Geomicrobial characterisation of a 60 m long permafrost core, taken over a future CO_2 storage site at Svalbard

Siren Fromreide

Supervisors: Laila Johanne Reigstad and Pål Tore Mørkved

Master's Thesis in Geobiology
Centre for Geobiology
Department of Biology
Faculty of Mathematics and Natural Sciences

University of Bergen
September 11th 2014

Abstract

In connection with a planned CO_2 storage pilot project in the Arctic, a 60 m long permafrost core was drilled in Adventdalen, Svalbard, representing the poorly studied deep permafrost ecosystems. The on-shore drilling was performed through deltaic, marine and glacial sediments, ending at the bedrock at 60 m. Here, seven different depths in the 3-60 m interval of the permafrost core were subjected to culture-independent methods such as 16S rRNA amplicon 454 pyrosequencing and functional and ribosomal gene quantifications to characterise the microbial community composition and abundance. Additionally, geochemical analyses of extracted pore water have been performed, as well as measurements of carbon content, major elements and grain size distributions.

The enumeration of bacterial and archaeal 16S rRNA genes showed high copy numbers in top sample at 3 m, a decrease at 4.5 m, and a further decrease from 54 to 60 m. Estimated prokaryotic cell numbers ranged between 3 x 10^5 and 1 x 10^8 cells g^{-1} sediment. Detection and quantification of selected functional marker genes indicated that microbial sulphate reduction is more pronounced than methanogenic and methanotrophic processes.

A 16S rRNA amplicon pyrosequencing library made with universal prokaryotic primers, revealed a dominance of poorly characterised microbial groups, such as Candidate division TM7, OP9, Chloroflexi Subdivision 11, Deep Sea Archaeal Group (DSAG), Miscellaneous Crenarchaeotic Group and Thermoplasmata. The bacterial and archaeal communities at 3 m were different from the other depths, dominated by Candidate division TM7 and Halobacteria, respectively. Low salt concentrations, high organic carbon content, high cell numbers and different community structure at 3 m suggested influence of surface-related processes such as migration of carbon and meteoric water at this depth.

Microbial community composition and geochemistry suggested an anaerobic habitat throughout the core. From 16.5 m to 57 m, the uncharacterised DSAG made up 40-99 % of the archaeal community, increasing their relative abundance with depth. This dominance was confirmed by DSAG specific qPCR. The primers used for amplicon pyrosequencing caused bias in the amplicon library, most likely due to preferential primer-template annealing of the degenerate forward primer. Both the geochemistry and microbial community composition had a marine signature from 16.5 m to 57 m, indicating that the marine depositional environment is the main factor determining microbial community structure, rather than the permafrost environment itself.

In relation to potential underground CO_2 storage at Svalbard, the results from this study serve as baseline information for future microbial monitoring if CO_2 should be stored at the site.

Acknowledgements

The work presented in this thesis was carried out at Centre for Geobiology (GCB) at the University of Bergen from 2013-2014 and was mainly funded by the project Subsurface CO_2 storage, Critical Elements and Superior Strategy (SUCCESS), a part of the Centre for environment-friendly energy research of the Research Council of Norway. This project is a part of CGBs work in SUCCESS WP 3, activity 4: "Marine monitoring, method development and baseline study". The drilling of the permafrost core was funded by the University Centre in Svalbard (UNIS) and Longyearbyen CO_2 lab.

I am very grateful for the guidance by my supervisors Laila Johanne Reigstad and Pål Tore Mørkved, who have been kind, patient and understanding through the whole project period. Thanks for having confidence in me, and providing me with the knowledge I needed to complete this Master's thesis. I have learned a lot from your supervision and expertise, both in the lab and in the writing process.

I would like to thank the Centre for Geobiology for providing me with a great working environment, and the tools and equipment needed to perform this research project. It has been two fun and educational years! Thanks to this Master's thesis, my supervisors and financial support from the Centre for Geobiology, I had the privilege to have two poster presentations in 2014; first at the "European Geosciences Union" (EGU) in Vienna under the Permafrost Open Session, and second at the "Biosignatures across Space and Time" conference here at the University of Bergen. I also got the opportunity to give an oral presentation of my work at a joint partner meeting at Centre for Permafrost (CENPERM) at the University of Copenhagen. During the Master's period, I also had the opportunity to participate in the "Nordic field course in Geobiology" at Iceland, through funding from the Department of Biology. This has been unique experiences that I will take with me in the future.

Gratitude also goes to a number of people at the University of Bergen who have assisted in technical analysis during this investigation: Hildegunn Almelid for analysis of ion concentrations, Ole Tumyr for X-ray Fluorescent element analysis, Professor Ingunn Hindenes Thorseth for analysis of nutrient concentrations and Pål Tore Mørkved for grain size distribution analysis. Special thanks to Steffen Leth Jørgensen for doing Deep Sea Archaeal Group (DSAG) quantification and phylogeny, and for participating in discussions and data interpretation.

I would like to thank my fellow master students for support and advice during this period. It has been nice to have someone who understands the frustrations and challenges that have appeared during these two years. Thanks to Jan Vander Roost, Julie Nikolaisen and Terje Johansen for proofreading the thesis manuscript. Lastly, I would like to thank my friends, family and boyfriend for providing me with support and motivation, and for telling me to relax when I needed to hear it.

Table of contents:

1. Introduction

1.1 Permafrost as a microbial habitat

Permafrost covers about 26 % of terrestrial soil ecosystems, and is defined as permanently frozen soil or rock that remains at or below 0 °C for at least 2 consecutive years (Williams and Smith, 1989). Permafrost is found in higher latitudes and elevations, where mean annual soil temperature is below freezing. It can extend from less than 1 meter to over 1000 m into the subsurface (Williams and Smith, 1989). This extreme environment is characterized by stable low temperatures, low water and nutrient availability due to freezing, and continuous gamma radiation from soil minerals (Ponder et al., 2004).

The upper layer of the permafrost that experiences thaw each summer is referred to as the active layer (Williams and Smith, 1989). Active layer thickness is influenced by seasonal air temperature, snow cover, vegetation, summer precipitation and topography (Hinkel and Nelson, 2003). The depth of the active layer can range from only a few centimetres up to several meters (Steven et al., 2006). The boundary between the active layer and the permafrost is called the permafrost table (Gilichinsky, 2002). This is a physical and biogeochemical barrier that restricts water infiltration and solute penetration from the surface into the permafrost layer. Microorganisms tend to accumulate on the permafrost table after each freezing-thawing cycle (Gilichinsky, 2002).

About 93-98 % of the water in permafrost is present as ice (Gilichinsky et al., 2007). Increasing amounts of ice reduces microbial activity and abundance. In the absence of liquid water, the microbial cells could become damaged by the formation of ice-crystals. Additionally, pure ice greatly limits outflow of metabolic end products and inflow of nutrients to the cell (Steven et al., 2006). The remaining 2-7 % of water is present in an unfrozen state as nanometer-thin films surrounding soil and ice particles. The water can remain unfrozen because of particle-water adsorption forces stabilising the liquid phase of the water (Lock, 1990). Fine grained sediments typically have more unfrozen water than coarse grained sediments, because the smaller particles have a larger surface area that interacts with water molecules (Gilichinsky, 2002). Salts and nutrients will concentrate in the thin films as the water freezes, which will further depress the freezing point of water (Steven et al., 2006). The unfrozen water films of permafrost are considered to be the main ecological niche where the microorganisms might survive, because they protect the cells against freezing-thawing stresses, contain dissolved nutrients and metabolites, and enables mass transfer of metabolic end-products (Gilichinsky, 2002).

The ability to successfully cultivate viable cells from permafrost depends on the age of the permafrost. It is assumed that the age of the microbial cells corresponds to the age of the permanently frozen state of the soils or sediments (Gilichinsky et al., 2007). The oldest cells recovered from permafrost date back to approximately 3 million years in the Arctic, and 5 million years in the Antarctic (Gilichinsky et al., 2008).

1.2 Microbial activity and adaptations to cold temperatures

Significant numbers of viable bacteria have been isolated from permafrost. Isolated organisms include aerobic heterotrophs, anaerobic bacteria and archaea, nitrogen-fixing bacteria, sulphur-oxidising and sulphur-reducing bacteria (Steven et al., 2006). Isolated organisms are rarely psychrophilic, but predominantly psychrotrophic (Rivkina et al., 2004). Permafrost microorganisms are often halotolerant, and it has been suggested that there could be a link between psychrotolerance and halotolerance (Vishnivetskaya et al., 2000).

It was previously assumed that microorganisms in permafrost were not biologically active, but in a frozen resting stage (Vorobyova et al., 1997). Several studies have detected microbial activity in permafrost environmental samples at subzero temperatures by measuring incorporation of labelled substrates (Rivkina et al., 2004, Rivkina et al., 2000, Steven et al., 2008) or gas fluxes (Kato et al., 2005, Rivkina et al., 2002). Measurements of *in situ* CO_2 flux in Canadian tundra yielded evidence for low levels of active microbial respiration (Wilhelm et al., 2012). Net production of CO_2 was observed in late winter at -9 to -16 °C, which demonstrates the microbial communities' ability to respire at subzero temperatures (Wilhelm et al., 2012). Recently, Tuorto and co-workers showed that bacterial genome replication occur in Alaskan permafrost by measuring incorporation of ^{13}C-acetate into bacterial DNA, during a six months incubation experiment at temperatures from 0 to -20 °C (Tuorto et al., 2014). Phylogenetic analysis revealed that the active bacteria were members of the Acidobacteria, Actinobacteria, Chloroflexi, Gemmatimonadetes and Proteobacteria phyla. A surprising outcome from this research was that some members of the Acidobacteria, Actinobacteria and Proteobacteria only synthesised DNA at temperatures below -6 °C (Tuorto et al., 2014).

Limiting factors for microbial activity at subzero temperatures include low availability of liquid water, high solute concentrations within the liquid water, nutrient limitation and decrease in molecular motion due to the cold temperatures (Mykytczuk et al., 2013). Some microorganisms might remain viable by entering a dormant resting stage in response to the conditions in the permafrost. Dormant cells will over time accumulate damage in their DNA and other internal biomolecules, due to exposure to background radiation and spontaneous chemical reactions within the cells (Lindahl, 1993). Some cells in permafrost maintain an active metabolism and cellular repair by having special adaptations for their cell components to remain functional at subzero temperatures (Bakermans et al., 2009). Examples are given below.

Low temperatures reduce the fluidity of cell membranes, shifting the membrane from a liquid-crystalline phase to a gel phase (Denich et al., 2003). Cold-adapted microorganisms modify the lipid composition of their cell membranes by the synthesis of more short chained, unsaturated and branched fatty acids to maintain membrane fluidity at low temperatures (Denich et al., 2003). These membrane changes allow transport across the

membrane to continue, by maintaining interactions with membrane proteins (Ponder et al., 2004).

Enzymes become less active at low temperatures, due to decreased flexibility of protein structure (Bakermans et al., 2009). Cold-adapted microorganisms reduce weak stabilising interactions in their enzymes, such as ion pairs, hydrogen bonds and hydrophobic interactions, increase solvent interactions with non-polar or interior amino acid residues and reduce proline and arginine content (Bakermans et al., 2009). These adaptations destabilise the protein, allowing the active site to be flexible at temperatures that decrease molecular motion (Jansson and Tas, 2014).

Low temperatures stabilise the secondary structures of nucleic acids, causing inhibition of DNA replication. Cold-adapted microorganisms minimise the formation of nucleic acids secondary structures through RNA chaperones (Bakermans et al., 2009). RNA chaperones prevent the formation of RNA secondary structures, and thereby enable correct transcription and translation of DNA (Jansson and Tas, 2014). Antifreeze proteins facilitate microbial survival at low temperatures, because they reduce the freezing point of water within the cell by modifying ice crystal structure and inhibit re-crystallisation of ice (Gilbert et al., 2004).

Ice formation in permafrost increases solute concentration in the unfrozen water films (Steven et al., 2006). This creates an environment with low water activity and high salt concentrations. One mechanism to allow microbial survival at high salinity and subzero temperatures is the production of compatible solutes. Compatible solutes are organic osmolytes such as sugars, polyols, free amino acids and betaines that microorganisms accumulate in their cytoplasm to prevent cell dehydration (Kempf and Bremer, 1998). These compounds can reach high intracellular concentrations without disturbing vital cellular functions. The production or uptake of compatible solutes excludes environmental solutes from the cell, while the compatible solutes accumulate and thereby control the cells osmotic balance (Kempf and Bremer, 1998).

1.3 Microbial diversity in permafrost

Origin, age and physiochemical characteristics of permafrost are highly variable, resulting in variations in microbial abundance and diversity between different locations. Early studies investigating microbial diversity in permafrost relied mostly on culture-dependent methods (Shi et al., 1997, Vishnivetskaya et al., 2000, Vorobyova et al., 1997), and knowledge of the overall composition of permafrost microbial communities was therefore limited. Since it is well known that only a small part of the total microbial communities can be cultured, there has been increased focus on using culture-independent methods to characterise the microbial communities in permafrost in recent years. Despite the extreme conditions within permafrost, molecular-based estimates of microbial diversity in permafrost are high (Jansson and Tas, 2014). It has been suggested that the microbial diversity in permafrost was comparable to the diversity of the active layer (Gilichinsky et al., 2008). A metagenomic

study of the active layer and 2 m permafrost of the Canadian High Arctic reported similar functions and phylogenetic community compositions in both the active layer and permafrost samples, but the permafrost microbial community had reduced diversity compared to the active layer (Yergeau et al., 2010). Reduced permafrost microbial diversity compared to active layer diversity was also observed in an Alaskan boreal forest (Tas et al., 2014). It is suggested that harsh conditions in permafrost are selective for cells that can survive subzero temperatures, low water and nutrient availability for prolonged periods of time (Yergeau et al., 2010).

Microbial abundance estimated by direct microscopic counts vary between 10^5 and 10^8 cells per gram dry weight in different permafrost areas (Gilichinsky et al., 2008). It is assumed that the diversity of Bacteria in permafrost is higher than the archaeal diversity (Jansson and Tas, 2014). Bacterial phyla that are frequently detected in permafrost using molecular-based techniques include Actinobacteria, Proteobacteria, Firmicutes, Chloroflexi, Acidobacteria and several uncharacterised phyla (Gilichinsky et al., 2007, Hansen et al., 2007, Steven et al., 2007, Steven et al., 2008, Tas et al., 2014, Vishnivetskaya et al., 2006, Wilhelm et al., 2011, Yergeau et al., 2010). To our knowledge, there are few studies describing the archaeal diversity in permafrost using molecular methods. Archaeal sequences detected in permafrost include those from the Crenarchaeota (Steven et al., 2007, Steven et al., 2008, Wilhelm et al., 2011, Yergeau et al., 2010), the Euryarchaeota (Steven et al., 2007, Steven et al., 2008, Tas et al., 2014) including sequences belonging to methanogenic archaea (Gittel et al., 2014, Mackelprang et al., 2011, Yergeau et al., 2010) and the Thaumarchaeota (Gittel et al., 2014) It should be noted that detection of DNA sequences in permafrost do not indicate that the microorganisms are metabolically active or even viable, as the subzero temperatures in permafrost are well suited for long-term preservation of dead cells or naked DNA (Willerslev et al., 2004).

Most research on microbial diversity in permafrost has been performed using clone libraries (Gilichinsky et al., 2007, Hansen et al., 2007, Steven et al., 2007, Steven et al., 2008, Vishnivetskaya et al., 2006, Wilhelm et al., 2011). Clone libraries provide a more limited sampling of the microbial communities, compared to next-generation sequencing approaches that can generate millions of sequences in one run (Mardis, 2008). In recent years there have been a few studies that have applied next-generation sequencing methods when investigating both bacterial and archaeal communities in the active layer and permafrost at different locations. These studies have analysed the microbial functional potential of the active layer and 2 m permafrost in the Canadian High Arctic (Yergeau et al., 2010), the microbial community's response to thaw (Mackelprang et al., 2011), the *active* microbial community involved in organic carbon transformations in the active layer at Svalbard (Tveit et al., 2013), the impact of fire on microbial community compositions in the active layer and 1 m permafrost of Alaska (Tas et al., 2014) and the microbial community structure in Siberian tundra (Gittel et al., 2014). To our knowledge, this present study represents the first comprehensive characterisation of microbial communities in a vertical

depth profile of permafrost samples recovered from more than 2 m depth, using next-generation in-depth sequencing in combination with geological and geochemical characterisation of the habitats.

Research on microbial ecology in permafrost may reveal the lower temperature limits of life, and the mechanisms that enable the microorganisms to survive in this extreme environment over thousands to millions of years. Furthermore, the microorganisms present in permafrost on Earth provide terrestrial analogues for potential extraterrestrial microbial life that may exist in permafrost on Mars or in other frozen parts of the Universe. Lastly, it is important to have knowledge about the distribution and functions of microorganisms in permafrost in order to understand how they will respond to global warming and permafrost thaw in the future.

1.4 Introduction to CO_2 capture and storage

The rise in average global temperatures observed during the last century is likely to be caused by the release of anthropogenic greenhouse gases (IPCC, 2013). Therefore, there is large focus on developing technologies that can reduce CO_2 emissions into the atmosphere. The capture of CO_2 from large emission sources and the subsequent storage of CO_2 in geological structures (CCS) is a promising technology for reducing CO_2 emissions in large quantities (IPCC, 2005).

Over 40 % of anthropogenic CO_2 emissions come from coal fuelled power plants (Quadrelli and Peterson, 2007). The use of coal for electricity production will increase in the next 20-30 years, as countries like China and India become more industrialised (Quadrelli and Peterson, 2007). For this reason, the capture of CO_2 emissions from coal fuelled power plants will become increasingly important in the future. Norway has only one coal fuelled power plant which is located in Longyearbyen, Svalbard. The University Centre in Svalbard (UNIS) and the Longyearbyen CO_2 lab have taken the initiative to use the favourable subsurface geology of Svalbard to develop a pilot project, which will test and demonstrate new technology for CCS from the coal fuelled power plant in Longyearbyen. The vision is to make Longyearbyen into a community with almost no man-made CO_2 emissions (Braathen et al., 2012).

Before CO_2 storage can start at Svalbard, detailed knowledge about the subsurface geology need to be obtained. In Adventdalen, about 5 km outside of Longyearbyen, several wells have been drilled in order to investigate the subsurface conditions for CO_2 storage. A suitable CO_2 storage reservoir has been identified at 670-970 m depth, overlain by 400 m of cap-rock (Braathen et al., 2012). At 150-200 m there is a sandstone section with a shale overburden that contains over-pressurised methane of biogenic origin (Pål Tore Mørkved, personal communication). The permafrost goes down to approximately 120 m depth at the site, with sediments going down to 60 m, and the remaining 60 m being shale. It is assumed that the permafrost can act as a secondary seal to keep the CO_2 below ground (Braathen et al., 2012). Several environmental baseline studies are currently ongoing, to assess the initial

conditions at the site that will act as reference points for future monitoring if CO_2 should be stored below ground at Svalbard in the future.

1.5 Impacts of CO_2 leakage on microbial communities

Monitoring of CCS project is essential for public acceptance and for early detection of potential leaks from the CO_2 storage reservoir. Geophysical and geochemical methods are widely used for monitoring, but the impact of CO_2 storage and eventual leakage on microbial abundance and community structure is poorly understood. CO_2 leakage can induce changes in pH, salinity and other geochemical factors that will influence the microbial populations (Noble et al., 2012). CO_2 leakage will cause anaerobic conditions, because the CO_2 replaces the oxygen. Anaerobic conditions are likely to favour methanogenic archaea and sulphate-reducing bacteria (SRB). These microbial groups are probably the most important groups to consider in microbial monitoring of CCS projects (Noble et al., 2012).

Microbial monitoring of a deep, terrestrial subsurface CO_2 storage reservoir has been performed in Ketzin, Germany, using Fluorescent In Situ Hybridization (FISH), Single-Strand-Conformation Polymorphism (PCR-SSCP) and Denaturing Gradient Gel Electrophoresis (DGGE) on recovered aquifer samples (Morozova et al., 2010, Morozova et al., 2011). This research showed that microbial numbers decreased after CO_2 injection, but after 5 months of CO_2 storage the population recovered and an enhanced activity (measured by FISH) of the total microbial population was observed (Morozova et al., 2010). However, the composition of the microbial community was altered as response to the CO_2 exposure. After CO_2 injection there was an increase in the abundance of methanogenic archaea, because the consequential decrease in pH gave more favourable conditions for these organisms. Further monitoring revealed that SRB were the dominating microbial group over time (Morozova et al., 2011).

Quantitative PCR analysis of bacterial and archaeal abundance at a terrestrial CO_2 vent in Germany, demonstrated significant reduced bacterial abundance near the CO_2 leak, compared to background, while the archaeal abundance significantly increased (Krüger et al., 2009). Microbial analyses of a terrestrial, natural CO_2 vent in Latera, Italy revealed more activity of methanogenic archaea and SRB at the CO_2 seep, while the activity of methane oxidising bacteria had a negative response to high CO_2 concentrations (Beaubien et al., 2008, Oppermann et al., 2010).

1.6 The 60 m permafrost core from Svalbard

A 60 m long permafrost core, named well DH8, was drilled at Svalbard, in connection with investigations by the Longyearbyen CO_2 lab in the Adventdalen well park (**Fig. 1**). The core represents the upper 60 m of the permafrost, from approximately 1.5 m below surface to the bedrock at 60 m.

The permafrost core is undergoing extensive multidisciplinary studies due to collaboration between Center for Permafrost (CENPERM) (Copenhagen, Denmark), Aarhus University (Denmark), Nordic Centre for Luminesence Research (Risø, Denmark), The University Centre in Svalbard (UNIS), University of Oslo and the Centre for Geobiology, University of Bergen. Research conducted on the core includes sedimentological, stratigraphical and cryostratigraphical work, measurements of CO_2 production from the core, isotopic analysis of methane gas, Optically Stimulated Luminescence (OSL) dating of core sediments, Foraminifera dating of core sediments, microbial activity, turnover times and analysis of endospores. The present study has investigated microbial abundance, functional potential, community structure and geological and geochemical parameters in seven selected depths of the 60 m permafrost core.

1.7 Objectives

The main objectives of this study were

1. To characterise the microbial abundance and community composition along a depth profile in a 60 m long permafrost core using culture-independent methods.

2. To examine potential factors that can influence the microbial abundance and community composition. Such factors can be geochemistry, depositional environment, nutrient availability, age, and potential contamination from drilling fluid.

3. To assess the present microbial abundance, functional gene abundance and community composition as a baseline study for future monitoring if CO_2 should be stored underground at Svalbard.

4. To search for microbial methane metabolisms through marker genes for methanotrophy and methanogenesis. High concentration of biogenic, over-pressured methane has previously been found in permeable sandstones at 150-200 m depth. The overpressure could cause potential methane seeps through the upper 150 m. The quantification of selected microbial methane metabolisms could be used to detect methane seep from below, and thus assess the sealing properties of the shale and permafrost.

2. Materials and methods

2.1 Site description and core retrieval

On September 3rd 2012, a 60 meter long permafrost core, named Well DH8, was drilled in Adventdalen, Svalbard (78°12'N, 15°49'E, **Fig. 1**), as part of the Longyearbyen CO_2 lab project.

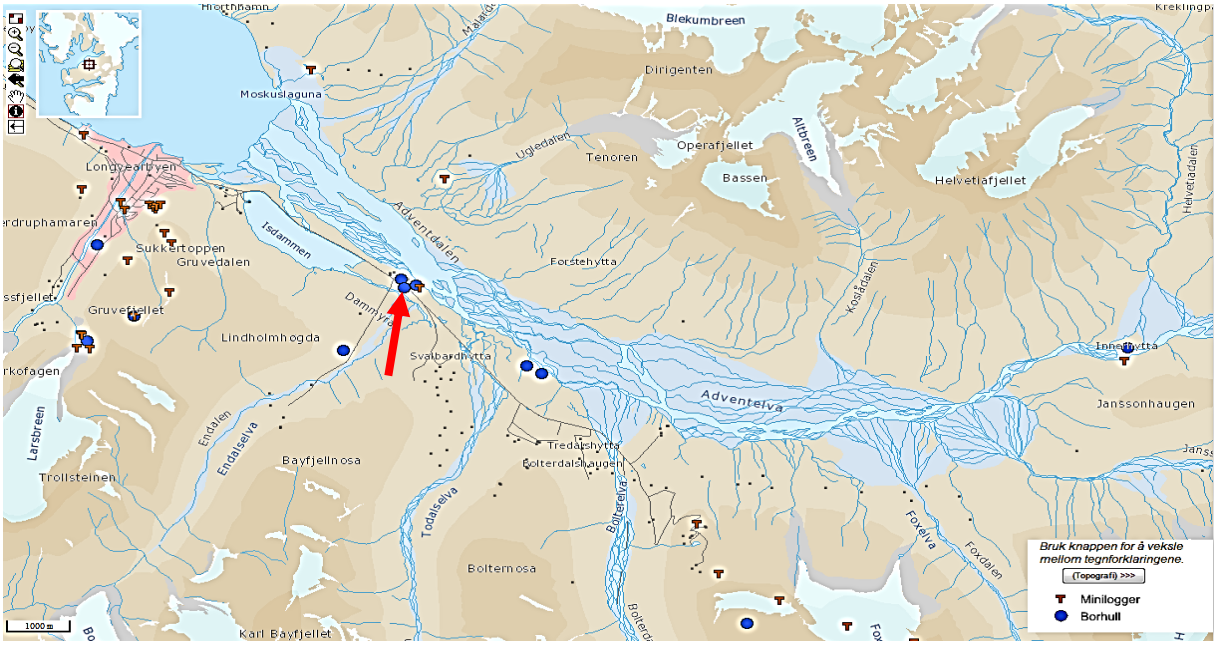

Figure 1: Location of drill site in Adventdalen, Svalbard. The red arrow shows the location of well DH8 (Permafrost Observatory Project).

Drilling was conducted using an industrial scaled drill rig. Core sections were retrieved in 1.5 m intervals of 6 cm diameter. Local river water from Adventelva was used as drilling fluid. During drilling, the core sections were embedded in a plastic tube, in order to maintain the integrity of the core during the drilling and the recovering process, and thereby reduce contamination from the drilling fluid. The core was processed immediately at the drill site after it was brought to the surface.

Field sampling was performed by Pål Tore Mørkved and Graham L. Gilbert. Approximately 5 cm pieces (including plastic tube) were cut off from the 1.5 m core sections, using a saw. The tools used in the field were not sterilised. The cut off pieces were sealed in plastic, and kept frozen by immediate storage at -10 °C and later stored at -20 °C. Only pieces that showed no sign of thawing were selected. The rest of the core was used for other purposes, which limited the number of samples available for this geomicrobial study. Samples from seven depths of the core were chosen for this project (**Table 1**). These samples were chosen to cover the main depth intervals and deposition types, and on basis of the integrity and preservation of the available samples (e.g. no visible cracks) to avoid contamination from the drilling fluid.

The thickness of the active layer in Adventdalen is approximately 1 m (**Fig. 2**).

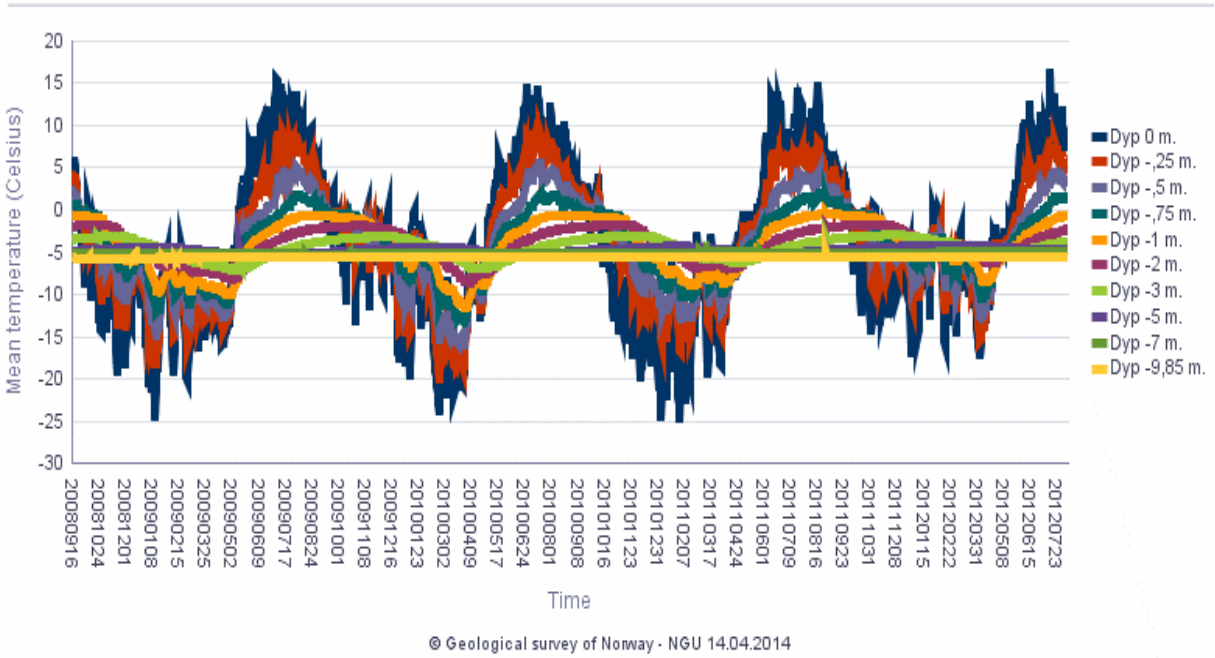

Figure 2: Active layer depth in Adventdalen from a bore hole temperature logger during the periode September 2008 to July 2012, showing that depths below 1 m are constantly below 0 °C (Permafrost Observatory Project, 2012).

Sedimentary stratigraphy analysis (Gilbert, 2014) revealed mainly a marine origin of the DH8 permafrost core, with aeolian deposits in the top few metres, extending through delta top, delta front and prodelta deposits. Below this were marine sediments and till, ending in bedrock (shale) at 60 m (**Fig. 3**). Optically Stimulated Luminescence (OSL) dating of the core sediments (Christine Thiel, In: Gilbert, 2014) showed that the core was of Holocene age. The shale at 60 m is part of the Carolinefjellet formation and was deposited during the early Cretaceous (Parker, 1967). Permafrost aggradation occurred during the late Holocene (approximately 3000 years ago) (Gilbert, 2014).

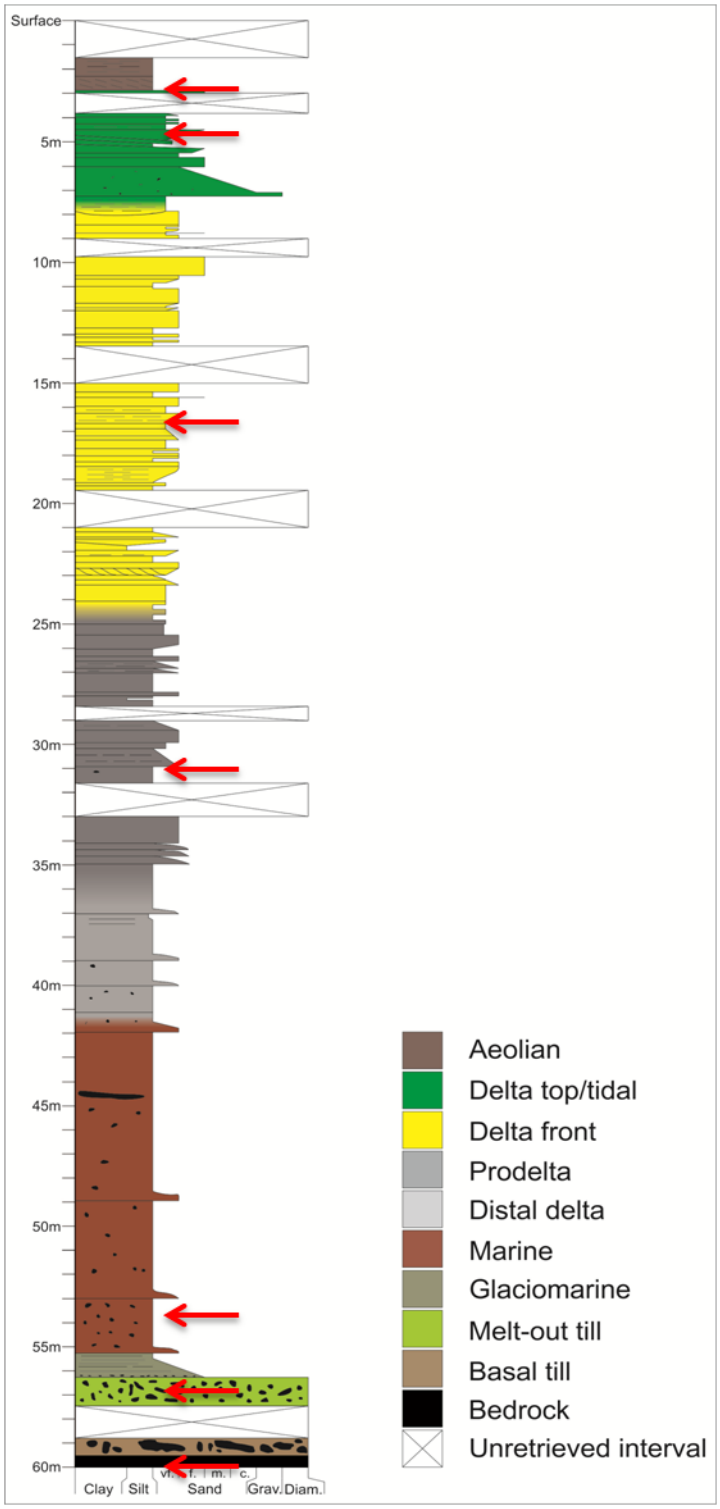

Figure 3: Geomorphological description of the 60 m permafrost core. Red arrows highlight the seven selected sampling depths for the geomicrobial studies of this Master's thesis (modified from Gilbert, 2014).

2.2 Sub-sampling of core

Upon analysis, the outer 1-2 cm of the core was removed in order to reach the centre of the core that had not been in contact with the drilling fluid. This was done in a freezing room holding a temperature of -20 °C. The bench was covered with aluminium foil, and sterilized with 70 % ethanol. Gloves were used while working with the samples to avoid contamination. A sterilised (70 % ethanol) saw, hammer and chisel were used to cut off the outer edges of the samples (**Fig. 4**). Some of the potentially drill fluid contaminated cut offs from all seven depths were kept, in order to compare the microbial communities in the outer edge samples with the inner core samples. Two of the outer edge samples were used in further analyses (**Table 1**).

Figure 4: Sterile sub-sampling of the frozen permafrost core using a sterilised chisel on an aluminium foil covered bench in the freezing room (-20 °C).

Table 1: Overview of samples and geomicrobial analyses used in this thesis.

Depth (m)	Grain size	TOC [a]	XRF [b]	Geoche-mistry	qPCR [c] 16S rRNA Bacteria	qPCR 16S rRNA Archaea	qPCR 16S rRNA DSAG [d]	qPCR functional genes	Amplicon 454 Pyro-sequencing	Cloning
3	x	x	x	x	x	x	x	x	x	
4.5	x	x	x	x	x	x	x	x	x	
16.5	x	x	x	x	x	x	x	x	x	
31.5	x	x	x	x	x	x	x	x	x	
54	x	x	x	x	x	x	x	x	x	x
57	x	x	x	x	x	x	x	x	x	
60				x	x	x	x	x	x	
3 *									x	
16.5 *									x	

[a] Total Organic Carbon

[b] X-ray Fluorescent

[c] quantitative PCR

[d] Deep Sea Archaeal Group

* Sample from the outer edge of the core

2.3 Geology and geochemistry

2.3.1 Grain size distribution

A grain size analysis was performed on the six sedimentary samples from 3 m to 57 m by Pål Tore Mørkved, using Malvern Mastersizer 2000 at Institute for Energy Technology (IFE). This instrument uses laser diffraction and the principle that the angle of the scattered light from the laser beam is directly related to the particle size. The grain sizes were classified according to the International Organization for Standardization (ISO) as clay (< 2 µm), silt (2-63 µm) or sand (63-2000 µm) (ISO 14688-1:2002).

2.3.2 Organic carbon content

Total carbon (TC) and total inorganic carbon (TIC) were measured with a Multi EA4000 Elementary analyser (AnalytikJena), in the six samples between 3 m and 57 m. From the difference in TC and TIC, the amount of total organic carbon (TOC) was calculated. Before the carbon analysis the samples were dried over night at 80 °C. After drying, the samples were ground into powder using a mortar. The samples were stored in an exicator until further analysis, to avoid taking up moisture from the air.

Two ceramic boats (sample containers) were prepared for each sample with 35-38 mg of the dried sediment in each boat; one for TIC analysis and one for TC analysis. Two standards with known amounts of carbon were used. These standards were calcium carbonate (= Standard 1), which contains 12 % carbon, and a carbonate standard (= Standard 2) with 50.16 g quartz and 47.86 g $CaCO_3$. Standard 1 was used to correct the measurement results on a daily basis. No organic carbon standard was used. Standard 1 was always run in the beginning of the day-long runs, and was further repeated for about every sixth sample, to ensure there was no drift during analysis. Standard 2 was run at the end of the analysis. The samples from 3 m and 4.5 m were run in triplicates, otherwise samples were run only once.

After weighing the samples and standards into the ceramic boats, the weight and name of the samples were written in the sample list of the MultiWin software, and then the boats were loaded to the machine. First, TIC was measured by adding acid (H_3PO_4, 30-40 %) to the sample. The adding of acid was done by an automatic dispenser mounted on the Elementary Analyser. The addition of acid dissolved the carbonates in the sample, and the generated CO_2 was measured by the instrument.

The second boat with dried sediment was for TC measurement, and here the boat was transferred completely to the combustion oven of the analyser where it was dissolved and oxidized in oxygen flow at 1100 °C. The CO_2 content in this gas was analysed by the instrument.

The TIC, TC and TOC results were corrected by a “day factor”, calculated daily from the Standard 1 measurements. All measurement results were multiplied by this factor. The day factor is calculated with the equation:

$$F = \frac{C_{nominal}}{C_{actual}}$$

,where $C_{nominal}$ is the carbon content of Standard 1 which consists of 120 g/kg, and C_{actual} is the carbon content of Standard 1 that was measured. The carbon content from the analysis (g/kg) was calculated to weight percent carbon (wt %).

2.3.3 XRF - major elements analysis

X-Ray Fluorescence (XRF) was used to identify and quantify the major elements in the six samples between 3 m and 57 m. The principle for this method is that when an atom gets irradiated with high-energy X-rays, electrons from an inner orbital of the atom get ejected, making the atom unstable. To compensate for this, electrons from an outer orbital fall down to the lower orbital to replace the electron that was ejected, and this releases energy as fluorescent X-rays. Each element has a characteristic fluorescent X-ray signature due to the differences in energy between the outer and inner orbital. By measuring the intensities of the emitted energy it is also possible to quantify each element in the sample.

First, the samples were dried over night at 70 °C. Then the samples were ground into powder using a silica based mortar. Approximately 2 gram of dried sample was weighed into a crucible, and the samples were heated to 1000 °C for 1 hour to burn off all organic material and carbonates in the samples. The crucibles were weighed before and after heating, and the equation

$$Loss\ on\ ignition = \frac{(A - B)}{C} \times 100\ \%$$

was used to determine how much material that was lost during heating, where A is the weight of the crucible with sample before heating, B is the weight of the crucible with sample after heating, and C is the weight of the sample after drying.

Next, the samples were melted into glass disks that were used in the XRF analysis. This was done by mixing 0.96 g of sample with 6.72 g Spectromelt A-10, which binds the sample material and lowers the melting temperature of the minerals in the sample. This was heated to about 800 °C, to melt the mixture and make a homogenous glass disk. The major elements were analyzed on a Philips PW1404 X-ray fluorescence spectrometer by Ole Tumyr.

2.3.4 Geochemical analyses of extracted ions

The pore water ions were extracted by sequential extraction using suprapure H_2O and 0.02 M $CaCl_2$. First, the water soluble ions were extracted with H_2O. After this, 0.02 M $CaCl_2$ was added to the sediment, to potentially extract a higher amount of cations that are associated with mineral particles. Two samples from each depth were analysed. The shale sample (60 m) was included in the analysis, by pulverising the rock in a mortar before undergoing the same treatment as the sedimentary samples. First, 18 ml suprapure H_2O was added to 2 g sample in a 50 ml Falcon tube (Houba et al., 2000). The sediment-water solution was shaken for 2 hours in room temperature. After this, the solution was centrifuged for 10 minutes at 5000 rpm at 4 °C to remove the sediment from the solution. The supernatant was filtrated through 0.2 µm syringe filter. Next, 18 ml 0.02 M $CaCl_2$ was added to the same sediment that had been treated with H_2O. This solution was also shaken for 2 hours, centrifuged and filtrated through 0.2 µm syringe filter, as performed on the water extracted sample above.

2.3.4.1 Cations

The major cations were analysed by Hildegunn Almelid at an ICP optical emission spectrometer (Thermo Elemental Iris) on both the H_2O extracted ions and $CaCl_2$ extracted ions. The cation concentrations (given in ppm from the instrument) were calculated to ppm/ml pore water using the following equation:

$$ppm/ml\ pore\ water = \frac{Concentration\ of\ cation\ (ppm)}{ml\ pore\ water} \times Total\ volume\ (pore\ water + 18\ ml)$$

This was converted to mM by dividing the result from the equation above with the molar mass of the element of interest. The pore water volume was found by subtracting the dry weight sediment (after drying over night at 80 °C) from the wet weight sediment. The cation concentrations from the H_2O extraction and $CaCl_2$ extraction were summed.

2.3.4.2 Anions

Cl^-, Br^- and SO_4^{2-} concentrations were analyzed with ion chromatography (Metrohn) by Hildegunn Almelid. The results from the instrument were calculated to mM as described above.

2.3.4.3 Nutrients

The concentrations of the nutrients NH_4^+, NO_3^- and PO_4^{2-} were measured using the QUAATRO continuous flow analyser (Seal Analytical) by Professor Ingunn Hindenes Thorseth. The results from the instrument were calculated to µM as described above.

2.3.4.4 pH and alkalinity

pH and total alkalinity was measured immediately after pore water extraction on a 826 pH mobile (Metrohn) and a 888 Titrando (Metrohn).

2.4 Molecular methods

2.4.1 DNA extraction

DNA from the seven samples was extracted using the FastDNA SPIN Kit for soil (MP Biomedicals, Solon, OH), following the manufacturers' protocol. Two to four parallel DNA extractions were performed at the top six depths (3 – 57 m), while one extraction was performed on the 60 m shale sample and in the two outer edge samples (**Table 1**). Approximately 0.5 g of sediment was weighed into a Lysing Matrix E tube. 978 µl Sodium Phosphate buffer and 122 µl MT Buffer was added and homogenized in the FastPrep instrument (MP Biomedicals) for 40 seconds at a speed setting of 6.0, in order to lyse the cells and free the DNA to the buffer solution. This was centrifuged at 14 000 x g for 5 minutes to pellet sediment waste, cell debris and lysing matrix. The supernatant was transferred to a clean 2.0 ml microcentrifuge tube. 250 µl Protein Precipitation Solution was added to precipitate the proteins in the DNA containing solution. The solution was mixed by shaking the tube by hand 10 times, and centrifuged for 5 minutes at 14 000 x g. Next, the DNA containing supernatant was transferred to a 15 ml Falcon tube and 1 ml binding matrix was added to the supernatant, in order to bind the DNA and remove it from the solution. The binding was carried out using a horizontal rotator for 2 minutes.

The binding matrix (with DNA) was allowed to settle for 3 minutes, before 500 µl of the uppermost supernatant without the binding matrix was removed. The binding matrix was resuspended in the remaining supernatant. 600 µl of the mixture was transferred to a SPIN Filter, and centrifuged at 14 000 x g for 1 minute to get the DNA from the solution down to the filter. The catch tube was emptied and this procedure was repeated until all binding matrix solution had passed through the SPIN filter.

To wash the DNA and binding matrix, 500 µl of SEWS-M (with ethanol) was added to the SPIN Filter, and the pellet was resuspended using the force of the liquid from the pipet tip. This was centrifuged for 1 minute at 14 000 x g, the catch tube was emptied, and centrifuged again for 2 minutes to make sure that the matrix was dried for wash solution. The catch tube was discarded and replaced with a new clean catch tube. The SPIN Filter was air dried for 5 minutes at room temperature, to evaporate any remaining ethanol from the wash solution. The binding matrix was resuspended in 60 µl distilled H_2O, and centrifuged for 1 minute at 14 000 x g to elute the DNA from the SPIN filter to the catch tube. The SPIN Filter was discarded, and the DNA was stored in the freezer at -20 °C until PCR and qPCR analysis.

The shale sample (60 m) was first crushed in a metal mortar before DNA extraction. The mortar was sterilised in 70 % ethanol and burned off, before adding the sample to the mortar. DNA from the shale sample was extracted using the same procedure as described above, but in the first step 5 µl of PolyA was added to the Lysing Matrix E tube to bind to the minerals that can inhibit the DNA extraction, and 973 µl Sodium Phosphate buffer was added. Subsequent steps are as described for the sediment samples above.

2.4.2 Quantification of 16S rRNA genes using qPCR

Quantitative Real-Time polymerase chain reaction (qPCR) was used to detect and quantify bacterial, archaeal and Deep Sea Archaeal Group (DSAG) 16S rRNA genes, using StepOne Real Time PCR system (Applied Biosystems). For bacterial and archaeal 16S rRNA quantification, two to four of the parallel DNA extractions from the six depths between 3 m and 57 m were analysed, and the single sample from 60 m was analysed. For DSAG 16S rRNA quantification one sample from each depth was run in the qPCR analysis. All samples and standards were run in triplicates, and positive and negative controls were included in all qPCR runs. To confirm product specificity, a melting curve for the obtained DNA products was performed after each run for all experiments. The following equation was used to calculate gene copies per gram wet sediment from the qPCR results:

$$Gene\ copies/g\ wet\ sediment = \frac{gene\ copies\ \times dilution\ factor\ \times DNA\ dilution}{sediment\ weight\ (g)}$$

2.4.2.1 Quantification of bacterial 16S rRNA

For quantification of bacterial 16S rRNA genes the primer bac341F (5´-CCTACGGGWGGCWGCA-3´ , where W=weak (A or T)) (Jorgensen et al., 2012) and 518R (5´-ATTACCGCGGCTGCTGG-3´) (Muyzer et al., 1993) were used. The DNA samples were diluted 1:10 in distilled water. The standard curve was made from *E.coli* DNA, and ranged from 1.31 x 10^7 to 1.31 x 10^2 copies of the 16S rRNA gene per µl. Each reaction (20 µl) contained 10 µl SYBR Green 2 x master mixture (Qiagen), 0.15 µl of each primer (100 µM stock), 8.7 µl H_2O and 1 µl diluted template. The qPCR program was initiated with a hot start denaturation step for 15 minutes at 95 °C to activate the polymerase and to completely denature the DNA double strand, followed by 35 cycles of denaturation for 15 seconds at 95 °C, primer annealing for 30 seconds at 58 °C and an elongation step for 30 seconds at 72 °C, where the polymerase add nucleotides to the single-stranded DNA. The melting curve created after the 35 cycles, started with 15 seconds denaturation at 95 °C, followed by 60 °C for 1 minute. This was increased to 95 °C (plate read after every 0.5 °C between 60 °C and 95 °C), and finished with 95 °C for 15 seconds. R^2 for the standard curve was 0.996 and the amplification efficiency (Eff%) was 117 %.

The bacterial 16S rRNA gene copy numbers were calculated to approximate cell numbers by dividing the gene copy numbers with 4.2, which currently is the average number of 16S rRNA genes per genome for Bacteria (Vetrovsky and Baldrian, 2013).

2.4.2.2 Quantification of archaeal 16S rRNA

For archaeal 16S rRNA gene amplification the primers Un515F (5´-CAGCMGCCGCGGTAA-3´, where M= amino (A or T)) (Lane *et al*., 1991) and Arc908R (5´-CCCGCCAATTCCTTTAAGTT-3´) (Jorgensen et al., 2012) were used. The DNA samples were diluted 1:10 in distilled water. The standard curve was made from the *Thaumarchaeota* fosmid 54d9 (Treusch et al., 2005), and had copy numbers of archaeal 16S rRNA from 10^7 to 10^2 copies/µl. Each reaction

contained the same components and volumes as described for bacterial 16S rRNA quantification. The qPCR program was initiated with 15 minutes at 95 °C, followed by 40 cycles of denaturation for 30 seconds at 95 °C, annealing for 30 seconds at 60 °C and elongation for 45 seconds at 72 °C. The melting curve had the same thermal conditions as for bacterial 16S rRNA quantification. R^2 for the standard curve was 0.995 and Eff% was 112 %.

2.4.2.3 Quantification of Deep Sea Archaeal Group (DSAG) 16S rRNA

The microbial community composition obtained by 16S rRNA amplicon 454 pyrosequencing showed that the Deep Sea Archaeal Group (DSAG) was the most abundant archaeal group from 16.5m to 57 m (see "Results", Section 3.6.2). To verify these results, the 16S rRNA gene of the DSAG was quantified by Steffen Leth Jørgensen using the primers DSAG535f (5'-ACCAGCTCTTCAAGTGG-3') (Jorgensen et al., 2013) and Arc908r (5'-CCCGCCAATTCCTTTAAGTT-3') (Jorgensen et al., 2012). The samples were diluted 1:10 in distilled water. A purified PCR amplicon obtained with M13 primers from a DSAG positive clone was used as standard. Each reaction contained the same components and volumes as described for bacterial 16S rRNA quantification. The qPCR program started with 15 min at 95 °C, then 35 cycles of denaturation at 95 °C for 30 seconds, annealing at 59 °C for 30 seconds and elongation at 72 °C for 45 seconds. The melting curve had the same thermal conditions as for bacterial 16S rRNA quantification. R^2 for the standard curve was 0.992 and Eff% was 102 %.

2.4.3 Quantification of the marker genes *dsrB*, *mcrA* and *pmoA*

To look into what functions the microbial cells have in the environment, one can identify and enumerate marker genes for selected metabolisms. In this project, the functional marker genes *dsrB* (dissimilatory sulfite reductase β-subunit), *mcrA* (methyl coenzyme M-reductase) and *pmoA* (particulate methane monooxygenase) were quantified using StepOne Real Time PCR system (Applied Biosystems). The *mcrA* and *pmoA* genes were first run in a conventional PCR to see if the genes were present in the samples. Two of the parallel DNA extractions from the six depths between 3 m and 57 m were analysed, and the single sample from 60 m was analysed.

2.4.3.1 Quantification of the dsrB gene

The *dsrB* gene is a marker gene for microbial sulphate reduction and encodes the enzyme dissimilatory sulphite reductase. This gene was quantified using the primers DSRp2060F (5'-CAACATCGTYCAYACCCAGGG-3', where Y= pyrimidine (C or T)) (Geets et al., 2006) and DSR-4R (5'-GTGTAGCAGTTACCGCA-3') (Wagner et al., 1998). The samples were diluted 1:2 in distilled water and the standard was from 10^7 to 10^2 copies of the *dsrB* gene per µl. Each reaction contained the same components and volumes as described for bacterial 16S rRNA quantification. The qPCR program for *dsrB* was initiated with a denaturation step for 15 min at 95 °C, followed by 36 cycles of denaturation for 35 seconds at 95 °C, annealing for 35 seconds at 55 °C and elongation for 30 seconds at 72 °C and 75 °C for 10 seconds. The

melting curve had the same thermal conditions as for bacterial 16S rRNA quantification. R^2 for the standard curve was 0.997 and Eff% was 87 %.

2.4.3.2 Quantification of the mcrA gene

The *mcrA* gene is a functional marker gene for methanogenesis and encodes the enzyme methyl coenzyme-M reductase. The *mcrA* gene was quantified using the primers ME3MF (5'-ATGTCNGGTGGHGTMGGSTTYAC-3', where N can be all four bases, H= not G (A, C or T), M= amino (A or C), S= Strong (C or G) and Y= pyrimidine (C or T)) and ME2R (5'-TCATBGCRTAGTTDGGRTAGT-3', where B= not A, R= purine (A or G) and D= not C) (Nunoura et al., 2008). The samples were diluted 1:2 in distilled water. The standard was generated from PCR products amplified from a methane leaking well in the North Sea. The standard was diluted from 1.33 x 10^6 to 13.3 copies of the *mcrA* gene per µl. Each reaction contained the same components and volumes as described for bacterial 16S rRNA quantification. The qPCR program for *mcrA* was initiated with a denaturation step for 15 min at 95 °C, followed by 50 cycles of denaturation for 30 seconds at 95 °C, annealing for 30 seconds at 54 °C and elongation for 45 seconds at 72 °C. The melting curve had the same thermal conditions as for bacterial 16S rRNA quantification.

2.4.3.3 Quantification of the pmoA gene

The *pmoA* gene is a marker gene for aerobic methanotrophs and encodes the enzyme methane monooxygenase. A PCR gradient (from 56.5 to 59 °C) was performed on the *pmoA* gene using the primers A189F (5'-GGNGACTGGGACTTCTGG-3', where N can be all four bases) and A650R (5'-ACGTCCTTACCGAAGGT-3') (Bourne et al., 2001) to see if this gene was present in the samples. Two samples (3 m and 16.5 m) were set up in a test run qPCR. The samples were diluted 1:2, 1:5 and 1:10 in distilled water. The standard was made from genomic DNA from *Methylococcus Capsulatus* (Bath), and ranged between 3.04 x 10^5 and 3.04 genomes per µl. The PCR program started with 15 minutes at 95 °C, followed by 40 cycles of 1 minute at 95 °C, primer annealing at 57 °C for 1 minute, and an elongation step for 35 seconds at 72 °C. The rest of the samples were not analysed.

2.4.4 16S rRNA amplicon 454 pyrosequencing

To determine the microbial community composition in the DH8 permafrost core, the 454 pyrosequencing method was used to sequence the 16S rRNA amplicons representing the prokaryotic community, using the GS FLX+/Titanium system (Roche). Amplicons from two parallel DNA samples from the six depths between 3 m and 57 m were sequenced, and one amplicon sample was sequenced from 60 m. Two samples (3 m and 16.5 m) from the potentially drill fluid contaminated outer edges of the core were included in the sequencing analysis, in order to investigate whether there had been any contamination by the drilling fluid.

2.4.4.1 Amplification of 16S rRNA using PCR

The first step of preparing the samples for amplicon 454 pyrosequencing was to amplify the 16S rRNA genes with PCR, using the carefully evaluated primer set (Jorgensen et al., 2012) Uni787F (5´-ATTAGATACCCNGGTAG-3´, where N can be all four bases) (Roesch et al., 2007) and Uni1391R (5´-ACGGGCGGTGWGTRC-3´, where W=weak (A or T), and R= purine (A or G)) (Jorgensen et al., 2012), targeting the V5-V8 region. Each reaction (20 µl) contained 10 µl 2x HotStar Taq master mixture (Qiagen), 0.2 µl of each primer (100 µM stock) and 8.4 µl distilled H_2O. To increase the PCR yield, 0.2 µl Bovine Serum Albumin (BSA) (20 mg/ml, Thermo Scientific) was added. The DNA was diluted 1:50 in distilled water for the 3 m samples, and 1:5 in the rest of the samples. The samples were run in triplicates, in order to minimise PCR drift. The PCR program was initiated with a hot start activation step for 5 minutes at 95 °C followed by 30 cycles of denaturation for 45 seconds at 95 °C, primer annealing for 45 seconds at 53 °C and an elongation step for 60 seconds at 72 °C. Before cooling down to 4 °C, a final elongation step for 7 minutes was enclosed to the program. The length of the PCR products was verified by 1.5 % agarose gel electrophoresis stained with GelRed (Biotium) at 60 V for 40 minutes.

2.4.4.2 Purification of PCR products

The triplicate PCR products were pooled to minimise PCR drifting, and purified using QIAquick PCR purification kit (Qiagen), in order to remove any contaminants, primers and enzyme leftovers in the PCR products. Five volumes of Buffer PB was added to 1 volume PCR sample, and mixed together. This was added to the QIAquick spin column and centrifuged for 60 seconds to bind the DNA. The flow through was discarded. 0.75 ml Buffer PE was added and centrifuged for 60 seconds, to wash the DNA on the filter. The flow through was discarded and the column was centrifuged for additional 60 seconds to completely remove the ethanol from the wash solution. After this, the column was placed in a clean 1.75 ml eppendorf tube and the DNA was eluted in 30-40 µl distilled H_2O. The purified PCR products were verified with 1 % gel electrophoresis stained with GelRed (Biotium) at 60 V for 40 minutes. If some of the bands were too weak, they were run again in PCR with modified dilutions.

2.4.4.3 Gel extraction

After PCR purification, the two parallel samples from the 3 m inner core and the single sample from 3 m outer edge had double bands on the agarose gel. This means that in addition to the gene of interest, a non-specific DNA product had been amplified. The PCR products from these samples were run on a 1 % Agarose gel made with 0.25 g low melting temperature Agarose (SeaPlaque GTG Agarose), 0.25 g Agarose (Sigma) in 50 ml 1 x TAE buffer. The gel was run at 60 V for 50 minutes. Under UV-light the desired bands were cut out from the gel using a sterile scalpel, and transferred to a sterilised Eppendorf tube. The gel slice was weighed, and the PCR products were extracted and purified from the gel using QIAquick Gel Extraction Kit (Qiagen), following the manufacturers' protocol. First, 3 volumes of Buffer QG were added to one volume of gel, and this was incubated at 50 °C for 10

minutes to dissolve the gel. Then, 1 gel volume of isopropanol was added to increase the yield of DNA fragments. This solution was added to a QIAquick spin column and centrifuged for 1 minute in order to bind the DNA. The flow through was discarded and 0.5 ml Buffer QG was added to the column and centrifuged for 1 minute to make sure that all agarose was removed. Next, 0.75 ml Buffer PE was added to the column and centrifuged for 1 minute to wash the DNA. The flow through was discarded, and the column was centrifuged for additional 1 minute to completely remove all ethanol from the wash solution. Then, the QIAquick column was placed in a clean 1.75 ml eppendorf tube and the DNA was eluted in 10 µl distilled H_2O and 25 µl of the product from a previous gel cleansing from the same sample (where the bands after purification turned out to be too weak), in order to obtain a higher DNA concentration. The purified PCR products were verified with 1 % gel electrophoresis stained with GelRed (Biotium) at 60 V for 40 minutes.

All samples had too low concentrations of PCR products after the first PCR run. To get higher concentrations, the PCR products were re-amplified in an additional PCR run, with dilutions from 1:90 to 1:250 and 20-25 cycles on the same PCR program as described above. The components of the PCR master mixture were the same as for the first 16S rRNA PCR amplification for 454 pyrosequencing, except that no BSA was added.

2.4.4.4 Attaching of Multiplex Identifiers

A second PCR round of 7 cycles was performed using the same primers and thermal conditions as described above for the 16S rRNA amplification, but linked to the 454 Life Sciences fusion primers, with the primer A and primer B adaptor sequences (**Fig. 5**). The forward primer was labelled with Multiplex Identifiers (MIDs), which is a unique barcode (different for each sample) that makes it possible to mix several samples into a single sequencing reaction (454 Life Sciences Corp, 2011).

Forward primer (Primer A, Lib-L):
5'-CCATCTCATCCCTGCGTGTCTCCGACTCAG-*{MID}*-*{template-specific sequence}*-3'

Reverse primer (Primer B, Lib-L):
5'-CCTATCCCCTGTGTGCCTTGGCAGTCTCAG-*{template-specific sequence}*-3'

Figure 5: Fusion primers for 454 amplicon sequencing (454 Life Sciences Corp, 2011).

Each reaction (20 µl) contained 12.5 µl 2x HotStarTaq® mastermix (Qiagen), 0.25 µl primer 1391R-B-Key and 2 µl 787F MID primer (different for each sample). A total amount of 10 µl H_2O and amplicon template was added, but the template volume was modified according to the concentration that showed at the agarose gel image. All samples were run in duplicates.

2.4.4.5 Bead purification

The PCR amplicons were purified again using AMPure XP bead purification (Agencourt), following the protocol from the Norwegian Sequencing Centre (2011). First, AMPure XP magnetic beads were added to the PCR products in the ratio 0.7:1. This solution was vortexed for 10 seconds, and incubated for 5 minutes at room temperature to allow binding of the PCR amplicons to the beads. Then the bead mix solution was placed on a magnet for 5 minutes to separate the beads from the solution. The clear solution was removed and 500 µl 70 % ethanol was added to the beads, and incubated for 1 minute at room temperature on the magnet. The ethanol solution was removed from the beads, and another 500 µl 70 % ethanol was added to the beads, incubated for 1 minute on the magnet, and the ethanol solution was removed. After this, the microtube was air dried for 5 minutes to evaporate any remaining ethanol. The beads were eluted in 25 µl 10 mM Tris-Cl, and vortexed for 20 seconds to release the PCR amplicons from the beads. This solution was placed on a magnet for 5 minutes to separate the beads from the solution. The clear solution, containing the PCR amplicons, was transferred to a clean microtube, and the purification was verified on a 1 % agarose gel at 60 V for 30 minutes.

2.4.4.6 DNA concentration

The concentration in the DNA amplicons was measured at Quantus Fluorometer (Promega). Some of the samples were additionally measured at Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). All amplicons (from 15 different samples) were pooled in a 1:1 ratio based on the measured DNA concentration (20 ng/µl from each sample), purified again with AMPure XP bead purification (Agencourt) and eluted in 30 µl 10 mM Tris-Cl. 18 µl of this solution, with a concentration of 24-27 ng/µl, were sequenced using multiplex GS FLX+/Titanium pyrosequencing (Roche) at the Norwegian High-Throughput Sequencing Centre (NSC) in Oslo, Norway.

2.4.5 Bioinformatic tools and procedures for the 454 pyrosequencing data

2.4.5.1 Filtering and removal of noise and chimeras from amplicon sequence data

The sequenced 16S rRNA amplicon dataset was returned from NSC as Standard Flowgram Format (SFF) files. Because of the large number of reads in the 16S rRNA amplicon 454 pyrosequencing method, it is important to distinguish noise introduced by the procedure, from the real sequence diversity. The SFF files were checked and treated with the algorithm AmpliconNoise using the script RunPreSplit_700.sh. This script removes primers, adaptors, too short/long sequences, chimeric sequences, sequencing errors and errors introduced in the PCR amplification (Quince et al., 2011).

2.4.5.2 Taxonomic classification

The filtered, de-noised and chimera-filtered output fasta files were submitted to the online taxonomic classification tool Classification Resources for Environmental Sequence Tags (CREST), using SilvaMod as a reference database. CREST is developed by researchers from

the University of Bergen, and is an alignment-based taxonomic classification method that uses the lowest common ancestor algorithm (Lanzen et al., 2012).

2.4.6 Clone library construction and sequencing

Three different clone libraries (two 16S rRNA clone libraries and one *dsrB* clone library) were made from the 54 m sample, to verify the amplicon 454 pyrosequencing results. In gene cloning, a gene fragment is inserted into a plasmid vector and then the plasmid gets inserted into *E.coli* cells that will be cultivated. The first step is to amplify the gene of interest with PCR using the enzyme *Taq* polymerase, which adds an A-overhang to the 3'-end of both strands of the PCR product. These single stranded A-overhangs base pairs with the 5'- U-overhangs of the plasmid vector. After the plasmid vectors have been inserted into the *E.coli* cells, the cells are cultivated on agar plates containing X-gal. X-gal is used to test if the cells have the enzyme β-galactosidase, with a technique called blue-white screening. Cell colonies that contain the plasmid with the correct gene insert will become white, and cells without the insert will become blue.

2.4.6.1 PCR amplification of 16S rRNA genes and dsrB genes

One 16S rRNA clone library was made with the primer pair used for 454 pyrosequencing of the 16S rRNA amplicons (Uni787F/Uni1391R, see "Materials and Methods", Section 2.4.4), and the same thermal conditions in PCR. The second 16S rRNA clone library was made with the primers Un515F (5´-CAGCMGCCGCGGTAA-3´, where M= amino (A or T)) (Lane *et al*., 1991) and Uni1391R (Jorgensen et al., 2012). The thermal conditions were 5 minutes at 95 °C, then 35 cycles of 95 °C for 30 seconds, 58 °C for 35 seconds and 72 °C for 1 minute, followed by 72 °C for 10 minutes. The *dsrB* clone library was constructed using the same primers that were used in qPCR (DSRp2060F/DSR-4R, see Section 2.4.3), and the same PCR thermal conditions, except that the enzyme activation was for 5 minutes at 95 °C. The sample was run in triplicate in all PCR runs, and a positive and a negative control was always included. The sample was diluted 1:5 in distilled water for 787F/1391R clone library, and 1:2 in distilled water for 515F/1391R and dsrB clone libraries. The PCR products were verified by 1 % agarose gel electrophoresis stained with GelRed (Biotium) for 30 minutes at 60 V.

2.4.6.2 Cloning

The cloning was performed with StrataClone PCR cloning kit (Agilent Technologies) following the manufacturers' protocol. The ligation reaction mixture was prepared by combining 3 µl StrataClone Cloning Buffer with 2 µl undiluted PCR product and 1 µl StrataClone Vector Mix amp/kan. This was incubated at room temperature for 5 minutes to allow the PCR products to be inserted into the vector. Next, 2 µl of the ligation reaction mixture was added to the StrataClone SoloPack competent cells and this was gently mixed together and incubated on ice for 20 minutes, to transfer the vectors into the cells. Then this mixture was heat-shocked at 42 °C for 45 seconds, followed by 2 minutes incubation on ice. After this, 250 µl LB medium (pre-warmed to 42 °C) was added to the competent cells and incubated for 1 hour

at 37 °C with agitation. After incubation, 100 µl and 50 µl of the transformation mixture was plated out on LB-ampicillin plates with X-gal, and incubated over night at 37 °C.

2.4.6.3 Amplification of gene inserts

The gene inserts were amplified in a PCR reaction using the primers M13-F (5'-ACTGGCCGTCGTTTTACAA-3') and M13-R (5'-GGAAACAGCTATGACCATG-3'). These primers amplify a part of the plasmid vector, as well as the PCR insert. White colonies were picked with a pipette tip and transferred to a PCR master mixture containing 10 µl HotStar Taq master mix, 0.1 µl of each primer (100 µM stock) and 8.8 µl distilled H_2O. One blue colony was picked as a negative control. The PCR program started with 5 minutes at 95 °C, then 30 cycles of 95 °C for 1 minute, 55 °C for 1 minute and 72 °C for 2 minutes, followed by 72 °C for 10 minutes. The PCR products were verified with 1 % agarose gel electrophoresis for 30 minutes at 60 V.

2.4.6.4 Sanger sequencing

In the Sanger sequencing reaction the amplicons from the M13 PCR reaction are copied by DNA polymerase. The BigDye 3.1 mixture contains nucleotides, DNA polymerase and fluorescently tagged dideoxynucleotides. Nucleotides are added to the single stranded template by DNA polymerase, and when a dideoxynucleotide gets incorporated the reaction is terminated. The dideoxynucleotide can be incorporated at any position, so a mixture of DNA strands with different lengths will be generated. The sequencing reaction contained 1 µl BigDye 3.1, 1 µl primer M13-F (33 pmol), 1 µl sequencing buffer, 6 µl distilled H_2O and 1 µl PCR product (1:20 diluted). This was run in PCR starting with 96 °C for 3 minutes, then 35 cycles of 96 °C for 12 seconds, 50 °C for 8 seconds and 60 °C for 4 minutes. A final extension step for 2 minutes was included, before cooling down to 4 °C. The sequences were analyzed at the Sequencing Facility at the University of Bergen, using capillary-based Applied Biosystem 3730XL Analyzer.

A total of 40 sequences from each primer set were obtained, and the sequences were trimmed from primer sequences and low quality sequencing ends using the software BioEdit. The sequenced 16S rRNA clones were taxonomically classified using CREST (Lanzen et al., 2012).

2.4.7 Phylogenetic analysis of *dsrB* sequences

The sequenced *dsrB* clones from the 54 m sample were translated into the correct amino acid reading frame using the Six Frame Translation tool of zbio.net (http://molbiol.ru/eng/scripts/01_13.html). The amino acid sequences were aligned using the online protein alignment tool ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). A rough phylogenetic analysis on amino acid level was performed using the program MEGA6 (Tamura et al., 2013) on the 39 permafrost generated *dsrB* clones, as well as most similar *dsrB* sequences obtained by search in BLASTp search facility of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/blast/), and *dsrB* sequences from known bacterial full-sequenced genomes, using maximum likelihood algorithms on an

alignment with no gaps. The analysed permafrost *dsrB* protein sequences were 102 amino acids long (*E.coli* position 165-267), which is about 1/3 of full length *dsrB* proteins. Thermodesulfovibrio and Archaeoglobus were used as outgroups for the bacterial *dsrB* tree.

2.4.8 Phylogenetic analysis of DSAG sequences

Phylogenetic analysis of the DSAG sequences obtained from the two 16S rRNA clone libraries and from 16S rRNA amplicon 454 pyrosequencing was performed by Steffen Leth Jørgensen. The DSAG sequences from this study were added to a previously published tree (Jorgensen et al., 2013) using the parsimony tool in ARB (Ludwig et al., 2004). The tree was generated from 497 published 16S rRNA gene sequences available in the Silva database release 111, and is based on neighbor joining and maximum likelihood algorithms (Jorgensen et al., 2013).

2.5 Statistics

Standard deviations were calculated using Microsoft Office Excel 2007. Pearson correlation and regression analyses were done in SPSS Statistics v22.0 (IBM). Multiple linear regression and model selection was done with a forward stepwise method, with significance level 0.05 for both model and individual predictors.

3. Results

3.1 Grain size distribution and carbon content

Initial descriptions of the top six sedimentary depths of the DH8 permafrost core included the grain size distributions with three main particle classes: clay (< 2 µm), silt (2-63 µm) and sand (63-2000 µm). The three top samples (3, 4.5 and 16.5 m) contained more sand than the deeper samples (47.1-88.5 % sand), with the highest sand content at 4.5 m (88.5 %) (**Table 2**). The three lower sedimentary samples (31.5, 54 and 57 m) contained more silt (82.6-98.1 %). The clay content was low in all samples ranging between 0.02 and 1.7 % (**Table 2**). See Appendix A for grain size distribution curves.

Table 2: Grain size distribution, Total Organic Carbon (TOC) and Total Inorganic Carbon (TIC) content in the six sedimentary permafrost horizons.

Depth (m)	Geomorphology	Grain Size			Carbon	
		% Clay	% Silt	% Sand	TOC (Weight %)	TIC (Weight %)
3	Delta top/tidal	0,02	52,92	47,06	1,76	0,93
4,5	Delta top/tidal	0,04	11,52	88,45	1,00	0,94
16,45	Delta front	0,15	43,40	56,45	1,37	1,01
31,5	Prodelta	0,11	94,11	5,78	1,11	1,05
54	Marine	0,52	98,09	1,39	1,23	1,02
57	Melt-out till	1,70	82,57	15,73	1,15	1,14

Further characterisation of the six sedimentary permafrost horizons included measurements of carbon content in the solid phase of the sediment. Total Organic Carbon (TOC) was calculated from the difference in Total Carbon (TC) and Total Inorganic Carbon (TIC). TOC ranged from 1.76 ± 0.2 weight % carbon at 3 m depth to 1.15 weight % carbon at 57 m (**Table 2**). TOC was lowest at 4.5 m (1.00 ± 0.1 weight % carbon). The Total Inorganic Carbon (TIC) content was between 0.93 ± 0.05 and 1.14 weight % carbon throughout the core.

3.2 Elemental composition

The major element composition in the solid phase of the six sedimentary samples was analysed by X-ray fluorescent (XRF) spectroscopy. The MgO, Al_2O_3 and K_2O content were lower from 3 m to 16.5 m than from 31.5 m to 57 m (**Table 3**). The SiO_2 content was higher at 3 m, compared to the other depths, and low at 31.5 m and 54 m. CaO showed lower content from 3 m to 31.5 m, than at 54 m and 57 m. The rest of the elements revealed minor differences with depth (**Table 3**).

Table 3: Major element composition analysed by X-ray fluorescent spectroscopy (XRF).

Depth (m)	Na_2O (%)	MgO (%)	Al_2O_3 (%)	SiO_2 (%)	P_2O_5 (%)	K_2O (%)	CaO (%)	TiO_2 (%)	MnO (%)	Fe_2O_3 (%)
3	1,46	0,69	10,37	72,62	0,14	2,25	0,31	0,51	0,03	4,97
4.5	1,30	0,98	11,89	69,74	0,18	2,51	0,19	0,58	0,03	6,17
16.5	1,74	0,88	11,66	69,03	0,15	2,52	0,30	0,58	0,02	4,75
31.5	1,86	1,28	14,91	63,52	0,20	3,07	0,31	0,76	0,03	5,96
54	1,69	1,52	15,77	61,22	0,21	3,27	0,62	0,76	0,03	5,85
57	1,36	1,16	12,33	68,07	0,15	2,64	0,61	0,65	0,04	5,16

3.3 Concentration of extracted pore water ions

To access the pore water of the permafrost, extraction by Rhizon samplers, commonly used for sediment cores, was unsuccessful due to very dry sediment. Because of this, the pore water ions were extracted through a two-step procedure; first H_2O extraction and then $CaCl_2$ extraction of the very same sediment. The $CaCl_2$ extraction was performed to potentially extract a higher amount of cations that are associated with mineral particles (see "Materials and Methods", Section 2.3.4). The concentrations measured in the extracts were calculated back to represent the concentrations in the original pore water volume. For the cations (except Ca^{2+}) the results in **Fig. 6**/**Table 4** are presented as the sum of the H_2O and $CaCl_2$ extractions.

The general trend from the combination of subsequent H_2O and $CaCl_2$ extractions was that the monovalent cations K^+ and Na^+, but also Si^{4+} were more extractable with H_2O than the divalent cations. For K^+, Na^+ and Si^{4+}, 33-88 % of the total ion concentration (H_2O extracted ions + $CaCl_2$ extracted ions) was extracted with H_2O (**Appendix B**). The exception is the K^+ concentration in the shale sample (60 m), where only 13 % of the total K^+ concentration was extracted with H_2O. For the divalent cations such as Mn^{2+}, Mg^{2+} and Sr^{2+}, 74-100 % of the total ion concentration got extracted with $CaCl_2$. Fe^{2+}, however, showed some variations with depth regarding how much of the ions that got extracted with H_2O or $CaCl_2$ (**Appendix B**).

The Cl^- concentrations were more similar from 4.5 m (327 mM) to 57 m (521 mM) depth, but showed lower values at top layer of 3 m (71 mM) (**Fig. 6**, left panel). This was also the trend for the Na^+ concentration (concentration from H_2O extraction + $CaCl_2$ extraction), which varied between 450 mM and 629 mM from 4.5 m to 57 m, but at 3 m the Na^+ concentration was lower (99 mM). In the shale sample (60 m), the Na^+ concentration was much higher than in any of the other samples (2148 mM) (**Fig. 6,** left panel). The SO_4^{2-} concentration was relatively low throughout the core (1-16 mM), except at 16.5 m where the concentration was 56 mM (**Fig. 6**).

The concentration of NO_3^- was low from 3 m to 57 m (0.02-0.5 mM). In the shale sample (60 m), the NO_3^- concentration was higher (3 mM) (**Fig. 6**). At 3 m and 4.5 m the NH_4^+ concentration was lower (0.7 and 0.5 mM) than in the rest of the core. From 16.5 m to 57 m the NH_4^+ concentration varied between 4 and 7 mM, and at 60 m the NH_4^+ concentration was higher (25 mM) (**Fig. 6**). The Mn^{2+} and Fe^{2+} concentrations (concentration from H_2O extraction + $CaCl_2$ extraction) showed a similar profile, with higher concentrations in top layer of 3 m and in the shale sample at 60 m (**Fig. 6**, right panel).

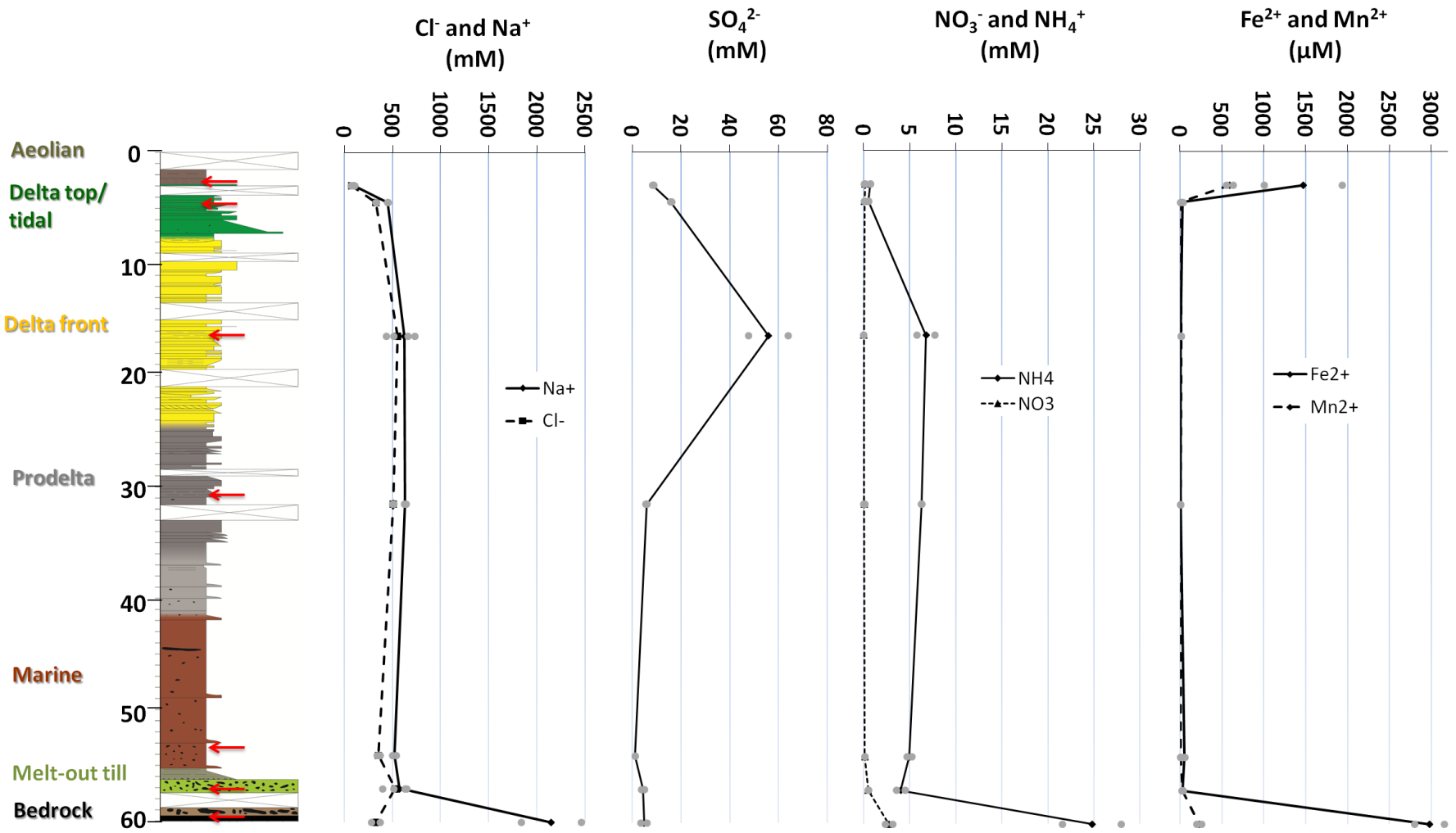

Figure 6: Concentrations of Cl^-, Na^+, SO_4^{2}, NO_3^-, NH_4^+, Mn^{2+} and Fe^{2+} (mM and µM) from extracted pore water, together with geomorphological description of the DH8 permafrost core (Gilbert, 2014). Na^+, Mn^{2+} and Fe^{2+} were extracted with H_2O and $CaCl_2$, and the concentration shown here is the sum of the two extractions. The remaining ions were extracted with H_2O only. The plotted lines and black symbols are the mean of two measurements from each depth, and the grey symbols represent the individual data points. Red arrows highlight the seven selected sampling depths.

pH and total alkalinity was also measured in the extracted pore water. In general both the pH and alkalinity increased with depth (**Table 4**). The pH varied between 6.4 and 9.3 throughout the core, and the alkalinity was between 0.4 and 1.7 (**Table 4**). Most of the ion concentrations in **Fig. 6** and **Table 4** show higher values in the shale sample (60 m) than in the sedimentary samples. The concentration of Ca^{2+} varied between 0.4 and 2.4 mM. The K^{+} concentration showed lower values at 3 m (18.7 mM), than from 4.5 m to 57 m (45.8-55.8 mM). The Mg^{2+}, Si^{4+} and Sr^{2+} concentrations showed little variation from 3 m to 54 m, but had higher concentrations at 57 m and 60 m (**Table 4**). The concentration of Br^{-} ranged between 0.1 µM and 0.7 µM. PO_4^{2-} showed lower concentrations from 3 m to 31.5 m (4.8-54.2 µM), than from 54 m to 60 m (112.6-1218.6 µM) (**Table 4**).

Table 4: pH, alkalinity and ion concentrations from extracted pore water. K^{+}, Mg^{2+}, Si^{4+} and Sr^{2+} were extracted with H_2O and $CaCl_2$, and the concentration shown here is the sum of the two extractions. The remaining ions were extracted with H_2O only.

Depth (m)	pH	Alkalinity	Ca^{2+} (mM)	K^{+} (mM)	Mg^{2+} (mM)	Si^{4+} (mM)	Sr^{2+} (mM)	Br^{-} (µM)	PO_4^{2-} (µM)
3	6,4	--	0,7	18,7	37,1	1,4	0,2	0,1	17,2
4,5	6,8	0,4	0,4	45,8	45,4	1,2	0,2	0,4	54,2
16,5	6,6	BD	2,4	46,4	65,2	1,2	0,4	0,7	4,8
31,5	7,1	0,5	0,4	53,0	48,9	0,7	0,3	0,7	28,4
54	8,5	1,7	1,1	51,2	46,1	1,0	0,3	0,5	112,6
57	7,8	1,0	1,1	55,8	70,5	2,2	1,0	0,2	197,8
60	9,3	1,5	0,4	475,7	361,2	101,5	10,9	BD	1218,6

-- Not measured

BD: Below detection limit

3.4 Microbial abundance determined by quantitative PCR (qPCR)

All molecular analyses were based on several parallel DNA extractions from each of the seven selected permafrost depths. Initial analysis focussed on the numbers of bacterial and archaeal 16S rRNA genes using quantitative PCR (qPCR) by amplification of bacterial and archaeal 16S rRNA genes separately. The overall trend in bacterial and archaeal 16S rRNA copies g^{-1} sediment showed high copy numbers in top sample at 3 m, a decrease in copy numbers at 4.5 m, and a decrease in 16S RNA copies from 54 m to 60 m (**Fig. 7**, left panel). Bacterial 16S rRNA copy numbers ranged between 5.84×10^8 ($\pm 5.35 \times 10^7$) copies g^{-1} sediment at 3 m depth to 2.42×10^5 ($\pm 4.38 \times 10^4$) copies g^{-1} sediment in the 60 m shale sample (**Fig. 7,** left panel). Archaeal 16S rRNA copies g^{-1} sediment were in the range of 2.98×10^6 ($\pm 5.55 \times 10^4$) at the top of the core, ending at 2.48×10^5 ($\pm 8.84 \times 10^4$) at 60 m (**Fig. 7,** left panel). The highest archaeal 16S rRNA copy numbers g^{-1} were 1.51×10^7 ($\pm 2.65 \times 10^6$) at 54 m depth. Archaeal 16S rRNA gene copy numbers made up 0.6-54 % of total 16S rRNA genes.

Since the uncharacterised Deep Sea Archaeal Group (DSAG) was one of the most abundant classes in the microbial community composition in this study (see Section 3.6.2), a qPCR

analysis specifically targeting 16S rRNA genes of DSAG was performed by Steffen Leth Jørgensen. Organisms belonging to DSAG were detected from 16.5 m to 57 m. The shape of the plotted line for DSAG 16S rRNA follows the plotted line for archaeal 16S rRNA nicely, with 16S rRNA copy numbers g^{-1} sediment varying between 1.56 x 10^6 (± 4.44 x 10^5) and 1.42 x 10^7 (± 6.55 x 10^6) (**Fig. 7,** left panel). Interestingly, the DSAG 16S rRNA gene copy numbers made up 5-50 % of total 16S rRNA genes and 45-188 % of archaeal 16S rRNA genes from 16.5 to 57 m.

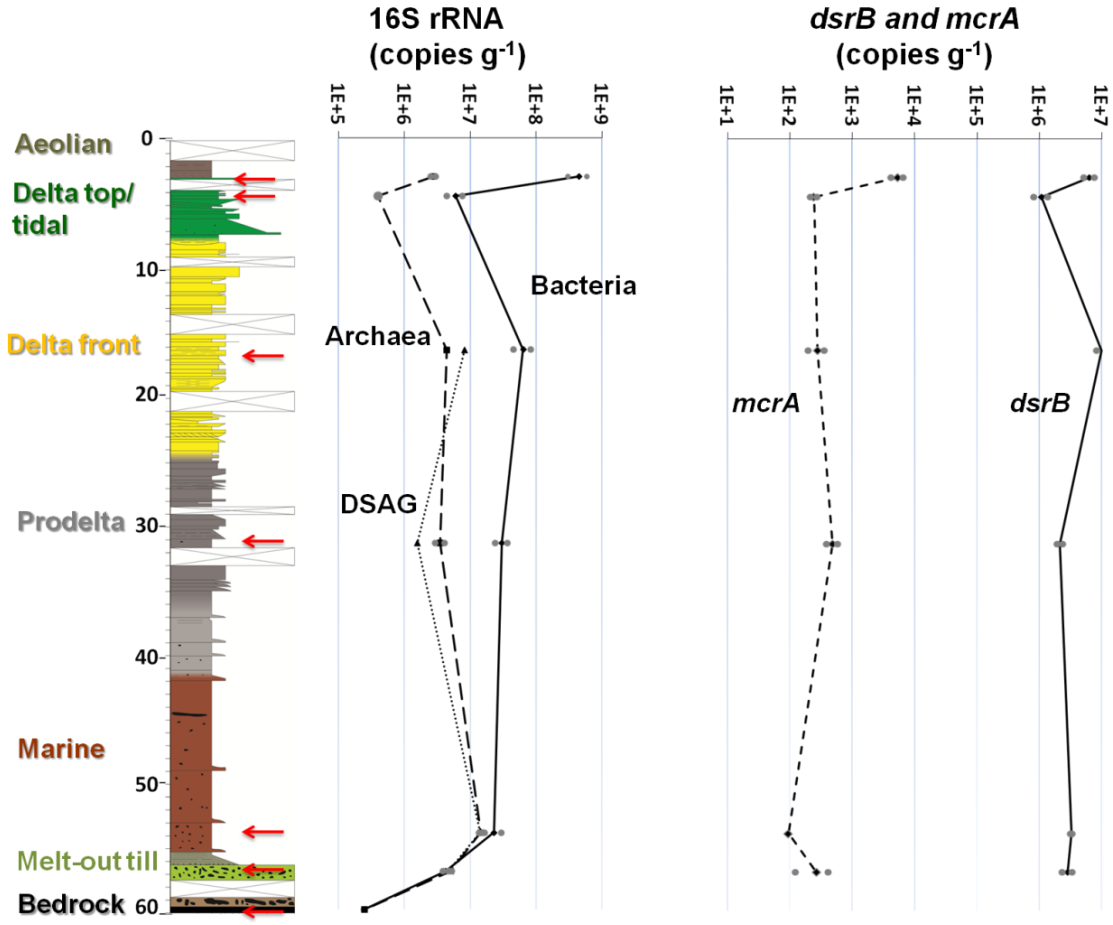

Figure 7: Quantitative PCR (qPCR) data (logarithmic scale) plotted against depth for 16S rRNA Bacteria, Archaea and Deep Sea Archaeal Group (DSAG), *mcrA* (methyl coenzyme M-reductase) and *dsrB* (dissimilatory sulfite reductase β-subunit), together with the geomorphological description (Gilbert, 2014) of the 60 m permafrost core. No DSAG 16S rRNA genes were detected at 3 m and 4.5 m depth. The plotted lines and black symbols are the mean of two measurements from each depth, and the grey symbols represent the individual data points. Red arrows highlight the seven selected sampling depths. See Appendix C for standard deviations.

As most bacteria have more than one copy of the 16S rRNA gene in their genome, approximate bacterial cell numbers were calculated by dividing the 16S rRNA gene copy number with 4.2, being the average number of 16S rRNA copies per bacterial genome (Vetrovsky and Baldrian, 2013). The archaeal 16S rRNA gene copy numbers were not corrected, as most archaea were found to belong to the DSAG (**Fig. 7**), where preliminary

genome data show that they most likely have only one 16S rRNA gene per genome (Steffen Leth Jørgensen, personal communication). When correcting the bacterial 16S rRNA copy numbers to approximate cell numbers, the relationship between Bacteria and Archaea becomes highly similar. While Bacteria dominate in the 3 m to 16.5 m portion of the core, the abundance is nearly similar at 31.5 m, with archaeal dominance throughout the rest of the deeper core including the shale layer at 60 m (**Fig. 8**). Total prokaryotic cell numbers (Bacteria + Archaea) were between 3 x 10^5 and 1 x 10^8 cells g^{-1} sediment. Archaea cell numbers made up 2.5-83 % of total cell numbers.

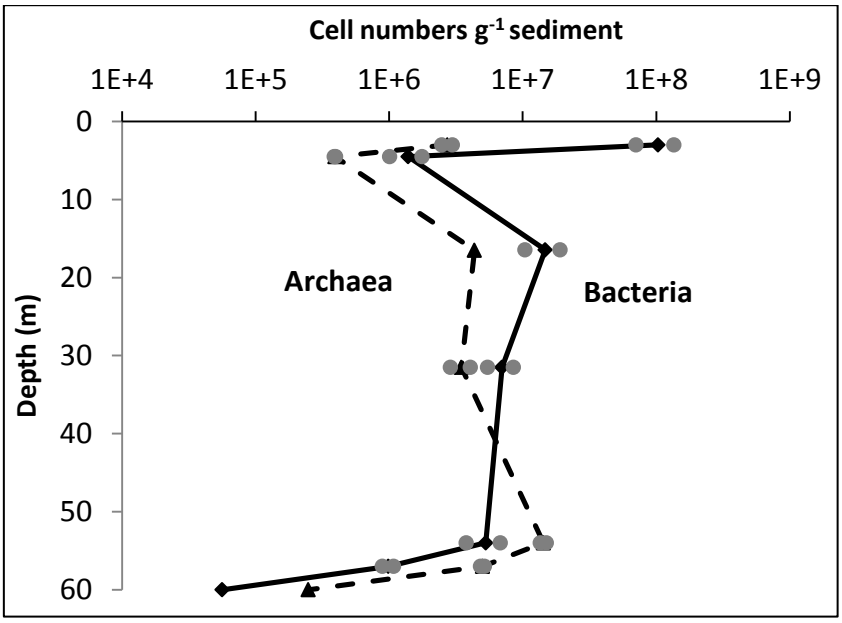

Figure 8: Bacterial and archaeal cell numbers per gram sediment plotted logarithmically against depth (m). Bacterial 16S rRNA copies (squares) have been divided by 4.2 (average number of 16S rRNA genes per bacterial genome, see text). The plotted lines and black symbols are the mean of two measurements from each depth, and the grey symbols represent the individual data points.

3.5 Abundance of *dsrB*, *mcrA* and *pmoA* genes

By using primers targeting prokaryotic marker genes encoding enzymes involved in key metabolic pathways, the genetic potential of that microbial function or metabolism can be assessed. Here, microbial sulphate reduction, methanogenesis and methanotrophy were targeted through quantification of the marker genes *dsrB* (dissimilatory sulfite reductase β-subunit), *mcrA* (methyl coenzyme M-reductase) and *pmoA* (particulate methane monooxygenase), respectively.

The *dsrB* gene was detected and quantified in all samples, except in the shale sample at 60 m. The shape of the plotted *dsrB* curve follows the shape of the curve for bacterial 16S rRNA, with higher gene copy numbers at 3 m (7.57 x 10^6 ± 1.01 x 10^6), a decrease at 4.5 m (8.10 x 10^5 ± 3.64 x 10^5) and then an increase at 16.5 m (1.17 x 10^7 ± 1.99 x 10^6) (**Fig. 7**, right panel). *dsrB* gene copy numbers made up 1-31 % of total 16S rRNA genes and 1-67 % of bacterial 16S rRNA genes from 3 m to 57 m.

The *mcrA* gene was detected from 3 m to 57 m, with highest copy numbers at 3 m (6.54 x 10^3 ± 1.34 x 10^3) and lowest copy numbers at 54 m (9.44 x 10^1 ± 2) (**Fig. 7,** right panel). These are very low numbers and were on the limit of detection for the qPCR setup.

pmoA was detected at 3 m and 16.5 m using conventional PCR, but the PCR products showed double bands when analysed with gel electrophoresis in all temperatures of the tested PCR gradient (56.5 °C to 59 °C), indicating amplification of an unspecific DNA product. At 59 °C the *pmoA* PCR products were very weak so increasing the temperature any further was not an option. The two mentioned samples were run in *pmoA* specific qPCR setup, but the results showed very low copy numbers (< 10^2), and the melting curve showed that two products had been amplified. Because of this, qPCR analysis for *pmoA* was not done for the other depths.

3.6 Microbial community composition and diversity

To investigate the microbial community composition and diversity among the different depths, a 16S rRNA gene amplicon library was generated with a primer set targeting the prokaryotic V5-V8 region (amplifying both bacterial and archaeal taxa). Two samples were sequenced from each depth, except at 60 m and in the outer edge samples where only one sample was sequenced. The results presented here are the sums of the two samples. 16S rRNA amplicon 454 pyrosequencing yielded a total of 82 126 raw reads. After treatment with AmpliconNoise (Quince et al., 2011), to remove short sequences, noise and chimeric sequences, this number decreased to 25 106 reads. The read numbers were unevenly distributed with depth, ranging from 55 to 6788 reads (**Table 5**).

Table 5: Overview of read numbers from 16S rRNA amplicon 454 pyrosequencing.

Depth (m)	Number of reads before AmpliconNoise	Number of reads after AmpliconNoise	Bacterial reads	Archaeal reads	% Bacteria	% Archaea
3	14486	3741	3234	507	86,4	13,6
4.5	5650	1383	898	485	64,9	35,1
16.5	19632	6788	714	6075	10,5	89,5
31.5	12216	3840	636	3204	16,6	83,4
54	15202	5178	74	5104	1,4	98,6
57	2948	1036	2	1034	0,2	99,8
60	3126	730	25	705	3,4	96,6
3*	8684	2355	1898	457	80,6	19,4
16.5* [a]	182	55	5	50	9,1	90,9

* Sample from the outer edge of the DH8 core (see “Materials and Methods”, Section 2.2).

[a] Was not included in further analysis due to low read numbers.

In the 16S rRNA gene amplicon library, the microbial community composition was dominated by bacterial reads at 3 m and 4.5 m. From 16.5 m to 60 m there was a dominance of archaeal reads. At 54 m and 57 m only 1.4 % and 0.2 % of the reads belonged to Bacteria (**Table 5**). This is in contrast to the qPCR data, showing similar 16S rRNA copy numbers for Bacteria and Archaea from 54 m to 60 m (**Fig. 7**).

3.6.1 Bacterial diversity and community composition

A total of 25 bacterial phyla were identified in the samples. A significant part of the reads belonged to poorly categorised groups like Candidate division TM7, OP9, and Chloroflexi (**Fig. 9**). Actinobacteria and Proteobacteria were also detected in most of the samples. The bacterial community at 3 m was different from the other depths with Candidate division TM7 as the most abundant phylum (67 % of the bacterial community). This phylum was not present in any of the other samples. From 4.5 m to 57 m Chloroflexi dominated the bacterial community composition. The bacterial community in the 60 m shale sample was also different from the other samples, dominated by Thermodesulfobacteria (**Fig. 9**).

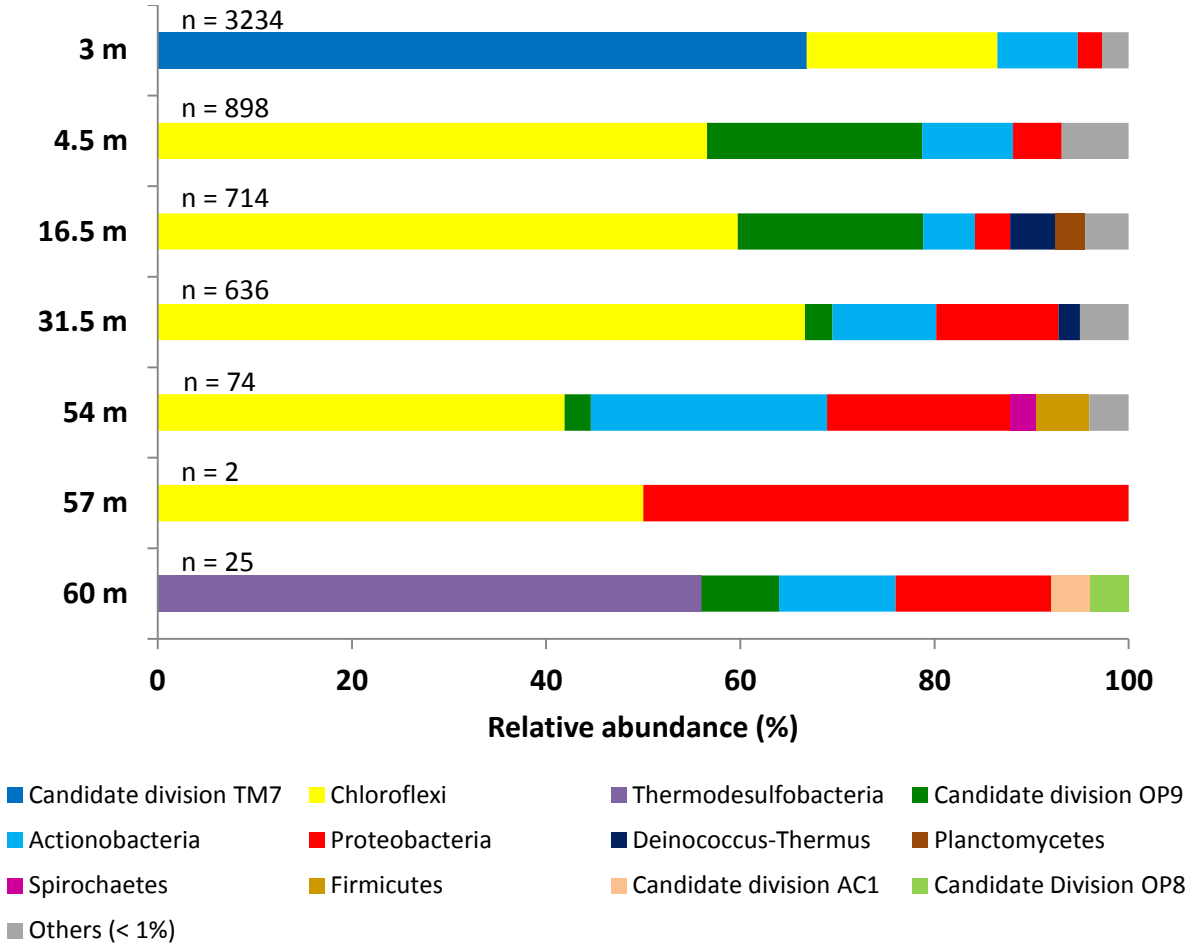

Figure 9: Relative abundance distribution of bacterial phyla obtained by 16S rRNA amplicon 454 pyrosequencing. The sediment depths are indicated to the left. The group "Others (< 1%)" is the collection of reads on phylum level that singlehanded constituted less than 1 % of the total bacterial amplicon reads. n= number of reads.

Moving to class level, in total 47 bacterial classes were identified. The most abundant bacterial classes were Chloroflexi Subdivision 11, Anaerolineae (Chloroflexi), Thermoleophilia (Actinobacteria), Gammaproteobacteria and Alphaproteobacteria (**Fig. 10**). Chloroflexi Subdivision 11 dominated down to 31.5 m, and Anaerolineae was the dominating bacterial class at 54 m and 57 m. The shale sample from 60 m was dominated by Thermodesulfobacteria (class).

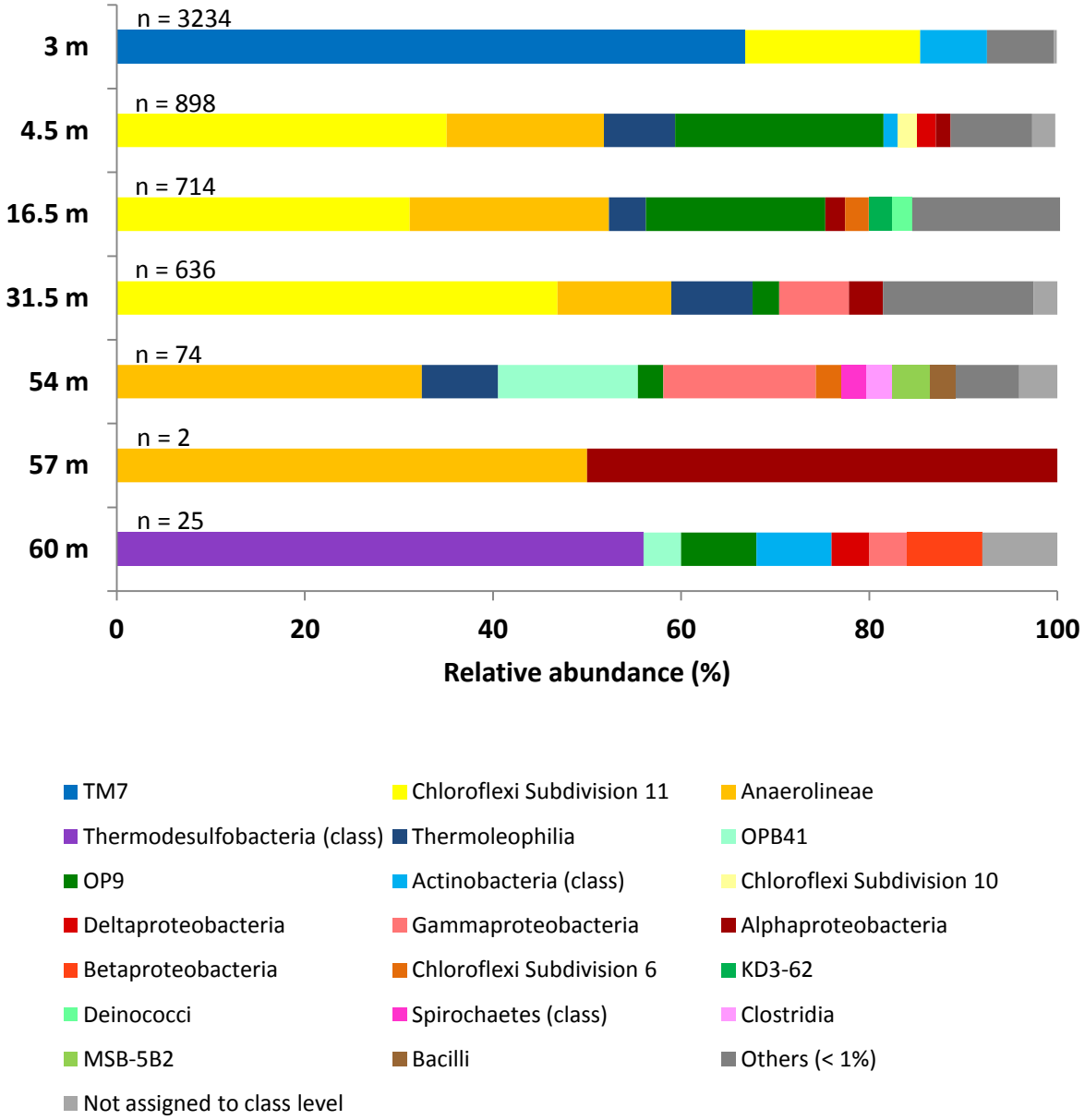

Figure 10: Relative abundance distribution of bacterial classes obtained by 16S rRNA amplicon 454 pyrosequencing. The sediment depths are indicated to the left. The group "Others (< 1%)" is the collection of reads on class level that singlehanded constituted less than 1 % of the total bacterial amplicon reads. The group "Not assigned to class level" refers to the reads that could not be classified by CREST to a deeper level than phylum. n = number of reads.

3.6.2 Archaeal diversity and community composition

Three archaeal phyla, the Crenarchaeota, Euryarchaeota and Thaumarchaeota, were detected in the permafrost samples. There was an increase in the abundance of Crenarchaeota from 3 m to 57 m (**Fig. 11**). At 3 m and 60 m Euryarchaeota dominated the overall archaeal community.

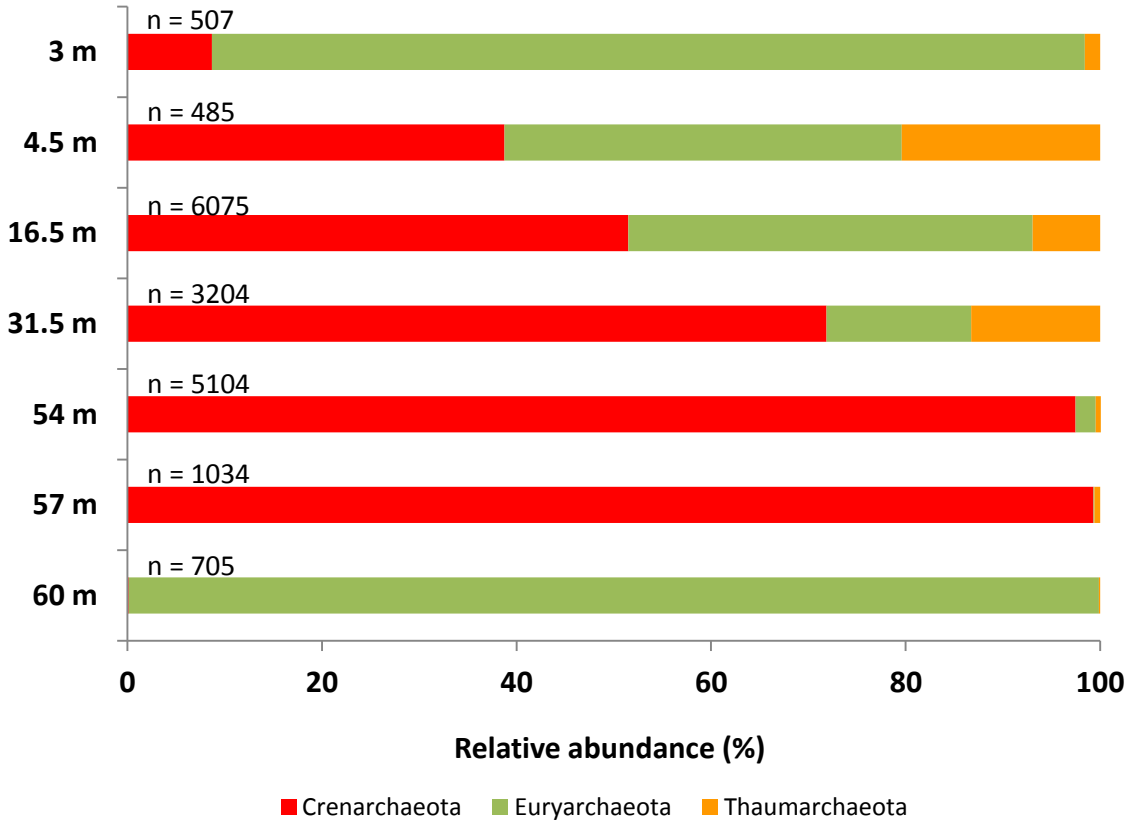

Figure 11: Relative abundance distribution of archaeal phyla obtained by 16S rRNA amplicon 454 pyrosequencing. The sediment depths are indicated to the left. n = number of reads.

Twelve different archaeal classes were identified. As for Bacteria, the archaeal community was different at 3 m, when compared to the other depths, where the Euryarchaeotal classes Halobacteria, Methanobacteria and Methanomicrobia were the most abundant (**Fig. 12**). At 4.5 m the Miscellaneous Crenarchaeotic Group (MCG) and Methanomicrobia (Euryarchaeota) were the dominating the archaeal classes. From 16.5 m to 57 m the uncharacterised DSAG (Crenarchaeota) comprised 40-99 % of the archaeal community. In the shale sample (60 m) Thermoplasmata (Euryarchaeota) and Archaeoglobi (Euryarchaeota) were the dominating archaeal classes (**Fig. 12**).

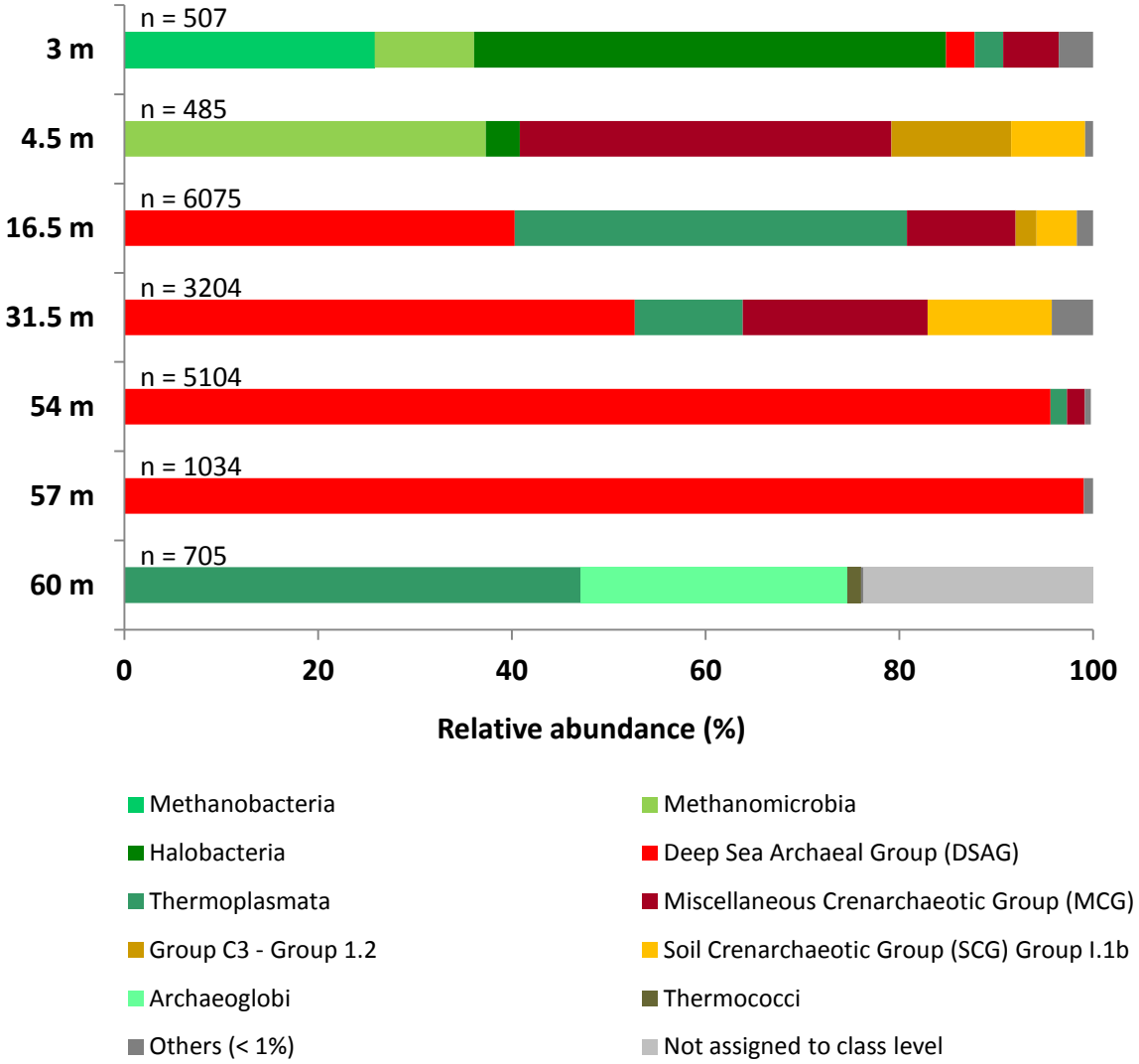

Figure 12: Relative abundance distribution of archaeal classes obtained by 16S rRNA amplicon 454 pyrosequencing. The sediment depths are indicated to the left. The group "Others (< 1%)" is the collection of reads on class level that singlehanded constituted less than 1 % of the total archaeal amplicon reads. The group "Not assigned to class level" refers to the reads that could not be classified by CREST to a deeper level than phylum. n= number of reads.

3.6.3 Microbial community composition in outer edge sample from 3 m

Two samples (3 m and 16.5 m) from the potentially drill-fluid (river water from Adventelva) contaminated outer edges of the core were sequenced to compare the microbial communities in the outer edge samples with the inner core samples. The outer edge sample from 16.5 m obtained only 182 raw reads after amplicon 454 pyrosequencing (**Table 5**), and this sample was not included in any further analyses. The distribution of bacterial phyla in the outer edge sample from 3 m compared to the inner core sample of the same depth was highly similar (**Fig. 13**). Candidate division TM7 was the dominating phyla, with a slightly higher relative abundance in the outer edge sample compared to the inner core.

The distribution of archaeal phyla was also highly similar between the outer edge and inner core sample from 3 m (**Fig. 13**). The only present-absent between the inner and outer cores on phylum level were the Thaumarchaeota. Thaumarchaeota was present with 1.6 % relative abundance in the inner core sample, but not detected in the outer edge sample.

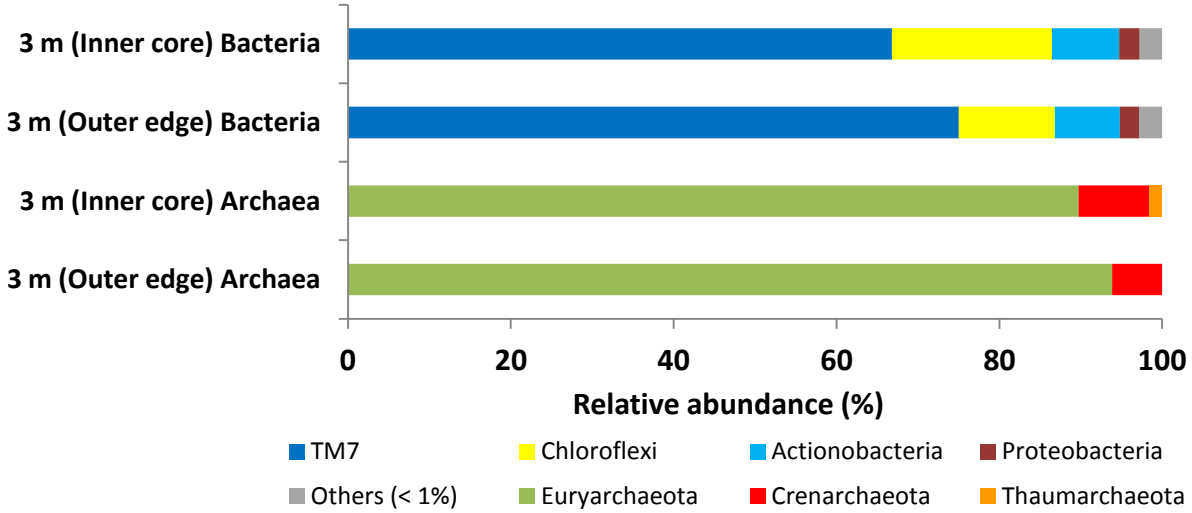

Figure 13: Relative abundance distribution of bacterial and archaeal phyla at 3 m obtained by 16S rRNA amplicon 454 pyrosequencing. The outer edge sample is the potentially drill fluid contaminated cut-off from the core. The inner core sample is assumed to be free from contamination. The group "Others (< 1%)" is the collection of reads on phylum level that singlehanded constituted less than 1 % of the total bacterial or archaeal amplicon reads.

On class level the bacterial distribution was also highly similar in the outer edge sample from 3 m compared to the inner core sample of the same depth (**Fig. 14**). The class Acidimicrobiia was only detected in the outer edge sample, but this class accounted for just 1.2 % of the overall bacterial community. The archaeal classes DSAG and Thermoplasmata were only detected in the inner core sample with together 6 % relative abundance (**Fig. 14**). Halobacteria was more abundant in the outer edge sample than in the inner core sample.

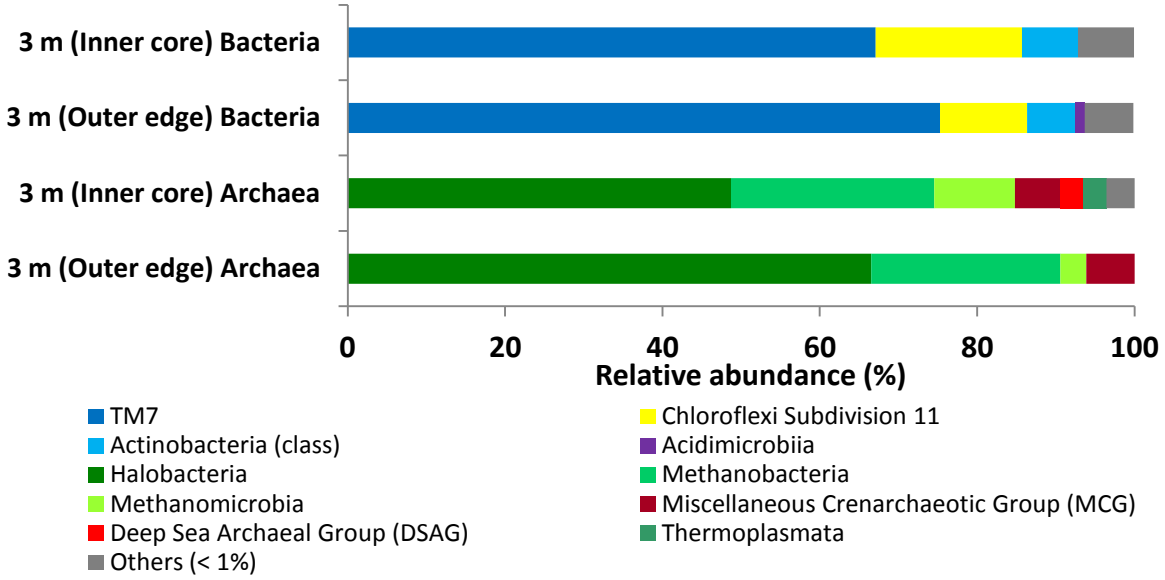

Figure 14: Relative abundance distribution of bacterial and archaeal classes at 3 m obtained by 16S rRNA amplicon 454 pyrosequencing. The outer edge sample is the potentially drill fluid contaminated cut-off from the core. The inner core sample is assumed to be free from contamination. The group "Others (< 1%)" is the collection of reads on class level that singlehanded constituted less than 1 % of the total bacterial or archaeal amplicon reads.

3.7 Microbial community composition and diversity from 16S rRNA clone libraries

The relationship between Archaea and Bacteria from the 16S rRNA amplicon library did not correspond with the relationship between Archaea and Bacteria from the qPCR results. Additionally, no known sulphate-reducing microbial groups were found in the 16S rRNA amplicon library, while the qPCR showed that the *dsrB* gene was abundant from 3 m to 57 m. Because of this, two 16S rRNA clone libraries were made of the sample from 54 m, using two different sets of universal prokaryotic primers (787F/1391R and 515F/1391R). This was done in order to investigate if the primers used for amplicon 454 pyrosequencing had underestimated the bacterial diversity, even though these primers theoretically should cover 98 % of all known prokaryotes (Jorgensen et al., 2012). The taxonomic classification of the sequences from the two 16S rRNA clone libraries obtained from Classification Resources for Environmental Sequence Tags (CREST) was plotted in bar charts for phylum and class level, together with the taxonomic classification from the 16S rRNA amplicon library. Bacterial and archaeal sequences are plotted together in the same diagram here, because of the low sequence number in the clone libraries.

A total of 39 clones were sequenced from the 16S rRNA clone library where the primers 787F/1391R were used. Of these 39 sequences, 34 belonged to the Crenarchaeota phylum (**Fig. 15**). Two sequences affiliated with Euryarchaeota, and three sequences were bacterial and belonged to Actinobacteria. In the 16S rRNA clone library where the primers 515F/1391R were used, 40 clones were sequenced. Only three of the 40 sequences affiliated with Archaea, in contrast to what was found in the 787F/1391R clone library. These three

sequences belonged to Crenarchaeota. The remaining 37 sequences belonged to five different bacterial phyla, dominated by Actinobacteria and Proteobacteria (**Fig. 15**).

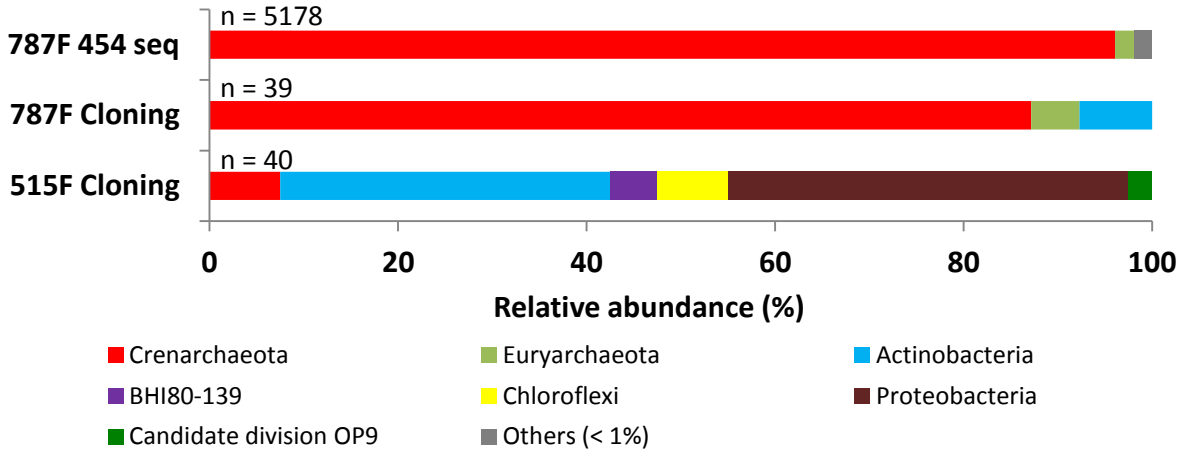

Figure 15: Relative abundance distribution of bacterial and archaeal phyla in the 16S rRNA clone libraries from the sample at 54 m. From the top going down: Taxonomic classification obtained from 16S rRNA amplicon 454 pyrosequencing, using the primers 787F/1391R. Middle bar: taxonomic classification obtained from 16S rRNA clone libraries using the primers 787F/1391R. Bottom bar: taxonomic classification obtained from 16S rRNA clone libraries using the primers 515F/1391R. The group "Others (< 1%)" is the collection of reads that singlehanded constituted less than 1 % of the total amplicon reads. n = number of reads/sequences.

On class level, 33 of the 34 Crenarchaeotal sequences from the clone library made with the primers 787F/1391R, belonged to DSAG and one sequence affiliated with MCG (**Fig. 16**). The two Euryarchaeotal sequences were further classified as Thermoplasmata. The three sequences belonging to Actinobacteria were classified as OPB41 and Thermoleophilia. In the 16S rRNA clone library made with the primers 515F/1391R, the three Crenarchaeotal sequences belonged to DSAG (**Fig. 16**). The dominating bacterial classes were Thermoleophilia (Actinobacteria), Gammaproteobacteria and OPB41 (Actinobacteria).

When comparing the taxonomy of these two different 16S rRNA clone libraries with the taxonomy from 16S rRNA amplicon 454 pyrosequencing, it is evident that the taxonomy from the 16S rRNA clone library made with 787F/1391R primers is similar to the 454 pyrosequencing taxonomy, with DSAG as the dominating group (**Fig. 16**). The taxonomy from the 16S rRNA clone library made with 515F/1391R primers is different and dominated by Bacteria, but the dominance by DSAG within Archaea is confirmed (**Fig. 16**).

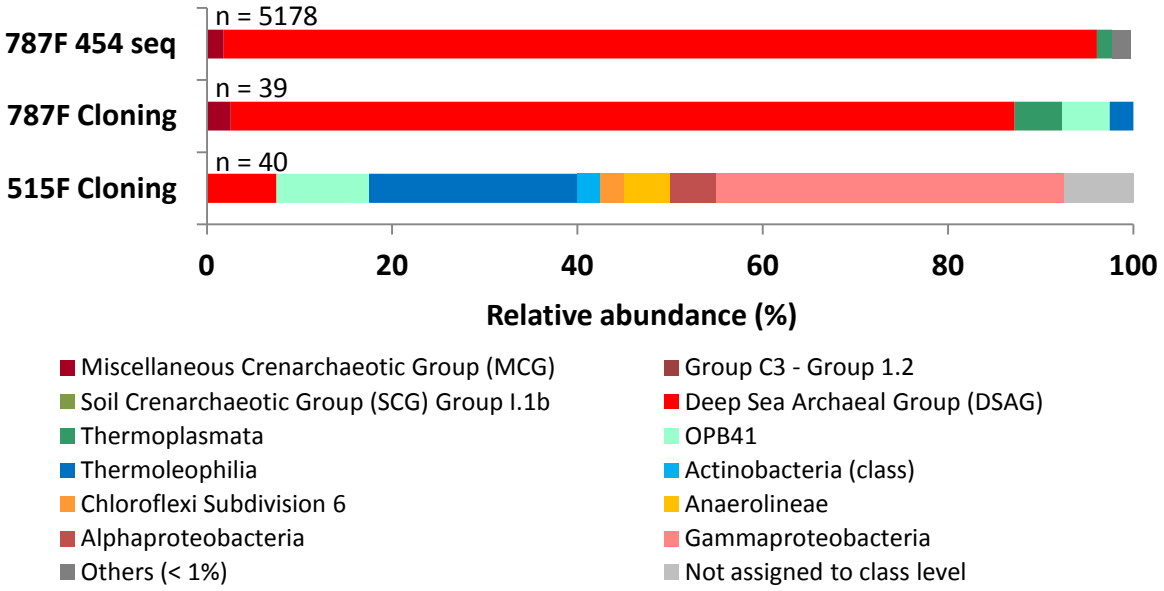

Figure 16: Relative abundance distribution of bacterial and archaeal classes in the 16S rRNA clone libraries from the sample at 54 m. From the top going down: Taxonomic classification from 16S rRNA amplicon 454 pyrosequencing, using the primers 787F/1391R. Middle bar: taxonomic classification obtained from 16S rRNA clone libraries using the primers 787F/1391R. Bottom bar: taxonomic classification obtained from 16S rRNA clone libraries using the primers 515F/1391R. The group "Others (< 1%)" is the collection of reads on class level that singlehanded constituted less than 1 % of the total amplicon reads. The group "Not assigned to class level" is a collection of reads that could not be classified by CREST to a deeper level than phylum. n = number of reads/sequences.

3.8 Phylogenetic analysis of *dsrB* sequences

The qPCR analysis showed high copy numbers of the *dsrB* gene from 3 m to 57 m, but no known sulphate-reducing microbial groups were found in the amplicon 454 pyrosequencing library. To investigate if the sulphate-reducing microbial community belonged to uncharacterised groups or known sulphate-reducing groups, the phylogenetic distribution of the sequenced *dsrB* clones from the 54 m sample was investigated. A rough phylogenetic tree on amino acid level was constructed using maximum likelihood algorithms (ML) via MEGA6 (Tamura et al., 2013). The ML tree was generated from the 39 sequenced *dsrB* clones, as well as on the most similar *dsrB* amino acid sequences found by BLASTp search and representatives from full-sequenced bacterial genomes (**Appendix D**). The *dsrB* phylogeny is highly complex and the overall topology obtained with the permafrost sequences of this study was therefore compared to a previously published tree by Kaneko et al. (2007), as proper tree constructions would require a lot of manual work and detailed insight into phylogenetics, a scope beyond this Master's thesis.

The phylogenetic analysis of the 39 sequenced *dsrB* clones from the 54 m sample revealed high diversity of the *dsrB* protein sequences, all clustering with sequences found in marine environments (**Appendix D**). 15 sequences (39 %) affiliated with dsr Group V (**Fig. 17**), which is a group with no cultured representatives (Kaneko et al., 2007). Six sequences (15 %)

affiliated with dsr Group III, which is closely related to the non-sulphate-reducing Gram-positive firmicute *Pelotomaculum* (Kaneko et al., 2007). Four sequences (10 %) clustered with dsr Group II (**Fig. 17**), closely related to the Deltaproteobacterial family Syntrophobacteraceae (Kaneko et al., 2007), and 12 sequences (31 %) clustered with dsr Group I, affiliating with the deltaproteobacterial family *Desulfobacteraceae* (Kaneko et al., 2007). Two sequences (5 %) affiliated with a group consisting of only uncultured representatives (**Appendix D**).

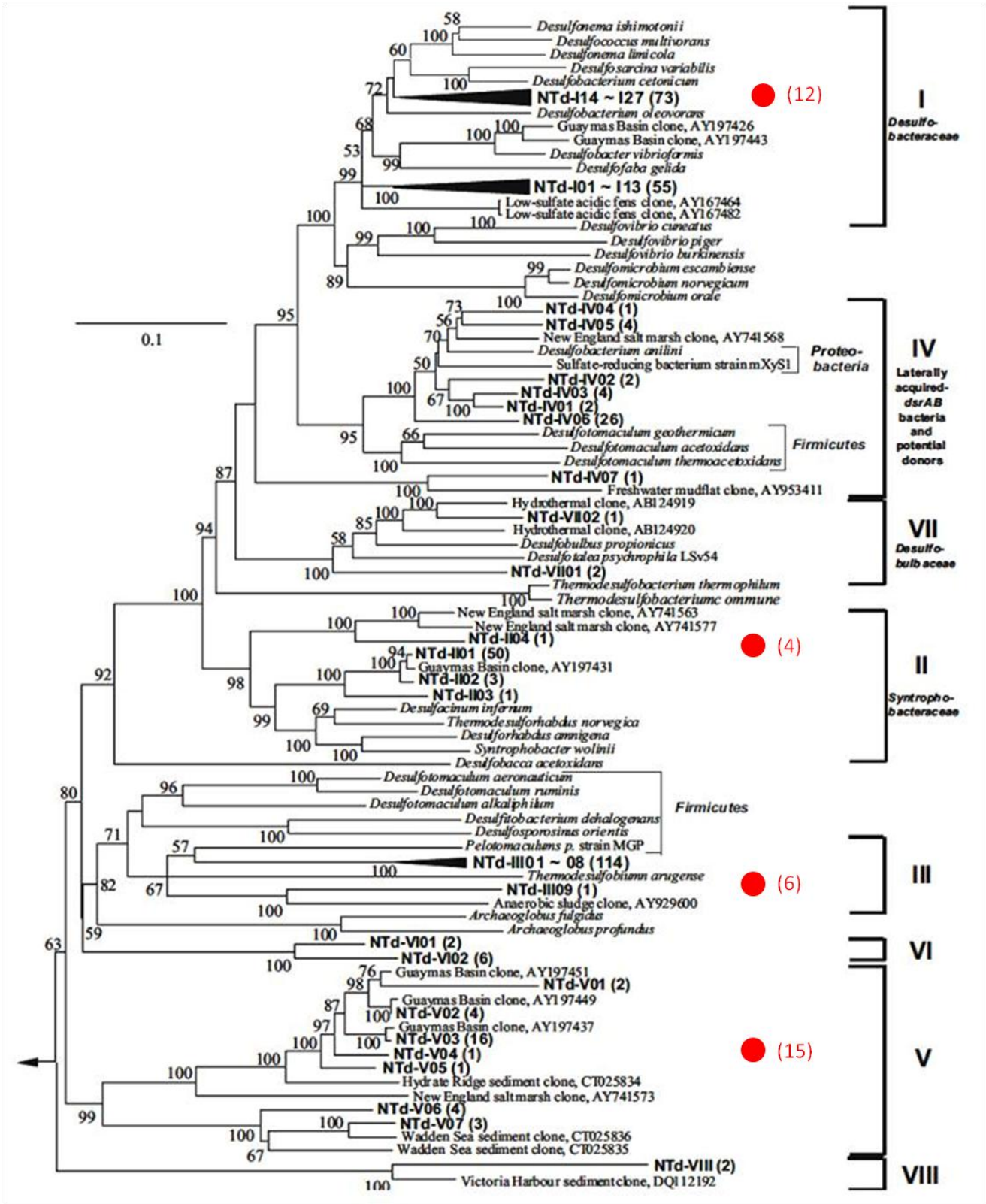

Figure 17: DsrAB phylogenetic tree (approximately 600 amino acids lenght) from the Nankai Trough deepsea sediment cores (Kaneko et al., 2007). The 39 amino acid sequences from the 54 m sample in the Adventdalen DH8 permafrost core were added to the tree based on the results of the preliminary phylogenetic tree in Appendix D, and are marked as red circles. The numbers in parentheses behind the red circles are the number of permafrost sequences within that group. Modified from Kaneko et al., 2007.

3.9 Phylogenetic analysis of DSAG sequences

To assess the phylogenetic distribution of the DSAG sequences, the DSAG reads obtained from 16S rRNA amplicon 454 pyrosequencing and the DSAG sequences obtained from the two 16S rRNA clone libraries (altogether 11093 reads) were added to a previously published tree (Jorgensen et al., 2013), using the parsimony tool in ARB (Ludwig et al., 2004). The tree is divided in three distinct monophyletic DSAG lineages (**Fig. 18**), termed Alpha, Beta and Gamma (Jorgensen et al., 2013).

Only 8 sequences affiliated with the beta cluster, and these sequences belonged to the 54 m sample. The rest of the 11085 sequences clustered within the Gamma lineage (**Fig. 18**). 98.2 % of the sequences affiliated within one single group of the Gamma cluster, together with sequences obtained from marine and freshwater sediments, and hypersaline environments (**Fig. 18**). The rest of the sequences (1.7 %) affiliated with DSAG sequences found in Guerro Negro hypersaline ponds (**Fig. 18**).

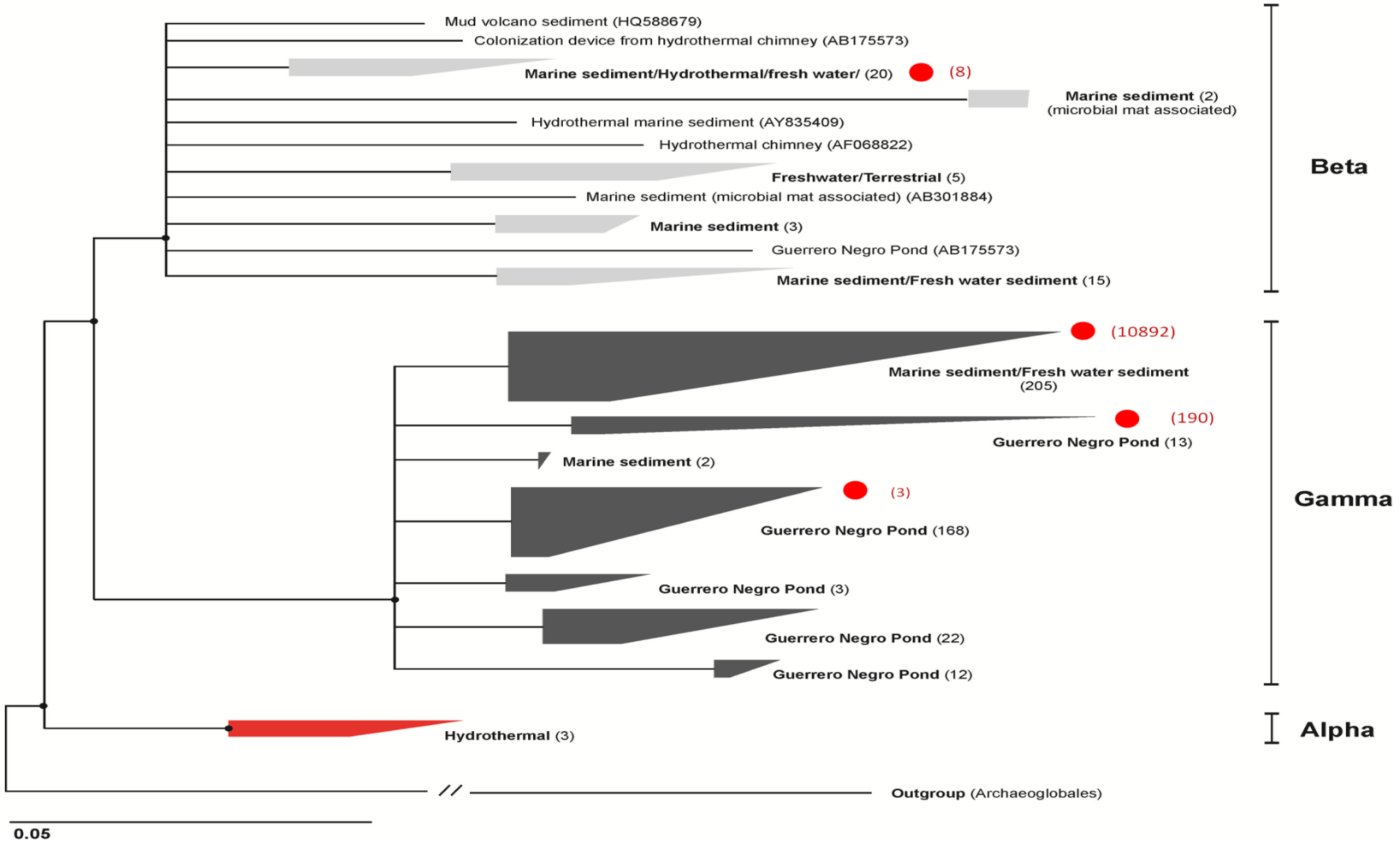

Figure 18: Phylogenetic tree of the Deep Sea Archaeal Group (DSAG) (Jorgensen et al., 2013). The DSAG sequences from this study were added to the tree using the parsimony tool in ARB, and are marked as red circles. The numbers in parentheses behind the red circles are the number of reads within that group.

3.10 Statistical analyses

Pearsons correlation analyses were performed to search for correlations between microbial data and geological and geochemical parameters. Initial analyses investigated if variations in pH could be explained by variations in mineral composition or dissolved ions. pH was significantly correlated ($p < 0.05$) with CaO ($R^2 = 0.935$) and MgO content ($R^2 = 0.849$) (**Fig. 19**). There was also a correlation between pH and K_2O ($R^2 = 0.718$), but not significant (**Appendix E**). None of the other minerals or dissolved ions revealed any significant correlations with pH.

Since CaO, MgO and K_2O are correlated, the linear regression model explaining pH based on mineral content would include CaO and Fe_2O_3 as variables with $p < 0.05$ for both the model end the independent variables (adjusted $R^2 = 0.98$) (**Appendix E**). Thus the pH variability can be largely explained by the amount of "base cations" in the minerals.

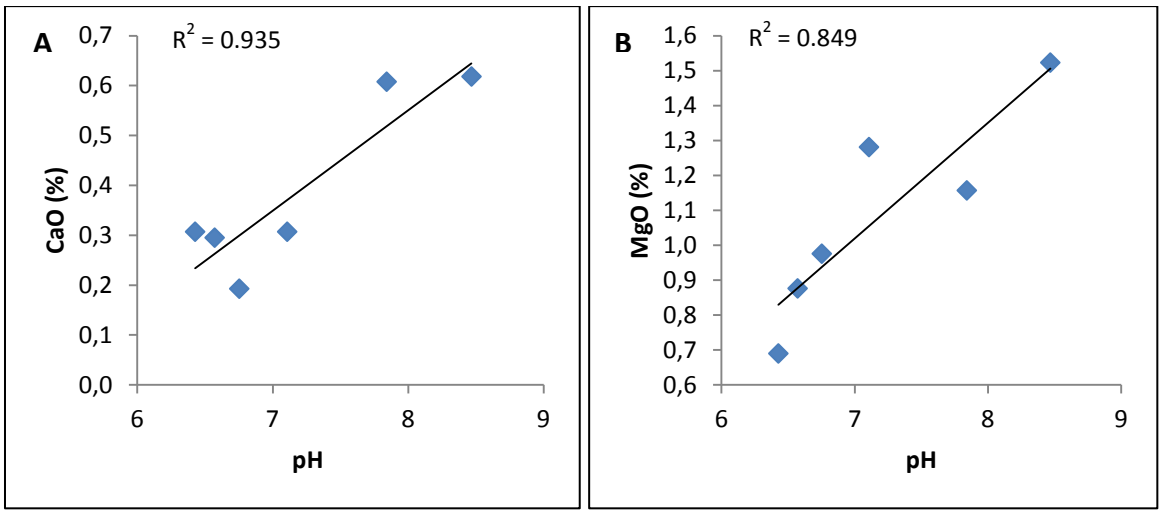

Figure 19: Pearson correlations between pH and (**A**) Calsium oxide and (**B**) Magnesium oxide, showing a significant ($p < 0.05$) correlation between these parameters.

Further analyses investigated if there were correlations between bacterial and archaeal 16S rRNA copy numbers and total organic carbon (TOC), dissolved nutrients and electron acceptors (K^+, Mg^{2+}, NH_4^+, PO_4^{2-}, NO_3^-, SO_4^{2-}, Fe^{2+} and Mn^{2+}). Bacterial 16S rRNA copy numbers were significantly ($p < 0.05$) correlated with TOC ($R^2 = 0.917$) and concentration of dissolved K^+ ($R^2 = -0.881$) (**Fig. 20**). Bacterial 16S rRNA copy numbers were negatively correlated with dissolved Mg^{2+}, NH_4^+, PO_4^{2-} and NO_3^-, but the correlations were not significant (**Appendix E**). Regarding potential electron acceptors, bacterial 16S rRNA were significantly ($p < 0.05$) correlated with the dissolved Fe^{2+} ($R^2 = 0.990$) and Mn^{2+} ($R^2 = 0.952$) concentrations (**Fig. 20**), but only when the concentrations at 60 m were removed. Archaeal 16S rRNA copy numbers showed no correlations with the parameters above (**Appendix E**).

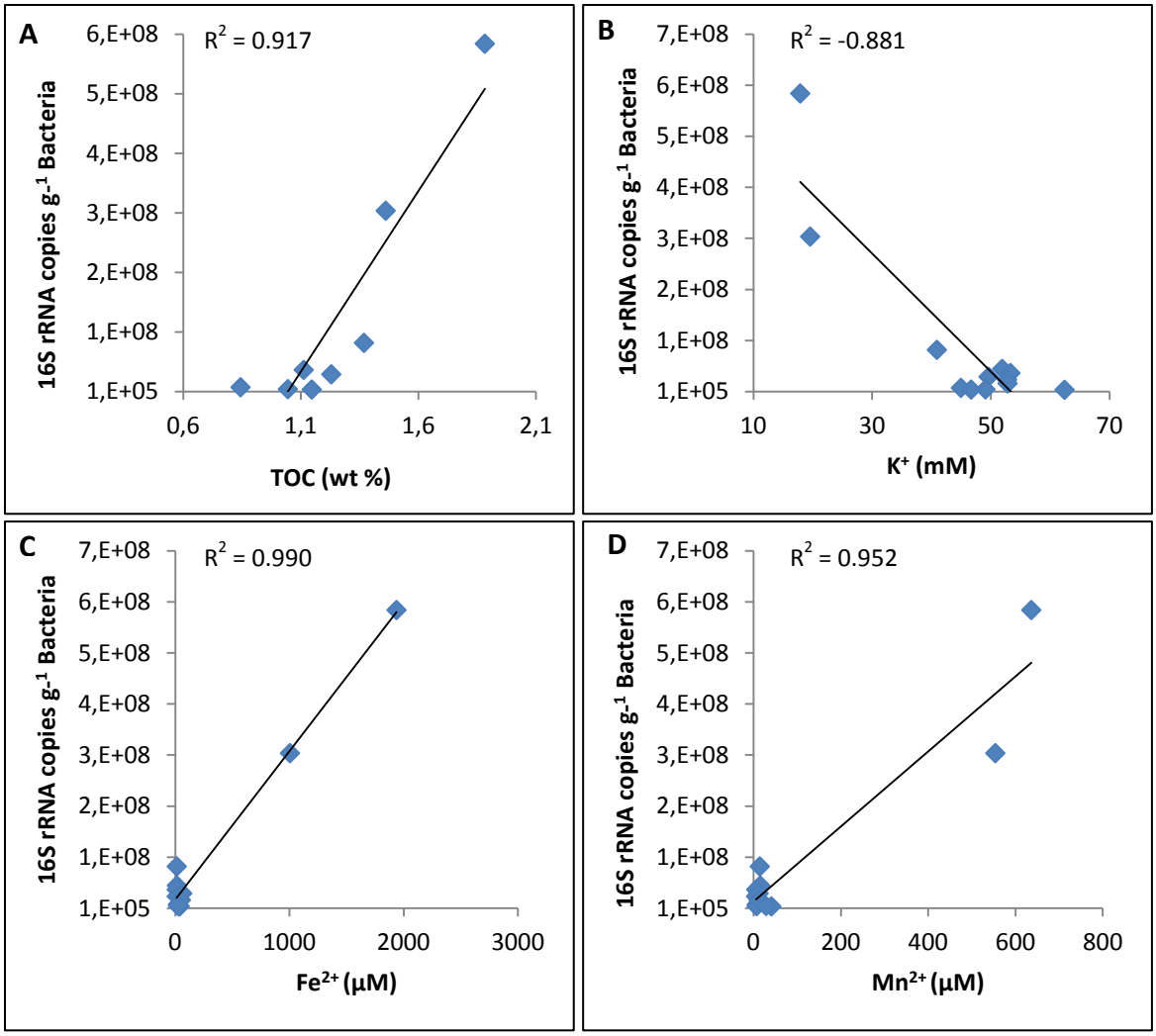

Figure 20: Pearson correlations between bacterial 16S rRNA copies g^{-1} and (**A**) Total Organic Carbon (TOC), (**B**) dissolved K^+, (**C**) dissolved Fe^{2+} and (**D**) dissolved Mn^{2+}showing significant ($p < 0.05$) correlations for these parameters. K^+, Fe^{2+} and Mn^{2+} data at 60 m have been removed.

4. Discussion of methods

4.1 Core retrieval

Several studies working with microbial diversity in permafrost have used fluid-less drilling techniques, combined with bacterial tracers or fluorescent microspheres, to trace potential contamination in the samples (Juck et al., 2005, Shi et al., 1997). The retrieval of the DH8 permafrost core was done with the use of local river water as drilling fluid and with careful, but not 100 % sterile handling in the field, as this was not possible because of technical reasons (see "Materials and Methods", Section 2.1). The exposure of the core material with the drilling fluid was limited by a plastic tube encasing the core during drilling. Furthermore, the outer centimetres of the core were removed by microbiologists before harvesting the inner parts of the core for microbial analysis, to ensure that the core material was free from drill fluid contamination (see "Materials and methods", Section 2.2). The microbial community composition in the potentially contaminated outer edge of the 3 m sample, showed no distinct differences from the inner core sample of the same depth (**Figs. 13 and 14**). Although no analyses of drill fluid microbial community composition is included in this study, a higher degree of one or more phyla in the outer edge of the core as compared to inner part would be expected, in case of contamination. Further, the microbial community composition showed variations with depth, and major phyla from one depth are not present at other depths. If there was significant contamination from the drilling fluid throughout the core, a more similar microbial community structure with depth would be expected. A dominance of anaerobic or fermentative microbial groups, and groups affiliating with marine environments were found in all samples (see "Discussion of results", Section 5.3), which are not expected to dominate the community composition in the river water used as drilling fluid. These findings suggest that there has been no significant drill fluid contamination in the seven depths of the DH8 permafrost core analysed in this study. It would have been interesting to investigate the microbial community composition in the river water used as drilling fluid, to compare its community composition with the microbial community of the sedimentary permafrost samples. This could give further indications if there was contamination from the river water in the permafrost samples. Unfortunately, samples of the river water were not available at the time of analysis. Adding known DNA or fluorescent molecules to the drilling fluid, and then specifically search for these within the core sediments, would have been an even more robust test to exclude drill fluid contamination. However, this was not possible at the time of drilling.

4.2 DNA extraction and inhibition of PCR reactions

All of the molecular analyses used in this study depend on the success of the DNA extraction. Some cells might have more robust cell walls and could be more resistant to the lysing method that was used (bead beating method using Lysing Matrix E (MP Biomedicals)). This could lead to underrepresentation of some microbial groups. However, it has been demonstrated that the FastDNA Spin Kit for Soil (MP Biomedicals) had the highest lysis

efficiency, when comparing the effectiveness of four commercial DNA extraction kits in permafrost samples (Vishnivetskaya et al., 2014). Although the lysis efficiency in these sediments cannot be tested within the framework of this project, the arguably most suitable protocol available has been used.

Another common problem of DNA extraction from sediment is the co-extraction of humic substances. Humic substances can inhibit downstream PCR reactions by interacting with the DNA strand, inhibit the DNA polymerase, or the polymerase cofactors. Inhibition by humic substances can make it problematic to obtain good quality PCR products for sequencing, and could cause underestimation of the actual gene copy numbers in qPCR (Albers et al., 2013). The DNA extracts from some of the samples used in this study had a yellow and brownish colour, presumably due to high concentrations of humic substances. There was no apparent inhibition observed in qPCR analysis, while inhibition was problematic in conventional PCR. Betaine was added to the PCR master mixture to improve the yield in PCR amplification. It is suggested that Betaine improves the amplification of genes by reducing the formation of secondary structures caused by GC-rich regions and, therefore, may be generally applicable to ameliorate the amplification of GC-rich DNA sequences (Henke et al., 1997). However, the addition of Betaine had no positive effect on PCR amplification in our samples.

Bovine Serum Albumin (BSA) is commonly used to enhance PCR product yield. BSA binds to inhibitory substances and thereby prevent their binding and inactivation of DNA polymerase (Kreader, 1996). Here, addition of BSA increased the yield in PCR amplification. Nevertheless, obtaining good quality PCR products for downstream 454 pyrosequencing remained difficult. The low read numbers obtained from 454 pyrosequencing (**Table 5**) are most likely due to inhibition in PCR amplification, and the resulting low concentrations of PCR products. Addition of BSA had no apparent effect on qPCR analysis, but this was only tested on *mcrA* and *pmoA* quantification.

When comparing the efficiency of different DNA extraction kits, it was demonstrated that genomic DNA obtained from the FastDNA Spin Kit for Soil (MP Biomedicals) needed additional purification before PCR amplifications, due to inhibitory substances in the DNA extracts (Vishnivetskaya et al., 2014). Purification of the DNA extracts could have increased the DNA quality in our samples, and thereby enhancing the PCR product yield.

4.3 Presence of dead cells or naked DNA in the permafrost samples

Detection of 16S rRNA sequences does not necessarily indicate that these sequences belong to viable and metabolically active organisms. Subzero temperatures in permafrost make this a well suited system for long term preservation of dead cells and free nucleic acids (Willerslev et al., 2004). This has implications for the interpretation of the microbial abundance and community structure, as the DNA analysed could originate from dormant, inactive or dead cells where the DNA have been preserved. The presence of dead cells or free DNA may lead to false conclusions when describing the abundance and function of the microbial communities, as a high portion of the diversity could be non-functional cells

without ecological significance. For example, Hansen et al. (2007) demonstrated that 74 % of the total microscopic cell count in permafrost samples from Adventdalen had compromised cell walls, and were considered non-viable. However, studies have shown that slow, but continuous metabolic activity occur in permafrost (see "Introduction", Section 1.2), and that there is evidence of DNA repair mechanisms in ancient microbial cells living in permafrost (Johnson et al., 2007). Research by Hovgaard (2014) on the Adventdalen DH8 permafrost core indicated that slow microbial activity occurs throughout the core, by estimating turnover times for microbial biomass using a D:L amino acid model. It has not been an objective of this project to investigate microbial activity, but it would be an interesting follow-up for future studies.

4.4 PCR bias in microbial quantification and community structure analysis

At 16.5 m DSAG 16S rRNA made up 188 % of total archaeal 16S rRNA copies (**Fig. 7**). Ideally, the DSAG 16S rRNA copies should constitute 100 % or less of the total archaeal 16S rRNA copies. Differing DSAG and archaeal 16S rRNA copy numbers illustrates the uncertainty with this method. It is recommended to be cautious when comparing absolute gene copy numbers from different qPCR runs, due to variations in standard curves, amplification efficiencies and variations between researchers that can result in different gene copy numbers for the same environmental sample (Smith et al., 2006). The archaeal and DSAG 16S rRNA copy numbers are, however, within the same order of magnitude, which is the level that gene copy numbers can be compared (Laila Reigstad, personal communication).

There are several factors regarding the use of primers in conventional PCR and qPCR that may bias the results. It has been demonstrated that taxonomic assignment is very sensitive to the region of the 16S rRNA gene that is targeted (Liu et al., 2008). Moreover, primer-template mismatch is a recognised problem that can affect the PCR efficiency. It is not possible to create a perfectly matching universal primer, due to single nucleotide variations even in highly conserved regions of the 16S rRNA genes (Schmalenberger et al., 2001, Sipos et al., 2007). Primer mismatches may lead to bias in qPCR, resulting in an underestimation of the actual gene copy number. In analysis of microbial community composition, primer-template mismatches can cause underestimation of taxonomic groups that do not 100 % match the primer sequence (Bru et al., 2008, Wu et al., 2009). 2010). Use of primers with degenerate positions can also bias results. Polz and Cavanaugh (1998) demonstrated that differences in GC content at degenerate positions in the primer target site of the 16S rRNA caused PCR bias. Their results showed over-amplification of target template containing G/C in the degenerate position, compared to templates containing A/T. This was due to different binding energies of the degenerate primers, where there was higher primer affinity to G/C regions. The unequal amplification efficiency increased with cycle number (Polz and Cavanaugh, 1998).

The results from 16S rRNA amplicon 454 pyrosequencing showed a bacterial dominated community at 3 m and 4.5 m, but from 16.5 m to 57 m Archaea constituted between 83 %

and 99.8 % of the relative abundance of reads (**Table 5**). In contrast to this relationship, the qPCR data showed bacterial dominance down to 31.5 m (90-99 % of total 16S rRNA genes belonged to Bacteria), and more similar bacterial and archaeal 16S rRNA copy numbers from 54 to 60 m (46-61 % of total 16S rRNA genes belonged to Bacteria) (**Fig. 7**). Reasons behind the observed differences in the relationship between Bacteria and Archaea could be that the bacterial abundance was overestimated in qPCR, or that the primers used in 454 pyrosequencing underestimated the bacterial population. To assess this, two different 16S rRNA clone libraries were constructed from the sample at 54 m, using the universal prokaryotic primers 787F/1391R and 515F/1391R. The two different 16S rRNA clone libraries revealed markedly different microbial communities, with a higher bacterial diversity in the 515F/1391R clone library (**Figs. 15 and 16**). These results strongly indicate that the 787F/1391R primers used in 454 pyrosequencing have underestimated the bacterial diversity in the samples. This was somewhat surprising as the 787F/1391R primers have been theoretically evaluated to be the optimal primer combination for amplicon 454 pyrosequencing (Jorgensen et al., 2012), considering the limitations regarding amplicon length in the 454 GS FLX technology. The 787F/1391R primer combination should theoretically cover 87-94 % of all known prokaryotes without mismatch. With one mismatch, the coverage of both primers increased to 98 % (Jorgensen et al., 2012). However, when evaluating the primer coverage *in silico* (computationally), the computer does not take into consideration that the primers might have preferential binding to some templates, even though the primers are matching all template sequences. It is therefore possible that the underestimation of bacterial taxa observed in the 454 pyrosequencing results could be due to primer degeneracy effects, where the primers had higher affinity for target templates with G/C in the degenerate primer position, leading to underrepresentation of groups with A/T in the same position (Polz and Cavanaugh, 1998).

Lanzén and co-workers (2011) analysed the microbial community composition at a hydrothermal vent site using amplicon 16S rRNA 454 pyrosequencing with the 787F/1391R primers, combined with shotgun sequencing (Lanzen et al., 2011). They found that some taxa were consistently underrepresented in datasets obtained with 16S rRNA 454 pyrosequencing. These underrepresented taxa contained an A or T in the degenerate position of the 787F primer. This observation led to the suggestion that the 787F primer has preferential binding to templates with G/C in the degenerate position (Lanzen et al., 2011).

Primer 787F (5´-ATTAGATACCCNGGTAG-3´) contains a single degeneracy at position 12, where all 4 bases can be present. To investigate what nucleotides were present in the degenerate position of primer 787F in the clone library sequences, a search after the 787F primer combination was performed within the sequences from the two different 16S rRNA clone libraries. In the 515F/1391R clone library the 787F primer combination was present in 36 of the 40 sequences. The 787F primer should thus theoretically be able to cover these 36 sequences. Of these 36 matching sequences, only 4 sequences (11 %) had a G/C in the

degenerate position of primer 787F, while the rest of the sequences had an A/T. Of the four sequences with a G/C in the degenerate primer position, 3 belonged to DSAG, and one belonged to Chloroflexi Subdivision 6.

A search after the 787F primer sequence within the sequences from the 787F/1391R clone library revealed that 36 of the 39 sequences (93 %) contained a G/C in the degenerate primer position, and only 3 sequences (7 %) had an A/T in the degenerate position. Additionally, all of the 497 DSAG sequences in the phylogenetic three of DSAG (**Fig. 18**) contain a G/C in that position, which could explain the dominance of DSAG in the amplicon 454 pyrosequencing library. These observations strongly suggest that the observed underestimation of bacterial taxa in the 454 pyrosequencing data could be due to primer degeneracy effects, where the 787F primer has preferential binding to templates with G/C in the degenerate primer position. This is in agreement with findings by Polz and Cavanaugh (1998) and Lanzén et al., (2011). It is possible to check if the reads obtained from 454 pyrosequencing were dominated by reads with G/C in the target site of the degenerate primer position. Unfortunately, that is a difficult and time-consuming task, which was not possible to investigate within the time-frame of this Master's thesis.

It is recommended to use small number of cycles in PCR amplification to diminish the effects of preferential primer binding to G/C regions (Polz and Cavanaugh, 1998) and to avoid chimera formation. For 16S rRNA amplicon 454 pyrosequencing it is recommended to use 20-25 cycles in PCR when preparing the samples for sequencing. Here, 30 cycles were used, due to problems with high inhibition in the DNA extracts, making it difficult to obtain strong enough PCR products for downstream 16S rRNA amplicon 454 pyrosequencing. Additionally, the PCR products were used as templates for a second PCR reaction, to obtain higher concentration of the PCR products (see "Materials and Methods", Section 2.4.4). Degeneracy in the forward primer and the high number of cycles used in PCR may thus have caused over-amplification of microbial taxa possessing a G or C in the target site of the degenerate primer position.

Clearly, cautions need to be taken when interpreting the 16S rRNA amplicon 454 pyrosequencing results, as the bacterial diversity was underestimated. However, the detection of the different taxa confirms their presence, but we need to be careful when looking at the relative abundance distribution. A different strategy could have been to use bacterial and archaeal specific primers and then sequence separately, for example by using the same primers that were used for bacterial and archaeal quantification in qPCR. This could have given sequencing results that were more comparable with the qPCR results. However, with this approach it is not possible to determine the relative abundance of Bacteria and Archaea in the sequencing results, and the time and cost for sample preparation and sequencing doubles. Although the DSAG could have been over-amplified due to the presence of G/C in the target site of the degenerate primer position, both the

DSAG specific qPCR (**Fig. 7**) and the archaeal community composition from the 515F/1391R clone library (**Fig. 16**) confirmed the dominance of DSAG within Archaea.

4.5 Geological and geochemical methods

Pore water cations were extracted by two sequential extractions. First, H_2O was added to the sediment to extract the water-soluble ions. Then, 0.02 M $CaCl_2$ was added to the same sediment that had first been treated with H_2O, to extract cations that are electrostatically adsorbed to negatively charged sediment particles. The added Ca^{2+} ions will replace the adsorbed cations on the sediment particles, increasing their concentration in the solution (Thomas, 1977). This procedure releases a larger amount of the bio-available cations. The fact that more cations were extracted with the $CaCl_2$ solution demonstrated that there are more cations available for the microorganisms than the cations that are dissolved in the pore water, especially divalent cations (**Appendix B**).

An uncertainty of the extraction method is that the ion concentrations were measured in extracted pore water, and then scaled up to represent the concentration in the actual pore water volume. The pore water volume was found by subtracting the dry weight of the sediment from the wet weight of the sediment. Small errors in weighing the wet and dry sediment would give large effects when estimating the ion concentrations in the pore water volume, as the pore waters are a relatively small part of the water involved during extraction. Most of the ion concentrations were a lot higher in the 60 m shale sample (**Fig. 6 and Table 4**). This sample was first crushed in a mortar before undergoing the same extraction procedure as the sedimentary samples, resulting in fresh mineral surfaces and dust, which will provide easily extractable ions. Furthermore, the ion concentrations were calculated to represent the concentration in an approximate pore water volume. As the shale has hardly any pore volume, the ion concentrations become artificially high.

Regarding the Total Organic Carbon (TOC) content, the largest uncertainty is that TOC was not measured directly; it was measured by subtracting Total Inorganic Carbon (TIC) from Total Carbon (TC). Small errors in TIC and TC measurements could result in large errors in the TOC estimation.

5. Discussion of results

5.1 Geological and geochemical parameters of the DH8 permafrost core

Sedimentary statigraphy analysis of the DH8 permafrost core (Gilbert, 2014) revealed that the core was of marine origin with bedrock (shale) at the bottom 60 m depth, overlain by till and marine deposits (**Fig. 3**). A thick package of deltaic sediments was subsequently deposited, followed by aeolian deposits in the top few meters of the core (Gilbert, 2014). Optically Stimulated Luminescence (OSL) dating of the core (Christine Thiel, In: Gilbert, 2014) indicated depositional age of the core from 8.7 thousand years ago (kya) at 49.65 m depth to 3.0 kya at 3 m. The shale from 60 m was deposited during early Cretaceous (Parker, 1967). Permafrost aggradation at this site began approximately at 3 kya (Gilbert, 2014).

The geomorphological description of the DH8 core characterise the 3 m sample to the delta top deposits (**Fig. 3**). The boundary between the aeolian deposits and delta top deposits was estimated to be at 2.9 m (Gilbert, 2014). The bacterial and archaeal communities at 3 m were different from the rest of the core samples (**Figs. 9-12**), which could indicate that the 3 m sample represents the aeolian depositional environment. It is also possible that the 3 m sample was in the surface when the aeolian sediments were deposited, and are thus more influenced by surface processes. The predominance of poorly described microbial groups that are associated with marine environments from 16.5 m to 57 m reflects the marine origin of the core (**Figs. 9-12**) (see Section 5.3 for a full discussion of the microbial groups). Furthermore, the Na^+ and Cl^- concentrations from extracted pore water corresponded with marine geochemical environment (Na^+: 481 mM and Cl^-: 559 mM, Pilson, 2013) downwards from 16.5 m (**Fig. 6**). Concentrations of $Na^+$and Cl^- (**Fig. 6**) were lower at 3 m and 4.5 m, suggesting influence of meteoric water at these depths. The concentrations of SO_4^{2-} were lower than what is expected from a marine environment (28.93 mM, Pilson, 2013) in all samples, except at 16.5 m (**Fig. 6**). These low SO_4^{2-} concentrations could be caused by microbial sulphate reduction, rather than influence of meteoric water below 4.5 m, since a corresponding reduction in the concentration of Cl^- was not found. This is strengthened by high copy numbers of the *drsB* gene from 3 m to 57 m (**Fig. 7**), suggesting that microbial sulphate reduction is, or have been, a dominant metabolism at these depths.

The grain size distribution analysis showed that the samples from 3 m to 16.5 m had a coarser texture (47-88 % sand) than the samples from 31.5 m to 57 m (82-98 % silt) (**Table 2**). The coarser texture in the three top samples is probably related to the deltaic depositional material at these depths (**Fig. 3**), as the coarser materials are being deposited at the delta top and delta front. At 57 m the sediments contained more sand than the sample from 54 m (**Table 2**), which agrees with the classification of the 57 m sample as melt-out till (**Fig. 3**). Relatively low MgO and Al_2O_3 content through the core (**Table 3**) indicated low content of clay minerals (Professor Ingunn Hindenes Thorseth, personal communication), which was confirmed in the grain size distribution analysis (**Table 2**).

The total organic carbon (TOC) content (**Table 2**) was relatively low from 4.5 m to 57 m (1.0-1.4 %), but higher at 3 m (1.8 %), possibly due to influence from the active layer or previous freezing and thawing cycles enabling migration of carbon from above. TOC was within the range described for Adventdalen permafrost (Hansen et al., 2007), Siberian permafrost (Vishnivetskaya et al., 2006), but higher than in Antarctic permafrost (Gilichinsky et al., 2007). The pH was slightly acidic (pH 6.4 to 6.8) in the top three sedimentary permafrost samples (**Table 4**). Presence and degradation of organic matter can lower the pH by formation of organic acids (Paul and Clark, 2007). However, no correlation between current TOC content and pH was found (**Appendix E**). Microbial respiration can also lower the pH in the pore water by releasing CO_2 that will form carbonic acid in water. The pH could be further reduced by microbial nutrient uptake, because microorganisms exchange H^+ with base cations. The lower pH at 3 m and 16.5 m (pH 6.4 and 6.6) (**Table 4**) could be attributed to the higher TOC content at these depths (1.8 and 1.4 %) (**Table 2**). There was an increase in pH from 31.5 m to 57 m (**Table 4**), which could be connected to the marine origin of the core, as seawater normally ranges between pH 7.5 and 8.5 (Pilson, 2013). The high pH measured in the crushed shale sample from 60 m could be related to the different depositional material of this sample, resulting in different water-mineral interactions than in the other sedimentary samples. The pH was significantly correlated with the CaO and MgO content of the sediment minerals (**Fig. 19**), but not with dissolved Ca^{2+} and Mg^{2+} ions. This could indicate that there is a connection between mineral composition and pH, where the minerals may release base cations to the pore water, causing increased pH.

Major shifts in mineral composition (**Table 3**) or pore water chemistry (**Fig. 6 and Table 4**) that could explain the shifts in microbial abundance or community composition were not found. However, the decrease in microbial abundance from 3 m to 4.5 m (**Fig. 7**) follows a change to coarser grain size and reduced organic carbon content (**Table 2**). One main geological parameter that significantly correlated (R^2 = 0.917) with microbial abundance was organic carbon content (**Fig. 20**), indicating that organic carbon is, or have been a major limitation for microbial growth in the DH8 permafrost core. Further, the bacterial abundance was significantly negatively correlated (R^2 = -0.881) with the concentration of dissolved K^+, and negatively correlated with other nutrients such as Mg^{2+}, NH_4^+, PO_4^{2-} and NO_3^- (**Appendix E)**, although the latter correlations were not significant ($p>0.05$). This could indicate that the bacterial population use these nutrients, causing subsequent low concentrations.

The bacterial abundance was significantly correlated with dissolved Fe^{2+} and Mn^{2+} concentrations, when 60 m data were removed (**Fig. 20**). However, when the 3 m data were removed in the statistical analysis, the correlations were no longer significant. Fe^{2+} and Mn^{2+} had very high concentrations at 3 m (**Fig. 6**), which is most likely due to geochemical processes rather than microbial reduction. Migration of organic acids with percolating water could increase solubility of iron and manganese oxides as Fe^{2+} and Mn^{2+} by reducing the pH (and potentially oxygen concentrations). It is therefore likely that organic carbon content is

more important for bacterial abundance, than the availability of iron and manganese as electron acceptors.

The archaeal abundance showed no correlations with the dissolved nutrients or electron acceptors (**Appendix E**). To improve the strength of the statistical test, one could have analysed more than seven depths, or analysed more samples from the same depths. Other aspects of the geochemistry will be discussed with the microbial data below.

5.2 Microbial abundance

5.2.1 16S rRNA copy numbers

The overall trend in bacterial and archaeal 16S rRNA copy numbers from qPCR analysis showed high copy numbers at 3 m, a decease at 4.5 m and a decrease from 54 m to 60 m (**Fig. 7**). High numbers in the uppermost sample at 3 m could be linked to the active layer above (0 - 1 m) (**Fig. 2**), where the sediments at this depth could have been influenced by migration of surface water and organic carbon. Additionally, the active layer of today has most probably reached down to below 3 m under warmer climate conditions (Hanne Christiansen, personal communication). The low microbial abundance at 4.5 m could be linked to the high sand content and the low organic carbon content at this depth (**Table 2**). Coarse grained sediments typically have less unfrozen water than fine grained sediments, because the larger particles have a smaller surface that interacts with water molecules (Gilichinsky, 2002). Unfrozen water films are important for microbial survival and activity because they protect against freezing-thawing stress, contains dissolved metabolites and nutrients and facilitate transfer of metabolic waste products (Gilichinsky, 2002). It is therefore likely to assume that the coarse sediment at 4.5 m contains less unfrozen water and therefore yield lower microbial numbers. The decrease in microbial abundance from 54 m to 60 m could be due to the higher age of the sediments at these depths (around 8700 years at 54 and 57 m, and 100 million years or more in the 60 m shale sample) (Gilbert, 2014, Parker, 1967) and the low organic carbon content (**Table 2**).

The DNA extraction from the shale sample (60 m) was successful, and qPCR analysis revealed relatively high bacterial and archaeal 16S rRNA copy numbers per gram shale, with 2.4 x 10^5 copies g^{-1} for Bacteria and 2.5 x 10^5 copies g^{-1} for Archaea (**Fig. 7**). Meslé and co-workers reported low genomic DNA yield extracted from shales, but managed to obtain enrichment cultures of methanogenic archaea (Mesle et al., 2013). Others have found microbial cell numbers in shale by direct microscopic counts, ranging from 5.2 x 10^3 to 2.4 x 10^4 cells g^{-1} (Zhang et al., 2005) and 2 x 10^6 cells g^{-1} (Kovacik et al., 2006).

The archaeal abundance obtained by qPCR quantification was high below 4.5 m (6-54 % of total 16S rRNA genes), and increased with depth (**Fig. 7**). This high archaeal abundance is in contrast to other studies working with active layer or shallow permafrost, where Archaea constituted a minor part of the total prokaryotic communities. In Canadian High Arctic permafrost the archaeal abundance was from 0.01 % (Yergeau et al., 2010) to 0.1 %

(Wilhelm et al., 2011) of total 16S rRNA copies. In the active layer of Svalbard peat-soils the relative abundance of Archaea was <1 % of the total microbial community (Tveit et al., 2013). In Siberian tundra the archaeal 16S rRNA genes accounted for <2 % of total 16S rRNA copies (Gittel et al., 2014). These relative abundances correspond to the archaeal abundance found in the uppermost permafrost sample (3 m) in the DH8 core, where archaeal 16S rRNA copies constituted 0.6 % of the total prokaryotic 16S rRNA copy numbers (**Fig. 7**). The mentioned studies above are all from shallow parts of the permafrost (1-2 m) or from the active layer.

5.2.2 Estimated cell numbers from qPCR analysis

For Bacteria, the number of 16S rRNA copies per genome varies between 1 and 15 copies, with an average of 4.2 16S rRNA genes per bacterial genome (Vetrovsky and Baldrian, 2013). In this project, the bacterial abundance found by qPCR 16S rRNA quantification was calculated to approximate cell numbers by dividing the bacterial 16S rRNA copy numbers with the average number of 16S rRNA genes per bacterial genome (**Fig. 8**). The average number (4.2 copies per genome) was used, because the taxonomic classification showed a dominance of poorly characterised microbial groups where the actual 16S rRNA gene copy number per genome is unknown.

Total prokaryotic cell numbers in the Adventdalen DH8 permafrost core was high in the top sample at 3 m, low at 4.5 m and there was a decreasing trend in cell numbers from 54 m to 60 m. The cell numbers varied between 3 x 10^5 and 1 x 10^8 (**Fig. 8**). Direct microscopic cell counts in Siberian permafrost revealed decreased cell numbers with depth (Rivkina et al., 1998, Vishnivetskaya et al., 2006), agreeing well with our estimated cell numbers from the qPCR analysis (**Fig. 21**). Cell numbers in a 9 m permafrost sample from Canada (Steven et al., 2007) was comparable to DH8 cell numbers at similar depths (**Fig. 21**). Hansen et al. (2007) counted the bacterial numbers in the shallow permafrost (1.4 - 1.6 m) of Adventdalen with epifluorescence microscopy. They found total bacterial numbers of 1.7 x 10^9 cells per gram dry weight (gdw^{-1}) (**Fig. 21**). This is one order of magnitude higher than the highest cell number found in the DH8 core. However, Hansen et al. operated with cells gdw^{-1}, resulting in cell numbers that are not directly comparable. Interestingly, they found that only 26 % of the total bacterial cells could be considered viable (Hansen et al., 2007).

Yergeau and colleagues measured bacterial and archaeal abundance in the active layer and at 2 m permafrost depth of Canadian High Arctic using qPCR (Yergeau et al., 2010). The ribosomal gene numbers found in the 2 m permafrost sample corresponded with DH8 cell numbers at 4.5 m (**Fig. 21**), while the microbial abundance was higher in the Canadian High Arctic active layer sample (Yergeau et al., 2010). Another study from a different site in the Canadian High Arctic (Wilhelm et al., 2011) found ribosomal gene numbers in a 1 m permafrost sample that were similar to DH8 cell numbers in top sample at 3 m (**Fig. 21**). Ribosomal gene numbers in the active layer and shallow parts of the permafrost in Siberia (Gittel et al., 2014), corresponded to our 4.5 m sample (**Fig. 21**).

To our knowledge, this project on the Adventdalen DH8 core is the first study of a permafrost core deeper than 2 m, where qPCR has been used to quantify bacterial and archaeal abundance. The estimated cell numbers in the DH8 core are comparable to cell numbers found in other deep permafrost cores (Rivkina et al., 1998, Steven et al., 2007, Vishnivetskaya et al., 2006), with exception of our sample from 4.5 m (**Fig. 21**). It is important to remember that there are methodological differences when comparing these cell numbers, where the different methods have various limitations. Also, the permafrost study sites may differ in environmental conditions and physical characteristics, like for example origin of sediments/soils, permafrost age, temperature, water content, ice content, TOC content and grain size. These are all factors that are likely to influence microbial cell numbers.

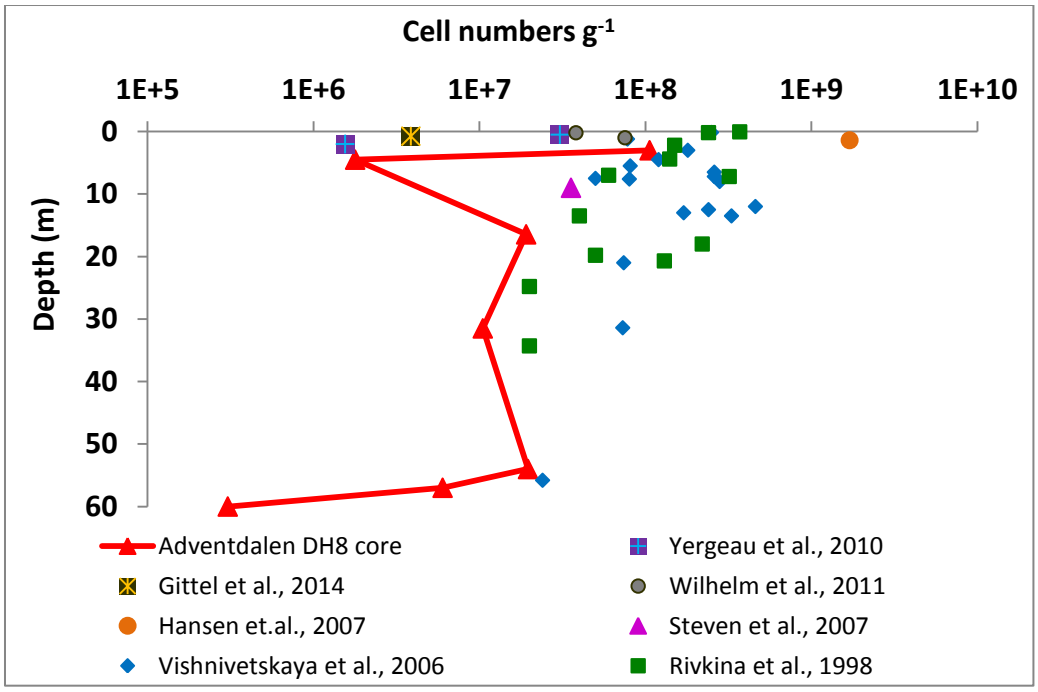

Figure 21: Comparison of cell numbers g^{-1} in a variety of permafrost locations. Yergeau et al., 2010, Gittel et al., 2014 and Wilhelm et al., 2011 have analysed the microbial abundance with qPCR, and the results are plotted as ribosomal gene copies per gram (gene copies per gram dry weight for Wilhelm et al., 2011). The rest of the studies have analysed the microbial cell numbers with direct microscopic cell counts. Cell number from the Hansen et al., 2007 study is given in cells per gram dry weight. Lines could not be drawn between the different samples from the Rivikina et al. 1998 and Vishnivetskaya et al. 2006 studies as they are from different permafrost cores, but published together.

5.3 Microbial community composition

The microbial community composition in the DH8 permafrost core was analysed by 16S rRNA amplicon 454 pyrosequencing using universal prokaryotic primers. The microbial communities at all depths were dominated by fermentative or anaerobic microbial groups (**Figs. 9-12**). Candidate division TM7 and OP9 are believed to be fermentative bacterial groups (Albertsen et al., 2013, Dodsworth et al., 2013, Kantor et al., 2013). Members of the

class Anaerolineae are recognised anaerobic heterotrophs (Podosokorskaya et al., 2013, Yamada et al., 2006). The methanogenic archaeal classes Methanomicrobia and Methanobacteria have also an anaerobic lifestyle (Bonin and Boone, 2006, Garcia et al., 2006). The Miscellaneous Crenarchaeotic Group (MCG), Deep Sea Archaeal Group (DSAG) and Marine Bentic Group D (Thermoplasmata) have been suggested to be anaerobic heterotrophs (Jorgensen et al., 2013, Kubo et al., 2012, Lloyd et al., 2013). Microbial sulphate reduction, which is an anaerobic metabolism, appears to be important from 3 m to 57 m, as this gene was found in high abundance at these depths (**Fig. 7**). Thus, the microbial community structure at all depths indicates the presence of anoxic zones throughout the core. The high Fe^{2+} and Mn^{2+} concentrations at 3 m together with high NH_4^+ and low NO_3^- concentrations throughout the core, provides further indications of an anoxic environment (**Fig. 6**). The presence of mainly heterotrophic organisms together with a strong correlation (R^2 = 0.917) between bacterial abundance and organic carbon content (**Fig. 20**), suggests that anaerobic degradation of organic carbon is the main energy source for the microbial community.

5.3.1 Bacterial community composition

The bacterial community at 3 m was very different compared to the other depths, dominated by Candidate division TM7 (67 % of bacterial community) (**Fig. 9**). This group was not found in the other depths investigated. Candidate division TM7 is a widely distributed bacterial division, but with no cultured representatives. Sequencing of the complete genome of single representatives from Candidate division TM7 predicted a fermentation-based metabolism for this group (Albertsen et al., 2013, Kantor et al., 2013). Members of this Candidate division have been detected in terrestrial, aquatic and clinical environments (Hugenholtz et al., 2001), as well as in the shallow permafrost of Adventdalen (Hansen et al., 2007).

The bacterial communities from 4.5 m to 57 m were dominated by mainly Chloroflexi, but Actinobacteria, Candidate division OP9 and Proteobacteria were also present (**Fig. 9**). Actinobacteria have shown to be one of the dominating phyla in previous permafrost studies (Hansen et al., 2007, Steven et al., 2007, Steven et al., 2008, Vishnivetskaya et al., 2006, Yergeau et al., 2010). This dominance is thought to be caused by their potential of being metabolically active at low temperatures and the presence of DNA repair mechanisms that may help reduce DNA damage under low temperatures and high salt concentrations (Johnson et al., 2007). Members of the phylum Proteobacteria have also been found to be a major component of permafrost microbial communities in the literature (Blanco et al., 2012, Gilichinsky et al., 2007, Steven et al., 2007, Steven et al., 2008, Tas et al., 2014, Wilhelm et al., 2011, Yergeau et al., 2010). In many of these studies, the Betaproteobacteria showed a higher abundance than any other Proteobacteria divisions (Blanco et al., 2012, Steven et al., 2008, Wilhelm et al., 2011, Yergeau et al., 2010). Here, the majority of Proteobacteria reads affiliated with Gammaproteobacteria and Alphaproteobacteria. Betaproteobacteria was only found in the shale sample (60 m) at 8 % relative abundance (**Fig. 10**). From 3 m to 31.5 m,

Chloroflexi Subdivision 11 was the dominating bacterial class (**Fig. 10**). Uncharacterised Chloroflexi representatives have been found in several permafrost studies (Mackelprang et al., 2011, Tas et al., 2014, Wilhelm et al., 2011), but their ecological role remains unknown.

The bacterial community composition was different in the 60 m shale sample, compared to the other depths investigated (**Figs. 9 and 10**). This depth was dominated by Thermodesulfobacteria, which are thermophilic sulphate-reducing bacteria (Garrity et al., 2001). High sulphate-reducing activity was measured at a sandstone-shale interface in a rock formation from Cerro Negro, New Mexico, while there was little sulphate-reducing activity in the actual shales, presumably due to their restrictive pore size (Krumholz et al., 1997). Investigation of microbial communities in organic-rich black shales resulted in the isolation of Gammaproteobacteria, Firmicutes and Actinobacteria, which were able to use the organic compounds in the shale as the sole source of carbon and energy (Matlakowska and Sklodowska, 2009). Actinobacteria was also present in the shale sample of the DH8 permafrost core (**Fig. 9**). Others have obtained enrichment cultures of methanogenic archaea, Proteobacteria, Actinobacteria, Bacteriodetes and Firmicutes from organic-rich shales (Mesle et al., 2013). Whether the bacterial community detected in the 60 m shale sample is an active community or just a preserved DNA signature remains unknown, but in light of the mentioned studies above it is reasonable to assume that there might be viable microorganisms in the shale.

5.3.2 Archaeal community composition

To our knowledge, the description of archaeal diversity in permafrost is limited to a few studies, and includes the detection of methanogenic archaea (Gittel et al., 2014, Mackelprang et al., 2011, Yergeau et al., 2010), ammonia oxidizing archaea (Thaumarchaeota) (Gittel et al., 2014, Wilhelm et al., 2011, Yergeau et al., 2010), the Soil Crenarchaeotic Group (Group I.1b, Thaumarchaeota) (Steven et al., 2008, Wilhelm et al., 2011), halophilic archaea (Euryarchaeota) (Steven et al., 2007, Steven et al., 2008) and sequences affiliating with Euryarchaeota (Tas et al., 2014). The DH8 permafrost microbial community revealed high abundance of archaeal taxa that have not been reported from other permafrost habitats, such as the DSAG, MCG and Thermoplasmata (**Fig. 12**). These archaeal groups are commonly found in marine sediments (Teske and Sorensen, 2008). The detection of archaeal taxa that have not been found in other permafrost studies could be related to the marine origin of the DH8 permafrost core. Furthermore, most research on microbial diversity in permafrost have been performed using clone libraries, which provide a more limited sampling of the microbial communities compared to next-generation sequencing methods, such as 454 pyrosequencing that was used in this study.

As for Bacteria, the archaeal community composition at 3 m was different from the communities found below, and was dominated by Halobacteria, Methanobacteria and Methanomicrobia (**Fig. 12**). Halobacteria was also present at 4.5 m, but with low abundance. The different archaeal (and bacterial) community compositions at 3 m could indicate that

this sample is representing the aeolian depositional environment (**Fig. 3**), as discussed in Section 5.1. Further, the 3 m sample is close to the active layer (0-1 m), and may have been seasonally thawed under warmer climate conditions (Hanne Christiansen, personal communication). The occurrence of heterotrophic microorganisms together with lower salinity, slightly higher TOC (1.8 %) and high cell numbers at this depth, supports this by indicating influence of surface processes such as migration of meteoric water and organic carbon.

The Euryarchaeotal class Halobacteria was abundant at 3 m (**Fig. 12**). Organisms belonging to this group are strictly dependent on high salt concentrations for growth and cellular activity. They grow optimally at salt concentrations above 150 g L^{-1}, and lyse at concentrations below 100 g L^{-1} (Oren, 2006). All members of the class Halobacteria could be further classified to the family Halobacteriaceae, and these are aerobic heterotrophs (Oren, 2006). However, some species of Halobacteriaceae have shown to be able to grow anaerobically by fermentation or by the use of nitrate, dimethylsulfoxide, fumarate or trimethylalamine N-oxide as electron acceptors (Oren, 2006). Halobacteria have also been detected in a 9 m permafrost core and in a 14 m permafrost/ground ice core from the Canadian High Arctic (Steven et al., 2007, Steven et al., 2008).

The presence of halophilic archaea at 3 m is in contrast with our pore water data, where the Na^+ and Cl^- concentrations at this depth were below classical seawater values (**Fig. 6**). However, under frozen conditions solutes get enriched in thin unfrozen water films, and it is within these films we would expect microbial activity. If we assume that 3 % of the water at 3 m is in a liquid state, which is within the range described by Gilichinsky et al. (2007), and that all NaCl is within the liquid water, the calculated NaCl concentration at 3 m becomes around 150 g L^{-1} (**Appendix B**), which is the optimal salt concentration for Halobacteria (Oren, 2006). Steven and colleagues demonstrated that permafrost microbial communities were adapted to high salt concentrations, as many permafrost isolates could grow on medium with 7 % NaCl (Steven et al., 2007). It has also been suggested that there might be a link between psychrotolerance and halotolerance (Vishnivetskaya et al., 2000). However, we have found no data describing the salt concentration ranges in water films in relevant systems/habitats.

Methanobacteria and Methanomicrobia were present in the 3 m sample of the DH8 core (**Fig. 12**), and these groups are recognised methanogens. The presence of methanogens at 3 m was confirmed by qPCR, where the highest *mcrA* (marker gene for methanogenic archaea) copy numbers per gram were found at this depth (6.5×10^3 copies g^{-1}) (**Fig. 7**). In the rest of the samples the *mcrA* copy numbers were low (9.4×10^1 to 5.8×10^2 copies g^{-1}) and close to the limit of detection for qPCR. Methanogens play an important role in the global carbon cycle, where they catalyze the terminal step in anaerobic degradation of organic carbon when all electron acceptors except CO_2 are used (Bonin and Boone, 2006). Methanogenic archaea have been found in studies of the active layer of Svalbard (Hoj et al., 2005, Tveit et

al., 2013) and in a 2 m permafrost sample from the Canadian High Arctic (Yergeau et al., 2010). Additionally, methanogenic archaea have been isolated from Holocene and Pliocene permafrost sediments in Siberia (Rivkina et al., 2007). The methanogens detected at 3 m and 4.5 m form a minor part of the total microbial communities at these depths, however they might be ecologically important as they are the last step of anaerobic degradation of organic matter, and can be essential for the thermodynamics of these processes. A change in the environmental conditions, for example increasing temperature causing permafrost thaw below 3 m, could increase the abundance and activity of these groups and lead to methane emissions from this site.

At 4.5 m the archaeal community was dominated by Methanomicrobia and MCG (**Fig. 12)**. From 16.5 m to 57 m, the archaeal community composition was less diverse than in the two top samples, with the Deep Sea Archaeal Group (DSAG) making up 40-99 % of the archaeal community, increasing their relative abundance with depth (**Fig. 12**). The dominance of DSAG was confirmed by DSAG specific qPCR, where DSAG 16S rRNA made up 45-188 % of archaeal 16S rRNA genes from 16.5 m to 57 m (**Fig. 7**). The DSAG is commonly found in marine sediments (Jorgensen et al., 2013). They are restricted to anaerobic or microaerophilic environments, and have been proposed (but debated) to be involved in anaerobic methane oxidation (Biddle et al., 2006, Knittel et al., 2005). Recent studies by Jørgensen and colleagues showed a significant correlation between the relative abundance of DSAG and concentrations of organic carbon and iron oxide (Jorgensen et al., 2012, Jorgensen et al., 2013). These observations lead to the suggestion that DSAG is involved in iron and/or manganese reduction while oxidising organic carbon (Jorgensen et al., 2013). No correlations between organic carbon, iron oxide, dissolved Fe^{2+} and the DSAG abundance from qPCR were found in this study. However, very few data points were included in the statistical analysis, as the DSAG was only detected from 16.5 m to 57 m, which could be the reason for the lack of statistical significance. Another explanation could be that the DSAG is not metabolically active in the permafrost, and then they do not influence the organic carbon and iron oxide concentrations.

Phylogenetic analyses of the DSAG sequences obtained from 454 pyrosequencing and from the two different 16S rRNA clone libraries, revealed that 98 % of the sequences fall within one single group of the Gamma lineage (Jorgensen et al., 2013), together with sequences recovered from a diverse range of marine settings (**Fig. 18**). It is also worth mention that a significant amount of the sequences in that group originated from hypersaline environments (Steffen Leth Jørgensen, personal communication). 1.7 % of the permafrost sequences clustering with the Gamma lineage affiliated with sequences exclusively originating from hypersaline ponds in Guerrero Negro in Baja California, Mexico (Jahnke et al., 2008, Ley et al., 2006, Robertson et al., 2009). The presence of DSAG lineages that can tolerate high salt concentrations strengthens the assumption that salts accumulate in thin unfrozen water films in permafrost (Steven et al., 2006), and that halotolerance is important for permafrost microorganisms (Vishnivetskaya et al., 2000).

Thermoplasmata was present at all depths, except at 4.5 m and 57 m (**Fig. 12**). Most of the reads affiliating with this class could be further classified to the Marine Benthic Group D. Single-cell genomic sequencing of three cells belonging to this group, indicated that they are capable of exogenous protein degradation in cold anoxic environments (Lloyd et al., 2013). The Soil Crenarchaeotic Group (SCG) was detected from 4.5 m to 31.5 m, but at low relative abundance compared to the other classes present at these depths (**Fig. 12**). This class was the most abundant archaeal phylotype found in the active layer and permafrost in Axel Heiberg wetlands of the Canadian High Arctic (Wilhelm et al., 2011), and was also the majority of archaeal clones found in a 14 m permafrost/ground ice core at a different site in the Canadian High Arctic (Steven et al., 2008). It has been reported that SCG appeared to have a competitive advantage over other archaeal groups at low temperatures (Hoj et al., 2008), suggesting that organisms belonging to the SCG could be of likely significance to cryoenvironments (Wilhelm et al., 2011). Although other archaeal taxa were found to be more abundant than the SCG in the DH8 permafrost core, it is still possible that members of this group are ecologically important in this system.

As for Bacteria, the archaeal community composition of the 60 m shale sample differed greatly from the other samples. Here, the archaeal community was dominated by Thermoplasmata and Archaeoglobi (**Fig. 12**). All reads affiliating with Thermoplasmata could be further classified to the Deep Sea Hydrothermal Vent Group 2 (DHVEG-2). Isolation and cultivation of a representative from this group revealed that the isolate was thermoacidophilic, sulfur- or iron-reducing heterotroph (Reysenbach et al., 2006). The Archaeoglobi affiliated reads could be further classified to the family Archaeoglobaceae. Representatives of this family are hyperthermophiles that grow by anaerobic oxidation of organic compounds or molecular hydrogen (Huber and Stetter, 2001). The shale at 60 m is part of the Carolinefjellet formation, deposited during the early Cretaceous (Parker, 1967). The presence of thermophilic bacterial and archaeal taxonomic groups in the shale sample could be due to the fact that the shale has been buried and subjected to warmer temperatures, before being uplifted to present level.

To sum up the study of the microbial community compositions, a dominance of anaerobic or fermentative microbial groups using organic carbon as an energy source was found throughout the core. Both the archaeal and bacterial communities at 3 m and 60 m were different from the rest of the core, most likely due to influence of surface processes at the top and the different depositional material at 60 m. From 16.5 m to 57 m the DSAG made up 40-99 % of the archaeal communities. The detection of microbial groups affiliating with marine environments demonstrates that the marine origin of the DH8 core influence the microbial community composition, even 5000-6000 years after deposition of the core material, and 3000 years with permafrost conditions. Permafrost represents an extreme environment where the microorganisms must survive constant subzero temperatures, low water and nutrient content, and continuous exposure to background radiation over thousands of years (Ponder et al., 2004). Over time, the unique conditions within the

permafrost may have preferentially selected for certain microbial groups that can survive in this harsh environment, leading to reduced microbial diversity. It is intriguing to speculate that the DSAG might be well adapted to long-term survival in such an extreme environment, as this group was clearly the dominating microbial group from 16.5 m to 57 m. However, it might be that this group was over-amplified due to primer degeneracy problems, as discussed in Section 4.4. Yet, the dominance by DSAG within Archaea was confirmed by DSAG specific qPCR (**Fig. 7**) and in the clone libraries (**Fig. 16**), and it is therefore possible that members of this group are well adapted for long-term survival in permafrost.

5.4 Microbial sulphate reduction

The low SO_4^{2-} concentration found at 3 m, 4.5 m and from 31.5 m and below is probably due to microbial sulphate reduction. This is supported by high copy numbers (qPCR) of *dsrB* marker genes throughout the core (**Fig. 7**). It remains unknown if this is a preserved geochemical signature, or if active sulphate reduction currently is happening. However, it has been demonstrated that slow microbial activity occur throughout the DH8 permafrost core (Hovgaard, 2014), making it likely to assume that some of this activity includes microbial sulphate reduction. The *dsrB* profile follows the bacterial 16S rRNA profile, indicating that sulfate reduction is, or has been, an important bacterial metabolism throughout the core (**Fig. 7**). No correlation between *dsrB* copy numbers and SO_4^{2-} concentrations was found. Interestingly, the 16S rRNA amplicon library showed no dominance of known sulphate-reducing groups (**Figs. 9 and 10**), in contrast to the qPCR data. Possible explanations for these conflicting results could be that the sulphate reducing community belongs to the poorly characterised Candidate divisions (TM7, OP9, Chloroflexi Subdivision 11), or that the primers used for sequencing do not pick up the sulphate reducing prokaryotes in the DH8 permafrost core. However, the primers should theoretically cover all known sulphate reducing prokaryotes (Laila Reigstad, personal communication). To investigate if the sulphate-reducing microbial community consisted of known sulphate-reducing phylogenetic groups or if it was dominated by uncultured representatives, a *dsrB* clone library was made from the 54 m sample, and then a rough phylogenetic analysis was performed on the sequenced *dsrB* clones.

Phylogenetic analyses of the sequenced *dsrB* clones on amino acid level revealed high diversity and bacterial origin for all sequences, affiliating with sequences found in marine environments (**Appendix D**). 39 % of the sequences fall into the dsr Group V (**Fig. 17**) (Kaneko et al., 2007), which exclusively consists of uncultured sulphate reducing bacteria (SRB), that are widely present in marine environments (Kaneko et al., 2007). These sequences were also related to six clones from Guaymas Basin hydrothermal vent area (AAO38145, AAO38115, AAO38111, AAO38123, AAO38135, AAO38121) (Dhillon et al., 2003), and to clones obtained from an unpublished study of Svalbard marine sediments (AHK06587, AHK06561, AHK06619, AHK06595, AHK06639, AHK06623). 31 % of the sequences belonged to dsr Group I (**Fig. 17**) (Kaneko et al., 2007), consisting of uncultured representatives from the deltaproteobacterial family *Desulfobacteriaceae*, and these

sequences were also related to a Baltic Sea clone (AFI98570) (de Rezende et al., 2013), a salt marsh clone (AAV68667) (Bahr et al., 2005), and to clones obtained from an unpublished study of Svalbard marine sediments (AHK06567, AHK06679, AHK06715, AHK06631, AHK06717, AHK06677) (**Appendix D**).

15 % of the sequences affiliated with dsr Group III (**Fig. 17**), consisting of no known cultured species, but closely related to *Pelotomaculum* sp. (non-sulphate reducing Firmicutes), and the thermophilic sulphate-reducer *Thermodusulfobium narugense* (Kaneko et al., 2007). Ten % of the *dsrB* protein sequences belonged to dsr Group II (Kaneko et al., 2007), together with uncultured representatives related to the deltaproteobacterial family *Syntrophobacteraceae*. These sequences were also related to a clone from an unpublished study on Svalbard marine sediments (AHK06565). Two sequences (5 %) clustered with a group consisting of only uncultured representatives, together with clones from unpublished studies from marine sediments at Svalbard (AHK06659), Changjiang estuary (AHG51095) and harbour sediments in China (AAZ15797) (**Appendix D**).

The presence of *dsrB* protein sequences in the 54 m sample that were closely related to sequences associated with marine sediments, gives a further indication that the marine origin of the DH8 permafrost core is determining the microbial community composition. Although this was only confirmed in the 54 m sample, it is likely to assume that this is also true for the rest of the core, as the microbial community composition showed a dominance of taxonomic groups associated with marine environments. The phylogenetic distribution of the *dsrB* protein sequences from the 54 m sample showed that 59 % of the sequences affiliated with groups with no known cultured representatives (dsr Group V and Group III) (**Fig. 17 and Appendix D**). This could explain why there are no known sulphate-reducing microbial groups in the 454 pyrosequencing library. However, 41 % of the permafrost *dsrB* sequences were closely related families belonging to Deltaproteobacteria (**Fig. 17**), but there were no deltaproteobacterial reads in the 454 pyrosequencing library (**Fig. 10**), or in the clone library made with the 515F/1391R primers at 54 m (**Fig. 16**). The absence of deltaproteobacterial reads in the 454 pyrosequencing library could be related to the primer degeneracy effects that were discussed in Section 4.4, or that the deltaproteobacterial sulphate reducers are present at such low numbers that they do not come up when analysing 16S rRNA genes.

5.5 Geomicrobial characterisation of the DH8 permafrost core in relation to CO_2 storage at Svalbard.

Monitoring of Carbon Capture and Storage (CCS) projects is crucial for early detection of an eventual CO_2 leakage and associated changes in the surrounding environment. Microorganisms present in soil or sediments will react to changes in their environment, and monitoring their numbers and community composition could give early warnings of CO_2 leakage (Noble et al., 2012). One of the objects of this Master's thesis was to perform a

baseline study of the DH8 60 m permafrost core, to assess the present state of the microbial communities in the permafrost before eventual CO_2 injection starts at the site.

In this study, microbial abundance and the abundance of selected functional marker genes have been analysed with qPCR, and the microbial community composition was determined using 16S rRNA amplicon 454 pyrosequencing. The data obtained in this Master's thesis will be submitted to the online database Virtual CO_2 Laboratory Project (VIRCOLA), making the data available for others working with the CO_2 storage project at Svalbard. Future monitoring of the microbial communities at this site could possibly detect an eventual CO_2 leakage, by comparing the new microbial results to our baseline study. However, there is limited knowledge about which microbial parameters that would change in response to a potential CO_2 leakage. It has been suggested that methanogenic archaea and sulphate-reducing bacteria are important groups to consider in microbial monitoring of CCS projects (Noble et al., 2012). These groups were quantified with qPCR in this study, and it could be important to continue monitoring their numbers if CO_2 should be stored at Svalbard in the future.

The methods and techniques in microbial research are constantly evolving, and it is very likely that future studies will employ different methods. This will result in data that are not directly comparable to our baseline. And, as we have seen in this study, it is challenging to compare different community characterisation methods even when using the same samples and same DNA extracts (**Figs. 15 and 16**). Great care should thus be taken when interpreting community structure data for such purposes. Even though direct data comparison could be problematic in the future, the detection of the different microbial groups in this study confirms their presence, but it does not imply that the microbial groups that are not detected are not present in this system. New and improved methods could give a deeper sampling of the microbial communities, but the results from our study provide important reference points for future microbial analyses.

5.6 Potential methane seep from sandstone reservoir

Another CO_2 storage related aspect is the cap rock integrity. High content of biogenic, over-pressured methane is present in permeable sandstones at 150-200 m depth at the study site (Pål Tore Mørkved, personal communication). This observation could possibly be used to assess the sealing properties of the shale and permafrost overburden, to evaluate their functioning as a secondary seal for the deeper (approximately 900 m) CO_2 storage. If there was a diffuse seep of methane from the sandstone reservoir reaching the permafrost above, it would be expected to find microorganisms using this methane as an energy source. To assess this, microbial methane metabolisms were investigated in this project through quantification of the functional marker genes *pmoA* and *mcrA*. Microorganisms that have the gene *pmoA* are able to perform methane oxidation, and thus serve as methane sinks in the environment (Bourne et al., 2001). The *mcrA* is a marker gene for archaeal methane

production (methanogenesis), and these organisms are a source of methane to the environment (Nunoura et al., 2008).

Methanotrophic bacteria have been identified at low abundance in the active layer of Svalbard wetlands (Wartiainen et al., 2006, Wartiainen et al., 2003). In this study methanotrophic microorganisms were first targeted using primers for the Methanol Dehydrogenase gene (*mxaF*) (Lau et al., 2013), as methane monooxygenase (*pmoA*) genes have shown to be highly similar to ammonia monooxygenase (*amoA*) genes, and are thus less specific. *mxaF* is ubiquitous in methanotrophs, in contrast to other methanotroph-specific genes, such as *pmoA* genes, which are absent in some methanotrophic proteobacteria (Lau et al., 2013). However, the *mxaF* primers showed a negative result in PCR analysis. PCR and qPCR analysis using primers for the *pmoA* gene showed very weak PCR products for this gene at 3 m and 16.5 m (see "Results", Section 3.5), and the microbial community composition revealed no known methanotrophic organisms throughout the core (**Figs. 9-12**). Consequently, we can conclude that microbial methanotrophy is not a major metabolism in the analysed depths of the DH8 core.

Although the *pmoA* quantification and microbial community analysis do not provide any indications of a methane seep, it is difficult to reject the hypothesis with only this observation. Analysis of microbial metabolisms and community composition is not the preferred method to search for gas seeps. However, if methanotrophic microorganisms were present at high abundance, and if there were no methanogenic archaea present that produce methane, it could have given indications of a methane seep. Clearly, more analyses of for example gas content and isotopic signature of eventual methane gas in the permafrost are needed to confirm if there is a seep of methane through the permafrost.

6. Conclusions

A 60 m long permafrost core from Svalbard was analysed in this study. The core is mainly of marine origin, with aeolian deposits in the top, extending through deltaic, marine and glacial deposits, ending in shale at 60 m (**Fig. 3**).

Geomicrobial analyses of seven depths (3, 4.5, 16.5, 31.5, 54, 57 and 60 m) of the permafrost core revealed different geochemistry, microbial abundance and community composition with depth. At 3 m, lower salt concentrations (**Fig. 6**), higher organic carbon content (**Table 2**), higher microbial abundance (**Fig. 7**) and different microbial community structure (**Figs. 9-12**) than below, indicated influence of surface processes such as migration of meteoric water, organic carbon and possible freezing/thawing cycles in earlier times.

From 16.5 m to 57 m the geochemistry (**Fig. 6**) and microbial community composition (**Figs. 9-12**) had a marine signature that was not found above, indicating that the marine depositional environment is influencing the geochemistry and microbial community structure, even 5000-6000 years after deposition and 3000 years with permafrost conditions. A high abundance of archaeal groups that have not been detected in other permafrost studies were found in the 60 m permafrost core (**Fig. 12**). The reason for this could be related to the marine origin of the core, or the higher sequencing depth that was used here.

A dominance of anaerobic or fermentative microbial groups was detected at all seven depths (**Figs. 9-12**). This observation, together with high Fe^{2+} and Mn^{2+} concentrations at 3 m, and high NH_4^+, but low NO_3^- concentrations throughout the core (**Fig. 6**), indicated an anoxic environment. The presence of mainly heterotrophic microorganisms together with a significant correlation ($R^2 = 0.917$) between bacterial abundance and organic carbon content, indicated that organic carbon is the main energy source for the microbial community. The bacterial abundance was negatively correlated with dissolved nutrients such as K^+, Mg^{2+}, NH_4^+, PO_4^{2-} and NO_3^-, suggesting that the bacterial population may have consumed these ions, causing subsequent low concentrations. Other shifts in mineral composition or pore water chemistry that could explain shifts in bacterial and archaeal abundance were not found.

Estimated total cell numbers from quantitative PCR (qPCR) analysis corresponded to cell numbers found in other deep permafrost cores (**Fig. 21**). Low SO_4^{2-} concentrations (compared to seawater) at 3 m and 4.5 m and from 31.5 m to 57 m (**Fig. 6**), and high abundance of *dsrB* genes from 3 m to 57 m (**Fig. 7**) indicated that sulphate reduction is, or have been, an important metabolism in the permafrost core. Phylogenetic analysis of *dsrB* sequences from the 54 m sample revealed bacterial origin for all sequences, clustering with sequences found in marine environment (**Fig. 17 and Appendix D**).

Bacterial and archaeal quantification by qPCR (**Fig. 7**) showed bacterial dominance down to 31.5 m, and similar bacterial and archaeal 16S rRNA copies per gram from 54 m to 60 m. In

contrast to this relationship, the 16S rRNA amplicon 454 pyrosequencing library revealed dominance of bacterial reads in the two top samples, but from 16.5 m to 60 m archaeal reads accounted for 83-99.8 % of the total reads (**Table 5**). These conflicting results are believed to be caused by primer degeneracy problems in the amplicon library, where the universal prokaryotic primers that were used for sequencing had preferential binding to target templates with G/C in the degenerate primer position, leading to underrepresentation of microbial groups containing A/T in that position.

The Deep Sea Archaeal group (DSAG) was the dominating microbial group from 16.5 m to 57 m (**Fig. 12**). Although this group could have been over-amplified in the pyrosequencing library due to primer-degeneracy problems, the dominance of DSAG within Archaea was confirmed in DSAG specific qPCR (**Fig. 7**) and in the 16S rRNA clone library made with a different primer set (**Fig. 16**). This could indicate that the DSAG is well adapted for long-term survival in permafrost, although detection of DNA sequences in permafrost does not necessarily indicate that they belong to viable organisms.

In relation to potential underground CO_2 storage at Svalbard, the results from this study serve as important reference points for future microbial monitoring during eventual CO_2 storage at the site. However, great care should be taken when comparing microbial community data obtained by different methods. No microbial indicators of a diffuse seep of methane from the sandstone formation at 150-200 m were found, although further analyses are needed to confirm this.

7. Future work

The first priority for future microbial studies on the DH8 permafrost core would be to re-run the microbial community analysis, due to the observed underrepresentation of bacterial taxa (see “Discussion of methods”, Section 4.4). This work is currently ongoing, using a new Ion Torrent sequencer at the University of Bergen. Unfortunately, these new results were too late for deadline of this thesis.

To further analyse complex interactions between geological, geochemical and microbial data, statistical analysis using Principal Component Analysis (PCA) could be done. Since the bacterial diversity was underestimated in the 454 pyrosequencing results, PCA was not performed on this dataset. Instead, PCA will be carried out on the Ion Torrent sequenced samples. To improve the strength of statistical tests, one could analyse more samples per depth.

A common approach in microbial ecology is to characterise the community biodiversity by estimating diversity indices with for example Chao species richness estimator (Chao et al., 2005), rarefaction curves (Gotelli and Colwell, 2001), Shannon index and Simpson index (Hill et al., 2003). These indices were not assessed in this project, due to the underrepresentation of bacterial taxa observed in the 454 pyrosequencing library. With the new Ion Torrent sequenced dataset it would hopefully be possible to estimate diversity indices for the permafrost microbial community and compare to other studies.

As the 16S rRNA clone library made with the primers 787F/1391R showed a dominance of sequences with G/C in the target site of the degenerate primer position (see see “Discussion of methods”, Section 4.4), it would have been interesting to investigate if this is also the case for the 454 pyrosequencing reads that were generated with the same primers. If this is also true for the pyrosequencing reads, cautions need to be taken when using these primers in future studies.

To ensure that there has been no contamination from the drilling fluid, it would have been useful to investigate the microbial community composition of the river water used as drilling fluid. CENPERM in Copenhagen have samples of the river water, which is possible to get hold of for future analyses. It would also be very interesting to obtain samples from the active layer in this area, to see if there are differences in geochemical and microbial parameters between the active layer and permafrost samples. Samples of the active layer from a bore hole close to the DH8 drill site are also available at CENPERM.

The molecular analyses performed in this study were performed on DNA level. It is not possible to know if the isolated DNA originated from dead cells or naked DNA preserved in the permafrost, or from viable cells. Yergeau et al. (2010) used Propidium Monoazide (PMA) to differentiate naked DNA and dead cells from viable cells. Intact cell membranes are impermeable to PMA, and the addition of this substance allows distinguishing whether the analysed DNA originated from dead cells and/or naked DNA, or from viable cells.

Hovgaard (2014) suggested that slow metabolic activity occur throughout the DH8 permafrost core by using a D:L amino acid model. To further verify if there are active microbial cells present, it would have been interesting to try to isolate RNA from the samples, and do microbial abundance and community analysis on RNA level. Other methods to detect active microbial cells include Fluorescent In Situ Hybridization (FISH), measurements of gas fluxes from the core and measurements of incorporation of radiocarbon labelled substrates. It would also be interesting to try to cultivate and isolate microorganisms in the permafrost samples, in order to evaluate potential microbial growth on a selection of nutrient media.

References

454 LIFE SCIENCES CORP. 2011. *454 Sequencing System Guidelines for Amplicon Experimental Design* [Online]. Roche. Available: http://my454.com/downloads/my454/applications-info/454SequencingSystem_GuidelinesforAmpliconExperimentalDesign_July2011.pdf [Accessed 19.11 2013].

ALBERS, C. N., JENSEN, A., BÆLUM, J. & JACOBSEN, C. S. 2013. Inhibition of DNA Polymerases Used in Q-PCR by Structurally Different Soil-Derived Humic Substances. *Geomicrobiology Journal,* 30**,** 675-681.

ALBERTSEN, M., HUGENHOLTZ, P., SKARSHEWSKI, A., NIELSEN, K. L., TYSON, G. W. & NIELSEN, P. H. 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotechnology,* 31**,** 533-+.

BAHR, M., CRUMP, B. C., KLEPAC-CERAJ, V., TESKE, A., SOGIN, M. L. & HOBBIE, J. E. 2005. Molecular characterization of sulfate-reducing bacteria in a New England salt marsh. *Environ Microbiol,* 7**,** 1175-85.

BAKERMANS, C., BERGHOLZ, P. W., AYALA-DEL-RÍO, H. & TIEDJE, J. 2009. Genomic Insights into Cold Adaptation of Permafrost Bacteria. *In:* MARGESIN, R. (ed.) *Permafrost Soils, Soil Biology.* Berlin Heidelberg, Germany: Springer-Verlag**,** 159.

BEAUBIEN, S., CIOTOLI, G., COOMBS, P., DICTOR, M., KRUGER, M., LOMBARDI, S., PEARCE, J. & WEST, J. 2008. The impact of a naturally occurring CO_2 gas vent on the shallow ecosystem and soil chemistry of a Mediterranean pasture (Latera, Italy). *International Journal of Greenhouse Gas Control,* 2**,** 373-387.

BIDDLE, J. F., LIPP, J. S., LEVER, M. A., LLOYD, K. G., SORENSEN, K. B., ANDERSON, R., FREDRICKS, H. F., ELVERT, M., KELLY, T. J., SCHRAG, D. P., SOGIN, M. L., BRENCHLEY, J. E., TESKE, A., HOUSE, C. H. & HINRICHS, K. U. 2006. Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc Natl Acad Sci U S A,* 103**,** 3846-3851.

BLANCO, Y., PRIETO-BALLESTEROS, O., GOMEZ, M. J., MORENO-PAZ, M., GARCIA-VILLADANGOS, M., RODRIGUEZ-MANFREDI, J. A., CRUZ-GIL, P., SANCHEZ-ROMAN, M., RIVAS, L. A. & PARRO, V. 2012. Prokaryotic communities and operating metabolisms in the surface and the permafrost of Deception Island (Antarctica). *Environ Microbiol,* 14**,** 2495-510.

BONIN, A. & BOONE, D. 2006. The Order Methanobacteriales. *In:* DWORKIN, M., FALKOW, S., ROSENBERG, E., SCHLEIFER, K.-H. & STACKEBRANDT, E. (eds.) *The Prokaryotes.* New York: Springer 231-243.

BOURNE, D. G., MCDONALD, I. R. & MURRELL, J. C. 2001. Comparison of pmoA PCR Primer Sets as Tools for Investigating Methanotroph Diversity in Three Danish Soils. *Appl Environ Microbiol,* 67**,** 3802-3809.

BRAATHEN, A., BÆLUM, K., CHRISTIANSEN, H. H., DAHL, T., EIKEN, O., ELVEBAKK, H., HANSEN, F., HANSSEN, T. H., JOCHMANN, M., JOHANSEN, T. A., JOHNSEN, H., LARSEN, L., LIE, T., MERTES, J., MØRK, A., MØRK, M. B., NEMEC, W., OLAUSSEN, S., OYE, V., RØD, K., TITLESTAD, G. O., TVERANGER, J. & VAGLE, K. 2012. The Longyearbyen CO_2 Lab of Svalbard, Norway – initial assessment of the geological conditions for CO_2 sequestration. *Norwegian Journal of Geology,* 92**,** 353-376.

BRU, D., MARTIN-LAURENT, F. & PHILIPPOT, L. 2008. Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Appl Environ Microbiol,* 74**,** 1660-1663.

CHAO, A., CHAZDON, R. L., COLWELL, R. K. & SHEN, T. J. 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters,* 8**,** 148-159.

DE REZENDE, J. R., KJELDSEN, K. U., HUBERT, C. R., FINSTER, K., LOY, A. & JORGENSEN, B. B. 2013. Dispersal of thermophilic Desulfotomaculum endospores into Baltic Sea sediments over thousands of years. *ISME J,* 7**,** 72-84.

DENICH, T. J., BEAUDETTE, L. A., LEE, H. & TREVORS, J. T. 2003. Effect of selected environmental and physico-chemical factors on bacterial cytoplasmic membranes. *Journal of Microbiological Methods,* 52**,** 149-182.

DHILLON, A., TESKE, A., DILLON, J., STAHL, D. A. & SOGIN, M. L. 2003. Molecular Characterization of Sulfate-Reducing Bacteria in the Guaymas Basin. *Appl Environ Microbiol,* 69**,** 2765-2772.

DODSWORTH, J. A., BLAINEY, P. C., MURUGAPIRAN, S. K., SWINGLEY, W. D., ROSS, C. A., TRINGE, S. G., CHAIN, P. S. G., SCHOLZ, M. B., LO, C.-C., RAYMOND, J., QUAKE, S. R. & HEDLUND, B. P. 2013. Single-cell and metagenomic analyses indicate a fermentative and saccharolytic lifestyle for members of the OP9 lineage. *Nature Communications,* 4.

GARCIA, J.-L., OLLIVIER, B. & WHITMAN, W. 2006. The Order Methanomicrobiales. *In:* DWORKIN, M., FALKOW, S., ROSENBERG, E., SCHLEIFER, K.-H. & STACKEBRANDT, E. (eds.) *The Prokaryotes.* New York: Springer 208-230.

GARRITY, G. M., HOLT, J. G., HATCHIKIAN, E. C., OLLIVIER, B. & GARCIA, J.-L. 2001. Phylum BIII. Thermodesulfobacteria phy. nov. *In:* BOONE, D. R., CASTENHOLZ, R. W. & GARRITY, G. M. (eds.) *Bergey's Manual of Systematic Bacteriology.* New York: Springer 389-393.

GEETS, J., BORREMANS, B., DIELS, L., SPRINGAEL, D., VANGRONSVELD, J., VAN DER LELIE, D. & VANBROEKHOVEN, K. 2006. DsrB gene-based DGGE for community and diversity surveys of sulfate-reducing bacteria. *J Microbiol Methods,* 66**,** 194-205.

GILBERT, G. L. 2014. *Sedimentology and geocryology of an Arctic fjord head delta (Adventdalen, Svalbard),* Oslo, Department of Geosciences, University of Oslo. Master's thesis.

GILBERT, J. A., HILL, P. J., DODD, C. E. R. & LAYBOURN-PARRY, J. 2004. Demonstration of antifreeze protein activity in Antarctic lake bacteria. *Microbiology,* 150**,** 171-180.

GILICHINSKY, D. 2002. Permafrost. *In:* BITTON, G. (ed.) *Encyclopedia of environmental microbiology.* New York: Wiley**,** pp 2367–2385.

GILICHINSKY, D., VISHNIVETSKAYA, T., PETROVA, M., SPIRINA, E. V., MAMYKIN, V. & RIVKINA, E. 2008. Bacteria in permafrost. *In:* MARGESIN, R., SCHINNER, F., MARX, J.-C. & GERDAY, C. (eds.) *Psychrophiles: from Biodiversity to Biotechnology.* Berlin Heidelberg, Germany: Springer-Verlag**,** pp 83–102.

GILICHINSKY, D. A., WILSON, G. S., FRIEDMANN, E. I., MCKAY, C. P., SLETTEN, R. S., RIVKINA, E. M., VISHNIVETSKAYA, T. A., EROKHINA, L. G., IVANUSHKINA, N. E., KOCHKINA, G. A., SHCHERBAKOVA, V. A., SOINA, V. S., SPIRINA, E. V., VOROBYOVA, E. A., FYODOROV-DAVYDOV, D. G., HALLET, B., OZERSKAYA, S. M., SOROKOVIKOV, V. A., LAURINAVICHYUS, K. S., SHATILOVICH, A. V., CHANTON, J. P., OSTROUMOV, V. E. & TIEDJE, J. M. 2007. Microbial populations in Antarctic permafrost: biodiversity, state, age, and implication for astrobiology. *Astrobiology,* 7**,** 275-311.

GITTEL, A., BARTA, J., KOHOUTOVA, I., MIKUTTA, R., OWENS, S., GILBERT, J., SCHNECKER, J., WILD, B., HANNISDAL, B., MAERZ, J., LASHCHINSKIY, N., CAPEK, P., SANTRUCKOVA, H., GENTSCH, N., SHIBISTOVA, O., GUGGENBERGER, G., RICHTER, A., TORSVIK, V. L., SCHLEPER, C. & URICH, T. 2014. Distinct microbial communities associated with buried soils in the Siberian tundra. *ISME J,* 8**,** 841-53.

GOTELLI, N. J. & COLWELL, R. K. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters,* 4**,** 379-391.

HANSEN, A. A., HERBERT, R. A., MIKKELSEN, K., JENSEN, L. L., KRISTOFFERSEN, T., TIEDJE, J. M., LOMSTEIN, B. A. & FINSTER, K. W. 2007. Viability, diversity and composition of the bacterial community in a high Arctic permafrost soil from Spitsbergen, Northern Norway. *Environmental Microbiology,* 9**,** 2870-2884.

HENKE, W., HERDEL, K., JUNG, K., SCHNORR, D. & LOENING, S. A. 1997. Betaine Improves the PCR Amplification of GC-Rich DNA Sequences. *Nucleic Acids Research,* 25**,** 3957-3958.

HILL, T. C. J., WALSH, K. A., HARRIS, J. A. & MOFFETT, B. F. 2003. Using ecological diversity measures with bacterial communities. *FEMS Microbiol Ecol,* 43**,** 1-11.

HINKEL, K. M. & NELSON, F. E. 2003. Spatial and temporal patterns of active layer thickness at Circumpolar Active Layer Monitoring (CALM) sites in northern Alaska, 1995-2000. *Journal of Geophysical Research-Atmospheres,* 108.

HOJ, L., OLSEN, R. A. & TORSVIK, V. L. 2005. Archaeal communities in High Arctic wetlands at Spitsbergen, Norway (78 degrees N) as characterized by 16S rRNA gene fingerprinting. *FEMS Microbiol Ecol,* 53**,** 89-101.

HOJ, L., OLSEN, R. A. & TORSVIK, V. L. 2008. Effects of temperature on the diversity and community structure of known methanogenic groups and other archaea in high Arctic peat. *ISME J,* 2**,** 37-48.

HOUBA, V. J. G., TEMMINGHOFF, E. J. M., GAIKHORST, G. A. & VAN VARK, W. 2000. Soil analysis procedures using 0.01 M calcium chloride as extraction reagent. *Communications in Soil Science and Plant Analysis,* 31**,** 1299-1396.

HOVGAARD, J. 2014. *Organic – geochemical studies of microbial activity in permafrost,* Aarhus, Center of Geomicrobiology, Aarhus University, Master's thesis.

HUBER, H. & STETTER, K. O. 2001. Order I. Archaeoglobales ord. nov. *In:* BOONE, D. R. & CASTENHOLZ, R. W. (eds.) *Bergey's Manual of Systematic Bacteriology Volume 1: The Archaea and the deeply branching and phototrophic Bacteria (2nd ed.).* New York: Springer Verlag**,** 169.

HUGENHOLTZ, P., TYSON, G. W., WEBB, R. I., WAGNER, A. M. & BLACKALL, L. L. 2001. Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. *Appl Environ Microbiol,* 67**,** 411-9.

IPCC 2005. IPCC Special Report on Carbon Dioxide Capature and Storage. *In:* PREPARED BY WORKING GROUP III OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE [METZ, B., O. DAVIDSON, H. C. DE CONINCK, M. LOOS, AND L. A. MEYER]. (ed.). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press**,** 442 pp.

IPCC 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change *In:* STOCKER, T. F., D. QIN, G.-K. PLATTNER, M. TIGNOR, S.K. ALLEN, J. BOSCHUNG, A. NAUELS, Y. XIA, V. BEX AND P.M. MIDGLEY (ed.). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press**,** 1535 pp.

ISO 14688-1:2002. 2002. *Geotechnical investigation and testing -- Identification and classification of soil -- Part 1: Identification and description*. Geneva, Switzerland, International Organization for Standardization (ISO).

JAHNKE, L. L., ORPHAN, V. J., EMBAYE, T., TURK, K. A., KUBO, M. D., SUMMONS, R. E. & DES MARAIS, D. J. 2008. Lipid biomarker and phylogenetic analyses to reveal archaeal biodiversity and distribution in hypersaline microbial mat and underlying sediment. *Geobiology,* 6**,** 394-410.

JANSSON, J. K. & TAS, N. 2014. The microbial ecology of permafrost. *Nat Rev Microbiol,* 12**,** 414-25.

JOHNSON, S. S., HEBSGAARD, M. B., CHRISTENSEN, T. R., MASTEPANOV, M., NIELSEN, R., MUNCH, K., BRAND, T., GILBERT, M. T., ZUBER, M. T., BUNCE, M., RONN, R., GILICHINSKY, D., FROESE, D. & WILLERSLEV, E. 2007. Ancient bacteria show evidence of DNA repair. *Proc Natl Acad Sci U S A,* 104**,** 14401-5.

JORGENSEN, S. L., HANNISDAL, B., LANZEN, A., BAUMBERGER, T., FLESLAND, K., FONSECA, R., OVREAS, L., STEEN, I. H., THORSETH, I. H., PEDERSEN, R. B. & SCHLEPER, C. 2012. Correlating microbial community profiles with geochemical data in highly stratified sediments from the Arctic Mid-Ocean Ridge. *Proc Natl Acad Sci U S A. ,* 109**,** E2846-E2855.

JORGENSEN, S. L., THORSETH, I. H., PEDERSEN, R. B., BAUMBERGER, T. & SCHLEPER, C. 2013. Quantitative and phylogenetic study of the Deep Sea Archaeal Group in sediments of the Arctic mid-ocean spreading ridge. *Front Microbiol,* 4**,** 299.

JUCK, D. F., WHISSELL, G., STEVEN, B., POLLARD, W., MCKAY, C. P., GREER, C. W. & WHYTE, L. G. 2005. Utilization of fluorescent microspheres and a green fluorescent protein-marked strain

for assessment of microbiological contamination of permafrost and ground ice core samples from the Canadian High Arctic. *Appl Environ Microbiol,* 71**,** 1035-1041.
KANEKO, R., HAYASHI, T., TANAHASHI, M. & NAGANUMA, T. 2007. Phylogenetic diversity and distribution of dissimilatory sulfite reductase genes from deep-sea sediment cores. *Mar Biotechnol (NY),* 9**,** 429-36.
KANTOR, R. S., WRIGHTON, K. C., HANDLEY, K. M., SHARON, I., HUG, L. A., CASTELLE, C. J., THOMAS, B. C. & BANFIELD, J. F. 2013. Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *MBio,* 4**,** e00708-13.
KATO, T., HIROTA, M., TANG, Y. H., CUI, X. Y., LI, Y. N., ZHAO, X. Q. & OIKAWA, T. 2005. Strong temperature dependence and no moss photosynthesis in winter CO_2 flux for a Kobresia meadow on the Qinghai-Tibetan plateau. *Soil Biology & Biochemistry,* 37**,** 1966-1969.
KEMPF, B. & BREMER, E. 1998. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Archives of Microbiology,* 170**,** 319-330.
KNITTEL, K., LOSEKANN, T., BOETIUS, A., KORT, R. & AMANN, R. 2005. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl Environ Microbiol,* 71**,** 467-79.
KOVACIK, W. P., JR., TAKAI, K., MORMILE, M. R., MCKINLEY, J. P., BROCKMAN, F. J., FREDRICKSON, J. K. & HOLBEN, W. E. 2006. Molecular analysis of deep subsurface Cretaceous rock indicates abundant Fe(III)- and S(zero)-reducing bacteria in a sulfate-rich environment. *Environ Microbiol,* 8**,** 141-55.
KREADER, C. A. 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Appl Environ Microbiol,* 62**,** 1102-1106.
KRUMHOLZ, L. R., MCKINLEY, J. P., ULRICH, F. A. & SUFLITA, J. M. 1997. Confined subsurface microbial communities in Cretaceous rock. *Nature,* 386**,** 64-66.
KRÜGER, M., WEST, J., FRERICHS, J., OPPERMANN, B., DICTOR, M.-C., JOULIAND, C., JONES, D., COOMBS, P., GREEN, K., PEARCE, J., MAY, F. & MÖLLER, I. 2009. Ecosystem effects of elevated CO2 concentrations on microbial populations at a terrestrial CO_2 vent at Laacher See, Germany. *Energy Procedia,* 1**,** 1933-1939.
KUBO, K., LLOYD, K. G., J, F. B., AMANN, R., TESKE, A. & KNITTEL, K. 2012. Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *ISME J,* 6**,** 1949-65.
LANZEN, A., JORGENSEN, S. L., BENGTSSON, M. M., JONASSEN, I., OVREAS, L. & URICH, T. 2011. Exploring the composition and diversity of microbial communities at the Jan Mayen hydrothermal vent field using RNA and DNA. *FEMS Microbiol Ecol,* 77**,** 577-89.
LANZEN, A., JORGENSEN, S. L., HUSON, D. H., GORFER, M., GRINDHAUG, S. H., JONASSEN, I., OVREAS, L. & URICH, T. 2012. CREST--classification resources for environmental sequence tags. *PLoS One,* 7**,** e49334.
LAU, E., FISHER, M. C., STEUDLER, P. A. & CAVANAUGH, C. M. 2013. The methanol dehydrogenase gene, mxaF, as a functional and phylogenetic marker for proteobacterial methanotrophs in natural environments. *PLoS One,* 8**,** e56993.
LEY, R. E., HARRIS, J. K., WILCOX, J., SPEAR, J. R., MILLER, S. R., BEBOUT, B. M., MARESCA, J. A., BRYANT, D. A., SOGIN, M. L. & PACE, N. R. 2006. Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl Environ Microbiol,* 72**,** 3685-3695.
LINDAHL, T. 1993. Instability and decay of the primary structure of DNA. *Nature,* 362**,** 709-715.
LIU, Z., DESANTIS, T. Z., ANDERSEN, G. L. & KNIGHT, R. 2008. Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Research,* 36.
LLOYD, K. G., SCHREIBER, L., PETERSEN, D. G., KJELDSEN, K. U., LEVER, M. A., STEEN, A. D., STEPANAUSKAS, R., RICHTER, M., KLEINDIENST, S., LENK, S., SCHRAMM, A. & JORGENSEN, B. B. 2013. Predominant archaea in marine sediments degrade detrital proteins. *Nature,* 496**,** 215-+.
LOCK, G. S. H. 1990. *The Growth and decay of Ice,* Cambridge, Cambridge University press, p 434.
LUDWIG, W., STRUNK, O., WESTRAM, R., RICHTER, L., MEIER, H., YADHUKUMAR, BUCHNER, A., LAI, T., STEPPI, S., JOBB, G., FORSTER, W., BRETTSKE, I., GERBER, S., GINHART, A. W., GROSS, O.,

GRUMANN, S., HERMANN, S., JOST, R., KONIG, A., LISS, T., LUSSMANN, R., MAY, M., NONHOFF, B., REICHEL, B., STREHLOW, R., STAMATAKIS, A., STUCKMANN, N., VILBIG, A., LENKE, M., LUDWIG, T., BODE, A. & SCHLEIFER, K. H. 2004. ARB: a software environment for sequence data. *Nucleic Acids Research,* 32**,** 1363-1371.

MACKELPRANG, R., WALDROP, M. P., DEANGELIS, K. M., DAVID, M. M., CHAVARRIA, K. L., BLAZEWICZ, S. J., RUBIN, E. M. & JANSSON, J. K. 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature,* 480**,** 368-71.

MARDIS, E. R. 2008. The impact of next-generation sequencing technology on genetics. *Trends Genet,* 24**,** 133-41.

MATLAKOWSKA, R. & SKLODOWSKA, A. 2009. The culturable bacteria isolated from organic-rich black shale potentially useful in biometallurgical procedures. *J Appl Microbiol,* 107**,** 858-66.

MESLE, M., PERIOT, C., DROMART, G. & OGER, P. 2013. Biostimulation to identify microbial communities involved in methane generation in shallow, kerogen-rich shales. *J Appl Microbiol,* 114**,** 55-70.

MOROZOVA, D., WANDREY, M., ALAWI, M., ZIMMER, M., VIETH, A., ZETTLITZER, M. & WÜRDEMANN, H. 2010. Monitoring of the microbial community composition in saline aquifers during CO_2 storage by fluorescence in situ hybridisation. *International Journal of Greenhouse Gas Control,* 4**,** 981-989.

MOROZOVA, D., ZETTLITZER, M., LET, D. & WÜRDEMANN, H. 2011. Monitoring of the microbial community composition in deep subsurface saline aquifers during CO_2 storage in Ketzin, Germany. *Energy Procedia,* 4**,** 4362-4370.

MUYZER, G., DEWAAL, E. C. & UITTERLINDEN, A. G. 1993. Profiling of Complex Microbial Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain-Reaction-Amplified Genes-Coding for 16S Ribosomal RNA. *Appl Environ Microbiol,* 59**,** 283-290.

MYKYTCZUK, N. C. S., FOOTE, S. J., OMELON, C. R., SOUTHAM, G., GREER, C. W. & WHYTE, L. G. 2013. Bacterial growth at-15 degrees C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1. *ISME J,* 7**,** 1211-1226.

NOBLE, R. R. P., STALKER, L., WAKELIN, S. A., PEJCIC, B., LEYBOURNE, M. I., HORTLE, A. L. & MICHAEL, K. 2012. Biological monitoring for carbon capture and storage – A review and potential future developments. *International Journal of Greenhouse Gas Control,* 10**,** 520-535.

NUNOURA, T., OIDA, H., MIYAZAKI, J., MIYASHITA, A., IMACHI, H. & TAKAI, K. 2008. Quantification of mcrA by fluorescent PCR in methanogenic and methanotrophic microbial communities. *FEMS Microbiol Ecol,* 64**,** 240-7.

OPPERMANN, B. I., MICHAELIS, W., BLUMENBERG, M., FRERICHS, J., SCHULZ, H. M., SCHIPPERS, A., BEAUBIEN, S. E. & KRÜGER, M. 2010. Soil microbial community changes as a result of long-term exposure to a natural CO_2 vent. *Geochimica et Cosmochimica Acta,* 74**,** 2697-2716.

OREN, A. 2006. The Order Halobacteriales. *In:* DWORKIN, M., FALKOW, S., ROSENBERG, E., SCHLEIFER, K.-H. & STACKEBRANDT, E. (eds.) *The Prokaryotes.* New York: Springer 113-164.

PARKER, J. R. 1967. The Jurassic and Cretaceous Sequence in Spitsbergen. *Geological Magazine,* 104**,** 487-505.

PAUL, E. A. & CLARK, F. E. 2007. *Soil microbiology, ecology, and biochemistry, third edition,* Burlington, USA, Academic Press Elsevier.

PERMAFROST OBSERVATORY PROJECT. *A Contribution to the Thermal State of Permafrost in Norway and Svalbard (TSP Norway)* [Online]. Trondheim, Norway: The Norwegian Permafrost Database, Geological Survey of Norway (NGU). Available: http://geo.ngu.no/kart/permafrost_svalbard/?lang=English [Accessed 14.04 2014].

PERMAFROST OBSERVATORY PROJECT. 2012. *A Contribution to the Thermal State of Permafrost in Norway and Svalbard (TSP Norway), Old Auroral Station 2, AS-B-2* [Online]. Trondheim, Norway: The Norwegian Permafrost Database, Geological Survey of Norway (NGU). Available: http://aps.ngu.no/pls/oradb/minres_bo_fakta.boho?p_id=126&p_spraak=E [Accessed 14.04 2014].

PILSON, M. E. Q. 2013. *An introduction to the chemistry of the sea: Michael E.Q. Pilson,* Cambridge, Cambridge University Press.
PODOSOKORSKAYA, O. A., BONCH-OSMOLOVSKAYA, E. A., NOVIKOV, A. A., KOLGANOVA, T. V. & KUBLANOV, I. V. 2013. Ornatilinea apprima gen. nov., sp. nov., a cellulolytic representative of the class Anaerolineae. *Int J Syst Evol Microbiol,* 63**,** 86-92.
POLZ, M. F. & CAVANAUGH, C. M. 1998. Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol,* 64**,** 3724-3730.
PONDER, M., VISHNIVETSKAYA, T., MCGRATH, J. & TIEDJE, J. 2004. Microbial life in permafrost: extended times in extreme conditions. *In:* FULLER, J., LANE, N. & BENSON, E. (eds.) *Life in the frozen state.* Boca Raton, Florida: CRC Press**,** 151–170.
QUADRELLI, R. & PETERSON, S. 2007. The energy-climate challenge: Recent trends in CO_2 emissions from fuel combustion. *Energy Policy,* 35**,** 5938-5952.
QUINCE, C., LANZEN, A., DAVENPORT, R. J. & TURNBAUGH, P. J. 2011. Removing noise from pyrosequenced amplicons. *BMC Bioinformatics,* 12**,** 38.
REYSENBACH, A. L., LIU, Y., BANTA, A. B., BEVERIDGE, T. J., KIRSHTEIN, J. D., SCHOUTEN, S., TIVEY, M. K., VON DAMM, K. L. & VOYTEK, M. A. 2006. A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature,* 442**,** 444-7.
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.
RIVKINA, E., LAURINAVICHIUS, K., MCGRATH, J., TIEDJE, J., SHCHERBAKOVA, V. & GILICHINSKY, D. 2004. Microbial life in permafrost. *Advances in Space Research,* 33**,** 1215-1221.
RIVKINA, E., SHCHERBAKOVA, V., LAURINAVICHIUS, K., PETROVSKAYA, L., KRIVUSHIN, K., KRAEV, G., PECHERITSINA, S. & GILICHINSKY, D. 2007. Biogeochemistry of methane and methanogenic archaea in permafrost. *FEMS Microbiol Ecol,* 61**,** 1-15.
RIVKINA, E. M., FRIEDMANN, E. I., MCKAY, C. P. & GILICHINSKY, D. A. 2000. Metabolic Activity of Permafrost Bacteria below the Freezing Point. *Appl Environ Microbiol,* 66**,** 3230-3233.
RIVKINA, E. M., LAURINAVICHUS, K. S., GILICHINSKII, D. A. & SHCHERBAKOVA, V. A. 2002. Methane generation in permafrost. *Doklady Akademii Nauk,* 383**,** 830-833.
ROBERTSON, C. E., SPEAR, J. R., HARRIS, J. K. & PACE, N. R. 2009. Diversity and Stratification of Archaea in a Hypersaline Microbial Mat. *Appl Environ Microbiol,* 75**,** 1801-1810.
ROESCH, L. F., FULTHORPE, R. R., RIVA, A., CASELLA, G., HADWIN, A. K., KENT, A. D., DAROUB, S. H., CAMARGO, F. A., FARMERIE, W. G. & TRIPLETT, E. W. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J,* 1**,** 283-90.
SCHMALENBERGER, A., SCHWIEGER, F. & TEBBE, C. C. 2001. Effect of primers hybridizing to Different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Appl Environ Microbiol,* 67**,** 3557-3563.
SHI, T., REEVES, R. H., GILICHINSKY, D. A. & FRIEDMANN, E. I. 1997. Characterization of viable bacteria from Siberian permafrost by 16S rDNA sequencing. *Microbial Ecology,* 33**,** 169-179.
SIPOS, R., SZEKELY, A. J., PALATINSZKY, M., REVESZ, S., MARIALIGETI, K. & NIKOLAUSZ, M. 2007. Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiol Ecol,* 60**,** 341-350.
SMITH, C. J., NEDWELL, D. B., DONG, L. F. & OSBORN, A. M. 2006. Evaluation of quantitative polymerase chain reaction-based approaches for determining gene copy and gene transcript numbers in environmental samples. *Environ Microbiol,* 8**,** 804-15.
STEVEN, B., BRIGGS, G., MCKAY, C. P., POLLARD, W. H., GREER, C. W. & WHYTE, L. G. 2007. Characterization of the microbial diversity in a permafrost sample from the Canadian high Arctic using culture-dependent and culture-independent methods. *FEMS Microbiol Ecol,* 59**,** 513-23.
STEVEN, B., LEVEILLE, R., POLLARD, W. H. & WHYTE, L. G. 2006. Microbial ecology and biodiversity in permafrost. *Extremophiles,* 10**,** 259-67.

STEVEN, B., POLLARD, W. H., GREER, C. W. & WHYTE, L. G. 2008. Microbial diversity and activity through a permafrost/ground ice core profile from the Canadian high Arctic. *Environ Microbiol,* 10**,** 3388-403.

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.

TAS, N., PRESTAT, E., MCFARLAND, J. W., WICKLAND, K. P., KNIGHT, R., BERHE, A. A., JORGENSON, T., WALDROP, M. P. & JANSSON, J. K. 2014. Impact of fire on active layer and permafrost microbial communities and metagenomes in an upland Alaskan boreal forest. *ISME J*.

TESKE, A. & SORENSEN, K. B. 2008. Uncultured archaea in deep marine subsurface sediments: have we caught them all? *ISME J,* 2**,** 3-18.

THOMAS 1977. Historical Developments in Soil Chemistry: Ion Exchange. *Soil Science Society of America journal,* 41**,** 230.

TREUSCH, A. H., LEININGER, S., KLETZIN, A., SCHUSTER, S. C., KLENK, H. P. & SCHLEPER, C. 2005. Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol,* 7**,** 1985-95.

TUORTO, S. J., DARIAS, P., MCGUINNESS, L. R., PANIKOV, N., ZHANG, T., HAEGGBLOM, M. M. & KERKHOF, L. J. 2014. Bacterial genome replication at subzero temperatures in permafrost. *ISME J,* 8**,** 139-149.

TVEIT, A., SCHWACKE, R., SVENNING, M. M. & URICH, T. 2013. Organic carbon transformations in high-Arctic peat soils: key functions and microorganisms. *ISME J,* 7**,** 299-311.

VETROVSKY, T. & BALDRIAN, P. 2013. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One,* 8**,** e57923.

VISHNIVETSKAYA, T., KATHARIOU, S., MCGRATH, J., GILICHINSKY, D. & TIEDJE, J. M. 2000. Low-temperature recovery strategies for the isolation of bacteria from ancient permafrost sediments. *Extremophiles,* 4**,** 165-173.

VISHNIVETSKAYA, T. A., LAYTON, A. C., LAU, M. C., CHAUHAN, A., CHENG, K. R., MEYERS, A. J., MURPHY, J. R., ROGERS, A. W., SAARUNYA, G. S., WILLIAMS, D. E., PFIFFNER, S. M., BIGGERSTAFF, J. P., STACKHOUSE, B. T., PHELPS, T. J., WHYTE, L., SAYLER, G. S. & ONSTOTT, T. C. 2014. Commercial DNA extraction kits impact observed microbial community composition in permafrost samples. *FEMS Microbiol Ecol,* 87**,** 217-30.

VISHNIVETSKAYA, T. A., PETROVA, M. A., URBANCE, J., PONDER, M., MOYER, C. L., GILICHINSKY, D. A. & TIEDJE, J. M. 2006. Bacterial community in ancient Siberian permafrost as characterized by culture and culture-independent methods. *Astrobiology,* 6**,** 400-414.

VOROBYOVA, E., SOINA, V., GORLENKO, M., MINKOVSKAYA, N., ZALINOVA, N., MAMUKELASHVILI, A., GILICHINSKY, D., RIVKINA, E. & VISHNIVETSKAYA, T. 1997. The deep cold biosphere: facts and hypothesis. *FEMS Microbiol Reviews,* 20**,** 277-290.

WAGNER, M., ROGER, A. J., FLAX, J. L., BRUSSEAU, G. A. & STAHL, D. A. 1998. Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J. Bacteriol,* 180**,** 2975-2982.

WARTIAINEN, I., HESTNES, A. G., MCDONALD, I. R. & SVENNING, M. M. 2006. Methylobacter tundripaludum sp nov., a methane-oxidizing bacterium from Arctic wetland soil on the Svalbard islands, Norway (78 degrees N). *International Journal of Systematic and Evolutionary Microbiology,* 56**,** 109-113.

WARTIAINEN, I., HESTNES, A. G. & SVENNING, M. M. 2003. Methanotrophic diversity in high arctic wetlands on the islands of svalbard (Norway) - denaturing gradient gel electrophoresis analysis of soil DNA and enrichment cultures. *Canadian Journal of Microbiology,* 49**,** 602-612.

WILHELM, R. C., NIEDERBERGER, T. D., GREER, C. & WHYTE, L. G. 2011. Microbial diversity of active layer and permafrost in an acidic wetland from the Canadian High Arctic. *Can J Microbiol,* 57**,** 303-15.

WILHELM, R. C., RADTKE, K. J., MYKYTCZUK, N. C., GREER, C. W. & WHYTE, L. G. 2012. Life at the wedge: the activity and diversity of arctic ice wedge microbial communities. *Astrobiology,* 12**,** 347-60.

WILLERSLEV, E., HANSEN, A. J. & POINAR, H. N. 2004. Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol Evol,* 19**,** 141-7.

WILLIAMS, P. J. & SMITH, M. W. 1989. *The Frozen Earth: Fundamentals of Geocryology,* Cambridge, Cambridge University Press, 306 p.

WU, J.-H., HONG, P.-Y. & LIU, W.-T. 2009. Quantitative effects of position and type of single mismatch on single base primer extension. *Journal of Microbiological Methods,* 77**,** 267-275.

YAMADA, T., SEKIGUCHI, Y., HANADA, S., IMACHI, H., OHASHI, A., HARADA, H. & KAMAGATA, Y. 2006. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the bacterial phylum Chloroflexi. *Int J Syst Evol Microbiol,* 56**,** 1331-40.

YERGEAU, E., HOGUES, H., WHYTE, L. G. & GREER, C. W. 2010. The functional potential of high Arctic permafrost revealed by metagenomic sequencing, qPCR and microarray analyses. *ISME J,* 4**,** 1206-1214.

ZHANG, G., DONG, H., XU, Z., ZHAO, D. & ZHANG, C. 2005. Microbial diversity in ultra-high-pressure rocks and fluids from the Chinese Continental Scientific Drilling Project in China. *Appl Environ Microbiol,* 71**,** 3213-27.

Appendix A: Grain size distribution curves

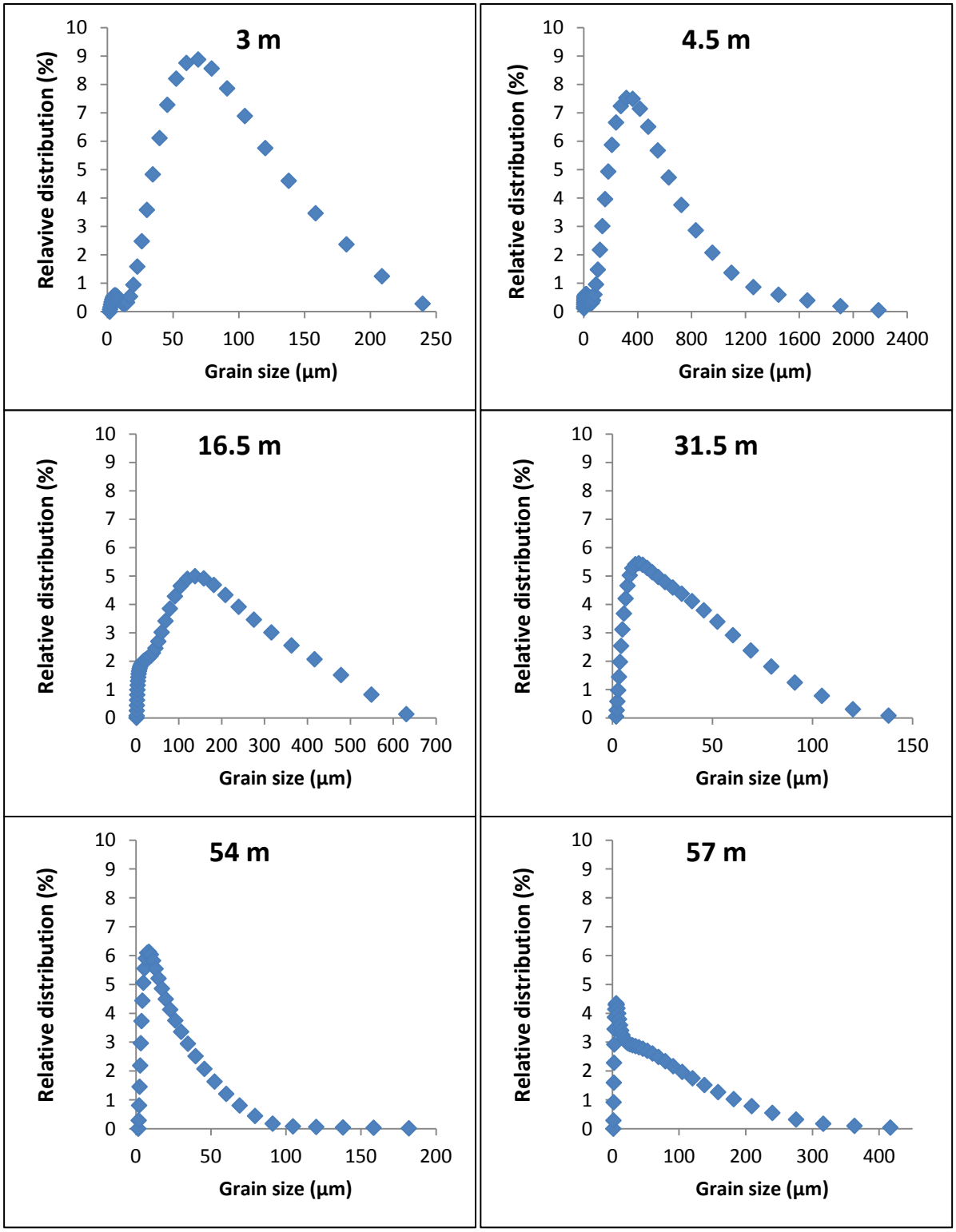

Figure A.1: Relative distribution of grain sizes at six sedimentary permafrost depths.

Appendix B: Ion concentratons

Table B.1: Ion concentrations from pore water extracted with H_2O. The % of the ions that were extracted with H_2O is relative to the total ion concentration ($CaCl_2$ extracted ions + H_2O extracted ions).

Depth (m)	Fe^{2+} (µM)	% Fe extracted with H_2O	Mn^{2+} (µM)	% Mn extracted with H_2O	Mg^{2+} (mM)	% Mg extracted with H_2O	Sr^{2+} (mM)	% Sr extracted with H_2O	K^{+} (mM)	% K extracted with H_2O	Na^{+} (mM)	% Na extracted with H_2O	Si^{4+} (mM)	% Si extracted with H_2O
3	33,9	3,4	13,4	2,4	1,3	3,3	0,0	2,0	6,4	32,9	81,8	74,8	0,9	58,8
3	75,7	3,9	21,3	3,4	1,5	4,4	0,0	2,6	6,5	36,4	68,5	76,7	0,8	60,0
4,5	15,7	70,0	BD	0,0	1,8	4,1	0,0	2,1	16,4	36,5	345,3	77,4	0,5	45,4
4,5	29,0	80,9	BD	0,0	1,9	4,0	0,0	2,2	16,6	35,6	346,8	76,3	0,6	45,8
16,5	BD	0,0	1,1	7,4	8,8	15,7	0,0	9,0	21,4	52,3	440,4	86,5	0,5	45,6
16,5	BD	0,0	1,7	11,2	19,0	25,6	0,1	15,6	30,2	58,1	639,7	87,6	0,6	45,7
31,5	BD	0,0	BD	0,0	3,1	6,3	0,0	3,5	21,6	40,5	498,3	78,0	0,2	34,6
31,5	BD	0,0	BD	0,0	3,0	6,2	0,0	3,4	21,7	41,2	483,5	78,0	0,3	35,0
54	54,4	89,5	BD	0,0	2,3	5,1	0,0	3,4	19,9	40,2	380,2	76,2	0,6	54,8
54	41,4	85,5	BD	0,0	2,6	5,4	0,0	3,7	21,1	39,9	408,1	76,1	0,5	55,1
57	27,5	100,0	BD	0,0	3,4	4,3	0,0	3,6	24,3	38,9	499,4	79,4	1,1	48,8
57	17,0	100,0	BD	0,0	2,4	3,9	0,0	3,2	18,3	37,4	404,4	79,4	1,1	51,0
60	3094,0	100,0	BD	0,0	2,3	0,8	0,0	0,1	52,7	12,9	998,5	54,4	68,7	72,1
60	2721,9	100,0	BD	0,0	2,0	0,5	0,0	0,0	67,7	12,5	1338,3	54,4	72,2	67,1

BD: Below detection limit

Table B.2: Ion concentrations from pore water extracted with $CaCl_2$. The % of the ions that were extracted with $CaCl_2$ is relative to the total ion concentration ($CaCl_2$ extracted ions + H_2O extracted ions).

Depth (m)	Fe^{2+} (µM)	% Fe extracted with $CaCl_2$	Mn^{2+} (µM)	% Mn extracted with $CaCl_2$	Mg^{2+} (mM)	% Mg extracted with $CaCl_2$	Sr^{2+} (mM)	% Sr extracted with $CaCl_2$	K^+ (mM)	% K extracted with $CaCl_2$	Na^+ (mM)	% Na extracted with $CaCl_2$	Si4+ (mM)	% Si extracted with $CaCl_2$
3	969,6	96,6	540,4	97,6	38,8	96,7	0,2	98,0	13,1	67,1	27,6	25,2	0,6	41,2
3	1861,0	96,1	614,8	96,6	32,6	95,6	0,2	97,4	11,4	63,6	20,8	23,3	0,5	40,0
4,5	6,7	30,0	7,1	95,1	42,2	95,9	0,2	97,9	28,5	63,5	101,1	22,6	0,6	54,6
4,5	6,8	19,1	6,8	94,7	44,9	96,0	0,2	97,8	30,0	64,4	107,7	23,7	0,7	54,2
16,5	9,6	100,0	13,3	92,6	47,5	84,3	0,3	91,0	19,5	47,7	68,5	13,5	0,6	54,4
16,5	10,7	100,0	13,7	88,8	55,2	74,4	0,3	84,4	21,7	41,9	90,9	12,4	0,7	54,3
31,5	8,3	100,0	5,8	94,9	45,8	93,7	0,3	96,5	31,7	59,5	140,2	22,0	0,4	65,4
31,5	7,9	100,0	5,4	93,5	45,9	93,8	0,3	96,6	31,0	58,8	136,4	22,0	0,5	65,0
54	6,4	10,5	10,9	93,7	42,9	94,9	0,3	96,6	29,7	59,8	118,6	23,8	0,5	45,2
54	7,0	14,5	10,5	92,8	44,4	94,6	0,3	96,3	31,7	60,1	128,2	23,9	0,4	44,9
57	BD	0,0	40,0	97,9	74,9	95,7	1,1	96,4	38,2	61,1	129,4	20,6	1,2	51,2
57	BD	0,0	28,0	97,5	60,3	96,1	0,9	96,8	30,8	62,6	104,7	20,6	1,0	49,0
60	BD	0,0	188,4	94,4	307,4	99,2	9,3	99,9	356,2	87,1	838,1	45,6	26,6	27,9
60	BD	0,0	251,2	95,7	410,8	99,5	12,4	100,0	474,8	87,5	1120,6	45,6	35,5	32,9

BD: Below detection limit

Table B.3: Ion concentrations from extracted pore water calculated for dry weight sediment. Fe^{2+}, Mn^{2+}, Na^{+}, K^{+}, Mg^{2+}, Si^{4+} and Sr^{2+} were extracted with H_2O and $CaCl_2$, and the concentration shown here is the sum of the two extractions. The remaining ions were extracted with H_2O only.

Depth (m)	Ca^{2+} (µg/g dry weight)	Fe^{2+} (µg/g dry weight)	K^{+} (µg/g dry weight)	Mg^{2+} (µg/g dry weight)	Mn^{2+} (µg/g dry weight)	Na^{+} (µg/g dry weight)	Si^{4+} (µg/g dry weight)	Sr^{2+} (µg/g dry weight)	Cl^{-} (µg/g dry weight)	Br^{-} (µg/g dry weight)	SO_4^{2-} (µg/g dry weight)	NaCl (g/L)	NaCl (g/L) when 3% of water is liquid
3	6,2	14,4	196,8	250,9	7,8	647,8	10,8	5,3	696,8	2,4	220,6	4,8	159,3
3	9,0	30,3	195,7	231,9	9,8	575,3	10,7	5,4	657,7	2,2	221,0	4,0	133,5
4,5	3,7	0,3	481,6	292,9	0,1	2812,1	8,7	4,0	3190,7	10,2	414,9	20,2	672,7
4,5	3,8	0,5	481,0	299,5	0,1	2754,1	9,4	4,3	3039,9	9,6	407,5	20,3	675,5
16,45	17,0	0,1	442,5	378,8	0,2	3237,4	8,9	8,0	4282,8	13,8	1270,1	25,7	858,0
16,45	34,5	0,2	538,0	477,8	0,2	4452,5	9,1	9,6	6195,9	19,0	1628,3	37,4	1246,2
31,5	4,8	0,1	583,3	333,0	0,1	4109,4	5,4	8,4	5094,6	18,2	163,0	29,1	970,8
31,5	4,7	0,1	586,9	338,0	0,1	4056,5	5,9	8,3	5016,3	18,7	153,1	28,3	942,0
54	11,2	1,0	569,4	322,2	0,2	3366,4	8,5	7,9	3517,2	12,7	31,1	22,2	740,6
54	12,7	0,7	565,1	312,6	0,2	3374,0	7,4	7,7	3562,4	12,3	34,5	23,8	794,9
57	6,8	0,2	318,1	248,0	0,3	1884,5	8,6	13,2	2983,6	7,8	61,5	29,2	972,9
57	6,3	0,2	318,9	253,5	0,3	1944,9	9,6	13,7	2344,6	7,9	59,2	23,6	787,8
60	0,2	3,3	308,7	145,4	0,2	815,3	51,7	15,7	190,6	0,0	6,7	58,4	1945,1
60	0,3	2,2	307,4	145,4	0,2	819,3	43,8	15,8	191,2	0,0	8,3	78,2	2607,0

Appendix C: qPCR data

Table C.1: Quantitative PCR (qPCR) data with standard deviations. Each sample was run in triplicates in qPCR analysis, and the standard deviations are calculated from the qPCR triplets.

Depth	16S rRNA copies g^{-1} BACTERIA	Standard deviation from qPCR triplets	16S rRNA copies g^{-1} ARCHAEA	Standard deviation from qPCR triplets	16S rRNA copies g^{-1} DSAG	Standard deviation from qPCR triplets	*dsrB* copies g^{-1}	Standard deviation from qPCR triplets	*mcrA* copies g^{-1}	Standard deviation from qPCR triplets
3	3,0E+08	2,7E+07	2,5E+06	1,9E+05	BD	0	5,2E+06	7,1E+05	4164,4	161,1
3	5,8E+08	5,4E+07	3,0E+06	5,5E+04	NA	NA	7,6E+06	1,0E+06	6539,7	1335,4
4,5	7,6E+06	9,6E+05	4,0E+05	7,3E+04	BD	0	1,4E+06	1,8E+05	273,0	53,1
4,5	4,3E+06	1,9E+05	3,9E+05	1,4E+05	NA	NA	8,1E+05	3,6E+05	211,7	74,8
16,5	8,2E+07	4,8E+06	4,4E+06	9,0E+04	8,2E+06	7,3E+05	1,2E+07	2,0E+06	195,8	49,1
16,5	4,5E+07	4,5E+06			NA	NA	8,3E+06	8,6E+05	356,4	104,6
31,5	3,7E+07	4,7E+06	4,1E+06	1,6E+05	1,6E+06	4,4E+05	2,4E+06	2,5E+05	586,1	
31,5	2,4E+07	3,0E+06	2,9E+06	5,0E+05	NA	NA	1,9E+06	1,0E+06	383,8	39,0
54	2,9E+07	1,3E+06	1,5E+07	2,7E+06	1,4E+07	6,6E+06	3,4E+06	1,1E+06	94,4	2,2
54	1,6E+07	3,3E+06	1,4E+07	3,2E+06	NA	NA	3,1E+06	5,0E+05	BD	0
57	3,8E+06	1,1E+06	5,2E+06	3,6E+05	4,7E+06	3,7E+05	2,3E+06	1,6E+06	408,5	
57	4,7E+06	1,4E+06	4,8E+06	7,7E+05	NA	NA	3,4E+06	2,6E+05	122,5	31,9
60	2,4E+05	4,4E+04	2,5E+05	8,8E+04	BD	0	BD	0	BD	0

NA: Not analysed

BD: Below detection limit

Appendix D: *dsrB* phylogenetic tree

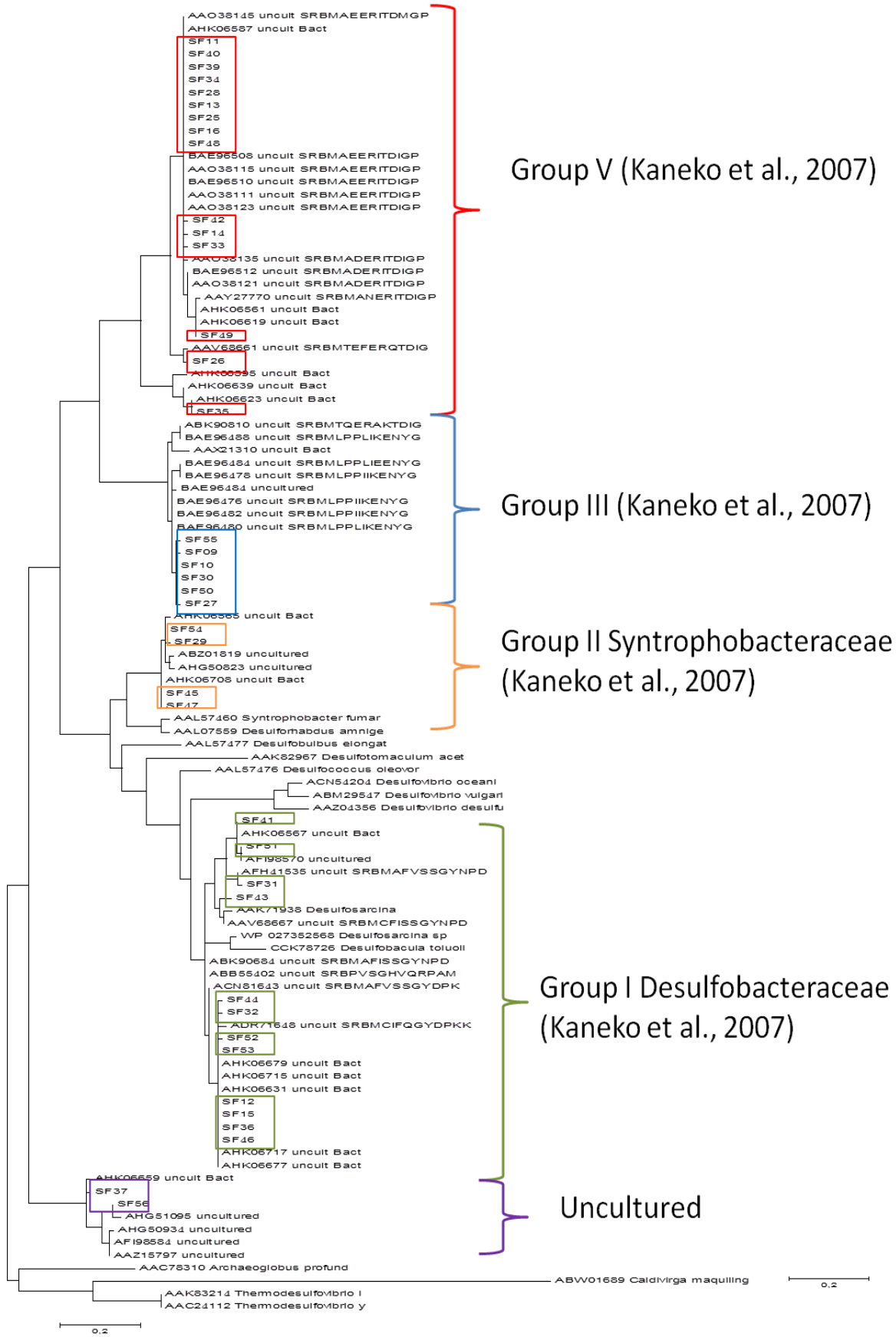

Figure D.1: Phylogenetic tree based on *dsrB* amino acids from the 54 m sample of the DH8 permafrost core. The tree was constructed using the maximum-likelyhood algorithms in MEGA 6. The *dsrB* amino acid sequences of the genus Thermodesulfovibrio (AAK83214, AAC241122) and Archaeoglobus (AAC78310) were used as outgroups. The clones from this study are marked with squares, and related to the phylogenetic tree from Kaneko et al. (2007).

Appendix E: Statistics

Table E.1: pH correlated with Total Organic Carbon (TOC), Calsium oxide (CaO), Magnesium oxide (MgO) and iron oxide (Fe_2O_3). Correlations marked in yellow are significant ($p < 0.05$).

		pH	wt% TOC	CaO	MgO	Fe2O3
Pearson Correlation	pH	1.000	-.132	.935	.849	.249
	wt% TOC	-.132	1.000	.179	-.332	-.807
	CaO	.935	.179	1.000	.646	-.094
	MgO	.849	-.332	.646	1.000	.580
	Fe2O3	.249	-.807	-.094	.580	1.000
Sig. (1-tailed)	pH	.	.401	.003	.016	.317
	wt% TOC	.401	.	.367	.260	.026
	CaO	.003	.367	.	.083	.430
	MgO	.016	.260	.083	.	.114
	Fe2O3	.317	.026	.430	.114	.

Model 1: CaO

Model 2: CaO + Fe2O3

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics			
					R Square Change	F Change	df1	df2
1	.935[a]	.874	.842	.270261	.874	27.677	1	4
2	.994[b]	.988	.980	.095874	.114	28.785	1	3
3	.996[c]	.993	.981	.092720	.004	1.208	1	2
4	.999[d]	.998	.991	.065128	.006	3.054	1	1

Model	Change Statistics	Durbin-Watson
	Sig. F Change	
1	.006	
2	.013	
3	.386	
4	.331	3.108

Table E.2: 16S rRNA Bacteria correlated with dissolved nutrients and Total Organic Carbon (TOC). K^+ and Mg^{2+} have been extracted with H_2O and $CaCl_2$. The remaining ions were extracted with H_2O only. Ion concentration data from 60 m were removed in the correlation analysis. Correlations marked in yellow are significant ($p < 0.05$).

		16S rRNA copies g-1 BACTERIA	K mM H2O + CaCl2	Mg mM H2O + CaCl2	NH4 mM H20	PO4 µM H20
Pearson Correlation	16S rRNA copies g-1 BACTERIA	1.000	-.881	-.567	-.432	-.453
	K mM H2O + CaCl2	-.881	1.000	.744	.585	.664
	Mg mM H2O + CaCl2	-.567	.744	1.000	.539	.751
	NH4 mM H20	-.432	.585	.539	1.000	.200
	PO4 µM H20	-.453	.664	.751	.200	1.000
	NO3 mM H20	-.108	.290	.689	-.039	.832
	wt% TOC	.917	-.755	-.369	-.127	-.357
Sig. (1-tailed)	16S rRNA copies g-1 BACTERIA	.	.002	.072	.143	.130
	K mM H2O + CaCl2	.002	.	.017	.064	.036
	Mg mM H2O + CaCl2	.072	.017	.	.084	.016
	NH4 mM H20	.143	.064	.084	.	.318
	PO4 µM H20	.130	.036	.016	.318	.
	NO3 mM H20	.399	.243	.029	.463	.005
	wt% TOC	.001	.015	.184	.382	.192

Table E.3: 16S rRNA Archaea correlated with nutrients and Total Organic Carbon (TOC). K^+ and Mg^{2+} have been extracted with H_2O and $CaCl_2$. Ion concentration data from 60 m were removed in the correlation analysis. The remaining ions were extracted with H_2O only.

		16S rRNA copies g-1 ARCHAEA	K mM H2O + CaCl2	Mg mM H2O + CaCl2	NH4 mM H20	PO4 µM H20
Pearson Correlation	16S rRNA copies g-1 ARCHAEA	1.000	.274	.112	.542	.385
	K mM H2O + CaCl2	.274	1.000	.744	.585	.664
	Mg mM H2O + CaCl2	.112	.744	1.000	.539	.751
	NH4 mM H20	.542	.585	.539	1.000	.200
	PO4 µM H20	.385	.664	.751	.200	1.000
	NO3 mM H20	.033	.290	.689	-.039	.832
	wt% TOC	.087	-.755	-.369	-.127	-.357
Sig. (1-tailed)	16S rRNA copies g-1 ARCHAEA	.	.256	.396	.083	.173
	K mM H2O + CaCl2	.256	.	.017	.064	.036
	Mg mM H2O + CaCl2	.396	.017	.	.084	.016
	NH4 mM H20	.083	.064	.084	.	.318
	PO4 µM H20	.173	.036	.016	.318	.
	NO3 mM H20	.469	.243	.029	.463	.005
	wt% TOC	.419	.015	.184	.382	.192

Table E.4: 16S rRNA Archaea correlated with potential electron acceptors. Fe^{2+} and Mn^{2+} have been extracted with H_2O and $CaCl_2$. The remaining ions were extracted with H_2O only. Ion concentration data from 60 m were removed in the correlation analysis.

		16S rRNA copies g-1 ARCHAEA	NO3 mM H20	SO42- mM H20	Fe µM H2O + CaCl2	Mn µM H2O + CaCl2
Pearson Correlation	16S rRNA copies g-1 ARCHAEA	1.000	-.272	-.284	-.350	-.255
	NO3 mM H20	-.272	1.000	-.257	.771	.093
	SO42- mM H20	-.284	-.257	1.000	-.189	-.107
	Fe µM H2O + CaCl2	-.350	.771	-.189	1.000	.640
	Mn µM H2O + CaCl2	-.255	.093	-.107	.640	1.000
Sig. (1-tailed)	16S rRNA copies g-1 ARCHAEA	.	.196	.185	.132	.212
	NO3 mM H20	.196	.	.210	.002	.386
	SO42- mM H20	.185	.210	.	.279	.370
	Fe µM H2O + CaCl2	.132	.002	.279	.	.013
	Mn µM H2O + CaCl2	.212	.386	.370	.013	.

Table E.5: 16S rRNA Bacteria copies per gram correlated with dissolved potential electron acceptors. Fe^{2+} and Mn^{2+} have been extracted with H_2O and $CaCl_2$. The remaining ions were extracted with H_2O only. Ion concentration data from 60 m were removed in the correlation analysis. Correlations marked in yellow are significant ($p < 0.05$).

		16S rRNA copies g-1 BACTERIA	NO3 mM H20	SO42- mM H20	Fe µM H2O + CaCl2	Mn µM H2O + CaCl2
Pearson Correlation	16S rRNA copies g-1 BACTERIA	1.000	-.144	-.066	.990	.952
	NO3 mM H20	-.144	1.000	-.444	-.060	.019
	SO42- mM H20	-.066	-.444	1.000	-.165	-.161
	Fe µM H2O + CaCl2	.990	-.060	-.165	1.000	.965
	Mn µM H2O + CaCl2	.952	.019	-.161	.965	1.000
Sig. (1-tailed)	16S rRNA copies g-1 BACTERIA	.	.328	.420	.000	.000
	NO3 mM H20	.328	.	.074	.427	.477
	SO42- mM H20	.420	.074	.	.305	.308
	Fe µM H2O + CaCl2	.000	.427	.305	.	.000
	Mn µM H2O + CaCl2	.000	.477	.308	.000	.