

*MICROBIAL DIVERSITY AND METABOLIC POTENTIAL OF THE SERPENTINITE
SUBSURFACE ENVIRONMENT*

By

Katrina Irene Twing

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Microbiology and Molecular Genetics – Doctor of Philosophy

2015

ABSTRACT

MICROBIAL DIVERSITY AND METABOLIC POTENTIAL OF THE SERPENTINITE SUBSURFACE ENVIRONMENT

By

Katrina Irene Twing

Serpentinization is the hydrous alteration of mafic rocks to form serpentine minerals and magnetite. The reactions of this alteration result in elevated pH of the surrounding fluids, abiotic generation of H₂, CH₄ (and other organic molecules), and depletion of dissolved inorganic carbon. Thus, serpentinization has implications for the origin of life on Earth and possibly Mars and other planetary bodies with water. The microbial diversity of continental serpentinite systems consistently shows communities that are dominated by two major taxa – microaerophilic Betaproteobacteria and anaerobic Clostridia. Previous studies relied on few samples collected from natural springs or seeps, meaning that the flow path of fluids from the subsurface process of serpentinization was unknown. The Coast Range Ophiolite Microbial Observatory (CROMO), a set of wells drilled into the actively serpentinizing subsurface environment in northern California, was established in northern California to gain a better understanding of the habitability and microbial functions within the serpentinite subsurface environment.

This dissertation represents a culmination of microbiological investigations into the serpentinite subsurface environment at CROMO to identify the microbial inhabitants of subsurface fluids, rocks, and *in situ* colonization experiments using molecular methods and high-throughput sequencing. The CROMO wells represent a broad range of geochemical gradients and pH and the concentrations of carbon monoxide and methane have the strongest correlation with microbial community composition. The most extremely high pH wells were inhabited exclusively by a single operational taxonomic unit (OTU) of Betaproteobacteria and a few OTUs

of Clostridia, while more moderate pH wells exhibited greater diversity. Genes involved in the metabolism of hydrogen, carbon monoxide, and carbon fixation were abundant in the extreme pH fluids, while genes for metabolizing methane were exclusively in the moderate pH wells.

The subsurface environment is an amalgamation of fluids and rocks, and as such, studying fluids alone only gives half the story. CROMO represents the first drill campaign into the continental serpentinite environment and the microbial diversity of serpentinite cores to a depth of 45 meters below surface suggests that specific geological features harbor different microbial communities. Archaea, previously undetected at CROMO, dominated cores containing magnetite-bearing serpentine, while bacteria were more abundant in layers containing clay particles. Additionally, organisms involved in the cycling of nitrogen and methane were found associated with core materials, indicating core-associated communities may have strong biogeochemical roles within the serpentinite subsurface environment. Given that microbial communities appear to vary with geological composition and that serpentinite fluids are a challenging habitat for life, depleted in inorganic carbon and electron acceptors, microbe-mineral interactions within the serpentinite subsurface environment through the use of *in situ* colonization devices to see if communities were able to utilize inorganic carbon in calcite or ferric iron as a terminal electron acceptor from magnetite. In the highest pH well, calcite led to an increased abundance of Clostridia and Deinococcus, while magnetite led to an increase in diversity, including Alphaproteobacteria, Gammaproteobacteria, and Actinobacteria, suggesting further that mineralogical composition of solids within the subsurface impact community composition. The data discussed here further our understanding of life associated with serpentinite fluids and minerals within the subsurface environment.

ACKNOWLEDGEMENTS

My scientific education and research has been a great journey, along which I have received advice and mentorship from many people I would like to thank now. A huge thank you to my advisor and mentor, Dr. Matt Schrenk for his unwavering support and guidance over the last few years. While in his laboratory, I have always felt like I am a valued and contributing member of a team. I would also like to thank the rest of CROMO team, Dr. Dawn Cardace, Dr. Tori Hoehler, Dr. Tom McCollom, Mike Kubo, and Dan Carnevale for great discussion and guidance through this interdisciplinary project and for a memorable field experience. Thank you to Christopher Thornton and Alex Hyer for technical assistance and guidance on data analysis. Thank you to Drs. Melitza Crespo-Medina and Dani Morgan-Smith for being great lab mates and even better friends. Thank you to Dr. Billy Brazelton for his mentorship over the years. I've thoroughly enjoyed working with and learning from him and am eagerly looking forward to the next stage of my career with him as a postdoctoral researcher in his laboratory.

I would like to acknowledge the NASA Astrobiology Institute for funding the establishment of CROMO, where the majority of my dissertation research took place, and the Deep Carbon Observatory for research funding and community networking opportunities. Thank you to my dissertation guidance committee, including Drs. Matt Schrenk, Kazem Kashefi, Ashely Shade, Warren Wood, and Tori Hoehler, for their insightful and constructive input on my research project and dissertation. Having transferred to MSU during the middle of my degree, I greatly appreciated the welcoming guidance of my committee and Dr. Bob Hausinger and Ms. Betty Miller in navigating my new department and community. Thanks to these people, I felt truly welcome in the MMG department and consider it an honor to now be an alumna of it.

Thank you to the strong female educators and scientists who have mentored me throughout my scientific education. I had a strong interest in the biological sciences from an early age, thanks to my middle and high school science teachers, Ms. Laubach and Ms. Brunette, who introduced my inquisitive mind to the scientific process. While perusing my B.A. at Clark University, Dr. Deborah Robertson mentored me through my first laboratory experience and introduced me to the wonders of marine microbiology. During my M.S. research at the University of Delaware, my advisor Dr. Barbara Campbell helped me become an independent researcher. I owe a very big thank you to Dr. Jennifer Biddle for being a great friend and mentor over the years – and particularly for introducing me to Matt and setting me on this exciting journey into subsurface microbiology.

Finally, I would like to thank my incredibly supportive family who have always believed in me and helped me achieve my dreams. Thank you to my mother, Ms. Lori Twing, for encouraging me to always do my best and reach for the stars. Her unwavering belief that I can achieve anything I put my mind helped to mold me into the dedicated scientist I am today. I would like to thank both my father Mr. Jeff Twing and stepfather Mr. Mike Marrone for encouraging me to look at challenges from different angles in order to find the best solution. A big thank you to my loving brother, Mr. Peydon Twing, for always being my biggest cheerleader and best friend. A very special thank you to my love and my partner, Mr. Eric Gardner, for standing beside me through both the good and stressful times of my Ph.D. education and supporting my dreams for the future.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
KEY TO ABBREVIATIONS	xi
CHAPTER 1 – Introduction	1
<i>Serpentinization</i>	<i>1</i>
<i>Habitability and Challenges of the Extreme Environment</i>	<i>2</i>
<i>Lost City Hydrothermal Field.....</i>	<i>4</i>
<i>Continental Ophiolites</i>	<i>6</i>
<i>Coast Range Ophiolite Microbial Observatory</i>	<i>10</i>
REFERENCES.....	14
CHAPTER 2 – Serpentinizing fluids from a subsurface observatory harbor extremely low diversity microbial communities adapted to high pH	19
Abstract.....	19
Introduction.....	20
Results and Discussion.....	22
<i>Sampling Site and Geochemistry</i>	<i>22</i>
<i>Bacterial 16S rRNA Gene Diversity and Community Composition.....</i>	<i>26</i>
<i>Biogeochemical Relationships</i>	<i>32</i>
<i>Metabolic Potential.....</i>	<i>34</i>
Conclusions.....	39
Materials and Methods.....	41
<i>Site Description and Sample Collection</i>	<i>41</i>
<i>Geochemistry</i>	<i>42</i>
<i>Microbial Cell Counts.....</i>	<i>42</i>
<i>DNA Extraction.....</i>	<i>43</i>
<i>16S rRNA Tag Sequencing and Data Analysis</i>	<i>43</i>
<i>Metagenomic Sequencing and Data Analysis.....</i>	<i>44</i>
<i>Statistical Analyses</i>	<i>45</i>
<i>Bioenergetic Calculations.....</i>	<i>45</i>
REFERENCES.....	48
CHAPTER 3 – Microbial diversity of serpentinite cores from the Coast Range Ophiolite Microbial Observatory	53
Abstract.....	53
Introduction.....	54
Results and Discussion.....	56
<i>Drill Core Depth Profiles</i>	<i>56</i>
<i>DNA Yield, Organic Carbon, and 16S rRNA Gene Abundance</i>	<i>58</i>
<i>Core-Enriched Bacterial Communities.....</i>	<i>60</i>
<i>Core-Enriched Archaeal Communities.....</i>	<i>66</i>
<i>Biogeochemical Correlations</i>	<i>68</i>
Conclusions.....	70

Materials and Methods	74
<i>Site Description and Sample Collection</i>	74
<i>DNA Extraction</i>	74
<i>Quantitative-PCR</i>	75
<i>16S rRNA Gene Sequencing and Data Analysis</i>	75
<i>Mineralogical and Geochemical Analyses</i>	76
<i>Statistical Analyses</i>	76
REFERENCES	77
CHAPTER 4 – Microbe-mineral interactions in the serpentinite subsurface environment	82
Abstract	82
Introduction	82
Results and Discussion	84
<i>In Situ Colonization Devices and CROMO Wells</i>	84
<i>16S rRNA Gene Abundance</i>	85
<i>Bacterial Community Composition</i>	85
<i>Mineral-Enriched Bacteria</i>	89
Conclusions	96
Materials and Methods	97
<i>In Situ Colonization Devices</i>	97
<i>DNA Extraction</i>	97
<i>Quantitative-PCR</i>	98
<i>16S rRNA Gene Sequencing and Data Analysis</i>	98
REFERENCES	100
CHAPTER 5 – Summary and Future Perspectives	104
Summary and Future Perspectives.....	104
REFERENCES	108
APPENDICES	110
APPENDIX A – Example DESeq commands for identifying group specific sequences .	111
REFERENCES	119
APPENDIX B – Bioinformatic analyses for Crespo-Medina <i>et al.</i>, 2014	121
REFERENCES	126

LIST OF TABLES

<i>Table 2.1 – Environmental and geochemical parameters associated with samples collected in August 2012</i>	24
<i>Table 2.2 – Significant correlations between environmental parameters according to Pearson’s correlations analysis</i>	25
<i>Table 2.3 – Summary of significant correlations (p-value < 0.05) between top OTUs (making up > 25% of the sample in which they are most abundant) and environmental parameters</i>	30
<i>Table 2.4 – Microbial communities at CROMO are dominated by bacteria</i>	31
<i>Table 2.5 – Summary of significant correlations between gene hits and OTUs</i>	36
<i>Table 2.6 – Bioenergetic calculations for CROMO groundwaters</i>	36
<i>Table 3.1 – Soil and core sample names and 16S rRNA sequence abundances</i>	59
<i>Table 3.2 – Relative abundance of individual core-enriched bacterial taxa within samples</i>	64
<i>Table 3.3 – Relative abundance of individual core-enriched archaeal taxa within samples</i>	69
<i>Table 3.4 – Classification of lithostratigraphic units and samples belonging to them</i>	70
<i>Table 4.1 – Geochemical parameters of well fluids measured at in situ device deployment and retrieval</i>	84
<i>Table 4.2 – DESeq comparisons for determining fluid- and mineral-specific taxa</i>	89
<i>Table 4.3 – Relative-abundance of fluid-enriched bacterial taxa in different sample types</i>	94
<i>Table 4.2 – Relative-abundance of magnetite-enriched bacterial taxa in different sample types</i>	94
<i>Table 4.3 – Relative-abundance of calcite-enriched bacterial taxa in different sample types</i>	94

Table B.1 – Percent identity between end point sequences from March microcosms, most abundant OTUs from tag sequence data, and reference sequences124

Table B.2 – Percent identity between two representative end point sequences from the August microcosms, most abundant OTUs from tag sequence data, and reference sequences124

LIST OF FIGURES

<i>Figure 1.1 – Conceptual model of biogeochemical environment at marine serpentinite sites, such as the Lost City Hydrothermal Field</i>	5
<i>Figure 1.2 – Conceptual model of biogeochemical environment at continental serpentinite sites</i>	9
<i>Figure 1.3 – Map of northern California, depicting the location of CROMO</i>	11
<i>Figure 2.1 – Microbial community structure</i>	27
<i>Figure 2.2 – Inter-well community composition comparison from August 2012</i>	28
<i>Figure 2.3 – Network diagram of significant correlations between OTUs and environmental variables</i>	32
<i>Figure 2.4 – Heatmap of most abundant OTUs and geochemical parameters across all samples</i>	33
<i>Figure 2.5 – Abundance of protein-encoding genes in metagenomes</i>	35
<i>Figure 3.1 – Depth profile of CSW1.1</i>	56
<i>Figure 3.2 – Depth profile of QV1.1</i>	57
<i>Figure 3.3 – DESeq comparisons between CSW core and CROMO fluid samples</i>	61
<i>Figure 3.4 – Relative abundance of core-, soil-, and fluid-enriched bacterial sequences</i>	62
<i>Figure 3.5 – Bacterial community structure of unique sequences</i>	63
<i>Figure 3.6 – Relative abundance of core-, soil-, and fluid-enriched archaeal sequences</i>	66
<i>Figure 3.7 – Archaeal community structure of unique sequences</i>	67
<i>Figure 3.8 – Correlation network of bacteria and lithostratigraphic units</i>	71

<i>Figure 3.9 – Negative correlation between iron concentration and DNA yield for select taxa</i>	71
<i>Figure 3.10 – Correlation network of archaea and lithostratigraphic units</i>	72
<i>Figure 4.1 – In situ colonization devices and schematic of deployment down well</i>	86
<i>Figure 4.2 – 16S rRNA gene copy abundance per gram of substrate, as determined by q-PCR</i> .	87
<i>Figure 4.3 – Rarefaction curve of unique sequences from in situ colonization device and well fluid sequencing</i>	88
<i>Figure 4.4 – Community similarity dendrogram of unique sequences from in situ colonization devices, well fluids, and core material</i>	90
<i>Figure 4.5 – Community diversity bar chart for CROMO fluids and in situ colonization devices</i>	92
<i>Figure 4.6 – Relative abundances of fluid-, glass-, magnetite-, and calcite-enriched bacteria taxa per each sample in all four wells</i>	95
<i>Figure B.1 – Diversity of 16S rRNA gene OTUs (97% sequence similarity) from fluids collected at CROMO in March 2013 was analyzed using mothur</i>	121
<i>Figure B.2 – Proportion of bacterial classes present in fluids collected at CROMO in March 2013</i>	122
<i>Figure B.3 – Diversity of OTUs (97% similarity) belonging to the classes Betaproteobacteria and Clostridia within CSW1.1, N08B, and CSWold</i>	123

KEY TO ABBREVIATIONS

- ANAMMOX – anaerobic ammonium oxidation
- ANOSIM – analysis of similarity
- BLASTP – basic local alignment search tool protein
- C/N – carbon to nitrogen ratio
- CH₄ – methane
- CO – carbon monoxide
- CO₂ – carbon dioxide
- cooS* – carbon monoxide dehydrogenase (anaerobic)
- CORK – Circulation Obviation Retrofit Kit
- coxL* – carbon monoxide dehydrogenase, large chain (aerobic)
- CROMO – Coast Range Ophiolite Microbial Observatory
- CSW – Core Shed Well
- CVA – Cabeço de Vide Aquifer
- DAPI – 4',6-diamidino-2-phenylindole
- DIC – dissolved inorganic carbon
- DNA – deoxyribonucleic acid
- DO – dissolved oxygen
- DOE – Department of Energy
- EA-IRMS – elemental-analysis isotope-ratio mass spectrometry
- H₂ - hydrogen
- IODP – International Ocean Discovery Program
- JGI – Joint Genome Institute

LCHF – Lost City Hydrothermal Field

LCMS – Lost City Methanosarcinales

Mbs – meters below surface

mcrA – methyl coenzyme-M reductase

MG-RAST – Metagenomics Rapid Annotation using Subsystem Technology

mxoF – methanol dehydrogenase

ORP – oxidation reduction potential

OTU – operational taxonomic unit

Pfam – protein family

pmoA – particulate monooxygenase

qPCR – quantitative polymerase chain reaction

RDP – Ribosomal Database Project

rRNA – ribosomal ribonucleic acid

RuBisCO – ribulose-1,5-bisphosphate carboxylase/oxygenase

SIMPROF – similarity profile analysis

TOC – total organic carbon

TN – total nitrogen

QV – Quarry Valley

VAMPS – Visualization and Analysis of Microbial Population Structures

XRD – x-ray diffraction

XRF – x-ray fluorescence

CHAPTER 1

Introduction

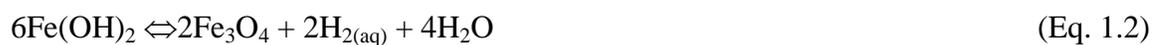
Serpentinization

Serpentinization is a widespread geochemical process, involving the water-rock alteration of ultramafic rocks, rich in the mineral olivine (forsterite and fayalite) and originating within the Earth's mantle (Equation 1.1).



fayalite + forsterite + water \rightarrow serpentine + brucite + iron (II) hydroxide

The reaction of ferrous iron (Fe^{2+}) in the resulting iron (II) hydroxide oxidizing to ferric iron (Fe^{3+}) results in the production of the mineral magnetite and release of hydrogen gas (Equation 1.2).



iron (II) hydroxide \Leftrightarrow magnetite + hydrogen gas + water

At temperatures below 150°C, this process results in high pH (>10) and the released of H_2 that, under certain circumstances, can react with CO_2 or CO to abiotically produce methane and small-chain hydrocarbons through Fischer-Tropsch Type reactions (Equation 1.3; 11, 27, 34).



Serpentinite fluids are typically enriched in Ca^{2+} ions, released from Ca-silicate minerals associated with olivine in ultramafic rocks (15). At high pH, the predominant dissolved inorganic carbon (DIC) species is carbonate (CO_3^{2-}), which readily precipitates out of solution with the Ca^+ ions in the fluids, producing calcium carbonate and effectively scrubbing all DIC from the fluids (2, 22). Thus, serpentinization creates a highly reduced, high pH environment enriched in H_2 and methane and depleted in DIC.

On the early Earth, before the crust had fully differentiated, ultramafic rocks were likely pervasive, and serpentinization may have been even more common than it is today (41). It has been speculated that abiogenic organic molecules formed in alkaline environments, such as those derived from serpentinization, were involved in prebiotic chemistry (36, 37) and could have led to the molecules necessary for the development of life on Earth. On Mars, mineralogical data from remote sensing indicate the presence of olivine (26) and serpentinite (14), providing a strong possibility that serpentinization has taken place there as well. Furthermore, possible methane fluxes into the Martian atmosphere (33, 51) indicate that serpentinization may be continuing under the Martian surface today. These similarities indicate that an understanding of microbiology and habitability within the serpentinite environment not only informs our understanding of present-day subsurface life, but also can have implications for the origin of life here on Earth and elsewhere in the universe.

Habitability and Challenges of the Extreme Environment

The habitability of an environment, or its potential to host life, can be modeled by calculating the balance of energy produced by an environment and the energy required by life forms to grow and function (19). Organisms can exist in different levels of activity: growth,

where cells are actively replicating and thriving; maintenance, where cells are supporting only the most basic metabolic functions; or survival, a state of hibernation where cells are not metabolically active. Each of these activity levels requires a different amount of energy for sustenance (18, 34). Thermodynamic models have indicated that the serpentinite environment provides enough potential energy, in the form of hydrogen (H_2) and methane (CH_4), to be habitable for microbial growth (29).

There are additional challenges to life in the serpentinite environment beyond the supply of energy required for life. The high pH of the environment not only challenges the use of a proton-motive force to generate cellular energy (ATP) (23, 39), but it also leads to the depletion of available DIC in the system. At high pH, the most abundant form of DIC is carbonate ion, which readily precipitates out of solution as calcium carbonate minerals (1, 22). Furthermore, the energy supply in the form of electron donors (i.e. H_2 and CH_4) is only bioavailable if there are electron acceptors present to be oxidized, and thus allow for microbial metabolism to proceed. While the serpentinite environment is rich in electron donors, in the form of H_2 and CH_4 , there is a depletion of corresponding electron acceptors in the environment for microbial growth and metabolism.

Some microbes are able to obtain energy from inorganic mineral phases in their environment, through a process called chemolithotrophy (40). A recent study looking at the diversity of microbes from the glacial environment cultured *in situ* on various mineral substrates has helped elucidate the microbe-mineral interactions within the system by identifying minerals that enriched for unique microbial communities (30). While the serpentinite subsurface environment is rich in electron donors (which can serve as potential energy sources), it lacks inorganic carbon and electron acceptors necessary for microbial life. Some of these chemical

constituents are found in mineral forms in the serpentinite environment, such as carbonate ions in calcite and ferric iron in magnetite, and could potentially be utilized by chemolithotrophic microorganisms.

Lost City Hydrothermal Field

The most well-characterized serpentinite ecosystem to date is the Lost City Hydrothermal Field (LCHF), which was discovered in 2000 on the Atlantis Massif at 30°N near the Mid-Atlantic Ridge (21, 22). The vent represents a 1.5 million year old serpentinite system that has been actively venting for at least 30,000 years (16). LCHF is characterized by large calcium carbonate chimneys extending up to 60 m from the seafloor and venting hot (40-90°C), high pH (9-11) fluids containing elevated concentrations of hydrogen (> 14 mM), methane (1-2 mM), and short chain organic molecules (21, 22). Isotopic analysis indicates that the most abundant organic molecules at LCHF, such as methane and formate, are of abiogenic origin (24, 35). The Lost City also exhibits increased concentrations of acetate relative to background seawater, which is thought to be of microbial origin, either as a product of autotrophy or the byproduct of the heterotrophic breakdown of larger organic molecules (24). Whether microbes are able to utilize mantle-derived carbon is a major question in a system where organic molecules are abiotically generated. Lang *et al.* (25) demonstrated that up to 50% of the carbon in biomass from actively venting chimney samples was mantle-derived, indicating that microbes in the serpentinite environment may be utilizing the carbon from the deep Earth.

The deep, highly reducing fluids from serpentinization mixing with ocean water, containing oxidants and CO₂, results in a chemical disequilibrium that microorganism in the environment can exploit for energy. The high abundance of methane at the LCHF has clearly impacted the microbial communities at the vents, as the target gene for methanogenesis and

anaerobic methane-oxidation, methyl-coenzyme M reductase (*mcrA*), was detected in chimney samples (22). There are distinct differences in community composition between young, actively venting chimneys and inactive or extinct chimneys. The anoxic interior of the actively venting chimneys was dominated by a single archaeal phylotype capable of metabolizing methane, termed the Lost City Methanosarcinales (LCMS; 6, 38). Active chimneys also contained methylotrophic *Gammaproteobacteria* (3). Meanwhile, older, inactive chimneys were dominated by anaerobic methane-oxidizing archaea from the ANME-1 group (3, 5). While the LCMS exhibited little diversity among 16S rRNA genes, metagenomic analysis displayed unprecedented abundance and diversity among transposase genes, which are involved in gene duplication and transfer (4) and were experimentally shown to exhibit diverse physiology with regards to the production and oxidation of methane (6).

Hydrogen-oxidizing Betaproteobacteria of the order Burkholderiales and potentially

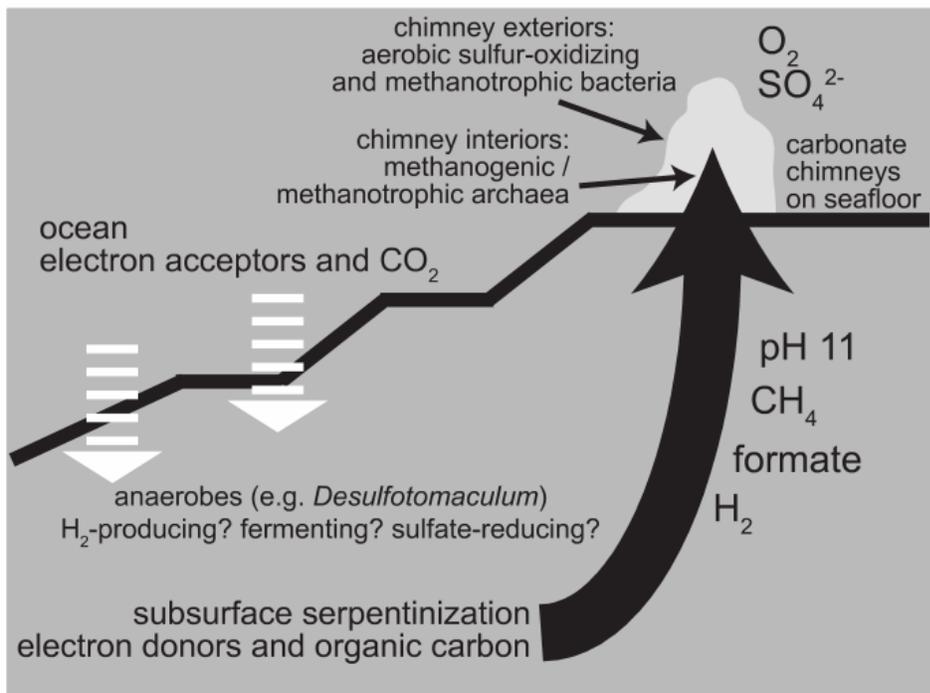


Figure 1.1 – Conceptual model of biogeochemical environment at marine serpentinite sites, such as the Lost City Hydrothermal Field. Reproduced from Schrenk, et al., 2013.

fermentative and/or sulfate-reducing Clostridia were also identified at the LCHF (5, 7). Additionally, genes required for carbon-fixation, both via the Calvin-Benson-Bassham pathway (RuBisCO) and the Wood-Ljungdahl pathway were detected in samples from the Lost City (5). Figure 1.1 depicts a conceptual hydrogeological model that provides a transport mechanism to sustain life at the LCHF, where oxic seawater (rich in electron acceptors) mixes with highly reducing serpentinite fluids rich in electron donors, providing a zone of high energy potential for microbial life (Fig. 1.1; 39). Both chemical and biological data from the Lost City Hydrothermal Field suggest that microorganisms in this ecosystem are capable of metabolizing the geochemical products of serpentinization (3, 25).

Continental Ophiolites

Serpentinization takes place when mantle-derived ultramafic rocks are exposed to the surface along slow and ultraslow spreading ridges (22), as is evident with the LCHF. Through tectonic processes, such as obduction, the ocean crust and mantle associated with it are emplaced onto the continents crust. Obduction, allows for the relatively easy access to the mantle and its geochemical weathering processes.

The Tablelands Ophiolite in Newfoundland, Canada, is a 500 million year old serpentinite, characterized by high pH springs, enriched in H₂ (0.5 mM) and CH₄ (18.8 μM) (45). Metagenomic analysis of one of the end member springs yielded some insights into the metabolic potential and extremely low biodiversity of the continental subsurface environment (7). Genes encoding enzymes involved in both the oxidation and production of H₂, [NiFe]-hydrogenase and [FeFe]-hydrogenase, respectively, were identified within the sample, indicating that organism within the community have the genetic potential not only to consume H₂ gas, but produce it as well. Phylogenetic analysis of the [NiFe]-hydrogenase identified the gene as

belonging to members of the Burkholderiales. Furthermore, the largest contig formed from the sequence data contained operons for the [NiFe]-hydrogenase, as well as carbon monoxide dehydrogenase (*coxL*) and RuBisCo (7) and the *coxL* gene bore strong homology to *Hydrogenophaga pseudoflava* (order Burkholderiales). The presence of these genes on the same contig, combined with the fact that *H. pseudoflava* is a facultative anaerobe, capable of autotrophic growth on H₂ and carbon monoxide (50), are evidence that these microbes may be inhabiting a mixing zone, where H₂ and O₂ are both present. Additionally, microcosm experiments from the Tablelands found carbon monoxide-oxidation to take place in samples with an abundance of *Hydrogenophaga* (32). Phylogenetic analyses of the [FeFe]-hydrogenase indicate that it comes from members of the Clostridiales, who are known to be anaerobic fermenters. A recent interdisciplinary study at Tablelands using correlation networks to determine relationships between microbiology and geochemical factors found that *Hydrogenophaga* was likely to inhabit H₂-rich transition zones, while a member of the Firmicutes, related to Clostridia, was more likely to inhabit the anoxic end-member fluids (8). The data indicate that the ecosystem is made up of two main players: aerobic H₂-fueled, autotrophic Burkholderiales inhabiting the oxygen mixing-zone and anaerobic H₂-producing Clostridia/Firmicutes, that may be autotrophically fermenting the abiogenic hydrocarbons produced by serpentinization, in end-member fluids (7, 8).

Cabeço de Vide is a high pH (11.4), serpentinite-influenced spring in Portugal, which has been studied for its microbiology (46, 47, 48). The spring has extremely low microbial biomass, with cell abundances of roughly 6×10^2 cells mL⁻¹ (46, 48). Similar to findings at the Tablelands Ophiolite (7, 8), 16S rRNA gene clone libraries and tag sequencing analyses indicate that the bacterial community is dominated by Clostridia, some of which are closely related to species

identified in other high-hydrogen, subsurface ecosystems (12, 48). The sample showed relatively high bacterial diversity, compared to other serpentinite sites (8), which included the Betaproteobacterium *Hydrogenophaga*, and low archaeal diversity, including methane-oxidizers of the ANME-1 group (48). Analysis of functional genes targeting carbon-fixation pathways indicated that the Calvin-Benson-Bassham cycle was the only pathway used. Additionally, genes for sulfur- and methane-oxidation were identified (48).

In recent years the microbiology at The Cedars, a site of serpentinization located in Northern California on the Franciscan Subduction Complex, has been explored. The Cedars hosts highly reducing (-656 to -585 mV), high pH fluids (pH 11-12) of non-marine origin (31). Studying two springs at The Cedars over three years, researchers observed stable microbial communities strongly correlated with the source of spring fluids (43). The deeper fluids, characterized by pH 11.9, redox potential of -700 mV, contained Clostridia, Chloroflexi, members of the candidate-phylum OD-1 and Euryarchaea (43). In contrast, the shallower spring, thought to be a mixing zone of deeply sourced serpentinite fluids and surface groundwater, were dominated by Betaproteobacteria closely related to the genus *Hydrogenophaga* (43). These findings are consistent with previous studies of the serpentinite system, suggesting that Clostridia (and other taxa) live in the deeper more end-member serpentinite fluids and Betaproteobacteria live at anoxic/oxic interfaces (7, 8, 39).

A novel Betaproteobacterium closely related to *Hydrogenophaga* was isolated from The Cedars and given the proposed genus *Serpentinomonas* (44). The three strains of *Serpentinomonas* (A1, B1, and H1) are all alkilaphilic (optimum pH of 11) and autotrophic with growth on hydrogen, oxygen, and calcium carbonate (44). The strains differ in their electron acceptor utilization, with B1 and H1 using nitrate and A1 using thiosulfate (44). According to

genomic data, all three strains contain the CO₂-fixation gene, RuBisCO and the [NiFe]-hydrogenase, needed for hydrogen-oxidation, while only strain A1 possesses carbon monoxide dehydrogenase, *coxL* (44). Comparisons of the 16S rRNA genes of the strains to previously published studies of continental serpentinites suggests that Betaproteobacteria previously described as *Hydrogenophaga* at CVA (48) and Tablelands (8) are 99-100% identical to *Serpentinomonas* strains H1 and A1, respectively (44).

Microbiological studies of continental serpentinites from around the world show striking similarities in community composition, dominated by anaerobic Clostridia and microaerophilic, hydrogen-oxidizing Betaproteobacteria (Fig. 1.2; 39). A member of the Comamonadaceae, within the Betaproteobacteria, has dominated 16S rRNA gene libraries from serpentinite fluids around the world (7, 8, 13, 44, 48). This organism, initially described as *Hydrogenophaga*-like (8, 43, 48), was isolated from the Cedars by Suzuki et al (44) who proposed that it belongs to a

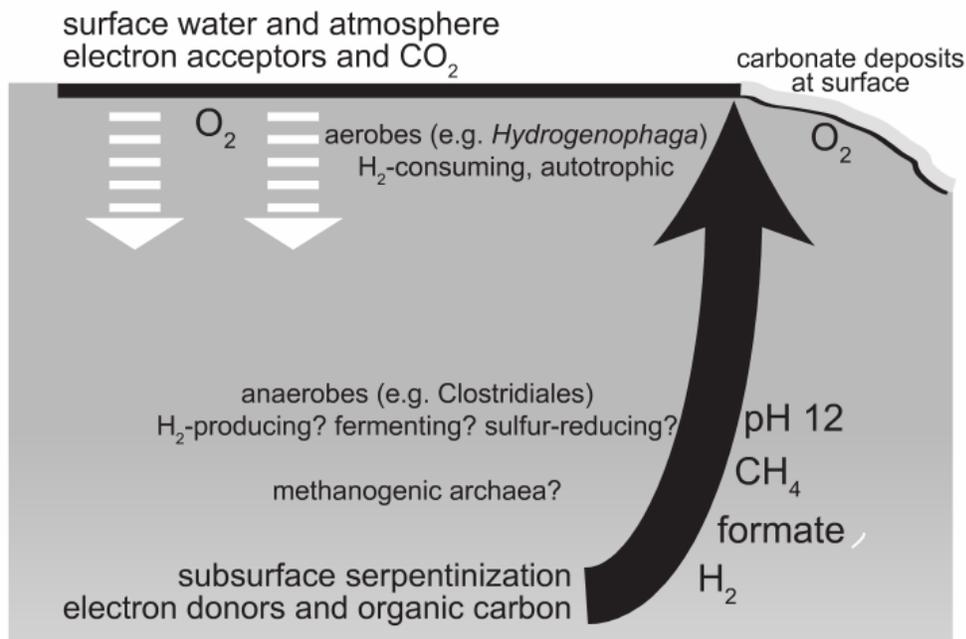


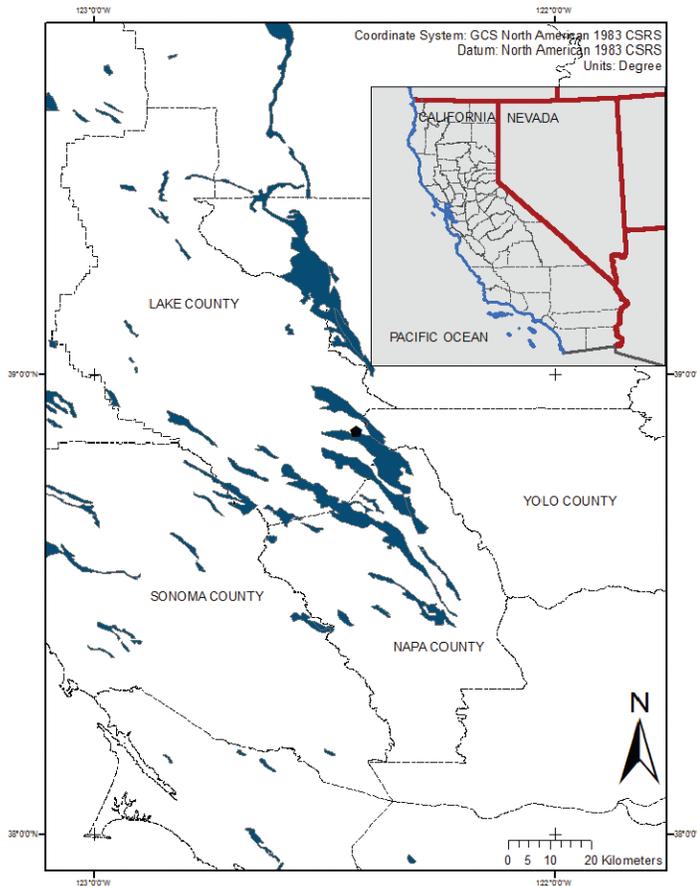
Figure 1.2 – Conceptual model of biogeochemical environment at continental serpentinite sites. Question marks represent yet resolved aspects of the system. Reproduced from Schrenk, *et al.*, 2013.

novel genus, *Serpentinomonas* (44). Betaproteobacteria seen at CVA and the Tablelands have >99% sequence identity across the 16S rRNA gene with the newly proposed *Serpentinomonas* (44). According to metagenomic evidence from the Tablelands (7) and genomic and isolate information from the Cedars (44), this organism is capable of aerobic H₂-oxidation and carbon fixation, via the Calvin-Benson-Bassham (CBB) pathway. Genes for carbon monoxide oxidation have been observed in both metagenomes (7) and genomes (44) and CO-oxidation has taken place in microcosms dominated by this organism (32). These complimentary data from multiple serpentinite sites support that one of the dominant members of microbial communities from serpentinite fluids is an aerobic, H₂-oxidizing Betaproteobacterium with potential for CO₂-fixation and CO-oxidation. Less is known about the Clostridia found in serpentinite samples because they have yet to be isolated. Microcosms from CROMO have yielded Clostridia closely related to *Dethiobacter alkaliphilus*, a facultative autotroph capable of H₂-oxidation and sulfate-reduction (13, 42). Evidence from the Tablelands suggests that the Clostridia from the serpentinite subsurface environment are capable of H₂-production, based on the phylogeny of [FeFe]-hydrogenases found in metagenomes (7).

Coast Range Ophiolite Microbial Observatory

Prior studies of continental serpentinites have sampled surface seeps that represent an interface between end-member fluids and the atmosphere (7, 43, 48) and therefore may not represent the actual conditions encountered deep within the subsurface. Furthermore, these samples may be indicative of life existing in intermediate high-energy zones, which form as the volatile-rich fluids from serpentinite systems mix with oxidized surface waters, as opposed to the highly reducing environments where serpentinitization takes place. The Coast Range Ophiolite (CRO) is a 155-170 million year old ophiolite located in northern California, containing

numerous calcium-hydroxide rich springs, indicating serpentinizing activity below the surface (1) and was chosen as the site for a microbial observatory, created with the expressed purpose of studying life and habitability within the continental serpentinite environment. The Coast Range Microbial Observatory allows direct access and study of the actively serpentinizing subsurface environment. CROMO was established at the UC-Davis McLaughlin Nature Reserve in Lower Lake, CA (Fig 1.3; 9) in August 2011 and consists of three groups of wells located within a one-mile radius of each other: the Core Shed Wells (CSW), Quarry Valley (QV), and the N-wells. The CSW and QV wells were drilled using clean drilling techniques in 2011 to enable



*Figure 1.3 – Map of northern California, depicting the location of CROMO. Serpentinites are highlighted in blue and the hexagon represents the CROMO site at the McLaughlin Natural Reserve in Lower Lake, CA. Reproduced from Cardace *et al.*, 2013.*

subsequent monitoring of the microbial communities and associated geochemistry within the serpentinite subsurface over time (9). CSW consists of five wells, drilled to depths of 9-27 meters. QV consists of three wells, drilled to depths of 15-23 meters. There are two main wells, CSW1.1 and QV1.1, which were drilled to 31 m and 45 m, respectively, in a manner that yielded core material (9). The N-wells, ranging in depth from 14-40 meters, depth from 14-40 meters, represent previously existing wells that were drilled in 1983 for environmental monitoring of mining impacts on groundwater, not with the specific purpose of monitoring microbiology and organic geochemistry (9). The establishment of CROMO allows for the geological, geochemical, and microbiological exploration of serpentinite core material, as well as fluids from the serpentinite environment over time.

One study looking at the formation temperatures of methane through the use of methane isotopologues suggested that the methane at CROMO was not of *in situ* origin, potentially sourced from deeper, higher temperature fluids (49). Meanwhile, methane from the nearby Cedars site showed clear evidence of being microbially generated (49). Microcosm experiments from CROMO well fluids, which resulted in the enrichment of dominant community members from the wells, suggested that the microbial communities are not nutrient limited and favored the addition of organic carbon (acetate) or inorganic sources (13). The main taxon grown in the microcosms were Betaproteobacteria of the family Comamonadaceae, which exhibited 100% sequence identity of the 16S rRNA gene to *Serpentinomonas* strain B1, isolated from The Cedars (13, 44). The other organism grown in the microcosms, particularly those enriched with the sulfur compounds, was *Dethiobacter* of the class Clostridia, which shared >99% sequence identity to clones from CVA (13, 48). These data are consistent with the model of the continental serpentinite subsurface environment hosting Clostridia and Betaproteobacteria (39).

One of the main objectives in establishing the Coast Range Ophiolite Microbial Observatory was to help elucidate microbial communities within the serpentinite subsurface environment, associated both with fluids and minerals (9). The research discussed in this dissertation directly investigates photosynthesis-independent microbial communities within the serpentinite subsurface environment by analyzing sequence data from both fluids and cores from CROMO. By combining 16S rRNA amplicon sequencing, metagenomes, and geochemical analyses to study serpentinite fluids from wells across geochemical gradients, the fluid-associated microbial communities, their metabolic potential, and geochemical drivers were determined. 16S rRNA gene analyses of CROMO core materials allowed for the determination of distinct core-associated microbial communities within the serpentinite subsurface environment. Finally, the use of *in situ* colonization experiments has allowed for further understanding microbe-mineral interactions at CROMO. The research presented here aids in our understanding of the serpentinite subsurface environment and habitability globally.

REFERENCES

REFERENCES

1. Barnes I, O'Neil JR (1971) Calcium-magnesium carbonate solid solutions from Holocene conglomerate cements and travertines in the Coast Range of California. *Geochim Cosmochim Acta* 35:699-718.
2. Barnes I, O'Neil JR, Trescases JJ (1978) Present day serpentinization in New Caledonia, Oman, and Yugoslavia. *Geochim Cosmochim Acta* 42:144-145.
3. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA (2006) Methane- and sulfur-metabolizing communities dominate the Lost City Hydrothermal Field ecosystem. *Appl Environ Microbiol* 72: 6257-6270.
4. Brazelton WJ, Baross JA (2009) Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. *ISME J* 3:1420-1424.
5. Brazelton WJ, et al. (2010) Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proc Natl Acad Sci* 107:1612-1617.
6. Brazelton WJ, Mehta MP, Kelley DS, Baross JA (2011) Physiological differentiation within a single-species biofilm fueled by serpentinization. *mBio* 2: doi:10.1128/mBio.00127-11.
7. Brazelton WJ, Nelson B, Schrenk MO (2012) Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2: doi: 10.3389/fmicb.2011.00268.
8. Brazelton WJ, Morrill PL, Szponar N, Schrenk MO (2013) Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl Environ Microbiol* 79:3906-3916.
9. Cardace D, et al. (2013) Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Scientific Drilling* 16:45-55.
10. Carnevale, DC (2013) Carbon sequestration potential of the Coast Range Ophiolite in California. (Master's thesis). Retrieved from: Open Access Master's Theses. Paper 46.
11. Charlou JL, Donval JP, Fouquet Y, Jean-Baptiste P, Holm N (2002) Geochemistry of high H₂ and CH₄ vent fluids issuing from ultramafic rocks at the Rainbow hydrothermal field (36°14'N, MAR). *Chem Geol* 191:345-359.
12. Chivian D, et al. (2008) Environmental genomics reveals a single-species ecosystem deep within Earth. *Science* 322:275-278.

13. Crespo-Medina M, et al. (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604.
14. Ehlmann BL, Mustard JF, Murchie SL (2010) Geologic setting of serpentine deposits on Mars. *Geophys Res Lett* 37: doi:10.1029/2010GL042596.
15. Frost BR, Beard JS (2007) On silica activity and serpentinization. *J Petrol* 48:1351-1368.
16. Früh-Green GL, et al. (2003) 30,000 years of hydrothermal activity at the Lost City Vent Field. *Science* 301:495-498.
17. Herzog H (2009) Carbon dioxide capture and storage. *In: Economics and Politics of Climate Change* (Oxford University Press, Oxford, UK) pp. 263-283.
18. Hoehler TM (2004) Biological energy requirements as quantitative boundary conditions for life in the subsurface. *Geobiology* 2:205-215.
19. Hoehler TM (2007) An energy balance concept for habitability. *Astrobiology* 7:824-838.
20. Kelemen PB, Matter J (2008) In situ carbonation of peridotite for CO₂ storage. *Proc Natl Aca Sci* 150:17295-17300.
21. Kelley DS, et al. (2001) An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30°N. *Nature* 412:145-149.
22. Kelley DS, et al. (2005) A serpentinite-hosted ecosystem: The Lost City Hydrothermal Field. *Science* 307:1428-1434.
23. Krulwich TA, Sachs G, Padan E (2011) Molecular aspects of bacterial pH sensing and homeostasis *Nature Rev Microbiol* 9:330-343.
24. Lang SQ, et al. (2010) Elevated concentrations of formate, acetate and dissolved organic carbon found at the Lost City hydrothermal field. *Geochim Cosmochim Acta* 74:941-952.
25. Lang SQ, et al. (2012) Microbial utilization of abiogenic carbon and hydrogen in a serpentinite-hosted system. *Geochim Cosmochim Acta* 92:82-99.
26. Longhi J, Knittle E, Holloway JR, Wänke H (1992) The bulk composition and internal structure of Mars. *In Mars* University of Arizona. Press, Tucson) pp. 184–208.
27. Matter J, Kelemen PB (2009) Permanent storage of carbon dioxide in geological reservoirs by mineral carbonation. *Nature Geosciences* 2:837-841.

28. McCollom TM, Seewald JS (2001) A reassessment of the potential for reduction of dissolved CO₂ to hydrocarbons during serpentinization of olivine. *Geochim Cosmochim Acta* 65: 3769-3778.
29. McCollom TM (2007) Geochemical constraints on sources of metabolic energy for chemolithotrophy in ultramafic-hosted deep-sea hydrothermal systems. *Astrobiology* 7: DOI: 10.1089/ast.2006.0119
30. Mitchell AC, Lafreniere MJ, Skidmore ML, & Boyd ES (2013) Influence of bedrock mineral composition on microbial diversity in a subglacial environment. *Geology* G34194.1 DOI:10.1130/G34194.1
31. Morrill PL, et al. (2013) Geochemistry and geobiology of a present-day serpentinization site in California: The Cedars. *Geochim Cosmochim Acta* 109:222-240.
32. Morrill PL, et al. (2014) Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. *Front Microbiol* 5:613.
33. Mumma MJ, et al. (2009) Strong release of methane on Mars in northern summer 2003. *Science* 323:1041-1045.
34. Price PB & Sowers T (2004) Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *PNAS* 101:4631-4636.
35. Proskurowski G, et al. (2008) Abiogenic hydrocarbon production at Lost City Hydrothermal Field. *Science* 319:604-607.
36. Russell MJ (2003) The importance of being alkaline. *Science* 302:580-581
37. Russell MJ, Hall AJ, Martin W (2010) Serpentinization as a source of energy at the origin of life. *Geobiology* 8:355-371.
38. Schrenk MO, Kelley DS, Bolton SA, Baross JA (2004) Low archaeal diversity linked to subsurface geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6: 1086-1095.
39. Schrenk MO, Brazelton WJ, Lang SQ (2013) Serpentinization, carbon and deep life. *Rev Mineral Geochem* 75:575-606.
40. Shock EL (2009) Minerals as energy sources for microorganisms. *Econ Geol* 104: 1235-1248.
41. Sleep NH, et al. (2004) H₂-rich fluids from serpentinization: Geochemical and biotic implications. *PNAS* 101:12818-12823.

42. Sorokin DY, Tourova TP, Musmann M, Muyzer G (2008) *Dethiobacter alkaliphilus* gen. nov. sp. nov. and *Desulfovibrio alkaliphilus* gen. nov. sp. nov.: two novel representatives of reductive sulfur cycle from soda lakes. *Extremophiles* 12: 431-439.
43. Suzuki S, et al. (2013) Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proc Natl Acad Sci USA* 110:15336-15341.
44. Suzuki S, et al. (2014) Physiological and genomic features of highly alkaliphilic hydrogen-utilizing *Betaproteobacteria* from a continental serpentinizing site. *Nat Commun* 5:3900.
45. Szponar N, et al. (2012) Geochemistry of a continental site of serpentinization in the Tablelands Ophiolite, Gros Morne National Park: a Mars analogue. *Icarus* doi:10.1016/j.icarus.2012.07.004.
46. Tiago I, Chung AP, Veríssimo A (2004) Bacterial diversity in a nonsaline alkaline environment: Heterotrophic aerobic populations. *Appl Environ Microbiol* 70:7378-7387.
47. Tiago I, Mendez V, Pires C, Morais PV, Veríssimo A. (2006) *Chimaericella alkaliphila* gen. nov., sp. nov., a Gram-negative alkaliphilic bacterium isolated from a nonsaline alkaline groundwater. *Syst Appl Microbiol* 29:100-108.
48. Tiago I, Veríssimo A (2013) Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environ Microbiol* 15:1687-1706.
49. Wang DT, et al. (2015) Nonequilibrium clumped isotope signals in microbial methane. *Science* 348:428-431.
50. Willems A, et al. (1989) *Hydrogenophaga*, a new genus of hydrogen-oxidizing bacteria that includes *Hydrogenophaga flava* comb. nov. (formerly *Pseudomonas flava*), *Hydrogenophaga palleronii* (formerly *Pseudomonas palleronii*), *Hydrogenophaga pseudoflava* (formerly *Pseudomonas pseudoflava* and “*Pseudomonas caboxydoflava*”) and *Hydrogenophaga taeniospiralis* (formerly *Pseudomonas taeniospiralis*). *Int J Syst Bacteriol* 39:319-333.
51. Zahnle K, Freedman RS, Catling DC (2011) Is there methane on Mars? *Icarus* 212:493-5

CHAPTER 2

Serpentinizing fluids from a subsurface observatory harbor extremely low diversity microbial communities adapted to high pH¹

Abstract

Serpentinization is a widespread geochemical process associated with the aqueous alteration of ultramafic rocks, resulting in abundant fuels (H₂ and CH₄), but potentially challenging conditions for life, including high pH and limited availability of terminal electron acceptors and low concentrations of dissolved inorganic carbon. As a consequence, past studies of serpentinites have reported low cellular abundances and microbial diversity. Establishment of the Coast Range Ophiolite Microbial Observatory well network has allowed, for the first time, a comparison of microbial communities and physicochemical parameters directly within a serpentinization-influenced subsurface aquifer. Samples were collected from seven wells and subjected to a range of analyses, including solute and gas chemistry, microbial diversity by 16S rRNA gene sequencing, and metabolic potential by metagenomics, in an attempt to elucidate what factors drive microbial activities in serpentinite habitats. The strongest geochemical influences on biodiversity in these samples were pH and the concentrations of methane and carbon monoxide. A single operational taxonomic unit (OTU) of Betaproteobacteria and a few OTUs of Clostridia almost exclusively inhabited fluids exhibiting the most serpentinized character. Metagenomes from these extreme samples contained abundances of sequences encoding proteins associated with hydrogen metabolism, carbon monoxide oxidation, and carbon

¹ The work described in this chapter is currently in submission to the ISME Journal for publication: K.I. Twing, D. Cardace, M.D. Kubo, W.J. Brazelton, T.M. Hoehler, T.M. McCollom, and M.O. Schrenk. (In Submission) Serpentinizing fluids from a subsurface observatory harbor extremely low diversity microbial communities adapted to high pH.

fixation. Samples from more moderate pH wells were characterized by higher bacterial diversity and the presence of genes required for the metabolism of methane. These data contribute to a growing body of evidence that serpentinite environments are characterized by limited taxonomic diversity, are fueled by hydrogen and carbon metabolisms, and may significantly contribute to subsurface fluxes of these important elements.

Introduction

The Earth's subsurface is predicted to be an expansive habitat for microorganisms. It is one or more steps removed from photosynthetic sources of carbon and energy and is instead influenced by geological sources (1, 2, 3). However, given the inherent lack of accessibility, direct sampling of subsurface environments has been limited. In continental settings, researchers have used caves (4, 5), mines (6, 7), or springs (8, 9, 10, 11), as windows into the subsurface environment. These features grant access to an otherwise inaccessible environment, but they represent opportunistic sampling at locations where the subsurface environment interacts with the surface. In the present study, wells were drilled directly into serpentinization-influenced aquifers of the Coast Range Ophiolite, a portion of ancient seafloor in northern California, USA, to sample microbial communities in serpentinizing rocks and groundwaters. This observatory represents the first investigation of microbial communities with direct access to the range of conditions in the serpentinizing subsurface (12).

Serpentinization is a widespread geochemical process involving the aqueous alteration of peridotite to serpentine minerals, resulting in an abundance of potential reductants, in the form of hydrogen, methane, and small organic molecules (13, 14). The geochemical process also releases hydroxyl ions, which creates extremely high pH fluids. At high pH, bicarbonate and carbonate are the dominant species of dissolved inorganic carbon (DIC), and the latter can precipitate out

of solution as carbonate minerals when in the presence of divalent cations, such as Ca^{2+} , commonly found in serpentinite fluids. Thus, serpentinitized fluids are characteristically low in DIC, particularly dissolved CO_2 . Compared to the abundance of reductants, in the form of H_2 and CH_4 typically found in these system, there are a lack of corresponding oxidants, probably limiting the potential for microbial metabolism. Thus the subsurface serpentinite environment is characterized by challenges to life such as extreme pH (> 10), limited availability of dissolved carbon, and a lack of potential terminal electron acceptors.

The best-characterized serpentinite-hosted microbial ecosystem to date is the Lost City Hydrothermal Field, located 15 km from the Mid-Atlantic Ridge (15). The tall carbonate chimneys at Lost City are dominated by methane-cycling archaea in the anoxic chimney interiors (16) and by methanotrophic and sulfur-oxidizing bacteria in the chimney exteriors (17). More recently, researchers have started exploring life within continental serpentinite environments by using natural springs, such as the Tablelands Ophiolite in Newfoundland, Canada (8, 9) and The Cedars site in northern California (11, 18), or previously established wells, such as the Cabeço de Vide Aquifer (CVA) in Portugal (19), to gain access to the serpentinite subsurface. In these previous studies of continental serpentinite sites, microbial communities were dominated by the bacterial taxa Betaproteobacteria and Firmicutes (20).

Metagenomic surveys of the Tablelands Ophiolite suggest that subsurface serpentinite communities are dominated by Firmicutes in the deep, anoxic source-waters and microaerophilic H_2 -oxidizing Betaproteobacteria at the shallow, oxic/anoxic interface (8, 9, 20). Microcosm experiments from the Coast Range Ophiolite Microbial Observatory (CROMO), the location of this study, have indicated that Betaproteobacteria closely related to *Hydrogenophaga pseudoflava* and Clostridia (phylum Firmicutes) closely related to *Dethiobacter alkaliphilus*, are

stimulated by small organic molecules that are expected to be available in the serpentinite environment (21). Furthermore, recently published genomes of cultivated isolates of the proposed genus *Serpentinomonas*, which are most closely related to the genus *Hydrogenophaga*, are consistent with a role for these organisms in oxic/anoxic interfaces in serpentinizing systems (18).

This study combines high-throughput 16S rRNA gene sequencing, metagenomic analyses, and geochemical monitoring across a range of physicochemical conditions in order to relate patterns in microbial diversity and metabolic potential to underlying geochemical parameters and processes in the serpentinite subsurface environment. This work improves our understanding of physiological adaptation and ecology in these ubiquitous ecosystems and facilitates our integration of these systems into models of carbon cycling.

Results and Discussion

Sampling Site and Geochemistry

Fluids were collected from seven wells within the Coast Range Ophiolite Microbial Observatory (CROMO), drilled for the purpose of monitoring biogeochemistry and microbial community dynamics with temporal and spatial resolution (12). The wells, ranging in depth from 9-27 meters, are located within a 1.5 km radius and exhibit a wide range of geochemical characteristics. To tease apart which environmental parameters contribute to microbial community composition, geochemical and microbiological data from these seven wells were compared.

The geochemical data associated with the well fluids from August 2012 are summarized in Table 2.1. Samples from wells CSW1.1 and QV1.1 are characterized by extremely high pH (12.2 and 11.5, respectively) and low redox potential (Table 2.1). These wells are depleted in

dissolved inorganic carbon (DIC), containing one to two orders of magnitude less DIC than a well with circumneutral pH, CSW1.4 (Table 2.1). The elevated pH and low DIC at CSW1.1 and QV1.1 are typical of actively serpentinizing environments and are considered to be representative of end-member fluids. CSW1.1 also exhibited higher concentrations of H₂ and organic acids relative to the other wells (Table 2.1). Meanwhile, wells CSW1.4 and QV1.2 exhibited circumneutral pH, low conductivity and higher DIC (Table 2.1), suggesting they contain a significant component of non-serpentinized groundwater. Nevertheless, these wells have elevated concentrations of H₂ and CO comparable to the wells with higher pH, indicating some input from serpentinization. CSW1.2 (pH 9.3), a moderately high pH well, defined here as pH 8-10, contained the highest concentration of methane (1.6 mM; Table 2.1) and appears to represent a mixing zone between serpentinite end-member fluids and surface waters.

Unsurprisingly, many of the environmental parameters of the system were correlated with one another (Table 2.2), as determined by pairwise Pearson's correlation analyses of samples in Table 2.1. pH was negatively correlated with oxidation-reduction potential (ORP), carbon monoxide, dissolved oxygen (DO), and DIC, and it was positively correlated with depth and organic acid concentration (Table 2.2). The concentrations of the organic acids acetate, formate, propionate, and butyrate were positively correlated with one another and they were negatively correlated with ORP (Table 2.2). Carbon monoxide concentration was positively correlated with ORP (i.e. positively correlated with a more positive ORP value) and DO, and it was negatively correlated with conductivity and H₂ concentration (Table 2.2). The concentration of methane was not significantly correlated with any other environmental parameters.

Table 2.1 – Environmental and geochemical parameters associated with samples collected in August 2012.

	CSW1.1	CSW1.2	CSW1.3	CSW1.4	CSW1.5	QV1.1	QV1.2
Depth (mbs)	19.5	19.2	23.2	8.8	27.4	23.0	14.9
Temp (°C)	17.2	18.5	16.9	15.2	16.2	17.9	18.4
pH	12.2	9.3	10.1	7.9	9.7	11.5	7.9
ORP (mV)	-284	-32	-83	-35	-121	-155	-30
DO (mg/L)	0.05	0.41	0.06	1.05	0.03	0.03	0.03
Conductivity (µS/cm)	5200	3710	4500	1560	4220	2068	1655
DIC (µM)	253 ± 8	605 ± 268	172 ± 16	5046 ± 531	545 ± 13	96 ± 2	979 ± 32
Dissolved H ₂ (µM)	0.289 ± 0.004	0.140 ± 0.001	0.283 ± 0.018	0.271 ± 0.013	0.138 ± 0.020	0.075 ± 0.001	0.076 ± 0.009
Dissolved CH ₄ (mM)	0.524 ± 0.132	1.625 ± 0.055	0.969 ± 0.529	0.002 ± 0.0003	1.266 ± 0.032	0.301 ± 0.021	0.303 ± 0.030
Dissolved CO (µM)	0.089 ± 0.002	0.158 ± 0.001	0.115 ± 0.013	0.187 ± 0.005	0.124 ± 0.008	0.142 ± 0.004	0.150 ± 0.008
Acetate (µM)	70.79 ± 1.26	< 1.55	< 1.55	< 3.04	< 3.04	10.20 ± 0.33	< 2.01
Formate (µM)	15.74 ± 0.99	< 1.39	< 1.39	< 1.79	< 1.79	< 1.39	< 2.23
Propionate (µM)	3.49 ± 0.003	< 0.01	< 0.01	< 0.01	< 0.01	0.16 ± 0.05	0.22 ± 0.01
Butyrate (µM)	20.99 ± 0.45	< 1.11	< 1.11	< 2.75	< 2.75	5.97 ± 0.30	< 1.89
Microbial cells (cells/mL)	1.8 × 10 ⁵	6.6 × 10 ⁵	2.3 × 10 ⁵	1.0 × 10 ⁵	3.9 × 10 ⁵	1.0 × 10 ⁶	9.5 × 10 ⁵

Table 2.2 – Significant correlations between environmental parameters according to Pearson’s correlation analysis.

		Correlation	R	p-value
Depth	pH	+	0.79	< 0.001
	DO	-	0.59	0.01
	DIC	-	0.59	0.01
Temp	H ₂	-	0.54	0.02
pH	Butyrate	+	0.77	< 0.001
	Formate	+	0.73	< 0.001
	Propionate	+	0.66	0.004
	Acetate	+	0.64	0.005
	ORP	-	0.87	< 0.001
	CO	-	0.75	< 0.001
	DIC	-	0.58	0.01
	DO	-	0.55	0.02
ORP	CO	+	0.87	< 0.001
	Butyrate	-	0.94	< 0.001
	Acetate	-	0.92	< 0.001
	Propionate	-	0.89	< 0.001
	Formate	-	0.88	0.03
	Conductivity	-	0.63	< 0.001
	H ₂	-	0.53	0.006
DO	DIC	+	0.80	< 0.001
	CO	+	0.65	0.004
Conductivity	H ₂	+	0.77	< 0.001
	Formate	+	0.69	0.002
	Propionate	+	0.67	0.003
	Acetate	+	0.66	0.004
	Butyrate	+	0.59	0.01
	CO	-	0.80	< 0.001
DIC	CO	+	0.57	0.02
H ₂	Formate	+	0.66	0.004
	Propionate	+	0.64	0.005
	Acetate	+	0.61	0.009
	Butyrate	+	0.55	0.02
	CO	-	0.60	0.01
CO	Acetate	-	0.82	< 0.001
	Formate	-	0.81	< 0.001
	Propionate	-	0.81	< 0.001
	Butyrate	-	0.80	< 0.001
Formate	Acetate	+	0.99	< 0.001
	Propionate	+	0.99	< 0.001
	Butyrate	+	0.97	< 0.001
Acetate	Propionate	+	0.99	< 0.001
	Butyrate	+	0.99	< 0.001
Propionate	Butyrate	+	0.98	< 0.001

Bacterial 16S rRNA Gene Diversity and Community Composition

Bacterial diversity was assessed in fluids collected in August 2012 from the seven CROMO wells. Three nearby previously established wells (N08A, N08B, N08C) were also included in this analysis in order to address potential concerns of post-drilling stabilization within the new wells. Environmental sequences of 16S rRNA gene amplicons were obtained with an Illumina MiSeq platform, yielding between 78,000 and 226,000 merged paired-end sequences per sample, for a total of 4,354,377 16S rRNA sequences in this study. These sequences were clustered into 11,454 OTUs at a 97% sequence similarity threshold, and only 71 of these OTUs comprised greater than 1% of the sequences in any of the samples analyzed. All diversity analyses in this study were conducted with OTUs, instead of relying solely on taxonomic annotations, in order to avoid the biases and limitations inherent to database-dependent analyses that are magnified when studying poorly characterized microbial communities. Field replicates of samples were collected and analyzed in parallel and were statistically indistinguishable from one another, as determined by a SIMPROF test of the whole bacterial community similarities among all samples (Fig. 2.1). The community compositions of samples from different wells were clearly distinct from each other (ANOSIM, $R = 0.96$, $p\text{-value} = 0.001$). Alpha diversity of the samples, as measured by the inverse Simpson diversity index, decreased with increasing pH (Fig. 2.2).

The wells with the highest pH, CSW1.1 and QV1.1, exhibited extremely low diversity, containing almost exclusively Betaproteobacteria and Firmicutes (Fig. 2.2), which is consistent with previous reports of communities in high-pH serpentinite fluids (9, 11, 19). CSW1.1 was dominated by a single betaproteobacterial OTU (OTU001), classified as a member of the family Comamonadaceae and 100% identical over 250 bp of the 16S rRNA gene's V4 region

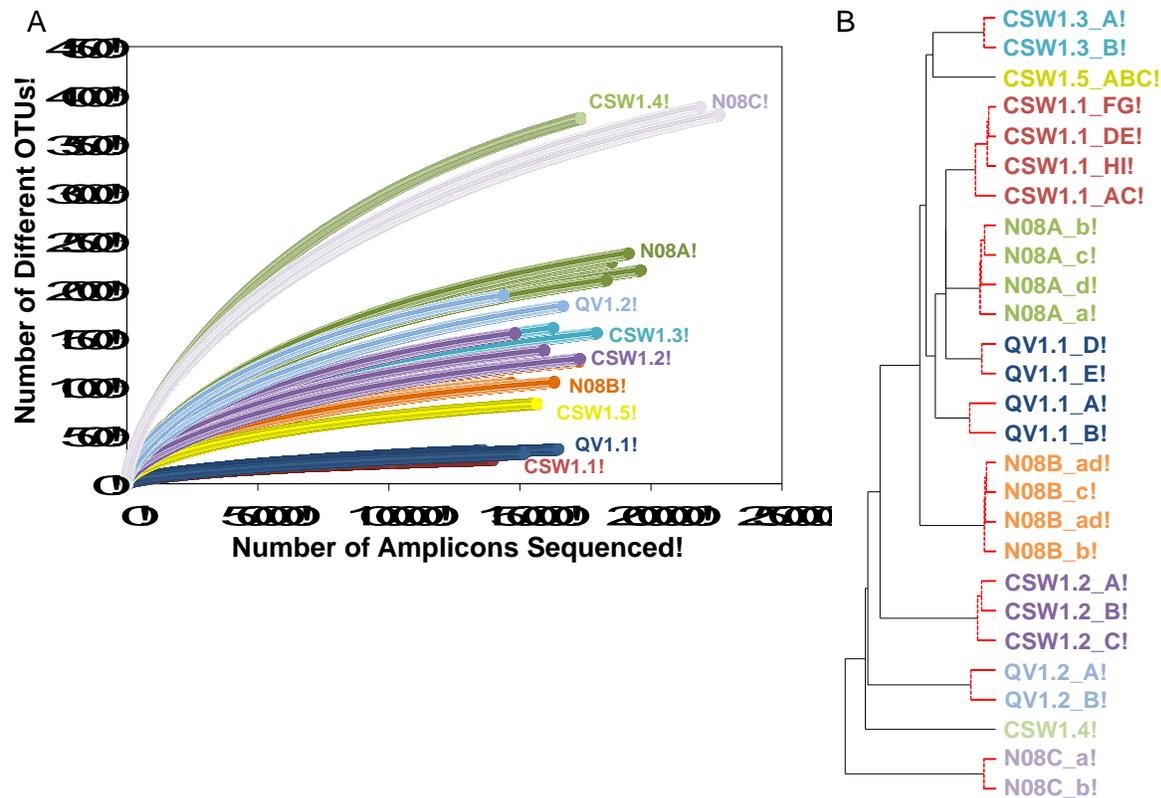
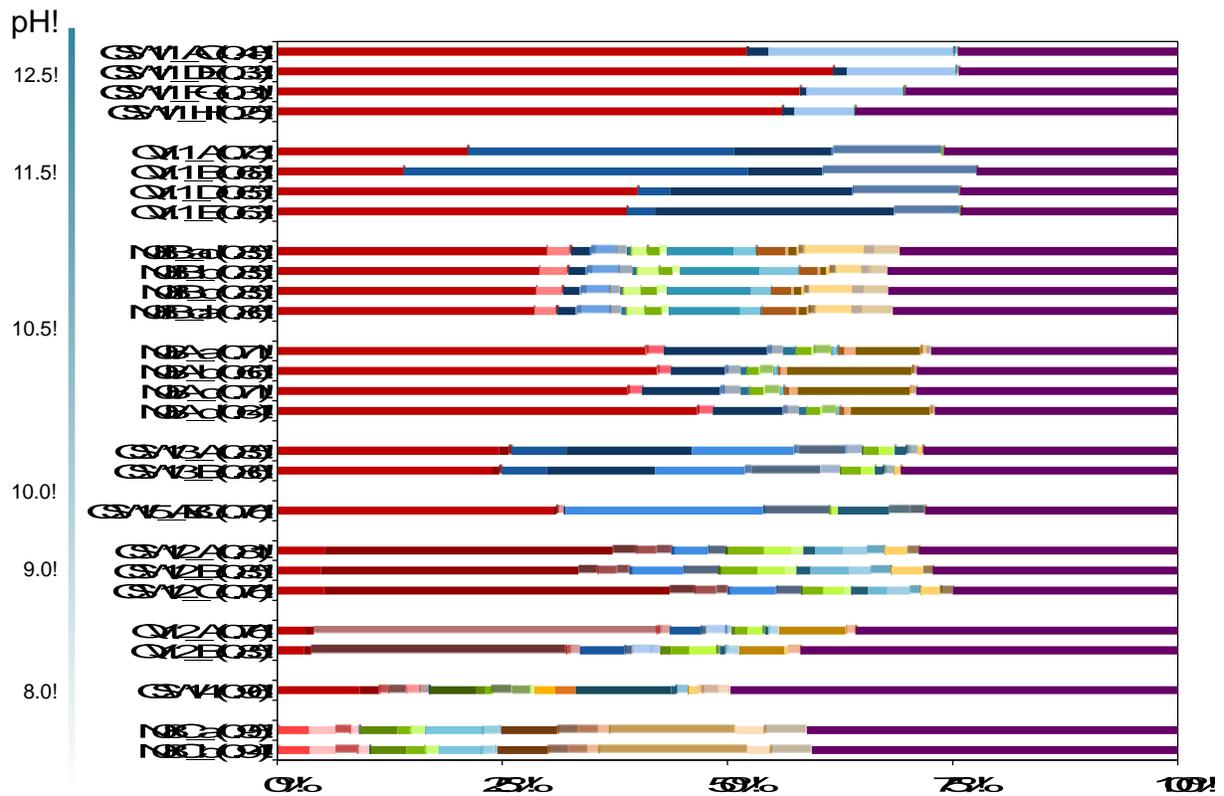


Figure 2.1 – Microbial community structure. A) Rarefaction analysis of 16S rRNA amplicon sequences. Multiple samples from the same well represent field replicates. B) Community similarity dendrogram calculated from Bray-Curtis index. Samples connected by red lines are not distinguishable from one another by a SIMPROF test and ANOSIM analysis indicated that there is a significant difference in community composition between wells ($R = 0.9$, $p\text{-value} < 0.05$)

(calculated with MatGat) to strain B1 from the proposed genus *Serpentinomonas* isolated from The Cedars serpentinite site (18; Table 2.3). The second most abundant OTU in CSW1.1 (OTU018) comprised $12.5 \pm 5.8\%$ of the sequences from that well, was classified as Thermoanaerobacterales SRB-2, and exhibited 99% sequence identity to a Clostridia clone from a well in Cabeço de Vide (CVA) in Portugal (19; Table 2.3). The third OTU detected in CSW1.1 (OTU002) accounted for only 1% of the sequences from that well, was classified as *Dethiobacter*, and shared 100% sequence identity to a clone from CVA (19; Table 2.3), a microcosm isolate from CROMO (21), and a clone from the deep groundwater site at The Cedars



OTU (97% sequence similarity):!

- | | | |
|----------------------|----------------------|-----------------------|
| ■ 001-Chloroflexi | ■ 001-Chloroflexi | ■ 003-Axostylax |
| ■ 002-Firmicutes | ■ 003-Firmicutes | ■ 007-Lactobacilli |
| ■ 003-Chloroflexi | ■ 004-Nitrospirae | ■ 008-Chloroflexi |
| ■ 004-Firmicutes | ■ 005-Nitrospirae | ■ 010-Verrucomicrobia |
| ■ 0092-Firmicutes | ■ 002-Firmicutes | ■ 003-SPI2 |
| ■ 002-Lactobacilli | ■ 005-Lactobacilli | ■ 003-SPI2 |
| ■ 007-Citricellulose | ■ 007-Erythrobacter | ■ 005-SPI2 |
| ■ 001-Citricellulose | ■ 008-Erythrobacter | ■ 001-SPI2 |
| ■ 005-Lactobacilli | ■ 016-Citricellulose | ■ 004-Spirochetes |
| ■ 005-Nitrospirae | ■ 001-Firmicutes | ■ 001-Firmicutes |
| ■ 007-Firmicutes | ■ 010-Firmicutes | ■ 005-Spirochetes |
| ■ 005-Nitrospirae | ■ 005-Spirochetes | ■ 001-Firmicutes |
| ■ 007-Erythrobacter | ■ 003-Lactobacilli | ■ 005-Lactobacilli |
| ■ 005-Lactobacilli | ■ 004-Spirochetes | ■ 003-Nitrospirae |
| ■ 003-Firmicutes | ■ 006-Axostylax | ■ 005-Firmicutes |
| ■ 009-Firmicutes | ■ 007-Lactobacilli | ■ 005-Nitrospirae |
| ■ 012-Axostylax | ■ 001-Firmicutes | ■ 017-Axostylax |
| ■ 016-Axostylax | ■ 005-Spirochetes | ■ 005-Chloroflexi |
| ■ 007-Chloroflexi | ■ 006-Firmicutes | ■ 005-Axostylax |
| ■ 001-Nitrospirae | ■ 001-Spirochetes | ■ 003-Lactobacilli |
| ■ 009-Spirochetes | ■ 005-Axostylax | ■ 002-Chloroflexi |
| ■ 001-Nitrospirae | ■ 005-Axostylax | ■ 002-Firmicutes |
| ■ 002-Firmicutes | ■ 021-Nitrospirae | ■ 007-Firmicutes |
| ■ 002-Firmicutes | ■ 001 | |

Figure 2.2 – Inter-well community composition comparison from August 2012. OTUs were formed at the 97% similarity level in mothur using the average-neighbor algorithm. Bars indicate individual OTUs and are color-coded by Phylum/Class: β -proteobacteria (red), Firmicutes (blue), α -proteobacteria (green), δ -proteobacteria (yellow), γ -proteobacteria (teal), and other classes of bacteria (brown). Other (purple bars) represents any OTUs making up < 1% of all samples in the dataset. Samples are organized by pH (ANOSIM: $R = 0.66$, p -value < 0.05). The numbers in parentheses beside the sample names represent the inverse-Simpson diversity index.

(11). The remaining $29.0 \pm 5.3\%$ of the CSW1.1 microbial community was made up of rare species, defined as OTUs comprising less than 1% of the total sequences in any sample. QV1.1 was dominated by three Clostridia OTUs (OTU003, OTU002, and OTU007) that together account for $47.1 \pm 13.4\%$ of the bacterial community (Fig. 2.2). Both OTU003 (classified as Thermoanaerobacterales SRB-2) and OTU002 (classified as *Dethiobacter*) exhibited 99-100% sequence identity to clones from CVA in Portugal (19; Table 2.3). The same single betaproteobacterial OTU001 made up $28.5 \pm 12.9\%$ of the QV1.1 community, and the remaining bacterial taxa were rare, and accounted for $24.3 \pm 1.6\%$ of the community (Fig. 2.2). A time-series analysis of samples collected from QV1.1 over the course of a year indicate that community composition within wells is relatively constant over time, with no significant difference between time points ($R=0.2$, $p\text{-value}=0.902$).

While Betaproteobacteria made up a large proportion of all samples above neutral pH, the diversity and composition of the Betaproteobacteria shifted with pH (Fig. 2.2). As expressed above, OTU001 made up $42.7 \pm 17.6\%$ of the extremely high pH wells. However, in samples with $\text{pH} \leq 10$, OTU001 was replaced by OTU008 (classified as *Azonexus hydrophilus*; Table 2.3), OTU005 (classified as Methylophilaceae) and OTU004 (classified as Comamonadaceae and 100% identical to *Alicyclophilus denitrificans*; Table 2.3) as the dominant taxa.

Clostridia, which accounted for up to 64% of the bacteria in the highest pH fluids, were also found in the moderately high pH wells. *Dethiobacter* OTUs made up $13.9 \pm 10.1\%$ of samples with a pH 9.5-11.0. Erysipelotrichia (another class of the phylum Firmicutes), which was enriched in subsurface conditions at the Tablelands Ophiolite (9), made up 8% and 2% of CSW1.5 (pH 9.7) and CSW1.2 (pH 9.3), respectively, but was not detected in any other CROMO samples. OTUs classified as Thermoanaerobacterales SRB-2 were detected in N08A

Table 2.3 – Summary of significant correlations (p -value < 0.05) between top OTUs (making up > 25% of the sample in which they are most abundant) and environmental parameters.

OTU	Variable	R	Corr*	Strength	Max Sample	Max Abundance (% of sample)	Class	Order	Family	Closest Relative (NCBI accession number)	% Identity [†]
OTU001	Butyrate	0.88	+	S	CSW1.1	61.7	Betaproteobacteria	Burkholderiales	Comamonadaceae	<i>Serpentinomonas</i> B1 (AP014569.1) ¹	100%
	pH	0.86	+	S							
	Acetate	0.86	+	S							
	Propionate	0.83	+	S							
	Formate	0.82	+	S							
	Conductivity	0.63	+	M							
	Hydrogen	0.54	+	M							
	ORP	0.91	-	VS							
CO	0.86	-	S								
OTU004	Methane	0.81	+	S	CSW1.2	38.4	Betaproteobacteria	Burkholderiales	Comamonadaceae	<i>Alicyclophilus denitrificans</i> (NR_074585.1) ²	100%
	DO	0.53	+	M							
OTU003	Depth	0.54	+	Mo	QV1.1	38.2	Clostridia	Thermoanaerobacterales	SRB2	<i>CVCloAm3Ph15</i> (AM778006) ³	99.6%
OTU008	pH	0.61	-	M	QV1.2	38.0	Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	<i>Azonexusv Hydrophilus</i> (EF158391.1) ⁴	100%
OTU002	Depth	0.67	+	M	QV1.1	26.5	Clostridia	Clostridiales	Syntrophomonadaceae	<i>CVCloAm2Ph135</i> (AM777954) ³	100%
OTU006	Methane	0.60	+	M	CSW1.5	21.9	Clostridia	Clostridiales	Syntrophomonadaceae	<i>CVCloAm2Ph135</i> (AM777954) ³	98.2%
OTU018	Formate	0.91	+	VS	CSW1.1	20.6	Clostridia	Thermoanaerobacterales		<i>CVCloAm3Ph98</i> (AM778028) ³	99.1%
	Propionate	0.91	+	VS							
	Acetate	0.90	+	VS							
	Butyrate	0.88	+	S							
	Conductivity	0.64	+	M							
	Hydrogen	0.60	+	M							
	pH	0.60	+	M							
	ORP	0.85	-	S							
CO	0.74	-	S								

Correlation relationship is defined by correlation coefficient: > 0.9 = very strong (VS), 0.7-0.9 = strong (S), 0.5-0.7 = moderate (M)

This is a subset of the data used to make the correlation network (Fig. 2.3)

* Corr = correlation relationship

[†]Percent identity, as determined by MatGat (29)

Closest relative references: ¹Suzuki *et al*, 2014; ²Oosterkamp *et al.*, 2011; ³Tiago and Veríssimo, 2013; ⁴Chou *et al.*, 2008

and QV1.2, as well as CSW1.1 and QV1.1. No Firmicutes OTUs were detected in the wells with pH less than 9 (Fig. 2.2). In addition to Betaproteobacteria and Firmicutes, the moderately high pH wells also contained Bacteroidetes as well as Alpha-, Delta-, and Gammaproteobacteria (Fig. 2.2). The circumneutral wells contained a greater complement of rare taxa and many taxa that were not present in the high pH wells (Fig. 2.2). The main taxa found within the highest pH samples, predominately Betaproteobacteria and Clostridia, are consistent with those found at other serpentinite sites (9, 11, 19). These bacterial 16S rRNA gene data suggest a core microbial community found within the serpentinite-influenced fluids, that shifts as the pH and serpentinite end-member water decreases.

No archaea were detected in any of the 16S rRNA amplicon libraries, which were created with the universal primers targeting the V4 region of the 16S rRNA gene used by the DOE Joint Genome Institute (22). To further investigate the potential presence of archaea in CROMO fluids, the relative abundance of archaea was also assessed by counting archaeal sequences in the metagenomic datasets (description of metagenomic data below). The number of metagenomic sequences classified as archaea by MG-RAST (23) did not exceed 1% of the total sequences in any sample (Table 2.4). Furthermore, none of these archaeal metagenomic reads included a 16S rRNA gene. These combined data indicate that archaea, if present, are not major contributors of the serpentinite subsurface microbiome at CROMO, and it is therefore not surprising that archaea

Table 2.4 – Microbial communities at CROMO are dominated by bacteria. Relative abundance metagenomic sequence reads assigned to different domains of life via the M5NR database in MG-RAST (Meyer *et al.*, 2008).

Domain	CSW1.1	QV1.1	CSW1.3	QV1.2
Bacteria	98.3	97.9	98.8	99.1
Archaea	0.2	1.1	0.7	0.2
Eukaryotes	1.2	0.8	0.3	0.5
Viruses	0.1	0.0	0.0	0.1
Unassigned	0.2	0.1	0.1	0.1

were not detected via amplicon sequencing with the universal primers.

Biogeochemical Relationships

One of the main goals of this study was to identify the geochemical drivers of microbial community composition within the serpentinite subsurface environment. A combination of pH, carbon monoxide, and CH₄ concentrations could explain 83% of bacterial community composition variability across wells (as determined by BEST in Primer-6; 24). Therefore, we visualized pairwise Pearson correlations among these three environmental parameters and the relative abundances of all associated OTUs with a correlation network diagram (Fig. 2.3).

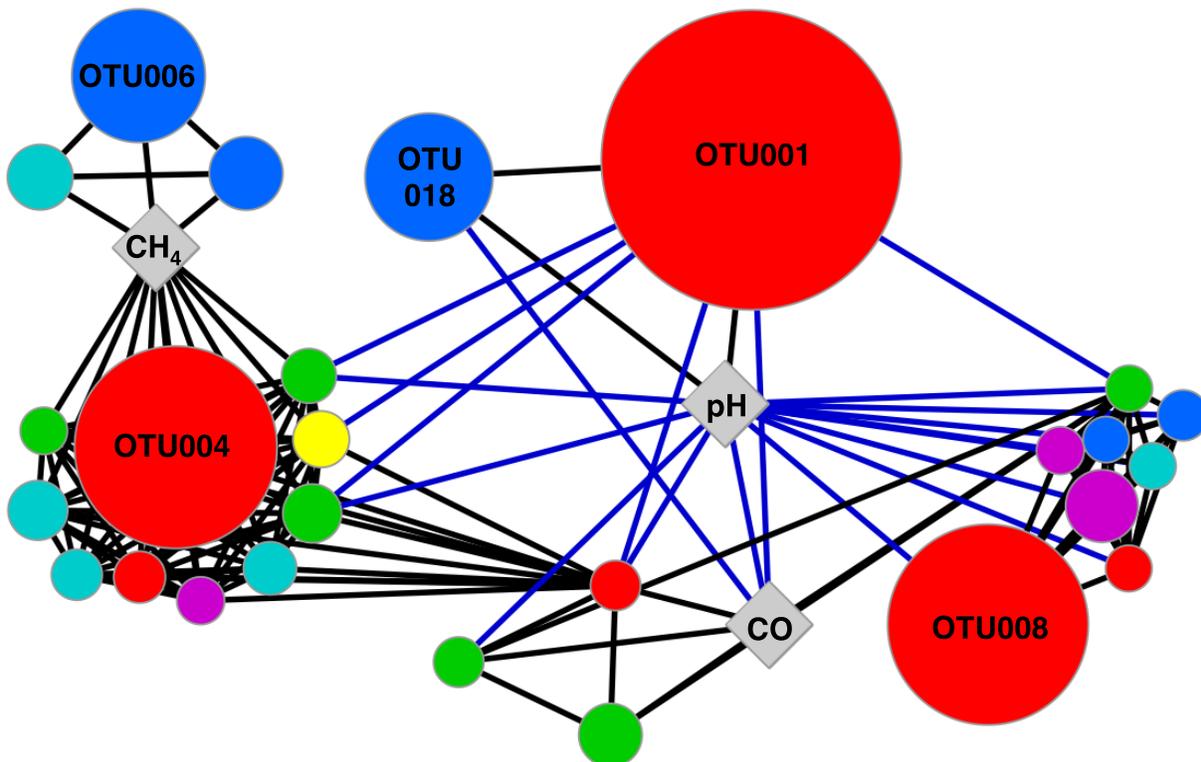


Figure 2.3 – Network diagram of significant correlations between OTUs and environmental variables identified in BEST analysis as accounting for 83% of community composition (p-value = 0.001). OTU node size is relative to the maximum abundance of the OTU across the samples. Node color represents the taxonomic assignment of the OTU at the Phylum/Class level: β-proteobacteria (red), Firmicutes (blue), α-proteobacteria (green), γ-proteobacteria (teal), Actinobacteria (yellow), and Bacteroidetes (pink). Nodes represent OTUs with a relative abundance > 1% of any sample, while OTUs making up > 10 % of any sample are labeled with OTU IDs. Positive and negative correlations are represented with black and blue lines, respectively.

pH was positively correlated with the two OTUs most abundant in CSW1.1 and negatively correlated with OTU008, classified as *Azonexus hydrophilus* (Table 2.3) and the dominant betaproteobacterial OTU in wells with a pH below 10 (Fig. 2.3). The top OTUs from CSW1.1, OTU001 (classified as Comamonadaceae) and OTU018 (classified as Thermoanaerobacteriales SRB-2), were negatively correlated with carbon monoxide concentration. Except for the betaproteobacterial OTU033, all other OTUs that positively correlate with carbon monoxide concentration belong to the alphaproteobacterial order Sphingomonadaceae (Fig. 2.3). Among the OTUs positively correlated with the abundance of

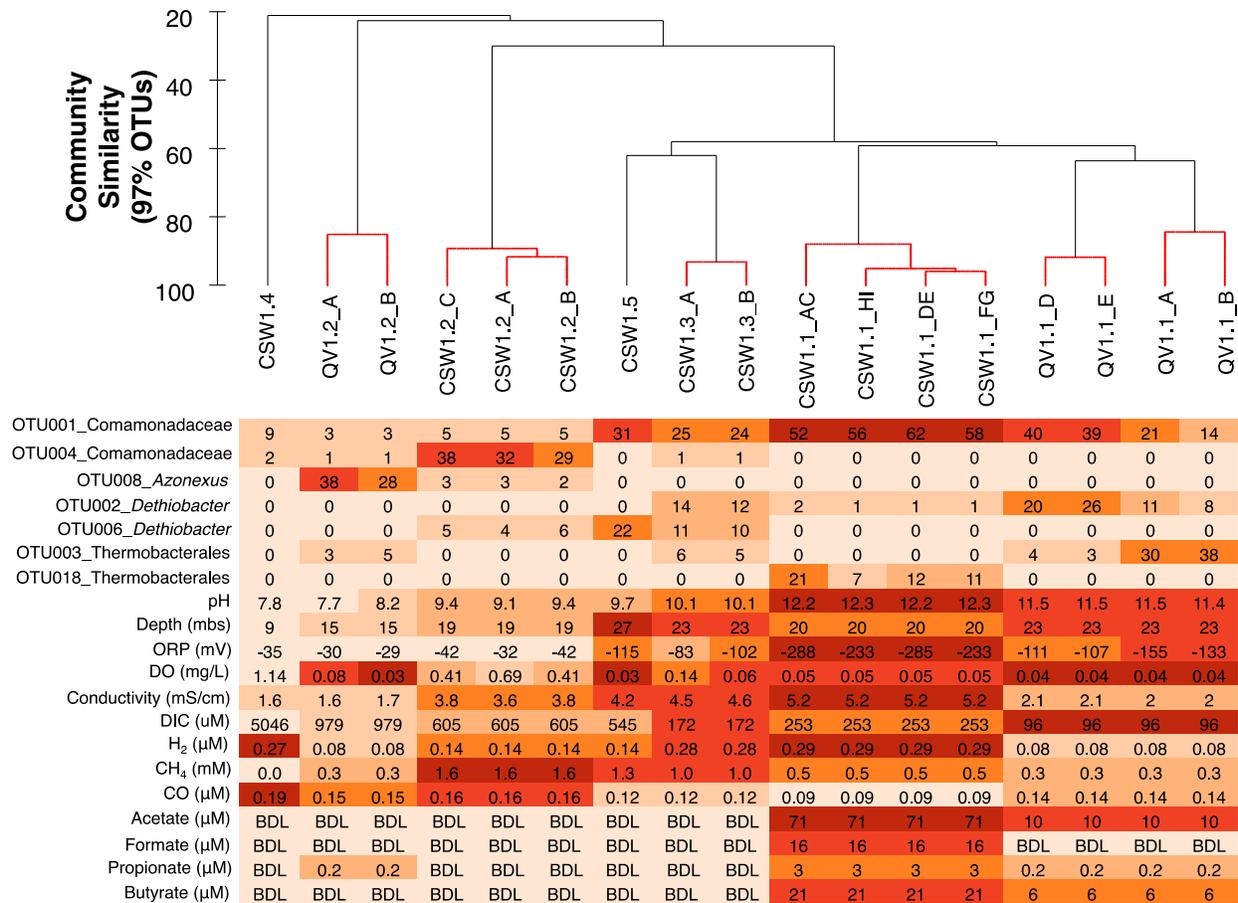


Figure 2.4 – Heatmap of most abundant OTUs and geochemical parameters across all samples. Dendrogram at the top represents community similarity between samples and red lines indicate no statistical difference between field replicates, as determined by SIMPROF (38).

methane were Betaproteobacteria OTU004, most abundant in CSW1.2, and Clostridia OTU006, most abundant in CSW1.5 (Table 2.3, Fig. 2.3). Additionally, five Gammaproteobacteria OTUs.

The most abundant OTUs in CSW1.1, OTU001 (classified as Comamonadaceae) and OTU018 (classified as Thermoanaerobacterales SRB-2), were also positively correlated with conductivity and the concentrations of organic acids and H₂, and were negatively correlated with ORP (Table 2.3). Two of the most abundant Clostridia OTUs (OTU002 and OTU003), both dominant in QV1.1, were significantly correlated only with well depth (Table 2.3). While Table 2.3 and the discussion above denote the sample in which each OTU was most abundant, it should be mentioned that many of those top OTUs were found in multiple samples, though at lower abundances (Fig. 2.4).

Metabolic Potential

To elucidate whether microbes within the serpentinite subsurface environment are capable of metabolizing the geochemical products of serpentinization, specifically hydrogen, methane, and carbon monoxide, assembled metagenomes from four of the wells were searched for sequences predicted to encode proteins diagnostic of targeted metabolic processes. The protein targets were: [Fe-Fe]-hydrogenase and [Ni-Fe]-hydrogenase for hydrogen metabolism, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo) for autotrophic CO₂ fixation via the Calvin-Benson-Bassham cycle, carbon monoxide dehydrogenase specific for aerobic (*coxL*) and anaerobic (*cooS*) carbon monoxide oxidation, particulate methane monooxygenase (*pmoA*) and methanol dehydrogenase (*mxoF*) for bacterial aerobic methane oxidation, and methyl coenzyme-M reductase (*mcrA*) for archaeal methanotrophy or methanogenesis. The numbers of matching proteins were normalized to total kilobases per metagenome and are reported as normalized

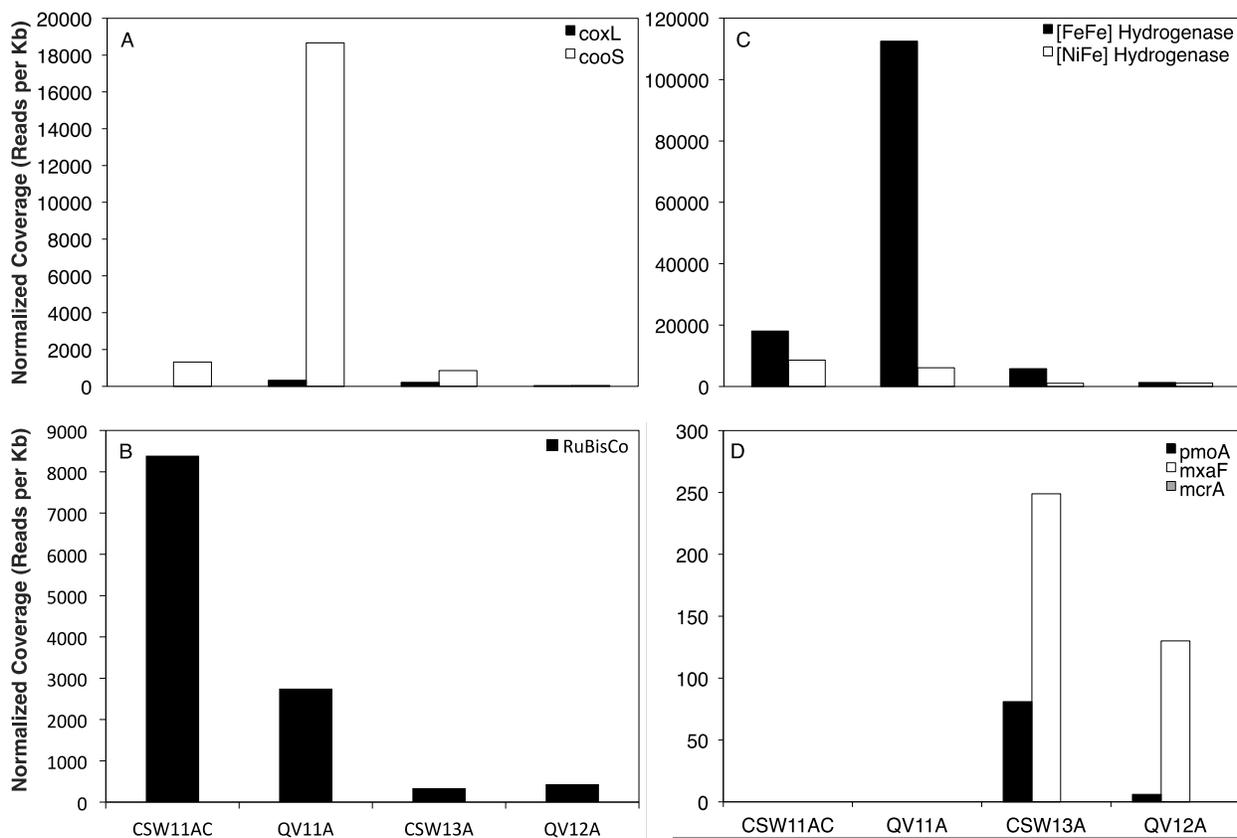


Figure 2.5 – Abundance of protein-encoding genes in metagenomes. Numbers of matching proteins are normalized to Mb per metagenome. Proteins searched for include *coxL* (A), *cooS* (A), RuBisCo (B), Fe-Fe hydrogenase (C), Ni-Fe hydrogenase (C), *pmoA* (D), *mxoF* (D), and *mcrA*(D).

coverage (Fig. 2.5). The only significant correlation between the abundance of protein targets and any of the geochemical parameters was a positive correlation between RuBisCo and the concentration of the organic acid butyrate ($R=0.99$, $p\text{-value}=0.002$). OTUs with significant correlations to gene targets are summarized in Table 2.5. The high pH of the serpentinite environment limits the availability of CO_2 , potentially leading to the use of carbon monoxide as an inorganic carbon source in carbon fixation. To assess carbon monoxide metabolism at CROMO, two forms of carbon monoxide dehydrogenase were searched for in the metagenomes: *coxL* and *cooS* were used to identify aerobic and anaerobic carbon monoxide oxidation,

Table 2.5 – Summary of significant correlations between gene hits and OTUs.

Gene	OTU	Max Sample	Max Abundance	Phylum	Class	Order	Family	Genus
<i>cooS</i>	Otu007	QV1.1B	17.1	Firmicutes	Clostridia			
<i>pmoA</i>	Otu006	CSW1.3A	11.3	Firmicutes	Clostridia	Clostridiales	Syntrophomonadaceae	Dethiobacter
<i>pmoA</i>	Otu017	CSW1.3B	7.61	Firmicutes	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelothrix
<i>pmoA</i>	Otu083	CSW1.3B	2.06	Firmicutes	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelothrix
<i>pmoA</i>	Otu186	CSW1.3B	1.04	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Arenimonas

Table 2.6 – Bioenergetic calculations for CROMO groundwaters: total energy available (Joules of energy per kg of modeled CROMO groundwater) based on the limiting reactant for various metabolic reactions and overall Gibbs Energy (kJ per electron transferred) for the reaction.

Metabolic Reaction	Limiting Reactant	Energy Available (Joules of energy/kg water)				Gibbs Energy (kJ/electron transferred)				
		QV1.2	CSW1.3	QV1.1	CSW1.1	QV1.2	CSW1.3	QV1.1	CSW1.1	
Aerobic hydrogen oxidation	$\text{H}_2(\text{aq}) + 0.5\text{O}_2(\text{aq}) = \text{H}_2\text{O}(\text{l})$	H_2	1.6×10^{-2}	6.0×10^{-2}	1.6×10^{-2}	6.1×10^{-2}	-100	-110	-100	-110
Aerobic carbon monoxide oxidation	$\text{CO}(\text{aq}) + 0.5\text{O}_2(\text{aq}) = \text{CO}_2(\text{aq})$	CO	3.6×10^{-2}	3.0×10^{-2}	4.0×10^{-2}	2.5×10^{-2}	-120	-130	-140	-140
Aerobic methane oxidation	$\text{CH}_4(\text{aq}) + 2\text{O}_2(\text{aq}) = \text{CO}_2(\text{aq}) + 2\text{H}_2\text{O}(\text{l})$	O_2	3.7×10^{-1}	7.7×10^{-1}	3.9×10^{-1}	6.6×10^{-1}	-97	-100	-100	-100
Anaerobic oxidation of methane	$\text{CH}_4(\text{aq}) + \text{SO}_4^{2-} = \text{HCO}_3^- + \text{HS}^-(\text{aq}) + \text{H}_2\text{O}(\text{l})$	SO_4^{2-}	1.0×10^{-4}	1.1×10^{-3}	1.0×10^{-4}	2.6×10^{-3}	-8.1	-8.4	-8.1	-11
Hydrogenotrophic methanogenesis	$\text{CO}_2(\text{aq}) + 4\text{H}_2(\text{aq}) = \text{CH}_4(\text{aq}) + 2\text{H}_2\text{O}(\text{l})$	H_2	4.0×10^{-4}	1.6×10^{-3}	1.0×10^{-4}	2.0×10^{-4}	-4.5	-5.0	-1.6	-2.4
Hydrogenotrophic methanogenesis	$\text{CO}(\text{aq}) + 3\text{H}_2(\text{aq}) = \text{CH}_4(\text{aq}) + \text{H}_2\text{O}(\text{l})$	H_2	1.5×10^{-3}	7.1×10^{-3}	1.7×10^{-3}	7.3×10^{-3}	-13	-15	-14	-15
Methanogenesis by carbon monoxide disproportionation	$4\text{CO}(\text{aq}) + 2\text{H}_2\text{O}(\text{l}) = \text{CH}_4(\text{aq}) + 3\text{CO}_2(\text{aq})$	CO	6.7×10^{-3}	6.7×10^{-3}	1.0×10^{-2}	6.6×10^{-3}	-33	-41	-50	-52

respectively. Both forms of the gene were most abundant in the QV1.1A metagenome (CO = 0.142 μ M; Table 2.1) with the anaerobic form of the gene having an abundance more than fifty-times that of the aerobic form (Fig. 3A). The *coxL* was detected in all the metagenomes, except for CSW1.1AC at low abundance. There is evidence of aerobic carbon monoxide oxidation at other sites of serpentinization, including the presence of *coxL* in the Tablelands metagenome (7) and the detection of carbon monoxide oxidation during microcosm experiments with fluids from the Tablelands (23). While *coxL* has been detected in one *Serpentinomonas* genome (strain A1; 16), it was not seen in *Serpentinomonas* strain B1 (18), the strain that shares 100% sequence identity to OTU001 in this study.

Sequences encoding the RuBisCo enzyme of the Calvin-Benson-Bassham cycle were detected in all four metagenomes, but it was most abundant in the CSW1.1 metagenome (Fig. 2.5B). The RuBisCo gene has been detected in the serpentinite environment previously and showed sequence similarity to *Betaproteobacteria* (8, 19). Furthermore, the genomes of all three *Serpentinomonas* strains contain the genes encoding for RuBisCo and other genes required for CO₂ fixation (18). Hydrogenase sequences were present in all four metagenomes, and the abundance of both hydrogenases was highest in the extreme-pH wells compared to moderate-pH wells (Fig. 2.5C). The [Fe-Fe]-hydrogenase gene was most abundant in the QV1.1 metagenome with a six-fold increase in gene hits compared to the CSW1.1 metagenome (Fig. 2.5C). An abundance of [Fe-Fe]-hydrogenases closely related to Clostridia was also detected in the Tablelands metagenome (8). The [Ni-Fe]-hydrogenase gene was more abundant in the CSW1.1 metagenome, and [Ni-Fe]-hydrogenase sequences from *Betaproteobacteria* were present in the Tablelands metagenome (8) and identified in the *Serpentinomonas* genomes (18).

We searched for *pmoA* and *mxoF*, the first and second proteins in the methane oxidation pathway of aerobic methanotrophs, within the CROMO metagenomes. Both *pmoA* and *mxoF* were detected in the moderate pH wells, with a higher abundance of *mxoF* hits (Fig. 2.5D), which is consistent with previous findings that *pmoA* is absent in some methanotrophic bacteria (26). Neither *pmoA* nor *mxoF* was detected in the CSW1.1 or QV1.1 metagenomes (Fig. 2.5D). No *mcrA* sequences were detected in any of the metagenomes, which is not surprising given the low abundance of archaea in the metagenomes (Table 2.4). The lack of genomic evidence for the presence of methanogens, as well as the energetic unfavorability of methanogenesis to occur in the fluids at CROMO, points to a non-*in situ* source of the methane at CROMO. This interpretation is consistent with methane isotopologue data that may suggest a higher-temperature (deeper) origin of methane at CROMO (27).

To assess the chemical energy available to support metabolic activity in the system, thermodynamic models were used to estimate the molar Gibbs energy change and potential total energy yield for several possible metabolic reactions in fluids from CSW1.1, CSW1.3, QV1.1, and QV1.2 (Table 2.6). The calculations focused in particular on chemolithoautotrophic metabolisms that might be supported by products of serpentinization. Consistent with metagenomic data, the calculations suggest that the most energetically favorable metabolisms across all wells are hydrogen oxidation, aerobic carbon monoxide oxidation, and aerobic methane oxidation (Table 2.6). Although anaerobic methanotrophy and methanogenesis are energetically favorable, the energy yields from these reactions may be too low to support metabolic activities (Table 2.6). The small amount of energy available from these reactions may explain the absence of archaea and the *mcrA* gene in our samples.

These metagenomic results indicate metabolic differences among the microbial communities at CROMO. CSW1.1 contained the greatest abundance of RuBisCo (Fig. 2.5B) and was dominated by *Serpentinomas* strain B1 (OTU001), whose genome contains a diversity of carbon utilizing genes. Given the limited carbon availability in a serpentinite environment, a diversity of strategies for obtaining carbon can allow for a more opportunistic life style. Furthermore, it has been proposed that the Betaproteobacteria within the serpentinite environment inhabit mixing zones between the end-member serpentinite fluids and oxic surface conditions (8, 20). Therefore, microorganisms in this niche may be able to access chemicals provided from meteoric input. Conversely, QV1.1 contained the most protein sequences associated with the oxidation of carbon monoxide and hydrogen (which are likely to be geochemical products of serpentinization) as well as the greatest abundance of Clostridia, organisms thought to be inhabitants of end-member serpentinite fluids (8, 11, 20). The metagenomes from the more moderate pH wells, CSW1.3 (pH 10.1) and QV1.2 (pH 7.9), contained more sequences associated with the metabolism of methane, which could have to do with the less challenging nature of these fluids and access to oxidants provided by the mixing with oxic surface waters.

Conclusions

Despite the abundance of potential energy sources (i.e. H₂ and CH₄) within the serpentinite environment, a lack of electron acceptors, as well as extreme pH and depletion of DIC, makes it a challenging habitat for life. While point-source studies have noted that the serpentinite environment is generally dominated by Betaproteobacteria and Firmicutes (9, 11, 19, 20), the capability to sample across a gradient of geochemical parameters at CROMO allowed us to assess how these groups and others are distributed with respect to the measured environmental

conditions. The wells considered to be most representative of the serpentinite end-member fluids, based on pH, ORP, and DIC concentrations, exhibit extremely low diversity. Roughly 75% of the microbial community in CSW1.1 (pH 12.2) and QV1.1 (pH 11.5) were made up of two and four OTUs, respectively, all within the Betaproteobacteria and Clostridia (Fig. 2.2). Microbial communities within the lower pH wells were still dominated by Betaproteobacteria and Clostridia, but different OTUs within these classes. The remarkable similarities in community composition, including 100% sequence identity of the 16S rRNA gene for many of the most abundant OTUs, between samples obtained from the cleanly drilled CROMO wells and the near-surface samples collected from the Cedars (11) and the Tablelands Ophiolite (9) support two main conclusions: 1) Clostridia are dominant inhabitants of the serpentinite subsurface environment and 2) Betaproteobacteria live at oxic/anoxic interfaces associated with sites of serpentinization.

The current study, using a comprehensive database of genomic and geochemical data, highlights the occurrence of Betaproteobacteria and Clostridia in the serpentinizing subsurface and the roles that high pH and associated parameters play in shaping the taxonomic and functional diversity of these communities (Fig. 2.3, Table 2.3). Microbial communities in end-member wells are more likely to encode proteins associated with hydrogen metabolism, carbon monoxide oxidation, and carbon fixation compared to those in more moderate wells (Fig. 2.5). Meanwhile, methane was most strongly correlated with the community compositions of the moderate pH wells (Fig. 2.5). These data suggest that there are different ecological niches within the serpentinite subsurface environment, with hydrogen metabolism and carbon monoxide oxidation taking place in the most end-member fluids and methane oxidation occurring in mixing zones between serpentinite end-member fluids and surface waters. The findings of this study add

to a growing body of evidence that serpentinizing subsurface aquifers are dominated by carbon-fixing, hydrogen-oxidizing Betaproteobacteria and anaerobic Clostridia. These data serve as a foundation for future studies investigating the activities of subsurface microbial populations and their influence on the flux of hydrogen and carbon from subsurface environments.

Materials and Methods

Site Description and Sample Collection

Samples were collected from seven wells at the Coast Range Ophiolite Microbial Observatory (CROMO), located at the UC-Davis McLaughlin Nature Reserve in Lower Lake, CA, which consists of three groups of wells located within a 1.4 km radius of each other: the Core Shed Wells (CSW), Quarry Valley (QV), and the N-wells. The CSW and QV wells were drilled using clean drilling techniques in 2011 to enable subsequent monitoring of the microbial communities and associated geochemistry within the serpentinite subsurface over time (10). CSW consists of five wells, drilled to depths of 9-27 meters. QV consists of three wells, drilled to depths of 15-23 meters. The N-wells, ranging in depth from 14-40 meters, represent previously existing wells that were drilled in 1983 for environmental monitoring of mining impacts on groundwater, not with the specific purpose of monitoring microbiology and organic geochemistry (10).

Well fluids were collected using positive displacement Teflon bladder pumps (Geotech Environmental Equipment, Denver, CO) and pumped through a YSI 3059 flow cell fitted with a YSI 556 multiprobe (Yellow Springs, OH), which measured water temperature, specific conductance, pH, dissolved oxygen (DO) and oxidation-reduction potential (ORP) once the DO measurement stabilized at a minimum value. The outlet of the YSI flow cell was fitted with silicon tubing, allowing for the collection of fluids for geochemical and microbiological samples.

In the field, samples for dissolved gas analyses (CH_4 , CO , and H_2) were extracted from the fluids and aqueous phase samples (DIC and organic acids) were collected anaerobically, as previously described (19).

Prior to fluid collection, the peristaltic pump tubing was flushed for three minutes with well water to remove any contaminants from previous wells. Fluids (ranging between 0.5 to 10 L, depending on the well) were filtered through a 0.22 μm Sterivex filter cartridge (Millipore, Billerica, MA) using an inline peristaltic pump for DNA analyses. Field replicate samples, ranging between two to eight filters per well, were collected. Sterivex cartridges were flash frozen with liquid nitrogen and stored at -80°C until DNA extraction. For microbial cell quantification, replicate samples of 50 mL of fluids were preserved at a final concentration of 3.7% formaldehyde and stored at 4°C . All publically available data generated from this project can be found at: <http://cromo.arc.nasa.gov>.

Geochemistry

Dissolved gases (H_2 , CH_4 , and CO) were extracted into an inert gas phase of known volume and analyzed for CH_4 via a SRI 8610C GC-FID and dissolved H_2 and carbon monoxide via a Trace Analytical RGA3 Reduced Gas Analyzer. DIC was measured by acidifying a known volume of well fluid within a sealed vial, and analyzing the concentration of liberated CO_2 in the headspace by GC-FID (SRI 8610) following passage through a “methanizer”. Organic acid samples were analyzed by HPLC with UV/VIS detection, following derivatization with 2-nitrophenylhydrazide (34). All sample vials were analyzed with duplicate injections.

Microbial Cell Counts

Fluids preserved for cell counts were filtered through a 0.2 μm filter. The cells were stained with 1 $\mu\text{g}/\text{ml}$ of 4',6-diamidino-2-phenylindole (DAPI) and were counted by

epifluorescence microscopy using appropriate filter sets according to previously published protocols (35, 36).

DNA Extraction

DNA extractions from sterivex filters were performed by lysis via freeze/thaw cycles and lysozyme/Proteinase K treatment and purified with phenol-chloroform extractions, precipitation in ethanol, and further purification with QiaAmp (Qiagen, Hilden, Germany) columns according to the manufacturer's instructions for purification of genomic DNA, as described previously in Brazelton *et al.*, 2013 (8).

16S rRNA Tag Sequencing and Data Analysis

Samples were submitted to the DOE Joint Genome Institute (JGI) for 16S rRNA amplicon sequencing of the V4 region on an Illumina MiSeq instrument, as described in Caporaso *et al.* (20). Sequence reads were aligned to the SILVA SSURef alignment (v119), and taxonomic classification was assigned using mothur (27, 37). Sequences were clustered into operational taxonomic units (OTUs) at the 3% distance threshold using the cluster.split command and the average-neighbor clustering algorithm in mothur (28). Beta diversity (between sample) of the microbial communities was assessed by calculation of the Bray-Curtis index and displayed in a dissimilarity dendrogram with statistical significance determined by the SIMPROF test run in Primer-6 (22, 38). Alpha diversity (within sample) was assessed by rarefaction analysis and the Simpson diversity index, which was calculated after subsampling the data to the sample with the fewest sequences (77,580). Sequence identity of reads belonging to the top OTUs compared with 16S rRNA sequences from other serpentinite studies (8, 9, 17) was performed using MatGAT with the default settings (29). The 16S rRNA sequence data are

publically available in the NCBI Sequence Read Archive under the accession number SRA280854.

Metagenomic Sequencing and Data Analysis

Samples were submitted to JGI for metagenomic sequencing on an Illumina HiSeq2000 instrument, as described in Hawley *et al.* (26). These data are publically available in the JGI IMG/M database (www.img.jgi.doe.gov) under the project IDs: 1021918, 1021921, 1021924, and 1021927; and in the MG-RAST database (www.metagenomics.anl.gov) under the following sample IDs: 4569549.3, 4569550.3, 4569551.3, and 4569552.3. Metagenomes were searched for protein-encoding genes indicative of metabolic pathways of interest using similar methods as those described by Brazelton *et al.* (7). Representative protein sequences of the genes-of-interest were used as queries in BLASTP (30) searches against the assembled metagenomes. Unique metagenomic sequence reads that matched at least one query were quantified as reads per kilobase and normalized to total kilobases in each metagenome. The queries for each BLASTP search were representative ('seed') protein sequences obtained from Pfam (31) for the large subunits of [Fe-Fe] hydrogenase (PF02906) and [Ni-Fe] hydrogenase (PF00374), particulate methane monooxygenase *pmoA* (PF02461), methanol dehydrogenase *mxoF* (PF01011), methyl co-enzyme-M reductase A *mcrA* (PF02249), and the large subunit of the RuBisCo gene (PF00016). The nickel-dependent carbon monoxide dehydrogenase gene, *ccoS* (TIGR01702), was obtained from the TIGRFAM database (39). Representative sequences for the large subunit of carbon monoxide dehydrogenase, *coxL*, were obtained from Cunliffe *et al.* (40). The assembled metagenomes were annotated with Prokka (41), and these annotations were compared with the results of the BLASTP searches described above. Specifically, any BLASTP hits with annotations that did not match the protein target were considered to be false positives. None of

the BLASTP searches had false positive rates greater than 20%, except for *coxL*. Therefore, the normalized coverages for *coxL* reported in Figure 2.5 were obtained by searching for annotations matching UniProt ID P19913, corresponding to *coxL* in *Hydrogenophaga pseudoflava*. The coverage (as reads per kilobase) for the contigs associated with these annotations were summed and normalized to total kilobases in each metagenome.

Statistical Analyses

Correlation network analyses were constructed from statistically significant pairwise Pearson correlations among environmental variables and sequence data (32) and visualized in Cytoscape v 2.8.3 (33). A matrix containing environmental data and relative OTU (97% similarity) abundance for each sample was used as input for pairwise Pearson correlation analysis computed with the `rcor.test` function in the R package `lmt` (42). The false-discovery rate (q-value) was computed for the distribution of Pearson p-values to account for multiple tests. Pairwise correlations with both p- and q-values of <0.05 were considered significant and included in network analyses. Network models of significant correlations were created using Cytoscape v2.8.3 (33).

The ANOSIM test using a Bray-Curtis similarity index was used to test whether individual environmental parameter categories had significant effects on the community composition of samples (22). To statistically determine which combinations of numerical environmental variables best described the community composition variation within the dataset, the BEST analysis was performed in PRIMER-6 (22, 38).

Bioenergetic Calculations

Gibbs energy (ΔG_r) calculations identify which reactions are thermodynamically favorable (i.e., exergonic), and therefore good candidates to supply energy for microbial

metabolic processes. This is done by evaluating the Gibbs energy of reaction (ΔG_r), which is the sum of the standard Gibbs energy (ΔG_r°) and a term that reflects the chemical composition of the system. In the latter, the impact of different chemistries on ΔG_r is taken into account; interstitial fluid chemistries differ across the four modeled wells, thus ΔG_r also differs. Measured concentrations of interstitial fluids are transformed into *in situ* activities, which are used directly in calculating ΔG_r , using the thermodynamic modeling package of EQ3/6 (43).

ΔG_r can be calculated at *in situ* conditions using the expression (Equation 2.1):

$$\Delta G_r = \Delta G_r^\circ + RT \ln Q_r \quad (\text{Eq. 2.1})$$

where ΔG_r is the Gibbs energy of reaction, ΔG_r° is the standard Gibbs energy, R and T represent the gas constant and temperature in Kelvin, respectively, and Q_r stands for the activity product, discussed below. ΔG_r° can be determined at the appropriate temperature and pressure for the aqueous species and minerals can be calculated using established equations of state (44 and references therein).

Q_r , the activity product, can be computed from environmental data as shown in (Equation 2.2):

$$Q_r = \prod a_i^{v_{i,r}} \quad (\text{Eq. 2.2})$$

where a_i represents the activity of the i th species, and $v_{i,r}$ represents the stoichiometric reaction coefficient. Values of a_i are generated from concentration data (Table 2.1) and activity coefficients, using the geochemical speciation code EQ3 (43). In this code, activity coefficients are calculated using a variant (B-dot equation) of the Extended Debye-Hückel activity coefficient formalism (45), with reference to the SUPCRT92 (46) thermodynamic database.

For comparative purposes, values of ΔG_r are standardized to the amount of energy per mole of electron transferred (44). As an estimate of the amount of metabolic energy available in

the environment at the time of sampling, the total potential energy yield for each reaction was calculated by multiplying ΔG_r by the concentrations of the limiting reactant in the reaction (see 46, 47). The result is an estimate of the amount of metabolic energy that was available from the reaction at the time of sampling if all of the limiting reactant was consumed in the specified reaction.

REFERENCES

REFERENCES

1. Whitman WB, Coleman DC, Weibe WJ (1998) Prokaryotes: the unseen majority. *Proc Natl Acad Sci USA* 95:6578-6583.
2. Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S (2012) Global distribution of microbial abundance and biomass in seafloor sediments. *Proc Natl Acad Sci USA* 109:16213-16216.
3. Northrup DE, Lavoie KH (2001) Geomicrobiology of caves: a review. *Geomicrobiol J* 18:199-222.
4. Engel AS, Stern LA, Bennett PC (2004) Microbial contribution to cave formation: new insights into sulfuric acid speleogenesis. *Geology* 32:369-372.
5. Onstott TC, et al. (2003) Indigenous and contaminant microbes in ultradeep mines. *Environ Microbiol* 5:1168-1191.
6. Chivian D, et al. (2008) Environmental genomics reveals a single-species ecosystem deep within Earth. *Science* 322:275-278.
7. Brazelton WJ, Nelson B, Schrenk MO (2012) Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2:268.
8. Brazelton WJ, Morrill PL, Szponar N, Schrenk MO (2013) Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl Environ Microb* 79:3906-3916.
9. Suzuki S, et al. (2013) Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinitizing ecosystem. *Proc Natl Acad Sci USA* 110:15336-15341.
10. Cardace D, et al. (2013) Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Scientific Drilling* 16:45-55.
11. McCollom TM, Seawald JS (2007) Abiotic synthesis of organic compounds in deep-sea hydrothermal environments. *Chem Rev* 107:382-401.
12. Proskurowski G, et al. (2008) Abiogenic hydrocarbon production at Lost City Hydrothermal Field. *Science* 319:604-607.
13. Kelley DS, et al. (2005) A serpentinite-hosted ecosystem: the Lost City Hydrothermal Field. *Science* 307:1428-1434.

14. Schrenk MO, Kelley DS, Bolton SA, Baross JA (2004) Low archaeal diversity linked to seafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6:1086-1095.
15. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA (2006) Methane- and sulfur-metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ Microbiol* 72:6257-6270.
16. Suzuki S, et al. (2014) Physiological and genomic features of highly alkaliphilic hydrogen-utilizing *Betaproteobacteria* from a continental serpentinizing site. *Nat Commun* 5:3900.
17. Tiago I, Veríssimo A (2013) Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environ Microbiol* 15:1687-1706.
18. Schrenk MO, Brazelton WJ, Lang SQ (2013) Serpentinization, carbon and deep life. *Rev Mineral Geochem* 75:575-606.
19. Crespo-Medina M, et al. (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604.
20. Caporaso JG, et al. (2011) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621-1624.
21. Meyer F, et al. (2008) The metagenomics RAST server: a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386.
22. Clarke KR, Gorley RN (2006) PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.
23. Morrill PL, et al. (2014) Investigations of potential microbial methanogenic and carbon monoxide utilization pathways in ultra-basic reducing springs associated with present-day continental serpentinization: the Tablelands, NL, CAN. *Front Microbiol* 5:613.
24. Lau E, Fisher MC, Steudler PA, Cavanaugh CM. (2013) The methanol dehydrogenase gene, *mxhF*, as a functional and phylogenetic marker for proteobacterial methanotrophs in natural environments. *PLoS ONE* 8:e56993.
25. Wang DT, et al. (2015) Nonequilibrium clumped isotope signals in microbial methane. *Science* 348:428-431.
26. Hawley, ER, et al. (2014) Metagenomic analysis of microbial consortium from nature crude oil that seeps into the marine ecosystem offshore Southern California. *Stand Genomic Sci* 9:1259-1274.

27. Pruesse E, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188-7196.
28. Schloss PD, Westcott SL (2011) Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl Environ Microb* 77:3219-3226.
29. Campanella JJ, Bitincka L, Smalley J (2003) MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* 4:29.
30. Altschult SF, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402.
31. Finn RD, et al. (2008) The Pfam protein families database. *Nucleic Acids Res* 36:D281-D288.
32. Fuhrman J, Steele J (2008) Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat Microb Ecol* 53:69-81.
33. Shannon P, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498-2504.
34. Albert D, Martens C (1997) Determination of low-molecular-weight organic acid concentrations in seawater and pore-water samples via HPLC. *Mar Chem* 56:27-37.
35. Hobbie JE, Daley RJ, Jasper S (1977) Use of nucleopore filters for counting bacteria by fluorescence microscopy. *Appl Environ Microb* 33:1225-1228.
36. Schrenk MO, Kelley DS, Delaney JR, Baross JA (2003) Incidence and diversity of microorganisms within the walls of an active deep-sea sulfide chimney. *Appl Environ Microbiol* 69:3580-3592.
37. Schloss PD, et al. (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microb* 75:7537-7541.
38. Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* 18:117-143.
39. Haft HH, et al. (2001) TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids Res* 29:41-43.
40. Cunliffe M, et al. (2008) Phylogenetic and functional gene analysis of the bacterial and archaeal communities associated with the surface microlayer of an estuary. *ISME J* 2:776-789.

41. Seeman, T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 14:2069-2069.
42. Rizopoulos D (2006) ltm: an R package for latent variable modeling. *J Stat Softw* 17:5. 43. Wolery TJ (1992) EQ6: EQ3NR, A Computer Program for Geochemical Aqueous Speciation-Solubility Calculations: Theoretical Manual, User's Guide, and Related Documentation (Version 7.0) Lawrence Livermore National Lab. UCRL-MA-110662 PT III
44. Amend JP, Shock EL (2001) Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and Bacteria. *FEMS Microbiol Rev* 25:175-243.
45. Helgeson HC (1969) Thermodynamics of hydrothermal systems at elevated temperatures and pressures. *Am J Sci* 267:729-804.
46. Johnson JW, Oelkers EH, Helgeson HC (1992) SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000C. *Computers & Geosciences* 18:899-947.
47. McCollom TM, Shock EL (1997) Geochemical constraints on chemolithoautotrophic metabolism by microorganisms in seafloor hydrothermal systems. *Geochim Cosmochim Acta* 61:4375-4391.
48. Oosterkamp MJ, et al. (2011) Genome sequences of *Alicyclophilus denitrificans* strains BC and K601. *J Bacteriol.* 193:5028-5029.
49. Chou JH, et al. (2008) *Azonexus hydrophilus* sp. nov., a *nifH* gene-harboring bacterium isolated from freshwater. *Int J Syst Evol Micr* 58:946-951.

CHAPTER 3

Microbial diversity of serpentinite cores from the Coast Range Ophiolite Microbial Observatory²

Abstract

Serpentinization is a geochemical process in which ultramafic rocks are transformed to serpentine minerals and magnetite in the presence of water. The reactions create high pH fluids with an abundance of hydrogen and methane that can serve as potential energy sources for microbial communities. The subsurface environment is an amalgamation of rock matrix and the groundwaters permeating it, creating a complex habitat for microorganisms, which can be free-living in fluids or particle-attached in biofilms. Fluids from serpentinites have consistently contained two dominant taxa of bacteria, Betaproteobacteria and Clostridia. Little is known, however, about microbe-mineral interactions within the serpentinite subsurface environment. The Coast Range Ophiolite Microbial Observatory (CROMO), a set of wells drilled into an actively serpentinizing ophiolite, was established with the aim of characterizing the habitability of the serpentinite environment and from it cores were obtained. This study represents the first look at the microbial communities associated with cores from the serpentinite subsurface environment. Different core-enriched taxa were correlated with distinct lithostratigraphic zones within the heterogeneous cores, suggesting that mineralogy may impact community composition. Archaea, previously undetected in CROMO fluid samples, were abundant in core samples

² The work described in this chapter is being developed for submission to *Applied and Environmental Microbiology*: K.I. Twing, W.J. Brazelton, M.D. Kubo, M. Tominga, D. Cardace, T.M. Hoehler, T.M. McCollom, and M.O. Schrenk. (In Preparation) Microbial diversity of serpentinite cores from the Coast Range Ophiolite Microbial Observatory. *Appl Environ Microbiol.*

containing magnetite-bearing serpentinite, while bacteria were more abundant in samples containing clay particles. The taxonomy of sequences significantly enriched in core samples suggest that serpentinite cores contain an abundance of bacteria and archaea involved in the cycling of nitrogen and methane, indicating that a more complex biogeochemical cycling may be taking place in the serpentinite subsurface environment than can be detected by investigating fluid samples alone.

Introduction

Serpentinization is the hydrous alteration of ultramafic rocks from the seafloor into magnetite and serpentine minerals (e.g. lizardite, chrysotiles, and antigorite; 14). The series of reactions creates high pH fluids, containing an abundance of hydrogen and methane. The precise mineral products of the reaction depend on a number of factors, including temperature, pressure, initial rock and mineral composition, and water to rock ratio (24). While this process begins on the subseafloor, sites of serpentinization can be accessed on land where the mantle has been obducted onto the continental crust and serpentinization continues, making it more accessible for research.

Fluids from continental serpentinite sites from around the world have yielded surprisingly consistent microbial community composition, containing low diversity and a dominance of Betaproteobacteria and Clostridia (3, 4, 10, 41, 44, 45, 46). The combination of interdisciplinary studies, integrating genetic and environmental data (4, 43), metagenomics (3), and cultured isolates (10, 43), has allowed for the development of models suggesting that anaerobic Clostridia are truly endemic to the extreme serpentinite subsurface environment, while Betaproteobacteria reside in the more hospitable oxic/anoxic mixing zones (41).

The serpentinite subsurface environment is a rock-hosted system. However, the majority of samples collected from these systems to-date have been fluid samples. Chimney samples from the Lost City Hydrothermal Field were dominated by methane-cycling archaea of the Methanosarcinales in anoxic chimney interiors (2, 40) and sulfur-cycling and methanotrophic bacteria at the exterior of the chimneys (1). In the continental setting, the only serpentinite minerals samples examined for microbiology are from the Leka ophiolite in Norway, where minerals from fractures, ranging from 15-160 cm deep were studied (11). Clone libraries of mineral samples from 155-160 cm were dominated by Actinobacteria, Acidobacteria, Firmicutes, Alpha- and Betaproteobacteria, and archaea from the Soil Crenarcheotic Group (11).

Serpentine soils are characterized by high levels of heavy metals (Ni, Cr, and Co) and low nutrients, which leads them to host very specialized flora (48). Due to this unique macroecology, many of the studies regarding the microbiology of serpentinite soils have focused on the rhizosphere (microbes associated with plant roots) and heavy metal tolerant microorganisms (19, 27, 30, 31, 34). The bacterial communities defined in these sources include: Acidobacteria, Actinobacteria, Bacteroidetes, Firmicutes, Gemmatimonadetes, Proteobacteria (Alpha-, Beta-, and Gamma-), and Verrucomicrobia. There have not been reports of archaea in serpentinite soil.

The serpentinite subsurface environment is a complex system, the microbiology of which is just beginning to be understood. The Coast Range Ophiolite Microbial Observatory (CROMO) was established to help address the question of microbe-mineral interactions in the serpentinite environment (7). This paper represents the microbial investigation, through the use of high-throughput 16S rRNA sequencing, of cores obtained from the serpentinite subsurface to a depth of 45 m and suggests that there are unique communities associated with solid surfaces in the serpentinite subsurface environment.

Results and Discussion

Drill Core Depth Profiles

Recovered cores indicate differences in lithology between the CSW1.1 and QV1.1 (Figs. 3.1 & 3.2), which may result in the variation in core recovery: 46% for CSW1.1 and 64% for QV1.1 (7). At CSW1.1, detrital serpentine soil was observed for 0-4 m, followed by 20 m of poor core recovery with a few sections of serpentine gravel in clay (Fig. 3.1A); then, below 24 mbs (meters below surface), magnetite, serpentine, and altered mafic rocks were found (Fig.

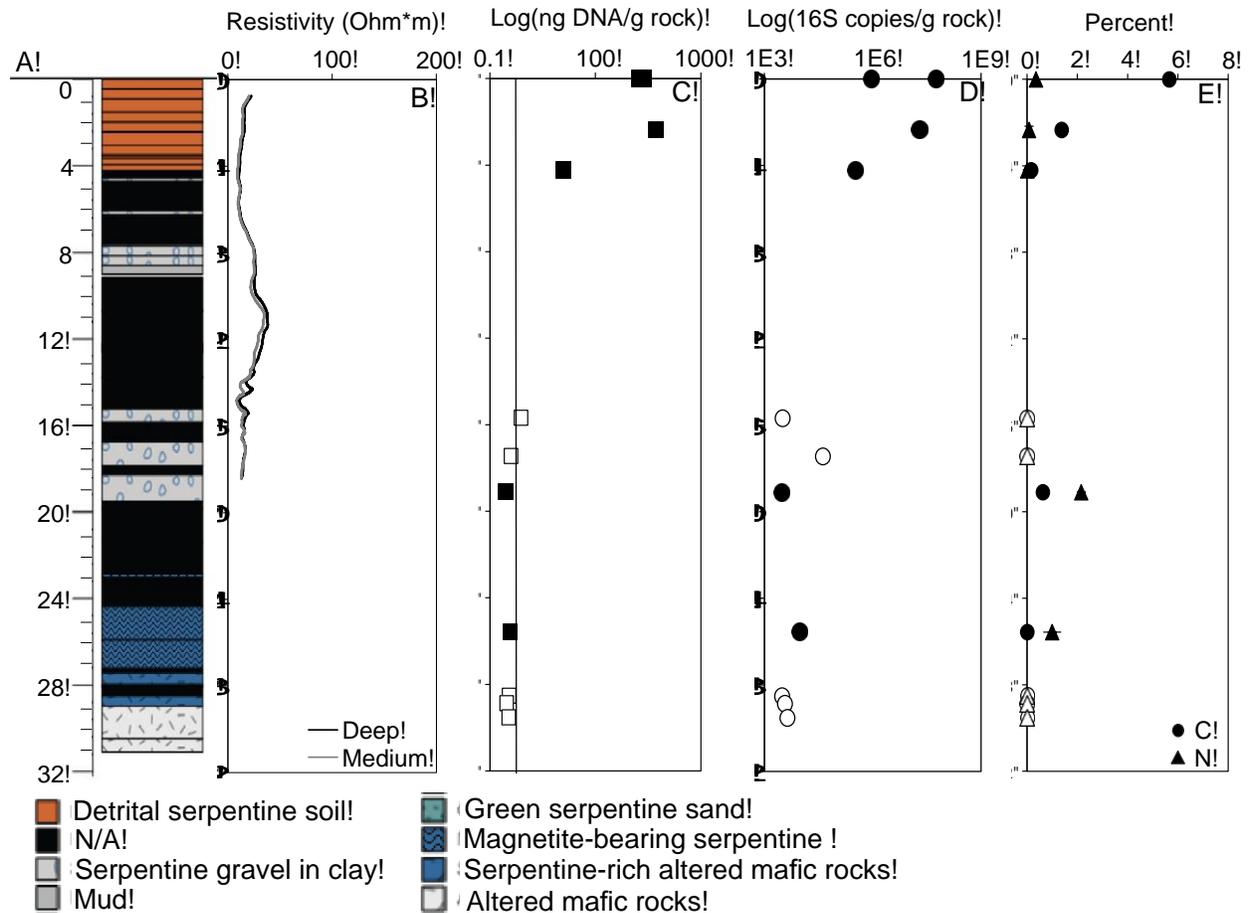


Figure 3.1 – Depth profile of CSW1.1. (A) Lithostratigraphic units, as described in Cardace *et al.*, 2013; (B) Resistivity data from down-hole logging; (C) DNA yield from DNA extractions, as quantified by fluorometric methods; (D) Bacterial 16S rRNA gene copies, as determined by domain-specific quantitative-PCR; (E) Percent organic carbon and percent nitrogen, as determined by EA-IRMS. N/A: sections with no core recovery. Filled symbols represent samples that yielded 16S rRNA sequences, samples represented by closed symbols failed sequencing.

3.1A). In contrast, QV1.1 only contained detrital serpentine soil for the first meter, after which there were various layers of magnetite, serpentine, and chlorite throughout the depth of the core (Fig. 3.2A). QV1.1 contained regions with no serpentine, such as at 31 mbs where magnetite-bearing clay with albite was found and 34-36 mbs where rocks containing albite, quartz, and chlorite were seen (Fig. 3.2A). Core recovery was higher in QV1.1 than in CSW1.1, with only two large sections between 2-5 mbs and 11-14 mbs where there was no core recovery (Fig.

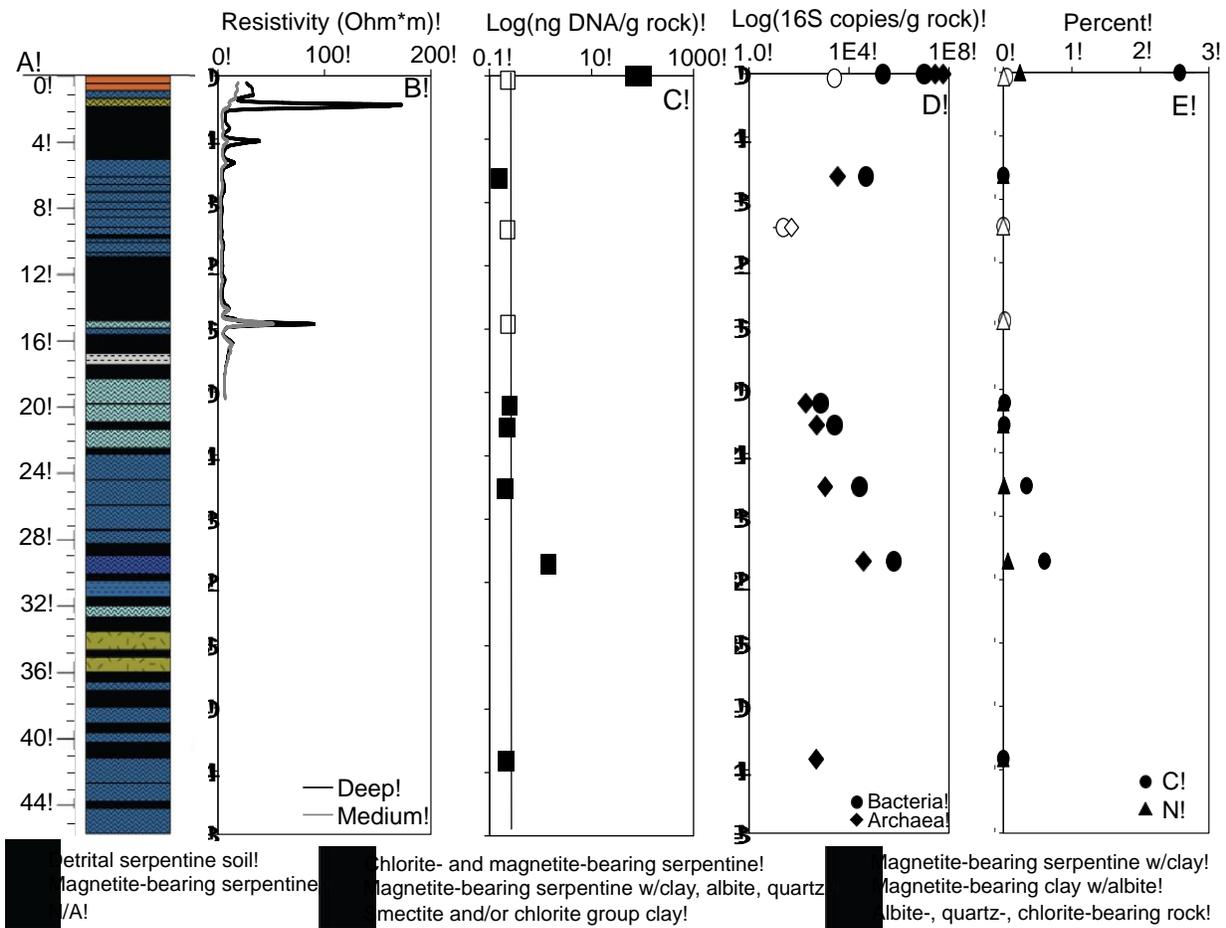


Figure 3.2 – Depth profile of QV1.1. (A) Lithostratigraphic units, as described in Cardace *et al.*, 2013; (B) Resistivity data from down-hole logging; (C) DNA yield from DNA extractions, as quantified by fluorometric methods; (D) 16S rRNA gene copies, as determined by domain-specific quantitative-PCR; (E) Percent organic carbon and percent nitrogen, as determined by EA-IRMS. N/A: sections with no core recovery. Filled symbols represent samples that yielded 16S rRNA sequences, samples represented by closed symbols failed sequencing.

3.2A).

Resistivity logging was conducted as part of the post-drilling, down-hole geophysical logging (Figs. 3.1B & 3.2B; 7). Due to the instrument limitations, the boreholes were only matrix is comprised of electrically resistive mineralogy (e.g. clay, sulfides) and/or more lithified, logged to 18 m (Figs. 3.1B & 3.2B). A broad peak of relatively high resistivity in CSW1.1 at the depth range of 5-15 m (Fig. 3.1B) coincides with the low core recovery in this zone (Fig. 3.1A). In QV1.1, there are two sharp spikes in resistivity at 2 and 15 m (Fig. 3.2B), which coincide with magnetite-bearing serpentine (Fig. 3.2A). High resistivity measurements indicate that the rock less porous formations. Other than the above-mentioned sections, most of the logged interval of both cores is very low resistivity, possibly suggesting where water resides in pore spaces, cracks, and faults within the wells (Figs. 3.1B & 3.2B).

DNA Yield, Organic Carbon, and 16S rRNA Gene Abundance

In order to conserve DNA for downstream analyses, only 2 μL of the 50 μL DNA extract was sacrificed for DNA quantification, making the limit of detection of the fluorometric quantification method 0.1 ng/ μL . DNA was extracted from roughly 20 g of core material for each sample, making the limit of detection for DNA quantification 0.2 ng/g, which many of the samples fell below (Fig. 3.1C & 3.2C). It should be noted, however, that despite unquantifiable DNA concentrations, many of the samples still were amplifiable via domain-specific q-PCR (Figs. 3.1D & 3.2D) and yielded 16S rRNA sequences (Table 3.1). Therefore, amplifiability, as opposed to DNA concentration was used to determine whether a sample was fit for submission to the sequencing facility. Due to the potential of contamination in very low biomass samples, blank control samples were run at every step of the DNA extraction and purification process. The

control samples were neither quantifiable by fluorometric methods nor amplifiable via domain-specific q-PCR, and therefore were considered to be negative of any contamination.

The total organic carbon (TOC) and total nitrogen (TN) were measured for each core sample by elemental-analysis isotope-ratio mass spectrometry (EA-IRMS; Figs. 3.1E & 3.2E). Other than the top 4 m of CSW1.1, which contained more organic carbon than nitrogen and had a C/N ratio representative of decaying organic matter (11:1), the two elements were below detection throughout the core, except for an increase of TN between 19-26 mbs (Fig. 3.1E), which is of unknown origin. This high C/N spike at ~20 m depth could indicate fresh biological organic matter, which has yet to be consumed by microorganisms. In QV1.1, the TN remained

Table 3.1 – Soil and core sample names and 16S rRNA gene sequence abundances

Sample	Depth (mbs)	DNA Yield (ng/g)	Bacterial 16S Sequences*	Archaeal 16S Sequences*
CSWsoil1	0	67.0	42,570	151,538
CSWsoil2	0	82.1	82,363	236,134
CSW1R	2.4	138.1	246,412	89,017
CSW2R	4.2	2.4	147,623	105,289
CSW16R	15.7	0.4	18	0
CSW17R	17.4	0.3	16	0
CSW10R	19.1	< 0.2	285,703	12,020
CSW14R	25.6	< 0.2	225,988	0
CSW22R	28.5	< 0.2	9	0
CSW23R	28.9	< 0.2	9	0
CSW24R	29.5	< 0.2	11	0
QVsoil1	0	66.4	44,333	193,560
QVsoil2	0	106.0	83,590	242,101
QV3R	0.3	< 0.2	1,919	2,806
QV7R	6.5	0.2	78,108	4,683
QV11R	9.7	< 0.2	4,635	266
QV13R	15.7	< 0.2	4,568	1,440
QV18R	20.8	< 0.2	9,225	504
QV21R	22.2	< 0.2	10,827	998
QV25R	26.1	< 0.2	8,495	2,486
QV30R	30.8	1.4	25,322	103,561
QV42R	43.3	< 0.2	11,414	130,842

*Sequence counts in bold were successfully sequenced, while those in regular font failed sequencing (i.e. produced < 5000 sequence reads).

low below the soil level, while the TOC spiked between 25 -30 mbs (Fig. 3.2E). The increased TOC with low TN could suggest abiotic input of carbon molecules (as from serpentinization) or older organic matter that has already been processed by microorganisms in the system.

There does not appear to be a consistent relationship between DNA concentration, 16S rRNA gene copies (obtained from q-PCR), and successful 16S rRNA sequencing (Figs. 3.1 & 3.2; Table 3.1). The spike in TN (~20 mbs) in CSW1.1 corresponds to the only samples that yielded 16S rRNA sequences at depth from that core, even though samples above and below that were amplified via q-PCR. The TOC spike in QV1.1 coincides with an increase in DNA concentrations and 16S rRNA genes (Fig. 3.2). The resistivity data (Figs. 3.1B & 3.2B), a proxy for how hard the rock material is, appears to correspond with whether quantifiable, amplifiable, and sequenceable DNA could be obtained from a core sample. This could be due to poor core recovery from highly resistive regions or due to fluid flow in conductive zones making them more habitable. Ongoing efforts are working to resolve the complex hydrogeology of the CROMO subsurface, which will help address this hypothesis more fully.

Core-Enriched Bacterial Communities

When dealing with natural samples, it is important to keep in mind that different “environments” may have been in communication with each other over relatively short time scales. For instance, this study is focused on describing the core-enriched community of microorganisms in the serpentinite subsurface environment and contains sequences from core-, soil-, and fluid-associated microbes. It would be naïve to think there is not natural flow of microbes between the fluids within the fractures of the rocks and vice versa. However, that is precisely how most methods of controlling for “contamination” in sequence data work (e.g. 37) – they rely on removing any sequence that occurs in both the sample (in this case core) and

background (in this case fluids and soils), assuming that the sequence must have come from the background and contaminated the true sample. Instead of using this overly conservative method, which potentially removes true inhabitants of the sample, this study used a statistical method modeled after differential expression in RNA-sequencing studies (26) to identify which sequences are significantly enriched in a particular type of sample, and therefore more likely to have originated there. In this fashion, core samples were compared to fluid and soil samples to identify the “core-enriched”, “fluid-enriched”, and “soil-enriched” sequences (sequences significantly more abundant in that sample type compared to the others), as opposed to the “other” sequences, which had no statistically defined habitat (Fig. 3.3).

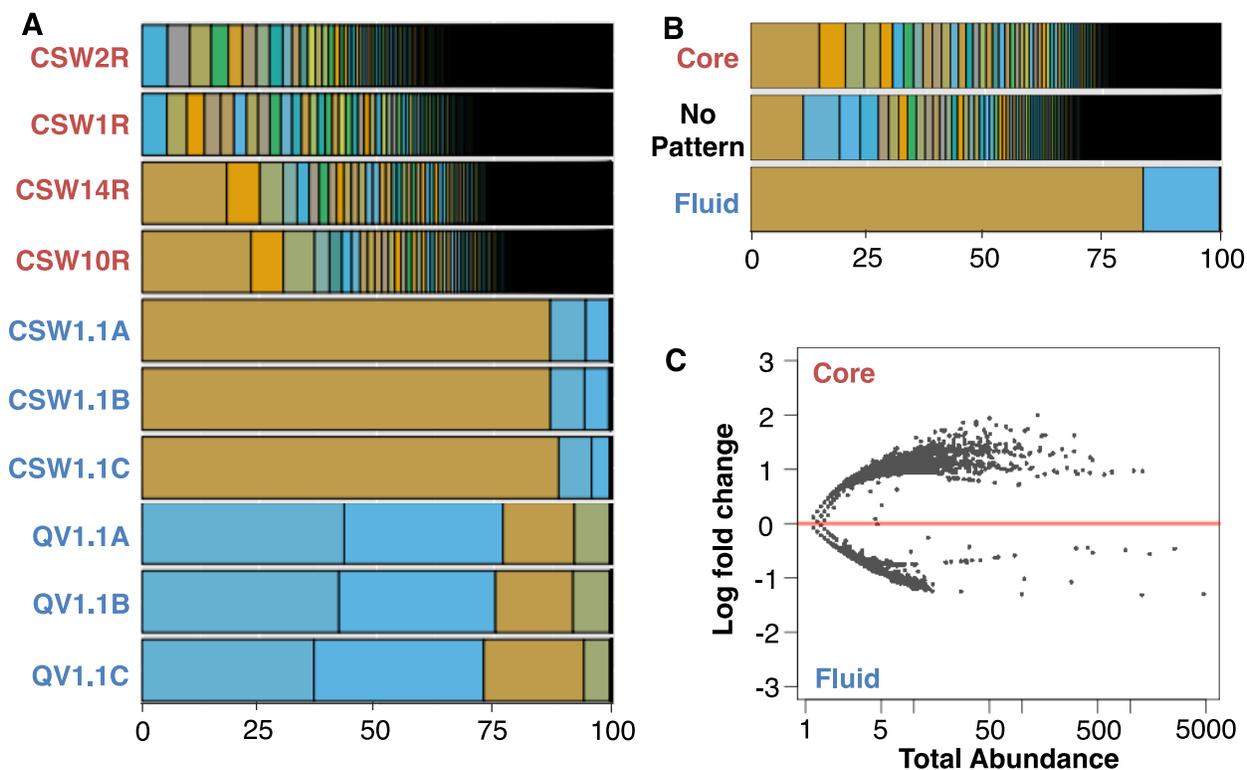


Figure 3.3 – DESeq comparison between CSW core and CROMO fluid samples. Core samples are depicted in red, fluid samples are depicted in blue. (A) Relative abundance of sequences found in individual core and fluid samples; (B) Sequences enriched in core and fluid samples versus those not statistically enriched in either group (No Pattern); (C) Plot of log fold change of each sequence between the two sample types over the total count (abundance) for each sequence across the data set.

The fluid-enriched taxa, which make up 20% of QV1.1 and 50% of CSW1.1 fluid communities (Fig. 3.4), are made up exclusively of Betaproteobacteria and Clostridia. The identification and abundances of these taxa are consistent with the findings in Chapter 2 (46), supporting the theory that Betaproteobacteria and Clostridia dominate fluids in the subsurface serpentinite environment (41). The remaining 80 and 50% of the communities, respectively, were made up of sequences not determined to be “fluid-enriched”, meaning there was no statistically significant increase in those taxa in the fluids compared to the soil or core samples.

Alternatively, the “soil-enriched” sequences were diverse, including members of the Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Planctomycetes,

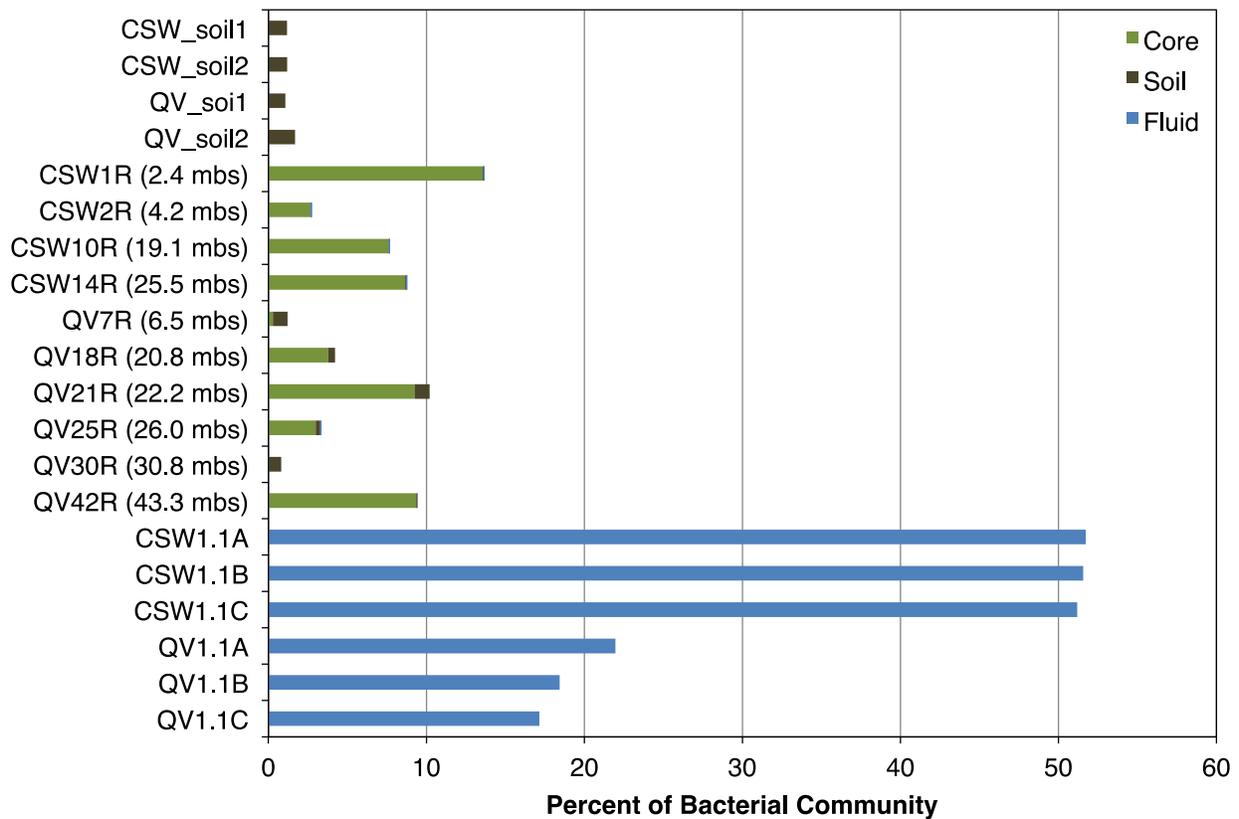


Figure 3.4 – Relative abundance of core-, soil-, and fluid-enriched bacterial sequences.

Protobacteria, and Verrucomicrobia and made up between 1-2% of the soil communities (Fig. 3.4). The soil-enriched taxa found here are consistent with previous reports of serpentine soil and rhizosphere microbial communities (19, 27, 31), with the exception of the Chloroflexi and Planctomycetes, which were not seen in previous studies. The members of the Chloroflexi identified as soil-enriched belong to the anaerobic, non-phototrophic, filamentous classes (16). The soil-enriched Planctomycetes sequences belonged to the order Planctomycetales, which contains deeply-branching bacteria capable of anaerobic ammonium oxidation (ANAMMOX; 13) and they have been isolated from a variety of soil environments (6), though not serpentine soils. The remaining 99% of the soil sequences were not significantly enriched in the soils, compared to the cores and fluids.

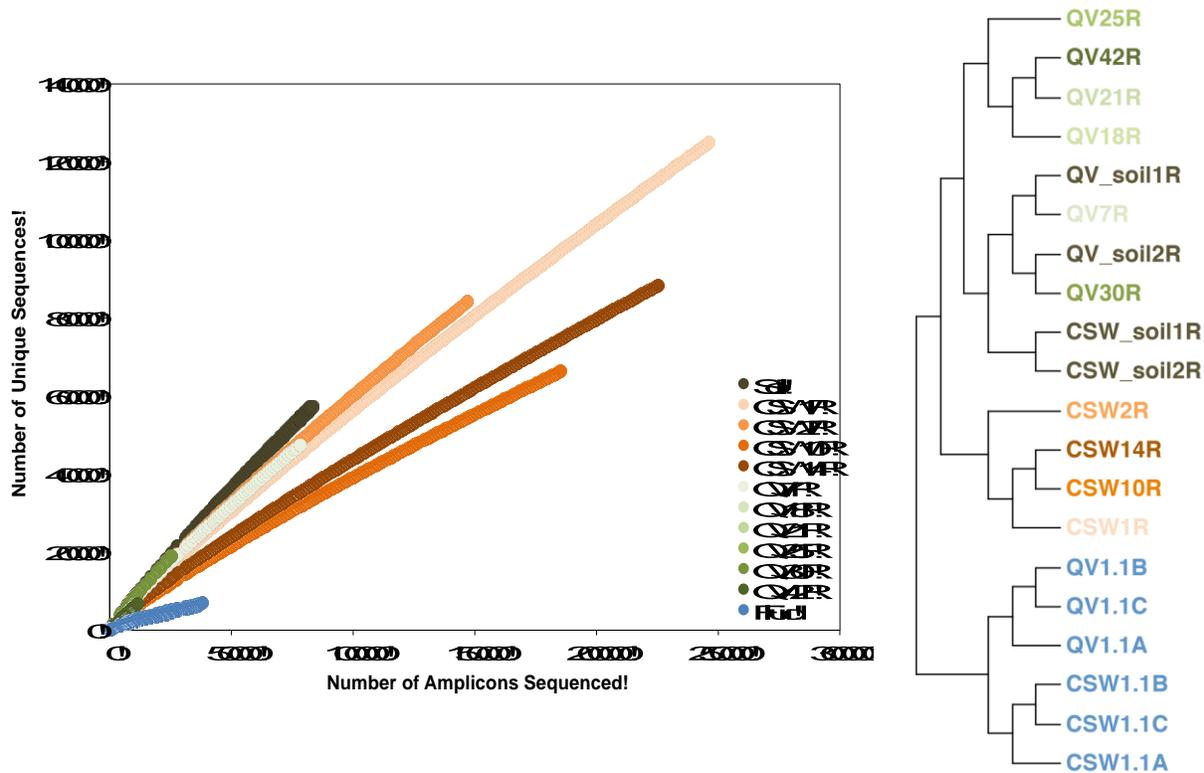


Figure 3.5 – Bacterial community structure of unique sequences. (A) Rarefaction analysis of unique 16S rRNA amplicon sequences. (B) Community similarity dendrogram calculated from Morisita-Horn dissimilarity index.

Table 3.2 - Relative abundance of individual core-enriched bacterial taxa within samples.

Phylum	Class	CSW1R	CSW2R	CSW10R	CSW14R	QV7R	QV18R	QV21R	QV25R	QV30R	QV42R
Actinobacteria	Actinobacteria	0	0.003	0.112	0.163	0.201	3.39	5.43	2.92	0.019	7.97
Proteobacteria	Deltaproteobacteria	2.52	0.532	1.27	1.81	0	0	0	0	0	0
Acidobacteria	Acidobacteria	2.35	0.128	1.33	1.51	0	0	0	0	0	0
Firmicutes	Bacilli	0	0.003	0.215	0.157	0.024	0.367	3.78	0.024	0.008	0.228
Nitrospirae	Nitrospira	1.43	1.26	0.653	0.647	0	0	0	0	0	0
Proteobacteria	Betaproteobacteria	1.37	0.062	0.715	0.833	0	0	0	0	0	0
Proteobacteria	Gammaproteobacteria	1.39	0.008	0.633	0.799	0	0	0	0	0	0
Acidobacteria	Subgroup_22	0.829	0.169	0.184	0.320	0	0	0	0	0	0
Spirochaetae	Spirochaetes	0.504	0.033	0.268	0.363	0	0	0	0	0	0
Chlorobi	Ignavibacteria	0.425	0.025	0.422	0.251	0	0	0	0	0	0
Proteobacteria	Alphaproteobacteria	0.391	0.033	0.136	0.251	0	0	0	0	0	0
WS3		0.313	0.055	0.198	0.205	0	0	0	0	0	0
Chloroflexi	Anaerolineae	0.399	0.083	0.080	0.199	0	0	0	0	0	0
Gemmatimonadetes	Gemmatimonadetes	0.226	0.055	0.162	0.217	0	0	0	0	0	0
Chloroflexi	Ktedonobacteria	0.221	0.002	0.142	0.131	0	0	0	0	0	0
Acidobacteria	Holophagae	0.069	0.083	0.282	0.036	0	0	0	0	0	0
OD1		0.177	0	0.054	0.108	0	0	0	0	0	0
Planctomycetes	Phycisphaerae	0.171	0.005	0.091	0.066	0	0	0	0	0	0
Chloroflexi	Dehalococcoidia	0.101	0.010	0.111	0.069	0	0	0	0	0	0
Chloroflexi		0.074	0	0.059	0.074	0	0	0	0	0	0
TM6		0.079	0.001	0.097	0.019	0	0	0	0	0	0
Planctomycetes	Pla4_lineage	0.101	0.001	0.038	0.029	0	0	0	0	0	0
Chloroflexi	P2-11E	0.058	0.0007	0.061	0.042	0	0	0	0	0	0
Chloroflexi	JG30-KF-CM66	0.032	0	0.028	0.099	0	0	0	0	0	0
Chloroflexi	Thermomicrobia	0.030	0.0007	0.034	0.042	0	0	0	0	0	0
Chloroflexi	TK10	0.024	0.0007	0.019	0.021	0	0	0	0	0	0
Thermotogae	Thermotogae	0.024	0.001	0.031	0.008	0	0	0	0	0	0
Chloroflexi	S085	0.008	0.010	0.010	0.016	0	0	0	0	0	0
Unclassified Bacteria		0.240	0.078	0.173	0.175	0.068	0.033	0.037	0.059	0	1.15
TOTAL		13.6	2.64	7.60	8.66	0.293	3.79	9.25	3.01	0.028	9.35

In the CSW cores, the “core-enriched” taxa make up between 2-14% of the bacterial communities and no soil-enriched taxa were detected (Fig. 3.4). The QV core samples, on the other hand, exhibited different community patterns. QV7R (4.2 mbs) and QV30R (30.8 mbs) contained mostly soil-enriched sequences, while the deepest sample, QV42R contained 10% core-enriched sequences and no soil-enriched sequences. The core samples from the middle of QV (20-26 mbs) contained mostly core-enriched sequences with ~1% soil-enriched sequences. These results are consistent with the results depicted in the community similarity dendrogram, suggesting that there are differences in the communities in CSW cores and QV cores and that QV7R and QV30R are more similar to soil samples than other core samples (Fig. 3.5).

A list of the core-enriched taxa can be found in Table 3.2. Among them are Chloroflexi and OD-1, which are suggested to be representative of the serpentinite subsurface environment at The Cedars (44), but not detected in CROMO fluids (Chapter 2; 46). The CSW core samples contained an abundance of nitrogen cycling core-enriched taxa, including nitrite-oxidizing *Nitrospira* (12) and Nitrospinae (23) and ammonium-oxidizing Nitrosomonadaceae (32). CSW cores also contained iron-oxidizing Betaproteobacteria of the family Gallionellaceae (15). The most abundant core-enriched sequence in QV cores belonged to the Mycobacteriaceae, a taxon that has also been detected in deep subsurface samples from the Homestake Gold Mine in South Dakota (35) and chromium contaminated sites (20). The only sequences found previously in serpentinite mineral samples were from clone libraries (11) and determined to be core-enriched in this study are H₂-oxidizing Actinobacteria (42) of the family Mycobacterium, abundant in cores from QV1.1 and nitrite-oxidizing *Nitrospira* (12) found in CSW1.1 cores.

Core-Enriched Archaeal Communities

Between 25-40% of the soil archaeal communities were made up of soil-enriched sequences (Fig. 3.6), which exclusively belonged to the Soil Crenarcheotic Group. None of the fluid samples from CSW1.1 or QV1.1 submitted for 16S archaeal sequencing yielded sequences. This is consistent with findings from Chapter 2, where no archaeal sequences were detected when universal 16S rRNA primers were used and metagenomic datasets contained only 1% archaeal sequences (46). Contrary to findings at other serpentinite sites, which have detected Methanosarcinales (44) and archaea from the SAGMEG-2 and ANME-1 groups (45), to date no archaea have been detected in the serpentinite fluids at CROMO.

While 18 core samples were submitted for 16S rRNA archaeal sequencing, only four of them were successfully sequenced. As with bacteria, the success of sequencing does not correlate with q-PCR results or DNA yield per g rock (Fig. 3.1 & 3.2). From CSW, only the shallowest

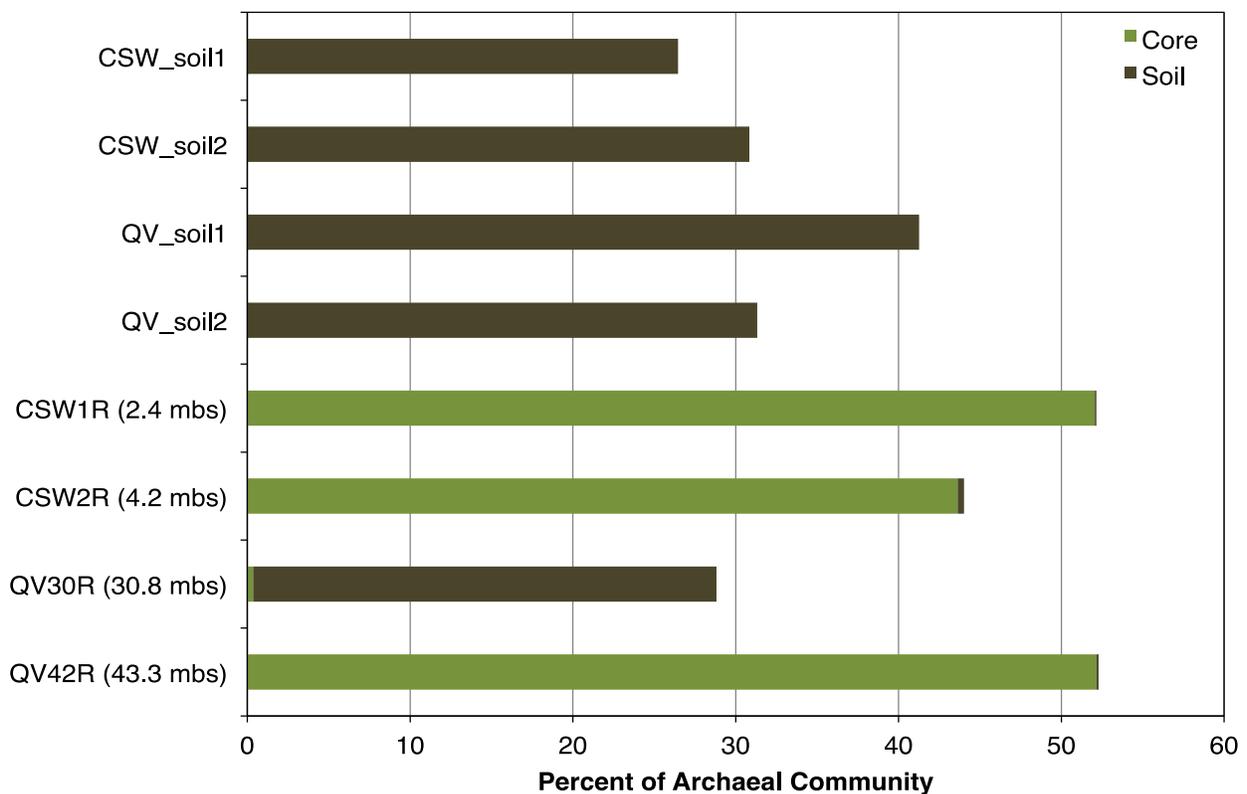


Figure 3.6 – Relative abundance of core-, soil-, and fluid-enriched archaeal sequences.

samples, CSW1R (2.4 mbs) and CSW2R (4.2 mbs) yielded sequences. While from QV, only the deepest samples, QV30R (30.8 mbs) and QV42R (43.3 mbs) were successfully sequenced. Between 45-50% of CSW1R, CSW2R, and QV42R consisted of core-enriched archaeal sequences (Fig. 3.6), which contain a diverse group of Euryarchaea, Crenarchaea, and Thaumarchaea (Table 3.3), which contain a diverse group of Euryarchaea, Crenarchaea, and Thaumarchaea (Table 3.3). The most abundant core-enriched archaeal taxa were from the Miscellaneous Crenarchaeotic Group, making up roughly 50% of the core-specific archaeal sequences in each of the three samples (Table 3.3).

Almost 17% of the archaeal community from QV42R is composed of methanogens (Methanomicrobia, Methanococci, Methanobacteria; Table 3.3), while these same taxa cumulatively made up less than 0.5% of the CSW1R and CSW2R communities and were not detected in QV30R. The microbial communities of the anoxic interiors of chimneys at the Lost City Hydrothermal Field were almost exclusively composed of Methanosarcinales (2, 40), an order that made up almost 11% of the archaeal community in QV42R. Similar to the bacterial

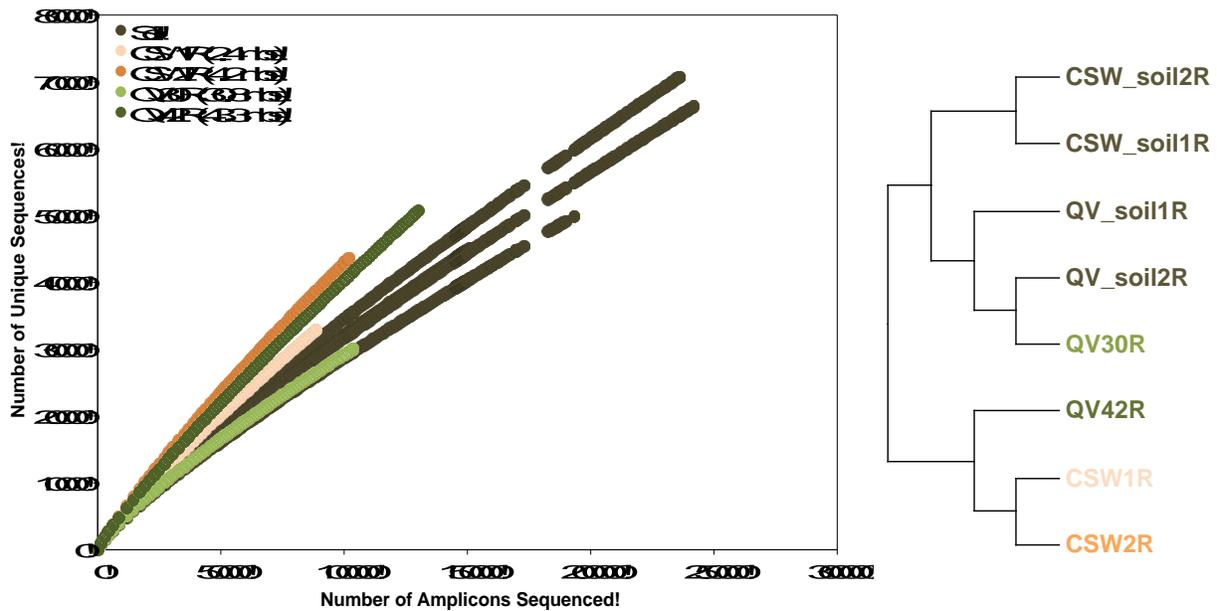


Figure 3.7 – Archaeal community structure of unique sequences. (A) Rarefaction analysis of 16S rRNA amplicon sequences. (B) Community similarity dendrogram calculated from Morisita-Horn dissimilarity index.

findings, the QV30R archaeal community looks much more like the soil samples than other core samples (Fig. 3.7). Almost 30% of QV30R was made up of the soil-enriched Soil Crenarcheotic Group, meanwhile less than 0.5% of the community was made up of core-enriched taxa (Fig. 3.6).

Biogeochemical Correlations

The abundance of core-enriched bacterial and archaeal sequences were compared with XRD (mineralogy), EA-IRMS (carbon and nitrogen abundance and stable isotope composition), and core lithology data to determine what might be driving community composition within the core samples. The lithostratigraphic zones of the cores (Figs. 3.1A & 3.2A; Table 3.4) had an impact on the microbial communities. With the exception of detrital serpentine soils (LS_1; table 3.4), which had significant correlations with both bacterial and archaeal sequences, the lithostratigraphic regions appeared to have correlations with members of one domain or the other. Within the detrital serpentine soil (LS_1), there was an abundance of bacteria, from the phyla Acidobacteria, Nitrospirae, Chloroflexi (class Anaerolineae), and Proteobacteria (class Deltaproteobacteria), and archaea, from the Miscellaneous Crenarchaeal Group, Thermoplasmata, and Thaumarchaeal Group C3 (Figs. 3.8 & 3.9A). Serpentine gravel in a clay matrix (LS_2; Fig. 3.1A), had a positive correlation with percent nitrogen and bacterial members of the Acidobacteria, Proteobacteria (Beta-, Alpha-, and Delta-), Chlorobi, and Gemmatimonadetes (Fig. 3.8B). Magnetite-bearing serpentine (LS_3), contained a diversity of archaeal groups, including the methanogenic orders Methanosarcinales, Methanomicrobiales, and Methanococcales (Fig. 3.9B). Magnetite-bearing serpentine with clay, albite, and quartz (LS_4) and was positively correlated to C/N ratio and Actinobacteria and Firmicutes (Fig. 3.8C).

Table 3.3 – Relative abundance of individual core-enriched archaeal taxa within samples.

Phylum	Class	CSW1R	CSW2R	QV30R	QV42R
Miscellaneous Crenarchaeotic Group		29.3	17.9	0	20.1
Euryarchaeota	Thermoplasmata	14.4	12.4	0.141	11.1
Euryarchaeota	Methanomicrobia	0.281	0.137	0	10.6
Thaumarchaeota	Group C3	4.59	0.720	0	1.56
Euryarchaeota	Methanococci	0.077	0.014	0	3.81
Thaumarchaeota	Marine Group I	1.58	2.20	0.140	0.057
Thaumarchaeota	Soil Crenarchaeotic Group	0.182	2.83	0	0.429
Thaumarchaeota	AK31	0.005	2.96	0	0
Euryarchaeota	Methanobacteria	0.039	0.050	0	2.36
Thaumarchaeota	Marine Benthic Group A	0.369	1.53	0	0.100
Thaumarchaeota	South African Gold Mine Group 1	0.017	1.23	0.019	0
Thaumarchaeota	Marine Benthic Group B	0.001	0.019	0	0.967
Unclassified Thaumarchaeota		0.026	0.863	0	0.005
Aenigmarchaeota	Deep Sea Euryarchaeotic Group	0.415	0.081	0	0.209
Woesearchaeota		0.183	0	0	0.144
Euryarchaeota	Archaeoglobi	0	0.024	0	0.211
Thaumarchaeota	AK56	0	0.169	0	0
Crenarchaeota	Thermoprotei	0	0.061	0	0.084
Thaumarchaeota	AK59	0	0.111	0	0
Diapherotrites		0	0	0	0.103
Thaumarchaeota	pSL12	0.035	0	0	0
Euryarchaeota	Thermococci	0	0	0	0.032
Miscellaneous Euryarchaeotic Group		0	0	0.028	0.031
Parvarchaeota		0	0	0.025	0
Aigarchaeota	Terrestrial Hot Spring Group	0	0.015	0	0
Unclassified Aenigmarchaeota		0	0	0	0.008
Unclassified Archaea		0.466	0.149	0	0.280
TOTAL		52.063	43.635	0.354	52.169

Table 3.4 – Classification of lithostratigraphic units and samples belonging to them.

LS Classification	Lithostratigraphic Units*	Samples
LS_1	Detrital serpentine soil	CSW1R, CSW2R
LS_2	Serpentine gravel in clay	CSW10R
LS_3	Magnetite-bearing serpentine	CSW14R, QV7R, QV25R, QV40R
LS_4	Magnetite-bearing serpentine w/clay, albite, quartz	QV18R, QV21R
LS_5	Magnetite-bearing serpentine w/clay	QV30R

*Lithostratigraphic units were defined in Cardace *et al.*, 2013 and depicted in Figs. 1 & 2.

Magnetite-bearing serpentine with clay (LS_5), which was found in the deep outlier sample QV30, contained the archaeal groups Thermoplasmata and Marine Group I (Fig. 3.9C). These data suggest that the geological make up of a sample has strong influence on the microbiology living within it and that different geological horizons favor bacteria over archaea and vice versa.

A diverse group of bacterial phyla exhibited significant negative correlations with iron concentration (from XRF data) and significant positive correlations with DNA yield, indicating an inverse relationship between the two parameters (Fig. 3.10). This inverse relationship could have to do with the adsorption of nucleic acids to iron clay minerals (17, 47). While iron oxide clays can decrease the yield of high molecular weight DNA, it is not suspected that they alter the community composition, as the DNA adsorption is non-selective with regard to nucleotide patterns (9). Given the potential for iron oxides, such as magnetite, to adsorb nucleic acid particles, it is possible using a DNA extraction method to specifically address this issue, such as the one developed by Hurt *et al.* (17) would yield different results.

Conclusions

The methods employed in this study successfully allowed for the identification of core-, soil-, and fluid-enriched taxa from CROMO samples. The fluid- and soil-enriched sequences were consistent with previous studies of continental serpentinite fluids (3, 10, 41, 44, 45, 46) and serpentine rhizospheres (19, 27, 30, 31, 34), respectively. Archaea were previously undetected at

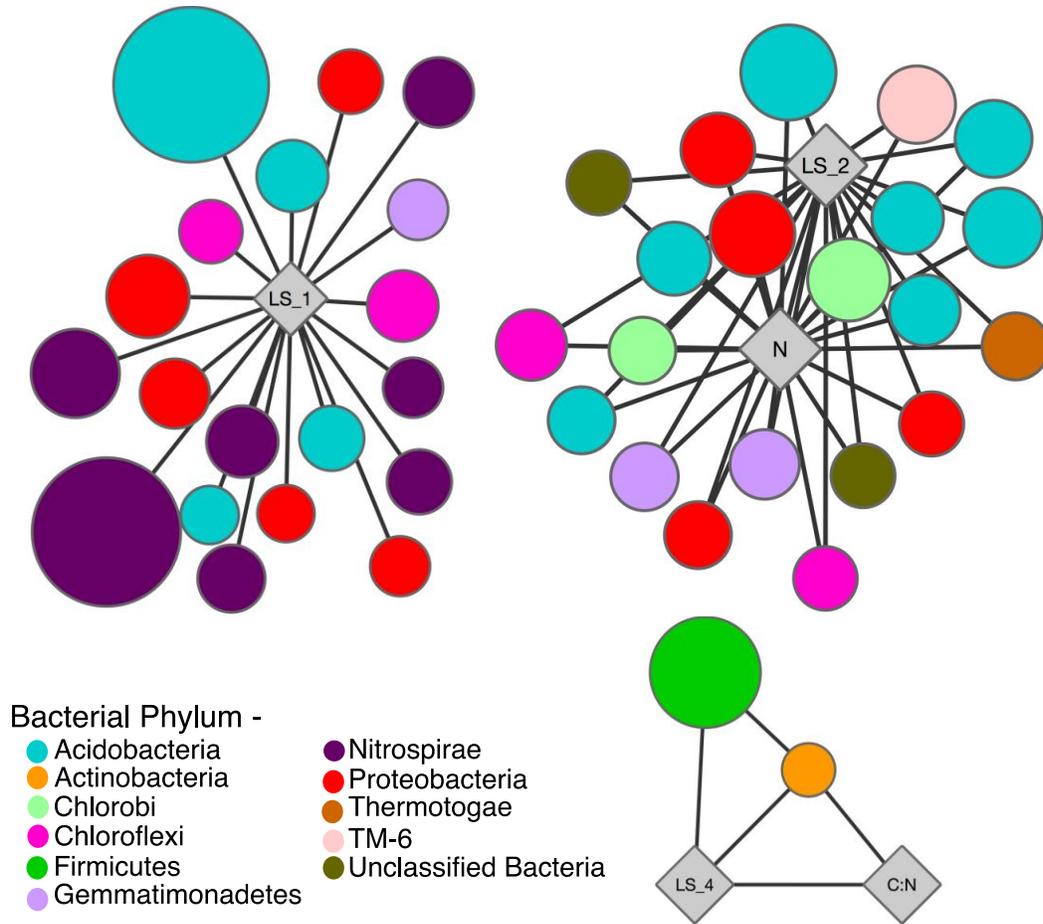


Figure 3.8 – Correlation network of bacteria and lithostratigraphic units. The size of the bacterial nodes is proportional to the relative abundance of the sequence in the sample where it is most abundant (ranging from 0.01-1% of the community).

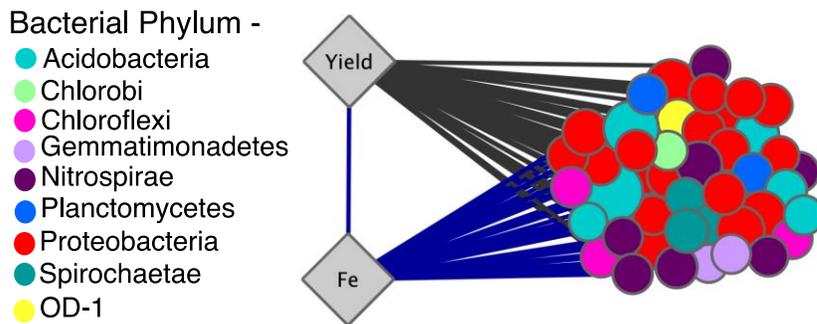


Figure 3.9 – Negative correlation between iron concentration and DNA yield for select taxa. Positive correlations in black and negative correlations in black. The size of the bacterial nodes is proportional to the relative abundance of the sequence in the sample where it is most abundant (ranging from 0.01-1% of the community).

CROMO (46), however this study found them to be abundant and diverse within core samples. The core-enriched sequences identified in this study represent a diversity of archaea and bacteria, not previously within the serpentinite environment, containing taxa important in the cycling of nitrogen-containing molecules and methane. Contrary to other sites of continental serpentinization (28, 45), fluids at CROMO contain elevated nitrogen species, particularly ammonium (10). The findings here indicate that bacteria and archaea associated with cores aid in the subsurface nitrogen cycling, through ammonium- and nitrite-oxidation within the serpentinite subsurface environment, which could supply this essential nutrient to the fluid-associated community.

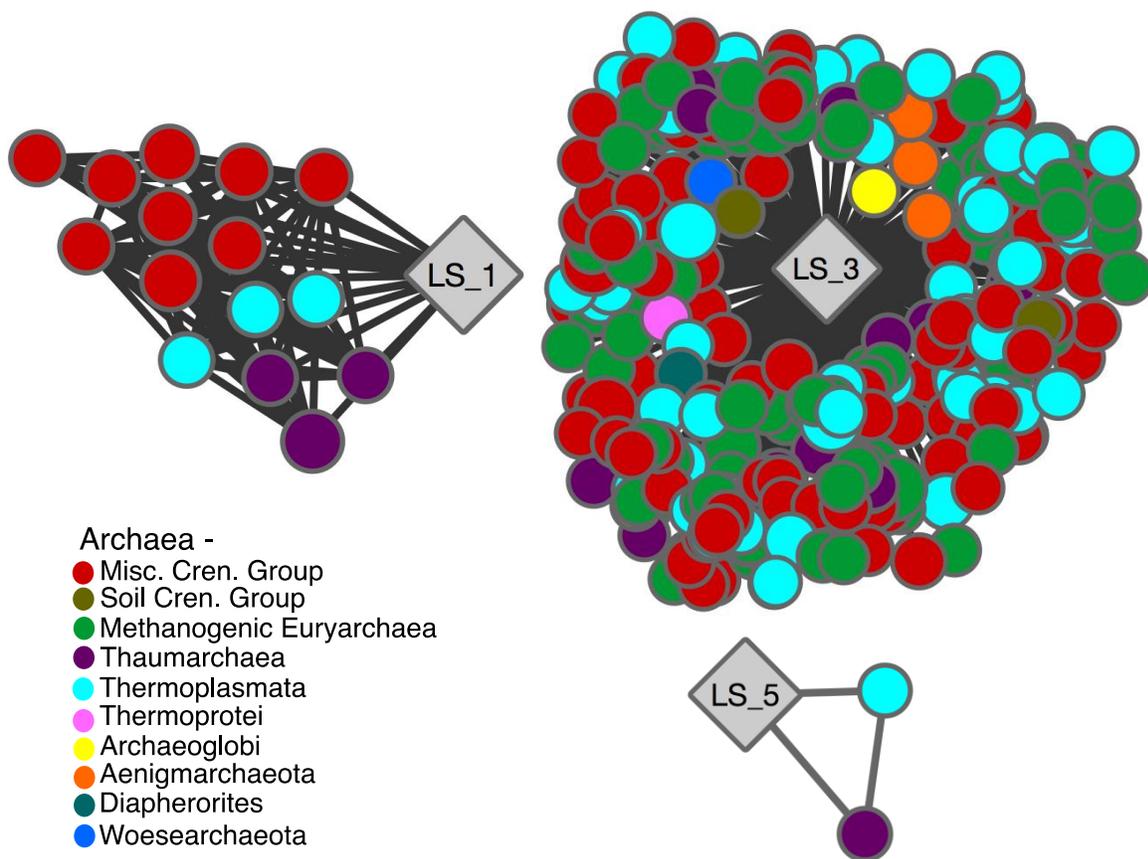


Figure 3.10 – Correlation network of archaea and lithostratigraphic units. The size of the bacterial nodes is proportional to the relative abundance of the sequence in the sample where it is most abundant (ranging from 0.01-2% of the community).

While methanogens have not previously been detected in fluid samples from CROMO (46), methane isotopologue data from the site suggests that the methane at CROMO is from a combination of thermogenic and microbial sources (49). Here methane-cycling archaea associated with core samples from QV were detected. These data suggest that particle-associated microbial communities are contributing to the cycling of nitrogen and methane within the serpentinite subsurface environment and that more complex biogeochemical cycles exist than can be seen from studying fluids alone.

The correlation analyses suggest a relationship between the lithostratigraphic data and the core-enriched taxa. Of particular intrigue was the shift from bacterial- to archaeal-dominated communities in the regions with magnetite-bearing serpentine both without clay (LS_3) and with clay (LS_5), though not in magnetite-bearing serpentine with clay, albite, and quartz (LS_4; Table 3.1; Fig. 3.9). These data suggest that the specific mineralogy of this region may favor bacterial taxa over archaeal.

Within the subsurface, fluids and solid-substrates interact naturally, and therefore, both must be thought of as components of the environment. Previous studies have focused solely on the microorganisms detected within fluid samples and neglected the components of the communities associated with subsurface solids. The data presented here, combined with the study of serpentinite fluids from CROMO (46), provide a more comprehensive depiction of the serpentinite subsurface environment than has been seen in previous studies (4, 44, 45) and indicates that core-associated communities are distinct within different lithologies and could be contributing to cryptic biogeochemical cycling within the serpentinite subsurface environment.

Materials and Methods

Site Description and Sample Collection

The Coast Range Ophiolite Microbial Observatory (CROMO) was established in August 2011, when eight wells were drilled into an actively serpentinizing subsurface environment at the UC Davis Donald and Sylvia McLaughlin Natural Reserve in Lower Lake, California. Details of the drilling operations can be found in Cardace *et al* (7). Briefly, two main wells, CSW1.1 and QV1.1, were drilled 1.4 km apart to depths of 31 m and 45 m, respectively, using HQ wireline coring with an inner diameter of 63.5 mm. To mitigate the potential of contamination from drill fluids used during drilling, as has been seen in ocean drilling studies (21), clean water (filtered through 0.1µm filter and ozonated) was used as drilling fluid and fluorescent microbead tracers were included to track contamination of exterior portions of the core. Cores from these wells were cataloged and preserved for complimentary mineralogical, geochemical, and microbiological analyses. Microbiological samples were wrapped in combusted aluminum foil, placed in sterile bags, frozen with liquid nitrogen and stored at -80°C until DNA extraction.

DNA Extraction

Thawed core and soil material was homogenized using autoclaved and ethanol sterilized percussion mortars and ceramic mortars and pestles. DNA was extracted from 2 x 10 g of each homogenized core using the MoBio PowerMaxSoil Kit (Carlsbad, CA), per the manufacturers instructions. The resulting DNA suspensions were pooled from replicate extractions and concentrated in an Amicon Ultra-2 Centrifugal Filter Unit with Ultracel-30 membrane (Millipore, Darmstadt, Germany) to a volume of 50 µL. Filtering of the well fluids and DNA extractions of the filters were conducted as described previously (4). DNA was quantified using the High

Sensitivity reagents for a Qubit® 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA), which has a detection limit of 0.1ng/μL DNA .

Quantitative-PCR

Copies of the 16S rRNA gene were determined by quantitative polymerase chain reaction (q-PCR) using domain-specific primers targeting the V6V4 hypervariable region of the 16S rRNA gene for archaea and bacteria, respectively (43). Samples were run on a BioRad C-1000 thermo-cycler with a q-PCR module using the SsoAdvanced SybrGreen Assay (Biorad, Hercules, CA).

16S rRNA Gene Sequencing and Data Analysis

Purified DNA samples from core, soils, and well fluids were submitted to the Josephine Bay Paul Center at the Marine Biological Laboratory for sequencing of the V4V5 region of the 16S rRNA gene on an Illumina MiSeq instrument as part of the Census of Deep Life project (29). The paired-end reads were merged and subjected to MBL's post-processing quality control for removal of low quality reads and chimera checking (18). The samples yielded between 8,493 and 246,345 merged, quality-filtered sequences each. Any sample that yielded less than 5,000 sequences was considered a failed sequencing run and not included in the analysis.

Using Mothur (39), the sequence data was clustered into unique sequences and assigned taxonomy compared by alignment to the SILVA database (v119 for bacteria and v123 for archaea; 31). Significant differences in the abundances of unique sequences between groups of samples were tested with the aid of the Phyloseq (25) and DESeq (22) packages in R. A complete list of R commands for an example comparison of two groups of samples is provided in Appendix A. For DESeq comparisons, samples were pooled into categories, such as “core”, “soil”, “fluids”, “CSW-core”, or “QV-core”. It was determined that QV30R and QV7R were

outliers (Fig. 5) and were therefore removed from the “core” category for comparisons between categories and were instead analyzed separately as individual core samples against the fluids and soils. Sequences were only categorized as “core-enriched” if they were enriched compared to both fluids and soils.

Mineralogical and Geochemical Analyses

Subsamples of each core were analyzed to obtain total nitrogen (TN) and total organic carbon (TOC), as well as $\delta^{13}\text{C}$ of TOC by elemental analysis isotope ratio mass spectroscopy (EA-IRMS). Core samples were acidified via the adapted capsule method (5) prior to analysis on a Carlo Erba 1108 elemental analyzer coupled to a ThermoFinnigan Delta Plus XL isotope ratio mass spectrometer. Descriptions of the core lithostratigraphy, down-hole logging, X-ray diffraction (XRD), and X-ray fluorescence (XRF) methods can be found in refs 7 and 8.

Statistical Analyses

A matrix containing environmental geological and geochemical data and relative abundances of core-enriched sequences for each core sample was used as input for pairwise Pearson’s correlation analysis, computed with the `rcor.test` function in the R package `lmt` (36). The false-discovery rate (q-value) was computed for the distribution of Pearson p-values to account for multiple tests. Pairwise correlations with both p- and q-values of <0.05 were considered significant and included in network analyses. Network models of significant correlations were created using Cytoscape v2.8.3 (38).

REFERENCES

REFERENCES

1. Brazelton WJ, Schrenk MO, Kelley DS, Baross JA (2006) Methane- and sulfur-metabolizing microbial communities dominate the Lost City hydrothermal field ecosystem. *Appl Environ Microbiol* 72:6257-6270.
2. Brazelton WJ, Mehta MP, Kelley DS, Baross JA (2011) Physiological differentiation within a single-species biofilm fueled by serpentinization. *mBio* 2: doi:10.1128/mBio.00127-11.
3. Brazelton WJ, Nelson B, Schrenk MO (2012) Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2:268.
4. Brazelton WJ, Morrill PL, Szponar N, Schrenk MO (2013) Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl Environ Microb* 79:3906-3916.
5. Brodie, C. R., M. J. Leng, J. S. L. Casford, C. P. Kendrick, J. M. Lloyd, Z. Yongqiang, and M. I. Bird. 2011. Evidence for bias in C and N concentrations and delta13C composition of terrestrial and aquatic organic materials due to per-analysis acid preparation methods. *Chemical Geology*. 282: 67-83.
6. Buckley DH, et al. (2006) Diversity of *Planctomycetes* in soil in relation to soil history and environmental heterogeneity. *Appl Environ Microb* 72:4522-4531.
7. Cardace D, et al. (2013) Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Scientific Drilling* 16:45-55.
8. Carnevale, DC (2013) Carbon sequestration potential of the Coast Range Ophiolite in California. (Master's thesis). Retrieved from: Open Access Master's Theses. Paper 46.
9. Cleaves HJ, et al. (2011) The adsorption of short single-stranded DNA oligomers to mineral surfaces. *Chemosphere* 88:1560-1567.
10. Crespo-Medina M, et al. (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604.
11. Daae FL, et al. (2013) Microbial life associated with low-temperature alteration of ultramafic rocks in the Leka ophiolite complex. *Geobiology* 11:318-339.
12. Ehrich S, et al. (1995) A new obligately chemolithoautotrophic, nitrite-oxidizing bacterium, *Nitrospira moscoviensis* sp. nov. and its phylogenetic relationship. *Arch Microbiol* 164:16-23.

13. Fuerst JA, et al. (2006) Anammoxosomes of anaerobic ammonium-oxidizing Planctomycetes. *Microbiol Monogr* 2:260-283.
14. Früh-Green GL, et al. (2004) Serpentinization of oceanic peridotites: Implications for geochemical cycles and biological activity. *In: The Subseafloor Biosphere at Mid-Ocean Ridges.* (American Geophysical Union, Washington, D.C.) pgs. 119-136.
15. Hallbeck L, Pedersen K (2014) The family *Gallionellaceae*. *The Prokaryotes*, (Springer-Verlag, Berlin) pp. 853-858.
16. Hanada S (2014) The phylum Chloroflexi, the Family *Chloroflexaceae*, and the related phototrophic families *Oscillochloridaceae* and *Roseiflexaceae*. *In: The Prokaryotes – Other Major Lineages of Bacteria and Archaea.* (Springer-Verlag, Berlin).
17. Hurt RA, et al. (2014) Improved yield of high molecular weight DNA coincides with increased microbial diversity access from iron oxide cemented sub-surface clay environments. *PLOS One* 9:e102826.
18. Huse SM, et al. (2014) VAMPS: a website for visualization and analysis of microbial population structures. *BMC Bioinformatics* 15:41.
19. Idris R, et al. (2004) Bacterial communities associated with flowering plants of the Ni hyperaccumulator *Thalpi goesingense*. *Appl Environ Microb* 70:2667-2677.
20. Katsaveli K, Vayenas D, Tsiamis G, Bourtzis K (2012) Bacterial diversity in Cr(VI) and Cr(III)-contaminated industrial wastewaters. *Extremophiles* 16:285-296.
21. Lever MA, et al. (2006). Trends in basalt and sediment core contamination during IODP Expedition 301. *Geomicrobiol J* 23:517-530.
22. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol* 15:550.
23. Lückner S, Daims H (2014) The family *Nitrospinaceae*. *In: The Prokaryotes*, (Springer-Verlag, Berlin) pp. 231-237.
24. McCollom TM, Bach W (2009) Thermodynamic constraints on hydrogen generation during serpentinization of ultramafic rocks. *Geochim Cosmochim Acta* 73:856-875.
25. McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS One* 8:e61217.
26. McMurdie PJ & Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531. doi:10.1371/journal.pcbi.100353126.

27. Megoni A, et al. (2004) Genetic diversity of bacterial communities of serpentine soil and of rhizosphere of the nickel-hyperaccumulator plant *Alyssum bertolonii*. *Microbial Ecol* 48:209-217.
28. Morrill PL, et al. (2013) Geochemistry and geobiology of a present day serpentinization site in California: the Cedars. *Geochim Cosmochim Acta* 109:222-249.
29. Morrison HG, Grim SL, Vineis JH, Sogin ML (2013) 16S amplicon fusion primers and protocol for Illumina platform sequencing. Figshare. Available online at: <http://dx.doi.org/10.6084/m9.figshare.833944>
30. Oline DK (2006) Phylogenetic comparisons of bacterial communities from serpentine and nonserpentine soils. *Appl Environ Microb* 72:6965-6971.
31. Pal A, Wauters G, Paul AK (2007) Nickel tolerance and accumulation by bacteria from rhizosphere of nickel hyperaccumulators in serpentine ecosystem of Andaman, India. *Plant Soil* 293:37-48.
32. Prosser JI, Head IM, Stein LY (2014) The family *Nitrosomonadaceae*. In: *The Prokaryotes*, (Springer-Verlag, Berlin) pp. 901-918.
33. Pruesse E, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188-7196.
34. Rajkumar M, Prasad MNV, Freitas H, Ae N (2009) Biotechnological applications of serpentine soil bacteria for phytoremediation of trace metals. *Crit Rev Biotechnol* 29:120-130.
35. Rastogi G, et al. (2010) Microbial and mineralogical characterization of soils collected from the deep biosphere of the former Homestake Gold Mine, South Dakota. *Microb Ecol* 60:539-550.
36. Rizopoulos D (2006) ltm: an R package for latent variable modeling. *J Stat Softw* 17:5.
37. Salter SJ, et al. (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology* 12:87.
38. Shannon P, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498-2504.
39. Schloss PD, Westcott SL (2011) Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl Environ Microb* 77:3219-3226.

40. Schrenk MO, Kelley DS, Bolton SA, Baross JA (2004) Low archaeal diversity linked to seafloor geochemical processes at the Lost City Hydrothermal Field, Mid-Atlantic Ridge. *Environ Microbiol* 6: 1086-1095.
41. Schrenk MO, Brazelton WJ, Lang SQ (2013) Serpentinization, carbon, and deep life. *Rev Mineral Geochem* 75:575-606.
42. Schwartz E, et al. (2006) The H₂-metabolizing prokaryotes. *In: The Prokaryotes*, (Springer, New York, NY) pp. 496-563.
43. Sogin ML, et al. (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. USA*. 103:12115-12120.
44. Suzuki S, et al. (2013) Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proc Natl Acad Sci USA* 110:15336-15341.
45. Tiago I, Veríssimo A (2013) Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environ Microbiol* 15:1687-1706.
46. Twing KI, et al. (In Submission) Serpentinizing fluids from a subsurface observatory harbor extremely low diversity microbial communities adapted to high pH. *ISME J*.
47. Vermeer AWP, van Riemsdijk WH, Koopal LK (1998) Adsorption of humic acid to mineral particles: Specific and electrostatic interactions. *Langmuir* 14:2810-2819.
48. Walker RB (1954) Factors affecting plant growth on serpentine soils. *Ecology* 35:259-266.
49. Wang DT, et al. (2015) Nonequilibrium clumped isotope signals in microbial methane. *Science* 348:428-431.

CHAPTER 4

Microbe-mineral interactions in the serpentinite subsurface environment

Abstract

Serpentinization creates a challenging habitat for microorganisms, with high pH fluids depleted in terminal electron acceptors and dissolved inorganic carbon (DIC). Findings from Chapter 3 suggest that there are specialized microbial communities associated with minerals in the serpentinite subsurface environment compared to fluids. *In situ* colonization devices containing mineral substrates were deployed into microbial observatory wells to test whether microbes in the serpentinite subsurface were capable of utilizing inorganic carbon in calcite and ferric iron in magnetite to make up for the lack of DIC and electron acceptors in surrounding fluids. Mineral substrates deployed into the most extreme serpentinite well (CSW1.1, pH 12.5) selected for unique microbial communities, compared with well fluids collected at the start of the experiment. Devices containing calcite resulted in an abundance of *Syntrophomonas* (a clostridium also found in starting fluids) and *Deinococcus*. Magnetite led to an increase in diversity, including Alphaproteobacteria, Gammaproteobacteria, and Actinobacteria, in addition to the Betaproteobacteria commonly found in CSW1.1 fluids. It is not clear, however, if or how these enriched microbial communities are utilizing the calcite and magnetite. These data suggest that specific mineral composition may impact microbial diversity, particularly in the harshest environmental conditions.

Introduction

Most microbial metabolisms (catabolic) are based on redox reactions, for which they need both electron donors and electron acceptors. Some microbes are able to metabolize chemicals found within mineral forms to obtain energy through chemolithotrophy (25). Microbes

can form complex communities and micro-niches within biofilms on hard surfaces (7). Given that the subsurface environment is rock-hosted, there is a great deal of interest in understanding microbe-mineral interactions. Previous studies have yielded successful results with *in situ* colonization of mineral substrates deployed into the International Ocean Discovery Program (IODP) boreholes in the basaltic ocean crust in borehole packers (otherwise known as circulation obviation retrofit kits or CORKS; e.g. 9, 20, 21) or as mineral chips in sub-glacial systems (17), suggesting that mineral composition leads to differences in microbial biomass and community composition.

The geochemical process of serpentinization creates a challenging habitat for life, not only in terms of the extreme pH of fluids, but also with limited availability of dissolved inorganic carbon (DIC) and scarcity of electron acceptors. The high pH shifts the DIC species in fluids to predominantly carbonate, which precipitates out of solution in the presence of abundant Ca^{2+} ions, to form calcium carbonate (1, 11). Characterization and genomic analyses of an isolate of *Serpentinomonas*, the Betaproteobacterial genus that dominates serpentinite fluids, indicated that the organism is capable of autotrophic growth on calcium carbonate (31). The abundant H_2 and methane frequently generated during serpentinization are potential electron donors for microbial metabolisms, but this potential energy can only be harnessed if they can be paired with electron acceptors. There is a dearth of electron acceptors in serpentinite fluids, relative to the abundance of electron donors, particularly in the anoxic end-member serpentinite fluids. While not detected in the serpentinite subsurface environment before, iron-reducing bacteria, such as *Shewanella putrefaciens* are able to reduce ferric iron found in the mineral magnetite (12).

This study explores the potential for microorganisms native to the serpentinite subsurface environment to utilize solid inorganic carbon and the potential electron acceptor ferric iron from mineral phases calcite and magnetite, respectively, to make up for the lack of these necessary chemicals in serpentinite fluids. *In situ* colonization devices containing mineral substrates were incubated in four wells spanning a range of geochemical gradients at the Coast Range Ophiolite Microbial Observatory (CROMO; 5) over the course of a year and the resulting microbial communities were analyzed by 16S rRNA gene sequencing.

Results and Discussion

In Situ Colonization Devices and CROMO Wells

Colonization devices containing mineral substrates included borosilicate glass as an inert substrate for control, magnetite (as a source of ferric iron), or calcite (as a source of carbon) were deployed near the bottom (~15 mbs) of the CSW1.1, CSW1.3, QV1.1, and QV1.2 wells for one year to assess the microbe-mineral interactions within the well. DNA was extracted from the mineral substrates and used for 16S rRNA gene quantitative-PCR and amplicon sequencing. For experimental replication, the devices were deployed in triplicate, in alternating order (Fig. 4.1). The bladder pump used to sample well fluids sits right above the bottom of the well and the *in situ* devices were deployed directly above it (Fig. 4.1D). Since the pump pulls water from above,

Table 4.1 - Geochemical parameters of well fluids measured at in situ device deployment and retrieval.

Well	Depth (mbs)	Sampling Trip	pH	ORP (mV)	Temp (°C)	Conductivity ($\mu\text{S}/\text{cm}^3$)	DO (mg/L)
CSW1.1	19.5	Aug. 2013	12.4	-254	15.8	4460	0.10
		July 2014	12.3	-297	16.3	4455	0.43
CSW1.3	23.2	Aug. 2013	10.3	-234	17.2	4610	0.03
		July 2014	10.3	-328	15.6	4597	0.19
QV1.1	23.0	Aug. 2013	11.6	-92	16.4	2610	0.15
		July 2014	11.5	-239	17.2	2841	0.27
QV1.2	14.9	Aug. 2013	9.1	-73	17.1	2800	0.22
		July 2014	9.4	-168	16.9	3042	0.20

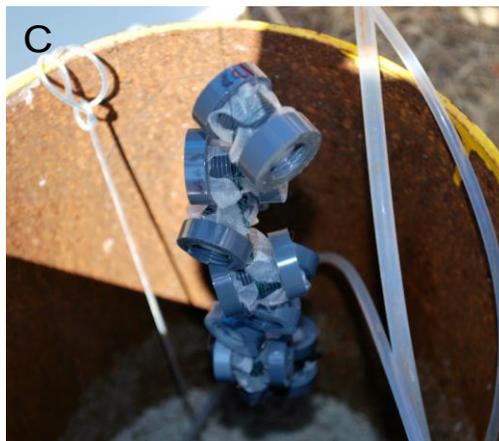
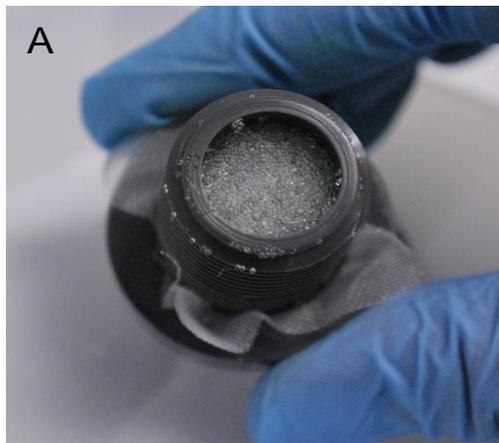
the devices were deployed two full days after sampling the wells in August 2013, so as to allow for fluid recharge in the well. A preliminary time-series study of *in situ* colonization at CROMO suggested that one year was optimum for microbial growth (data not shown). Table 4.1 summarizes the environmental parameters of the wells waters at the beginning (referred to henceforth as 2013 fluids) and end (referred to henceforth as 2014 fluids) of the experiment of the one-year colonization experiment, which are consistent with prior measurements at of these wells at CROMO (Chapter 2; 33).

16S rRNA Gene Abundance

The 16S rRNA gene abundances of microbial communities from the *in situ* devices were determined by bacteria-specific quantitative-PCR (Fig. 4.2). There were significantly more 16S rRNA gene copies per g substrate in the glass-treatment than on the magnetite or calcite-treatments in CSW1.1 ($F_{2,6} = 6.35$, p-value = 0.03) and fewer gene copies on magnetite than either glass or calcite in QV1.1 ($F_{2,6} = 6.93$, p-value = 0.03). No other significant differences in gene copy abundance were seen either within wells or between wells (Fig. 4.2), likely due to the high variation between replicates (*e.g.* CSW1.3 calcite or QV1.2 magnetite).

Bacterial Community Composition

Across all wells, magnetite yielded communities with greater diversity than did calcite (Fig. 4.3). Glass, on the other hand, yielded communities of variable diversity, with the most diverse glass samples coming from QV1.1 and QV1.2 and the least diverse glass samples coming from CSW1.1 (Fig. 4.3). Microbial communities associated with the *in situ* colonization devices were more similar to communities from CROMO well fluids than from CROMO cores or soils (Fig. 4.4). Overall, community composition was affected to a much greater extent by which well the colonization devices were deployed in than by the specific mineral matrix within the device,



D

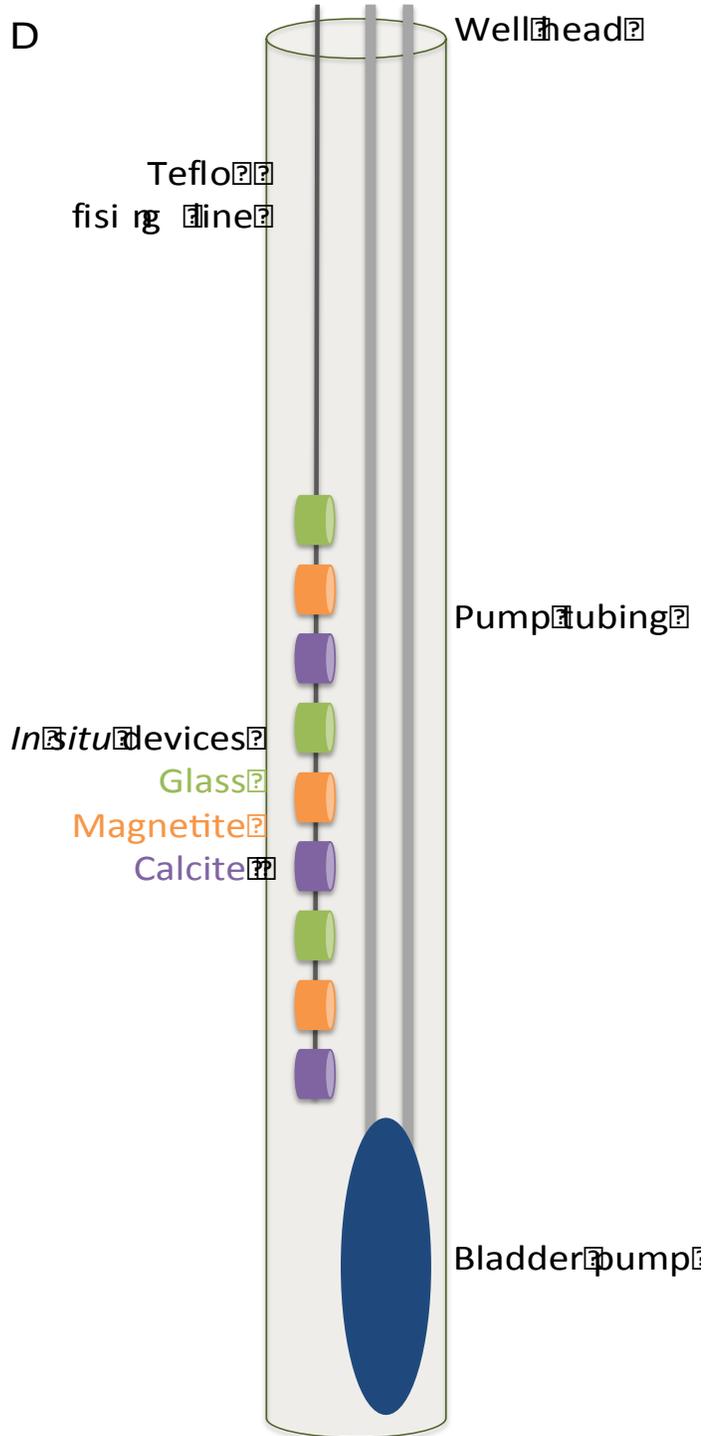


Figure 4.1 – In situ colonization devices and schematic of deployment down well. A) Open devices filled with borosilicate glass beads for inert substrate control. B) Fully assembled devices prior to autoclave sterilization. C) Devices tied together with Teflon fishing line being deployed into QV1.1. D) Diagram of down-well deployment. The *in situ* devices were tied together with Teflon fishing line and secured to the top of the well-head and deployed just above the bladder pump, which sits at the bottom of the well.

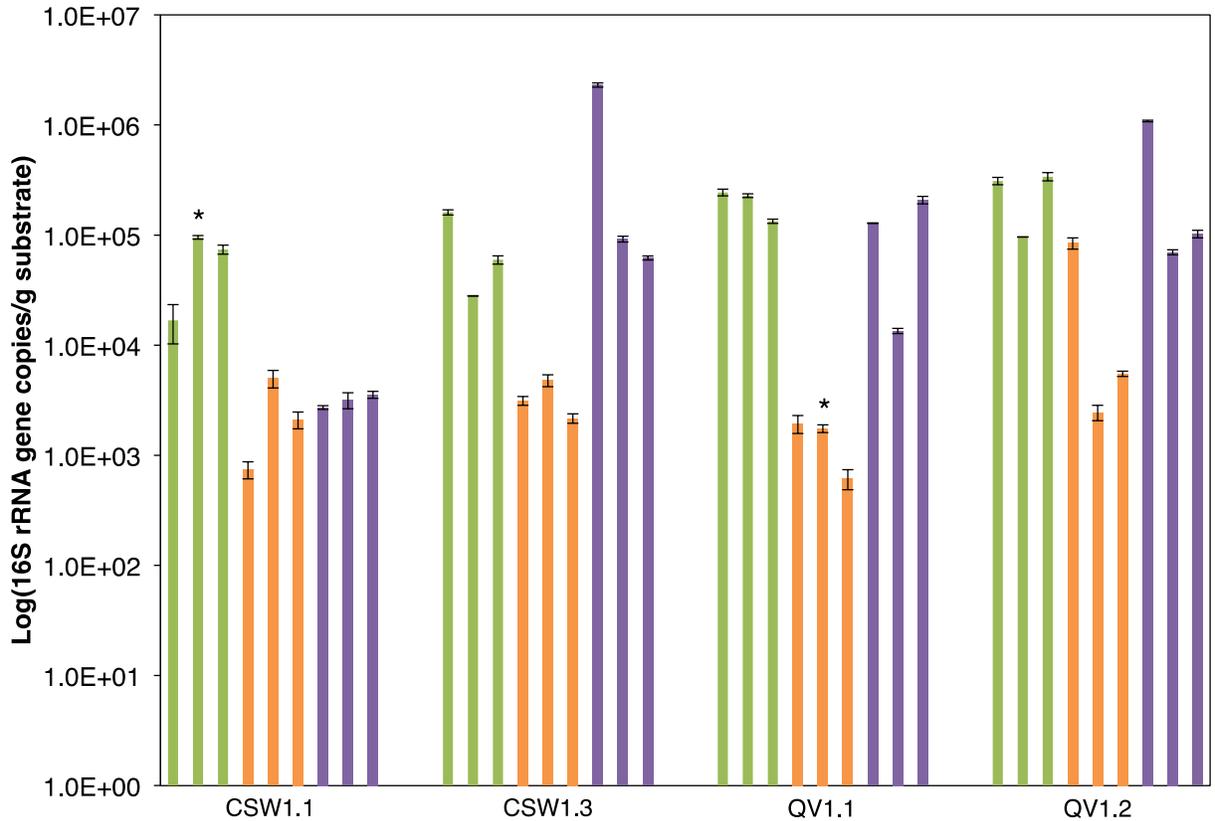


Figure 4.2 – 16S rRNA gene copy abundance per gram of substrate, as determined by q-PCR. Mineral-substrates are color-coded: glass, green; magnetite, orange; calcite, purple. Significant variations in 16S rRNA gene copy number, as determined with ANOVA, are represented with asterisks.

with all the communities from QV1.2 and CSW1.3 treeing together, respectively (Fig. 4.4). Samples from the extreme pH wells, CSW1.1 (pH 12.4) and QV1.1 (pH 11.6) were interspersed with each other in the community similarity dendrogram (Fig. 4.4). The exceptions to this are two of the magnetite treatments CSW1.1, which were more similar to communities on magnetite from CSW1.3 than other CSW1.1 or QV1.1 communities (Fig. 4.4). With regard to mineral substrates, glass and calcite tree together and separately from magnetite for all of the well. These data suggest that, similar to the findings in from CROMO fluids, fluid source (and pH) has an impact on microbial community composition (33), with mineral substrate being secondary.

Similar to previous studies of the serpentinite subsurface, and as shown for CROMO in Chapter 2 (33), fluids were dominated by Comamonadaceae (Betaproteobacteria) and Clostridia

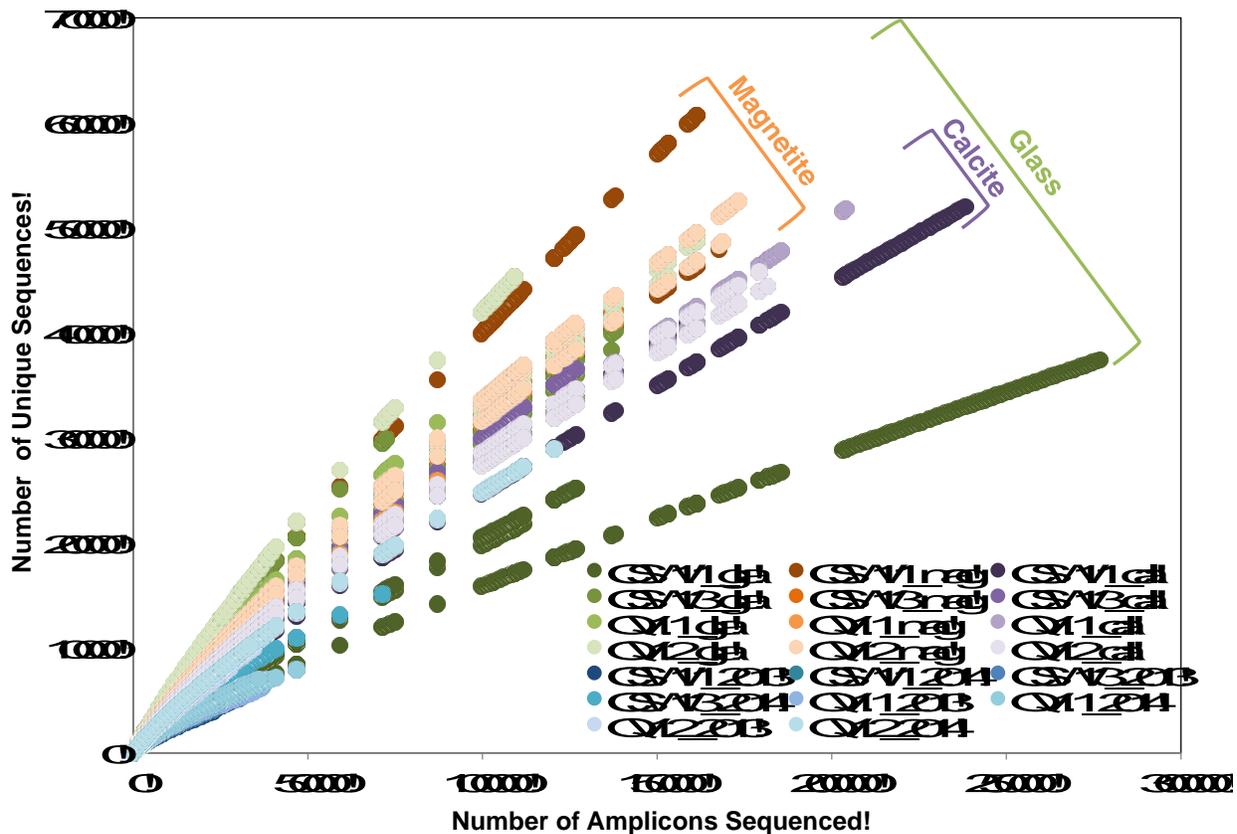


Figure 4.3 – Rarefaction curve of unique sequences from in situ colonization device and well fluid sequencing

sequences (Fig. 4.5). CSW1.1 fluids from 2013 (start of the experiment) contained an abundance of Betaproteobacteria and QV1.1 fluids contained a mix of different Clostridia species (Fig. 4.5; 33). Similarly, the lower pH wells were more diverse than the high pH wells (Fig. 4.5). In CSW1.1, there is a shift in community composition between 2013 fluids (start of the experiment) and 2014 fluids (end of the experiment), namely the spike in *Deinococcus-Thermus* sequences, that can perhaps be explained by sloughing off from glass and calcite colonization devices or the stimulated growth of those organisms within the fluids due to the presence of calcite. Similarly, QV1.1 end fluids contained Bacteroidetes sequences that were not apparent in the starting fluid community. The community composition on the mineral substrates from CSW1.1 look distinct

from starting fluids. In contrast, the fluid and mineral samples from other wells (CSW1.3, QV1.1, QV1.2) share similar community composition at the family level (Fig. 4.5).

Mineral-Enriched Bacteria

Due to the difficulty in determining the source of sequences in comparative datasets, the statistical method of DESeq (13), modeled after differential expression in RNA studies (16), was used to determine sequences that were statistically enriched in a particular sample type compared to the others (Table 4.2). Given the potential for transfer of organisms between fluids at the end of the experiment and the mineral substrates, either from microbes being knocked off the *in situ* colonization devices or stimulation of microbial growth within fluids by the addition of minerals, only sequences from fluids from the start of the experiment, collected in August 2013, were used as the ‘fluids’ group for comparisons with mineral substrate communities.

Communities associated with mineral substrates incubated in CSW1.1 were the most divergent from well fluids compared to the other three wells investigated (Fig. 4.5) and therefore more sequences were determined to be enriched for different minerals within that well (Fig. 4.6).

Table 4.2 – DESeq comparisons for determining fluid- and mineral-specific taxa.

Enriched Community	Sample Group 1	Sample Group 2
CSW1.1 Fluid start	CSW1.1 2013 fluids	CSW1.1 2014 fluids
	CSW1.1 2013 fluids	CSW1.1 <i>in situ</i> devices (pooled)
	CSW1.1 2013 fluids	CSW1.1 glass devices
	CSW1.1 2013 fluids	CSW1.1 magnetite devices
	CSW1.1 2013 fluids	CSW1.1 calcite devices
Glass	CSW1.1 glass devices	CSW1.1 2013 fluids
	CSW1.1 glass devices	CSW1.1 magnetite devices
	CSW1.1 glass devices	CSW1.1 calcite devices
Magnetite	CSW1.1 magnetite devices	CSW1.1 2013 fluids
	CSW1.1 magnetite devices	CSW1.1 glass devices
	CSW1.1 magnetite devices	CSW1.1 calcite devices
Glass	CSW1.1 calcite devices	CSW1.1 2013 fluids
	CSW1.1 calcite devices	CSW1.1 glass devices
	CSW1.1 calcite devices	CSW1.1 magnetite devices

* Comparisons were repeated for all four wells

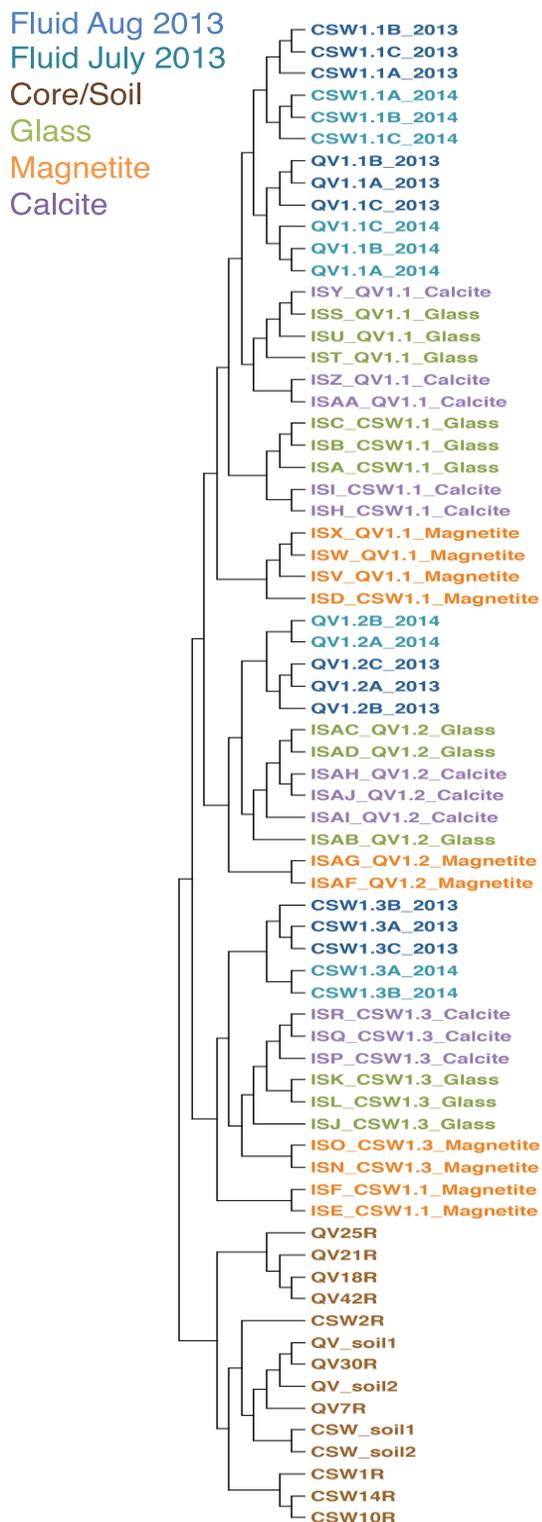
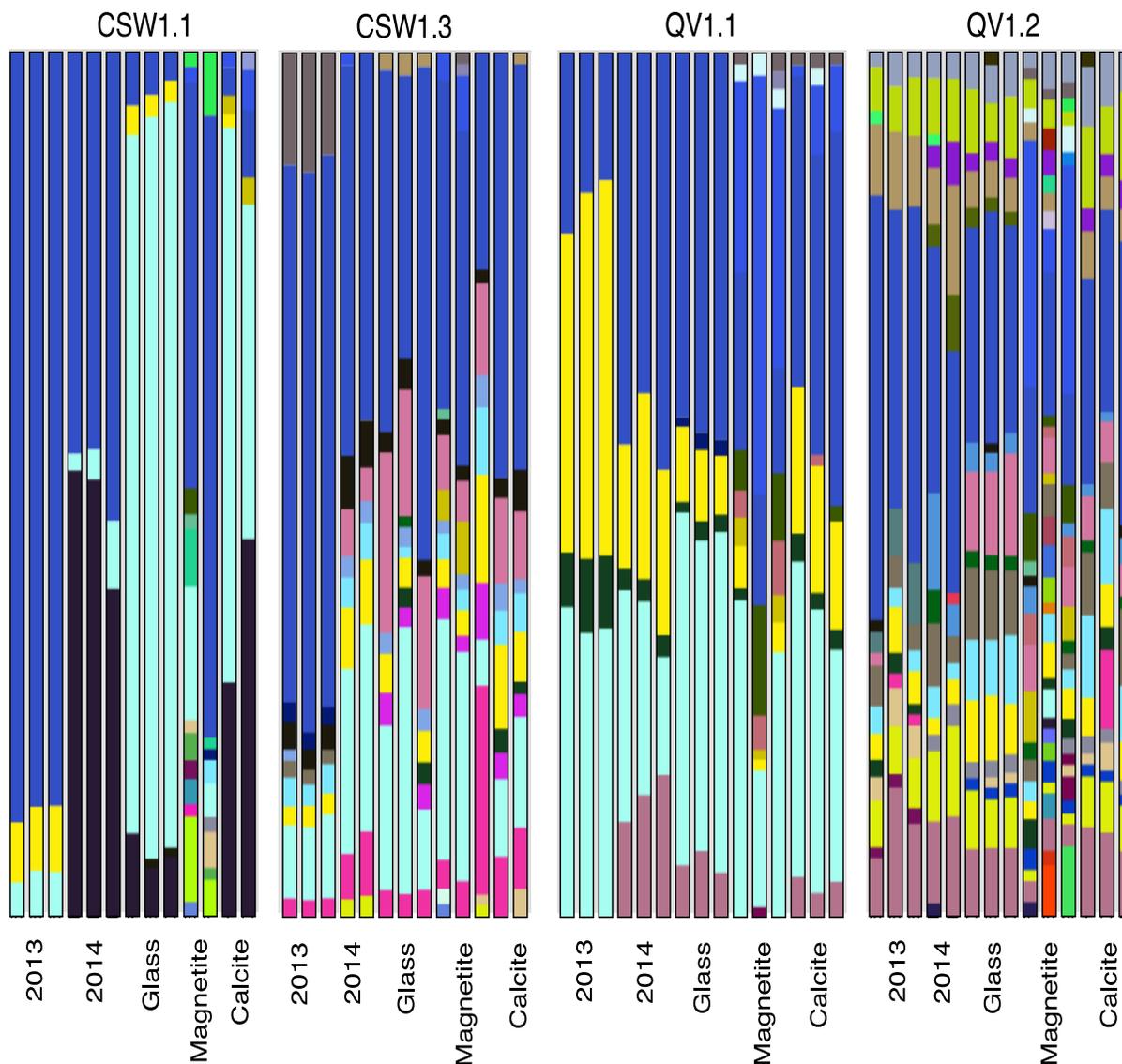


Figure 4.4 – Community similarity dendrogram of unique sequences from in situ colonization devices, well fluids, and core material. The sorclass dissimilarity index was used to make the dendrogram in Mothur (24).

The fluid-enriched taxa in CSW1.1 are consistent with previous findings from the serpentinite environment (3, 4, 29, 31) and that well, in particular (Chapter 2; 7, 32), with a dominance of Comamonadaceae and Clostridia (Table 4.3). For the fluid samples (and specific sampling dates) described in Chapter 2, these taxa made up greater than 70% of the microbial communities from CSW1.1 (33), however only 20% of the community was identified as “fluid-enriched” in this study (Fig. 4.6A). This could either be due to a shift in community structure or, more likely, sequences belonging to these major groups were also found in *in situ* devices and therefore not determined to be strictly “fluid-enriched”.

Magnetite-enriched taxa made up 25% of the magnetite-filled *in situ* devices (Fig. 4.6A) and contained an abundance of Comamonadaceae sequences (Table 4.4). It should be noted that these are not the same Comamonadaceae that make up the fluid-enriched sequences, as they were not detected in the well fluids (Table 4.4). Additionally, members of the Sporichthyaceae (class Actinobacteria) and LD12-freshwater-group of the SAR11 were detected (Table 4.4). None of the taxa enriched in the magnetite *in situ* devices are known to be capable of iron reduction (19, 36).

The calcite-enriched microbial communities consisted of Synthrophomonadaceae (class Clostridia) and Deinococci (Table 4.5) and made up 50-70% of the communities within calcite devices, but also 10-20% of glass devices and 50-70% of fluids from the end of the experiment (Fig. 4.6A). Given the abundance of these organisms in fluids from July 2014, it is possible that they naturally occurred in the fluids and were just trapped in the *in situ* devices, instead of being enriched by them. However, given that neither taxa has been detected in appreciable abundances in CROMO fluids before and the stability of microbial communities at CROMO over time (Chapter 2; 33), it is more likely that they were enriched by the addition of the calcite substrate



Family-level

- Microbacteriaceae (Actinobacteria)
- Sporichthyaceae (Actinobacteria)
- Bacteroidales
- Deinococcaceae
- Clostridiales (Clostridia)
- Peptostreptococceae (Clostridia)
- Syntrophomonadaceae (Clostridia)
- Thermoanaerobacterales (Clostridia)
- Erysipelotrichaceae (Firmicutes)
- Caulobacteraceae (A-pb)
- Methylocystaceae (A-pb)
- SAR11 (A-pb)
- Burkholderiaceae (B-pb)
- Comamonadaceae (B-pb)
- Methylophilaceae (B-pb)
- Methylococcaceae (G-pb)

Figure 4.5 – Community diversity bar chart for CROMO fluids and in situ colonization devices. Taxonomy is defined at the family level and chart was made with data from VAMPS (10).

or sloughed off into the fluids during sampling. Members of the Syntrophomonadaceae can oxidize carbon monoxide (26, 27, 28) and are generally found in syntrophic relationships with methanogens (28). It should be noted that the DNA from *in situ* devices was only sequenced for bacterial 16S rRNA genes and therefore archaea were not looked for in this study, so whether or not a syntrophic relationship was taking place in this instance cannot be addressed. *Deinococcus-Thermus* have been isolated from arid deserts and exhibit adaptations to resist challenges, such as UV-radiation and desiccation (14). While it is not clear what their role in the serpentinite system is, they have been previously detected in low abundances (0.5-1.5%) of 16S rRNA clone and amplicon libraries from fluids (32) and carbonate structures (2, 35, 22) within the serpentinite environment. It should be noted that both of these organisms are known spore-formers and their abundance in the *in situ* devices and 2014 fluids compared to 2013 fluids suggests either they grow within the samples or that the minerals help trap rare spore and recalcitrant microbes within the system.

Fluid-enriched sequences from CSW1.3 belonged to the Comamonadaceae and Gammaproteobacteria (Table 4.3), consistent with findings in Chapter 2 (33). No calcite-enriched taxa were detected in CW1.3 (Fig. 4.6B), a well with pH 10.3 fluids containing low DIC (172 μ M; 33). The magnetite-enriched bacterial community made up ~3% of the magnetite *in situ* devices communities (Fig. 4.6B) and contained a variety of Actinobacter, Firmicutes, and Proteobacteria, however none of the taxa made up more than 1% of the communities. The Comamonadaceae sequence enriched in magnetite samples from CSW1.1 was also enriched in this sample, but much less abundant (0.3% versus 14% of the magnetite-enriched communities; Table 4.5). The microbial communities from fluids and *in situ* devices from the QV1.1 showed

Table 4.3 - Relative-abundance of fluid-enriched bacterial taxa in different sample types

Well	Phylum	Class	Order	Family	2013	2014	Glass	Magnetite	Calcite
CSW1.1	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	42 ± 0.44	8.4 ± 1.6	0.40 ± 0.2	0.12 ± 0.09	0.33 ± 0.28
	Firmicutes	Clostridia	Thermoanaerobacterales	SRB2	0.98 ± 0.07	0.04 ± 0.02	0.03 ± 0.01	0	0.02 ± 0.01
CSW1.3	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	1.4 ± 1.2	0.94 ± 1.3	0.02 ± 0.006	0	0.03 ± 0.01
	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	0.71 ± 0.51	0.57 ± 0.71	0.01 ± 0.002	0.01 ± 0.01	0.02 ± 0.004

*No fluid-enriched taxa were found in QV1.1 or QV1.2.

**Values represent average and standard deviation of relative abundance for that treatment in the specified well.

Table 4.4 – Relative-abundance of magnetite-enriched bacterial taxa in different sample types

Well	Phylum	Class	Order	Family	2013	2014	Glass	Magnetite	Calcite
CSW1.1	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	0.01 ± 0.01	0	0.04 ± 0.01	14 ± 6.5	0.02 ± 0
	Actinobacteria	Actinobacteria	Frankiales	Sporichthyaceae	0	0	0	1.8 ± 1.1	0.001 ± 0
	Proteobacteria	Alphaproteobacteria	SAR11	LD12 freshwater group	0	0	0	1.2 ± 1.1	0
QV1.2	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	0	0	0.09 ± 0.03	6.2 ± 0.54	0.07 ± 0.04
	Actinobacteria	Actinobacteria	Propionibacteriales	Propionibacteriaceae	0	0	0.01 ± 0.01	2.5 ± 2.8	0.01 ± 0.01
	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	0.02 ± 0.01	0.02 ± 0.01	0.02 ± 0.01	1.2 ± 0.23	0.02 ± 0.01
	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	0	0	0.01 ± 0	1.1 ± 0.31	0.01 ± 0.01
	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	0	0	0.01 ± 0	1.0 ± 0.08	0.01 ± 0.01

*No magnetite-enriched taxa were found in CSW1.3 or QV1.1.

**Values represent average and standard deviation of relative abundance for that treatment in the specified well.

Table 4.5 – Relative-abundance of calcite-enriched bacterial taxa in different sample types

Well	Phylum	Class	Order	Family	2013	2014	Glass	Magnetite	Calcite
CSW1.1	Deinococcus-Thermus	Deinococci	Deinococcales	Trueperaceae	0.001 ± 0.002	33 ± 5.9	5.6 ± 1.5	0.26 ± 0.08	25 ± 7.4
	Firmicutes	Clostridia	Clostridiales	Syntrophomonadaceae	0.001 ± 0.002	32 ± 5.8	5.3 ± 1.4	0.25 ± 0.08	24 ± 7.1

*No calcite-enriched taxa were found in CSW1.3, QV1.1, or QV1.2.

**Values represent average and standard deviation of relative abundance for that treatment in the specified well.

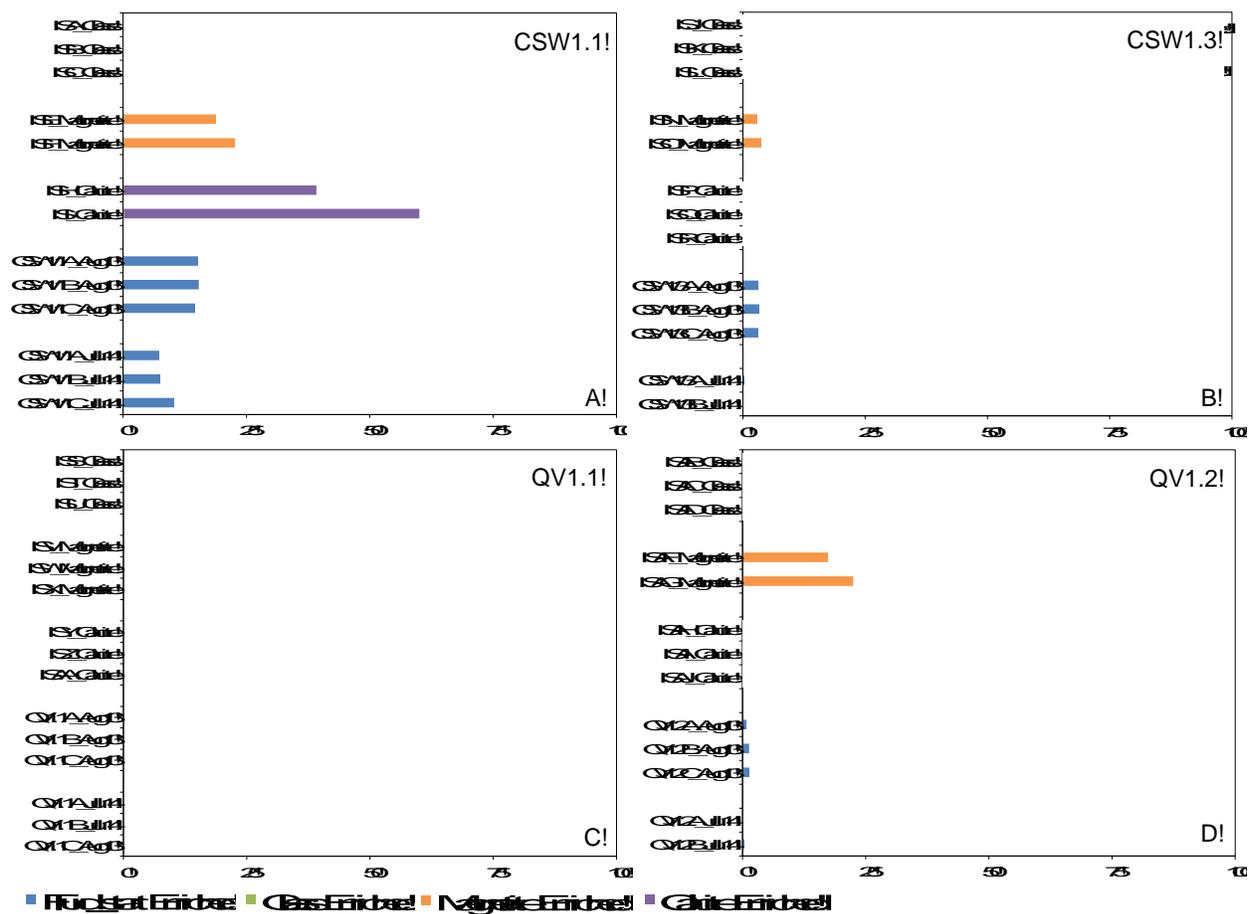


Figure 4.6 – Relative abundances of fluid-, glass-, magnetite-, and calcite-enriched bacterial taxa per each sample in all four wells. Calcite-enriched taxa were only detected in CSW1.1, while magnetite-enriched taxa were detected in CSW1.1, CSW1.3, and QV1.2. No mineral- or fluid-enriched taxa were determined in QV1.1.

similar compositions (Fig. 4.5) and therefore no taxa were determined to be statistically enriched in one type of sample compared to the others (Fig. 4.6C). The fluid-enriched taxa in QV1.2 constituted 1% of the fluid samples (Fig. 4.6D) and contained members of the Bacteroidetes, Actinobacteria, and Betaproteobacteria. Magnetite-enriched taxa from this well made up 20% of the communities from magnetite-containing *in situ* devices (Fig. 4.6D) and consisted of various families within the Betaproteobacteria, Actinobacteria, and methylotrophic Alphaproteobacteria (Table 4.4).

Conclusions

The purpose of this study was to explore microbe-mineral interactions within the serpentinite subsurface environment on environmentally relevant mineral substrates. Both calcite and magnetite are prevalent throughout CROMO cores (6). Additionally, we aimed to assess the microbiological capabilities to access inorganic carbon (from calcite) and terminal electron acceptors (in the form of ferric iron from magnetite) from mineral sources, given that such vital solutes are depleted in serpentinite fluids. The mineral-substrate *in situ* colonization devices yielded deployed into four different wells at CROMO yielded different microbial communities, depending on the inoculum source (i.e. well fluids; Figs. 4.5 & 4.6), with the most drastic results coming from CSW1.1, the most extreme serpentinite well. The addition of calcite to CSW1.1 resulted in the significant enrichment of *Deinococcus-Thermus* sequences, previously detected at serpentinite sites in low abundances (2, 22, 31, 33) and the addition of magnetite in the same well resulted in enrichment of Comamonadaceae, common inhabitants of the serpentinite the serpentinite subsurface environment. The results from more moderate CROMO wells were less clear, indicating that enrichment of select communities on minerals in this environment may be dependent on the start inoculum (i.e. well fluids). CSW1.1 fluids are depleted in DIC and exhibit extreme pH (Table 4.1) and therefore the mineral substrates may have offered rare taxa a refuge from the harsh environment. The lack of evident relationships between the mineral-enriched taxa and the mineral substrates (e.g. known ability to fix CO₂ from carbonate or iron-reduction of magnetite) raises the question of why specific taxa were enriched by the addition of certain mineral substrates. These data, along with the findings from CROMO cores (34), suggest that the mineralogy of the subsurface environment can impact community composition, potentially offering rare taxa refuge from challenging environmental.

Materials and Methods

In Situ Colonization Devices

To assess the growth of microorganisms on solid surfaces within the serpentinite subsurface environment, *in situ* colonization devices containing mineral substrates were developed and deployed into select wells. The *in situ* colonization devices were made of ethanol-washed, 1 inch Schedule 80 PVC tubing cut into 1.5 inch pieces and capped with 143 μm Teflon mesh (Small Parts Inc, Logansport, IN), allowing for the transport of fluids and small particles within the device (Fig. 4.1B). The devices were filled with ~ 7.5 g of 500 μm diameter borosilicate glass beads (Sigma-Aldrich, St. Louis, MO), magnetite ($\text{Fe}^{3+}_2\text{Fe}^{2+}\text{O}_4$) (Ward's Natural Science, Rochester, NY), or calcite (CaCO_3) (Ward's Natural Science, Rochester, NY), which were crushed and sieved to insure uniform mineral size. Post-assembly, the devices were wrapped in aluminum foil and autoclaved. The devices were deployed in triplicate into the CROMO wells CSW1.1 (pH 12.4), CSW1.3 (pH 10.3), QV1.1 (pH 11.6), and QV1.2 (pH 9.1) in August 2013. The devices were tied together with Teflon fishing line, with 10 cm between devices, and deployed to a depth of 15 mbs. The devices were left undisturbed in the wells for one year, until recovery in July 2014. Upon retrieval, the *in situ* devices were placed into individual sterile Whirlpak bags (Nasco, Fort Atkinson, WI), flash frozen in liquid nitrogen, and stored at -80°C until DNA extraction. Well fluids were collected during both deployment and retrieval, as previously described (8).

DNA Extraction

DNA was extracted from mineral substrates using the MoBio PowerMaxSoil Kit (Carlsbad, CA), per the manufacturers instructions. The resulting DNA was concentrated in an Amicon Ultra-2 Centrifugal Filter Unit with Ultracel-30 membrane (Millipore, Darmstadt,

Germany) to a volume of 50 μ L. Filtering of the well fluids and DNA extractions of the filters were conducted as described previously (4). DNA was quantified using the High Sensitivity reagents for a Qubit® 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA), which has a detection limit of 0.1ng/ μ L DNA.

Quantitative-PCR

Copies of the 16S rRNA gene were determined by quantitative polymerase chain reaction (q-PCR) using domain-specific primers targeting the V6V4 hypervariable region of the 16S rRNA gene of bacteria (29). Samples were run on a BioRad q-PCR instrument using the SsoAdvanced SybrGreen Assay (Biorad, Hercules, CA).

The statistical package StatPlus (AnalystSoft, Inc.) was used in Excel 2011 to perform one-way ANOVA tests, to determine if the different mineral substrates in the *in situ* devices, and two-way ANOVA test, to determine if the interaction between well and mineral substrates in microcosms, resulted in a significant difference in 16S rRNA gene abundance (from q-PCR data).

16S rRNA Gene Sequencing and Data Analysis

Purified DNA samples from *in situ* device minerals, CROMO well fluids, microcosm minerals, and microcosm fluids were submitted to the Josephine Bay Paul Center at the Marine Biological Laboratory for sequencing of the V4V5 region of the 16S rRNA gene on an Illumina MiSeq instrument as part of the Census of Deep Life project (18). The paired-end reads were merged and subjected to MBL's post-processing quality control for removal of low quality reads and chimera checking (10). The *in situ* device and microcosm samples yielded between 15,753 and 277,722 merged, quality-filtered sequences each. Any sample that yielded less than 5,000 sequences was considered a failed sequencing run and not included the analysis.

Community diversity bar charts were generated using the Visualization and Analysis of Microbial Populations Structures (VAMPS) interface, which assigns taxonomy of sequences against SILVA and RDP (10). The charts presented here were formed using Family-level taxonomic assignment and represent relative abundances of sequences.

Using Mothur (24), the sequence data was clustered into unique sequences and assigned taxonomy compared by alignment to the SILVA database v119 (23). Rarefaction analyses and creation of a community similarity dendrogram, using the sorclass dissimilarity index, were performed in Mothur (24). Significant differences in the abundances of unique sequences between groups of samples were tested with the aid of the Phyloseq (15) and DESeq (13) packages in R. More details about this method can be found in Chapter 3 (34) and Appendix A.

REFERENCES

REFERENCES

1. Barnes I, O'Neil JR (1971) Calcium-magnesium carbonate solid solutions from Holocene conglomerate cements and travertines in the Coast Range of California. *Geochim Cosmochim Acta* 35:699-718.
2. Brazelton WJ, et al. (2010) Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proc Natl Aca Sci* 107:1612-1617.
3. Brazelton WJ, Nelson B, Schrenk MO (2012) Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2: doi: 10.3389/fmicb.2011.00268.
4. Brazelton WJ, Morrill PL, Szponar N, Schrenk MO (2013) Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl Environ Microb* 79:3906-3916.
5. Cardace D, et al. (2013) Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Scientific Drilling* 16:45-55.
6. Carnevale, DC (2013) Carbon sequestration potential of the Coast Range Ophiolite in California. (Master's thesis). Retrieved from: Open Access Master's Theses. Paper 46.
7. Costerton JW, et al. (1994) Biofilms, the customized microniche. *J Bacteriol* 176:2137-2142.
8. Crespo-Medina M, et al. (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604.
9. Edwards KJ, McCollom TM, Konishi H, Buseck PR. (2003) Seafloor bioalteration of sulfide minerals: results from in situ incubation studies. *Geochim Cosmochim Acta* 67:2843-2856.
10. Huse SM, et al. (2014) VAMPS: a website for visualization and analysis of microbial population structures. *BMC Bioinformatics* 15:41.
11. Kelley DS, et al. (2005) A serpentinite-hosted ecosystem: The Lost City Hydrothermal Field. *Science* 307:1428-1434.
12. Kostka JE, Nealson KH (1995) Dissolution and reduction of magnetite by bacteria. *Environ Sci Technol* 29:2535-2540.
13. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol* 15:550.

14. Makarova KS, et al. (2001) Genome of the extremely radiation-resistant bacterium *Deinococcus radiodurans* viewed from the perspective of comparative genomics *Microbiol Mol Biol Rev* 65:44-79.
15. McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS One* 8:e61217.
16. McMurdie PJ & Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531. doi:10.1371/journal.pcbi.100353126.
17. Mitchell AC, Lafrenière MJ, Skidmore ML, Boyd ES (2013) Influence of bedrock mineral composition on microbial diversity in a subglacial environment. *Geology* DOI: 10.1130/G34194.1
18. Morrison HG, Grim SL, Vineis JH, Sogin ML (2013) 16S amplicon fusion primers and protocol for Illumina platform sequencing. Figshare. Available online at: <http://dx.doi.org/10.6084/m9.figshare.833944>
19. Normad P (2006) The families Frankiaceae, Geodermatophilaceae, Acidothermaceae, and Sporichthyaceae. *In: The Prokaryotes*. 3:669-681.
20. Orcutt BN, et al. (2010) Colonization of subsurface microbial observatories deployed in young ocean crust. *ISME J* 5:692-703.
21. Orcutt BN, Wheat CG, Edwards KJ (2010) Subseafloor ocean crust microbial observatories: development of FLOCS (Flow-through Osmo Colonization System) and evaluation of borehole construction materials. *Geomicrobiol J* 27:143-157.
22. Postec A, et al. (2015) Microbial diversity in a submarine carbonate edifice from the serpentinizing hydrothermal system of the Prony Bay (New Caledonia) over a 6-year period. *Front Microbiol* 6:857.
23. Pruesse E, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188-7196.
24. Schloss PD, Westcott SL (2011) Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl Environ Microb* 77:3219-3226.
25. Shock EL (2009) Minerals as energy sources for microorganisms. *Economic Geology* 104:1235-1248.

26. Slepova TV, et al. (2006) *Carboxydocella sporoproducens* sp. nov., a novel anaerobic CO-utilizing/H₂-producing thermophilic bacterium for a Kamchatka hot spring. *Int J Syst Evol Micro* 56:797-800.
27. Sokolova TG, et al. (2002) *Carboxydocella thermautotrophica* gen. nov., sp. nov., a novel anaerobic CO-utilizing thermophile from Kamchatkan hot spring. *Int J Syst Evol Micro* 52:1961-1967.
28. Sokolova TG, et al. (2009) Diversity and ecophysiological features of thermophilic carboxydrotrophic anaerobes. *FEMS Microbiol Rev* 68:131-141.
29. Sogin ML, et al. (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. USA*. 103:12115-12120.
30. Suzuki S, et al. (2013) Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proc Natl Acad Sci USA* 110:15336-15341.
31. Suzuki S, et al. (2014) Physiological and genomic features of highly alkaliphilic hydrogen-utilizing *Betaproteobacteria* from a continental serpentinizing site. *Nat Commun* 5:3900.
32. Tiago I, Veríssimo A (2013) Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinization. *Environ Microbiol* 15:1687-1706.
33. Twing KI, et al. (In Review) Serpentinizing fluids from a subsurface observatory harbor extremely low diversity microbial communities adapted to high pH. *Proc Natl Acad Sci USA*.
34. Twing KI, et al. (In Preparation) Microbial diversity of serpentinite cores from the Coast Range Ophiolite Microbial Observatory. *Appl Environ Microbiol*.
35. Woycheese KM, et al. (2015) Out of the dark: transitional subsurface-to-surface microbial diversity in a terrestrial serpentinizing seep (Manleluag, Pangasianan, the Philippines). *Front Microbiol* 6:44.
36. Zaremba-Niedzwiedzka K, et al. (2013) Single-cell genomics reveal low recombination frequencies in freshwater bacteria of the SAR11 clade. *Genome Biology* 14:R130.

CHAPTER 5

Summary and Future Perspectives

Prior to this research, our understanding of life in the continental subsurface of active serpentinite environments was limited to a few studies based on opportunistic sampling of natural springs and seeps (1, 2, 10). While these studies yielded surprisingly similar results that helped develop a model of the system (9), it can be difficult to constrain what is driving microbiological trends from relatively few samples in time and space. The establishment of CROMO, a long-term microbial observatory within an actively serpentinizing subsurface environment, presently consists of two main wells with end-member characteristic fluids. In addition of the two main wells, it has a set of six monitoring wells, drilled to varying depths across geochemical gradients, which allowed for the identification of geochemical drivers (pH, carbon monoxide, and methane concentrations) of the abundant taxa (Chapter 2; 11). The most extreme pH wells exclusively contained a single OTU of Betaproteobacteria and a few OTUs of Clostridia, in addition to genes involved in carbon fixation and the cycling of methane and carbon monoxide. Wells with more moderate pH (9-11) contained a higher diversity of bacteria and genes for the aerobic oxidation of methane. These data suggest that there are different ecological niches within the serpentinite subsurface environment, with hydrogen metabolism and carbon monoxide oxidation taking place in the most end-member fluids and methane oxidation occurring in mixing zones between serpentinite end-member fluids and surface waters (Chapter 2; 11).

CROMO represents the first drill campaign into the continental serpentinite environment and cores of CSW1.1 and QV1.1 permitted microbiological exploration of subsurface serpentinite rocks for the first time. The taxa that appear to be specifically enriched in cores

relative to fluids suggest that there may be cryptic chemical cycling (e.g. nitrogen and methane) taking place on subsurface rocks, which may not be evident from fluid samples alone (Chapter 3, 12). Differences in lithostratigraphic zonation within the homogenous cores led to differences in community composition, with archaea (previously undetected at CROMO; Chapter 2, 11) dominating magnetite-bearing serpentine minerals and diverse bacteria dominating clay-containing layers (Chapter 3, 12). Given that unique microorganisms are associated with core materials in comparison to fluids at CROMO, the aim of Chapter 4 was to further explore microbe-mineral interactions through *in situ* experimentation down well. The results from *in situ* colonization on mineral substrates within the pH 12.5 well suggest that magnetite and calcite may select for different microbial taxa. These data suggest that core-associated communities are distinct within different lithologies and could be contributing to cryptic biogeochemical cycling within the serpentinite subsurface environment and that specific minerals within those lithologies may be driving community composition (Chapter 3; 12; Chapter 4).

Bioinformatic methods, such as DESeq (7), have long been used in to determine the differential expression of RNA data, however, they have not previously been employed on environmental DNA data (8). By statistically assessing differential relative abundances of 16S rRNA sequences, taxa that were particularly enriched in a given sample type, such as fluids, cores, or mineral-substrates, could be assessed without having to remove all overlapping sequences from the datasets in an effort to control for contamination. This method allowed for the identification of mineral-enriched taxa, which could be further explored to help decipher what biogeochemical cycling may be taking place in distinct geological features within the subsurface environment.

Despite all of the great technological advances in molecular biology and sequencing technologies over the last decade, it is still essential to culture and characterize microorganisms to fully understand their metabolic capabilities. There are ongoing culturing efforts to isolate relevant organisms from the serpentinite subsurface environment at CROMO and elsewhere. The data presented here can help inform those efforts by defining what the environmentally-relevant taxa are, as was seen in the study by Crespo-Medina, *et al.* (4), and identifying geochemical parameters that control their abundance, which may be the key to getting them to grow in the laboratory.

The Coast Range Microbial Observatory represents the first drill campaign into the continental subsurface serpentinite environment and allows for direct access and study of the actively serpentinizing subsurface environment. With few exceptions (3, 4, 13), the research presented here represents some of the first data to come out of the CROMO project and makes a major contribution to understanding life within serpentinite fluids and minerals. As CROMO is a large, interdisciplinary project, as more geochemical, geological, and hydrological data from the site become available, it will yield a better understanding of the connectivity of the wells and the system as a whole, which will likely clarify some of the questions still remaining about the distribution and constraints on life within the serpentinite subsurface environment.

In the coming months, two large-scale drilling projects will take place to further explore the extent of life within the serpentinite subsurface environment. The International Ocean Discovery Program (IODP) Expedition 357 will drill into the Atlantis Massif, the location of the LCHF, to investigate the extent, diversity, and activity of microbes living within the serpentinizing ocean crust (5). In Oman, the Samail Ophiolite will be drilled to investigate life in the continental subsurface environment there (6). As CROMO was the first drill campaign in the

continental serpentinite environment (3), the work completed there, both in terms of the methodology and findings can serve to inform both of these upcoming drilling projects. A great deal of troubleshooting in this project was put into determining the best nucleic acid extraction method for getting DNA out of serpentinite cores, which will help inform molecular methodologies used on upcoming drilling projects. Furthermore, the data described in Chapter 3 regarding the microbial diversity of serpentinite cores can serve as a comparison to rock-hosted organisms found in other continental and marine sites of serpentinization, such as Oman and the Atlantis Massif.

REFERENCES

REFERENCES

1. Brazelton WJ, Nelson B, Schrenk MO (2012) Metagenomic evidence for H₂ oxidation and H₂ production by serpentinite-hosted subsurface microbial communities. *Front Microbiol* 2:268.
2. Brazelton WJ, Morrill PL, Szponar N, Schrenk MO (2013) Bacterial communities associated with subsurface geochemical processes in continental serpentinite springs. *Appl Environ Microb* 79:3906-3916.
3. Cardace D, et al. (2013) Establishment of the Coast Range ophiolite microbial observatory (CROMO): drilling objectives and preliminary outcomes. *Scientific Drilling* 16:45-55.
4. Crespo-Medina M, et al. (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604.
5. Früh-Green G, et al. Serpentinization and life: biogeochemical and tectono-magmatic processes in young mafic and ultramafic seafloor. *IODP Proposal*. 758-Full2.
6. Kelemen P, et al. (2013) Scientific drilling and related research in the Samail Ophiolite, Sultanate of Oman. *Scientific Drilling* 15:64-71.
7. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol* 15:550.
8. McMurdie PJ & Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531. doi:10.1371/journal.pcbi.100353126.
9. Schrenk MO, Brazelton WJ, Lang SQ (2013) Serpentinization, carbon and deep life. *Rev Mineral Geochem* 75:575-606.
10. Suzuki S, et al. (2013) Microbial diversity in The Cedars, an ultrabasic, ultrareducing, and low salinity serpentinizing ecosystem. *Proc Natl Acad Sci USA* 110:15336-15341.
11. Twing KI, et al. (In Review) Serpentinizing fluids from a subsurface observatory harbor extremely low diversity microbial communities adapted to high pH. *Proc Natl Acad Sci USA*.
12. Twing KI, et al. (In Preparation) Microbial diversity of serpentinite cores from the Coast Range Ophiolite Microbial Observatory. *Appl Environ Microbiol*.
13. Wang DT, et al. (2015) Nonequilibrium clumped isotope signals in microbial methane. *Science* 348:428-431.

APPENDICES

APPENDIX A

Example DESeq commands for identifying group specific sequences³

The input files are (all input files must be saved with unix linebreaks):

1. AllCoreArch.final.count_table file from Mothur (4)
2. AllCoreArch.final.taxonomy table from Mothur (4)
3. Arch_SampleInfo.csv - Sample info file with exact same sample names as count table
assigning samples to groups

```
## Edit the taxonomy table from mothur with this script to make it more readable downstream
```

```
taxonomy_edit.py [filename].taxonomy
```

```
## Output: AllCoreArch.final.taxonomy.renamed.txt
```

```
## Create phyloseq (2) otus-tax-sample object in R
```

R

```
library(phyloseq)
```

```
otus = read.table("AllCoreArch.final.count_table", header=TRUE, row.names=1)
```

```
otus = otu_table(otus, taxa_are_rows=TRUE)
```

```
otus = subset(otus, select=-c(1:1))
```

```
tax = read.table("AllCoreArch.final.taxonomy.renamed.txt", sep='\t', header=FALSE,  
row.names=1)
```

```
tax = tax_table(as.matrix(tax))
```

³ Workflow developed in the laboratory of Dr. William Brazelton at the University of Utah, Department of Biology.

```

sam = read.csv("Arch_SampleInfo.csv")

sam = sample_data(sam)

row.names(sam) = sample_names(otus)

merged = merge_phyloseq(otus,tax,sam)

merged

## Create a bar chart of whole-sample taxonomic classifications.
## "V6" corresponds to family level.

merged_props = transform_sample_counts(merged, function(x) 100 * x/sum(x))

merged_props_glomV6 = tax_glom(merged_props, "V6")

merged_props_glomV6

## Output:
## phyloseq-class experiment-level object
## otu_table() OTU Table: [ 155 taxa and 8 samples ]
## sample_data() Sample Data: [ 8 samples by 8 sample variables ]
## tax_table() Taxonomy Table: [ 155 taxa by 8 taxonomic ranks ]
## Use number of taxa (155) in next command

library(ggplot2)

cbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2",
"#D55E00", "#CC79A7")

png(filename = "predeseq-barplot-glomV6.png", width = 480, height = 480, units = "px",
pointsize = 18)

```

```
plot_bar(merged_props_gloomV6, fill="V6") + coord_flip() + ylab("Percent Sequences") +  
scale_fill_manual(values=colorRampPalette(cbPalette)(155)) +  
theme(legend.position='none')
```

```
## Identify sample-specific taxa with the DESeq package (1):
```

```
library("DESeq2")
```

```
deseqd = phyloseq_to_deseq2(merged, ~CSW_vs_woOL)
```

```
deseqd = DESeq(deseqd)
```

```
res = results(deseqd, contrast=c("CSW_vs_woOL", "CSW", "QV"))
```

```
res = res[order(res$padj), ]
```

```
write.csv(res, file='deseq-results.csv')
```

```
## Plot of positively and negatively enriched sequences
```

```
## Make deseql plot with red points as significant
```

```
png(filename = "plotMA.png", width = 480, height = 480, units = "px", pointsize = 18)
```

```
plotMA(res, ylim = c(-3,3))
```

```
dev.off()
```

```
## red points are "significant at 10% false discovery rate (padj)" - DESeq-vignette.pdf
```

```
## Select the group-specific taxa, find the taxonomy information associated with them, and
```

```
## record this information as R objects and write them to files so that we can open them as tables
```

```
## First: select significantly up- or down-"regulated" taxa, including very large differences where
```

```
## padj=NA:
```

```

bigpos = res[(res$log2FoldChange > 0.9), ]
bigneg = res[(res$log2FoldChange < -0.9), ]
res_no_na = res[order(res$padj, na.last=NA),]
sigres = res_no_na[(res_no_na$padj < 0.1),]
sigrespos = sigres[(sigres$log2FoldChange > 0),]
sigresneg = sigres[(sigres$log2FoldChange < 0),]
bigpos_names = rownames(res) %in% rownames(bigpos)
bigneg_names = rownames(res) %in% rownames(bigneg)
sig_names = rownames(res) %in% rownames(sigres)
all_names = bigpos_names | bigneg_names | sig_names
other = res[!all_names,]

## Add taxonomy data to the selected results

bigpos_tax = cbind(as(bigpos, "data.frame"), as(tax_table(merged)[rownames(bigpos), ],
"matrix"))
bigneg_tax = cbind(as(bigneg, "data.frame"), as(tax_table(merged)[rownames(bigneg), ],
"matrix"))
sigrespos_tax = cbind(as(sigrespos, "data.frame"),
as(tax_table(merged)[rownames(sigrespos), ], "matrix"))
sigresneg_tax = cbind(as(sigresneg, "data.frame"),
as(tax_table(merged)[rownames(sigresneg), ], "matrix"))
other_tax = cbind(as(other, "data.frame"), as(tax_table(merged)[rownames(other), ],
"matrix"))

```

```

## Add count table information to the selected results, for future reference in the output ## files

bigpos_tax_counts = cbind(as(bigpos_tax, "data.frame"),
as(otu_table(merged)[rownames(bigpos_tax), ], "matrix"))

bigneg_tax_counts = cbind(as(bigneg_tax, "data.frame"),
as(otu_table(merged)[rownames(bigneg_tax), ], "matrix"))

sigrespos_tax_counts = cbind(as(sigrespos_tax, "data.frame"),
as(otu_table(merged)[rownames(sigrespos_tax), ], "matrix"))

sigresneg_tax_counts = cbind(as(sigresneg_tax, "data.frame"),
as(otu_table(merged)[rownames(sigresneg_tax), ], "matrix"))

other_tax_counts = cbind(as(other_tax, "data.frame"),
as(otu_table(merged)[rownames(other_tax), ], "matrix"))

## Tables of DESeq results broken down by positive, negative, and other

## Merge all sig results including NAs with huge log2changes, skipping duplicate rows

## Write the results for future reference

duprows_pos = rownames(bigpos_tax_counts) %in% rownames(sigrespos_tax_counts)

duprows_neg = rownames(bigneg_tax_counts) %in% rownames(sigresneg_tax_counts)

allpos =

rbind(data.frame(bigpos_tax_counts[!duprows_pos,]),data.frame(sigrespos_tax_counts))

allneg =

rbind(data.frame(bigneg_tax_counts[!duprows_neg,]),data.frame(sigresneg_tax_counts))

write.csv(allpos,file='deseq-allpos-results.csv')

```

```

write.csv(allneg,file='deseq-allneg-results.csv')

write.csv(other_tax_counts,file='deseq-other-results.csv')

## Make phyloseq objects with OTU counts and taxonomy for manipulation in R later

otus_pos = otus[row.names(allpos), ]

otus_neg = otus[row.names(allneg), ]

otus_other = otus[row.names(other), ]

physeq_pos = merge_phyloseq(otus_pos,tax)

physeq_neg = merge_phyloseq(otus_neg,tax)

physeq_other = merge_phyloseq(otus_other,tax)

## Rename sample names and merge into a single phyloseq object

sample_names(physeq_pos) = paste("pos-", sample_names(physeq_pos), sep="")

sample_names(physeq_neg) = paste("neg-", sample_names(physeq_neg), sep="")

sample_names(physeq_other) = paste("other-", sample_names(physeq_other), sep="")

physeq_all = merge_phyloseq(physeq_neg, physeq_other)

## The next set of commands will merge the counts for the positive, negative, and other
## categories of samples. Edit the first command so that it lists the correct number of
## pos, neg, and other. Check sample_names(physeq_all) to be sure:

sample_names(physeq_all)

```

```

types =
c('pos','pos','pos','pos','pos','pos','pos','pos','neg','neg','neg','neg','neg','neg','neg','neg','other','other','other','other','other','other','other')

sampledata = sample_data(data.frame(Type = types, size = nsamples(physeq_all), replace =
TRUE, row.names = sample_names(physeq_all), stringsAsFactors = FALSE))

sampledata

physeq_all_merged = merge_phyloseq(physeq_all,sampledata)

physeq_all_merged = merge_samples(physeq_all_merged, 'Type')

physeq_all_merged_props = transform_sample_counts(physeq_all_merged, function(x) 100
* x/sum(x))

## Create Bar charts of taxonomy specific to positive, negative, and other sequences

## Family-level plot:

physeq_all_merged_props_glom6 = tax_glom(physeq_all_merged_props, "V6")

library(ggplot2)

cbPalette <- c("#999999", "#E69F00", "#56B4E9", "#009E73", "#F0E442", "#0072B2",
"#D55E00", "#CC79A7")

png(filename = "postdeseq-barplot-glom6.png", width = 480, height = 480, units = "px",
pointsize = 18)

## Use the same number of taxa as before (155)

plot_bar(physeq_all_merged_props_glom6, fill="V6") + coord_flip() + ylab("Percent
Sequences") + scale_fill_manual(values=colorRampPalette(cbPalette)(155)) +
theme(legend.position='none')

```

dev.off()

Output 6: Tables of results corresponding to bar charts above

Write csv file containing otu counts and taxonomy to aid in interpretation of the bar plot. The

t() transposes the otu_table so that species are rows, as in the tax table.

final_tax_otu_props_glom6 = cbind(as(t(otu_table(physeq_all_merged_props_glom6)),

"matrix"), as(tax_table(physeq_all_merged_props_glom6), "matrix"))

write.csv(final_tax_otu_props_glom6,file='final_tax_otu_props_glom6.csv')

REFERENCES

REFERENCES

1. Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *bioRxiv*, 1–21. doi:10.1101/002832
2. McMurdie PJ & Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*, 8(4), e61217. doi:10.1371/journal.pone.0061217
3. McMurdie PJ & Holmes S (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Computational Biology*, 10(4), e1003531. doi:10.1371/journal.pcbi.1003531
4. Schloss PD, et al. (2009) Introducing mothur: Open Source, Platform-independent, Community-supported Software for Describing and Comparing Microbial Communities. *Appl Environ Microbiol* AEM.01541–09. doi:10.1128/AEM.01541-09

APPENDIX B

Bioinformatic analyses for Crespo-Medina *et al.*, 2014⁴

Summary of Crespo-Medina et al., 2014 and my bioinformatic contributions

In Crespo-Medina *et al.* (2), fluids from CROMO wells were incubated in microcosms with various carbon sources and the addition of nutrient and/or electron acceptors. Methane and acetate additions led to the enrichment of a Betaproteobacterium of the family Comamonadaceae. While the addition of electron acceptors was not necessary for growth, it did and in serpentinite subsurface systems as a whole, the 16S rRNA sequences from the

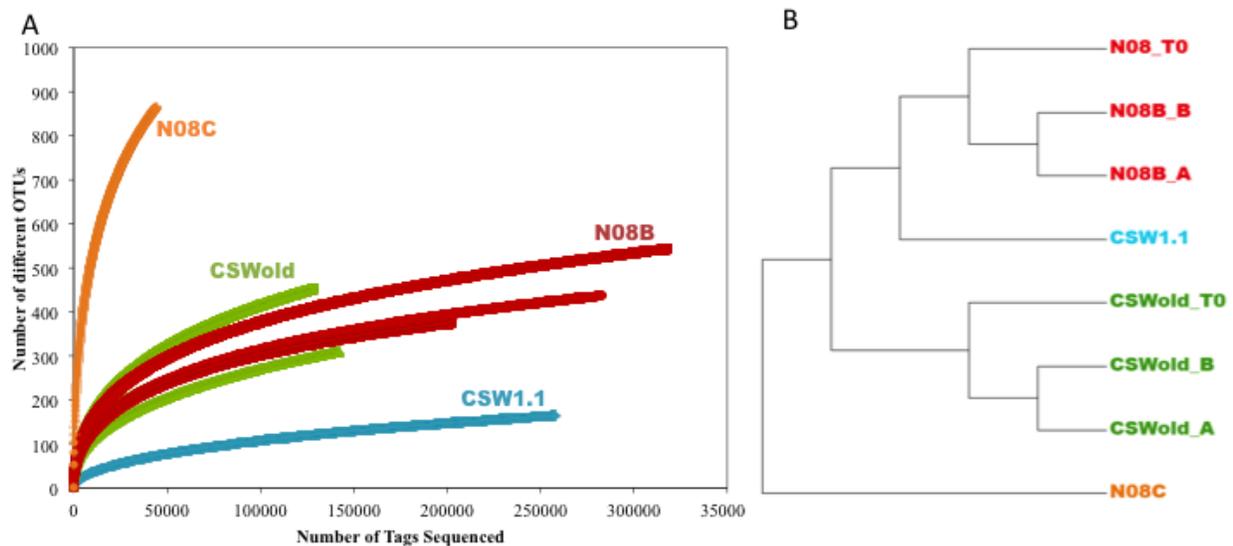


Figure B.1 – Diversity of 16S rRNA gene OTUs (97% sequence similarity) from fluids collected at CROMO in March 2013 was analyzed using mothur (7). A) Alpha-diversity, as displayed by rarefaction, indicates the greatest within sample diversity in the control sample of N08C and the least within sample diversity in the extreme sample of CSW1.1. Within sample diversity of field replicates and T0 from both N08B and CSWold show similar patterns. B) Beta-diversity, displayed in a community-dissimilarity dendrogram calculated from the Morisita-Horn index, indicates that samples from a single well are more similar to each other than samples from different wells.

⁴ The work described here was published in *Frontiers in Microbiology* in 2014: Crespo-Medina M, Twing KI, Kubo MDY, Hoehler TM, Cardace D, McCollom T, Schrenk MO (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604

microcosms needed to be compared to environmental sequences. I performed the bioinformatic analyses to assess the bacterial diversity of environmental samples (Fig. B.1, B.2, & B.3) and comparisons of 16S rRNA genes to sequences from other serpentinite sites to determine closest relatives (Tables B.1 & B.2).

Bioinformatic Results

The microbial diversity at the class level varies by sampling well, with the most extreme well, CSW1.1, exhibiting the lowest diversity and the moderate well N08C exhibiting high diversity (Fig. B.1A). Field replicate samples from the wells N08B and CSWold show nearly identical community composition to one another, while the T0 samples from both those wells, filtered merely days later immediately before the beginning of the microcosm experiments,

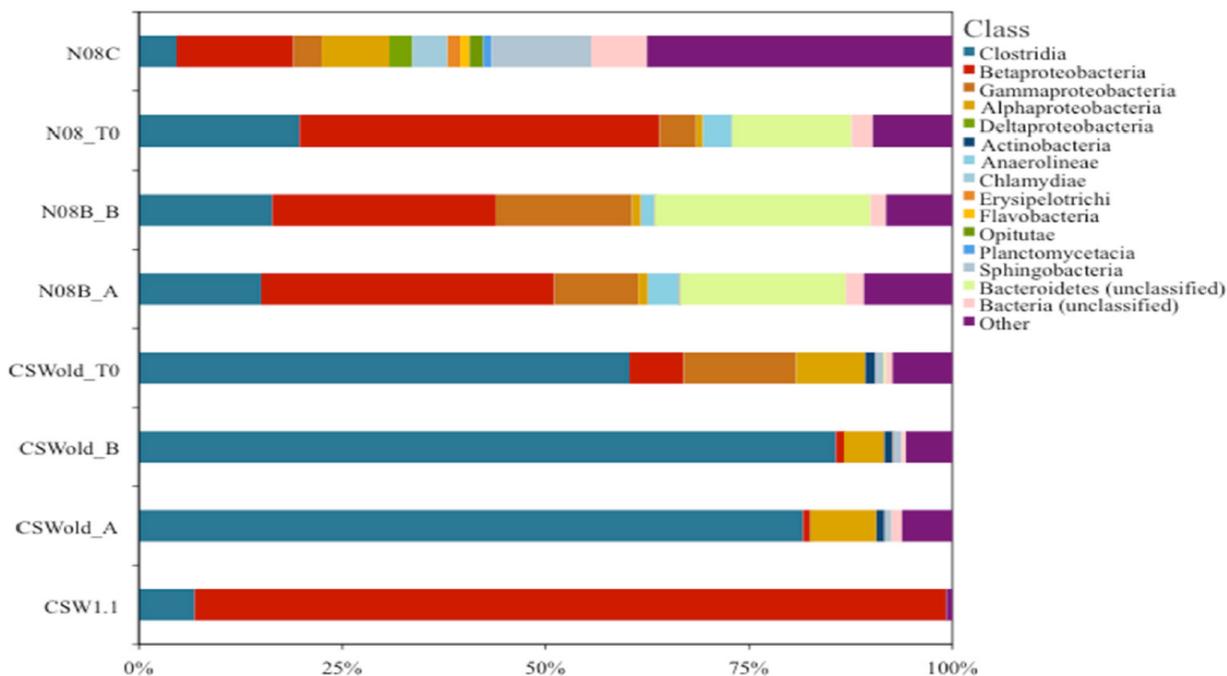


Figure B.2 – Proportion of bacterial classes present in fluids collected at CROMO in March 2013. OTUs were made at 97% similarity level and assigned taxonomy from the SILVA database using mothur (Pruesse *et al*, 2007; Schloss & Westcott, 2011). The CSWold and N08B T0 samples show a slight shift in community composition compared to their field replicate counterparts.

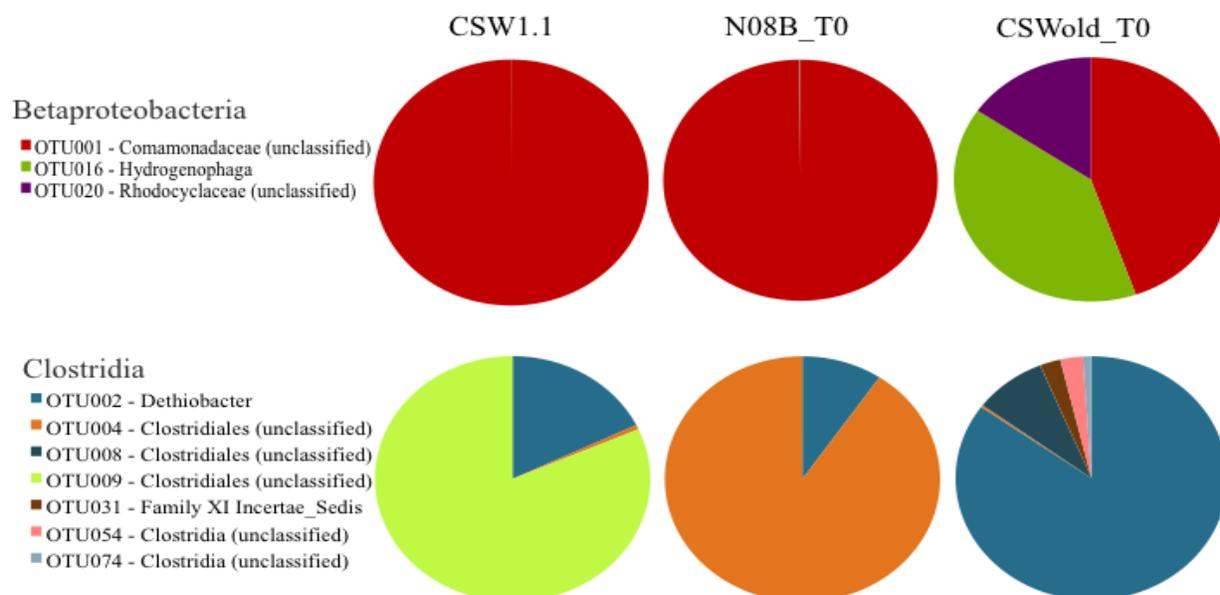


Figure B.3 – Diversity of OTUs (97% similarity) belonging to the classes *Betaproteobacteria* and *Clostridia* within *CSW1.1*, *N08B*, and *CSWold*. The *Betaproteobacteria* in *CSW1.1* and *N08B* are dominated by a single OTU (OTU001) belonging to the family *Comamonadaceae*. The same OTU makes up 45% of the *Betaproteobacteria* in *CSWold*.

display slight shifts in community composition (Fig. B.1B).

The total community of *CSW1.1* is 92% of a single OTU (OTU001) of *Betaproteobacteria* belonging to the family *Comamonadaceae* (Figs. B.2 and B.3). Forty-four percent of the total community of *N08B_T0* is made up of *Betaproteobacteria*, also all belonging to OTU001. Only 6.5% of the total community of *CSWold_T0* is made up of *Betaproteobacteria* (Fig. B.2) and among them 45% are OTU001, 40% are *Hydrogenophaga* (also of the family *Comamonadaceae*) and 15% are an OTU belonging to the family *Rhodocyclaceae* (Figs. B.2 and B.3).

Clostridia, which make up 60% of *CSWold_T0* total community, are dominated by a single OTU of *Dethiobacter* (OTU002) (Figs. B.2 and B.3). The same OTU makes up 18% and

9% of the clostridial OTUs in CSW1.1 and N08B_T0, respectively (Fig.B.3). Instead 81% of sequences related to Clostridia in CSW1.1 belong to OTU009, which can only be classified to Clostridia belong to a different OTU of unclassified Clostridiales (Fig. B.3).

Table B.1 – Percent identity between end point sequences from March microcosms, most abundant OTUs from tag sequence data, and reference sequences

	CSWold	N08B	<i>Dethiobacter alkaliphilus</i> (NR044205)	CVCloAm3Ph44 ^a (AM778015.1)
CSWold				
N08B	97.9			
<i>Dethiobacter alkaliphilus</i> (NR044205)	98.4	96.8		
CVCloAm3Ph44 (AM778015.1)	97.3	98.9	96.8	
OTU002	98.1	99.7	97.1	99.2

^a Clone from serpentinization-driven Cabeco de Vide Aquifer in Portugal (Tiago and Veríssimo, 2013).

The [microcosm] sequences from CSWold and N08B were 97.3 and 98.9 similar, respectively, to an environmental sequence from another continental serpentinizing site in Portugal (10) (Table B.1). Two representative sequences [from the August microcosms], one from CSW1.1 and one from N08B microcosm, with CO₂ and acetate respectively, shared 100% identity to Comamonadaceae bacterium B1, isolated from the Cedars (9; Table B.2).

Table B.2 – Percent identity between two representative end point sequences from the August microcosms, most abundant OTUs from tag sequence data, and reference sequences

	CSW1.1 CO ₂	N08B Acetate	Bacterium B1 ^a (AP014569)	Bacterium A1 ^b (AP014568)	CVCloAm1Ph9 ^c (AM777999)
CSW1.1_CO ₂					
N08B_Acetate	100				
Bacterium_B1 (AP014569)	100	100			
Bacterium_A1 (AP014568)	99.7	99.7	99.7		
CVCloAm1Ph9 (AM777999)	100	100	100	99.7	
OTU001	100	100	100	99.7	100

^a Comamonadaceae bacterium B1 and ^b Comamonadaceae bacterium A1 from continental serpentinizing site at The Cedars (Suzuki et al., 2014).

^c Clone from serpentinization-driven Cabeco de Vide Aquifer in Portugal (Tiago and Veríssimo, 2013).

16S rRNA Tag Sequencing and Data Analysis Methods

DNA samples were submitted to the Marine Biological Laboratory's (MBL) Josephine Bay Paul Center for 16S rRNA amplicon sequencing of the V4V5 region of the 16S rRNA gene on an Illumina MiSeq instrument (3, 4). Sequences were subjected to quality control and preprocessing in accordance with MBL's standard methods on their VAMPS server (5).

Sequence reads were aligned to the SILVA SSURef alignment (v102) and taxonomic classification was assigned using mothur (6; 8). Sequences were clustered into operational taxonomic units (OTUs) at the 3% distance threshold using the cluster.split command and the average-neighbor clustering algorithm in mothur (7). Alpha-diversity (within sample) was assessed by rarefaction analysis, while beta-diversity (between sample) of the microbial communities was assessed by calculation of the Morisita-Horn index and displayed in a dissimilarity dendrogram produced with the tree.shared command in mothur (7).

MatGAT (1) was used to determine the percent sequence identify of the most abundant OTUs compared with 16S rRNA sequences, trimmed to the V4V5 region of the gene, from isolates within this study and from previous studies of the serpentinite subsurface (9, 10).

REFERENCES

REFERENCES

1. Campanella JJ, Bitincka L, Smalley J (2003) MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. *BMC Bioinformatics* 4:29.
2. Crespo-Medina M, et al. (2014) Insights into environmental controls on microbial communities in a continental serpentinite aquifer using a microcosm-based approach. *Front Microbiol* 5:604.
3. Huse SM, et al. (2014). Comparison of brush and biopsy sampling methods of the ileal pouch for assessment of mucosa-associated microbiota of human subjects. *Microbiome*. 2, doi: 10.1186/2049-2618-2-5.
4. Huse SM, et al. (2014). VAMPS: a website for visualization and analysis of microbial population structures. *BMC Bioinformatics*. 15, doi: 10.1186/1471-2105-15-41.
5. Morrison HG, Grim SL, Vineis JH, Sogin ML. (2013). 16S amplicon fusion primers and protocol for Illumina platform sequencing. *figshare*.
<http://dx.doi.org/10.6084/m9.figshare.833944>
6. Pruesse E, et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acid Research*. 35, 7188-7196.
7. Schloss PD, et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *AEM*. 75, 7537-7541.
8. Schloss PD, Westcott SL (2011). Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *AEM*. 77, 3219-3226.
9. Suzuki S, et al. (2014) Physiological and genomic features of highly alkaliphilic hydrogen-utilizing *Betaproteobacteria* from a continental serpentinitizing site. *Nat Commun* 5:3900.
10. Tiago I, Veríssimo A (2013) Microbial and functional diversity of a subterrestrial high pH groundwater associated to serpentinitization. *Environ Microbiol* 15:1687-1706.