0371428_1057691	0.03	-0.46	-1.06	NA	NA	Unbinned	KO:K03043	rpoB	DNA-directed RNA polymerase subunit beta	Transcription
Ga0371428_1114431	-1.07	0.30	1.43	NA	0.39	Unbinned	KO:K07315	rsbU_P	phosphoserine phosphatase RsbU/P	Transcription
Ga0371428_1132821	-0.84	0.60	-0.45	-1.38	0.25	Bacteroidales	KO:K01883	CARS, cysS	cysteinyl-tRNA synthetase	Translation
Ga0371429_11098151	-0.87	0.17	-0.35	-0.84	-1.26	NA	KO:K03231	EEF1A	elongation factor 1-alpha	Translation
Ga0371427_11071181	-0.27	0.10	-0.30	NA	NA	NA	KO:K03233	EEF1G	elongation factor 1-gamma	Translation
Ga0371428_1074933	0.19	-0.10	-1.29	NA	NA	Phormidium	KO:K02356	efp	elongation factor P	Translation
Ga0371437_10520791	NA	-0.19	-0.89	NA	1.18	Desulfobacteraceae	COG1234	elaC	Ribonuclease BN, tRNA processing enzyme	Translation
Ga0371428_1029151	NA	0.48	0.31	NA	0.65	Unbinned	KO:K02838	frr	ribosome recycling factor	Translation
Ga0371428_1065611	-2.38	0.44	0.34	NA	0.90	Gammaproteobacteria	KO:K02355	fusA, GFM, EFG	elongation factor G	Translation
Ga0371435_12621271	-0.45	-0.15	-1.84	NA	NA	Unbinned	KO:K02519	infB, MTIF2	translation initiation factor IF-2	Translation

Ga0371436_11038001	-0.72	-0.57	-1.04	NA	NA	Proteobacteria	COG0456	rimI	Ribosomal protein S18 acetylase RimI and related acetyltransferases	Translation
Ga0371428_10108174	-0.06	-0.01	-0.06	NA	NA	Unbinned	KO:K02863	RP-L1, MRPL1, rplA	large subunit ribosomal protein L1	Translation
Ga0371428_1050872	-0.26	0.06	0.34	1.14	0.83	Unbinned	KO:K02863	RP-L1, MRPL1, rplA	large subunit ribosomal protein L1	Translation
Ga0371428_1138623	-0.71	0.12	0.58	-0.96	0.20	Unbinned	KO:K02864	RP-L10, MRPL10, rplJ	large subunit ribosomal protein L10	Translation
Ga0371429_10939732	0.19	0.30	-1.84	NA	NA	NA	KO:K02864	RP-L10, MRPL10, rplJ	large subunit ribosomal protein L10	Translation
Ga0371428_1023948	-0.03	0.27	-3.62	NA	NA	Phormidium	KO:K02864	RP-L10, MRPL10, rplJ	large subunit ribosomal protein L10	Translation
Ga0371428_1023946	0.49	-0.28	-4.76	NA	NA	Phormidium	KO:K02867	RP-L11, MRPL11, rplK	large subunit ribosomal protein L11	Translation
Ga0371428_10200632	-0.41	0.30	-2.81	NA	NA	Phormidium	KO:K02876	RP-L15, MRPL15, rplO	large subunit ribosomal protein L15	Translation
Ga0371428_1049578	0.41	-0.70	-2.46	NA	NA	Phormidium	KO:K02887	RP-L20, MRPL20, rplT	large subunit ribosomal protein L20	Translation
Ga0371428_10200644	0.05	0.16	-3.04	NA	NA	Phormidium	KO:K02890	RP-L22, MRPL22, rplV	large subunit ribosomal protein L22	Translation
Ga0371429_10628481	-0.85	0.35	-0.68	NA	NA	NA	KO:K02891	RP-L22e, RPL22	large subunit ribosomal protein L22e	Translation
Ga0371428_10770737	-0.92	0.45	-0.46	NA	NA	Unbinned	KO:K02892	RP-L23, MRPL23, rplW	large subunit ribosomal protein L23	Translation
Ga0371428_10200647	-0.02	-0.10	-2.46	NA	NA	Phormidium	KO:K02892	RP-L23, MRPL23, rplW	large subunit ribosomal protein L23	Translation
Ga0371428_1077883	-0.61	0.53	0.11	NA	NA	NA	KO:K02895	RP-L24, MRPL24, rplX	large subunit ribosomal protein L24	Translation
Ga0371428_1063716	0.32	1.22	-0.04	NA	NA	Gammaproteobacteria	KO:K02897	RP-L25, rplY	large subunit ribosomal protein L25	Translation
Ga0371427_10625753	-0.12	-0.33	-0.62	NA	NA	NA	KO:K02901	RP-L27e, RPL27	large subunit ribosomal protein L27e	Translation
Ga0371428_1120034	NA	0.71	0.25	NA	0.80	Gammaproteobacteria	KO:K02902	RP-L28, MRPL28, rpmB	large subunit ribosomal protein L28	Translation
Ga0371428_10770739	-1.55	0.45	0.00	NA	NA	Unbinned	KO:K02906	RP-L3, MRPL3, rplC	large subunit ribosomal protein L3	Translation

Ga0371431_10193591	NA	0.28	0.60	NA	0.25	Unbinned	KO:K02906	RP-L3, MRPL3, rplC	large subunit ribosomal protein L3	Translation
Ga0371428_1049579	-0.05	-0.48	-1.71	NA	-0.31	Phormidium	KO:K02916	RP-L35, MRPL35, rpmI	large subunit ribosomal protein L35	Translation
Ga0371428_1047431	0.27	-0.24	0.34	NA	NA	Unbinned	KO:K02926	RP-L4, MRPL4, rplD	large subunit ribosomal protein L4	Translation
Ga0371429_10472722	-1.17	0.58	0.28	NA	NA	Unbinned	KO:K02926	RP-L4, MRPL4, rplD	large subunit ribosomal protein L4	Translation
Ga0371437_11599363	-0.30	-2.35	NA	NA	0.56	Unknown	KO:K02931	RP-L5, MRPL5, rplE	large subunit ribosomal protein L5	Translation
Ga0371428_1060731	0.64	-0.16	0.32	-0.33	0.26	Unbinned	KO:K02933	RP-L6, MRPL6, rplF	large subunit ribosomal protein L6	Translation
Ga0371436_10433123	0.42	0.40	0.04	-1.19	NA	Unbinned	KO:K02933	RP-L6, MRPL6, rplF	large subunit ribosomal protein L6	Translation
Ga0371428_10200635	0.45	-0.27	-3.27	NA	NA	Phormidium	KO:K02933	RP-L6, MRPL6, rplF	large subunit ribosomal protein L6	Translation
Ga0371428_1024065	-0.34	0.19	-0.49	NA	NA	Desulfobacteraceae	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371428_10108175	-2.35	0.56	-1.45	NA	-0.54	Unbinned	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371428_1138622	-1.10	0.20	0.43	-0.04	0.49	Unbinned	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371429_10416144	NA	-0.93	-0.08	-0.75	0.19	Pseudanabaena	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371429_10749222	0.38	-0.35	0.15	NA	0.78	Spirulina	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371429_10202112	1.17	-0.34	-2.05	NA	-1.09	Unbinned	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371431_108033518	0.79	-0.44	0.33	NA	0.85	Planktothrix	KO:K02935	RP-L7, MRPL12, rplL	large subunit ribosomal protein L7/L12	Translation
Ga0371428_1099701	-0.03	-0.47	0.18	0.28	-0.03	Unbinned	KO:K02945	RP-S1, rpsA	small subunit ribosomal protein S1	Translation
Ga0371428_10328021	0.10	1.23	0.38	NA	NA	Desulfobacteraceae	KO:K02946	RP-S10, MRPS10, rpsJ	small subunit ribosomal protein S10	Translation

Ga0371429_10111562	-0.28	0.76	0.77	NA	0.02	Gammaproteobacteria	KO:K02358, KO:K02946	RP-S10, MRPS10, rpsJ	small subunit ribosomal protein S10	Translation
Ga0371429_10227612	-0.10	NA	-1.72	NA	-1.30	Unbinned	KO:K02946	RP-S10, MRPS10, rpsJ	small subunit ribosomal protein S10	Translation
Ga0371428_10200627	-0.19	-0.04	-1.53	NA	NA	Phormidium	KO:K02948	RP-S11, MRPS11, rpsK	small subunit ribosomal protein S11	Translation
Ga0371428_10200628	-0.21	-0.35	-2.62	NA	NA	Phormidium	KO:K02952	RP-S13, rpsM	small subunit ribosomal protein S13	Translation
Ga0371428_10225033	-0.22	-0.51	-2.49	NA	NA	Phormidium	KO:K02954	RP-S14, MRPS14, rpsN	small subunit ribosomal protein S14	Translation
Ga0371428_1085282	NA	-0.02	0.51	NA	0.44	Gammaproteobacteria	KO:K02959	RP-S16, MRPS16, rpsP	small subunit ribosomal protein S16	Translation
Ga0371428_10770749	-0.55	-0.37	NA	NA	1.01	Unbinned	KO:K02959	RP-S16, MRPS16, rpsP	small subunit ribosomal protein S16	Translation
Ga0371429_10120644	-0.18	-0.60	-1.83	NA	-0.83	Unknown	KO:K02959	RP-S16, MRPS16, rpsP	small subunit ribosomal protein S16	Translation
Ga0371428_1117263	0.57	0.28	0.51	1.43	1.22	Gammaproteobacteria	KO:K02961	RP-S17, MRPS17, rpsQ	small subunit ribosomal protein S17	Translation
Ga0371428_1057203	NA	1.02	0.44	NA	1.76	Gammaproteobacteria	KO:K02967	RP-S2, MRPS2, rpsB	small subunit ribosomal protein S2	Translation
Ga0371428_1061403	NA	-0.19	0.35	-0.60	0.58	Gammaproteobacteria	KO:K02967	RP-S2, MRPS2, rpsB	small subunit ribosomal protein S2	Translation
Ga0371428_10108151	-0.23	0.24	-0.53	NA	NA	Unbinned	KO:K02967	RP-S2, MRPS2, rpsB	small subunit ribosomal protein S2	Translation
Ga0371428_1033212	0.15	-0.73	-2.87	NA	NA	Unbinned	KO:K02967	RP-S2, MRPS2, rpsB	small subunit ribosomal protein S2	Translation
Ga0371476_11685875	-0.39	0.08	-3.62	NA	NA	Unbinned	KO:K02968	RP-S20, rpsT	small subunit ribosomal protein S20	Translation
Ga0371429_10187752	-0.22	-0.02	-2.99	NA	NA	Oscillatoria	KO:K02968	RP-S20, rpsT	small subunit ribosomal protein S20	Translation
Ga0371428_10013718	0.20	-1.21	-2.16	NA	NA	Phormidium	KO:K02970	RP-S21, MRPS21, rpsU	small subunit ribosomal protein S21	Translation
Ga0371427_10468711	-0.10	0.04	-0.69	NA	NA	NA	KO:K02974	RP-S24e, RPS24	small subunit ribosomal protein S24e	Translation
Ga0371433_10103893	NA	0.33	0.48	NA	1.28	Unbinned	KO:K02982	RP-S3, rpsC	small subunit ribosomal protein S3	Translation

Ga0371427_10552251	-1.22	0.32	-0.84	NA	NA	Unbinned	KO:K02984	RP-S3Ae, RPS3A	small subunit ribosomal protein S3Ae	Translation
Ga0371428_1142753	-0.76	-0.05	0.23	-0.32	0.98	Unbinned	KO:K02988	RP-S5, MRPS5, rpsE	small subunit ribosomal protein S5	Translation
Ga0371428_10200633	0.34	0.08	-2.03	NA	NA	Phormidium	KO:K02988	RP-S5, MRPS5, rpsE	small subunit ribosomal protein S5	Translation
Ga0371428_1065612	0.85	-0.04	0.77	0.05	1.23	Gammaproteobacteria	KO:K02992	RP-S7, MRPS7, rpsG	small subunit ribosomal protein S7	Translation
Ga0371428_1014176	0.12	-0.43	-2.44	NA	0.11	Phormidium	KO:K02992	RP-S7, MRPS7, rpsG	small subunit ribosomal protein S7	Translation
Ga0371437_10956332	-2.26	0.09	-1.08	NA	-1.53	Unbinned	KO:K02994	RP-S8, rpsH	small subunit ribosomal protein S8	Translation
Ga0371428_10200636	0.50	0.27	-1.37	NA	NA	Phormidium	KO:K02994	RP-S8, rpsH	small subunit ribosomal protein S8	Translation
Ga0371428_10200622	-0.14	-0.03	-2.37	NA	NA	Phormidium	KO:K02996	RP-S9, MRPS9, rpsI	small subunit ribosomal protein S9	Translation
Ga0371428_1004701	1.79	NA	0.61	3.87	0.05	Unbinned	COG0724	RRM	RNA recognition motif (RRM) domain	Translation
Ga0371428_10189114	-0.07	0.00	-2.78	NA	NA	Phormidium	COG0724	RRM	RNA recognition motif (RRM) domain	Translation
Ga0371428_10310410	0.43	-0.68	-2.95	NA	NA	Phormidium	COG0724	RRM	RNA recognition motif (RRM) domain	Translation
Ga0371428_10584314	0.25	-0.73	NA	NA	-0.96	Phormidium	COG0724	RRM	RNA recognition motif (RRM) domain	Translation
Ga0371428_10108150	-2.04	0.49	-0.31	NA	NA	Unbinned	KO:K02357	tsf, TSFM	elongation factor Ts	Translation
Ga0371428_1018911	-0.36	-0.14	-1.36	NA	NA	Phormidium	KO:K02357	tsf, TSFM	elongation factor Ts	Translation
Ga0371478_1026371	2.35	0.33	0.15	NA	NA	Unbinned	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371428_1097943	-0.24	-0.35	-0.10	NA	NA	Bacteroidales	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371428_1012692	-0.75	-0.09	-0.06	-0.96	0.48	Desulfobacteraceae	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371428_1113971	1.70	0.19	1.00	2.34	1.57	Unbinned	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371428_1174801	NA	-0.37	0.27	0.75	0.52	Unbinned	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371429_10948443	0.30	-2.12	0.54	1.14	0.88	Spirulina	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371429_10111562	-0.28	0.76	0.77	NA	0.02	Gammaproteobacteria	KO:K02358, KO:K02946	tuf, TUFM	elongation factor Tu	Translation
Ga0371429_10531621	2.13	NA	1.26	2.06	1.15	Unbinned	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371430_10071441	-1.23	0.15	0.15	NA	NA	Unbinned	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371436_10050854	0.78	0.08	-2.39	NA	NA	Unknown	KO:K02358	tuf, TUFM	elongation factor Tu	Translation
Ga0371428_1087054	-0.22	0.66	0.60	NA	NA	Gammaproteobacteria	COG0845	acrA	Multidrug efflux pump subunit AcrA (membrane-fusion protein)	Transporter
Ga0371435_10307442	0.31	-0.32	-0.70	NA	NA	Unknown	COG0841	acrB	Multidrug efflux pump subunit AcrB [Defense mechanisms]	Transporter
Ga0371429_10132721	-0.48	-0.45	-4.71	NA	NA	Unbinned	COG2931, COG2931	COG2931	Ca2+-binding protein, RTX toxin-related [Secondary metabolites biosynthesis, transport and catabolism]	Transporter
Ga0371429_10658801	0.41	0.61	-3.07	NA	NA	Unbinned	COG2931	COG2931	Ca2+-binding protein, RTX toxin-related [Secondary metabolites biosynthesis, transport and catabolism]	Transporter
Ga0371439_10422483	0.19	0.98	0.20	NA	NA	Unbinned	KO:K07213	copZ, golB	copper chaperone	Transporter

Figure SI 4.1 Map of MIS sampling scheme in July 2016. Shown in red is the transect along which sampling was conducted in July 2016. The spokes from the origin closest to the Alcove are, clockwise, D, C, B, and A. Radiating from the origin O are cross lines 1, 2, 3, 4, 5, and 6. Phormidium-dominated flat mat was collected along the D line. Giraffe mat was observed from B4-A6. Other samples of flat purple mat and flat white mat were collected along the A line, and from A5-C5.

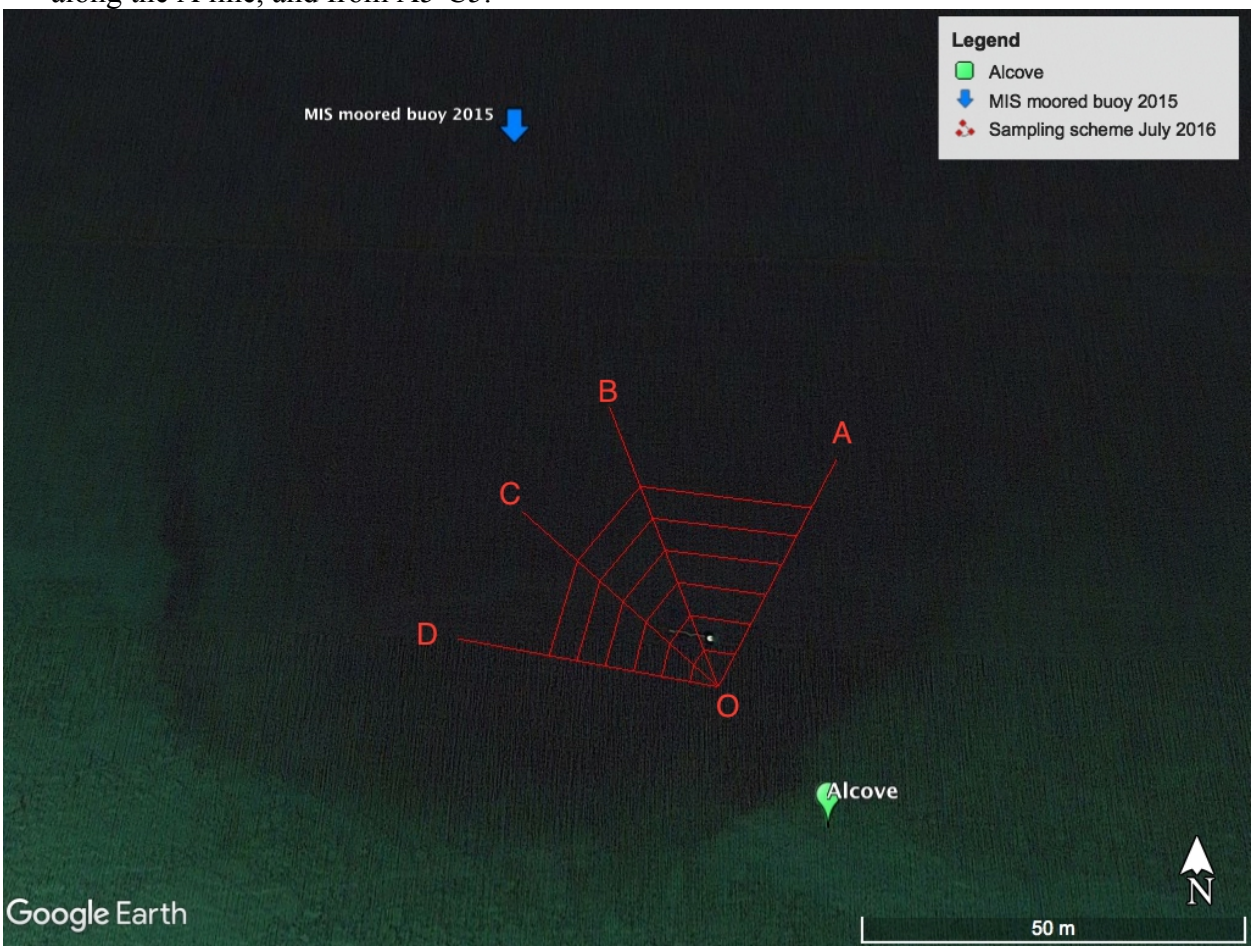


Figure SI 4.2 Nonmetric multidimensional scaling plot (NMDS) of the bacterial communities of different mat types. NMDS was conducted on a Morisita-Horn dissimilarity matrix calculated from the relative bacterial abundances. Samples are categorized by month (symbol outline color), year (symbol shape), and mat type (symbol fill color). Giraffe mat samples and white mat samples clustered with each other. Flat mat and finger mat collected in 2016 and 2017 were variable in composition and dissimilar to 2015 material.

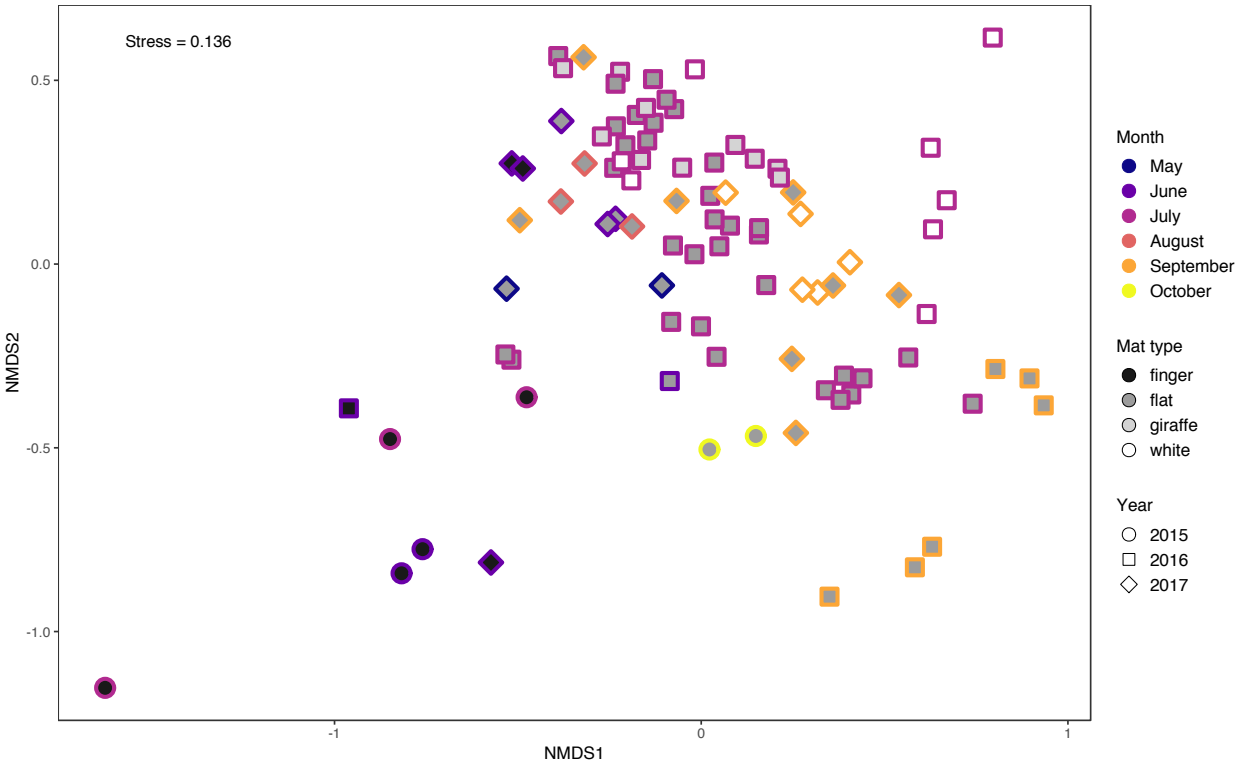


Figure SI 4.3 Ternary diagram categorizing flat mat samples by relative abundances of Phormidium, epsilonproteobacterial SOB (ESOB), and the sum of Spirulina and Pseudanabaena, among the sum of these members. Samples in the "Phormidium" corner had a higher relative abundance of Phormidium compared to the other 2 groups, and samples from pre-July 2016 typically fell in this triangle. Samples from 2016 are distributed throughout the ternary diagram. Samples from 2017 are typically <65% Phormidium, <50% ESOB, and >50% Spirulina and Pseudanabaena.

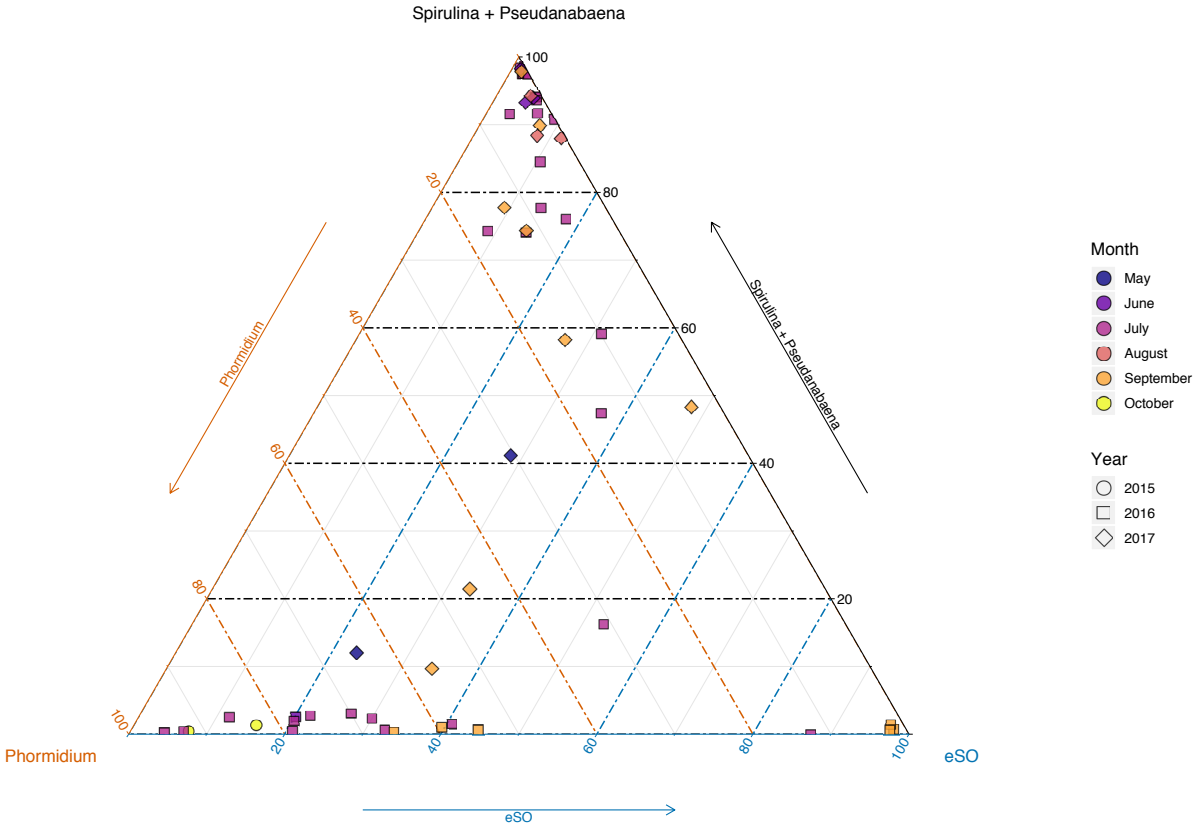


Figure SI 4.4 Histogram of the nearest-taxon-index (NTI) of the bacterial communities of different mat morphotypes. The average NTI for each community is plotted as the solid vertical line.

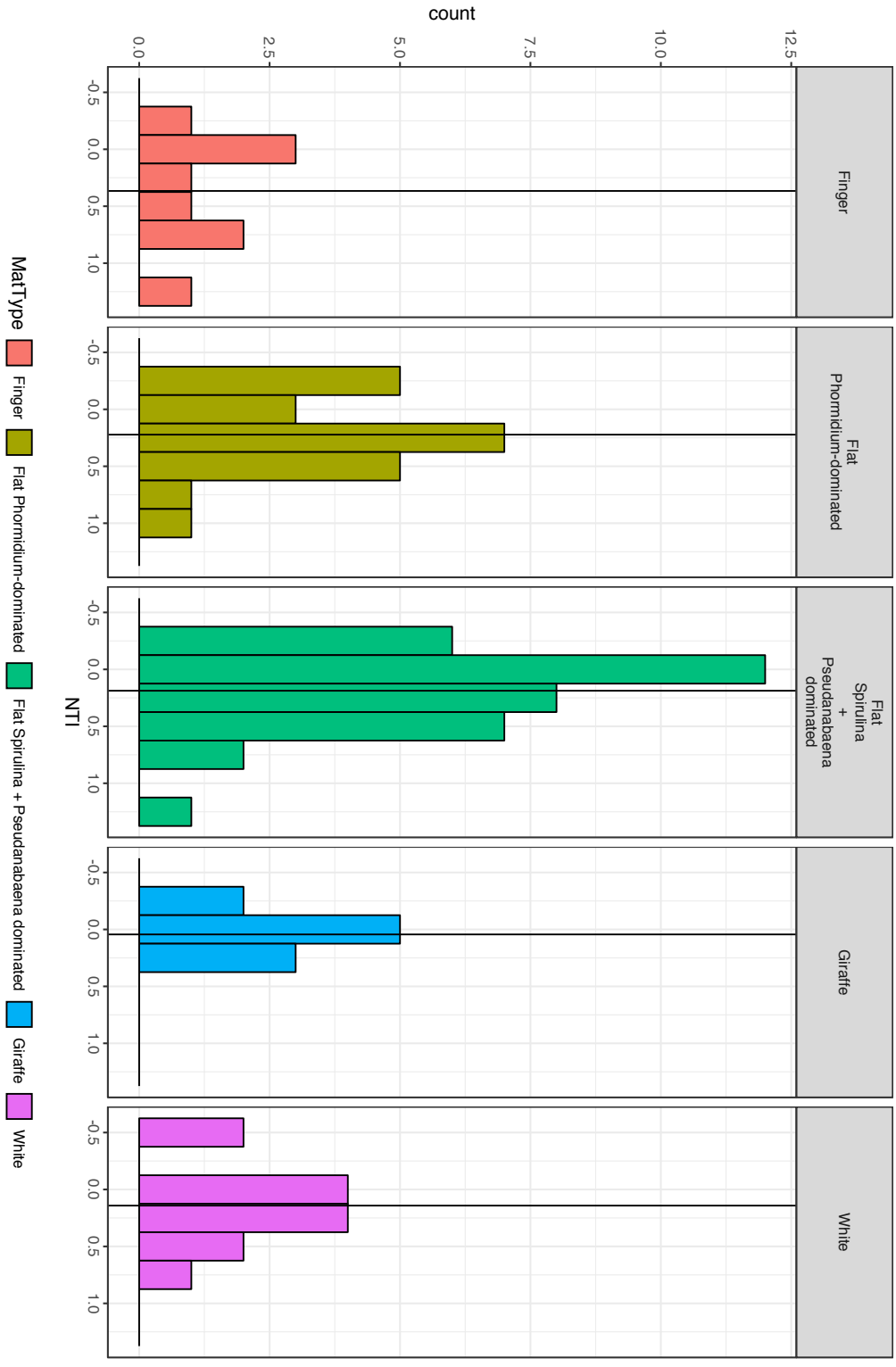


Figure SI 4.5 Phylogeny of sqr genes evaluated from MAGs. They are classified according to (Marcia et al. 2010).

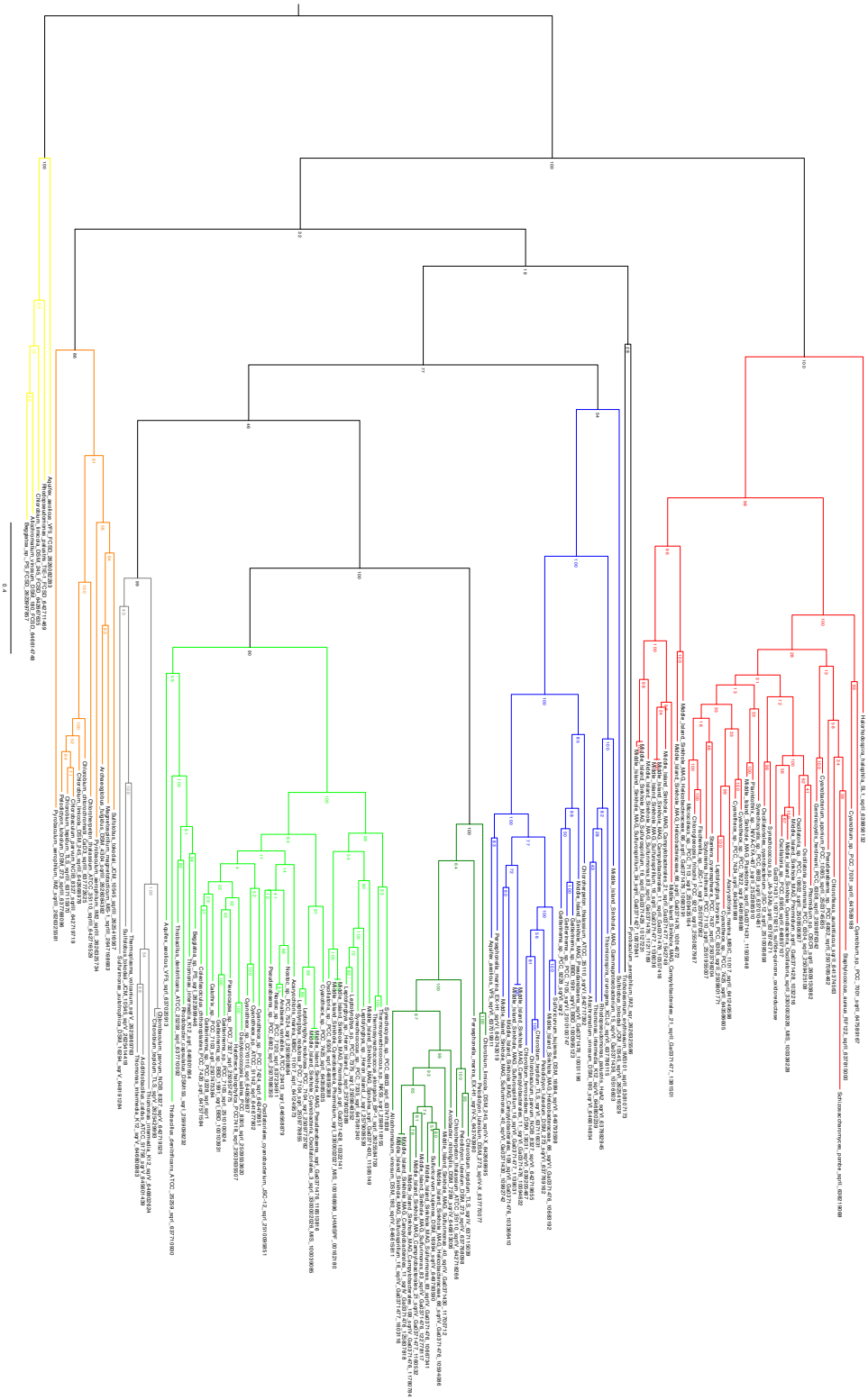
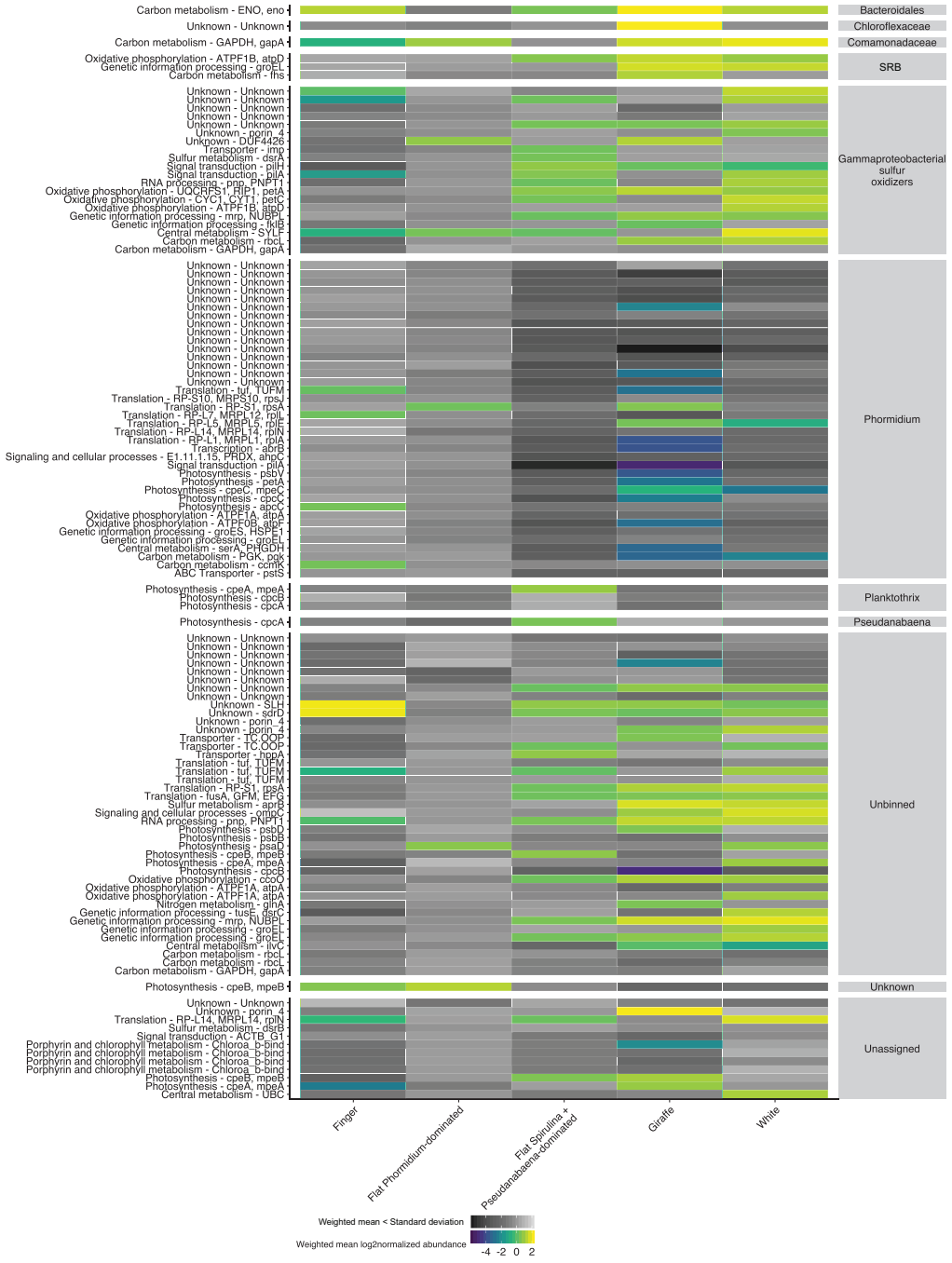


Figure SI 4.6 Heatmap of significantly differentially abundant proteins in mat morphotypes grouped by taxa/bin and metabolic function. Proteins with higher weighted mean log2-normalized abundances are plotted in warm/light grey colors, whereas proteins with lower abundances are plotted in cool/dark grey colors. Proteins whose weighted means are larger than their standard deviations are plotted in the color gradient, whereas proteins that are variably represented in samples of the same mat type, and thus have weighted means are less than their standard deviations, are shown in the grey gradient.



Chapter V: Conclusions and Future Directions

5.1 Introduction

In this chapter, I summarize the major conclusions of the dissertation chapters, and outline potential future directions of research. My dissertation leveraged 'omics approaches to answer questions about the genetics and ecology of anoxygenic photosynthetic cyanobacteria in culture and in natural communities: What is the genomic foundation for metabolic flexibility in an anoxygenic photosynthetic cyanobacterium in a variably sulfidic, low-oxygen environment? What is the impact of seasonality on the community structure and functioning of a natural AP cyanobacterial mat? What are the relationships between microbial community functioning, geochemical cycling, and the morphological appearance of microbial mats?

To answer these questions, **Chapter II** described the genome of a model organism for studying anoxygenic photosynthesis, *Geitlerinema* sp. PCC 9228, with a focus on genes involved in AP, OP, and nitrogen fixation. *Geitlerinema* was the first described culture of an AP cyanobacterium, but its genome was not sequenced and described until this research. This was the first study to identify genes indicative of a low-O_2 lifestyle previously described only in non-cyanobacteria. **Chapter III** evaluated the shift in physicochemical environment and community structure and functioning of the Middle Island Sinkhole microbial mat over multiple years and seasons. This study revealed that there are substantial seasonal shifts in the abundance and functioning of sulfur cycling microbes in the mat, including putative AP cyanobacteria, sulfate-reducing deltaproteobacteria, and sulfide-oxidizing gammaproteobacteria, and provided evidence

that these shifts are linked to light and groundwater chemistry. **Chapter IV** described the metagenomes and metaproteomes of distinct mat morphological types observed in Middle Island Sinkhole. This study relates the functioning of key sulfur cycling bacteria to the morphotypes of mats that appear similar to those observed in the fossil record.

5.2 Metabolic genes and pathways for life at low-O_2 in anoxygenic photosynthetic cyanobacteria

A crucial advancement from this dissertation is the establishment of the genetic repertoire for AP cyanobacteria to thrive in a variable sulfidic, low-O_2 habitat (Grim & Dick, 2016). Genome sequences provide a molecular record to reconstruct the evolutionary history of cyanobacterial AP through Earth history (Dick et al., 2018; Soo et al., 2017). Previous physiological studies determined the sulfide-induced changes in photosynthetic mode, nitrogen metabolism, and sulfide metabolism in *Geilterinema* (Belkin & Padan, 1978; Cohen et al., 1986). In **Chapter II**, I phylogenetically described the key genes involved in *Geitlerinema*'s photosynthetic modes. *Geitlerinema* is not unusual in that it has multiple *psbA* genes for the D1 protein, the water-oxidation site of photosystem II: one copy of an anaerobic group 2 *psbA*, one copy of microoxic group 3 *psbA*, and two copies of a standard oxic/high-light *psbA*. However, my contextualization of *psbA* genes in terms of phylogeny and genomic neighborhood identified a potential relationship between the use of variant *psbA* genes in AP. Select other cyanobacterial genomes with *sqr* genes had groups 2 and 3 *psbA* genes in close chromosomal proximity to their *sqr*, including two other *Geitlerinema* species (Den Uyl et al., 2016) and a metagenome-assembled-genomic contig from Middle Island Sinkhole metagenomes (Voorhies et al., 2012). Due to the draft status of its genome and the difficulty in bridging genomic contigs, it is unknown if *Geitlerinema* sp. PCC 9228 has collocated *psbA* and *sqr* genes. However, this study

presents the possibility of alternative D1 proteins being involved in AP, such as to inactive PSII or to participate in electron transfer from donors other than water, *e.g.*, sulfide (Becraft et al., 2015; Murray, 2012; Olsen et al., 2015). Testing this potential concerted metabolic reworking between OP and AP, using e.g. transcriptomics and proteomics, is an intriguing hypothesis for future studies with *Geitlerinema* and other AP cultures.

In addition to its well-described high affinity sulfide-quinone reductase (SQR) (Arieli et al., 1994; Bronstein et al., 2000; Grim & Dick, 2016; Marcia et al., 2010; Pham et al., 2008), *Geitlerinema* also has a second putative sulfide-donation site on the immediate donor side of PSI that responds to high (~mM) H_2S concentrations (Arieli et al., 1991; Shahak et al., 1987), but the role of this metabolism and the genes involved (including whether it is SQR) have not been described. In **Chapter II**, I identified multiple *sqr* versions in *Geitlerinema* and other cyanobacterial genomes. *Geitlerinema* holds a type I cyanobacterial *sqr* with a high H_2S affinity (K_m = 44 μM) (Shahak et al., 1987), as well as a type VI *sqr* most similar to green sulfur bacterial *sqr* transcribed at mM H_2S concentrations (Chan et al., 2009). Given that so few cyanobacteria possess it, the role of type VI *sqr* in cyanobacterial sulfur metabolism is not well described. In *Leptolyngbya* sp. strain hensonii, type VI *sqr* is likely involved in AP and has an apparent H_2S affinity of K_m = 0.05 to 0.2 mM (Hamilton et al., 2018). The elevated K_m of *Leptolyngbya*'s *sqr* type VI compared to type I *sqr* hints at the use of *Geitlerinema*'s type VI *sqr* at elevated H_2S levels. In its natural environment, *Geitlerinema* thrives in the hypolimnion of Solar Lake, with low-light mM H_2S. In those conditions, light may be the limiting factor for AP growth rather than H_2S availability. However, when mixed in the water column and exposed to different light and sulfide levels, *Geitlerinema* may employ its variable *sqr* and/or *psbA* genes to maintain homeostatic photosynthetic growth. This additional *sqr* is an attractive target for future

physiological studies to elucidate the shifts in AP cyanobacterial metabolism with shifting redox conditions.

Geitlerinema's ability to fix nitrogen under AP has been described (Belkin & Padan, 1978; Belkin et al., 1982), but the combination of AP and diazotrophy has not been further explored, especially at a genetic level, in cyanobacteria. Chapter II identified unexpected nitrogen regulation genes and potential metabolic strategies in AP cyanobacteria. *Geitlerinema* and another AP cyanobacterium, *Coleofaciscus chthonoplastes*, have *nifI*$_1$*I*$_2$ nitrogenase regulatory genes that have been previously observed only in diazotrophic archaea and anaerobic bacteria that regulate their nitrogenase activity post-translation (Boyd et al., 2015; Forchhammer, 2004). Additionally, the proximity of *Geitlerinema*'s group 2 *psbA* gene to *ntcA*, a transcriptional regulator of nitrogen metabolism genes (Forchhammer, 2004), presents an unexplored interface between carbon fixation and nitrogen acquisition under O_2-limiting conditions. Very little is known about genetic regulation of cyanobacterial nitrogen acquisition under sulfidic conditions. *Pseudanabaena* FS39, a strain isolated from sulfidic Frasassi Springs, assimilates nitrate when conducting AP but cannot grow without provided ammonium or nitrate (Klatt, Al-Najjar, et al., 2015). Previous physiological studies of AP growth in *Geitlerinema, C. chthonoplastes,* and *Leptolyngbya* sp. strain hensonii have occurred in nitrogen-replete media, suggesting nitrate assimilation occurs under AP conditions (Cohen et al., 1986; Hamilton et al., 2018). In *Geitlerinema*, translation of its group 2 *psbA* under anaerobic conditions may prevent O_2 production and allow for nitrogen fixation, which in turn may necessitate regulation of nitrogen metabolic genes via *ntcA*. These findings motivate future investigations into carbon, nitrogen, and sulfur metabolic fine-tuning in AP cyanobacteria.

5.3 Seasonal dynamics in the functioning and abundance of sulfur-cycling microbial mat

Genomic analyses of the dominant *Phormidium* and *Planktothrix* members in the low-O_2 Middle Island Sinkhole microbial mat indicated the potential for them to perform the observed anoxygenic photosynthesis in the mat (Voorhies et al., 2012). *Phormidium* possesses a type I, high H_2S affinity *sqr* which is involved in AP in cyanobacteria, whereas *Planktothrix* encodes for a type II *sqr* typically employed for sulfide detoxification (Grim & Dick, 2016). *Phormidium* and *Planktothrix* have groups 3 and 4 *psbA* genes for microaerobic/dynamic redox conditions as well as standard growth, but *Phormidium* additionally has a group 2 *psbA* for anaerobic processes (Grim & Dick, 2016). A cyanobacterial *nifHDK* nitrogenase gene suite was also observed in the metagenomes and new binning efforts assign it to *Phormidium*, suggesting that it may fix nitrogen to alleviate nitrogen stress that may occur due to nitrogen loss during denitrification, which is active in the mat (Voorhies, 2014). During anoxic conditions or to prevent O_2 production, *e.g.*, for the purpose of facilitating nitrogen fixation, *Phormidium* may utilize group 2 *psbA* to disable water oxidation in photosystem II, a mechanism not likely available to *Planktothrix* due to the absence of this variant in the genome.

The combination of these gene suites suggests a greater degree of metabolic versatility through varying O_2, H_2S, and nitrogen stress conditions in *Phormidium* compared to *Planktothrix*. Though *sqr* in *Phormidium* has not been confirmed to participate in AP, its phylogenetic grouping with known-AP cyanobacterial *sqr* points to *Phormidium* being a key AP organism. In contrast, *Planktothrix* and its type II *sqr* may survive sulfidic conditions through sulfide detoxification or sulfide oxidation without phototrophic growth.

An important conclusion from my dissertation is that the seasonally-changing physicochemical environment affects the functioning and abundance of sulfur cycling microbes in a natural community. In **Chapter III**, 16S rRNA gene level analyses revealed temporal

patterns in bacterial populations in MIS microbial mats. Cyanobacteria, dominated by putative AP members *Phormidium* and sulfide-tolerant *Planktothrix*, were most abundant in early and late summer months. Sulfate-reducing bacteria were most abundant in late summer, and in autumn the bacterial community was dominated by sulfide-oxidizing bacteria, primarily *Beggiatoa*. Additionally, *Planktothrix* was the more abundant cyanobacterial member in autumn. This shift in the community structure is most likely due to changing cyanobacterial growth conditions: when light levels were highest, and when the water chemistry was most enriched in high-sulfate groundwater, *Phormidium* cyanobacteria were more abundant.

Chapter III presents the first description of proteomic signatures of discrete populations of AP cyanobacteria, sulfide-oxidizing bacteria, and sulfate-reducing bacteria in a low-oxygen, sulfidic cyanobacterial mat. The linkage of these community dynamics to light and/or geochemical conditions has implications for the balance between different cyanobacteria with distinct functional niches, the functioning of photosynthetic modes in AP cyanobacteria, and shifting redox environment. The proteomic and 16S rRNA gene data in **Chapter III** corroborates the proposed metabolic flexibility of *Phormidium* compared to *Planktothrix*, as seen in their genetic content. Of the cyanobacteria, it is the most abundant in 16S rRNA gene and proteomic profiles, especially in summer months. The higher significant abundance of *Phormidium* allophycocyanin (a key light harvesting protein), thioredoxin, and chaperone proteins in summer months correlates with better growth conditions, namely high light levels. These observations in a natural community are consistent with results from cyanobacterial cultures that growth rates and abundance of ribosomal proteins are positively related (Jahn et al., 2018).

This study also revealed the potential for cyanobacterial niche adaptation in MIS mat, an ecological strategy not previously considered in this low-diversity system. The improved

metagenomic binning and proteomic surveys in **Chapter III** showed that not only were phycobiliproteins the most abundant protein in the system, but that *Phormidium* and many different cyanobacteria have different types of phycobiliproteins that are differentially abundant throughout the seasons. *Planktothrix* and other cyanobacterial photosynthetic proteins were significantly higher in summer, not lower as one might expect should they be directly competing with *Phormidium* for optimal light conditions. Additionally, the multiple phycocyanin and phycoerythrin proteins that were significantly variably abundant in the metagenome-assembled-genome of *Phormidium* potentially reflect multiple strains and/or the ability for *Phormidium* to conduct complementary chromatic adaptation (CCA). During CCA, cyanobacteria that encode for phycoerythrin are capable of modulating synthesis of PE alone or phycoerythrin along with phycocyanin (Bryant, 1982). Cyanobacterial cultures have furthered our understanding of the regulation and environmental conditions behind CCA (Bezy et al., 2011; Bryant, 1982; Gan et al., 2014), and in natural mat communities, different cyanobacteria position themselves in the mat for optimal light conditions (Hawes & Schwarz, 2001; Oberhaus et al., 2007; Sumner et al., 2015). Though the MIS cyanobacteria remain unculturable, laboratory manipulations of light availability on microbial mat could untangle possible light adaptation, including CCA, in the dominant cyanobacteria. Our research positions MIS as an intriguing natural setting in which to explore niche complementarity, light adaptation, and the interspecific interaction of metabolically flexible cyanobacteria potentially capable of CCA.

While we were unable to extract proteomic evidence of AP, several of the significantly abundant *Phormidium* proteins are related to both modes of AP and OP (light harvesting and carbon fixation), as well as protein synthesis and growth. Thioredoxin, a key protein involved in the Calvin cycle and oxidative pentose phosphate cycle (Blankenship, 2014), and pigment

protein allophycocyanin are required for both photosynthetic modes to operate, and were most abundant in summer. The shift in *Phormidium* proteomic abundance between summer and autumn is tied to less favorable growth conditions in autumn, which manifests in the 16S rRNA gene relative abundance data. In autumn when light levels are lower, the relative abundances of *Phormidium* proteins and marker genes are lower than those of other mat members. At that time, while *Planktothrix* is relatively more abundant in the mat, it is likely due to a lower relative abundance of *Phormidium* rather than better growth conditions for *Planktothrix*, given that the dominant autotroph proteomic signatures in autumn are *Beggiatoa* sulfide-oxidizing bacteria and diatoms.

In MIS sulfide-oxidizing bacteria present an additional biological intersection for the sulfur and oxygen cycles, in that they rely upon O_2 from the photosynthesizers or overlying lake water to oxidize H_2S. *Beggiatoa* are pivotal members of chemosynthetic communities that produce organic carbon, fix nitrogen, and remove sulfide from the environment (Flood et al., 2014). Likely when light conditions are poor for cyanobacteria, such as during low light levels at the beginning and end of the day, in autumn and by extension winter, *Beggiatoa* in MIS may perform the role of AP cyanobacteria in depleting sulfide.

Diatoms are present in the mat surface throughout the year (Nold, Zajack, et al., 2010), but the impacts of diatoms in the MIS community functioning and O_2 budget were previously unexplored. The proteomics data in Chapter III highlight the diatoms as integral primary producers, with photosynthetic and carbon fixation proteins especially abundant in autumn. The proteomics data also provided evidence for an active sulfate-reducing bacterial community in the cyanobacterial mat. O_2 tolerance is an essential trait for such microbes that thrive among AP/OP cyanobacteria and diatoms (Canfield & Marais, 1991). Further, active sulfate reduction within

the mat is potentially a local and important source of sulfide to sulfide-oxidizing members, as opposed to sulfide diffusing from sediment sulfate reduction. These results reveal how the interplay of metabolic activities of specific populations shapes community function. The proteomics datasets also provide a rich framework for studying functional mechanisms by which mat-forming cyanobacteria in extreme environments deal with changing light levels and geochemistry.

5.4 Functional redundancy within cyanobacteria, sulfate-reducing bacteria, and sulfide-oxidizing bacteria across different mat morphotypes

Microbial mats are present in the geologic record for over 3.5Ga in a variety of forms in stromatolites, microbialites, and microbially induced sedimentary structures (Noffke et al., 2013; Nutman et al., 2016; Tice & Lowe, 2004). Studies of modern mats indicate that different mechanisms of formation and microbial metabolisms may influence their appearance. Physical processes such as tides and flooding, and metabolisms that precipitate minerals can determine biofilm thickness (Stal, 2012). Additionally, gas-producing metabolisms such as oxygenic photosynthesis can uplift mats from underlying sediment (Bosak et al., 2009; Noffke, 2010). However, distinct microbial mat types are rarely observed at the same time, in the same place under homogenous environmental conditions. **Chapter IV** presents a concerted metagenomic, metaproteomic, and geochemical approach to target the relationship between geochemical cycling, microbial functioning, and the morphology of coexisting microbial mats that appear similar to those that appear in the geologic record. My multifaceted characterization of four geologically-relevant mat morphotypes (finger, giraffe, flat purple, and flat white mats) observed in a short spatiotemporal frame in Middle Island Sinkhole, revealed unexpected microbial and functional diversity in these microbial mats. The subtle differences in microbial functioning and

the subsequent influences on their interactions with geochemical cycling may play a role in the distinct appearances of the mats.

A key advancement from **Chapter IV** is the expansion of our knowledge of the metabolic capacity of cyanobacterial mats in low-O_2 environments. My investigations documented a core shift in the cyanobacterial community of MIS from 2015-2017. Previously, the genome of a dominant strain of *Phormidium*, and partial genomes of a less abundant strain of *Phormidium* as well as *Planktothrix* and other oscillatorial members, were described in MIS fingers and mats (Voorhies et al., 2012; 2016). In the current study, the contemporary MIS mats are dominated by new strain(s) of *Phormidium*, *Planktothrix*, and two previously rare cyanobacteria, *Pseudanabaena* and *Spirulina*. 16S rRNA gene surveys indicated that either *Phormidium* or the pair of *Pseudanabaena* and *Spirulina* were dominant in the mats, but not both at the same time. Further coupled 'omics and geochemical measurements on mats characterized by the dominance of these different cyanobacteria may uncover the impact of this shifting phototrophic community on sulfide and oxygen cycling in MIS and other analog systems. Extreme environments tend to restrict diversity in microbial mats (Stal, 2012), and previous surveys have determined MIS mats to have relatively low diversity (Kinsman-Costello et al., 2017; Nold, Pangborn, et al., 2010; Voorhies et al., 2012). Interspecific competition occurs between phylogenetically related organisms that perform the same metabolism and prefer the same habitat niche (Stegen et al., 2012; Violle et al., 2011). Results from comparative genomics suggest that *Phormidium* and *Pseudanabaena* may occupy similar functional niches. They both have cyanobacterial *sqr* genes implicated in anoxygenic photosynthesis, variant *psbA* genes that may be synthesized at different O_2 levels, and *nifHDK* nitrogenase gene suites for fixing nitrogen. A crucial difference between the two genomes is that *Pseudanabaena* lacks

phycoerythrin to absorb green light, and has a substantially smaller genome than *Phormidium.* These differences are also paralleled in the genome comparison of *Spirulina* and *Planktothrix*, with *Spirulina* having a smaller genome and synthesizing phycoerythrocyanin instead of phycoerythrin. A smaller genome is a strategy observed in genome streamlining, and improves competitiveness of organisms due to a reduced requirement for resources for replication (Giovannoni et al., 2014). While *Pseudanabaena* and *Spirulina* may have similar niches to *Phormidium* and *Planktothrix*, their smaller genomes may impart an energetic advantage that manifests in the abundance data. Thus with environmental conditions putatively limiting available niche space in this low-diversity system, the emergence of these new cyanobacterial members, particularly of *Pseudanabaena* and *Spirulina,* begs further investigation into the competition and/or ecological succession that occurred in the MIS cyanobacterial mat.

Chapter IV also presents novel insights into the metabolisms and diversity of sulfide-oxidizing epsilonproteobacteria and gammaproteobacteria in the microbial mats. *Beggiatoa* and other gammaproteobacteria have been implicated in sulfur cycling in MIS and related mats, (Kinsman-Costello et al., 2017; Nold, Pangborn, et al., 2010; Sharrar et al., 2017). However, epsilonproteobacteria are relatively rare in these mats, thus their genome sequences and metabolic potential are largely unknown. Metagenomes of MIS mat morphotypes yielded two abundant metagenome-assembled-genome bins (MAGs) of gammaproteobacterial sulfide oxidizers, and eight MAGs of epsilonproteobacterial sulfide oxidizers. The dominant gammaproteobacterium in giraffe mat, Gamma_bin_8, possesses the bidirectional hydrogenase analogous to cyanobacteria, and reverse dissimilatory sulfite reductase pathway and related genes, whereas the gammaproteobacterium sourced from diverse mat types, Gamma_bin_15 may perform sulfide oxidation and sulfur disproportionation along with hydrogen oxidation.

All the epsilonproteobacterial MAGs encode 1-3 *sqr* genes, and at least one cytochrome and hydrogenase, but vary in their putative affinities for H_2S and O_2, and ability to metabolize H_2. Two finger-endemic *Sulfurospirillum* MAGs have one *sqr* each, and are equipped for denitrification and dissimilatory nitrate reduction, aerobic respiration at both high and low [O_2], and H_2 production. Epsilonproteobacteria primarily observed in white and flat-*Spirulina* and *Pseudanabaena* mat have 2-3 *sqr* genes potentially expressed for different H_2S conditions and/or metal detoxification, 1-2 hydrogenases, typically one cytochrome, and pathways for denitrification, nitrogen fixation, or nitrate reduction. In other systems with gammaproteobacteria and epsilonproteobacteria, the availability of O_2 (Macalady et al., 2006) and the stage of ecological succession in the mat (Patwardhan et al., 2018) were related to which sulfide-oxidizing group was dominant. Higher O_2 and more established mats tend to have higher relative abundances of gammaproteobacteria than of epsilonproteobacteria. These sulfide oxidizers appear to differ in their ability to metabolize O_2, H_2, and forms of sulfur and nitrogen, which may be related to their distribution and abundance in different mat types. While it is unclear if the age of the observed MIS mats varied, geochemical measurements indicated photosynthetic O_2 production was nonexistent in white mat, and of similar rates in the other mats. This research invites follow up into the metabolic importance of epsilonproteobacteria, and their ecological interactions with gammaproteobacteria and other microbes, in MIS and other phototrophic mats.

Our investigations lead to the hypothesis that the shift in dominant cyanobacteria, sulfide oxidizers, and sulfate reducers in the microbial mats is related to differences in biogeochemical cycling. In addition to the geochemical environments varying by mat type, key proteins varied significantly in abundance. Measurements indicate that mat-derived O_2 flux and depth-integrated

oxygenic photosynthetic rates are comparable between all mats except white mat. Proteins belonging to *Phormidium* were significantly higher in abundance in finger and select flat mats in which *Phormidium* and *Planktothrix* were most abundant, and in lower abundance in giraffe, white and other mats in which *Pseudanabaena* and *Spirulina* were abundant. In contrast, other cyanobacterial proteins were more abundant in giraffe, white, and *Spirulina* and *Pseudanabaena* dominated flat mats. Because these proteins were related to photosynthesis and/or basic metabolic functions, these proteomes provide evidence for more favorable growth conditions for *Phormidium* in fingers and select flat mats, and better growth conditions for other cyanobacteria in giraffe, white and other flat mats. The functional similarity between *Phormidium/Pseudanabaena* and *Planktothrix/Spirulina*, as seen in their gene content, may help explain similar geochemical measurements in cyanobacterially-dominated mats.

This research highlights the importance of sulfate-reducing bacteria in the mat locally providing sulfide to the community. The two different sulfate-reducing MAGs may have different O_2 tolerances, based on their cytochromes: the *Desulfonema* primarily found in giraffe and purple mat dominated by *Spirulina* and *Pseudanabaena* has both high- and low-O_2 affinity cytochromes, whereas the Desulfobacteraceae MAG found throughout many mat types has a high-O_2 affinity cytochrome. Proteins belonging to sulfate-reducing bacteria were also more abundant in white mat, and cyanobacterial mat in which *Spirulina* and *Pseudanabaena* were dominant, than in fingers and flat-*Phormidium* dominated mat. Total sulfide (S_{tot}) is less available in fingers and more available in *ex situ* flat purple mat, but these values in flat and giraffe mats were not wildly variable, and their similar ranges to O_2 flux from the mat indicate near complete aerobic consumption with photosynthetic O_2. Laboratory experiments measured AP in fingers and flat-*Phormidium* dominated mat, though we were unable to confirm AP *in situ*

in flat and giraffe mat through microsensors or proteomics. But just like *Phormidium*, *Pseudanabaena* and potentially *Spirulina* are equipped for AP. The low S_{tot} in fingers compared to the sediment-attached mats, the similar magnitude of S_{tot} and O_2 fluxes in flat and giraffe mats, and the documented AP in mats, reflects a disconnect between sulfide sources and sinks: if all the measured sulfide can be consumed by gamma- and epsilonproteobacterial aerobic sulfide oxidation, but AP is occurring, then another sulfide source must be available. Thus cyanobacterial AP but may be partially driven from cryptic sulfide sources, such as sulfur disproportionation in the capable sulfide-oxidizing MAG members, or likely more significantly, localized sulfate reduction from the deltaproteobacterial MAG members.

Geochemical and molecular characterization of white mat in **Chapter IV** sheds new light into the role of chemotrophs in forming microbial mats. Generally, white mat has little to no O_2 and a wide S_{tot} range, and hosts primarily sulfide-oxidizing epsilonproteobacteria and the more versatile Gamma_bin_15. Many proteins belonging to gamamproteobacterial sulfide oxidizers were significantly more abundant in giraffe, white, and *Spirulina* and *Pseudanabaena*-dominant mats, suggesting favorable growth conditions in these mat types compared to finger and *Phormidium*-dominated mats. From their genomes, these sulfur-cycling bacteria are equipped to handle and variable H_2S and O_2, and may rely upon water-column derived O_2 more so than photosynthetically-sourced O_2, or other electron acceptors such as nitrogen or sulfur species. Because sulfide oxidizers and cyanobacteria can generate similar mat features (Flood et al., 2014), this research presents a counterpoint to the importance of cyanobacteria in generating mat structures similar to those observed in the geologic record

New metagenomes coupled with quantitative metaproteomes of distinct mat morphotypes enabled unprecedented insight into the potential and actual functioning of these mats. While the

MAGs suggested key mat-specific genetic differences in cyanobacterial AP and OP, levels of aerotolerance in sulfate reducers, and different strategies of sulfur metabolism in sulfide-oxidizing bacteria, we observed nearly 1/3 of proteins were shared across multiple mat types and invariable in abundance. Many of these identifiable proteins belonged to *Phormidium* and were involved in photosynthesis, but just as many were sourced from Gammaproteobacteria, or were involved in carbon metabolism. Additionally, sulfur metabolism proteins such as subunits of dissimilatory sulfite reductase (in sulfate-reducing bacteria) and rDsr (in sulfide-oxidizing bacteria) were widely observed in the mats. This core proteome confirms that despite their morphological differences, these mats are hotspots of photosynthesis, sulfate reduction, and sulfide oxidation, and the identity (and thus functional potential) of key phototroph(s) and sulfide oxidizer(s) may subtly influence biogeochemistry.

5.5 Synthesis

For over 40 years, our understanding of anoxygenic photosynthetic cyanobacteria has been limited to physiological and molecular studies in cultured representatives. Thanks to methodological and technological improvements in molecular and geochemical techniques in the last ten years, our understanding of the ecology of AP cyanobacteria in natural assemblages has begun to improve (de Beer et al., 2017; Hamilton et al., 2018; Klatt, de Beer, et al., 2016; Klatt, Meyer, et al., 2016). The research in this dissertation combines 'omics methods with geochemical measurements to describe the genetic signature of low-O_2 lifestyles in AP cyanobacteria (**Chapter II**), characterize the role of seasonality in the functioning and community composition of low-O_2 cyanobacterial mats (**Chapter III**), and headlined an investigation of the relationship between microbial mat appearance, metabolisms, and

geochemistry (**Chapter IV**). Phylogenetic contextualization of metabolic genes in oxygenic and anoxygenic photosynthesis, and nitrogen and sulfur metabolism in *Geitlerinema* sp. PCC 9228 in **Chapter II** provided a foundation for analyzing the protein signatures of metagenome-assembled-genomic bins (MAGs) from seasonally-influenced mats in **Chapter III**, as well as the differential cyanobacterial MAGs and their protein abundances in distinct microbial mats described in **Chapter IV**. Seasonally changing light and geochemistry cascades into the protein signatures of cyanobacteria, sulfate-reducing deltaproteobacteria, and sulfide-oxidizing gammaproteobacteria. **Chapter III** highlights the need to explore cyanobacterial mats on temporal scales longer than diel cycles, such as seasonal or yearly cycles. Microbial mats of distinct morphological appearances are observed throughout Earth history in coastal ecosystems, but the ecological factors responsible for mat morphotypes in early-Earth analog systems are little known. **Chapter IV** presents a multifaceted metagenomics, metaproteomics, and geochemical investigation to explore the interaction of microbial metabolism, geochemistry, and community composition in microbial mat morphotypes that have been observed in the geologic record. These results invite additional investigation into the ecology of sulfur cycling members—namely anoxygenic photosynthetic cyanobacteria, sulfate-reducing bacteria, and sulfide-oxidizing bacteria—in microbial mats, past and present.

5.6 References

Arieli, B., Padan, E., & Shahak, Y. (1991). Sulfide-induced Sulfide-Quinone Reductase Activity in Thylakoids of Oscillatoria limnetica. Journal of Biological Chemistry, 266(1), 104–111.

Arieli, B., Shahak, Y., Taglicht, D., Hauska, G., & Padan, E. (1994). Purification and characterization of sulfide-quinone reductase, a novel enzyme driving anoxygenic photosynthesis in Oscillatoria limnetica. Journal of Biological Chemistry, 269(8), 5705–5711.

Becraft, E. D., Wood, J. M., Rusch, D. B., Kühl, M., Jensen, S. I., Bryant, D. A., et al. (2015). The molecular dimension of microbial species: 1. Ecological distinctions among, and homogeneity within, putative ecotypes of Synechococcus inhabiting the cyanobacterial mat of Mushroom Spring, Yellowstone National Park. Frontiers in Microbiology, 6. https://doi.org/10.3389/fmicb.2015.00590

Belkin, S., & Padan, E. (1978). Sulfide-dependent hydrogen evolution in the cyanobacterium Oscillatoria limnetica. FEBS Letters, 94(2), 291–294. https://doi.org/10.1016/0014-5793(78)80959-4

Belkin, S., Arieli, B., & Padan, E. (1982). Sulfide dependent electron transport in Oscillatoria limnetica. Israel Journal of Botany. https://doi.org/10.1080/0021213X.1982.10676943

Bezy, R. P., Wiltbank, L., & Kehoe, D. M. (2011). Light-dependent attenuation of phycoerythrin geneexpression reveals convergent evolution of greenlight sensing in cyanobacteria. Proceedings of the National Academy of Sciences, 108(45), 18542–18547. https://doi.org/10.1073/pnas.1107427108/-/DCSupplemental

Blankenship, R. E. (2014). Molecular mechanisms of photosynthesis (1st ed.). Blackwell Science Ltd.

Bosak, T., Liang, B., Sim, M. S., & Petroff, A. P. (2009). Morphological record of oxygenic photosynthesis in conical stromatolites. Proceedings of the National Academy of Sciences of the United States of America, 106(27), 10939–10943. https://doi.org/10.1073/pnas.0900885106

Boyd, E. S., Costas, A. M. G., Hamilton, T. L., Mus, F., & Peters, J. W. (2015). Evolution of Molybdenum Nitrogenase during the Transition from Anaerobic to Aerobic Metabolism. Journal of Bacteriology, 197(9), 1690–1699. https://doi.org/10.1128/JB.02611-14

Bronstein, M., Schütz, M., Hauska, G., Padan, E., & Shahak, Y. (2000). Cyanobacterial Sulfide-Quinone Reductase: Cloning and Heterologous Expression. Journal of Bacteriology, 182(12), 3336–3344. https://doi.org/10.1128/JB.182.12.3336-3344.2000

Bryant, D. A. (1982). Phycoerythrocyanin and Phycoerythrin: Properties and Occurrence in Cyanobacteria. Journal of General Microbiology, 128, 835–844.

Canfield, D. E., & Marais, Des, D. J. (1991). Aerobic sulfate reduction in microbial mats. Science, 251, 1471–1473.

Chan, L.-K., Morgan-Kiss, R. M., & Hanson, T. E. (2009). Functional Analysis of Three Sulfide:Quinone Oxidoreductase Homologs in Chlorobaculum tepidum. Journal of Bacteriology, 191(3), 1026–1034. https://doi.org/10.1128/JB.01154-08

Cohen, Y., Jørgensen, B. B., Revsbech, N. P., & Poplawski, R. (1986). Adaptation to Hydrogen Sulfide of Oxygenic and Anoxygenic Photosynthesis among Cyanobacteria. Applied and Environmental Microbiology, 51(2), 398–407.

de Beer, D., Weber, M., Chennu, A., Hamilton, T. L., Lott, C., Macalady, J., & M Klatt, J. (2017). Oxygenic and anoxygenic photosynthesis in a microbial mat from an anoxic and sulfidic spring. Environmental Microbiology, 19(3), 1251–1265. https://doi.org/10.1111/1462-2920.13654

Den Uyl, P. A., Richardson, L. L., Jain, S., & Dick, G. J. (2016). Unraveling the Physiological Roles of the Cyanobacterium Geitlerinema sp. BBD and Other Black Band Disease Community Members through Genomic Analysis of a Mixed Culture. PloS One, 11(6), e0157953. https://doi.org/10.1371/journal.pone.0157953

Dick, G. J., Grim, S. L., & Klatt, J. M. (2018). Controls on O2 Production in Cyanobacterial Mats and Implications for Earth's Oxygenation. Annual Review of Earth and Planetary Sciences, 46(1), 123–147. https://doi.org/10.1146/annurev-earth-082517-010035

Flood, B. E., Bailey, J. V., & Biddle, J. F. (2014). Horizontal gene transfer and the rock record: comparative genomics of phylogenetically distant bacteria that induce wrinkle structure formation in modern sediments. Geobiology, 12(2), 119–132. https://doi.org/10.1111/gbi.12072

Forchhammer, K. (2004). Global carbon/nitrogen control by P IIsignal transduction in cyanobacteria: from signals to targets. FEMS Microbiology Reviews, 28(3), 319–333. https://doi.org/10.1016/j.femsre.2003.11.001

Gan, F., Zhang, S., Rockwell, N. C., Martin, S. S., Lagarias, J. C., & Bryant, D. A. (2014). Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. Science, 345(6202), 1312–1317. https://doi.org/10.1126/science.1256963

Giovannoni, S. J., Thrash, J. C., & Temperton, B. (2014). Implications of streamlining theory for microbial ecology. The ISME Journal, 8(8), 1553–1565. https://doi.org/10.1038/ismej.2014.60

Grim, S. L., & Dick, G. J. (2016). Photosynthetic Versatility in the Genome of Geitlerinema sp. PCC 9228 (Formerly Oscillatoria limnetica “Solar Lake”), a Model Anoxygenic Photosynthetic Cyanobacterium. Frontiers in Microbiology, 7(590), 1144. https://doi.org/10.3389/fmicb.2016.01546

Hamilton, T. L., Klatt, J. M., de Beer, D., & Macalady, J. L. (2018). Cyanobacterial photosynthesis under sulfidic conditions: insights from the isolate Leptolyngbya sp. strain hensonii. The ISME Journal, 12(2), 568–584. https://doi.org/10.1038/ismej.2017.193

Hawes, I., & Schwarz, A. M. J. (2001). Absorption and utilization of irradiance by cyanobacterial mats in two ice-covered antarctic lakes with contrasting light climates. Journal of Phycology, 37(1), 5–15. https://doi.org/10.1046/j.1529-8817.1999.014012005.x

Jahn, M., Vialas, V., Karlsen, J., Maddalo, G., Edfors, F., Forsström, B., et al. (2018). Growth of Cyanobacteria Is Constrained by the Abundance of Light and Carbon Assimilation Proteins. CellReports, 25(2), 478–486.e8. https://doi.org/10.1016/j.celrep.2018.09.040

Kinsman-Costello, L. E., Sheik, C. S., Sheldon, N. D., Allen Burton, G., Costello, D. M., Marcus, D., et al. (2017). Groundwater shapes sediment biogeochemistry and microbial diversity in a submerged Great Lake sinkhole. Geobiology, 15(2), 225–239. https://doi.org/10.1111/gbi.12215

Klatt, J. M., Al-Najjar, M. A. A., Yilmaz, P., Lavik, G., de Beer, D., & Polerecky, L. (2015). Anoxygenic photosynthesis controls oxygenic photosynthesis in a cyanobacterium from a sulfidic spring. Applied and Environmental Microbiology, 81(6), 2025–2031. https://doi.org/10.1128/AEM.03579-14

Klatt, J. M., de Beer, D., Häusler, S., & Polerecky, L. (2016). Cyanobacteria in Sulfidic Spring Microbial Mats Can Perform Oxygenic and Anoxygenic Photosynthesis Simultaneously during an Entire Diurnal Period. Frontiers in Microbiology, 7(116), 440. https://doi.org/10.3389/fmicb.2016.01973

Klatt, J. M., Meyer, S., Häusler, S., Macalady, J. L., de Beer, D., & Polerecky, L. (2016). Structure and function of natural sulphide-oxidizing microbial mats under dynamic input of light and chemical energy. The ISME Journal, 10(4), 921–933. https://doi.org/10.1038/ismej.2015.167

Macalady, J. L., Lyon, E. H., Koffman, B., Albertson, L. K., Meyer, K., Galdenzi, S., & Mariani, S. (2006). Dominant microbial populations in limestone-corroding stream biofilms, Frasassi cave system, Italy. Applied and Environmental Microbiology, 72(8), 5596–5609. https://doi.org/10.1128/AEM.00715-06

Marcia, M., Ermler, U., Peng, G., & Michel, H. (2010). A new structure-based classification of sulfide:quinone oxidoreductases. Proteins: Structure, Function, and Bioinformatics, 78(5), 1073–1083. https://doi.org/10.1002/prot.22665

Murray, J. W. (2012). Sequence variation at the oxygen-evolving centre of photosystem II: a new class of "rogue" cyanobacterial D1 proteins. Photosynthesis Research, 110(3), 177–184. https://doi.org/10.1007/s11120-011-9714-5

Noffke, N. (2010). Geobiology: Microbial mats in sandy deposits from the Archean Era to today (pp. 1–198). Springer Science & Business Media.

Noffke, N., Christian, D., Wacey, D., & Hazen, R. M. (2013). Microbially Induced Sedimentary Structures Recording an Ancient Ecosystem in the ca.3.48 Billion-Year-Old Dresser Formation, Pilbara, Western Australia. Astrobiology, 13(12), 1103–1124. https://doi.org/10.1089/ast.2013.1030

Nold, S. C., Pangborn, J. B., Zajack, H. A., Kendall, S. T., Rediske, R. R., & Biddanda, B. A. (2010). Benthic bacterial diversity in submerged sinkhole ecosystems. Applied and Environmental Microbiology, 76(1), 347–351. https://doi.org/10.1128/AEM.01186-09

Nold, S. C., Zajack, H. A., & Biddanda, B. A. (2010). Eukaryal and archaeal diversity in a submerged sinkhole ecosystem influenced by sulfur-rich, hypoxic groundwater. Journal of Great Lakes Research, 36(2), 366–375. https://doi.org/10.1016/j.jglr.2010.02.014

Nutman, A. P., Bennett, V. C., Friend, C. R. L., Van Kranendonk, M. J., & Chivas, A. R. (2016). Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. Nature, 1–12. https://doi.org/10.1038/nature19355

Oberhaus, L., Briand, J. F., Leboulanger, C., Jacquet, S., & Humbert, J. F. (2007). Comparative effects of the quality and quantity of light and temperature on the growth of Planktothrix agardhii and P. rubescens. Journal of Phycology, 43(6), 1191–1199. https://doi.org/10.1111/j.1529-8817.2007.00414.x

Olsen, M. T., Nowack, S., Wood, J. M., Becraft, E. D., LaButti, K., Lipzen, A., et al. (2015). The molecular dimension of microbial species: 3. Comparative genomics of Synechococcus strains with different light responses and in situ diel transcription patterns of associated putative ecotypes in the Mushroom Spring microbial mat. Frontiers in Microbiology, 6, 604. https://doi.org/10.3389/fmicb.2015.00604

Patwardhan, S., Foustoukos, D. I., Giovannelli, D., Yücel, M., & Vetriani, C. (2018). Ecological Succession of Sulfur-Oxidizing Epsilon- and Gammaproteobacteria During Colonization of a Shallow-Water Gas Vent. Frontiers in Microbiology, 1–16. https://doi.org/10.3389/fmicb.2018.02970

Pham, V. H., Yong, J.-J., Park, S.-J., Yoon, D.-N., Chung, W.-H., & Rhee, S.-K. (2008). Molecular analysis of the diversity of the sulfide : quinone reductase (sqr) gene in sediment environments. Microbiology, 154, 3112–3121. https://doi.org/10.1099/mic.0.2008/018580-0

Shahak, Y., Arieli, B., Binder, B., & Padan, E. (1987). Sulfide-dependent photosynthetic electron flow coupled to proton translocation in thylakoids of the cyanobacterium Oscillatoria limnetica. Archives of Biochemistry and Biophysics, 259(2), 605–615. https://doi.org/10.1016/0003-9861(87)90527-3

Sharrar, A. M., Flood, B. E., Bailey, J. V., Jones, D. S., Biddanda, B. A., Ruberg, S. A., et al. (2017). Novel Large Sulfur Bacteria in the Metagenomes of Groundwater-Fed Chemosynthetic Microbial Mats in the Lake Huron Basin. Frontiers in Microbiology, 8, 2104. https://doi.org/10.3389/fmicb.2017.00791

Soo, R. M., Hemp, J., Parks, D. H., Fischer, W. W., & Hugenholtz, P. (2017). On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. Science, 355(6332), 1436–1440. https://doi.org/10.1126/science.aal3794

Stal, L. J. (2012). Cyanobacterial Mats and Stromatolites. In B. A. Whitton (Ed.), Ecology of Cyanobacteria II (pp. 65–125). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_4

Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities, 6(9), 1653–1664. https://doi.org/10.1038/ismej.2012.22

Sumner, D. Y., Hawes, I., Mackey, T. J., Jungblut, A. D., & Doran, P. T. (2015). Antarctic

microbial mats: A modern analog for Archean lacustrine oxygen oases. Geology, 43(10), 887–890. https://doi.org/10.1130/G36966.1

Tice, M. M., & Lowe, D. R. (2004). Photosynthetic microbial mats in the 3,416-Myr-old ocean. Nature, 431(7008), 549–552. https://doi.org/10.1038/nature02888

Violle, C., Nemergut, D. R., Pu, Z., & Jiang, L. (2011). Phylogenetic limiting similarity and competitive exclusion. Ecology Letters, 14(8), 782–787. https://doi.org/10.1111/j.1461-0248.2011.01644.x

Voorhies, A. A. (2014, January 24). Investigation of microbial interactions and ecosystem dynamics in a low O2 cyanobacterial mat.

Voorhies, A. A., Biddanda, B. A., Kendall, S. T., Jain, S., Marcus, D. N., Nold, S. C., et al. (2012). Cyanobacterial life at low O2: community genomics and function reveal metabolic versatility and extremely low diversity in a Great Lakes sinkhole mat. Geobiology, 10(3), 250–267. https://doi.org/10.1111/j.1472-4669.2012.00322.x