University of Nevada, Reno

Microbial Ecology, Biogeochemistry, and Signatures of Life at Low Temperature in Arctic Thermokarst Lake Sediments and High Sierra Snow Fields

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

by

Paula B. Matheus Carnevali

Dr. Alison Murray/Dissertation Advisor

May, 2015

THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

PAULA B. MATHEUS CARNEVALI

Entitled

Microbial Ecology, Biogeochemistry, And Signatures Of Life At Low Temperature In Arctic Thermokarst Lake Sediments And High Sierra Snow Fields

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

Alison E. Murray, Ph.D., Advisor

Gary Blomquist, Ph.D., Committee Member

David Shintani, Ph.D., Committee Member

Henry Sun, Ph.D., Committee Member

Kevin Hand, Ph.D., Committee Member

Jerry Qualls, Ph.D., Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

May, 2015

Abstract

The subjects of biological methane (CH_4) production and microbial community diversity and structure in extreme cold environments were at the core of this dissertation. CH_4 is a potent greenhouse gas that contributes to the warming of the planet and has chemical properties that allow its detection by instruments that can go into space. In this sense, CH_4 is also a biosignature, because its presence could be indicative of the presence of life. Mechanisms that lead to the formation of this gas on Earth include reactions mediated by biological enzymes that evolved early in life's evolutionary time. Therefore, microbial communities composed of different metabolic guilds that act in synergy to decompose organic matter are key elements of understanding biological CH_4 production. Two chapters of this dissertation were devoted to studying the potential for biological CH_4 production from arctic thermokarst lake sediments. Thermokarst lakes are abundant in the area of continuous permafrost on the Coastal Plain on the North Slope of Alaska, and they are sensitive to climate warming. In Chapter 2 we detected biological CH_4 production from sediments of two lakes by studying the isotopic composition of C in CH_4 ($\delta^{13}C_{CH4}$) and its concentration in porewaters, and by conducting sediment incubation experiments with no substrate additions. At CH_4 concentrations no higher than ~ 2.2 μmoles CH_4 g^{-1} dry sediment, an average $\delta^{13}C_{CH4}$ signature ~ 72 $_0/_{00}$ in the surface sediment intervals of Siqlukaq Lake (Siq) was indicative of biogenic CH_4, while a $\delta^{13}C_{CH4}$ signature ~ 51 $_0/_{00}$ indicated that in Sukok Lake (Suk) there was a mixed signature in the range of thermogenic CH_4. We also determined that the highest amount of CH_4 produced in the lab at 10 °C (~ 7 μmoles CH_4 g^{-1} dry sediment) occurred in Siq, the lake where we detected four times more organic carbon than in the other two sites (SukS and SukB). These findings were supported

by high amounts in Siq sediments of archaeal lipid biomarkers and by copies of the *mcrA* gene, which is specific to the pathway of methanogenesis. Furthermore, we determined that the geochemical environment of these lakes was different, with higher concentrations of dissolved iron (Fe) in Siq (51.6 - 855 μM) than in SukB (0.1 - 482 μM) and SukS (33.8 - 141 μM); higher concentrations of sulfate (SO_4^{2-}) in SukB (0 - 752 μM) and SukS (0 - 329 μM) than in Siq (0.0 - 23.9 μM); and similar concentrations of nitrate (NO_3^-) at all sites (0.0 - 213 μM) except for a couple of depths in SukB where NO_3^- was much higher (up to 2.3 mM).

These differences in sediment geochemistry led us to investigate what the diversity and structure of the bacterial and archaeal assemblages was in the sediments of these two lakes. In Chapter 3 we sequenced the small subunit (SSU) ribosomal RNA (rRNA) gene and the SSU rRNA molecule of Bacteria and Archaea by Sanger and iTag sequencing (Illumina MiSeq). iTag sequence analyses showed that the microbial assemblage was highly diverse (maximum Simpson's reciprocal index of 148) and that there was a difference in operational taxonomic units (OTUs) predicted for DNA and aRNA sequences between Siq, SukB, and SukS, as reflected by non-metric multidimensional scaling (NMDS). OTUs related to taxa that are capable of organic matter degradation were found among the OTUs comprising ≥ 1 % of the aRNA sequences, which were taken as a proxy for members of the community that may have been potentially active in the sediments at the time of sampling. These taxa included members of the phyla Actinobacteria and Bacteroidetes and represented up to 10 % of the sequences found in all the sites. There was an important presence (> 5 % of the sequences) of sulfate reducing bacteria (SRB) affiliated with the orders Syntrophobacterales and Desulfobacterales at SukB. OTUs at this site were also related to Gammaproteobacteria of the order Pseudomonadales, and to a few groups of unicellular Cyanobacteria. In Siq, OTUs related to the phylum Nitrospirae comprised ~ 5 %

of the sequences, and the same group comprised ~ 2.5 % in SukB. Further investigation of the functional potential of the community using phylogenetic investigation of communities by reconstruction of unobserved state (PICRUSt) supported the potential for microbial mediated transformations of nitrogen and sulfur, and for CH_4 cycling. One of the most noteworthy findings was that we observed distinct differences in the distribution of Archaea between the two lakes, in which methanogenic Archaea were mostly found in Siq. Through phylogenetic inference of all archaeal sequences we determined that archaea in these arctic thermokarst lakes are evolutionarily related to other archaea in the Arctic. As an appendix to this work, we also conducted a culturing effort to enrich for methanogenic Archaea adapted to this cold environment. This effort led to enrichment of several methanogens in co-culture with bacteria, augmenting the short list of cold-adapted methanogens in culture.

The last chapter of this dissertation focused on an entirely different kind of cold tolerant, ice loving microbial community – those inhabiting high altitude snow packs. In these habitats, a variety of microorganisms including unicellular algae, bacteria, and fungi can be found under freezing conditions during the seasonal snowpack. Here, we aimed to uncover the diversity and structure of the microbial community in summer snowpacks where conspicuous red algae were present. We also aimed to understand the mechanisms that drive community assembly in this ephemeral habitat. Snow samples were collected from Mt. Conness (CNS) in the eastern Sierra Nevada of California (as our primary study site), and two other geographically distant sites: the Pacific Crest Trail (PCT) and Wheeler Peak (WP). Chemical analyses of inorganic and organic nutrients and Chlorophyll *a* (Chl*a*) were conducted in CNS, and the snow microbiota in samples from all three sites was studied by iTag sequencing of the SSU rRNA gene of the three domains of life. Detailed analysis of

the resampled data set of bacterial and archaeal iTag sequences indicated that the bacterial assemblage had a unique structure. In addition to commonly observed 'singletons' (OTUs with just 1 sequence) that dominate NGS datasets, we found OTUs that were only present at one site and comprised a high number of sequences ('singleton OTUs'). Interestingly, these OTUs were mostly affiliated with two different families of Bacteroidetes in the genera *Polaromonas* and *Herminiimonas*. Furthermore, even though the bacterial assemblage was up to 80 % dissimilar between sites at the OTU level, there was a resemblance in taxonomic composition at the Order level among all three sites, and a few 'core' OTUs affiliated with the Phyla Bacteroidetes and Actinobacteria were found in all samples. Variation in structure of the abundant OTUs between sites (and depths) was reflected in the proportion of sequences related to two families of Betaproteobacteria (the Comamonadaceae and the Oxalobacteraceae) and groups of Bacteroidetes. To elucidate possible mechanisms for microbial community assembly and to explain bacterial assemblage structure we tested two hypotheses: (i) that the presence of snow algae determined bacterial assemblage structure, and (ii) that environmental factors affected the structure. A Spearman's rank correlation between chlorophyll a (Chl*a*; as a proxy to snow algae) and bacterial diversity, indicated that bacterial diversity decreased with Chl*a* concentration. This finding was consistent with three different kinds of Speraman's rank correlations (positive, logarithmic, and negative) between specific algal OTUs (from a parallel data set) and specific bacterial OTUs, including OTUs that were present in all three sites. A relationship between environmental factors and bacterial OTUs was found for only two of the environmental factors studied (DOC and PN). Further testing using multivariate statistical approaches may allow us to formulate possible explanations to how geographically distant regions share similar microbial community

composition at this high taxonomic level, and determine what mechanisms drive community assembly in these snow packs.

Together, the chapters of this dissertation allowed me to address first order questions concerning life in extremely cold environments that have yet to be well-studied, and that are particularly sensitive to climate warming. Gaining a better understanding of biogeochemical cycling in these ecosystems, may allow us to predict the impact of environmental change on the resident microbial communities, and ultimately how will that affect the cycling of carbon and other major nutrients. Furthermore, these ecosystems are of great interest, because they may be used as analogs to study biosignatures relevant to the search for life in other icy worlds in the Solar System.

Acknowledgements

There are not enough words to express my gratitude to everyone who made this dissertation possible, but I want them to know how grateful I am.

Alison, my Ph.D. advisor, who along this process became a true mentor. I believe she will have a lasting positive impact on my career as a scientist for years to come. From her I learned how science in done with the highest levels of professionalism, and for that I respect and admire her. I thank her for pushing me past my limits, because now I know what I am truly capable of. I thank her for giving me unimaginable and countless opportunities to learn and grow, and most importantly, for having the patience to let me do it. She made me stronger, and for that I will always be grateful.

Kevin Hand, who led the NASA Astrobiology of Icy Worlds Investigation 3 team. I thank him for being a great leader, for creating amazing opportunities for all of us, and for inspiring me to dream even bigger. His hard work has already made my dreams come true!

Megan Rorhssen, a true warrior and collaborator. I thank her for letting me learn from her, for giving me the strength and courage to make things happen, and for working by my side when we were already exhausted. Mostly, I thank her for being my friend.

Dan Berisford, who helped me overcome several of the technical challenges of sampling the arctic lakes, and even taught me how to use a drill. His friendship and generous support made the most difficult things look better and brighter to me.

Andy Klesh and John Leitchy, I thank them for all the support they provided in the field, even at the expense of their own sleep. Their work and dedication is admirable, and it was most inspiring to me!

John Priscu, I thank him for showing me what many years of experience in polar field work looks like, and for giving me a new goal to reach. I also thank him for all of his support, and for always keeping things interesting!

Pamela Santibanez, Alex Michaud, Yonqin Liu, Ralph Lorenz, and everyone whose participation in field work made it possible to do this research. Special thanks to Heather Adams who organized the first few field campaigns, and started to uncover what was interesting about the different arctic lakes we surveyed.

Jeremy Dodsworth and Brian Hedlund, who taught me everything I know about anaerobic work, and provided enormous amounts of support that made the culturing effort possible. I especially thank Jeremy for being a true teacher, he was always there for me, patient and caring, and I will always be grateful for everything I learned from him.

Everyone who was part of the Murray Lab during my tenure as a Ph.D. student. I thank them for the technical help, moral support, and their friendship: Vivian Peng, Ema Kuhn, Gary Trubl, Olivia Rassuchine, Emily Ulrich, Michael Blayberg, Robert Read, Charlotte Tyler, Protima Wagh, and the newest additions to the lab, Mary Highman and Samantha Mayer.

Chris Fritsen and his group, especially Jeramie Memmott, for all the help and support he provided. I thank Chris for giving me great opportunities to learn!

DRI staff, who were always helpful: Barbara Jackson, Cindy Littlefield, Petra Bartella, Russ Bergin, Wendy Long, Ryan Coots, Maria Vasquez, Melissa McCoy, Martha McRae, Denisse

Hamond, Paul Neely, John Karlas, Glenn Wilson, Mitch Goldfin, Lisa Wable, John Ford, James Rios, Melanie Scott, Wayne, Mike, Jeff…and the list goes on!

Other faculty and researchers at DRI who were there when I needed them most: John Mejia, Mary Miller, Hans Moosemuller, Duane Moser, Sophie Baker, and Bill Coulomb. Also, Jay Arnone and Alan Gertler, for providing other opportunities to develop my career as a scientist.

Grad students at DRI who were friends and companions on this journey: Kathy Bywaters, KC King, Megan Johnson, Eric Wirthlin, and Clint Davis.

My committee members, Gary Blomquist, Jerry Qualls, David Shintani, Henry Sun, and Kevin Hand, who were very supportive of me, and gave me encouragement and great advice along the way.

Dr. Glenn Miller, who was most helpful with GC work, and was always there when I needed help.

Everyone at BASC and UMIAQ, for keeping us safe in the field, and for providing all the logistics support and good stories.

DEES directors over the years, who supported Alison and made it possible for me to be at DRI all these years. Thanks to DRI for its financial support, and for being an institution that trains students under the highest of expectations, so that they leave prepared for the challenges of being a scientist.

The NASA Astrobiology Institute, which provided the funding for this work through the Jet Propulsion Lab, and all the other members of the Astrobiology of Icy Worlds Projects.

My friends and family, who have supported me unconditionally over the years, and always give me a reason to want to reach for the stars.

Leo, my best friend, husband, and partner, the only person who knows all the sacrifices we have made to reach this goal, and whose support has made reaching this point possible. Simply put, I could not have done this without him; he kept me going when I felt like giving up, he was the ever present light that guided my journey, and his unconditional love and support never cease to amaze me. I thank him for being in my life all of these years.

Table of Contents

List of Tables

List of Figures

Chapter 1

General Introduction

There are two underlying themes of my dissertation research: signatures of life in cold environments that could potentially be used to search for life elsewhere, and the importance of microbial communities involved in carbon cycling as it influences the greenhouse gas inventory in our planet. These topics are related to the goals described in the Astrobiology Roadmap, a document created by the NASA Astrobiology Institute to guide the design and execution of research in the context of Astrobiology, which also provides a framework for my dissertation.

Astrobiology seeks to answer some of the most fundamental questions that humankind has yet to solve, what is the origin of life on Earth? and what is the extent of life in the universe? To answer these questions we must first establish what forms of life exist on Earth and what the limits are for life as we know it. Life is defined as a sum of characteristic features that include the presence of a membrane, metabolic activity, and most importantly, the ability to reproduce. Studying how these features have co-evolved with our planet may be the key to understanding what we need to look for in other planets, and it is among the goals outlined by the Astrobiology Roadmap (Des Marais et al., 2008).

My dissertation was consistent with goals 4, 6, and 7 as described in the Astrobiology roadmap. Goal 6 outlines the need to understand how microbial metabolic activities will respond to environmental change in light of increasing evidence of human-induced change of geochemical cycles and their associated microbial communities (Des Marais et al., 2008). Chapter 2 characterizes the geochemical environment of arctic thermokarst lake sediments that are known to

produce methane (CH_4), a potent greenhouse gas critical in carbon (C) cycling. Rates of biological CH_4 formation (methanogenesis) in lake sediments were obtained, which could be used in current estimates of CH_4 emissions from the Arctic. The effect of a temperature increase on methanogenesis was also addressed (relevant to goal 4 in the Astrobiology Roadmap too), and a baseline for prediction of scenarios of future change was provided.

Goal 4 in the Astrobiology Roadmap recommends to perform comparative genomic studies of microbial diversity, to understand the metabolic requirements of these microbes; to study how microorganisms affect the cycling of essential nutrients (*e.g.*, C, N, S, O), and what sort of record is left by their metabolic activities on Earth (Des Marais et al., 2008). Chapter 3 addressed the issues of microbial diversity in thermokarst lake sediments, the metabolic potential of the microbial community inhabiting these sediments, and the likely involvement of these microbes in geochemical cycling.

Goal 7 in the Astrobiology Roadmap is to identify biosignatures that are indicative of past or present life on Earth, and that can be used to detect life in extraterrestrial samples or that could be measured remotely (Des Marais et al., 2008). Biosignatures are defined as features that can emerge from life, and include atmospheric gases and molecular structures that can be found on a planet's surface (Pilcher, 2003; Seager et al., 2012). Overall, my dissertation work looked at CH_4 as a biosignature that results from a chain of microbial metabolic activities codependent on each other, and at photosynthetic pigments that are found in microbial life forms

and that are easily detected by remote-sensing technologies. Additionally, I believe that a molecular marker used in this work to investigate microbial community diversity and structure *i.e.*, the SSU rRNA gene could be considered another kind of biosignature.

It was this goal of identifying biosignatures that incentivized the overall project of which my dissertation was only a small part of. Our investigation within the Astrobiology of Icy Worlds project, led by the NASA Jet Propulsion Lab (JPL, Pasadena, CA), aimed to study the detectability of spectral signatures beyond photosynthetic pigments that are of potential relevance to the search for life in icy worlds of the Solar System. Studying CH_4 emissions from arctic thermokarst lakes in Alaska, allowed our collaborators at JPL to work on the issue of detectability, and allowed us to determine what the potential for biological CH_4 production from these lakes was (Chapter 2 & 3). Chapter 4 was motivated by the same theme of detectability, in this case of spectral signatures in summer snow fields, potentially laden with algal biomass that can be remotely sensed.

In the three main chapters of the dissertation, I addressed first order questions concerning life and associated processes in habitats that are sensitive to climate warming and relevant analogs to other Icy Worlds in our Solar System. This work contributes to the fundamental understanding of microbial life in high latitude and high altitude environments where freezing temperatures are prevalent and life experiences environmental extremes.

The next few sections in this general introduction provide background information concerning: (i) why CH_4 is of so much interest to us as a signature of life, (2) describe how CH_4 is produced by microorganisms, (3)what other sources of CH_4 are there (or its pitfalls as a biosignature), (4) how important CH_4 is in the context of the greenhouse gas inventory, (5) arctic lakes as an example of ecosystem that are major sources of CH_4 on Earth, and (6) what kind of life thrives in this and other kinds of cold environments. This chapter ends with the main goals of my dissertation research.

1. How to search for life? Methane as a biosignature

The transmission spectra of certain atmospheric gases are potentially detectable using remote-sensing technologies, a desirable feature of biosignatures. What determines the detectability of atmospheric gases is a combination of factors that affect their absolute concentration in the atmosphere, including the accumulation time, global mixing with other gases, and their atmospheric lifetime, which is partially affected by radiation from the parent star (Seager, 2013). Mixing processes require time from the initial step in the chain of metabolic reactions that lead to gas formation, to allow for extensive accumulation in the atmosphere, this last step being limited by the rate of destruction under atmospheric conditions. Extensive atmospheric accumulation is essential for detectability because biosignature gases could be locally produced and therefore undetectable in a given atmospheric mix.

Oxygen (O_2) is currently considered the dominant biosignature on Earth, under the assumptions that biosignature gases result from specific steps along metabolic pathways (as the byproduct of a set of redox reactions), and that they accumulate at a rate faster than their destruction (McKay, 2011; Seager et al., 2012). Oxygen makes up 21 % of the atmosphere and it is the byproduct of photosynthesis, a ubiquitous metabolic process that captures the energy of sunlight and enables the formation of cellular components. Our current understanding is that O_2 started to accumulate in the atmosphere after the Great Oxidation Event ~ 2.4 billion years ago (Ga), when oxygenic photosynthesis became a widespread process among photosynthetic bacteria. Enough time passed to allow for atmospheric accumulation, which also meant a change in the redox state of the planet from reducing to oxidizing. However, under certain planetary conditions, for instance, planets with high atmospheric concentrations of water vapor, O_2 levels could increase due to photodissociation instead of photosynthesis (Seager, 2014). Therefore, one important aspect of the use of atmospheric gases to infer the presence of life is the understanding of their environmental or planetary context (Segura et al., 2007).

Similarly, the presence of CH_4 in an atmosphere must be interpreted in its planetary context, because some gases that are byproducts from metabolic reactions could also have alternative sources (Seager, 2014). For instance, even though ~ 95 % of atmospheric CH_4 on Earth is microbially produced, there are other sources of CH_4 that do not involve biology. CH_4 can be formed by break down of buried organic

matter at high temperature and pressure (thermogenesis), or as a result of serpentinization reactions in hydrothermal vents (*e.g.* Lost City Hydrothermal Field in the mid-Atlantic ocean). During serpentinization, ferrous iron–rich rocks are oxidized by hot water, thus reducing the water, and producing hydrogen (H_2), which in the presence of CO_2 would form CH_4 independently from biology (Schrenk et al., 2013). The trace amount of CH_4 recently detected by the Curiosity rover on Mars would be another good example of alternative sources of CH_4. CH_4 there seems to be generated by a combination of mechanisms that may include photochemical degradation of exogenous organics, release from clathrates, or perhaps biology (Webster et al., 2015; Atreya et al., 2007).

Furthermore, as with other biosignature gases, CH_4 is not easily detected remotely using current spectroscopy-based technologies. At the current mixing ratio in Earth's atmosphere the absorption band of CH_4 at 7.6 μm would be overshadowed by the strong water absorption at 6 μm, making remote detection very difficult (Pilcher, 2003). CH_4 in Mars was measured *in situ* using tunable laser absorption in a closed sample cell (Mahaffy et al., 2012) and was detected at the 3.3-μm band in the infrared (IR) spectrum. In places like Europa, an icy moon of Jupiter, the search for CH_4 may be even harder, because CH_4 would have to be detected in ice using the 2.27-2.46 μm bands in the near-IR spectrum, where CH_4 shows weaker features (Hand et al., 2010).

Finally, the search for biosignature gases in the solar system and in remote exoplanets is constrained by the parameters that define a "habitable zone" around

the parent star(s). On Earth, when we think of life, we think of water, and it is precisely water that determines the potential for life in a given planet, because liquid water only occurs in a certain temperature range and a given pressure (McKay, 2011). Without water, biology as we know it would not exist, O_2 would not be produced by photosynthesis, and perhaps CH_4 would still be the dominant gas in the atmosphere. That is why an atmospheric biosignature gas would be in complete chemical disequilibrium with the atmosphere, *i.e.*, it would be several orders of magnitude higher than if generated by photochemical disequilibrium (Seager, 2014), even though that also means that it could be a false positive. This challenge has lead Segura (2005) to propose the use of both O_2 and CH_4 together as biosignatures, with the caveat that these atmospheric gases could only be found at detectable concentrations in UV radiation environments lower than Earth's radiation environment (Seager, 2014).

2. Methanogenesis

Before Earth's atmosphere became oxidized, it was full of CH_4. Methanogenesis is the process in which microorganisms of the domain Archaea obtain minimal amounts of energy for growth (as low as - 3.2 KJ/reaction under specific environmental conditions, see below) and release CH_4 as a result, and it is one of the most ancient forms of microbial metabolism. Methanogenic Archaea are among the oldest microorganisms in the tree of life, and methanogenesis seems to have originated between 4.11 Ga (3.31-4.49 Ga) and 3.78 Ga (3.05-4.16 Ga) (Battistuzzi et

al., 2004; Ueno et al., 2006). That means that Archaea may have arisen between the Hadean and the Archaean, during the time of the "Faint Young Sun Paradox", when the sun was 20-25 % less bright than today but the Earth was kept warm by greenhouse gases (*e.g.*, CO_2 and CH_4) that may have been formed by microorganisms like methanogenic Archaea (Kasting & Ono, 2006).

In modern times, to obtain substrates for methanogenesis, organic matter must be first broken down by primary fermenters, including bacteria and fungi that are able to decompose complex organic molecules such as cellulose and lignin. The products of this step are short-chain fatty acids (*e.g.*, acetate and formate), some H_2 and some CO_2, which can be utilized by methanogens and other kinds of bacteria in anaerobic reactions. The remaining partially decomposed organic matter with three or more carbon atoms (*e.g.*, butyrate and propionate) is then taken by secondary fermenters that make more acetate, H_2 and CO_2. However, this step is only possible at low concentrations of H_2 (< 1 μM) and formate (< 100 μM) that would enable regeneration of the NAD^+ consumed during acetate formation by fermentation. If the bacteria carrying out this fermentation require the presence of methanogens to keep the levels of H_2 and formate low, the relationship between these microorganisms is considered to be a form of obligate syntrophy, otherwise it would be considered a form of non-obligate interspecies electron transfer (Whitman et al., 2006).

An example of obligate syntrophy between a bacterium and an archaeaon that may be relevant in the context of this dissertation research follows.

Syntrophomonas sp., a fatty acid-oxidizing bacterium oxidizes butyrate to acetate and H_2 under thermodynamically unfavorable conditions ($\Delta G^{0'}$ = + 48.2 KJ/reaction), which means that it cannot grow on its own; but in the presence of hydrogenotrophic methanogens that use this H_2 (or formate) to reduce CO_2 ($\Delta G^{0'}$ = - 130.7 KJ/reaction), the combined reactions will be thermodynamically favorable ($\Delta G^{0'}$ = - 82.5 KJ) (Madigan et al., 2014) and inhibition of fatty acid oxidation by end-product formation can be prevented.

Another form of methanogenesis, acetoclastic methanogenesis, is a disproportionation reaction in which the methyl C of acetate acts as the electron acceptor in the oxidation of the carboxyl C to CO_2, turning it into CH_4. Similarly, a variety a C_1 compounds such as methanol, trimethylamine, dimethylsulfide, and some alcohols can be disproportionated to form CH_4. The standard change in free energy of some of these reactions is enough to make 1 ATP ($G^{0'}$= -31.8 KJ/mol) or less per mole of CH_4, especially under environmental conditions which tend to include lower concentration of substrates. For instance, the free energy change under standard conditions (ΔG; 1 M solutes, 1 atm gases, 25 °C) for the reduction of CO_2 with H_2 to CH_4 is -136 KJ/reaction, and -31 KJ/reaction for the conversion from acetate; but in a typical anoxic freshwater environment the concentrations of substrates are much lower (1 mM fatty acids, 20 mM HCO_3^-, 10 μM glucose, 10^{-4} atm H_2) and the free energy change for hydrogenotrophic methanogenesis is -3.2 KJ/reaction and -24.7 KJ/reaction for acetaoclastic methanogenesis (Zinder, 1984).

The different pathways for methanogenesis involve several coenzymes and cofactors (see Ferry & Kastead, 2007) for a complete review of the biochemistry), but two of those coenzymes are common to the different pathways: coenzyme M and coenzyme B (Fig. 1.1). Coenzyme M (HS-CoM) accepts the methyl group from acetate in acetoclastic methanogenesis (CH_3-S-CoM) and acts as the substrate for a key enzyme in all methanogenic pathways, the methyl coenzyme M reductase (Mcr), which catalyzes the reductive demethylation of CH_3-S-CoM to CH_4 and the subsequent reduction of coenzyme B (HS-CoB). The resulting heterodisulfide CoMS-SCoB is then reduced by a heterodisulfide reductase to the active sulfhydryl forms of the coenzymes, generating an electrochemical proton gradient that involves an A_1-A_0-type ATP synthase and driving ATP synthesis. However, the electrons participating in this reduction originate from the carbonyl group of acetate in the acetoclastic pathway, or from the oxidation of H_2 or formate in the hydrogenotrophic pathway (Ferry, 2010). It is important to note that there are two mechanisms by which methanogens conserve energy: a sodium motive force associated to hydrogenotrophic methanogenesis, during transfer of the methyl group from methanopterin (one of the coenzymes that functions as a C_1 carrier) to CoM; and a proton motive force during both acetoclastic and hydrogenotrophic methanogenesis, linked to the reaction catalyzed by the heterodisulfide reductase (Ferry, 2007; Madigan, 2014}.

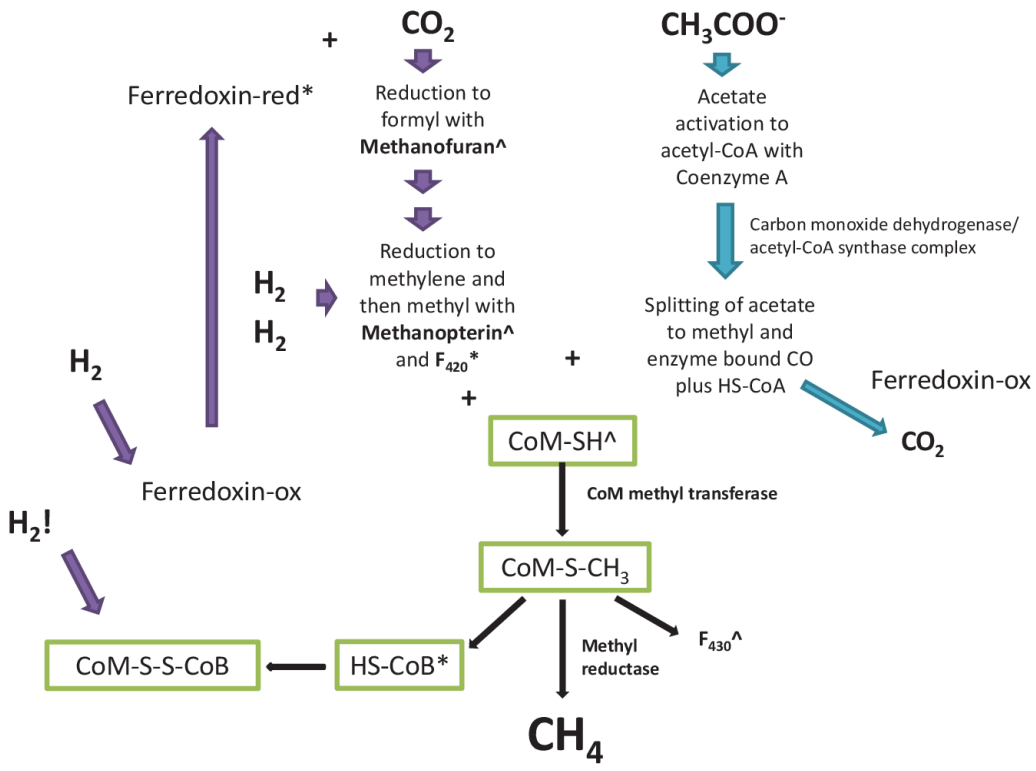


Figure 1.1 Two of the main pathways of methanogenesis and their common steps. In the CO_2-reduction pathway reduced ferredoxin (Ferredoxin-red) donates electrons to CO_2 in a reaction catalyzed by an enzyme containing methanofuran, the first C_1 carrier (^) in this pathway. In the next step, a tetrahydromethanopterin (methanopterin^) coenzyme receives the formyl group from methanofuran, and the formyl group undergoes dehydration to methenyl, then reduction to methylene coupled to F_{420} (*electron donor) and to methyl, with electrons from H_2. Coenzyme M (CoM-SH^) methyltransferase relocates the methyl group from methanopterin to CoM, the first common step between hydrogenotrophic and acetoclastic methanogenesis. The last step, leading to CH_4 formation involves the key enzyme methylreductase, which is a highly complex system (containing subunit α coded by the *mcrA* gene) that interacts with coenzyme B (HS-CoB*) and cofactor F_{430}^. The resultant disulfide complex (CoM-S-S-CoB) is regenerated with electrons from H_2, the same way oxidized ferredoxin (Ferredoxin-ox) is regenerated. The first two steps during acetoclastic methanogenesis convert acetate to acetyl-CoA. Then, the carbon monoxide dehydrogenase/acetyl-CoA synthase complex catalyze the cleavage of the C-C and C-S bonds, the transfer of the methyl group to methanopterin^, and the oxidation of the carbonyl group to CO_2 while reducing ferredoxin-ox. This last set of reactions provides the electrons to reduce CoM-S-CH_3 to CH_4 (Whitman et al., 2006; Ferry & Kastead, 2007; Ferry, 2010).

Methanogenic Archaea can be outcompeted by other microorganisms like sulfate-reducing bacteria or iron-reducing bacteria, which in comparison carry out more energy efficient reactions. The main issue is that many of these microorganisms can use the same electron donors (by products of organic matter decomposition) for their respective metabolic pathways. For example, Group II marine sulfate-reducing bacteria (complete oxidizers that are able to oxidize acetate and other long chain fatty) can obtain energy from acetate oxidation (except *Desulfobacter sp.*): $CH_3COO^- + SO_4^{2-} + 3H^+ \rightarrow 2\ CO_2 + H_2S + 2H_2O$ ($\Delta G^{0'}$ = - 57.5 KJ/reaction) (Madigan et al., 2014). Another group of organisms that can use acetate as an electron donor belong to the family Geobacteraceae in the Deltaproteobacteria. Members of the genera *Geobacter*, *Desulfuromonas*, and *Desulfuromusa* can oxidize acetate during dissimilative ferric iron reduction (Madigan et al., 2014).

3. How is biogenic CH_4 distinguished from abiogenic CH_4?

CH_4 can also be formed by alternative mechanisms that are abiogenic. Biology, more specifically, enzymes, discriminate against substrates with heavier isotopic compositions in their atoms, a process known as isotopic fractionation. In the case of the substrates for methanogenesis, compounds with the lighter carbon isotope (^{12}C) would be strongly preferred by microbial enzymes over compounds with the heavier carbon isotope (^{13}C), resulting in lighter CH_4 relative to the parent substrates. We can use $\delta^{13}C$, that is the ^{13}C to ^{12}C ratio compared to the Pee Dee

belemnite standard (a 65- to 150-million year old limestone formation and the geologic standard for stable carbon isotope analysis) to determine the origin of the CH_4. ^{13}C-depleted CH_4 is indicative of biological CH_4 production, while ^{13}C-enriched CH_4 is associated to thermogenic CH_4 formation (Whiticar, 1999).

However, to confirm the origin of CH_4 it is useful to look at the stable isotopes of both C and H atoms in CH_4. The same way the carbon in CH_4 originates from the substrates for methanogenesis, some of the hydrogen atoms in CH_4 are derived from the water and substrates available to methanogens (Whiticar, 1999; Boone, 2000), and there is a preference for the lighter isotope (^1H) over deuterium (^2H). The $\delta^{13}C_{CH4}$ can be plotted against the ^2H to ^1H ratio (δD) relative to the standard mean ocean water (SMOW standard) in a CD-diagram to identify the source of CH_4 (Whiticar, 1999). Furthermore, in combination with molecular or compositional data on short-chain hydrocarbons (C_2-C_4), the origin of CH_4 can be determined more accurately. This is especially useful when the isotopic signature of CH_4 is borderline thermogenic, because there is overlap with the δD_{CH4} of certain kinds of biogenic CH_4 (Whiticar, 1999). Noticeably, the following factors affect the isotopic composition of CH_4: the isotopic fractionation in the parent organic matter, which is ultimately determined by the isotopic composition of CO_2; kinetic isotope effects that determine the isotopic fractionation of CH_4 relative to the parent substrates, and in the case of atmospheric CH_4 sources, it is also important to consider the effect of CH_4 oxidation on isotopic fractionation (Whiticar, 1999; Whiticar, 2000).

Additional information can be gleaned from the δD_{CH4} and the source of hydrogen, because the aceticlastic pathway uses three hydrogens from the methyl groups and one from water, while the hydrogenotrophic pathway extracts all the hydrogen from water or protons (Whiticar, 1999). Moreover, one can use the isotopic fractionation of C in CO_2 to determine the isotope separation factor (εc) between $\delta^{13}C_{CO2}$ and $\delta^{13}C_{CH4}$, which is relative to the pathway of CH_4 formation (Whiticar, 1999).

4. Methane as a greenhouse gas

There is an equally important reason to look at CH_4 in Earth's atmosphere besides its astrobiological relevance. CH_4 is the third main contributor to the greenhouse effect, but it is ~ 30 times more potent as a greenhouse gas compared to CO_2 over a 100-year time horizon (Lelieveld, 1998; IPCC, 2001). Moreover, CH_4 is rising at a rate of 0.4 % yr^{-1} (Kirschke et al., 2013) although the Intergovernmental Panel for Climate Change (IPCC) estimated a total source of 598 Tg CH_4 yr^{-1} in 2001, and later adjusted that value to 582 Tg CH_4 yr^{-1} in 2007 (IPCC, 2007) due to changes in CH_4 sinks over time. These sources of CH_4 are classified as natural (wetlands, termites, ocean, hydrates, geological sources, wild animals and wild fires) or anthropogenic sources (energy, coal mining, gas and oil industry, landfills and waste, ruminants, rice agriculture, biomass burning, and vegetation). Thus, both kinds of CH_4 sources can be of biogenic (microbial) or non-biogenic origin. In particular, natural gases

can be formed by microorganisms, thermogenic processes, hydrothermal and geothermal activity, and in the mantle (Whiticar, 2000).

A distinction must be made between current biological CH_4 formation and biological CH_4 formation that happened in the past. The ^{14}C content of CH_4 can be used to determine the age of the C source, because only modern organic carbon that is < 35,000 yrs old contains ^{14}C (radiocarbon). Furthermore, the age of the C source is a useful proxy to detect thermogenic CH_4, which is formed from ^{14}C-depleted organic carbon (Kvenvolden & Rogers, 2005). Based on the isotopic composition of C in the atmosphere (reviewed by Stevens & Wahlen, 2000; Whiticar, 2000) biogenic or young CH_4 contributes the most to the CH_4 budget (Kvenvolden & Rogers, 2005). However, a high degree of uncertainty in the total CH_4 budget is reflected in the most current estimates for global CH_4 emissions that attribute > 70 % of the total CH_4 sources to modern biochemical sources including: wildfires (~ 3 Tg CH_4 yr^{-1}), oceans (~ 9 Tg CH_4 yr^{-1}), wild animals (~ 15 Tg CH_4 yr^{-1}), termites (~ 22 Tg CH_4 yr^{-1}), and wetlands (~ 175 Tg CH_4 yr^{-1}); and place the geologic sources at ~ 30 % of the total CH_4 sources (~ 175 Tg CH_4 yr^{-1}) (IPCC, 2007; Etiope, 2008). According to these estimates, geologic CH_4 sources comprise ~ 53 Tg CH_4 yr^{-1} and include emissions from volcanoes (< 1 Tg CH_4 yr^{-1}), geothermal sources (2.5-6.3 Tg CH_4 yr^{-1}), mud volcanoes (6-9 or 10-20 Tg CH_4 yr^{-1}), seeps (3-4 Tg CH_4 yr^{-1}), microseepage (10-25 Tg CH_4 yr^{-1}), marine seepage (~ 20 Tg CH_4 yr^{-1}), and serpentinization (unknown Tg CH_4 yr^{-1}). However, process-based modeling (bottom-up estimates) place geologic CH_4 estimates at 60 Tg CH_4 yr^{-1}; and inverse

modeling based on concentration, isotopic atmospheric composition, aircraft and satellite, spatially distributed and uninterrupted observations (top-down estimates) place geologic CH_4 estimates at 80 Tg CH_4 yr^{-1} (Etiope, 2012). Additionally, anthropogenic fossil emissions are estimated to be ~ 90 Tg CH_4 yr^{-1} (Denman et al., 2007), which leaves a residual fossil emission source (~ 32 Tg CH_4 yr^{-1}) yet to be identified (Etiope et al., 2008). Each one of these sources of CH_4 is regulated by mechanisms of CH_4 consumption or by CH_4 sinks that ultimately determine the net CH_4 flux to the atmosphere.

This work was focused on a particular kind of natural CH_4 source, northern wetlands, which have been estimated to contribute < 6 to 40 Tg CH_4 yr^{-1}(IPCC, 2001; Zhuang et al., 2006). Furthermore, IPCC (2007) established that Polar Regions and perennially frozen environments are the most vulnerable to climate change. For instance, the North Slope of Alaska, located entirely above the Arctic Circle (Kittel et al., 2011) has experienced a temperature increase of 0.5 °C $decade^{-1}$ from 1966 to 1995 (Serreze et al., 2000), and the mean annual temperature in Barrow, the northernmost town in Alaska, has increased 2.4 °C $decade^{-1}$ from 1949 to 2008 based on the historical record (ACRC, 2008). This trend of increasing temperatures have impacted permafrost temperatures in the Arctic by > 3 °C (Zhang et al., 2003; Osterkamp, 2007) affecting the landscape in different ways and specifically impacting CH_4 emissions from thermokarst lakes (Rouse et al., 1997). Current estimates that rely on bottom-up estimates suggest that the Arctic contributes 15-50 Tg CH_4 yr^{-1} to the total CH_4 budget, or 3-9 % of the net source (excluding CH_4 sinks;

Denman et al., 2007; McGuire et al., 2009). Top-down estimates include a 15-35 Tg CH_4 yr^{-1} contribution from thermokarst lakes (Walter et al., 2006)), of which 0.72 Tg CH_4 yr^{-1} correspond to Alaska lakes emissions (Sepulveda-Jauregui et al., 2014), increasing the contribution from the Arctic to 32-112 Tg CH_4 yr^{-1} (McGuire et al., 2009).

Figure 1. 2 Satellite image of the western portion of the Coastal Plain in the North Slope of Alaska, showing terrain occupied by thermokarst lakes and drained lake basins.

5. Thermokarst lakes

Permafrost is defined as ground that remains frozen (≤ 0 ˚C) for a minimum of two years (Davis, 2000). Permafrost underlies ~ 25 % of Earth's terrestrial surface, and recent estimates suggest that there are between 1400 and 1850 Gt of organic carbon

stored in arctic soil stocks (McGuire et al., 2009; Tarnocai et al., 2009; Graham et al., 2012). The top layer of permafrost may thaw seasonally and it is thus known as 'the active layer', with depths that range from 30-60 cm at high latitudes (*e.g.,* Point Barrow, Alaska at 71° 23' N, 156° 28' W) to 300 cm at lower latitudes (*e.g.*, Anchorage, Alaska at 61° 13' N, 149° 54' W) (Brewer, 1958; Péwé, 1982; Washburn, 1980). The bottom of the active layer is known as 'the permafrost table', and the permafrost itself can be hundreds of meters deep, particularly in regions of continuous permafrost that comprise the northern halves of Canada and Alaska, the northern two-thirds of Greenland, and grand parts of Siberia, China and Tibet (Davis, 2000). Depending on the amount of ice in permafrost and its spatial extent, permafrost can be thick and continuous, discontinuous, sporadic, or isolated (Péwé, 1975; Péwé, 1983).

Permafrost is generally found at high latitudes (lowlands, highlands, intra- or inter-mountane depressions) including shallow continental shelves under the polar seas (Davis, 2000), or high altitudes at mid-latitudes (mountains, highlands, plateaus) (Péwé, 1983; Heginbottom & Dubreuil, 1993; Davis, 2000; Brown et al., 1997). In Siberia, Alaska, and western Canada permafrost occurs in unglaciated regions with ice-rich deposits (ice wedges, ice lenses, and pingo ice cores) or in previously glaciated regions with remnant ice in the ground (Grosse et al., 2013). During the last deglaciation, in the transition from the Pleistocene to the Holocene, temperatures warmed, inducing a form or permafrost degradation that led to thermokarst lake formation in lowland regions with ice-rich unconsolidated

sediment deposits (Zimov et al., 1997; Walter Anthony et al., 2014). Thermokarst lakes are mostly found in Arctic, sub-Arctic, and Boreal regions (Lehner & Doll, 2004; Grosse et al., 2013) and Grosse et al. (2013) estimated a total thermokarst lake and pond area in the pan-Arctic between 250,000 Km^2 and 380,000 Km^2.

In northern permafrost, both localized disturbances to the ground's thermal regime and large climatic changes promote thawing of massive ice wedges in the frozen ground (French, 1976). Competing forces such as climate, hydrology, geology, and short-term ecosystem alterations may affect the thermal stability of permafrost (Jorgenson et al., 2010; Grosse et al., 2013), inducing ice wedge melting. Melted ice-wedges coalescence through superficial ponds and the water that accumulates on top of the permafrost increases water-sediment interface temperatures up to 10 °C above the mean annual air temperatures, due to water properties such as low albedo, absorption of long wave radiation, and a higher heat storage capacity in comparison to ice and dry ground (Walter Anthony et al., 2014). A positive feedback loop is then created in which the warmer temperatures promote even more permafrost degradation and lake bottom expansion, to the point where water column depth exceeds winter ice formation enhancing the growth of the so called thaw bulb (Grosse et al., 2013). This process may be also enhanced by substrates with higher thermal diffusivity like ice-rich mineral substrates, according to a modeling scenario by West and Plug (2008). Lateral expansion of the lakes is aided by thermal and mechanical erosion of the shorelines as described by Grosse *et al*. (2013), with lake sizes ranging between 100 m and 15 Km in diameter, and lake

depths of up to 25 m depending on the permafrost region. For instance, in the Coastal Plain of the North Slope of Alaska there are shallow ice wedges in the permafrost and lakes in this area tend to be shallow (< 3 m deep), but in Siberia and the Northwestern Canadian territories, there are deeper and larger ice wedges and massive glacial ice, and the lakes are much deeper (Grosse et al., 2013).

In the Coastal Plain of the North Slope of Alaska, an area that extends 100 Km to the South from the Arctic Ocean (Morrissey & Livingston, 1992), an estimated 50-70 % of the region is occupied by thermokarst lakes and drained lake basins. Thermokarst lakes in this region present an elliptical shape, and are oriented perpendicular to the prevailing wind direction according to the most widely accepted, but still controversial hypothesis. Wind-induced waves create circulation cells in the ice-free period of the year, causing lateral shoreline erosion 50 ° from the wave approach, deposition of sediments in other shores, and the formation of littoral shelves. Therefore, this hypothesis suggests that the prevailing wind direction would dictate the orientation of the long axis of the lakes and enhance thermokarst erosion at the ends (Carson & Hussey, 1962).

Thermokarst lakes on the North Slope are frozen for ~ 9 months of the year (Phelps et al., 1998) and some lakes freeze all the way to the bottom due to shallow depths. During the summer months, mean daily temperatures above freezing favor lake thaw in late spring, and continuous sunlight reaches the water column over 24 hr periods. During this time of the year CH_4 is released from the sediments into the water column where it may be oxidized to CO_2, and then released to the atmosphere.

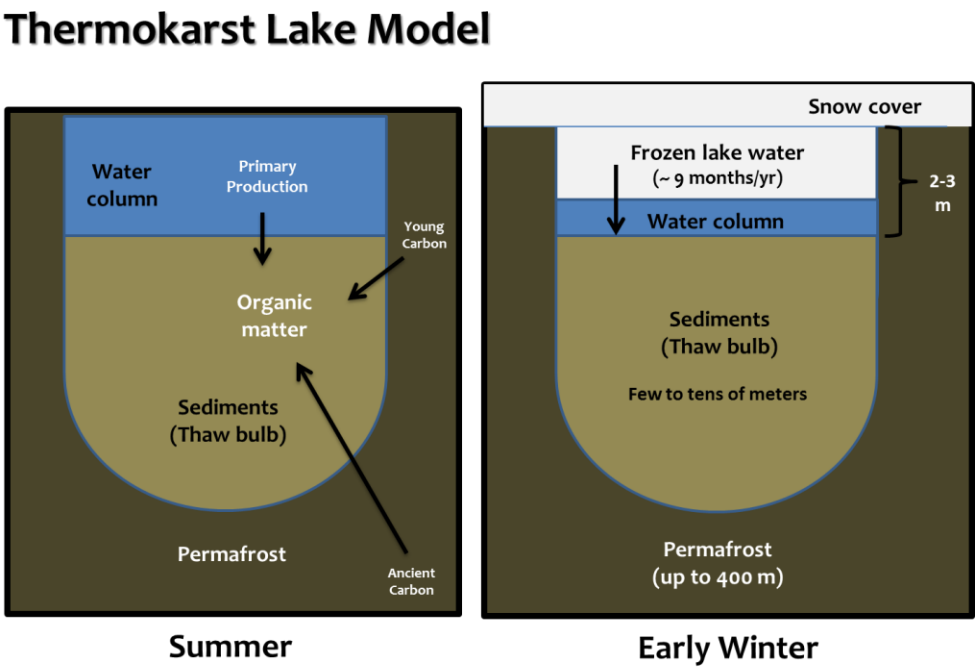


Figure 1. 3 Model of a thermokarst lake with open water in the summer and frozen water in the early winter.

An estimated 0.5 to 11.3 g CH_4 m^{-2} from high-latitude lakes and ponds is contributed to the atmospheric CH_4 budget (Whalen et al., 1990; Bartlett et al., 1992; Roulet et al., 1997)) during the ice-free season, but there is also evidence of a CH_4 pulse during spring thaw from lakes on the North Slope (Phelps, 1998). Thermokarst lakes have played different roles on CH_4 emissions over geologic time. After formation during the Holocene, lakes in Siberia and Alaska may have shifted from carbon sources to sinks {Walter Anthony, 2014), but currently they are considered to be CH_4 sources.

6. Microbial life at the limits: examples from two cold environments

Microorganisms that inhabit permafrost and thermokarst lakes must be able to cope with a variety of environmental stressors such as low temperatures, low water activity, high levels of radiation, and freeze-thaw cycles. However, these arctic ecosystems are just one kind of extreme cold environment where microbial communities thrive; other examples of cold environments are glaciers, ice sheets, marine sediments, and snow. Snowpacks are also of interest for my dissertation research, because they represent a contrasting ecosystem that imposes a different set of stressors on microbial life adapted to the cold.

Methanogenic Archaea are unique in the microbial world because of their ability to produce CH_4. The majority of methanogens in culture were isolated from warm environments like wetlands and rice-paddies; but despite their ability to thrive in warm environments, methanogens have also colonized extremely cold places. Psychrophilic and psychrotrophic methanogens not only survive in these conditions, but are also able to maintain an active metabolism and to reproduce. I dedicated two chapters of my dissertation research to study the activity of methanogens in arctic thermokarst lakes, and their role in the microbial community inhabiting the sediments of these lakes. My ultimate goal was to further the understanding of cold-adapted methanogens and their impact on the greenhouse gas budget of our planet. For the same reason, I attempted to expand the short list of cold-adapted methanogens in culture (Table 1.1; and see appendix), which in the

future would allow us to learn more about their genetic adaptations to different environmental stresses, and what we could look for in other icy bodies of the solar system.

Table 1.1 Some cold-adapted methanogens in culture

Scientific name	Temperature preference	Place of origin	Substrate	Reference
Methanococcoides burtonii sp. nov.	Psychrotolerant > 0 (min), 20 (optimal), 30 °C (max)	Ace Lake, Antarctica	Methanol, methylamines	Franzmann, P. D. *et a*l. 1992. Syst Appl Microbiol 15.
Methanogenium frigidum sp. nov.	**Psychrophilic** < 0 (min), 15 (optimal), 20 °C (max)	**Ace Lake, Antarctica**	**H_2-using**	**Franzmann, P. D. *et al*. 1997. Int J Syst Bacteriol 47.**
Methanogenium marinum sp. nov.	Psychrotolerant > 0 (min), 20 (optimal), 30 °C (max)	Skan Bay Alaska	H_2-using	Chong, S. C., *et al*. 2002. A Van Leeuw J Microb 81.
Methanococcoides alaskense sp. nov.	Psychrotolerant > 0 (min), 20 (optimal), 30 °C (max)	Skan Bay Alaska	Trimethylamines (TMAs)	Singh, N., *et al*. 2005. Int J Syst Evol Micr 55.
Methanogenium boonei sp. nov	Psychrotolerant > 0 (min), 20 (optimal), 30 °C (max)	Skan Bay, Alaska	H_2 or formate	Kendall, M. M. , *et al*. 2007. Appl Environ Microb 73.
Methanobacterium veterum sp. nov	Psychrotolerant > 0 (min), 20 (optimal), 30 °C (max)	Ancient Siberian permafrost	H_2, methylamine, methanol	Krivushin, K. V., *et al*. 2010. Int J Syst Evol Micr 60.

(Franzmann et al., 1997; Franzmann et al., 1992; Chong et al., 2002; Singh et al., 2005; Kendall et al., 2007; Krivushin et al., 2010)

Some environments that are favorable for methanogenesis occur in cold climates (northern wetlands in Siberia, Alaska, Canada, Scandinavia), and are dominated by permafrost, where seasonal variation in ground temperature determines the availability of pore water (Wagner, 2008). However, methanogens can be metabolically active in permafrost (Mikucki et al., 2011) when certain

conditions are met. For instance, the freezing point of water can be depressed during freezing of pure water because of solute exclusion, which allows the presence of thin films of water at temperatures as low as -20 °C (Price, 2000; Bakermans, 2008). Incorporation of radiolabeled acetate at – 20 °C over a 550-day period as well as microbial growth at -10 °C has been detected in permafrost sediments (Rivkina et al., 2000). These findings suggest that microbial life is metabolically active in permafrost at subzero temperatures, and that it is able to continue the cycling of nutrients under this conditions (Mikucki et al., 2011), though at very slow reactions rates.

Of all the environmental factors that can be considered extreme (*i.e.*, departure from optimal conditions), low water activity due to freezing temperatures is perhaps one of the most stressful to microorganisms. To be active at subzero temperatures microbes have to deal with limited amounts of water (the solvent for enzymes), and the resulting decrease in catalysis and transport; the decrease in membrane fluidity, ice crystal formation, and the destabilization of important molecules like DNA and RNA (Cavicchioli et al., 2000; Bakermans, 2008). Some of the adaptations to this type of stress include modification of enzyme kinetics to enable high enzymatic activity at low temperatures; synthesis of glycoproteins and peptides that prevent freezing by dropping the freezing point of water inside the cell; increased membrane fluidity by increasing the degree of unsaturation of fatty acids, modifying branching patterns, and shortening fatty acid chain length; regulation of ion channel permeability; synthesis of cold-shock proteins (Csps) that

stabilize the mRNA to allow protein synthesis, and synthesis of cold acclimation proteins (Caps) that are synthesized during prolonged exposure to cold temperatures (Pikuta et al., 2007).

Methanogens are just one example of the kind of microorganisms that are adapted to live in cold environments, in conditions of freezing temperatures and freeze-thaw cycles, low water activity, high or low levels of salinity or pH, high levels of UVB radiation during the summer, and complete darkness during the winter (Hoover et al., 2002; Pikuta et al., 2007). In snow-covered regions in the Arctic, the Antarctic, and in high altitude environments, microorganisms belonging to several different phyla from the three domains of life must also adapt to conditions of low water activity due to low temperature. One of the most studied eukaryotic microorganisms in snow is the unicellular alga *Chlamydomonas nivalis*, which produces pigmented red cells (cysts and zygotes) giving this coloration to the snow. *C. nivalis* uses these pigments (flavonoids and astaxanthins) as antioxidants, an adaptation strategy that makes it resistant to high levels of UVA and UVC radiation (Hoover et al., 2002; Pikuta et al., 2007). As a consequence, the red biomass in snow changes the albedo, which induces snow melt and allows other microorganisms to thrive because of increased water availability (Pikuta et al., 2007). This 'red snow' commonly found in the Sierra Nevada of California is a signature of life occupying a contrasting habitat of interest for my dissertation.

The very high radiation levels the algae experience at high altitudes decrease with snow depth, however, snowpack temperatures at depth can be well below the

constant 0-1 °C characteristic of wet snow (Remias, 2012). Wet snow mostly occurs during the spring melt, when the snowfall from the previous winter starts to thaw, but persist in a water-logged state until the summer, favoring snow algae development (Remias, 2012). Unicellular algae like *C. nivalis* require higher amounts of liquid water than bacteria in permafrost because of the larger cell sizes (10-30 μm), and old summer snow allows the formation of a continuous water film around macroscopic snow crystals that is the ideal microhabitat for the algae (Jones, 2001; Remias, 2012).

The life cycle of *C. nivalis* as described by Remias (2012) is tied to annual environmental changes in snowpack, reflecting adaptation to the harsh conditions of this cold environment. In the cyst phase, cells are dormant, *i.e.,* non-motile, stop reproducing, preserve energy and nutrients, rearrange their chloroplast, produce lipids to decrease the water content in the cell, and accumulate UV-protecting pigments. All of these changes allow the algae to withstand conditions of high radiation, low nutrient availability, and freeze-thaw cycles during the summer. Some discrepancies remain in the understanding of cyst physiology, as Remias (2005) determined that mature cysts of *C. nivalis* from the Alps were photosynthetically active and produced detectable levels of oxygen, indicating at least low levels of metabolism. In the spring of the next season, cells germinate (reproduce asexually) and release four-to-eight daughter cells into the snowpack that has accumulated over the winter. In the early summer, the flagellated algal cells that emerge from cyst forms can propel upwards seeking adequate light conditions for sexual

reproduction and growth. At this stage, these cells are most sensitive to environmental conditions; therefore, they remain deep enough in the snowpack to be protected from high irradiation and freeze-thaw cycles (Remias, 2012). When temperatures rise, cellular growth and metabolism ensues, and the nutrients can be exhausted. Flagellates then turn into vegetative cysts or generative zygotes, loosing motility and the ability to reproduce. In this state, the algae become more resistant to the harsher environmental conditions associated to the snowmelt period (Remias, 2012). Likewise, concurrent irradiation levels on the surface of the snowpack become increasingly high stimulating carotenoid production and chloroplast size reduction so that the cells finally enter the dormant state that lasts until next season (Remias et al., 2005).

During their flagellated phase, snow algae are photosynthesizing, which makes them primary producers and CO_2 sinks under favorable conditions (Painter et al., 2001; Williams et al., 2009). Additionally, they create conditions that promote the growth of heterotrophic bacteria. These two types of metabolism (photoautotrophic and chemoheterotrophic) are different from chemolithoautotrophic metabolism common to several species of cold-adapted methanogens, which is less favorable in cold environments (Hoover et al.; Pikuta et al., 2007). Nevertheless, the microbial community associated with snow algae has only been described in a few studies (Segawa et al., 2005; Liu et al., 2009; Zhang et al., 2010). Much less attention has been paid to these communities in comparison to microbial communities in permafrost (Zinger et al., 2009; Larose et al., 2010;

Harding et al., 2011; Hamilton et al., 2013; Moller et al., 2013)) or other icy environments including glacial ice, subglacial systems, lake ice, sea ice, and ice-sheet environments (Mikucki et al., 2011).

An interesting characteristic of algae in snow is that they constitute a detectable biosignature due to carotenoid pigmentation. Painter (2001) have used imaging spectroscopy to analyze algal spectral features enabling detection by remote-sensing which can be calibrated to determine algal concentrations. Unlike methanogens, snow algae can be detected directly using spectroscopy-based techniques. While spectroscopy alone would not tell us whether methanogens are present, a combination of spectroscopic and spectrometric techniques can detect the metabolic byproducts formed by methanogens. However, the likelihood of encountering evolutionarily advanced photosynthetic microorganisms elsewhere in the Solar System is questionable, whereas the presence of a primitive form of metabolism like methanogenesis is much more likely, particularly in Mars.

7. Overall objectives

My dissertation aimed to study microbial life in two different kinds of cold environments with astrobiological relevance: the sediments of thermokarst lakes in permafrost at high-latitudes, and snow fields at high-altitude environments. Understanding how life thrives in these environments will aid our efforts to search for life in the Solar System, especially in places like the permafrost regions in Mars and the icy surface of Europa. The microbial communities inhabiting these cold

environments are also very relevant for the global cycling of C, and with the current trend of climate change, we must fully understand their role in nutrients cycling and how they may respond to change.

With these goals in mind my dissertation aimed to:

1. Determine the sources of CH_4 in sediments of arctic thermokarst lakes in the North Slope of Alaska, and what geochemical factors favor biological CH_4 production from these lakes.
2. Describe the diversity and structure of the microbial community that supports organic matter decomposition leading to methanogenesis, and to analyze how organic carbon affects microbial community composition.
3. Uncover the microbial communities associated with snow algae in the snow fields of Mt. Conness and on the Pacific Crest Trail near Mt. Lincoln in the Sierra Nevada of California, and on Wheeler Peak in the Great Basin National Park in Nevada.

References

Acrc (2008) The Alaska Climate Research Center.

Atreya SK, Mahaffy PR, Wong AS (2007) Methane and related trace species on Mars: Origin, loss, implications for life, and habitability. *Planetary and Space Science,* **55**, 358-369.

Bakermans C (2008) *Limits for microbial life at subzero temperatures*. In: *Psychrophiles: From biodiversity to biotechnology* (eds Margesin R, Schinner F, Marx J-C, Gerday C). Springer Berlin Heidelberg, pp. 17-28.

Bartlett KB, Crill PM, Sass RL, Harriss RC, Dise NB (1992) Methane emissions from tundra environments in the Yukon-Kuskokwim delta, Alaska. *Journal of Geophysical Research: Atmospheres,* **97**, 16645-16660.

Battistuzzi FU, Feijao A, Hedges SB (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evolutionary Biology,* **4**, 1-14.

Boone DR (2000) Biological formation and consumption of methane. In: *Atmospheric Methane* (ed Khalil MAK). Springer, New York, pp. 42-62.

Brewer MC (1958) Some results of geothermal investigations of permafrost in northern Alaska. *Eos, Transactions American Geophysical Union,* **39**, 19-26.

Brown J, Ferrians OJ, Heginbottom J, Melnikov E (1997) *Circum-Arctic map of permafrost and ground-ice conditions,* US Geological Survey Reston.

Carson CE, Hussey KM (1962) The oriented lakes of Arctic Alaska. *The Journal of Geology,* **70**, 417-439.

Cavicchioli R, Thomas T, Curmi PMG (2000) Cold stress response in Archaea. *Extremophiles,* **4**, 321-331.

Chong S, Liu Y, Cummins M, Valentine D, Boone D (2002) Methanogenium marinum sp. nov., a H_2-using methanogen from Skan Bay, Alaska, and kinetics of H2 utilization. *Antonie Van Leeuwenhoek,* **81**, 263-270.

Davis TN (2000) *Permafrost : a guide to frozen ground in transition,* University of Alaska Press, Fairbanks.

Denman KL, Brasseur G, Chidthaisong A, Ciais P, Cox PM, Dickinson RE, Hauglustaine D, Heinze C, Holland E, Jacob D (2007) Couplings between changes in the climate system and biogeochemistry. In: *Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change* (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL). Cambridge University Press, Cambridge pp. 541-584.

Des Marais DJ, Nuth JaR, Allamandola LJ, Boss AP, Farmer JD, Hoehler TM, Jakosky BM, Meadows VS, Pohorille A, Runnegar B, Spormann AM (2008) The NASA astrobiology roadmap. *Astrobiology,* **8**, 715-730.

Etiope G (2012) Climate science: Methane uncovered. *Nature Geoscience,* **5**, 373-374.

Etiope G, Lassey KR, Klusman RW, Boschi E (2008) Reappraisal of the fossil methane budget and related emission from geologic sources. *Geophysical Research Letters,* **35**, 1-5.

Ferry JG (2010) The chemical biology of methanogenesis. *Planetary and Space Science,* **58**, 1775-1783.

Ferry JG, Kastead KA (2007) Methanogenesis. In: *Archaea* (ed Cavicchioli R). American Society for Microbiology, Washington DC.

Franzmann PD, Liu Y, Balkwill DL, Aldrich HC, Conway De Macario E, Boone DR (1997) Methanogenium frigidum sp. nov., a psychrophilic, H_2-using methanogen from Ace Lake, Antarctica. *International Journal of Systematic Bacteriology,* **47**, 1068-1072.

Franzmann PD, Springer N, Ludwig W, Demacario EC, Rohde M (1992) A methanogenic archaeon from ace lake, Antarctica Methanococcoides burtonii sp. nov. *Systematic and Applied Microbiology,* **15**, 573-581.

French HM (1976) *The periglacial environment,* Longman, Inc., New York.

Graham DE, Wallenstein MD, Vishnivetskaya TA, Waldrop MP, Phelps TJ, Pfiffner SM, Onstott TC, Whyte LG, Rivkina EM, Gilichinsky DA, Elias DA, Mackelprang R, Verberkmoes NC, Hettich RL, Wagner D, Wullschleger SD, Jansson JK (2012) Microbes in thawing permafrost: the unknown variable in the climate change equation. *The ISME Journal,* **6**, 709-712.

Grosse G, Jones B, Arp C (2013) Thermokarst lakes, drainage, and drained basins. In: *Treatise on Geomorphology* (eds Shroeder F, Giardino R, Harbor J). Academic Press, San Diego, pp. 325-353.

Hamilton TL, Peters JW, Skidmore ML, Boyd ES (2013) Molecular evidence for an active endogenous microbiome beneath glacial ice. *The ISME Journal,* **7**, 1402-1412.

Hand KP, Mckay CP, Pilcher CB (2010) Spectroscopic and spectrometric differentiation between abiotic and biogenic material on icy worlds. *Proceedings of the International Astronomical Union,* **6**, 165-176

Harding T, Jungblut AD, Lovejoy C, Vincent WF (2011) Microbes in High Arctic snow and implications for the cold biosphere. *Applied and Environmental Microbiology,* **77**, 3234-3243.

Heginbottom J, Dubreuil M (1993) A new permafrost and ground ice map for the national atlas of Canada. In: *Proceedings of the Sixth International Conference on Permafrost,* pp. 255-260.

Hoover RB, Pikuta EV, Marsic D, Ng J (2002) Anaerobic psychrophiles from Alaska, Antarctica, and Patagonia: implications to possible life on Mars and Europa. In: *International Symposium on Optical Science and Technology*. International Society for Optics and Photonics, pp. 313-324.

IPCC (2001) Climate change 2001: The scientific basis. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change. (eds Houghton J, Ding Y, Griggs D, Noguer M, Van Der Linden P, Dai X, Maskell K, Johnson CA), Cambridge.

IPCC (2007) Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H). Cambridge University Press, Cambridge; New York.

Jones HG (2001) *Snow ecology : an interdisciplinary examination of snow-covered ecosystems,* Cambridge University Press, Cambridge.

Jorgenson MT, Romanovsky V, Harden J, Shur Y, O'donnell J, Schuur EAG, Kanevskiy M, Marchenko S (2010) Resilience and vulnerability of permafrost to climate change. *Canadian Journal of Forest Research,* **40**, 1219-1236.

Kasting JF, Ono S (2006) Palaeoclimates: the first two billion years. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **361**, 917-929.

Kendall MM, Wardlaw GD, Tang CF, Bonin AS, Liu Y, Valentine DL (2007) Diversity of Archaea in marine sediments from Skan Bay, Alaska, including cultivated methanogens, and

description of Methanogenium boonei sp nov. *Applied and Environmental Microbiology,* **73**, 407-414.

Kirschke S, Bousquet P, Ciais P, Saunois M, Canadell JG, Dlugokencky EJ, Bergamaschi P, Bergmann D, Blake DR, Bruhwiler L, Cameron-Smith P, Castaldi S, Chevallier F, Feng L, Fraser A, Heimann M, Hodson EL, Houweling S, Josse B, Fraser PJ, Krummel PB, Lamarque J-F, Langenfelds RL, Le Quere C, Naik V, O'doherty S, Palmer PI, Pison I, Plummer D, Poulter B, Prinn RG, Rigby M, Ringeval B, Santini M, Schmidt M, Shindell DT, Simpson IJ, Spahni R, Steele LP, Strode SA, Sudo K, Szopa S, Van Der Werf GR, Voulgarakis A, Van Weele M, Weiss RF, Williams JE, Zeng G (2013) Three decades of global methane sources and sinks. *Nature Geoscience,* **6**, 813-823.

Kittel TGF, Baker BB, Higgins JV, Haney JC (2011) Climate vulnerability of ecosystems and landscapes on Alaska's North Slope. *Regional Environmental Change,* **11**, S249-S264.

Krivushin KV, Shcherbakova VA, Petrovskaya LE, Rivkina EM (2010) Methanobacterium veterum sp nov., from ancient Siberian permafrost. *International Journal of Systematic and Evolutionary Microbiology,* **60**, 455-459.

Kvenvolden KA, Rogers BW (2005) Gaia's breath—global methane exhalations. *Marine and Petroleum Geology,* **22**, 579-590.

Larose C, Berger S, Ferrari C, Navarro E, Dommergue A, Schneider D, Vogel TM (2010) Microbial sequences retrieved from environmental samples from seasonal arctic snow and meltwater from Svalbard, Norway. *Extremophiles,* **14**, 205-212.

Lehner B, Doll P (2004) Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology,* **296**, 1-22.

Liu Y, Yao T, Jiao N, Kang S, Xu B, Zeng Y, Huang S, Liu X (2009) Bacterial diversity in the snow over Tibetan Plateau Glaciers. *Extremophiles,* **13**, 411-423.

Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA (2014) *Brock biology of microorganisms,* Benjamin-Cummings Publishing Company, Boston.

Mahaffy PR, Webster CR, Cabane M, Conrad PG, Coll P, Atreya SK, Arvey R, Barciniak M, Benna M, Bleacher L, Brinckerhoff WB, Eigenbrode JL, Carignan D, Cascia M, Chalmers RA, Dworkin JP, Errigo T, Everson P, Franz H, Farley R, Feng S, Frazier G, Freissinet C, Glavin DP, Harpold DN, Hawk D, Holmes V, Johnson CS, Jones A, Jordan P, Kellogg J, Lewis J, Lyness E, Malespin CA, Martin DK, Maurer J, Mcadam AC, Mclennan D, Nolan TJ, Noriega M, Pavlov AA, Prats B, Raaen E, Sheinman O, Sheppard D, Smith J, Stern JC, Tan F, Trainer M, Ming DW, Morris RV, Jones J, Gundersen C, Steele A, Wray J, Botta O, Leshin LA, Owen T, Battel S, Jakosky BM, Manning H, Squyres S, Navarro-Gonzalez R, Mckay CP, Raulin F, Sternberg R, Buch A, Sorensen P, Kline-Schoder R, Coscia D, Szopa C, Teinturier S, Baffes C, Feldman J, Flesch G, Forouhar S, Garcia R, Keymeulen D, Woodward S, Block BP, Arnett K, Miller R, Edmonson C, Gorevan S, Mumm E (2012) The sample analysis at Mars investigation and instrument suite. *Space Science Reviews,* **170**, 401-478.

Mcguire AD, Anderson LG, Christensen TR, Dallimore S, Guo L, Hayes DJ, Heimann M, Lorenson TD, Macdonald RW, Roulet N (2009) Sensitivity of the carbon cycle in the Arctic to climate change. *Ecological Monographs,* **79**, 523-555.

Mckay CP (2011) The search for life in our Solar System and the implications for science and society. *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences,* **369**, 594-606.

Mikucki J, Han SK, Lanoil B (2011) Ecology of psychrophiles: Subglacial and permafrost environments. In: *Extremophiles Handbook* (ed Horikoshi K). Springer Japan, pp. 755-775.

Moller AK, Soborg DA, Abu Al-Soud W, Sorensen SJ, Kroer N (2013) Bacterial community structure in High-Arctic snow and freshwater as revealed by pyrosequencing of 16S rRNA genes and cultivation. *Polar Research,* **32**, 1-11.

Morrissey LA, Livingston GP (1992) Methane emissions from Alaska Arctic tundra: An assessment of local spatial variability. *Journal of Geophysical Research: Atmospheres,* **97**, 16661-16670.

Osterkamp TE (2007) Characteristics of the recent warming of permafrost in Alaska. *Journal of Geophysical Research-Earth Surface,* **112**, 1-10.

Painter TH, Duval B, Thomas WH, Mendez M, Heintzelman S, Dozier J (2001) Detection and quantification of snow algae with an airborne imaging spectrometer. *Applied and Environmental Microbiology,* **67**, 5267-5272.

Péwé TL (1975) Quaternary geology of Alaska. In: *U.S. Geological Survey Professional Paper*.

Péwé TL (1982) *Geologic hazards of the Fairbanks area, Alaska,* Alaska Division of Geological & Geophysical Surveys, Fairbanks.

Péwé TL (1983) Alpine Permafrost in the Contiguous United States: A Review. *Arctic and Alpine Research,* **15**, 145-156.

Phelps AR, Peterson KM, Jeffries MO (1998) Methane efflux from high-latitude lakes during spring ice melt. *Journal of Geophysical Research,* **103**, 29029-29036.

Pikuta EV, Hoover RB, Tang J (2007) Microbial extremophiles at the limits of life. *Critical Reviews in Microbiology,* **33**, 183-209.

Pilcher CB (2003) Biosignatures of early earths. *Astrobiology,* **3**, 471-486.

Price PB (2000) A habitat for psychrophiles in deep Antarctic ice. *Proceedings of the National Academy of Sciences,* **97**, 1247-1251.

Remias D (2012) Cell structure and physiology of alpine snow and ice algae. In: *Plants in Alpine Regions* (ed Lütz C). Springer Vienna, pp. 175-185.

Remias D, Lutz-Meindl U, Lutz C (2005) Photosynthesis, pigments and ultrastructure of the alpine snow alga Chlamydomonas nivalis. *European Journal of Phycology,* **40**, 259-268.

Rivkina EM, Friedmann EI, Mckay CP, Gilichinsky DA (2000) Metabolic activity of permafrost bacteria below the freezing point. *Applied and Environmental Microbiology,* **66**, 3230-3233.

Roulet NT, Crill PM, Comer NT, Dove A, Boubonniere RA (1997) CO_2 and CH_4 flux between a boreal beaver pond and the atmosphere. *Journal of Geophysical Research-Atmospheres,* **102**, 29313-29319.

Rouse WR, Douglas MSV, Hecky RE, Hershey AE, Kling GW, Lesack L, Marsh P, Mcdonald M, Nicholson BJ, Roulet NT, Smol JP (1997) Effects of climate change on the freshwaters of Arctic and Subarctic North America. *Hydrological Processes,* **11**, 873-902.

Schrenk MO, Brazelton WJ, Lang SQ (2013) Serpentinization, carbon, and deep Life. *Reviews in Mineralogy & Geochemistry,* **75**, 575-606.

Seager S (2013) Exoplanet Habitability. *Science,* **340**, 577-581.

Seager S (2014) The future of spectroscopic life detection on exoplanets. *Proceedings of the National Academy of Sciences,* **111**, 12634-12640.

Seager S, Schrenk M, Bains W (2012) An astrophysical view of earth-based metabolic biosignature gases. *Astrobiology,* **12**, 61-82.

Segawa T, Miyamoto K, Ushida K, Agata K, Okada N, Kohshima S (2005) Seasonal change in bacterial flora and biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time PCR. *Applied and Environmental Microbiology,* **71**, 123-130.

Segura A, Kasting JF, Meadows V, Cohen M, Scalo J, Crisp D, Butler R, Tinetti G (2005) Biosignatures from earth-like planets around M dwarfs. *Astrobiology,* **5**, 706-725.

Segura A, Meadows VS, Kasting JF, Crisp D, Cohen M (2007) Abiotic formation of O_2 and O_3 in high-CO_2 terrestrial atmospheres. *Astronomy & Astrophysics,* **472**, 665-679.

Sepulveda-Jauregui A, Walter Anthony KM, Martinez-Cruz K, Greene S, Thalasso F (2014) Methane and carbon dioxide emissions from 40 lakes along a north–south latitudinal transect in Alaska. *Biogeosciences Discussions,* **11**, 13251-13307.

Serreze MC, Walsh JE, Chapin FS, Osterkamp T, Dyurgerov M, Romanovsky V, Oechel WC, Morison J, Zhang T, Barry RG (2000) Observational evidence of recent change in the northern high-latitude environment. *Climatic Change,* **46**, 159-207.

Singh N, Kendall MM, Liu YT, Boone DR (2005) Isolation and characterization of methylotrophic methanogens from anoxic marine sediments in Skan Bay, Alaska: description of Methanococcoides alaskense sp nov., and emended description of Methanosarcina baltica. *International Journal of Systematic and Evolutionary Microbiology,* **55**, 2531-2538.

Stevens CM, Wahlen M (2000) The isotopic composition of atmospheric methane and its sources. *Atmospheric Methane*, 25-41.

Tarnocai C, Canadell JG, Schuur EA, Kuhry P, Mazhitova G, Zimov S (2009) Soil organic carbon pools in the northern circumpolar permafrost region. *Global Biogeochemical Cycles,* **23**, 1-11.

Ueno Y, Yamada K, Yoshida N, Maruyama S, Isozaki Y (2006) Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature,* **440**, 516-519.

Wagner D (2008) Microbial communities and processes in arctic permafrost environments. In: *Microbiology of Extreme Soils* (eds Dion P, Nautiyal C). Springer Berlin Heidelberg, pp. 133-154.

Walter Anthony KM, Zimov SA, Grosse G, Jones MC, Anthony PM, Chapin FS III, Finlay JC, Mack MC, Davydov S, Frenzel P, Frolking S (2014) A shift of thermokarst lakes from carbon sources to sinks during the Holocene epoch. *Nature,* **511**, 452-456.

Walter KM, Zimov SA, Chanton JP, Verbyla D, Chapin FS III (2006) Methane bubbling from Siberian thaw lakes as a positive feedback to climate warming. *Nature,* **443**, 71-75.

Washburn AL (1980) *Geocryology: a survey of periglacial processes and environments,* Wiley, New York.

Webster CR, Mahaffy PR, Atreya SK, Flesch GJ, Mischna MA, Meslin PY, Farley KA, Conrad PG, Christensen LE, Pavlov AA, Martin-Torres J, Zorzano MP, Mcconnochie TH, Owen T, Eigenbrode JL, Glavin DP, Steele A, Malespin CA, Archer PD, Sutter B, Coll P, Freissinet C, Mckay CP, Moores JE, Schwenzer SP, Bridges JC, Navarro-Gonzalez R, Gellert R, Lemmon MT, Team MSLS (2015) Mars methane detection and variability at Gale crater. *Science,* **347**, 415-417.

West JJ, Plug LJ (2008) Time-dependent morphology of thaw lakes and taliks in deep and shallow ground ice. *Journal of Geophysical Research: Earth Surface,* **113**, F01009.

Whalen SC, Reeburgh WS, Sandbeck KA (1990) Rapid methane oxidation in a landfill cover soil. *Applied and Environmental Microbiology,* **56**, 3405-3411.

Whiticar M (2000) Can stable isotopes and global budgets be used to constrain atmospheric methane budgets? In: *Atmospheric Methane* (ed Khalil MAK). Springer Berlin Heidelberg, pp. 63-85.

Whiticar MJ (1999) Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chemical Geology,* **161**, 291-314.

Whitman W, Bowen T, Boone D (2006) The methanogenic bacteria. In: *The Prokaryotes* (eds Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E). Springer New York, pp. 165-207.

Williams MW, Seibold C, Chowanski K (2009) Storage and release of solutes from a subalpine seasonal snowpack: soil and stream water response, Niwot Ridge, Colorado. *Biogeochemistry,* **95**, 77-94.

Zhang S, Yang G, Wang Y, Hou S (2010) Abundance and community of snow bacteria from three glaciers in the Tibetan Plateau. *Journal of Environmental Sciences-China,* **22**, 1418-1424.

Zhang Y, Chen WJ, Cihlar J (2003) A process-based model for quantifying the impact of climate change on permafrost thermal regimes. *Journal of Geophysical Research,* **108**, 1-16.

Zhuang Q, Melillo JM, Sarofim MC, Kicklighter DW, Mcguire AD, Felzer BS, Sokolov A, Prinn RG, Steudler PA, Hu S (2006) CO_2 and CH_4 exchanges between land ecosystems and the atmosphere in northern high latitudes over the 21st century. *Geophysical Research Letters,* **33**, 1-5.

Zimov SA, Voropaev YV, Semiletov IP, Davidov SP, Prosiannikov SF, Chapin FS III, Chapin MC, Trumbore S, Tyler S (1997) North Siberian lakes: A methane source fueled by Pleistocene carbon. *Science,* **277**, 800-802.

Zinder SH (1984) Microbiology of anaerobic conversion of organic wastes to methane: recent developments. *ASM News,* **50**, 294-298.

Zinger L, Shahnavaz B, Baptist F, Geremia RA, Choler P (2009) Microbial diversity in alpine tundra soils correlates with snow cover dynamics. *The ISME Journal,* **3**, 850-859.

Chapter 2

Methane sources in arctic thermokarst lake sediments on the North Slope of Alaska

Co-authorship statement

P. Matheus Carnevali, M. Rohrssen, M. R. Williams , A. B. Michaud, H. Adams , D. Berisford, G. D. Love, J. C. Priscu, O. Rassuchine, K. P. Hand, A. E. Murray.

PM, MR and AEM designed the sampling approach and the experiments. PM, MR, AEM, DB, HA, JP, AM, KH collected samples for all of the analyses. PM executed the CH_4 production experiment, qPCR, and analyzed the data generated from CH_4 production experiment, CH_4 concentration, $\delta^{13}C_{CH4}$ and $\delta^{13}C_{CO2}$, pore water chemistry of anions and metals, and qPCR under the supervision of AEM. OR helped with CH_4 concentration and $\delta^{13}C$ analyses, and with DNA extractions. MR and MW performed analyses for TOC and all of the lipid biomarkers under the supervision of GL. AM and JP performed O_2 analyses. HA and JP performed limnological analyses. PM, MR and AEM wrote this chapter in collaboration with GL, JP, AM, KH, DB and HA. AEM, JP, GL and KH secured the funding for this research.

The contents of this chapter were published in:

Geobiology **Volume 13, Issue 2, March 2015, Pages: 181-197**

Abstract

The permafrost on the North Slope of Alaska is densely populated by shallow lakes that result from thermokarst erosion. These lakes release methane (CH_4) derived from a combination of ancient thermogenic pools and contemporary biogenic production. Despite the potential importance of CH_4 as a greenhouse gas, the contribution of biogenic CH_4 production in arctic thermokarst lakes in Alaska is not currently well understood. To further advance our knowledge of CH_4 dynamics in these lakes, we focused our study on (i) the potential for microbial CH_4 production in lake sediments, (ii) the role of sediment geochemistry in controlling biogenic CH_4 production, and (iii) the temperature dependence of this process. Sediment cores were collected from one site in Siqlukaq Lake and two sites in Sukok Lake in late October to early November. Analyses of pore water geochemistry, sedimentary organic matter and lipid biomarkers, stable carbon isotopes, results from CH_4 production experiments, and copy number of a methanogenic pathway-specific gene (*mcrA*) indicated the existence of different sources of CH_4 in each of the lakes chosen for the study. Analysis of this integrated data set revealed that there is biological CH_4 production in Siqlukaq at moderate levels, while the very low levels of CH_4 detected in Sukok had a mixed origin, with little to no biological CH_4 production. Furthermore, methanogenic Archaea exhibited temperature dependent use of *in situ* substrates for methanogenesis, and the amount of CH_4 produced was directly related to the amount of labile organic matter in the sediments. This study constitutes an

important first step in better understanding the actual contribution of biogenic CH_4 from thermokarst lakes on the coastal plain of Alaska to the current CH_4 budgets.

1. Introduction

Thermokarst lakes, resulting from ground ice melting in permafrost regions (Howard & Prescott, 1973; French, 1976; Kokelj & Jorgenson, 2013) may be significant contributors to the global CH_4 budget (Oechel et al., 1993; Phelps et al., 1998; Walter et al., 2007; Schuur et al., 2009). A large proportion of this CH_4 may be derived from ancient thermogenic CH_4 trapped deep within or under the permafrost (referred to as the "cryosphere cap", Walter-Anthony *et al.*, 2012). However, the large amount of organic matter stored in the thaw layer (talik) between the water column and the permafrost table (Ling & Zhang, 2003; Pedersen et al., 2011; Parsekian et al., 2013), either from interglacial or contemporary photosynthesis, also serves as a significant source of carbon for *in situ* methanogenesis. Identifying and deconvolving the production sources of CH_4 will improve our ability to generate accurate predictions about the changing climate in the Arctic.

The North Slope of Alaska has extensive continuous permafrost (~ 60-75 % ice by volume and ~ 400 m deep; Sellmann et al., 1975; Hinkel et al., 2003; Jorgenson et al., 2008), and is occupied by thousands of shallow (~ 2-3 m deep), relic (*i.e.,* drained), and contemporary thermokarst lakes (Hussey & Michelson, 1966; Frohn et al., 2005; Jorgenson & Shur, 2007) that are ice-covered or frozen to the ground for at least 9 months of the year. Northern Alaska lake sediments may be

gradually eroded through the lake thaw cycle and preferentially re-deposited at the upwind and downwind lake margins (Carson & Hussey, 1962; Hinkel et al., 2003). Relative contributions from allochthonous or autochthonous sources of organic matter have not been well studied in the sediments of this area. Allochthonous organic matter can be transported via fluvial or aeolian processes, and derive from modern active layer soils or Pleistocene-aged terrigenous organic matter from permafrost (Repenning, 1983). Autochthonous organic matter can be produced by present-day lacustrine autotrophs (Ramlal et al., 1994; Hecky & Hesslein, 1995; Bonilla et al., 2005) and is potentially more labile than allochthonous organic matter.

The coastal plain in the North Slope of Alaska also contains an estimated 53 billion cubic feet of natural gas (Houseknecht et al., 2010). Radiocarbon analyses have indicated that gas seeps in the area may be sourced from the gas reservoirs at depth and/or laterally from thermogenic CH_4 trapped under the permafrost ice cap, rather than from present-day microbial activity within the lake, as is common in Siberia (Walter Anthony et al., 2012). A distinction between microbial and thermogenic CH_4 can be made by combining isotopic ratios (*e.g.*, $\delta^{13}C_{CH4}$, $\delta^{13}C_{CO2}$ and $\delta D\text{-}CH_4$) and C2-C4 hydrocarbon ratios (Whiticar, 1999). Microbial CH_4 production, which $\delta^{13}C$ ranges between -110 and -50 $^0/_{00}$ (Quay et al., 1988; Whiticar, 1999), results from anaerobic decomposition of organic matter in sediments. Thermogenic CH_4 has a range between -52 and -20 $^0/_{00}$ (Whiticar, 1999; Judd, 2000; Kvenvolden & Rogers, 2005) and it is generated at subsurface depths between 1 and 4 Km by

decomposition of residual organic matter under high pressure and temperature, during coal formation or thermal alteration of oil (Judd, 2000). $\delta^{13}C_{CH4}$ signatures at the boundary between biogenic and thermogenic CH_4 could result from mixed sources, including CH_4 oxidation, advanced stage of parent organic matter decomposition, contributions from different methanogenic pathways, or a combination of thermogenic and biogenic signatures (Whiticar, 1999). Lastly, abiogenic CH_4 originates in the mantle, and it has a $\delta^{13}C$ between -45 and -5 $^{0}/_{00}$ (Judd, 2000).

Our study focused on biological CH_4 production in two Alaskan thermokarst lakes. Specifically we examined: i) *in situ* CH_4 concentrations and carbon isotope compositions of CH_4 in sediments; ii) temperature response of methanogenesis at natural substrate levels; iii) archaeol lipid biomarkers (archaeol) and the methyl coenzyme reductase alpha subunit (*mcrA*) gene, which is a key enzyme in the pathway for methanogenesis; and iv) description of the substrates available for methanogenesis.

2. Methods

2.1. Sampling sites

Siqlukaq Lake (Siq) and Sukok Lake (Suk), two arctic thermokarst lakes near the town of Barrow, Alaska (Fig. 2.1), were sampled on late October to early November field campaigns. Two sites were sampled at Sukok, one near an active, submerged natural gas seep (Sukok Seep - SukS), and another about 1 km southwest from the

seep site (Sukok B site - SukB), to determine the effects of localized CH_4 flux on biological CH_4 production within these sediments. A total of 16 sediment cores were recovered for various geochemical and biological analyses over the course of four field campaigns spanning 4 years (Table 1). Sediments in both lakes lack well-defined sedimentological features, such as laminations, and the lakes possess taliks of at least 1.1 m depth (the maximum sediment thickness penetrated in coring).

Suk is located ~ 29 km south of Barrow and 12.7 km east-southeast of the mouth of Walakpa Bay, in the Walakpa gas field, a natural gas field approximately 600 m deep (the permafrost base nearby the lake is ~ 280 m; Glenn & Allen, 1992). An east-southeast trending fault occurs in the subsurface north of Sukok, however no such feature is identified beneath the lake itself (Glenn & Allen, 1992). Openings in autumn ice cover resulting from active CH4 ebullition in the lake were observed in April 2010 and the late October-early November 2010-2013 field campaigns. Satellite imagery indicates that Suk consists of at least three coalesced thermokarst lakes and lies within a portion of the arctic coastal plain that has seen repeated thermokarst episodes (Fig. 2.1). Suk is approximately 4.2 km wide and 5.5 km long (not including the slightly-adjoined southern basin). At the time of sampling, total water depth for Suk was ~ 0.80-1.35 m, with ~ 0.10-0.25 m of ice, and lake water temperatures were relatively uniform with depth between 0.7 and 1.3 °C, as determined with a portable Orion 5 star multi meter (Thermo Scientific, Waltham, MA, USA).

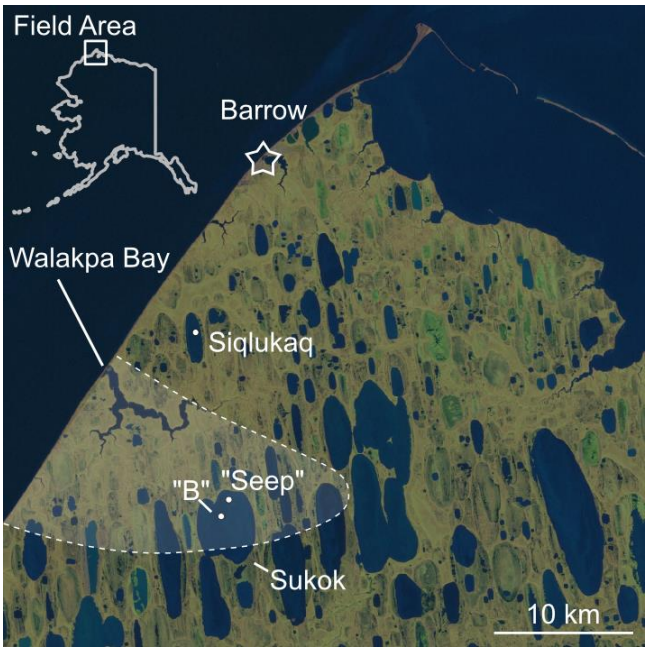


Figure 2.1 Landsat 7 image of the Arctic Coastal Plane near Barrow, AK (L7 ETM+ SLC-on, 31 August 2000). Shaded field and hatched lines approximate subsurface boundaries of the Walakpa Gas Field (Glenn & Allen, 1992). Dots indicate sites described in the main text: Siqlukaq (Siq), and the two Sukok Lake sites are "Seep" proximal to an active, ebullient gas seep (SukS), and "B" distal from the area of active CH_4 seepage (SukB).

Siq, located 6.6 km east-northeast of the mouth of Walakpa Bay, outside of the gas field, has no visible open holes in winter lake ice due to gas ebullition. Satellite photography shows that Siq likely drains into Walakpa Bay at high-stand and shows evidence of previously higher lake levels, but a less complex hydrologic history than that of Suk (Fig. 2.1). Siq is smaller and more elongated than Suk, measuring approximately 1.0 km wide by 3.8 km long. During the sampling period, Siq was ~ 1.5-1.6 m deep, had ~ 0.15-0.25 m of ice, and lake water temperatures between 0.2 and 2.0 °C.

2.2. Sediment geochemistry

Sediment cores were retrieved from all sampling sites using a universal percussion corer (Aquatic Research Instruments), ~ 10 cm in diameter polycarbonate coring tube, and plug caps on both ends to decrease oxidation. Sediment core lengths ranged between 20 and 110 cm.

2.2.1. Oxygen (O_2) microelectrode profiles

Shallow sediment cores (20 cm) were collected at each sampling site (Siq13, SukB13 and SukS13, Table 1). Overlying water (4.5 cm) was left on top of the sediment to minimize atmospheric O_2 influence on the microprofile in the sediment (Boetius & Damm, 1998). Microelectrode O_2 profiles were conducted within 30 min of core collection using a Clark-style oxygen microelectrode (Unisense, Aarhus, Denmark) with a tip diameter of 500 µm. These microelectrodes respond in a linear fashion to O_2 concentration (Revsbech, 1989) and a two point calibration curve was used to standardize the instrument. The microelectrode was attached to a manual micromanipulator and lowered through the water and sediment column at 100 µm increments. Profiling was conducted in a darkened tent which provided a thermal barrier to prevent freezing and to maintain core temperatures (2.7-6.0 °C) during profiling.

Table 2.1 Sampling sites, dates, and analyses performed in each core. Cores were labeled using the first three letters to indicate the lake, the next letter to indicate the site (only in the case of Sukok Lake), followed by the last 2 digits of the year, and replicate cores indicated by a, b, c, and d.

Coordinates (N, W)	Sampling dates	O_2 profile	Pore water chemistry	Total Carbon	Lipid Biomarkers	*mcrA* qPCR	$[CH_4]$ & $\delta^{13}C_{CH4}$	Porosity	CH_4 production	Sediment texture
Siqlukaq										
71° 10.486' 156° 53.891'	29-Oct-2010	_	_	Siq10-a	Siq10-a	_	_	_	_	_
71° 10.481' 156° 53.910'	22-Oct-2011	_	_	Siq11-a	Siq11-a	Siq11-a	Siq11-a^	_	Siq11-b	Siq11-b
71° 10.482' 156° 53.900'	30-Oct-2012	_	Siq12-a*	_	_	_	Siq12-a	_	_	_
71° 10.457' 156° 54.004'	02-Nov-2012	_	Siq12-c*	_	_	_	Siq12-c	_	_	_
71° 10.451' 156° 53.872'	01-Nov-2013	Siq13-a	_	_	_	_	_	_	_	_
Sukok B										
71° 04.006' 156° 49.841'	24-Oct-2011	_	_	SukB11-a	SukB11-a	SukB11-a	SukB11-a	_	SukB11-b	SukB11-b
71° 04.006' 156° 49.841'	02-Nov-2013	SukB13-a	SukB13-b	_	_	_	SukB13-c	SukB13-c	_	_
71° 03.997 156° 49.918'	06-Nov-2013	_	SukB13-d	_	_	_	_	_	_	_
Sukok S										
71° 04.455' 156° 49.250'	27-Oct-2010	_	_	SukS10-a	SukS10-a	_	_	_	_	_
71° 04.519' 156° 49.208'	20-Oct-2011	_	_	SukS11-a	SukS11-a	SukS11-a	SukS11-a	_	SukS11-b	SukS11-b
71° 04.513' 156° 49.202'	31-Oct-2013	SukS13-a	SukS13-b	_	_	_	_	_	_	_
71° 04.519' 156° 49.204'	05-Nov-2013	_	SukS13-c	_	_	_	SukS13-d	SukS13-d	_	_

* Acetate concentration, Formate concentration, and ^ $\delta^{13}C_{CO2}$ were also analyzed in these cores

Depth-integrated aerobic O_2 consumption (IOC) was calculated using Fick's second law of diffusion assuming zero-order kinetics (Nielsen et al., 1990; Rasmussen & Jørgensen, 1992). The corrected diffusion coefficient (D_s) was calculated by adjusting the O_2 diffusion coefficient in fresh water (at sediment temperature during profiling) for porosity and tortuosity, based on measured porosity values and sediment type (Broecker & Peng, 1974; Rasmussen & Jørgensen, 1992).

2.2.2. Pore water chemistry

Dissolved gas and chemical gradients in the sediments, were determined for deeper cores (50-70 cm) collected from Siq (Siq12-a and Siq12-c) in 2012, and from Suk (SukS13-b, SukS13-c, SukB13-b, and SukB13-d) in 2013 (Table 1). Pore waters were sampled through pre-drilled holes in the core liners (sealed from the surrounding environment until samples were taken) using Rhizons (Seeberg-Elverfeldt et al., 2005). The ~ 0.15 μm porous membrane of each Rhizon was conditioned before sampling by rinsing with milli-Q water. Samples were drawn into 10 mL sterile syringes connected to the Rhizons.

Pore water samples for analysis of low molecular weight organic acids including acetate and formate, and anions SO_4^{2-} and NO_3^-, were collected in 2012 from Siq. Samples were collected (5 mL HDPE bottles), frozen, and then transported to the Biogeochemistry Laboratory at Indiana State University. A Dionex ICS-2000 with an AS11-HC column (Sunnyvale, CA) was utilized to measure the concentration

of each compound following Johnson et al. (2012) and Baker and Vervier (2004). Due to technical issues, acetate and formate were only measured in the surficial samples at Siq. Samples for analyses of the anions SO_4^{2-} and NO_3^- collected (15 mL polypropylene tubes) from Suk in 2013 were frozen, and then transported to the University of Tennessee, Knoxville. Pore waters were analyzed using a Dionex ICS-2100 RFIC fitted with an ASRS-4mm suppressor column, an AS18 analytical column, and an AG18 guard column (Sunnyvale, CA) following methods similar to Banihani et al. (2009).

Pore water samples for metals analysis collected (15 mL polypropylene tubes) from Siq and Suk, were acidified with UHP HNO_3 to a final concentration of 1 % HNO_3 (vol/vol) in the field, and stored at room temperature for < 60 days. The 5 mL samples were brought to 10 mL with 1 % UHP HNO_3 before analysis. A Thermo Element II high resolution inductively coupled plasma-mass spectrometer with a PFA-ST concentric Teflon nebulizer (ESI, Inc. Portland, OR) and a spray chamber of cyclonic glass (ESI, Inc. Portland, OR) in the Ultratrace Chemistry Laboratory at the Desert Research Institute in Reno, NV was used to quantify the metals. Low (LR) and medium (MR) resolutions were used as needed for isotopic separations. Standards were made from mixed stock standard from Inorganic Ventures, Inc. (Christiansburg, Virginia, USA) in a 1 % UHP HNO_3 matrix, and all blanks were made of 1 % UHP HNO_3.

2.3. Sediment CH_4

2.3.1. CH_4 concentrations and stable carbon isotope analyses

One sediment core per sampling site was collected in 2011 (Siq11-a, SukB11-a, and SukS11-a) to determine CH_4 concentrations and stable isotope signatures. The cores were sampled on site following Riedinger et al. (2010) and Koch et al. (2009). Samples were preserved at ~ 4 °C and analyzed at the University of California Santa Barbara (UCSB), for CH_4 concentration, $\delta^{13}C_{CH4}$, and $\delta^{13}C_{CO2}$, following methods by Kinnaman et al. (2007), with the exceptions that the bottle headspaces were displaced with 1-5 mL degassed water containing NaCl (35 g L^{-1}), and 1 mL sample was injected onto a 250 µL sample loop for quantitation. CH_4 concentration (µmoles CH_4 g^{-1} sediment dry weight) was estimated from the molar fraction of CH_4 in the headspace using equation 1 from D'Hondt et al. (2003), excluding the terms for porosity and sediment volume and including sediment dry weight.

To determine inter-annual variability of CH_4 concentrations and C isotopic composition, additional cores were collected from Siq in 2012 (Siq12-a and Siq12c) and from Suk (SukB13-c and SukS13-d) on November 2013. Parallel sediment plugs were collected for porosity analysis within 1.5 cm of samples for CH_4 and CO_2 analysis, and were stored at ~ 4 °C (Riedinger et al., 2010). The approximate ratio of sediment mass to volume in each syringe was obtained by measuring the total volume and wet weight of sediment. Water content was determined using dry

weight after heating at 105 °C for 24 hours. Porosity was calculated as water volume divided by wet sediment volume.

For comparison, CH_4 concentration was also estimated in μM following equation 1 in D'Hondt et al. (2003) and using an average porosity from depths sampled in SukB13-c and SukS13-d (data was not available for Siq), and an average sediment volume of 2.7 ± 0.7 mL. Sediment volume was estimated from the average bottle headspace, the known volume of NaCl solution, and the average known volume of the bottles with stoppers.

2.3.2. CH_4 production experiment

Samples were obtained for CH_4 production experiments from sediment cores collected in 2011 (Siq11, SukS11, and SukB11) at the same time and within ~ 15 cm of those for *in situ* CH_4 concentration and stable carbon isotope analysis (Table 1). Sediment cores for these experiments were transported at < 4 °C to a cold room, where they were maintained at 2 °C for one month. Each core was cut in three sections of approximately equal size, subsampled inside an anaerobic chamber (3-4 % H_2:N_2 atmosphere), and mixed with approximately equal volumes of cold, sterile, anoxic, deionized water by stirring, to eliminate any gas that may have been 'trapped' in the sediments (Kiene & Capone, 1985). Sediment slurries in 10 mL aliquots were distributed among 18 sterile, 125 mL-serum bottles (capped with butyl rubber stoppers) per depth. Negative controls with the same water and bottles were also prepared in the anaerobic chamber. The headspace of the bottles was

exchanged with ultra-high purity N_2 for 5 min using a manifold with 0.2 μm filters and sterile needles.

The sediment slurries were incubated upside down at 2 °C or 10 °C for the duration of the experiment. Time zero samples were collected after 2 hr incubation. For all time points, headspace samples were collected with a gas tight syringe (Hamilton, Reno, NV, USA) following vigorous shaking, before and after autoclaving of sediment slurries. Samples were analyzed with a Mini 2 gas chromatograph (GC; Shimadzu, Kyoto, Japan) equipped with a flame ionization detector (FID), a Poropak T column, and the following settings: oven temperature 20 °C, injection port temperature 210 °C, and ultra-high purity nitrogen as carrier gas at a flow rate of 40 mL min^{-1}. The injection volume was 100 μL. Peak area was quantified with a Hewlett-Packard 3390A integrator. CH_4 standards were used for calibration. The headspace of triplicate samples was analyzed every 5 to 10 days (depending on when CH_4 was first detected), for a period of 25 to 60 days depending on the lake.

At the end of the experiment, the dry weight of the sediments was determined after drying for 24 h at 105 °C. The CH_4 concentration in the headspace (μmoles CH_4 g^{-1} of dry weight) was determined as explained in section 2.3.1. Additionally, average CH_4 production (accumulation) over time was estimated. Methane production rates were calculated from a linear regression of three consecutive data points. The temperature coefficient (Q_{10}) was calculated following Duc et al. (2010).

2.4. Quantitative PCR (qPCR) of the methyl-coenzyme reductase alpha subunit gene (mcrA)

Sediment subsamples were obtained from the first 30 cm of the 2011 cores Siq11-a, SukS11-a, and SukB11-a (Table 1). Subsamples that were stored at -80 °C were later transferred to sucrose lysis buffer to preserve the integrity of the nucleic acids (SLB, 40 mM EDTA, 50 mM Tris-HCl, 0.75 M sucrose) prior to nucleic acid extractions. Community genomic DNA from surface sediment subsamples from each lake was extracted using a modified protocol for the power soil DNA isolation kit (MoBio, Carlsbad, CA, USA). Samples in SLB were thawed on ice for 45 min and centrifuged at 10,000 x g for 10 min. The supernatant was removed; ~ 0.5 g sediment subsamples (for a total weight of ~ 1 g per sample) were extracted following manufacturer's instructions, and later quantified using a standard picogreen assay (Life Technologies, Grand Island, NY, USA).

To quantify *mcrA* gene fragment copy number, qPCR was carried out using the ML primer pair (Luton et al., 2002) and the following conditions: 1X SYBR Green PCR master mix, 0.1 µM of each primer, 0.1 $\mu g.\mu L^{-1}$ of bovine serum albumin (BSA) in a 25 µL final volume. One µL of template DNA from Siq samples was used in 2-4 replicate reactions, and 4 µL of SukS or SukB samples was used in another set of 4 replicate reactions. A standard curve was prepared using *Methanocaldococcus jannaschii* genomic DNA with 1 µL of 10-fold dilutions covering 5 orders of magnitude (6.4×10^6 to 6.4×10^2 copies of *mcrA* gene assuming 1 copy of *mcrA* per genome), 4 replicates each. qPCR was performed using an Applied Biosystems 7500

Fast system (Life Technologies, Grand Island, NY, USA) in standard mode, and following PCR conditions by Luton et al. (2002). PCR efficiency was 75.7 % and amplification of standards was linear (r^2 = 0.993) from 10^2 to 10^6 copies of the template per µL. *McrA* gene copy numbers were expressed per g of wet sediment and per ng DNA. To confirm amplification specificity a melt curve analysis was performed immediately after qPCR using standard instrument settings, and agarose gel electrophoresis was used to confirm expected amplicon size.

2.5. Organic carbon and lipid biomarkers analyses

2.5.1. Sediment texture, total carbon, and inorganic carbon content

Sediment texture was determined following a simple texture analysis chart (Thien, 1979), with the same cores (Siq11-b, SukB11-b, and SukS11-b) used for the CH_4 production experiment. Total carbon and total inorganic carbon content was determined from five cores: Siq10-a, SukS10-a, Siq11-a, SukS11-a, and SukB11-a. Frozen cores were sectioned every 5 cm to a depth of 30 cm (2011 only) and every 10 cm for the remaining length of each core. These samples were collected into furnaced (550 °C, 8 hours) glass vials with foil lids that were refrozen (-20 °C) for transportation to the University of California, Riverside. Samples were lyophilized and then analyzed using an Eltra CS-500 carbon-sulfur analyzer, yielding total carbon (TC) and total inorganic carbon (TIC). Total organic carbon (TOC) was obtained by subtraction.

2.5.2. Lipid biomarker analyses

Lipid biomarkers were extracted from lyophilized samples obtained from the same cores sampled above using a Microwave Accelerated Reaction System (CEM Corp.) with 9:1 vol/vol dichloromethane: methanol to yield total lipid extracts (TLEs). One aliquot of each TLE was derivatized with N,O-bis(trimethylsilyl)trifluoroacetamide (BSFTA) in pyridine before gas chromatography-mass spectrometry (GC-MS). A second aliquot of total extract was fractionated via solid phase extraction columns to yield neutral lipids that were also subjected to GC-MS. Catalytic hydropyrolysis (HyPy; Love et al., 2005) was applied to an additional aliquot of freeze-dried sediment. HyPy conditions were 5 wt. % molybdenum sulfide catalyst, 150 bar H_2 flowing at 5 L min^{-1}, and programmed temperature ramp of 100 ^{o}C min^{-1} to 250 ^{o}C, followed by ramping at 8 ^{o}C min^{-1} to 480 ^{o}C. Gas chromatography-mass spectrometry (GC-MS) analyses of freely extractable and kerogen-bound saturated hydrocarbons, and derivatized total extracts were performed with an Agilent 5973 MSD mass spectrometer interfaced to an Agilent 7890A GC, equipped with a DB-1MS capillary column (60 m x 0.32 mm, 0.25μm film) and run with He as carrier gas . The temperature program for GC-MS full scan and selected ion monitoring was 60 °C (2 min), ramp to 250 °C at 20 °C min^{-1}, to 325 °C at 2 °C min^{-1}, and held at 325 °C for 20 min. Lipid biomarkers were identified by comparison with published mass spectra and retention times, and quantified using a d4-□□□-24-ethylcholestane internal standard. Archaeol was quantified using m/z 130 in selected ion monitoring mode with a d14-*para*-terphenyl internal standard (m/z 244) and calculated

response factor. The response factor was obtained by comparison of detector response areas between replicate analyses of known amounts of archaeol and d14-*p*-terphenyl.

3. Results

3.1. Sediment geochemistry

3.1.1. Oxygen profiles and depth integrated aerobic O_2 consumption

Dissolved O_2 concentration at the sediment-water interface in freshly collected cores from Siq13, SukB13, and SukS13 was 212.4, 110.0, and 6.5μmol L^{-1}, respectively (Fig. 2.2). Oxygen decreased rapidly with depth in all cores, becoming depleted by 1.0, 10.0 and 0.5 mm in the same respective cores. The depth of O_2 depletion typically coincided with concentration at the sediment-water interface in the Suk cores; cores from Siq did not fit this trend indicating that O_2 dynamics in Siq were different from that in Suk. Based on the shape of the O_2 depletion profiles, and assuming steady state conditions, average (± SD) rates of metabolic O_2 consumption were estimated to be 503.5-892.9, 40.8 ± 14.4, and 72.6 ± 19.1 μmol O_2 m^{-2} h^{-1} in cores from Siq, SukB, and SukS respectively.

3.1.2. Pore water chemistry

Pore water chemistry from Siq12 sediment cores was generally reproduced in replicate cores. The average concentration of NO_3^- (76.62 ± 73.18 μM) and SO_4^{2-} (8.74 ± 8.47 μM) in the upper 8 cm of the two replica cores was higher than the average concentration of NO_3^- (1.40 ± 2.51 μM) and SO_4^{2-} (0.96 ± 0.70 μM) below 8

cm (Figs. 3A and 3J). The concentration of total dissolved Fe in both cores was high throughout the depth profile, reaching a maximum (854.94 μM at 20 cm in Siq12-a)

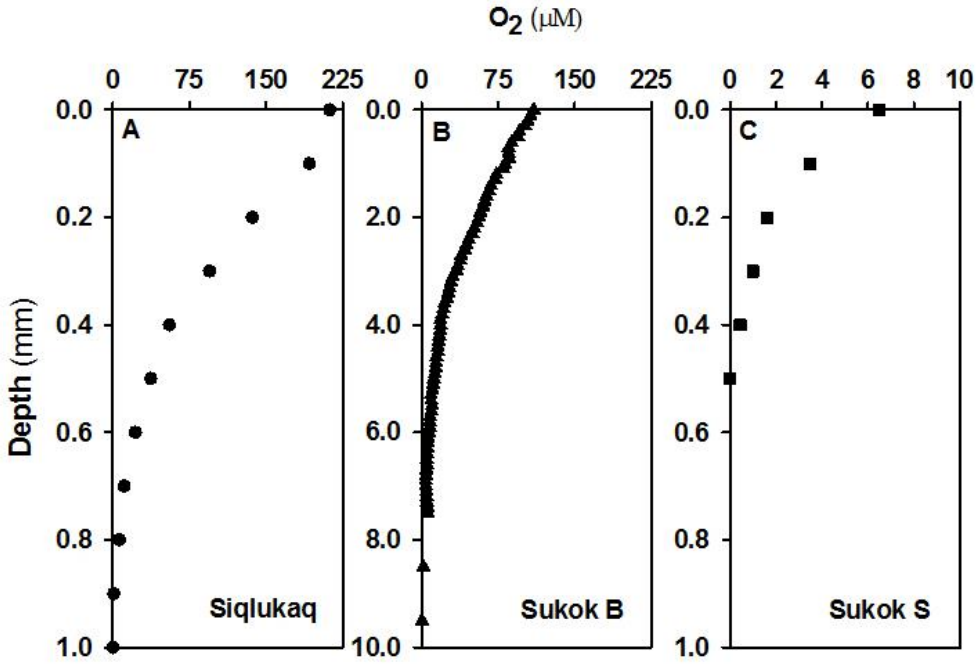


Figure 2.2 Oxygen microprofiles measured at 100 μm resolution from Siqlukaq, Sukok B, and Sukok S. (A) Siq13-a, (B) SukB13-a, (C) SukS13-a.

between 12 and 20 cm below the sediment-water interface. Additionally, the concentration of total dissolved Mn in the pore waters was about 44-fold lower than the concentration of total dissolved Fe, but showed a similar depth profile to Fe (Figs. 3D and 3G).

Overall, the concentration of NO_3^- and SO_4^{2-} in Suk was higher than the concentration of these anions in Siq. Reproducibility between cores at the same site was poor. Two maxima were detected in SukB13-b (2.34 mM at 10 cm and 745.73

μM at 18 cm) but the concentrations were much lower (0.00-27.47 μM) in the rest of the core (sampled to 30 cm), as well as in SukB13-d (0.87-52.58 μM; Fig. 2.3B). In SukS13-b increasing NO_3^- was detected with increasing depth (0.00-126.25 μM), but in SukS13-c, NO_3^- peaked at 11 cm (1.07 mM) and then decreased (Fig. 2.3C). The SO_4^{2-} concentration increased with depth in SukB13-b (0.00 μM at 2 cm to 751.68 μM at 30 cm), and in SukB13-d there was an average of 275.27 ± 112.74 μM throughout the core (sampled to 17 cm) with two maxima of ~ 450 μM (at 5 cm and 11 cm; Fig. 2.3K). The SO_4^{2-} concentration mostly increased with depth in the first ~ 20 cm (0.00-329.29 μM) of SukS13-b. A similar pattern was observed in SukS13-c down to ~ 17 cm (0.00-296.42 μM; Fig. 2.3L).

Depth profiles of dissolved Fe and Mn in SukB13 (Figs. 3E and 3H) were somewhat similar to Siq12. The concentration of dissolved Fe in SukB increased by 300-fold from 1 cm below the surface to 12 cm below the surface, and then decreased with depth; the concentration of dissolved Mn was relatively constant throughout the profile in SukB. The concentration of these metals throughout the sediment profile in SukS (Figs. 3F and 3I) showed a lack of a curve and was lower than in Siq (dissolved Fe ~ 8-fold and Mn ~ 3-fold) and SukB (dissolved Fe ~ 5-fold and Mn ~ 2-fold).

3.2. CH_4 concentrations and stable carbon isotope analyses

3.2.1. In situ CH_4 concentration and stable carbon isotopes

Average (± SD) CH_4 concentration was orders of magnitude higher in the upper intervals from Siq (2.18 ± 0.24 μmoles CH_4 g^{-1} dry sediment) than in SukB (3.12 ± 2.30 x 10^{-4} μmoles CH_4 g^{-1} dry sediment) and SukS (3.21 ± 1.44 x 10^{-3} μmoles CH_4 g^{-1} dry sediment; Fig 4A). CH_4 concentration decreased precipitously with depth in Siq sediments, from 0.84 μmoles CH_4 g^{-1} dry sediment at ~ 38 cm to 0.02 μmoles CH_4-g^{-1} sediment at 88 cm. Conversely, CH_4 concentration remained relatively constant throughout the SukB sediment profile (7.00 ± 4.10 x 10^{-4} μmoles CH_4 g^{-1} sediment), while CH_4 concentration increased with depth in the SukS sediments, with the highest amount of CH_4 observed at 75 cm (0.11 μmoles CH_4-g^{-1} dry sediment). The CH_4 level detected in the deepest interval of SukS core was 2 orders of magnitude higher than the amount of CH_4 detected at a similar depth in SukB.

Methane from Siq sediments was more ^{13}C-depleted than in SukS sediments (Fig. 2.4B). The most negative $\delta^{13}C_{CH4}$ values (-76.7 to -79.2 ± 0.2 $^0/_{00}$) were detected in Siq surface sediment, between 8 and 38 cm. However, the signal became less depleted in ^{13}C at 58 cm (-64.3 ± 1.6 $^0/_{00}$), and even less depleted at 88 cm (-43.3 ± 6 $^0/_{00}$). The $\delta^{13}C_{CH4}$ values in SukS sediments were less depleted at 35 cm (-55.5 ± 1.4 $^0/_{00}$) and at 55 cm (-47.3 ± 0.2 $^0/_{00}$) than those for Siq. The least depleted $\delta^{13}C_{CH4}$ value in SukS was observed at 75 cm (- 43.4 ± 0.2 $^0/_{00}$). The surface CH_4 concentrations for SukS sediments along with the entire depth profile in SukB were insufficient for the analysis of isotopes.

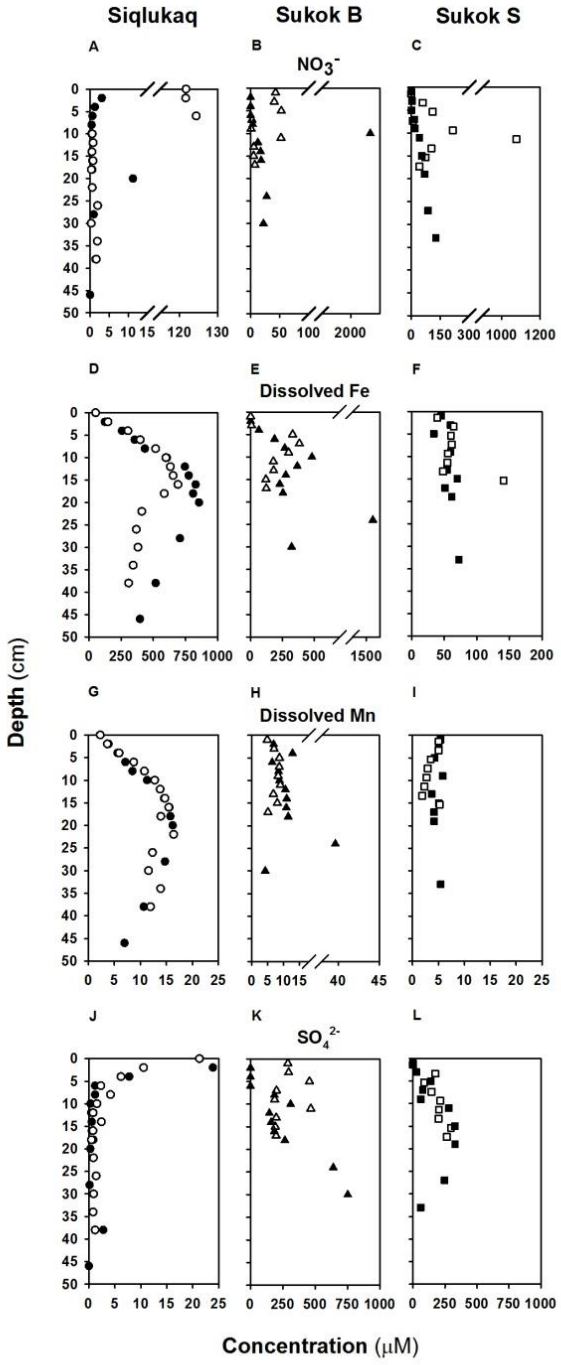


Figure 2.3 Pore water chemistry profiles from replicate sediment cores. (A, D, G, J) Siqlukaq Lake: Siq12-a: black, Siq12-c: white. (B, E, H, K) Sukok site distal to the seep: SukB13-b: black, SukB13-d: white. (C, F, I, L) Sukok Lake, seep site: SukS13-b: black, SukS13-c: white. Note different scales on different axes.

3.2.2. CH4 production experiment

Biological CH_4 production was observed in sediments from Siq11 at both temperatures (2 and 10 °C), and the upper sediments from SukB11 at 10 °C (Fig. 2.5). The highest amount of CH_4 produced was 7.4 ± 1.2 μmoles CH_4 g^{-1}dry weight (n=3) in the 2-20 cm interval of Siq sediments at 10 °C, and 3.7 ± 0.0 μmoles CH_4 g^{-1}dry weight (n=2) at 2 °C after 25 days of incubation (Fig. 2.5A). The next deeper interval of Siq sediment (26-47 cm) produced 0.9 ± 0.2 μmoles CH_4 g^{-1}dry weight (n=6) at 10 °C, and 0.2 ± 0.0 μmoles CH_4 g^{-1}dry weight (n=6) at 2 °C (Fig. 2.5B). CH_4 production from the 2-30 cm SukB-11 sediment interval was not detected until day 10 of incubation at 10 °C. After 50 days of incubation, 0.3 ± 0.1 μmoles CH_4 g^{-1}dry weight (n=6) were produced at 10 °C and < 0.1 μmoles CH_4 g^{-1}dry weight (n=6) were produced at 2 °C in this SukB upper interval (Fig. 2.5D).

Upper Siq and SukB sediments produced more CH_4 than any other interval of sediment sampled. CH_4 production < 0.1 μmoles CH_4 g^{-1}dry weight was observed in the deepest interval from Siq (49-76 cm; Fig. 2.5C), and no CH_4 production was observed in the deeper intervals from SukB (34-62 cm and 64-90 cm) during the incubation period (Figs. 5E and 5F). Furthermore, no CH_4 production was observed from the SukS sediments, although very small amounts of CH_4 were detected at all depths sampled.

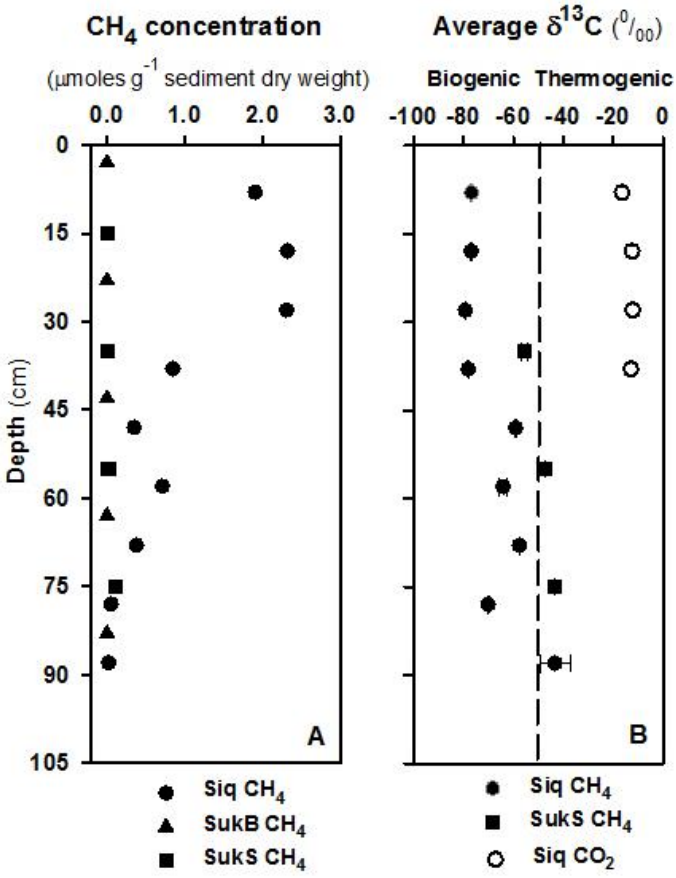


Figure 2.4 Methane concentration, $\delta^{13}C_{CH4}$ and $\delta^{13}C_{CO2}$ in arctic thermokarst lakes. (A) CH_4 concentration per gram of dry sediment (Siq11-a, SukB11-a, and SukS11-a). (B) $\delta^{13}C_{CH4}$ (Siq11-a and SukS11-a) and $\delta^{13}C_{CO2}$ (Siq11-a) in pore waters. The approximate boundary between biogenic and thermogenic CH_4 (Whiticar, 1999) is indicated with a dashed line for visual convenience. The error for the average CH_4 concentration (analytical replicates) was ± 2 %. The error for the average $\delta^{13}C_{CH4}$ (analytical replicates) was ± 0.2 $^0/_{00}$, except at 58 cm (± 1.6 $^0/_{00}$) and 88 cm (~ ± 6.0 $^0/_{00}$) in Siq, and at 55 cm (± 1.4 $^0/_{00}$) in SukS. For comparison to other data sets, CH_4 concentration in pore waters as µM was estimated in Section 4.5 (main text).

Temperature had a distinct influence on the rate of CH_4 production (obtained as the slope of a linear regression of 3 consecutive data points). CH_4 was produced

at a rate of 2.2 μmoles CH_4 day^{-1} g^{-1} dry weight at 10 °C and at a rate of 0.8 μmoles CH_4 day^{-1} g^{-1} dry weight at 2 °C in the upper Siq sediments. In the next deeper interval of Siq sediment, CH_4 was produced at a much lower rate: 0.3 μmoles CH_4 day^{-1} g^{-1} dry weight at 10 °C and < 0.1 μmoles CH_4 day^{-1} g^{-1} dry weight at 2 °C. The temperature coefficient (Q_{10}) in Siq was 3.7 for the upper sediment interval, and 9.5 for the next interval down indicating different metabolic responses of the microbial community through the sediment core. The rate of CH_4 production in the upper sediments of SukB could only be estimated at 10 °C (0.9 μmoles CH4 day^{-1} g^{-1}dry weight), given that CH_4 did not show a linear increase over time at 2 °C, hence, no Q_{10} value was computed.

3.3. Proxies for Methanogen abundance

The *mcrA* gene was detected in the surface sediments from the 3 sites by qPCR amplification, but only Siq11 gene copy numbers were at or above a conservative limit of detection that was established for this assay (100 copies *mcrA* μL^{-1}). The highest *mcrA* gene copy number detected was 1.9 x 10^4 *mcrA* copies g^{-1} sediment (5.4 x 10^2 *mcrA* copies ng^{-1} DNA) for Siq samples between ~ 14-15 cm below the surface, while the copy number between ~ 6-7 cm was 6.4 x 10^3 *mcrA* copies g^{-1} sediment (1.3 x 10^2 *mcrA* copies ng^{-1} DNA), and between 23-24 cm was 5.0 x 10^3 *mcrA* copies g^{-1} sediment (4.7 x 10^2 *mcrA* copies ng^{-1} DNA).

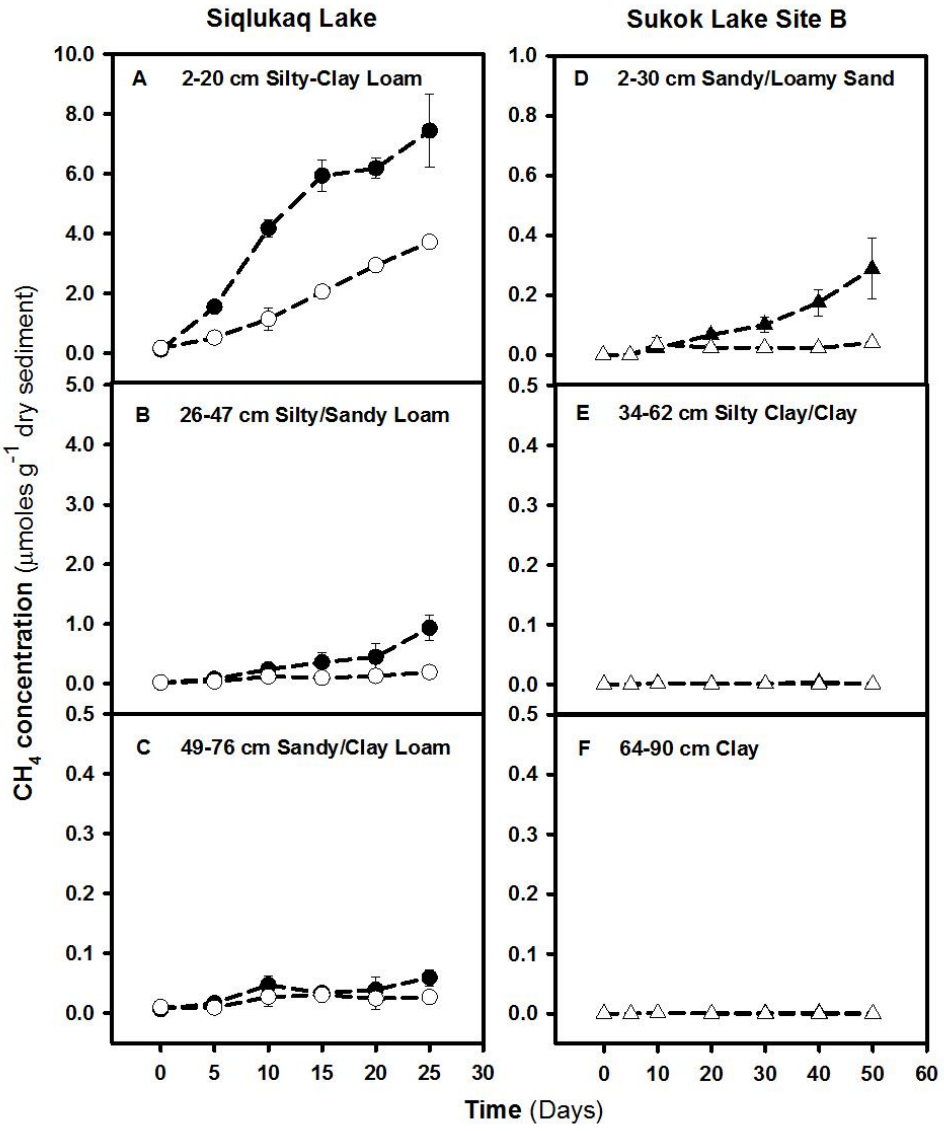


Figure 2.5 CH_4 production from sediment incubations at 2 temperatures using *in situ* organic matter. CH_4 production at 10 °C (black) and at 2 °C (white) at 3 different depths. Siq11-b (A) 2-20 cm, (B) 26-47, and (C) 49-76 cm. SukB11-b (D) 2-30 cm, (E) 34-62 cm, and (F) 64-90 cm. Note the different scales. CH_4 production was an average of 2 to 6 replica samples (see methods). Soil texture from the upper and lower layer of each sediment depth is described in each panel. CH_4 was detected in some incubations from SukS, but no pattern of CH_4 production was observed.

Archaeol was highest in Siq where two maxima were encountered: 45 and 37 µg g^{-1} sediment at ~ 12.5 cm and ~17.5 cm, but was lower outside these intervals

(Fig. 2.6B). Absolute archaeol concentration was greater in the distal seep site, SukB, beginning at 0.1 μg g^{-1} sediment at 7.0 cm, reaching a maximum of 1.4 μg g^{-1} sediment at ~ 26.5 cm, and then below detection for the lowermost portion of the studied samples. In the SukS11 core, the absolute archaeol concentration was slightly elevated in the upper ~ 23.5 cm but consistently low (≤ 0.07 μg g^{-1} sediment) throughout the entire 1 m core (2-3 orders of magnitude lower than at Siq and SukB). The archaeol to *n*-C_{31} ratio (Fig. 2.6C) was consistently higher in Siq (average 0.10 ± 0.05) and SukB (average 0.08 ± 0.06) than in the SukS site (0.01 ± 0.01).

3.4. Sediment properties, organic matter content and composition

Total organic carbon (TOC) measurements varied widely among and within cores analyzed in the three sites (Fig. 2.6A). TOC was highest in the upper 40 cm of Siq11 sediment (avg. 14.9 wt. %), then decreased to an average of 2.1 wt. % at depths greater than 40 cm below the sediment-water interface. The Siq11 sediment core had an overall average of 7.2 ± 2.9 wt. % TOC. SukB11 exhibited relatively low and consistent carbon contents down core (overall average 1.9 ± 3.7 wt. %), with the exception of a high TOC interval around 23 cm sediment depth (12.9 wt. %). Excluding this interval, the SukB11 core had an average of 0.9 ± 0.5 wt. % TOC. SukS sediments contained the lowest amounts of organic carbon, averaging 0.5 ± 0.3 wt. % over the entire SukS11 sediment core. Sediment TOC profiles from the 2011 samples reported here were consistent with those from a similar sample set collected in 2010 at the same locations (Table S2). Sediment texture analyses of

2011 sediment cores indicated that SukS and Siq have similar grain size profiles, with a dominance of clay and silt (90-95%) in the upper portion of the core, and 60-80% sand in the lower portions. SukB has the opposite profile, with >70% sand in the uppermost portion, and 90% silt and clay in the remainder of the core. SukB sediments contained carbonate, which was most likely detrital in origin.

Organic matter sources and composition were evaluated through analysis of saturated hydrocarbons, BSTFA-derivatized total lipid extracts, and catalytic hydropyrolysis products (functionalized free- and kerogen-bound lipids that have been converted to hydrocarbons). The proportions and amounts of these compounds differed from site to site and with depth at a given site (Tables S2 and S3). Both the extractable and the kerogen-bound hydrocarbons in Siq10 sediments were dominated by intermediate chain length *n*-alkanes. The most abundant *n*-alkane in free hydrocarbons was *n*-C_{23} and in the kerogen-bound fraction was *n*-C_{24}. The proportion of *n*-alkanes derived from aquatic plants (P_{aq} values, average 0.77 for the free hydrocarbons) was consistent with this intermediate chain length. These samples also contained low short to long *n*-alkane ratios, and relatively low sterane to hopane ratios (free sterane to hopane average 0.03, kerogen-bound sterane to hopane average 2.27). SukS10 organic matter showed a pronounced difference between free hydrocarbons and kerogen-bound hydrocarbons. The most abundant *n*-alkane in the free fraction was *n*-C_{31}, but the most abundant *n*-alkane in the bound-fraction was *n*-C_{16}. The short to long *n*-alkane ratio average was 0.02 in the free fraction, and the short to long *n*-alkane ratio average in the bound fraction

was 7.35. Free *n*-alkanes also had lower P_{aq} values (average 0.18) in this lake. Carbon preference indices (CPI) for both lakes showed strong odd-over-even predominance in free *n*-alkanes (from decarboxylation of free fatty acids) and even-over-odd in bound *n*-alkanes (from reduction of functionalized lipids).

4. Discussion

Methane emissions have received extensive attention in numerous environments (Wagner et al., 2007; Liu et al., 2013; Negandhi et al., 2013), where as much as 80-90 % of the atmospheric CH_4 is microbially derived (Whiticar, 1999). Furthermore, atmospheric CH_4 is a potent greenhouse gas that is currently rising (Hoehler & Alperin, 2014). Considering that the Arctic is highly sensitive to climate change (Kittel et al., 2011), accurate estimates of CH_4 emissions are of the utmost importance for the global CH_4 budget. Methane has also been recognized as a biosignature for life beyond Earth (Tazaz et al., 2013). In this section, we discuss the implications of our findings for North American Arctic CH_4 budgets.

4.1. CH4 sources and sediment biogeochemistry in Siqlukaq Lake

4.1.1. Sediment CH4 profiles

Detection of CH_4 (highest between 8-18 cm below the surface) in Siq was consistent with rapid O_2 depletion (~ 1 mm, Fig. 2.2) and pore water biogeochemistry (Fig. 2.3).

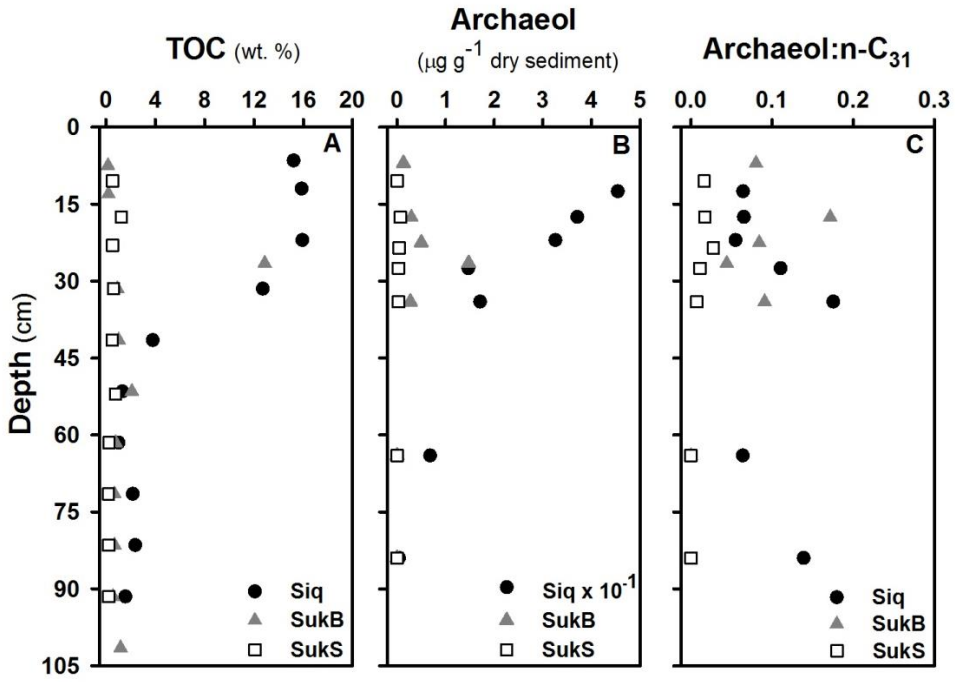


Figure 2.6 October 2011 sediment depth profiles of (A) total organic carbon, (B) archaeol concentration, and (C) archaeol to *n*-C_{31} ratios. Note differences in archaeol scales among sites.

The O_2 concentration in the water-sediment interface was within the range of concentrations detected in other shallow arctic lakes (Whalen et al., 2013), but O_2 was depleted in Siq at a shallower depth. High amounts of surficial TOC (Fig. 2.6) in combination with a silty clay loam, suggested that Siq sediments would become anoxic at shallow sediment depths. In fact, within the first 1 mm in the sediment profile O_2 was consumed at a rate higher than that computed in the other sediment cores we studied. This rate is at the low end of those observed in a eutrophic lake (Lake Zug, Switzerland; Maerki et al., 2009), and it is similar to the rate measured in a meso-eutrophic lake (Lake Vechten, The Netherlands; Sweerts et al., 1991). Moreover, SO_4^{2-} and NO_3^- were consumed within the first 10 cm below the

sediment-water interface, while dissolved Fe and Mn were chemically reduced, indicating thermodynamic conditions favorable for methanogenesis deeper in the profile. At depths > 48 cm, CH_4 concentration decreased in concert with TOC values. Overall, a positive correlation was found between the *in situ* CH_4 concentration and the TOC content in the sediments ($r^2 = 0.80$, $p = 0.008$), partly explaining the decrease in biological CH_4 production with depth.

4.1.2. $\delta^{13}C_{CH4}$

The stable isotope signature of C in CH_4 ($\delta^{13}C_{CH4}$) from Siq sediment suggested a biogenic source (Fig. 2.4 and Fig. S2*)*. The most depleted $\delta^{13}C_{CH4}$ values (-79.2 to -57.6 ± 0.2 $^{0}/_{00}$) fell within the range of $\delta^{13}C_{CH4}$ values recorded in the literature for biogenic CH_4 production in other arctic studies (Quay et al., 1988; Walter et al., 2008). Only the deepest sediment layer (~ 88 cm) in Siq showed a borderline thermogenic (c.f. Walter Anthony *et al.*, 2012) signal. At this depth, however, substrate depletion due to extensive organic matter decomposition (Pedersen et al., 2011) could result in biogenic CH_4 with heavier $\delta^{13}C_{CH4}$ values (Whiticar, 1999).

Biological CH_4 production pathways may be inferred by estimating the isotope separation factor (ε_C) between $\delta^{13}C_{CO2}$ and $\delta^{13}C_{CH4}$ and the apparent C fractionation factor (α_C) (Whiticar, 1999). For the upper 38 cm of Siq sediments, we calculated ε_C between 60.4 and 67.1, and α_C between 1.065 and 1.073. This corresponds to CH_4 production by CO_2 reduction (Whiticar, 1999), and is in accordance to the values observed by Walter et al. (2008) for ebullient CH_4 sources

in Siberian lakes. Alternative pathways, such as acetate fermentation, have been found to occur in freshwater sediments that are rich in organic carbon. Although acetate was present in the first few centimeters of Siq sediment (0.3-0.6 μM), it is possible that other forms of anaerobic respiration outcompeted methanogenesis in the use of acetate, leaving CO_2 reduction as the dominant pathway. For instance, the similarity between the dissolved Fe (Fig. 2.3D) and CH_4 (Fig. 2.4A) profiles is not currently understood in these lakes, though could be explained by a syntrophic relationship between iron reducing bacteria and methanogens (Zhou et al., 2014), in which iron-reducing bacteria oxidize acetate to CO_2, and the CO_2 is reduced by methanogenic Archaea to CH_4 (via the hydrogenotrophic pathway). Despite the little we know about iron reduction in the sediments of these lakes, this process has been connected to ecosystem respiration in drained-lake basins on the coastal plain of Alaska (Lipson et al., 2013).

4.1.3. CH_4 production

Methane was biologically produced from organic matter present in the first ~ 47 cm of Siq11 sediment at 2 °C and at 10 °C. Coincidentally, the amounts of CH_4 produced at 2 °C were similar to pore water CH_4 levels, indicating that CH_4 present in the sediments was most likely derived from *in situ*, present-day methanogenesis. The rate of CH_4 production at 10 °C in the upper sediments of Siq was in the range of the CH_4 production rates reported for the 0.4-9.0 cm sediment depth of three shallow (4.1-6.7 m) arctic lakes in the Alaskan Foothill (Bretz & Whalen, 2014).

These results also suggest that CH_4 was produced in a temperature-dependent fashion. This kind of temperature dependence is expected for biological reactions where enzymes are involved (Hochachka & Somero, 1973), and corroborates the temperature dependence of methanogenesis at the microbial community-level recently shown by Yvon-Durocher et al. (2014). As observed in surface (2-20 cm) Siq sediment, a 10 °C temperature rise can be correlated to a 2.5 to 3.5-fold increase in CH_4 production (Conrad & Schutz, 1988); however the 3-fold increase in the rate of CH_4 production in the next deeper (26-47 cm) Siq interval, supports the idea that decomposing structurally complex, aromatic molecules requires higher activation energies, causing enzymatic reactions to be more sensitive to temperature (Mikan et al., 2002; Davidson & Janssens, 2006). Below the methanogenic zone in Siq sediment, older, more recalcitrant organic matter may be found.

4.2. CH_4 sources and sediment biogeochemistry in Sukok Lake

4.2.1. Sediment CH_4 profiles

CH_4 was not retained in the sediments of Sukok Lake. The low CH_4 concentrations in both SukB and SukS sediments (Fig. 2.4) were surprising, considering the proximity of the SukS sediments to an ebullient gas seep (Fig. 2.1). SukS sediments transitioned from silty clay at the upper intervals, to a combination of sandy clay loam and loamy sand in the deeper intervals, perhaps reflecting current- or ebullition-induced winnowing that facilitated CH_4 channeling through the seep. Additionally, lower TOC contents in Suk compared to Siq (by a factor of ~ 4) may

have indirectly affected O_2 levels in the sediments, supporting lower O_2 consumption rates in SukB and in SukS versus Siq. TOC concentration was slightly higher in SukS than in SukB, and O_2 concentration was lower at the SukS water-sediment interface, but comparable in magnitude to another arctic lake in the Alaskan Foothill (Bretz & Whalen, 2014). The lower TOC content at both Suk sites may also explain the higher concentrations of alternative electron acceptors (NO_3^- and SO_4^{2-}) in the sediments of this lake, and establishes the conditions for other biogeochemical transformations to take place in the upper layers of SukB and SukS (*i.e.*, anaerobic methane oxidation).

4.2.2. $\delta^{13}C_{CH4}$

Physical mixing of CH_4 from different sources (*e.g.* microbial and thermogenic) could explain the "transitional" isotope signature observed in SukS. The primary source of CH_4 in the deeper sediment layers was consistent with thermogenic CH_4, though there may also be some biological CH_4 production in discrete layers of these sediments. Considering that Sukok Lake is located in a gas field, and that the sample from SukS was collected nearby an active CH_4 seep, a thermogenic CH_4 signal is plausible. Additional measures of C2-C4 hydrocarbon ratios or δD-CH_4 could be used to further substantiate this observation. Also, methanogenesis could be thermodynamically constrained by high CH_4 partial pressures (like the ones observed at the nearby CH_4 seep) in combination with low concentrations of methanogenic substrates (Chong et al., 2002), or by high redox potentials. Data from the CH_4 production experiment with SukS samples however ruled out the possibility

of thermodynamic inhibition imposed by CH_4 itself. Note that the $\delta^{13}C_{CH4}$ for SukB samples could not be determined due to insufficient CH_4 levels.

4.2.3. CH_4 production

Biological CH_4 production was not observed from any SukS11 interval at 2 or 10 °C, which is consistent with the isotopically heavy $\delta^{13}C_{CH4}$ in SukS11. However, CH_4 production at 10 °C was detected in the upper SukB11 sediments at a rate 23-fold lower than Siq11. Temperature had a weaker influence on CH_4 production in Suk than in Siq. According to Davidson and Janssens (2006), when substrate is abundant, temperature increases affect the maximum reaction rate (V_{max}), but under substrate-limiting conditions the substrate concentration at which the reaction rate equals V_{max} $[2\ (K_m)]^{-1}$ also increases with temperature, leading to a lower apparent temperature dependence. Therefore, the lag phase of CH_4 production observed in SukB surface sediments, may be consistent with the observation that organic matter in Suk has a higher vascular plant input and thus, may not only be less abundant, but also less labile than in Siq.

The fact that there was more TOC in the surface sediments of SukS than in SukB, but there was no CH_4 production in SukS at 10 °C indicates that biological CH_4 production in SukS sediments may be subject to other environmental constraints. Alternatively, longer incubation times or substrate additions could lead to CH_4 production from SukS, as reported in other ecosystems (*e.g.* Tibetan plateau lakes; Liu et al., 2013).

4.3. Proxies for Methanogen abundance as a control on CH_4 concentration in thermokarst lake sediments

Copy numbers of the *mcrA* gene in lake sediments were used to study the relationship between methanogen biomass and CH_4 concentration in sediment pore waters (Colwell et al., 2008; Freitag & Prosser, 2009; Liu et al., 2011). Assuming the presence of 1 copy of *mcrA* gene per genome of methanogenic Archaea, the copy number of the *mcrA* gene may be proportional to the number of cells with potential for CH_4 production (Luton et al., 2002; Steinberg & Regan, 2008) or consumption, because *mcrA* genes are also found in anaerobic CH_4 oxidizers (Raghoebarsing et al., 2006; Beal et al., 2009; Ettwig et al., 2010).

In Siq surface sediment the *mcrA* gene copy numbers were low but detectable, while *mcrA* copy numbers in Suk were below the limits of detection established for the assay at both sites. This pattern may at least partly explain the observed difference in CH_4 concentration between the 2 lakes and is a good approximation of relative differences between the lakes. In comparison to *mcrA* copy numbers found in a sample of active layer (permafrost) from the Canadian High Arctic (Yergeau et al., 2010), the *mcrA* copy numbers detected in Siq11 were 1 to 2 orders of magnitude higher. Also, our results are quite comparable to two active layer samples from the Western Canadian Arctic, where *mcrA* copies g^{-1} wet soil were between 10^3 and 10^6 (Frank-Fahle et al., 2014), though our highest copy number was 1.86×10^4 *mcrA* copies g^{-1} sediment in Siq.

Methanogen/archaeal biomass was corroborated by the detection of archaeol in the lake sediments (McCartney et al. 2013). Archaeol is an isoprenoid membrane lipid produced by Archaea that has been attributed in freshwater sediments and peats to methanogenic Archaea, (Pancost et al., 2011). The presence of archaeol may reflect active shallow sedimentary Archaea (Parkes et al., 2007), or preserved dead microbial biomass (Pancost et al., 2011; Bischoff et al., 2013). In either case, not only was the concentration of archaeol much higher in Siq than in SukB and in SukS, but it also tracked the trends of *mcrA* copy numbers per gram of sediment, amounts of TOC, and CH_4 concentrations observed in the depth profile of Siq, supporting the use of this compound as a proxy for methanogenic Archaea in lacustrine environments. Compared to the concentrations of archaeol detected in an area of continuous permafrost in Siberia (Bischoff et al., 2013), where CH_4 has been measured *in situ* and in incubations at 10 °C (Table S1), archaeol concentrations detected in the upper intervals of Siq were three orders of magnitude higher, and the concentrations detected in Suk were within the range of concentrations or slightly higher.

Archaeol abundance was also normalized to the C_{31} *n*-alkane abundance, a lipid likely to derive primarily from diagenesis of land plant leaf waxes, to correct for terrigenous inputs and to serve as a means of assessing which lake had the highest *in situ* archaeol production. Archaeol to *n*-C_{31} ratio was consistently higher in Siq and SukB than in the SukS site, indicating higher contributions from lake sedimentary Archaea, above baseline values transported in by soil lipid inputs.

4.4. Amount of organic matter, sources, and its relevance to CH_4 production

Within and among the studied thermokarst lakes, organic matter content and composition was heterogeneous. Unlike Siberian permafrost where the main source of organic matter is from the Pleistocene (Zimov et al., 1997), lacustrine sedimentary organic matter in these lakes appears to have a contribution from on-going *in* situ primary production, as observed in other lakes of the North Slope of Alaska (Bretz & Whalen, 2014). Although ubiquitous in the environment and of moderately low source specificity, *n*-alkane chain length has been shown to differ between vascular plant waxes (typically odd carbon number *n*-alkanes greater than C_{22}; Killops & Killops, 2005) and microbial lipids (approximately C_{16}-C_{24}; reviewed in Meyers & Ishiwatari, 1993). Additionally, *n*-alkane chain length may reflect the proportion of submerged versus emergent or terrigenous macrophytes, in which submerged macrophytes produce larger proportions of shorter (C_{23}, C_{25}) *n*-alkanes (Ficken et al., 2000).

The dominance of intermediate chain length *n*-alkanes in Siq sediments indicated a substantial organic contribution from a mixture of aquatic microbial lipids (bacteria) and algae, and terrestrial plants and soils. Siq P_{aq} values were consistent with a source from submerged macrophytes or microbial primary producers. Sterane to hopane ratios reflected the relative contributions of plants and algae versus hopanoid-producing bacteria, with a larger proportion of eukaryotic material present in the bound fraction.

Suk organic matter showed a greater contribution from allochthonous, perhaps more resistant or reworked organic matter. Free hydrocarbons in SukS were dominated by terrigenous leaf waxes (larger amounts of long-chain *n*-alkanes with high CPI; Tables S2 and S3) and trace amounts of mature diastereoisomers of steranes and hopanes, indicating a small contribution from petroleum at this site. P_{aq} values for the free *n*-alkanes in SukS sediments also reflected supply from terrigenous plants including bryophytes (*e.g.*, *Pogonatum sp.* which produces low amounts of C_{25}, C_{27}, and C_{23} *n*-alkanes; Haas, 1982), while the longer chain *n*-alkanes more likely derived from graminoids (*e.g.*, *Arctophila fulva, Carex spp.* and *Eriophorum spp.* which produce predominantly C_{27}, C_{31}, and C_{29} *n*-alkanes; Oros et al., 2006; Ronkainen et al., 2013). However, high sterane to hopane ratios (with C_{29} steranes the most abundant) and elevated short-chain *n*-alkanes in the kerogen-bound hydrocarbons showed significant inputs of likely microbially-derived *n*-alkanes. These findings from bound organics from the kerogen phase most likely reflect additional, though proportionately less abundant inputs from C_{29} steroid-producing algae or macrophytes (*e.g*. chlorophytes) in comparison to Siq.

Microbial CH_4 production rates in lake sediments can be controlled by the amount of dissolved or total organic carbon (Kelly & Chynoweth, 1981; Avery et al., 2003; Bergman et al., 2000; Liu et al., 2011). In this study, organic matter composition and proxies for methanogen abundance also seemed to influence CH_4 production. For instance, the maximum amount of organic carbon observed in SukB sediments was comparable to that of the methanogenic zone in Siq, however organic

carbon in Suk was more recalcitrant than in Siq, and archaeol concentrations in Suk were much lower, perhaps explaining the low to non-existent CH_4 production in this lake. Moreover, in Siq sediments TOC in excess of 10 wt. % extended to 42 cm, but archaeol concentrations diminished below ~ 22 cm, suggesting depletion of readily available fermentation products with depth as well. TOC content may not be used as sole predictor of short term CH_4 production in these lakes; the lability of sedimentary organic matter influences the amount and composition of substrates ultimately available for methanogenesis.

4.5. Implications for CH_4 production from permafrost in the North American Arctic

In comparison to permafrost samples from other arctic environments, the amount of CH_4 produced in incubations of Siq and SukB sediments was 2 to 3 orders of magnitude higher; while the amount of CH_4 accumulated in the sediments of the lakes was within the range of CH_4 concentrations detected in other arctic locations (Table S1). Note that we have not taken seasonal or inter-annual variation in this study (beyond dissolved CH_4 concentrations and the carbon isotopes), which may introduce some degree of uncertainty upon our observations. Assuming an average porosity of 0.28 ± 0.07, an average headspace volume of 0.009 ± 0.001 L , and an average sediment volume of 0.003 ± 0.001 L , CH_4 concentrations detected in Siq11 ranged between ~121.40 μM and ~4.89 mM. Conversely, CH_4 concentrations in SukB11 ranged between ~0.75 μM and ~5.80 μM, and CH_4 concentrations in SukS11 ranged between ~12.62 μM and ~405.12 μM. These concentrations are comparable

to CH_4 concentrations detected 60 miles to the south in Qalluuraq Lake, which has active gas seeps, although the concentrations in the first ~ 30 cm sediments of Siq11 were ~ 2.5 X higher than in Qalluuraq Lake (He et al., 2012). Moreover, CH_4 concentrations in Siq exceeded by approximately an order of magnitude the maximum CH_4 concentration detected in another shallow (6 m) arctic lake (GTH 112) in the Alaskan Foothills, which displayed a similar trend in the pore water CH_4 profile (Bretz & Whalen, 2014).

Our study demonstrated that methanogenic Archaea present in Alaska's North Slope thermokarst lakes are able to use *in situ* substrates for methanogenesis in a temperature-dependent fashion, and that the amount of CH_4 produced is proportional to the *mcrA* copy number, the concentration of archaeol, and the amount of labile organic matter in the sediments. These findings are particularly important when considering possible scenarios of climate change (Yvon-Durocher et al., 2014). The effect of increasing the temperature by 8 °C (from 2 to 10 °C) on CH_4 production rates was substantial for Siq. Currently, the largest source of CH_4 in this region of the North Slope of Alaska is release of thermogenic CH_4 (Walter Anthony et al., 2012). In scenarios of warming climate, our data lead us to contend that biological CH_4 production may play a larger role in CH_4 emissions in the future, though here, we have not considered CH_4 sinks (*e.g.* methanotrophy), which will also likely respond in parallel with temperature (Lofton et al., 2014).

A point of caution is that this study focused on the interior of the lakes, although it is possible CH_4 production varies throughout the lakes, and may be

strongest at organic-rich thermokarst lake margins. Therefore, future estimates of CH_4 emissions should comprise spatial characterization and include the organic-rich shelf area, to be an adequate predictor of CH_4 release from Alaskan thermokarst lakes. CH_4 emission estimates would also benefit from including annual components of the production cycle to account for the impact of temperature shifts. With these cautions in mind, this study constitutes an important first step in determining the contribution of biogenic CH_4 to CH_4 budgets in the changing Alaskan arctic environment in proximal, yet contrasting thermokarst lake ecosystems.

Acknowledgements

We especially thank A. Klesh, J. Leichty, and P. Santibañez, for assistance in the field; Katey Walter-Anthony for sharing her observations about the study area; and Frank Löffler for analyses of pore waters in the University of Tennessee. We are very grateful to N. Riedinger, J. Memmott, G. Miller, E. Ulrich, M. Miller, G. Trubl, J. Dodsworth, and J. Qualls for invaluable technical support. Likewise, we appreciate the efforts of the Barrow Arctic Science Consortium (BASC) and the UMIAQ Corporation in Barrow, AK for providing logistical support and insight into the local region. Special thanks to the anonymous reviewers of the manuscript for their comments and suggestions, and to Life Technologies for use of the Applied Biosystems 7500 Fast system to conduct qPCR. PMC was supported in part by the Division of Earth and Ecosystem Sciences, DRI. Funding for ABM was provided in part by NSF IGERT Program in Geobiological Systems (DGE 0654336). KPH and DB

acknowledge support through the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). Financial support for this work was provided in part by the NASA Astrobiology Institute, Astrobiology of Icy Worlds program at JPL, and a NASA Astrobiology Science and Technology for Exploring Planets (ASTEP) award (Project Narvak, NNN13D036T). Support from these programs is gratefully acknowledged.

References

Avery GB, Shannon RD, White JR, Martens CS, Alperin MJ (2003) Controls on methane production in a tidal freshwater estuary and a peatland: methane production via acetate fermentation and CO_2 reduction. *Biogeochemistry,* **62**, 19-37.

Baker MA, Vervier P (2004) Hydrological variability, organic matter supply and denitrification in the Garonne River ecosystem. *Freshwater Biology,* **49**, 181-190.

Banihani Q, Sierra-Alvarez R, Field J (2009) Nitrate and nitrite inhibition of methanogenesis during denitrification in granular biofilms and digested domestic sludges. *Biodegradation,* **20**, 801-812.

Beal EJ, House CH, Orphan VJ (2009) Manganese- and iron-dependent marine methane oxidation. *Science,* **325**, 184-187.

Bergman I, Klarqvist M, Nilsson M (2000) Seasonal variation in rates of methane production from peat of various botanical origins: effects of temperature and substrate quality. *FEMS microbiology ecology,* **33**, 181-189.

Bischoff J, Mangelsdorf K, Gattinger A, Schloter M, Kurchatova AN, Herzschuh U, Wagner D (2013) Response of methanogenic archaea to Late Pleistocene and Holocene climate changes in the Siberian Arctic. *Global Biogeochemical Cycles,* **27**, 305-317.

Boetius A, Damm E (1998) Benthic oxygen uptake, hydrolytic potentials and microbial biomass at the arctic continental slope. *Deep-Sea Research Part I-Oceanographic Research Papers,* **45**, 239-275.

Bonilla S, Villeneuve V, Vincent WF (2005) Benthic and planktonic algal communities in a High Arctic lake: Pigment structure and contrasting responses to nutrient enrichment. *Journal of Phycology,* **41**, 1120-1130.

Bretz KA, Whalen SC (2014) Methane cycling dynamics in sediments of Alaskan Arctic Foothill lakes. *Inland Waters,* **4**, 65-78.

Broecker WS, Peng TH (1974) Gas-exchange rates between air and sea. *Tellus,* **26**, 21-35.

Carson CE, Hussey KM (1962) The oriented lakes of Arctic Alaska. *The Journal of Geology,* **70**, 417-439.

Chong S, Liu Y, Cummins M, Valentine D, Boone D (2002) Methanogenium marinum sp. nov., a H_2-using methanogen from Skan Bay, Alaska, and kinetics of H_2 utilization. *Antonie Van Leeuwenhoek,* **81**, 263-270.

Colwell FS, Boyd S, Delwiche ME, Reed DW, Phelps TJ, Newby DT (2008) Estimates of biogenic methane production rates in deep marine sediments at Hydrate Ridge, Cascadia margin. *Applied and Environmental Microbiology,* **74**, 3444-3452.

Conrad R, Schutz H (1988) Methods for studying methanogenic bacteria and methanogenic activities in aquatic environments. In: *Methods in Aquatic Bacteriology.* (ed Austin B). Wiley, Chichester, UK, pp. 301-343.

D'hondt SL, Jørgensen BB, Miller DJ (2003) Chapter 5. Explanatory Notes. In: *Proceedings of the Ocean Drilling Program,* pp. 102.

Davidson EA, Janssens IA (2006) Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature,* **440**, 165-173.

Duc NT, Crill P, Bastviken D (2010) Implications of temperature and sediment characteristics on methane formation and oxidation in lake sediments. *Biogeochemistry,* **100**, 185-196.

Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MMM, Schreiber F, Dutilh BE, Zedelius J, De Beer D, Gloerich J, Wessels H, Van Alen T, Luesken F, Wu ML, Van De Pas-Schoonen KT, Den Camp H, Janssen-Megens EM, Francoijs KJ, Stunnenberg H, Weissenbach J, Jetten MSM, Strous M (2010) Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature,* **464**, 543-548.

Ficken KJ, Li B, Swain DL, Eglinton G (2000) An *n*-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. *Organic Geochemistry,* **31**, 745-749.

Frank-Fahle BA, Yergeau E, Greer CW, Lantuit H, Wagner D (2014) Microbial functional potential and community composition in permafrost-affected soils of the NW Canadian Arctic. *PLoS One,* **9**, e84761.

Freitag TE, Prosser JI (2009) Correlation of methane production and functional gene transcriptional activity in a peat soil. *Applied and Environmental Microbiology,* **75**, 6679-6687.

French HM (1976) *The periglacial environment,* Longman, Inc., New York.

Frohn RC, Hinkel KM, Eisner WR (2005) Satellite remote sensing classification of thaw lakes and drained thaw lake basins on the North Slope of Alaska. *Remote Sensing of Environment,* **97**, 116-126.

Glenn RK, Allen WW (1992) Geology, reservoir engineering and methane hydrate potential of the Walakpa Gas Field, North Slope, Alaska. In: *Final technical report (under grant DE-FG21-91MC28131) submitted to the U.S. Department of Energy.* North Slope Borough, Barrow, AK., pp. 1-26.

Haas K (1982) Surface wax of *Andreaea* and *Pogonatum* species. *Phytochemistry,* **21**, 657-659.

He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB (2012) Diversity of active aerobic methanotrophs along depth profiles of arctic and subarctic lake water column and sediments. *Isme Journal,* **6**, 1937-1948.

Hecky RE, Hesslein RH (1995) Contributions of benthic algae to lake food webs as revealed by stable isotope analysis. *Journal of the North American Benthological Society,* **14**, 631-653.

Hinkel KM, Eisner WR, Bockheim JG, Nelson FE, Peterson KM, Dai XY (2003) Spatial extent, age, and carbon stocks in drained thaw lake basins on the Barrow Peninsula, Alaska. *Arctic Antarctic and Alpine Research,* **35**, 291-300.

Hochachka PW, Somero GN (1973) *Strategies of Biochemical Adaptation,* Saunders, Philadelphia.

Hoehler T, Alperin MJ (2014) Methane minimalism. *Nature,* **507**, 436-437.

Houseknecht DW, Bird KJ, Schuenemeyer JH, Attanasi ED, Garrity CP, Schenk CJ, Charpentier RR, Pollastro RM, Cook TA, Klett TR (2010) 2010 updated assessment of undiscovered oil and gas resources of the National Petroleum Reserve in Alaska (NPRA). In: *U.S. Geological Survey Fact Sheet 2010,* pp. 4.

Howard HH, Prescott GW (1973) Seasonal variation of chemical parameters in Alaskan tundra lakes. *American Midland Naturalist,* **90**, 154-164.

Hussey KM, Michelson RW (1966) Tundra relief features near Point Barrow Alaska. *Arctic,* **19**, 162.

Johnson LT, Royer TV, Edgerton JM, Leff LG (2012) Manipulation of the dissolved organic carbon pool in an agricultural stream: responses in microbial community structure, denitrification, and assimilatory nitrogen uptake. *Ecosystems,* **15**, 1027-1038.

Jorgenson MT, Shur Y (2007) Evolution of lakes and basins in northern Alaska and discussion of the thaw lake cycle. *Journal of Geophysical Research,* **112**.

Jorgenson T, Yoshikawa K, Kanevskiy M, Shur Y, Romanovsky VE, Marchenko S, Grosse G, Brown J, Jones B (2008) Permafrost characteristics of Alaska. Institute of Northern Engineering, University of Alaska Fairbanks.

Judd AG (2000) Geological sources of methane. In: *Atmospheric methane. Its role in the global environment.* (ed Khalil MaK). Springer, New York, pp. 280-303.

Kelly CA, Chynoweth DP (1981) The contributions of temperature and the input of organic-matter in controlling rates of sediment methanogenesis. *Limnology and Oceanography,* **26**, 891-897.

Kiene RP, Capone DG (1985) Degassing of pore water methane during sediment incubations. *Applied and Environmental Microbiology,* **49**, 143-147.

Killops SD, Killops VJ (2005) *Introduction to organic geochemistry,* Blackwell Publishing, Ltd., Oxford.

Kinnaman FS, Valentine DL, Tyler SC (2007) Carbon and hydrogen isotope fractionation associated with the aerobic microbial oxidation of methane, ethane, propane and butane. *Geochimica et Cosmochimica Acta,* **71**, 271-283.

Kittel TGF, Baker BB, Higgins JV, Haney JC (2011) Climate vulnerability of ecosystems and landscapes on Alaska's North Slope. *Regional Environmental Change,* **11**, S249-S264.

Koch K, Knoblauch C, Wagner D (2009) Methanogenic community composition and anaerobic carbon turnover in submarine permafrost sediments of the Siberian Laptev Sea. *Environmental Microbiology,* **11**, 657-668.

Kokelj SV, Jorgenson MT (2013) Advances in Thermokarst Research. *Permafrost & Periglacial Processes,* **24**, 108-119.

Kvenvolden KA, Rogers BW (2005) Gaia's breath—global methane exhalations. *Marine and Petroleum Geology,* **22**, 579-590.

Ling F, Zhang T (2003) Numerical simulation of permafrost thermal regime and talik development under shallow thaw lakes on the Alaskan Arctic Coastal Plain. *Journal of Geophysical Research-Atmospheres,* **108**, 1-11.

Lipson DA, Raab TK, Goria D, Zlamal J (2013) The contribution of Fe(III) and humic acid reduction to ecosystem respiration in drained thaw lake basins of the Arctic Coastal Plain. *Global Biogeochemical Cycles,* **27**, 399-409.

Liu DY, Ding WX, Jia ZJ, Cai ZC (2011) Relation between methanogenic archaea and methane production potential in selected natural wetland ecosystems across China. *Biogeosciences,* **8**, 329-338.

Liu Y, Yao T, Gleixner G, Claus P, Conrad R (2013) Methanogenic pathways,^{13}C isotope fractionation, and archaeal community composition in lake sediments and wetland soils on the Tibetan Plateau. *Journal of Geophysical Research: Biogeosciences,* **118**, 650-664.

Lofton DD, Whalen SC, Hershey AE (2014) Effect of temperature on methane dynamics and evaluation of methane oxidation kinetics in shallow arctic Alaskan lakes. *Hydrobiologia,* **721**, 209-222.

Love GD, Bowden SA, Jahnke LL, Snape CE, Campbell CN, Day JG, Summons RE (2005) A catalytic hydropyrolysis method for the rapid screening of microbial cultures for lipid biomarkers. *Organic Geochemistry,* **36**, 63-82.

Luton PE, Wayne JM, Sharp RJ, Riley PW (2002) The *mcrA* gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfill. *Microbiology,* **148**, 3521-3530.

Maerki M, Muller B, Dinkel C, Wehrli B (2009) Mineralization pathways in lake sediments with different oxygen and organic carbon supply. *Limnology and Oceanography,* **54**, 428-438.

Mccartney CA, Bull ID, Waters SM, Dewhurst RJ (2013) Technical note: Comparison of biomarker and molecular biological methods for estimating methanogen abundance. *Journal of Animal Science,* **91**, 5724-5728.

Meyers PA, Ishiwatari R (1993) Lacustrine organic geochemistry - an overview of indicators of organic-matter sources and diagenesis in lake-sediments. *Organic Geochemistry,* **20**, 867-900.

Mikan CJ, Schimel JP, Doyle AP (2002) Temperature controls of microbial respiration in arctic tundra soils above and below freezing. *Soil Biology & Biochemistry,* **34**, 1785-1795.

Negandhi K, Laurion I, Whiticar MJ, Galand PE, Xu X, Lovejoy C (2013) Small thaw ponds: an unaccounted source of methane in the Canadian High Arctic. *PLoS ONE,* **8**, e78204.

Nielsen LP, Christensen PB, Revsbech NP, Sorensen J (1990) Denitrification and oxygen respiration in biofilms studied with a microsensor for nitrous-oxide and oxygen. *Microbial Ecology,* **19**, 63-72.

Oechel WC, Hastings SJ, Vourlrtis G, Jenkins M, Riechers G, Grulke N (1993) Recent change of arctic tundra ecosystems from a net carbon dioxide sink to a source. *Nature,* **361**, 520-523.

Oros DR, Abas MRB, Omar NYMJ, Rahman NA, Simoneit BRT (2006) Identification and emission factors of molecular tracers in organic aerosols from biomass burning: Part 3. Grasses. *Applied Geochemistry,* **21**, 919-940.

Pancost RD, Mcclymont EL, Bingham EM, Roberts Z, Charman DJ, Hornibrook ERC, Blundell A, Chambers FM, Lim KLH, Evershed RP (2011) Archaeol as a methanogen biomarker in ombrotrophic bogs. *Organic Geochemistry,* **42**, 1279-1287.

Parkes RJ, Cragg BA, Banning N, Brock F, Webster G, Fry JC, Hornibrook E, Pancost RD, Kelly S, Knab N, Jorgensen BB, Rinna J, Weightman AJ (2007) Biogeochemistry and

biodiversity of methane cycling in subsurface marine sediments (Skagerrak, Denmark). *Environmental Microbiology,* **9**, 1146-1161.

Parsekian AD, Grosse G, Walbrecker JO, Muller-Petke M, Keating K, Liu L, Jones BM, Knight R (2013) Detecting unfrozen sediments below thermokarst lakes with surface nuclear magnetic resonance. *Geophysical Research Letters,* **40**, 535-540.

Pedersen JA, Simpson MA, Bockheim JG, Kumar K (2011) Characterization of soil organic carbon in drained thaw-lake basins of Arctic Alaska using NMR and FTIR photoacoustic spectroscopy. *Organic Geochemistry,* **42**, 947-954.

Phelps AR, Peterson KM, Jeffries MO (1998) Methane efflux from high-latitude lakes during spring ice melt. *Journal of Geophysical Research,* **103**, 29029.

Quay PD, King SL, Lansdown JM, Wilbur DO (1988) Isotopic composition of methane released from wetlands: Implications for the increase in atmospheric methane. *Global Biogeochemical Cycles,* **2**, 385-397.

Raghoebarsing AA, Pol A, Van De Pas-Schoonen KT, Smolders AJ, Ettwig KF, Rijpstra WI, Schouten S, Damste JS, Op Den Camp HJ, Jetten MS, Strous M (2006) A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature,* **440**, 918-921.

Ramlal PS, Hesslein RH, Hecky RE, Fee EJ, Rudd JWM, Guildford SJ (1994) The organic-carbon budget of a shallow arctic tundra lake on the Tuktoyaktuk Peninsula, N. W. T., Canada *Biogeochemistry,* **24**, 145-172.

Rasmussen H, Jørgensen BB (1992) Microelectrode studies of seasonal oxygen-uptake in a coastal sediment - Role of molecular diffusion. *Marine Ecology Progress Series,* **81**, 289-303.

Repenning CA (1983) New evidence for the age of the Gubik Formation Alaskan North Slope. *Quaternary Research,* **19**, 356-372.

Revsbech NP (1989) An oxygen microsensor with a guard cathode. *Limnology and Oceanography,* **34**, 474-478.

Riedinger N, Brunner B, Lin YS, Voßmeyer A, Ferdelman TG, Jørgensen BB (2010) Methane at the sediment–water transition in Black Sea sediments. *Chemical Geology,* **274**, 29-37.

Ronkainen T, Mcclymont EL, Väliranta M, Tuittila E-S (2013) The *n*-alkane and sterol composition of living fen plants as a potential tool for palaeoecological studies. *Organic Geochemistry,* **59**, 1-9.

Schuur EaG, Vogel JG, Crummer KG, Lee H, Sickman JO, Osterkamp TE (2009) The effect of permafrost thaw on old carbon release and net carbon exchange from tundra. *Nature,* **459**, 556-559.

Seeberg-Elverfeldt J, Schlüter M, Feseker T, Kölling M (2005) Rhizon sampling of pore waters near the sediment/water interface of aquatic systems. *Limnology and oceanography,* **Methods 3**, 361-371.

Sellmann PV, Bown J, Lewellen RI, Mckim H, Merry C (1975) The classification and geomorphic implications of thaw lakes on the arctic coastal plain, Alaska, U.S. In: *Army CRREL Research Report,* pp. 21.

Steinberg LM, Regan JM (2008) Phylogenetic comparison of the methanogenic communities from an acidic, oligotrophic fen and an anaerobic digester treating municipal wastewater sludge. *Applied and Environmental Microbiology,* **74**, 6663-6671.

Sweerts J, Bargilissen MJ, Cornelese AA, Cappenberg TE (1991) Oxygen-consuming processes at the profundal and littoral sediment water interface of a small meso-

eutrophic lake (Lake Vechten, the Netherlands). *Limnology and Oceanography,* **36**, 1124-1133.

Tazaz AM, Bebout BM, Kelley CA, Poole J, Chanton JP (2013) Redefining the isotopic boundaries of biogenic methane: Methane from endoevaporites. *Icarus,* **224**, 268-275.

Thien SJ (1979) A flow diagram for teaching texture-by-feel analysis. *Journal of Agronomic Education.,* **8**, 54-55.

Wagner D, Gattinger A, Embacher A, Pfeiffer E-M, Schloter M, Lipski A (2007) Methanogenic activity and biomass in Holocene permafrost deposits of the Lena Delta, Siberian Arctic and its implication for the global methane budget. *Global Change Biology,* **13**, 1089-1099.

Walter Anthony KM, Anthony P, Grosse G, Chanton J (2012) Geologic methane seeps along boundaries of arctic permafrost thaw and melting glaciers. *Nature Geoscience,* **5**, 419-426.

Walter KM, Chanton JP, Chapin FS, Schuur EaG, Zimov SA (2008) Methane production and bubble emissions from arctic lakes: Isotopic implications for source pathways and ages. *Journal of Geophysical Research,* **113**.

Walter KM, Smith LC, Chapin FS, 3rd (2007) Methane bubbling from northern lakes: present and future contributions to the global methane budget. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences,* **365**, 1657-1676.

Whalen SC, Lofton DD, Mcgowan GE, Strohm A (2013) Microphytobenthos in shallow arctic Lakes: fine-scale depth distribution of chlorophyll a, radiocarbon assimilation, irradiance, and sissolved O_2. *Arctic, Antarctic, and Alpine Research,* **45**, 285-295.

Whiticar MJ (1999) Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chemical Geology,* **161**, 291-314.

Yergeau E, Hogues H, Whyte LG, Greer CW (2010) The functional potential of High Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *The ISME journal,* **4**, 1206-1214.

Yvon-Durocher G, Allen AP, Bastviken D, Conrad R, Gudasz C, St-Pierre A, Thanh-Duc N, Del Giorgio PA (2014) Methane fluxes show consistent temperature dependence across microbial to ecosystem scales. *Nature,* **507**, 488-491.

Zhou S, Xu J, Yang G, Zhuang L (2014) Methanogenesis affected by the co-occurrence of iron(III) oxides and humic substances. *FEMS microbiology ecology,* **88**, 107-120.

Zimov SA, Voropaev YV, Semiletov IP, Davidov SP, Prosiannikov SF, Chapin FS, Chapin MC, Trumbore S, Tyler S (1997) North Siberian lakes: A methane source fueled by Pleistocene carbon. *Science,* **277**, 800-802.

Supplementary Tables and Figures

Table S2.1 CH_4 concentration and CH_4 production rates reported in the literature for arctic permafrost samples.

Location	Type of sample	Depth (cm)	% Total Organic Carbon (mg g^{-1})	CH_4 (µmol g-1)	CH4 production rate (µmoles CH4 g-1 day-1)	Annotated citation
Lena Delta, Siberia	Holocene permafrost deposits	0-150	4.2 % (42)	0.282	< ~ 0.0036 at 5 °C	Wagner et al. (2007). Sampled in the summer of 2001. TOC is the median for the specific time unit (Unit I).
Lena Delta, Siberia	Low-centered ice-wedge permafrost polygon	0-100	< ~ 4-18 % (40-180)	0.004-3.600	< ~ 0.0022 at 10 °C	Bischoff et al. (2013). Sampled in the summer of 2002.
Laptev Sea coast	Polygon center, Samoylov Island	0-40	3.6-18.3 % (36-183)	0.150-542.0	< ~0.0144 at 5 °C	Ganzert et al. (2007).
	Floodplain, Samoylov Island	0-52	0.8-3.1 % (8-31)	0.004-0.411	≤ ~ 0.0144	
Laptev Sea	Submarine permafrost sediments	40.5-63.4 (mbsf)	0.03-8.7 % (0.3-87)	0.000-0.284		Koch et al. (2009). Sampled in the spring of 2005.
Cold foot, Alaska	Permafrost-protected soil	35-80	0.063 ± 0.041- 0.087 ± 0.07 (mg DOC. g^{-1})		~ 0.001-0.010 at 5 °C	Waldrop et al. (2010). Sampled in the spring of 2007. Location North of the Arctic Circle at 67° N, 150° W. Organic-mineral active layer between 35-50 cm, and continuous mineral permafrost between 65-80 cm.
Hess Creek, Alaska	Permafrost-protected soils	35-100	0.46 ± 0.031-0.72 ± 0.28 (mg DOC. g^{-1})		~ 0.010-0.100 at 5 °C	Waldrop et al. (2010). Lowland south of Yukon River, below Arctic Circle at 65° N, 149° W. Samples composed of organic active layer and organic permafrost
Herschel Island, Canadian Arctic	Drained lake, low centered polygon	0-35	23-28 % (230-280)		0.0336 -0.924 at 10 °C	Barbier et al. (2012). Samples collected in the summer of 2010. Location at 69° N, 138° W. Active layer (0-36 cm). No CH4 production was detected between 0-5 cm.

Table S2.2 Free hydrocarbon biomarker extracted from Siqlukaq and Sukok sediments.

Sample	Depth (cm)	TOC (wt. %)[1]	Sterane: Hopane[2]	Short:Long *n*-alkanes[3]	ACL (odd)[4]	Most abundant *n*-alkane	CPI[5]	P_{aq}[6]
Siq10-a	6.5	14.23	0.01	0.13	27	23	3.06	0.77
	21.5	4.84	0.02	0.04	27	23	2.45	0.82
	38.0	0.89	0.05	0.01	27	27	3.33	0.72
SukS10-a	5.0	1.03	0.29	0.01	29	31	9.94	0.27
	15.0	1.17	0.27	0.01	29	31	13.31	0.18
	25.0	0.97	0.30	0.02	30	31	13.41	0.13
	35.0	1.11	0.43	0.01	29	31	12.84	0.18
	45.0	1.06	0.34	0.03	30	31	11.24	0.15
	55.0	0.77	0.36	0.00	30	31	14.30	0.11
	65.0	1.62	0.35	0.01	30	31	14.94	0.10
	75.0	0.80	0.44	0.06	28	27	5.05	0.39
	85.0	1.82	0.33	0.02	30	31	14.40	0.17
	99.0	1.04	0.36	0.00	29	31	14.28	0.13

[1]TOC, Total Organic Carbon

[2]Calculated as: all isomers of C_{27}-C_{29} steranes/all isomers of C_{29}-C_{35} desmethylhopanes

[3]Calculated as: (C_{16}-C_{20})/(C_{27}-C_{31})

[4]Average chain length calculated as: (25*area C_{25} + 27*area C_{27} + 29*area C_{29} + 31*area C_{31} + 33*area C_{33})/(area C_{25} + area C_{27} + area C_{29} + area C_{31} + area C_{33})

[5]Carbon preference index calculated as: (0.5)*((C_{25} + C_{27} + C_{29} + C_{31} + C_{33})/(C_{24} + C_{26} + C_{28} + C_{30} + C_{32})+(C_{25} + C_{27} + C_{29} + C_{31} + C_{33})/(C_{26} + C_{28} + C_{30} + C_{32} + C_{34}))

[6]Proportion of *n*-alkanes derived from aquatic plants calculated as (C_{23} + C_{25})/(C_{23} + C_{25} + C_{29} + C_{31})

Table S2.3 Lipid biomarker ratios from catalytic hydropyrolysis hydrocarbon products.

Sample	Depth (cm)	Sterane: Hopane [2]	Short:Long *n*-alkanes[3]	ACL (even)[4]	Most abundant *n*-alkane	CPI[5]	P_{aq}[6]
Siq10-a	6.5	2.4	0.59	27	24	0.54	0.74
	21.5	1.3	0.56	27	24	0.42	0.56
	38.0	3.1	1.13	27	24	0.34	0.67
SukS10-a	5.0	25.8	5.73	25	24	0.29	0.91
	15.0	17.9	3.42	26	24	0.28	0.85
	25.0	18.0	6.73	25	24	0.32	0.91
	35.0	8.8	4.02	26	24	0.40	0.85
	45.0	31.0	3.95	26	24	0.36	0.82
	55.0	14.8	13.12	25	18	0.39	0.96
	65.0	30.1	2.60	26	24	0.40	0.75
	75.0	6.3	16.83	25	20	0.49	1.00
	85.0	46.3	7.01	25	24	0.41	0.94
	99.0	57.2	10.11	25	18	0.36	0.93

[1]TOC, Total Organic Carbon

[2]Calculated as: all isomers of C_{27}-C_{29} steranes/all isomers of C_{29}-C_{35} desmethylhopanes

[3]Calculated as: (C_{16}-C_{20})/(C_{27}-C_{31})

[4]Calculated as: (24*area C_{24} + 26*area C_{26} + 28*area C_{28} + 30*area C_{30} + 32*area C_{32})/(area C_{24} +area C_{26} + area C_{28} + area C_{30} + area C_{32})

[5]Carbon preference index calculated as: (0.5)*((C_{25} + C_{27} + C_{29} + C_{31} + C_{33})/(C_{24}+ C_{26} + C_{28} + C_{30}+ C_{32})+(C_{25} + C_{27} + C_{29} + C_{31} + C_{33})/(C_{26} + C_{28} + C_{30} + C_{32} + C_{34}))

[6]Proportion of *n*-alkanes derived from aquatic plants calculated as (C_{23} + C_{25})/(C_{23} + C_{25} + C_{29} + C_{31}).

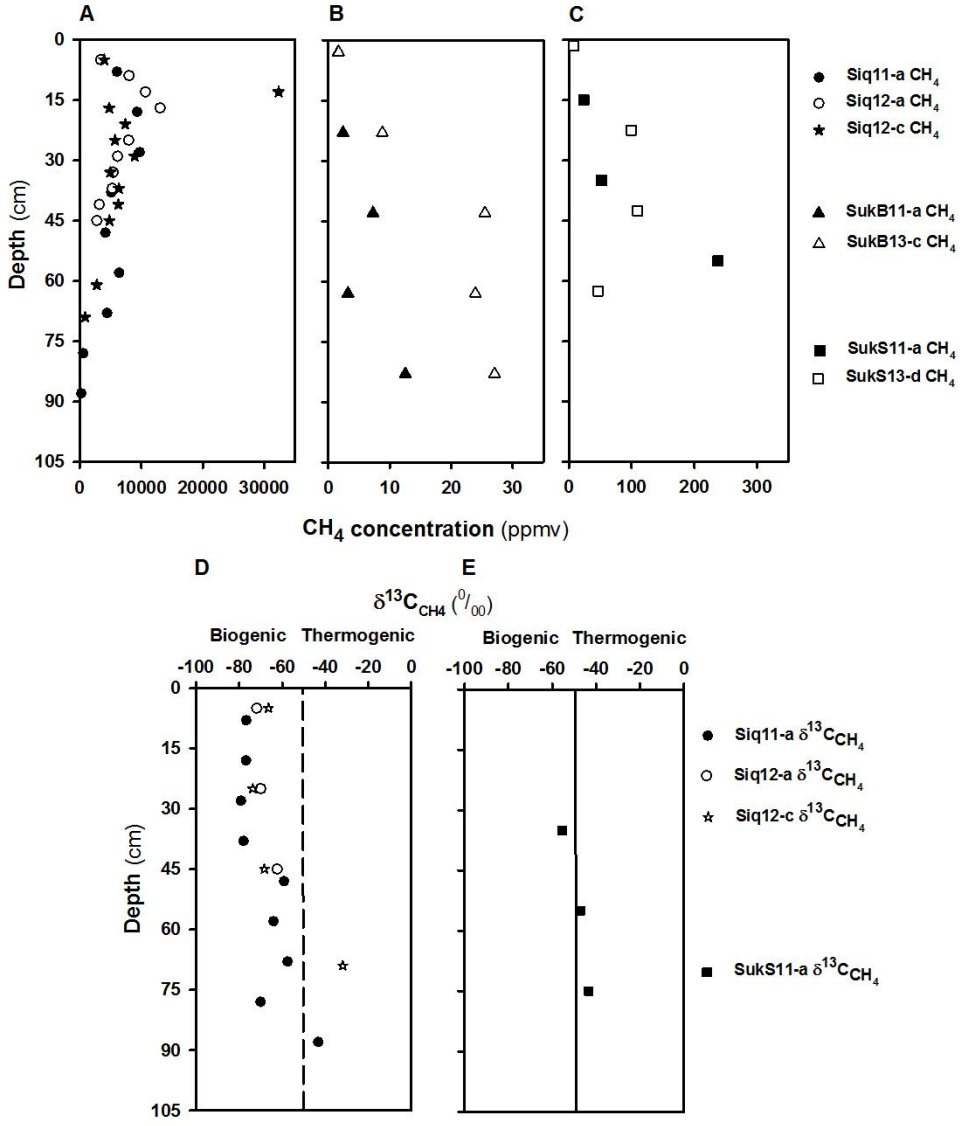


Figure S2.1 Methane concentration and $\delta^{13}C_{CH4}$ from all the cores collected for these analyses. (A-C) CH_4 concentration and (D,E) $\delta^{13}C_{CH4}$ in pore waters. Note the different x-axis scales. $\delta^{13}C_{CH4}$ could not be measured for Sukok lake samples collected in 2013 due to CH_4 concentrations too low for the analysis. The approximate boundary between biogenic and thermogenic (Whiticar, 1999) is indicated with a dashed line. The error for the average CH_4 concentration (analytical replicates) was ± 2 %. The error for the average $\delta^{13}C_{CH4}$ (analytical replicates) was ± 0.2 $^0/_{00}$, except: 58 cm (± 1.6 $^0/_{00}$) and 88 cm (~ ± 6.0 $^0/_{00}$) in Siq11-a; 25 cm (± 0.4 $^0/_{00}$) and 45 cm (± 0.8 $^0/_{00}$) in Siq12a; 5 cm (± 0.4 $^0/_{00}$), 25 cm (± 0.3 $^0/_{00}$), and 69 cm (~ ± 8.1 $^0/_{00}$) in Siq12-c; and 55 cm (± 1.4 $^0/_{00}$) in SukS.

Supplementary References

Barbier BA, Dziduch I, Liebner S, Ganzert L, Lantuit H, Pollard W, Wagner D (2012) Methane-cycling communities in a permafrost-affected soil on Herschel Island, Western Canadian Arctic: active layer profiling of mcrA and pmoA genes. *FEMS microbiology ecology,* **82**, 287-302.

Bischoff J, Mangelsdorf K, Gattinger A, Schloter M, Kurchatova AN, Herzschuh U, Wagner D (2013) Response of methanogenic archaea to Late Pleistocene and Holocene climate changes in the Siberian Arctic. *Global Biogeochemical Cycles,* **27**, 305-317.

Ganzert L, Jurgens G, Munster U, Wagner D (2007) Methanogenic communities in permafrost-affected soils of the Laptev Sea coast, Siberian Arctic, characterized by 16S rRNA gene fingerprints. *FEMS microbiology ecology,* **59**, 476-488.

Koch K, Knoblauch C, Wagner D (2009) Methanogenic community composition and anaerobic carbon turnover in submarine permafrost sediments of the Siberian Laptev Sea. *Environmental Microbiology,* **11**, 657-668.

Wagner D, Gattinger A, Embacher A, Pfeiffer E-M, Schloter M, Lipski A (2007) Methanogenic activity and biomass in Holocene permafrost deposits of the Lena Delta, Siberian Arctic and its implication for the global methane budget. *Global Change Biology,* **13**, 1089-1099.

Waldrop MP, Wickland KP, White R, Berhe AA, Harden JW, Romanovsky VE (2010) Molecular investigations into a globally important carbon pool: permafrost-protected carbon in Alaskan soils. *Global Change Biology,* **16**, 2543-2554.

Chapter 3

Diversity, structure, functional potential of the microbial community, and phylogeny of Archaea in sediments of Alaskan Arctic thermokarst lakes

Co-authorship statement

P. B. Matheus Carnevali, C. Herbold, K. P. Hand, and A. E. Murray

PM and AM collected samples and designed the experiments. PM executed the experiments and data analyses with mentoring by AM. CH performed statistical analyses in R. PM wrote the chapter under the supervision of AM. AM and KH secured support for the research program and AM secured support for the NGS.

Abstract

Thermokarst lakes sediments provide ideal conditions for organic matter degradation through fermentation and respiration processes, which ultimately lead to the production of greenhouse gases such as CO_2 and CH_4. Microorganisms inhabiting the sediments are the engine behind some of these processes, and their metabolic potential determines the fate of important nutrients in the environment. Two lakes (Siqlukaq Lake –Siq, and Sukok Lake – Suk) on the Coastal Plain of Alaska contain 0.2 to 15 wt. % of total organic carbon (TOC) in their sediments, and the lake with higher TOC content (Siq) exhibited higher concentrations and a higher potential for biological CH_4 production (Chapter 2). Likewise, there are different geochemical gradients in the sediments of Siq and Suk; and in one site in Suk, there was an active thermogenic CH_4 seep at the time of sampling. Our goals for this study were to understand the effect that the sediment geochemistry (and vice versa) may have on the diversity and structure of the active microbial community (Bacteria and Archaea), and to determine the phylogeny of Archaea inhabiting these arctic thermokarst lake sediments. To study the microbial community as represented by the small subunit (SSU) ribosomal RNA (rRNA) gene and the SSU rRNA, we used a combination of molecular approaches, including MiSeq next-generation sequencing (iTag sequencing), archaeal variable region 3 (V3) SSU rRNA gene denaturing gradient gel electrophoresis (DGGE), and archaeal V3 SSU rRNA gene cloning. We focused on the analyses of rRNA iTag sequences as proxies for the potentially active

microorganisms in the sediments at the time of sampling, which allowed us to investigate the functional potential of the community as well. The findings from iTag sequencing indicated that the microbial community in the sediments was highly diverse (Inverse Simpson's reciprocal index 33.4-148.6), and that it was composed of 20 abundant taxa of Bacteria (≥ 1 % of amplified RNA – aRNA or RNA sequences or more) across three depths surveyed from cores collected in Siq and two sites in Suk (SukB and SukS). These phyla encompassed groups capable of organic matter degradation, sulfate reduction, nitrogen cycling, and photosynthesis. A deeper look into the functional potential of the community using phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) supported these findings. The archaeal assemblage was found to be portioned between lakes, and methanogenic Archaea were mostly found in the sediments of Siq and were almost entirely absent in Suk. To know the status of the microbial community in these of climate-sensitive ecosystems is essential, because it will allow for predictive changes in community structure that could lead to changes in greenhouse gas emissions.

1. Introduction

Arctic regions are among the most climate-sensitive ecosystems on the planet (McGuire et al., 2009; Kittel et al., 2011; Graham et al., 2012), particularly because of the high amount of partially degraded organic matter that is stored in permafrost (frozen ground), both in continental and oceanic territories (Zimov et al., 2006).

Thermokarst lakes occupy an estimated 250,000 to 380,000 km^2 area in the pan-Arctic (Grosse et al., 2013), and are perhaps even more sensitive to warming than permafrost itself. The large heat storage capacity of the water column combined with long term heat flux to the sediments, maintains lake bottoms above freeze temperatures and result in expansion of the thaw bulb underlying the lake bottom (Grosse et al., 2013).

The thaw bulb underneath thermokarst lakes contains organic matter that was stored in permafrost, in addition to fresh organic matter produced in the water column during the long days of summer. These two sources of carbon can fuel microbial processes such as anaerobic respiration, fermentation, and methanogenesis. The sediments of two lakes in the area (Siqlukaq Lake –Siq, and Sukok Lake – Suk) contain 0.2 to 15 wt. % of total organic carbon (TOC), and the lake with higher TOC content (Siq) exhibited a higher potential for biological CH_4 production (Chapter 2; Matheus Carnevali et al., 2015). Biological CH_4 production occurs as a result of these microbial activities in arctic thermokarst lake sediments on the Coastal Plain of Alaska. Therefore, we hypothesized that the microbial community in Siq would be different from the microbial community in SukB and SukS, considering the differences in total organic carbon (TOC) content detected in these sediments.

Microbial communities in arctic rivers, wetlands, permafrost, and permafrost-affected soils have been studied with DNA profiling techniques and clone libraries of the SSU rRNA gene (Høj et al., 2005; Galand et al., 2006; Galand et

al., 2008; Steven et al., 2007; Barbier et al., 2012). Next generation sequencing (NGS) technologies that provide the ability to parallelize the analysis of multiple samples, and to sequence at a higher sequencing depth at a faster pace, are being used more frequently as their cost decreases rapidly. In arctic environments, the use of NGS to study microbial communities has been limited to the Arctic Ocean (Galand et al., 2009; Kirchman et al., 2014), peat soils (Tveit et al., 2013), permafrost (Yergeau et al., 2010), permafrost-affected soils (Frank-Fahle et al., 2014), and wetlands (Serkebaeva et al., 2013; Stoeva et al., 2014). Only recently, Neghandi (2013) studied CH_4 sources in thaw ponds located in the Canadian High Arctic.

Most studies about thermokarst lakes in the North Slope of Alaska have focused on the limnology of the lakes (Howard & Prescott, 1973), on CH_4 fluxes (Morrissey & Livingston, 1992; Phelps et al., 1998), CH_4 cycling (Bretz & Whalen, 2014; Lofton et al., 2014), and other aspects. A few studies have looked at microbial communities in the sediments of thermokarst lakes on the North Slope of Alaska (He et al., 2012a; He et al., 2012b; He et al., 2012c). To know the status of the microbial community in this kind of climate-sensitive ecosystems is of utmost importance, because it allows for informed predictive changes in community structure and function that could lead to changes in greenhouse gas emissions to be estimated. Thus, knowing the microbial community ultimately improves our ability to estimate current and future greenhouse gas emissions in the Arctic. To advance our understanding of these lake ecosystems, we investigated the diversity, composition,

and functional potential of the microbial communities inhabiting two lakes on the Coastal Plain of Alaska.

Our goal was to study the bacterial and archaeal assemblages in the sediments of thermokarst lakes in light of the differences in geochemistry between lakes. A combination of molecular approaches, including MiSeq next-generation sequencing (iTag sequencing) of the small subunit (SSU) ribosomal RNA (rRNA) gene and the SSU rRNA was used. Ribosomal RNA sequences were taken as a proxy for potentially active microorganisms in the sediments at the time of sampling (Putkinen et al., 2009), which allowed us to investigate the functional potential of the community as well. Furthermore, the phylogeny of Archaea in these arctic lakes was determined.

2. Methods

2.1. Sampling sites and samples collection

Artic thermokarst lakes nearby the town of Barrow, Alaska, were sampled during three different field campaigns (April and Oct 2010 and Oct 2011). Sediment cores were collected using a universal percussion corer (Aquatic Research Instruments) and stored frozen at -80 °C until subsampling. Samples for the primary iTag sequencing analyses described in this study were collected in 2011 from Siqlukaq Lake (Siq11; 71° 10.481’ N, 156° 53.910’ W), and from two sites in Sukok Lake: a site nearby an open natural gas seep (SukS11; 71° 04.519’ N, 156° 49.208’ W) and another site ~ 1 km southwest from the seep (Sukok B site - SukB; 71° 04.006’ N,

156° 49.841’ W). Lakes Ikroavik (Ikv10; 71° 14.036’ N, 156° 38.349’ W), Siqlukaq (Siq10; 71° 10.486’ N, 156° 53.891’ W), Sukok (Sukok Seep site – SukS10; 71° 04.455’ N, 156° 49.250’ W), and Qalluuraq Lake (70° 22.672’ N, 157° 20.927' W; 60 miles to the south) had been also sampled in 2010. The 2010 samples were used to survey the archaeal diversity by profiling the V3 region of the SSU rRNA gene with denaturing gradient gel electrophoresis (DGGE), and to determine the phylogeny of Archaea in these lakes.

2.2. Nucleic acid extraction for iTag sequencing of the SSU rRNA gene and SSU rRNA of Bacteria and Archaea

Over the course of this study three DNA extraction protocols (one described in this section and two described in section 2.3) and one RNA extraction protocol were executed. Efforts were made to optimize DNA extraction procedures to ensure amplifiability of the DNA. Extracts were used for different downstream analyses and were not intercompared.

To study the microbial community diversity and structure and the ‘potentially active’ members of the community, genomic DNA and total RNA were extracted from three sections (U-upper, M-middle, L-lower) of the first 30 cm of the 2011 cores (Siq11, SukS11, and SukB11). Sediment subsamples that had been stored at -80 °C and later transferred to sucrose lysis buffer (SLB, 40 mM EDTA, 50 mM Tris-HCl, 0.75 M sucrose) to better preserve the integrity of the nucleic acids, were used to extract community genomic DNA as previously described in (Chapter 2) (Matheus Carnevali et al., 2015). In brief, samples in SLB were thawed on ice for

45 min and centrifuged at 10,000 x g for 10 min. The supernatant was removed; 2 x ~ 0.5 g sediment subsamples from each depth interval (for a total weight of ~ 1 g per sample) were extracted using a Power Soil DNA Isolation Kit (MoBio, Carlsbad, CA, USA) following manufacturer's instructions, and later quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, Grand Island, NY, USA).

To ultimately obtain total RNA, nucleic acids were also extracted from each section of Siq11, SukS11, and SukB11 using a modified RNA PowerSoil Total RNA Isolation Kit (MoBio, Carlsbad, CA, USA). Two grams of sediment (in duplicate) were bead beat with MoBio's proprietary solutions and phenol:chloroform:isoamyl alcohol (pH 6.5-7.0) by vortexing at maximum speed for 15 min. The organic and the aqueous phases were separated by centrifugation at 1,040 x g for 24 min, followed by a secondary precipitation step at 4 °C for 10 min to further purify the aqueous phase, and another centrifugation step under the same conditions. The resulting supernatant was split in 1.7 ml Eppendorf tubes and the extracted nucleic acids were precipitated in isopropanol at room temperature for 30 min, followed by cold centrifugation at 13,000 x g for 15 min, and a clean-up step. This last step involved 70 % ethanol at room temperature, cold centrifugation at 13,000 x g for 15 min, air-drying of the pellets (or using a Speed Vac Vacuum Concentrator at medium temperature), and final suspension of the total RNA in 45-50 µL of nuclease-free water (Life Technologies, Carlsbad, CA, USA). All reagents, disposable supplies and instruments used for this protocol were RNAse-free or treated with RNaseZap

decontamination solution (Life Technologies, Carlsbad, CA, USA). Reagent controls were also processed to create reagent extraction controls.

In order to isolate the total rRNA, nucleic acid extracts were treated two to three times with 4 U RNase-free DNase I (Life Technologies, Carlsbad, CA, USA) at 37 °C in 100 µL reaction volumes for 30 min. The enzyme was later extracted with phenol:chloroforma:isoamyl alcohol (pH 6.5-7.0). Total RNA was precipitated at -20 °C overnight and suspended in 25 µL of nuclease-free water. These extracts were quantified using a Quant-iT RiboGreen RNA Assay Kit (Life Technologies, Carlsbad, CA, USA).

2.3. Reverse transcription of the SSU rRNA gene and RNA amplification

To determine if the total RNA could reverse transcribed to complementary DNA (cDNA) in sufficient amounts for sequencing, reverse transcription of the total RNA was directed with a SSU rRNA-specific primer (806R) (Caporaso et al., 2010). Total RNA was incubated with the dNTPs (Qiagen, Valencia, CA, USA) for 5 min at 65 °C, transferred to 4 °C, and then incubated at 55 °C for 60 min with 200 U of SuperScript III Reverse Transcriptase (Life Technologies, Carlsbad, CA, USA) in a 20 µL reaction, following manufacturer's instructions. Some of the total RNA samples (Siq11-U, -M, and -L) were diluted 1/10 before the reverse transcription step to prevent downstream reactions inhibition, due to presence of humic and fulvic acids in the extracts. Final products were checked on an agarose gel. cDNA was quantified using a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, Grand Island, NY, USA).

Reagent controls were subjected to DNase treatments and reverse transcription with no subsequent amplification.

Only Siq11-U, Siq11-M, SukB11-U, SukB11-M, and SukB11-L samples could be reverse transcribed to cDNA directly from the RNA in sufficient amounts for sequencing. The remaining samples (Siq11-L; SukS11-U, SukS11-M, and SukS11-L) had to be first amplified using the MessageAmp II aRNA Amplification Kit (Life Technologies, Carlsbad, CA, USA) according to manufacturer's instructions. Additionally, RNA obtained from Siq11-U, Siq11-M, SukB11-U, SukB11-M, and SukB11-L samples was amplified for comparisons of cDNA derived from amplified RNA (aRNA) from all samples. The resulting aRNA was quantified using a Quant-iT RiboGreen RNA Assay Kit (Life Technologies, Carlsbad, CA, USA), and reverse transcribed to cDNA with primer 806R (Caporaso et al., 2010). DNA and cDNA from both RNA and aRNA samples were sent to the Joint Genome Institute (JGI, Walnut Creek, CA) for library preparation and paired-end (2 x 250 bp) MiSeq Illumina sequencing of the variable region 4 (V4) of the SSU rRNA using primers 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') following Caporaso et al. (2010).

2.4. DNA extraction for DGGE profiling and cloning of the SSU rRNA gene of Archaea

To profile the SSU rRNA gene of Archaea using DGGE and to make clone libraries, 1 g sediment subsamples from 2010 cores were collected every 10 cm from the top ~ 50 cm of Siq10, the top ~10-50 cm of Ikv10, and the 2-106 cm length of the SukS10

core. Community genomic DNA was extracted using a phenol-chloroform method (Massana, 1997) and quantified with a NanoDrop. To generate a second round of clone libraries of the SSU rRNA gene of Archaea from cores obtained in 2011, DNA was extracted from ~ 1 g sediment approximately every 10 cm of the first 30 cm of each core (Siq11, SukS11, and SukB11). A Power Soil DNA isolation kit (MoBio, Carlsbad, CA, USA) was utilized for genomic DNA extractions following manufacturer's instructions.

2.5. Archaeal PCR and denaturing gradient gel electrophoresis analysis of 2010 samples

In order to get a rapid snapshot of the archaeal assemblage diversity and structure, primers targeting the archaeal SSU rRNA gene: GC_347F (5'-CGCCCGCCGCGCCCCGCG CCCGTCCCGCCGCCCCCGCTCCGGGCGCAGCAGGCGMGAA-3'); and UNIV519R (Muyzer et al., 1993) were used to amplify the variable region 3 (V3) with the following reaction mix: 1X Standard Taq Buffer (New England BioLabs, Ipswich, MA, USA), 2 mM $MgCl_2$, 0.2 mM of each dNTP (Qiagen, Valencia, CA, USA), 0.5 µM of each primer, and 1.25 U Taq DNA polymerase (New England BioLabs, Ipswich, MA, USA) in a final volume of 50 µL. Thermocycler conditions were: 94 °C for 5 min, followed by 10 cycles of denaturation at 94 °C for 45 sec, touch-down primer annealing with the temperature decreasing from 65 °C to 55 °C (1 °C per cycle) for 30 sec, extension at 72 °C for 30 sec; 18 cycles of denaturation at 94 °C for 30 sec, primer annealing at 55 °C for 30 sec, extension at 72 °C for 30 sec, and a final extension step at 72 °C for 10 min. Amplification products were precipitated overnight at -20 °C with two volumes

of absolute ethanol and 1/10 of the volume of 3 M sodium acetate pH 5. After centrifugation in cold for 30 min, the DNA was washed with 70% ethanol, centrifuged for 10-15 min, air-dried, or dried in a DNA speed vac (Savant), and suspended in 18 µl of sterile DI water. Two to seven PCR reaction products were pooled to reach 300 ng of DNA per sample. A 30 to 65 % gradient of denaturants (7 M urea and 40% deionized formamide in 100% concentration of denaturants) was used. Electrophoresis was run at 60 °C for 16 hrs at 62 V (Murray et al., 1996). Analyses of the banding patterns resulting from DGGE were done with the GelCompar II software (Applied Maths).

2.6. Clone libraries construction and analyses

To study the phylogeny of Archaea in these thermokarst lake sediments, 5 µL of each extract obtained in 2010 from the different sediment intervals were pooled for amplification of the SSU rRNA gene using primers 109F (Großkopf et al., 1998) and 915R (Stahl & Amann, 1991) using the following reaction mix: 1X Standard Taq Buffer (New England BioLabs, Ipswich, MA, USA), 2 mM $MgCl_2$, 0.2 mM of each dNTP (Qiagen, Valencia, CA, USA), 0.5 µM of each primer, and 1U Taq DNA polymerase (New England BioLabs, Ipswich, MA, USA) in a final volume of 20 µL. Thermocycler conditions were: 94 °C for 5 min, followed by 28 cycles of denaturation at 94 °C for 1 min, primer annealing at 52 °C for 1 min, extension at 72 °C for 1 min 30 sec, and a final extension step at 72 °C for 6 min.

The amplification products (bp) were gel purified using a QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA), pooled, and cloned using the TOPO TA cloning kit (Life Technologies, Carlsbad, CA, USA) according to manufacturer's recommendations. Plasmids were extracted with the Plasmid Miniprep$_{96}$ Kit (EMD Millipore, Temecula, CA, USA) and screened using a restriction digest as follows: 25 U EcoRI (Life Technologies, Carlsbad, CA, USA), 1X Buffer H (Life Technologies, Carlsbad, CA, USA), 0.1 mg mL^{-1} bovine serum albumin (Promega, Sunnyvale, CA, USA), and 1 µL of plasmid DNA at 37 °C for 1 hr. The enzyme was deactivated at 65 °C for 20 min. Insert size was checked in 1 % agarose gels stained with GelRed (Biotium, Hayward, CA, USA). Plasmids with the desired insert were sequenced in one direction with the ABI BigDye Terminator Cycle Sequencing Ready Reaction Kit V3.1 using vector primer T7 on a ABI3730 DNA Analyzer (Life Technologies, Carlsbad, CA, USA).

In 2011, the SSU rRNA gene was amplified using primers 20F-958R (DeLong, 1992) using the following reaction mix: 1X Ampli Taq Buffer II (Life Technologies, Carlsbad, CA, USA), 2.5 mM $MgCl_2$ (Life Technologies, Carlsbad, CA, USA), 0.2 mM of each dNTP (Qiagen, Valencia, CA, USA), 0.5 µM of each primer, and 1.25 U AmpliTaq DNA polymerase (Life Technologies, Carlsbad, CA, USA) in a final volume of 25 µL. Thermocycler conditions were: 94 °C for 10 min, followed by 35 cycles of denaturation at 94 °C for 1 min, primer annealing at 55 °C for 1 min, extension at 72 °C for 1 min, and a final extension step at 72 °C for 7 min. The amplification products from each depth were gel purified using a QIAquick Gel Extraction Kit (Qiagen,

Valencia, CA, USA), pooled with other amplification products from the same depth, and cloned using the TOPO TA cloning kit (Life Technologies, Carlsbad, CA, USA) according to manufacturer's recommendations. Plasmids were extracted with a Qiagen 3000 BioRobot and a R.E.A.L prep (Qiagen, Valencia, CA, USA), quantified with the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, Carlsbad, CA, USA), then sequenced in one direction with the ABI BigDye Terminator Cycle Sequencing Ready Reaction Kit V3.1 using vector primer T7 on a ABI3730 DNA Analyzer (Life Technologies, Carlsbad, CA, USA).

Only high quality Sanger sequences from both libraries were used for sequence analyses, including trimming with BioEdit version 7.2.5, and chimera check with Decipher (Wright et al., 2012). To study the evolutionary relatedness between thermokarst sediment Archaea and known Archaea lineages, the SSU rRNA gene sequences were aligned with their closest matches from the GenBank, several sequences from related cultivated Archaea, and other relevant sequences from the Silva database. Sequences were aligned against the Archaea variability profile using Silva's SINA aligner (Pruesse et al., 2012). Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6 (Tamura et al., 2013).

2.7. iTag sequence processing and statistical analyses

JGI initially processed the sequences to remove contaminants, barcodes, and primers. The quality of these sequences was evaluated using the FastQC software (Babraham Bioinformatics). Trimmomatic (Bolger et al., 2014) was then used in the

single end mode to filter the low quality sequences, and 3,834,911 (85.8 %) sequences out of 4,469,505 sequences passed the quality filtering step. Sequences were then separated into forward and reverse reads with an in-house script and used to form contigs with Mothur (Kozich et al., 2013), which was also the platform used to remove chimeras, classify the sequences against the Greengenes gg_13_8_99 taxonomy, and to cluster them into operational taxonomic units (OTUs) at a distance of 0.03. These sequences were used to create an OTU database containing OTU abundance, a representative sequence, and the consensus taxonomy (Kozich et al., 2013). Mothur was also used to do alpha-diversity analyses on a resampled OTU table. Beta-diversity analyses were done after bootstrap resampling using R (R Core Development Team, 2015) and methods and custom scripts developed by Herbold et al. (2014). To ensure that the data was free of sequencing error, OTUs with 5 sequences or less among all of the samples (DNA, aRNA, and RNA) were removed prior to resampling. To determine between site similarity for DNA and aRNA sequences, a Mann-Whitney test was used to test pairwise comparisons of the Bray-Curtis distances between sites, and the significant thresholds were defined by the Benjamini-Hochberg procedure with a false discovery rate of 0.05 in each sample. To determine whether the abundance of any aRNA-OTU at any site was greater or less than the abundance of any aRNA-OTU at any site, abundances were log-transformed and compared across sites at the phylum-level (or the class level for Proteobacteria) using a Tukey's test, where phyla and site were interacting factors

and depth was the independent variable. The median *p*-value was obtained by repeating the Tukey's test for all resamplings.

2.8. Functional profiling

To determine the functional potential of the community represented in the aRNA samples, as a proxy for 'potentially active' microorganisms (Putkinen et al., 2009) in the lake sediments at the time of sampling, phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) (Langille et al., 2013) was implemented through the Hunttenhover lab modules in the Galaxy server. Sequences were reclassified with Mothur against the Greengenes gg_13_5_99 taxonomy and the OTU table was formatted to match PICRUSt software requirements. In simple terms, PICRUSt predicts community functional composition by associations of SSU rRNA gene sequence data (corrected by copy number) with reference genomes based on relatedness (80 %). The Kyoto Encyclopedia of Gene and Genomes (KEGG) Orthology (KO) was used in this study to annotate the table of predicted gene family counts per sample. Statistical analyses of differences between predicted gene content was performed using Statistical Analysis of Metagenomic (and other) Profiles, STAMP (Parks & Beiko, 2010; Parks et al., 2014).

3. Results

3.1. α-diversity and β-diversity of the microbial assemblage (Bacteria and Archaea)

A total of 1,213,686 filtered sequences comprising DNA, aRNA and RNA were used to analyze the diversity and structure of the microbial community. Out of this total, 290,351 sequences were unique, with an average of 62,297 ± 7,068 DNA sequences (n=9), an average of 47,783 ± 4,817 aRNA sequences (n=9), and average of 44,783 ± 8,699 RNA sequences (n=5) (Table 3.1). The sequences were clustered into 39,244 operational taxonomic units (OTUs) at a distance of 0.03. DNA sequences formed an average of 4,325 ± 1,116 OTUs, of which 2,928 ± 838 OTUs were doubletons and singletons; the aRNA sequences formed and average of 3,832 ± 1,119 OTUs including 2,434 ± 787 doubletons and singletons; and the RNA sequences formed an average of 4,096 ± 1,043 OTUs including 2,649 ± 959 doubletons and singletons. Therefore, a great majority of the OTUs were comprised of rare sequences.

After 1000 iterations of subsampling to less than the lowest number of sequences (28,000) in a given sample, iTag sequence coverage, richness and diversity were estimated (Table 3.2). Good's coverage (Good, 1953) indicated sampling over 92 %. Simpson's diversity index indicates the probability that any two individuals of the same species would be randomly drawn from a community. The index is inversely proportional to diversity, but the reciprocal of the index rises as the community becomes more even (Magurran, 2003). Therefore, Simpson's reciprocal index (Table 3.2) suggested that the samples were highly diverse (33-132 for DNA, 53-223 for aRNA, and 68-149 for RNA). Diversity among the DNA iTags was similarly high (104-133) in the upper and middle intervals of Siq and SukB, and it decreased with depth. The diversity of RNA iTags was higher in Siq than in SukB,

Table 3.1 iTag sequence statistics. Data presented here was not normalized to correct for iTag pool sizes. Sediment intervals correspond to three sections in the first 30 cm of sediment: upper (U), middle (M), and lower (L).

Lake	Sediment Interval	Number of filtered sequences (1,213,686)			Number of OTUs at 0.03 distance (singletons and doubletons removed)			Number of singleton or doubleton OTUs at 0.03 distance		
		DNA	**aRNA**	**RNA**	**DNA**	**aRNA**	**RNA**	**DNA**	**aRNA**	**RNA**
Siq11	**U**	66,997	**46,989**	51,281	1,669	**1,681**	1,645	4,152	**3,406**	3,954
	M	71,199	**42,585**	48,254	1,766	**1,346**	1,463	3,997	**2,260**	3,064
	LB	67,282	**47,140**	-------	1,443	**1,685**	-------	3,102	**2,961**	-------
SukB11	**U**	63,393	**46,861**	47,983	1,564	**1,251**	1,341	3,204	**2,324**	2,767
	M	63,461	**48,210**	29,397	1,449	**1,279**	1,111	3,061	**2,056**	1,951
	L	66,139	**41,519**	46,050	1,163	**934**	1,673	2,305	**1,072**	1,509
SukS11	**U**	58,683	**47,850**	-------	1,099	**1,440**	-------	2,156	**1,996**	-------
	M	48,521	**50,672**	-------	1,533	**2,045**	-------	2,805	**3,634**	-------
	L	54,999	**58,221**	-------	890	**916**	-------	1,566	**2,198**	-------

and the highest diversity for both was found in the middle interval. Among the aRNA iTags the diversity in SukS-M was the highest of all (223).

For β-diversity analyses, the read depth between different samples was normalized by bootstrap resampling (500 bootstraps). A non-metric multidimensional scaling (NMDS) ordination analysis of Bray-Curtis dissimilarity with 20 resamplings for each sample (Fig. 3.1) revealed the separation between DNA sequences from all the three sites (Fig. 3.1a). However, this pattern changed for the aRNA sequences, because the NMDS analysis (Fig. 3.1b) showed a greater separation between SukB-L and SukS-L and the rest of the samples. In either case, the Siq samples from different depths clustered closely together, and in the case of Siq the aRNA sequences were also clustered more tightly with SukB-U and –M, and with SukS-U and –M. Based on a Mann-Whitney test, DNA sequences from Siq were more similar to SukB than to SukS ($p = 0.0001$, significant in 1000/1000 resamples) and DNA sequences from SukS were more similar to SukB than to Siq ($p = 0.0053$, significant in1000/1000 resamples). The aRNA sequences from the three sites were not significantly different (0/1000 resamples).

3.2. Thermokarst lake sediment bacterial assemblage composition

In chapter 2, the three sampling sites presented different gradients of nitrate, sulfate, dissolved iron, and dissolved manganese concentrations (Matheus Carnevali et al., 2015). We were interested in how these environmental parameters influence the microbial community structure (or viceversa), and what microbial guilds inhabit

Table 3.2 Alpha diversity of samples subsampled to 28,000 sequences per group. Sediment intervals correspond to three sections in the first 30 cm of sediment: upper (U), middle (M), and lower (L). aRNA data is in bold.

Lake	Depth		Coverage	S_{obs}	Simpson's Reciprocal Index	Simpson's Reciprocal Index low confidence interval	Simpson's Reciprocal Index high confidence interval
Siq	**U**	DNA	0.93	3473.1	132.8	129.1	136.7
		aRNA	**0.93**	**3874.5**	**130.8**	**125.8**	**136.2**
		RNA	0.92	3942.4	112.5	107.5	117.9
	M	DNA	0.93	3372.8	112.5	108.9	116.3
		aRNA	**0.95**	**2951.5**	**72.7**	**70.2**	**75.3**
		RNA	0.93	3359.7	148.6	143.8	153.6
	L	DNA	0.94	2784.3	76.5	74.2	78.9
		aRNA	**0.93**	**3586.7**	**129.2**	**125.1**	**133.6**
SukB	**U**	DNA	0.94	3050.1	138.8	134.3	143.5
		aRNA	**0.95**	**2722.8**	**68.9**	**66.5**	**71.3**
		RNA	0.94	3061.5	67.7	65.2	70.3
	M	DNA	0.94	2863.2	104.1	100.7	107.7
		aRNA	**0.95**	**2547.2**	**116.3**	**113.0**	**119.8**
		RNA	0.95	2989.0	117.3	113.8	121.1
	L	DNA	0.96	2185.7	36.0	34.7	37.3
		aRNA	**0.97**	**1714.7**	**53.0**	**51.2**	**54.9**
		RNA	0.96	2666.4	86.4	83.1	90.0
SukS	**U**	DNA	0.96	2193.8	33.5	32.4	34.7
		aRNA	**0.96**	**2722.1**	**86.3**	**83.4**	**89.3**
	M	DNA	0.94	3282.1	72.0	69.6	74.6
		aRNA	**0.92**	**4205.3**	**223.0**	**215.0**	**231.6**
	L	DNA	0.97	1741.0	40.5	39.4	41.7
		aRNA	**0.96**	**2018.3**	**65.5**	**65.8**	**69.4**

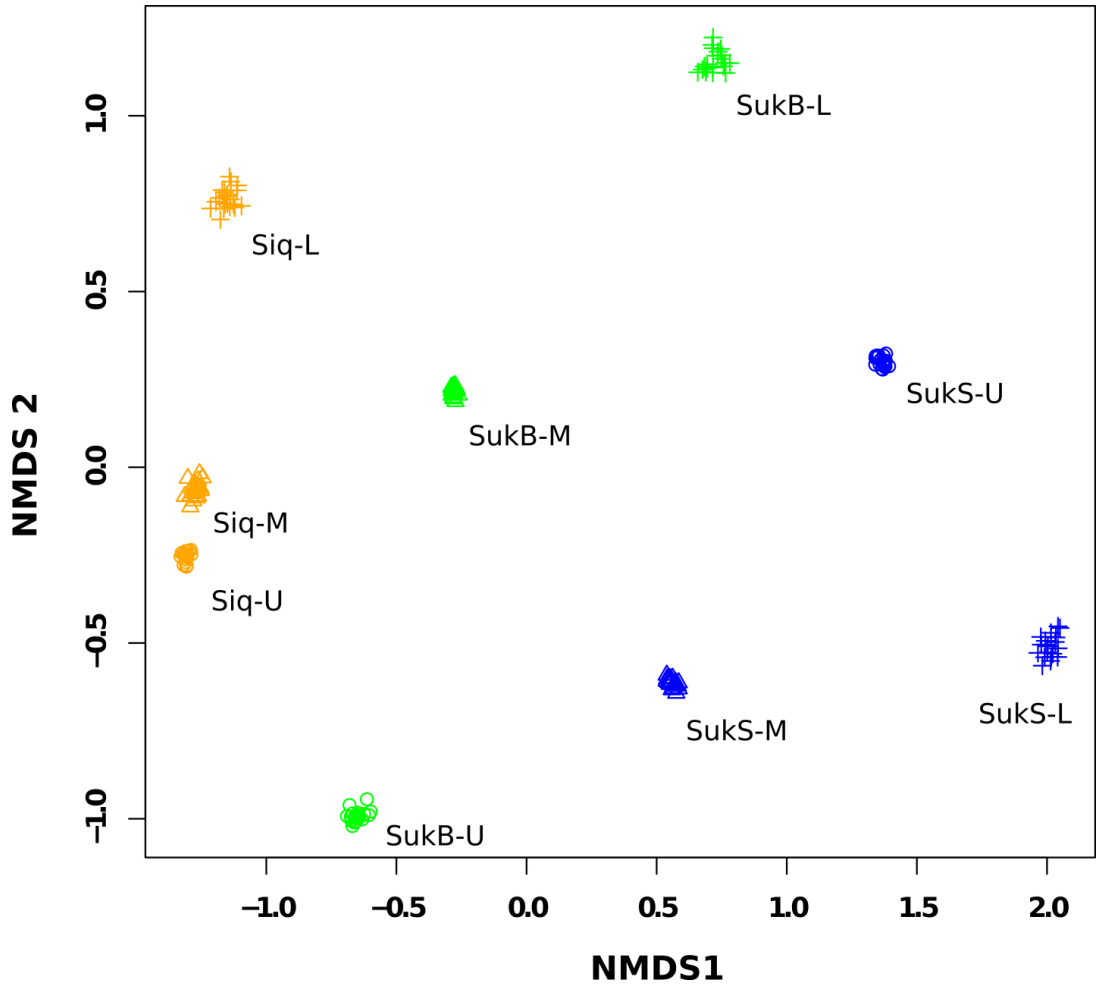


Figure 3.1a Non-metric multidimensional scaling (NMDS) analysis of DNA iTag sequences. Calculated stress (0.08) using Bray-Curtis similarity.

the sediments. We used the aRNA sequences as a proxy for the members of the community that may be potentially active in the sediments. The RNA sequences were also included in the analysis as a reference for the results obtained by RNA amplification, and we observed a relatively high degree of similarity (42-68 %) between these two sets of sequences for paired samples.

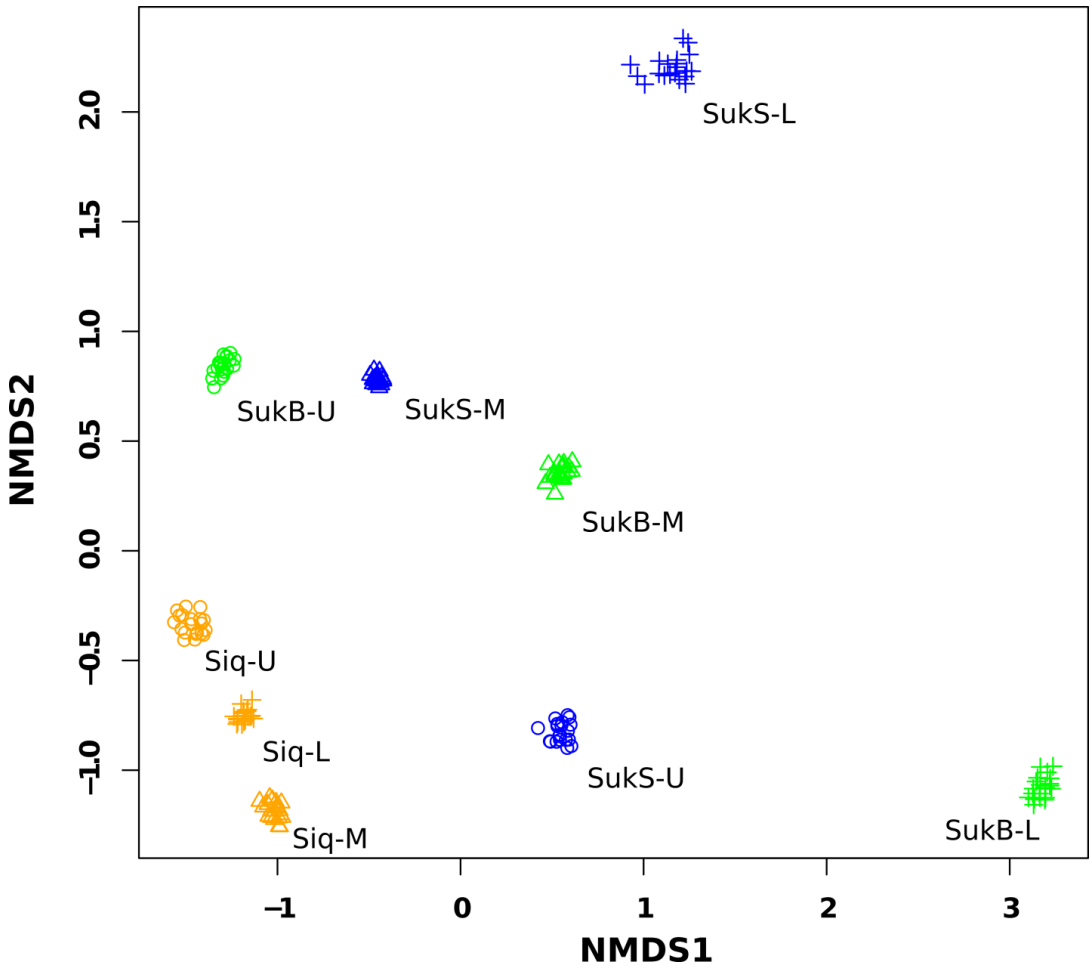


Figure 3.1b Non-metric multidimensional scaling (NMDS) analysis of aRNA iTag sequences. Calculated stress (0.138) using Bray-Curtis similarity.

In the database normalized by bootstrap resampling, aRNA and RNA OTUs comprising 1 % of sequences or more accounted for 85 OTUs and represented at least 15 Bacteria phyla. These phyla included: Acidobacteria (1 OTU), Actinobacteria (11 OTUs), Proteobacteria (34) with the classes Alphaproteobacteria (4 OTUs), Betaproteobacteria (5 OTUs), Deltaproteobacteria (14 OTUs), and Gammaproteobacteria (11 OTUs) represented among the OTUs; Armatimonadetes

(2 OTUs), Bacteroidetes (13 OTUs), Chlorobi (2 OTUs), Chloroflexi (5 OTUs), Cyanobacteria (4 OTUs), Fibrobacteres (1 OTU), Firmicutes (2 OTUs), Nitrospirae (2 OTUs), Aminicenantes (OP8; 1 OTU), Artribacteria (OP9; 1 OTU), Planctomycetes (1 OTU), Verrucomicrobia (3 OTUs), and two unclassified OTUs (Fig. 3.2).

OTUs related to the order Cytophagales in the Bacteroidetes were the most abundant among the three sites, followed by OTUs related to the Deltaproteobacteria. OTUs related to the order Syntrophobacterales and the order Desulfobacterales were more abundant in SukB than in Siq and SukS. Similarly, the Cyanobacteria OTUs of the orders Gloebacterales, Synechococcales, and Chroococcales were more abundant in SukB, as well as OTUs related to the order Pseudomonadales in the Gammaproteobacteria. Nitrospira OTUs were more abundant in Siq than in SukB and SukS, and the SB-45 Artribacteria (OP9) were mostly present in Suk.

Significant differences ($p < 0.05$) in the abundance of phyla and Proteobacteria classes were observed. Log-transformed abundances of aRNA-OTUs across sites were compared using a Tukey's test, where phyla and site were interacting factors and depth was the independent variable. In Siq, Actinobacteria, Bacteroidetes, Deltaproteobacteria, and Gammaproteobacteria were more abundant than Cyanobacteria; in SukB and SukS Deltaproteobacteria were more abundant than Nitrospira; and in SukS Bacteroidetes were also more abundant than Nitrospirae. The major difference between Siq and SukB and SukS was the

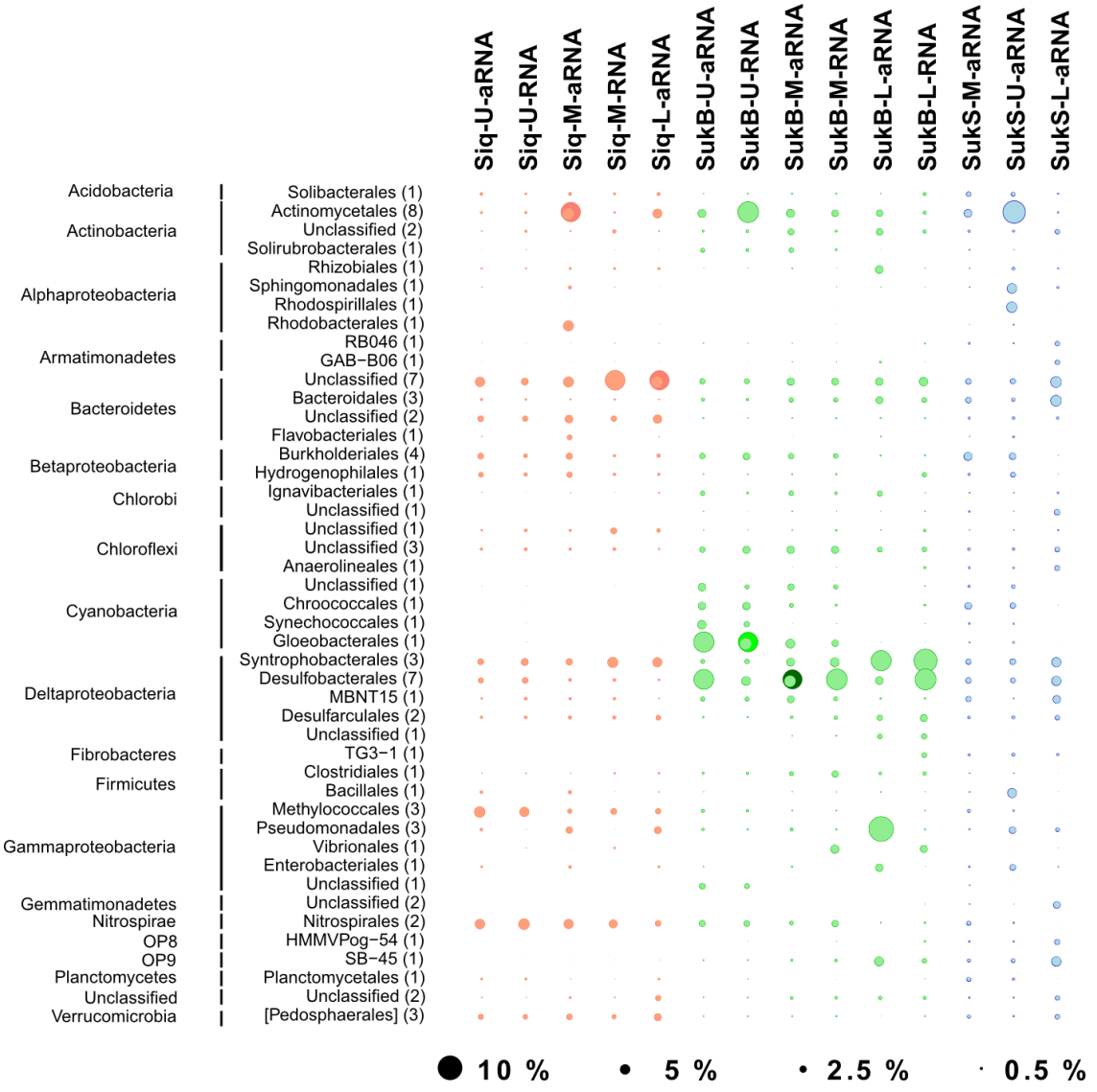


Figure 3.2 Relative abundance of bootstrap resampled OTUs observed with ≥1% relative abundance in either the aRNA or the RNA datasets, representing the potentially active microorganisms. Circle area is proportional to relative abundance (r^2x relative abundance). There are three circles plotted on top of one another: the largest circle is the darkest, and represents the maximum resampled relative abundance (upper 95% confidence interval), the next largest circle represents median and is a moderate shade of color. The smallest circle is also the lightest and represents the minimum relative abundance over the resampling (lower 95% confidence interval). Confidence intervals were close to the median because these are the abundant OTUs. Names on the far left represent Bacteria phyla or class for the Proteobacteria, and are followed by Orders, and number of related OTUs in parentheses.

abundance of Cyanobacteria in SukB and SukS. Conversely, the abundance of Nitrospirae in SukB was lower than the abundance of Bacteroidetes in Siq; and the abundance of Nitrospirae in SukS was lower than the abundance of Deltaproteobacteria in Siq. Therefore, the abundance of Nitrospirae in Siq was a major difference between this lake and the two sites in Suk.

3.3. Thermokarst lake sediment-associated archaeal diversity

Exceedingly divergent between site patterns of archaeal OTUs were evident upon two-dimensional hierarchical clustering (average linkage) analysis based on non-parametric Spearman's rank correlations (Fig. 3.3). Only DNA, aRNA, and RNA OTUs related to Archaea that had 95% or greater confidence of being observed in the bootstrap resampled community are shown in Figure 3.3.

OTUs related to the Miscellaneous Crenarchaeal Group (MCG) were present among the DNA sequences in all three sites, but only among the aRNA sequences in Siq-M, and the RNA sequences in Siq-M, SukB-M and –L. Methanogen-related OTUs dominated Siq OTUs, with OTUs related to Methanosarcinales present among the DNA, aRNA, and RNA sequences at all depths (Fig. 3.3). There were also OTUs related to the order Methanomicrobiales among DNA sequences from all three depths, and among the aRNA or RNA from each depth; and OTUs related to the order Methanobacteriales among DNA sequences from Siq-U and –L, and among Siq-U aRNA and Siq-M RNA. Methanogenic Archaea were absent from most SukB and SukS samples, which were populated by different kinds of Archaea. Notably, the only

methanogen-related OTUs that were present in SukB were found among the DNA sequences.

Other OTUs that could be possibly involved in CH_4 cycling (*i.e.*, potentially in anaerobic CH_4 oxidation) were found among the DNA and aRNA sequences in SukB–L, and were related to ANME-2a (DNA) and ANME-2d (aRNA). Furthermore, among the SukS-U DNA and aRNA; SukB-M and SukS-M aRNA, and SukB-L RNA, some OTUs in the Class Thermoplasmata were related to the family Methanomassiliicoccaceae, which includes methanogenic members (Iino et al., 2013). OTUs related to the putative group, Parvarchaea YLA114 were found among SukS-U DNA and aRNA sequences; and SukS-M, SukB-L, and SukB-M aRNA sequences as well. Among SukS-M DNA and aRNA sequences there were OTUs related to the order Nitrososphaerales in the Class Thaumarchaeota.

Based on the aRNA sequences, Archaeal groups that could have been potentially active at the different sites include OTUs related to *Methanosaeta* in Siq; OTUs related to ANME-2d (*Candidatus Methanoperedens nitroreducens*), and the putative Parvarchaea YLA114in SukB; the Candidate division WCHD3-30 and the Methanomassillicoccaceae in SukB and SukS; and the miscellaneous crenarcheal group clone pGrfC26 at all three sites.

The diversity profile of Archaea obtained by DGGE of the hypervariable region 3 (V3) of the SSU rRNA gene (Fig. S3.1 and S3.2) from Siq10, SukS10, and Ikr10, supported the partition of Archaea among Siq and SukS (Fig. S3.3); and

indicated that there was a high number of archaeal phylotypes per site (18-43). Likewise, phylotypes from the other two lakes (Qaalluraq Lake and Ikroavik Lake) that were sampled as part of the diversity survey carried out in 2010, clustered with Siq (NW in the Fig. S3.1, S3.3) and separate from Suk.

Phylogenetic analysis of the archaeal SSU rRNA gene clone libraries showed that archaeal sequences were closely related to members of the Classes Methanomicrobia and Thermoplasmata in the phylum Euryarchaeota, and to members of the phylum Thaumarchaeota and the Miscellaneous Crenarchaeotic Group (MCG), which forms a separate clade from the Thaumarchaeota and the Crenarchaeota (Kubo et al., 2012). Interestingly, there were three clusters of sequences that appeared to be only related to candidate archaeal phyla or novel uncultured microorganisms (Fig. 3.3). One of these groups clustered with sequences related to the Order Themorplasmaates, Class Thermoplasmata. Another one of these groups was related to sequences in the Class Methanomicrobia, and included matches to members of the so-called Rice Cluster II (Großkopf et al., 1998). The remaining group formed a cluster within the Euryarchaeota, which was separate from the rest. We refer to this cluster as an 'Arctic Cluster' because the closest matches to our sequences were other sequences retrieved from the Arctic (Galand et al., 2006; Galand et al., 2008), but it also included groups previously identified as the Rice Cluster V (Großkopf et al., 1998), the Lake Dagow Sediment cluster (Glissman et al., 2004) and one member of the Candidate division WCHD3-30.

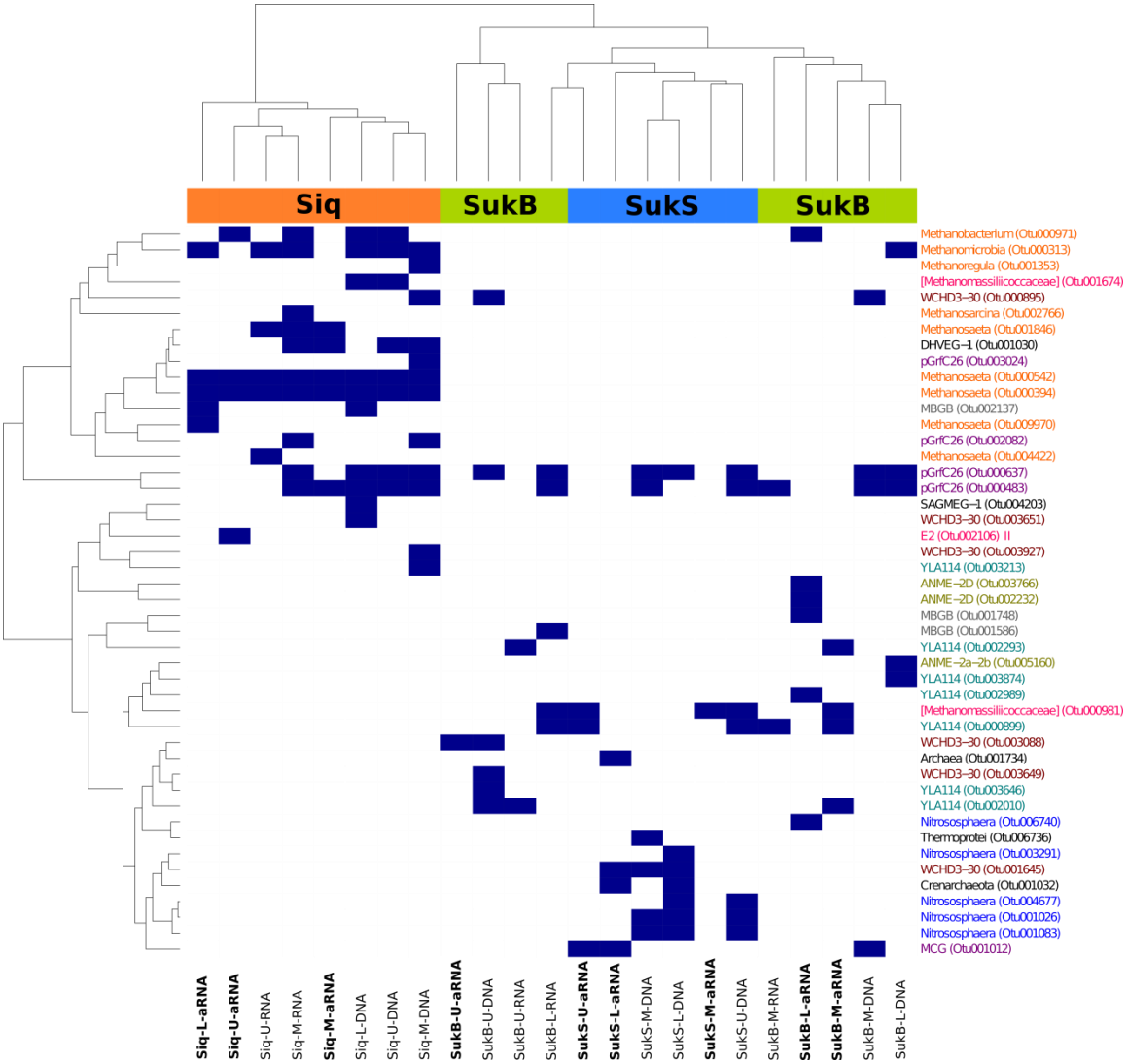


Figure 3.3 Archaeal OTUs that were observed with >95 % confidence over all the resamplings. The abundances were rank-transformed within each resampling and then average rank of OTUs within each sample over all resamplings was used to cluster samples and OTUs together using Spearman's correlation. Distances were calculated as 1-Spearman and clustering is average neighbor. Colors on the right represent the following archaeal groups: genera of Methanogens (orange), Thermoplasmatales (fuchsia), Marine Benthic Group B (MBGB; gray), Parvarchaea clone YLA114 (teal), Anaerobic Methanotrophs (ANME; olive), Miscellaneous Crenarchaeal Group (MCG; purple), Candidate division WCHD3-30 (maroon), Candidatus genus of Thaumarchaeota (blue), other Archaea groups (black).

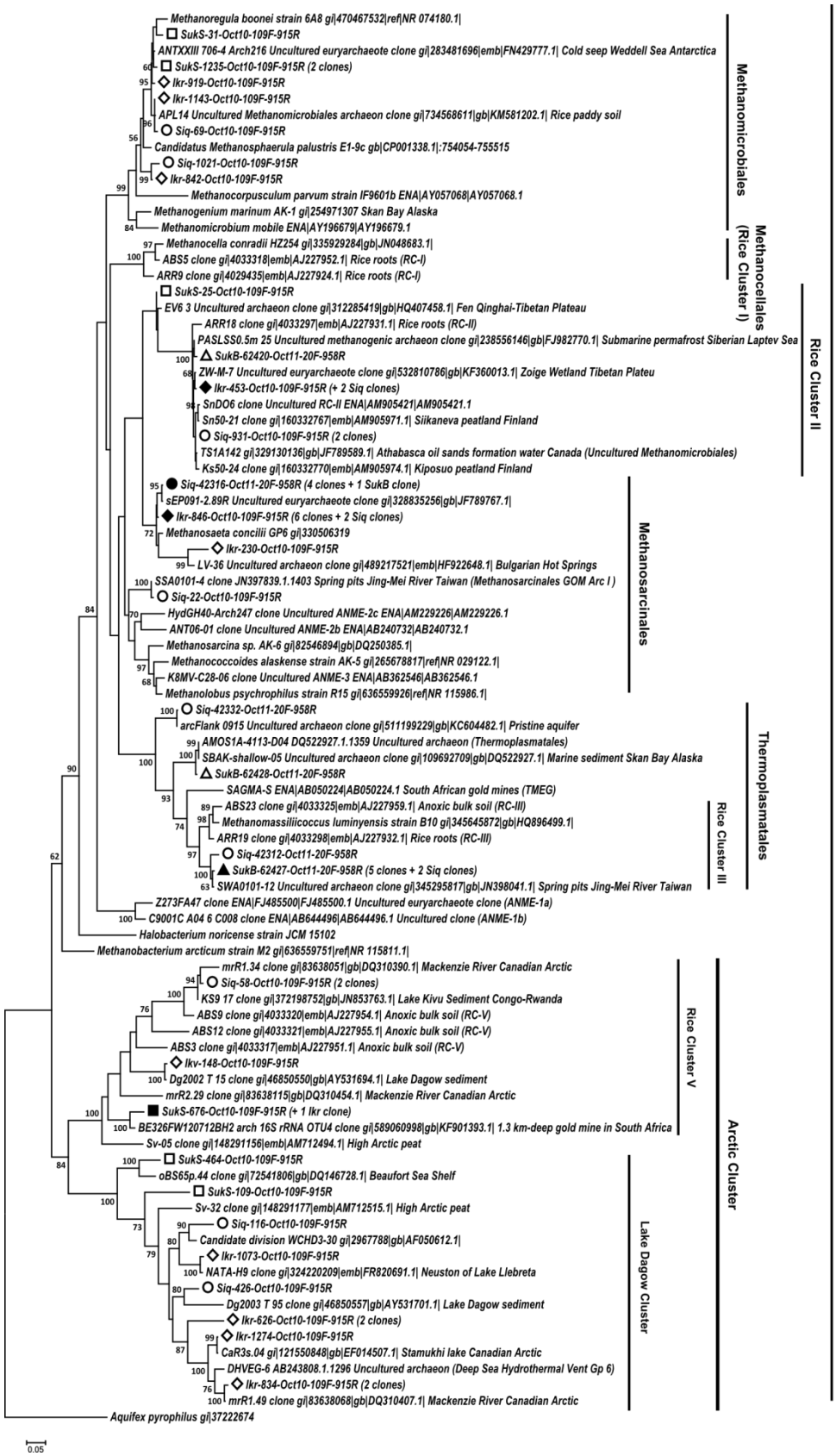

Methanoregula boonei strain 6A8 gi|470467532|ref|NR 074180.1|
SukS-31-Oct10-109F-915R
ANTXXIII 706-4 Arch216 Uncultured euryarchaeote clone gi|283481696|emb|FN429777.1| Cold seep Weddell Sea Antarctica
SukS-1235-Oct10-109F-915R (2 clones)
Ikr-919-Oct10-109F-915R
Ikr-1143-Oct10-109F-915R
APL14 Uncultured Methanomicrobiales archaeon clone gi|734568611|gb|KM581202.1| Rice paddy soil
Siq-69-Oct10-109F-915R
Candidatus Methanosphaerula palustris E1-9c gb|CP001338.1|:754054-755515
Siq-1021-Oct10-109F-915R
Ikr-842-Oct10-109F-915R
Methanocorpusculum parvum strain IF9601b ENA|AY057068|AY057068.1
Methanogenium marinum AK-1 gi|254971307 Skan Bay Alaska
Methanomicrobium mobile ENA|AY196679|AY196679.1
Methanocella conradii HZ254 gi|335929284|gb|JN048683.1|
ABS5 clone gi|4033318|emb|AJ227952.1| Rice roots (RC-I)
ARR9 clone gi|4029435|emb|AJ227924.1| Rice roots (RC-I)
SukS-25-Oct10-109F-915R
EV6 3 Uncultured archaeon clone gi|312285419|gb|HQ407458.1| Fen Qinghai-Tibetan Plateau
ARR18 clone gi|4033297|emb|AJ227931.1| Rice roots (RC-II)
PASLSS0.5m 25 Uncultured methanogenic archaeon clone gi|238556146|gb|FJ982770.1| Submarine permafrost Siberian Laptev Sea
SukB-62420-Oct11-20F-958R
ZW-M-7 Uncultured euryarchaeote clone gi|532810786|gb|KF360013.1| Zoige Wetland Tibetan Plateau
Ikr-453-Oct10-109F-915R (+ 2 Siq clones)
SnDO6 clone Uncultured RC-II ENA|AM905421|AM905421.1
Sn50-21 clone gi|160332767|emb|AM905971.1| Siikaneva peatland Finland
Siq-931-Oct10-109F-915R (2 clones)
TS1A142 gi|329130136|gb|JF789589.1| Athabasca oil sands formation water Canada (Uncultured Methanomicrobiales)
Ks50-24 clone gi|160332770|emb|AM905974.1| Kiposuo peatland Finland
Siq-42316-Oct11-20F-958R (4 clones + 1 SukB clone)
sEP091-2.89R Uncultured euryarchaeote clone gi|328835256|gb|JF789767.1|
Ikr-846-Oct10-109F-915R (6 clones + 2 Siq clones)
Methanosaeta concilii GP6 gi|330506319
Ikr-230-Oct10-109F-915R
LV-36 Uncultured archaeon clone gi|489217521|emb|HF922648.1| Bulgarian Hot Springs
SSA0101-4 clone JN397839.1.1403 Spring pits Jing-Mei River Taiwan (Methanosarcinales GOM Arc I)
Siq-22-Oct10-109F-915R
HydGH40-Arch247 clone Uncultured ANME-2c ENA|AM229226|AM229226.1
ANT06-01 clone Uncultured ANME-2b ENA|AB240732|AB240732.1
Methanosarcina sp. AK-6 gi|82546894|gb|DQ250385.1|
Methanococcoides alaskense strain AK-5 gi|265678817|ref|NR 029122.1|
K8MV-C28-06 clone Uncultured ANME-3 ENA|AB362546|AB362546.1
Methanolobus psychrophilus strain R15 gi|636559926|ref|NR 115986.1|
Siq-42332-Oct11-20F-958R
arcFlank 0915 Uncultured archaeon clone gi|511199229|gb|KC604482.1| Pristine aquifer
AMOS1A-4113-D04 DQ522927.1.1359 Uncultured archaeon (Thermoplasmatales)
SBAK-shallow-05 Uncultured archaeon clone gi|109692709|gb|DQ522927.1| Marine sediment Skan Bay Alaska
SukB-62428-Oct11-20F-958R
SAGMA-S ENA|AB050224|AB050224.1 South African gold mines (TMEG)
ABS23 clone gi|4033325|emb|AJ227959.1| Anoxic bulk soil (RC-III)
Methanomassiliicoccus luminyensis strain B10 gi|345645872|gb|HQ896499.1|
ARR19 clone gi|4033298|emb|AJ227932.1| Rice roots (RC-III)
Siq-42312-Oct11-20F-958R
SukB-62427-Oct11-20F-958R (5 clones + 2 Siq clones)
SWA0101-12 Uncultured archaeon clone gi|345295817|gb|JN398041.1| Spring pits Jing-Mei River Taiwan
Z273FA47 clone ENA|FJ485500|FJ485500.1 Uncultured euryarchaeote clone (ANME-1a)
C9001C A04 6 C008 clone ENA|AB644496|AB644496.1 Uncultured clone (ANME-1b)
Halobacterium noricense strain JCM 15102
Methanobacterium arcticum strain M2 gi|636559751|ref|NR 115811.1|
mrR1.34 clone gi|83638051|gb|DQ310390.1| Mackenzie River Canadian Arctic
Siq-58-Oct10-109F-915R (2 clones)
KS9 17 clone gi|372198752|gb|JN853763.1| Lake Kivu Sediment Congo-Rwanda
ABS9 clone gi|4033320|emb|AJ227954.1| Anoxic bulk soil (RC-V)
ABS12 clone gi|4033321|emb|AJ227955.1| Anoxic bulk soil (RC-V)
ABS3 clone gi|4033317|emb|AJ227951.1| Anoxic bulk soil (RC-V)
Ikv-148-Oct10-109F-915R
Dg2002 T 15 clone gi|46850550|gb|AY531694.1| Lake Dagow sediment
mrR2.29 clone gi|83638115|gb|DQ310454.1| Mackenzie River Canadian Arctic
SukS-676-Oct10-109F-915R (+ 1 Ikr clone)
BE326FW120712BH2 arch 16S rRNA OTU4 clone gi|589060998|gb|KF901393.1| 1.3 km-deep gold mine in South Africa
Sv-05 clone gi|148291156|emb|AM712494.1| High Arctic peat
SukS-464-Oct10-109F-915R
oBS65p.44 clone gi|72541806|gb|DQ146728.1| Beaufort Sea Shelf
SukS-109-Oct10-109F-915R
Sv-32 clone gi|148291177|emb|AM712515.1| High Arctic peat
Siq-116-Oct10-109F-915R
Candidate division WCHD3-30 gi|2967788|gb|AF050612.1|
Ikr-1073-Oct10-109F-915R
NATA-H9 clone gi|324220209|emb|FR820691.1| Neuston of Lake Llebreta
Siq-426-Oct10-109F-915R
Dg2003 T 95 clone gi|46850557|gb|AY531701.1| Lake Dagow sediment
Ikr-626-Oct10-109F-915R (2 clones)
Ikr-1274-Oct10-109F-915R
CaR3s.04 gi|121550848|gb|EF014507.1| Stamukhi lake Canadian Arctic
DHVEG-6 AB243808.1.1296 Uncultured archaeon (Deep Sea Hydrothermal Vent Gp 6)
Ikr-834-Oct10-109F-915R (2 clones)
mrR1.49 clone gi|83638068|gb|DQ310407.1| Mackenzie River Canadian Arctic
Aquifex pyrophilus gi|37222674
0.05
Methanomicrobiales
Methanocellales (Rice Cluster I)
Rice Cluster II
Methanosarcinales
Thermoplasmatales
Rice Cluster III
Rice Cluster V
Arctic Cluster
Lake Dagow Cluster
Euryarchaeota

Figure 3.4a Maximum likelihood phylogenetic tree of SSU rRNA gene Sanger sequences related to Euryarchaeota. Hollow symbols indicate that the given clone represents other clones from the same lake. Filled symbols indicate that the given clone represents other clones from the same lake and from other lakes as well. Lakes are coded by symbol as follows: Siq – circles, SukB – triangles, SukS – squares. Only bootstrap values of 50 or more are shown. The evolutionary history was inferred by using the Maximum Likelihood method based on the Jukes-Cantor model. The tree with the highest log likelihood (0.0000) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 88 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 574 positions in the final dataset. Evolutionary analyses were conducted in MEGA6. Clusters correspond to Groβkopf et al. (1998), Glissman et al. (2004), and Galand et al. (2006; 2008).

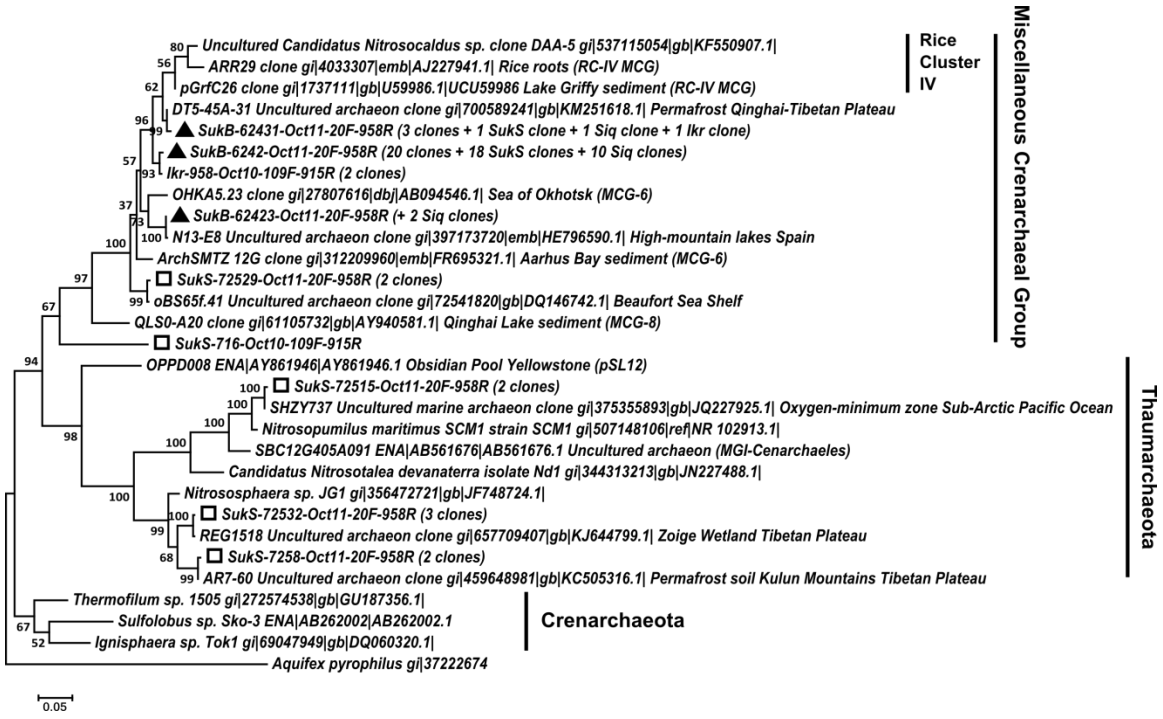


Figure 3.4b Maximum likelihood phylogenetic tree of SSU rRNA gene Sanger sequences related to other Archaea phyla. The analysis was done as described in Figure 3.4a and it involved 30 nucleotide sequences. There were a total of 540 positions in the final dataset.

Sequences from the phylogenetic tree (Fig. 3.4) were also aligned with OTUs related to Archaea among the iTag sequences (Fig. S3.3). Except for those OTUs related to the Crenarchaea Marine Benthic Group B, the phylogeny presented in Figure 3.4 was supported by clustering observed in Fig. S3.3, and the topology of the two trees was very similar to each other.

3.4. Functional potential of the microbial community

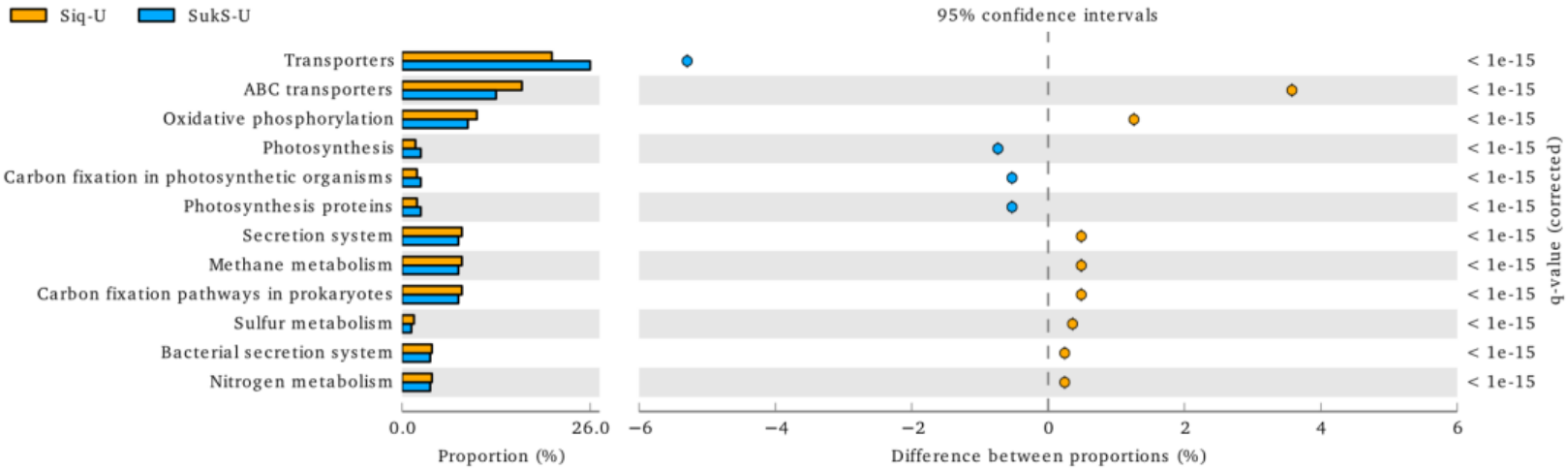
To investigate the functional potential of the microbial community and to look into differences in the potential for organic matter degradation at the different sites we employed a new tool (PICRUSt; Langille et al., 2013), which is based on function inference for related, completely sequenced genomes. Energy metabolism, membrane transport, and signal transduction KO pathways were targeted as a means to identify processes that may be related to the different geochemical environments previously described for each site (Chapter 2; Matheus Carnevali et al., 2015).

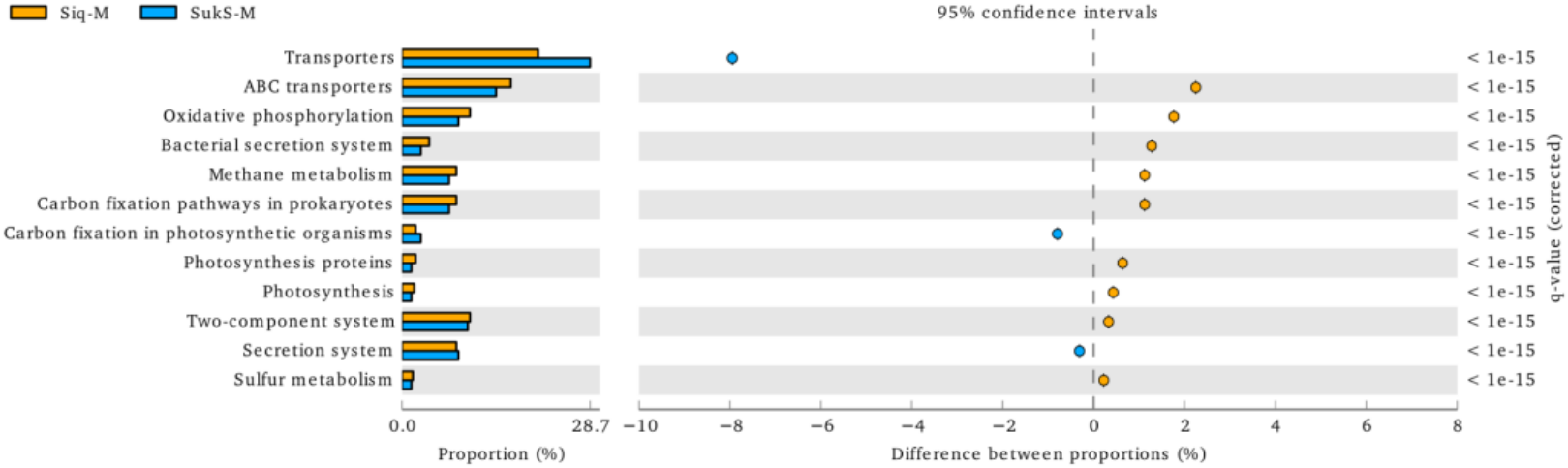
Consistently, predicted gene counts related to transporters, two-component systems, oxidative phosphorylation, and ABC transporters were significantly different between sites in the upper sediment layer (Fig. 3.5). ABC transporter and oxidative phosphorylation predicted gene counts were found in higher proportion (~ 1.5 and ~ 4 % respectively) in Siq, whereas other transporters were more abundant in SukS. A similar trend was observed in the middle sediment layers, except that Siq-M estimated gene counts related to bacterial secretion systems,

oxidative phosphorylation, methane metabolism, and carbon fixation pathways in prokaryotes were also significantly higher in Siq than in SukS-M. At the bottom sediment intervals there were significant differences between proportions in predicted gene counts for secretion systems, CH_4 metabolism, carbon fixation pathways in prokaryotes, nitrogen metabolism, and transporters, all of which were higher in proportion in Siq-B. Oxidative phosphorylation and ABC transporters predicted gene counts were more abundant in SukS-L.

Differences in proportions between Siq and SukB were reflected in estimated gene counts related to transporters (~ 6-8 % higher in SukB-U and -M), two-component systems (~2.5 % higher in Siq-U), and oxidative phosphorylation (~1-2.5 % higher in Siq-U and -M). In the bottom sediment interval, secretion systems, CH_4 metabolism, carbon fixation pathways, and other transporters were ~ 1-3 % higher in Siq-L than in SukB-L, whereas oxidative phosphorylation, ABC transporters, and two-component systems were ~ 1-3 % higher in SukB-L than in Siq-B (Fig. S3.5). Methane, nitrogen, and sulfur metabolism estimated gene counts were generally higher in proportion (~ 0.2-2 %) in Siq than in SukB, except that in SukB-L sulfur metabolism was higher in proportion (~ 0.2 %) than in Siq. Photosynthesis-related gene counts were also higher in proportion in SukB-L.

The most notable differences in proportion between sites within Sukok Lake (Fig. S3.6) were in estimated gene counts related to ABC transporters and other transporters, which were higher in proportion (~ 2-3 %) in SukB-U, and higher in proportion (~ 1-2.5 %) in SukS-M. Two-component system estimated gene counts


Siq-U
SukS-U
95% confidence intervals
Transporters
ABC transporters
Oxidative phosphorylation
Photosynthesis
Carbon fixation in photosynthetic organisms
Photosynthesis proteins
Secretion system
Methane metabolism
Carbon fixation pathways in prokaryotes
Sulfur metabolism
Bacterial secretion system
Nitrogen metabolism
< 1e-15
q-value (corrected)
0.0
26.0
Proportion (%)
−6
−4
−2
0
2
4
6
Difference between proportions (%)


Siq-M
SukS-M
95% confidence intervals
Transporters
ABC transporters
Oxidative phosphorylation
Bacterial secretion system
Methane metabolism
Carbon fixation pathways in prokaryotes
Carbon fixation in photosynthetic organisms
Photosynthesis proteins
Photosynthesis
Two-component system
Secretion system
Sulfur metabolism
< 1e-15
q-value (corrected)
0.0
28.7
Proportion (%)
−10
−8
−6
−4
−2
0
2
4
6
8
Difference between proportions (%)

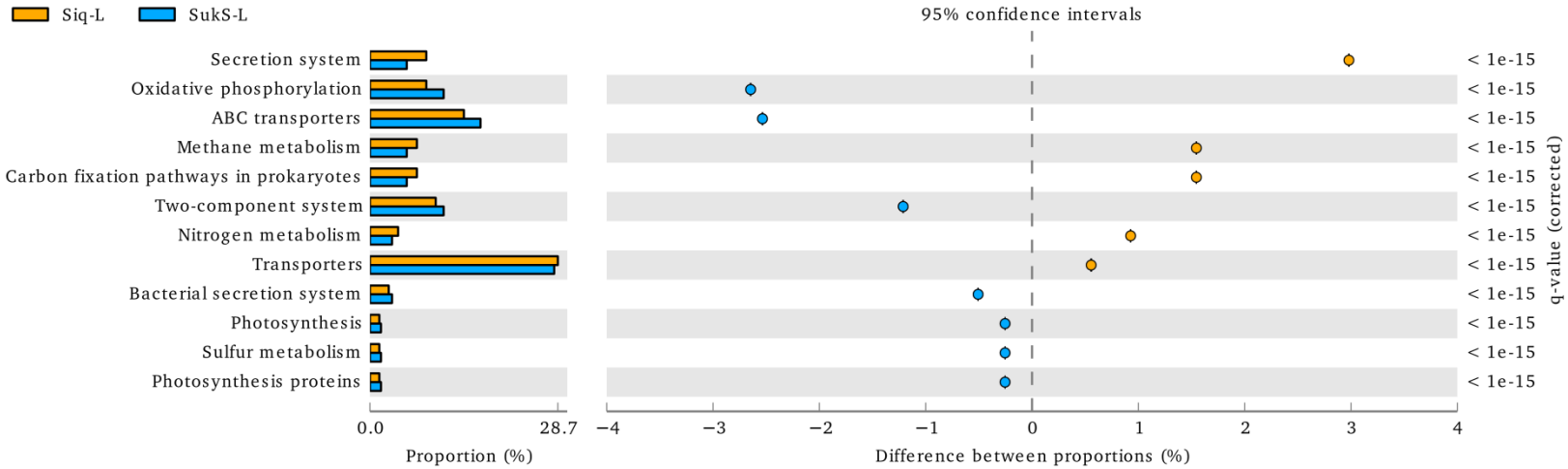


Figure 3.5 Extended error bar plots for the proportion of Siq and SukS sequences related to energy metabolism and environmental information processing KEGG categories filtered by effect size (> 0.2 %). KEGG Orthologs were ordered according to the difference between proportions (note the different scales). The Benjamini-Hocherg's FDR approach was used as a multiple test correction at a false discovery rate of 0.05 % (*q*-value > 0.05 %).

were higher in proportion (~ 2.5 %) in SukS-U; and photosynthesis proteins, bacterial secretion systems, methane metabolism, and carbon fixation pathways in prokaryotes were ~ 1 % higher in proportion in SukB-M. Nitrogen metabolism estimated gene counts were slightly higher in proportion in SukB-L.

4. Discussion

To fully understand the functioning of ecosystems at risk of undergoing dramatic changes due to climate warming, it is necessary to know the diversity and structure of the microbial communities that inhabit them. Yet, in the case of arctic thermokarst lakes, which are sensitive to the warming trend in the Arctic (Osterkamp, 2005; McGuire et al., 2009; Thienpont et al., 2013) very few studies have been conducted. Our work aimed (i) to understand how different geochemical environments in two arctic thermokarst lake sediments on the Coastal Plain of Alaska influence the diversity and composition of the microbial community (and vice versa) and, (ii) to identify what members of the community were potentially involved in organic matter degradation processes that ultimately lead to methanogenesis

4.1. Bacterial assemblage diversity, composition, and function

Based on an NMDS analysis of the Bray-Curtis dissimilarity between DNA sequences determined by iTag sequencing (DNA iTags) (Fig. 3.1a), the OTU abundance of the subsampled microbial community was different between sites, particularly between

Siq and SukS. Furthermore, analysis of aRNA iTags (Fig. 3.1b) suggested that the portion of the community that was potentially active in the sediments was more similar between depths within each site than between sites. These results were also supported by hierarchical clustering of archaeal SSU rRNA gene bands analyzed by DGGE in 2010 (Fig. S3.3). The differences in geochemical profiles (Chapter 2) may drive the differences observed here between sites, while at the same time the interaction between the microbial community and the environment may be altering these differences.

Microbial diversity as estimated by Simpson's reciprocal index was high in comparison to other environments (*e.g.,* snowpacks in Chapter 3, where the highest value for this index estimated for DNA was 19), particularly in Siq and SukB. Diversity among the DNA iTag sequences decreased with depth, following the trend of decrease in TOC with depth in Siq and the low amounts of TOC present at higher depths in all three sites (Chapter 2). The high levels of diversity observed among the aRNA iTags in SukS-M were surprising, and could perhaps be explained by a more efficient amplification process of the total RNA extracted from this sediment interval. The lower amounts of TOC present at this site suggest a lower presence of humic and fulvic acids that may interfere with the amplification reaction.

Given that we were interested in bacterial OTUs that may be actively involved in organic matter degradation processes, we used the aRNA sequences to indicate potentially active bacteria and to study the functional potential of the community. The ribosome to cell ratio is proportional to growth rates, and RNA

molecules are less stable than DNA when exposed to the environment, therefore rRNA is considered a reflection of active populations (Putkinen et al., 2009). Good correlation ($r > 0.8$) has also been found between distributions of SSU rRNA sequences, metatranscriptomes, and metagenomes from permafrost soils (Hultman et al., 2015). In view of the similarity between the aRNA and RNA sequences (~ 42-68 % Bray-Curtis similarity), especially in the upper sediment intervals, we looked at the most frequently occurring OTUs in the aRNA and RNA data sets (Fig. 3.2). In only in a few cases the aRNA may have overrepresented a particular OTU as a result of the linear amplification process. The aRNA and RNA OTUs representing potentially active microorganisms in the sediments were related to three main bacterial phyla: Actinobacteria, Bacteroidetes, and Proteobacteria of the classes Deltaproteobacteria and Gammaproteobacteria. At least two of these phyla have been associated to organic matter degradation processes (Madigan et al., 2014).

Though in different proportions and based on different experimental approaches, a similar composition of bacteria (including all the taxa related to OTUs comprising 1 % of sequences or more) has been observed in other arctic environments: permafrost, active layer and a thermokarst bog around Fairbanks, AK (Hultman et al., 2015); permafrost-affected polygonal soil in the Canadian Arctic (Frank-Fahle et al., 2014); permafrost-affected soils of Northeast Greenland (Ganzert et al., 2014); peat layers of a wetland in European North Russia (Serkebaeva et al., 2013); High-Arctic peat soils in Svalbard (Tveit et al., 2013); active layer and permafrost from a wetland in the Canadian High Arctic (Wilhelm et

al., 2011); permafrost from Nunavut, Canada (Steven et al., 2007); and most importantly in Toolik Lake on the North Slope of Alaska (Crump et al., 2012). Similar Archaea composition has also been found in the freshwater lakes of the Yunnan Plateau (Zhang et al., 2015).

Among the OTUs that comprised 1 % of the aRNA or RNA sequences or more, the Bacteroidetes were dominant members in the Siq bacterial assemblage, and may include obligate anaerobic, fermentative, and saccharolytic bacteria that can ferment sugars and proteins, and make acetate and succinate (Madigan et al., 2014). Other members of this group specialize in the degradation of complex polysaccharides like cellulose (*e.g.*, *Cytophagas sp.*), though others are obligate aerobes with limited fermentative capabilities (Madigan et al., 2014). Members of this phylum have also been found to reduce NO_3^-, degrade starch, and to be psychrophilic or psychrotolerant (*e.g.*, many members of the order Flavobacteriales) (Madigan et al., 2014).

Another phylum that was common to all sites and that could potentially participate in organic matter degradation processes is the Actinobacteria, which includes members that are facultative anaerobes and are able to degrade cellulose, lignin and chitin (*e.g., Actinomyces sp.*) (Madigan et al., 2014). Likewise, the phylum Verrucomicrobia and the Class Betaprotebacteria include chytinolytic species (*e.g., Chitinomonas sp.*).Together, members of these bacterial groups may lead the first steps of organic matter decomposition in artic lake sediments.

Other abundant OTUs that were commonly found in all three sites were related to the Deltaproteobacteria (Fig.3.2). Members of this class often have defined biogeochemical characteristics that are phylogenetically constrained, and mostly involved in sulfate reduction. This is an important finding, because there was at least 50 times more sulfate in upper SukB and SukS sediments than in Siq sediments (Chapter 2), and the highest aRNA Deltaproteobacteria signal that were potentially active were found in SukB. In SukB-L, members of the order Syntrophobacterales (Deltaproteobacteria) comprised ~ 10 % of the aRNA and RNA sequences. In Siq, the Syntrophobacterales comprised 2.5 – 5 % of the sequences, and OTUs related to the genus *Desulfobacca*, which is an acetate-degrading sulfate-reducing bacterium (Oude Elferink et al., 1999) were most common. The presence of this group in Siq, may explain the trend of decrease in SO_4^{-2} concentration with depth in the upper sediment intervals in Siq.

Among the Syntrophobacterales there are species capable of fermenting pyruvate, fumarate, and malate (Madigan et al., 2014), *i.e.* they are able to carry out the final fermentative steps that precede methanogenesis, and could help explain the findings from Siq regarding CH_4 as well. There are also species that are commonly found in syntrophic association with methanogenic Archaea in the deep biosphere (*e.g., Smithella sp.*), and participate in methanogenic degradation of petroleum hydrocarbons (Kallmeyer & Wagner, 2014). Other species have been found to oxidize propionate to make acetate, CO_2, and H_2, and can also live in syntrophy with H_2-consuming bacteria (Madigan et al., 2014). Yet, other species

(*e.g., Syntrophomonas wolfei*) can oxidize butyrate to produce either H_2 or formate, both substrates for hydrogenotrophic methanogenesis. In our previous work with these sediments, we detected a small petroleum contribution in the organic matter of Suk (Matheus Carnevali et al., 2015). Petroleum is usually composed of various kinds of hydrocarbons (*e.g.*, alkanes such as ethane, propane, and butane) with different degrees of susceptibility to bacterial degradation, a process that can be affected by low temperature and binding of hydrocarbons to other soil components (Barathi & Vasudevan, 2001; Olajire & Essien, 2014). These factors may affect the ability of microorganisms like the Syntrophobacterales-related OTUs to degrade organic matter. This may help explain the lag phase observed prior to biological CH_4 production in SukB sediments under experimental conditions, as well as the lower amount of CH_4 produced, and what looks like the lack of utilization of the high amounts of SO_4^{2-} available in SukB sediments.

Another group that was abundant in SukB and SukS (2.5-10 % of the sequences) was Desulfobacterales-related OTUs. Sulfate-reducing bacteria (SRB) in this group were related to members of the families Desulfobacteraceae, Desulfobulbaceae, and Desulfarculaceae. Some species (*e.g., Desulfococcus oleovorans*) are capable of utilizing C_{12}-C_{20} alkanes as growth substrates, producing fatty acids (Aeckersberg et al., 1998; So et al., 2003), and others like members of the genus *Desulfobulbus* are considered "power cords" (Pfeffer et al., 2012). Most importantly, this last genus has been found in association with anaerobic methanotrophs (ANME) (Orphan et al., 2001; Orphan et al., 2002; Holmes et al.,

2004). These bacteria can use alternative electron acceptors like iron (Holmes et al., 2004; Beal et al., 2009), which was found in high concentrations in the sediments of SukB and Siq (Chapter 2), suggesting that in these sediments these microorganisms may be involved in iron reduction, and could potentially support CH_4 oxidation this way. The trend of SO_4^{2-} consumption in SukS remains to be explained, as well as the possible mechanisms for its formation by other members of the community (*e.g.,* green sulfur bacteria in the phylum Chlorobi) in both SukB and SukS.

Other notable groups of bacteria that were abundant in specific samples like SukS-U include the Sphingomonadales (2.5 % of sequences), some of which can metabolize a broad range of organic and aromatic compounds; and the Rhodospiralles (2.5 % of sequences), which are a metabolically versatile group of anoxygenic phototrophs, N_2-fixing bacteria, denitrifiers, methylotrophs and acetic acid bacteria (Madigan et al., 2014). In addition, Pseudomonadales-related OTUs in SukB-L (10 % of sequences) are known chemoorganotrophs that can use NO_3^- for anaerobic respiration (Madigan et al., 2014), and several can perform complete denitrification.

Among the OTUs with the highest number of sequences in Siq (~ 5 %), those related to the phylum Nitrospirae and to the order Methylococcales (2.5-5 % of the sequences) stood out. Some members of the phylum Nitrospirae are known for the autotrophic oxidation of NO_2^- to NO_3^-. This group of bacteria, together with Rhodospiralles- and Pseudomonadales-related OTUs, may be involved in the cycling of N detected in all three sites. The Methylococcales are a well-known group of type

I methanotrophs (Madigan et al., 2014) and may be indicative of mixing of lake water and sediments, because they typically perform aerobic CH_4 oxidation and were also found in abundance in the water column (Priscu, *unpublished*). In SukB, the presence of unicellular Cyanobacteria of the orders Gloebacterales, Synechococcales, and Chroococcales was also surprising. These bacteria are oxygenic photoautotrophs and require sunlight as their main source of energy.

4.2. Arctic thermokarst lake Archaea in the North Slope of Alaska

Given the essential role that methanogenic Archaea play in biological CH_4 production, which was detected in the sediments of Siq and SukB (Chapter 2), we investigated the previously unknown structure, composition, and phylogeny of Archaea in the sediments of these thermokarst lakes. Based on iTag sequencing and DGGE, different groups of Archaea were distributed between these arctic lakes. OTUs related to methanogens were mostly found in Siq, and OTUs related to other groups of Archaea were mostly found in Suk. This partition of methanogens is highly consistent with the differences in biological CH_4 production observed in these lakes.

Among the groups found in SukB there were aRNA OTUs related to *Candidatus Methanoperedens nitroreducens* (ANME-2d), and Parvarchaea YLA114. Microorganisms in the ANME-2d group have been found to use NO_3^- as the terminal electron acceptor in the anaerobic oxidation of CH_4 in consortia with anaerobic ammonium oxidizing bacteria (Haroon et al., 2013). Sequences that were related to the Parvarchaea YLA114, an environmental sequence retrieved from Yellowstone

Lake (Kan et al., 2011) were part of a small group of sequences that does not have cultured representatives in the databases, thus supporting the novelty of this group. Parvarchaeota are part of an archaeal superphylum that comprises Archaea of small cell sizes and genomes (Rinke et al., 2013; Castelle et al., 2015; Baker et al., 2010). The only group that was consistently found in all three sites despite their biogeochemical differences was composed of sequences related to the MCG, perhaps supporting the idea that these Archaea are metabolically diverse (Kubo et al., 2012).

To study the phylogeny of Archaea from these unexplored arctic lake sediments, we cloned ~800-900 bp fragments of the SSU rRNA genes using primers that targeted Archaea and that spanned overlapping regions of the gene (either 109F-915R or 20F-958R). Most of the sequences in the clone libraries were related to the MCG described by Kubo et al. (2012) as a separate clade from the Thaumarchaeota and the Crenarchaeota. The MCG is a very diverse (76% - 87.7% intragroup similarity of the SSU rRNA gene) and ubiquitous group of Archaea that includes clones isolated from freshwater environments (*e.g.*, pRfc26), deep sea hydrothermal vents, mud volcanoes, methane hydrates in marine sediments, among several others. They were thought to be involved in CH_4 cycling, but new evidence suggests that does not seem to be the case (Kubo et al., 2012), and that they thrive in low energy subsurface environments. Further, the archaeal assemblage in sediments with abundant MCG has been shown to use fossil organic matter as a carbon source (Biddle et al., 2006). The high number of sequences that were related to this group may be a result of primer bias (20F-958R) when compared to Archaea

detected with the other primer set (109F-915R). However, their presence was supported by the iTaq sequences (Fig. S3.4), and some of the closest relatives in the GenBank and the Silva databases were obtained in geographical regions that are also cold environments (*e.g.*, other arctic environments, the Tibetan Plateau, wetlands and other freshwater bodies, and permafrost environments).

Our findings also corresponded with the phylogeny of Archaea described by Großkopf et al. (1998) and Glissman et al. (2004), whose classification system has been used as a reference by other authors studying permafrost and water bodies in the Arctic. Sequences related to the Euryarchaeota in our clone libraries were related to methanogens in the Class Methanomicrobia, mostly with members of the family Methanoregulaceae in the Methanomicrobiales, and the family Methanosaetacea in the Methanosarcinales. The remaining sequences formed clusters with uncultured members of the Euryarchaeota that have been previously identified as Rice Clusters I-V (Großkopf et al., 1998). Rice Cluster II (RC-II) also included sequences from sediments in the Tibetan Plateau, biodegraded soil from a Canadian reservoir, and submarine permafrost from the Laptev Sea. Sequences retrieved from polygonal ponds of northeastern Canada (Negandhi et al., 2013) and Finnish fen-type peatlands (Putkinen et al., 2009) also fall in this cluster. Other sequences were related to the Rice Cluster III (RC-III) (Großkopf et al., 1998) and belonged to a different group of Euryarchaeota, the family Methanomasiilicoccaceae in the Thermoplasmatales, which includes a new group of methanogens that reduces methanol with H_2 (Dridi et al., 2012).

A cluster composed of uncultivated sequences only, which we refer to here as an 'Arctic Cluster', comprised two groups of sequences: i) sequences mostly related to the Rice Cluster V (RC-V), many of which have been reported from Arctic habitats (Großkopf et al., 1998; Galand et al., 2006; Galand et al., 2008; Høj et al., 2008), and ii) sequences retrieved from an arctic river, a stamukhi lake, and the Arctic Ocean (Galand et al., 2006; Galand et al., 2008), Lake Dagow Sediment (LDS) in Germany (Glissman et al., 2004), and the Candidate division WCHD3-30, which fall into the LDS cluster (Galand et al., 2006; Høj et al., 2008). Sequences in this cluster were also related to sequences from deep sea hydrothermal vents (DHVEG, Deep Sea Hydrothermal Vent Euryarchaeal Group-6) (Takai & Horikoshi, 1999), in correspondence with sequences related to Parvarchaea YLA114 (Kan et al., 2011) and findings by Galand et al. (2006). In parallel to the last study, the diversity of Archaea in these arctic lakes was unusually high (Fig. 3.3, Fig. S3.1 and S3.2), in which many phylotypes were apparent upon DGGE analysis of amplified DNA, archaeal SSU rRNA gene sequences and iTags. Sequences in this cluster, though common to a variety of arctic environments, are cosmopolitan as well, yet their functional significance remains unknown.

4.3. Functional potential of the microbial community

We also studied the functional potential of the microbial community in a separate analysis that involved the use of predicted metagenomes that may be associated to the SSU rRNA sequences we obtained, and the abundance of these sequences was a proxy for the abundance of the KEGG orthologs that may be expressed by the

community. A high proportion of predicted transporter genes and oxidative phosphorylation genes were observed at all sites, followed by two-component signal transduction systems, secretion system, methane metabolism, and carbon fixation genes. The prominent proportion of predicted transporter genes found in this study correlated with recent findings of abundant protein transporters in the proteome of a thermokarst bog (Hultman et al., 2015). Predicted nitrogen metabolism (ammonia oxidation, nitrite reduction, nitrate reduction, nitric oxide reduction, and nitrogen fixation), photosynthesis, and sulfur metabolism (sulfite reduction and sulfate transport) genes were observed in a minor proportion.

Significant differences in proportion (> 0.2 %) between Siq-M/B and SukS-M/B (Fig. 3.5), and SukB-M and SukS-M (Fig. S3.5) were observed for CH_4 metabolism, supporting previous observations of major differences in biological CH_4 production potential between the three sites. Furthermore, the proportion of predicted genes related to CH_4 metabolism (~ 9 %) was usually greater than the proportion of predicted genes associated to nitrogen and sulfur metabolisms, except for the upper sediment layers of both Suk sites. There was a significant difference (*q-value* < 0.0001) between proportions of predicted nitrogen metabolism genes between Siq-L (> 1 %) and SukS-L, and between SukB-L and SukS-L (Fig. S3.6), suggesting that this kind of energy metabolism may be more active in Siq and SukB. Predicted sulfur metabolism genes were consistently present at all three sites.

These differences in predicted gene proportion also reflect the observed distribution of Archaea in the lakes, where the majority of methanogen-related

OTUs were found in Siq (Fig. 3.3), and the majority of OTUs related to archaeal ammonia oxidizers of the phylum Thaumarchaeota were found in Suk. Conversely, the abundance of bacterial OTUs (Fig. 3.3) related to the phylum Nitrospirae in Siq are indicative of the occurrence of nitrite oxidation to NO_3^- uncoupled to ammonia oxidizers. The higher abundance of SRB in SukB may be explained by the higher concentrations of SO_4^{2-} available there.

Even though the abundance of Cyanobacteria detected in Suk was significantly higher than the abundance of Cyanobacteria in Siq, there was only a minor difference in proportion (< 0.5 %) in predicted genes related to photosynthesis and carbon fixation in photosynthetic organisms. The presence of a substantial Cyanobacteria signal is a puzzling finding that warrants further interrogation, and their presence in the sediments may be explained by mixing with the water column. Alternatively, there may be potentially long turnover times resulting in preservation of cyanobacterial RNA molecules.

Together, our findings indicate that the composition of the microbial community inhabiting these arctic thermokarst lakes may be influenced by the geochemical environment that characterizes the sediments of each lake. Likewise, the microbial community may be altering the cycling of nutrients in these lakes, as shown by the abundance of methanogens in Siq sediments and the higher potential for biological CH_4 production observed in this lake. However, further studies are necessary to establish the degree of correlation with the environment. Overall, the bacterial assemblage has the potential to efficiently degrade organic matter of mixed

composition, and together with the Archaea, to contribute to the cycling of nitrogen, sulfur, and CH_4. Finally, the phylogeny of Archaea indicates the presence of novel groups of Archaea that merit further investigation.

Acknowledgements

Special thanks to Megan Rorhssen, Pamela Santibañez, Dan Berisford, Heather Adams, John Priscu, Andy Klesh, John Leichty, and the Barrow Arctic Science Consortium for assistance in the field. We are very grateful to Olivia Rassuchine, Vivian Peng, Ema Kuhn, and Gareth Trubl for technical assistance in the lab; and to John Mejia for writing scripts for iTag sequence processing. We specially thank the Joint Genomic Institute for sequencing effort. iTag sequencing was supported by an award to AE Murray, JGI-634 by the DOE Joint Genome Institute. PMC was supported in part by the Division of Earth and Ecosystem Sciences, DRI. KPH acknowledges support through the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). Financial support for this research was provided by the NASA Astrobiology Institute, Astrobiology of Icy Worlds program at JPL.

References

Aeckersberg F, Rainey FA, Widdel F (1998) Growth, natural relationships, cellular fatty acids and metabolic adaptation of sulfate-reducing bacteria that utilize long-chain alkanes under anoxic conditions. *Archives of Microbiology,* **170**, 361-369.

Baker BJ, Comolli LR, Dick GJ, Hauser LJ, Hyatt D, Dill BD, Land ML, Verberkmoes NC, Hettich RL, Banfield JF (2010) Enigmatic, ultrasmall, uncultivated Archaea. *Proceedings of the National Academy of Sciences,* **107**, 8806-8811.

Barathi S, Vasudevan N (2001) Utilization of petroleum hydrocarbons by Pseudomonas fluorescens isolated from a petroleum-contaminated soil. *Environment International,* **26**, 413-416.

Barbier BA, Dziduch I, Liebner S, Ganzert L, Lantuit H, Pollard W, Wagner D (2012) Methane-cycling communities in a permafrost-affected soil on Herschel Island, Western Canadian Arctic: active layer profiling of *mcrA* and *pmoA* genes. *FEMS Microbiology Ecology,* **82**, 287-302.

Beal EJ, House CH, Orphan VJ (2009) Manganese- and iron-dependent marine methane oxidation. *Science,* **325**, 184-187.

Biddle JF, Lipp JS, Lever MA, Lloyd KG, Sørensen KB, Anderson R, Fredricks HF, Elvert M, Kelly TJ, Schrag DP (2006) Heterotrophic archaea dominate sedimentary subsurface ecosystems off Peru. *Proceedings of the National Academy of Sciences,* **103**, 3846-3851.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* **30**, 2114-2120.

Bretz KA, Whalen SC (2014) Methane cycling dynamics in sediments of Alaskan Arctic Foothill lakes. *Inland Waters,* **4**, 65-78.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, Mcdonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods,* **7**, 335-336.

Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ, Frischkorn KR, Tringe SG, Singh A, Markillie LM, Taylor RC, Williams KH, Banfield JF (2015) Genomic expansion of domain Archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Current Biology,* **25**, 690-701.

Crump BC, Amaral-Zettler LA, Kling GW (2012) Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *The ISME Journal,* **6**, 1629-1639.

Delong EF (1992) Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences,* **89**, 5685-5689.

Dridi B, Fardeau M-L, Ollivier B, Raoult D, Drancourt M (2012) Methanomassiliicoccus luminyensis gen. nov., sp. nov., a methanogenic archaeon isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology,* **62**, 1902-1907.

Frank-Fahle BA, Yergeau E, Greer CW, Lantuit H, Wagner D (2014) Microbial functional potential and community composition in permafrost-affected soils of the NW Canadian Arctic. *PLoS ONE,* **9**, e84761.

Galand PE, Casamayor EO, Kirchman DL, Potvin M, Lovejoy C (2009) Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *The ISME Journal,* **3**, 860-869.

Galand PE, Lovejoy C, Pouliot J, Garneau M-È, Vincent WF (2008) Microbial community diversity and heterotrophic production in a coastal arctic ecosystem: A stamukhi lake and its source waters. *Limnology and Oceanography,* **53**, 813-823.

Galand PE, Lovejoy C, Vincent WF (2006) Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aquatic Microbial Ecology,* **44**, 115.

Ganzert L, Bajerski F, Wagner D (2014) Bacterial community composition and diversity of five different permafrost-affected soils of Northeast Greenland. *FEMS Microbiology Ecology,* **89**, 426-441.

Glissman K, Chin K-J, Casper P, Conrad R (2004) Methanogenic pathway and archaeal community structure in the sediment of eutrophic Lake Dagow: effect of temperature. *Microbial Ecology,* **48**, 389-399.

Good IJ (1953) The population frequencies of species and the estimation of population parameters. *Biometrika,* **40**, 237-264.

Graham DE, Wallenstein MD, Vishnivetskaya TA, Waldrop MP, Phelps TJ, Pfiffner SM, Onstott TC, Whyte LG, Rivkina EM, Gilichinsky DA, Elias DA, Mackelprang R, Verberkmoes NC, Hettich RL, Wagner D, Wullschleger SD, Jansson JK (2012) Microbes in thawing permafrost: the unknown variable in the climate change equation. *The ISME Journal,* **6**, 709-712.

Grosse G, Jones B, Arp C (2013) Thermokarst lakes, drainage, and drained basins. In: *Treatise on Geomorphology* (eds Shroeder F, Giardino R, Harbor J). Academic Press, San Diego, pp. 325-353.

Großkopf R, Stubner S, Liesack W (1998) Novel euryarchaeotal lineages detected on rice roots and in the anoxic bulk soil of flooded rice microcosms. *Applied and Environmental Microbiology,* **64**, 4983-4989.

Haroon MF, Hu S, Shi Y, Imelfort M, Keller J, Hugenholtz P, Yuan Z, Tyson GW (2013) Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature,* **500**, 567-570.

He R, Wooller MJ, Pohlman JW, Catranis C, Quensen J, Tiedje JM, Leigh MB (2012a) Identification of functionally active aerobic methanotrophs in sediments from an arctic lake using stable isotope probing. *Environmental Microbiology,* **14**, 1403-1419.

He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB (2012b) Diversity of active aerobic methanotrophs along depth profiles of arctic and subarctic lake water column and sediments. *The ISME Journal,* **6**, 1937-1948.

He R, Wooller MJ, Pohlman JW, Quensen J, Tiedje JM, Leigh MB (2012c) Shifts in identity and activity of methanotrophs in arctic lake sediments in response to temperature changes. *Applied and Environmental Microbiology,* **78**, 4715-4723.

Herbold CW, Lee CK, Mcdonald IR, Cary SC (2014) Evidence of global-scale aeolian dispersal and endemism in isolated geothermal microbial communities of Antarctica. *Nature communications,* **5,** 1-10.

Høj L, Olsen RA, Torsvik VL (2005) Archaeal communities in High Arctic wetlands at Spitsbergen, Norway (78 N) as characterized by 16S rRNA gene fingerprinting. *FEMS Microbiology Ecology,* **53**, 89-101.

Høj L, Olsen RA, Torsvik VL (2008) Effects of temperature on the diversity and community structure of known methanogenic groups and other archaea in High Arctic peat. *The ISME Journal,* **2**, 37-48.

Holmes DE, Bond DR, O'neil RA, Reimers CE, Tender LR, Lovley DR (2004) Microbial communities associated with electrodes harvesting electricity from a variety of aquatic sediments. *Microbial Ecology,* **48**, 178-190.

Howard HH, Prescott GW (1973) Seasonal variation of chemical parameters in Alaskan tundra lakes. *American Midland Naturalist,* **90**, 154-164.

Hultman J, Waldrop MP, Mackelprang R, David MM, Mcfarland J, Blazewicz SJ, Harden J, Turetsky MR, Mcguire AD, Shah MB (2015) Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature*, **521**, 208-212.

Iino T, Tamaki H, Tamazawa S, Ueno Y, Ohkuma M, Suzuki K-I, Igarashi Y, Haruta S (2013) Candidatus Methanogranum caenicola: a Novel Methanogen from the Anaerobic Digested Sludge, and Proposal of Methanomassiliicoccaceae fam. nov. and Methanomassiliicoccales ord. nov., for a Methanogenic Lineage of the Class Thermoplasmata. *Microbes and Environments,* **28**, 244-250.

Kallmeyer J, Wagner D (2014) *Microbial Life of the Deep Biosphere,* De Gruyter, Boston.

Kan J, Clingenpeel S, Macur RE, Inskeep WP, Lovalvo D, Varley J, Gorby Y, Mcdermott TR, Nealson K (2011) Archaea in Yellowstone Lake. *The ISME Journal,* **5**, 1784-1795.

Kirchman DL, Hanson TE, Cottrell MT, Hamdan LJ (2014) Metagenomic analysis of organic matter degradation in methane-rich Arctic Ocean sediments. *Limnology and Oceanography,* **59**, 548-559.

Kittel TGF, Baker BB, Higgins JV, Haney JC (2011) Climate vulnerability of ecosystems and landscapes on Alaska's North Slope. *Regional Environmental Change,* **11**, S249-S264.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology,* **79**, 5112-5120.

Kubo K, Lloyd KG, F Biddle J, Amann R, Teske A, Knittel K (2012) Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *The ISME Journal,* **6**, 1949-1965.

Langille MGI, Zaneveld J, Caporaso JG, Mcdonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Thurber RLV, Knight R, Beiko RG, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology,* **31**, 814-821.

Lofton DD, Whalen SC, Hershey AE (2014) Effect of temperature on methane dynamics and evaluation of methane oxidation kinetics in shallow Arctic Alaskan lakes. *Hydrobiologia,* **721**, 209-222.

Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA (2014) *Brock biology of microorganisms,* Benjamin-Cummings Publishing Company, Boston.

Magurran AE (2003) *Measuring Biological Diversity,* John Wiley & Sons.

Matheus Carnevali PB, Rohrssen M, Williams MR, Michaud AB, Adams H, Berisford D, Love GD, Priscu JC, Rassuchine O, Hand KP, Murray AE (2015) Methane sources in arctic thermokarst lake sediments on the North Slope of Alaska. *Geobiology,* **13**, 181-197.

Mcguire AD, Anderson LG, Christensen TR, Dallimore S, Guo L, Hayes DJ, Heimann M, Lorenson TD, Macdonald RW, Roulet N (2009) Sensitivity of the carbon cycle in the Arctic to climate change. *Ecological Monographs,* **79**, 523-555.

Morrissey LA, Livingston GP (1992) Methane emissions from Alaska arctic tundra: An assessment of local spatial variability. *Journal of Geophysical Research: Atmospheres,* **97**, 16661-16670.

Murray AE, Hollibaugh JT, Orrego C (1996) Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Applied and Environmental Microbiology,* **62**, 2676-2680.

Muyzer G, De Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology,* **59**, 695-700.

Negandhi K, Laurion I, Whiticar MJ, Galand PE, Xu X, Lovejoy C (2013) Small thaw ponds: an unaccounted source of methane in the Canadian High Arctic. *PLoS ONE,* **8**, e78204.

Olajire A, Essien J (2014) Aerobic degradation of petroleum components by microbial consortia. *Journal of Petroleum & Environmental Biotechnology,* **5**, 1-22.

Orphan VJ, House CH, Hinrichs K-U, Mckeegan KD, Delong EF (2001) Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science,* **293**, 484-487.

Orphan VJ, House CH, Hinrichs K-U, Mckeegan KD, Delong EF (2002) Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proceedings of the National Academy of Sciences,* **99**, 7663-7668.

Osterkamp TE (2005) The recent warming of permafrost in Alaska. *Global and Planetary Change,* **49**, 187-202.

Oude Elferink SJWH, Akkermans-Van Vliet WM, Bogte JJ, Stams AJM (1999) Desulfobacca acetoxidans gen. nov., sp. nov., a novel acetate-degrading sulfate reducer isolated from sulfidogenic granular sludge. *International Journal of Systematic Bacteriology,* **49**, 345-350.

Parks DH, Beiko RG (2010) Identifying biologically relevant differences between metagenomic communities. *Bioinformatics,* **26**, 715-721.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics,* **30**, 3123-3124.

Pfeffer C, Larsen S, Song J, Dong M, Besenbacher F, Meyer RL, Kjeldsen KU, Schreiber L, Gorby YA, El-Naggar MY, Leung KM, Schramm A, Risgaard-Petersen N, Nielsen LP (2012) Filamentous bacteria transport electrons over centimetre distances. *Nature,* **491**, 218-221.

Phelps AR, Peterson KM, Jeffries MO (1998) Methane efflux from high-latitude lakes during spring ice melt. *Journal of Geophysical Research,* **103**, 29029-29036.

Pruesse E, Peplies J, Glöckner FO (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics,* **28**, 1823-1829.

Putkinen A, Juottonen H, Juutinen S, Tuittila E-S, Fritze H, Yrjälä K (2009) Archaeal rRNA diversity and methane production in deep boreal peat. *FEMS Microbiology Ecology,* **70**, 87-98.

R Core Development Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing.

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.

Serkebaeva YM, Kim Y, Liesack W, Dedysh SN (2013) Pyrosequencing-based assessment of the bacteria diversity in surface and subsurface peat layers of a northern wetland, with focus on poorly studied phyla and candidate divisions. *PLoS ONE,* **8**, e63994.

So CM, Phelps CD, Young LY (2003) Anaerobic transformation of alkanes to fatty acids by a sulfate-reducing bacterium, strain Hxd3. *Applied and Environmental Microbiology,* **69**, 3892-3900.

Stahl DA, Amann R (1991) Development and application of nucleic acid probes. 205-248. In: *Nucleic acid techniques in bacterial systematics* (eds Stackebrandt E, Goodfello M). John Wiley & Sons Ltd., Chichester.

Steven B, Briggs G, Mckay CP, Pollard WH, Greer CW, Whyte LG (2007) Characterization of the microbial diversity in a permafrost sample from the Canadian High Arctic using culture-dependent and culture-independent methods. *FEMS Microbiology Ecology,* **59**, 513-523.

Stoeva MK, Aris-Brosou S, Chételat J, Hintelmann H, Pelletier P, Poulain AJ (2014) Microbial community structure in lake and wetland sediments from a High Arctic polar desert revealed by targeted transcriptomics. *PLoS ONE,* **9**, e89531.

Takai K, Horikoshi K (1999) Genetic diversity of Archaea in deep-sea hydrothermal vent environments. *Genetics,* **152**, 1285-1297.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution,* **30**, 2725-2729.

Thienpont JR, Rühland KM, Pisaric MFJ, Kokelj SV, Kimpe LE, Blais JM, Smol JP (2013) Biological responses to permafrost thaw slumping in Canadian Arctic lakes. *Freshwater Biology,* **58**, 337-353.

Tveit A, Schwacke R, Svenning MM, Urich T (2013) Organic carbon transformations in High-Arctic peat soils: key functions and microorganisms. *The ISME Journal,* **7**, 299-311.

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. *Canadian Journal of Microbiology,* **57**, 303-315.

Wright ES, Yilmaz LS, Noguera DR (2012) DECIPHER, a aearch-based approach to chimera identification for 16S rRNA sequences. *Applied and Environmental Microbiology,* **78**, 717-725.

Yergeau E, Hogues H, Whyte LG, Greer CW (2010) The functional potential of High Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *The ISME Journal,* **4**, 1206-1214.

Zhang J, Yang Y, Zhao L, Li Y, Xie S, Liu Y (2015) Distribution of sediment bacterial and archaeal communities in plateau freshwater lakes. *Applied Microbiology and Biotechnology,* **99**, 3291-3302.

Zimov SA, Schuur EaG, Chapin FS III (2006) Permafrost and the global carbon budget. *Science,* **312**, 1612-1613.

Supplementary Tables and Figures

Table S3.1 Tukey's test comparing the log transformed abundances of phyla (based on aRNA sequences comprising ≥ 1 % of the sequences). Phyla and sites were interacting factors and depths act as an independent variable.

X-Y variables	Diference	*p-value*
Cyanobacteria:SIQ-Actinobacteria:SIQ	2.901	0.041
Cyanobacteria:SIQ-Bacteroidetes:SIQ-	3.807	0.002
Nitrospirae:SUKB-Bacteroidetes:SIQ	-2.981	0.032
Nitrospirae:SUKS-Bacteroidetes:SIQ	-3.399	0.008
Deltaproteobacteria:SIQ-Cyanobacteria:SIQ	3.601	0.004
Gammaproteobacteria:SIQ-Cyanobacteria:SIQ	3.092	0.022
Actino:SUKB-Cyanobacteria:SIQ	3.458	0.006
Bacteroidetes:SUKB-Cyanobacteria:SIQ	3.267	0.012
Delta:SUKB-Cyanobacteria:SIQ	3.962	0.001
Actino:SUKS-Cyanobacteria:SIQ	3.114	0.020
Bacteroidetes:SUKS-Cyanobacteria:SIQ	3.558	0.004
Deltaproteobacteria:SUKS-Cyanobacteria:SIQ	3.657	0.003
Nitrospirae:SUKS-Deltaproteobacteria:SIQ	-3.191	0.016
Nitrospirae:SUKS-Actinobacteria:SUKB	-3.050	0.026
Nitrospirae:SUKS-Bacteroidetes:SUKB	-2.856	0.048
Nitrospirae:SUKB-Deltaproteobacteria:SUKB	-3.137	0.019
Nitrospirae:SUKS-Deltaproteobacteria:SUKB	-3.554	0.004
Nitrospirae:SUKS-Bacteroidetes:SUKS	-3.148	0.018
Nitrospirae:SUKS-Deltaproteocbacteria:SUKS	-3.247	0.013

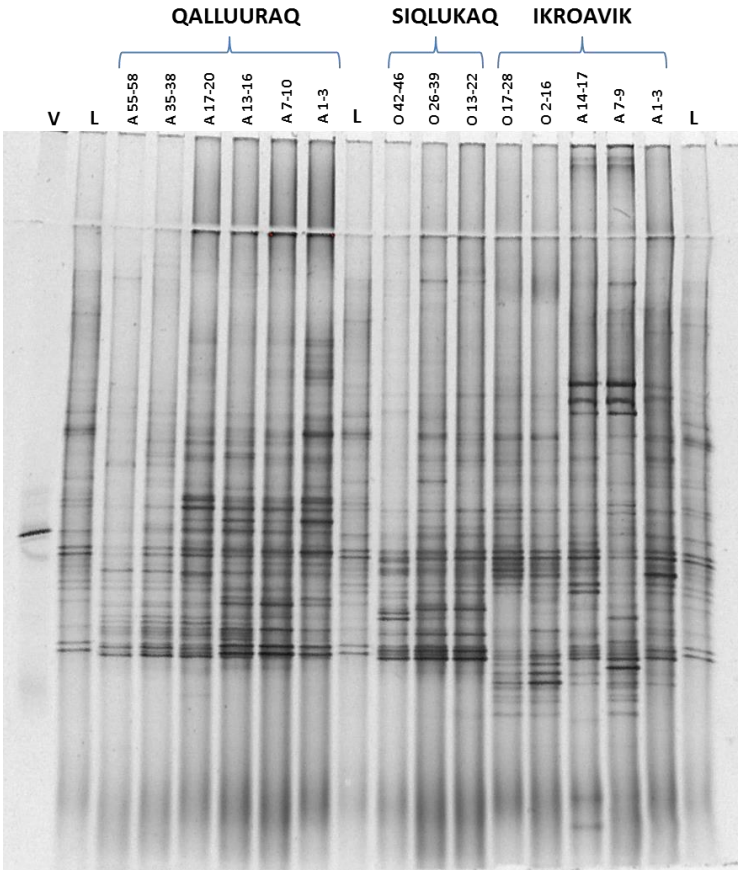


Figure S3.1 Denaturing gradient gel electrophoresis (DGGE) of the variable region 3 (V3) in the SSU rRNA gene of Archaea from different sediment intervals in Ikr10 (A, April 2010 samples, O, October 2010 samples) and Siq10. Genomic DNA was amplified with the Archaea-specific primer GC_358F and the UNIV517R. PCR reactions were pooled to obtain 600 ng of DNA. The denaturing gradient on the gel was between 30 % and 65% (top to bottom). The electrophoresis was performed at 1000V/hr for 16 hr at 60 °C. Most phylotypes were grouped by depth and at least one phylotypes was common to all lakes.

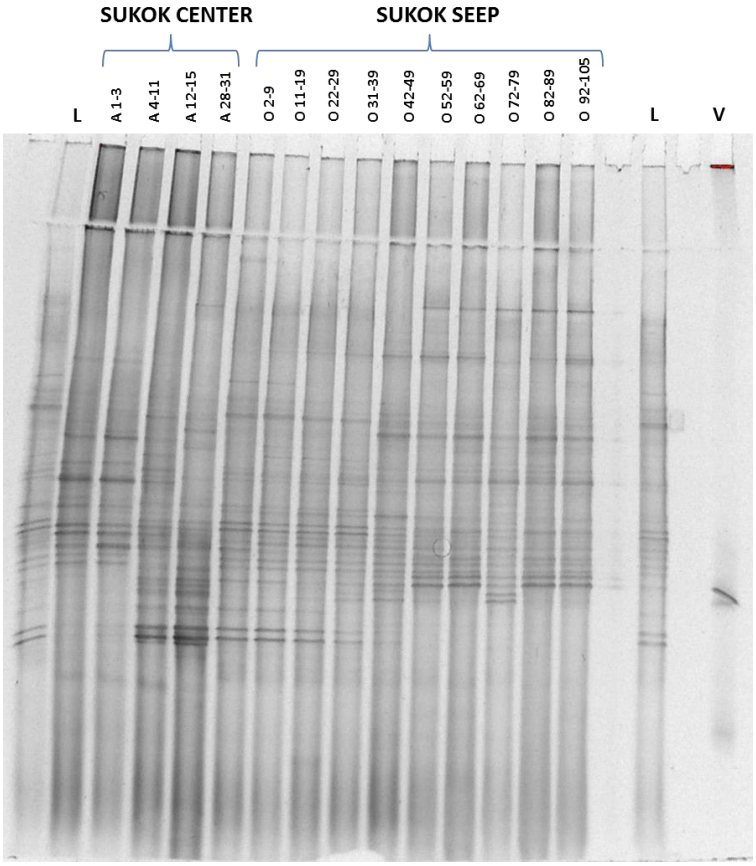


Figure S3.2 Denaturing gradient gel electrophoresis (DGGE) of the variable region 3 (V3) in the SSU rRNA gene of Archaea from different sediment intervals in SukS10. Genomic DNA was amplified with the Archaea-specific primer GC_358F and the UNIV517R. PCR reactions were pooled to obtain 600 ng of DNA. The denaturing gradient on the gel was between 30 % and 65% (top to bottom). The electrophoresis was performed at 1000V/hr for 16 hr at 60 °C. Most phylotypes were grouped by depth and at least one phylotypes was common to all lakes.

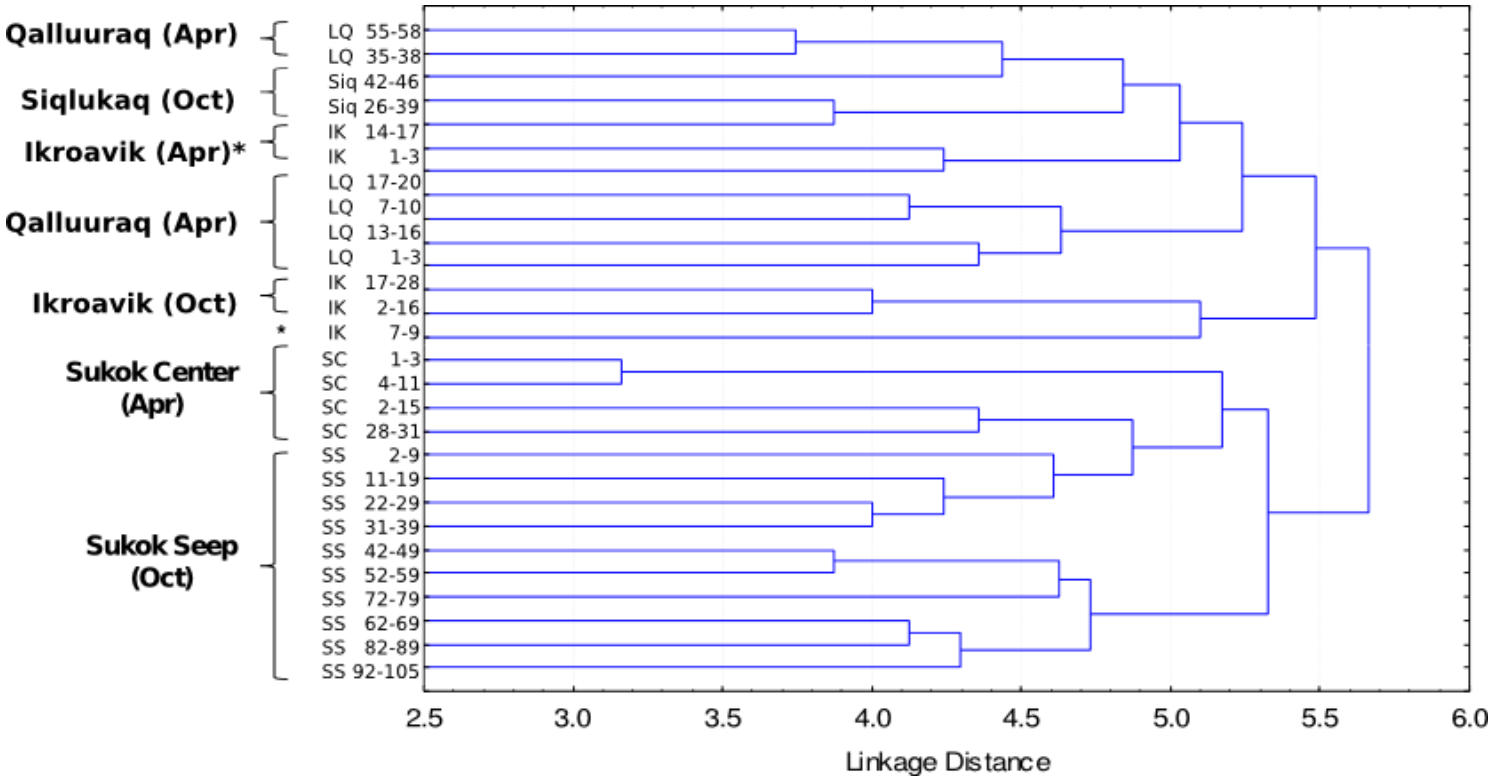


Figure S3.3 Unweighted pair group method with Arithmetic Mean (UPGMA) hierarchical clustering analysis of DGGE phylotypes based on Euclidean distances. Samples for this analysis were collected in 2010. Numbers next to the lake initials are sediment intervals in cm. *Ik 7-9 corresponds to April 2010.

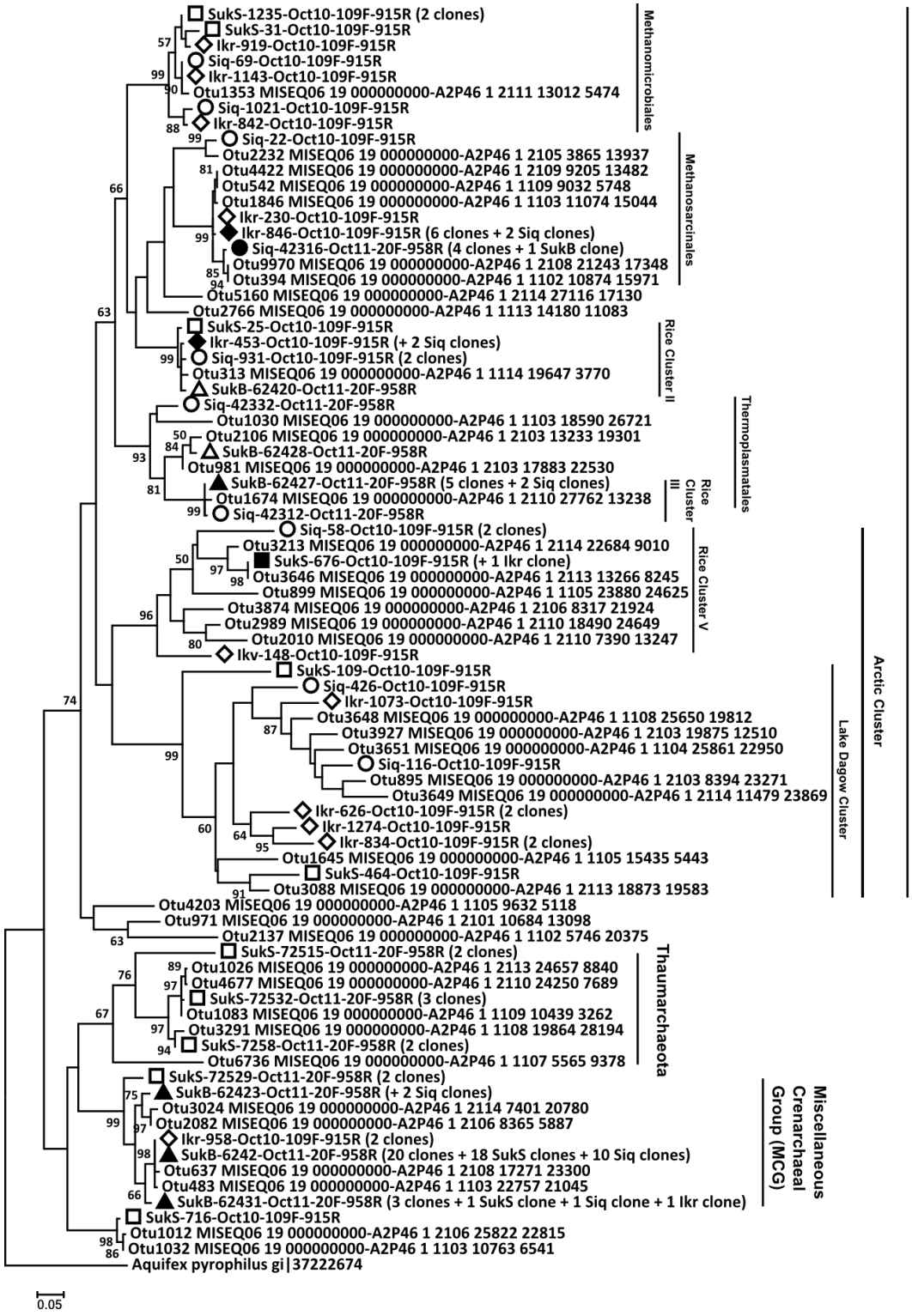
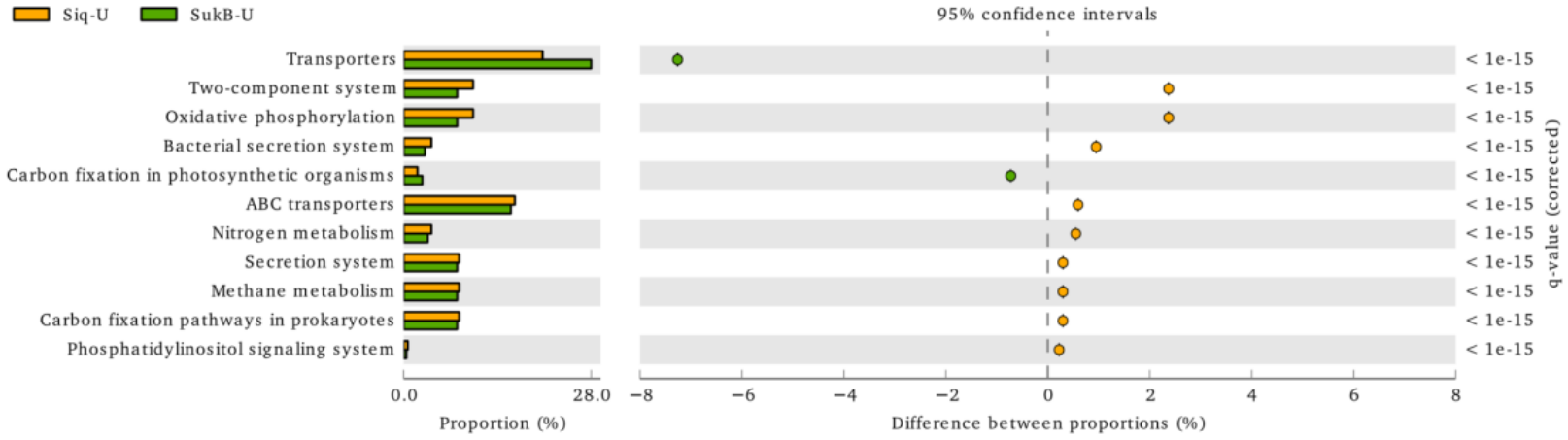
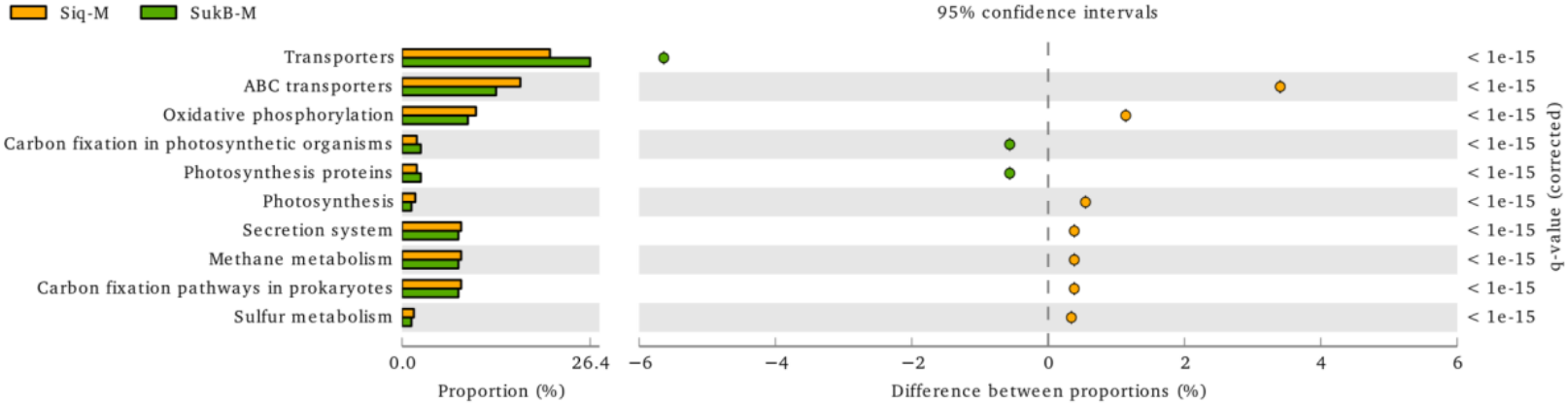

SukS-1235-Oct10-109F-915R (2 clones)
SukS-31-Oct10-109F-915R
Ikr-919-Oct10-109F-915R
Siq-69-Oct10-109F-915R
Ikr-1143-Oct10-109F-915R
Otu1353 MISEQ06 19 000000000-A2P46 1 2111 13012 5474
Siq-1021-Oct10-109F-915R
Ikr-842-Oct10-109F-915R
Siq-22-Oct10-109F-915R
Otu2232 MISEQ06 19 000000000-A2P46 1 2105 3865 13937
Otu4422 MISEQ06 19 000000000-A2P46 1 2109 9205 13482
Otu542 MISEQ06 19 000000000-A2P46 1 1109 9032 5748
Otu1846 MISEQ06 19 000000000-A2P46 1 1103 11074 15044
Ikr-230-Oct10-109F-915R
Ikr-846-Oct10-109F-915R (6 clones + 2 Siq clones)
Siq-42316-Oct11-20F-958R (4 clones + 1 SukB clone)
Otu9970 MISEQ06 19 000000000-A2P46 1 2108 21243 17348
Otu394 MISEQ06 19 000000000-A2P46 1 1102 10874 15971
Otu5160 MISEQ06 19 000000000-A2P46 1 2114 27116 17130
Otu2766 MISEQ06 19 000000000-A2P46 1 1113 14180 11083
SukS-25-Oct10-109F-915R
Ikr-453-Oct10-109F-915R (+ 2 Siq clones)
Siq-931-Oct10-109F-915R (2 clones)
Otu313 MISEQ06 19 000000000-A2P46 1 1114 19647 3770
SukB-62420-Oct11-20F-958R
Siq-42332-Oct11-20F-958R
Otu1030 MISEQ06 19 000000000-A2P46 1 1103 18590 26721
Otu2106 MISEQ06 19 000000000-A2P46 1 2103 13233 19301
SukB-62428-Oct11-20F-958R
Otu981 MISEQ06 19 000000000-A2P46 1 2103 17883 22530
SukB-62427-Oct11-20F-958R (5 clones + 2 Siq clones)
Otu1674 MISEQ06 19 000000000-A2P46 1 2110 27762 13238
Siq-42312-Oct11-20F-958R
Siq-58-Oct10-109F-915R (2 clones)
Otu3213 MISEQ06 19 000000000-A2P46 1 2114 22684 9010
SukS-676-Oct10-109F-915R (+ 1 Ikr clone)
Otu3646 MISEQ06 19 000000000-A2P46 1 2113 13266 8245
Otu899 MISEQ06 19 000000000-A2P46 1 1105 23880 24625
Otu3874 MISEQ06 19 000000000-A2P46 1 2106 8317 21924
Otu2989 MISEQ06 19 000000000-A2P46 1 2110 18490 24649
Otu2010 MISEQ06 19 000000000-A2P46 1 2110 7390 13247
Ikv-148-Oct10-109F-915R
SukS-109-Oct10-109F-915R
Siq-426-Oct10-109F-915R
Ikr-1073-Oct10-109F-915R
Otu3648 MISEQ06 19 000000000-A2P46 1 1108 25650 19812
Otu3927 MISEQ06 19 000000000-A2P46 1 2103 19875 12510
Otu3651 MISEQ06 19 000000000-A2P46 1 1104 25861 22950
Siq-116-Oct10-109F-915R
Otu895 MISEQ06 19 000000000-A2P46 1 2103 8394 23271
Otu3649 MISEQ06 19 000000000-A2P46 1 2114 11479 23869
Ikr-626-Oct10-109F-915R (2 clones)
Ikr-1274-Oct10-109F-915R
Ikr-834-Oct10-109F-915R (2 clones)
Otu1645 MISEQ06 19 000000000-A2P46 1 1105 15435 5443
SukS-464-Oct10-109F-915R
Otu3088 MISEQ06 19 000000000-A2P46 1 2113 18873 19583
Otu4203 MISEQ06 19 000000000-A2P46 1 1105 9632 5118
Otu971 MISEQ06 19 000000000-A2P46 1 2101 10684 13098
Otu2137 MISEQ06 19 000000000-A2P46 1 1102 5746 20375
SukS-72515-Oct11-20F-958R (2 clones)
Otu1026 MISEQ06 19 000000000-A2P46 1 2113 24657 8840
Otu4677 MISEQ06 19 000000000-A2P46 1 2110 24250 7689
SukS-72532-Oct11-20F-958R (3 clones)
Otu1083 MISEQ06 19 000000000-A2P46 1 1109 10439 3262
Otu3291 MISEQ06 19 000000000-A2P46 1 1108 19864 28194
SukS-7258-Oct11-20F-958R (2 clones)
Otu6736 MISEQ06 19 000000000-A2P46 1 1107 5565 9378
SukS-72529-Oct11-20F-958R (2 clones)
SukB-62423-Oct11-20F-958R (+ 2 Siq clones)
Otu3024 MISEQ06 19 000000000-A2P46 1 2114 7401 20780
Otu2082 MISEQ06 19 000000000-A2P46 1 2106 8365 5887
Ikr-958-Oct10-109F-915R (2 clones)
SukB-6242-Oct11-20F-958R (20 clones + 18 SukS clones + 10 Siq clones)
Otu637 MISEQ06 19 000000000-A2P46 1 2108 17271 23300
Otu483 MISEQ06 19 000000000-A2P46 1 1103 22757 21045
SukB-62431-Oct11-20F-958R (3 clones + 1 SukS clone + 1 Siq clone + 1 Ikr clone)
SukS-716-Oct10-109F-915R
Otu1012 MISEQ06 19 000000000-A2P46 1 2106 25822 22815
Otu1032 MISEQ06 19 000000000-A2P46 1 1103 10763 6541
Aquifex pyrophilus gi|37222674
0.05
Methanomicrobiales
Methanosarcinales
Rice Cluster II
Thermoplasmatales
Rice Cluster III
Rice Cluster V
Lake Dagow Cluster
Arctic Cluster
Euryarchaeota
Thaumarchaeota
Miscellaneous Crenarchaeal Group (MCG)

Figure S3.4 Maximum likelihood phylogenetic tree of archaeal SSU rRNA gene Sanger sequences and iTag sequences. Hollow symbols indicate that the given clone represents other clones from the same lake. Filled symbols indicate that the given clone represents other clones from the same lake and from other lakes as well. Lakes are coded by symbol as follows: Siq – circles, SukB – triangles, SukS – squares. Only bootstrap values of 50 or more are shown.The evolutionary history was inferred by using the Maximum Likelihood method based on the Jukes-Cantor model. The tree with the highest log likelihood (0.0000) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 81 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 250 positions in the final dataset. Evolutionary analyses were conducted in MEGA6.


Siq-U
SukB-U
95% confidence intervals
Transporters
Two-component system
Oxidative phosphorylation
Bacterial secretion system
Carbon fixation in photosynthetic organisms
ABC transporters
Nitrogen metabolism
Secretion system
Methane metabolism
Carbon fixation pathways in prokaryotes
Phosphatidylinositol signaling system
< 1e-15
q-value (corrected)
0.0
28.0
Proportion (%)
−8
−6
−4
−2
0
2
4
6
8
Difference between proportions (%)


Siq-M
SukB-M
95% confidence intervals
Transporters
ABC transporters
Oxidative phosphorylation
Carbon fixation in photosynthetic organisms
Photosynthesis proteins
Photosynthesis
Secretion system
Methane metabolism
Carbon fixation pathways in prokaryotes
Sulfur metabolism
< 1e-15
q-value (corrected)
0.0
26.4
Proportion (%)
−6
−4
−2
0
2
4
6
Difference between proportions (%)

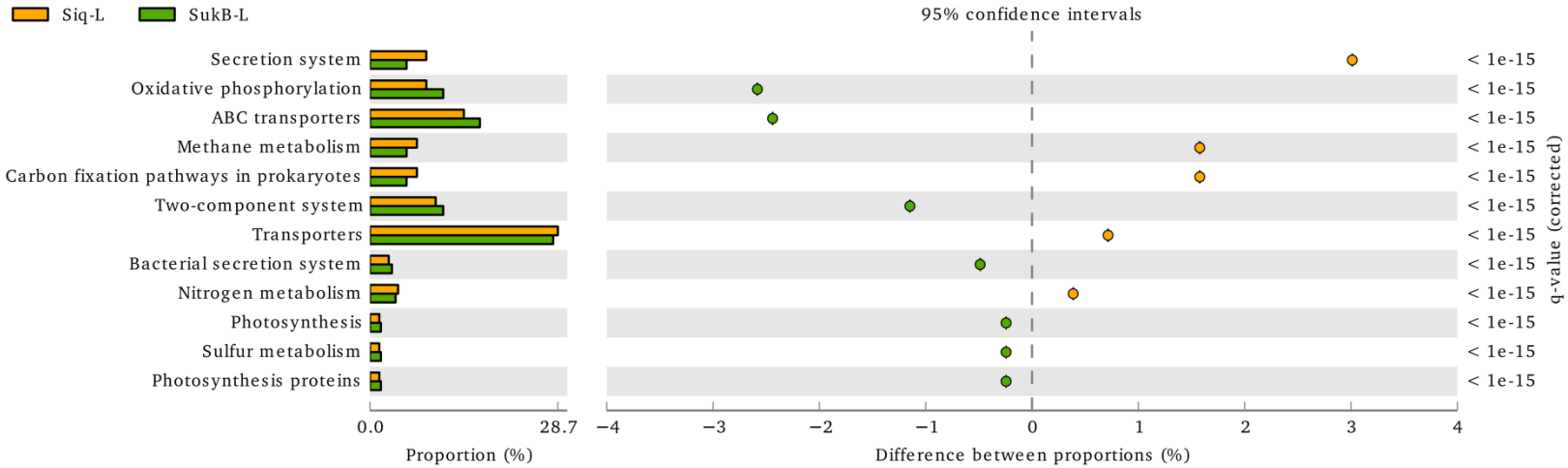
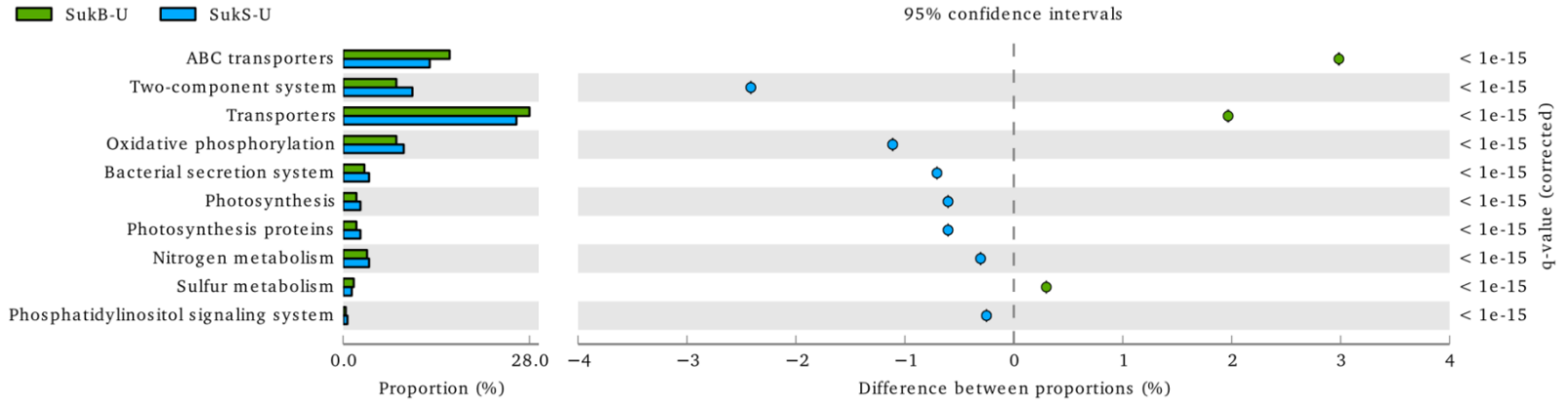


Figure S3.5 Extended error bar plots for the proportion of Siq and SukB sequences related to energy metabolism and environmental information processing KEGG categories filtered by effect size (> 0.2 %). KEGG Orthologs were ordered according to the difference between proportions (note the different scales). The Benjamini-Hocherg's FDR approach was used as a multiple test correction at a false discovery rate of 0.05 % (*q*-value > 0.05 %).



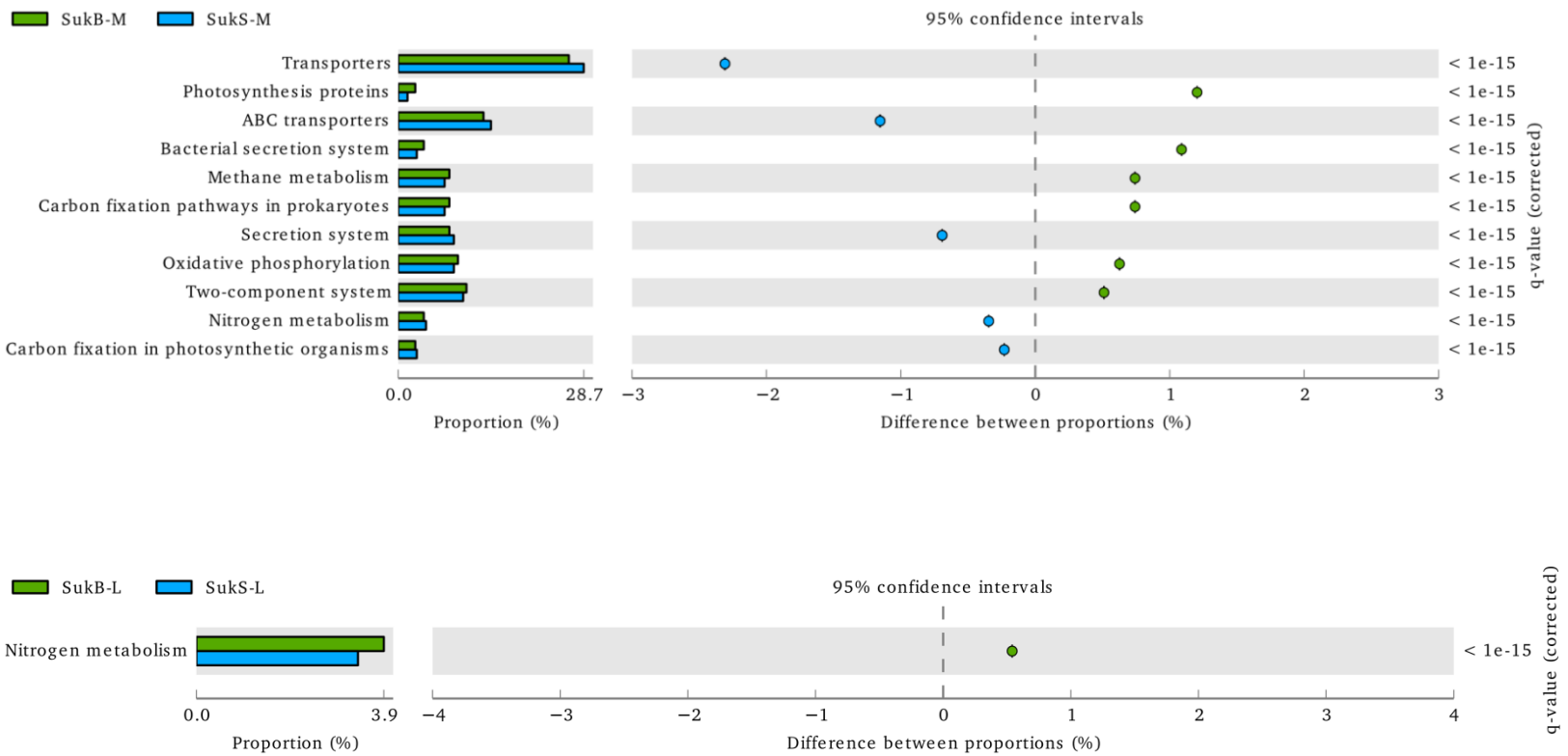


Figure S3.6 Extended error bar plots for the proportion of SukB and SukS sequences related to energy metabolism and environmental information processing KEGG categories filtered by effect size (> 0.2 %). KEGG Orthologs were ordered according to the difference between proportions (note the different scales). The Benjamini-Hocherg's FDR approach was used as a multiple test correction at a false discovery rate of 0.05 % (*q*-value > 0.05 %)

Chapter 4

Unique structure of the microbial community associated with snow algae in the Western U.S.

Co-authorship statement

P. B. Matheus Carnevali, T. H. Painter, K. P. Hand, and A. E. Murray

PM, TP, KH, and AM designed the sampling strategy and collected samples. AM designed the experiments and performed statistical analyses. PM and AM executed the experiments. PM analyzed the data, and wrote this chapter under the supervision of AM. TP, KH, AM secured funding for the research at CNS, and AM was part of a field course in which samples at WP and PCT were collected. AM secured support for the NGS.

Abstract

The snowpack in high mountain ranges across the world is an ephemeral environment that allows the exchange of nutrients and microorganisms between the atmosphere, the ground, and water bodies that are fed by meltwater. A variety of microorganisms including unicellular algae, bacteria, fungi, and ciliates inhabit the snowpack; and during spring and summer, when liquid water is available algal blooms (*e.g.*, *Chlamydomonas nivalis*) are observed as red snow. Studies about the diversity, composition and function of the microorganisms inhabiting seasonal snow cover are sparse. The present study aimed to uncover the diversity and structure of the microbial community in snowpacks with visible snow algae, and to understand the mechanisms that drive community assembly. Snow samples were collected from Mt. Conness (CNS) in the Sierra Nevada of California, which was considered our primary study site. Chemical analyses of these samples to determine concentrations of dissolved inorganic nutrients, dissolved organic carbon (DOC), particulate carbon (POC), particulate nitrogen (PN) and Chlorophyll *a* (Chl*a*) were conducted. The snow microbiota was studied by quantifying algal and bacterial cells, and by MiSeq (Illumina) next-generation sequencing (NGS) of the SSU rRNA gene (iTag sequencing) of the three domains of life from some of these samples. An additional site in the Sierra Nevada (the Pacific Crest Trail, PCT), and another site in the Snake Range in eastern Nevada (Wheeler Peak, WP1) were sampled for comparison of the microbial communities. Analysis of a resampled dataset of bacterial iTag sequences

indicated that the bacterial assemblage had a unique structure. In addition to the traditional 'singletons' (OTUs with just 1 sequence) that dominate NGS datasets, OTUs that comprised a high number of sequences were found at only one site ('singleton OTUs'). Interestingly, these OTUs were mostly affiliated to two different families of Betaproteobacteria (the Comamonadacea and the Oxalobacteraceae) in the genera *Polaromonas* and *Herminiimonas*. Even though at the OTU level the bacterial assemblage was only 5 % similar between some sites, there was a high resemblance in taxonomic composition among all three sites, and 7 'core' OTUs were affiliated to the Phyla Bacteroidetes and Actinobacteria. Variations in structure of the abundant OTUs between sites (and depths) were reflected in the proportion of sequences related to the same families of Betaproteobacteria and groups of Bacteroidetes. To elucidate possible mechanisms for microbial community assembly, and to explain this unique bacterial assemblage structure we tested two hypotheses: that the presence of snow algae determined bacterial assemblage structure, and that environmental factors affected the structure. An inverse relationship between Chl*a* and bacterial diversity was found, and this finding was supported by significant, positive, and negative Spearman's rank correlations between algal OTUs (from a parallel dataset) and bacterial OTUs, particularly between core bacterial OTUs and OTUs affiliated to the Chlorophyceae. Only two significant correlations were found between core bacterial OTUs and environmental factors (DOC and PN). Further testing using multivariate statistical approaches may allow us to formulate possible explanations to similar microbial community

composition geographically distant regions, and determine what mechanisms drive community assembly in these snowpacks.

1. Introduction

The so called "red snow", commonly found in snowpacks in high mountain ranges across the world, owes its coloration to the UV-resistant carotenoid pigments produced by green algae such as *Chlamydomonas nivalis* (Duval & Hoham, 2000; Thomas & Duval, 1995; Thomas, 1972). These algae inhabit an environment that is around freezing temperatures (Kawecka & Drake, 1978) most of the year, but during the spring and the summer, when liquid water is available, nutrients entrapped in the snow become available and promote algal growth (Komárek & Nedbalová, 2007). These conditions are also suitable for the presence of prokaryotes and a variety of uni- and multi-cellular eukaryotic organisms (*e.g.*, fungi, ciliates, insects, rotifers); and some bacteria have even been found to be attached to the algal cells (Weiss, 1983; Segawa et al., 2005; Takeuchi & Kohshima, 2004).

Studies about the diversity, composition and function of the microorganisms inhabiting seasonal snow cover are sparse (Larose et al., 2013b) and most studies have focused on the snow algae (Takeuchi & Kohshima, 2004; Duval & Hoham, 2000; Thomas & Duval, 1995; Muller et al., 1998; Kviderova, 2010; Lutz et al., 2014) which are composed mostly of different species of green alga (*e.g., C. nivalis,*

Chloromonas spp., Mesotaenium berggrenii, and *Raphidonema nivale*) (Takeuchi & Kohshima, 2004; Hoham, 1989). Studies about the bacterial assemblage inhabiting the snow in cold environments have been conducted, but not necessarily in connection with snow algae, and these studies encompass a range of alpine and polar environments such as glaciers in the Tibetan plateau (Liu et al., 2009; Zhang et al., 2010), the Tateyama Mountains in Japan (Segawa et al., 2005), Antarctica (Fujii et al., 2010), and the Arctic (Larose et al., 2010; Harding et al., 2011; Moller et al., 2013). Only recently, a study looked at microbial communities in red snow as a whole in Antarctica, and detected the presence of green algae, fungi, and bacteria by analysis of the SSU rRNA gene, with a predominant signal of bacteria in the genus *Hymenobacter* (Fujii et al., 2010). Also, next-generation sequencing was used to study the microbial community in High-Arctic snow, where the top and middle layers of the snowpack were dominated by bacteria in the phyla Proteobacteria, Bacteroidetes and Cyanobacteria (Moller et al., 2013).

The ephemeral nature of the snowpack adds a layer of complexity to the study of snow as an ecosystem, because the snowpack is an interphase that lasts for varying amounts of time (< 1 to 6 months in the Sierra Nevada depending on elevation and aspect) and allows the exchange of nutrients, microorganisms, and contaminants between the atmosphere, the ground, and water bodies that are fed by meltwater (Larose et al., 2013a). Little is known about how environmental factors affect the microbial community overall. However, light, temperature, snowmelt, and nutrients affect the life cycle of green algae flagellated stages, and other factors like

pH can affect algal growth in the lab (Hoham, 1989). Conversely, the presence of snow algae can have an effect on the snow albedo (Thomas & Duval, 1995) and on CO_2 fluxes from snowfields (Sommerfeld et al., 1991). Therefore, studying the microbial community inhabiting the snowpack would allow: i) to understand how the presence of microorganisms in the snow may contribute to cooling or warming of the climate, and how the drought conditions that the Western U.S. is currently experiencing may ultimately affect this ecosystem; ii) to aid future studies using remote sensing technologies to estimate microbial biomass in snowfields, because they may benefit from taking into consideration the microbial community composition and the variety of cell wall properties of the different microorganisms that make up the community. This knowledge would also contribute to expanding the database of spectral biosignature that can be detected remotely.

The present study aimed to uncover the diversity and structure of the microbial community in snowpacks with visible snow algae (red snow), and to understand the mechanisms that drive community assembly. For the purposes of this chapter, we will refer to all members of the snow bacteria occupying a given sample as an assemblage, given that we do not *a priori* know about the interactions between these organisms, though the concept of community assembly will be discussed. Using the conventional terminology, community is a group of potentially interacting organisms that co-occur in space and time (Magurran, 2003). To address these issues, we looked at assemblage diversity and structure and tested two hypotheses: (1) that an algal bloom (visible red snow) would be usually associated

with the same groups of microorganisms, even in geographically distant locations, and (2) that environmental selection could determine the structure and composition of the entire community. In order to test these hypotheses we sampled two sites in the Sierra Nevada of California and another site in the Snake Range in eastern Nevada. Chemical analyses of snow samples were conducted to determine dissolved inorganic nutrients, dissolved organic carbon (DOC), particulate carbon (POC), particulate nitrogen (PN) and Chlorophyll *a* (Chl*a*) concentrations. The snow microbiota were studied by quantifying algal and bacterial cells, and by next-generation sequencing of the SSU rRNA gene of the three domains of life from some of these samples.

2. Methods

2.1. Area of study and samples collection

Mount Conness (CNS), located north of Tioga Pass (37° 58′01″ N, 119° 19′17″ W), on the border between Inyo National Forest and Yosemite National Park, is one of the peaks (12,649 ft) of the Sierra Nevada range in California. Samples were collected on 7 July 2009 in the snow fields of CNS at a location of a long term study site where previous surveys tracked the presence of Chlamydomonadaceae-related algae, *e.g*. *Chlamydomonas nivalis,* on ground-based and airborne imaging spectrometer data (Painter et al., 2001). At the base of the mountain, a ~12 m long horizontal transect (Fig. S4.6a) was established along the snow field. Between 0.05 cm^3 and 0.1 cm^3 of snow were collected at surface using sterile instruments, at 1, 2, 3, 4, 5, 7, 9 and 11

m from one end of the transect, and then stored in sterile Whirl Pak bags. Between meter 4 and meter 5, a snow pit was excavated and samples were collected along a vertical profile at the surface (0 m), and 0.2, 0.3, 0.6 and 0.9 m below the surface, in the same way samples were collected from the horizontal transect (Fig. S4.6b). At the time of collection three whirlpak bags (~ 1 L ea.) of snow were obtained from each of the sites for subsequent analysis (*e.g.*, chemical analyses, cell counts, nucleic acids extractions). All samples were stored under cold conditions during transportation to the lab either on the same or subsequent day. Other samples were collected in a similar fashion in July 2010 from the Pacific Crest Trail (PCT; 39° 17' 06"N, 120° 19' 10'' W; 18 July), which is another sampling site in the Sierra Nevada of California, and from Wheeler Peak (WP1; 38° 59' 34'' N, 114° 18' 26" W; 14 July) in Great Basin National Park, which is the second tallest peak in Nevada. Samples from the PCT were collected at 2,406 m.a.s.l. and samples from WP (WP1) were collected at 3,292 m.a.s.l. CNS was our primary sampling site, and the other two sites were mainly used for microbial assemblage comparisons.

2.2. Confocal microscopy

In preparation for microscopy, snow samples from all three sites were thawed at room temperature, mixed well, and aliquoted in 50 mL conical tubes. Ten percent of the volume of the sample was then added as formalin 37 % to fix the cells. Staining was performed using 1 mL of 4',6'-diamidino-2-phenylindole (DAPI, 0.1 mg mL^{-1}) and 10 mL of each sample. Samples were kept in the dark for at least 10 min with

rotation, and subsequently passed through a 0.2 µm GTBP filter and a 0.45 µm HAWP filter using a vacuum pump. Finally, the filters were placed on clean glass slides with immersion oil and cover slips. The slides were examined by confocal microscopy (Olympus, FluoView™ FV1000). Visualization of the algae under the microscope was done at a magnification of 200X, and bacteria were visualized at a magnification of 600X. Only CNS samples counts are presented here.

2.3. Chemical analyses

Snow samples from CNS were thawed and 20-60 mL were passed through combusted GF/F filters and collected into clean bottles. The filtrates were used for analysis of dissolved organic carbon (DOC), nitrite ($N\text{-}NO_2$), nitrate ($N\text{-}NO_3$), ammonia ($N\text{-}NH_3$), and phosphate ($O\text{-}PO_4$) concentrations, according to procedures outlined in Standard Methods for the Examination of Water and Wastewater (Greenberg et al., 1992). Filters were kept for particulate organic carbon (POC) and particulate nitrogen (PN) analysis, following Karl et al. (1991). Another 6-50 mL portion of each sample (from CNS and from WP1) was filtered through GF/F filters and kept at -20 °C until processing. Filters were then incubated in 10 mL of 90 % acetone in the dark at -20 °C, for less than 24 hr in order to extract chlorophyll. Chlorophyll *a* was measured following the method by Welschmeyer (1994) using a fluorometer (Tuner Designs Model 10-AU).

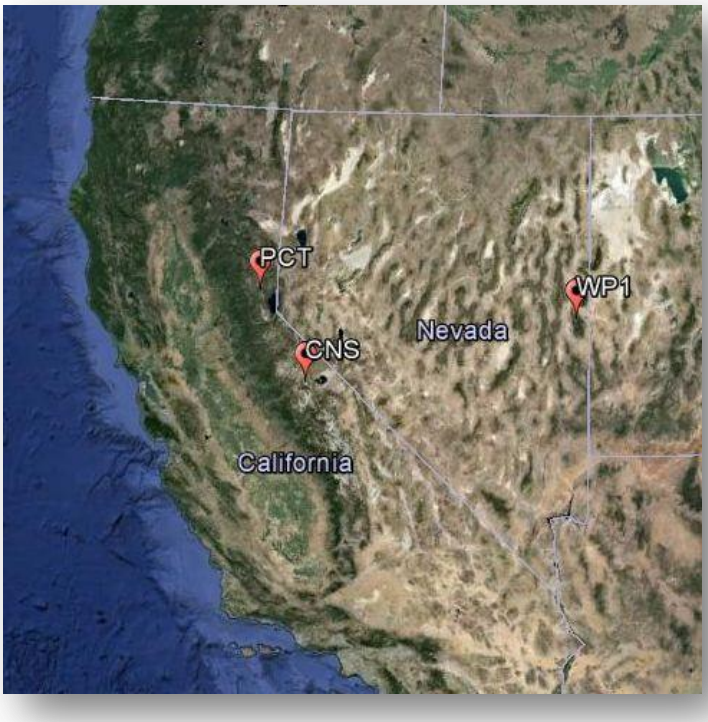


Figure 4.1 Location of the CNS and PCT sampling sites on the Sierra Nevada (CA) and WP1 on the Snake Range (NV). Source: Google Earth.

2.4. Nucleic acids extraction and next generation sequencing

The snow samples that had been kept at -80 °C were thawed for two days at 4 °C and filtered through a 0.2 µm pore size Sterivex filter (Millipore) using a peristaltic pump. Sucrose lysis buffer (1.8 mL; 40 mM EDTA, 50 mM Tris-HCl, 0.75 M sucrose) was added to each filter, and the filters were preserved at -20 °C until further processing. Cell lysis was completed by adding lysozyme (1 mg mL^{-1}) and the filters were incubated at 37 °C for 30 minutes, combined with an addition of proteinase K (0.5 mg mL^{-1}) and 20 % sodium dodecyl sulfate (SDS), and incubation at 55 °C for 2

hr. A standard phenol:chloroform nucleic acids extraction was performed on the lysates following methods in (Massana et al., 1997), with the exception that two rinses with chloroform-isoamylalcohol (24:1) were done in place of just one. The aqueous phase was finally passed through a Microcon filter to clean the DNA using sterile DI water and concentrate it to a volume of ~ 80 µL. DNA quantification was done using a NanoDrop 100. DNA extracts were preserved at -80 °C.

DNA extracts from selected samples and all three sampling sites was used to amplify and sequence the variable region 4 (V4) of the SSU rRNA gene of Bacteria and Archaea, with primers 515F-806R (Caporaso et al., 2010) and the V4 region of the SSU rRNA gene of Eukarya with primers V4Forward – 5'-CCAGCASCYGCGGTAATTCC-3' (*Saccharomyces cerevisiae* positions 565-584) and V4Reverse – 5'- ACTTTCGTTCTTGATYRA-3' (*S. cerevisiae* positions 964-981) (Hadziavdic et al., 2014). Extracts were processed in the Joint Genomic Institute (JGI) for paired-end sequencing (2 x 300 bp) using the MiSeq Illumina platform (iTaq sequencing), and contaminants, barcodes, and primers were removed from the raw fastq files by JGI.

2.5. iTag sequences analyses and statistics

Sequences were analyzed following the MiSeq SOP outlined for Mothur (Kozich et al., 2013). In brief, sequences were separated into individual sample files containing forward and reverse reads using deinterleave_fastq.sh (https://gist.github.com/nathanhaigh/3521724). These sequences were joined into

contigs, and quality filtered by removing sequences with any ambiguous bases, or that were longer than 325 bp. Unique sequences were identified and aligned against the newest release of the Greengenes database (August, 2013), preclustered, classified, and clustered into operational taxonomic units (OTUs) at a distance of 0.03 following the MiSeq SOP (Kozich et al., 2013). The remaining sequences were used to create a database (using the mothur script, create.database) containing representative sequences for each OTU formed at a distance of 0.03 and the OTUs were classified as well. A perl script (Daisychopper.pl) was used to subsample the sequences to the lowest number of sequences in a given sample (103,385 sequences). The subsampled data set was used for alpha-diversity and beta-diversity statistical analyses. Spearman's rank order correlations were calculated with the statistical package Statistica Data Miner.

3. Results

3.1. Cell counts

Confocal microscopy was used to observe and quantify algal and bacterial cells in samples collected from CNS. The number of both kinds of cells was on the same order of magnitude (10^3 cells mL^{-1}), but the number of bacterial cells was at least 2.5-fold higher. The highest numbers of algae were observed at meter 4 on the horizontal transect, and at the surface (0.0 m) of the vertical profile (between meters 4 and 5), which correlated with field observations of the highest amounts of "red snow" in the area sampled. The number of algal cells was higher between

meters 7 and 11 (~ 1.4 x 10^3 cells mL^{-1}) than at meters 1, 2, 3 and 5 (< 1.3 x 10^3 cells mL^{-1}; Fig. 1A). The opposite pattern was observed for bacteria, whose highest numbers were detected between meters 1 and 3, particularly at meter 2, where the highest number of bacteria of the whole horizontal transect was found. The number of both, algal and bacterial cells decreased with depth, but it was evident that the algae showed similar numbers of cells between the horizontal transect and the vertical profile (1.1-1.8 x 10^3 cells mL^{-1}).

3.2. Chemical analyses

Snow samples from all sites were collected in July, a time when the snowmelt was well underway and the snowpack was changing. At CNS, a suite of chemical parameters that could potentially affect or be affected by microbial life were determined. The concentration of chlorophyll *a* (Chl*a*) was measured as a proxy for the presence of photosynthetic microbial life, particularly green algae, in addition to performing cell counts, mentioned in the previous section. The Chl*a* concentration was relatively constant along the horizontal transect (0.011-0.164 μg L^{-1}), except at meter 9, where it was the highest (1.000 μg L^{-1}) (Fig. 1B). A similar trend was observed for NO_2-N (1-2 μg L^{-1}), as well as the concentration of PO_4-P, although at meters 1 and 9 there were slightly higher concentrations (5 μg L^{-1}) of the later (Fig. 1B). The highest DOC concentration (18 μg L^{-1}) was also observed at meter 9 (Fig. 1A), as well as the highest concentration of Chl*a*. At this same location, an increase of NO_3-N and NH_3-N concentration was observed, which was only surpassed by the

concentrations of these nutrients observed at the end (11 m) of the horizontal transect (33 and 24 μg L^{-1} respectively). Particulate organic carbon (POC) was found in low amounts (< 11 mg L^{-1}) along the horizontal transect (Fig. 4.2A), and the highest concentration was detected at meter 3, where the concentration of Chl*a* also showed a small peak, and the number of bacterial cells was higher than in other locations (Fig. 4.2A). The concentration of PN was consistently lower than that of POC (≤ 0.16 mg L^{-1}) along the horizontal transect. Only samples collected at meters 2 and 11 on the horizontal transect, and at 0.2 m below surface, had N to P ratios that were close to the Redfield ratio 16:1 (Redfield, 1934) with values of 14.00, 14.74, and 15.50 respectively. The C to N ratios that were closest to the Redfield ratio (6:625) were found at meters 5 (5.845) and 11 (5.96) on the horizontal transect and at 0.9 m below surface (7.17). Only at these sites where the Redfield ratios were close to the expected values nutrients may not have been limiting.

Nitrogen in the form of nitrite (NO_2-N) was the only nutrient whose concentrations (1 μg L^{-1}) were low and consistent throughout the different depths. Across the depth profile, the highest amount of POC (37.88 mg L^{-1}) was detected at the surface, with the concentration being ~ 4 times the concentration of POC detected at meter 3 of the horizontal transect. This result is not surprising, because at this site, abundant red snow was observed at the time of collection, and high numbers of algae were observed under the microscope, explaining the high content of particulate matter. However, at a nearby location (meter 4 of the horizontal transect) the concentration of POC was much lower (1.87 mg L^{-1}), as well as the

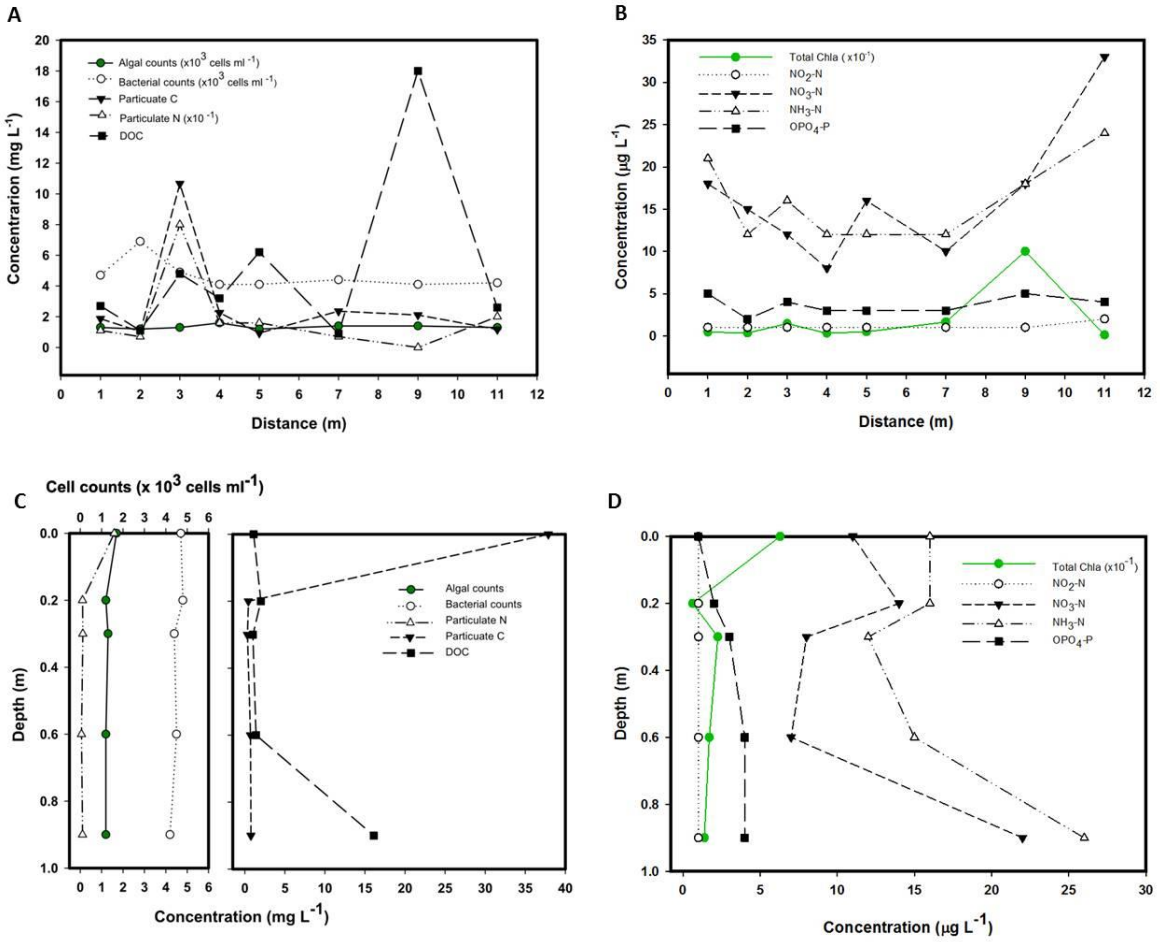


Figure 4.2 Cell counts and snow chemistry along the horizontal transect and the vertical profile in CNS on July 2009. Algal and bacteria counts, particulate carbon (POC), particulate nitrogen (PN) and dissolved organic carbon (DOC) concentrations along the transect (A) and the vertical profile (C). Concentration of chlorophyll a (Chl*a*), nitrogen in the form of nitrate (NO_3-N), ammonia (NH_3-N), and nitrite (NO_2-N), and orthophosphate (OPO_4-P) along the transect (B) and the vertical profile (D).

concentration of DOC and inorganic nutrients, exemplifying a relatively patchy distribution of the chemical constituents measured. The concentration of PN across the vertical profile was consistently low (≤ 1.59 mg L^{-1}).

3.3. Microbial community structure and diversity based on iTag sequences

To gain deeper insights into the structure of the microbial community inhabiting snow, iTag next-generation sequencing (MiSeq, Illumina) of the bacterial and archaeal SSU rRNA gene (V4 region) was applied to samples collected during summer from three locations CNS, PCT and WP1. Surface or near-surface snow algae were present at each location at varying levels, enabling us to compare the structure of the community between areas dominated by snow algae and areas with less visible snow algae. After sequence processing following MiSeq SOP (Kozich et al., 2013) there was between 103,385 and 147,860 sequences per sample. A subset of the CNS samples (10) were sequenced, and only samples from the first 40 cm in the snowpack were compared between sites at WP1 and PCT. These sequences initially clustered into 70,326 OTUs at a distance of 0.03, then following subsampling to the lowest number of sequences (103,385), resulted in a total of 45,992 OTUs.

Given that the majority of OTUs were singletons (35,679 OTUs were only represented by a single sequence), doubletons and tripletons, all OTUS with 3 or few sequences were removed prior to comparative analyses. The remaining OTUs (4,416) encompassed between 99,048 and 101,406 sequences, and fell into three groups: those OTUs that comprised 1.00 % of the sequences or more (127 OTUs),

OTUs that comprised between 0.01 and 0.99 % of the sequences (1,947 meso-abundant OTUs), and those with less than 0.0099 % of the sequences (2,411 OTUs). This last group corresponds to our definition for rare OTUs, which was within the range used by Hugoni et al. (2013). To simplify the analysis of the meso-abundant and the rare OTUs, only OTUs that were present in 3 samples or more were retained for analysis.

The structure of the OTU data set was distinctive such that 71 % of all OTUs were present in only one sample (*e.g.* 55 OTUs out of the 127 OTUs comprising ≥ 1 % of the sequences were singleton OTUs; Fig. 4.3). These are referred to as 'singleton OTUs' and could even dominate a given snow sample, *i.e.*, in 5 of the 15 samples singleton OTUs represented > 10% of the sequences per sample. Similar assemblage structures were observed among the meso-abundant and rare 'singleton OTUs' (Fig. S4.1). Among the abundant, meso-abundant and rare groups there usually was 1 abundant OTU that comprised a higher percentage of sequences (0.60 - 0.95 % in the last two groups), particularly at the surface snow layers of all three sites, and the outer sites on the CNS horizontal transect (Fig. S4.1). Strikingly, even though the most abundant OTUs were unique, the taxonomic composition of 'singleton OTUs' was similar among the abundant, mesoabundant and rare OTU groups (data not shown). In the CNS samples the majority of 'singleton OTUs' were related to *Herminiimonas* (Class Betaproteobacteria, Order Burkholderiales, family Oxalobacteraceae). In the PCT and the WP1 samples there was a mixture of *Herminiimonas*, and *Polaromonas*, another genus of Burkholderiales in the family

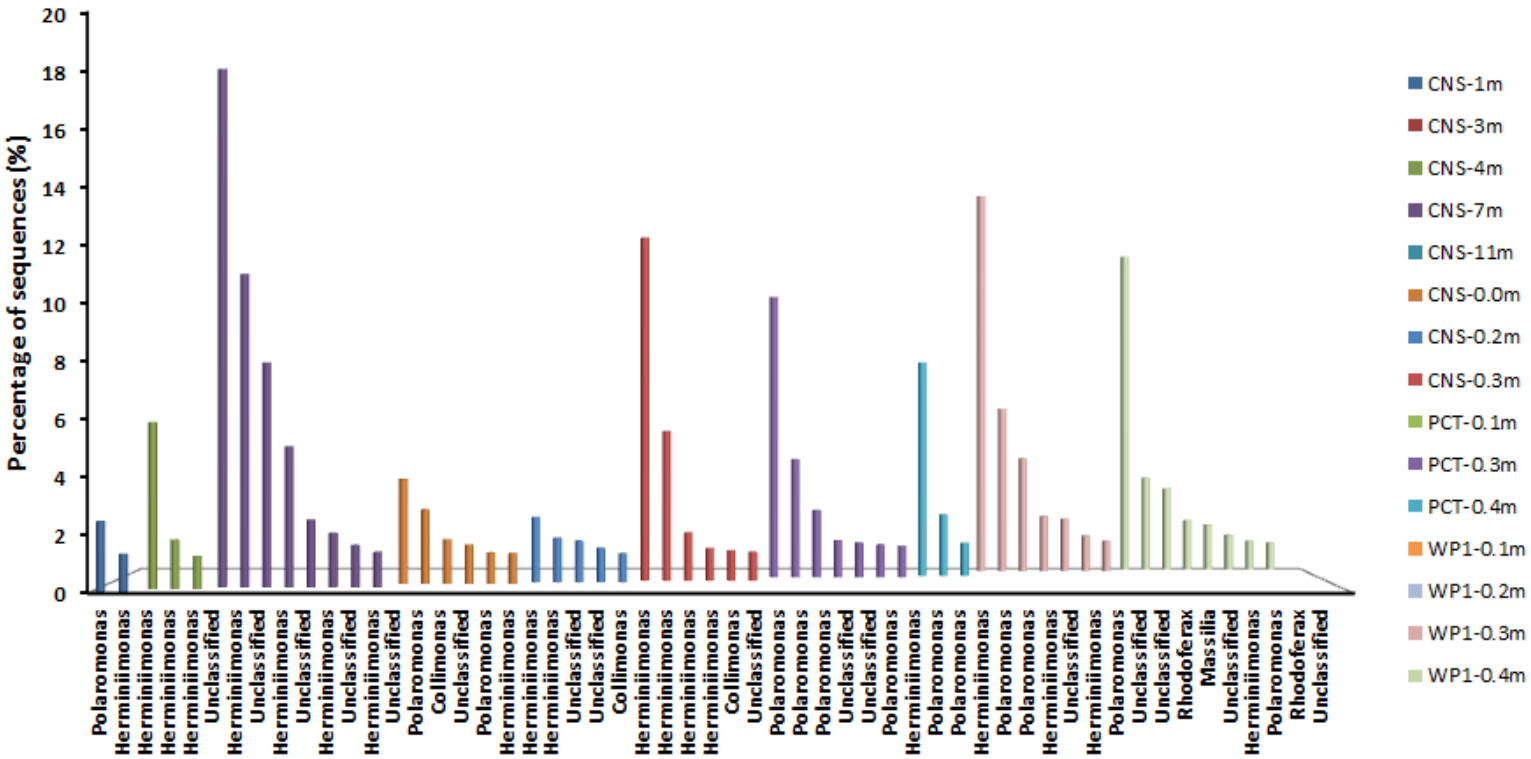


Figure 4.3 OTUs found in only one sample (singleton OTUs) that comprised ≥ 1 % of sequences (subsampled sequence dataset including OTUs with 3 sequences or more). Only 10 out of the 15 samples were represented on this figure. Missing samples did not have 'singleton OTUs' comprising 1 % of the sequences or more, but they have representative 'singleton OTUs' that comprised < 1 % of the sequences (Fig. S4.1). Coloring scheme matches Fig. S4.1 for comparative purposes.

Comamonadaceae. Other well-represented genera observed in these sites were *Rhodoferax* (Comamonadaceae) and *Massilia* (Oxalobacteraceae).

When considering the assemblage as a whole, OTUs related to the families Comamonadaceae and Oxalobacteraceae were among the abundant OTUs, and comprised the largest number of sequences in all samples (Fig. 4.4a). In the family Comamonadaceae, OTUs related to *Polaromonas* were more abundant in the depth profiles of PCT and WP1, with increasing numbers of sequences with depth. In CNS, *Collimonas*, another genus within this family was more prevalent among samples. However, except for the sample collected at meter 7 on the CNS transect, most samples from CNS were dominated by OTUs related to members of the Chitinophagaceae, Sphingomonadaceae and other unclassified OTUs in the phylum Bacteroidetes.

There was a common "core" set of OTUs related to Bacteroidetes (OTUs 1-4 and OTU8) and Actinobacteria (OTUs 6 and 7) present in all samples, though the proportions varied widely amongst the 15 samples (*e.g.* ranging from 0.01 - 32 % for a single OTU; Fig. S4.2). Notably, there was a sharp difference in abundance between the core OTUs present in the Sierra Nevada and those present in Wheeler Peak. WP1 was dominated by Bacteroidetes-related OTUs, OTU2 (average 24.5 ± 13.0 %; Class Cytophagia) and mostly OTU1 (average 4.7 ± 3.2 %; Class Saprospirae), as opposed to a more even distribution of the other OTUs observed in the Sierra sites. Across the Sierra sites Actinobacteria, classes Saprospirae (OTU1) and Sphingobacteria (OTU4), and other unclassified Bacteroidetes (OTU3) averaged

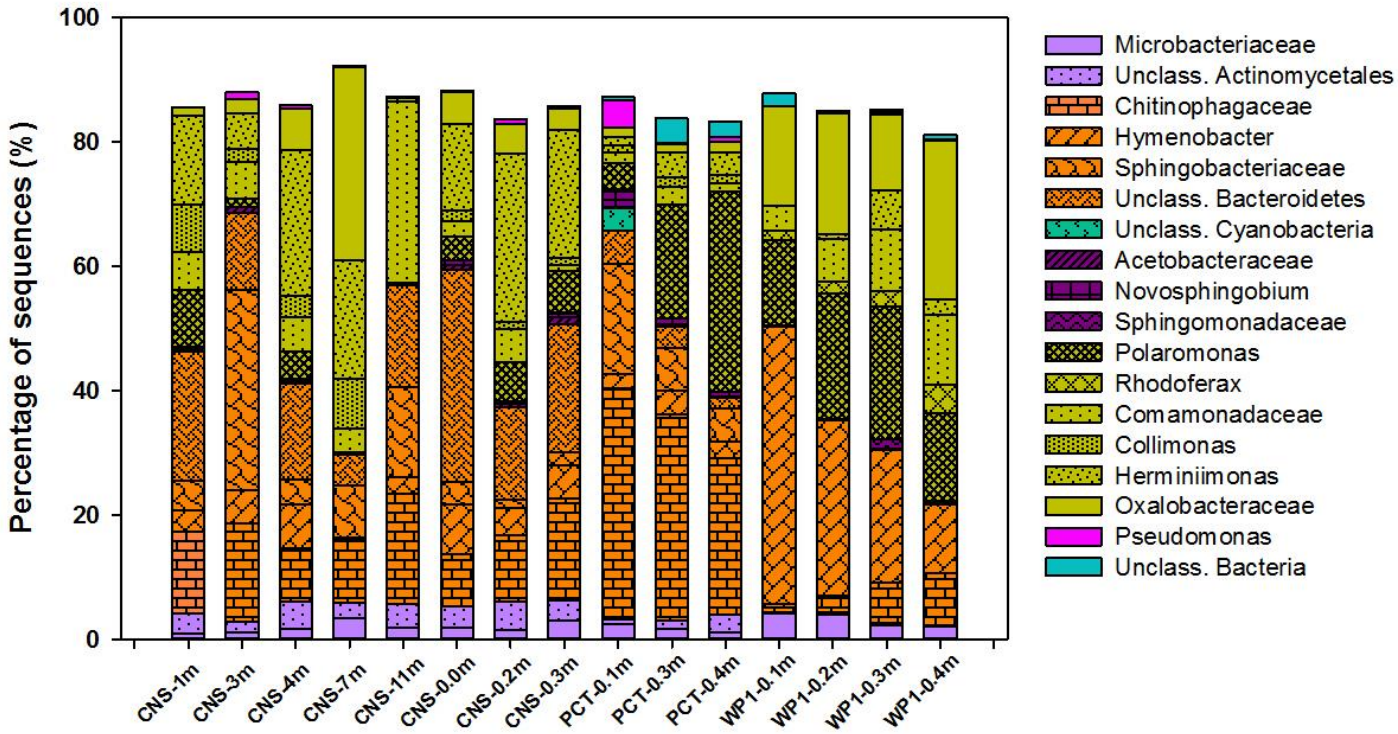


Figure 4.4a Relative abundance of the abundant OTUs (≥ 1 % of the sequences) grouped by the highest taxon possible. Taxa were grouped by color, each color representing a different phylum or class where appropriate. From top to bottom: Phylum Actinobacteria (Family Microbacteriaceae, Order Actinomycetales); Phylum Bacteroidetes (Family Chitinophagaceae, Genus *Hymenobacter*, Family Sphingobacteriaceae, Unclassified Bacteroidetes); Phylum Cyanobacteria, Class Alphaproteobacteria (Family Acetobacteraceae, Genus *Novosphingobium*, Family Sphingomonadaceae); Class Betaproteobacteria (Genera *Polaromonas* and *Rhodoferax*, Family Comamonadaceae, Genera *Collimonas* and *Herminiimonas*, Family Oxalobacteraceae); Class Gammaproteobacteria (Genus *Pseudomonas*). Sample IDs include the location in capital letters (CNS: Mt. Conness, PCT: Pacific Crest Trail, WP1: Wheeler Peak) followed by distance from one end of the horizontal transect (1 m, 3 m, 4 m, 7 m, 11 m) or by snow pit depth. Depths represent the bottom of the depth range, except for CNS-0.0 m that represents the upper end.

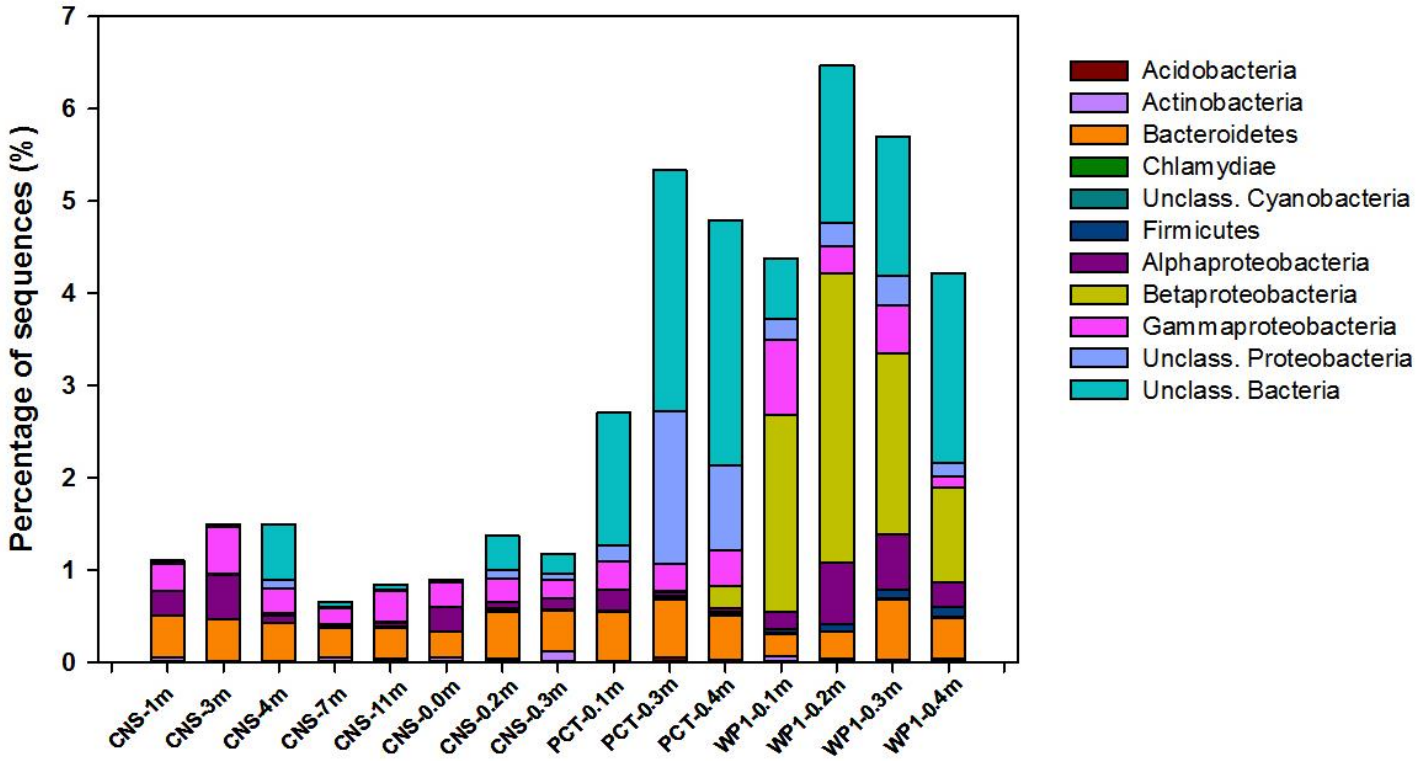


Figure 4.4b Relative abundance of the meso-abundant OTUs (0.01 – 0.99 % of the sequences) grouped by Phylum or Class (same color scheme as Fig. 4.4a) to show degree of resemblance to abundant OTUs at the indicated taxonomic level. Sample IDs include the location in capital letters (CNS: Mt. Conness, PCT: Pacific Crest Trail, WP1: Wheeler Peak) followed by distance from the end of one horizontal transect (1 m, 3 m, 4 m, 7 m, 11 m) or by snow pit depth. Depths represent the bottom of the depth range, except for CNS-0.0 m that represents the upper end.

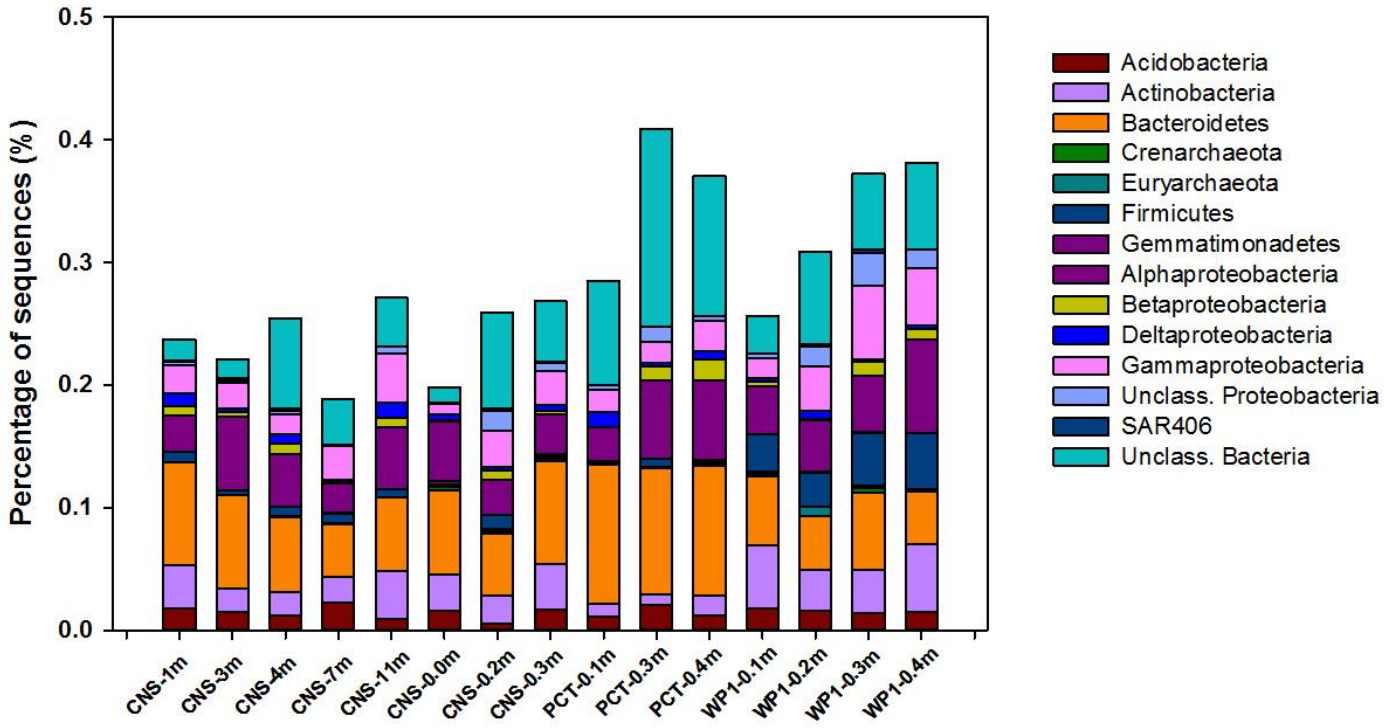


Figure 4.4c Relative abundance of the rare OTUs (< 0.0099 % of the sequences) grouped by Phylum or Class (same color scheme as Fig. 4.4a, b) to show degree of resemblance to abundant OTUs at the indicated taxonomic level. Sample IDs include the location in capital letters (CNS: Mt. Conness, PCT: Pacific Crest Trail, WP1: Wheeler Peak) followed by distance from one end of the horizontal transect (1 m, 3 m, 4 m, 7 m, 11 m) or by snow pit depth. Depths represent the bottom of the depth range, except for CNS-0.0 m that represents the upper end. Note: the % of Gemmatimonadetes was negligible, and color of the bars cannot be distinguished.

11.7 ± 8.4 %. The main different between the two Sierra sites was the abundance of OTU1, which in the PCT site was ~2.4-fold higher than at CNS.

Most OTUs in the meso-abundant group were 'singleton OTUs', and the taxonomic composition of meso-abundant OTUs was different from the other two groups of OTUs (Fig. 4.4b). OTUs in this group were consistently related to the phylum Bacteroidetes, the same phylum containing the core OTUs, in which they comprised 0.24-0.66 % of the sequences. Additionally, there were OTUs related to Gammaproteobacteria present in all samples (0.13-0.82 % sequences).Also, in WP1, OTUs related to the Betaproteobacteria were most abundant in comparison to the other OTUs (1-3 % sequences).

In the group of rare OTUs (Fig. 4.4c), there was a much lower percentage of sequences related to the Betaproteobacteria (0.001-0.02 % sequences), but OTUs related to the Bacteroidetes were consistently present across samples, in addition to Acidobacteria, Actinobacteria, Alphaproteobacteria and Gammaproteobacteria. Some other groups that were not detected in the other two groups appeared in this group, the Deltaproteobacteria, SAR406 and OTUs related to phyla of the domain Archaea.

Sequence coverage was > 96 % for all the samples, as determined using Good's coverage (Good). The bacterial community presented low alpha-diversity, and the inverse Simpson's index (Table S4.1) indicated that the diversity within samples was low (CNS: 7.6-15.7; PCT: 11.4-12.3; WP1: 5.6-18.8; CNS-Transect: 7.4-

16.2). The sample collected at meter 4 was the most diverse (Inverse Simpson's Index of 16.16) on the CNS transect, while the 0.2 m horizon in the CNS profile had highest diversity (Inverse Simpson's Index of 15.66) in the CNS vertical profile. However, the deepest (0.4 m) samples in the PCT and WP1 were the most diverse in those profiles (Inverse Simpson's Index of 12.31 and 18.80 respectively). Intra-sample rarefaction curves for the observed number of OTUs reflected similar results (Fig. S4.3). Furthermore, there were low levels of similarity at the OTU level (Fig. 4.5) among samples. Similarity between CNS and PCT was 23 to 29 %, between CNS and WP1 it was 8 to13 %, and between PCT and WP1 it was 5 to 27%, in all cases increasing with depth. The Unifrac weighted approach was used to test the probability that the dissimilarity among samples was not greater than expected by chance. Results from this test indicated that this probability was low, and that dissimilarity among vertical profile samples from the different sites was significant for all three sites ($p < 0.001$).

3.4. Factors that drive community assembly

The most important questions we intended to answer regarding microbial community diversity and structure in summer snow with varying levels of snow algae was whether there were specific relationships between algal signals and certain groups of bacteria, or if there was a relationship that leads to higher diversity levels (Connell & Orias, 1964). Alternatively, the prevalence of certain bacterial groups over others may be the result of the effect of the environment

(selection by environmental factors) on the entire microbial community (Bacteria and Eukaryotes).

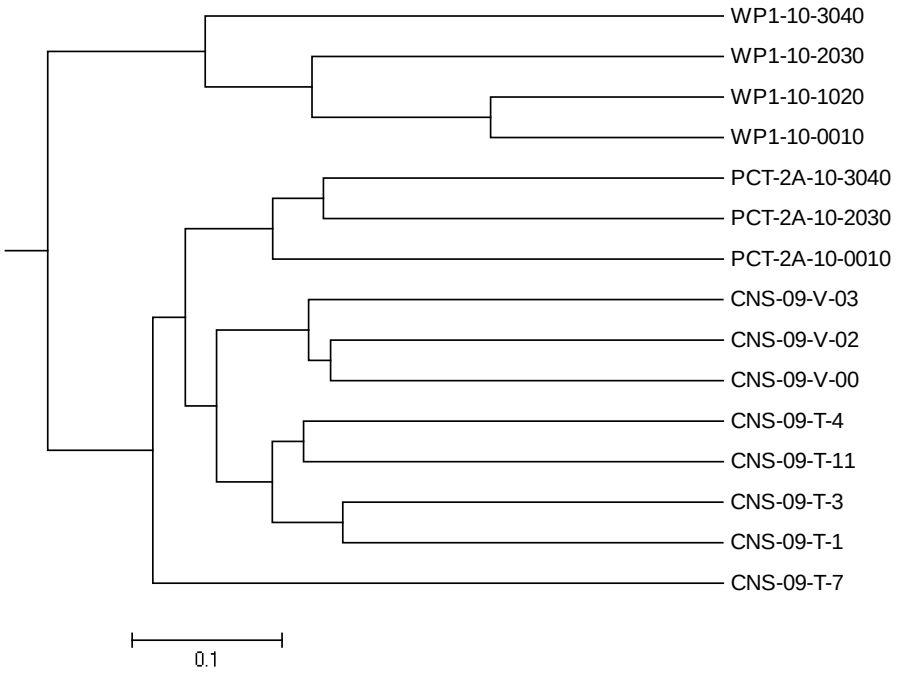


Figure 4.5 Dendrogram of Bray-Curtis dissimilarities clustered with the unweighted pair group method using the arithmetic averages (UPGMA) algorithm.

The relationship between algae and bacteria was first tested by studying the Spearman's rank correlation between Chl*a* (as a proxy for green algae in the Class Chlorophyceae) and the inverse Simpson's diversity index (Fig. S4.4). A negative relationship between snow samples with higher algal signals and bacterial diversity was observed in the vertical profiles from CNS and WP1, where Chl*a* data was obtained. There was a significant negative linear relationship ($p < 0.05$) in

which diversity decreased, while at the same time Chl*a* increased. Furthermore, in WP1 diversity clearly increased with depth (Table S1) while Chl*a* decreased in concentration (Fig. S4.7). A similar pattern was observed for CNS, but not as strong as in WP1.

To further test the hypothesis that there was a relationship between specific algal and bacterial OTUs we used the iTag sequencing data of the SSU rRNA of Eukayotes that were obtained for some of the samples and analyzed by JGI (CNS-4m, CNS-7m, CNS-0.2m, CNS-0.3m, and all the samples from PCT and WP1) in parallel with the bacterial data set described here. Spearman's rank correlations were calculated with the most abundant bacterial OTUs (≥ 1 % of sequences or more) and 8 OTUs affiliated with the Chlorophyceae class of green algae in the Eukaryota iTag data set. A positive correlation ($r > 0.693$, $p < 0.05$) was found between core OTUs (OTU 1, 3, 6 and 8) and some of the algae-related OTUs (Fig. S4.5). Additional positive correlations were found with other abundant bacterial OTUs, most of which were found in higher proportion in WP1 than in the other two sites. Relationships observed (Fig. S4.5): (i) positive relationships, *e.g.* OTU65397 increased in abundance with the increase in abundance of a Chlamydomonadaceae-related iTag (OTU000033; $r=0.937$, $p< 0.000$); (ii) Bacterial OTUs co-occurred with chlorophyte OTUs following a "threshold" of bacterial OTU occurrence. For example, OTU00001 (affiliated with Saprospirae) was present at levels between 5000 to 10,000 tags with little corresponding chlorophyte signal, then increased in samples where the abundance of chlorophyte OTUs related to the families

Chlamydomonadaceae (OTU000067; r=0.706, p= 0.015) and Mamiellophyceae (OUT000112; r=0.814, p=0.002) were detected at high levels ~ 20,000 tags. OTU3 had a similar pattern, in which the iTag abundance increased once a *Chloromonas nivalis*-related OTU (OTU000037) was at a level > 15,000 tags. (iii) OTU4638 (affiliated to Comamonadaceae) appeared to be inversely related (r= -0.956, p= 0.044) to the abundance of *Chloromonas hohamii* (OTU000039).

To test the second hypothesis, the relative abundance of the same 'core OTUs' was used for Spearman's rank order correlations with inorganic nutrients, Chl*a*, DOC, POC, and PN from the CNS transect. Only two significant correlations were found, a correlation between OTU8 (affiliated with Saprospirae) and PN (r= 0.9; *p = 0.037*), and between OTU2 (affiliated with Cytophagia) and DOC (r= 0.9; *p = 0.037*).

4. Discussion

4.1. Cell counts from Mt. Conness

Mean algal concentration on the top 10 cm of the snowpack in the east drainages of Mt. Conness was similar to those estimated by Painter et al. (2001) using imaging spectrometry (*e.g.*, 2.5 x 10^3 cells mL^{-1}). Thomas & Duval (1995) studied standing crops of algae and bacteria at locations nearby during the late spring of 1994, and the numbers of bacterial cells they encountered were greater than the numbers of algae, particularly in snow samples where algal blooms were observed. Similar

findings were reported in our study, even though the numbers of algal cells obtained were one order of magnitude smaller than the numbers obtained here. In comparison to another site in the Southern Hemisphere, algal cell counts were an order of magnitude lower than the cell concentration observed for snow algae in a Patagonian glacier, where cell numbers also decreased with altitude. The authors attributed this trend to the occurrence of an earlier snowmelt at lower altitudes, leading to higher exposure to the harsh environmental conditions during the reproductive season, and to lower nutrient content in their sites (Takeuchi & Kohshima, 2004).

The snow in CNS contained bacterial cell counts that were well within the range of bacterial abundance reported for other snow samples from the Antarctic, the Arctic, and other high mountain regions including the Tibetan Plateau (Amato et al., 2007; Liu et al., 2009; Segawa et al., 2005; Zhang et al., 2010). Bacterial abundance was similar throughout the vertical profile, while numbers of algal cells showed a sharp drop from the surface to the greater depths, as Thomas (1972) previously reported. Additionally, bacterial counts were higher than the number of algal cells at all the depths at CNS, suggesting that there may have been ample nutrients to support growth in all the snow layers. Note that growth rates were not directly determined. However, this trend of higher abundance of bacteria than snow algae in CNS is inverted relative to what other authors found in nearby locations, where algae in red snow were tens to thousands of times more abundant than bacteria in red snow, and 6 to 32 times more abundant than bacteria in white snow;

and both algae and bacteria were more abundant in red snow than in white snow (Thomas & Duval, 1995). Studying succession through the stages of snow algal bloom development and demise is likely required to best address these dynamics.

4.2. Snowpack dynamics and abiotic parameters related to the microbial community

The snowpack is a heterogeneous environment, composed of grains of snow in a range of sizes, ice layers, and different water saturation levels, all of which reduce the permeability to air and water (Larose et al., 2013a). Major mobilization of nutrients and particles only happens during snowmelt, late in the spring and early in the summer, when the temperature of the snowpack pack rises above 0 °C, or when there is intense solar radiation (Larose et al., 2013a; Kuhn, 1987). The first 30 % of meltwater removes about 80 % of solutes from the snowpack via percolation of meltwater, and repeated melting and freezing increases solute concentration, especially in the deeper layers of the snowpack (Larose et al., 2013a).

Snow algae are known to thrive only in wet snow that lasts throughout the summer months, as well as in old summer snow that has undergone cycles of freeze-thaw. This old summer snow contains snow crystals surrounded by a continuous water film, thick enough to create a microenvironment around the algae (Remias, 2012). The thin water layer between the algae and the snow crystals is the main source of inorganic nutrients, and inorganic elements that accumulate on cell surfaces may be taken and stored inside the cells (Lütz-Meindl & Lütz, 2006). The red-pigmented snow algae represent the climax stage of the life cycle of green algae

(Remias, 2012), in which the accumulation of carotenoids has been attributed not only to high levels of radiation in the surface snow, but also to N-, P-, S-, K- and Fe-deficiency (Bidigare et al., 1993; Muller et al., 1998). However, Kviderova (2010) found that phosphorous input from surrounding vegetation led to algal blooms in snowfields of the Giant Mountains in the Czech Republic. Algae can also deplete the nutrient content in the snow, particularly the pool of nitrogen, which has been correlated with lower conductivity in the snow (Hoham, 1989; Hoham et al., 1993; Jones, 1991).

Our samples were collected during the summers (early July) of 2009 and 2010 in areas where visible red snow algae were present (*e.g.* vertical profile from all three sites) in addition to surrounding areas with low concentrations of snow algae (some samples in the CNS transect). Considering the low the levels of nutrients in CNS samples (most samples were far from the Redfield ratio)in comparison to other snow samples adjacent to green snow algae (*e.g.* 0.13 mg L^{-1} O-PO_4, 31.24 mg L^{-1} NO_3^--N, 0.049 mg L^{-1} NH_3-N; Hoham & Mullet, 1977), it is likely that the CNS samples had undergone several rounds of snow melt, perhaps also explaining the spikes of DOC, NO_3^--N and NH_3-N in the deepest samples. The concentration of nutrients in the transect samples was also, in general, much lower than the concentration of the same nutrients measured by Kviderova (2010). For instance, this study of snow algae in the Giant Mountains, Czech Republic (June 2008) detected 1,000 to 3,000 times more Chl*a* than we did in areas covered with snow algae, although algal cell counts also were two orders of magnitude higher in

that study. Additionally, Thomas & Duval (1995), detected much higher levels of Chl*a* (200 μg L^{-1}) in patches of snow algae in the area around Mt. Conness.

Lack of correlation between the environmental factors measured in CNS and the cell abundance, suggested that other factors (*e.g.,* pH, altitude, slope, time of sampling) may have a significant effect on the abundance of microorganisms in the snow. However, abundance is perhaps not the most indicative proxy for use in determining snow ecological relationships; functional ecology and species-specific interactions stand more informative. Therefore, we tested the relationship between Chl*a* concentrations and diversity in the vertical profiles from CNS, and a site adjacent to WP1.

4.3. Diversity of the microbial community in comparison to other places

MiSeq next-generation sequencing of the SSU rRNA gene indicated that the microbial community inhabiting the snow presented low levels of diversity, and it was composed of only 4 abundant phyla of organisms. These phyla corresponded to the same groups reported by other studies of snow samples from all over the world, and include the Proteobacteria (Classes: Betaproteobacteria, Alphaproteobacteria, and Gammaproteobacteria), Bacteroidetes, Actinobacteria, and Cyanobacteria (Harding et al., 2011; Liu et al., 2009; Fujii et al., 2010; Larose et al., 2010; Moller et al., 2013). Archaea were only present in the rare OTUs data set and were never more than 0.007 % of the sequences in a given sample. Even at higher taxonomic levels there was correspondence with groups reported by other studies, for

instance, *Hymenobacter*-related phylotypes (Phylum Bacteroidetes) were the dominant group detected in Antarctic snow (Fujii et al., 2010), and were among the most abundant OTUs detected in WP1 surface samples. This group is also commonly found in other kinds of environments, of which air samples (Buczolits et al., 2002) are of interest to this study. Some of the genera that were observed in glaciers on the Tibetan Plateau (Liu et al., 2009; Zhang et al., 2010), like *Sphingomonas* (Phylum Alphaproteobacteria)*, Polaromonas,* and *Rhodoferax* (these last two in the family Comamonadaceae, Order Burkholderiales, Class Betaproteobacteria), were also observed in this study, and the structure of the microbial community described in those glaciers also showed variation in the relative abundance of these groups among different glaciers. Similar to this study, Harding et al. (2011) also reported high abundance of OTUs related to the order Burkholderiales, constituting 64 % of the bacterial clone sequences detected in High Arctic snow, in addition to OTUs related to the Bacteroidetes. Interestingly, the genus *Herminiimonas* in the family Oxalobacteraceae (Order Burkhorderiales) contains bacteria that have been isolated from spring and mineral waters, but have not been reported in other snow ecosystems to the best of our knowledge.

4.4. Theory of community assemblage and sources of microorganisms

According to Vellend's framework for community assembly (Nemergut et al., 2013) there are two forces that may bring new organisms into communities: speciation and dispersal. These can be contrasted with processes such as ecological drift and

selection by environmental factors that may affect presence/absence, as well as the relative abundance of microorganisms over time.

In this study, we hypothesized that in the snow ecosystem, the presence of snow algae may be an important factor driving bacterial community structure. Snow algae may provide a protective layer against the high irradiation at the snow surface (Remias, 2012). Furthermore, they are considered essential primary producers in the ecosystem, and CO_2 sinks (Painter et al., 2001; Williams et al., 2009), which may favor the presence of groups of heterotrophic microorganisms. Thomas & Duval (1995) detected more organic matter in red snow than in white snow, which was favorable for bacterial growth in the late summer, supporting the hypothesis that snow algae provide autochthonous sources of carbon for bacteria. Also, Fujii et al. (2010) suggested that bacteria associated with snow algae in Antarctica are secondary inhabitants of the snow. Our study showed that there was an inverse relationship between Chl*a* concentration in the CNS and WP1 vertical profiles and bacterial diversity, which decreased with higher Chl*a* concentrations, indicating that only certain groups of bacteria were favored by the presence of the algae. Furthermore, we found that certain bacterial OTUs were correlated with certain algal OTUs, supporting the idea that snow algae may create a microenvironment that is favorable for other microorganisms to grow.

We also tested the hypothesis that environmental factors could explain the relative abundance of the core bacterial OTUs, however we only found two significant correlations between an OTU affiliated to Saprospirae with PN, and

between an OTU affiliated to Cytophagia with DOC. Saprospirae and Cytophagia are unicellular gliders that may be able to degrade complex organic matter (Reichenbach, 2006). One aspect of microbial communities in snow ecosystems is the patchiness of its distribution, which may explain the overall lack of correlations observed here with environmental parameters. This kind of distribution could be due to eolian transport (Harding et al., 2011), because snowpacks receive inorganic nutrients (*e.g.*, CO_2, N_2O, salt particles, mineral dust, and sulfates), organic matter (*e.g.* pollen and dead foliage), and contaminants by wet and dry deposition, *i.e.*, either incorporated into falling precipitation or transferred directly to the snow surface (Larose et al., 2010).

Wet and dry deposition may drive community assembly by dispersal processes. Part of this deposition may be spore-forming bacteria, gram-positive bacteria, and fungi that are adapted to mobilization via aerial transport (Harding et al., 2011; Brinkmeyer et al., 2003). Segawaet al. (2005) also observed that a great proportion of bacteria detected in surface snow samples from the Tateyama Mountains in Japan, was delivered by mineral particles; and similar results were reported by Liu et al. (2009) for the Guoqu Glacier on the Tibetan Plateau. Harding et al. (2011) sampled air and snow in the High Arctic, and reported 71 SSU rRNA gene clones from air samples related to Bacteroidetes, Acidobacteria, Firmicutes, Alpha-, Beta- and Gamma-proteobacteria. Among these clones, the genera *Rhodoferax* and *Polaromonas* were detected in both snow and air samples, suggesting adaptation of these groups to aerial distribution, which makes them very

ubiquitous. Both were found in our study – in particular, microdiverse *Polaromoas*-affiliated iTags were among the dominant OTUs at all sites (up to 32 % of the sequences in WP1). Darcy et al. (2011) have suggested a mechanism by which bacteria in the genus *Polaromonas* may withstand the stresses of atmospheric transport.

Different processes may be operating in conjunction and at different scales to determine community structure, for instance, a combination of dispersal and selection by environmental factors may explain the observed distribution of 'singleton OTUs'. The singleton OTUs could also be a result of stochastic processes influencing the relative abundance of these groups; features such as weak selection, low alpha diversity, in addition to low numbers of community members have been described as characteristic features leading to ecological drift in macro-organisms (Chase & Myers, 2011). There is a possibility that the bacterial assemblage may be enabling algal growth as well, and an alternative mechanism by which bacteria might create a favorable microenvironment for algae should be investigated.

4.5. Future directions and closing remarks

To properly test hypotheses regarding mechanisms that may lead to snow community assembly and to analyze the relationship between environmental parameters and indices of alpha and beta diversity of the microbial community, more extensive spatial and temporal data sets should be examined with multivariate statistical approaches. This may lead to formulation of possible explanations of how

geographically distant regions share the same taxonomic groups, such as we found in this study. What is more, the uniqueness of the microbial community structure observed in this study, demands further investigation, to explain both the resemblance in taxonomic composition between the abundant and the rare OTUs, and the differences in structure between these two groups of OTUs. These differences included the high levels of dissimilarity at the OTU level between the snow samples.

Moreover, other groups of microorganisms in the kingdom Eukarya co-inhabit the snow with green algae, including other unicellular flagellates and ciliates, fungi, and rotifers (Segawa et al., 2005; Harding et al., 2011). To obtain a more complete picture of the structure of the microbial community associated to snow algae, and to identify the specific groups of algae that constitute the red snow, analysis of the sequenced V4 region of the SSU rRNA gene of the snow eukaryal assemblage in parallel with the bacterial and archaeal dataset will offer a significant layer of information. The next stage of this work involves analyses of these data using similar procedures, and the microbial assemblage described here will be re-analyzed in the context of those microorganisms.

Future studies of the microbial community inhabiting snow fields flourishing with snow algae should encompass the temporal aspect of these ecosystems, and sampling of possible sources of microorganisms, like air and soil. We also recommend to sample i) the snowpack at its highest point of accumulation, just before snowmelt starts, and several times during the snowmelt season; ii) to collect

meltwater as well as snow because it should incorporate most taxa that were present in the snow that season, as well as the soil in contact with the snow (Larose et al., 2010); and iii) to analyze multiple environmental parameters of the snow for the different kinds of samples.

Acknowledgements

We appreciate the guidance of KP Hand and TH Painter, Jet Propulsion Laboratory (JPL), California Institute of Technology, and the field teams at CNS in 2009, and members of the Nevada Geobiology Summer Short Course in 2010 that was led by CH Fritsen and D Crowther. We also appreciate field assistance and technical support provided by Vivian Peng. Financial Support for this research was provided by the NASA Astrobiology Institute, Astrobiology of Icy Worlds program at JPL, NASA Cooperative Agreement/Grant # NNGO5G587H to D Crowther and CH Fritsen, and iTag sequencing was supported by an award to AE Murray, JGI-634 by the DOE Joint Genome Institute.

References

Amato P, Hennebelle R, Magand O, Sancelme M, Delort A-M, Barbante C, Boutron C, Ferrari C (2007) Bacterial characterization of the snow cover at Spitzberg, Svalbard. *FEMS Microbiology Ecology,* **59**, 255-264.

Bidigare RR, Ondrusek ME, Kennicutt MC, Iturriaga R, Harvey HR, Hoham RW, Macko SA (1993) Evidence for a photoprotective function for secondary carotenoids of snow algae. *Journal of Phycology,* **29**, 427-434.

Brinkmeyer R, Knittel K, Jurgens J, Weyland H, Amann R, Helmke E (2003) Diversity and structure of bacterial communities in arctic versus antarctic pack ice. *Applied and Environmental Microbiology,* **69**, 6610-6619.

Buczolits S, Denner EBM, Vybiral D, Wieser M, Kampfer P, Busse HJ (2002) Classification of three airborne bacteria and proposal of Hymenobacter aerophilus sp nov. *International Journal of Systematic and Evolutionary Microbiology,* **52**, 445-456.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, Mcdonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods,* **7**, 335-336.

Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical transactions of the Royal Society B: Biological sciences,* **366**, 2351-2363.

Connell JH, Orias E (1964) The ecological regulation of species diversity. *American Naturalist,* **98**, 399-414.

Darcy JL, Lynch RC, King AJ, Robeson MS, Schmidt SK (2011) Global distribution of polaromonas phylotypes - evidence for a highly successful dispersal capacity. *Plos One,* **6**, e23742.

Delong EF (1992) Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences of the United States of America,* **89**, 5685-5689.

Duval B, Hoham RW (2000) Snow algae in the northeastern US: Photomicrographs, observations, and distribution of Chloromonas spp. (Chlorophyta). *Rhodora,* **102**, 365-372.

Fujii M, Takano Y, Kojima H, Hoshino T, Tanaka R, Fukui M (2010) Microbial community structure, pigment composition, and nitrogen source of red snow in antarctica. *Microbial Ecology,* **59**, 466-475.

Good IJ (1953) The population frequencies of species and the estimation of population parameters. *Biometrika,* **40**, 237-264.

Greenberg A, Clesceri LS, Eaton AD (1992) Standard methods for the analysis of water and wastewater. *American Public Health Association*.

Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C (2014) Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers. *PLoS ONE,* **9**, e87624.

Harding T, Jungblut AD, Lovejoy C, Vincent WF (2011) Microbes in High Arctic snow and implications for the cold biosphere. *Applied and Environmental Microbiology,* **77**, 3234-3243.

Hoham R, Laursen A, Clive S, Duval B (1993) Snow algae and other microbes in several alpine areas in New England. In: *Proceedings of the 50th Annual Eastern Snow Conference,* pp. 165-173.

Hoham RW (1989) Snow as a habitat for microorganisms.

Hoham RW, Mullet JE (1977) The life history and ecology of the snow alga Chloromonas cryophila sp. nov.(Chlorophyta, Volvocales)*. *Phycologia,* **16**, 53-68.

Hugoni M, Taib N, Debroas D, Domaizon I, Dufournel IJ, Bronner G, Salter I, Agogué H, Mary I, Galand PE (2013) Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proceedings of the National Academy of Sciences,* **110**, 6004-6009.

Jones H (1991) Snow chemistry and biological activity: a particular perspective on nutrient cycling. In: *Seasonal Snowpacks*. Springer, pp. 173-228.

Karl D, Dore J, Hebel D, Winn C (1991) Procedures for particulate carbon, nitrogen, phosphorus and total mass analyses used in the Us-JGOFS Hawaii ocean time-series program. In: *Marine particles: Analysis and characterization* (eds Hurd D, Spencer D). American Geophysical Union, Washington, D. C., pp. 71-77.

Kawecka B, Drake B (1978) Biology and ecology of snow algae 1. The sexual reproduction of Chlamydomonas nivalis (Bauer) Wille (Chlorophyta, Volvocales). *Acta Hydrobiologica,* **20**, 111-116.

Komárek J, Nedbalová L (2007) Green cryosestic algae. In: *Algae and cyanobacteria in extreme environments*. Springer, Netherlands, pp. 321-342.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology,* **79**, 5112-5120.

Kuhn M (1987) Micro-meteorological conditions for snow melt. *Journal of Glaciology,* **33**, 24-26.

Kviderova J (2010) Characterization of the community of snow algae and their photochemical performance in situ in the Giant Mountains, Czech Republic. *Arctic Antarctic and Alpine Research,* **42**, 210-218.

Larose C, Berger S, Ferrari C, Navarro E, Dommergue A, Schneider D, Vogel TM (2010) Microbial sequences retrieved from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway. *Extremophiles,* **14**, 205-212.

Larose C, Dommergue A, Vogel T (2013a) The dynamic arctic snow pack: An unexplored environment for microbial diversity and activity. *Biology,* **2**, 317-330.

Larose C, Prestat E, Cecillon S, Berger S, Malandain C, Lyon D, Ferrari C, Schneider D, Dommergue A, Vogel TM (2013b) Interactions between snow chemistry, mercury inputs and microbial population dynamics in an arctic snowpack. *Plos One,* **8**.

Liu Y, Yao T, Jiao N, Kang S, Xu B, Zeng Y, Huang S, Liu X (2009) Bacterial diversity in the snow over Tibetan Plateau Glaciers. *Extremophiles,* **13**, 411-423.

Lütz-Meindl U, Lütz C (2006) Analysis of element accumulation in cell wall attached and intracellular particles of snow algae by EELS and ESI. *Micron,* **37**, 452-458.

Lutz S, Anesio AM, Villar SEJ, Benning LG (2014) Variations of algal communities cause darkening of a Greenland glacier. *FEMS Microbiology Ecology,* **89**, 402-414.

Magurran AE (2003) *Measuring Biological Diversity,* John Wiley & Sons.

Massana R, Murray AE, Preston CM, Delong EF (1997) Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Applied and Environmental Microbiology,* **63**, 50-56.

Moller AK, Soborg DA, Abu Al-Soud W, Sorensen SJ, Kroer N (2013) Bacterial community structure in High-Arctic snow and freshwater as revealed by pyrosequencing of 16S rRNA genes and cultivation. *Polar Research,* **32**, 1-11.

Muller T, Bleiss W, Martin CD, Rogaschewski S, Fuhr G (1998) Snow algae from northwest Svalbard: their identification, distribution, pigment and nutrient content. *Polar Biology,* **20**, 14-32.

Murray A, Hollibaugh J, Orrego C (1996) Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Applied and Environmental Microbiology,* **62**, 2676-2680.

Muyzer G, Dewaal EC, Uitterlinden AG (1993) Profiling of complex microbial-populations by denaturing gradient gel-electrophoresis analysis of polymerase chain reaction-amplified genes-coding for 16s ribosomal-RNA. *Applied and Environmental Microbiology,* **59**, 695-700.

Nemergut DR, Schmidt SK, Fukami T, O'neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S (2013) Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews,* **77**, 342-356.

Painter TH, Duval B, Thomas WH, Mendez M, Heintzelman S, Dozier J (2001) Detection and quantification of snow algae with an airborne imaging spectrometer. *Applied and Environmental Microbiology,* **67**, 5267-5272.

Redfield AC (1934) *On the proportions of organic derivatives in sea water and their relation to the composition of plankton,* University Press, Liverpool.

Reichenbach H (2006) The order cytophagales. In: *The prokaryotes*. Springer, pp. 549-590.

Remias D (2012) Cell structure and physiology of alpine snow and ice algae. In: *Plants in alpine regions* (ed Lütz C). Springer Vienna, pp. 175-185.

Segawa T, Miyamoto K, Ushida K, Agata K, Okada N, Kohshima S (2005) Seasonal change in bacterial flora and biomass in mountain snow from the Tateyama Mountains, Japan, analyzed by 16S rRNA gene sequencing and real-time PCR. *Applied and Environmental Microbiology,* **71**, 123-130.

Sommerfeld RA, Musselman RC, Reuss JO, Mosier AR (1991) Preliminary measurements of CO2 in melting snow. *Geophysical Research Letters,* **18**, 1225-1228.

Takeuchi N, Kohshima S (2004) A snow algal community on Tyndall Glacier in the Southern Patagonia Icefield, Chile. *Arctic Antarctic and Alpine Research,* **36**, 92-99.

Thomas WH (1972) Observations on snow algae in California 1, 2. *Journal of Phycology,* **8**, 1-9.

Thomas WH, Duval B (1995) Sierra-Nevada, California, USA, snow algae - snow albedo changes, algal bacterial interrelationships, and ultraviolet-radiation effects. *Arctic and Alpine Research,* **27**, 389-399.

Weiss RL (1983) Fine structure of the snow alga (*Chlamydomonas nivalis*) and associated bacteria. *Journal of Phycology,* **19**, 200-204.

Welschmeyer NA (1994) Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. *Limnology and Oceanography,* **39**, 1985-1992.

Williams MW, Seibold C, Chowanski K (2009) Storage and release of solutes from a subalpine seasonal snowpack: soil and stream water response, Niwot Ridge, Colorado. *Biogeochemistry,* **95**, 77-94.

Zhang S, Yang G, Wang Y, Hou S (2010) Abundance and community of snow bacteria from three glaciers in the Tibetan Plateau. *Journal of Environmental Sciences-China,* **22**, 1418-1424.

Supplementary Tables and Figures

Table S4.1 Alpha diversity statistics of the snow samples.

Sample	Nº Sequences	Good's coverage	OTU richness (S_{obs})	Simpson's reciprocal Index (SRI)	SRI low confidence interval	SRI high confidence interval
CNS-1 m	103,385	0.97	4,293	13.44	13.29	13.60
CNS-3 m	141,637	0.98	3,615	7.47	7.40	7.55
CNS-4 m	119,606	0.96	5,753	16.16	16.00	16.32
CNS-7 m	110,130	0.96	5,309	13.88	13.75	14.01
CNS-11 m	132,940	0.97	5,559	13.22	13.12	13.33
CNS-0.0 m	127,275	0.97	4,295	7.57	7.48	7.66
CNS-0.2 m	121,526	0.96	5,892	15.66	15.50	15.82
CNS-0.3 m	121,133	0.97	5,053	12.77	12.64	12.90
PCT-0.1 m	124,141	0.98	3,483	11.43	11.33	11.53
PCT-0.3 m	115,174	0.97	4,603	11.95	11.80	12.11
PCT-0.4m	119,547	0.97	4,943	12.31	12.18	12.44
WP1-0.1	130,802	0.97	4,567	5.55	5.49	5.61
WP1-0.2 m	139,437	0.97	6,021	10.77	10.66	10.89
WP1-0.3 m	128,686	0.96	6,006	13.86	13.72	14.01
WP1-0.4 m	147,860	0.96	7,223	18.80	18.65	18.94

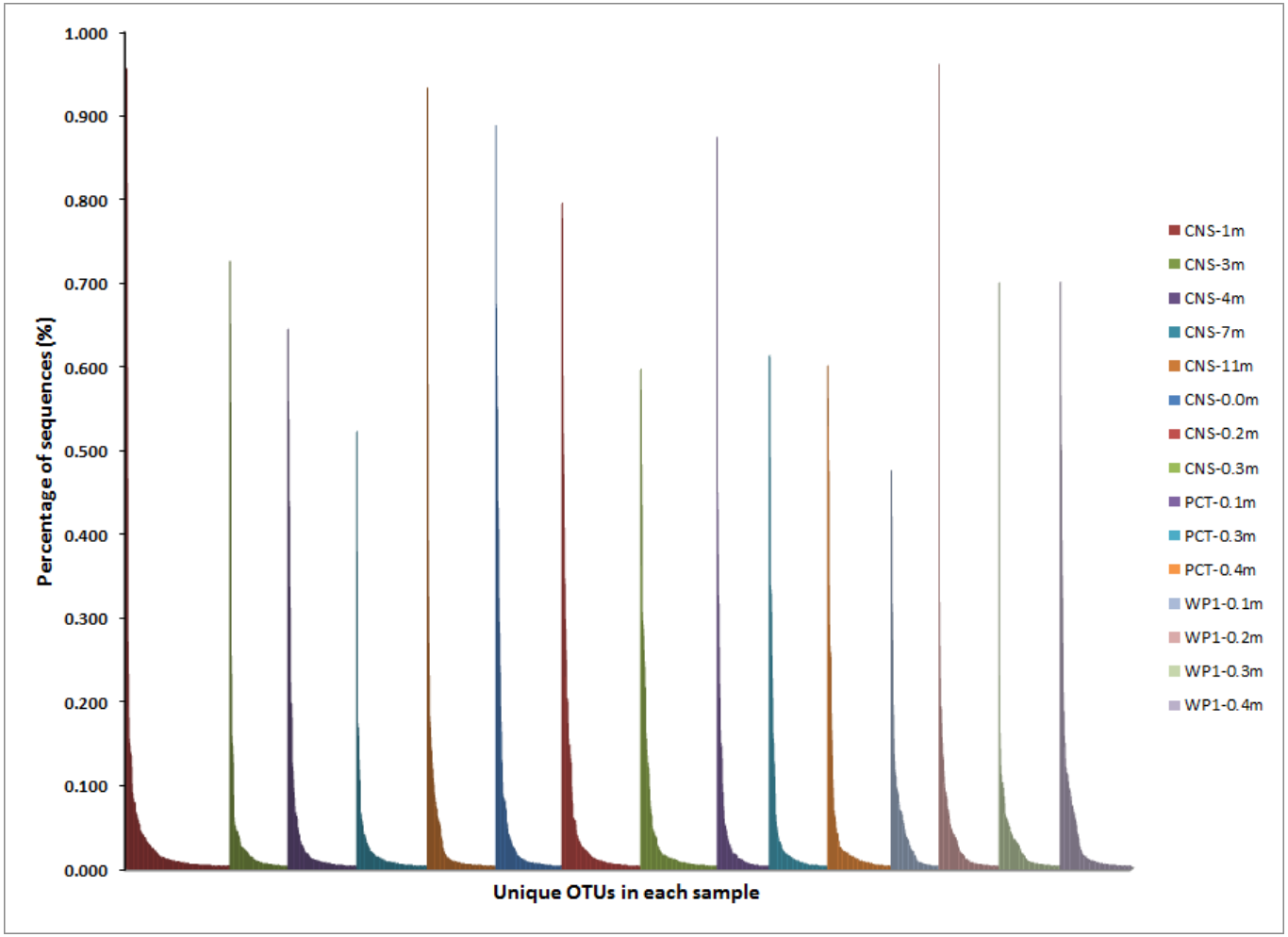


Figure S4.1 Abundance distribution of the 'singleton OTUs' comprising < 1 % of the sequences. Each bar represents a different OTU.

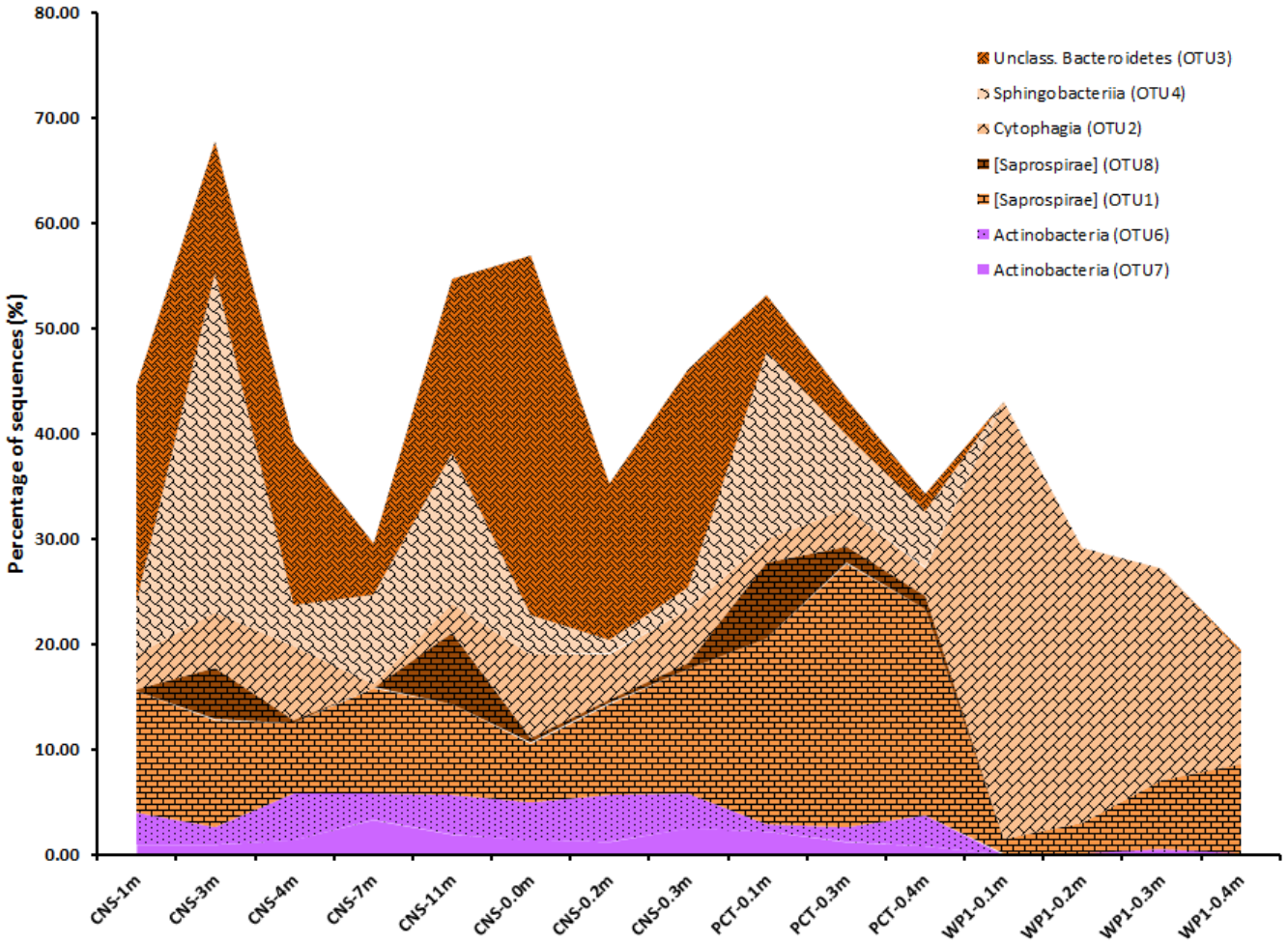


Figure S4.2 Relative abundance of the 'core' OTUs that comprised ≥ 1 % of the sequences and that were present in all the samples.Coloring scheme according to Figure 4.4a.

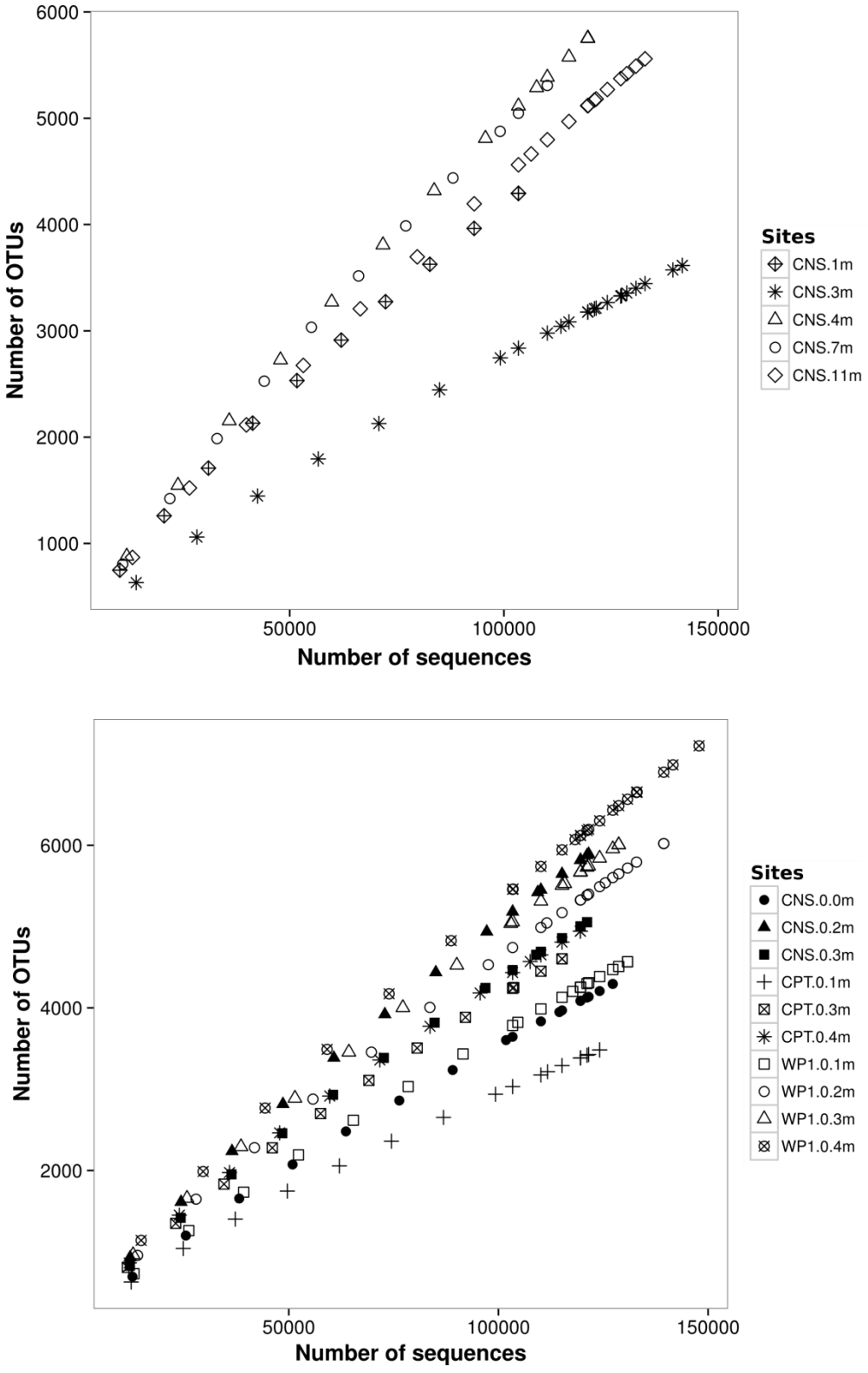


Figure S4.3 Intra-sample rarefaction curves for the observed number of OTUs.

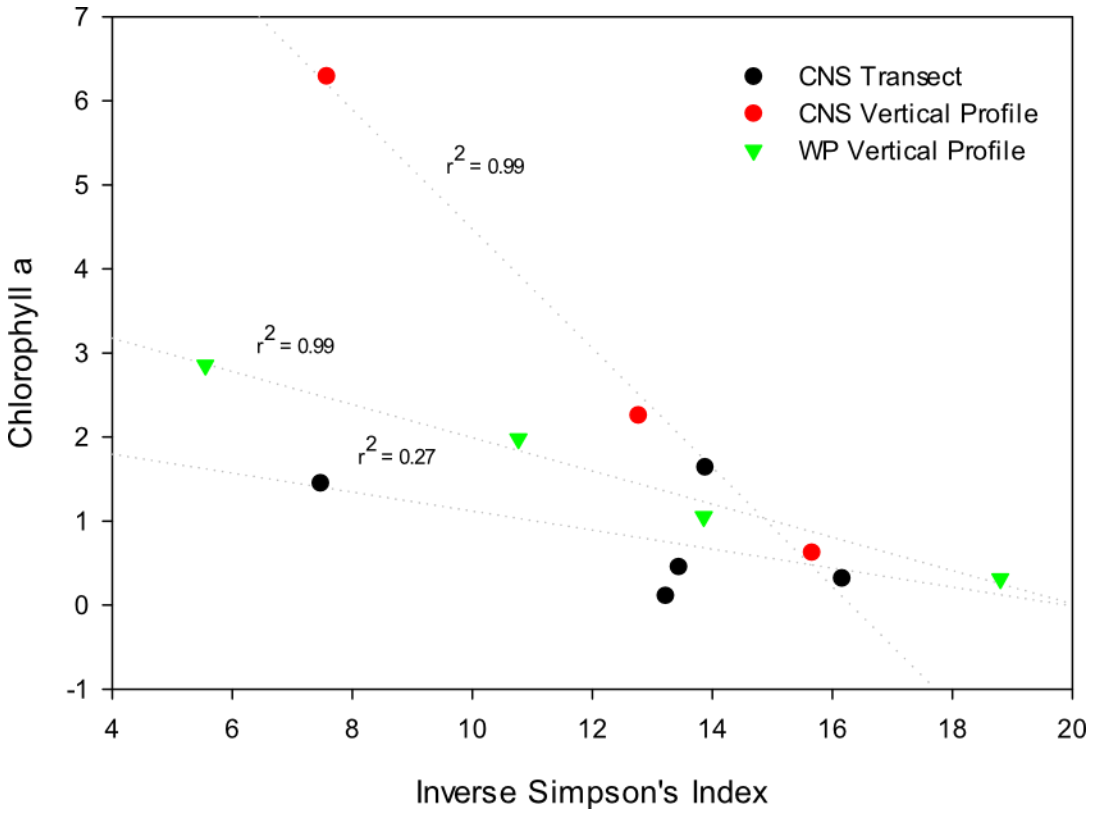


Figure S4.4 Spearman's rank correlation between chlorophyll a and the inverse Simpson's diversity index.

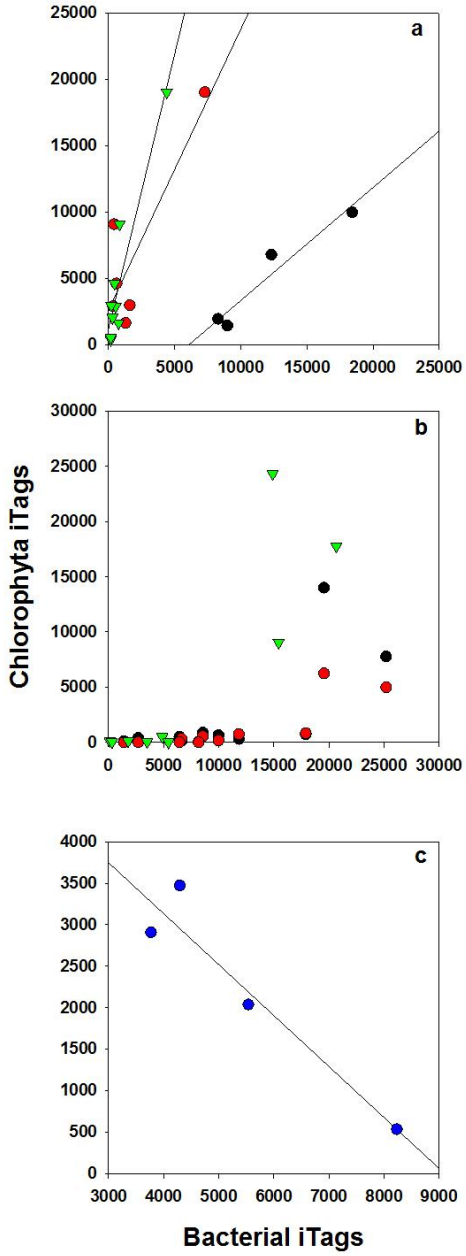


Figure S4.5 Spearman's rank correlations between algal OTUs and bacterial OTUs. a) Correlation (r=0.937, p< 0.000) between algal OTU33 (Chlamydomonadaceae-related iTag) and bacterial OTU65397 (*Polaromonas*-related iTag - Black circles, from WP1). b) Correlation between algal OTU67 (Chlamydomonadaceae-related iTag) and bacterial OTU1 (Saprospirae-related iTag - Black circles) (r=0.706, p= 0.015), algal OUT112 (Mamiellophyceae-related iTag) and bacterial OTU1 (Red circles) (r=0.814, p=0.002), and correlation between algal OTU37 (*Chloromonas nivalis*-related iTag) and bacterial OTU3 (Bacteroidetes-related iTag – Green triangles) (r=0.864, p=0.003). c) Correlation (r= -0.956, p= 0.044) between algal OTU39 (*Chloromonas hohamii*-related iTag) and bacterial OTU4638 (Comamonadaceae-related iTag – Blue circles).

Figure S4.6 Snow sampling sites. a) Mt. Conness (CNS) snow field; transect sampled is indicated with the black dashed line. b) Closer up view of the transect sampled with meter flags indicating sampling sites. c) Pacific Crest Trail (PCT) snow field. d) Close up look at visible snow algae patch showing melted zones near densely colored areas. e) Wheeler Peak (WP1) snow pit top of snow was 77 cm above melt layer interface.

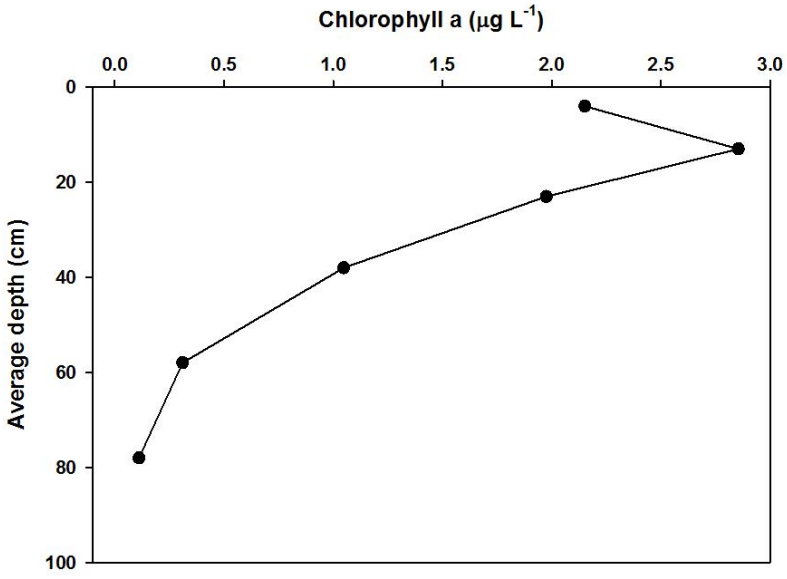


Figure S4.7 Wheeler Peak (WP) chlorophyll a concentration in the vertical profile.

Supplementary Information

S1. Supplementary Methods

S1.1 Amplification of the SSU rRNA gene of Bacteria, Archaea and Eukarya

To obtain a diversity profile of the microbial assemblage in the CNS samples, primers BactGC358f and Univ 517r were used to amplify the variable region 3 (V3) of the SSU rRNA gene, a ~220-240 bp fragment (Murray et al., 1996), including a 40 bp GC clamp added on to the forward primer. The PCR reaction mixture consisted of 1X PCR buffer with 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5 µM of forward and reverse primers, 1 µg μl^{-1} of BSA, 10 ng of template DNA, and 2.5 U of TaqDNA polymerase, in a 50 µl final reaction volume. Primers Euk GC960bf and Univ1200r with (Gast et al., 2004; Muyzer et al., 1993) were used to amplify the variable 6 (V6) region of the SSU rRNA gene, a ~ 300 bp fragment. PCR conditions were the same as above. The thermal-cycling program utilized to amplify the bacterial and eukaryal sequences was as follows: initial denaturation at 94 °C for 3 min, 10 cycles of 94 °C for 30 sec, touch-down primer annealing with the temperature decreasing from 65 °C to 55 °C (1 °C per cycle) for 30 sec, and 72 °C for 30 sec. Next, 18 cycles of 94 °C for 30 sec, 55 °C for 30 sec and 72 °C for 30 sec, and finally primer extension at 72 °C for 7 min. To screen for Archaea sequences, primers Arch20f and Univ958r were used (Delong, 1992) The expected size of the fragment was 950 bp. PCR conditions were the same as above, except that each reaction contained a 2 mM concentration of $MgCl_2$. Thermal cycling was done as follows: initial denaturation at 94 °C for 3

min, 38 cycles of 94 °C for 1 min, primers annealing at 55 °C for 1 min, 72 °C for 1 min 30 sec, and final extension at 72 °C for 7 min. Another set or primers for the variable region 3 of the SSU rRNA was used to screen for Archaea too, AR-GC_347F and UNIV519R. The PCR reaction mixture consisted of 1X PCR buffer, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5 µM of forward and reverse primers, 10 ng of template DNA, and 2.5 U of TaqDNA polymerase, in a 50 µl final reaction volume. The same touch-down protocol utilized to amplify bacterial and eukaryal DNA was applied here.

S1.2. Diversity profiling of CNS samples using DGGE

Bacteria and Eukarya PCR amplicons from 5 x 50 µl PCR reactions were pooled and precipitated over night at -20 °C with two volumes of absolute ethanol and 1/10 of the volume of 3 M sodium acetate pH 5. After centrifugation in cold for 30 min, the DNA was washed with 70% ethanol, centrifuged for 5 to 8 min, air-dried, or dried in a DNA speed vac (Savant), and resuspended in 16 µl of sterile DI water. DNA was quantified using fluorometry and the PicoGreen dye. DNA samples (600 ng of Eukarya and 800 ng of Bacteria) were loaded onto two different 8 % polyacrylamide-bisacrylamide (37.5:1) gels. A 25 to 75 % gradient of denaturants (7 M urea and 40 % deionized formamide in 100 % concentration of denaturants) was used for Eukarya DNA, and a 25 % to 60 % gradient was used for Bacteria DNA. Electrophoresis was run at 60 °C for 15 h and 30 min at 63 V (Eukarya) or 16 h at 63 V (Bacteria). To prepare these gels and run the electrophoresis methods from

Murray (1996) were followed. Band identification and analyses of the banding patterns resulting from DGGE were done with the Quantity One software (BioRad).

S2. Supplementary Results

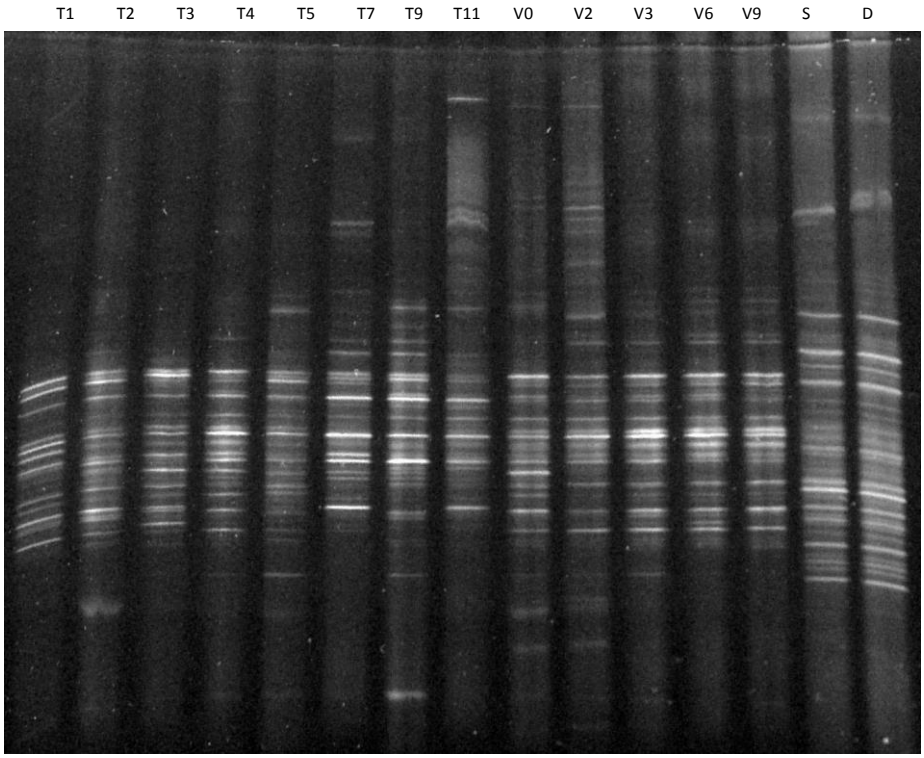


Figure S4.8a DGGE of PCR amplified fragments using general primers for Bacteria. (T) Horizontal transect. (V) Vertical profile. (T1) 1 m; (T2), 2 m; (T3) 3 m; (T4) 4 m; (T5) 5 m; (T7) 7 m; (T9) 9 m; and (T11) 11 m from one end of the transect. (V.0) Surface; (V.2) 0.2 m; (V.3) 0.3 m; (V.6) 0.6 m and (V.9) 0.9 m below surface. (S) Surface water sample from Alaska. (D) Deep water sample from Alaska.

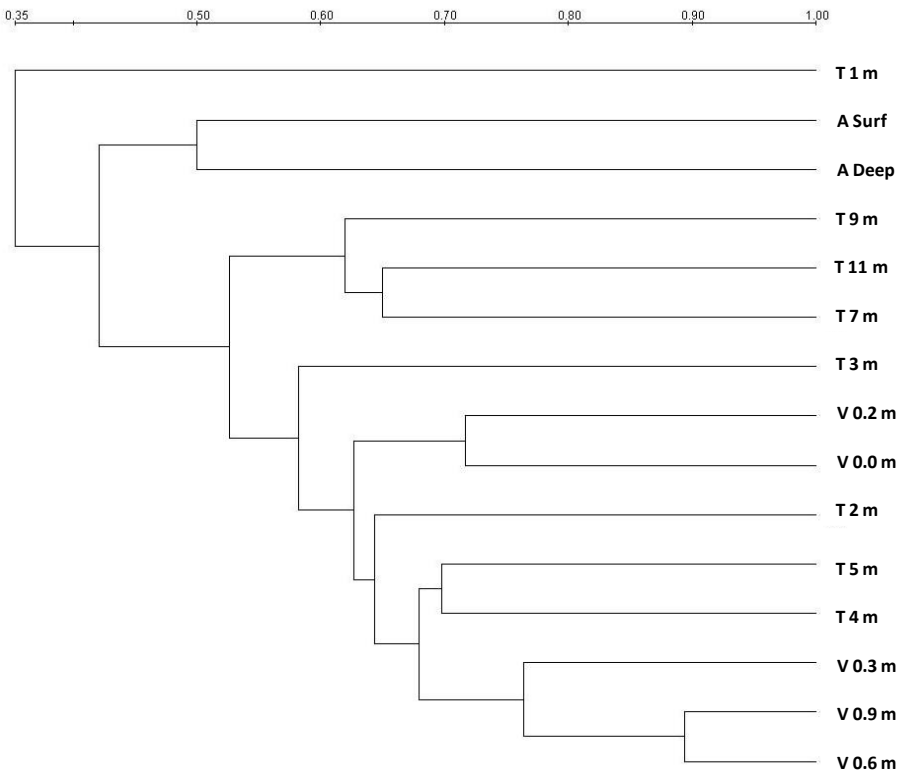


Figure S4.8b Dendrogram of phylotypes observed by DGGE analysis of the SSU rRNA gene of Bacteria. Generated by UPGAMA (unweighted pair group method using arithmetic averages) cluster analysis using the Quantity One software (BioRad). Similarity scale: 0.0 being no similarity, and1.0 being identical. (T) Horizontal transect. (V) Vertical profile. (T1) 1 m; (T2), 2 m; (T3) 3 m; (T4) 4 m; (T5) 5 m; (T7) 7 m; (T9) 9 m; and (T11) 11 m from one end of the transect. (V.0) Surface; (V.2) 0.2 m; (V.3) 0.3 m; (V.6) 0.6 m and (V.9) 0.9 m below surface. (A) Water samples collected in Alaska for comparison. (Surf) Surface. (Deep) Deep.

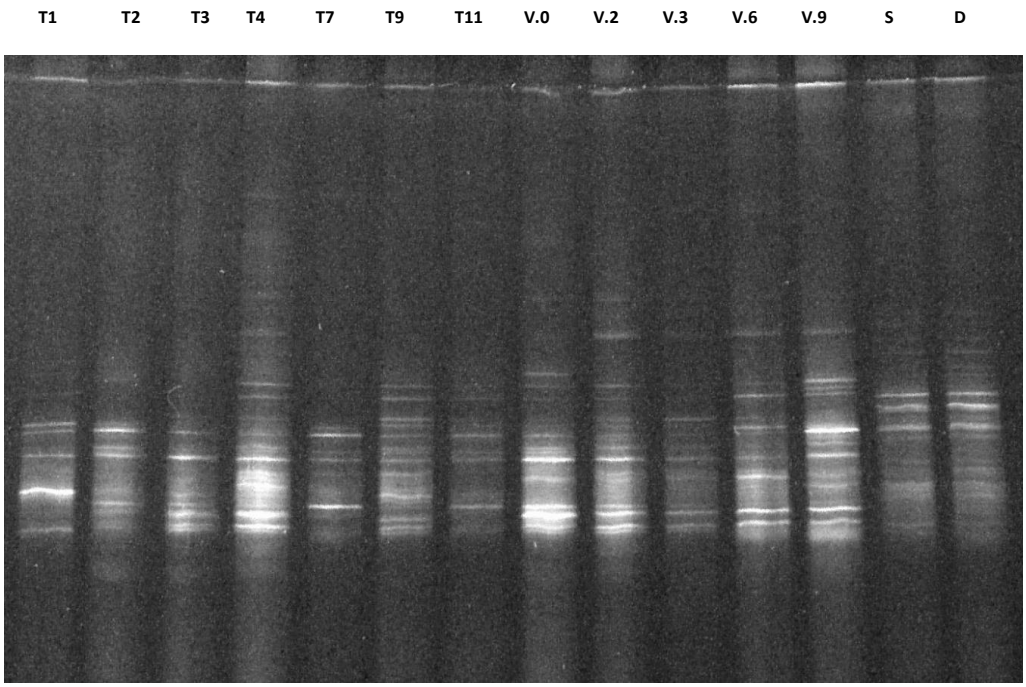


Figure S4.9a DGGE of PCR-amplified fragments using general primers for Eukarya. (T) Horizontal transect. (V) Vertical profile. (T1) 1 m; (T2), 2 m; (T3) 3 m; (T4) 4 m; (T5) 5 m; (T7) 7 m; (T9) 9 m; and (T11) 11 m from one end of the transect. (V.0) Surface; (V.2) 0.2 m; (V.3) 0.3 m; (V.6) 0.6 m and (V.9) 0.9 m below surface. (S) Surface water sample from Alaska. (D) Deep water sample from Alaska.

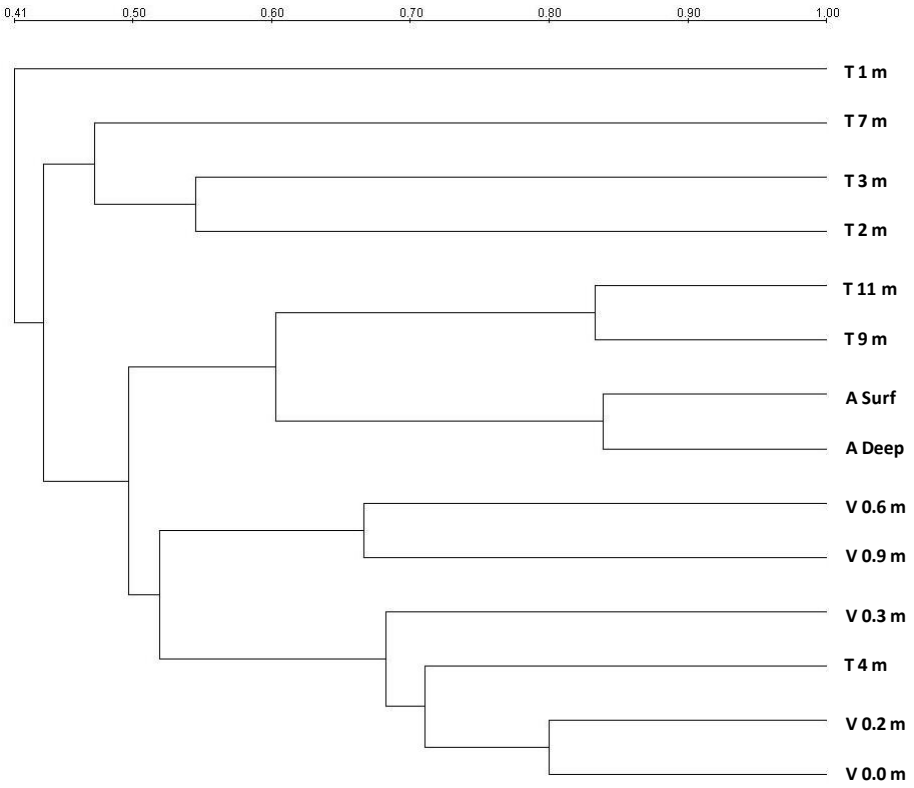


Figure S4.9b. Dendrogram of phylotypes observed by DGGE analysis of the SSU rRNA gene of Eukarya. Generated by UPGAMA (unweighted pair group method using arithmetic averages) cluster analysis using the Quantity One software (BioRad). Similarity scale: 0.0 being no similarity, and1.0 being identical. (T) Horizontal transect. (V) Vertical profile. (T1) 1 m; (T2), 2 m; (T3) 3 m; (T4) 4 m; (T5) 5 m; (T7) 7 m; (T9) 9 m; and (T11) 11 m from one end of the transect. (V.0) Surface; (V.2) 0.2 m; (V.3) 0.3 m; (V.6) 0.6 m and (V.9) 0.9 m below surface. (A) Sea water samples collected in Alaska for comparison. (Surf) Surface. (Deep) Deep.

Supplementary References

Delong EF (1992) Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences of the United States of America,* **89**, 5685-5689.

Gast RJ, Dennett MR, Caron DA (2004) Characterization of protistan assemblages in the Ross Sea, Antarctica, by denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology,* **70**, 2028-2037.

Massana R, Murray AE, Preston CM, Delong EF (1997) Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Applied and Environmental Microbiology,* **63**, 50-56.

Moller AK, Soborg DA, Abu Al-Soud W, Sorensen SJ, Kroer N (2013) Bacterial community structure in High-Arctic snow and freshwater as revealed by pyrosequencing of 16S rRNA genes and cultivation. *Polar Research,* **32**, 1-11.

Muller T, Bleiss W, Martin CD, Rogaschewski S, Fuhr G (1998) Snow algae from northwest Svalbard: their identification, distribution, pigment and nutrient content. *Polar Biology,* **20**, 14-32.

Murray A, Hollibaugh J, Orrego C (1996) Phylogenetic compositions of bacterioplankton from two California estuaries compared by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Applied and Environmental Microbiology,* **62**, 2676-2680.

Muyzer G, Dewaal EC, Uitterlinden AG (1993) Profiling of complex microbial-populations by denaturing gradient gel-electrophoresis analysis of polymerase chain reaction-amplified genes-coding for 16s ribosomal-RNA. *Applied and Environmental Microbiology,* **59**, 695-700.

Chapter 5

General Conclusions

What is life? Where did it come from? Where can we find it? We will continue the search for answers to these questions, not only as a quest for truth, but also as a quest for knowledge, because we- humans- yearn "to know", and because we want to be surprised by the answers to these questions.

1. Positive and negative feedback loops as a property of self-regulating systems

I started my dissertation, thinking that there was something very counterintuitive about finding methanogens in not cold, but extremely cold environments. These microorganisms of the domain Archaea excrete CH_4 as a way to reach chemical equilibrium, and can operate with amounts of energy that are barely enough to make 1 ATP molecule ($G°' = -31.8$ KJ mol^{-1}) (Madigan et al., 2014). CH_4 turns out to be a potent greenhouse gas (IPCC, 2007), though estimating total CH_4 emissions to the atmosphere is no easy task, considering the different sources and mechanisms for CH_4 formation. Water logged sediments of wetlands, rice paddies, and lakes provide anoxic environments with low redox potentials ideal for methanogenesis, and if in addition to that, there is high availability of labile organic matter, as in permafrost, all of these habitats are among the best for biological CH_4 production on the planet.

Thermokarst lakes on the North Slope of Alaska are frozen for months at a time, preserving for thousands of years organic matter that had only been partially degraded; and when permafrost thaws, this organic matter can be decomposed and utilized by microorganisms inhabiting the sediments. These lakes are also exposed

to 24 hours of sunlight during the summer, when temperatures are maintained well above 0 °C, which promotes primary production in the water column contributing fresh organic matter to the sediments as well. Moreover, thermokarst erosion at the bottom and edges of the lakes, releases more of the organic matter trapped in permafrost; and all of these features make of thermokarst lakes hotspots for CH_4 production (Walter et al., 2007).

In chapter 2 my research demonstrated that there is one lake (Siqlukaq) on the North Slope of Alaska where biological CH_4 production takes place at both, *in situ* sediment temperatures and at higher temperatures. Another lake (Sukok) is a conduit for thermogenic gas formed deep under the permafrost, and its sediments are composed of one-fourth of the organic matter detected in the surface sediments of Siqlukaq Lake. Surprisingly, we detected biological CH_4 production from another area (SukB) in Sukok Lake, when temperatures were increased 8-10 degrees above the *in situ* surface sediment temperatures. This result suggested that even though there were low amounts of organic matter in Sukok, microorganisms that are present in this lake can mineralize it under warmer conditions. Together, these findings led us to contend that under scenarios of climate change and warmer temperatures (also experienced during summer), there is a potential for increased biological CH_4 production in sediments of arctic thermokarst lakes on the North Slope of Alaska (Matheus Carnevali et al., 2015).

However, the same way there are processes that lead to CH_4 production, there are processes that lead to CH_4 consumption (He et al., 2012; Lofton et al.,

2014; Bretz & Whalen, 2014). Other members of our team detected CH_4 oxidation in the water column of the lakes, together with microorganisms that are able to carry out this process; and we detected the presence of aerobic type I methanotrophs among the bacteria OTUs, and anaerobic methane oxidizing archaea among the Archaea OTUs (Chapter 3). Future work on CH_4 emissions from these lakes will benefit from including CH_4 flux calculations that account for CH_4 production and consumption in each compartment of the lakes (*i.e.*, sediments, water column, and the ice layer). Also the understanding of CH_4 cycling should include sample collections across different seasons (*e.g.*, summer) and different areas (*e.g.*, shores) of the lakes. This information would enhance future scenario development and allow determining (i) if CH_4 emissions from arctic thermokarst lakes on the North Slope of Alaska have an impact on the warming trend that is currently observed in the Arctic, and (ii) what the consequences are for the greenhouse gas budget of our planet. It is important to remember that in addition to biological CH_4 production sensitivity to temperature, CH_4 consumption (oxidative processes), are also biological processes mediated by enzymes, and are responsive to temperature (Lofton et al., 2014). Therefore, it is possible that an increase in one process may be counteracted by an increase in the other. Chapter 2 of this dissertation constitutes a first approximation to the potential for biological CH_4 production from lakes on the Coastal Plain of Alaska, and I hope it will help constraining estimates of CH_4 emissions from the Arctic.

It is undeniable that the Arctic is changing; we can see it in pictures of icebergs breaking apart, permafrost slumping on the coastal areas, and changes in polar animals' behavior. However, the effects of climate warming go beyond what we can see with our own eyes, and may have deeper consequences that we anticipate. Microorganisms are the key players of biogeochemical cycling on the planet, and microbial interactions with essential elements in the environment (*e.g.*, C, N, S, Fe) determine their fate globally. Understanding the structure of microbial communities inhabiting climate-sensitive environments like permafrost, will allow us to predict how ecosystem functioning could be affected under different scenarios of climate change in the future.

2. Microbial communities in cold environments and their global distribution

2.1. Microbial communities in thermokarst lake sediments

Microorganisms that decompose organic matter in arctic thermokarst lakes set up the conditions for biological CH_4 production, a process that results from a chain of redox reactions that rely on a variety of electron donors and electron acceptors. These microorganisms are individual engines that act in concert, and as a whole, to regulate nutrient cycling on the planet. However, the diversity and structure of the microbial communities that inhabit different environments are determined in large part by the environment itself. Methanogens and their enzymes may be more efficient at warm temperatures, but the fact that there is active biological CH_4 production in the sediments of arctic thermokarst lakes indicates that these

enzymes must be adapted to function under the stresses imposed by cold temperatures as well. Additionally, these microorganisms may have developed mechanisms to remain inactive for long periods of time, as demonstrated by Rivkina et.al. (1998) who found viable methanogens in thousands of years old permafrost samples. This may also explain why we were able to obtain methanogens from enrichment cultures inoculated with sediments from Sukok Lake (Appendix), and why the *mcrA* gene was detected by PCR amplification of DNA extracted from this lake, although it was found at levels that were too low for our qPCR assay (Chapter 2).

In Chapter 3 we uncovered the phylogenetic and functional diversity of microorganisms inhabiting the sediments of thermokarst lakes, and demonstrated that the availability of certain carbon and energy sources in the sediments (*e.g.*, organic matter, sulfate, nitrate, and dissolved iron) was paralleled by the presence of microorganisms that can use these substrates and nutrients as electron donors or electron acceptors in a variety of metabolic reactions. Furthermore, we detected unexpected groups of microorganisms like cyanobacteria, which are obligate phototrophs and require sunlight as their primary energy source. This finding caused me to question the biological-geological-physical mechanisms that would allow these types of microorganisms to survive in darkness in the sediments of these shallow lakes. For example, can these organisms exist in dormant states? What is the life-span or turnover time of ribosomal RNA in cold and frozen sediments? What are the important scales (both physical and temporal) of thermokarst erosion?

And, what is the influence of wind on rearrangement of these shallow lake sediments?

Perhaps one of the main contributions from chapter 3 lies in the use of MiSeq next-generation sequencing to uncover the diversity and structure of the bacterial and archaeal assemblages in these lakes. This aspect of the biogeochemistry of the lakes had not been studied until now, and it allowed exposing striking differences in the structure of the archaeal assemblage in the two lakes. Clearly, methanogens play a bigger role in sediments of Siqlukaq, as determined by the abundance of ribosomal RNA, but what determines the almost absence of methanogenic SSU rRNA in the sediments of Sukok remains to be explained. Future studies could investigate if there is an effect of active gas seepage on the structure of the microbial assemblage, and search for mechanisms to explain that influence.

By combining the sensitivity of next-generation sequencing with the nearly full length archaeal SSU rRNA gene sequences, we were also able to study the phylogeny of Archaea inhabiting the sediments of these lakes. Remarkably, we found a high degree of resemblance with the phylogeny of Archaea identified in other kinds of sediments, and most importantly in other places in the Arctic. Furthermore, expanding the short list of cold-adapted methanogens in culture was a driving force that sparked my interest on this line of work. The culturing effort allowed a glimpse into the cultivable portion of the microbial assemblage, and reinforced findings from the use of molecular methods alone. Even though the culture work was not included in this dissertation as a chapter on its own, it

constituted a major endeavor during my tenure as a Ph. D. student, and I expect to continue developing this work in the future in collaboration with my Ph.D. advisor.

In comparison to the snow samples collected from two different mountain ranges at lower latitude, the diversity of the microbial community in thermokarst lake sediments was very high. Even though the surface of these lakes is frozen for about 9 months of the year, this difference in diversity was expected, given the broad range of nutrients that are available for microorganisms in the lakes and the burst of activity that is expected to take place in the lakes during the summer months.

2.2. Microbial communities in snow samples

The snowpack is an ephemeral environment that not only experiences extreme diurnal temperature variations, but also undergoes changes within seasons with variability in air mass movement and storm systems resulting in different levels of snow deposition. The snow ecosystem is perhaps more heterogeneous from the biogeochemical point of view than lake sediments are, because of the geochemical gradients that are established in sediments, mostly by depth. Conversely, snowpacks accumulate over a short season and present a range of depths that are determined by factors such as topography, the amount of snow accumulation during the season, the direction of exposure to sunlight, and the slope of the terrain. This contributes to the formation of areas where the deepest layers of the snowpack may be anaerobic, and shallower zones can be exposed to more air circulation. Additionally, the main

source of nutrients in this ecosystem are likely dominated by allochthonous sources, and may be carried by wet and dry deposition, or originate from the surface of the soil where the snowpack resides (Larose et al., 2013; Williams et al., 2009). However, late in spring when the snowmelt begins, the snow becomes a more dynamic ecosystem where nutrients are transferred from the surface layers to increasingly high depths of the snowpack, and autotrophic microorganisms produce higher amounts of autochthonous organic matter.

Chapter 4 revealed incredible parallels in composition between snow samples that were collected on opposite ends of the state of Nevada, even though at the OTU level there was high dissimilarity and differences in structure. Likewise, we exposed a unique structure of the microbial community composed of OTUs with different levels of abundance but composed of similar taxa, some of which were also dominant in all three sites. Even more interesting perhaps, was the finding of unique OTUs that were highly abundant in individual samples, and that were also common to all sites (*e.g.,* OTUs related to *Polaromonas* and *Herminiimonas*). Explaining the driving forces behind this unique community structure and finding mechanisms that may lead to community assemblage in these ephemeral snow ecosystems, opens up the door to a number of questions about the interconnectivity of similar ecosystems around the world and the adaptations that allow specific groups of organisms to inhabit them.

3. Astrobiological implications of the study of life in cold environments on Earth

The two primary elements that microbial communities inhabiting arctic thermokarst lake sediments and snowpacks may have in common are the molecular adaptations that allow microbial life forms to thrive in the cold, and the global distribution patterns of certain phylogenetic groups. However, microorganisms inhabiting all kinds of environments are one engine, whose interaction with the environment shape the face of the planet and determine what the Earth looks like from space. In the search for life elsewhere, we have to think of what life looks like from afar, under conditions that are not necessarily amenable to our standards for life. Only the study of life in extreme environments in our planet will allow us to expand our definition of life, and to predict how life could exist under analog planetary conditions. The identification of signatures that could be used as indicators of life is also essential to constrain our search for life, and to identify the most likely candidates that could harbor life today. The more we learn about the common features that microbial life forms share in extreme environments, the better chances we have to identify life when we find it. If I could only chose one thing to take with me from this work, it would be that life is extraordinarily versatile; and that we cannot let our preconceived limits for what life is to determine what we look for in other worlds, because even cold environments that may seem lifeless harbor some of the most powerful life forms on the planet.

References

Bretz KA, Whalen SC (2014) Methane cycling dynamics in sediments of Alaskan Arctic Foothill lakes. *Inland Waters,* **4**, 65-78.

He R, Wooller MJ, Pohlman JW, Catranis C, Quensen J, Tiedje JM, Leigh MB (2012) Identification of functionally active aerobic methanotrophs in sediments from an arctic lake using stable isotope probing. *Environmental Microbiology,* **14**, 1403-1419.

Ipcc (2007) Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. (eds Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt K, Tignor M, Miller H). Cambridge University Press, Cambridge; New York.

Larose C, Dommergue A, Vogel T (2013) The dynamic arctic snow pack: An unexplored environment for microbial diversity and activity. *Biology,* **2**, 317-330.

Lofton DD, Whalen SC, Hershey AE (2014) Effect of temperature on methane dynamics and evaluation of methane oxidation kinetics in shallow arctic Alaskan lakes. *Hydrobiologia,* **721**, 209-222.

Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA (2014) *Brock biology of microorganisms,* Benjamin-Cummings Publishing Company, Boston.

Matheus Carnevali PB, Rohrssen M, Williams MR, Michaud AB, Adams H, Berisford D, Love GD, Priscu JC, Rassuchine O, Hand KP, Murray AE (2015) Methane sources in arctic thermokarst lake sediments on the North Slope of Alaska. *Geobiology,* **13**, 181-197.

Rivkina E, Gilichinsky D, Wagener S, Tiedje J, Mcgrath J (1998) Biogeochemical activity of anaerobic microorganisms from buried permafrost sediments. *Geomicrobiology Journal,* **15**, 187-193.

Walter KM, Smith LC, Chapin FSI (2007) Methane bubbling from northern lakes: present and future contributions to the global methane budget. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* **365**, 1657-1676.

Williams MW, Seibold C, Chowanski K (2009) Storage and release of solutes from a subalpine seasonal snowpack: soil and stream water response, Niwot Ridge, Colorado. *Biogeochemistry,* **95**, 77-94.

Appendix

Table A1. Summary of relevant enrichment cultures from arctic lake sediments

Lake	Depth (cm)	Medium	T (°C)	Sequence ID	No clones	Min. % similarity (Silva database)	Tentative order	Tentative family	Tentative Genus	Tentative species
Ikroavik	02--18	Acetate + Antibiotics	~ 20	A1	3	99	Methanosarcinales	Methanosarcinaceae	*Methanosarcina*	*lacustris*
				B1	7	99	Enterobacteriales	Enterobacteriaceae	*Yersinia*	--
Ikroavik	02--18	H_2/CO_2 + Antibiotics	~ 20	A2						
				B2						
Siqlukaq	68--96	H_2/CO_2 + Antibiotics	~ 20	A3	8	97	Methanobacteriales	Methanobacteriaceae	*Methanobacterium*	--
				B3	--	--	--	--	--	--
Siqlukaq	02--28	Methanol + Antibiotics	~ 10	A6	9	97	Methanomicrobiales		*Methanosphaerula*	--
				B6	1	93	Clostridiales	XII Incertae Sedis	--	--
Siqlukaq	02--28	H_2/CO_2	~ 10	A4						
				B4						
Siqlukaq	02--28	H_2/CO_2	~ 10	A7						
				B7						
Siqlukaq	02--28	H_2/CO_2	~ 10	A8	--	--	--	--	--	--
				B8	1	98	Clostridiales	Clostridiaceae	*Clostridium*	--
Siqlukaq	02--28	H_2/CO_2	~ 10	A9	--	--	--	--	--	--
				B9	7	96	Clostridiales	Eubacteriaceae	*Acetobacterium*	--
Sukok Seep	02--16	Acetate + Antibiotics	~ 10	A5						
				B5						
Sukok Seep	02--16	Acetate	~ 20	A10	10	99	Methanosarcinales	Methanosarcinaceae	*Methanosarcina*	*lacustris*
				B10	3	95	Clostridiales	Peptostreptococcaceae	*Proteocatella*	--
Sukok Seep	02--16	Acetate	~ 20	A11	10	99	Methanosarcinales	Methanosarcinaceae	*Methanosarcina*	*lacustris*
				B11	11	97	Lactobacillales	Carnobacteriaceae	*Trichococcus*	--