# MICROBIAL DIVERSITY AND EVOLUTION IN GULF OF MEXICO HYPERSALINE ENVIRONMENTS

Lisa Marie Nigro

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in particial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Marine Sciences.

Chapel Hill
2015

Approved by:

Andreas Teske

Samantha Joye

Barbara MacGregor

Carol Arnosti

Brian White

## ABSTRACT

Lisa Marie Nigro: MICROBIAL DIVERSITY AND EVOLUTION IN GULF OF MEXICO HYPERSALINE ENVIRONMENTS
(Under the direction of Andreas Teske)

My thesis investigated the microbial diversity and taxonomic composition of hypersaline and non-extreme sediment environments in the Gulf of Mexico to explore environmental controls on microbial community structure and function. Environmental conditions and resource availability are considered key drivers of microbial diversity and evolution. Extreme environments are traditionally thought to select for more specialized individuals and lower overall diversity, while non-extreme environments foster diverse communities of generalists that are capable of occupying a wider niche. However, extreme environments can also be heterogeneous and complex systems, with a wide variety of energetic resources, providing both challenges and opportunities for diversification.

The sediment microbiomes of this study represented natural hydrocarbon seeps, a deep-sea hypersaline anoxic basin (Orca Basin, 26% salinity), Deepwater Horizon (DWH) oil spill-impacted sediments, and Continental Slope background sediments, all collected in November 2010. These samples were analyzed with ~5 million 16S rRNA gene amplicon sequences. Shannon diversity estimates indicated that Orca Basin hypersaline sediments, Continental Slope sediments, and DWH-contaminated surface sediments had similar species diversity, while hydrocarbon seep sediment diversity was significantly lower. UNIFRAC beta diversity analysis indicated that microbial communities inhabiting Orca Basin hypersaline sediments and

hydrocarbon seep sediments had taxa unique to each of these site types, while sediments from the Continental Slope and DWH area were not statistically more similar to each other than other sediments. Taxonomic analysis showed that seep cores contained higher abundances of ANME-1 and Candidate Division JS1, while Orca Basin hypersaline sediment-associated sequences were dominated by Marine Group I Archaea, Bacteroidetes and halotolerant Deltaproteobacteria.

The microbial composition of the Orca Basin brine was also investigated with 16S rRNA Sanger sequencing, which indicated that Candidate Division KB1-related sequences were the most abundant. A partial KB1 genome sequenced from a single sorted cell revealed that its proteome is dominated by acidic amino acids; therefore, KB1 likely incorporates potassium ions for osmoregulation. KB1 also contains complete enzymes for the acetyl-CoA pathway, and has genes consistent with a glycine betaine transporter cassette as well as genes that may allow for use of glycine betaine as an energy source.

I dedicate this work to my parents, Linda and Dominic Nigro.

## ACKNOWLEDGMENTS

I would first like to thank my thesis advisory committee for providing excellent guidance on this project. To my primary advisor, Andreas Teske for introducing me to the project as well providing funding, space and resources in his lab to conduct experiments. He also gave me guidance and suggestions for the thesis and manuscripts. I am incredibly grateful to Mandy Joye. She provided everything from enthusiasm and project guidance to temporary employment when funding was scarce. Mandy also introduced me to the experience of collecting samples with a submarine and provided me with space and resources in her lab to conduct some of the geochemical measurements. She was also an amazing mentor and provided the encouragement and resources to develop my career. To Barbara MacGregor for being there, even at 1:00 am, to help with my thesis and share stories of life in New England. She also provided detailed feedback and pushed for high standards. To Carol Arnosti for always being a constant voice of reason when things seemed out of hand or unclear, as they often are in graduate school. Lastly, to Brian White who served on my committee as a member that had a very different expertise, yet always attended my committee meetings and encouraged discussion.

I'd also like to thank my "lab family" that was always there for me, both professionally and personally: Tingting Yang, JP Balmonte, Zena Cardman, Luke McKay, Cassandre Lazar, Lindsay D'Ambrosio and Andrew Hyde. Also to the other MASC students in the program, especially Natalie Cohen, Carly Moreno, Jill Arriola, Isaac Westfield and Anna Jalowska.

I also received support from other laboratories that helped on my project. While I was at the University of Georgia, Kim Hunter trained me on several of the geochemical experiments and provided all around experimental support. Kimi Takagi provided great discussions about life, both professional and otherwise, and was instantly there for me when I was a guest in the lab. Matt Saxton provided a great working atmosphere in the molecular group at UGA and was always there to give advice. Also thanks to Siobain Duffy at Rutgers University, who is both an amazing friend and someone I could always rely on to help me find my way.

Most importantly, I'd like to thank my family. My parents, Linda Nigro and Dominic Nigro for not pushing me to go to college, but also encouraging me to do so once I wanted to persue an advanced degree. You were there supporting me the whole way. Also to my brother William Nigro for being a great friend as well as a great sibling.

I would also like to thank the crew and scientists on the R/V Atlantis (leg AT18-2) and R/V Pelican (leg PE22-12) as well as Chief Scientists Mandy Joye (Atlantis) and Mark Pagani and Courtney Warren (Pelican). UNC Research Computing allowed access to their servers, and UNC's bioinformatics provided excellent support, especially by Jeff Roach. Sean Gibbons at University of Chicago gave me initial bioinformatics advice. This project was funded by NSF Award EF-0801741. The UNC Off-Campus Dissertation Fellowship and UNC Dissertation Completion Fellowship provided tuition and stipend assistance.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

AOA: Ammonia oxidizing archaea

AOM: Anaerobic oxidation of methane

DIC: Dissolved inorganic carbon

DOC: Dissolved organic carbon

DWH: Deepwater Horizon

GOM: Gulf of Mexico

MG1: Marine Group 1

OTU: Operational taxonomic unit

PCR: Polymerase Chain Reaction

SAG: Single amplified genome

SRB: Sulfate reducing bacteria

# CHAPTER 1. INTRODUCTION

Understanding the interaction between organisms and their environment has always been a fundamental goal of ecologists. For centuries, scientists have endeavored to understand the abundance and distribution of organisms, energy flow through ecological systems, evolutionary adaptation, succession, and the effects these processes have on the environment (Benson, 2000). Microorganisms arguably make the most significant impact on Earth's ecosystems. Having critical roles in biogeochemical cycling of elements, their abundance and distribution greatly affects the Earth's climate (Singh et al., 2010). They also have incredible evolutionary significance, providing the precursor for multicellular life and the platform for large complex organisms (Olsen et al., 1994). Despite their importance, the ecology of microorganisms is largely understudied. Their small size, lack of distinguishable morphological characters, and tendency to be difficult to isolate and culture are among the reasons they have historically been underappreciated. Advances in molecular cloning and DNA sequencing developed by Sanger (Sanger et al., 1977) greatly changed this perspective. Analysis of the small ribosomal subunit of the rRNA gene suggested that prokaryotic organisms make up two Domains of life (Archaea and Bacteria) (Woese and Fox, 1977) and dominate life's diversity (Pace, 1997). 16S rRNA gene surveys of a wide variety of environments showed that diverse microorganisms are found almost everywhere, even in the Earth's most extreme habitats, including hot springs over 100$^{o}$C, hydrothermal fluids, hot and cold deserts, the deep subsurface, and in salt-saturated lakes (Rappé and Giovannoni, 2003). Extraterrestrial life is surely dominated by microorganisms

(Cavicchioli, 2002) and those found in extreme environments on Earth provide examples of the possible diversity on other planetary bodies. Many microorganisms also consume or neutralize pollutants, providing important environmental ecosystem services (Watanabe, 2001).

While nucleic acid sequencing has provided some insights into microbial communities in diverse environments, molecular sequencing has also been also very costly. Therefore, scientists were limited in the number of sequences they could analyze, often looking at just a few hundred microbial 16S rRNA gene sequences per sample. More recently, "next generation" sequencing has provided a means to sequence millions of sequences from the environment at comparatively low cost, increasing the ability to explore microbial community. When combined with techniques for whole genome isolation and amplification and/or computational algorithms that group sequences into OTUs from metagenomes , next-generation sequencing has also allowed for the exploration of the potential biogeochemical and metabolic roles of the uncultured majority (Baker et al., 2006; Rappé and Giovannoni, 2003; Schuster, 2008; Stepanauskas, 2012). My dissertation combines these new techniques to explore microorganisms in a very dynamic environment, the Gulf of Mexico seafloor.

Deeply buried beneath the Gulf of Mexico are expansive hydrocarbons and a kilometer-thick layer of Jurassic-derived salt. In a process known as salt tectonics, sediment loading and compaction causes movement of the salt layer, which travels to the surface via subsurface conduits, releasing salt and/or hydrocarbons to the seafloor. More rapid expulsion results in mud volcanoes that emit sediment and salt saturated fluid along with oil and gases, while slower expulsion allows salt-saturated fluids to collect in seafloor depressions, forming underwater lakes

(Aharon et al., 1992; Joye et al., 2005; Murray, 1966). The Gulf of Mexico's expansive subsurface hydrocarbon preserves have also attracted deep-sea oil drilling and extraction activities which resulted in the largest and deepest oil spill in U.S. history when the Deepwater Horizon oil platform exploded in April 2010. This rapid expulsion of hydrocarbons into the environment differed greatly from natural seepage and provided another unexplored environment for microorganisms (Ryerson et al., 2012).

In **Chapter 2**, I provide an overview of the Gulf of Mexico sediment microbiome, providing evidence of the microorganisms that inhabit sediments in diverse environments, including hydrocarbon seeps, an anoxic NaCl-saturated brine basin, areas contaminated by the Deepwater Horizon fallout six months after the oil spill, and a site from the Continental Slope, with no apparent hydrocarbon seepage or extreme features. This data contains over 5 million sequences representing an estimated 65,000 OTUs. I provide evidence that Gulf of Mexico seep sites contain microbial taxa reported in other seeps globally, but whose communities, and likely community interactions, also differ from each other. I also provide further evidence that the sediments contaminated with hydrocarbons from the Deepwater Horizon oil spill contain microbial communities that are diverse and similar to uncontaminated sediments, with potentially only a small percentage of oil-degrading microorganisms. This supports evidence from other studies that have claimed that some recalcitrant hydrocarbon material may persist for long periods of time on the Gulf of Mexico seafloor (Kimes et al., 2013; Mason et al., 2014). Lastly, I comment on the microorganisms that are potentially active in the sediments of Orca Basin, a hypersaline anoxic basin (Shokes et al., 1977).

**Chapter 3** further explores the Orca Basin hypersaline environment. I expand on the specific microbial communities that may inhabit the extreme system and also further report the results from a sequencing survey of the red and pink sediments of the basin's slope. In addition to next generation sequencing, I used Sanger sequencing methods that provided longer sequence reads for more accurate taxonomic classification, and present an analysis of potential differences resulting from the two techniques. I also discuss data that resulted from collaboration with several universities, combining sequencing data produced at UNC with geochemistry data provided or processed at the University of Georgia and the University of Bremen, Germany. This combined information indicates that Orca Basin has high microbial diversity that is influenced by geochemistry. Geochemical and sequencing evidence suggested that the basin is a source of biogenic methane produced by methanogens that are known to be active at high salinities and have the potential ability to withstand high concentrations of ammonia (Kadam and Boone, 1996). The amplicons from the hypersaline sediments also contained a surprising abundance of Marine Group I related sequences, potentially active in anaerobic ammonia oxidation. Sequences related to other taxa previously detected in hypersaline environments, including halotolerant Deltaproteobacteria, some of which are likely sulfate reducers, as well as taxa with no known metabolic or biogeochemical function, including uncultured Bacteroidetes and Candidate Division KB1 phyla, were also observed.

The KB1 Candidate Division (Eder et al., 1999) was further explored in **Chapter 4** where I describe preliminary results from a partial genome of an Orca Basin single cell-sorted bacterium that give insights into the evolutionary and metabolic adaptations of this taxon. I analyzed the phylogenetic relationship of KB1 to its closest relatives, the Candidate Division

OP1 (Hugenholtz et al., 1998) and the MSBL 6 group (Daffonchio et al., 2006), which includes taxa that have been detected in high salinity and non-extreme methane and hydrocarbon impacted environments. My results support the evidence that KB1 is related to 3 other taxonomic groups, including the initial Obsidian Pool hot spring clone of the Candidate Division OP1 bacteria, a group containing geothermal spring bacterium *Candidatus Aceotothermum autotrophicum* (NCBI AP011801) and a clade that includes the MSBL6 group (e.g. NCBI AY592717). I further explored the phylogeography of the KB1 Candidate division and determined that all sequences in the ARB/SILVA database that were monophyletic to the initial described KB1 sequences (e.g. AJ133617) were from hypersaline environments. I provide evidence that KB1 has undergone evolutionary changes that allowed it to thrive in a hypersaline environment by showing that KB1 contains an acidic proteome that is consistent with the ability to accumulate high concentration of $K^+$ in its cytoplasm for osmoprotection (Oren, 2005). Through genome analysis, I support for previous claims that KB1 can use glycine betaine as an energy source (Yakimov et al., 2013). I also present evidence that KB1 contains genes encoding for enzymes of the acetyl-CoA pathway that are likely orthologous with OP1. Lastly, I provide a hypothesis of the potential metabolic processes of KB1, and the possibility that KB1 can oxidize glycine betaine by utilizing the acetyl-CoA pathway in reverse.

# CHAPTER 2: MICROBIAL DIVERSITY OF DIVERSE GULF OF MEXICO SEDIMENTS

## 2.1 Introduction

Beneath the Gulf of Mexico seafloor lie expansive hydrocarbon stores and a thick Jurassic-derived salt layer. In a process known as salt tectonics, sediment compaction causes the salt layer to move, shifting the overlying sediments. In some areas, salt and/or hydrocarbons are released to the seafloor via fault conduits, resulting in such features as salt and/or hydrocarbon-emitting mud volcanoes, brine pools, and oil/gas seeps. These systems are a source of natural oil and gas to the Gulf of Mexico sediments and waters and an extreme and challenging environment for life (Aharon et al., 1992; Murray, 1966; Oren, 2005).

While natural salt and hydrocarbon seepage have been observed in many locations on the Gulf of Mexico slope, much of the hydrocarbon reserve is trapped in the subsurface (Kennicut et al., 1988). The Gulf of Mexico is, therefore, also the focus of extensive off-shore drilling, with the risk of introducing large amounts of oil into the ocean. The Deepwater Horizon oil spill was such an event where more than 4 million barrels of oil were released into the Gulf of Mexico (McNutt et al., 2012), some of which was deposited on the seafloor (Passow et al., 2012; Ryerson et al., 2012).

The microbial communities responsible for biogeochemical cycling in these hydrocarbon-impacted systems are poorly understood. Less than 1% of all microorganisms have been successfully cultured, making species identity and function difficult to correlate (Rappé and Giovannoni, 2003; Stewart, 2012). The use of 16S rRNA gene amplification and sequencing as a molecular barcode to determine microbial species diversity and identity revolutionized the field

of microbial ecology (Rappé and Giovannoni, 2003). Several properties of the 16S rRNA gene make it ideal for microbial taxonomic analysis, including presence in all organisms and large size (>1400 bases) with regions that evolve at different rates (Woese, 1987). However, due to high costs of molecular cloning and Sanger sequencing, studies were limited to analyzing only a few samples at once and to obtaining just hundreds of sequences from an environment, accessing only a fraction of the microbial diversity. Recent advances in 'next-generation' sequencing have made it possible to analyze millions of sequences at the same cost as a few hundred Sanger sequences (e.g. MacLean et al., 2009; Schuster, 2008).

**2.2 Site Descriptions**

Samples were collected from a variety of sites in the Gulf of Mexico with a shipboard multicorer. Sampling was performed around 3 brine pools, 3 mud volcanoes, a deep anoxic hypersaline basin and the Deepwater Horizon oil spill site (Mississippi Canyon lease block 252). (Figure 2.1, Table 2.1).


**Figure 2.1 Sampling site locations.** Samples were collected on the R/V Atlantis expedition AT-18-2 in November to December 2010 from site AC601 to MC252.

*Garden Banks 425:* The Garden Banks 425 (GB425) lease block is on the western edge of the Auger basin and contains an area of salt bodies on its western border. GB425 also contains a large active mud volcano. The site frequently experiences mud flows and contains a small basin where brine collects. The mud volcano fluid has an estimated salinity of 133, a sediment composition of 65%, a pH of 7.4, and temperatures similar to the overlying bottom water (~5$^{o}$C). Mud flow events prevent the long-term establishment of chemosynthetic macroorganisms, but patches of microbial mats have been

observed throughout the area. Sediment in the area contains high molecular weight hydrocarbons and gas hydrates (Joye, 2010; Joye et al., 2005; MacDonald et al., 2000).

*Garden Banks 697 Hot Site***:** The GB697 lease block is characterized by gas and oil seepage and active salt tectonics. It contains a small, active mud volcano that discharges gas and fluidized briny mud, with a salinity of ~210. Temperature of the fluid was above bottom water (~25$^{o}$C). Microbial mats and clams were also observed at the site (Joye, 2010).

*Orca Basin***:** Orca Basin is a hypersaline anoxic depression located on the Texas-Louisiana continental slope (26 56'N, 91 19'W) that is approximately 2400 m deep. The basin is comprised of two deep depressions containing anoxic brine pools (North and South Basin) with a salinity of ~260 separated by a saddle region (Sheu, 1990). The salt is mainly composed of halite and originates from the Louann salt formation that resulted from evaporitic conditions during the Jurassic (Murray, 1966). The strong density gradient at the seawater-brine interface prevents exchange of solutes, keeping the bottom of the basin anoxic (Figure 2.2). The brine-seawater interface also serves as a particle trap, with total suspended matter concentrations as high as 880 µg/L, up to 8 times higher than in the overlying water (Trefry et al., 1984). In comparison to seawater, the brine is enriched in $Na^{+}$, $K^{+}$, $Ca^{2+}$, $Cl^{-}$ and $SO_4^{2-}$ and slightly depleted in $Mg^{2+}$ (Sheu, 1990), and is more acidic, with a pH of 6.5. The sediment beneath the brine pools is black, while the saddle region dividing the North and South basins has been described as laminated with red sediments that are mainly hematite (Sheu, 1990; Sheu and Presley, 1986). Orca Basin is discussed in more detail in Chapter 3.

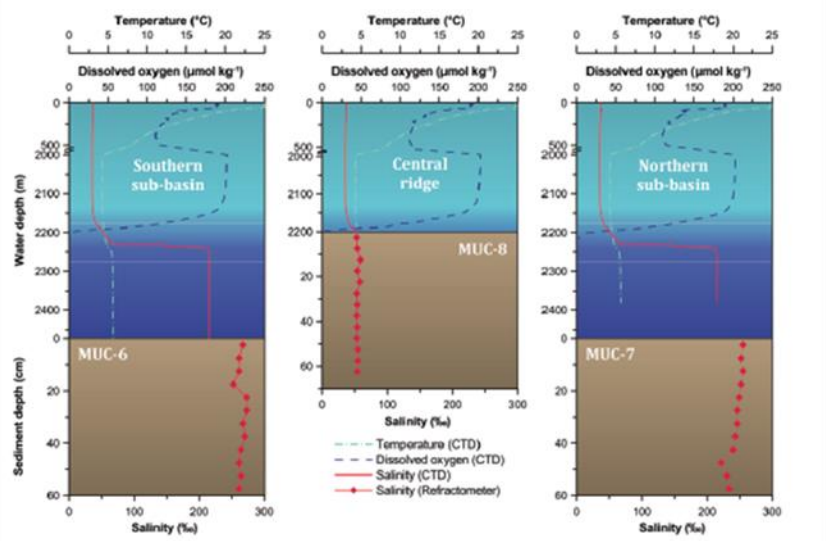



**Figure 2.2 Orca Basin brine and seawater CTD profiles.** Figure provided by Felix Elling, University of Bremen.

*Green Canyon 600:* The Green Canyon 600 (GC600) site is on the upper continental slope of the Gulf of Mexico and is among the largest natural seeps in the Gulf of Mexico. GC600 is characterized by a large fractured carbonate ridge, a large amount of oil and gas expulsion, and many chemosynthetic fauna, including clams, tube worms, and mussels. Microbial mats are widespread and sediment cores retrieved from the site often contain visible oil. Oil slicks are also sometimes observed at the ocean surface (Joye et al., 2005).

*Mississippi Canyon 118:* Mississippi lease block 118 (MC118) is located on Louisiana's continental slope, and has been one of the most studied sites on the Gulf of Mexico continental shelf after the implementation of a seafloor observatory near gas hydrate mounds (McGee et al., 2009). Besides the hydrate mounds, the site features hydrocarbon and gas seeps, brine flows, carbonate pavements and mud volcanoes. Chemosynthetic animals, corals, crabs and *Beggiatoa* mats have all been observed and studied at this site (Lutken, 2013).

| *Core* | *Site Name* | *Site Description* | *Latitude* | *Longitude* | *Date* | *Depth (mbsf)* | *Core Description* |
|---|---|---|---|---|---|---|---|
| MUC 3 | GB425 | seep area | 27 33.153 N | 92 32.497 W | 15-Nov-10 | 567 | ochre-brown color surface to 3 cm, chunks of carbonate to bottom of core. |
| MUC 4 | GB425 | seep area white microbial mats | 27 33.140 N | 93 32.473 W | 15-Nov-10 | 567 | very gassy and oily, no visible zonation, reported to hit microbial mat |
| MUC 5 | GB697 | small mud volcano area | 27 19.2197 N | 92 6.6406 W | 17-Nov-10 | 1014 | no description |
| MUC 6 | Orca Basin | Orca Basin South brine pool | 26 54 .141 N | 91 21.793 W | 18-Nov-10 | 2500 | Black in color throughout core |
| MUC 7 | Orca Basin | Orca Basin North brine pool | 27 0.007 N | 91 17.000 W | 19-Nov-10 | 2500 | Black in color to 45 cm, then gray |
| MUC 8 | Orca Basin | Orca Basin slope sediment | 26 56.260 N | 91 17.100 W | 20-Nov-10 | 2500 | Red with brown laminations |
| MUC 12 | GC600 | hydrocarbon-rich cold seep | 27 21.91 N | 90 33.85 W | 26-Nov-10 | 1250 | carbonate-rich, oil inclusions especially below 20 cm |
| MUC 13 | GC246 | Continental Slope | 27 7.443 N | 90 17.239 W | 26-Nov-10 | 867 | ochre-brown, darker in color from 25 cm to bottom |
| MUC 14 | MC118 | gas hydrate area | 28 51.0 N | 88 29.5 W | 28-Nov-10 | 1000 | shallow upper brown debris, possibly DWH oil-impacted |
| MUC 18 | MC252 | DWH well head area | 28 40.917 N | 18.584 W | 29-Nov-10 | 1621 | Potential oil floculent material 0-3 cm on top of core |
| MUC 19 | MC252 | DWH well head area | 28 43.444 N | 88 21.854 W | 29-Nov-10 | 1621 | Potential oil floculent material from 0 to 4 cm on top of core, slight oil sheen |
| MUC 22 | MC343 | South of DWH well head | 28 42.364 N | 88 21.917 W | 1-Dec-10 | 1700 | No description |

**Table 2.1 Site and Collection Information**

*Mississippi Canyon 252:*MC252 is on the Gulf of Mexico continental slope south of Louisiana. The site is the location of the Deepwater Horizon drilling platform that exploded on April 20, 2010, releasing 4.1 million barrels of liquid oil (McNutt et al., 2012) and ca. 1.7 x $10^8$ kg of gas (Reddy et al., 2012). The complex mixture of oil and gas migrated to the surface or was

trapped in underwater currents. After physical, chemical and microbial weathering (Ryerson et al., 2012, Passow et al. 2012), some of the oil remnants were observed as a dark flocculent layer on the top of sediment cores in the area (Figure 2.3).

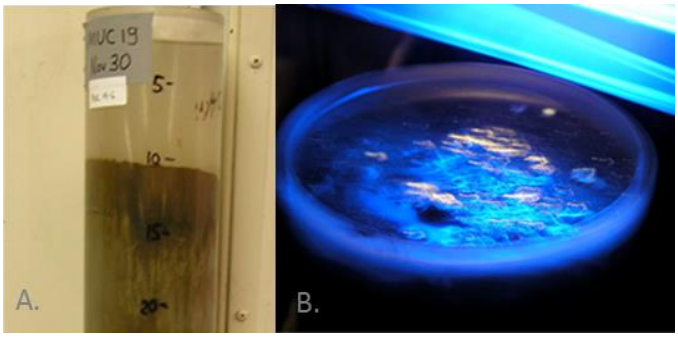



**Figure 2.3 DWH Site Core MUC 19. A.** Flocculent rusty brown sediment was observed in the first 3 cm of the core. **B.** Potential oil droplets flouresce under UV light.

## 2.3 Materials and Methods

### 2.3.1 Sample collection

Sediment core samples were collected on the *R/V Atlantis* (expedition AT18-2) in November to December 2010 (Table 2.1) with a shipboard multicorer. Sediment cores were immediately capped and stored at 4°C until they were processed within 12 hours of collection. Multicore samples were divided into 2.5 to 5 cm sections. Extruding materials were either autoclaved or treated with ethanol prior to sectioning. Samples were immediately frozen at -80$^{o}$C until processing for DNA extraction.

### 2.3.2 Nucleic acid extraction and sequencing

Sequencing samples were thawed on ice and 0.5 g of sediment was extracted with a 1 mL syringe and placed in a 96 well MoBio bead-beating plate for subsequent extraction with the Powersoil-htp 96 Well DNA Isolation Kit (MoBio). DNA extraction, 16S rRNA gene amplification of the V4 fragment (prokaryotic universal), and sequencing on the Miseq2000 platform was performed at Joint Genome Institute (Walnut Creek, CA) according to the Earth Microbiome Project standard protocol (Caporaso et al., 2012; Gilbert et al., 2010).

### 2.3.3 Data analysis

Amplicon sequences were analyzed with the QIIME (Caporaso et al., 2010a). Sequence quality scores were plotted and sequences truncated at a cutoff of quality score of 25. Paired ends

were joined with the fastq-join method (Aronesty, 2011) and sequence data demultiplexed and identified by the 12 base golay error correcting barcodes. Sequences were grouped into operational taxonomic units (OTUs) and putative chimeras removed with UPARSE (Edgar, 2013). Singletons were removed. Taxonomy was assigned to OTU sequences using the Greengenes 13.8 classifier (McDonald et al., 2012). Sequences were aligned with the python-implemented NAST alignment algorithm (Caporaso et al., 2010b). Phylogenetic trees relating OTUs were constructed using the FastTree method (Price et al., 2009). Alpha diversity was calculated using Shannon diversity (Shannon, 1948) and Faith's PD whole tree (Faith, 2006) metrics based on multiple rarefactions of the data. Box plots were created for each designated group and significant differences estimated with a pairwise t-test using nonparametric Monte Carlo permutations to determine p-values. Beta diversity was estimated using UniFrac weighted and unweighted metrics (Hamady et al., 2009) and visualized with Principal Coordinate Analysis (PCoA).

**2.4 Results and Discussion**

A total of 5,092,497 amplicon reads from sediment core samples (Table 1) were obtained after demultiplexing and quality control filtering, and a total of 64,960 OTUs were analyzed after removal of singletons and potentially chimeric sequences. Several sequences, especially from the extremely saline samples, were removed after BLAST searching (Altschul et al., 1990). These sequences were associated with putative human microbiome contamination (100% identity and coverage in 100 top hits of genera associated with *Corynebacterium*, *Staphylococcus,* and *Streptococcus* spp.). However, it should be noted that other explanations for the presence of these OTUs, including preservation of DNA in the hypersaline sediments (Boere et al., 2011) or

unknown role of these genera in Orca Basin. Several cores did not yield sequences for many of the sediment layers of MC252 site cores (e.g. MUC 17) and were not plotted to show taxonomy distribution downcore, but samples containing at least 5000 amplicon reads were included in diversity analyses.

The 16S rRNA amplicon sequences indicated diverse taxonomic representation in all Gulf of Mexico sediment cores processed (Figures 2.4 and 2.5). Amplicons assigned to the Proteobacteria were the most abundant in almost all samples and at nearly all sediment depths, from the surface up to 60 cm downcore. There were notable differences in the presence of particular taxa based on the environment the sample was collected from. Methanomicrobia-related sequences were more prominent in sediment core samples associated with seep sites, while Bacteroidetes-related sequences were more abundant in hypersaline sediments. Chloroflexi-related sequences were present in higher abundance in "background" slope sediments and cores taken from near the Deepwater Horizon wellhead.

### 2.4.1 Hypersaline environment associated taxa

The Orca basin hypersaline sediment cores (MUC 6 and MUC 7) had noticeably higher abundances of Bacteroidetes (up to 28% total amplicon reads), typical of hypersaline environments (e.g. Ley et al., 2006). Amplicon abundance of phylum-level taxa did not change drastically downcore. There was a surprisingly high abundance of Thaumarchaeota, specifically Marine Group I- related sequences (up to 38%), which persisted throughout the core (Figure 2.4). The Orca Basin cores had among the lowest proportions of Gammaproteobacteria-related amplicons (Figure 2.5), and proportionally higher Deltaproteobacteria-related sequences at the surface (0-5 cm), likely due to the anoxic environment. The hypersaline cores also contain sequences related to taxa not observed in the other Gulf of Mexico cores, including members of

the Family Desulfohalobiaceae, which are among the most abundant of the Proteobacteria sequences, comprising of up to 6.6% of the total sequences. Desulfobacteraceae were also more abundant in the Orca Basin cores (Figure 2.5). Detailed analysis of the Orca Basin sediment samples is presented in Chapter 3.

### 2.4.2 Methane cycle-associated taxa

Most cores from seep sites, in which the presence of gas bubbles and/or oil was observed (MUC 5 and MUC 12), had expected higher abundances of Methanomicrobia-related archaeal sequences, while the other sights had low abundances of known methane cycling organisms (Figure 2.4). The Methanomicrobia include most of the known methanogens as well as archaea that are capable of anaerobic oxidation of methane (ANME-1, ANME-2, and ANME-3) (reviewed in Offre et al., 2013). The seep samples were dominated by sequences related to ANME-1, with few methanogen-related sequences detected (Figure 2.5). One exception was a core from the GB425 mud volcano and seep area (MUC 4) contained visible oil and had a strong gas smell, but contained less than 1% of amplicon sequences associated with Methanomicrobia. However, it is possible that sequences associated with *Methylococcales* that were abundant at the GB425 seep site may represent putative methane oxidizers. The majority of these sequences identified as *Methylococcales*, had high BLAST identity (99%) to *Methyloprofundus sedimenti*, an aerobic methane and methanol-oxidizing species in the 'deep-sea 1' clade of methanotrophs (Tavormina et al., 2015). Interestingly, this OTU is observed even at the deepest depths (55 cmbs) of the core, well below the likelihood of available oxygen (Figure 2.7). Potentially, the dominant *Methylooccales*-associated OTU contains broader or alternative metabolic capabilities than reported for *Methyloprofundus sedimenti* (Tavormina et al., 2015), and this group's biogeochemical role in GB425 remains unknown.

Additional explanations for low abundance of Methanomicrobia in MUC 4 could also be due to methodological bias in the amplification and sequencing. Very few (<2%) ANME-2 related sequences were detected in any of the samples. The primers used in this study were analyzed for percent match with the SILVA NR database using TestPrime 1.0 (Klindworth et al., 2012) , and 75% of ANME-2 sequences have 100% match to the primer sequence, however, the low representation ANME-2 was surprising based on previous studies of Gulf of Mexico sediments (Table 2). Future investigation with alternative primers or metagenomic/transcriptomic/proteomic studies could provide greater insights to methane cycling at these sites.

### 2.4.3 Japan Sea Group 1 (JS1)

Sequences associated with the uncultured JS1 group were also proportionally abundant in several of the cores, including the cold seep area samples and especially in the red sediments of the Orca Basin ridge sediments (Figure 2.4). JS1 was also observed in lower abundances (<5%, see Figure 3.5, Chapter 3) in Orca Basin hypersaline sediments. Initially, JS1 was considered part of the OP9 Candidate Division that was first detected in Yellowstone National Park's Obsidian Pool (Hugenholtz et al., 1998). It was subsequently detected by 16S rRNA gene sequencing in a wide variety of anaerobic marine deep-sea sediments and was proposed to be its own candidate phylum group (Webster et al., 2004). JS1-associated sequences have been among the most abundant in next generation sequencing 16S rRNA gene surveys of hydrocarbon seeps and sulfate methane transition zones (SMTZ) (Ruff et al., 2015). JS1 has also been detected in hypersaline environments, including the Guerrero Negro hypersaline mats (Ley et al., 2006). Near full-length 16S rRNA gene analysis of JS1 and OP1-associated taxa, as well as phylogenetic analysis of conserved protein-coding (housekeeping) genes in single cell amplified

genomes (SAGs), supported the monophyly of the combined JS1/OP1 group, and also indicated that JS1 was distinct from OP9 (Nobu et al., 2015). However, given the genetic distance between JS1 and OP1 is smaller (<75%) than what generally delineates separate phyla, the authors suggested they be one phylum, Artibacteria, with separate class affiliation.

The broad metabolic role of JS1 was also indicated by comparative genome analysis (Nobu et al., 2015). JS1-associated taxa had genes consistent with anaerobic fermentation of carbohydrates and organic acids like acetate or propionate. It is possible that the higher percent abundance of JS1 in the Orca Basin red sediment and some sections of the seep cores (MUC 4, 5 and 12) may indicate utilization of organic acids, but these were not measured at any of the core sites in this study, and the presence of molecular machinery does not necessarily confer function. However, JS1 was among the bacterial groups that did incorporate $^{13}$C-labeled acetate in a stable isotope probing study of sulfate-reducing sediments of Aarhus Bay (Na et al., 2015), indicating that at least some JS1 taxa can utilize acetate. Syntrophy was proposed as a possible metabolic strategy of JS1 since the multiple genome analysis indicated a lack of respiratory pathways (Nobu et al., 2015). While syntrophy cannot be assessed purely by amplicon abundance, discussion of potential syntrophy is discussed in the next section.

| Sample Type | Sample Description | Dominant Bacterial Sequences | Dominant Archaeal sequences | Method | Reference |
|---|---|---|---|---|---|
| Cold seep microbial mat | sediment below Beggitoa mat 0-3 cmbsf | Eel-2 group | ANME-1b, ANME-2a, ANME-2c, DHVE-8 | 16S rRNA clone library/Sanger | (Lloyd et al., 2010) |
| | sediment below Beggiatoa mat 12-15 cmbsf | Eel-2 group | ANME-2a, ANME-2c, ANME-1b, DHVE-8 | | |
| | sediment edge of Beggiatoa mat 0-3 cmbsf | SEEP-SRB1, Desulfobulbaceae | DHVE-8, ANME-2a, ANME-2c | | |
| | sediment edge of Beggiatoa mat 12-15 cmbsf | Eel-2 group, SEEP-SRB1 | ANME-1b | | |
| Cold seep area | sediment outside Beggiatoa mat 0-3 cmbsf | diverse Gammaproteobacteria, diverse Alphaproteobacteria, Acidobacteria | MG1 | | |
| | sediment outside Beggiatoa mat 12-15 cmbsf | SEEP-SRB1, Desufobacterium anilini group | ANME-2a | | |
| Cold seep | oily sediments at 4 Gulf of Mexico seep sites | SEEP-SRB3, SEEP-SRB4, Desulfobacterium anilini group, JS1 | ANME-1, ANME-2 | 16S rRNA clone library/Sanger | (Orcutt et al., 2010) |
| Gas hydrate | Green Canyon, within gas hydrate | SEEP-SRB1, Syntrophaceae, Desulfobacter, Sulfurovum, Firmicutes | ANME-2C, ANME-1B, Methanomicrobiaceae | 16S rRNA clone library/Sanger | (Mills et al., 2005) |
| Hydrocarbon-influenced brine lake | brine pool methane seep | Eel-2 group ( SILVA v121 SEEP-SRB2), JS1 | ANME-1b | 16S rRNA clone library/Sanger | (Lloyd et al., 2006) |
| DWH oil contaminated | contaminated surface September 2010-October 2010 | uncultured Gammaproteobacteria (GG OTU ID 248394), Colwellia, Rhodobacteraceae | n.d. | 16S Illumina amplicon, metagenomics | (Mason et al., 2014) |
| DWH oil contaminated | contaminated surface1.5-3 cmbsf September 2010-October 2010 | Rhizobiales, Rhodobacterales, Desulforomondales, Desulfovibrionales, Desulfobacterales, Alteromondales | Euryarchaeota, Thaumarchaeota | metagenomics | (Kimes et al., 2013) |
| DWH oil contaminated | contaminated surface 0-2 cm 'May 2011 | Methylococcus, Methylobacter, Actinobacteria, Firmicutes and Chloroflexi | n.d. | 16S Illumina amplicon | (Liu et al., 2012) |

**Table 2. 2 Microbial Taxa Associated with Gulf of Mexico Sediments**

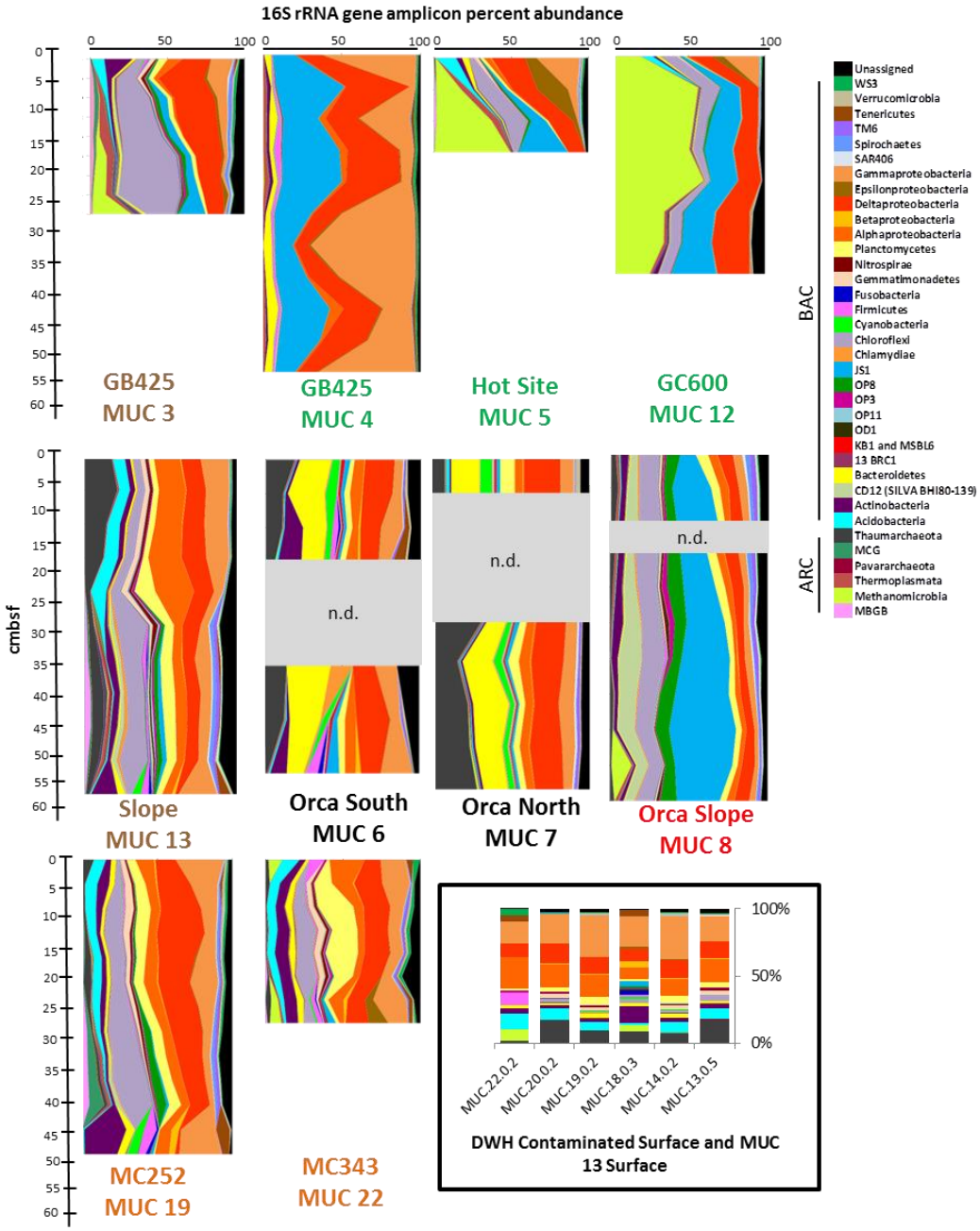



**Figure 2.4 Relative taxonomic abundances of Gulf of Mexico samples inferred from 16S rRNA amplicon frequency.** Samples were from seep sites (green), hypersaline Orca Basin (black), the red sediments of the Orca Basin slope (red), potential Deepwater Horizon contaminated sediments (orange) and continental shelf sediments (brown) with no visible hydrocarbon presence or otherwise known extreme environmental condition.

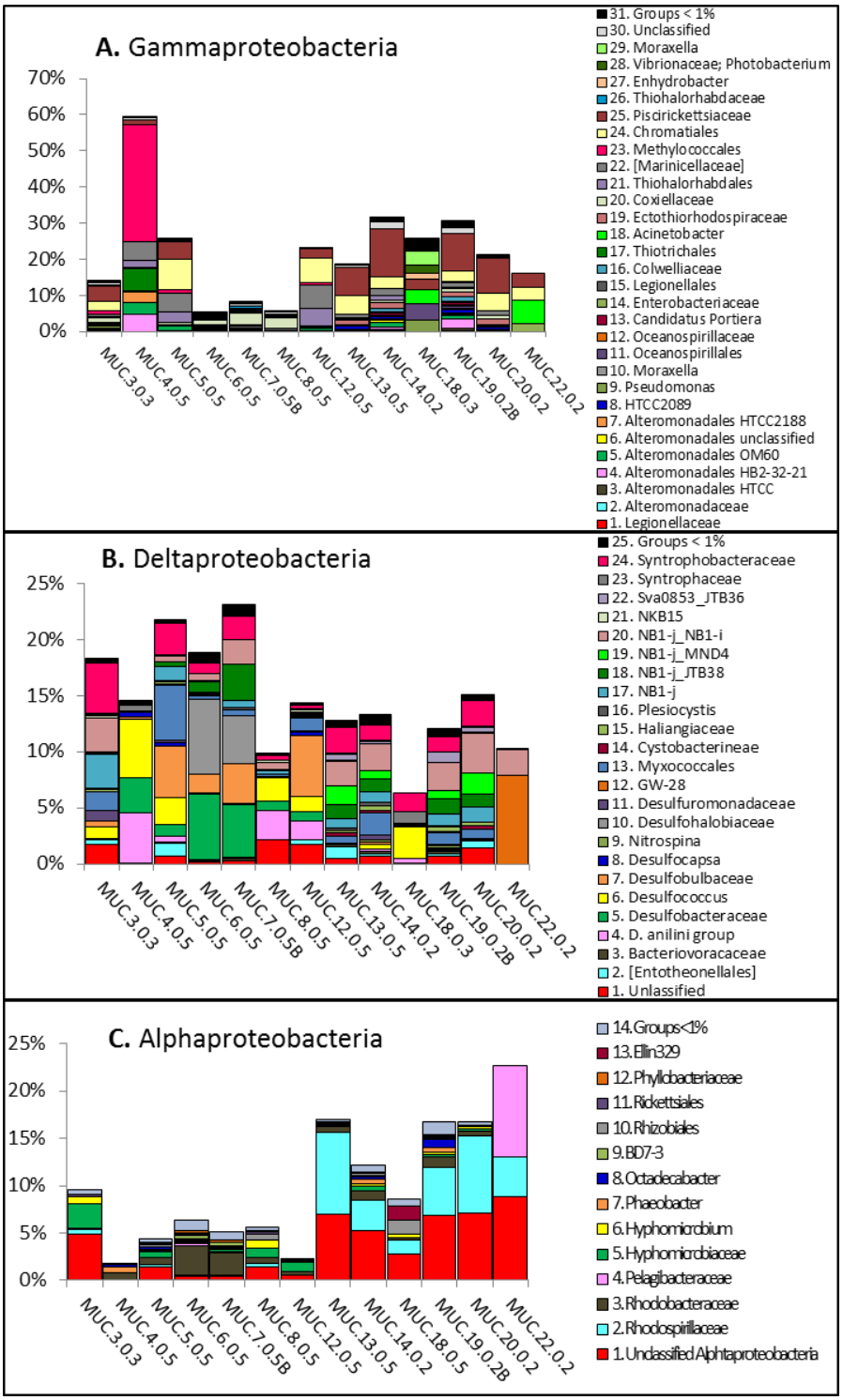



**Figure 2.5 Relative abundance of Proteobacterial-related amplicon sequences from surface sediments.** Percent abundance is given as a proportion of total bacterial and archaeal-associated amplicon reads. **A.** Gammaproteobacteria, **B.** Deltaproteobacteria and **C.** Alphaproteobacteria.

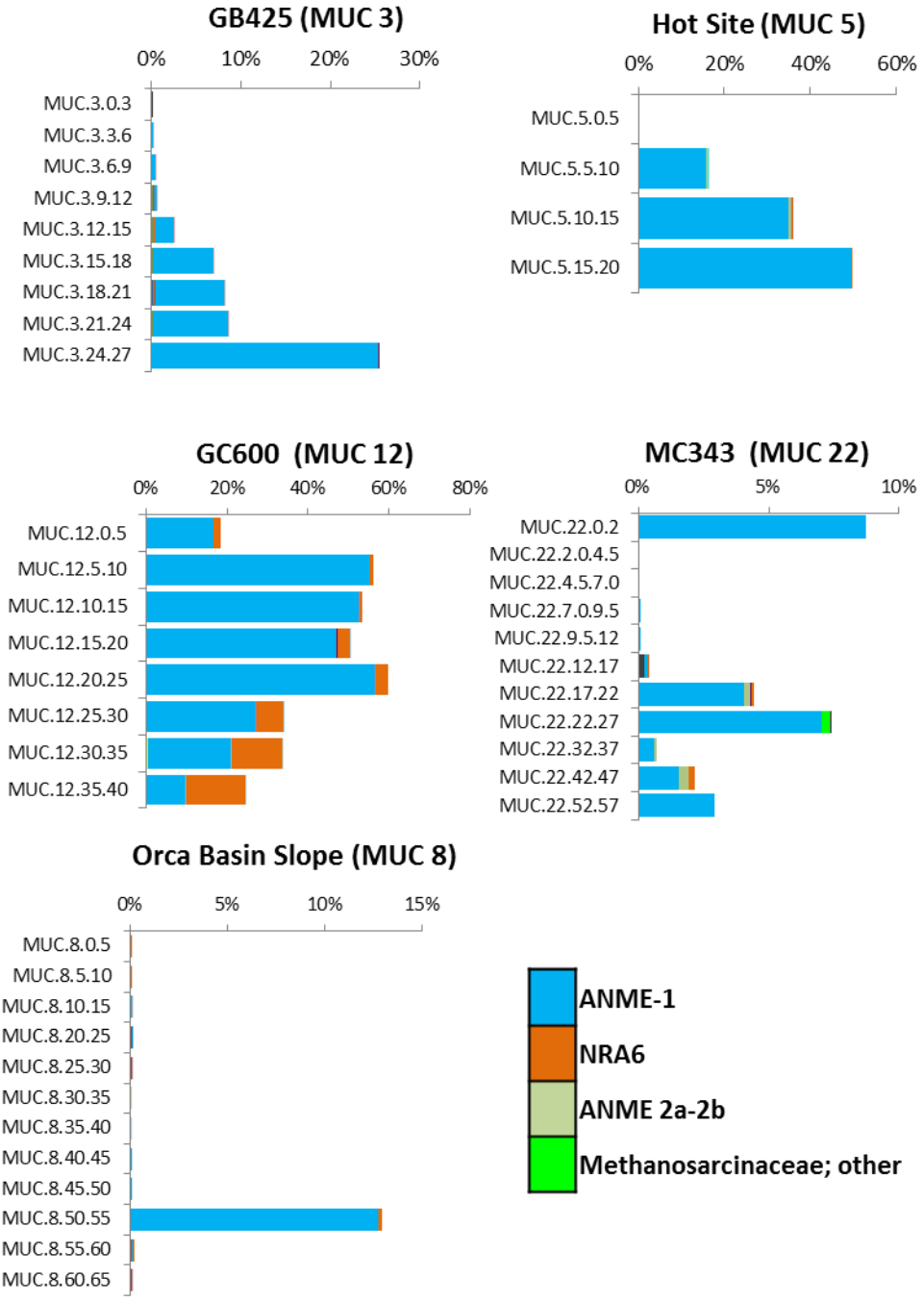



**Figure 2.6 Relative abundance of Methanomicrobia.** Methanomicrobia-related sequence abundances for selected cores. Percent abundance of total bacterial and archaeal- associated amplicon sequences.

### 2.4.4 Patterns of microbial taxa in seep-associated cores

The GB425 core that was described as containing oil and gas (MUC 4; Table 2.1) had high abundances of JS1, Deltaproteobacteria and Gammaproteobacteria, bacterial lineages that were also among the most abundant in the other seep cores, and have been noted to be common phylotypes associated with seeps globally (Ruff et al., 2015). JS1 and ANME-associated sequences correlated in abundance in several of the cores, including MUC 3 (carbonate-rich core) as well as MUC 5 and MUC 12 (seep cores), (Figure 2.6). JS1-associated 16S rRNA gene sequences were previously noted to correlate in abundance with ANME-associated sequences, but no suggestion of syntrophy with ANME was detected with fluorescent in-situ hybridization (FISH) (Vigneron et al., 2014). However, syntrophy between JS1 and ANME- SRB consortia, or other mutualistic interactions between the groups, can also not be ruled out.

Interestingly, in MUC 4, JS1-related amplicon sequences had a similar abundance pattern downcore to Deltaproteobacteria, while Gammaproteobacteria-related sequences had the opposite pattern and were inversely correlated to both Deltaproteobacteria and JS1 sequence abundance (Figure 2.7 and Figure 2.8). The dominant Deltaproteobacteria sequences in MUC 4 were 'Desulfarculaceae' (Greengenes and SILVA classification) and Desulfococcus (SILVA SEEP-SRB1), which were nearly equally abundant downcore. On further inspection, 'Desulfarculaceae' sequences were monophyletic with *Desulfobacterium anilini* (Schnell et al., 1989), a sulfate reducing species that was observed to degrade aromatic and aliphatic compounds. This species was recently reclassified into the genus *Desulfatiglans*, along with an isolate monophyletic with *D. anilini* that was shown to degrade 4-chlorophenol as well as other aromatic compounds, acetate, butyrate and pyruvate (Suzuki et al., 2014). However, higher phylogenetic classification has not been fully investigated. The Deltaproteobacterial group

SEEP-SRB1 have been noted to be ubiquitous in seep environments, and have been known to function syntrophically with ANME (Knittel and Boetius, 2009). It is difficult to assess the potential relationship, if any, between these taxa without further investigation. Potentially, JS1 correlation with Deltaproteobacteria in MUC 4 (Figure 2.7 and Figure 2.8) may indicate syntophy or preference for use of byproducts of sulfate reduction for energy. The significance of the inverse correlation of JS1-related sequences (and Deltaproteobacteria) with Gammaproteobacteria-related sequences in MUC 4 also remains unknown. The majority of these sequences were identified as *Methylococcales*, and as discussed previously, OTUs from all depths had high BLAST identity (99%) to cultured *Methyloprofundus sedimenti*, an aerobic methane and methanol-oxidizing species (Tavormina et al., 2015). However, the uncultured MUC 4 *Methylococcales* OTU may possess alternative energy sources and anaerobic lifestyle, as suggested by its abundance well beyond the depth expected to be anoxic (55 cm).

### 2.4.5 Deepwater Horizon area associated taxa

Microbial communities from MC252 surface sediments (MUC 17, 18, 19) were consistently diverse, with sequences representing many different phyla (Figure 4). Cores that were taken further away from the well head (MUC 14, 20 and 22) resembled the Continental Slope core (MUC 13) at the phylum or class taxonomic level. One notable difference is that the MUC 22 surface sample did contain ANME-1 related sequences (~8%) – an unlikely component of the surface microbial community, considering the oxidized and non-methanogenic character of this sample - while the other cores had Thaumarchaeota as the most abundant archaea-associated sequences.

Proteobacteria were the most abundant phylum in all surface cores, with Acidobacteria often the next most abundant. Within the Proteobacteria, Alpha- Delta- and Gammaproteobactria

associated sequences were the most abundant. One study that combined metagenomics and 16S rRNA gene amplicon sequencing with geochemical analysis (Mason et al., 2014), observed that sequences related to an uncultured Gammaproteobacteria phylotype in the Family *Altermonadaceae* (NCBI EU287169) were abundant (up to 13%) in sediment samples collected near the wellhead in September to October 2010. This phylotype was associated with the DWH deep-sea plume during the oil spill (Kessler et al., 2011) and has been associated with Gulf of Mexico hydrocarbon-impacted sediments (Kleindienst et al., 2012). The Altermonadaceae phylotype (100% identity) was also detected in our surface samples, with the highest percent abundance at 5% in MUC 4 (seep core) and 2.6% in MUC 19 (MC 252). It was also observed in other seep cores in this study, but at a much lower (<0.1%) sequence abundance (Figure 5). Members of the Colwelliaceae were also detected in one of the DWH-affected cores (MUC 19), some sequences of which had high identity (100%) to sequences from DWH surface slicks (NCBI KF786789), but again at a lower percent abundance than samples reported in September-October 2010 (1.2 % in MUC 19 vs 6% in Mason *et al.* 2014). A highly abundant Rhodobacteraceae sequence that was observed in the Mason *et al.* 2014 study (up to 13% of 16S rRNA gene amplicon reads) was detectable in MUC 19 as well, but only a few hundred reads were associated with the phylotype (0.5%). This was similar to the percent abundance observed at MUC 14 (MC118, potential minor DWH oil impacted) and MUC 4 (0.6%), a natural seep site.

MUC 18 had additional sequences with potential relation to oil degrading taxa, including *Pseudomonas* (3.2% total amplicon reads) and Oceanospirillales (4.3% amplicon reads). However, *Pseudomonas* species are also widespread and are common contaminants (La Duc et al., 2009), and the Oceanospirillales sequences were mainly associated with thiotrophic symbionts (e.g. NCBI KF657323) or uncultured sediment bacteria. The Oceanospirillales

initially reported in the deep-sea plume (Hazen et al., 2010) were not detected. Kimes et al. (2013) reported higher abundances of Desulfuromonadales, Desulfovibrionales, and Desulfobacterales in cores with higher DWH petrocarbon contamination (especially in the affected layer below 1 cmbsf) in comparison to cores that were likely uncontaminated, but this was not observed in this study (Figure 2.5). MUC 22 (MC 343) did have abundant (in comparison to other sediment samples) Deltaproteobacteria-related amplicon sequences (7.9%) in an uncultured lineage that had sequence similarity to a peat soil clone (NCBI AB364749) and PCB-spiked Ohio River sediments (GU180168). MUC 18 (MC 252) also had higher abundance of Tenericutes (5%) that are commonly associated with infections in farm animals but also are found in a wide variety of marine and freshwater environments, as well as higher abundance (3.5%) of an uncultured taxon that had the highest similarity (99%) to a mud volcano (NCBI KF440323) cloned sequence and sequences from methane seep sediments (e.g. NCBI FJ264786).

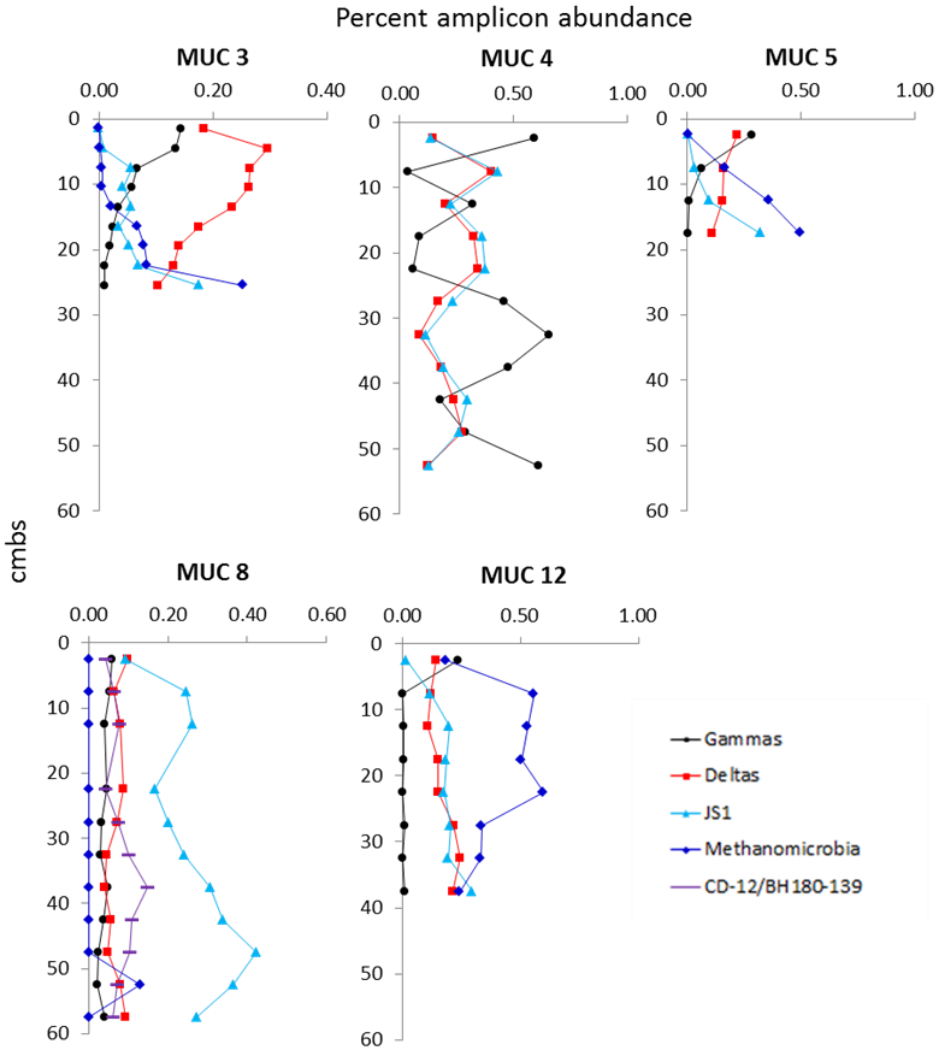



**Figure 2.7 Percent abundance patterns of dominate taxonomic groups of selected cores.** Taxonomic group was not plotted if the percent abundance was less than 1%.

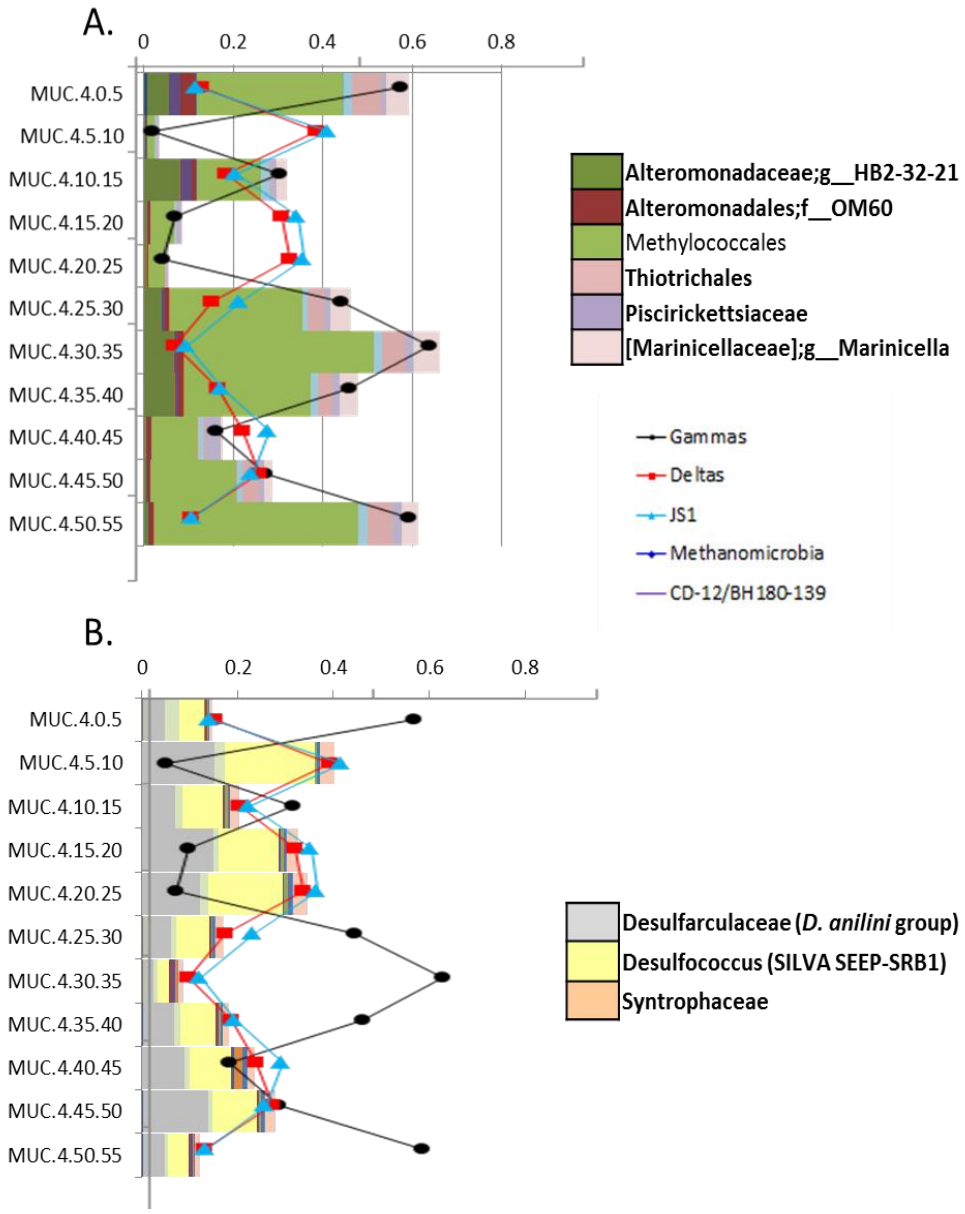



**Figure 2.8 Percent abundance patterns of JS1, Gammaproteobacteria, and Deltaproteobacteria.** Barcharts of more defined taxonomic resolution are provided for Gammaproteobacteria (A.) and Deltaproteobacteria (B.).

### 2.4.6 Beta diversity analysis

Differences and similarities among the microbial communities of the Gulf of Mexico sediment samples were assessed using UniFrac diversity (Hamady et al., 2009; Lozupone and Knight, 2005) followed by principle coordinate analysis (PCoA). UniFrac uses the phylogenetic distance between groups of taxa to assess similarities between groups. Unweighted UniFrac analysis, which assesses each OTU with the same weight regardless of how many times the OTU occurs in the sample, is shown in Figure 2.9. Samples were categorized as hydrocarbon seep (orange), hypersaline (black), Orca Basin slope (MUC 8, red) DWH contaminated (MUC 17, 18, 19 sections with visible oil, dark blue), DWH potentially contaminated (MUC 14, 20, 21, 22, yellow), and carbonate-rich slope (purple), and slope (continental slope, light blue). None of the sample types were statistically different (P values need to be reanalyzed) as a group, although removing 2 depths from the hypersaline group (Orca Basin South 50-55 cm and Orca North 45-60 cm), did result in the hypersaline sample being significantly different than other samples. Hydrocarbon seep sequences differed by the presence of ANME-1, both between cores, with MUC 4 having very low abundance of ANME-1 and MUC 12 having nearly 50% of the sequences of this phylotype, and by depth, with MUC 5 ANME-1 sequences increasing with depth (Figure 2.4).

The bacterial and archaeal sequences were also parsed out and analyzed separately. The archaeal-related sequences were mainly associated with ANME-1 (in cold seep cores and in some deep layers of slope cores) and MG1. The DWH-contaminated sequences were most-closely clustered in the archaea unweighted UniFrac analysis. The hypersaline sequences were very closely associated with each other, but were also related to some of the Continental Slope samples. DWH-contaminated cores were also more closely associated in the unweighted Archaea

UniFrac analysis. The unweighted analysis of only the bacterial-related sequences also indicated a close clustering of hypersaline sequences, though less association was observed among DWH contaminated samples. The seep sites showed some close association, but there was also some variance, especially with the MUC 5, the multicore near the small gassy mud volcano. Weighted UniFrac diversity analysis (Figure 2.10) showed closer association of samples from the Orca Basin slope (red sediment), but overall, samples showed little association by site, with the exception of the seep samples.

The weighted analysis with only archaea-related sequences showed that hypersaline sequences were very similar, and in a tight cluster with some of the slope samples, including some of the putative DWH-contaminated samples. This indicates that the hypersaline sequences, mainly associated with Marine Group 1, have similar sequence identity and relative abundance to Continental Slope and DWH-contaminated samples. The seep samples also cluster tightly in this analysis. It should be noted that MUC 4, with very low percentage of archaeal amplicon reads, was not included in the analysis. The rest of the archaeal-related sequences from seep sites were dominated by ANME-1. The weighted analysis of the bacteria-related sequences indicated some similarity in hypersaline samples, though the deepest samples from each basin did show some taxonomic abundance differences. Sequences were otherwise not clustered by environment type, though again, some seep sequences did form a tight cluster, indicating similar sequences and abundances at these sites. Overall, the diversity analysis indicates that sediment samples from each site were diverse, and vary downcore. Unweighted UniFrac analysis indicated similar bacterial groups inhabit the hypersaline sediment samples, though the weighted analysis indicated that the taxonomic abundances vary downcore and between samples. Deepwater

Horizon-contaminated sections did not cluster into similar groups, though may have similar MG-1 associated Archaea.

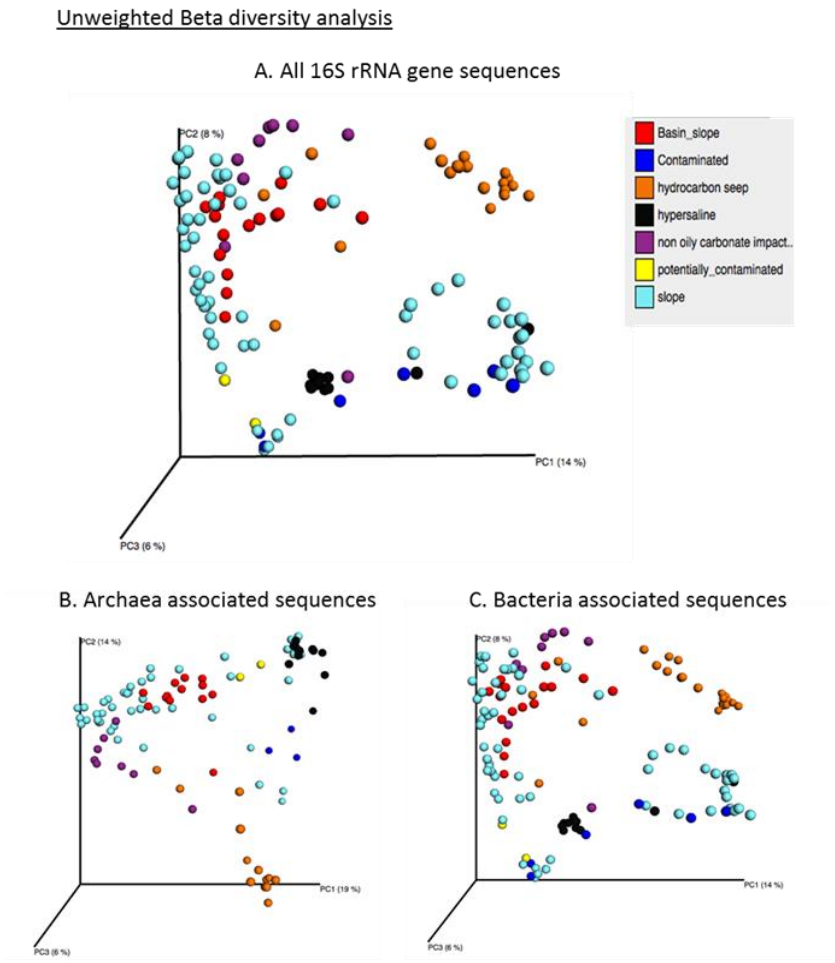



**Figure 2.9 Weighted UniFrac beta diversity analysis.** Samples were categorized as hydrocarbon seep (orange), hypersaline (black), Orca Basin slope (MUC 8) DWH contaminated (MUC 17, 18, 19 sections with visible oil, dark blue), DWH potentially contaminated (MUC 14, 20, 21, 22, yellow), and carbonate-rich slope (purple), and continental slope (light blue).

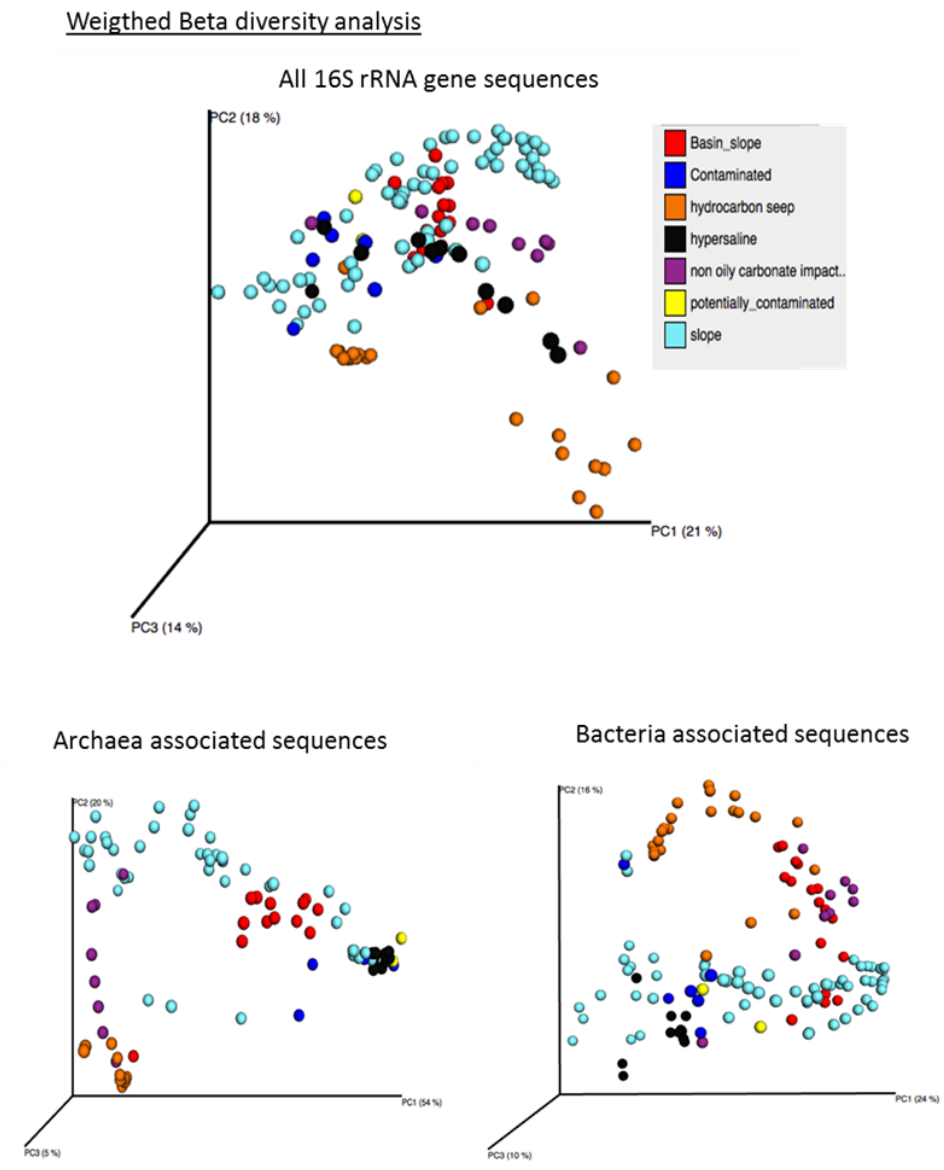



**Figure 2.10 Unweighted UniFrac beta diversity analysis.** Samples were categorized as hydrocarbon seep (orange), hypersaline (black), Orca Basin slope (MUC 8, red) DWH contaminated (MUC 17, 18, 19 sections with visible oil, dark blue), DWH potentially contaminated (MUC 14, 20, 21, 22, yellow), and carbonate-rich slope (purple), and continental slope (light blue).

## 2.5 Conclusions

Over 5 million sequences and 65,000 operational taxonomic units (OTUs) of Gulf of Mexico sediments were analyzed from cold seep sites, a hypersaline basin, Deep Water Horizon-contaminated surface sediments, and many sediments from the Continental slope down to 60 cmbsf. Taxonomic abundance diversity estimated by amplicon abundance indicated some similarities between sites, but also many differences. Seep sites (MUC 4, 5 and 12) varied in their distribution of ANME-1 sequences, and MUC 4 had <2% detected Methanomicrobia, potentially due to primer or sequencing bias. MUC 4 was also the only core to contain negligible Chloroflexi sequences, and contained sequences from the Gammaproteobacteria, Deltaproteobacteria and JS1 groups that had consistent patterns of correlation downcore. MUC 8, the multicore taken on the red-sedimented slope of Orca Basin, contained sequences that were similar to slope sediments and seep sediments, with JS1 being abundant throughout the core. Sequences from the hypersaline Orca Basin had taxonomic similarities to each other, and differed from the other sediment profiles. Furthermore, sediments with visible oil material contamination showed diversity similar to continental slope cores. Though DWH-contaminated samples did have some sequences that were indicative of oil degraders (Altermonadaceae, Colwelliaceae, oil degrading Rhodobacteraceae), they occurred at much lower abundances than sequencing surveys taken one month prior. This indicates fairly rapid consumption of oil compounds followed by slow consumption of 'recalcitrant' material and diversification of microbial communities as other carbon and energy sources become more favorable. The diversity analyses indicated a complex system with similarities in taxa in some sample types (hypersaline environment, seep site), but with differences in sequence abundance. Understanding

these trends will require further methodological (assessment of method bias), biogeochemical, and microbiological exploration.

# CHAPTER 3: MICROBIAL DIVERSITY OF HYPERSALINE SEDIMENTS ASSOCIATED WITH ORCA BASIN

## 3.1 Introduction

Orca Basin is a seafloor anoxic brine basin located on the Texas-Louisiana continental slope (26°56'N, 91°19'W). The basin has two deep brine pools located in the North and South regions of the basin at approximately 2400 m depth that are separated by a saddle region (Shokes et al. 1977) (Appendix A, Figure S1). Unlike hypersaline basins in the Red Sea and Mediterranean Sea that are affected by plate tectonic activity, Orca Basin formed as a result of salt tectonics, or movement of buried salt layers due to differential loading of sediments. The strong density gradient across the brine/seawater interface prevents advective mixing, keeping the brine and the sediments within the basin anoxic. The brine-seawater interface also serves as a particle trap, with total suspended matter as high as 880 µg/L, up to 8 times higher than in the overlying water (Trefry et al., 1984). In comparison to seawater, the brine is slightly acidic (pH = 6.5), enriched by factor 7.5 and 8.5 in $Na^+$ and $Cl^-$, respectively, moderately enriched by factors 1.3 to 2.6 in $K^+$, $Ca^{2+}$, $SO_4^{2-}$ and slightly depleted in $Mg^{2+}$ (Sheu, 1990). This composition is consistent with an origin from halite-rich evaporate salt layers within the Louann salt formation that resulted from evaporitic conditions during the Jurassic (Sheu, 1990). The brine is also enriched in biogenic compounds such as silica, phosphate, ammonia and methane introduced by sedimentation from the upper water column, by compaction-induced seepage, or potentially produced by the in-situ microbial communities (Shokes et al. 1977; Wiesenburg et al. 1985; Van Cappellen et al., 1998; Dickins and Van Vleet, 1992). Since the extreme salinity attenuates

microbial consumption, these compounds accumulate in brine lakes in very high concentrations that are sufficient to generate a diffusive flux across the density barrier into the water column (Wankel et al., 2010). Nitrate and oxygen in the overlying seawater are consumed near the halocline of Orca Basin and remain below detection limit in the brine, indicating that they are microbially consumed near the brine/seawater interface (Dickins and Van Fleet 1992; Shokes et al. 1977).

The hypersaline sediments of Orca Basin are dark in color, ranging from jet black to gray (Figure 3.2a). Porewater salinities were observed to reflect the overlying brine, averaging 260 ppt for 100 cm downcore, and then gradually decreasing (reviewed in Sheu, 1990). Buried fronds of *Sargassum* have been found intact with only selective degradation (Harvey et al., 1992) indicating attenuated microbial activity. The top 750 cm of sediment beneath the South basin brine pool contained predominately marine-derived amorphous organic matter that may be degraded due to long residence times and high microbial activity at the brine/seawater interface (Tribovillard et al., 2008, 2009). Reduced sulfur analysis suggested that black sediments contained higher concentrations of FeS, while the gray sediments were mainly composed of pyrite (Sheu and Presley, 1986). Sediment on the slope of the basin has been described as brown with red laminations (Sheu, 1990). Red sediment was determined both by X-ray mineralology and chemical analysis to have high concentrations of iron oxide in the form of hematite (Sheu and Presley, 1986). Pink sediments were also detected on the slopes of the basin with abundant sponges observed at one site (Figure 3.1b, this study).

Previous porewater geochemical measurements of the hypersaline sediment revealed high carbon and sulfur content. Sulfate was measured to be ~44-55 mM at the sediment surface and

decreased slowly, with sulfide detected in some hypersaline cores (up to ~150 µM) (Sheu, 1990), despite no measurable sulfide in the deep brine (Wiesenberg et al., 1985).


**Figure 3.1 Sediment characteristics. A.** Photographs of multicores taken from (a) South basin brine pool (MUC 6), (b) North basin brine pool (MUC 7), and (c) red sediments of the basin slope (MUC 8). **B:** Site where push cores were taken with the *DSV Alvin* at 2198 m (left) and near sponges at 2167 m (right).

In contrast to microbial community structure analyses of deep hypersaline anoxic basins (DHABs) in the Mediterranean and Red Sea (Antunes et al., 2011), microbial community surveys of Orca Basin based on 16S rRNA gene or functional gene sequencing are currently missing. So far, studies in Orca Basin have focused on microbial activity at the halocline. Measurements of adenosine 5´-triphosphate and uridine uptake indicated a very active microbial community within an approx. 50 m layer (2173 to 2221 m) where oxygen became depleted from approx. 71 to 13% of deep seawater concentrations, while microbial activity in the deep brine decreased substantially (LaRock et al., 1979). Archaeal isoprenoid lipids showed a distinct peak that coincided with the halocline depth interval at ca. 2260 to 2270 m depth, indicating that

archaea thrive or at least accumulate in this interface horizon (Dickins and Van Fleet 1992). Subsequent work showed abundance peaks of cultured iron- and manganese-oxidizing and reducing bacteria near 2170 and 2220 m depth, respectively, above the steepest halocline layer between 2230 to 2250 m (Van Cappellen et al., 1998). As a first step towards a comprehensive microbial census of Orca Basin, we have analyzed the bacterial and archaeal community composition of Orca Basin sediments by 16S rRNA gene sequencing on both Sanger and Miseq platforms from extracted DNA of geochemically characterized sediment cores underlying the deep brine pools and the slope of the basin. In addition, we analyzed 16S rRNA gene sequences from amplified DNA of single cells from the brine pool obtained by fluorescence activated cell sorting.

## 3.2 Materials and Methods

### 3.2.1 Sample collection and processing

Sediment core samples were collected on the *R/V Atlantis* (expedition AT18-2) in November 2010. Sediment samples from the anoxic brine-filled southern and northern basins (MUC 6 and MUC 7) as well as the central ridge area of Orca Basin (MUC 8) were collected by a shipboard multicorer; maximal core length was 70 cm. Sediment from the steep flanks of Orca Basin near the redoxcline and halocline was also collected with push cores on the *Alvin* during dive 4650 (cores 4650-24 and 4650-4) (Appendix A, Figure S1). For molecular analysis, brine from the South basin brine pool was collected with Niskin bottles attached to a rosette sampler equipped with a CTD on the *R/V Atlantis*. Brine for single cell sorting was collected by the same method on a subsequent cruise on the *R/V Pelican* (expedition 12-22) in April 2012.

Sediment cores were immediately capped and stored at 4°C until they were processed within 12 hours of collection. Overlying water was removed with a sterile syringe. Multicore samples

were divided into 5 cm sections, while push core samples were sectioned in 1 to 5 cm intervals. All core extruding materials were either autoclaved or treated with ethanol prior to sampling. Water for molecular analysis was transferred from Niskin bottles into sterile containers, and approximately 500 mL was vacuum-filtered onboard using a sterile 47 mm diameter polycarbonate filter with a 0.22 µm pore size. All samples for molecular analysis were immediately frozen at -80°C after processing. Brine collected for single cell sorting was transferred to an argon-flushed sterile canning jar and aliquoted into a $N_2$ flushed stoppered serum vial before sending to the Bigelow Single Cell Genomics Center (East Boothbay, Maine) for cell sorting.

### 3.2.2 Geochemical analysis preparation

Geochemistry subsamples were taken from each Alvin core and multicore during core slicing. For measurement of methane concentration and $\delta^{13}C$-signatures, 5 ml syringe samples of sediment were added to 2 ml of 1.0 M NaOH solution in 30 ml serum vials, and capped with butyl rubber stoppers and crimped shut.

For multicore sulfate ($SO_4^{2-}$) and sulfide ($H_2S$) measurements, porewater samples were obtained by filling a 50 ml falcon tube with sediment and centrifuging it at 3000 rpm for 15 minutes. Porewater samples were filtered through a 0.22 µm syringe filter into sterile falcon tubes. For hydrogen ($H_2$) concentration determination, 3 mL of sediment was sampled with a cut-off syringe and transferred to He-purged vials, capped with a rubber septum, crimped and re-purged with He. The vials were then stored for headspace equilibration at 4°C for four days (Hoehler et al., 1994) until analysis. Samples for sulfide concentration measurements were processed by aliquoting 1 ml of porewater to 0.1 ml of 0.1M zinc acetate in a 2 mL microcentrifuge vial, and stored subsequently at 4°C. For samples prepared for sulfate

concentration determinations, 1.0 ml of porewater was added to 50 µl of 50% HCl in a 2 ml microcentrifuge tube, bubbled with $N_2$ for three minutes to remove sulfide, and stored at 4°C until analysis by ion chromatography. The remaining porewater was stored at -20$^{o}$C for salinity measurements.

For all other multicore porewater analyses and for all *Alvin* pushcore gas and porewater analyses, cores were processed following Joye et al. (2004). Cores were extruded and sliced in 2.5 to 5 cm sections. Methane samples were processed by adding 5mL of each sediment section extracted by a 30 mL cut-off syringe to a He-flushed serum vial amended with 4 mL of 2M NaOH, capped with a butyl rubber septa, crimped, shaken and stored upside down at room temperature. Remaining sediment was then transferred to argon-flushed squeezer cups, collected with an argon-purged syringe, and passed through 0.22 µm Target filters before being aliquoted for pore water analyses. For subsamples to be processed for cation ($Na^+$, $K^+$, $Mg^{2+}$) and anion ($Cl^-$, $SO_4^{2-}$) concentration analysis, porewater was preserved with 1 µM final concentration of nitric acid until analysis. For subsamples preserved for sulfide concentration analysis, 0.5 to 2 mL of porewater was transferred to a vial containing 1 mL 20% zinc acetate. The remaining filtered porewater was either analyzed immediately for ammonium ($NH_4^+$) concentration or frozen until processed for dissolved organic carbon (DOC) and salinity.

### 3.2.3 Geochemical measurements

Porewater and brine salinity was determined with an Atago Master-S / Millα salinity refractometer. Samples were filtered through a 0.45µm syringe filter and diluted by weight with deionized water. Measurements had an accuracy of 2 ppt. Methane concentrations from multicore samples were analyzed by injecting 1 mL of headspace gas into a SRI 8610C GC (SRI Instruments, USA) equipped with a flame ionization detector. Concentrations were determined

by comparing to a certified standard with a 1% mixture of methane, ethane, n-propane, iso-butane, n-butane, and n-pentane in helium. The precision was calculated to be 1.2% by repeated measurements. Methane samples from *Alvin* push core sediments were measured following Joye *et al.* (2004) with a Shimadzu gas chromatograph. Samples were measured on the ship with a Peak Performer 1 Reduced Gas Analyzer equipped with mercury oxide detector (detection limit 800 parts per trillion; Peak Laboratories, USA). Calibrations were performed with a 1% $H_2$ standard balanced in helium (Scott Specialty Gases, USA). Repeated measurements were performed with a standard to determine precision (>1.5%).

Major cations and anions, NOx, and DOC were measured as previously described in Joye et al. (2004) with some modification for high salt concentration samples. $SO_4^{2-}$, $Cl^-$, sulfide, and $NH_4^+$ concentrations were measured colorimetrically using Cline's reagent (Cline, 1969) for sulfide and the indo-phenol method for $NH_4^+$ (Solórzano, 1969). NOx concentrations were determined colorimetrically using the Cd-reduction followed by diazotization method with an Alpkem 2- channel autoanalyzer. Precision of these analyses was ~2%.

### 3.2.4 DNA extraction and sequencing

For Sanger sequencing, DNA from sediment was extracted with the PowerSoil DNA isolation kit according to the manufacturer's instructions (MoBio, Carlsbad, CA) except for the following modifications: 0.5 g of sediment was centrifuged at maximum speed (16000 x g) for 5 minutes and overlying water decanted prior to processing. After bead beating, the supernatants from 3 tubes were combined into one column, and the final eluate was reduced to 50 µL. For the brine sample, DNA was extracted from the filter with the PowerSoil DNA isolation kit according to the manufacturer's recommendations except the final volume was adjusted to 50 µL.

The bacterial and archaeal 16S rDNA genes were amplified with Qiagen's HiFidelity polymerase according to the manufacturer's recommendations for buffers, polymerase, and Q solution. In addition, $MgCl_2$ was used to a final concentration of 2mM. Primers B8f and B1492R for bacterial 16S rRNA genes (Caporaso et al., 2012) (Teske et al. 2002) and A8f and A1492 for archaeal 16S rRNA genes (Teske et al. 2002) were used at a final concentration of 2 µM each. Amplification was performed in a thermal cycler under the following conditions: an initial denaturation 95°C for 5 min, 35 cycles of 95°C for 30 s, 52°C for 1 min and 72°C for 2 minutes, and a final extension at 72°C for 10 minutes. PCR products were electrophoresed on a 1% agarose gel amended with GelRed (Biotum, Hayward, CA) and visualized under UV light. Positive PCR products were purified with the minElute PCR Purification Kit (Qiagen) or, if multiple banding was observed, the MinElute Gel Extraction kit (Qiagen). Bacterial 16S rRNA gene PCR products were ligated to the TOPO TA PCR 2.1 vector (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions, with the exception that reactions were incubated for 16 hours. Archaeal 16S rRNA gene PCR products were ligated to the pDrive cloning vector (Qiagen) according to the manufacturer's recommendations at a ligation temperature of 8°C. All vector reactions were transformed into *Escherichia coli* TOP10 chemically competent cells (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Transformed bacteria were plated on LB agar with 50 µg/mL kanamyacin and 40 µg/mL X-gal. White colonies were arbitrarily selected, picked and re-streaked on another LB agar plate for bidirectional Sanger sequencing using vector primers M13 F and M13 R (Genewiz, South Plainfield, NJ).

For Illumina sequencing, samples were thawed on ice and 0.5 g of sediment was extracted with a 1 mL syringe and placed in a 96 well MoBio bead beating plate for subsequent extraction with the Powersoil-htp 96 Well DNA Isolation Kit (MoBio). DNA extraction, 16S rRNA gene amplification of the V4 fragment, and sequencing on the Miseq2000 platform was performed at Joint Genome Institute (Walnut Creek, CA) according to the Earth Microbiome Project standard protocol (Caporaso et al., 2012).

For single cell analysis, Brine was sent to the Bigelow Single Cell Genomics Center (East Boothbay, Maine) for sorting, followed by multiple displacement amplification of cellular DNA (Stepanauskas 2012) and PCR screening for bacterial and archaeal 16S rRNA genes as described in detail in the methods file SCGC_Services_Description.pdf, available for download at scgc.bigelow.org/services/#br=services_description_table).

### 3.2.5 Sequencing analysis

Sanger sequence reads were assembled and edited in Sequencher (Genecodes, Ann Arbor, MI). Sequences were aligned with SINA 1.2.11 (Pruesse et al., 2012) and were imported into ARB (Ludwig et al., 2004) and manually corrected. Neighbor-joining trees were constructed in MEGA6 with a Jukes Cantor correction (Jukes and Cantor, 1969). Statistical support of nodes was assessed by 1000 bootstrap replicates. Illumina sequences were analyzed with QIIME, v1.8. Sequence quality scores were plotted and sequences truncated at a cutoff of quality score of 25. Paired ends were joined with the fastq-join method (Aronesty, 2011) and sequence data demultiplexed and identified by the 12 base golay error correcting barcodes. Sequences were grouped into operational taxonomic units (OTUs) and putative chimeras removed with UPARSE (Edgar, 2013). Singletons were removed. Taxonomy was assigned to OTU sequences using the SILVA taxonomy classifier (Pruesse et al., 2007). Sequences were aligned with the python

implemented NAST alignment algorithm (Caporaso et al., 2010b). Phylogenetic trees relating OTUs were constructed using the FastTree method (Price et al., 2009). Alpha diversity was calculated using Shannon diversity (Shannon, 1948) and Faith's PD whole tree (Faith, 2006) metrics based on multiple rarefactions of the data. Box plots were created for each designated group and significant differences estimated with a pairwise t-test using nonparametric Monte Carlo permutations to determine p-values. Beta diversity was estimated using UniFrac weighted and unweighted metrics (Hamady et al., 2009) and visualized with Principal Coordinate Analysis (PCoA).

## 3.3 Results

### 3.3.1 Site description and physical characteristics of sediment

The sediment core from the South basin (MUC 6) was black throughout, and the North basin multicore (MUC 7) was similarly black with a transition to gray starting at 45 cm. The multicore of the basin slope (MUC 8) was red with brownish-red laminations throughout. The multicorer overpenetrated the extremely soft sediments, and no overlying water was recovered. Therefore, the exact depths are unknown and the true sediment surface was not represented in these cores. The *Alvin* push core samples were pink with reddish-brown laminations. Unlike the multicore samples, they were not over-penetrated and contained overlying water (Figure 3.1).

### 3.3.2 Salinity and major anions and cations

Refractometry measurements indicated that the sediment beneath the basin's brine pools reflected the salinity of the overlying brine. The salinity of the deep brine was measured to be 258 ± 6‰. South basin sediment porewater salinities decreased from 267‰ at the sediment surface to 261‰ at 60 cm. The North basin sediment porewater salinities decreased from 255‰

at the top of the core to 234‰ at the bottom of the core. The Orca ridge multicore porewater salinity ranged between 53 and 59‰. Porewater salinity was not measured in the Alvin push cores.

Porewater measurements of major anions and cations indicated little concentration change with depth (Figure 3.2). In the North Basin core, porewater $SO_4$ concentrations decreased slightly with depth, from near 30 mM in the 0-5 cm section to 23 mM at 45-50 cm depth. In the South Basin core, porewater $SO_4$ concentrations were elevated, starting with 43 mM at 0-5cm and decreasing only slightly to 39 mM at 55-60 cm depth. $Na^+$ and $Cl^-$ concentrations were elevated in the sediments under the brine, with each ~5 M in the South basin sediment and ~4.5 M in the North basin sediment, ~8x the concentrations of the seawater overlying the basin. In comparison to the background site core, $Na^+$ and $Cl^-$ were slightly elevated (~1.5x) in the ridge multicore (MUC 8) and the Pink Jello site core, while the Sponge Garden core had concentrations similar to seawater. Cores from the North and South basin also had $K^+$ and $Ca^{2+}$ concentrations that were slightly (~2x) more elevated than seawater concentrations (Figure 3.2).

**3.3.3 Dissolved gases, nitrogen, carbon and sulfide.**

Porewater $CH_4$ concentrations were < 2 µM in the basin slope push cores and in the MUC 13 background sediment core. The basin slope multicore sample MUC-8 contained methane in concentrations between 3 and 12 µM, while the methane concentrations in the brine-impacted MUC 6 (South Basin) and MUC 7 (North Basin) sediments were much higher, up to 156 µM in the South Basin core and 1400 µM in the North Basin core. The $\delta^{13}C$-$CH_4$ values in the South Basin core (-75.5 to -76.7 ‰) and the North Basin core (-89.6 to -83.1‰) indicated a biogenic source. Ammonium concentrations were high in the brine-impacted sediments, reaching 480 µM in the South basin sediment core and 323 µM in the North basin sediment core (Figure

3.2). The ridge push cores had $NH_4$ concentrations that ranged from 12 to 70 µM, while the basin slope multicore concentrations ranged from 66 µM at the surface to ~300 µM within 15 to 30 cm downcore. Combined nitrite and nitrate (NOx) concentrations were 0.81 µM in the North Basin core, and 2 and 8 µM in the South Basin core. Dissolved organic carbon (DOC) was highly elevated in the sediments beneath the brine (up to 5862 µM in the North Basin sediment and 2201 µM in the South Basin sediment), and ranged from 300-1000 µM in the basin slope core (Figure 2). Sulfide was below detection level in all cores except the Orca Basin North brine pool sediments, where sulfide concentrations up to 64 µM were measured.

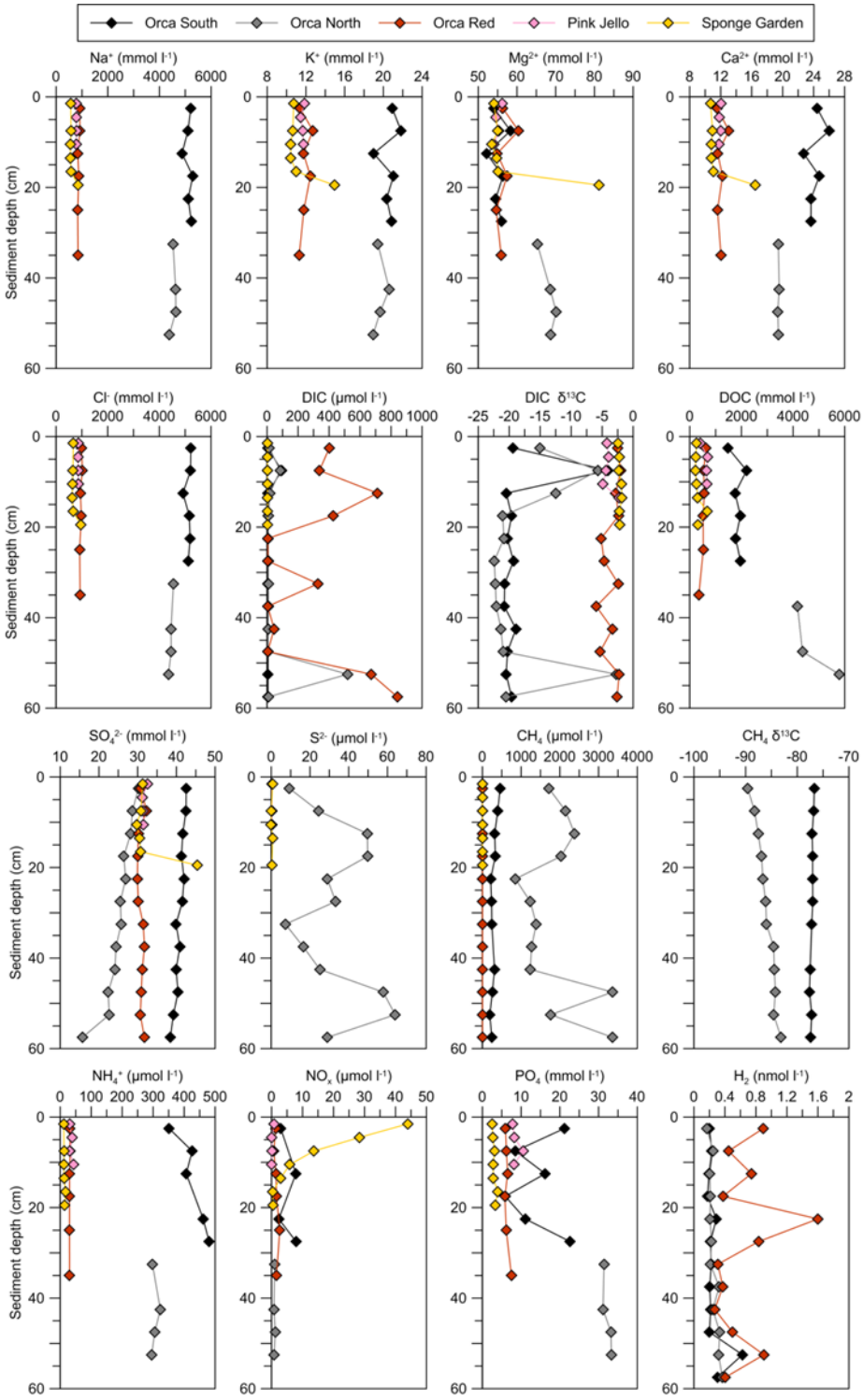



**Figure 3.2 Geochemical Measurements.**

### 3.3.4 Sanger bacterial 16S rRNA gene analysis

The bacterial 16S rRNA gene library from the deep anoxic brine and brine sediments were dominated by sequences related to Deltaproteobacteria, Candidate Division KB1 and Bacteroidetes (Figure 3.3 and 3.4). The dominant phylotype in the deep brine and in the underlying sediments was affiliated with the uncultured Candidate Division KB1. This group was originally detected in the hypersaline Red Sea basin Kebrit Deep (Eder et al., 1999) by 16S rRNA gene sequencing and has subsequently been found predominately in high salinity environments (e.g. Antunes et al., 2011). Sequences affiliated with Deltaproteobacteria and Bacteroidetes were also abundant in the hypersaline clone libraries. Bacteroidetes sequences had the strongest similarity to sequences from a hypersaline microbial mat in the Candeleria lagoon of Puerto Rico (Isenbarger et al., 2007) and an evaporitic crust community of Guerrero Negro (Mexico) (Kirk Harris et al., 2012).

Deltaproteobacterial sequences were also related to clones from hypersaline sites, including a Mediterranean solar saltern (NCBI accession FJ536432) and the brine-seawater interface of the hypersaline basin Shaban Deep in the Red Sea (Eder et al., 2002). The sulfate-reducing bacteria that have been detected by 16S rRNA gene sequencing in the brine and sediments (Appendix A, Figure S3) are in part members of halophilic and/or halotolerant lineages and genera (*Desulfocella halophila;* Brandt et al., 1999), or their closest known sister lineages are halophilic (*Desulfohalobium retbaense*, Ollivier et al., 1991; *Desulfovermiculus halophilus,* Belyakova et al., 2006; *Desulfohalobium utahense,* Jakobsen, 2006). The clone libraries also contain phylotypes, such as Cyanobacteria, that likely represent cells from sinking particles originating in the upper water column that have been preserved in the brine (Boere et al., 2011).

The clone library sequences from the non-hypersaline slope of the basin represented diverse bacterial groups. In the red slope multicore sample (MUC 8), dominant bacterial-associated sequences included the uncultured candidate divisions OP8 and JS1, as well as Chloroflexi, Fusobacteria, Planctomycetes, and Deltaproteobacteria (Figure 3.3). The Pink Jello sediment sample had no detectable sequences related to either JS1 or OP8, indicating some taxonomic differences in the red multicore and pink *Alvin* core sediments of the basin slope.

Planctomycetes were present in all sections of the basin slope cores, and Fusobacteria were present in all slope samples except the Sponge Garden core 0-1 cm section. The Pink Jello and Sponge Garden cores contained similar phyla, including *Nitrospira* sp., Candidate Division SR1, as well as Cyanobacteria, and Alpha-, Delta- and Gammaproteobacteria (Figure 3.3). The Deltaproteobacteria were from diverse taxonomic families in the Pink Sediment sample core, including the Desulfbacteraceae, Desulfurellaceae and Desulfobulbaceae, as well as unclassified groups, and were most closely related to uncultured organisms (Appendix, Figure S2). Most Sponge Garden site clone sequences were affiliated with unclassified groups within the Deltaproteobacteria. A single clone was closely related (95% similar) to *Nitrospina gracilis*, a nitrite-oxidizing bacterium (Watson and Mandel, 1971). While several sequences from the hypersaline sediments and brine fell within the same OTU (Figure 3.4), no shared OTUs were observed between sequences from hypersaline and non-hypersaline samples.

### 3.3.5 Sanger archaeal 16S rRNA gene sequences

Primers targeting the archaeal 16S rRNA gene yielded fewer distinct archaeal lineages than did bacterial primers (Figure 3.3). The hypersaline brine and sediment clone libraries were mainly dominated by members of the Marine Group I clade (MG1),whose cultured representatives have thus far been ammonia oxidizers (Pester et al., 2011), and by members of

the genus *Methanohalophilus*. Cultured *Methanohalophilus* species are methylotrophic and halotolerant to halophilic methanogens that have been isolated previously from hypersaline environments, for example sediments of the Great Salt Lake in Utah (Paterek and Smith, 1985), Wadi Natrun in Egypt (Mathrani et al., 1988) and a saline aquifer in the Oregon desert (Liu et al., 1990) (Appendix A, Figure S2). In addition, the South Orca Basin 55-60 cm clone library contained sequences representing putative methane-oxidizing ANME-1a, ANME-1b and ANME-2c (Knittel and Boetius, 2009) as well as Marine Group II and Marine Group III archaea. In the North Basin 0-5 cm section, one sequence associated with ANME-2a as well as several sequences associated with Marine Group II were detected. Sequences from the Orca Basin slope sediment mainly contained MG1 sequences in addition to sequences related to the Deep-Sea Hydrothermal Vent Group 5 (DHVEG-5), and unclassified Thermoplasmatales (Figure 3 and S3). The red multicore samples (MUC 8) yielded few archaeal clones (N=5). These were related to unclassified uncultured Methanomicrobia, Thermoplasmatales and Marine Group I archaea (Appendix A, Figure S3). Most sequences obtained with archaeal 16S rRNA gene primers were distantly related (<70% similarity) to uncultured eukaryotic sediment organisms (results not shown).

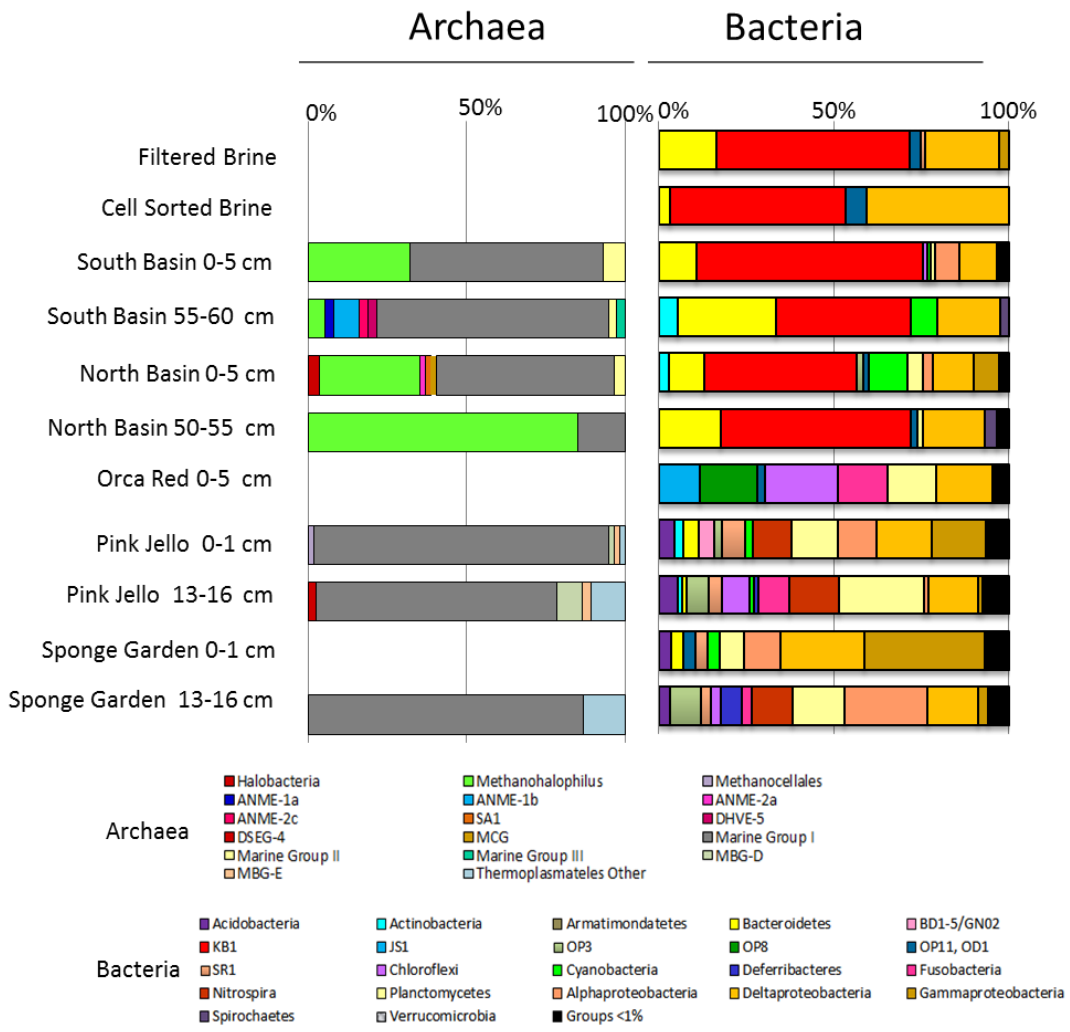



**Figure 3.3 Relative abundance of Sanger sequencing associated taxa.**

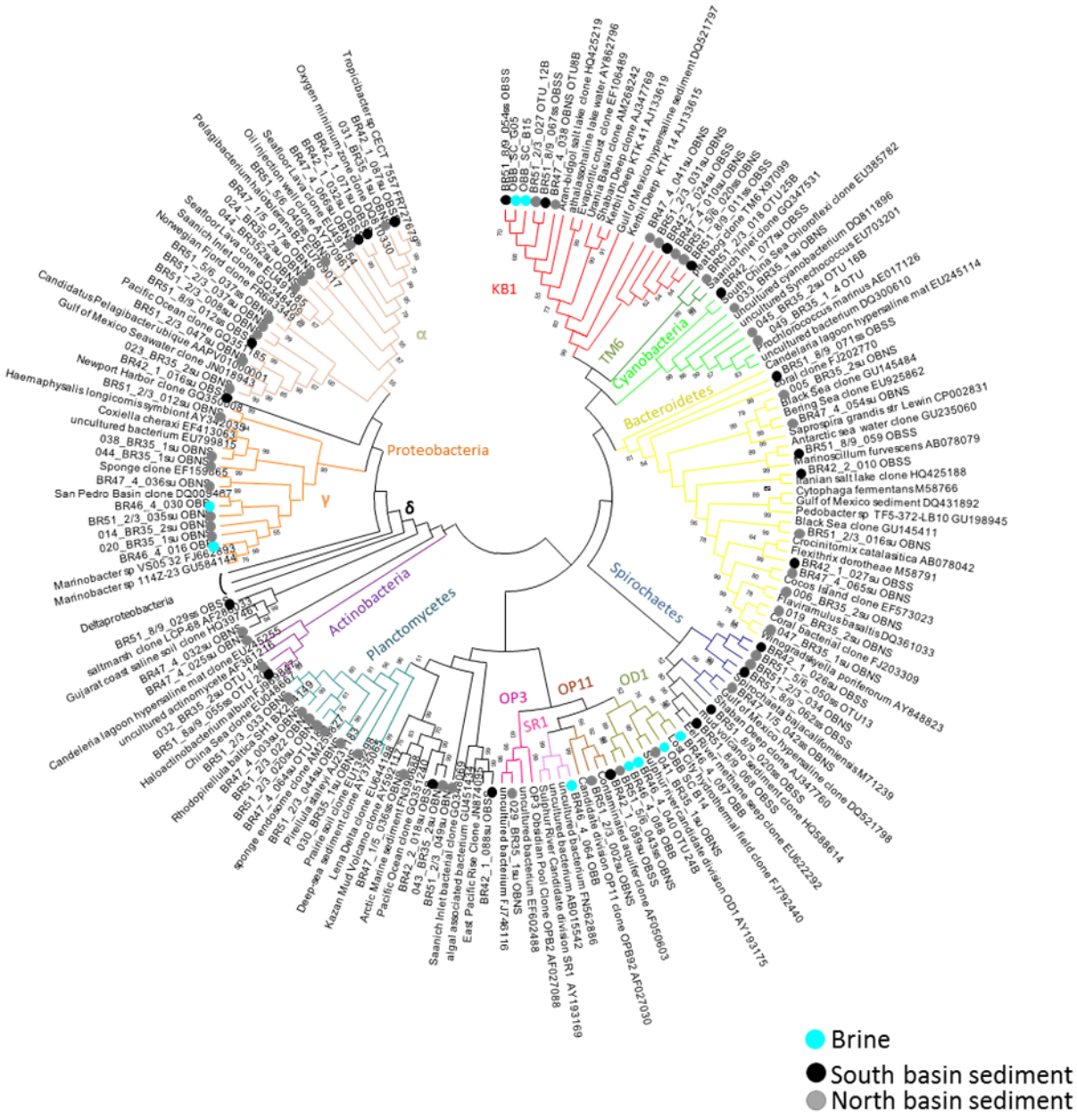



**Figure 3.4 NJ tree of Orca Basin brine and hypersaline sediment** 16S rRNA gene clone sequences. Bootstrap values (1000 replicates) are displayed next to nodes. Only bootstrap values of >50% shown. Deltaproteobacteria sequences are displayed in Appendix A.

### 3.3.6 Illumina sequencing analysis

A total of 1,974,676 reads were obtained after demultiplexing and quality control filtering (read length 253 bases). Several samples failed to produce sequences. After chimera detection and USEARCH OTU clustering, a total of 65,415 OTUs were obtained, of which 29,927 were singletons and removed from downstream analysis. Several South Basin hypersaline sediment sections (10-15 cm and 15-20 cm) had some putative anthropogenic sequences of *Corynebacterium*, *Staphylococcus*, and *Streptococcus* spp. based on BLAST (Altschul et al., 1990) searches; these likely contaminants were removed for downstream analyses, but are retained in the raw sequencing uploaded files in VAMPs and the SRA.

The hypersaline sediment of the North and South brine pools contained archaeal sequences associated with similar phyla. Sequences associated with MG1 were the dominant archaeal-related sequences, comprising up to 25% of the total reads per sample. Members of the Euryarchaeota-associated DHVE-5 and DSEG groups were also indicated, as well as members of the Thermoplasmata. Sequences related to Methanomicrobia had similar taxonomic affiliation as those observed in the Sanger clone libraries; they accounted for <0.5% of all sequences. Sequences related to *Methanohalophilus* sp. were detected in both the North and South basin, and anaerobic methane-oxidizing archaea (ANME-1a, ANME-1b, ANME 2) were also observed in the sequence reads (Supplemental File 1). Also similar to the clone library data, sequences related to *Bacteroidetes* and *Deltaproteobacteria* were among the most numerous bacterial groups. The *Deltaproteobacteria*-related sequences were most similar to sequences detected in other hypersaline systems, including members of the genus *Desulfovermiculus* within the family *Desulfohalobiaceae* (Belyakova et al., 2006) and the Mediterranean Sea Brine Group 7 (MSBL-7) (Borin et al., 2009). Cyanobacteria-related sequences were also detected in all the core

sections, suggesting the presence of preserved cells or DNA (Boere et al., 2011). Sequences related to the Candidate Division KB1 were present, but in low abundance (<1% in most core sections, Figure 3.5). Archaeal sequences in the basin slope sediment (MUC 8) were also dominated by MG1, with the exception of the 50-55 cm sample that contained a higher abundance of Methanomicrobia (ANME-1) (Supplemental File 1, Figure 2.5).

The red slope multicore sample was dominated by the Candidate Division JS1 (Webster et al., 2004). Members of the Chloroflexi and Nitrospirae phyla as well as Deltaproteobacteria and Gammaproteobacteria were also present. The Pink Jello *Alvin* core sample yielded fewer JS1-related sequences in comparison with the red multicore, MUC 8. Taxa identified as BH180-139 (also known as CD12) and members of the OP8, as well as Deltaproteobacteria and Gammaproteobacteria were also common. Marine Group I was the dominant archaeal group, except in the 50-55 cm section which was dominated by ANME-1b related sequences (Figures 3.5 and Appendix A, Figure S7).

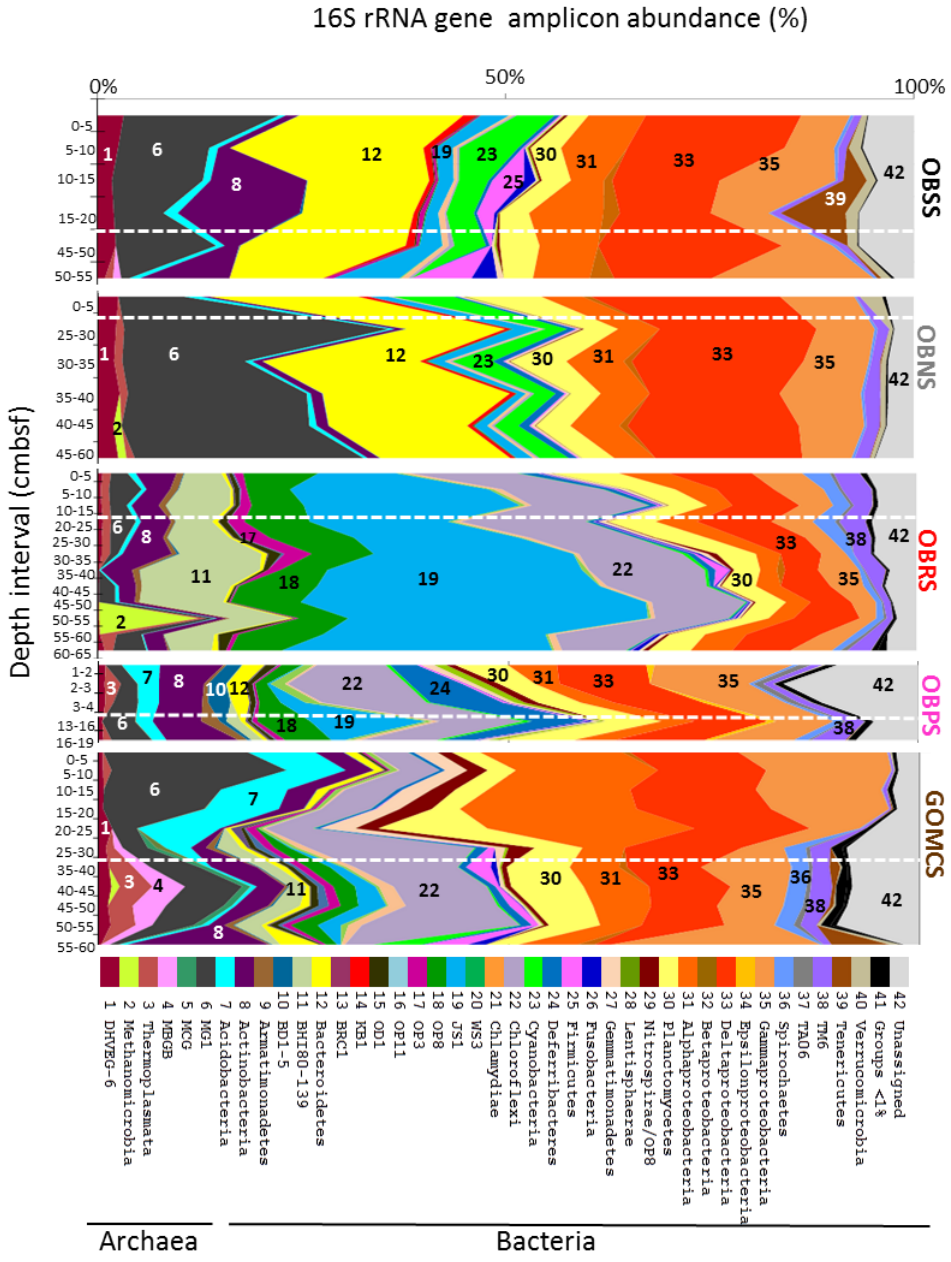



**Figure 3.5 Relative taxonomic composition of archaeal and bacterial 16S rRNA gene amplicons.** Taxonomic classifications were determined using QIIME's uclust method and the most recent available SILVA reference and taxonomy mapping files (SILVA 111 release), in addition to manual evaluation of the most abundant Candidate Division lineages (see materials and methods). Samples (top to bottom) include: OBSS, Orca Basin South brine pool sediment (MUC 6); OBNS, Orca Basin North brine pool sediment (MUC 7); OBRS, Orca Basin slope red sediment (MUC 8); OBPS, Orca Basin slope pink sediment (Alvin core 4650-24). GOMCS, Gulf of Mexico Continental Slope sediment (MUC 13). Broken white lines indicate a break in sequencing at continuous depths downcore.

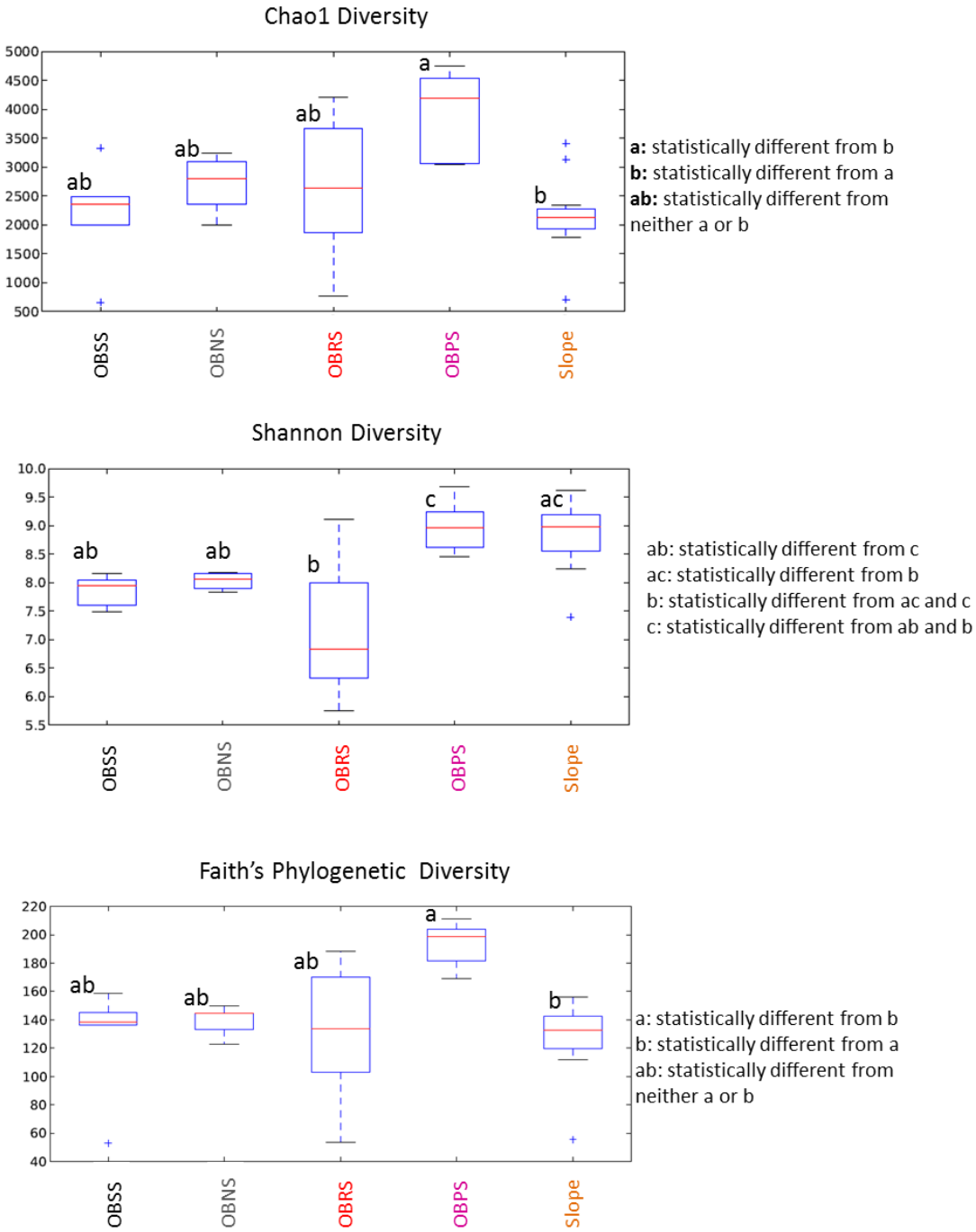



**Figure 3.6 Box plots of alpha diversity estimators.** Calculations are based on ten iterations at a sequence depth of 15000 per sample for each core. The box indicates the upper and lower quartiles, and the line in the middle of box represents the median. Lines above and below the box indicate 1.5 interquartile ranges, and the + symbol indicates outliers. Significant differences were determined by a non-parametric t-test using Monte Carlo permutations to calculate p-values in QIIME. Samples considered significantly different at $P < 0.05$. Orca Basin South Sediment (OBSS, N=5), Orca Basin North Sediment (OBNS, N=5), Orca Basin Red sediment (OBRS, N=12), Orca Basin Pink Jello core (N=5), Gulf of Mexico background sediment core (Slope, N=11).

### 3.3.7 Comparison between Sanger and Illumina sequences

The Sanger and Illumina sequencing surveys produced many similarities, but also some notable differences. Many of the same bacterial and archaeal phyla were observed in both surveys, with 62% (Pink Jello sample, 4650-24 1-2 cm) to 85% (Orca South sediment sample, 0-5 cm) of Sanger bacterial OTUs present in the Illumina sequences. The two methods indicated different relative taxon abundances, especially in the hypersaline samples. The Sanger clone libraries indicated that the bacterial communities in these sediments were dominated by sequences related to Candidate Division KB1, while sequences related to this group represented a small fraction (~1%) of the sediment community in the Miseq reads.

### 3.3.8 Diversity analyses

Chao 1, Shannon, and Faith's Phylogenetic Diversity analyses indicated that Orca Basin hypersaline samples were not significantly different than background slope sediment ($P>0.05$). However, both PD Whole Tree and Shannon diversity indices indicated that the Pink Jello slope sample contained higher species diversity than the Orca Basin hypersaline sediment samples (Figure 3.6). The Orca red slope sediment sample had a large range of species diversity downcore. Unweighted UniFrac beta-diversity analysis indicated that North and South hypersaline sediment sequences were not statistically different from each other except for the deepest sediment layer of the South Basin, but were statistically different from the Orca slope, sequences from a mud volcano (Hot Site), a hydrocarbon seep and mud volcano area (GB425), and background sediment in the deep Gulf of Mexico (MUC 13).

However, weighted Unifrac analysis indicated more differences within Orca Basin hypersaline samples. Analysis of only the archaeal-related sequences indicated that Orca Basin hypersaline archaea were similar to archaea from continental slope sediments, likely due to the

prevalence of Marine Group I archaea in both samples (Figure 3.6). Red Sea Atlantis II and Discovery Deep sequences (Siam et al., 2012) grouped separately from Orca Basin (Figure 3.7). The Red Sea archaea-related sequences were also reported to contain high numbers of Marine Group I organisms in most sediment sections, but also higher abundances of ANME-1 and Marine Benthic Group E and Terrestrial Hot Spring Group II Crenarchaeota (Thaumarchaeota; Pester et al., 2011) in the deepest layer of the Atlantis II Deep brine sediment. Similar to Orca Basin, the bacteria-related sequences of the Red Sea brine sediments almost always contained high percentages of Proteobacteria, but Alphaproteobacteria and Gammaproteobacteria tended to dominate the sequence abundances in the Red Sea (Siam et al., 2012), while Orca Basin Illumina amplicon reads contained higher percentages of Deltaproteobacteria-sequences (Figure 3.5). The deepest (high total sulfur content) layer of the Red Sea contained sequences that were indicative of OP1 (shown to be KB1), Fusobacteria, and Chloroflexi (Siam et al., 2012).

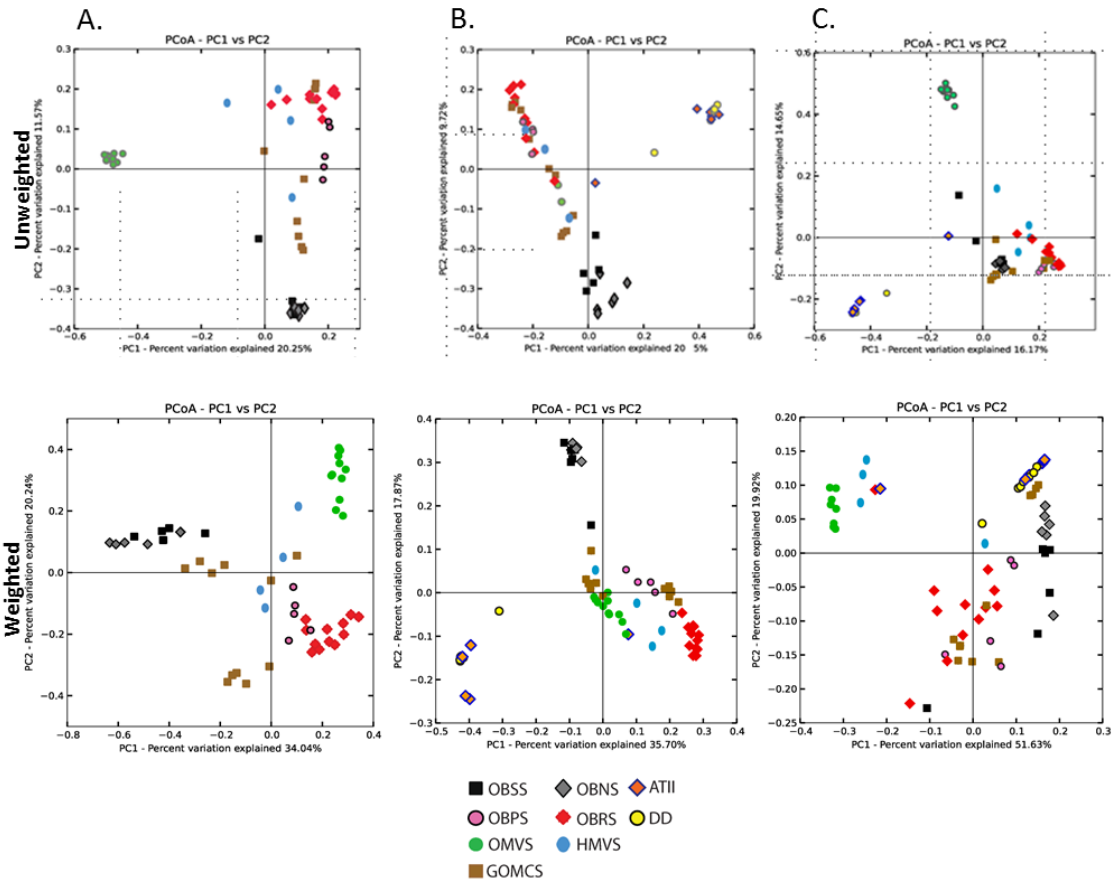



**Figure 3.7 Unweighted (top) and weighted (bottom) UniFrac PCoA analyses.** Analysis includes Orca Basin Illumina Miseq (this study) and Red Sea brine basin 454 sequences (Siam et al., 2012). A. All Gulf of Mexico (GOM) amplicon reads, B. Bacteria from GOM and Red Sea Atlantis II Deep (ATII), and Discovery Deep (DD). C. Archaeal-associated reads from GOM and Red Sea. GOM key descriptions: OBSS, Orca Basin South Sediment; OBNS Orca Basin North Sediment, OBPS, Orca Basin Pink Sediment; OBRS, Orca Basin Red Sediment; OMVS, Hydrocarbon and mud volcano sediment (MUC 4); HMVS, hypersaline mud volcano and seep site (MUC 5).

## 3.4 Discussion

### 3.4.1 Sulfur cycle dynamics in Orca Basin

The measured $SO_4^{2-}$ concentrations of sediments beneath the southern brine pool of Orca Basin were ~42 mM with slight depletion downcore (Figure 3.2), consistent with values from previous studies (Sheu 1990, Hurtgen, 1999). The Orca Basin North brine pool and red oxic ridge cores had lower sulfate concentrations (30 mM at surface), with the North basin core showing greater depletion with depth (Figure 3.2). Sulfide concentrations were below the

detection limit (<3 uM) in the South basin and Orca slope sediment cores, and were between 7-64 μM in the North Basin core. Potential sulfate reduction rates (SRR) in South basin sediments were previously observed to be relatively low (maximum 38 nmol/$cm^3$/day), but higher than in reddish-brown sediments from the non-hypersaline flanks of the basin (15 nmol/$cm^3$/day) (Hurtgen et al., 1999), consistent with evidence for gradual sulfate depletion, black sediment color due to iron monosulfide formation and low concentrations of porewater sulfide. Higher sulfide concentrations in the North basin core may indicate basin-specific differences in sulfate reduction rates or sulfide complexation with iron, and requires further study.

Despite potential differences in sulfur cycle dynamics between the North and South brine pool sediment cores, differences in abundance or diversity of sequences related to the Deltaproteobacteria, a group that comprises many of the known sulfate reducers, were not observed (Figure 3.5). Furthermore, taxonomic analysis of the Sanger 16S rRNA gene amplicons revealed Deltaproteobacteria-related sequences that were phylogenetically similar to cultured halotolerant sulfate reducers or uncultured organisms that are typically found in other hypersaline environments (e.g. members of the Desulfobulbaceae, Desulfbacteriaceae, and MSBL-7 classified by SILVA in Desulfobacteraceae; Appendix A, Figure S4, Supplementary File 2). Therefore, while further evidence from activity assays are required to fully link 16S rRNA gene sequences to function, it is likely that at least some of these phylotypes represent functional sulfate reducers.

### 3.4.2 Archaeal-associated sequences and potential biogeochemical activities

*Methanomicrobia.* The clone library evidence and the porewater methane concentration and biogenic signature of the South (91-156.8 μM; $\delta^{13}CH_4$ -75.5 to -76.7 ‰) and North (564.97-1393.44 μM; $\delta^{13}CH_4$ -89.6 to -83.1‰) Orca Basin hypersaline sediments suggested that halophilic or halotolerant methanogens were present and were metabolically active. *Methanohalophilus* spp., the most abundant methanogenic phylotype-associated sequence found in the clone libraries (Figure 3.3), have been detected in the Mediterranean Sea and Red Sea DHABs (Antunes et al., 2011; Siam et al., 2012), and are common in hypersaline environments due to the availability of noncompetitive methylated substrates that are byproducts of fermentation of compatible solutes (Liu and Whitman, 2008). However, the Illumina 16S rRNA gene amplicon sequencing detected very low abundance of Methanomicrobia (<1%), including *Methanohalophilus sp.*, in Orca Basin hypersaline sediments; instead sequences were dominated by Bacteria and Marine Group I Archaea. Differences between the Sanger and Illumina taxa abundances of Methanomicrobia may be due to methodological bias, as indicated by the general low abundances of Methanomicrobia in a variety of Gulf of Mexico sediment environments, including one seep environment (Chapter 2).

ANME organisms were detected in Orca Basin through Sanger sequencing, though sequences were low in percent abundance (Figure 3.3, Appendix A,Figure S3). Anaerobic oxidation of methane has been proposed to be primarily mediated by methane oxidizing archaea (ANME) in consortia with sulfate-reducing bacteria (SRB) (Knittel and Boetius, 2009). Few studies have reported on ANME organisms in hypersaline environments. Only ANME-1b archaea were detected by molecular methods in a Gulf of Mexico brine pool methane seep (Lloyd et al., 2006b), indicating that this group may tolerate hypersaline conditions to some

extent. ANME-1 was also observed to dominate a hypersaline mud volcano (Maignien et al., 2013) and fluorescence *in situ* hybridization (FISH) indicated no associated bacterial partner. Putative AOM-associated ANME 1 and ANME 2 archaea were also detected by pyrotag sequencing in sediments of the Atlantis II and Discovery Deep DHABs of the Red Sea, with the majority of ANME-related sequences at a depth with higher sulfur content (19.1% S) in comparison to other depths (<0.1% to 9% S). ANME-1 were detected in very low abundaces (<0.01%) in the Orca Basin Illumina sequences (Supplemental File 1).

In Orca Basin hypersaline sediments, most Deltaproteobacteria taxa were phylogenetically distinct from those recovered in non-hypersaline sediments (Appendix A, Figure S3), consistent with observations in other high salinity environments. The Illumina sequencing data also suggested that putative sulfate-reducing syntrophs were present. Sequences related to the SEEP-SRB1 group (Knittel et al., 2003) were observed in all Orca Basin cores, on the slope and in both of the brine pools (Appendix A, Figure S3). SEEP-SRB1 sequences were particularly abundant in the South Basin core at 50-55 cm, comprising 23.4% of all Deltaproteobacteria sequences. However, not all SEEP-SRB1 have been observed to be associated with ANME-2 clade, and all SEEP-SRB have also been shown to be free living cells (Schreiber et al., 2010). Two other Deltaproteobacteria groups that have been reported to be physically associated with ANME, the SEEP-SRB2 (Kleindienst et al, 2012) and sequences related to *Desulfobulbaceae* (Losekann et al., 2007) were detected. It is also possible that unidentified partners of ANME evolved at high salinity.

*Marine Group I.* MG1 archaea-related sequences were abundant in Orca Basin sediment, being the most abundant archaea in almost every Miseq amplicon sample (Figure 3.5), and the most abundant 16S rRNA gene group in the Orca Basin Sanger sequencing dataset, with the sole

exception of the North Basin 50-55 cm layer (Figure 3.3). Similar observations were made in sediment cores of the Atlantis II and Discovery Deep of the Red Sea (Siam et al. 2012), where the most abundant archaeal pyrotags (73.3 to 96.1%) from every sediment section were MG1-affiliated, except the deepest (3.5 m), in which MG1 comprised of 1.9% of total archaeal reads in Atlantis II Basin and 34.3% in Discovery Deep Basin. These sections were higher in sulfur (19.1% S) than other sections of the core (0.1% to 9% S) in Atlantis II and higher in nitrogen (0.22% N) than other sections of the core (<0.10%) in Discovery Deep.

These data suggest either accumulation and preservation of MG1 cells and/or their DNA under hypersaline conditions (Boere et al., 2011) or their survival and growth in Orca Basin sediments. Possible energy sources could be ammonia oxidation or potentially heterotrophy linked to nitrate reduction, as suggested previously for MG1 archaea in anoxic, but nitrate-replete marine sediments (Durbin and Teske 2011; Jorgensen et al. 2012). The Orca Basin hypersaline sediments were enriched in $NH_4$ (>200 µM) and contained detectable amounts of NOx (ca. 1 to 7 µM), despite undetectable concentrations in the overlying brine (Figure 2b). To survive in Orca Basin, MG1 Archaea would also require physiological adaptations to hypersaline conditions. A recent study of a *Nitrosopumilus*-related MG1 genome analyzed from a Red Sea brine basin (salinity 182) indicated potential for osmotolerance genes (specifically ectoine production, glutamate transporters, and a proline-glutamate conversion enzymes) , as well as genes involved in ammonium oxidation (amoA) (Kamanda Ngugi et al., 2015).

### 3.5 Conclusions

The hypersaline sediments of Orca Basin host diverse microbial taxa. The 16S rRNA gene sequencing data indicated that the community composition in the North and South basin hypersaline sediments were similar to each other, but differ from the red sediments of the basin slope and sediment collected from the Continental Slope. Salinity was likely an influencing factor, with both North and South basin brine sediments containing Illumina amplicon sequences most abundant in Bacteroidetes, Marine Group I and halotolerant-associated Deltaproteobacteria, taxa consistent with other investigations in hypersaline environments (Antunes et al., 2011).

Sanger sequencing yielded a higher relative abundance of Candidate Division KB1 sequences from hypersaline sediments, but otherwise indicated similar taxonomic composition. In the red sediment of the Orca Basin slope (MUC 8 sample), sequences related to Candidate Division JS1 bacteria, the BHI180-138 uncultured bacterial group and Chloroflexi were most prominent. Another sample taken on the Orca Basin slope ('pink Jello' Alvin core 4650-24) had less abundant JS1 sequences and contained higher percentages of Proteobacteria, similar to the Continental Slope core, and *Defferibacteres* sequences, indicating diversity of microbial communities along the Orca Basin slope (at a depth above the influence of brine). Orca Basin hypersaline sediment porewater contained high concentrations of $CH_4$ (500-3500 uM) and $NH_4^+$ (300-500 uM). *Methanohalophilus* spp. related sequences were detected in both the South and North basins, and were likely contributors to the production of $CH_4$, while no large percentage of sequences related to known anaerobic $CH_4$ oxidizers were detected by either sequencing method.

Marine Group I related sequences were abundant in hypersaline sediments with both Sanger and Illumina sequencing methods, similar to sequences analyzed from the Red Sea brine

basins (Siam et al., 2012), and may be a possible sink for $NH_4$ in the basin. Further analyses into the microbial community, including transcriptomics, metagenomics and genome analysis will give more insight to microbial dynamics in extreme hypersaline environments.

# CHAPTER 4: THE PHYLOGEOGRAPHY AND POTENTIAL METABOLIC FUNCTIONS OF THE CANDIDATE DIVISION KB1 BACTERIA

## 4.1 Introduction

Hypersaline environments characterized by extremely high salt concentrations are abundant on Earth (e.g. Boetius and Joye, 2009) and include hypersaline lakes, such as the Dead Sea or the Great Salt Lake, evaporative lagoons, salterns, and hypersaline marine basins where salt deposits dissolve into deep-sea water and create distinct "lakes" of high-density brines on the seafloor of the Red Sea, Mediterranean Sea and the Gulf of Mexico. These deep-sea hypersaline anoxic basins (DHABs) have been the recent focus of molecular microbial investigation (Eder et al., 2001; van der Wielen, 2005; La Cono et al., 2011; Yakimov et al., 2013). In spite of their harsh conditions, 16S rRNA gene sequencing of DHABs indicate that they contain abundant and diverse microbial life that has adapted to these extreme environments. Several uncultured groups have been initially observed in DHAB microbial diversity studies, including Candidate Division KB1 (Bacteria) and MSBL-1 (Archaea), and other 'MSBL' bacterial groups within the phylum Proteobacteria (Daffonchio et al., 2006; Eder et al., 2001; van der Wielen, 2005). Due to the difficulties of culturing environmental organisms, especially from extreme environments, few of these microbes have been cultured and formally described (reviewed in Antunes et al., 2011).

High salinity environments are particularly challenging to microorganisms because biological membranes are permeable to water, requiring cells to maintain a cytoplasmic solute concentration higher than that of the surrounding brine in order to maintain turgor pressure and prevent water loss. There are two known strategies microorganisms use to prevent dehydration:

1) the "salt in" strategy, where organisms actively pump salt into the cell (usually $K^+$ and $Cl^-$) and 2) the "salt out" strategy, where organisms produce or transport compatible organic solutes into the cytoplasm. Microorganisms that predominantly use the "salt in" strategy also require special enzymatic adaptations, namely a higher proportion of acidic amino acids and a lower proportion of hydrophobic amino acids. This same bias prevents normal enzymatic function at lower salinities and, therefore, "salt in" strategists have been documented to be specialists and true halophiles. On the other hand, organisms predominantly using the "salt out" strategy produce organic-compatible solutes that do not require special intracellular protein adaptations (Oren, 2008) Compatible solute production or transport can be regulated by the cell, allowing "salt out" strategists to be generalists, capable of living at a range of salinities. However, compatible solute production is more energetically costly than maintaining high intracellular salt concentrations (Oren, 1999).

Of particular interest are the Candidate Division KB1 bacteria, which have been found exclusively in high salinity environments, and of which little is known. KB1 was initially detected by 16S rRNA gene sequencing from the Kebrit Deep brine basin in the Red Sea (Eder et al., 1999). The 16S rRNA gene sequence indicates KB1 is closely related to but phylogenetically distinct from the Candidate Division OP1 phylum, initially detected from the Obsidian Pool hot spring in Yellowstone National Park (Hugenholtz et al., 1998), and was considered a distinct Candidate phylum. Subsequently, sequences from a variety of habitats (cold seeps, hypersaline environments, high temperature environments, hydrocarbon-contaminated environments, marine and lake sediments) that are phylogenetically related to KB1 and OP1 have been detected by 16S rRNA gene sequencing (Fig 4.1). This has led to confusion of the phylogenetic status of KB1 and OP1 as separate phyla, with conflicting taxonomic classification from the major taxonomic

classifiers with SILVA (Ref 119) classifying all OP1- and KB1-related organisms under one phylum, KB1, and undefined lower taxonomic classification. Greengenes (ver. 13.8 ) recently classified OP1 and KB1-related sequences into one phylum, OP1 (McDonald et al., 2012), with distinct class levels of OP1, KB1, Acetothermia, and MSBL6. Acetothermi has also been suggested as a phylum name for the entire OP1/KB1 complex (Rinke et al., 2013), based on the first sequenced genome of a bacterium from a geothermal acidic stream (Takami et al., 2012). While the metabolic functions of KB1 are largely unknown, enrichments containing predominately KB1 and the archaeal putative halophilic methanogen MSBL1 incorporated high concentrations of $^{14}$C-labeled glycine betaine (GB), which resulted in increased concentrations of trimethylamine (TMA) detected in the media. This suggested possible cleavage of GB into acetate and TMA, and syntrophy with TMA-consuming methanogens (Yakimov et al., 2013).

This chapter reports the initial analyses of the phylogeography of KB1 and its phylogenetic relationships to Candidate Division OP1 Bacteria. Potential metabolic roles and osmotic stress strategies based on a partial single cell amplified genome (SAG) from Orca Basin (Chapter 3) are also discussed and support the hypothesis that KB1 is a "salt in" strategist that has the potential to use glycine betaine (GB) as a carbon and energy source and utilizes the acetyl-CoA pathway.

## 4.2 Materials and Methods

### 4.2.1 Sample collection and processing

Samples were collected in April 2012 onboard the *R/V Pelican* (Expedition PE22-12) at Orca Basin (Lat 26.5467 N, Lon 91.2165 W, 2410 mbsf) with Niskin bottles attached to a rosette sampler equipped with a CTD. Brine was transferred to an argon-flushed sterile canning jar and aliquoted into a $N_2$-flushed sterile stoppered serum vial. The sample was stored at 4$^{o}$C and sent to the Bigelow Single Cell Genomics Center (East Boothbay, ME). The sample was processed for single cell sorting followed by multiple displacement amplification of cellular DNA (Stepanauskas, 2012) and PCR screening of bacterial 16S rRNA genes followed by genome sequencing of the selected cell as described in detail in the methods file SCGC_Services_Description.pdf, available for download at scgc.bigelow.org/services/#br=services_description_table).

### 4.2.2 Genome Analysis

The assembled genome was uploaded and analyzed with the RAST annotation server (Aziz et al., 2008). Selected putative protein-encoding genes were further investigated by BLASTP searching the NCBI nonredundant protein database (Altschul et al., 1990). Protein sequences were also acquired for comparison through Uniprot (http://www.uniprot.org). Amino acids were aligned using MUSCLE (Edgar, 2004). NJ trees were constructed using a Poisson correction with 1000 bootstrap replicates using MEGA6 (Tamura et al., 2013).

### 4.2.3 16S rRNA gene phylogenetic analysis

Sequences classified as KB1 (which includes both the KB1 and OP1 Candidate Divisions) in the SILVA nonredundant database (ver. 119) (Pruesse et al., 2007) were analyzed with the ARB

phylogenetic software package (Ludwig et al., 2004) using RaxML maximum likelihood (Stamatakis, 2006).

### 4.3 Results and Discussion

#### 4.3.1 Phylogenetics and phylogeography of KB1

Phylogenetic analysis of sequences related from the most recent nonredundant SILVA near-full length 16S rRNA gene database (ver. 119) found four distinct phylogenetic groups related to the KB1 and OP1 candidate phyla with high (>80%) bootstrap statistical support (Figure 4.1). The initial Obsidian Pool sequence (Hugenholtz et al., 1998) formed a monophyletic group with sequences from other high-temperature (e.g. NCBI accession EU645931) and hydrocarbon-influenced (e.g. NCBI accession EU542465) environments. The first sequenced genome of the OP1 Candidate Division, from the acidic hot spring isolate *C. Acetothermum autotrophicum*, was in a cluster separate from the Obsidian Pool OP1 clone sequences that included sequences from both high-temperature and hydrocarbon-impacted environments.

KB1-classified sequences from Orca Basin brine and hypersaline sediments formed a monophyletic group with the initial KB1 sequences from the Red Sea Kebrit Deep brine basin (Eder et al., 1999) and sequences from a wide range of other hypersaline habitats, including high salinity microbial mats, hypersaline lakes, solar salterns, and other deep-sea brine basins (Tkavc et al., 2011; Makhdoumi-Kakhki et al., 2012; van der Wielen, 2005) (Fig 1).

The fourth monophylectic cluster observed was comprised mainly of sequences from hydrocarbon-impacted environments, including cold seeps, sediment near sulfate methane transition zones (SMTZ), and a methane-influenced brackish lake from which a partial single-cell genome was obtained (Rinke et al., 2013). This group also included the Mediterranean Brine

Lake 6 (MSBL6) group that was initially observed in 16S rRNA gene sequences from various Mediterranean brine basins (van der Wielen, 2005).

Preliminary phylogenetic analysis of all sequences in the SILVA 119 database identified as KB1 (which includes all published sequences analyzed in Fig. 1) indicated that all 348 sequences that were monophyletic with the initial Kerbit Deep brine basin sequence (Eder et al., 1999) were from high-salinity environments.

### 4.3.2 General overview of the KB1 genome

The Orca Basin KB1-classified SAG genome was 709508 bp and was estimated to be 20% complete. The sequenced genome contained sequences for at least 12 tRNAs. Putative genes were identified for vitamin and cofactor biosynthesis and metabolism, cell wall biosynthesis, RNA protein metabolism, DNA metabolism and repair, fatty acid and lipid metabolism, genes potentially related to oxidative stress response, and genes coding for putative hot and cold shock proteins (Supplemental File 1). Genes coding for ABC transporter proteins for glycine betaine (GB) uptake and the complete acetyl-CoA cycle were also observed (Figures 4.2 and 4.3).

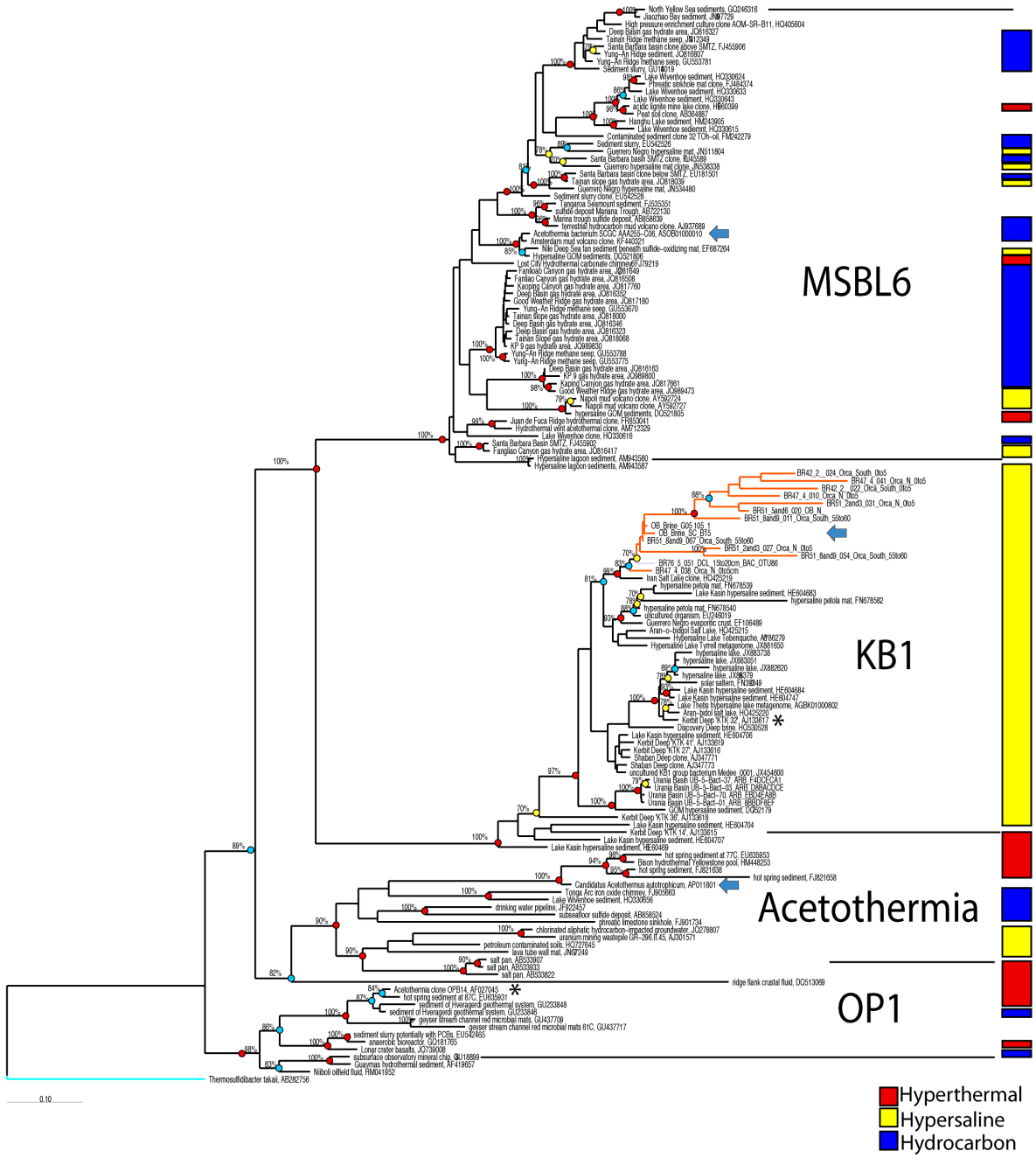



**Figure 4.1 RaxML tree of near full length 16S rRNA genes of Candidate Division KB1 and related bacteria.** The GTRMIX rate distribution model was utilized with the rapid bootstrap algorithm. Circles indicate bootstrap support of 1000 replicates, red 90-100%, turquois 80-89% and yellow 70-79%. Orange branches indicate Orca Basin sequences. Blue arrows indicate sequences from partial or near-complete genome sequences. Initial clones of the OP1 and KB1 Candidate Phyla are indicated with an asterisk.

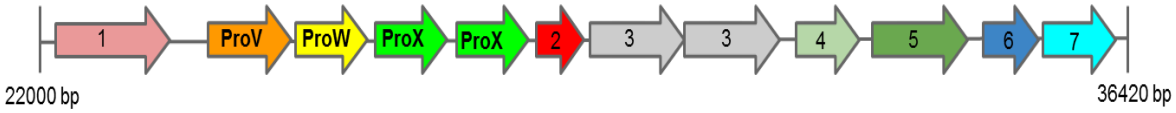



**Figure 4.2 Chromosomal region of the putative PROU ATP binding cassette operon for glycine betaine transport.** Enzyme Commission number (EC) and protein encoding gene (peg) identification number in RAST noted when availalble. The PROU region (PROVWX) (T.C. 3.A.1.12.1 ,peg 571-574) is nearby putative protein encoding genes (peg) of 1) hypothetical protein (putative methyltransferase) (peg 570), 2) 5-methyl-THF:corrinoid methyltransferase; E.C. 2.1.1.258 (peg 575), 3) Glycine decarboxylase (P1 and P2 proteins), 4) 5-methyltetrahydrofolate--homocysteine methyltransferase EC 2.1.1.13 (peg 578), 5) formyl-THF-ligase; E.C. 6.3.4.3 (peg 579), 6) GTP cyclohydrolase I (EC 3.5.4.16) type 1, 7) methylene-THF dehydrogenase; E.C. 1.5.1.5 (peg 581)

### 4.3.3 Osmoregulation and metabolism

The osmotic regulation strategy of the sequenced KB1 is of particular interest since initial phylogeographic analysis indicates that the KB1 phylogenetic clade contains sequences solely from hypersaline environments. KB1 contains genes potentially coding for compatible solute uptake (specifically GB or proline betaine) and $K^+$ uptake (Supplemental File 1), indicating either intracellular accumulation of compatible solutes ("salt out" strategist), accumulation of potassium salts ("salt in" strategist) or a combination of these strategies.

One indication may come from the distribution of predicted protein isoelectric points in the KB1 genome. Accumulation of salts, primarily $K^+$, for osmotic regulation was first observed in halophilc Archaea in the phylum Halobacteria, and subsequently in the bacterium *Salinibacter ruber* in the phylum Bacteroidetes (Antón et al., 2002; Mongodin et al., 2005). The predicted proteins inferred from these organisms' genomes contain a higher proportion of acidic amino acids than do proteins from microorganisms that do not accumulate intracellular salts (Mongodin et al., 2005). The isolelectric point abundance of inferred proteins from the KB1 SAG (Figure 4.4) indicated that the organism contained proteins with higher abundances of acidic amino acids (30% pI 5.0), similar to *Salinibacter ruber*, but contained less proportional

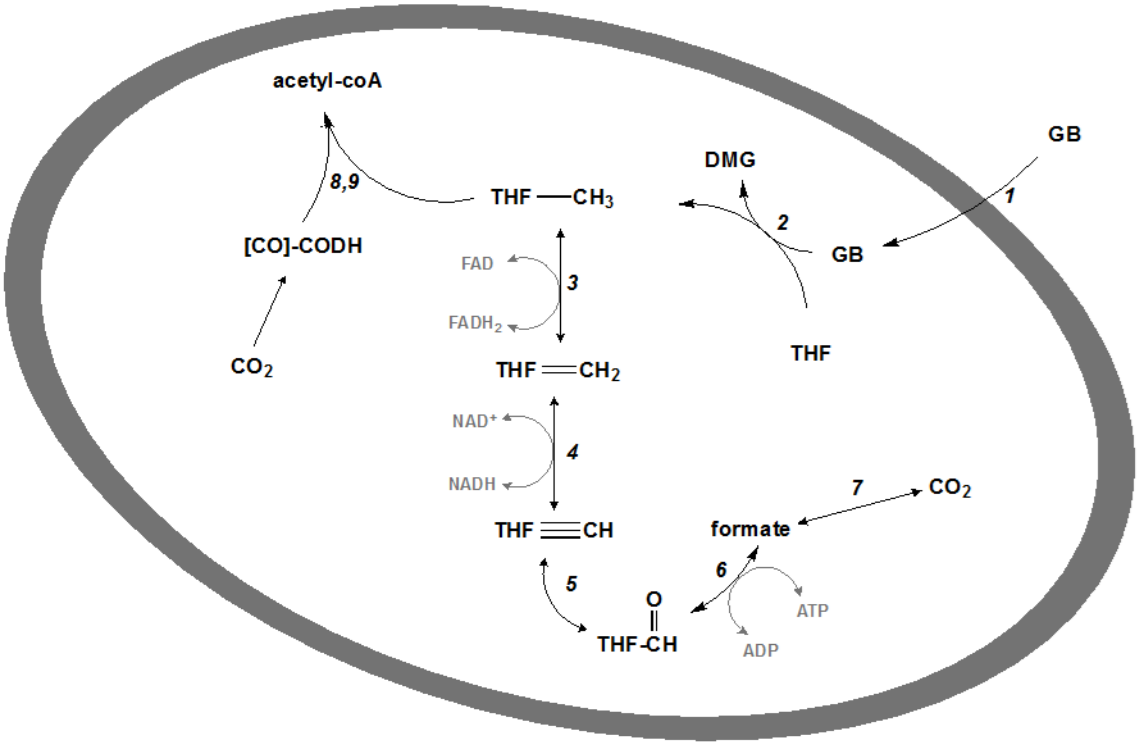



**Figure 4.3 : Proposed fate of glycine betaine (GB) in KB1.** Enzyme Commision number (EC) and protein encoding gene (peg) identification number in RAST noted when appropriate. (1) GB is imported into the cell via the ATP binding cassette, ProU; T.C. 3.A.1.12.1 (2) Trimethylamine methyltransferase then transfers one of the methyl groups of GB to tetrahydrofolate (THF) resulting in dimethylglycine (DMG) and 5-methyl-THF. The methylated THF can now enter the Wood-Ljungdahl pathway (3-9)
Reactions 3-7 are bidirectional: (3) methylene-THF reductase; E.C. 1.5.1.20 (peg 555) (4) methylene-THF dehydrogenase; E.C. 1.5.1.5 (peg 581) (5) methenyl-THF-cyclohydrolase; E.C. 3.5.4.9 (peg 581) (6) formyl-THF-ligase; E.C. 6.3.4.3 (peg 579) (7) formate dehydrogenase; E.C. 1.2.1.43 (peg 563); EC 1.2.1.2 ( peg 234) (8) 5-methyl-THF:corrinoid methyltransferase; E.C. 2.1.1.258 (peg 575) (9) carbon monoxide dehydrogenase/acetyl-coA synthase complex (CODH/ACS), CODH EC.1.2.99.2 (peg 248); acetyl-CoA synthase (peg 243, 244, 246).

acidic amino acids than the extreme archaeal halophile, *Halobacterium* NCR-1 (~42% pI 4.5).

Conversely, *C. Acetothermum autotrophicum* and the partial SAG genome *Acetothermia* SCGC

AAA255-C06 in the MSBL6 clade (Rinke et al., 2013) had similar protein isoelectric point

abundance patterns to organisms that do not accumulate high concentrations of intracellular salts,

including *Escherichia coli*, *C. Maribeggiatoa,* and *Desulfohalobium retbaense,* a sulfate-reducing bacterium isolated from the hypersaline Retba Lake (Senegal, West Africa) thought to accumulate compatible solutes as its primary mechanism for osmoregulation (Spring et al., 2010).

Amino acids were further analyzed to determine which acidic amino acids (glutamate or aspartate) may be contributing to the pI pattern in KB1. Analysis of individual amino acid abundances in proteins inferred from ORFs from the Orca Basin SAG indicated that glutamate was more abundant than aspartate, which was only marginally higher in abundance than the nonhalophilic or "salt out" strategist bacteria. However, bias for glutamate was not universal among the "salt in" strategists, as *Halabacterium* contained higher concentrations of aspartate, while *Salinibacter ruber* contained elevated but near equal abundances of glutamate and aspartate (Figure 4.5).

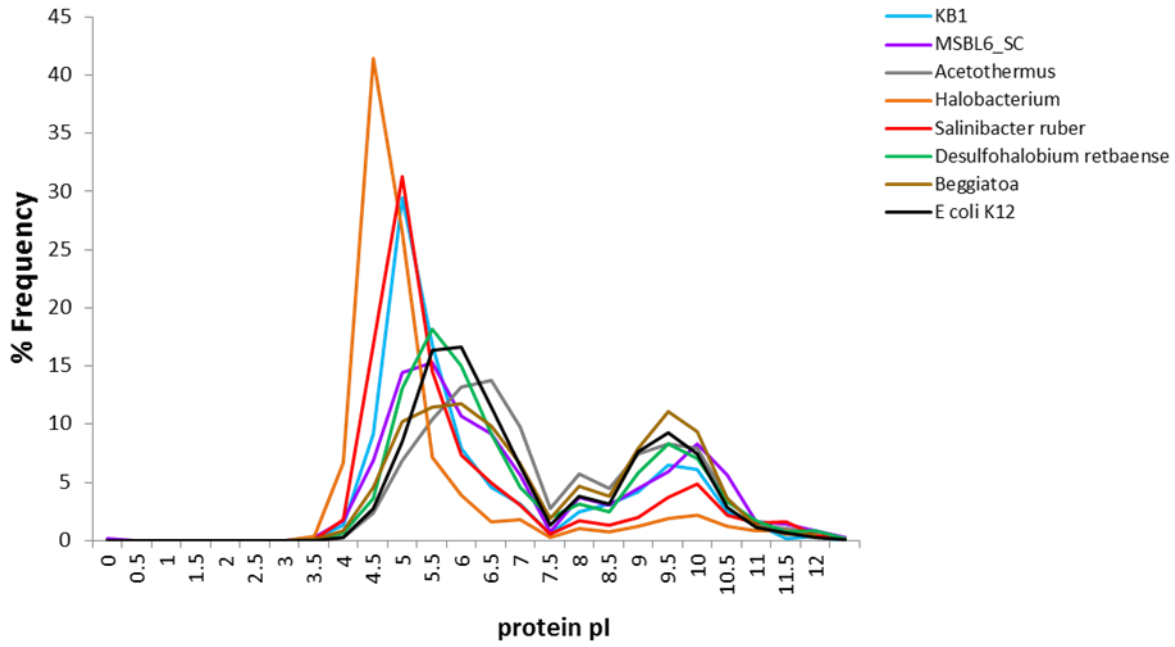



**Figure 4.4 Frequency of predicted Isoelectric point of putative proteins in the genomes of the OP1/KB1 phylum complex**. Genomes included: Orca Basin KB1, MSBL6 (Greengenes taxonomy) SAG from a methane and sulfidic zone of a brackish lake (SCGCAAA255-C06) (Rinke et al., 2013), C. Acetothermum autotrophicum (Takami et al.), Halobacterium NRC-1 a halophilic "salt in" archaeal halophile, Salinibacter ruber, a "salt in" bacterial halophile, Desulfohalobium retbaense a "salt out" bacterial halophile, C. Maribeggiatoa, a giant sulfur bacterium, and Escherichia coli K12.

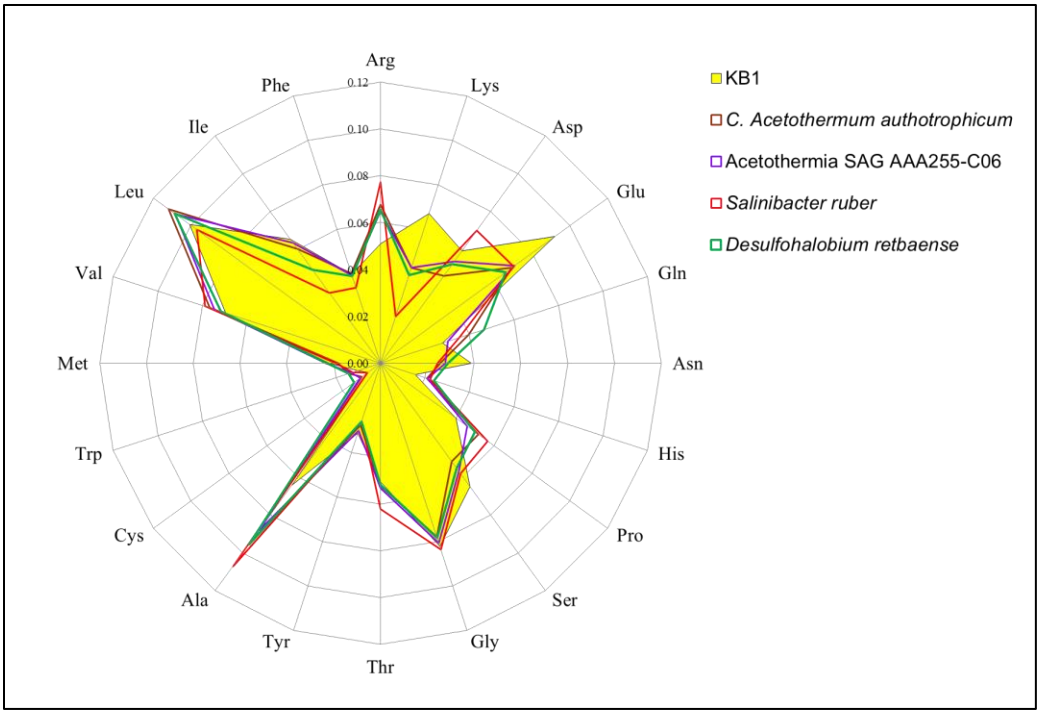



**Figure 4.5 Frequency of amino acids of putative proteins.** Genomes analyzed include memberes of the OP1/KB1 phylum complex (Orca Basin KB1, MSBL6 (Greengenes taxonomy) SAG from a methane and sulfidic zone of a brackish lake (SCGCAAA255-C06) (Rinke et al., 2013), *C. Acetothermum autotrophicum* (Takami et al.), *Halobacterium* NRC-1 a halophilic "salt in" archaeal halophile, *Salinibacter ruber*, a "salt in" bacterial halophile, *Desulfohalobium retbaense* a "salt out" bacterial halophile, *C. Maribeggiatoa*, a giant sulfur bacterium, and *Escherichia coli* K12.

### 4.3.4 Potential fate of glycine betaine in KB1

Glycine betaine (GB) is among the most common compatible solutes produced by "salt out" strategists in response to osmotic stress. GB transport systems can aid cells in maintaining osmotic balance and are well studied in cultured isolates, including *Escherichia* and *Salmonella* species. The Orca Basin SAG KB1 genome contains putative genes that encode for a 3-protein proline or glycine betaine ABC transporter system, PROU (genes *proVWX*) (Fig. 2). Once transported into the cytoplasm, GB has several possible fates (Heijthuijsen and Hansen, 1989; Möller et al., 1984; Müller et al., 1981; Naumann et al., 1983; Watkins et al., 2014): 1) It can be retained as a compatible solute, 2) It can be cleaved to acetate and trimethylamine (TMA), or 3)

it can be demethylated. In Archaea, GB demethylation has been observed to be connected directly to methanogenesis (Watkins et al., 2014).

Demethylation to glycine has been considered the dominant pathway in aerobic bacteria, while cleavage has been considered the dominant anaerobic pathway of GB catabolism. Demethylation does occur in anaerobic bacteria and methanogenic archaea, but the process has not been extensively studied. The KB1 genome includes up to 6 methyltransferases, including corrinoid-containing methyltransferases. Ticak et al. (2014) recently provided evidence that a family of corrinoid-containing methyltransferases that are related to trimethylamine transferases in methanogen *Methanosarcina* spp, demethlyate GB in *Desulfitobacterium hafniense*. The TMA methyltransferase protein was upregulated during growth of *D. hafniesnse* on GB, and further spectrophotometric evidence suggested that the recombinant protein converted GB and cob(l)alamin to DMG and methylcobalamin. KB1 contains a similar *Methanosarcina*-like corrinoid methyltransferase, though further investigation would be necessary to determine if these methyltransferases confer methylation of particular substrates (Figure 4.8). If the corrinoid protein does serve as a methyl transferase in KB1, one potential fate of the methyl group is its transfer to tetrahydrofolate (THF), producing methyl-THF, a key compound in the acetyl-CoA (Wood-Ljungdahl) pathway. The KB1 genome contains evidence of a complete acetyl-CoA pathway (Figure 4.3); THF-$CH_3$ could potentially be oxidized to $CO_2$ by utilizing the pathway in reverse.

Alternatively, KB1 may utilize one of the alternative pathways of GB catabolism (Figure 4.6). Yakimov et al. (2013) suggested that KB1 produces both acetate and TMA, and may be a syntroph of methylotrophic methanogens, as enrichments with abundant KB1 and the putative methanogen MSBL1 produced TMA following GB addition. Syntrophy between GB-

metabolizing bacteria and methylotrophic archaea has been documented (e.g. King, 1984). The cleavage enzyme (glycine betaine reductase) that traditionally produces acetate and TMA was not detected in KB1 (Figure 4.6), but cannot be discounted. The fate of DMG is also unknown, and no enzymes were detected to demethylate DMG in the genome, but potential enzymes were detected for oxidation of sarcosine (*N*-methylglycine) (SOX, EC 1.5.3.1) and glycine decarboxylase indicating the possibility of complete demethylation to glycine (Figure 4.6).

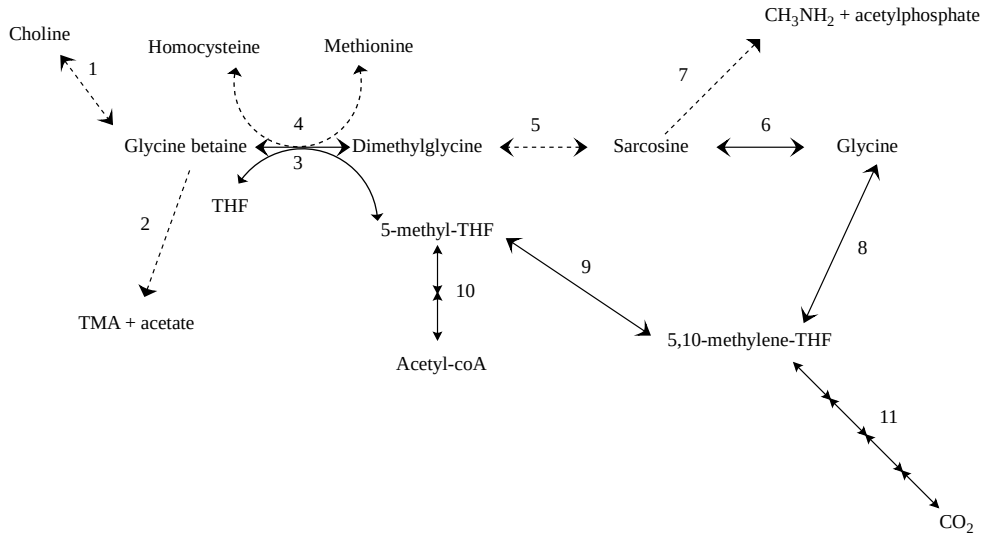



**Figure 4.6 Potential Metabolic fates of glycine betaine and its derivatives.** (1) choline oxidase; E.C. 1.1.3.17 (2) glycine betaine reductase (3) trimethylamine methyltransferase (4) betaine-homocysteine S-methyltransferase; E.C. 2.1.1.5 (5) dimethylglycine oxidase; E.C. 1.5.3.10 (6) sarcosine oxidase; E.C. 1.5.3.1 (7) sarcosine reductase; E.C. 1.21.4.3 (8) glycine dehydrogenase; E.C. 1.4.4.2 (9) methylene-THF-reductase; E.C. 1.5.1.20 (10) subsequent reactions involving a corrinoid:methyltransferase and the carbon monoxide dehydrogenase/acetyl-coA synthetase complex result in the formation of acetyl-coA-bidirectional (11) the bidirectional conversion of $CO_2$ to acetyl-coA (Wood-Ljungdahl pathway). Dotted lines indicate enzymatic reactions not found in the KB1 genome; solid lines are putatively present in the KB1 genome.

The protein sequences of the PROU system for glycine betaine uptake, methyltransferases, and acetyl-coA pathway enzymes were analyzed to investigate relatedness of proteins to other organisms. NJ trees of the PROU proteins (PROV,PROW, PROX) yielded different closest relatives, with PROW being monophyletic to an unidentified protein of *Halomonas sp.* BC04 and PROW to Actinobacteria, with weak (64%) bootstrap support. The two copies of PROX were different but monophyletic. Their relatedness to other in PROX enzymes could not be resolved (Appendix B). Putative methyltransferase sequences in KB1 were monophyletic to proteins of other bacteria that have been observed to demethylate glycine betaine (though function of these specific methyltransferases is largely unknown) and were closest to sequences from Peptococcaceae and *Acetohalobium arabiticum*. No GB uptake genes or methyltransferases were observed in the OP1-related genome of *C. Acetothermum autotrophicum* (Takami et al., 2012), indicating that GB uptake may not be universal to all members of the KB1/OP1 superphylum.

Analysis of proteins potentially involved in the acetyl-CoA pathway, including the COOS (acetyl- and ASCB (acetyl-CoA synthase proteins) indicate monophyly with *C. Acetothermum autotrophicum* (Figures 4.7 and 4.8). Complete enzymes of the acetyl-coA pathway have been recognized in the genome of *C. Acetothermum autotrophicum* (Acetothermi group Figure 4.1), indicating a potentially orthologous pathway. It was proposed that *C. Acetothermum autotrophicum* utilizes the pathway to fix $CO_2$ with $H_2$. If this assumption is true, KB1 may have applied the pathway in reverse after acquisition of GB and methyltransferase proteins. Conversely, KB1 may still utilize the pathway in the forward direction to process pyruvate, producing acetyl-CoA, or utilize the pathway in both directions (e.g. Hattori et al., 2005).

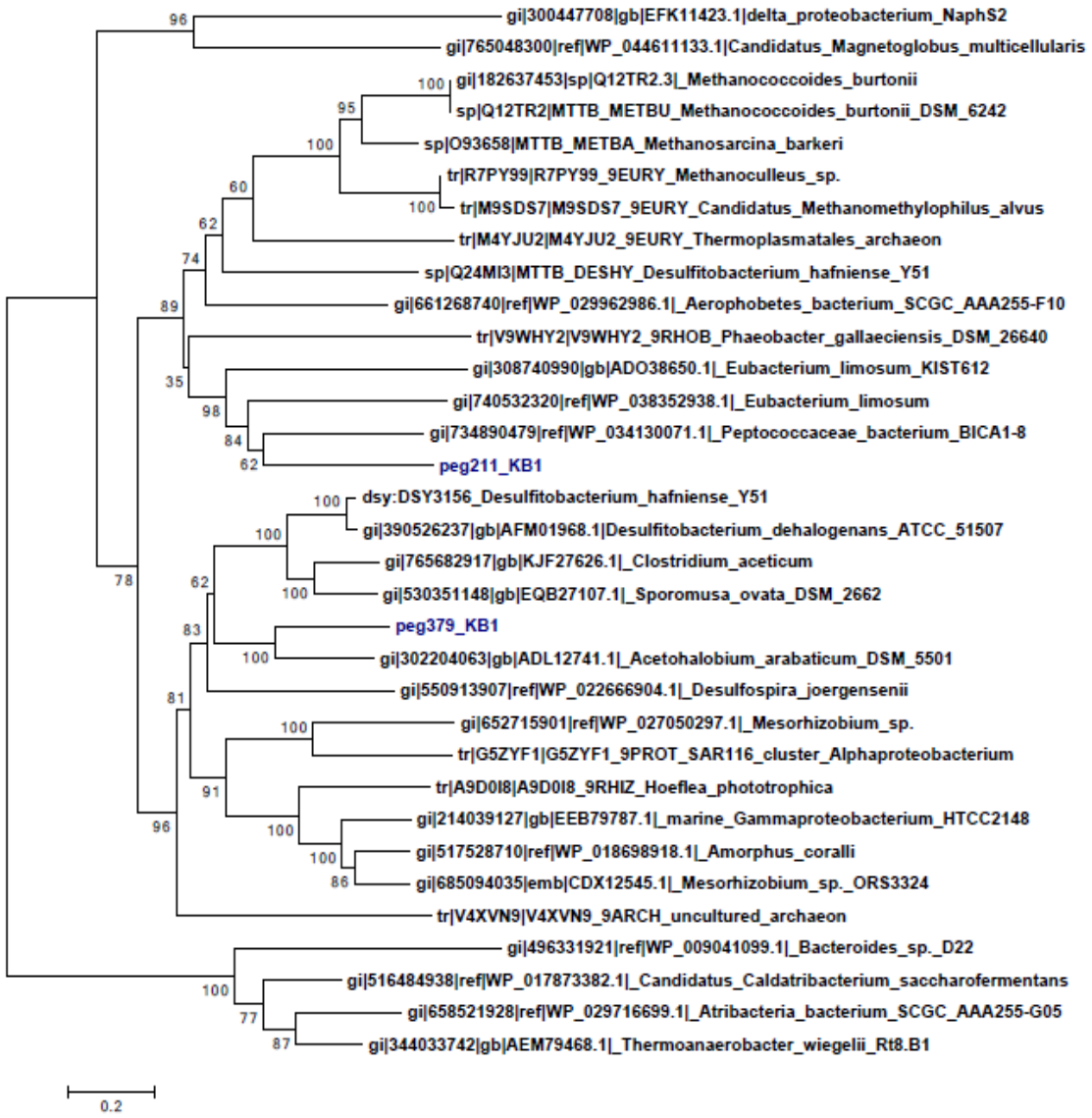



**Figure 4.7 NJ tree of putative corinnoid methyltransferases.** KB1 sequences are in blue. Tree was constructed with a Poisson model and 1000 bootstrap replicates. . Amino acid alignment positions with less than 20% informative information (including gaps) were not considered. Bootstrap statistical support (>50%) is displayed next to the node.

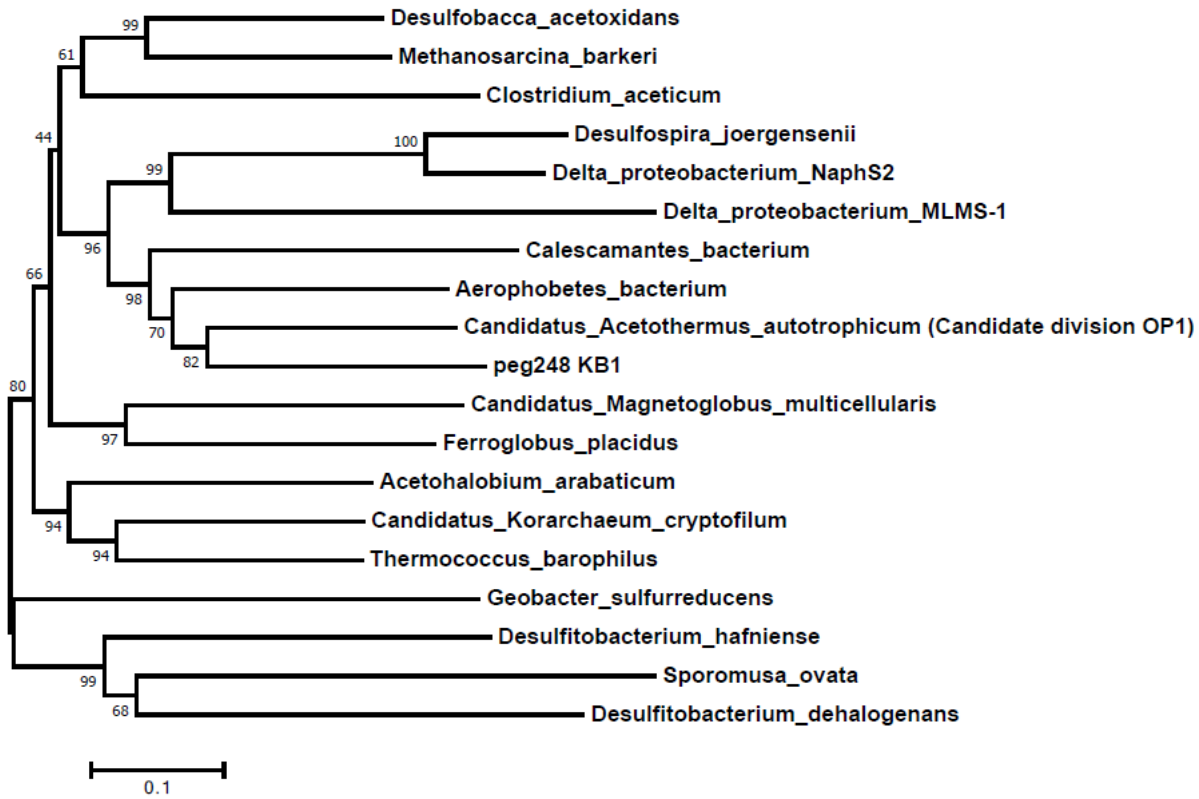



**Figure 4.8 NJ tree of COOS protein sequences in the acetyl-CoA pathway.** Tree was constructed with a Poisson model and 1000 bootstrap replicates. . Amino acid alignment positions with less than 20% informative information (including gaps) were not considered. Bootstrap statistical support (>50%) is displayed next to the node.

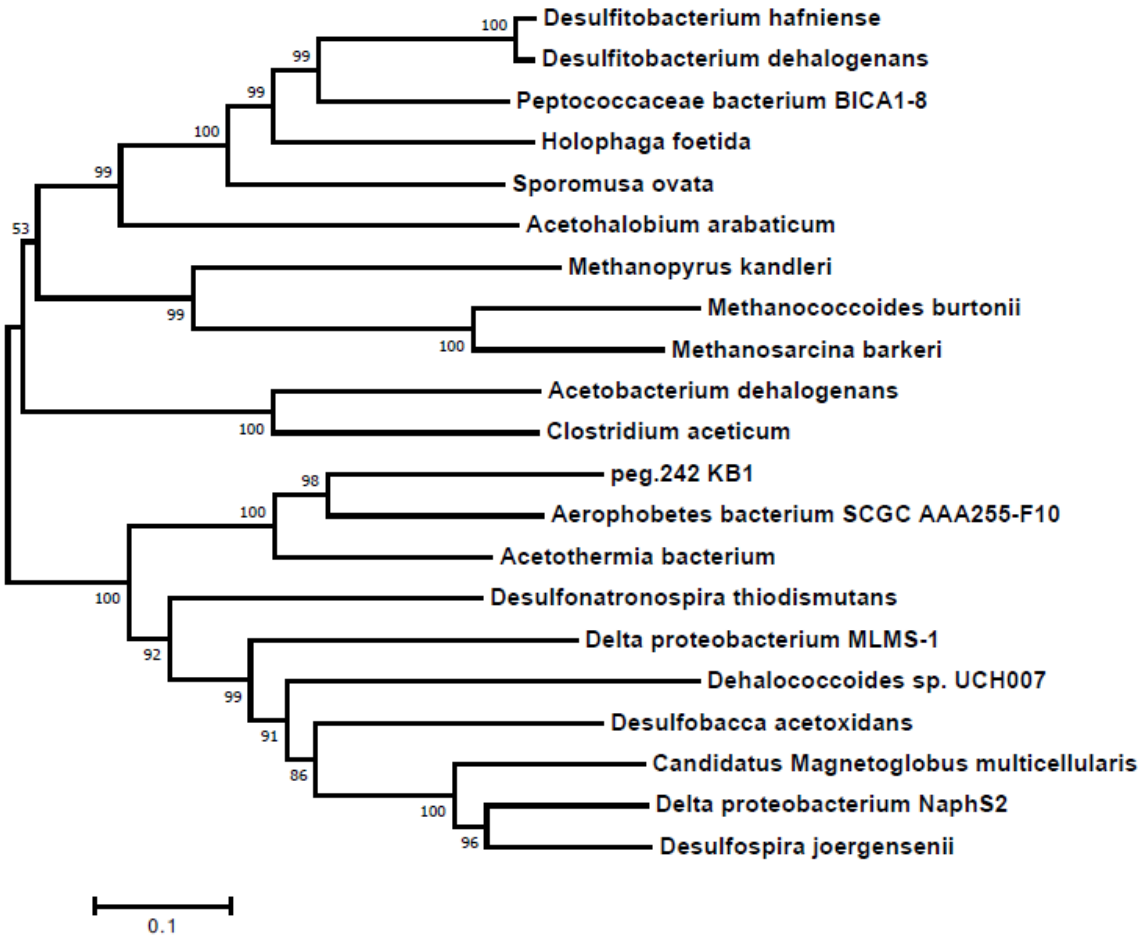



**Figure 4.9 NJ tree of ACSB proteins of the acetyl-CoA pathway.** Tree was constructed with a Poisson model and 1000 bootstrap replicates. Amino acid alignment positions with less than 20% informative information (including gaps) were not considered. Bootstrap statistical support (>50%) is displayed next to the node.

## 4.4 Conclusions

The Candidate Division OP1/KB1 bacteria are deeply rooted on the tree of life (Takami et al., 2012), and are phylogenetically related (Figure 4.1). The 4 distinct groups of the OP1/KB1 complex were consistent with taxonomic class distinctions of the Greengenes classification. Therefore, it is proposed to either combine OP1, KB1, MSBL6 (and relatives) and Acetothermum *autotrophicum* into one phylum and retain the four separate groups as class distinctions, or to split them into four different phyla. The acidic amino acid bias indicated evolutionary change in KB1 from its common ancestor with OP1, allowing it to thrive in a

hypersaline system (and potentially prohibiting it from living at low salt concentration). The acetyl CoA pathway is potentially orthologous, but GB and methyltransferases were potentially horizontally transferred. KB1 evolved to utilize a unique combination of pathways to survive in hypersaline environments.

# APPENDIX A: SUPPLEMENTAL MATERIAL FOR CHAPTER 3

## Supplemental Materials and Methods

The geochemical measurements were performed by a combination of laboratories and people. $Na^{+}$, $K^{+}$, $Mg^{2+}$, $Ca^{2+}$, $Cl^{-}$ and sulfate ($SO_4^{2-}$), were measured in the laboratory of Samantha Joye at the University of Georgia (UGA) by members of her group or by the author. All dissolved organic carbon (DOC), nitrate+nitrite ($NO_x$), phosphate and ($PO_4$) measurments were provided by UGA. Dissolved organic carbon (DIC) and isotopes (DIC $\delta^{13}C$) were measured in the laboratory of Christopher Martens at UNC or at UGA. Hydrogen and methane measurements were processed by Felix Elling at the University of Bremen, Germany in the laboratory of Kai-Uwe Hinrichs or provided by UGA. Methane isotopes ($CH_4$ $\delta^{13}C$) were measured by Felix Elling at the University of Bremen. Sulfide concentrations were provided by UGA or processed by the author at UGA or UNC.

The Illumina Miseq amplicon reads were provided by the laboratory of Jack Gilbert at the University of Chicago.

Chapter 3 Suppplemental figures are displayed below.

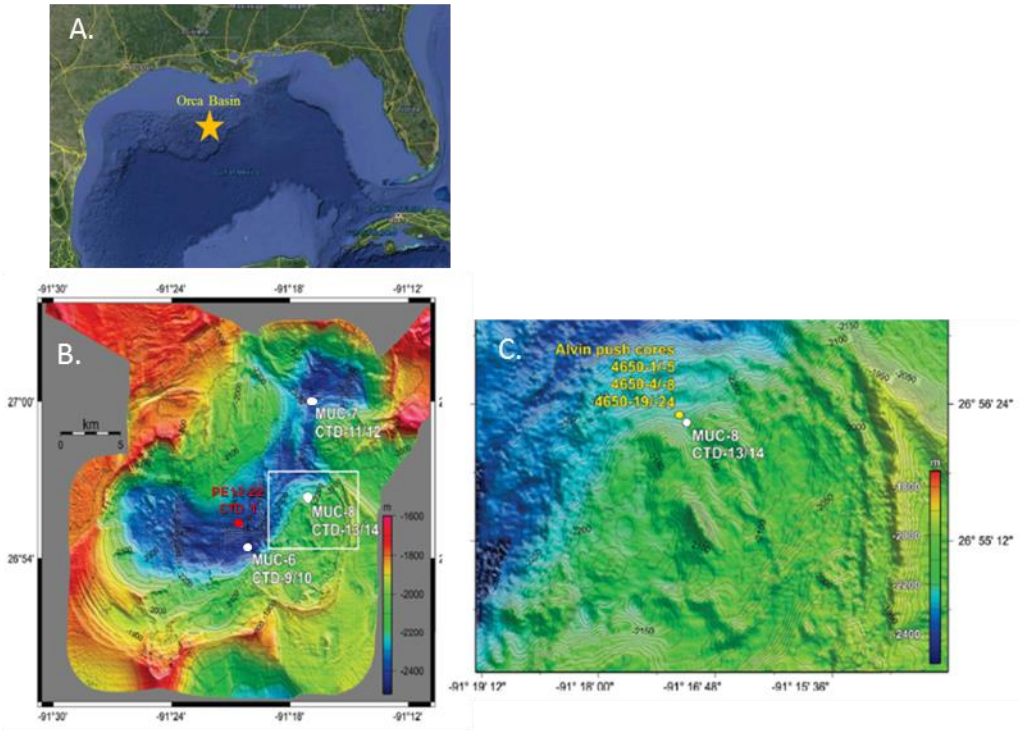



**Figure S1:** A. Position of Orca Basin. B. Sampling locations within Orca Basin displaying multicore and CTD locations from the *R/V Atlantis* expedition 18-2 (white) and the CTD location from the *R/V Pelican* expedition PE12-22 (red). C. The region indicated by the white box in B, showing the location of *DSV Alvin* push cores from expedition *AT18-2*.

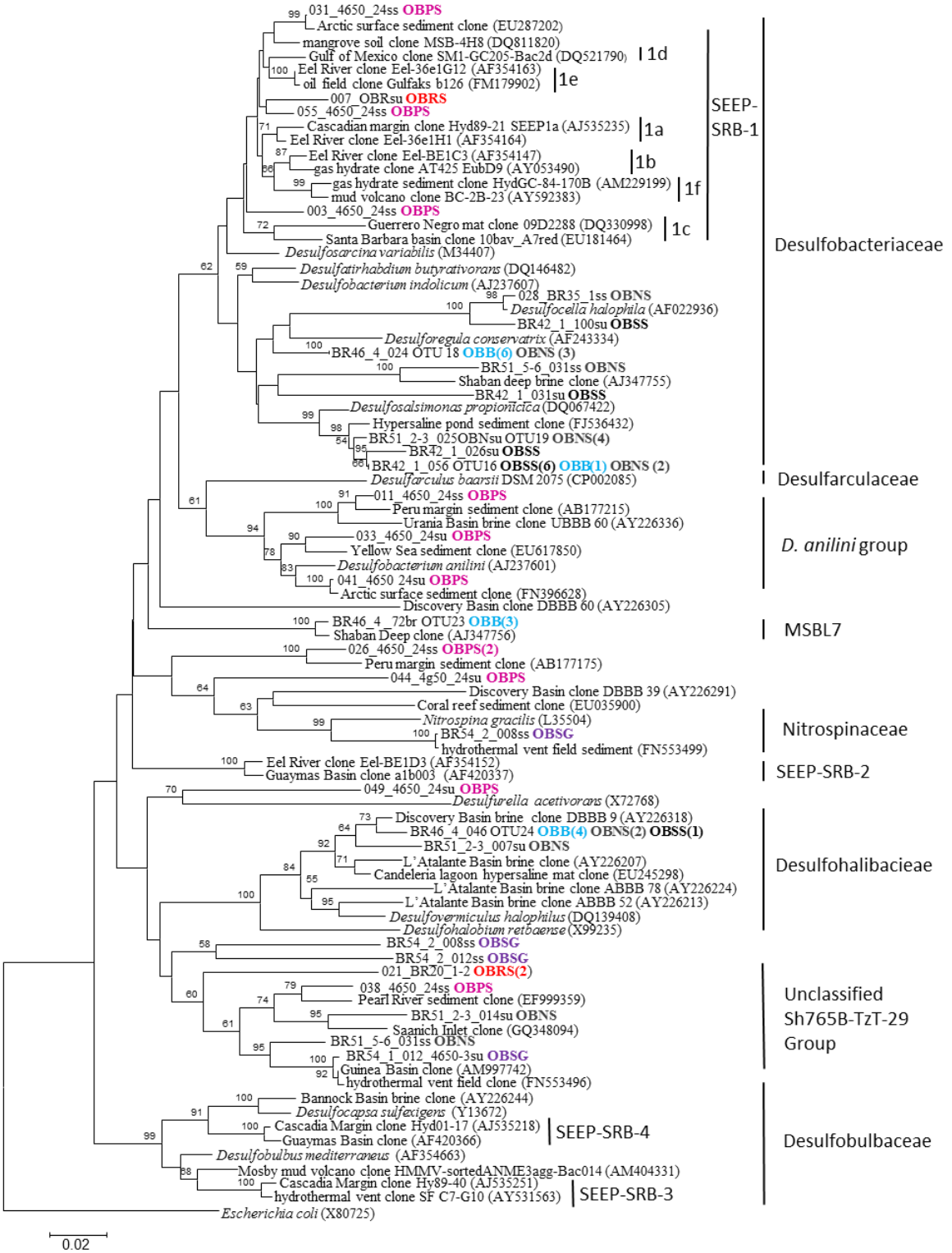



**Figure S2**: NJ tree of Orca Basin Deltaproteobacteria 16S rRNA gene clone sequences. Bootstrap values are displayed next to nodes. Only bootstrap values of >50% shown. OBB: Orca basin South brine pool, **OBSS**: Orca Basin South brine pool sediment (MUC 6), **OBNS**: Orca Basin North brine pool sediment (MUC 7), **OBRS**: Orca Basin slope red sediment (MUC 8), **OBPS**: Orca Basin slope pink sediment (core 4650-24). **OBSG**: Orca Basin sponge garden (core 4650-3) For sites with multiple sequences > 97% , number of clone sequences per OTU are indicated in parentheses.

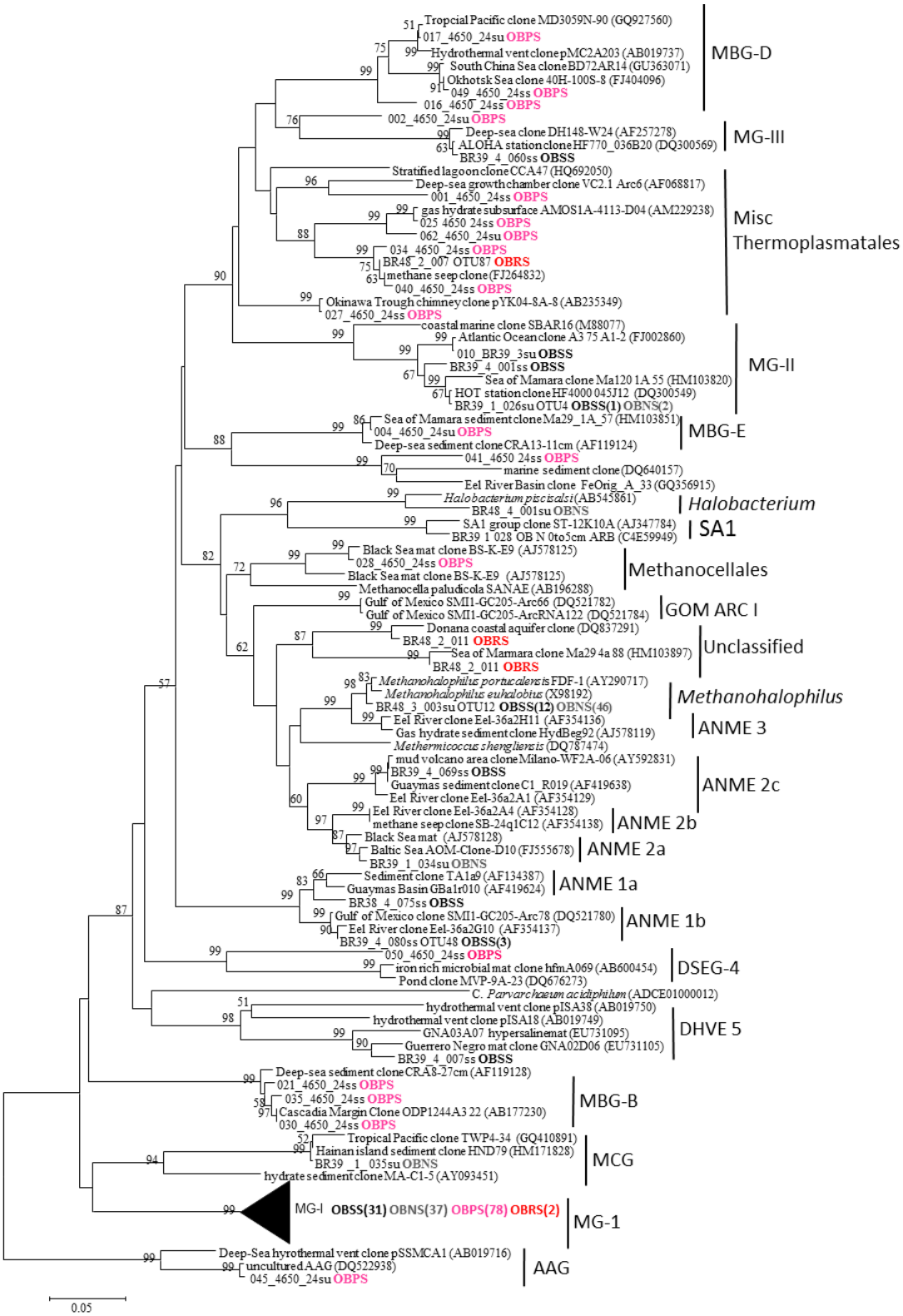



**Figure S3**: NJ tree of Orca Basin archaeal 16S rRNA gene clone sequences. Bootstrap values are displayed next to nodes. Only bootstrap values of >50% shown. **OBSS**: Orca Basin South brine pool sediment (MUC 6), **OBNS**: Orca Basin North brine pool sediment (MUC 7), **OBRS**: Orca Basin slope red sediment (MUC 8), **OBPS**: Orca Basin slope pink sediment (core 4650-24). For sites with multiple sequences > 97% , number of clone sequences per OTU are indicated in parentheses.

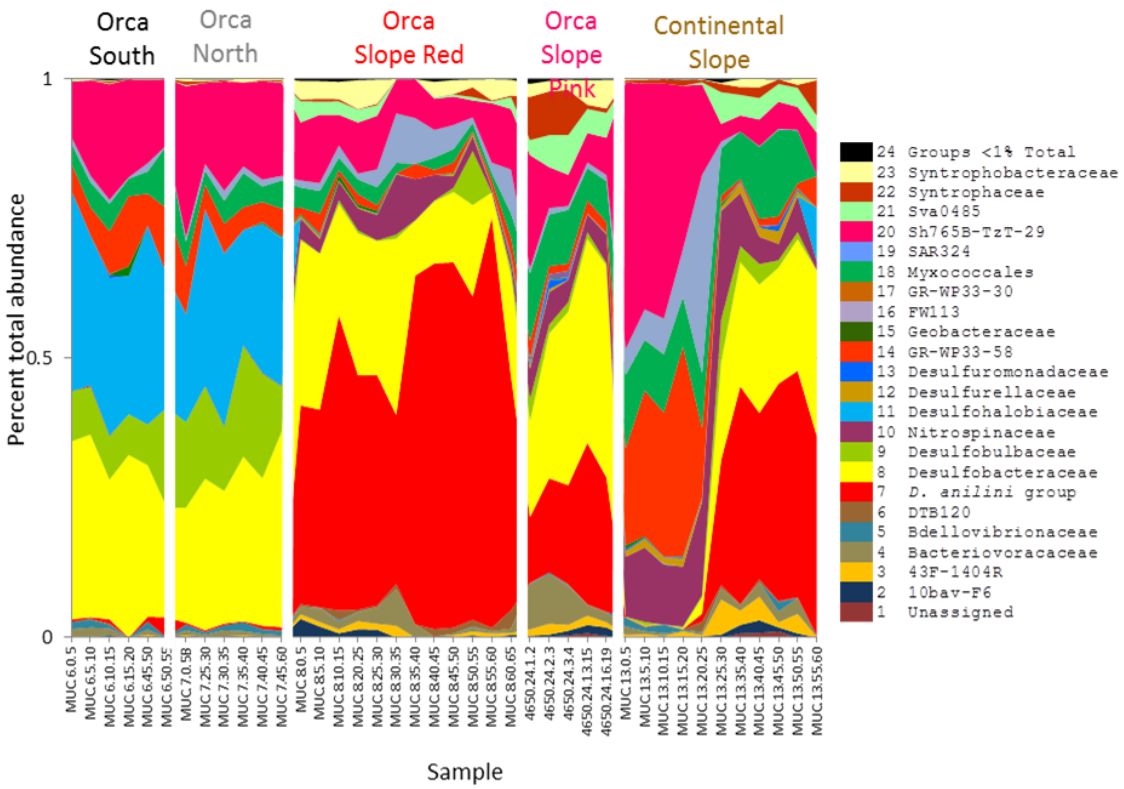



**Figure S4:** Relative abundance of Deltaproteobacteria.

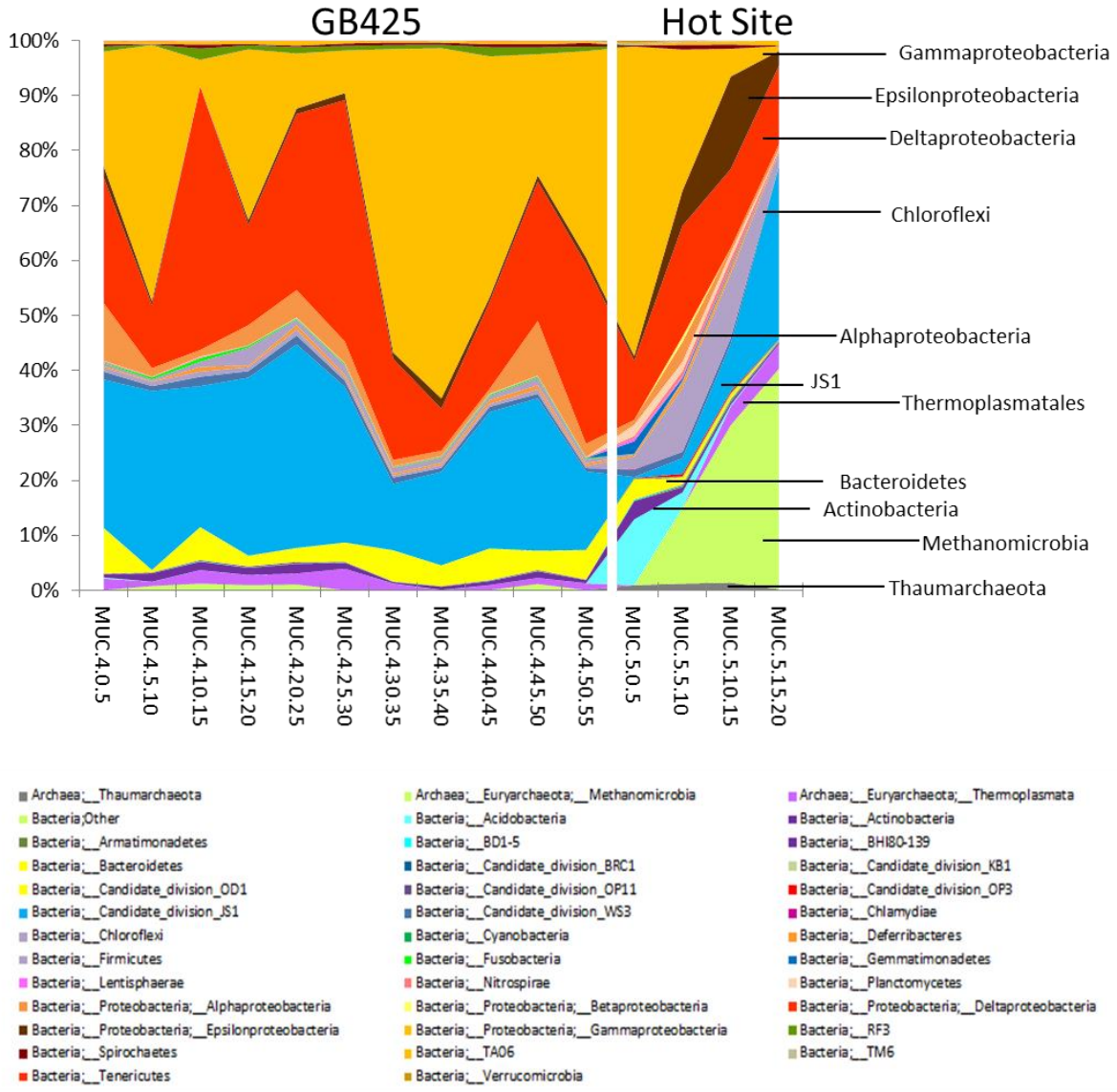



**Figure S5** : Relative abundance microbial groups MUC 4 and 5 in Beta diversity plots.

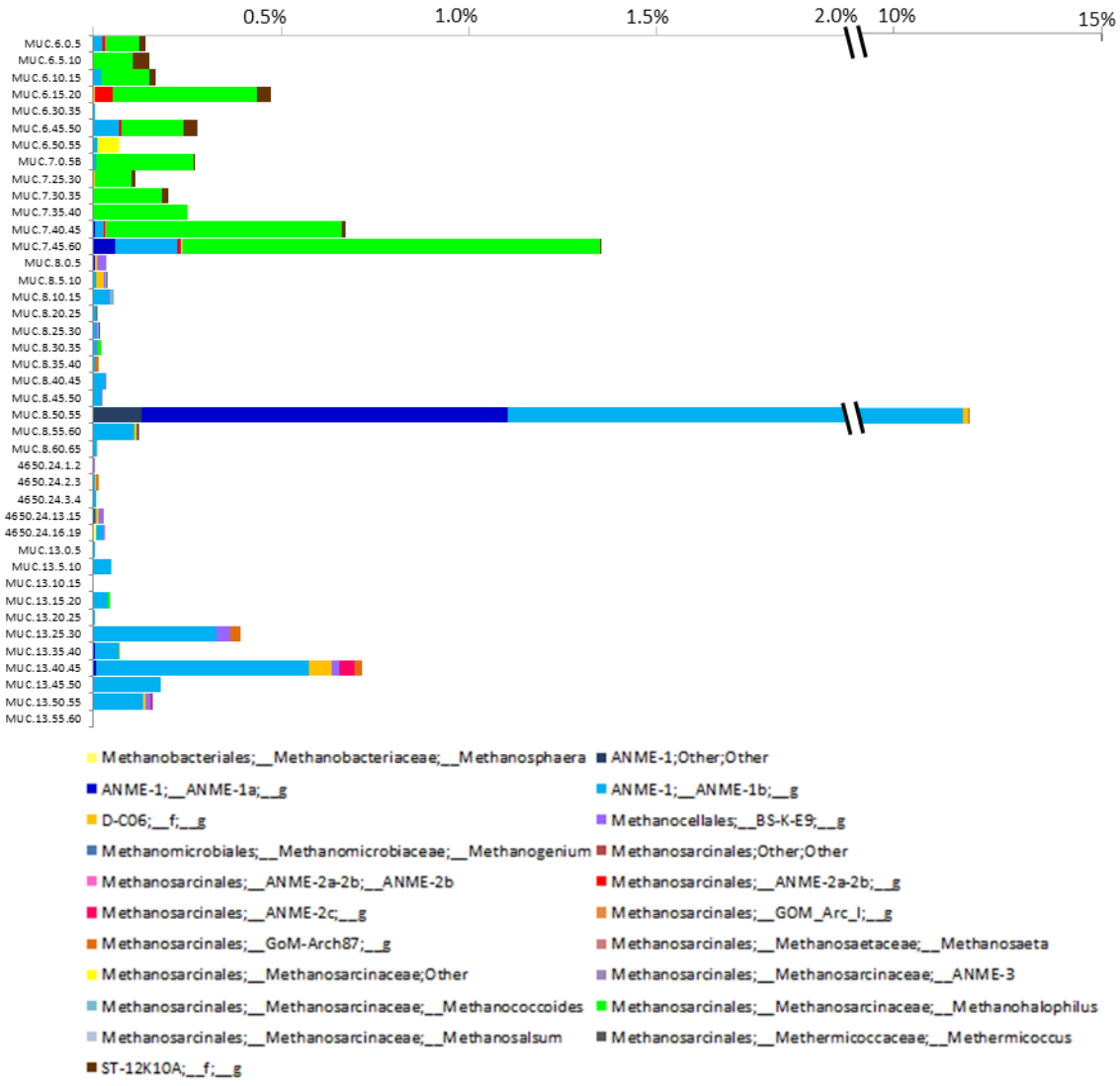
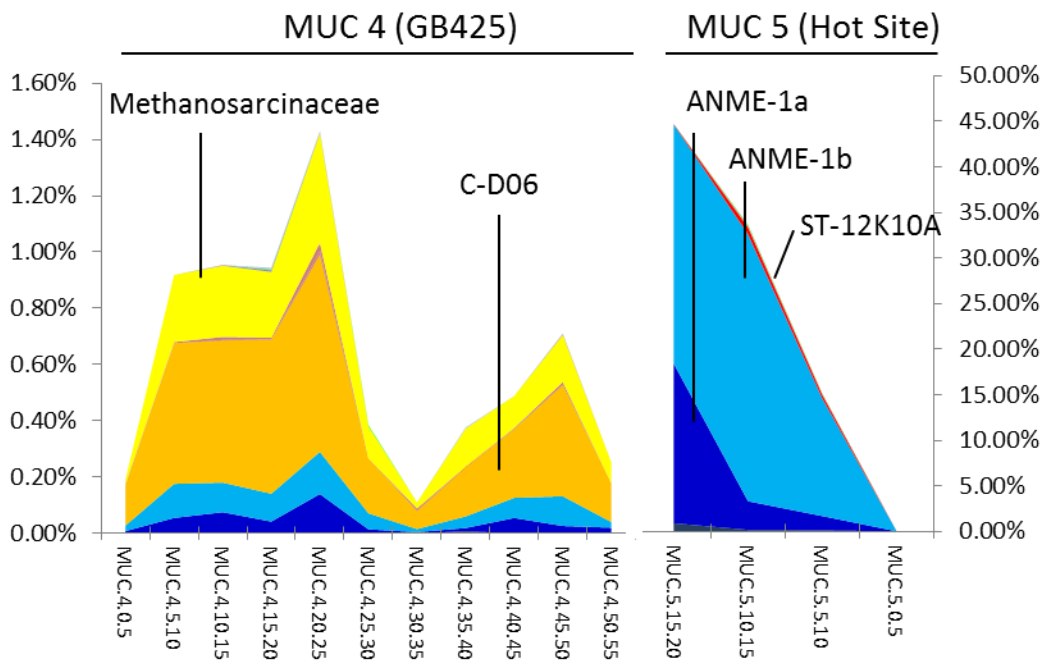



**Figure S6:** Relative abundance of Methanomicrobia

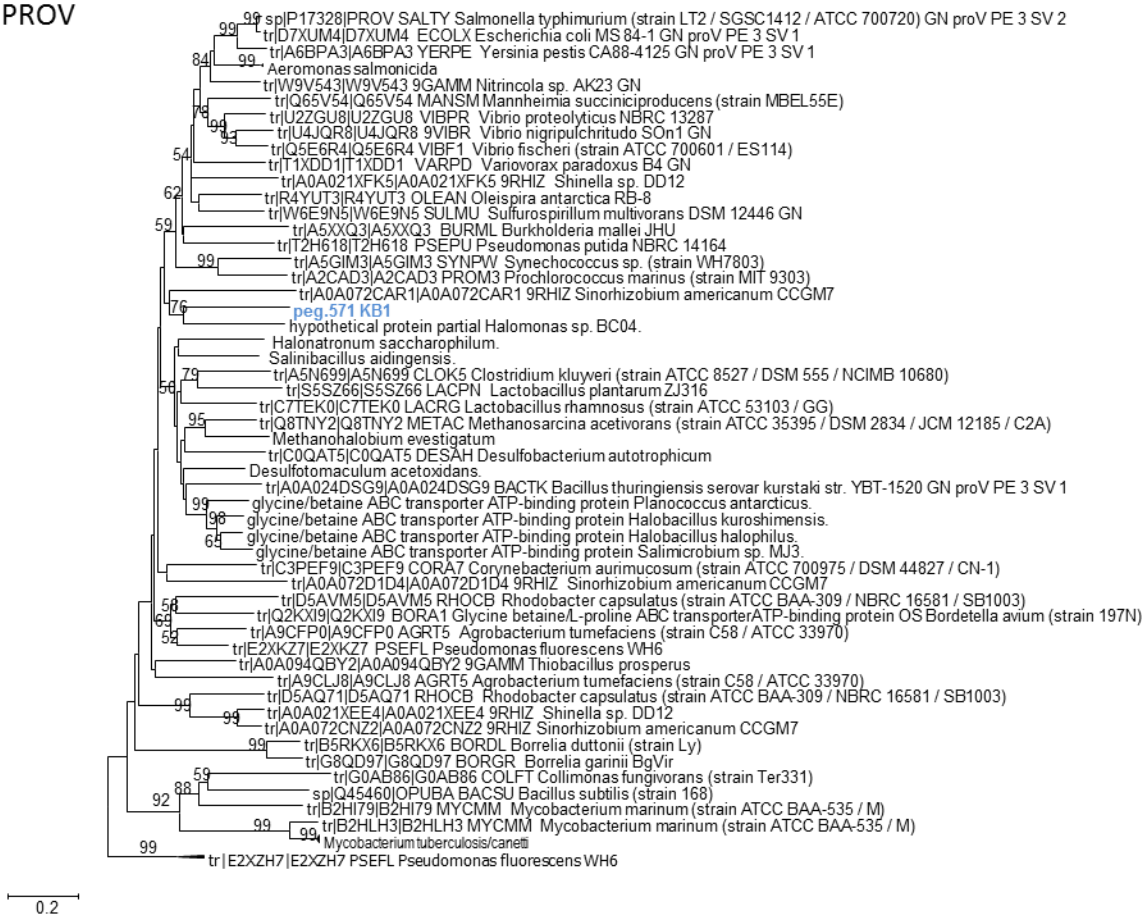



**Supplemental Figure (S7).** Relative abundances of Metanomicrobia-associated Illumina sequences downcore (left to right) of MUC4 (GB425) and MUC 5 (Hot site) areas. Abundance is displayed as percent of total prokaryotic-associated sequences. Note difference in % abundance scale in MUC 4(left) and MUC 5 (right).

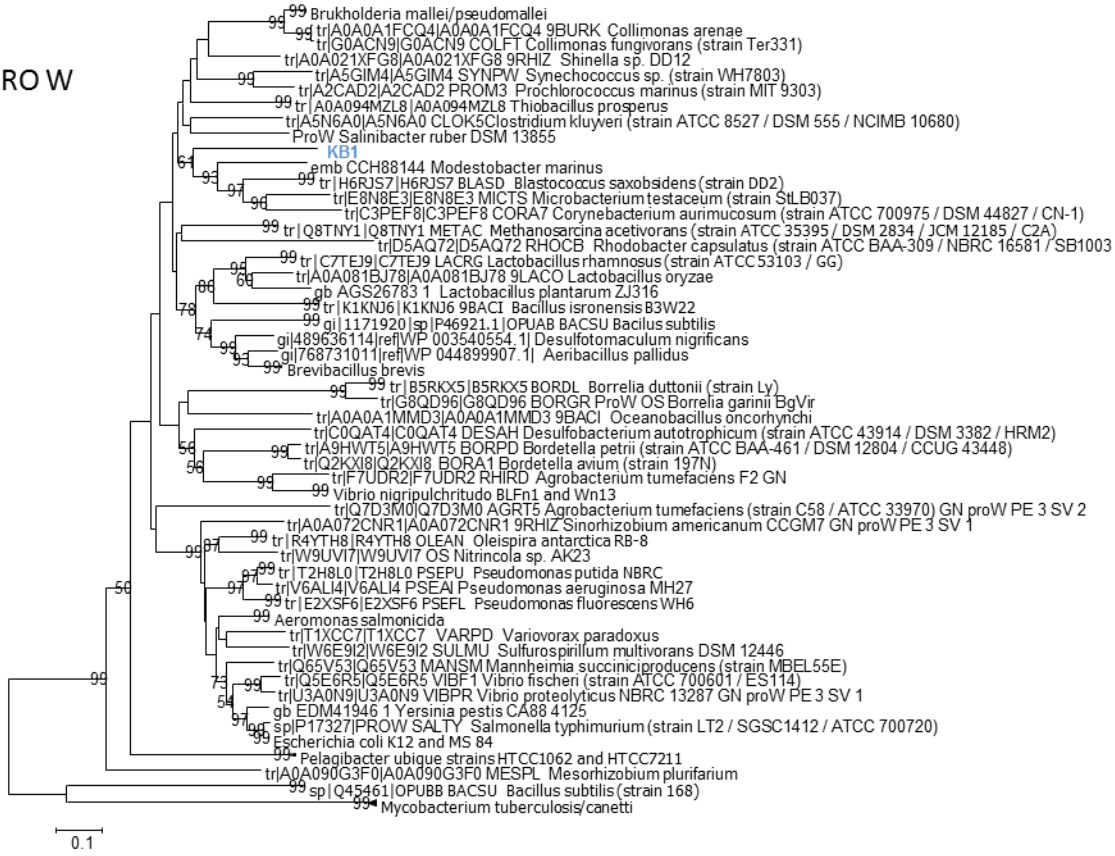
## APPENDIX B: SUPPLEMENTAL MATERIAL FOR CHAPTER 4

Phylogenetic trees of each of the PROU proteins are displayed below. Sequences from KB1 are in blue.

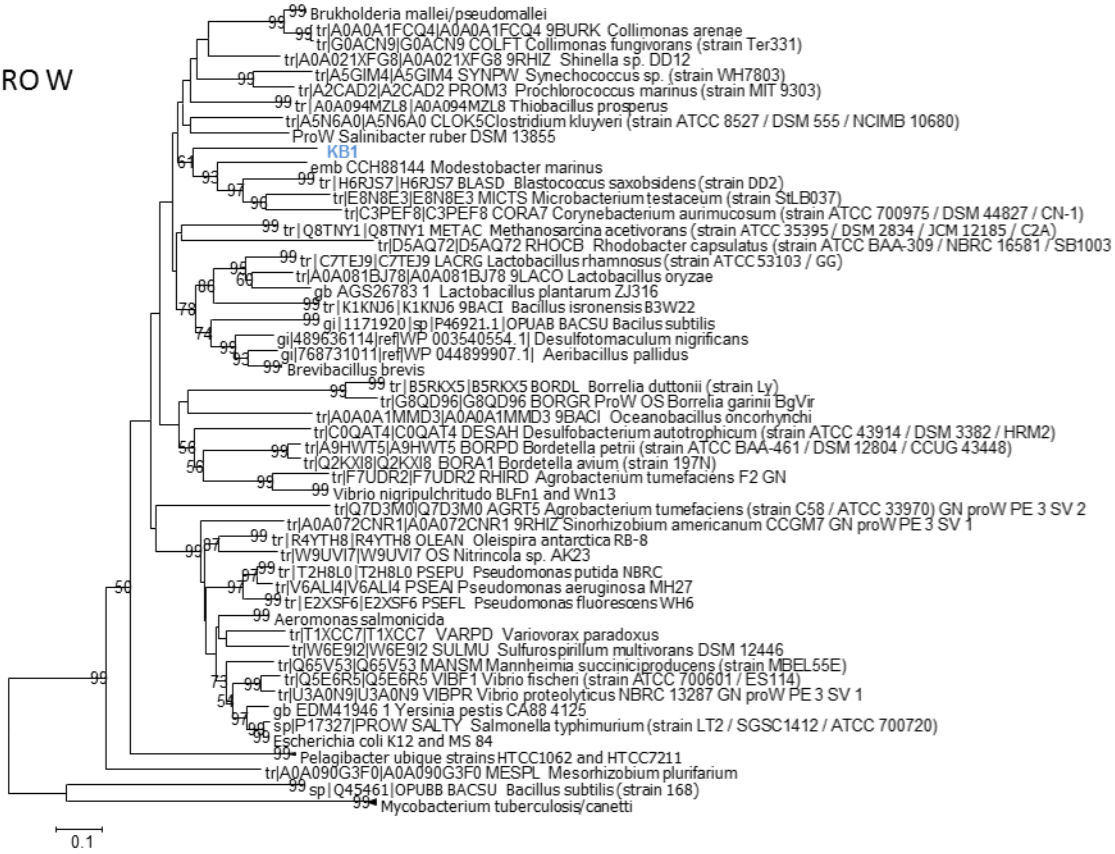
A. PROV



B. PRO W


Brukholderia mallei/pseudomallei
tr|A0A0A1FCQ4|A0A0A1FCQ4 9BURK Collimonas arenae
tr|G0ACN9|G0ACN9 COLFT Collimonas fungivorans (strain Ter331)
tr|A0A021XFG8|A0A021XFG8 9RHIZ Shinella sp. DD12
tr|A5GIM4|A5GIM4 SYNPW Synechococcus sp. (strain WH7803)
tr|A2CAD2|A2CAD2 PROM3 Prochlorococcus marinus (strain MIT 9303)
tr|A0A094MZL8|A0A094MZL8 Thiobacillus prosperus
tr|A5N6A0|A5N6A0 CLOK5Clostridium kluyveri (strain ATCC 8527 / DSM 555 / NCIMB 10680)
ProW Salinibacter ruber DSM 13855
KB1
emb CCH88144 Modestobacter marinus
tr|H6RJS7|H6RJS7 BLASD Blastococcus saxobsidens (strain DD2)
tr|E8N8E3|E8N8E3 MICTS Microbacterium testaceum (strain StLB037)
tr|C3PEF8|C3PEF8 CORA7 Corynebacterium aurimucosum (strain ATCC 700975 / DSM 44827 / CN-1)
tr|Q8TNY1|Q8TNY1 METAC Methanosarcina acetivorans (strain ATCC 35395 / DSM 2834 / JCM 12185 / C2A)
tr|D5AQ72|D5AQ72 RHOCB Rhodobacter capsulatus (strain ATCC BAA-309 / NBRC 16581 / SB1003
tr|C7TEJ9|C7TEJ9 LACRG Lactobacillus rhamnosus (strain ATCC 53103 / GG)
tr|A0A081BJ78|A0A081BJ78 9LACO Lactobacillus oryzae
gb AGS26783 1 Lactobacillus plantarum ZJ316
tr|K1KNJ6|K1KNJ6 9BACI Bacillus isronensis B3W22
gi|1171920|sp|P46921.1|OPUAB BACSU Bacilus subtilis
gi|489636114|ref|WP 003540554.1| Desulfotomaculum nigrificans
gi|768731011|ref|WP 044899907.1| Aeribacillus pallidus
Brevibacillus brevis
tr|B5RKX5|B5RKX5 BORDL Borrelia duttonii (strain Ly)
tr|G8QD96|G8QD96 BORGR ProW OS Borrelia garinii BgVir
tr|A0A0A1MMD3|A0A0A1MMD3 9BACI Oceanobacillus oncorhynchi
tr|C0QAT4|C0QAT4 DESAH Desulfobacterium autotrophicum (strain ATCC 43914 / DSM 3382 / HRM2)
tr|A9HWT5|A9HWT5 BORPD Bordetella petrii (strain ATCC BAA-461 / DSM 12804 / CCUG 43448)
tr|Q2KXI8|Q2KXI8 BORA1 Bordetella avium (strain 197N)
tr|F7UDR2|F7UDR2 RHIRD Agrobacterium tumefaciens F2 GN
Vibrio nigripulchritudo BLFn1 and Wn13
tr|Q7D3M0|Q7D3M0 AGRT5 Agrobacterium tumefaciens (strain C58 / ATCC 33970) GN proW PE 3 SV 2
tr|A0A072CNR1|A0A072CNR1 9RHIZ Sinorhizobium americanum CCGM7 GN proW PE 3 SV 1
tr|R4YTH8|R4YTH8 OLEAN Oleispira antarctica RB-8
tr|W9UVI7|W9UVI7 OS Nitrincola sp. AK23
tr|T2H8L0|T2H8L0 PSEPU Pseudomonas putida NBRC
tr|V6ALI4|V6ALI4 PSEAI Pseudomonas aeruginosa MH27
tr|E2XSF6|E2XSF6 PSEFL Pseudomonas fluorescens WH6
Aeromonas salmonicida
tr|T1XCC7|T1XCC7 VARPD Variovorax paradoxus
tr|W6E9I2|W6E9I2 SULMU Sulfurospirillum multivorans DSM 12446
tr|Q65V53|Q65V53 MANSM Mannheimia succiniciproducens (strain MBEL55E)
tr|Q5E6R5|Q5E6R5 VIBF1 Vibrio fischeri (strain ATCC 700601 / ES114)
tr|U3A0N9|U3A0N9 VIBPR Vibrio proteolyticus NBRC 13287 GN proW PE 3 SV 1
gb EDM41946 1 Yersinia pestis CA88 4125
sp|P17327|PROW SALTY Salmonella typhimurium (strain LT2 / SGSC1412 / ATCC 700720)
Escherichia coli K12 and MS 84
Pelagibacter ubique strains HTCC1062 and HTCC7211
tr|A0A090G3F0|A0A090G3F0 MESPL Mesorhizobium plurifarium
sp|Q45461|OPUBB BACSU Bacillus subtilis (strain 168)
Mycobacterium tuberculosis/canetti
0.1

B. PRO W


Brukholderia mallei/pseudomallei
tr|A0A0A1FCQ4|A0A0A1FCQ4 9BURK Collimonas arenae
tr|G0ACN9|G0ACN9 COLFT Collimonas fungivorans (strain Ter331)
tr|A0A021XFG8|A0A021XFG8 9RHIZ Shinella sp. DD12
tr|A5GIM4|A5GIM4 SYNPW Synechococcus sp. (strain WH7803)
tr|A2CAD2|A2CAD2 PROM3 Prochlorococcus marinus (strain MIT 9303)
tr|A0A094MZL8|A0A094MZL8 Thiobacillus prosperus
tr|A5N6A0|A5N6A0 CLOK5 Clostridium kluyveri (strain ATCC 8527 / DSM 555 / NCIMB 10680)
ProW Salinibacter ruber DSM 13855
KB1
emb CCH88144 Modestobacter marinus
tr|H6RJS7|H6RJS7 BLASD Blastococcus saxobsidens (strain DD2)
tr|E8N8E3|E8N8E3 MICTS Microbacterium testaceum (strain StLB037)
tr|C3PEF8|C3PEF8 CORA7 Corynebacterium aurimucosum (strain ATCC 700975 / DSM 44827 / CN-1)
tr|Q8TNY1|Q8TNY1 METAC Methanosarcina acetivorans (strain ATCC 35395 / DSM 2834 / JCM 12185 / C2A)
tr|D5AQ72|D5AQ72 RHOCB Rhodobacter capsulatus (strain ATCC BAA-309 / NBRC 16581 / SB1003
tr|C7TEJ9|C7TEJ9 LACRG Lactobacillus rhamnosus (strain ATCC 53103 / GG)
tr|A0A081BJ78|A0A081BJ78 9LACO Lactobacillus oryzae
gb AGS26783 1 Lactobacillus plantarum ZJ316
tr|K1KNJ6|K1KNJ6 9BACI Bacillus isronensis B3W22
gi|1171920|sp|P46921.1|OPUAB BACSU Bacilus subtilis
gi|489636114|ref|WP 003540554.1| Desulfotomaculum nigrificans
gi|768731011|ref|WP 044899907.1| Aeribacillus pallidus
Brevibacillus brevis
tr|B5RKX5|B5RKX5 BORDL Borrelia duttonii (strain Ly)
tr|G8QD96|G8QD96 BORGR ProW OS Borrelia garinii BgVir
tr|A0A0A1MMD3|A0A0A1MMD3 9BACI Oceanobacillus oncorhynchi
tr|C0QAT4|C0QAT4 DESAH Desulfobacterium autotrophicum (strain ATCC 43914 / DSM 3382 / HRM2)
tr|A9HWT5|A9HWT5 BORPD Bordetella petrii (strain ATCC BAA-461 / DSM 12804 / CCUG 43448)
tr|Q2KXI8|Q2KXI8 BORA1 Bordetella avium (strain 197N)
tr|F7UDR2|F7UDR2 RHIRD Agrobacterium tumefaciens F2 GN
Vibrio nigripulchritudo BLFn1 and Wn13
tr|Q7D3M0|Q7D3M0 AGRT5 Agrobacterium tumefaciens (strain C58 / ATCC 33970) GN proW PE 3 SV 2
tr|A0A072CNR1|A0A072CNR1 9RHIZ Sinorhizobium americanum CCGM7 GN proW PE 3 SV 1
tr|R4YTH8|R4YTH8 OLEAN Oleispira antarctica RB-8
tr|W9UVI7|W9UVI7 OS Nitrincola sp. AK23
tr|T2H8L0|T2H8L0 PSEPU Pseudomonas putida NBRC
tr|V6ALI4|V6ALI4 PSEAI Pseudomonas aeruginosa MH27
tr|E2XSF6|E2XSF6 PSEFL Pseudomonas fluorescens WH6
Aeromonas salmonicida
tr|T1XCC7|T1XCC7 VARPD Variovorax paradoxus
tr|W6E9I2|W6E9I2 SULMU Sulfurospirillum multivorans DSM 12446
tr|Q65V53|Q65V53 MANSM Mannheimia succiniciproducens (strain MBEL55E)
tr|Q5E6R5|Q5E6R5 VIBF1 Vibrio fischeri (strain ATCC 700601 / ES114)
tr|U3A0N9|U3A0N9 VIBPR Vibrio proteolyticus NBRC 13287 GN proW PE 3 SV 1
gb EDM41946 1 Yersinia pestis CA88 4125
sp|P17327|PROW SALTY Salmonella typhimurium (strain LT2 / SGSC1412 / ATCC 700720)
Escherichia coli K12 and MS 84
Pelagibacter ubique strains HTCC1062 and HTCC7211
tr|A0A090G3F0|A0A090G3F0 MESPL Mesorhizobium plurifarium
sp|Q45461|OPUBB BACSU Bacillus subtilis (strain 168)
Mycobacterium tuberculosis/canetti
0.1

## REFERENCES


Aharon, P., Roberts, H.H., and Snelling, R. (1992). Submarine venting of brines in the deep Gulf of Mexico: Observations and geochemistry. Geology *20*, 483–486.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Antón, J., Oren, A., Benlloch, S., Rodríguez-Valera, F., Amann, R., and Rosselló-Mora, R. (2002). Salinibacter ruber gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria from saltern crystallizer ponds. Int. J. Syst. Evol. Microbiol. *52*, 485–491.

Antunes, A., Ngugi, D.K., and Stingl, U. (2011). Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. Environ. Microbiol. Rep.

Aronesty, E. (2011). ea-utils: Command-line tools for processing biological sequencing data.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., et al. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics *9*, 75.

Baker, B.J., Tyson, G.W., Webb, R.I., Flanagan, J., Hugenholtz, P., Allen, E.E., and Banfield, J.F. (2006). Lineages of acidophilic archaea revealed by community genomic analysis. Science *314*, 1933.

Belyakova, E.V., Rozanova, E.P., Borzenkov, I.A., Tourova, T.P., Pusheva, M.A., Lysenko, A.M., and Kolganova, T.V. (2006). The new facultatively chemolithoautotrophic, moderately halophilic, sulfate-reducing bacterium Desulfovermiculus halophilus gen. nov., sp. nov., isolated from an oil field. Microbiology *75*, 161–171.

Benson, K.R. (2000). The emergence of ecology from natural history. Endeavour *24*, 59–62.

Boere, A.C., Sinninghe Damsté, J.S., Rijpstra, W.I.C., Volkman, J.K., and Coolen, M.J.L. (2011). Source-specific variability in post-depositional DNA preservation with potential implications for DNA based paleoecological records. Org. Geochem. *42*, 1216–1225.

Boetius, A., and Joye, S. (2009). Thriving in Salt. Science *324*, 1523–1525.

Borin, S., Brusetti, L., Mapelli, F., D'Auria, G., Brusa, T., Marzorati, M., Rizzi, A., Yakimov, M., Marty, D., De Lange, G.J., et al. (2009). Sulfur cycling and methanogenesis primarily drive microbial colonization of the highly sulfidic Urania deep hypersaline basin. Proc. Natl. Acad. Sci. *106*, 9151.

Brandt, K.K., Patel, B.K.C., and Ingvorsen, K. (1999). Desulfocella halophila gen. nov., sp. nov., a halophilic, fatty-acid-oxidizing, sulfate-reducing bacterium isolated from sediments of the Great Salt Lake. Int. J. Syst. Bacteriol. *49*, 193–200.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., and Gordon, J.I. (2010a). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., and Knight, R. (2010b). PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics *26*, 266–267.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. *6*, 1621–1624.

Van Cappellen, P., Viollier, E., Roychoudhury, A., Clark, L., Ingall, E., Lowe, K., and Dichristina, T. (1998). Biogeochemical cycles of manganese and iron at the oxic-anoxic transition of a stratified marine basin (Orca Basin, Gulf of Mexico). Environ. Sci. Technol. *32*, 2931–2939.

Cavicchioli, R. (2002). Extremophiles and the Search for Extraterrestrial Life. Astrobiology *2*, 281–292.

Cline, J.D. (1969). Spectrophotometric Determination of Hydrogen Sulfide in Natural Waters1. Limnol. Oceanogr. *14*, 454–458.

La Cono, V., Smedile, F., Bortoluzzi, G., Arcadi, E., Maimone, G., Messina, E., Borghini, M., Oliveri, E., Mazzola, S., L'Haridon, S., et al. (2011). Unveiling microbial life in new deep-sea hypersaline Lake Thetis. Part I: Prokaryotes and environmental settings. Environ. Microbiol. *13*, 2250–2268.

Daffonchio, D., Borin, S., Brusa, T., Brusetti, L., van der Wielen, P.W.J.J., Bolhuis, H., Yakimov, M.M., D'Auria, G., Giuliano, L., Marty, D., et al. (2006). Stratified prokaryote network in the oxic–anoxic transition of a deep-sea halocline. Nature *440*, 203–207.

Dickins, H.D., and Van Vleet, E.S. (1992). Archaebacterial activity in the Orca Basin determined by the isolation of characteristic isopranyl ether-linked lipids. Deep Sea Res. Part Oceanogr. Res. Pap. *39*, 521–536.

La Duc, M.T., Osman, S., Vaishampayan, P., Piceno, Y., Andersen, G., Spry, J.A., and Venkateswaran, K. (2009). Comprehensive census of bacteria in clean rooms by using DNA microarray and cloning methods. Appl. Environ. Microbiol. *75*, 6559–6567.

Eder, W., Ludwig, W., and Huber, R. (1999). Novel 16S rRNA gene sequences retrieved from highly saline brine sediments of Kebrit Deep, Red Sea. Arch. Microbiol. *172*, 213–218.

Eder, W., Jahnke, L.L., Schmidt, M., and Huber, R. (2001). Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. Appl. Environ. Microbiol. *67*, 3077.

Eder, W., Schmidt, M., Koch, M., Garbe-Schönberg, D., and Huber, R. (2002). Prokaryotic phylogenetic diversity and corresponding geochemical data of the brine–seawater interface of the Shaban Deep, Red Sea. Environ. Microbiol. *4*, 758–763.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods *10*, 996–998.

Faith, D.P. (2006). The role of the phylogenetic diversity measure, PD, in bio-informatics: getting the definition right. Evol. Bioinforma. Online *2*, 277.

Gilbert, J., Meyer, F., Jansson, J., Gordon, J., Pace, N., Tiedje, J., Ley, R., Fierer, N., Field, D., Kyrpides, N.C., et al. (2010). The Earth Microbiome Project: Meeting report of the "1st EMP meeting on sample selection and acquisition" at Argonne National Laboratory October 6th 2010. Stand. Genomic Sci. *3*, 249–253.

Hamady, M., Lozupone, C., and Knight, R. (2009). Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. ISME J. *4*, 17–27.

Harvey, H.R., Kennicutt, M.C., and others (1992). Selective alteration of Sargassum lipids in anoxic sediments of the Orca Basin. Org. Geochem. *18*, 181–187.

Hattori, S., Galushko, A.S., Kamagata, Y., and Schink, B. (2005). Operation of the CO Dehydrogenase/Acetyl Coenzyme A Pathway in both Acetate Oxidation and Acetate Formation by the Syntrophically Acetate-Oxidizing Bacterium Thermacetogenium phaeum. J. Bacteriol. *187*, 3471–3476.

Hazen, T.C., Dubinsky, E.A., DeSantis, T.Z., Andersen, G.L., Piceno, Y.M., Singh, N., Jansson, J.K., Probst, A., Borglin, S.E., Fortney, J.L., et al. (2010). Deep-Sea Oil Plume Enriches Indigenous Oil-Degrading Bacteria. Science *330*, 204–208.

Heijthuijsen, J.H.F.G., and Hansen, T.A. (1989). Betaine Fermentation and Oxidation by Marine Desulfuromonas Strains. Appl. Environ. Microbiol. *55*, 965–969.

Hugenholtz, P., Pitulle, C., Hershberger, K.L., and Pace, N.R. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. J. Bacteriol. *180*, 366–376.

Hurtgen, M.T., Lyons, T.W., Ingall, E.D., and Cruse, A.M. (1999). Anomalous enrichments of iron monosulfide in euxinic marine sediments and the role of H2S in iron sulfide transformations: Examples from Effingham Inlet, Orca Basin, and the Black Sea. Am. J. Sci. *299*, 556–588.

Isenbarger, T.A., Finney, M., Rios-Velazquez, C., Handelsman, J., and Ruvkun, G. (2007). Miniprimer PCR, a New Lens for Viewing the Microbial World. Appl. Environ. Microbiol. *74*, 840–849.

Jakobsen, T.F. (2006). Desulfohalobium utahense sp. nov., a moderately halophilic, sulfate-reducing bacterium isolated from Great Salt Lake. Int. J. Syst. Evol. Microbiol. *56*, 2063–2069.

Joye, S.B. (2010). Gulf of Mexico Seafloor Brine Ecosystems.

Joye, S.B., Boetius, A., Orcutt, B.N., Montoya, J.P., Schulz, H.N., Erickson, M.J., and Lugo, S.K. (2004). The anaerobic oxidation of methane and sulfate reduction in sediments from Gulf of Mexico cold seeps. Chem. Geol. *205*, 219–238.

Joye, S.B., MacDonald, I.R., Montoya, J.P., and Peccini, M. (2005). Geophysical and geochemical signatures of Gulf of Mexico seafloor brines. Biogeosciences Discuss. *2*, 637–671.

Jukes, T., and Cantor, C. (1969). {Evolution of protein molecules}. In Mammalian Protein Metabolism, M. Munro, ed. (Academic Press), pp. 21–132.

Kadam, P.C., and Boone, D.R. (1996). Influence of pH on Ammonia Accumulation and Toxicity in Halophilic, Methylotrophic Methanogens. Appl. Environ. Microbiol. *62*, 4486–4492.

Kamanda Ngugi, D., Blom, J., Alam, I., Rashid, M., Ba-Alawi, W., Zhang, G., Hikmawan, T., Guan, Y., Antunes, A., Siam, R., et al. (2015). Comparative genomics reveals adaptations of a halotolerant thaumarchaeon in the interfaces of brine pools in the Red Sea. ISME J. *9*, 396–411.

Kessler, J.D., Valentine, D.L., Redmond, M.C., Du, M., Chan, E.W., Mendes, S.D., Quiroz, E.W., Villanueva, C.J., Shusta, S.S., Werra, L.M., et al. (2011). A Persistent Oxygen Anomaly Reveals the Fate of Spilled Methane in the Deep Gulf of Mexico. Science.

Kimes, N.E., Callaghan, A.V., Aktas, D.F., Smith, W.L., Sunner, J., Golding, B., Drozdowska, M., Hazen, T.C., Suflita, J.M., and Morris, P.J. (2013). Metagenomic analysis and metabolite profiling of deep–sea sediments from the Gulf of Mexico following the Deepwater Horizon oil spill. Front. Microbiol. *4*.

King, G.M. (1984). Metabolism of Trimethylamine, Choline, and Glycine Betaine by Sulfate-Reducing and Methanogenic Bacteria in Marine Sediments. Appl. Environ. Microbiol. *48*, 719–725.

Kirk Harris, J., Gregory Caporaso, J., Walker, J.J., Spear, J.R., Gold, N.J., Robertson, C.E., Hugenholtz, P., Goodrich, J., McDonald, D., Knights, D., et al. (2012). Phylogenetic stratigraphy in the Guerrero Negro hypersaline microbial mat. ISME J.

Kleindienst, S., Ramette, A., Amann, R., and Knittel, K. (2012). Distribution and in situ abundance of sulfate-reducing bacteria in diverse marine hydrocarbon seep sediments. Environ. Microbiol. *14*, 2689–2710.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. gks808.

Knittel, K., and Boetius, A. (2009). Anaerobic Oxidation of Methane: Progress with an Unknown Process. Annu. Rev. Microbiol. *63*, 311–334.

Knittel, K., Boetius, A., Lemke, A., Eilers, H., Lochte, K., Pfannkuche, O., Linke, P., and Amann, R. (2003). Activity, Distribution, and Diversity of Sulfate Reducers and Other Bacteria in Sediments above Gas Hydrate (Cascadia Margin, Oregon). Geomicrobiol. J. *20*, 269–294.

LaRock, P.A., Lauer, R.D., Schwarz, J.R., Watanabe, K.K., and Wiesenburg, D.A. (1979). Microbial biomass and activity distribution in an anoxic, hypersaline basin. Appl. Environ. Microbiol. *37*, 466.

Lee, M., Woo, S.-G., and Ten, L.N. (2012). Characterization of novel diesel-degrading strains Acinetobacter haemolyticus MJ01 and Acinetobacter johnsonii MJ4 isolated from oil-contaminated soil. World J. Microbiol. Biotechnol. *28*, 2057–2067.

Ley, R.E., Harris, J.K., Wilcox, J., Spear, J.R., Miller, S.R., Bebout, B.M., Maresca, J.A., Bryant, D.A., Sogin, M.L., and Pace, N.R. (2006). Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. Appl. Environ. Microbiol. *72*, 3685–3695.

Liu, Y., and Whitman, W.B. (2008). Metabolic, Phylogenetic, and Ecological Diversity of the Methanogenic Archaea. Ann. N. Y. Acad. Sci. *1125*, 171–189.

Liu, Y., Boone, D.R., and Choy, C. (1990). Methanohalophilus oregonense sp. nov., a Methylotrophic Methanogen from an Alkaline, Saline Aquifer. Int. J. Syst. Bacteriol. *40*, 111–116.

Liu, Z., Liu, J., Zhu, Q., and Wu, W. (2012). The weathering of oil after the Deepwater Horizon oil spill: insights from the chemical composition of the oil from the sea surface, salt marshes and sediments. Environ. Res. Lett. *7*, 035302.

Lloyd, K.G., Lapham, L., and Teske, A. (2006a). An Anaerobic Methane-Oxidizing Community of ANME-1b Archaea in Hypersaline Gulf of Mexico Sediments. Appl. Environ. Microbiol. *72*, 7218–7230.

Lloyd, K.G., Lapham, L., and Teske, A. (2006b). An Anaerobic Methane-Oxidizing Community of ANME-1b Archaea in Hypersaline Gulf of Mexico Sediments. Appl. Environ. Microbiol. *72*, 7218–7230.

Lloyd, K.G., Albert, D.B., Biddle, J.F., Chanton, J.P., Pizarro, O., and Teske, A. (2010). Spatial Structure and Activity of Sedimentary Microbial Communities Underlying a Beggiatoa spp. Mat in a Gulf of Mexico Hydrocarbon Seep. PLoS ONE *5*, e8738.

Lozupone, C., and Knight, R. (2005). UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. Appl. Environ. Microbiol. *71*, 8228–8235.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Buchner, A., Lai, T., Steppi, S., Jobb, G., and Förster, W. (2004). ARB: a software environment for sequence data. Nucleic Acids Res. *32*, 1363–1371.

Lutken, C. (2013). HYDRATE RESEARCH ACTIVITIES THAT BOTH SUPPORT AND DERIVE FROM THE MONITORING STATION/SEA-FLOOR OBSERVATORY, MISSISSIPPI CANYON 118, NORTHERN GULF OF MEXICO.

MacDonald, I.R., Buthman, D.B., Sager, W.W., Peccini, M.B., and Guinasso, N.L. (2000). Pulsed oil discharge from a mud volcano. Geology *28*, 907–910.

MacLean, D., Jones, J.D.G., and Studholme, D.J. (2009). Application of "next-generation" sequencing technologies to microbial genetics. Nat. Rev. Microbiol. *7*, 287–296.

Maignien, L., Parkes, R.J., Cragg, B., Niemann, H., Knittel, K., Coulon, S., Akhmetzhanov, A., and Boon, N. (2013). Anaerobic oxidation of methane in hypersaline cold seep sediments. FEMS Microbiol. Ecol. *83*, 214–231.

Makhdoumi-Kakhki, A., Amoozegar, M.A., Kazemi, B., PaiC, L., and Ventosa, A. (2012). Prokaryotic Diversity in Aran-Bidgol Salt Lake, the Largest Hypersaline Playa in Iran. Microbes Environ. *27*, 87–93.

Mason, O.U., Scott, N.M., Gonzalez, A., Robbins-Pianka, A., Bælum, J., Kimbrel, J., Bouskill, N.J., Prestat, E., Borglin, S., Joyner, D.C., et al. (2014). Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill. ISME J. *8*, 1464–1475.

Mathrani, I.M., Boone, D.R., Mah, R.A., Fox, G.E., and Lau, P.P. (1988). Methanohalophilus zhilinae sp. nov., an Alkaliphilic, Halophilic, Methylotrophic Methanogen. Int. J. Syst. Bacteriol. *38*, 139–142.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6*, 610–618.

McGee, T., Macelloni, L., Lutken, C., Bosman, A., Brunner, C., Rogers, R., Dearman, J., Sleeper, K., and Woolsey, J.R. (2009). Hydrocarbon gas hydrates in sediments of the Mississippi Canyon area, Northern Gulf of Mexico. Geol. Soc. Lond. Spec. Publ. *319*, 29–49.

McNutt, M.K., Camilli, R., Crone, T.J., Guthrie, G.D., Hsieh, P.A., Ryerson, T.B., Savas, O., and Shaffer, F. (2012). Review of flow rate estimates of the Deepwater Horizon oil spill. Proc. Natl. Acad. Sci. *109*, 20260–20267.

Mills, H.J., Martinez, R.J., Story, S., and Sobecky, P.A. (2005). Characterization of Microbial Community Structure in Gulf of Mexico Gas Hydrates: Comparative Analysis of DNA- and RNA-Derived Clone Libraries. Appl. Environ. Microbiol. *71*, 3235–3247.

Möller, B., Oßmer, R., Howard, B.H., Gottschalk, G., and Hippe, H. (1984). Sporomusa, a new genus of gram-negative anaerobic bacteria including Sporomusa sphaeroides spec. nov. and Sporomusa ovata spec. nov. Arch. Microbiol. *139*, 388–396.

Mongodin, E.F., Nelson, K.E., Daugherty, S., Deboy, R.T., Wister, J., Khouri, H., Weidman, J., Walsh, D.A., Papke, R.T., and Perez, G.S. (2005). The genome of Salinibacter ruber: convergence and gene exchange among hyperhalophilic bacteria and archaea. Proc. Natl. Acad. Sci. U. S. A. *102*, 18147–18152.

Müller, E., Fahlbusch, K., Walther, R., and Gottschalk, G. (1981). Formation of N,N-Dimethylglycine, Acetic Acid, and Butyric Acid from Betaine by Eubacterium limosum. Appl. Environ. Microbiol. *42*, 439–445.

Murray, G.E. (1966). Salt Structures of Gulf of Mexico Basin--A Review. AAPG Bull. *50*, 439–478.

Na, H., Lever, M.A., Kjeldsen, K.U., Schulz, F., and Jørgensen, B.B. (2015). Uncultured Desulfobacteraceae and Crenarchaeotal group C3 incorporate 13C-acetate in coastal marine sediment. Environ. Microbiol. Rep. n/a – n/a.

Naumann, E., Hippe, H., and Gottschalk, G. (1983). Betaine: New Oxidant in the Stickland Reaction and Methanogenesis from Betaine and l-Alanine by a Clostridium sporogenes-Methanosarcina barkeri Coculture. Appl. Environ. Microbiol. *45*, 474–483.

Nobu, M.K., Dodsworth, J.A., Murugapiran, S.K., Rinke, C., Gies, E.A., Webster, G., Schwientek, P., Kille, P., Parkes, R.J., Sass, H., et al. (2015). Phylogeny and physiology of candidate phylum “Atribacteria” (OP9/JS1) inferred from cultivation-independent genomics. ISME J.

Offre, P., Spang, A., and Schleper, C. (2013). Archaea in Biogeochemical Cycles. Annu. Rev. Microbiol. *67*, 437–457.

Ollivier, B., Hatchikian, C.E., Prensier, G., Guezennec, J., and Garcia, J.-L. (1991). Desulfohalobium retbaense gen. nov., sp. nov., a Halophilic Sulfate-Reducing Bacterium from Sediments of a Hypersaline Lake in Senegal. Int. J. Syst. Bacteriol. *41*, 74–81.

Olsen, G.J., Woese, C.R., and Overbeek, R. (1994). The winds of (evolutionary) change: breathing new life into microbiology. J. Bacteriol. *176*, 1–6.

Orcutt, B.N., Joye, S.B., Kleindienst, S., Knittel, K., Ramette, A., Reitz, A., Samarkin, V., Treude, T., and Boetius, A. (2010). Impact of natural oil and higher hydrocarbons on microbial diversity, distribution, and activity in Gulf of Mexico cold-seep sediments. Deep Sea Res. Part II Top. Stud. Oceanogr. *57*, 2008–2021.

Oren, A. (1999). Bioenergetic aspects of halophilism. Microbiol. Mol. Biol. Rev. *63*, 334–348.

Oren, A. (2005). Halophilic Microorganisms: Physiology and Phylogeny. Origins 413–426.

Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Syst. *4*, 2.

Pace, N.R. (1997). A Molecular View of Microbial Diversity and the Biosphere. Science *276*, 734–740.

Passow, U., Ziervogel, K., Asper, V., and Diercks, A. (2012). Marine snow formation in the aftermath of the Deepwater Horizon oil spill in the Gulf of Mexico. Environ. Res. Lett. *7*, 035301.

Paterek, J.R., and Smith, P.H. (1985). Isolation and Characterization of a Halophilic Methanogen from Great Salt Lake. Appl. Environ. Microbiol. *50*, 877–881.

Pester, M., Schleper, C., and Wagner, M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. Curr. Opin. Microbiol. *14*, 300–306.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Mol. Biol. Evol. *26*, 1641–1650.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. *35*, 7188–7196.

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

Rappé, M.S., and Giovannoni, S.J. (2003). The Uncultured Microbial Majority. Annu. Rev. Microbiol. *57*, 369–394.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.-F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. Nature *499*, 431–437.

Ruff, S.E., Biddle, J.F., Teske, A.P., Knittel, K., Boetius, A., and Ramette, A. (2015). Global dispersion and local diversification of the methane seep microbiome. Proc. Natl. Acad. Sci. *112*, 4015–4020.

Ryerson, T.B., Camilli, R., Kessler, J.D., Kujawinski, E.B., Reddy, C.M., Valentine, D.L., Atlas, E., Blake, D.R., Gouw, J. de, Meinardi, S., et al. (2012). Chemical data quantify Deepwater Horizon hydrocarbon flow rate and environmental distribution. Proc. Natl. Acad. Sci. *109*, 20246–20253.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. *74*, 5463–5467.

Schnell, S., Bak, F., and Pfennig, N. (1989). Anaerobic degradation of aniline and dihydroxybenzenes by newly isolated sulfate-reducing bacteria and description of Desulfobacterium anilini. Arch. Microbiol. *152*, 556–563.

Schuster, S.C. (2008). Next-generation sequencing transforms today's biology. Nat. Methods *5*, 16–18.

Shannon, C.E. (1948). Bell System Tech. J. 27 (1948) 379; CE Shannon. Bell Syst. Tech J *27*, 623.

Sheu, D. (1990). The Anoxic Orca Basin (Gulf of Mexico): Geochemistry of Brines and Sediments. Aquat. Sci. *2*, 491–507.

Sheu, D.D., and Presley, B.J. (1986). Formation of hematite in the euxinic Orca Basin, northern Gulf of Mexico. Mar. Geol. *69*, 309–321.

Shokes, R.F., Trabant, P.K., Presley, B.J., and Reid, D.F. (1977). Anoxic, Hypersaline Basin in the Northern Gulf of Mexico. Science *196*, 1443–1446.

Siam, R., Mustafa, G.A., Sharaf, H., Moustafa, A., Ramadan, A.R., Antunes, A., Bajic, V.B., Stingl, U., Marsis, N.G.R., Coolen, M.J.L., et al. (2012). Unique Prokaryotic Consortia in Geochemically Distinct Sediments from Red Sea Atlantis II and Discovery Deep Brine Pools. PLoS ONE *7*, e42872.

Singh, B.K., Bardgett, R.D., Smith, P., and Reay, D.S. (2010). Microorganisms and climate change: terrestrial feedbacks and mitigation options. Nat. Rev. Microbiol. *8*, 779–790.

Solórzano, L. (1969). DETERMINATION OF AMMONIA IN NATURAL WATERS BY THE PHENOLHYPOCHLORITE METHOD 1 1 This research was fully supported by U.S. Atomic Energy Commission Contract No. ATS (11-1) GEN 10, P.A. 20. Limnol. Oceanogr. *14*, 799–801.

Spring, S., Nolan, M., Lapidus, A., Glavina Del Rio, T., Copeland, A., Tice, H., Cheng, J.-F., Lucas, S., Land, M., Chen, F., et al. (2010). Complete genome sequence of Desulfohalobium retbaense type strain (HR100T). Stand. Genomic Sci. *2*, 38–48.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688–2690.

Stepanauskas, R. (2012). Single cell genomics: an individual look at microbes. Curr. Opin. Microbiol. *15*, 613–620.

Stewart, E.J. (2012). Growing Unculturable Bacteria. J. Bacteriol. *194*, 4151–4160.

Suzuki, D., Li, Z., Cui, X., Zhang, C., and Katayama, A. (2014). Reclassification of Desulfobacterium anilini as Desulfatiglans anilini comb. nov. within Desulfatiglans gen. nov., and description of a 4-chlorophenol-degrading sulfate-reducing bacterium, Desulfatiglans parachlorophenolica sp. nov. Int. J. Syst. Evol. Microbiol. *64*, 3081–3086.

Takami, H., Noguchi, H., Takaki, Y., Uchiyama, I., Toyoda, A., Nishi, S., Chee, G.-J., Arai, W., Nunoura, T., Itoh, T., et al. (2012). A Deeply Branching Thermophilic Bacterium with an Ancient Acetyl-CoA Pathway Dominates a Subsurface Ecosystem. PLoS ONE *7*, e30559.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. *30*, 2725–2729.

Tavormina, P.L., Hatzenpichler, R., McGlynn, S., Chadwick, G., Dawson, K.S., Connon, S.A., and Orphan, V.J. (2015). Methyloprofundus sedimenti gen. nov., sp. nov., an obligate methanotroph from ocean sediment belonging to the "deep sea-1" clade of marine methanotrophs. Int. J. Syst. Evol. Microbiol. *65*, 251–259.

Ticak, T., Kountz, D.J., Girosky, K.E., Krzycki, J.A., and Ferguson, D.J. (2014). A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. Proc. Natl. Acad. Sci. *111*, E4668–E4676.

Tkavc, R., Gostinčar, C., Turk, M., Visscher, P.T., Oren, A., and Gunde-Cimerman, N. (2011). Bacterial communities in the "petola" microbial mat from the Sečovlje salterns (Slovenia). FEMS Microbiol. Ecol. *75*, 48–62.

Trefry, J.H., Presley, B.J., Keeney-Kennicutt, W.L., and Trocine, R.P. (1984). Distribution and chemistry of manganese, iron, and suspended particulates in Orca Basin. Geo-Mar. Lett. *4*, 125–130.

Tribovillard, N., Bout-Roumazeilles, V., Algeo, T., Lyons, T.W., Sionneau, T., Montero-Serrano, J.C., Riboulleau, A., and Baudin, F. (2008). Paleodepositional conditions in the Orca Basin as inferred from organic matter and trace metal contents. Mar. Geol. *254*, 62–72.

Tribovillard, N., Bout-Roumazeilles, V., Sionneau, T., Serrano, J.C.., Riboulleau, A., and Baudin, F. (2009). Does a strong pycnocline impact organic-matter preservation and accumulation in an anoxic setting? The case of the Orca Basin, Gulf of Mexico. Comptes Rendus Geosci. *341*, 1–9.

Vigneron, A., Cruaud, P., Pignet, P., Caprais, J.-C., Gayet, N., Cambon-Bonavita, M.-A., Godfroy, A., and Toffin, L. (2014). Bacterial communities and syntrophic associations involved in anaerobic oxidation of methane process of the Sonora Margin cold seeps, Guaymas Basin. Environ. Microbiol. n/a – n/a.

Wankel, S.D., Joye, S.B., Samarkin, V.A., Shah, S.R., Friederich, G., Melas-Kyriazi, J., and Girguis, P.R. (2010). New constraints on methane fluxes and rates of anaerobic methane oxidation in a Gulf of Mexico brine pool via in situ mass spectrometry. Deep Sea Res. Part II Top. Stud. Oceanogr. *57*, 2022–2029.

Watanabe, K. (2001). Microorganisms relevant to bioremediation. Curr. Opin. Biotechnol. *12*, 237–241.

Watkins, A.J., Roussel, E.G., Parkes, R.J., and Sass, H. (2014). Glycine Betaine as a Direct Substrate for Methanogens (Methanococcoides spp.). Appl. Environ. Microbiol. *80*, 289–293.

Webster, G., Parkes, R.J., Fry, J.C., and Weightman, A.J. (2004). Widespread Occurrence of a Novel Division of Bacteria Identified by 16S rRNA Gene Sequences Originally Found in Deep Marine Sediments. Appl. Environ. Microbiol. *70*, 5708–5713.

Van der Wielen, P.W.J.J. (2005). The Enigma of Prokaryotic Life in Deep Hypersaline Anoxic Basins. Science *307*, 121–123.

Woese, C.R. (1987). Bacterial evolution. Microbiol. Rev. *51*, 221–271.

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

Yakimov, M.M., La Cono, V., Slepak, V.Z., La Spada, G., Arcadi, E., Messina, E., Borghini, M., Monticelli, L.S., Rojo, D., Barbas, C., et al. (2013). Microbial life in the Lake Medee, the largest deep-sea salt-saturated formation. Sci. Rep. *3*.