Microbial evolution and ecology in subzero hypersaline environments

Zachary Shane Cooper

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Jody W. Deming, Chair

John A. Baross

Robert M. Morris

Program Authorized to Offer Degree:

Oceanography

University of Washington

Abstract

Microbial evolution and ecology in subzero hypersaline environments

Zachary Shane Cooper

Chair of the Supervisory Committee:
Professor Jody W. Deming
Oceanography

Microbial life, particularly prokaryotic life, is prevalent in the extreme polar settings of cryopeg and sea ice brines. Cryopegs are unfrozen layers, found in Arctic permafrost below the active layer, that are composed of relict marine sediments and their seawater-derived brines. Sea ice brines are the inhabitable liquid fraction of the contemporary ice that forms from seawater annually across the Arctic Ocean and along the Antarctic coastline. These environments each host microbial communities that contend with subzero temperatures and high salt concentrations, which present significant challenges to metabolic reaction rates and cellular stability. A major contrast between these environments is the period of exposure to stressors experienced by the inhabitants. Cryopegs remain geophysically stable under subzero hypersaline conditions over millennia while sea ice is relatively ephemeral with its brines generally lasting less than a year. The evolutionary processes that have allowed organisms to prevail in these extreme environments are deeply intertwined with the ecology of these ecosystems. This dissertation explores this entanglement by quantifying the genetic and genomic patterns of these ecosystems. Chapter 1 introduces the ideas that motivated the hypotheses driving the subsequent work aimed at understanding prokaryotic life in subzero brine environments. Chapter 2 uses environmental amplicon sequencing to explore the biodiversity of microbial communities inhabiting sea ice and cryopeg brines and assesses the observed patterns in the context of the defining environmental

physics and chemistry. Chapter 3 employs cutting-edge metagenomic sequencing technology to explore evolutionary patterns of microdiversity that are prevalent in the reconstructed genomes of bacterial populations *in situ*, and Chapter 4 uses pangenomics to assess the evolutionary history and metabolic capabilities of *Marinobacter* species, the predominant bacteria inhabiting cryopeg brines.

Moving through ecological scales, from communities to populations to species, this dissertation has produced the first characterization of microbial communities inhabiting cryopeg brines using environmental sequencing of 16S rRNA genes, finding them to be characterized by low levels of alpha diversity (dominated by the genus *Marinobacter*) and distinct by beta diversity from sea ice communities despite similarities in saline and temperature extremes. It has uncovered the importance of the mobilome (those genes involved in genome rearrangements) in the evolution of populations under constant extreme stress and quantified low levels of nucleotide diversity that are signatures of clonal expansion and selection for well-adapted genotypes in cryopeg brines. Focusing on the primary bacterial species that has flourished in cryopeg brines, it has illustrated the tremendous genetic diversity within the genus *Marinobacter* and identified evolutionary tactics, particularly horizontal gene transfer, and functional capabilities that have likely allowed these bacteria to prevail under constant extremes. Herein, a thorough exploration of the evolution and ecology of microbial communities inhabiting cryopeg and sea ice brines is contained, and the lessons learned from these studies is illustrated through a tome of discussions on the relevance and importance of life in these environments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

My advisor, Jody Deming, deserves my sincere appreciation for taking a chance by taking me on as a graduate student. Her guidance has led me through many projects and ideas, and she has fortified my passion for science. I have had the incredible opportunity to travel to and work in the Arctic as her graduate student, and I have gained an appreciation for the polar environment that I would have never known otherwise. Our scientific discussions always make me wish for an endless amount of time, funding, and energy to find the answers to questions about everything. I have gained tremendous experiences in the field and in the lab as a doctoral student along with a lifetime love for the Arctic and for the microbial life that inhabits it. My committee members, John Baross, Bob Morris, Alex Gagnon, Drew Gorman-Lewis, and Steven Roberts have all guided me through my graduate career, and I thank them all for their help and guidance through my dissertation.

I have had the immense pleasure of working with fantastic members of the Deming Lab over the years. Shelly Carpenter has been the greatest guide and support I ever could have asked for in the lab and in the field, and I owe her a huge amount of gratitude. Max Showalter, Frida the Gecko, Anders Torstensson, Josephine Rapp, Annie Shoemaker, and Georges Kanaan have all been tremendous lab mates, and I have learned a lot from working with each of them, particularly Max, whose cheer and helpfulness have always made days in the office better. I have had the great fortune and pleasure of joining the lab of Jodi Young for an expedition to Antarctica, and I thank her, Hannah Dawson, and Susan Rundell for making that an excellent experience and for teaching me about the wonders of sea ice algae.

Beyond the lab, my friends and colleagues in Oceanography and Astrobiology have always been a source of fun and support. Thank you, Robert, Sasha, Isaiah, Ann, Claire, David, Hannah, Amy, Anna, Katy, Erik, Ryan, and Sarah, for being wonderful friends and finding fun ways to spend time outside of work either going to trivia nights, bowling, writing bizarre zine articles, or going through tabletop game adventures. My 2016 cohort and the Oceanography graduate student community have been excellent, and I am extremely happy to have gone through graduate school with you all.

I have been very fortunate to have received funding and support from the Vetlesen and Moore Foundations as well as from NASA and NSF. The projects that I have worked on have been impactful on my ideas for the future, and I am grateful for the opportunity to have worked on these scientific endeavors.

I would like to extend thanks to Mike Adams at the University of Georgia for bringing me into the lab that got me started in research as an undergraduate student and for introducing me to the world of extremophiles and encouraging me to pursue a career in science, and I thank Gina Lipscomb for tirelessly guiding me through several semesters of laboratory molecular biology and biochemistry that gave me the foundations I needed to be a successful scientist.

Finally, I'd like to thank my parents for always supporting me and for fostering a love of science in me from a young age, and I thank my partner Karina for her unending love, support, and excitement through moving across the country, raising our pets Toph and Opal, travelling the world, and for helping me pursue my dreams.

I could not have completed this dissertation without the people in my life who have made it happen, and I thank you all.

DEDICATION

For Karina, whose love and support has guided me through all my endeavors into the unknown.

Chapter 1. INTRODUCTION TO BACTERIAL PREVALENCE IN SUBZERO BRINES

1.1 DERIVATION AND EVOLUTION OF SUBZERO BRINE ECOSYSTEMS FROM THE MARINE ENVIRONMENT

At the base of all ecosystems are the physical, geological, and chemical characteristics that define their habitability. To date, Earth is the only known home to life in the Universe, which is broadly considered to be due to the prevalence of liquid water on a global scale (Styczinski *et al.*, in review). In fact, nearly every body of liquid water occurring naturally on Earth is inhabited by life, particularly microbial (primarily referring to prokaryotes throughout this dissertation; Fenchel and Finlay, 2004), with exceptions occurring only at the greatest physicochemical extremes or under biogeographical isolation (Stevenson *et al.*, 2015; Cockell, 2021). The biomolecules that compose living organisms are subject to the laws of physics and chemistry, and their functional rates and stability are constrained by these parameters in relation to their geological setting. The extreme environment is one with physical and chemical conditions that impart limits on stability and reaction rates of biomolecules near the edges of known physicochemical habitable parameter space (Harrison *et al.*, 2013; Lever *et al.*, 2015; Nunn *et al.*, 2015). Such environments occur throughout Earth, particularly in polar environments (Jansson & Taş, 2014; Boetius *et al.*, 2015; Malard & Pearce, 2018).

Multiple extreme conditions often occur simultaneously in nature. One major limiting condition is high salt concentration which imposes osmotic stress on the membranes of cells and poses the risk of functional impairment to proteins and nucleotides (Nunn *et al.*, 2015; Firth *et al.*, 2016). Hypersaline conditions often co-occur with temperatures below the freezing point of

water (≤ 1.9˚C for standard seawater; Cox & Weeks, 1983), and temperature plays a key role in the physical stability and catalytic rate of enzymes in a cell (Price & Sowers, 2004). When seawater freezes, non-water molecules and particles in the solution, including living organisms (e.g., bacteria, algae, and their viruses), are excluded from crystallization and become concentrated in the remaining fluid-filled space (Wells & Deming, 2006; Collins & Deming, 2011; Deming & Collins, 2016). This freezing and partitioning process is a defining feature of the sea ice environment (Golden *et al*., 2007; Petrich & Eicken, 2017), one of the largest and most dynamic ecosystems on Earth. In 2021, sea ice covered 14.7 million km^2 at its maximum extent and 4.7 million km^2 at its minimum (Fetterer *et al*., 2017, data accessed November 2021), with the majority of sea ice melting and reforming on an annual basis. This environment, which serves as a major hub of ecological interactions in polar oceans, is presently undergoing major reduction in volume and extent due to anthropogenically induced warming of the oceans (Boetius *et al*., 2015; Overland *et al*., 2019; IPCC, 2021).

Sea ice is not alone in containing subzero brines. They can also be found in ice-covered lakes (Murray *et al*., 2012), subglacial aquifers (Mikucki *et al*., 2009), and saline terrestrial ponds and springs (Perrault *et al*., 2007; Niederberger *et al*., 2010). Another less-well-studied, cold and saline environment is that of cryopegs, which are layers in permafrost that remain unfrozen at subzero temperatures due to the presence of dissolved solutes despite the perennially subzero setting of the environment (Gilichinsky *et al*., 2003). Cryopegs can be derived from coastal marine sediments that, during a freezing climate, were exposed to the atmosphere after a drop in sea level and became buried within permafrost over geological time (Iwahana *et al*., 2021). This process partitions the inhabitants and solutes of the seawater-saturated marine sediments into a brine fraction determined by temperature-driven dynamics analogous, but with a greater

sediment load, to the formation process of sea ice. Cryopeg brines become isolated from atmospheric and other external influences over time, remaining stable at subzero hypersaline conditions for millennia. Like sea ice, the permafrost environment is also experiencing major thaw due to anthropogenic climate warming, changing rapidly with positive feedbacks that accelerate warming (Jansson & Taş, 2014; Overland *et al*., 2019; IPCC, 2021).

Exploration of the microbial inhabitants of cryopeg brines as a unique feature of permafrost has accelerated in recent years (Gilichinsky *et al*., 2005; Colangelo-Lillis *et al*., 2016; Cooper *et al*., 2019; Zhong *et al*., 2020; Rapp *et al*., 2021). Chapter 2 of this dissertation (Cooper *et al*., 2019) advances this work by providing the first characterization of the microbial communities inhabiting cryopeg brines using environmental DNA sequencing (16S rRNA gene amplicons targeting prokaryotes) to identify the structure and biodiversity of Bacteria and Archaea present *in situ,* in contrast with the better studied sea ice environment. From this community characterization emerged questions on the mechanisms of evolution that would allow bacterial populations to prevail in these extreme environments (Chapter 3) and on the functional capability and characteristics of the remarkably abundant and unique bacterial species of the genus *Marinobacter* found to inhabit the cryopeg brines we sampled (Chapter 4).

1.2 GENOMIC EVOLUTION IN THE FACE OF MULTIPLE EXTREME STRESSORS

Evolution by natural selection is a universal process for life (Darwin, 1859; Elena & Lenski, 2003). It allows for organisms that are most fit in an environment to outcompete others, i.e. for those with suitable adaptations to overcome natural stressors and persist over evolutionary time. Given the nature of competition and survival, the ecological role of an organism is deeply intertwined with its evolutionary trajectory (Schloter *et al*., 2000; Cordero & Polz, 2014;

Brockhurst *et al*., 2019), such that the two concepts must be considered simultaneously to interpret the history of life. The entanglement of ecological competition and evolutionary adaptation is inherent for extreme environments where an organism must simultaneously maintain adaptations to stressors and competitive functionality against other community members (Harrison *et al*., 2013; Li *et al*., 2014).

The process of selection occurs by allowing organisms with a phenotypic advantage to reproduce more efficiently than other members of its community (Schloter *et al*., 2000; Li *et al*., 2014; Moulana *et al*., 2020), an advantage that results in the propagation of the genomic material of the competitive individual through future generations. Selection of the phenotype therefore results in genotypic signatures. Following the central dogma of molecular biology (Jafari *et al*., 2017), genes that are transcribed efficiently and that produce proteins that provide functional advantages are the sources of the selective process. This process is complicated by stochastic error introduction at the single nucleotide level during genome replication, which may change a genotype for better, worse, or to no effect at all (Elena & Lenski, 2003; Schloissnig et al., 2013; Good et al., 2017). For prokaryotes, the mode of gene evolution by replicative errors from mother to daughter cells that results in beneficial phenotypes is known as vertical inheritance. This process is undergone by all living beings, but it can require many generations for a novel phenotype to be introduced, as genes span thousands of nucleotide bases and a single mutation may not change its encoding. The process of vertical evolution can be too slow in some cases for an organism to overcome a newly introduced stressor in its environment, but alternative mechanisms for evolution exist to overcome this challenge.

Horizontal gene transfer (HGT) is the process of the integration and retention of an entire new gene into a genome (Galtier, 2007; Gillings, 2017; Oliveira *et al*., 2017). This evolutionary

mechanism allows for the rapid acquisition of a novel phenotype, overcoming the time constraint intrinsic to vertical inheritance. HGT is widespread within the empire of Prokaryota, which are morphologically simple (mostly lacking intracellular compartmentalization and defined organelles; Harris and Theriot, 2018) but genetically diverse organisms (Schloter *et al*., 2000; Brockhurst *et al*., 2019; Maistrenko *et al*., 2020; Moulana et al., 2020). Bacteria and Archaea, along with their viruses, represent the earliest lineages of life on Earth (Weiss *et al*., 2016). They have spent eons trading genes (Fuchsman *et al*., 2017) and undergoing evolution that has allowed them to occupy every ecological niche available, including in nearly all of those environments with the most extreme physicochemical conditions (Fenchel & Finlay, 2004; Harrison *et al*., 2013; Cockell, 2021). Multiple modes of HGT are used by prokaryotes, including transformation where extracellular DNA is drawn into a cell, transduction where viruses transfer genes from one host to another, and conjugation where cells form physical structures to connect to other cells and transfer genes (Gillings, 2017). Other mechanisms for gene exchange exist (e.g., genetic transfer agents; Lang *et al*., 2012), and new aspects of HGT are still being uncovered. At the genomic level, HGT depends on the utilization of genetic machinery that allows for the mobility of genes. The genes involved in the production and movement of mobile genetic elements are collectively referred to as the mobilome (Toussaint & Chandler, 2012; Oliveira *et al*., 2017). The role of HGT and the mobilome in microbial adaptation to extreme environments, particularly subzero brine environments, is explored further in this dissertation, where the evolutionary history of bacterial populations inhabiting sea ice and cryopeg brines (Chapter 3) and the dominant species of cryopeg brines (Chapter 4) are considered. As evolution is a fundamental aspect of life, the mechanisms that enable it, in any environment, are critical to understanding of the history, extent, and future of life in the Universe.

1.3 TERRESTRIAL AND ASTROBIOLOGICAL IMPLICATIONS OF LIFE IN SUBZERO BRINES

At the crux of this dissertation is the intertwinement of the physical, chemical, and geological setting of the subzero brine environment with the ecology and evolutionary modes of the endemic microbial inhabitants. Chapter 2 considers the physical, chemical, and geological parameters that define and influence the microbial community structure and diversity in sea ice and cryopeg brine environments. The following chapters build upon these analyses to interrogate the evolutionary and metabolic processes that allow microbial life to prevail in these extreme environments, which provide fundamental insight into the role of environmental constraints on microbial adaptation. Earth is a microbial planet, inhabited and numerically dominated by prokaryotes since the earliest geological epochs (Weiss *et al*., 2016). They have controlled the atmospheric composition with their metabolic processes (Ilbert & Bonnefoy, 2013), exerting feedbacks on global temperature and biogeochemical cycling (Maier, 2015). The history, present, and future of Earth are all shaped by microorganisms that will continue to evolve in the face of environmental changes, forging new niches until the end of this Solar System (Beech, 2011), and perhaps beyond.

Outside Earth, life has not yet been identified, though considerable progress on exploring and characterizing habitable environments in the Solar System and beyond is underway (Styczinski *et al*., in review). Considering habitability based on life as we know it, habitable environments within the Solar System are widely characterized by the presence of liquid water, and most of these settings are cold and saline. Our ability to identify life in these environments is facilitated by our ability to understand and identify life in analogous environments on Earth (Gilichinsky *et al*., 2003; Murray *et al*., 2012; Cockell *et al*., 2018). Exploration and discovery have always been

at the frontier of scientific advancement, and the search for life beyond Earth falls fully in line with this perspective. The field of Astrobiology is dedicated to understanding the origin, evolution, and extent of life in the Universe; explorations of the most extreme settings on Earth help to define the boundaries of that extent and provide important insights on the potential of life.

The climate of Earth is changing due to anthropogenic influences, creating dramatic and unprecedented changes to the environment (IPCC, 2021). Human communities are being displaced (Romero Manrique *et al.*, 2018), massive ecosystems are being lost (Canadell & Jackson, 2021), and the entire biosphere of Earth is shifting to an unknown and potentially unstable state (Bonan & Doney, 2018). Microorganisms play primary roles in biogeochemical cycling (Maier, 2015) and in the lives of every other organism through symbiotic interactions (Germerodt *et al.*, 2016; Fricker *et al.*, 2019). They have persisted through numerous faunal mass extinction events (Xie *et al.*, 2005) and will continue to evolve and fill any ecological niches made available by changing environmental conditions (e.g., permafrost thaw resulting in niches for mesophilic soil microbes). The cryosphere, in particular, is undergoing rapid change (Boetius *et al.*, 2015; Fetterer *et al.*, 2017; Overland *et al.*, 2019), diminishing the opportunities to understand the lives of the fascinating communities adapted to live within a frozen environment. It is in the spirit of fundamental understanding of fragile ecosystems and the future of life on Earth and beyond that we explore life at its limits and that I have worked to explore the evolution and ecology of the microbial inhabitants of subzero brines.

References

Beech, M. (2011). The past, present and future supernova threat to Earth's biosphere. *Astrophysics and Space Science 2011 336:2*, *336*(2), 287–302. https://doi.org/10.1007/S10509-011-0873-9

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., & Rapp, J. Z. (2015). Microbial ecology of the cryosphere: Sea ice and glacial habitats. *Nature Reviews Microbiology*, *13*(11), 677–690. https://doi.org/10.1038/nrmicro3522

Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. *Science*. American Association for the Advancement of Science. https://doi.org/10.1126/science.aam8328

Brockhurst, M. A., Harrison, E., Hall, J. P. J., Richards, T., McNally, A., & MacLean, C. (2019). The Ecology and Evolution of Pangenomes. *Current Biology*. 29(20), R1094–R1103. https://doi.org/10.1016/j.cub.2019.08.012

Canadell, J. G., & Jackson, R. B. (2021). *Ecosystem Collapse and Climate Change*. (J. G. Canadell & R. B. Jackson, Eds.) (Vol. 241). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-71330-0_1

Cockell, C. S., Schwendner, P., Perras, A., Rettberg, P., Beblo-Vranesevic, K., Bohmeier, M., Rabbow, E., Moissl-Eichinger, C., Wink, L., Marteinsson, V., Vannier, P., Gomez, F., Garcia-Descalzo, L., Ehrenfreund, P., Monaghan, E. P., Westall, F., Gaboyer, F., Amils, R., Malki, M., Pukall, R., Cabezas, P., Walter, N. (2018). Anaerobic microorganisms in astrobiological analogue environments: From field site to culture collection. *International Journal of Astrobiology*, *17*(4), 314–328. https://doi.org/10.1017/S1473550417000246

Cockell, C. S. (2021). Are microorganisms everywhere they can be? *Environmental Microbiology*. https://doi.org/10.1111/1462-2920.15825

Colangelo-Lillis, J., Eicken, H., Carpenter, S. D., & Deming, J. W. (2016). Evidence for marine origin and microbial-viral habitability of sub-zero hypersaline aqueous inclusions within permafrost near Barrow, Alaska. *FEMS Microbiology Ecology*, *92*(5). https://doi.org/10.1093/femsec/fiw053

Collins, R. E., & Deming, J. W. (2011). Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up. *Polar Biology*, *34*(12), 1831–1841. https://doi.org/10.1007/s00300-011-1008-z

Cooper, Z. S., Rapp, J. Z., Carpenter, S. D., Iwahana, G., Eicken, H., & Deming, J. W. (2019). Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiology Ecology*, *95*(12). Z. Cooper dissertation, Chapter 2. https://doi.org/10.1093/femsec/fiz166

Cordero, O. X., & Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Reviews Microbiology*. Nature Publishing Group. https://doi.org/10.1038/nrmicro3218

Cox, G. F. N., & Weeks, W. F. (1983). Equations for determining the gas and brine volumes in sea ice samples. *Journal of Glaciology*, *29*(102), 306–316. https://doi.org/10.3189/s0022143000008364

Darwin, C. (1859) *On the origin of species by means of natural selection*. London: J. Murray. [PDF] Retrieved from the Library of Congress, https://www.loc.gov/item/06017473/.

Deming, J. W., & Eric Collins, R. (2016). Sea ice as a habitat for Bacteria, Archaea and viruses. In D. N. Thomas (Ed.), *Sea Ice: Third Edition* (pp. 326–351). https://doi.org/10.1002/9781118778371.ch13

Durot, M., Bourguignon, P. Y., & Schachter, V. (2009). Genome-scale models of bacterial metabolism: Reconstruction and applications. *FEMS Microbiology Reviews*, *33*(1), 164–190. https://doi.org/10.1111/j.1574-6976.2008.00146.x

Elena, S. F., & Lenski, R. E. (2003). Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, *4*(6), 457–469. https://doi.org/10.1038/nrg1088

Fenchel, T., & Finlay, B. J. (2004). The ubiquity of small species: patterns of local and global diversity. *BioScience*, *54*(8), 777–784. https://doi.org/10.1641/0006-3568(2004)054[0777:TUOSSP]2.0.CO;2

Fetterer, F., Knowles, K., Meier, W., Savoie, M., & Windnagel., A. K. (2017). Sea Ice Index, Version 3. Extent Anomaly Graphs, updated daily. Accessed 1 November, 2021. https://doi.org/10.7265/N5K072F8

Firth, E., Carpenter, S. D., Sørensen, H. L., Collins, R. E., & Deming, J. W. (2016). Bacterial use of choline to tolerate salinity shifts in sea-ice brines. *Elementa*, *2016*. https://doi.org/10.12952/journal.elementa.000120

Fricker, A. M., Podlesny, D., & Fricke, W. F. (2019). What is new and relevant for sequencing-based microbiome research? A mini-review. *Journal of Advanced Research*, *19*, 105–112. https://doi.org/10.1016/j.jare.2019.03.006

Fuchsman, C. A., Collins, R. E., Rocap, G., & Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ*, *2017*(9), e3865. https://doi.org/10.7717/peerj.3865

Galtier, N. (2007). A model of horizontal gene transfer and the bacterial phylogeny problem. *Systematic Biology*, *56*(4), 633–642. https://doi.org/10.1080/10635150701546231

Germerodt, S., Bohl, K., Lück, A., Pande, S., Schröter, A., Kaleta, C., Schuster, S., Kost, C. (2016). Pervasive Selection for Cooperative Cross-Feeding in Bacterial Communities. *PLoS Computational Biology*, *12*(6), e1004986. https://doi.org/10.1371/journal.pcbi.1004986

Gilichinsky, D., Rivkina, E., Shcherbakova, V., Laurinavichuis, K., & Tiedje, J. (2003). Supercooled water brines within permafrost - An unknown ecological niche for microorganisms: A model for astrobiology. *Astrobiology*, *3*(2), 331–341. https://doi.org/10.1089/153110703769016424

Gilichinsky, D., Rivkina, E., Bakermans, C., Shcherbakova, V., Petrovskaya, L., Ozerskaya, S., Ivanushkina, N., Kochkina, G., Laurinavichuis, K., Pecheritsina, S., Fattakhova, R., Tiedje, J. M. (2005). Biodiversity of cryopegs in permafrost. In *FEMS Microbiology Ecology*, *53*, 117–128. https://doi.org/10.1016/j.femsec.2005.02.003

Gillings, M. R. (2017). Lateral gene transfer, bacterial genome evolution, and the Anthropocene. *Annals of the New York Academy of Sciences*, *1389*(1), 20–36. https://doi.org/10.1111/nyas.13213

Golden, K. M., Eicken, H., Heaton, A. L., Miner, J., Pringle, D. J., & Zhu, J. (2007). Thermal evolution of permeability and microstructure in sea ice. *Geophysical Research Letters*, *34*(16), 2–7. https://doi.org/10.1029/2007GL030447

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., & Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature 2017 551:7678*, *551*(7678), 45–50. https://doi.org/10.1038/nature24287

Harris, L. K., & Theriot, J. A. (2018). Surface area to volume ratio: a natural variable for bacterial morphogenesis. *Trends in Microbiology*, *26*(10), 815–832. https://doi.org/10.1016/j.tim.2018.04.008

Harrison, J. P., Gheeraert, N., Tsigelnitskiy, D., & Cockell, C. S. (2013). The limits for life under multiple extremes. *Trends in Microbiology*, *21*(4), 204–212. https://doi.org/10.1016/j.tim.2013.01.006

Ilbert, M., & Bonnefoy, V. (2013). Insight into the evolution of the iron oxidation pathways. *Biochimica et Biophysica Acta - Bioenergetics*, *1827*(2), 161–175. https://doi.org/10.1016/j.bbabio.2012.10.001

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press. In Press.

Iwahana, G., Cooper, Z. S., Carpenter, S. D., Deming, J. W., & Eicken, H. (2021). Intra-ice and intra-sediment cryopeg brine occurrence in permafrost near Utqiaġvik (Barrow). *Permafrost and Periglacial Processes*, *32*(3), 427–446. https://doi.org/10.1002/ppp.2101

Jafari, M., Ansari-Pour, N., Azimzadeh, S., & Mirzaie, M. (2017). A logic-based dynamic modeling approach to explicate the evolution of the central dogma of molecular biology. *PLOS ONE*, *12*(12), e0189922. https://doi.org/10.1371/JOURNAL.PONE.0189922

Jansson, J. K., & Taş, N. (2014). The microbial ecology of permafrost. *Nature Reviews Microbiology*, *12*(6), 414–425. https://doi.org/10.1038/nrmicro3262

Lang, A. S., Zhaxybayeva, O., & Beatty, J. T. (2012). Gene transfer agents: Phage-like elements of genetic exchange. *Nature Reviews Microbiology*. Nature Publishing Group. https://doi.org/10.1038/nrmicro2802

Lever, M. A., Rogers, K. L., Lloyd, K. G., Overmann, J., Schink, B., Thauer, R. K., Hoehler, T.M., Jørgensen, B. B. (2015). Life under extreme energy limitation: A synthesis of laboratory- and field-based investigations. *FEMS Microbiology Reviews*, *39*(5), 688–728. https://doi.org/10.1093/femsre/fuv020

Li, S. J., Hua, Z. S., Huang, L. N., Li, J., Shi, S. H., Chen, L. X., Kuang, J.L., Liu, J., Hu, M., Shu, W. S. (2014). Microbial communities evolve faster in extreme environments. *Scientific Reports*, *4*(1), 1–9. https://doi.org/10.1038/srep06205

Maier, R. M. (2015). Biogeochemical Cycling. In *Environmental Microbiology: Third Edition* (pp. 339–373). Academic Press. https://doi.org/10.1016/B978-0-12-394626-3.00016-8

Maistrenko, O. M., Mende, D. R., Luetge, M., Hildebrand, F., Schmidt, T. S. B., Li, S. S., Coelho, L.P., Huerta-Cepas, J., Sunagawa, S., Bork, P. (2020). Disentangling the impact of environmental and phylogenetic constraints on prokaryotic strain diversity. *ISME Journal*, 735696. https://doi.org/10.1101/735696

Malard, L. A., & Pearce, D. A. (2018). Microbial diversity and biogeography in Arctic soils. *Environmental Microbiology Reports*, *10*(6), 611–625. https://doi.org/10.1111/1758-2229.12680

Mikucki, J. A., Pearson, A., Johnston, D. T., Turchyn, A. V., Farquhar, J., Schrag, D. P., Anbar, A.D., Priscu, J.C., Lee, P. A. (2009). A contemporary microbially maintained subglacial ferrous "Ocean." *Science*, *324*(5925), 397–400. https://doi.org/10.1126/science.1167350

Moulana, A., Anderson, R. E., Fortunato, C. S., & Huber, J. A. (2020). Selection Is a Significant Driver of Gene Gain and Loss in the Pangenome of the Bacterial Genus *Sulfurovum* in Geographically Distinct Deep-Sea Hydrothermal Vents. *MSystems*, *5*(2). https://doi.org/10.1128/mSystems.00673-19

Niederberger, T. D., Perreault, N. N., Tille, S., Lollar, B. S., Lacrampe-Couloume, G., Andersen, D., Greer, C.W., Pollard, W., Whyte, L. G. (2010). Microbial characterization of a subzero, hypersaline methane seep in the Canadian high arctic. *ISME Journal*, *4*(10), 1326–1339. https://doi.org/10.1038/ismej.2010.57

Nunn, B. L., Slattery, K. V., Cameron, K. A., Timmins-Schiffman, E., & Junge, K. (2015). Proteomics of *Colwellia psychrerythraea* at subzero temperatures - a life with limited movement, flexible membranes and vital DNA repair. *Environmental Microbiology*, *17*(7), 2319–2335. https://doi.org/10.1111/1462-2920.12691

Oliveira, P. H., Touchon, M., Cury, J., & Rocha, E. P. C. (2017). The chromosomal organization of horizontal gene transfer in bacteria. *Nature Communications*, *8*(1), 1–10. https://doi.org/10.1038/s41467-017-00808-w

Overland, J., Dunlea, E., Box, J. E., Corell, R., Forsius, M., Kattsov, V., Olsen, M. S., Pawlak, J., Reiersen, L. O., & Wang, M. (2019). The urgency of Arctic change. *Polar Science*, 21, 6–13. https://doi.org/10.1016/j.polar.2018.11.008

Perreault, N. N., Andersen, D. T., Pollard, W. H., Greer, C. W., & Whyte, L. G. (2007). Characterization of the prokaryotic diversity in cold saline perennial springs of the Canadian high arctic. *Applied and Environmental Microbiology*, *73*(5), 1532–1543. https://doi.org/10.1128/AEM.01729-06

Petrich, C., & Eicken, H. (2017). Overview of sea ice growth and properties. In D. N. Thomas (Ed.), *Sea Ice: Third Edition* (pp. 1–41). Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118778371.ch1

Price, P. B., & Sowers, T. (2004). Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(13), 4631–4636. https://doi.org/10.1073/pnas.0400522101

Rapp, J. Z., Sullivan, M. B., & Deming, J. W. (2021). Divergent genomic adaptations in the microbiomes of Arctic subzero sea-ice and cryopeg brines. *Frontiers in Microbiology*, *12*, 1–21. https://doi.org/10.3389/fmicb.2021.701186

Romero Manrique, D., Corral, S., & Guimarães Pereira, Â. (2018). Climate-related displacements of coastal communities in the Arctic: Engaging traditional knowledge in adaptation strategies and policies. *Environmental Science and Policy*, *85*, 90–100. https://doi.org/10.1016/j.envsci.2018.04.007

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Kultima, J. R., Martin, J., Kota, K., Sunyaev, S. R., Weinstock, G. M., & Bork, P. (2013). Genomic variation landscape of the human gut microbiome. *Nature 2012 493:7430*, *493*(7430), 45–50. https://doi.org/10.1038/nature11711

Schloter, M., Lebuhn, M., Heulin, T., & Hartmann, A. (2000). Ecology and evolution of bacterial microdiversity. *FEMS Microbiology Reviews*, *24*(5), 647–660. https://doi.org/10.1111/j.1574-6976.2000.tb00564.x

Stevenson, A., Burkhardt, J., Cockell, C. S., Cray, J. A., Dijksterhuis, J., Fox-Powell, M., Key, T. P., Kminek, G., McGenity, T. J., Timmis, K. N., Timson, D. J., Voytek, M. A., Westall, F., Yakimov, M. M., & Hallsworth, J. E. (2015). Multiplication of microbes below 0.690 water activity: implications for terrestrial and extraterrestrial life. *Environmental Microbiology*, *17*(2), 257–277. https://doi.org/10.1111/1462-2920.12598

Styczinski, M.J., Cooper, Z.S., Glaser, D.M., Lehmer, O., Mierzejewski, V., Tarnas, J. Chapter 7: Assessing habitability beyond Earth. In NASA Astrobiology Primer 3.0. *Astrobiology* (in review).

Toussaint, A., & Chandler, M. (2012). Prokaryote genome fluidity: Toward a system approach of the mobilome. *Methods in Molecular Biology*, *804*, 57–80. https://doi.org/10.1007/978-1-61779-361-5_4

Weiss, M. C., Sousa, F. L., Mrnjavac, N., Neukirchen, S., Roettger, M., Nelson-Sathi, S., & Martin, W. F. (2016). The physiology and habitat of the last universal common ancestor. *Nature Microbiology*, *1*(9), 1–8. https://doi.org/10.1038/nmicrobiol.2016.116

Wells, L. E., & Deming, J. W. (2006). Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environmental Microbiology*, *8*(6), 1115–1121. https://doi.org/10.1111/j.1462-2920.2006.00984.x

Xie, S., Pancost, R. D., Yin, H., Wang, H., & Evershed, R. P. (2005). Two episodes of microbial change coupled with Permo/Triassic faunal mass extinction. *Nature*, *434*(7032), 494–497. https://doi.org/10.1038/nature03396

Zhong, Z.-P., Rapp, J. Z., Wainaina, J. M., Solonenko, N. E., Maughan, H., Carpenter, S. D., Deming, J. W., Sullivan, M. B. (2020). Viral ecogenomics of Arctic cryopeg brine and sea ice. *MSystems*, *5*(3), e00246-20. https://doi.org/10.1128/msystems.00246-20

Chapter 2. DISTINCTIVE MICROBIAL COMMUNITIES IN SUBZERO HYPERSALINE BRINES FROM ARCTIC COASTAL SEA ICE AND RARELY SAMPLED CRYOPEGS

This chapter has previously been published:

Cooper, Z. S., Rapp, J. Z., Carpenter, S. D., Iwahana, G., Eicken, H., & Deming, J. W. (2019). Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiology Ecology*, *95*(12), 1–15. https://doi.org/10.1093/femsec/fiz166

2.1 ABSTRACT

Hypersaline aqueous environments at subzero temperatures are known to be inhabited by microorganisms, yet information on community structure in subzero brines is very limited. Near Utqiaġvik, Alaska, we sampled subzero brines (–6°C, 115–140 ppt) from cryopegs, i.e. unfrozen sediments within permafrost that contain relic (late Pleistocene) seawater brine, as well as nearby sea ice brines to examine microbial community composition and diversity using 16S rRNA gene amplicon sequencing. We also quantified the communities microscopically and assessed environmental parameters as possible determinants of community structure. The cryopeg brines harbored surprisingly dense bacterial communities (up to 10^8 cells mL^{-1}) and millimolar levels of dissolved and particulate organic matter, extracellular polysaccharides, and ammonia. Community composition and diversity differed between the two brine environments by alpha- and beta-diversity indices, with cryopeg brine communities appearing less diverse and dominated

by one strain of the genus *Marinobacter*, also detected in other cold, hypersaline environments, including sea ice. The higher density and trend towards lower diversity in the cryopeg communities suggest that long-term stability and other features of a subzero brine are more important selective forces than *in situ* temperature or salinity, even when the latter are extreme.

2.2 INTRODUCTION

Microorganisms inhabit nearly all aqueous systems on Earth, including those with extreme temperatures. In highly saline environments, the freezing point of water can be depressed well below 0˚C, enabling liquid brine to exist within the system. Such subzero hypersaline brines serve as microbial habitats in many natural environments (Boetius *et al.* 2015) including sea ice (Brown and Bowman 2001; Junge *et al.* 2001; Eronen-Rasimus *et al.* 2016; Deming and Collins 2017; Yergeau *et al.* 2017; Rapp *et al.* 2018), cryopegs in permafrost (Gilichinsky *et al.* 2003, 2005; Spirina *et al.* 2017), ice-covered saline lakes and aquifers (Ward and Priscu 1997; Mikucki *et al.* 2009; Murray *et al.* 2012), and saline terrestrial ponds and springs (Meyer *et al.* 1962; Perrault *et al.* 2007; Niederberger *et al.* 2010). These brines are similar with respect to characteristic temperature and salinity but vary in other physical and chemical parameters. Two types of these subzero hypersaline habitats exist within kilometers of each other in the coastal region of the Alaskan Arctic: the brines of landfast sea ice (e.g., Ewert and Deming 2014) and the brines of cryopegs (Yoshikawa *et al.* 2004; Colangelo-Lillis *et al.* 2016). They serve as natural systems for studying the influence of extreme environmental conditions on microbial community structure and diversity.

Cryopegs are subsurface features within a stable permafrost matrix, where fluids exist in discrete layers and do not freeze at the subzero temperatures due to high concentrations of salts and other dissolved materials (Gilichinsky *et al.* 2003, 2005). Cryopegs remain at nearly the same temperature year-round, observed to range from –9 to –11°C in Siberian cryopegs (Gilichinsky *et al.* 2005) and –6 to –8°C in Alaskan cryopegs (Colangelo-Lillis *et al.* 2016; K. Yoshikawa, personal communication), with the minor variations related to season and distance from the

surface. Several cryopegs have been categorized as thalassohaline (having a salt composition indicative of a seawater source) and are proposed to have formed from saturated marine sediments that were exposed to the atmosphere with falling sea level during glaciation periods, becoming desiccated and incorporated into permafrost along with surrounding material over thousands of years (Gilichinsky *et al.* 2003, 2005; Meyer *et al.* 2010a, 2010b; Colangelo-Lillis *et al.* 2016). As thalassohaline permafrost features, cryopegs serve as natural systems for studying the effects of long-term exposure of marine microbial communities to subzero, hypersaline conditions in isolation from other aqueous environments and solar radiation.

The formation of cryopegs is analogous to the hypothesized scenario for recession of an ancient ocean on Mars. During desiccation of an ancient Martian ocean and cooling of the atmosphere, dissolved materials would have concentrated into highly saline brines that saturated surface materials. Hypersaline brines would have been forced underground as the Martian surface became cryotic, forming solid ice complexes and rejecting brines to the subsurface (Knauth and Burt 2002; Ehlmann *et al.* 2011; McEwen *et al.* 2011). These cryopeg-like systems could conceivably provide sanctuaries for life on Mars where the surface is too cold and irradiated to support life as known on Earth (Jakosky *et al.* 2003; McKay *et al.* 2013; Westall *et al.* 2013; Hassler *et al.* 2014). Recent evidence suggests that subsurface brines of a more contemporary nature may exist on Mars (Ojha *et al.* 2015). Analog environments like cryopegs can thus serve as models for understanding the adaptations required and community structures that may result from long-term exposure to subzero, hypersaline conditions.

Microorganisms have been cultured from East Siberian and North Alaskan cryopegs (Gilichinsky *et al.* 2005; Kochkina *et al.* 2007; Pecheritsyna *et al.* 2007, 2012; Shcherbakova *et al.* 2009; Spirina *et al*. 2017), including isolates of *Psychrobacter*, *Desulfovibrio*, and *Brevibacterium*, bacterial genera with cold-adapted and halotolerant species, as well as some eukaryotic yeast and mycelial fungi. Near Utqiaġvik, Alaska, cryopeg brines occur directly beneath a permafrost tunnel that was excavated to study freshwater ice wedge formation in the area (Yoshikawa *et al.* 2004; Meyer *et al.* 2010a, 2010b). Paleoclimate and geological research date the sediment directly below the tunnel as deposited more than 14,000 years ago and likely earlier (Meyer *et al.* 2010a, 2010b). Today cryopeg brines exist roughly 8 m below the surface, and directly below a massive ice formation, previously characterized as a complex of ice wedges of meteoric origin (here we use the more general geological term "massive ice", following more recent geological work; G. Iwahana, unpublished data). Measurements of ^{18}O from the cryopeg brines and the massive ice indicate separate origins, with the brines having originated most likely from seawater before the ice formed and been isolated for a minimum of 14,000 years (Colangelo-Lillis *et al.* 2016). Colangelo-Lillis *et al.* (2016) collected a limited volume of brine from below the tunnel floor in 2009 that was subjected to high-throughput DNA sequencing of the *in situ* viral community, which indicated low diversity relative to other aquatic environments, while counts of bacteria and virus-like particles (VLP) suggested an active microbial community that likely included *Marinobacter* based on detection of *Marinobacter* genes in the virome. Another thalassohaline cryopeg system has also been identified near Utqiaġvik, at a different location and greater depth (40 m; Yoshikawa *et al*. 2004). From a brine sample recovered from this system, Spirina *et al.* (2017) cultured representatives of four bacterial phyla (seven genera), many of which were halotolerant, though no *Marinobacter* were reported. Information on the

structure and diversity of the microbial community using non-cultivation methods is not available for this or any known cryopeg system.

During the formation of sea ice, the much better studied of the two subzero brine systems we examined, salts and other dissolved or suspended materials in seawater are concentrated into (sub)millimeter-scale brine channels and pockets that exist within the ice matrix (Petrich and Eicken 2017). The resulting brines increase in salinity and concentration of other materials as the brine volume shrinks with decreasing temperature during winter, allowing the brines to remain liquid well below –20˚C (Cox and Weeks 1983). The presence of brine in sea ice takes the form of a liquid-filled network inhabited by microbial communities (Junge *et al*. 2004), where microbial motility along nutritional and salinity gradients may occur and nutrients can be exchanged with the underlying water column (Showalter and Deming 2018). Inhabiting this brine network over the lifetime of the ice comes with the stressors of fluctuating extremes in temperature and salinity. Temperatures in sea ice can range from about –1˚C to below –20˚C annually, with seasonal, diurnal and storm-induced fluctuations, while brine salinities vary as a function of temperature from near fresh conditions (during melt season) to greater than 220 ppt in winter (Ewert and Deming 2014; Deming and Collins 2017). The nature of sea ice thus exposes the microbial community initially entrained within the ice to extreme physicochemical conditions, including stressful fluctuation events (Collins *et al.* 2010; Petrich and Eicken 2017). Specific adaptations to survive the combined extremes of subzero temperature and hypersalinity include exopolysaccharide production for cryoprotection (Krembs *et al*. 2011; Deming and Young 2017) and organic compatible solute uptake and metabolism for osmotolerance (Firth *et al.* 2016; Torstensson *et al.* 2019).

The classes Gammaproteobacteria, Alphaproteobacteria, and Bacteroidia (including members of the former class Flavobacteriia, reassigned to class Bacteroidia in release 132 of the SILVA database) are abundant in sea ice and include members well adapted to cold and saline conditions (Bowman *et al.* 1997; Groudieva *et al.* 2004; Boetius *et al.* 2015). Adaptations to fluctuating extreme conditions are less well known (Ewert and Deming 2014), nor can we find community structure analyses of brines collected directly from sea ice without first melting the ice and thus altering brine salinity and composition. In 2018, Arctic sea ice covered 14.5 million km^2 at its maximum extent and 4.6 million km^2 at its minimum extent (Fetterer *et al.* 2017), demonstrating both the spatial scale of subzero hypersaline environments and their ephemeral and highly variable nature. These fluctuations induce a taxonomic shift, resulting in differently composed communities present in the ice as selective conditions change seasonally (Collins *et al.* 2010; Bowman *et al.* 2012; Eronen-Rasiumus *et al.* 2015). As climate warming continues, the timing and extent of these seasonal fluctuations will likely change. Climate warming impacts can also be seen in permafrost, as the rate of permafrost thaw increases (Grosse *et al.* 2016). Reports of permafrost thaw and brine seepage into ice cellars across the North Slope of Alaska suggest a possible threat to food security for indigenous people of the region (Nyland *et al.* 2017). Cryopeg brines may also be affected by climate change and contribute to these threats. As the brines warm and migrate through the permafrost, their microbial communities may benefit from new conditions and food sources. Changes in the microbial communities of sea ice, permafrost and cryopegs will be important to monitor to understand the biogeochemical impacts of a warming climate.

In this study, we collected multiple cryopeg brines from below the permafrost tunnel near Utqiaġvik, Alaska, to investigate resident prokaryotic community structure and diversity between borehole locations and between years. We also sampled the massive ice deposit above the cryopeg brines, as well as nearby coastal sea ice brines, to compare microbial community structure and diversity in cryopeg brines with that in adjacent environmental material and in another marine-derived subzero brine system. All samples were subjected to high-throughput amplicon sequencing of the V3–V4 hypervariable regions of the 16S rRNA gene, along with analyses of selected physicochemical conditions to consider possible selective forces acting *in situ* on the brine communities. The *in situ* composition and diversity of microbial communities in cryopeg brines is unknown to date — how such rarely sampled communities are structured in relation to those of surrounding and nearby extreme environments can help to reveal ecological mode and history. The influence of prolonged exposure of microbial communities to extremely low temperatures and high salinities, coupled with geophysical isolation, has not been fully explored. Here, we lay the foundation to explore these open questions, leading the way for future gene-based studies to uncover specific functions and interactions, including with viruses, that allow life to adapt and flourish in subzero hypersaline brines.

2.3 MATERIALS AND METHODS

Site descriptions

We collected samples near Utqiaġvik, Alaska, in May 2017 and May 2018 as part of a larger project to investigate gene exchange in subzero hypersaline environments. Cryopeg brines were obtained from the Barrow Permafrost Tunnel, which lies 6 m below the surface at 71.2944 °N, 156.7153 °W. The tunnel was excavated by the United States Army Cold Regions Research and Engineering Laboratory in the early 1960s, entirely within a massive ice formation in the permafrost (Meyer *et al.* 2010a). Working inside the tunnel at –6°C, we collected cryopeg brines and selected massive ice samples from discrete boreholes that generally penetrated about 2 m below the tunnel floor, through the massive ice into permafrost and unfrozen (brine-saturated) cryopeg sediment (Figure 2.1). Samples were obtained from both previously established boreholes (Colangelo-Lillis *et al.* 2016) and new boreholes, drilled with a cleaned and ethanol-rinsed ice auger or SIPRE corer (as in Colangelo-Lillis *et al.* 2016) until unfrozen sediment or liquid brine was encountered. This effort was time-consuming (taking 4–6 hours to enter the tunnel, drill a borehole that yielded brine, and collect a new sample) and challenging, given the low temperature and limited space in the tunnel (sharp ice crystals have been growing inward from the tunnel walls and ceiling since its excavation), but every precaution was taken to keep pristine any cryopeg brine encountered. We installed newly purchased, ethanol-rinsed PVC or ABS pipe in each new borehole, wearing ethanol-rinsed nitrile gloves for these and all proximal operations.

Landfast sea ice sampling sites were located at 71.3730 ˚N, 156.5047 ˚W in May 2017 and 71.4730 ˚N, 156.7294 ˚W in May 2018 near the Barrow Sea Ice Mass Balance site operated by

the University of Alaska Fairbanks (Druckenmiller *et al*. 2009). Air temperature during sampling in both years varied between –6° and –4°C. Sites were cleared of snow prior to drilling sackholes: snow depth was 16–19 cm in 2017 and 6–10 cm in 2018. Sea ice brine (SB) was collected by drilling multiple sackholes, 1 m apart, into the sea ice, following Eicken *et al.* (2009). In May 2017, the sea ice was 117 cm thick, and sackholes were drilled to 75 cm; in May 2018, ice thickness was 110 cm, and sackholes were drilled to 55 cm. Vertical temperature profiles were measured from two sea ice cores for each sampling year, following Eicken *et al.* (2009), to estimate sackhole brine temperature at relevant depths (and avoid inserting a thermoprobe into the pristine brine). Upper sea ice temperatures (above 75 cm in 2017, above 55 cm in 2018) were –4˚C in 2017 and –3˚C in 2018.

Sample collection and processing

Cryopeg brine was collected using an apparatus composed of hand pump, acid-washed 2-L vacuum flask (initially autoclaved, then acid-washed/ ethanol-rinsed between samples), and 1.5-m length of Masterflex and Teflon tubing. Prior to a day of cryopeg sampling, multiple lengths of this tubing were acid-washed, rinsed with MilliQ water, rinsed with ethanol, dried to remove excess liquid that might freeze in the tunnel, and sealed with sterile parafilm in the laboratory at the Barrow Arctic Research Center (BARC) for transport to the tunnel. Each cryopeg brine was collected from a discrete borehole (named accordingly; Figure 2.1) and returned to the BARC immediately for storage at –6˚C until processing. Massive ice samples were collected by using an ethanol-cleaned spatula to gather ice shavings during ice auger drilling (CB4_IW1_17 and CB4_IW2_17) or as whole ice cores (CBIW_IW3_17 and CBIW_IW7_17) that were placed into sterile Whirl-Pak® bags. Massive ice samples were placed at –20˚C upon return to the BARC

and kept frozen until later processing at the University of Washington. Further details on each borehole and sampling operation are provided in supplemental information (Text S2.1).

Sea ice brines were collected in both years after allowing brine to drain from the surrounding sea ice into (covered) sackholes over a period of 3–5 hours. Brine was pooled in the field from multiple sackholes to obtain 20-L samples by hand-pumping the brines into an acid-washed cubitainer first rinsed with sample brine. Sea ice from the drained area was not collected. In 2018, we collected four additional (unpooled) 500-mL brine samples (SB1–4_18), each from a separate sackhole, to compare with the pooled sea ice brine communities.

All brine samples were transported in insulated containers to subzero cold rooms at BARC (set to *in situ* brine temperatures) and processed as soon as possible after collection (within 2–8 hours). Samples collected from the permafrost tunnel and filtered for DNA extraction were limited in volume: 25–500 mL for cryopeg brine and 12.5–50 mL for melted massive ice. For sea ice brines, a volume of 250 mL was filtered for each individual sackhole brine sample in 2018, while 2500 mL were filtered for the brine pooled from multiple sackholes each year. These samples were filtered onto 0.22 µm filters (either Sterivex, without a prefilter, or 47-mm in-line GTTP filters with a 3.0-µm GFA prefilter; Table 2.1) in a BARC cold room at 4ºC. The filters were stored immediately at –80ºC until shipped (surrounded by –80°C ice packs) to the University of Washington in Seattle for similar storage until further processing.

Designation of sample types

The samples collected for this study fell into three general categories: the two types of subzero brines, cryopeg brine and sea ice brine; and the massive ice found above the cryopeg brine. The massive ice and cryopeg brine samples from below the permafrost tunnel were divided further into subtypes (Table 2.1), distinguished by salinity or locational source, as described below. Sea ice brine samples were not separated further, as they were not similarly distinguishable. Table 2.1 lists all of the samples and whether or not the sample was prefiltered. Prefiltration was performed for selected samples, processed for later metagenomic analyses, to separate eukaryotic and prokaryotic cells to the extent possible. Both filtered and prefiltered samples were included in this study, however, as the prokaryotic 16S rRNA gene was targeted for amplification and amplicons assigned to eukaryotes were removed *in silico*.

Massive ice

The massive ice samples were grouped by salinity (Table 2.1). Samples CB4_IW1_17 and CB4_IW2_17 were collected as auger shavings during the drilling of CB4 and had no detectable salinity (as measured on the melted sample); they were grouped as “massive ice”. Samples CBIW_IW3_17 and CBIW_IW7_17 were collected as core sections of the massive ice while drilling CBIW. They had detectable brine content (measurable salinity on the melted sample) and were grouped as “massive ice with some brine”.

Cryopeg brines

Where multiple samples were available from the same cryopeg borehole (CB), the cryopeg brine samples were grouped by locational source; i.e. “intra-ice brine” encountered *in situ* within the

massive ice and “intra-sediment brine” encountered in contact with sediment. Remaining intra-sediment brines, represented by a single sample per location, were grouped as “intra-sediment brine (other)”. Although brine sampling in 2017 was limited (as in 2009; Colangelo-Lillis *et al.* 2016), collections from new boreholes drilled in 2018 increased the overall spatial coverage of this subsurface brine system. They also allowed consideration of potential changes over time, as intra-sediment brine was collected in 2018 from the same borehole sampled in 2009 (designated CB1 by Colangelo-Lillis *et al.* 2016), while intra-ice brine was collected from another borehole (CBIW, originally designated IWC by Colangelo-Lillis *et al.* 2016) in successive years, 2017 and 2018. CBIW also yielded brine samples 4 days apart in 2018 (Table 2.1). A cryopeg brine sample that had been collected and filtered in-line (3.0 µm prefilter, then 0.2 µm filter, with both filters stored at –70°C) in August 2009 by Colangelo-Lillis *et al.* (2016) was also included in this study. All individual samples with descriptors are listed in Table 1; further details about drilling circumstances and boreholes are provided in the supporting information (Text S2.1).

Enumeration of cells and virus-like particles

Samples were collected for bacterial and viral enumeration, fixed using 0.2-µm filtered 37% formaldehyde at a final concentration of 2%, and stored at 4°C until returned to the University of Washington. Total cell counts and the percentage of dividing cells were obtained by filtering the fixed samples onto 0.2 µm filters, cross-staining with 4′,6-diamidino-2-phenylindole (DAPI) and acridine orange, and analyzing by epifluorescence microscopy at 1562.5x (12.5x objective X 1.25x optivar X 100x oil immersion lens) on a 1953 Zeiss Universal microscope, following Sherr and Sherr (1983) and Ewert and Deming (2014). Total VLPs were enumerated at The Ohio State University using the wet-mount method described by Cunningham *et al.* (2016) using SYBR

Gold. Counts below the detection limit of that method (1 x 10^6 VLP mL^{-1}) were made at the University of Washington by filtering samples onto 0.02 µm anodisc filters and staining with SYBR Gold, following Noble and Fuhrmann (1998).

Quantification of physicochemical conditions, nutrients, and organic matter

We measured temperature in the field using a digital thermometer, brine salinity via refractometry, and pH using pH paper. Where sample volume allowed, concentrations of particulate organic carbon and nitrogen (POC, PN), particulate and dissolved extracellular polysaccharides (pEPS, dEPS), dissolved organic carbon (DOC), and the nutrients PO_4, NO_3, NO_2, and NH_4 were determined. For POC (and PN), samples were collected on pre-combusted (500˚C) Whatman GF/F filters. Sample filters were fumed with HCL to remove inorganic carbon, and CHN was measured on a Leeman Labs Model CEC440 Elemental Analyzer in the Marine Chemistry Laboratory (MCL) at the University of Washington. DOC was measured on 0.2 µm sample filtrate using a Shimadzu TOC-VCSH DOC analyzer according to standard protocols in the MCL. For EPS measurements we used the phenol-sulfuric acid method following Krembs *et al.* (2011), with conversion from glucose-equivalents (standards based on glucose concentrations) to carbon-equivalents. Samples for pEPS were collected on 0.4 µm polycarbonate filters, with the filtrate (< 0.4 µm) representing dEPS. Nutrients were analyzed in the MCL using a Technicon AutoAnalyzer II (Knap *et al.* 1996).

DNA extraction and sequencing

We removed all Sterivex filters from their plastic housings using ethanol-rinsed, flame-sterilized pliers. The Sterivex and all 47-mm filters were cut into smaller pieces using similarly sterilized

scalpels and forceps. DNA was extracted from cut filters using the DNeasy PowerSoil Extraction kit (QIAGEN, Hilden, Germany), quantified using Qubit dsDNA HS assay (Invitrogen, Carlsbad, CA), and checked for quality (A260/280 and A260/230) using a NanoDrop One spectrophotometer (ThermoFisher Scientific, Waltham, MA). Aliquots of extracted DNA containing at least 50 ng DNA were sent to GENEWIZ, Inc. (South Plainfield, NJ, USA) for library preparation and sequencing.

The preparation of next generation sequencing libraries and Illumina MiSeq sequencing was conducted by GENEWIZ, Inc. (South Plainfield, NJ, USA), where DNA samples were quantified using a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and DNA quality was checked on a 0.6% agarose gel. The sequencing library was constructed using a MetaVx™ Library Preparation kit (GENEWIZ, Inc., South Plainfield, NJ, USA). Briefly, 50 ng of DNA were used to generate amplicons that cover the V3 and V4 hypervariable regions of the 16S rRNA gene of Bacteria and Archaea. The V3 and V4 regions were amplified using proprietary primer sets developed at GENEWIZ, Inc. (South Plainfield, NJ, USA), where the forward primers contain the sequence ‘‘CCTACGGRRBGCASCAGKVRVGAAT” (similar to 341F primers) and the reverse primers contain the sequence ‘‘GGACTACNVGGGTWTCTAATCC” (similar to 806R primers; Huang et al. 2014). Indexed adapters were added to the ends of the amplicons by limited cycle PCR. Sequencing libraries were validated using an Agilent 4200 TapeStation (Agilent Technologies, Palo Alto, CA, USA), and quantified by Qubit and real time PCR (Applied Biosystems, Carlsbad, CA, USA). DNA libraries were multiplexed and loaded on an Illumina MiSeq instrument according to manufacturer’s instructions (Illumina, San Diego, CA, USA). Sequencing was performed using a 2 x 250 paired-end (PE) configuration; image

analysis and base calling were conducted by the MiSeq Control Software (MCS) on the MiSeq instrument.

Sequence processing

We received demultiplexed paired-end reads from GENEWIZ, Inc. (South Plainfield, NJ, USA), which have been deposited in the Sequence Read Archive under BioProject accession number PRJNA540708 or individually under accession numbers SRR9003057– SRR9003078. Reads were assembled, quality-filtered, and aligned following the MiSeq SOP for *mothur* v1.40.1 (Schloss *et al.* 2009; Kozich *et al.* 2013). To accommodate processing of the V3 and V4 regions of the 16S rRNA gene, we deviated from the SOP as follows: assembled reads shorter than 400 bp and longer than 480 bp were discarded; the SILVA release 132 non-redundant alignment reference provided through the *mothur* wiki (https://mothur.org/wiki/Silva_reference_files) was trimmed to accommodate the regions sequenced here by using the pcr.seqs command given the start and end positions of our sequences within the alignment; preclustering was set to allow up to four differences; classification of unique sequences was performed using the SILVA taxonomy file; and unique reads assigned to chloroplasts, mitochondria, Eukaryota, and unknown domains were discarded. OTU clustering (at 97% similarity) was performed using the *VSEARCH* distance-based greedy clustering algorithm (Rognes *et al.* 2016) within *mothur* (updated here to v1.40.5 to fix an error in the *mothur* implementation of *VSEARCH* clustering) rather than the default *Opticlust* algorithm (Westcott and Schloss 2017) because of its lower memory requirement and higher speed on this size of dataset. We conducted taxonomic classification of OTUs using the SILVA (release 132) reference alignment provided in the *mothur* wiki. Statistical analyses and figures were generated using the *R* (version 3.4.4; R Core

Team 2018) package *Phyloseq* (version 1.22.3; McMurdie and Holmes 2013). Singletons were removed from the OTU data before all analyses to avoid including sequencing error. A log of mothur commands used is available in supporting information (Text S2.2).

Biodiversity and statistical calculations

The alpha diversity metrics calculated include the number of observed OTUs and Chao1 index to measure richness and the Shannon and Inverse Simpson indices to measure diversity. All reads but singletons (OTUs with a single read across samples) were kept to avoid creating statistical miscalculations that can propagate when rarefying reads. OTU abundances were $\log_{10}(x + 1)$ transformed for beta-diversity analyses using base *R* to normalize read counts, stabilize variance, and reduce error due to heteroscedasticity (McMurdie and Holmes 2014). Using the *R* package *vegan* (version 2.5.3; Oksanen *et al.* 2018), we performed an analysis of variance (ANOVA) along with a post-hoc pairwise significance (Tukey's honestly significant difference) test to assess the significance ($p \leq 0.05$) of variance in the richness and diversity indices between the three main sample types: massive ice (n = 4), cryopeg brine (n = 10), and sea ice brine (n = 8). To assess similarity of the communities, Bray-Curtis dissimilarity-based ordinations, based on transformed OTU abundances, were generated using non-metric multidimensional scaling (NMDS) with the ordinate function in *Phyloseq* set to iterate 1000 times to find the lowest stress ordination possible, along with an analysis of similarity (ANOSIM) using the anosim function in vegan, on samples grouped further into subtypes (Table 2.1). NMDS ordinations were plotted in a split format showing sample similarities in one panel with OTU contributions, colored by genus, displayed adjacently. A principal component analysis (PCA) was performed in base *R* using the prcomp function for the subset of brine samples that were accompanied by the full suite

of collected environmental data to assess which of those parameters might drive differences between cryopeg and sea ice brines. Scripts used for biodiversity calculations and PCA are available in supporting information (Text S2.2).

Minimal entropy decomposition

To assess variations in strain composition, we used the *oligotyping* program to conduct a minimal entropy decomposition (MED; version 2.0; Eren *et al.* 2015) analysis on sequences belonging to the genus *Marinobacter* from samples where they were the most relatively abundant taxa. We processed *mothur* taxonomy, count, and fasta files generated after unique sequence classification to obtain only *Marinobacter* sequences. First, samples not dominated by *Marinobacter* were removed from the dataset in *mothur* using the remove.groups function. Then, the fasta file was screened using the screen.seqs function to select for sequences that were exactly 462 bp long (the median length of all sequences present after the initial sample removal), leaving a total of 529,228 sequences of equal length assigned to the genus *Marinobacter*. All alignment gaps were removed from the fasta file using the function degap.seqs. The processed *mothur* files were then used as input in the shell script mothur2oligo.sh, found at https://github.com/DenefLab/MicrobeMiseq/tree/master/mothur2oligo, which allows for sequences from a specific taxon to be extracted and reformatted for use in the MED analysis. Using the properly formatted *Marinobacter* sequence file, MED analysis was run with default parameters, where sequences that are dissimilar are iteratively split into nodes of similar sequences until the Shannon entropy of each base within all nodes reaches below a threshold of 0.0965.

2.4 RESULTS

Environmental conditions

During the May campaigns of both years, the temperature at the tunnel floor was –6°C. The temperature of cryopeg brine *in situ* was not measured directly to limit disturbance of such rarely accessed fluids; however, based on *in situ* thermistors deployed through a separate, dedicated borehole (CB3) in the tunnel between December 2010 and March 2012, temperature at cryopeg depth is relatively stable, oscillating between –8°C and –6°C (K. Yoshikawa, personal communication). Sea ice temperatures at the sackhole depth and time of sampling were slightly warmer, –4°C in 2017 and –3°C in 2018, though the ice (and its inhabitants) would have experienced much lower temperatures prior to May (e.g., upper sea ice temperatures near Utqiaġvik can fall below –20°C during winter; Ewert and Deming 2014, their Figure 1a; Petrich and Eicken 2017, their Figure 1.10). The sea ice sackhole brines of this study had salinities of 78 ppt in 2017 and 75 ppt in 2018, as generally expected based on ice temperature (Cox and Weeks 1983), while the colder cryopeg brines had higher salinities of 140 ppt in 2017 and 112–122 ppt in 2018.

Additional environmental parameters measured on the subset of brine samples with sufficient volume showed that cryopeg brines were extremely rich in organic compounds, with concentrations of POC, DOC, pEPS and dEPS (in C equivalents) measured at the millimolar level, 2–4 orders of magnitude higher than the micromolar concentrations measured in sea ice brine (precise values provided in Table 2.2). Most of the organic carbon in both types of brine was in the dissolved form: 30–102 mM DOC and 9–20 mM dEPS in cryopeg brines compared to 0.20–0.45 mM DOC and 0–0.20 mM dEPS in sea ice brines. The particulate organic component

was overall smaller but still much higher in the cryopeg brines: 2–12 mM POC and 0.09–14 mM pEPS in cryopeg brines compared to 0.020–0.022 mM POC and 0.001–0.002 mM pEPS in sea ice brines. Cryopeg brines also contained much higher levels of NO_2 and NH_4^+ relative to sea ice brines, while NO_3^- and PO_4^{3-} concentrations were similarly low in the two environments (Table 2.2). C:N ratios were generally low (5.8–6.4 in cryopeg and 7.1–9.0 in sea ice brine), and pH values approached neutral (6.6 in cryopeg and 7.2 in sea ice brine).

The PCA, based on physical and biogeochemical data (Table 2.2), resulted in a clear separation of cryopeg and sea ice brines along PC1, which explained 53.2% of the variation between sample types (Figure 2.2). Each of the factors enriched in cryopeg brines relative to sea ice brines contributed nearly equally to explaining the differences between brine types. No single driving factor among those measured could be identified, though VLP abundance and concentrations of nitrate and nitrite appeared less influential than other factors.

DNA-containing cell and VLP abundance

The abundance of DAPI-stained cells in cryopeg brine was remarkably high (Table 2.3), ranging from 5.70×10^6 to 1.70×10^8 mL^{-1}, compared to $1.11–2.22 \times 10^5$ mL^{-1} in sea ice brine. In samples evaluated for the number of dividing cells (Table 2.3), 1.1–5.5% of the total were noted as dividing in cryopeg brine samples collected from boreholes (none were observed in CBIA_surf_18 from the tunnel floor), compared to 4.7% in sea ice sackhole brine. A striking exception was cryopeg brine sample CB4_18, where 29% or 3.31×10^6 of the 1.14×10^7 cells mL^{-1} were observed as dividing. Although cellular abundance in this sample fell on the lower end of the range for cryopeg brines, the absolute number of dividing cells was twice that found in

other cryopeg brines (1.30–1.82 x 10^6) and orders of magnitude higher than observed in sea ice brine (5.22 x 10^3). VLP concentrations ranged from 3.54×10^6 to 3.89×10^8 mL^{-1} in cryopeg brine compared to $1.53–6.16 \times 10^5$ mL^{-1} in sea ice brine (Table 2.3). Ratios of VLP to cells in cryopeg brines collected in 2017 and 2018 ranged from 0.31 to 3.3 (below the ratio of 10 reported by Colangelo-Lillis *et al.* 2016, for cryopeg brine sampled in 2009) but generally comparable to ratios in sea ice brine, which ranged from 0.69 to 3.00 (Table 2.3). The lowest ratio was observed for cryopeg brine sample CB4_18.

Subzero brine community structure

After quality control, sequence processing, and removal of singletons, a total of 1,789,043 unique reads across the 22 sequenced samples could be assigned to the domains Bacteria and Archaea. Read totals for individual samples ranged from 29,543 to 143,814 (Table S2.1). Rarefaction curves of OTUs vs sampling depth appear nearly saturated for all samples (Figure S2.1). The sampling blank (MilliQ rinse water from between-sample cleaning procedures for the cryopeg sampling tubing) contained a low number of reads (Figure S2.1). As these reads primarily matched the most abundant OTUs from the *Marinobacter*-dominated brine samples, with some exogenous members present exclusively in the blank (MilliQ source water was virtually devoid of cells by epifluorescence microscopy), we attribute them to low-abundance cell retention in the sample tubing, insufficient to influence the high-density brine samples. Samples were grouped according to sample type, massive ice (n = 4); cryopeg brine (n = 10); and sea ice brine (n = 8), for richness and diversity analyses to compare between environments (Figure 2.3). Average Shannon diversity (H') values for the massive ice, cryopeg brine, and sea ice brine samples were 3.17, 1.92, and 2.49, respectively. ANOVA of all measures showed

significant differences between sample types ($p < 0.0001$ for richness indices and $p < 0.01$ for diversity indices). Post hoc Tukey HSD calculations were generated from the ANOVA for each measure to allow pairwise sample comparisons. Communities in massive ice samples (much lower in bacterial abundance) were significantly richer ($p < 0.001$, Observed/Chao1 index) and more diverse ($p < 0.01$, Shannon/Inverse Simpson index) than those in either cryopeg or sea ice brines (Table S2.1). Sea ice brine communities appeared more diverse than those in cryopeg brines, but without statistical significance ($p = 0.079$, Shannon diversity index; $p = 0.294$, Inverse Simpson index). Cryopeg brines appeared less rich than sea ice brines, but without statistical significance in number of observed OTUs ($p = 0.679$) or Chao1 richness index ($p = 0.851$). NMDS ordination (based on Bray-Curtis dissimilarity) of the samples by $\log_{10}(x+1)$-transformed OTU composition (Figure 2.4) had a stress value of 0.075 and displayed a clear separation of cryopeg brine and massive ice samples from sea ice brine samples, as also supported statistically by ANOSIM ($R = 0.79$, $p = 9.9 \times 10^{-4}$).

Taxonomic classifications

Across all samples, the classes Gammaproteobacteria and Bacteroidia dominated the microbial communities by relative abundance (Figure 2.5A). Cryopeg brines were numerically dominated by Gammaproteobacteria (average relative abundance of 57%) followed by Bacteroidia (average relative abundance of 22%). Sea ice brines were dominated by Bacteroidia (average relative abundance of 51%) followed by Gammaproteobacteria (average relative abundance of 30%). Archaea were rare, with all reads assigned to the domain making up only 0.094% of total reads across all samples. Only 22 OTUs (0.5%) were assigned to Archaea out of 3833 OTUs across all samples. The most abundant archaeal OTU (83% of all archaeal reads) belongs to the class

Methanomicrobia genus *Methanosaeata*. Of the reads in this OTU, 89% came from a single massive ice sample (CB4_IW1_17). The class Methanomicrobia made up 10 of the 22 Archaeal OTUs, found primarily in massive ice samples; the class Nitrososphaeria made up 8 of the 22 Archaeal OTUs, found primarily in sea ice brines. Few archaeal OTUs were identified in cryopeg brines, e.g. one in CB4_18 (0.063% of reads) and another in CBIW_18 (0.011% of reads), while others were even less abundant. Archaea were not considered further. No cyanobacterial reads, not attributed to chloroplasts, were detected in the sea ice samples.

At the genus level, the differences in taxonomic composition increase between massive ice, cryopeg brines, and sea ice brines (Figure 2.5B). All but one of the cryopeg brines were dominated by *Marinobacter* and *Gillisia* with average relative abundances of 49% and 15%, respectively.

The exception was the cryopeg brine community from CB4, which was dominated by *Psychrobacter* at 54% and *Bacteroides* at 11% (and also anomalous for its high percentage of dividing cells and some environmental properties). The general relative abundances of the cryopeg brine community members were validated by comparison to a shotgun metagenome of a 2017 sample from CBIW that depicts similar abundance values for the major taxa (Text S2.3, Figure S2.2). Massive ice samples, collected from boreholes during drilling to access cryopegs, were also dominated by *Marinobacter* and *Gillisia* with average relative abundances of 35% and 18%, respectively. A small number of reads (27–243) of the main OTU of *Marinobacter* from the cryopeg brines was also found in each of the sea ice brine samples. In contrast, sea ice brine communities were dominated by *Polaribacter* and *Octadecabacter* with average relative abundances of 46% and 8%, respectively (Figure 2.5B). Although some minor differences were

observed between individual sackhole brine communities that were spatially distinct, and pooling the brines appeared to favor the dominant genus (by de-emphasizing the more diverse rare members), the sea ice brine communities were largely similar in composition.

Given the high relative abundances of *Marinobacter* in all but one of the cryopeg brine and massive ice samples at both genus and OTU levels, a MED analysis was performed to assess whether strain variation occurs between these samples. Approximately 80% of *Marinobacter* sequences could be attributed to a single representative oligotype. BLAST analysis of this sequence against the NCBI 16S rRNA database showed 99.6% similarity with *Marinobacter psychrophilus* strain i20041 cultured from sea ice in the Canadian Basin (Zhang *et al.* 2008).

2.5 DISCUSSION

The cryopeg brine system underlying the Barrow Permafrost Tunnel provides a unique opportunity to consider microbial adaptation to subzero hypersaline conditions in a geophysically stable and isolated environment. Our efforts to characterize the cryopeg brines and their resident microbial communities and contrast them with the more variable and less isolated environment of sea ice brines represent a first step in this direction. Based on radiocarbon dating, the cryopeg system we examined has been isolated for at least 14,000 years (Colangelo-Lillis *et al*. 2016) and, based on local stratigraphy, possibly from the last major transgression in 123 ka (Eisner et al., 2005). Cryopeg brines may thus share the selective pressures of subzero temperature and hypersalinity with contemporary sea ice brines, but they diverge widely in age, physical environmental interactions, and the magnitude of seasonal fluctuations in temperature and salinity (Ewert and Deming 2014; K. Yoshikawa, personal communication).

They also diverge in organic content, as the cryopeg brines of this study were highly enriched in dissolved and particulate organic carbon and extracellular polysaccharides relative to sea ice brines (millimolar versus micromolar levels; Table 2.2). The DOC levels in cryopeg brine (range of 30–102 mM) were also several orders of magnitude greater than those reported for other subzero brines from Blood Falls, Antarctica (80–450 uM DOC; Mikucki and Priscu 2007). Although more detailed chemical analyses are needed, much of the organic carbon in our cryopeg brine samples was in the form of dissolved carbohydrate (e.g. 9–20 mM dEPS), suggesting ample labile or semi-labile carbon to drive microbial heterotrophy. Millimolar levels of ammonia (1.2–4.5 mM NH_4) in the brines and low C:N ratios (5.8–6.7) are also consistent with the remineralization of labile carbon by marine bacteria. The high concentrations of total and dividing cells and their viruses, and the organic-rich conditions of the cryopeg brines presented in this study, show that these subsurface thalassohaline environments can support remarkable microbial communities, despite continuously extreme temperature and salinity. The cell densities measured, which far exceeded those in the sampled sea ice brines, resemble values for late-logarithmic growth phase in a nutrient-replete bacterial culture, suggesting that the cryopeg brines are near carrying capacity but that the organisms are not under growth-inhibiting nutrient limitation (Eiler *et al.* 2003). The *in situ* microbial communities in these subsurface brines thus appear to be well-adapted to the conditions of their subzero hypersaline habitat conventionally considered to be stressful.

Cryopeg brines are likely anoxic environments *in situ* due to their physical residence in a solid permafrost matrix at depth (in the dark, removed from the atmosphere). We had no means to

measure oxygen concentration *in situ* without risking contamination of the brines, nor were we able to prevent oxygen exposure during sampling despite best efforts. We collected and processed samples as rapidly as possible to prevent alteration of the brine communities before filtration and freezing. Chua *et al.* (2018) estimated a doubling time of 9.9 days at –4.5˚C for *Marinobacter gelidimuriae*, a close relative of the dominant *Marinobacter* identified here. Mykytczuk *et al.* (2013) reported a doubling time of 28 days at –5°C and 180 ppt (NaCl) and about 144 days at –10°C and 180 ppt (NaCl) for their low-temperature, record-holding permafrost isolate, *Planococcus halocryophilus* Or1. Similarly low growth rates have been reported for other psychrophilic bacteria found in subzero brine environments (Huston 2003; Methé *et al.* 2005; Ewert and Deming 2014). Given the short time between sampling and processing (2–8 hours) and strict maintenance of samples at *in situ* temperature and salinity, no significant growth or alteration of community structure can be expected during this period. We thus suggest that the taxonomic abundances and diversity indices presented here represent the best possible insight to the *in situ* community structure of these rarely sampled cryopeg brines.

The alpha-diversity indices of this study indicate that these dense cryopeg brine communities tend to be less diverse than those in sea ice brines, as well as significantly less diverse than in the overlying massive ice. Higher diversity (H’ = 2.49) in sea ice brines may be attributed to the broad fluctuations in environmental conditions that select for both phenotypic plasticity and different taxa during different seasons (Collins *et al.* 2010; Ewert and Deming 2014; Deming and Collins 2017). Lower diversity (H’ = 1.92) in cryopeg brines may be because the magnitude of fluctuations experienced is much lower, requiring less plasticity and favoring fewer taxa. Over the long period of geological isolation of cryopeg brines from the ocean, the extreme conditions

may have forced taxa typically competitive under warmer, lower salinity conditions to become rare members of the community or be lost from it, lowering the overall diversity. An alternative (non-exclusive) explanation for lower diversity in cryopeg brines is that the measured high concentrations of organic substrates have allowed copiotrophic microbial taxa to outcompete others of oligotrophic mode. None of these explanations preclude the possibility that even relatively minor seasonal or annual fluctuations in environmental conditions over long periods of time may have played a role in structuring the microbial community. The observation of highest diversity (H' = 3.17) in the low-abundance microbial assemblage in massive ice likely results from the depositional nature of this feature at the time of its formation. A very limited liquid phase and thus low biological activity in the massive ice would favor preservation of deposited cells, similar to what has been observed in other permafrost features (Jansson *et al.* 2014). The presence of dominant cryopeg brine taxa in the massive ice samples may result from thermal migration of cryopeg fluids or interactions at the brine/ice boundary where partial melting and refreezing occurs as temperatures oscillate seasonally between –8 and –6°C. Massive-ice refreezing in cold seasons may entrain cryopeg brine taxa, causing major taxa to appear similar between these sample types while minor taxa differ.

Beta diversity analyses suggest that communities in sea ice brines have distinct structures that differ from those present in the subterranean system of massive ice and cryopeg brines. Within that system, communities in massive ice samples with undetectable bulk ice salinity (measured by refractometry) separate farther from cryopeg brines than do those in massive ice samples that clearly contained some brine, implying some physical connectivity between massive ice and cryopeg brine. In line with this implication, communities in intra-ice brine (from CBIW) exhibit

higher similarity to massive ice communities than to intra-sediment brine communities (from CB1; Figure 2.4). Other samples of cryopeg brines (single samples per location) with confirmed or suspected connectivity with massive ice, i.e. by upward migration through the ice between sampling years (CBIA) or by drilling circumstances that may have partially mixed ice, sediment and brine (G2), also appear similar to the massive ice samples that contained some brine. The community in intra-sediment brine CB4_18, observed to be anomalous in other biological and chemical characteristics (Tables 2.2, 2.3), appears unique in the NMDS plot relative to all other brine samples. When CB4 was drilled to a depth of 187 cm in 2017 (and left covered), no brine was encountered, but by 2018 brine had infiltrated into the void (and up the borehole to within 29 cm of the tunnel floor) created the previous year (Text S2.1). The history and location of this brine (the most westerly of other intra-sediment brines; Figure 2.1) thus differ from the other cryopeg brines, potentially explaining the separation observed.

The common occurrence and high abundance of the classes Gammaproteobacteria and Bacteroidia in the landfast sea ice and cryopeg brines of this study, as well in other studies of Arctic and Antarctic sea ice (Brown and Bowman 2001; Bowman *et al.* 2012; Yergeau *et al.* 2017; Rapp *et al.* 2018), Baltic sea ice (Eronen-Rasimus *et al.* 2015), Arctic terrestrial springs (Niederberger *et al.* 2010), ice-covered saline lakes and aquifers (Ward and Priscu 1997; Mikucki and Priscu 2007; Murray *et al.* 2012), and other cryopeg systems in Siberia and Alaska (Gilichinsky *et al.* 2005; Pecheritsyna *et al.* 2007; Spirina *et al.* 2017), highlight the consistent and dominant occupation of subzero hypersaline niches by these classes. Genus-level differences between the brines studied here demonstrate the effects of local selective pressures on the specific adaptive capabilities found within each class.

The cryopeg brines of this study (with one exception) were dominated by the genus *Marinobacter*, which appears as a cosmopolitan member of cold, saline aqueous environments. For example, *Marinobacter* is abundant in a cold saline spring on Axel Heiberg Island in the Canadian High Arctic (Niederberger *et al.* 2010), a system that shares physicochemical characteristics with cryopeg brines, including temperature, hypersalinity, residence within or below permafrost, and isolation from light and the atmosphere (Perreault *et al.* 2007). *Marinobacter* is also abundant in Blood Falls, the saline subglacial outflow from Taylor Glacier, Antarctica (Mikucki and Priscu 2007). Cultivation studies have yielded isolates of *Marinobacter* from both the cold spring system on Axel Heiberg Island (Niederberger *et al.* 2010) and Blood Falls (Chua *et al.* 2018), as well as from a variety of other saline (if warmer) environments, including coastal sediments, oil-polluted seawater, low-temperature hydrothermal sediments, and the phycospheres of diatoms and dinoflagellates (Handley and Lloyd 2013). *Marinobacter* isolates show metabolisms for utilizing a range of organic compounds, from simple amino acids and glucose to complex hydrocarbons, as well as for nitrate reduction, chemolithoautotrophic Fe(II) oxidation, and other aerobic and anaerobic capabilities (Edwards *et al.* 2003; Kaye *et al.* 2011; Handley and Lloyd 2013). This wide variety of metabolic functions, along with the general traits of psychrotolerance, halotolerance, phenotypic plasticity, flexible heterotrophic metabolism, and an r-strategist, copiotrophic lifestyle, make the *Marinobacter* genus highly successful across a variety of saline environments.

Two psychrophilic strains of *Marinobacter* have been isolated: *M. psychrophilus* strain i20041 from arctic sea ice (Zhang *et al.* 2008) and *M. gelidimuriae* from Blood Falls in Antarctica (Chua

et al. 2018). *M. gelidimuriae* shares high 16S sequence similarity (97.8%) with *M. psychrophilus* strain i20041 (Chua *et al.* 2018), as does the one *Marinobacter* strain that dominates in our cryopeg brine samples (99.6%). If future work yields a cryopeg brine isolate with similar physiological characteristics, then together these organisms may represent a distinct clade of *Marinobacter* uniquely adapted to subzero, hypersaline brine environments. In the one cryopeg brine sample in our study where *Marinobacter* did not dominate (CB4_18), the genus was nevertheless present. Given the lower total bacterial density, higher percentage of dividing cells, lowest VBR, and other anomalies of this sample, we speculate that its current *Psychrobacter*-dominated community was in a different stage of development, relative to those spatially distant from it (Figure 2.1), and that *Marinobacter* may yet become dominant, a testable hypothesis with future sampling. The similarity in *Marinobacter*-dominated communities in other cryopeg locations over time, from 2009 to 2018 (Figure 2.5B), lends support to this hypothesis.

The abundance and dominance of *Marinobacter* across the sampled cryopeg brines raises questions not only about the adaptability of this genus but also the potential interconnectivity of brine pockets within this subterranean massive ice-cryopeg system. Although individual brines varied in their chemical characteristics and composition of low-abundance members of the community, MED analysis indicates that a single *Marinobacter* strain dominated the samples. Several possible scenarios may explain this dominance across seemingly isolated borehole samples. Partial freeze-thaw cycles (between –8 and –6 °C) at the brine/ice boundaries in a cryopeg system may separate the brine into distinct pools in colder seasons but allow subvolumes to connect in warmer seasons for repeat periods of homogenization. Alternatively, the sampled brines may have been physically isolated *in situ* yet contained communities that

shared a common highly-adapted strain due to a common successional history driven by the same temperature and salinity conditions that favor this organism. The high level of relatedness of this *Marinobacter* strain to the contemporary sea ice isolate, *M. psychrophilus*, suggests a selection for similar phenotypic traits important to competing successfully in both cryopeg and sea ice brines, despite isolation of the cryopeg habitat over recent geologic time. The potential for phenotypic plasticity mediated by viral processes under the relevant *in situ* conditions remains an open question.

Subzero hypersaline aqueous environments are widespread at high latitudes on Earth and long studied in sea ice given visible microbial colonization of its brine network. Despite the much smaller number of available studies for cryopegs, these systems are thought to be widespread in coastal permafrost settings (Yoshikawa *et al.* 2004; Gilichinsky *et al.* 2005). Local observations suggest that cryopeg brines may be seeping into traditional deep ice cellars, posing a possible threat to Indigenous food security (Nyland *et al*. 2017). As Earth's warming atmosphere and oceans continue to reduce the extent of polar sea ice and permafrost, microbial communities in subzero brines may be altered in ways that are difficult to predict yet may have strong influences not only on their immediate surroundings, but also on general ecosystem functions in the future. Uncovering microbial community composition and diversity in these extreme environments is thus urgent, and is now possible through interdisciplinary collaboration, modern sampling and DNA sequencing techniques, as exemplified here. Metagenomic, -viromic, and -transcriptomic studies (ongoing from this study) will help to reveal details of metabolic function and adaptive mechanisms. Information on these systems may be valuable not only to understanding ecological dynamics during a changing climate on Earth, but also in guiding considerations of potentially

habitable environments beyond Earth, including subsurface brine features that have been identified on Mars (Ojha *et al.* 2015) and icy moons, such as Europa and Enceladus (McCord *et al.* 2002; Zolotov 2007). The ecological patterns observed in terrestrial analog environments like cryopegs can inform expectations and features to anticipate when searching for life in similarly extreme environments throughout the Solar System.

2.6 CONCLUSIONS

The subzero hypersaline cryopeg and sea ice brines sampled near Utqiaġvik, Alaska, harbored sizeable microbial communities, including dividing cells, despite being characterized by conditions considered biologically stressful. The cryopeg brines, isolated in a subsurface environment for at least 14,000 years, hosted particularly dense microbial assemblages. The lower diversity of the microbial communities in the cryopeg brines reflects the dominance of highly adapted bacterial members, particularly of the genus *Marinobacter*, that have thrived at the very low temperatures and hypersaline conditions of these brines. The more diverse sea ice microbial communities, containing a canonical bacterial membership, are indicative of the seasonally fluctuating and ephemeral nature of this environment. Both brine environments contained highly abundant members from the same taxonomic classes but differed markedly at the genus level. We suggest that shared extreme temperature and salinity are sufficient to explain selection at the class level, but that genus-level distinctions are driven by other environmental parameters, including the magnitude and frequency of physical fluctuations experienced over time and the sources and levels of carbon and nitrogen, all of which remain to be fully explored. Ongoing research seeks to interrogate the function of these microbial ecosystems and the adaptive mechanisms used for exchanging genetic material when rapid adaptation is

advantageous, in order to further the understanding of how life adapts to and thrives under extreme conditions.

Funding

This work was supported by the Gordon and Betty Moore Foundation [grant number GBMF5488].

Conflicts of interest

There are no conflicts of interest to declare.

2.7 ACKNOWLEDGMENTS

Kenji Yoshikawa and Jerry Brown kindly supplied information and unpublished data regarding the Barrow Permafrost Tunnel and local geology. We thank Zhiping Zhong for providing total VLP counts for this study along with other members of the larger team for support in the field: Max Showalter, Anders Torstensson, Hannah Dawson, and Jodi Young. The Ukpeagvik Iñupiat Corporation Science team provided essential logistical support and helped with access to the tunnel. Metagenome sequencing was kindly performed by Rachel Mackelprang and Christopher Chabot at California State University Northridge.

References

Boetius A, Anesio AM, Deming JW *et al.* Microbial ecology of the cryosphere: Sea ice and glacial habitats. *Nat Rev Microbiol* 2015;**13**:677–90.

Bowman JP, McCammon SA, Brown M V *et al.* Diversity and association of psychrophilic bacteria in Antarctic sea ice. *Appl Environ Microbiol* 1997;63:3068–78.

Bowman JS. The relationship between sea ice bacterial community structure and biogeochemistry: A synthesis of current knowledge and known unknowns. *Elem Sci Anth* 2015;**3**:000072.

Bowman JS, Rasmussen S, Blom N *et al.* Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene. *ISME J* 2012;**6**:11–20.

Brown M V., Bowman JP. A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiol Ecol* 2001;35:267–75.

Chua MJ, Campen RL, Wahl L *et al*. Genomic and physiological characterization and description of *Marinobacter gelidimuriae* sp. nov., a psychrophilic, moderate halophile from Blood Falls, an antarctic subglacial brine. *FEMS Microbiol Ecol* 2018;94:1–15.

Colangelo-Lillis J, Eicken H, Carpenter SD *et al.* Evidence for marine origin and microbial-viral habitability of sub-zero hypersaline aqueous inclusions within permafrost near Barrow, Alaska. *FEMS Microbiol Ecol* 2016;**92**:1–15.

Collins RE, Rocap G, Deming JW. Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Environ Microbiol* 2010;**12**:1828–41.

Cox GFN, Weeks WF. Equations for determining the gas and brine volumes in sea-ice samples. *J Glaciol* 1983;**29**:306–16.

Cunningham BR, Brum JR, Schwenck SM, Sullivan MB, John SG. Wet-mount Method for Enumeration of Aquatic Viruses. *protocols.io* 2016, DOI: 10.17504/protocols.io.c8pzvm

Deming JW, Collins RE. Sea ice as a habitat for Bacteria, Archaea and viruses. In: Thomas DN (ed), *Sea Ice – An Introduction to its Physics, Chemistry, Biology and Geology*, Third Ed, John Wiley and Sons, Ltd, 2017, 326–51.

Deming JW, Young JN. The role of exopolysaccharides in microbial adaptation to cold habitats. *Psychrophiles: From Biodiversity to Biotechnology*, Second Ed, Cham: Springer International Publishing, 2017, 259–84.

Druckenmiller ML, Eicken H, Johnson M, Pringle D & Williams C. Towards an integrated coastal sea-ice observatory: system components and a case study at Barrow, Alaska. *Cold Reg Sci Technol* 2009;**56**:61–72.

Edwards KJ, Rogers DR, Wirsen CO *et al*. Isolation and characterization of novel psychrophilic, neutrophilic, Fe-oxidizing, chemolithoautotrophic α- and γ-Proteobacteria from the deep sea. *Appl Environ Microbiol* 2003;**69**:2906–13.

Ehlmann BL, Mustard JF, Murchie SL *et al.* Subsurface water and clay mineral formation during the early history of Mars. *Nature* 2011;**479**:53–60.

Eicken H, Gradinger R, Salganek M *et al. Field Technique for Sea Ice Research*. University of Alaska Press, 2009.

Eiler A, Langenheder S, Bertilsson S *et al.* Heterotrophic bacterial growth efficiency and community structure at different natural organic carbon concentrations. *Appl Environ Microbiol* 2003;**69**:3701–9.

Eisner WR, Bockheim JG, Hinkel KM *et al.* Paleoenvironmental analyses of an organic deposit from an erosional landscape remnant, Arctic Coastal Plain of Alaska. *Palaeogeogr Palaeoclimatol Palaeoecol* 2005;**217**:187–204.

Eren AM, Morrison HG, Lescault PJ *et al*. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. ISME J 2015;**9**:968–79.

Eronen-Rasimus E, Lyra C, Rintala JM *et al.* Ice formation and growth shape bacterial community structure in Baltic Sea drift ice. *FEMS Microbiol Ecol* 2015;**91**:1–13.

Eronen-Rasimus E, Piiparinen J, Karkman A *et al.* Bacterial communities in Arctic first-year drift ice during the winter/spring transition. *Environ Microbiol Rep* 2016;8:527–35.

Ewert M, Deming JW. Bacterial responses to fluctuations and extremes in temperature and brine salinity at the surface of Arctic winter sea ice. *FEMS Microbiol Ecol* 2014;**89**:476–89.

Fetterer, F., K. Knowles, W. Meier, M. Savoie and AKW. *Sea Ice Index, Version 3.0*., 2017.

Firth E, Carpenter SD, Sørensen HL *et al.* Bacterial use of choline to tolerate salinity shifts in sea-ice brines. *Elem Sci Anthr* 2016;**4**:000120.

Gilichinsky D, Rivkina E, Shcherbakova V *et al.* Supercooled water brines within permafrost—An unknown ecological niche for microorganisms: A model for Astrobiology. *Astrobiology* 2003;**3**:331–41.

Gilichinsky D, Rivkina E, Bakermans C *et al.* Biodiversity of cryopegs in permafrost. *FEMS Microbiol Ecol* 2005;**53**:117–28.

Grosse G, Goetz S, McGuire AD *et al.* Changing permafrost in a warming world and feedbacks to the Earth system. *Environ Res Lett* 2016;**11**:040201.

Groudieva T, Kambourova M, Yusef H *et al.* Diversity and cold-active hydrolytic enzymes of culturable bacteria associated with Arctic sea ice, Spitzbergen. *Extremophiles* 2004;8:475–88.

Handley KM, Lloyd JR. Biogeochemical implications of the ubiquitous colonization of marine habitats and redox gradients by *Marinobacter* species. *Front Microbiol* 2013;**4**:1–10.

Hassler DM, Zeitlin C, Wimmer-Schweingruber RF *et al.* Mars' surface radiation environment measured with the Mars Science Laboratory's Curiosity rover. *Science* 2014;**343**:1244797.

Huang J, Zhou Z, Duan H, Wu X, and S Li. Methods and kits for identifying microorganisms in a sample. US9745611B2. Alexandria, VA, USA: United States Patent and Trademark Office, 2014.

Huston AL. Bacterial adaptation to the cold: Isolation and characterization of the psychrophilic marine γ Proteobacterium *Colwellia psychrerythraea* strain 34H. PhD dissertation, University of Washington, 2003.

Jakosky BM, Nealson KH, Bakermans C *et al.* Subfreezing activity of microorganisms and the potential habitability of Mars' polar regions. *Astrobiology* 2003;**3**:343–50.

Jansson JK, Taş N. The microbial ecology of permafrost. *Nat Rev Microbiol* 2014;**12**:414–25.

Junge K, Krembs C, Deming J *et al.* A microscopic approach to investigate bacteria under in situ conditions in sea-ice samples. *Ann Glaciol* 2001;33:304–10.

Junge K, Eicken H, Deming JW. Bacterial activity at -2 to -20 degrees C in Arctic wintertime sea ice. *Appl Environ Microbiol* 2004;70:550–7.

Kaye JZ, Sylvan JB, Edwards KJ *et al*. *Halomonas* and *Marinobacter* ecotypes from hydrothermal vent, subseafloor and deep-sea environments. *FEMS Microbiol Ecol* 2011;**75**:123–33.

Kellogg CTE, Deming JW. Comparison of free-living, suspended particle, and aggregate-associated bacterial and archaeal communities in the Laptev Sea. *Aquat Microb Ecol* 2009;**57**:1–18.

Knap A, Michaels A, Close A *et al.* Protocols for the Joint Global Ocean Flux Study (JGOFS) Core Measurements, UNESCO, 1996.

Knauth LP, Burt DM. Eutectic brines on Mars: Origin and possible relation to young seepage features. *Icarus* 2002;**158**:267–71.

Kochkina GA, Ivanushkina NE, Akimov VN *et al.* Halo-and psychrotolerant *Geomyces* fungi from arctic cryopegs and marine deposits. *Microbiology* 2007;**76**:31–8.

Kozich JJ, Westcott SL, Baxter NT *et al.* Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl Environ Microbiol* 2013;**79**:5112–20.

Krembs C, Eicken H, Deming JW. Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic. *Proc Natl Acad Sci* 2011;**108**:3653–8.

McCord TB, Teeter G, Hansen GB *et al.* Brines exposed to Europa surface conditions. *J Geophys Res* 2002;**107**:5004.

McEwen AS, Ojha L, Dundas CM *et al.* Seasonal flows on warm Martian slopes. *Science (80-)* 2011;**333**:740–3.

McKay CP, Stoker CR, Glass BJ *et al.* The Icebreaker Life Mission to Mars: A search for biomolecular evidence for life. *Astrobiology* 2013;**13**:334–53.

McMurdie PJ, Holmes S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 2013;**8**:e61217.

McMurdie PJ, Holmes S. Waste Not, Want Not: Why rarefying microbiome data is inadmissible. *PLoS Comput Biol* 2014;10:e1003531.

Methé BA, Nelson KE, Deming JW *et al.* The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and proteomic analyses. *Proc Natl Acad Sci U S A* 2005;**102**:10913–8.

Meyer GH, Morrow MB, Wyss O *et al.* Antarctica: The microbiology of an unfrozen saline pond. *Science (80-)* 1962;**138**:1103–4.

Meyer H, Schirrmeister L, Andreev A *et al.* Late glacial and Holocene isotopic and environmental history of northern coastal Alaska - Results from a buried ice-wedge system at Barrow. *Quat Sci Rev* 2010a;**29**:3720–35.

Meyer H, Schirrmeister L, Yoshikawa K *et al.* Permafrost evidence for severe winter cooling during the Younger Dryas in northern Alaska. *Geophys Res Lett* 2010b;**37**:1–5.

Mikucki JA, Priscu JC. Bacterial diversity associated with Blood Falls, a subglacial outflow from the Taylor Glacier, Antarctica. *Appl Environ Microbiol* 2007;73:4029–39.

Mikucki JA, Pearson A, Johnston DT *et al.* A contemporary microbially maintained subglacial ferrous "Ocean." *Science (80-)* 2009;**324**:397–400.

Murray AE, Kenig F, Fritsen CH *et al.* Microbial life at –13 C in the brine of an ice-sealed Antarctic lake. *Proc Natl Acad Sci* 2012;**109**:20626–31.

Mykytczuk NCS, Foote SJ, Omelon CR *et al.* Bacterial growth at −15 °C; molecular insights from the permafrost bacterium *Planococcus halocryophilus* Or1. *ISME J* 2013;7:1211–26.

Niederberger TD, Perreault NN, Tille S *et al.* Microbial characterization of a subzero, hypersaline methane seep in the Canadian high arctic. *ISME J* 2010;**4**:1326–39.

Noble RT, Fuhrman JA. Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquat Microb Ecol* 1998;14:113–8.

Nyland KE, Klene AE, Brown J *et al.* Traditional Iñupiat ice cellars (SIĠḷUAQ) in Barrow, Alaska: Characteristics, temperature monitoring, and distribution. *Geogr Rev* 2017;**107**:143–58.

Ojha L, Wilhelm MB, Murchie SL *et al.* Spectral evidence for hydrated salts in recurring slope lineae on Mars. *Nat Geosci* 2015;**8**:829–32.

Oksanen, J, Blanchet, FG, Kindt, R, Legendre, P *et al*. *vegan: Community Ecology Package*. R Package Version 2.5-3. 2018. https://CRAN.R-project.org/package=vegan

Pecheritsyna SA, Shcherbakova VA, Kholodov AL *et al.* Microbiological analysis of cryopegs from the Varandei Peninsula, Barents Sea. *Microbiology* 2007;**76**:614–20.

Pecheritsyna SA, Rivkina EM, Akimov VN *et al. Desulfovibrio arcticus* sp. nov., a psychrotolerant sulfate-reducing bacterium from a cryopeg. *Int J Syst Evol Microbiol* 2012;**62**:33–7.

Perreault NN, Andersen DT, Pollard WH *et al.* Characterization of the prokaryotic diversity in cold saline perennial springs of the Canadian high Arctic. *Appl Environ Microbiol* 2007;**73**:1532–43.

Petrich C, Eicken H. Overview of sea ice growth and properties. *Sea Ice*. Chichester, UK: John Wiley & Sons, Ltd, 2017, 1–41.

R Core Team. *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing 2018, http://www.R-project.org/.

Rapp JZ, Fernández-Méndez M, Bienhold C and Boetius A. Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Front Microbiol* 2018;**9**:1035.

Rognes T, Flouri T, Nichols B *et al.* VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 2016;**4**:e2584.

Schloss PD, Westcott SL, Ryabin T *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009;**75**:7537–41.

Shcherbakova VA, Chuvil'skaya NA, Rivkina EM *et al.* Novel halotolerant bacterium from cryopeg in permafrost: Description of *Psychrobacter muriicola* sp. nov. *Microbiology* 2009;**78**:84–91.

Sherr EB and BF Sherr. Double-staining epifluorescence technique to assess frequency of dividing cells and bacteriovory in natural populations of heterotrophic microprotozoa. *Appl Environ Microbiol* 1983;46:1388-93.

Showalter GM, Deming JW. Low-temperature chemotaxis, halotaxis and chemohalotaxis by the psychrophilic marine bacterium *Colwellia psychrerythraea* 34H. *Environ Microbiol Rep* 2018;**10**:92–101.

Spirina EV, Durdenko EV, Demidov NE *et al.* Halophilic-psychrotrophic bacteria of an Alaskan cryopeg—a model for astrobiology. *Paleontol J* 2017;**51**:1440–52.

Ward BB, Priscu JC. Detection and characterization of denitrifying bacteria from a permanently ice-covered Antarctic Lake. *Hydrobiologia* 1997;347:57–68.

Westall F, Loizeau D, Foucher F *et al.* Habitability on Mars from a microbial point of view. *Astrobiology* 2013;**13**:887–97.

Westcott SL, Schloss PD. OptiClust, an improved method for assigning amplicon-based sequence data to operational taxonomic units. *mSphere* 2017;**2**:e00073-17.

Yergeau E, Michel C, Tremblay J *et al.* Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Sci Rep* 2017;7:42242.

Yoshikawa K, Romanovsky V, Duxbury N *et al.* The use of geophysical methods to discriminate between brine layers and freshwater taliks in permafrost regions. *J Glaciol Geocryol* 2004;**26**:301–9.

Zhang DC, Li HR, Xin YH *et al. Marinobacter psychrophilus* sp. nov., a psychrophilic bacterium isolated from the Arctic. *Int J Syst Evol Microbiol* 2008;**58**:1463–6.

Zolotov MY. An oceanic composition on early and today's Enceladus. *Geophys Res Lett* 2007;**34**:L23203, doi:10.1029/2007GL031234.

Table 2.1. Description of individual samples in this study by sample subtype

Sample subtype	#	Sample[a]	T (˚C)[b]	S (ppt)	Description	Size fraction (µm)[c]
Massive ice	1	CB4_IW1_17	–6	0	Ice shavings, upper borehole	> 0.2
	2	CB4_IW2_17	–6	0	Ice shavings, lower borehole	> 0.2
Massive ice with brine	3	CBIW_IW3_17	–6	7	Ice core section, upper borehole	> 0.2
	4	CBIW_IW7_17	–6	22	Ice core section, lower,borehole	> 0.2
Intra-ice brine	5	CBIW_17	–6	140	Brine, bottom of borehole	0.2–3.0
	6	CBIW_18	–6	121	Brine, bottom of borehole	> 0.2
	7	CBIW_re_18	–6	120	Recharged brine, 4 days later	> 0.2
Intra-sediment brine	8	CB1_0.2_09	–6	115	Brine, upper borehole	0.2–3.0
	9	CB1_3.0_09	–6	115	Brine, upper borehole	> 3.0
	10	CB1_18	–6	122	Brine, lower borehole	> 0.2
Intra-sediment brine (other)	11	CBIA_surf_18	–6	120	Brine, borehole surface seepage	> 0.2
	12	CBIA_18	–6	112	Brine, lower borehole, slurried with ice crystals, sulfidic odor	> 0.2
	13	G2_18	–6	109	Brine, lower borehole, slurried with sediment	> 0.2
	14	CB4_18	–6	121	Brine, lower borehole	> 0.2
Sackhole brine from sea ice	15	SB_17	–4	78	Brine, pooled from sackholes	0.2–3.0
	16	SB_0.2_18	–3	75	Brine, pooled from sackholes	0.2–3.0
	17	SB_3.0_18	–3	75	Brine, pooled from sackholes	> 3.0
	18	SB_18	–3	75	Brine, pooled from sackholes	> 0.2
	19	SB1_18	–3	75	Brine, sackhole #1	> 0.2
	20	SB2_18	–3	75	Brine, sackhole #2	> 0.2
	21	SB3_18	–3	75	Brine, sackhole #3	> 0.2
	22	SB4_18	–3	75	Brine, sackhole #4	> 0.2

[a] CB indicates cryopeg borehole; G, borehole intended for geology; SB, sackhole brine; last two numbers of sample name, sampling year (2009, 2017, 2018). See Text S2.1 for details of borehole terminology and drilling circumstances; Figure 2.1 for tunnel locations of massive ice, permafrost, boreholes, and encountered brine and cryopeg sediment.
[b] Temperature in tunnel recorded at tunnel floor; in sea ice, at sackhole depth.
[c] Size fractionation indicates whether the sample was collected directly on a 0.2 µm filter (> 0.2) or first prefiltered with a 3.0 µm filter (0.2–3.0); > 3.0 indicates the sample material collected on the prefilter.

Table 2.2. Concentrations of organic matter and nutrients for the subset of samples with sufficient volume.

#	Sample	POC (µM)	DOC (µM)	pEPS (µM C)	dEPS (µM C)	PO_4 (µM)	NO_3 (µM)	NO_2 (µM)	NH_4 (µM)
5	CBIW_17	1.98×10^{3}	3.00×10^{4}	1.44×10^{4}	8.55×10^{3}	1.94	5.56	2.03	1.75×10^{3}
6	CBIW_18	2.57×10^{3}	8.22×10^{4}	1.14×10^{2}	1.25×10^{4}	0.68	0.82	0.89	1.17×10^{3}
10	CB1_18	1.24×10^{4}	1.02×10^{5}	1.63×10^{3}	1.94×10^{4}	1.22	–[a]	16.2	3.35×10^{3}
14	CB4_18	4.15×10^{3}	8.50×10^{4}	8.91×10^{1}	1.98×10^{4}	0.60	13.6	2.96	4.52×10^{3}
15	SB_17	2.03×10^{1}	4.46×10^{2}	1.90×10^{0}	2.61×10^{2}	1.80	0.15	0.02	5.50×10^{-1}
18	SB_18	2.25×10^{1}	2.00×10^{2}	9.67×10^{-1}	–[a]	1.60	3.63	0.11	3.42×10^{0}

[a] Below detection

Table 2.3. Bacterial and viral abundances, percent dividing cells, and virus to bacteria ratio (VBR)

#	Sample	Cells (mL^{-1})	Dividing cells (%)	VLP (mL^{-1})	VBR
1	CB4_IW1_17	4.04×10^5	–[a]	–	–
2	CB4_IW2_17	4.08×10^5	–	–	–
4	CBIW_IW7_17	1.40×10^7	–	–	–
5	CBIW_17	1.39×10^8	–	1.22×10^8	0.9
6	CBIW_18	1.22×10^8	1.3	3.52×10^8	2.9
7	CBIW_re_18	1.18×10^8	1.1	3.89×10^8	3.3
8	CB1_0.2_09[b]	5.70×10^6	5.5[c]	5.70×10^7	10
10	CB1_18	9.57×10^7	1.9	1.24×10^8	1.3
11	CBIA_surf_18	1.70×10^8	0.0	–	–
12	CBIA_18	7.31×10^7	2.5	5.50×10^7	0.8
14	CB4_18	1.14×10^7	29	3.54×10^6	0.3
15	SB_17	2.22×10^5	–	$1.53 \times 10^{5,d}$	0.7
18	SB_18	1.11×10^5	4.7	$6.16 \times 10^{5,d}$	5.5

[a] Not determined

[b] Data from Colangelo-Lillis et al. (2016)

[c] Estimated from epifluorescent imagery in Colangelo-Lillis et al. (2016)

[d] Determined by the anodisc method of Nobel and Fuhrman (1998), as values were below detection limit of the wet-mount method of Cunningham *et al.* (2006) used for the higher counts.

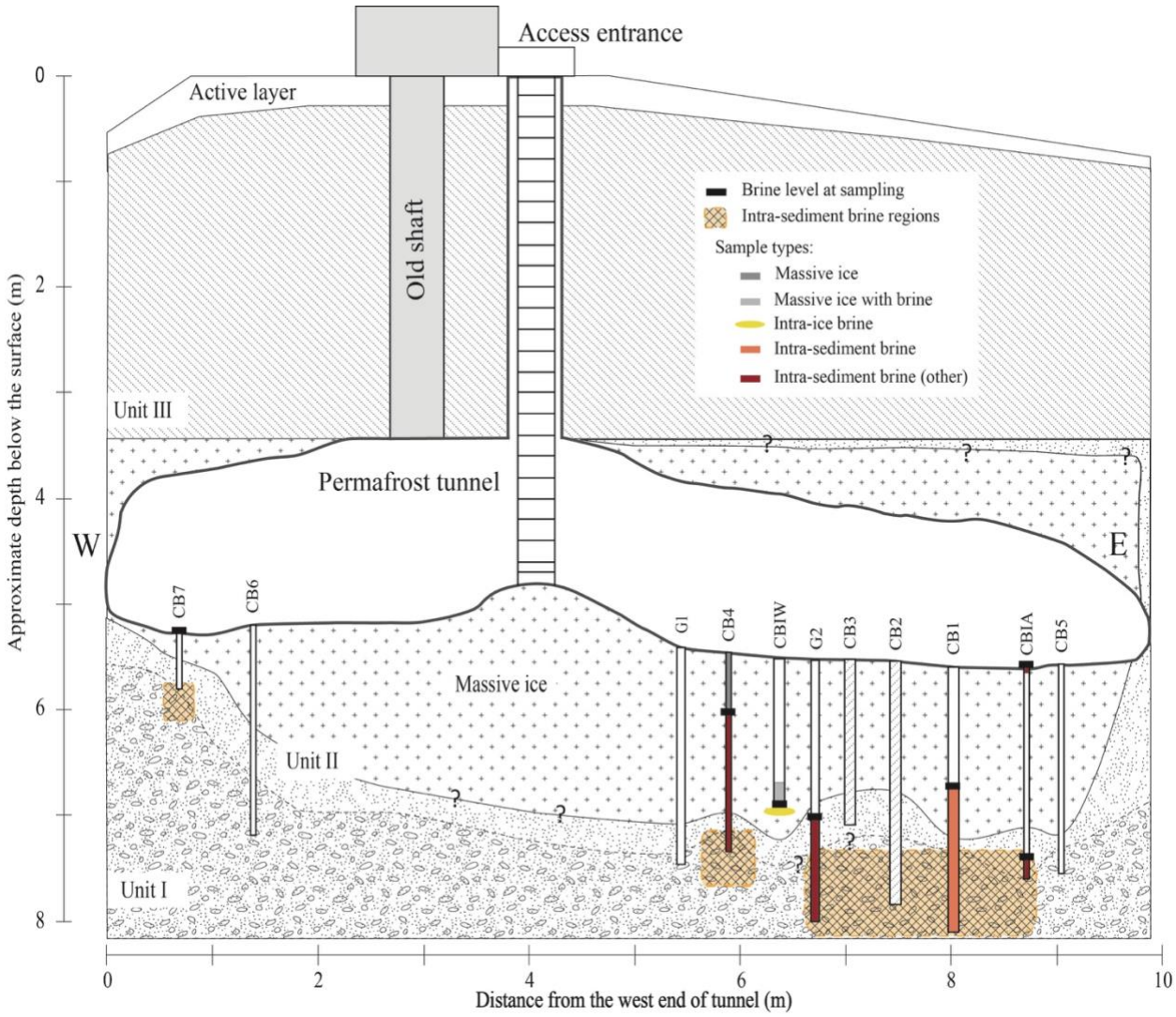

Figure 2.1. North-facing cross-sectional diagram of the Barrow Permafrost Tunnel system. The tunnel is entered via the access ladder, and equipment and supplies are lowered through the same space. Cryopeg boreholes CB1, CB2, CB3, and the upper (dry) portion of CBIW were drilled prior to this study; CB2 and CB3 were sealed and unavailable to this study. The other boreholes were drilled in 2017 and 2018, with pipes installed to preserve the holes and facilitate sampling. Crosshatched sections indicate confirmation of cryopeg sediments (intra-sediment brine regions). Units I–III refer to permafrost regions, following Meyer et al. (2010a). W and E indicate West and East directions, respectively. Question mark symbols "?" indicate unresolved boundary position. See Table 2.1 for sample details, and Text S2.1 for borehole terminology and drilling details.

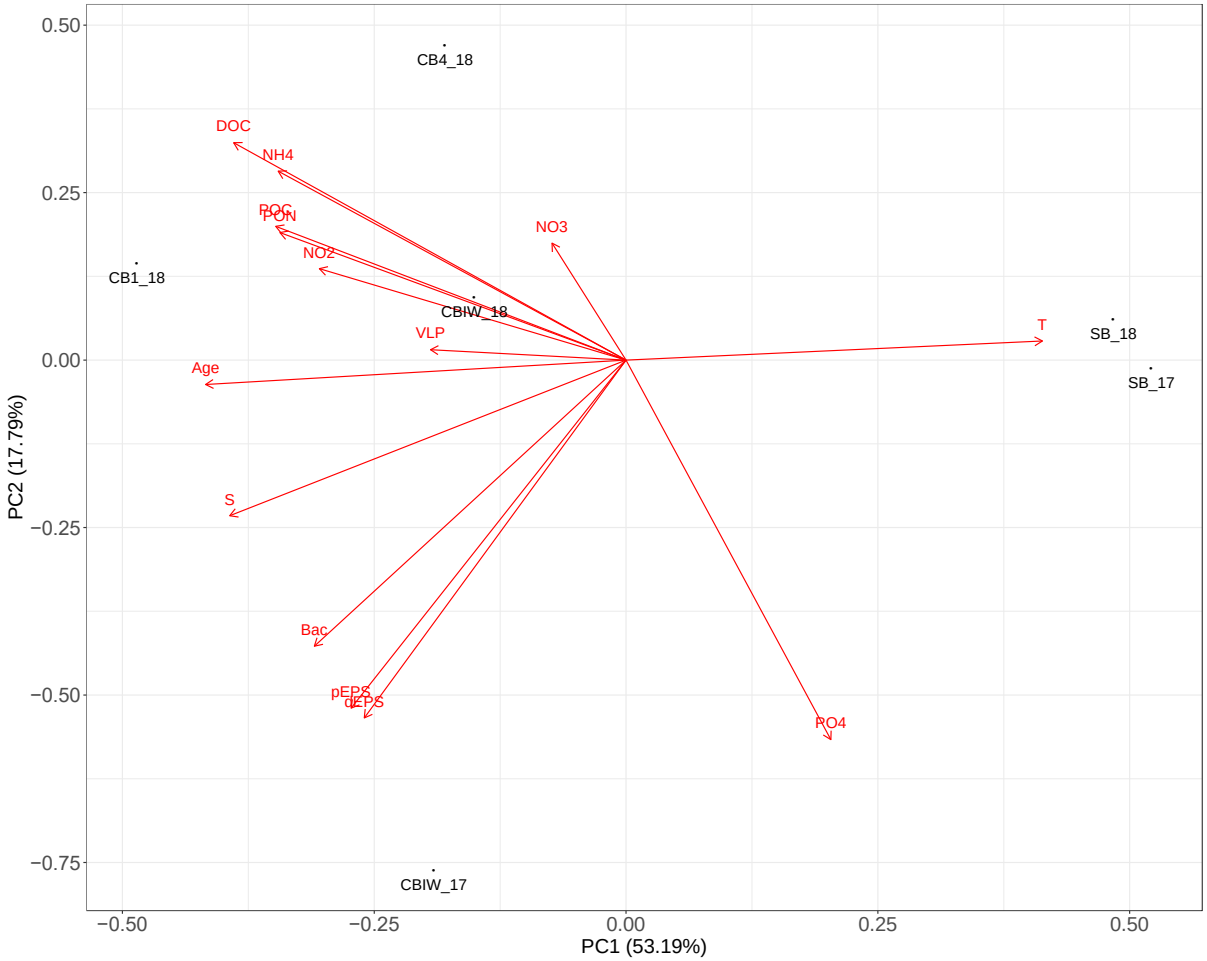

Figure 2.2. Principal component analysis of cryopeg and sea ice brine samples that had a full suite of measured environmental factors.
Vectors indicate contribution of each variable to each principal component.

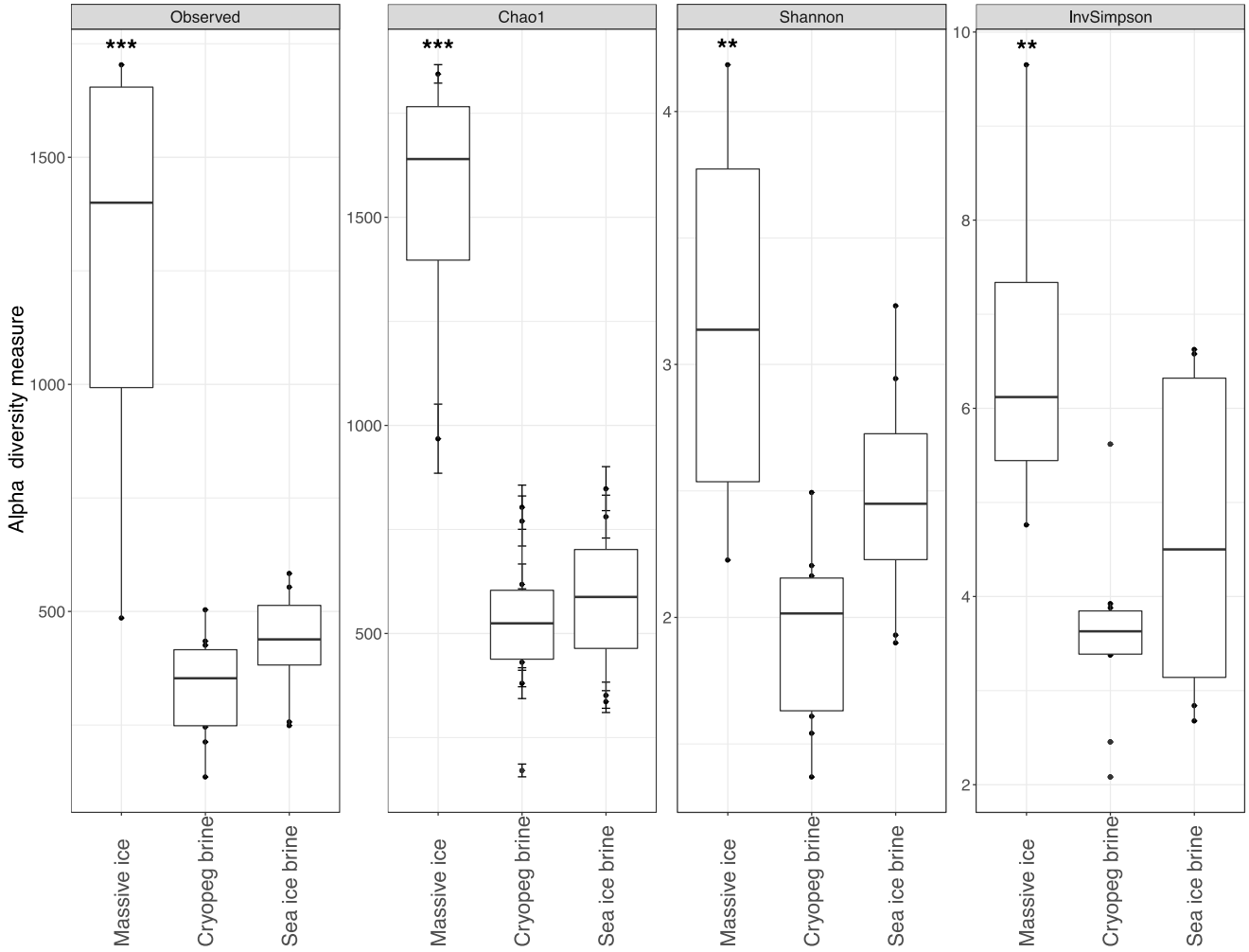

Figure 2.3. Indices of microbial community richness (left panels) and diversity (right panels) for the three main sample types: massive ice, cryopeg brine, and sea ice brine.
Massive ice samples (n = 4) differ significantly from cryopeg brines (n = 10) by every measure (** indicates $p < 0.01$ and *** indicates $p < 0.001$). Indices for sea ice brines (n = 8) tend to be higher than cryopeg brines, also by every measure, but not significantly (e.g. $p = 0.0796$ for Shannon diversity index).

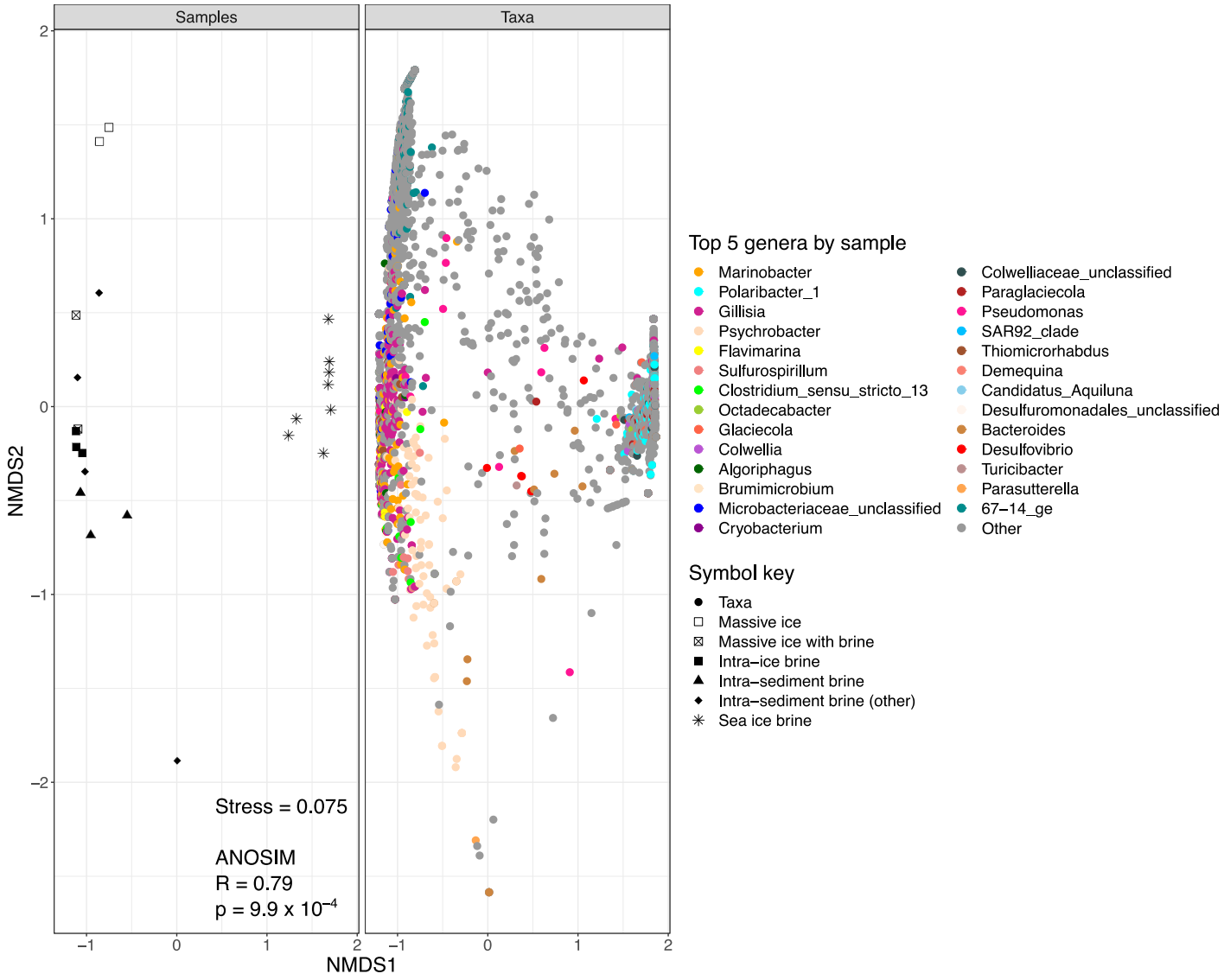

Figure 2.4. Split non-metric multidimensional scaling (NMDS) ordination, using Bray-Curtis dissimilarities, showing relatedness of each sample and OTU contributions (colored by genus) to sample ordination.
Left panel: black and white symbols indicate sample subtype; right panel: colors indicate taxonomy.

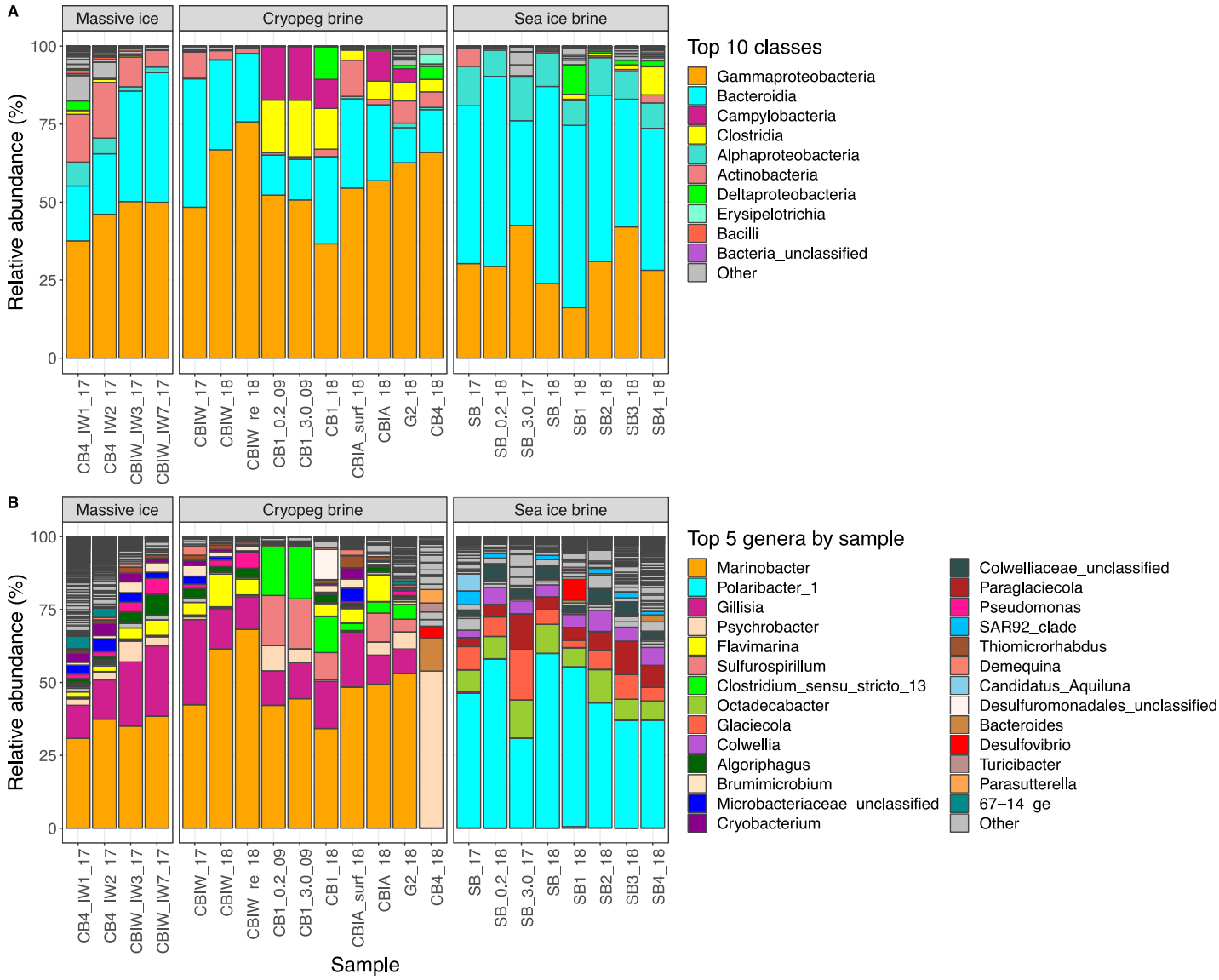

Figure 2.5. Relative abundance plots of taxa in each sample, grouped by the three main sample types.
A) The top ten classes across all samples are color-coded; lower abundance classes are in gray.
B) The top five genera from each sample are color-coded and shown in each sample where they occur; lower abundance genera are in gray. Sample designations are further described in Table 2.1.

Chapter 3. MOBILOME-DRIVEN EVOLUTION IN BACTERIAL POPULATIONS FROM SUBZERO BRINES

3.1 ABSTRACT

The cryosphere on Earth contains liquid environments, such as cryopeg (unfrozen marine-derived brine-containing layers in permafrost) and sea ice brines, that are characterized by multiple extremes, particularly subzero temperatures and hypersalinities. These physicochemical characteristics limit growth rates for microorganisms while presenting a setting that can be concentrated with organic and inorganic compounds along with biological entities. Evolution via vertical inheritance is challenging when generation times are long relative to the period of exposure to selective pressures experienced by an individual organism, and the physically concentrated setting of subzero brines can create competition for resources and an increased risk of viral predation. To explore the evolutionary processes that have allowed life to persist in subzero brines, we sampled stably extreme (over millennia) cryopeg and ephemeral (< 1 year) sea ice brines from the Alaskan Arctic coast, conducted metagenomic sequencing with a combination of Nanopore long- and Illumina short-read sequencing techniques, curated a collection of high-quality metagenome-assembled genomes (MAGs), and characterized the genomic microdiversity and evolutionary signatures of the *in situ* populations. We found that cryopeg brine MAGs have significantly ($p \leq 4.0 \times 10^{-3}$) fewer single nucleotide variants (SNVs) and lower base entropy per gene than sea ice brine MAGs, indicating a lower rate of vertically inherited variation in the cryopeg brine system. Cryopeg MAG genes associated with the mobilome (NCBI COG category X) have a significantly ($p \leq 1.5 \times 10^{-2}$) lower average pN/pS and number of SNVs per gene than the populations in sea ice. Specifically, transposase genes, which make up approximately 80% of the mobilome in populations from both environments,

have fewer ($p = 2.5 \times 10^{-4}$) SNVs per gene in cryopeg populations than sea ice. From a gene sharing analysis, we found that cryopeg MAGs share a significantly ($p = 5.6 \times 10^{-3}$) higher proportion of mobilome genes with their communities than observed for sea ice MAGs. These findings illustrate the relative importance of the mobilome to evolution in stably extreme environmental settings like cryopegs as a mechanism for introducing adaptive genotypes where vertical inheritance may be too slow relative to the duration of the extreme conditions.

3.2 INTRODUCTION

Prokaryotes have evolved to inhabit nearly every aqueous environment on Earth (Fenchel & Finlay, 2004), including those with extreme physical and chemical characteristics that push the molecular machinery of life to the limits of functionality (Harrison et al., 2013). Cold (< 5˚C) environments on Earth are particularly prevalent, and microorganisms are well equipped to survive if not thrive at low, even subzero, temperatures (Morita, 1975; Rivkina et al., 2000; Methé et al., 2005; Boetius et al., 2015; Nikrad et al., 2016; Rapp et al., 2021). In the marine environment, dissolved salts depress the freezing point of water, allowing it to remain liquid below 0˚C. When sea ice forms from seawater, salts concentrate further within the pore space of the ice, where liquid brines remain at temperatures much lower (< –20˚C) than the freezing point of seawater (approximately –1.9˚C; Cox & Weeks, 1983). Microorganisms inhabiting subzero, hypersaline brines must contend with depressed molecular reaction rates due to the temperature, and with cellular instability created by the osmolarity of the environment (Nunn et al., 2015; Firth et al., 2016; Deming & Collins, 2017; Czajka et al., 2018). The combination of these conditions results in long generation times which make the rate of evolution via vertical inheritance quite slow relative to the duration of selective pressures are experienced (Schloter et al., 2000; Elena and Lenski, 2003; Li et al., 2014; Bendall et al., 2016; Moulana et al., 2020). Life under multiple extremes is challenging, and the microorganisms that are endemic to subzero brines must contend with these challenges to prevail.

The sea ice environment is vast, covering nearly 10% of the surface area of the global ocean, and it hosts primary producers that compose the base of an enormous ecosystem supporting much of the life in polar oceans (Boetius et al., 2015). Sea ice is porous, with brine channels and pockets

characterizing the interior space (Golden et al., 2007; Petrich & Eicken, 2017); the volume and salinity of this liquid fraction can change rapidly as a function of atmospheric temperature on a scale of hours to weeks in time (Ewert & Deming, 2014). Most sea ice melts within a year of formation, and the brine contents of the ice are returned to the surface layer of the ocean. Despite the ephemeral nature of this environment, sea ice brines are inhabited by distinct communities of bacteria, archaea, and their viruses (Bowman, 2015; Deming & Collins, 2016; Cooper et al., 2019, Chapter 2) among other microscopic organisms, particularly the microalgae (Torstensson et al., 2015; Dawson et al., 2020). Sea ice bacteria are known to employ motility to actively seek and inhabit the interior brines of sea ice (Showalter & Deming, 2018), where they cycle organic matter at subzero temperatures (Vancoppenolle et al., 2013; Showalter & Deming, 2021). The freezing process that creates elevated nutrient concentrations (relative to source seawater) makes this environment attractive to microorganisms, but it also increases the concentration of cells and their contact rate with viruses which can lead to increased resource competition and chances of infection (Wells & Deming, 2006b; Collins & Deming, 2011b; Showalter, 2020). To combat the challenges of life under these extremes, sea ice microorganisms produce extracellular polysaccharides to favorably alter their environment (Krembs et al., 2011), adjust their lipids to maintain membrane fluidity (Morita, 1975; Methé et al., 2005; Casanueva et al., 2010), transport compatible solutes for osmotolerance (Firth et al., 2016; Torstensson et al., 2019), and combat infection from viruses using membrane modifications and by producing EPS to act as a diffusive barrier (Wells & Deming, 2006a; Luhtanen et al., 2018; Showalter, 2020; Zhong et al., 2020).

Cryopegs are unfrozen layers of permafrost that can be derived from seawater-saturated marine sediment and that are isolated from the ocean, atmosphere, and other external influences

(Gilichinsky et al., 2003; Streletskaya & Leibman 2003; Iwahana et al., 2021). They have been identified in permafrost around the coastal regions of the Arctic (Gilichinsky et al., 2003; Streletskaya & Leibman 2003; Yoshikawa et al., 2004; Pecheritsyna et al., 2007; Spirina et al., 2017; Iizuka et al., 2019; Iwahana et al., 2021). We studied a cryopeg system located in Northern Alaska, near Utqiaġvik, that formed at least 40 ka BP and that has been isolated from external influences, maintaining subzero hypersaline conditions, since formation (Meyer et al., 2010a; 2010b; Colangelo-Lillis et al., 2016; Cooper et al., 2019, Chapter 2; Iwahana et al., 2021). We discovered that these cryopeg brines are home to low-diversity microbial communities (relative to communities in sea ice brines), composed primarily of bacteria and viruses at remarkable densities (approximately 10^8 cells and viruses mL^{-1}; Cooper et al., 2019, Chapter 2) and well-adapted to their environment (Zhong et al., 2020; Rapp et al., 2021). Some cryopeg brine pools in this system have been entrained entirely within a massive ice formation, completely isolated for approximately 11 ka BP (Iwahana et al., 2021), with a resident bacterial community that today is dense and presumably active (Rapp et al., 2021; Appendix C). Low temperatures, high salt concentrations, and high-density communities with increased potential for resource competition and viral infection are challenges faced by the microbial communities inhabiting both cryopeg and sea ice brines. Despite these similar challenges, the communities that inhabit these spaces are taxonomically (Cooper et al., 2019, Chapter 2) and metabolically distinct (Rapp et al., 2021), and we hypothesize that these dissimilarities may be driven by the difference in duration and frequency of exposure to subzero hypersaline conditions.

Simultaneously contending with an extreme environment, competing for space and resources, and defending against viral infection would challenge any living being. Natural selection allows

the most fit member of a population (a group of very closely related organisms, usually the same species, that co-occur in an environment) to outcompete others (Darwin, 1859), and this process is tightly controlled and highly competitive when stressful conditions are persistent, as adaptations can be metabolically expensive (Methé et al., 2005; Czajka et al., 2018; Rapp et al., 2021). Bacterial evolution via vertical inheritance begins as a mother cell replicates its genome into two copies that each become the genome of a daughter cell. The process of genome replication can create errors at the level of single nucleotides, resulting in single nucleotide variants (SNVs) that increase microdiversity (nucleotide-level genomic variability) within a population of related bacteria (Bendall et al., 2016; Anderson et al., 2017). An SNV may or may not result in a change to a codon that translates to a different amino acid than originally encoded. An amino acid change that affects protein function can provide an adaptive advantage or be deleterious to the competitive functioning of the cell (Schloissnig et al., 2013; Delmont et al., 2019). Because SNVs are introduced stochastically, the result of their introduction is more likely to have no effect or a deleterious effect than to be beneficial for the cell. Thus, acquiring adaptation via vertical inheritance can be quite slow.

An alternative mechanism for evolution that is widespread within the Prokaryotes is horizontal gene transfer (HGT), where an entire gene can be transferred between genomes at once. HGT can occur via several processes, including viral transduction, conjugation between cells, and natural transformation where extracellular genes in the environment (e.g., those released after cell lysis) are taken up by a cell before they degrade (Galtier, 2007; Gillings, 2017). With the introduction of a new gene, a new protein can be tested by selection in the next generation, rather than the countless generations required for favorable SNVs to accrue and result in a

phenotypically novel protein (Collins & Deming, 2013; Li et al., 2014; Fuchsman et al., 2017; Good et al., 2017). The genes that allow the mobility of genes between genomes make up the mobilome (Toussaint & Chandler, 2012). Mobilome genes are involved in multiple modes of genome rearrangement, including the transposition of genes between locations on a genome (i.e., by transposases; McClintock, 1987; Blot, 1994) and the integration of new genetic material (i.e., by integrases; Mazel, 2006). These genes provide an important utility for organisms that can benefit from quickly acquiring novel genotypes, and they themselves can serve as signatures of gene mobility in a population. We hypothesize that genomic flexibility, via the mobilome, plays a significant role in the evolution of microorganisms inhabiting extreme environments where conditions slow the rate of vertical evolution, such as subzero brines, and particularly where those conditions are persistent over millennia.

We had previously collected samples of sea ice and cryopeg brines from the Alaskan Arctic coast, near Utqiaġvik, to investigate the diversity and genomic potential of communities inhabiting subzero brines (Cooper et al., 2019, Chapter 2). Using these samples, we also characterized the geological history and physical processes that define cryopeg brines in the region (Iwahana et al., 2021) and uncovered the biodiversity and functional potential of microbial communities in cryopeg and sea ice brines, identifying genomic signatures of adaptation to the subzero brine environment in both bacteria and viruses (Zhong et al., 2020; Rapp et al., 2021; Cooper et al., Chapter 3). Here, we expand this work by conducting additional metagenomic sequencing that improves the recoverability of genomes from the environmental samples to better represent *in situ* bacterial populations using metagenome assembled genomes (MAGs), and we use these data to examine patterns of microdiversity and gene sharing in both

cryopeg and sea ice brine communities. Although this study focuses on bacterial populations in subzero marine brines, I suggest that the lessons on evolution in these environments characterized by multiple extremes are broadly applicable to a global understanding of the evolution of life at its physical and chemical limits.

3.3 Materials and Methods

Collection of subzero brines

Cryopeg and sea ice brines were collected near Utqiaġvik, Alaska, USA in May 2017 and May 2018 as part of an interdisciplinary project aimed to understand the role of gene exchange for microbial life in subzero brine environments. Cryopeg brines were collected from boreholes in the Barrow Permafrost Tunnel (71.2944 °N, 156.7153 °W; see Figure 2.1 in Cooper et al., 2019, Chapter 2, for a diagram of the tunnel and borehole locations). The tunnel entrance is located in the tundra, approximately 2 km from the modern Arctic coastline. The tunnel is approximately 6 m below the ground surface, originally excavated by the United States Army Cold Regions Research and Engineering Laboratory in the early 1960s. It has been used for several decades of research on local geocryology (Yoshikawa et al., 2004; Meyer et al., 2010a; 2010b; Iizuka et al., 2019; Iwahana et al., 2021) and, more recently, microbiology (Colangelo-Lillis et al., 2016; Cooper et al., 2019, Chapter 2; Zhong et al., 2020; Rapp et al., 2021; Cooper et al., Chapter 3). The tunnel is narrow (~ 1 m) with low ceilings (~ 1.5 m) and the ambient air temperature during sampling was –6˚C. Despite these operational challenges, we sampled cryopeg brines as cleanly as possible using a hand-pumped vacuum apparatus with acid- and ethanol-washed (between each sampling) tubing collecting into a sterile flask. Ethanol-cleaned nitrile gloves were worn during all biological sampling. Cryopeg brine samples were returned to the Barrow Arctic

Research Center (BARC) after sampling where they were kept in cold rooms set to –6˚C and processed within 2–8 hours of collection.

Sea ice brines were collected from landfast sea ice located at 71.3730 °N, 156.5047 °W in May 2017 and 71.4730 °N, 156.7294 °W in May 2018 near the Barrow Sea Ice Mass Balance site operated by the University of Alaska Fairbanks (Druckenmiller et al., 2009). Sea ice brines were collected directly by drilling sackholes (boreholes drilled partially through the depth of the ice) and leaving them to fill while covered for a period of 3–5 hours. Ice thickness varied between 2017 and 2018 with depths of 117 and 110 cm (sackholes were drilled to 75 and 55 cm), respectively. Air temperature ranged from –4 to –6˚C both years, and snow depth was 16–19 cm in 2017 and 6–10 cm in 2018. Upper sea ice temperatures (above sackhole depth) were –4˚C in 2017 and –3˚C in 2018. Sea ice brines were processed at the BARC within 8 hours of collection.

Volumes of cryopeg brine collected for DNA extraction ranged from 25 to 500 mL, depending on brine volume recovered during sampling. For sea ice brines, 2500 mL of pooled brine from multiple sackholes were collected each year. All samples were filtered through 0.22 µm filters (47-mm in-line GTTP Isopore filters with a 1.6-µm Whatman GF/A or a 3.0 µm polycarbonate prefilter or without a prefilter through a Sterivex filter [all filters from Merck-Millipore, Burlington, MA]) in a BARC cold room at 4˚C. Filters were frozen at –80˚C and shipped to the University of Washington in Seattle for further processing. Subsamples were also collected for ancillary measurements including salinity, pH, and organic and inorganic nutrients. These data have been reported previously and are available in Table 2.2 of Cooper et al. (2019, Chapter 2).

Metagenome sequencing, assembly, and binning

DNA from cryopeg brine (n = 4) and sea ice brine (n = 2) samples, concentrated onto 0.22 µm Sterivex filters in the field and kept frozen at –80°C until processing, was extracted using the DNeasy PowerSoil kit (QIAGEN, Germantown, MD) as previously described (Cooper et al., 2019, Chapter 2). Samples collected for DNA extraction for this study are identical to samples CBIW_17, CBIW_18, CBIA_18, CB1_18, SB_17, and SB_18 described by Cooper et al. (2019, Chapter 2). Samples CBIA_18 and CB1_18 are referred to simply as CBIA and CB1 in this study as parallel samples from 2017 were not available for these analyses. Metagenomic sequencing was conducted using both Illumina short (2 x 150 bp) and ONT long reads. The Illumina sequencing library was prepared, sequenced, and analyzed at the DOE Joint Genome Institute as previously described (Rapp et al., 2021), and the ONT sequencing library was prepared, sequenced, and quality-controlled following the VirION2 pipeline (Zablocki et al., 2021). Long reads were first assembled using Flye v2.5 (Kolmogorov et al., 2019), and short reads were subsequently used for error correction using Pilon v1.23 (Walker et al., 2014), to obtain high-quality metagenome assemblies. Each metagenome was assembled from a single sample of cryopeg brine, allowing for recovery of MAGs from discrete samples. The raw Illumina metagenomic reads were produced previously and are described in detail by Rapp et al. (2021); they are available through IMG/M under accession numbers 3300031836 (CBIW17), 3300032129 (CBIW18), 3300032135 (CBIA), and 3300034171 (CB1). The hybrid metagenomic assemblies generated here will be available publicly after journal submission and peer review.

Assembled metagenomic contigs were processed using anvi'o v6.1 (Eren et al., 2015a). Gene calling was performed using Prodigal v2.6.3 (Hyatt et al., 2010) to identify open reading frames (ORFs). Each ORF was annotated with National Center for Biotechnology Information (NCBI) Clusters of Orthologous Genes (COG) annotations (Tatusov et al., 1997; Galperin et al., 2019) which were assigned using DIAMOND v0.9.22.123 (Buchfink et al., 2014). Bacterial and archaeal single copy universal marker genes (Lee, 2019) were identified using HMMER v3.3 (Eddy, 2011). Automated binning was conducted using MetaBAT 2.0 v2.15 (Kang et al., 2019), MaxBin2 v2.2.7 (Wu et al., 2016), and Concoct v1.1.0 (Alneberg et al., 2014). The best bins from each tool were selected automatically using DAS Tool v1.1.2 (Sieber et al., 2018). After binning, the bins were refined manually in anvi'o using coverage, gene clustering, and GC content as guides. After binning, each MAG was assigned taxonomy using GTDB-tk v1.1.1 (Parks et al., 2018), and MAG completeness was estimated using CheckM v1.1.2 (Parks et al., 2015). Abbreviated IDs that reference taxonomy and the corresponding original sample name are used to refer to the MAGS in this study (Table 3.1). The MAG sequences will be available publicly after journal submission and peer review.

Microdiversity analyses

Microdiversity metrics were calculated for each MAG using anvi'o v6.1 (Eren et al., 2015a). We used the anvi'o command "anvi-gen-variability-profile" to calculate single nucleotide variants (SNVs), single amino acid variants (SAAVs), and single codon variants (SCVs) for each MAG. The SNV calculation returned a Shannon entropy value, H, for each base position in each MAG, providing information on the distribution of bases likely to be found at any position (Eren et al.,

2015b). Values of H = 0 indicate that a base position is always occupied by a single nucleotide, while a value of H = 1 means that there is an equal probability of any nucleotide being present at a base position. For further SNV calculations, we required a coverage of at least 10X and a minimum departure from the consensus base of 0.05 for an SNV to be counted, following Anderson et al. (2017). SAAVs and SCVs were used to calculate pN/pS for each gene call in each MAG using the anvi'o command "anvi-script-calculate-pn-ps-ratio". The pN/pS value for each gene is the ratio of nonsynonymous to synonymous polymorphisms present for each codon in a gene (Schloissnig et al., 2013). Ratios of pN/pS that equal 1 indicate neutral genetic drift, values greater than 1 indicate Darwinian evolution (accumulation of novel and competitive phenotypes) allowing for the production of new genes with new protein sequences, and values approaching 0 indicate clonal selection for genotypes that do not introduce new variation. Mean entropy, mean pN/pS, and the number of SNVs were calculated for each gene using a custom python script. All calculations and scripts will be available publicly after journal submission and peer review.

Shared gene analyses

How many and which genes are shared between each MAG and each metagenome was assessed using DIAMOND v2.0.11.149 (Buchfink et al., 2014). Two custom databases were produced using the DIAMOND "makedb" program. The first database consists of all of the protein sequences for each gene call from each MAG (n = 24). A second database was produced using all of the protein sequences for each gene call in each metagenome (n = 6), which includes the sequences for each MAG. The "blastp" method was used in DIAMOND, where each MAG was run as a query against the two databases described above. For a hit to count, a minimum E-value

of 1 x 10^{-10}, a minimum query coverage of 50%, and a minimum identity of 80% were required (Fuchsman et al., 2017). Custom python scripts were used to associate DIAMOND hits with gene calls and to estimate the relationships between MAGs and between MAGs and metagenomes. The number of genes (represented as a proportion of total genes in each MAG) found between MAGs and each other MAG or metagenome was used to create a network using the "networkx" (Hagberg et al., 2008) package in python. Similarity of the MAGs in either case was assessed using PCA in the scikit learn (Pedgregosa et al., 2011) python package. All network/ordination plots were created using matplotlib (Hunter, 2007) and seaborn (Waskom, 2021) visualization packages in python. All calculations and scripts will be available publicly after journal submission and peer review.

3.4 Results and Discussion

Description of MAGs generated

Using cutting-edge sequencing and binning technology, we generated a collection of 24 high-quality MAGs from individually assembled metagenome samples (Table 3.1). The number of contigs comprising each MAG ranges from 6 to 151 (mean ± standard deviation = 48 ± 31), and the N50 (weighted median of contig lengths) for each MAG ranges from 24,409 to 2,386,693 bp (mean = 231,318 ± 464,145 bp). Mean coverage for each MAG ranged from 21 to 7,351X (mean = 1,368 ± 2,213X). Statistics and additional information for each MAG can be found in Table 3.1. The quantity and quality of MAGs generated here highlights the benefits of new methods for combining short-read and long-read metagenomic sequences from individual samples (Walker et al., 2014; Kolmogorov et al., 2019; Zablocki et al., 2021) and of the utility of using multiple automated binning techniques (Alneberg et al., 2014; Wu et al., 2016; Kang et al., 2019) with

additional manual refinement (Sieber et al., 2018). Given the nature of their production, these MAGs are each representative of the *in situ* population of bacteria inhabiting a discrete sample of cryopeg or sea ice brine (Table 3.1). The nucleotide variation uncovered by the depth of coverage for each MAG allowed us to investigate microdiversity within the *in situ* population.

The MAGs in this study belong to taxa that were identified previously in these samples, using 16S rRNA gene amplicon sequencing (Cooper et al., 2019, Chapter 2), and represent both the most abundant and some less abundant of these taxa (Table 3.1). Four of the cryopeg brine MAGs, each from a different field sample, were identified as *Marinobacter*, the most abundant genus previously identified in these brines. All four of the MAGs have identical 16S rRNA gene sequences (matching the V3–V4 region; Cooper et al., in prep, Chapter 4). Other abundant taxa represented in the cryopeg brine MAG collection include *Leeuwenhoekiella* (n = 3; based on GTDB taxonomy; identified as *Flavimarina* based on NCBI taxonomy and by Cooper et al., 2019, Chapter 2) and *Sulfurospirillum* (n = 1). For sea ice brines. the dominant genus identified by Cooper et al. (2019, Chapter 2) was *Polaribacter*, and two of the MAGs, one from each of the two sea ice metagenomes, are also *Polaribacter*. Other abundant taxa from the sea ice brine MAG collection include *Glaciecola* (n = 2) and *Octadecabacter* (n = 1). The remaining taxa represented in the overall MAG collection are variably less abundant, but still represent a spread of members of the communities identified by amplicon sequencing. With MAGs representing taxa across a range of abundances for both cryopeg and sea ice brines, we considered that the MAGs well-represent these distinctive *in situ* communities, though a limitation to consider is our inability to consider the total community assemblage using the available data. Although each MAG is expected to have evolutionary signatures that align with their lineage, we hypothesized

that patterns of microdiversity across MAGs of variable abundances would align with the environment.

Microdiversity signatures of evolution

Microdiversity provides information on the amount of variation that occurs in the genomes of population members at a nucleotide level. Genome replication can create random errors, some of which may alter the function of a gene. If the new function is beneficial, then it may be selected for during propagation; if the new function is deleterious or simply not as efficient, then it may be selected against, causing it to disappear from a population. A mutation that has no effect on the resultant protein results in neutral genetic drift. Each of these pathways leaves an observable signature in a genome that can be used to understand the evolutionary history of a population. We assessed levels of microdiversity in each MAG in this study by counting single nucleotide, codon, and amino acid variants (SNVs/SCVs/SAAVs). We used these data to calculate the amount of variation at any position as Shannon entropy, H (Eren et al., 2015b), and the level of selection on a gene using pN/pS (Schloissnig et al., 2013). Entropy describes the level of single nucleotide variability at a base position, with H values closer to 0 indicating little deviation from the consensus base and H values closer to 1 indicating an even distribution of all nucleotides. Higher entropy indicates a scenario where SNVs are accumulating, while lower entropy indicates selection against new variation. The pN/pS value describes the ratio of nonsynonymous to synonymous polymorphisms in a gene (Schloissnig et al., 2013). When a base is changed, it alters the codon and becomes an SCV. Multiple codons can encode for the same amino acid, so an SCV does not always result in an SAAV. When the codon change does not result in an SAAV, then that is counted as a synonymous polymorphism, but when it does generate an

SAAV, the polymorphism is nonsynonymous. A pN/pS value less than 1 indicates clonal selection for a gene (very little functional variation), while a value greater than 1 indicates that Darwinian evolution has occurred (multiple functional variants are present and selected for). A pN/pS value of 1 indicates that genetic drift is occurring (new variants are being produced but no selective pressure is acting). Given the nature of stochastic mutations, a pN/pS value much less than 1 will likely be associated with lower entropy, and *vice versa*. Diverging from this general rule is the possibility to find lower entropy values associated with pN/pS greater than 1 if the diversity that is present is selected for and beneficial for a population, but finding higher entropy values associated with pN/pS values much less than 1 is mathematically unlikely given the probability of highly entropic mutations resulting in a nonsynonymous codon change.

Considering all genes in each MAG (Figure 3.1), we observed a significantly ($p = 1.2 \times 10^{-3}$) higher proportion of SNVs per kb in MAGs from sea ice brines (72.9 SNVs per kb) than in cryopeg brine MAGs (38.1 SNVs per kb). We also observed a significantly ($p = 4.0 \times 10^{-3}$) greater mean entropy in sea ice brine MAGs ($H = 0.056$) than in cryopeg brine MAGs ($H = 0.033$). Not all genes in a genome are necessarily functional nor do they necessarily play a role in competition or adaptation, but observations of patterns in a whole genome can be informative of general trends in error tolerance during replication. The lower SNV density and entropy observed for cryopeg MAGs indicates that errors have generally been less tolerated and that the functions of some genes in each genome are important enough for their sequences to be maintained. Measurements of mean pN/pS for all genes in MAGs from cryopeg brines (pN/pS = 1.04) and sea ice brines (pN/pS = 1.13) were not significantly different ($p = 0.241$), but whole genome values near 1 for both environments suggest that the populations are generally neutrally evolving

and well adapted to their environments. Although individual functional genes may undergo selection separately from the entire genome (Lan & Reeves, 2000; Fuchsman et al., 2017; Moulana et al., 2020), these results provide a clear signal that SNV density is generally greater in MAGs from sea ice brines than cryopeg brines, even as both communities appear to be undergoing neutral evolution in their environments. The persistent long-term stressors of the cryopeg brine present a scenario where specific functions must be met, and minor errors may be unforgiving for mutations that do not provide a benefit.

Subzero brine environments that are microbially dense present many opportunities for interactions that might result in gene flow (Wells & Deming, 2006b; Pecheritsyna et al., 2007; Collins & Deming, 2011a,b; Murray et al., 2012; Collins & Deming, 2013; Boetius et al., 2015; Cooper et al., 2019, Chapter 2; Zhong et al., 2020; Rapp et al., 2021). Genes involved in genome rearrangements and genetic transfers constitute the mobilome, which includes genes for transposases, integrases, and recombinases, as well as virus-related genes (McClintock, 1987; Blot, 1994; Mazel, 2006; Toussaint & Chandler, 2012). Because mobilome genes are generally retained in the genome in proximity to the new genes they introduced, the microdiversity of these genes can be used to understand the evolutionary forces driving these mechanisms and to observe the selection of the genes they move. We found that mobilome genes in MAGs (Figure 3.2) from cryopeg brines on average have significantly fewer SNVs per kb (35.4 vs. 96.6, $p = 2.0 \times 10^{-4}$), lower mean entropy (0.041 vs. 0.074, $p = 5.5 \times 10^{-2}$), and lower pN/pS (0.822 vs. 1.06, $p = 1.5 \times 10^{-2}$) than in MAGs from sea ice brines. Following the pattern for the whole MAGs, the mobilome genes from cryopeg brine MAGs have lower SNV density, which indicates that specific gene sequences have been conserved, and the lower pN/pS of the mobilome genes in

these MAGs suggests that clonal selection is ongoing. In both cryopeg and sea ice brine MAGS, approximately 80% of the genes in the mobilome are transposases (Table 3.2; 3.3). Only two MAGs (representing one low-abundance genus from each environment) contained no identifiable transposase genes, while the number of transposases in all other MAGs ranged from 2 to 78. For these transposase genes (Figure 3.3), we found that cryopeg brine MAGs had significantly fewer SNVs per kb (31.1 vs. 78.3, $p = 2.5 \times 10^{-4}$). Though the proportion of transposases in each MAG is similar, the lower SNV density observed for these genes in cryopeg MAGs provides further evidence of selection for a specific functionality of the mobilome in the evolution of cryopeg populations. Additionally, cryopeg MAG integrase genes have lower SNVs per kb (27 vs 103, $p = 1.8 \times 10^{-4}$), though they are less abundant than transposases in the mobilomes of the 15 MAGs where they are present (Table 3.3). This evidence supports our hypothesis that the mobilome, and in turn gene flow, is an important mechanism for the evolution of life in the persistently stressful and microbially dense environment of the cryopeg brines we sampled.

Gene sharing networks in subzero brines

We checked each MAG for genes that it shares with other MAGs and with whole metagenomes to look for signatures of gene flow and conservation. Using a reciprocal best hit method, we searched for genes that were at least 80% identical and covered at least 50% of the length of the query to identify related genes in each MAG (Fuchsman et al., 2017). At this level of identity, we were searching for genes that may be indicative of a shared evolutionary history but not necessarily exact copies of genes. Searches of genes in each MAG against genes in all other MAGs depict an image of gene sharing between populations in both cryopeg and sea ice brines (Figure 3.4). We observed no clear distinction of genes shared between MAGs based on

environment. The spatial separation of MAGs in the PCA ordination of Figure 3.4 instead more closely follows taxonomic relationships; e.g., Gammaproteobacteria such as *Marinobacter* and *Glaciecola* appear on the right of PC1 and Flavobacteriia such as *Polaribacter* and *Leeuwenhoekiella* on the left. Interestingly, we observed that taxa that are abundant in their respective environments appear to share very few genes with each other. *Marinobacter* and *Leeuwenhoekiella* are both abundant in cryopeg brines (Cooper et al., 2019, Chapter 2), yet *Marinobacter* shares fewer genes with *Leeuwenhoekiella* than with any other populations ($p = 3.8 \times 10^{-4}$). Similarly, *Polaribacter* and *Glaciecola* are abundant in sea ice brines (Cooper et al., 2019, Chapter 2), yet *Polaribacter* also shares fewer genes with *Glaciecola* than with any other populations ($p = 5.3 \times 10^{-2}$). As *Marinobacter* and *Polaribacter* are each the most abundant genus in their respective environments, this lack of gene sharing with direct competitors may be a signal of either genotypic competition or cooperation (Turner et al., 1996; Germerodt et al., 2016; Nadell et al., 2016). By having a set of genes distinct from their competitors, abundant taxa in these environments likely occupy unique niches where they flourish that enable their numerical dominance. By this logic, it would be advantageous for abundant bacteria to obtain genes that are unique from their competitors, perhaps acquiring them from bacteria that are less abundant in their environment.

The nature of the generation of MAGs using presently available technologies means that they cannot fully represent the total diversity present in sea ice and cryopeg brine communities (Cooper et al., 2019, Chapter 2). To obtain a more complete view of genes that are shared between populations (represented by each MAG) and other members of their community, we performed an additional shared gene analysis between each MAG and each entire metagenome

used in this study. A principal component analysis paired with a network analysis depicts the network of genes shared between each MAG and each community (Figure 3.5). This depiction shows a clear separation of each MAG based on environment, as expected, though some signals of genes shared between both environments are evident. We found a greater mean proportion of genes from (sea ice-dominant) *Polaribacter* in cryopeg brines than of genes from (cryopeg-dominant) *Marinobacter* in sea ice brines (0.1% vs 0.05% of total genes from the corresponding genome that are shared, $p = 1.2 \times 10^{-4}$). This difference could indicate that cryopeg brine bacteria use some of the same genes required for adaptation to the sea ice environment but that there are additional adaptations to the cryopeg brine environment that are not observed in sea ice. No clear difference was observed between general patterns of gene sharing between environments, but the proportion of genes shared between MAGs and all environments show a clear distinction exclusively for the mobilome. By checking the proportion of total genes belonging to each NCBI COG category that are shared between each MAG and each metagenome, we found that the proportion belonging to the mobilome (COG category X) was significantly higher for cryopeg than for sea ice brine (average of 16.0% of all shared genes for each cryopeg MAG vs. 3.6% for sea ice brine, $p = 5.6 \times 10^{-3}$). No proportion of shared genes belonging to any other COG category differed significantly between the two environments. Of the mobilome genes, the most abundantly shared genes are for transposases, which are once again shared significantly more between cryopeg brine MAGs and the metagenomes than for sea ice brine MAGs (14.4% vs. 3.25% of total genes in each MAG, $p = 4.7 \times 10^{-3}$). These data reinforce the importance of the mobilome in cryopeg MAGs observed in the microdiversity analyses. They also support the hypothesis that populations in cryopeg brines use the mobilome as an avenue for evolution where

environmental stressors are constant and vertical inheritance may be insufficient for adaptation due to long generation times relative to rates of selection.

Population evolution in cryopeg and sea ice brines

Both sea ice and cryopeg brines are extreme environments that present multiple avenues of stress for the endemic microbial communities. Osmotic pressure of a highly saline environment provides extreme stress on the membrane of a cell, and the difficulty of moderating membrane fluidity is intensified as low temperatures decrease molecular reaction rates and membrane fluidity, making survival challenging (Nunn et al., 2015; Deming & Collins, 2016; Firth et al., 2016; Czajka et al., 2018). While combating these forces, subzero brine inhabitants must simultaneously contend with high contact rates with viruses (Wells & Deming, 2006b; Collins & Deming, 2011b) and other bacteria in competition for resources (Cooper et al., 2019, Chapter 2; Rapp et al., 2021). Previous work has identified that the density of viruses and the quantity of free DNA in sea ice brines is conducive to horizontal gene transfer (Collins & Deming, 2010a,b), and this hypothesis has been further supported by evidence of horizontally transferred genes that play a role in osmotolerance in the genomes of sea ice bacteria (Collins & Deming, 2013). Sea ice is geophysically dynamic, fluctuating regularly in temperature and salinity (Ewert & Deming, 2014; Petrich & Eicken, 2017), and the inhabitants must repopulate the ice each year as it melts extensively in the spring and reforms from seawater in the autumn (Grossman & Dieckmann, 1994; Eronen-Rasimus et al., 2014; Collins & Deming, 2016; Showalter & Deming, 2018). These fluctuations provide both extreme stress and regular opportunities for moderate conditions where growth and evolution can proceed more rapidly. In contrast, cryopeg brines have remained below freezing and at high salinities over millennia (Iwahana et al., 2021), and the resident

community gets no reprieve of moderation from the stressors of subzero brines. The dominant genus in the cryopeg brines that we have studied is *Marinobacter* (Cooper et al., 2019, Chapter 2), and we have found that the species endemic to these brines contains tremendous genetic diversity and has undergone substantial horizontal gene transfer (Cooper et al., Chapter 4). To survive in the microbially dense cryopeg brines, we have hypothesized that bacteria must favor HGT as a mechanism for evolution to overcome the challenges of long generation times against constant stressors that select for competitive members of a population on shorter time scales. In addition to HGT, we hypothesize that transposable elements may also play an important role in genetic regulation in cryopeg brine populations by rearranging genomes such that a subset of necessary genes for a given condition are actively transcribed when needed and moved to an inactive state otherwise to conserve energy. Here, we have identified additional evidence that mobilome-mediated gene flow is likely prevalent in both sea ice and cryopeg brines, while also identifying a clear signature of the greater importance of the mobilome in sharing genes between cryopeg populations and their communities.

The consistent differences in SNV density at the whole MAG, gene category, and gene levels between sea ice and cryopeg brine populations illustrate the different evolutionary strategies and consequences of each environment. Given that sea ice brine bacteria have opportunity for growth and replication in less extreme conditions when cryopeg brine bacteria are always under growth-limiting stress, sea ice brine bacteria simply may have gone through more generations since the geophysical separation of these two environments around 40 thousand years ago (Iwahana et al., 2021). The lack of vertically inherited SNVs in cryopeg brine populations, however, is more likely the result of a selective force over this extended period for the strains with specific

genotypes that are competitive in their niches. Cryopeg brines receive limited input from their surroundings, with only minimal input from immediately adjacent permafrost material that may undergo minor freeze-thaw cycles through interactions with brine (Iwahana et al., 2021). This lack of external influence may have created highly specific niches that leave little opportunity for deviation for a population to remain competitive. Low diversity in cryopeg brines has previously been observed at the community level by Cooper et al. (2019, Chapter 2) and at the gene level by Rapp et al. (2021), and here we add to that story by illustrating the low level of microdiversity that is prevalent for the populations studied here. In contrast, we expect that moderate environments with higher biodiversity (e.g., pelagic seawater) will harbor populations with higher levels of microdiversity than observed here that demonstrate a greater degree of evolutionary freedom that is less restricted by extreme stressors. Low levels of biodiversity are known in extreme environments and are indicative of the highly specialized nature of the microorganisms that inhabit them (Pikuta et al., 2007). This study adds to that paradigm and pushes the boundaries of our consideration of the evolutionary mechanisms (i.e., horizontal gene transfer) that are used in extreme environments (Brazelton & Baross, 2009; Li et al., 2014). Subzero brine environments on Earth are vulnerable to loss due to climate change, and the highly specialized inhabitants of these communities may undergo unforeseen changes with exposure to new conditions. Whether studying these communities on Earth or searching for them in analogous environments on other habitable worlds, we can understand and observe that they are capable of utilizing evolutionary strategies that can overcome complex physical and chemical limitations.

3.5 CONCLUSION

Acquiring high-quality, nearly complete MAGs generated from a combination of long- and short-read sequencing technologies gave us a unique opportunity to examine the evolutionary dynamics of microbial populations in natural extreme environments. By sequencing metagenomes from individual samples of cryopeg and sea ice brines, we generated MAGs that represent a snapshot of bacterial populations that allow us to observe the variety of evolutionary states that occur in extreme environments. Previous studies have discussed the biodiversity and metabolic potential of bacterial communities inhabiting cryopeg and sea ice brines. Here, we used the collection of MAGs representing *in situ* populations to quantify genomic microdiversity and explore networks of gene sharing that illustrate evolutionary and ecological strategies in these extreme environments. Cryopeg brine populations contain signatures of clonal expansion and genome evolution via horizontal gene transfer, particularly in genes related to the mobilome. Comparing genes that are shared between cryopeg and sea ice brine populations indicates that genotypic exclusivity may be important between abundant taxa in each environment and that cryopeg brine MAGs share more mobilome genes in general than sea ice MAGs. These findings illustrate the competitive and highly selective nature of the cryopeg brine environment and demonstrate the utility of the mobilome as an evolutionary mechanism in a persistently extreme environment. Cryopeg populations have a clear signal of clonal selection and sharing of genes related to the mobilome (significantly greater than for sea ice populations), and the process of horizontal gene transfer via the mobilome likely plays a role in the success of microorganisms under persistent pressure from multiple extreme conditions.

References

Alneberg, J., Bjarnason, B. S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N.J., Andersson, A.F., & Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*(11), 1144–1146. https://doi.org/10.1038/nmeth.3103

Anderson, R. E., Reveillaud, J., Reddington, E., Delmont, T. O., Eren, A. M., McDermott, J. M., Seewald, J.S., & Huber, J. A. (2017). Genomic variation in microbial populations inhabiting the marine subseafloor at deep-sea hydrothermal vents. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-01228-6

Bendall, M. L., Stevens, S. L. R., Chan, L. K., Malfatti, S., Schwientek, P., Tremblay, J., Schackwitz, W., Martin, J., Pati, A., Bushnell, B., Froula, J., Kang, D., Tringe, S. G., Bertilsson, S., Moran, M. A., Shade, A., Newton, R. J., McMahon, K. D., & Malmstrom, R. R. (2016). Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. *ISME Journal*, *10*(7), 1589–1601. https://doi.org/10.1038/ismej.2015.241

Blot, M. (1994). Transposable elements and adaptation of host bacteria. *Genetica 1994 93:1*, *93*(1), 5–12. https://doi.org/10.1007/BF01435235

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., & Rapp, J. Z. (2015). Microbial ecology of the cryosphere: Sea ice and glacial habitats. *Nature Reviews Microbiology*, *13*(11), 677–690. https://doi.org/10.1038/nrmicro3522

Bowman, J. S. (2015). The relationship between sea ice bacterial community structure and biogeochemistry: A synthesis of current knowledge and known unknowns. *Elementa*, *3*(0), 000072. https://doi.org/10.12952/journal.elementa.000072

Brazelton, W. J., & Baross, J. A. (2009). Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME Journal*, *3*(12), 1420–1424. https://doi.org/10.1038/ismej.2009.79

Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59-60. https://doi.org/10.1038/nmeth.3176

Casanueva, A., Tuffin, M., Cary, C., & Cowan, D. A. (2010). Molecular adaptations to psychrophily: The impact of "omic" technologies. *Trends in Microbiology*. https://doi.org/10.1016/j.tim.2010.05.002

Colangelo-Lillis, J., Eicken, H., Carpenter, S. D., & Deming, J. W. (2016). Evidence for marine origin and microbial-viral habitability of sub-zero hypersaline aqueous inclusions within permafrost near Barrow, Alaska. *FEMS Microbiology Ecology*, *92*(5). https://doi.org/10.1093/femsec/fiw053

Collins, R. E., & Deming, J. W. (2011a). Abundant dissolved genetic material in Arctic sea ice Part I: Extracellular DNA. *Polar Biology*, *34*(12), 1819–1830. https://doi.org/10.1007/s00300-011-1041-y

Collins, R. E., & Deming, J. W. (2011b). Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up. *Polar Biology*, *34*(12), 1831–1841. https://doi.org/10.1007/s00300-011-1008-z

Collins, R. E., & Deming, J. W. (2013). An inter-order horizontal gene transfer event enables the catabolism of compatible solutes by *Colwellia psychrerythraea* 34H. *Extremophiles*, *17*(4), 601–610. https://doi.org/10.1007/s00792-013-0543-7

Cooper, Z. S., Rapp, J. Z., Carpenter, S. D., Iwahana, G., Eicken, H., & Deming, J. W. (2019). Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiology Ecology*, *95*(12). Z. Cooper dissertation, Chapter 2. https://doi.org/10.1093/femsec/fiz166

Cooper, Z.S., Rapp, J. Z., Shoemaker, A.M.D., Anderson, R.E., Zhong, Z.P., Sullivan, M.B., and Deming, J.W. (in prep) Evolutionary divergence of *Marinobacter* strains in cryopeg brines as revealed by pangenomics. Z. Cooper dissertation, Chapter 4.

Cox, G. F. N., & Weeks, W. F. (1983). Equations for determining the gas and brine volumes in sea ice samples. *Journal of Glaciology*, *29*(102), 306–316. https://doi.org/10.3189/s0022143000008364

Czajka, J. J., Abernathy, M. H., Benites, V. T., Baidoo, E. E. K., Deming, J. W., & Tang, Y. J. (2018). Model metabolic strategy for heterotrophic bacteria in the cold ocean based on *Colwellia psychrerythraea* 34H. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(49), 12507–12512. https://doi.org/10.1073/pnas.1807804115

Darwin, C. (1859) *On the origin of species by means of natural selection*. London: J. Murray. [PDF] Retrieved from the Library of Congress, https://www.loc.gov/item/06017473/.

Dawson, H. M., Heal, K. R., Boysen, A. K., Carlson, L. T., Ingalls, A. E., & Young, J. N. (2020). Potential of temperature- And salinity-driven shifts in diatom compatible solute concentrations to impact biogeochemical cycling within sea ice. *Elementa*, *8*(23). https://doi.org/10.1525/ELEMENTA.421/112775

Delmont, T. O., Kiefl, E., Kilinc, O., Esen, O. C., Uysal, I., Rappé, M. S., Giovanonni, S., & Eren, A. M. (2019). Single-amino acid variants reveal evolutionary processes that shape the biogeography of a global SAR11 subclade. *ELife*, *8*. https://doi.org/10.7554/eLife.46497

Deming, J. W., & Eric Collins, R. (2016). Sea ice as a habitat for Bacteria, Archaea and viruses. *Sea Ice: Third Edition*, 326–351. https://doi.org/10.1002/9781118778371.ch13

Druckenmiller, M. L., Eicken, H., Johnson, M. A., Pringle, D. J., & Williams, C. C. (2009). Toward an integrated coastal sea-ice observatory: System components and a case study at Barrow, Alaska. *Cold Regions Science and Technology*, *56*(2–3), 61–72. https://doi.org/10.1016/J.COLDREGIONS.2008.12.003

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

Elena, S. F., & Lenski, R. E. (2003). Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, *4*(6), 457–469. https://doi.org/10.1038/nrg1088

Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015a). Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ*, *2015*(10), 1–29. https://doi.org/10.7717/peerj.1319

Eren, A. M., Morrison, H. G., Lescault, P. J., Reveillaud, J., Vineis, J. H., & Sogin, M. L. (2015b). Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME Journal*, *9*(4), 968–979. https://doi.org/10.1038/ismej.2014.195

Eronen-Rasimus, E., Kaartokallio, H., Lyra, C., Autio, R., Kuosa, H., Dieckmann, G. S., & Thomas, D. N. (2014). Bacterial community dynamics and activity in relation to dissolved organic matter availability during sea-ice formation in a mesocosm experiment. *MicrobiologyOpen*, *3*(1), 139–156. https://doi.org/10.1002/mbo3.157

Ewert, M., & Deming, J. W. (2014). Bacterial responses to fluctuations and extremes in temperature and brine salinity at the surface of Arctic winter sea ice. *FEMS Microbiology Ecology*, *89*(2), 476–489. https://doi.org/10.1111/1574-6941.12363

Fenchel, T., & Finlay, B. J. (2004). The Ubiquity of Small Species: Patterns of Local and Global Diversity. *BioScience* (Vol. 54). Oxford Academic. https://doi.org/10.1641/0006-3568(2004)054[0777:TUOSSP]2.0.CO;2

Firth, E., Carpenter, S. D., Sørensen, H. L., Collins, R. E., & Deming, J. W. (2016). Bacterial use of choline to tolerate salinity shifts in sea-ice brines. *Elementa*, *2016*. https://doi.org/10.12952/journal.elementa.000120

Fuchsman, C. A., Collins, R. E., Rocap, G., & Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ*, *2017*(9), e3865. https://doi.org/10.7717/peerj.3865

Galperin, M. Y., Kristensen, D. M., Makarova, K. S., Wolf, Y. I., & Koonin, E. V. (2019). Microbial genome analysis: The COG approach. *Briefings in Bioinformatics*, *20*(4), 1063–1070. https://doi.org/10.1093/bib/bbx117

Galtier, N. (2007). A model of horizontal gene transfer and the bacterial phylogeny problem. *Systematic Biology*, *56*(4), 633–642. https://doi.org/10.1080/10635150701546231

Germerodt, S., Bohl, K., Lück, A., Pande, S., Schröter, A., Kaleta, C., Schuster, S., & Kost, C. (2016). Pervasive Selection for Cooperative Cross-Feeding in Bacterial Communities. *PLOS Computational Biology*, *12*(6), e1004986. https://doi.org/10.1371/JOURNAL.PCBI.1004986

Gilichinsky, D., Rivkina, E., Shcherbakova, V., Laurinavichuis, K., & Tiedje, J. (2003). Supercooled water brines within permafrost - An unknown ecological niche for microorganisms: A model for astrobiology. *Astrobiology*, *3*(2), 331–341. https://doi.org/10.1089/153110703769016424

Gillings, M. R. (2017). Lateral gene transfer, bacterial genome evolution, and the Anthropocene. *Annals of the New York Academy of Sciences*, *1389*(1), 20–36. https://doi.org/10.1111/nyas.13213

Golden, K. M., Eicken, H., Heaton, A. L., Miner, J., Pringle, D. J., & Zhu, J. (2007). Thermal evolution of permeability and microstructure in sea ice. *Geophysical Research Letters*, *34*(16), 2–7. https://doi.org/10.1029/2007GL030447

Good, B. H., McDonald, M. J., Barrick, J. E., Lenski, R. E., & Desai, M. M. (2017). The dynamics of molecular evolution over 60,000 generations. *Nature 2017 551:7678*, *551*(7678), 45–50. https://doi.org/10.1038/nature24287

Grossmann, S., & Dieckmann, G. S. (1994). Bacterial standing stock, activity, and carbon production during formation and growth of sea ice in the Weddell Sea, Antarctica. *Applied and Environmental Microbiology*, *60*(8), 2746–2753. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16349347

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring network structure, dynamics, and function using NetworkX. In *7th Python in Science Conference (SciPy 2008)* (pp. 11–15). Retrieved from http://conference.scipy.org/proceedings/SciPy2008/paper_2/

Harrison, J. P., Gheeraert, N., Tsigelnitskiy, D., & Cockell, C. S. (2013). The limits for life under multiple extremes. *Trends in Microbiology*. Elsevier Current Trends. https://doi.org/10.1016/j.tim.2013.01.006

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

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. *BMC Bioinformatics*, *11*, 119. https://doi.org/10.1186/1471-2105-11-119

Iizuka, Y., Miyamoto, C., Matoba, S., Iwahana, G., Horiuchi, K., Takahashi, Y., Kanna, N., Suzuki, K., & Ohno, H. (2019). Ion concentrations in ice wedges: An innovative approach to reconstruct past climate variability. *Earth and Planetary Science Letters*, *515*, 58–66. https://doi.org/10.1016/j.epsl.2019.03.013

Iwahana, G., Cooper, Z. S., Carpenter, S. D., Deming, J. W., & Eicken, H. (2021). Intra-ice and intra-sediment cryopeg brine occurrence in permafrost near Utqiaġvik (Barrow). *Permafrost and Periglacial Processes*, *32*(3), 427–446. https://doi.org/10.1002/ppp.2101

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, *2019*(7), e7359. https://doi.org/10.7717/peerj.7359

Kolmogorov, M., Yuan, J., Lin, Y., & Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology*, *37*(5), 540–546. https://doi.org/10.1038/s41587-019-0072-8

Krembs, C., Eicken, H., & Deming, J. W. (2011). Exopolymer alteration of physical properties of sea ice and implications for ice habitability and biogeochemistry in a warmer Arctic. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(9), 3653–3658. https://doi.org/10.1073/pnas.1100701108

Lan, R., & Reeves, P. R. (2000). Intraspecies variation in bacterial genomes: The need for a species genome concept. *Trends in Microbiology*, 8(9), 396-401. https://doi.org/10.1016/S0966-842X(00)01791-1

Lee, M. D. (2019). GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics*, *35*(20), 4162–4164. https://doi.org/10.1093/bioinformatics/btz188

Li, S.-J., Hua, Z.-S., Huang, L.-N., Li, J., Shi, S.-H., Chen, L.-X., Kuang, J.L., Liu, J., Hu, M., & Shu, W.-S. (2014). Microbial communities evolve faster in extreme environments. *Scientific Reports 2014 4:1*, *4*(1), 1–9. https://doi.org/10.1038/srep06205

Luhtanen, A. M., Eronen-Rasimus, E., Oksanen, H. M., Tison, J. L., Delille, B., Dieckmann, G. S., Rintala, J. M., & Bamford, D. H. (2018). The first known virus isolates from Antarctic sea ice have complex infection patterns. *FEMS Microbiology Ecology*, *94*(4), 28. https://doi.org/10.1093/FEMSEC/FIY028

Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nature Reviews Microbiology 2006 4:8*, *4*(8), 608–620. https://doi.org/10.1038/nrmicro1462

McClintock, B. (1987). *The discovery and characterization of transposable elements: the collected papers of Barbara McClintock.* New York, New York: Garland Publishing.

Methé, B. A., Nelson, K. E., Deming, J. W., Momen, B., Melamud, E., Zhang, X., Moult, J., Madupu, R., Nelson, W. C., Dodson, R. J., Brinkac, L. M., Daugherty, S. C., Durkin, A. S., DeBoy, R. T., Kolonay, J. F., Sullivan, S. A., Zhou, L., Davidsen, T. M., Wu, M., Huston, A. L., Lewis, M., Weaver, B., Weidman, J. F., Khouri, H., Utterback, T. R., Feldblyum, T. V., & Fraser, C. M. (2005). The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(31), 10913–10918. https://doi.org/10.1073/pnas.0504766102

Meyer, H., Schirrmeister, L., Andreev, A., Wagner, D., Hubberten, H. W., Yoshikawa, K., Bobrov, A., Wetterich, S., Opel, T., Kandiano, E., & Brown, J. (2010a). Lateglacial and Holocene isotopic and environmental history of northern coastal Alaska - Results from a buried ice-wedge system at Barrow. *Quaternary Science Reviews*, *29*(27–28), 3720–3735. https://doi.org/10.1016/j.quascirev.2010.08.005

Meyer, H., Schirrmeister, L., Yoshikawa, K., Opel, T., Wetterich, S., Hubberten, H. W., & Brown, J. (2010b). Permafrost evidence for severe winter cooling during the Younger Dryas in northern Alaska. *Geophysical Research Letters*, *37*(3), 1–5. https://doi.org/10.1029/2009GL041013

Morita, R. Y. (1975). Psychrophilic bacteria. *Bacteriological Reviews*. American Society for Microbiology. https://doi.org/10.1128/mmbr.39.2.144-167.1975

Moulana, A., Anderson, R. E., Fortunato, C. S., & Huber, J. A. (2020). Selection is a significant driver of gene gain and loss in the pangenome of the bacterial genus *Sulfurovum* in geographically distinct deep-sea hydrothermal vents. *MSystems*, *5*(2). https://doi.org/10.1128/mSystems.00673-19

Murray, A. E., Kenig, F., Fritsen, C. H., McKay, C. P., Cawley, K. M., Edwards, R., Kuhn, E., McKnight, D. M., Ostrom, N. E., Peng, V., Ponce, A., Priscu, J. C., Samarkin, V., Townsend, A. T., Wagh, P., Young, S. A., Yung, P. T., & Doran, P. T. (2012). Microbial life at -13°C in the brine of an ice-sealed Antarctic lake. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(50), 20626–20631. https://doi.org/10.1073/pnas.1208607109

Nadell, C. D., Drescher, K., & Foster, K. R. (2016). Spatial structure, cooperation and competition in biofilms. *Nature Reviews Microbiology 2016 14:9*, *14*(9), 589–600. https://doi.org/10.1038/nrmicro.2016.84

Nikrad, M. P., Kerkhof, L. J., & Häggblom, M. M. (2016). The subzero microbiome: Microbial activity in frozen and thawing soils. *FEMS Microbiology Ecology*, *92*(6), 1–16. https://doi.org/10.1093/femsec/fiw081

Nunn, B. L., Slattery, K. V., Cameron, K. A., Timmins-Schiffman, E., & Junge, K. (2015). Proteomics of *Colwellia psychrerythraea* at subzero temperatures - a life with limited movement, flexible membranes and vital DNA repair. *Environmental Microbiology*, *17*(7), 2319–2335. https://doi.org/10.1111/1462-2920.12691

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

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*, *36*(10), 996. https://doi.org/10.1038/nbt.4229

Pecheritsyna, S. A., Shcherbakova, V. A., Kholodov, A. L., Akimov, V. N., Abashina, T. N., Suzina, N. E., & Rivkina, E. M. (2007). Microbiological analysis of cryopegs from the Varandei Peninsula, Barents Sea. *Mikrobiologiia*, *76*(5), 694–701. https://doi.org/10.1134/S0026261707050153

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. Retrieved from http://jmlr.org/papers/v12/pedregosa11a.html

Petrich, C., & Eicken, H. (2017). Overview of sea ice growth and properties. In *Sea Ice: Third Edition* (pp. 1–41). Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118778371.ch1

Pikuta, E. V., Hoover, R. B., & Tang, J. (2007). Microbial extremophiles at the limits of life. *Critical Reviews in Microbiology*, 33(3), 183-209. https://doi.org/10.1080/10408410701451948

Rapp, J. Z., Sullivan, M. B., & Deming, J. W. (2021). Divergent genomic adaptations in the microbiomes of Arctic subzero sea-ice and cryopeg brines. *Frontiers in Microbiology*, *12*(July), 1–21. https://doi.org/10.3389/fmicb.2021.701186

Rivkina, E. M., Friedmann, E. I., McKay, C. P., & Gilichinsky, D. A. (2000). Metabolic activity of Permafrost Bacteria below the freezing point. *Applied and Environmental Microbiology*, *66*(8), 3230–3233. https://doi.org/10.1128/AEM.66.8.3230-3233.2000

Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Kultima, J. R., Martin, J., Kota, K., Sunyaev, S. R., Weinstock, G. M., & Bork, P. (2013). Genomic variation landscape of the human gut microbiome. *Nature 2012 493:7430*, *493*(7430), 45–50. https://doi.org/10.1038/nature11711

Schloter, M., Lebuhn, M., Heulin, T., & Hartmann, A. (2000). Ecology and evolution of bacterial microdiversity. *FEMS Microbiology Reviews*, *24*(5), 647–660. https://doi.org/10.1111/j.1574-6976.2000.tb00564.x

Showalter, G. M., & Deming, J. W. (2018). Low-temperature chemotaxis, halotaxis and chemohalotaxis by the psychrophilic marine bacterium *Colwellia psychrerythraea* 34H. *Environmental Microbiology Reports*, *10*(1), 92–101. https://doi.org/10.1111/1758-2229.12610

Showalter, G. M. (2020). *Acquisition, Degradation, and Cycling of Organic Matter Within Sea-Ice Brines by Bacteria and Their Viruses*. *ProQuest Dissertations and Theses*. http://hdl.handle.net/1773/46535

Showalter, G. M., & Deming, J. W. (2021). Extracellular enzyme activity of model cold-adapted bacteria and Arctic sea-ice microbial communities under subzero hypersaline conditions. *Aquatic Microbial Ecology*, *87*, 99–111. https://doi.org/10.3354/AME01974

Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., & Banfield, J. F. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nature Microbiology*, *3*(7), 836–843. https://doi.org/10.1038/s41564-018-0171-1

Spirina, E. V., Durdenko, E. V., Demidov, N. E., Abramov, A. A., Romanovsky, V. E., & Rivkina, E. M. (2017). Halophilic-psychrotrophic bacteria of an Alaskan cryopeg—a model for astrobiology. *Paleontological Journal*, *51*(13), 1440–1452. https://doi.org/10.1134/S0031030117120036

Streletskaya, I. D., & Leibman, M. O. (2003). Cryogeochemical model of tabular ground ice and cryopegs formation in central Yamal, Russia. *Proc. Int. Conf. on Permafrost*, 1111–1115.

Tatusov, R. L., Koonin, E. V., & Lipman, D. J. (1997). A genomic perspective on protein families. *Science*, *278*(5338), 631–637. https://doi.org/10.1126/science.278.5338.631

Torstensson, A., Dinasquet, J., Chierici, M., Fransson, A., Riemann, L., & Wulff, A. (2015). Physicochemical control of bacterial and protist community composition and diversity in Antarctic sea ice. *Environmental Microbiology*, *17*(10), 3869–3881. https://doi.org/10.1111/1462-2920.12865

Torstensson, A., Young, J. N., Carlson, L. T., Ingalls, A. E., & Deming, J. W. (2019). Use of exogenous glycine betaine and its precursor choline as osmoprotectants in Antarctic sea-ice diatoms 1. *Journal of Phycology*, *55*(3), 663–675. https://doi.org/10.1111/jpy.12839

Toussaint, A., & Chandler, M. (2012). Prokaryote genome fluidity: Toward a system approach of the mobilome. *Methods in Molecular Biology*, *804*, 57–80. https://doi.org/10.1007/978-1-61779-361-5_4

Turner, P. E., Souza, V., & Lenski, R. E. (1996). Tests of ecological mechanisms promoting the stable coexistence of two bacterial genotypes. *Ecology*, *77*(7), 2119–2129. https://doi.org/10.2307/2265706

Vancoppenolle, M., Meiners, K. M., Michel, C., Bopp, L., Brabant, F., Carnat, G., Delille, B., Lannuzel, D., Madec, G., Moreau, S., Tison, J. L., & van der Merwe, P. (2013). Role of sea ice in global biogeochemical cycles: emerging views and challenges. *Quaternary Science Reviews*, *79*, 207–230. https://doi.org/10.1016/J.QUASCIREV.2013.04.011

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., & Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, *9*(11), e112963. https://doi.org/10.1371/journal.pone.0112963

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/JOSS.03021

Wells, L. E., & Deming, J. W. (2006a). Characterization of a cold-active bacteriophage on two psychrophilic marine hosts. *Aquatic Microbial Ecology*, *45*(1), 15–29. https://doi.org/10.3354/ame045015

Wells, L. E., & Deming, J. W. (2006b). Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environmental Microbiology*, *8*(6), 1115–1121. https://doi.org/10.1111/j.1462-2920.2006.00984.x

Wu, Y. W., Simmons, B. A., & Singer, S. W. (2016). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. https://doi.org/10.1093/bioinformatics/btv638

Yoshikawa, K., Romanovsky, V., Duxbury, N., Brown, J., & Tsapin, A. (2004). The use of geophysical methods to discriminate between brine layers and freshwater taliks in permafrost regions. *Journal of Glaciology and Geocryology*, *26*, 301–309.

Zablocki, O., Michelsen, M., Burris, M., Solonenko, N., Warwick-Dugdale, J., Ghosh, R., Pett-Ridge, J., Sullivan, M. B., & Temperton, B. (2021). VirION2: a short- and long-read sequencing and informatics workflow to study the genomic diversity of viruses in nature. *PeerJ*, *9*, e11088. https://doi.org/10.7717/peerj.11088

Zhong, Z.-P., Rapp, J. Z., Wainaina, J. M., Solonenko, N. E., Maughan, H., Carpenter, S. D., Cooper, Z. S., Jang, H. B., Bolduc, B., Deming, J. W., & Sullivan, M. B. (2020). Viral ecogenomics of Arctic cryopeg brine and sea ice. *MSystems*, *5*(3). https://doi.org/10.1128/msystems.00246-20

Table 3.1. MAG statistics

Binning program ID	Genus	Sample	ID	Total Length (bp)	Number of contigs	N50	GC content (%)	Mean Coverage (X)	Completeness (%)	Contamination (%)
concoct_046_1_1	Marinobacter	CBIW17	m17	3,085,923	32	175,875	54.28	4017.41	79.01	1.80
metabat2_002_1	Marinobacter	CBIW18	m18	4,143,481	42	130,109	54.42	6154.52	84.82	0.92
metabat2_013_1	Marinobacter	CB1	m1	3,588,131	41	130,140	54.44	2996.23	82.29	0
concoct_032_1	Marinobacter	CBIA	mia	4,242,133	43	141,017	54.23	5973.60	92.03	0.61
metabat2_009_1	Leeuwenhoekiella	CBIW17	l17	4,579,413	26	209,893	40.38	318.27	89.41	0.63
concoct_051_1_1	Leeuwenhoekiella	CBIW18	l18	4,716,813	49	206,289	40.49	209.31	95.10	0.70
concoct_093_1_1	Leeuwenhoekiella	CBIA	lia	5,016,789	57	155,341	40.37	822.83	96.58	1.65
maxbin2_011_1	Psychrobacter	CBIW17	ps17	3,598,623	6	2,386,693	42.95	96.76	98.21	1.03
concoct_011	Psychrobacter	CBIA	psia	3,510,628	25	176,028	43.01	197.20	97.8	1.12
concoct_012	UBA2294	CB1	ub1	3,760,202	108	58,208	50.62	721.46	96.77	0.65
concoct_071_1	UBA2294	CBIA	ubia	2,858,326	151	24,409	50.62	86.20	80.97	0.65
concoct_043	Sulfurospirillum	CBIA	suia	1,784,900	73	30,380	34.94	506.53	79.49	1.03
metabat2_028_1	Thiomicrorhabdus	CBIW17	t17	2,490,835	44	76,445	42.11	20.99	84.00	1.22
metabat2_027	Salegentibacter	CBIW17	sa17	3,577,183	60	77,021	37.95	64.39	85.88	1.08
concoct_007_1_1	Pseudomonas	CBIW17	pseu17	2,919,532	48	84,964	55.61	76.23	87.99	2.23
concoct_041_1_1	Algoriphagus	CBIW17	al17	6,261,446	71	146,871	40.05	78.50	91.60	1.09
maxbin2_015_1	Methylophaga	CBIW17	me17	3,095,353	32	215,644	39.93	54.29	99.53	1.31
concoct_021_1_1	Polaribacter	SB17	po17	2,801,699	19	195,336	32.90	7351.09	92.16	1.71
metabat2_025	Polaribacter	SB18	po18	2,744,059	20	177,044	32.80	1597.05	92.08	1.87
concoct_025_1_1	Glaciecola	SB17	g17	3,569,280	22	249,496	40.88	440.96	95.62	1.36
metabat2_001_1	Glaciecola	SB18	g18	3,198,037	51	101,843	41.03	32.52	83.66	2.89
metabat2_015	HTCC2207	SB18	h18	2,349,403	49	64,042	50.83	41.29	76.86	0.56
metabat2_010_1	Octadecabacter	SB17	o17	2,934,704	62	67,120	55.41	456.67	78.02	1.56
concoct_011_1_1	ASP10-02a	SB17	as17	2,907,160	18	271,428	46.53	514.70	93.33	1.03

Table 3.2. MAG microdiversity data

MAG ID	Genus	Whole MAG				Mobilome			
		Number of genes	Mean entropy	Mean SNVs per kb	Mean pN/pS	Number of genes	Mean entropy	Mean SNVs per kb	Mean pN/pS
m17	*Marinobacter*	3,237.00	0.029	38.26	1.06	8.00	0.026	38.53	0.94
m18	*Marinobacter*	4,333.00	0.022	16.60	0.99	19.00	0.025	25.59	0.73
m1	*Marinobacter*	3,577.00	0.023	18.32	1.02	20.00	0.032	36.92	1.05
mia	*Marinobacter*	4,233.00	0.029	32.27	1.17	26.00	0.161	40.68	0.95
l17	*Leeuwenhoekiella*	4,782.00	0.017	16.82	0.97	30.00	0.030	35.20	0.87
l18	*Leeuwenhoekiella*	4,550.00	0.017	16.69	0.97	41.00	0.032	37.38	0.74
lia	*Leeuwenhoekiella*	4,629.00	0.013	12.21	1.06	28.00	0.037	40.40	0.93
ps17	*Psychrobacter*	3,109.00	0.009	15.64	1.01	13.00	0.016	24.76	0.56
psia	*Psychrobacter*	3,050.00	0.006	2.59	1.09	2.00	0.009	9.13	0.55
ub1	UBA2294	3,777.00	0.010	6.42	1.11	28.00	0.050	30.10	0.46
ubia	UBA2294	2,918.00	0.006	1.90	1.12	24.00	0.042	24.62	0.94
suia	*Sulfurospirillum*	2,029.00	0.007	4.37	0.85	7.00	0.047	48.74	0.86
t17	*Thiomicrorhabdus*	3,346.00	0.047	53.05	1.08	7.00	0.037	46.10	1.25
sa17	*Salegentibacter*	3,510.00	0.010	8.43	1.00	6.00	0.043	63.84	0.49
pseu17	*Pseudomonas*	2,983.00	0.011	10.15	1.20	0.00	–	–	–
al17	*Algoriphagus*	5,917.00	0.012	20.06	0.97	56.00	0.016	20.93	0.73
me17	*Methylophaga*	3,116.00	0.009	11.80	0.85	17.00	0.050	43.87	1.09
po17	*Polaribacter*	3,123.00	0.031	28.09	0.88	17.00	0.043	52.00	0.97
po18	*Polaribacter*	2,929.00	0.046	38.62	0.91	23.00	0.051	64.01	1.04
g17	*Glaciecola*	3,792.00	0.053	74.46	1.11	4.00	0.042	53.38	1.07
g18	*Glaciecola*	3,652.00	0.057	75.17	1.11	6.00	0.150	188.49	1.04
h18	HTCC2207	2,646.00	0.067	96.17	1.57	5.00	0.050	65.43	1.03
o17	*Octadecabacter*	3,806.00	0.069	95.86	1.22	92.00	0.086	113.48	1.05
as1	ASP10-02a	3,304.00	0.048	66.36	1.04	4.00	0.096	139.21	1.22

Table 3.3. MAG mobilome microdiversity data

MAG ID	Genus	Transposase				Integrase			
		Number of genes	Mean entropy	Mean SNVs per kb	Mean pN/pS	Number of genes	Mean entropy	Mean SNVs per kb	Mean pN/pS
m17	*Marinobacter*	12.00	0.019	23.50	1.40	2.00	0.012	10.03	1.52
m18	*Marinobacter*	12.00	0.021	22.60	0.99	3.00	0.023	22.05	1.21
m1	*Marinobacter*	13.00	0.021	20.88	0.85	1.00	0.017	16.82	2.21
mia	*Marinobacter*	15.00	0.121	32.91	0.96	3.00	0.199	27.36	0.48
l17	*Leeuwenhoekiella*	23.00	0.028	34.05	0.84	1.00	0.015	7.97	1.91
l18	*Leeuwenhoekiella*	37.00	0.033	38.28	0.77	3.00	0.019	23.12	0.46
lia	*Leeuwenhoekiella*	25.00	0.036	38.26	0.72	0.00	–	–	–
ps17	*Psychrobacter*	12.00	0.017	26.04	0.59	2.00	0.012	21.92	0.74
psia	*Psychrobacter*	2.00	0.009	9.13	0.55	0.00	–	–	–
ub1	UBA2294	26.00	0.050	26.75	0.46	0.00	–	–	–
ubia	UBA2294	21.00	0.043	22.66	0.81	1.00	0.072	80.56	2.84
suia	*Sulfurospirillum*	7.00	0.047	48.74	0.86	0.00	–	–	–
t17	*Thiomicrorhabdus*	6.00	0.031	41.64	1.45	0.00	–	–	–
sa17	*Salegentibacter*	5.00	0.030	51.36	0.50	0.00	–	–	–
pseu17	*Pseudomonas*	0.00	–	–	–	0.00	–	–	–
al17	*Algoriphagus*	53.00	0.013	18.18	0.66	5.00	0.032	39.16	0.54
me17	*Methylophaga*	13.00	0.057	43.11	1.05	1.00	0.012	21.08	0.92
po17	*Polaribacter*	14.00	0.046	55.95	1.04	0.00	–	–	–
po18	*Polaribacter*	23.00	0.051	64.01	1.04	1.00	0.051	87.00	0.28
g17	*Glaciecola*	2.00	0.035	44.79	1.36	1.00	0.057	74.52	0.30
g18	*Glaciecola*	2.00	0.096	141.88	0.73	0.00	–	–	–
h18	HTCC2207	4.00	0.042	51.38	0.70	0.00	–	–	–
o17	*Octadecabacter*	78.00	0.086	111.98	1.04	8.00	0.077	101.93	0.62
as1	ASP10-02a	0.00	–	–	–	3.00	0.105	147.55	0.97

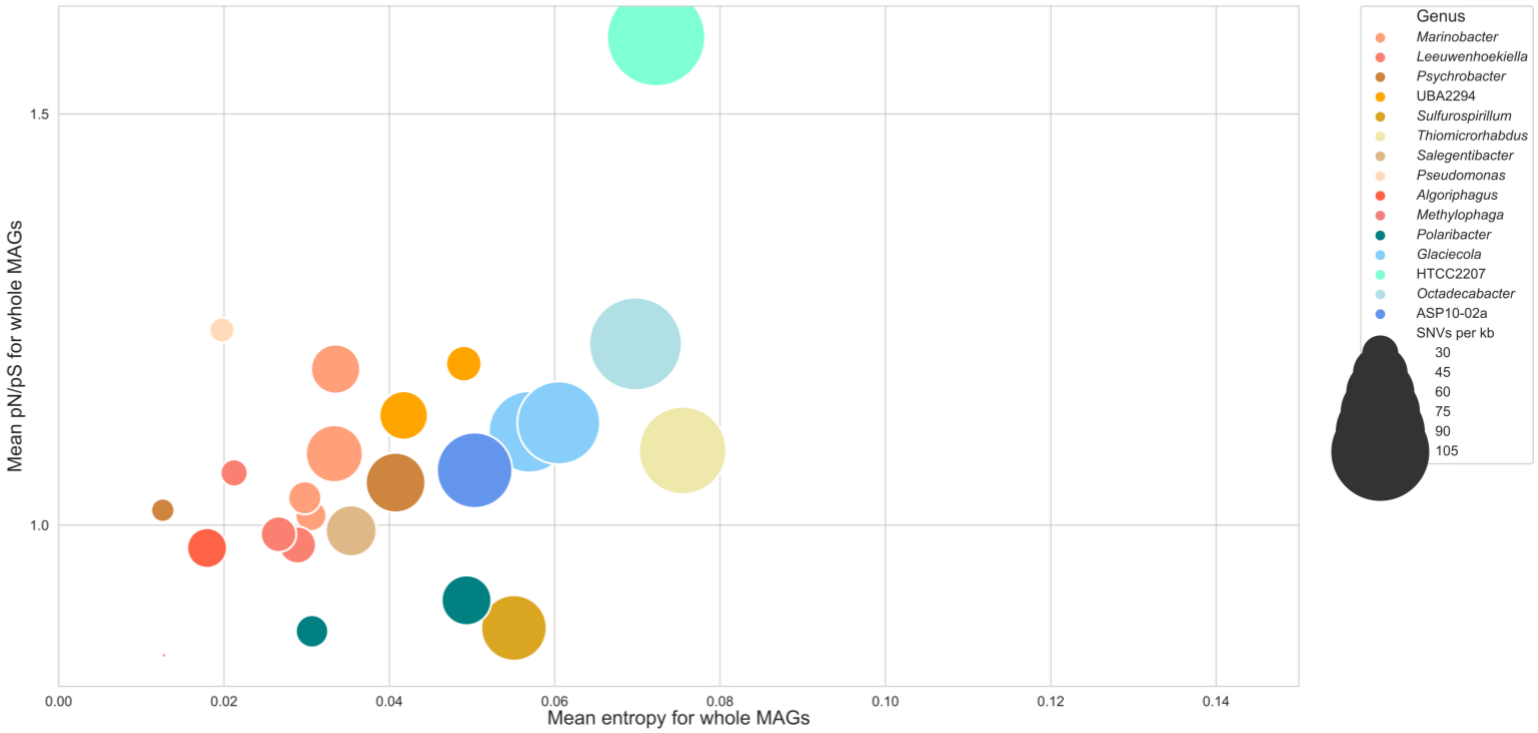

Figure 3.1. Scatter plot of mean values of microdiversity metrics for all genes in each MAG. Each circle represents a single MAG color-coded by taxonomy and scaled by mean number of single nucleotide variants per kilobase.

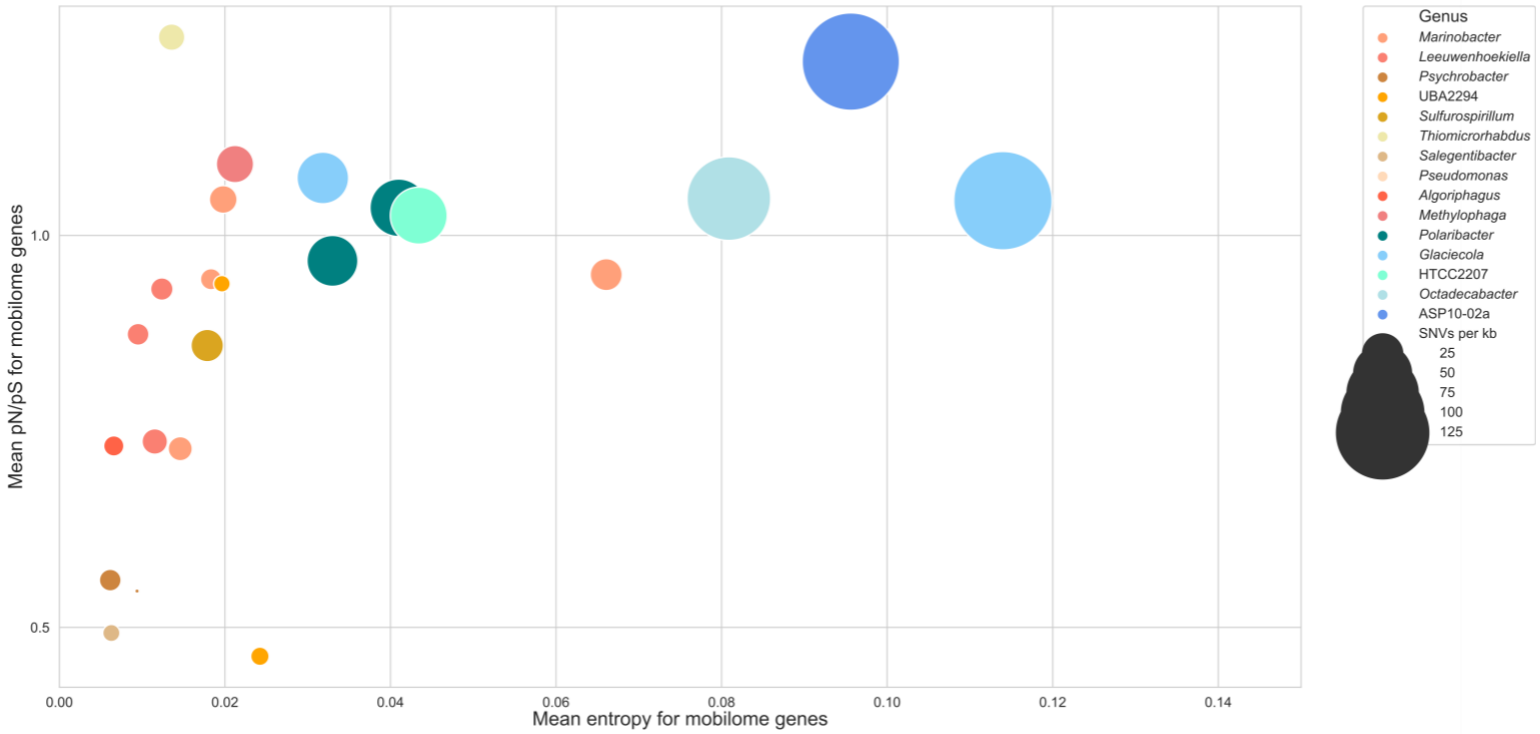

Figure 3.2. Scatter plot of mean values of microdiversity metrics for all mobilome (COG category X) genes in each MAG.
Each dot represents a single MAG color-coded by taxonomy and scaled by mean number of single nucleotide variants per kilobase.

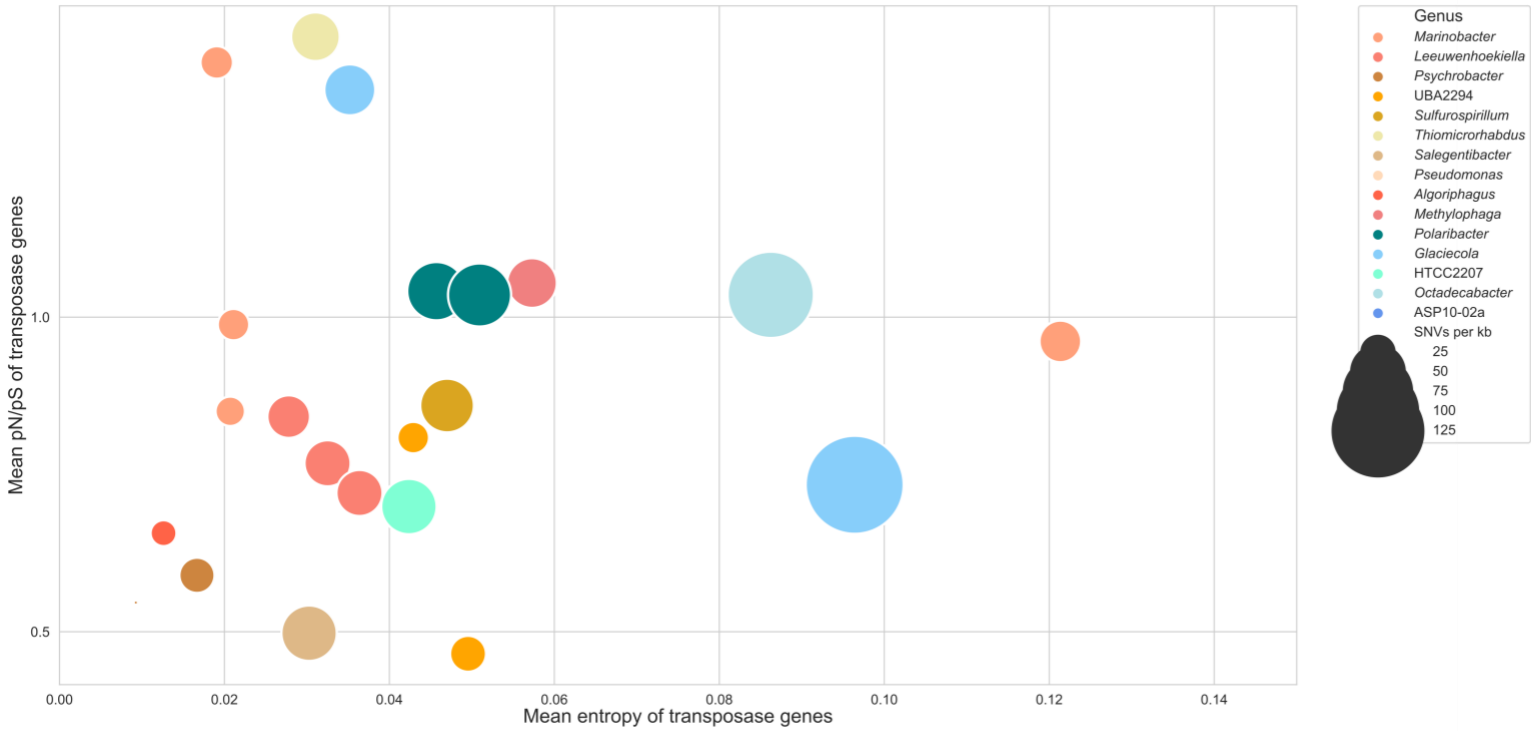

Figure 3.3. Scatter plot of mean values of microdiversity metrics for all transposase genes in each MAG.
Each dot represents a single MAG color-coded by taxonomy and scaled by mean number of single nucleotide variants per kilobase.

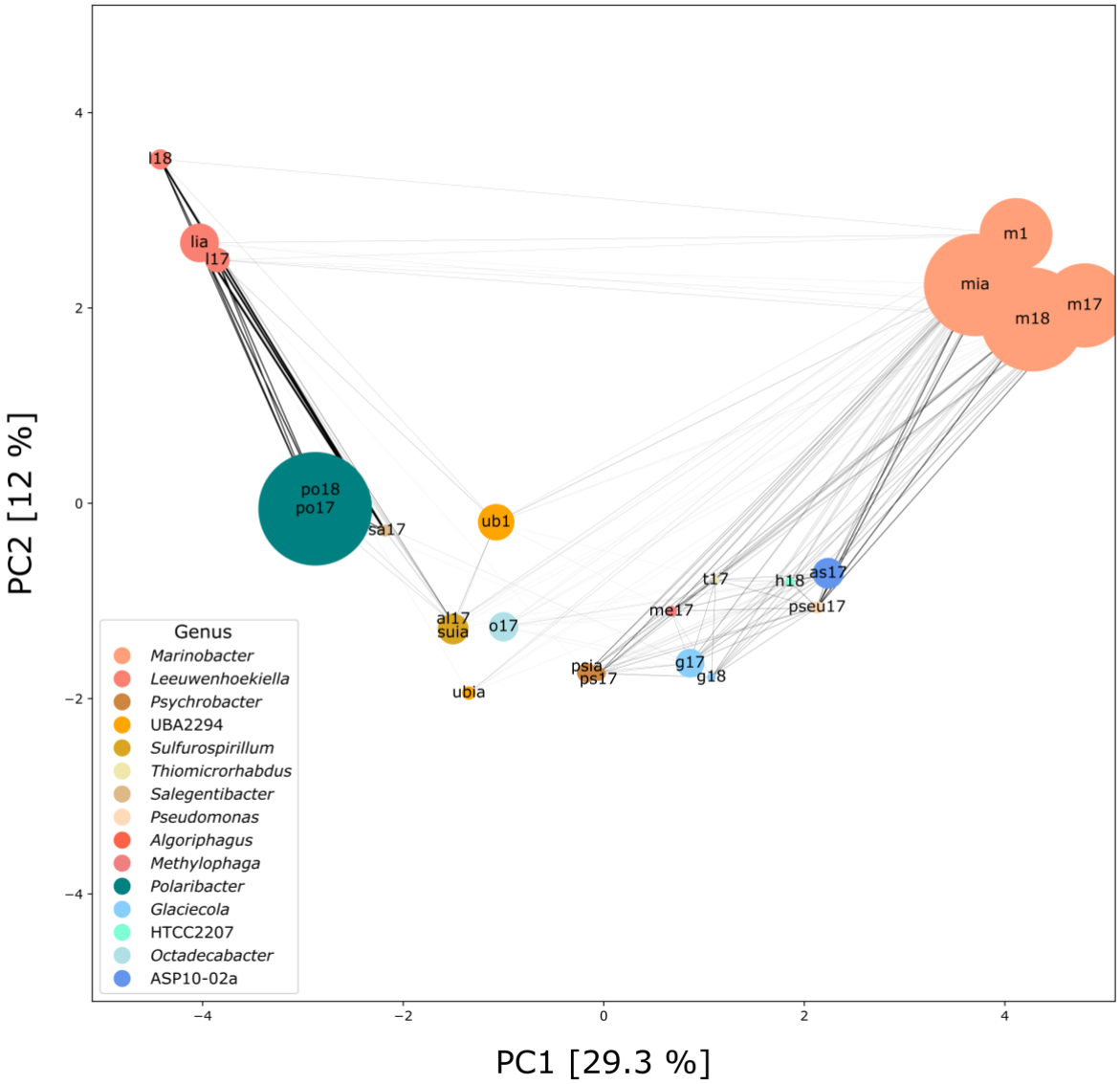

Figure 3.4. Network diagram and PCA ordination of each MAG based on genes shared between MAGs.
The thickness of the line between each MAG indicates the relative number of genes shared between those two MAGs. Each MAG is color-coded according to its taxonomy and size-scaled based on mean coverage. Each circle is labelled with a MAG identifier that can be found in Table 3.1.

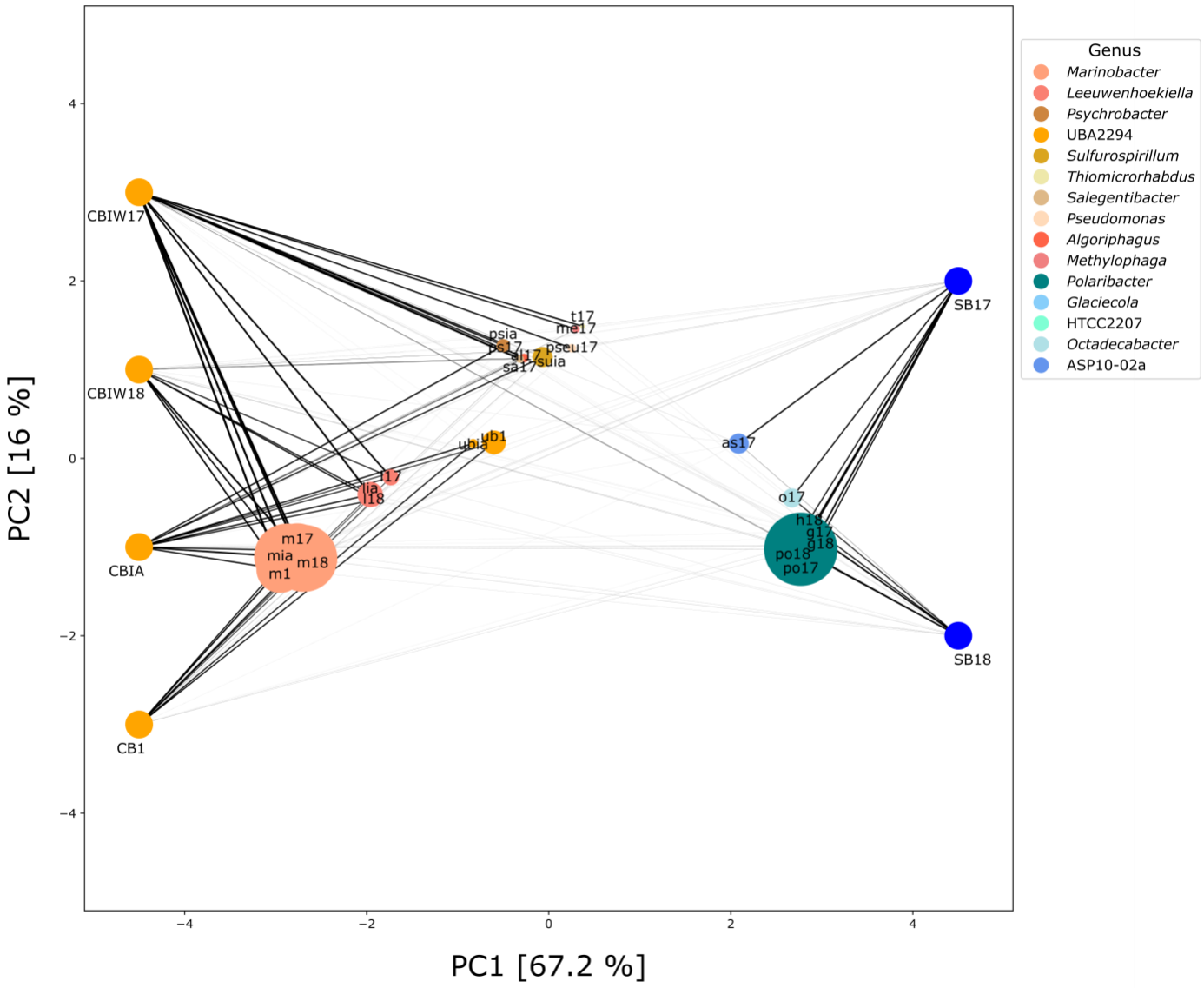

Figure 3.5. Network diagram and PCA ordination of each MAG based on genes shared between MAGs and the metagenomes.
The thickness of the line between each MAG indicates the relative number of genes shared between that MAG and the metagenome. Each MAG is color-coded according to its taxonomy and size-scaled based on mean coverage. Each circle is labelled with a MAG identifier that can be found in Table 3.1. Metagenome circles (4 for cryopeg brines on the left; 2 for sea ice brines on the right) were placed manually and are not based on the PCA.

Chapter 4. EVOLUTIONARY DIVERGENCE OF *MARINOBACTER* STRAINS IN CRYOPEG BRINES AS REVEALED BY PANGENOMICS

4.1 ABSTRACT

Marinobacter spp. are cosmopolitan in saline environments, displaying a diverse set of metabolisms that allow them to competitively occupy such niches, some of which can be extreme in both salinity and temperature. Although *Marinobacter* spp. have been characterized physiologically, no genus-wide genomic analysis is available. Here, we introduce 8 new *Marinobacter* genomes (4 complete genomes from purified isolates and 4 metagenome-assembled genomes) obtained from the extreme environment of subzero, hypersaline cryopeg brines, relic seawater-derived liquid habitats within permafrost, sampled near Utqiaġvik, Alaska. Using these new genomes and 45 representative publicly available genomes of *Marinobacter* spp. from other settings, we assembled a pangenome to gain an overview of the genus and how the new extremophile members fit into it evolutionarily and ecologically, based on genetic potential and environmental source. We found that *Marinobacter* spp. in general encode metabolic pathways that are thermodynamically favored at low temperature, cover a broad range of organic compounds, and optimize protein usage, e.g., the Entner-Doudoroff pathway, the glyoxylate shunt, and amino acid catabolism. We also identified a distinct clade that encompasses only the subzero brine-dwelling *Marinobacter* spp., including each of our new genomes, and that appears to have diverged genotypically and phylogenetically from all other members of the named genus. We found that this clade displays genomic characteristics that may explain adaptations imparting competitive ecological advantages within the extreme

environments they inhabit, including more abundant membrane transport systems (e.g., for organic substrates, compatible solutes and ions) and stress-induced transcriptional regulatory mechanisms (e.g., for cold and salt stress), than in the other *Marinobacter* clades. We also identified more abundant signatures of potential horizontal transfer of genes involved in transcription, the mobilome, and a variety of metabolite exchange systems, which led to consideration of how this evolutionary mechanism may confer adaptation to the extreme environment of subzero hypersaline brines, e.g., by providing genome diversification mechanisms where adaptation via vertical evolution is physiologically rate-limited). Assessing our new extremophile genomes in a pangenomic context has provided a unique view into the ecological and evolutionary history of the genus *Marinobacter*, particularly with regards to its remarkable diversity and its opportunism in extremely cold and saline environments.

4.2 INTRODUCTION

Evolution and ecology are fundamentally entangled biological principles. Since life originated on Earth, the genomic variation that drives evolution has been introduced vertically, through the production of stochastic mutations at the individual nucleotide level, as well as horizontally, via horizontal gene transfer (HGT) between organisms recognizable at the whole gene level. Those mutations or introductions that confer adaptation are selected by ecological conditions (Schloter et al., 2000; Brockhurst et al., 2019). Ubiquitous on Earth, including extreme environments, are the prokaryotes (Fenchel and Finlay, 2004), morphologically simple organisms (single-celled, generally lacking internal compartmentalization) that maximize metabolic efficiency by optimizing surface area to volume ratios (Harris and Theriot, 2018). Their lack of morphological variety has not been a limitation but rather a strength, allowing for extremely diverse genotypes encompassing enormous metabolic potential (Durot et al., 2008). Metabolic diversity and flexibility are essential components of global microbial ubiquity, but the mechanisms for acquiring the components that enable habitation of the extreme environments on Earth are not well understood. Evidence for HGT can be obtained from an extreme environment (Collins and Deming, 2013; Fuchsman et al., 2017; Moulana et al., 2020), but evaluations of the potential importance of this mechanism compared to its role in a moderate environment are rare.

Extreme environments are those with physical or chemical characteristics that are understood to limit the ability of life to function, such as temperatures near the freezing or boiling point of water (Harrison et al., 2013). The majority of inhabited environments on Earth are marine and cold, at temperatures below 5˚C which slow the rate of metabolic reactions (Price and Sowers, 2004). Across the climate-threatened cryosphere (below 0˚C) where water can remain liquid due

to increased concentrations of salt ions and other impurities (Cox and Weeks, 1983), microbial life remains ubiquitous (Boetius et al., 2015). Though extreme environments present unique challenges to life, microorganisms can be highly abundant in them (> 10^8 mL^{-1} in subzero brines; Cooper et al., 2019, Chapter 2), with niche competition continuing as an important selective pressure (Rapp et al., 2021).

A bacterial group globally distributed across the dominant cold and saline environments of Earth is the genus *Marinobacter* (Figure 4.1). A total of 45 genomes of species representatives of *Marinobacter* were publicly available at the start of this analysis. This genus presents a potentially useful model for examining evolutionary mechanisms responsible for adaptation to extreme environments because its various members thrive under conditions ranging from moderate to extreme, offering that comparison. Physiological studies have shown that this group of heterotrophic Gammaproteobacteria is metabolically diverse, capable of consuming a wide variety of organic compounds, including hydrocarbons, carbohydrates, and amino acids, and are broadly characterized as facultatively anaerobic (Gauthier et al., 1992; Green et al., 2006; Liu et al., 2012; Handley and Lloyd, 2013). The type species, *M. hydrocarbonoclasticus*, was isolated from hydrocarbon-polluted seawater near the mouth of an oil refinery (Gauthier et al., 1992). Other species have since been obtained from a variety of saline environments characterized by warm to extremely cold temperatures (precise *in situ* temperatures are not always reported in this literature). The species used in our study (listed in Table 4.1, with geographic source indicated on Figure 4.1) were originally obtained from seawater (temperate and polar, Shivaji et al., 2005; Vaidya et al., 2015), marine sediments (Montes et al., 2008; Han et al., 2017), hypersaline lakes (Bagheri et al., 2013; Zhong et al., 2015), deep-sea hydrothermal vent settings (Kaye et al., 2011;

Wang et al., 2012), deep-sea temperate and cold brine seeps (Antunes et al., 2007; Sun et al., 2020), deep subsurface saline deposits (Bonis and Gralnick, 2015), and the polar subzero habitats of sea ice (Canadian Basin; Zhang et al., 2008), hypersaline springs in permafrost (Axel Hieberg Island, Canadian High Arctic; Niederberger et al., 2010), Arctic sea ice (Zhang et al., 2008), and Antarctic brines from Lake Vida (Kuhn et al., 2014) and Blood Falls (Chua et al., 2018). This genus thus highlights an opportunistic lifestyle that enables life to persist and even thrive in saline and frequently cold environments, including those under the most extreme conditions in polar regions.

Recently, we identified *Marinobacter* as the dominant genus inhabiting cryopeg brines found near Utqiaġvik, Alaska (Cooper et al., 2019, Chapter 2). Cryopegs are liquid-saturated sediment layers within permafrost that remain unfrozen perennially at subzero temperatures due to their salt content (Gilichinsky et al., 2003; 2005). The cryopeg system we sampled is approximately 2 km from the modern Arctic coastline, and most likely derived from marine sediment that was exposed to subzero temperatures and incorporated into permafrost at least 40 ka BP based on radiocarbon dating (Meyer et al., 2010a; 2010b; Iwahana et al., 2021). For the last 40 ka, the cryopeg brines have been isolated from external influences, with observations indicating relatively stable subzero temperatures (–8 to –6˚C) and hypersalinities (109 to 140 ppt; Cooper et al., 2019, Chapter 2; Iwahana et al., 2021). The microbial communities inhabiting these subzero brines are dense (approximately 10^8 cells mL^{-1}), with *Marinobacter* accounting for, on average, a remarkable 49% of the bacterial community (relative abundance by 16S rRNA gene amplicon sequencing; Cooper et al., 2019, Chapter 2). Though not dominant in the sea ice brines we sampled, *Marinobacter* has been found to be abundant in Arctic sea ice, composing

approximately 22% of the Gammaproteobacteria present (Brinkmeyer et al., 2003). Although the *Marinobacter* spp. previously isolated from other subzero brines in Arctic and Antarctic settings were not reported as dominant *in situ* (Zhang et al., 2008; Niederberger et al., 2010; Kuhn et al., 2014; Chua et al., 2018), the common occurrence of *Marinobacter* spp. in subzero hypersaline environments in Earth's polar regions raises questions about the evolutionary history and adaptations of this genus that give it a competitive advantage under such extreme conditions.

Here, we use pangenomics (Tettelin et al., 2005; Tettelin et al., 2008; Delmont and Eren, 2018; Brockhurst et al., 2019) to examine the genomic potential of the genus *Marinobacter*, with a particular focus on newly obtained members from cryopeg brines. Pangenomics is a method of genomic analysis that compares the gene content of a set of closely related genomes, generally at the genus or species level and often with a focus on "core" and "variable" genes. It allows for exploration of the characteristics that unify a group as a whole and for discovery of characteristics that may distinguish a subset of the group. Pangenomics has been applied to help understand the ecology and evolution of many microbial taxa around the globe, including important taxa such as the globally distributed bacterial genus *Prochlorococcus*, a photosynthesizing genus cosmopolitan to ocean surface waters (Delmont and Eren, 2018). It has also been applied to *Sulfurovum*, a mesophilic chemolithoautotrophic genus that is ubiquitous in hydrothermal vent environments, to investigate the influences of biogeography and environmental conditions on genomic differentiation (Moulana et al., 2020). Motivating this study was the successful acquisition of new genomes of *Marinobacter* from the Alaskan cryopeg brines we have been studying (Cooper et al, 2019, Chapter 2; Iwahana et al., 2021; Rapp et al., 2021) using both classical cultivation methods and sample recovery of metagenome-assembled

genomes (MAGs), in the latter case using short-read (Illumina) and long-read (Nanopore) assembly approaches to maximize their quality. These genomes add novelty to the known diversity of this genus and allow a new focus on the cryosphere. Using the pangenome generated from these cryopeg genomes and the publicly available representative genomes of other reported *Marinobacter* species, we explored the evolutionary history, functional capabilities, and pangenomic diversity of this genus, including the potential role of horizontal gene transfer in adapting to the extreme conditions of subzero, hypersaline environments.

4.3 Materials and Methods

Isolation of Marinobacter *strains from cryopeg brines*

Cryopeg brines were collected near Utqiaġvik, Alaska, in May 2018 as previously described (Cooper et al., 2019, Chapter 2). Samples were frozen at –20˚C in the field laboratory and returned to the University of Washington (Seattle, WA, USA). Cryopeg brine, labelled as CBIW_18 (as described by Cooper et al., 2019, Chapter 2), was used as the inoculum for bacterial isolation, after thawing from –20˚C to –1˚C over several days. The brine was serially diluted to a final dilution of 10^{-9} in a variety of defined media, based on ONR7a, each containing a single organic substrate (L-alanine, L-leucine, Na-acetate, kerosene, and 1-butanol) at 10 mM concentrations. ONR7a is a minimal medium used in the isolation of bacteria that degrade hydrocarbons and other individual organic substrates (Dyksterhouse et al., 1995). ONR7a is composed of three solutions that are combined after preparation: 1) 22.79 NaCl, 3.98 g Na_2SO_4, 0.72 g KCl, 83.00 mg NaBr, 31.00 mg $NaHCO_3$, 27.00 mg H_3BO_3, 2.60 mg NaF, 0.27 g NH_4Cl, 89.00 mg Na_2HPO_4 x $7H_2O$, 1.30 g TAPSO in 500.00 mL H_2O with pH adjusted to 7.6 with NaOH; 2) 11.18 g $MgCl_2$ x $6H_2O$, 1.46 g $CaCl_2$ x $2H_2O$, 24.00 mg $SrCl_2$ x $6H_2O$, in 450.00 mL

H_2O; and 3) 2.00 mg $FeCl_2$ x $4H_2O$, in 50.00 mL H_2O. All of the dilutions were incubated at 2˚C for 66 days when growth was observed in all of the 10^{-7} dilutions and not in the 10^{-8} or 10^{-9} dilutions. From each of the 10^{-7} dilution cultures, 100 µL was transferred to solid plate media consisting of 1.2% agar and half-organic-strength Difco Marine Broth 2216 (Fisher, Waltham, MA). The salinity of the marine broth medium was amended to full seawater strength with artificial seawater composed of 24 g NaCl, 0.7 g KCl, 5.3 g $MgCl_2$ x $6H_2O$, $MgSO_4$ x $7H_2O$, and 1.3 g TAPSO in 1 L H_2O with pH adjusted to 7.5. Colonies were picked from agar plates and restreaked three times until a final colony was transferred to liquid Marine Broth for growth.

Screening of genus-level taxonomic identity of isolates (two from each ONR7a + organic substrate medium) was conducted using Sanger sequencing of the 16S rRNA gene. For 16S sequencing, cell pellets from liquid cultures were sent to GENEWIZ, Inc. (South Plainfield, NJ) where DNA was extracted and sequenced. Isolates that were confirmed as *Marinobacter* came from media containing alanine (n = 2), acetate (n = 1), and kerosene (n = 1) each as the sole organic carbon source. The two isolates from the alanine-amended medium were designated M1C and M2C, and the isolates from kerosene- and acetate-amended media were designated M3C and M4C, respectively.

Genome sequencing

Pelleted cultures of the four selected isolates were extracted and sequenced at the Microbial Genome Sequencing Center (MiGS; Pittsburgh, PA, USA). DNA extraction, library preparation, genome sequencing, quality control, and assembly were performed at MiGS. Briefly, the genomes were sequenced using both Illumina short-read and Oxford Nanopore Technologies

(ONT) long-read sequencing technologies. Quality control of the raw Illumina and ONT reads was performed using bcl2fastq (Illumina) and Porechop (Wick et al., 2017a), respectively. Hybrid assembly of the reads was conducted using Unicycler (Wick et al., 2017b). After receiving the assembled genomes, we checked genome completeness using CheckM v1.1.2 (Parks et al., 2015), and assigned the taxonomy of each genome using GTDB-tk v1.1.1 (Parks et al., 2018). Gene calling was performed in anvi'o v6.1 (Eren et al., 2015) using Prodigal v2.6.3 (Hyatt et al., 2010) to identify open reading frames (ORFs). Each ORF was annotated with National Center for Biotechnology Information (NCBI) Clusters of Orthologous Genes (COG) annotations (Tatusov et al. 1997, Galperin et al., 2019) which were assigned using DIAMOND v0.9.22.123 (Buchfink et al., 2014). These genome sequences will be available publicly after journal submission and peer review.

Metagenome sequencing, assembly, and binning

DNA from cryopeg brine samples (n = 5), concentrated onto 0.22 μm Sterivex filters in the field and kept frozen at –80°C until processing, was extracted using the DNeasy PowerSoil kit (QIAGEN, Germantown, MD) as previously described (Cooper et al., 2019, Chapter 2). Samples collected for DNA extraction for this study are identical to samples CBIW_17, CBIW_18, CBIA_18, and CB1_18 described by Cooper et al. (2019, Chapter 2). Samples CBIA_18 and CB1_18 are referred to as CBIA and CB1 in this study since they are not temporally replicated. Metagenomic sequencing was conducted using both Illumina short (2 x 150 bp) and ONT long reads. The Illumina sequencing library was prepared, sequenced, and analyzed at the DOE Joint Genome Institute as previously described (Rapp et al., 2021), and the ONT sequencing library was prepared, sequenced, and quality-controlled following the VirION2 pipeline (Zablocki et al.,

2021). Long reads were first assembled using Flye v2.5 (Kolmogorov et al., 2019), and short reads were subsequently used for error correction using Pilon v1.23 (Walker et al., 2014). Each metagenome was assembled from a single sample of cryopeg brine, allowing for recovery of MAGs from discrete samples. The raw Illumina metagenomic reads were produced previously and are described in detail by Rapp et al. (2021) and are available through IMG/M under accession numbers 3300031836 (CBIW17), 3300032129 (CBIW18), 3300032135 (CBIA), and 3300034171 (CB1); the hybrid metagenomic assemblies generated here will be available publicly after journal submission and peer review.

Assembled metagenomic contigs were processed using anvi'o v6.1 (Eren et al., 2015). Gene calling for the metagenomes was performed in anvi'o as described above for the whole isolate genomes. Bacterial and Archaeal single copy universal marker genes (Lee, 2019) were identified using HMMER v3.3 (Eddy, 2011). Automated binning was conducted using MetaBAT 2.0 v2.15 (Kang et al., 2019), MaxBin2 v2.2.7 (Wu et al., 2016), and Concoct v1.1.0 (Alneberg et al., 2014). The best bins from each tool were automatically selected using DAS Tool v1.1.2 (Sieber et al., 2018). After binning, the bins were manually refined in anvi'o using coverage, gene clustering, and GC content as guides (https://merenlab.org/data/refining-mags/, accessed 19 May 2020). After binning, each MAG was assigned taxonomy using GTDB-tk v1.1.1 (Parks et al., 2018), and MAG completeness was estimated using CheckM v1.1.2 (Parks et al., 2015). The MAGs in this study are referred to as CBIW17, CBIW18, CBIA, and CB1 referencing the corresponding original sample name. The MAG sequences will be available publicly after journal submission and peer review.

Pangenomic analysis

Along with the four MAGs (CB1, CBIA, CBIW17, and CBIW18) and four isolate genomes (M1C, M2C, M3C, and M4C; isolated from sample CBIW_18 described in Cooper et al., 2019, Chapter 2) we obtained for this study, 45 representative genomes of *Marinobacter* spp. were acquired from the NCBI Assembly database (on 16 September 2020) as genomic FASTA files. References and accession numbers for each genome are listed in Table 4.1. We followed the standard pangenomic workflow in anvi'o v6.1 (Delmont and Eren 2018) to build and analyze the *Marinobacter* pangenome from 53 total genomes. We created a contig database for each isolate and representative genome using the command "anvi-script-FASTA-to-contigs-db"; this script performs gene calling as described in the sections above. Each genome was annotated with NCBI COGs as described above. Original metagenomic contig databases were used for each MAG. We combined all genomic data by creating a genome storage database using the command "anvi-gen-genomes-storage" and specifying paths to internally stored contig databases for the MAGs (using data already processed at the metagenome level) and external paths to the rest of the *Marinobacter* contig databases. Then, the *Marinobacter* pangenome was assembled from the genome storage database using the command "anvi-pan-genome" with the "--use-ncbi-blast" flag which directs the program to use BLASTp (Altschul et al., 1990) for calculating amino acid similarities between related gene calls. Using the default settings, the "anvi-pan-genome" program employed a Markov Cluster Algorithm (MCL; van Dongen and Abreu-Goodger 2012) which clustered highly similar gene calls using the BLASTp search results to allow for assessment of gene overlap between individual genomes. The pangenome database produced contains information on gene cluster presence/absence and abundance for each genome and gene cluster similarity across all genomes.

Phylogenetic inference

To assess the phylogenetic relationships among the members of the *Marinobacter* genus, we constructed phylogenetic trees using three approaches. In the first, we calculated average nucleotide identity (ANI) across the *Marinobacter* pangenome using the anvi'o program "anvi-compute-genome-similarity" which employed the program pyANI v0.2.10 (Pritchard et al., 2016) using the ANIb option. The ANIb option uses BLASTN+ to align 1020 nt fragments of input data and calculate similarity. The program produced a matrix of ANI values for shared genomic regions, and Anvi'o produced a hierarchically clustered phylogenetic tree using Euclidean distance and Ward clustering methods by default, which we call the "ANI" phylogenetic tree.

To construct the two additional phylogenetic trees, which we call the "core" and "rooted universal" trees, we used, respectively, a set of 108 single copy genes found in all genomes from the *Marinobacter* pangenome (the "core" genes) and a suite of universal bacterial single copy marker genes (Lee, 2019) found in all genomes across the pangenome, with the model psychrophilic bacterium *Colwellia psychrerythraea* 34H (Methé et al., 2005) as an outgroup. To generate the core phylogenetic tree, genes were selected from the *Marinobacter* pangenome using the program "anvi-get-sequences-for-gene-clusters" filtering for gene clusters that have a maximum functional homogeneity of 0.90 and a minimum geometric homogeneity of 0.95 that occur only once in each genome and that occur in a minimum of 53 of the genomes. The "--concatenate-gene-clusters" flag was used to specify the output format appropriate for phylogenetic tree generation. To generate the rooted universal tree, the genome for *C. psychrerythraea* 34H was first downloaded from NCBI, then processed similarly to all other

genomes in the *Marinobacter* pangenome. HMM hits from all *Marinobacter* genomes and the *Colwellia* genome were aggregated using the program "anvi-get-sequences-for-hmm-hits" and specifying "--hmm-source Bacteria_71" and that the genes occur in at least 54 of the input genomes. The flags "--return-best-hit", "--get-aa-sequences", and "--concatenate" were used to get the best possible protein sequences for phylogenetic tree construction. Both the core and rooted universal phylogenetic trees were constructed using FastTree v2.1.10 (Price et al., 2010), which uses the Jones-Taylor-Thorton evolutionary model and Shimodaira-Hasegawa test (1,000 resamplings) to optimize tree topology. All trees were visualized using FigTree v1.4.4 (Rambaut, 2009). Clades of *Marinobacter* were defined operationally based on the branching patterns of the core phylogenetic tree.

Principal component analysis

We conducted a principal component analysis (PCA) by using the number of genes belonging to each gene cluster in each genome in the pangenome. The PCA was conducted in python v3.7.8 using scikit-learn v0.23.2 (Pedregosa et al., 2011) and pandas v1.1.2 (McKinney, 2010; Reback et al., 2020). Two components were used for the PCA analysis. The resulting data were plotted and visualized using seaborn v0.11.0 (Waskom, 2021). Custom scripts for processing and visualization will be available publicly after journal submission and peer review.

Metabolic inference

Metabolic pathway completeness was measured for each genome in the pangenome using KofamScan v1.3.0 (Aramaki et al., 2020) to assign Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthologs (KOs) to each gene. Protein sequences for each gene were recovered for each

genome using the anvi'o program "anvi-get-sequences-for-gene-calls" with the "--get-aa-sequences" flag. KofamScan was run for each genome using KO profiles and lists downloaded on 19 January 2021. KO annotation results were aggregated and analyzed using KEGG-Decoder v1.2.2 (Graham et al., 2018). A heatmap of the results was produced using a custom python script.

Functional enrichment based on COG annotations for each clade was assessed using the "anvi-get-enriched-functions-per-pan-group" program in anvi'o (Eren et al., 2015). The program uses a generalized linear model and applies the logit link function to compute an enrichment score and p-value for each gene cluster in the pangenome. The program then applies a false detection rate correction to the p-values to obtain a q-value for inference of significance (Storey and Tibshirani, 2003).

Detection of horizontal gene transfer

We identified putative horizontally transferred genes using HGTector v2.0b3 (Zhu et al., 2014). Briefly, this program searches each gene in a genome, compares it to a taxonomy-informed database, and assesses the distribution of best hits to closely and more distantly phylogenetically related taxa to determine whether a gene is likely the result of vertical inheritance or HGT. To accomplish this task, HGTector uses a database consisting of NCBI RefSeq protein sequences; we downloaded the pre-built default database of reference and representative microbial genomes (dated as of 21 October 2019) found on the program manual website. After downloading and decompressing the database, it was compiled using DIAMOND v2.0.6.144 (Buchfink et al., 2014). As input for HGTector2, we provided gene calls from each genome in the *Marinobacter*

pangenome, the same as used for metabolic inference above. The "hgtector search" function was run on each genome individually using DIAMOND to score each gene against the database. After searching for best gene matches, "hgtector analyze" was run using a folder of all search results as input. For the analysis, the "--self-tax" was set to "2742" (NCBI taxon ID for the genus *Marinobacter*), the "--self-rank" was set to "genus", the "--self-low" flag was set to "low", the "--close-tax" was set to "72275" (NCBI taxon ID for the family Alteromonadaceae), and the "--bandwidth" method was set to "grid". These settings were implemented in order to specify that best gene hits to *Marinobacter* or within Alteromonadaceae are not counted as horizontally transferred and to carefully identify cutoff for HGT determination. HGTector2 output was combined with NCBI COG (Galperin et al., 2019) annotations for each gene call assigned during pangenome assembly. Genes that were not annotated with COGs but scored as putative horizontally transferred genes were counted as "None". The output data was analyzed and visualized using custom python v3.7.8 scripts.

4.4 RESULTS

Marinobacter sp. nov. genome characteristics

Following the successful cultivation of new *Marinobacter* spp. and recovery of high-quality MAGS from our subzero cryopeg brines, we first examined the basic characteristics of these new genomes for comparative purposes (Table 4.2). For three of the new isolates, M1C, M2C, and M4C, each genome assembled into a single contig with lengths ranging from 4,745,310 to 4,747,818 bp, while M3C assembled into one large contig (4,762,857 bp) and three smaller contigs (1,983 to 210,472 bp). All four of these genomes were scored as 100% complete with CheckM (Parks et al., 2015). The four MAGs generated from the cryopge brines (CBIW17,

CBIW18, CBIA, and CB1) ranged in length from 3,085,923 to 4,242,133 bp and in completeness from 79 to 92% complete as scored by CheckM (Parks et al., 2015). GC content ranged from 54.25 to 54.45% for all eight of these novel genomes. The isolate genomes encoded 4,395 to 4,687 genes. Across the *Marinobacter* pangenome, the mean (± standard deviation) genome length was 4,112,097 ± 444,972 bp, GC content was 56.94 ± 2.15%, and number of genes was 3,826 ± 418. Genome statistics for all genomes included in the pangenome are provided in Table 4.2.

Phylogeny and genomic similarity of the genus

To evaluate the evolutionary relatedness of the newly generated genomes of *Marinobacter* to other members of the pangenome presented here, we used three methods of phylogenetic inference. To follow patterns of phylogeny in the pangenome, we assigned each genome to a clade, where clades were based on branching patterns of the core phylogenetic tree (based on single copy core genes; Figure 4.2B). This unrooted tree, consisting of the 108 genes found in gene clusters that had a geometric homogeneity score of at least 0.95 and a functional homogeneity of less than 0.90, ensures a high level of both sequence alignment and functional diversity. For comparison, we constructed a rooted universal tree using a subset of the universal bacterial single copy marker genes (Lee, 2019) that were found in every genome in the pangenome, with *Colwellia psychrerythraea* 34H as the outgroup (Figure 4.2A). A maximum likelihood phylogenetic tree was also constructed based on ANI values using Euclidean distance and Ward clustering (Figure 4.2C). The rooted universal phylogenetic tree (Figure 4.2A) indicates that Clade III, which consists of the novel *Marinobacter* genomes from cryopeg brines and the genomes from *Marinobacter* spp. previously isolated from subzero brines in the Arctic

(Zhang et al., 2008) and Antarctic (Chua et al., 2018), branches before the rest of the genus. Clade III is paraphyletic in the rooted universal tree and monophyletic in both core (Figure 4.2B) and ANI (Figure 4.2C) trees. Clades IV, VI, VII, and VIII also consistently cluster monophyletically, while clades I, II, and V cluster inconsistently across trees. Clade assignments generally correspond with the environment of isolation of each species representative, particularly for Clades I and III (Table 4.1).

To assess relatedness of *Marinobacter* members, ANI was calculated between each genome in the pangenome based on sequence similarity of shared gene content. ANI and alignment coverage values for the pangenome can be found in Supplementary Tables S4.1 and S4.2 (available upon publication or request). The mean (± standard deviation), median, minimum, and maximum percent identity of shared genes across the pangenome were 77.1 (± 6.1), 75.2, 72.7, and 100%, respectively. Figure 4.3 displays a heatmap of ANI across the pangenome. The novel isolate genomes and MAGs in Clade III have a mean ANI of 99.5 (± 0.47)% to each other, 90.7 (± 1.2)% to *M. psychrophilus*, and 93.1 (± 1.6)% to *M. gelidimuriae*. Clade III has a mean ANI of 74.7 (± 0.81)% to all other genomes. Alignment coverage is low and highly variable across the pangenome with a mean of 44.0 (± 16.4)%, median of 39.5%, minimum of 19.1%, and maximum of 100%. Histograms of the distribution of percent identity and alignment coverage scores can be found in Supplementary Figure S4.1 (available upon publication or request).

Ribosomal genes recovered from the genomes, each isolated from different field samples, indicate that the novel cryopeg isolates all share a 100% identical 16S rRNA gene. The closest species relatives by 16S rRNA gene similarity to these isolates are *M. gelidimuriae* and *M.*

psychrophilus with 99.29% and 99.09% similarity, respectively. The remaining species of *Marinobacter* have a 16S rRNA gene similarity to the isolates of ≤ 97.22%. Alignment coverage is generally low between species of *Marinobacter* (Supplementary Figure S4.1), and the cryopeg genomes have an alignment coverage of 61.4% to *M. gelidimuriae* and 71.3% to *M. psychrophilus*.

Scope and distribution of the Marinobacter pangenome

Pangenomic analysis allows for an evaluation of conserved genes within a group of closely related genomes. We used this analysis to explore the diversity of genetic pathways available to this genus and to identify the genomic features that might set apart the newly produced genomes from the rest of the genus. The *Marinobacter* pangenome constructed here consists of all of the genes found across 53 genomes: 45 species representatives for the genus, 4 newly sequenced genomes from isolates of *Marinobacter* from cryopeg brines (M1C, M2C, M3C, and M4C), and 4 MAGs assembled from metagenomes sequenced from cryopeg brines (CBIW17, CBIW18, CBIA, and CB1). Isolation environments for these *Marinobacter* spp. are distributed across the globe, primarily in marine or terrestrial hypersaline settings (Figure 4.1). As the pangenome was constructed, genes from each genome were clustered to identify shared content between genomes. The pangenome contained a total of 24,580 gene clusters containing 202,751 total genes across all 53 genomes. Here, the distribution of gene content in the pangenome is described by the number of gene clusters shared by a range of genomes, which were generally divided into decreasing sets of ten (Figure 4.3). Across all or almost all (50–53) *Marinobacter* genomes, 1,465 gene clusters were shared, approximating the "core" of the pangenome. With decreasing sets of genomes, 579 gene clusters were shared across 40–49 genomes, 532 across

30–39 genomes, 561 across 20–29 genomes, 1,243 across 10–19 genomes, and 7,428 across 2–9 genomes. The largest number of gene clusters, 12,772, appeared in only a single genome (singletons). Gene cluster annotations and genome contributions can be found in Supplementary Table S4.3 (available upon publication or request). Genes involved in translation, ribosomal structure and biogenesis (COG category J) are the most abundant category and comprise 13.04% of the gene clusters in the core range of the pangenome. Category J gene clusters diminish in abundance consistently into the regions of the pangenome where fewer genomes share genes, and comprise only 0.70% of the singleton gene clusters. Mobilome genes (COG category X) are not abundant in the core but increase in abundance in the regions of the pangenome with fewer genomes sharing genes, making up 0.07% of the core and 2.03% of all singleton gene clusters. The number of gene clusters without functional annotations increases steadily from 4.78% of the core to 56.9% of the singleton gene clusters (Supplementary Table S4.3; available upon publication or request).

The PCA identified differences between members of the pangenome based on gene content (Figure 4.4; Supplementary Table S4.4). Principal component 1 (PC1) explains 12.3% of the variance in the ordination, with PC2 explaining 6.7%. The ordination shows a separation of the novel genomes from cryopeg brines (in Clade III) along PC1, while the majority of the other members of the genus cluster tightly together.

Metabolic potential of Marinobacter *spp.*

Marinobacter spp. are known from physiological and environmental studies to perform a broad suite of metabolic functions (Handley and Lloyd, 2013), but a complete survey of metabolic

potential encoded in the genomes of *Marinobacter* has not been published. We performed this survey to assess the variety of metabolisms that *Marinobacter* spp. are genetically capable of utilizing and to identify metabolisms likely relevant to inhabiting in cryopeg brines. Metabolic pathway completeness was assessed by assigning KOs to each gene in each genome and then using KEGG Decoder (Graham et al., 2018) to assess the proportion of KOs required for a pathway to be present. Values for KEGG pathway completeness for all genomes in the *Marinobacter* pangenome are listed in Supplementary Tables S4.5 and S4.6; a subset, where at least one genome has 50% of a pathway complete, is displayed in Figure 4.5. Here, we highlight metabolic pathways that appear as important to the entire genus of *Marinobacter* and those that represent features unique to Clade III. Nearly all *Marinobacter* genomes (except only *M. gelidimuriae*) encode the genes required for the Entner-Doudoroff Pathway, while none of the genomes encodes a complete glycolysis pathway. All *Marinobacter* genomes encode for the glyoxylate shunt. Almost all *Marinobacter* spp. encode metabolisms for each of the twenty standard amino acids. An exception is the metabolism of asparagine, which is nevertheless encoded by 41 of the 53 genomes in the pangenome.

Clade III genomes (with the exception of the MAGs which are not complete genomes) encode complete pathways (as defined by Graham et al., 2018) for nitrite oxidation and reduction, dissimilatory nitrate reduction, nitric oxide reduction, and nitrous oxide reduction. Pathways for dissimilatory nitrate reduction to ammonium are not present in Clade III genomes, and no nitrogen fixation pathways are present in any *Marinobacter* genome. The potential to oxidize thiosulfate appears in the cryopeg genomes, while the pathways involved in this process are less common in other Clade III genomes (*M. psychrophilus* and *M. gelidimuriae*) or in any other

clades of *Marinobacter*. Whereas all *Marinobacter* genomes encode for the copper transporter CopA, and nearly all encode for the ATP-binding cassette (ABC)-type substrate-binding protein AfuA for ferric iron (the exceptions are three of the cryopeg MAGs and *M. similis*), only a subset of genomes encode for the cobalt transporter CorA (with most of Clade III except *M. gelidimuriae*) and several encode for the ferrous iron transporter FeoB (including M1C, M2C, M4C, and *M. psychrophilus*). Cryopeg isolates M1C, M2C, and M4C and MAG CB1 each encode a complete pathway for Type I secretion systems, while this system is mostly absent in nearly all other *Marinobacter* genomes. Type II and Type VI secretion systems are nearly complete for the majority of *Marinobacter* genomes, while Type III and Type IV secretion systems appear to be absent from the pangenome. Clade III genomes M1C, M2C, M4C and *M. psychrophilus* encode complete pathways for retinal biosynthesis which involves carotenoid and rhodopsin biosynthesis; this pathway is less complete in other Clade III genomes and nearly absent in most other *Marinobacter* genomes. Clade III genomes (except for some MAGs and *M. gelidimuriae*) encode for nearly complete carbon-phosphorous (C-P) lyase system. Clade III genomes also encode a nearly complete pathway for polyhydroxybutyrate synthesis, which has variable levels of completion for all other genomes in the pangenome.

Enrichment of transporters and transposases in subzero brine Marinobacter *spp.*

In a continuation of our effort to identify distinguishing features of the newly acquired *Marinobacter* genomes, we conducted an analysis using anvi'o to detect which gene clusters, annotated with COG functions, were enriched in each clade relative to their presence in other clades. This analysis revealed a variety of genes that were enriched solely in Clade III, the top 17 of which have an adjusted q-value < 0.01. The top three functions enriched in Clade III are the

permease, periplasmic, and ATPase components of the ABC-type uncharacterized transport system YnjBCD (COG4134, COG4135, and COG4136). Other functions enriched solely in Clade III include stress response protein SCP2 (COG2310), bacteriorhodopsin (COG5524), tellurite resistance protein TerB (COG3793), the 4Fe-4S-binding domain of the Fe-S cluster biogenesis protein NfuA (COG0694), the gamma subunit of sarcosine oxidase (COG4583), recombination-promoting DNA endonuclease RpnC/YadD (COG5464; previously annotated as a predicted transposase), and a putative component of the toxin-antitoxin plasmid stabilization module (COG3657). The complete results of this functional enrichment analysis are provided in Supplementary Table S4.7.

Horizontal gene transfer

We searched each gene in each genome for signs of HGT using HGTector (Zhu et al., 2014), which employs a statistical approach to estimate the likelihood of a gene being derived from a close or distant relative that would indicate its acquisition via HGT. The patterns of HGT are of particular interest to life in cryopeg brines given their biogeographical isolation and constant exposure to selective and stressful conditions. We calculated the proportion of genes in each genome (percentage of total genes) that were scored as putative HGTs (Figure 4.6; see Supplementary Table S4.8) and used this information along with NCBI COG annotations to score total proportion of HGT for each COG category in each genome. The mean proportion of total genes derived from HGT for the entire pangenome is 6.9 (± 3.7)%, the median is 5.3%, the minimum is 2.5%, and the maximum is 16.4%. Clade III alone contained a mean of 12.5 (± 2.9)% total genes derived from HGT; all other clades had lower means, in descending order: 8.8

(± 2.9)% for Clade I, 5.5 (± 2.9)% for Clade IV, 4.9 (± 2.9)% for Clade II, 4.8 (± 2.9)% for Clade VIII, 4.5 (± 2.9)% for Clade VII, 4.2 (± 2.9)% for Clade V, and 3.6 (± 2.9)% for Clade VI.

Compared to other clades, Clade III had notably high proportions of putative HGTs from COG categories for transcription (K), carbohydrate transport and metabolism (G), amino acid transport and metabolism (E), energy production and conversion (C), and the mobilome (X), particularly prophages and transposons. These categories each contribute a mean proportion of total genes between 0.87 and 1.26% (Figure 4.6). HGT-derived genes without a COG annotation comprise a mean proportion of 0.87% of total genes for Clade III genomes. The most abundant HGT-derived genes found in the new cryopeg genomes were components of the TRAP-type C4-dicarboxylate transport system (COG1593, COG1638, and COG3090; category G), LysR family DNA-binding transcriptional regulators (COG0583; category K), and transposases (COG2801 and COG3039; category X). For other Clade III genomes, the most abundant HGTs found in the *M. gelidimuriae* genome were transposases (COG1943; category X), 5-methylcytosine-specific restriction endonuclease McrA (COG1403; category V), and site-specific recombinase XerD (COG4974; category L), while the most abundant HGTs found in the *M. psychrophilus* genome were DNA-binding transcriptional regulators from the LysR and GntR families (COG0583 and COG1802; category K), an ABC-type amino acid transport system (COG0765; category E), and components of the TRAP-type C4-dicarboxylate transport system (COG1593, COG1638, and COG3090; category G).

4.5 DISCUSSION

The Marinobacter pangenome

The phenotypically diverse genus *Marinobacter* demonstrates the benefits of metabolic versatility to evolutionary success in saline environments (Handley and Lloyd, 2013). Our analysis of the *Marinobacter* pangenome, including complete genomes from novel isolates and MAGs from the extreme cryopeg brine environment, as well as publicly available representative genomes of *Marinobacter* species, shows the immense diversity of genes encompassed by this taxon (Figure 4.3). The majority of the 24,580 detected gene clusters occurred as singletons (52% present in only one genome) among the 53 genomes analyzed here, and an additional 30% of gene clusters could only be detected in 2–9 genomes, illustrating the tremendous diversity of genes in the *Marinobacter* pangenome as it currently exists. While the majority of singleton genes are of unknown function, genes related to the mobilome are more abundant among the singleton genes than in the core, where they are negligible, and thus reflect the role of genome modification processes that allow each genome to contain a unique set of genes. We approximated the core region of the pangenome to include gene clusters that are found in 50–53 of the genomes to avoid bias against core genes that may not be complete in the MAGs. This approximated core, which includes genes involved in common metabolic processes and genomic utility, contained < 6% (1,465 gene clusters) of the total gene clusters. By comparison, a study of the pangenome of *Prochlorococcus* indicated that this genus of cyanobacteria has a diverse pangenome with singletons and core constituting approximately 30 and 10% of all gene clusters, respectively (Delmont and Eren, 2018). The *Prochlorococcus* pangenome is discussed as having a distribution of genes that correlates to the widespread success of this genus in occupying different regimes of light availability in the ocean (Delmont and Eren, 2018). Although

Marinobacter is not as clearly differentiated into ecotypes as *Prochlorococcus* (Biller et al., 2014), the even greater diversity and open structure of the *Marinobacter* pangenome helps to explain the successful habitation of this genus in a wide variety of environments.

Biogeography and phylogeny of Marinobacter

Plotting the isolation locations of the *Marinobacter* species discussed here (Figure 4.1; Table 4.1) reveals that they are globally distributed and derived primarily from coastal marine environments. A few exceptions to the coastal marine setting include *M. halophilus* (Zhong et al., 2015) and *M. persicus* (Bagheri et al., 2014) isolated from terrestrial hypersaline lakes in China and Iran, respectively, and *M. subterrani* (Bonis and Gralnick, 2015) isolated from iron-rich saline fluids below the continental surface in Soudan Mine, Minnesota, USA. Species of *Marinobacter* have also been isolated from anthropogenic nearshore, saline environments, including oil-polluted wastewater, the source of *M. gudaonensis* (Gu et al., 2007) and *M. shengliensis* (Luo et al., 2015), wine-barrel-decalcification wastewater, the source of *M. vinifirmus* (Liebgott et al., 2006), and salted anchovies, which yielded *M. piscensis* (Hedi et al., 2015). The environments of isolation also generally correspond to their clade assignments in this study (Table 4.1), though some associations are clearer than others. Clades I and IV contain isolates primarily from marine sediments; Clade II, from salterns and estuaries; Clade III, from subzero brines; Clades V and VIII, from a mix of deep and hydrothermal and seawater environments; and Clades VI and VII, from seawater (Figure 4.1; Table 4.1). Exceptions to these descriptions of clade environments exist (e.g., pollution- and protist-associated isolates are dispersed throughout), but generally, *Marinobacter* spp. appear to occupy the niche for opportunistic heterotrophs in nutrient-rich saline environments across the globe.

The phylogenetic tree based on a concatenation of core genes from the pangenome (Figure 4.2B) represents a comprehensive genomic perspective for assessing genome relatedness in the pangenome. In using this method to assign the genomes into clades, we were able to recognize patterns of species distributions by environment or common genomic characteristics. In this core tree (Figure 4.2B), the novel cryopeg genomes clustered together in Clade III with the two other genomes from subzero brine environments (Figure 4.1): *M. psychrophilus*, isolated from sea ice in the Canadian Basin of the Arctic Ocean (Zhang et al., 2008), and *M. gelidimuriae*, isolated from subglacial brine flowing from Blood Falls in Antarctica (Chua et al., 2018). This initial grouping of the novel genomes along with other *Marinobacter* genomes from subzero brines of both polar regions suggests that environmental setting plays a role in the differentiation of this genus.

The traditional rooted universal phylogenetic tree (Figure 4.2A), constructed to validate organization of the core tree (Figure 4.2B), upheld the clustering of subzero brine *Marinobacter* genomes in Clade III. It also showed that this group branches from the rest of the genus earlier than any other clade. Within Clade III, the MAGs form their own early branch on the rooted universal tree (Figure 4.2A), though this feature may be an artifact of their reconstruction from natural samples that contain a greater sequence diversity from the *in situ* population than the genomes from clonal isolates. In this tree, except for Clades I and II, the other clades mostly remained intact, although neighboring clade organization differed between the trees. A third tree based on ANI scores (Figure 4.2C) also supports the distinct clustering and separate branching of Clade III from the rest of the genus. The support of Clade III by these multiple methods indicates

that this branch of subzero brine-dwelling *Marinobacter* spp. evolved from a common lineage, and that they have successfully and independently occupied this niche through competitive survival, if not growth, under extreme conditions over geological time.

The consistent clustering of Clade III genomes, along with their shared origins in subzero brines from both polar regions (Figure 4.1), raises questions about the relatedness of this clade to the rest of the genus. Using ANI (Supplementary Table S4.1), we observed that the new genomes from cryopegs are all highly related (> 99%), though the MAGs might vary more with higher completion. The closest relatives to the cryopeg genomes are *M. gelidimuriae* and *M. psychrophilus* (also in Clade III) with ANI scores of approximately 93.1% and 90.7%, respectively. A threshold of 95–96% has been suggested as a cutoff for species delineation by ANI, along with a cutoff of 98.65% similarity of the 16S rRNA gene (Kim et al., 2014). *M. gelidimuriae* and *M. psychrophilus* have alignment coverage (AC) values of 61.4% and 71.3% with the cryopeg genomes, which is notably higher than the AC of the rest of the *Marinobacter* genomes to the Clade III genomes (Supplementary Table S4.2). *M. gelidimuriae* and *M. psychrophilus* also share 99.29% and 99.02% 16S rRNA gene similarity with the novel cryopeg genomes, respectively. Though the 16S similarity is high relative to both of these species, *M. gelidimuriae* and *M. psychrophilus* themselves are sufficiently distinct by ANI and 16S similarity (Chua et al., 2018) to delineate them as separate species. Given that the ANI scores of the new cryopeg isolates fall below the defined cutoff and their AC to the closest relative species is low, we suggest that, at a minimum, these new isolates represent a novel *Marinobacter* species.

As a group, Clade III *Marinobacter* genomes have much higher ANI with each other than with other *Marinobacter* species; i.e., ≥ 90.6% ANI between Clade III members versus 72.7–76.4% ANI between Clade III members and other clades (Figure 4.3; Supplementary Table S4.1). Along with low ANI between *Marinobacter* spp. (median ANI of approximately 75%), we also observed generally low AC values between all species, with the majority ranging between approximately 30 and 50% (Supplementary Figure S4.1). The variability in ANI and AC observed here for the *Marinobacter* genus falls within the variability discussed for defining genus demarcation (Barco et al., 2020). However, the early branching of clade III in the rooted phylogenetic tree (Figure 4.2A), along with the high level of ANI between Clade III species and lower levels of ANI between Clade III and all other species (Figure 4.2C; Figure 4.3), calls into question whether this clade truly belongs within the genus *Marinobacter*. We suggest that Clade III may represent a novel genus of Bacteria, related to but separate from *Marinobacter*.

Metabolic potential of the Marinobacter *pangenome*

From previous studies that have explored the physiological diversity of cultured species of *Marinobacter* (Handley and Lloyd, 2013), the genus is well known for its ability to metabolize a large variety of hydrocarbons, carbohydrates, and amino acids as sources of carbon (Gauthier et al., 1992; Green et al., 2006; Liu et al., 2012; Handley and Lloyd, 2013). Few of the species are considered strict aerobes, but several are facultatively anaerobic, with an ability to use alternative electron acceptors, such as nitrate or fumarate (Takai et al., 2005), and to use arsenic, iron, and manganese for redox cycling (Handley et al., 2009; Wang et al., 2012). In this study, we explored the functional potential of the genus using COG and KEGG annotations. Each gene cluster in the pangenome was annotated with a COG, and the number of copies of each gene in

each genome contributing to that cluster was determined. A principal component analysis of the gene cluster COG annotations (Figure 4.4) demonstrated that the majority of *Marinobacter* genomes share similar functional potential and cluster together; however, a signal of differentiation by Clade III is observed along PC1 where each member of this clade falls outside of the main cluster of the genus. The MAGs fall farthest from the main cluster, which may be explained by their inclusion of genes from the *in situ* population of *Marinobacter* and by the incomplete nature of their assembly and potential lack of genes found abundantly in the main group of genomes. The results of the PCA further demonstrates the distinction of this clade and adaptive potential to the subzero brine environment that does not appear to be present in the majority of the genomes from other environments.

After annotating each gene in each genome with KOs and processing further to determine which metabolisms were available for each species of *Marinobacter* (Figure 4.5; Supplementary Tables S4.5 and S4.6), we found that the glycolysis pathway was not complete for any species of *Marinobacter*. Rather, the Entner-Doudoroff (ED) pathway, an alternative to glycolysis, was complete in all but one of the genomes (*M. gelidimuriae*). The ED pathway has been discussed as an adaptive alternative to glycolysis, particularly for psychrophilic bacteria (Czajka et al., 2018), because it requires less total protein mass and fewer temperature-constrained enzymatic reactions to convert glucose to pyruvate (Czajka et al., 2018), even though it nets one less ATP (Conway, 1992). Additionally, each genome encodes the glyoxylate shunt (GS) pathway, which is an alternative to the tricarboxylic acid (TCA) cycle, using five of the eight enzymes in the TCA cycle and two unique enzymes (net one less enzyme). The GS bypasses the carbon dioxide-producing steps of the TCA cycle and has been hypothesized to play a role in oxidative stress

response. It is also essential for acetate and fatty acid metabolism in bacteria (Ahn et al., 2016), both of which have been observed as organic substrates for *Marinobacter* (Handley and Lloyd, 2013). The presence in all genomes of complete metabolic pathways for each of the 20 common amino acids (with the exception of asparagine metabolism, for which some genes are missing in 12 of the genomes) suggests that *Marinobacter* has a competitive advantage over other heterotrophic psychrophiles with more limited metabolisms to meet their nitrogen requirements. As most *Marinobacter* species also encode flagella and chemotactic pathways, they appear to have the capacity to actively seek such labile organic resources (as has been demonstrated for *Colwellia psychrerythraea* 34H under subzero hypersaline conditions; Showalter et al., 2018). The genetic potential to seek resources and utilize a broad range of organic compounds are clear indicators of the diverse and opportunistic lifestyles for which *Marinobacter* spp. are known, while the discovery of streamlined metabolic pathways despite lower energetic yield highlights the metabolic tradeoffs for survival in an extreme environment.

Distinct characteristics of the subzero brine clade

Clade III genomes encoded several KEGG pathways that were less well represented in the rest of the pangenome and may indicate adaptation to their extreme environment (Figure 4.5). Among these are nitrogen and sulfur oxidation and reduction pathways, metal transport pathways, and Type I secretion systems. The availability of this variety of redox reactions adds support to previous observations of the facultatively anaerobic lifestyle of *Marinobacter* (Takai et al., 2005; Handley et al., 2009; Handley and Lloyd, 2013; Chua et al., 2018) and fits with models of redox potential available in anoxic marine sediments (Zobell, 1946; Lam et al., 2019), which are likely the source material for cryopegs before incorporation into permafrost (Iwahana et al., 2021).

Clade III *Marinobacter* spp. also appear well equipped for ion and metal transport across the membrane, befitting the concentrated ionic and likely anoxic or hypoxic environments of subzero brines. Transport mechanisms are particularly important in a brine setting where osmolarity and metal toxicity need to be closely monitored and regulated by the cell for survival (Csonka, 1989; Firth et al., 2016). In addition to the economical utility of the Entner-Doudoroff and glyoxylate shunt pathways available to the genus, Clade III genomes also encode a majority of the components necessary for the C-P lyase system which allows for the foraging of phosphorous from organophosphates when phosphorous may not be readily available in the biologically preferred inorganic form (Kamat and Raushel, 2013; Stosiek et al., 2020). The species in Clade III were each isolated from subzero, hypersaline environments where redox availability and membrane regulation are critical to survival and competition (Zhang et al., 2008; Chua et al., 2018; Cooper et al., 2019, Chapter 2). The metabolic characteristics encoded in their genomes thus reflect the flexible and adaptive strategies employed for their success in subzero brines.

An analysis of the COG annotations revealed several functions that appear to be enriched in Clade III compared to the other clades (Supplementary Table S4.7), including membrane transport systems, compatible solute metabolism, and genetic regulatory mechanisms. ABC-type membrane transport systems, the most enriched functions in Clade III, likely enable the essential transport of organic substrates, ions, and compatible solutes (Schneider and Hunke, 1998). The presence of bacteriorhodopsin in Clade III, along with the KEGG pathway for retinal biosynthesis (Figure 4.5), indicate that transmembrane pigments may also play a role in the fitness of these species. This functional capability is only observed in the isolates from cryopeg

brines and in their closest relative, *M. psychrophilus*, isolated from sea ice. Bacteriorhodopsin is known as a transmembrane proton-pumping protein that uses light energy to catalyze the process (Lanyi, 2004). The presence of this gene in the cryopeg genomes is enigmatic, as these species have been isolated from surface radiation for thousands of years (Iwahana et al., 2021). We hypothesize that this protein, if expressed, may play a role in ion exchange across the membrane in the hypersaline setting of the cryopeg. As the cryopeg ecosystem is under continuous osmotic pressure, more than one means to regulate ion exchange may increase fitness. The production of compatible solutes, such as glycine, is an essential means for maintaining cellular integrity. In cryopeg genomes, the enrichment of enzymes involved in the metabolism of sarcosine to glycine appears to reflect the importance of compatible solute production in a subzero brine habitat, as this metabolism has also been documented in sea ice bacteria (Collins and Deming, 2013; Firth et al., 2016).

In addition to functions that would modulate transmembrane transport and osmolarity under extreme conditions, we found an enrichment of genes in Clade III related to genetic regulation and modification. These genes may have particular evolutionary importance for bacteria in cryopeg brines where continuously low temperatures mean long generation times (Mykytczuk et al., 2013; Chua et al., 2018). To combat environmental stressors while managing slow rates of vertical evolution, we hypothesized that *Marinobacter* may need to employ more regulatory and recombinant strategies than may be typical to maintain a competitive state. Stress response proteins, which are transcriptional regulators that can be related to salt and cold stress (Starosta et al., 2014), were enriched in Clade III; they may play a role in managing gene expression under the extreme conditions of a subzero brine. A recombination-promoting DNA endonuclease and a

plasmid stabilization module were also found to be enriched in Clade III. These mobilome genes may play a role in maintaining genetic diversity and fluidity in the population (Toussaint and Chandler, 2012). Though enrichment in a clade does not confirm that traits are present specifically for their adaptive capacity to a given environment, we consider the potential of enriched mobilome genes in Clade III important to developing a more nuanced understanding of adaptive success in the subzero brine environment.

HGT signatures across the Marinobacter *pangenome*

Due to the extreme temperature and salinity conditions of subzero brines, bacterial growth rates in these environments are expected to be slow, yet the concentrating effect of freezing saline water means that infection pressure from viruses within the resulting brines may be much higher than in their source waters (Wells and Deming, 2006; Collins and Deming, 2011). In subzero cryopeg brines with high bacterial densities, cell-cell interactions are also expected to be high (Rapp et al., 2021). Given the anticipated low rates of vertical inheritance associated with slow growth rates, we hypothesized that HGT plays an essential role in the evolution and adaptation of bacteria in environments with multiple extreme stressors. We measured putative HGT events in each *Marinobacter* genome using an alignment-based approach that identifies genes that are more closely related to distant taxonomic relatives than to close relatives (Zhu et al., 2014). This method reveals genes that may have been transferred to a genome from another genome outside of its own taxonomic family (for the analytical settings used in this study). We observed an overall higher abundance of HGT genes in Clade III genomes relative to other clades and found that these genes belong to categories that may be involved in conferring adaptation to subzero brine conditions (Figure 4.6). Among the most abundant HGT genes were those in the COG

categories for transcription (K) and the mobilome: prophages, transposons (X). These transferred genes may play roles in genome flexibility under extreme conditions or simply be signatures of the mechanisms of HGT. Category K includes stress response transcription factors, such as cold and salt stress response proteins (Starosta et al., 2014), which would contribute to the adaptive preparedness displayed by *Marinobacter* spp. in Clade III. Category X includes a variety of genes, from transposases, integrases, and recombinases to phage-associated genes, each of which may play a role in genome modification or HGT (Toussaint and Chandler, 2012). We cannot disentangle the possibility of virally transduced or conjugation-based HGT from these data, but both are strong possibilities in the densely populated settings of subzero brines (Collins and Deming 2011; 2013' Rapp et al., 2021). The cryopeg brines that yielded the new *Marinobacter* isolates and MAGs contained high densities of both bacteria and virus like particles (approximately 10^8 mL^{-1} each; Cooper et al., 2019, Chapter 2), and our previous work on the same subzero brine system showed that virally-mediated gene transfer can directly affect microbial fitness (Zhong et al., 2020).

We also identified notable abundances of HGT events from COG categories for carbohydrate transport and metabolism (G), amino acid transport and metabolism (E), and energy production and conversion (C) in Clade III genomes. As these processes reflect the metabolic flexibility displayed by the genus *Marinobacter* as a whole (Handley and Lloyd, 2013), the abundance of HGT-derived genes in these categories and potentially those of unknown function highlights the importance of flexible metabolic potential in an extreme environment (Collins and Deming, 2013; Fuchsman et al., 2017). The cryopeg ecosystem is isolated from the surface environment, leaving bacteria to survive on the nutrients available *in situ*. Over a long timeframe, many of

these nutrients may derive from cryptic growth, where cytoplasmic contents released upon cell lysis (due to viral lysis or other forms of cell death) are recycled in support of a persistent, slow-growing community (as presented by Rapp et al., 2021). As labile organics are consumed, more recalcitrant (sediment-adsorbed or permafrost-derived) forms of organic substrates may become important sources of nutrients for the community. Extracellular enzymes that degrade organic matter have been demonstrated to be active under cryopeg conditions (Showalter and Deming, 2021), and the *Marinobacter* species characterized here is well equipped genetically to take advantage of the variety of substrates that can be produced by this general process. We suggest that *Marinobacter* spp. in subzero brines have succeeded at competitive opportunism by maintaining and expanding broad genomic potential for the consumption of a variety of metabolites using multiple redox sources (electron acceptors). The combination of genetic and metabolic flexibility potentially achieved through HGT highlights the key strategies employed by this genus that has allowed it to occupy its competitive position globally in saline environments and endemically in subzero brines.

4.6 CONCLUSION

Pangenomics provides a unique approach to explore the evolutionary history and ecological context of microorganisms. The pangenome of *Marinobacter,* a globally distributed genus of bacteria inhabiting saline environments, displays genetic versatility that complements previous physiological studies. There is a tremendous amount of rare (variable) gene content in this pangenome, and a relatively small, conserved (core) region. This pattern of gene distribution reflects the metabolically flexible lifestyle across a wide range of saline environments for which the genus is known. Here, we have produced a set of genomes belonging to a novel species of

Marinobacter, successfully cultivated from cryopeg brines and recovered as high-quality MAGs, that represents the dominant taxon in the sampled communities. This species is most closely related to other *Marinobacter* species found in sea ice and in subglacial brines from both Arctic and Antarctic settings; together they represent a unique clade of *Marinobacter*. Members of this subzero brine clade share genetic features thought to be important in adapting to extremely cold and hypersaline environments. Patterns of phylogeny and average nucleotide identity indicate that this clade may have diverged early from the genus *Marinobacter* and may even represent a novel genus of bacteria endemic to subzero brine environments, which will be explored in future analyses. The genetic inventory of this clade alongside ecological context gives a clear depiction of the competitive and selective forces that drive microbial evolution in these extreme environments, including metabolic versatility, broad membrane transport and environmental sensory functions, and genomic flexibility via the mobilome and horizontal gene transfer.

4.7 ACKNOWLEDGEMENTS

We acknowledge that the isolates of the novel *Marinobacter* species were obtained from cryopeg brines located in the land of the North Alaska Coast Iñupiat. The Ukpeagvik Iñupiat Corporation science team provided essential logistical support and helped with access to the permafrost tunnel, below which the subzero brines were discovered. We also thank members of the larger project team for support in the field and field-based laboratory: Shelly Carpenter, Max Showalter, Anders Torstensson, Hannah Dawson and Jodi Young. We also thank Olivier Zablocki, Natalie Solonenko, and Marie Burris for help in sequencing and assembling the metagenomes.

References

Ahn, S., Jung, J., Jang, I. A., Madsen, E. L., and Park, W. (2016). Role of glyoxylate shunt in oxidative stress response. *J. Biol. Chem.* 291, 11928–11938. doi:10.1074/jbc.M115.708149.

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.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. J. Mol. Biol. 215, 403–410. doi:10.1016/S0022-2836(05)80360-2.

Antunes, A., Francça, L., Rainey, F. A., Huber, R., Fernanda Nobre, M., Edwards, K. J., et al. (2007). *Marinobacter salsuginis* sp. nov., isolated from the brine-seawater interface of the Shaban Deep, Red Sea. *Int. J. Syst. Evol. Microbiol.* 57, 1035–1040. doi:10.1099/ijs.0.64862-0.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., et al. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* 36, 2251–2252. doi:10.1093/bioinformatics/btz859.

Bagheri, M., Amoozegar, M. A., Didari, M., Makhdoumi-Kakhki, A., Schumann, P., Spröer, C., et al. (2013). *Marinobacter persicus* sp. nov., a moderately halophilic bacterium from a saline lake in Iran. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 104, 47–54. doi:10.1007/s10482-013-9923-3.

Barco, R. A., Garrity, G. M., Scott, J. J., Amend, J. P., Nealson, K. H., and Emerson, D. (2020). A genus definition for bacteria and archaea based on a standard genome relatedness index. *MBio* 11. doi:10.1128/MBIO.02475-19.

Biller, S. J., Berube, P. M., Lindell, D., and Chisholm, S. W. (2014). *Prochlorococcus*: The structure and function of collective diversity. *Nat. Rev. Microbiol.* 13, 13–27. doi:10.1038/nrmicro3378.

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., and Rapp, J. Z. (2015). Microbial ecology of the cryosphere: Sea ice and glacial habitats. *Nat. Rev. Microbiol.* 13, 677–690. doi:10.1038/nrmicro3522.

Bonis, B. M., and Gralnick, J. A. (2015). *Marinobacter subterrani*, a genetically tractable neutrophilic Fe(II)-oxidizing strain isolated from the Soudan Iron Mine. *Front. Microbiol.* 6, 719. doi:10.3389/fmicb.2015.00719.

Brinkmeyer, R., Knittel, K., Jürgens, J., Weyland, H., Amann, R., and Helmke, E. (2003). Diversity and Structure of Bacterial Communities in Arctic versus Antarctic Pack Ice. *Appl. Environ. Microbiol.* 69, 6610–6619. doi:10.1128/AEM.69.11.6610-6619.2003.

Brockhurst, M. A., Harrison, E., Hall, J. P. J., Richards, T., McNally, A., and MacLean, C. (2019). The ecology and evolution of pangenomes. *Curr. Biol.* 29, R1094–R1103. doi:10.1016/j.cub.2019.08.012.

Buchfink, B., Xie, C., and Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi:10.1038/nmeth.3176.

Cao, J., Liu, P., Liu, R., Su, H., Wei, Y., Liu, R., et al. (2019). *Marinobacter profundi* sp. nov., a slightly halophilic bacterium isolated from a deep-sea sediment sample of the New Britain Trench. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 112, 425–434. doi:10.1007/s10482-018-1176-8.

Chua, M. J., Campen, R. L., Wahl, L., Grzymski, J. J., and Mikucki, J. A. (2018). Genomic and physiological characterization and description of *Marinobacter gelidimuriae* sp. nov., a psychrophilic, moderate halophile from Blood Falls, an antarctic subglacial brine. *FEMS Microbiol. Ecol.* 94, 1–15. doi:10.1093/femsec/fiy021.

Collins, R. E., and Deming, J. W. (2011). Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up. *Polar Biol.* 34, 1831–1841. doi:10.1007/s00300-011-1008-z.

Collins, R. E., and Deming, J. W. (2013). An inter-order horizontal gene transfer event enables the catabolism of compatible solutes by *Colwellia psychrerythraea* 34H. *Extremophiles* 17, 601–610. doi:10.1007/s00792-013-0543-7.

Conway, T. (1992). The Entner-Doudoroff pathway: history, physiology and molecular biology. *FEMS Microbiol. Rev.* 103, 1–28. doi:10.1111/j.1574-6968.1992.tb05822.x.

Cooper, Z. S., Rapp, J. Z., Carpenter, S. D., Iwahana, G., Eicken, H., and Deming, J. W. (2019). Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiol. Ecol.* 95, 1–15. doi:10.1093/femsec/fiz166. Z. Cooper Dissertation, Chapter 2.

Cox, G. F. N., and Weeks, W. F. (1983). Equations for determining the gas and brine volumes in sea ice samples. *J. Glaciol.* 29, 306–316. doi:10.3189/s0022143000008364.

Csonka, L. N. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiol. Rev.* 53, 121–147. doi:10.1128/mmbr.53.1.121-147.1989.

Cui, Z., Gao, W., Li, Q., Xu, G., and Zheng, L. (2013). Genome sequence of the polycyclic aromatic hydrocarbon-degrading bacterium *Marinobacter nanhaiticus* strain D15-8WT. *Genome Announc.* 1. doi:10.1128/genomeA.00301-13.

Cui, Z., Gao, W., Xu, G., Luan, X., Li, Q., Yin, X., et al. (2016). *Marinobacter aromaticivorans* sp. nov., a polycyclic aromatic hydrocarbon-degrading bacterium isolated from sea sediment. *Int. J. Syst. Evol. Microbiol.* 66, 353–359. doi:10.1099/ijsem.0.000722.

Czajka, J. J., Abernathy, M. H., Benites, V. T., Baidoo, E. E. K., Deming, J. W., and Tang, Y. J. (2018). Model metabolic strategy for heterotrophic bacteria in the cold ocean based on *Colwellia psychrerythraea* 34H. *Proc. Natl. Acad. Sci. U. S. A.* 115, 12507–12512. doi:10.1073/pnas.1807804115.

Delmont, T. O., and Eren, E. M. (2018). Linking pangenomes and metagenomes: The Prochlorococcus metapangenome. PeerJ 2018, 1–23. doi:10.7717/peerj.4320.

van Dongen, S., and Abreu-Goodger, C. (2012). Using MCL to extract clusters from networks. Methods Mol. Biol. 804, 281–295. doi:10.1007/978-1-61779-361-5_15.

Durot, M., Bourguignon, P. Y., and Schachter, V. (2009). Genome-scale models of bacterial metabolism: Reconstruction and applications. *FEMS Microbiol. Rev.* 33, 164–190. doi:10.1111/j.1574-6976.2008.00146.x.

Dyksterhouse, S. E., Gray, J. P., Herwig, R. P., Lara, J. C., and Staley, J. T. (1995). *Cycloclasticus pugetii* gen. nov., sp. nov., an Aromatic hydrocarbon- degrading bacterium from marine sediments. *Int. J. Syst. Bacteriol.* 45, 116–123. doi:10.1099/00207713-45-1-116.

Eddy, S. R. (2011). Accelerated profile HMM searches. PLoS Comput. Biol. 7, 1002195. doi:10.1371/journal.pcbi.1002195.

Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi'o: An advanced analysis and visualization platformfor 'omics data. *PeerJ* 2015, 1–29. doi:10.7717/peerj.1319.

Fenchel, T., and Finlay, B. J. (2004). The ubiquity of small species: patterns of local and global diversity. *BioScience* 54, 777-784. doi:10.1641/0006-3568(2004)054[0777:TUOSSP]2.0.CO;2.

Firth, E., Carpenter, S. D., Sørensen, H. L., Collins, R. E., and Deming, J. W. (2016). Bacterial use of choline to tolerate salinity shifts in sea-ice brines. *Elementa* 2016. doi:10.12952/journal.elementa.000120.

Fuchsman, C. A., Collins, R. E., Rocap, G., and Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ* 5, e3865. doi:10.7717/PEERJ.3865.

Gao, W., Cui, Z., Li, Q., Xu, G., Jia, X., and Zheng, L. (2013). *Marinobacter nanhaiticus* sp. nov., polycyclic aromatic hydrocarbon- degrading bacterium isolated from the sediment of the South China Sea. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 103, 485–491. doi:10.1007/s10482-012-9830-z.

Galperin, M. Y., Kristensen, D. M., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2019). Microbial genome analysis: The COG approach. *Brief. Bioinform.* 20, 1063–1070. doi:10.1093/bib/bbx117.

Gärdes, A., Kaeppel, E., Shehzad, A., Seebah, S., Teeling, H., Yarza, P., et al. (2010). Complete genome sequence of *Marinobacter adhaerens* type strain (HP15), a diatom-interacting marine microorganism. *Stand. Genomic Sci.* 3, 97–107. doi:10.4056/sigs.922139.

Gauthier, M. J., Lafay, B., Christen, R., Fernandez, L., Acquaviva, M., Bonin, P., et al. (1992). *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *Int. J. Syst. Bacteriol.* 42, 568–576. doi:10.1099/00207713-42-4-568.

Gilichinsky, D., Rivkina, E., Bakermans, C., Shcherbakova, V., Petrovskaya, L., Ozerskaya, S., et al. (2005). Biodiversity of cryopegs in permafrost. *FEMS Microbiol. Ecol.*, 117–128. doi:10.1016/j.femsec.2005.02.003.

Gilichinsky, D., Rivkina, E., Shcherbakova, V., Laurinavichuis, K., and Tiedje, J. (2003). Supercooled water brines within permafrost - An unknown ecological niche for microorganisms: A model for astrobiology. *Astrobiology* 3, 331–341. doi:10.1089/153110703769016424.

Gorshkova, N. M., Ivanova, E. P., Sergeev, A. F., Zhukova, N. V., Alexeeva, Y., Wright, J. P., et al. (2003). *Marinobacter excellens* sp. nov., isolated from sediments of the Sea of Japan. *Int. J. Syst. Evol. Microbiol.* 53, 2073–2078. doi:10.1099/ijs.0.02693-0.

Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2018). Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* 12, 1861–1866. doi:10.1038/s41396-018-0091-3.

Graumann, P., Wendrich, T. M., Weber, M. H. W., Schröder, K., and Marahiel, M. A. (1997). A family of cold shock proteins in *Bacillus subtilis* is essential for cellular growth and for efficient protein synthesis at optimal and low temperatures. *Mol. Microbiol.* 25, 741–756. doi:10.1046/j.1365-2958.1997.5121878.x.

Green, D. H., Bowman, J. P., Smith, E. A., Gutierrez, T., and Bolch, C. J. S. (2006). *Marinobacter algicola* sp. nov., isolated from laboratory cultures of paralytic shellfish toxin-producing dinoflagellates. *Int. J. Syst. Evol. Microbiol*. 56, 523–527. doi:10.1099/ijs.0.63447-0.

Grimaud, R., Ghiglione, J. F., Cagnon, C., Lauga, B., Vaysse, P. J., Rodriguez-Blanco, A., et al. (2012). Genome sequence of the marine bacterium *Marinobacter hydrocarbonoclasticus* SP17, which forms biofilms on hydrophobic organic compounds. *J. Bacteriol.* 194, 3539–3540. doi:10.1128/JB.00500-12.

Gu, J., Cai, H., Yu, S. L., Qu, R., Yin, B., Guo, Y. F., et al. (2007). *Marinobacter gudaonensis* sp. nov., isolated from an oil-polluted saline soil in a Chinese oilfield. *Int. J. Syst. Evol. Microbiol.* 57, 250–254. doi:10.1099/ijs.0.64522-0.

Guo, B., Gu, J., Ye, Y. G., Tang, Y. Q., Kida, K., and Wu, X. L. (2007). *Marinobacter segnicrescens* sp. nov., a moderate halophile isolated from benthic sediment of the South China Sea. *Int. J. Syst. Evol. Microbiol.* 57, 1970–1974. doi:10.1099/ijs.0.65030-0.

Han, J. R., Ling, S. K., Yu, W. N., Chen, G. J., and Du, Z. J. (2017). *Marinobacter salexigens* sp. nov., isolated from marine sediment. *Int. J. Syst. Evol. Microbiol.* 67, 4595–4600. doi:10.1099/ijsem.0.002337.

Handley, K. M., Héry, M., and Lloyd, J. R. (2009). *Marinobacter santoriniensis* sp. nov., an arsenaterespiring and arsenite-oxidizing bacterium isolated from hydrothermal sediment. *Int. J. Syst. Evol. Microbiol.* 59, 886–892. doi:10.1099/ijs.0.003145-0.

Handley, K. M., and Lloyd, J. R. (2013). Biogeochemical implications of the ubiquitous colonization of marine habitats and redox gradients by *Marinobacter* species. *Front. Microbiol.* 4, 1–10. doi:10.3389/fmicb.2013.00136.

Harris, L. K., and Theriot, J. A. (2018). Surface area to volume ratio: A natural variable for bacterial morphogenesis. *Trends Microbiol.* 26, 815–832. doi:10.1016/j.tim.2018.04.008.

Harrison, J. P., Gheeraert, N., Tsigelnitskiy, D., and Cockell, C. S. (2013). The limits for life under multiple extremes. *Trends Microbiol.* 21, 204–212. doi:10.1016/j.tim.2013.01.006.

Hedi, A., Cayol, J. L., Sadfi, N., and Fardeau, M. L. (2015). *Marinobacter piscensis* sp. nov., a moderately halophilic bacterium isolated from salty food in Tunisia. *Curr. Microbiol.* 70, 544–549. doi:10.1007/s00284-014-0754-x.

Huo, Y. Y., Wang, C. S., Yang, J. Y., Wu, M., and Xu, X. W. (2008). *Marinobacter mobilis* sp. nov. and *Marinobacter zhejiangensis* sp. nov., halophilic bacteria isolated from the East China Sea. *Int. J. Syst. Evol. Microbiol.* 58, 2885–2889. doi:10.1099/ijs.0.2008/000786-0.

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119. doi:10.1186/1471-2105-11-119.

Ivanova, E. P., Ng, H. J., Webb, H. K., Feng, G., Oshima, K., Hattori, M., et al. (2014). Draft genome sequences of *Marinobacter similis* A3d10T and *Marinobacter salarius* R9SW1T. *Genome Announc.* 2, 442–456. doi:10.1128/genomeA.00442-14.

Iwahana, G., Cooper, Z. S., Carpenter, S. D., Deming, J. W., and Eicken, H. (2021). Intra-ice and intra-sediment cryopeg brine occurrence in permafrost near Utqiaġvik (Barrow). *Permafr. Periglac. Process.* 1–20. doi:10.1002/ppp.2101.

Kaeppel, E. C., Gärdes, A., Seebah, S., Grossart, H. P., and Ullrich, M. S. (2011). *Marinobacter adhaerens* sp. nov., isolated from marine aggregates formed with the diatom *Thalassiosira weissflogii*. *Int. J. Syst. Evol. Microbiol.* 62, 124–128. doi:10.1099/ijs.0.030189-0.

Kamat, S. S., and Raushel, F. M. (2013). The enzymatic conversion of phosphonates to phosphate by bacteria. *Curr. Opin. Chem. Biol.* 17, 589–596. doi:10.1016/j.cbpa.2013.06.006.

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 2019, e7359. doi:10.7717/peerj.7359.

Kaye, J. Z., Sylvan, J. B., Edwards, K. J., and Baross, J. A. (2011). *Halomonas* and *Marinobacter* ecotypes from hydrothermal vent, subseafloor and deep-sea environments. *FEMS Microbiol. Ecol.* 75, 123–133. doi:10.1111/j.1574-6941.2010.00984.x.

Kim, M., Oh, H. S., Park, S. C., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 346–351. doi:10.1099/ijs.0.059774-0.

Kim, J. O., Lee, H. J., Han, S. I., and Whang, K. S. (2017). *Marinobacter halotolerans* sp. Nov., a halophilic bacterium isolated from a saltern crystallizing pond. *Int. J. Syst. Evol. Microbiol.* 67, 460–465. doi:10.1099/ijsem.0.001653.

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi:10.1038/s41587-019-0072-8.

Kuhn, E., Ichimura, A. S., Peng, V., Fritsen, C. H., Trubl, G., Doran, P. T., et al. (2014). Brine assemblages of ultrasmall microbial cells within the ice cover of Lake Vida, Antarctica. *Appl. Environ. Microbiol.* 80, 3687–3698. doi:10.1128/AEM.00276-14.

Lam, B. R., Barr, C. R., Rowe, A. R., and Nealson, K. H. (2019). Differences in applied redox potential on cathodes enrich for diverse electrochemically active microbial isolates from a marine sediment. *Front. Microbiol.* 10, 1979. doi:10.3389/fmicb.2019.01979.

Lanyi, J. K. (2004). Bacteriorhodopsin. *Annu. Rev. Physiol.* 66, 665–688. doi:10.1146/annurev.physiol.66.032102.150049.

Lee, M. D. (2019). GToTree: a user-friendly workflow for phylogenomics. *Bioinformatics* 35, 4162–4164. doi:10.1093/bioinformatics/btz188.

Li, G., Wang, S., Gai, Y., Liu, X., Lai, Q., and Shao, Z. (2020). *Marinobacter changyiensis*, sp. Nov., isolated from offshore sediment. *Int. J. Syst. Evol. Microbiol.* 70, 3004–3011. doi:10.1099/ijsem.0.004118.

Liebgott, P. P., Casalot, L., Paillard, S., Lorquin, J., and Labat, M. (2006). *Marinobacter vinifirmus* sp. nov., a moderately halophilic bacterium isolated from a wine-barrel-decalcification wastewater. *Int. J. Syst. Evol. Microbiol.* 56, 2511–2516. doi:10.1099/ijs.0.64368-0.

Liu, C., Chen, C. X., Zhang, X. Y., Yu, Y., Liu, A., Li, G. W., et al. (2012). *Marinobacter antarcticus* sp. nov., a halotolerant bacterium isolated from Antarctic intertidal sandy sediment. *Int. J. Syst. Evol. Microbiol.* 62, 1838–1844. doi:10.1099/ijs.0.035774-0.

Liu, Q., Xamxidin, M., Sun, C., Cheng, H., Meng, F. X., Wu, Y. H., et al. (2018). *Marinobacter fuscus* sp. nov., a marine bacterium of gammaproteobacteria isolated from surface seawater. *Int. J. Syst. Evol. Microbiol.* 68, 3156–3162. doi:10.1099/ijsem.0.002956.

Luo, Y. J., Xie, B. S., Lv, X. L., Cai, M., Wang, Y. N., Cui, H. L., et al. (2015). *Marinobacter shengliensis* sp. nov., a moderately halophilic bacterium isolated from oil-contaminated saline soil. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 107, 1085–1094. doi:10.1007/s10482-015-0401-y.

Mckinney, W. (2010). Data Structures for Statistical Computing in Python. *Proc. of the 9th Python in Science Conf.* 445, 51-56.

Methé, B. A., Nelson, K. E., Deming, J. W., Momen, B., Melamud, E., Zhang, X., et al. (2005). The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proc. Natl. Acad. Sci. U. S. A.* 102, 10913–10918. doi:10.1073/pnas.0504766102.

Meyer, H., Schirrmeister, L., Andreev, A., Wagner, D., Hubberten, H. W., Yoshikawa, K., et al. (2010a). Lateglacial and Holocene isotopic and environmental history of northern coastal Alaska - Results from a buried ice-wedge system at Barrow. *Quat. Sci. Rev.* 29, 3720–3735. doi:10.1016/j.quascirev.2010.08.005.

Meyer, H., Schirrmeister, L., Yoshikawa, K., Opel, T., Wetterich, S., Hubberten, H. W., et al. (2010b). Permafrost evidence for severe winter cooling during the Younger Dryas in northern Alaska. *Geophys. Res. Lett.* 37, 1–5. doi:10.1029/2009GL041013.

Montes, M. J., Bozal, N., and Mercadé, E. (2008). *Marinobacter guineae* sp. nov., a novel moderately halophilic bacterium from an Antarctic environment. *Int. J. Syst. Evol. Microbiol.* 58, 1346–1349. doi:10.1099/ijs.0.65298-0.

Moulana, A., Anderson, R. E., Fortunato, C. S., and Huber, J. A. (2020). Selection is a significant driver of gene gain and loss in the pangenome of the bacterial genus *Sulfurovum* in geographically distinct deep-sea hydrothermal vents. *mSystems* 5. doi:10.1128/mSystems.00673-19.

Musa, H., Kasim, F. H., Gunny, A. A. N., Gopinath, S. C. B., Chinni, S. V., and Ahmad, M. A. (2019). Whole genome sequence of moderate halophilic marine bacterium *Marinobacter litoralis* SW-45: Abundance of non-coding RNAs. *Int. J. Biol. Macromol.* 133, 1288–1298. doi:10.1016/j.ijbiomac.2019.05.003.

Mykytczuk, N. C. S., Foote, S. J., Omelon, C. R., Southam, G., Greer, C. W., and Whyte, L. G. (2013). Bacterial growth at –15°C; molecular insights from the permafrost bacterium *Planococcus halocryophilus* Or1. *ISME J.* 7, 1211–1226. doi:10.1038/ismej.2013.8.

Niederberger, T. D., Perreault, N. N., Tille, S., Lollar, B. S., Lacrampe-Couloume, G., Andersen, D., et al. (2010). Microbial characterization of a subzero, hypersaline methane seep in the Canadian High Arctic. *ISME J.* 4, 1326–1339. doi:10.1038/ismej.2010.57.

Ng, H. J., López-Pérez, M., Webb, H. K., Gomez, D., Sawabe, T., Ryan, J., et al. (2014). *Marinobacter salarius* sp. nov. and *Marinobacter similis* sp. nov., isolated from sea water. *PLoS One* 9, e106514. doi:10.1371/journal.pone.0106514.

Papke, R. T., de la Haba, R. R., Infante-Domínguez, C., Pérez, D., Sánchez-Porro, C., Lapierre, P., et al. (2013). Draft genome sequence of the moderately halophilic bacterium *Marinobacter lipolyticus* strain SM19. *Genome Announc.* 1. doi:10.1128/genomeA.00379-13.

Park, S., Kim, S., Kang, C. H., Jung, Y. T., and Yoon, J. H. (2015). *Marinobacter confluentis* sp. Nov, a lipolytic bacterium isolated from a junction between the ocean and a freshwater lake. *Int. J. Syst. Evol. Microbiol.* 65, 4873–4879. doi:10.1099/ijsem.0.000659.

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

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996. doi:10.1038/nbt.4229.

Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., et al. (2011). Scikit-learn: Machine Learning in Python. *J. Mach. Learn Res.* 12,2825-2830.

Price, P. B., and Sowers, T. (2004). Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proc. Natl. Acad. Sci. U. S. A.* 101, 4631–4636. doi:10.1073/pnas.0400522101.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490. doi:10.1371/journal.pone.0009490.

Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., and Toth, I. K. (2016). Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8, 12–24. doi:10.1039/c5ay02550h.

Qu, L., Zhu, F., Zhang, I., Gao, C., and Sun, X. (2011). *Marinobacter daqiaonensis* sp. nov., a moderate halophile isolated from a yellow sea salt pond. *Int. J. Syst. Evol. Microbiol.* 61, 3003–3008. doi:10.1099/ijs.0.028993-0.

Rambaut, A. (2009). FigTree, a graphical viewer of phylogenetic trees. Edinburgh: Institute of Evolutionary Biology University of Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/

Rani, S., Koh, H. W., Kim, H., Rhee, S. K., and Park, S. J. (2017). *Marinobacter salinus* sp. Nov., a moderately halophilic bacterium isolated from a tidal flat environment. *Int. J. Syst. Evol. Microbiol.* 67, 205–211. doi:10.1099/ijsem.0.001587.

Rani, S., Jeon, W. J., Koh, H. W., Kim, Y. E., Kang, M. S., and Park, S. J. (2017). Genomic potential of *Marinobacter salinus* Hb8T as sulfur oxidizing and aromatic hydrocarbon degrading bacterium. *Mar. Genomics* 34, 19–21. doi:10.1016/j.margen.2017.02.005.

Rapp, J. Z., Sullivan, M. B., and Deming, J. W. (2021). Divergent genomic adaptations in the microbiomes of Arctic subzero sea-ice and cryopeg brines. *Front. Microbiol.* 12, 1–21. doi:10.3389/fmicb.2021.701186.

Reback, J., McKinney, W., jbrockmendel, Bossche, J. Van den, Augspurger, T., Cloud, P., et al. (2020). pandas-dev/pandas: Pandas 1.1.2. doi:10.5281/ZENODO.4019559.

Schloter, M., Lebuhn, M., Heulin, T., and Hartmann, A. (2000). Ecology and evolution of bacterial microdiversity. *FEMS Microbiol. Rev.* 24, 647–660. doi:10.1111/j.1574-6976.2000.tb00564.x.

Schneider, E., and Hunke, S. (1998). ATP-binding-cassette (ABC) transport systems: Functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS Microbiol. Rev.* 22, 1–20. doi:10.1111/j.1574-6976.1998.tb00358.x.

Shieh, W. Y., Jean, W. D., Lin, Y. Te, and Tseng, M. (2003). *Marinobacter lutaoensis* sp. nov., a thermotolerant marine bacterium isolated from a coastal hot spring in Lutao, Taiwan. *Can. J. Microbiol.* 49, 244–252. doi:10.1139/w03-032.

Shivaji, S., Gupta, P., Chaturvedi, P., Suresh, K., and Delille, D. (2005). *Marinobacter maritimus* sp. nov., a psychrotolerant strain isolated from sea water off the subantarctic Kerguelen islands. *Int. J. Syst. Evol. Microbiol.* 55, 1453–1456. doi:10.1099/ijs.0.63478-0.

Showalter, G. M., and Deming, J. W. (2018). Low-temperature chemotaxis, halotaxis and chemohalotaxis by the psychrophilic marine bacterium *Colwellia psychrerythraea* 34H. *Environ. Microbiol. Rep.* 10, 92–101. doi:10.1111/1758-2229.12610.

Showalter, G. M., and Deming, J. W. (2021). Measured and modeled activity of extracellular enzyme leucine aminopeptidase among cultured and *in situ* sea ice bacteria. *Aquat. Microb. Ecol.* (in press).

Sieber, C. M. K., Probst, A. J., Sharrar, A., Thomas, B. C., Hess, M., Tringe, S. G., et al. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat. Microbiol.* 3, 836–843. doi:10.1038/s41564-018-0171-1.

Starosta, A. L., Lassak, J., Jung, K., and Wilson, D. N. (2014). The bacterial translation stress response. *FEMS Microbiol. Rev.* 38, 1172–1201. doi:10.1111/1574-6976.12083.

Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci.* 100, 9440–9445. doi:10.1073/PNAS.1530509100.

Stosiek, N., Talma, M., and Klimek-Ochab, M. (2020). Carbon-phosphorus lyase—the state of the art. *Appl. Biochem. Biotechnol.* 190, 1525–1552. doi:10.1007/s12010-019-03161-4.

Sun, Q.-L., Sun, Y.-L., Sun, Y.-Y., Luan, Z.-D., and Lian, C. (2020). *Marinobacter fonticola* sp. nov., isolated from deep sea cold seep sediment. *Int. J. Syst. Evol. Microbiol* 70, 1172–1177. doi:10.1099/ijsem.0.003895.

Takai, K., Moyer, C. L., Miyazaki, M., Nogi, Y., Hirayama, H., Nealson, K. H., et al. (2005). *Marinobacter alkaliphilus* sp. nov., a novel alkaliphilic bacterium isolated from subseafloor alkaline serpentine mud from Ocean Drilling Program Site 1200 at South Chamorro Seamount, Mariana Forearc. *Extremophiles* 9, 17–27. doi:10.1007/s00792-004-0416-1.

Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997). A genomic perspective on protein families. *Science* (80-.). 278, 631–637. doi:10.1126/science.278.5338.631.

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome." *Proc. Natl. Acad. Sci.* 102, 13950–13955. doi:10.1073/PNAS.0506758102.

Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: the bacterial pan-genome. *Curr. Opin. Microbiol.* 11, 472–477. doi:10.1016/J.MIB.2008.09.006.

Toussaint, A., and Chandler, M. (2012). Prokaryote genome fluidity: Toward a system approach of the mobilome. *Methods Mol. Biol.* 804, 57–80. doi:10.1007/978-1-61779-361-5_4.

Vaidya, B., Kumar, R., Korpole, S., Pinnaka, A. K., and Tanuku, N. R. S. (2015). *Marinobacter nitratireducens* sp. nov., a halophilic and lipolytic bacterium isolated from coastal surface sea water. *Int. J. Syst. Evol. Microbiol.* 65, 2056–2063. doi:10.1099/ijs.0.000218.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9, e112963. doi:10.1371/journal.pone.0112963.

Wang, H., Li, H., Shao, Z., Liao, S., Johnstone, L., Rensing, C., et al. (2012). Genome sequence of deep-sea manganese-oxidizing bacterium *Marinobacter manganoxydans* MnI7-9. *J. Bacteriol.* 194, 899–900. doi:10.1128/JB.06551-11.

Waskom, M. (2021). seaborn: statistical data visualization. *J. Open Source Softw.* 6, 3021. doi:10.21105/joss.03021.

Wells, L. E., and Deming, J. W. (2006). Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environ. Microbiol.* 8, 1115–1121. doi:10.1111/j.1462-2920.2006.00984.x.

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017a). Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb. Genomics* 3. doi:10.1099/mgen.0.000132.

Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017b). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput. Biol.* 13, e1005595. doi:10.1371/journal.pcbi.1005595.

Wu, Y. W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics.* doi:10.1093/bioinformatics/btv638.

Xu, X. W., Wu, Y. H., Wang, C. S., Yang, J. Y., Oren, A., and Wu, M. (2008). *Marinobacter pelagius* sp. nov., a moderately halophilic bacterium. *Int. J. Syst. Evol. Microbiol.* 58, 637–640. doi:10.1099/ijs.0.65390-0.

Xu, S., Wang, D., Wei, Y., Cui, Q., and Li, W. (2018). *Marinobacter bohaiensis* sp. nov., a moderate halophile isolated from benthic sediment of the Bohai Sea. *Int. J. Syst. Evol. Microbiol.* 68, 3534–3539. doi:10.1099/ijsem.0.003025.

Yoon, J. H., Shin, D. Y., Kim, I. G., Kang, K. H., and Park, Y. H. (2003). *Marinobacter litoralis* sp. nov., a moderately halophilic bacterium isolated from sea water from the East Sea in Korea. *Int. J. Syst. Evol. Microbiol.* 53, 563–568. doi:10.1099/ijs.0.02363-0.

Yoon, J. H., Yeo, S. H., kim, I. G., and Oh, T. K. (2004). *Marinobacter flavimaris* sp. nov. and Marinobacter daepoensis sp. nov., slightly halophilic organisms isolated from sea water of the Yellow Sea in Korea. *Int. J. Syst. Evol. Microbiol.* 54, 1799–1803. doi:10.1099/ijs.0.63151-0.

Yoon, J. H., Lee, M. H., Kang, S. J., and Oh, T. K. (2007). *Marinobacter salicampi* sp. nov., isolated from a marine solar saltern in Korea. *Int. J. Syst. Evol. Microbiol.* 57, 2102–2105. doi:10.1099/ijs.0.65197-0.

Zablocki, O., Michelsen, M., Burris, M., Solonenko, N., Warwick-Dugdale, J., Ghosh, R., et al. (2021). VirION2: a short- and long-read sequencing and informatics workflow to study the genomic diversity of viruses in nature. *PeerJ* 9, e11088. doi:10.7717/peerj.11088.

Zhang, D. C., Li, H. R., Xin, Y. H., Chi, Z. M., Zhou, P. J., and Yu, Y. (2008). *Marinobacter psychrophilus* sp. nov., a psychrophilic bacterium isolated from the Arctic. *Int. J. Syst. Evol. Microbiol.* 58, 1463–1466. doi:10.1099/ijs.0.65690-0.

Zhang, Y., Zhong, X. C., Xu, W., Lu, D. C., Zhao, J. X., and Du, Z. J. (2020). *Marinobacter vulgaris* sp. Nov., a moderately halophilic bacterium isolated from a marine solar saltern. *Int. J. Syst. Evol. Microbiol.* 70, 450–456. doi:10.1099/ijsem.0.003774.

Zhong, Z. P., Liu, Y., Liu, H. C., Wang, F., Zhou, Y. G., and Liu, Z. P. (2015). *Marinobacter halophilus* sp. nov., a halophilic bacterium isolated from a salt lake. *Int. J. Syst. Evol. Microbiol.* 65, 2838–2845. doi:10.1099/ijs.0.000338.

Zhong, Z.-P., Rapp, J. Z., Wainaina, J. M., Solonenko, N. E., Maughan, H., Carpenter, S. D., et al. (2020). Viral ecogenomics of Arctic cryopeg brine and sea ice. *mSystems* 5. doi:10.1128/msystems.00246-20.

Zhu, Q., Kosoy, M., and Dittmar, K. (2014). HGTector: An automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* 15. doi:10.1186/1471-2164-15-717.

ZoBell, C. E. (1946). Studies on redox potential of marine sediments. *Am. Assoc. Pet. Geol. Bull.* 30, 477–513. doi:10.1306/3d933808-16b1-11d7-8645000102c1865d

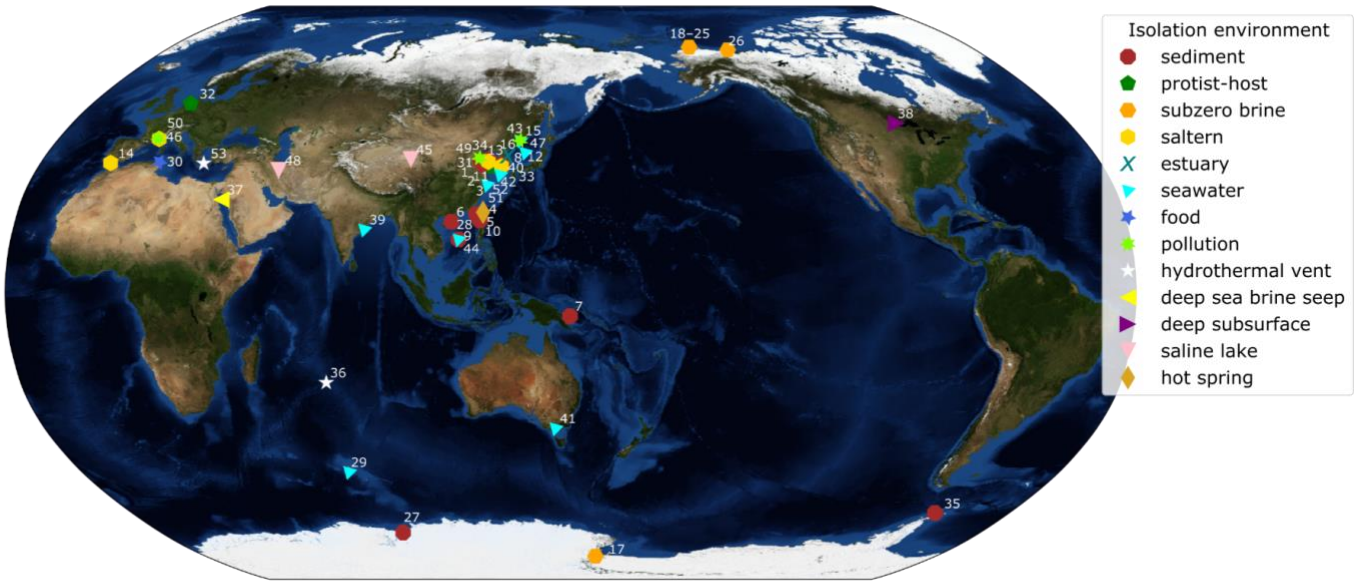

Figure 4.1. World map with isolation locations of each species of *Marinobacter* represented in the pangenome.
Each symbol and color represents the isolation environment of the species or genome used here. Each point on the map is annotated with an ID number linking that point to its genome description in Table 4.1. Annotations have been adjusted manually for readability and correspond to a point in the cluster if the area is too crowded. The background image used for the map is the Blue Marble: Next Generation Topography and Bathymetry image created by the NASA Earth Observatory.

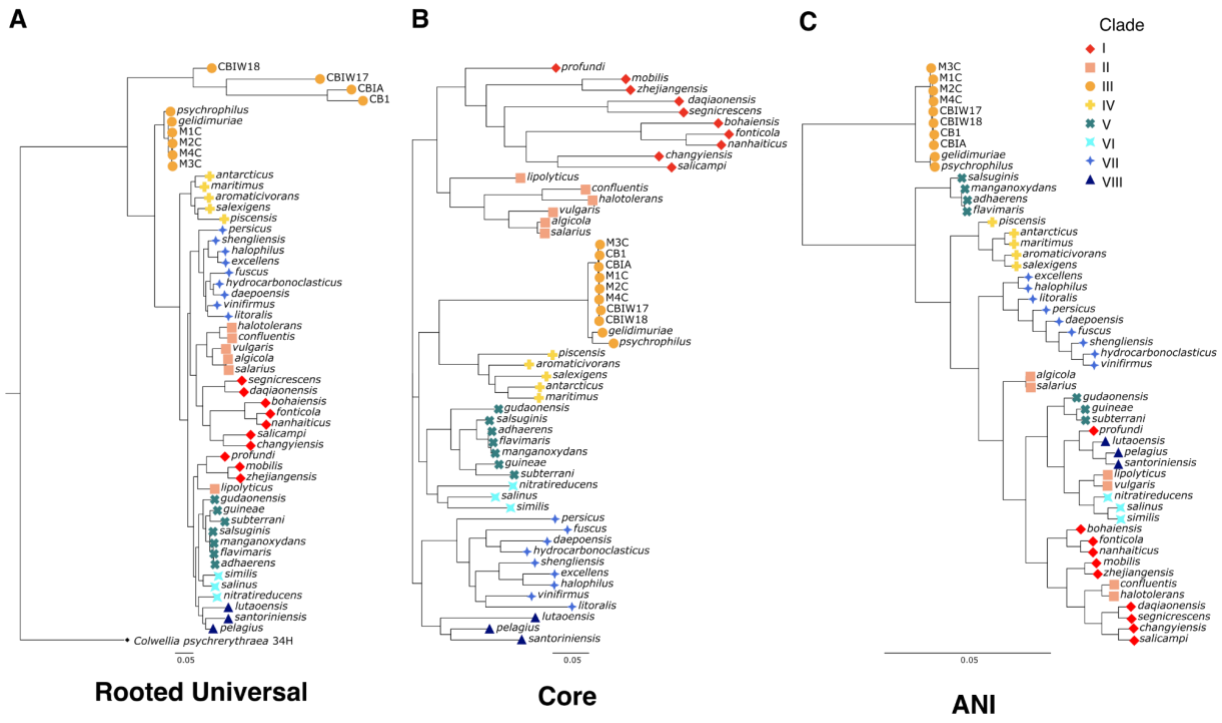

Figure 4.2. Phylogenetic trees for the genus *Marinobacter*
A) "Rooted universal", B) "Core", and C) "ANI" based phylogenetic trees for the 8 new genomes from cryopegs and 45 other representatives of *Marinobacter* species. The rooted universal tree (A) is constructed from a set of universal single copy bacterial marker genes and includes *Colwellia psychrerythraea* 34H as an outgroup. The core tree (B) is constructed from a set of genes found in single copies in each genome that have ≥ 95% geometric homogeneity and ≤ 90% functional homogeneity across the pangenome. The ANI tree (C) represents the relatedness of each genome based on percentage identity of shared genomic content. The leaves of each tree are coded by color and shape according to the clade to which they belong.

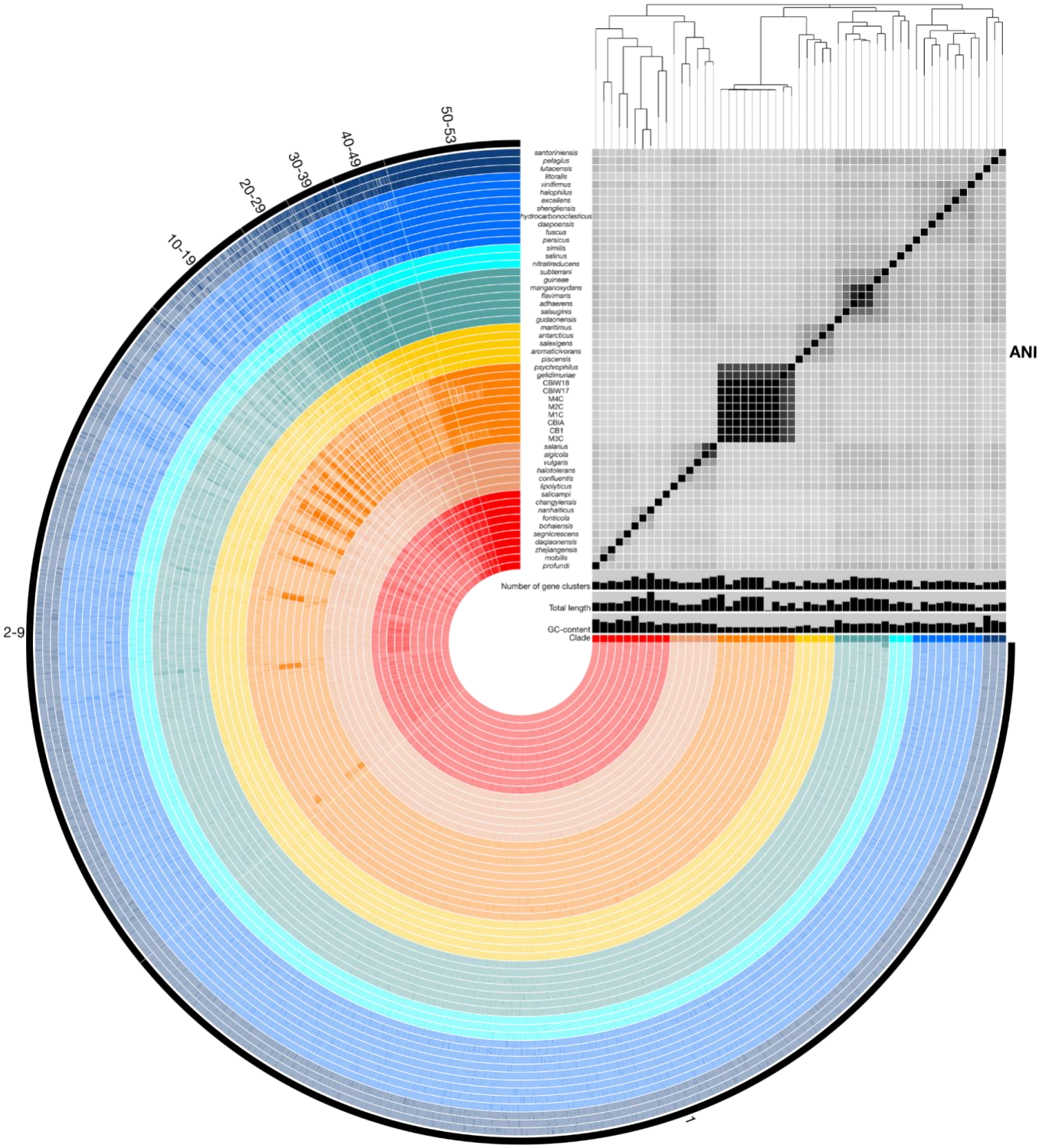

Figure 4.3. Representation of the *Marinobacter* pangenome with ANI heatmap. Interior circles are colored according to the clade to which the species belongs. Clades are ordered from I (red) to VIII (dark blue) from the innermost to outermost rings. Brighter lines in

these interior circles of the pangenome represent gene cluster presence in each genome. Numbered exterior bars indicate the number of genomes containing gene clusters in each region of the pangenome. The bar graphs display GC-content (50–65%), total genome length (3,000,000–5,358,909 bp), and number of gene clusters (2500–5000) per genome. The heatmap represents ANI values between shared regions of each genome (70–100%). The phylogenetic tree above the heatmap displays the core-gene phylogeny of the pangenome and serves as the basis for clade designations.

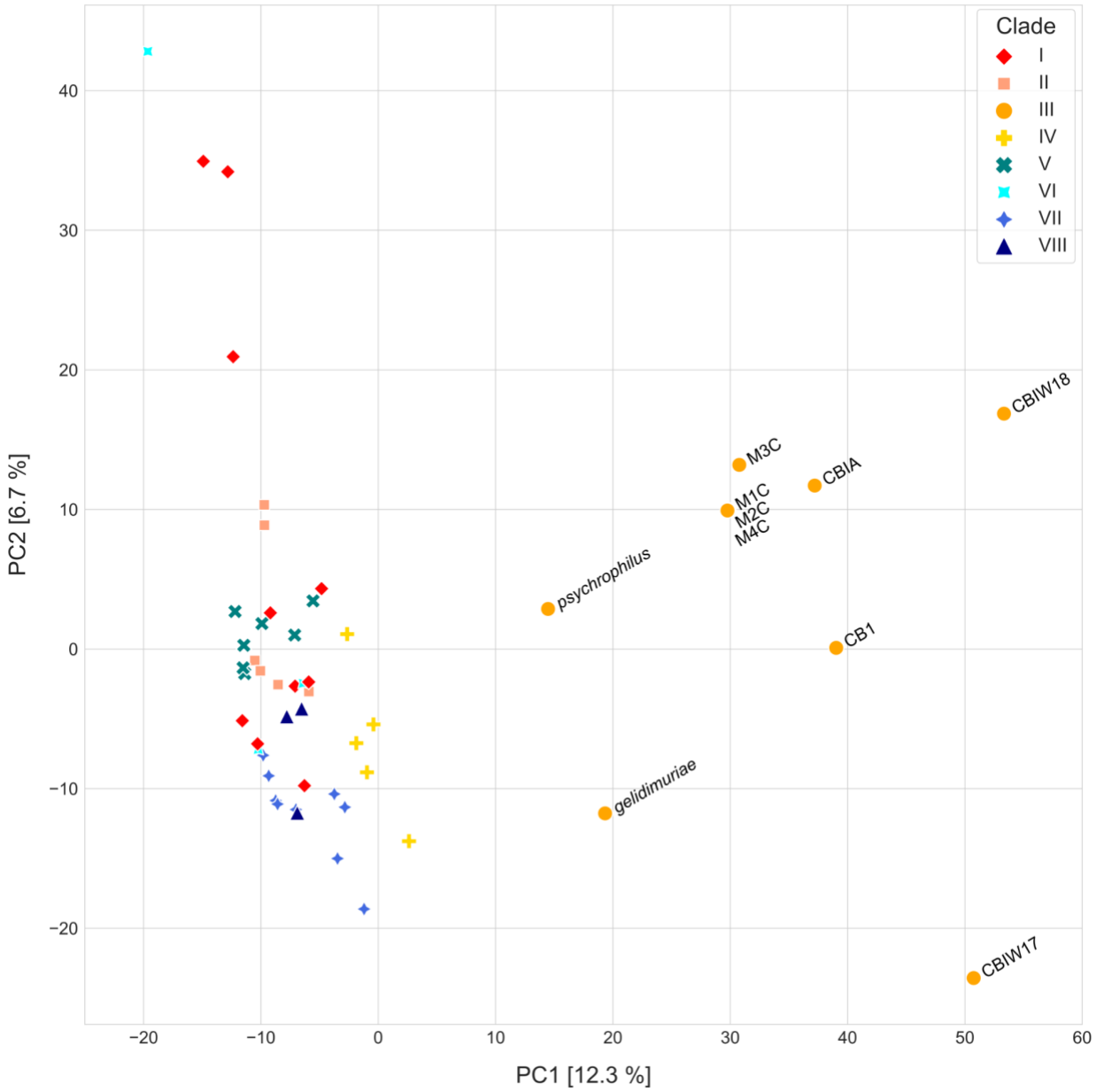

Figure 4.4. Principal component analysis of gene frequency in each genome in the *Marinobacter* pangenome.
Symbols are colored according to their phylogenetic clade assignment. Clade III genomes, appearing on the right side of PC1, are labelled by the *Marinobacter* species name or isolate number.

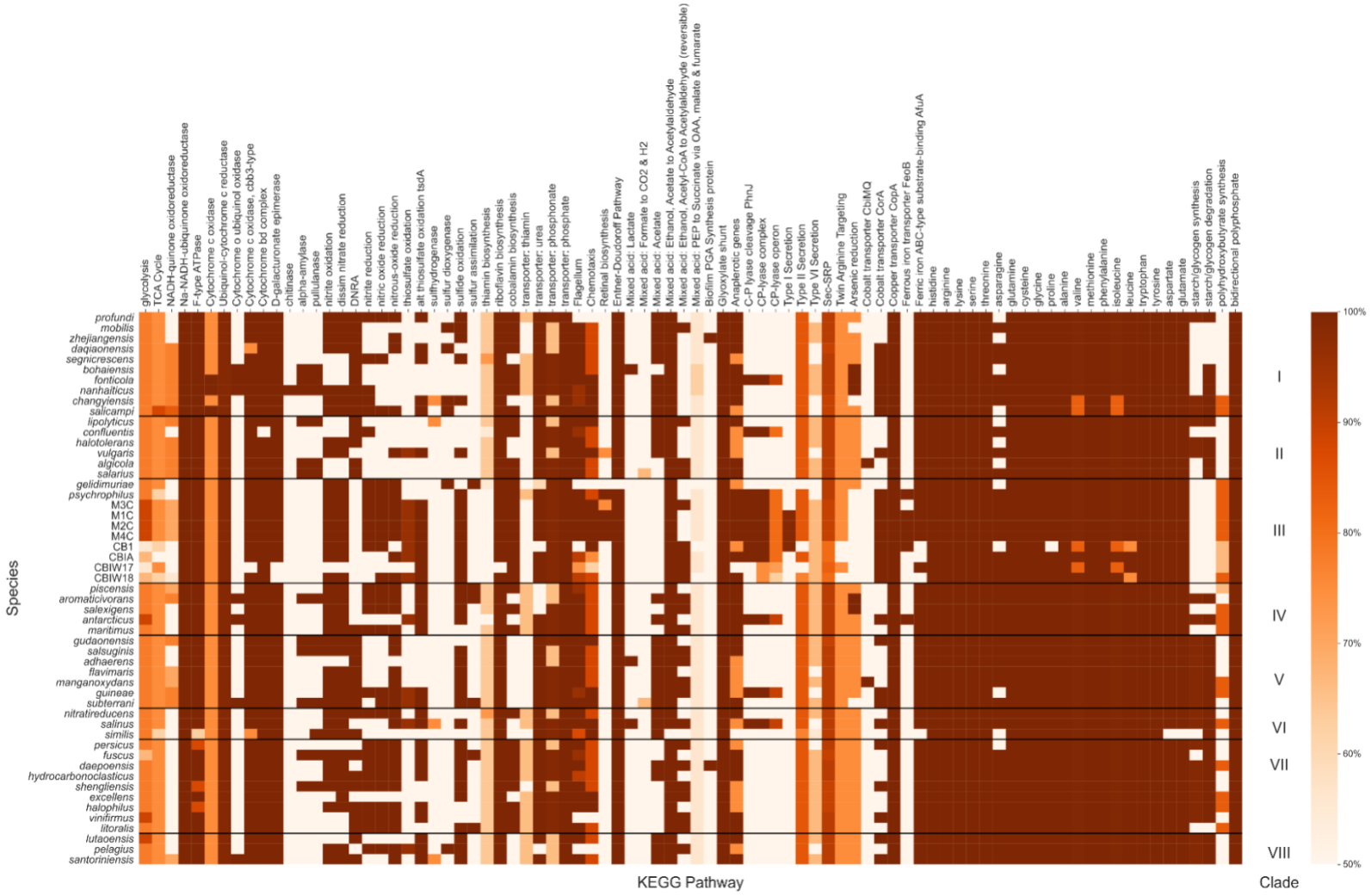

Figure 4.5. Heatmap of the completeness level of KEGG metabolic pathways identified by KEGG Decoder for *Marinobacter* genomes.
The colormap ranges from 50% (lightest orange) to 100% (darkest orange). Each row represents a single *Marinobacter* genome, and each column represents a metabolic pathway. Black horizontal lines separate the heatmap into clades with clade numbers listed between the heatmap and the scale bar.

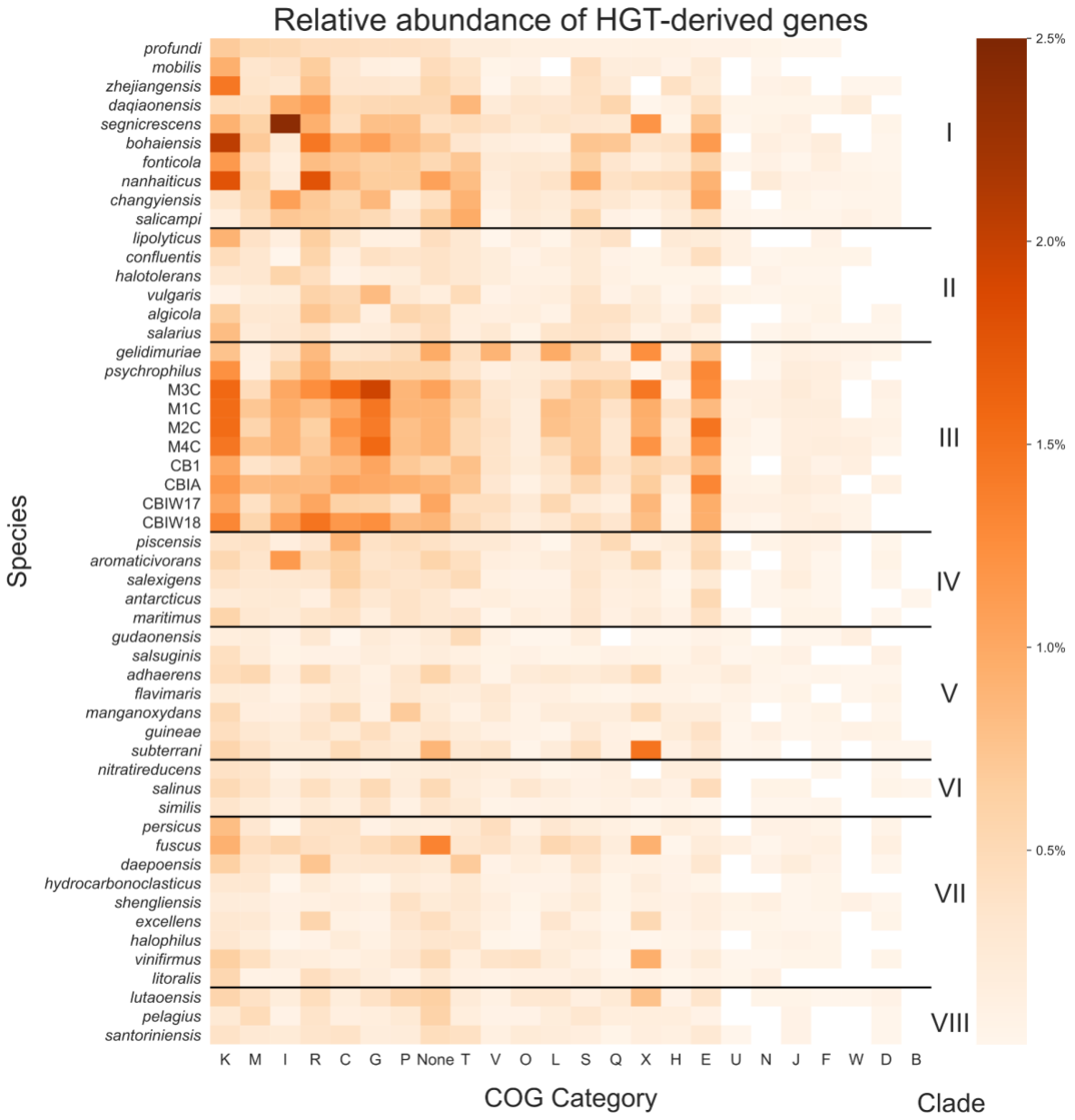

Figure 4.6. Heatmap of the relative abundance of genes derived by HGT in each *Marinobacter* genome.

Each row represents a single *Marinobacter* genome, and each column represents a COG category. Normalization was done by counting the number of potentially HGT derived genes, dividing by the total number of genes in a genome, and converting to a percentage. The colormap ranges from 0% (lightest orange) to 2.5% (darkest orange). Black horizontal lines separate the heatmap into clades in order with clade numbers listed between the heatmap and the scale bar.

COG Categories are defined as follows: K for transcription; M, cell wall/membrane/envelope biogenesis; I, lipid transport and metabolism; R, general function prediction only; C, energy production and conversion; G, carbohydrate transport and metabolism; P, inorganic ion transport and metabolism; T, signal transduction mechanisms; V, defense mechanisms; O, posttranslational modification, protein turnover, chaperones; L, replication, recombination and repair; S, function unknown; Q, secondary metabolites biosynthesis, transport and catabolism; X, mobilome: prophages, transposons; H, coenzyme transport and metabolism; E, amino acid transport and metabolism; U, intracellular trafficking, secretion, and vesicular transport; N, cell motility; J, translation, ribosomal structure and biogenesis; F, nucleotide transport and metabolism; W, extracellular structures; D, cell cycle control, cell division, chromosome partitioning; B, chromatin structure and dynamics; and None, no COG annotation available.

Chapter 5. CONCLUSION

The goal of this dissertation was to understand the evolution and ecology of microbial communities (primarily the Prokaryotes) that inhabit the subzero brine environments of cryopeg and sea ice. These extreme environments impart limits on life in a manner that is both highly relevant to our understanding of critically endangered cryosphere ecosystems on Earth and to our search for life in habitable environments on other celestial bodies. The background, motivation, and hypotheses defining this topic have been rendered in careful detail in this dissertation, and I use this passage to briefly discuss how each chapter works together to create a cohesive discussion of the evolutionary and ecological aspects of life in these extreme environments.

As an oceanographer and astrobiologist, I consider exploration of new frontiers to be paramount. These fields strive to push the boundaries of what we know about the natural world and to uncover the details of environments that are farthest from the reaches of humankind. Chapter 2 of this dissertation covers the initial findings of an expedition to the Alaskan Arctic coast to collect brines from sea ice and the rarely sampled cryopeg environment, and it sets the stage for follow-up research on the endemic communities of these extreme environments (Cooper et al., 2019, Chapter 2). Before this project, the microbiology of cryopegs had been studied primarily in the context of culturable organisms, and no *in situ* characterization of the microbial community structures was available (Gilichinsky et al., 2003; Pecheritsyna et al., 2007). We found the Alaskan cryopeg brines we sampled to be densely populated, characterized by low measures of biodiversity, and almost exclusively inhabited by bacteria and their viruses, and we discovered that the genus *Marinobacter* composed a remarkable majority of these communities. This genus had not previously been reported as an important member of the prokaryotic communities of cryopegs. We also found that these communities defied our working

hypothesis by differing significantly in taxonomy from similarly cold and saline sea ice brines. This chapter revealed that cryopegs are not only inhabited by bacteria and viruses that manage to survive, but that they harbor bacteria that are well adapted to and capable of flourishing in the continuously extreme conditions of subzero brines (Cooper et al., 2019, Chapter 2; Appendix C).

Building on the discovery of remarkably dense and well-adapted bacterial communities in cryopeg brines and observations of their differences from sea ice communities, I set out to gain a deeper understanding of the evolutionary patterns that govern the persistence of life in these subzero brine environments. Though cryopegs and sea ice can be similar in temperature and salinity, they are characterized by drastically different periods of geophysical stability along with different nutrient availability regimes. Sea ice fluctuates greatly in temperature over a matter of hours to months, which drives changes in salinity and connectivity of brines with the seawater below and the snow and atmosphere above, and it nearly completely melts and refreezes on an annual cycle. In contrast, cryopegs remain within a tight range of temperatures that remain below zero continuously over the course of years, and the system we studied has been isolated from the ocean and atmosphere within permafrost for approximately 40,000 years (Iwahana et al., 2021). With this difference in mind, we hypothesized that populations of bacteria inhabiting cryopeg brines would need to utilize evolutionary mechanisms alternative to vertical inheritance that account for decreased rates of evolution relative to the rate of selection by environmental factors. In Chapter 3, I found that populations that inhabit cryopeg brines have a greater selection for mobilome-associated genes and generally have lower levels of genomic microdiversity than the populations of sea ice brines. This finding indicates that cryopeg brine populations contain genes that are highly selected for and that allow little room for deviation, in contrast to sea ice brine populations that vary more within populations. Based on the conclusions

of this chapter, I have hypothesized that horizontal gene transfer is a crucial mechanism for the persistence of life under multiple extreme conditions. This pattern is consistent for most populations we observed in cryopegs and has been supported by a metagenomic survey conducted on these samples (Rapp et al., 2021). It also opens additional questions on the specific genomic characteristics that have allowed a single species of *Marinobacter* to dominate in these brines.

Metagenomics is an incredible technology that allows for the quantitative assessment of ecosystems *in situ* and is the foundation for the analyses I performed for Chapter 3. Metagenomics is not without limitations, however. Modern technologies are prone to errors at a low but not insignificant rate, and they fail to capture the complete genetic inventory of an ecosystem. In Chapter 4, I expanded our exploration of *Marinobacter* beyond metagenomics by bringing four strains of this genus into culture from cryopeg brine and obtaining complete sequences of their genomes. I combined these 4 complete genomes with 4 metagenome assembled genomes from cryopeg brines (as in Chapter 3) and with 45 representative genomes of other *Marinobacter* species to generate a pangenome of this genus. I found that the species of *Marinobacter* that inhabits cryopegs is a novel species with closest relation to the only other species of *Marinobacter* that inhabit subzero brines: *M. psychrophilus* from Arctic sea ice (Zhang et al., 2008) and *M. gelidimuriae* from an Antarctic subglacial aquifer at Blood Falls (Chua et al., 2018). Additionally, I identified a suite of genes that are enriched in this clade of *Marinobacter* that are clear signs of adaptation to the subzero brine environment (e.g. membrane transporters and cold shock transcriptional regulators), and I corroborated the findings from Chapter 3 that horizontal gene transfer is indeed quite important in the subzero brine environment. The pangenome of *Marinobacter* is remarkably diverse, and the species that

inhabit subzero brines appear to contain a plethora of physiological adaptations specifically selected for success in this extreme environment.

Altogether, the findings of this dissertation bring depth to the understanding of the evolution and ecology of the microbial communities that inhabit cryopeg and sea ice brines. Extremophiles are capable of tremendous amounts of adaptation and flexibility in the face of extremes. The changing climate of Earth will bring these communities that are already active within the cryosphere into a warmer and differently connected geophysical regime, where they will be competing for niches occupied by community members that perform better under less extreme conditions. The fate of these communities is currently unknown, but future research monitoring and modelling the dynamics of these ecosystems will be impactful for understanding the fate of microbial contributions to climate change. Additionally, these environments represent some of the best analogs for the Martian subsurface and for icy moons that exist on Earth, and making observations of these analog environments is informative to our expectations for the capabilities and persistence of life in expected habitable environments throughout the Solar System. The findings of this dissertation are encouraging. Bacteria are capable of tremendous adaptations through complex modes of evolution that have allowed them to flourish at the combined extremes of low temperatures and high salinities, as they appear to have adapted to other extremes encountered on Earth. If microbial life reminiscent of these communities has taken a foothold beyond Earth, I expect it to be capable of significant function and persistence by evolving to push the known biological limits imparted by physics and chemistry. The Universe is vast and ancient, and scientific exploration of the unknown has only just begun.

References

Cooper, Z. S., Rapp, J. Z., Carpenter, S. D., Iwahana, G., Eicken, H., & Deming, J. W. (2019). Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiology Ecology*, *95*(12), 1–15. https://doi.org/10.1093/femsec/fiz166

Chua, M. J., Campen, R. L., Wahl, L., Grzymski, J. J., & Mikucki, J. A. (2018). Genomic and physiological characterization and description of Marinobacter gelidimuriae sp. nov., a psychrophilic, moderate halophile from Blood Falls, an antarctic subglacial brine. *FEMS Microbiology Ecology*, *94*(3), 1–15. https://doi.org/10.1093/femsec/fiy021

Gilichinsky, D., Rivkina, E., Shcherbakova, V., Laurinavichuis, K., & Tiedje, J. (2003). Supercooled water brines within permafrost - An unknown ecological niche for microorganisms: A model for astrobiology. *Astrobiology*, *3*(2), 331–341. https://doi.org/10.1089/153110703769016424

Iwahana, G., Cooper, Z. S., Carpenter, S. D., Deming, J. W., & Eicken, H. (2021). Intra-ice and intra-sediment cryopeg brine occurrence in permafrost near Utqiaġvik (Barrow). *Permafrost and Periglacial Processes*, *32*(3), 427–446. https://doi.org/10.1002/ppp.2101

Pecheritsyna, S. A., Shcherbakova, V. A., Kholodov, A. L., Akimov, V. N., Abashina, T. N., Suzina, N. E., & Rivkina, E. M. (2007). Microbiological analysis of cryopegs from the Varandei Peninsula, Barents Sea. *Mikrobiologiia*, *76*(5), 694–701. https://doi.org/10.1134/S0026261707050153

Rapp, J. Z., Sullivan, M. B., & Deming, J. W. (2021). Divergent Genomic Adaptations in the Microbiomes of Arctic Subzero Sea-Ice and Cryopeg Brines. *Frontiers in Microbiology*, *12*(July), 1–21. https://doi.org/10.3389/fmicb.2021.701186

Zhang, D. C., Li, H. R., Xin, Y. H., Chi, Z. M., Zhou, P. J., & Yu, Y. (2008). Marinobacter psychrophilus sp. nov., a psychrophilic bacterium isolated from the Arctic. *International Journal of Systematic and Evolutionary Microbiology*, *58*(6), 1463–1466. https://doi.org/10.1099/ijs.0.65690-0

APPENDIX A: CHAPTER 2 SUPPLEMENTARY INFORMATION

Supplemental Text S2.1

Terminology. The historical naming of cryopeg boreholes (CBs) drilled in the Utqiaġvik Permafrost Tunnel includes some inconsistencies between publications. In particular, the boreholes drilled in 2004 and designated CB1 and CB3 in the geological work of Meyer *et al.* (2010a) were reversed and named CB3 and CB1 in the geomicrobiological work of Colangelo-Lillis *et al.* (2016). Because we are building on the latter work and using an archived brine sample from it, we have kept their designations. For the two new boreholes drilled by Colangelo-Lillis *et al.* (2016) in August 2009 and designated IWC (for ice wedge core) and IA (for ice auger), we converted IWC to CBIW and IA to CBIA to clarify that both are cryopeg boreholes (Figure 2.1). The specific term ice wedge is used in previous literature to describe the massive ice in which the tunnel was excavated. More extensive geological findings from the tunnel (G. Iwahana, unpublished data) do not validate the use of ice wedge for this setting, so we use the broader geological term "massive ice" throughout our text. For the geomicrobiological work done in 2009, 2017 and 2018, any new drilling or opening of existing holes was done using a carefully cleaned and ethanol-rinsed ice auger or SIPRE corer. Cleaning included scrubbing the drilling components with detergent, rinsing liberally with tap water, then with distilled water; after the final ethanol-rinse, drilling parts were bagged in sterile whirlpacks for transport to the tunnel and opened only at the point of drilling.

Sampling details by borehole

CBIW

In 2009, a borehole designated IWC was cored into the massive ice for geochemical and microbiological work using a SIPRE corer. This effort yielded a series of ice sections but no brine (Colangelo-Lillis *et al*. 2016). In 2017, we drilled deeper into IWC (renamed CBIW) and encountered brine but no permafrost or cryopeg sediment (only massive ice). A volume of 1.5 L of brine (maximum volume to date) was withdrawn using the sterile pumping technique, as described in Methods; the brine appeared yellow-brown in color. During this deeper drilling, ice core sections were collected for geochemical analyses (addressed in a separate paper). Although stringent sterile handling and controlled low-temperature melting of these ice sections was not followed post-sampling in the laboratory, we included two of the sections (3^{rd} and 7^{th}) in our study (designated CBIW_IW3 and CBIW_IW7), as they were later shown to contain brine of interest. In 2018, we again encountered brine in CBIW at the base of the borehole and withdrew two samples, separated in time by 4 days (designated CBIW_18 and CBIW_re_18, where "re" indicates *in situ* recharge of brine into the bottom of the borehole). After the last microbiological sampling, CBIW was drilled even deeper, finally encountering permafrost but no cryopeg sediment. The three brine samples obtained from CBIW (CBIW_17, CBIW_18, and CBIW_re_18) are called intra-ice brine, as the drilling results indicated the brine-yielding pocket was bounded by massive ice. Whether this brine may have been connected laterally to nearby intra-sediment brine is not known (question marks in Figure 2.1).

CB1

Two brine samples were obtained from this pre-existing borehole (Meyer et al. 2010a), one in August 2009 (Colangelo-Lillis et al. 2016) archived at –70°C since then, and the other in May 2018 (this study). As the 2009 sample was size-fractionated, three samples were available for 16S rRNA gene sequencing: CB1_0.2_09, CB1_3.0_09, and CB1_18, where the last number indicates sampling year (2009 or 2018) and the middle number indicates size fraction (0.2 for 0.2–3.0 µm, 3.0 for > 3.0 µm; Table 2.1). When drilled in 2004, this borehole (with PVC pipe installed and capped) had not yielded any mobile brine. In 2009, brine was present and sampled until no further liquid could be withdrawn; the brine was orange-brown in color. In 2017, we removed the cap from CB1 finding it filled with ice and did not sample, but in 2018, we removed the ice filling the top of the pipe and observed that brine had recharged to a depth of 113 cm into the borehole. This brine was sampled until no further liquid could be withdrawn. The brine from CB1 is called intra-sediment brine, as it was derived from cryopeg sediment at the base of the borehole.

CBIA

In 2009, this borehole was drilled using an ice auger (designated CBIA in this study; Figure 2.1) and found to be dry. In 2017, we found CBIA iced-over (no pipe had been installed earlier) and made no attempt to reopen it (re-drill it). In May 2018, we observed a discrete area (~ 4 cm^2) of brine seepage onto the tunnel floor at the point where CBIA had been previously drilled. This brine, orange-brown in color, was collected using a sterile syringe (sample designated CBIA_surf_18). We then re-opened CBIA by ice auger, encountered brine, and withdrew as much as possible. The sampled brine unavoidably contained some ice crystals that had formed in

the borehole since 2009). A sterile PVC pipe was installed for possible future sampling (beyond this study).

G1, G2

In 2018, additional boreholes, labeled simply G1, G2, etc., were drilled for geological work using cleaned (but not ethanol-rinsed) drilling components. G1 was dry, but G2 yielded brine that we sampled using our sterile withdrawal technique, as described in Methods. Brine withdrawn from G2 (sample designated G2_18) was unavoidably slurried with sediment and possible ice crystals and appeared chalky gray in color. Other boreholes in the G series are not shown on Figure 2.1, as they were drilled after all microbiological work in the tunnel had been completed.

CB4

In 2017, we drilled new borehole CB4, 75 cm west of CBIW (the most westerly sample in this study), through the massive ice and underlying permafrost into unfrozen cryopeg sediments (Figure 2.1), but no liquid brine was encountered. At two points during the drilling of CB4 (initial and deeper drilling), ice-auger shavings through the massive ice were collected with an ethanol-sterilized spatula (samples designated CB4_IW1 and CB4_IW2). These brine-free ice shavings contained a scattering of visible granular matter, which may account for the higher than expected bacterial counts for such freshwater ice samples (Table 2.3). In 2018, we found that brine had filled the borehole pipe to within 29 cm of the tunnel floor. We collected the brine (sample designated CB4_18) until no further liquid could be withdrawn. After sampling in 2018, a sterile PVC pipe was installed for future sampling.

CB2, CB3, CB5, CB6, CB7

These boreholes did not yield brine samples for this study but did provide other information. CB2 and CB3 were pre-existing boreholes (drilled in 2004 along with CB1; Meyer et al. 2010a), while CB5–7 were drilled in 2018 using sterile technique to the fullest extent possible. CB2 was dry in 2004 and had frozen over by 2009, when it was found filled with orange-brown ice crystals (Colangelo-Lillis et al. 2016). CB3 had yielded a small volume of brine in 2009 (no archived sample), but it had been fitted (and sealed) with thermoprobes for *in situ* temperature recordings between December 2010 and March 2012 (K. Yoshikawa, personal communication). We did not attempt to reopen or access CB2 or CB3 in 2017 or 2018. Boreholes CB5 and CB6 were drilled in 2018 in search of additional brine, but all were dry. Borehole CB7 yielded a small volume of brine, but none was collected for this study.

Supplemental Text S2.2 – Bioinformatic codes

Mothur commands for sequence processing

Mac version

mothur v.1.40.1

Last updated: 04/16/18

by

Patrick D. Schloss

#use set.dir() to set working directory

mothur > make.contigs(file=cryo.files)

mothur > screen.seqs(fasta=cryo.trim.contigs.fasta, group=cryo.contigs.groups,
summary=cryo.trim.contigs.summary, minlength=400, maxlength=480, maxambig=0)

mothur > unique.seqs(fasta=current)

mothur > count.seqs(name=current, group=current)

mothur > pcr.seqs(fasta=silva.nr_v132.align, start=6387, end=25319, keepdots=F) #trim the Silva v132 database to the region containing these sequences. Region was found after an initial attempt at aligning without trimming the database.

mothur > rename.file(input=silva.nr_v132.pcr.align, new=silva.v4.fasta) #I know the V4 in the name may be misleading at this point, but I was following the convention in the SOP.

mothur > align.seqs(fasta=cryo.trim.contigs.good.unique.fasta, reference=silva.v4.fasta)

mothur > screen.seqs(fasta=cryo.trim.contigs.good.unique.align,
count=cryo.trim.contigs.good.count_table,
summary=cryo.trim.contigs.good.unique.summary, maxhomop=9, start=15, end=18932)

mothur > filter.seqs(fasta=cryo.trim.contigs.good.unique.good.align, vertical=T, trump=.)

mothur > unique.seqs(fasta=current, count=current)

mothur > pre.cluster(fasta=cryo.trim.contigs.good.unique.good.filter.unique.fasta,
count=cryo.trim.contigs.good.unique.good.filter.count_table, diffs=4)

mothur >
chimera.vsearch(fasta=cryo.trim.contigs.good.unique.good.filter.unique.precluster.fasta,
count=cryo.trim.contigs.good.unique.good.filter.unique.precluster.count_table,
dereplicate=t)

mothur > remove.seqs(fasta=cryo.trim.contigs.good.unique.good.filter.unique.precluster.fasta,
accnos=cryo.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.accn
os)

mothur > classify.seqs(fasta=current, count=current, reference=silva.v4.fasta,
taxonomy=/Users/zac/mothur/silva.nr_v132.tax, cutoff=80)

mothur >
remove.lineage(fasta=cryo.trim.contigs.good.unique.good.filter.unique.precluster.pick.fas
ta, count=current,
taxonomy=cryo.trim.contigs.good.unique.good.filter.unique.precluster.pick.nr_v132.wan
g.taxonomy, taxon=Chloroplast-Mitochondria-Eukaryota-unknown)

mothur > summary.tax(taxonomy=current, count=current)

#To reduce computational load we used the Opticlust dgc method for OTU clusting instead of the default dist.seq() and cluster() method.

#To run the dgc clustering method, mothur had to be updated to v 1.40.5 due to a bug that wouldn't allow the dgc method to run without the unnecessary distance matrix.

Mac version

mothur v.1.40.5

Last updated: 06/19/18

by

Patrick D. Schloss

mothur >

cluster(fasta=cryo.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.fasta, count=cryo.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.count_table, method=dgc, cutoff=0.03)

mothur >

make.shared(list=cryo.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.dgc.list, count=cryo.trim.contigs.good.unique.good.filter.unique.precluster.denovo.vsearch.pick.pick.counttable, label=0.03)

mothur >

remove.rare(shared=cryo.trim.contigs.good.unique.good.filter.unique.precluster.pick.pick.dgc.shared, bygroup=t, nseqs=1) #removing singletons is not a default step in the SOP.

mothur > remove.rare(list=current, count=current, bygroup=t, nseqs=1)

mothur > classify.otu(list=current, count=current, taxonomy=cryo.trim.contigs.good.unique.good.filter.unique.precluster.pick.nr_v132.wang.pick.taxonomy, label=0.03)

mothur > count.groups(shared=current)

Alpha diversity calculations and plotting in Phyloseq (R script)

```
#initiate phyloseq/ggplot2
library("phyloseq");packageVersion("phyloseq")
library("ggplot2");packageVersion("ggplot2")

#import mothur data to physeq object; uses “shared” and “taxonomy” files from mothur
processing.
cryo <- import_mothur(mothur_shared_file = "cryo.shared", mothur_constaxonomy_file =
        "cryo.taxonomy")

#add sample data to physeq object; matrix of samples by sample type descriptor
samdat = sample_data(data.frame(read.table(file="expsam", header=TRUE, sep="\t"),
        row.names=1))
cryo<-`sample_data<-`(cryo, samdat)

#Remove “Blank” sample from phyloseq object
cryo_sam <- subset_samples(cryo, SampleType !="Blank")

#To avoid the risk of artificially inflated diversity estimates caused by erroneous reads, we
removed singletons from the dataset.
cryo_d <- filter_taxa(cryo_sam, function(x) sum(x)>1, TRUE)
sample_sums(cryo_d)
```

```
richest3 <- estimate_richness(cryo_d, measures=c("Observed", "Chao1", "Shannon",
        "InvSimpson"))

#Vector of sample by type in order that they exist in Phyloseq object
samtvec3 <- c('Cryopeg brine','Cryopeg brine','Massive ice','Massive ice','Cryopeg
        brine','Cryopeg brine','Cryopeg brine','Massive ice','Massive ice','Cryopeg brine','Cryopeg
        brine','Cryopeg brine','Cryopeg brine','Cryopeg brine','Sea ice brine','Sea ice brine','Sea
        ice brine','Sea ice brine','Sea ice brine','Sea ice brine','Sea ice brine','Sea ice brine')
richest3$SampleType <- as.factor(samtvec3)

#Run anovas for different diversity indices and output anova summaries
shannonanova = aov(Shannon ~ SampleType, data=richest3)
summary(shannonanova)
invsimanova = aov(InvSimpson ~ SampleType, data=richest3)
summary(invsimanova)
obsanova = aov(Observed ~ SampleType, data=richest3)
summary(obsanova)
chaoanova = aov(Chao1 ~ SampleType, data=richest3)
summary(chaoanova)

#run TukeyHSD to find pair-wise relationships
shpw = TukeyHSD(shannonanova)
ispw = TukeyHSD(invsimanova)
```

```
obpw = TukeyHSD(obsanova)

chpw = TukeyHSD(chaoanova)
```

NMDS and ANOSIM in Phyloseq (R script)

```
#initiate phyloseq/ggplot2

library("phyloseq");packageVersion("phyloseq")

library("ggplot2");packageVersion("ggplot2")

library("plyr"); packageVersion("plyr")

library("vegan"); packageVersion("vegan")


#import mothur data to physeq object

cryo <- import_mothur(mothur_shared_file = "cryo.shared", mothur_constaxonomy_file =
        "cryo.taxonomy")


#Fix imported taxonomy file to include taxonomic ranks

colnames(tax_table(cryo)) <- c("kingdom", "phylum", "class", "order", "family", "genus")


#Add sample data to Phyloseq object

samdat = sample_data(data.frame(read.table(file="samextra1.txt", header=TRUE, sep="\t"),
        row.names=1))

cryo<-`sample_data<-`(cryo, samdat)


#Create object without blank data
```

```
cryo_sam <- subset_samples(cryo, SampleType !="Blank")

#reorder levels by SampleType
sample_data(cryo_sam)$SampleType = factor(sample_data(cryo_sam)$SampleType,
        levels=c("Massive ice", "Massive ice with brine", "Intra-sediment brine", "Intra-ice
        brine", "Intra-sediment brine (other)", "Sea ice brine"))

#Remove singletons; apply log10(x + 1) transform of read counts
cryo_d <- filter_taxa(cryo_sam, function(x) sum(x)>1, TRUE)
t <- log10(cryo_d@otu_table + 1)
cryo_rel <-`otu_table<-`(cryo_d, t)
#Setting desired order of samples for plotting
ST = get_variable(cryo_rel, 'SampleType')
lev = c("Massive ice", "Massive ice with brine","Intra-ice brine", "Intra-sediment brine (other)",
        "Intra-sediment brine", "Sea ice brine")
ST <- factor(ST, levels=lev)
otus <- data.frame(cryo_rel@otu_table)

#Analysis of similarities
cryo_ano = anosim(t(otus), ST, permutations = 1000)
sig = cryo_ano$signif
stat = cryo_ano$statistic
```

```
#NMDS
ord <- ordinate(cryo_rel, "NMDS", "bray", iters=1000)
```

Principal component analysis

```
library('ggfortify');packageVersion('ggfortify')
library('ggplot2')
#Use matrix of samples by biogeochemical values
df <- data.frame(read.table(file='cryopca.txt', header=TRUE, sep='\t'),
                 row.names=1)
cryo.pca <- prcomp(df, center=TRUE, scale=TRUE)
pcasum <-summary(cryo.pca)
load <-cryo.pca$rotation
```

Supplemental Text S2.3

Metagenome sequencing and processing

A metagenomic library for cryopeg sample CBIW 2017 was generated, courtesy of the R. Mackelprang laboratory, from 2 ng input DNA using emulsion PCR methods described previously (Blow *et al.* 2008; Mackelprang *et al.* 2011, 2017). The library was sequenced on an Illumina HiSeq 4000, generating a total of 12,414,616 reads, with a length of 100 bp. We used *BBDuk* v35.68 (https://sourceforge.net/projects/bbmap) to decontaminate the raw metagenomic reads using kmers. For the removal of common contaminants, i.e. PhiX, we applied k=28, and for adapter trimming k=27 and mink=12. Reads were then quality-filtered using *Trimmomatic* v0.35 (Bolger et al. 2014) with a sliding window of 4:15, cutting a read once the average quality within a 4 base window falls below a quality threshold of 15 (based on Phred-33), and discarding reads shorter than 100 bp. This approach resulted in 9,367,618 remaining reads (75.5% of raw), with the majority being paired (95%). *Pear* v0.9.6 (Zhang et al. 2014) was used in default settings, and a minimum overlap of 10 bp as well as a minimum assembly length of 100 bp, to merge the paired reads. Next, the merged reads were combined with the non-merged and the remaining single reads, and all reads were filtered for ribosomal RNA genes using *SortMeRNA* v2.0 with default settings (Kopylova et al. 2012). A total of 25,600 rRNA gene fragments were detected and submitted to the NGS analysis pipeline of the SILVA rRNA gene database project (SILVAngs 1.3) (Quast et al. 2013). Each read was aligned using the SILVA Incremental Aligner SINA v1.2.10 (Pruesse et al. 2012) against the SILVA SSU rRNA SEED and quality-controlled with default settings. The pipeline clustered OTUs at 98% identity using cd-hit-est version 3.1.2 (www.bioinformatics.org/cd-hit) (Li and Godzik 2006).

Table S2.1. Alpha diversity statistics for each sample after singleton removal.

		Richness		Diversity		
#	Sample	Observed (# OTUs)	Chao1	Shannon	Inverse Simpson	# Reads
1	CB4_IW1_17	1704	1844	4.19	9.65	88621
2	CB4_IW2_17	1638	1739	3.64	6.57	100004
3	CBIW_IW3_17	1162	1540	2.64	5.67	113957
4	CBIW_IW7_17	486	968	2.23	4.76	86131
5	CBIW_17	504	803	2.04	3.88	91591
6	CBIW_18	368	546	1.54	2.46	143814
7	CBIW_re_18	246	380	1.37	2.08	92687
8	CB1_0.2_09	338	560	1.69	3.93	101462
9	CB1_3.0_09	256	460	1.61	3.63	97647
10	CB1_18	213	431	2.21	5.62	40943
11	CBIA_surf_18	435	770	2.16	3.74	31401
12	CBIA_18	385	618	1.99	3.63	139534
13	G2_18	426	502	2.49	3.38	37337
14	CB4_18	136	171	2.13	3.42	61497
15	SB_17	257	336	2.33	3.24	125456
16	SB_0.2_18	444	565	2.94	6.24	88115
17	SB_3.0_18	433	501	3.23	6.58	29543
18	SB_18	249	351	2.40	4.71	56130
19	SB1_18	584	675	2.50	4.29	41294
20	SB2_18	500	780	1.90	2.68	58607
21	SB3_18	554	848	1.93	2.84	74082
22	SB4_18	424	610	2.65	6.63	68657

References

Blow MJ, Zhang T, Woyke T et al. Identification of ancient remains through genomic sequencing. *Genome Res* 2008;**18**:1347–53.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20.

Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 2012;**28**:3211–7.

Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006;**22**:1658–9.

Mackelprang R, Waldrop MP, DeAngelis KM *et al*. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* 2011;**480**:368–71.

Mackelprang R, Burkert A, Haw M et al. Microbial survival strategies in ancient permafrost: insights from metagenomics. *ISME J* 2017;**11**:2305–18.

Pruesse E, Peplies J, Glöckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;**28**:1823–9.

Quast C, Pruesse E, Yilmaz P et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 2012;**41**:D590–6.

Zhang J, Kobert K, Flouri T et al. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 2014;**30**:614–20.

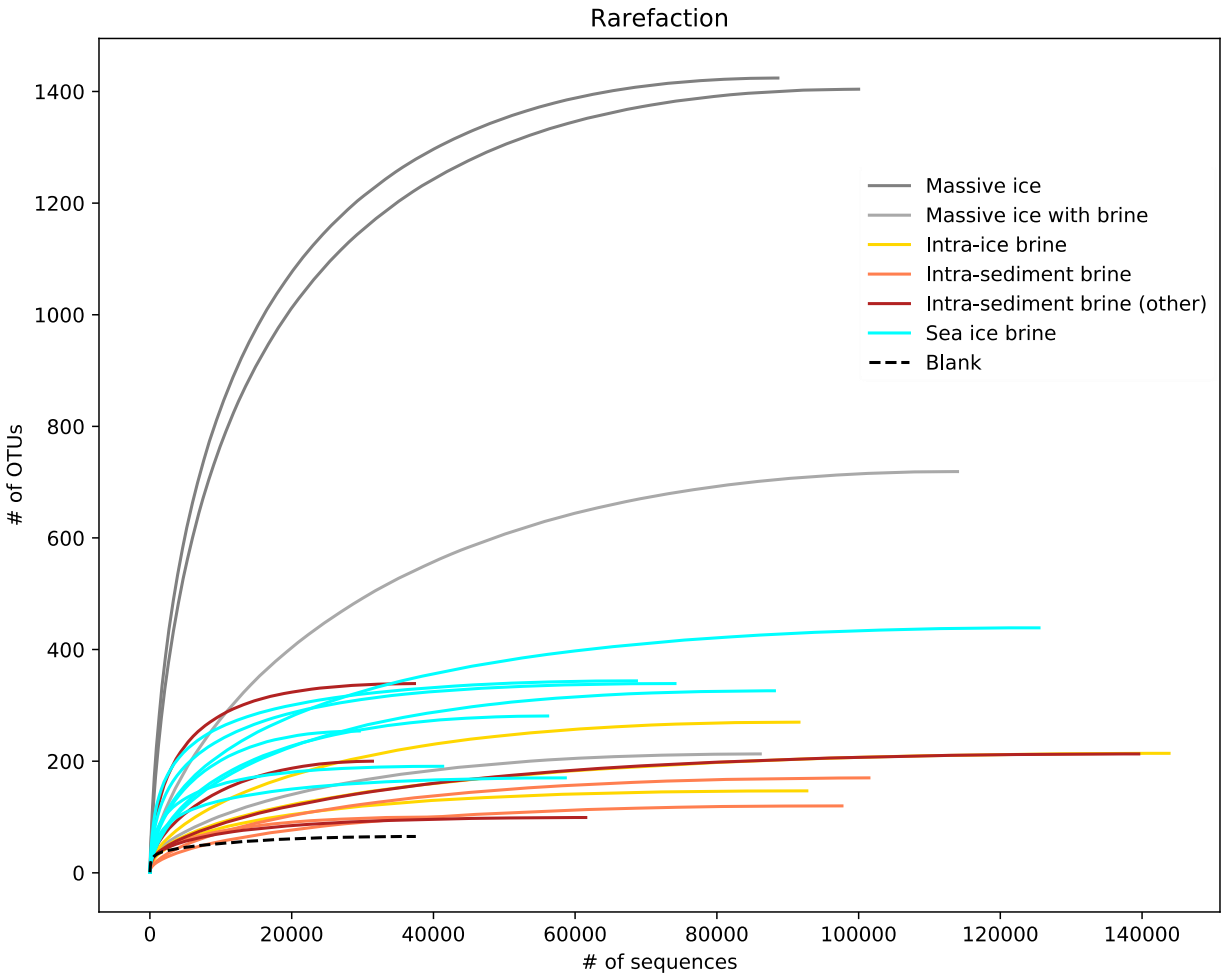

Figure S2.1. Rarefaction curves depicting number of OTUs (97% similarity) identified per number of sequences analyzed.

Sample subtypes match those in Table 2.1 and color code for cryopeg samples matches Figure 2.1.

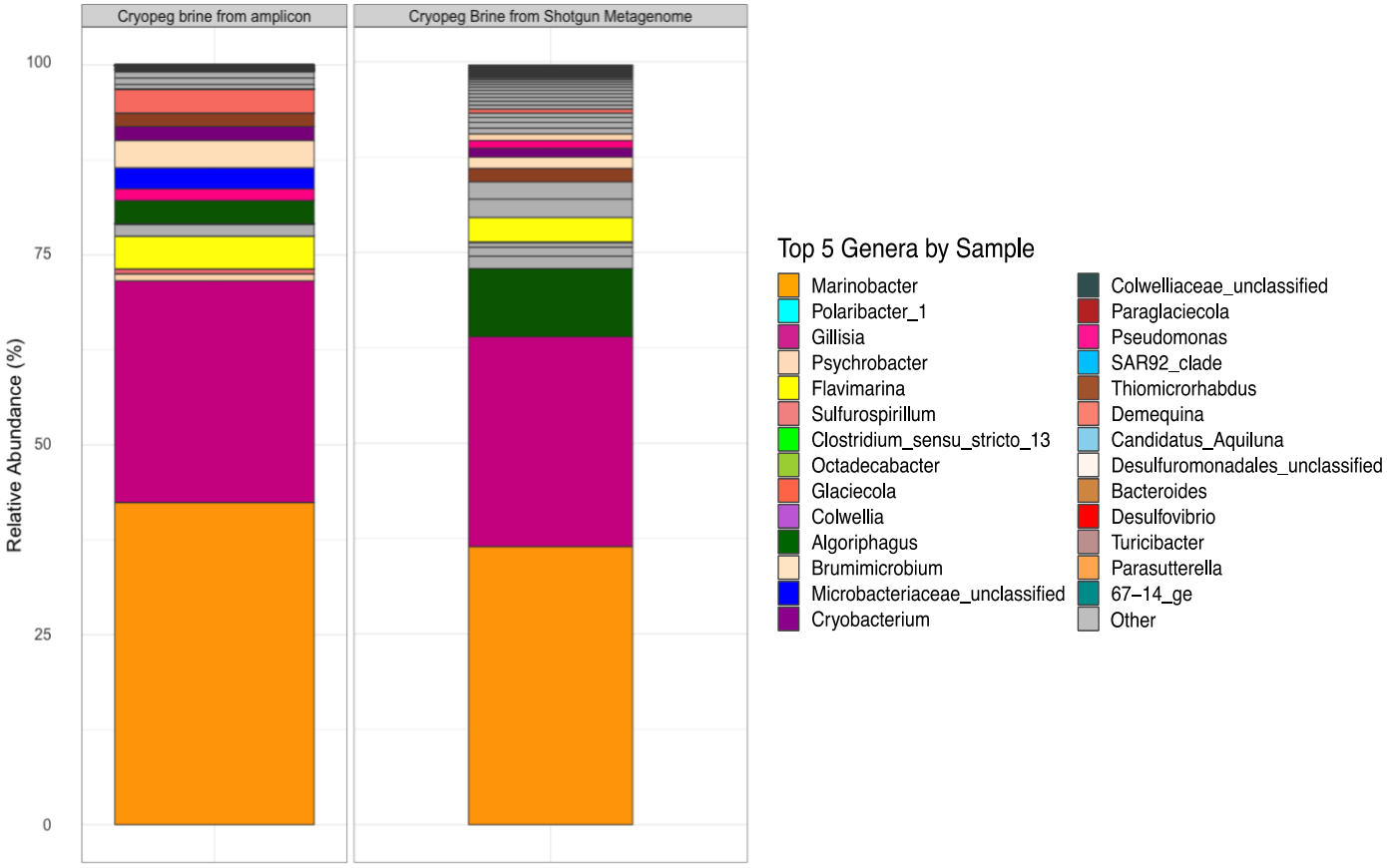

Figure S2.2. Relative abundance comparison of grouped genera in the cryopeg brine sample CBIW_17 from amplicon sequencing and from a shotgun metagenome.

The taxa are color-coded as in Figure 2.5.

APPENDIX B: CHAPTER 4 SUPPLEMENTARY INFORMATION

Supplemental Data Tables

The supplemental data tables for this chapter are too large to be formatted properly for this document. They can be found in an online repository using the following link: https://doi.org/10.5281/zenodo.5705982.

Supplementary Figure

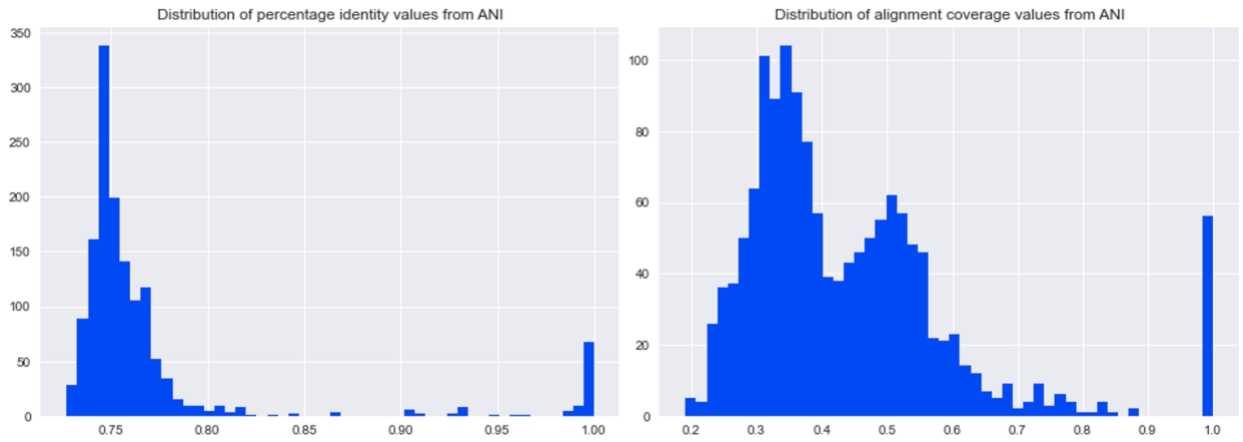

Figure S4.1. Histogram of average nucleotide identity (ANI) and alignment coverage values calculated pairwise for each genome in the *Marinobacter* pangenome.
The x-axis displays the percentage values of identity and alignment coverage, respectively. The y-axis displays the count of observations.

APPENDIX C: GROWTH CHARACTERISTICS OF *MARINOBACTER* STRAINS ISOLATED FROM CRYOPEG BRINES ACROSS TEMPERATURE AND SALINITY RANGES

Summary

Chapter 4 includes a discussion of the genomes of four strains of *Marinobacter* that were isolated from cryopeg brines. Details on the isolation of these strains are included in the Methods section of that chapter. Briefly, isolations were conducted using the dilution-to-extinction method by serially diluting cryopeg brine (from a recharged sample of borehole CBIW_18) into variations of a defined medium that included single organic sources. From the lowest dilution that yielded growth at 2˚C, we inoculated cultures on half-organic-strength 2216 Marine Broth (amended with artificial seawater) agar plates and picked colonies which were restreaked three times each for purification. The final isolates were identified as belonging to the genus *Marinobacter* using Sanger sequencing of the 16S rRNA gene, and taxonomic assignment was confirmed following whole genome sequencing by alignment to the Genome Taxonomy Database. The four strains are designated M1C, M2C, M3C, and M4C.

The *in situ* temperature and salinity conditions of the cryopeg brines that yielded these cultures were approximately –6˚C and 120 ppt. We aimed to quantify the growth rates of each strain across a range of temperatures and salinities, with a specific goal of determining their ability to grow under the *in situ* conditions. The following tables include optical density measurements and microscopic counts of the cryopeg *Marinobacter* either growing across a range of salinities (17.5 to 200 ppt; half strength 2216 Marine Broth amended with NaCl for salinities above 32 ppt) at 13˚C for all strains or a range of temperatures (–15 to 25˚C) in media

at 32 ppt (–1 to 25˚C) and 120 ppt (–15 to 25˚C) for strains M1C and M3C. For all conditions and time points in the salinity range experiment, measurements of optical density at 600 nm were taken without removing sample volume. A subset of time points were sampled for counts using epifluorescence microscopy and the DAPI stain at 1562.5X (as in Chapter 2). For all conditions and time points in the temperature experiment, samples were taken for counts as described for the salinity experiment.

The data presented in this appendix took many months to generate and have only been completed in the weeks leading up to submission of this dissertation. Calculations of growth rates across these temperature and salinity ranges will be included in a future manuscript, but selected growth curves are presented in Figure C.1. These data indicate that *Marinobacter* strains M1C, M2C, M3C, and M4C each demonstrate growth at salinities of 17 to 120 ppt and temperatures of –10 to < 20˚C for M1C and M3C (–15˚C growth observed for M3C; Table C.6). Although no growth was detected at salinities > 120 ppt, each of these strains can tolerate salinities up to 200 ppt, as evidenced by their growth at lower salinities (≤ 120 ppt) following transfer (subculturing) from incubations for 3 months at higher salinities (data not shown). Strains were capable of growth under the *in situ* conditions of cryopeg brines (–6˚C and 120 ppt). These data add to the information generated so far that indicate that these bacteria are well-adapted to the subzero brine environment. They also characterize these bacteria as polyextremophilic with regards to temperature and salinity and are here defined as hyperhalopsychrophiles or eutectophiles (as defined by Deming, 2002[1]). Given their ability to grow under conditions analogous to the brines of the Martian subsurface or the interior of icy moons, these novel strains add to the known diversity of astrobiologically-relevant extremophiles and expand the field of knowledge on the prevalence of life in extreme subzero brines.

[1] Deming, J. W. (2002). Psychrophiles and polar regions. *Current Opinion in Microbiology*. https://doi.org/10.1016/S1369-5274(02)00329-6

Table C.1. Optical density measurements across a range of salinities for M1C, M2C, M3C, and M4C at 13˚C

Time[a] (days)	Salinity (ppt)																							
	17				32				55				75				95				120			
	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C
0	0	0	0	0	0.000	-[b]	0.000	-	-	-	-	-	-	0.010	0.010	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0	0	0	0	0	0.000	-	0.000	-	-	-	-	-	-	0.010	0.010	-	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.08	0.07	0.055	0.045	0.070	0.085	0.040	0.050	0.045	0.060	0.030	0.040	0.025	0.030	0.020	0.030	0.015	0.000	0.010	0.000	0.000	0.007	0.020	0.010
4	0.08	0.075	0.055	0.06	0.075	0.090	0.045	0.050	0.050	0.080	0.035	0.050	0.030	0.015	0.015	0.025	0.000	0.002	0.010	0.000	0.005	0.000	0.015	0.010
5	0.115	0.09	0.1	0.09	0.120	0.100	0.075	0.090	0.115	0.090	0.070	0.120	0.090	0.065	0.035	0.055	0.050	0.017	0.015	0.013	0.020	0.000	0.005	0.015
5	0.12	0.09	0.1	0.07	0.115	0.100	0.100	0.085	0.120	0.100	0.120	0.080	0.090	0.080	0.040	0.115	0.040	0.010	0.010	0.020	-	0.000	0.005	0.010
6	0.135	0.145	0.15	0.125	0.100	0.140	0.120	-	0.110	0.120	0.085	0.095	0.067	0.100	0.080	0.090	0.055	0.000	0.021	0.035	0.005	0.005	0.013	0.015
6	0.14	0.115	0.13	0.11	0.090	0.135	0.125	0.120	0.112	0.135	0.110	0.130	0.075	0.090	0.060	0.085	0.050	0.04	0.040	0.055	0.005	0.005	0.015	0.005
8	0.115	0.15	0.17	0.145	0.100	0.200	0.155	0.140	0.145	0.125	0.125	0.175	0.130	0.090	0.085	0.135	0.087	0.04	0.055	0.070	0.015	0.015	0.010	0.010
8	0.13	0.145	0.205	0.13	0.090	0.215	0.160	0.145	0.090	0.145	0.150	0.175	0.150	0.090	0.090	0.125	0.085	0.03	0.045	0.070	0.015	0.020	0.010	0.015
10	0.24	0.205	0.22	0.19	0.195	0.215	0.215	0.175	0.170	0.195	0.160	0.160	0.130	0.190	0.140	0.200	0.135	0.100	0.100	0.110	0.020	0.020	0.015	0.010
10	0.215	0.19	0.215	0.19	0.180	0.205	0.220	0.185	0.225	0.195	0.170	0.185	0.150	0.200	0.145	0.170	0.140	0.060	0.100	0.120	0.020	0.015	0.010	0.005
12	0.23	0.265	0.315	0.25	0.220	0.260	0.275	0.230	0.215	0.210	0.200	0.200	0.200	0.205	0.170	0.230	0.120	0.115	0.135	0.150	-	0.050	0.035	0.100
12	0.23	0.275	0.275	0.235	0.210	0.245	0.290	0.240	0.245	0.230	0.215	0.225	0.185	0.185	0.180	0.205	0.125	0.120	0.125	0.135	0.060	0.055	0.025	0.075
13	0.25	0.26	0.285	0.27	0.250	0.320	0.325	0.290	0.225	0.220	0.245	0.220	0.175	0.170	0.205	0.250	0.120	0.130	0.200	0.155	0.070	0.050	0.075	0.070
13	0.28	0.25	0.3	0.25	0.250	0.300	0.335	0.260	0.215	0.250	0.210	0.235	0.205	0.200	0.205	0.220	0.135	0.130	0.165	0.165	0.070	0.055	0.045	0.090
15	0.28	0.3	0.31	0.285	0.280	0.310	0.370	0.285	0.240	0.245	0.280	0.295	0.180	0.195	0.230	0.215	0.140	0.145	0.175	0.175	0.070	0.055	0.080	0.100
15	0.28	0.285	0.32	0.28	0.245	0.340	0.400	0.300	0.240	0.265	0.300	0.270	0.210	0.210	0.215	0.245	0.150	0.145	0.180	0.175	0.080	0.075	0.070	0.085
19	0.325	0.33	0.345	0.33	0.315	0.365	0.420	0.340	0.315	0.290	0.335	0.300	0.210	0.210	0.275	0.260	0.200	0.150	0.210	0.200	0.110	0.065	0.125	0.140
19	0.37	0.32	0.365	0.33	0.410	0.350	0.420	0.340	0.280	0.330	0.335	0.350	0.225	0.225	0.265	0.260	0.185	0.170	0.210	0.205	0.110	0.100	0.115	0.110
22	0.34	0.34	0.355	0.38	0.340	0.390	0.435	0.365	0.355	0.320	0.350	0.305	0.230	0.230	0.300	0.290	0.175	0.175	0.215	0.205	0.110	0.105	0.130	0.105
22	0.335	0.325	0.385	0.34	0.400	0.365	0.420	0.375	0.325	0.350	0.340	0.315	0.245	0.260	0.285	0.295	0.225	0.195	0.220	0.215	0.120	0.140	0.135	0.140
26	-	-	-	-	-	-	-	-	-	-	-	-	0.250	0.265	0.335	0.330	0.185	0.220	0.245	0.230	0.125	0.110	0.200	0.165
26	-	-	-	-	-	-	-	-	-	-	-	-	0.310	0.270	0.380	0.310	0.215	0.210	0.240	0.250	0.135	0.150	0.155	0.180
29	0.39	0.39	0.45	0.41	0.380	0.420	0.490	0.410	0.420	0.350	0.385	0.360	0.300	0.290	0.380	0.370	0.260	0.245	0.260	0.240	0.150	0.180	0.170	0.175
29	0.44	0.38	0.48	0.45	0.490	0.410	0.470	0.450	0.380	0.400	0.380	0.370	0.340	0.310	0.470	0.370	0.240	0.230	0.270	0.310	0.175	0.165	0.180	0.180
34	-	-	-	-	-	-	-	-	-	-	-	-	0.310	0.335	0.420	0.410	0.225	0.230	0.220	0.230	0.150	0.190	0.175	0.195
34	-	-	-	-	-	-	-	-	-	-	-	-	0.320	0.330	0.520	0.325	0.280	0.280	0.255	0.310	0.150	0.215	0.195	0.240
41	-	-	-	-	-	-	-	-	-	-	-	-	0.325	0.335	0.440	0.460	0.250	0.250	0.290	0.280	0.170	0.175	0.235	0.265
41	-	-	-	-	-	-	-	-	-	-	-	-	0.345	0.360	0.600	0.365	0.255	0.255	0.335	0.340	0.165	0.195	0.205	0.260

[a]Separate duplicate tubes for each time point.

[b]Hyphen indicates measurement not available.

Table C.2. Microscopic counts of bacterial abundance (cells mL^{-1}) across a range of salinities for M1C, M2C, M3C, and M4C at 13˚C

Time[a] (days)	Salinity (ppt)											
	17				32				55			
	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C
0	2.88×10^6	1.60×10^5	5.00×10^6	3.13×10^5	1.39×10^5	1.65×10^5	1.83×10^5	1.91×10^5	1.51×10^5	1.62×10^5	1.58×10^5	1.66×10^5
5	5.51×10^7	2.76×10^7	4.12×10^7	4.04×10^7	4.76×10^7	3.23×10^7	5.16×10^7	5.18×10^7	2.39×10^7	1.91×10^7	3.49×10^7	2.44×10^7
5	2.23×10^7	4.59×10^7	3.61×10^7	3.34×10^7	3.52×10^7	3.86×10^7	4.78×10^7	2.70×10^7	3.09×10^7	1.06×10^7	2.56×10^7	3.29×10^7
13	-[b]	-	-	-	-	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-	-	-	-	-	-
19	1.14×10^8	2.67×10^8	-	4.75×10^7	3.20×10^8	3.03×10^8	-	3.61×10^8	1.05×10^8	2.52×10^8	2.17×10^8	2.06×10^8
19	-	-	-	-	-	-	-	-	-	-	-	-
21	-	-	-	-	-	-	-	-	-	-	-	-
21	-	-	-	-	-	-	-	-	-	-	-	-
29	-	-	-	-	-	-	-	-	-	-	-	-
29	-	-	-	-	-	-	-	-	-	-	-	-

[a]Separate duplicate tubes for each time point.

[b]Hyphen indicates measurement not available.

Table C.2 (extended)

Time[a] (days)	Salinity (ppt)											
	75				95				120			
	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C
0	1.31×10^5	1.63×10^5	1.54×10^5	1.86×10^5	1.49×10^5	-[b]	6.90×10^4	9.50×10^4	1.40×10^5	1.71×10^5	1.77×10^5	5.06×10^5
5	1.74×10^7	1.70×10^7	1.56×10^7	2.60×10^7	5.03×10^6	6.68×10^6	4.27×10^6	1.32×10^7	6.07×10^5	5.05×10^5	1.03×10^7	8.21×10^5
5	1.92×10^7	2.71×10^7	2.81×10^7	2.43×10^7	4.99×10^6	7.55×10^6	4.83×10^6	7.80×10^6	5.52×10^5	5.02×10^5	3.82×10^7	3.81×10^5
13	-	-	-	-	3.66×10^7	4.83×10^7	3.80×10^7	3.41×10^7	1.84×10^7	1.59×10^7	1.73×10^7	1.88×10^7
13	-	-	-	-	-	3.09×10^7	-	3.23×10^7	2.14×10^7	2.50×10^7		1.71×10^7
19	5.44×10^7	-	-	-	5.44×10^7	4.18×10^7	3.73×10^7	3.65×10^7	2.98×10^7	2.03×10^7	1.96×10^7	2.59×10^7
19	-	-	-	-	3.57×10^7	2.56×10^7	-	3.63×10^7	-	2.68×10^7	-	-
21	-	-	-	-	-	-	-	-	-	-	-	-
21	-	-	-	-	-	-	-	-	-	-	-	-
29	1.47×10^8	1.59×10^8	1.87×10^8	1.23×10^8	2.85×10^7	4.49×10^7	5.49×10^7	3.92×10^7	2.17×10^7	-	-	2.56×10^7
29	1.14×10^8	1.37×10^8	1.67×10^8	1.11×10^8	4.14×10^7	-	-	-	3.38×10^7	-	-	-

[a]Separate duplicate tubes for each time point.

[b]Hyphen indicates measurement not available.

Table C.2 (extended)

Time[a] (days)	Salinity (ppt)											
	140				170				200			
	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C	M1C	M2C	M3C	M4C
0	1.57×10^5	1.38×10^5	-	-	1.52×10^5	1.70×10^5	-	1.75×10^5	1.29×10^5	1.49×10^5	8.74×10^4	1.05×10^5
5	1.51×10^5	1.79×10^5	1.93×10^5	2.15×10^5	1.27×10^5	1.62×10^5	1.46×10^5	1.85×10^5	1.48×10^5	1.70×10^5	1.84×10^5	1.38×10^5
5	1.66×10^5	1.73×10^5	1.48×10^5	2.12×10^5	1.56×10^5	1.59×10^5	1.63×10^5	1.92×10^5	1.50×10^5	1.86×10^5	1.69×10^5	2.29×10^8
13	-	-	-	-	-	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-	-	-	-	-	-
19	-	-	-	-	-	-	-	-	-	-	-	-
19	-	-	-	-	-	-	-	-	-	-	-	-
21	4.12×10^5	2.16×10^5	1.80×10^5	3.95×10^5	1.29×10^5	1.53×10^5	2.03×10^5	2.30×10^5	1.15×10^5	1.41×10^5	1.52×10^5	1.38×10^5
21	3.45×10^5	2.58×10^5	1.98×10^5	3.67×10^5	1.58×10^5	1.30×10^5	1.43×10^5	1.98×10^5	7.77×10^4	1.36×10^5	8.20×10^4	1.65×10^5
29	-	-	-	-	-	-	-	-	-	-	-	-
29	-	-	-	-	-	-	-	-	-	-	-	-

[a]Separate duplicate tubes for each time point, except Day 0.

[b]Hyphen indicates measurement not available.

Table C.3. Microscopic counts of bacterial abundance (cells mL^{-1}) across a range of temperatures for M1C at 32 ppt

Time (days)[a]	Temperature (°C)				
	–1	**4**	**8**	**13**	**21**
0	1.20×10^7	1.20×10^7	1.20×10^7	1.20×10^7	1.20×10^7
0	1.20×10^7	1.20×10^7	1.20×10^7	1.20×10^7	1.20×10^7
1	-[b]	-	-	2.78×10^7	1.93×10^7
1	-	-	-	1.93×10^7	1.94×10^7
3	-	3.16×10^7	2.44×10^7	3.96×10^7	2.58×10^7
3	-	3.64×10^7	2.92×10^7	4.43×10^7	2.66×10^7
6	-	8.03×10^7	5.39×10^7	4.12×10^7	3.53×10^7
6	-	5.77×10^7	6.81×10^7	7.97×10^7	3.44×10^7
7	4.05×10^7	-	-	-	-
7	3.54×10^7	-	-	-	-
8	-	1.65×10^8	1.37×10^8	1.78×10^8	-
8	-	1.74×10^8	1.11×10^8	1.62×10^8	-
12	-	2.06×10^8	1.68×10^8	2.05×10^8	-
12	-	2.64×10^8	1.65×10^8	2.45×10^8	-
15	5.36×10^7	2.41×10^8	1.84×10^8	3.11×10^8	-
15	5.63×10^7	2.66×10^8	3.01×10^8	3.23×10^8	-
21	-	1.30×10^8	3.60×10^8	3.93×10^8	-
21	-	3.84×10^8	3.39×10^8	3.14×10^8	-
28	1.11×10^8	6.13×10^8	4.16×10^8	4.56×10^8	-
28	2.73×10^8	4.45×10^8	4.63×10^8	4.00×10^8	-
42	2.99×10^8	5.65×10^8	3.89×10^8	4.24×10^8	-
42	4.57×10^8	5.82×10^8	4.45×10^8	4.89×10^8	-
52	3.93×10^8	-	-	-	-
52	5.39×10^8	-	-	-	-
66	6.50×10^8	-	-	-	-
66	5.35×10^8	-	-	-	-

[a]Separate duplicate tubes for each time point. [b]Hyphen indicates measurement not available.

Table C.4. Microscopic counts of bacterial abundance (cells mL^{-1}) across a range of temperatures for M3C at 32 ppt

Time (days)[a]	Temperature (˚C)				
	–1	**4**	**8**	**13**	**21**
0	1.10×10^7	1.10×10^7	1.10×10^7	1.10×10^7	1.10×10^7
0	1.10×10^7	1.10×10^7	1.10×10^7	1.10×10^7	1.10×10^7
1	-[b]	-	-	2.05×10^7	1.85×10^7
1	-	-	-	1.79×10^7	1.70×10^7
3	-	3.80×10^7	3.08×10^7	3.64×10^7	2.10×10^7
3	-	3.57×10^7	3.74×10^7	3.56×10^7	2.87×10^7
6	-	5.30×10^7	4.11×10^7	5.53×10^7	3.18×10^7
6	-	6.75×10^7	4.79×10^7	5.81×10^7	2.60×10^7
7	3.04×10^7	-	-	-	-
7	3.13×10^7	-	-	-	-
8	-	1.78×10^8	1.44×10^8	1.70×10^8	-
8	-	1.68×10^8	1.15×10^8	1.81×10^8	-
12	-	2.43×10^8	1.54×10^8	2.36×10^8	-
12	-	2.06×10^8	2.14×10^8	1.84×10^8	-
15	4.90×10^7	2.30×10^8	2.11×10^8	2.84×10^8	-
15	7.21×10^7	2.83×10^8	2.30×10^8	3.20×10^8	-
19	-	-	-	-	-
19	-	-	-	-	-
21	-	2.46×10^8	2.31×10^8	3.17×10^8	-
21	-	3.57×10^8	3.29×10^8	3.47×10^8	-
28	1.29×10^8	5.18×10^8	3.77×10^8	3.80×10^8	-
28	1.71×10^8	4.24×10^8	4.37×10^8	4.22×10^8	-
42	2.77×10^8	6.62×10^8	6.57×10^8	4.16×10^8	-
42	3.72×10^8	6.69×10^8	6.22×10^8	4.83×10^8	-
52	3.59×10^8	-	-	-	-
52	4.99×10^8	-	-	-	-
66	5.09×10^8	-	-	-	-
66	4.84×10^8	-	-	-	-

[a]Separate duplicate tubes for each time point.; [b]Hyphen indicates measurement not available.

Table C.5. Microscopic counts of bacterial abundance (cells mL^{-1}) across a range of temperatures for M1C at 120 ppt

	Temperature (˚C)							
Time (days)[a]	-15	-10	-5	-1	4	8	13	21
0	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7
0	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7	1.63×10^7
1	-[b]	-	-	-	-	-	9.93×10^6	1.10×10^7
1	-	-	-	-	-	-	1.15×10^7	1.26×10^7
3	-	-	-	-	7.38×10^6	1.00×10^7	9.11×10^6	9.80×10^6
3	-	-	-	-	1.11×10^7	1.07×10^7	1.32×10^7	1.14×10^7
6	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-
7	-	-	-	1.46×10^7	-	-	-	-
7	-	-	-	9.93×10^6	-	-	-	-
8	-	-	1.42×10^7	-	2.48×10^7	2.21×10^7	1.65×10^7	1.14×10^7
8	-	-	1.18×10^7	-	2.50×10^7	1.81×10^7	2.06×10^7	1.22×10^7
12	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-
15	-	-	-	2.64×10^7	2.79×10^7	3.10×10^7	2.87×10^7	7.36×10^6
15	-	-	-	3.10×10^7	2.96×10^7	2.85×10^7	3.24×10^7	5.89×10^6
19	1.20×10^7	1.54×10^7	1.65×10^7	-	-	-	-	-
19	1.18×10^7	1.31×10^7	1.52×10^7	-	-	-	-	-
21	-	-	-	-	5.51×10^7	4.87×10^7	2.08×10^7	5.79×10^6
21	-	-	-	-	6.63×10^7	3.69×10^7	1.58×10^7	6.80×10^6
28	-	-	2.54×10^7	5.20×10^7	1.42×10^8	5.96×10^7	2.83×10^7	9.03×10^6
28	-	-	3.47×10^7	5.37×10^7	1.48×10^8	9.18×10^7	3.11×10^7	8.85×10^6
42	1.19×10^7	5.56×10^7	6.81×10^7	1.83×10^8	1.90×10^8	1.10×10^8	3.25×10^7	1.09×10^7
42	1.41×10^7	6.36×10^7	6.95×10^7	1.99×10^8	2.07×10^8	1.25×10^8	3.11×10^7	-
52	1.23×10^7	1.84×10^8	2.08×10^8	2.66×10^8	2.00×10^8	1.29×10^8	-	-
52	1.27×10^7	2.39×10^8	2.04×10^8	2.67×10^8	2.14×10^8	1.34×10^8	-	-
66	1.62×10^7	2.43×10^8	2.90×10^8	2.84×10^8	-	-	-	-
66	1.44×10^7	2.84×10^8	2.46×10^8	2.74×10^8	-	-	-	-

[a]Separate duplicate tubes for each time point.; [b]Hyphen indicates measurement not available.

Table C.6. Microscopic counts of bacterial abundance (cells mL^{-1}) across a range of temperatures for M3C at 120 ppt

Time (days)[a]	Temperature (˚C)							
	–15	–10	–5	–1	4	8	13	21
0	1.35×1^{07}	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7
0	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7	1.35×10^7
1	-[b]	-	-	-	-	-	1.05×10^7	1.56×10^7
1	-	-	-	-	-	-	9.54×10^6	1.21×10^7
3	-	-	-	-	1.44×10^7	1.47×10^7	1.40×10^7	1.45×10^7
3	-	-	-	-	1.46×10^7	1.15×10^7	1.47×10^7	1.37×10^7
6	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-
7	-	-	-	1.38×10^7	-	-	-	-
7	-	-	-	1.32×10^7	-	-	-	-
8	-	-	1.30×10^7	-	1.92×10^7	2.45×10^7	2.10×10^7	1.13×10^7
8	-	-	1.18×10^7	-	2.48×10^7	2.64×10^7	2.18×10^7	1.30×10^7
12	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-
15	-	-	-	2.77×10^7	3.89×10^7	2.64×10^7	2.02×10^7	7.98×10^6
15	-	-	-	3.17×10^7	3.71×10^7	2.88×10^7	2.45×10^7	6.52×10^6
19	1.40×10^7	1.97×10^7	1.54×10^7	-	-	-	-	-
19	1.24×10^7	1.47×10^7	1.55×10^7	-	-	-	-	-
21	-	-	-	-	5.62×10^7	4.90×10^7	3.45×10^7	5.46×10^6
21	-	-	-	-	4.63×10^7	3.70×10^7	5.46×10^6	-
28	-	-	2.60×10^7	4.29×10^7	1.82×10^8	5.17×10^7	3.41×10^7	-
28	-	-	2.73×10^7	5.77×10^7	6.08×10^7	4.30×10^7	3.16×10^7	-
42	1.34×10^7	7.77×10^7	6.29×10^7	1.51×10^8	9.09×10^7	1.34×10^8	3.69×10^7	1.22×10^7
42	1.84×10^7	5.44×10^7	6.10×10^7	1.62×10^8	9.66×10^7	9.84×10^7	3.90×10^7	1.26×10^7
52	2.04×10^7	2.33×10^8	2.40×10^8	2.28×10^8	2.14×10^8	1.35×10^8	-	-
52	2.15×10^7	2.10×10^8	2.38×10^8	2.33×10^8	1.99×10^8	1.40×10^8	-	-
66	2.09×10^7	3.56×10^8	2.54×10^8	2.67×10^8	-	-	-	-
66	2.37×10^7	2.79×10^8	3.03×10^8	2.63×10^8	-	-	-	-

[a]Separate duplicate tubes for each time point.; [b]Hyphen indicates measurement not available

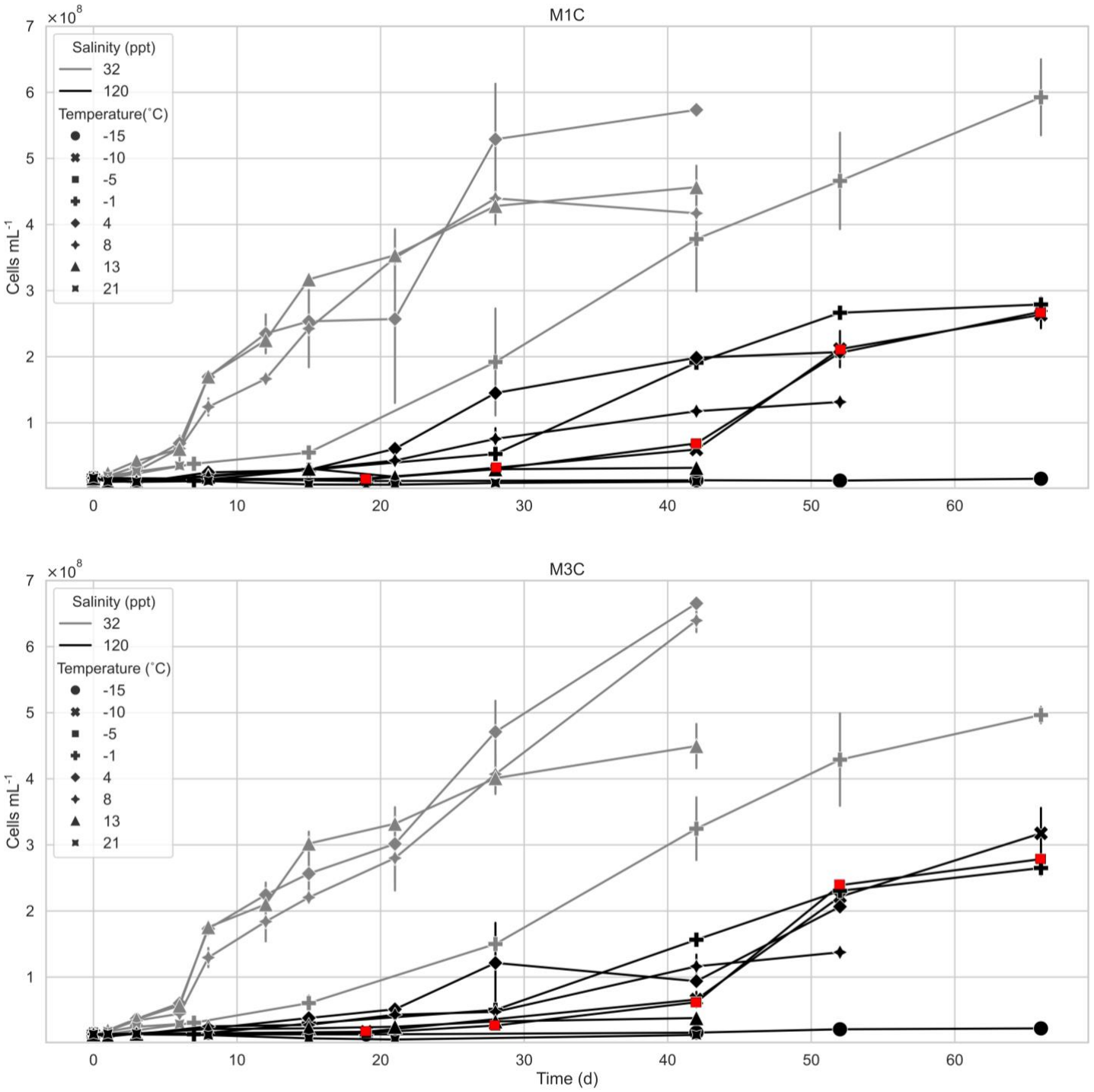

Figure C.1. Growth curves of *Marinobacter* strains M1C (top) and M3C (bottom) across a range of temperatures at 32 and 120 ppt.

Error bars indicate standard deviation (n = 2). Points without error bars are either not replicated or the error bars are too small to be seen behind the point. Red squares highlight growth near *in situ* conditions of the source cryopegs (T = –5˚C and S = 120 ppt). Note: the 32 ppt, 21˚C counts are truncated after 6 days due to contamination appearing thereafter. Data for this figure can be found in Tables C.3–C.6.

CURRICULUM VITAE

Zachary S. Cooper

University of Washington, School of Oceanography, Astrobiology Program
Marine Sciences Building, 1503 NE Boat St., Box 357940, Seattle, WA 98195

Education

2021	**Ph.D.**, Oceanography and Astrobiology University of Washington, Seattle, WA *Dissertation: Microbial evolution and ecology in subzero hypersaline environments*
2020	**Ph.D. Candidacy**, Oceanography and Astrobiology University of Washington, Seattle, WA *Dissertation proposal: Microbial evolution and ecology in subzero hypersaline environments*
2018	**M.S.**, Oceanography University of Washington, Seattle, WA *Thesis: Microbial community diversity in sea ice and cryopeg brines*
2016	**B.S.**, Biology, *Summa Cum Laude* University of Georgia, Athens, GA
2014	**A.S.**, Biology Middle Georgia State University, Cochran, GA

Employment and Research

2016-2021	Graduate Research Assistant, School of Oceanography, U. Washington Supervisor: Dr. Jody W. Deming Studying extremophile microbiology focusing on genomic evolution and ecology of psychrophilic microorganisms in subzero, hypersaline, astrobiological analog environments by conducting amplicon, genomic, and metagenomic bioinformatic analyses
2021	Predoctoral Research Associate II, Astrobiology Program, U. Washington Supervisor: Dr. Scott M. Perl, NASA Jet Propulsion Laboratory, California Institute of Technology Quarter-long research rotation. Used Raman spectroscopy and digital microscopy to measure biosignatures produced by psychrophilic, halotolerant bacteria in desiccated and frozen brines.

2015-2016	Undergraduate Research Assistant, U. Georgia, Dept. of Biochemistry and Molecular Biology Supervisor: Dr. Michael W.W. Adams Genetically engineered hyperthermophilic archaea for biofuel production, analyzed mutant strain productivity, and began developing a genetic marker tool for use in high-temperature, anaerobic systems.
2015-2016	Undergraduate Research Assistant, U. Georgia, Center for Applied Genetic Technology Supervisor: Dr. Wayne Parrott Subcultured and maintained genetically modified rice plants from embryonic to mature stages.

Field Research

Oct. – Dec. 2019	*R/V IB Nathaniel B. Palmer*, Southern Ocean, Western Antarctic Peninsula, 45 days. Collected sea ice algae samples for radioisotope primary production assays, metabolomics, and genomics. Main responsibilities included sea ice coring, laboratory maintenance, and ^{14}C-primary production assays. PI, Dr. Jodi Young.
Aug. 2018	*R/V Norseman II*, Bering Strait, Pacific-Arctic Ocean, 12 days. Assisted in mooring construction, deployments, and recoveries as well as operating CTD transects for physical measurements of water flux through the Bering Strait to the Arctic Ocean. PI, Dr. Rebecca Woodgate.
May 2018	Utqiaġvik, AK, 15 days. Expanded on May 2017 subzero brine collection from sea ice and cryopegs for microbiological analyses. Assisted in permafrost drilling to understand physical distribution of cryopegs. PI, Dr. Jody Deming.
Aug. 2017	*R/V Revelle*, N. Pacific, Axial Seamount, Pacific Ocean, 10 days. VISIONS'17. Collected samples from the deep sea for microbiology and logged data for the ROV *JASON*. PI, Dr. Deb Kelley.
May 2017	Utqiaġvik, AK, 10 days. Collected brine and seawater samples from coastal sea ice and permafrost cryopeg brines for microbiological analyses. Assisted in permafrost drilling to locate new cryopeg sites. PI, Dr. Jody Deming.

Teaching Experience

Spring 2019	Teaching Assistant: Arctic Change Supervisor: Dr. Rebecca Woodgate Lead weekly review sessions for the course of 45 students. Graded weekly homework assignments including problem sets and short essays. Proctored and graded exams.
Spring 2017	Teaching Assistant: Special Topics in Biological Oceanography Supervisor: Dr. John Baross. Lectured on cold-active bacteriophage and field work in the Arctic. Graded and organized written assignments and oral presentations for 22 undergraduate students.
2013-2014	Undergraduate Tutor, Middle Georgia State University Tutored students in General and Organic Chemistry, Pre-calculus and Calculus, and English Composition

Workshops

April 2019	University of Washington, Seattle, WA, USA. Astrobiology Program workshop on genomics including DNA extraction, library preparation, and sequencing of sediment samples from Death Valley, CA.
Oct. 2018	Death Valley, CA, USA. Washington Astrobiology Program workshop on paleontology and life in extreme environments. Collected samples for microbiology from desert soils and hypersaline streams and observed stromatolites and trilobites.
March 2018	Jet Propulsion Laboratory, Pasadena, CA. Washington Astrobiology Program workshop on geochemical astrobiology, space exploration, and exoplanet sensing satellite technology.
Nov. 2017	Argonne National Labs, IL. PATRIC/RAST Workshop. Introduction to comparative genomic analyses using PATRIC/RAST.
Oct. 2017	Ohio State University, Columbus, OH. Viromics workshop hosted by the Sullivan Lab. Introduction to viromics techniques using CyVerse and IMG/VR among other cloud-based HPC resources.
Sep. 2016	Mount Rainier, WA. U. Washington Astrobiology Program workshop on remote environmental sample collection and biosignatures. Developed protocol for biological field sampling of algae in Muir Snowfields in a team of four graduate students and presented protocol and results.

Outreach and Volunteering

Winter 2019	Polar Science Communication Fellow, Pacific Science Center, Seattle, WA. Attended workshops focused on learning science communication skills and developed a planetarium show with imagery from the Arctic to illustrate field science and environmental microbiology for public presentations to audiences of all ages.
March 2017 March 2018 March 2019 March 2020	Pacific Science Center, Seattle, WA. Polar Science Weekend. Exhibit on Life Below Freezing. Presented hands on educational activity demonstrating ecological niches in sea ice for museum visitors of all ages over two days.

Publications

1. **Cooper, Z.S.**, Carpenter, S.D., Shoemaker, A.M.D., and Deming, J.W. Characterization of *Marinobacter* sp. nov. isolated from cryopeg brines. (in prep, based on Appendix C).

2. **Cooper, Z.S.**, Anderson, R.E., Zhong, Z.P., and Deming, J.W. Mobilome-driven evolution in bacterial populations from subzero brines. (in prep, based on Chapter 3).

3. **Cooper, Z.S.**, Rapp, J.Z., Shoemaker, A.M.D., Anderson, R., Zhong, Z., Sullivan, M.B., Deming, J.W. Evolution and ecology of *Marinobacter* sp. nov. isolated from cryopeg brines as revealed by pangenomics. (in prep, based on Chapter 4).

4. Styczinski, M.J., **Cooper, Z.S.**, Glaser, D.M., Lehmer, O., Mierzejewski, V., Tarnas, J. Chapter 7: Assessing Habitability Beyond Earth. in NASA Astrobiology Primer 3.0. *Astrobiology*. (submitted).

5. Iwahana, G., **Cooper, Z.S.**, Carpenter, S.D., Deming, J.W., Eicken, H. Intra-ice and intra-sediment cryopeg brine occurrence in permafrost near Utqiaġvik (Barrow). *Permafrost and Periglacial Processes*. 2021; 1– 20. https://doi.org/10.1002/ppp.2101

6. Zhong, Z., Rapp, J.Z., Wainaina, J.M., Solonenko, N.E., Maughan, H., Carpenter, S.D., **Cooper, Z.S.**, Jang, H.B., Bolduc, B., Deming, J.W., Sullivan, M.B. 2020. Viral ecogenomics of Arctic cryopeg brine and sea ice. *mSystems*. 5(3), 1-17 https://doi.org/10.1128/msystems.00246-20

7. **Cooper, Z.S.**, Rapp, J.Z., Carpenter, S.D., Iwahana, G, Eicken, H, Deming, J.W. 2019. Distinctive microbial communities in subzero hypersaline brines from Arctic coastal sea ice and rarely sampled cryopegs. *FEMS Microbiology Ecology*, 95(12), 1–15. https://doi.org/10.1093/femsec/fiz166. *Editor's choice*

Presentations

1. **Cooper, Z.S.**, Anderson, R.E., Zhong, Z.P., and Deming, J.W. 2022. Mobilome-influenced evolution in bacterial populations from subzero brines. Oral presentation, Ocean Sciences Meeting, 27 February - 4 March 2022, virtual.

2. **Cooper, Z.S.** and Perl, S.M. 2021. Cryobrine habitability and Raman detections of biological salt-ice features. Oral Presentation, 17 June 2021, Jet Propulsion Laboratory Origins and Habitability Lab, Pasadena, CA, USA.

3. **Cooper, Z.S.**, Rapp, J.Z., Carpenter, S.D., and Deming, J.W. 2019. Microbial life in subzero hypersaline brines: relevance to the habitability of subsurface Mars and icy moons. Poster presentation, AbSciCon. 24-28 June 2019, Bellevue, WA, USA. *Student poster competition finalist*

4. Zhong, Z., Solonenko, N., Bolduc, B., Carpenter, S., **Cooper, Z.**, Jang, H.B., Deming, J., and Sullivan, M. 2019. Arctic viral ecogenomics reveals potential modes of cold and salt tolerance. Poster presentation, Ohio Supercomputer Center Statewide Users Group Meeting, 18 April 2019, Ohio State University, Columbus, Ohio, USA.

5. **Cooper, Z.S.**, Rapp, J.Z., Carpenter, S.D., Iwahana, G., Eicken, H., and Deming, J.W. 2019. Uncovering microbial community diversity in Alaskan cryopeg brines. Oral presentation, Polar and Alpine Microbiology Conference. 4-8 February 2019, University of Waikato, Hamilton, New Zealand.

6. Weaver, T., Berger, J., Biddle, J., Collins, R.E., **Cooper, Z.S.**, and Deming, J.W. 2018. Sonification of Deep Ocean Microbial Ecology (soniDOME). Poster presentation, POLAR 2018. 15-26 June 2018, Davos, Switzerland.

7. **Cooper, Z.S.**, Carpenter, S.D., Iwahana, G., Eicken, H., and Deming, J.W. 2018. Cryopeg and Sea Ice Brines: Comparison of Subzero, Hypersaline Aqueous Environments as Microbial Habitats. Poster Presentation, Ocean Sciences Meeting. 11-16 February, Portland, OR, USA.

8. **Cooper, Z.S.**, and Deming, J.W. 2017. On developing a genetic transformation system for the psychrophilic marine bacterium *Colwelllia psychrerythraea* 34H. Poster Presentation, Polar and Alpine Microbiology Conference. 8-12 September, Nuuk, Greenland.

Awards

2016-2017	Vetlesen Fellowship, U. Washington, School of Oceanography. Full coverage of tuition, benefits, and stipend for nine months.
2016	CURO Research Assistantship, U. Georgia. $1,000

Professional Associations

Association of Polar Early Career Scientists (APECS)

Association for the Sciences of Limnology and Oceanography (ASLO)

Academic and Recreational Graduate Oceanographers (ARGO), founding officer