12-2021

High Arctic Permafrost Microbial Characterizations: Siberian and Svalbard microbiology of ancient and active layer permafrost

Katie Sipes
University of Tennessee, Knoxville, ksipes@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Katie Sipes entitled "High Arctic Permafrost Microbial Characterizations: Siberian and Svalbard microbiology of ancient and active layer permafrost." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Microbiology.

Karen G. Lloyd, Major Professor

We have read this dissertation and recommend its acceptance:

Tatiana A. Vishnivetskaya, Jill A. Mikucki, Sarah L. Lebeis, Andrew D. Steen

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

High Arctic Permafrost Microbial Characterizations: Siberian and Svalbard microbiology of ancient and active layer permafrost

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Katie Sipes
December 2021

ACKNOWLEDGEMENTS

Thank you to my family and friends that have helped me succeed in graduate school and in life. Special thanks to my father for always giving advice, and my mother for calling to chat. My brother, Nicholas for telling me to learn how to code and always being my cheerleader and voice of reason.

My Knoxville friends and groups that have helped me remain mentally and physically sound; KyBra athletics (RIP), Dragonfly Circus Studio, Feral Feline Friends and Knoxville Vegans; SERF office people and Lloyd lab members. Special thanks to Raegan Paul, the best undergraduate researcher I had the privilege of mentoring and working alongside.

One million thanks to my primary advisor, Karen Lloyd, for allowing me to have intellectual freedoms over my project and investigate the rabbit holes but also redirecting. Additionally, thank you for letting me go to Svalbard with Joy Buongiorno in 2017, that trip was the catalyst that made me fall head over heels with the Arctic and environmental research. I am also so grateful for the funding you have provided to me to be able to focus on research and discovery. I have had an incredible graduate school experience thanks to you.

Thank you, Joy Buongiorno, for listening and being an incredible mentor. Also, for letting me help you in Svalbard in 2017 and for going into all the closed hot tubs with me. I would have never wanted to experience the midnight sun with anyone else.

Thank you, April Armes, for being my best friend and office buddy, you are the kindest person I know.

And of course, to all the incredible individuals from all over the world that have made me a better person and scientist. I am forever grateful for the lessons and knowledge I have gained, and I am eager to learn more.

ABSTRACT

Permafrost is soil that has remained frozen for at least two years. The active layer is the surface soil above the permafrost that experiences seasonal thaw and refreezing. The environmental characteristics of permafrost and active layer are different but are directly related to each other. As the climate continues to warm, the active layer will expand into the permafrost and the continuously frozen soil will be subjected to seasonal thawing. The organisms that inhabit both the active layer and the permafrost soil will respond differently to the climate based on where in the soil they are present and the soil characteristics. Moreover, the climate will either be inundated with a large amount of microbially mediated greenhouse gasses from increased metabolic activity or it will become a carbon sink with the increased viability for vegetation. Regardless, the microbial communities that are present in both the deep permafrost and the active layer will experience changes as the climate continues to warm. Those deep permafrost microbes are either adapted to their current location and can survive in the cold, nutrient-depleted soils, or they are in a hibernation like state; waiting for the warm climate until they can become metabolically active again. Opposingly, the microbes in the active layer could either be the main source of greenhouse gas emissions or they act as a carbon sink and use gasses in microbial metabolism. Either way, studying the microbial interactions in both deep permafrost and active layer are important when assessing how microbial interactions will play a role, and respond, in the changing climate.

This dissertation combines metagenomic data in the form of metagenome assembled genomes to assess how microbes at a sample location can interact with its environment. Genetic features in metagenome assembled genomes or from a metagenomic library are used to determine if the organisms that are sequenced are interacting with the geochemical characteristics of their habitat. In the chapters to follow, MAGs are analyzed from Siberian permafrost and Svalbard active layer. Annotated genes from Siberian MAGs show a suite of genes that demonstrate the ancient soils contain genetically adapted microbes. In Svalbard active layer, geochemical analyses are combined with culture independent methods to assess how the microbial community is active in their environment and how microbes will be able to metabolically respond to thaw. Lastly, analysis of MAGs from five Svalbard active layer cores shows how the phyla; *Acidobacteriota* and *Actinobacteriota* dominate different locations of the active layer stratigraphy, with the former having high abundance in the upper half of the frozen active layer and the latter dominating in the thawed active layer. These studies show how the microbial interactions with each other, and the environment will affect and be affected when the active layer has a long thawing season.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF SUPPLEMENTAL TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTAL FIGURES

CHAPTER 1 : INTRODUCTION

Permafrost and the active layer

Permafrost is sediment, soil, or ground that has been continuously frozen for at least two years. All permafrost soils cumulatively cover nearly 25% of Earth's terrestrial surface (1, 2). Permafrost in the Arctic latitudes alone is predicted to introduce 63 billion tons of organic carbon available for microbial degradation within the surface half meter alone (3). This thawing will be dramatic within the oldest, coldest, and deepest undisturbed permafrost in the world, which is located within the Kolyma-Indigirka Lowland Region in Northeastern Siberia. Permafrost here contain layers up to 3 million years old (4, 5). Microbial activity contributes biotically-generated greenhouse gases (for example: CO_2 or CH_4) to the atmosphere, making the metabolic activity of these microbial communities a critical factor when examining organic carbon reservoirs (2). However, the community dynamics, genetic capabilities, and potential for activity in microorganisms living in *current-state permafrost,* or as the permafrost is now, are not well understood. Specifically, microorganisms in permafrost may be unable to adapt to the inevitable permafrost thaw, thus new selective pressures not experienced in millions of years may impose a rapid shift in community structure.

The part of permafrost with the most immediate impact on the climate is active layer soil. The active layer is the upper portion of soil found on top of the permafrost that undergoes seasonal freezing and thawing. Although variable in depth by region, active layer soils have been increasing in thickness at rates 0.7 to 1.7 cm per year and are expected to accelerate to 2cm per year in the next decade (6). A large unknown in the active layer literature is how its microbial population can adapt to the changing environment. This is especially true in the active layer around Svalbard, Ny Ålesund, (79°N) where the population of microbes and how they function remains understudied. The climatic shifts around Svalbard have affected the active layer thickness to increase at least 1 cm a year (7) with some locations thawing up to 4 cm a year (8). As the Arctic Circle continues to warm at a faster rate than the rest of the globe, rapid fluctuations in environmental characteristics will be apparent in this location. Microbes that live in this soil must overcome a unique set of circumstances to mitigate such a highly variable and nutrient poor environment. The Arctic region is known to be nitrogen-limited and for the degradable organic carbon stocks to be 'locked away' in the frozen soil (9). Therefore, the newly thawed permafrost soils that incorporate into the active layer present newly available carbon compounds for microbial degradation. This expanding active layer can be investigated now to be used as a prediction of what will happen to the currently-frozen layers when permafrost soils turn to thawing active layer. Two different microbial populations are present in these soils: 1- there are microbes that may be adapted to the present conditions that will be outcompeted when the soil

thaws permanently and 2- there are microbes that are in statis that are less metabolically active now but when the soil thaws they will become more active.

Microbes from the active layer and permafrost

A wide range of approaches have been previously used to study microbial interactions with the permafrost and active layer. Despite the harsh environmental features, a wide variety of organisms have been found in all parts of permafrost (10). Some methods have shown microbial activity in the active layer. Stable isotope probing with ^{14}C-glucose added to *ex situ* frozen active layer produced two orders of magnitude higher CO_2 than autoclaved controls (11). More sources of microbial activity in the active layer and thawing permafrost come from metatranscriptomics. Although combined methods of metatranscriptomics and metagenomics have shown variable results between the genes that are present and genes that are active (12), it can show how the environment is selecting for microbial adaptations. Culturing and 16S rRNA gene analysis of permafrost and active layers have generated a catalog of commonly found organisms. These include most commonly found bacterial phyla (*Acidobacteria, Actinobacteria, Proteobacteria, Chloroflexi, Firmicutes, Bacteroidetes,* and many others) (10). Further physiological identification of bacterial and archaeal organisms is assisted by metagenomic methods.

To resolve activity on a cellular level once can assess a cultivated organism's transcription of genes that perform a task, coupled with the demonstration that the task occurs *in situ*. This presents a problem for assessing the microbial activity at *in situ* conditions, environment, and comprehendible rates. Many organisms are not only refractory to current laboratory isolation or community culturing efforts (13) but also operate within time scales that are not pragmatic to long-term investigations (14). The permafrost literature presents both culture-based and *ex situ* investigations to circumvent the inability to receive real-time microbial metabolic measurements. Most culture efforts, however, cultivate organisms within 4°C to 30°C to overcome the time barrier imposed when organisms are grown within their true environmental temperatures (15–17). Only five psychrophilic organisms have had transcripts studied at cold temperatures (<0°C): *Exiguobacterium sibiricum*, -2.5°C (18), *Psychrobacter arcticus,* -6°C (19), *Psychrobacter sp.* PAMC 2119, -5°C (20), *Planococcus halocryophilus* -15°C (21) and *Fragilariopsis cylindrus,* -1°C (22). The incomplete assessment of organismal activity based on cultivation approaches demands *in silico* approaches of investigation.

Two commonly used methods to overcome challenges associated with culture specific methods are metatranscriptomics and metagenomics. Metagenomics allows for a whole assessment of all the nucleic material in the environment, and with bioinformatic experiments (*in silico*), descriptions of those genetic features can be explored. This *in silico* approach allows estimation of a wide variety of organismal diversity, abundance, genetic potential and environmental interactions within extreme cryoenvironments (23). Metatranscriptomics and metagenomics paired together can demonstrate what genetic features are not only dominate, but those of which are active in the environment. These two methods, paired with curation of

metagenomic assembled genomes (MAGs), can lead to answering *which* organisms are present and *what* they are doing. Many studies pairing these two methods have been able to discern that there are microbial populations that are adapted to the current cryoenvironment they inhabit and are also transcribing some genes relating to that adaption (24, 25). Some of these adaptions included energy acquisition from aerobic and anaerobic processes as well as genetic capabilities to withstand freeze-thaw conditions.

Unfortunately, metatranscriptomes are difficult to retrieve from environments. This may be because free nucleic matter is consumed as an energy source by organisms that have scavenging proteins (12) or because mRNA studied in the lab degrades within minutes after creation (26) and probably faster in low-energy, frozen environments. This leaves metagenomics for inferring environmental functions, acknowledging the inability to claim *which* genes are active in the environment. In order to match the genes that are most likely to be active in the environment one must assess the environmental characteristics. Environmental geochemistry can therefore lend support to metagenomic inferences of microbial functions. Finding organisms that do not contain the genetic equipment to withstand a certain habitat could mean contamination, but it could also mean that the organism was transported to an environment that to which it is not adapted (i.e. thermophilic organisms in cryoenvironments or marine organisms in freshwater environment) (27). Endemic populations of living microorganisms can be studied by using genetic trends and comparative analysis between different locations. If a MAG contains the genetic features that would support environmental adaptation and subsistence in the specific cryoenvironment as compared to other environments then it is more likely that the organism, and its genes, are adapted for that environment.

Possible fate of permafrost through microbes

Permafrost can exist as “cold permafrost” or “warm permafrost”, where the difference is the temperature each remains frozen. Cold permafrost remains between -12°C and -8°C while warm permafrost is frozen at -2°C to 0°C (28). Svalbard permafrost is considered warm permafrost (29) while deep Siberian permafrost is mainly cold permafrost (9, 30, 31). This means that the Svalbard permafrost, among other high arctic permafrost, is more susceptible to thaw and incorporation into the active layer (32). Even though the Siberian permafrost is colder and would thaw after the warm permafrost, the total climate temperature increases will create an increased rate of greenhouse gas production from the active layer (30). The greenhouse gases (CO_2 and CH_4) that are expected to be released from the active layer are directly related to the microbiological activity that can be performed in the variable arctic soil types. Siberian permafrost peatlands will respond differently since their organic material content is higher than Svalbard permafrost which has high mineral content. Peatland carbon pools are poised to release hundreds of petagrams of greenhouse gases into the atmosphere via microbial degradation (33). Comparatively, mineral rich cryosols of Canada and Svalbard have been shown to consume atmospheric methane (34) and carbon dioxide (7), respectively. If the future of the soils is to act

as a carbon sink, then the increase in temperatures will stimulate plant productivity and carbon may be incorporated into plant biomass and change the Arctic landscape. If the thawing permafrost and increased activity increases microbial activity and the soil organic matter becomes newly degraded, then the environment may become a carbon source to the atmosphere (25). To understand the fate of the permafrost and active layer environments we must take a polyphasic assessment of the current microbiological population (10), the environmental characteristics (35), and the genetic assessment of the microbes' ability to interact with a dynamic system and adapt to change.

Overview of the research in this dissertation

In the following dissertation chapters, ancient Siberian permafrost and active layer Svalbard soils are investigated for microbial presence. First, Siberian permafrost aged up to 1,000,000-years-old was studied with metagenomic libraries to investigate if these ancient soils contained high quality nucleic matter. Eight metagenome assembled genomes were compared to 230 open access genomes and revealed genetic pathways for adaptations to liquid brines in the ancient permafrost. This chapter demonstrates that there are microbes adapted to the low-energy permafrost environment and can subsist within the environment.

The two following chapters analyze Svalbard active layer in two different ways. First, two short cores within the Leirhaugen glacier moraine are investigated geochemically to determine soil characteristics and how the isotopic signatures reveal indications of microbial metabolic presence. The carbon isotopic signatures of these two sites show that they differ in dominating microbial metabolisms where one site has carbon isotopic signatures that are indicative of autotrophy. Further, the carbon to nitrogen ratios of the soil displays that heterotrophy is the preferred metabolism in the location that contains more labile carbon. Bulk soil and ten cultured *Pseudomonas* isolates were assessed for their ability to degrade different compounds. These potential activities showed the microbial population was starved for carbon and nitrogen and was using amino acids to rectify that shortage. These enzymatic activities also showed adaptability in colder temperatures and ability to mitigate scarce environmental resources. Metagenomic libraries were queried for all annotations related to the enzymes and there were trends with the number of gene counts and the enzyme activities. That is to say that the lower activity enzymes had lowest gene counts in the libraries and vice versa.

The final research chapter of this dissertation analyzes 169 medium quality MAGs from five active layer cores in the same location around the Bayelva River in the Leirhaugen glacier moraine. The five cores were taken in two different seasons, winter (two cores, 44 sample depths) and summer (three cores 12 sample depths). In this investigation, we sequenced DNA from 2cm intervals of the active layer cores to investigate microbial correlation in high resolution stratification. The 2cm interval of the active layer core showed important correlations between MAGs that were not endemic to a certain depth. The network analyses displayed that the middle horizons of the core are most susceptible to cryoturbation and seasonal mixing from freezing and

thawing inclusions. Further, a population of *Actinobacteriota* were present in depths that the *Acidobacteriota* were in lower presence, suggesting a dominance shift the deeper from the surface soils.

This dissertation shows how *in silico* approaches allow us to investigate the low-DNA yielding environments and how those organisms subsist. MAGs from ancient Siberia have genetic adaptations that allow them to survive in the energy-depleted, frozen environment. Read mapping showed that the Svalbard MAGs were present in locations that they did not originate from, with temperature related patterns. There many *Actinobacteriota* MAGs that were present in the upper surface of one winter core and abundant in the three cores taken in the winter. These dissertation chapters are three parts to a continuing investigation of how microbes are adapted to the environment they are present yet contain the genetic ability to adapt to a warming world.

Reference

1. French HM. 2007. The Periglacial Environment. John Wiley & Sons Ltd,., West Sussex, England.
2. Tarnocai C, Canadell JG, Schuur EAG, Kuhry P, Mazhitova G, Zimov S. 2009. Soil organic carbon pools in the northern circumpolar permafrost region. Global Biogeochem Cycles 23:n/a-n/a.
3. Schuur EAG, Abbott B. 2011. Climate change: High risk of permafrost thaw. Nature 480:32–33.
4. Veremeeva A, Gubin S. 2009. Modern tundra landscapes of the Kolyma Lowland and their evolution in the Holocene. Permafr Periglac Process 20:399–406.
5. Péwé TL. 2006. permafrost -- Britannica Academic. Br Acad.
6. Åkerman HJ, Johansson M. 2008. Thawing permafrost and thicker active layers in sub-arctic Sweden. Permafr Periglac Process 19:279–292.
7. Jentzsch K, Schulz A, Pirk N, Foken T, Crewell S, Boike J. 2021. High levels of CO2 exchange during synoptic-scale events introduce large. Geophys Res Lett 1–9.
8. Christiansen, Hanne H., Gilbert, Graham L., Neumann, Ullrich, Demidov, Nikita, Guglielmin, Mauro, Isaksen, Ketil, Osuch, Marzena, Boike J. 2020. Ground ice content, drilling methods and equipment and permafrost dynamics in Svalbard 2016-2019 (PermaSval)SESS Report 2020-The State of Environmental Science in Svalbard.
9. Beermann F, Langer M, Wetterich S, Strauss J, Boike J, Fiencke C, Schirrmeister L, Pfeiffer EM, Kutzbach L. 2017. Permafrost Thaw and Liberation of Inorganic Nitrogen in Eastern Siberia. Permafr Periglac Process 28:605–618.
10. Jansson JK, Taş N. 2014. The microbial ecology of permafrost. Nat Rev Microbiol 12:414–425.
11. Panikov NS, Flanagan PW, Oechel WC, Mastepanov MA, Christensen TR. 2006. Microbial activity in soils frozen to below -39°c. Soil Biol Biochem 38:785–794.
12. Bird JT, Tague ED, Zinke L, Schmidt JM, Steen AD, Reese B, Marshall IPG, Webster G, Weightman A, Castro HF, Campagna SR, Lloyd KG. 2019. Uncultured microbial phyla suggest mechanisms for multi-thousand-year subsistence in Baltic Sea sediments. MBio 10:e02376-18.
13. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. 2018. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. mSystems 3:e00055-18.
14. Zinke LA, Mullis MM, Bird JT, Marshall IPG, Jørgensen BB, Lloyd KG, Amend JP, Kiel Reese B. 2017. Thriving or surviving? Evaluating active microbial guilds in Baltic Sea sediment. Environ Microbiol Rep 9:528–536.
15. Vishnivetskaya T, Kathariou S, McGrath J, Gilichinsky D, Tiedje JM. 2000. Low-temperature recovery strategies for the isolation of bacteria from ancient permafrost sediments. Extremophiles 4:165–173.
16. Sonjak S, Frisvad JC, Gunde-Cimerman N. 2006. Penicillium mycobiota in Arctic subglacial ice. Microb Ecol 52:207–216.
17. Finster KW, Herbert RA, Lomstein BA. 2009. Spirosoma spitsbergense sp. nov. and Spirosoma luteum sp. nov., isolated from a high Arctic permafrost soil, and emended description of the genus Spirosoma. Int J Syst Evol Microbiol 59:839–844.

18. Rodrigues DF, Ivanova N, He Z, Huebner M, Zhou J, Tiedje JM. 2008. Architecture of thermal adaptation in an Exiguobacterium sibiricum strain isolated from 3 million year old permafrost: A genome and transcriptome approach. BMC Genomics 9:547.
19. Bergholz PW, Bakermans C, Tiedje JM. 2009. Psychrobacter arcticus 273-4 Uses Resource Efficiency and Molecular Motion Adaptations for Subzero Temperature Growth. J Bacteriol 191:2340–2352.
20. Koh HY, Park H, Lee JH, Han SJ, Sohn YC, Lee SG. 2017. Proteomic and transcriptomic investigations on cold-responsive properties of the psychrophilic Antarctic bacterium *P sychrobacter* sp. PAMC 21119 at subzero temperatures. Environ Microbiol 19:628–644.
21. Mykytczuk NCSS, Foote SJ, Omelon CR, Southam G, Greer CW, Whyte LG. 2013. Bacterial growth at −15 °C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. ISME J 7:1211–1226.
22. Mock T, Hoch N. 2005. Long-term temperature acclimation of photosynthesis in steady-state cultures of the polar diatom Fragilariopsis cylindrus. Photosynth Res 85:307–317.
23. Raymond-Bouchard I, Whyte LG. 2017. From transcriptomes to metatranscriptomes: Cold adaptation and active metabolisms of psychrophiles from cold environments, p. 437–457. *In* Psychrophiles: From Biodiversity to Biotechnology: Second Edition. Springer International Publishing.
24. Hultman J, Waldrop MP, Mackelprang R, David MM, McFarland J, Blazewicz SJ, Harden J, Turetsky MR, McGuire AD, Shah MB, VerBerkmoes NC, Lee LH, Mavrommatis K, Jansson JK. 2015. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. Nature 521:208–212.
25. Mackelprang R, Saleska SR, Jacobsen CS, Jansson JK, Taş N. 2016. Permafrost Meta-Omics and Climate Change. Annu Rev Earth Planet Sci 44:439–462.
26. Rauhut R, Klug G. 1999. mRNA degradation in bacteria. FEMS Microbiol Rev 23:353–370.
27. Müller AL, De Rezende JR, Hubert CRJ, Kjeldsen KU, Lagkouvardos I, Berry D, Jørgensen BB, Loy A. 2014. Endospores of thermophilic bacteria as tracers of microbial dispersal by ocean currents. ISME J 8:1153–1165.
28. Keating K, Binley A, Bense V, Van Dam RL, Christiansen HH. 2018. Combined Geophysical Measurements Provide Evidence for Unfrozen Water in Permafrost in the Adventdalen Valley in Svalbard. Geophys Res Lett 45:7606–7614.
29. Biskaborn BK, Smith SL, Noetzli J, Matthes H, Vieira G, Streletskiy DA, Schoeneich P, Romanovsky VE, Lewkowicz AG, Abramov A, Allard M, Boike J, Cable WL, Christiansen HH, Delaloye R, Diekmann B, Drozdov D, Etzelmüller B, Grosse G, Guglielmin M, Ingeman-Nielsen T, Isaksen K, Ishikawa M, Johansson M, Johannsson H, Joo A, Kaverin D, Kholodov A, Konstantinov P, Kröger T, Lambiel C, Lanckman JP, Luo D, Malkova G, Meiklejohn I, Moskalenko N, Oliva M, Phillips M, Ramos M, Sannel ABK, Sergeev D, Seybold C, Skryabin P, Vasiliev A, Wu Q, Yoshikawa K, Zheleznyak M, Lantuit H. 2019. Permafrost is warming at a global scale. Nat Commun 10:1–11.
30. Nitzbon J, Westermann S, Langer M, Martin LCP, Strauss J, Laboor S, Boike J. 2020. Fast response of cold ice-rich permafrost in northeast Siberia to a warming climate. Nat Commun 11:1–11.

31. Zimov SA, Davydov SP, Zimova GM, Davydova AI, Schuur EAGG, Dutta K, Chapin FS, Chapin IS. 2006. Permafrost carbon: Stock and decomposability of a globally significant carbon pool. Geophys Res Lett 33:1–5.
32. Cohen J, Screen JA, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K, Entekhabi D, Overland J, Jones J. 2014. Recent Arctic amplification and extreme mid-latitude weather. Nat Geosci. Nature Publishing Group.
33. Schuur EAG, Bockheim J, Canadell JG, Euskirchen E, Field CB, Goryachkin S V., Hagemann S, Kuhry P, Lafleur PM, Lee H, Mazhitova G, Nelson FE, Rinke A, Romanovsky VE, Shiklomanov N, Tarnocai C, Venevsky S, Vogel JG, Zimov SA. 2008. Vulnerability of Permafrost Carbon to Climate Change: Implications for the Global Carbon Cycle. Bioscience 58:701–714.
34. Lau MCYY, Stackhouse BT, Layton AC, Chauhan A, Vishnivetskaya TA, Chourey K, Ronholm J, Mykytczuk NCSS, Bennett PC, Lamarche-Gagnon G, Burton N, Pollard WH, Omelon CR, Medvigy DM, Hettich RL, Pfiffner SM, Whyte LG, Onstott TC. 2015. An active atmospheric methane sink in high Arctic mineral cryosols. ISME J 9:1880–1891.
35. Taş N, Prestat E, Wang S, Wu Y, Ulrich C, Kneafsey T, Tringe SG, Torn MS, Hubbard SS, Jansson JK. Landscape topography structures the soil microbiome in arctic polygonal tundra.

CHAPTER 2 METAGENOME ASSEMBLED GENOMES FROM 1-MILLION-YEAR-OLD SIBERIAN PERMAFROST

A version of this chapter was originally published by Katie Sipes:
Katie Sipes, Abraham Almatari, Alexander Eddie, Daniel Williams, Elena Spirina, Elizaveta Rivkina, Renxing Liang, TC Onstott, Tatiana Vishnivetskaya, and Karen G. Lloyd. "Eight metagenome-assembled genomes provide evidence for microbial adaptation in 20,000 to 1,000,000-year-old Siberian permafrost" *Applied and Environmental Microbiology* (2021). https://doi.org/10.1128/AEM.00972-21

I contributed to this research chapter by processing the data files, generating hypothesis, analyzing the data, discerning results, and writing the publication.

Abstract

Permafrost microbes may be metabolically active in microscopic layers of liquid brines, even in ancient soil. Metagenomics can help discern whether permafrost microbes show adaptations to this environment. Thirty-three metagenome assembled genomes (MAGs) were obtained from six depths (3.5 m to 20 m) of freshly-cored permafrost from the Siberia Kolyma-Indigirka Lowland region. These soils have been continuously frozen for ~20,000 to 1,000,000 years. Eight of these MAGs were ≥80% complete with <10% contamination and were taxonomically identified as Aminicenantes, Atribacteria, Chloroflexi, and Actinobacteria within bacteria and Thermoprofundales within archaea. MAGs from these taxa have previously been obtained from non-permafrost environments and have been suggested to show adaptations to long-term energy-starvation, but they have never been explored in ancient permafrost. The permafrost MAGs had higher proportions of clusters of orthologous genes (COGs) from 'Energy production and conversion' and 'Carbohydrate transport and metabolism' than their non-permafrost counterparts. They also contained genes for trehalose synthesis, thymine metabolism, mevalonate biosynthesis and cellulose degradation that were less prevalent in non-permafrost genomes. Many of these genes are involved in membrane stabilization and osmotic stress responses, consistent with adaptation to the anoxic, high ionic strength, cold environments of permafrost brine films. Our results suggest that this ancient permafrost contains DNA in high enough quality to assemble MAGs from microorganisms with adaptations to subsist long-term freezing in this extreme environment.

Introduction

Arctic soils are often studied by sampling the upper seasonally-thawed active layer (1). Climate change is rapidly increasing the depths of permafrost active layers worldwide (2), making it necessary to learn more about the microbes living in the deeply-buried permafrost before it thaws completely. Evidence for microbes persisting in ancient permafrost has come in the form of revivable cultured isolates, microscopic visualization of intact cells, live/dead cell staining, low amino acid D/L ratios in cellular biomass, high quality metagenomic reconstructions of microbial genomes, and the presence of high-quality RNA molecules (3–8). Evidence from other studies of 33,000 year old permafrost in Alaska suggests that microbial communities are equipped to survive in this type of environment by using unique energy acquisition mechanisms and stress responses such as scavenging detrital biomass and environmental sensing (9). Variation in metagenomic functions of the in situ microbial populations in 30,000 year-old permafrost in the Kolyma-Indigirka lowland of Northeastern Siberia corresponded to local geochemistry, suggesting the populations were alive (8). Additionally, a study of >500,000 year-old permafrost samples from Northeast Russia contained microbial DNA with likely DNA repair mechanisms (10). Thus, some microbes in ancient permafrost may not be dormant (7, 11, 12) but instead maintain a continuous torpor-like state that allows them to subsist in the low energy environment.

The Yedoma geological suite originates from the late Pleistocene era in the Northeastern Siberia Kolyma-Indigirka Lowland region (13). Yedoma, or Ice Complex, permafrost deposits were frozen syngenetically (deposited and frozen simultaneously) ~20,000 to ~60,000 years ago (14). Yedoma spans approximately 1 million km2 (15) and is found, on average, up to 10 meters below the surface (14, 16). The next layer down, called Olyor, is unlike Yedoma since sediments were frozen epigenetically (deposited and then frozen) after being deposited ~0.6-1.6 million years before present (16, 17) and they contain few ice formations (18). Some of the perennially frozen soils of this region have been dated up to 3 million years old (19, 20). Even though the permafrost of the Kolyma area is of freshwater origin (13), other studies have shown that during freezing, small solutes were occluded from the surrounding mineral material, creating small brine films (5).

To live in the permanently frozen soils underneath the active layer without seasonal nutrient replenishment, a microbe must be able to sustain cellular integrity with a limited source of liquid water (19). Liquid water exists in permafrost in the form of cryopegs or small brine films (5) which are hypersaline liquid water films and lenses that can occur throughout the permafrost at temperatures of -10C and salt content of 140-300 g/L (19). These saline fluids can be in the shape of flattened films along sediment grains or in small pockets, depending on depositional characteristics (21).

While it is believed that the presence of liquid water in these brine films is sufficient to maintain live microbes within the thermostable deep permafrost, the interactions between microbial cells and the brine film have not been visualized at a microscopic level (22). However, plump cells with unique membrane adaptations have been visualized with electron microscopy in 200,000-year-old permafrost samples (23). Additionally, cells visualized from ice core and permafrost samples are notably very small (<1μm diameter) (24, 25), but not small enough to be completely immersed in the brine vein (19). Since these brine films are so small, ~1 nm wide (26), the cell membrane would have to make contact with the brine in order to act as a conduit for small molecule transport. Necessary microbial metabolic processes could occur in permafrost as long as physical mass transfer can occur in the environment (27). Since the cell's internal environment remains unfrozen (28) the cell could transfer small metabolites (27) between the intracellular matrix and the brine films.

Many organisms have been able to be cultured from Kolyma-Indigirka permafrost, including representatives from the *Carnobacterium* genus (29), *Gammaproteobacteria,* Actinobacteria and Firmicutes (30). Similar groups have been detected from 40,000 to 50,000-year-old permafrost (31). Microbes have also been isolated from Siberian permafrost in high salt concentrations (0.5% to 15% NaCl) (19, 32, 33) and in various media types (34), suggesting that they could survive in brines. Culture work, however, does not assess the entire microbial community, since many organisms have not been cultured, despite many attempts (35).

To supplement culturing studies, we used metagenomes to study the microbial populations in ~20,000 to 1,000,000-year-old permafrost formed near the freshwater Alazeya River in the

Kolyma-Indigirka Lowland Region of Northeastern Siberia. We determined the geological age of the upper Yedoma suite by 14C dating. The deeper soil age was determined to be 650,000 to up to 1,000,000-year-old in previous studies (13, 15, 17, 19, 36). Microbial communities, largely Firmicutes and Proteobacteria, have been identified by 16S rRNA gene analysis in permafrost at this location, and are likely alive since they have substantial amounts of intracellular DNA (4). In this study, we used metagenomes to determine if microbes in these ancient sediments: 1) produce metagenome assembled genomes (MAGs) which suggests the presence of high-quality nucleic matter, and 2) contain genes that suggest an adaptation to long-term subsistence under permafrost environmental conditions. We identified 8 MAGs with a ≥80% genome completeness and ≤10% contamination from 5 taxonomic groups. A unique set of functional genes were present in permafrost MAGs but absent in publicly available genomes of the same taxonomic groups in non-permafrost environments (henceforth referred to as non-permafrost genomes). These ancient permafrost MAGs give preliminary insight to the survival capabilities of these microbes in 20,000 to 1,000,000-year-old permafrost.

Materials and Methods

Field Sampling details

Permafrost cores were taken from the Northeastern Kolyma-Indigirka Lowland Region in Siberia, Russia, N69 20.376, E154 59.787 at the end of July 2015. A 20 m long core (AL3-15) was retrieved in a similar fashion as its sibling core (AL1-15), which was taken at the foothill in same location, as previously described (4). The temperature was measured inside the borehole at the end of each sampling day to allow the borehole to calibrate back to *in situ* temperature using an Onset® HOBO® Data Logger. The geological suites changed from Yedoma to Olyor depositional characteristics near 7.7 meters below surface.

Environmental characteristics

Permafrost samples were analyzed for pH, methane, and total carbon concentration. pH was measured with a benchtop pH meter (Mettler TOLEDO, Seven Easy pH). Conductivity was measured using a conductometer Ekspert-002 (Soil Services, Russia). To measure ions in the permafrost, pore water was extracted from 100g of permafrost following procedures from Van Reeuwijk, 2002 (80). Porewater analyses were performed following the standards of the Russian Federation GOST-R Certificate. Methane was liberated from the frozen permafrost soon after sample retrieval in the field as described previously (12) using a headspace method (81). Concentrations of the gas was analyzed on a Shimadzu GC-mini-2 (Tokyo, Japan) gas chromatograph with a hydrogen-flame ionization detector and argon as a carrier gas. The ^{14}C age of inorganic and organic carbon from three subsamples of the upper layers (2.9-3, 3.5 and 5.6 m)

was determined at Arizona Accelerator Mass Spectrometry Laboratory according to following procedures. The inorganic carbon was released by acidifying the bulk material using phosphoric acid (under vacuum) and collecting the CO_2. The residue was dried and combusted at 400˚C to release a relatively "younger" organic carbon fraction (82) (i.e., microbial cell material) from the permafrost sediment. The resulting residue was further combusted at 800˚C to oxidize any tightly bound carbon fraction that was potentially associated with clay particles in the sediments. This presumptive "older" carbon fraction could be inherited photosynthate or detrital carbon through geological time. An additional product of this gas collection is to be able to determine the $\delta^{13}C$ of the sample compared to Pee Dee Belemnite.

Metagenomic processing and analysis

Seven subsampled depths (3.5, 7.2, 14.1, 14.8, 16.6, 17.2, and 20 meters below the surface), spanning ~20,000 to 1,000,000-years-old, were used for metagenome sequencing. DNA was extracted from approximately 5 g of permafrost using 10 technical replicate extractions from the DNeasy PowerSoil Kit (Qiagen, Germantown, MD, USA) with 0.5 g of sample in each extraction. Ten extractions per sample were pooled together and DNA was ethanol precipitated and resuspended in 50 µl molecular-grade water. DNA concentrations ranged from 2.9 to 29.2 ng/µl. Metagenomic libraries were prepared using Nextera XT DNA Library Preparation Kit (Illumina, USA) following manufacture instructions and sequenced at the University of Tennessee Genomics Core on an Illumina MiSeq with 2 x 250 cycle protocol.

A diagram describing the following bioinformatic methods can be found in Supplemental Figure 1-S7. The raw metagenomic files were analyzed using KBase.us (83) and all links are in the Data Availability section (registration with https://www.kbase.us is free and required to access the narrative). The sample depths were kept separate for single sample assembling and metagenomic binning in order to avoid combining contigs from different sample depths into the same MAG. Trimmomatic v 0.36 (84) was used for poor-quality data trimming, removing barcodes and creating a paired-end assembly with parameters: leading_min_quality: 3, min_length: 36, sliding_window: 28, sliding_window_size: 10, trailing_min_quality: 3. MetaSPAdes v3.11.1 (85) was used for metagenomic assembly of minimum contig length of 2,000 base pairs. MaxBin2 v2.2.4 (86) default setting was used for placing contigs into metagenomic bins with parameters of prob_threshold: 0.8 and min_contig_length: 2000. Prodigal (87) was used for gene finding and DIAMOND (88) for gene translation within anvi'o v6 'esther' (89) using parameter 'anvi-run-ncbi-cogs' in the 'anvi-gen-contigs-database' command. Genome Taxonomy Data Base–Toolkit v 1.4.0 (GTDB-Tk) was used for initial taxonomic classification of the created bins (39) with the classify_wf command with a minimum alignment percent (min_perc_aa) of 10. A maximum likelihood phylogenetic tree was made in (Supplemental Datafile 6) CLC Workbench 6 using 139 shared conserved genes gathered from using 'anvi-get-sequences-for-hmm-hits' command with '—return-best-hit', '—get-aa-sequences', '—

concatenate' in anvi'o. MAGs were quality assessed with CheckM v 1.0.18 (90) for completeness and contamination percentages. Bowtie2's v2.3.3 (91) default settings was used to map the paired metagenomes against each of the MAGs to determine the presence of MAGs at other depths. This mapping also helps to find similarities and difference of the MAGs in each metagenomics samples. Mapping results were standardized by calculating reads per kilobase per million mapped reads (RPKM), which accounts for the gene length and each sample's library size (Supplemental Datafile 7). Non-permafrost genomes were downloaded from the Joint Genome Institute Integrated Microbial Genomes (JGI IMG) for comparison of genetic content (92). The five taxonomic groups the MAGs belonged to were *Actinobacteria:(n = 1, o = 57)*, *Aminicenantes:(n=1, o=54)*, *Thermoprofundales:(n=2, o=9)*, *Atribacteria:(n=3, o = 66)*, and *Chloroflexi:(n=1, o =42)*, where *n* is the number of permafrost MAGs and *o* is the number of non-permafrost genomes downloaded from JGI IMG.

The program anvi'o (89) was used to compare the annotated genes for similarity between genomes using DIAMOND (88) and NCBI database at an mcl inflation score of 2 within the groups. Comparisons were made between two experimental groups: 1) MAGs from this study against downloaded genomes (MAGs and SAGs) from the same phylogenetic group and 2) all the Siberian MAG groups against each other. These comparisons were made on the basis of gene annotations and, more broadly, the Clusters of Orthologous groups (COGs) (93). To investigate the KEGG pathways and metabolic pathways of the annotated Siberian MAGs the KO numbers from protein fasta files were compiled in GhostKoala and analyzed with KEGG-Decoder (94). All MAGs made from this study can be found on NCBI Accession BioProject PRJNA596250: Dimensions: MAGs from Siberian Permafrost.

Cell visualization

Cells were visualized by filtering and staining with SYBR Gold and DAPI from sonicated bulk soil with Nanopore 0.2μm filter (95) and imaged under a Zeiss Imager M2 with 100x magnification (Figure 2-S4). A Guava easyCyte 12HT Benchtop Cytometer was used with SYBR Green cell dye to examine cells as well. Only cells frozen at -80°C were available, without cell preservation in the field, so many of the cells may have lysed. Cell counting and live/dead staining methods did not yield quantifiable results. Additionally, single cell amplified genomes (SAGs) were also made from samples depths 5.6 m and 7.2 m at the Bigelow Single Cell Genomics Center in East Boothbay, Maine, USA (96). 10 single cells with the best whole genome amplification (Figure 2-S5) with Bigelow's Whole Genome Amplification (WGA) and Multiple displacement amplification (MDA) methods and were sequenced using an Illumina MiSeq 2 x 250 cycle protocol at the University of Tennessee, Knoxville.

Results

Site Characteristics

A 20 m deep permafrost core (AL3-15) from the Kolyma-Indigirka Lowland Region (69.339600, 154.996450) was subsectioned for microbiology and geology immediately upon retrieval in the field (Figure 2-1A). The permafrost geology enters a transition zone at about 6.2 m where it changes from the Yedoma geological suite to Olyor at 7.7 m. The temperature inside the borehole decreased from -2°C 1 m below the surface to -8.7°C 5 m below the surface and stayed at constant range of -8 to -8.6°C in the lower 10 meters of the borehole (Supplemental Figure 2-1). The δ13C of the inorganic carbon fraction (-3.4 to -6‰) was more 13C-enriched than the organic carbon (-28.8 to -24.3‰) at 2.9, 3.5 and 5.6 m (Supplemental Table 2-1). The 14C ages of the inorganic carbon increased from 21,760±120 to 33,900±550 yr with stratigraphic depth over the upper 5.6 m of the Yedoma suite. However, the 14C ages of the two organic carbon pools showed no stratigraphic trend with great variability at different depths (Supplemental Table 2-1). The large difference between 2.9 m (38,590±980 to 41,700±1,400) and 3.5 m (18,228±78 to 20,158±99 yr) suggested that the ice-rich layer at 3.5 m might be an ice wedge that was much younger than the surrounding strata. The underlying sediments of the Olyor suite were too old for radiocarbon dating. However, with other methods like fossil and palynology records, these soils were previously dated between 650,000 and 1,000,000-years-old (13, 15, 17, 19, 36). The porewater extracted from thawed subsamples of permafrost showed fluctuations in dissolved ions downcore but generally increased with depth (Figure 2-1A). Chloride had the greatest increase down the core, ranging from 1 mmol/100gsoil at 3.5 m to 2.7 mmol/100gsoil at 10.2 m. No chloride measurements were made below this depth. Methane increased from 1.2 to 17.6 mmol/100gsoil and sulfate increased from 0.019 to 2.5 mmol/100gsoil from depths 3 m to 17.2 m, respectively (Figure 2-1B). Methane was high only in the older Olyor deposits below about 13 m. The pH stayed at a constant 8-8.5 range throughout the core (Supplemental Figure 2-1). Cells were visualized with DAPI and SYBR Gold staining (Supplemental Figure 2-2A), but were below the quantification limit (37) for microscopic direct cell counts. Further evidence of intact cells comes from flow cytometry, which showed a population of similarly sized particles that stained with SYBR Gold (Supplemental Figures 2-2B-C). None of the 10 sequenced single cell amplified genomes (SAGs) were high enough quality to use in this study (0% to 45% completeness). The only Genome Taxonomy Database-Toolkit (GTDB-Tk) classifiable SAG was a Firmicute (45% complete, 5% contamination) from 7.2 m, matching the taxonomy of MAG_07_7.2m (which is not analyzed in this study due to its poor quality but can be found at the Data Accessibility NCBI Deposit Link).

Metagenomes

Of the seven depths from which DNA was extracted (3.5, 7.2, 14.1, 14.8, 16.6, 17.2, and 20 m), six samples produced metagenomic assemblies (Table 2-1). The metagenome from 17.2

m only produced ~1 MB of data that did not assemble, so it was not included in the analysis. The number of reads, assembled contigs, mean read length, N50, and the total number of MAGs did not trend with the sample depth or the total size of the retrieved data (Table 2-1). Every assembled metagenome had at least 91% of the contigs incorporated into the MAGs, suggesting that the MAGs represented most of the microbial community that was assembled (Table 2-1). The 6 metagenomes produced a total of 33 MAGs (Supplemental Table 2-2), 8 of which were ≥80% complete with <10% contamination (Supplemental Figure 2-2). GTDB-Tk and a maximum likelihood tree with 139 conserved proteins placed these MAGs into the following groups: Atribacteria (n = 3), Chloroflexi (n = 1), Aminicenantes (n = 1), Actinobacteria (n = 1), and Thermoprofundales (n = 2) among a total of 230 publicly-available reference genomes (Figure 2-2). A search through the public databases, National Center for Biotechnology Information (NCBI) and the Joint Genome Institute Integrated Microbial Genomes (JGI IMG), resulted in no MAGs from these groups from other ancient permafrost studies. The remaining 25 Siberian MAGs did not meet the quality standards for inclusion in our analysis. These were classified as *Caldatribacteriota* (n = 2), *Chloroflexota* (n = 2), *Aminicenantes* (n = 3), *Actinobacteriota* (n = 9), *Firmicutes* (n = 2), *Bathyarchaeota* (n = 2), *Planctomycetes* (n = 3), *Acidobacteriota* (n = 1), and *Thermoprofundales* (n = 1) (Supplemental Table 2-2).

Description of major taxonomic groups

Thermoprofundales. Two MAGs, one from 7.2 m (96% complete, 6% contamination) and another from 14.8 m (94% complete, 2% contamination) grouped with Thermoprofundales genomes (Figure 2-2; Supplemental Figure 2-2). Thermoprofundales is an order (38) of archaea within phylum *Euryarchaeota*, and class *Thermoplasmata* (Supplemental Table 2-2). *Thermoplasmata* has been redesignated as the phylum *Thermoplasmatota* based on GTDB taxonomy (39) and also the class *Izemarchaea* (40). It was previously called DHVEG, which stands for Deep-Sea Hydrothermal Vent Euryarchaeotal Group (41), and Marine Benthic Group–D (42). Thermoprofundales are uncultured archaea predominantly found in marine subseafloor seeps worldwide (38, 42–44), and were recently detected in anoxic freshwater sediments (45). All 12,042 of the *Thermprofundales* 16S rRNA gene sequences in the SILVA r138 were from anoxic marine sediments. To our knowledge, only 9 Thermoprofundales genomes are publicly available (Supplemental Datafile 1), originating from sediment from the White Oak River estuary in North Carolina (5 MAGs) and Aarhus Bay in Denmark (4 SAGs).

Atribacteria *Atribacteria* MAGs were retrieved from 7.2 m (90% complete, 0% contamination), 14.1 m (81% complete, 3% contamination), and 14.8 m (89% complete, 0% contamination) (Figure 2-2; Supplemental Figure 2-2). The phylum Atribacteria was previously called JS1/OP10 (46) and has been proposed to be renamed through GTDB taxonomy as Caldatribacteriota (39). The 66 non-permafrost Atribacteria MAGs (Supplemental Datafile 2) originate across a variety

of locations, including marine sediment from Aarhus Bay, Denmark; Great Boiling Springs, Nevada; soil from an acetate-fed aquifer in Rifle, Colorado, Walnut Creek; Green River sediment in Utah USA; Sakinaw Lake in Canada; a terephthalate-degrading reactor biofilm in Mexico; Etoliko Lagoon in Greece; and sediment in Irkutsk Lake Baikal in Russia. The 455 Atribacteria 16S rRNA genes in the SILVA r138 database are typically from temperate soil, deep marine sediment, and bioreactors. To our knowledge, neither 16S rRNA genes nor MAGs from this group have been previously reported from permafrost environments.

Chloroflexi One MAG from 20 m (81% complete 10% contamination) grouped within the Chloroflexi order Anaerolineales (Figure 2-2; Supplemental Figure 2-2). Anaerolineales is abundant in many different types of environments; a meta-analysis of 1,504 metagenomes showed Chloroflexi DNA sequences were in the top ten most abundant groups in metagenomes from marine sediment, host-associated environments, hypersaline environments, freshwater, hot springs, and terrestrial subsurface (35). Forty-two non-permafrost Chloroflexi (Supplemental Datafile 3) were chosen to represent the phylogenetic and environmental diversity available among the non-permafrost genomes from deep Pacific oceanic basalt-hosted subsurface hydrothermal fluid, White Oak River estuary sediment, Utah Grand County groundwater, a biological phosphorous bioreactor, marine samples from the TARA04 Ocean Project, groundwater from an acetate-fed Rifle, Colorado aquifer, Green River; Utah freshwater, sediment from Australia, and marine aquatic samples with unreported locations.

Actinobacteria One MAG (MAG_01_20m; 96% complete, 8% contamination) was in the Actinobacteria phylum, with family Dermatophilaceae and genus Cutibacterium, which is commonly cultured from human skin (47). A complete nuclease gene from this MAG had 100% similarity in a BLASTN search to *Cutibacterium acnes*, and the entire MAG had a 99.72% average nucleotide identity (ANI) (48) to *Propionibacterium, Cutibacterium sp. KPL2009*, from the human genome project (47). The other Siberian permafrost *Actinobacteria* MAGs were not closely related to the skin microbe (Supplemental Table 2-2), but since they had lower completeness and higher contamination scores, their phylogenetic relatedness could not be accurately determined.

Aminicenantes One MAG (MAG_02_14.8m) with 94% completeness and 3% contamination came from the 14.8 m sample and grouped within Aminicenantes genomes on a maximum likelihood tree based on concatenated conserved genes (Figure 2-2; Supplemental Figure 2-2). *Aminicenantes* was previously called OP8 and has been suggested to be in the *Acidobacteriota* phylum based on the GTDB reclassification (46). The *Aminicenantes* MAG was compared to 54 non-permafrost *Aminicenantes* genomes (Supplemental Datafile 5), chosen so that each available study site was represented: Etoliko Lagoon in Greece; Sakinaw Lake in Canada; soil in Rifle,

Colorado; hydrothermal fluid from the Juan de Fuca Ridge flank in the Pacific Ocean; and marine sediment from Baltimore Maryland, USA (Supplemental Datafile 5).
Aminicenantes has been found in a variety of marine and terrestrial environments, but only up to 10.2% maximum relative abundance (49). *Aminicenantes* comprised a slightly higher percentage of mapped reads in most of our permafrost metagenomes (3.5 m: 12.77%, 7.2 m: 15.77%, 14.1 m: 16.81%, 14.8 m: 14.7%, 16.6 m: 5.6% and 20 m: 0.14%; Figure 2-5 and Supplemental Datafile 7). This suggests that Aminicenantes may be in higher relative abundance in ancient permafrost, or the discrepancy may be due to primer bias (50) during PCR amplification in the published studies.

Comparison of Siberian MAGs to non-permafrost genomes

Genes that were present in Siberian MAGs that were rare or absent in the non-permafrost MAGs/SAGs included 21 major COG categories (Figure 2-3A), excluding COG categories 'Chromatin structure and dynamics', 'RNA processing', 'Secondary metabolites', and 'Cytoskeleton'. The COG categories with the largest number of genes unique to the Siberian MAGs were 'Carbohydrate metabolism and transport', 'Energy production and conversion' "Function unknown" and "General functional prediction only". The Siberian Thermoprofundales MAGs had the most genes that were absent in non-permafrost genomes (47 genes, Supplemental Datafile 6). The greatest numbers were 'Carbohydrate metabolism and transport', with four components of an ABC-type sugar transport system, ABC-type glycerol phosphate systems, cellobiose phosphorylase, and pyruvate kinase, 'Energy production and conversion', with FoF1-type ATP synthase and isocitrate dehydrogenase, and 'Inorganic ion transport and metabolism', with two components of a Ca2+/H+ antiporter and a Na+/H+ antiporter related arsenite permease (Figure 2-3A). The Siberian Thermoprofundales MAGs also had more genes per genome related to the 'Carbohydrate metabolism and transport' (2.5x) and 'Defense' COG Categories (3x) than the non-permafrost genomes (Supplemental Figure 2-3).
There were no genes that were present in all Siberian Atribacteria MAGs that were also absent from all the non-permafrost MAGs. Therefore, we investigated the genes in all Siberian Atribacteria MAGs found in only 10 to 33% of the non-permafrost Atribacteria MAGs (Supplemental Datafile 6). Of those 32 genes, 5 were in the 'Carbohydrate metabolism and transport' COG Category (Figure 2-3, Supplemental Datafile 2). The seven genes with the lowest representation (11 to 23%) in non-permafrost Atribacteria were pyruvate/oxaloacetate carboxyltransferase, trehalose-6-phosphate synthase/hydroxylamine reductase (hybrid-cluster protein), cellobiose phosphorylase, nucleoside diphosphate kinase, cation transport ATPase, and ribosomal protein L32 (Supplemental Datafile 2 and 6).

Thirty-three genes were present in at least one out of three Siberian Atribacteria MAGs, that were absent in all non-permafrost Atribacteria. Six of these were from the 'Energy production and conversion' COG category: four subunits of NADH:ubiquinone oxidoreductase and two

subunits of a heme/copper type cytochrome/quinol oxidase. Three Siberian Atribacteria MAG genes were in the 'Carbohydrate metabolism and transport' COG category, including mevalonate-3-phosphate 5-kinase, Aryl-phospho-beta-D-glucosidase, and a phosphoglycerate mutase. The three permafrost Atribacteria MAGs had twice the number of genes per genome related to the 'Energy production and conversion', 'Inorganic ion transport and metabolism', 'Replication, recombination and repair', 'Cell wall/membrane/envelope biogenesis' and 'Lipid transport and metabolism' COG categories as non-permafrost genomes (Supplemental Figure 2-3).

The Siberian Chloroflexi MAG had 20 genes that were absent in all non-permafrost genomes (Figure 2-3; Supplemental Datafile 6), including a Na+/alanine symporter, fumarate hydratase class II and nitric oxide reductase large subunit. A complete methyl-coenzyme M reductase contig (subunits alpha, beta, gamma, and operon protein D; 3483 amino acids long) binned in the Chloroflexi MAG, but this contig had no other genes on it (Supplemental datafile 3). The COG categories with the most genes that were only found in the Siberian Chloroflexi MAG, relative to non-permafrost MAGs, were 'Coenzyme transport and metabolism', 'Extracellular structures', and 'Cell motility' (Figure 2-3, Supplemental Datafile 3). The Siberian Chloroflexi MAG generally had 1-1.5x genes per genome in each COG Category than the 42 non-permafrost Chloroflexi (Supplemental Figure 2-3).

Eleven genes were unique to the Siberian Aminicenantes MAG relative to non-permafrost Aminicenantes MAGs, including two genes in the 'Replication, recombination and repair' and two in the "Function unknown" COG Categories (Supplemental Datafile 6). The Aminicenantes MAG had only one gene (endo-1,4-beta mannosidase) that was unique to the 'Carbohydrate metabolism and transport' COG category, but it had 2.7x more genes per genome than in the non-permafrost genomes (Supplemental Figure 2-3). While there were only few genes unique to the Aminicenantes Siberian MAG, it did have more genes per genome in many COG categories than the non-permafrost genomes (Supplemental Figure 2-3).

Since many of the genes that were unique to the Siberian MAGs were involved in transport, osmoregulation, and carbohydrate utilization, but spanned multiple COG categories, we made custom groups for these three functions to examine them together (Figure 2-3B, Supplemental Datafile 6, Supplemental Table 2-3). The transporters group had the most genes that were unique to the Siberian MAGs, led by Thermoprofundales and Atribacteria with nine and eight genes, respectively. Atribacteria and Thermoprofundales also had the most unique genes in the carbohydrate utilization group with five and four, respectively. 'Carbohydrate metabolism and transport' had the most numbers of genes, four of which were ABC-type sugar/transport system genes (Figure 2-3B). The average number of genes within each COG Category in the Actinobacteria MAG were less than or equal to the number of genes per genome in the non-permafrost genomes (Supplemental Figure 2-3). Thermoprofundales had the most unique genes in the osmoregulation group too, and all taxa except Aminicenantes had a Na+ symporter for osmoregulation. The only gene from Aminicenantes that fell within these three groups was a succinate-acetate transporter.

Comparison of Siberian MAGs to each other

While no genes had identical annotations between all eight of the Siberian MAGs (Figure 2-4), each MAG had some version of three Na+/H+ antiporters NhaD, MnhC and MnhG as well as a biotin transporter. Also, many genes had similar annotations in different unions (Supplemental Datafile 6). Forty genes were present in seven Siberian permafrost MAGs that were not in the Actinobacteria MAG (Figure 2-4 and Supplemental Datafile 6). These included pyruvate-formate lyase-activating enzyme, formate hydrogenlyase, carbamoyltransferase, Na+/H+ antiporter, M28 family peptidase, predicted nucleotidyltransferase component of viral defense system, protein-L-isoaspartate O-methyltransferase, and cellobiose phosphorylase.

The six bacterial Siberian MAGs shared 145 gene annotations with each other, which was more than they did with the Thermoprofundales, the only archaeal group (Figure 2-4). The Siberian Actinobacteria had the largest number of genes (243) that were uniquely annotated from the other seven Siberian MAGs (Figure 2-4). Some of these genes included a H+/Cl+, Na+/H+ and K+/H+ antiporters with C-terminal TrkAC and CorC domains (Supplemental Datafile 6). The Siberian Thermoprofundales MAGs contained 206 genes that were uniquely annotated relative to the other Siberian MAGs (Figure 2-4; Supplemental Datafile 6). These included acetate kinase, isocitrate dehydrogenase, Ca2+/H+ antiporter, HSP90 ATPase, and a Na+/proline symporter (Supplemental Datafile 4). Other genes specific to the Siberian Thermoprofundales MAGs included 21 different archaeal type regulatory proteins and an uncharacterized conserved protein related to pyruvate-formate lyase activating enzyme, many archaeal-type synthetases, and many genes with general or predicted functions (Supplemental Figure 2-6A through C, Supplemental Datafile 3).

Membrane stability genes for cellular integrity were shared between two or more groups. Multiple genes and enzyme kinases for the mevalonate pathway were found in the two Thermoprofundales, one Chloroflexi MAG, and one Atribacteria MAG (Supplemental Datafile 6). Apart from Thermoprofundales, the other Siberian permafrost MAGs shared membrane associated genes, such as: peptidoglycan biosynthesis protein MviN/MurJ, bacterial cell division protein FtsW, and energy-coupling factor transporter ATP-binding protein EcfA2 (Figure 2-4, Supplemental Datafile 6). The Aminicenantes MAG had 96 unique genes that were not present in any other Siberian MAG (Figure 2-4). This MAG also contained 111 genes similar to genes annotated in Actinobacteria and Chloroflexi MAGs, which are two other widely found soil and permafrost microbes (51, 52).

Discussion

Comparison of the Siberian MAGs to those found in other permafrost studies.

Apart from the 3.5 m depth, which was an ice wedge with few sediment inclusions, and the 17.2 m depth, which did not produce a metagenome for unknown reasons, DNA was present

in sufficient quantity to produce MAGs throughout the 20,000 to 1,000,000-year-old permafrost. The identities of the MAGs from ancient Siberian permafrost differed greatly from those found in active layer permafrost environments. In the active layer of the Stordalen Mire, Woodcroft et al, (2018) found 1,434 MAGs (>70% completeness and <10% contamination) comprised of 27% Actinobacteria, 27% Aminicenantes, 14% Proteobacteria, 4% Chloroflexi, and 6% Euryarchaeota (53). Our ancient Siberian permafrost contained members of the Chloroflexi, Actinobacteria, and Aminicenantes but also had MAGs from groups that were absent in the Stordalen Mire active layer, such as Atribacteria and Thermoprofundales. This suggests there could be large differences in microbial community composition between seasonally thawed active layers and ancient permafrost. Atribacteria and Thermoprofundales were absent in Alaskan ancient permafrost from 19,000 to 33,000 years old (6, 7). Another difference between our findings in Siberian permafrost and other studies is that the ancient Alaskan permafrost study showed an increase of Firmicutes 16S rRNA and a decrease of Actinobacteria with permafrost age (6). A similar trend was observed in the borehole adjacent to this study, AL3-15, called AL1-15 (4) where Firmicutes and Proteobacteria dominated 16S rRNA gene amplicon libraries. These taxa; Firmicutes, Planctomycetota, Crenarchaeota, and Bacteroidota were found in Siberian MAGs but were not analyzed due to insufficient quality. This discrepancy may be due to the differential biases of amplicon-based libraries and metagenomic-based libraries in low biomass samples (35, 50). The presence of groups such as Atribacteria and Thermoprofundales that are not commonly found in permafrost or freshwater environments, but are more commonly found in marine environments, suggest that these ancient freshwater Siberian deposits may contain organisms adapted to saline conditions. Genes common in Siberian permafrost MAGs indicate adaptation to cold, saline, low energy environment.

Saline regulatory genes

Many genes shared among two or more groups of the Siberian MAGs were involved in the transport of small molecules like sodium and carbohydrates. Having a variety of genes for ion and salt transportation could mean that these MAGs resemble organisms that are adapted to function in a high ionic environment, like brines. All groups shared a similar annotation for three Na+/H+ antiporters and osmoregulators (NhaD, MnhC and MnhG) as well as a biotin transporter. All Siberian MAGs had unique genes in the Transporter category as compared to the non-permafrost genomes (Figure 2-3B).

Trehalose

Many of the genes that were specific to Siberian MAGs, and were not present in non-permafrost outgroup genomes, were involved in transport, osmoregulation, and carbohydrate utilization. The prominence of these categories suggests the Siberian MAGs had unique adaptations to interact with their environment, including dealing with osmotic stress and using

carbohydrates as energy sources. Trehalose-6-phosphate synthase was found in the Siberian MAGs and lacked homologues in non-permafrost genomes. It was in both Thermoprofundales MAGs (22% of the non-permafrost outgroup), all three Atribacteria MAGs (13% of the non-permafrost outgroup), the Aminicenantes MAG (2% of the non-permafrost outgroup) and the Actinobacteria MAG (59% of the non-permafrost outgroup). The Chloroflexi MAG had a gene annotated as 'Trehalose and maltose hydrolase (possible phosphorylase)'. Trehalose-6-phosphate synthase has been suggested to help deep subsurface inhabitants maintain a low-energy state in marine sediments by producing trehalose, a disaccharide that prevents aggregation of degraded proteins, protects against osmotic stress, and increases cellular longevity (54, 55). Trehalose is a cryoprotectant and stabilizes a cellular membranes and DNA in low temperatures and high osmolarity to increase cell longevity and slow replication rates (56–58). Trehalose synthase has also been found in metagenomes from Antarctic soil and sediment samples (59).

Mevalonate pathway

Multiple genes and enzyme kinases for the mevalonate pathway were found in both Thermoprofundales MAGs, the Chloroflexi MAG, and the Atribacteria MAG. The mevalonate pathway uses acetic acid for the biosynthesis of isoprenoids, which have been shown to stabilize membranes, which increase an organism's survival in low temperatures (60–62) and have been found in bacteria (63, 64). A functional mevalonate pathway has also been identified in a Methanosarcina MAG from a deep Antarctic permafrost enrichment (65). The source of the acetate ions for the mevalonate pathway could be from fermentation of the surrounding organic matter from the cellobiose, arabinoxylans, and proteinaceous peptides. Genes to ferment these substrates were also identified in the Siberian MAGs. Additionally, Psychrobacter arcticus 273-4, cultured from Siberian permafrost, has phenotypic evidence of an acetate-based metabolism (66, 67) and all eight MAGs from this study had genes indicating a mixed acid acetate metabolism (Supplemental Figure 2-6).

Cellobiose phosphorylase and carbamate kinase

Metagenomes of thawed permafrost have shown an increase in degradation of cellulose (68). Cellobiose phosphorylase genes, which is an enzyme that aids in the degradation of cellulose as a carbon source (69–71), were identified in all our Siberian MAGs, except the Actinobacteria MAG. Cellobiose phosphorylase converts cellobiose into glucose and glucose-1-phosphate, offering an energetic advantage in anoxic environments. Cellobiose phosphorylase suggests that permafrost organisms may be able to slowly degrade available carbon substrates (44, 45). The Siberian Atribacteria MAGs had other genes indicative of adaptation to low energy environments. Further, two Siberian Atribacteria MAGs contained aconitase A, alpha-A-arabinofuranosidase, and alpha-amylase/alpha-mannosidase, that were in only 6%, 15% and 19% of non-permafrost genomes, respectively. These are important in cellulose degradation and carbohydrate metabolism (46). Carbon starvation protein (cstA) was in two Siberian Atribacteria MAGs and

30% of non-permafrost Atribacteria. This protein enhances peptide catabolism during carbon starvation (45). Carbamate kinase was in two Siberian Atribacteria MAGs and 25% of non-permafrost Atribacteria, suggesting that the permafrost Atribacteria can conserve energy by creating ADP and carbamoyl phosphate from the combination of ATP, CO2 and NH3 (47). This would benefit organisms in a low energy environment because this enzyme has roles within purine, glutamate, proline, and nitrogen metabolisms (48).

Biotic methane

Through their genome-centric analysis of functional genes in active layer soils, Woodcroft et al. (2018) found genetic evidence for organic matter decomposition into CO2 and CH4. Studies of 33,000-year-old permafrost (9) and recently thawed permafrost (68) from Alaska also generated a variety of methanogens. A gene for methyl coenzyme M reductase (mcr) was identified in another study's metagenome from a 30,696-year-old Siberian permafrost sample where methane was measured (8). Additionally, biogenic methane was detected in one of two boreholes in another study of Kolyma-Indigirka Lowland permafrost while methanogens were distributed throughout both boreholes (72). Other permafrost up to 33,000 years old, where methane was absent, did not yield 16S rRNA genes from known methanogenic groups (7), or any mcr genes (8). There was no evidence of methane metabolism in any of the Siberian Thermoprofundales MAGs, unlike their non-permafrost counterparts. We found four subunits of the mcr gene on a single contig from the 20 m metagenome sample in the Chloroflexi MAG, suggesting that the methane observed there could have been biotically produced (Figure 2-1B). The gene's top BLASTP hit was an mcr found in Candidatus "Methanoperedens ferrireducens" archeaon from an Australian marine sediment incubation (73). Since this MAG did not contain a full methanogenic pathway, and no bacteria have previously been shown to contain mcr it is likely that this mcr gene was part of the 10% contamination.

DNA scavenging

Ureidoglycolate dehydrogenase was present in two Siberian Atribacteria MAGs and 48% of non-permafrost Atribacteria. This is involved in the degradation of allantoin; a decomposition product of DNA and allows the use of DNA as a nitrogen source. Other Atribacteria have been suggested to use an allantoin degradation pathway to access detrital DNA as an energy source under extreme starvation conditions in deep subsurface marine sediment (39). The presence of DNA-foraging enzymes like those involved in allantoin degradation (55) and cellular debris recycling (31) may be another key to survival in low energy environments. The Siberian Aminicenantes MAG had a protein-degrading metabolic pathway similar to that inferred in other Aminicenantes (55), since they all had the same types of extracellular peptidases such as dipeptidase, dipeptidyl aminopeptidase, acylaminoacyl peptidase, RecA-mediated SOS response autopeptidase, D-alanyl-D-alanine carboxypeptidase and cyanophycinase exopeptidases (Supplemental Figure 2-6 and Supplemental Datafile 5). Atribacteria, Thermoprofundales and

Aminicenantes from subseafloor sediments also have been found to have these extracellular enzymes (74) Additionally, Thermoprofundales has been suggested to ferment proteinaceous organic matter since it has a high genomic content of extracellular peptidases whose activity could be measured in the bulk sediment (43). Therefore, the presence of genes for osmotic, cold, and energetic stress tolerance in our ancient Siberian permafrost MAGs suggests they may be adapted to this environment.

Conclusions

The Siberian permafrost MAGs belong to taxonomic groups that are commonly found in low-energy, anoxic, saline habitats. Finding MAGs related to marine-associated microbes like Thermoprofundales, Aminicenantes, and Atribacteria further supports that these MAGs are adapted to increased salinity (75) due to freezing of freshwater sediments, since the Yedoma and Olyor deposits in this area of Siberia have never been inundated by seawater (13, 14, 17). Concentrations of chloride increased with depth (Figure 2-1) up to 879.8 ppm (248 mM), which is hypersaline compared to accepted baseline tested agriculture soils, where 50 ppm is considered excessive (76). The high prevalence of Na+/H+, Ca+2/alanine antiporters, mechanosensitive channels, osmoprotectants, and other ion transporters (Figure 2-3B) relative to the non-permafrost genomes, may suggest an adaptation to the saline conditions observed in our geochemistry (Figure 2-1), since these MAGs are common in halotolerant organisms (19, 77). Additionally, microbial cultures from Siberian permafrost also have osmotic-specific adaptations and activity in -15°C permafrost (78, 79). Genes for osmotic stress tolerance have also been found within Antarctic (31) and Canadian (78) permafrost metagenomes.

We speculate that upon burial and freezing, these permafrost microbes became the dominant organisms because those not adapted to the high salinity liquid water films in permafrost died. If these organisms are adapted to surviving in saline water brine films, then it is unlikely that they will retain their dominance in these microbial communities upon permafrost thawing due to climate change. As the permafrost thaws, these native brine-adapted microbial communities will likely be replaced by freshwater-adapted organisms when surface waters penetrate the newly thawed permafrost.

Further evidence to support extant life in this ancient Siberian permafrost is that the DNA was intact enough to produce reads that assembled into MAGs with less than 9% of the assemblies remaining unbinned (Table 2-1) and high read recruitment in the MAGs (Figure 2-5). Each MAG had some DNA read recruitment from other sample depths, which suggests that, with even greater sequencing depth, a more diverse population of MAGs might have been observed.

Finding evidence for these organisms in ancient 1,000,000-year-old permafrost samples furthers the idea of a microbial community being able to persist in ancient permafrost within the high salinity brine films. The Siberian permafrost MAGs analyzed in this study demonstrate how individual organisms can be adapted to their environment relative to members of the same taxonomic groups in non-permafrost environments. Although it is impossible to have absolute

certainty that these organisms were alive at the time of sampling, their genome functions and unique functionality compared to non-permafrost genomes (Figure 2-3 and 2-4) coupled with the geochemistry of the environment (Figure 2-1) suggests adaptations to the liquid brine films in ancient permafrost.

References

1. Morgalev YN, Lushchaeva I V., Morgaleva TG, Kolesnichenko LG, Loiko S V., Krickov I V., Lim A, Raudina T V., Volkova II, Shirokova LS, Morgalev SY, Vorobyev SN, Kirpotin SN, Pokrovsky OS. 2017. Bacteria primarily metabolize at the active layer/permafrost border in the peat core from a permafrost region in western Siberia. Polar Biol 40:1645–1659.
2. Park H, Kim Y, Kimball JS. 2016. Widespread permafrost vulnerability and soil active layer increases over the high northern latitudes inferred from satellite remote sensing and process model assessments. Remote Sens Environ 175:349–358.
3. Abramov A, Vishnivetskaya T, Rivkina E. 2020. Are permafrost microorganisms as old as permafrost? FEMS Microbiol Ecol 1–12.
4. Liang R, Lau M, Vishnivetskaya T, Lloyd KG, Wang W, Wiggins J, Miller J, Pfiffner S, Rivkina EM, Onstott TC. 2019. Predominance of Anaerobic, Spore-Forming Bacteria in Metabolically Active Microbial Communities from Ancient Siberian Permafrost. Appl Environ Microbiol 85:AEM.00560-19.
5. Gilichinsky D, Rivkina E, Bakermans C, Shcherbakova V, Petrovskaya L, Ozerskaya S, Ivanushkina N, Kochkina G, Laurinavichuis K, Pecheritsina S, Fattakhova R, Tiedje JM. 2005. Biodiversity of cryopegs in permafrost. FEMS Microbiol Ecol 53:117–128.
6. Mackelprang R, Burkert A, Haw M, Mahendrarajah T, Conaway CH, Douglas TA, Waldrop MP. 2017. Microbial survival strategies in ancient permafrost: insights from metagenomics. ISME J 11:2305–2318.
7. Burkert A, Douglas TA, Waldrop MP, Mackelprang R. 2019. Changes in the active, dead, and dormant microbial community structure across a pleistocene permafrost chronosequence. Appl Environ Microbiol 85:AEM.02646-18.
8. Rivkina E, Petrovskaya L, Vishnivetskaya T, Krivushin K, Shmakova L, Tutukina M, Meyers A, Kondrashov F. 2016. Metagenomic analyses of the late Pleistocene permafrost - Additional tools for reconstruction of environmental conditions. Biogeosciences 13:2207–2219.
9. Mackelprang R, Burkert A, Haw M, Mahendrarajah T, Conaway CH, Douglas TA, Waldrop MP. 2017. Microbial survival strategies in ancient permafrost: Insights from metagenomics. ISME J 11:2305–2318.
10. Johnson SS, Hebsgaard MB, Christensen TR, Mastepanov M, Nielsen R, Munch K, Brand T, Gilbert MTP, Zuber MT, Bunce M, Rønn R, Gilichinsky D, Froese D, Willerslev E. 2007. Ancient bacteria show evidence of DNA repair. Proc Natl Acad Sci U S A 104:14401–14405.
11. Willerslev E, Hansen AJ, Poinar HN. 2004. Isolation of nucleic acids and cultures from fossil ice and permafrost.
12. Gilichinsky DAA, Wilson GSS, Friedmann EII, Mckay CPP, Sletten RSS, Rivkina EMM, Vishnivetskaya TAA, Erokhina LGG, Ivanushkina NEE, Kochkina GAA, Shcherbakova

VA, Soina VSS, Spirina EV V., Vorobyova EAA, Fyodorov-Davydov DG, Hallet B, Ozerskaya SMM, Sorokovikov VAA, Laurinavichyus KSS, Shatilovich AV V., Chanton JPP, Ostroumov VEE, Tiedje JMM, Scherbakova VA, Soina VSS, Spirina EV V., Vorobyova EAA, Frodorov-Davydov DG, Hallet B, Ozerskaya SMM, Sorokovikov VAA, Laurinavichyus KSS, Shatilovich AV V., Chanton JPP, Ostroumov VEE, Tiedje JMM. 2007. Microbial Populations in Antarctic Permafrost: Biodiversity, State, Age, and Implication for Astrobiology. Astrobiology 7:275–311.

13. Veremeeva A, Gubin S. 2009. Modern tundra landscapes of the Kolyma Lowland and their evolution in the Holocene. Permafr Periglac Process 20:399–406.
14. Strauss J, Schirrmeister L, Grosse G, Fortier D, Hugelius G, Knoblauch C, Romanovsky V, Schädel C, Schneider von Deimling T, Schuur EAG, Shmelev D, Ulrich M, Veremeeva A. 2017. Deep Yedoma permafrost: A synthesis of depositional characteristics and carbon vulnerability. Earth-Science Rev 172:75–86.
15. Grosse G, Robinson JE, Bryant R, Taylor MD, Harper W, DeMasi A, Kyker-Snowman E, Veremeeva A, Schirrmeister L, Harden J. 2013. Distribution of late Pleistocene ice-rich syngenetic permafrost of the Yedoma Suite in east and central Siberia, Russia. Epic US Geol Surv Open File Rep , USGS, 2013(1078), pp 1-37.
16. Soina, V.S., Vorobiova, E.A., Zvyagintsev, D.G., Gilichinsky DA. 1995. Preservation of cell structures in permafrost: A model for exobiology. Adv Sp Res 15:237–242.
17. Schirrmeister L, Kunitsky V, Grosse G, Wetterich S, Meyer H, Schwamborn G, Babiy O, Derevyagin A, Siegert C. 2011. Sedimentary characteristics and origin of the Late Pleistocene Ice Complex on north-east Siberian Arctic coastal lowlands and islands – A review. Quat Int 241:3–25.
18. Lupachev A V., Gubin S V., Gerasimova MI. 2019. Problems of the cryogenic soils' diagnostics in the recent Russian Soil Classification System. Eurasian Soil Sci 52:1170–1174.
19. 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 3:331–341.
20. Knelman JE, Legg TM, O'Neill SP, Washenberger CL, González A, Cleveland CC, Nemergut DR. 2012. Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. Soil Biol Biochem 46:172–180.
21. Tolstikhin NI, Tolstikhin ON. 1974. Groundwater and surface water in the permafrost regionUSSR Academy of Sciences, Siberian Branch, Novosibirsk.
22. Price PB, Sowers T. 2004. Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. Proc Natl Acad Sci U S A 101:4631–4636.
23. Soina VS, Mulyukin AL, Demkina E V., Vorobyova EA, El-Registan GI. 2004. The structure of resting bacterial populations in soil and subsoil permafrost. Astrobiology

4:345–358.
24. Sheridan PP, Miteva VI, Brenchley JE. 2003. Phylogenetic analysis of anaerobic psychrophilic enrichment cultures obtained from a greenland glacier ice core. Appl Environ Microbiol 69:2153–60.
25. Gilichinsky DA, Soina VS, Petrova MA. 1993. Cryoprotective properties of water in the Earth cryolithosphere and its role in exobiology. Orig Life Evol Biosph 23:65–75.
26. Anderson D. M. 1967. Ice Nucleation and the Substrate-ice Interface. Nature 216:563–566.
27. Ostroumov VE, Siegert C. 1996. Exobiological aspects of mass transfer in microzones of permafrost deposits. Adv Sp Res 18:79–86.
28. McGrath J, Wagener S GD. 1994. Cryobiological studies of ancient microorganisms isolated from the Siberian permafrost. Viable microorganisms in permafrost. Pushchino Pushchino Res Centre, Russ Acad.
29. Nicholson WL, Krivushin K, Gilichinsky D, Schuerger AC. 2013. Growth of Carnobacterium spp. from permafrost under low pressure, temperature, and anoxic atmosphere has implications for Earth microbes on Mars. Proc Natl Acad Sci U S A 110:666–671.
30. Vishnivetskaya TA, Petrova MA, Urbance J, Ponder M, Moyer CL, Gilichinsky DA, Tiedje JM. 2006. Bacterial community in ancient Siberian permafrost as characterized by culture and culture-independent methods. Astrobiology 6:400–414.
31. Goordial J, Davila A, Greer CW, Cannam R, DiRuggiero J, McKay CP, Whyte LG. 2017. Comparative activity and functional ecology of permafrost soils and lithic niches in a hyper-arid polar desert. Environ Microbiol 19:443–458.
32. Wilhelm RC, Niederberger TD, Greer C, Whyte LG. 2011. Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian High Arctic. Can J Microbiol 57:303–315.
33. Mitzscherling J, Winkel M, Winterfeld M, Horn F, Yang S, Grigoriev MN, Wagner D, Overduin PP, Liebner S. 2017. The development of permafrost bacterial communities under submarine conditions. J Geophys Res Biogeosciences 122:1689–1704.
34. Vishnivetskaya T, Kathariou S, McGrath J, Gilichinsky D, Tiedje JM. 2000. Low-temperature recovery strategies for the isolation of bacteria from ancient permafrost sediments. Extremophiles 4:165–173.
35. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. 2018. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. mSystems 3:e00055-18.
36. Kraev G, Rivkina E, Vishnivetskaya T, Belonosov A, van Huissteden J, Kholodov A, Smirnov A, Kudryavtsev A, Teshebaeva K, Zamolodchikov D. 2019. Methane in gas shows from boreholes in epigenetic permafrost of Siberian Arctic. Geosciences 9:67.
37. Buongiorno J, Turner S, Webster G, Asai M, Shumaker AK, Roy T, Weightman A, Schippers A, Lloyd KG. 2017. Interlaboratory quantification of Bacteria and Archaea in

deeply buried sediments of the Baltic Sea (IODP Expedition 347). FEMS Microbiol Ecol 93:1–16.
38. Zhou Z, Liu Y, Lloyd KG, Pan J, Yang Y, Gu J-D, Li M. 2019. Genomic and transcriptomic insights into the ecology and metabolism of benthic archaeal cosmopolitan, Thermoprofundales (MBG-D archaea). ISME J 13:885–901.
39. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996.
40. Adam PS, Borrel G, Brochier-Armanet C, Gribaldo S. 2017. The growing tree of Archaea: new perspectives on their diversity, evolution and ecology. ISME J 11:2407.
41. Takai K, Gamo T, Tsunogai U, Nakayama N, Hirayama H, Nealson KH, Horikoshi K. 2004. Geochemical and microbiological evidence for a hydrogen-based, hyperthermophilic subsurface lithoautotrophic microbial ecosystem (HyperSLiME) beneath an active deep-sea hydrothermal field. Extremophiles 8:269–282.
42. Vetriani C, Reysenbach A-L, DorÃ© J. 1998. Recovery and phylogenetic analysis of archaeal rRNA sequences from continental shelf sediments. FEMS Microbiol Lett 161:83–88.
43. Lloyd KG, Schreiber L, Petersen DG, Kjeldsen KU, Lever MA, Steen AD, Stepanauskas R, Richter M, Kleindienst S, Lenk S, Schramm A, Jørgensen BB, Jorgensen BB. 2013. Predominant archaea in marine sediments degrade detrital proteins. Nature 496:215–218.
44. Lauer A, Sørensen K, Teske A. 2016. Phylogenetic characterization of marine benthic archaea in organic-poor sediments of the Eastern Equatorial Pacific Ocean (ODP Site 1225). Microorganisms 4:32.
45. Borrel G, Lehours A-C, Crouzet O, Jézéquel D, Rockne K, Kulczak A, Duffaud E, Joblin K, Fonty G. 2012. Stratification of archaea in the deep sediments of a freshwater meromictic lake: vertical shift from methanogenic to uncultured archaeal lineages. PLoS One 7:e43346.
46. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu W-T, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431–437.
47. Corvec S. 2018. Clinical and biological features of Cutibacterium (Formerly Propionibacterium) avidum, an underrecognized microorganism. Clin Microbiol Rev 31:e00064-17.
48. Yoon S-H, Ha S, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286.
49. Farag IF, Davis JP, Youssef NH, Elshahed MS. 2014. Global patterns of abundance, diversity and community structure of the Aminicenantes (Candidate Phylum OP8). PLoS

One 9:e92139.
50. Eloe-Fadrosh EA, Ivanova NN, Woyke T, Kyrpides NC. 2016. Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. Nat Microbiol 1:15032.
51. Jansson JK, Taş N. 2014. The microbial ecology of permafrost. Nat Rev Microbiol 12:414–425.
52. Xue Y, Jonassen I, Øvreås L, Taş N. 2020. Metagenome-assembled genome distribution and key functionality highlight importance of aerobic metabolism in Svalbard permafrost. FEMS Microbiol Ecol 96.
53. Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAFF, Hoelzle RD, Lamberton TO, McCalley CK, Hodgkins SB, Wilson RM, Purvine SO, Nicora CD, Li C, Frolking S, Chanton JP, Crill PM, Saleska SR, Rich VI, Tyson GW. 2018. Genome-centric view of carbon processing in thawing permafrost. Nature 560:49–54.
54. Argüelles JC. 2000. Physiological roles of trehalose in bacteria and yeasts: A comparative analysis. Arch Microbiol.
55. Bird JT, Tague ED, Zinke L, Schmidt JM, Steen AD, Reese B, Marshall IPG, Webster G, Weightman A, Castro HF, Campagna SR, Lloyd KG. 2019. Uncultured microbial phyla suggest mechanisms for multi-thousand-year subsistence in Baltic Sea sediments. MBio 10:e02376-18.
56. Diniz-Mendes L, Bernardes E, de Araujo PS, Panek AD, Paschoalin VMF. 1999. Preservation of frozen yeast cells by trehalose. Biotechnol Bioeng 65:572–578.
57. Brauer MJ, Yuan J, Bennett BD, Lu W, Kimball E, Botstein D, Rabinowitz JD. 2006. Conservation of the metabolomic response to starvation across two divergent microbes. Proc Natl Acad Sci U S A 103:19302–19307.
58. Kyryakov P, Beach A, Richard VR, Burstein MT, Leonov A, Levy S, Titorenko VI. 2012. Caloric restriction extends yeast chronological lifespan by altering a pattern of age-related changes in trehalose concentration. Front Physiol 3:256.
59. Koo H, Hakim J, Morrow C, Crowley M, Andersen D, Bej A. 2018. Metagenomic analysis of microbial community compositions and cold-responsive stress genes in selected antarctic lacustrine and soil ecosystems. Life 8:29.
60. Lee M, Gräwert T, Quitterer F, Rohdich F, Eppinger J, Eisenreich W, Bacher A, Groll M. 2010. Biosynthesis of isoprenoids: crystal structure of the [4Fe-4S] cluster protein IspG. J Mol Biol 404:600–10.
61. Smit A, Mushegian A. 2000. Biosynthesis of isoprenoids via mevalonate in Archaea: the lost pathway. Genome Res 10:1468–84.
62. Dibrova D V., Galperin MY, Mulkidjanian AY. 2014. Phylogenomic reconstruction of archaeal fatty acid metabolism. Environ Microbiol 16:907–918.
63. Voynova NE, Rios SE, Miziorko HM. 2004. Staphylococcus aureus Mevalonate Kinase: Isolation and characterization of an enzyme of the isoprenoid biosynthetic pathway. J Bacteriol 186:61–67.

64. Wilding EI, Brown JR, Bryant AP, Chalker AF, Holmes DJ, Ingraham KA, Iordanescu S, So CY, Rosenberg M, Gwynn MN. 2000. Identification, evolution, and essentiality of the mevalonate pathway for isopentenyl diphosphate biosynthesis in gram-positive cocci. J Bacteriol 182:4319–4327.
65. Vishnivetskaya TA, Buongiorno J, Bird J, Krivushin K, Spirina E V., Oshurkova V, Shcherbakova VA, Wilson G, Lloyd KG, Rivkina EM. 2018. Methanogens in the Antarctic Dry Valley permafrost. FEMS Microbiol Ecol 94.
66. Ayala-Del-Río HL, Chain PS, Grzymski JJ, Ponder MA, Ivanova N, Bergholz PW, Bartolo G Di, Hauser L, Land M, Bakermans C, Rodrigues D, Klappenbach J, Zarka D, Larimer F, Richardson P, Murray A, Thomashow M, Tiedje JM. 2010. The genome sequence of psychrobacter arcticus 273-4, a psychroactive siberian permafrost bacterium, reveals mechanisms for adaptation to low-temperature growth. Appl Environ Microbiol 76:2304–2312.
67. Tribelli PM, López NI. 2018. Reporting key features in cold-adapted bacteria. Life. MDPI AG.
68. Mackelprang R, Waldrop MP, DeAngelis KM, David MM, Chavarria KL, Blazewicz SJ, Rubin EM, Jansson JK. 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. Nature 480:368–371.
69. Chomvong K, Kordić V, Li X, Bauer S, Gillespie AE, Ha S-J, Oh E, Galazka JM, Jin Y-S, Cate JHD. 2014. Overcoming inefficient cellobiose fermentation by cellobiose phosphorylase in the presence of xylose. Biotechnol Biofuels 7:85.
70. Schimz KL, Broll B, John B. 1983. Cellobiose phosphorylase (EC 2.4.1.20) of Cellulomonas: occurrence, induction, and its role in cellobiose metabolism. Arch Microbiol 135:241–249.
71. Haft RJF, Gardner JG, Keating DH. 2012. Quantitative colorimetric measurement of cellulose degradation under microbial culture conditions. Appl Microbiol Biotechnol 94:223–229.
72. Krivushin K, Kondrashov F, Shmakova L, Tutukina M, Petrovskaya L, Rivkina E. 2015. Two metagenomes from late pleistocene northeast siberian permafrost. Genome Announc 3:e01380-14.
73. Cai C, Leu AO, Xie GJ, Guo J, Feng Y, Zhao JX, Tyson GW, Yuan Z, Hu S. 2018. A methanotrophic archaeon couples anaerobic oxidation of methane to Fe(III) reduction. ISME J 12:1929–1939.
74. Chakraborty A, Ruff SE, Dong X, Ellefson ED, Li C, Brooks JM, McBee J, Bernard BB, Hubert CRJ. 2020. Hydrocarbon seepage in the deep seabed links subsurface and seafloor biospheres. Proc Natl Acad Sci 117:11029 LP – 11037.
75. Zheng S, Ponder MA, Shih JYJ, Tiedje JM, Thomashow MF, Lubman DM. 2007. A proteomic analysis of Psychrobacter articus 273-4 adaptation to low temperature and salinity using a 2-D liquid mapping approach. Electrophoresis 28:467–488.

76. Horneck DA, Sullivan DM, Owen JS, Hart JM. 2011. Soil test interpretation guide. Oregon State Univ Ext Serv 1–12.
77. Perreault NN, Andersen DT, Pollard WH, Greer CW, Whyte LG. 2007. Characterization of the prokaryotic diversity in cold saline perennial springs of the Canadian high arctic. Appl Environ Microbiol 73:1532–1543.
78. Mykytczuk NCSS, Foote SJ, Omelon CR, Southam G, Greer CW, Whyte LG. 2013. Bacterial growth at −15 °C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. ISME J 7:1211–1226.
79. Tuorto SJ, Darias P, McGuinness LR, Panikov N, Zhang T, Häggblom MM, Kerkhof LJ. 2014. Bacterial genome replication at subzero temperatures in permafrost. ISME J 8:139–149.
80. Reeuwijk LP van. 2002. Procedures for Soil Analysis, 6th ed. International Soil Reference and Information Center, Wageningen, The Netherlands.
81. Alperin MJ, Reeburgh WS. 1985. Inhibition experiments on anaerobic methane oxidation. Appl Environ Microbiol 50:940–945.
82. Donahue DJ, Linick TW, Jull AJT. 1990. Isotope-ratio and background corrections for accelerator mass spectrometry radiocarbon measurements. Radiocarbon 32:135–142.
83. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569.
84. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–20.
85. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834.
86. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607.
87. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119.
88. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using

DIAMOND. Nat Methods 12:59–60.
89. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319.
90. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–55.
91. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9.
92. Markowitz VM, Chen IMA, Chu K, Szeto E, Palaniappan K, Grechkin Y, Ratner A, Jacob B, Pati A, Huntemann M, Liolios K, Pagani I, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG/M: The integrated metagenome data management and comparative analysis system. Nucleic Acids Res 40:123–129.
93. Tatusov RL, Galperin MY, Natale DA, Koonin E V. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36.
94. Graham ED, Heidelberg JF, Tully BJ. 2018. Potential for primary productivity in a globally-distributed bacterial phototroph. ISME J 12:1861–1866.
95. Hobbie JE, Daley RJ, Jasper S. 1977. Use of nuclepore filters for counting bacteria by fluorescence microscopy. Appl Environ Microbiol 33:1225–1228.
96. Stepanauskas R, Fergusson EA, Brown J, Poulton NJ, Tupper B, Labonté JM, Becraft ED, Brown JM, Pachiadaki MG, Povilaitis T, Thompson BP, Mascena CJ, Bellows WK, Lubys A. 2017. Improved genome recovery and integrated cell-size analyses of individual uncultured microbial cells and viral particles. Nat Commun 8:1–10.

Appendix

Figures and Tables

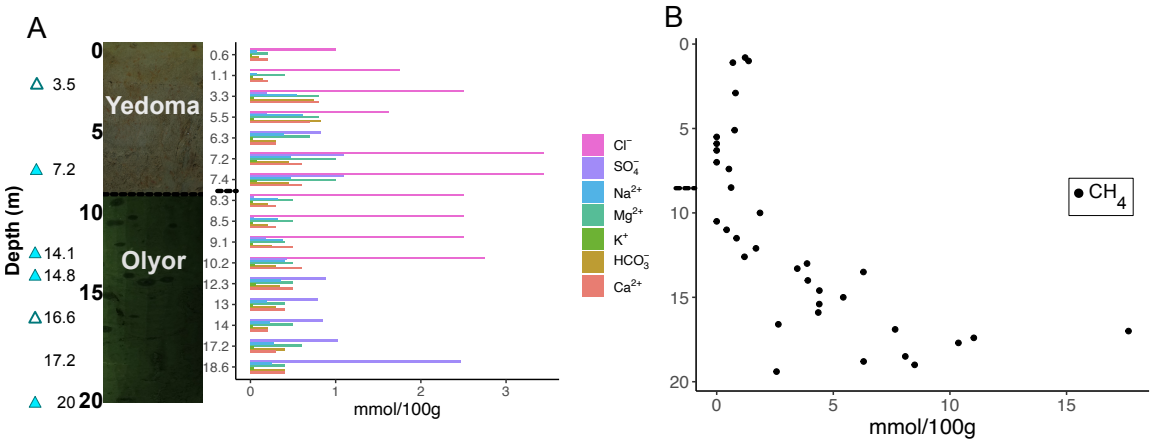


Figure 2-1 Downcore permafrost borehole characteristics

A) Photos of Yedoma and Olyor sediments and their relative depths, with triangles indicating depths that produced MAGs and filled triangles indicating depths with MAGs that are >80% complete and <10% contaminated. Concentrations of major ions are shown for each depth. Note that the depth intervals are discrete in panel B to spread the datapoints out evenly across the depths. B) Concentrations of methane for each depth. Other geochemical measurements including temperature, pH, and total carbon, are available in Supplemental Figure 2-1. Small, dashed line at 7.7 m shows the depth of the transition between Yedoma and Olyor layers.

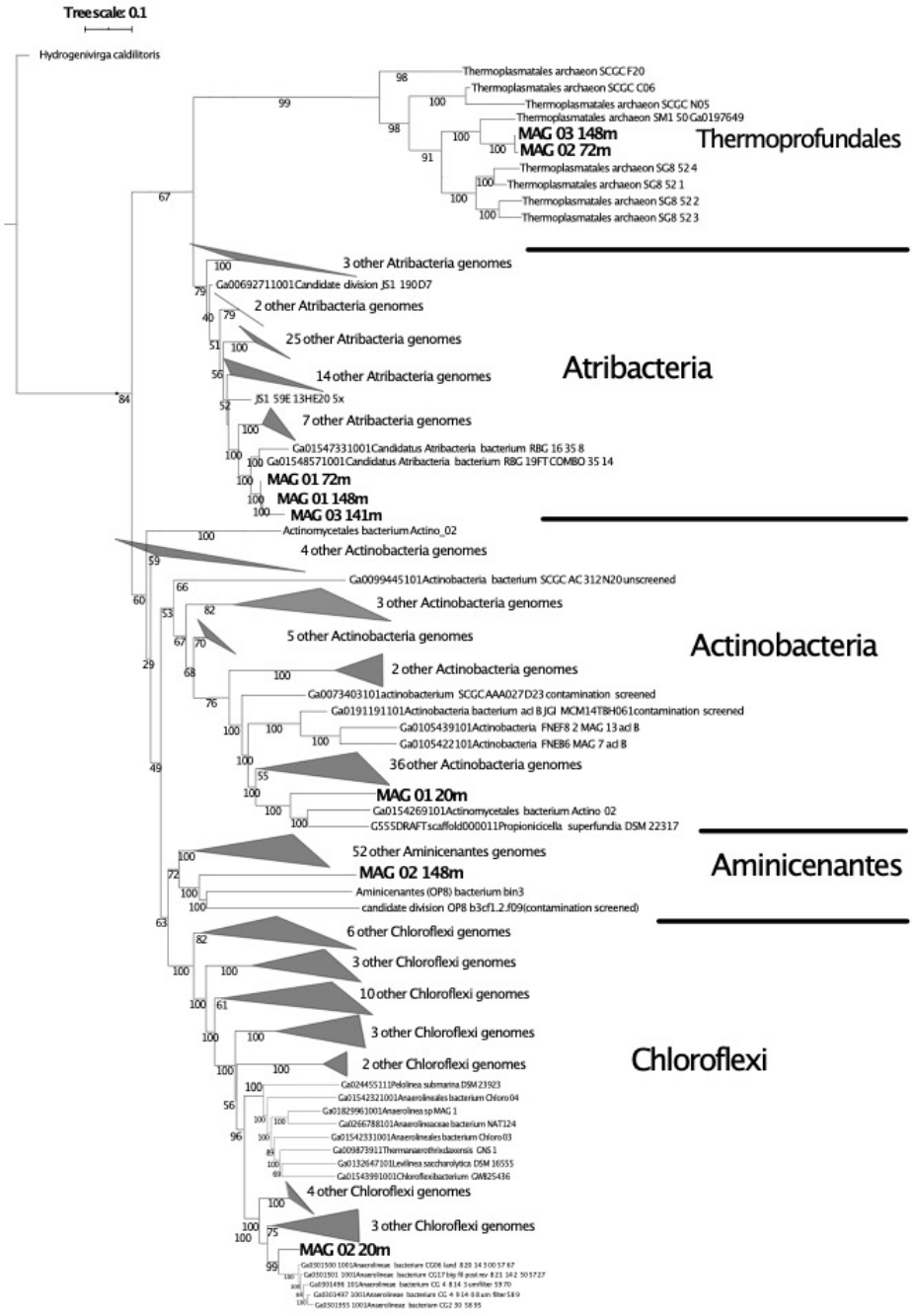


Figure 2-2: Phylogenetic tree of Siberian MAGs

Maximum Likelihood phylogenetic tree of 139 concatenated conserved genes for 8 Siberian MAGs and 230 total reference MAGs and genomes that were used in the comparative analysis. No reference MAGs or genomes from these groups were available from permafrost. Tree was visualized with the interactive Tree of Life (iTOL). A full list of shared genes can be found on Supplemental Datafile 6.

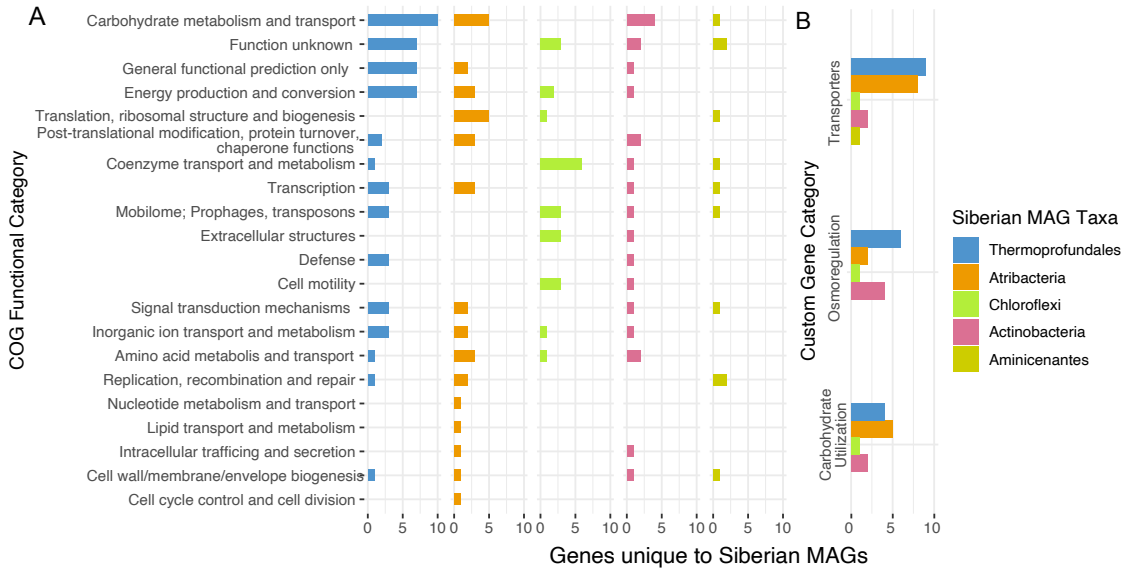


Figure 2-3: Unique genes to the Siberian MAGs

A) Number of genes in each COG Category present in the Siberian MAG groups that are absent in the reference genomes of the same group (except for Atribacteria, where the genes were in less than a third of the reference genomes). Exact names of genes can be found in the section of each of the Siberian MAGs throughout the main text and supplemental datafiles 1-5. B) Bold functional categories encompass genes from different COG categories but are associated with transport, osmoregulation, or carbohydrate utilization. Full list of genes can be found in Supplemental Datafile 6.

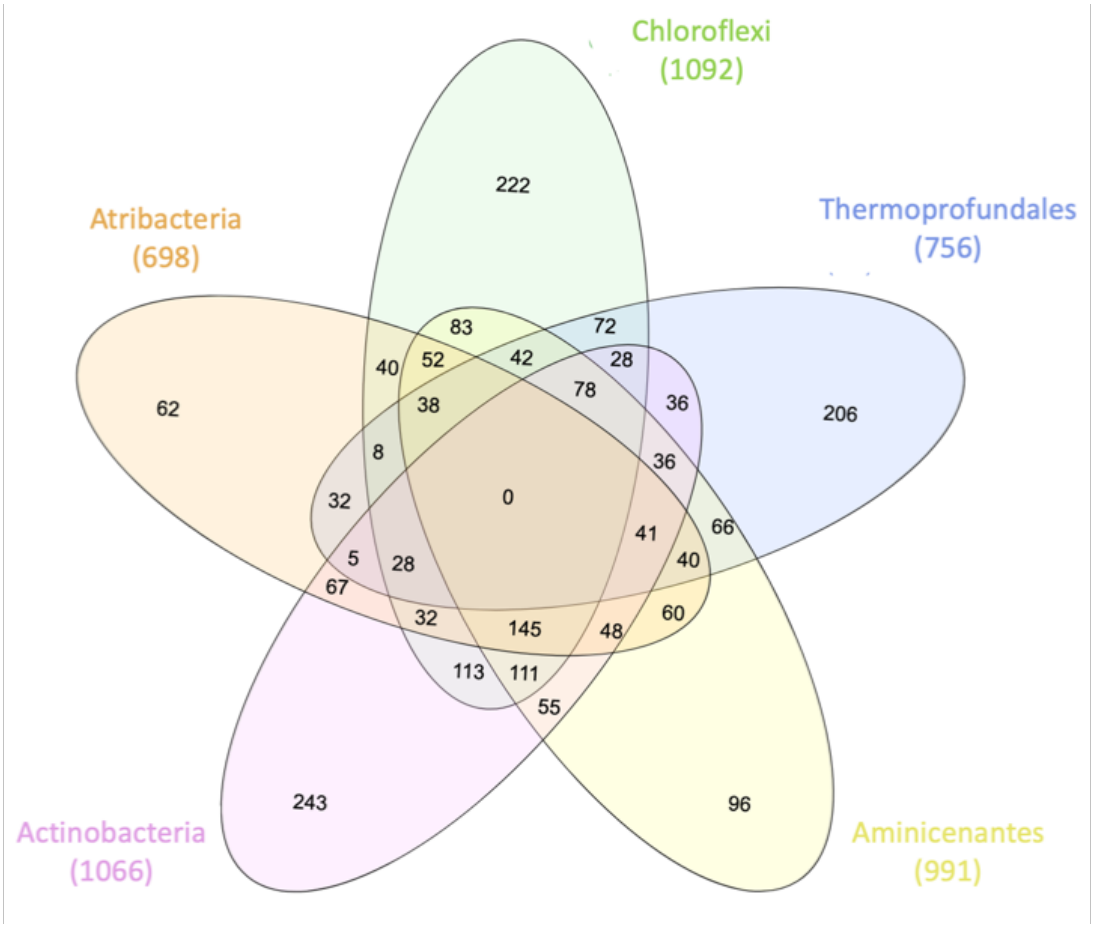


Figure 2-4: Number of unique genes shared between different Siberian MAG groups
The numbers of in parentheses are total number of genes present in all MAGs for that group. The complete list of genes in each MAG group and each union can be found in Supplemental Datafile 6.

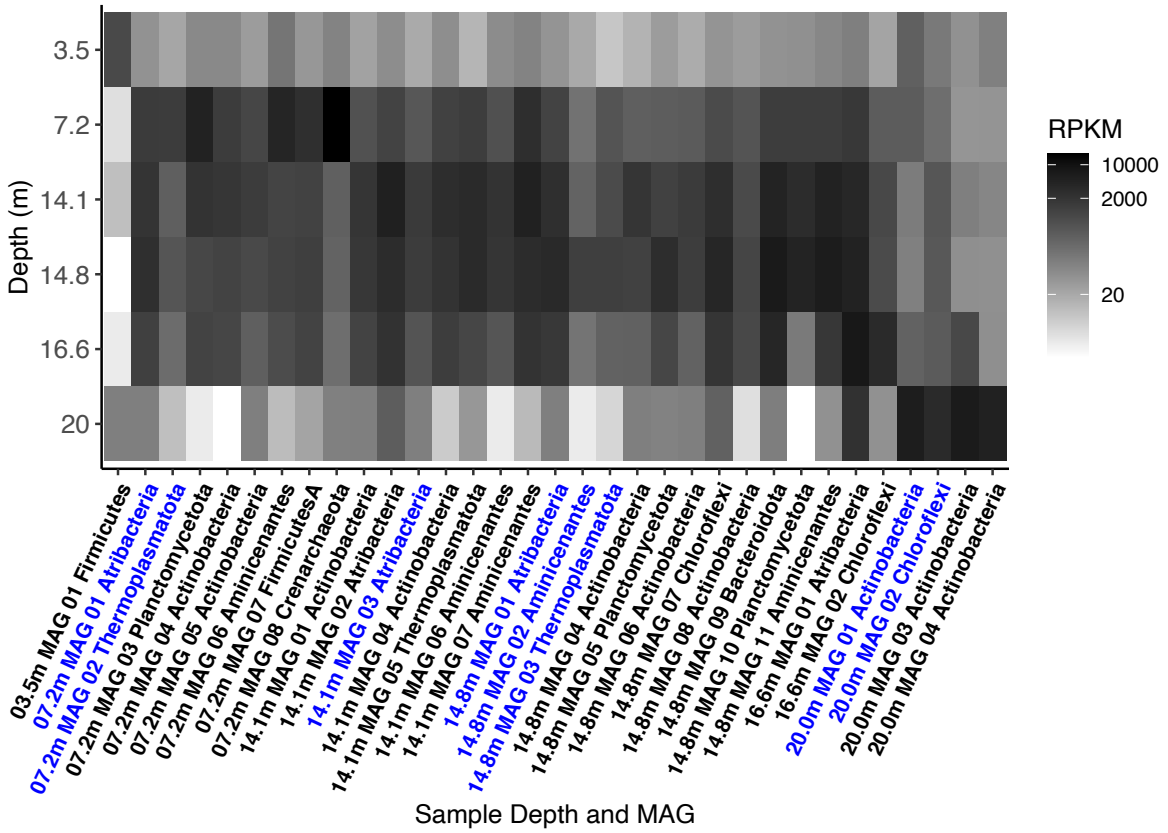


Figure 2-5: Metagenomic read recruitment for each MAG within each sample depth
Abundance is reported in 'reads per kilobase per million mapped reads' (RPKM). MAGs in blue text are MAGs analyzed in this study (≥80% completeness and <10% contamination), black text MAGs came from these samples but are not analyzed extensively in this study (but are available at the accession number provided in the Data Availability section). Calculations can be found in Supplemental Datafile 7.

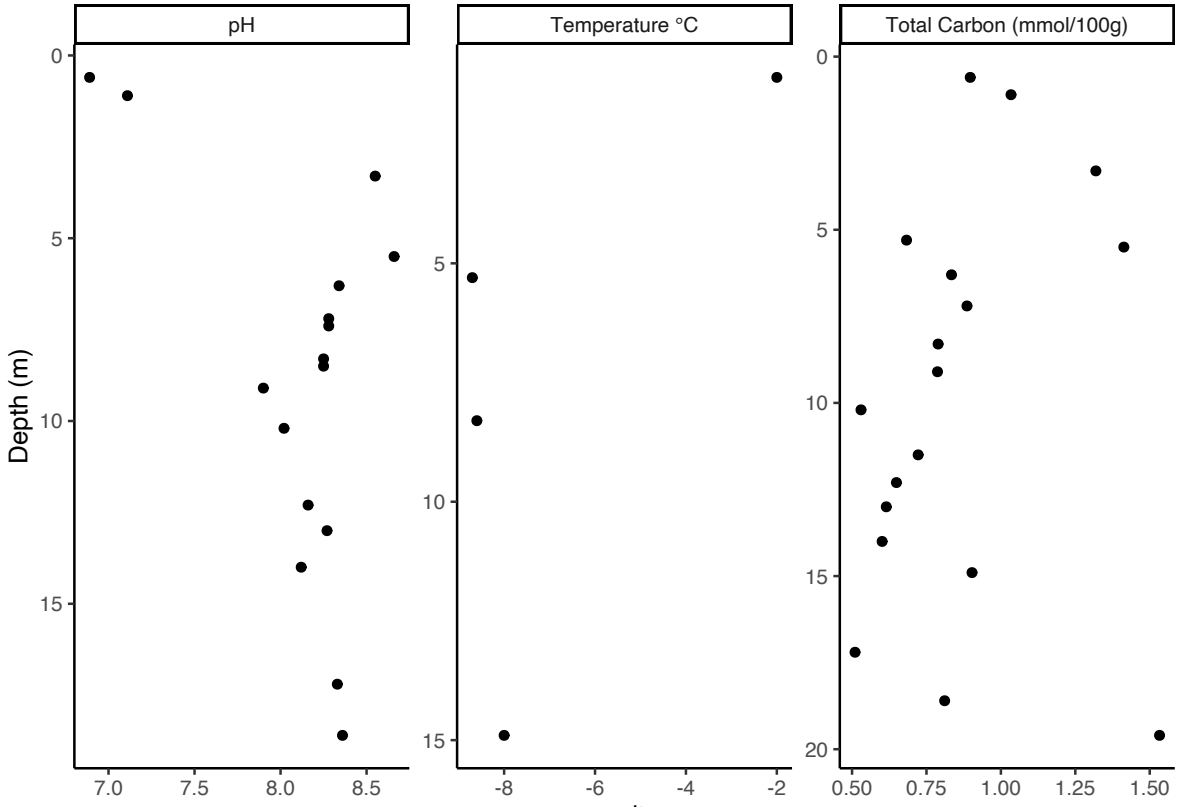
Table 2-1: Metagenome details from the seven sampling depths.
Sample depth 17.2m did not yield enough data to curate a metagenome.

Sample Depth (m)	Total unassembled reads	Total assembled reads	Total assembled contigs	N50	Total Number of MAGs	Percent of contigs binned
3.5	14,187,214	13,195,548	115	2842	1	91
7.2	9,621,340	9,045,678	7556	4323	8	94
14.1	11,513,548	10,956,796	4513	4527	7	93
14.8	13,544,854	12,728,740	7098	4625	11	93
16.6	5,698,300	5,115,016	900	2871	2	99
17.2	420	N/A	N/A	N/A	0	0
20	13,338,036	11,783,820	3058	4353	4	96

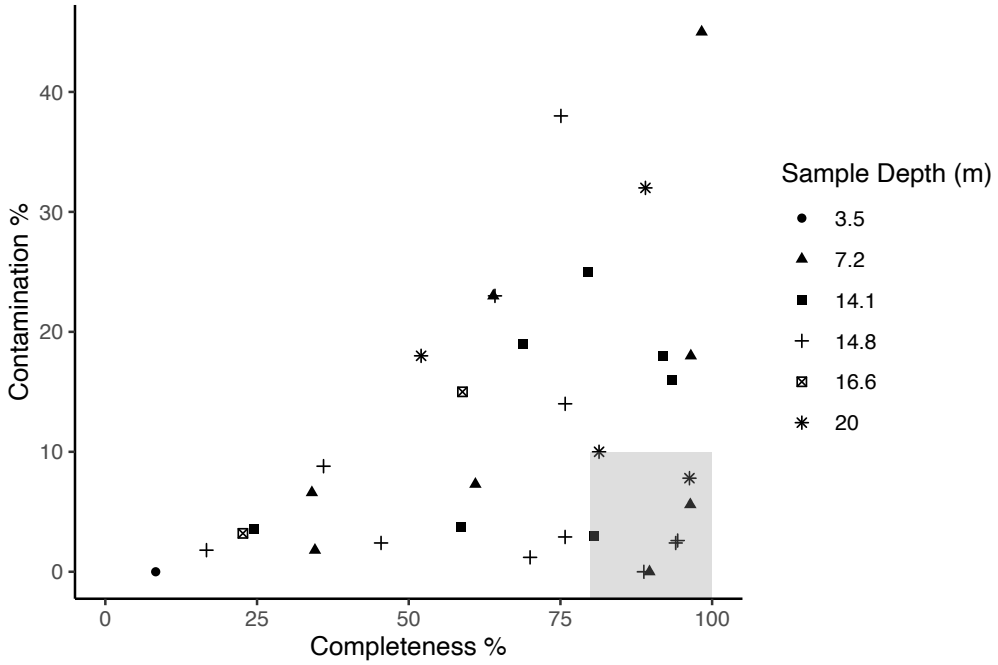
Supplemental Material

For supplemental datafiles referenced within this chapter, please refer to online version of this publication. Online version contains all supplemental materials to maintain brevity of this document https://doi.org/10.1128/AEM.00972-21

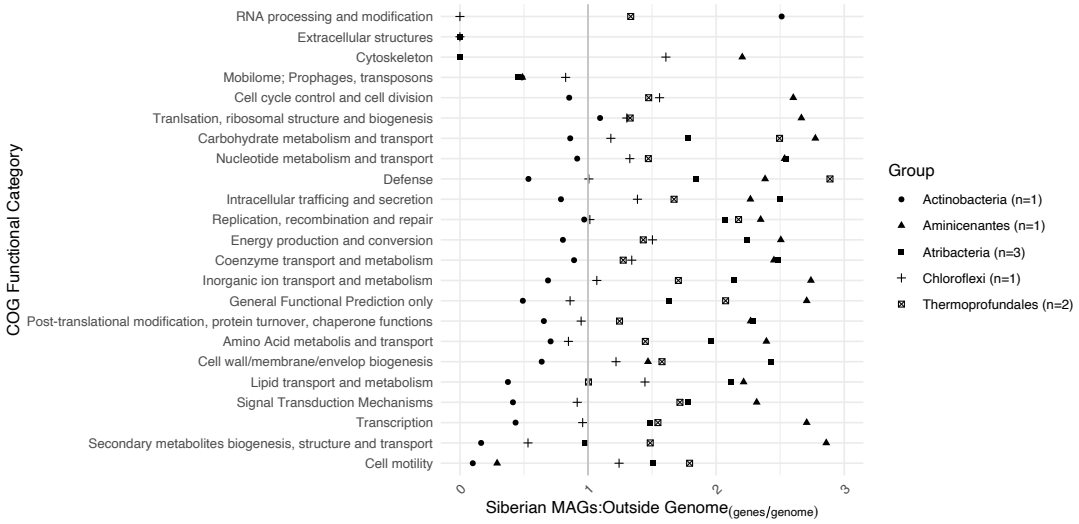
Supplemental Figures and Tables



Supplemental Figure 2-1: Extra field-tested parameters from borehole AL3-15



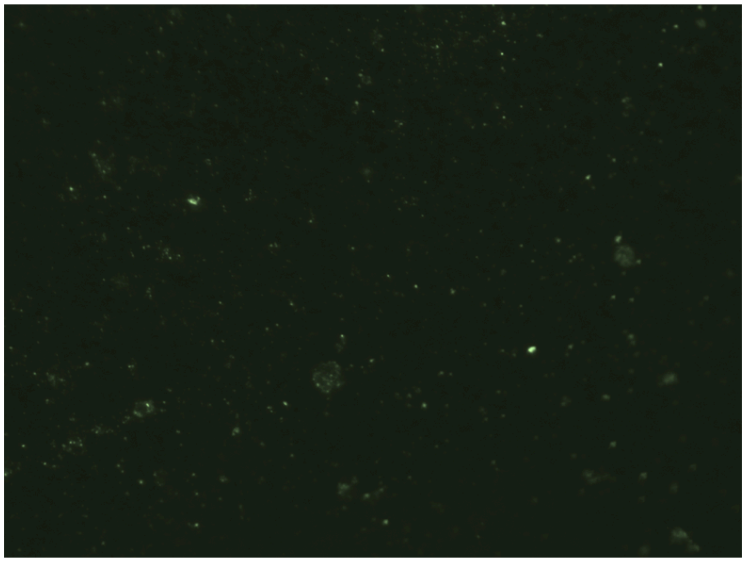
Supplemental Figure 2-2: Siberian permafrost MAG completeness and contamination, shown by the depth of the permafrost layer from which they originated. The shaded region denotes MAGs that were focused on for this study. (≥80% completeness, <10% contamination)



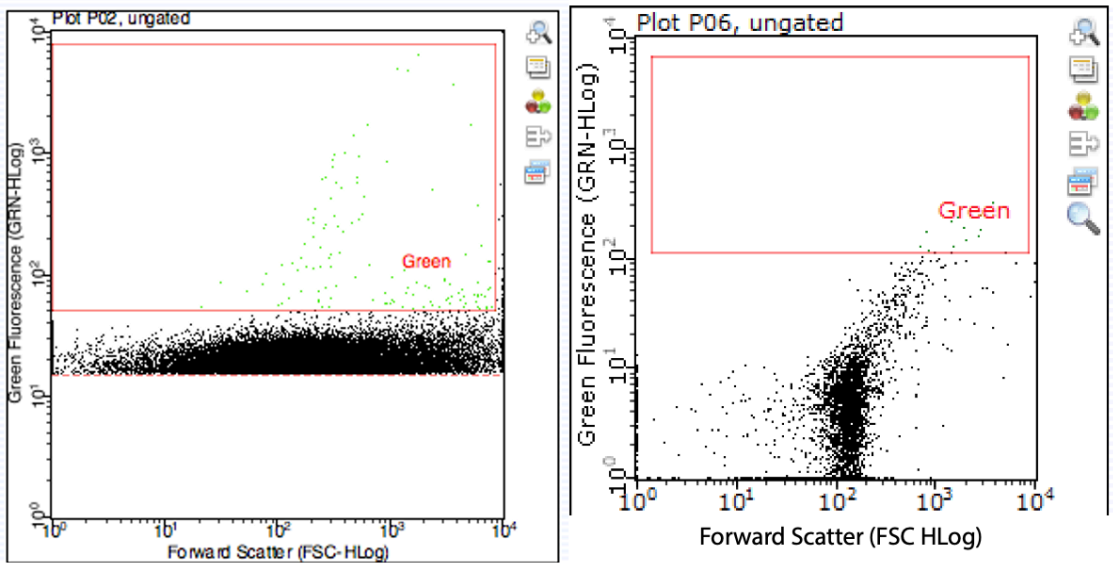
Supplemental Figure 2-3: Ratio of genes per MAG/genome for each group.

Genes were summed for each COG Category and divided by the number of MAGs/genomes for each of the eight Siberian MAGs and for the reference genomes. The values shown are the ratio of these genes per MAG/Genome for the eight Siberian MAGs to the ratio of these genes per MAG/Genome for the reference genomes of that group.

A



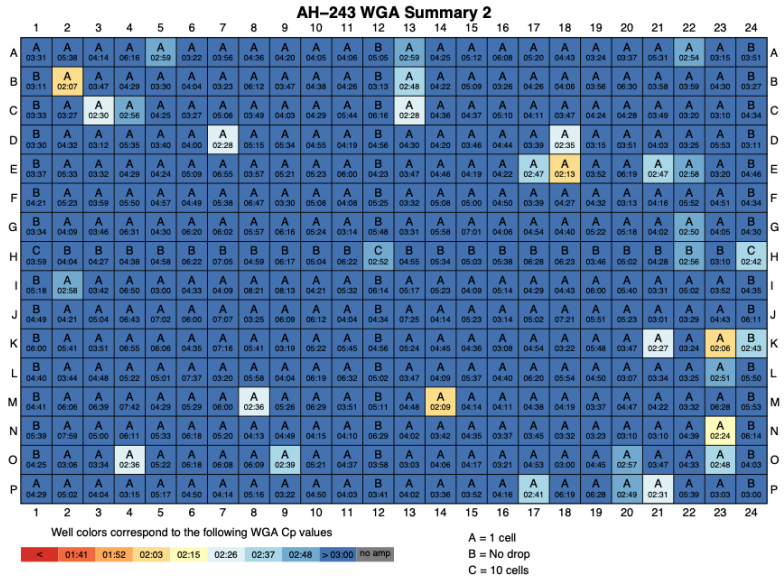
B C



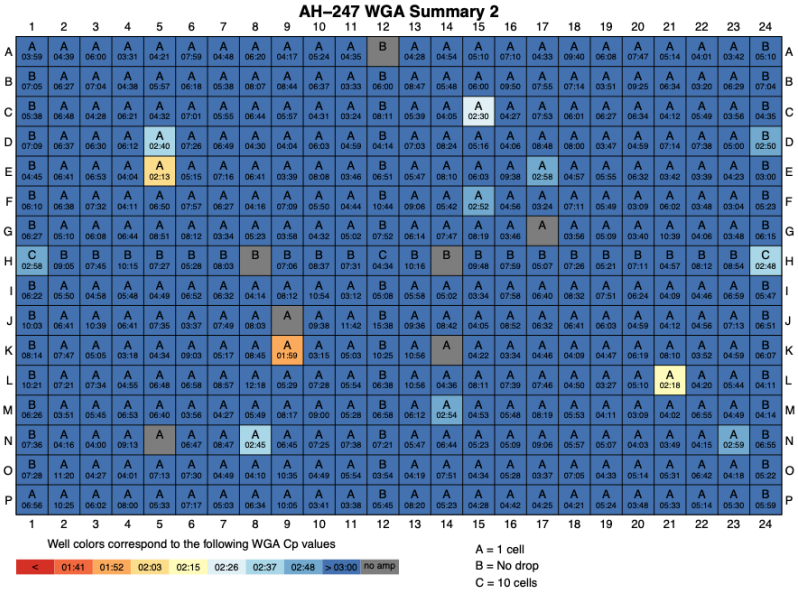
Supplemental Figure 2-4: Cell staining and flow cytometer.

A) Cells from the 5.6m sample, stained with SYBR Gold and viewed under Zeiss Imager M2 1000X magnification. B) Guava easyCyte 12HT Benchtop Flow Cytometer output with a gated size and fluorescence (green region, outlined in the red box) compared to the background noise (black) on 3.5m sample. C) Similar as panel B but with 7.2m sample.

A

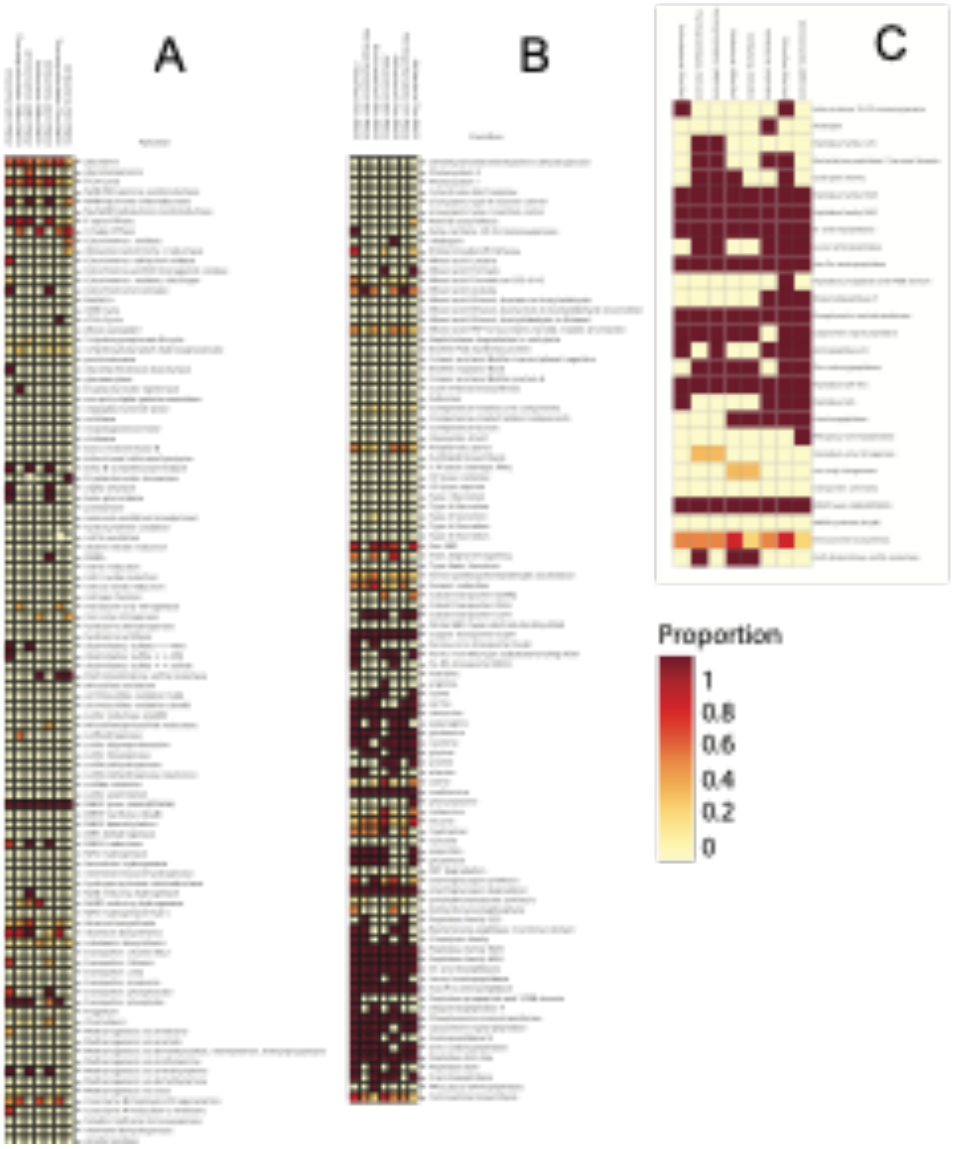


B



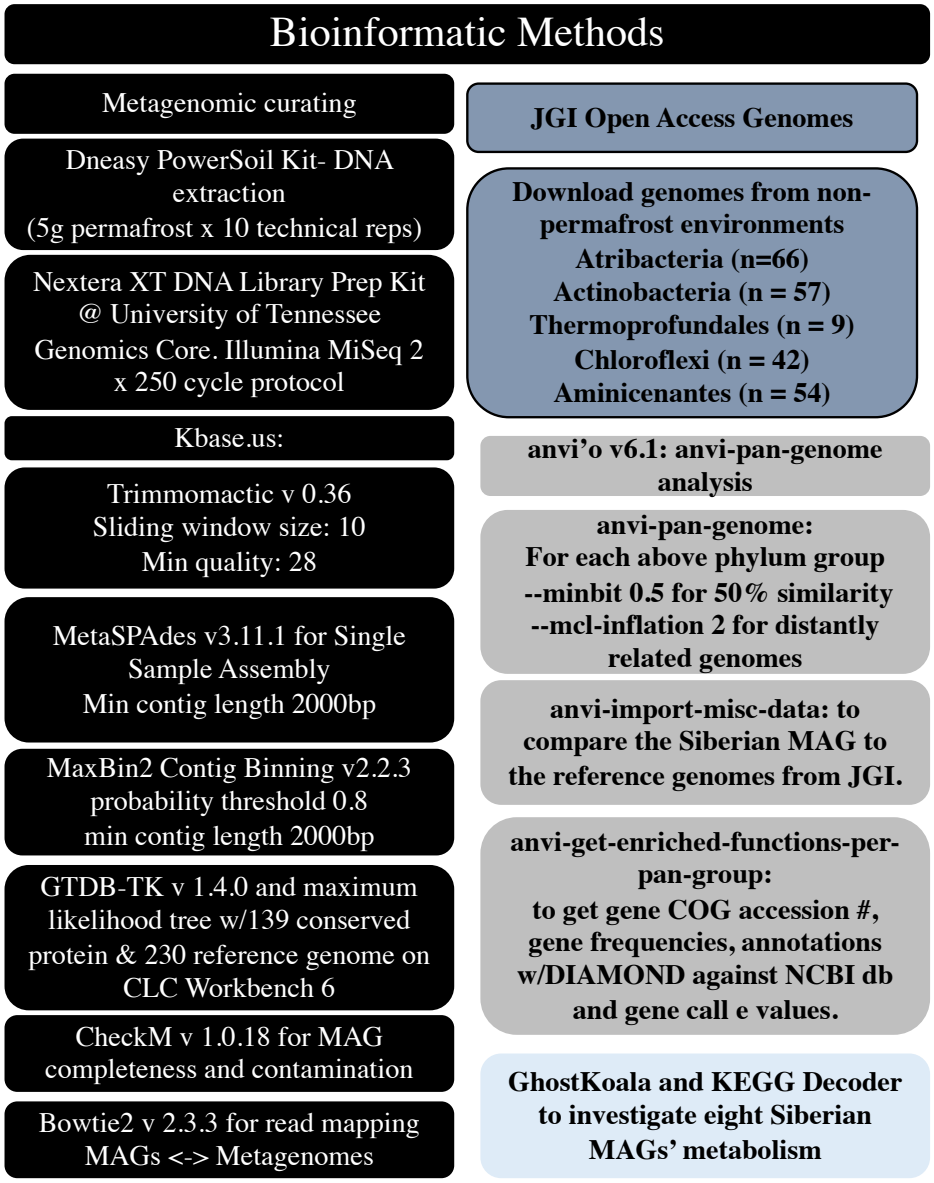
Supplemental Figure 2-5: Single cell whole genome amplification (WGA-X)

Soil from Borehole AL3-15 underwent WGA-X from Bigelow with SAG Generation 2 single cell sorting servicing. AH 243 plate (A) was created from soil 5.6m below the surface, the wells that were taken for further library preparation were B2, E18, K23, M14, and N23. AH247 (B) was from 7.2m below the surface of borehole AL3-15, the wells that were used for library prep were E5, K9, L21, N8, and C15.



Supplemental Figure 2-6: Functional pathways for each of the eight MAGs.

Code used to make this can be found on https://github.com/bjtully/BioData/tree/master/KEGGDecoder. A-B) Standard KEGG-Decoder output cut in half for readability. C) Output from KEGG-Expander code. Color saturation indicates the percentage of genes present from that metabolic pathway. Where single genes are listed instead of pathways (e.g., Leucyl aminopeptidase) the possible values are either 0% or 100%.



Supplemental Figure 2-7: Pictoral representation of the methods used to conduct this study, parameters included in each step.

Black coloration highlights metagenome processing, open access genomes taken from JGI Database is highlighted in blue. Grey color highlights pan genomic analysis done with Anvi'o v 6.1. Light blue highlights KEGG Decoder for metabolism. 'db' stands for database.

Supplemental Table 2-1: Age determination of inorganic and organic fractions using 14C dating from upper Yedoma Siberian permafrost samples.

The inorganic carbon fraction was released via reaction with phosphoric acid. Two different fractions of organic carbon (presumably "younger" and "older") were sequentially released by combustion at 400°C and 800˚C, respectively.

Sample Depth (m)	Treatment	C yield (mg)	$\delta^{13}C$‰ VPDB	^{14}C age BP (yr)
2.9-3	Inorganic C	0.73	-6.0	21,760±120
2.9-3	400˚C	1.01	-26.7	41,700±1,400
2.9-3	800˚C	1.41	-28.8	38,590±980
3.5	Inorganic C	1.64	-3.7	26,260±210
3.5	400˚C	1.21	-26.9	20,158±99
3.5	800˚C	2.34	-27.5	18,228±78
5.6	Inorganic C	2.36	-4.6	33,900±550
5.6	400˚C	0.97	-27.8	38,020±910
5.6	800˚C	0.27	-24.3	29,040±420

Supplemental Table 2-2: All metagenome assembled genomes (MAG's) completeness and contamination scores as reported by CheckM. GTDB-Tk taxonomy was used. Missing information is due to unavailable assignments. MAGs analyzed in the main text were those with ≥80% completeness and ≤10% contamination.

Sample Depth and MAG	Completeness %	Contamination %	Domain	Phylum	Class
03.5m MAG 01	8	0	Bacteria	Firmicutes	Bacilli
07.2m MAG 01	90	0	Bacteria	Caldiatribacteriota	JS1
07.2m MAG 02	96	6	Archaea	Thermoplasmatota	E2
07.2m MAG 03	96	18	Bacteria	Planctomycetota	
07.2m MAG 04	64	23	Bacteria	Actinobacteriota	UBA1414
07.2m MAG 05	35	2	Bacteria	Actinobacteriota	UBA1414
07.2m MAG 06	98	44	Bacteria	Aminicenantes	Aminicenantia
07.2m MAG 07	61	7	Bacteria	Firmicutes	Mahellia
07.2m MAG 08	34	7	Archaea	Bathyarchaeota	
14.1m MAG 01	25	4	Bacteria	Actinobacteriota	UBA1414
14.1m MAG 02	80	25	Bacteria	Caldiatribacteriota	JS1
14.1m MAG 03	81	3	Bacteria	Caldiatribacteriota	
14.1m MAG 04	59	4	Bacteria	Actinobacteriota	UBA1414
14.1m MAG 05	93	16	Archaea	Thermoplasmatota	E2
14.1m MAG 06	92	18	Bacteria	Aminicenantes	Aminicenantia
14.1m MAG 07	69	19	Bacteria	Aminicenantes	Luteitaleia
14.8m MAG 01	89	0	Bacteria	Caldiatribacteriota	JS1

Supplemental Table 2-2 continued

Sample Depth and MAG	Completeness %	Contamination %	Domain	Phylum	Class
14.8m MAG 02	94	3	Bacteria	Aminicenantes	
14.8m MAG 03	94	2	Archaea	Thermoplasmatota	E2
14.8m MAG 04	70	1	Bacteria	Actinobacteriota	UBA1414
14.8m MAG 05	36	9	Bacteria	Planctomycetota	Phycisphaerae
14.8m MAG 06	17	2	Bacteria	Actinobacteriota	UBA1414
14.8m MAG 07	76	14	Bacteria	Chloroflexota	Anaerolineae
14.8m MAG 08	45	2	Bacteria	Actinobacteriota	UBA1414
14.8m MAG 09	75	38	Archaea	Crenarchaeota	Bathyarchaeia
14.8m MAG 10	75	3	Bacteria	Planctomycetota	FEN-1346
14.8m MAG 11	64	23	Bacteria	Acidobacteriota	Vicinamibacteria
16.6m MAG 01	59	15	Bacteria	Caldiatribacteriota	JS1
16.6m MAG 2	23	3	Bacteria	Chloroflexota	UBA2235
20m MAG 01	96	8	Bacteria	Actinobacteriota	Actinobacteria
20m MAG 02	81	10	Bacteria	Chloroflexota	Anaerolineae
20m MAG 03	89	32	Bacteria	Actinobacteriota	Thermoleophilia
20m MAG 04	52	18	Bacteria	Actinobacteriota	Actinobacteria

Supplemental Table 2-3: List of genes that made up the created COG categories in Figure 2-3B.

Transporters	GeneName	Taxa	COG Letter	COG Category
	ABC-type transport system involved in Fe-S cluster assembly, permease and ATPase components	Actinobacteria	O	Post-translational modification, protein turnover, chaperone functions
	Na+-translocating ferredoxin:NAD+ oxidoreductase RNF, RnfG subunit	Actinobacteria	C	Energy production and conversion
	Succinate-acetate transporter protein	Aminicenantes	S	Function Unknown
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	T	Signal Transduction Mechanisms
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	E	Amino Acid metabolis and transport
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	T	Signal Transduction Mechanisms
	ABC-type amino acid transport/signal transduction system, periplasmic component/domain	Atribacteria	E	Amino Acid metabolis and transport
	Cation transport ATPase	Atribacteria	P	Inorganic ion transport and metabolism
	Lipopolysaccharide export system protein LptC	Atribacteria	M	Cell wall/membrane/envelop biogenesis
	Na+/H+ antiporter NhaD or related arsenite permease	Atribacteria	P	Inorganic ion transport and metabolism
	Predicted ABC-type sugar transport system, permease component	Atribacteria	R	General Functional Prediction only
	Na+/alanine symporter	Chloroflexi	E	Amino Acid metabolis and transport
	ABC-type glycerol-3-phosphate transport system, periplasmic component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type glycerol-3-phosphate transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, ATPase component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	Ca2+/H+ antiporter	Thermoprofundales	P	Inorganic ion transport and metabolism
	Na+/H+ antiporter NhaD or related arsenite permease	Thermoprofundales	P	Inorganic ion transport and metabolism
	Putative Ca2+/H+ antiporter, TMEM165/GDT1 family	Thermoprofundales	R	General Functional Prediction only

Supplemental Table 2-3 continued

	GeneName	Taxa	COG Letter	COG Category
	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	TctA family transporter	Thermoprofundales	R	General Functional Prediction only
Osmoregulation				
	Uncharacterized membrane protein YagU, involved in acid resistance, DUF1440 family	Actinobacteria	S	Function Unknown
	Outer membrane lipoprotein SlyB	Actinobacteria	M	Cell wall/membrane/envelop biogenesis
	Plasmid replication initiator protein	Actinobacteria	X	Mobilome; Prophages, transposons
	Na+-translocating ferredoxin:NAD+ oxidoreductase RNF, RnfG subunit	Actinobacteria	C	Energy production and conversion
	Cation transport ATPase	Atribacteria	P	Inorganic ion transport and metabolism
	Na+/H+ antiporter NhaD or related arsenite permease	Atribacteria	P	Inorganic ion transport and metabolism
	Na+/alanine symporter	Chloroflexi	E	Amino Acid metabolis and transport
	Uncharacterized membrane protein DedA, SNARE-associated domain	Thermoprofundales	S	Function Unknown
	Predicted signal transduction protein containing a membrane domain, an EAL and a GGDEF domain	Thermoprofundales	T	Signal Transduction Mechanisms
	Na+/H+ antiporter NhaD or related arsenite permease	Thermoprofundales	P	Inorganic ion transport and metabolism
	Putative Ca2+/H+ antiporter, TMEM165/GDT1 family	Thermoprofundales	R	General Functional Prediction only
	Ca2+/H+ antiporter	Thermoprofundales	C	Energy production and conversion
	Ca2+/H+ antiporter	Thermoprofundales	P	Inorganic ion transport and metabolism
Sugar Utilization				
	Phosphotransferase system cellobiose-specific component IIA	Actinobacteria	G	Carbohydrate metabolism and transport
	Phosphotransferase system cellobiose-specific component IIC	Actinobacteria	G	Carbohydrate metabolism and transport
	Cellobiose phosphorylase	Atribacteria	G	Carbohydrate metabolism and transport
	D-lyxose ketol-isomerase	Atribacteria	G	Carbohydrate metabolism and transport
	Rhamnose utilisation protein RhaD, predicted bifunctional aldolase and dehydrogenase	Atribacteria	G	Carbohydrate metabolism and transport

Supplemental Table 2-3 continued

Transporters	GeneName	Taxa	COG Letter	COG Category
	Trehalose-6-phosphate synthase	Atribacteria	G	Carbohydrate metabolism and transport
	Predicted ABC-type sugar transport system, permease component	Atribacteria	R	General Functional Prediction only
	Hydroxyethylthiazole kinase, sugar kinase family	Chloroflexi	H	Coenzyme transport and metabolism
	Cellobiose phosphorylase	Thermoprofundales	G	Carbohydrate metabolism and transport
	Ribose/xylose/arabinose/galactoside ABC-type transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, permease component	Thermoprofundales	G	Carbohydrate metabolism and transport
	ABC-type sugar transport system, ATPase component	Thermoprofundales	G	Carbohydrate metabolism and transport

CHAPTER 3 MICROBES IN ACTIVE LAYER NY ÅLESUND, SVALBARD (79°N) PERMAFROST SHOW HETEROTROPHIC AND AUTOTROPHIC METABOLISMS

A version of this chapter is awaiting publication at Frontiers of Microbiology, titled:
Permafrost active layer microbes from Ny Ålesund, Svalbard (79°N) show autotrophic and heterotrophic metabolisms with diverse carbon-degrading enzymes, by
Katie Sipes, Raegan Paul, Aubrey Fine, Peibo Li, Renxing Liang, Julia Boike, Tullis C. Onstott, Tatiana A. Vishnivetskaya, Sean Schaeffer, Karen G. Lloyd

I contributed to this novel research by creating the sampling plan and experimentation, preforming field sampling, DNA extraction, all sequencing, metagenome analysis, soil enzyme assay, soil extractable compounds, 16S rRNA gene and metagenome *in silico* analyses, mentoring author's RP and PL in all of their contributions to this chapter, and writing and editing the manuscript.

Reagan Paul performed the following relating to cultures; maintenance, enzyme assays, DNA extraction, 16S rRNA gene PCR and investigation with gel electrophoresis. Contribution to the original manuscript submission resulted in coauthorship.

Aubrey Fine contributed by assisting with enzyme assays and soil geochemical results.

Peibo Li performed soil preparation for elemental analysis.

Abstract

The active layer of permafrost in Ny Ålesund, Svalbard (79°N) around the Bayelva River in the Leirhaugen glacier moraine is measured as a small net carbon sink at the brink of becoming a carbon source. In many permafrost dominating ecosystems, microbes in the active layers have been shown to drive organic matter degradation and greenhouse gas production, creating positive feedback on climate change. However, the microbial metabolisms linking the environmental geochemical processes and the populations that perform them have not been fully characterized. In this paper, we present geochemical, enzymatic, and isotopic data paired with ten *Pseudomonas* sp. cultures and metagenomic libraries of two active layer soil cores (BPF1 and BPF2) from Ny Ålesund, Svalbard, (79°N). Relative to BPF1, BPF2 had statistically higher C/N ratios (15 ± 1 for BPF1 vs. 29 ± 10 for BPF2; n=30, $p < 10^{-5}$), statistically lower organic carbon (2% ± 0.6% for BPF1 vs 1.6% ± 0.4% for BPF2, $p < 0.02$), statistically lower nitrogen (0.1% ± 0.03% for BPF1 vs. 0.07% ± 0.02% for BPF2, $p < 10^{-6}$). The $\delta^{13}C$ values for inorganic carbon did not correlate with those of organic carbon in BPF2, suggesting lower heterotrophic respiration. An increase in the $\delta^{13}C$ of inorganic carbon with depth either reflects an autotrophic signal or mixing between a heterotrophic source at the surface and a lithotrophic source at depth. Potential enzyme activity of xylosidase and N-acetyl-β-D-glucosaminidase increases twofold at 15°C, relative to 25°C, indicating cold-adaptation in the cultures and bulk soil. Potential enzyme activity of leucine aminopeptidase across soils and cultures was two orders of magnitude higher than other tested enzymes, implying organisms use leucine as a nitrogen and carbon source in this nutrient limited environment. Besides demonstrating large variability in carbon compositions of permafrost active layer soils only ~84 meters apart, results suggest that the Svalbard active layer microbes are often limited by organic carbon or nitrogen availability and have adaptations to the current environment, as well as metabolic flexibility to adapt to the warming climate.

Introduction

Temperatures in the Arctic are increasing faster than they are at lower latitudes (Cohen et al., 2014, Pörtner et al 2019). The permafrost of Svalbard, in particular, is known as “warm

permafrost", since it is close to 0°C, making it very sensitive to warming (3). Globally, permafrost and active layer soils are estimated to contain carbon stocks that are twice as large as the current atmospheric carbon pool (4). Models have estimated about 1600 Pg (petagrams, 10^{15}) of carbon in permafrost regions (5) with an estimated 195 Pg of carbon projected to be released in the form of gaseous carbon compounds by the year 2100, which would increase global temperatures an additional 0.03°C to 0.23°C (6, 7). Microbial activity within soils, especially in the active layer that thaws every summer, is an important driver of nutrient and carbon cycling (8). Future climatic conditions are likely to increase the availability of carbon sources for microbial decomposition since temperature, water availability, and the rate of microbial carbon degradation in permafrost are related (5, 7, 9). The permafrost of the high Svalbard Arctic (79°N) is currently a small net annual sink for CO_2 (10). Nitrogen is often a limiting reagent for plant growth and microbial metabolism in the Arctic, and is derived from the decomposition of active layer organic matter (6, 11). Pathways for carbon degradation and the relationship of carbon degradation to nitrogen limitation have not been fully characterized in this region, nor have they been coupled to extracellular enzyme assays, isolations, and isotopic compositions of carbon and nitrogen. Due to increases in temperature and microbial activity, this warm permafrost location could be at the brink of becoming a net source of greenhouse gasses such as CH_4 and CO_2. Here, we combine soil carbon and nitrogen analysis, extracellular enzymes assays, isolate activities, and metagenomes to gain a broader view of current carbon degradation activities, and their relationship to nitrogen cycling, in Svalbard active layer soils.

The permafrost at the site near the Bayelva River in Ny Ålesund, Svalbard (79°N), has been continuously monitored for physical characteristics of the soil and snow since 1998, and reports a yearly average permafrost temperature of -2.5°C (Boike et al., 2018). Microbial interactions with the geochemical processes of the active layer have not yet been characterized at this site. However, analysis of nearby snow cover and fjord sediment found *Alphaproteobacteria*, *Betaproteobacteria*, *Gammaproteobacteria*, *Firmicutes* and *Actinobacteria* to be the most common phyla (12–14). Annual variation in active layer soil microbial communities in Adventdalen, Svalbard (78°N), a site that is ~615 km away from our study location, was observed to be largely driven by organic matter availability or sunlight (15). Metagenomic and gas flux analyses at Adventdalen showed that the active layer microbial

community is involved in many different carbohydrate degradation pathways (16, 17). While total DNA population studies allow for a broad characterization of the microbial community, culturing environmental organisms can provide direct evidence of the interactions between organisms and the environment. Arctic conditions are difficult to replicate in a laboratory setting due to low nutrient availability and seasonal freezing. Commonly, cultures have been grown in a wide range of temperatures (+4 to +30°C) to parallel Arctic seasonal variations and to investigate ubiquitous organisms (18–20). One commonly used medium for cultivating Arctic soil isolates is R2A (18, 20–22) which mimics the low-energy conditions of the active layer and has been found to be best suited for the isolation of slow growing oligotrophic organisms (18, 23).

We investigated the microbial influence on two active layer geochemical profiles by combining cultured isolates with metagenomic inferences, extracellular enzyme assays, and soil geochemistry such as carbon and nitrogen content and stable isotope ratios. This multifaceted approach showed heterotrophic metabolism dominating the location with higher labile carbon. Autotrophic signatures are more prevalent in the site with higher inorganic carbon content and a higher C/N ratio. Potential rates of extracellular enzymes and the gene counts of the enzymes were compared between measurements made on bulk soil vs. ten *Pseudomonas* sp. isolates. Enzymatic analyses over a range of temperatures displayed higher activity in colder temperatures in both the bulk soil and cultured isolates. Understanding the pathways of carbon degradation in natural microbial communities and cultured isolates from the active layer soils is important for determining how these communities will degrade natural organic matter as more of it becomes available due to thaw.

Materials and Methods

Field sampling: In April 2018, soil core samples were drilled from completely frozen active layer permafrost in Ny Ålesund, Svalbard (Fig. 1). These two drill sites were near Bayelva Monitoring Site (Boike et al. 2018). Bayelva Permafrost site 1 (BPF1) (N 78° 55.237’ E 011° 50.495’) is closest to the Bayelva Monitoring Site, 21 m above sea level. BPF2 (~84 m from BPF1, N 78° 55.261’ E 011° 50.294’) is near a summer glacial melt riverbank at an elevation of 20 m above sea level (Fig. 1A). Three boreholes were drilled at each of the two

sites with a SIPRE auger drill (NSF, USA) with ~0.6 m snow cover that was shoveled away prior to drilling (Fig. 1B, 1C). Depths of core samples were limited by the ability to recover intact material. Core depths for BPF1 were 21 cm, 20 cm and 58 cm, and core depths from BPF2 were 30 cm, 21 cm, and 21 cm below the surface (Fig. 1D). Freshly cored samples were removed from the drill and kept inside pre-sterilized polycarbonate core liners (Jon's Machine Shop, USA) used during drilling. Core liners were capped and stored inside a sterile lined cooler to maintain frozen temperature.

Core processing: Cores were removed from the core liner and sliced into 2 cm depth intervals using a sterile geological sampling hammer and chisel at the King's Bay AS Marine Laboratory (Ny Ålesund, Svalbard). A portion of the sample was weighed then dried in a 60°C oven for 24 hours to determine gravimetric water content. Bulk density of the cores was estimated by measuring volume and dry mass of one intact 2.8 cm core puck with a diameter of 8 cm collected from BPF1; the resulting uniform bulk density was applied to all core samples for bulk density.

Culturing: To study colony variation and CFU/mL we tested various soil dilutions and three different agar types on the BPF1 0-2 cm. Following previously published methods, (18) we tested R2A, TSA and ½-strength TSA with a Master Soil Mixture (MSM) made from 10 g soil and 100 mL of 1X phosphate buffered saline (PBS, pH 7.2). The following four final soil suspensions ratios were plated: MSM:1XPBS; 0.1:0.9 (1 mL plated), 0.5:0.5 (1 mL plated), 1:0 (1 mL plated), and 3:0 (3 mL plated) for each of the three media types, in triplicate. CFU/mL were counted after growth at 4°C for 3 weeks. Full-strength TSA yielded a lawn of colonies after just a few days, and ½ TSA yielded only a few colonies (Fig S1). R2A agar was chosen for our study since it had the greatest colony diversity and allowed for slower growing colonies to form. The MSM:PBS dilutions listed above were made for eight depth intervals between BPF1 and BPF2 core sites (Fig. S2). From BPF1, the depth intervals were 0-12 cm, 12-24 cm, 24-36 cm, 36-48 cm, 48-58 cm. From BPF2, the depth intervals were 0-12 cm, 12-20 cm, and 20-30 cm. A MSM from these 8 depth intervals were plated with the same soil suspension scheme and incubated at 4°C for 3 weeks. After CFU/mL and colony variants were measured on the experimental plates, we streaked for isolation on R2A. Of the hundreds of isolates obtained on plates, 10 were selected for continued analysis based on their morphological diversity. Isolates

were grown on R2A for four weeks at 4°C, after choosing R2A over TSA and 1/2-strength TSA because the diluted medium allowed for isolation of oligotrophic microbes (Figs S1 and S2). The exact depth intervals they came from are listed in Table 1.

Soil Geochemical Analyses: Total carbon, nitrogen, carbon isotopic signature ($\delta^{13}C$), and nitrogen isotopic signature ($\delta^{15}N$) were determined on completely dried soil that was ground using a mortar and pestle into a fine powder. Large stones were removed and final particle size of the fine powder was not measured before analysis. To quantify the organic fraction of carbon, 1 mL of 1 N HCl was added to 5.0 g of soil and oven dried at 45°C for 56 hours to volatilize inorganic carbon (24). The $\delta^{13}C$ and percentage of inorganic carbon were calculated through mass balance for samples with >0.7% inorganic carbon (25). Total carbon, nitrogen, $\delta^{13}C$, and $\delta^{15}N$ were determined on finely ground soil samples (30 mg for BPF1 and 40 mg for BPF2) using a Costech ECS4010 Elemental Analyzer coupled to a Thermo-Finnigan Delta+XL mass spectrometer via a Thermo-Finnigan Conflo III device. Helium was used as the carrier gas, and the oxidation furnace was operated at 1050°C and the reduction furnace at 650°C. These measurements were performed at the Stable Isotope Laboratory at the University of Tennessee, Knoxville, USA.

Extractable dissolved organic carbon (DOC), phosphate, and inorganic nitrogen were measured on each of the 2 cm intervals of core samples (Fig. S3). Extracts were prepared by combining 5.0 g soil with 20 mL of 0.5 M potassium sulfate and shaking at room temperature for 4 hours. Samples were then filtered through a Whatman GF/B glass microfiber filter (1.0 µm pore size) using a vacuum extraction manifold. Filtrate was collected and frozen at -20°C for 12 hours. DOC was quantified by reacting extracts with a 0.42 M potassium persulfate (26) to oxidize soil organic carbon to CO_2. A series of potassium hydrogen phthalate standards was included in each analysis to use for calculation of persulfate-oxidized organic carbon. Samples and standards were reacted overnight (80°C) in sealed glass vials with rubber septa for headspace gas sampling, which was conducted after samples had cooled to room temperature using an infrared CO_2 gas analyzer. Sample extracts were carried through three colorimetric assays to measure extractable phosphate (PO_4^{3-}), nitrate (NO_3^-) and ammonium (NH_4^+) concentrations. The sum of nitrate and ammonium concentrations is used to calculated total

inorganic nitrogen, and the sum of organic nitrogen and inorganic nitrogen is equal to total nitrogen. Inorganic phosphate was measured using the Malachite Green assay (27). Nitrate was determined using a vanadium (III) chloride reagent (28) and ammonium was quantified using the Berthelot reaction (29). All nitrogen and phosphorous assays were conducted using protocols modified for a 96-well microplate reader (Synergy H1 Hybrid Reader, Biotek Inc., Winooski, VT USA).

Additional measurements for electrical conductivity, pH, and labile carbon were performed on a separate core sample from each site that was divided as follows: BPF1: 0-12 cm, 12-24 cm, 24-36 cm, 36-48 cm, 48-58 cm. BPF2: 0-12 cm, 12-20 cm, and 20-30 cm (Fig. S3). Electrical conductivity and pH were measured on these samples using calibrated bench-top meters (15 g soil in 45 mL deionized water). Additionally, pH, labile carbon and electrical conductivity measurements were performed on eight soil sample increments (BPF1:0-12 cm, 12-24 cm, 24-36 cm, 36-48 cm, 48-58 cm. BPF2:0-12 cm, 12-20 cm, and 20-30 cm) using calibrated bench-top meters (15 g dry soil in 45 mL deionized water). Permanganate oxidizable carbon (POXC), used to analyze total labile carbon, was extracted and quantified according to the method of Weil et al (2003). Briefly, 2.5 g was reacted with 20 mL 0.02 M potassium permanganate ($KMnO_4$) + 0.1 M calcium chloride solution by shaking at 120 rpm for 2 minutes. After settling, extracts were diluted 1:100 with milliQ water then measured for absorbance at 550 nm wavelength using a microplate spectrophotometer. Standards of a known concentration of $KMnO_4$ were included with each plate and used to determine the moles of $KMnO_4$ oxidized upon reaction with soil organic carbon. Assuming 9000 mg carbon oxidized per mole $KMnO_4$ (Weil et al., 2003), the amount of POXC was corrected for soil water content and reported in µg POXC g dry $soil^{-1}$.

DNA extraction and 16S rRNA gene amplification: For DNA extraction from isolates, a pellet was formed with low-speed centrifuging (7000 x g) of isolates that had been grown on R2A at 4°C for 4 weeks and DNA was extracted with a Qiagen DNA Power Soil Kit (QIAGEN, Germany).To amplify the 16S rRNA gene of the isolates, a PCR Master Mix was prepared as follows: 0.25 µl of Speedstar Taq polymerase (TaKaRa Bio, USA), 4 µl of 2.5 mM dNTPs, 5 µl of 10x Fast Buffer 1 (TaKaRa Bio, USA), 10 µl of 27F primer (5'-AGAGTTTGATYMTGGCTCAG-3') 10 µl of 1492R primers (5'-

TACGGYTACCTTGTTACACTT-3') (30) (Eurofins Genomics, USA), 29.25 µl of dH_2O and 2 µl of DNA; totaling 200 µl of volume per sample. Samples underwent PCR thermocycling in a BioRad T100 ThermoCycler (BioRad, USA) for 95°C for 1 minute, 95°C for 5 seconds, and 65°C for 20 seconds. The last two steps were repeated 34 times. After the thermocycler, samples were dyed with 6X TriTrack DNA Loading Dye (ThermoFisher, USA). Visualization of the PCR product was compared to a GeneRuler 1kb Plus DNA Ladder (ThermoFisher, USA). The samples were then placed into a 1.5% agarose gel with midori green DNA stain at 90V for 45 minutes with BioRad PowerPac Basic (BioRad, USA) for PCR product verification.

Sanger Sequencing of 16S rRNA genes of isolates: Amplified 16S rRNA gene (27F, 1492R) PCR product was cleaned with QIAGEN PCR Clean Up Kit (QIAGEN, Germany) and Sanger sequenced at the Sequencing Core Facility at the University of Tennessee, Knoxville. Sequences were viewed and forward and reverse reads were combined from the chromatogram in 4Peaks (v.1.7.2), and DECIPHER v2.17.1 (31) was used to check for chimeras. The combined sequences were analyzed in nucleotide BLAST (v2.11.0) and the closest related organisms were downloaded for comparison. Output sequences were classified with SILVA Sina (v1.2.11). SINA Alignment (v1.2.11) was also used to compare these isolates to 16S rRNA genes cataloged in the SSU database and to make a RAxML tree. All sequenced isolates have been deposited on NCBI GenBank Accession Numbers MZ773212-MZ773221.

Whole Genome Sequencing: DNA extractions from the 10 cultured isolates were sequenced with an Illumina MiSeq V3, 600 cycles (2x300) at The University of Tennessee, Knoxville Center for Environmental Biotechnology. Whole genome sequences were retrieved from Illumina BaseSpace and were assembled with SPAdes v3.13.0 (32) on KBase (link in data availability section) (33). Prokka v. 1.14.6 was used for annotations (34). All whole genomes are available on NCBI Accession Number PRJNA649544.

Metagenomes: DNA was extracted from the longest core sample from each site (BPF1: 58 cm, BPF2: 30 cm) using the QIAGEN DNA PowerSoil DNeasy DNA Extraction kit (Qiagen, Germany) with 0.5 g of starting material at King's Bay AS Marine Laboratory, no more than 3 hours after removal from the ground. DNA was quantified on each 2 cm soil sample using a Qubit™ 4 Fluorometer (ThermoFisher, USA). To get a total of 10 ng/µL for each metagenome, BPF1 samples were pooled into two groups (0-30 cm and 30-58 cm) and all

BPF2 samples were pooled (0-30 cm). These three samples were sequenced on an Illumina MiSeq with a V3 flow cell, using a 600 Nextera cycle kit with 275-300 bases paired end reads. Data was downloaded from Illumina's BaseSpace platform and analyzed with KBase (Arkin et al. 2018) with default program settings. Forward and reverse fastq reads were assembled with 98% of the two reads surviving and adapters were trimmed with Trimmomatic v0.36 (Bolger, Lohse, and Usadel 2014), then all three metagenomes were assembled separately with MetaSPAdes v3.14.1 (Nurk et al. 2017) and annotated with Prokka v. 1.14.6 (34). Metagenome libraries of soil samples and whole genomes of isolates were compared with read mapping in terms of 'reads per kilobase per read library' through Bowtie2 (35) and in-house python scripts.

Cell counts: The same protocol was used for cell counts of the bulk soil intervals and the cultures. The soil intervals used a dilution of 1:20, (1X PBS:soil suspension) the cultures were diluted based on visual opacity and spectrophotometer reading but did not exceed a 1:40 dilution (1X PBS:culture broth) (See Supplemental Datafile 1). 5X SYBR Gold stain was added and filtered on a vacuum Hoefer box on 0.2 um Millipore round filters. Filters were adhered to microscope slides with Vecta Sheild©, and a cover slip was applied. Microscope used was a Zeiss Axio Imager M2 Epifluorescence Microscope (Oberkochen, Germany). Cells were counted in 30 random fields of view at 10X magnification in the singular grid hemocytometer for eyepiece PL 10X/23 in 23mm X 23mm. Total cell counts were calculated by the following equation:

$$\frac{Cells}{mL} = \frac{\overline{x_{cells}}}{samplefiltered} * \frac{\mathrm{A_{filter}}}{\mathrm{A_{grid}}} * dilutionfactor \qquad \text{(Equation 1)}$$

Where $\overline{x_{cells}}$ is the average of the cells counted, sample filtered is the total sample used in the dilution, $\mathrm{A_{filter}}$ is the area of the Milipore 0.2 µm filter used, $\mathrm{A_{grid}}$ is the area of the hemocytometer grid, and dilution factor is the initial dilution factor of the sample:1X PBS.

Potential Enzyme Activity:

Bulk soil: Maximum potential activities for seven major carbon, nitrogen, and phosphorous hydrolytic enzymes were assayed at three incubation temperatures (5°C, 15°C, and 25°C) using fluorometric methods in triplicate (36, 37). Fluorescently-labeled substrates were used to measure the activity of exogenously added small substrate proxies for the following

enzymes: α-glucosidase (AG), β-glucosidase (BG), β-D-cellubiosidase (CB), leucine aminopeptidase (LAP), N-Acetyl-β-D-glucosaminidase (NAG), phosphatase (PHOS), and β-xylosidase (XYL). The same 8 soil increments (2.75 g) used to measure soil extractables (above) were suspended with 50 mM Tris buffer (pH 7.7) matching the measured mean sample pH, using high-speed blending for 60 seconds. Added fluorescent labels MUC (7-amino-4-methylcoumarin) and MUB (4-methylumbelliferone) were used for standardization. Soil slurries were incubated with MUC and MUB standards and labeled substrates (200 µl of 200 µM solution) for 3 hours at 25°C, 6 hours at 15°C, and 24 hours at 5°C in triplicate. Fluorescence was measured using a microplate reader (Synergy H1 Hybrid Reader, Biotek Inc., Winooski, VT USA) with 365 nm excitation wavelength and 450 nm emission wavelength set at optimal gain. Enzyme activity was calculated in nmol g dry soil^{-1} h^{-1}, with higher activities indicating a greater amount of fluorescently labeled substrate that was degraded under the ideal conditions of incubation.

Pure Culture Enzyme Activity: We modified the above maximum enzymatic activity potential method to measure the ten cultured isolates' activity. Cultures grew in broth R2A for 24 hours at 25°C prior to the experiment to reach exponential growth phase (Fig. S4). Cultures at this time were also used for a cell count to determine the number of cells that were in solution (Fig S5A). Cultures were suspended in 35 mL 7.7 pH Tris buffer and then distributed across the 96-well plate and incubated with the small substrate proxies listed above.

Results

Soil geochemical analyses:

The organic carbon of BPF1 ranged from 1 to 3.5% (Fig. 2A) while inorganic carbon ranged between 0 and 0.5% (except for one point at 1.5%). BPF2 had a slightly decreasing trend with depth for organic carbon (2.0% to 1.0%) and slight increase in inorganic carbon from ~0 to 1.5% with depth (Fig. 2B, Supplemental Datafile 1). Total nitrogen decreased with depth in BPF2 and ranged from 0.05 and 0.10%. Total nitrogen was nearly twice as high in BPF1, ranging from 0.05 to 0.25% (Fig. 2C). The $\delta^{13}C$ of organic carbon ranged from -26‰ to -25‰ with the exception of two points (Fig. 2D). The inorganic $\delta^{13}C$ values for BPF1 varied between

-30 and -20‰, while the inorganic $\delta^{13}C$ of BPF2 contains more ^{13}C with depth (-20‰ to ~10‰) (Fig. 2E). $\delta^{15}N$ in BPF1 and BPF2 decreased with depth (2.5 to 1.5‰) in the top 30 cm only (Fig. 2F). Carbon to nitrogen ratios (C/N) in BPF1 ranged from ~13 to 18 and those of BPF2 increased with depth from ~19 to 50 (Fig. 2G). The water content per weight of the samples decreased with depth at both sites, except for BPF1 30-32 cm, which was mostly ice (Fig. 2I). Dissolved organic carbon (DOC) was similar in both cores, ranging from 0.3 to 3.2 $mg/g_{dry\ soil}$. Inorganic nitrogen (sum of the measured NH_4^+ and NO_3^-) ranged from 0.05 to 0.2 $\mu g/g_{dry\ soil}$ and was similar between the two cores (Fig. 3). BPF2 had a lower concentration of PO_4^{3-} at each depth compared to BPF1. However, both had negative values that were below the limit of detection of 0.0015 $\mu g/g_{dry\ soil}$. NH_4^+ was highest at the surface (0.07 $\mu g/g_{dry\ soil}$ for BPF1 and 0.18 $\mu g/g_{dry\ soil}$ for BPF2), and consistently close to 0.05 $\mu g/g_{dry\ soil}$ for both cores below the surface. NO_3^- was similar for both cores, varying from 0 to 0.04 $\mu g/g_{dry\ soil}$ (Fig. 3). The measured electrical conductivity was less than 0.05 mS in BPF1 and ranged from 0.025 mS to 0.17 mS in BPF2 (Fig. S3). The pH range of BPF1 and BPF2 were both between 7.4 and 8.1 (Fig. S3). Labile carbon (mg $POXC/g_{dry\ soil}$) was lower in BPF2 (200-250), while BPF1 had greater POXC and higher variability with depth (320-825) (Fig. S3).

Cultured Isolates:

Ten cultures were chosen from the plated 0.5 mL soil suspension on R2A based on preliminary culture tests (See Methods, Figs. S1, S2). Eight cultured isolates originated from BPF1 soil (Table 1) and two from BPF2 soil. All were gram negative rods of *Pseudomonas* sp. (Fig 4, Fig. S6). Isolates of *Pseudomonas* sp. have been previously found in soil from Thuringian Basin, Council Alaska, Livingston Island Antarctica and other Antarctica soils (Table 1 and references therein, Fig. 4). The 16S rRNA gene sequence of each isolate was ≥99% similar to a previously cultured 16S rRNA genes on NCBI (Fig. 4, Table 1).

Potential Enzyme Activities:

Bulk soil:

The highest maximum potential enzymatic activities in bulk soil were observed for LAP, PHOS, and BG, with the highest value from LAP at 25°C in BPF2 0-12 cm (225 nmol/g/hour)

(Fig. 5, Table S1). The lowest maximum enzymatic potential activities were from AG, XYL, and CB. PHOS was the only enzyme of these three to have higher activities at 25°C for all 8 soil intervals, with the other enzymes showing maximum potential activity at the lower 15°C temperature for at least one depth (Fig. 5A). One sample (BPF1, 36-48 cm) had higher XYL activity at 5°C than either 15°C or 25°C (~2 nmol/g/hour). In BPF1, NAG had the highest activity at 15°C for all intervals except 36-48 cm (~1-10 nmol/g/hour). In three depths of BPF1, XYL had a higher activity at the two lower temperatures. These increased activity at lower temperatures were only observed for BPF1. The 5°C treatment had the lowest enzyme activity for half of the tested soil intervals. BG in both sites had the lowest activity at 5°C. The 25°C treatment had the highest enzyme activity for 59% of the tested soil intervals. Temperature was correlated with LAP (spearman = 0.66) and PHOS (spearman = 0.64) (Fig. S7). Most enzymatic activities decreased with depth at both sites, except for PHOS, LAP, XYL in BPF1.

Culture isolates:

The ten cultured *Pseudomonas* sp. isolates displayed a wide range of maximum potential enzymatic activities (Fig. 5B, Table S2). The highest activities were at 25°C from G16 (3.5 nmol/mL/hour) and B7 (2.4 nmol/mL/hour). NAG and XYL had higher activity at 15°C for five of the cultures. PHOS and LAP had higher activity for each cultured organism at 25°C. Culture E5 had low activities for all enzymatic substrates measured, most likely because it had much lower cell abundance than the other cultures (Fig. S5A). Compared to measurements in bulk soil, the cultures generally had higher maximum potential activities for NAG, and lower values for PHOS, XYL, AG, and BG. The highest activities for cultures were in the LAP, NAG and PHOS, in descending order. Overall, the 25°C treatment had the highest activity in most of the enzymes, except for XYL and NAG, similarly to the soil suspensions. CB, AG, BG had very low activities in all cultures except B7, possibly because B7 had more cells than the other cultures at the time of measurement (Fig. S5A).

Metagenomes:

Extracted DNA was very low throughout the 44 individual depth samples with only three samples higher than 2 ng/μL in BPF1 (Fig. 2H). The three metagenomic libraries had between 8 million and 25 million reads that assembled into 7,000 to 13,000 contigs (Table 2).

Since all metagenomic bins were within the MIMAGs Low Quality or Medium quality standards (38), we analyzed the whole metagenomic assembly as an indicator of the environmental abilities. Whole genome sequencing (WGS) was used to obtain genomes for the cultured isolates. Metagenomic reads were then mapped to the cultured whole genome isolates (Fig. S8). Genomes of the cultured isolates recruited reads from at least one metagenome (7 to 22,500 reads per kilobase per metagenomic library). G17 and G19 recruited the most reads from all three metagenomes. The 10 culture whole genomes and the three metagenome libraries were analyzed for presence of the genes encoding the enzymes tested in the enzymatic activity measurements (Table 3, Table S3). LAP had the highest range of gene counts (2 to 71) and PHOS had the most even distribution of counts between all the samples (9 to 15). AG, BG and NAG had little to no gene counts while XYL had a sparse amount. Leucine aminopeptidase was the only enzyme that appeared to be related to the quantity of genes in the metagenomes. LAP had the cumulative highest number of gene counts and the highest range of activity for the bulk soil and the cultures at 25°C. PHOS genes were present in every sample and had highest activities in the 25°C treatment.

Discussion

Differentiating microbial metabolism between the two active layer sites displayed by geochemistry

Despite being less than 100 meters apart, BPF1 and BPF2 differed in the $\delta^{13}C$ signatures of inorganic carbon, as well as percentages of inorganic carbon, organic carbon, and total nitrogen. The organic carbon $\delta^{13}C$ of both cores showed signatures originating from plant material (-25‰ to -35‰) (39) (Fig. 2D). At BPF1, the $\delta^{13}C$ of inorganic carbon was similar to the $\delta^{13}C$ of organic carbon at all depths. This likely indicates that the inorganic carbon was produced from organic matter through microbial degradation, since heterotrophy expresses little to no kinetic isotopic effects. In BPF2, the inorganic carbon was ^{13}C-enriched ranging from -20‰ to 14‰ with depth. Three possibilities to explain this are 1) inorganic carbon becomes ^{13}C-enriched due to microbial carbon fixation, which can leave behind ^{13}C-enriched inorganic carbon due to kinetic fractionation (40); 2) a large source of ^{13}C-enriched inorganic carbon

diffuses upward in the core, mixing with ^{13}C-depleted heterotrophically produced inorganic carbon as it diffuses downward, and 3) both processes are occurring. Evidence that less microbial heterotrophy may be occurring in BPF2, which would allow autotrophy to have a greater effect on the $\delta^{13}C$ values for inorganic carbon, comes from the fact that there are lower amounts of labile organic matter, compared to BPF1. C/N ratios can indicate how labile carbon compounds are for microbial activity in the soil, with lower ratios (13-18) indicating a greater availability for metabolic use (41). BPF1 had significantly more organic carbon ($p < 0.015$) differences by comparable depths between sites and a statistically lower C/N ratio ($p < 10^{-4}$), largely driven by the statistically higher nitrogen content ($p < 10^{-5}$), suggesting that there may be greater activity of microbial carbon degradation in BPF1 due to higher carbon lability (Fig. 2A, 2G). POXC values, a proxy for organic matter lability (42), were also higher in BPF1 (582 mg POXC/$g_{dry\ soil}$ ± 109 vs. 225 mg POXC/$g_{dry\ soil}$ ± 31 for BPF2) (Fig. S3). Microbial biomass was also greater in BPF1, shown by higher DNA concentrations (Fig. 2H, 87% of samples at the same depth were higher in BPF1) and cell counts in the comparable soil increment depths (~80,000 cells/ g soil vs ~26,000 cells/ g soil, Fig. S5B, Supplemental Datafile 1) than BPF2. However, carbonates could have come from the wide range of carbonate deposits that are found in Svalbard (43). These could contribute to the proglacial active layer after glacial scouring and would be likely to have values ^{13}C-enriched by at least 20‰ relative to organic matter (44), so the possibility of the $\delta^{13}C$ trends in inorganic carbon could also reflect a carbonate source from below mixing with a heterotrophic source above. Given the evidence for a decrease in heterotrophy in BPF2, and the prevalence of ^{13}C-enriched carbonate deposits in the area, both processes likely contribute to the observed stable carbon isotope trends.

Nitrogen is often a limiting nutrient in Arctic environments and this can reduce the compounds that are available for microbial metabolism (45, 46). The inorganic nitrogen at these two sites (NH_4^+ and NO_3^-) makes up a small portion of the total nitrogen present (less than 10^{-8}%, Fig. 2C and Fig. 3) which means the limited nitrogen that is present here is organic nitrogen, most likely in the form of plant material and input from animals (47). NH_4^+ and NO_3^- are highest in the surface (0.05 μg/$g_{dry\ soil}$ and 0.04 μg/$g_{dry\ soil}$, respectively). At both sites and the lowest values of NH_4^+ (0.02 μg/$g_{dry\ soil}$) are higher than the highest values of NO_3^- (0.04 μg/$g_{dry\ soil}$, Fig. 3). This could be an indication of NH_4^+ oxidation or nitrate absorption by plants,

since the $\delta^{15}N$ signature of plant organic matter is 2-3‰ (48). This may occur in the upper few centimeters at both sites, as NH_4^+ decreases and NO_3^- slightly increases with depth. Therefore, the $\delta^{15}N$ signature that is decreasing with depth at both locations likely originates from plant processes and degraded proteins.

Phosphatase generally has the highest activity among the enzymes commonly measured in soils because most soil microbes are capable of creating extracellular phosphatases and when phosphorus is low, phosphatase enzyme expression is increased (49). Lastly, the low/below limit of detection values for phosphate in the soil profile (Fig. 3) and the increased activities for phosphatase (Fig. 5A) further emphasize how these organisms in these soils are limited in nutrients for microbial activity. Even though BPF1 has more labile carbon and higher microbial biomass than BPF2, it only had higher activities in two enzymes (AG and BG) and only in the upper section. Even with less heterotrophy in BPF2, the reason the microbial community maintains similar rates of carbon-degrading enzymes may be to access the nitrogen and phosphorous in organic compounds.

In both Bayelva active layer cores the highest measured enzymatic activity was for leucine aminopeptidase, and that activity was generally higher at 25°C than at 15°C or 5°C. Leucine aminopeptidases are critical cell maintenance enzymes that drive peptide turnover (50) and are often used as an indicator of total peptidase potential in an environment due to the non-specificity of the enzyme (51). There are many environments that report LAP activity to be the highest among this tested suite of enzymes (52). Bacteria have been shown to use leucine, along with other amino acids, as an alternative source of carbon and nitrogen in energy-limited environments (53). Nitrogen demand may explain why LAP is nearly two orders of magnitude higher in activity than the other six tested enzymes for both the bulk soil and cultured isolates (Fig. 5A-B).

Soil and cultured enzymes show cold adaptation

Characterizing environmental microbes is challenging as many organisms resist common culturing methods (54). Our isolates belong to a genus that has previously been found in active layers (Fig. 4, Table 1). *Pseudomonas* species are gram-negative, aerobic, bacilli

organisms (55). This genus contains over 140 species and most are saprophytic (55) and common in soil. Some *Pseudomonas* sp. can be psychrophilic while others are mesophilic (18–20). The isolates we obtained from the Bayelva sites are most closely related to species that have been found from similar cold regions such as Alaska, Antarctica, and Tibetan plateau (Table 1).

A wide range of enzyme classes, including peptidases and carbohydrate lyases, have been shown to be cold-adapted (56, 57). One of the best studied enzymes that has a lower K_m (and therefore greater catalytic efficiency) at cold temperatures is chitobiase, or N-acetyl-□-D-glucosaminidase, obtained from an *Arthobacter* sp. culture (56). NAG is also often used as an indicator of nitrogen mineralization within soils (6). This enzyme showed higher V_{max} values at 15°C, rather than 25°C in four depths of BPF1 and in five of the *Pseudomonas* sp. cultures. This implies that psychrophily in this class of enzymes may be widespread among active layer microbes since multiple isolates as well as the natural population had this property. The only other enzyme that appeared to be cold-adapted was xylosidase, which had its highest V_{max} at 15°C in three soil samples and in the same five *Pseudomonas* sp. isolates that had psychrophilic N-acetyl B-D glucosaminidase. One depth of BPF1 had the highest activity for this enzyme at 5°C. Xylosidase has been found to be cold adapted in a *Duganella* sp. isolate from Antarctic soil (56). Both enzymes are good evidence that the cold-adaptations observed in pure-culture experiments are upheld in natural microbial populations. This suggests that these enzymes are well-adapted to a cold environment but have the flexibility to continue functioning as temperatures warm.

Metagenomes and cultures parallel enzymatic results

The read mapping between the metagenomes and the whole genomes of the isolates showed that each isolate was present in the environment (Fig. S8). However, the low reads per kilobase per million mapped reads (RPKM) values suggests that these organisms were not the dominant community members; instead, they were the best at growing on the general medium used. There were more genes encoding LAP than other enzymes within both the metagenomes and isolates' whole genomes (Table 3). The second highest gene counts and activity were in

PHOS across the whole genomes and the metagenomes. The whole genomes had more gene counts for three of the four carbohydrate-degrading enzymes, (AG, BG, and XYL), than the metagenomes. However, enzymatic activity was often higher in bulk soil than in cultured isolates. This could indicate that microbes in the natural communities have enzymes with higher V_{max} values, or that the cultured *Pseudomonas* sp. produced lower amounts of enzymes when grown in culture, perhaps because they were not nutrient limited. Additionally, gene presence may not always mean a higher rate of enzymes exported for compound utilization. For instance, leucine can be cleaved by multiple types of enzymes, which may inflate the number of genes for LAP (52). This is reflected in the lack of correlation between gene dosage per genome or per metagenome and the measured activity for that enzyme class.

Ny Ålesund, Svalbard active layer microbes are ready for thaw

Dissolved organic carbon (DOC) is a potential source of carbon and energy for heterotrophic organisms (58, 59). Previous studies have shown high spatial variability in permafrost DOC values. For instance, active layers in the Eight Mile Lake soil in Alaska, USA, ranged from 0.063 mg/g to 0.46mg/g (5). This location does not reach maximum DOC, 46.38 mg/g, until the bottom of the active layer at 35-45 cm. Comparatively, the DOC in the two Svalbard active layer sites was highest at the surface (2 mg/g_{soil}, Fig. 3).

Nitrogen is the limiting nutrient for plant growth in arctic tundra, and an increase in nitrogen will increase primary production (7). When the active layer thaws, it provides the rooting zone, which is a region for plant roots and soil microbes to compete for nitrogen (6). Salmon *et al.* (2018) found that C/N decreased, %N decreased, and %C decreased with depth. The C/N of organic matter is fairly constant with depth at both BPF1 and BPF2, and the range of values from 20 to 28 is lower than the C/N in active layer soil of Eight Mile Lake, Alaska (21 to 46) (6). This indicates that the increase in nitrogen limitation with depth at the Alaskan site may not apply to our Bayelva, Svalbard active layer samples.

These Svalbard active layer soils differ from Taylor Valley in Antarctic McMurdo Dry Valley soils in DOC, water content, inorganic nitrogen, and labile carbon (60). Moisture from soils near McMurdo Dry Valley had twenty times the DOC, two times the water content, and three orders of magnitude higher inorganic nitrogen than the Svalbard active layer samples. The

McMurdo Dry Valley has a higher inorganic nitrogen contribution from NO_3^-, while Svalbard's inorganic nitrogen contribution is mainly from NH_4^+. Additionally, we found the Svalbard active layer to have a higher amount of organic nitrogen. Labile carbon (POXC) in Svalbard was up to a hundred times higher than the Antarctic active layer, even though the DOC was up to twenty times lower. These measurements could indicate that Arctic and Antarctic sites may have different outcomes for carbon degradation as permafrost thaws and microbial activity increases. Our work suggests that Svalbard will have higher activity, compared to McMurdo Dry Valley soils, due to the higher amounts of organic nitrogen and labile carbon based on POXC and C/N ratios.

Conclusion

The Svalbard active layer soils show unique trends in soil geochemistry between the two sites, despite being only 84 m apart. The two sites show differences in their measured carbon, nitrogen, and lability of the organic carbon present. These differences may be driving the different microbial metabolisms, where BPF1 has an environment and carbon isotopic signatures of microbial heterotrophy and BPF2 shows evidence that the low activities of heterotrophs allow microbial autotrophy to dominate the isotopic fractionation signal and/or heterotrophically produced inorganic carbon mixes with geological carbonate deposits. BPF1 is on an incline while BPF2 is in a seasonal river moraine and could be receiving relocated soil material from higher elevations (colluvial soil). These two sites had high values of LAP activity, which is commonly observed in soil environments, perhaps due to the large range of enzymes capable of hydrolyzing leucine (52). The four carbohydrate-degrading enzymes, AG, BG, XYL, CB, had the lowest activity for the soil and the cultures. Carbohydrates often are derived from plants, which are small, scarce, and seasonal in this environment (Fig. 1C). Some of these enzymes (XLY & NAG) had higher activity in 5°C or 15°C than in 25°C, suggesting that enzymatic adaptation to cold temperatures or an increase in enzyme production, may be a strategy for microbial cold adaptation (60).

The isolates found in the two sites are organisms that have been previously reported in cold soil environments, which suggests that they are from the study site and not contaminants

(Table 1). The recruitment of metagenomic reads to the genomes of these isolates further suggests that these ten *Pseudomonas* sp. are present in the whole *in situ* soil community, but not dominant. While enzymatic activities of bulk soil and isolated cultured are difficult to compare to true *in situ* environmental activities, there was agreement between the two. The culture and bulk soil enzymatic activities had highest values in LAP. PHOS activities were all higher in the 25°C while XYL and NAG showed some higher activities in the lower temperatures tested. The total cell counts for the bulk soil were an order of magnitude larger than the cultures. The cultures cell counts ranged from 100 to ~36,000 cell/mL and the bulk soil ranged from 10,000 to 180,000 cells/g_{soil} (Figure S5A and B, Supplemental Data File 1). Despite this, their potential maximum enzymatic activities were similar. This could mean that a smaller number of organisms are able to have the same enzymatic effect as a larger population.

Given the greater activity of most enzymes at higher temperatures, it is likely that Svalbard active layer soils will experience higher microbial activity as the temperatures increase in this warm permafrost. From the data presented here, it is likely that the microbes will become more active once the Svalbard active layer expands as compared to the microbes present in Taylor Valley of McMurdo Dry Valley sites (60). The measured higher quantity of labile carbon in areas of soil like BPF1 could indicate that microbial activity will spike during permafrost thaw and active layer expansion. While nitrogen is still a limiting nutrient for microbial activity in the Arctic, the increase of soil organic matter degradation will introduce more labile nitrogen and carbon compounds, allowing for a higher rate of microbial activity. If microbial respiration increases then stored nitrogen will be more available for plants to grow and perhaps could further lead to higher rates of plant derived carbon enzyme activity. This study demonstrates that the active layer soil near Bayelva in Ny Ålesund, Svalbard, will become more microbially active with different carbon degradation pathways, adaptable enzymatic activities, and utilization of the scarce resources.

Acknowledgments. The authors thank Andrew D. Steen, Lauren Mullen, Lisa Hubert and Alexander B. O. Michaud for help with field sampling. Madison Spradley assisted with lab work. Nicholas T. Sipes and Michael Tomaino assisted with in house scripts, can be found on https://github.com/sipesk/SvalbardActiveLayer. Funding came from the Simons Foundation

(404586 to K.G.L), NSF Dimensions of Biodiversity (DEB-1442262 to T.V., T.C.O and K.G.L), U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science Program (DE-SC0020369 to K.G.L, T.C.O and T.V.), and the ASPIRE program (funded through the NSF to R.P.)

Data Availability

Whole genomes and Genbank submissions can be found on NCBI Accession Number PRJNA649544 and MZ773212-MZ77322, respectively.

KBase with the workflow of the metagenomic analysis can be found on permanent links:

WGS: https://narrative.kbase.us/narrative/83182

MISEQ: https://narrative.kbase.us/narrative/56628

References

1. Cohen J, Screen JA, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K, Entekhabi D, Overland J, Jones J. 2014. Recent Arctic amplification and extreme mid-latitude weather. Nat Geosci. Nature Publishing Group.
2. H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama NMW (eds. . 2019. The ocean and cryosphere in a changing climate. A special report of the intergovernmental panel on climate change. Intergov Panel Clim Chang undefined.
3. Overland JE, Wang M, Walsh JE, Stroeve JC. 2014. Future Arctic climate changes: Adaptation and mitigation time scales. Earth's Futur 2:68–74.
4. Keating K, Binley A, Bense V, Van Dam RL, Christiansen HH. 2018. Combined Geophysical Measurements Provide Evidence for Unfrozen Water in Permafrost in the Adventdalen Valley in Svalbard. Geophys Res Lett 45:7606–7614.
5. Mann PJ, Sobczak W V., Larue MM, Bulygina E, Davydova A, Vonk JE, Schade J, Davydov S, Zimov N, Holmes RM, Spencer RGM. 2014. Evidence for key enzymatic controls on metabolism of Arctic river organic matter. Glob Chang Biol 20:1089–1100.
6. Waldrop MP, Wickland KP, White III R, Berhe AA, Harden JW, Romanovsky VE. 2010. Molecular investigations into a globally important carbon pool: permafrost-protected carbon in Alaskan soils. Glob Chang Biol 16:2543–2554.
7. Salmon VG, Schädel C, Bracho R, Pegoraro E, Celis G, Mauritz M, Mack MC, Schuur EAG. 2018. Adding depth to our understanding of nitrogen dynamics in permafrost soils. J Geophys Res Biogeosciences 123:2497–2512.
8. Beermann F, Langer M, Wetterich S, Strauss J, Boike J, Fiencke C, Schirrmeister L, Pfeiffer EM, Kutzbach L. 2017. Permafrost Thaw and Liberation of Inorganic Nitrogen in Eastern Siberia. Permafr Periglac Process 28:605–618.
9. Garnello A, Marchenko S, Nicolsky D, Romanovsky V, Ledman J, Celis G, Schädel C, Luo Y, Schuur EAG. 2021. Projecting Permafrost Thaw of Sub-Arctic Tundra With a Thermodynamic Model Calibrated to Site Measurements. J Geophys Res Biogeosciences 126:e2020JG006218.
10. Jentzsch K, Schulz A, Pirk N, Foken T, Crewell S, Boike J. 2021. High levels of CO2 exchange during synoptic-scale events introduce large. Geophys Res Lett 1–9.
11. Mu C, Zhang T, Zhang X, Cao B, Peng X, Cao L, Su H. 2016. Pedogenesis and physicochemical parameters influencing soil carbon and nitrogen of alpine meadows in permafrost regions in the northeastern Qinghai-Tibetan Plateau. Catena 141:85–91.
12. Hu G, Fang H, Liu G, Zhao L, Wu T, Li R, Wu X. 2014. Soil carbon and nitrogen in the active layers of the permafrost regions in the Three Rivers' Headstream. Environ Earth Sci 72:5113–5122.
13. Schimel JP, Schaeffer SM. 2012. Microbial control over carbon cycling in soil. Front

Microbiol. Frontiers Research Foundation.
14. Buongiorno J, Sipes K, Wasmund K, Loy A, Lloyd KG. 2020. Woeseiales transcriptional response to shallow burial in Arctic fjord surface sediment. PLoS One 15:e0234839.
15. Buongiorno J, Herbert LC, Wehrmann LM, Michaud A, Laufer K, Røy H, Jørgensen BB, Szynkiewicz A, Faiia A, Yeager KM, Schindler K, Lloyd KG. 2019. Complex microbial communities drive iron and sulfur cycling in Arctic fjord sediments. Appl Environ Microbiol AEM.00949-19.
16. Amato P, Hennebelle R, Magand O, Sancelme M, Delort AM, Barbante C, Boutron C, Ferrari C. 2007. Bacterial characterization of the snow cover at Spitzberg, Svalbard. FEMS Microbiol Ecol 59:255–264.
17. Schostag M, Stibal M, Jacobsen CS, Bælum J, Tas N, Elberling B, Jansson JK, Semenchuk P, Priemé A. 2015. Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. Front Microbiol 6:1–13.
18. Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L. 2018. Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. mSystems 3.
19. Vishnivetskaya T, Kathariou S, McGrath J, Gilichinsky D, Tiedje JM. 2000. Low-temperature recovery strategies for the isolation of bacteria from ancient permafrost sediments. Extremophiles 4:165–173.
20. Beyer A, Rzanny M, Weist A, Möller S, Burow K, Gutmann F, Neumann S, Lindner J, Müsse S, Brangsch H, Stoiber-Lipp J, Lonschinski M, Merten D, Büchel G, Kothe E. 2015. Aquifer community structure in dependence of lithostratigraphy in groundwater reservoirs. Environ Sci Pollut Res 22:19342–19351.
21. Finster KW, Herbert RA, Lomstein BA. 2009. Spirosoma spitsbergense sp. nov. and Spirosoma luteum sp. nov., isolated from a high Arctic permafrost soil, and emended description of the genus Spirosoma. Int J Syst Evol Microbiol 59:839–844.
22. Hansen AA, Herbert RA, Mikkelsen K, Jensen LL, Kristoffersen T, Tiedje JM, Lomstein BA, Finster KW. 2007. Viability, diversity and composition of the bacterial community in a high Arctic permafrost soil from Spitsbergen, Northern Norway. Environ Microbiol 9:2870–2884.
23. Medina D, Walke JB, Gajewski Z, Becker MH, Swartwout MC, Belden LK. 2017. Culture media and individual hosts affect the recovery of culturable bacterial diversity from Amphibian skin. Front Microbiol 8.
24. Sonjak S, Frisvad JC, Gunde-Cimerman N. 2006. Penicillium mycobiota in Arctic subglacial ice. Microb Ecol 52:207–216.
25. Harris D, Horwáth WR, van Kessel C. 2001. Acid fumigation of soils to remove carbonates prior to total organic carbon or CARBON-13 isotopic analysis. Soil Sci Soc Am J 65:1853–1856.
26. Komada T, Anderson MR, Dorfmeier CL. 2008. Carbonate removal from coastal

sediments for the determination of organic carbon and its isotopic signatures, δ13C and Δ14C: Comparison of fumigation and direct acidification by hydrochloric acid. Limnol Oceanogr Methods 6:254–262.
27. Doyle A, Weintraub MN, Schimel JP. 2004. Persulfate digestion and simultaneous colorimetric analysis of carbon and nitrogen in soil extracts. Soil Sci Soc Am J 68:669–676.
28. D'Angelo E, Crutchfield J, Vandiviere M. 2001. Rapid, Sensitive, Microscale Determination of Phosphate in Water and Soil. J Environ Qual 30:2206–2209.
29. Doane TA, Horwáth WR. 2003. Spectrophotometric determination of nitrate with a single reagent. Anal Lett 36:2713–2722.
30. Rhine ED, Mulvaney RL, Pratt EJ, Sims GK. 1998. Improving the Berthelot Reaction for Determining Ammonium in Soil Extracts and Water. Soil Sci Soc Am J 62:473.
31. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. 2008. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol 74:2461–2470.
32. Wright ES, Yilmaz LS, Noguera DR. 2012. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. Appl Environ Microbiol 78:717–725.
33. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol 19:455–477.
34. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569.
35. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
36. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9.
37. Saiya-Cork KR, Sinsabaugh RL, Zak DR. 2002. The effects of long term nitrogen

deposition on extracellular enzyme activity in an Acer saccharum forest soil. Soil Biol Biochem 34:1309–1315.
38. Bell CW, Fricks BE, Rocca JD, Steinweg JM, McMahon SK, Wallenstein MD. 2013. High-throughput fluorometric measurement of potential soil extracellular enzyme activities. J Vis Exp 1–16.
39. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol. Nature Publishing Group.
40. Mu C, Zhang T, Wu Q, Zhang X, Cao B, Wang Q, Peng X, Cheng G. 2014. Stable carbon isotopes as indicators for permafrost carbon vulnerability in upper reach of Heihe River basin, northwestern China. Quat Int 321:71–77.
41. Sinsabaugh RL, Hill BH, Follstad Shah JJ. 2009. Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment. Nature 462:795–798.
42. Weil RR, Islam KR, Stine MA, Gruver JB, Samson-Liebig SE. 2003. Estimating active carbon for soil quality assessment: A simplified method for laboratory and field useAmerican Journal of Alternative Agriculture.
43. Kou D, Yang G, Li F, Feng X, Zhang D, Mao C, Zhang Q, Peng Y, Ji C, Zhu Q, Fang Y, Liu X, Xu-Ri, Li S, Deng J, Zheng X, Fang J, Yang Y. 2020. Progressive nitrogen limitation across the Tibetan alpine permafrost region. Nat Commun 11:1–9.
44. Manzoni S, Taylor P, Richter A, Porporato A, Ågren GI. 2012. Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. New Phytol 196:79–91.
45. Solheim B, Endal A, Vigstad H. 1996. Nitrogen fixation in Arctic vegetation and soils from Svalbard, Norway. Polar Biol 16:35–40.
46. Gauthier PPG, Lamothe M, Mahé A, Molero G, Nogués S, Hodges M, Tcherkez G. 2013. Metabolic origin of δ15N values in nitrogenous compounds from Brassica napus L. leaves. Plant, Cell Environ 36:128–137.
47. Lee Taylor D, Sinsabaugh RL. 2015. The Soil Fungi, p. 77–109. *In* Soil Microbiology, Ecology and Biochemistry. Elsevier.
48. Matsui M, Fowler JH, Walling LL. 2006. Leucine aminopeptidases: Diversity in structure and function. Biol Chem.
49. Sinsabaugh RL, Lauber CL, Weintraub MN, Ahmed B, Allison SD, Crenshaw C,

Contosta AR, Cusack D, Frey S, Gallo ME, Gartner TB, Hobbie SE, Holland K, Keeler BL, Powers JS, Stursova M, Takacs-Vesbach C, Waldrop MP, Wallenstein MD, Zak DR, Zeglin LH. 2008. Stoichiometry of soil enzyme activity at global scale. Ecol Lett 11:1252–1264.

50. Steen ADA, Vazin JJP, Hagen SMS, Mulligan KHK, Wilhelm SSW. 2015. Substrate specificity of aquatic extracellular peptidases assessed by competitive inhibition assays using synthetic substrates. Aquat Microb Ecol 75:271–281.
51. Díaz-Pérez AL, Díaz-Pérez C, Campos-García J. 2016. Bacterial l-leucine catabolism as a source of secondary metabolites. Rev Environ Sci Biotechnol. Springer Netherlands.
52. Ramos JL, Levesque RC. 2006. PseudomonasPseudomonas. Springer US.
53. Kim DY, Kim J, Lee YM, Lee JS, Shin DH, Ku BH, Son KH, Park HY. 2021. Identification and Characterization of a Novel, Cold-Adapted d-Xylobiose- and d-Xylose-Releasing Endo-β-1,4-xylanase from an Antarctic Soil Bacterium, Duganella sp. PAMC 27433. Biomolecules 11:1–16.
54. Feller G, Gerday C. 2003. Psychrophilic enzymes: hot topics in cold adaptation. Nat Rev Microbiol 1:200–208.
55. Kaplan LA, Newbold JD. 2000. Surface and Subsurface Dissolved Organic Carbon, p. 237–258. *In* Streams and Ground Waters. Elsevier.
56. Fischer H, Sachse A, Steinberg CEW, Pusch M. 2002. Differential retention and utilization of dissolved organic carbon by bacteria in river sediments. Limnol Oceanogr 47:1702–1711.
57. Zeglin LH, Sinsabaugh RL, Barrett JE, Gooseff MN, Takacs-Vesbach CD. 2009. Landscape distribution of microbial activity in the mcmurdo dry valleys: Linked biotic processes, hydrology, and geochemistry in a cold desert ecosystem. Ecosystems 12:562–573.

Appendix

Figures and Tables

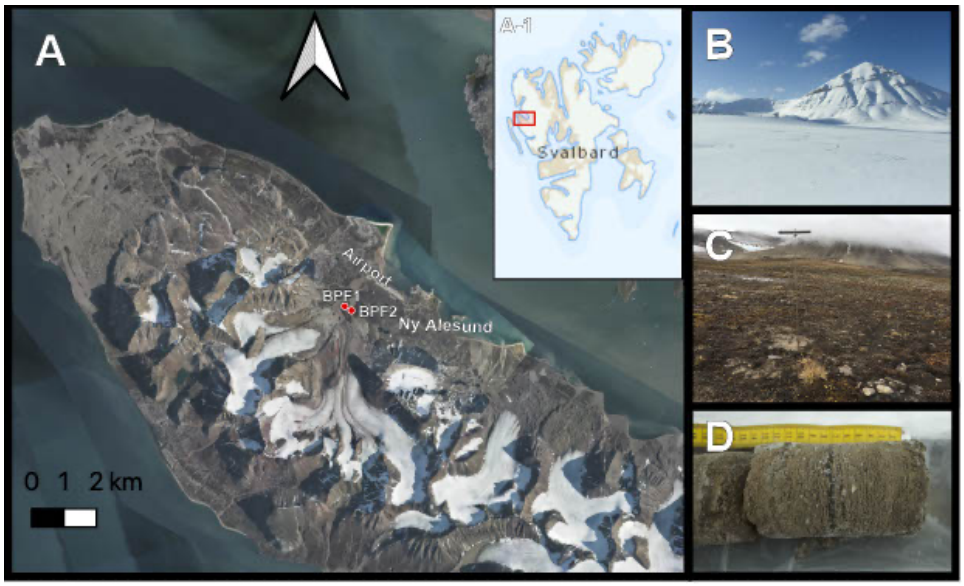


Figure 3-1: Sample site A) Active layer cores were taken from two permafrost sites near the Bayelva River in the Leirhaugen glacier moraine in Ny Ålesund, Svalbard, 79°N. Inset A-1 shows the location of the image in panel A within the Svalbard archipelago. B) Picture taken from the BPF1 borehole in April 2018 at the time of retrieval. C) Borehole location of BPF1 marked by metal permafrost probe in September 2019. D) Example of core sample retrieved from BPF2 site.

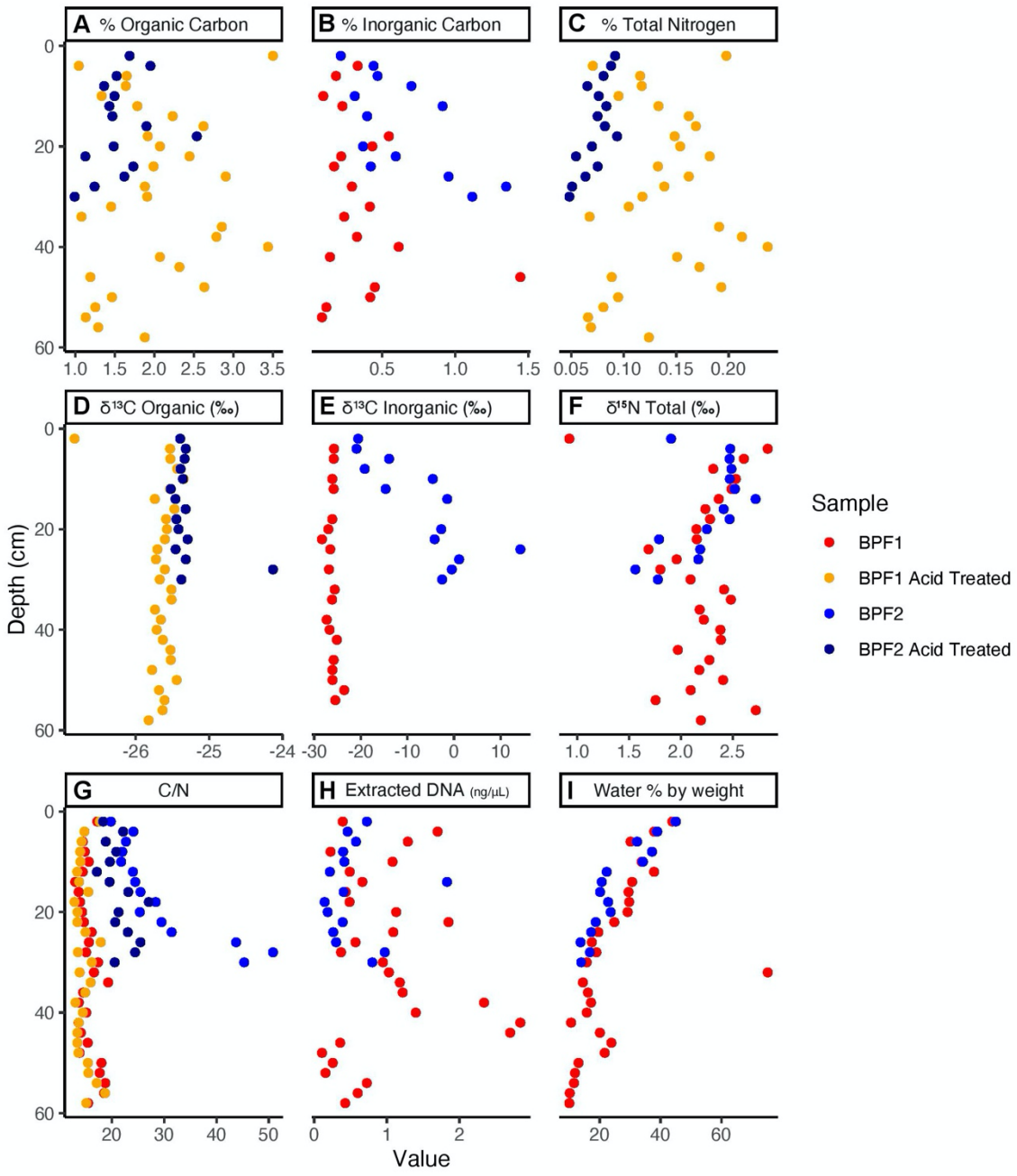


Figure 3-2: Elemental Analysis of acidified and non-acidified soil.

Elemental Analyzer data for BPF1 and BPF2 core samples before acid-treatment (red for BPF1 and blue for BPF2) and after acid-treatment (yellow for BPF1 and dark blue for BPF2).

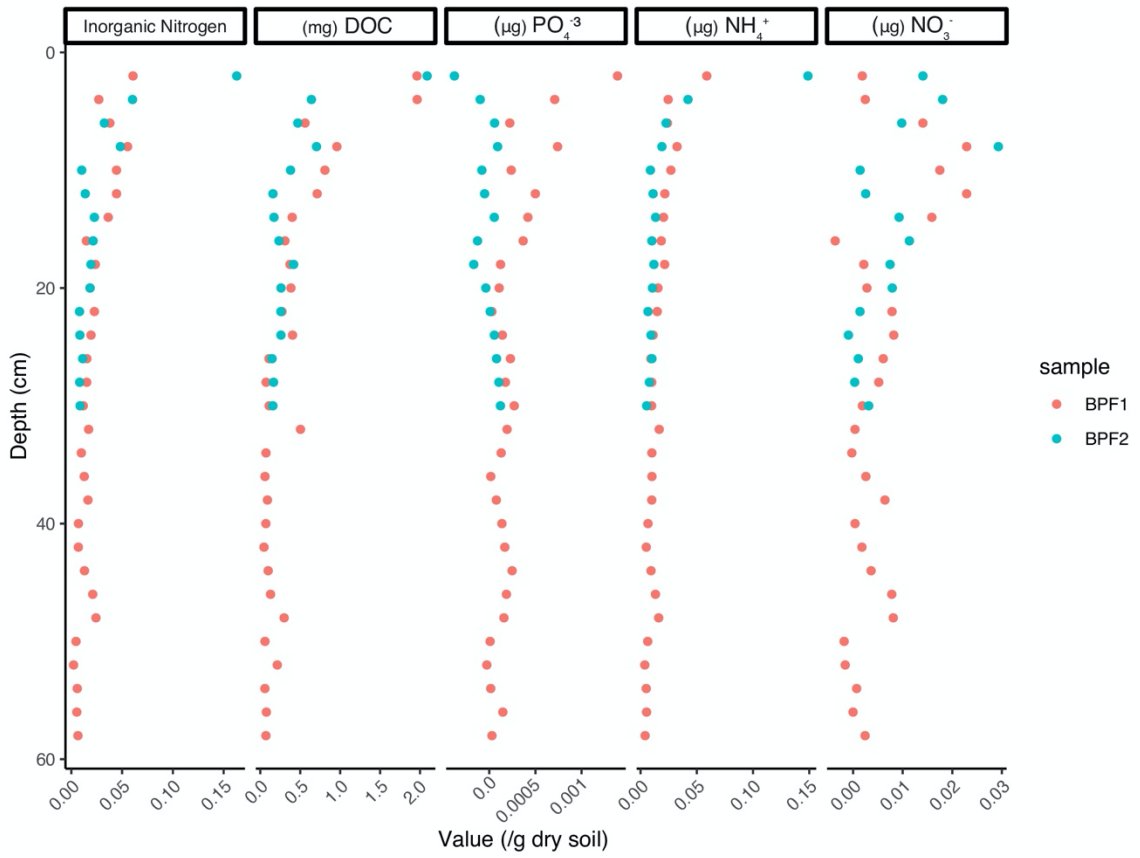


Figure 3-3: High resolution extractable compounds:

Values for PO_4^{3-}, NH_4^+, and NO_3^- measured in µg/g of dry soil, with the exception of DOC measured in mg/g of dry soil. Inorganic nitrogen is the sum of NH_4^+ and NO_3^-. Values below zero indicate an amount below the detectable level.

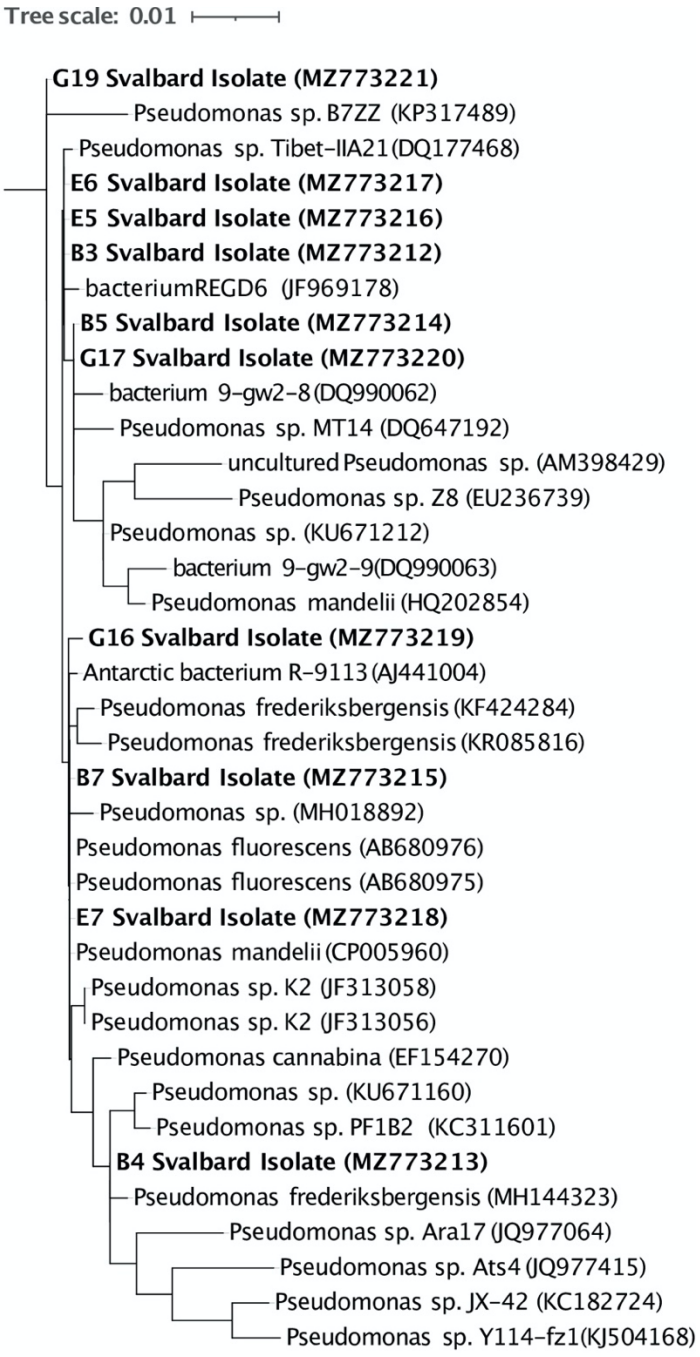
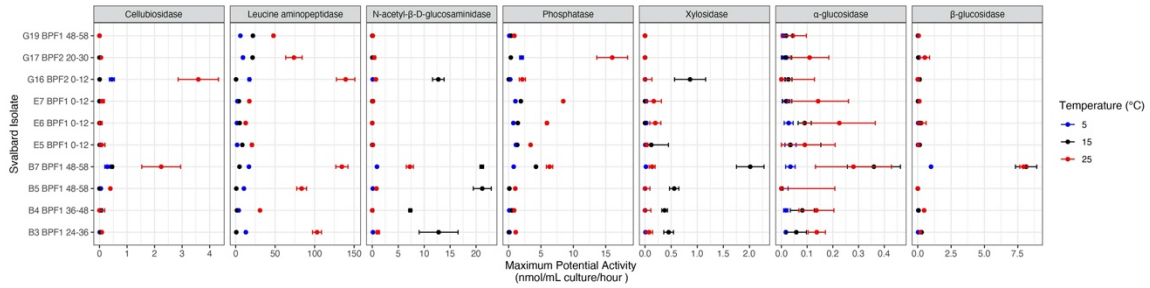


Figure 3-4: Full length 16S rRNA from Svalbard isolates phylogenetically compared to other Pseudomonas isolates

The 10 Svalbard isolates (in bold) were aligned using the Silva SINA (v1.2.11) and Arb-Silva was used to identify their closest relatives. A RAxML tree was visualized in iTOL (v5.7). The Genbank accession number for each organism is listed in parentheses. The tree is rooted with *Aurantimonas sp. (AB291857).*

A)



B)

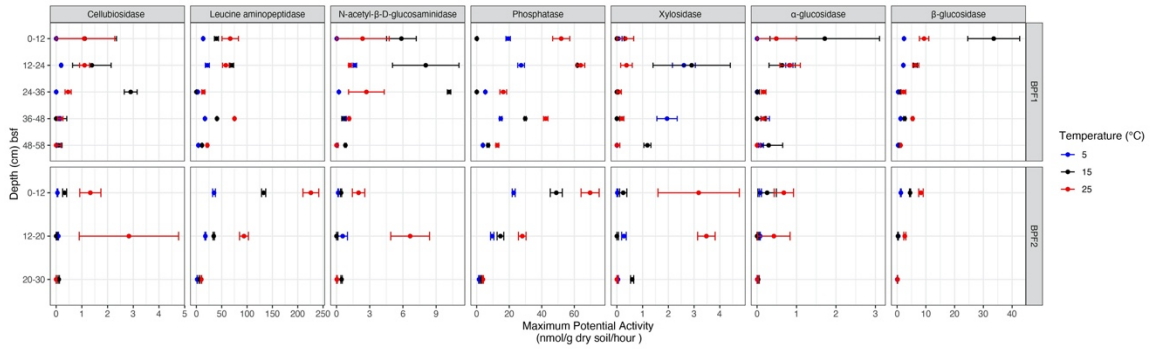


Figure 3-5: Maximum potential enzyme activity assay:

Each temperature (red for 25°C, black for 15°C, and bright blue for 5°C) for A) BPF1 and BPF2 and B) 10 cultured isolates were done in triplicates. Markers show the mean of triplicate measurements with error bars for one standard deviation. Enzymes are as follows: β-D-cellubiosidase (CB), leucine aminopeptidase (LAP), N-Acetyl-β-D-glucosaminidase (NAG), phosphatase (PHOS), β-xylosidase (XYL), α-glucosidase (AG), β-glucosidase (BG).

Table 3-1: Closest relative of Svalbard 16S rRNA genes from cultured *Pseudomonas* sp. isolated from bulk soil

Sample	Svalbard Site Origin	Percent Match	Organism Name	Location	First Author	NCBI Reference
B3	BPF1 24-36 cm	99%	*Pseudomonas silesiensis* strain ILQ215	Various soil samples from The Peruvian Andean Plateau	Carolina Chumpitaz-Segovia	(Published) June 23rd, 2020
E5	BPF1 36-48 cm	100%	*Pseudomonas* sp. strain PAMC 27331	Antarctic soil	H.J. Park	(Submitted) June 3rd, 2020
G17	BPF1 48 - 58 cm	99%	*Pseudomonas* sp. strain PAMC 27357	Frozen soil samples from Council, Alaska USA.	Hyoungseok Lee	(Sample Collected) June 29, 2012
B4	BPF1 48-58 cm	99%	*Pseudomonas mandelii* strain JZY4-67	QinLing Mountain China	R. Chen	(Accepted) February 26th, 2020
E6	BPF1 0-12 cm	99%	*Pseudomonas mandelii* strain UTB_118	Sediment samples from Station Juan Carlos I., Livingston Island, Antarctica	L. Ward-Bowie	(Sample Collected) February 25th, 2018
G19	BPF1 0-12 cm	99%	*Pseudomonas* sp. strain E1-4	Antarctic soil	M. Zhou	(Submitted) January 18th, 2017
B5	BPF1 0-12 cm	100%	*Pseudomonas* sp. strain PAMC 27292	Antarctic soil	H.J. Park	(Submitted) June 3rd, 2020
E7	BPF2 0-12 cm	100%	*Pseudomonas* sp. strain PAMC 27303	Antarctic soil	H.J. Park	(Submitted) June 3rd, 2020
B7	BPF2 20-30 cm	100%	*Pseudomonas mandelii strain BLH-Y1*	Qinghai-Tibet Plateau	Y. Wang	(Submitted) September 24th, 2013
G16	BPF1 48-58 cm	99%	*Pseudomonas mandelii* strain UTB_115	Sediment samples from Station Juan Carlos I., Livingston Island, Antarctica	L. Ward-Bowie	(Sample Collected) February 25th, 2018

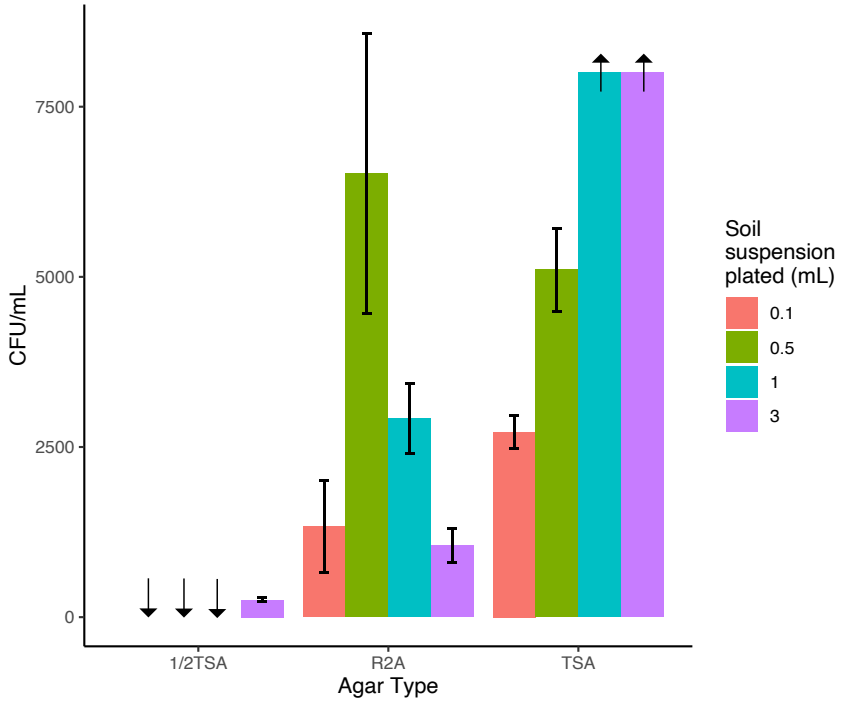
Table 3-2: MiSeq Metagenome Information

BPF Metagenome Library	BPF1 0-30cm	BPF1 30-58cm	BPF2 0-30cm
Number of Reads	25871532	8,366,690	13769728
# contigs	13,898	3,316	7525
# contigs > 10,000 bp	612	299	186
# contigs >= 100,000 bp	0	0	0
Largest contig	82,493	66,148	35,326
Total length	52,430,264	15,449,290	25,710,039
Total length >= 10,000 bp	9,664,112	5,805,672	2,525,799
Total length >= 100,000 bp	0	0	0
N50	3,705	6,093	3,278
N75	2,514	2,687	2,453
L50	3,591	545	2,309
L75	7,991	1,611	4,605
GC%	62.1	64	58.7
#N's	54,379	12,310	27,269
#N's per 100 kbp	103.7	79.7	106.1
Binned Contigs	10780(77.6%)	2870 (86.6%)	6763 (89.9%)
Unbinned contigs	3118(22.4%)	446 (13.4%)	762 (10.1%)
Contigs too short	0	0	0
Bins	11	3	6
High Quality Bins	0	0	0
Medium Quality Bins	2	0	0
Low Quality Bins	11	3	6

Table 3-3: Gene counts from whole metagenomes and whole genome sequences that encode for each enzyme.

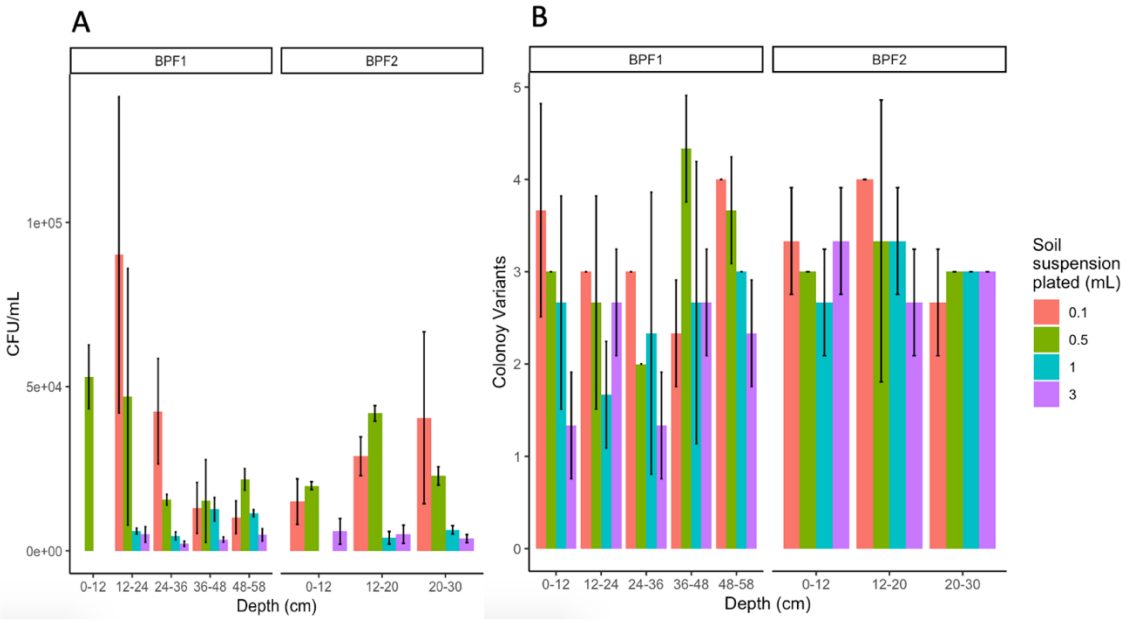
The list of genes contributing to these counts can be found in Supplemental Table 3-3

Enzyme	BPF10-30cm	BPF130-58cm	BPF20-30 cm	B3	B4	B5	B7	E5	E6	E7	G16	G17	G19
α-glucosidase	0	0	0	0	0	0	0	0	0	0	0	0	4
β-glucosidase	8	9	6	0	0	0	0	9	1	0	7	7	9
Cellubiosidase	6	6	6	2	1	1	2	2	0	2	2	2	5
Leucine Aminopeptidase	14	71	4	3	2	2	4	11	7	8	8	8	15
N-acetyl-β-D-glucosaminidase	4	4	2	1	1	1	1	2	1	2	1	1	1
Phosphatase	24	21	27	10	8	9	12	15	12	13	11	11	15
Xylosidase	0	0	6	2	2	2	4	4	1	3	3	3	14



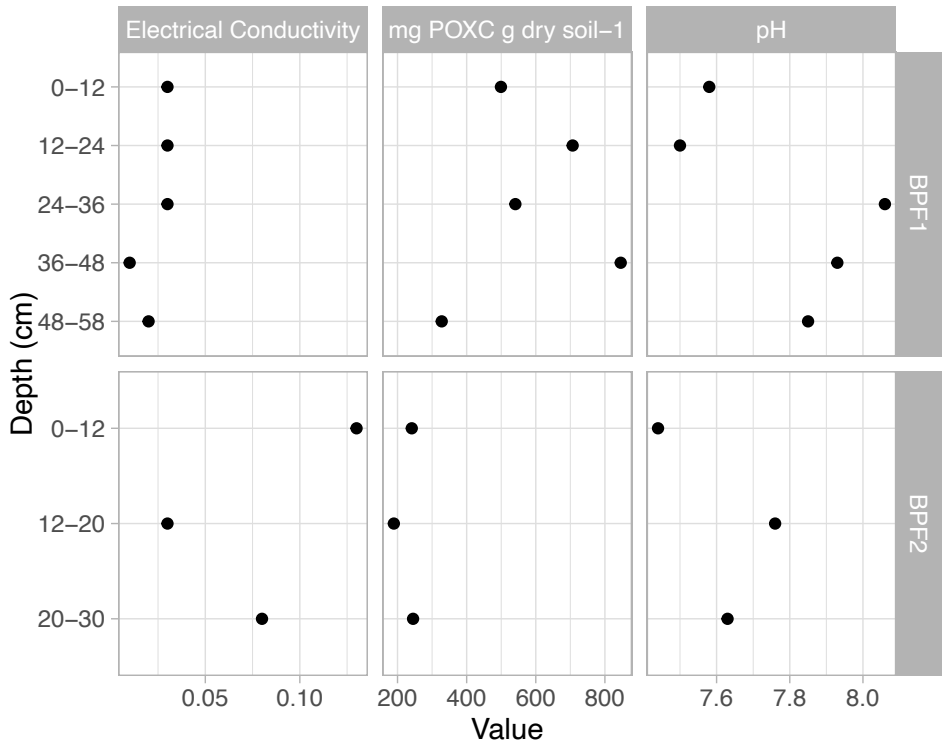
Supplemental Figure 3-1: Preliminary testing for mean CFU/mL from agars ½ TSA, R2A, and TSA with surface soil from BPF2.

Downward-facing arrows indicate samples that were too low to count, upward-facing arrows indicate samples that were too high to count.

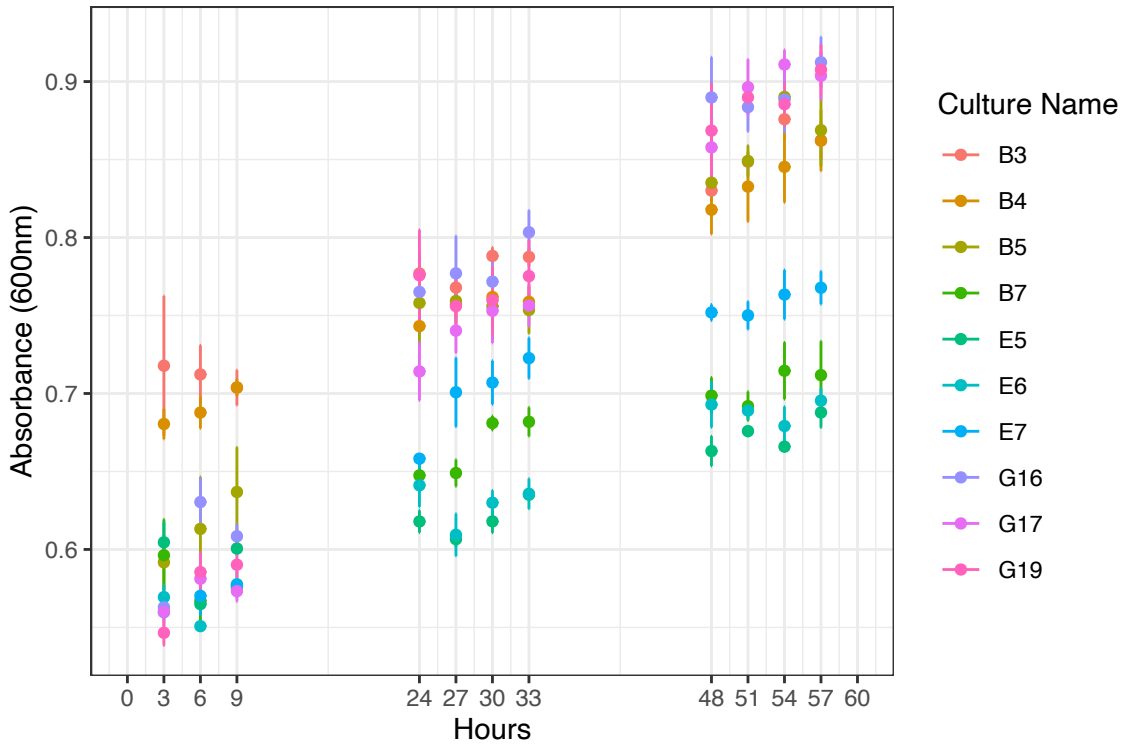


Supplemental Figure 3-2: Each interval from BPF1 and BPF2 were plated on R2A agar using the multiple dilution schemes.

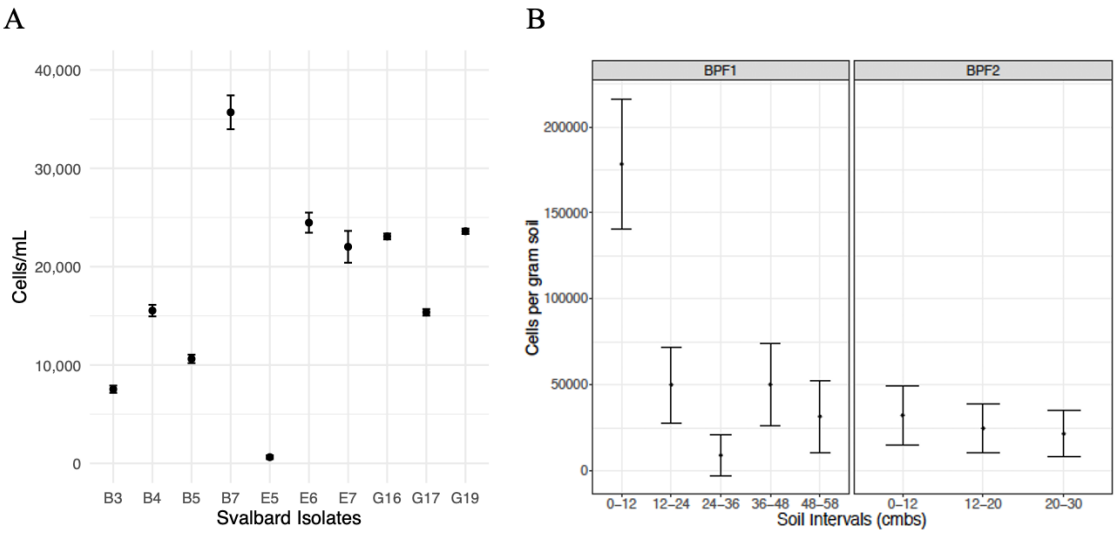
R2A plates were incubated at 4°C for at least three weeks. A) Mean CFU/mL values on each of the depth segments. B) Unique colony variations between the depth segments.



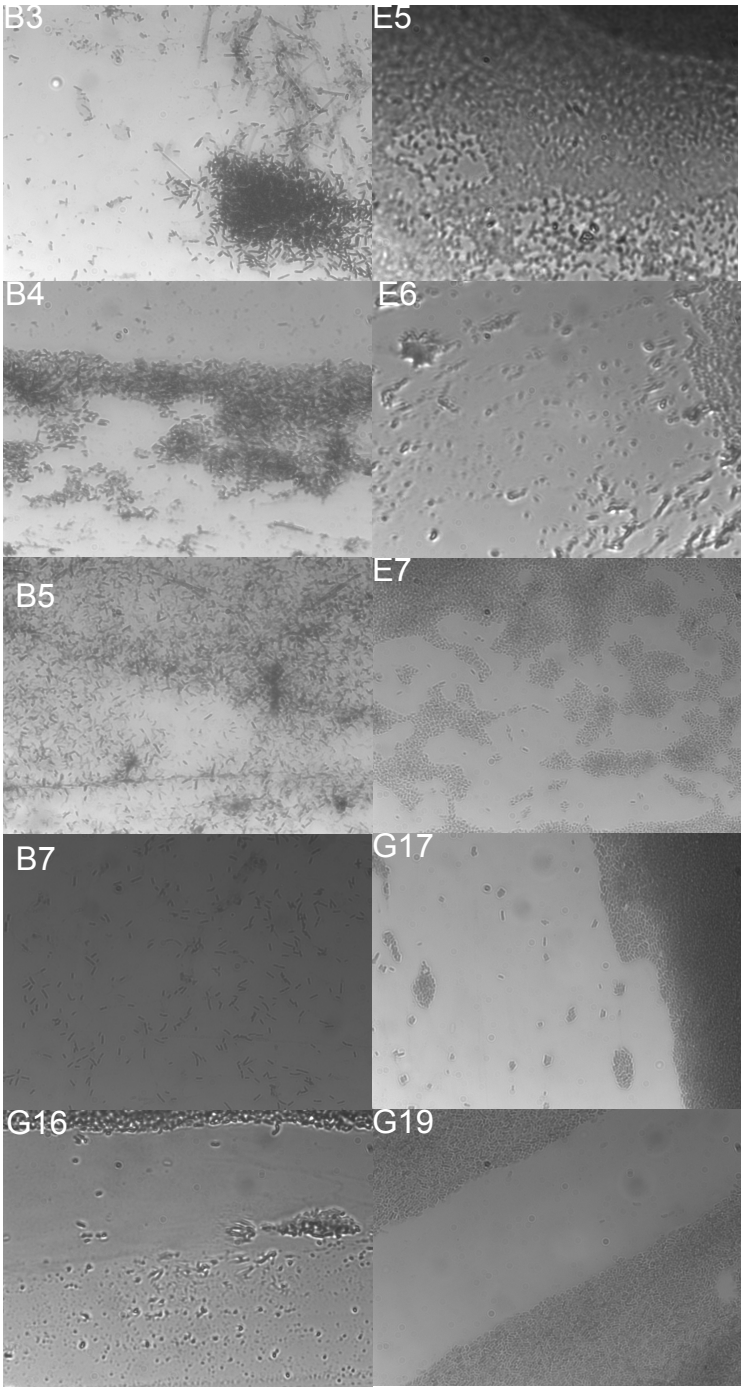
Supplemental Figure 3-3: Electrical conductivity, permanganate oxidizable carbon (POXC) for labile carbon, and pH for each soil interval.



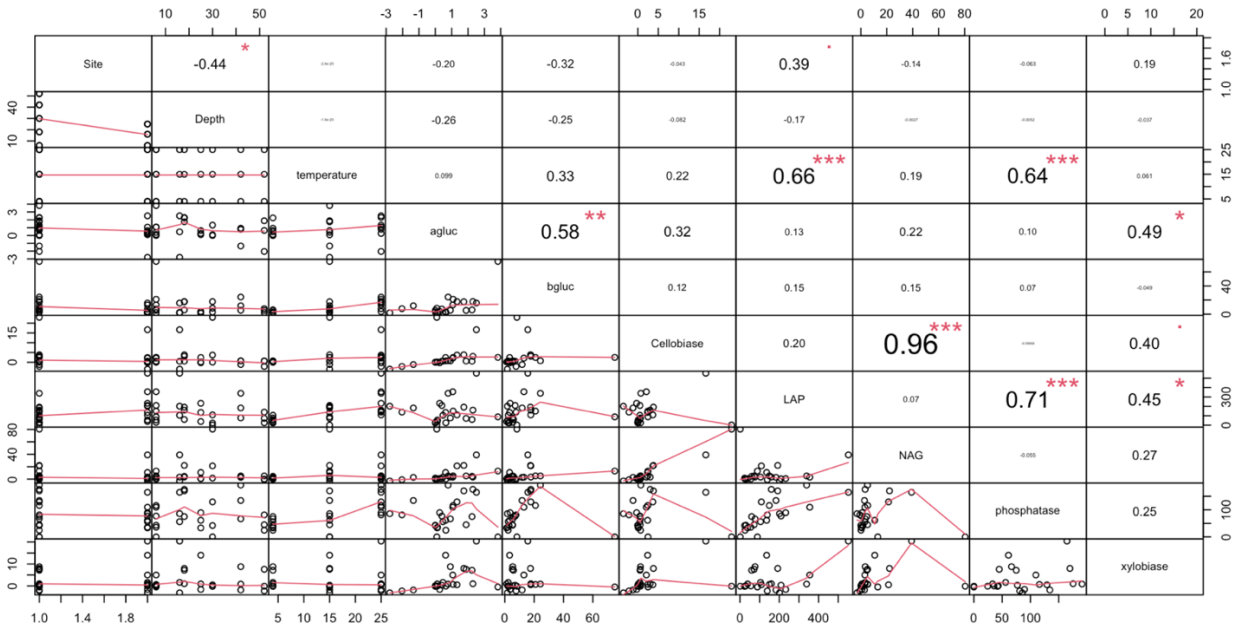
Supplemental Figure 3-4: 60-hour growth curve for the 10 isolates in R2A broth at 25°C.



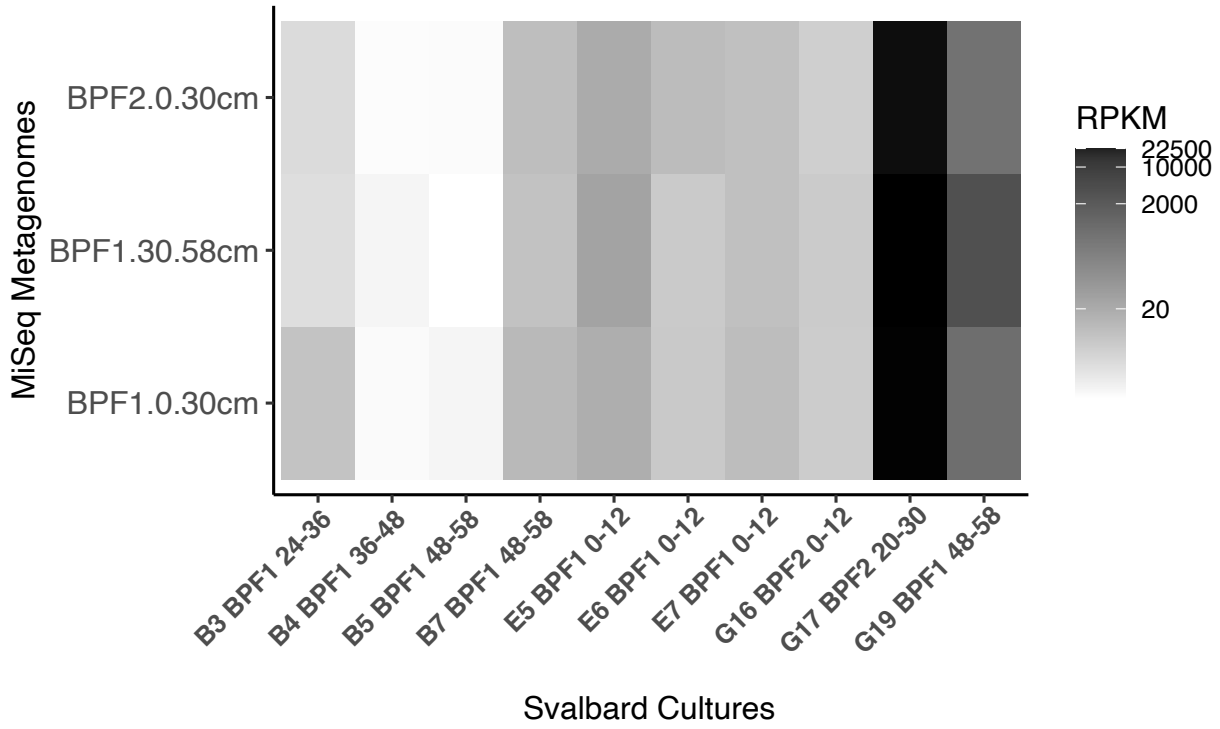
Supplemental Figure 3-5: Direct microscopic cell counts in A) 1mL of culture broth of Svalbard isolates, and B) Bulk soil. Cells were stained with SYBR Gold and 30 random fields of view were averaged for each sample. Error bars show one standard deviation about the mean.



Supplemental Figure 3-6: Each isolate at 10X under a Zeiss X2 Imager Microscope.



Supplemental Figure 3-7: Spearman correlations between all variables in the enzymatic activity experiment: Site, depth, temperature and the seven enzymes. The left side of the diagonal of the graph is the distribution of the corresponding variable as the x (above the graph) and the y (right of the graph). The correlation coefficient is reported by the number in the boxes in the right side of the diagonal. The red stars indicate p values of *=0.05, **=0.01, ***=0.001.



Supplemental Figure 3-8: Read mapping done in reads per kilobase per million mapped reads (RPKM) between the Whole Genome Sequenced Isolates and the MiSeq Metagenome Libraries.

Supplemental Figure 3-9: Enzyme activity was done in triplicate measurements at each temperature (color differences) for all samples.

A) Core site interval depths. B) 10 cultured isolates. Each enzyme's potential activity is plotted within its own detected range at nmols/sample/hour and normalized base on cell counts (see cell count supplemental methods) for each method. Negative activity values were determined to be zero.

Supplemental Table 3-1: Bulk soil average enzyme activity in nmol g dry soil^{-1} hr^{-1}.

AG; α-glucosidase, BG; β-glucosidase, CB; Cellubiosidase, LAP; Leucine aminopeptidase, NAG; N-acetyl-B-D-glucosaminadase, PHOS; Phosphatase, XYL; Xylosidase

Sample	°C	CB	LAP	XYL	AG	BG	NAG	PHOS
BPF1 0-12	4	-0.01891634	14.62690496	0.06712252	-0.000610205	2.610760061	0.05	43.13
BPF1 12-24	4	0.001571204	0.174834495	0.02098741	0.006594753	0.017448974	4.54	74.93
BPF1 24-36	4	3.93565E-05	0.332830972	-0.002263	-3.57181E-17	0.042898627	1.67	41.96
BPF1 36-48	4	0.002842535	0.336585927	0.03858498	0.004044226	0.023887859	3.42	66.59
BPF1 48-58	4	0.175759748	6.618544147	-0.0205911	0.167670387	0.603024985	0.16	24.50
BPF1 0-12	15	2.314098312	82.56388243	-0.4239038	3.571911254	69.95803283	13.08	0.08
BPF1 12-24	15	0.001014161	0.051088281	0.00212189	0.000464075	0.004664384	22.12	169.85
BPF1 24-36	15	0.940767113	0.042647753	-0.00404693	-3.57181E-17	0.340664943	80.45	0.30
BPF1 36-48	15	-0.0031029	0.481967099	-0.00475387	-0.003528958	0.031168132	2.83	134.37
BPF1 48-58	15	0.002586361	0.310925698	0.03289366	0.00815832	0.022739152	5.14	44.90
BPF1 0-12	25	0.004005994	0.245156642	0.00118744	0.001787691	0.034435884	5.25	115.94
BPF1 12-24	25	0.005146887	0.270595539	0.00174025	0.003900555	0.03021199	3.36	175.41
BPF1 24-36	25	0.00133276	0.039399493	0.00020545	0.000479372	0.00642148	21.38	129.23
BPF1 36-48	25	0.000752293	0.345531298	0.00090902	0.000797073	0.024754033	5.13	191.20
BPF1 48-58	25	0.034950642	-2.243781606	0.03614213	0.032964809	-0.124710218	-0.76	81.22
BPF2 0-12	4	0.009333702	6.486834494	0.00339629	0.01563761	0.241405681	0.24	35.78
BPF2 12-20	4	0.002694385	0.568340866	0.00875777	0.002515573	0.012870465	3.28	56.14
BPF2 20-30	4	0.004703593	1.632476249	0.04018685	0.004868634	0.052399687	0.72	33.20
BPF2 0-12	15	0.010434534	4.189273462	0.00777444	0.008059451	0.14229918	0.65	76.93
BPF2 12-20	15	0.014199689	-0.809565853	0.01241964	0.011199037	-0.007110017	-2.90	85.79
BPF2 20-30	15	0.003631818	0.205756807	0.02122036	0.000956156	0.005028672	10.75	61.27
BPF2 0-12	25	0.006086564	1.033641051	0.01454205	0.003103802	0.037937288	3.13	109.63
BPF2 12-20	25	0.002584444	0.085782999	0.0031755	0.000390073	0.002441777	39.10	164.57
BPF2 20-30	25	-0.00346513	2.713889143	-0.01810231	0.003106666	0.042537433	0.96	89.51

Supplemental Table 3-2: Culture average enzyme activity in nmol ml^{-1} hr^{-1}

AG; α-glucosidase, BG; β-glucosidase, CB; Cellubiosidase, LAP; Leucine aminopeptidase, NAG; N-acetyl-B-D-glucosaminadase, PHOS; Phosphatase, XYL; Xylosidase

Sample	°C	AG	BG	CB	LAP	NAG	PHOS	XYL
B3	5	5.889E-05	9.343E-05	3.538E-05	4.416E-02	4.552E-04	4.701E-04	3.342E-05
B4	5	5.837E-05	2.041E-04	-3.852E-04	1.322E-02	-1.535E-04	3.647E-04	-2.478E-04
B5	5	-9.722E-05	-2.794E-05	1.692E-04	3.581E-02	3.440E-04	4.429E-04	-7.986E-05
B7	5	1.201E-04	3.401E-03	9.608E-04	5.764E-02	3.106E-03	2.713E-03	5.822E-05
E5	5	0.0001213	6.7017E-05	3.5969E-05	0.00650309	0.00013235	0.00388765	2.073E-05
E6	5	9.649E-05	1.119E-04	1.265E-05	5.423E-03	3.963E-06	2.539E-03	8.462E-05
E7	5	6.120E-05	4.798E-05	4.670E-05	7.436E-03	6.354E-05	3.617E-03	7.175E-05
G16	5	-1.399E-04	-1.315E-04	1.538E-03	5.970E-02	3.001E-04	9.094E-04	-1.442E-04
G17	5	4.695E-05	2.281E-04	2.328E-05	3.164E-02	1.610E-04	6.859E-03	-1.876E-05
G19	5	1.905E-05	2.821E-05	-1.738E-05	2.054E-02	4.570E-05	3.840E-04	-3.008E-05
B3	15	4.9558E-05	0.00024412	3.9481E-05	0.00068525	0.01095167	4.2243E-05	0.00038771
B4	15	6.9572E-05	3.5669E-05	4.8562E-05	0.00139917	0.00627899	0.00043975	0.00032164
B5	15	-9.328E-05	3.8003E-06	-5.737E-05	0.00068897	0.01816254	0.00010051	0.00047831
B7	15	0.0003081	0.00696319	0.00039172	0.0041447	0.01807888	0.00362262	0.0017266
E5	15	2.9345E-05	0.00012179	6.9347E-06	0.00720835	-3.654E-05	0.00115197	0.00010263
E6	15	7.7473E-05	0.00012476	-3.65E-05	0.00432777	-5.304E-05	0.00122157	-8.357E-05
E7	15	1.9308E-05	2.6815E-05	-1.094E-05	0.00348942	2.6143E-05	0.00162423	-1.408E-05
G16	15	2.3261E-05	0.0001138	3.5238E-06	0.00040202	0.01090368	2.5603E-05	0.00073921
G17	15	1.641E-05	3.3509E-05	-9.295E-06	0.01834945	1.8086E-05	0.00029233	-9.137E-05
G19	15	1.6029E-05	2.3596E-05	-1.312E-05	0.01871441	1.4081E-05	0.00033731	-3.529E-05
B3	25	1.7005E-05	2.1256E-05	5.2323E-05	0.00618252	2.9759E-05	0.0003963	-2.191E-05
B4	25	0.00011102	0.00024589	-0.0001608	0.00959177	-2.009E-05	0.00077357	0.00011966
B5	25	5.3883E-05	8.9426E-05	-0.0001808	0.00493833	8.84E-05	0.00048847	7.2908E-05
B7	25	4.2544E-05	0.00184824	0.00117401	0.01776767	0.00154004	0.000628	0.0001662
E5	25	3.9015E-05	3.0836E-05	3.0413E-05	0.00881604	6.6514E-05	0.00145725	1.2786E-05
E6	25	-1.246E-05	3.3657E-05	-1.217E-05	0.00477814	4.1836E-05	0.00124977	-2.416E-05
E7	25	2.3731E-05	5.5964E-05	1.5751E-05	0.01315247	7.8339E-05	0.00138254	0.00014377
G16	25	7.9532E-05	8.1324E-05	2.6517E-05	0.01605988	4.2929E-05	0.00076292	4.338E-05
G17	25	4.7426E-05	9.8148E-05	0.00012399	0.00632102	0.00011569	0.00068852	0.00044066
G19	25	4.6859E-05	5.8478E-05	7.6177E-05	0.00772018	3.8212E-05	0.00072055	1.0921E-05

Supplemental Table 3-3: List of the gene names from UniProt and or CAZY that were used in the total counts for the metagenome and whole genome isolates.

Gene product name	Enzyme
cellobiose phosphorylase	Cellubiosidase
Cellobiose 2-epimerase	Cellubiosidase
pepA/Cytosol aminopeptidase	Leucine aminopeptidase
Putative Cytosol aminopeptidase	Leucine aminopeptidase
Aminopeptidase	Leucine aminopeptidase
Putative lipoprotein aminopeptidase LpqL	Leucine aminopeptidase
Cytosol Aminopeptidase	Leucine aminopeptidase
Aminopeptidase N	Leucine aminopeptidase
Dipeptidyl aminopeptidase 4	Leucine aminopeptidase
D-aminopeptidase	Leucine aminopeptidase
Dipeptidyl aminopeptidase BIII	Leucine aminopeptidase
Dipeptidyl aminopeptidase BII	Leucine aminopeptidase
Dipeptidyl aminopeptidase BI	Leucine aminopeptidase
N-acetylglucosaminyldiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	N-acetyl-B-D-glucosamididase
tagA	N-acetyl-B-D-glucosamididase
N-acetyltransferase domain-containing protein	N-acetyl-B-D-glucosamididase
Glycoside hydrolase superfamily	N-acetyl-B-D-glucosamididase
N-acetyl-beta-D glucosaminidase	N-acetyl-B-D-glucosamididase
Beta-N-acetylglucosaminidase	N-acetyl-B-D-glucosamididase
Nag	N-acetyl-B-D-glucosamididase
Phosphoserine phosphatase 1	Phosphatase
Phosphoglycolate phosphatase	Phosphatase
Alkaline phosphatase D	Phosphatase
Alkaline phosphatase	Phosphatase
Phosphoadenosine phosphosulfate reductase	Phosphatase
Phosphoserine phosphatase RsbU	Phosphatase
Phosphoserine phosphatase SerB2	Phosphatase
Endo-1,4-beta-xylanase A/xynA1	Xylosidase
Beta-xylosidase/xynB	Xylosidase
Bifunctional xylanase/deacetylase/xynD	Xylosidase
Xyloglucanase	Xylosidase
Beta-xylosidase	Xylosidase
Glucan 1,4-alpha-glucosidase SusB	Alpha glucosidase
Beta-glucosidase BoGH3A	Beta glucosidase
Thermostable beta-glucosidase B	Beta glucosidase
Beta-glucosidase A	Beta glucosidase

Supplemental Table 3-3 continued:

Gene product name	Enzyme
Beta-N-acetylglucosaminidase/beta-glucosidase	Beta glucosidase
Beta-glucosidase B	Beta glucosidase
Glucan endo-1,3-beta-glucosidase A1	Beta glucosidase
Beta-glucosidase BoGH3B	Beta glucosidase
Beta-glucosidase	Beta glucosidase
putative 6-phospho-beta-glucosidase	Beta glucosidase
Oligo-1,6-glucosidase	Beta glucosidase
cellobiose phosphorylase	Cellubiosidase
Cellobiose 2-epimerase	Cellubiosidase
pepA/Cytosol aminopeptidase	Leucine aminopeptidase
Putative Cytosol aminopeptidase	Leucine aminopeptidase

CHAPTER 4
METAGENOME ASSEMBLED GENOMES FROM ACTIVE LAYER IN NY ÅLESUND, SVALBARD (79°N) SHOW POPULATION DYNAMICS RELATED TO HIGH DEPTH RESOLUTION

This original research is being prepared for future publication submission with at least the following authors:
Katie Sipes, Michael Tamino, Tatiana A Vishnivetskaya, Karen G. Lloyd

I contributed to this research by taking field samples, extracting DNA, sequencing samples, organizing data, analyzing datafiles, generating hypothesis and assessing final conclusions.

MT assisted with python scripts

Abstract

Microbial diversity in the High Arctic (latitude >76°N) is of growing interest as habitats are being affected by microbial activity due to climate change. The microbial communities have not been characterized throughout the stratification of the active layer soil discretely. We tested if organismal abundance and dominance was related to location or depth by analyzing metagenome assembled genomes (MAGs) from 56 thin-sliced intervals of five active layer sites from Ny Ålesund, Svalbard (79°N). Two sites were collected when the active layer was completely frozen in April 2018 and three were collected when thawed in September 2019. From this, 169 medium quality MAGs were curated from 1.3TB of metagenomic libraries. Microbial abundance was assessed with read recruitment across all locations and showed two populations differentiating by depth. Different classes of the *Acidobacteriota* and *Actinobacteriota* phyla dominate different locales between the five sites. The *Acidobacteriota-Thermoanaerobaculia* MAGs dominate the upper 12 cm in the frozen core and in the 43-45 cm of thawed soil. Another *Acidobacteriota* class*, Blastocatellia,* is present near the depths where the *Thermoanaerobaculia* is but is the most abundant in the upper 16 cm of a different frozen core site, where no *Thermoanaerobaculia* are present. The other prevalent phylum, *Actinobacteriota*, highest abundance occurs inversely of where the *Acidobacteriota* are dominant. Abundance of *Chloroflexota* MAGs were higher in the thawed soils and correlate with the *Actinobacteriota* population. The high correlation coefficient between MAGs in the same site (>0.7 Spearman correlation) and high microbial diversity (Simpson D = 0.6 to 0.8) at the sites indicate that microbial interactions are best assessed in high depth resolution (2 cm increments). We show that microbial populations in the Svalbard active layer shift by depth and season, maintaining unique populations of microbes existing at all depths, year-round.

Introduction

Arctic soils are in danger of irreversible change due to the climate related crisis of this century. The high Arctic Archipelago, Svalbard (79°N), is affected by warming amplification as it receives warm water from the Atlantic Ocean and warm air from greenhouse gas intensification at the pole. As the global temperature continues to rise, permafrost in Svalbard will thaw faster than its current ~1 vertical cm a year (1–3). Geological models of how the destabilizing soil will affect commerce and the greenhouse gas budget often include microbial soil organic matter degradation (4, 5). The warm permafrost of Svalbard, which is permafrost that stays frozen at temperatures -2°C to 0°C (6), will be the first of the polar environments to thaw with the increasing temperatures. There are two main possibilities for the fate of microbes living within Arctic soil; i) the microbes exposed to the increasing temperatures will respond by becoming more active and thus release more greenhouse gases to the atmosphere, ii) this increased microbial activity will act as a carbon sink and metabolize greenhouse gases or iii) the

endemic soil microbes will perish due to habitat change and a different community will succeed it.

Cryoturbation occurs in the active layer, where soil predisposed to water saturation is moved down as it frosts and freezes seasonally (7). Cryoturbation in the Arctic was particularly intense during the mid-Holocene, when these environments experienced a period of warming and the metabolically accessible soil organic carbon from the surface was transferred to the thawing permafrost (7). This means that the increased temperatures due to anthropogenic climate change may result in soil organic material incorporation from the active layer to the permafrost through cryoturbation in the Arctic. Organisms at different soil strata will be able to access the redistributed soil organic carbon at that has been moved to a different layer due to this cryoturbation. The redistribution of the thawed carbon-rich soil into the colder temperature soils, near the permafrost, may slow microbial decomposition (7) or it may increase microbial activity and result in faster organic matter degradation (8). How organisms are vertically distributed throughout the active layer is essential for understanding how the community could respond to warming at different depths.

Previous bacterial taxonomy studies in Arctic cryosols have focused on the upper surface of the active layer at 1-5cm (9), <20cm (10, 11), 30cm (12, 13), 40cm (14), and 60cm (13, 15). All these studies have found the presence of bacterial taxa from *Actinobacteriota, Acidobacteriota* and *Chloroflexota* phyla. These ubiquitous organisms are most commonly found in soil habitats (16, 17) and have been shown to contain environmentally relevant genes, such as those for methanogenesis (12), nitrogen-fixation (18), and plant-degrading enzymes (19). These studies have established these three phyla as key players in the microbial response to global climate change in bulk soils. However, microbial community variations over the small spatial scales that are relevant to these vertically layered soil is unknown (20). Further, the microbial community of the whole active layer cannot be assumed based on the microbes present in the upper surface layers. This is because if the microbial community in the deeper active layer is adapted to those environmental conditions then its community composition and metabolic properties cannot be assumed based on studies of the microbial community found at the surface horizons (21). We examine these elements in detail with medium quality MAGs. Evaluating the microbial communities at a high depth resolution (1 or 2 cm increments) allows investigation of microbial population response to location, depth and seasonal differences.

Studying microbes within cryosols is difficult due to low DNA concentrations, dominance of difficult-to-culture organisms, and fluctuating environmental characteristics. These factors make retrieval of microbes for culture or *in situ* microbial activity investigation difficult. One way to analyze the extant microbial DNA in the environment is to use metagenomic-based approaches. In this way, the bulk DNA and the assembled organisms (metagenome assembled genomes – MAGs) can be analyzed to assess the entire microbial community's taxonomy and metabolic potential. Further, with bioinformatic analysis, we can

investigate the community structure at different sites and how the broader total DNA content compares to the MAGs.

Chapter three of this dissertation investigated potential energy substrates and genetic abilities of the soil microbes in the same Svalbard active layer through combined environmental metagenomes and ten isolated *Pseudomonas* cultures. From this investigation, two of these active layer sites (BPF1 and BPF2) have been characterized geochemically with carbon and nitrogen percentages, fractions of organic and inorganic carbon isotopes and carbon to nitrogen ratios. The isotopic fractionation of the inorganic carbon displayed signatures likely attributed to autotrophic microbial metabolism dominating in BPF2 and heterotrophic metabolism in BPF1. Further, enzymatic activities from bulk soil and culture isolates, paired with metagenomes from these sites showed paralleling results in ability to utilize amino acids as a carbon and nitrogen source. The enzymatic assays also showed activities in colder temperatures for enzymes that degrade nitrogen and plant polymers. This study lacked metagenome assembled genomes (MAGs) to examine how organisms were interacting with each other and the environment. In the study presented here, we sampled Svalbard active layer at thin sliced depth intervals (1 cm to 2 cm) to investigate the microbial communities throughout five sites, at two seasons: April 2018, when the active layer was frozen and September 2019 when the active layer was thawed. A total of 56 individual samples in the Bayelva plain in the Laugerhain glacier moraine were sequenced on Illumina NovaSeq, totaling 1.3TB of metagenomic reads. These microbes in the active layer are assumed to be responsible for the future degradation of the soil organic matter. Understanding how these communities fluctuate in vertical and seasonal space is important to estimate how microbes will influence the rate of soil thaw.

Materials and methods

Two cores, 84m apart, were collected in active layer soils April 2018 (BPF1:58cm and BPF2:30cm deep) as described in Sipes *et al.,* in review, from Ny Ålesund, Svalbard, (79°N). An additional three sites (BPF3, 4 and 5) were collected in September 2019 with a shovel ~150m to 225m from the first two sites (Figure 4-1). BPF1, 2 and 3 samples were divided in 2cm increments, while BPF4 and 5 were divided into 1cm intervals. BPF1 and 2 were sampled as described in Sipes *et al,* in review (Chapter Three), with sterile techniques in Kings Bay AS, Marine Bay Lab in Ny Ålesund, Svalbard. BPF3, 4 and 5 were removed from a wall of soil pit and put into sterile bags for DNA extraction at University of Tennessee, Knoxville. Approximately 0.5g of soil was taken from the center of the soil interval to minimize contamination for DNA extraction with with a QIAGEN PowerSoil DNA Kit (Hilden, Germany) and sequenced on an Illumina NovaSeq at SciLifeLab in Stockholm, Sweden. The metagenomic quality control, assembly, and binning computational work was performed on the resources at the Infrastructure for Scientific Applications and Advanced Computing (ISAAC) supported by the University of Tennessee, Knoxville, USA. The 56 libraries were assembled and binned separately. MetaWrap v1.3.2 (22) was used for read quality control, assembly with

metaspades (23), binning with maxbin2 (24) and metabat2 (25), refined on a medium MAG quality category (26) of ≥50% completeness and <10% contamination scores via Checkm v1.4.0 (27), and taxonomically identified with GTDB-Tk v1.5.1 (28). MAG distribution was computed with metaWRAP's quant_bins module which assesses the contigs coverage based on length-weighted averages, standardized per million reads for each sample. Total MAG abundance distribution across each of the 56 samples was compared using a Spearman correlation to construct a network analysis in python3 with matplotlib. Phyla level, given by GTDB-Tk, were grouped and total abundance across all sites were compared with a Spearman correlation for phyla level network analysis. Shannon and Simpson diversity indices at each site was calculated at the phyla and class level. MAGs were annotated with PROKKA v1.14.6 (29) and assessed for genes encoding enzymes tested in chapter three (Sipes et al, in review). These enzymes were: α-glucosidase (AG), β-glucosidase (BG), cellubiosidase (CB), leucine aminopeptidase (LAP), N-Acetyl-β-D-glucosaminidase (NAG), phosphatase (PHOS), and xylosidase (XYL). List of gene names used to infer transcription of these enzyme can be found in Table 3-S3.

Results

Metagenomic libraries

Metagenomic sequencing, assembly, and binning yielded 169 individual MAGs of medium quality (30) (Figure 4-2 and Supplemental Table 4-1). Of the 56 samples assembled, 52 yielded MAGs (exceptions were BPF2 (10-12cm and 18-20cm, BPF4 1-2 cm, and BPF5 43-45cm) (Figure 4-2). The percentage of reads from the assemblies that were binned into MAGs varied from 7.3% at BPF3 18-20 to 90.5% at BPF1 0-2cm (Figure 4-2). Neither MAG size, number of MAGs, nor MAG completeness trend with percent of assembled reads binned (R^2=<0.2, Table 4-2). MAG size overall increased with completeness ($R^2 = 0.61$, $p < 10^{-16}$, Supplemental Figure 4-2).

Taxonomy distribution

There were 12 GTDB-Tk identifiable phyla within the 169 medium quality MAGs (Figure 4-3). *Acidobacteriota* had the most individuals (n=60) from 5 of the 13 classes: *Acidobacteriae* n=1, *Blastocatellia* n=37, *Mor1* n=1, *Thermoanaerobaculia* n=8, *Vicinamibacteria* n =13. *Acidobacteriota* have been reclassified from *Acidobacteria* by RED scores (28). *Actinobacteriota* was the phylum with the second highest number of individuals (n=52). These 52 MAGS spanned four of the nine designated families: *Acidimicrobiia* n= 3, *Actinomycetia* n=2, *Thermoleophilia* n = 45, and *UBA4738* n=2. The phylum *Chloroflexota* had 24 MAGs in it from three classes; *Anaerolineae* n = 2, *Dehalococcoidia* = 1, *Ellin6529* n = 2. 16 MAGs were identified as *Proteobacteria from classes Alphaproteobacteria* n=5 and *Gammaproteobacteria* n=11. *Gemmatimonadota* had 3 MAGs all from the *Gemmatimonadetes*

class. *Planctomycetota* had 5 MAGs from two classes, *Phycisphaerae* n=3 and Planctomycetes n=2. *Verrucomicrobiota* phyla had three MAGs, all from the *Verrucomicrobiae* class. The only archaeal MAGs identified were 2 *Thermoproteota,* both in the *Nitrososphaeria* class. There were four singleton MAGs from the phyla *Bacteroidota, Cyanobacteria, Eisenbacteria, and Methlomirabilota.* (Figure 4-3). BPF1is the richest site at the phyla and class level, but the Simpson index is only the highest at the phyla level (D = 0.77) (Table 4-2). At the class level, BPF4 has the highest Simpson diversity (D = 0.91). BPF1 and BPF4 had the highest Shannon-Weiner diversities at the phyla and class level (phyla: H = 1.7 and 1.4, class: H= 2.2 and 2.1, respectively, Table 4-2).

MAG read mapping to all metagenomic samples and Spearman correlation

Almost all MAGs showed some read recruitment from metagenomes from other sample sites and/or depths. Every MAG's highest read recruitment (abundance) was from their origin sample depth. Only 20 MAGs (9 *Actinobacteriota*, 3 *Chloroflexota*, 1 *Proteobacteria*, 1 *Gemmatimonadota*, 2 *Plactomycetota*, 1 *Eisenbacteria* and 3 *Acidobacteriota*) had reads recruit to all 56 samples, while most MAGs were constrained to a location or depth (Figure 4-4). A spearman correlation coefficient was used to cluster MAGs based on their read abundances across all samples. Stronger correlation coefficients (higher than 0) between MAGs resulted in closer clustering in the dendrogram (Figure 4-4). Dendrogram nodes do not always cluster by phylum or MAG origin site. Most MAGs were present in at least two sites with few MAGs from the same phyla specific to a single site: 11 *Actinobacteriota-Blastocatellia* MAGs were only in BPF2; 12 *Acidobacteriota-Vicinamibateria* MAGs were only below 22 cm in BPF1; 8 *Actinobacteriota-Blastocatellia* MAGs were only in the upper 32 cm of BPF1 and below 46 cm; 6 *Actinobacteriota-Thermleophilia* MAGs were only below 12 cm in BPF1; 5 *Proteobacteriota-Gammaproteobacteria* MAGs were present 24-40 cm and 50-58 cm in BPF1; 3 *Planctomycetota-Phycisphaerae* MAGs were in low abundance and only present above 50cm in BPF1 (Table 4-S2).

MAG abundance distributions correlated to other MAGs from the same location or depth but not from different sites at a spearman correlation coefficient ≥ 0.7 (Figure 4-5A). Some 2 cm interval sample depths show high correlations (red color on Spearman correlation, Figure 4-5A) between the MAG abundance distributions, but not all MAGs have abundance distributions that correlate with other MAGs at their depth. At BPF1, depths between 16 and 48 cm contain a large network of 58 MAGS that are correlated to at least two other MAGS. Two *Acidobacteriota* in this cluster each correlate with 10 other MAGs. BPF2 had similar results, with a tighter cluster of co-correlating MAGs at middle depths (10-26 cm) while MAGs from samples above and below it do not correlate with MAGs from other depths. The 10 MAGs from BPF3 correlate within the same depth and this is the only site with correlations between MAGs at the surface and the bottom sample depths. On the contrary, BPF4 has three groups of correlating MAGs separated by depth with the tightest correlation between five MAGs at 43-45

cm. The 6 MAGs in BPF5 are only at the surface samples, with no MAGs coming out of the deepest sample depth. All six are not correlated together but one MAG from each depth is connected to another MAG in the neighboring sample depth.

At the phyla level, a co-correlated group contains *Actinobacteria, Chloroflexota,* and *Proteobacteria. Chloroflexota* and *Actinobacteria* have the highest correlation (Spearman coefficient = 0.9) (Figure 4-5B). Another co-correlated group contains *Actinobacteria, Chloroflexota,* and *Methylomirabilota*. The *Acidobacteriota, Thermoproteota*, and *Gemmatimonadota* are correlated, but the latter two phyla are not. *Verrucomicrobiota, Plantomycetota,* and *Bacteroidota* also make a co-correlating group with the latter two phyla not correlated (Figure 4-5B).

Top three phyla summed reads

The top three phyla with the most MAGs were *Actinobacteriota, Acidobacteriota* and *Chloroflexota* (Figure 4-3). The summation of the reads recruited to each of these phyla, separated by class, show how classes change between the thin slice depth increments in the sites (Figure 4-6). *Acidobacteriota-Thermoanaerobaculia* has the highest amount of reads above 12 cm in BFP1 (191 in 2-4 cm, 148 in 6-8 cm, 71 in 8-10 cm: Table 4-S1). In BPF2, 3 and 5, *Thermoanaerobacullia* has < 3 reads but has a spike (35 reads) in BPF4 43-45 cm. A different *Acidobacteriota* class, *Blastocatellia,* has a spike in read abundance in BPF1 18-20 cm (81 reads). *Blastocatellia* dominates the read abundance in the upper 16 cm at BPF2 (33 in 0-2 cm, 65 in 6-8 cm, 42 in 12-14 cm and 35 in 14-16cm). There were less than 1.5 reads in BPF3 that mapped to *Blastocatellia* but in BPF4 and BPF5 there are large spikes of read abundance between 0 and 3 cm depths. In BPF4, the highest read abundance is at 2-3 cm (28 reads), with the other depths at less than 7 reads. In BPF5, the highest read abundance is also at 2-3cm (21 reads) and then at 1-2 cm (14 reads). The other *Acidobacteriota* classes (*Acidobacteriae, Mor1,* and *Vicinamibacteria* make up less than 16% of the total *Acidobacteriota* recruited reads in BPF1, 3% in BPF2, 1.4% in BPF3, 3.4% in BPF4 and 0.7% in BPF5 (Table 4-S1).

At all five sites, the most dominant *Actinobacteriota* class was *Thermoleophilia.* In BPF1, *Thermoleophilia* has greatest read abundance below 24 cm, highest abundances were at 28-30 cm (75 reads), 38-20 cm (88 reads), and 52-54 cm (50 reads). In BPF2, the *Thermoleophilia* had a range of 13 to 32 reads (found at 6-8 cm and 24-26 cm, respectively) throughout the core. In BPF3, *Thermoleophilia* had more than 38 reads at each sample depth, with the deepest depth having the most reads (45 in 18-20 cm). In BPF4, depth 43-45cm had the greatest number of *Thermoleophilia* reads (30 reads). BPF4 and BPF5 had a similar amount of *Thermoleophilia* reads in the 0 -3 cm of surface sample (BPF4:BPF5 – 0-1 cm 23:28, 1-2 cm 16:26, 2-3 cm 12:26) (Figure 4-6, Supplemental Table 4-2). The other three *Actinobacteriota* classes (*Acidimicrobiia, Actinomycetia,* and *UBA4738)* cumulatively made up less than 38% of total *Actinobacteriota* reads (BPF1: 11%, BPF2: 29%, BPF3: 31%, BPF4: 37%, 17%).

The most abundant *Chloroflexota* class is *Ellin6529* in every site except BPF2, where *Anaerolineae* has nearly twice as many reads at every depth than *Ellin6529.* In BPF1, *Dehalococcoidia* has more total reads than *Anaerolineae* but only has reads mapping from 22 to 38 cm and then 46 to 58 cm. *Ellin6529* MAGs made up a majority of the mapped read in BPF1: 92%, BPF2: 34%, BPF3:97%, BPF4:56%, and BPF5: 69%. *Ellin6529* has the absolute maxima read abundance at BPF3 0-1 cm (69 reads).

Enzyme counts and autotrophy

All 169 MAGs contained some annotation for genes encoding for one of the seven enzymes tested in the Svalbard active layer soil and culture broths in chapter 3 (Figure 4-7, Supplemental Table 4-2). The smallest gene representation in the MAGs was in AG and CB. LAP had the highest gene counts per MAG out of the seven enzymes. BG had more MAGs from BPF1 with gene counts at deeper than 30 cm in the core. MAGs had <3 gene counts for NAG. Phosphatase encoding genes are present in MAGs at all sites in the upper 6cm. There were only seven MAGs that did not contain gene counts for LAP or PHOS, but they were not the same MAGs (Table 4-S2). Xylosidase genes were more present in BPF2 in the surface samples and decreased with depth (Figure 4-7 and Table 4-S2). There were 13 MAGs that had a carbon fixation pathway. The only pathway present which contained ≥60% of the pathway and the key enzyme (31) was the Calvin Benson (CBB) cycle. Ten MAGs were *Actinobacteriota,* two were *Chloroflexota* and one was *Cyanobacteria* (Supplemental Figure 4-3). BPF3 was the only site without any MAGs with the Calvin-Benson cycle present. BPF1 samples had reads recruit to all MAGs except for the *Cyanobacteria* (Supplemental Figure 4-3). The other sites have three or less MAGs with read recruitment present.

Discussion

Microbial abundance across sites

The microbial abundance variations demonstrate that there are organismal fluctuations at different locations based on the site, depth, or season. Mapping the reads from all whole metagenomic libraries to all the MAGs gives the distribution patterns of each MAG's relative abundance. Since single sample assembly was used to assemble the MAGs, the origin source of that DNA and reads are only from that location. Read mapping allows the analysis of MAG distribution, and therefore organismal abundance, in other samples, despite the reads coming from one location. There are two unique patterns visible when the MAGs were organized by phyla (Supplemental Figure 4-2). Some MAGs and taxa that come from the upper 16cm in the BPF1 have a decreasing read abundance down the core. This higher surface read recruitment in

BPF1 is seen for seven *Proteobacteria*, six *Chloroflexota,* ten *Actinobacteriota,* and ten *Acidobacteriota* (Figure 4-4). The group of 11 BPF3, 4 and 5 *Actinobacteriota* also had read recruitment from the upper 28 cm of BPF1 and most depths in BPF2. Seven *Actinobacteriota* that originate from other sample sites have decreasing abundances in BPF2. There are nine out of the 52 *Actinobacteriota* MAGs that have some reads recruiting at all sample depths. A group of 12 *Acidobacteriota* that originate from BPF2 have isolated read recruitment to that site only.

Since BPF1 and BPF2 were sampled while fully frozen in April 2018 and BPF3, BPF4 and BPF5 were sampled while thawed in September 2019, the 20 MAGs that were present in all samples may represent organisms that are present continuously and distributed widely across sites. Ther eight *Actinobacteriota-Thermoleophilia* MAGs, originating from one of the thawed samples that are also present in the upper surface of BPF1 that then decrease abundance with the soil core depth could indicate that these organisms are more abundant during the thawed season (Figure 4-4). The other group of 23 *Actinobacteriota-Thermoleophilia* that display abundance restricted to the lower 22cm of BPF1 and have little abundance elsewhere and the five *Actinobacteriota-Thermoleophilia* that have isolated abundance to the BPF2 site and aren't present in other depths demonstrates that some organisms from the same phylum inhabit different locations, perhaps based on the soil condition. Further, the three distinct abundance patterns from the *Acidobacteriota-Blastocatellia* demonstrate something similar, where abundance is localized in BPF1 to the 4 to 20 cm depths and below 4 cm in BPF2.
Given these variations of microbial presences at the sites, perhaps the organisms fill niches that relate to the soil stratigraphy.

The temperature of the Svalbard active layer fluctuates from -10°C to 10°C at 40cm between April and September (1). The warmest part of the frozen active layer would be the upper surface since the ice layer and ≥ 0.3 m snowpack acts as an insulating layer (1). Thus, the upper layer of BPF1, would the highest temperatures of this site. The three thawed sites would have been exposed to warming conditions for at least three months. This may be why these MAGs are present in the upper half of BPF1 and in the thawed BPF3, BPF4, and BPF5 sites. The *Actinobacteriota* MAGs that are present in both the upper 22 cm of frozen BPF1 and the upper 3 cm of the three thawed samples (BPF3-5) are potentially the population of microbes that prefer the warming soil conditions. Some MAGs from phyla *Acidobacteriota, Chloroflexota, and Proteobacteria* were only present in the deeper part (≥22 cm) of BPF1 and did not recruit reads from the thawed sample sites. These could be organisms that are best suited for the frozen conditions of the dark winter season (Nov-March) and become less dominant when the season shifts and the active layer thaws. Some other *Acidobacterota-Blastoctellia* MAGs were present in all depths in BPF2 except for the upper 4 cm, where vegetation is visible. If this active layer permanently thaws and becomes more available for vegetation growth, *Blastoctellia* organisms may retreat deeper in the soil as they are known to not reside with vegetation cover (9).

Dominating phyla, Acidobacteriota and Actinobacteriota

Active layer permafrost microbial communities can be assessed using thin slice stratification sampling (2 cm intervals). Network analyses and read mapping across multiple active layer locations display a unique trend based on location in the soil strata. Distribution patterns of the *Gemmatimonadota* and *Thermoproteota* MAGs correlated with the *Acidobacteriota* (Spearman coefficient = 0.5 and 0.7, respectively, Figure 4-5B). *Acidobacteriota* were also reported to become more dominant in thawing northern Sweden mineral rich palsa mire (fen) (32) and increased abundance throughout a range of sub-Antarctic and Antarctic soils with increased soil respiration (33). In the Svalbard active layer study here, the *Acidobacteriota* was the most abundant (n=60) and the *Blastocatellia* class had the most MAGs (n=37, 22%). This class has been described to be one of the most prevalent in tundra aeolian sands and is significantly associated with habitats without vegetation (9). The *Blastocatellia* MAGs originating in BPF1 and BPF2 have little to no read recruitment in the upper 4 cm of both BPF1 and upper 4 cm in BPF2 and in the thawed samples from BPF3, BPF4 and BPF5 (Figure 4-4). *Acidobacteriota-Thermoanaerobaculia*, showed an opposite pattern where the highest abundances was present above 12cm at BPF1 and had some read recruitment in the thawed soil samples (Figure 4-4). A *Thermoanaerobaculia* isolate, cultured from a hot spring, was characterized as a strict anaerobic chemoorganotroph that can grow on proteinaceous compounds and contains genes for reductive carboxylation (34). *Acidobacteriota* isolates from pyrite mine in Japan and clover pasture in Australia also contain genomic evidence for withstanding variable condition in soil such as nutrient fluctuations and water content variations (35).

Other studies have found the same classes of *Acidobacteriota* in a variety of locations with evidence for metabolisms related to their environment. MAGs from *Acidobacteriota-Mor1* class contained capabilities to dissimilate inorganic sulfur compounds in Svalbard fjord sediment (36). *Vicinamibacteria* and *Acidobacteriae* MAGs from activated sludge wastewater treatment plants have been reported to be involved in nitrogen and phosphorus removal and iron reduction and utilization of organic compounds like xylose, acetate and fatty acids. (37). Considering the wide variety of genetic features that *Acidobacteriota* geneomes and MAGs contain, their presence and dominance in the Svalbard active layer could mean that this phylum is the best suited for the variability in the environment.

Actinobacteriota had the second most individual MAGs present (n=52) and had the most cumulative reads in each of the BPF3, 4 and 5 cores. The *Actinobacteriota* phyla co-corelated in distribution with three other phyla (*Chloroflexota, Proteobacteria,* and *Methylomirabilota*: Spearman coeff : 0.9, 0.7, and 0.7; respectively, Figure 4-5B). The three former phyla have been found in active layer soils (12, 16, 32, 38) and *Methylomirabilota* has been found in bottom water of a permafrost thaw pond in Northern Canada (39) *Methylomirabiolta (previously called "Candidate phylum NC10")* has been isolated from paddy soil in China and demonstrated anaerobic oxidation of methane couple to nitrite reduction (40). In Svalbard cryosols,

Proteobacteriota MAGs contained genetic features involved in assimilatory and dissimilatory sulfate reduction (16), which is similar to the *Acidobacteriota* found in Svalbard fjord sediment (36). *Actinobacteriota* class *Acidimicrobila* only had one MAG in the active layer studied in this soil but has been found in thawed discontinuous permafrost in Sweden and identified to contribute to iron cycling (32). Cultured isolates of *Actinobacteriota* class *Actinomycetia* have been retrieved from sulfide ore deposits in Russia and were shown to uptake oxygen and emit carbon dioxide (41). The *Thermoleophilia* class is known to be abundant in soil and hot spring samples and play a role in environmental biogeochemical cycling and contained genes for carbon-nitrogen hydrolase (42) The last *Actinobacteriota* class, UBA4738, is a relatively new reclassification and MAGs from Antarctic Mackay Glacier soils contained genetic evidence to couple atmospheric H_2 and CO oxidation to fix carbon (43). The *Actinobacteriota* was the phylum with the most individual MAGs containing the CBB pathway (Supplemental Figure 4-3) and abundance decreased with depth of the BPF1 core. *Actinobacteriota* MAGs have previously been characterized to have other carbon fixation pathways, like Wood-Lunjdahl (44). While expected to have MAGs from both *Acidobacteriota* and Actinobacteriota, it further demonstrates how diverse these phyla are and how a variety of metabolisms are needed to overcome the energy-deplete characteristics of the soils.

Despite the three thawed sites having fewer sample depths than BPF1 and BPF2, there were similar trends in the species diversity and evenness. At the phylum and class level, BPF4 is second to BPF1 in the most diverse and even location (Table 4-2). This means that dispite the smaller sample size in the thawed soil samples, the MAGs present in these sites can serve as a smaller scale of what the frozen soils may look like when thawed. *Actinobacteriota* depth distribution between the five cores. In BPF1 and BPF2, *Acidobacteriota* had large spikes in abundance in the upper 20 cm, whereas *Actinobacteriota* were more evenly distributed with depth and were more abundant than *Acidobacteriota* below 20 cm (Figure 4-6). Further, the correlation of groups demonstrate that *Actinobacteria* and *Acidobacteriota* inhabit different niches (Figure 4-5B). Both phyla are common soil organisms that have been found in a variety of environments in the Arctic, including ancient permafrost (Chapter 2, (45)), thermokarst bog (13), Alaskan cave permafrost (46) and 50,000 year old Canadian permafrost (47). Other studies of Svalbard permafrost in a lower latitude showed a community shift between *Acidobacteriota* and *Proteobacteria* in the active layer to *Actinobacteria, Bacteroidetes, Chloroflexota* and *Proteobacteria* in the permafrost layer 2 m down (16). This study also showed these MAGs contained genes to regulate ammonium, sulfur, and phosphate. The diversity present in these active layer soils can be seen by the different dominating MAGs (*Actinobacteriota* and *Acidobacteriota).*

The enzymatic gene counts show higher amounts of carbon degrading polymers localized to the surface samples (Figure 4-7). As seen in chapter three of this dissertation, the enzymatic ability will determine if the soil organic matter in the active layer and thawed permafrost will be utilized. Even though the conclusion of chapter three marked BPF2 as a

dominating autotrophic microbial metabolism, there are more MAGs with autotrophy from BPF1. Since BPF2 is half the total depth of BPF1, and most of the autotrophic MAGs are below 22 cm at BPF1 perhaps the autotrophic organisms were not retrieved due to the shorter core. BPF2 is known to have lower total organic carbon than BPF1 but there could still be heterotrophic microbes present. The MAG enzymatic gene counts paralleled the results from the bulk soil and culture enzymatic activity analysis in chapter three where LAP gene counts are the highest, followed by PHOS. The high LAP enzyme gene counts per MAG are highest at the top of the sites, but are presentthroughout the entire depth of the sample. This shows that these organisms could utilize amino acids for energy wherever they are present in the soil and further emphasizes why understanding how organisms are present in the soil strata will matter for microbial activity.

Conclusion

Microbes in cryosols are expected to experience a dramatic environmental shift as their frozen habitat thaws. The active layer soil specifically, will experience a longer thawed season as the climate continues to change (1). Because of this, the active layer can serve as a preview of how the whole cryosol environment will change the microbial interactions in the warming world. Cryoturbation during the warm mid-Holocene era was known to inundate active layer soils with soil organic carbon from the permafrost (7). This warming activity is happening in these Svalbard crysoils now (1) and will affect the surface soils the most (51). Cryosol vertical depth stratification should be considered when evaluating microbial communities that are expected to respond to warmer soil.

Two population are present in different strata of the active layer are shown to have abundance related to the depth layer and frozen condition of the soil. The different classes in the *Acidobacteriota* phylum were shown to inhabit the upper 18 cm of the frozen cores and throughout the thawed active layer. The *Acidobacteriota-Blastocatellia* had the highest abundance in the upper 8cm of BPF1 then the highest abundance shifts to *Acidobacteriota-Thermoanaerobaculia* between 10 to 16 cm. The is more centralized to the soil in the upper half of BPF1 and is also present in the three thawed soil sites (BPF3, BPF4, BPF5). A population of *Actinobacteriota-Thermoleophilia* is found at highest abundance in the deeper depths of BPF1. Another population of organisms seem to be restricted to only the BPF2 site and without seasonal replication in the same location, it is difficult to determine if those organisms will persist in warmer soil. The *Acidobacteriota* a phylum that has been previously shown to represent up to 52% of the soil bacterial community and up to 20% of the whole soil community (48–50) and is one of the most abundant and diverse phyla on Earth (50).

The microbial community between the five active layer sites contain a variety of organisms throughout the 2cm intervals. The *Actinobacteriota* and *Acidobacteriota* are in two different microbial correlated networks and show dominance at different layers of the five sites.

Concurrently, the *Actinobacteriota* had higher abundance in the lower depths of the frozen cores and less abundance in the thawed soil. The *Chloroflexota* in the Svalbard active layer displays similar distribution as the *Actinobacteriota* and may dominate the soil when it experiences longer a thawing season. These organisms have been found in other soil and active layer studies and may drive soil organic matter to be continuously degraded even in the frozen season (15, 55). This means that instead of a population that is 'waiting' for the active layer to thaw before becoming active, there may be two different depth- and season-dependent populations that are responding to environmental changes. Therefore, when the active layer experiences a longer thawed season the population that is most adapted to the surface and summer temperatures will become the dominating phylum.

References

1. Boike J, Juszak I, Lange S, Chadburn S, Burke E, Paul Overduin P, Roth K, Ippisch O, Bornemann N, Stern L, Gouttevin I, Hauber E, Westermann S. 2018. A 20-year record (1998-2017) of permafrost, active layer and meteorological conditions at a high Arctic permafrost research site (Bayelva, Spitsbergen). Earth Syst Sci Data 10:355–390.
2. Cohen J, Screen JA, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K, Entekhabi D, Overland J, Jones J. 2014. Recent Arctic amplification and extreme mid-latitude weather. Nat Geosci. Nature Publishing Group.
3. H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama NMW (eds. . 2019. The ocean and cryosphere in a changing climate. A special report of the intergovernmental panel on climate change. Intergov Panel Clim Chang undefined.
4. Jorgenson MT, Racine CH, Walters JC, Osterkamp TE. 2001. Permafrost degradation and ecological changes associated with a warming climate in central Alaska. Clim Change 48:551–579.
5. Gittel A, BÃ¡rta J, KohoutovÃ¡ I, Schnecker J, Wild B, ÄŒapek P, Kaiser C, Torsvik VL, Richter A, Schleper C, Urich T. 2014. Site- and horizon-specific patterns of microbial community structure and enzyme activities in permafrost-affected soils of Greenland. Front Microbiol 5:541.
6. Keating K, Binley A, Bense V, Van Dam RL, Christiansen HH. 2018. Combined Geophysical Measurements Provide Evidence for Unfrozen Water in Permafrost in the Adventdalen Valley in Svalbard. Geophys Res Lett 45:7606–7614.
7. Bockheim JG. 2007. Importance of Cryoturbation in Redistributing Organic Carbon in Permafrost-Affected Soils. Soil Sci Soc Am J 71:1335–1342.
8. Turetsky MR, Abbott BW, Jones MC, Anthony KW, Olefeldt D, Schuur EAG, Grosse G, Kuhry P, Hugelius G, Koven C, Lawrence DM, Gibson C, Sannel ABK, McGuire AD. 2020. Carbon release through abrupt permafrost thaw. Nat Geosci 13:138–143.
9. Ivanova AA, Zhelezova AD, Chernov TI, Dedysh SN. 2020. Linking ecology and

systematics of acidobacteria: Distinct habitat preferences of the Acidobacteriia and Blastocatellia in tundra soils. PLoS One 15:1–19.

10. Schostag M, Stibal M, Jacobsen CS, Bælum J, Tas N, Elberling B, Jansson JK, Semenchuk P, Priemé A. 2015. Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. Front Microbiol 6:399.
11. Wilhelm RC, Niederberger TD, Greer C, Whyte LG. 2011. Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian high arctic. Can J Microbiol 57:303–315.
12. Mackelprang R, Waldrop MP, DeAngelis KM, David MM, Chavarria KL, Blazewicz SJ, Rubin EM, Jansson JK. 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. Nature 480:368–371.
13. Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, Hoelzle RD, Lamberton TO, McCalley CK, Hodgkins SB, Wilson RM, Purvine SO, Nicora CD, Li C, Frolking S, Chanton JP, Crill PM, Saleska SR, Rich VI, Tyson GW. 2018. Genome-centric view of carbon processing in thawing permafrost. Nature 560:49–54.
14. Morgalev YN, Lushchaeva I V., Morgaleva TG, Kolesnichenko LG, Loiko S V., Krickov I V., Lim A, Raudina T V., Volkova II, Shirokova LS, Morgalev SY, Vorobyev SN, Kirpotin SN, Pokrovsky OS. 2017. Bacteria primarily metabolize at the active layer/permafrost border in the peat core from a permafrost region in western Siberia. Polar Biol 40:1645–1659.
15. Hultman J, Waldrop MP, Mackelprang R, David MM, McFarland J, Blazewicz SJ, Harden J, Turetsky MR, McGuire AD, Shah MB, VerBerkmoes NC, Lee LH, Mavrommatis K, Jansson JK. 2015. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. Nature 521:208–212.
16. Xue Y, Jonassen I, Øvreås L, Taş N. 2020. Metagenome-assembled genome distribution and key functionality highlight importance of aerobic metabolism in Svalbard permafrost. FEMS Microbiol Ecol 96.
17. Taş N, Prestat E, Wang S, Wu Y, Ulrich C, Kneafsey T, Tringe SG, Torn MS, Hubbard SS, Jansson JK. Landscape topography structures the soil microbiome in arctic polygonal

tundra.

18. Schostag M, Stibal M, Jacobsen CS, Bælum J, Tas N, Elberling B, Jansson JK, Semenchuk P, Priemé A. 2015. Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. Front Microbiol 6:1–13.
19. Tveit A, Schwacke R, Svenning MM, Urich T. 2013. Organic carbon transformations in high-Arctic peat soils: Key functions and microorganisms. ISME J 7:299–311.
20. Kaiser C, Meyer H, Biasi C, Rusalimova O, Barsukov P, Richter A. 2007. Conservation of soil organic matter through cryoturbation in arctic soils in Siberia. J Geophys Res 112:G02017.
21. Fierer N, Nemergut D, Knight R, Craine JM. 2010. Changes through time: Integrating microorganisms into the study of succession. Res Microbiol 161:635–642.
22. Uritskiy G V., DiRuggiero J, Taylor J. 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 6:158.
23. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834.
24. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607.
25. Kang DD, 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.
26. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of

bacteria and archaea. Nat Biotechnol. Nature Publishing Group.

27. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–55.
28. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996.
29. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
30. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 35:725–731.
31. Probst AJ, Castelle CJ, Singh A, Brown CT, Anantharaman K, Sharon I, Hug LA, Burstein D, Emerson JB, Thomas BC, Banfield JF. 2017. Genomic resolution of a cold subsurface aquifer community provides metabolic insights for novel microbes adapted to high CO2 concentrations. Environ Microbiol 19:459–474.
32. Mondav R, McCalley CK, Hodgkins SB, Frolking S, Saleska SR, Rich VI, Chanton JP, Crill PM. 2017. Microbial network, phylogenetic diversity and community membership in the active layer across a permafrost thaw gradient. Environ Microbiol 19:3201–3218.
33. Yergeau E, Bokhorst S, Kang S, Zhou J, Greer CW, Aerts R, Kowalchuk GA. 2012. Shifts in soil microorganisms in response to warming are consistent across a range of Antarctic environments. ISME J 6:692–702.
34. Stamps BW, Losey NA, Lawson PA, Stevenson BS. 2014. Genome sequence of

Thermoanaerobaculum aquaticum MP-01T, the first cultivated member of Acidobacteria subdivision 23, isolated from a hot spring. Genome Announc 2:1–3.

35. Ward NL, Challacombe JF, Janssen PH, Henrissat B, Coutinho PM, Wu M, Xie G, Haft DH, Sait M, Badger J, Barabote RD, Bradley B, Brettin TS, Brinkac LM, Bruce D, Creasy T, Daugherty SC, Davidsen TM, DeBoy RT, Detter JC, Dodson RJ, Durkin AS, Ganapathy A, Gwinn-Giglio M, Han CS, Khouri H, Kiss H, Kothari SP, Madupu R, Nelson KE, Nelson WC, Paulsen I, Penn K, Ren Q, Rosovitz MJ, Selengut JD, Shrivastava S, Sullivan SA, Tapia R, Thompson S, Watkins KL, Yang Q, Yu C, Zafar N, Zhou L, Kuske CR. 2009. Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. Appl Environ Microbiol 75:2046–2056.
36. Flieder M, Buongiorno J, Herbold CW, Hausmann B, Rattei T, Lloyd KG, Loy A, Wasmund K. 2021. Novel taxa of Acidobacteriota implicated in seafloor sulfur cycling. ISME J 1–22.
37. Kristensen JM, Singleton C, Clegg LA, Petriglieri F, Nielsen PH. 2021. High Diversity and Functional Potential of Undescribed “Acidobacteriota” in Danish Wastewater Treatment Plants. Front Microbiol 12.
38. Jansson JK, Taş N. 2014. The microbial ecology of permafrost. Nat Rev Microbiol.
39. Crevecoeur S, Vincent WF, Comte J, Matveev A, Lovejoy C. 2017. Diversity and potential activity of methanotrophs in high methane-emitting permafrost thaw ponds. PLoS One 12:e0188223.
40. He Z, Cai C, Wang J, Xu X, Zheng P, Jetten MSM, Hu B. 2016. A novel denitrifying methanotroph of the NC10 phylum and its microcolony. Sci Rep 6:1–10.
41. Egorova D, Pyankova A, Shestakova E, Demakov V, Levin L, Maltsev S, Isaevich A, Grishin E, Kormshchikov D. 2021. Risk assessment of change in respiratory gas concentrations by native culturable bacteria in the air of sulfide ore mines. Environ Geochem Health 1–15.
42. Li C, He YQ, Cui LQ, Albuquerque L, Chen RW, Long LJ, Tian XP. 2021. Miltoncostaea marina gen. nov. sp. nov., and Miltoncostaea oceani sp. nov., a novel deep branching phylogenetic lineage within the class Thermoleophilia isolated from marine

environments, and proposal of Miltoncostaeaceae fam. nov. and Miltoncostaeales ord. nov. Syst Appl Microbiol 44.

43. Ortiz M, Leung PM, Shelley G, Goethem MW Van, Bay SK, Jordaan K, Vikram S, Hogg ID, Makhalanyane TP, Chown SL, Grinter R, Cowan DA, Greening C. 2020. A genome compendium reveals diverse metabolic adaptations of Antarctic soil microorganisms. bioRxiv 2020.08.06.239558.
44. Jiao JY, Fu L, Hua ZS, Liu L, Salam N, Liu PF, Lv AP, Wu G, Xian WD, Zhu Q, Zhou EM, Fang BZ, Oren A, Hedlund BP, Jiang HC, Knight R, Cheng L, Li WJ. 2021. Insight into the function and evolution of the Wood–Ljungdahl pathway in Actinobacteria. ISME J 15:3005–3018.
45. Sipes K, Almatari A, Eddie A, Williams D, Spirina E, Rivkina E, Liang R, Onstott TC, Vishnivetskaya T, Lloyd KG. 2021. Eight metagenome-assembled genomes provide evidence for microbial adaptation in 20,000 to 1,000,000-year-old Siberian permafrost. Appl Environ Microbiol.
46. Burkert A, Douglas TA, Waldrop MP, Mackelprang R. 2019. Changes in the active, dead, and dormant microbial community structure across a pleistocene permafrost chronosequence. Appl Environ Microbiol 85:AEM.02646-18.
47. Goordial J, Davila A, Greer CW, Cannam R, DiRuggiero J, McKay CP, Whyte LG. 2017. Comparative activity and functional ecology of permafrost soils and lithic niches in a hyper-arid polar desert. Environ Microbiol 19:443–458.
48. Dunbar J, Barns SM, Ticknor LO, Kuske CR. 2002. Empirical and theoretical bacterial diversity in four Arizona soils. Appl Environ Microbiol 68:3035–3045.
49. Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M. 2002. Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. Appl Environ Microbiol 68:2391–2396.
50. Kielak AM, Barreto CC, Kowalchuk GA, van Veen JA, Kuramae EE. 2016. The ecology of Acidobacteria: Moving beyond genes and genomes. Front Microbiol 7:1–16.
51. Jentzsch K, Schulz A, Pirk N, Foken T, Crewell S, Boike J. 2021. High levels of CO2 exchange during synoptic-scale events introduce large. Geophys Res Lett 1–9.

52. Gittel A, BÃ¡rta J, KohoutovÃ¡ I, Schnecker J, Wild B, ÄŒapek P, Kaiser C, Torsvik VL, Richter A, Schleper C, Urich T. 2014. Site- and horizon-specific patterns of microbial community structure and enzyme activities in permafrost-affected soils of Greenland. Front Microbiol 5:541.
53. Bockheim JG, Tarnocai C. 1998. Recognition of cryoturbation for classifying permafrost-affected soils. Geoderma 81:281–293.
54. Kreslavsky MA, Head JW, Marchant DR. 2008. Periods of active permafrost layer formation during the geological history of Mars: Implications for circum-polar and mid-latitude surface processes. Planet Space Sci 56:289–302.
55. Niederberger TD, Perreault NN, Tille S, Lollar BS, Lacrampe-Couloume G, Andersen D, Greer CW, Pollard W, Whyte LG. 2010. Microbial characterization of a subzero, hypersaline methane seep in the Canadian High Arctic. ISME J 4:1326–1339.

Appendix

Figures and Tables

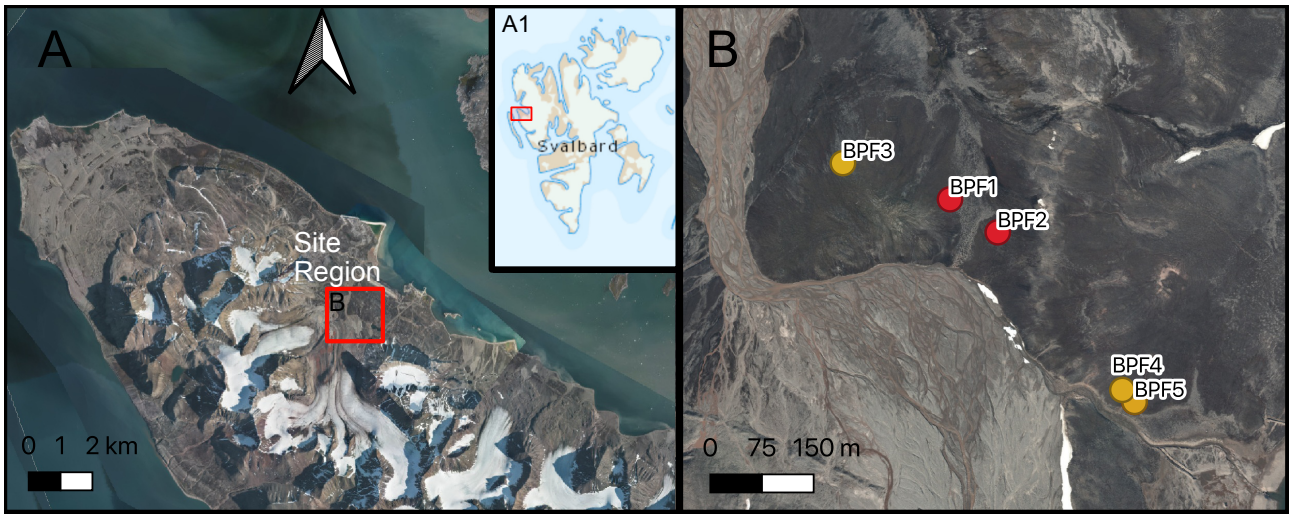


Figure 4-1: Site location near the research station in Ny Ålesund (A) on the Svalbard archipelago (A1). Site region is expanded in panel B. Sites colored in red were collected in April 2018 (Chapter 3 in this document), sites in gold were collected in September 2019.

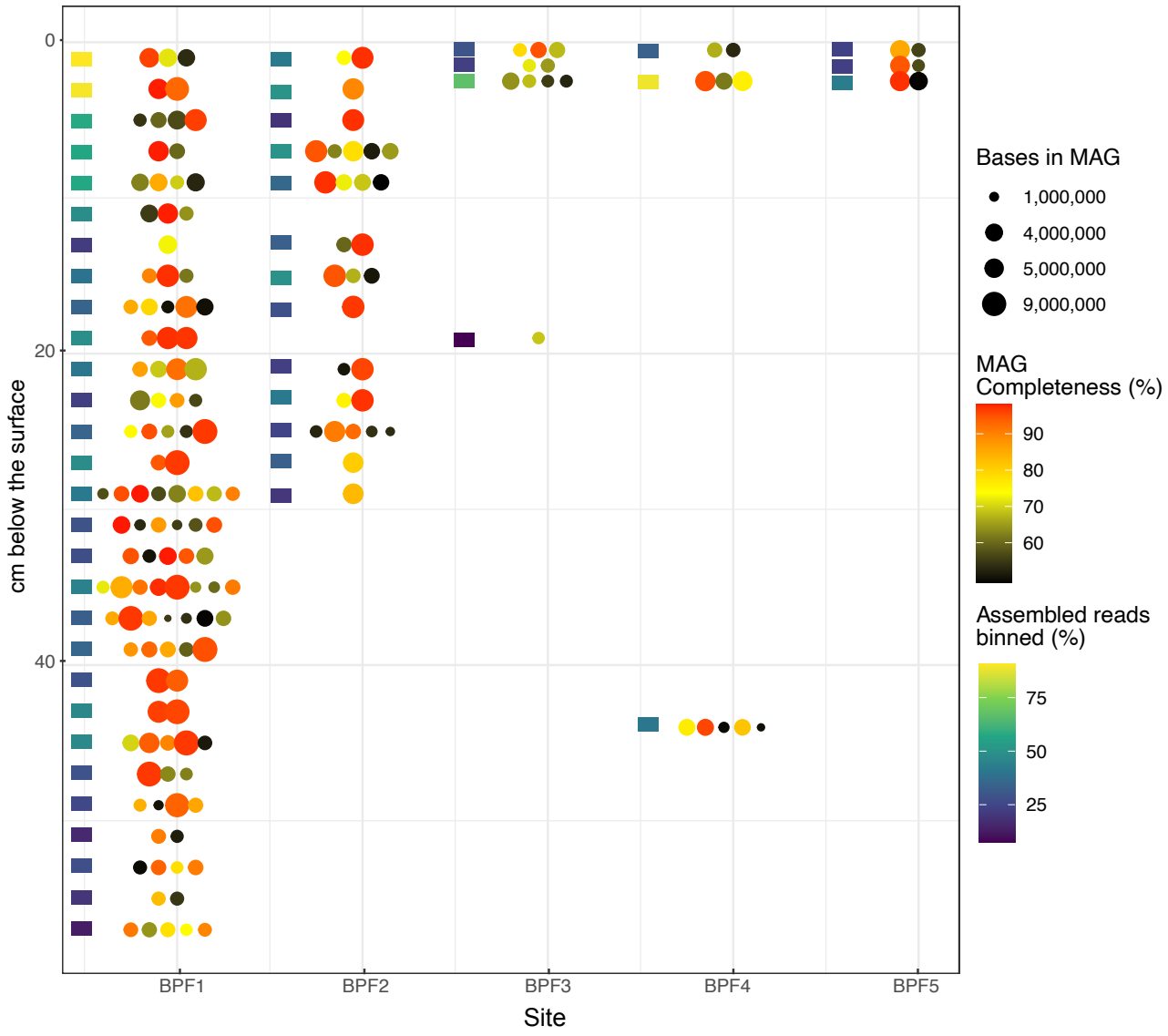


Figure 4-2: MAG size, completeness and assembled reads binned per sample depth.

MAG size (circle size), completeness (black, yellow, red circle) and individuals from each single sample assembly. Percentage of the assembled reads that went into bins is depicted by the box next to the circles. The four metagenomes that did not yield medium quality MAGs were excluded from this figure.

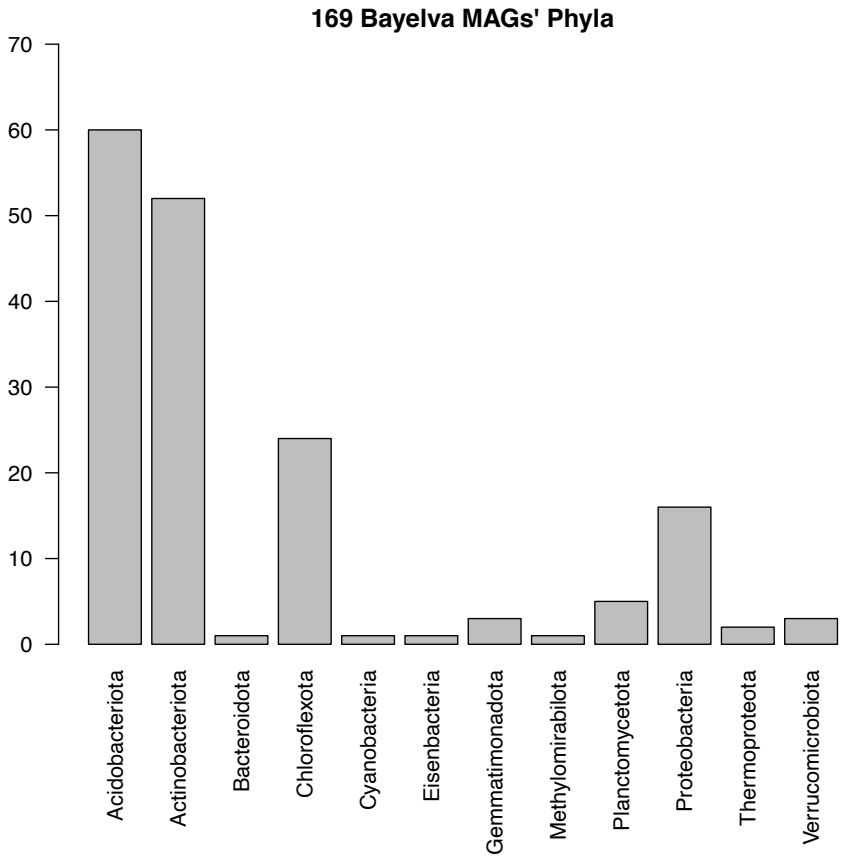


Figure 4-3: Phyla distribution of the 169 MIMAGS medium quality MAGs.

Figure 4-4: Heatmap for reads per million reads of MAGs (169) to the assembled samples (56).

Abundance has been natural log transformed for visual representation. MAGs are clustered based on spearman correlation coefficient. Sites are ordered from surface to bottom depth.

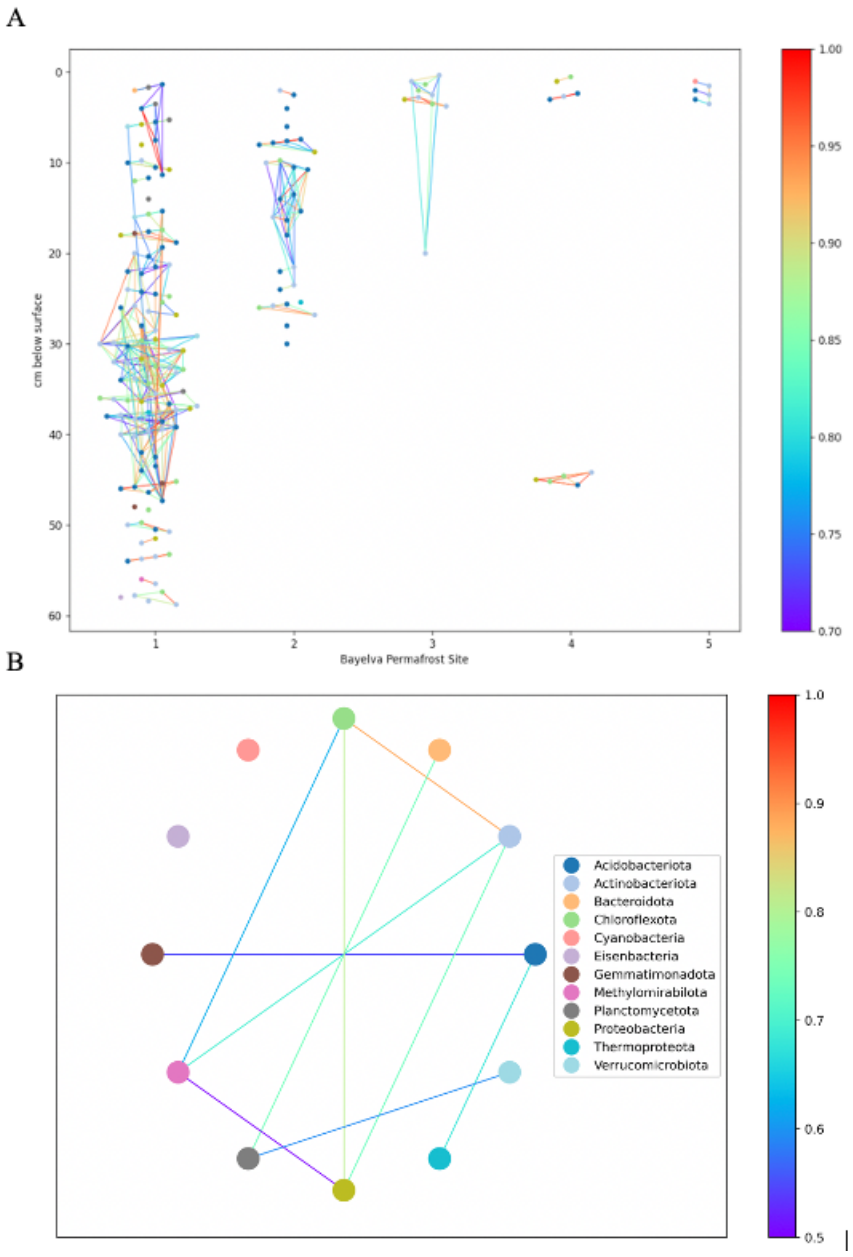


Figure 4-5: Network analysis A) Each MAG correlation based on the unique read abundance from Figure 4-3. Spearman correlation coefficient is represented with the line between MAGs and corresponds to the colors in the legend. MAGs originating from the same location are jittered along the 2cm depth interval to visualize correlations between MAGs from the same origin sample source B) Phyla level correlation of MAG relationships based on read abundances across all samples. Point colors use the same phyla key in panel B.

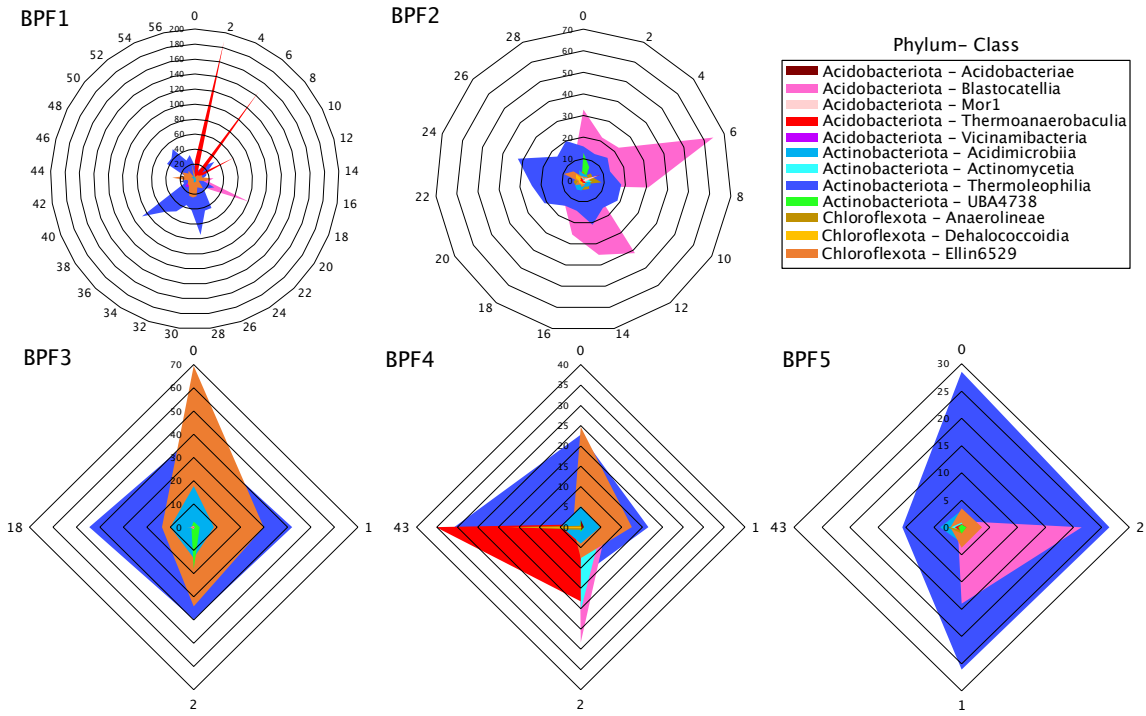


Figure 4-6: Top three phyla with the most MAGs summed read abundance (y axis) by site's top depth (perimeter).

All recruited reads for *Acidobacteriota* (shades of red), *Actinobacteriota* (shades of blue) or *Chloroflexota* (shades of orange) displayed at the class level are summed for every sample depth increment at each site. The center of each plot marks zero mapped reads. Individual site and read recruitments for all MAGs can be found on Table 3-S1.1

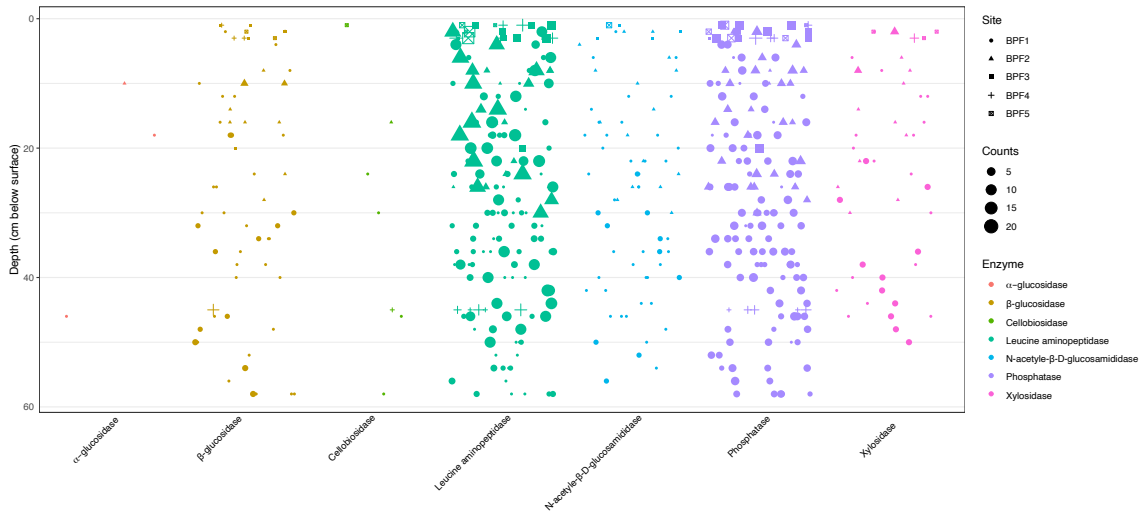


Figure 4-7: Gene counts for each enzyme per MAG by depth.

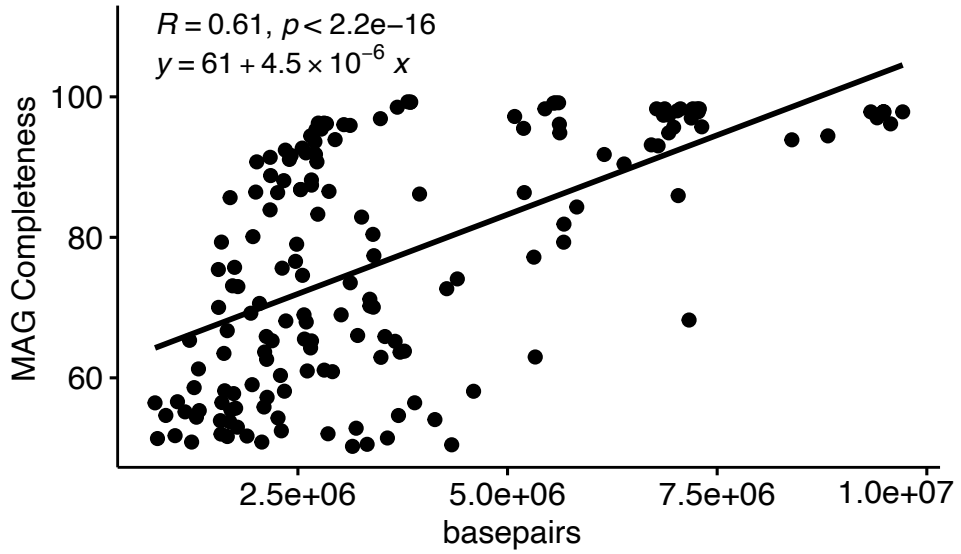
Number of genes that encode for an enzyme is represented by size of the symbol, color represents the enzyme type and shape represents the sample site.

Table 4-1: Linear regression between the percent of assembled reads binned per sample and the number of MAGs, MAG size or MAG completeness.

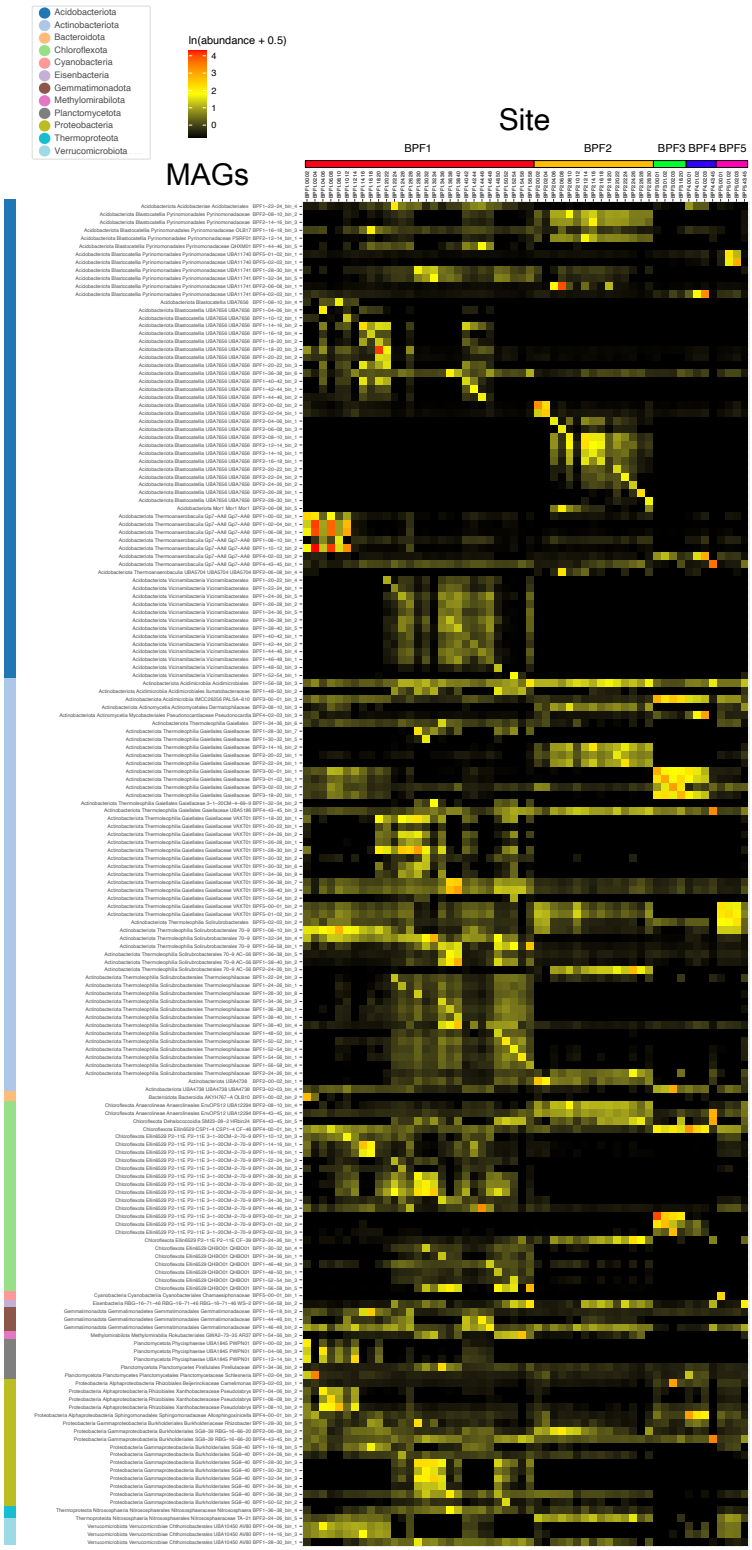
	Percentage of assembled reads binned per sample
Number of MAGs per sample	$R^2 = -0.031$ $p = 0.69$
MAG completeness	$R^2 = 0.012$ $p = 0.87$
MAG size (bp)	$R^2 = 0.2$ $p = 0.0082$

Table 4-2: Diversity indices of each Bayelva permafrost site based on the present MAGs at the phyla and class level. All MAGs annotate up to the class level.

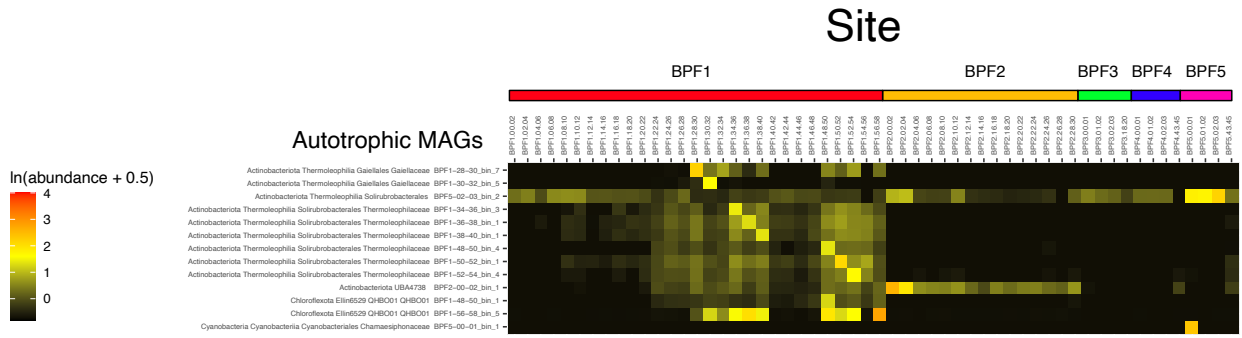
	Phyla		Class	
Site	Shannon Index	Simpson Diversity Index	Shannon Index	Simpson Diversity Index
BPF1	1.73869665	0.77296578	2.24462601	0.85237679
BPF2	1.32135738	0.53777778	1.52747017	0.64444444
BPF3	0.89794572	0.54	1.41848366	0.73
BPF4	1.36615885	0.74	2.16395566	0.91
BPF5	1.01140426	0.61111111	1.01140426	0.63888889



Supplemental Figure 4-1: MAG completeness and total base pairs with regression line and R2 value.



Supplemental Figure 4-2: Reorganization of Figure 4-3, phyla are grouped alphabetically



Supplemental Figure 4-3: Read recruitment for the MAGs with autotrophic metabolisms. The Calvin Benson Cycle was the only carbon fixation pathway present with ≥60% of the pathway complete and the key enzyme, Rubisco present.

Supplemental Table 4-1: Information on all 169 MAGs' completeness and contamination percentages, phylum, number of reads and total number of base pairs. The MAG column displays the site and the depth increment the MAG was assembled from.

MAG	Completeness (%)	Contamination (%)	Phyla	Reads	Base pairs
BPF1-00-02_bin_1	97.2	4.27	Acidobacteriota	83724	5085009
BPF1-00-02_bin_2	72.69	4.54	Bacteroidota	62024	4275032
BPF1-00-02_bin_3	54.64	1.14	Planctomycetota	53907	3699313
BPF1-02-04_bin_1	99.15	7.12	Acidobacteriota	91368	5569227
BPF1-02-04_bin_2	93.88	3.37	Planctomycetota	137848	8390584
BPF1-04-06_bin_1	55.68	2.41	Verrucomicrobiota	29403	1757510
BPF1-04-06_bin_2	60.88	5.66	Proteobacteria	48665	2912269
BPF1-04-06_bin_3	58.09	5.88	Planctomycetota	76502	4594825
BPF1-04-06_bin_4	97.39	3.89	Acidobacteriota	113938	6924428
BPF1-06-08_bin_1	99.15	5.41	Acidobacteriota	91971	5608833
BPF1-06-08_bin_2	61.1	1.27	Proteobacteria	47017	2813973
BPF1-08-10_bin_1	63.65	3.92	Acidobacteriota	61596	3718528
BPF1-08-10_bin_2	86.16	3.9	Proteobacteria	65448	3950780
BPF1-08-10_bin_3	70.59	7.47	Actinobacteriota	33679	2041770
BPF1-08-10_bin_4	54.06	4.6	Acidobacteriota	60402	4133684
BPF1-10-12_bin_1	56.45	3.53	Acidobacteriota	65075	3892517
BPF1-10-12_bin_2	99.1	4.56	Acidobacteriota	91023	5550592
BPF1-10-12_bin_3	65.28	7.27	Chloroflexota	31617	2192232
BPF1-12-14_bin_1	74.08	5.75	Planctomycetota	73213	4401270
BPF1-14-16_bin_1	91.09	0.99	Chloroflexota	39442	2398630
BPF1-14-16_bin_2	98.29	5.56	Acidobacteriota	119660	7287078
BPF1-14-16_bin_3	62.63	8.78	Verrucomicrobiota	35574	2127242
BPF1-16-18_bin_1	86.36	0	Chloroflexota	37177	2258351
BPF1-16-18_bin_2	80.42	9.34	Gemmatimonadota	56042	3395095
BPF1-16-18_bin_3	51.65	0	Acidobacteriota	27731	1658211
BPF1-16-18_bin_4	93.02	5.77	Acidobacteriota	111747	6795176
BPF1-16-18_bin_5	51.44	5.77	Proteobacteria	52163	3567515
BPF1-18-20_bin_1	95.32	0.86	Actinobacteriota	44569	2715467
BPF1-18-20_bin_2	98.29	3.42	Acidobacteriota	115954	7061658
BPF1-18-20_bin_3	98.01	3.42	Acidobacteriota	115222	7024570
BPF1-20-22_bin_1	87.43	1.72	Actinobacteriota	43764	2662467
BPF1-20-22_bin_2	70.23	4.27	Acidobacteriota	55648	3358736
BPF1-20-22_bin_3	93.16	2.99	Acidobacteriota	110235	6714501

Supplemental Table 4-2 continued

MAG	Completeness (%)	Contamination (%)	Phyla	Reads	Base pairs
BPF1-20-22_bin_4	68.23	3.42	Acidobacteriota	117715	7164573
BPF1-22-24_bin_1	62.97	2.56	Acidobacteriota	87587	5330317
BPF1-22-24_bin_2	74.59	1.98	Chloroflexota	42048	2554810
BPF1-22-24_bin_3	88.07	2.13	Actinobacteriota	33321	2331863
BPF1-22-24_bin_4	57.76	1.07	Acidobacteriota	24938	1733386
BPF1-24-26_bin_1	75.73	0.86	Actinobacteriota	28778	1744943
BPF1-24-26_bin_2	96.26	1.72	Actinobacteriota	39697	2812353
BPF1-24-26_bin_3	66.72	4.05	Chloroflexota	27423	1657415
BPF1-24-26_bin_4	55.42	7.12	Proteobacteria	28475	1702222
BPF1-24-26_bin_5	97.86	8.6	Acidobacteriota	155665	9485262
BPF1-26-28_bin_1	95.4	0.86	Actinobacteriota	45588	2778077
BPF1-26-28_bin_2	97.77	8.6	Acidobacteriota	155612	9479659
BPF1-28-30_bin_1	58.61	3.21	Verrucomicrobiota	21135	1262838
BPF1-28-30_bin_2	96.26	1.29	Actinobacteriota	45001	2743391
BPF1-28-30_bin_3	99.26	2.84	Proteobacteria	63155	3845301
BPF1-28-30_bin_4	58.1	0.95	Acidobacteriota	39020	2339490
BPF1-28-30_bin_5	63.8	1.09	Proteobacteria	62891	3770135
BPF1-28-30_bin_6	83.29	0	Chloroflexota	44936	2737534
BPF1-28-30_bin_7	68.97	1.72	Actinobacteriota	42267	2572961
BPF1-28-30_bin_8	91.39	2.76	Actinobacteriota	35799	2169401
BPF1-30-32_bin_1	99.29	2.37	Proteobacteria	62606	3810752
BPF1-30-32_bin_2	54.4	7.76	Actinobacteriota	21369	1289837
BPF1-30-32_bin_3	88.17	0	Chloroflexota	43685	2661212
BPF1-30-32_bin_4	56.6	3.44	Chloroflexota	17774	1063606
BPF1-30-32_bin_5	59.03	8.5	Actinobacteriota	32314	1954520
BPF1-30-32_bin_6	96.18	1.55	Actinobacteriota	40099	2840431
BPF1-32-34_bin_1	95.93	1.49	Chloroflexota	44065	3122888
BPF1-32-34_bin_2	51.72	3.45	Actinobacteriota	31167	1893215
BPF1-32-34_bin_3	99.22	2.37	Proteobacteria	63081	3840918
BPF1-32-34_bin_4	95.52	2.44	Actinobacteriota	38874	2749473
BPF1-32-34_bin_5	65.88	8.46	Acidobacteriota	51299	3538091
BPF1-34-36_bin_1	73.1	1.85	Chloroflexota	28315	1720263
BPF1-34-36_bin_2	85.94	4.79	Planctomycetota	116171	7035207
BPF1-34-36_bin_3	92.7	0.86	Actinobacteriota	42060	2551783
BPF1-34-36_bin_4	98.52	2.84	Proteobacteria	60557	3685288

Supplemental Table 4-3 continued

MAG	Completeness (%)	Contamination (%)	Phyla	Reads	Base pairs
BPF1-34-36_bin_5	97.86	8.6	Acidobacteriota	155600	9482575
BPF1-34-36_bin_6	65.34	1.72	Actinobacteriota	19887	1207756
BPF1-34-36_bin_7	61.27	2.78	Chloroflexota	21738	1314141
BPF1-34-36_bin_8	92.02	0.86	Actinobacteriota	42613	2595568
BPF1-36-38_bin_1	86.44	1.81	Actinobacteriota	32919	1997020
BPF1-36-38_bin_2	97.86	8.6	Acidobacteriota	153134	9332488
BPF1-36-38_bin_3	86.79	5.87	Proteobacteria	41914	2532115
BPF1-36-38_bin_4	56.44	0	Thermoproteota	13233	796143
BPF1-36-38_bin_5	55.16	0.6	Actinobacteriota	19015	1154487
BPF1-36-38_bin_6	50.54	5.18	Acidobacteriota	55505	3325559
BPF1-36-38_bin_7	65.26	2.59	Actinobacteriota	44110	2662034
BPF1-38-40_bin_1	88.79	0.86	Actinobacteriota	35865	2174848
BPF1-38-40_bin_2	93.91	1.75	Actinobacteriota	48404	2943805
BPF1-38-40_bin_3	86.54	1.72	Actinobacteriota	47352	2870636
BPF1-38-40_bin_4	60.34	0	Actinobacteriota	37757	2291851
BPF1-38-40_bin_5	96.15	8.6	Acidobacteriota	156980	9568184
BPF1-40-42_bin_1	97.86	9.03	Acidobacteriota	155606	9480392
BPF1-40-42_bin_2	94.87	3.85	Acidobacteriota	113623	6922424
BPF1-42-44_bin_1	97.38	3.85	Acidobacteriota	112796	6859524
BPF1-42-44_bin_2	97.01	8.72	Acidobacteriota	154441	9407608
BPF1-44-46_bin_1	71.19	3.4	Gemmatimonadota	55594	3357764
BPF1-44-46_bin_2	94.87	6.81	Acidobacteriota	92870	5622980
BPF1-44-46_bin_3	90.74	0	Chloroflexota	44741	2727255
BPF1-44-46_bin_4	97.86	8.93	Acidobacteriota	137048	9712036
BPF1-44-46_bin_5	52.46	4.32	Acidobacteriota	38466	2302678
BPF1-46-48_bin_1	97.86	8.6	Acidobacteriota	155714	9488971
BPF1-46-48_bin_2	64.26	2.59	Gemmatimonadota	44123	2651292
BPF1-46-48_bin_3	63.48	8.9	Chloroflexota	23400	1618096
BPF1-48-50_bin_1	85.65	0.31	Chloroflexota	27828	1691584
BPF1-48-50_bin_2	51.77	0.85	Actinobacteriota	17441	1035366
BPF1-48-50_bin_3	94.44	8.6	Acidobacteriota	144995	8820072
BPF1-48-50_bin_4	86.81	2.01	Actinobacteriota	36030	2532218
BPF1-50-52_bin_1	91.65	1.38	Actinobacteriota	39886	2419480
BPF1-50-52_bin_2	52.98	3.95	Proteobacteria	29764	1778409
BPF1-52-54_bin_1	50.86	0	Acidobacteriota	34218	2069668

Supplemental Table 4-4 continued

MAG	Completeness (%)	Contamination (%)	Phyla	Reads	Base pairs
BPF1-52-54_bin_2	94.44	3.82	Actinobacteriota	43655	2652387
BPF1-52-54_bin_3	79.32	2.44	Chloroflexota	26260	1588263
BPF1-52-54_bin_4	91.81	0.86	Actinobacteriota	38534	2709567
BPF2-26-28_bin_1	81.88	8.08	Acidobacteriota	94178	5672543
BPF2-28-30_bin_1	84.32	5.45	Acidobacteriota	96334	5825415
BPF3-00-01_bin_1	80.1	0.57	Actinobacteriota	32416	1963645
BPF3-00-01_bin_2	96.04	0	Chloroflexota	49981	3048075
BPF3-00-01_bin_3	68.97	0	Actinobacteriota	49749	3015741
BPF3-01-02_bin_1	72.99	0	Actinobacteriota	29378	1783989
BPF3-01-02_bin_2	65.89	0	Chloroflexota	34890	2122836
BPF3-02-03_bin_1	65.2	4.99	Proteobacteria	61151	3659402
BPF3-02-03_bin_2	69.2	6.58	Actinobacteriota	32027	1938534
BPF3-02-03_bin_3	56.49	0	Chloroflexota	26245	1593500
BPF3-02-03_bin_4	53.73	0.85	Actinobacteriota	28264	1688417
BPF3-18-20_bin_1	70.03	0.93	Actinobacteriota	25600	1554212
BPF4-00-01_bin_1	67.95	7.41	Chloroflexota	37376	2599343
BPF4-00-01_bin_2	54.29	7.07	Proteobacteria	32928	2262842
BPF4-02-03_bin_1	96.1	5.3	Acidobacteriota	92493	5618098
BPF4-02-03_bin_2	62.92	1.52	Acidobacteriota	58308	3490178
BPF4-02-03_bin_3	77.18	3.67	Actinobacteriota	88237	5314118
BPF4-43-45_bin_1	77.39	0.85	Acidobacteriota	56565	3406324
BPF4-43-45_bin_2	96.89	2.92	Proteobacteria	57351	3483897
BPF4-43-45_bin_3	50.86	0	Actinobacteriota	20314	1232642
BPF4-43-45_bin_4	82.88	2	Chloroflexota	54220	3259648
BPF4-43-45_bin_5	51.36	1.6	Chloroflexota	13815	825344
BPF5-00-01_bin_1	86.37	4.52	Cyanobacteria	75200	5199531
BPF5-00-01_bin_2	57.23	5.8	Actinobacteriota	30749	2132456
BPF5-01-02_bin_1	95.51	2.61	Acidobacteriota	85618	5191136
BPF5-01-02_bin_2	58.16	6.77	Actinobacteriota	23313	1626993
BPF5-02-03_bin_1	98.29	2.56	Acidobacteriota	89424	5445790
BPF5-02-03_bin_2	50.47	9.48	Actinobacteriota	63038	4334012

Supplemental Table 4-5: Summed recruited reads from all depths and sites for all phyla.

Site	Top Depth (cm)	Acidobacteriota	Actinobacteriota	Bacteroidota	Chloroflexota	Cyanobacteria	Eisenbacteria	Gemmatimonadota	Methylomirabilota	Planctomycetota	Proteobacteria	Thermoproteota	Verrucomicrobiota
BPF1	0	29.2585995	20.24519	13.690415	6.349621	0	0.060519	3.646957	0	18.153889	8.971393	1.002635	2.821798
BPF1	2	191.058643	14.1471365	1.145198	3.645899	0	0.05772	2.336896	0	30.326693	7.2496925	0.4749975	3.2856805
BPF1	4	27.397351	16.950588	0	5.4370615	0	0.077735	2.980226	0.069133	12.1908195	19.841008	0.545755	8.68934
BPF1	6	148.119565	27.067138	0.100481	6.533724	0	0.035265	0.738597	0	1.556396	20.358165	0.792676	4.216106
BPF1	8	21.671845	38.747036	0	7.7693555	0	0.024538	1.85816	0	3.880307	19.9331545	1.186389	3.279423
BPF1	10	71.930095	22.361783	0.118528	12.030773	0	0.083136	4.01992	0.100207	1.668747	6.669954	1.658229	4.316796
BPF1	12	5.573742	18.310523	0	10.131348	0	0.038843	4.5642345	0.143138	13.280584	12.579609	1.201896	4.682798
BPF1	14	27.7210545	14.327024	0	21.5471865	0	0.093796	5.331883	0.172553	0.502772	5.2986295	1.5121445	8.253328
BPF1	16	20.5765855	13.03503	0	15.69069	0	0.108885	13.41649	0.171822	0.755567	10.378858	1.876564	4.17226
BPF1	18	82.3959155	31.434754	0	6.3046425	0	0.086312	2.579485	0.142778	2.785769	7.6516515	1.461124	4.680562
BPF1	20	30.1703465	24.0460515	0	8.826724	0	0.095014	3.311387	0.163093	0.377561	4.9557395	1.28742	4.382264
BPF1	22	15.894291	22.6272915	0	18.038386	0	1.844152	3.535357	0.196413	0.342081	5.6607056	0.682939	1.22817
BPF1	24	16.6852245	48.948515	0	14.5672735	0	0.47707	3.19417	0.148382	0.465223	8.4159465	1.024184	0.923671
BPF1	26	19.899001	48.640995	0	12.6745475	0	0.401116	6.214885	0.137688	0.7901125	6.2369555	1.186008	2.1043715
BPF1	28	10.485609	80.377744	0	24.613467	0	0.535477	0.71363	1.141075	1.073365	34.7349175	0.941423	4.47184
BPF1	30	11.9389935	46.878539	0	25.986482	0	1.309839	0.68031	1.934228	2.039456	33.082081	1.062258	1.434621

Supplemental Table 4-6 continued

Site	Top Depth (cm)	Acidobacteriota	Actinobacteriota	Bacteroidota	Chloroflexota	Cyanobacteria	Eisenbacteria	Gemmatimonadota	Methylomirabilota	Planctomycetota	Proteobacteria	Thermoproteota	Verrucomicrobiota
BPF1	32	11.694542	41.3917055	0	25.4716995	0	0.76174	0.61952	2.797716	1.013368	32.8552665	1.294683	1.541959
BPF1	34	15.3849995	53.409834	0	16.899544	0	0.958328	0.719865	0.444291	9.4224905	20.247193	1.257854	0.995219
BPF1	36	25.567682	61.888494	0	12.356264	0	0.044255	0.610875	0.207011	1.640399	11.5237635	5.788111	0.726206
BPF1	38	22.367095	91.115744	0	13.3286855	0	0.220099	1.332819	0.762385	1.427822	7.284286	2.839142	1.151009
BPF1	40	26.809948	19.2693805	0	10.6038235	0	0.105661	6.539424	0.065306	1.411271	4.6787605	1.692277	2.1432365
BPF1	42	23.1450455	19.618831	0	9.736005	0	0.084023	4.451926	0.078824	1.295801	4.6131666	1.55438	2.073787
BPF1	44	27.424493	11.821926	0	31.8450095	0	0.094027	10.678027	0.084144	1.186615	5.2639785	1.637217	1.297817
BPF1	46	22.0400495	16.301629	0	15.535354	0	0.185657	9.7780695	0.046253	1.0399395	4.568505	0.9035395	1.0194475
BPF1	48	9.6593	50.3629185	0	22.54257	0	0.478171	0.613099	2.477425	0.721848	15.1131505	1.121504	0.960899
BPF1	50	4.428331	47.229242	0	15.193432	0	1.23892	0.672447	2.816983	0.685334	11.2891655	0.579635	1.668517
BPF1	52	8.9525955	55.8381385	0	17.782493	0	3.396803	0.92475	3.168154	0.593993	5.0872675	0.576043	1.32761
BPF1	54	3.346652	31.334771	0	7.0274985	0	2.316847	0.647476	13.219396	0.271747	4.657605	0.135561	1.850464
BPF1	56	13.859534	42.507878	0	21.764861	0	8.019787	1.218371	1.181816	0.914672	8.939946	0.634054	1.110061
BPF2	0	33.170782	32.6932165	0.107491	6.075164	0	0.150123	2.034166	0.089804	0.725472	6.706944	3.182694	0.641092
BPF2	2	22.033516	24.011431	0	4.187432	0	0.110026	0.853839	0.039453	1.066479	6.687869	2.164107	0.11697
BPF2	4	24.96057	21.6019605	0	8.1440845	0	0.11284	1.612404	0.060999	0.379017	6.876642	2.187559	0.60217

Supplemental Table 4-7 continued

Site	Top Depth (cm)	Acidobacteriota	Actinobacteriota	Bacteroidota	Chloroflexota	Cyanobacteria	Eisenbacteria	Gemmatimonadota	Methylomirabilota	Planctomycetota	Proteobacteria	Thermoproteota	Verrucomicrobiota
BPF2	8	33.441344	26.65968	0.062207	11.833803	0	0.290185	1.045398	0.071354	0.525358	6.613404	2.2844975	1.000324
BPF2	10	7.292991	25.870147	0	5.399241	0	0.098918	2.738901	0	0.3868655	4.7896325	1.658972	0.506858
BPF2	12	43.6117715	21.419121	0	7.250192	0	2.394505	2.02879	0.13858	0.60267	3.7178455	2.8934105	1.312648
BPF2	14	36.4196665	26.2535585	0	7.319338	0	1.846684	2.929076	0.027287	0.436072	3.543886	3.25285	0.388778
BPF2	16	26.995638	21.3150095	0	9.684523	0	2.852032	2.213391	0.121029	0.499384	2.9548095	1.877028	1.053116
BPF2	18	15.201607	22.0179665	0	8.991635	0	1.974963	1.884722	0.160627	0.523511	2.4554092	2.5241235	1.348882
BPF2	20	20.031752	28.3010815	0	10.6782125	0	0.534678	2.321971	0.293384	0.407246	2.026117	2.549789	0.489567
BPF2	22	19.375536	29.31845	0	9.321304	0	3.553839	2.503498	0.115281	0.404206	2.218407	1.9751105	0.409839
BPF2	24	13.3693885	38.368785	0	16.542291	0	1.420373	2.020015	0.506716	0.249269	1.7354115	4.7327995	0.271178
BPF2	26	15.7894985	19.776378	0	13.05786	0	1.876172	1.464348	0.987556	0.158534	1.01456	0.8914225	0
BPF2	28	13.158285	28.898231	0	7.088714	0	1.219626	1.405831	0	0.44141	3.572885	1.499509	0.673012
BPF3	0	2.581719	59.735102	0	69.75589	0	0.106425	0.413789	0	0.469364	3.811525	0	0.000716
BPF3	1	3.3364335	53.747764	0	30.076896	0	0.052089	0.415327	0	0.261564	2.601734	0	0
BPF3	2	2.157811	71.1379995	0	34.347143	0	0.053297	0.274432	0	0.339629	21.83609	0	0.072071
BPF3	18	2.247421	55.576564	0	16.914497	0	0.086198	0.542781	0	0.386255	2.9818865	0.327934	0
BPF4	0	4.714847	30.0578925	0.092399	25.0241675	0	0.234381	0.420615	0	3.533617	23.729149	0	0.888364

Supplemental Table 4-8 continued

Site	Top Depth (cm)	Acidobacteriota	Actinobacteriota	Bacteroidota	Chloroflexota	Cyanobacteria	Eisenbacteria	Gemmatimonadota	Methylomirabilota	Planctomycetota	Proteobacteria	Thermoproteota	Verrucomicrobiota
BPF4	2	47.2817115	37.572468	0.09515	8.4418315	0	0.191772	0.503462	0	1.045591	10.1581905	0	1.2649505
BPF4	43	35.568203	35.909237	0	36.6840905	0	2.586438	2.527498	0	0.198002	29.33575	0	0.058919
BPF5	0	1.904506	33.058779	0	3.9108415	9.94552	0.063781	0.837032	0	0.392057	3.251972	1.1995225	1.948371
BPF5	1	14.753151	30.4254335	0	4.869738	0	0.097456	0.951986	0	0.45323	3.236063	2.123443	2.5211535
BPF5	2	22.246243	30.493321	0	4.4571445	0	0.132954	1.317202	0	0.397846	3.574559	1.849261	2.47772
BPF5	43	2.1233915	17.0359235	0	5.299296	0	7.084969	2.478711	0.235052	0.417407	3.451339	0.6499995	0.090215

Supplemental Table 4-9: Number of individual gene counts present for each MAG that had annotated genes for the seven enzymes: α-glucosidase (AG), β-glucosidase (BG), β-D-cellubiosidase (CB), leucine aminopeptidase (LAP), N-Acetyl-β-D-glucosaminidase (NAG), phosphatase (PHOS), and β-xylosidase (XYL).

Site and Sample Depth	Phyla	Class	AG	BG	CB	LAP	NAG	PHOS	XYL
BPF1-22-24_bin_2	Chloroflexota	Ellin6529	0	1	0	3	2	6	0
BPF1-28-30_bin_5	Proteobacteria	Gammaproteobacteria	0	1	1	1	2	3	1
BPF1-28-30_bin_8	Actinobacteriota	Thermoleophilia	0	0	0	1	2	5	0
BPF1-30-32_bin_3	Chloroflexota	Ellin6529	0	2	0	3	2	7	0
BPF1-32-34_bin_1	Chloroflexota	Ellin6529	0	1	0	3	2	7	0
BPF1-34-36_bin_3	Actinobacteriota	Thermoleophilia	0	1	0	1	2	5	0
BPF1-38-40_bin_4	Actinobacteriota	Thermoleophilia	0	0	0	1	2	8	0
BPF1-48-50_bin_4	Actinobacteriota	Thermoleophilia	0	1	0	1	2	5	0
BPF1-50-52_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	1	2	5	0
BPF1-54-56_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	1	2	4	0
BPF5-00-01_bin_1	Cyanobacteria	Cyanobacteriia	0	0	1	2	2	6	0
BPF1-00-02_bin_3	Planctomycetota	Phycisphaerae	0	0	0	0	1	1	0
BPF1-04-06_bin_3	Planctomycetota	Phycisphaerae	0	0	0	2	1	6	0
BPF1-04-06_bin_4	Acidobacteriota	Blastocatellia	0	0	0	14	1	6	1
BPF1-10-12_bin_1	Acidobacteriota	Blastocatellia	0	0	0	6	1	2	1
BPF1-10-12_bin_3	Chloroflexota	Ellin6529	0	1	0	2	1	6	0
BPF1-12-14_bin_1	Planctomycetota	Phycisphaerae	0	0	0	1	1	5	0
BPF1-14-16_bin_1	Chloroflexota	Ellin6529	0	1	0	3	1	6	0
BPF1-14-16_bin_2	Acidobacteriota	Blastocatellia	0	0	0	17	1	6	1
BPF1-16-18_bin_1	Chloroflexota	Ellin6529	0	1	0	3	1	6	0
BPF1-16-18_bin_4	Acidobacteriota	Blastocatellia	0	0	0	18	1	7	1
BPF1-16-18_bin_5	Proteobacteria	Gammaproteobacteria	0	2	0	2	1	1	1
BPF1-18-20_bin_2	Acidobacteriota	Blastocatellia	0	0	0	16	1	6	1
BPF1-18-20_bin_3	Acidobacteriota	Blastocatellia	0	0	0	16	1	4	0
BPF1-20-22_bin_2	Acidobacteriota	Blastocatellia	0	0	0	10	1	3	0
BPF1-20-22_bin_3	Acidobacteriota	Blastocatellia	0	0	0	17	1	5	1
BPF1-20-22_bin_4	Acidobacteriota	Vicinamibacteria	0	0	0	13	1	5	3
BPF1-22-24_bin_1	Acidobacteriota	Vicinamibacteria	0	0	0	9	1	2	1
BPF1-22-24_bin_3	Actinobacteriota	Thermoleophilia	0	0	0	1	1	5	0
BPF1-24-26_bin_3	Chloroflexota	Ellin6529	0	1	0	1	1	6	0
BPF1-24-26_bin_5	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-26-28_bin_2	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3

Supplemental Table 4-10 continued

Site and Sample Depth	Phyla	Class	AG	BG	CB	LAP	NAG	PHOS	XYL
BPF1-32-34_bin_4	Actinobacteriota	Thermoleophilia	0	0	0	1	1	5	0
BPF1-34-36_bin_2	Planctomycetota	Planctomycetes	0	0	0	5	1	6	0
BPF1-34-36_bin_5	Acidobacteriota	Vicinamibacteria	0	0	0	15	1	5	3
BPF1-34-36_bin_7	Chloroflexota	Ellin6529	0	0	0	2	1	3	0
BPF1-36-38_bin_2	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-36-38_bin_6	Acidobacteriota	Blastocatellia	0	0	0	11	1	2	0
BPF1-38-40_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	1	1	5	0
BPF1-38-40_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	1	1	2	0
BPF1-38-40_bin_5	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-40-42_bin_1	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-40-42_bin_2	Acidobacteriota	Blastocatellia	0	0	0	17	1	6	1
BPF1-42-44_bin_1	Acidobacteriota	Blastocatellia	0	0	0	16	1	6	1
BPF1-42-44_bin_2	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-44-46_bin_2	Acidobacteriota	Blastocatellia	0	0	0	10	1	6	1
BPF1-44-46_bin_3	Chloroflexota	Ellin6529	0	2	0	2	1	7	0
BPF1-44-46_bin_4	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-46-48_bin_1	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-48-50_bin_3	Acidobacteriota	Vicinamibacteria	0	0	0	14	1	5	3
BPF1-52-54_bin_4	Actinobacteriota	Thermoleophilia	0	0	0	1	1	5	0
BPF2-00-02_bin_1	Actinobacteriota	UBA4738	0	1	0	1	1	2	0
BPF2-00-02_bin_2	Acidobacteriota	Blastocatellia	0	0	0	20	1	7	4
BPF2-02-04_bin_1	Acidobacteriota	Blastocatellia	0	0	0	18	1	5	0
BPF2-04-06_bin_1	Acidobacteriota	Blastocatellia	0	0	0	22	1	4	1
BPF2-06-08_bin_3	Acidobacteriota	Blastocatellia	0	0	0	18	1	6	1
BPF2-06-08_bin_5	Acidobacteriota	Mor1	0	0	0	2	1	3	0
BPF2-08-10_bin_1	Acidobacteriota	Blastocatellia	0	0	0	23	1	4	1
BPF2-12-14_bin_2	Acidobacteriota	Blastocatellia	0	0	0	24	1	4	1
BPF2-14-16_bin_1	Acidobacteriota	Blastocatellia	0	1	0	25	1	4	1
BPF2-16-18_bin_1	Acidobacteriota	Blastocatellia	0	0	0	24	1	4	1
BPF2-20-22_bin_2	Acidobacteriota	Blastocatellia	0	0	0	26	1	4	1
BPF2-22-24_bin_2	Acidobacteriota	Blastocatellia	0	0	0	24	1	4	0
BPF2-24-26_bin_1	Chloroflexota	Ellin6529	0	0	0	3	1	4	0
BPF2-24-26_bin_2	Acidobacteriota	Blastocatellia	0	0	0	19	1	4	1
BPF2-26-28_bin_1	Acidobacteriota	Blastocatellia	0	1	0	14	1	3	1
BPF2-28-30_bin_1	Acidobacteriota	Blastocatellia	0	0	0	18	1	4	1

Supplemental Table 4-11 continued

Site and Sample Depth	Phyla	Class	AG	BG	CB	LAP	NAG	PHOS	XYL
BPF3-01-02_bin_2	Chloroflexota	Ellin6529	0	1	0	3	1	3	0
BPF3-02-03_bin_3	Chloroflexota	Ellin6529	0	1	0	6	1	1	0
BPF3-02-03_bin_4	Actinobacteriota	UBA4738	0	0	0	2	1	3	0
BPF1-00-02_bin_1	Acidobacteriota	Thermoanaerobaculia	0	1	0	13	0	7	1
BPF1-00-02_bin_2	Bacteroidota	Bacteroidia	0	0	0	3	0	0	0
BPF1-02-04_bin_1	Acidobacteriota	Thermoanaerobaculia	0	1	0	14	0	7	0
BPF1-02-04_bin_2	Planctomycetota	Planctomycetes	0	0	0	7	0	6	0
BPF1-04-06_bin_1	Verrucomicrobiota	Verrucomicrobiae	0	0	0	0	0	1	0
BPF1-04-06_bin_2	Proteobacteria	Alphaproteobacteria	0	0	0	2	0	0	0
BPF1-06-08_bin_1	Acidobacteriota	Thermoanaerobaculia	0	1	0	15	0	7	1
BPF1-06-08_bin_2	Proteobacteria	Alphaproteobacteria	0	0	0	2	0	0	0
BPF1-08-10_bin_1	Acidobacteriota	Thermoanaerobaculia	0	1	0	10	0	2	0
BPF1-08-10_bin_2	Proteobacteria	Alphaproteobacteria	0	0	0	2	0	1	0
BPF1-08-10_bin_3	Actinobacteriota	Thermoleophilia	0	0	0	1	0	4	0
BPF1-08-10_bin_4	Acidobacteriota	Blastocatellia	0	0	0	0	0	4	0
BPF1-10-12_bin_2	Acidobacteriota	Thermoanaerobaculia	0	1	0	15	0	7	1
BPF1-14-16_bin_3	Verrucomicrobiota	Verrucomicrobiae	0	0	0	2	0	1	0
BPF1-16-18_bin_2	Gemmatimonadot a	Gemmatimonadetes	1	2	0	5	0	2	0
BPF1-16-18_bin_3	Acidobacteriota	Blastocatellia	0	0	0	2	0	2	0
BPF1-18-20_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	2	0	6	0
BPF1-20-22_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	2	0	7	0
BPF1-22-24_bin_4	Acidobacteriota	Acidobacteriae	0	0	1	3	0	0	0
BPF1-24-26_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	0	0	4	0
BPF1-24-26_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	2	0	6	0
BPF1-24-26_bin_4	Proteobacteria	Gammaproteobacteria	0	0	0	2	0	1	0
BPF1-26-28_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	2	0	7	0
BPF1-28-30_bin_1	Verrucomicrobiota	Verrucomicrobiae	0	0	0	1	0	0	0
BPF1-28-30_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	2	0	6	0
BPF1-28-30_bin_3	Proteobacteria	Gammaproteobacteria	0	0	0	3	0	4	0
BPF1-28-30_bin_4	Acidobacteriota	Blastocatellia	0	0	0	1	0	3	0
BPF1-28-30_bin_6	Chloroflexota	Ellin6529	0	1	0	3	0	6	0
BPF1-28-30_bin_7	Actinobacteriota	Thermoleophilia	0	2	0	4	0	4	0
BPF1-30-32_bin_1	Proteobacteria	Gammaproteobacteria	0	0	0	3	0	4	0

Supplemental Table 4-12 continued

Site and Sample Depth	Phyla	Class	AG	BG	CB	LAP	NAG	PHOS	XYL
BPF1-30-32_bin_6	Actinobacteriota	Thermoleophilia	0	0	0	2	0	7	0
BPF1-32-34_bin_2	Actinobacteriota	Thermoleophilia	0	2	0	4	0	4	0
BPF1-32-34_bin_3	Proteobacteria	Gammaproteobacteria	0	0	0	3	0	4	0
BPF1-32-34_bin_5	Acidobacteriota	Blastocatellia	0	1	0	3	0	4	0
BPF1-34-36_bin_1	Chloroflexota	Ellin6529	0	2	0	2	0	4	0
BPF1-34-36_bin_4	Proteobacteria	Gammaproteobacteria	0	0	0	3	0	4	0
BPF1-34-36_bin_6	Actinobacteriota	Thermoleophilia	0	0	0	1	0	4	0
BPF1-34-36_bin_8	Actinobacteriota	Thermoleophilia	0	0	0	2	0	6	0
BPF1-36-38_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	1	0	6	0
BPF1-36-38_bin_3	Proteobacteria	Gammaproteobacteria	0	0	0	2	0	2	0
BPF1-36-38_bin_5	Actinobacteriota	Thermoleophilia	0	0	0	0	0	2	0
BPF1-36-38_bin_7	Actinobacteriota	Thermoleophilia	0	1	0	4	0	4	1
BPF1-38-40_bin_3	Actinobacteriota	Thermoleophilia	0	1	0	4	0	4	1
BPF1-44-46_bin_1	Gemmatimonadota	Gemmatimonadetes	1	1	0	10	0	3	1
BPF1-44-46_bin_5	Acidobacteriota	Blastocatellia	0	0	1	2	0	0	0
BPF1-46-48_bin_2	Gemmatimonadota	Gemmatimonadetes	0	2	0	6	0	4	0
BPF1-46-48_bin_3	Chloroflexota	Ellin6529	0	1	0	1	0	2	0
BPF1-48-50_bin_1	Chloroflexota	Ellin6529	0	3	0	2	0	3	0
BPF1-48-50_bin_2	Actinobacteriota	Acidimicrobiia	0	0	0	1	0	1	0
BPF1-50-52_bin_2	Proteobacteria	Gammaproteobacteria	0	0	0	1	0	4	0
BPF1-52-54_bin_1	Acidobacteriota	Vicinamibacteria	0	0	0	4	0	4	0
BPF1-52-54_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	2	0	6	0
BPF1-52-54_bin_3	Chloroflexota	Ellin6529	0	3	0	2	0	2	0
BPF1-54-56_bin_2	Methylomirabilota	Methylomirabilia	0	0	0	4	0	7	0
BPF1-56-58_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	1	0	4	0
BPF1-56-58_bin_2	Eisenbacteria	RBG-16-71-46	0	1	1	3	0	2	0
BPF1-56-58_bin_3	Actinobacteriota	Acidimicrobiia	0	1	0	3	0	5	0
BPF1-56-58_bin_4	Actinobacteriota	Thermoleophilia	0	1	0	1	0	4	0
BPF1-56-58_bin_5	Chloroflexota	Ellin6529	0	3	0	2	0	4	0
BPF2-06-08_bin_1	Acidobacteriota	Blastocatellia	0	1	0	13	0	5	3
BPF2-06-08_bin_2	Proteobacteria	Gammaproteobacteria	0	0	0	2	0	3	0

Supplemental Table 4-13 continued

Site and Sample Depth	Phyla	Class	AG	BG	CB	LAP	NAG	PHOS	XYL
BPF2-14-16_bin_2	Actinobacteriota	Thermoleophilia	0	1	0	2	0	4	0
BPF2-14-16_bin_3	Acidobacteriota	Blastocatellia	0	1	1	3	0	2	0
BPF2-20-22_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	2	0	6	0
BPF2-22-24_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	2	0	6	0
BPF2-24-26_bin_3	Actinobacteriota	Thermoleophilia	0	0	0	1	0	3	0
BPF2-24-26_bin_4	Actinobacteriota	Thermoleophilia	0	0	0	1	0	4	0
BPF3-00-01_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	3	0	6	0
BPF3-00-01_bin_3	Actinobacteriota	Acidimicrobiia	0	1	0	3	0	7	0
BPF3-01-02_bin_1	Actinobacteriota	Thermoleophilia	0	0	0	3	0	7	0
BPF3-02-03_bin_1	Proteobacteria	Alphaproteobacteria	0	0	0	5	0	6	0
BPF3-02-03_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	1	0	4	0
BPF3-18-20_bin_1	Actinobacteriota	Thermoleophilia	0	1	0	3	0	7	0
BPF4-00-01_bin_1	Chloroflexota	Ellin6529	0	1	0	3	0	2	0
BPF4-00-01_bin_2	Proteobacteria	Alphaproteobacteria	0	0	0	7	0	0	0
BPF4-02-03_bin_1	Acidobacteriota	Blastocatellia	0	1	0	9	0	2	3
BPF4-02-03_bin_2	Acidobacteriota	Thermoanaerobaculia	0	1	0	5	0	2	0
BPF4-02-03_bin_3	Actinobacteriota	Actinomycetia	0	0	0	4	0	8	0
BPF4-43-45_bin_1	Acidobacteriota	Thermoanaerobaculia	0	0	1	7	0	1	0
BPF4-43-45_bin_2	Proteobacteria	Gammaproteobacteria	0	0	0	2	0	3	0
BPF4-43-45_bin_3	Actinobacteriota	Thermoleophilia	0	0	0	1	0	2	0
BPF4-43-45_bin_4	Chloroflexota	Anaerolineae	0	6	0	5	0	5	0
BPF4-43-45_bin_5	Chloroflexota	Dehalococcoidia	0	0	0	1	0	1	0
BPF5-00-01_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	1	0	1	0
BPF5-01-02_bin_1	Acidobacteriota	Blastocatellia	0	1	0	12	0	2	1
BPF5-01-02_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	1	0	2	1
BPF5-02-03_bin_1	Acidobacteriota	Blastocatellia	0	1	0	14	0	2	1
BPF5-02-03_bin_2	Actinobacteriota	Thermoleophilia	0	0	0	1	0	5	0
BPF1-36-38_bin_4	Thermoproteota	Nitrososphaeria	0	0	0	1	0	2	0
BPF2-24-26_bin_5	Thermoproteota	Nitrososphaeria	0	0	0	0	0	2	0

CHAPTER 5 CONCLUSION

Siberia and Svalbard

The permafrost in Siberia, Russia, and the active layer of Ny Ålesund, Svalbard are dramatically different environments, despite both being Arctic cryosoils. Microbes in permafrost from Siberia Russia, analyzed in Chapter two of this dissertation, originated from the most ancient frozen soils in the world. While the active layer from Svalbard, analyzed in chapter three and four, is at the cusp at being thawed permanently due to increased global temperatures. The future of permafrost affected soils is dependent upon thaw rates and microbial activity. The time is now to understand the microbial presence in these environments, both in aspects of cataloging an endangered ecosystem and to gain insight of how the present community will respond to the changing climate.

The permafrost in Siberia Russia has been aged up to 3 million years old. Through metagenomic-based investigation, the MAGs from 1 million-year-old permafrost samples contain genes that would make them well suited for the low-energy and frozen environment. When these Siberian MAGs are compared to genomes from other locations, there was a high proportion of genes from COG categories related to energy production and conversion, carbohydrate transport and metabolism and osmoregulation. These MAGs also had pathways that were energetically advantageous in the energy limited environment, such as the mevalonate pathway and trehalose synthase. These pathways would allow organisms to maintain membrane integrity and subsist over longer time scales without requiring a lot of energy. Through this genetic assessment of ancient Siberian MAGs compared to genomes of other environments, it was shown that the MAGs contained features that allowed for long term survival and subsiting in the extreme ancient permafrost.

The two Svalbard active layer soil cores, analyzed in chapter three, displayed geochemical variability, despite being 84 m apart in the same glacier moraine plane. BPF1 and BPF2 cores differ in depth, carbon and nitrogen content, carbon:nitrogen ratios and carbon isotopic signatures. The carbon isotopic signatures in BPF1 display signatures indicative of microbial heterotrophic metabolism. BPF2, however, shows carbon isotopic fractionation signatures of microbial autotrophy the deeper below the surface of the core. These differences in the geological evidence could point towards different dominating microbial metabolisms throughout the soil. Microbial autotrophy isotopic fractionation signatures dominate in BPF2 where there is also less organic carbon and less energetically favorable ratios of carbon:nitrogen. Further, the enzymatic activity assays of the bulk soil and isolated cultures show similar conclusions. Both assays showed the enzymatic activity of leucine aminopeptidase to be the highest. This parallels the energy limitations of the active layer, which would force organisms to use leucine and other amino acids as carbon and nitrogen sources.

Further, the higher activity of N-acetyl-beta-D-glucosaminidase (NAG) at 15°C demonstrates that NAG is adaptable to temperature changes, especially in lower temperatures. The metagenome libraries of these two active layer cores also had gene counts for these enzymes that paralleled the level of activity seen. The ten isolated cultures and the bulk soil had high individual counts of leucine aminopeptidase and little to no counts for the four plant polymer degrading enzymes. In this environment, where there is little plant organic matter, these organisms must utilize other resources. These lines of evidence further support how the microbes from this active layer are adapted to where they live and the resources available now, but also could adapt to higher yearly soil temperatures.

Lastly, the 169 MAGs analyzed from five active layer cores in Svalbard in chapter four show that microbial community structure is best investigated at a higher stratification refinement level. With single sample metagenomic assembly, MAGs were able to be compared across all five sample sites. Patterns that appeared from the 169 MAGs' read mapping showed that there was microbial presence of two classes of *Acidobacteriota (Blastocatellia* and *Thermoanaeobaculia)* most abundant at the surface of BPF1 and the thawed soil of BPF3, BPF4 and BPF5 cores. *Actinobacteriota-Thermoleophilia* is most abundant in the bottom half of the frozen cores and appears evenly in the thawed soil. The network analysis displays that the middle of the BPF1 and BPF2 cores show high correlations with each other based on read mapping across all five depths (Spear. Corr = > 0.7). At a phyla level network analysis, the *Actinobacteriota* and *Acidobacteriota* correlate with different phyla based on distribution in the soil. The *Chloroflexota* phyla co-correlated with the *Actinobacteriota* and was the most dominant phylum in BPF3 top layer thawed soil. The difference in microbial abundance in these Svalbard active layer sites show that MAG depth location matter when assessing how the microbial community will respond to warming temperatures.

In conclusion, Svalbard active layer will warm faster than the Siberian permafrost and the microbes there will be able to metabolize only the material that is present in the soil. As the environment warms, and cryoturbation rate increases, the organisms in the upper parts of the active layer will be faced with variable amounts of soil organic carbon. The enzymatic activity assays show that these soils are capable of activity in the current temperatures and that the enzymes will be able to adapt to a wider range of temperatures. The Siberian permafrost organisms, to the contrary, contain genetic features that are adapted to their current ancient frozen environment. This means that when the permafrost thaws and incorporates into the active layer the Siberian organisms may be outcompeted by organisms that are adapted for the highly variable temperatures in the environment. The future for the microbes inhabiting Svalbard active layer and the ancient Siberian permafrost depend on the rate of temperature change and how microbial activity will respond to the influx of soil organic matter.

VITA

Katie Marie Sipes was born to John and Janet Sipes on October 6th, 1993. Katie went through many ideas throughout her childhood of what profession she wanted to become. Veterinarian and medical laboratory scientist were at the top of the realistic occupation list; actress and circus performer were the top of the not-practical list. At James Madison University, Katie took courses that would fulfill veterinarian school requirements and met Dr. Eva Strawbridge in the mathematics department. Eva saw Katie's determination and decided to convert her to the mathematical dark side. Katie then started undergraduate research in mathematics as a freshman and stayed with Eva until she graduated in 2016. Katie presented mathematical modeling research at a total of 29 conferences all over the United States, won a few presentation awards and realized that this lifestyle was much more fun than being a vet. After attending a vet summer school intensive, Katie decided that doctors that eat their own patients was a moral conundrum and decided that graduate school would be a less morally straining career path. After joining University of Tennessee's microbiology department in 2016, Katie thought that medical laboratory scientist would be a good career path, since the people on ABC's "*House"* had such crazy lives. Katie then met Dr. Karen G. Lloyd and decided to assist a senior graduate student, Joy Buongiorno, in the lab. Karen could not assist Joy on a fieldwork trip to Ny Ålesund, Svalbard, so Karen graciously allowed Katie to be Joy's assistant. This trip to the Arctic would change Katie's life forever.

After this Svalbard trip in 2017, Katie would find her own way to continue fieldwork by developing a project studying active layer and permafrost around Ny Ålesund. Katie fell so head over heels with the Arctic's midnight sun, fieldwork, and scientific comradery that environmental researcher now became the top of her career list. In true stubborn passion, Katie set out to learn all the things (just a little bit) so that she could make herself a good teammate. She learned how to code, learned Swedish and Chinese, took pilot lessons, became scuba certified, and learned how to weld. She even joined the Knoxville Circus arts group and learned aerial silks and hula hooping – just to make sure that circus performer wasn't a desirable occupation still. After recognizing that being in a circus was kind of like being in a cult, Katie reevaluated how to both do science and flip around. Katie realized that she wants to become an astronaut, which takes the first-place prize for both realistic and not-practical occupations. To add to the crazy, Katie purchased 800 pounds of Epsom salt to make a sensory deprivation float tank where she can meditate and pretend that she is in the International Space Station. Katie's shorter-long-term goals include getting ice diver certified and selling her house before she moves to Denmark for her postdoctoral research position with Alexandre Anesio and the Deep Purple research group.

Katie truly loves learning, selling other people's junk on Facebook Market place and being in a team. Katie does not understand dry humor, slight sarcasm and most 90's movie references. Katie has been referred to as an alien by roommates and ex-boyfriends alike and proudly wears this as a badge of honor. One day, she hopes to be reunited with her stardust origins in space.