BACTERIAL COMMUNITIES ASSOCIATED WITH NATURAL AND ANTHROPOGENIC PETROLEUM IMPACTED ANAEROBIC ENVIRONMENTS

By

JAMES PAUL DAVIS

Bachelor of Science in Microbiology
University of Oklahoma
Norman, Oklahoma
2006

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2012

BACTERIAL COMMUNITIES ASSOCIATED WITH NATURAL AND ANTHROPOGENIC PETROLEUM IMPACTED ANAEROBIC ENVIRONMENTS

Dissertation Approved:

Dr. Mostafa S. Elshahed

Dissertation Adviser

Dr. Robert V. Miller

Dr. Wouter D. Hoff

Dr. Babu Z. Fathepure

Dr. Ulrich Melcher

Outside Committee Member

Dr. Sheryl A. Tucker

Dean of the Graduate College

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

Preface

Advances in microbial ecology in the past three decades have shown that the size and diversity of bacterial communities within various environments are extensive. Techniques centered on the use of the 16S rRNA gene as a phylogenetic marker in bacterial diversity surveys, have allowed for a depiction of the distribution, diversity, phylogeny, and abundance of various microbial populations in every type of habitat. The use of culture-independent molecular biological approaches for diversity analysis of bacterial communities is at the core of this dissertation. Understanding the composition and size of microbial populations is the first step towards revealing the role of specific microbial groups in nutrient cycling, ecosystem productivity, and shaping the environment itself. This dissertation focuses on microbial populations associated with petroleum (specifically natural gas) environments and is composed of two distinct research projects. The first project (Chapter II) is an investigation of a specific uncultured bacterial lineage in a petroleum-impacted spring in southwest Oklahoma. The second project (Chapter III) was a broad-scale bacterial survey of natural gas production wells from north-central Texas.

Chapter I was written to provide the pertinent information for an introduction to the general topic of Molecular Microbial Ecology, as well as other topics covered in this dissertation. It is not an exhaustive review of literature. This field is rapidly changing, and so any discourse on the current state of the research of the time would likely be outdated within a few years. For example, at the beginning of my tenure as a graduate student in 2007, Sanger sequencing of 16S rRNA gene clone libraries was the most popular method for studying a bacterial population. In 2012, direct sequencing of PCR amplicons using next-generation sequencing technologies (*e.g.* 454-pyrosequencing, Illumina) have become the

norm due to the lower cost and the increased volume of data generated when compared to Sanger sequencing. Furthermore, '-omics' studies (*i.e.* Proteomics, Transcriptomics, Metagenomics, *etc.*), whether individually or in tandem, are rapidly becoming the new standard of base ecology strategy in the near future, due to the relatively lower costs of next-gen sequencing technologies.

The work presented in Chapter II is on Candidate Division SR1, a member of the uncultured majority of bacteria. Prior to this study, there were only a handful of SR1-affiliated sequences available in public databases, with the exception of the environment sampled, nothing was known about the distribution, diversity, abundance, or morphology of this group. Using the 16S rRNA sequences classified as candidate division SR1, oligonucleotide primers and probes were designed to selectively amplify this yet-uncultured bacterial group. A survey of multiple environments including petroleum-impacted anaerobic Zodletone Spring, indicated that SR1 was only found in anaerobic environments, and has a wide range of population size, diversity and richness, relative to the total bacterial population. Fluorescence *in situ* hybridization (FISH) was used to determine the morphology of SR1 and using this technique, a link between cellular morphology and phylogeny was shown. Dr. Noha H. Youssef conducted the FISH, and it is published in the journal *Applied and Environmental Microbiology* (Davis et al. 2009).

The temporal and spatial changes in the overall bacterial population associated with natural gas wells, was the focus of work in Chapter III. This study was conducted to better understand bacterial community assembly and succession in newly drilled natural gas wells. Despite an increase in the number of recently drilled natural gas wells, very little is known about the bacterial communities associated with natural gas reservoirs and well components.

For this study, two newly drilled natural gas wells located in the Barnett Shale in the Fort Worth Basin were sampled. Prior to this study, there had been instances of biocorrosion and sulfide gas in some of the wells of the Barnett Shale. Hydrogen sulfide is created by various sulfate-, thiosulfate-, and sulfur-reducers: members of Deltaproteobacteria, Synergistetes, and the order Halanaerobiales, as well as others. Hydrogen sulfide gas is toxic and can corrode metal. Further, before natural gas can be used, it must be treated to remove sulfides so it will burn efficiently and not release extra pollutants. The Barnett Shale bacterial communities were tracked over time from two well components (separator and tank) at two wells. The bacterial communities detected were comprised mostly of moderate halophiles or halotolerant members of the phyla Firmicutes and Proteobacteria (Classes: Beta-, Gamma-, and Epsilonproteobacteria) in both wells at all sampling times and locations. Many of the lineages detected have been detected in prior investigations of the bacterial communities from other petroleum reservoirs and production facilities. Sulfidogenic lineages (sulfate-, sulfur-, and thiosulfate-reducing lineages) were detected in all of the samples. The bacterial communities from the separator and tank samples had very little similarity to the bacterial communities in the drilling mud and hydraulic-fracture waters that were used to drill these wells, suggesting that the *in situ* development of the unique bacterial communities in such well components was in response to the prevalent geochemical conditions. Conversely, temporal and spatial community comparisons display the establishment of a core microbial community in each sampled location. Dr. Christopher G. Struchtemeyer aided in the water sampling, DNA extraction, PCR, and the manuscript editing. This work is published in the journal *Microbial Ecology* (Davis et al. 2012).

CHAPTER I

CULTURE-INDEPENDENT MOLECULAR BIOLOGICAL APPROACHES IN MICROBIAL ECOLOGY

Microbial Ecology. It is well documented that microorganisms are ubiquitous in nature and are very abundant (6, 7, 123). Not only are bacteria found in virtually all soils and water sources, they are also found in more extreme environments including near-boiling hot springs, ocean sediments, highly anaerobic sulfur springs, and hypersaline soils and sediments (26, 44, 47, 75). One gram of soil is estimated to contain 10^9 bacterial cells, with a total planetary microbial population estimated at 5×10^{30} cells (111). Bacteria are arguably the most important organisms involved in nutrient cycling of global carbon, nitrogen, phosphorous, and sulfur (98, 119). Much of the Earth's biomass is comprised of microbial carbon, estimated at 5 Pg (or 5×10^{12} kg) (111). Given the abundance and omnipresence of microorganisms in the environment, studying microbial populations is key to understanding the local and global biosphere of Earth. The study of bacteria from the environment (specifically the soil) and how they are involved in nutrient cycling has been attributed to the late 19^{th}/early 20^{th} century scientists Sergei Winogradsky and Martinus Beijerinck (15, 24). Their methods (as well as the scientists that followed) relied completely on culturing microorganisms. While extremely valuable, subsequent studies have indicated that these techniques do not accurately represent the total bacterial population from an environmental source (50, 81). Independent studies showed that the majority of the bacterial populations in the environment were metabolically active, and further studies showed that nearly 99% of bacteria from the environment would not grow using standard culturing techniques; these discrepancies were collectively called "The Great Plate Count Anomaly" (61, 97, 122). It was not until the late 1970's and early 1980's that new molecular techniques were developed to prove that indeed the majority of organisms have not yet been cultured (46, 79). These new

molecular techniques centered on the ubiquitous 16S rRNA gene of the ribosomal subunit (41, 70, 109).

The 16S rRNA Gene. In the late 1970's Carl Woese et al. developed a technique that uses the 16S rRNA gene sequence to compare the evolutionary relationships among organisms (113). The 16S rRNA gene product is involved in translation of messenger RNA, and so the gene is universally distributed among prokaryotes and eukaryotes. The 16S rRNA gene sequence has highly conserved regions with interspersed hypervariable regions (41, 70, 74). The conserved regions are considered to be 'universal' because the DNA sequences in these regions are conserved among the three domains of life: Eukarya, Bacteria, and Archaea (36, 70). Comparison of the hypervariable regions from multiple 16S rRNA gene sequences can be used to distinguish organisms from each other (104, 115). It has been found that there are gene sequences that are specific for each domain, and oligonucleotide primers have been developed that allow the amplification of sequences specific for each domain or universal among all three domains (36, 42, 114). As well, these 'fingerprint' sequences have been found specific for the classes of Proteobacteria, and even down to the genus and species level (57, 82, 96). Sequencing the 16S rRNA gene was first used to identify cultured organisms, but was soon found to be useful in describing and discovering novel, uncultured microbes (47). Subsequently, culture-independent approaches were developed for directly identifying members of bacterial communities present in the environment (75, 76). In such approaches, bulk DNA is extracted from the environment, universal oligonucleotide primers are used to amplify (via PCR) the 16S genes from the DNA. The PCR products are cloned and then sequenced. The sequences are used to infer taxonomy and distribution of the microbes

within a given environment through comparative phylogenetic analysis. This culture-independent method was expanded further when whole-cell hybridization methods (using fluorescently-labeled probes) were used to link 16S gene sequences with cellular morphology, allowing visualization and quantification of specific bacterial lineages within various ecosystems (2).

The uncultured majority. The ability to isolate and grow a pure culture of a bacterium from an environmental sample is a fundamental step in basic microbiology, and is the second of Koch's Postulates (33). However, the Great Plate Count Anomaly indicated that the majority of bacteria cannot be cultured (97). The culture-independent method developed by Pace et al. was widely utilized by microbial ecologists to construct multiple bacterial 16S rRNA gene clone libraries from multiple environments, it became clear that many of the uncultured bacteria represented multiple novel lineages (34, 54, 95, 110). By 2002, there were 36 bacterial phyla with 16S rRNA gene sequences in public databases, which was three times as many that was described in Carl Woese's seminal studies in the late 1970s (46, 114). Nearly one third of these new phyla did not have a cultured representative, and were labeled 'candidate phyla' or candidate divisions (46). By 2003, there were 52 total phyla, of which, 26 were candidate division (79). As of 2012, the Greengenes (http://greengenes.lbl.gov/) taxonomic framework has 84 recognized bacterial phyla, of which 52 have no pure cultured representative, and nine are represented by only a few isolates (22). These numbers do not include additional uncultured sub-phyla groups within known phyla, which when taken in account shows that there truly is an uncultured majority.

Anaerobic Petroleum Environments. Cultivating bacteria from anaerobic habitats requires special techniques and equipment, and even more than aerobic microorganisms, the majority of anaerobes elude cultivation (48, 62, 67). Anaerobic environments include deep sediments and eutrophic water habitats (*e.g.* ocean, lake, and spring sediments), animal and insect digestive tracts, and petroleum reservoirs (10, 12, 56). Metabolically, anaerobic bacteria are diverse and include multiple lineages of heterotrophs, autotrophs, and phototrophs (67, 116). Among others, petroleum-related anaerobic environments include naturally formed (*e.g.* mud volcanoes, gas vents, and petroleum reservoirs), as well as anthropogenic (*e.g.* petroleum production facilities and hydrocarbon contaminated water and soil) ecosystems.

Geological fissures that form due to a divergence of two tectonic plates are known as gas vents or methane seeps, and in the ocean they are usually located along Mid-ocean ridges (*e.g.* Mid-Atlantic Ridge and East Pacific Rise) (30, 31, 93). Hydrothermal vents are gas vents that release super-heated water (>300°C), gases (H_2, SO_4^{2-}, CO_2), and gaseous petroleum (*e.g.* CH_4), (30, 31, 93). These areas can host 10^4 to 10^5 more bacterial cells than the surrounding ocean water due to nutrient richness and elevated temperatures (30, 31, 93). The anaerobic bacteria found in these sites are typically autotrophic and heterotrophic sulfur- and sulfate-reducers (*e.g.* Deltaproteobacteria), H_2-oxidizing denitrifiers (*e.g.* Epsilonproteobacteria), and heterotrophic fermenters (*e.g.* order Thermotogales).

Surface methane seeps and gas vents can be found in eutrophic bodies of water (lakes and fresh water springs) (88). One such site, Zodletone Spring, is located in southwest Oklahoma and it is a mesophilic, freshwater spring that is highly anaerobic

with concentrations of zero-valent sulfur at 0.1mM and sulfide at 8-10 mM. The spring has abundant concentrations of gaseous hydrocarbons: methane, ethane, and propane (26, 88). An early bacterial 16S rRNA clone library survey revealed active members of the sulfur-cycle including anaerobic sulfur- and sulfate-reducers and photosynthetic purple and green-sulfur bacterial groups (26). Several novel sub-phyla in the proteobacterial classes Gamma- and Epsilonproteobacteria, were also detected, and there were multiple unclassified bacterial lineages with high sequence similarity to sequences from other studies of hydrocarbon contaminated environments (26). A later study of Zodletone Spring expanded the view of the diversity of the bacterial community present in the sediment, revealing the presence of 60 bacterial phyla and candidate phyla present (117). The majority of bacteria encountered with bacterial primers in this large-scale, 454-pyrosequencing study were unclassifiable (nearly 30% of the sequences) and there were many 'rare biosphere' groups (candidate divisions OP8, WS3, SR1, and others) present as well (117). Other major lineages included the anaerobic phyla Bacteroidetes, Firmicutes, and sulfate-reducing Deltaproteobacteria.

Natural gas reservoirs can include conventional sources of natural gas (*e.g.* gas sands) where the gas can be easily obtained with traditional drilling methods. However, the largest amount of natural gas is trapped in unconventional sources such as coal beds and shale (5). One of the largest producers of shale gas in the U.S. is the Barnett Shale, located in north-central Texas (5, 64). To obtain natural gas from geological shale unconventional drilling techniques (hydrofracturing and horizontal drilling) are required (27). Shale gas can be formed by biotic activities or abiotically, and the natural gas of the Barnett Shale was formed abiotically when the kerogen and long-chain hydrocarbons

cracked into short-chain gaseous hydrocarbons due high pressure and temperatures in the formation (64). Despite the abiotic origin of the gas, multiple instances of bio-corrosion (caused by biogenic sulfide-gas) have been reported in several wells of the Barnett Shale (28). A 454-pyrosequencing study of the bacterial communities associated with the drilling fluids used to drill several new wells revealed that the bio-corrosive sulfide-gas could be linked to the communities found in the drilling mud, and is potentially exacerbated by the sulfur-containing components of the drilling mud (100). This study found that the drilling mud bacterial communities were comprised mostly of the phyla Firmicutes and Gammaproteobacteria, and there were multiple sulfate-reducing Deltaproteobacteria lineages, and several thermophilic lineages including Thermodesulfobacteria and Thermotogae present (100).

Sequencing Strategies. One of the major steps in the culture-independent method is DNA sequencing (41). Nucleotide sequencing started in the early 1970's with the first sequencing efforts focusing on RNA (63). In the 1970's, Walter Gilbert and Allen Maxam were the first to successfully sequence DNA, but another method, developed by Frederick Sanger, was less labor-intensive and used less toxic reagents, and hence became the primary and most reliable method for DNA sequencing (59, 84, 85). The method is based on the enzymatic machinery involved in DNA replication (DNA polymerases and deoxynucleoside triphosphates [dNTPs]), with added dideoxynucleoside triphosphates (ddNTPs) (84, 85). The ddNTPs are the four nucleotides bases (A, C, T, G) without the 3'-OH group on the deoxyribose sugar. Briefly, an oligonucleotide primer (complementary to the template DNA to be sequenced) is annealed to a known portion of the DNA (*e.g.* a conserved region of the 16S rRNA gene); this provides the starting point

for the DNA polymerase. The DNA polymerase begins adding single dNTPs one base at a time (complementary to the template DNA) extending from the 3'-OH group of the primer (32, 85). The DNA chain is extended until a ddNTP is complementarily added, and since the 3'-OH group is lacking, the DNA polymerase cannot continue adding dNTPs (*i.e.* chain termination). Originally, four separate reactions would be set up and each would have: the template DNA (to be sequenced), primer, dNTPs, and a different ddNTP (*e.g.* ddATP, ddCTP, *etc.*) for each reaction. Later this method was modified further so that each of the four ddNTPs were fluorescently-labeled (with a different color for each base) and the process could be accomplished in one reaction, and automated by optically recording addition of each fluorescent ddNTP (91, 92). The Sanger method was instrumental in the Human Genome Project (HGP) because Sanger sequencing is accurate, and produces long reads (~1 kb) (53, 71). Despite its enormous contribution to science, Sanger sequencing does have some drawbacks. First, the sequence quality is low in the first part and last part of the read (59, 60). Second, Sanger sequencing is costly, especially when compared to other sequencing technologies that have been developed (53).

Next generation sequencing (NGS) is a broad term applied to all high-throughput sequencing techniques that have been developed since 2005 (60, 90). 454-pyrosequencing was the first of the NGS techniques to be successfully commercialized (38, 69). Pyrosequencing was developed in the late 1990's, and differs from Sanger sequencing by detection of light after a nucleotide base is added during DNA synthesis, rather than chain-terminating ddNTPs (69, 90). By 2008, 454 Life Sciences (Bradford, CT) developed and adapted commercialized pyrosequencing reagents and technologies,

which made this technology highly accessible (3, 83). Briefly, this sequencing method relies on having an biotinylated adaptor that is ligated to the DNA to be sequenced, and the adaptor binds to a DNA capture bead (with one DNA molecule per bead) (83). The beads with the necessary PCR reagents are emulsified in an oil immersion so there is one bead with the necessary reagents in a single water droplet. This clonal colony amplifies this single DNA molecule (83). These amplified bead droplets are deposited onto a PicoTiter sequencing plate where there will be one bead per well (1.6 million wells per plate); each well is an addressable position where the light that is emitted from the sequencing can be associated with an addressable position (90). The sequencing involves steps of flowing the sequencing reagents (DNA polymerase, sequencing primer, and other enzymes) across the plate with a single dNTP (*e.g.* dATP) at a time (3, 60, 90). If the base is incorporated to the newly synthesized DNA strand a pyrophosphate (PP_i) is released. ATP sulfurylase converts PP_i to ATP, and the ATP drives the conversion of luciferin to oxyluciferin, which generates visible light. A camera reads the light, and the intensity of the light correlates with the number of bases added. So, a single dGTP added would generate less light than three dGTPs added in a row. Unincorporated dNTPs and ATP are degraded by apyrase, and then the next dNTP can be flowed across the plate (3, 60, 90).

As of 2011, 454-pyrosequencing is approximately 80% cheaper than Sanger sequencing, and provides more data per run (1 million bases per run) (11, 38). It has been shown that 454-pyrosequencing gives a deeper resolution of the microbial ecology of environments sampled (25, 93, 103). As an example, two separate studies (Sanger sequenced clone library and a 454-pyrosequenced set) conducted at the same site

(sediment from an anaerobic sulfide-rich spring) revealed the overwhelming advantage of 454-pyrosequencing in encountering new and unclassified phylotypes (26, 117). In 2003, the Sanger sequencing study (of a 16S rRNA gene clone library) resulted in: 116 clones with 66 OTUs and 16% of the sequences unaffiliated/unclassified (26). In 2010, the 454-pyrosequencing study using general bacterial-specific 16S primers resulted in 77361 sequences with 18265 OTUs and nearly 30% of the sequences classified as unaffiliated/unclassified (117).

References

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

2. **Ansorge, W.** 2009. Next-generation DNA Sequencing Techniques. New Biotech. **25:**195-203.

3. **Arthur, J., B. Bohm, B. Coughlin, and M. Layne.** 2008. Hydraulic fracturing considerations for natural gas wells of the Marcellus Shale. The Groundwater Protection Council Annual Forum, Cincinnatti, OH.

4. **Bowers, R. M., S. McLetchie, R. Knight, and N. Fierer.** 2011. Spatial variability in airborne bacterial communities across land-use types and their relationship to potential sources. ISME J **5:**601-612.

5. **Bowers, R. M., A. P. Sullivan, E. K. Costello, J. L. Collett, R. Knight, and N. Fierer.** 2011. Sources of bacteria in outdoor air across cities in the midwestern United States. Appl. Environ. Microbiol. **77:**6350-6356.

6. **Brauman, A., M. D. Kane, M. Labat, and J. A. Breznak.** 1992. Genesis of Acetate and Methane by Gut Bacteria of Nutritionally Diverse Termites. Science **257:**1384-1387.

7. **Bybee, S. M., H. Bracken-Grissom, B. D. Haynes, R. A. Hermansen, R. L. Byers, M. J. Clement, J. A. Udall, E. R. Wilcox, and K. A. Crandall.** 2011. Targeted Amplicon Sequencing (TAS): A Scalable Next-Gen Approach to Multilocus, Multitaxa Phylogenetics. Genome Biol. Evol. **3:**1312-1323.

8. **Callaway, T. R., S. E. Dowd, T. S. Edrington, R. C. Anderson, N. Krueger, N. Bauer, P. J. Kononoff, and D. J. Nisbet.** 2010. Evaluation of bacterial diversity in the rumen and feces of cattle fed different levels of dried distillers grains plus solubles using bacterial tag-encoded FLX amplicon pyrosequencing. Journal of Animal Science **88:**3977-3983.

9. **Chung, K. T., and D. H. Ferris.** 1996. Martinus Willem Beijerinck (1851-1931): Pioneer of General Microbiology. ASM News **62:**539-543.

10. **DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen.** 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl. Environ. Microbiol. **72:**5069-5072.

11. **Dworkin, M.** 2012. Sergei Winogradsky: A founder of modern microbiology and the first microbial ecologist. FEMS Microbiol. Rev. **36:**364-379.

12. **Edwards, R. A., B. Rodriguez-Brito, L. Wegley, M. Haynes, M. Breitbart, D. M. Peterson, M. O. Saar, S. Alexander, E. A. Alexander Jr, and F. Rohwer.** 2006. Using pyrosequencing to shed light on deep mine microbial ecology. BMC Genomics. **7:**57.

13. **Elshahed, M. S., J. M. Senko, F. Z. Najar, S. M. Kenton, B. A. Roe, T. A. Dewers, J. R. Spear, and L. R. Krumholz.** 2003. Bacterial Diversity and Sulfur Cycling in a Mesophilic Sulfide-Rich Spring. Appl. Environ. Microbiol. **69:**5609-5621.

14. **Entrekin, S., M. Evans-White, B. Johnson, and E. Hagenbuch.** 2011. Rapid expansion of natural gas development poses a threat to surface waters. Front. Ecol. Environ. **9:**503-511.

15. **Fichter, J., K. Johnson, K. French, and R. Oden.** 2009. Biocides control Barnett Shale fracturing fluid contamination. Oil Gas J. **107:**38-44.

16. **Flores, G. E., J. H. Campbell, J. D. Kirshtein, J. Meneghin, M. Podar, J. I. Steinberg, J. S. Seewald, M. K. Tivey, M. A. Voytek, Z. K. Yang, and A.-L. Reysenbach.** 2011. Microbial community structure of hydrothermal deposits from geochemically different vent fields along the Mid-Atlantic Ridge. Environmental Microbiology **13:**2158-2171.

17. **Flores, G. E., M. Shakya, J. Meneghin, Z. K. Yang, J. S. Seewald, C. Geoff Wheat, M. Podar, and A. L. Reysenbach.** 2012. Inter-field variability in the microbial communities of hydrothermal vent deposits from a back-arc basin. Geobiology **10:**333-346.

18. **Franca, L. T. C., E. Carrilho, and T. B. L. Kist.** 2002. A review of DNA sequencing techniques. Quart. Rev. Biophys. **35:**169-200.

19. **Fredericks, D. N., and D. A. Relman.** 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. Clinical Microbiology Reviews **9:**18-33.

20. **Fuhrman, J. A., K. McCallum, and A. A. Davis.** 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. Applied and Environmental Microbiology **59:**1294-1302.

21. **Galvez, A., M. Maqueda, M. Martinez-Bueno, and E. Valdivia.** 1998. Publication rates reveal trends in microbiological research. ASM News **64:**269-275.

22. **Giovannoni, S. J., E. F. DeLong, G. J. Olsen, and N. R. Pace.** 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. J. Bacteriol. **170:**720-726.

23. **Glenn, T. C.** 2011. Field guide to next-generation DNA sequencers. Molec. Ecol. Resourc. **11:**759-769.

24. **Head, I. M., J. R. Saunders, and R. W. Pickup.** 1998. Microbial Evolution, Diversity, and Ecology: A Decade of Ribosomal RNA Analysis of Uncultivated Microorganisms. Micro. Ecol. **35:**1-21.

25. **Hicks, R. E., R. I. Amann, and D. A. Stahl.** 1992. Dual staining of natural bacterioplankton with 4',6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. Applied and Environmental Microbiology **58:**2158-2163.

26. **Hollister, E. B., A. S. Engledow, A. J. M. Hammett, T. L. Provin, H. H. Wilkinson, and T. J. Gentry.** 2010. Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. ISME J **4:**829-838.

27. **Hugenholtz, P.** 2002. Exploring prokaryotic diversity in the genomic era. Genome Biol. **3:** reviews0003.1–reviews0003.8.

28. **Hugenholtz, P., B. M. Goebel, and N. R. Pace.** 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J. Bacteriol. **180:**4765-4774.

29. **Hungate, R. E.** 1969. A Roll Tube Method for Cultivation of Strict Anaerobes. *In* J. R. Norris and D. W. Ribbons (ed.), Methods in Microbiology, vol. 3B. Academic Press Inc., New York.

30. **Jannasch, H. W., and G. E. Jones.** 1959. Bacterial populations in sea water as determined by different methods of enumeration. Limnol. Oceanogr. **4:**128-129.

31. **Lander, E. S.** 2011. Initial impact of the sequencing of the human genome. Nature **470:**187-197.

32. **Liesack, W., and E. Stackebrandt.** 1992. Occurrence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. Journal of Bacteriology **174:**5072-5078.

33. **Magot, M., B. Ollivier, and B. K. Patel.** 2000. Microbiology of petroleum reservoirs. Antonie Van Leeuwenhoek **77:**103-116.

34. **Manz, W., R. Amann, W. Ludwig, M. Wagner, and K.-H. Schleifer.** 1992. Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. Syst. Appl. Microbiol. **15:**593-600.

35. **Metzker, M. L.** 2005. Emerging technologies in DNA sequencing. Genome Res. **15:**1767-1776.

36. **Metzker, M. L.** 2009. Sequencing technologies - the next generation. Nat Rev Genet **11:**31-46.

37. **Meyer-Reil, L.-A.** 1978. Autoradiography and Epifluorescence Microscopy Combined for the Determination of Number and Spectrum of Actively Metabolizing Bacteria in Natural Waters. Appl. Environ. Microbiol. **36:**506-512.

38. **Miller, T. L., and M. J. Wolin.** 1974. A Serum Bottle Modification of the Hungate Technique for Cultivating Obligate Anaerobes. Appl. Microbiol. **27:**985-987.

39. **Min Jou, W., G. Haegeman, M. Ysebaert, and W. Fiers.** 1973. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. Nature **237:**82-88.

40. **Montgomery, S., D. Jarvie, K. Bowker, and R. Pollastro.** 2005. Mississippian Barnett Shale, Fort Worth basin, north-central Texas: gas-shale play with multi-trillion cubic foot potential. AAPG Bull. **89:**155–175.

41. **Nealson, K. H.** 1997. SEDIMENT BACTERIA: Who's There, What Are They Doing, and What's New? Annu. Rev. Earth Planet. Sci. **25:**403-434.

42. **Nyrén, P.** 2007. The History of Pyrosequencing. Methods Mol. Biol. **373:**1-14.

43. **Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace.** 1986. Microbial Ecology and Evolution: A Ribosomal RNA Approach. Ann. Rev. Microbiol. **40:**337-365.

44. **Olson, M. V.** 1993. The Human Genome Project. Proc. Natl. Acad. Sci. U.S.A. **90:**4338-4344.

45. **Pace, N., D. Stahl, D. Lane, and O. GJ.** 1985. Analyzing natural microbial populations by rRNA sequences. ASM News **51:**4-12.

46. **Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen.** 1986. The analysis of natural microbial populations by ribosomal RNA sequences. Adv Microb Ecol **9:**1-55.

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

48. **Rappé, M. S., and S. J. Giovannoni.** 2003. The Uncultured Microbial Majority. Ann. Rev. Microbiol. **57:**369-394.

49. **Razumov, A. S.** 1932. The direct method of calculation of bacteria in water: comparision with the Koch method. Mikrobiol. **1:**131-146.

50. **Rossi, F., S. Torriani, and F. Dellaglio.** 1999. Genus- and Species-Specific PCR-Based Detection of Dairy Propionibacteria in Environmental Samples by Using Primers Targeted to the Genes Encoding 16S rRNA. Applied and Environmental Microbiology **65:**4241-4244.

51. **Rothberg, J., and J. Leamon.** 2008. The development and impact of 454 sequencing. Nat Biotech **26:**1117-1124.

52. **Russo, E.** 2003. Special Report: The birth of biotechnology. Nature **421:**456-457.

53. **Sanger, F., and A. R. Coulson.** 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94:**441-448.

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

55. **Schleifer, K.-H., and W. Ludwig.** 1994. Molecular taxonomy: classification and identification, p. 1-15. *In* F. G. Priest, A. Ramos-Cormenzana, and B. J. Tindall (ed.), Bacterial diversity and systematics. Plenum Press, New York.

56. **Senko, J. M., B. S. Campbell, J. R. Henriksen, M. S. Elshahed, T. A. Dewers, and L. R. Krumholz.** 2004. Barite deposition resulting from phototrophic sulfide-oxidizing bacterial activity. Geochimica et Cosmochimica Acta **68:**773-780.

57. **Shendure, J., and H. Ji.** 2008. Next-generation DNA sequencing Nat. Biotech. **26:**1135-1145.

58. **Smith, L., S. Fung, M. Hunkapiller, T. Hunkapiller, and L. Hood.** 1985. The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. Nucleic Acids Res. **13:**2399-2412.

59. **Smith, L., J. Sanders, R. Kaiser, P. Hughes, C. Dodd, C. Connell, C. Heiner, S. Kent, and L. Hood.** 1986. Fluorescence detection in automated DNA sequence analysis. Nature **321:**674-679.

60. **Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl.** 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere. Proc. Natl. Acad. Sci. U.S.A. **103:**12115-12120.

61. **Stackebrandt, E., W. Liesack, and B. M. Goebel.** 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rDNA analysis. The FASEB Journal **7:**232-6.

62. **Stahl, D. A., and R. I. Amann.** 1991. Development and application of nucleic acid probes in bacterial systematics, p. 205-248. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, Chichester, England.

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

64. **Stevenson, F. J., and M. A. Cole.** 1999. Cycles of Soil: Carbon, Nitrogen, Phosphorus, Sulfur, Micronutrients. John Wiley and Sons, Inc., New York.

65. **Struchtemeyer, C. G., J. P. Davis, and M. S. Elshahed.** 2011. Influence of the Drilling Mud Formulation Process on the Bacterial Communities in Thermogenic Natural Gas Wells of the Barnett Shale. Applied and Environmental Microbiology **77:**4744-4753.

66. **Swofford, D., and G. J. Olsen.** 1990. Phylogeny Reconstruction, p. 411-501. *In* D. Hillis and C. Moritz (ed.), Molecular systematics. Sinauer Associates, Sunderland, MA.

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

68. **Van de Peer, Y., S. Chapelle, and R. De Wachter.** 1996. A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res. **24:**3381-3391.

69. **Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane.** 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. **173:**697-703.

70. **Weller, R., M. M. Bateson, B. K. Heimbuch, E. D. Kopczynski, and D. M. Ward.** 1992. Uncultivated cyanobacteria, Chloroflexus-like inhabitants, and spirochete-like inhabitants of a hot spring microbial mat. Applied and Environmental Microbiology **58:**3964-3969.

71. **Whitman, W. B., D. C. Coleman, and W. J. Wiebe.** 1998. Prokaryotes: The unseen majority. Proc. Natl. Acad. Sci. U.S.A. **95:**6578-6583.

72. **Woese, C., and G. Fox.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U.S.A. **74:**5088-5090.

73. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221-271.

74. **Woese, C. R.** 1994. There must be a prokaryote somewhere: microbiology's search for itself. Microbiological Reviews **58:**1-9.

75. **Youssef, N., M. S. Elshahed, and M. J. McInerney.** 2009. Microbial Processes in Oil Fields: Culprits, Problems, and Opportunities, p. 141-251. *In* S. S. Allen I. Laskin and M. G. Geoffrey (ed.), Advances in Applied Microbiology, vol. Volume 66. Academic Press.

76. **Youssef, N. H., M. B. Couger, and M. S. Elshahed.** 2010. Fine-Scale Bacterial Beta Diversity within a Complex Ecosystem (Zodletone Spring, OK, USA): The Role of the Rare Biosphere. PLoS ONE **5:**e12414.

77. **Zehr, J. P., and B. B. Ward.** 2002. Nitrogen Cycling in the Ocean: New Perspectives on Processes and Paradigms. Appl. Environ. Microbiol. **68:**1015-1024.

78. **Zimmermann, R., R. Iturriaga, and J. Becker-Birck.** 1978. Simultaneous determination of the total number of aquatic bacteria and the number thereof involved in respiration. Appl. Environ. Microbiol. **36:**926-935.

79. **Zinger, L., L. A. Amaral-Zettler, J. A. Fuhrman, M. C. Horner-Devine, S. M. Huse, D. B. M. Welch, J. B. H. Martiny, M. Sogin, A. Boetius, and A. Ramette.** 2011. Global Patterns of Bacterial Beta-Diversity in Seafloor and Seawater Ecosystems. PLoS ONE **6:**e24570.

CHAPTER II

ASSESSMENT OF THE DIVERSITY, ABUNDANCE, AND ECOLOGICAL DISTRIBUTION OF MEMBERS OF CANDIDATE DIVISION SR1 REVEALS A HIGH LEVEL OF PHYLOGENETIC DIVERSITY BUT LIMITED MORPHOTYPIC DIVERSITY

Davis, J.P., N.H. Youssef, and M.S. Elshahed. 2009. Appl. Environ. Microbiol. 75:4139-4148.

DOI:10.1128/AEM.00137-09

Abstract

A combination of 16S rRNA gene clone library surveys, quantitative PCR (qPCR) analysis, and fluorescent *in situ* hybridization was used to investigate the diversity, abundance, and distribution of members of candidate division SR1 in multiple habitats. Using SR1-specific 16S rRNA gene primers, multiple novel SR1 lineages were identified in four different anaerobic environments: sediments from Zodletone Spring, a sulfide- and sulfur-rich spring in southwestern Oklahoma; inner layers of microbial mats obtained from Sperm Pool, a high-temperature, low-pH pool (55°C, pH 2.5) in Yellowstone National Park; fresh bovine ruminal contents; and anaerobic freshwater pond sediments (Duck Pond) in Norman, Oklahoma. qPCR analysis indicated that SR1 members constitute a small fraction (<0.01%) of the microbial communities in Duck Pond and ruminal samples but constitute a significant fraction (11.6 and 48.7%) of the total number of bacterial 16S rRNA genes in Zodletone Spring and the inner layers of Sperm Pool microbial mat samples, respectively. By using SR1-specific fluorescent probes, filamentous cells were identified as the sole SR1 morphotype in all environments examined, with the exception of Sperm Pool, where a second bacillus morphotype was also identified. Using a full-cycle 16S rRNA approach, it was shown that each of these two morphotypes corresponds to a specific phylogenetic lineage identified in the Sperm Pool clone library. This work greatly expands the intralineage phylogenetic diversity within candidate division SR1 and provides valuable quantification and visualization tools that could be used for investigating the ecological roles, dynamics, and genomics of this as-yet-uncultured bacterial phylum.

Introduction

Gene-based surveys using the 16S rRNA gene conducted during the last two decades have convincingly demonstrated that the scope of bacterial diversity is much broader than previously implied using culture-based approaches (39, 57, 62). Remarkably, many of the novel lineages discovered using 16S rRNA surveys represent deep phylum-level branches within the domain Bacteria (21, 37), necessitating coining the term "candidate divisions" to describe such lineages (36). As of November 2008, the number of recognized candidate divisions varies in different taxonomical schemes (*e.g.*, between 43 in NCBI taxonomic outline and 61 in Hugenholtz taxonomic outline in Greengenes web server [18]). These estimates will undoubtedly continue to rise with the implementation of novel sequencing technologies in microbial diversity studies (48), as well as with the recent availability of curated databases and rapid alignments and classification tools for 16S rRNA gene clone libraries (17, 18). With the exception of the 16S rRNA sequences and description of the environment from which they were encountered, little is usually known regarding the physiological properties, energy conservation pathways, and ecological significance of the members of the majority of these novel candidate divisions. This is especially true for members of novel candidate divisions that have always been encountered as a minor component within environmental clone libraries (*e.g.*, candidate divisions AC1, AD3, LD1, NC10, SC3, SC4, SPAM, TM6, OD1, and WS6) (21–23, 44), as well as those that are deposited in public databases but have not yet been described in peer-reviewed publications (*e.g.*, candidate division ctg-CGOF). Clearly, the development of lineage-specific oligonucleotide 16S rRNA primers and probes could enhance our understanding of the breadth of phylogenetic

diversity within these groups (12, 20, 25, 28, 33, 38) and aid in the implementation of targeted genomics investigation (59–61). One of these yet-uncultured lineages is candidate division SR1, which has frequently been reported in 16S rRNA gene clone libraries, especially those derived from anaerobic habitats. Sequences belonging to members of candidate division SR1 as recognized today were first encountered in a survey of a hydrocarbon-contaminated aquifer (21), in which they were classified as members of candidate division OP11 (21, 36). Shortly afterwards, 16S rRNA sequences belonging to candidate division SR1 as recognized today were reported from deep-sea sediments (15, 43), hydrothermal vents (63), oral cavity (GenBank accession number AF125207), termite gut (34), diseased coral tissue (27), near-boiling, silica-depositing thermal springs (7), and a mesophilic sulfide and sulfur rich spring (22). SR1 sequences were referred to in these studies as unknown or unaffiliated (7, 43), candidate division VC2 (22), Aquificales-related (15), or OP11 (34). Harris et al. (32) recognized the polyphyletic nature of candidate division OP11 as originally proposed (36, 37) and suggested the name candidate division SR1 (after Sulfur River, KY, where additional SR1 sequences were encountered) to describe members of the OP11-4 group. Since then, multiple SR1 sequences have been reported from diverse environments, *e.g.*, sulfur-rich springs and caves (31, 46), microbial mats in various hot springs in Yellowstone National Park (50, 68), deep-sea sediments (73), hydrothermal vents (16, 51, 55, 58), and insect and animal guts and alimentary tracts (42, 52, 75).

In this study, multiple SR1-specific primers and probes were developed based on SR1 sequences currently available in public databases and were used for the following: (i) an in-depth investigation of the intralineage phylogenetic diversity within candidate

division SR1 in multiple habitats; (ii) enumeration of SR1 16S rRNA genes using quantitative PCR (qPCR); and (iii) visualization of SR1 cells using fluorescently labeled probes and linking the observed morphologies to specific SR1 lineages encountered in this study using a full-cycle 16S rRNA approach. This study described multiple novel lineages and a high level of phylogenetic diversity within candidate division SR1. As well, this study also demonstrated that in contrast to this high phylogenetic diversity, SR1 cells appear to have limited morphotypic diversity, with only two morphotypes identified in all ecosystems examined.

Materials and Methods

Site description, sampling, and geochemical measurements. The diversity, abundance, and morphology of members of candidate division SR1 were investigated in multiple habitats: (i) sediments from an anaerobic, sulfide- and sulfur-rich spring (Zodletone Spring) in southwestern Oklahoma (see references 22 and 66 for a detailed description of the spring); (ii) the middle layer (1 cm) of a multispecies microbial mat (approximately 5 cm thick) collected from the outfall channel of Sperm Pool, a high-temperature (55°C), acidic (pH 2.5) pool in Yellowstone National Park; (iii) fresh bovine rumen contents from grass-fed fistulated cows in Oklahoma State University Animal Nutritional Physiology Center (Stillwater, OK); (iv) sediments from an anaerobic, mesophilic freshwater pond (Duck Pond) in Norman, OK; (v) soil samples from Kessler farm biological station in central Oklahoma (see reference 23 for description of the site and soil properties); and (vi) two crude oil-impacted soil samples and one pristine aerobic surface soil sample collected from a tallgrass prairie preserve in Osage County in Oklahoma from a site adjacent to an oil pipeline break.

Samples collected for 16S rRNA gene analysis and qPCR analysis from Zodletone Spring, Kessler farm biological station soil, Duck Pond, and cow ruminal contents were stored on ice or dry ice until being transferred to the laboratory where they were frozen at 20°C till further analysis. Sperm Pool and prairie soil samples were kindly provided by Babu Fathepure (Oklahoma State University) and Kathleen E. Duncan (University of Oklahoma), respectively. These samples were shipped on ice and stored in the laboratory at -20°C. Samples for fluorescent *in situ* hybridization (FISH) from Zodletone Spring, rumen contents, and Duck Pond were fixed on-site (1:3,

sediment:fixant buffer) in 4% (wt/vol) paraformaldehyde in phosphate-buffered saline (PBS) (154 mM NaCl, 1.69 mM KH_2PO_4, 5 mM Na_2HPO_4) and stored on ice. Yellowstone microbial mat samples were shipped on dry ice and fixed upon arrival to the laboratory as described above.

For geochemical measurements, samples for sulfide analysis were added directly to an equal volume of 10% zinc acetate solution on-site to fix sulfide. Sulfide was quantified as previously described (35). Sulfate and nitrate were quantified by ion chromatography as described before (66). Zero-valent sulfur was determined as previously described in detail in reference 66. Briefly, samples were acidified to pH 1.5 to 2 on-site to precipitate sulfane sulfur and were quantified as sulfide following a Cr(II) extraction procedure (72).

Primers and probe design. Oligonucleotides selectively targeting members of candidate division SR1 were designed using the probe design function on SR1 affiliated sequences available in Greengenes May 2007 database in ARB software package (45). Candidate oligonucleotides obtained were further evaluated against a more detailed SR1 list that includes, in addition to sequences available in the ARB database, partial SR1 sequences (1,200 nucleotides) available in GenBank database, and SR1 sequences that were recently deposited (from June 2007 to October 2008) in GenBank (Data Not Shown). Redundancies were added to increase coverage when necessary, and the probes obtained were tested for specificity using BLASTnr search (2), as well as RDP probe match function within the Ribosomal Database Project web server (13). Four SR1-specific oligonucleotides were obtained and used as primers for diversity analysis, qPCR analysis, and/or as FISH probes in this study (Table 2-1). Probes specific to SR1

subgroup I and SR1 subgroup V lineages (see Results below) were designed in our attempt to link observed cell morphologies to certain phylogenetic lineages in Sperm Pool samples using a full cycle 16S rRNA gene approach (see Results below) (4). Lineage-specific probes were designed by importing Sperm Pool community sequences into the ARB database using the ARB probe design function (45). The two probes obtained (Table 2-1) were evaluated for specificity and coverage as described above.

DNA extraction, construction, and sequencing of 16S rRNA gene clone libraries. DNA was extracted from all environments using the FastDNA spin kit for soil (MP Biomedicals, Solon, OH). 16S rRNA gene of members of candidate division SR1 was selectively amplified using the Bacteria-specific forward primer 27F and the SR1-specific primer 914R (Table 2-1). Compared to other primer combinations (SR1-427F/SR1-668R, SR1-427F/SR1-1075R, and 27F/SR11075R), this primer pair combination gave a near full-length gene coverage (ca. 900 bp), and had absolute specificity for SR1 sequences. PCR was conducted in a 50-l reaction mixture containing the following (final concentrations are given): 2 l of extracted DNA, 1 PCR buffer (Promega), 2.5 mM $MgSO_4$, 0.2 mM deoxynucleoside triphosphate mixture, 2.5 U of GoTaq Flexi DNA polymerase (Promega, Madison, WI), and 10 μM of each of the forward and reverse primers. PCR amplification was carried out according to the following protocol: initial denaturation for 5 min at 95°C, followed by 30 cycles, with 1 cycle consisting of denaturation at 95°C for 45 s, annealing at 52°C for 45 s, and elongation at 72°C for 1.5 min. A final elongation step at 72°C for 15 min was included. PCR products obtained were cloned into a TOPO-TA cloning vector according to the

manufacturer's instructions (Invitrogen Corp., Carlsbad, CA), and sequenced as previously described (22).

Phylogenetic analysis. To check the phylum level affiliations of clones obtained in this study, sequences were initially compared to entries in the GenBank database using BLASTnr (2). In addition, sequences were aligned in Greengenes NAST aligner to a 7,862-character global alignment (17) and run through Greengenes classifier (18). In addition to Greengenes classifier output, the NAST-aligned sequences were imported to Greengenes May 2007 ARB database and added to the ARB universal dendrogram using the ARB parsimony function to determine their position in the global phylogenetic tree (45).

For operational taxonomic unit (OTU) assignment and phylogenetic tree construction, SR1 sequences were aligned using ClustalX program (71), and the alignments were exported to PAUP (version 4.01b10; Sinauer Associates, Sunderland, Mass). A pair-wise distance matrix generated in PAUP was exported to DOTUR (64) and used for assignment of OTUs at 97% sequence similarity cutoff. Basic diversity measurements, *e.g.*, Shannon-Weaver diversity index, average nucleotide diversity, and Good's coverage were calculated as previously described (29, 47, 49). Phylogenetic trees were constructed using SR1 OTUs from this study and representative closely related reference sequences. Distance neighbor-joining trees with no corrections, F-84 corrections, and Jukes-Cantor corrections were constructed using PAUP and gave similar tree topologies.

qPCR. qPCR was used to quantify members of SR1 in multiple environments by using a MyiQ thermocycler (Bio-Rad Laboratories, Hercules, CA) and B-R SYBR green SuperMix for iQ (Quanta Bioscience, Inc., Gaithersburg, MD). The primer pair SR1-427F/SR1-668R was used for SR1 quantification, and the primer pair EUB-338/UNI518R (23, 24) was used to amplify the total bacterial community. Specificity of primer pair SR1-427F/SR1-668R was initially confirmed by cloning the PCR product obtained and sequencing 12 clones, all of which were affiliated with SR1 (data not shown).

The 25 μl PCR reaction mixture contained 0.3 μM of each forward and reverse primers (final concentration), 3 μl extracted template DNA, and 12.5 μl B-R SYBR green SuperMix. The reactions were heated at 95°C for 3 min, followed by 55 cycles, with one cycle consisting of 20 s of 95°C and 30 s at 52°C or 54°C. A pCR 4-TOPO (Invitrogen) plasmid with an SR1 16S rRNA gene insert generated using primer pair 27F and 914R was used as a positive control, as well as to generate a standard curve for both reactions. The efficiency of the amplification of the standards (*E*) was calculated from the slope of the standard curve using the formula $E = (10^{-1/\text{slope}}) - 1$.

FISH. (i) Sample preparation, sediment removal, and fixation. Paraformaldehyde-fixed samples were centrifuged at high speed (14,000 x *g* for 10 min) to remove the fixant. For Zodletone Spring and Duck Pond sediments, the pellet was resuspended in PBS containing 100 mM sodium pyrophosphate followed by vigorous shaking for 10 min at room temperature. Large sediment particles were removed by centrifugation at 2,000 x *g* for 10 min. The supernatant containing the cells and finer sediment particles was subjected to a higher speed centrifugation (6,000 x *g* for 5 min) to

remove fine sediment particles followed by high-speed centrifugation at 14,000 x *g* for 10 to 15 min to collect cells. The pellet was then resuspended in a minimal volume of PBS, and the presence of cells was confirmed by phase-contrast microscopy. Fixed cells were stored at 20°C in PBS-ethanol (1:1) until their use for FISH. For the Sperm Pool mat samples, which contained no sediments, paraformaldehyde was removed by centrifugation at 14,000 x *g* for 10 to 15 min. The microbial mat was resuspended in PBS and dispersed using VDI 12 sonicator (VWR Corp., West Chester, PA) pulses (two or three pulses [each pulse 30 s]) at position 2 and then stored in PBS-ethanol (1:1) at -20°C until use.

(ii) Hybridization and visualization. Alexa Fluor 488-labeled probes S-P-SR10427-a-A-18 (1) (SR1-427) and S-P-SR1-1075-a-A-18 (SR1-1075), targeting all members of candidate division SR1, as well as Alexa Fluor 488-labeled probes S-P-SR1I-0232-a-A-18 (SR1-232), and probe S-P-SR1II-0575-a-A-18 (SR1-575), which target subgroups I and V in BD2-14 lineage within SR1, respectively, were synthesized by Invitrogen. All probes were named using the probe nomenclature scheme by Alm et al. (1). All hybridizations were done on Cel-Line slides (Thermoscientific, Portsmouth, NH) with Adcell bioadhesive coating. Six-well slides were used for independent sample positioning. Hybridizations were carried out as previously described (3). Briefly, 10-ml aliquots of fixed cells in PBS were put into individual wells and allowed to air dry. Cells were then dehydrated in an increasing series of ethanol solutions (50, 80, and 100%) for 3 min in each solution. Hybridization buffers contained 900 mM NaCl, 0.1% sodium dodecyl sulfate, 20 mM Tris HCl, the required percentage of formamide (see below), and a final concentration of 5 ng/ml of oligonucleotide probe (when two probes were used, a

final concentration of 5 ng/ml of each was used) in a final volume of 10 ml. All hybridizations were carried out at 46°C for 3 h in humid chambers. Slides were then rinsed with prewarmed (48°C) washing buffer, followed by a 25-min immersion in washing buffer at 48°C. Washing buffers contained 0.1% sodium dodecyl sulfate, 20 mM Tris HCl, and depending on the formamide concentration in the hybridization buffer, between 10 and 900 mM NaCl (Table 2-1).

Slides were then rinsed in nanopure water, allowed to air dry, followed by a 10-min incubation at room temperature in the dark with the DNA-binding dye 4,6 - diamidino-2-phenylindole (DAPI) (final concentration of 10 g/ml). Excess DAPI was then rinsed off, and the slides were air dried in the dark and then mounted in 4% n-propyl gallate in 90% (vol/vol) glycerol in PBS. Slides were visualized using an Olympus BX51 microscope (Olympus, Center Valley, PA), equipped with Brightline fluorescein isothiocyanate (FITC) and tetramethyl rhodamine isothiocyanate (TRITC) filter sets for Alexa Fluor 488 and Alexa Fluor 546 fluorescence, as well as a Brightline DAPI high-contrast filter set for DAPI fluorescence. Photomicrographs were taken with a DP71 digital camera (Olympus). Exposure times were 10 ms for DAPI fluorescence and 100 ms for Alexa Fluor fluorescence.

FISH controls. Due to the unavailability of pure culture representatives of candidate division SR1, a clone FISH technique was used to determine the optimum formamide concentration for probe hybridizations (65). Clones containing the target sequence for the SR1-427 probe were obtained from our ruminal fluid clone library. Several clones were sequenced to identify a clone with the insert in the forward direction. Since the SR1-1075 probe targets a region that is not amplified with SR1 primers used in

this study, a synthetic 362-bp 16S rRNA gene segment corresponding to nucleotide positions 810 to 1163 in an uncultured SR1 clone from Sulfur River in Kentucky was used (32), GenBank accession number AY193201). Gene synthesis was performed by Genscript Corp. (Piscataway, NJ), and the insert (supplied in pUC57 vector that lacks a T7 priming site) was cloned into PCR-4-Topo vector, and subsequently transformed into TOP10 chemically competent host cells (Invitrogen), according to the manufacturer's instructions. Several clones were then sequenced to identify a clone with the insert in the forward direction. These clones, now carrying the target sequence for either the SR1-427 or SR1-1075 probe and a T7 priming site (while the host cells lack a T7 RNA polymerase) could be used for clone FISH following an overnight incubation with chloramphenicol (170 mg/liter) to increase the plasmid copy number and generate high levels of target rRNA through leaky transcription from the T7 priming site (65). The clones were then fixed in 4% paraformaldehyde (overnight at 4°C), washed twice in PBS, and then stored at 20°C in PBS-ethanol until subjected to FISH. Formamide concentrations of 35 to 65% were tested in 5% increments to determine the optimum concentration for use with environmental samples (Table 2-1).

Negative controls. The RDP database probe check identified cultured microorganisms with one mismatch to SR1 probes to be used as negative controls. *Desulfotomaculum geothermicum* strain B2T (DSMZ 3669; GenBank accession number AJ621886) has one mismatch to the SR1-427 probe, and *Vulcanisaeta souniana* strain IC059T (DSMZ 14430; GenBank accession no AB063645) has one mismatch to the SR1-1075 probe. Both microorganisms were obtained from the German Resource Centre for Biological Material (DSMZ, Braunschweig, Germany) and cultured in the appropriate

culturing medium. Cells were fixed and stored as described above until they were used for FISH. To ensure that negative hybridization with the SR1 probes in the negative controls was due to target sequence mismatch and not due to the inaccessibility of the probes to the target site, the Alexa Fluor 546-labeled universal probe Univ1390 (76) and *Crenarchaeota* probe Cren499 (11) were used as FISH-positive controls for *D. geothermicum* and *V. souniana* strains, respectively (the universal probe Univ1390 has a single mismatch with *V. souniana* 16S rRNA gene and hence could not be used). The non-EUB-338 probe (4) was also included as an additional control to monitor nonspecific probe binding. For *V. souniana*, cell wall permeabilization was required (by incubation with proteinase K 4U/ml at 37°C for 1 h; Sigma, St. Louis, MO) due to its protein-rich cell wall (70, 74).

Nucleotide sequence accession numbers. Sequences generated in this study were deposited in GenBank under accession numbers FJ479804 to FJ480103.

Results

Phylogenetic diversity of members of candidate division SR1 in multiple anaerobic habitats. SR1 sequences were identified in Zodletone Spring, Sperm Pool microbial mat, Duck Pond, and cow ruminal samples. SR1 sequences, however, were not detected in the soil samples examined in this study. This is in agreement with the fact that SR1-affiliated sequences have never been encountered in any of the thousands of soil clone libraries constructed and analyzed so far (39) and suggests that this is due to the unavailability of suitable conditions in soils, rather than inadequate sampling.

A total of 300 clones were sequenced from these four environments, with the aim of achieving high coverage (>95%, Table 2-2) within each clone library. The number of clones, OTUs, and various sequence diversity estimates per environment (*e.g.* Shannon-Weaver diversity index, mean sequence divergence, average nucleotide divergence) are shown in Table 2-2. Clearly, Zodletone Spring and Sperm Pool microbial mats were the most diverse, while rumen samples and duck pond were the least diverse, with only two OTUs in each of these two libraries (Table 2-2).

Within the Hugenholtz taxonomic outline in Greengenes database, candidate division SR1 is currently divided into two subphylum level lineages, BH1 and BD2-14, after clones BH1 (GenBank accession number AF352532) and BD2 to BD14 (GenBank accession number AB015542) from Japan deep-sea sediments and Black Pool in Yellowstone National Park, respectively (7, 43). Members of BH-1 lineage have been encountered only in geothermal habitats, mainly hot springs and pools at Yellowstone National Park (7, 50, 56, 68), as well as California (GenBank accession number

EU942246), and Tibet (GenBank accession number EF205568). Members of this monophyletic lineage have high sequence divergence (30.0 to 33.7%) from their closest SR1 clones belonging to BD2-14 lineages. Indeed, the BH1 and BD2-14 lineages are not always monphyletic in trees constructed using taxa belonging to phyla loosely related to SR1 (*e.g.*, OP11 and OD1) using parsimony and neighbor-joining algorithms (data not shown). Future availability of sequences belonging to the BH1 lineage might warrant designating this group as an independent candidate division.

Compared to the relatively limited ecological distribution of the BH-1 lineage, members of the BD2-14 lineage are relatively more widely distributed, being detected in geothermal environments as well as low-temperature terrestrial and marine environments with various degrees of salinity (Fig. 2-1). All sequences retrieved in this study (157 clones from Zodletone Spring, 71 from Sperm Pool, 30 from cow rumen, and 33 from Duck Pond) belonged to various subgroups within this lineage. Phylogenetic analysis (Fig. 2-1) identified nine bootstrap-supported subgroups within the BD2-14 lineage. Some of these lineages (subgroups I, VI, and IX) are exclusively represented by sequences retrieved in this study and hence represent novel lineages within candidate division SR1 (Fig. 2-1).

All bovine ruminal fluid sequences belonged to two OTUs, both of which were closely related and belonged to BD2-14 subgroup III. It is interesting to note that this lineage is exclusively composed of SR1 clones retrieved from human (*e.g.*, oral cavity, esophagus), mammalian (cow rumen, rhinoceros feces), or insect (termite) origins (Fig. 2-1).

Sequences from Sperm Pool microbial mats (71 clones) belonged to seven OTUs and two distinct SR1 subgroups. One group (34 clones, two OTUs) belonged to subgroup V and was closely related to sequences from Zodletone Spring, Duck Pond, and deep-sea clone BD2-14 (Fig. 2-1). The second group (37 clones, six OTUs) formed a bootstrap-supported, deep branching, and distinct novel SR1 lineage (subgroup I), together with a single OTU retrieved from a Duck Pond clone library. Members of this lineage have extremely low sequence similarity (78 to 80%) to their closest SR1 relative outside this group.

The Zodletone Spring SR1 community was clearly the most diverse of all environments tested, judging by nucleotide sequence diversity parameters (Table 2-2), as well as the fact that Zodletone Spring SR1 sequences belonged to seven out of the nine SRI-BD2-14 lineages (subgroups II, IV, V, VI, VII, VIII, and IX). Three lineages (subgroups II, VI, and IX) are exclusively formed by sequences encountered in Zodletone Spring, either only in this study (subgroups VI and IX), or in this study as well as in a clone library generated using general bacterial primers (subgroup II, clone ZB18 in Fig. 2-1) in a previous study (22).

Candidate division SR1 quantification. The primer pair SR1-427F and SR1-668R was used to quantify SR1 16S rRNA gene copies in all four environments examined in this study. Members of candidate division SR1 have the lowest 16S rRNA gene copy numbers in Duck Pond, followed by bovine rumen, and in both samples, members of SR1 represented a small fraction (less than 0.01%) of the total 16S rRNA gene copy number (Table 2-3). In contrast to the low SR1 16S rRNA gene copy numbers in these two environments, SR1 16S rRNA genes were present in much higher numbers

and represented a higher fraction of the community in Zodletone Spring sediments and Sperm Pool microbial mat sample (Table 2-3). These results clearly indicate that under appropriate conditions, members of candidate division SR1 could form a significant fraction and an integral part of a microbial community and could potentially fulfill an as-yet-unidentified but potentially crucial role(s) within the mat ecosystem.

Visualization of candidate division SR1 using FISH. Alexa Fluor-labeled phylum-specific SR1 probes were used to visualize cells belonging to candidate division SR1 *in situ*. The persistence of solid particulates after successive centrifugation, coupled with the low proportion of SR1 cells (Table 2-3), and the auto-fluorescence of micro-eukaryotes within the bovine rumen, prevented us from effectively visualizing SR1 within this habitat. However, members of SR1 were identified in Duck Pond, Zodletone Spring, and Sperm Pool. Clone FISH suggested that a formamide concentration of 55% was ideal for both SR1-427 and SR1-1075 probes. No signal was detected when these two probes were used against pure cultures of *D. geothermicum*, and *V. souniana* at the appropriate formamide concentration. On the other hand, cells of *D. geothermicum* and *V. souniana* were successfully labeled and visualized using Univ1390 and Cren499 probes but not the nonEUB-338 probe (data not shown). These controls ensure that cells identified using SR1-427 and SR1-1075 probes in natural habitats with a formamide concentration of 55% in the hybridization buffer belong to candidate division SR1.

In Zodletone Spring sediments, Duck Pond sediments, and Sperm Pool mats, FISH using both SR1-227 and SR1-1075 probes separately or in combination revealed a filamentous morphotype with a highly variable length of 2.7 to 137.5 μm in Zodletone Spring (n = 15), 6.36 to 32 μm in Duck Pond (n = 32), and 6.4 to 109 μm (n = 44) in

Sperm Pool, but a constant cell width of 0.7 to 0.8 μm in all environments (Fig. 2-2). Cells sometimes appeared segmented, especially with DAPI staining (Fig. 2-2A). This cell morphology is strikingly similar to TM7 cells previously visualized in the oral cavity (56). The filamentous morphotype described above was the only cell morphology visualized in Zodletone Spring and Duck Pond sediments. However, within Sperm Pool microbial mats, a second SR1 cell morphotype was observed. These cells were bacilli with round ends (2.7 to 5.5 μm in length, 1.8 μm in width) (n = 28), which appeared mostly as single cells, but sometimes as doubles or chains (Fig. 2-2E and F).

It is interesting to note that in addition to observing two SR1 morphotypes in Sperm Pool mats, 16S rRNA analysis indicated that the SR1 community in the mat sample belonged to two distinct lineages (subgroups I and V in subphylum BD2-14 [Fig. 2-1]). This observation led us to hypothesize that each of these two lineages corresponds to a distinct cell morphotype. To examine this hypothesis, Alexa Fluor-labeled probes SR1-232 and SR1-575 were used to selectively target subgroups I and V, respectively. Using clone FISH, optimum formamide concentrations of 55 and 60% were determined for probes SR1-232 and SR1-575, respectively. Probe SR1-232, targeting subgroup I hybridized only to bacilli, indicating that members of SR1 BD2-14 group I are bacilli, while probe SR1-575 hybridized only to filaments, indicating that members of SR1 BD2-14 group I are filamentous (Fig. 2-3).

Table 2-1. Oligonucleotide primers or probes used in this study

Primer or probe	*E. coli* positions	Sequence (5′ to 3′)	Reference(s)	Function(s)	Optimized conditions	
					PCR annealing temp (°C)[a]	Formamide concn/NaCl concn[b]
SR1-427F	427-445	GAAGAMGMATGACGGTAC	This study	Clone library, qPCR	52[c]	NA
SR1-427R	427-445	GTACCGTCATKCKTCTTC	This study	FISH	NA	55/20
SR1-668R	668-686	CCACCKGAAATTCCACTA	This study	qPCR	52[c]	NA
SR1-914R	914-932	GYTCCCCCGCCTATCCYT	This study	Clone library	58[d]	NA
SR1-1075R	1075-1093	TTAACYRGACACCTTGCG	This study	Clone library, qPCR, FISH	52[e]	55/20
SR1I-232R	232-249	TAGCTGGTGGTCCGCGCC	This study	FISH	NA	55/20
SR1V-575R	575-592	TATGTCGGGCTACGGACA	This study	FISH	NA	60/14
Non-EUB-338F	320-338	ACATCCTACGGGAGGC	3, 4	qPCR, FISH	54[f]	35/80
EUB-518R	518-537	CGTATTACCGCGGCTGCTGG	53	qPCR	54[f]	NA
Bact-27F	9-27	GAGTTTGATCMTGGCTCAG	41	Clone library	58[d]	NA
Univ1390	1390-1410	GACGGGCGGTGTGTACAA	76	FISH	NA	0/900
Cren499	499-516	CCAGRCTTGCCCCCCGCT	11	FISH	NA	0/900

[a]Optimized PCR annealing temperatures are for primer pairs. NA, not applicable.
[b]The optimum formamide concentration in the hybridization buffer is shown as a percentage (vol/vol), and the corresponding optimum NaCl concentration in the washing buffer is shown in millimolar concentration. NA, not applicable.
[c]Primer pair SR1-427F/SR1-668R.
[d]Primer pair 27F/SR1-914R.
[e]Primer pair SR1-427F/SR1-1075R.
[f]Primer pair EUB-338/518R.

Table 2-2. Diversity within SR1 clone libraries generated in this study

Environment	No. of clones	No. of OTUs	Shannon-Weaver diversity index	Maximum sequence divergence	Avg nucleotide diversity (θ)	% Coverage
Zodletone Spring	158	18	2.26	0.25	0.12	98.7
Sperm Pool mat	79	8	1.37	0.26	0.11	94.9
Duck Pond	30	2	0.69	0.21	0.09	100.0
Bovine rumen	33	2	0.14	0.05	0.01	97.0

Table 2-3. SR1 quantification using qPCR in multiple environments

Environment	No. of 16S rRNA gene copies[a]		% SR1
	SR1	Total	
Zodletone Spring	1.50×10^7	1.29×10^8	11.6
Bovine rumen	1.24×10^6	1.55×10^9	0.08
Sperm Pool mat	1.22×10^{10}	2.51×10^{10}	48.7
Duck Pond	6.28×10^4	6.74×10^8	0.009

[a]Values are expressed as the number of 16S rRNA

Table 2-4. Basic geochemistry characteristics of anaerobic ecosystems examined in this study

Environment	Salinity (%)	pH	Temp (°C)	Sulfide concn (mM)	Sulfur conc. (mM)	Sulfate conc. (mM)	Nitrate conc. (mM)
Zodletone Spring	0.2	6.8	20	15.6	4.2	0.04	0
Bovine rumen	3	6	41	0.31	0.18	0.006	0
Sperm Pool mat	ND[a]	2.5	53	ND	ND	ND	0.085
Duck Pond	0	7	20	5.3	0.47	1.1	ND

[a]ND, not determined.

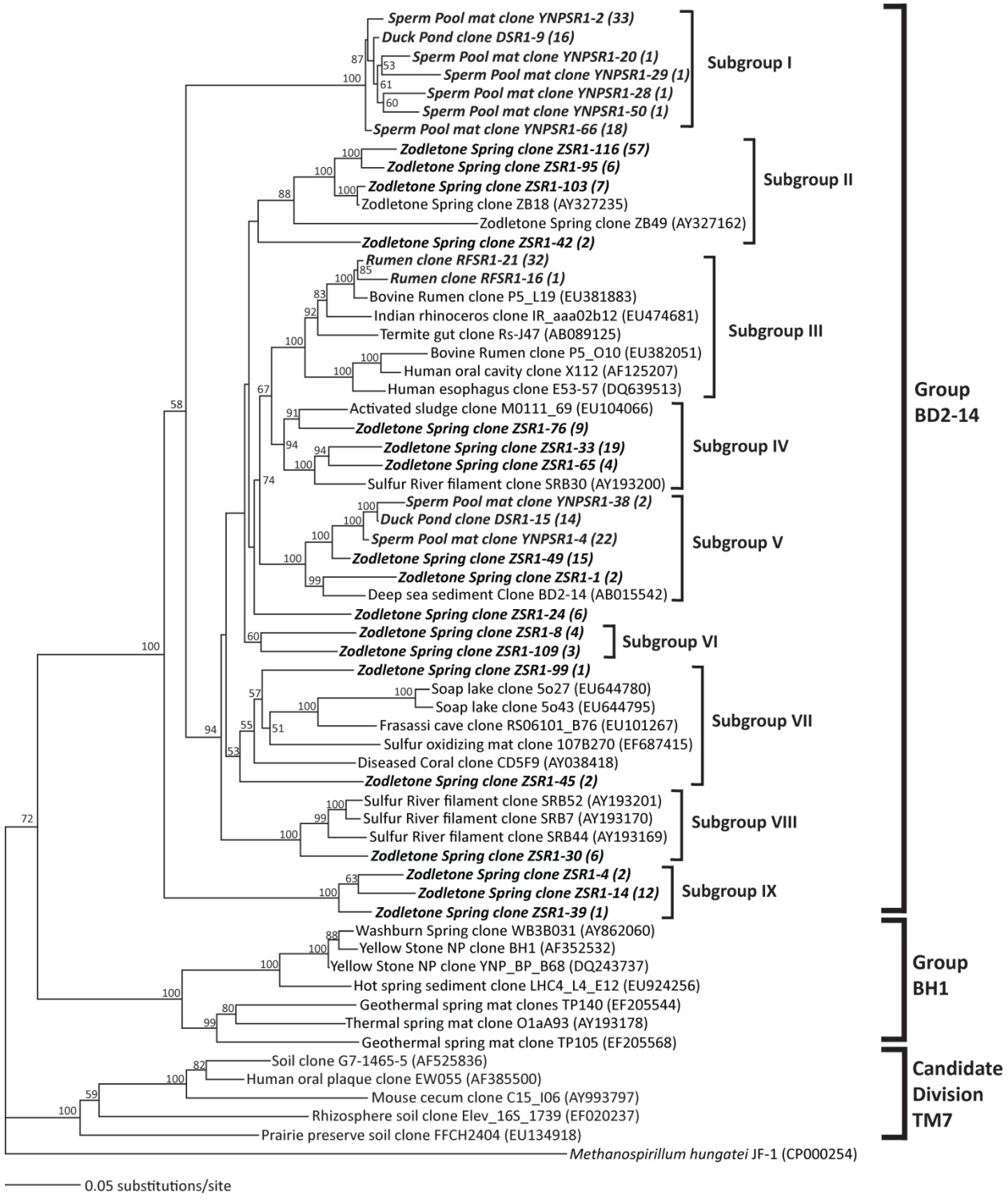


Figure 2-1. Phylogenetic tree based on 16S rRNA gene sequences of members of candidate division SR1 encountered in this study. Bootstrap values (expressed as percentages) are based on 1,000 replicates and are shown for branches with bootstrap values of more than 50%. Sequences generated in this study are in

boldface type, with the number of clones in each OTU reported in parentheses. GenBank accession numbers are shown in parentheses for other clones. The tree was constructed with the neighbor-joining algorithm with Jukes-Cantor corrections as described in Materials and Methods.

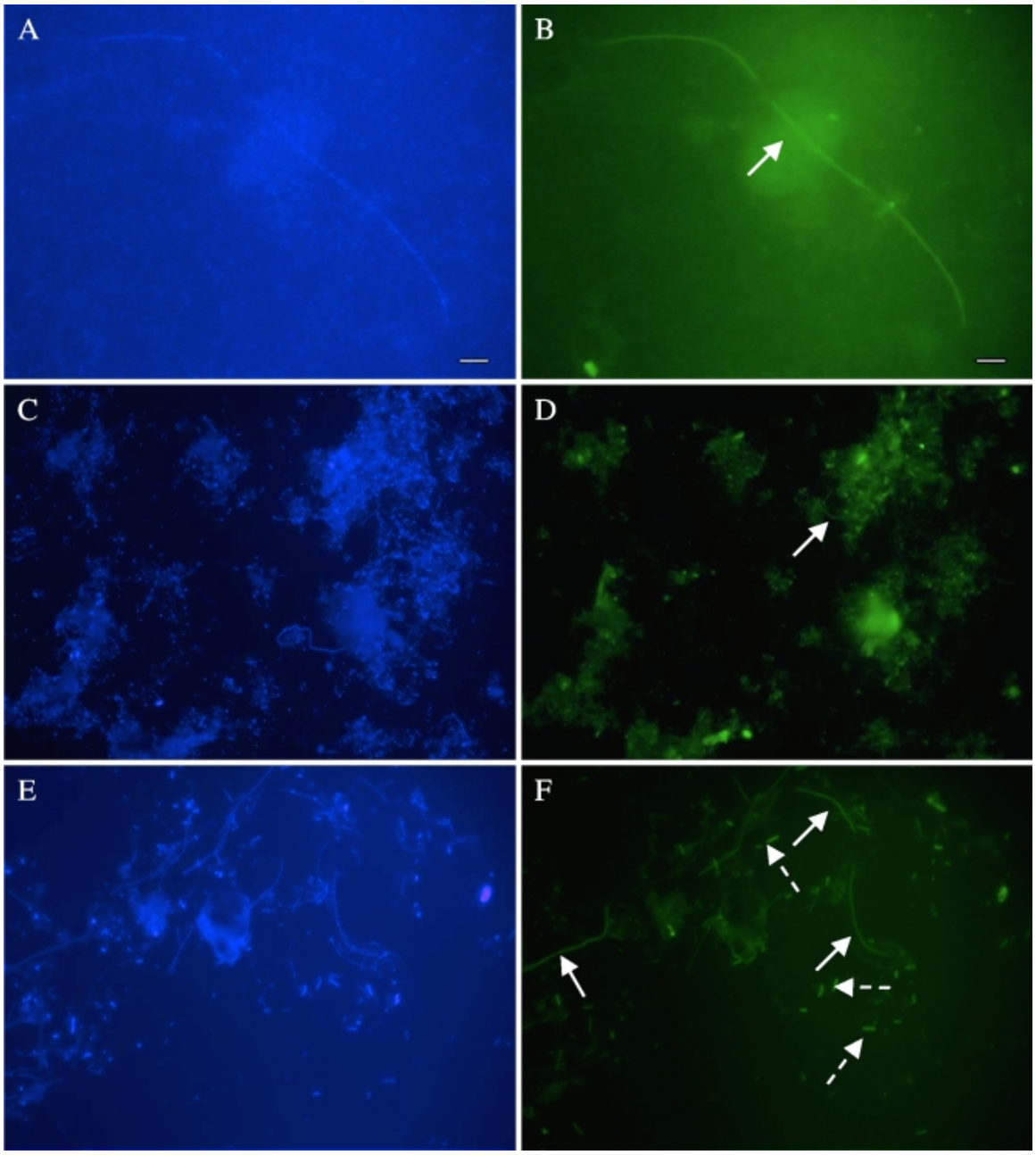


Figure 2-2. Whole-cell hybridization of paraformaldehyde-fixed cells with Alexa Fluor 488-labeled candidate division SR1-427 probe. Panels A and B, C and D, and E and F depict DAPI-stained cells (A, C, and E) versus FISH-labeled cells (B, D, and F) of anaerobic sulfur spring (Zodletone Spring) source sediment sample (A and B), anaerobic freshwater sediment sample (Duck Pond) (C and D), and Yellowstone National Park Sperm Pool microbial mat sample (E and F). Note the scarcity of SR1 cells in Zodletone Spring (B) and Duck Pond (D) ecosystems, as opposed to the abundance of SR1 cells in Sperm Pool ecosystem (F). The solid white arrows point to the filamentous SR1 morphotypes in all three environments, while the white broken arrows point to the bacillus morphotype in Sperm Pool. The bars (10 μm) in panels A and B applies to all panels of the figure.

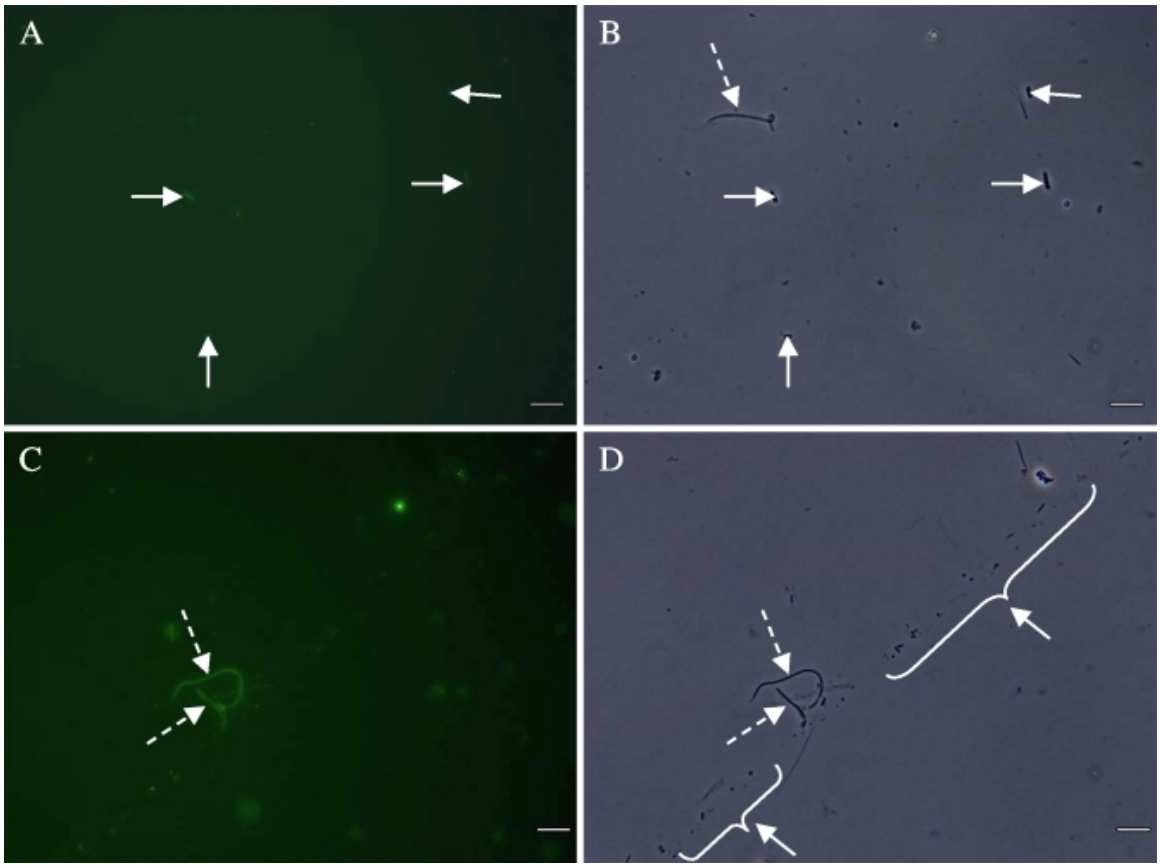


Figure 2-3. Whole-cell hybridization of Yellowstone microbial mat paraformaldehyde-fixed cells with Alexa Fluor 488-labeled probe SR1-232 (specific for SR1 BD2-14 subgroup I) (A) and Alexa Fluor 488-labeled probe SR1-575 (specific for SR1 BD2-14 subgroup V) (C). Panels B and D are the phase-contrast images for the same microscopic fields shown in panels A and C, respectively. Solid white arrows point to bacillus-like cells labeled with probe SR1-232 but not with probe SR1-575, while white broken arrows point to filamentous cells labeled with probe SR1-575 but not with probe SR1-232. Bars, 10 μm.

Discussion

In this study, multiple primers and probes were designed, evaluated, and utilized targeting candidate division SR1 to investigate the phylogenetic diversity, abundance, and cell morphologies of this as-yet-uncultured lineage in multiple habitats. It was shown that the scope of phylogenetic diversity within candidate division SR1 is much broader than previously implied based on abundance of SR1 clones in clone libraries constructed using general bacterial primers. This study also successfully developed and implemented qPCR and FISH protocols for quantification and visualization of SR1 cells *in situ* and demonstrated that members of candidate division SR1 have a limited morphotypic diversity in all environments examined, with either one or two morphotypes encountered in all environments examined.

The development of primers and probes targeting a specific microbial lineage allowed for in-depth, targeted diversity surveys of the lineage (5, 6, 9, 38), spatial and temporal monitoring of the target lineage within a specific ecosystem, as well as between various ecosystems (9, 10) and detection of cell morphologies and quantification through FISH and qPCR was also demonstrated (8, 24, 25, 38). Further, group-specific oligonucleotides are crucial for targeted genomic and metagenomic-based approaches, *e.g.*, for the screening of metagenomic libraries (61), and the implementation of single-cell-based genomic approaches (60), as well as for targeted enrichment and isolation of novel representatives of these lineages (14, 28, 40, 67, 69). The large amount of information regarding the ecology, metabolic abilities, and growth characteristics of lineages for which primers and probes have been developed and implemented is in stark contrast to the dearth of information regarding lineages for which no similar effort have

been made. This study thus represents a useful first step that aims to provide the tools necessary for better targeting SR1 cells and genomic fragments in natural environments, enrichments, as well as in metagenomic libraries.

This study greatly expands our knowledge of the phylogenetic diversity within candidate division SR1 and identifies multiple novel lineages within this candidate division. With the exception of the Sperm Pool microbial mats, which were selectively chosen for this study based on prior knowledge of the high proportion of the SR1 cells within this specific layer of this microbial mat, diversity (as measured by richness, average nucleotide diversity, nucleotide range, and affiliation with SR1 lineages), and abundance (number of 16S rRNA gene copies/gram or milliliter) was highest in Zodletone Spring sediments and lowest in Duck Pond samples (Table 2-3). Factors controlling diversity and abundance of members of SR1 in various ecosystems have not yet been elucidated. However, it has previously been speculated, based on ecological distribution, that members of SR1 are involved in sulfur transformation (32), *e.g.*, chemolithotrophic sulfide oxidation (58). Sulfide, elemental sulfur, and sulfate measurements (Table 2-4) in all three environments show that sulfur and sulfide levels (but not sulfate) were also highest in Zodletone Spring than in Duck Pond and bovine rumen fluid samples. This positive correlation between sulfur and sulfide levels on one side and SR1 numbers further attests to the potential role of members of SR1 in sulfur transformation. Information regarding Sperm Pool geochemistry is not available, but sulfur cycling and high sulfide and sulfur levels are known to be associated with various pools in Yellowstone National Park (7, 50, 68). Indeed, in clone libraries constructed using general bacterial primers from Sperm Pool microbial mats, SR1 sequences were

associated with bacteria known to be involved in sulfur transformation within the mat (B. Fathepure, personal communication).

On the basis of this information, it is hypothesized that members of candidate division SR1 have lower *in situ* growth rates than other more-competitive sulfur-metabolizing microorganisms (mainly within the Proteobacteria). A constant supply of fairly high levels of sulfur and sulfide (*e.g.*, in Zodletone Spring), especially when coupled to extreme conditions (*e.g.*, low pH and high temperature in Sperm Pool), will result in easing the competition between members of SR1 and other sulfur-metabolizing microorganisms and SR1 abundance and diversity will then increase. On the other hand, in environments with limited supply and low levels of sulfate and elemental sulfur (*e.g.*, bovine rumen, Duck Pond), members of SR1 will be outcompeted by more efficient sulfur metabolizers, resulting in lower abundance and reduced diversity. While a sulfur-based metabolism for members of candidate division SR1 appears plausible, the nature of such sulfur-based metabolism (whether members of SR1 are chemolithotrophic sulfide or sulfur oxidizers as suggested by Perner et al. (58) or chemoorganotrophic or autotrophic sulfur reducers) is not yet clear. In addition, it is plausible that various members of candidate division SR1 might have multiple distinct metabolic abilities. Further, it is entirely possible that similar to certain bacterial and archaeal genera (*e.g.*, *Beggiatoa* and *Thermoproteus*), some (or all) SR1 strains might grow mixotrophically, switching between chemolithotrophy and heterotrophy depending on the surrounding environmental conditions (26, 30).

Acknowledgments. I would like to thank Kathleen E. Duncan, Babu Z. Fathepure, and James Cullison for soil, Sperm Pool, and rumen samples, respectively. I thank Yao Tan, Yiran Dong, and Elizabeth Butler for help with the sulfide, sulfur, and sulfate analyses; Cody Sheik for help with qPCR analysis; and Fares Z. Najar for DNA sequencing. This work was supported by NSF MO grants MCB_0240683 and EF0801858 and OSU start-up funds to M.S.E. 5609–5621.

References

1. **Alm, E. W., D. B. Oerther, N. Larsen, D. A. Stahl, and L. Raskin.** 1996. The oligonucleotide probe database. Appl. Environ. Microbiol. **62:**3557–3559.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **Amann, R., L. Krumholz, and D. Stahl.** 1990. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. J. Bacteriol. **172:**762–770.
4. **Amann, R. I., W. Ludwig, and K.-H. Schleifer.** 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143–169.
5. **Barns, S. M., E. C. Cain, L. Sommerville, and C. R. Kuske.** 2007. Acidobacteria phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. Appl. Environ. Microbiol. **73:**3113–3116.
6. **Blackwood, C. B., A. Oaks, and J. S. Buyer.** 2005. Phylum- and class-specific PCR primers for general microbial community analysis. Appl. Environ. Microbiol. **71:**6193–6198.
7. **Blank, C. E., S. L. Cady, and N. R. Pace.** 2002. Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. Appl. Environ. Microbiol. **68:**5123–5135.

8. **Boetius, A., K. Ravenschlag, C. J. Schubert, D. Rickert, F. Widdel, A. Gleseke, R. Amann, B. B. Jorgensen, U. Witte, and O. Pfannkuche.** 2000. A marine microbial consortium apparently mediating anaerobic oxidation of methane. Nature **407:**623–626.

9. **Buckley, D. H., V. Huangyutitham, T. A. Nelson, A. Rumberger, and J. E. Thies.** 2006. Diversity of Planctomycetes in soil in relation to soil history and environmental heterogeneity. Appl. Environ. Microbiol. **72:**4522–4531.

10. **Buckley, D. H., and T. M. Schmidt.** 2001. Environmental factors influencing the distribution of rRNA from Verrucomicrobia in soil. FEMS Microbiol. Ecol. **35:**105–112.

11. **Burggraf, S., T. Mayer, R. Amann, S. Schadhauser, C. R. Woese, and K. O. Stetter.** 1994. Identifying members of the domain Archaea with rRNAtargeted oligonucleotide probes. Appl. Environ. Microbiol. **60:**3112–3119.

12. **Chouari, R., D. Le Paslier, C. Dauga, P. Daegelen, J. Weissenbach, and A. Sghir.** 2005. Novel major bacterial candidate division within a municipal anaerobic sludge digester. Appl. Environ. Microbiol. **71:**2145–2153.

13. **Cole, J. R., R. J. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity, and J. M. Tiedje.** 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res. **35:**D169–D172.

14. **Connon, S. A., and S. J. Giovannoni.** 2002. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. Appl. Environ. Microbiol. **68:**3878–3885.

15. **Corre, E., A.-L. Reysenbach, and D. Prieur.** 2001. ε-Proteobacterial diversity from a deep-sea hydrothermal vent on the Mid-Atlantic Ridge. FEMS Microbiol. Lett. **205:**329–335.

16. **Daffonchio, D., S. Borin, T. Brusa, L. Brusetti, P. W. J. J. van der Wielen, H. Bolhuis, M. M. Yakimov, G. D'Auria, L. Giuliano, D. Marty, C. Tamburini, T. J. McGenity, J. E. Hallsworth, A. M. Sass, K. N. Timmis, A. Tselepides, G. J. de Lange, A. Hu¨bner, J. Thomson, S. P. Varnavas, F. Gasparoni, H. W. Gerber, E. Malinverno, and C. Corselli.** 2006. Stratified prokaryote network in the oxic/anoxic transition of a deep-sea halocline. Nature **440:**203–207.

17. **DeSantis, T. Z., P. Hugenholtz, K. Keller, E. L. Brodie, N. Larsen, Y. M. Piceno, R. Phan, and G. L. Andersen.** 2006. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. Nucleic Acids Res. **34:**W394–W399.

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

19. Reference deleted by AEM.

20. **Dojka, M. A., J. K. Harris, and N. R. Pace.** 2000. Expanding the known diversity and environmental distribution of an uncultured phylogenetic division of bacteria. Appl. Environ. Microbiol. **66:**1617–1621.

21. **Dojka, M. A., P. Hugenholtz, S. K. Haack, and N. R. Pace.** 1998. Microbial diversity in a hydrocarbon- and chlorinated-solvent-contaminated aquifer undergoing intrinsic bioremediation. Appl. Environ. Microbiol. **64:**3869–3877.

22. **Elshahed, M. S., J. M. Senko, F. Z. Najar, S. M. Kenton, B. A. Roe, T. A. Dewers, J. R. Spear, and L. R. Krumholz.** 2003. Bacterial diversity and sulfur cycling in a mesophilic sulfide-rich spring. Appl. Environ. Microbiol. **69:**5609–5621.

23. **Elshahed, M. S., N. H. Youssef, A. M. Spain, C. Sheik, F. Z. Najar, L. O. Sukharnikov, B. A. Roe, J. P. Davis, P. D. Schloss, and L. R. Krumholz.** 2008. Novelty and uniqueness patterns of rare members of the soil biosphere. Appl. Environ. Microbiol. **74:**5422–5428.

24. **Fierer, N., J. A. Jackson, R. Vilgalys, and R. B. Jackson.** 2005. Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. Appl. Environ. Microbiol. **71:**4117–4120.

25. **Fieseler, L., M. Horn, M. Wagner, and U. Hentschel.** 2004. Discovery of the novel candidate phylum "Poribacteria" in marine sponges. Appl. Environ. Microbiol. **70:**3724–3732.

26. **Fischer, F., W. Zillig, K. O. Stetter, and G. Schreiber.** 1983. Chemolithoautotrophic metabolism of anaerobic extremely thermophilic archaebacteria. Nature **301:**511–513.

27. **Frias-Lopez, J., A. L. Zerkle, G. T. Bonheyo, and B. W. Fouke.** 2002. Partitioning of bacterial communities between seawater and healthy, black band diseased, and dead coral surfaces. Appl. Environ. Microbiol. **68:**2214–2228.

28. **Girguis, P. R., A. E. Cozen, and E. F. DeLong.** 2005. Growth and population dynamics of anaerobic methane-oxidizing archaea and sulfate-reducing bacteria in a continuous-flow bioreactor. Appl. Environ. Microbiol. **71:**3725–3733.

29. **Good, I. J.** 1953. The population frequencies of species and the estimation of population parameters. Biometrika **40:**237–264.

30. **Gude, H., W. Strohl, and J. Larkin.** 1981. Mixotrophic and heterotrophic growth of Beggiatoa alba in continuous culture. Arch. Microbiol. **129:**357–360.

31. **Hall, J. R., K. R. Mitchell, O. Jackson-Weaver, A. S. Kooser, B. R. Cron, L. J. Crossey, and C. D. Takacs-Vesbach.** 2008. Molecular characterization of the diversity and distribution of a thermal spring microbial community by using rRNA and metabolic genes. Appl. Environ. Microbiol. **74:**4910–4922.

32. **Harris, J. K., S. T. Kelley, and N. R. Pace.** 2004. New perspective on uncultured bacterial phylogenetic division OP11. Appl. Environ. Microbiol. **70:**845–849.

33. **Hongoh, Y., P. Deevong, S. Hattori, T. Inoue, S. Noda, N. Noparatnaraporn, T. Kudo, and M. Ohkuma.** 2006. Phylogenetic diversity, localization, and cell morphologies of members of the candidate phylum TG3 and a subphylum in the phylum Fibrobacteres, recently discovered bacterial groups dominant in termite guts. Appl. Environ. Microbiol. **72:**6780–6788.

34. **Hongoh, Y., M. Ohkuma, and T. Kudo.** 2003. Molecular analysis of bacterial microbiota in the gut of the termite Reticulitermes speratus (Isoptera; Rhinotermitidae). FEMS Microbiol. Ecol. **44:**231–242.

35. **Hsieh, Y. P., and Y. N. Shieh.** 1997. Analysis of reduced inorganic sulfur by diffusion methods: improved apparatus and evaluation for sulfur isotopic studies. Chem. Geol. **137:**255–261.

36. **Hugenholtz, P., B. M. Goebel, and N. R. Pace.** 1998. Impact of cultureindependent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. **180:**4765–4774.

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

38. **Hugenholtz, P., G. W. Tyson, R. I. Webb, A. M. Wagner, and L. L. Blackall.** 2001. Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. Appl. Environ. Microbiol. **67:**411–419.

39. **Janssen, P. H.** 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Appl. Environ. Microbiol. **72:**1719–1728.

40. **Konneke, M., A. E. Bernhard, J. R. De La Torre, C. B. Walker, J. B. Waterbury, and D. A. Stahl.** 2005. Isolation of an autotrophic ammoniaoxidizing marine archaeon. Nature **437:**543–546.

41. **Lane, D. J.** 1991. 16S/23S rRNA sequencing, p. 115–174. In E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, Chichester, United Kingdom.

42. **Ley, R. E., M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight, and J. I.**

Gordon. 2008. Evolution of mammals and their gut microbes. Science **320:**1647–1651.

43. **Li, L., C. Kato, and K. Horikoshi.** 1999. Bacterial diversity in deep-sea sediments from different depths. Biodivers. Conserv. **8:**659–677.

44. **Lipson, D. A., and S. K. Schmidt.** 2004. Seasonal changes in an alpine soil bacterial community in the Colorado Rocky Mountains. Appl. Environ. Microbiol. **70:**2867–2879.

45. **Ludwig, W., O. Strunk, R. Westram, L. Richter, H. Meier, Y. Kumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A. W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. Ko¨nig, T. Liss, R. Lu¨mann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, and K.-H. Schleifer.** 2004. ARB: a software environment for sequence data. Nucleic Acids Res. **32:**1363–1371.

46. **Macalady, J. L., S. Dattagupta, I. Schaperdoth, D. S. Jones, G. K. Druschel, and D. Eastman.** 2008. Niche differentiation among sulfur-oxidizing bacterial populations in cave waters. ISME J. **2:**590–601.

47. **Magurran, A. E.** 2004. Measuring biological diversity. Blackwell Science Ltd., Oxford, United Kingdom.

48. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y.-J. Chen, Z. Chen, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380.

49. **Martin, A. P.** 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. Appl. Environ. Microbiol. **68:**3673–3682.

50. **Meyer-Dombard, D. R., E. L. Shock, and J. P. Amend.** 2005. Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. Geobiology **3:**211–227.

51. **Nakagawa, S., K. Takai, F. Inagaki, H. Chiba, J.-I. Ishibashi, S. Kataoka, H. Hirayama, T. Nunoura, K. Horikoshi, and Y. Sako.** 2005. Variability in microbial community and venting chemistry in a sediment-hosted backarc hydrothermal system: impacts of subseafloor phase-separation. FEMS Microbiol. Ecol. **54:**141–155.

52. **Nakajima, H., Y. Hongoh, S. Noda, Y. Yoshida, R. Usami, T. Kudo, and M. Ohkuma.** 2006. Phylogenetic and morphological diversity of Bacteroidales members associated with the gut wall of termites. Biosci. Biotechnol. Biochem. **70:**211–218.

53. **Nogales, B., E. R. B. Moore, W.-R. Abraham, and K. N. Timmis.** 1999. Identification of the metabolically active members of a bacterial community in a polychlorinated biphenyl-polluted moorland soil. Environ. Microbiol. **1:**199–212.

54. Reference deleted by AEM.

55. **Omoregie, E. O., V. Mastalerz, G. de Lange, K. L. Straub, A. Kappler, H. Roy, A. Stadnitskaia, J.-P. Foucher, and A. Boetius.** 2008. Biogeochemistry and community composition of iron- and sulfur-precipitating microbial mats at the Chefren mud volcano (Nile Deep Sea Fan, Eastern Mediterranean). Appl. Environ. Microbiol. **74:**3198–3215.

56. **Ouverney, C. C., G. C. Armitage, and D. A. Relman.** 2003. Single-cell enumeration of an uncultivated TM7 subgroup in the human subgingival crevice. Appl. Environ. Microbiol. **69:**6294–6298.

57. **Pace, N. R.** 1997. A molecular view of microbial diversity and the biosphere. Science **276:**734–740.

58. **Perner, M., R. Seifert, S. Weber, A. Koschinsky, K. Schmidt, H. Strauss, M. Peters, K. Haase, and J. F. Imhoff.** 2007. Microbial CO2 fixation and sulfur cycling associated with low-temperature emissions at the Lilliput hydrothermal field, southern Mid-Atlantic Ridge (9° S). Environ. Microbiol. **9:**1186–1201.

59. **Pernthaler, A., A. E. Dekas, C. T. Brown, S. K. Goffredi, T. Embaye, and V. J. Orphan.** 2008. Diverse syntrophic partnerships from deep-sea methane vents revealed by direct cell capture and metagenomics. Proc. Natl. Acad. Sci. USA **105:**7052–7057.

60. **Podar, M., C. B. Abulencia, M. Walcher, D. Hutchison, K. Zengler, J. A. Garcia, T. Holland, D. Cotton, L. Hauser, and M. Keller.** 2007. Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl. Environ. Microbiol. **73:**3205–3214.

61. **Quaiser, A., T. Ochsenreiter, C. Lanz, S. C. Schuster, A. H. Treusch, J. Eck, and C. Schleper.** 2003. Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. Mol. Microbiol. 50:563–575.

62. **Rappe, M. S., and S. J. Giovannoni.** 2003. The uncultured microbial majority. Annu. Rev. Microbiol. **57:**369–394.

63. **Reysenbach, A. L., K. Longnecker, and J. Kirshtein.** 2000. Novel bacterial and archaeal lineages from an *in situ* growth chamber deployed at a MidAtlantic Ridge hydrothermal vent. Appl. Environ. Microbiol. **66:**3798–3806.

64. **Schloss, P. D., and J. Handelsman.** 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. **71:**1501–1506.

65. **Schramm, A., B. M. Fuchs, J. L. Nielsen, M. Tonolla, and D. A. Stahl.** 2002. Fluorescence *in situ* hybridization of 16S rRNA gene clones (Clone-FISH) for probe validation and screening of clone libraries. Environ. Microbiol. **4:**713–720.

66. **Senko, J. M., B. S. Campbell, J. R. Henricksen, M. S. Elshahed, T. A. Dewers, and L. R. Krumholz.** 2004. Barite deposition mediated by photootrophic sulfide-oxidizing bacteria. Geochim. Cosmochim. Acta **68:**773–780.

67. **Simon, H. M., C. E. Jahn, L. T. Bergerud, M. K. Sliwinski, P. J. Weimer, D. K. Willis, and R. M. Goodman.** 2005. Cultivation of mesophilic soil crenarchaeotes in enrichment cultures from plant roots. Appl. Environ. Microbiol. **71:**4751–4760.

68. **Spear, J. R., J. J. Walker, T. M. McCollom, and N. R. Pace.** 2005. Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. Proc. Natl. Acad. Sci. USA **102:**2555–2560.

69. **Stevenson, B. S., S. A. Eichorst, J. T. Wertz, T. M. Schmidt, and J. A. Breznak.** 2004. New strategies for cultivation and detection of previously uncultured microbes. Appl. Environ. Microbiol. **70:**4748–4755.

70. **Stoica, E., and G. J. Herndl.** 2007. Contribution of Crenarchaeota and Euryarchaeota to the prokaryotic plankton in the coastal northwestern Black Sea. J. Plankton Res. 29:699–706.

71. **Thompson, J. D., T. J. Gibson, F. Pleiomiak, F. Jeanmougin, and D. G. Higgins.** 1997. The CLUSTAL_X interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:**4876–4882.

72. **Ulrich, G. A., L. R. Krumholz, and J. M. Suflita.** 1997. A rapid and simple method for estimating sulfate reduction activity and quantifying inorganic sulfides. Appl. Environ. Microbiol. **63:**1627–1630.

73. **van der Wielen, P. W., H. Bolhuis, S. Borin, D. Daffonchio, C. Corselli, L. Giuliano, G. D'Auria, G. J. de Lange, A. Huebner, S. P. Varnavas, J. Thomson, C. Tamburini, D. Marty, T. J. McGenity, and K. N. Timmis.** 2005. The enigma of prokaryotic life in deep hypersaline anoxic basins. Science **307:**121–123.

74. **Wilhartitz, I., R. L. Mach, E. Teira, T. Reinthaler, G. J. Herndl, and A. H. Farnleitner.** 2007. Prokaryotic community analysis with CARD-FISH in comparison with FISH in ultra-oligotrophic ground- and drinking water. J. Appl. Microbiol. **103:**871–881.

75. **Xenoulis, P. G., B. Palculict, K. Allenspach, J. M. Steiner, A. M. Van House, and J. S. Suchodolski.** 2008. Molecular-phylogenetic characterization of microbial communities imbalances in the small intestine of dogs with inflammatory bowel disease. FEMS Microbiol. Ecol. **66:**579–589.

76. **Zheng, D., E. W. Alm, D. A. Stahl, and L. Raskin.** 1996. Characterization of universal small rRNA hybridization probes for quantitative molecular microbial ecology studies. Appl. Environ. Microbiol. **62:**4504–4513.

CHAPTER III

BACTERIAL COMMUNITIES ASSOCIATED WITH PRODUCTION FACILITIES OF TWO NEWLY DRILLED THERMOGENIC NATURAL GAS WELLS IN THE BARNETT SHALE (TEXAS, USA)

Davis, J.P., C.G. Struchtemeyer, and M.S. Elshahed. 2012. Microbial Ecology

DOI: 10.1007/s00248-012-0073-3.

Abstract

The bacterial communities in the gas-water separator and water storage tank of two newly drilled natural gas wells in the Barnett Shale in north central Texas, were monitored using a 16S rRNA gene pyrosequencing approach over a period of six months. Overall, the communities were composed mainly of moderately halophilic and halotolerant members of the phyla Firmicutes and Proteobacteria (Classes: Beta-, Gamma-, and Epsilonproteobacteria) in both wells at all sampling times and locations. Many of the observed lineages were encountered in prior investigations of microbial communities from various fossil fluid formations and production facilities. In all of the samples, multiple H_2S-producing lineages were encountered; belonging to the sulfate- and sulfur-reducing class Deltaproteobacteria, order Clostridiales, and phylum Synergistetes, as well as the thiosulfate-reducing order Halanaerobiales. The bacterial communities from the separator and tank samples bore little resemblance to the bacterial communities in the drilling mud and hydraulic-fracture waters that were used to drill these wells, suggesting the *in situ* development of the unique bacterial communities in such well components was in response to the prevalent geochemical conditions present. Conversely, comparison of the bacterial communities on temporal and spatial scales suggested the establishment of a core microbial community in each sampled location. The results provide the first overview of bacterial dynamics and colonization patterns in newly drilled, thermogenic natural gas wells, and highlights patterns of spatial and temporal variability observed in bacterial communities in natural gas production facilities.

Introduction

It is well understood that the current consumption rate of crude oil is leading to an economic impasse, and the increased awareness of climate change has spurred an increase for more abundant and cleaner alternative fuel sources (112). Natural gas is viewed as a favorable alternative fuel because of its abundance, relatively low cost, and because its combustion results in lower greenhouse gas emissions compared to other fossil fuels (45, 120). Since 2000, the contribution of natural gas from geological shale to total U.S. natural gas supplies has increased from 1% to approximately 20%, and is expected to increase to nearly 50% by 2025 (51, 77).

The Barnett Shale in north central Texas is one of the most important shale gas reservoirs in the United States, and accounts for over a third of the total U.S. shale gas production (77). The gas in the Barnett Shale is completely thermogenic in origin (43). The gas was formed when formation temperatures exceeded 175°C, which caused cracking of the kerogen and petroleum that was present in the formation (43). Due to these extreme conditions, no indigenous bacterial populations are present in the Barnett Shale (51). Current temperatures of natural gas wells in the Barnett Shale have cooled to around 82°C, which could theoretically support the growth of microorganisms (8) . However repopulation of the shale by bacteria was likely prevented by the nanodarcy permeability and extremely small average pore throat size (typically < 0.005 μm) of the shale (8, 9).

Although the gas from the Barnett Shale is abiotic in origin, there have been several reports of biogenic sulfide production and biocorrosion in various production

facilities *e.g.* gas water separator and water-storage tanks in well sites located throughout the Barnett Shale (29). There are several possible sources of bacterial populations that could account for such effects including: drilling mud utilized during well drilling process, waters used during the hydraulic-fracturing processes, as well as secondary microbial development and colonization by above surface, airborne microorganisms (55, 99, 101). However in spite of the increased importance of natural gas in the United States, only few studies have examined the microbial ecology of natural gas wells and their production components (gas-water separators and production-water storage tanks) (49). Further, all previous studies viewed the microbial communities in a single time point and at one specific location, which only provides a snapshot of the microbial community and hence does not offer spatiotemporal insights on the stability and dynamics of microbial communities in natural gas production facilities.

In this study, the bacterial communities associated with production facilities in two newly drilled wells were investigated to better understand this little-studied, yet vital system. Natural gas emerges from the well in a gas-water mixture, and the water portion is known as produced or production water. Production water was collected at multiple time points over a period of six months from two locations at each well. The first location that was sampled was the gas-water separator (hereafter referred to as the separator), which separates the production water from natural gas after it emerges from the wellhead. The second location that was sampled was the production water storage tank, which is used to house production water after natural gas and water separation occurs (hereafter referred to as the tank). These two wells have also been the subject of prior investigations

in which the microbial communities in the drilling mud utilized in the drilling process, as well as the microbial community introduced during the hydraulic-fracturing processes were characterized using pyrosequencing, allowing us to address questions regarding the role of such processes in establishing microbial communities in production natural gas facilities (99, 101). Our goals in this study are: (1) to identify and document the origins and phylogenetic diversities of bacterial communities that developed in separators and production water storage tanks at these two newly drilled wells, (2) to compare the bacterial communities in the separator samples and production water tank samples (from both wells) to determine if the communities were highly similar to one another and therefore representative of the bacterial communities in other separators and water storage tanks located throughout the Barnett Shale, (3) to compare the bacterial community from a single location at various time points to document temporal dynamics associated with natural gas production within a single location, and (4) to identify the presence, nature, and proportion of sulfidogenic lineages in the bacterial communities from these production facilities.

Materials and Methods

Description of sampling sites. Production water samples were collected from the separator and tank at two newly drilled natural gas wells located in the Barnett Shale. The two natural gas wells are located in Denton (SM well) and Johnson (AI well) counties in north central Texas. Both of these wells were drilled and hydro-fractured during the summer of 2009 and started actively producing natural gas in October 2009. Water samples from the gas-water separator and the produced water storage tank were collected from October 2009 through March 2010. The separator and tank were located on the surface, near (approximately 20 meters) the well-head. The separators were vertical units with an approximate diameter of 0.7 m and 2.5 to 3 m in height. The tanks had a total capacity of approximately 75 m^3. All samples were collected from sampling ports located at the bottom of separators and tanks with sterile Nalgene bottles that were filled to capacity, frozen on dry ice while in transit to the laboratory, and stored at -20°C upon arrival at the laboratory. The temperature of the water (from the separator and tank) was near ambient environmental temperatures when collected (data not shown). Drilling mud and hydro-fracturing fluid (frac-water) were collected and described in detail elsewhere (99, 101).

Geochemistry of the production waters. The geochemical properties of the produced water from the tanks and separators were measured approximately 24 hours after returning to the laboratory. Samples were not filtered prior to geochemical analysis. Total dissolved solids (TDS) and pH were measured using the ExStik® II pH/conductivity meter (Extech Instruments Corp., Waltham, MA). Salinity was measured

with the VWR Portable Refractometer (VWR International, LLC, West Chester, PA). Alkalinity, ferrous iron, total iron, sulfate, nitrate, and nitrite were all measured using Hach test kits (Hach Co., Loveland, CO). Samples for sulfide analysis were sampled separately and fixed using zinc acetate immediately on site to precipitate soluble and most insoluble sulfides (*e.g.* FeS) as zinc sulfide, and sulfide was subsequently measured using the methylene blue assay (16).

DNA extraction, 16S rRNA gene amplification, and pyrosequencing. Five-hundred milliliters of each water sample was centrifuged for 45 minutes at 10,000 RPM at 4°C. The cell pellets were suspended in 200 mL of sterile 1x TE buffer and DNA was extracted with FastDNA® Spin kit for soil (QBiogene, Carlsbad, CA). The 16S rRNA gene from the production water DNA was amplified using bacterial-specific, barcoded primers with FLX platform adaptors. The forward primer was modified so that it contained the 454 Roche adapter A (GCCTCCCTCGCGCCATCAG) followed by an 8 bp barcode sequence, a two base linker sequence (CA), and the conserved bacterial primer 338F (40, 66). A unique 8 bp barcode sequence was used for each sample (data not shown). The reverse primer was modified so that it contained the 454 Roche adapter B (GCCTTGCCAGCCCGCTCAG) followed by a 2 bp linker (TC) and the conserved bacterial primer 518R(66). PCR was conducted in a 50 mL volume containing (final concentration): 0.15-0.2 ng/mL template DNA, 1x goTaq PCR Buffer (Promega, Madison, WI), 2.5 mM MgSO4, 0.2 mM dNTPs mixture, 0.4 mM each of both the forward and reverse primers, and 2.5U of goTaq Flexi DNA Polymerase (Promega). PCR amplification was conducted using the following cycling conditions: initial denaturing

step for 5 min. at 95°C, followed by a 30 cycle step with denaturing at 95°C for 45 sec., annealing at 54°C for 45 sec., and elongation at 72°C for 1.5 min, and followed by a final step of elongation at 72°C for 15 min. Positive PCR products were pooled, then purified using Purelink™ PCR purification kit (Invitrogen, Carlsbad, CA). DNA was sequenced at the Environmental Genomics Core Facility (EnGenCore) at the University of South Carolina, Columbia, SC using FLX technology.

Sequence Processing. The software package, mothur, was used for processing the obtained pyrosequencing reads (86). The raw sequences were screened and low quality sequences were removed based on: quality (>25 quality score threshold), minimum nucleotide length (>80 bp), maximum homopolymers (>8 bases), and sequences with ambiguous bases (N). In addition, sequences without an exact match to the primer sequence were also removed. The remaining high-quality sequences were aligned in mothur using a furthest neighbor algorithm against Greengenes database template as described in reference (21). A pair wise distance matrix was created from the alignment with mothur, and the distance matrix was then used to assign the sequences into Operational Taxonomic Units (OTUs), using a 97% sequence similarity cutoff from the furthest neighbor. The OTUs were then classified using Greengenes classifier program (20, 21). Chao and ACE species richness indices and Good's coverage were also calculated using mothur (13, 14, 108).

Spatial and temporal comparisons of bacterial populations in tanks and separator samples. Multiple pair wise diversity estimates were used to compare community membership and structure between all the datasets examined in this study.

Shared OTUs between all possible pairs of samples were identified by creating a joint distance matrix of all sequences within all datasets in mothur and using this matrix to generate a shared OTUs file. The file was used to conduct pair wise comparisons between every possible pair of samples using both qualitative similarity indices (those that use presence/absence data) *e.g.* Sørensen index, as well as quantitative indices (those that take OTUs abundance or relative abundance into consideration) *e.g.* abundance-based Sørensen index (94). Non-metric multidimensional scaling (NMDS) plots for communities using Sørensen similarity indices were created using the NMDS algorithm in mothur.

Three different groups of pair wise comparisons were conducted. First, a comparison was conducted of the bacterial communities obtained from both tanks and separators at all time points to the microbial community characterized in drilling mud formulations used in the well drilling process in each well, as well as to the microbial community characterized in water obtained from the hydraulic fracturing procedure at both well locations (99, 101). These comparisons were conducted to identify the potential contribution of microorganisms introduced during the drilling and hydraulic-fracturing process on the microbial community developing in above ground facilities post-production. Second, temporal comparisons of sequences obtained from a specific location (*e.g.* AI tank in October Vs AI tank in November) were conducted to quantify changes within a specific location over time. Finally, spatial comparisons of sequences obtained at the same sampling time (*e.g.* AI tank Vs AI separator samples in October 2009, or AI

tank Vs SM tank in October 2009) were conducted to quantify relatedness of bacterial communities at two distinct locations at the same time within each well.

Results

Sampling and geochemical characterization. A total of seventeen samples were collected and analyzed. Six tank samples and four separator samples were collected from the AI well, and five tank samples and 2 separator samples were collected from the SM well. Results of geochemical analysis from all samples are shown in Table 3-1. All samples were characterized by a slightly acidic pH. AI well separator and tank samples had salinities levels around 8% with a gradual increase over time. Similarly, within SM tank samples, even though had lower salinity levels compared to AI well, salinity levels also increased over time. No change in salinity was observed in the two SM separator samples. Nitrate and nitrite levels were below detection limits (50 mg/L for nitrate and 0.5 mg/L for nitrite) in all separator and tank samples examined. Elevated concentrations of total iron and ferrous iron were detected in all of the tank and separator samples. Elevated concentrations of sulfate were detected at every sampling event in the separator and tank samples from the SM well. However, sulfate was only detected sporadically in separator and tank samples from the AI well. Sulfide concentrations were undetectable (detection limit: 0.1 mg/L) in all samples. The slow flow of water from the separators and tanks combined with the introduction of outside air during the sampling process likely prevented long-term accumulation of sulfide in sampled locations.

Pyrosequencing and diversity estimates. A total of 50,209 high-quality sequences were obtained in this study, with an average of 2900 sequences per sample, and an average length of 203 bp (Table 3-2). Coverage estimates (Table 3-2) suggested that the majority of the bacterial community has been encountered. Multiple diversity

estimates were used to gauge diversity within and across sites. There was a general trend (with few exceptions) of a gradual decrease in Chao and ACE values were observed over time in AI and SM tanks, but not in separator samples (Table 3-2). Samples from the more saline, sulfate-depleted AI tanks and separators were in general less diverse than the less saline SM tanks and separators (ACE: tanks $p = 0.23$; separators $p = 0.09$). Finally, samples within the tanks were more diverse than the separator samples, except for SM separator samples, which were only collected twice (ACE: AI $p = 0.12$; SM $p = 0.22$).

Phylogenetic analysis. Microbial community structure and phylogenetic affiliation of major microbial lineages identified in all 17 samples in the four sampled locations are described below.

AI Tank. Within the AI tank, Proteobacterial-affiliated sequences represented the majority of the sequences in all six samples from the AI tank (ranging between 76 to 94% of sequences in examined samples), followed by members of the phylum Firmicutes (ranging between 6 to 23% of all AI samples examined) (Table 3-1a). The absolute majority (95-99% in examined samples) of the Proteobacteria sequences belonged to the Gammaproteobacteria order Alteromonadales and the Epsilonproteobacteria order Campylobacterales. The absolute majority of Firmicutes sequences (99%) belonged to the orders Bacillales, Clostridiales, and Halanaerobiales. Although these five order-level lineages represented the majority of the community, the proportion of each of these lineages could increase or decrease significantly within a sampling time frame (Figures 3-1, 3-2, and 3-3).

Alteromonadales sequences were nearly all (99% of the Alteromonadales fraction in various datasets) affiliated with the genus *Marinobacter*. The closest cultured relative of these sequences, with 96-98% sequence similarity, was *Marinobacter hydrocarbonoclasticus*, a known halotolerant, aerobic aliphatic hydrocarbon degrader, and iron oxidizer that was first isolated from petroleum contaminated seawater (35).

Sequences affiliated with the Epsilonproteobacteria order Campylobacterales, were most closely related (97-98%) to members of the genus *Arcobacter*, with the closest cultured relative (98% sequence similarity) being *A. mytili*, a mesophilic, halotolerant, sulfide-oxidizer (17). *Arcobacter* sequences have often been detected in other petroleum studies, and are capable of microaerophilic chemolithoautotrophic growth by utilizing sulfide as an electron donor (39).

As mentioned above, the majority of the Firmicutes sequences identified in AI tank belonged to the orders Bacillales, Clostridiales, and Halanaerobiales. The majority of sequences affiliated with the order Bacillales belonged to the family Bacillaceae, with members of the families Alicyclobacillaceae, Paenibacillaceae, Planococcaceae, and Staphylococcaceae, representing a small fraction of the Bacillales community (Figure 3-3). Bacillaceae sequences mostly belonged to the genera *Bacillus* and *Geobacillus*. The sequences classified as members of the genus *Bacillus* were closely related (97% sequence similarity) to *B. amyloliquefaciens*, a halotolerant fermenter (107). The sequences classified as members of the genus *Geobacillus* were closely related (97-98% sequence similarity) to *G. thermoleovorans*, a thermophilic, hydrocarbon-degrading,

facultative anaerobe, frequently encountered in various petroleum-impacted environments, hydrothermal vents, and hot springs (23, 58).

All the Clostridiales affiliated sequences belonged to the family Clostridiaceae, and most of the sequences were affiliated with the genus *Caminicella.* The closest cultured relative (96-98% sequence similarity) to the *Caminicella*-related sequences was *C. sporogenes*, which is a halophilic, thermophilic, anaerobic fermenter, and is capable of reducing elemental sulfur, L-cystine, and thiosulfate to H_2S (1).

The Halanaerobiales related sequences belonged to families Halanaerobiaceae and Halobacteroidaceae (Figure 3-3a). All of the Halanaerobiaceae sequences were classified as members the genus *Halanaerobium*. Most of the sequences were closely related (97-98%) to *H. congolense*, a moderate halophile and strict anaerobe that produces sulfide as a result of the oxidation of sugars and proteinaceous substrates in ecosystems that contain sulfur or thiosulfate. *H. congolense* was isolated from an African oil field and has been detected in 16S rRNA gene surveys from natural gas pipeline, oil pipeline biofilm, Brazilian oil reservoir, and a water-oil tank battery (68, 80, 89, 105, 121). The sequences classified as members of the family Halobacteroidaceae were all classified as the genus *Orenia*, and were closely related (94-98% sequence similarity) to clones (GenBank accession no. DQ647102) from the high-temperature Troll oil field in the Norwegian sector of the North Sea (19). *O. marismortui* (previously *Sporohalobacter marismortui*) was the closest cultured relative (93-94% sequence similarity), and is a halophilic, strict anaerobic fermenter, and a moderate thermophile isolated from the Dead Sea (72, 78).

AI Separator. Richness estimates conducted at a putative species level ($OTU_{0.03}$ sequence divergence cutoff, Table 3-2), indicated that AI tank and AI separator have similar levels of diversity at the species-level. However, detailed phylogenetic analysis revealed that at higher taxonomic levels (phyla, classes, and orders), AI separators were less diverse than the AI tank samples (Table 3-1a and b). Sequences affiliated with the phylum Firmicutes accounted for almost all (94-99%) of the sequences in every AI separator sample that was collected (Table 3-1b). Various proportions of Clostridiales (1.4% to 82%) and Halanaerobiales (ranging between 16% and 98%) dominated the sequences. Most of the Clostridiales sequences were members of the genus *Caminicella*, while the Halanaerobiales sequences were most closely related to *H. congolense*.

SM Tank. The SM tank water samples had similar diversity indices to the AI Tank samples, overall (Table 3-2). Furthermore, the SM tank samples exhibited similarities in phylogenetic composition with the AI tank samples (Table 3-1c). The SM tank sample, like the AI tank sample, contained a core community that consisted primarily of sequences that were affiliated with the Gammaproteobacteria, Epsilonproteobacteria, and Firmicutes. The populations of Gammaproteobacteria in the SM tank samples were similar to those observed in the AI tank and consisted primarily of sequences, which were affiliated with the order Alteromonadales and closely related (96-97% sequence similarity) to *Marinobacter hydrocarbonoclasticus*. In addition, a fraction (0-31.9% of sequences in various time points) of the Gammaproteobacteria sequences in the SM tank samples were affiliated with the orders Oceanospirillales and Pseudomonadales, which were not detected in the AI tank samples (Figures 3-2a and 3-

2c). These Oceanospirillales sequences belonged to the genus *Chromohalobacter* and were most closely related (99-100% sequence similarity) to *C. salexigens* (a strict aerobic, mesophilic, halophilic microorganism) (4). The Pseudomonadales sequences were 97-100% similar to *P. stutzeri*, a thiosulfate-oxidizing, polycyclic aromatic hydrocarbon-degrader, that has previously been detected in petroleum reservoirs in the U.S. and Canada (39, 52, 73, 76). The populations of Epsilonproteobacteria in the SM tank samples were very similar to those observed in the AI tank samples and consisted primarily of sequences that were affiliated with the order Campylobacterales and were closely related (96-97% sequence similarity) to *A. mytili*.

The populations of Firmicutes in the SM tank samples were also very similar to those observed in AI tank samples and consisted primarily of sequences that were affiliated with the orders Halanaerobiales, Bacillales, and Clostridiales. The majority of Firmicutes sequences in the SM tank were members of the order Halanaerobiales (Figure 3-3c). Halanaerobiales sequences in SM tanks were closely related (98-99% sequence similarity) to *H. congolense*. The Bacillales sequences were either related (94-96% sequence similarity) to *B. amyloliquefaciens*, or to *Geobacillus pallidus* (97-98% sequence similarity), a thermophile isolated from waste-water (87). The majority of the Clostridiales sequences were related (94-96% sequence similarity) to the genus *Caminicella*. The identification of sulfide-producing lineages within the Firmicutes population in SM tank is a reflection of the higher sulfate levels in SM tank compared to AI tank. In addition to Gram positive sulfate-reducing lineages, the SM tank samples also contained several groups of sequences that were affiliated with sulfide-producing bacteria from the class Deltaproteobacteria (0.5 - 7% range at various sampling times) and

phylum Synergistetes (0 - 3.4% range at various sampling times), and were detected in very low numbers (0 - 0.07% of the community) in the AI tank samples. The sequences that were affiliated with sulfide-producing Deltaproteobacteria were classified as members of the orders Desulfobacterales, Desulfovibrionales, and Desulfuromonadales, members of which have been frequently detected in multiple petroleum environments (19, 37, 39). The Desulfobacterales classified sequences were related (92-93% sequence similarity) to *Desulfobacter vibrioformis*. The Desulfovibrionales sequences were related (98% sequence similarity) to *Desulfovibrio capillatus,* which was isolated from a crude oil storage tank (65). The closest relatives (97-98% sequence similarity) to the Desulfuromonadales classified sequences were clones (GenBank accession no. FJ941600) from an oil-water separation tank (105). Sequences affiliated with the phylum Synergistetes represented 0-3% of the community in SM tanks. Synergistetes affiliated sequences were closely related (96-97% sequence similarity) to *Dethiosulfovibrio acidaminovorans*, which is a sulfur and thiosulfate-reducer (102).

SM Separator. The SM Separator datasets were the most diverse out of all the sampled locations (Table 3-2 and Table 3-1). The majority of the sequences belonged to the Gammaproteobacteria (orders Pseudomonadales and Alteromonadales), Betaproteobacteria (order Burkholderiales), and Firmicutes (orders Bacillales, Clostridiales, and Halanaerobiales). The Pseudomonadales sequences belonged to the genus Pseudomonas (97-98% sequence similarity), and the Alteromonadales sequences belonged to the genus *Marinobacter* (96-98% sequence similarity). The Bacillales sequences were classified as members of the genus *Bacillus*. Other Bacillales sequences

were classified as the genus *Planococcus*. The Clostridiales sequences were mostly related to the genus *Caminicella* (as described above). All of the Halanaerobiales sequences were 97-98% similar to *Halanaerobium congolense*. The Betaproteobacteria sequences were found to be in the order Burkholderiales, and the sequences were quite diverse at the genus level, but the majority of the sequences were related to the genera *Alcaligenes*, *Comamonas*, and *Ralstonia*.

Spatial and temporal comparisons of bacterial populations in tanks and separator samples. The microbial community structure was compared between various tanks and separator samples (from this study) on temporal and spatial scales. Also, tank and separator communities were compared to drilling mud and frac-water communities (previously characterized in both wells) to examine the role of drilling and hydraulic-fracturing processes in post-production microbial community establishment (99, 101). The results indicate that the production (tank or separator) communities from both wells bear little resemblance to the mud as well as the frac-water communities at the same well, with average shared numbers ranging between 0.06% and 3.15% (Table 3-3, Figure 3-4). This clearly indicates that the communities in tank and separator samples are completely distinct from communities introduced to the well during well drilling and hydraulic-fracturing processes. Spatial comparison of microbial communities between separator and tank samples of the same well at the same sampling time showed sequence similarities ranging between 25-68% (Table 3-4, Figure 3-4).

On the other hand, both spatial and temporal comparisons showed relatively higher levels of similarities between communities when compared to drilling mud or frac-

water communities. Spatial similarities appear to be lowest at the early sampling stages, which correspond to the early stages of production within sampled wells, and increase by time (Table 3-4, Figure 3-4). The highest level of similarities obtained were observed in temporal comparisons of microbial communities within the same locations, with percentage shared sequences values as high as 88, 91, and 94% in AI tanks SM tank, and AI separator samples, respectively (Table 3-5, Figure 3-4). Similarities between microbial communities on a temporal scale appear to be low at the early sampling stages, which corresponds to early stages of well production. Gradually, these percentages increase over time, suggesting the establishment of more stable microbial communities in those locations over time.

Table 3-1. Geochemical analysis of the tanks and separators for the two studied wells[a]

Sample[b]	Date Sampled[c]	SO_4^{2-}	Total Iron	Ferrous Iron	Salinity	TDS	Alkalinity	pH
AI Tank	19-Oct-2009	0	115	8	8%	10160	280	6.38
	13-Nov-2009	0	110	14	10%	12220	380	6.28
	18-Dec-2009	0	105	55	11%	10430	340	6.32
	19-Jan-2010	0	120	50	11%	8890	360	6.27
	10-Feb-2010	0	95	22	11%	ADL[d]	280	6.15
	12-Mar-2010	5	65	50	11%	ADL	200	6.44
AI Separator	19-Oct-2009	20	180	11	8%	7880	280	6.53
	13-Nov-2009	0	80	11	10%	11520	340	6.23
	10-Feb-2010	0	120	24	10%	ADL	400	6.13
	12-Mar-2010	3	100	70	11%	ADL	380	6.12
SM Tank	07-Oct-2009	52	25	3	6%	9310	360	6.56
	23-Nov-2009	10	35	12	6%	9690	300	6.34
	11-Dec-2009	14	65	16	8%	10260	340	6.38
	22-Feb-2010	38	15	12	8%	ADL	160	6.41
	31-Mar-2010	36	23	16	8%	ADL	180	6.38
SM Separator	11-Dec-2009	18	50	16	8%	10050	320	6.37
	22-Feb-2010	18	36	10	8%	ADL	200	6.23

[a] With the exception of salinity (values in % w/v) and pH, all values shown are in mg/l.
[b] Sample refers to the origin of each sample. Two wells were sampled (AI and SM) and two sites were sampled per well (storage tank, and gas-water separator).
[c] The dates correspond to the day-month-year when the sample was collected. Sampling points range between 2 (SM separator) to 6 (AI tank).
[d]ADL: Above detection limit

Table 3-2. Alpha diversity estimates for all samples studied.

Sample	Date Sampled	No. of Sequences	OTUs	Chao[a]	ACE[a]	Coverage
AI Tank	19-Oct-2009	1273	183	619	740	92%
	13-Nov-2009	3779	287	615	906	96%
	18-Dec-2009	887	126	243	439	92%
	19-Jan-2010	2082	227	473	734	94%
	10-Feb-2010	1730	197	411	699	94%
	12-Mar-2010	4496	252	417	581	98%
AI Separator	19-Oct-2009	1193	128	335	433	94%
	13-Nov-2009	2595	248	382	395	96%
	10-Feb-2010	5466	310	562	734	97%
	12-Mar-2010	7399	163	334	465	99%
SM Tank	07-Oct-2009	1154	270	580	1045	86%
	23-Nov-2009	1697	202	349	564	94%
	11-Dec-2009	3729	298	624	931	96%
	22-Feb-2010	2909	230	612	845	96%
	31-Mar-2010	4541	265	455	700	97%
SM Separator	11-Dec-2009	1369	295	1383	736	89%
	22-Feb-2010	3910	676	582	1737	91%

[a] See references 13 and 14

Table 3-3. Percentage shared sequences between drilling mud and frac-water communities and the bacterial communities in the studied samples[a].

Studied well	Sample location	Communities compared	
		Frac-water	Mud
AI	Tank	2.55 ± 0.015	0.07 ± 0.001
	Separator	3.15 ± 0.023	0.15 ± 0.002
SM	Tank	0.12 ± 0.003	0.1 ± 0.001
	Separator	0.25 ± 0.002	2.35 ± 0.002

[a] Numbers are the averages ± standard deviations of all shared sequences percentages for all sampling points at each location (6 sampling points for AI tank, 4 sampling points for AI separator, 5 sampling points for SM tank, and 2 sampling points for SM separator).

Table 3-4. Effect of sampling time on the bacterial communities in the samples studied.

Sampling location	Sampling times compared	Sørensen index[a]		% Shared sequences[b]
		Incidence-based	Abundance-based	
AI Tank	Oct-Nov	0.33	0.62	53%
	Nov-Dec	0.35	0.95	76%
	Dec-Jan	0.27	0.90	61%
	Jan-Feb	0.51	0.92	88%
	Feb-Mar	0.45	0.88	85%
SM Tank	Oct-Nov	0.06	0.38	13%
	Nov-Dec	0.48	0.94	91%
	Dec-Feb	0.34	0.85	74%
	Feb-Mar	0.42	0.91	74%
AI Separator	Oct-Nov	0.32	0.61	56%
	Nov-Feb	0.40	0.92	85%
	Feb-Mar	0.45	0.96	94%
SM Separator	Dec-Feb	0.29	0.63	52%

[a] numbers are pair wise Sørensen similarity index between the bacterial communities of the samples shown in the first column at the 2 sampling times shown in the second column. Incidence-based Sørensen index uses only the presence and absence data while the abundance-based index takes the abundance into account.
[b] numbers are the percentages of shared sequences between the bacterial communities of the samples shown in the first column at the 2 sampling times shown in the second column.

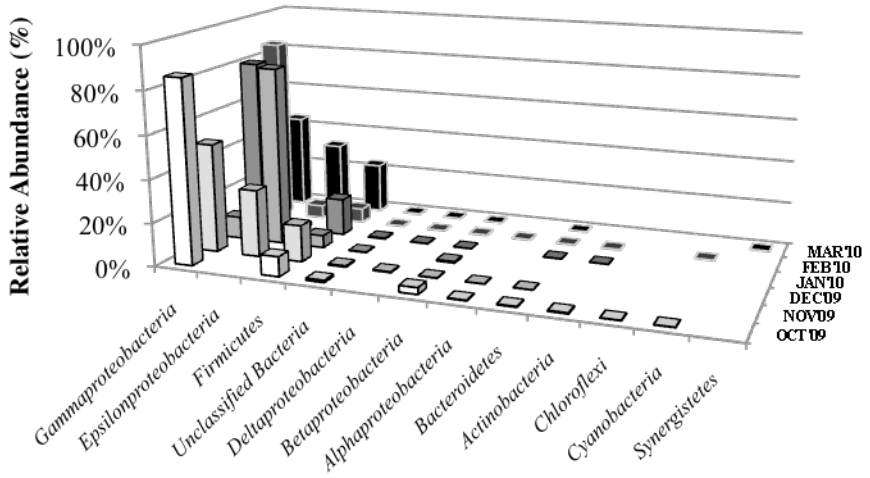
Table 3-5. Effect of sampling location on the bacterial communities in the samples studied. The tank and separator communities are compared for each of the wells studied.

Sample	Sampling time[a]	Sørensen index		% Shared Sequences
		Incidence-based	Abundance-based	
AI	Oct	0.12	0.13	34%
	Nov	0.15	0.25	28%
	Feb	0.10	0.12	59%
	Mar	0.27	0.35	68%
SM	Dec	0.07	0.52	25%
	Feb	0.22	0.33	49%

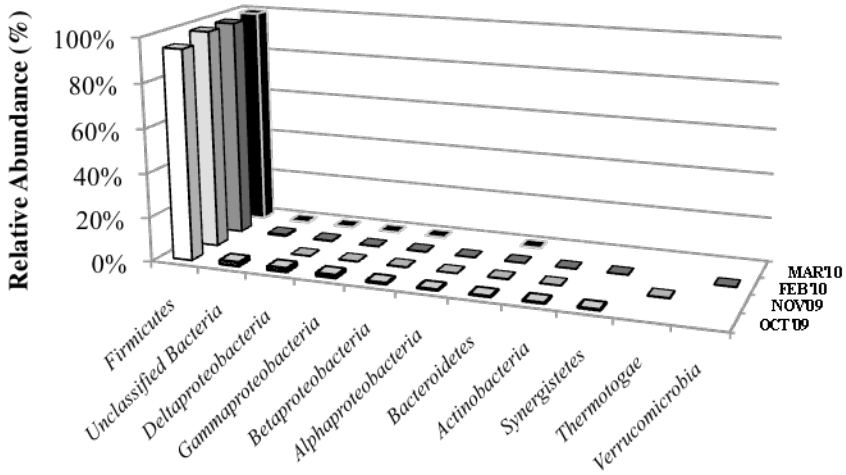
[a] Sampling time refers to the sampling month at which the tank and separator communities of each well are compared.

Figure 3-1. Phylum/Class-level classifications of the bacterial 16S rRNA gene pyrosequences obtained from (a) AI tank, (b) AI separator, (c) SM tank, and (d) SM separator samples. The time of sampling is shown on the Z-axis for each sample.

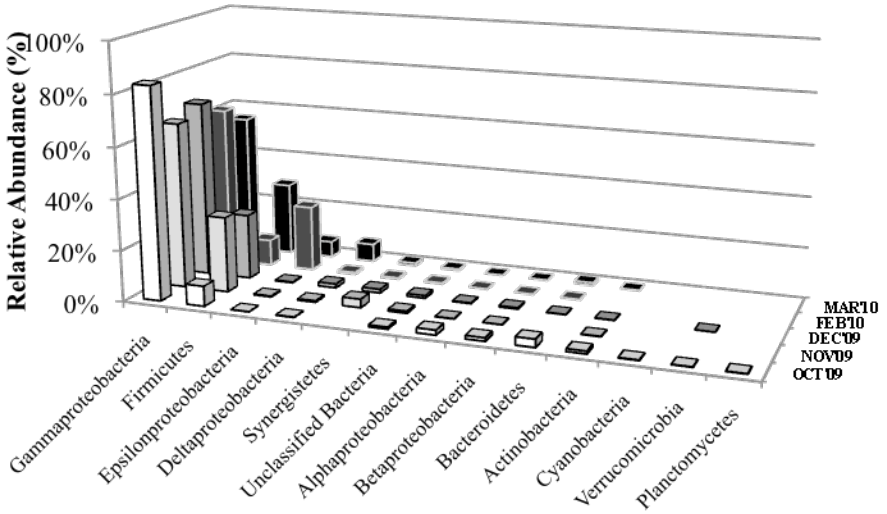
a.



b.



c.


Relative Abundance (%)
100%
80%
60%
40%
20%
0%
MAR'10
FEB'10
DEC'09
NOV'09
OCT'09
Gammaproteobacteria
Firmicutes
Epsilonproteobacteria
Deltaproteobacteria
Synergistetes
Unclassified Bacteria
Alphaproteobacteria
Betaproteobacteria
Bacteroidetes
Actinobacteria
Cyanobacteria
Verrucomicrobia
Planctomycetes

d.

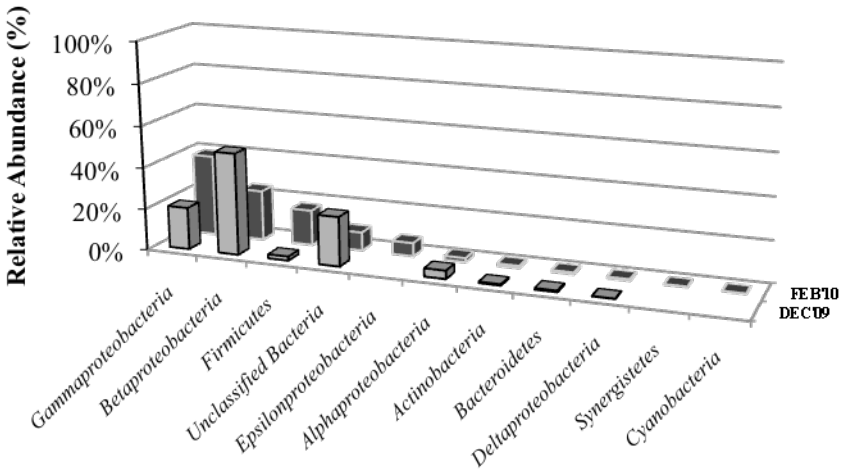
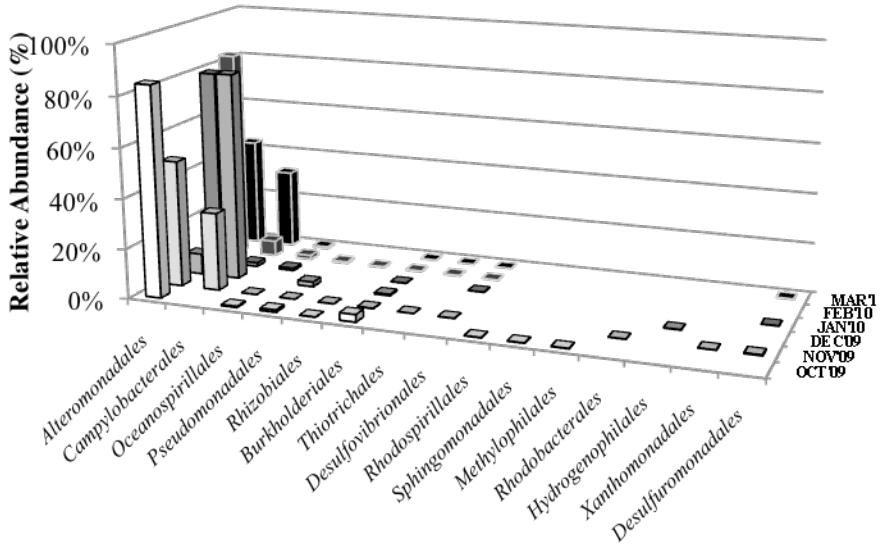
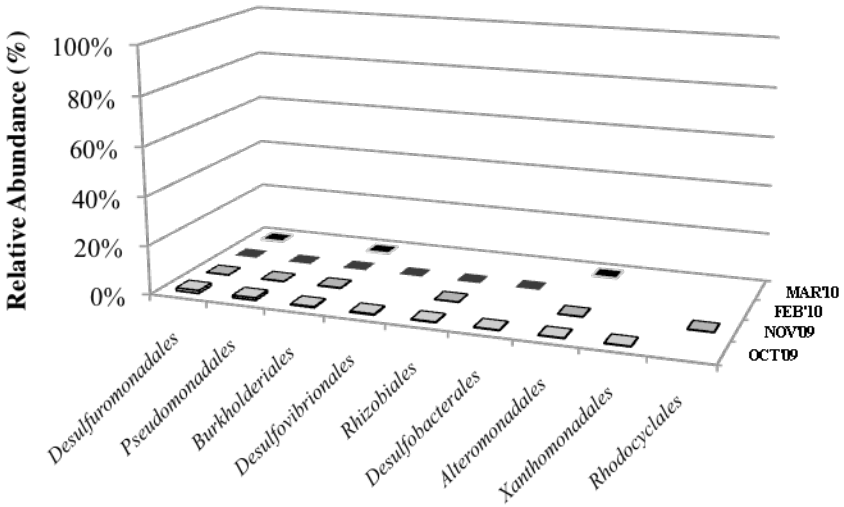

Relative Abundance (%)
100%
80%
60%
40%
20%
0%
FEB'10
DEC'09
Gammaproteobacteria
Betaproteobacteria
Firmicutes
Unclassified Bacteria
Epsilonproteobacteria
Alphaproteobacteria
Actinobacteria
Bacteroidetes
Deltaproteobacteria
Synergistetes
Cyanobacteria

Figure 3-2. Order-level classifications of the Proteobacteria sequences obtained from (a) AI tank, (b) AI separator, (c) SM tank, and (d) SM separator samples. The time of sampling is shown on the Z-axis for each sample.

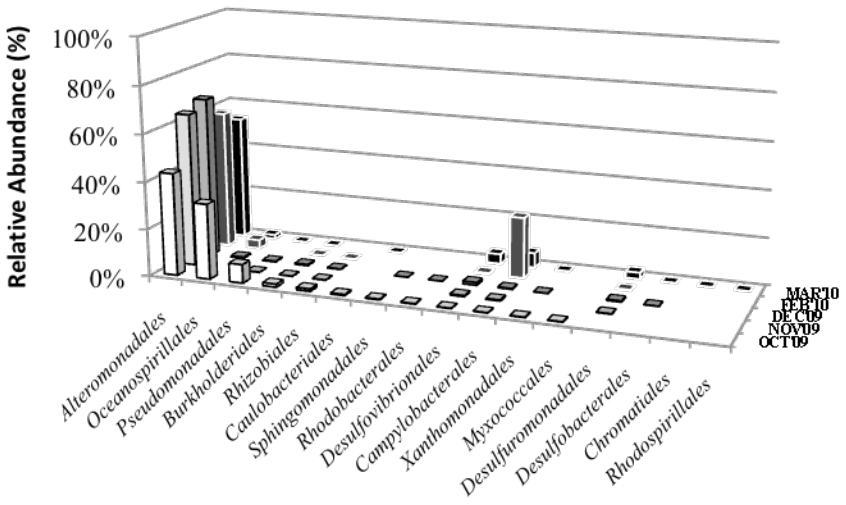
a.



b.



c.


Relative Abundance (%)
100%
80%
60%
40%
20%
0%
MAR'10
FEB'10
DEC'09
NOV'09
OCT'09
Alteromonadales
Oceanospirillales
Pseudomonadales
Burkholderiales
Rhizobiales
Caulobacteriales
Sphingomonadales
Rhodobacterales
Desulfovibrionales
Campylobacterales
Xanthomonadales
Myxococcales
Desulfuromonadales
Desulfobacterales
Chromatiales
Rhodospirillales

d.

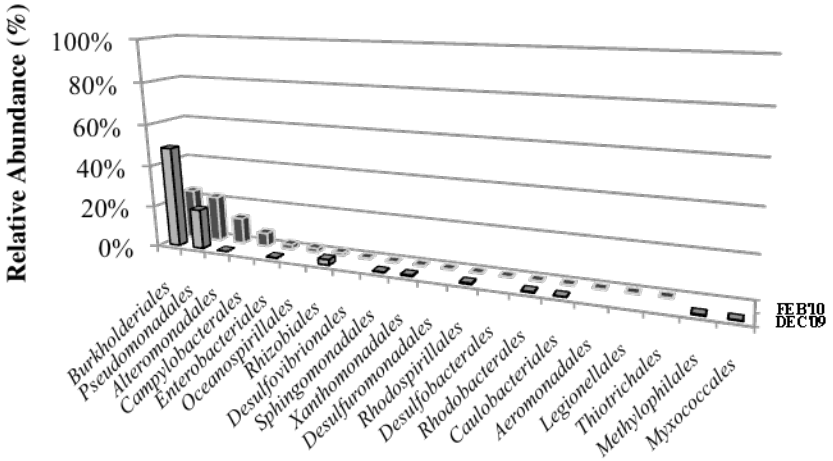
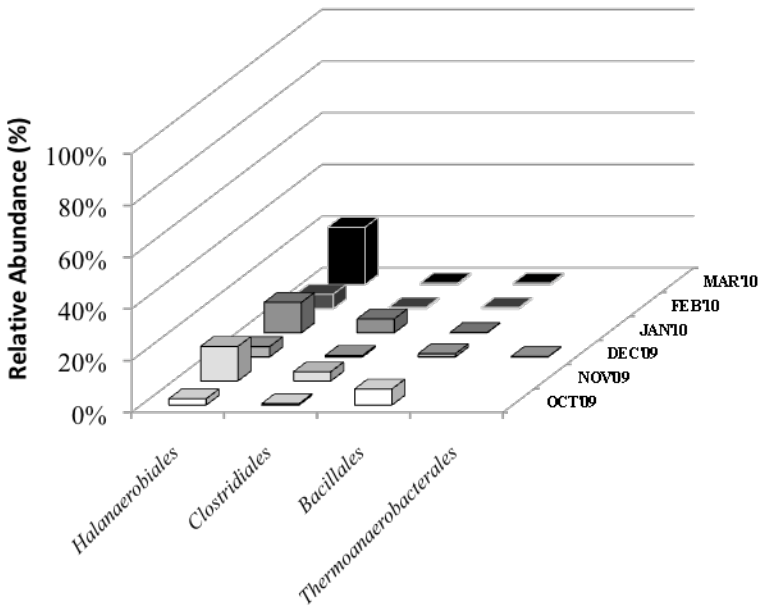
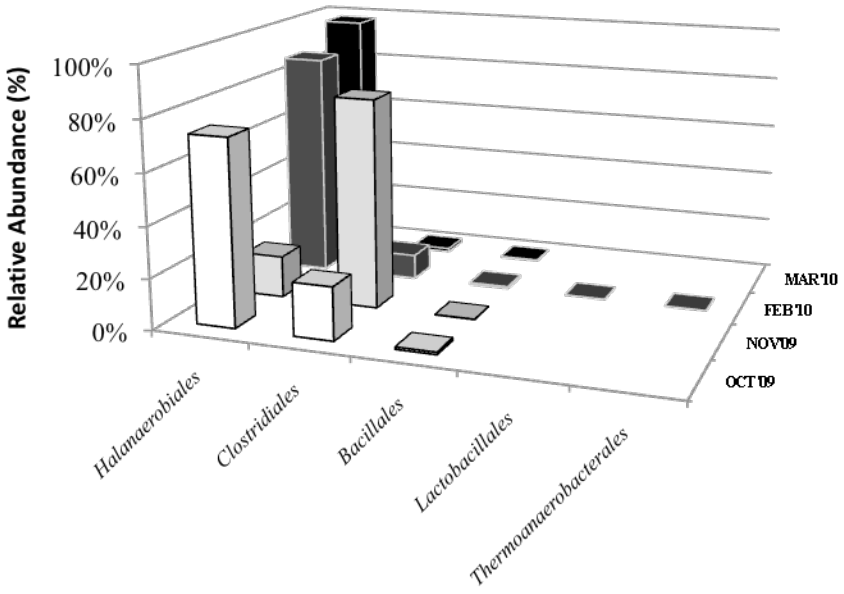

Relative Abundance (%)
100%
80%
60%
40%
20%
0%
FEB'10
DEC'09
Burkholderiales
Pseudomonadales
Alteromonadales
Campylobacterales
Enterobacteriales
Oceanospirillales
Rhizobiales
Desulfovibrionales
Sphingomonadales
Xanthomonadales
Desulfuromonadales
Rhodospirillales
Desulfobacterales
Rhodobacterales
Caulobacteriales
Aeromonadales
Legionellales
Thiotrichales
Methylophilales
Myxococcales

Figure 3-3. Order-level classifications of the Firmicutes sequences obtained from (a) AI tank, (b) AI separator, (c) SM tank, and (d) SM separator samples. The time of sampling is shown on the Z-axis for each sample.

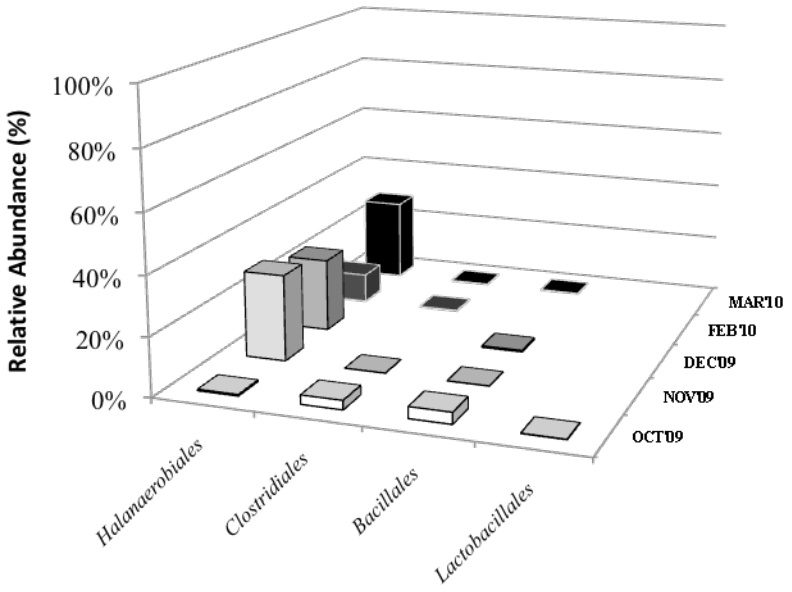
a.



b.



c.


Relative Abundance (%)
100%
80%
60%
40%
20%
0%
MAR'10
FEB'10
DEC'09
NOV'09
OCT'09
Halanaerobiales
Clostridiales
Bacillales
Lactobacillales

d.

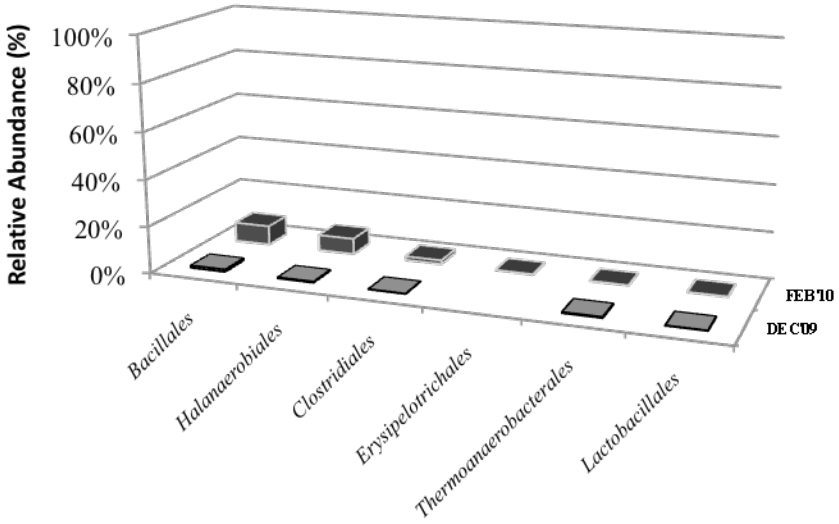
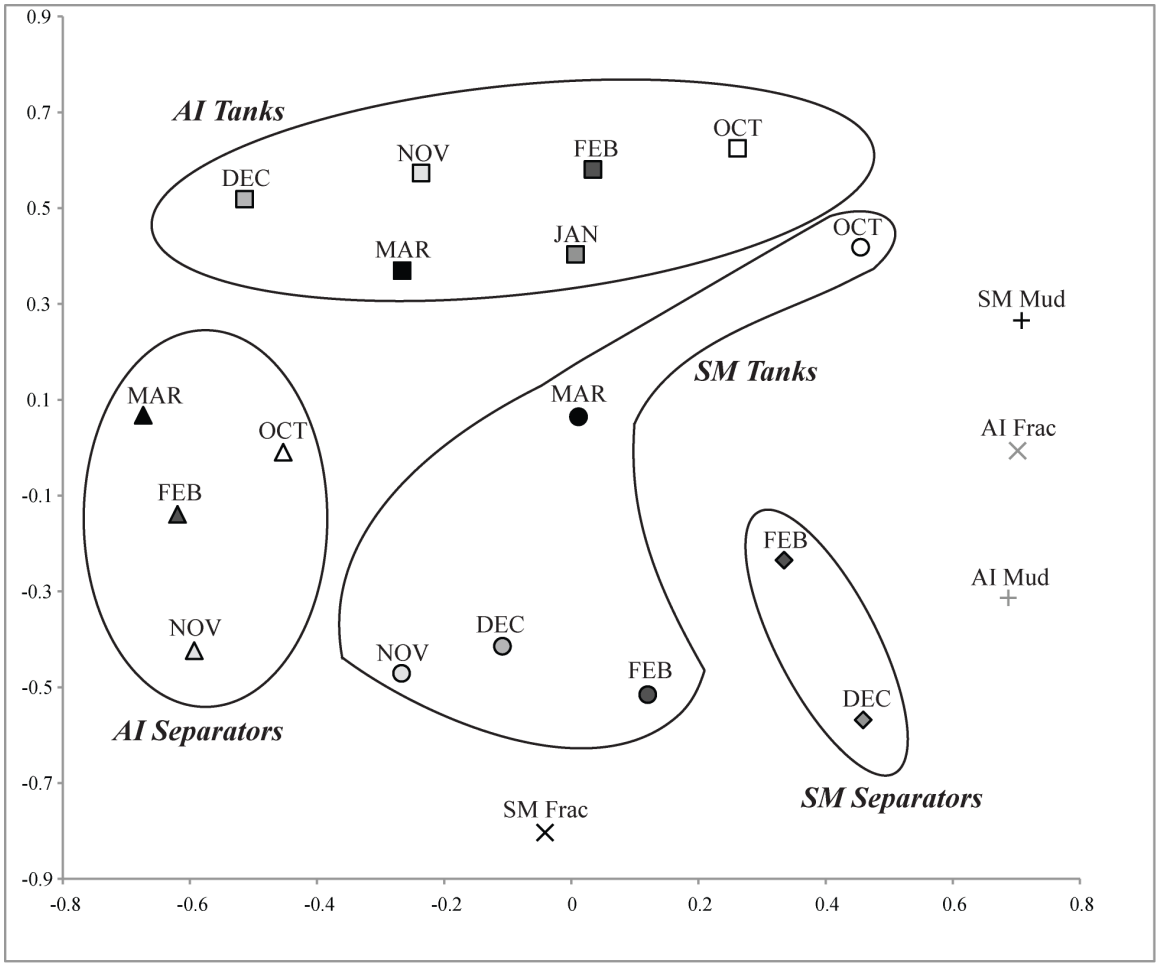

Relative Abundance (%)
100%
80%
60%
40%
20%
0%
FEB'10
DEC'09
Bacillales
Halanaerobiales
Clostridiales
Erysipelotrichales
Thermoanaerobacterales
Lactobacillales

Figure 3-4. Non-metric multidimensional scaling plot of all samples studied (SM tanks (closed squares), SM separators (closed stars), AI tanks (closed polygons), AI separators (closed circles) compared to mud and frac-water (open circles) samples. Pair wise abundance-based Sørensen indices were used to construct the plot.



Discussion

The predominance of specific lineages within the bacterial communities in AI and SM tanks and separators could best be understood by correlating the putative roles of these lineages to the observed geochemical conditions in such locations. The relatively high salinity within all samples (6-8% in SM and 8-11% in AI; Table 3-1) could explain the abundance of lineages that are exclusively halophilic (*e.g. Marinobacter, Halanaerobium, Chromohalobacter, and Caminicella*) in the bacterial communities from 14 out of 17 samples that were analyzed. Further, sequences affiliated with lineages that contain multiple halophilic species (but are not exclusive to halophilic microorganisms, *e.g.* order Clostridiales) were also abundant members of the community in AI separator datasets as well.

Obligate halophilic and halotolerant members of the community were observed in lineages that contain a variety of metabolic and physiological capabilities ranging from chemolithoautotrophs (*Arcobacter* sp.), aerobic heterotrophs (*Pseudomonas* sp.), anaerobic fermenters (*Halanaerobium* sp.), to anaerobic respiratory sulfate-reducers (*Desulfovibrio* sp.), which is consistent with previous studies of the microbial communities in petroleum production facilities (39, 105, 106). Lineages capable of growth at a wide range of redox potentials have also been observed in this study, including aerobes, facultative anaerobes, microaerophiles, and strict anaerobes (*e.g.* *Desulfovibrio* sp. and *Halanaerobium* sp.). Identification of such a wide, and seemingly contradictory, array of oxygen preference profiles in members of the community, often within a single sampling event, emphasizes the dynamic nature of sampled habitats. The flow of water from the well into the separators and then to the tanks is not constant, and

the water often rests in the production equipment between the periods of flow. This may create temporary anaerobic conditions in these areas, especially in the separators. Also, since the water flow is not constant, the water at the bottom of storage tanks should be more anaerobic due to the lack of circulation, and more aerobic at the top. Indeed, the coexistence of aerobic, microaerophilic, and anaerobic bacteria seems to be a hallmark of petroleum production facilities studies (*e.g.* oil-water separation tank samples taken from the Berkel oil field) (105).

Detailed phylogenetic analysis identified multiple members of the microbial communities within all samples that bear close resemblance to those observed in prior investigation of various fossil fluid production formations and facilities. Examples include sequences identified that are affiliated with genera *Pseudomonas*, *Arcobacter*, *Marinobacter*, *Geobacillus*, *Caminicella*, and *Halanaerobium*. These genera have previously been isolated or detected via 16S rRNA gene sequencing surveys from a high temperature oil-field of the San Joaquin Basin, oil formations (Troll, Dan, and Halfdan) of the North Sea, Pelican Lake oil field in Canada, a natural gas pipeline, and oil-water separator tanks in the Netherlands (19, 37, 39). The general geochemical and environmental reasons rationalizing the selection of lineages with such metabolic capabilities and physiological characteristics in this study, as well as in prior fossil formation diversity studies is fairly well understood *e.g.* elevated salinity, availability of SO_4^{2-}, sulfide, and elemental sulfur as described above. However the reasons for the observed repeated selection of a relatively limited number of genera and species in petroleum formations and facilities, from the larger pool of hydrocarbon degraders, halophiles and sulfur-metabolizers are not yet fully understood.

The wide range of sulfur metabolizing lineages within this dataset are of importance to the oil industry due to deleterious effects of sulfide (18, 118). In general, sulfidogenic lineages could be divided into two main groups: obligate respiratory sulfate-, sulfur-, and thiosulfate-reducers that utilize these compounds as a terminal electron acceptor during anaerobic respiration, and microorganisms that are capable of sulfide production from sulfur and thiosulfate while growing fermentatively. Obligate sulfate, sulfur, and thiosulfate reducing bacteria (members of the orders Desulfovibrionales and Desulfuromonadales) were more prevalent (range 0.2-7%) in SM tank samples, where sulfate levels were higher (10-52 mg/L, Table 3-1), and were present in much lower numbers in AI (0-1.6% separator and tank) samples where sulfate is limited (Table 3-1). The SRB sequences do increase over time in the SM Tank samples (Table 3-1c), which implies an increase to the biocorrosive potential of the bacterial communities present in the production equipment. As well, sequences affiliated with the sulfur-reducing genus *Dethiosulfovibrio,* within the phylum Synergistetes was detected in all sampled locations (though not in all the sampling times, 8 out of the 17 samples), and had the highest percentage (3.4%) in SM tank in November (Table 3-1). Facultative sulfidogenic microorganisms *i.e.* those capable of facultative sulfur and thiosulfate reduction to sulfide, were also identified in this study, and were close relatives of the thiosulfate and elemental sulfur-reducing *H. congolense* (order Halanaerobiales), and of the sulfur and sulfate-reducing genus *Caminicella* (order Clostridiales) (80). Sequences identified as close relatives of *H. congolense* were identified in most samples, and represented the majority of the sequences in most of the AI Separator samples (Figure 3-3b). Sequences related to the genus *Caminicella* were detected in many (13 out of 17) of the samples (0-

80% range), but were highest in SM separator in November (Figure 3-3d). Further, thiosulfate is completely soluble, emits no odor, and therefore tends to be overlooked in the petroleum industry (55). The large proportion of sequences affiliated with thiosulfate-reducing bacterial lineages (AI Separator, Figure 3-3b) indicates the possibility of deleterious, biocorrosive activity. Therefore, our results suggest that a consortium of both sulfate- and sulfur-reducing bacteria, in addition to facultative sulfidogenic bacteria could contribute to the incidents of sulfidogenesis and corrosion that has been reported in separators and tanks from the Barnett Shale (29). Multiple sources of sulfate and sulfur could be present in such locations, including barite and sulfonates that were present in drilling fluids (99). A variety of sulfur containing compounds (*e.g.* pyrite - FeS_2) were detected that could be converted to sulfide by air exposure that could subsequently by oxidized by microaerophilic conditions by chemolithoautotrophic sulfide oxidizers *e.g. Arcobacter*. These compounds were identified at various sampling points and locations (AI and SM tanks, Figures 3-2a and 3-2c).

Finally, pair-wise beta diversity comparisons revealed that the bacterial communities identified within the tanks and separators (at all time points) bore little resemblance to the communities of the drilling mud and frac-water (used to drill the wells). This indicates that microbial communities in these locations are not simply a carryover of microorganisms from the drilling mud and frac-water, but are distinct communities that appear to develop *in situ* in response to the prevalent conditions at each sampling site. Several members of the observed communities in the tanks and separators could have been minor components of the microbial communities in the drilling mud and frac-water, and adapted to the conditions in the tank or separator. For example,

Halanaerobiales sequences represented a small fraction of the total sequences in the SM and AI mud, but were in the majority in the AI separators and AI/SM tanks (99). Alternatively, soil, surface, and airborne microorganisms at these locations could provide such an inoculum for the developing microbial community. Regardless of the origin of microbial communities developing at these locations, it appears that a core microbial community eventually developed in all sampled production equipment. Within specific locations, while the membership of the microbial communities remains similar, its community structure and relative proportions of various taxa appear to fluctuate overtime. For example, the microbial communities in AI tank always contained Gammaproteobacteria and Firmicutes sequences (Table 3-1). However, the proportion of the orders (Alteromonadales, Oceanospirillales, Halanaerobiales, and Clostridiales) increases and decreases over time (Figures 3-2 and 3-3). This shows the dynamic nature in response to the sporadic influx of produced water from the formation.

In conclusion, the bacterial communities in the above ground production facilities (gas water separators and tanks) in two thermogenic, newly drilled, natural gas wells in the Barnett Shale in north central Texas, were surveyed over a period of 6 months. Analysis revealed that the bacterial communities from these locations: (1) Reflect the geochemical properties of waters in the production facilities (salinity, availability of multiple electron acceptors, and availability of various sulfur-containing chemical species), (2) bear clear resemblance to communities identified in prior studies of bacterial communities in fossil fluid production formations and above ground facilities in such locations, (3) harbor multiple obligate and facultative sulfidogenic lineages, and (4) bears little resemblance to communities previously identified in fluids utilized in well drilling

and hydraulic-fracturing, suggesting that such communities have developed *in situ* post production.

Acknowledgments

I would like to thank Reema B. Davis and Dr. Noha H. Youssef for assistance critiquing the manuscript and technical assistance.

References

1. **Alain, K., P. Pignet, M. Zbinden, M. Quillevere, F. Duchiron, J.-P. Donval, F. O. Lesongeur, G. R. Raguenes, P. Crassous, J. l. Querellou, and M.-A. Cambon-Bonavita.** 2002. Caminicella sporogenes gen. nov., sp. nov., a novel thermophilic spore-forming bacterium isolated from an East-Pacific Rise hydrothermal vent. Int J Syst Evol Microbiol **52:**1621-8.
2. **Amann, R. I., W. Ludwig, and K.-H. Schleifer.** 1995. Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143-169.
3. **Ansorge, W.** 2009. Next-generation DNA Sequencing Techniques. New Biotech. **25:**195-203.
4. **Arahal, D. R., M. T. Garcia, C. Vargas, D. Canovas, J. J. Nieto, and A. Ventosa.** 2001. Chromohalobacter salexigens sp. nov., a moderately halophilic species that includes Halomonas elongata DSM 3043 and ATCC 33174. International Journal of Systematic and Evolutionary Microbiology **51:**1457-62.
5. **Arthur, J., B. Bohm, B. Coughlin, and M. Layne.** 2008. Hydraulic fracturing considerations for natural gas wells of the Marcellus Shale. The Groundwater Protection Council Annual Forum, Cincinnatti, OH.
6. **Bowers, R. M., S. McLetchie, R. Knight, and N. Fierer.** 2011. Spatial variability in airborne bacterial communities across land-use types and their relationship to potential sources. ISME J **5:**601-612.

7. **Bowers, R. M., A. P. Sullivan, E. K. Costello, J. L. Collett, R. Knight, and N. Fierer.** 2011. Sources of bacteria in outdoor air across cities in the midwestern United States. Appl. Environ. Microbiol. **77:**6350-6356.

8. **Bowker, K. A.** 2007. Barnett Shale gas production, Fort Worth Basin: Issues and discussion. AAPG Bulletin **91:**523-533.

9. **Bowker, K. A.** 2003. Recent development of the Barnett Shale play. Fort Worth Basin. West Texas Geological Society Bulletin **42:**4-11.

10. **Brauman, A., M. D. Kane, M. Labat, and J. A. Breznak.** 1992. Genesis of Acetate and Methane by Gut Bacteria of Nutritionally Diverse Termites. Science **257:**1384-1387.

11. **Bybee, S. M., H. Bracken-Grissom, B. D. Haynes, R. A. Hermansen, R. L. Byers, M. J. Clement, J. A. Udall, E. R. Wilcox, and K. A. Crandall.** 2011. Targeted Amplicon Sequencing (TAS): A Scalable Next-Gen Approach to Multilocus, Multitaxa Phylogenetics. Genome Biol. Evol. **3:**1312-1323.

12. **Callaway, T. R., S. E. Dowd, T. S. Edrington, R. C. Anderson, N. Krueger, N. Bauer, P. J. Kononoff, and D. J. Nisbet.** 2010. Evaluation of bacterial diversity in the rumen and feces of cattle fed different levels of dried distillers grains plus solubles using bacterial tag-encoded FLX amplicon pyrosequencing. Journal of Animal Science **88:**3977-3983.

13. **Chao, A., and S.-M. Lee.** 1992. Estimating the Number of Classes via Sample Coverage. J Am Stat Assoc **87:**210-217.

14. **Chao, A., M. C. Ma, and M. C. K. Yang.** 1993. Stopping rules and estimation for recapture debugging with unequal failure rates. Biometrika **80:**193-201.

15. **Chung, K. T., and D. H. Ferris.** 1996. Martinus Willem Beijerinck (1851-1931): Pioneer of General Microbiology. ASM News **62:**539-543.

16. **Cline, J. D.** 1969. Spectrophotometric determination of hydrogen sulfide in natural waters. Limnol. Oceanogr. **14:**454-458.

17. **Collado, L., I. Cleenwerck, S. Van Trappen, P. De Vos, and M. J. Figueras.** 2009. Arcobacter mytili sp. nov., an indoxyl acetate-hydrolysis-negative bacterium isolated from mussels. Int J Syst Evol Microbiol **59:**1391-1396.

18. **Cord-Ruwisch, R., W. Kleinitz, and F. Widdel.** 1987. Sulfate-reducing bacteria and their activities in oil production. J. Petrol. Technol. **1:**97 – 106.

19. **Dahle, H., F. Garshol, M. Madsen, and N.-K. Birkeland.** 2008. Microbial community structure analysis of produced water from a high-temperature North Sea oil-field. Antonie van Leeuwenhoek **93:**37-49.

20. **DeSantis, T. Z., P. Hugenholtz, K. Keller, E. L. Brodie, N. Larsen, Y. M. Piceno, R. Phan, and G. L. Andersen.** 2006. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. Nucleic Acids Res **34:**W394-399

21. **DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen.** 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl. Environ. Microbiol. **72:**5069-5072.

22. **DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen.** 2006. Greengenes, a Chimera-

Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl. Environ. Microbiol. **72:**5069-5072.

23. **Dinsdale, A. E., G. Halket, A. Coorevits, A. Van Landschoot, H.-J. Busse, P. De Vos, and N. A. Logan.** 2011. Emended descriptions of Geobacillus thermoleovorans and Geobacillus thermocatenulatus. Int J Syst Evol Microbiol **61:**1802-1810.

24. **Dworkin, M.** 2012. Sergei Winogradsky: A founder of modern microbiology and the first microbial ecologist. FEMS Microbiol. Rev. **36:**364-379.

25. **Edwards, R. A., B. Rodriguez-Brito, L. Wegley, M. Haynes, M. Breitbart, D. M. Peterson, M. O. Saar, S. Alexander, E. A. Alexander Jr, and F. Rohwer.** 2006. Using pyrosequencing to shed light on deep mine microbial ecology. BMC Genomics. **7:**57.

26. **Elshahed, M. S., J. M. Senko, F. Z. Najar, S. M. Kenton, B. A. Roe, T. A. Dewers, J. R. Spear, and L. R. Krumholz.** 2003. Bacterial Diversity and Sulfur Cycling in a Mesophilic Sulfide-Rich Spring. Appl. Environ. Microbiol. **69:**5609-5621.

27. **Entrekin, S., M. Evans-White, B. Johnson, and E. Hagenbuch.** 2011. Rapid expansion of natural gas development poses a threat to surface waters. Front. Ecol. Environ. **9:**503-511.

28. **Fichter, J., K. Johnson, K. French, and R. Oden.** 2009. Biocides control Barnett Shale fracturing fluid contamination. Oil Gas J. **107:**38-44.

29. **Fichter, J. K., K. Johnson, K. French, and R. Oden.** 2009. Biocides control Barnett shale fracturing fluid contamination. Oil & Gas J. **107:**38-44.

30. **Flores, G. E., J. H. Campbell, J. D. Kirshtein, J. Meneghin, M. Podar, J. I. Steinberg, J. S. Seewald, M. K. Tivey, M. A. Voytek, Z. K. Yang, and A.-L. Reysenbach.** 2011. Microbial community structure of hydrothermal deposits from geochemically different vent fields along the Mid-Atlantic Ridge. Environmental Microbiology **13:**2158-2171.

31. **Flores, G. E., M. Shakya, J. Meneghin, Z. K. Yang, J. S. Seewald, C. Geoff Wheat, M. Podar, and A. L. Reysenbach.** 2012. Inter-field variability in the microbial communities of hydrothermal vent deposits from a back-arc basin. Geobiology **10:**333-346.

32. **Franca, L. T. C., E. Carrilho, and T. B. L. Kist.** 2002. A review of DNA sequencing techniques. Quart. Rev. Biophys. **35:**169-200.

33. **Fredericks, D. N., and D. A. Relman.** 1996. Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. Clinical Microbiology Reviews **9:**18-33.

34. **Fuhrman, J. A., K. McCallum, and A. A. Davis.** 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. Applied and Environmental Microbiology **59:**1294-1302.

35. **Gauthier, M. J., B. Lafay, R. Christen, L. Fernandez, M. Acquaviva, P. Bonin, and J.-C. Bertrand.** 1992. Marinobacter hydrocarbonoclasticus gen. nov., sp. nov., a New, Extremely Halotolerant, Hydrocarbon-Degrading Marine Bacterium. Int J Syst Bacteriol **42:**568-576.

36. **Giovannoni, S. J., E. F. DeLong, G. J. Olsen, and N. R. Pace.** 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. J. Bacteriol. **170:**720-726.

37. **Gittel, A., K. B. Sorensen, T. L. Skovhus, K. Ingvorsen, and A. Schramm.** 2009. Prokaryotic Community Structure and Sulfate Reducer Activity in Water from High-Temperature Oil Reservoirs with and without Nitrate Treatment. Appl. Environ. Microbiol. **75:**7086-7096.

38. **Glenn, T. C.** 2011. Field guide to next-generation DNA sequencers. Molec. Ecol. Resourc. **11:**759-769.

39. **Grabowski, A., O. Nercessian, F. Fayolle, D. Blanchet, and C. Jeanthon.** 2005. Microbial diversity in production waters of a low-temperature biodegraded oil reservoir. FEMS Microbiol Ecol **54:**427-443.

40. **Hamady, M., J. J. Walker, J. K. Harris, N. J. Gold, and R. Knight.** 2008. Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods **5:**235-237.

41. **Head, I. M., J. R. Saunders, and R. W. Pickup.** 1998. Microbial Evolution, Diversity, and Ecology: A Decade of Ribosomal RNA Analysis of Uncultivated Microorganisms. Micro. Ecol. **35:**1-21.

42. **Hicks, R. E., R. I. Amann, and D. A. Stahl.** 1992. Dual staining of natural bacterioplankton with 4',6-diamidino-2-phenylindole and fluorescent oligonucleotide probes targeting kingdom-level 16S rRNA sequences. Applied and Environmental Microbiology **58:**2158-2163.

43. **Hill, R. J., D. M. Jarvie, J. Zumberge, M. Henry, and R. M. Pollastro.** 2007. Oil and gas geochemistry and petroleum systems of the Fort Worth Basin. AAPG Bulletin **91:**445-473.

44. **Hollister, E. B., A. S. Engledow, A. J. M. Hammett, T. L. Provin, H. H. Wilkinson, and T. J. Gentry.** 2010. Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. ISME J **4:**829-838.

45. **Hu, E., Z. Huang, B. Liu, J. Zheng, X. Gu, and B. Huang.** 2009. Experimental investigation on performance and emisssions of a spark-ignition engine fuelled with natural gas-hydrogen blends combined with EGR. Int J Hydrogen Energy **34:**528-539.

46. **Hugenholtz, P.** 2002. Exploring prokaryotic diversity in the genomic era. Genome Biol. **3:**REVIEWS0003. Epub 2002 Jan 29.

47. **Hugenholtz, P., B. M. Goebel, and N. R. Pace.** 1998. Impact of Culture-Independent Studies on the Emerging Phylogenetic View of Bacterial Diversity. J. Bacteriol. **180:**4765-4774.

48. **Hungate, R. E.** 1969. A Roll Tube Method for Cultivation of Strict Anaerobes. *In* J. R. Norris and D. W. Ribbons (ed.), Methods in Microbiology, vol. 3B. Academic Press Inc., New York.

49. **Jan-Roblero, J., A. Posadas, J. Zavala Díaz de la Serna, R. García, and C. Hernández-Rodríguez.** 2008. Phylogenetic characterization of bacterial consortia obtained of corroding gas pipelines in Mexico. World J Microbiol Biotechnol **24:**1775-1784.

50. **Jannasch, H. W., and G. E. Jones.** 1959. Bacterial populations in sea water as determined by different methods of enumeration. Limnol. Oceanogr. **4:**128-129.

51. **Kerr, R. A.** 2010. Natural Gas From Shale Bursts Onto the Scene. Science **328:**1624-1626.

52. **Lalucat, J., A. Bennasar, R. Bosch, E. García-Valdés, and N. J. Palleroni.** 2006. Biology of Pseudomonas stutzeri. Microbiol Mol Biol Rev **70:**510-547.

53. **Lander, E. S.** 2011. Initial impact of the sequencing of the human genome. Nature **470:**187-197.

54. **Liesack, W., and E. Stackebrandt.** 1992. Occurrence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. Journal of Bacteriology **174:**5072-5078.

55. **Magot, M.** 2005. Indigenous microbial communities in oil fields, p. 21-33. *In* B. Ollivier and M. Magot (ed.), Petroleum microbiology. ASM Press, Washington D.C.

56. **Magot, M., B. Ollivier, and B. K. Patel.** 2000. Microbiology of petroleum reservoirs. Antonie Van Leeuwenhoek **77:**103-116.

57. **Manz, W., R. Amann, W. Ludwig, M. Wagner, and K.-H. Schleifer.** 1992. Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. Syst. Appl. Microbiol. **15:**593-600.

58. **Maugeri, T. L., C. Gugliandolo, D. Caccamo, and E. Stackebrandt.** 2002. Three novel halotolerant and thermophilic Geobacillus strains from shallow marine vents. Syst Appl Microbiol **25:**450-455.

59. **Metzker, M. L.** 2005. Emerging technologies in DNA sequencing. Genome Res. **15:**1767-1776.

60. **Metzker, M. L.** 2009. Sequencing technologies - the next generation. Nat Rev Genet **11:**31-46.

61. **Meyer-Reil, L.-A.** 1978. Autoradiography and Epifluorescence Microscopy Combined for the Determination of Number and Spectrum of Actively Metabolizing Bacteria in Natural Waters. Appl. Environ. Microbiol. **36:**506-512.

62. **Miller, T. L., and M. J. Wolin.** 1974. A Serum Bottle Modification of the Hungate Technique for Cultivating Obligate Anaerobes. Appl. Microbiol. **27:**985-987.

63. **Min Jou, W., G. Haegeman, M. Ysebaert, and W. Fiers.** 1973. Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. Nature **237:**82-88.

64. **Montgomery, S., D. Jarvie, K. Bowker, and R. Pollastro.** 2005. Mississippian Barnett Shale, Fort Worth basin, north-central Texas: gas-shale play with multi-trillion cubic foot potential. AAPG Bull. **89:**155–175.

65. **Mori, K., H. Tsurumaru, and S. Harayama.** 2010. Iron corrosion activity of anaerobic hydrogen-consuming microorganisms isolated from oil facilities. Journal of Bioscience and Bioengineering **110:**426-430.

66. **Muyzer, G., E. C. deWaal, and A. G. Uitterlinden.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. **59:**695-700.

67. **Nealson, K. H.** 1997. SEDIMENT BACTERIA: Who's There, What Are They Doing, and What's New? Annu. Rev. Earth Planet. Sci. **25:**403-434.

68. **Neria-González, I., E. T. Wang, F. Ramírez, J. M. Romero, and C. Hernández-Rodríguez.** 2006. Characterization of bacterial community associated to biofilms of corroded oil pipelines from the southeast of Mexico. Anaerobe **12:**122-133.

69. **Nyrén, P.** 2007. The History of Pyrosequencing. Methods Mol. Biol. **373:**1-14.

70. **Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace.** 1986. Microbial Ecology and Evolution: A Ribosomal RNA Approach. Ann. Rev. Microbiol. **40:**337-365.

71. **Olson, M. V.** 1993. The Human Genome Project. Proc. Natl. Acad. Sci. U.S.A. **90:**4338-4344.

72. **Oren, A., H. Pohla, and E. Stackebrandt.** 1987. Transfer of Clostridium lortetii to a new genus Sporohalobacter gen. nov. as Sporohalobacter lortetii comb. nov. and description of Sporohalobacter marismortui sp. nov. Syst Appl Microbiol **9:**239-246.

73. **Orphan, V. J., L. T. Taylor, D. Hafenbradl, and E. F. Delong.** 2000. Culture-Dependent and Culture-Independent Characterization of Microbial Assemblages Associated with High-Temperature Petroleum Reservoirs. Appl Environ Microbiol **66:**700-711.

74. **Pace, N. R., D. A. Stahl, D. J. Lane, and G. J. Olsen.** 1986. The analysis of natural microbial populations by ribosomal RNA sequences. Adv Microb Ecol **9:**1-55.

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

76. **Pham, V. D., L. L. Hnatow, S. Zhang, R. D. Fallon, S. C. Jackson, J.-F. Tomb, E. F. DeLong, and S. J. Keeler.** 2009. Characterizing microbial diversity in production water from an Alaskan mesothermic petroleum reservoir with two independent molecular methods. Environ Microbiol **11:**176-187.

77. **Pollastro, R. M.** 2007. Total petroleum system assessment of undiscovered resources in the giant Barnett Shale continuous (unconventional) gas accumulation, Fort Worth Basin, Texas. AAPG Bulletin **91:**551-578.

78. **Rainey, F. A., T. N. Zhilina, E. S. Boulygina, E. Stackebrandt, T. P. Tourova, and G. A. Zavarzin.** 1995. The Taxonomic Status of the Fermentative Halophilic Anaerobic Bacteria: Description of Haloanaerobiales ord. nov., Halobacteroidaceae fam. nov., Orenia gen. nov. and further Taxonomic Rearrangements at the Genus and Species Level. Anaerobe **1:**185-199.

79. **Rappé, M. S., and S. J. Giovannoni.** 2003. The Uncultured Microbial Majority. Ann. Rev. Microbiol. **57:**369-394.

80. **Ravot, G., M. Magot, B. Ollivier, B. K. C. Patel, E. Ageron, P. A. D. Grimont, P. Thomas, and J. L. Garcia.** 1997. Haloanaerobium congolense sp. nov., an anaerobic, moderately halophilic, thiosulfate- and sulfur-reducing bacterium from an African oil field. FEMS Microbiol Lett **147:**81-88.

81. **Razumov, A. S.** 1932. The direct method of calculation of bacteria in water: comparision with the Koch method. Mikrobiol. **1:**131-146.

82. **Rossi, F., S. Torriani, and F. Dellaglio.** 1999. Genus- and Species-Specific PCR-Based Detection of Dairy Propionibacteria in Environmental Samples by Using Primers Targeted to the Genes Encoding 16S rRNA. Applied and Environmental Microbiology **65:**4241-4244.

83. **Rothberg, J., and J. Leamon.** 2008. The development and impact of 454 sequencing. Nat Biotech **26:**1117-1124.

84. **Sanger, F., and A. R. Coulson.** 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. **94:**441-448.

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

86. **Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber.** 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. **75:**7537-7541.

87. **Scholz, T., W. Demharter, R. Hensel, and O. Kandler.** 1987. Bacillus pallidus sp. nov., a new thermophilic species from sewage. Systematic and Applied Microbiology **9:**91-96.

88. **Senko, J. M., B. S. Campbell, J. R. Henriksen, M. S. Elshahed, T. A. Dewers, and L. R. Krumholz.** 2004. Barite deposition resulting from phototrophic sulfide-oxidizing bacterial activity. Geochimica et Cosmochimica Acta **68:**773-780.

89. **Sette, L., K. Simioni, S. Vasconcellos, L. Dussan, E. Neto, and V. Oliveira.** 2007. Analysis of the composition of bacterial communities in oil reservoirs from a southern offshore Brazilian basin. Antonie van Leeuwenhoek **91:**253-266.

90. **Shendure, J., and H. Ji.** 2008. Next-generation DNA sequencing Nat. Biotech. **26:**1135-1145.

91. **Smith, L., S. Fung, M. Hunkapiller, T. Hunkapiller, and L. Hood.** 1985. The synthesis of oligonucleotides containing an aliphatic amino group at the 5' terminus: synthesis of fluorescent DNA primers for use in DNA sequence analysis. Nucleic Acids Res. **13:**2399-2412.

92. **Smith, L., J. Sanders, R. Kaiser, P. Hughes, C. Dodd, C. Connell, C. Heiner, S. Kent, and L. Hood.** 1986. Fluorescence detection in automated DNA sequence analysis. Nature **321:**674-679.

93. **Sogin, M. L., H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, and G. J. Herndl.** 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere. Proc. Natl. Acad. Sci. U.S.A. **103:**12115-12120.

94. **Sørensen, T.** 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Biologiske Skrifter/Kongelige Danske Videns-kabernes Selskab **5:**1-34.

95. **Stackebrandt, E., W. Liesack, and B. M. Goebel.** 1993. Bacterial diversity in a soil sample from a subtropical Australian environment as determined by 16S rDNA analysis. The FASEB Journal **7:**232-6.

96. **Stahl, D. A., and R. I. Amann.** 1991. Development and application of nucleic acid probes in bacterial systematics, p. 205-248. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, Chichester, England.

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

98. **Stevenson, F. J., and M. A. Cole.** 1999. Cycles of Soil: Carbon, Nitrogen, Phosphorus, Sulfur, Micronutrients. John Wiley and Sons, Inc., New York.

99. **Struchtemeyer, C. G., J. P. Davis, and M. S. Elshahed.** 2011. Influence of the Drilling Mud Formulation Process on the Bacterial Communities in Thermogenic Natural Gas Wells from the Barnett Shale. Appl. Environ. Microbiol. **77:**4744-4753.

100. **Struchtemeyer, C. G., J. P. Davis, and M. S. Elshahed.** 2011. Influence of the Drilling Mud Formulation Process on the Bacterial Communities in Thermogenic Natural Gas Wells of the Barnett Shale. Applied and Environmental Microbiology **77:**4744-4753.

101. **Struchtemeyer, C. G., and M. S. Elshahed.** 2011. Bacterial communities associated with hydraulic fracturing fluids in thermogenic natural gas wells in North Central Texas, USA. FEMS Microbiol Ecol.

102. **Surkov, A. V., G. A. Dubinina, A. M. Lysenko, F. O. Gl√∂ckner, and J. Kuever.** 2001. Dethiosulfovibrio russensis sp. nov., Dethosulfovibrio marinus sp. nov. and Dethosulfovibrio acidaminovorans sp. nov., novel anaerobic, thiosulfate-

and sulfur-reducing bacteria isolated from 'Thiodendron' sulfur mats in different saline environments. Int J Syst Evol Microbiol **51:**327-37.

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

104. **Van de Peer, Y., S. Chapelle, and R. De Wachter.** 1996. A quantitative map of nucleotide substitution rates in bacterial rRNA. Nucleic Acids Res. **24:**3381-3391.

105. **van der Kraan, G. M., J. Bruining, B. P. Lomans, M. C. M. Van Loosdrecht, and G. Muyzer.** 2010. Microbial diversity of an oil–water processing site and its associated oil field: the possible role of microorganisms as information carriers from oil-associated environments. FEMS Microbiol Ecol **71:**428-443.

106. **Voordouw, G., S. M. Armstrong, M. F. Reimer, B. Fouts, A. J. Telang, Y. Shen, and D. Gevertz.** 1996. Characterization of 16S rRNA genes from oil field microbial communities indicates the presence of a variety of sulfate-reducing, fermentative, and sulfide-oxidizing bacteria. Appl Environ Microbiol **62:**1623-9.

107. **Wang, L.-T., F.-L. Lee, C.-J. Tai, and H.-P. Kuo.** 2008. Bacillus velezensis is a later heterotypic synonym of Bacillus amyloliquefaciens. Int J Syst Evol Microbiol **58:**671-675.

108. **Weaver, W., and C. E. Shannon.** 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL.

109. **Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane.** 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. **173:**697-703.

110. **Weller, R., M. M. Bateson, B. K. Heimbuch, E. D. Kopczynski, and D. M. Ward.** 1992. Uncultivated cyanobacteria, Chloroflexus-like inhabitants, and

spirochete-like inhabitants of a hot spring microbial mat. Applied and Environmental Microbiology **58:**3964-3969.

111. **Whitman, W. B., D. C. Coleman, and W. J. Wiebe.** 1998. Prokaryotes: The unseen majority. Proc. Natl. Acad. Sci. U.S.A. **95:**6578-6583.

112. **Witze, A.** 2007. Energy: That's oil, folks. Nature **445:**14-17.

113. **Woese, C., and G. Fox.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U.S.A. **74:**5088-5090.

114. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221-271.

115. **Woese, C. R.** 1994. There must be a prokaryote somewhere: microbiology's search for itself. Microbiological Reviews **58:**1-9.

116. **Youssef, N., M. S. Elshahed, and M. J. McInerney.** 2009. Microbial Processes in Oil Fields: Culprits, Problems, and Opportunities, p. 141-251. *In* S. S. Allen I. Laskin and M. G. Geoffrey (ed.), Advances in Applied Microbiology, vol. Volume 66. Academic Press.

117. **Youssef, N. H., M. B. Couger, and M. S. Elshahed.** 2010. Fine-Scale Bacterial Beta Diversity within a Complex Ecosystem (Zodletone Spring, OK, USA): The Role of the Rare Biosphere. PLoS ONE **5:**e12414.

118. **Youssef, N. H., M. S. Elshahed, and M. J. McInerney.** 2009. Microbial processes in oil fields: culprits, problems, and opportunities. Adv. Appl. Microbiol. **66:**141-251.

119. **Zehr, J. P., and B. B. Ward.** 2002. Nitrogen Cycling in the Ocean: New Perspectives on Processes and Paradigms. Appl. Environ. Microbiol. **68:**1015-1024.

120. **Zhang, H., J. Chen, and S. Guo.** 2008. Preparation of natural gas adsorbents from high-sulfur petroleum coke. Fuel **87:**304-311.

121. **Zhu, X. Y., J. Lubeck, and J. J. Kilbane, II.** 2003. Characterization of Microbial Communities in Gas Industry Pipelines. Appl. Environ. Microbiol. **69:**5354-5363.

122. **Zimmermann, R., R. Iturriaga, and J. Becker-Birck.** 1978. Simultaneous determination of the total number of aquatic bacteria and the number thereof involved in respiration. Appl. Environ. Microbiol. **36:**926-935.

123. **Zinger, L., L. A. Amaral-Zettler, J. A. Fuhrman, M. C. Horner-Devine, S. M. Huse, D. B. M. Welch, J. B. H. Martiny, M. Sogin, A. Boetius, and A. Ramette.** 2011. Global Patterns of Bacterial Beta-Diversity in Seafloor and Seawater Ecosystems. PLoS ONE **6:**e24570.

VITA

James Paul Davis

Candidate for the Degree of Microbiology

Doctor of Philosophy

Thesis:
BACTERIAL COMMUNITIES ASSOCIATED WITH NATURAL AND ANTHROPOGENIC PETROLEUM IMPACTED ANAEROBIC ENVIRONMENTS

Major Field: Microbial Ecology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Microbiology at Oklahoma State University, Stillwater, Oklahoma in December 2012.

Completed the requirements for the Bachelor of Science in Microbiology at University of Oklahoma, Norman, Oklahoma in 2006.

Experience:

Teaching Assistant and Research Assistant - Department of Microbiology and Molecular Genetics, Oklahoma State University, Stillwater, Oklahoma August 2006 through December 2012

Professional Memberships:

American Society of Microbiology
Oklahoma Academy of Science

Name: James Paul Davis Date of Degree: December, 2012

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: BACTERIAL COMMUNITIES ASSOCIATED WITH NATURAL AND ANTHROPOGENIC PETROLEUM IMPACTED ANAEROBIC ENVIRONMENTS

Pages in Study: 117 Candidate for the Degree of Doctor of Philosophy

Major Field: Microbiology

Scope and Method of Study: The aim of this dissertation was to determine the diversity, abundance, distribution, and morphology of candidate division SR1 from various habitats (including a natural gas impacted fresh water spring), and to determine the spatial and temporal changes in the bacterial communities associated with two newly drilled natural gas wells. *Methods:* Oligonucleotide primer and probe design; bulk DNA extraction from environmental samples; Polymerase Chain Reaction (PCR); quantitative PCR; Phylogenetic and diversity analysis; Fluorescence *in situ* hybridization; Beta-diversity analysis

Findings and Conclusions: This work demonstrates the complexity of the bacterial communities associated with natural gas environments. Candidate division SR1 appears to be only found in anaerobic environments, and certain environmental conditions allow for this 'rare biosphere' division to be a more prominent member of the bacterial community. It has been shown that SR1 has two morphotypes: filamentous and rod, and that cellular morphology was linked to phylogeny. The second half of this work demonstrated that the bacterial populations associated with natural gas well equipment fluctuate over time and space. The majority of the sequences were classified in the phyla: Firmicutes and Proteobacteria (classes Gamma- and Epsilonproteobacteria). Overall, the bacterial communities are halotolerant/halophilic with many hydrocarbon-degrading lineages. There was a significant sulfidogenic bacterial consortia present that increased over time. The bacterial communities in each well showed an increase in community similarity over time in each component. Spatially, the bacterial communities in the tanks were significantly different from the separators. The drilling mud and frac-water communities were not significantly similar to the production water communities in both wells.

ADVISER'S APPROVAL: Mostafa S. Elshahed, Ph.D.