Characterization of bacterial community structure in deep subsurface sedimentary core samples from Michigan Basin, Ontario

Dimitri Ilin

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the MSc degree in the Chemical and Environmental Toxicology Program

Department of Earth Sciences

Faculty of Science

University of Ottawa

Abstract

Deep subsurface rock samples from Upper Ordovician strata in the Michigan Basin were analyzed for the presence of microbial communities. High concentrations of biogenic methane were observed in the Upper and Middle Ordovician formations. Total porosity values for the shale, shale hard bed and limestone samples were 7.4%, 2.5% and 1.9%, respectively. Hydrocarbon presence ranged from petroliferous shale, to bituminous layering in shale hard beds, to hydrocarbon odour in limestone. Organic carbon content ranged from 0.5 to 2.5%, highest amount being present in the shale. Environmental DNA was extracted from core samples and PCR amplified using 16S rDNA bacterial primers. PCR performed with archaeal 16S rDNA and methanogen-specific (*mcrA*) primers did not yield DNA amplification. Gene analysis indicated that bacterial sequences similar to *Proteobacteria*, *Cyanobacteria*, *Firmicutes*, and *Actinobacteria* were present. Most sequences were not related to known cultivated species. *Proteobacteria* was the most dominant phyla at all depths and included heterotrophic, lithotrophic, acidophilic, radiotolerant, and sulphate-reducing species of bacteria. This study concludes that the observed biogenic methane is a product of ancient methanogenesis.

Résumé

Des échantillons de roches souterraines de l'Ordovicien supérieur du bassin du Michigan ont été analysés afin de détecter la présence de communautés microbiennes. Des concentrations élevées en méthane biogénique ont été observées dans les formations Ordoviciennes Haute et Moyenne. Les valeurs de porosité pour le schiste, le schiste dur et les échantillons de calcaire étaient de 7,4%, 2,5% et 1,9% respectivement. Des hydrocarbures étaient présents sous forme de schiste pétrolifère et de bitume en couches dans des lits de schiste dur; de plus, une odeur d'hydrocarbure se dégageait des échantillons calcaire. La teneur en carbone organique variait de 0,5 à 2,5%, la valeur la plus élevée ayant été observée dans le schiste. L'ADN environnemental a été extrait d'échantillons à 4 profondeurs et amplifié par PCR en utilisant des amorces bactériennes ciblant le gène 16SADNr. Les PCR réalisées avec des amorces spécifiques aux archées méthanogèniques ainsi qu'avec des amorces spécifiques à la méthanogénèse (gène *mcrA*) se sont révélées non conclusives et suggèrent l'absence d'une méthanogénèse contemporaine. L'analyse des séquences des gènes 16S indique que la présence d'ADN bactérien similaire à celui des *Protéobactéries*, *Cyanobactéries*, *Firmicutes* et d'*Actinobacteria*. La plupart des séquences ne sont pas liés à des espèces cultivées connues. Les protéobactéries représentent le phylum dominant à toutes les profondeurs et représentent des organismes hétérotrophes, lithotrophes, acidophiles, radiotolérant ainsi que des espèces de bactéries sulfatoréductrices. Cette étude conclut que le méthane biogénique observé est le produit d'une méthanogénèse ancienne.

Acknowledgments

I would like to thank my supervisors, Dr. Ian Clark and Dr. Alexandre Poulain, for their guidance, support and patience throughout the research project. Without their dedication, expertise, and willingness to continually go above and beyond, this thesis work would not have been possible. I also would like to thank the members of my committee, Dr. Danielle Fortin and Dr. Fred Michel, for their valuable comments and support.

I would like to thank my co-researchers at the University of Ottawa for their ongoing support, especially Michelle Brazeau, Philip Pelletier, Hardeep Gill and Emmanuel Yumvihoze. I owe a special thank you to Julien Godding, a dedicated young scientist, for his inspirations and positive outlook. I feel fortunate to have had the full support of everyone at the University during every stage of the research and writing process.

Table of Contents

List of Tables

List of Figures

List of Key Acronyms

bp: base pair

BSA: Bovine serum albumin

DMSO: Dimethyl sulfoxide

DGR: Deep Geologic Repository

DNA: Deoxyribonucleic acid

ETC: electron transport chain

mBGS: meters Below Ground Surface

OTU: Operational taxonomic unit

PCR: Polymerase chain reaction

RFLP: Restriction fragment length polymorphism

rRNA: ribosomal Ribonucleic acid

TOC: total organic carbon

1. General Introduction

Recent studies looking into microbial life in the deep subsurface environments have demonstrated a wide presence of heterotrophic and lithotrophic microorganisms, in what was previously believed inhospitable conditions. These microorganisms have evolved to exist without oxygen and light and are commonly found in anoxic pockets within the deep subsurface. They can use a wide array of organic and inorganic molecules as a source of energy and can employ a range of electron acceptors.

Early microbiological studies on the deep subsurface have been hampered by issues of contamination and inability to fully identify isolated microorganisms. These studies were carried out using culture based methods and primarily focused on comparing colony and cellular morphology with a goal of identifying viability and diversity of microorganism communities. In general, culture based methods are very limited in growth of microorganisms from environmental samples. Advancements in the field of molecular biology allowed for development of non-culture based techniques and more in-depth characterization of these communities based on their DNA composition. Construction of DNA sequence databases facilitated the identification of microorganisms and allowed for identification of biological contaminants. In addition, construction of stringent DNA extraction protocols aimed at eliminating biological contamination significantly increased the validity of obtained results.

DNA extraction and amplification from deep subsurface sediments remains a challenge due to the low biomass and the inhibitory effect of rock matrix. Protocols designed to account for these issues are often unreliable because they work on case-to-case basis. Therefore, successful extraction and amplification of DNA requires significant improvement and rigorous testing of the developed protocols. Although challenging, studies looking at deep subsurface microorganisms are of tremendous benefit to the field of microbiology. These studies contribute to our understanding of the bacterial and archaeal diversity in the subsurface and their survival mechanisms. Specific interest lies in bacteria involved in degradation of petroleum and toxic substances for the use in the clean-up of oil spills and various ground remediation procedures. Heat tolerant bacteria are of interest as well, as they produce thermostable enzymes and antibiotics. There is a lot to gain from studying microbiology of the deep subsurface.

2. Literature Review

2.1 *Previous work on deep subsurface microorganisms*

Until recently, deep subsurface environment has been eluding the focus of scientific research, because of primary scientific interest in near-surface and terrestrial environments (Chandler, 1998). Studies were mainly conducted on freshly deposited or highly permeable sediments, instead of subsurface rock, because microbial life was thought to be supported solely by freshly deposited or migrant organic matter. Recently, novel mechanisms focusing on the survival of microbial life within the deep subsurface rock were proposed. These mechanisms demonstrate that microorganisms can survive in environments with limited nutrient availability and growth space (Krumholz, 2000). Interest in the deep subsurface environment was ignited by a study claiming life in the meteorite from Mars (McKay, 1996). Abundant polycyclic aromatic hydrocarbons and biogenically produced carbonate globules with magnetite and iron sulphide precipitates were found within the meteorite. Additionally, survival of microorganisms during ejection from the surface of Mars was theorized to be unaffected by extreme acceleration conditions and thus found to be possible (Mastrapa, 2001). These findings raised speculation of past or present life on Mars. Due to high solar radiation and cold temperatures currently found on Martian surface, it is likely that most surface terrain is uninhabitable to all forms of life. However, microbes could theoretically survive in the warmer, yet hypersaline subsurface of the planet, where habitable conditions are not as extreme.

Here on Earth, habitable conditions are far more prevalent and life on the surface of the planet is full of diversity. Recently, studies were published claiming high abundance and diversity of microbial life in the deep subsurface environments (Krumholz, 2000). In fact, researchers at the University of Georgia proposed that the majority of global biomass is found beneath, in the deep subsurface (Whitman, 1998). This prediction was based on the estimated global amount of carbon and prokaryotes, and should be treated as an estimate only. However, the chance of finding life, in what was previously thought to be an inhospitable environment, is appropriate. Microorganisms can potentially survive to a depth of 5 to 10 kilometres, survival only being limited on the adaptation to high temperatures (Gold, 1992). At that depth, upper temperature limit for the survival of microorganisms is reached. Distribution of microorganisms could be mediated by growth along fractures and fluid movement though pore spaces. Chemical

energy released through liquefaction of solid materials could support microbial growth via heat exchange.

One of the earlier studies aimed at microbial characterization of deep subsurface was conducted on a 21 m^3 section of rock, obtained from a 400 m deep underground tunnel at the Nevada Test Site (Haldeman, 1993). This work was largely aimed at determining the distribution of microorganisms on a three-dimensional scale within the rock section using culture based techniques. Early microbiological studies have primarily focused on comparing colony and cellular morphology, physiological profiles and antibiotic and metal resistance of the recovered isolates. Identification of isolated organisms however was of a challenge, since microbiological identification systems and databases did not include much information on environmental microorganisms (Haldeman, 1993). Although, identification of microorganisms was not possible, the study provided information of the viability of bacterial communities, high diversity of these communities, and the homogenous distribution of bacteria within the rock section. Another study looked at microbial diversity in aquifers that run through faults or fracture systems in deep granitic rock at 450 to 650 meter depth in Aspo tunnel, Sweden (Pedersen, 1997). In this study, microbial communities from the groundwater were analyzed using culture-based techniques to determine the number of cultivable organisms, and non-culture based techniques were used to assess the diversity of the microbial populations. 16S rRNA genes were sequenced from the cultivated bacteria and phylogenetic analysis demonstrated a rich diversity of bacterial and archaeal species in the groundwater. In addition, various geochemical processes were linked to the discovered bacterial and archaeal species.

In a recent study, drill core samples obtained from Ultra-High-Pressure Dabie-Sulu metamorphic orogenic belt in China, yielded diverse bacterial and archaeal communities (Zhang, 2005). Drill core samples were obtained from the depth of 500 to 2000 meters and were analyzed through 16S rRNA gene sequencing and phylogenetic analysis. In addition the study provided petrophysical and geochemical data of the rock formations including anion, TOC and ^{13}C isotope content. Another study looking at microbial diversity in carbonate rocks from cave environments, overcame complications with high concentration of calcium, and low-biomass estimated to be $2x10^2$ cells/g (Barton, 2006). The authors were able to isolate and amplify 16S rRNA genes and use DNA fingerprinting methods to characterize microbial communities.

2.2 *How can microorganisms exist in deep subsurface rock formations?*

Most of life on the surface of the Earth is primarily supported by oxygen and the products of photosynthetic activity. In the deep subsurface environment, microorganisms have evolved to exist without oxygen and light. They comprise diverse heterotrophic and lithotrophic microbial communities able to exist in previously thought inhospitable conditions. These microorganisms are commonly found in anoxic pockets within the deep subsurface. Such anoxic conditions were found within the reducing atmosphere of the early Earth where oxygen was confined to small pockets within the anoxic ocean. It is possible that microorganism present during the anoxic conditions of the Earth could have survived and adapted to exist in the anoxic pockets of the subsurface.

2.2.1 *Heterotrophic microorganisms*

Heterotrophic microorganisms are distinguished from other microorganisms by their inability to fix carbon, but require organic carbon for growth. They must rely on a source of organic carbon produced from inorganic carbon by autotrophic organisms in order to survive. Areas rich in organic carbon are of high interest when searching for microbial life in the deep subsurface environment. The deep subsurface is primarily an anoxic environment, meaning that microorganisms must use anaerobic metabolism involving anaerobic respiration or fermentation. Since oxygen is unavailable, anaerobic organisms can use other electron acceptors such as sulphate (SO_4^{2-}), nitrate (NO^{3-}), ferric iron (Fe_2O_3) or sulphur. In general, anaerobic respiration is considered less energy efficient than aerobic respiration due to lower reduction potential of anaerobic ETC (electron transport chain) electron acceptors. One study compared concentrations of sulphate reducing activity to total organic carbon (TOC) in Cretaceous rock formations, deposited 90-93 million years ago (Krumholz, 1997). Authors observed sulphate reducing activity on the interface between permeable sandstone and low-permeability shales with high TOC content. Additionally, they detected sulphur reducing and acetogenic activity when shales and sandstone were incubated together and no activity when incubated separately, demonstrating a syntrophic relationship. Subsequently, two new species of sulphate reducing and acetogenic bacteria were isolated from the sandstone. It was postulated that TOC was diffusing out of the shales into the sandstone-shale boundary, where acetogenic bacteria were using it as a source of

energy. Metabolic products made by acetogens provided the necessary nutrients for sulphate reducing bacteria.

The diversity of heterotrophic microorganisms was investigated in Late Cretaceous and Early Tertiary formations within the Piceance Basin in the USA (Colwell, 1997). Iron reducing, sulphate reducing and fermenting bacteria were isolated from deep sandstone samples. It was determined that sulphate reducing bacteria were capable of reducing Fe(III) and could obtain energy by using this process (Colwell, 1997). Many metal reducers have various metabolic pathways for deriving energy. Adaptation of sulphate reducers to use iron is only one of the ways microorganisms employ to survive. In a separate study, a novel *Bacillus infernus* species was isolated from 2,700 meter subsurface (Boone, 1995). This was the first known species of genus *Bacillus* to be a strict anaerobe. *Bacillus infernus* species used Fe(III), Mn(IV), trimethylamine oxide or nitrate as an electron acceptor and grew in thermophilic, halophilic, and slightly alkaliphilic conditions. Clearly, heterotrophic microorganisms are well adapted to extreme environments and likely there are many more undiscovered pathways for long-term survival within the deep subsurface.

2.2.2 *Lithotrophic microorganisms*

Lithotrophic microbial communities inhabit deep subsurface environments, but unlike heterotrophic microorganisms discussed above, they do not rely on deposited organic carbon as a source of energy. Instead, they can obtain energy from reduced equivalents derived from mineral or molecular substrates found in the rock. Substrates can be produced biotically by other microorganisms or abiotically through chemical interactions. These substrates can include sulphur, sulphate, phosphate, hydrogen, or carbon dioxide. Obtained energy is used to fix carbon and provide solid carbon compounds for the use in biosynthesis. Lithotrophs can belong to bacterial or archaeal domains and can produce energy through aerobic or anaerobic respiration. Most common ETC electron acceptors include oxygen, nitrate, sulphate (SO_4^{2-}), carbon dioxide and water. One study found rich microbial diversity with presence of lithotrophic microorganisms in deep granitic aquifers in Sweden (Pedersen, 1990). Anaerobic archaea were enriched from groundwater samples with high concentrations of hydrogen, carbon dioxide, hydrogen sulphide and methane. It was posited that archaea were utilizing hydrogen gas as an electron donor, carbon dioxide as an electron acceptor and producing methane as a by-product.

Presence of sulphate reducing bacteria was detected in the groundwater accounting for high concentrations of hydrogen sulphide. In a separate study, acetogenic bacteria and acetotrophic methanogens were enriched from deep granitic groundwaters with high concentrations of hydrogen and methane (Kotelnikova, 1998). It was proposed that acetogenic bacteria can utilize hydrogen gas and carbon dioxide to generate acetate, which can be used as an energy source by acetotrophic methanogens. Sulphate reducing and iron reducing bacteria can also utilize generated acetate for their metabolic activities.

Hydrogen gas is believed to be the primary energy source in deep granite aquifers and is thought to be produced through reaction of ferrous iron in basalt with water (Krumholz, 2000). Basalts from various locations show similar hydrogen gas production, indicating that this activity is wide spread. Hydrogen production from basalt may not be unique to Earth and that under the right conditions, hydrogen production could occur on other planets, such as Mars. Recently, NASA's Mars Reconnaissance Orbiter photographed channels or gullies on the Hellas impact basin demonstrating seasonal presence of water. If there is water on the surface, then water is likely to be present in the subsurface of Mars. The simple reaction of water with basalt could theoretically be enough to support extraterrestrial life.

2.2.3 *Growth space as a limiting factor in deep subsurface environments*

Deep subsurface environments contain a diversity of viable microorganisms, but their growth is significantly limited by the imposed extreme conditions. Growth space is a major limiting factor for proper colony formation. Microorganism can survive in the pores of permeable formations or in microfractures. Sandstone is a good candidate for hosting viable microorganisms due to high porosity (average of 22%), large pore size (average more than 20 μm) and large pore-throat diameter (average of 2 μm) (Nelson, 2009). Shales have a much lower permeability and smaller pore-throat sizes (from 0.005 to 0.1 μm) (Nelson, 2009); a less ideal microbial environment. However, shales can contain high concentrations of organic carbon, methane, carbon dioxide and other gases, making them an ideal nutrient source for subsurface microorganisms. Study by Capuano (1993), indicated that permeability of microfractures in shales is similar in magnitude to that of permeability of sandstone and provided evidence of fluid flow through the shales. The average diameter (opening) of the fractures was estimated to be 10

µm. In these environments, if the average size of bacteria is assumed to be 0.5 µm, small bacterial communities could be present.

Recently, a nano-scale bacterium with a size range of 0.08 to 0.5 µm has been isolated and classified as *Nanobacterium sanguineum* (Kajander, 1997; Kajander, 1998). Nanobacteria were observed to occur in clusters, produce biofilms, and to be highly resistant to heat, gamma-irradiation and antibiotic treatment (Kajander, 1997). Their growth rate was estimated to be one hundredth that of ordinary bacteria, and they were characterised as alpha-2 subgroup of Proteobacteria. Nanobacteria has been observed in carbonate sediments and rocks (Folk, 1993), limestone (Folk, 1997a), hotsprings (Folk, 1994), and has even been proposed as an analogue to nano-fossils found on Martian meteorite ALH84001 (Thomas-Keprta, 1997). Although, the concept of nanobacteria is novel and highly controversial, it is plausible to hypothesise that under conditions of restricted growth space, such as in the microfractures or pore space, bacteria and archaea could adapt by decreasing their size. For example, in response to stress from starvation, bacteria can survive through decrease in size and activity (Roszak, 1987). It is then possible that microorganisms or their "nano" relatives could survive in shale microfractures or permeable sandstone, given the rich source of nutrients, water flow and adequate growth space.

2.3 *Geoscientific subsurface characterization of Deep Geologic Repository*

The Nuclear Bruce Site located in Tiverton, Ontario, is a proposed site for the construction of Deep Geologic Repository (DGR), which will function as storage for low-level and medium-level radioactive waste. DGR is proposed to be constructed at 680 meter depth within the Cobourg Formation composed of low-permeability limestone and overlain by 200 meters of low-permeability shale (Intera, 2011). As a part of Geoscientific Site Characterization Plan multiple boreholes were drilled to varying depths to assess the composition and chemistry of the underlying formations. Geoscientific data and models included in this section were reproduced with permission from Nuclear Waste Management Organization.

2.3.1 *Core sample location with reference to stratigraphy and mineral composition of bedrock*

Location of six core samples (S595-S801) chosen for further analysis is marked in relation to stratigraphic column of Bruce Nuclear Site bedrock in Figure 2.1. Their respective depths were 595, 643, 654, 663, 704 and 801 mBGS. Identification of major and minor minerals

in formations containing core samples is summarized in Table 2.1 (Summarized from Intera, 2011).

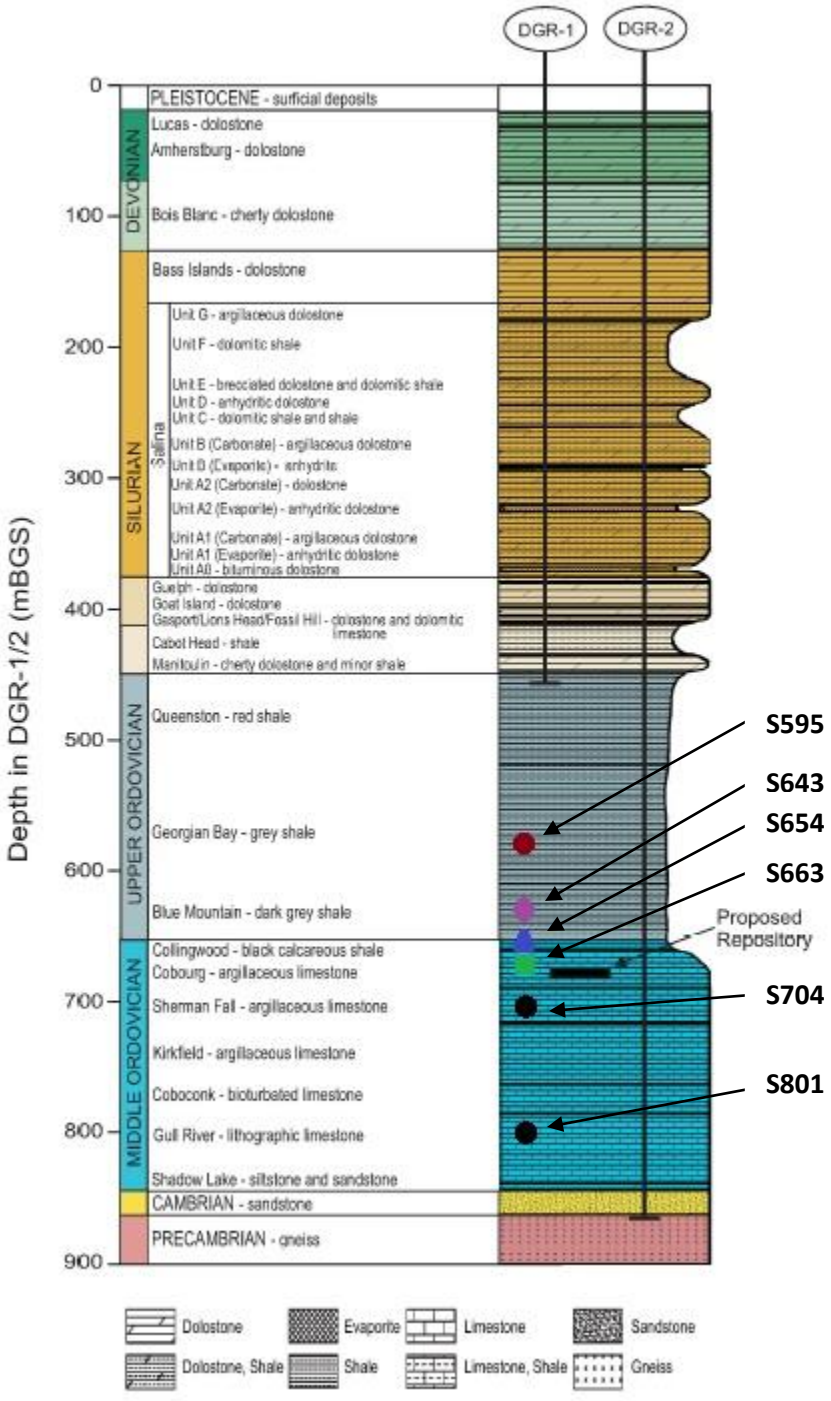

Figure 2.1: Stratigraphic sequence of bedrock based on DGR-1 and DGR-2 borehole data. Formations are identified by name, respective age and rock type. Location of six core samples (S595-S801) analyzed are color coded and labelled. Proposed depth of DGR is shown at bottom right. Modified from Intera, 2011.

Table 2.1: Summary of stratigraphic location and identification of rock type, major and minor minerals in bedrock formations for core samples S595-S801

Sample	Stratigraphic Location	Rock Type	Composition of major minerals (%wt) *						Presence of minor minerals			
			Calcite	Dolomite	Quartz	Silicates	Illite	Chlorite	Gypsum	Anhydrite	Pyrite	Halite
S595	Georgian Bay	grey shale	9	11	29	41	26	12	●	●	●	●
S643	Blue Mountain	dark grey shale	6	3	32	49	34	14			●	●
S654	Collingwood	black calcareous shale	73	9	7	10	4	2			●	
S663	Cobourg	argillaceous limestone	81	8	3	6	2	2		●	●	●
S704	Sherman Fall	argillaceous limestone	75	10	3	2	2	1	●	●	●	●
S801	Gull River	lithographic limestone	53	38	2	6	2	1		●	●	●

*based on arithmetic formation average data for six boreholes DGR1-DGR6 (Intera, 2011)

2.3.2 *Petrophysical properties of bedrock*

Rock density and porosity values are provided in this section. Data are grouped into Ordovician shales and Ordovician limestone categories. Ordovician shales include Georgian Bay, Blue Mountain and Collingwood formations representing samples S595, S643, and S654, respectively. Ordovician limestone included Cobourg, Sherman Fall and Gull River formations and represent samples S663, S704 and S801, respectively.

2.3.2.1 *Rock density for Ordovician shale and limestone*

Rock density data presented in this section include wet bulk density, dry bulk density and grain density. Wet bulk density is the ratio of total rock mass to the bulk rock volume. Dry bulk density is the ratio of solid phase of rock, without the pore water content, to the volume of bulk rock. Grain density is a ratio of solid rock content, dry bulk mass, to the total volume of grains in the rock sample. Summary of density data is provided in Table 2.2 (Summarized from Intera, 2011).

Table 2.2: Mean wet bulk, dry bulk and grain density for Ordovician shale and limestone

Rock type	Sample #	Wet Bulk density g/cm^3	Dry Bulk density g/cm^3	Grain density g/cm^3
shale	S595, S643, S654	2.65	2.59	2.76
limestone	S663, S704, S801	2.69	2.66	2.71

The mean wet bulk density of the Ordovician shale is 2.65 g/cm^3 and mean dry bulk density is 2.59 g/cm^3, indicating a difference of 0.06 g/cm^3. Ordovician limestone has a mean wet bulk density of 2.69 g/cm^3 and mean dry bulk density of 2.66 g/cm^3, with a difference of 0.03 g/cm^3. Higher difference of wet to dry bulk densities indicates the difference in porewater content of the two rock types. According to the data, porewater content of the shale is twice that of the limestone. Higher dry bulk density of the shale indicates its higher porosity, as compared to the limestone.

Mean grain densities are 2.76 g/cm^3 for the shale and 2.71 g/cm^3 for the limestone. The slight difference in grain density is a reflection of difference in the mineral composition. Limestone is composed primarily of calcite (2.71 g/cm^3) and dolomite (2.85 g/cm^3), while shale composition is primarily of illite (2.75 g/cm^3) and chlorite (2.6-2.8 g/cm^3). Sheet silicates (variable density), gypsum (2.30 g/cm^3), anhydrite (2.97 g/cm^3), halite (2.17 g/cm^3) and pyrite (5.01 g/cm^3) can influence the mean densities as well. Georgian Bay and Gull River formations were found to have the highest mean grain densities due to high content of pyrite and hematite (5.25 g/cm^3).

2.3.2.2 *Porosity of Ordovician shale and limestone*

In this section, total porosity, liquid porosity and water-loss porosity data for Ordovician shale (S595 & S643), shale hard beds (S654 & S663) and limestone (S704 and S801) are summarized. Total porosity is the ratio of rock volume not occupied by mineral grains to the total volume of the rock. Liquid porosity is the ratio of liquid, including water with dissolved solutes and oil, to total volume of the rock. Water-loss porosity is the ratio of volume of pure water to total rock volume. Arithmetic porosity means are summarized in Table 2.3 (Summarized from Intera, 2011).

Table 2.3: Mean total, liquid and water porosity for Ordovician shale and limestone

Rock type	Sample #	Total porosity % volume	Liquid porosity % volume	Water-loss porosity % volume
Shale	S595, S643	7.4	8.0	7.0
shale hard beds	S654, S663	2.5	1.7	1.5
limestone	S704, S801	1.9	1.8	1.6

Total porosity values are highest for Ordovician shale with 7.4%, followed by shale hard beds with 2.5% and limestone with 1.9%. Differences in porosity are dictated by the mineralogical composition of the formation. Shale is more porous, composed primarily of illite and chlorite-phylosilicate sheets, while shale hard beds are composed of less porous limestone and siltstone. Liquid porosity and water-loss porosity values of 8.0% and 7.0% for shale, 1.7% and 1.5% for shale hard beds and 1.8% and 1.6% for limestone respectively, indicate that shale has much higher pore content of water and oil. In general, total porosity includes liquid porosity and any volume of gas present. Higher liquid than total porosity of shale is likely due to hydrated water content of gypsum, which is released during experimental heating. (Intera, 2011)

2.3.2.3 *Pore size distribution of Ordovician shale and limestone*

Pore size of Ordovician shale and limestone is illustrated as median pore-throat radii (r_{50}) of pores in Figure 2.2. There is an increasing trend of pore-throat radii from the Cobourg Formation to Georgian Bay shales. However, pore-throat radii for Ordovician shale and limestone, within the depth of S595 to S801 samples, stay within the 3 to 10 nm range. Pore-throat radii for the limestones are not as clear due to lack of data points; however, r_{50} values are not expected to exceed this range due to low porosity of the limestone.

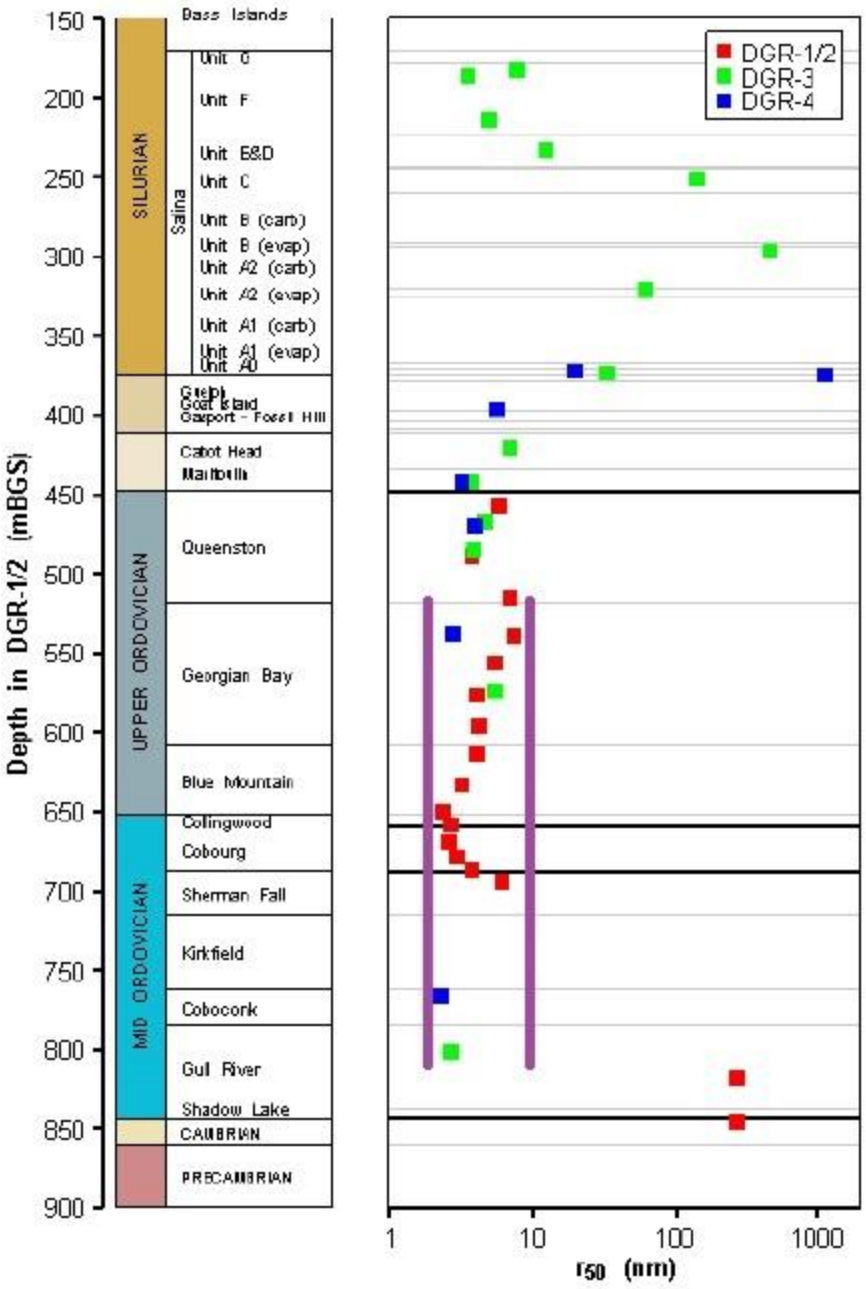

Figure 2.2: Pore-throat radii (r50) of bedrock based on DGR-1, DGR-2, DGR-3and DGR-4 borehole data. Range of pore-throat radii for Ordovician shale and limestone formations within the depth S595-S801 samples are confined to within 3-10 nm range (purple lines). Formations are identified by name and their respective age. Modified from Intera, 2011.

2.3.3 *Hydrocarbon presence in Ordovician shale and limestone*

Occurrence of hydrocarbons in Ordovician shale and limestone is variable depending on the formation. Hydrocarbon distribution for selected Ordovician formations is summarized in Table 2.4 (Summarized from Intera, 2011).

Table 2.4: Hydrocarbon occurrence in Ordovician shale and limestone formations

Formation	Sample #	Hydrocarbon occurrence
Georgian Bay	S595	Petroliferous shale with hydrocarbon odour
Blue Mountain	S643	Petroliferous shale with hydrocarbon odour
Collingwood	S654	Petroliferous shale with hydrocarbon odour
Cobourg	S663	Bituminous laminations
Sherman Fall	S704	Hydrocarbon odour
Gull River	S801	Hydrocarbon odour

Oil saturation profile for boreholes DGR1-DGR6 is illustrated in Figure 2.3. It is clear that higher concentrations of oil are found in Ordovician formations, in particular between Georgian Bay and Sherman Fall formations. In any particular formation oil saturation varies, however, most of values are found within 2% to 10% range. Highest oil saturation occurs in Collingwood Formation with a value of 17%. Differences in oil saturation could be related to oil heterogeneity. As was described previously, oil represents an excellent energy source for heterotrophic microorganisms.

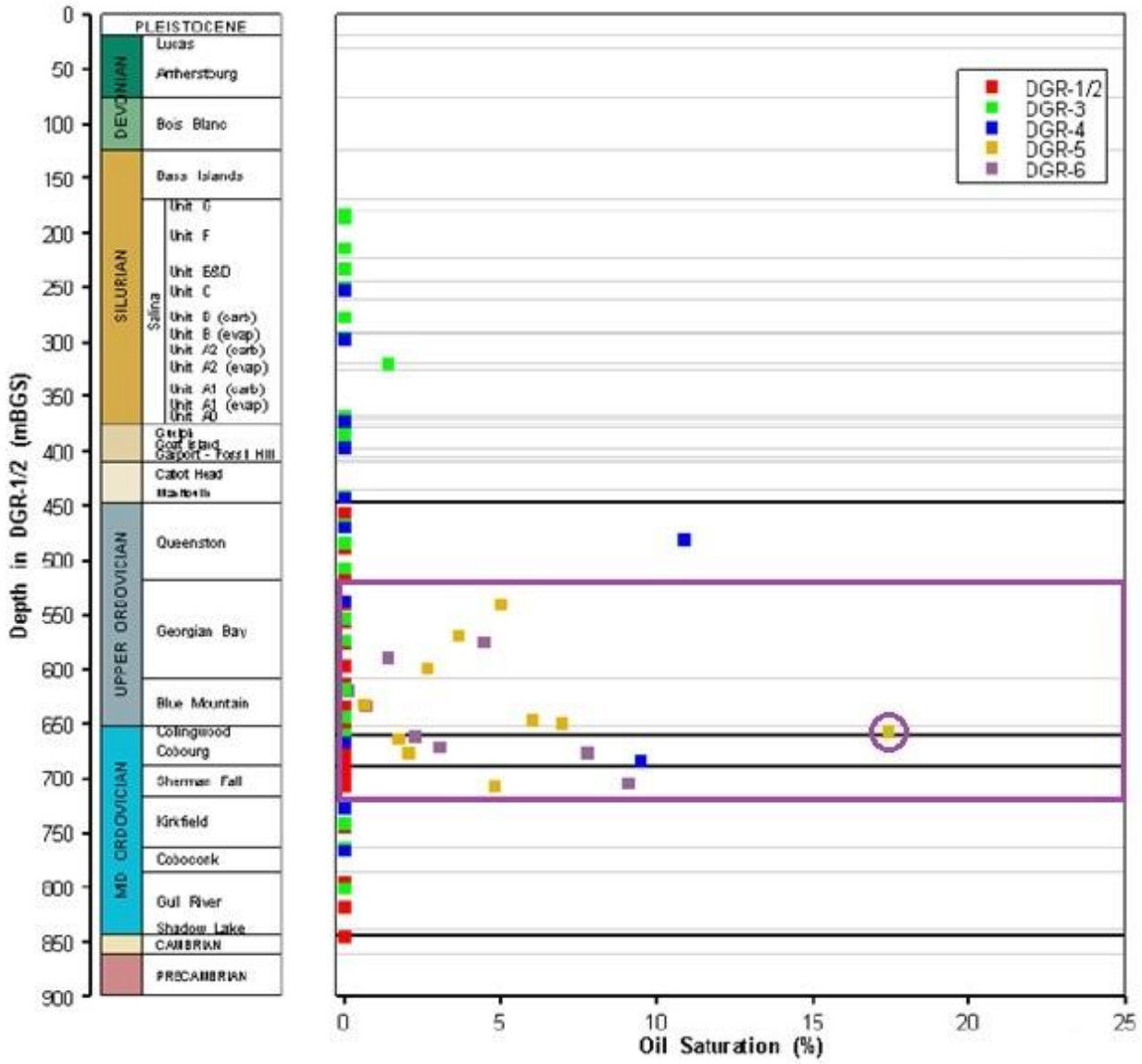

Figure 2.3: Oil saturation profile derived from boreholes DGR1-DGR6 for various formations. Area of interest corresponding to Ordovician shale and limestone with high oil saturation is included within the purple perimeter. Purple circle highlights maximum oil saturation value of 17% (purple circle). Modified from Intera, 2011.

2.3.4 *Organic carbon and sulphur content in Ordovician shale and limestone*

Organic carbon is an essential energy source for heterotrophic bacteria and sulphate reducing activity was previously correlated to high TOC content (Krumholz, 1997). Organic carbon content and total sulphur determined from core samples in boreholes DGR1 to DGR6 are illustrated in Figure 2.4. Organic carbon content in Upper Ordovician shale and Middle Ordovician limestone falls in the 0.1 to 0.5% range with an exception of organic carbon peak at the bottom of Blue Mountain and Collingwood formation with a maximum value of 2.5%. Total sulphur content found in Ordovician shale and limestone ranges between 0.1 and 1.4%. The presence of total sulphur can be related to sulphide-bearing minerals such as pyrite in shale and anhydrite in limestone. The total sulphur peak is located in the zone of Blue Mountain and Collingwood formation. Coincidence of the organic carbon and sulphur peaks is likely due to the presence hydrocarbons such petroleum and bitumen that contain sulphur.

Acetogenic bacteria that use organic carbon as an energy source can produce acetate and hydrogen sulphide in the absence of oxygen. Hydrogen sulphide is a reducing agent and can react with water in pores to produce elemental sulphur or sulphate. In turn, produced sulphur and sulphate can support sulphur-reducing and sulphate-reducing bacteria, respectively.

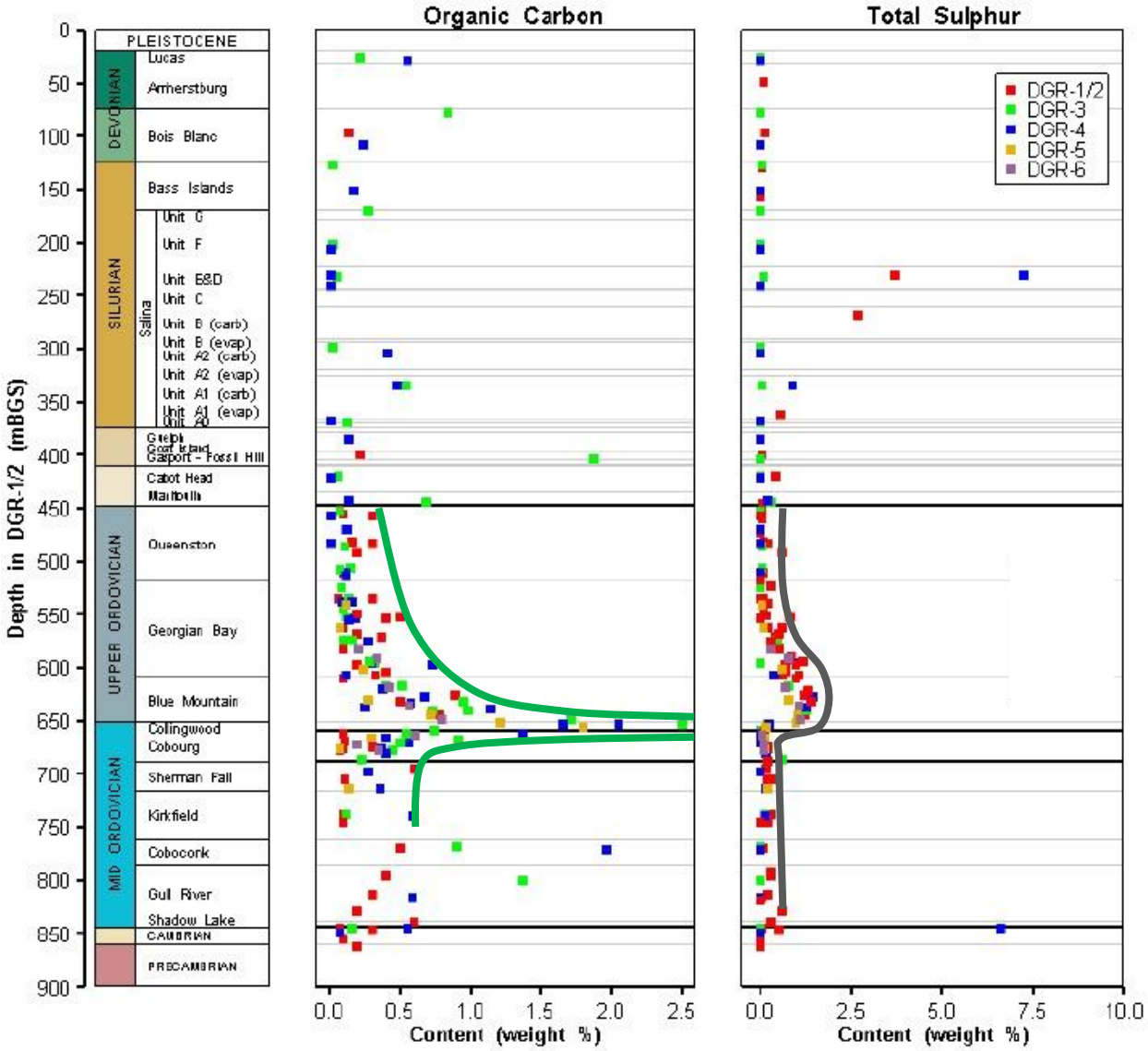

Figure 2.4: Content of organic carbon and total sulphur obtained from DGR1-DGR6 boreholes. Organic carbon peak is outlined with via green and total sulphur trend via dark grey contour. Modified from Intera, 2011.

2.3.5 *Methane and carbon dioxide concentrations in Ordovician shale and limestone*

Presence of methane and carbon dioxide of biogenic origin are important indicators of past or present microorganism activity. Carbon can be used by acetogenic bacteria to generate acetate that can be used by methanogenic archaea to produce methane.

2.3.5.1 *Methane and $\delta^{13}C$*

Methane concentrations, normalized to water content (mm kgw^{-1}) and the $\delta^{13}C$ isotope profile are illustrated in Figure 2.5. The methane concentration increases from the Queenston Formation to the bottom of the Collingwood Formation, with the maximum methane concentration detected in the Collingwood Formation. Methane concentration drops off sharply below the Cobourg Formation, forming a prominent methane peak. A smaller methane peak can be observed at the Coboconk Formation. The Collingwood and Cobourg formations have the highest measured concentrations of methane forming the tip of the peak. This peak, however, is augmented by the low water content of the Cobourg limestone, and would occur at the base of the Blue Mountain Formation if methane yields were normalized to mass of rock. $\delta^{13}C$ isotope concentrations within methane show a depletion trend from Queenston Formation down to the top of the Blue Mountain Formation, below which the $\delta^{13}C$ concentration values remain fairly constant down to the top of the Cobourg formation. Below the Cobourg Formation the $\delta^{13}C$ isotope trend is significantly enriched forming a gradient; enrichment of $\delta^{13}C$ isotope continues to increase with depth. Both the highest methane concentration and the lowest concentration of $\delta^{13}C$ isotope coincide at the Collingwood Formation.

The $\delta^{13}C$ isotope depletion trend indicates that methane contained within Upper Ordovician shale is likely of biogenic origin due to low $\delta^{13}C$ values. The Blue Mountain and Collingwood formations contain the highest concentration of biogenic methane. Process of acetate fermentation is thought to be responsible for the production of biogenic methane. On the other hand, $\delta^{13}C$ isotope enrichment for Middle Ordovician limestone is indicative of a thermocatalytic origin of methane. Thermocatalytic methane is generated through cracking of hydrocarbons at high temperatures. (Intera, 2011)

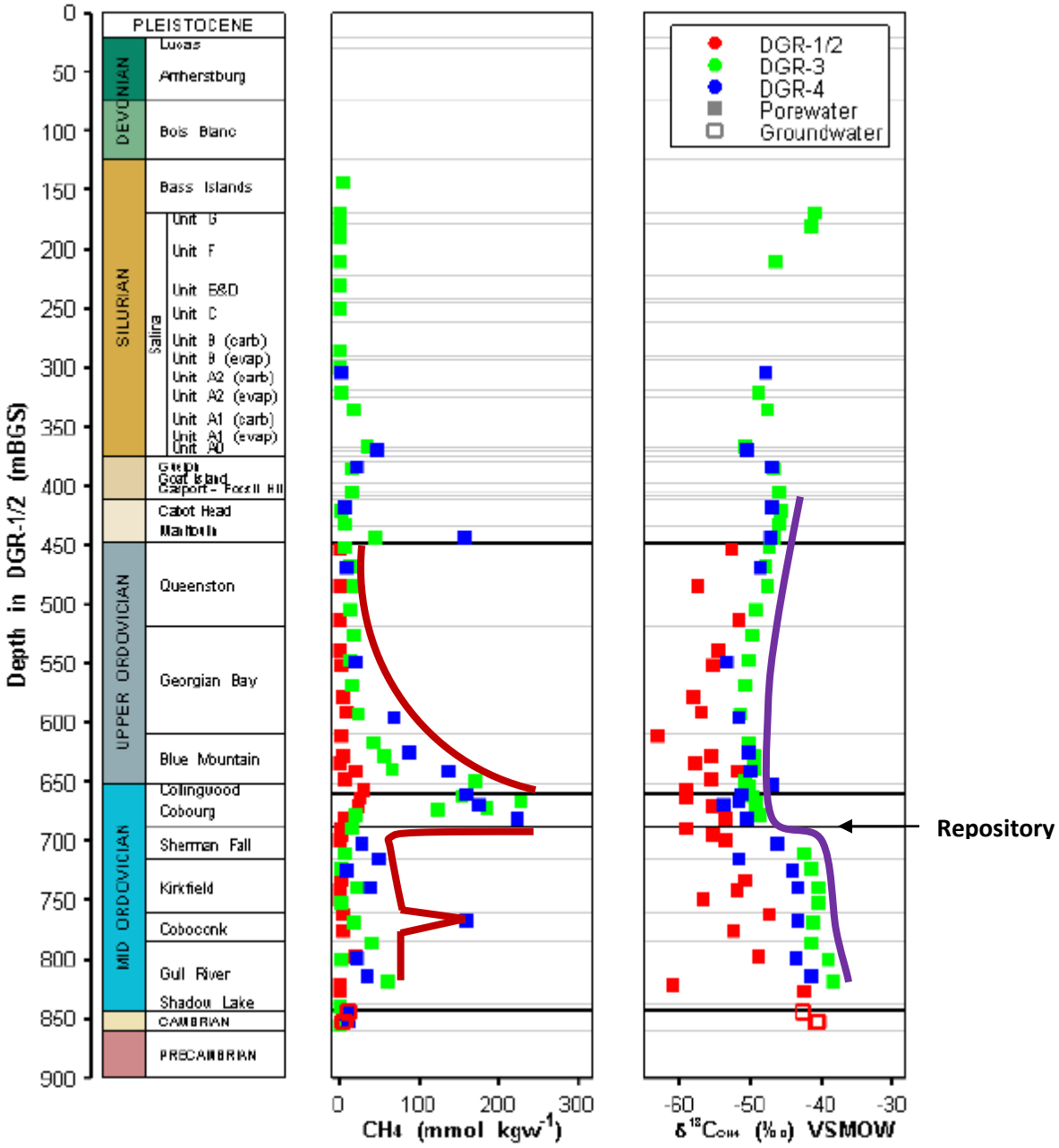

Figure 2.5: Concentration of methane (CH_4) and $\delta^{13}C$ isotope concentration within methane obtained from DGR1-DGR4 boreholes. Concentrations of methane were obtained from porewater and normalized to porewater mass. Methane peak is outlined via dark red and $\delta^{13}C$ isotope trend via violet contour. Note: VSMOW should be PDB. Modified from Intera, 2011.

2.3.5.2 *Carbon dioxide and $\delta^{13}C$*

Concentrations of carbon dioxide and the corresponding $\delta^{13}C$ isotope profile are illustrated in Figure 2.6. Concentrations of carbon dioxide in the Upper Ordovician shale are low, but the concentrations begin to increase in the Collingwood shale. Carbon dioxide concentrations become significantly higher in the Middle Ordovician limestone, where the high concentrations persist and another carbon dioxide peak is observed in the Coboconk Formation. This creates a significant CO_2 gradient below the Collingwood Formation. The $\delta^{13}C$ isotope enrichment peak from middle of the Georgian Bay to the bottom of Blue Mountain Formation indicates lower abundance of ^{12}C within the measured CO_2 as compared to the rest of DGR profile. Such low abundance of ^{12}C could be a result of biological methanogenesis, because inorganic exchange or reaction with the carbonate rock cannot produce such high $\delta^{13}C$ enrichments (Intera, 2011). The $\delta^{13}C$ isotope enrichment peak is located above the methane peak and the area of high carbon dioxide concentration.

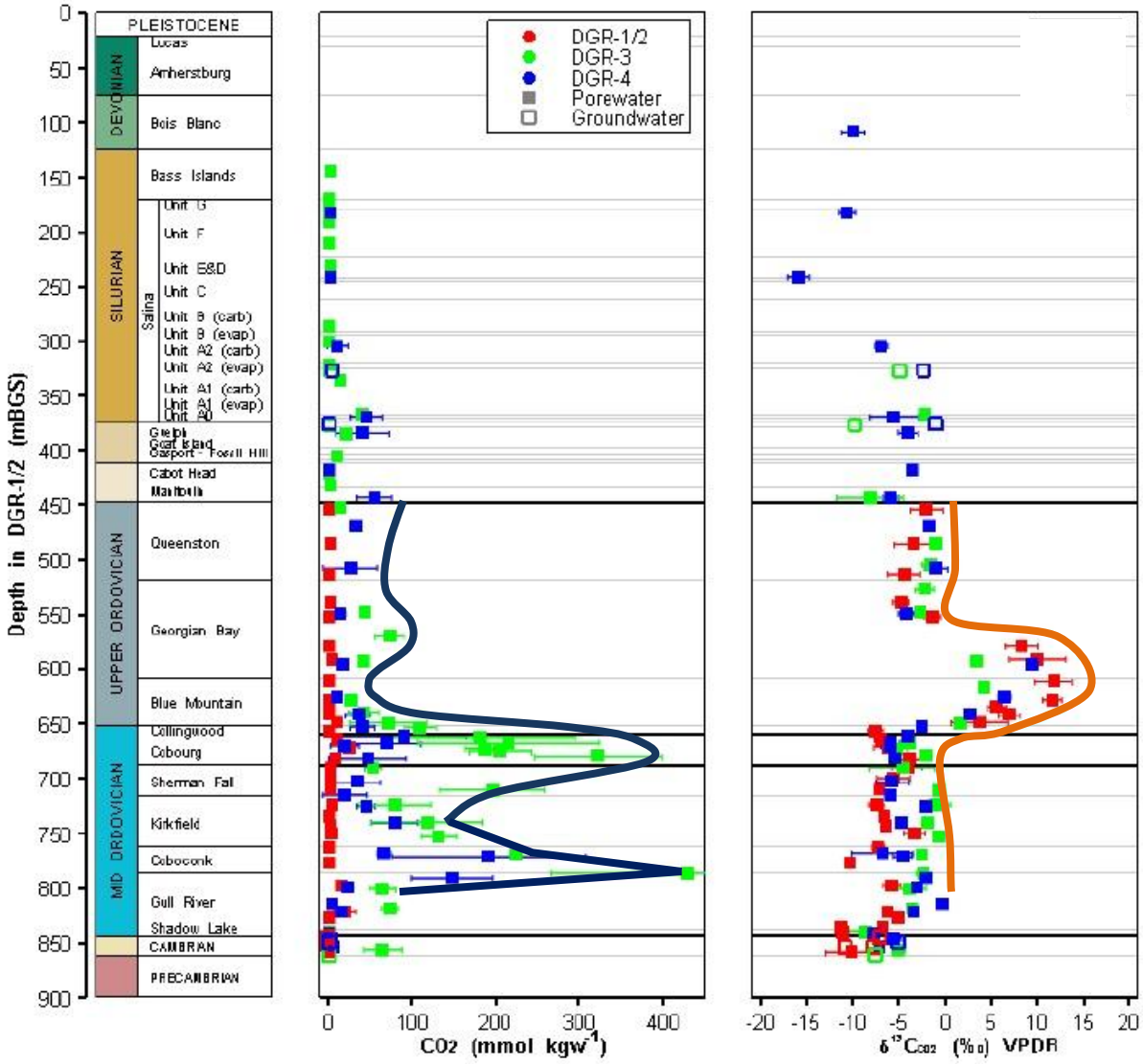

Figure 2.6: Concentration of carbon dioxide (CO_2) and $\delta^{13}C$ isotope concentration within carbon dioxide obtained from DGR1-DGR4 boreholes. Concentrations of carbon dioxide were obtained from porewater and normalized to porewater mass. Profile of carbon dioxide concentration is outlined via dark blue and $\delta^{13}C$ isotope peak via orange contour. Please note that these are approximate tread lines and uncertainties in the CO2 concentrations below 650 mBGS are very significant, although relatively higher as compared to above 650 mBGS. Modified from Intera, 2011.

2.4 *Hypothesis and objectives*

The geological system presented in this study contains unique geochemical features that could support viable microorganisms. Organic carbon found within limestone and shale, together with the enrichments in the $\delta^{13}C$ of carbon dioxide and the occurrence of the biogenic methane peak could indicate biological methanogenesis.

The hypothesis is that the biogenic methane peak with $\delta^{13}C$ isotope depletion and low carbon dioxide concentrations with a $\delta^{13}C$ isotope enrichment are geochemical features created through past or current biological processes.

In this project, the overarching objective is to develop/construct a model of how microbes contributed to the observed geochemical gradients.

To do so, a set of objectives was designed:

The first objective will be to test for the presence of bacterial and archaeal 16S rRNA genes in zones containing biogenic methane (S595, S643, S654, S663) and in zones with thermocatalytic methane (S704, S801). This will be done by using DNA extraction and DNA amplification techniques discussed in the following section. We expect to amplify both bacterial and archaeal DNA from the biogenic methane zone and none from the thermocatalytic zone.

The second objective will be to identify microorganisms based on the amplified 16S rRNA genes. This will be done through cloning and sequencing steps discussed in the following section. We expect to find heterotrophic and lithotrophic bacteria, along with methanogenic archaea.

The third objective will be to classify microbial communities through molecular phylogenetic analysis, and to determine if the distribution and diversity of microbes varies or remains constant with depth. Distribution of microbes will be related to observed geochemical features for the purpose of relating geochemistry and microbiology.

3. Methods

3.1 *Core Sampling Methods*

Six cores were sampled along the borehole depth profile from Bruce Nuclear Site near Tiverton, Ontario. Cores selected for DNA phylogenetic analysis were strategically selected at 595, 643, 654, 663, 704, and 801 mBGS based on geochemical data. Each core was sectioned using clean instruments (70% ethanol washed) and sectioned pieces were handled using powder-free sterile gloves. Core pieces were picked from the inner part of the core to avoid any surface or drilling fluid contaminants. During transport between laboratories at the University of Ottawa, core pieces were placed in sterile polypropylene Falcon tubes.

3.2 *DNA Fingerprinting methods*

Due to the constraints posed by the compositional nature of the core samples, low-biomass, and scarce information available on the subject of DNA Fingerprinting methods in deep geological formations, many DNA extraction and amplification methods were tested. Majority of the methods undertaken generated negative results; however, much was learned from these methods. Various parts of the protocol were modified, and through trial and error, a working procedure was established. Only the methods that generated positive results are thus explained in this section. All other methods that were attempted but failed to produce positive results are included in the appendix.

3.2.1 *Sample preparation for DNA extraction*

Approximately 10 g of the sectioned core pieces were pulverized using sterile mortar and pestle. At all times powder-free gloves were worn and sample preparation area and gloves were sterilized using 70% ethanol and 10% bleach. Mortar and pestle were sterilized by washing with 70% ethanol and 10% bleach, and flaming thoroughly using a Bunsen burner. Sterile mortar and pestle were placed by the side of the Bunsen burner to prevent any airborne contamination. Prior to being pulverized, each of the selected core pieces was flamed all around using a Bunsen

burner and placed in the sterile mortar. Core pieces were pulverized for approximately 15-20 minutes until no visible coarse grains remained and sample resembled a fine powder. Powdered sample was transferred to a 50 mL pre-weighed sterile polycarbonate centrifuge tube. The tube was weighed using a microbalance and powder mass was determined.

The powdered samples were subsequently washed with 40 mL of a pre-extraction solution containing 50 mM Tris-HCl, 50 mM Sodium Phosphate, 10 mM EDTA, 20 mM EGTA. This washing step was included to remove cationic contaminants that would hinder subsequent DNA extraction and amplification. The centrifuge tubes were vortexed at maximum speed for 30 seconds to re-suspend the powder and was subsequently centrifuged for 10 minutes at 10,000 RPM and 4^{o}C. Pre-extraction solution was carefully removed by pipetting, as not to disturb the pellet. Washing procedure was repeated four more times. Prior to use, pre-extraction solution was adjusted to a pH of 7.6 and sterilized by autoclaving at 121^{o}C for 60 minutes. All steps of the washing procedure, apart from the centrifugation steps, were carried out in a laminar flow hood. Laminar flow hood and all the necessary equipment were sterilized with 70% ethanol, 10% bleach and UV lights. During the transfer from centrifuge to the laminar flow hood, the outside of the centrifuge tubes was sterilized using 70% ethanol and 10% bleach. Aerosol resistant filter tips were used for the washing step with pre-extraction solution.

3.2.2 *DNA extraction from core powder*

DNA extraction was performed using the PowerMax Soil DNA Isolation Kit (# 12988-10) from MoBio Laboratories. The DNA extraction steps were carried out following the manufacturer's instructions. Concentration of the final DNA product was determined using a NanoDrop spectrophotometer.

All steps of the DNA extraction procedure, apart from the centrifugation steps, were performed in a laminar flow hood. Laminar flow hood and all the necessary equipment were sterilized using 70% ethanol, 10% bleach and UV decontamination lights. Aerosol resistant filter tips were used throughout DNA extraction procedure.

3.2.3 *Polymerase Chain Reaction from extracted DNA*

PCR reactions were carried out in a final volume of 25 uL using Invitrogen Platinum Taq DNA polymerase (# 10966). All PCR reactions were prepared in the laminar flow hood using aerosol resistant filter tips. Laminar flow hood and all the necessary equipment were sterilized using 70% ethanol, 10% bleach and UV lights.

Specific PCR conditions were developed in order to amplify extracted DNA product. It was determined through testing that the following PCR reagents and concentrations yielded the best DNA amplification: 2.5 uL of 10X PCR Buffer, 0.2 mM dNTP, 3.0 mM, $MgCl_2$, 1 uM of each primer, 1.0 unit of Invitrogen Platinum Taq DNA Polymerase, 5% (v/v) DMSO, 0.5 ug/uL BSA, 1 uL of final extracted DNA product, and PCR grade water to a final volume of 25 uL. Addition of DMSO and BSA was found to significantly improve DNA amplification.

In order to amplify bacterial DNA, we used universal bacterial primers 27F (5'- AGA GTT TGA TCM TGG CTC AG -3') and 907R (5'- CCG TCA ATT CMT TTR AGT TT -3'), known to successfully amplify bacterial 16S rRNA gene fragments from environmental samples (Hery *et al*. 2007). To test for the presence of archaeal DNA, we used universal archaeal primers A571F (5'- GCY TAA AGS RIC CGT AGC -3') and UA1204R (5'- TTM GGG GCA TRC IKA CCT -3'), which were shown to amplify archaeal 16S rRNA gene fragments from all types of archaea (Baker *et al*. 2003). We also tested for the presence of genes encoding for the alpha subunit of the methyl coenzyme-M reductase, responsible for the last step of methanogenesis using primer pair ME1 (5'- GCM ATG CAR ATH GGW ATG TC -3') and ME2 (5'- TCA TKG CRT AGT TDG GRT AGT -3') (Hales *et al*. 1996). For all genes of interest in this study, the thermocycler program was set to 94^{o}C for 10 minutes initial denaturation, followed by 35 cycles of 94^{o}C for 30 seconds, 48.8^{o}C for 60 seconds, and 72^{o}C for 60 seconds with a final extension at 72^{o}C for 5 minutes.

Presence of nanobacterial DNA was tested using nanobacteria-specific primers NanoBacF (5'- ACA ATG GTG GTG ACA GTG GG -3') and NanoBacR (5'- CGG GTA AAA CCA ACT CCC AT -3') (Integrated DNA Technologies) that were previously used to amplify nanobacterial gene fragments from human patients (Kim *et al*. 2011). In order to test for presence of nanoarchaeal DNA, we used two nanoarchaea-spefic primers. First primer pair was

A571F (5'- GCY TAA AGS RIC CGT AGC -3') and N961R (5'- CMA TTA AAC CGC RCA CCC -3') (Integrated DNA Technologies), and second primer pair was 9bF (5'- CCC GTT GAT CCT GCG GGA G -3') and 511mcR (5'- CTT GCC CAC CGC TT -3') (Integrated DNA Technologies). Both of these primer pairs were previously used to amplify nanoarchaeal DNA gene fragments from Chinese and New Zealand hydrothermal sediments (Casanueva *et al*. 2008).

For nanobacterial and nanoarchaeal DNA amplification Biometra TProfessional Basic Gradient thermal cycler was used, with the following PCR parameters: 94^{o}C for 10 minutes initial denaturation, followed by 35 cycles of 94^{o}C for 30 seconds, 53^{o}C for 60 seconds, and 72^{o}C for 60 seconds with a final extension at 72^{o}C for 5 minutes.

PCR products were loaded onto 1% agarose gel with ethidium bromide staining and separated by electrophoresis at 80V for 50 minutes. DNA bands were visualized using a Fusion Fx5 Gel DOC (Vilbert Lourmat). PCR products were frozen at -20^{o}C for subsequent analysis.

3.2.4 *PCR product clean-up for cloning purposes*

PCR products amplified for cloning purposes were never frozen, and after PCR amplification were loaded directly onto a 1% agarose gel with ethidium bromide staining. In order to achieve best possible DNA yields, 50 uL of PCR products were loaded per sample. Gel electrophoresis was performed at 80V for 50 minutes. DNA bands were confirmed using Fusion Fx5 Gel DOC and bands were cut-out from the gel using sterile scalpel and tweezers that were cleaned with 70% ethanol, 10% bleach and flamed briefly prior to the start and in between each band isolation. Each gel piece containing a specific DNA band was then placed in a sterile pre-weighed 1.5 mL centrifuge tube. DNA was then extracted from the gel pieces and purified using Qiagen Quiex II Gel Extraction Kit (# 1054486). Purified PCR product was eluted using 40 uL of 10 mM Tris-Cl and DNA concentration was measured using a NanoDrop spectrophotometer.

3.2.5 *Cloning PCR product into E. Coli vector*

PCR products were gel purified and cloned using StrataClone PCR Cloning Kit using pSC-A as a cloning vector (# 240205). Cloning was performed with StrataClone SoloPack competent cells expressing Cre recombinase, which were provided with the kit and following the manufacturer's recommendations. Information on procedural details and required materials can be found in the StrataClone PCR Cloning Kit manual. For maximum cloning efficiency, freshly purified DNA products were used for cloning.

PCR products that were amplified using Taq Polymerase contain Adenine (A) overhangs that can complementary bind to specifically designed Thymine (T) overhangs found on either side of the cloning vector. Pairing of A to T allows for the ligation of PCR product into the vector, which is subsequently taken up by the completent *E. Coli* cells through heat shock treatment. Competent cells were allowed to recover for 90 minutes at 37^{o}C and were subsequently plated in 25 uL, 50 uL, and 100 uL amounts on LB agar plates containing 10 mL/mg ampicilin and 2% (v/v) X-gal. Ampicilin allows only *E. Coli* cells containing a plasmid vector to grow, while X-Gal differentiates cells with vectors containing the insert and vectors that do not. LB plates were incubated at 37^{o}C for 17-20 hours to allow for proper size and number of colonies to grow. White colonies, each containing a unique ligated DNA sequence, were picked for subsequent analysis.

3.2.6 *Clone Libraries*

All procedures in this section were performed on a bench top cleaned with 70% ethanol and 10% bleach, near a Bunsen burner. To prevent cross-contamination, aerosol-free filter tips were used and all equipment was cleaned with 70% ethanol and 10% bleach.

3.2.6.1 *Colony extraction*

For each sample, 96 colonies were picked at random using autoclaved wooden picks and transferred into separate wells in a sterile 96 deep-well plate containing 1 mL of LB broth with

10 mg/mL ampicilin. Wooden picks were allowed to rest in the wells for about 5 minutes, after which time they were carefully taken out. Deep well plates were sealed using Progene air permeable plate sealers (# 24-BF-400-S) and placed in a shaking incubator at 139 RPM, 37^{o}C for 18 hours. After the 18 hour incubation period, wells containing growing *E. Coli* cells were opaque in appearance.

3.2.6.2 *Cryostock*

Freshly grown *E. Coli* cells were transferred from a 96 deep-well plate to a Progene 96 well round bottom assay plate (# P-96-450R-C) used for storing cells at very low temperatures. Here it is necessary to use aerosol-free filter tips to transfer the cultures, since the chance of cross-contamination between wells is very high. 100 uL of *E. Coli* cells were combined with 100 uL of autoclaved 1:1 Glycerol:H_2O solution in each well to minimize cell damage during freezing and thawing of the cryostock. Cryostock plates were sealed with Sarstedt Alu-sealing tape (# 95.1995) and placed at -80^{o}C for long term storage. Remaining cells in the 96 deep-well plate were used for M13 PCR and stored at 4^{o}C for further analysis.

3.2.6.3 *M13 PCR*

M13 PCR was performed in order to amplify DNA insert located within pSC-A cloning vector from the transformed *E. coli* cells. PCR reactions were carried out in a total volume of 25 uL using freshly grown *E. Coli* cells as a template and Lucigen EconoTaq PLUS 2X Master Mix (# 30035-1) that contained all the necessary reagents. The primer pair used was M13F (5'- GTA AAA CGA CGC CCA G -3') and M13R (5'- CAG GAA ACA GCT ATG AC -3') (Integrated DNA Technologies). PCR reactions were carried out in the Biometra TProfessional Basic Gradient thermal cycler with the following PCR parameters: 94^{o}C for 10 minutes initial denaturation, followed by 35 cycles of 94^{o}C for 30 seconds, 44.5^{o}C for 30 seconds, and 72^{o}C for 90 seconds with a final extension at 72^{o}C for 12 minutes. Success of the PCR reaction was checked by gel electrophoresis. Samples were subsequently kept at 4^{o}C or frozen at -20^{o}C for further analysis.

3.2.6.4 *Amplified rDNA restriction analysis (ARDRA)*

For 47 out of 96 M13 PCR products from each sample: 0.5 uL of the PCR products was added to 1 uL of New England BioLabs 10X NEbuffer 4 (# B7004S), 1.25 U of New England BioLabs AluI restriction enzyme (# R0137S), 1.25 U New England EcoRI restriction enzyme (# R0101S) and topped with molecular grade water to a final volume of 10 uL. RFLP reactions were enclosed in sterile PCR tubes and incubated in the Biometra TProfessional Basic Gradient thermal cycler with set at 37^{o}C for 3 hours, which allowed DNA to be fully cut by the restriction enzymes. Reactions were subsequently transferred into an incubator set at 65^{o}C for 20 minutes to inactivate restriction enzymes. Reactions were then stored at -20^{o}C for further analysis.

3.2.6.5. *ARDRA using gel electrophoresis*

Digested PCR products were loaded onto a 2% agarose gel, made by mixing 200 mL of 0.5X TBE buffer with 4 g of Invitrogen UltraPure Agarose (# 16500-500). Agarose mixture was heated in the microwave until it fully melted and started to boil. Subsequently agarose solution was allowed to cool on the bench top to around 55^{o}C, at which point 15 uL of 1% ethidium bromide solution was added and carefully mixed into the solution to prevent formation of air bubbles. Agarose solution was then poured into the tray of FisherBiotech Electrophoresis System (# FB-SB-1316) and four 20-well combs (20 uL well capacity) were used per each gel. Gel was allowed to stand for 45 minutes in order to fully solidify. Once the gel was ready, 0.5X TBE running buffer was poured into the electrophoresis unit until the gel was fully covered. The first and last well of every comb was loaded with 4 uL of New England BioLabs 100 bp ladder (# N3231L). Additionally, undigested M13 PCR product was loaded once per comb to serve as a negative control. Gel was run at 80V for 75 minutes to allow for full separation of DNA bands and the ladder. Gel was visualized using Fusion Fx5 Gel DOC (Montreal Biotech) under UV exposure and digital images were taken for subsequent analysis. ARDRA profiles were visually analyzed in order to group identical digestion products together according to patterns produced by size of the bands.

3.2.7 *Sequencing*

Identical restriction patterns were grouped together through careful analysis of the gel images and up to two clones from each pattern group were selected and re-amplified using the M13 PCR protocol with higher grade reagents. PCR reaction was carried out in a final volume of 25 uL using 1.0 U of Invitrogen Platinum Taq DNA polymerase (# 10966), 2.5 uL of 10X PCR Buffer –Mg, 0.2 mM dNTP mixture, 3.0 mM $MgCl_2$, 1 uM of each M13 primer, PCR grade water and 1 uL of transformed *E. Coli* cells containing the plasmid insert as a template. Success of the M13 PCR was checked by gel electrophoresis and DNA amount of each PCR reaction was determined using a NanoDrop spectrophotometer to be over 20 ng/uL. Two clones representative of each restriction patterns were sent for Sanger sequencing using M13F primer at the McGill University and Genome Quebec Innovation Centre.

3.3 *Analysis of Results*

3.3.1 *Construction of phylogenetic trees*

Sequencing results were converted to Fasta format and parts of sequences coding for StrataClone PCR Cloning Vector pSC-A and M13 primer pair were deleted using Mega5 Alignment Explorer. Additionally, sequences that were thought to be chimeric were removed from the database. Sequences were aligned using Ribosomal Database Project release #10 (http://rdp.cme.msu.edu) (RDP). Aligned sequences were then matched to known sequences using the Seqmatch program from the RDP website. Phylogenetic tree was constructed from the aligned sequence data set using the RDP Tree Builder tool. The phylogenetic tree is constructed using Weighbor tree topology, which is a weighted version of Neighbor Joining that gives significantly less weight to the longer distances in the distance matrix. DNA distance matrix is calculated using the Jukes-Cantor Correction and the phylogenetic tree was tested with 100 bootstrap re-samplings. Bacterial phylogenetic tree was rooted with *Methanocaldococcus jannaschii* (L77117) as an archaeal outgroup. Phylogenetic tree was further edited using Mega5 Tree Explorer program and was inserted as an image into Microsoft Powerpoint where topological names and markers were added.

3.3.2 *Statistical analysis on clone libraries*

Statistical analysis was conducted on each of the four clone libraries based on the RFLP profiles. Distinct RFLP profiles containing one or more clones were considered to be a distinct species.

Rarefaction curves were constructed in Microsoft Excel using data set obtained from San Diego State University FastGroupII bioinformatics platform (http://biome.sdsu.edu/fastgroup/).

3.3.2.1 *Diversity analysis on clone libraries*

As a part of the analysis, two diversity indices were calculated. The first index, known as the Simpson's diversity index (*D*) was calculated as follows:

$$D = \sum (n / N)^2 \qquad [1]$$

In this formula, "*n*" represents the total number of organisms of a particular species and "*N*" represents the total number of organisms of all species. Therefore, "*n*/*N*" is the fraction of all organisms that belong to a particular species. Lower values obtained with this index indicate higher diversity and higher values indicate lower diversity. For example, a value of 0 would indicate an infinite diversity, while a value of 1 would indicate no diversity.

The second index used in the diversity analysis is called the Shannon-Wiener diversity index (H) and was calculated as follows:

$$H = -\sum (p_i (\ln p_i)) \qquad [2]$$

In this formula, "p_i" is the relative abundance of each species, calculated as the proportion of organisms of a particular species to the total number of organisms of all species, or "*n*/*N*". The total number of species is called species richness (*S*), and it is the total number of species "$_i$" in the equation. Calculated values of the index typically fall within the range of 1.5 to 3.5. Lower values indicate low species richness and evenness, while higher values indicate high species richness and evenness. The advantage of this index is that it takes into account the number of the

species and the evenness of the species. This index can be increased with higher number of unique species or with higher species evenness.

Species evenness (E_H) refers to the proportion of the number of organisms of each species in a community and is thus taken as a measure of biodiversity. The more even the number of organisms in each species the more even is the community. Species evenness is calculated as follows:

$$E_H = H / (\ln S) \qquad [3]$$

In this formula, "*H*" is the value obtained from the Shannon-Weiner diversity index and "*S*" is the richness of the species or the total number of species in the community. Species evenness falls in the range of 0 to 1, with lower values indicating higher evenness of the community.

3.3.2.2 *Similarity analysis between clone libraries*

Two similarity indexes were calculated in order to assess the similarity between the clone libraries and compare clone libraries in terms of variability. The first similarity index, known as the Bray-Curtis index (*BC*) (Bray and Curtis, 1957) was calculated using the following equation:

$$BC = \{ \sum |n_{1i} - n_{2i}| \ / \ \sum (n_{1i} + n_{2i}) \ \} \quad [4]$$

In this equation, "n_{1i}" is the number of individuals of species "*i*" in clone library "*1*" and "n_{2i}" is the number of individuals of species "*i*" in clone library "*2*". Values generated by the index fall within 0 to 1 range, with higher values indicating higher species turnover and species richness in between two clone libraries. Bray-Curtis index can be sensitive to the sample size (Wolda, 1981) and therefore another similarity index was calculated to support the results.

The second similarity index, known as the Morisita-Horn index of overlap (C_λ) (Horn, 1966) is a simplified Morisita similarity index (Morisita, 1959) and was calculated using the following equation:

$$C_\lambda = \{ 2 \sum n_{1i} \, n_{2i} \ \} \ / \ \{ \ (\lambda_1 + \lambda_2) \, N_1 N_2 \ \} \qquad [5], \quad \lambda_x = \{ \sum (n_{xi})^2 \ \} \ / \ \{ \ (N_x)^2 \ \}$$

In this equation, "n_{xi}" stands for number of individuals of species "*i*" in clone library "*x*", and "N_x" is the total number of individual in clone library "*x*". For two samples, if species are found in the same proportion then the index will equal to 1. If the index value is close to zero, then the two samples are completely different. Morisita-Horn index was chosen because it provides a good "fit" when approximating the relation between maximum value of the index and sample size and is best used when number of individuals in each samples size, *ie.* clone library, is relatively the same. (Wolda, 1981)

Similarity indices were calculated using EstimateS Ver. 8.2.0, a freeware program able to provide statistical estimation of species richness and shared species from samples (Colwell, 2006). Additional information about EstimateS is can be found in an online user manual at: http://viceroy.eeb.uconn.edu/EstimateSPages/EstSUsersGuide/EstimateSUsersGuide.htm.

4. Results

4.1 *Description of the core samples*

Six core samples were selected for testing and analysis from DGR4 borehole at the Bruce Nuclear Site, Ontario. The relative depths of the core samples and the experiments performed are summarized in Table 4.1.

Table 4.1: Description of the core samples used in the study

Sample name	Core depth	Experiments performed
S595	595 mBGS	DNA extraction, PCR, clone libraries, RFLP, sequencing
S643	643 mBGS	DNA extraction, PCR, clone libraries, RFLP, sequencing
S654	654 mBGS	DNA extraction, PCR, clone libraries, RFLP, sequencing
S663	663 mBGS	DNA extraction, PCR, clone libraries, RFLP, sequencing
S704	704 mBGS	DNA extraction, PCR
S801	801 mBGS	DNA extraction, PCR

4.2 *DNA extraction on six core samples and a positive spike*

Extraction of DNA was performed as per section 3.2.2 on core samples listed in Table 4.1. Sensitivity of the DNA extraction procedure was further tested by spiking a rock core with E. coli cells at approximately 30,000 cells per gram of the core. Spiking results are comparable to the DNA concentrations obtained from our low biomass samples. DNA concentrations on the extraction products are summarized in Table 4.2.

Table 4.2: DNA concentrations of core samples and positive spike extraction products

Sample name	DNA (ng/uL)	260 nm / 280 nm	260 nm / 230 nm
S595	5.0	2.52	1.22
S643	3.7	3.28	1.12
S654	4.5	2.29	1.29
S663	4.7	2.04	1.25
S704	7.8	2.27	2.09
S801	5.3	2.51	1.43
(+) spike	4.7	2.87	1.00

4.3 *16S rRNA PCR results on DNA extraction products*

PCR was conducted as per section 3.2.3 on the DNA extraction products from the six core samples and the positive spike. PCR was conducted on samples S595 to S663 together with a positive spike at an earlier time, while samples S704 and S801 were analyzed at a later date.

4.3.1 *PCR optimization with PlatinumTaq, DMSO and BSA*

PCR optimization was performed on DNA extraction products from samples S595 to S663 using higher grade PlatinumTaq, dimethyl sulfoxide (DMSO) and bovine serum albumin (BSA). Results of PCR optimization are illustrated in Figure 4.1.

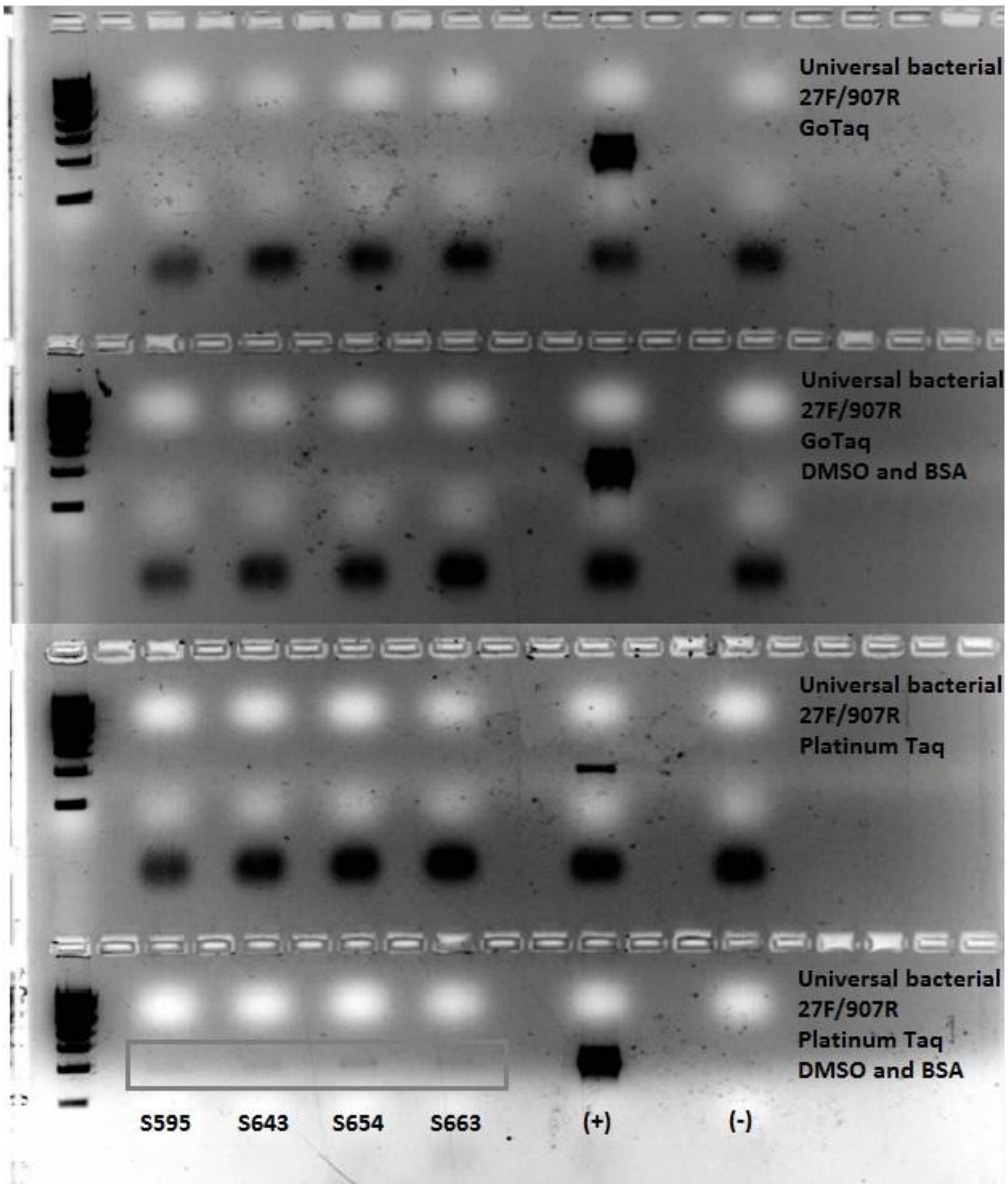

Figure 4.1: PCR optimization results on samples S595, S643, S654, S663 using universal bacterial primers, GoTaq polymerase (top), GoTaq polymerase & DMSO & BSA (middle top), PlatinumTaq (middle bottom), PlatinumTaq & DMSO & BSA (bottom). From left to right: 1 kb DNA ladder (Fisher BioReagents, # BP2582-200), S595, S643, S654, S663, PCR positive control, and PCR negative control. Amplification bands can be observed in the grey box at the bottom left of the image.

PCR optimization results in Figure 4.1 illustrate that an optimal combination of reagents must be added to the PCR reaction in order to obtain 16S rRNA bacterial gene amplification. PCR performed with GoTaq or Platinum Taq alone does not produce any DNA amplification. Same can be said for DMSO and BSA in combination with GoTaq, meaning that DMSO and

BSA alone do not yield any DNA amplification. PlatinumTaq with DMSO and BSA leads to the production of DNA bands at approximately 900 bp. Due to this result, all future PCR reactions were performed with PlatinumTaq, 5% (v/v) DMSO and 0.5 ug/uL (f.c) BSA.

PCR results in Figure 4.1 show that only universal bacterial (27F/907R) primers provided amplification of extracted DNA from S595, S643, S654, and S663. Size of the PCR products using universal bacterial primers was the same for the four samples and the spike sample at approximately 900 bp.

4.3.2 *Core samples S595, S643, S654, S663 and positive spike*

PCR results for samples S595 to S663 and a positive spike are illustrated using an electrophoresis gel image in Figure 4.2.

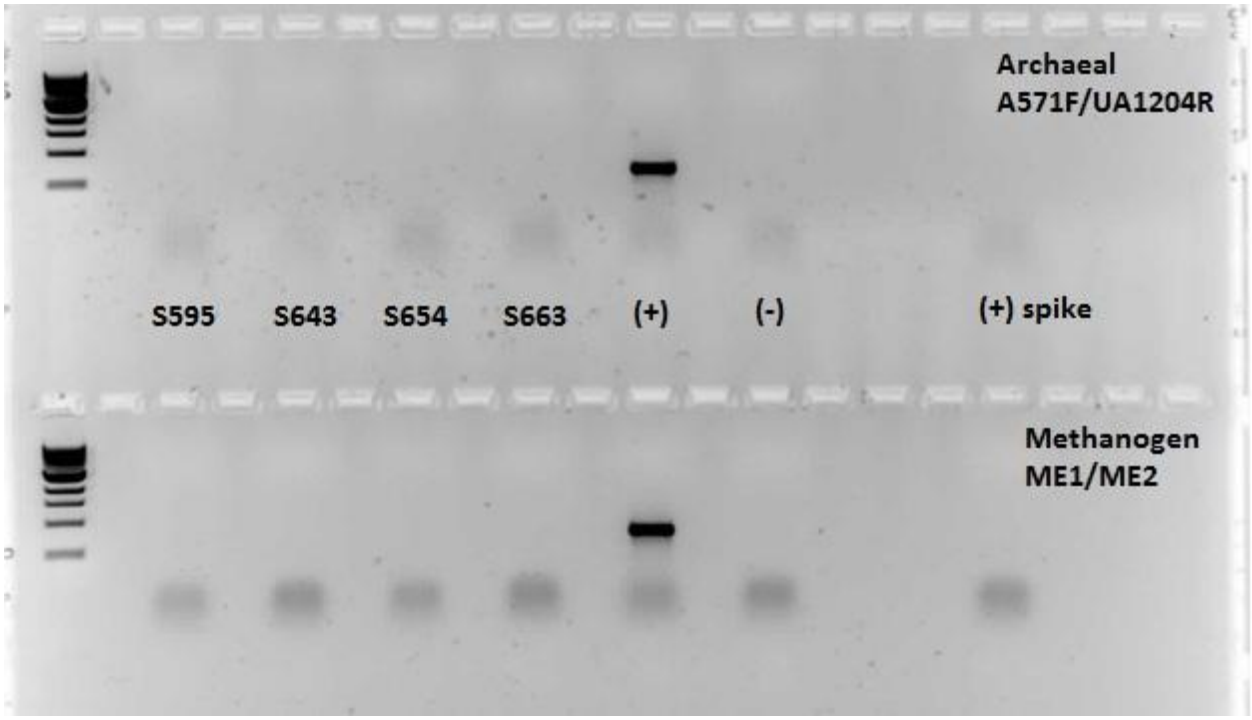

Figure 4.2: PCR results on samples S595, S643, S654, S663 and positive spike performed with universal archaeal (top) and methanogen specific *mcrA* gene (bottom) primers. From left to right: 1 kb DNA ladder (Fisher BioReagents, # BP2582-200), S595, S643, S654, S663, PCR positive control, PCR negative control, and positive spike (with *E.coli DNA*) sample.

PCR using universal archaeal (A571F/UA1204R) and methanogen specific *mcrA* gene (ME1/ME2) primers in Figure 4.3 show no DNA amplification for the four core samples. Positive spike sample with *E. coli* DNA was not amplified with archaeal primers.

4.3.3 *Core samples S704 and S801*

PCR results for samples S704 and S801 are illustrated using an electrophoresis gel image in Figure 4.3.

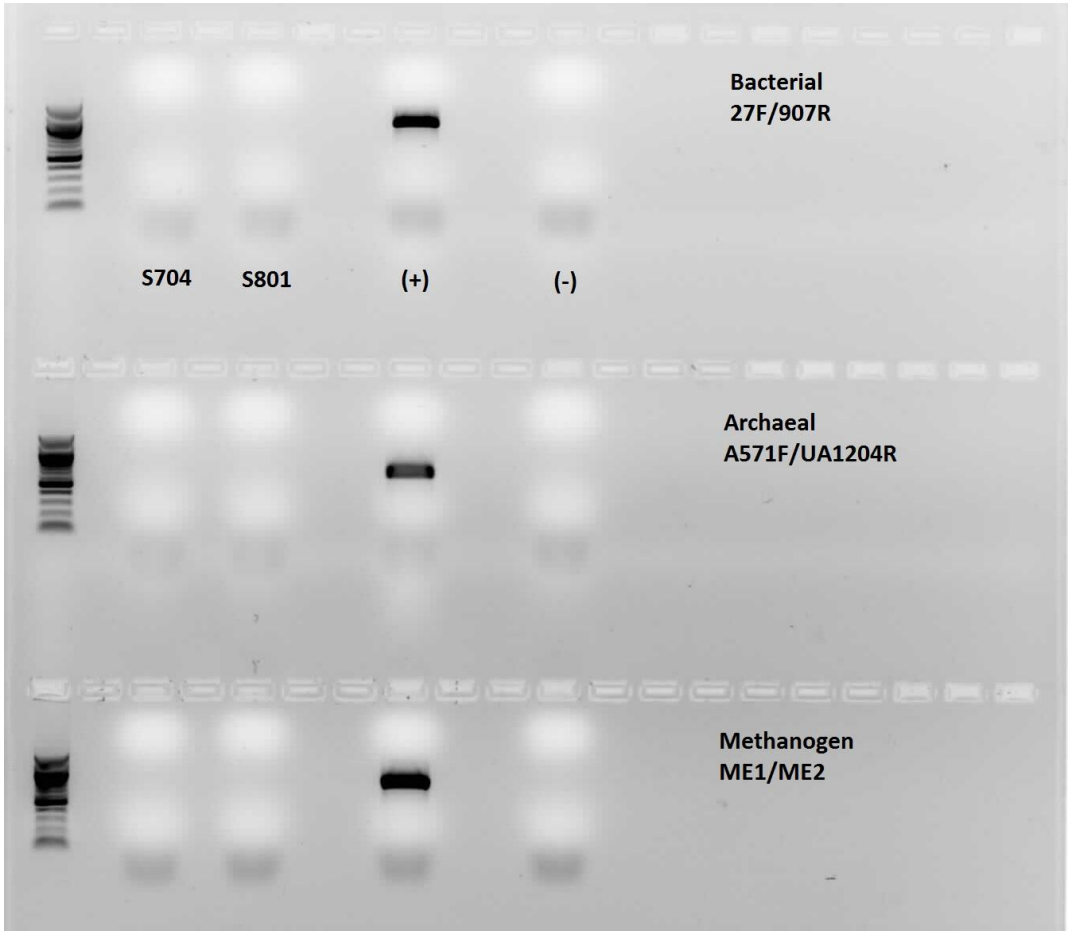

Figure 4.3: PCR results on samples S704 and S801 performed with universal bacterial (top), universal archaeal (middle) and methanogen specific *mcrA* gene (bottom) primers. From left to right: 100 bp DNA ladder (New England BioLabs, # N3231L), S704, S801, PCR positive control, and PCR negative control.

PCR results in Figure 4.3 illustrate that PCR using universal bacterial (27F/907R), universal archaeal (A571F/UA1204R) and methanogen specific *mcrA* gene (ME1/ME2) primers did not yield any DNA amplification for S704 and S801 samples.

4.3.4 *Nanobacterial and nanoarchaeal PCR on core samples S595 to S801*

Six core samples S595 to S801 were tested for the presence of nanobacterial and nanoarchaeal DNA. PCR results can be seen on an electrophoresis gel image in Figure 4.4.

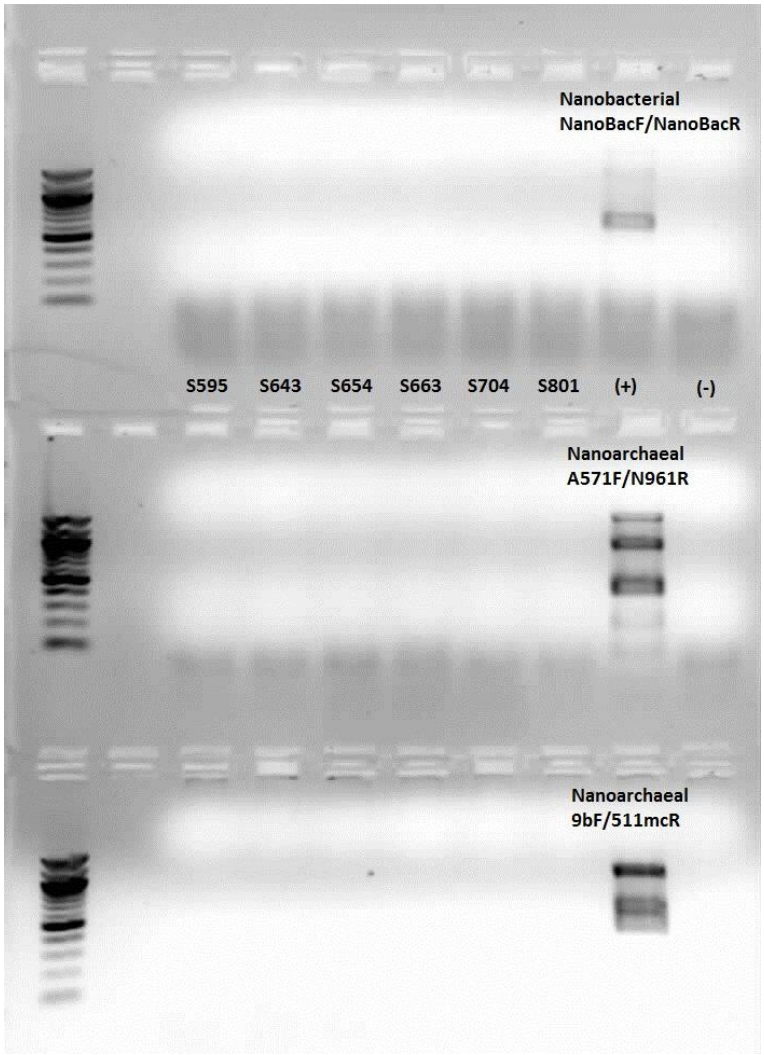

Figure 4.4: PCR results on samples S595 to S801 performed with nanobacterial (top), universal nanoarchaeal set 1 (middle) and universal nanoarchaeal set 2 (bottom) primers. From left to right: 100 bp DNA ladder (New England BioLabs, # N3231L), S595, S643, S654, S663, S704, S801, PCR positive control (*E. coli* or archaeal DNA), and PCR negative control.

PCR using nanobacterial (NanoBacF/NanoBacR) and nanoarchaeal (A571F/N961R; 9bF/511mcR) did not yield any DNA amplification, as seen by the lack of bands for samples S595 to S801. Amplification of bacterial and archaeal positive controls with nano primers is of concern, since only nanobacteria and nanoarchaea should be amplified. Additionally, bands for samples S595 to S663 are not present. This is likely due to different PCR conditions that were used for the nanobacterial and nanoarchaeal DNA amplification and differ from the optimal PCR conditions required to amplify our bacterial DNA.

4.4 *Post gel-extraction DNA concentrations for samples S595, S643, S654, and S663*

Samples S595 to S663 were selected for further DNA fingerprinting analysis and amplified using universal bacterial primers (27F/907R) as per protocol in Section 3.2.3. After amplification of 16S rDNA genes were confirmed on the gel, they were subsequently extracted and DNA was purified for cloning purposes. Concentrations of post-gel extraction DNA for samples S595 to S663 are summarized in Table 4.3.

Table 4.3: Post gel-extraction DNA concentrations for samples S595 to S663

Sample name	DNA (ng/uL)	260 nm / 280 nm	260 nm / 230 nm
S595	13.3	1.80	1.48
S643	9.9	1.64	2.65
S654	8.2	1.68	4.98
S663	11.9	1.54	2.07

4.5 *RFLP analysis on clone libraries S595, S643, S654, and S663*

For clone libraries S595 to S663, 47 clones from each sample were chosen for further DNA fingerprinting analysis using RFLP. M13 PCR was conducted on fresh cell culture and obtained PCR products were digested using EcoRI and AluI restriction enzymes. Digestion results depicting digestion patterns for clone libraries S595 to S663 are discussed in the following sections.

4.5.1 *RFLP results for clone library S595*

RFLP analysis on clone library S595 yielded 7 distinct digestion patterns (A, B, C, D, E, F, G), including 1 unique digestion pattern (B). RFLP patterns for clone library S595 are illustrated in Figure 4.5.

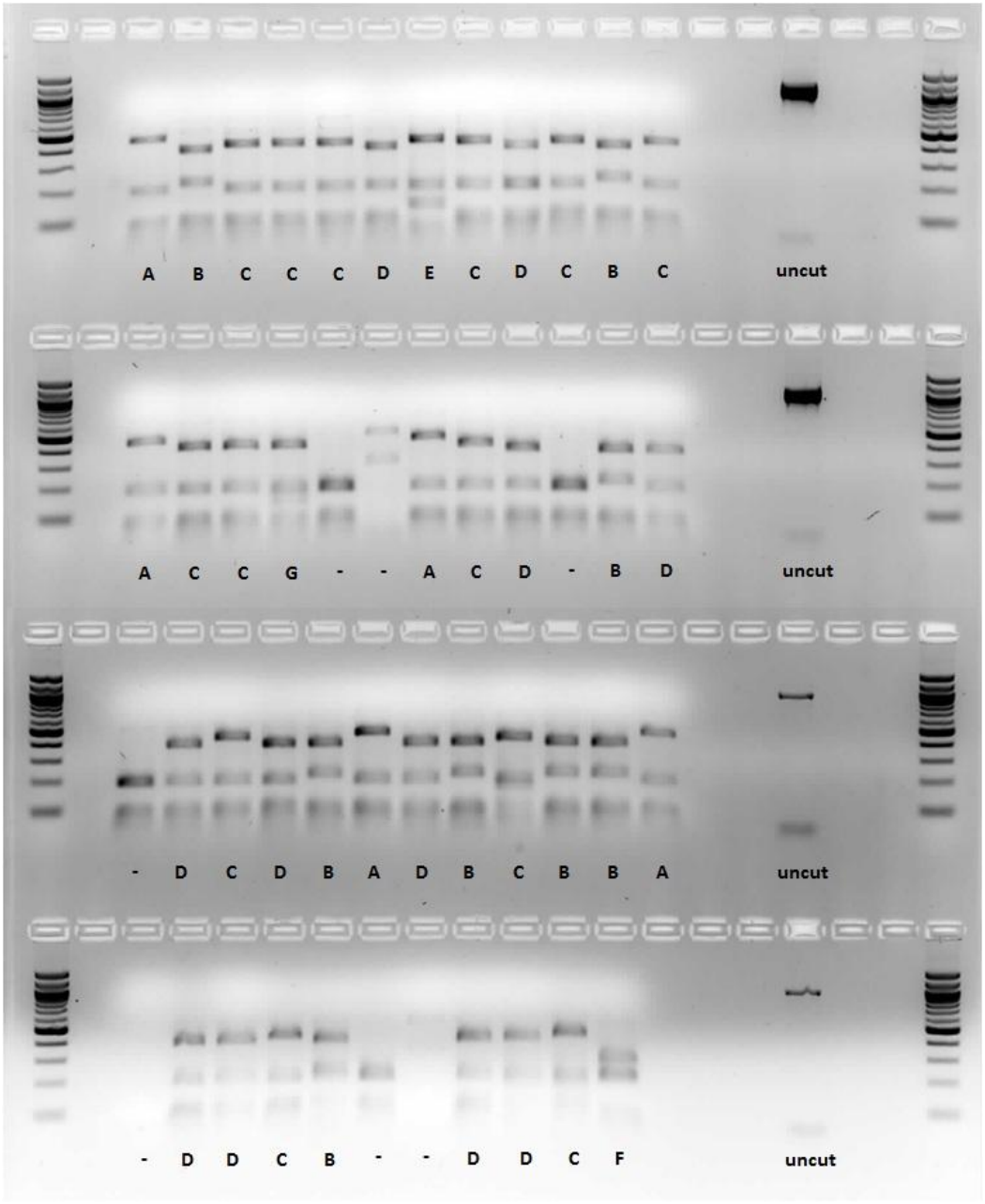

Figure 4.5: RFLP digestion results on clone library S595, performed with EcoRI and AluI restriction enzymes depicting various digestion patterns. Uncut M13 PCR product and 100 bp DNA ladder (New England BioLabs, # N3231L) are included for reference purposes.

4.5.2 *RFLP results for clone library S643*

RFLP analysis on clone library S643 yielded 8 distinct digestion patterns (A, C, D, E, F, G, H, I), including 1 unique digestion pattern (I). RFLP patterns for clone library S643 are illustrated in Figure 4.6.

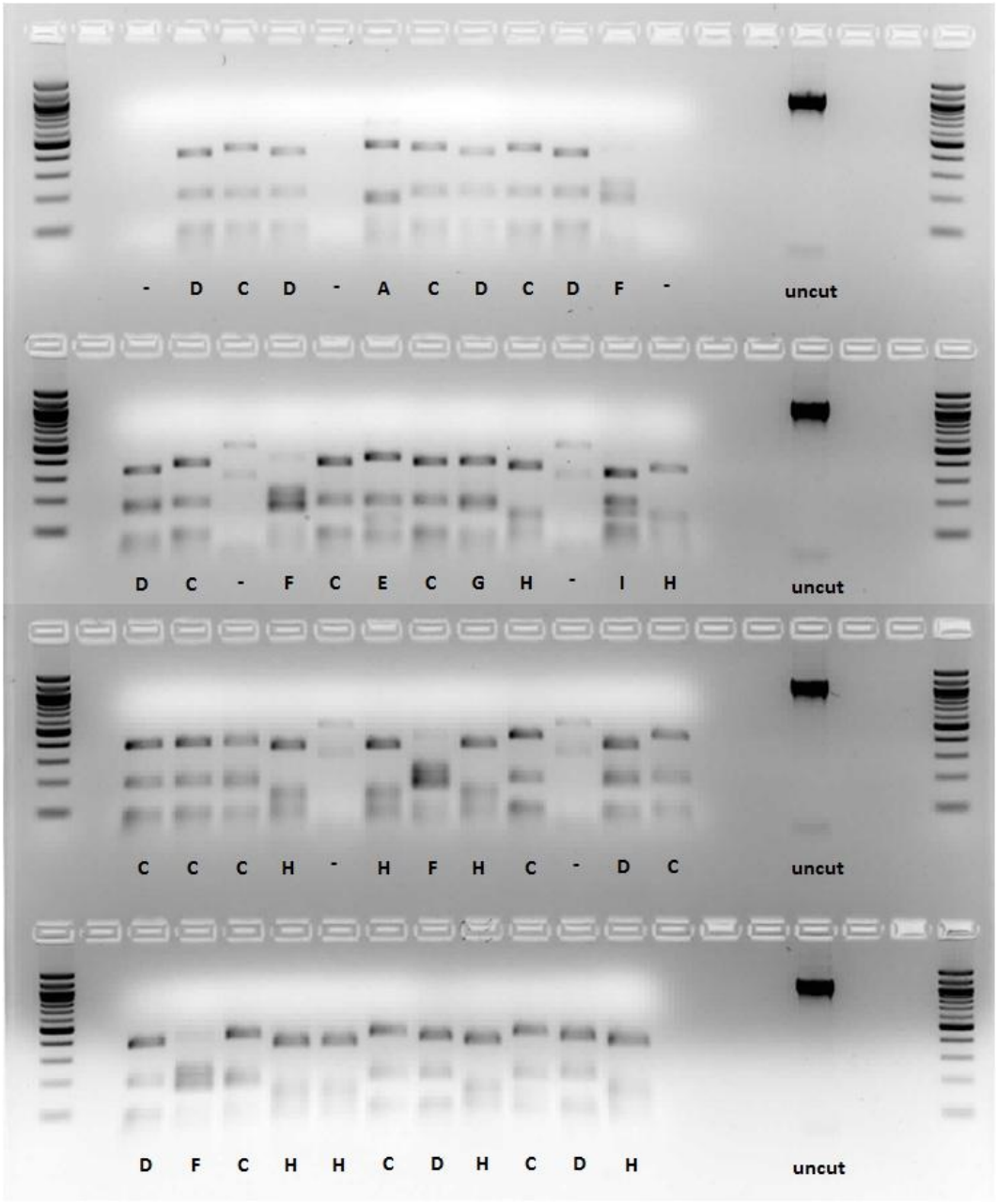

Figure 4.6: RFLP digestion results on clone library S643, performed with EcoRI and AluI restriction enzymes depicting various digestion patterns. Uncut M13 PCR product and 100 bp DNA ladder (New England BioLabs, # N3231L) are included for reference purposes.

4.5.3 *RFLP results for clone library S654*

RFLP analysis on clone library S654 yielded 11 distinct digestion patterns (A, C, D, F, G, H, J, K, L, M, N), including four unique digestion patterns (J, K, L, M). RFLP patterns for clone library S654 are illustrated in Figure 4.7.

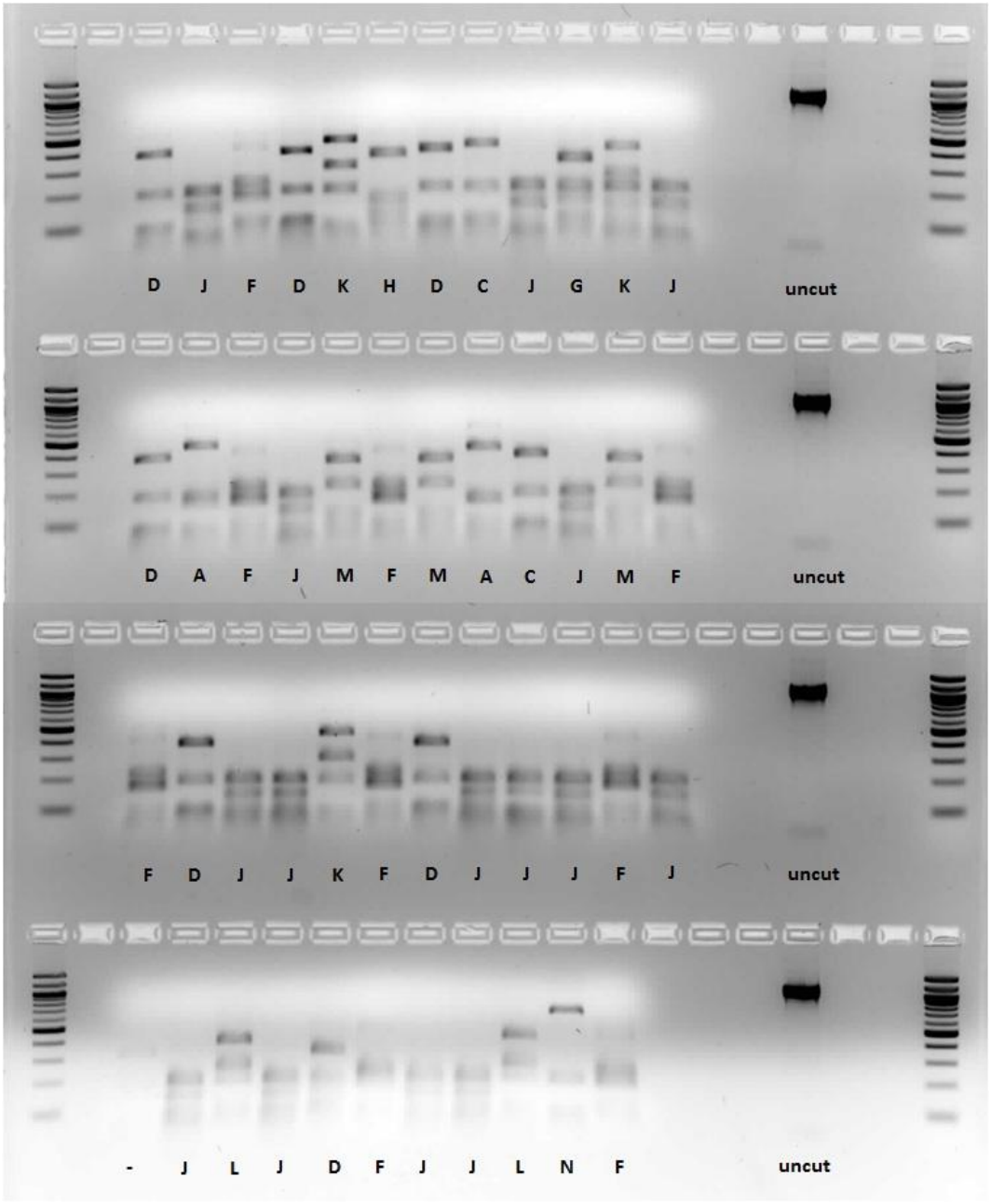

Figure 4.7: RFLP digestion results on clone library S654, performed with EcoRI and AluI restriction enzymes depicting various digestion patterns. Uncut M13 PCR product and 100 bp DNA ladder (New England BioLabs, # N3231L) are included for reference purposes.

4.5.4 *RFLP results for clone library S663*

RFLP analysis on clone library S663 yielded 12 distinct digestion patterns (A, C, D, F, G, N, O, P, Q, R, S, T), including 6 unique digestion patterns (O, P, Q, R, S, T). RFLP patterns for clone library S663 are illustrated in Figure 4.8.

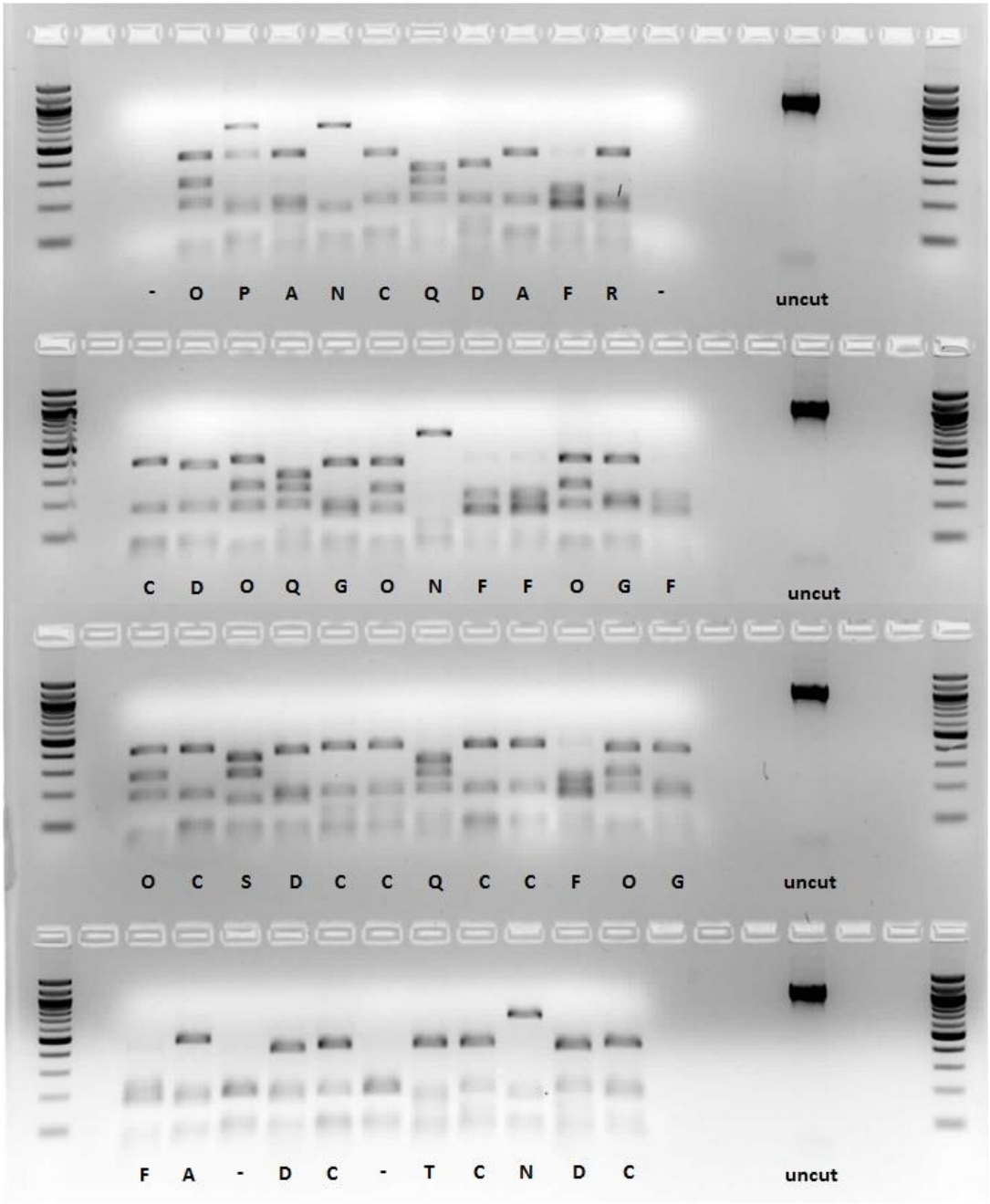

Figure 4.8: RFLP digestion results on clone library, S663 performed with EcoRI and AluI restriction enzymes depicting various digestion patterns. Uncut M13 PCR product and 100 bp DNA ladder (New England BioLabs, # N3231L) are included for reference purposes.

4.5.5 *Summary of RFLP analysis for clone libraries S595 to S663*

Overall, 20 different RFLP profiles (A to T) were obtained for clone libraries S595 to S663. In general, restriction profiles vary across clone libraries and some distinct restriction patterns are unique. Results of restriction profile analysis are summarized in Table 4.4.

Table 4.4: Summary of RFLP analysis on S595 to S663 clone libraries

Clone library	Number of patterns per distinct RFLP profile																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	**Total**
S595	5	8	13	11	1	1	1														**40**
S643	1		14	9	1	4	1	9	1												**40**
S654	2		2	7		9	1	1		15	2	3	3	1							**46**
S663	3		10	5		6	3							3	6	1	3	1	1	1	**43**

*Shaded RFLP numbers represent unique restriction pattern profile

4.6 *Statistical analysis of clone libraries*

Statistical analysis of clone libraries was performed using distinct RFLP profiles of the clones discussed in Section 4.5. For the purpose of the analysis, each distinct RFLP pattern profile was considered to be a distinct operational taxonomic unit (OTU).

4.6.1 *Diversity analysis on clone libraries*

Diversity analysis was conducted on clone libraries in order to differentiate clone libraries based on species richness and abundance. Summary of diversity analysis can be found in Table 4.5. Statistical error for calculations was determined by the difference in total abundance that ranges between 40 and 46. Based on this range, standard deviation of 2.5 or 6% was determined and a confidence interval for rarefaction curves was set at 94%. True diversity of the clone libraries must be discussed by taking statistical error into account, since total abundance values can greatly influence diversity indices.

Table 4.5: Summary of diversity analysis on clone libraries S595 to S663 with standard deviation values of ±6%

Clone library	Simpson's Diversity Index (*D*)	Shannon-Wiener Diversity Index (*H*)	Species Evenness (E_H)	Species richness	Total abundance
S595	0.239 ± 0.014	1.57 ± 0.09	0.811 ± 0.048	7.0	40
S643	0.236 ± 0.014	1.64 ± 0.09	0.788 ± 0.047	8.0	40
S654	0.183 ± 0.010	1.99 ± 0.11	0.828 ± 0.049	11.0	46
S663	0.128 ± 0.007	2.23 ± 0.13	0.898 ± 0.053	12	43

According to Table 4.5, clone library S663 is the most species rich with 12 OTUs and is followed closely by clone library S654 with 11 OTUs. Clone libraries S643 and S654 have lower species richness with 8 OTUs and 7 OTUs respectively.

Simpson's Diversity Index confirms the species richness estimates with most diverse clone library being S663 with a *D* of 0.128, followed by clone library S654 with a *D* of 0.183, and clone libraries S643 and S595 with *D* values of 0.236 and 0.239 respectively. However, *D* values for clone libraries S595 and S643 fall within the statistical error range, and therefore are statistically the same. Shannon-Wiener Diversity Index produces similar results, with higher *H* value of 2.23 belonging to clone library S663, followed by S654 with 1.99, S643 with 1.64 and S595 with 1.57. However, clone libraries S595 and S643 fall within the statistical error range, as do clone libraries S654 and S663, and therefore their diversity should be considered statistically the same. Therefore, *D* and *H* index can only conclusively determine that clone libraries S663 and S654 are more diverse than clone libraries S643 and S595.

Species evenness values are highest for clone library S663 with E_H of 0.898, followed by S654 with 0.828, S595 with 0.811 and S643 with 0.788. By taking standard deviation into account, clone libraries S595, S654 and S663 have the same species evenness, as do clone libraries S643, S595 and S564.

Rarefaction curves were constructed in order to visualize species richness and to assess whether true species diversity was captured though RFLP analysis. Rarefaction curves are illustrated in Figure 4.9.

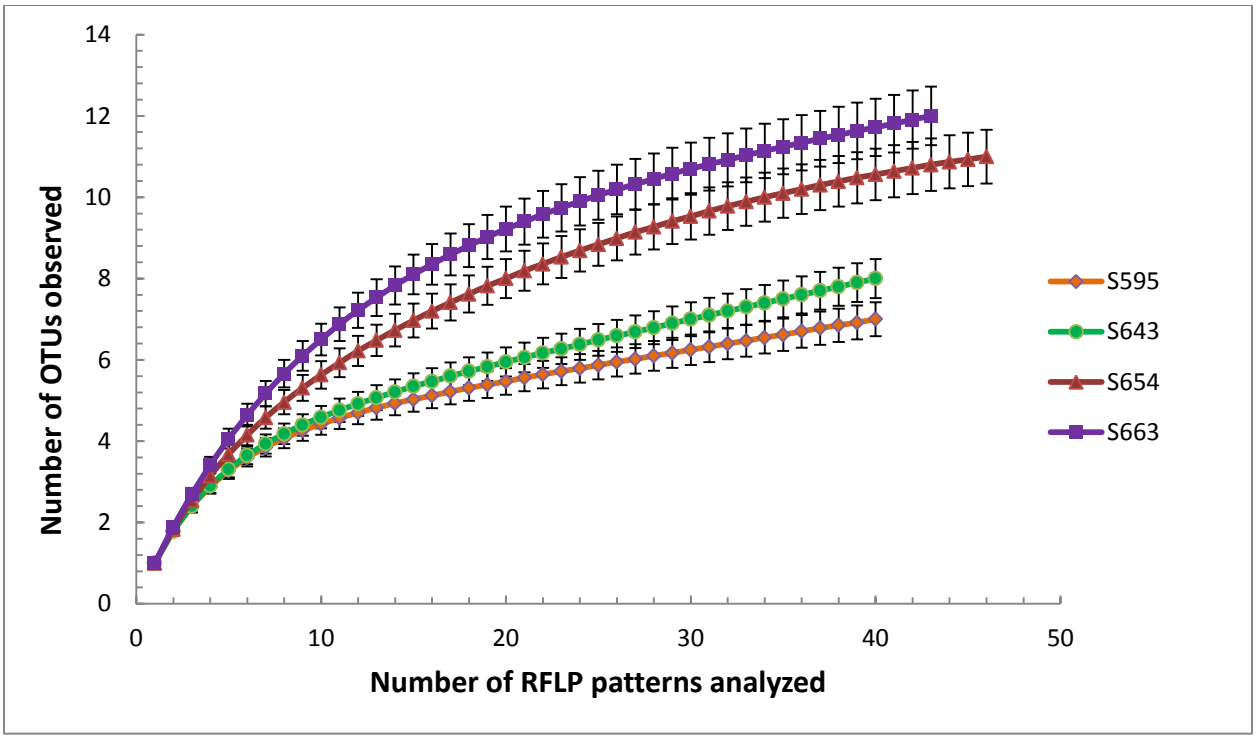

Figure 4.9: Rarefaction curves for clone libraries S595, S643, S653 and S663. Error bars represent 94% confidence interval.

Rarefaction curves indicate that there is a difference in species richness between clone libraries. Confidence intervals overlap for clone libraries S663 and S654 at maximum number of analyzed RFLP patterns, indicating that clone libraries have no difference in species richness. As well, confidence intervals overlap for clone libraries S595 and S643 for the majority of data points, once again indicating that species richness is the same. Therefore, rarefaction curves support results generated by the diversity indices: clone libraries S663 and S654 are more species diverse than clone libraries S643 and S595.

Rarefaction curves demonstrate that true species diversity was likely captured for clone library S654, since the curve begins to plateau at the end. However, for clone libraries S595, S643 and S663, rarefaction curves decrease in slope but do not begin to plateau at the maximum number of RFLP patterns, indicating that true species diversity likely was not captured. If the resolution of the experiment can be improved by the analysis of a greater number of RFLP patterns, it is likely that species richness between clone libraries can be examined more closely and that true species diversity can be captured.

4.6.2 *Similarity analysis on clone libraries*

Clone libraries S595 to S663 were analyzed for patterns of similarity and variability using Bray-Curtis index of similarity and Morisita-Horn index of overlap. Results of the similarity calculations are summarized in Table 4.6. The number of clones selected for the calculation of similarity indices was the same as for the diversity analysis.

Table 4.6: Similarity comparison of clone libraries S595 to S663 using Bray-Curtis index of similarity and Morisita-Horn index of overlap

Clone Library	S595	S643	S654	S663
		Bray-Curtis (*BC*) index of similarity		
S595		0.65	0.302	0.481
S643	0.768		0.372	0.506
S654	0.316	0.360		0.382
S663	0.662	0.686	0.392	
	Morisita-Horn (C_λ) index of overlap			

*Upper right area of the table provides Bray-Curtis index of similarity and bottom right area provides Morisita-Horn index of overlap. Grey-shaded area separates the two indices and colors link indices for the same two clone libraries.

The Bray-Curtis (*BC*) index of similarity demonstrated that similarity values ranged from 0.65 to 0.302 with an average similarity index of 0.449. According to the index, the order of similarity between clone libraries is as follows: S595 and S643 (0.65), S643 and S663 (0.506), S595 and S663 (0.481), S654 and S663 (0.382), S643 and S654 (0.372), S595 and S654 (0.302). The Morisita-Horn (C_λ) index of overlap provided similar results with values ranging from 0.768 to 0.316, with an average index value of 0.531. The trend of similarity for clone libraries was the same: S595 and S643 (0.768), S643 and S663 (0.686), S595 and S663 (0.662), S654 and S663 (0.392), S643 and S654 (0.360), S595 and S654 (0.316). Both *BC* and C_λ similarity indices indicate that most similar clone libraries were S595 and S643 and most dissimilar clone libraries were S595 and S654. Average similarity indices for *BC* (0.449) and C_λ (0.531) show that overall, no clear similarity exists between the clone libraries.

4.7 *Phylogenetic analysis of clone libraries S595 to S663*

Clone libraries for phylogenetic analysis were constructed from data obtained through DNA sequencing of one or two clones of each RFLP pattern profile that was identified in Section 4.5. Number of RFLP pattern profiles and number of sequencing results obtained for each clone

library is summarized in Table 4.7. Sequencing was conducted on bacterial 16S rRNA genes, since archaeal, nanobacterial and nanoarchaeal DNA could not be amplified with PCR. Therefore phylogenetic analysis focuses solely on bacterial 16S rRNA genes.

Table 4.7: Number of RFLP pattern profiles and sequencing results for S595 to S663 clone libraries

16S rRNA clone library	# of RFLP pattern profiles	# of sequencing results obtained
S595	7	10
S643	8	12
S654	11	16
S663	12	18

4.7.1 *Phylogenetic composition of S595 to S663 clone libraries*

Phylogenetic composition of clone libraries was determined using phylogentic data obtained though Seqmatch program from the RDP website. Phylogenetic composition of clone libraries S595 to S663 is illustrated in Figure 4.10.

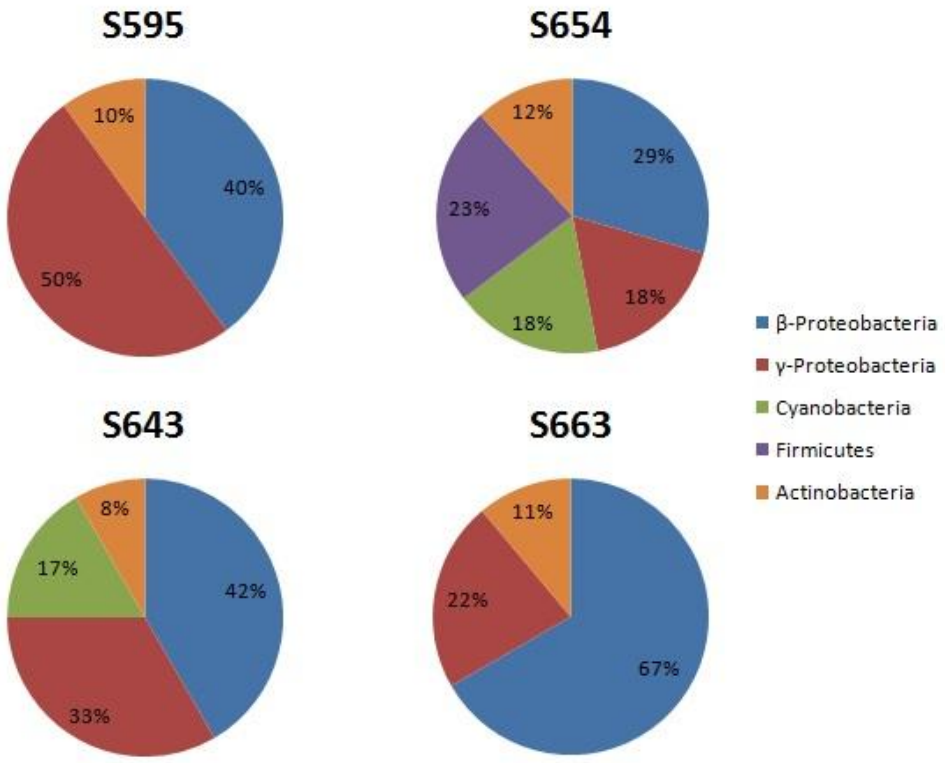

Figure 4.10: Phylogenetic composition of clone library S595, S643, S653 and S663 derived using RDP Seqmatch.

Clone library S595 was composed of 40 clones with 7 distinct RFLP pattern profiles, yielding 10 sequencing results. According to RDP Seqmatch (Figure 4.10) and phylogenetic tree branching (Figure 4.11), three distinct phylogenetic groups were identified: *γ-Proteobacteria* (50%), *β-Proteobacteria* (40%) and *Actinobacteria* (10%). Class of *γ-Proteobacteria* was represented by: an uncultured bacterium (AY860755) with closest match to DNA sequence isolated from environmental sample from Qing-Zang, China; two uncultured proteobacterium, one uncultured *Escherichia* species, and one *Escherichia vulneris*. Class of *β-Proteobacteria* was composed of four uncultured bacterium. Uncultured bacterium (HQ221494) was previously isolated from a supraglacial stream system on the Cotton Glacier in Antarctica. Uncultured bacterium (GQ339248) was previously found to be present in freshwater iron-cycling (Bruun, 2010). Uncultured bacterium (HM068921) was isolated from acid mining lake in Germany (Kampe, 2010). *Actinobacteria* (GU200912) was previously isolated from carbon dioxide rich environment.

Clone library S643 was composed of 40 clones with 8 distinct RFLP pattern profiles that yielded 12 sequencing results. Sequencing results showed that four phylogenetic groups were present: *β-Proteobacteria* (42%), *γ-Proteobacteria* (33%), *Cyanobacteria* (17%) and *Actinobacteria* (8%). Class of *β-Proteobacteria* was composed of two uncultured beta-proteobacterium, *Delftia acidovorans*, uncultured *Delftia* species and *Ralstonia* species. Uncultured beta-proteobacterium (DQ295343) was previously isolated from environmental sample from salt marshes in Jiuduansha wetlands, China (Wang, 2007). Class of *γ-Proteobacteria* was represented by two uncultured bacterium, an uncultured proteobacterium and uncultured *Escherichia* species. Class of *Cyanobacteria* included two uncultured bacterium with the same sequence. Class of *Actinobacteria* contained one uncultured *Corynebacterium* species.

Clone library S654 was composed of 46 clones with 11 distinct RFLP pattern profiles and 17 sequencing results. The phylogenetic composition of the clone library was 5 phylogenetic groups: *β-Proteobacteria* (29%), *Firmicutes* (23%), *γ-Proteobacteria* (18%), *Cyanobacteria* (18%) and *Actinobacteria* (12%). Class of *β-Proteobacteria* included three uncultured bacterium, one uncultured beta-proteobacterium, and *Ralstonia* species. Uncultured bacterium (AB376601) was previously isolated from anaerobic environmental seed sludge sample in Japan (Date, 2009). Class of *Firmicutes* was only found in clone library S654, and was composed of

two uncultured bacterium and two uncultured *Bacilli* bacterium. Uncultured bacterium (AB330606) was previously isolated from methane environment from methophilic fermentation reactor. Class of *γ-Proteobacteria* included uncultured proteobacterium, uncultured *Photorhabus* species and *Shigella boydi*. Class of *Cyanobacteria* was composed of three uncultured bacterium. Class of *Actinobacteria* contained an uncultured actinobacterium and an uncultured *Corynebacterium* species. Uncultured actinobacterium (EF220289) was previously isolated from Antarctic terrestrial habitats (Yargeau, 2007).

Clone library S663 was composed of 43 clones with 12 distinct RFLP pattern profiles, producing 18 sequencing results. Phylogenic composition of clone library included three phylogenetic groups: *β-Proteobacteria* (67%), *γ-Proteobacteria* (22%) and *Actinobacteria* (11%). Class of *β-Proteobacteria* was composed of seven uncultured bacterium, two uncultured beta-proteobacterium, *Comamonas* species, *Delftia* species and *Ralstonia* species. Uncultured bacterium (EF507039) was classified as a biphenyl-utilizing bacteria derived from soil sample. Uncultured bacterium (AB222591) was isolated from environmental sample containing high levels of iron (Ha, 2005). Class of *γ-Proteobacteria* one included uncultured bacterium, two *Acinetobacter* species, and one *Pseudomonas* species. *Acinetobacter* species (EU260166) were previous isolated from oligotrophic lakes (Pondes, 2009), an environment that offers little to sustain life. Class of *Actinobacteria* was composed of one uncultured bacterium and *Kocuria* species. Uncultured bacterium (GU200912) was found in a high concentration carbon dioxide soil environment in Mojave Desert (Nguen, 2011).

Phylogenetic tree was constructed as per methodology in Section 3.3.1 using sequencing data for clone libraries S595 to S663. Phylogenetic tree illustrating relationships between various clones is illustrated in Figure 4.11.

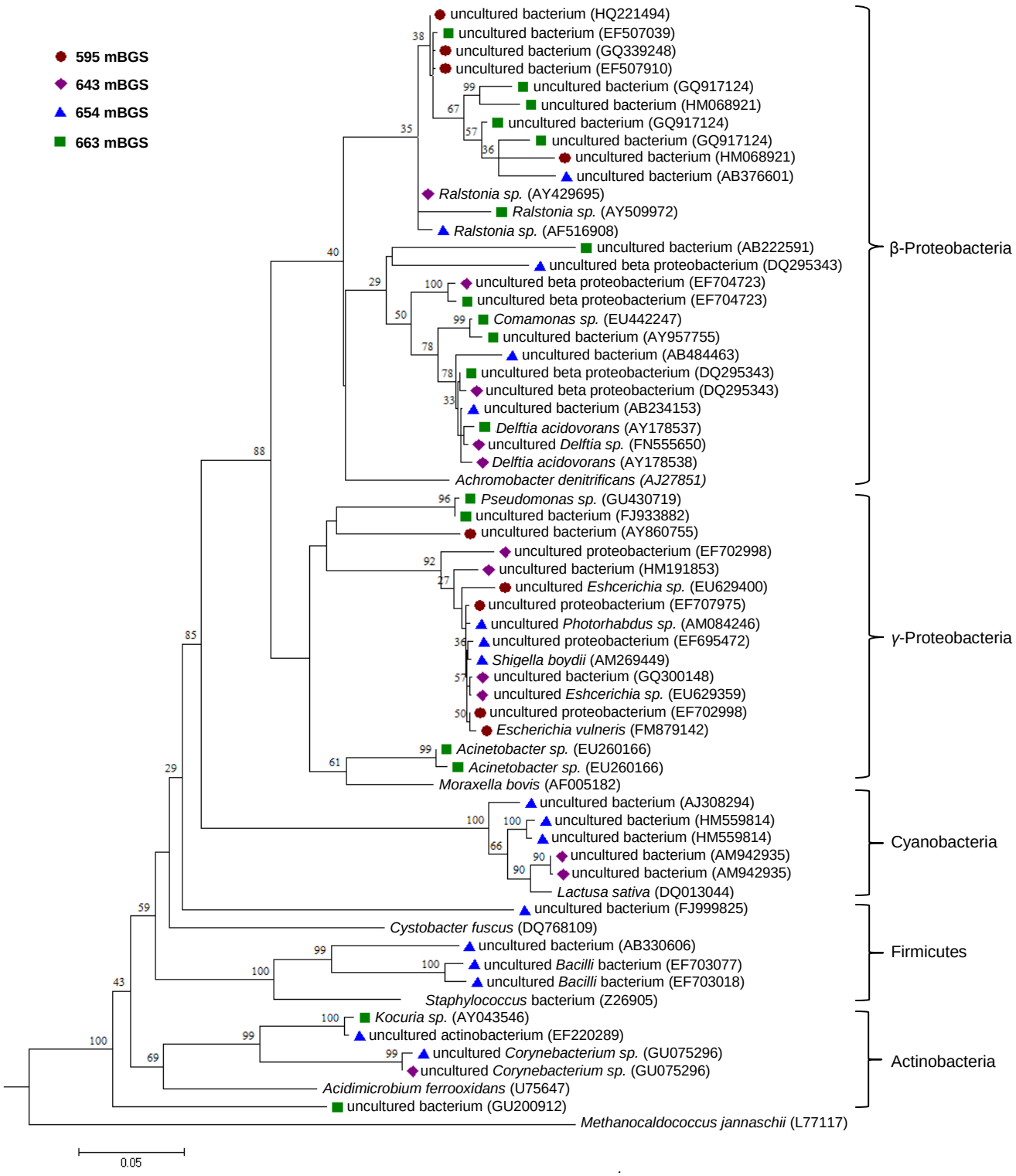

Figure 4.11: Phylogenetic relationships of bacterial 16S rRNA gene sequences obtained from S595, S643, S654 and S663 clone libraries. Clones are represented by their closest relative based on aligned sequence similarity. Phylogentic tree was constructed using RDP Tree Builder tool using distance-based Weighbor tree topology, with minimum analysis of 200 homologous positions of sequence from each clone. *Methanocaldococcus jannaschii* was used as the outgroup. Bootstrap values were calculated based on 100 re-samplings and are reported as a percentage. Scale bar at the bottom left indicates the estimated number of base changes per nucleotide sequence position.

4.7.2 *Phylogenetic bacterial rRNA gene clone type analysis from S595 to S663 clone libraries*

Complete phylogenetic clone type analysis of bacterial rRNA gene clone libraries is summarized in Table 4.7. Each of the aligned sequencing results obtained from a distinct RFLP pattern profile is matched with a related organism obtained through RDP Seqmatch. Clone sequences were compared with related organism though percent similarity, which is percent sequence identity over all pair wise comparable positions. Search was performed on both cultured and uncultured organisms present in the library.

Table 4.8: Phylogenetic bacterial rRNA gene clone type analysis from S595 to S663 libraries

16S rRNA clone library	RFLP profile	Phylogenetic group	Related organism or clone (GenBank accession #)	% Similarity	# of identical clones/ total #
S595	A	*β-Proteobacteria*	uncultured bacterium (GQ339248)	99.9	1/10
	B	*γ-Proteobacteria*	uncultured proteobacterium (EF702998)	100	1/10
	B	*γ-Proteobacteria*	*Escherichia vulneris* (FM879142)	100	1/10
	C	*β-Proteobacteria*	uncultured bacterium (HM068921)	99.8	1/10
	C	*β-Proteobacteria*	uncultured bacterium (HQ221494)	100	1/10
	D	*γ-Proteobacteria*	uncultured proteobacterium (EF707975)	100	1/10
	D	*γ-Proteobacteria*	uncultured *Escherichia sp.* (EU629400)	99.5	1/10
	E	*γ-Proteobacteria*	uncultured bacterium (AY860755)	100	1/10
	F	*Actinobacteria*	uncultured bacterium (GU200912)	98.4	1/10
	G	*β-Proteobacteria*	uncultured bacterium (EF507910)	100	1/10
S643	A	*β-Proteobacteria*	*Delftia acidovorans* (AY178538)	99.7	1/12
	C	*β-Proteobacteria*	*Ralstonia sp.* (AY429695)	99.7	1/12
	C	*γ-Proteobacteria*	uncultured *Escherichia sp.* (EU629359)	100	1/12
	D	*γ-Proteobacteria*	uncultured proteobacterium (EF702998)	100	1/12
	D	*γ-Proteobacteria*	uncultured bacterium (GQ300148)	99.8	1/12
	E	*β-Proteobacteria*	uncultured beta proteobacterium (EF704723)	99.8	1/12
	F	*β-Proteobacteria*	uncultured beta proteobacterium (DQ295343)	99.8	1/12
	F	*β-Proteobacteria*	uncultured *Delftia sp.* (FN555650)	100	1/12
	G	*γ-Proteobacteria*	uncultured bacterium (HM191853)	99.8	1/12
	H	*Cyanobacteria*	uncultured bacterium (AM942935)	100	2/12
	H	*Cyanobacteria*	uncultured bacterium (AM942935)	100	
	I	*Actinobacteria*	uncultured *Corynebacterium sp.* (GU075296)	99.0	1/12
	A	*β-Proteobacteria*	uncultured bacterium (AB376601)	100	1/16

S653	A	*β-Proteobacteria*	uncultured beta proteobacterium (DQ295343)	100	1/16
	C	*β-Proteobacteria*	*Ralstonia sp.* (AF516908)	100	1/16
	C	*γ-Proteobacteria*	*Shigella boydii* (AM269449)	100	1/16
	D	*γ-Proteobacteria*	uncultured *Photorhabdus sp.* (AM084246)	100	1/16
	D	*γ-Proteobacteria*	uncultured proteobacterium (EF695472)	100	1/16
	F	*β-Proteobacteria*	uncultured bacterium (AB484463)	99.6	1/16
	F	*β-Proteobacteria*	uncultured bacterium (AB234153)	100	1/16
	G	*Actinobacteria*	uncultured *Corynebacterium sp.* (GU075296)	98.2	1/16
	H	*Cyanobacteria*	uncultured bacterium (AJ308294)	100	1/16
	K	*Firmicutes*	uncultured *Bacilli* bacterium (EF703018)	100	1/16
	K	*Firmicutes*	uncultured *Bacilli* bacterium (EF703077)	100	1/16
	L	*Firmicutes*	uncultured bacterium (AB330606)	99.7	1/16
	M	*Cyanobacteria*	uncultured bacterium (HM559814)	99.8	2/16
	M	*Cyanobacteria*	uncultured bacterium (HM559814)	99.8	
	N	*Firmicutes*	uncultured bacterium (FJ999825)	98.6	1/16
S663	A	*β-Proteobacteria*	uncultured bacterium (GQ917124)	100	3/18
	C	*β-Proteobacteria*	uncultured bacterium (EF507039)	100	1/18
	C	*β-Proteobacteria*	uncultured beta proteobacterium (EF704723)	99.7	1/18
	D	*β-Proteobacteria*	*Ralstonia sp.* (AY509972)	99.7	1/18
	D	*γ-Proteobacteria*	*Acinetobacter sp.* (EU260166)	100	1/18
	F	*β-Proteobacteria*	uncultured beta proteobacterium (DQ295343)	100	1/18
	F	*β-Proteobacteria*	*Delftia acidovorans* (AY178537)	100	1/18
	G	*γ-Proteobacteria*	*Acinetobacter sp.* (AJ534870)	100	1/18
	G	*β-Proteobacteria*	uncultured bacterium (GQ917124)	100	
	N	*γ-Proteobacteria*	uncultured bacterium (FJ933882)	100	1/18
	N	*β-Proteobacteria*	uncultured bacterium (AB222591)	100	1/18
	O	*β-Proteobacteria*	uncultured bacterium (AY957755)	99.8	1/18
	O	*β-Proteobacteria*	*Comamonas sp.* (EU442247)	100	1/18
	Q	*Actinobacteria*	uncultured actinobacterium (EF220289)	99.9	1/18
	Q	*Actinobacteria*	*Kocuria sp.* (AY043546)	99.8	1/18
	R	*β-Proteobacteria*	uncultured bacterium (GQ917124)	100	
	S	*γ-Proteobacteria*	*Pseudomonas sp.* (GU430719)	100	1/18
	T	*β-Proteobacteria*	uncultured bacterium (HM068921)	99.8	1/18

5. Discussion

5.1 *The question of viable organisms versus ancient DNA*

The results obtained in this study indicate that 16S bacterial rDNA was successfully extracted and amplified from core samples S595-S663. Presence of genetic material however, does not necessarily indicate viable bacteria. Core samples used in this study come from Upper and Middle Ordovician formations deposited approximately 450 Myrs ago. Can DNA survive for such an extended period of time?

DNA can be subject to degradation and might not remain as a suitable template for sensitive reactions such as PCR. One study places the upper age limit of microbial DNA suitable for PCR amplification at 400 to 600 Kyrs (Willersev, 2004). PCR amplification is dependent on the quality of the ability to separate DNA into single strands. Longer preservation time of DNA leads to severe DNA inter- and intra-strand cross-linking, which prevents proper separation of DNA strands during PCR. In addition, prolonged preservation time can lead to other types of DNA damage, including oxidation or double stranded breaks.

Resistance of DNA to degradation is known to be variable, largely due to differences in bacterial hardness and DNA base composition. Endospores are regarded as the hardest of all bacterial types and are known to survive in extreme environmental conditions (Nicholson, 2000). Endospores have been found in a diversity of environments, including oligotrophic conditions, under which they survived by staying dormant and relying on their slow metabolism. *Bacillus* spores of ancient origin were revived, cultured and sequenced from an amber sample thought to be 25 to 50 Myrs old (Cano, 1995). In a different study, Siberian permafrost samples estimated to be 1.8 to 3.0 Myrs old contained viable spore forming beta-proteobacteria, gamma-proteobacteria, and Actinobacteria (Shi, 1997). High content of guanine and cytosine in the DNA of Actinobacteria allows for stronger intrastrand DNA bonding and longer DNA persistence. One of the oldest spore forming bacteria of genus *Bacillus* was isolated and grown from brine sample derived from the inclusion in a 250 Myrs old salt crystal (Vreeland, 2000). These studies show that DNA is able to survive for a longer period of time than estimated by Willersev (2004).

The ability of bacteria to stay viable for millions of years in samples analyzed by Cano (1995) and Vreeland (2000), allowed for DNA survival beyond the maximum theoretical age of

bare DNA. Survival of bacteria in extreme conditions depends on existence or steady supply of nutrients and sufficient growth space. DGR rocks analyzed in this study could support heterotrophic and lithotrophic microorganisms, based on presence of organic compounds and other data mentioned in Section 2.3. Growth space required for microorganism communities in the shale and limestone formations can be of concern, because the porethroat diameter averages from 6 to 20 nm. Diameter of the pores is usually bigger than the porethroat, but is not expected to be significantly larger. Such small pores would not be able to host any viable microorganisms, given the average bacterial size of 0.5 μm. Gas and fluid inclusions in the shale and limestone on the other hand can provide adequate space for microbial growth. Survival of microorganisms in ultra-high-pressure rocks has been previously attributed to their location within fluid and gas inclusions (Zhang, 2005). Microfractures within shales could also serve as a possible growth space as demonstrated by Capuano (1993), however DGR shale samples do not contain connecting microfractures and do not display such high permeability.

Given the old age of Ordovician formations and the comparatively small theoretical age of DNA survival outside of a viable host, conclusion can be drawn that the bacterial DNA isolated from shale and limestone samples must truly come from viable, possibly dormant, bacteria. Metabolic activity of dormant bacteria is very slow, but DNA repair mechanisms still work and keep DNA from degrading.

No archaeal DNA was amplified from our samples, indicating that it was either not present or that molecular techniques were not sensitive enough to amplify the archaeal DNA present. In a typical community, archaea comprises from 5 to 10% of a microorganism population, meaning that archaeal DNA is less abundant. We tested for the presence of archaeal DNA using both archaeal and methanogen specific *mcrA* gene primers, with no successful amplification. These primers were able to amplify archaeal and methanogen specific (*mcrA*) DNA in the positive control of the PCR reaction. Note that these primers yielded successful amplification of archaeal and methanogen specific DNA (*mcrA*) when applied to more recent lake sediment samples, deposited over the last hundred years.

From methane data in Figure 2.5, observation can be made that the methane peak, located in the Collingwood and Cobourg formations, is of biogenic origin due to low $\delta^{13}C$ isotope concentrations. This would indicate that methane was indeed produced by a living methanogenic

archaea. So why do we not detect any archaeal DNA? We need to consider the possibility that methane is not a product of extant archaeal activity, and could be as old as the formations themselves. Archaea has since disappeared and its DNA was not preserved, due to relatively quick DNA degradation time outside a viable host. Discussion of probable past archaeal activity and presently dormant bacterial communities can be found in Section 5.2. These bacterial communities can be up to hundreds of millions years old, and that are seen together with present-day geochemistry in a geological "snapshot".

5.2 *Ancient methanogenesis*

To better understand why the observed methane peak is of ancient origin, geological history of Southern Ontario must be considered, with a closer look at Ordovician limestone and shale. Ordovician carbonates, including Black River Group and Trenton Group, were deposited in a major marine transgression that followed the uplift and erosion of the early Ordovician and Cambrian rocks. The Black River Group and Trenton Group include limestone from Gull River Formation and limestone from Sherman Fall, Cobourg and Collingwood formations, respectively. Middle Ordovician carbonates are characterized by a succession from tidal carbonates to lagoonal carbonates and offshore shallow water and deep shelf carbonates. Ordovician shales were deposited on top of carbonates during the Taconic Orogeny around 440 million years ago. During the orogeny, carbonate platform of the Trenton Group collapsed and was covered with marine clastic shale sediments resulting in the formation of Upper Ordovician shales including Blue Mountain and Georgian Bay. Upper Ordovician shales end with Queenston Formation, the top of which provides evidence of global sea level drop and forms a disconformity. Marine transgressions that followed sea level drop are likely responsible for deposition of formations above the Ordovician. Algonquin arch uplift during the Silurian Acadian Orogeny caused subsidence of the Michigan Basin, restricting marine conditions that caused sea water evaporation and evolution of high salinity (Intera, 2011).

Archaeal activity in the organic-rich shales resulted in the production of methane gas. Restricting marine conditions in the Michigan Basin caused increased brine concentrations due to the evaporation of water. Concentrated brine diffused downwards into underlying formations, reaching methanogenic archaea. The increasing salt concentration caused gradual elimination of methanogens and a decrease in the production of methane (Credit: I. Kenell, NWMO).

Methanogens were completely eliminated when salt concentrations became very high, causing an end to methanogenesis. The biogenic methane was subsequently trapped and largely retained within these low-permeability formations, being deposited on organic-rich substrates (Intera, 2011). This is similar to the observation for in-situ helium produced in the shales that has been retained for over 200 Myrs (Intera, 2011). Methane was likely sorped and trapped on organic rich kerogen layer forming a biogenic methane peak.

Survival of organisms at high salt concentrations requires high energy expenditure to achieve osmotic equilibrium. Methanogenesis from carbon dioxide and hydrogen, or acetate generate very little energy and accumulate organic osmotic solutes that must be removed through energy expenditure (Oren, 2011). One study demonstrated that growth of methanogenic archaea was significantly decreased at 16.6 g/L content of Cl^- (Gomec, 2005), which would represent salinity of 0.5M. The salinity in Ordovician limestone and shale range from 2.5 to 7.5 M (Intera, 2011), which is higher than the tolerance limit of methanogens at 0.5 M. Therefore, methanogens would not be expected to survive in this hypersaline environment. Elimination of methanogen communities along with the sorption of biogenic methane onto organic-rich substrates can account for the presence of biogenic methane peak and lack of methanogen or archaeal DNA. Based on these findings, biogenic methane present within the formations is a product of past archaeal activity.

This study could not obtain evidence of acetogenic or sulphur-reducing bacteria working together with ancient methanogenic archaea to produce biogenic methane. Therefore, no conclusions can be drawn regarding whether methanogens used acetate or carbon dioxide as a carbon source for methane production, or if products of methanogenesis were further used by associated bacteria. One line of evidence separating these processes is the relative location of biogenic methane (Figure 2.5, left) and an enrichment peak for δ^{13} C of carbon dioxide (Figure 2.6, right). If these processes were occurring currently or were dependant, we would expect to see highest concentrations of methane and carbon dioxide at approximately equal depths. However, what we observe is a small carbon dioxide peak with a wide δ^{13} C enrichment above a significantly narrower biogenic methane peak. This can suggest that the biogenic methane was produced in the carbon dioxide δ^{13} C enrichment zone and has migrated down-section to the high organic carbon zone, where it has remained due to sorption onto the organics.

5.3 *Classification of bacteria currently residing in Ordovician shale and limestone*

In deep subsurface environments under extreme conditions, bacteria with extremophilic properties are expected to be found. These bacteria will not use light as a source of energy, and anaerobic conditions will prevent the use of oxygen as a terminal electron acceptor. Additionally, bacteria would need to adjust to high salinity conditions. In this section, bacterial diversity found at four depths within Ordovician shale and limestone is described and correlated with environmental conditions and ecological functions of the bacteria. Differences in bacterial diversity detected in the samples could be a result of constraints imposed by varying nutrient availability or growth conditions specific to those samples.

5.3.1 *Bacterial diversity in Ordovician shale and limestone*

Diversity of bacteria in the Ordovician shale and limestone differs with depth. According to the Simpsons' and Shannon-Weiner diversity indices bacterial diversity at 654 and 663 mBGS (black calcareous shale and argillaceous limestone) is higher than at 595 and 643 mBGS (grey shale and dark grey shale). These indices are based on RFLP profiles and indicate that bacteria at 654 and 663 mBGS contain higher abundance of unique DNA sequences. In addition, species richness was higher at 654 and 663 mBGS. Bray-Curtis and Morisita-Horn indices of similarity indicate that bacteria are most similar at 595 and 643 mBGS and most dissimilar at 595 and 654 mBGS. These results are supported well by phylogenetic analysis (Figure 4.10), where genus of *Proteobacteria* account for 90% and 75% of bacterial composition at 595 and 643 mBGS, respectively, making these two depths most similar. Phylogenetic composition at 654 mBGS is the most diverse, accounting for five classes of bacteria and being most dissimilar from 595 mBGS, which contains only two classes of bacteria. Four bacterial phylum identified were *Proteobacteria, Cyanobacteria, Firmicutes,* and *Actinobacteria*.

5.3.2 *Proteobacteria*

Proteobacteria are characterized as Gram-negative, mainly being facultative or obligate anaerobes and classified as chemoautotrophs or chemoheterotrophs. Chemoautotrophs are able to synthesize their biomolecules from carbon dioxide by using inorganic energy sources, such as hydrogen sulphide, sulphur and hydrogen gas. In contrast hemoheterotrophs can utilize organic carbon and sometimes sulphur as an energy source.

Proteobacteria were present in all four samples and represented the largest fraction of identified bacteria. *Proteobacteria* phylum is divided into six classes, however only *β-Proteobacteria* and *γ-Proteobacteria* were present in our samples. Both classes were found at all four depths, being most dominant at 595 mBGS. Overall, we identified 16 *Proteobacteria* of interest, including genus *Ralstonia*, *Delftia*, *Commonas*, *Shigella*, *Escherichia* and *Acinetobacter*. *Ralstonia* was the most dominant genus of *Proteobacteria*.

5.3.2.1 *Ralstonia species*

We identified *Ralstonia sp.* (HM068921) that was previously isolated from an acidic mine lake in Germany (Kampe, 2010). This acidophile was found at 595 mBGS and 663 mBGS within our samples. Presence of acidophile is not surprising, since the pH of porewater in analyzed DGR4 rock samples ranges from 5 to 6 (Intera, 2011), with lower pH being found at Cobourg and Collingwood formations. It is likely that the lower pH is due to presence of carbon dioxide, which dissolves in porewater producing carbonic acid.

Two *Ralstonia* species (EF507910 and AY509972) identified were previously found in radioactive environments containing uranium and technetium (Michalsen, 2009; Chicote, 2005). These two species were located in samples at 595 and 663 mBGS, which coincides with higher uranium concentrations in Ordovician shale and limestone. It was previously demonstrated that sulphate-reducing-bacteria can reduce uranium and technetium, especially under high sulphate concentrations (Spear, 2000). In DGR geological profile, elevated concentrations of uranium (up to 4 ppm) (Intera, 2011) coincide with high total sulphur. In non-reductive anaerobic conditions, uranium precipitation and biomineralization by *γ-Proteobacteria* was demonstrated to be most active at pH 5 to 7 (Beazley, 2009), similar to the slight acidic conditions found in DGR4 samples.

One detected *Ralstonia* species (EF507039) can utilize biphenyl as an energy source (Leigh, 2007). Biphenyl is an aromatic organic hydrocarbon, which has a colourless crystal appearance, and is commonly found in coal tar, crude oil and natural gas. This species was present at 663 mBGS in argillaceous limestone where oil saturation was measured from 10 to 17% (Figure 2.3). Such high oil concentrations could have been derived from kerogen, which is forms a peak where high concentrations of oil are found. Strain of *Raltonia* was previously

grown in media containing dibenzofuran (Becher, 2000) or carbazole (Waldau, 2009), both derivatives of biphenyl. These bacteria employed enzymes with dioxygenase activity to produce nucleophilic attack on one of the biphenyl rings and break it down into salicylic acid. Additionally, *Ralstonia* species (AY429635) was detected at 643 mBGS, which is able to grow in the presence of propane. This depth is characterized by high concentrations of methane and other gases such as propane that can be also absorbed onto kerogen.

One oligotrophic species of *Ralstonia* (HQ221494) that is able to survive in conditions without organic carbon was identified at 595 mBGS. This depth is characterized by lower abundance of organic carbon and higher concentrations of total sulphur and carbon dioxide. Sulphur can be used as an inorganic source of energy to synthesise biomolecules from carbon dioxide. Additionally, *Ralstonia* species (AB376601) was identified at 653 mBGS that is able to survive in anaerobic environments.

5.3.2.2 *Delftia species*

One of the more interesting bacteria identified in this study was a *Delftia* species (AB234153) that is associated with the metagenome of Scleractinian corals (Yokouchi, 2006). This species was found at 653 mBGS, once again corresponding to a depth characterized by high concentrations of kerogen and organic carbon. Schleractinia are exclusively marine in origin and are found in all regions of the ocean. Colonial and solitary corals, two groups of Scleractinian corals, are found primarily in shallow tropical and deep polar marine waters. Neither of these conditions is currently present in Western Ontario. These conditions were present during Cambrian to Mississippian periods when present Eastern North America was located in tropical latitudes covered by inland seas. An important point is that Sclerantinian corals first appeared around 200 to 250 million years ago, around 200 million years after deposition of Ordovician shales. However, they could have been introduced during the Silurian Acadian Orogeny when the Michigan Basin subsided, causing brine infiltration into underlying formations. Another possibility is that this bacterial species was associated with ancient corals that are currently extinct. Corals have been previously dated back to Early Devonian period, around 416 million years ago (Ferbandez-Matinez, 2010), and even earlier to the Ordovician and Cambrian periods (Munnecke, 2010). Bacteria associated with ancient corals would have similar characteristics to Scleractinian coral bacteria, and thus would be classified as such.

5.3.2.3 *Other species of interest from phylum of Proteobacteria*

Comamonas species (EU442247), previously isolated from deep terrestrial subsurface environments (Brown, 2008), were identified at 663 mBGS. This bacterial species, although previously not being exposed to anthropogenic antibiotic sources, possessed antibiotic resistance.

Shigella boydii (AM269449) was identified at 653 mBGS and has been previously discovered in marine setting in East China Sea. This bacterium is thought to be restricted to the Indian subcontinent being mainly endemic in that region (Yang, 2005); thus, presence of this bacterial species is of interest due to halotolerant properties of the bacteria.

Acinetobacter species (EU260166) was identified at 663 mBGS, and has been previously isolated from oligotrophic lake in Brasil (Pontes, 2009). Oligotrophic lakes are typically unproductive lakes characterized by low nutrient availability. Microorganisms living in these conditions typically have very slow metabolism. Strains of *Acinetobacter* genus have been characterized by their ability to utilize a range of aromatic compounds as a sole source of energy (Thangaraj, 2008; Pessione, 1997). As mentioned previously, 663 mBGS is characterized by elevated concentrations of organic carbon, including aromatic hydrocarbons. Presence of *Acinetobacter* genus in samples from this study is questionable, since it is strictly aerobic and non-spore forming bacterium. It is possible that small amounts of oxygen could be trapped in the pores, allowing for survival of this bacterium.

Escherichia vulneris (FM879142) was identified at 595 mBGS. This particular species was isolated from a coal mine sample and was characterized as a producer of hydrogen sulphide gas. Variants of *Enterobacteriaceae* family are known to produce hydrogen sulphide gas (Barrett, 1987; Bouvet, 1985; Maker, 1974). *Escherichia coli* are known to perform sulphate reduction using sulphate as the ultimate electron acceptor during anaerobic respiration (Fujimoto, 1961). Given high concentrations of total sulphur present at 595 mBGS, it is possible that *Escherichia vulneris* could employ organic compounds as a carbon source producing carbon dioxide (peak for $\delta^{13}C$ isotope of CO_2, Figure 2.6) and producing hydrogen sulphide gas from sulphate as a by-product of its metabolism.

5.3.3 *Cyanobacteria*

Cyanobacteria is a phylum of bacteria that obtains energy for growth through oxygenic photosynthesis. *Cyanobacteria* have been dated back to the Archaean era at around 2.8 billion years ago and are thought to have changed the atmosphere of Earth from reducing to oxidizing conditions (Olson, 2006). Presence of calcified algae and *Cyanobacteria* has also been observed in Ordovician limestone samples (Riding, 2001), indicating their abundance during the Ordovician period. Interestingly, earliest plant species are thought to have evolved from algal mats during the Ordovician period (Strother, 1996). According to the endosymbiotic theory, chloroplasts in plants could have originated from engulfed *Cyanobacteria* (Archibald, 2003).

Cyanobacteria were present at 643 and 654 mBGS sample depths comprising 17% and 18% of bacterial composition, respectively. These depths correspond with the high abundance of organic carbon (Figure 2.4), primarily derived from kerogen. During the Ordovician period, kerogen would be produced from organic matter such as *Cyanobacteria*, algae or early plant species. Presence of *Cyanobacterial* DNA cannot suggest long term survival of *Cyanobacteria* in kerogen without sunlight. Similarly, DNA would not survive for such an extended period of time, due to complete degradation. Interestingly, *Cyanobacteria* has been previously isolated from 529 meter deep ultra-high-pressure rocks (Zhang, 2005). Due to strict contamination control measures, the likelihood of contamination with foreign DNA or bacteria was found to be very low (Zhang, 2005). A possible explanation for the identification of this phylum in DGR4 samples is that *Cyanobacteria* was indeed engulfed by another bacterial cell that could survive in the kerogen layer, using organic carbon as an energy source. This would allow preservation of *Cyanobacterial* DNA in a living cell and account for *Cyanobacteria* presence in these extreme conditions. On the other hand, we could have identified a heterocyst that is produced by *Cyanobacteria*. It contains thick cell walls producing hydrophobic barrier to oxygen and creating microanaerobic environment inside. Heterocysts and *Cyanobacteria* are mutually dependant. They are produced at regular intervals along the *Cynobacterial* filament and contain oxygen-sensitive machinery used for nitrogen fixation. Heterocyst species are able to fix nitrogen into ammonia, nitrate or nitrite. During the conversion of reducing to oxidizing atmosphere of the Earth, oxygen-sensitive machinery was segregated in the heterocyst from the oxygen-evolving photosynthetic apparatus in the neighbouring viable cells (Haselkorn, 1978).

Thick cell wall of the heterocysts could allow for longer periods of survival, although maximum survival age of heterocysts could not be determined.

5.3.4 *Firmicutes*

Firmicutes are a phylum of gram positive bacteria characterized by low guanine and cytosine content within DNA. *Firmicutes* can be rod- or cocci-shaped and have the ability to form endospores that can survive in extreme conditions. Their main genera falls into two classes: *Bacilli* and *Clostridia*. Class of *Bacilli* are obligate or facultative anaerobes, while *Clostridia* are obligate anaerobic fermentative bacteria possessing the ability to reduce sulphite.

Firmicutes were present at 654 mBGS and accounted for 23% of bacterial composition at that depth. Both classes of *Bacilli* and *Clostridia* were identified. *Firmicutes* belonging to class of *Bacilli* were previously identified in a 2.7 billion year old deep subsurface fossil sample (Gerard, 2009). Detection of *Clostridia* species, a fermentative organism, is not surprising since high concentrations of total sulphur and very low concentrations of oxygen are present.

Survival of *Firmicutes* can be attributed to their ability to form spores and their entrapment in kerogen where they would stand a higher chance of preservation.

5.3.5 *Actinobacteria*

Unlike *Firmicutes*, *Actinobacteria* are characterized as having high content of guanine and cytosine, increasing the structural integrity of the DNA. *Actinobacteria* are thought to play a role in the carbon cycle by decomposition of organic matter (Bull, 2005). Some species of *Actinobacteria* can survive in anaerobic conditions (Wintzingerode, 2001; Kampfer, 2000) and form spores (Conn, 2004).

Actinobacteria was present at 595, 643, 654, and 663 mBGS accounting for 10%, 8%, 12%, and 11% of bacterial composition, respectively. *Actinobacteria* were detected at depths characterized by a higher abundance of organic carbon. Previously, it was demonstrated that *Actinobacteria* can use various organic carbon compounds as a source of energy (Groth, 2002), and can inhabit hydrocarbon rich environments under limited oxygen conditions (Bjorklof, 2009). *Actinobacteria* has been previously isolated from extreme environments such as Ultra-High-Pressure rocks in China (Zhang, 2005).

Actinobacteria (GU200912) of *Rubrobacter* genus was present at 595 mBGS, previously isolated from environment with elevated concentration of carbon dioxide (Nguyen, 2011). *Rubrobacter* genus is extremophilic, known for being radiotolerant, giving it the ability to survive extreme gamma-radiation (Chen, 2004). *Actinobacteria* (EF220289) of *Rothia* species was identified at 663 mBGS and was previously isolated from terrestrial habitats in the Antarctic (Yergeau, 2007). This bacterium is an extremophile, able to withstand cold temperatures and survive in the extreme conditions of Antarctica.

5.3.6 *Summary of bacterial diversity*

Bacterial species belonging to the phylum of *Proteobacteria*, *Cyanobacteria*, *Firmicutes,* and *Actinobacteria* were identified. *Proteobacteria* was the most dominant phyla at all depths and included heterotrophic, lithotrophic, acidophilic, radiotolerant, and sulphate-reducing species of bacteria. *Cyanobacteria* were found at depth characterized by high concentrations of organic carbon. We explain their survival using endosymbiotic theory, where *Cyanobacteria* was engulfed by another bacterium able to grow on organic carbon. *Firmicutes* comprised the smallest fraction of identified bacteria and were present in kerogen-rich layers. We attribute survival of *Firmicutes* to their ability to form spores and better preservation in kerogen. *Actinobacteria* was present at all depths, being the second most dominant phyla and included thermophilic and radiotolerant species. Survival of *Actinobacteria* can be attributed to their ability to degrade organic carbon, form spores and possess extremophilic properties.

5.4 *Theoretical model*

The biogenic methane peak and low concentrations of carbon dioxide with a $\delta^{13}C$ isotope enrichment are the result of past biological processes. Methanogenic archaea or acetogenic bacteria were not detected in this study. It is possible that these processes took place during diagenesis, with subsequent trapping of methane and carbon dioxide. Trapped microorganisms responsible for these processes would eventually die off as well, degrading over time and destroying any DNA evidence. This would explain why biogenic signatures are present within the geological profile, but microorganisms responsible for their formation are not detected.

The following is a model, based on geochemical profiles, of bacterial and archaeal activity during the Ordovician time period that lead to the production of biogenic signatures.

The rationale is that heterotrophic acetogenic bacteria utilized organic carbon found in Collingwood/Blue Mountain shale to produce acetate.

$$CH_2O \text{ (organic hydrocarbon)} \xrightarrow{\text{acetogenic bacteria}} CH_3COOH \text{ (acetate)}$$

Acetate was further consumed by heterotrophic sulphur-reducing bacteria (SRB) producing carbon dioxide. Acetate could also have been utilized by lithotrophic methanogenic archaea to produce methane (CH_4) and carbon dioxide (CO_2). Sulphur-reducing bacteria would reduce sulphur to hydrogen sulphide (H_2S) that could react with iron sulphide (FeS), releasing hydrogen gas and producing pyrite (FeS_2). This reaction would account for pyrite content found within Ordovician shale and limestone.

$$CH_3COOH \text{ (acetate)} \xrightarrow[SO_4 \rightarrow H_2S]{\text{SRB}} CO_2$$

$$CH_3COOH \text{ (acetate)} \xrightarrow{\text{methanogenic archaea}} CO_2 + CH_4$$

$$H_2S + FeS \longrightarrow FeS_2 + H_2$$

Resulting carbon dioxide was consumed by lithotrophic archaea with presence of hydrogen gas to produce methane. In the process, carbon dioxide was depleted and its $\delta^{13}C$ isotope signature was enriched.

$$CO_2 + 4H_2 \xrightarrow{\text{methanogenic archaea}} CH_4 + 2H_2O$$

Methane then diffused into underlying formations and absorbed onto zones with high organic carbon content (type III gas-prone kerogen) in the Collingwood and Cobourg formations. Methane located in the Middle Ordovician limestone was produced by a thermocatalytic reaction and is not biogenic in nature.

6. Conclusion

In this study, bacterial DNA was extracted, amplified and characterized from deep subsurface shale and limestone samples. Bacterial composition in the samples was diverse and four distinct phylogenetic groups of *Proteobacteria*, *Actinobacteria*, *Cyanobacteria*, and *Firmicutes* were present. Majority of identified bacterial species were uncultured extremophilic heterotrophs, some with halotolerant properties. We believe that identified bacteria have survived in the rocks for a prolonged period of time by utilizing available energy sources that were trapped during deposition or by assuming a state of dormancy.

Presence of biogenic methane peak in the Upper Ordovician shales can be attributed to the activity of ancient methanogenic archaea. Methanogenesis was active in the Cobourg and Collingwood carbonates up until the Silurian Acadian Orogeny that was responsible for the subsidence of the Michigan Basin. Restricting marine conditions of the Michigan Basin and brine evaporation resulted in increased salt concentrations. Diffusion and subsequent evaporation of salt water into the underlying formations, containing methanogenic archaea, caused gradual elimination of methanogens. As brine concentrations became very high, archaeal communities were fully eliminated. Methane was trapped in the formations and diffused downwards under pressure, eventually depositing on organic-rich kerogen layers and forming a prominent peak.

This study contributes to the knowledge of environmental microbiology through identification of subsurface bacteria and development of DNA extraction and amplification methods. A more in-depth study employing a larger quantity of rock samples can perhaps detect higher species richness. A follow-up study employing culture-dependent methods can further classify bacterial communities that are present with specific focus on correlation between geochemistry and microbiology. Characterization of these bacteria, especially with extremophilic properties, can help in the discovery of novel survival pathways that bacteria can employ in extreme subsurface environments.

7. References

Archibald, J., Keeling, P. 2003. Plant genomes: cyanobacterial genes revealed. *Heredity* **90**: 2-3.

Baker, G.C. *et al*. 2003. Review and re-analysis of domain specific 16S primers. *Journal of Microbiological Methods* **55**: 541-555.

Barrett, E., Clark, M. 1987. Tetrathionate Reduction and Production of Hydrogen Sulphide from Thiosulfate. *Microbiological Reviews* **51** (2): 192-205.

Barton, H.A. *et al*. 2006. DNA Extraction from low-biomass carbonate rock: An improved method with reduced contamination and the low-biomass contaminant database. *Journal of Microbiological Methods* **66**: 21-31.

Beazley, M. 2009. Nonreductive biomineralization of uranium (VI) as a result of microbial phosphatase activity. *Georgia Institute of Technology*.

Becher, D., Specht, M., Hammer, E., Francke, W., Schauer, F. 2000. Cometabolic Degradation of Dibenzofuran by Biphenyl-Cultivated *Ralsonia* sp. Strain SBUG 290. *Applied Environmental Microbiology* **66** (10): 4528-4531.

Bjorklof, K., Karlsson, S., Frostegard, A., Jorgensen, K. 2009. Presence of Actinobacterial and Fungal Communities in Clean and Petroleum Hydrocarbon Contaminated Subsurface Soil. *Open Microbiol J.* **3**: 75-86.

Boone, D., *et al*. 1995. *Bacillus infernus* sp. nov., an Fe(III)- and Mn(II)-Reducing Anaerobe from the Deep Terrestrial Subsurface. International *Journal of Systematic Bacteriology* **45** (3): 441-448.

Bouvet, O., Grimont, P., Richard, C., Aldova, E., Hausner, O., Gabrhelova, M. 1985. Budvicia aquatica gen. nov., sp. nov.: a Hydrogen Sulfide-Producing member of the Enterobacteriaceae. *IJSEM* **35** (1): 60-64.

Bray J.R., Curtis J.T. 1957. An ordination of upland forest communities in southern Wisconsin. *Ecology Monograms* **27**: 325-349.

Brown, M., Balkwill, D. 2009. Antibiotic Resistance of Bacteria Isolated from Deep Terrestrial Subsurface. *Environmental Microbiology* **57** (3): 484-493.

Bruun, A., Finster, K., Gunnlaugsson, H., Nornberg, P., Friedrich, M. 2010. A Comprehensive Investigation on Iron Cycling in a Freshwater Seep Including Microscopy, Cultivation and Molecular Community Analysis. *Geomicrobiology Journal* **27** (1): 15-34.

Bull, A., Stach, J., Ward, A., Goodfellow, M. 2005. Marine actinobacteria: perspectives, challenges, future directions. *Antonie van Leeuwenhoek* **87**: 65-79.

Capuano, R. 1993. Evidence of Fluid Flow in Microfractures in Geopressured Shales. *American Association of Petroleum Geologists Bulletin* **77** (8): 1303-1314.

Cano, R., Boruki, M. 1995. Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. *Science* **268** (5213): 1060-1064.

Casanueva, A. *et al*. 2008. Nanoarchaeal 16S rRNA gene sequences are widely dispersed in hyperthermophilic and mesophilic halophilic environments. *Extremophiles* **12**: 651-656.

Chandler, D., Brockman, F., Bailey, T., Fredrickson, J. 1998. Phylogenetic Diversity of Archaea and Bacteria in Deep Subsurface Paleozol. *Microbial Ecology* **36**: 37-50.

Chen, M. *et al*. 2004. *Rubrobacter taiwanensis* sp. nov., a novel thermophilic, radiation-resistant species isolated from hot springs. *IJSEM* **54**: 1849-1855.

Chicote, E., *et al*. 2005. Isolation and identification of bacteria from spent nuclear fuel pools. *Journal of Industrial Microbiology and Biotechnology* **32** (4): 155-162.

Colwell, R. 2006. *Estimate S*: Statistical estimation of species richness and shared species from samples. Version 8. Persistent URL <purl.oclc.org/estimates>.

Colwell, F. *et al*. 1997. Microorganisms from deep, high temperature sandstones: constraints on microbial colonization. *FEMS Microbiology Reviews* **20** (3-4): 425-435.

Conn, V., Franco, C. 2004. Effect of Microbial Inoculants on the Indigenous Actinobacterial Endophyte Population in the Roots of Wheat as Determined by Terminal Restriction Fragment Length Polymorphism. *Applied and Environmental Microbiology* **70** (11): 6407-6413.

Date, Y., Isaka, K., Ikuta, H., Sumino, T., Kaneko, N., Yoshie, S., Tsuneda, S., Inamori, Y. 2009. Microbial diversity of anammox bacteria enriched from different types of seed sludge in an anaerobic continuous-feeding cultivation reactor. *Journal of Bioscience and Bioengineering* **107** (3): 281-286.

Ferbandez-Martinez, E., Fernandez, L., Mendez-Bedia, I., Soto, F., Mistiaen, B. 2010. Earliest Pragian (Early Devonian) corals and stromatoporoids from reefal settings in the Cantabrian Zone (N Spain). Geologica Acta 8 (3): 301-323.

Folk, R. 1997a. Nanobacteria: surely not figments, but what under heaven are they?. *Journal of Sedimentary Petrology* **63**: 990-999.

Folk, R. 1994. Interaction between bacteria, nanobacteria, and mineral precipitation in hot springs of central Italy. *Géographie Physique et Quaternaire* **48**: 223-246.

Folk, R. 1993. SEM imaging of bacteria and nanobacteria in carbonate sediments and rocks. *Journal of Sedimentary Petrology* **63**: 990-999.

Fujimoto, D., Ishimoto, M. 1961. Sulphate Reduction in *Escherichia coli*. *Journal of Biochemistry* **50** (6): 533-537.

Gerard, E., Moreira, D., Philippot, P., Kranendonk, M., Lopez-Garcia, P. 2009. Modern Subsurface Bacteria in Pristine 2.7 Ga-Old Fossil Stromatolite Drillcore Samples from Fortescue Group, Western Australia. *PLoS ONE* **4** (4): e5298.

Gold, T. 1992. The hot deep biosphere. *Proceedings of the National Academy of Science* **89**: 6054-6049.

Gomec, C., Gonuldinc, S., Eldem, N., Ozturk, I. 2005. Behaviour of an Up-flow Anaerobic Sludge Bed (UASB) reactor at extreme salinity. *Water Science & Technology* **51** (11): 115-120.

Groth, I., Schumann, P., Schutze, B., Augsten, K., Stackebrandt, E. 2002. *Knoellia sinensis* gen. nov., sp., nov., and *Knoellia subterranea* sp. nov., two novel actinobacteria isolated from cave. *IJSEM* **52**: 77-86.

Ha, D., Kusumoto, R., Koyama, T., Fujii, T., Furukawa, K. 2005. Evaluation of the Swim-Bed Attached-Growth Process for Nitrification of Hanoi Groundwater Containing High Levels of Iron. *Japanese Journal of Water Treatment Biology* **41** (4): 181-192.

Haldeman, D., Amy, P., Ringelberg, D., White, D. 1993. Characterization of the Microbiology Within a 21 m^3 Section of Rock from the Deep Subsurface. *Microbial Ecology* **26**: 145-159.

Hales, B.A. *et al*. 1996. Isolation and identification of methanogen-specific DNA from blanket bog peat by PRC amplification and sequence analysis. *Applied Environmental Microbiology* **62** (2): 668-675.

Haselkorn, R. 1978. Heterocysts. *Annual Review Plant Physiology* **29**: 319-344.

Hery, M. *et al*. 2007. Effect of earthworms on the community structure of active methanotrophic bacteria in a landfill cover soil. *The ISME Journal* **2**: 92-104.

Horn, H. 1966. Measurement of "overlap" in comparative ecological studies. *American Naturalist* **100**: 419-242.

Intera Engineering Ltd., "Descriptive Geosphere Site Model", Intera Engineering Ltd. report for the Nuclear Waste Management Organization NMWO DGR-TR-2011-24 R000, 2011, Toronto, Canada.

Kajander , O., Kuronen, I., Akerman, K., Pelttari, A., Ciftcioglu, N. 1997. Nanobacteria from blood, the smallest culturable autonomously replicating agent on Earth. *Proceeding of SPIE* **3111**: 420-428.

Kajander, O., Ciftcioglu, N. 1998. Nanobacteria: An alternative mechanism for pathogenic intra- and extracellular calcification and stone formation. *Proceedings of the National Academy of Sciences USA* **95**: 8274-8279.

Kampe, H., Dziallas C., Grossart H., Kamjunke, N. 2010. Similar Bacterial Community Composition in Acidic Mining Lakes with Different pH and Lake Chemistry. *Microbial Ecology* **60** (3): 618-627.

Kampfer, P. *et al*. 2000. Frigoribacterium faeni gen. nov., sp. nov., a novel psychrophilic genus of the family Microbacteriaceae. *IJSEM* **50** (1): 355-363.

Kim, Tae-Hyoung *et al*. 2011. Detection of Nanobacteria in Patients with Chronic Prostatitis and Vaginitis by Reverse Transcriptase Polymerase Chain Reaction. *Korean Journal of Urology* **52**: 194-199.

Kotelnikova, S., Pedersen, K. 1998. *FEMS Microbiology Ecology* **26** (2): 121-134.

Krumholz, L., McKinley, J., Ulrich G., Sulfita, J. 1997. Confined subsurface microbial communities in Cretaceous rock. *Nature* 386: 64-66.

Krumholz, L. 2000. Microbial communities in the deep subsurface. *Hydrogeology Journal* **8**: 4-10.

Leigh, M.B., *et al*. 2007. Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). *ISME* **1**: 134-148.

Maker, M., Washington, J. 1974. Hydrogen Sulphide-Producing Variants of Escherichia coli. *Applied Microbiology* **28** (2): 303-305.

Mastrapa, R. 2001. Survival of bacteria exposed to extreme acceleration: implications for panspermia. *Earth and Planetary Science Letters* **189** (1-2): 1-8.

Michalsen, M., *et al*. 2009. Treatment of Nitric Acid-, U(VI)-, and Tc(VII)-Contaminated Groundwater in Intermediate-Scale Physical Models of an In Situ Biobarier. *Environmental Science and Technology* **43** (6): 1952-1961.

Morisita, M. 1959. Measuring of interspecific association and similarity between communities. *Memoir Faculty of Science, Kyushu University, Ser E. Bio.* **3**: 65-80.

Munnecke, A., Calner, M., Harper, D., Servais, T. 2010. Ordovician and Silurian sea-water chemistry, sea level, and climate: A synopsis. *Palaeogeography, Palaeoclimatology, Palaeoecology* **296** (3-4): 389-413.

Nelson, P. 2009. Pore-throat sizes in sandstones, tight sandstones, and shales. *American Association of Petroleum Geologists Bulletin* **93** (3): 329-340.

Nguen, L., Buttner, M., Cruz, P., Smith, S., Robleto, E. 2011. Effects of Elevated Atmospheric CO(2) on Rhizosphere Soil Microbial Communities in a Mojave Desert Ecosystem. *Journal of Arid Environments* **75** (10): 917-925.

Nicholson, W., Munakata, N., Horneck, G., Melosh, H., Setlow, P. 2000. Resistance of Bacillus Endospores to Extreme Terrestrial and Extraterrestrial Environments. *Microbiology and Molecular Biology Reviews* **64** (3): 548-572.

Olson, J. 2006. Photosynthesis in the Archean era. *Photosynthesis Research* **88** (2): 109-117.

Oren, Aharon. 2011. Thermodynamic limits of microbial life at high salt concentrations. *Environmental Microbiology* **13** (8): 1908-1923.

Pedersen, K. 1997. Microbial life in deep granitic rock. *FEMS Microbiology Reviews* **20**: 399-414.

Pessione, E., Giunta, C. 1997. Acinetobacter radioresistens metabolizing aromatic compounds. 2. Biochemical and microbiological characterization of the strain. *Microbios* **89** (359): 105-117.

Pontes, D., Pinheiro, F., Lima-Bittencourt C., Guedes, R., Cursino, L., Barbosa, F., Santos, F., Chartone-Souza, E., Nascimento, A. 2009. *Microbial Ecology* **58** (4): 762-772.

Riding, R., Fan, J. 2001. Ordovician calcified algae and cyanobacteria, Northern Tarim Basin subsurface, China. *Palaeontology* **44** (4): 783-810.

Roszak, D., Colwell, R. 1987. Survival Strategies for Bacteria in the Natural Environment. *Microbiological Reviews* **51** (3): 365-379.

Shi, T., Reeves, R., Gilichinsky, D., Friedmann, E. 1997. Characterization of viable bacteria from Siberian permafrost by 16S rDNA sequencing. *Microbial Ecology* **33** (3): 169-179.

Spear, J., Figueroa, L., Honeyman, B. 2000. Modelling reduction of Uranium U(VI) under Variable Sulphate Concentrations by Sulphate-Reducing Bacteria. *Applied Environmental Microbiology* **66** (9): 3711-3721.

Strother, P., Al-Hajri, S., Traverse, A. 1996. New evidence for land plants from the lower Middle Ordovician of Saudi Arabia. *Geology* **24** (1): 55-58.

Thangaraj, K., Kapley, A., Purohit, H. 2008. Characterization of diverse Acinetobacter isolates for utilization of multiple aromatic compounds. *Bioresource Technology* **99** (7): 2488-2494.

Thomas-Keprta, K. *et al*. 1997. The search for terrestrial nanobacteria as possible analogs for purported Martian nanofossils in Martian meteorite ALH84001. Conference Paper, 28th Annual LunKand Planetary Science Conference, p. 435.

Vreeland, R., Rosenzweig, W., Powers, D. 2000. Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. *Nature* **407**: 897-900.

Waldau, D., Mikolasch, A., Lalk, M., Schauer, F. 2009. Derivatization of bioactive carbazoles by the biphenyl-degrading bacterium *Ralstonia* sp. strain SBUG 290. *Applied Microbiology and Biotechnology* **83** (1): 67-75.

Wang, M., Chen, J., Li, B. 2007. Characterization of Bacterial Community Structure and Diversity in Rhizosphere Soils of Three Plants in Rapidly Changing Salt Marshes Using 16S rDNA. *Pedosphere* **17** (5): 545-556.

Whitman, W., Coleman, D., Wiebe, W. 1998. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America* **95** (12): 6578-6583.

Willersev, E., *et al*. 2005. Long-term persistence of bacterial DNA. *Current Biology* **14** (1): R9-R10.

Wintzingerode, F. *et al*. 2001. Salana multivorans gen. nov., sp. nov., a novel actinobacterium isolated from an anaerobic bioreactor and capable of selenate reduction. *IJSEM* **51** (5): 1653-1661.

Wolda, H. 1981. Similarity Indices, Sample Size and Diversity. *Oecologia (Berl)* **50**: 296-302.

Yang, F., *et al*. 2005. Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. *Nucleic Acids Research* **33** (19): 6445-6458.

Yergeau, E., Newsham, K., Pearce, D., Kowalchuk, G. 2007. Pattern of bacteria diversity across range of Antarctic terrestrial habitats. *Environmental Microbiology* **9** (11): 2670-2682.

Yokouchi, H., Fukuoka, Y., Mukoyama, D., Calugay, R., Takeyama, H., Matsunaga, T. 2006. Whole-metagenome amplification of microbial community associated with scleractinian coral by multiple displacement amplification using phi29 polymerase. *Environmental Microbiology* **8** (7): 1155-1163.

Zhang, G., Dong, H., Xu Z., Zhao, Z., Zhang, C. 2005. Microbial Diversity in Ultra-High-Pressure Rocks and Fluids from the Chinese Continental Scientific Drilling Project in China. *Applied and Environmental Microbiology* **71** (6): 3213-3227.

8. Appendix – DNA extraction using PCI

Prior to successful DNA extraction from core samples 595 to 663, multiple attempts were made to extract DNA using Phenol-Chloroform-Isoamyl Alcohol (PCI) procedure. This procedure did not yield any successful bacterial or archaeal DNA amplification from a range of samples.

8.1 Experimental Set-up

In order to test for the efficiency of DNA extraction procedure using PCI, an experiment was conducted using the set-up illustrated in Table 8.1. Rock sample at 595 mBGS (sample A and B) and Endostromatolite rock sample (sample C) known to have been exposed to room bacterial contaminants for more than a year were chosen for the analysis. Rock sample 595 was analyzed as is (sample A) and also by the addition of viable *E.coli* cells into the crucible prior to the crushing or rock core sample (sample B). *E.coli* cells were added for the purpose of observing the matrix effect of the sample and to test the low-end efficiency of PCI extraction procedure.

Table 8.1: Sample description for DNA extraction procedure using PCI

Sample	A	B	C
Sample Description	S595 rock powder	S595 rock powder + *E.coli* cells (10,000 cells/g)	Endostromatolite rock sample, previously exposed to room and hand bacterial contaminants
Dry rock powder weight	9.98 g	9.73 g	9.48 g

8.2 DNA extraction procedure using PCI

Samples were prepared for DNA extraction using procedures as described in Section 3.2.1. All reagent volumes given below were added as a certain volume per 1 gram of rock powder. Rock mass was determined by taking the mass of rock powder once rock powder was transferred from crucible into a tube. Tubes were weighted prior to the start of the procedure to determine the empty weight. All subsequent steps were performed under sterile conditions.

In a laminar flow hood, 1 mL/g of 2x-AE buffer and 0.525 mL/g of lysozyme were added to the 50 mL Falcon tubes containing washed rock powder. Tubes were capped, briefly vortexed, and incubated in water bath at 37^{o}C for 30 minutes.

After the incubation, 0.275 mL/g of Proteinase K (10mg/mL) and 38.6 mL/g of SDS (20%) were added to the mixture in the tubes. Tubes were subsequently vortexed for 30 seconds and subsequently incubated at 50^{o}C for 30 minutes. Subsequently, 0.650 mL/g of SDS (20%) was added to each tube.

To perform separation of DNA from other cellular debris, mixture containing phenol-chloroform-isoamyl alcohol (pH = 7.9) with volumetric ratio of 25:24:1 was added at 1:1 v/v to the rock powder mixture within the tubes. If 25 mL of rock powder mixture was present, then 25 mL of PCI mixture was added.

Tubes were subsequently vortexed using falcon tube adapters at medium setting for 5 minutes, followed by vortex at high setting for 2 minutes. Tubes were centrifuged at 4,000 for 10 minutes at 4^{o}C. During this step, extracted DNA would typically end up in the upper aqueous phase and cell debris would be found in the lower organic phase.

Aqueous supernatant from each tube was carefully transferred into a new sterile 50 mL Falcon tube, without disturbing the lower organic layer. Chloroform was added to the tube containing the supernatant at a volumetric ratio of 1:1, in order to wash the supernatant and to achieve better separation of organic and aqueous phases. Tubes were centrifuged at 4,000 RPM for 10 minutes at 4^{o}C.

Aqueous supernatant from each tube was carefully transferred into a new sterile 50 mL Falcon tube, without disturbing the lower organic phase. Supernatant volume in each tube was measured and 3.0 M Sodium Acetate and 99% Ethanol were added at respective volumetric ratios to the supernatant at 1:10 and 2.5:1.

Tubes were briefly vortexed and DNA was precipitated at -20^{o}C overnight (12+ hours) in order to achieve the best possible precipitation of DNA. Subsequently, tubes were centrifuged at 4,000 RPM for 2 hours at 4^{o}C.

Ethanol was removed from the tubes by careful decanting without disturbance of the pellet (later on determined to be a combination of salts and little if any DNA). Pellet was washed by adding 1 mL of cold 70% Ethanol to the tube. Tubes were subsequently centrifuged at 4,000 RPM for 45 minutes at 4^{o}C.

Supernatant containing Ethanol was carefully decanted and pellet was air-dried within the laminar flow hood for approximately 30 minutes or until no visible traces of ethanol were present.

Pellet was subsequently dissolved in 10 mM Tris solution (pH=7.5) or molecular grade water and stored at -20^{o}C for further analysis.

8.3 DNA amplification using Polymerase Chain Reaction

PCR reactions were carried out in a final volume of 25 uL using Invitrogen Platinum Taq DNA polymerase (# 10966). All PCR reactions were prepared in the laminar flow hood using aerosol resistant filter tips. Laminar flow hood and all the necessary equipment were sterilized using 70% ethanol, 10% bleach and UV lights.

Specific PCR conditions were used in order to try the amplification of extracted DNA: 2.5 uL of 10X PCR Buffer, 0.2 mM dNTP, 3.0 mM, $MgCl_2$, 1 uM of each primer, 1.0 unit of Invitrogen Platinum Taq DNA Polymerase, 1 uL of final extracted DNA product (at 1:1, 1:10 or 1:100 DNA template concentrations), and PCR grade water to a final volume of 25 uL.

In order to amplify bacterial DNA, we used universal bacterial primers 27F (5'- AGA GTT TGA TCM TGG CTC AG -3') and 907R (5'- CCG TCA ATT CMT TTR AGT TT -3'), known to successfully amplify bacterial 16S rRNA gene fragments from environmental samples (Hery *et al*. 2007). To test for the presence of archaeal DNA, we used universal archaeal primers A571F (5'- GCY TAA AGS RIC CGT AGC -3') and UA1204R (5'- TTM GGG GCA TRC IKA CCT -3'), which were shown to amplify archaeal 16S rRNA gene fragments from all types of archaea (Baker *et al*. 2003). We also tested for the presence of genes encoding for the alpha subunit of the methyl coenzyme-M reductase, responsible for the last step of methanogenesis using primer pair ME1 (5'- GCM ATG CAR ATH GGW ATG TC -3') and ME2 (5'- TCA TKG CRT AGT TDG GRT AGT -3') (Hales *et al*. 1996). For all genes of interest in this study, the

thermocycler program was set to 94^{o}C for 10 minutes initial denaturation, followed by 35 cycles of 94^{o}C for 30 seconds, 48.8^{o}C for 60 seconds, and 72^{o}C for 60 seconds with a final extension at 72^{o}C for 5 minutes.

PCR products were loaded onto 1% agarose gel with ethidium bromide staining and separated by electrophoresis at 80V for 50 minutes. DNA bands were visualized using a Fusion Fx5 Gel DOC (Vilbert Lourmat). PCR products were frozen at -20^{o}C for subsequent analysis.

8.4 DNA amplification results

Figures 8.1 to 8.3, demonstrate DNA amplification results with universal bacterial (27F/907R), universal archaeal (A571F/UA1204R) and methanogen specific *mcrA* gene (ME1/ME2) primers. The only DNA amplified was from a sample 595 with *E.coli* spike that generated bands at 1:1 and 1:10 concentrations of PCR template (Figure 8.1). DNA amplification using archaeal specific primers (Figure 8.2) and methanogen specific primers (Figure 8.3) was not successful.

DNA amplification using PCI procedure likely failed due to two primary reasons. First, PCI method might not be sensitive enough to extract DNA from a low biomass environment such as our samples. Second, it is possible that organic and inorganic impurities introduced during the DNA extraction procedure could inhibit PCR reaction. For these reasons, PCI extraction procedure was not used any further.

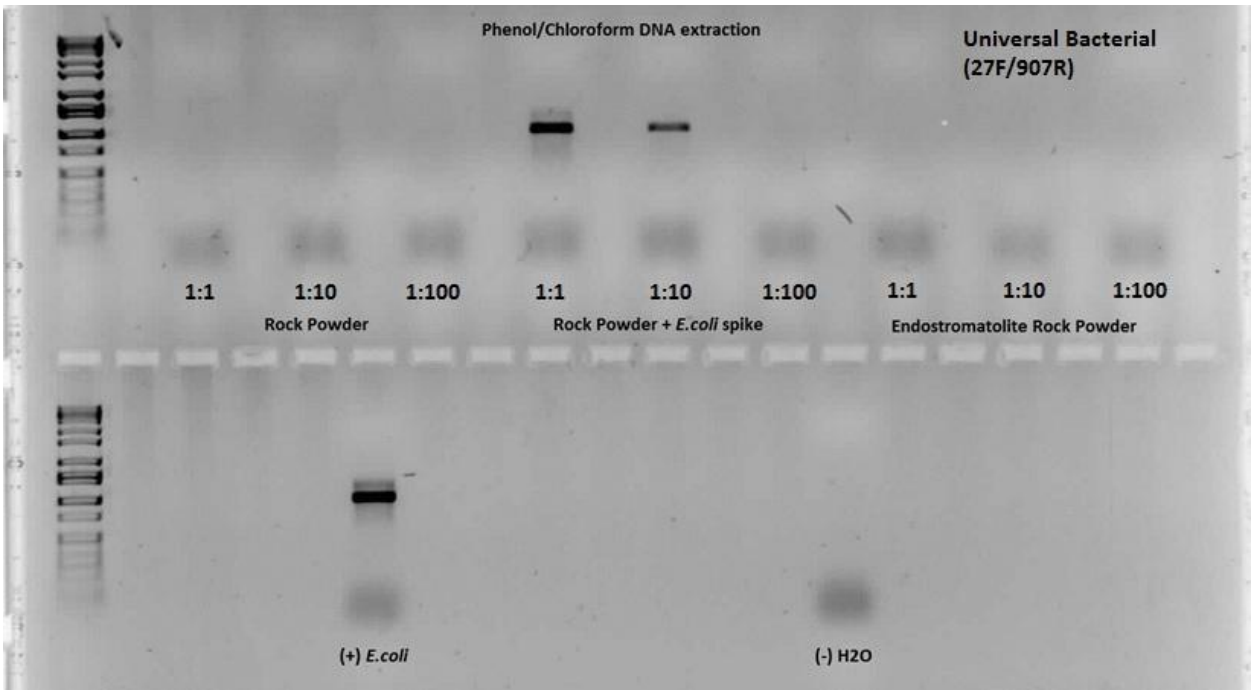

Figure 8.1 PCR results (1:1, 1:10 and 1:100 DNA template concentrations) on samples S595, S595 with *E.coli* spike, and Endostromatolite rock performed with universal bacterial 27F/907R primers. From left to right: DNA ladder (Fisher BioReagents, # BP2575-100), S595, S595 + *E.coli* spike, PCR positive control (bottom), PCR negative control (bottom).

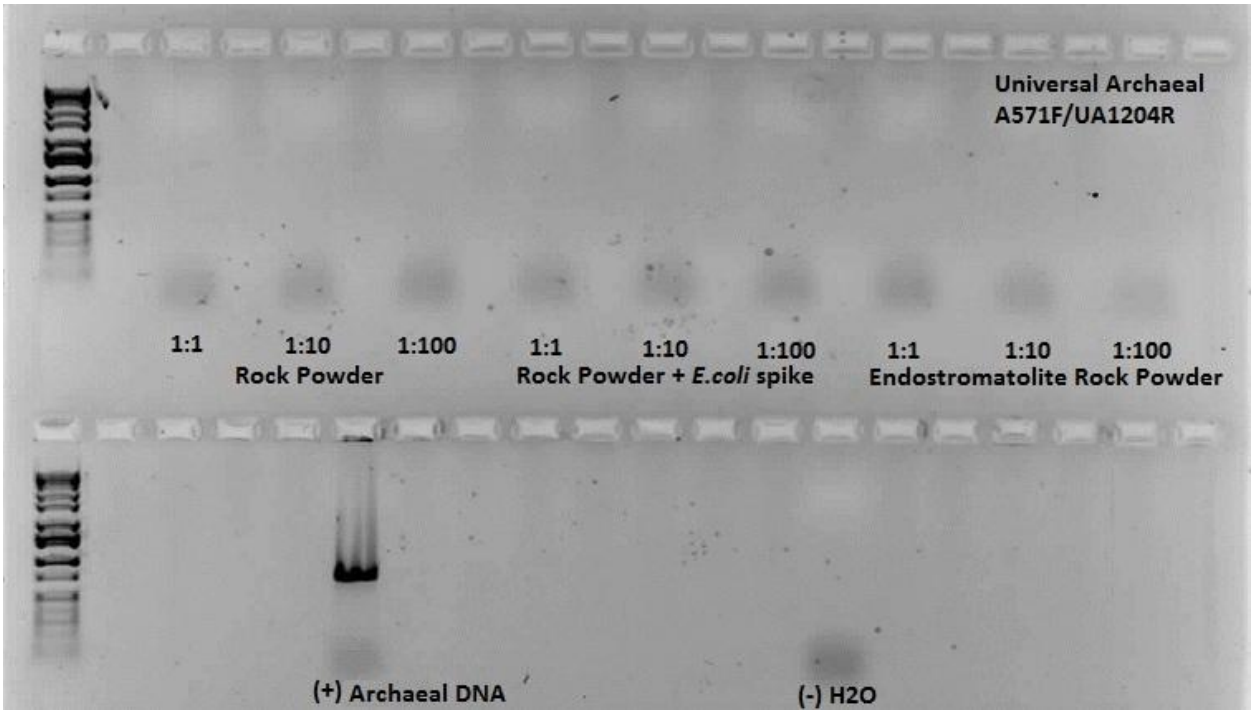

Figure 8.2 PCR results (1:1, 1:10 and 1:100 DNA template concentrations) on samples S595, S595 with *E.coli* spike, and Endostromatolite rock performed with universal archaeal A571F/UA1204R primers. From left to right: DNA ladder (Fisher BioReagents, # BP2575-100), S595, S595 + *E.coli* spike, PCR positive control (bottom), PCR negative control (bottom).

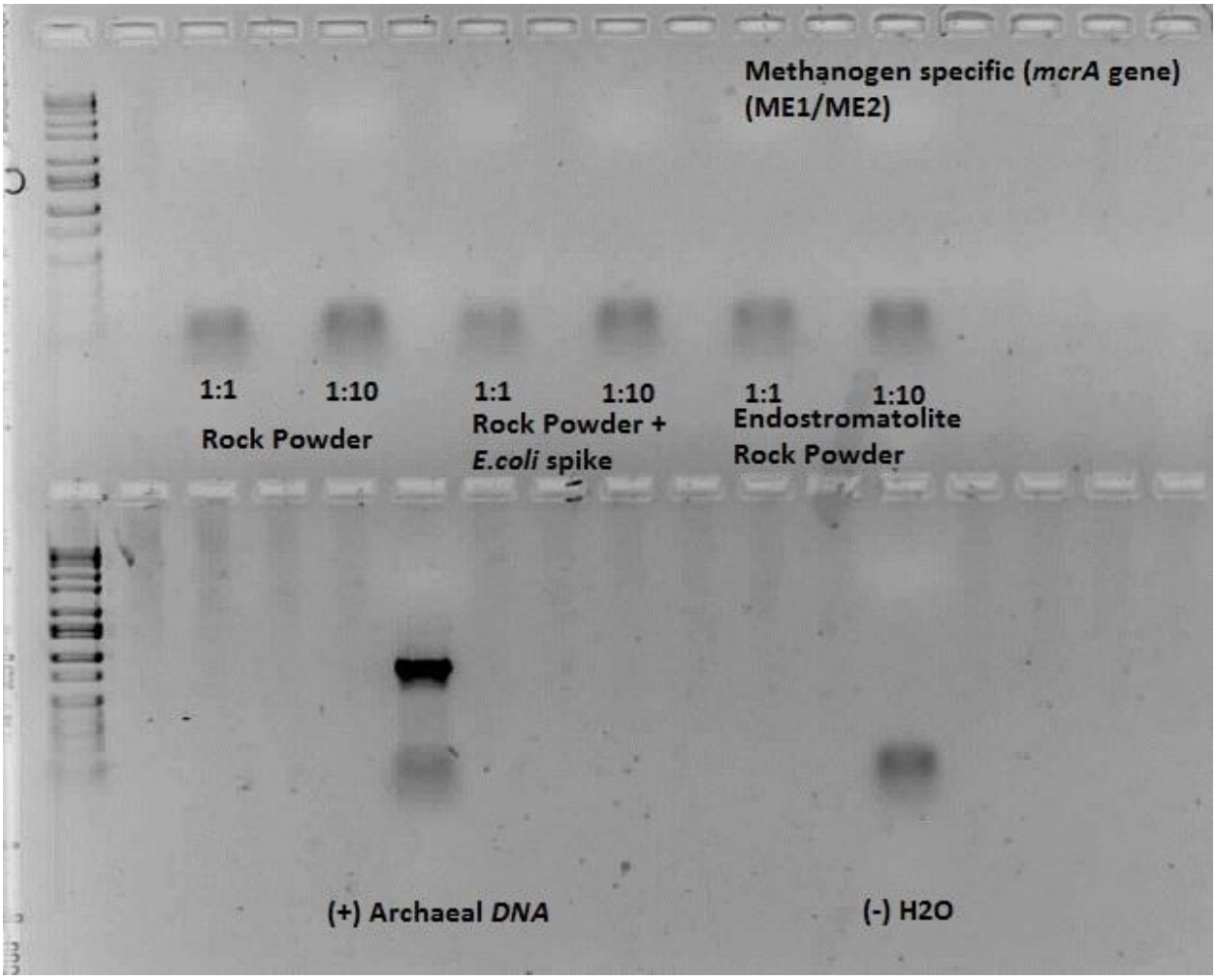

Figure 8.2 PCR results (1:1 and 1:10 DNA template concentrations) on samples S595, S595 with *E.coli* spike, and Endostromatolite rock performed with methanogen specific *mcrA* gene A571F/UA1204R primers. From left to right: DNA ladder (Fisher BioReagents, # BP2575-100), S595, S595 + *E.coli* spike, PCR positive control (bottom), PCR negative control (bottom).