Development of Microbial Characterization Techniques for Low-Permeability Sedimentary Rocks

NWMO TR-2013-17 December 2013

Greg F. Slater[1], Duane P. Moser[2], Barbara Sherwood Lollar[3]

[1]McMaster University
[2]Desert Research Institute
[3]University of Toronto

nwmo
NUCLEAR WASTE MANAGEMENT ORGANIZATION
SOCIÉTÉ DE GESTION DES DÉCHETS NUCLÉAIRES

Nuclear Waste Management Organization
22 St. Clair Avenue East, 6th Floor
Toronto, Ontario
M4T 2S3
Canada

Tel: 416-934-9814
Web: www.nwmo.ca

Development of microbial characterization techniques for low-permeability sedimentary rocks

NWMO TR-2013-17

December 2013

Greg F. Slater[1], Duane P. Moser[2], Barbara Sherwood Lollar[3]
[1]McMaster University, [2]Desert Research Institute, [3]University of Toronto

This report has been prepared under contract to NWMO. The report has been reviewed by NWMO, but the views and conclusions are those of the authors and do not necessarily represent those of the NWMO.

All copyright and intellectual property rights belong to NWMO.

Document History

Title:	Development of microbial characterization techniques for low-permeability sedimentary rocks		
Report Number:	NWMO TR-2013-17		
Revision:	R000	Date:	December 2013
Author Company(s)			
Authored by:	Greg F. Slater (McMaster University) Duane P. Moser (Desert Research Institute)		
Verified by:	Barbara Sherwood Lollar (University of Toronto)		
Approved by:	Barbara Sherwood Lollar (University of Toronto)		
Nuclear Waste Management Organization			
Reviewed by:	Jennifer McKelvie Laura Kennell Monique Hobbs		
Accepted by:	Mark Jensen		

ABSTRACT

Title: **Development of microbial characterization techniques for low-permeability sedimentary rocks**
Report No.: **NWMO TR-2013-17**
Author(s): Greg F. Slater[1], Duane P. Moser[2], Barbara Sherwood Lollar[3]
Company: [1]McMaster University, [2]Desert Research Institute, [3]University of Toronto
Date: December 2013

Abstract
The development of techniques for the characterization of potential indigenous microbial communities in low-permeability sedimentary rock was undertaken for six core samples of shale, interbedded shale and limestone, argillaceous dolostone and argillaceous limestone using phospholipid fatty acid (PLFA), molecular genetic and traditional microbiological approaches. To circumvent intrinsic properties of these cores (e.g., low biomass and high salinity/clay content), extensive methods development and control comparisons were undertaken. To characterise likely sources of contamination, samples of drilling water were collected and analyzed concurrently with each core segment. In addition, outer surfaces of collected cores were analyzed for PLFA and DNA and the results compared to those from drilling water and core interiors. Pristine subsamples from core interiors were collected by splitting off a minimum of 5 mm of outer core surfaces using a hydraulic splitter and pulverizing interior samples under sterile conditions. Flow cytometry, microscopy and cultivation, as well as PLFA and DNA analysis, resulted in negligible detection of microorganisms in the drilling water. It is hypothesised that the high salinity (1100 kg/m^3) of this water may have lysed most microbial cells originally present in the Lake Huron freshwater from which the drilling fluid was prepared, resulting in near-aseptic drilling.

PLFA were quantified in all core interior samples, corresponding to cell densities of ~1 – 3 x 10^5 cells/gram of rock, approximately an order of magnitude above the calculated detection limit for PLFA analysis of 3 x 10^4 cells/gram of rock. These cell abundances are consistent with the lowest abundances reported for oligotrophic systems, with the exception of some recent reports from oligotrophic marine sediments. PLFA concentrations for estimated rock volumes sampled during surface rinses were comparable to those observed for core interiors, with two exceptions that were one and two orders of magnitude higher. PLFA distributions for core interior and outer rinse samples were unique for all samples analyzed, indicating distinct sources for the depths sampled, as well as for outer rinses as compared to core interiors. Likewise, unique distributions of organic molecules present as alkanes and/or unresolved complex mixtures in each core indicate distinction between samples from different lithologies. Conversely, unlike the case of PLFA analysis, amplified DNA based surveys (16S rRNA gene) failed to detect evidence for indigenous microorganisms in any of the drilling water or core samples. These samples either contained microbial DNA at below detection limits (or were free of it entirely), or, alternatively, possess some feature that degrades or binds free DNA, thereby preventing its detection. After employing optimised DNA extraction and aggressive PCR amplification protocols (e.g. high numbers of PCR cycles, nested PCR in some cases, or use of highly degenerate primers in pyrotag analysis), traces of amplified bacterial DNA were detected from all samples, including the controls. Exhaustive comparisons of 454 pyrotag and traditional clone library results revealed that this DNA was likely derived from low-level reagent contamination. Such amplification of trace level reagent/process contaminants is expected to be an ongoing challenge that must be addressed during the analysis of low biomass rocks. Inhibition of DNA recovery was also observed in these samples as detection limit calibrations

utilizing added bacterial cells indicated that DNA could only be recovered from samples containing 4 x 10^5 cells/gram or higher. The cause of this inhibition was not discernible. It may be that strong sorption of DNA to clay particles or the presence of organic compounds inhibiting DNA amplification in our assays were the cause of the poor detection limits.

Overall, the results observed in this sample set indicate that these low-porosity rocks contain, at most, very low levels of biomass, comparable to cell densities observed in oligotrophic marine sediments. However, there are two caveats that must be recognized. First, the poorer detection limits observed for DNA-based methods as compared to the PLFA results means that the two methods cannot be directly compared in this current study. Further optimization of DNA extraction and purification protocols may overcome this limitation. Secondly, the PLFA-based estimates of cell numbers are based on turnover rates developed in experimental studies of surface and shallow subsurface systems. PLFA turnover rates specific to these low biomass, high salinity, low porosity deep subsurface environments have yet to be determined. If PLFA turnover rates are significantly lower in these systems, the estimated cell densities would be an overestimate of viable cells present. Beyond addressing this specific sample set, the experimental and analytical approach developed for this study represents an effective means to assess low biomass systems and should assist in the development of future site investigation plans to detect and characterize life in low-permeability, low biomass environments under consideration for the development of a deep geological repository.

TABLE OF CONTENTS

		Page
ABSTRACT		**iii**
1.	**INTRODUCTION**	**1**
2.	**METHODS AND RESULTS**	**3**
2.1	**SAMPLE COLLECTION AND DRILLING WATER CHARACTERIZATION**	**3**
2.1.1	**Overview**	**3**
2.1.2	**Materials and Methods (Sampling and Drilling Fluid)**	**3**
2.1.2.1	Core Collection	3
2.1.2.2	Drilling Water Collection, Field Measurements and Chemistry	6
2.1.2.3	Drilling Water Cell Enumerations: Flow Cytometry and Microscopy	6
2.1.2.4	Cultivation-based Assays and Controls (Drilling Water and Core)	7
2.1.2.5	Drilling Water PLFA Analysis	7
2.1.3	**Drilling Water Results**	**8**
2.1.3.1	Drilling Water Field Measurements and Chemistry	8
2.1.3.2	Drilling Water Cell Enumerations: Flow Cytometry and Microscopy	8
2.1.3.3	Cultivation-based Assays and Controls (Drilling Water and Core)	9
2.1.3.4	Drilling Water PLFA Analysis	9
2.2	**PLFA ANALYSES OF ROCK CORE MATERIAL**	**11**
2.2.1	**Materials and Methods (PLFA)**	**11**
2.2.1.1	Overview	11
2.2.1.2	PLFA Analysis	12
2.2.1.3	Rock Core Rinses	12
2.2.1.4	Rock Core Interiors	12
2.2.1.5	F1 Hydrocarbon Characterization	13
2.2.2	**PLFA Results**	**13**
2.2.2.1	PLFA Concentrations	13
2.2.2.2	Cell Density Estimates	14
2.2.2.3	PLFA Distributions	15
2.2.2.4	Rock Core Appearances	18
2.2.3	**Hydrocarbon Profiles**	**18**
2.3	**DNA ANALYSES OF ROCK CORE MATERIAL**	**19**
2.3.1	**Overview**	**19**
2.3.2	**Materials and Methods (DNA)**	**20**
2.3.2.1	DNA Extraction	20
2.3.2.2	Molecular Methods and Statistical Analysis	21
2.3.2.3	Sensitivity Calibration and Controls	23
2.3.3	**DNA Results**	**23**
2.3.3.1	PCR Assay Detection of bacteria and Archaea	23
2.3.3.2	454 Pyrotag Analysis of Archaea	25
2.3.3.3	454 Pyrotag Analysis of bacteria	25
2.3.3.4	Bacterial Diversity Confirmations using 16S rRNA Gene Libraries	34
2.3.3.5	Sensitivity Analysis: Microbial Detection Limits and Controls	34
3.	**DISCUSSION**	**37**

3.1 PLFA DISCUSSION 37
3.1.1 Viable Microbial Communities 37
3.1.1.1 PLFA Abundance 37
3.1.1.2 PLFA Distribution 37
3.1.1.3 Hydrocarbon Results 39
3.1.1.4 Further Questions Related to PLFA Results 39
3.2 DNA DISCUSSION 40
3.2.1 Context for the DNA Study 40
3.2.2 Direct Detection and DNA-Based Life Detection in Drilling Water 41
3.2.3 Targeted Microbial Investigations 42
3.2.4 Indigenous vs. Contaminating Microorganisms: 454 Pyrotag Analysis 43
3.2.4.1 454 Pyrotag: Abundant Bacterial Lineages 44
3.2.4.2 454 Pyrotag: Rare bacterial Lineages 46
3.2.5 16S rRNA bacterial Gene Libraries 47
3.2.6 Sensitivity Analysis: Bacterial Detection Limits and Controls 48
3.2.7 Further Questions Related to DNA Results 48

4. CONCLUSIONS 50

4.1 PLFA CONCLUSIONS 50
4.2 DNA CONCLUSIONS 50
4.3 SUMMARY 51

ACKNOWLEDGEMENTS 52

REFERENCES 53

APPENDIX A: SUPPLEMENTAL DISCUSSION 65

LIST OF TABLES

Page

Table 1: Summary of Project Samples and Associated Drilling Water Field Variables 5
Table 2: Major Chemistry Variables for Drilling Water 5
Table 3: Filtered Drilling Water Samples - Total PLFA Detected and Cell Density Estimates....10
Table 4: PLFA Results for Rock Core Interior Samples 15
Table 5: PLFA Results for Rock Core Rinses 15
Table 6: Bacterial Pyrotag Summary: Most Abundant Species Detected 32
Table 7: Least Abundant Bacterial Species Detected in Select Samples 33

LIST OF FIGURES

Page

Figure 1: Relative Abundances (mol%) of PLFA Detected in a 1 L Drilling Water Sample, DWR DGR8-014 10
Figure 2: Total concentrations of PLFA (µg of PLFA per gram of rock) detected from core rinses and core interiors; (*) Actual 529o_GeorgianBay concentration is off the scale of this graph at 76.2 ng/g 14
Figure 3: Relative abundances (mol%) of PLFA detected in rock core interiors and rock core rinses 17
Figure 4: Appearance of Rock Core Samples – (A) Individual slices of 334_SalinaA1. The two darker pieces in the bottom right corner resulted from a break along a natural fracture in the core. (B) Interior of 661_Cobourg, illustrating the dark material covering the surfaces along a natural fracture in the rock..18
Figure 5: Stained Agarose Gel Illustrating Presence or Absence of Putative Archaeal PCR Products Under Conditions Tested: (A) 32 Cycles; (B) 35 Cycles; and (C) Half-nested reaction. 24
Figure 6: Stained Agarose Gel Illustrating Presence or Absence of Putative Bacterial PCR Products Under Conditions Tested: (A) 32 Cycles; (B) 35 Cycles 25
Figure 7: Bacterial Phylum-Level Overview of the Entire 16S rRNA Gene Pyrotag Dataset Illustrating the Manner in Which Samples Cluster by Groupings Submitted to Sequencing Provider 26
Figure 8: Bacterial Phylum-Level Dendrogram Based on Figure 7 Data, Showing How Patterns of Microbial Diversity are Controlled by Batch Sent to Sequencing Provider. 27
Figure 9: Bacterial Species-Level Dendrogram Based on Data for all Cores and Controls Affiliated with This Project. 27
Figure 10: Summary Major Bacterial Species Detected in the Initial of Three Sample Batches Sent to the Sequencing Provider (“Batch I”). 28
Figure 11: Summary of the Most Abundant Bacterial Species Detected in DNA Extracts from the Largest Batch Sent to the Sequencing Provider (“Batch-II”). 29
Figure 12: Principle Coordinate Analysis (Species-level) Illustrating Patterns of Diversity Among Abundant Bacteria. Light Blue Circles Represent Drilling Water; Dark Blue Circles Represent Outer Core Subsamples; Black Circles Represent Inner Core Subsamples; Green Circles are Controls. Lowercase “S” Designates “Small Kit”.............29
Figure 13: Principle Coordinate Analysis (Species-level) Illustrating Rare Bacterial Diversity (e.g. Dominant Taxa Removed). Light Blue Circles Represent Drilling Water; Dark Blue

Circles Represent Outer Core; Black Circles Represent Inner Core; Green Circles are Controls. Lowercase “S” Designates “Small Kit” Extracts. 31
Figure 14: Neighbour-joining Phylogenetic Tree (1000 bootstraps). Scale bar = 2 Nucleotide Substitutions Per 100 Base Pairs 35
Figure 15: Sensitivity Calibration Experiment (Gel Electrophoresis) Showing Recovery of DNA From Known Numbers of Added Bacterial Cells From Baked Rock Controls). Limit of Detection ~ 4 x 10^5 Cells Per Gram. 36

1. INTRODUCTION

This report details the development and application of methods to characterise potential indigenous microbial communities present in low-permeability sedimentary rocks. It is recognised that microbial metabolisms and their by-products have the potential to affect the geochemical conditions of used nuclear fuel repositories (Stroes-Gascoyne et al., 2007). As such, assessment of the far-field microbial community and processes is an important component of site assessment. An example of such an assessment is the study of the Opalinus Clay, a candidate host rock for high-level nuclear waste repository in Switzerland (Mauclaire et al., 2007; Stroes-Gascoyne et al., 2007). The report presented here builds upon the state-of-science review presented in NWMO TR-2011-09 by B. Sherwood Lollar (Sherwood Lollar, 2011) by applying the knowledge and methods reviewed therein to a set of sedimentary core samples.

Specifically, this report describes the coordinated application of phospholipid fatty acid (PLFA) analysis and molecular genetic (DNA) analysis to assess the presence and identity of indigenous microbial communities in core samples obtained during drilling of a borehole in the Michigan Basin. This study had three primary objectives:

1) Constraining potential sources of microbial contamination during drilling, sampling and downstream analysis, such that microorganisms detected can be attributed to indigenous microbial communities with confidence;
2) Assessment and characterization of potential indigenous microbial communities by phospholipid fatty acid (PLFA) analysis; and
3) Assessment and characterization of potential indigenous microbial communities by molecular genetic (DNA) analysis.

As noted in Sherwood Lollar (2011 and references therein), it is becoming increasingly apparent that microbial life may permeate deep within the subsurface of the Earth. Studies of subsurface microbiological systems have ranged from wells drilled to hundreds of metres (Chapelle et al., 2002), sometimes specifically for the study of subsurface microbiology (Pedersen, 2001), to studies of petroleum reservoirs and deep terrestrial subsurface systems up to 4 – 5 km depth over a range of rock types (Moser et al., 2005; Hallmann et al., 2008; Sherwood Lollar, 2011). Such microbial life can utilise a variety of chemolithotrophic processes, taking advantage of the stored energy present in chemical disequilibria that may occur in subsurface systems (Pedersen, 2000). Microorganisms in these subsurface systems are generally found to be tolerant of both elevated salinity and temperature (Krumholz et al., 1997; Liu, 1997; Moser et al., 2003) and to utilise a range of metabolisms including reduction of sulphate (Krumholz et al., 1997; Moser et al., 2005; Chivian et al., 2008) or metals, as well as fermentation and production and/or consumption of methane and acetate (Kotelnikova, 1998; Kieft et al., 1999),

Although there have been some assessments of the presence of indigenous subsurface microbial communities in solid core samples (Sherwood Lollar, 2011), most work to-date has focused on groundwater produced from boreholes (Kotelnikova, 1998; Lin et al., 2006; Sherwood Lollar et al., 2006). However, the low-permeability rocks considered as candidates for geologic nuclear fuel repositories do not allow for sample collection via groundwater, and thus, direct sampling of core materials must be undertaken in order to assess the potential presence of indigenous microbial communities.

Microbial cell densities are expected to be very low within low-permeability rocks. Studies of planktonic cells in groundwater samples have yielded cell density estimates of 10^2 to 10^6 cells per mL in most deep subsurface systems (Pedersen, 1993; Whitman, 1998). Though cell abundance has not been widely assessed for solid rock samples, assessment of microbial biomass in marine seafloor sediments has yielded estimates of ~10^5 (and recently as few as 10^3) cells/gram sediment (Parkes et al., 1994; Parkes et al., 2000; Kallmeyer et al., 2012). It is reasonably expected that low-permeability rocks will have cell abundances comparable, or lower, than those obtained by these studies. Thus, a key requirement for low-permeability rock studies is sensitivity of the analytical approaches used. The studies of marine seafloor sediments discussed above relied upon direct cell counting techniques (Parkes et al., 1994; Parkes et al., 2000; Kallmeyer et al., 2012). However, this approach cannot be applied to consolidated material, because it would rely on physical separation of the cells from the rock surfaces. The preparation of consolidated materials, which requires pulverizing the rock, would be expected to severely compromise microbial cells, ultimately resulting in an underrepresentation of the total cell abundance.

Key approaches for detecting subsurface organisms in hard rock habitats fall into three classes: molecular genetic characterization, characterization of biomarkers such as PLFAs, and the culturing of organisms from the substrates. In the case of the latter, if an organism consistent with site conditions and absent from known sources of contamination can be cultured, this is the strongest possible evidence that viable organisms are indigenous to a sample. However, this approach is of limited utility because it is generally accepted that less than 1% of microorganisms in environmental samples are culturable with known techniques (Staley and Konopka, 1985; Amann et al., 1995). Thus, as with any environmental sample, while culturing should be a component of analysis, molecular genetics and biomarker approaches are probably the most appropriate methods with which to assess microbial populations in low-permeability rocks.

Here, we report the development and application of DNA- and PLFA-based approaches to analyze microbial communities in low-permeability sedimentary rock core samples. A key component of this project was to constrain potential sampling- and laboratory-derived microbial contamination. It has long been recognised that, although intact core is one of the few means available to survey for life in solid rock (Haldeman and Amy, 1993; Chapelle and Bradley, 1996; Colwell et al., 1997; Fredrickson et al., 1997), the potential for contamination remains a constant challenge to be overcome (Haldeman and Amy, 1993; Onstott, 1998). Thus, to constrain potential contamination, we utilised a combination of direct counting, cultivation, DNA and PLFA approaches to assess potential contamination from drilling water during sample collection. We further analysed outer core surfaces and compared the results obtained to samples from core interiors to assess whether distinct microbial communities were being added to samples during collection. This assessment focused on the outer 1 – 5 mm of core material, based on previous results from this group and others, whereby dye or fluorescent microsphere tracers, or assessments of extractable hydrocarbons, have shown that penetration into low-permeability rocks is on the order of 1 – 5 mm (Onstott et al., 2003; Kallmeyer et al., 2006; Sherman et al., 2007). Comparison between outer surface samples and internal samples can be used to confirm that removal of the outer >5 mm of surface has, in fact, removed any sampling influence.

A second fundamental challenge, due to the low biomass expected for these systems (Stroes-Gascoyne et al., 2007; Colwell and D'Hondt, 2013), is the sensitivity of analysis. Methods that

rely on direct detection of microbial cells or biomarker compounds can be limited in their ability to detect very low concentrations of cells or cellular components obscured by inhibitors. To overcome this limitation, large sample sizes must be extracted and concentrated to achieve maximum sensitivity. While this can be a logistical challenge, the advantage of such approaches is that they assess the presence of indicators directly, without amplification. This allows techniques, such as phospholipid fatty acid (PLFA) analysis, to not only detect the presence of microbial indicators, but also to estimate the cell abundances present. In contrast, molecular genetic approaches generally rely on amplification of genetic material to assess the presence of organisms. These approaches are extremely powerful and, in principle, can detect very small numbers of target cells. However, the amplification steps must be undertaken with care in order to avoid contamination. Further, matrix effects, from mineral (e.g. clay (Stroes-Gascoyne et al., 2007) and/or organic components (Alaeddini, 2012) or due to high salinity (Bessetti, 2007) of samples, can cause interferences that drastically reduce the effectiveness of amplification.

Thus, this work was vulnerable to two fundamental errors from its inception: 1) that life is present in the inner core samples, but due to challenges inherent to clay-rich rock samples, cannot be detected; and 2) that the core interiors are, in fact, sterile, but due to the exquisite sensitivity of amplified DNA procedures, life signatures resulting from trace contaminants in drilling water or laboratory reagents would be detected and reported as indigenous. This latter case is especially difficult to avoid with confidence and motivated the methods development efforts detailed here.

2. METHODS AND RESULTS

2.1 SAMPLE COLLECTION AND DRILLING WATER CHARACTERIZATION

2.1.1 Overview

The collection of cores and the associated potential for contamination applies to both PLFA- and DNA-based characterizations that appear later in this report. Thus, a description of the sampling protocol and characteriztions of the major potential source of microbial contamination – drilling water – are detailed in this section to set the stage for the detailed analyses of the cores that follow. These characterizations detail the major aquatic chemistry features of the drilling water and the results of several complementary approaches that were applied in an effort to constrain the bioburden of drilling fluid.

2.1.2 Materials and Methods (Sampling and Drilling Fluid)

2.1.2.1 Core Collection

Core and comparative drilling fluids samples were collected by Geofirma Engineering Ltd. as part of Ontario Power Generation's (OPG) Deep Geological Repository Project for Low and Intermediate Level Waste at the Bruce Nuclear Site near Tiverton, Ontario (Borehole DGR-8). Water used as drilling fluid was prepared from Lake Huron water that was adjusted to 1100 kg/m^3 salinity via the addition of 90 kg/m^3 of NaCl and 45 kg/m^3 $CaCl_2$. These salts were added to prevent dissolution of anhydrite and halite zones and to prevent weathering of shales rather than for microbiological purposes, but as discussed below, the increased salinity appears to have been very significant in reducing microbial presence in the drilling water. Subcores from

prescribed depths were logged, bagged, and frozen on-site prior to shipment for analysis. Subcores were handled with sterile nitrile gloves, and chosen to be visually representative of their respective stratigraphic units. Cores were broken using a flamed hammer and chisel to achieve required lengths (20 to 30 cm), and stored in Ziploc bags to minimise contamination. Sample analysis described herein focused on six sample sets from the Salina A1 Unit, and the Queenston, Georgian Bay, Cobourg (2) and Sherman Fall formations, as indicated in Table 1.

Table 1: Summary of Project Samples and Associated Drilling Water Field Variables

NWMO Sample ID	Method	Fate/ Project Sample ID	Geologic Formation	Lithology	Depth (m)	Drilling Water Temp (^{o}C)	Drilling Water Conductivity (mS/cm)	Drilling Water pH	Drilling Water [a]DO (ppm)
DGR8-334.42	PLFA	334"i/o"_Salina A1	Salina A1	Argillaceous dolostone	334.4	19.8	162.2	7.34	3
DGR8-334.72	DNA	334"i/o"_Salina A1	Salina A1		334.7	19.8	162.2	7.34	3
DGR8-464.94	PLFA	464"i/o"_Queenston	Queenston	Shale	464.9	20.6	164.6	6.96	<1
DGR8-464.70	DNA	464"i/o"_Queenston	Queenston		464.7	20.6	164.6	6.96	<1
DGR8-529.94	DNA	529"i/o"_Georgian Bay	Georgian Bay	Inter-bedded shale and limestone	529.9	15	166.9	6.97	--
DGR8-531.57	PLFA	529"i/o"_Georgian Bay	Georgian Bay		531.6	15	166.9	6.97	--
DGR8-661.28	DNA	661"i/o"_Cobourg	Cobourg	Argillaceous limestone	661.3	16.1	143.3	6.99	1
DGR8-661.63	PLFA	661"i/o"_Cobourg	Cobourg		661.6	16.1	143.3	6.99	1
DGR8-682.66	DNA	682"i/o"_Cobourg	Cobourg	Argillaceous limestone	682.7	14.1	146	7.19	1.5
DGR8-682.86	PLFA	682"i/o"_Cobourg	Cobourg		682.9	14.1	146	7.19	1.5
DGR8-694.93	DNA	694"i/o"_Sherman Fall	Sherman Fall	Argillaceous limestone	694.9	17.5	143.2	6.89	1
DGR8-694.57	PLFA	694"i/o"_Sherman Fall	Sherman Fall		694.6	17.5	143.2	6.89	1

[a]Dissolved oxygen

Table 2: Major Chemistry Variables for Drilling Water

Sample Depth (m) (NWMO ID)	Geologic Unit	[b]Cl-	SO_4^{2-}	Na^+	K^+	Ca^{2+}	Mg^{2+}	[c]NO_3^-	$NH3^+$	DOC	Fe
334.72 (DGR8-003)	Salina A1	79600	666	36500	543	12700	53.1	0.24	1.28	3.8	1.1
464.7 (DGR8-004)	Queenston	76300	580	34600	554	12300	56.5	0.23	2.51	3.4	4.3
495.68 (DGR-005)	Queenston	78300	559	34500	618	12700	77.6	0.14	4.26	4.4	2.2
531.57 (DGR-007)	Georgian Bay	94000	417	42900	704	14400	72.9	0.22	4.33	6	5
661.63 (DGR-012)	Cobourg	70900	882	32200	559	12100	48.9	0.44	5.61	4.1	1.8
670.1 (DGR-013)	Cobourg	71900	888	30800	560	11800	52.5	0.26	5.8	4	2.5
682.86 (DGR-014)	Cobourg	69500	926	31600	514	11200	48.8	0.25	5.22	3.6	6.3
694.57 (DGR-015)	Sherman Fall	74500	961	33200	570	12100	57.5	0.23	5.89	4.3	6.2
[a]DGR-00Xa	Blank Test	12.3	87.5	13.4	6.94	103	8.58	0.04	0.06	NA	<0.1

[a]Tap water control, [b]Units = mg/L, [c]F-, NO_2^-, ~ detection limit 0.01 mg/L; F^- ≤ detection limit of 0.5 mg/L

2.1.2.2 Drilling Water Collection, Field Measurements and Chemistry

As noted above, drilling fluids were prepared using Lake Huron freshwater adjusted to 1100 kg/m^3 salinity via the addition of 90 kg/m^3 of NaCl and 45 kg/m^3 $CaCl_2$. This drilling fluid was recycled in a closed loop system where drilling fluid losses were tracked. Only minor losses of drilling fluid were observed during the depth intervals relevant to this study. Drilling fluid (~50 mL) was sampled for aqueous chemistry at depth intervals corresponding to the sampling of cores in 60 mL Nalgene polypropylene bottles (pre-soaked for 24 hours in nanopure water) and frozenon site. Drilling water samples were numbered sequentially on site and collected from depths most closely corresponding to the geologic units of the core samples analyzed. The drilling water analytical results for those closest to the cores analyzed in this study are presented in Table 2. Physical measurements, including temperature, conductivity and pH, were performed on-site by the sponsor. Dissolved oxygen concentrations were measured from flowing sample lines using Chemets visual colourimetric ampules (Chemetrics, K-7501 and K-7512). Unfiltered samples for cellular enumeration were collected in 50 mL polypropylene centrifuge tubes (2% glutaraldehyde preservative) and stored on ice or at 4°C. Drilling water was sampled for microbial DNA as well, at each of the core sampling points, by passing 60 – 180 mL through Sterivex (Millipore) 0.2. micron filter units with a sterile 60 cc syringe until plugged. Filters were purged of excess water, frozen immediately on dry ice, and subsequently stored at -80°C. In four cases, 1 litre of drilling water was collected and frozen on-site and then transported frozen to McMaster University for large-volume filtration and analysis using the same filter systems.

For major ion analysis, 30 mL aliquots of drilling water were removed from 50 mL tubes used for storage at -20°C, filtered using a Millex GP 0.22 µm polyethersulfone membrane syringe filter (Millipore) and shipped to the Desert Research Institute's (DRI's) US Environmental Protection Agency- certified (EPA) Water Analysis Laboratory in Reno, Nevada. The remaining 20 mL was held back for archival purposes and cultivation assays (see 2.1.2.4). Despite "frozen" storage conditions, these samples were liquid due to their high salinity. Specific analytes were obtained using US Environmental Protection Agency (USEPA, 1979, 1993) and Standard Methods protocols (Rice, 2012). EPA protocols used included: sulphate and chloride by ion chromatography (EPA 351.2 and EPA 300). Standard Methods protocols used included: nitrate and nitrite by automated colourimetric cadmium-reduction methodologies (SM4500-NO3F), ammonium by the phenate methods (SM-4500-NH4F), fluoride by ion-selective electrode (SM-4500-F-C), major cations and iron by atomic absorption spectroscopy (SM3111B), and dissolved organic carbon (DOC) by persulfate digestion/IR spectroscopy (DOC, SM5310C).

2.1.2.3 Drilling Water Cell Enumerations: Flow Cytometry and Microscopy

Direct microbial cell enumeration was undertaken by several complementary approaches. Flow cytometry was performed using a MicroPro flow cytometer (Benton Dickenson), according to manufacturer's protocols. Glutaraldehyde-preserved (2.5%) drilling water was pre-filtered using a 40 µm cell strainer (BD Falcon, USA) to remove large particles prior to analysis. Two-millilitre samples, undiluted and diluted (1:10 and 1:100 with 3X marine salts solution to approximately match the original salinity), were stained with Syto62 Reagent (product # AA-3000-03, BD) and analysed using standard instrument protocols for total and viable cells, and with a custom assay for autofluorescent cells (e.g., small algae).

Direct microscopic counts (Hobbie, 1977) were performed on 1 mL of the preserved drilling water samples using the nucleic acid binding stain 4', 6-diamidino-2-phenylindole (DAPI). Duplicate preparations of stained cells (3 µg/mL DAPI for ten minutes) for all samples were filtered onto 0.22 µm black polycarbonate membranes, rinsed by vacuum filtration to remove unbound DAPI, and dried with a light vacuum. Rinses included filtered (0.2 micron) MilliQ, tap water, and 3X artificial sea salt H_2O to verify that cells were not lost due to plasmolysis. Filters were examined by epifluorescence microscopy under UV illumination at 1000X (Zeiss Axioskop 2 Plus), from which cell densities per mL of drilling water can be back-calculated from average counts per field of at least 30 counted (note, no cells were actually visualised). To confirm the unusual result of no cells being visible in these preparations, all of the slides were independently assessed by three different staff members including, Dr. Duane Moser (an experienced user of epifluorescence microscopy). To verify the lack of visible cells noted for epifluorescence microscopy, a variation of direct counting was also performed with a Petroff-Hausser Counting Chamber (Hausser Scientific, 5 µL aliquots) under both bright field and phase contrast illumination at 400X. Although no cells were detected with this method, counts would be expressed as averages of cells/mL of drilling water calculated over twenty fields.

2.1.2.4 Cultivation-based Assays and Controls (Drilling Water and Core)

Several focused efforts were made to cultivate informative classes of microorganisms from certain samples as independent verification of the DNA and PLFA results. The most likely source of microbial contamination to the cores would have been non-halophilic aerobic heterotrophs from the surface water source for the drilling fluid. To survey for these microorganisms, all of the drilling water samples were tested by traditional plate count methodologies on agar plates using the US EPA's standard plate count medium, R2A (Reasoner, 1985). To perform this assay, 50-µL aliquots of drilling fluid from each sample were streaked for isolation on the surface of R2A plates and incubated at room temperature in the dark for three weeks, checking for growth every several days. An attempt was made to cultivate sulphate-reducing bacteria (SRBs) from cores 334i_Salina A1 and 334o_Salina A1 also. Crushed inner core material (0.1 g) was added to 60 mL stoppered glass serum vials containing 30 mL of liquid media (pH 7.0) and containing per litre: 25 g NaCl, 3.5 g $MgSo_4$•$7H_2O$, 2.7 g $MgCl_2$•$6H_2O$, 0.33 g KCl, 0.25 g NH_4Cl, 0.14 g $CaCl_2$•H_2O, 1 mL ATCC Trace Vitamins and Minerals (ATCC, Manassas,VA), 100 µL resazurin (5 mg/mL), 2 mL 1 M phosphate buffer, 0.5 g cysteine-HCl, 0.5 g Na_2S, and 1 g $NaHCO_3$, all filter sterilised. Cultures were incubated under H_2/CO_2 2:20 (the balance in N_2) at room temperature or 45°C for several months. Variations of this medium were prepared using seawater salts at 1X, 3X and 5X concentrations as well. Culture bottles were tested for microbial growth weekly by visual examination for turbidity, and microscopically in the same manner as with the drilling fluid samples (e.g., using a Petroff-Hausser Counting Chamber under phase contrast at 400X). These cultures were assessed for the production of sulfide as well, using lead acetate strips.

2.1.2.5 Drilling Water PLFA Analysis

The PLFA content of the drilling water was analysed to assess a potential external source of contamination during sampling. A set of syringe filters collected during drilling from 60 mL of drilling water were extracted using the PLFA method (see Section 2.2.1.2 for detailed description of methods). No PLFA were detected from these extracts; however, the small volume of water sampled resulted in relatively poor sensitivity. Detection limits were calculated based on the lowest mass of PLFA that could be quantified by the gas chromatograph mass

spectrometer to be in the 4 x 10^7 cells/ml. Once this issue was identified, four one litre samples of drilling water, from depths of 662, 670, 682, and 695 metres, were collected and frozen on site. These samples were returned to McMaster University where they were filtered in the lab using the same syringe filter system. The filters were subsequently extracted and analysed for PLFA. The larger water volumes for these samples decreased the calculated detection limit of this analysis to 3 x 10^4 cells/ml.

2.1.3 Drilling Water Results

2.1.3.1 Drilling Water Field Measurements and Chemistry

Field parameters were measured by Geofirma Ltd. for each of the twenty-two depth intervals for which a core subsample was obtained. The parameters corresponding to the six samples analyzed in this study are reported in Table 1. Temperature measurements ranged from a low of 6.8°C (e.g., DGR8-574.53) to as high as 20.6°C in sample DGR8-464.70. No particular pattern of temperature vs. depth was apparent in this dataset. As might be expected, considering the salinity modification that was performed on the drilling water, conductivity was very high (~ 3X that of seawater) for all of the samples and constrained to a relatively limited range, from a low of 125.5 to a maximum of 172.5 mS/cm. Values for pH ranged from a low of 6.28 in the deepest sample (DGR8-723.75) to 7.75 in the shallowest (DGR8-287) and generally clustered around neutrality. Dissolved oxygen was detected in all samples, but was relatively low (e.g., < 1 mg/L to a maximum of 3 mg/L in DGR8-334.72). Values were not obtained for six samples from 519 – 574 metres depth and, as with the temperature and conductivity datasets, no obvious trends with depth were apparent.

Select aquatic and nutrient chemistry measurements were obtained for the subset of samples for which DNA analysis was performed (Table 2). One additional sample of DRI tap water from the Las Vegas Valley Water District was submitted for analysis as a control or benchmark for a typical surface water source in recognition of the unusual nature (e.g., high salinity) of the drilling water sample set. The major contributors of salinity were Cl^- (69,500 – 94,000 mg/L), Na^+ (30,800 – 42,900 mg/L) and Ca^{2+} (11,200 – 14,400 mg/L). As with the physical variables, no obvious trends were apparent in relation to depth between chemical species. The potentially bioactive anion, sulphate (417 – 961 mg/L), was generally high and displayed a pattern of increasing concentration with depth. Whereas K^+ was high (514 – 704 mg/L), F^- was below the detection limit of 0.5 mg/L in all samples. NO_3^-, a potential dissimilatory electron acceptor for microorganisms, was detectable but relatively low in all samples, ranging from 0.04 and 0.14 mg/L; NO_2^- was low (approximately at the 0.01 mg/L detection limit). Conversely, ammonium was present at high concentrations (1.28 – 5.89 mg/L) and, unlike the other analytes, displayed a clear increase in concentration with depth. Dissolved organic carbon (DOC, 3.4 – 6.0 mg/L) was generally high across the dataset and showed no obvious pattern with depth. Soluble iron concentrations, ranging from 1.1 – 6.3 mg/L, were high and loosely tracked the ammonium depth profile.

2.1.3.2 Drilling Water Cell Enumerations: Flow Cytometry and Microscopy

Direct cell counting was only attempted for the drilling fluid. Three methods were used: flow cytometry, epifluorescence microscopy, and phase contrast microscopy. For the flow cytometry dataset, despite the use of various dilutions and salinities of diluent, all samples proved to be uncountable. Rather than producing typical fluorescence vs. scatter plots with populations of cells clearly evident as clusters of datapoints, these plots produced very high counts of

randomly positioned particles, possibly consistent with cell debris (data not shown). Similarly, with one exception, epifluorescence and visible light microscopy failed to reveal intact microorganisms when evaluated by multiple experienced users. This single exception, one of the duplicate preparations for DGR8-382.61 collected from 382.61 m depth, contained a very large number of brilliantly-stained intact cells when prepared for microscopy on two different occasions separated by many months. As this result only appeared in one of the replicate preparations from this sample, the cells detected are likely not likely inherent to the sample, but rather were probably added to the sample after collection. However, the appearance of these cells also can be regarded as a positive control for the method (e.g., cells are easily detected when present in significant numbers). Very large amounts of ground mineral particles were visible in both forms of microscopy and could have obscured microbial cells. Collectively, the results of the three complementary direct counting methods indicate an absence of cells or were inconclusive.

2.1.3.3 Cultivation-based Assays and Controls (Drilling Water and Core)

An attempt was made to enrich several classes of microorganisms of particular relevance to nuclear waste management or which would be predicted to exist as contaminants in the drilling water. Serum bottle enrichments targeting halophilic sulphate-reducing microorganisms were established from core 334i_Salina A1 and 334o_Salina A1 in high-salt media under a range of relevant conditions of temperature and headspace gas compositions. These bottle enrichments were monitored for several months and no indications of growth (e.g., turbidity, visible cells via microscopy) or sulfide production were ever noted (data not shown). Likewise, streaking of undiluted drilling fluid for isolation on R2A agar plates failed to provide evidence for the survival of freshwater microorganisms that would have been present in the water used for drilling fluid. Thus, no evidence for the presence of cultivable microorganisms was obtained under the conditions used for this study.

2.1.3.4 Drilling Water PLFA Analysis

The results for the PLFA analyses of filtered drilling water are shown in Table 3. Three of the 1 L filtered drilling water samples did not contain any quantifiable PLFA, though the 1 L samples from 661.63 m and 694.57 m had PLFA present that were below the level of quantification (denoted as n.q. in Table 3). One of the samples, from 682.86 m depth, contained a total of 0.4 µg of PLFA. Based on a conversion factor of $6x10^4$ cells per picomole of PLFA, this PLFA concentration represents a microbial cell density of around $8x10^4$ cells per mL of drilling water (Green and Scow 2000). The specific PLFA present in this sample are shown in Figure 1.

Table 3: Filtered Drilling Water Samples - Total PLFA Detected and Cell Density Estimates

Depth (m) (Sample ID)	Geologic Unit	Volume of filtered water (mL)	Total mass of PLFA detected (µg)	Cell density estimate (cells/mL water)
n/a (DWR DGR8 001–Blank)	System blank	60	n.d.	n.d.
n/a (DWR DGR8 – 001)	System blank	60	n.d.	n.d.
334.42 (DWR DGR8 – 003)	Salina A1	60	n.d.	n.d.
531.57 (DWR DGR8 – 007)	Georgian Bay	60	n.d.	n.d.
661.63 (DWR DGR8 – 012)	Cobourg	1000	n.q.	n.q.
670.10 (DWR DGR8 – 013)	Cobourg	1000	n.d.	n.d.
682.86 (DWR DGR8 – 014)	Cobourg	1000	0.4	$8x10^4$
694.57 (DWR DGR8 – 015)	Sherman Fall	1000	n.q.	n.q.

n/a – depth not available, blank sample from drilling system; n.d. - No PLFA Detected; n.q. - Non-Quantifiable

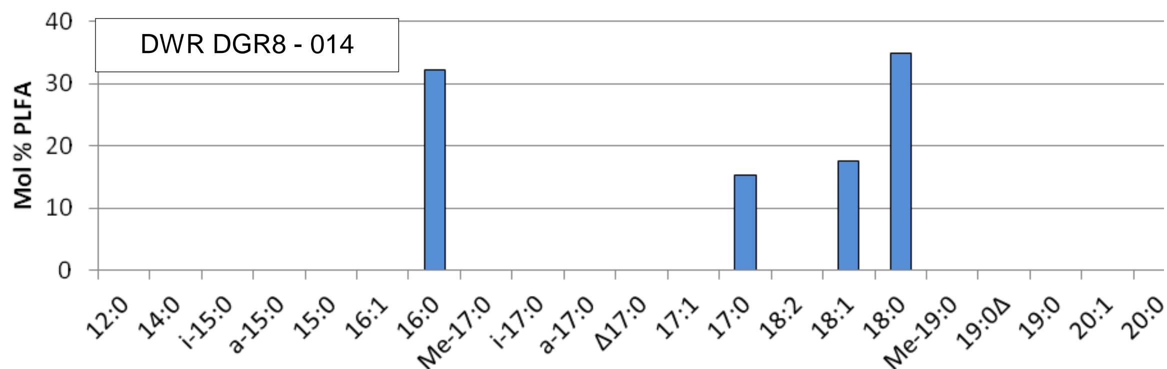

Figure 1: Relative Abundances (mol%) of PLFA Detected in a 1 L Drilling Water Sample, DWR DGR8-014

The PLFAs present in the DWR DGR8-014 sample were 16:0, 17:0, 18:1 and 18:0 which are not diagnostic. The PLFAs C16:0, C18:0 and C18:1 are consistent with observations in the drill core samples (see Section 3.2); however, C17:0 was not observed in the corresponding rock core (682i_Cobourg).

2.2 PLFA ANALYSES OF ROCK CORE MATERIAL

Phospholipids are a major component of bacterial and eukaryotic microbial cell membranes and have been shown to hydrolyze within days-to-weeks upon cell death under ambient surface conditions (White et al., 1979). Therefore, phospholipid fatty acids (PLFA) collected from environmental samples are considered to represent a snapshot of the viable bacterial and eukaryotic members of the microbial community within a particular system (White et al., 1979). PLFA analysis of environmental samples can provide two primary types of information about the microbial community at a site. Firstly, total PLFA concentrations provide estimates for the number of viable microbial biomass present. Cell density estimates are based on generic conversion factors relating PLFA concentrations to microbial cell densities (Balkwill et al., 1988; Green and Scow, 2000). The total amount of PLFA composing a microbial cell membrane is determined by the size of the microbial cell and, therefore, the size of the cell membrane. Cells present in low permeability environments may be expected to be small due to the limiting nature of the environment and thus conversion factors based on larger cell sizes may result in underestimates of cell density, nonetheless, , this approach remains valuable as a means to estimate microbial cell mass in environments where cells cannot be quantified by direct counting techniques (White and Ringelberg, 1997a).

Secondly, PLFA can serve as biomarkers for particular microbial groups, or for responses of microorganisms to environmental conditions or stresses. Using literature records of PLFA produced by specific bacteria and eukarya, PLFA biomarkers can be used to identify particular microbial groups in an environment (Ringleberg et al., 1997; Green and Scow, 2000; Fang et al., 2007; Weiss and Cozzarelli, 2008). Alternatively, the presence or ratio of certain PLFA can indicate physiological conditions of the cells (e.g., stress responses). For instance, in response to stress or resource depletion, cells can convert between unsaturated and cyclopropyl PLFA, or can alter the ratio of cis to trans unsaturated (Guckert et al., 1986; Kieft et al., 1994). This intent of this study was to identify and quantify PLFA, if possible, from the six low-permeability sedimentary rock core samples, and to evaluate if the presence of PLFA could or could not be due to sampling artefacts, with the goal of determining whether or not there was evidence for indigenous microbial communities.

2.2.1 Materials and Methods (PLFA)

2.2.1.1 Overview

Six rock cores collected from the priority geologic units (see Section 2.1.2.1) and covering a range of depths and lithologies (Table 1) from DGR8, were analysed for PLFA to assess the presence of viable microbial communities indigenous to the subsurface rocks. As described in Section 2.1.2.4, drilling water blanks were analysed to determine whether the water used for drilling contained detectable PLFA. In order to assess whether or not PLFA were introduced into the rock cores during the drilling process or during sample handling, the outer surfaces of each rock core sample were rinsed with extraction solvents prior to cutting or crushing, and these rinses were analysed for PLFA. Based on the work of Sherman et al. (2007), it was assumed that contamination penetration would be limited to the outer 5 mm of the rock. Analysis of rock core interiors to assess the presence of indigenous microbial populations involved removal of the outer surfaces of the core (> 5 mm) to avoid any surface penetration effects. In order to obtain sufficient interior sample mass, only intact core segments were analysed, as broken segments would have lost too much mass during sub-sampling to remove the potentially impacted outer surfaces.

2.2.1.2 PLFA Analysis

PLFA were extracted from all samples using a modified Bligh and Dyer PLFA extraction method (Bligh, 1959) The resulting total lipid extracts (TLE) were separated into (F1) non-polar, (F2) neutral and (F3) polar fractions using silica gel chromatography (Guckert et al., 1985). Phospholipids recovered from the polar fractions were converted to fatty acid methyl esters (FAMEs) via mild-alkaline methanolysis and subsequently purified by a secondary silica gel chromatography (Guckert et al., 1985). Microbial FAMEs were identified and quantified using gas chromatography-mass spectrometry (GC-MS) on an Agilent GC-MS (Agilent Technologies Inc., Santa Clara, CA, USA) with a DB-5MS capillary column (30m x 0.25µm film thickness). The temperature programme for the GC was maintained at 50°C for one minute, then increased the temperature at 20 °C/min to a temperature of 130°C, then increased at 4°C/min to 160°C, and finally increased at a rate of 8°C/min to a temperature of 300°C which was held for 5 minutes. In order to optimise sensitivity and allow the maximum sample mass to be injected via the auto-sampler system on the GC-MS, samples were run in 50 µL or 100 µL volumes with 2 µL injections. The limit of quantification on this system was determined to be the lowest concentration standard that can be reliably analysed, which is 2 ng/µL for a 1 µL injection on the GC-MS. Sample peaks with area responses lower than this standard could be detected, but not quantified. All reported concentrations consider only quantifiable peaks.

Identities of the fatty acids are listed with the following nomenclature: total number of carbon atoms followed by the total number of double bonds (e.g., 16:1 represents a 16 carbon monounsaturated fatty acid). Terminal-branching fatty acids are indicated by the prefixes *i (iso)* and *a (anteiso)*. Mid-branching positions are represented by the number of carbon atoms from the carboxyl group to the methyl group (e.g., 10Me16:0). Cyclopropyl fatty acids are represented by the prefix *cy*.

2.2.1.3 Rock Core Rinses

Rinses of the outer core material were analysed for PLFA to assess the presence and abundance of potential contaminants on the outer surfaces of the rock cores. The cores were individually rinsed with Bligh and Dyer extraction solvents (45 mL DCM, 90 mL methanol and 36 mL phosphate buffer) prior to cutting and crushing. These solvents were analysed for PLFA. Total concentrations of PLFA and cell density estimates were calculated assuming that the outer 3 mm of rock core material was extracted during rinsing based on Sherman et al. (2007). Sensitivity of cell density estimates to the depth of extraction was examined and is presented in Appendix A1.

2.2.1.4 Rock Core Interiors

The outer surfaces of the six cores were removed using a hydraulic jack equipped with stainless steel chisels (Onstott et al., 2003). The purpose of this step was to remove, at a minimum, the outer 5 mm of rock material that may have been impacted during drilling. The thickness of material removed was based on the work of Sherman et al. (2007). The interior pieces of the cores were crushed using a stainless steel mortar and pestle, and then powdered using a stainless steel puck mill, both modeled after Sherman et al. (2007). The stainless steel mortar and pestle and puck mill were washed and rinsed with distilled water, and solvent rinsed

with dichloromethane, methanol, hexane and acetone in between each sample preparation. Blank analysis of pre-combusted sand crushed in the stainless steel mortar and pestle and puck mill were found to contain no detectable PLFA.

2.2.1.5 F1 Hydrocarbon Characterization

Petroleum hydrocarbons present as distributions of individual alkanes and/or an unresolved complex mixture (UCM) "hump" are potential carbon and energy sources for microbial communities living within deep subsurface rocks. Further, the comparison of hydrocarbons present in rinse versus interior samples provided a further means of assessing the potential for drilling related impacts on the cores. Analysis of the non-polar (F1) fractions of the PLFA silica gel fractions was carried out via GC-MS in order to further assess the potential for drilling related impacts.

2.2.2 PLFA Results

2.2.2.1 PLFA Concentrations

All six rock core interiors and their exterior rinses were found to contain PLFA (Figure 2). In all six cases, the total concentrations of PLFA in the core interiors were equal to or lower than those in the rock core rinses. Interior PLFA concentrations across the six rock core samples were similar, ranging from 0.4 to 1.5 ng of PLFA per gram of rock. PLFA concentrations in four of the exterior rinses were comparable to these values, ranging from 1.0 to 3.4 ng of PLFA per gram of rock. Exterior rinses from 464o_Queenston and 529o_GeorgianBay yielded very high concentrations of PLFA, 17.8 and 76.2 ng of PLFA per gram of rock, respectively, relative to the other four core rinses.

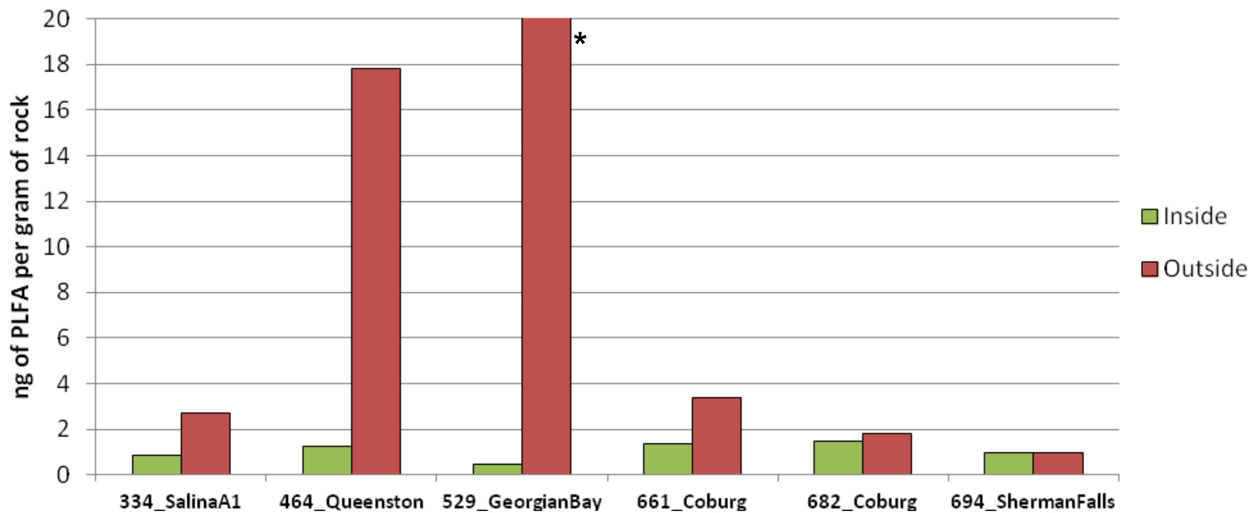

Figure 2: Total concentrations of PLFA (µg of PLFA per gram of rock) detected from core rinses and core interiors; (*) Actual 529o_GeorgianBay concentration is off the scale of this graph at 76.2 ng/g

2.2.2.2 Cell Density Estimates

Total amounts of PLFA detected, and cell density estimates based on PLFA concentrations for the rock core interiors and the rock core rinses, are shown in Tables 4 and 5, respectively. Cell density estimates are based on a conversion factor of $6x10^4$ cells per picomole of PLFA (Green and Scow 2000). For the rock core interiors, cell density estimates ranged from $1x10^5$ to $3x10^5$ cells per gram of rock (Table 4). Cell density estimates for the rock core rinses ranged from $2x10^5$ to $2x10^7$ cells per gram of rock (Table 5). The cell density estimate for drilling water sample DGR8-014, at $8x10^4$ cells per mL, is low compared to cell density estimates for the interiors of the rock core samples, which are on the order of 10^5 cells per gram of rock. A low cell density estimate for this drilling water sample, in combination with a lack of quantifiable PLFA in the three other drilling water samples, suggests that the drilling water is likely not a major source of bacterial cells, and thus PLFA, to the rock core samples.

Table 4: PLFA Results for Rock Core Interior Samples

Core	Mass of interior material extracted for PLFA (g)	Total mass of PLFA detected (µg)	PLFA concentration (ng PLFA/ g of rock)	Cell density estimates (cells/ g of rock)
334i_SalinaA1	1108.4	1.3	0.8	$2x10^5$
464i_Queenston	875.0	1.1	1.2	$3x10^5$
529i_GeorgianBay	1257.8	0.6	0.4	$1x10^5$
661i_Cobourg	972.3	1.3	1.4	$3x10^5$
682i_Cobourg	1008.3	1.5	1.5	$3x10^5$
694i_ShermanFall	859.4	0.8	1.0	$2x10^5$

Table 5: PLFA Results for Rock Core Rinses

Core Rinse	Mass of rock material extracted during rinsing (g)	Total mass of PLFA detected (µg)	PLFA concentration (ng PLFA/g of rock)	Cell density estimates (cells/g of rock)
334o_SalinaA1	525.5	1.4	2.7	$6x10^5$
464o_Queenston	523.3	9.3	17.8	$4x10^6$
529o_GeorgianBay	445.2	33.9	76.2	$2x10^7$
661o_Cobourg	456.8	1.5	3.4	$8x10^5$
682o_Cobourg	459.2	0.8	1.8	$4x10^5$
694o_ShermanFall	429.4	0.4	1.0	$2x10^5$

2.2.2.3 PLFA Distributions

The relative abundances (in mole percentage) of individual PLFAs detected in the extracts of the rock core interiors and rock core rinses are illustrated in Figure 3. The three most abundant PLFA were 16:0, 18:1 and 18:0 in all rock core rinses and all rock core interiors. Generally, the samples contained relatively low abundances of other PLFAs, with the exception of 334o_SalinaA1 and 464i_Queenston, both of which only contained 16:0, 18:1 and 18:0. These three major PLFA are not diagnostic for particular microbial groups. Of the other PLFA observed, there are a number of branched and cyclopropyl PLFA. Branched PLFAs are interpreted to represent gram-negative bacteria (Green and Scow 2000) and are often related to sulphate-reducing microorganisms in these types of subsurface environments (Mauclaire et al., 2007; Stroes-Gascoyne et al., 2007); however, they can be produced by other organisms. Cyclopropyl PLFA have been shown to be related to stress conditions during cell growth (Kieft

et al. 1994; Petersen and Klug 1994), consistent with the expectation that cells in such an environment would be under stress. A number of saturated PLFA were observed in these samples; however, these PLFA are not considered diagnostic.

PLFA patterns in the outer rinse samples showed some similarities. PLFA distributions in 464o_Queenston and 529o_GeorgianBay were highly similar, and the pattern of major PLFA was similar for 334o_SalinaA1. The PLFA distribution in 661o_Cobourg and 682o_Cobourg were similar. 694o_ShermanFall was distinct from the other rinses. In contrast, PLFA distributions for core interiors were distinct from the corresponding rinses in all cases. Further, interior PLFA patterns varied between all core samples. The variation in PLFA distributions between core interior samples indicates that there are differences in either the microbial community composition or environmental conditions experienced by the cells that produced them.

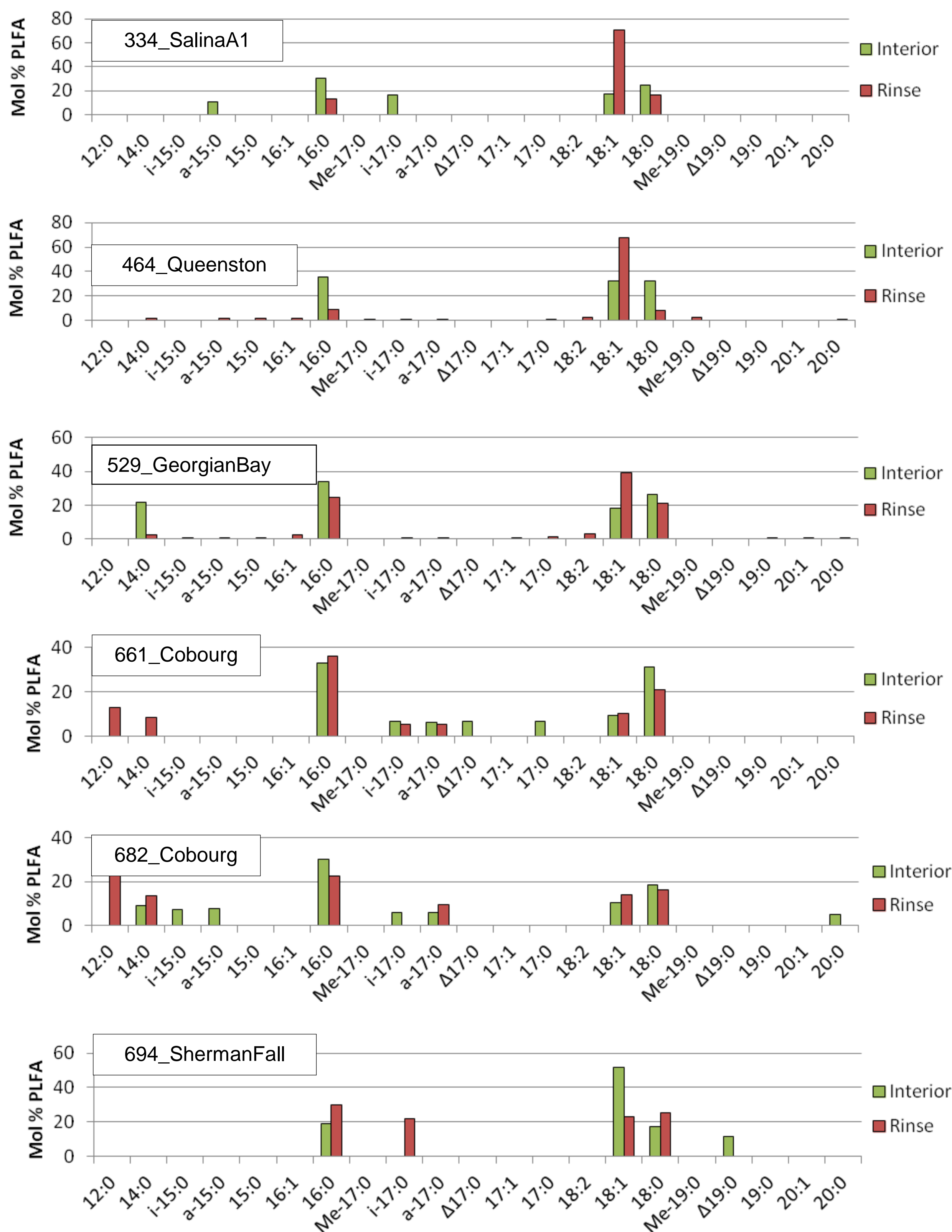

Figure 3: Relative abundances (mol%) of PLFA detected in rock core interiors and rock core rinses

2.2.2.4 Rock Core Appearances

334_SalinaA1 and 661_Cobourg were found to contain very thin natural fractures, along which the cores broke once the pressure of the hydraulic jack was applied (Figure 4). The material along these fractures was very dark and oily in nature. For 334_SalinaA1, the darker material was initially analysed separately from the remainder of the core. The PLFA composition and concentrations did not differ between the dark material and the remainder of the 334_SalinaA1sample. For 661_Cobourg, the fracture ran along the entire length of the core; therefore, the darker material was analysed in combination with the rest of the interior material.

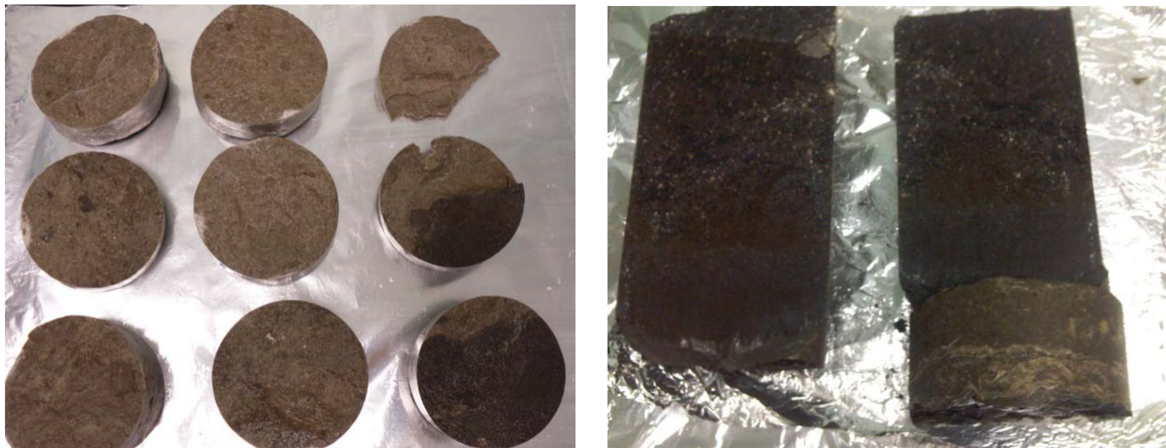

Figure 4: Appearance of Rock Core Samples – (A) Individual slices of 334_SalinaA1. The two darker pieces in the bottom right corner resulted from a break along a natural fracture in the core. (B) Interior of 661_Cobourg, illustrating the dark material covering the surfaces along a natural fracture in the rock

2.2.3 Hydrocarbon Profiles

The non-polar (F1) silica gel fractions of the PLFA method contain any hydrocarbons extracted into the total lipid extract. Analysis of the F1 fractions of all core rinse and interior samples yielded distinct responses between core rinses and interiors, as well as distinct patterns related to the geologic units from which the samples came. This observation supported the indigenous origins of the PLFA with which the hydrocarbons were originally extracted. Only two drilling water F1 fractions were analysed. UCM was observed in the F1 of drilling water sample DWR DGR8-012, while no hydrocarbons were observed in the drilling water sample DWR DGR8-015.

2.3 DNA ANALYSES OF ROCK CORE MATERIAL

2.3.1 Overview

Amplified DNA-based approaches are based on extraction of DNA from microbial cells present within a sample and subsequent amplification and characterization of the DNA. These approaches offer a combination of very high theoretical sensitivity (e.g., ~ 5 microbial cells under optimal conditions or ~100 *E. coli* cells added to soil (Sabat et al., 2000)) and unequaled potential for exploration of microbial diversity and functional potential across a range of specificities (e.g., species – domain levels). In recent decades, these technologies have become the standard approach for the tracking of microbial population structure in environmental samples. The method was first developed to circumvent the so-called "Great Plate Count Anomaly", where ~99% percent of microorganisms visible by microscopy in seawater proved undetectable by cultivation-based approaches (Staley and Konopka, 1985; Amann et al., 1995). Subsequent work has revealed that this same challenge applies to essentially all environmental samples, including those from deep subsurface and hard rock environments (Fredrickson and Onstott, 1996; Pedersen, 1996; Colwell et al., 1997; Takai et al., 2001; Baker et al., 2003; Onstott et al., 2003; Chivian et al., 2008). Thus, during the design phase of this project, the decision was made to rely primarily upon DNA-based techniques as an independent verification, as well as to provide community compositional detail in support of the PLFA dataset.

Although a variety of powerful new "omics"-enabled approaches (e.g., "genomics") now exist for assessing the details of microbial communities in the environment, they are costly and reserved for special cases (Nealson and Venter, 2007; Rusch et al., 2007; Yooseph et al., 2010; Schonknecht et al., 2013). From a practical perspective, the main molecular target for explorations of the genetic diversity of life in the environment remains the same as it has been for the past several decades: the small subunit of the ribosomal RNA gene (16S rRNA or SSU rRNA). This gene encodes a portion of the central "housekeeping" apparatus of all cellular life, the ribosome. Variations in the sequence of the gene encoding this structure have proven extremely useful for inferring evolutionary relationships between organisms of all types (Woese and Fox, 1977; Woese et al., 1990; Rappe and Giovannoni, 2003). As a result, a unified scheme for the organization of cellular life into three domains: Archaea, Bacteria and Eukaryote, has come into broad usage and vast databases (e.g., the ribosomal database project, RDP (Cole et al., 2009)) are curated, enabling the base-by-base comparisons of query sequence against up to 3 million reference sequences, in the case of RDP. The relationships of the DNA sequences detected from core and reference materials over the course of this project form the basis of the life identifications in this report.

More specifically, over the past several decades, the standard tool for 16S rRNA gene assessments of prokaryotic life in any environmental sample (Pace et al., 1985; Giovannoni et al., 1990; Barns et al., 1996; Onstott et al., 2003; Moser et al., 2005) has been the amplified rRNA gene library. This approach relies upon polymerase chain reaction (PCR) (Saiki et al., 1988) to amplify target DNA from mixed DNA extracts, generating millions of copies of a target sequence, which can easily be studied independently through cloning. Although the approach offers both advantages and disadvantages, it remains the gold standard for obtaining DNA sequences of phylogenetically-informative length from prokaryotic cells (e.g., full-length of the ~1,500 base pair [bp] 16S rRNA gene from bacteria and archaea). However, this approach is labor-intensive and the number of sequences obtained tends to be modest (e.g., ≤ 96 with the most popular platform); thus, the cost per sequence is relatively high.

More recently, several high-throughput, next generation ("next gen") approaches for assessing the diversity of 16S rRNA genes in environmental samples have become available. The most common in use today is the 454 Pyrotag Library (Parameswaran et al., 2007). This pyrotag method utilises a bar-coding approach to enable the multiplex sequencing of much larger numbers of PCR product molecules. The major advantage of this approach is that a very large number of sequences (e.g., ~3,000 – 20,000) can be obtained at 10- to 100-fold lower cost per base than with clone libraries (Sogin et al., 2006; Tringe and Hugenholtz, 2008). Disadvantages include the potential erroneous inflation of diversity assessments due to higher error rates (Kunin et al., 2010) and relatively short read-lengths (~500 bp). All things considered, most workers feel that the advantages of next gen analyses outweigh the disadvantages for surveys of environmental microbial diversity. For this study, because the major objectives are to assess 1) the presence or absence of microbial life, and 2) variations in community structure between putative uncontaminated inner core samples and various controls, the much deeper coverage afforded by pyrotag analysis is indicated. As an independent validation and assessment of contaminants that may have been introduced during the pyrotag analysis itself, clone libraries were also constructed in parallel for representative samples (see DNA Results).

This report details the systematic optimization of methods enabling the detection of DNA-based archaeal and bacterial signatures in hard rock samples and their subsequent application to clay-rich, very low biomass, low-permeability samples. However, unlike the case of PLFA analysis, these samples either contained microbial DNA at below detection limits (or were free of it entirely), or, alternatively, possess some feature that degrades or binds free DNA, thereby preventing its detection. Thus, it became necessary to push DNA extraction and PCR-based amplification protocols to their technical limits. PCR-based technologies are highly sensitive to false positives under these conditions. Unfortunately, the differentiation of reagent/method contamination from indigenous environmental signatures is extremely difficult. The results presented here are, thus, considered in the context of a parallel calibration of method sensitivity through the addition of known numbers of microbial cells to sterilised rock and an exhaustive analysis of PCR and DNA extraction kit reagent contamination. This work represents what may be the most exhaustive examination of potential DNA contamination of continental rock samples ever undertaken using multiple approaches. The experimental and analytical approach developed specifically for this sample set should assist in the development of future site investigation plans to detect and characterise life in low-permeability, low biomass environments under consideration for development of a deep geological repository.

2.3.2 Materials and Methods (DNA)

2.3.2.1 DNA Extraction

For drilling fluid samples, Sterivex (Millipore) housings were opened with sterilised pliers and filter membranes removed using flamed hand tools (forceps and razor blade). One quarter of each filter (ripped into ~pea-sized pieces) was used for extraction and the remainder archived for potential future use. DNA was extracted with the MoBio Power Soil DNA Isolation Kit (Cat# 1288-100). The manufacturer's instructions were followed with the addition of a freeze-thaw step, as described in (Onstott et al., 2003), and the resulting extracts stored at -80°C. Attempts were made to quantify these DNA extracts using a NanoDrop 2000 spectrophotometer (Thermo Scientific, OH), but none produced absorption spectra indicative of nucleic acids (data not shown). To test the effect of increased phosphate, one quadrant of the filter from drilling water

sample 334_DW was extracted using the same extraction kit, but with a solution of 1M phosphate buffer in 15% ethanol substituted for the kit's solution C1 (0.2M phosphate), as described in (Wankel et al., 2012). As this modification failed to provide detectably stronger amplifications by polymerase chain reaction (PCR), this variation to the sample extraction procedure was not pursued further.

Upon receipt at DRI, the test and experimental cores were stored at 4°C and -20°C, respectively. The test cores were unpreserved cores from another borehole that were utilised for early method development and are not discussed further in this report. In the laboratory, cores were processed in a scrupulously cleaned laminar-flow hood and handled with fresh nitrile gloves and autoclaved tools, frequently re-sterilised with ethanol and flame. Gloves were changed often and cleaned with ethanol every few minutes. An ethanol-soaked, flamed aluminum-foil-lined box was used as a work area utilizing dry ice (under the foil) to keep core material frozen. To obtain pristine inner subsamples, cores were scored ~3 cm deep and 6 cm from the end with a Dremel rotary tool and diamond cutting wheel (ethanol-soaked). To permanently mark all outer surfaces so that they would not be erroneously included as part of the pristine samples for further analysis, the cores were lightly spray painted (Painters' Touch, Rust-Oleum, USA). All samples with any paint present were excluded from further analysis. Freshly-painted core was wrapped in several layers of plastic wrap (Saran Wrap). Previous experimentation has shown this product to be functionally sterile inside the roll (DeFlaun et al., 2007). A section of each core was removed by driving a sterile cold chisel into the core, thus liberating a puck-like piece of the core. The chisel was cleaned with ethanol and flamed every two strikes during this process. The outsides of the puck were subsequently pared away using an ethanol- and flame-sterilised chisel. Inner and outer pieces were separately ground to a fine powder (≤10 μm, verified by light microscopy) using a "dolly pot"-type 3-inch diameter rock crusher (Prospectinggear.com) that was washed and ethanol soaked/flamed between samples. To verify that the paint was not a source of microbial contamination, a sterilised 45 mm Supor PES filter (Pall Corporation, USA) was painted in the same manner as the cores as a control.

Samples of rock powder (10 g) and a paint-covered filter control (to verify that the paint used did not contain contaminating DNA) were transferred to individual bead tubes provided by the MoBio Power Max Mega Prep Soil DNA Kit (Cat # 12988-10), and DNA isolated following the manufacturer's instructions – with additional steps to enhance recovery and concentrate the DNA as previously described. These modifications included: 1) the inclusion of a freeze-thaw step, 2) reduction in the final elution volume (5 mL instead of the prescribed 15 mL), and 3) isopropanol precipitation followed by re-suspension in a small volume (50 μL) of nuclease-free water (Molecular Biologicals International Inc.). As with the drilling water samples, attempts to quantify these DNA extracts using a NanoDrop 2000 Spectrometer failed to produce absorption spectra indicative of nucleic acids (data not shown).

2.3.2.2 Molecular Methods and Statistical Analysis

DNA amplifications were carried out both at DRI (for extraction efficiency testing and DNA library construction) and at a contract sequencing laboratory used for 454 pyrotag analysis. In-house PCR amplification for drilling water, core and the various control extracts were performed using 1 μL of DNA extract/template in a 20 μL reaction volume, with 10 μM forward and reverse primers for Archaea or Bacteria, 2X GC Taq Buffer, 2.5 mM dNTPs, LA Taq (TaKaRa), and nuclease free water. To test for the possibility that PCR inhibitors co-eluted with template DNA, the same reactions were carried out on diluted DNA extracts as well (1:2, 1:20, 1:200) and no amplicons were obtained for any. For Archaeal and Bacterial reactions, 21F (5'-

TTCCGGTTGATCCYGCCGGA- 3') and 8F (5'- AGAGTTTGATCCTGGCTCAG- 3') forward primers were used, respectively. The universal 1492R reverse primer (5'-TACCTTGTTACGACTT- 3') was used for both amplifications. Reaction conditions included: a hot start (94°C for 5 min); followed by 30 - 35 cycles of 94°C for 1 min, 50°C for 30 s, 72°C for 2 min; and a final elongation step of 72°C for 5 min. All PCR reactions were conducted in a GeneAmp PCR System thermalcycler (Model 9700, PerkinElmer, USA). In some cases, a second round of PCR (half-nested) was required to obtain visible products (Takai et al., 2001). This involved using 1 µL of the full-length PCR product after an initial 32 cycles, 10 µM of the previously used forward primers (21F for Archaea and 8F for Bacteria) and 10 µM of an internal reverse primer, 926R (5'- CCGTCAATTCCTTTRAGTTT- 3'), for Bacteria and 912R (5'-CTCCCCCGCCAATTCCTTTA – 3') for Archaea. Otherwise, the nested reactions utilised the same reagents and conditions (32 cycles) as the full-length PCR amplifications. Although included in the interest of completeness, bacterial nested reactions were used only for methods development and none of these results are shown in this report.

Pyrotag sequencing services were provided by Research and Testing Laboratory (Lubboc, TX), partially analysed as described in their "Data Analysis Methodology" document (McGowan, 2006). This provider utilised bar-tagged versions of the same primers noted above for their amplifications. However, after failing to amplify archaea from any of the extracts, they noted pyrotag sequences from three samples using a highly degenerate tagged Archaeal primer, Arch349F (5'-GYGCASCAGKCGMGAAW-3', see Section 2.3.3.2). Briefly, after amplification, raw sequence data was quality checked, de-noised and chimera checked by in-house procedures. After conversion to the FASTA format, taxonomic relationships for the remaining sequences were determined by grouping into operational taxonomic units (OTUs) with 100% identity (0% divergence) using USEARCH (Edgar, 2010). For each cluster, the seed sequence was put into a FASTA-formatted sequence file, which was queried against a database of high-quality sequences derived from the US' National Center for Biotechnology Information (NCBI) using a distributed.NET algorithm that utilises BLASTN+ (KrakenBLAST www.krakenblast.com). Percent identities of the sequences were classified at the appropriate taxonomic levels (using well-characterised database sequences) based upon the following criteria. Sequences with identity scores greater than 97% identity (<3% divergence) were resolved at the species level; between 95% and 97% were resolved at the genus level; between 90% and 95% were resolved at the family level; between 85% and 90% were resolved at the order level; between 80 and 85% were resolved at the class level; and 77% to 80% were resolved at phyla level. Any match below this percent identity was discarded.

For 16S rRNA gene libraries, PCR amplicons were purified (UltraClean™ GelSpin™ DNA Purification Kit, MoBio Laboratories, Inc.) and subjected to molecular cloning methodology using TOPO®-TA kits (Invitrogen; Carlsbad, CA). Selected clones were unidirectionally sequenced (Functional BioSciences; Madison, WI) and aligned, matched with nearest cultured neighbours, and checked for chimeras using Greengenes (DeSantis et al., 2006). Alignments were refined and phylogenetic relationships determined using MEGA (Tamura K et al., (2007)). Statistical analysis of pyrotag DNA sequence datasets was based on percentages of bacterial groups at phylum- and species- levels. Principle coordinate and cluster analyses were performed using the PAST software package (Hammer, 2001). Plots were generated based on the dissimilarity calculated by the Bray-Curtis method (Anderson et al., 1990). Partial 16S rRNA sequences (~1,500 bp) obtained in these libraries were submitted to Genbank and assigned accession numbers KF383957 – KF383969.

2.3.2.3 Sensitivity Calibration and Controls

Gram-positive (*Bacillus subtilis*) and Gram-negative (*Proteus vulgaris*) bacterial test strains were grown aerobically with shaking (150 RPM) to late exponential phase in liquid Luria Broth (LB, DIFCO). Cells from 30 mL aliquots were collected by centrifugation (3,200 relative centrifugal force (rcf) for 15 min), washed, and re-suspended or diluted in sterile phosphate-buffered saline to achieve desired cell densities (Grice et al., 2008). Culture densities were determined by flow cytometry and verified by direct microscopic counts using a Petroff-Hausser counting chamber. Known numbers of test bacterial cells were added to 0.25 g portions of baked (500°C for 4 hours) rock powder from core 494 to achieve final concentrations of 0 – 10^7 cells per gram of rock in 10-fold increments. Total added control DNA was extracted using MoBio Power Soil kits (Cat# 1288-100, MoBio, Salano Beach, CA) in the same manner as for the drilling water filter samples. In this case, the Mega Prep kit that was used on the actual test cores was not employed due to cost and labor that would have been required for the large number of control samples processed. This deviation in protocols was deemed justifiable because the kit chemistries are the same and one sample set (334i_Salina A1 and 334o_Salina A1, (Fig. 5) performed similarly between the two kits. A “Large kit” vs. “small kit” calibration with added cells was also performed (see Figure 15) and performed nearly identically. PCR was performed under the same conditions as for the other test extracts.

2.3.3 DNA Results

2.3.3.1 PCR Assay Detection of bacteria and Archaea

A convenient and sensitive assay to detect a given group of microorganisms is simply to perform a PCR reaction followed by gel electrophoresis – testing for the presence or absence of a product of the expected size. Figure 5 shows the results of attempts to obtain archaeal PCR reactions from the samples considered for this project and the effect of several different variations on the PCR protocol. As evidenced from Figure 5A, 32 cycles of PCR were insufficient to produce full-length (~1,500 bp) products for any of the DNA extracts (drilling water, inner and outer core splits). An additional three cycles of PCR (35 total, 5B), likewise, failed to produce visible products from any of the samples. Although it has the potential to introduce artifacts into DNA libraries (e.g., propagation of primer bias (Suzuki and Giovannoni, 1996)), nested PCR is sometimes used to boost the sensitivity of PCR used for rRNA gene libraries (Takai et al., 2001) and was attempted here. As shown in Figure 5C, this approach failed to produce archaeal PCR products for the majority of samples as well. Although three weak products were obtained (464o_Queenston, 670i_Cobourg and 670o_Cobourg), none were of the expected size and further analyses were not pursued. Taken together, these results indicate that, although a limit of detection was not determined for this domain of life, very low densities (or a complete absence) of archaea are likely in these samples.

Figure 5: **Stained Agarose Gel Illustrating Presence or Absence of Putative Archaeal PCR Products Under Conditions Tested: (A) 32 Cycles; (B) 35 Cycles; and (C) Half-nested reaction.**

Bacterial PCR tests were also conducted and the results are summarised in Figure 6. Although bacterial nested PCRs were performed for drilling fluid samples and bands of various sizes were obtained, these were not analysed further because modest bands of the correct size were also obtained without resorting to this bias-sensitive technique. For example, weak banding of the correct size can be seen in Figure 6A (PCR amplification of extracts from samples 334i_Salina A1, 494i_Queenston, and 592i_Georgian Bay). Although the presence of these bands did initially support the possibility that indigenous bacterial DNA was amplified, an additional three cycles of PCR did not significantly strengthen their performance (Figure 6B). In fact, when 35 cycles of PCR were employed, all of the extracts amplified with a similar, but weak, intensity, including baked control extracts from cores 494_Queenston and 682_Cobourg, and a reagent/process blank for our DNA extraction procedure (see Figure 6B).

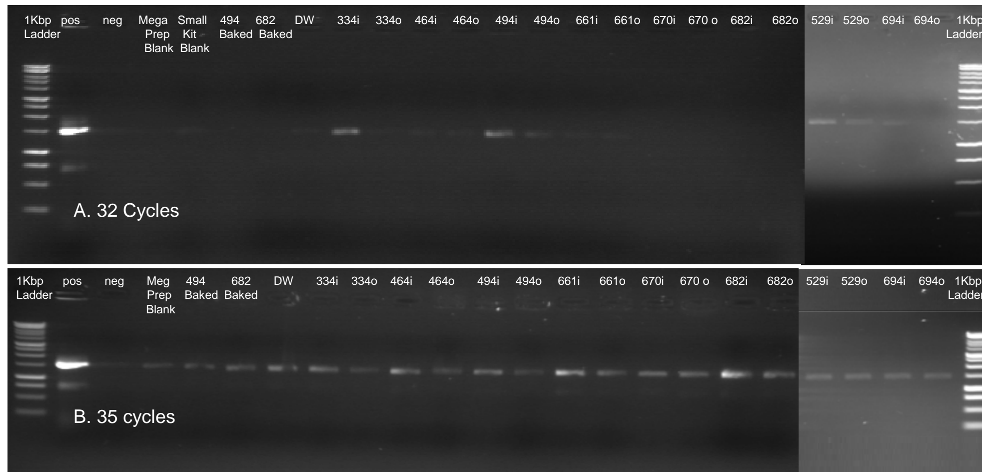

Figure 6: Stained Agarose Gel Illustrating Presence or Absence of Putative Bacterial PCR Products Under Conditions Tested: (A) 32 Cycles; (B) 35 Cycles

2.3.3.2 454 Pyrotag Analysis of Archaea

None of the sample and control extracts produced amplifiable archaeal DNA for either in-house PCR assays or with the standard 454 pyrotag sequencing method utilised by the subcontracted provider. However, a limited number of archaeal positives were obtained when the provider of the 454 dataset utilised a highly degenerate internal forward primer, Arch349F (Takai and Horikoshi, 2000) instead of the their usual 28F. The results of this procedure are summarised in Table A2.2 (Appendix A). Using this approach, archaeal sequences were obtained from three of the samples. All were from extracts obtained with the “small” DNA extraction kit (MoBio 1288-100) and affiliated with a symbiont of marine sponges, *Cenarchaeum* sp. (Preston et al., 1996): 334DW$_s$_Salina A1 at 78.7%, 334o$_s$_Salina A1 at 97.5% and 464DW$_s$_Queenston at 99.4% of the total archaeal signature. The same three samples also had minor contributions of a marine Thaumarchaeote, *Nitrosopumilus* sp. (Konneke et al., 2005) at 21.1, 2.4, and 0.14% of the total, respectively. Other than these results, no archaea were detected in any samples examined for this project.

2.3.3.3 454 Pyrotag Analysis of bacteria

Figure 7 provides a phylum-level overview of bacterial community structure across the entire sample set from pyrotag libraries. From this broad view of bacterial diversity, it was observed that this collection falls into three distinct groupings, denoted here as Batch I, II and III. In this dataset, the major pattern determinant, which overpowers all other factors, is into which of the three batches sent to the sequencing provider a given extract happened to fall. The most abundant specific taxa in these various groupings are considered elsewhere in the report, but the major phyla detected, depending upon the sample batch, include Proteobacteria, Firmicutes, Actinobacteria, and Spirochaetes. These groupings are further examined by cluster analysis at the phylum- and species- level (dendrograms in Figures 8 and 9). In both cases, the patterns into which the samples fall (e.g., dendrogram clades), perfectly preserve the qualitative pattern from Figure 7. The data presented in Figure 9 appears to retain a minor impact of sample-specific information in the Batch II clade. This clade is clearly divided into two

subclades, with all of the small kit samples occupying one and the remainder of the samples occupying the other (label denoted in italics in the figure).

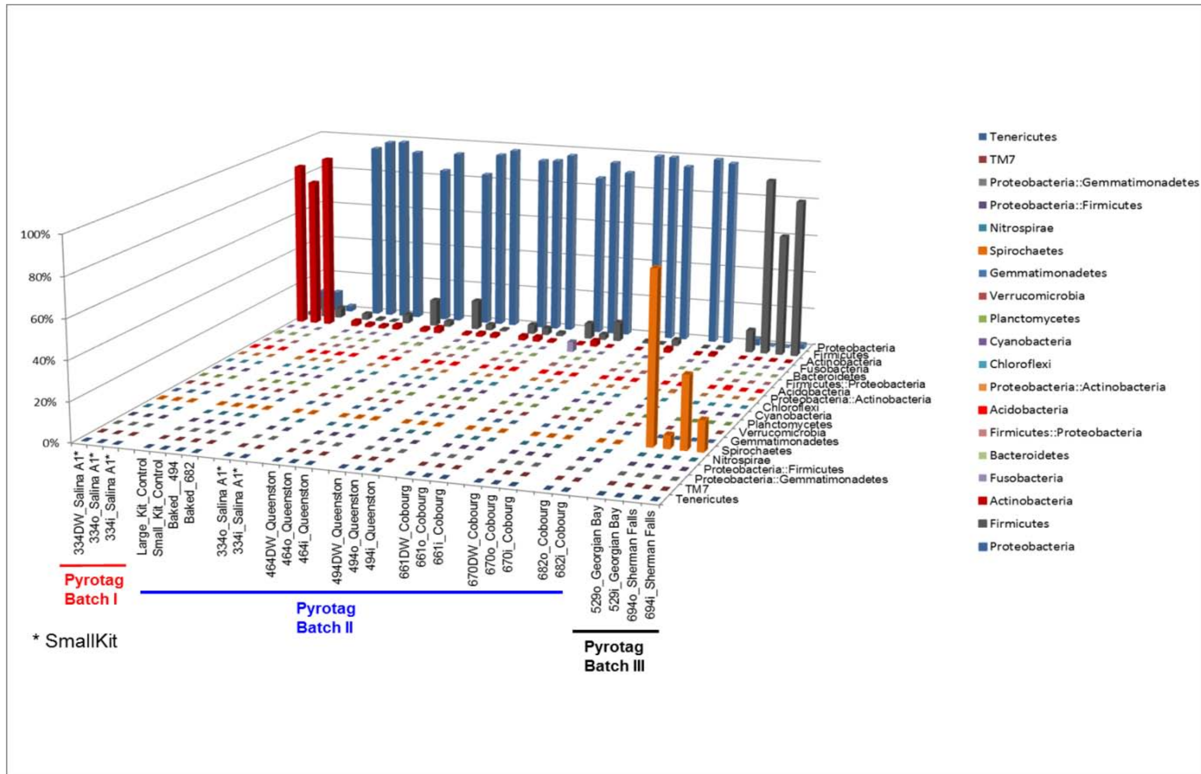

Figure 7: Bacterial Phylum-Level Overview of the Entire 16S rRNA Gene Pyrotag Dataset Illustrating the Manner in Which Samples Cluster by Groupings Submitted to Sequencing Provider

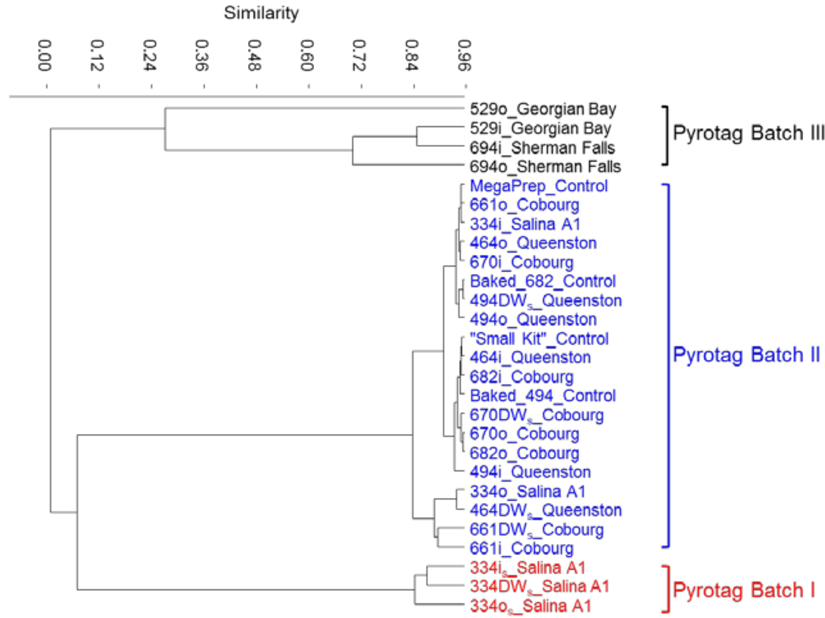

Figure 8: Bacterial Phylum-Level Dendrogram Based on Figure 7 Data, Showing How Patterns of Microbial Diversity are Controlled by Batch Sent to Sequencing Provider.

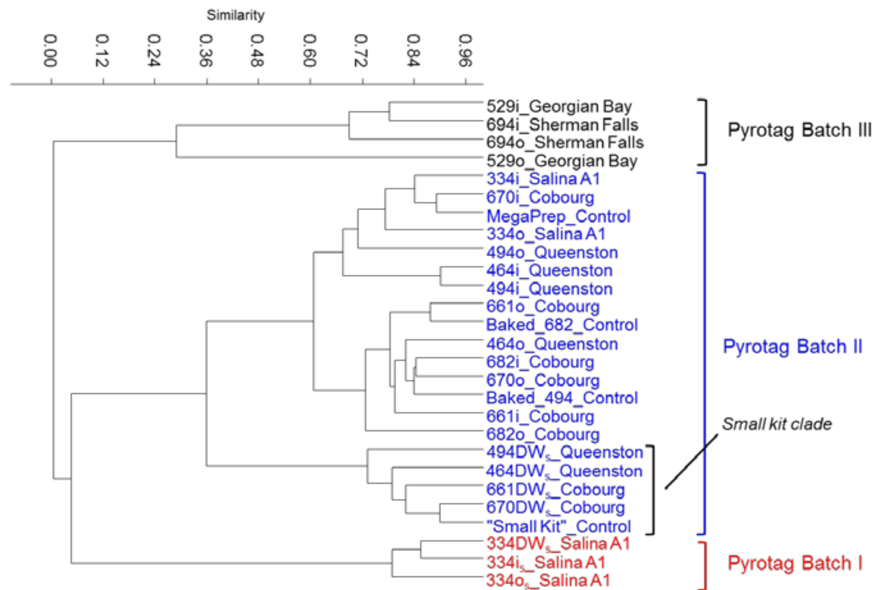

Figure 9: Bacterial Species-Level Dendrogram Based on Data for all Cores and Controls Affiliated with This Project.

Figures 10, 11, and A2.1 (Appendix A) explore patterns of bacterial species abundance for major community members in each of the three batches of DNA extracts handled by the sequencing laboratory. In each of these figures, a rank abundance curve of the bacterial

lineages present at more than 1% of the total is presented. For each batch, this top 1% cut-off produced a relatively restricted number of taxa: 12 (of 136 total in Batch I), 8 (of 271 total in Batch II) and 15 (of 28 total in Batch III). In each case, the most abundant microorganisms were different, but the overall patterns were similar. Most strikingly, for all samples in each batch, one-to-three taxa comprise most of the total sequences detected, regardless of the type and origin of sample. For example, for Batch I (samples 334DW$_s$_Salina A1_334i$_s$_Salina A1, and 334o$_s$__Salina A1, Fig. 6), the most abundant microorganism, comprising over half of the entire population for each sample was affiliated with *Mycobacterium bolletii*. This set of three samples was originally submitted as a trial run to evaluate the 454 pyrotag procedure for this purpose.

Batch II was a set of eighteen samples sent to the pyrotag sequencing lab that comprises the bulk of the samples considered for this project. Although Batch II showed greater variation, more than half of all sequences detected were comprised of some combination of *Brevundimonas diminuta*, *Burkholderia cepacia*, or a yet-to-be-described species of *Burkholdaria* (Figure 11). This sample set did, however, display one sample-specific trend in that the most abundant taxon in all core samples and core controls, *Brevundimonas diminuta*, was not present in the associated drilling water samples. As these samples were extracted using the smaller version of the Power Soil kit, rather than the large kit employed for most of the core samples, a kit-associated trend may be indicated here (see Figure 9).

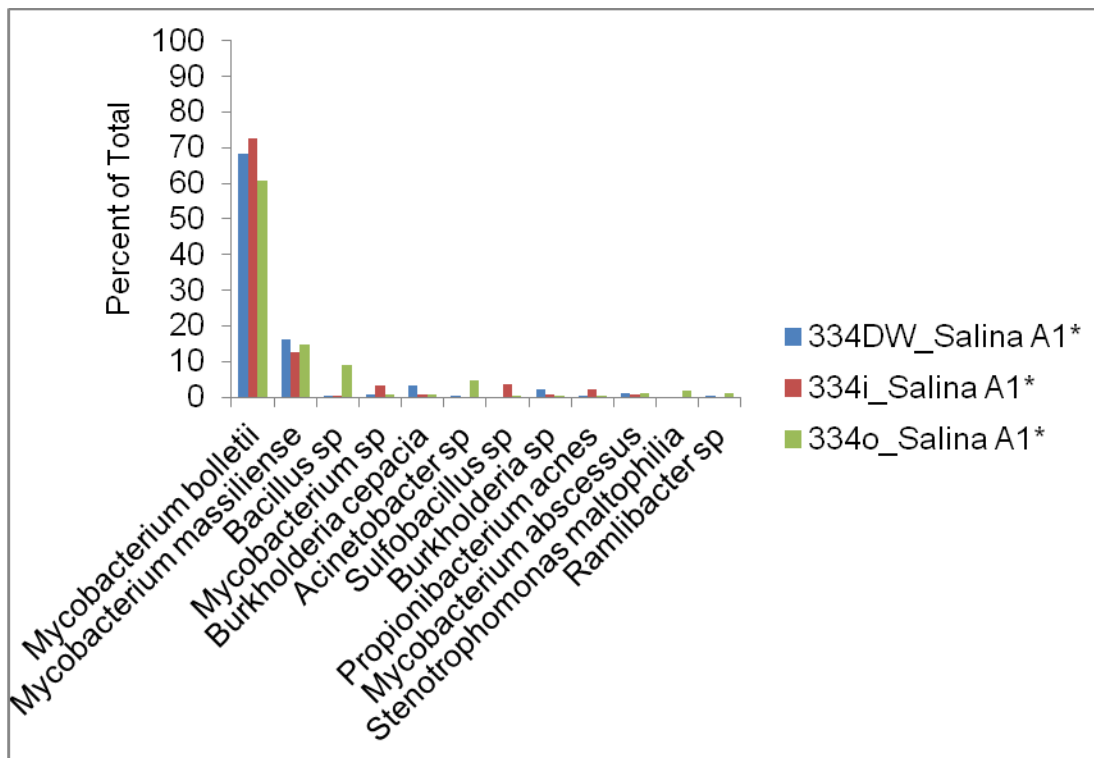

Figure 10: Summary Major Bacterial Species Detected in the Initial of Three Sample Batches Sent to the Sequencing Provider ("Batch I").

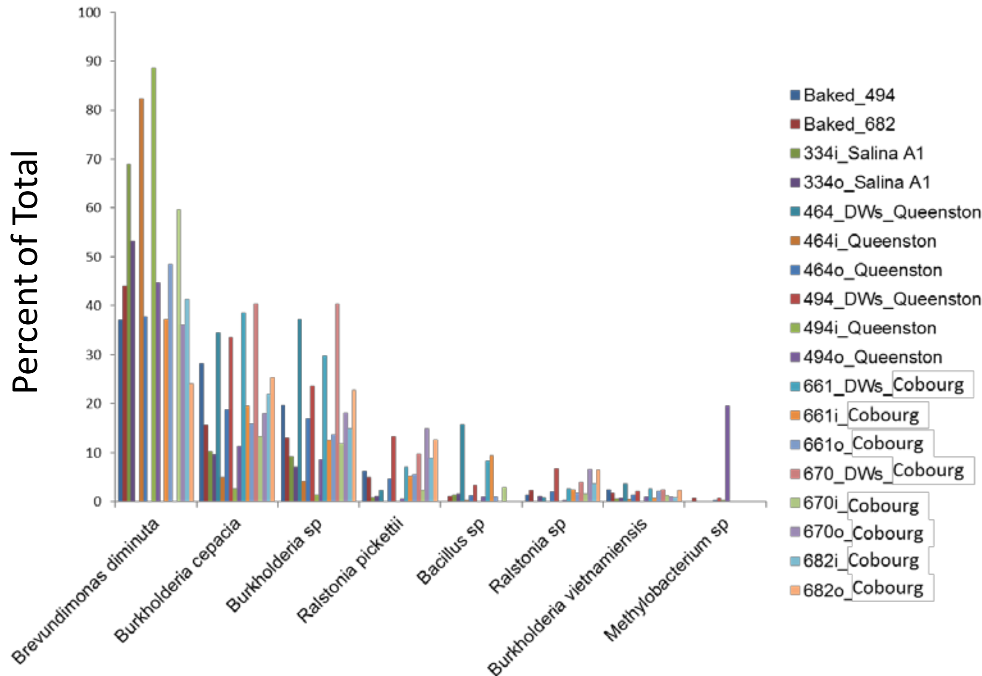

Figure 11: Summary of the Most Abundant Bacterial Species Detected in DNA Extracts from the Largest Batch Sent to the Sequencing Provider ("Batch-II").

Finally, more than half of all the sequences in Batch III, from 529i_Georgian Bay, 529o_Georgian Bay, 694i_Sherman Fall and 694o_Sherman Fall, are affiliated with *Sporanaerobacter acetigenes* or *Leptospira broomii* (Figure A2.1 in Appendix A).

The species-level (highest sensitivity) principal coordinate analyses in Figures 12 and 13 were designed to be compared to one another and were constructed to constrain the impact of early PCR bias effects (e.g., the possibility that disproportionately abundant taxa resulting from this effect might skew apparent population structure). Figure 12 provides an analysis containing all taxa detected in the project. Figure 13 displays the results of the same analysis, but with the dominant taxa removed.

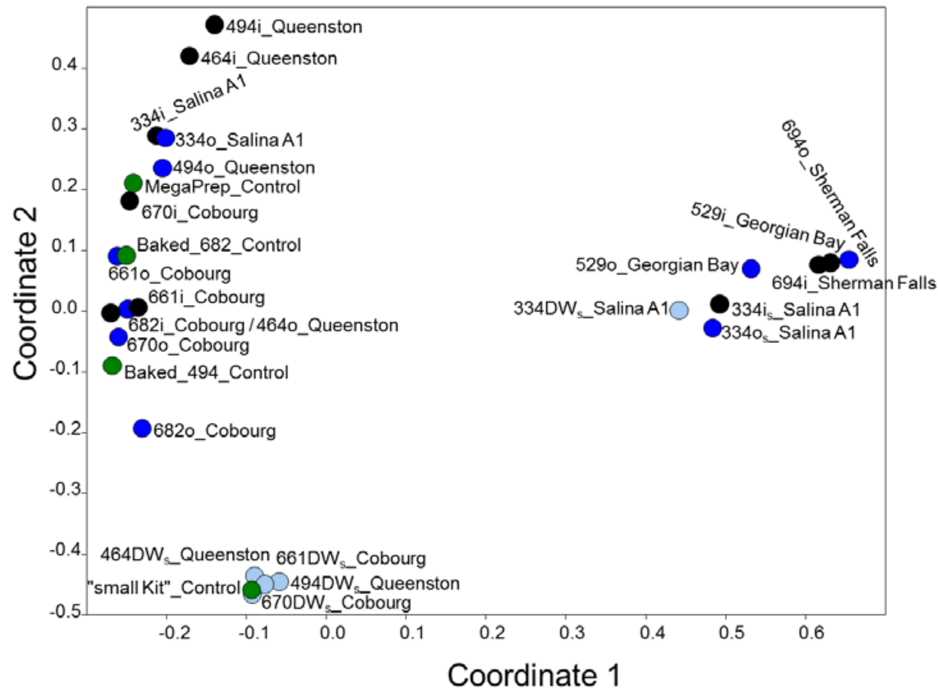

Figure 12: Principle Coordinate Analysis (Species-level) Illustrating Patterns of Diversity Among Abundant Bacteria. Light Blue Circles Represent Drilling Water; Dark Blue Circles Represent Outer Core Subsamples; Black Circles Represent Inner Core Subsamples; Green Circles are Controls. Lowercase "S" Designates "Small Kit".

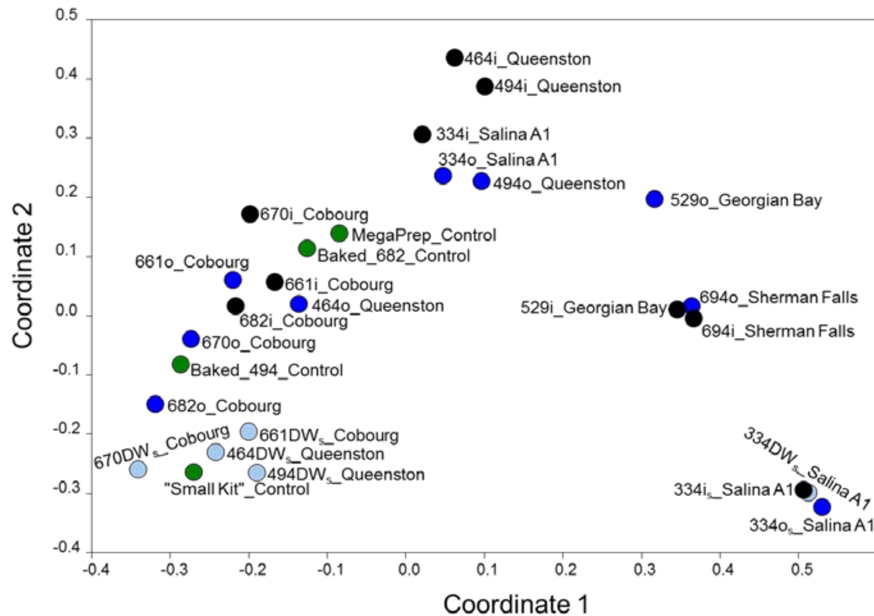

Figure 13: Principle Coordinate Analysis (Species-level) Illustrating Rare Bacterial Diversity (e.g. Dominant Taxa Removed). Light Blue Circles Represent Drilling Water; Dark Blue Circles Represent Outer Core; Black Circles Represent Inner Core; Green Circles are Controls. Lowercase "S" Designates "Small Kit" Extracts.

The unfiltered dataset in Figure 12 shows three strong groupings of samples from the entire dataset. In this figure, all of the samples from Batch II group far to the left of the figure, whereas samples from batches I and III form adjacent groupings far to the right side of the plot. With the exception of a drilling water sample from 334_Salina A1, all of the small kit extracts and controls group tightly near the bottom left of the plot. The overall pattern, as illustrated in Figure 12, is grouping of samples based on the order in which samples were sent to the sequencing laboratory. In marked contrast, however, the removal of the three most abundant taxa from each dataset (see Figure 13) abolishes most of the tendency to group that is shown in Figure 12. In this case, controls and inner and outer samples all mix together more or less randomly across all four quadrants of the plot. Once again, the one trend that does persist in this treatment, however, is that all of the small kit extracts are distributed across the bottom half of the figure, indicating that the effect of kit selection did induce some degree of community partitioning that carried through the sequencing lab batch artifact. All things considered, no sample-defined pattern (e.g., inner vs. outer, or depth of sample) was apparent at the most sensitive (i.e., species) level of taxonomic specificity, even for the large datasets (Table 6) obtained here.

Table 6: Bacterial Pyrotag Summary: Most Abundant Species Detected

Sample	Most abundant taxon (% of total)	Phylum	Major physiology
"Small Kit"_Control	Burkholderia cepacia 40.5%	Proteobacteria	Pathogen/Aerobe
MegaPrep_Control	Brevundimonas diminuta 57.1%	Proteobacteria	Pathogen/commensal, Aerobe from air, water systems
Baked_494_Control	Brevundimonas diminuta 37.2%	Proteobacteria	" "
Baked_682_Control	Brevundimonas diminuta 44.1%	Proteobacteria	" "
334DW$_s$_Salina A1	Mycobacterium bolletii 68.4%	Actinobacteria	Pathogen/commensal from plants, water systems
334o$_s$_Salina A1	Mycobacterium bolletii 60.7%	Actinobacteria	" "
334i$_s$_Salina A1	Mycobacterium bolletii 72.7%	Actinobacteria	" "
334o_Salina A1	Brevundimonas diminuta 53.3%	Proteobacteria	" "
334i_Salina A1	Brevundimonas diminuta 69.1%	Proteobacteria	" "
464DW$_s$_Queenston	Burkholderia sp 37.3% Burkholderia cepacia 34.6%	Proteobacteria	" "
464o_Queenston	Brevundimonas diminuta 37.7%	Proteobacteria	" "
464i_Queenston	Brevundimonas diminuta 82.4%	Proteobacteria	" "
494DW$_s$_Queenston	Burkholderia cepacia 33.8%	Proteobacteria	" "
494o_Queenston	Brevundimonas diminuta 44.7%	Proteobacteria	" "
494i_Queenston	Brevundimonas diminuta 88.6%	Proteobacteria	" "
529o_Georgian Bay	Leptospira broomii 84.0%	Spirochaetes	Aerobe, Opportunistic Pathogen
529i_Georgian Bay	Sporanaerobacter acetigenes 67.4%	Firmicutes	Rumen anaerobe, RDX degrader
661DW$_s$_Cobourg	Burkholderia cepacia 38.6%	Proteobacteria	" "
661o_Cobourg	Brevundimonas diminuta 48.6%	Proteobacteria	" "
661i_Cobourg	Brevundimonas diminuta 37.4%	Proteobacteria	" "
670DW$_s$_Cobourg	Burkholderia sp. 40.5% Burkholderia cepacia 40.4%	Proteobacteria	" "
670o_Cobourg	Brevundimonas diminuta 36.2%	Proteobacteria	" "
670i_Cobourg	Brevundimonas diminuta 59.8%	Proteobacteria	" "
682o_Cobourg	Burkholderia cepacia 25.3%	Proteobacteria	" "
682i_Cobourg	Brevundimonas diminuta 41.4%	Proteobacteria	" "
694o_Sherman Fall	Sporanaerobacter acetigenes 46.3%	Firmicutes	" "
694i_Sherman Fall	Sporanaerobacter acetigenes 48.3%	Firmicutes	" "

Finally, to examine the possibility that indigenous microorganisms were present in the DNA extracts, but obscured by more potent contaminant overprints, a very detailed manual examination of the *least abundant* sequences in a subset of 454 libraries was undertaken. The aim here was to identify lineages that would represent credible presences in the inner core samples, consistent with the deep or saline biosphere or lacking from the various borehole and reagent contaminant controls. Table 7 contains the rarest community members (always << 1% of the total), as determined by the pyrotag pipeline for samples associated with the 334_Salina A1 depth. A second set of samples (Table A2.3 in Appendix A) was examined for samples and controls associated with 494_Queenston. In both cases, the microorganisms detected were

loosely related to the most abundant and were predicted aerobes, cyanobacteria, and human-associated pathogens and commensals.

Table 7: Least Abundant Bacterial Species Detected in Select Samples

334DW$_s$_Salina A1	334o$_s$_Salina A1	334o_Salina A1	334i$_s$_Salina A1	334i_Salina A1
Acidobacterium capsulatum	*Actinomadura sp*	*Burkholderia multivorans*	*Acidithiobacillus ferrooxidans*	*Acinetobacter sp*
Acidobacterium sp	*Burkholderia ambifaria*	*Chloroflexus sp*	*Massilia sp*	*Blastochloris sp*
Anabaena sp	*Burkholderia vietnamiensis*	*Enterococcus durans*	*Pseudomonas aeruginosa*	*Brevundimonas terrae*
Aphanocapsa sp	*Clostridium straminisolvens*	*Enterococcus faecium*	*Staphylococcus lugdunensis*	*Burkholderia ambifaria*
Aurantimonas sp	*Corynebacterium sp*	*Escherichia coli*	*Terrahaemoph.. aromaticivorans*	*Clostridium orbiscindens*
Bacillus licheniformis	*Corynebacterium urealyticum*	*Eubacterium hallii*		*Corynebacterium sp*
Corynebacterium xerosis	*Eggerthella sinensis*	*Klebsiella pneumoniae*		*Cyanospira rippkae*
Cryptosporangium japonicum	*Gordonia bronchialis*	*Methylobacterium sp*		*Escherichia sp*
Enterococcus faecalis	*Gordonia sp*	*Mycobacterium sp*		*Finegoldia sp*
Frankia sp	*Gordonia sputi*	*Phormidium priestleyi*		*Neisseria sp*
Geitlerinema sp	*Kineococcus sp*	*Rhizobium sp*		*Thermomicrobium sp*
Isosphaera sp	*Kineococcus xinjiangensis*	*Rubrobacter sp*		
Klebsiella granulomatis	*Kineosporia aurantiaca*	*Staphylococcus aureus*		
Labrys wisconsinensis	*Methylophilus leisingeri*	*Staphylococcus sp*		
Methylibium petroleiphilum	*Mycobacterium conceptionense*	*Streptococcus salivarius*		
Methylophilus leisingeri	*Mycobacterium terrae*	*Streptomyces armeniacus*		
Methylophilus sp	*Nocardioides sp*		**Small Kit Control**	**Mega Kit Control**
Nocardioides sp	*Olsenella sp*		*Bacillus sp*	*Azospirillum zeae*
Pannonibacter phragmitetus	*Oxalophagus oxalicus*		*Burkholderia glumae*	*Blastococcus sp*
Prochlorococcus sp	*Streptomyces sp*		*Propionibacterium sp*	*Caulobacter sp*
Propionibacterium sp	*Vibrio sp*		*Proteus vulgaris*	*Clostridium sp*
Pseudomonas sp			*Serratia entomophila*	*Escherichia sp*
Pseudonocardia sp				*Mycoplana bullata*
Rhizobium sp				*Mycoplana sp*
Roseomonas sp				*Peptoniphilus sp*
Streptomyces griseocarneus				*Propionibacterium granulosum*
Synechococcus sp				
Xanthomonas axonopodis				
Xanthomonas oryzae				

2.3.3.4 Bacterial Diversity Confirmations using 16S rRNA Gene Libraries

To better constrain the source of DNA that had appeared in the pyrotag dataset, the subset of DNA extracts that were the basis of 454 Batch II were used to construct small 16S rRNA gene libraries. For this analysis, 12 library clones from each sample were unidirectionally sequenced and a phylogenic tree (Figure 14) produced. This analysis indicated that the two most abundant sequences obtained from across the entire pyrotag dataset (e.g., corresponding to *Brevundimonas diminuta* and *Burkholdaria* sp.) were also the most abundant clones in the gene libraries. Further, as revealed in the phylogenetic tree, the sequences from both analyses are near perfect matches to one another, indicating a very high probability that these sequences derive from the same or similar sources. Several other sequences that appeared in this library include a sequence identical to *Pseudomonas stutzeri* in extracts from 334i_Salina A1 and the Baked 494 Control. This strain was also present in the 454 libraries from most samples, including that of 334i-Salina A1. However, the strain was not present in the 494 Baked control. There were also some differences in results obtained by the two approaches, for example, *Acinetobacter* sp. B43 did not appear in any of the 454 libraries but did appear in the 661i_Cobourg library.

2.3.3.5 Sensitivity Analysis: Microbial Detection Limits and Controls

As spectrophotometric measurements of DNA extracts from across this project (not shown) indicated low to non-existant measurable DNA, a sensitivity calibration exercise was conducted to ascertain the lower detection threshold for bacterial DNA from these core materials under the specific conditions applied here. For this test, the appearance of a visible PCR product of the appropriate size on agarose gels was scored as detection. Figure 15 illustrates the results of this assay for the recovery of PCR-amplifiable DNA from known concentrations of added bacteria from the two major structural classes (e.g., *Bacillus* = Gram-postitive and *Proteus* = Gram-negative). This experiment utilised the same extraction chemistry as the experimental core samples (except that it employed a small kit format) applied to a baked rock control. Appropriate controls (e.g., postive and negative PCR reagent controls) were also included, and performed as expected. Overall, this experiment revealed that both Gram-positive and Gram-negative cells displayed similar detection thresholds under the conditions employed here, with the first traces of amplifiable DNA being recovered after the addition of ~ 4×10^5 to 4×10^6 cells per gram of rock (note that the gel image in Figure 15 fails to capture the faintest of bands that could be visualised). Sensitivity was similar regardless of whether 32 or 35 cycles of PCR were applied.

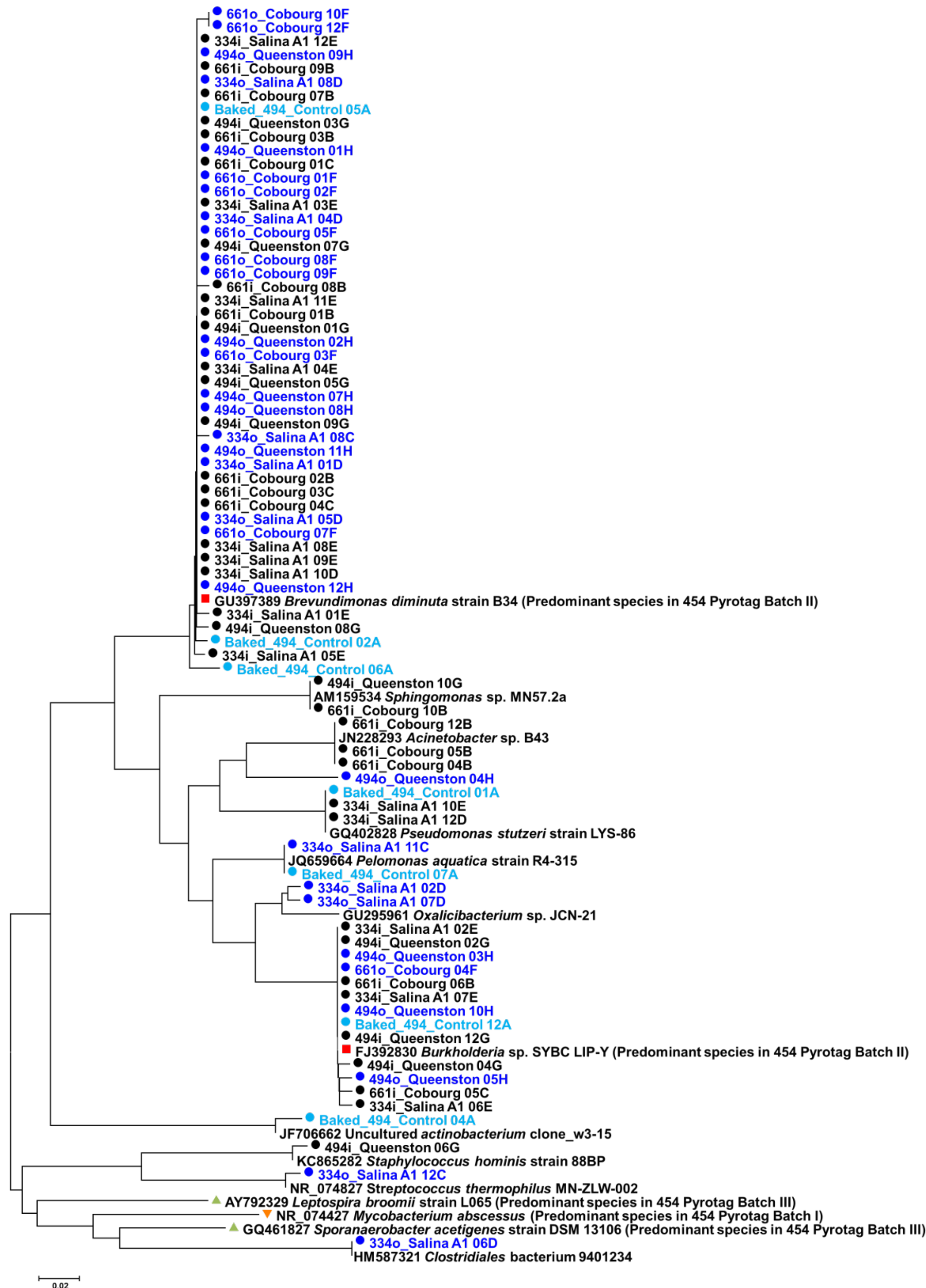

Figure 14: Neighbour-joining Phylogenetic Tree (1000 bootstraps). Scale bar = 2 Nucleotide Substitutions Per 100 Base Pairs

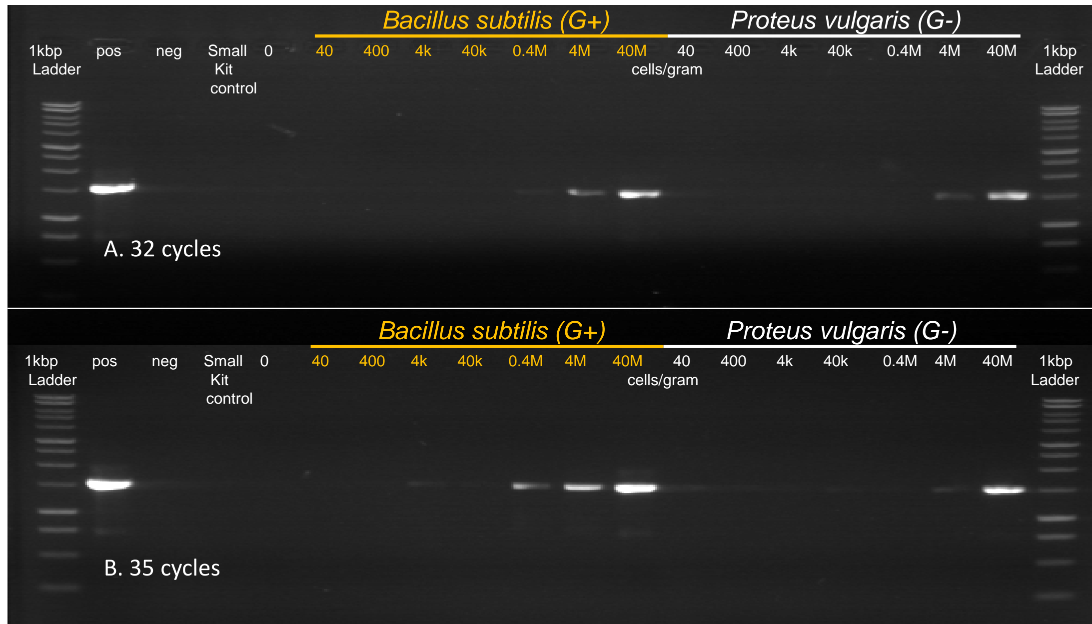

Figure 15: Sensitivity Calibration Experiment (Gel Electrophoresis) Showing Recovery of DNA From Known Numbers of Added Bacterial Cells From Baked Rock Controls). Limit of Detection ~ 4 x 10^5 Cells Per Gram.

3. DISCUSSION

3.1 PLFA DISCUSSION

3.1.1 Viable Microbial Communities

3.1.1.1 PLFA Abundance

PLFA unique to the rock core interiors were observed at very low abundances, indicating the presence of a very low biomass in these systems. Estimated cell densities of 1 – $3x10^5$ cells per gram or mL were an order of magnitude above the calculated limit of quantification of 3 $x10^4$ cells/gram and are comparable to some of the lowest estimated cell densities reported in the literature which have been observed in oligotrophic marine sediments (Parkes et al. 1995, Kallemeyer 2012). These cell densities are comparable to the estimates of 10^4 - 10^5 cells/gram observed by Stroes-Gascoyne et al. (2007) in the Opalinus clay. However, the observation of PLFA indicative of 10^5 cells/gram in the Cobourg limestone contrasts with the previous lack of detection of PLFA in this unit reported by Stroes-Gascoyne (Stroes-Gascoyne, 2008). Conversely, the observation of 10^5 cells/gram in the Queenston shale sample is lower than the previous observation of 10^6 cells/gram reported by Stroes-Gascoyne (Stroes-Gascoyne, 2008). It is worth noting that in both the previous study by Stroes-Gascoyne and the current study, the observed PLFA concentrations were extremely low. The differences between the two studies may relate to the volume of material extracted, which has implications both to the sensitivity of the analysis and also the influence of any potential heterogeneity associated with the samples coming from different boreholes. The exact mass of material extracted for PLFA analysis in the previous study was not reported, however it was indicated to be subset of the prepared material from one core, a significant amount of which was used for other studies. In the case of the current study, the entire core was allocated to PLFA analysis. By extracting circa 1 kg of powdered material in the current study, our limit of quantification using gas chromatography mass spectrometry was calculated to be 3 x 10^4 cells/gram based on the lowest concentration standard run. The extraction of such a large mass of material was undertaken to improve quantification limits, but it may also have avoided influence of heterogeneity within the samples related to the noted variations in sample colour and presence of fractures that were observed during splitting of the rock cores. It is very possible that the differences between the two studies related to heterogeneity within the units and the subsampled cores being from different boreholes.

PLFA abundances in the outer core rinses were of the same order of magnitude as the interior samples, with the exception of 464o_Queenston and 529o_GeorgianBay, which had abundances that were one and two orders of magnitude higher than their respective interior samples. The cause of these increased abundances is not clear but it may relate to contamination of the outer core surface post-sampling and is discussed further below.

3.1.1.2 PLFA Distribution

Interior Core samples:

The three most abundant PLFA in all interior core samples were C16:0, C18:0 and C18:1. These PLFA are ubiquitous and cannot be used to indicate either the identity or environmental

condition of the microbial communities. However, beyond these three PLFA, the observation of unique PLFA distribution in each of the rock core interiors, distinct from the patterns observed for the rinses, supports the presence of an indigenous microbial community. As noted in the results section, a number of the PLFA were branched and/or cyclopropyl. The former PLFA are associated with Gram-negative microorganisms and often are associated with sulphate-reducing microorganisms (Mauclaire et al., 2007; Stroes-Gascoyne et al., 2007). Increases in branched PLFA have also been previously observed within core interiors and interpreted as representing indigenous microbial communities (White and Ringelberg, 1997b). Cyclopropyl PLFA are considered to indicate cell stress (e.g., starvation; (White and Ringelberg, 1997b), which would be anticipated in these high salinity, oligotrophic, low water activity environments.

Outer rinse samples:

PLFA distributions in the outer core rinses were distinct from the corresponding interior samples, indicating that they were sampling a distinct source of the PLFA. It was noted that there were similarities amongst the outer core rinse distributions, particularly between 464o_Queenston and 529o_GeorgianBay, and again between 661o_Cobourg and 682o_Cobourg. This suggested that a similar source may have contributed to the outer surfaces of these samples.

The origin of the PLFA observed in the rinses is difficult to assign. Only one drilling water sample (DWR DGR8 – 014) was observed to contain PLFA, and this was at a very low concentration equivalent to $8x10^4$ cells per mL. All other drilling water samples were observed to have no quantifiable PLFA. The lack of PLFA in the drilling water is hypothesised to be due to the lysis of microbial cells present in the source water upon the addition of salt during drilling. This result is consistent with the failure to detect intact cells by flow cytometry, epifluorescence microscopy, phase contrast microscopy, and as viable cells with the ability to form colonies on R2A agar. This suggests that lipid contributions from the drilling water to the rock cores were likely negligible. The source of the observed PLFA in DWR DGR8 – 14 cannot be determined based on the limited PLFA profile. However, the PLFA present in this sample are the three primary PLFA observed in all samples, with the addition of C17:0. One possible explanation for the PLFA observed in DWR DGR8 – 14 is that these PLFA were, in fact, extracted from the rock during the drilling process; however, the corresponding core sample, 682_Cobourg, did not contain any C17:0 PLFA, suggesting this was not the case.

The estimated cell abundances for 464o_Queenston and 529o_GeorgianBay core samples were one and two orders of magnitude higher than the remainder of the samples, respectively. Furthermore, these rinses had the most diverse PLFA profiles, with many PLFA present at trace levels (refer to Figure 3). The high diversity observed in these two samples may be a result of the order of magnitude higher PLFA abundances resulting in the detection of PLFA that were below detection limits in the other samples. In this case, the similarity between the samples may be an artifact of the higher PLFA abundances. Alternatively, this higher diversity may reflect the microbial community present on the sample surfaces. The source of the elevated PLFA concentrations and elevated cell abundance estimates is not well constrained. It may indicate that the surfaces of these cores were contaminated during sample handling. The negligible PLFA present in the drilling water samples and lack of consistent observation of elevated PLFA on other core samples surfaces indicates that drilling water contamination is an unlikely explanation. However, post-drilling contamination is also possible and may be the explanation for PLFA detection in these samples. Regardless of the source of these elevated PLFA, the levels of PLFA interior to the cores are neither elevated compared to other samples,

nor do they show particularly high diversity, indicating that this surface effect did not impact the interior analyses and cell density estimates.

3.1.1.3 Hydrocarbon Results

The hydrocarbon analysis results support the conclusion that the interior samples were not impacted during drilling. The F1 fraction of the silica gel chromatography from the PLFA method contained hydrocarbons extracted at the same time as the PLFAs. Overall, a low abundance of hydrocarbons (alkanes, unresolved complex mixture (UCM)) was observed in the rock core rinse samples. In contrast, interior samples contained both alkane and UCM distributions, generally in much higher concentrations than the rinses. The presence of higher concentrations of distinct hydrocarbon patterns in the core interiors indicates that these compounds were indigenous. If they were contributed during drilling, they would have been expected to be higher in the rinses. It is more likely that the converse is true – hydrocarbon concentrations in the outer samples were reduced as a result of loss to the drilling water during sample collection. Based on literature reports (Sherman et al., 2007) and our observation, the outer surfaces of the cores are permeable to extracting solvents during rinsing to depth of circa 3 mm. This permeability might enable such extraction to occur.

Furthermore, the hydrocarbons observed in the samples can be related to the geologic units/formations sampled. There is a high degree of similarity between the lower three cores from the argillaceous limestone. These cores had relatively abundant alkanes and some UCM, showing some evidence of biodegradation over the lifetime of the organics. In the upper Salina argillaceous dolostone, only UCM was observed, indicating that a high degree of biodegradation had affected the hydrocarbons. The Queenston Formation had the lowest abundance of organics and the least evidence of biodegradation, followed by the Georgian Bay Formation. The unique patterns in hydrocarbon distribution between cores, and between interiors and exteriors, support the idea that these extractions were sampling indigenous materials when the PLFA were extracted.

The presence of these hydrocarbons is also important as they represent a potential food source for microbial communities in these low permeability sediments. Further, they represent one possible explanation for the low DNA yields as organic compounds can interfere with DNA replication.

3.1.1.4 Further Questions Related to PLFA Results

The very low cell abundance estimates reported herein are based on the assumption that PLFA hydrolyze within days-to-weeks after cell death. The hydrolysis of free esters in water is expected to be very rapid, potentially within hours (White et al., 1979; Stroes-Gascoyne et al., 2007). Hydrolysis rates may be longer in the case of intact cells, but studies have shown that degradation does occur within days-to-weeks in the presence of an active biological community, such as in marine sediments. However, abiotic controls have shown that phospholipid hydrolysis is much slower without biological activity (Harvey et al., 1986; Logemann et al., 2011). In the deep, saline subsurface systems being studied here, where biological activity is expected to be low or absent, the lifetime of PLFA may be longer than in the surface systems used to generate these estimates. Could the PLFA observed in these samples in fact be relict PLFA? The resolution of this question is critical within the context of the failure of the DNA

portion of this project to detect microbial life. If all or part of the PLFA detected were, in fact, of ancient origin, there would be no reason to expect DNA to be present.

Unfortunately, the degradation rates of PLFA under low water activity and/or in low biomass environments have yet to be reported. Thus, while it can be recognised that preservation of PLFA should be considered in this these low permeability systems, the time period over which this may occur is not well constrained.

Another potential source of the PLFA observed in this system may be microbial spores and or organisms in a viable but non-culturable (VBNC) state (Stroes-Gascoyne et al., 2007). There is the potential for cells in such states to yield PLFA upon extraction, but remain unculturable, and potentially to significantly reduce the success of DNA extraction. A potential means to assess the role of spores in these systems would be dipicnolyic acid analysis. Recently, increased sensitivity techniques have enabled assessment of spores in marine subsurface sediments (Lomstein et al., 2012). Application of similar approaches to terrestrial subsurface systems would elucidate the potential role of spores in both the PLFA yields and in the potential for microbial survival in far-field environments.

The lack of agreement between PLFA and DNA based assessment of bacteria in this study also raises the question as to whether or not the lack of detection of archea by DNA analysis indicates they are absent. Addition of archeal lipid analysis to future research would enable the same dual perspective (geochemical and DNA) based assessment to be made for archea as was made for bacteria in this study. Analysis and interpretation of the presence of archeal lipids will need to consider the increased lifetime expected for these compounds, as noted above.

3.2 DNA DISCUSSION

3.2.1 Context for the DNA Study

The "double-edged sword" of DNA-based life detection is the extreme sensitivity of the supporting technologies. This approach is reliable and reproducible when reasonable yields of high-quality DNA are available (normally the case with environmental samples). However, when very low yields of quality DNA are all that are available, trace contaminants, normally obscured, can be detected and misinterpreted as being derived from sample-associated microorganisms. Unfortunately, this is a common problem with rock samples from the deep subsurface , where mine or borehole contamination (Onstott et al., 2003; Moser, 2012) or reagent-associated contamination may represent the major source of DNA for analysis. When potentially spurious life is detected in sample extracts, a detailed analysis of sources of possible contamination is required. This process normally involves careful comparative analysis of patterns of detected microorganisms across sample sets which include controls such as drilling water and reagent blanks. A complementary approach usually involves an assessment of the suitability of detected lineages for the habitat being examined. For example, a high proportion of predicted aerobes from an anaerobic sample would represent evidence of probable contamination. As a result, DNA assessments from low-biomass or potentially sterile subsurface samples can ultimately reveal nothing but extensive details about reagent or drilling water contamination. This unavoidable scenario, although unsatisfying from a discovery science perspective, is essential to establish a limited-life outcome and appears to be the case with this dataset. Thus, the results obtained here, while very comprehensive, represent a detailed examination of commercial reagent contamination and detection-limit calibrations.

Measurements of environmental variables at the time of sampling are also included to enable an assessment of the compatibility of detected life for the habitat of these deep cores.

This work builds upon a sparse but important legacy of research projects, going back a number of decades, which have targeted life in the lithospheric inner space of Earth. In particular, this project targets the most challenging, but most directly applicable class of deep biosphere sampling – that of low porosity rock cores. Where more commonly sampled deep groundwater embodies a number of significant advantages as a vehicle for the detection of deep life, ultimately these fluids are sampled with the intention of learning more about life in the deep rocks themselves. In the case of these samples, because there is no groundwater associated with the formations being sampled, sampling of the rock is the only option for a study of deep life.

Biological sampling using deep cores is not a trivial endeavor. As cores can only be obtained via diamond drilling, bits must be cooled with liquid or air. Drilling fluids tend to contain a variety of potential contaminants (chemical and biological) that often compromise the integrity and ultimately the informative value of samples collected. Such contaminants (or added substances, such as NaCl in this case) can also serve as fortuitous tracers for assessing the quality of core subsamples. Based on the work of others (Story et al., 1995; Krumholz et al., 1997) and this group (Moser et al., 2003; Onstott et al., 2003; Davidson et al., 2011), a fairly solid understanding of how drilling-associated microbial contaminants become entrained into hard rock samples now exists. One common theme to all of these studies is that drilling fluids usually contain relatively high numbers of both bacteria (Onstott et al., 2003; Davidson et al., 2011) and archaea (Takai et al., 2001). In fact, in some cases, drilling fluid has been invoked as a mobile phase, enabling the sampling of the formation for indigenous microorganisms (Zhang et al., 2005; Zhang et al., 2006).

3.2.2 Direct Detection and DNA-Based Life Detection in Drilling Water

The results of this study appear to be very different from those of previous work in hard rock coring projects (Takai et al., 2001; Onstott et al., 2003; Moser, 2010; Davidson et al., 2011) in that both PLFA- and DNA-based analyses indicate an absence or, at best, very low numbers of microorganisms in any of the drilling water samples. This unusual conclusion was validated by the absence of detectable microorganisms via cultivation, flow cytometry, light microscopy or rRNA gene detection on agarose gels (refer to Figures 5 and 6). Although interpretations of the DNA dataset are complicated by the appearance of PCR products across the dataset (including negative controls), the absence of detectable microorganisms in the drilling water was further substantiated by the lack of distinction between any drilling water sample and its associated inner (or outer) corresponding core subsample (Figures 7, 8, 9, 10, 11, 12, 13 and Table 6). Other factors, especially the order in which samples were submitted to the DNA sequencing provider, appear to control patterns of diversity in this dataset. This result held true at both the phylum- (Figures 7 and 8) and species-levels (Figures 9 – 13) of diversity and for both bacterial (Figures 7 – 13) and archaeal datasets (Table A2.2).

A credible explanation for the apparent lack of life in the drilling water is obtained when one considers one of the best known principles of food preservation – high concentrations of salt are bacteriocidal (Wijnker et al., 2006; Shafiur, 2007). As in the case of the drilling water employed here, with salinities ranging from ~ 140 – 170 mS/cm (e.g., up to ~ 3X seawater), the salt concentrations were high. As the source of water for drilling was ultimately fresh surface water, it seems likely that native microorganisms exposed to such salinities would have been

physically destroyed by an extreme manifestation of the process of plasmolysis (e.g., the irreversible contraction of cellular membranes due to the loss of water by exosmosis (Korber et al., 1996; Random House Dictionary, 2013). This conclusion is supported by the fact that no culturable microorganisms were detected by traditional microbial plate count applied to all of the drilling fluid samples using a freshwater medium (R2A). Thus, the drilling water appears to have been functionally sterile; arguably the first example of "aseptic drilling" yet achieved during a subsurface microbiology project. One advantageous, although unintended, consequence of the saline drilling employed here should be that the normally-pervasive problem of drilling water contamination was probably negated and, thus, if any indigenous life was present in the formation, it would have been easy to detect.

3.2.3 Targeted Microbial Investigations

One target of this project was to see if evidence for indigenous microorganisms that produce methane could be found. As all known microbial methane producers are confined to the domain Archaea (e.g., Euryarchaeota, "Methanogens" (Luo et al., 2009)), our first priority was to determine if the expected taxa of methanogens were associated with inner core materials. As shown in Figure 5, archaeal PCR amplifications from across the sample set were unsuccessful in producing products of the expected size, even with high numbers of cycles applied (e.g., 35). Very low abundances of archaea in subsurface samples are not unusual (Moser et al., 2005; Gihring et al., 2006) and the application of nested PCR has, at times, enabled the detection of putative deep life (Takai et al., 2001). In this case, however, relying upon this adaptation was not warranted because, even with the incorporation of nested PCR, amplicons of the expected size were not produced from cores or drilling water. It should be noted that three sample extracts (464o_Queenston, 670i_Cobourg, and 670o_Cobourg) did produce bands, although they ran as much larger fragments on an agarose gel than expected. Normally, a result such as this indicates non-specific priming and can be discounted. However, it is conceivable that these large bands are real and actually represent versions of the 16S rRNA gene that contain introns (Takai and Horikoshi, 1999). The cloning and sequencing of these amplicons might represent a target for future work.

Although none of the samples produced archaeal amplicons with 454 pyrotag analysis, when an extremely degenerate forward primer, Arch349F (Takai and Horikoshi, 2000), was applied to a subset of samples, archaeal products were obtained (Table A2.2). None of the sequences detected, however, bore any affiliation with known methanogens. The majority of archaeal sequences detected (up to 97.5% of sequences detected) from all three samples were affiliated with an obligate symbiont of marine sponges, *Cenarchaeum* sp. The same three samples also had minor contributions of a marine Thaumarchaeote, *Nitrosopumilus* sp. (Konneke et al., 2005). Both of these microorganisms are obligate aerobes and most likely represent reagent contaminants. The expected physiology of these sequences, coupled with their low diversity and appearance in both drilling water and core samples, indicates that they are probably not defensible subsurface microorganisms.

Also notable for their absence in this dataset are predicted halophiles, many of which are also archaea (e.g., "*Halobacteria* spp." (Oren, 1999)). Because viable haloarchaea have been detected in subsurface halite deposits millions of years old (Vreeland, 1999; McGenity et al., 2000; Satterfield et al., 2005), this group should represent one of the more likely targets for indigenous microorganisms in these salty shale deposits. Our complete failure to detect these forms by cultivation and by molecular techniques, suggests that they are most likely not present, although more work would be required to completely rule out their presence.

Collectively, these independent lines of evidence all support the conclusion that neither extant methanogens, nor any other group of archaea, are likely to be present in these samples in any abundance.

A final group of possible interest to the programme are anaerobic respiratory microorganisms, such as sulphate- or iron-reducing bacteria (SRBs and IRBs). Of the various classes of microorganisms that are common in the subsurface, bacterial sulphate reducers and iron reducers could potentially adversely affect used fuel containers due to their production of corrosive H_2S, which can cause microbiologically influenced corrosion. It is noteworthy that not a single obvious SRB- or IRB-associated sequence was detected amongst the thousands of DNA sequences considered over the course of this study. However, it is never possible to prove function from phylogenetic information and at least one of the genera detected, *Bacillus*, does contain known iron reducers (Nealson and Saffarini, 1994), including strains from the deep subsurface (Boone et al., 1995; Kanso et al., 2002). Attempts to cultivate SRBs from a variety of drilling fluid and crushed core samples were unsuccessful. Collectively, these data appear to indicate that this sample set is most likely devoid of, or contains very low numbers of, methanogens, sulphate reducers, iron reducers and halophiles.

3.2.4 Indigenous vs. Contaminating Microorganisms: 454 Pyrotag Analysis

Because, almost by definition, the process of obtaining samples from the deep subsurface will contaminate them with surface-derived life signatures, one of the great challenges of subsurface microbial ecology research is knowing when indigenous life has been detected. One increasingly accepted means to circumvent this challenge has been to perform a subtractive analysis between known sources of contamination and samples expected to contain indigenous microorganisms. In cases where drilling/service water contains distinctive microbial populations, the discovery of unrelated populations in, for example, flowing boreholes, provides confidence that such sites represent “windows into deep microbial populations” (Baker et al., 2003; Moser et al., 2003). Because the best indications are that drilling fluid in this case lacks an intact bacterial population, this intrinsic tracer is not available for the differentiation of putative contaminant and indigenous bacterial populations. However, bacterial DNA was detected in 16S clone and pyrotag libraries for most of the samples, indicating a close examination of the patterns of diversity across this large sample set is required.

Patterns of bacterial diversity were examined by a number of complementary approaches. Routine PCR amplification (Figure 6) set the stage for all subsequent analyses. As with the archaeal portion of the study, 32 cycles of PCR, using proven primer sets for bacteria and a very efficient DNA polymerase (e.g., LA Taq, TaKaRa), failed to produce convincing products for most of the extracts tested. However, 35 PCR cycles did yield products of the correct size for all of the samples and three of the four negative controls (e.g., the DNA extraction blank and two baked rock controls). The most obvious interpretation, based on this result, would be that the products obtained were reagent-derived contamination rather than anything that could be traced back to the samples. The problem of trace contamination in laboratory and PCR reagents has been known for over a decade (Tanner et al., 1998), but normally is not insurmountable when reasonable yields of DNA are obtained. These results represented a crossroads for the project. Recognizing that the entire dataset may contain nothing but post-drilling reagent contamination, it was decided that the sequencing of the entire dataset was the only way to arrive at a defensible conclusion.

Figure 7 shows a simplified rank abundance overview of the patterns of diversity at the phylum-level obtained from the analysis of many thousands of DNA sequences. All subsequent DNA figures in this report serve to validate the unexpected conclusion supported by this figure. This conclusion being that, of the many factors with the potential to control these groupings (e.g., geologic formation, Mega vs. small DNA extraction kit, baked vs. non-sterilised nature of rock material), the sequencing batch into which a particular sample fell proved to be the primary determinant of community structure. There may be a number of ways to interpret this pattern, but the most obvious appears to be that, in the absence of significant quantities of DNA in our extracts, the sequencing laboratory detected their own reagent-associated or other in-house bacterial contamination. It was possible, however, that useful details were hidden within this high-level assessment of the data, and, thus, further analyses of the data were performed.

Cluster analyses at the phylum- and species-levels (Figures 8 and 9) produced tree topologies that exactly match the batch-defined patterns obtained from the rank-abundance curve. However, at the species level, the Batch II node clearly divided into two clades, one of which corresponded to samples (all drilling water) and a reagent control from the small kit variation of DNA extraction chemistry. This result suggests that buried within the sequencing laboratory's contamination signature might be some sample-specific information. This clade contains a mixture of inner, outer and drilling water samples from several core depths and there is no indication that this clade reveals anything other than an extraction artifact traceable back to the laboratory or a reagent used at DRI.

Deeper examinations of the species composition of the three sequencing batches are provided in Figures 10 and 11 and in Table 6. These figures examine the most abundant bacterium detected for each of the core and control samples. What is most remarkable about this dataset is the low diversity and relatively consistency of this most abundant bacterium across the dataset, again defined by sequencing batch instead of any other factor. The nature of the nearest neighbours to these abundant taxa clearly do not represent the ecophysiological classes of microorganisms that one would expect in a hypersaline, anaerobic, subsurface setting. Rather, the unmistakable dominance of ubiquitous, aerobic commensals and pathogens strongly supports origins in the laboratory or factory rather than the interiors of the deep cores.

3.2.4.1 454 Pyrotag: Abundant Bacterial Lineages

Brevundimonas diminuta-like sequences were the most common detected in our pyrotag libraries (Table 6). The strain detected in this project did not bear a close phylogenetic relationship to seqeunces in GenBank. Only two sequences in the database bore > 97% sequence identity to library clones (e.g., were of the same predicted species (Stackebrandt et al., 2002)); however, a very large number of environmental DNA sequences just outside this cutoff were present in the public database. The closest matches were from marine sediments associated with deep-sea hydrothermal vents (Wankel et al., 2012). However, matches nearly as close (e.g., at ~97% sequence identity) were derived from human skin (Grice et al., 2008), airborne wastewater plant emissions (Genbank accession #: JX855294.1), drinking water systems (Genbank accession #: JX085570), a parasitic wasp (Brucker and Bordenstein, 2012), zeolite-biological aerated filters from industrial wastewater (HQ682045.1), tropical soils (HQ445635.1), metal-contaminated estuarine sediment (HQ132463.1), and the roots of the tea plant (HM234006.1). *Brevundimonas diminuta* strains are also potential human pathogens, causing a variety of infections (Han and Andrade, 2005) and endocarditis (Yang et al., 2006). Formerly known as *Pseudomonas diminuta*, these strict aerobes are being increasingly

recognised as oppourtunistic and often antibiotic-resistant (Poole, 2004) human pathogens. Thus, this most abundant microbe detected over the course of the project does not appear to be a reasonable fit for the habitat being assessed (e.g., saline, devoid of eukarytotic host life and anaerobic) and should be considered a probable contaminant.

Second most abundant across the dataset is the *Burkholderia cepacia*-like sequences (Table 6), aligned with > 99% identity to a large group of abundant uncultivated environmental clones from a wide range of habitats including seawater (e.g., KC002545.1), soil (JQ806479.1) and the rhizosphere (JN975047.1), field corn (JX174211.1) ginger farm soil (KC172354.1), a spacecraft life support module (DQ447812.1) and *Trigoniulus corallines* (rusty millipede) gut (JN208912.1). Other *B. cepacia*-like sequences from this data are closely related to sequences detected in spacecraft cleanrooms (Moissl, 2007), the nector of wild plants (Alvarez-Perez, 2012), acid mine drainage (Petrie et al., 2003), and urban aerosols (Brodie et al., 2007). *B. cepacia* are a ubiquitous group of obligately aerobic bacteria composed of at least 17 different described species (e.g., the *Burkholderia cepacia complex* (BCC)) (Lipuma, 2005). *B. cepacia* is an important human pathogen which most often causes pneumonia in immunocompromised individuals with underlying lung disease (such as cystic fibrosis or chronic granulomatous disease) (Mahenthiralingam et al., 2005). It also attacks onion and tobacco plants, as well as displaying a remarkable ability to digest petroleum. The bacterium is well known as a very serious cause of hospital-acquired infections and is of concern medically due to its natural resistance to many common antibiotics (McGowan, 2006) and disinfectants (Anderson et al., 1990), representing a range of ecological roles not indicative of a subsurface lifestyle.

The third most abundant bacterium detected over the course of this work, *Mycobacterium bolletii* (also known as *M. abscessus* and *M. massiliiense* (Raiol et al., 2012); see Table 6), is part of a large complex of closely-related environmental clones from GenBank with > 99% sequence identity. This complex is also known as the *Mycobacterium chelonae-abscessus* complex (MCAC). Members are antibiotic-resistant (Adekambi et al., 2006) opportunistic pathogens in patients with underlying pulmonary disorders (Brown-Elliott and Wallace, 2002; Adekambi et al., 2006). The closest match is an unpublished cystic fibrosis pathogen (NR_074427.1). However, sequences nearly identical to those detected here have been detected in chlorinated water distribution systems (JX284540.1, JX084974.1, JN198972.1, JX284519.1), as fish pathogens (AY489137.1), and in constructed environments, such as hospital hot water systems, aerosols from showers, ice machines, swimming pools, dental unit water and medical devices (Thomas et al., 2006).

Finally, the appearance of *Sporanaerobacter acetigenes* as the most abundant sequence detected in several samples may be noteworthy. Unlike most bacteria detected over the course of this work, this strain is an obligate anaerobe (belonging to cluster XII of the *Clostridiales*). *S. acetigenes* is best known as an elemental sulfur reducer and amino acid fermenter and the only microorganism known to break down the explosive, RDX, when grown in pure culture (Hernandez-Eugenio et al., 2002). Of possible relevance, tolerating up to 5% NaCL, the strain is known as a moderate halophile. Unfortunately, lipid profiles (e.g., PLFA) of this microorganism have not been published and it is not possible to correlate the appearance of *S. acetigenes* with any of the PLFA detected during this project. Its prevalence in deep subsurface or hypersaline environments, however, remains to be determined.

Although the dominance of a small number of strains across the dataset is striking, it is possible that the appearance of these common forms merely reflects early PCR (Suzuki and Giovannoni, 1996) or template-to-product (Polz and Cavanaugh, 1998) bias in association with

low DNA concentrations. Thus, it is possible that sample-specific information might still be present as a component analogous to the so-called rare biosphere, a possibility the 454 pyrotag method employed here is well suited to reveal (Sogin et al., 2006). The differences in clustering between the principal coordinate analyses in Figures 12 and 13, for example, appear to indicate that patterns of bacterial diversity across the entire dataset are affected by the contributions of a small number of abundant taxa, and that when these abundant sequences are removed from the analysis, other factors appear to control clustering (refer to Figure 13).

3.2.4.2 454 Pyrotag: Rare bacterial Lineages

To eliminate the possibility that information concerning indigenous microorganisms might be hidden by PCR bias, the identities of the *least* abundant taxa in pyrotag libraries were examined (Table 7). Unlike the most abundant single taxon, which represented from ~ 25 – 88% (Table 6) of all bacteria detected in a given sample, the least abundant were truly miniscule components of the community (always a fraction of 1%, data not shown). The aim here was to identify lineages that would represent credible presences in the inner core samples, consistent with the deep or saline biosphere, or absent from the various borehole and reagent contaminant controls. The identities of these rare taxa form two major sample groupings containing controls, drilling water and inner and outer cores and are detailed in Tables 8 and 9. Overall, the dataset is remarkable for a general character that would be consistent with, for example, fresh or marine surface water (e.g., cyanobacteria like *Prochlorococcus* sp., *Anabaena* sp., *Synechococcus* sp. or *Phormidium* sp., and *Vibrio* sp.) or soil environments (e.g. *Rhizobia* sp., *Streptomyces* sp., and *Frankia* sp.). Other notable detections include a variety of human-associated lineages, including *Staphylococcus aureus*, *Enterococci*, and *E. coli* (outer sample), and *Enterococcus faecalis* (drilling water). Most importantly in this category, both inner samples ("small" and Mega Soil DNA Isolation Kit extractions) contain human-associated strains: the human pathogen *Massilia* sp. (La Scola et al., 1998), skin commensal/pathogen *Staphylococcus lugdunensis* (Frank et al., 2008) and fecal anaerobes *Clostridium orbiscindens* (Schoefer et al., 2003) and *Escherichia* sp. Other notable strains in the 334 inner samples include: *Burkholderia ambifaria,* a common plant commensal and human pathogen (Coenye et al., 2001); a heterocystous cyanobacterium, *Cyanospira rippkae* (Florenzano, 1983); and *Acinetobacter* spp., an omnipresent constituent from, amongst other habitats, drilling fluid samples from South African deep mine studies (Moser et al., 2003; Onstott et al., 2003). The presence of *Acidithiobacillus ferrooxidans* may, at first glance, suggest a subsurface microorganism, but, in fact, this microorganism is recognised as an obligatorily aerobic iron- and sulfur-oxidiser and a major source of acid mine drainage (Kelly and Wood, 2000).

As an independent verification for the results shown in Table 7, a second set of samples was examined for suggestions of indigenous rare biosphere constituents that otherwise might be obscured by a strong contaminant overprint (Supplemental DiscussionTable A2.3). This dataset overall resembled that of the 334 m depth samples in Table 7, being dominated by a range of common photosynthetic (e.g., *Cylindrospermum* sp., *Synechococcus* sp.), soil (e.g. *Bradyrhizobium* sp., *Mesorhizobium* sp.) and human-associated (e.g., *Enterobacter* sp., *Staphylococcus aureus*, *Mycobacterium chelonae* (Giulieri et al., 2012)) microorganisms. *Pseudomonas* strains were detected in all relevant samples, including the inner core sample. The inner sample possessed one pathogen, *Streptococcus mitis*, a prominent component of human oral flora and occasional cause of endocarditis (Lamas and Eykyn, 2003). As with the 334_Salina sample set, at least one sequence offers the suggestion of a rock-associated microbe in the inner sample, but the appearance here of *Ferrithrix thermotolerans* suggests the presence of another obligatorily-aerobic iron oxidiser associated with acid mine drainage

(Johnson et al., 2009). The presence of members of the *Dehalococcoides* sp. does suggest the possibility of anaerobic metabolism. This genus is generally known for coupling the reductive dehalogenation of chlorinated environmental contaminants (e.g., trichlorethane (Maymo-Gatell et al., 1997)), a specialty metabolism that would seem to be of little relevance in deep shale deposits. In the final analysis, neither the abundant nor rare representative microbial communities from these extracts by high throughout pyrotag analyses indicate anything other than common laboratory contamination.

3.2.5 16S rRNA bacterial Gene Libraries

A complementary assessment of patterns of bacterial diversity across this sample set is summarised by the phylogenetic tree shown in Figure 14. This dataset represents direct sequencing of 16S rRNA gene library clones obtained from DNA extracts that never left DRI; thus, it is impossible that contract laboratory contamination was the source of sequences detected in this analysis. Despite this constraint, the most abundant bacterial sequences detected by the cloned gene library approach were near-perfect DNA matches for the two most abundant sequences obtained from the 454 analysis (e.g., *Brevundimonas diminuta* and *Burkholdaria* sp., Table 6).

One of the few candidates identified by substractive analysis was *Acinetobacter* sp. B43, a strain that did not appear in any of the 454 libraries, but did appear in the 661i_Cobourg sample gene library. Relatives of this *Acinetobacter* strain are exceedingly common in environmental samples, producing hundreds of Genbank hits at > 99% identity from a wide array of habitats when queried by BLAST (Altschul et al., 1997). Illustrating this point, the first four BLAST hits with the best homology to the library clone from this dataset include sequences assiciated with human dermatitis (Kong et al., 2012), an Antartic air sample (JX559175.1), the estuarine bacterioneuston ((Azevedo et al., 2012) and household air (JN082488.1). Despite the fact that this sequence was only found in an inner core subsample, a case that the bacterium is anything but a ubiquitous environmental contaminant is difficult to make. Another sequence type, found in two inner core samples (494i_Queenston 10G and 661i_Cobourg 10B), was nearly identical to an commonly detected *Sphingomonas* strain associated with petroleum degradation in soil (e.g., JQ919161.1), the skin of freshwater fish (FR853730.1), marine mud volcanoes (FJ936934.1) and human skin (Grice et al., 2009); such pedigrees, once again, fail to support a subsurface origin.

The striking degree to which bacterial DNA detected in these two apparently contaminant-derived datasets resemble one another begs an explanation. Since both the 454 and DNA library approaches provided evidence of bacterial detection, even in the baked rock and reagent blank controls, the source of bacterial DNA in both must derive from some part of the DNA extraction or PCR amplification process. The fact that each of the three batches of DNA extracts submitted by DRI for 454 analysis distinctly group together supports the possibility that the contamination event occurred at some point after the samples had left DRI. This result suggests that the contamination was likely derived from the PCR amplification step at the sequencing provider. However, this interpretation is challenged by the fact that our independently-produced gene libraries detected very similar populations of bacterial DNA. This apparent inconsistency can be explained in one of several ways. Since 454 pyrotag uses a PCR-based methodology, it is possible, for example, that the PCR steps performed in both laboratories incorporated contaminating DNA and that these contaminant profiles just happened to be very similar to one another. The implication in this case is that PCR in general is prone to trace reagent contamination and that this contamination tends to be very similar from batch to

batch and manufacturer to manufacturer. A second possibility is that the contract lab was detecting trace contamination from our DNA extraction kit reagents, but for some reason, obtained a batch-specific bias (e.g. the results highlighted in Figures 7, 8, and 9). In the end, it is probably not possible to distinguish between these or some unrecognized alternative scenario for the origin of this contamination, but the result remains the same: a lack of evidence for indigenous life detection in any of these samples by DNA-based methods.

3.2.6 Sensitivity Analysis: Bacterial Detection Limits and Controls

The DNA dataset presented in this report, for the most part, represents a detailed analysis of what was *not* detected. However, the failure to detect something does not prove it is not there. For all analytical work there is a limit of detection below which something may be present, but not detected. Thus, a final and possibly more important analytical task was the determination of the limit of detection for the DNA-based assessments of these core materials. An initial assumption underlying this work was that DNA-based methodologies would enable the detection of a relatively small number of bacterial or archaeal cells. In the case of the DNA component of this study, a major conclusion was the inability to detect evidence for indigenous life, and, thus, the determination of method sensitivity was essential. A variety of workers, including members of this group, have reported the detection of microbial DNA from core samples in the past, although rarely has a limit of detection been established. As life is almost always detected in molecular DNA assessments of environmental samples, doped controls to determine minimum detection thresholds are generally not indicated. In this study, however, particularly because PLFA results appear to indicate the presence of life, a follow-up analysis to determine DNA detection limits became essential. Overall, the results of the follow-up experiment (Figure 15) revealed that both Gram-positive and Gram-negative bacterial cells displayed similar detection thresholds, with the first traces of amplifiable DNA being recovered after the addition of ~ $4 x 10^5$ to $4 x 10^6$ cells per gram of rock.

These values were unexpectedly high, but several aspects of the results, as illustrated in Figure 15, may shed some light upon the mechanisms supporting these observations. First, sensitivity of these tests was similar, regardless of whether 32 or 35 cycles of PCR were applied. This suggests that the recovery, as revealed in Figure 15, was determined by the availability of amplifiable DNA and not simply a PCR performance effect. Moreover, at least in a qualitative sense, it appears that the relative strength of PCR product formation increased very quickly once a threshold of detectability was reached. This result is most consistent with DNA binding sites on the crushed core surfaces reaching saturation at a certain cell number, above which DNA becomes readily extractable. If the core material was intrinsically hostile to DNA, it is likely that PCR detection would not have been possible, no matter how much was added. Such difficulties in the detection of microbial DNA in low-permeability rock samples have been noted before (e.g., (Stroes-Gascoyne et al., 2007), lending credence to these observations.

3.2.7 Further Questions Related to DNA Results

During the methods development phase, DNA recoveries were substantially improved by “fine-tuning” available technologies (e.g., through the use of high-phosphate extraction buffers to compete with DNA for access to charged mineral sites, and through the use of techniques to concentrate and purify DNA after extraction) and it is anticipated that additional methods development would further improve the limits of detection. Because DNA did not appear to be damaged by its interaction with crushed rock, methods for liberating DNA from charged binding

sites, such as the use of RNA blockers, may substantially improve recovery. Also, given the high organic content of these cores, it is also possible that co-purifed substances were also inhibitory to PCR. Thus, the application of surfactants, for example, to dispurse hydrophobic organics may enable the further purification of DNA. While it will never be possible to establish that a given sample does not contain some very small numbers of microorganisms, lower detection limits (by several orders of magnitude) would be ideal and the DNA dataset would be able to contribute to the resolution of questions concerning the origins of PLFA (e.g., extant microorganisms vs. preserved cellular components). Also, the availability of DNA detection protocols of higher sensitivity for cores would increase confidence in assessments of lithological units as being either sterile or hosting biologic communities.

4. CONCLUSIONS

4.1 PLFA CONCLUSIONS

The interiors of six rock core samples were found to contain PLFAs indicative of viable microbial communities living in the natural pore spaces of the rock cores. PLFA concentrations across the six cores were relatively similar, representing microbial cell densities in the range of $1x10^5$ to $3x10^5$ cells per gram of rock. These cell densities are comparable to the observations reported for the Opalinus clay by Stroes-Gascoyne (2007) and to the estimated cell density of 10^6 cells/gram observed for the Queenston Shale reported by Stroes-Gascoyne (Stroes-Gascoyne, 2008), though the observation of PLFA in the Cobourg Limestone contrasted the previous lack of detection of PLFA by Stroes-Gascoyne (Stroes-Gascoyne, 2008). PLFAs were detected on the surfaces of the rock cores as well, with similar PLFA compositions detailing the rock core interiors. 464o_Queenston and 529o_GeorgianBay were found to contain higher concentrations of PLFA, representing microbial cell densities of $4x10^6$ and $2x10^7$ cells per gram of rock, respectively. These values likely represent surface contamination, either via the drilling process or during sample handling. In any case, this evidence of external contamination does not appear to have affected the interior PLFA concentrations and supports the assumption that outer surface contaminants do not reach further than a few millimetres into the rock core interior. Low concentrations of PLFA in one drilling water sample, and a lack of PLFAs in the others, indicates that the drilling water is likely not a significant source of contamination to the rock cores under the conditions employed for this study. The presence of UCMs of alkanes in the rock core interiors suggests that microbial communities living in the microfractures and pore spaces of these rocks could potentially utilise these organic compounds as a food source.

For PLFA analysis, we recognise that the low biological activity expected for these environments might limit the rate of phospholipid hydrolysis. Currently the persistence of PLFA in low water activity, low biomass conditions is not well constrained. We propose that spores and/or VBNC cells may be contributing to PLFA detection and may represent a community of organisms that could become active in these systems if stimulated by disturbance. The testing of this hypothesis could involve assessment of indicators of bacterial spores (e.g., dipicolinic acid) and/or assessment of organic hydrocarbon fractions to determine if there is evidence of biological processing.

4.2 DNA CONCLUSIONS

Despite a great deal of invested effort, in the end, two complementary DNA-based technologies (PCR-enabled 16S rRNA gene library construction and 454 pyrotag analysis) failed to provide evidence for indigenous life in any core or drilling water sample. Rather, these methods, when applied to inner and outer core samples and relevant controls, produced a large number of low-diversity bacterial and archaeal sequences that confirmed the extreme sensitivity of these procedures and the almost certain origin of detected sequences as laboratory contaminants. These results were corroborated by the coincident failure of cultivation-, flow cytometric and microscopic analyses to detect microorganisms in crushed rock cores and drilling water. Calibration experiments utilizing known concentrations of added microorganisms, despite use of the best generally-accepted approaches, revealed a significant deficiency in the ability to recover bacterial DNA from core material. For this reason, the main conclusion of the DNA

study is that the cores tested contain less than about $4x10^5$ cells per gram of rock and that the sedimentary system under study is of low biomass.

Potential means to improve DNA recovery discussed in Section 3.2.7 would enable this hypothesis to be tested. However, the lack of agreement between the PLFA and DNA based assessment of bacteria also leads to the hypothesis that the lack of detection of archea by DNA based techniques is due to a similar issue with DNA based approaches, and that, in fact, archea may be present at low abundance in these formations. Future work would benefit by the addition of archeal ether lipid analysis in combination with optimised DNA based analysis focussed on overcoming the unknown inhibition that affected amplification in these samples in order to test this hypothesis.

4.3 SUMMARY

The results of this study demonstrate that, if present, microbial biomass is extremely low in these low-permeability materials. This low biomass, or lack thereof, is consistent with the low water activity and porosity/permeability of these cores. Unfortunately, the two primary approaches applied to these samples failed to corroborate one another, as the cell density estimates based on PLFA data of 1 to $3x10^5$ cells/gram are just below the detection limit determined for DNA analysis of $4x10^5$ cells/gram. DNA-based interpretations were further complicated by the detection of commercial reagent-associated contamination in most of the samples, which underscores the difficulty of extracting and analyzing microorganisms from low-biomass samples. Despite development of new intensive methods to analyze archea and bacteria in the rock cores, the challenges associated with the low biomass in these rock could not be overcome. As such, the DNA results can neither confirm nor refute the presence of microbial cells indicated by the PLFA results. Notably, a similar situation was reported by Stroes-Gascoyne et al. (2007) for the Opalinus clay, where PLFA analysis yielded concentrations greater than those observed in this study, and yet DNA could not be amplified from the samples. Stroes-Gascoyne et al. (2007) proposed that the highly sorptive nature of the clay may be related both to PLFA preservation and to DNA amplification inhibition. In the current study, it is likewise proposed that DNA amplification was inhibited by interference from either mineral or organic components within the host rock. Potential means to improve DNA recovery have been considered and would enable this hypothesis to be tested.

ACKNOWLEDGEMENTS

We gratefully acknowledge the key work done by the following people that contributed to this report. At McMaster University Danielle Simkus completed the PLFA extractions and analyses and was assisted by Jennie Kirby. At Desert Research Institute, Dr. Gaosen Zhang and Mr. Adrian Escobar completed the DNA extraction and analysis with assistance from Dr. Scott Hamilton-Brehm. At University of Toronto Katrina Chu and Georges Lacrampe-Coulome provided logistical support and assistance. We also acknowledge the thorough review and helpful comments from Dr. Gideon Wolfaardt (Ryerson University) and Prof. Darren Korber (University of Saskatchewan). Their efforts substantially improved our interpretation of this dataset.

REFERENCES

Adekambi, T., P. Berger, P, D. Raoult and M. Drancourt, 2006. rpoB gene sequence-based characterization of emerging non-tuberculous mycobacteria with descriptions of *Mycobacterium bolletii* sp. nov., *Mycobacterium phocaicum* sp. nov. and *Mycobacterium aubagnense* sp. nov. Int J Syst Evol Microbiol 56, 133-143.

Alaeddini, R., 2012. Forensic implications of PCR inhibition—A review. Forensic Science International: Genetics 6, 297-305.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J.H. Zhang, Z. Zhang, W. Miller and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.

Alvarez-Perez, S., C.M. Herrera and C. de Vega, 2012. Zooming-in on floral nectar: A first exploration of nectar-associated bacteria in wild plant communities. FEMS Microbiol Ecol 80, 591-602.

Amann, R.I., W. Ludwig and K.H. Schleife, 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59, 143-169.

Anderson, R.L., R.W. Vess, A.L. Panlilio and M.S. Favero, 1990. Prolonged survival of *Pseudomonas cepacia* in commercially manufactured povidone-iodine. Appl Environ Microbiol 56, 3598-3600.

Azevedo, J.S.N., I. Ramos, S. Araujo, C.S. Oliveira, A. Correia and I.S. Henriques, 2012. Spatial and temporal analysis of estuarine bacterioneuston and bacterioplankton using culture-dependent and culture-independent methodologies. Anton Leeuw Int J G 101, 819-835.

Baker, B.J., D.P. Moser, B.J. MacGregor, S. Fishbain, M. Wagner, N.K. Fry, B. Jackson, N. Speolstra, S. Loos, K. Takai, B. Sherwood Lollar, J. Fredrickson, D. Balkwill, T.C. Onstott, C.F. Wimpee and D.A. Stahl, 2003. Related assemblages of sulphate-reducing bacteria associated with ultradeep gold mines of South Africa and deep basalt aquifers of Washington State. Environmental Microbiology 5, 267-277.

Balkwill, D.L., F.R. Leach, J.T. Wilson, J.F. Mcnabb and D.C. White, 1988. Equivalence of Microbial Biomass Measures Based on Membrane Lipid and Cell-Wall Components, Adenosine-Triphosphate, and Direct Counts in Subsurface Aquifer Sediments. Microbial Ecol 16, 73-84.

Barns, S.M., C.F. Delwiche, J.D. Palmer, S.C. Dawson, K.L. Hershberger and N.R. Pace, 1996. Phylogenetic perspective on microbial life in hydrothermal ecosystems, past and present. In, Ciba Foundation Symposium 202-Evolution of Hydrothermal Ecosystems on Earth (And Mars?), 24-39.

Bessetti, J., 2007. An introduction to PCR inhibitors. J. Microbiol. Meth 28, 159-167.

Bligh, E.G. and W.J. Dyer, 1959. A rapid method of total lipid extraction and purification. Can. J. Biochem. Phys. 37, 911-917.

Boone, D.R., Y. Liu, Z.-J. Zhao, D.L. Balkwill, G.R. Drake, T.O. Stevens and H.C. Aldrich, 1995. *Bacillus infernus* sp. nov., an Fe (III)-and Mn (IV)-reducing anaerobe from the deep terrestrial subsurface. International journal of systematic bacteriology 45, 441-448.

Brodie, E.L., T.Z. DeSantis, J.P. Parker, I.X. Zubietta, Y.M. Piceno and G.L. Andersen, 2007. Urban aerosols harbor diverse and dynamic bacterial populations. Proc Natl Acad Sci U S A 104, 299-304.

Brown-Elliott, B.A. and R.J. Wallace Jr., 2002. Clinical and taxonomic status of pathogenic nonpigmented or late-pigmenting rapidly growing mycobacteria. Clinical microbiology reviews 15, 716-746.

Brucker, R.M and S.R. Bordenstein, 2012. The roles of host evolutionary relationships (genus: Nasonia) and development in structuring microbial communities. Evolution 66, 349-362.

Chapelle, F.H. and P.M. Bradley, 1996. Microbial acetogenesis as a source of organic acids in ancient Atlantic Coastal Plain sediments. Geology 24, 925-928.

Chapelle, F.H., K. O'Neill, P.M. Bradley, B.A. Methé, S.A. Ciufo, L.L. Knobel and D.R. Lovley, 2002. A hydrogen-based subsurface microbial community dominated by methanogens. Nature 415, 312-315.

Chivian, D., E.L. Brodie, E.J. Alm, D.E. Culley, P.S. Dehal, T.Z. Desantis, T.M. Gihring, A. Lapidus, L.H. Lin, S.R. Lowry, D.P. Moser, P.M. Richardson, G. Southam, G. Wanger, L.M. Pratt, G.L. Andersen, T.C. Hazen, F.J. Brockman, A.P. Arkin and T.C. Onstott, 2008. Environmental genomics reveals a single-species ecosystem deep within Earth. Science 322, 275-278.

Coenye, T., E. Mahenthiralingam, D. Henry, J.J. LiPuma, S. Laevens, M. Gillis, D.P. Speert and P. Vandamme, 2001. *Burkholderia ambifaria* sp. nov., a novel member of the *Burkholderia cepacia* complex including biocontrol and cystic fibrosis-related isolates. Int J Syst Evol Microbiol 51, 1481-1490.

Cole, J.R., Q. Wang, E. Cardenas, J. Fish, B. Chai, R.J. Farris, A.S. Kulam-Syed-Mohideen, D.M. McGarrell, T. Marsh, G.M. Garrity and J.M. Tiedje, 2009. The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37, D141-145.

Colwell, F.S. and S. D'Hondt, 2013. Nature and extent of the deep biosphere. Rev Mineral Geochem 75, 547-574.

Colwell, F.S., T.C. Onstott, M.E. Delwiche, D. Chandler, J.K. Fredrickson, Q.J. Yao, J.P. McKinley, D.R. Boone, R. Griffiths, T.J. Phelps, D. Ringelberg, D.C. White, L. LaFreniere, D. Balkwill, R.M. Lehman, J. Konisky and P.E. Long, 1997. Microorganisms from deep, high temperature sandstones: Constraints on microbial colonization. Fems Microbiol Rev 20, 425-435.

Davidson, M.M., B.J. Silver, T.C. Onstott, D.P. Moser, T.M. Gihring, L.M. Pratt, E.A. Boice, B. Sherwood Lollar, J. Lippmann-Pipke, S.M. Pfiffner, T.L. Kieft, W. Seymore and C. Ralston, 2011. Capture of Planktonic Microbial Diversity in Fractures by Long-Term Monitoring of Flowing Boreholes, Evander Basin, South Africa. Geomicrobiology Journal 28, 275-300.

DeFlaun, M.F., J.K. Fredrickson, H. Dong, S.M. Pfiffner, T.C. Onstott, D.L. Balkwill, S.H. Streger, E. Stackebrandt, S. Knoessen and E. van Heerden, 2007. Isolation and characterization of a *Geobacillus thermoleovorans* strain from an ultra-deep South African gold mine. Syst Appl Microbiol 30, 152-164.

DeSantis, T.Z., P. Hugenholtz, N. Larsen, M. Rojas, E.L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu and G.L. Andersen, 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72, 5069-5072.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460-2461.

Fang, J.S., S.T. Hasiotis, S. Das Gupta, S.S. Brake and D.A. Bazylinski, 2007. Microbial biomass and community structure of a stromatolite from an acid mine drainage system as determined by lipid analysis. Chem Geol 243, 191-204.

Florenzano, G., C.Sili, E. Pelosi and M. Vincenzini, 1983. *Cyanospira rippkae* and *Cyanospira capsulata* (gen. nov. and spp. nov.): New filamentous heterocystous cyanobacteria from Magadi lake (Kenya). Archives of Microbiology 140, 301-306.

Frank, K.L., J.L. Del Pozo and R. Patel, 2008. From clinical microbiology to infection pathogenesis: how daring to be different works for *Staphylococcus lugdunensis*. Clinical microbiology reviews 21, 111-133.

Fredrickson, J.K., J.P. McKinley, B.N. Bjornstad, P.E. Long, D.B. Ringelberg, D.C. White, L.R. Krumholz, J.M. Suflita, F.S. Colwell, R.M. Lehman, T.J. Phelps and T.C. Onstott, 1997. Pore-size constraints on the activity and survival of subsurface bacteria in a late Cretaceous shale-sandstone sequence, northwestern New Mexico. Geomicrobiology Journal 14, 183-202.

Fredrickson, J.K. and T.C. Onstott, 1996. Microbes deep inside the earth. Sci Am 275, 68-73.

Gihring, T.M., D.P. Moser, L.H. Lin, M. Davidson, T.C. Onstott, L. Morgan, M. Milleson, T.L. Kieft, E. Trimarco, D.L. Balkwill and M.E. Dollhopf, 2006. The distribution of microbial taxa in the subsurface water of the Kalahari Shield, South Africa. Geomicrobiology Journal 23, 415-430.

Giovannoni, S.J., T.B. Britschgi, C.L. Moyer and K.G. Field, 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature 345, 60-63.

Giulieri, S., M. Cavassini and K. Jaton, 2012. *Mycobacterium chelonae* illnesses associated with tattoo ink. The New England journal of medicine 367, 2357; author reply 2357-2358.

Green, C.T. and K.M. Scow, 2000. Analysis of phospholipid fatty acids (PLFA) to characterize microbial communities in aquifers. Hydrogeol J 8, 126-141.

Grice, E.A., H.H. Kong, S. Conlan, C.B. Deming, J. Davis, A.C. Young, G.G. Bouffard, R.W. Blakesley, P.R. Murray, E.D. Green, M.L. Turner, J.A. Segre and N.C.S. Progra, 2009. Topographical and Temporal Diversity of the Human Skin Microbiome. Science 324, 1190-1192.

Grice, E.A., H.H. Kong, G. Renaud, A.C. Young, G.G. Bouffard, R.W. Blakesley, T.G. Wolfsberg, M.L. Turner and J.A. Segre, 2008. A diversity profile of the human skin microbiota. Genome research 18, 1043-1050.

Guckert, J., M. Hood and D. White, 1986. Phospholipid ester-linked fatty acid profile changes during nutrient deprivation of Vibrio cholerae: Increases in the trans/cis ratio and proportions of cyclopropyl fatty acids. Applied and Environmental Microbiology 52, 794-801.

Guckert, J.B., C.P. Antworth, P.D. Nichols and D.C. White, 1985. Phospholipid, ester-linked fatty acid profiles as reproducible assay for change in prokaryotic community structure of estuarine sediments. FEMS Microbiology Ecology 31, 147-158.

Haldeman, D.L. and P.S. Amy, 1993. Bacterial Heterogeneity in Deep Subsurface Tunnels at Rainier Mesa, Nevada Test Site. Microbial Ecol 25, 183-194.

Hallmann, C., L. Schwark and K. Grice, 2008. Community dynamics of anaerobic bacteria in deep petroleum reservoirs. Nature Geoscience 1, 588-591.

Hammer, Ø., D.A.T. Harper and P.D. Ryan, 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4, 9.

Han, X.Y. and R.A. Andrade, 2005. *Brevundimonas diminuta* infections and its resistance to fluoroquinolones. The Journal of antimicrobial chemotherapy 55, 853-859.

Harvey, H.R., R.D. Fallom and J.S. Patton, 1986. The effect of organic matter and oxygen on the degradation of bacterial membrane lipids in marine sediments. Geochimica and Cosmochimica Acta 50, 795-805.

Hernandez-Eugenio, G., M.L. Fardeau, J.L. Cayol, B.K. Patel, P. Thomas, H. Macarie, J.L. Garcia and B. Ollivier, 2002. *Sporanaerobacter acetigenes* gen. nov., sp. nov., a novel acetogenic, facultatively sulfur-reducing bacterium. Int J Syst Evol Microbiol 52, 1217-1223.

Hobbie, J.E., R.J. Daley and S. Jasper, 1977. Use of nuclepore filters for counting bacteria by fluorescence microscopy. Appl Environ Microbiol 33, 1225-1228.

Johnson, D.B., P. Bacelar-Nicolau, N. Okibe, A. Thomas and K.B. Hallberg, 2009. *Ferrimicrobium acidiphilum* gen. nov., sp. nov. and *Ferrithrix thermotolerans* gen. nov., sp. nov.: heterotrophic, iron-oxidizing, extremely acidophilic actinobacteria. Int J Syst Evol Microbiol 59, 1082-1089.

Kallmeyer, J., K. Mangelsdorf, B. Cragg and B. Horsfield, 2006. Techniques for contamination assessment during drilling for terrestrial subsurface sediments. Geomicrobiology Journal 23, 227-239.

Kallmeyer, J., R. Pockalny, R.R. Adhikari, D.C. Smith and S. D'Hondt, 2012. Global distribution of microbial abundance and biomass in subseafloor sediment. P Natl Acad Sci USA 109, 16213-16216.

Kanso, S., A.C. Greene and B.K. Patel, 2002. *Bacillus subterraneus* sp. nov., an iron-and manganese-reducing bacterium from a deep subsurface Australian thermal aquifer. International journal of systematic and evolutionary microbiology 52, 869-874.

Kelly, D.P. and A.P. Wood, 2000. Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. Int J Syst Evol Microbiol 50 Pt 2, 511-516.

Kieft, T., D. Ringelberg and D. White, 1994. Changes in ester-linked phospholipid fatty acid profiles of subsurface bacteria during starvation and desiccation in a porous medium. Applied and Environmental Microbiology 60, 3292-3299.

Kieft, T.L., J.K. Fredrickson, T.C. Onstott, Y.A. Gorby, H.M. Kostandarithes, T.J. Bailey, D.W. Kennedy, S.W. Li, A.E. Plymale, C.M. Spadoni and M.S. Gray, 1999. Dissimilatory reduction of Fe(III) and other electron acceptors by a Thermus isolate. Appl Environ Microbiol 65, 1214-1221.

Kong, H.H., J. Oh, C. Deming, S. Conlan, E.A. Grice, M.A. Beatson, E. Nomicos, E.C. Polley, H.D. Komarow, P.R. Murray, M.L. Turner and J.A. Segre, 2012. Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. Genome research 22, 850-859.

Konneke, M., A.E. Bernhard, J.R. de la Torre, C.B. Walker, J.B. Waterbury and D.A. Stahl, 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature 437, 543-546.

Korber, D.R., A. Choi, G.M. Wolfaardt and D.E. Caldwell, 1996. Bacterial plasmolysis as a physical indicator of viability. Appl Environ Microbiol 62, 3939-3947.

Kotelnikova, S. and K. Pedersen, 1998. Distribution and activity of methanogens and homoacetogens in deep granitic aquifers at Aspo hard rock laboratory, Sweden. FEMS Microbiol. Ecol. 26, 121–134.

Krumholz, L.R., J.P. McKinley, F.A. Ulrich and J.M. Suflita, 1997. Confined subsurface microbial communities in Cretaceous rock. Nature 386, 64-66.

Kunin, V., A. Engelbrektson, H. Ochman and P. Hugenholtz, 2010. Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environmental Microbiology 12, 118-123.

La Scola, B., R.J. Birtles, M.N. Mallet and D. Raoult, 1998. *Massilia timonae* gen. nov., sp. nov., isolated from blood of an immunocompromised patient with cerebellar lesions. Journal of Clinical Microbiology 36, 2847-2852.

Lamas, C.C. and S.J. Eykyn, 2003. Blood culture negative endocarditis: Analysis of 63 cases presenting over 25 years. Heart 89, 258-262.

Lin, L.H., P.L. Wang, D. Rumble, J. Lippmann-Pipke, E. Boice, L.M. Pratt, B. Sherwood Lollar, E.L. Brodie, T.C. Hazen, G.L. Andersen, T.Z. DeSantis, D.P. Moser, D. Kershaw and T.C. Onstott, 2006. Long-term sustainability of a high-energy, low-diversity crustal biome. Science 314, 479-482.

Lipuma, J.J., 2005. Update on the *Burkholderia cepacia* complex. Current opinion in pulmonary medicine 11, 528-533.

Liu, Y., T.M. Karnauchow, K.E. Jarrell, D.L. Balkwill, G.R. Drake, D. Ringelberg, R. Clarno and D.R. Boone, 1997. Description of two new thermophilic species of *Desulfotomaculum*, *Desulfotomaculum putei* sp. nov. from the deep terrestrial subsurface and *Desulfotomaculum luciae* sp. nov. from a hot spring. Int. J. Syst. Bacteriol. 47, 615–621.

Logemann, J., J. Graue, J. Koster, B. Engelen, J. Rullkotter and H. Cypionka, 2011. A laboratory experiment of intact polar lipid degradation in sandy sediments. Biogeosciences 8, 2547-2560.

Lomstein, B.A., A.T. Langerhuus, S. D'Hondt, B.B. Jorgensen and A.J. Spivack, 2012. Endospore abundance, microbial growth and necromass turnover in deep sub-seafloor sediment. Nature 484, 101-104.

Luo, H.W., Z.Y. Sun, W. Arndt, J. Shi, R. Friedman and J.J. Tang, 2009. Gene Order Phylogeny and the Evolution of Methanogens. Plos One 4.

Mahenthiralingam, E., T.A. Urban and J.B. Goldberg, 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. Nature reviews. Microbiology 3, 144-156.

Mauclaire, L., J.A. McKenzie, B. Schwyn and P. Bossart, 2007. Detection and cultivation of indigenous microorganisms in Mesozoic claystone core samples from the Opalinus Clay Formation (Mont Terri Rock Laboratory). Physics and Chemistry of the Earth 32, 232-240.

Maymo-Gatell, X., Y. Chien, J.M. Gossett and S.H. Zinder, 1997. Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene. Science 276, 1568-1571.

McGenity, T.J., R.T. Gemmell, W.D. Grant and H. Stan-Lotter, 2000. Origins of halophilic microorganisms in ancient salt deposits. Environmental Microbiology 2, 243-250.

McGowan, J.E., Jr., 2006. Resistance in nonfermenting gram-negative bacteria: multidrug resistance to the maximum. American journal of infection control 34, S29-37; discussion S64-73.

Moissl, C., S. Osman, M.T. La Duc, A. Dekas, E. Brodie, T. DeSantis and K. Venkateswaran, 2007. Molecular bacterial community analysis of clean rooms where spacecraft are assembled. FEMS Microbiol. Ecol. 61, 509-521.

Moser, D.P., B. Sherwood Lollar, G. Slater, T.C. Onstott, J.C. Bruckner, J.C. Fisher and J. Reihle, 2012. Finding the Biotic Fringe in the Continental Deep Subsurface. General Meeting American Society for Microbiology. San Francisco, CA.

Moser, D.P., T.M. Gihring, F.J. Brockman, J.K. Fredrickson, D.L. Balkwill, M.E. Dollhopf, B. Sherwood Lollar, L.M. Pratt, E. Boice, G. Southam, G. Wanger, B.J. Baker, S.M. Pfiffner, L.H. Lin and T.C. Onstott, 2005. *Desulfotomaculum* and *Methanobacterium* spp. dominate a 4-to 5-kilometer-deep fault. Applied and Environmental Microbiology 71, 8773-8783.

Moser, D.P., T.C. Onstott, J.K. Fredrickson, F.J. Brockman, D.L. Balkwill, G.R. Drake, S.M. Pfiffner, D.C. White, K. Takai, L.M. Pratt, J. Fong, B. Sherwood Lollar, G. Slater, T.J. Phelps, N. Spoelstra, M. Deflaun, G. Southam, A.T. Welty, B.J. Baker and J. Hoek, 2003. Temporal shifts in the geochemistry and microbial community structure of an ultradeep mine borehole following isolation. Geomicrobiology Journal 20, 517-548.

Moser, D.P., S. Bang, T.L. Jones, D. Boutt, T. Kieft, B. Sherwood Lollar, L.C. Murdoch, S.M. Pfiffner, J. Bruckner, J.C. Fisher, J. Newburn, A. Wheatley and T.C. Onstott, 2010. First Microbial Community Assessment of Borehole Fluids from the Deep Underground Science and Engineering Laboratory (DUSEL). American Geophysical Union, Fall Meeting 2010, Abstract #H13F-1049.

Nealson, K.H. and D. Saffarini, 1994. Iron and manganese in anaerobic respiration: environmental significance, physiology, and regulation. Annual Reviews in Microbiology 48, 311-343.

Nealson, K.H. and J.C. Venter, 2007. Metagenomics and the global ocean survey: What's in it for us, and why should we care? The ISME Journal 1, 185-187.

Onstott, T., D.P. Moser, S.M. Pfiffner, J.K. Fredrickson, F.J. Brockman, T. Phelps, D. White, A. Peacock, D. Balkwill and R. Hoover, 2003. Indigenous and contaminant microbes in ultradeep mines. Environmental Microbiology 5, 1168-1191.

Onstott, T.C., T.J. Phelps, F.S. Colwell, D. Ringelberg, D.C. White, D.R. Boone, J.P. McKinley, T.O. Stevens, D.L. Balkwill, T. Griffin and T. Kieft, 1998. Observations pertaining to the origin and ecology of microorganisms recovered from the deep subsurface of Taylorsville Basin, Virginia. Geomicrobiol. J. 15, 353–385.

Oren, A., 1999. Bioenergetic aspects of halophilism. Microbiol Mol Biol R 63, 334-348.

Pace, N.R., D.A. Stahl, D.J. Lane and G.J. Olsen, 1985. Analyzing natural microbial populations by rRNA sequences ASM News 51, 4-12.

Parameswaran, P., R. Jalili, L. Tao, S. Shokralla, B. Gharizadeh, M. Ronaghi and A.Z. Fire, 2007. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. Nucleic Acids Res 35, e130.

Parkes, R.J., B.A. Cragg, S.J. Bale,J.M. Getliff, K. Goodman, P.A. Rochelle, J.C. Fry, A.J. Weightman and S.M. Harvey, 1994. Deep bacterial biosphere in Pacific-Ocean sediments. Nature 371, 410-413.

Parkes, R.J., B.A. Cragg and P. Wellsbury, 2000. Recent studies on bacterial populations and processes in subseafloor sediments: A review. Hydrogeol J 8, 11-28.

Pedersen, K., 1993. The deep subterranean biosphere. Earth-Sci Rev 34, 243-260.

Pedersen, K., 1996. Investigations of subterranean bacteria in deep crystalline bedrock and their importance for the disposal of nuclear waste. Canadian Journal of Microbiology 42, 382-391.

Pedersen, K., 2000. Exploration of deep intraterrestrial microbial life: Current perspectives. Fems Microbiol Lett 185, 9-16.

Pedersen, K., 2001. Diversity and activity of microorganisms in deep igneous rock aquifers of the Fennoscandian shield. John Wiley New York.

Petrie, L., N.N. North, S.L. Dollhopf, D.L. Balkwill and J.E. Kostka, 2003. Enumeration and characterization of iron(III)-reducing microbial communities from acidic subsurface sediments contaminated with uranium(VI). Appl Environ Microbiol 69, 7467-7479.

Polz, M.F. and C.M. Cavanaugh, 1998. Bias in template-to-product ratios in multitemplate PCR. Appl Environ Microbiol 64, 3724-3730.

Poole, K., 2004. Efflux-mediated multiresistance in Gram-negative bacteria. Clinical microbiology and infection: The official publication of the European Society of Clinical Microbiology and Infectious Diseases 10, 12-26.

Preston, C.M., K.Y. Wu, T.F. Molinski and E.F. DeLong, 1996. A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. Proc Natl Acad Sci USA 93, 6241-6246.

Raiol, T., G.M. Ribeiro, A.Q. Maranhao, A.L. Bocca, I. Silva-Pereira, A.P. Junqueira-Kipnis, M. Brigido Mde and A. Kipnis, 2012. Complete genome sequence of *Mycobacterium massiliense*. J Bacteriol 194, 5455.

Random House Dictionary, R.H., Inc., 2013.

Rappe, M.S. and S.J. Giovannoni, 2003. The uncultured microbial majority. Annual Review of Microbiology 57, 369-394.

Reasoner, D.J. and E.E. Geldreich, 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl. Environ. Microbiol. 49, 1-7.

Rice, E.W., R.B. Baird, L.S. Clesceri and A.D. Eaton (Ed.), 2012. Standard Methods For The Examination Of Water and Wastewater. American Public Health Association 800 I Street, NW, Washington, DC 20001.

Ringleberg, D.B., S. Sutton and D.C. White, 1997. Biomass, bioactivity and biodiveristy: Microbial ecology of the deep subsurface: analysi of ester-linked phospholipid fatty acids. FEMS Microbiology Ecology 20, 371-377.

Rusch, D.B., A.L. Halpern, G. Sutton, K.B. Heidelberg, S. Williamson, S. Yooseph, D. Wu, J.A. Eisen, J.M. Hoffman, K. Remington, K. Beeson, B. Tran, H. Smith, H. Baden-Tillson, C. Stewart, J. Thorpe, J. Freeman, C. Andrews-Pfannkoch, J.E. Venter, K. Li, S. Kravitz, J.F. Heidelberg, T. Utterback, Y.H. Rogers, L.I. Falcon, V. Souza, G. Bonilla-Rosso, L.E. Eguiarte, D.M. Karl, S. Sathyendranath, T. Platt, E. Bermingham, V. Gallardo, G. Tamayo-Castillo, M.R. Ferrari, R.L. Strausberg, K. Nealson, R. Friedman, M. Frazier and J.C. Venter, 2007. The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through eastern tropical Pacific. PLoS Biology 5, e77.

Sabat, G., P. Rose, W.J. Hickey and J.M. Harkin, 2000. Selective and sensitive method for PCR amplification of Escherichia coli 16S rRNA genes in soil. Appl Environ Microbiol 66, 844-849.

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis and H.A. Erlich, 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239, 487-491.

Satterfield, C.L., T.K. Lowenstein, R.H. Vreeland, W.D. Rosenzweig and D.W. Powers, 2005. New evidence for 250 Ma age of halotolerant bacterium from a Permian salt crystal. Geology 33, 265-268.

Schoefer, L., R. Mohan, A. Schwiertz, A. Braune and M. Blaut, 2003. Anaerobic degradation of flavonoids by Clostridium orbiscindens. Appl Environ Microbiol 69, 5849-5854.

Schonknecht, G., W.H. Chen, C.M. Ternes, G.G. Barbier, R.P. Shrestha, M. Stanke, A. Brautigam, B.J. Baker, J.F. Banfield, R.M. Garavito, K. Carr, C. Wilkerson, S.A. Rensing, D. Gagneul, N.E. Dickenson, C. Oesterhelt, M.J. Lercher and A.P. Weber, 2013. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339, 1207-1210.

Shafiur, R., 2007. Handbook of Food Preservation. CRC Press New York.

Sherman, L.S., J.R. Waldbauer and R.E. Summons, 2007. Improved methods for isolating and validating indigenous biomarkers in Precambrian rocks. Org Geochem 38, 1987-2000.

Sherwood Lollar, B., 2011. Far-field Microbiology Considerations Relevant to a Deep Geological Repository – State of Science Review. Nuclear Waste Management Organization. NWMO TR-2011-09.

Sherwood Lollar, B., G. Lacrampe-Couloume, G.F. Slater, J. Ward, D.P. Moser, T.M. Gihring, L.-H. Lin and T.C. Onstott, 2006. Unravelling abiogenic and biogenic sources of methane in the Earth's deep subsurface. Chem Geol 226, 328-339.

Sogin, M.L., H.G. Morrison, J.A. Huber, D.M. Welch, S.M. Huse, P.R. Neal, J.M. Arrieta and G.J. Herndl, 2006. Microbial diversity in the deep sea and the underexplored “rare biosphere”. Proceedings of the National Academy of Sciences 103, 12115-12120.

Stackebrandt, E., W. Frederiksen, G.M. Garrity, P.A. Grimont, P. Kampfer, M.C. Maiden,X. Nesme, R. Rossello-Mora, J. Swings, H.G. Truper, L. Vauterin, A.C. Ward and W.B. Whitman, 2002. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 52, 1043-1047.

Staley, J.T. and A. Konopka, 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu Rev Microbiol 39, 321-346.

Story, S.P., P.S. Amy, C.W. Bishop and F.S. Colwell, 1995. Bacterial transport in volcanic tuff cores under saturated flow conditions. Geomicrobiology Journal 13, 249-264.

Stroes-Gascoyne, S. and C.J. Hamon, 2008. Preliminary microbial analysis of limestone and shale rock samples. Nuclear Waste Management Organization. NWMO TR-2008-09.

Stroes-Gascoyne, S., A. Schippers, B. Schwyn, S. Poulain, C. Sergeant, M. Simonoff, C. Le Marrec, S. Altmann, T. Nagaoka, L. Mauclaire, J. McKenzie, S. Daumas, A. Vinsot, C. Beaucaire and J.M. Matray, 2007. Microbial community analysis of Opalinus Clay drill core samples from the Mont Terri Underground Research Laboratory, Switzerland. Geomicrobiology Journal 24, 1-17.

Suzuki, M.T. and S.J. Giovannoni, 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl Environ Microbiol 62, 625-630.

Takai, K. and K. Horikoshi, 1999. Molecular phylogenetic analysis of archaeal intron-containing genes coding for rRNA obtained from a deep-subsurface geothermal water pool. Appl Environ Microbiol 65, 5586-5589.

Takai, K. and K. Horikoshi, 2000. Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. Appl Environ Microbiol 66, 5066-+.

Takai, K., D.P. Moser, M. DeFlaun, T.C. Onstott and J.K. Fredrickson, 2001. Archaeal diversity in waters from deep South African gold mines. Applied and Environmental Microbiology 67, 5750-5760.

Tamura K, J. Dudley and M.S.K. Nei, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.

Tanner, M.A., B.M. Goebel, M.A. Dojka and N.R. Pace, 1998. Specific ribosomal DNA sequences from diverse environmental settings correlate with experimental contaminants. Appl Environ Microbiol 64, 3110-3113.

Thomas, V., K. Herrera-Rimann, D.S. Blanc and G. Greub, 2006. Biodiversity of amoebae and amoeba-resisting bacteria in a hospital water network. Appl Environ Microbiol 72, 2428-2438.

Tringe, S.G. and P. Hugenholtz, 2008. A renaissance for the pioneering 16S rRNA gene. Curr Opin Microbiol 11, 442-446.

USEPA, 1979. Methods for chemical analysis of water and wastes. In: Environmental Monitoring and Support Laboratory, Cincinnati, Ohio, USA.

USEPA, 1993. Methods for the determination of inorganic substances in environmental samples. In: United States Environmental Protection Agency, Washington DC, USA.

Vreeland, R.H. and D.W. Powers (Ed.), 1999. Considerations for microbiological sampling of crystals from ancient salt formations, CRC Press, Boca Raton.

Wankel, S.D., M.M. Adams, D.T. Johnston, C.M. Hansel, S.B. Joye and P.R. Girguis, 2012. Anaerobic methane oxidation in metalliferous hydrothermal sediments: Influence on carbon flux and decoupling from sulfate reduction. Environmental Microbiology 14, 2726-2740.

Weiss, J.V. and I.M. Cozzarelli, 2008. Biodegradation in contaminated aquifers: Incorporating microbial/molecular methods. Ground Water 46, 305-322.

White, D. and D. Ringelberg, 1997a. Utility of the signature lipid biomarker analysis in determining the in situ viable biomass, community structure, and nutritional/physiologic status of deep subsurface microbiota. The Microbiology of the terrestrial deep subsurface, 119-136.

White, D.C., W.M. Davis, J.S. Nickels, J.D. King and R.J. Bobbie, 1979. Determination of the sedimentary microbial biomass by extractible lipid phosphate. Oceologica 40, 51-62.

White, D.C. and D. Ringelberg, 1997b. Utility of the Signature Lipid Biomarker Analysis in Determingint the In Situ Viable Biomass, Community Structure, and Nutritional/Physiological Status of Deep Subsurface Microbiota. In: Amy, P.S., Haldeman, D.L. (Eds.), The Microbiology of the Terrestrial Deep Subsurface. Lewis Publishers, Boca Raton.

Whitman, W.B., 1998. Procaryotes: The unseen majority. Proc Natl Acad Sci USA 95, 6578-6583.

Wijnker, J.J., G. Koop and L.J.A. Lipman, 2006. Antimicrobial properties of salt (NaCl) used for the preservation of natural casings. Food Microbiol 23, 657-662.

Woese, C.R. and G.E. Fox, 1977. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc Natl Acad Sci USA 74, 5088-5090.

Woese, C.R., O. Kandler and M.L. Wheelis, 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87, 4576-4579.

Yang, M.L., Y.H. Chen, T.C. Chen, W.R. Lin, C.Y. Lin and P.L. Lu, 2006. Case report: Infective endocarditis caused by *Brevundimonas vesicularis*. BMC infectious diseases 6, 179.

Yooseph, S., Nealson, K.H., Rusch, D.B., McCrow, J.P., Dupont, C.L., Kim, M., Johnson, J., Montgomery, R., Ferriera, S., Beeson, K., Williamson, S.J., Tovchigrechko, A., Allen, A.E., Zeigler, L.A., Sutton, G., Eisenstadt, E., Rogers, Y.H., Friedman, R., Frazier, M. and Venter, J.C., 2010. Genomic and functional adaptation in surface ocean planktonic prokaryotes. Nature 468, 60-66.

Zhang, G.X., Dong, H.L., Jiang, H.C., Xu, Z.Q. and Eberl, D.D., 2006. Unique microbial community in drilling fluids from Chinese Continental Scientific drilling. Geomicrobiology Journal 23, 499-514.

Zhang, G.X., Dong, H.L., Xu, Z.Q., Zhao, D.G. and Zhang, C.L., 2005. Microbial diversity in ultra-high-pressure rocks and fluids from the Chinese Continental Scientific Drilling Project in China. Appl Environ Microbiol 71, 3213-3227.

APPENDIX A: SUPPLEMENTAL DISCUSSION

TABLE OF CONTENTS

Page

A.1 CELL DENSITY ESTIMATES FOR ROCK CORE RINSES 67
A.2 SUPPLEMENTAL DNA DISCUSSION 68
A.2.1 SPECIES LEVEL RANK ABUNDANCE CURVE FOR PYROTAG BATCH III 69
A.2.2 ACHAEAL PYROTAG SUMMARY 70
A.2.3 LEAST ABUNDANT BACTERIAL SEQUENCES IN 494 QUEENSTON SAMPLES 71

LIST OF TABLES

Table A.1.1: Comparison of PLFA concentrations and cell density estimates for varying depth of penetration of rock core rinses 68
Table A.2.2: Archaeal pyrotag summary: Most abundant species detected 70
Table A.2.3: Least abundant bacteria in samples associated with 494-Queenston 71

LIST OF FIGURES

Figure A.2.1: Rank abundance curve at species level for Pyrotag Batch III 69

A.1 Cell Density Estimates for Rock Core Rinses

PLFA concentrations and cell density estimates calculated for the rock core rinses were calculated based on the assumption that the outer 3 mm of rock material are extracted during rinsing with extraction solvents. This assumption is based on the observation of a 3 mm outer rim of solvent-saturated rock material after rinsing. In support of this observation, Sherman et al. (2007) reported the presence of external contaminants within the outer 3-5 mm rim of rock core material. The PLFA concentrations and cell density estimates calculated for the rock core rinses will change depending on this assumption. In order to assess the extent of this change, PLFA concentrations and cell density estimates were re-calculated using alternative assumptions (i.e., that the outer 1 mm or the outer 5 mm of rock material are extracted during rinsing) (Table A1.1). The total mass of rock material extracted during rinsing was calculated as follows: the total volume of rock material composing the interior of the rock core (excluding the outer 1, 3 or 5 mm of material) was first calculated, using $V = L\pi r^2$, where L is the length of the rock core minus the outer 1, 3, or 5 mm and r is the radius of the rock core minus the outer 1, 3, or 5 mm. This volume was subsequently subtracted from the total volume of the entire rock core sample, in order to calculate the total volume of rock material within the outer 1, 3, or 5 mm of the rock core. This volume was multiplied by the density of the rock core sample to determine the mass of rock material within the outer 1, 3 or 5 mm.

The results from these calculations indicate that, even if the solvents extract up to 5 mm into the core, the concentrations of PLFA on the outer rock core surfaces are greater than or equal to the concentrations in the interiors. There are no cases where the PLFA concentrations in the rock core interiors are higher than the concentrations on the core surfaces. This observation is expected, as any PLFA indigenous to the interiors of the cores are likely to be present on the surfaces of the cores as well.

Table A.1.1: Comparison of PLFA Concentrations and Cell Density Estimates for Varying Depth of Penetration of Rock Core Rinses: Assuming That the Outer 1, 3 or 5 mm of Rock Material are Extracted During Rinsing

Core Rinse	PLFA concentration (ng/g of rock)			Cell density estimate (cells/g of rock)		
	1 mm	3 mm	5 mm	1 mm	3 mm	5 mm
334o_SalinaA1	7.8	2.7	1.7	$2x10^6$	$5x10^5$	$3x10^5$
464o_Queenston	51.7	17.8	11.0	$1x10^7$	$4x10^6$	$2x10^6$
529o_GeorgianBay	220.7	76.2	47.4	$5x10^7$	$2x10^7$	$1x10^7$
661o_Cobourg	9.8	3.4	2.1	$2x10^6$	$8x10^5$	$5x10^5$
682o_Cobourg	5.3	1.8	1.1	$1x10^6$	$4x10^5$	$3x10^5$
694o_ShermanFall	2.8	1.0	0.6	$6x10^5$	$2x10^5$	$1x10^5$

A.2 Supplemental DNA Discussion

To preserve the flow of the main narrative, several supplementary figures and tables supporting the same conclusions as those used in the preceding chapters are appended here. Figure B1 shows a rank abundance curve for bacterial sequences detected by 454 pyrotag analysis Batch-II. The overall pattern in this figure is strikingly similar to the other two batch summaries. Table A2.2 shows the species-level results of an attempt by the DNA sequencing provider to obtain archaeal sequences from DNA extracts from 334DWs_Salina A1, 334os_Salina A1 and 464DWs_Queenston. The experimental procedure and the significance of these results are described in the main body of the report, but overall, these results are most consistent with the detection of some form of reagent-associated contamination. Finally, Table B3 shows the results of an exploration of the least abundant bacterial sequences obtained from 454 pyrotag sequencing of DNA extracts from the 494 Queenston inner core sample. As the result was essentially identical to the other version of this test that was performed, this figure is shown in the Appendix.

A.2.1. SPECIES LEVEL RANK ABUNDANCE CURVE FOR PYROTAG BATCH III.

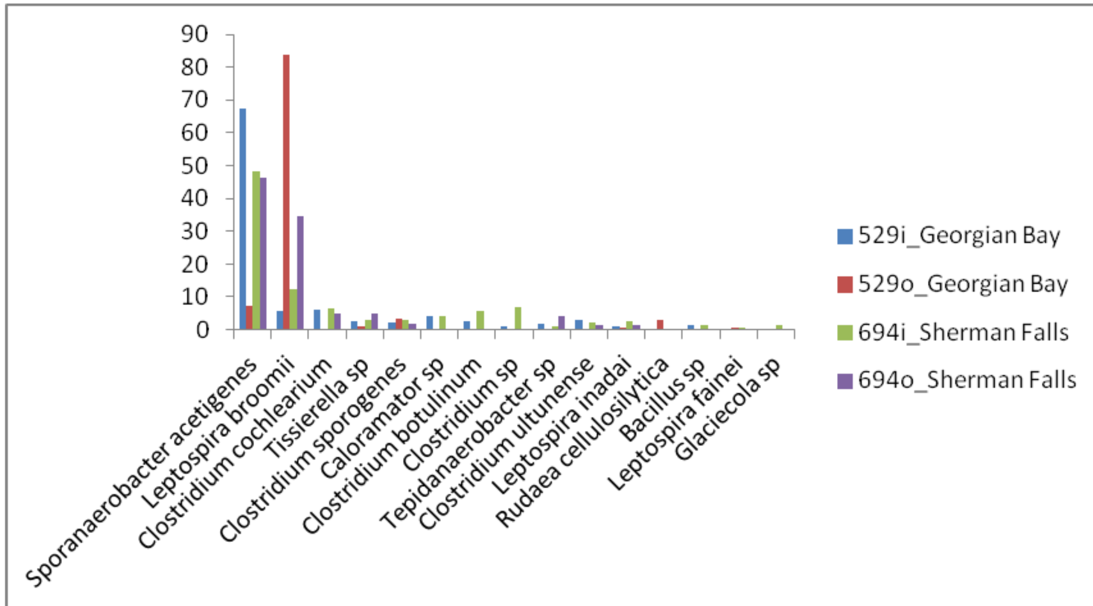

Figure A.2.1: Rank abundance curve at species level for Pyrotag Batch III

A.2.2 ARCHAEAL PYROTAG SUMMARY

Table A.2.2: Archaeal Pyrotag Summary: Most Abundant Species Detected

	Cenarchaeum sp. (% to total)	*Candidatus Nitrosopumilus* sp. (% to total)	*Nitrosopumilus maritimus* (% to total)
"Small Kit"_Control	[a]---	---	---
Mega Kit_Control	---	---	---
Baked_494_Control	---	---	---
Baked_682_Control	---	---	---
334DW$_s$_Salina A1	78.7	21.1	0.03
334o$_s$_Salina A1	97.5	2.4	
334i$_s$_Salina A1	---	---	---
334o_Salina A1	---	---	---
334i_Salina A1	---	---	---
464DW$_s$_Queenston	99.4	0.14	
464o_Queenston	---	---	---
464i_Queenston	---	---	---
494DW$_s$_Queenston	---	---	---
494o_Queenston	---	---	---
494i_Queenston	---	---	---
529o_Georgian Bay	---	---	---
529i_Georgian Bay	---	---	---
661DW$_s$_Cobourg	---	---	---
661o_Cobourg	---	---	---
661i_Cobourg	---	---	---
670DW$_s$_Cobourg	---	---	---
670o_Cobourg	---	---	---
670i_Cobourg	---	---	---
682o_Cobourg	---	---	---
682i_Cobourg	---	---	---
694o_Sherman Fall	---	---	---
694i_Sherman Fall	---	---	---

[a]Failed to amplify/below detection limit

A.2.3 LEAST ABUNDANT BACTERIAL SEQUENCES IN 494 QUEENSTON SAMPLES

Table A.2.3: Least Abundant Bacteria in Samples Associated with 494-Queenston

Baked_494_Control	494DW$_s$_Queenston	494o_Queenston	494i_Queenston	Mega Kit_Control
Blastochloris sp	*Alicyclobacillus sp*	*Anabaena augstumalis*	*Aquaspirillum sp*	*Azospirillum zeae*
Burkholderia glumae	*Bradyrhizobium sp*	*Bacillus cereus*	*Brevundimonas terrae*	*Blastococcus sp*
Enterobacter hormaechei	*Bradyrhizobium sp*	*Brevundimonas terrae*	*Clostridium sp*	*Caulobacter sp*
Janibacter melonis	*Hafnia sp*	*Cylindrospermum sp*	*Conexibacter sp*	*Clostridium sp*
Mycoplana sp	*Prevotella sp*	*Dolichospermum circinale*	*Dehalococcoides sp*	*Escherichia sp*
Pseudomonas sp	*Prevotella sp*	*Mesorhizobium sp*	*Ferrithrix thermotolerans*	*Mycoplana bullata*
Pseudomonas sp	*P.fluorescens*	*Methylobacterium oryzae*	*Pseudomonas sp*	*Mycoplana sp*
	P. fluorescens	*Microbacterium maritypicum*	*Streptococcus mitis*	*Peptoniphilus sp*
	Variovorax paradoxus	*Microbacterium sp*		*Propionibact.. granulosum*
		Mycobacterium chelonae		
		Mycoplana bullata		
		Staphylococcus epidermidis		
		Streptococcus sp		
		Synechococcus sp		