12-2012

# Natural and Lignocellulose-Enriched Microbial Communities in Great Boiling Spring, NV

Jessica Cole
*University of Nevada, Las Vegas*

# NATURAL AND LIGNOCELLULOSE-ENRICHED MICROBIAL COMMUNITIES IN GREAT BOILING SPRING, NV

by

Jessica K. Cole

Bachelor of Science in Biology
University of Nevada - Las Vegas
2010

A thesis submitted in partial fulfillment
of the requirements for the

Master of Science in Biological Sciences

School of Life Sciences
College of Sciences
The Graduate College

University of Nevada, Las Vegas
December 2012

THE GRADUATE COLLEGE

We recommend the thesis prepared under our supervision by

Jessica Cole

entitled

Natural and Lignocellulose-Enriched Microbial Communities in Great Boiling Spring, NV

be accepted in partial fulfillment of the requirements for the degree of

**Master of Science in Biological Sciences**
School of Life Sciences

Brian Hedlund, Ph.D., Committee Chair

Helen Wing Ph.D., Committee Member

Eduardo Robleto, Ph.D., Committee Member

Ernesto Abel-Santos, Ph.D., Graduate College Representative

Tom Piechota, Ph.D., Interim Vice President for Research &
Dean of the Graduate College

**December 2012**

ABSTRACT

**Natural and Lignocellulose-Enriched Microbial Communities in Great Boiling Spring, NV**

by

Jessica K. Cole

Dr. Brian Hedlund, Committee Chair
Professor of Biology
University of Nevada, Las Vegas

The natural microbial communities present in Great Boiling Spring were investigated and contrasted against those present after *in situ* enrichment with lignocellulose. High-throughput cultivation-independent DNA sequencing of the V8 region of the small subunit (SSU) rRNA gene generated a total of 274,119 quality-filtered pyrosequencing fragments. Twelve natural spring samples were analyzed, including four high-temperature water samples and eight sediment samples ranging from 87 - 62 °C. Eight lignocellulosic enrichments incubated in the spring sediment and water at two high-temperature sites were analyzed. The natural water communities were found to be extremely uneven but relatively constant throughout time. The natural sediment communities clustered based according to sample temperature and location. This study demonstrates the strong influence of temperature on microbial diversity and composition of both natural and enriched samples, with a negative correlation between temperature and richness. A thermophilic and cellulolytic bacterium was isolated from one of the lignocellulosic enrichments. It was aerobic and able to utilize a wide variety of substrates as sole carbon and energy sources, including both amorphous and crystalline cellulose.

The results of phylogenetic, chemotaxonomic, and morphological analyses determined the strain to be representative of a new order-level lineage within the class "*Chloroflexi*".

ACKNOWLEDGEMENTS

I would like to thank Dr. Brian Hedlund for giving me the opportunity to perform such interesting research. I appreciate the guidance and advice he provided throughout the years I spent working in his laboratory. Also, I owe much to David and Sandy Jamieson, who generously allowed access to Great Boiling Spring, making my research possible.

My committee members, Drs. Eduardo Robleto, Helen Wing, and Ernesto Abel-Santos, deserve substantial thanks for the consideration and recommendations that they provided. I must thank everyone who provided feedback during my presentations, and Keala Kiko and Nicholle Booker for essential administrative support.

My fellow laboratory members were responsible for improving many aspects of my projects and publications with their support and advice. Some of these people include Tara Edwards, Mandy Williams, Jeremy Dodsworth, Joseph Peacock, Brandon Gieler, and Senthil Murugapiran. It was a great pleasure to work with each and every one of you.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

The cultivation-independent exploration of high-temperature microbial communities has historically been conducted using low-throughput technologies. However, such studies were inherently limited by the relatively few sequences processed, and most experiments did not include temporal or spatial sampling. In general, this resulted in only the most dominant members of the microbial communities being uncovered, and changes in community structure over space and time were not examined. The recent development of affordable high-throughput DNA sequencing technologies has allowed for the exploration of microbial communities with a depth and breadth previously unheard of. However, the application of this technology to the study of the spatial and temporal changes in the structure of thermophilic microbial communities has been extremely limited. Studies have investigated the spatial or temporal changes in a microbial community at temperatures up to 73 °C, the photosynthetic temperature limit, but only included members of the domain Bacteria in their censuses. No studies have examined the microbial community of hot spring water. Therefore, the knowledge garnered from such studies is limited in scope, describing only the bacteria existing in phototrophic communities located in the spring sediment, and excluding examination of the dynamics of the chemotrophic communities at higher temperatures in spring water and sediment.

Thermophilic microorganisms, such as those found in hot springs, are of significant interest to the biofuels industry. In particular, thermophiles are considered a

potential resource not only for their cellulolytic abilities, but also their thermostable enzymes. Thermostable enzymes are sought for their potential ability to economize the currently expensive process of converting energy-rich but recalcitrant materials into biofuels because they would increase the efficiency of the process by allowing it to proceed at higher temperatures. However, most studies of cellulolytic activity in thermophiles have focused on the exploration of the abilities of individual isolates, or that of cellulolytic communities that exist at temperatures of no more than 60 °C.

This study reports the use of high-throughput DNA sequencing of the SSU rRNA gene to examine both natural and lignocellulose-enriched microbial communities in the water and sediment of Great Boiling Spring (GBS). The spatial and temporal changes in the composition and diversity of the microbial community, including bacteria and archaea, were examined. In addition, the microbial community of a group of *in situ* lignocellulosic enrichments is described, and contrasted to the natural microbial communities, resulting in shifts in the thermophilic microbial communities incurred due to lignocellulosic enrichment.

As much as culture-independent studies have augmented human knowledge of the diversity of microorganisms, the isolation of microbes in pure culture is still crucial to describing a novel organism using traditional microbiological techniques. Pure cultures allow the exploration of the physiology of one organism without interference from other microbes that would be present in a mixed culture. Here, the isolation of a previously unknown group of closely-related cellulolytic and thermophilic bacterial strains is described, along with the characterization of one of the strains, JKG1$^{T}$. As a whole, this research helps to explore potential diversity and variation of the natural and cellulolytic

microbial communities present in GBS, reveals a strong, negative, and linear relationship between temperature and the richness of microbial communities, and describes the discovery of a previously undetected lineage of cellulolytic bacteria.

CHAPTER 2

# SEDIMENT MICROBIAL COMMUNITIES IN GREAT BOILING SPRING ARE CONTROLLED BY TEMPERATURE AND DISTINCT FROM WATER COMMUNITIES


## Abstract

Great Boiling Spring is a large, circumneutral, geothermal spring in the U.S. Great Basin. Twelve samples were collected from water and four different sediment sites on four different dates. Microbial community composition and diversity were assessed by PCR amplification of a portion of the small subunit (SSU) rRNA gene using an improved, universal primer set followed by pyrosequencing of the V8 region. Analysis of 164,178 quality-filtered pyrotags clearly distinguished sediment and water microbial communities. Water communities were extremely uneven and dominated by *Thermocrinis*. Sediment microbial communities grouped according to temperature and sampling location, with a strong, negative, linear relationship between temperature and richness at all taxonomic levels. Two sediment locations, Site A (87-80 °C) and Site B (79 °C), were predominantly composed of single phylotypes of the bacterial lineage GAL35 ($\bar{p}$=36.1%), *Aeropyrum* ($\bar{p}$=16.6%), the archaeal lineage pSL4 ($\bar{p}$=15.9%), the archaeal lineage NAG1 ($\bar{p}$=10.6%), and *Thermocrinis* ($\bar{p}$=7.6%). The ammonia-oxidizing archaeon "*Candidatus* Nitrosocaldus" was relatively abundant in all sediment samples <82 °C ($\bar{p}$=9.51%), delineating the upper temperature limit for chemolithotrophic ammonia oxidation in this spring. This study underscores the distinctness of water and

sediment communities in GBS and the importance of temperature in driving microbial diversity, composition, and, ultimately, the functioning of biogeochemical cycles.

## Introduction

Over the last twenty years, our understanding of life at high temperature has benefitted immensely from cultivation-independent censuses of microbial communities in continental geothermal springs, particularly those in Yellowstone National Park (YNP). The majority of these studies applied low-throughput technologies, relying on PCR-amplification, cloning, and Sanger sequencing of the small subunit (SSU) rRNA gene as a molecular identifier (Barns *et al.* 1994; Blank *et al.* 2002; Hugenholtz *et al.* 1998; Meyer-Dombard *et al.* 2005; Reysenbach *et al.* 2000 1994; Spear *et al.* 2005). In aggregate, these studies revealed that phyla that dominate most aquatic systems (*e.g.* *Proteobacteria*, *Firmicutes*, *Actinobacteria*, and *Bacteroidetes*) are replaced with thermophilic specialists at temperatures ≥80 °C and the surprising abundance of thermophilic bacteria, particularly *Aquificae* (Reysenbach *et al.* 2005; Spear *et al.* 2005). In addition, this research uncovered unexpectedly high diversity of yet-uncultivated lineages of archaea, in spite of the predominance of *Thermoprotei* among cultivated terrestrial thermophiles (Spear *et al.* 2005). Despite their utility, these censuses were limited by the absence of spatial or temporal sampling and their relative shallowness; thus, they provided only a snapshot of the community and an incomplete understanding of factors that influence microbial community composition and diversity.

A few studies have addressed the relationship between temperature and microbial community structure by studying temperature gradients along geothermal outflows. Two studies examined $S^0$ and $As^{5+}/Fe^{3+}$-precipitating geothermal outflows from sulfuric acid-

buffered springs in YNP. Jackson *et al.* (Jackson *et al.* 2001) noted a transition from a community dominated by *Hydrogenobaculum* and *Desulfurella* in the 62.1 °C source and $S^0$ precipitation zone of Dragon Spring to a community with abundant novel archaea in the cooler $As^{5+}/Fe^{3+}$ precipitation zone below 55 °C. A similar study of a diverted outflow from Perpetual Spring documented the presence of *Stygioglobus*, *Caldococcus*, *Caldisphaera*, *Thermocladium*, and *Hydrogenobaculum* in mature biofilms in $S^0$-depositing zones of ~80 to 70 °C, yielding to possibly less diverse communities in the cooler $As^{5+}/Fe^{3+}$-precipitating zones (Macur *et al.* 2004). Other studies focused on geothermal outflows in circumneutral springs in YNP. Meyer-Dombard *et al.* (Meyer-Dombard *et al.* 2011) examined geothermal outflows from four springs in Sentinel Meadows in the Lower Geyser Basin, focusing on comparing microbial communities in similar springs with and without streamer biofilms. All springs contained abundant *Thermocrinis* at the source, with changes in diversity that were interpreted as a possible ecotone (transitional area of overlap between two communities containing species characteristic of both communities) in the transitional area between hotter chemosynthetic and cooler photosynthetic systems in two of the springs.

Recent advances in DNA sequencing technologies have allowed for much deeper microbial community censuses through pyrosequencing of SSU rRNA gene fragments (pyrotags); however, to our knowledge, only one study has employed pyrotag sequencing to study the relationship between temperature and microbial community composition and diversity in terrestrial geothermal systems (Miller *et al.* 2009). The study examined photosynthetic communities in two silica-depositing geothermal outflows, White Creek and Rabbit Creek, in the Lower Geyser Basin of YNP between 73 °C and 39 °C.

Operational taxonomic unit (OTU) richness at 99% identity was negatively related to temperature and community similarity decreased exponentially with temperature difference, demonstrating that temperature was an important factor in controlling community composition.

Great Boiling Spring (GBS) is a large circumneutral spring in the U.S. Great Basin, the microbiology and geochemistry of which have been studied in some detail (Costa *et al.* 2009; Dodsworth *et al.* 2011; Dodsworth *et al.* 2012; Hedlund *et al.* 2011; Huang *et al.* 2007; Lefèvre *et al.* 2010; Miller-Coleman *et al.* 2012). GBS is sourced with ~85 °C $Na^+/Cl^-$-dominated hydrothermal fluid that accumulates in the source pool for ~1.5 days, leading to oxic conditions in the bulk water (Costa *et al.* 2009). The geothermal source contains ammonia as the dominant form of dissolved inorganic nitrogen, which supports a highly active nitrogen cycle in which ammonia is oxidized to nitrite by a close relative of the ammonia-oxidizing archaeon "*Candidatus* Nitrosocaldus yellowstonii" (Dodsworth *et al.* 2011). In turn, oxidized nitrogen supports both dissimilatory nitrate reduction to ammonium and denitrification, which is leaky at the step of nitrous oxide reduction (Dodsworth *et al.* 2011; Hedlund *et al.* 2011). Cultivation-independent censuses of the microbial community in GBS performed using low-throughput Sanger sequencing revealed a high proportion of novel lineages of archaea and bacteria. (Costa *et al.* 2009). This high degree of novelty was further supported by pyrosequencing of single samples of water and sediment communities, which also defined a stark contrast between the two communities (Dodsworth *et al.* 2011).

The present study uses SSU rRNA gene pyrotag sequencing to examine microbial community composition and diversity in GBS within the context of spatial and temporal

sampling. The primary goals were: (i) to determine the effect of temperature on community composition, diversity, and inferred function (ii) to determine whether microbial communities in sediments and bulk water are similar or distinct, and (iii) to evaluate the temporal stability of the microbial community in GBS in the face of temperature fluctuations. The current study assesses both archaea and bacteria at all taxonomic levels and includes analysis of water-borne cells as well as sediment samples above, near, and below the photosynthetic temperature limit at several different sampling times.

## Methods

Sample collection, DNA extraction, PCR, and pyrosequencing

GBS is located at N40°39'41" W119°21'58". Five sample locations, Sites A, B, C, D, and E, (Figure B.1) were identified within GBS for sediment sampling. The sediment/water interface (top ~1 cm) was sampled using sterile technique. Bulk water samples (Site W) were collected using either normal filtration, with a 0.2 μm Supor filter (hydrophilic polyethersulfone, Pall Corporation, Port Washington, NY) or tangential flow filtration (TFF; Prep/Scale filter with 30 kDa molecular weight cut-off, Millipore, Billerica, MA, USA). Sampling was performed at all sites in Feb., 2010, with additional samples taken at a subset of the sites in July, 2007, Dec., 2008, June, 2009, and July, 2010 (Table 1). Samples are identified in the text by a four number code indicating the month and year of sampling (in YYMM format) followed by a letter indicating the sample site, *e.g*. 0706W refers to the bulk water sample taken during June 2007. Sediment and normal flow filters were placed on dry ice immediately and transported to the laboratory where they were stored at -80 °C until processing. TFF concentrates were

stored on ice for <24 hours after which cells were pelleted by centrifugation at 6,000 rcf for 10 minutes and stored at -80 °C until processing. Temperature and pH were measured in the bulk water using a LaMotte pH5 Series pH/temperature meter (LaMotte, Chestertown, MD, USA).

DNA was extracted from sample 0812W using the Joint Genome Institute's (JGI) CTAB protocol (Department of Energy Joint Genome Institute & California 2010). DNA was extracted from all other samples using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA) using modifications specified previously (Dodsworth *et al.* 2011). DNA was precipitated with 70% ethanol and resuspended in 0.5 x TE (5 mM Tris, 0.5 mM EDTA, pH 8). DNA retrieved from the sediment sample collected at Site D was insufficient for PCR and therefore Site D was excluded from further analysis. DNA was shipped on dry ice to the JGI, where a portion of the SSU rRNA gene, including the V6-V8 hypervariable regions, was amplified using JGI's universal PCR primers 926F (5'- cct atc ccc tgt gtg cct tgg cag tct cag AAA CTY AAA KGA ATT GAC GG- 3') and 1392R (5' - cca tct cat ccc tgc gtg tct ccg act cag - <XXXXX> - ACG GGC GGT GTG TRC – 3'). Primer sequences were modified by the addition of 454 A or B adapter sequences (lower case). In addition, the reverse primer included a 5 bp barcode for multiplexing of samples during sequencing. Twenty µL PCR reactions were performed in duplicate and pooled to minimize PCR bias using 0.4 µL Advantage GC 2 Polymerase Mix (Advantage-2 GC PCR Kit, Clontech, Mountain View, CA, USA), 4 µL 5X GC PCR buffer, 2 µL 5M GC Melt Solution, 0.4 µL 10mM dNTP mix (MBI Fermentas, Glen Burnie, MD, USA), 1.0 µL of each 25 nM primer, and 10 ng sample DNA. The thermal cycler protocol was 95°C for 3 min, 25 cycles of 95°C for 30 s, 50°C for 45 s, and 68°C

for 90 s, and a final 10-min extension at 68°C. PCR amplicons were purified using SPRI Beads and quantified using a Qubit flurometer (Invitrogen, Carlsbad, CA, USA). Samples were diluted to 10 ng/μL and mixed in equal concentrations. Emulsion PCR and sequencing of the PCR amplicons were performed following the Roche 454 GS FLX Titanium technology (454 Life Sciences, Branford, CT, USA) manufacturer's instructions. Results obtained with the universal JGI primer set used for amplification indicated a lack of detection of archaea known to be present in GBS (data not shown). Comparison of the standard JGI forward primer with near full-length SSU rRNA gene sequences obtained from GBS (Costa *et al.* 2009) and other geothermal environments revealed that the corresponding sequence in most archaea had at least one mismatch to the forward primer (data not shown). Therefore, a revised forward primer, 926F454TitFNew (AAA CTY AAA KGA ATT GRC GG), was designed and used for pyrotag amplification. Choice of primer pair during PCR can heavily influence estimates of microbial richness and evenness due to varying degrees of binding efficiency between primers and the sequences of the various SSU rRNA gene primer binding sites present in the sample (Engelbrektson *et al.* 2010). The use of the newly designed primer pair led to a dramatic increase in species richness compared to the dataset obtained from the standard primer set (data not shown).

Sequence processing

285,719 pyrotags from 12 samples were quality filtered, aligned, and analyzed using the mothur software package v. 1.20.2 (Schloss *et al.* 2009). Quality filtering discarded sequences with ambiguous base calls, homopolymers of seven nt or longer, and sequences with an average quality score of less than 27 (as recommended by Kunin &

Hugenholtz 2010) over a window size of 50 nt. The mothur-provided bacterial and archaeal Silva reference files were concatenated into a single file against which quality-filtered sequences were aligned using default Needleman-Wunch algorithm parameters. The alignment was manually curated and sequences that did not end at nt 1,386 in the *E. coli* SSU rRNA gene were discarded. Pre-clustering was performed at a level of 1% difference. The alignment was checked for chimeras using mothur's chimera.slayer command (modeled after ChimeraSlayer by Broad Institute, Cambridge, MA, USA; http://microbiomeutil.sourceforge.net/#A_CS) with abundant sequences in the sample set serving as the reference set. Sequences with very low abundance or those that did not align appropriately were searched against the NCBI BLAST database using blastn to further evaluate sequence quality (Zhang *et al.* 2000). Confirmed and possible chimeric sequences, eukaryotic sequences, chloroplast sequences, and low-abundance sequences with little similarity to any known organism were discarded, resulting in a curated set of 164,178 sequences, 200-230 nt in length, approximately covering bases 1,156-1,377 (including the V8 hypervariable region) of the *E. coli* SSU rRNA gene. Sequences passing quality filters have been submitted to the NCBI Sequence Read Archive.

The Quantitative Insights into Microbial Ecology (QIIME) software package v 1.3.0 (Caporaso *et al.* 2010) was used to designate OTUs using the UCLUST algorithm (Edgar 2010). Taxonomy was assigned according to the QIIME-compatible GreenGenes database-derived OTU set released on 4$^{th}$ Feb, 2011 (DeSantis *et al.* 2006) using the BLAST algorithm, and also the QIIME-compatible pre-built Ribosomal Database Project (RDP) using the RDP classifier algorithm (Cole *et al.* 2009). Each OTU was assigned to the most detailed lineage in the database matching ≥90% of the sequences within the

OTU. Where the taxonomy assignments of the 97% OTUs derived from the GreenGenes database and the RDP classifier disagreed, the representative sequences of those OTUs were submitted to NCBI BLAST to further assess the taxonomic assignment (Johnson *et al.* 2008). The GAL35 group is currently one of two candidate class-level groups within the candidate phylum OP1 but was given a phylum-level designation in a previous edition of the GreenGenes phylogeny (29 November, 2010). In our own phylogenetic analyses, the position of the GAL35 group is uncertain, but may be affiliated with the *Chloroflexi* (Costa *et al.* 2009 ;Dodsworth *et al.* unpublished results). We therefore chose the conservative approach of retaining its designation as a distinct lineage rather than grouping it with other previously named phylum-level groups.

Statistical analyses

Rarefaction and collector's curves were generated for OTUs observed at the 97% OTU level using mothur. Rarefaction curves were prepared at sequencing depth of 915 and below for archaea and 4,410 and below for bacteria. QIIME was used to calculate alpha diversity measures at OTU definitions of 80, 85, 92, 95 and 97% sequence identity, including OTUs observed, Chao1 (Chao 1984), the Shannon Index, and Simpson's Index of Diversity. Simpson's Evenness was calculated in Microsoft Excel 2007.

QIIME was used to perform beta diversity analyses. Bray-Curtis dissimilarity (Bray & Curtis 1957) was calculated at the 97% OTU level and principal coordinates analysis (PCoA) and hierarchical cluster trees were constructed from the resulting matrix. A hierarchical cluster tree at high taxonomic level (phylum level for bacteria and class level for archaea) was constructed based upon taxonomy assigned to 97% OTUs.

Analysis of Similarity (ANOSIM) and Similarity Percentage (SIMPER) calculations were conducted using PAST v. 2.11 (Hammer *et al.* 2001). ANOSIM was used to test for significant differences between sample groups defined by the 97% OTU cluster tree. Holm's sequential Bonferroni correction was used to correct for multiple comparisons (Holm 1979). Those groups found to be significantly different according to uncorrected p-values were subjected to SIMPER analysis based on Bray-Curtis dissimilarity in order to define the OTUs primarily responsible for the differences between the groups.

Pearson's $\chi^2$ goodness-of-fit calculations and linear regressions were performed using the R Project for Statistical Computing (R Development Core Team 2011). In regressions of diversity versus temperature, sediment community samples collected at the same location in different years were treated as independent samples because the temperatures differed between years and the time intervals between sampling were sufficiently long that communities could adjust to changed temperature conditions.

Mineralogy

Samples were dried at ambient temperature in an anaerobic chamber (Coy, Gross Lake, MI, USA) with an atmosphere of 90% $N_2$, 5% $CO_2$, and 5% $H_2$. Samples were then scanned with a Scintag X1 X-ray powder diffractometer (Scintag, Cupertino, CA, USA) using CuKα wavelength, a fixed slit scintillation detector, and a power of 1400 W (40 kV, 35 mA). Scans were performed from 2 to 70$^o$ 2-theta stepping at 0.02$^o$ with a count time of 2 s per step. Qualitative identification of mineral phase was made by using the Jade 7 program (MDI, Livermore, CA, USA). This program utilized the International Center for

Diffraction Data Powder diffraction File database (ICDD PDF-2, Sets 1-46, 1996) as a reference source.

## Results and Discussion

Overview of pyrotag dataset

A total of 164,178 quality-filtered SSU rRNA gene pyrotags were obtained from twelve samples collected on up to four sampling dates from bulk water (Site W) and four sediment sites (Sites A, B, C, and E; Figure B.1). Total community OTU abundance ranged from 25 to 252 at the 97% level and from 10 to 27 at the 80% level, roughly corresponding to species and phyla, respectively (Table A.1). The V8 region, which possesses less variation than the unmasked full-length SSU rRNA gene, was the only hypervariable region contained within the quality-filtered pyrotags. Use of only the V8 region likely caused an underestimation of the richness of the communities sampled in this study (Youssef *et al.* 2009). Collector's curves suggested that sampling did not approach completion at any sample location (Figure B.2).

Water and sediment communities are distinct

Distinct clustering of microbial communities from water and sediment samples was revealed by cluster analysis calculated at the phylum/class level (Figure 2.1a) and the species level (Figure 2.2a). Furthermore, the water and sediment communities were separated by the primary axis of variation by PCoA performed at the species level (Figure 2.2b). Highly significant differences were confirmed between water and sediment communities by ANOSIM ($p<0.01$, corrected $p<0.05$; Figure B.3, Node 1). In order to isolate differences between water and sediment communities from variation in sample temperature, the microbial communities in the four water samples were compared against

sediment samples collected within the same temperature range (82-80 °C; 0812A, 0906A, 1002A) using ANOSIM. This analysis demonstrated a significant difference between the communities in the water and equivalent-temperature sediment samples ($p<0.01$, corrected $p<0.05$; Figure B.3, AB-W). Similar richness was found in the water and sediment communities at all taxonomic levels by alpha diversity calculations with the same temperature-equivalent sample set; however, the water community was significantly less even at OTU levels from 95%, 92%, and 85%, roughly corresponding to the level of genus, family, and order, respectively, as indicated by Student's t-tests (Figure 2.3; Tables A.1-A.3).

Species that contributed to the dissimilarity between water and sediment samples were identified using SIMPER analysis (Table 2.2). Two OTUs, a *Thermocrinis* sp. ($\bar{p}$=89.47%) and *Pyrobaculum caldifontis* ($\bar{p}$=8.63%), dominated water communities, yet each was present at low relative abundance in sediment communities. The high relative abundance of *Thermocrinis* and *Pyrobaculum* in the water column could result from these populations shedding from sediment communities; however, the lower evenness and higher abundance of cells in the GBS water column compared with a nearby spring with a much shorter water residence time (Dodsworth *et al.* 2011) suggest that these two genera live and reproduce in the water column. Other taxa that contributed to dissimilarity of water and sediment communities were more abundant in sediments, including an *Aeropyrum* sp. and an OTU related to the novel archaeal lineage NAG1 (Kozubal *et al.* 2012; Table A.4).

The distinctness of water and sediment communities is well known in most natural aquatic systems and has been identified in a large meta-analysis as a primary

driver delineating microbial community composition (Lozupone & Knight 2007). These differences segregate biogeochemical functions in natural aquatic systems. For example, in marine systems, the photic zone of the water column is characterized by primary productivity and highly dynamic recycling of both elements and energy (*i.e.* the "microbial loop"), whereas the deep sea water column and sediments are essentially heterotrophic systems that are responsible for mineralization of sinking material from above (reviewed in Fenchel, King, and Blackburn 1998). While biologists studying thermophilic viruses focus entirely on water-borne material, the majority of research on the biology of geothermal systems focuses entirely on sediments or mineral precipitates and potentially overlooks an important community of organisms in the water column.

Temperature controls richness and composition of sediment communities at all taxonomic levels

The effects of temperature on microbial communities were explored using two different approaches: first, by determining whether microbial composition clusters concordant with differences in temperature, and, second, by determining whether quantitative relationships exist between measures of OTU diversity and temperature (Table A.1). Cluster analysis calculated at the phylum/class level (Figure 2.1a) and the species level (Figure 2.2a) showed that sediment samples clustered according to temperature and PCoA at the species level showed that temperature corresponded with the secondary axis of variation (Figure 2.2b). A significant difference between sediment communities that were divided into three temperature groups, 87 to 79 °C (Sites A and B), 76 to 72 °C (Site C), and 62 °C (Site E ;$p<0.05$, corrected $p<0.05$) was confirmed using ANOSIM (Figure B.3b).

Within the sediment samples, an inverse relationship between microbial diversity and temperature was exhibited by both Simpson's and Shannon Diversity Indices (Tables A.1-A.3). For example, both Simpson's and Shannon Diversity values for 97% OTUs displayed negative, linear relationships with temperature ($R^2$=0.60 and $R^2$=0.76, respectively).

To disentangle the components of alpha diversity, the relationships between temperature and evenness and richness were assessed separately. Using the combined archaea and bacteria dataset, richness at every OTU level showed a highly significant, negative, linear relationship with temperature both before and after correction for multiple comparisons ($R^2 \geq 0.88$, $p<0.001$, corrected $p<0.01$; Figure 2.4a). The estimated slopes (OTUs per °C ± S.E.) ranged from -8.57 ± 1.37 for 97% OTUs to -0.73 ± 0.09 for 80% OTU richness (Figure 2.4a). When assessed separately, bacterial richness showed a significant, negative, linear relationship with temperature at all OTU definitions ($R^2 \geq 0.88$, $p<0.001$, corrected $p<0.001$; Figure B.4b), whereas archaeal richness only showed a significant relationship with temperature at the 80% ($R^2$=0.89 $p<0.001$, corrected $p<0.01$) and 85% OTU levels ($R^{2=}$0.64, $p<0.05$, corrected $p<0.1$ ;Figure B.4a). A similar negative relationship between estimated richness and temperature was shown by the Chao1 richness estimator; however, the relationship was weaker than that observed for OTUs (Table A.4).

Many factors typically covary with temperature in geothermal systems, due to processes such as degassing, mineral precipitation, evaporation, autotrophy, and oxidation (Fouke 2011; Nordstrom *et al.* 2005). In GBS, cooler temperature covaried negatively with dissolved oxygen concentration (r=-0.88 ;Figure B.5). Alpha diversity

measurements also covaried to some degree with pH and oxygen (Appendix C). However, multiple regressions of combined OTU richness against the chemical variables showed that temperature accounted for more variation in richness than any chemical variable. In multiple regressions of 80%, 85%, and 95% OTU levels, richness against physicochemical variables, there were highly significant, negative relationships between richness and temperature after adjusting for other environmental variables ($R^2$ =0.99, p=0.006; $R^2$ =0.99, p=0.006; $R^2$ =0.97, p=0.009; respectively). Although other unmeasured physicochemical variables cannot be ruled out as factors that influence microbial community composition, these results suggest temperature is the major factor driving community composition and structure. The homogeneous mineralogy of GBS sediments (Table 2.1) eliminated sediment heterogeneity as a factor.

Statistical analyses performed with the combined dataset (Figure 2.4b) or the bacteria-only dataset (Figure B.5d) revealed no significant relationships between temperature and evenness at any OTU definition. However, a significant, positive, linear relationship between evenness and temperature was evident in the archaeal dataset at OTU levels 80, 85, 92, and 95% before correction and 80 and 95% after correction (Figure B.5c).

To further investigate the relationship between temperature and diversity, we examined the community similarity as a function of pairwise temperature difference between samples. This revealed an exponential decay in community similarity as the temperature difference between samples increased (Figure 2.5). Our results were similar to those observed by Miller *et al.* (Miller *et al.* 2009) in the photosynthetic zones of two alkaline geothermal outflows in YNP, ranging from 74 to 39 °C. However, we observed a

much steeper decline in species richness as temperature increased within the temperature range we studied, 87 to 62 °C, suggestive of a much stronger selective force at high temperature. To our knowledge, diversity decay dynamics along a temperature gradient had not rigorously been addressed at temperatures above the photosynthetic limit.

Composition of sediment communities in relation to temperature

SIMPER was used to identify taxa that contributed most to differences between sediment communities in the three temperature groups confirmed to be significant by ANOSIM, 87 to 79 °C (Sites A and B), 76 to 72 °C (Site C), and 62 °C (Site E ;Table 2.3). The highest temperature sites, A (87-80 °C) and B (79 °C), were predominantly composed of single phylotypes of the bacterial lineage GAL35 ($\bar{p}$=36.1%), *Aeropyrum* ($\bar{p}$=16.6%), the archaeal lineage pSL4 ($\bar{p}$=15.9%), the archaeal lineage NAG1 ($\bar{p}$=10.6% ;Figure B.6c), and *Thermocrinis* ($\bar{p}$=7.6%). These five species were present at lower temperature Sites C (76-72 °C) and E (62 °C), but were significantly less abundant ($\chi^2$=77909, p<0.0001; Figure 2.1b, Figure B.7). Bacteria made up a significantly higher percentage of the communities at cooler temperature Sites C and E (73.7% versus 59.9%; $\chi^2$=2076, p<0.0001; Figure 2.1b). Species that were particularly abundant at Site C (76-72 °C) included a phylotype in the *Armatimonadetes* ($\bar{p}$=15.9%), a distinct phylotype in the archaeal lineage pSL4 ($\bar{p}$=15.9%), and "*Ca*. N. yellowstonii" ($\bar{p}$=8.89%). Although most phylum- and class-level OTUs were represented by a single abundant species in the pyrotag dataset (>1%), *Armatimonadetes* and pSL4 were exceptional in that temperature specialization was apparent (Figure B.6a Figure B.6b). The recently discovered phylum *Armatimonadetes* (Stott *et al.* 2008) was represented primarily by three 97% OTUs, one of which was abundant at Site C ($\bar{p}$=15.9%), and the others which were only present at

Site E. pSL4 was also represented primarily by three OTUs with one dominant at high temperature sites (Figure B.7b, blue) and another dominant at lower temperature sites (Figure B.7b, purple).

Inferences on upper temperature limits of biological functions

The distribution and abundance of pyrotags can offer clues to the spatial distribution of biogeochemical functions within the spring. Two species-level lineages were identified as "*Ca.* N. yellowstonii". The more abundant of these OTUs comprised 5-15% of all pyrotags at Sites A, B, and C when sampled between 80 °C and 72 °C and shared 99% identity with "*Ca.* N. yellowstonii" (EU239960) over the full length of the pyrotag. The less abundant OTU represented ~6% of pyrotags at Site E (62 °C) and was 97% identical to "*Ca.* N. yellowstonii" (EU239960) over the full length of the pyrotag. Previous work has documented relatively high rates of ammonia oxidation in GBS sediment (82 and 81 °C) and extremely high abundance of "*Ca.* N. yellowstonii" (*amoA* 3.5–3.9 x $10^8$ and SSU rRNA gene 6.4–9.0 x $10^8$ copies $g^{-1}$ sediment ;Dodsworth *et al.* 2011). These results delineate an upper temperature limit of ~82 °C for ammonia oxidation by "*Ca.* N. yellowstonii" in GBS, yet "*Ca.* N. yellowstonii" has heretofore only been cultivated at temperatures up to 74 °C in the laboratory (de la Torre *et al.* 2008). The cause of the ~8 °C difference in upper temperature for laboratory growth and abundance and activity *in situ* remains to be elucidated.

In GBS, ammonia oxidation limits denitrification, which seems to be heterotrophic at 81 and 79 °C and is carried out at least in part by *Thermus thermophilus* strains that produce nitrous oxide as the terminal denitrification product (Dodsworth *et al.* 2011; Hedlund *et al.* 2011). *T. thermophilus* was most abundant in pyrotags at Sites B (79

°C) and C (76 °C ;$\bar{p}$=2.2%), but was replaced by other *Thermus* species, primarily *T. scotoductus*, at lower temperatures ($\bar{p}$=9.1% ;Figure B.8). However, the low temperature sites likely hosted other denitrifiers as well. The temperature specialization of *T. thermophilus* at relatively high temperature sites in GBS, in combination with the low species richness at those sites, may drive the observed high nitrous oxide flux at high temperature in GBS sediment (Hedlund *et al.* 2011).

Site E (62 °C) was well below the known upper temperature limit for photosynthesis (Brock 1967; Cox *et al.* 2011) and contained OTUs belonging to known photosynthetic genera. *Synechococcus* and *Chloroflexus* were each represented by at least one 97% OTU of ~3.57% relative abundance at Site E. However, neither organism was detected at Site C (76 and 72 °C), possibly due to temperature fluctuations that were common in GBS (data not shown). An OTU identified as a member of an unknown genus of *Rhodospirillaceae* was present at Site E ($p$=1.41%); however, phototrophy is paraphyletic within that family.

Abundance of novel lineages

Novel phylum- and class-level groups had highest relative abundance in high temperature sediments (Sites A,B $\bar{p}$=55.64%), decreasing at Site C ($\bar{p}$=48.90%), and lowest at Site E ($p$=12.41% ;Figure B.9). The novel bacterial lineage GAL35 and the novel archaeal lineage pSL4 dominated this trend. GAL35 has been detected in geothermal springs in YNP, Chile, Bulgaria, and China, and it has never been detected in non-thermal environments (DeSantis *et al.* 2006). In GBS, GAL35 had highest relative abundance at Site A and Site B when they were sampled at 80 and 79 °C ($\bar{p}$=52.41%), but it was relatively abundant in all pyrotag datasets except Site E. pSL4 is affiliated with the

recently-proposed phylum "*Candidatus* Aigarchaeota", an uncultivated group defined by environmental 16S rRNA gene sequences found in a variety of thermal environments and dominant in some anaerobic hydrothermal sediments and mats (de la Torre *et al.* 2008; Marteinsson *et al.* 2001; Nunoura *et al.* 2005; Nunoura *et al.* 2010; Nunoura *et al.* 2011). A composite genome of a representative of "*Candidatus* Aigarchaeota", "*Candidatus* Caldiarchaeum subterraneum", was found to contain an uptake hydrogenase and aerobic carbon monoxide dehydrogenase, suggesting this organism could exhibit a chemolithotrophic metabolism (Nunoura *et al.* 2005). In GBS, pSL4 was most abundant at Site A at 87 °C ($\bar{p}$=54.13%) and was present in at least 3.85% relative abundance in all pyrotag datasets.

Novel genus-level groups had highest relative abundance in high temperature sediments (Site A,B $\bar{p}$=70.66%), decreasing slightly at Site C ($\bar{p}$=67.58%), and lowest at Site E ($p$=35.32% ;Figure B.10). The novel archaeal lineage NAG1 and the novel bacterial lineage OS-L were the most abundant lineages unidentified at the genus level. NAG1 sequences have been obtained from One Hundred Springs Plain Spring in YNP (Kozubal *et al.* 2012). In GBS, NAG1 had highest relative abundance in Site A at 82 °C ($\bar{p}$=20.47%) and was not detectable at temperatures below 79 °C. The novel bacterial lineage OS-L is a member of the recently-proposed phylum *Armatimonadetes,* formerly known as OP10*,* cultivated members of which include aerobic heterotrophs (Lee *et al.* 2011; Stott *et al.* 2008; Tamaki *et al.* 2011). In GBS, OS-L was most prevalent at samples taken between 76 and 72 °C ($\bar{p}$=8.50%) and was undetectable at temperatures above 76 °C.

## Conclusions

The current study underscores a need to consider water and sediment communities of geothermal springs separately, particularly in long residence time springs, which permit growth of a specific water-borne population. In GBS, the segregation of nitrogen-cycle functions between sediment and water communities has been clearly documented (Dodsworth *et al.* 2011; Hedlund *et al.* 2011) and other functions are likely segregated as well. Although a number of studies have examined microbial community composition and structure along geothermal outflows, the relative depth of the pyrotag approach employed here supported a strong statistical framework and incisive analyses. We observed a strong, negative, linear relationship between temperature and richness at all taxonomic levels and illustrated the strong selective force of temperature within the range that was sampled. Temperature also affects ecosystem functioning. In this study, we documented the upper temperature limit of chemolithotrophic ammonia oxidation at ~82 °C in situ. In addition, we identified the thermal niche of *Thermus thermophilus*, which is likely responsible for high nitrous oxide flux from high temperature GBS sediments (Dodsworth *et al.* 2011; Hedlund *et al.* 2011). This study also revealed a high abundance of novel lineages in GBS, particularly in the highest temperature sediments. Five species-level OTUs comprised >82.17% of all pyrotags in sediment samples ≥82 °C, and three of these OTUs were unidentified at the phylum or class level (GAL35, pSL4 and NAG1). This high degree of novelty underscores the need to uncover the roles of these organisms in order to understand the biology of this ecosystem.

**Table 2.1.** Physical, chemical, and mineralogical data and sequencing statistics.

| Sample Name | Site[1] | Date (m-y) | Temp. (°C)[2] | pH | Predominant Minerals[3] | Oxygen (μM) | $NH_4^+$ (μM) | $NO_2^-$ (μM) | $NO_3^-$ (μM) | N (Initial)[4] | N (Final)[5] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0706W | W | Jun-07 | 80 | 7.40 | n/a | n/a | 71 | 1.90 | 14.0 | 15,003 | 9,781 |
| 0812W | W | Dec-08 | 80 | 7.20 | n/a | 0.9 | 70 | 0.79 | 1.7 | 35,116 | 17,624 |
| 0906W | W | Jun-09 | 82 | 6.85 | n/a | 1.0 | 93 | 1.04 | 8.8 | 12,266 | 8,046 |
| 1002W | W | Feb-10 | 82 | 6.63 | n/a | 0.6 | 69 | (bd)[7] | 7.2 | 21,841 | 13,891 |
| 0812A | A | Dec-08 | 80 | 7.20 | Al, Am, Mu, Q | 0.9 | 70 | 0.79 | 1.7 | 57,106 | 37,890 |
| 0906A | A | Jun-09 | 82 | 6.85 | As, Mu,Q | 1.0 | 93 | 1.04 | 8.8 | 10,511 | 5,706 |
| 1002A | A | Feb-10 | 82 | 6.63 | An, L, Mu, Q | 0.8 | 69 | (bd) | 7.2 | 27,235 | 17,944 |
| 1007A | A | Jul-10 | 87 | n/a | Mu, Q | n/a | n/a | n/a | n/a | 26,687 | 6,689 |
| 1002B | B | Feb-10 | 79 | 6.81 | Mu, Q | 0.8 | 25 | (bd) | 3.0 | 30,813 | 20,720 |
| 1002C | C | Feb-10 | 72 | 7.05 | Mi, Mu, Pa, Ps, Q | 1.8 | 15 | 0.80 | 1.6 | 15,877 | 9,768 |
| 1007C | C | Jul-10 | 76 | n/a | Mu, Q | n/a | n/a | n/a | n/a | 16,742 | 5,907 |
| 1002E | E | Feb-10 | 62 | 7.24 | Mu, O, Q | 1.8 | 47 | 0.80 | 9.1 | 16,522 | 10,212 |
| | | | | | | | | | Sums: | 285,719 | 164,178 |

[1] See Figure S1 for site locations in GBS.

[2] Temperature for sediment samples recorded in water directly above sample sites.

[3] Al, albite; Am, analcime; An, andalusite; As, andesine; L, labradorite; Mi, microcline; Mu, muscovite; O, orthoclase; Pa, paragonite; Ps, pseudomalachite; Q, quartz.

[4] Pyrotags generated by 454 sequencing

[5] Pyrotags remaining after chimera check and other quality filters.

[6] n/a, Data not collected.

[7] (bd) Below detection limit (<0.1μM).

**Table 2.2.** SIMPER analysis results displaying top ten 97% OTUs responsible for dissimilarity between water and sediment samples[a].

| Taxon[b] | Genus | Contrib. (%)[c] | Mean abund.[d] Wat. (%) | Mean abund. Sed (%) |
|---|---|---|---|---|
| *Aquificae* (B) | *Thermocrinis* | 45.31 | 89.47 | 6.73 |
| GAL35 (B) | Unidentified GAL35 | 14.77 | 0.20 | 27.18 |
| *Thermoprotei* (A) | *Aeropyrum* | 6.14 | 0.07 | 11.27 |
| pSL4 (A) | Unidentified pSL4 | 5.88 | 0.04 | 10.77 |
| *Thermoprotei* (A) | *Pyrobaculum* | 4.19 | 8.63 | 1.87 |
| *Thermoprotei* (A) | Unidentified NAG1 | 3.64 | 0.02 | 6.68 |
| *Thaumarchaeota* (A) | *Candidatus* Nitrosocaldus | 2.78 | 0.02 | 5.09 |
| pSL4 (A) | Unidentified pSL4 | 2.00 | 0.00 | 3.64 |
| *Deinococcus-Thermus* (B) | *Thermus* | 1.90 | 0.28 | 3.43 |
| *Armatimonadetes* (B) | Unidentified OS-L | 1.15 | 0.01 | 2.10 |

[a]Pooled water samples compared to pooled sediment samples.
[b]Phylum-level taxa for bacteria and class-level taxa for archaea.
[c]Contribution of OTU to overall dissimilarity between groups.
[d]Average abundance of OTU in each group.

**Table 2.3.** SIMPER analysis results displaying top ten 97% OTUs responsible for dissimilarity between high temperature (A,B) vs. medium temperature (C) vs. low temperature (E) sediment[a].

| Taxon[b] | Genus | Contrib.[c] (%) | Mean abund. A,B (%)[d] | Mean abund. C (%) | Mean abund. E (%) |
|---|---|---|---|---|---|
| GAL35 (B) | Unidentified GAL35 | 17.15 | 36.16 | 17.76 | 0.95 |
| *Thermoprotei* (A) | *Aeropyrum* | 8.85 | 16.63 | 3.51 | 0.00 |
| pSL4 (A) | Unidentified pSL4 | 8.38 | 15.91 | 3.27 | 0.01 |
| *Deinococcus-Thermus* (B) | *Thermus* | 6.66 | 0.08 | 3.04 | 21.06 |
| Thermoprotei (A) | Unidentified NAG1 | 6.20 | 10.64 | 0.08 | 0.00 |
| pSL4 (A) | Unidentified pSL4 | 5.40 | 0.57 | 11.30 | 3.80 |
| *Thaumarchaeota* (A) | *Candidatus* Nitrosocaldus | 3.95 | 4.35 | 8.89 | 1.22 |
| *Armatimonadetes* (B) | Unidentified OS-L | 3.82 | 0.02 | 8.02 | 0.68 |
| *Aquificae* (B) | *Thermocrinis* | 3.66 | 7.57 | 6.84 | 2.25 |
| *Chlorobi* (B) | Unidentified *Chlorobi* | 2.21 | 0.65 | 4.49 | 3.88 |

[a]Sites A and B sediment samples compared to Site C sediment samples compared to Site E sediment sample.
[b]Phylum-level taxa for bacteria and class-level taxa for archaea.
[c]Contribution of OTU to overall dissimilarity between groups.
[d]Average abundance of OTU in each group.

**Figure 2.1** Visual representation of the similarity of samples based on their community composition, considered at the phylum level for bacteria and the class level for archaea. (a) Cluster tree calculated using Bray-Curtis dissimilarity, with jackknife values ≥80.0 % displayed. (b) Bar chart displaying the community composition of each sample, including the 15 most abundant taxa in all samples with remaining taxa included as "Others". "(A)" or "(B)" after taxon name in legend designates an archaeal or bacterial taxon, respectively.

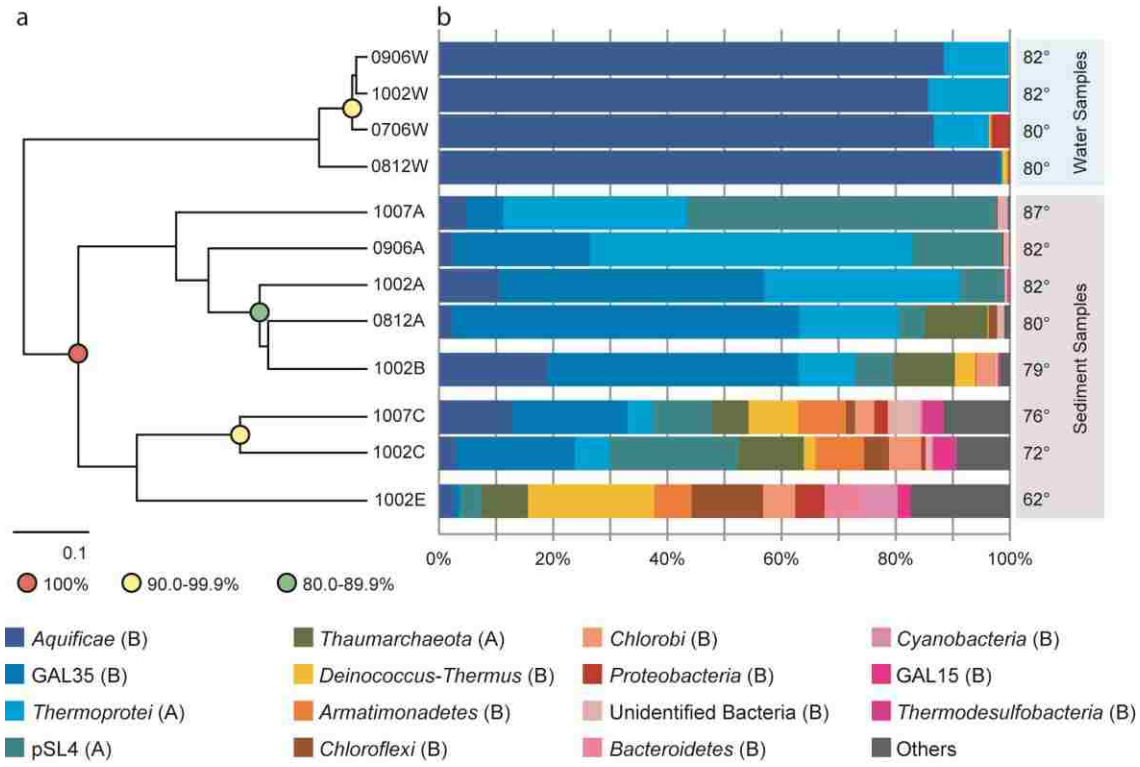


**Figure 2.2.** Visual representation of the similarity of samples based on their community composition, considered at the 97% OTU level. (a) Cluster tree calculated using Bray-Curtis dissimilarity, with jackknife values ≥80.0 % displayed. (b) PCoA constructed with Bray-Curtis dissimilarity. Principal coordinate 1, P1, and principal coordinate 2, P2, plotted against each other, with a total of 71.85% of variation explained.

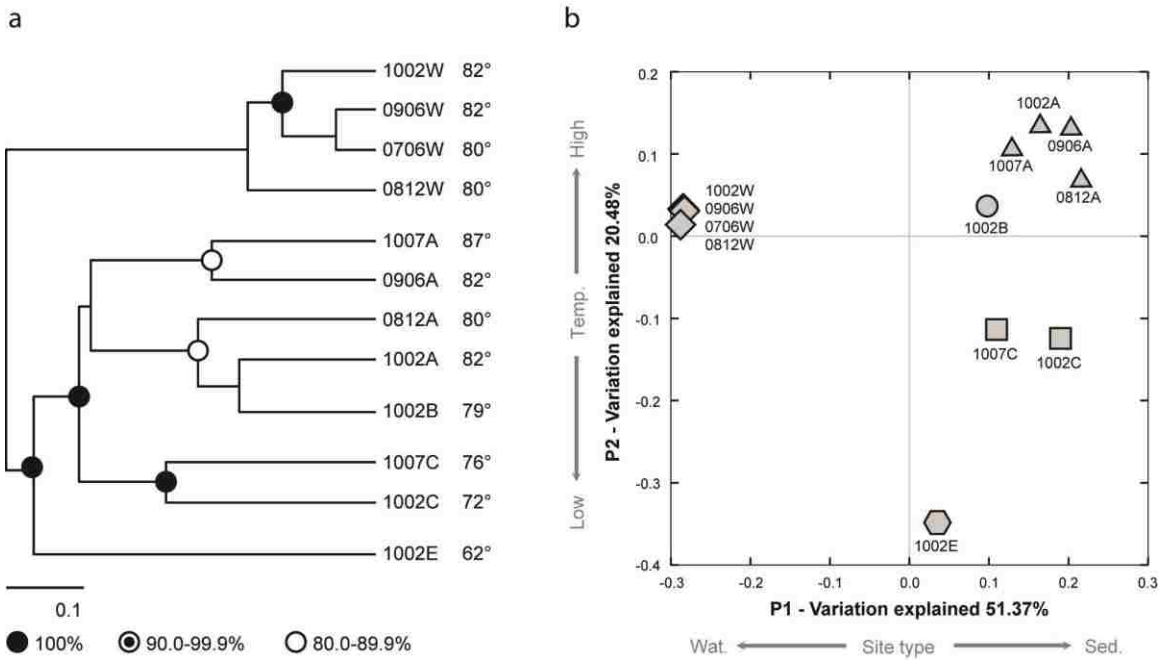


**Figure 2.3.** Comparison of alpha diversity measures of complete communities (both Archaea and Bacteria) in water samples to sediment samples of comparable temperature (82-80 °C). Calculations were performed with OTUs determined at five sequence identity levels and errors bars indicate standard error of the mean. Student's t-tests were performed for each pair of comparisons, water vs. sediment, at each OTU level to determine significance. ***, significant at α=0.001; **, significant at α=0.01; *, significant at α=0.05; ~, significant at α=0.1. (a) Richness measured in the water samples compared to that of sediment samples of comparable temperature. (b) Simpson's evenness measured in the water samples compared to that of the sediment samples.

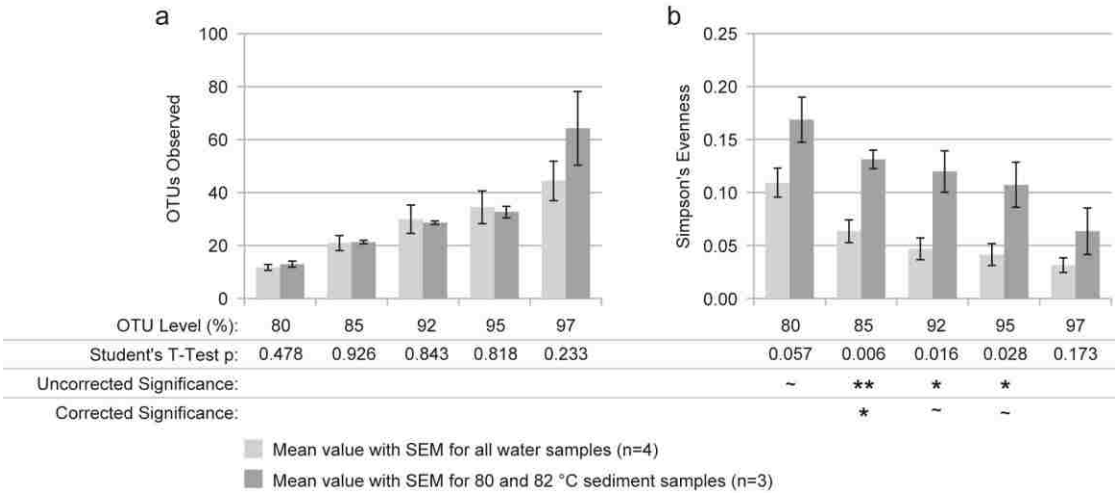
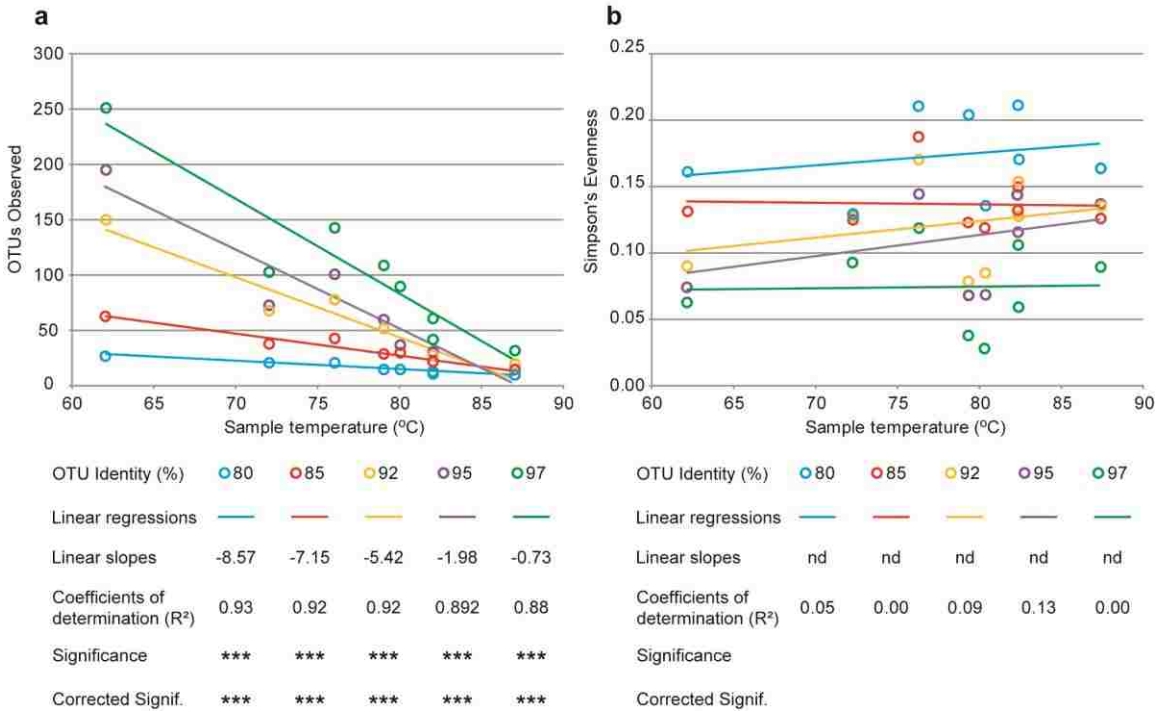


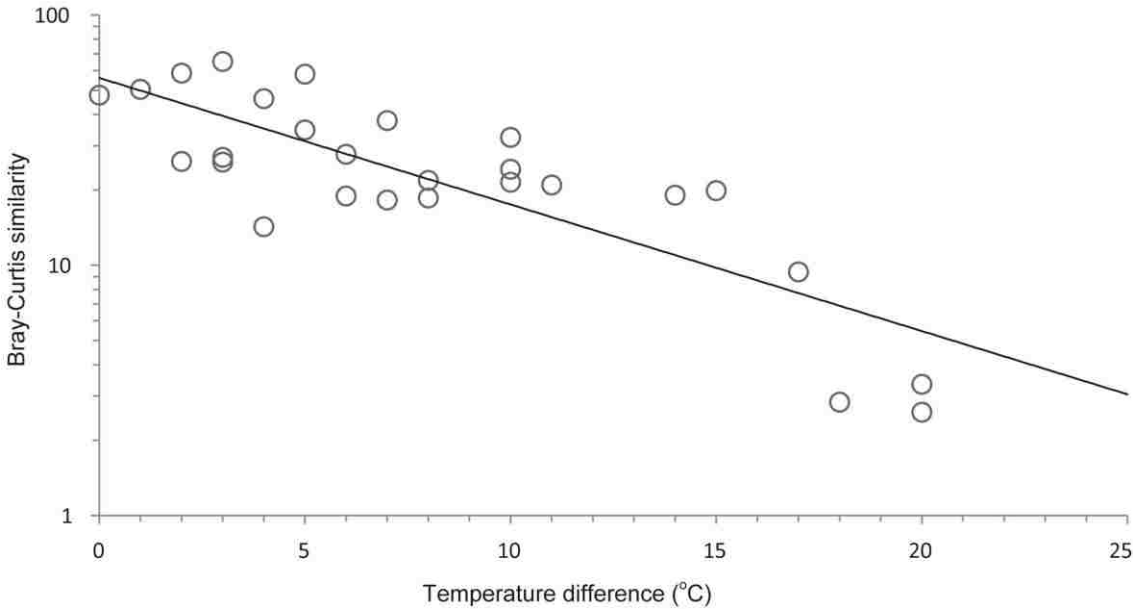
**Figure 2.4.** Alpha diversity measures of complete communities (combined Archaea and Bacteria) for all sediment samples, plotted against sample temperature, with regression lines for each OTU level. ***, significant at α=0.001; **, significant at α=0.01; *, significant at α=0.05; ~, significant at α=0.1. (a) Richness vs. sample temperature. (b) Simpson's evenness vs. sample temperature. nd, slope not significantly different from zero.




| OTU Identity (%) | 80 | 85 | 92 | 95 | 97 |
|---|---|---|---|---|---|
| Linear regressions | | | | | |
| Linear slopes | -8.57 | -7.15 | -5.42 | -1.98 | -0.73 |
| Coefficients of determination ($R^2$) | 0.93 | 0.92 | 0.92 | 0.892 | 0.88 |
| Significance | *** | *** | *** | *** | *** |
| Corrected Signif. | *** | *** | *** | *** | *** |

| OTU Identity (%) | 80 | 85 | 92 | 95 | 97 |
|---|---|---|---|---|---|
| Linear regressions | | | | | |
| Linear slopes | nd | nd | nd | nd | nd |
| Coefficients of determination ($R^2$) | 0.05 | 0.00 | 0.09 | 0.13 | 0.00 |
| Significance | | | | | |
| Corrected Signif. | | | | | |

**Figure 2.5.** Exponential decay of community similarity as the difference between sample temperatures increases. The linear regression between log-transformed Bray-Curtis similarity and the difference between sample temperatures was highly significant ($y = 56.021e^{-0.116x}$, $R^2 = 0.73$, $p < 0.001$).



# CHAPTER 3

# PYROSEQUENCING ANALYSIS OF HIGH-TEMPERATURE CELLULOLYTIC MICROBIAL CONSORTIA DEVELOPED IN GREAT BOILING SPRING AFTER *IN SITU* LIGNOCELLULOSE ENRICHMENT


## Abstract

To characterize high-temperature, cellulolytic microbial communities, two lignocellulosic substrates, ammonia fiber-explosion-treated corn stover and aspen shavings, were incubated at average temperatures of 77 and 85 °C in the sediment and water column of Great Boiling Spring, Nevada. Comparison of 109,941 quality-filtered 16S rRNA gene pyrosequences (pyrotags) from eight enrichments to 37,057 quality-filtered pyrotags from corresponding natural samples revealed distinct enriched communities dominated by known cellulolytic and hemicellulolytic organisms, *Thermotogae* and *Dictyoglomi*, possibly cellulolytic *Thermoprotei*, sugar-fermenting *Thermosphaera*, and hydrogenotrophic *Thermodesulfobacteria* and *Archaeoglobales*. Minor enriched populations included the candidate bacterial phylum OP9 and candidate archaeal groups C2 and DHVE3. Enrichment temperature was the major factor influencing community composition within enrichments, with a negative correlation between temperature and richness. Lignocellulosic substrate composition also contributed significantly to community structure. This study demonstrates the importance of known microbial groups in the natural degradation of lignocellulose at high temperatures and suggests that most of the diversity of cellulolytic thermophiles may be limited to lineages with representatives in pure culture.


## Introduction

Growing human populations and expanding industrialization have led to an increasing global demand for the finite supply of fossil fuels. Global petroleum consumption has increased by 12.9% in the last decade alone (BP 2012a) and the global demand for energy is projected to increase by 38.6% between 2010 and 2030 (BP 2012b), prompting growing interest in alternative fuel sources. Liquid biofuels, such as bioethanol, which are compatible with modern vehicles and the extant fuel delivery and supply infrastructure, are an appealing supplement to petroleum-based fuel supplies (Rubin 2008). Because biofuels are derived from renewable plant material, they represent an opportunity to move toward carbon neutrality and, in the process, reduce the scale of anthropogenic climate change. However, current bioethanol production methods, considered "first-generation" biofuel technology, are not without their drawbacks. Because first-generation biofuels extract fermentable sugars from plants traditionally utilized as food crops, their production directly competes with the supply of food for human populations. In the United States, 34 to 41% of the corn crop is diverted to ethanol production (USDA 2012), resulting in pressure on arable land availability and higher food prices (International Energy Agency 2008; Sims *et al.* 2010). The negative consequences of existing biofuel production techniques have stimulated interest in the development of so-called "second-generation" biofuel technologies, which derive fermentable sugars from dedicated crops or the abundant lignocellulosic waste produced by agriculture, forestry, and other industries.

The structural complexity and low aqueous solubility of lignocellulosic biomass creates a significant barrier to its use and the cost of ethanol production from lignocellulosic sources is currently prohibitive (Himmel *et al.* 2007). Existing production

methods involve chemical, thermal, and mechanical pretreatment of plant tissues to increase the availability of the structural carbohydrates for hydrolysis, which is carried out by cellulolytic microorganisms or their purified cellulolytic enzymes (cellulases). However, the high unit cost to produce these cellulases has limited their application to ethanol production, creating a demand for more cost-effective enzymes that would make second-generation biofuels an economically feasible alternative to fossil fuels and first-generation biofuels (Alizadeh *et al.* 2005).

Thermostable cellulases offer several potential benefits to mitigate the high costs of enzymatic saccharification of lignocellulose for biofuel production. Thermostable cellulases tend to have much greater activity at their optimal temperature than the cellulases obtained from mesophilic organisms, because each 10 °C increase in reaction temperature increases enzymatic rates two- to three-fold (Mozhaev 1993). Conducting reactions at higher temperature increases the solubility of substrates, increasing the yield of the end products (Mozhaev 1993) and reduces the viscosity of the reaction mixture, reducing water demand (Haki and Rakshit 2003). Additionally, thermostable enzymes are resistant to denaturation from other factors and highly stable for long-term storage, lengthening their shelf life and operational life during lignocellulose digestion (Mozhaev 1993; Viikari *et al.* 2007).

Plant material, including lignocellulose, is rarely abundant in high-temperature ecosystems. Nevertheless, several enzymes for cellulolysis have been isolated from thermophilic microorganisms (Bronnenmeier *et al.* 1995, 1995; Hreggvidsson *et al.* 1996; Zverlov *et al.* 1998; Bok *et al.* 1998; Bauer *et al.* 1999; Evans *et al.* 2000; Ando *et al.* 2002) and putative genes for many more have been identified (Te'o *et al.* 1995; Nelson *et*

*al.* 1999; Chhabra *et al.* 2002; Blumer-Schuette *et al.* 2008; Brumm *et al.* 2011; Frock *et al.* 2012). A family 5 endoglucanase with preference for cellulose substrates at 85 °C was cloned from the marine archaeon *Pyrococcus horikoshii* (Ando *et al.* 2002) and the closely related *P. furiosus* encodes a family 12 endoglucanase that has been shown to act upon several forms of highly crystalline cellulose at 95 °C (Bauer *et al.* 1999). Graham *et al*. (2011) extended the known upper temperature limit of cellulolysis by describing a multi-domain cellulase with maximal activity at 109 °C, encoded by an *Ignisphaera*-like archaeon. The only other thermophilic archaeon confirmed to degrade crystalline cellulose is *Desulfurococcus fermentans*, which grows optimally between 80 and 82 °C (Perevalova *et al.* 2005). Thermophilic bacteria, including several species of *Thermotogales* and *Clostridiales* are known to use cellulose, hemicellulose, and other β-glucans, but have lower optimal growth temperatures than the cellulolytic archaea (Blumer-Schuette *et al.* 2008). Notably, several species of *Caldicellulosiruptor*, with growth temperature optima between 70 and 78 °C, have been found to utilize cellulose with varying degrees of crystallinity (Blumer-Schuette *et al.* 2008; Hamilton-Brehm *et al.* 2010; Yang *et al.* 2010), including four capable of hydrolyzing Whatman no. 1 filter paper (Blumer-Schuette *et al.* 2010).

Most studies of cellulolysis in thermophiles have involved investigation of the saccharolytic activity of individual species, with little investigation into whole thermophilic communities acting upon lignocellulosic materials. The lignocellulolytic community studies that have been performed have focused on environments at or below 60 °C and have found enriched communities dominated by members of *Firmicutes*, *Actinobacteria*, *Proteobacteria*, *Bacteroidetes*, and *Chloroflexi* (Allgaier *et al.* 2010;

Sizova *et al.* 2011; Gladden *et al.* 2012). These studies are valuable for their characterization of moderately thermophilic communities acting upon a variety of lignocellulosic substrates, but the dominant phyla of high-temperature (>70 °C) environments tend to be completely distinct from those of lower-temperature systems (Barns *et al.* 1994; Reysenbach *et al.* 1994, 2000; Hugenholtz *et al.* 1998; Blank *et al.* 2002; Meyer-Dombard *et al.* 2005; Spear *et al.* 2005; Costa *et al.* 2009). It is therefore likely that cellulolytic consortia in high-temperature systems are different from the more moderately thermophilic communities already described.

Here, we use 16S rRNA gene pyrotag sequencing to examine *in-situ* changes in the microbial community of Great Boiling Spring (GBS), Nevada, USA, in response to enrichment with lignocellulose. GBS is a large, circumneutral hot spring located in the U.S. Great Basin that receives a regular influx of allochthonous plant material due to low rates of mineral precipitation (Anderson 1978) and has been found to harbor a rich assortment of novel microorganisms (Costa *et al.* 2009; Dodsworth, Hungate, and Hedlund 2011). We established eight *in situ* enrichments, each containing one of two different lignocellulosic substrates, in both the sediment and water column of the source pool and outflow channel of the spring. The enriched microbial communities were compared to those in natural sediment samples from the same sites in order to examine the effect of lignocellulosic enrichment on the microbial community of GBS and better understand lignocellulose-degrading organisms and communities.

## Methods

Sample incubation and collection

GBS is located at N40° 39.684' W119° 21.978'. Site 85, denoted for its average water temperature ~85 °C, is at the northwest side of the main spring pool (N40° 39.686' W119° 21.979'). Site 77, average water temperature ~77 °C, is ~11 m from Site 85 and ~0.5 m from the outflow (N40° 39.682', W119° 21.973'). Site 77 averaged 8.3 °C (range: 6.2 - 14.1 °C; standard deviation: 1.4 °C) cooler than Site 85 during the incubation period. Site 85 corresponds to Site A and Site 77 to Site C, as described by Hedlund *et al.* (2011).

Enrichment packets were prepared by enclosing 20 g of either aspen shavings, commercially available as pet litter (Kaytee Products, Chilton, WI), or AFEX-treated corn stover (kindly provided by Bruce Dale, Michigan State University). Each packet was constructed by sewing together two ~10 cm squares of nylon mesh with a 100 μm pore size (#NM0100P3 Pentair Industrial, Hanover Park, IL) with nylon thread.

Sediment incubations were placed approximately 1 cm below the sediment-water interface. Samples incubated in the water column were suspended ~10 cm below the air-water interface, enclosed in 20 x 12 x 5 cm polypropylene boxes punctured with ~100 holes, each with a 0.5 cm diameter, to maintain their position in the water column and allow for exchange of bulk spring water with the enrichment packets.

All packets were deployed on 29 August, 2009. Those at Site 77 were harvested 1 November, 2009 and those at Site 85 were harvested 29 November, 2009. Immediately after retrieval, the contents of each packet were divided into sterile, conical, polypropylene tubes using sterile forceps. The subsamples were then frozen on dry ice for transport and stored at -80 °C prior to analysis.

Non-incubated samples of aspen shavings and corn stover were prepared for DNA extraction and quantification by soaking ~5 g of each substrate in 50 ml conical tubes containing 40 ml 0.5x TE buffer (5 mM Tris, 0.5 mM EDTA, pH 8) at 80 °C for 4 hours. The buffer was decanted and the remaining wetted substrate stored at -80 °C.

Natural water and sediment sample collection

Briefly, sterile, 50 ml conical polypropylene tubes were used to collect the top ~1 cm of sediment at each site for natural sediment samples. Sediment samples were collected at Site 85 on three dates between June 2009 and July 2010 and Site 77 on two dates in February 2010 and July 2010. Additionally, bulk water samples were collected from the main spring pool near Site 85 by tangential flow filtration or 0.2 μm normal filtration on three dates between June 2006 and February 2010.

Environmental data collection

The water temperature and pH at each site was measured at the beginning and end of incubation with a handheld pH 5 meter (LaMotte, Chestertown, MD). Temperature data logger iButtons (DS1922T; Maxim, Sunnyvale, CA), set to record temperature every 2 hours, were sealed in 50 ml conical tubes and suspended in the spring water at the two incubation sites from September 15, 2009 until the end of incubation (Table 3.1, Figure B.12).

Substrate lignocellulose content analysis

All substrate content analyses were carried out by Dairy One, Inc. of Ithaca, New York.

Acid detergent fiber (ADF) content was determined by digesting 0.5 g samples in Ankom Technology FAD20CB acid detergent solution (2.93% w/v sulfuric acid and 2%

w/v cetyltrimethylammonium bromide; Ankom Technology, Macedon, NY) for 75 minutes in an Ankom A200 digestion unit. Samples were rinsed three times in boiling water for five minutes, and then soaked in acetone for three minutes before drying at 100 °C for two hours. Samples were weighed before and after digestion to determine ADF content.

Lignin and cellulose content were determined by digesting the ADF residue for three hours in an Ankom Daisy II incubator at ambient temperature, using 72% w/w sulfuric acid. After rinsing and drying, the remaining solid was considered the lignin portion of the original sample. The portion dissolved away by this treatment was considered cellulose.

Neutral detergent fiber (NDF) content was determined by digesting 0.5 g samples for 75 minutes in an Ankom A200 digestion unit with Ankom Technology FND20 neutral detergent solution (3.0% sodium lauryl sulfate, 1.86% EDTA disodium dehydrate, 0.68% sodium borate decahydrate, 0.46% anhydrous dibasic sodium phosphate, and 1.0% triethylene glycol), amended with 0.2% α-amylase and 1% w/v sodium sulfite. Samples were then rinsed twice for five minutes with a boiling α-amylase solution and once for five minutes with boiling water, then soaked in acetone for 3 minutes before drying at 100 °C for two hours. Samples were weighed before and after digestion to determine NDF content.

Hemicellulose content was calculated by subtracting the ADF content from the NDF portion.

Ash content was determined according to the Association of Official Analytical Chemists (AOAC) Official Method 942.05. Pre-weighed samples were held at 600 °C for

two hours to incinerate organic materials. The remaining mass after treatment was weighed to determine the ash portion of the original sample.

DNA extraction, PCR amplification and pyrosequencing

DNA extractions were performed using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH) as described previously (Dodsworth, Hungate, de la Torre, *et al.* 2011). Samples were divided into approximately 0.5 g portions for DNA extraction. DNA was extracted from each portion, pooled per sample, and precipitated with 70% ethanol. DNA was resuspended in sterile 0.5 x TE (5 mM Tris, 0.5 mM EDTA, pH 8) and quantified using a Nanodrop 1000 (Thermo Scientific, Waltham, MA, USA).

Extracted DNA from all samples was shipped on dry ice to the United States Department of Energy's Joint Genome Institute (JGI) for pyrosequencing. The V8 and a portion of the V7 hypervariable regions of the 16S rRNA gene were amplified using primers 926F454TitFNew (AAA CTY AAA KGA ATT GRC GG) and 1392R (ACG GGC GGT GTG TRC) and sequenced using 454 GS-FLX Titanium pyrosequencing (454 Life Sciences, Branford, CT, USA).

Pyrotag data filtering and preparation

146,162 pyrotags from eight natural samples and 158,427 pyrotags from eight enrichment samples were trimmed short of their first ambiguous base. Sequences shorter than 120 nt, including 5 nt multiplex ID barcodes and primers, were excluded from further processing. Quantitative Insights into Microbial Ecology (QIIME) software package v 1.5.0 (Caporaso *et al.* 2010) was used to denoise sequences and all pyrotags were truncated to a maximum length of 204 nt, based on minimum quality scores. OTUs at 97%, 95%, 92%, 85%, and 80% minimum identity levels were identified by QIIME,

using the uclust algorithm, and the most abundant representative sequence within each OTU was inspected for chimeras with ChimeraSlayer (Haas *et al.* 2011) against the NAST-aligned "gold" alignment file supplied with the ChimeraSlayer package using QIIME. Each OTU was BLASTed against the GreenGenes database (DeSantis *et al.* 2006) and assigned the most detailed taxonomy that matched ≥ 90% of the sequences within the OTU. Sequences that could not be assigned taxonomic identification were investigated using the NCBI BLAST database using the megablast algorithm (Zhang *et al.* 2000). Eukaryotic sequences and potential chimeras were omitted from further processing. The GAL35 lineage was removed from the OP1 phylum and maintained as its own phylum, consistent with previous versions of the GreenGenes phylogeny database. Sequences, taxonomies, and OTU definitions were imported into a MySQL Community Edition version 5.5.24 (Oracle Corporation 2012) database. In order to create representative natural sediment and water samples, averaged samples were created by randomly subsampling equal proportions of each natural sample collected at a site. Sample UW consisted of one third of each of the natural water samples collected. Sample U85 consisted of one third of each of the sediment samples collected at Site 85. Sample U77 consisted of one half of each of the sediment samples collected at Site 77. Aggregate samples were exported from the database as QIIME-compatible OTU tables for community analyses.

Community composition statistical analyses

QIIME was used to calculate OTUs observed (S) and Simpson's Index of Diversity (1-D) (Simpson 1949) at each OTU identity level measured. Simpson's index of evenness was calculated in Microsoft Excel 2010, based on S and 1-D.

A subsample of 4,842 pyrotags from each sample was identified by random sorting in MySQL to produce a pyrotag subset, with all samples rarefied to match the smallest sample (85AS). The rarefied data was exported and QIIME was used to calculate the Bray-Curtis dissimilarity scores (Bray and Curtis 1957) between each rarefied sample. The Bray-Curtis dissimilarities were used within QIIME to produce a sample cluster tree and principal coordinates (PCoA) graph. Permutational multivariate analysis of variance (PERMANOVA; also called non-parametric MANOVA, NPMANOVA) was used to calculate the statistical significance of community composition differences between groups of samples (Anderson 2001). PERMANOVA statistics were calculated for sample groups identified by the cluster tree and for groups separated by experimental conditions using PAST v. 2.11 (Hammer *et al.* 2001), based on Bray-Curtis dissimilarity with 9,999 permutations. For groups that were found to be significantly different, PAST was used to perform SIMPER analysis to identify the most discriminatory OTUs between the sample groups.

Enrichment calculation

Genus fold enrichment was calculated in MySQL. All samples were normalized by multiplying the total hits per genus in the sample by a sample-specific multiplier in order to increase each sample to match the largest sample (85CW: 21,349 pyrotags). One hit was then added to each genus in every sample in order to eliminate divide-by-zero errors. Adjusted genus hits in enrichment samples were then divided by the adjusted genus hits in the corresponding natural sediment sample in order to calculate approximate fold enrichment. This technique slightly underestimated actual enrichment, especially for low-abundance genera. The base-10 logarithm of the fold enrichment ratios for highly-

abundant genera and specific OTUs of interest were exported from the database. The exported file was imported into the R statistical package (R. Development Core Team 2011) which was used to produce a heatmap of enrichment.

## Results and Discussion

Lignocellulose enrichments in Great Boiling Spring

Because of its size, topography, and low flow rate compared to the source pool size, GBS contains temperature zones that differ by more than 20 °C (Costa *et al.* 2009). These various zones harbor distinct microbial communities, with an inverse, linear relationship between community richness and temperature. Additionally, distinct microbial communities have been observed between the water and sediment of GBS (Dodsworth, Hungate, and Hedlund 2011). For this study, two sites along the perimeter of GBS were selected for lignocellulosic substrate incubation and sample collection (Table 3.1, Figure B.11). Long-term temperature loggers were deployed during incubation to record the temperature at each incubation site (Figure B.12). The hotter site, designated "Site 85," is close to the geothermal source and had an average temperature of ~85 °C, approximately equal to the source water. The cooler site, designated "Site 77," was near the spring's outflow and, with an average temperature of ~77 °C, is the coolest site in the spring that is normally above the photosynthetic temperature limit of 73 - 75 °C (Brock 1967).

Two enrichment packets of aspen shavings (samples designated with "A") and two packets of ammonia fiber explosion (AFEX)-treated corn stover (samples designated with "C") were incubated at each site. One packet of each substrate was buried in the sediment (samples designated with "S") and one was suspended in the water column

(samples designated with "W") at each site. Each of the eight enrichment samples is indicated here by the incubation site, substrate type, and incubation location, e.g. 85CS is corn stover incubated in the spring sediment at Site 85 (85CS = Site 85, Corn stover, Sediment).

To confirm growth on the lignocellulosic substrates, we extracted and quantified the DNA in non-incubated cellulosic material and in incubated samples. Post-incubation DNA yields were significantly higher compared to non-incubated samples for both corn stover (2.2. to 5.7-fold increase) and aspen (12.7 to 22.3-fold increase) substrates (Table A.5), indicating an increase in biomass and enrichment on the cellulosic materials *in situ*.

To compare the enrichment communities to the natural communities of GBS, we produced average samples by pooling and subsampling several of the natural sediment and water samples. Samples U85 and U77 represent the natural sediment communities at Site 85 and Site 77, respectively. Sample UW represents the natural community of the bulk water of GBS. This *in-silico* aggregation of samples created generalized natural communities for comparison to the enriched communities and minimized the effects of short-term community fluctuations due to temperature variations in GBS.

Effects of incubation on composition of lignocellulosic materials

The material used in the enrichments was composed mainly of fiber (cellulose, hemicellulose, and lignin), with smaller amounts of ash, protein, and bioavailable sugars (Table A.6). Lignocellulosic material was visibly recognizable after incubating for 64 days and 92 days at Site 77 and Site 85, respectively, with only slight physical modification occurring as a result of the incubation. Samples buried in the sediment were black in color, likely due in part to precipitation of metal sulfides on the surface of the

cellulosic substrates under anaerobic conditions. Except for the lack of dark coloration, the suspended packets were visually similar to the packets buried in the sediment after incubation, including a compact structure, suggesting an anoxic core. Though maintained in the normally oxic conditions of the bulk water of GBS (Costa *et al.* 2009; Miller-Coleman *et al.* 2012), a primarily anaerobic community developed within the suspended enrichment packets (described below). This was likely due to microbial respiration exceeding the diffusion of oxygenated water into the compact material in the packets. The presence of an enriched anaerobic community in the packets suspended in the water, very similar to that of the enrichments in the sediment, led us to consider all enrichment samples to be analogous to the sediment sample at the same site.

The change in ash content (mineral content) between incubated and control substrates was used as a proxy for organic matter consumption, with a greater proportion of ash corresponding to greater consumption of organic material. The ash content was at least two-fold higher in all incubated samples than in the controls, indicating that a portion of the original organic material was consumed (2.26- to 6.56-fold increase in corn stover samples; 5.24- to 15.3-fold increase in aspen samples; Figure 3.1a). However, based on the observations of mineral precipitation described above, the ash measurements may slightly overestimate the true organic matter consumption. No clear trend in the ash content of samples incubated in the sediment versus the water column was apparent, supporting the assertion that the change in the percent ash was due mainly to microbial consumption of organic matter and not to sediment accumulation or mineral precipitation. The increase in ash content was greater in the Site 85 samples, consistent with the longer incubation time and higher DNA content of samples incubated at that site.

The major component of both the aspen shavings and corn stover was cellulose, which ranged from ~40 to 70% of the total mass of the samples (Table A.6). The cellulose fraction increased in all incubated samples relative to the controls, with the fractional change being significantly higher in the corn stover samples than in the aspen samples. However, the total *mass* of cellulose likely decreased from the original amount due to consumption of the organic fraction as indicated by the increase in ash content. Hemicellulose was preferentially degraded over cellulose in all samples, but especially so in the corn stover samples, where the cellulose to hemicellulose ratio increased from 1.20 in the non-incubated substrate to 5.38 to 8.83 in the samples incubated in the spring. The aspen material was more cellulose-rich, with an initial cellulose to hemicellulose ratio of 3.84 and showed a more moderate increase to 4.61 to 6.90 in incubated samples (Figure 3.1b).

The reasons why overall degradation of organic material was higher in corn stover samples than in aspen samples and why hemicellulose was more preferentially degraded in the corn stover samples remain unclear. Rates of degradation were likely influenced by the initial composition of the substrate materials and the AFEX pretreatment applied to the corn stover. AFEX has been shown to alter the lignin structure, solubilize lignin, decrystallize cellulose, and increase the surface area of lignocellulosic material, optimizing conditions for enzymatic hydrolysis of cellulose and hemicellulose (Mosier *et al.* 2005).

Lignin decreased in almost all incubations, although percent lignin slightly increased in the 77AW and 85AW treatments, and generally mirrored hemicellulose depletion (Table A.6). Because lignin is covalently bonded to hemicellulose, the

degradation of hemicellulose alone, as opposed to direct biological metabolism of lignin, could result in the solubilization and loss of lignin from the sample. In general, greater changes were observed in the corn stover incubations for most of the components analyzed, suggesting that this substrate is more labile than the aspen shavings. No correlations between utilization of a particular component and either temperature or incubation location were observed, although more dramatic changes often occurred in the sediment treatments and the 85 °C incubations.

Community diversity changes in response to lignocellulose enrichment

In order to assess the differences between the natural sediment communities and the cellulolytic communities, we analyzed the operational taxonomic units (OTUs) found in each sample. OTUs were defined based on ≥97% pyrotag sequence identity, which approximates the species level, although use of the V8 region of the 16S rRNA gene likely resulted in underestimations of the true richness (Youssef *et al.* 2009). 601 OTUs were identified among all enrichment and natural sediment samples. The natural sediment communities were comprised of 228 OTUs, while 460 OTUs were found in enrichment samples. Only 87 OTUs were shared between enrichment samples and natural sediment communities and only 3 OTUs (#C236, #C600, and #C603) were found in greater than 5% relative abundance in either natural sediment community and any enrichment sample (Table A.7), demonstrating enrichment of distinct microbial communities on the lignocellulosic substrates.

All samples from Site 85 had a lower species richness than Site 77 samples, consistent with earlier findings of a temperature-driven richness gradient in GBS sediments. Observed richness was similar in all samples at Site 85, including the

aggregate natural sediment sample, U85. Three of the Site 77 enrichment samples had fewer OTUs than the natural sediment sample at that site, U77, and the fourth had approximately the same richness as U77 (Figure 3.2a). Similar results were observed from genus to phylum level and no significant pattern in richness was observed when comparing lignocellulosic substrate or sample incubation location (Figure B.13a). No significant correlation was found between sequencing depth and species richness ($R^2$ = 0.0999; p = 0.344), indicating that the observed differences in species richness represent actual differences in the samples, rather than artifacts of incomplete sampling.

The OTU count provides an estimation of species richness in a sample, but because high-abundance and low-abundance OTUs are counted the same, they do not provide a complete picture of diversity, which includes both the richness and evenness of a community. In order to assess community evenness, we calculated Simpson's index of evenness (Simpson 1949; Krebs 1998). No trends pertaining to community evenness were observed between enrichments and the natural sediment communities and community evenness was not significantly altered by enrichment (Site 85, p = 0.499; Site 77, p = 0.260; Figure 3.2b). Enrichments from Site 77 tended to be less even at all taxonomic levels than enrichments from Site 85, although the evenness of enrichment communities at Site 85 and Site 77 were not significantly different (p = 0.188; Figure B.13b).

Comparison and clustering of sample communities

Hierarchical clustering and principal coordinates analysis (PCoA) of the enrichment and natural sediment samples revealed the enrichment samples to have community compositions distinct from the natural sediment communities, with the

exception of 77CS (Figure 3.3). The most significant variable influencing community composition was whether the sample was natural or enriched with lignocellulose. Further segregation of samples was influenced by average incubation temperature, followed by the type of lignocellulosic material in the sample. Incubation of samples in the water column or sediment did not contribute significantly to any differences in community composition.

In order to test the significance of the variables influencing community composition, we performed PERMANOVA tests on all nodes in the sample cluster tree (Figure 3.3a) that joined three or more samples and on groups defined by experimental conditions (Table A.8). Tree node 2, which separated the enrichment samples from the natural sediment samples and 77CS, represented a highly significant division ($p = 0.0084$). The significance of the presence of lignocellulosic material was confirmed by a comparison of all enrichment samples (including 77CS) to the natural sediment samples ($p = 0.0199$). Incubation site/temperature was also confirmed as a significant factor in community composition, whether considering all enrichment and natural sediment samples at each site ($p = 0.0152$) or only enrichment samples (tree node 3; $p = 0.0275$). Neither the type of lignocellulosic material included in the enrichment packets, nor whether the packets were suspended in the bulk water or buried in the sediment of the spring, were significant ($p = 0.1765$, 0.8306, respectively).

77CS was anomalous in that it clustered with the natural Site 77 sediment sample, U77, rather than the other enrichment samples. 77CS was the least dissimilar sample to U77, with a Bray-Curtis dissimilarity score of 0.508. However, 77CS was slightly more similar to 77CW than it was to U77, with a score of 0.462 between 77CW and 77CS.

Other water/sediment sample pairs had average scores of 0.217, however, indicating that 77CS and 77CW are substantially more dissimilar to each other than the other water/sediment enrichment pairs. Despite its similarity to the natural samples based upon overall community composition, 77CS did show strong evidence of enrichment. Notably, *Thermotoga*, which was highly enriched in all samples, was the most abundant genus in 77CS. The anomalous nature of the 77CS sample may be explained by less enrichment of the microbial community in this sample during the incubation period or by penetration of the packet by spring sediment during incubation or collection of the sample. However, the nutritional analysis of the incubated material from that sample suggested strong cellulolytic activity and the ash content was not anomalously high.

Specific microbial taxa enriched on lignocellulosic substrates

Similarity percentage (SIMPER) analysis was used to identify the organisms most responsible for the differences between the communities identified by PERMANOVA as significantly different (Tables A.9-A.12). Examination of OTUs identified by SIMPER to be the most discriminating between enrichments and natural sediment communities (Table 3.2) revealed that lignocellulose amendment led to significant enrichment of the known cellulolytic and hemicellulolytic organisms *Thermotoga* (#C529, #C782) and *Dictyoglomus* (#C692), a potentially heterotrophic and hydrogenotrophic member of the *Archaeoglobaceae* (#C903) and an unidentified *Ignisphaera*-like archaeon (#C359). Conversely, there was a significant decrease in the relative populations of OTUs that dominated the natural sediment communities, such as candidate groups GAL35 (OTU #C603), "*Aigarchaeota*" (#C056, #C487; Nunoura *et al.* 2011), and NAG1 (#C136; Kozubal *et al.* 2012) and *Aeropyrum* (#C199).

Calculation of the fold enrichment of the genera identified within the samples (Figure 3.4) reinforced the distinctiveness of the communities formed within the enrichment packets. The most abundant genera in the enrichment samples (*Thermotoga*, *Ignisphaera*-like *Desulfurococcaceae*, *Archaeoglobus*-like *Archaeoglobaceae*, *Dictyoglomus*, and *Thermofilum*) were highly enriched over the natural samples. The genera that dominated the natural sediment communities (*Aeropyrum* and members of the uncultivated lineages "*Aigarchaeota*", GAL35, and NAG1) showed significantly lower relative abundance in the enrichment sample populations.

The natural sediment communities of GBS consisted primarily of uncultured organisms. Only 28.3% of the community in U85 and 32.0% of the community in U77 were comprised of organisms belonging to described families. The enrichment communities, however, averaged 77.2% (minimum 77CS, 58.5%; maximum 85CW, 97.2%; standard deviation 12.8%) identified at the family level. The predominance of families with cultivated representatives in the lignocellulose-enriched communities may be due to the relative ease with which organisms with heterotrophic metabolisms are cultivated *in vitro*. It may also reflect the history of investigation into organisms with potential cellulolytic activity.

Enrichment of highly carbohydrate-active *Thermotoga*

The difference in the dominant OTUs between natural and lignocellulose-enriched communities was reflected at broader taxonomic levels. Considering the relative abundance of the bacterial phyla and archaeal classes that comprised each community (Figure B.14), *Thermotogae* represented at least 25% of each enrichment sample from Site 77 and three of the four enrichment samples from Site 85, but less than 1% of either

natural sediment sample, U85 or U77. Highly enriched *Thermotogae* consisted of two OTUs, #C529 and #C782. OTU #C529 was 100% identical to *Thermotoga petrophila* and *T. naphthophila* and 99% identical to *T. maritima* and *T. neapolitana* over the pyrotag length. This OTU was highly enriched in all enrichment samples except 77AS and 77AW, showing 179- to 7480-fold enrichment over natural sediment communities and a relative abundance that ranged from 14.1% to 35.1% of the total enriched communities. The Site 77 aspen enrichment communities, where OTU#529 was not abundant, had a high representation of OTU #C782, which shared 99% sequence identity with *T. thermarum* and 98% sequence identity with *T. hypogea*. OTU #C782 represented 22.0% to 33.6% of the Site 77 aspen enrichment communities and displayed a 101- to 154-fold increase over U77. OTU #C782 also represented 22.7% of the 77CW sample, but less than 4% of any other community.

No *Thermotoga* species has been shown to grow on crystalline cellulose, but some are known to grow on amorphous cellulose (Windberger *et al.* 1989; Takahata *et al.* 2001) and a variety of other carbohydrate polymers, including other β-linked glucans, α-linked glucans, hemicellulose, and pectin (Blumer-Schuette *et al.* 2008). The four cultivated *Thermotoga* species that are similar to OTU #C529 share ≥99% identity in their full-length 16S rRNA sequences and 77 to 83% of the protein-encoding open reading frames in their respective genomes, but have slightly different carbohydrate metabolisms (Frock *et al.* 2012). The best characterized of these is *T. maritima*, which has nearly seven percent of the predicted coding sequences in its genome (AE000512) dedicated to the catabolism and uptake of sugars and sugar polymers, including genes for two endoglucanases and several enzymes in the xylan degradation pathway (Nelson *et al.*

1999). Carbohydrate-active enzymes, including two cellulases, laminarase, xylanase, two possible β-D-xylosidases, α-D-glucuronidase, and α-L-arabinosidase have been purified from *T. maritima* (Bronnenmeier *et al.* 1995). Many of these genes have also been identified in the genomes of *T. petrophila* (CP000702; Zhaxybayeva *et al.* 2009), *T. naphthophila* (CP001839), *T. neapolitana* (CP000916), and *T. thermarum* (CP002351). Other *Thermotoga* enzymes have been isolated and expressed recombinantly in *E. coli*, including pectate lyase (Kluskens *et al.* 2003), exopolygalacturonase (Kluskens *et al.* 2005), and α-L-arabinofuranosidase (Xue *et al.* 2006), though these genes may not all be expressed by *Thermotoga* and expression is dependent on substrate availability (Chhabra *et al.* 2002, 2003). In addition to cellulose and hemicellulose depolymerization activity, *Thermotoga* species also express enzymes capable of hydrolyzing the disaccharide cellobiose (Liebl 2001) and the trisaccharide cellotriose (Bronnenmeier *et al.* 1995). They also ferment simple sugars, releasing $H_2$, $CO_2$, lactate, and acetate (Huber *et al.* 1986), suggesting that the *Thermotogae* were potentially involved in every step of decomposition of our lignocellulosic substrates. The biology of *Thermotoga* implicates them in primary hemicellulolysis, but their inability to grow on crystalline cellulose suggests that other groups in the enrichments might be more important to the key step of primary cellulolysis.

Enrichment of heterotrophic *Thermoprotei*

Although the archaeal class *Thermoprotei* was abundant in both the enrichment samples (32.2 to 61.1% relative abundance) and natural sediment sample (21.8% relative abundance) at Site 85, the *Thermoprotei* present in the enrichment samples were distinct from those found in U85. The *Thermoprotei* in the natural sediment sample U85 were

primarily affiliated with a single OTU, #C199, which shared 98% sequence identity over the pyrotag length with *Aeropyrum pernix* (NC_000854.2), a marine hyperthermophile that grows aerobically on a variety of proteinaceous compounds. This OTU accounted for 20.3% of the natural sediment community at Site 85, but no more than 4% of the pyrotags present in any enrichment sample.

The *Thermoprotei* present in the enrichment samples at Site 85 included members of both the *Desulfurococcaceae* and *Thermofilaceae.* Two OTUs comprised most of the *Desulfurococcaceae*, #C011 and #C359. #C011 was 3.9 to 9.8% of the microbial community in the Site 85 enrichment samples, where it showed a 821- to 2,080-fold enrichment over U85. OTU #C359 represented 15.6% to 23.5% of the community of the Site 85 enrichment samples, a 3,330- to 5,270-fold enrichment. #C359 shared 100% pyrotag sequence identity with an organism ("*Ignisphaera*-like archaeon", JF509453) previously identified within a lignocellulolytic consortium enriched from a ~94 °C geothermal spring ~170 meters north of GBS and within the Great Boiling Spring system (Graham *et al.* 2011). The consortium was the first documented to degrade crystalline cellulose at temperatures exceeding 90 °C, and the *Ignisphaera*-like archaeon was identified as the microorganism responsible for the strong cellulolytic ability of the consortium, and encoded a multi-domain cellulase with maximal activity at 109 °C. The enrichment of these two OTUs on lignocellulosic material at Site 85 and the fact that the two OTUs shared high identity to an organism with demonstrated strong cellulolytic ability suggests that these microorganisms were likely participating directly in the primary decomposition of the lignocellulosic materials.

Members of *Thermofilaceae* at Site 85 included OTU #C867, which was 99% identical to *Thermofilum pendens* (CP000505) over the pyrotag length. OTU #C867 constituted less than 4.0% of the communities present in the aspen samples (85AS and 85AW), but represented 12.4% to 18.9% of the community of the corn stover samples (85CS and 85CW). *T. pendens* is a strictly anaerobic hyperthermophile that grows chemoorganotrophically using sulfur-based anaerobic respiration. The genome of *T. pendens* contains a large number of highly-expressed ABC transporters responsible for carbohydrate uptake, including a transporter with high similarity to a characterized cellobiose transporter, and genes dedicated to carbohydrate metabolism, including a secreted family of glycosyl hydrolases with weak similarity to known cellulases (Anderson *et al.* 2008). The organism represented by OTU #C867 may be responsible for some of the cellulolytic activity performed by the microbial communities in 85CS and 85CW, but based on the pronounced ability of *T. pendens* for carbohydrate uptake, it is more likely that this organism was enriched due to the release of saccharides by the cellulolytic and hemicellulolytic activities of other community members.

Enrichment of hemicellulolytic and possibly cellulolytic *Dictyoglomi*

Substrate-preferential enrichment was also observed for an OTU assigned to the phylum *Dictyoglomi*, the members of which are known to produce a variety of thermostable enzymes with significant biotechnological applications, including xylanases, amylases, and mannases (Patel 2010). The 16S rRNA genes of the only two recognized species of *Dictyoglomus*, *D. turgidum* (CP001251) and *D. thermophilum* (CP001146), share 99% sequence identity. The genus *Dictyoglomus* was represented in our samples almost entirely by a single OTU, #C692, which also shared 99% sequence

identity to both *D. turgidum* and *D. thermophilum* over the length of the pyrotag. *D. turgidum* and *D. thermophilum* are strictly anaerobic and thermophilic chemoorganotrophs that can ferment a variety of carbohydrates. #C692 represented less than 0.4% of the U77 natural sediment sample, but was abundant in the enrichment samples at Site 77, except the anomalous 77CS. OTU #C692 showed greater enrichment in the Site 77 aspen samples, 77AW and 77AS, where it represented 18.3% to 20.6% of the community, than in the corn stover samples, where it represented only 9.7% of the 77CW community and 1.0% of the 77CS community. OTU #C692 was not detected in U85 or the corn stover enrichments at Site 85, but did represent ~2% of the aspen samples at Site 85, further illustrating this organism's preference for the aspen material. Although the representative sequence of this OTU shared equal identity to both species of *Dictyoglomus*, its appearance in Site 85 samples is more consistent with the growth temperature range for *D. turgidum* (86 °C maximum, 72 °C optimum ;Svetlichnii and Svetlichnaya 1988), than with that of *D. thermophilum* (80 °C maximum, 73 - 78 °C optimum ;Saiki *et al.* 1985). *D. thermophilum* has not been shown to express cellulases, but does produce thermostable xylanases and can grow on a variety of fermentable carbohydrates, including several monohexoses and monopentoses, as well as cellobiose (Saiki *et al.* 1985). However, *D. turgidum* has been shown to utilize carboxymethylcellulose and microcrystalline cellulose as carbon sources (Svetlichnii and Svetlichnaya 1988), and the genes for 54 carbohydrate-active enzymes were annotated in its genome (Brumm *et al.* 2011).

Enrichment of other taxa

Also enriched were taxa with no or minimal known saccharolytic capability, including two OTUs within *Thermodesulfobacteria* (#C240 and #C707). Characterized *Thermodesulfobacteria* are known to use glycolysis intermediates and fermentation products such as pyruvate, lactate, and $H_2$ as electron donors in the reduction of sulfate (Zeikus *et al.* 1983). The organisms represented by these two OTUs were likely enriched by the metabolic products of other community members acting directly upon the enrichment substrates. OTU #903, with ≥98% sequence identity to several members of the *Archaeoglobaceae* (AE000782, AJ299218, FJ216404, NR_028166, CP002588, CP001899, NR_041788, CP001857), was also enriched in all samples, showing an average 5.79-fold enrichment at Site 77 and representing an average 7.30% of the enriched community at Site 85, where it did not appear in the natural sediment sample. Characterized members of *Archaeoglobaceae* are known to use glucose as well as fermentation products lactate and formate in the reduction of oxidized sulfur compounds (Schönheit and Schäfer 1995), suggesting this OTU may have been acting at multiple steps in the decomposition of lignocellulosic material. Also, since all known *Thermodesulfobacteria* and *Archaeoglobaceae* consume $H_2$, these organisms might play a role in consortial lignocellulose fermentation by keeping concentrations of $H_2$ low, reducing the inhibitory effects of $H_2$ on fermentative processes.

Several members of candidate taxa were also enriched by the lignocellulosic substrates. Notably, OTU #C758 represented 7.1 to 9.5% of samples 77AS, 77AW, and 77CW, where it showed a 51.4- to 69.4-fold increase over its representation in U77. This OTU is belonged to the candidate bacterial phylum OP9. OTU #C216, belonging to the candidate order pGrfC26 within the candidate archaeal class C2, was enriched in all

aspen samples (56.8- to 235-fold enrichment) but did not appear in most corn samples. OTU #C896, identified as a member of archaeal candidate phylum DHVE3, showed a 12.8- to 16.3-fold enrichment in the aspen samples at Site 77 but did not appear in any other enrichments. Because these candidate taxa have not yet been described, their roles in the enriched community remain unclear, but their enrichment suggests that they are involved in consortial lignocellulose degradation and may represent unexplored sources of novel carbohydrate-active enzymes.

## Conclusions

Lignocellulose incubations in GBS stimulated the formation of distinct communities dominated by organisms found only in low abundances in the natural environment. The closest characterized relatives of the dominant enriched OTUs are organisms known to be capable of diverse metabolisms that together engage in consortial lignocellulose digestion, including cellulose and hemicellulose degraders, sugar fermenters, and hydrogenotrophs. Enriched communities were different at Site 85 and Site 77, primarily due to higher representation of the archaeal classes *Thermoprotei* and *Archaeoglobi* at the higher-temperature site. In addition, species richness was inversely correlated with incubation temperature, which is consistent with findings for natural sediments within GBS and a strong role of high temperature in limiting microbial diversity. *Thermotoga* were highly enriched in all samples at both sites, consistent with the known utilization of a variety of carbohydrates by members of the genus. Two different species of *Thermotoga* were enriched in different samples and demonstrated some apparent mutual exclusion. Several taxa showed preference for a specific lignocellulosic material; *Thermosphaera* and *Thermofilum* were more highly enriched on

corn stover, while *Dictyoglomus*, *Thermodesulfobacteria*, and the archaeal group C2 showed marked preference for the aspen substrate.

Aspen and corn stover were consumed differently, with a higher ratio of hemicellulose degradation and possibly higher overall consumption of organic material in the corn stover samples than in the aspen samples. The specific reasons for this disparity are unclear, but the AFEX pretreatment and the higher initial ratio of hemicellulose to cellulose in the pretreated corn stover than in the aspen may have contributed to the differential degradation of each portion of the materials. Additional investigation, using a broader variety of lignocellulosic substrates, is necessary to fully elucidate the various factors leading to preferential decomposition of different components of lignocellulose by thermophilic communities.

The enzymatic conversion of lignocellulosic material to industrially useful materials such as ethanol and $H_2$ involves several steps, carried out by many enzymes and typically performed by several different organisms. This study characterized the thermophilic, cellulolytic and carbohydrate-active communities that form under several lignocellulose enrichment conditions in GBS and demonstrated the differential response of communities to different lignocellulose sources and incubation temperatures. Such factors are important for the optimization of industrial biofuel production processes. The differential decomposition of lignocellulosic stocks of varying cellulose and hemicellulose compositions and the temperature-specific formation of syntrophic communities comprised of mutually compatible organisms should inform the ongoing pursuit of systems for production of second-generation biofuels.

**Table 3.1.** Sample incubation conditions and pyrotag yields

| Sample Name | Avg. Temp. (°C)[a] | Site | Enrichment | Incubation Location | $N_i$[c] | $N_f$[d] |
|---|---|---|---|---|---|---|
| UW | 81 | Water | None | na[b] | na[b] | 11,233 |
| U77 | 74 | 77 | None | na[b] | na[b] | 11,308 |
| U85 | 83 | 85 | None | na[b] | na[b] | 14,516 |
| 77AS | 77 | 77 | Aspen | Sediment | 29,188 | 21,006 |
| 77AW | 77 | 77 | Aspen | Water | 25,528 | 18,924 |
| 77CS | 77 | 77 | Corn Stover | Sediment | 17,266 | 12,128 |
| 77CW | 77 | 77 | Corn Stover | Water | 14,527 | 10,565 |
| 85AS | 85 | 85 | Aspen | Sediment | 7,094 | 4,842 |
| 85AW | 85 | 85 | Aspen | Water | 12,167 | 8,259 |
| 85CS | 85 | 85 | Corn Stover | Sediment | 20,416 | 12,868 |
| 85CW | 85 | 85 | Corn Stover | Water | 32,241 | 21,349 |

[a] Average temperature of natural samples or as recorded during incubation (Figure B.11)

[b] na, not applicable

[c] Number of pyrotags generated by pyrosequencing

[d] Number of quality-filtered pyrotags used in analysis

**Table 3.2.** Significant OTUs discriminating between enrichment and natural samples

| OTU | Identity | Δ[a] | Contrib. (%)[b] | Natural Samples (%)[c] | Enrichment Samples (%)[d] |
|---|---|---|---|---|---|
| C529 | *Thermotoga* sp. | + | 10.76 | 0.04 | 19.25 |
| C603 | GAL35 | - | 10.36 | 21.89 | 3.47 |
| C056 | "*Aigarchaeota*" | - | 9.18 | 16.67 | 0.28 |
| C199 | *Aeropyrum* sp. | - | 6.22 | 11.61 | 0.51 |
| C359 | *Ignisphaera*-like | + | 6.08 | 0.00 | 10.84 |
| C782 | *Thermotoga* sp. | + | 5.81 | 0.10 | 10.41 |
| C136 | NAG1 | - | 3.97 | 7.10 | 0.06 |
| C903 | *Archaeoglobus* sp. | + | 3.95 | 0.65 | 7.70 |
| C692 | *Dictyoglomus* sp. | + | 3.66 | 0.13 | 6.61 |
| C487 | "*Aigarchaeota*" | - | 3.62 | 6.46 | 0.42 |

[a] Difference between natural sediment and enrichment populations. +, OTU has greater representation in enrichment samples than natural samples. -, OTU has lower representation in enrichment samples than natural samples.

[b] Percent contribution to community composition difference

[c] Average percent representation in natural sediment communities

[d] Average percent representation in enrichment sample communities

**Figure 3.1.** Composition of non-incubated and incubated lignocellulosic substrates. (a) Ash content. (b) Cellulose to hemicellulose ratio.

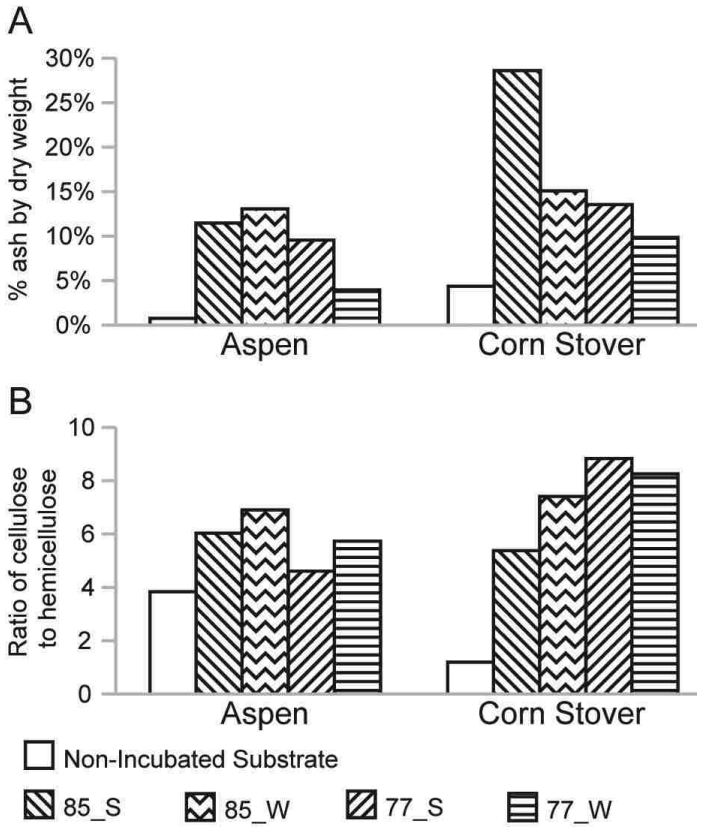


**Figure 3.2.** Alpha diversity measurements, comparing natural sediment and enrichment samples. (a) OTUs (≥97% identity) observed in each sample. (b) Simpson's index of evenness in each sample.

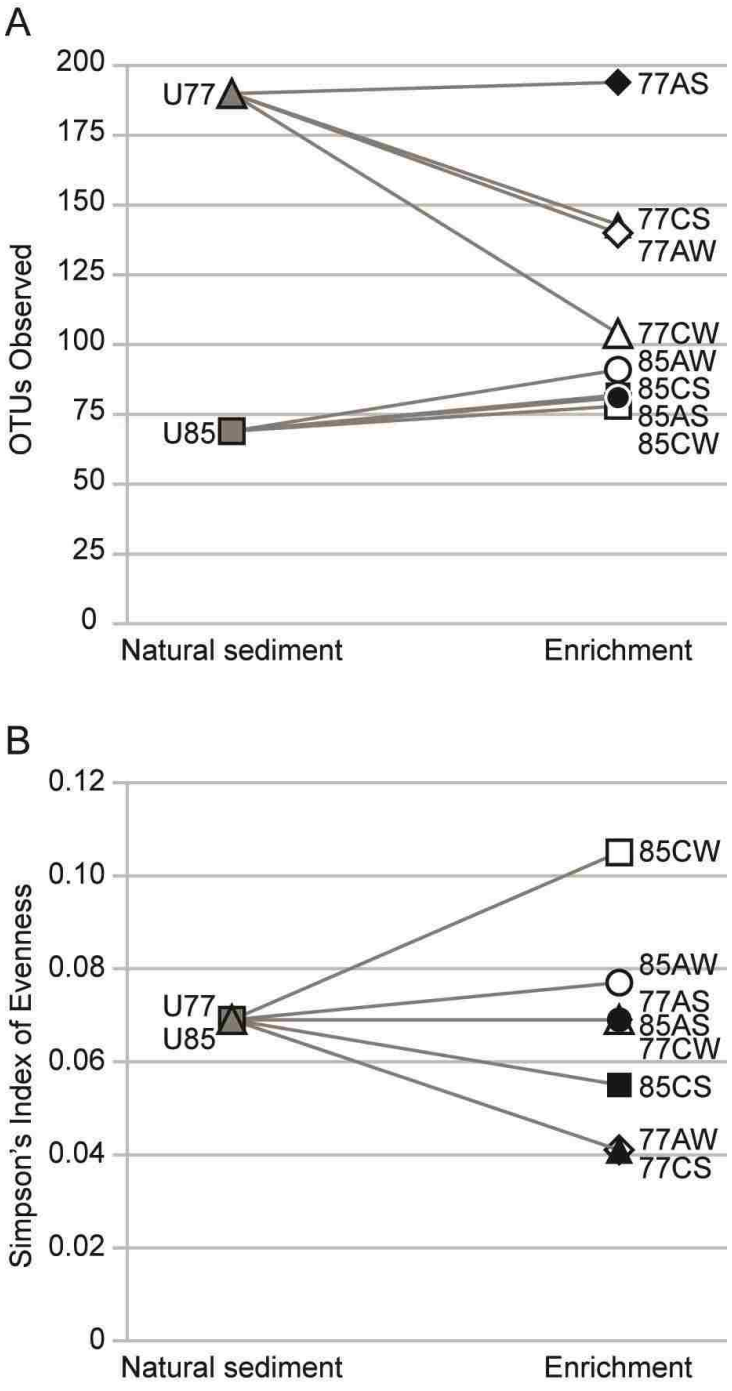


**Figure 3.3.** Natural and enriched samples clustered based on Bray-Curtis dissimilarity calculations of rarefied samples. (a) Cluster tree with samples grouped according to the similarity of the community composition of the samples. All nodes were supported by jackknife scores ≥ 99.9% after 1000 permutations. (b) PCoA of sample distances on principal coordinate 1 (P1) and principal coordinate 2 (P2), with a total of 63.85% of variation explained. (c) PCoA showing sample distances on principal coordinate 2 (P2) and principal coordinate 3 (P3), with a total of 46.29% of variation explained.

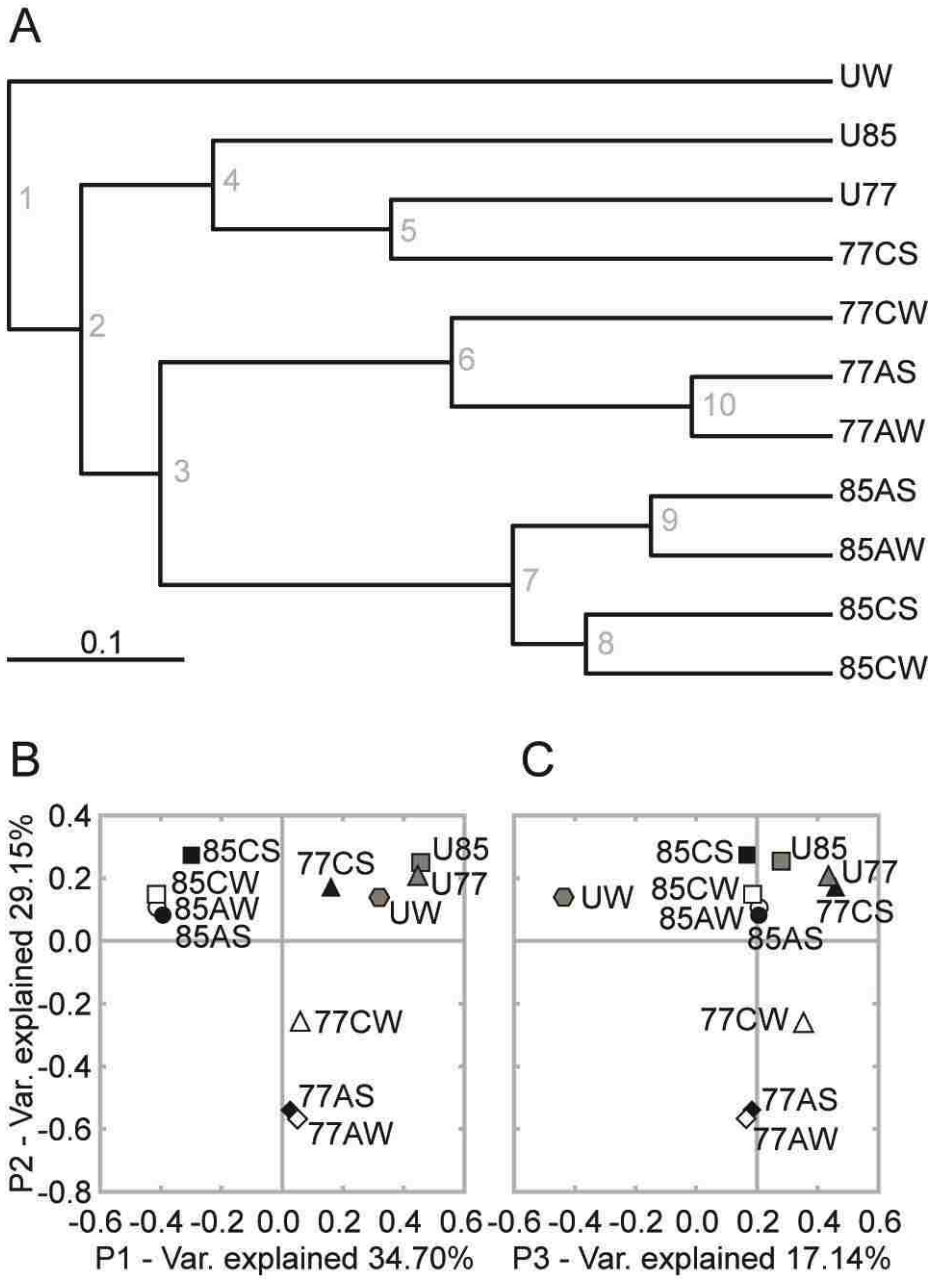
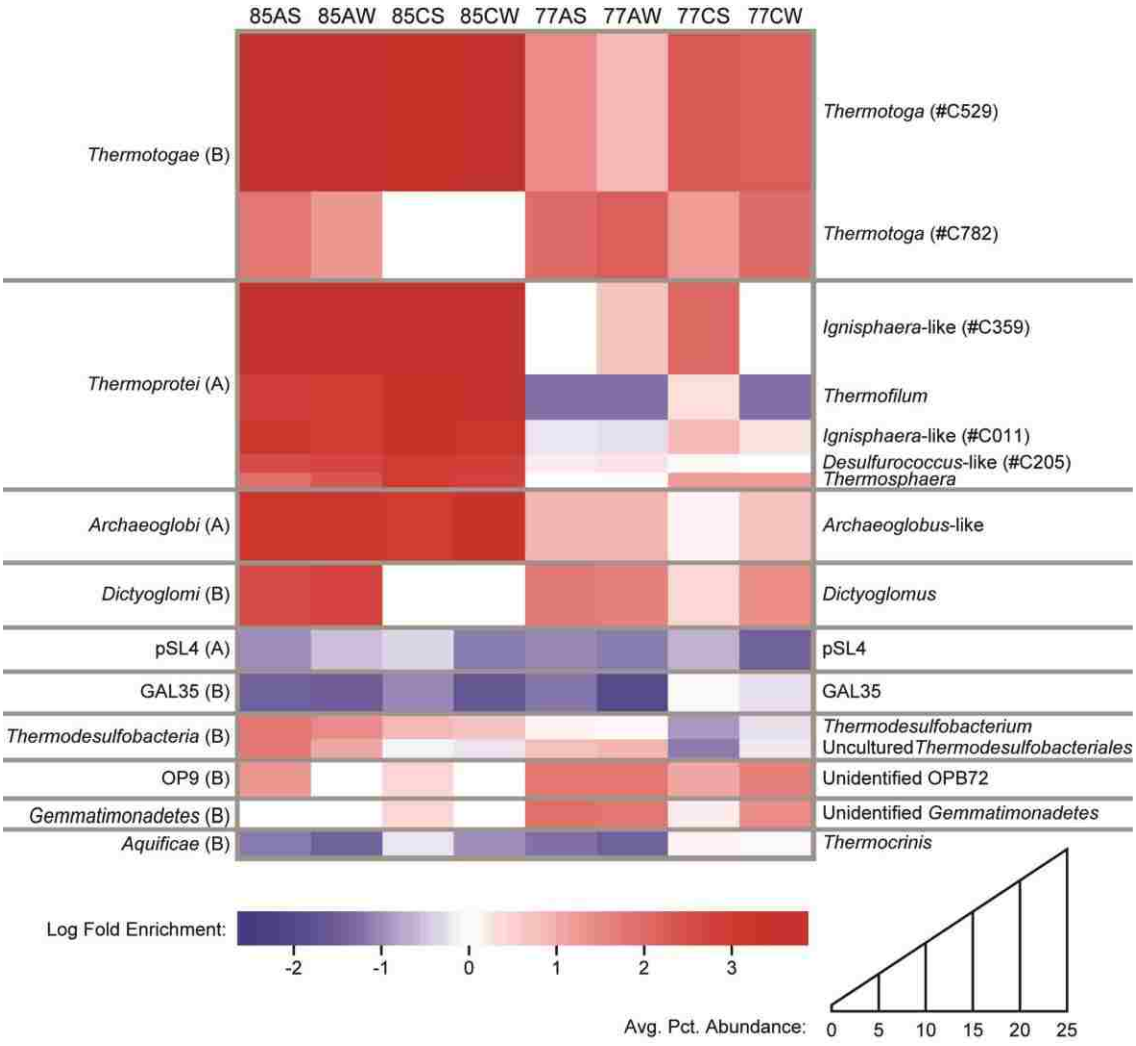


**Figure 3.4.** Heatmap showing log fold enrichment of highly abundant genera and OTUs of specific interest. Taxa are scaled vertically based on percent representation in all enrichment samples, as shown in average percent abundance key. Red, increased relative abundance over natural sediment community at same sampling site; white, no change; blue, decreased relative abundance.



# CHAPTER 4

# *KALLOTENUE PAPYROLYTICUM* GEN. NOV., SP. NOV., A CELLULOLYTIC AND FILAMENTOUS THERMOPHILE ISOLATED FROM GREAT BOILING SPRING THAT REPRESENTS A NOVEL LINEAGE (*KALLOTENUACEAE* ORD. NOV., *KALLOTENUALES* FAM. NOV.) WITHIN THE CLASS *CHLOROFLEXI*


## Abstract

Several closely-related, novel, thermophilic, and cellulolytic bacterial strains, designated JKG1$^T$, JKG2, JKG3, JKG4, and JKG5, were isolated from a cellulolytic enrichment (corn stover) incubated in the water column of Great Boiling Spring, NV. Strain JKG1$^T$ had cells of a diameter of 0.7 - 0.9 µm and length of 1.0 - 2.0 µm that formed unbranched multicellular filaments up to 600 µm. Spores were not formed and dense liquid cultures were red. The temperature range for growth was 45-65 °C, with an optimum of 55 °C. The pH range for growth was 6.5-9.5, with an optimum of 7.5. JKG1$^T$ grew as an aerobic heterotroph, could utilize glucose, cellobiose, xylose, arabinose, sodium pyruvate, sodium lactate, yeast extract, carboxymethylcellulose, filter paper, microcrystalline cellulose, xylan, and starch as its sole carbon and energy sources, and was not observed to photosynthesize. The cells stained Gram-negative. Phylogenetic analysis using 16S rRNA gene sequences placed the new isolates in the class *"Chloroflexi"*, but distant from other cultivated members, with the highest sequence identity 82.5% to *Roseiflexus castenholzii*. The results of chemotaxonomic analyses supported placement of the novel strain within the class *"Chloroflexi"*. Specifically, the predominant quinone was menaquinone 9; no ubiquinones were detected. The major cellular fatty acids were $C_{18:0}$, anteiso-$C_{17:0}$, iso-$C_{18:0}$, and iso-$C_{17:0}$. C16:0, iso-$C_{16:0}$, $C_{17:0}$,

and anteiso-$C_{19:0}$. The peptidoglycan amino acids were alanine, ornithine, glutamic acid, serine, and asparagine. Whole-cell sugars included mannose, rhamnose, glucose, galactose, ribose, arabinose, and xylose. Morphological, phylogenetic, and chemotaxonomic results suggest that JKG1$^T$ is representative of a new lineage within the class *"Chloroflexi"*, which we propose to designate *Kallotenue papyrolyticum* gen. nov., sp. nov., *Kallotenuales* fam. nov., *Kallotenuaceae* ord. nov.

## Introduction

The phylum "*Chloroflexi"* is a deeply-branching lineage of bacteria composed of a limited number of cultivated representatives that display diverse metabolic strategies and phenotypes. The class *"Chloroflexi*" contains gliding filamentous bacteria divided into two orders, *Chloroflexales* and *Herpetosiphonales*. *Chloroflexales* is comprised of anoxygenic phototrophs that have bacteriochlorophyll and may or may not have chlorosomes (Dubinina & Gorlenko 1975; Hanada *et al.* 2002; Keppen *et al.* 1994; Pierson & Castenholz 1974; Pierson *et al.* 1985). *Herpetosiphonales* includes only one genus, *Herpetosiphon*, the two described species of which, *H. geysericola* and *H. aurantiacus,* are mesophilic and aerobic heterotrophs. *H. aurantacus* utilizes carbohydrates as its carbon source, with 6.5% of its genome devoted to carbohydrate transport and metabolism (Kiss *et al.* 2011). However, *H. geysericola*, reported to degrade cigarette paper, is the only member of the class *"Chloroflexi"* known to hydrolyze either soluble or insoluble cellulose (Lewin 1970).

## Methods

Habitat and enrichment.

Five bacterial strains (JKG1$^T$, JKG2, JKG3, JKG4, and JKG5) were isolated from a lignocellulosic enrichment incubated in Great Boiling Spring (GBS). GBS is a circumneutral geothermal spring, located in northwestern Nevada at N40°39'41" W119°21'58", USA (Costa *et al.* 2009). *In situ* enrichments composed of either corn stover or aspen shavings were incubated in two locations in GBS, Site 77 (77 °C) and Site 85 (85 °C ;Peacock *et al.* 2012, submitted). The enrichments were composed of 20 g of either ammonia fiber explosion (AFEX)-treated corn stover (abbreviated "C"; generously provided by Bruce Dale, Michigan State University) or aspen shavings (abbreviated "A"; Kaytee Products, Chilton, WI) enclosed within individual mesh nylon liquid filter bags of 100 μm pore size (Pentair Industrial, Hanover Park, IL) sewn shut with nylon thread. Bags were either suspended in the spring water column ~10 cm below the air-water interface or buried approximately 1 cm below the sediment surface at each incubation site. All eight enrichments were placed in the spring on August 29th, 2009, with those at Site 77 harvested on November 1, 2009 and Site 85 on November 29, 2009. Samples were maintained on dry ice during transport back to the laboratory.

Cultivation and isolation conditions

A modified version of Castenholz Medium D (Castenholz 1969), designated Castenholz medium D VN (abbreviated VN medium), was prepared with the addition of 1X Wolfe's Vitamins (Balch & Wolfe 1976) and 0.027 g $NH_4Cl$. Ten mL of VN medium was prepared in 30-mL glass screw-top culture tubes with 1% w/v filter paper (Whatman Filter Paper Grade 1, GE Healthcare, Piscataway, NJ) as the sole carbon and energy source and air as headspace (abbreviated VNF medium). Material (corn stover or aspen shavings) from each of the eight enrichments was obtained using sterile forceps and used

to inoculate three screw-top tubes, each of which was incubated at 60 °C, 70 °C, or 80 °C. After 8 weeks of incubation at 60 °C a biofilm was visible in the culture inoculated from the corn stover enrichment located in the water column of GBS at Site 85. The biofilm was composed of a mixed culture and was maintained on VNF medium until pure cultures could be obtained. Traditional approaches of plating or dilution to extinction failed to yield cellulolytic isolates Therefore, single cells were procured by loading the mixed culture onto a sterile microfluidic device, in which individual filaments were isolated using optical tweezers (Youssef *et al.* 2011) and recovered from the chip by flushing with VNF medium. Five strains (JKG1, JKG2, JKG3, JKG4, and JKG5) of similar morphology were cultivated from single filaments and strain JKG1$^T$ was used for further characterization.

Morphology

The colony morphology of strain JKG1$^T$ was observed after seven days of incubation at 55 °C on VN medium with 0.2% w/v glucose and buffered at pH 7.5 with 15 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES; JT Baker, Avator Performance Materials, Center Valley, PA ;abbreviated VNG medium). VNG medium was solidified by the addition of 1.5% w/v agar (BD Difco agar granulated, Franklin Lakes, NJ). Filament morphology was observed using phase-contrast microscopy with an Olympus BX51 phase-contrast microscope and Olympus V-TV1X-2 camera.

Temperature and pH.

Growth ranges of pH and temperature were tested using VNG medium. Cultures were prepared in glass 25-mL Balch tubes stoppered with butyl rubber septa (Wheaton, Millville, NJ) with an air headspace. The following pH values were tested with 15 mM of

the indicated buffer: pH 5.5, 6.0, and 6.5, 2-(N-Morpholino)ethanesulfonic acid (MES); pH 7.0, 7.5, and 8.0, HEPES buffer; and pH 8.5, 9.5, and 10.0 2-(N-cyclohexylamino)-ethanesulfonic acid (CHES). Temperature was tested in increments of 5 °C from 35 to 75 °C.

Nutritional and physiological tests

VN medium buffered to pH 7.5 with 15 mM HEPES was used to test the utilization of a variety of substrates with an air headspace. Substrates were tested in Balch tubes incubated in the dark at 55 ºC for 36 hours at a 20° angle with orbital shaking at 100 RPM. Substrates were tested at 0.2% w/v and included the following: glucose (BD, Franklin Lakes, NJ), sucrose (EMD, Billerica, MA), D (+) cellobiose (Calbiochem, EMD, Billerica, MA), xylose (Spectrum, Gardena, CA), arabinose (Calbiochem, EMD, Billerica, MA), and sodium pyruvate (JT Baker, Avator Performance Materials, Center Valley, PA), casamino acids (EMD, Billerica, MA), tryptone (BD Bacto, Franklin Lakes, NJ), peptone (BD Bacto, Franklin Lakes, NJ), yeast extract (EM Science, EMD, Billerica, MA), acetate (EM Science, VWR, Radnor, PA), citrate (EM Science, VWR, Radnor, PA), glycerol ( EMD, Billerica, MA), lactate (EM Science, EMD, Billerica, MA), carboxymethylcellulose sodium (Spectrum, Gardena, CA), filter paper (Whatman Filter Paper Grade 1, GE Healthcare, Piscataway, NJ), microcrystalline cellulose (extra pure 90 μm, Acros Organics, Thermo Fisher Scientific, Fair Lawn, NJ), xylan (birch wood, Sigma-Aldrich, St. Louis, MO), and starch (soluble potato starch powder, JT Baker, Avator Performance Materials, Center Valley, PA). Three consecutive transfers were conducted for each substrate, along with an inoculated negative control that contained no carbon source. The third transfer was performed in triplicate. Substrate

utilization was determined to be positive if the average OD of the final transfers exceeded that of the average OD of the negative control final transfers. Optical density was determined using a Spectronic 20D spectrophotometer (Milton Roy).

Growth under anaerobic conditions was tested using Balch tubes with $N_2$-gassed VN medium buffered to pH 7.5 with 15 mM HEPES, and a headspace of $N_2$. Tubes were incubated in the dark at 55 ºC for four days with orbital shaking at 100 RPM at a 20° angle. The substrates glucose, casamino acids, and yeast extract were each tested at 0.2% w/v.

Phototrophic growth was tested using $N_2$-gassed VN medium buffered to pH 7.5 with 15 mM HEPES. Cultures were prepared in Balch tubes with $N_2$ as headspace and 0.3 mM $Na_2S$ as an electron donor. Tubes were incubated under the illumination of a full-spectrum LED bulb (Grow Lite, Stimulus Brands, Burnsville, MN) at 55 ºC for four days with orbital shaking at 100 RPM at a 20° angle. Photoheterotrophic growth was tested using the substrates glucose, citrate, and yeast extract at 0.2% w/v. Photoautotrophic growth was tested by the addition of 50 mM $NaHCO_3$.

Chemotaxonomic analyses

Cultures for chemotaxonomic analyses were grown aerobically in R2A medium at 55 °C with constant orbital shaking of 100-150 RPM (Reasoner & Geldreich 1985). Glucose was prepared separately as a sterile solution and added after autoclaving. Lyophilzed pellets were produced by centrifugation of cultures at 18,514 rcf and lyoplilization of pellets with a Virtis lyophilizer (The Virtis Co.,Inc., Gardiner, NY) under a vacuum of 12 Pascals. Fatty acid methyl ester analysis was performed on a lyophilized pellet by MIDI Labs (Newark, DE ;Sasser, 2006). Lyophlilized pellets were

analyzed to determine whole-cell sugars, peptidoglycan amino acids, and quinones. Preparation of purified cell-wall and peptidoglycan analysis were performed by TLC and HPLC as described by Schleifer (1985) and Tang *et al*. (2009). The whole-cell sugars were prepared and analyzed according to the method as described by Tang *et al*. (2009). Menaquinones were extracted from lyophilized cells as described by Collins *et al*. (1977) and Minnikin *et al*. (1984) and then analyzed by HPLC (Tamaoka *et al.* 1983).

DNA extraction, PCR, and DNA sequencing

DNA from all five strains was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH) using modifications specified previously (Dodsworth *et al.* 2011). PCR was performed using primers 9bF (gRg ttt gat cct ggc tca g ;Eder *et al.* 1999) and 1512uR (acg gHt acc ttg tta cga ctt ;Eder *et al.* 2001). PCR products were sequenced with the PCR primers and with 704bR (tct acg Yat ttc acY gct) and 516uF (tgB cag cMg ccg cgg taa) to obtain overlapping reads. Cycling conditions were as follows: 95 °C for 4 min; 29 steps of 95 °C for 30 s, 55 °C for 1 min, 72 °C for two min; and a final extension step of 7 min at 72 °C. PCR products were shipped to Functional Biosciences (Madison, WI) and sequenced using the ABI 3730xl DNA Sequencer. The ends of the 16S rRNA gene sequences were trimmed to remove bases with quality scores of less than 20, aligned against the mothur-provided SILVA alignment in the program mothur v1.20.2 (Schloss *et al.* 2009), and trimmed to the shortest sequence, resulting in an alignment of the near-full length 16S rRNA gene sequences of all five strains 1,355 nt in length.

Phylogenetic analysis

An alignment of 16S rRNA gene sequences of members of the phylum "*Chloroflexi*", including strain JKG1$^T$ and reference sequences obtained from the NCBI, was prepared by aligning all sequences against the mothur-provided SILVA alignment in the program mothur. Bioedit v7.0.5.3 was used to manually curate the alignment and calculate the pairwise 16S rRNA gene sequence identities (Hall 1999). The mothur-provided SILVA-compatible 1,349-position Lane mask was applied (Lane, D. J. 1991). PHYLIP v3.69 was used to produce neighbor-joining, maximum parsimony, and maximum likelihood phylogenetic trees with 100 bootstraps per tree (Felsenstein, 2005). The following parameters in PHYLIP were modified as indicated: input order of species was randomized, outgroup root was set to *B. subtilis*, jumbles were 1, global rearrangements on, speedier but rougher analysis off, and more thorough search on. Trees were visualized and exported using Dendroscope v2 (Huson *et al.* 2007).

## Results and Discussion

### Morphology

Strain JKG1$^T$ was isolated from a lignocellulosic enrichment incubated in Great Boiling Spring, NV. The isolate formed small white rhizoid colonies. Gliding motility was observed. The isolate was an unbranched, multicellular, and filamentous bacterium with a diameter of 0.7 - 0.9 μm and filament lengths up to 600 μm (Figure 4.1). Filaments contained individual cells of 1.0 - 2.0 μm in length, and transparent sections resembling the sleeves observed in *Herpetosiphon* sp. were sometimes observed between cells or at the ends of the filaments (Lee & Reichenbach 2006). Dense liquid cultures were red and spores did not form. The isolate stained Gram-negative (Leboffe & Pierce 2006).

### Physiological properties

Strain JKG1$^T$ was able to grow heterotrophically under aerobic conditions utilizing glucose, cellobiose, xylose, sodium pyruvate, sodium lactate, yeast extract, carboxymethylcellulose, filter paper, microcrystalline cellulose, xylan, and starch,as sole carbon sources. Optimal growth occurred at 55 °C and pH 7.5. Growth occurred from 45 - 65 °C and pH 6.5 - 9.5. Photoheterotrophic, photoautotrophic, and fermentative growth were not observed.

Chemotaxonomic analyses

The major cellular fatty acids of strain JKG1$^T$ were $C_{18:0}$ (26.30% of total fatty acids), anteiso-$C_{17:0}$ (15.02%), iso-$C_{18:0}$ (12.68%), iso-$C_{17:0}$ (11.45%), $C_{16:0}$ (8.94%), iso-$C_{16:0}$ (6.17%), $C_{17:0}$ (4.96%), and anteiso-$C_{19:0}$ (3.91%). The minor fatty acids were anteiso-$C_{15:0}$ (1.87%), iso-$C_{19:0}$ (1.44%), ω9c-$C_{18:1}$ (1.37%), $C_{19:0}$ (1.04%), 3OH-$C_{17:0}$ (0.88%), ω7c-$C_{18:1}$ (0.75%), iso-$C_{15:0}$ (0.72%), iso-ω9c-$C_{17:1}$ (0.56%), anteiso-ω9c-$C_{17:1}$ (0.54%), anteiso-$C_{13:0}$ (0.51%), $C_{15:0}$ (0.49%), and $C_{14:0}$ (0.39%).

The peptidoglycan amino acids were alanine (ala), ornithine (orn), glutamic acid (glu), serine (ser), and asparagine (asn). Meso-diaminopimelic acid was not present, and peptide cross-links are likely formed between ornithine and D-alanine, as in other members of the phylum "*Chloroflexi*"(Cavaletti *et al.* 2006; Jürgens *et al.* 1987 1989; Yabe *et al.* 2010 2011). In particular, ornithine has been found to replace DAP as the diamino acid that forms the cross-peptide bonds (Garrity & Holt 2001). The predominant quinone was MK-9(H6) (84.1% relative abundance), with a minor concentration of MK-8(H2) (15.9%). Whole-cell sugars were found to be mannose (man), rhamnose (rham), glucose (glc), galactose (gal), ribose (rib), arabinose (ara), xylose (xyl), and two unknown compounds.

Phylogenetic analysis

The 16S rRNA gene sequences of all five strains were identical over the length of the alignment shared by all five sequences (1,355 nt). Sequence identity between strain JKG1$^T$ and other members of the class "*Chloroflexi*" was low, with the organism most closely related to strain JKG1$^T$, *R. castenholzii,* sharing 82.5% identity. Phylogenetic analysis of JKG1$^T$ using the 16S rRNA gene sequence placed the strain in the class "*Chloroflexi*", with bootstrap values of at least 90% at the node defining the class in all three phylogenetic tree-building methods utilized (maximum likelihood, neighbor-joining, and maximum parsimony; Figure 4.2).

## Conclusions

Members of the class "*Chloroflexi"* share a nonbranching filamentous morphology, with the majority of the isolates capable of aerobic respiration (Table 4.1). However, strain JKG1$^T$ is distinguishable from these other members of the class because it is the only thermophilic chemoorganotroph not capable of photosynthesis. In addition, strain JKG1$^T$ harbors distinct profiles of fatty acids, quinones, whole-cell sugars, and cell wall amino acids, along with low 16S rRNA gene sequence identity to other cultivated members of the class. Altogether, the unique phylogenetic, phenotypic, and chemotaxonomic features of strain JKG1$^T$ suggest that it is representative of a novel genus, family, and order within the class "*Chloroflexi*". The designations *Kallotenue* gen. nov., *Kallotenales* fam. nov., and *Kallotenuaceae* ord. nov., are proposed.

**Description of *Kallotenuaceae* ord. nov.**

*Kallotenuaceae*: *Kal.lo.te.nu.a.ce.a'e*. N.L. n. *Kallotenue*, type genus of the order; suff. -*aceae*, ending denoting an order; N.L. fem. pl. n. *Kallotenuaceae*, the order of the genus Kallotenue.

The description is the same as for the genus *Kallotenue*. The order is a member of the class "*Chloroflexi*". The type family is *Kallotenuales*.

**Description of *Kallotenales* fam. nov.**

*Kallotenuales*: *Kallotenuales*. N.L. n. *Kallotenue*, type genus of the family; suff. -ales, ending denoting an family; N.L. fem. pl. n. *Kallotenuales*, the family of the genus *Kallotenue*.

The description is the same as for the genus *Kallotenuae*. The family is a member of the order *Kallotenuaceae*. The type genus is *Kallotenue*.

**Description of *Kallotenue* gen. nov.**

*Kallotenue*: *Kal.lo.te'nu.e*. Gr. neut. n. *kallos*, beauty, grace; L. adj. *tenuis* -is -e, slender, fine, thin; N.L. neut. n. *Kallotenue*, a thin beauty.

Filamentous, nonbranching, and nonsporeforming bacteria. Gram-stain negative. Thermophilic, aerobic chemoorganoheterotrophs that grow at 45 - 65 ºC and pH 6.5 - 9.5. Peptidoglycan amino acids include alanine, ornithine, glutamine serine, and asparagine. Major whole-cell sugars comprise mannose, rhamnose, and glycogen. The predominant quinone is menaquinone 9. The type species is *Kallotenue papyrolyticum*.

**Description of *Kallotenue papyrolyticum* sp. nov.**

*Kallotenue papyrolyticum*: *Pa.py.ro.ly'ti.cum*. Gr. n. *papuros*, *papyrus*, paper; N.L. neut. adj. *lyticum* (from Gr. neut. adj. *lutikon*), able to loose, able to dissolve; N.L. neut. adj. *papyrolyticum*, paper-dissolving.

The following properties are displayed, in addition to those specified for the genus. Growth occurs at optimum temperature of 55 ºC and optimum pH of 7.5. Colonies are small, white, and rhizoid. The following substrates support growth: glucose, cellobiose, xylose, sodium pyruvate, sodium lactate, yeast extract, carboxymethylcellulose, filter paper, microcrystalline cellulose, xylan, and starch. The major cellular fatty acids are $C_{18:0}$, anteiso-$C_{17:0}$, iso-$C_{18:0}$, iso-$C_{17:0}$, $C_{16:0}$, iso-$C_{16:0}$, $C_{17:0}$, and anteiso-$C_{19:0}$.

**Table 4.1.** Phenotypic characteristics of strain JKG1$^T$ and related members of the class "*Chloroflexi*"

The table was constructed based in part upon Table 1 in Hanada and Pierson (2006). -, Negative;±, variable;+, positive; nd, not determined. [a]Ara, Arabinose; Gal, Galactose; Glc, Glucose; Man, Mannose; Rham,Rhamnose ; Rib, Ribose; Unk, Unknown; Xyl, Xylose. [b] Ala, Alanine; Asn, Asparagine; Glu, Glutamic Acid; Gly, Glycine; His, Histidine; Orn, Ornithine; Ser, Serine.

| Physiological subgroup | Aerobic chemotrophs | | | Chlorosome-less phototrophs | Chlorosome-containing phototrophs | |
|---|---|---|---|---|---|---|
| Taxon | | *Herpetosiphon aurantiacus* | *Herpetosiphon geysericola* | *Roseiflexus castenholzii* | *Chloroflexus aurantiacus* | *Oscillochloris trichoides* |
| Type strain | JKG1$^T$ | DSM 785$^T$ | ATCC23076$^T$ | DSM 13941$^T$ | J-10-fl$^T$ | DG-6$^T$ |
| Cell diameter (μM) | 0.7-0.9 | 1.0-1.5 | 1.0 | 0.8-1.0 | 0.7-1.2 | 1.0-1.5 |
| Sleeves | + | + | + | - | ± | ± |
| Gas vesicles | - | - | - | - | - | + |
| Gram stain results | Negative | Negative | Negative | Negative | Negative | Variable |
| Optimum temp. (°C) | 55 | 30-37 | 30-37 | 55 | 55 | 28-30 |
| Metabolism: | | | | | | |
| Photoheterotrophy | - | - | - | + | + | + |
| Photoautotrophy | - | - | - | - | ± | + |
| Aerobic respiration | + | + | + | + | + | - |
| Fermentation | - | nd | nd | - | nd | nd |
| Cellulolytic ability: | | | | | | |
| carboxymethyl cellulose | + | - | nd | nd | nd | nd |
| insoluble cellulose | + | - | + | nd | nd | nd |
| Major cellular fatty acids | $C_{18:0}$, anteiso-$C_{17:0}$, iso-$C_{18:0}$, and iso-$C_{17:0}$. C16:0, iso-$C_{16:0}$, $C_{17:0}$, and anteiso-$C_{19:0}$ | $C_{16:1}$, $C_{16:0}$, $C_{18:1}$ | $C_{16:1}$, $C_{16:0}$, $C_{18:1}$ | $C_{16:0}$, $C_{14:0}$, $C_{15:0}$ | $C_{18:0}$. $C_{16:0}$, $C_{18:1}$, $C_{17:0}$ | $C_{18:1}$, $C_{16:0}$, $C_{16:1}$ |
| Predominant quinone | MK-9 | MK-6 | nd | MK-11 | MK-10 | MK-10 |
| Whole-cell or [cell wall] sugars[a] | Man, Rham, Glc, Gal, Rib, Ara, Xyl, Unk (x2) | [Man, Gal, Rham, Ara, Xyl] | nd | nd | [Rham, Man, Glc, Gal, Xyl, Ara] | nd |
| Cell wall amino acids[b] | Ala, Orn, Glu, Ser, Asn | Glu, Ala, Orn, Gly, His | nd | nd | Glu, Ala, Orn, Gly, His | nd |
| Isolation environment | Lignocellulosic enrichment in terrestrial hot spring | Slimy coating of *Chara* sp. in freshwater lake | Rock surface beside hot spring | Bacterial mat in hot spring | Hot spring | Hydrogen sulfide-containing spring |
| References | Present study | (Holt & Lewin 1968; Jürgens *et al.* 1989; Kiss *et al.* 2011) | (Lewin 1970) | (Hanada *et al.* 2002) | (Jürgens *et al.* 1987; Knudsen *et al.* 1982; Pierson & Castenholz 1974) | (Berg *et al.* 2005; Keppen *et al.* 1994) |

[a]Ara, Arabinose; Gal, Galactose; Glc, Glucose ; Man, Mannose; Rham,Rhamnose ; Rib, Ribose; Unk, Unknown; Xyl, Xylose

[b] Ala, Alanine; Asn, Asparagine; Glu, Glutamic Acid; Gly, Glycine; His, Histidine; Orn, Ornithine; Ser, Serine

**Figure 4.1** Phase-contrast micrograph of strain JKG1$^{T}$.

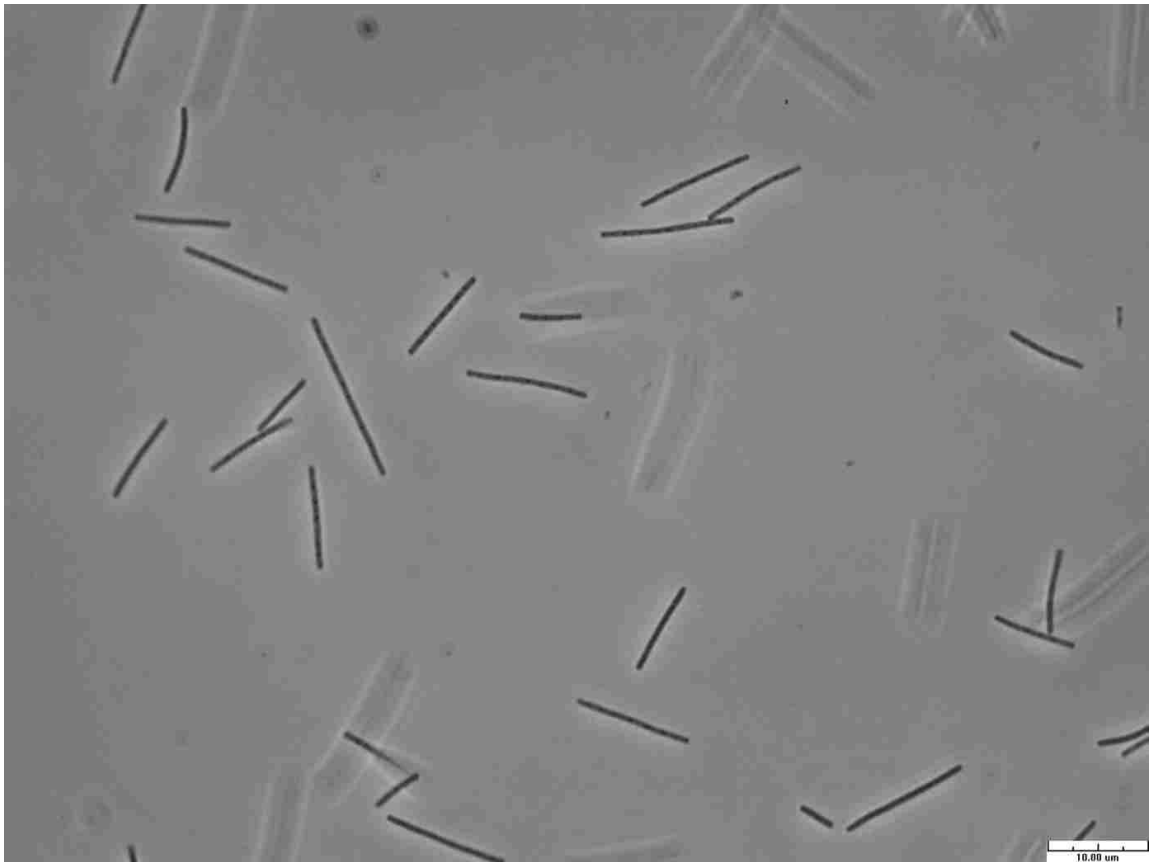
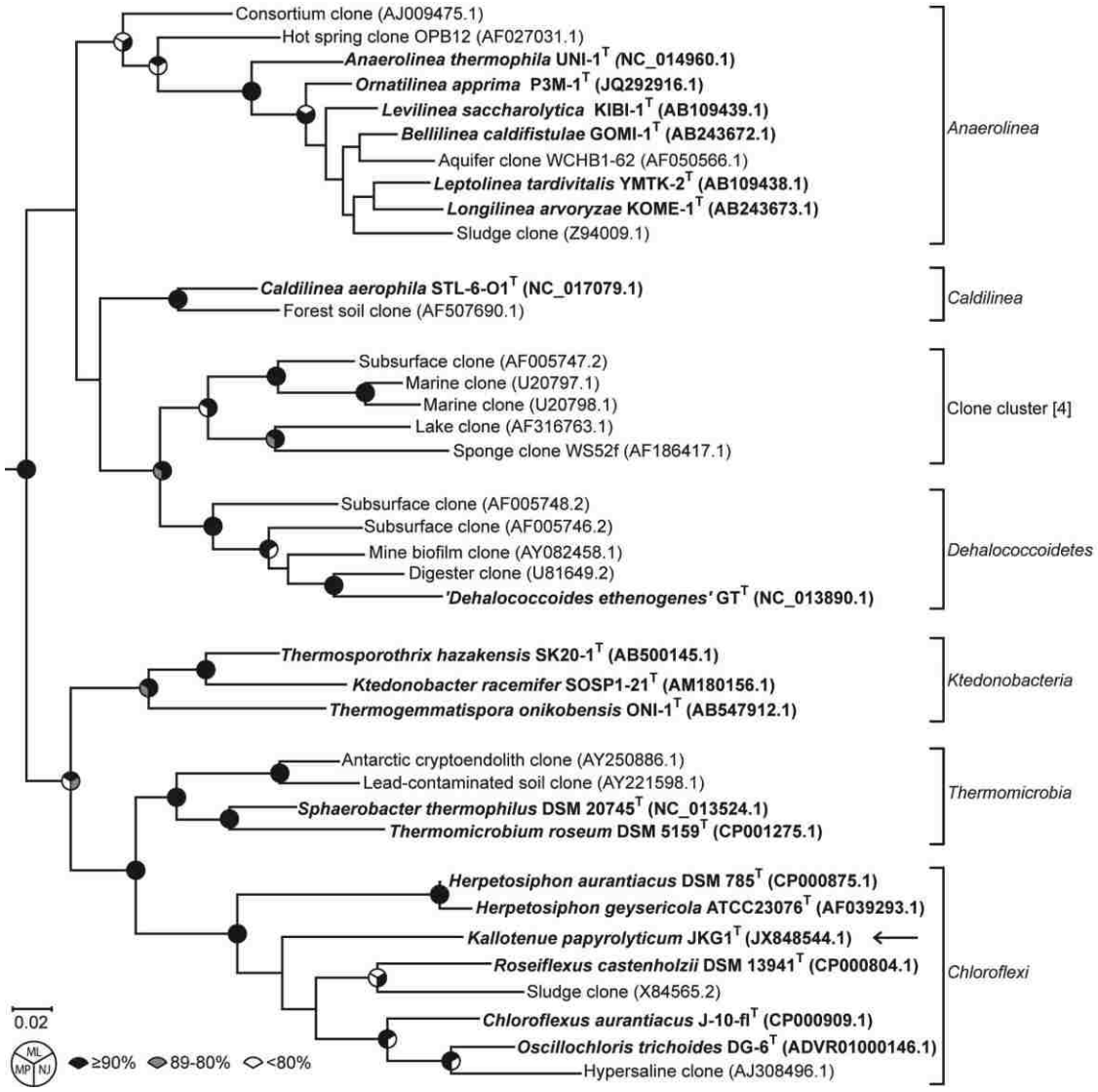


**Figure 4.2** Maximum likelihood phylogenetic tree of the phylum "*Chloroflexi"*, with strain JKG1$^T$ indicated by an arrow. The tree was constructed based upon 1,183 aligned positions that remained after the application of the Lane mask to the 16S rRNA gene sequences. Bootstrap values are out of 100 replicates for each tree-building method represented (maximum likelihood (ML), maximum parsimony (MP), and neighbor-joining (NJ)) and are indicated at each node using black (≥90% recovery), grey (89 - 80% recovery), or no shading (<80% recovery). Scale bar indicates 0.02 changes per nucleotide. Accession numbers are included in parentheses and strains are included for cultivated representatives. Outgroups were *Bacillus subtilis* 168$^T$ (D26185.1), *Escherichia coli* ATCC 11775$^T$ (X80725.1), and *Corynebacterium diphtheriae* NCTC 11397$^T$ (X84248.1).



# CHAPTER 5

# CONCLUSIONS

This study explored the composition and structure of both natural and cellulolytic microbial communities in GBS, and described the isolation and characterization of a novel cellulolytic isolate representative of an order-level group. The natural sediment and water communities were found to be distinct, substantiating the importance of considering the water and sediment communities of hot springs separately. Natural high-temperature sediment was found to contain an abundance of novel taxonomic groups, including those at the phylum- and class-levels, indicating that the bacteria and archaea present in high-temperature environments such as GBS have yet to be sufficiently described. Further investigation into the functions of the novel phylotypes will be necessary for the hot spring understood as a whole.

Novel microorganisms associated with lignocellulose-degrading communities could yield novel enzymes useful to biofuel production. Second-generation biofuels technology is currently too costly to implement on a large scale, primarily due to the large quantities of enzyme necessary to hydrolyze the lignocellulose into fermentable sugars. This has prompted a search for more efficient and stable enzymes (for a review see Carroll *et al.* 2009), with thermostable enzymes considered a viable solution. The work presented here addressed the search for new enzymes by investigating the microbial composition of high-temperature lignocellulose enrichments. Fundamental knowledge of the composition of high-temperature lignocellulolytic microbial communities is a primary foundation for future high-temperature biofuels work. Enrichment with lignocellulose

promoted a pronounced shift in the natural microbial community, resulting in communities dominated by organisms relatively rare in the natural spring communities. However, the microbes found to be dominant members of the enriched communities were in many cases related to cultivated organisms, indicating that the most abundant microbial lineages involved in the degradation of cellulose may already be known to science. Therefore, members of these lineages may be the most promising candidates for the discovery of lignocellulolytic enzymes.

Strain JKG1$^T$ was isolated from the corn stover enrichment in the water column of GBS at 85 °C, but was not found in any of the natural or enriched pyrosequencing datasets. Since primer sequences were confirmed to match the near-full length 16S rRNA gene sequence of strain JKG1$^T$, it is likely that the absence of the novel isolate in the pyrosequencing datasets is due to a simple lack of abundance of this organism in GBS and the lignocellulosic enrichments. In addition, the enrichment that strain JKG1$^T$ was isolated from was primarily anaerobic and incubated at an average temperature of 85 °C. However, strain JKG1$^T$ was only able to be cultivated in the laboratory under strictly aerobic conditions and at temperatures no greater than 65 °C. In regard to metabolism, it is possible that this strain can perform anaerobic respiration utilizing combinations of electron donors and acceptors that we did not test. In addition, the strain may be able to perform fermentation but produces hydrogen as a waste product, thereby needing a hydrogenotroph to maintain homeostasis. In regard to temperature, the lack of growth at temperatures greater than 65 °C could be an example of an unexplained phenomenon sometimes observed when comparing the *in vitro* to *in situ* growth temperatures of microorganisms. The temperature range of some organisms has been found to be

significantly lower during laboratory cultivation than that observed in the organism's natural environment.

A thermophilic lifestyle requires every part of the cell to be stable and functional at high temperature. The thermophile must possess mechanisms to moderate challenges raised by high temperatures, such as increased membrane fluidity and permeability and increased levels of damage to genetic material (for a review see Charlier *et al.* 2005). Difficulties such as these are likely responsible for the negative relationship we have demonstrated between temperature and the richness of microbial communities, both natural and lignocellulose enriched. Additional surveys into the diversity along thermal gradients in other environments, such as geothermal soils, hydrothermal vents, mud pots, and acidic or alkaline hot springs, would be informative as to the universality of these observations into the relationship between temperature and microbial richness.

# APPENDIX A: ADDITIONAL TABLES

**Table A.1.** Alpha diversity calculations of combined archaea and bacteria communities at OTU percent definitions 80, 85, 92, 95, 97.

| Archaea & Bacteria | | 0906W | 1002W | 0706W | 0812W | 1007A | 0906A | 1002A | 0812A | 1002B | 1007C | 1002C | 1002E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Temp (°C) | | 82 | 82 | 80 | 80 | 87 | 82 | 82 | 80 | 79 | 76 | 72 | 62 |
| No. of OTUs | 80% | 11 | 11 | 10 | 15 | 10 | 11 | 13 | 15 | 15 | 21 | 21 | 27 |
| | 85% | 14 | 21 | 21 | 28 | 15 | 22 | 22 | 20 | 29 | 43 | 38 | 63 |
| | 92% | 17 | 28 | 32 | 43 | 20 | 28 | 28 | 30 | 52 | 78 | 68 | 151 |
| | 95% | 18 | 33 | 40 | 47 | 20 | 30 | 31 | 37 | 60 | 101 | 73 | 196 |
| | 97% | 25 | 44 | 48 | 61 | 32 | 42 | 61 | 90 | 109 | 143 | 103 | 252 |
| Chao1 | 80% | 16 | 11 | 10 | 16 | 10 | 14 | 23 | 15 | 16 | 21 | 21 | 37 |
| | 85% | 21 | 32 | 25 | 30 | 17 | 31 | 45 | 21 | 36 | 47 | 39 | 72 |
| | 92% | 29 | 68 | 55 | 94 | 22 | 32 | 37 | 35 | 262 | 86 | 81 | 194 |
| | 95% | 27 | 75 | 69 | 101 | 23 | 36 | 57 | 41 | 125 | 112 | 84 | 270 |
| | 97% | 51 | 107 | 103 | 173 | 37 | 61 | 100 | 139 | 184 | 170 | 143 | 402 |
| Simpson's Evenness | 80% | 0.115 | 0.121 | 0.132 | 0.070 | 0.161 | 0.167 | 0.207 | 0.133 | 0.200 | 0.205 | 0.127 | 0.158 |
| | 85% | 0.090 | 0.063 | 0.063 | 0.037 | 0.124 | 0.130 | 0.147 | 0.117 | 0.120 | 0.183 | 0.123 | 0.129 |
| | 92% | 0.074 | 0.047 | 0.041 | 0.024 | 0.133 | 0.150 | 0.125 | 0.084 | 0.077 | 0.167 | 0.125 | 0.089 |
| | 95% | 0.070 | 0.040 | 0.033 | 0.022 | 0.133 | 0.141 | 0.113 | 0.068 | 0.067 | 0.141 | 0.126 | 0.073 |
| | 97% | 0.051 | 0.030 | 0.028 | 0.017 | 0.087 | 0.104 | 0.058 | 0.028 | 0.038 | 0.117 | 0.091 | 0.062 |
| Simpson Diversity | 80% | 0.207 | 0.248 | 0.242 | 0.043 | 0.378 | 0.454 | 0.628 | 0.499 | 0.666 | 0.768 | 0.625 | 0.766 |
| | 85% | 0.207 | 0.248 | 0.243 | 0.043 | 0.461 | 0.650 | 0.691 | 0.572 | 0.713 | 0.873 | 0.786 | 0.876 |
| | 92% | 0.207 | 0.248 | 0.243 | 0.043 | 0.623 | 0.763 | 0.715 | 0.601 | 0.750 | 0.923 | 0.882 | 0.925 |
| | 95% | 0.207 | 0.248 | 0.243 | 0.043 | 0.625 | 0.763 | 0.715 | 0.601 | 0.752 | 0.930 | 0.892 | 0.930 |
| | 97% | 0.209 | 0.249 | 0.244 | 0.046 | 0.642 | 0.770 | 0.719 | 0.604 | 0.756 | 0.940 | 0.894 | 0.936 |
| Shannon Diversity | 80% | 0.549 | 0.626 | 0.764 | 0.197 | 1.201 | 1.234 | 1.676 | 1.271 | 1.888 | 2.519 | 1.838 | 2.598 |
| | 85% | 0.551 | 0.631 | 0.796 | 0.209 | 1.456 | 1.801 | 1.967 | 1.768 | 2.266 | 3.483 | 2.859 | 3.669 |
| | 92% | 0.562 | 0.641 | 0.804 | 0.218 | 1.894 | 2.315 | 2.220 | 2.056 | 2.687 | 4.350 | 3.750 | 4.826 |
| | 95% | 0.563 | 0.642 | 0.811 | 0.222 | 1.911 | 2.323 | 2.226 | 2.078 | 2.743 | 4.562 | 3.875 | 5.083 |
| | 97% | 0.579 | 0.659 | 0.824 | 0.248 | 2.060 | 2.440 | 2.311 | 2.155 | 2.855 | 4.909 | 3.991 | 5.267 |

**Table A.2.** Alpha diversity calculations of archaea at OTU percent definitions 80, 85, 92, 95, 97.

| Archaea | | 0906W | 1002W | 0706W | 0812W | 1007A | 0906A | 1002A | 0812A | 1002B | 1007C | 1002C | 1002E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Temp (°C) | | 82 | 82 | 80 | 80 | 87 | 82 | 82 | 80 | 79 | 76 | 72 | 62 |
| No. of OTUs | 80% | 2 | 1 | 1 | 2 | 3 | 3 | 3 | 4 | 4 | 6 | 8 | 9 |
| | 85% | 2 | 2 | 2 | 3 | 4 | 4 | 5 | 5 | 6 | 11 | 9 | 10 |
| | 92% | 3 | 4 | 6 | 4 | 7 | 7 | 9 | 9 | 11 | 26 | 16 | 17 |
| | 95% | 3 | 4 | 8 | 4 | 6 | 9 | 12 | 13 | 16 | 39 | 17 | 19 |
| | 97% | 7 | 8 | 11 | 5 | 15 | 19 | 30 | 41 | 40 | 53 | 25 | 23 |
| Chao1 | 80% | 2 | 1 | 1 | 2 | 3 | 3 | 3 | 4 | 4 | 6 | 8 | 12 |
| | 85% | 2 | 2 | 2 | 3 | 4 | 4 | 5 | 5 | 6 | 12 | 9 | 12 |
| | 92% | 3 | 4 | 6 | 4 | 7 | 7 | 9 | 9 | 11 | 26 | 16 | 25 |
| | 95% | 3 | 4 | 9 | 4 | 6 | 10 | 17 | 13 | 16 | 39 | 17 | 30 |
| | 97% | 13 | 8 | 14 | 5 | 15 | 31 | 57 | 56 | 48 | 56 | 28 | 38 |
| Simpson's Evenness | 80% | 0.502 | 1.000 | 1.000 | 0.559 | 0.405 | 0.358 | 0.409 | 0.272 | 0.272 | 0.198 | 0.141 | 0.124 |
| | 85% | 0.502 | 0.501 | 0.503 | 0.497 | 0.351 | 0.447 | 0.425 | 0.387 | 0.238 | 0.151 | 0.146 | 0.213 |
| | 92% | 0.342 | 0.254 | 0.177 | 0.895 | 0.287 | 0.417 | 0.330 | 0.429 | 0.353 | 0.175 | 0.178 | 0.130 |
| | 95% | 0.342 | 0.254 | 0.133 | 0.895 | 0.338 | 0.325 | 0.248 | 0.301 | 0.268 | 0.171 | 0.197 | 0.134 |
| | 97% | 0.148 | 0.128 | 0.097 | 0.842 | 0.141 | 0.160 | 0.103 | 0.099 | 0.109 | 0.137 | 0.137 | 0.121 |
| Simpson Diversity | 80% | 0.004 | 0.000 | 0.000 | 0.105 | 0.177 | 0.069 | 0.186 | 0.082 | 0.082 | 0.157 | 0.116 | 0.102 |
| | 85% | 0.004 | 0.002 | 0.006 | 0.329 | 0.287 | 0.441 | 0.530 | 0.483 | 0.298 | 0.399 | 0.240 | 0.530 |
| | 92% | 0.026 | 0.014 | 0.060 | 0.721 | 0.503 | 0.657 | 0.664 | 0.741 | 0.743 | 0.780 | 0.648 | 0.549 |
| | 95% | 0.026 | 0.014 | 0.060 | 0.721 | 0.506 | 0.658 | 0.664 | 0.745 | 0.767 | 0.850 | 0.702 | 0.606 |
| | 97% | 0.032 | 0.023 | 0.067 | 0.762 | 0.528 | 0.671 | 0.678 | 0.755 | 0.770 | 0.862 | 0.708 | 0.640 |
| Shannon Diversity | 80% | 0.022 | 0.000 | 0.000 | 0.310 | 0.504 | 0.255 | 0.554 | 0.297 | 0.300 | 0.530 | 0.391 | 0.421 |
| | 85% | 0.022 | 0.012 | 0.031 | 0.879 | 0.764 | 1.003 | 1.178 | 1.203 | 0.938 | 1.255 | 0.782 | 1.310 |
| | 92% | 0.114 | 0.071 | 0.251 | 1.912 | 1.234 | 1.707 | 1.774 | 2.082 | 2.239 | 2.823 | 1.863 | 1.609 |
| | 95% | 0.114 | 0.071 | 0.263 | 1.912 | 1.251 | 1.719 | 1.783 | 2.149 | 2.433 | 3.359 | 2.102 | 1.822 |
| | 97% | 0.155 | 0.122 | 0.309 | 2.182 | 1.409 | 1.875 | 1.930 | 2.311 | 2.516 | 3.589 | 2.212 | 1.938 |

**Table A.3.** Alpha diversity calculations of bacteria at OTU percent definitions 80, 85, 92, 95, 97.

| Bacteria | | 0906W | 1002W | 0706W | 0812W | 1007A | 0906A | 1002A | 0812A | 1002B | 1007C | 1002C | 1002E |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Temp (°C) | | 82 | 82 | 80 | 80 | 87 | 82 | 82 | 80 | 79 | 76 | 72 | 62 |
| No. of OTUs | 80% | 9 | 10 | 9 | 13 | 7 | 8 | 10 | 11 | 11 | 15 | 13 | 18 |
| | 85% | 12 | 19 | 19 | 25 | 11 | 18 | 17 | 15 | 23 | 32 | 29 | 53 |
| | 92% | 14 | 24 | 26 | 39 | 13 | 21 | 19 | 21 | 41 | 52 | 52 | 134 |
| | 95% | 15 | 29 | 32 | 43 | 14 | 21 | 19 | 24 | 44 | 62 | 56 | 177 |
| | 97% | 18 | 36 | 37 | 56 | 17 | 23 | 31 | 48 | 68 | 90 | 78 | 229 |
| Chao1 | 80% | 19 | 10 | 9 | 14 | 7 | 11 | 20 | 11 | 12 | 15 | 13 | 19 |
| | 85% | 23 | 34 | 23 | 27 | 13 | 32 | 35 | 16 | 30 | 34 | 30 | 58 |
| | 92% | 26 | 84 | 56 | 90 | 15 | 26 | 30 | 24 | 251 | 61 | 63 | 166 |
| | 95% | 24 | 82 | 59 | 97 | 17 | 26 | 34 | 28 | 136 | 77 | 67 | 235 |
| | 97% | 32 | 124 | 85 | 168 | 21 | 30 | 42 | 78 | 187 | 119 | 111 | 357 |
| Simpson's Evenness | 80% | 0.112 | 0.101 | 0.121 | 0.080 | 0.322 | 0.162 | 0.147 | 0.102 | 0.180 | 0.208 | 0.118 | 0.193 |
| | 85% | 0.084 | 0.053 | 0.057 | 0.042 | 0.224 | 0.073 | 0.087 | 0.081 | 0.098 | 0.198 | 0.142 | 0.125 |
| | 92% | 0.072 | 0.042 | 0.042 | 0.027 | 0.216 | 0.063 | 0.078 | 0.058 | 0.055 | 0.171 | 0.115 | 0.084 |
| | 95% | 0.067 | 0.035 | 0.034 | 0.024 | 0.201 | 0.063 | 0.078 | 0.051 | 0.051 | 0.149 | 0.107 | 0.068 |
| | 97% | 0.056 | 0.028 | 0.030 | 0.019 | 0.171 | 0.058 | 0.048 | 0.025 | 0.034 | 0.121 | 0.078 | 0.057 |
| Simpson Diversity | 80% | 0.007 | 0.008 | 0.084 | 0.039 | 0.557 | 0.228 | 0.320 | 0.107 | 0.495 | 0.680 | 0.348 | 0.713 |
| | 85% | 0.007 | 0.008 | 0.084 | 0.039 | 0.594 | 0.240 | 0.326 | 0.173 | 0.554 | 0.842 | 0.756 | 0.848 |
| | 92% | 0.007 | 0.008 | 0.084 | 0.039 | 0.644 | 0.242 | 0.327 | 0.175 | 0.558 | 0.887 | 0.833 | 0.912 |
| | 95% | 0.007 | 0.008 | 0.084 | 0.039 | 0.645 | 0.242 | 0.327 | 0.176 | 0.558 | 0.892 | 0.833 | 0.916 |
| | 97% | 0.009 | 0.009 | 0.085 | 0.042 | 0.656 | 0.244 | 0.332 | 0.179 | 0.567 | 0.908 | 0.836 | 0.923 |
| Shannon Diversity | 80% | 0.040 | 0.046 | 0.340 | 0.176 | 1.400 | 0.715 | 0.796 | 0.386 | 1.318 | 2.100 | 1.190 | 2.289 |
| | 85% | 0.043 | 0.050 | 0.372 | 0.187 | 1.618 | 0.808 | 0.845 | 0.683 | 1.596 | 3.145 | 2.655 | 3.385 |
| | 92% | 0.043 | 0.051 | 0.357 | 0.194 | 1.861 | 0.825 | 0.848 | 0.680 | 1.678 | 3.775 | 3.413 | 4.664 |
| | 95% | 0.044 | 0.053 | 0.364 | 0.197 | 1.873 | 0.825 | 0.852 | 0.680 | 1.681 | 3.876 | 3.458 | 4.927 |
| | 97% | 0.057 | 0.063 | 0.373 | 0.223 | 1.969 | 0.840 | 0.891 | 0.715 | 1.803 | 4.263 | 3.577 | 5.121 |

**Table A.4.** SIMPER analysis results displaying top ten 97% OTUs responsible for dissimilarity between water and sediment samples of comparable temperature (82 – 80 °C).

| Taxon[b] | Genus | Contrib. (%)[c] | Mean abund.[d] Wat. (%) | Mean abund. A,B (%) |
|---|---|---|---|---|
| *Aquificae* (B) | *Thermocrinis* | 45.44 | 89.47 | 7.57 |
| GAL35 (B) | Unidentified GAL35 | 19.96 | 0.20 | 36.16 |
| *Thermoprotei* (A) | *Aeropyrum* | 9.18 | 0.07 | 16.63 |
| pSL4 (A) | Unidentified pSL4 | 8.81 | 0.04 | 15.91 |
| *Thermoprotei* (A) | Unidentified NAG1 | 5.89 | 0.02 | 10.64 |
| *Thermoprotei* (A) | *Pyrobaculum* | 4.08 | 8.63 | 2.47 |
| *Thaumarchaeota* (A) | *Candidatus* Nitrosocaldus | 2.40 | 0.02 | 4.35 |
| pSL4 (A) | Unidentified pSL4 | 0.56 | 0.00 | 0.99 |
| Unidentified Bacteria (B) | Unidentified Bacteria | 0.44 | 0.00 | 0.80 |
| *Chlorobi* (B) | Unidentified *Chlorobi* | 0.36 | 0.00 | 0.65 |

[a]Pooled water samples compared to pooled sediment samples.

[b]Phylum-level taxa for bacteria and class-level taxa for archaea.

[c]Contribution of OTU to overall dissimilarity between groups.

[d]Average abundance of OTU in each group.

**Table A.5.** Quantification of DNA extracted from non-incubated and incubated substrates

| Substrate and incubation | DNA (ng/g wet weight) | Fold increase over non-incubated [a] |
|---|---|---|
| Non-incubated Aspen | 66 | na[b] |
| 77AW | 840 | 12.7 |
| 77AS | 940 | 14.3 |
| 85AW | 1050 | 15.9 |
| 85AS | 1470 | 22.3 |
| Non-incubated Corn Stover | 1370 | na[b] |
| 77CW | 3420 | 2.5 |
| 77CS | 4550 | 3.3 |
| 85CW | 3050 | 2.2 |
| 85CS | 7760 | 5.7 |

[a] Compared within the given substrate type (aspen or corn stover)

[b] na, not applicable

**Table A.6.** Complete measurements of nutritional data for all samples.

| Label[1] | CP[2] | ADF[3] | NDF[4] | Lignin | ESC[5] | Ash[6] | Cell[7] | Hemicell[8] | NFC[9] | Other[10] |
|---|---|---|---|---|---|---|---|---|---|---|
| 77AS | 1.9 | 77.8 | 91.8 | 13.3 | 0.1 | 9.55 | 64.5 | 14 | na[11] | 0 |
| 77AW | 1.9 | 80.1 | 91.8 | 13 | 0.1 | 3.93 | 67.1 | 11.7 | na[11] | 2.27 |
| 85AS | 1.9 | 75.5 | 85.7 | 14 | 1.3 | 11.47 | 61.5 | 10.2 | na[11] | 0 |
| 85AW | 1.5 | 77.9 | 87.8 | 9.6 | 0.6 | 13.06 | 68.3 | 9.9 | na[11] | 0 |
| UA | 1.7 | 73.5 | 89.3 | 12.9 | 0.9 | 0.75 | 60.6 | 15.8 | na[11] | 7.35 |
| 77CS | 4.7 | 69.4 | 77 | 2.3 | 0 | 13.54 | 67.1 | 7.6 | 0 | na[11] |
| 77CW | 4.6 | 80 | 88.7 | 8.2 | 0.7 | 9.85 | 71.8 | 8.7 | 5 | na[11] |
| 85CS | 3.9 | 57.4 | 67.7 | 2 | 1.4 | 28.62 | 55.4 | 10.3 | 6.2 | na[11] |
| 85CW | 4.8 | 67.9 | 76.4 | 4.9 | 3 | 15.09 | 63 | 8.5 | 0 | na[11] |
| UC | 7 | 49.4 | 81.7 | 10.6 | 0.6 | 4.36 | 38.8 | 32.3 | 9.8 | na[11] |

[1] Label = temperature (77 or 85 °C), substrate (A= aspen shavings, C= corn stover), incubation location (S= sediment, W= water)

[2] Crude protein; in all samples, crude protein was equal to true protein

[3] Acid detergent fiber = cellulose + lignin

[4] Neutral Detergent Fiber = hemicellulose + cellulose + lignin

[5] Ethanol-soluble sugars (mostly mono- and di-saccharides)

[6] Mineral content or non-organic component

[7] Cellulose; calculated from ADF-Lignin

[8] Hemicellulose; calculated from NDF-ADF

[9] Non-fiber carbohydrates; %NFC = 100% - (%CP + (%NDF - %NDICP) + %Fat + %Ash).

[10] Other = components not analyzed, including fat, pectin, organic acids, and soluble fiber.

[11] na: Data not available.

**Table A.7.** Counts of OTUs (≥ 97% identity) in each sample, including taxonomic identification of each OTU. Only OTUs with at least 1% relative abundance in at least one sample are included.

| OTU | UW | U85 | U77 | 77AS | 77AW | 77CS | 77CW | 85AS | 85AW | 85CS | 85CW | Identity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C004 | 0 | 177 | 229 | 1 | 1 | 26 | 4 | 0 | 0 | 0 | 0 | "*Aigarchaeota*" |
| C008 | 0 | 0 | 3 | 0 | 0 | 128 | 5 | 0 | 0 | 0 | 0 | *Thermus* sp. |
| C011 | 0 | 0 | 4 | 4 | 3 | 34 | 8 | 389 | 318 | 1256 | 1392 | *Ignisphaera*-like *Desulfurococcaceae* |
| C035 | 592 | 154 | 113 | 3 | 2 | 14 | 5 | 4 | 4 | 288 | 62 | *Pyrobaculum calidifontis* |
| C036 | 0 | 0 | 6 | 1 | 6 | 63 | 0 | 133 | 587 | 1317 | 373 | "*Aigarchaeota*" |
| C056 | 16 | 4053 | 513 | 0 | 0 | 31 | 19 | 0 | 0 | 244 | 35 | "*Aigarchaeota*" |
| C063 | 1 | 0 | 181 | 0 | 0 | 24 | 1 | 0 | 0 | 0 | 0 | GN03 (WS3) |
| C064 | 0 | 0 | 11 | 0 | 0 | 207 | 2 | 0 | 0 | 0 | 0 | *Rhodothermus* sp. |
| C121 | 0 | 0 | 199 | 210 | 37 | 0 | 0 | 26 | 26 | 0 | 0 | GAL35 |
| C136 | 5 | 2123 | 17 | 0 | 0 | 0 | 1 | 1 | 0 | 36 | 5 | Novel Archaeal Group I |
| C199 | 7 | 2940 | 418 | 0 | 0 | 33 | 2 | 0 | 5 | 408 | 32 | *Aeropyrum* sp. |
| C205 | 0 | 0 | 7 | 20 | 21 | 8 | 6 | 100 | 203 | 602 | 680 | *Desulfurococcus*-like *Desulfurococcaceae* |
| C206 | 0 | 0 | 0 | 33 | 42 | 76 | 51 | 33 | 49 | 28 | 84 | *Thermotoga* sp. |
| C216 | 0 | 0 | 6 | 900 | 621 | 0 | 5 | 53 | 23 | 0 | 0 | pGrfC26 (*Crenarchaeota* C2) |
| C235 | 0 | 1 | 70 | 58 | 20 | 0 | 0 | 3 | 2 | 0 | 0 | NRP-J (*Crenarchaeota*) |
| C236 | 10278 | 887 | 630 | 60 | 33 | 833 | 608 | 19 | 14 | 448 | 120 | *Thermocrinis* sp. |
| C240 | 1 | 18 | 55 | 385 | 287 | 22 | 80 | 396 | 279 | 105 | 128 | Unidentified Bacterium in *Thermodesulfobacteriaceae* |
| C242 | 0 | 0 | 0 | 56 | 105 | 8 | 17 | 5 | 9 | 0 | 0 | *Dictyoglomus* sp. |
| C245 | 1 | 85 | 329 | 0 | 0 | 9 | 4 | 0 | 1 | 1 | 0 | Unidentified Bacterium |
| C249 | 0 | 0 | 17 | 81 | 135 | 0 | 0 | 1 | 2 | 3 | 2 | *Geoglobus*-like *Archaeoglobaceae* |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C253 | 0 | 0 | 16 | 0 | 1 | 2 | 2 | 13 | 6 | 152 | 62 | *Aeropyrum* sp. |
| C301 | 0 | 0 | 0 | 0 | 0 | 8 | 8 | 18 | 92 | 520 | 620 | *Thermosphaera aggregans* |
| C341 | 84 | 13 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Unidentified Bacterium in Aquificaceae |
| C359 | 0 | 0 | 0 | 0 | 3 | 70 | 0 | 758 | 2046 | 2906 | 5026 | *Ignisphaera*-like *Desulfurococcaceae* |
| C369 | 98 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | *Tepidimonas* sp. |
| C422 | 0 | 1 | 127 | 1 | 1 | 1026 | 129 | 0 | 0 | 0 | 0 | Unidentified Bacterium in *Thermomicrobia* |
| C487 | 0 | 0 | 1437 | 1 | 3 | 390 | 42 | 0 | 1 | 0 | 0 | "*Aigarchaeota*" |
| C528 | 0 | 0 | 70 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | LCP-6 (*Thermodesulfovibrionaceae*) |
| C529 | 0 | 0 | 12 | 678 | 147 | 2602 | 2100 | 1618 | 2900 | 1815 | 5633 | *Thermotoga* sp. |
| C543 | 0 | 0 | 152 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | Unidentified Bacterium in GAL15 |
| C589 | 0 | 0 | 548 | 1 | 1 | 521 | 99 | 0 | 0 | 0 | 0 | Unidentified Bacterium in *Chlorobi* |
| C600 | 2 | 4 | 942 | 3 | 4 | 677 | 249 | 0 | 0 | 0 | 0 | *Candidatus* “Nitrosocaldus” sp. |
| C603 | 11 | 3727 | 2016 | 25 | 2 | 1765 | 995 | 13 | 33 | 270 | 103 | GAL35 |
| C628 | 0 | 0 | 10 | 376 | 231 | 0 | 0 | 5 | 75 | 0 | 1 | "*Aigarchaeota*" |
| C666 | 1 | 0 | 27 | 0 | 0 | 185 | 40 | 0 | 0 | 0 | 0 | *Thermus thermophilus* |
| C678 | 0 | 0 | 63 | 23 | 61 | 0 | 2 | 0 | 0 | 0 | 0 | *Fervidobacterium* sp. |
| C692 | 0 | 0 | 43 | 4327 | 3467 | 119 | 1022 | 91 | 184 | 0 | 0 | *Dictyoglomus* sp. |
| C707 | 0 | 12 | 46 | 491 | 573 | 3 | 25 | 280 | 83 | 8 | 9 | Unidentified Bacterium in *Thermodesulfobacteriales* |
| C716 | 0 | 0 | 72 | 118 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | Unidentified Bacterium in *Syntrophobacteraceae* |
| C734 | 0 | 0 | 0 | 1 | 0 | 114 | 86 | 2 | 0 | 11 | 4 | GAL35 |
| C742 | 0 | 0 | 23 | 307 | 306 | 0 | 48 | 0 | 0 | 0 | 0 | *Caldicellulosiruptor* sp. |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C743 | 0 | 0 | 3 | 0 | 1 | 0 | 0 | 77 | 34 | 0 | 8 | Unidentified Archaeon in *Thermoprotei* |
| C745 | 3 | 0 | 940 | 1 | 0 | 604 | 100 | 0 | 0 | 0 | 0 | OS-L (*Armatimonadetes*) |
| C758 | 0 | 0 | 15 | 1997 | 1801 | 179 | 745 | 4 | 0 | 1 | 0 | OPB72 (OP9) |
| C782 | 0 | 0 | 24 | 4613 | 6355 | 454 | 2394 | 15 | 7 | 0 | 0 | *Thermotoga* sp. |
| C790 | 0 | 0 | 12 | 2041 | 1370 | 19 | 342 | 0 | 0 | 0 | 0 | Unidentified Bacterium in *Gemmatimonadetes* |
| C800 | 0 | 0 | 111 | 2 | 2 | 0 | 0 | 2 | 1 | 0 | 0 | Unidentified Bacterium in *Thermodesulfobacteriaceae* |
| C840 | 0 | 0 | 93 | 32 | 17 | 1 | 5 | 30 | 67 | 0 | 0 | OP1 |
| C853 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 5 | 23 | 70 | 189 | *Thermofilum pendens* |
| C859 | 23 | 0 | 551 | 8 | 2 | 1226 | 137 | 0 | 0 | 1 | 0 | *Thermus* sp. |
| C867 | 0 | 0 | 5 | 0 | 0 | 21 | 0 | 142 | 309 | 1590 | 4027 | *Thermofilum pendens* |
| C887 | 0 | 1 | 0 | 4 | 2 | 2 | 0 | 124 | 43 | 6 | 16 | Unidentified Archaeon in DHVE3 |
| C890 | 0 | 0 | 6 | 546 | 355 | 2 | 43 | 25 | 46 | 0 | 0 | Unidentified Bacterium in EM3 |
| C896 | 0 | 0 | 6 | 154 | 177 | 0 | 0 | 0 | 0 | 0 | 0 | Unidentified Archaeon in DHVE3 |
| C898 | 0 | 0 | 216 | 0 | 0 | 86 | 8 | 0 | 0 | 0 | 0 | TK17 (*Chloroflexi*) |
| C903 | 0 | 0 | 157 | 2282 | 2025 | 222 | 937 | 306 | 594 | 556 | 2429 | *Archaeoglobus* sp. |

**Table A.8.** PERMANOVA statistics indicating the significance of the difference between samples indicated by tree nodes and other logical differences between sample groups.

| Variation | F[a] | p-value |
|---|---|---|
| Node 2: Natural sediment and 77CS vs. Enrichment samples | 4.375 | 0.0084 ** |
| Site: 85 vs. 77, incl. natural sediment samples. | 3.475 | 0.0152 * |
| Enriched :Enrichment vs. natural samples | 3.606 | 0.0199 * |
| Node 3: 77 (exc. 77CS) vs. 85, enrichment samples only | 15.63 | 0.0275 * |
| Node 1: UW vs. all others | 2.077 | 0.0892 ~ |
| Enrichment Corn stover enrichments vs. aspen enrichments) | 1.454 | 0.1765 |
| Node 7: Site 85 aspen vs. Site 85 corn stover | 3.525 | 0.3245 |
| Node 6: Site 77 aspen vs. 77CW | 9.623 | 0.3354 |
| Node 4: U85 vs. U77 and 77CS | 2.321 | 0.3413 |
| Location: Sediment enrichments vs. water enrichments | 0.2102 | 0.8306 |

** Significant at $\alpha = 0.01$

* Significant at $\alpha = 0.05$

~ Significant at $\alpha = 0.1$

[a] PERMANOVA F-Statistic

**Table A.9.** SIMPER results for comparison of samples segregated at tree node #2, including U85, U77, and 77CS in one group and all other enrichment samples in the other group. Only OTUs contributing at least 1% of the difference of the community compositions are included.

| OTU | Identity | Δ[a] | Contrib. (%)[b] | Avg. Rep. (%) U85, U87, 77CS [c] | Avg. Rep. (%) Enr. Exc. 77CS [d] |
|---|---|---|---|---|---|
| C603 | GAL35 | - | 10.31 | 19.64 | 1.82 |
| C529 | *Thermotoga* sp. | + | 9.428 | 7.08 | 18.98 |
| C359 | *Ignisphaera*-like *Desulfurococcaceae* | + | 7.087 | 0.25 | 12.29 |
| C782 | *Thermotoga* sp. | + | 6.629 | 1.38 | 11.34 |
| C056 | *"Aigarchaeota"* | - | 6.381 | 11.17 | 0.29 |
| C199 | *Aeropyrum* sp. | - | 4.407 | 7.83 | 0.55 |
| C903 | *Archaeoglobus* sp. | + | 4.367 | 1.01 | 8.55 |
| C692 | *Dictyoglomus* sp. | + | 4.176 | 0.43 | 7.39 |
| C867 | *Thermofilum pendens* | + | 3.165 | 0.10 | 5.49 |
| C487 | *"Aigarchaeota"* | - | 3.07 | 5.33 | 0.05 |
| C859 | *Thermus* sp. | - | 2.953 | 5.16 | 0.18 |
| C136 | Novel Archaeal Group I | - | 2.743 | 4.73 | 0.06 |
| C236 | *Thermocrinis* sp. | - | 2.668 | 6.15 | 1.61 |
| C600 | *Candidatus* "Nitrosocaldus" sp. | - | 2.592 | 4.61 | 0.36 |
| C745 | OS-L (*Armatimonadetes*) | - | 2.557 | 4.46 | 0.13 |
| C011 | *Ignisphaera*-like *Desulfurococcaceae* | + | 2.302 | 0.09 | 3.99 |
| C758 | OPB72 (OP9) | + | 2.093 | 0.52 | 3.57 |
| C422 | Unidentified Bacterium in *Thermomicrobia* | - | 1.857 | 3.24 | 0.18 |
| C036 | *"Aigarchaeota"* | + | 1.803 | 0.19 | 3.16 |
| C589 | Unidentified Bacterium in *Chlorobi* | - | 1.713 | 3.02 | 0.14 |
| C790 | Unidentified Bacterium in *Gemmatimonadetes* | + | 1.708 | 0.08 | 2.93 |
| C240 | Unidentified Bacterium in *Thermodesulfobacteriaceae* | + | 1.309 | 0.23 | 2.50 |
| C707 | Unidentified Bacterium in *Thermodesulfobacteriales* | + | 1.021 | 0.17 | 1.86 |

[a] Difference between natural sediment and enrichment populations. + OTU has greater representation in enrichment samples (exc. 77CS). - OTU has lower representation in enrichment samples (exc. 77CS).

[b] Percent contribution to community composition difference
[c] Average percent representation in natural sediment sample communities and 77CS
[d] Average percent representation in enrichment sample communities exc. 77CS

**Table A.10.** SIMPER results for comparison of samples from Site 85 to samples from Site 77, including natural sediment community samples at each site. Only OTUs contributing at least 1% of the difference of the community compositions are included.

| OTU | Identity | Δ[a] | Contrib. (%)[b] | Avg. Rep. (%) Site 85 [c] | Avg. Rep. (%) Site 77 [d] |
|---|---|---|---|---|---|
| C529 | *Thermotoga* sp. | - | 10.56 | 21.69 | 9.05 |
| C359 | *Ignisphaera*-like *Desulfurococcaceae* | - | 10.44 | 17.20 | 0.15 |
| C782 | *Thermotoga* sp. | + | 10.08 | 0.09 | 16.60 |
| C603 | GAL35 | + | 6.205 | 5.89 | 8.47 |
| C692 | *Dictyoglomus* sp. | + | 5.768 | 0.80 | 9.83 |
| C867 | *Thermofilum pendens* | - | 4.662 | 7.68 | 0.06 |
| C056 | "*Aigarchaeota*" | - | 3.958 | 6.07 | 1.03 |
| C011 | *Ignisphaera*-like *Desulfurococcaceae* | - | 3.385 | 5.58 | 0.06 |
| C758 | OPB72 (OP9) | + | 3.208 | 0.02 | 5.29 |
| C903 | *Archaeoglobus* sp. | + | 2.97 | 5.89 | 6.69 |
| C199 | *Aeropyrum* sp. | - | 2.947 | 4.71 | 0.75 |
| C036 | "*Aigarchaeota*" | - | 2.641 | 4.40 | 0.12 |
| C790 | Unidentified Bacterium in *Gemmatimonadetes* | + | 2.54 | 0.00 | 4.17 |
| C859 | *Thermus* sp. | + | 2.044 | 0.00 | 3.35 |
| C236 | *Thermocrinis* sp. | + | 1.989 | 2.09 | 3.86 |
| C487 | "*Aigarchaeota*" | + | 1.988 | 0.00 | 3.26 |
| C600 | *Candidatus* "Nitrosocaldus" sp. | + | 1.986 | 0.01 | 3.26 |
| C136 | Novel Archaeal Group I | - | 1.771 | 2.91 | 0.03 |
| C745 | OS-L (*Armatimonadetes*) | + | 1.741 | 0.00 | 2.85 |
| C240 | Unidentified Bacterium in *Thermodesulfobacteriaceae* | - | 1.46 | 2.64 | 0.99 |
| C205 | *Desulfurococcus*-like *Desulfurococcaceae* | - | 1.422 | 2.38 | 0.09 |
| C422 | Unidentified Bacterium in *Thermomicrobia* | + | 1.338 | 0.00 | 2.19 |
| C589 | Unidentified Bacterium in *Chlorobi* | + | 1.222 | 0.00 | 2.00 |
| C707 | Unidentified Bacterium in *Thermodesulfobacteriales* | - | 1.143 | 1.44 | 1.26 |
| C301 | *Thermosphaera aggregans* | - | 1.019 | 1.69 | 0.02 |

[a] Difference between populations at different sites. + OTU has greater representation in Site 77 samples. - OTU has greater representation in Site 85 samples.

[b] Percent contribution to community composition difference

[c] Average percent representation in Site 85 samples
[d] Average percent representation in Site 77 samples

**Table A.11.** SIMPER results for comparison of natural sediment communities, U85 and U77, to all enrichment samples. Only OTUs contributing at least 1% of the difference of the community compositions are included.

| OTU | *Identity* | Δ[a] | Contrib. (%)[b] | Avg. Env. Rep. (%)[c] | Avg. Enr. Rep. (%)[d] |
|---|---|---|---|---|---|
| C529 | *Thermotoga* sp. | + | 10.76 | 0.04 | 19.25 |
| C603 | GAL35 | - | 10.36 | 21.89 | 3.47 |
| C056 | *"Aigarchaeota"* | - | 9.18 | 16.67 | 0.28 |
| C199 | *Aeropyrum* sp. | - | 6.22 | 11.61 | 0.51 |
| C359 | *Ignisphaera*-like *Desulfurococcaceae* | + | 6.08 | 0.00 | 10.84 |
| C782 | *Thermotoga* sp. | + | 5.81 | 0.10 | 10.41 |
| C136 | Novel Archaeal Group I | - | 3.97 | 7.10 | 0.06 |
| C903 | *Archaeoglobus* sp. | + | 3.95 | 0.65 | 7.70 |
| C692 | *Dictyoglomus* sp. | + | 3.66 | 0.13 | 6.61 |
| C487 | *"Aigarchaeota"* | - | 3.62 | 6.46 | 0.42 |
| C867 | *Thermofilum pendens* | + | 2.71 | 0.02 | 4.83 |
| C745 | OS-L (*Armatimonadetes*) | - | 2.45 | 4.38 | 0.69 |
| C600 | *Candidatus* "Nitrosocaldus" sp. | - | 2.37 | 4.25 | 0.98 |
| C236 | *Thermocrinis* sp. | - | 2.17 | 5.78 | 2.27 |
| C011 | *Ignisphaera*-like *Desulfurococcaceae* | + | 1.98 | 0.00 | 3.53 |
| C758 | OPB72 (OP9) | + | 1.84 | 0.09 | 3.28 |
| C859 | *Thermus* sp. | - | 1.77 | 2.44 | 1.49 |
| C036 | *"Aigarchaeota"* | + | 1.58 | 0.02 | 2.83 |
| C790 | Unidentified Bacterium in *Gemmatimonadetes* | + | 1.45 | 0.06 | 2.58 |
| C589 | Unidentified Bacterium in *Chlorobi* | - | 1.34 | 2.40 | 0.65 |
| C240 | Unidentified Bacterium in *Thermodesulfobacteriaceae* | + | 1.10 | 0.28 | 2.21 |
| C245 | Unidentified Bacterium | - | 1.03 | 1.85 | 0.02 |

[a] Difference between natural sediment and enrichment populations. + OTU has greater representation in enrichment samples. - OTU has lower representation in enrichment samples.

[b] Percent contribution to community composition difference

[c] Average percent representation in natural sediment sample communities

[d] Average percent representation in enrichment sample communities

**Table A.12.** SIMPER results for comparison of samples segregated at tree node #3, including 77AS, 77AW, and 77CW in one group and all Site 85 enrichment samples in the other group. Only OTUs contributing at least 1% of the difference of the community compositions are included.

| OTU | Identity | Δ[a] | Contrib. (%)[b] | Avg. Rep. (%) Site 85 [c] | Avg. Rep. (%) Site 77 [d] |
|---|---|---|---|---|---|
| C782 | *Thermotoga* sp. | + | 16.37 | 0.11 | 26.23 |
| C359 | *Ignisphaera*-like *Desulfurococcaceae* | - | 13.43 | 21.48 | 0.01 |
| C529 | *Thermotoga* sp. | - | 12.58 | 27.26 | 8.01 |
| C692 | *Dictyoglomus* sp. | + | 9.34 | 1.00 | 15.94 |
| C867 | *Thermofilum pendens* | - | 6.00 | 9.60 | 0.00 |
| C758 | OPB72 (OP9) | + | 5.16 | 0.03 | 8.28 |
| C011 | *Ignisphaera*-like *Desulfurococcaceae* | - | 4.36 | 6.98 | 0.01 |
| C790 | Unidentified Bacterium in *Gemmatimonadetes* | + | 4.29 | 0.00 | 6.86 |
| C036 | *"Aigarchaeota"* | - | 3.43 | 5.51 | 0.01 |
| C903 | *Archaeoglobus* sp. | + | 2.16 | 7.35 | 10.16 |
| C603 | GAL35 | + | 2.12 | 0.77 | 3.22 |
| C205 | *Desulfurococcus*-like *Desulfurococcaceae* | - | 1.80 | 2.97 | 0.10 |
| C240 | Unidentified Bacterium in *Thermodesulfobacteriaceae* | - | 1.64 | 3.28 | 1.45 |
| C216 | pGrfC26 (*Crenarchaeota* C2) | + | 1.50 | 0.34 | 2.54 |
| C236 | *Thermocrinis* sp. | + | 1.43 | 1.17 | 2.19 |
| C707 | Unidentified Bacterium in *Thermodesulfobacteriales* | + | 1.40 | 1.78 | 1.96 |
| C301 | *Thermosphaera aggregans* | - | 1.30 | 2.11 | 0.03 |

[a] Difference between populations at different sites. + OTU has greater representation in Site 77 samples. - OTU has greater representation in Site 85 samples.

[b] Percent contribution to community composition difference

[c] Average percent representation in Site 85 samples

[d] Average percent representation in Site 77 samples

APPENDIX B: ADDITIONAL FIGURES

**Figure B.1.** Photograph of GBS with sampling sites identified. Temperature range displayed at each site represents the temperatures recorded during sample collection. Temperature of sediment samples measured in water above each sample site.

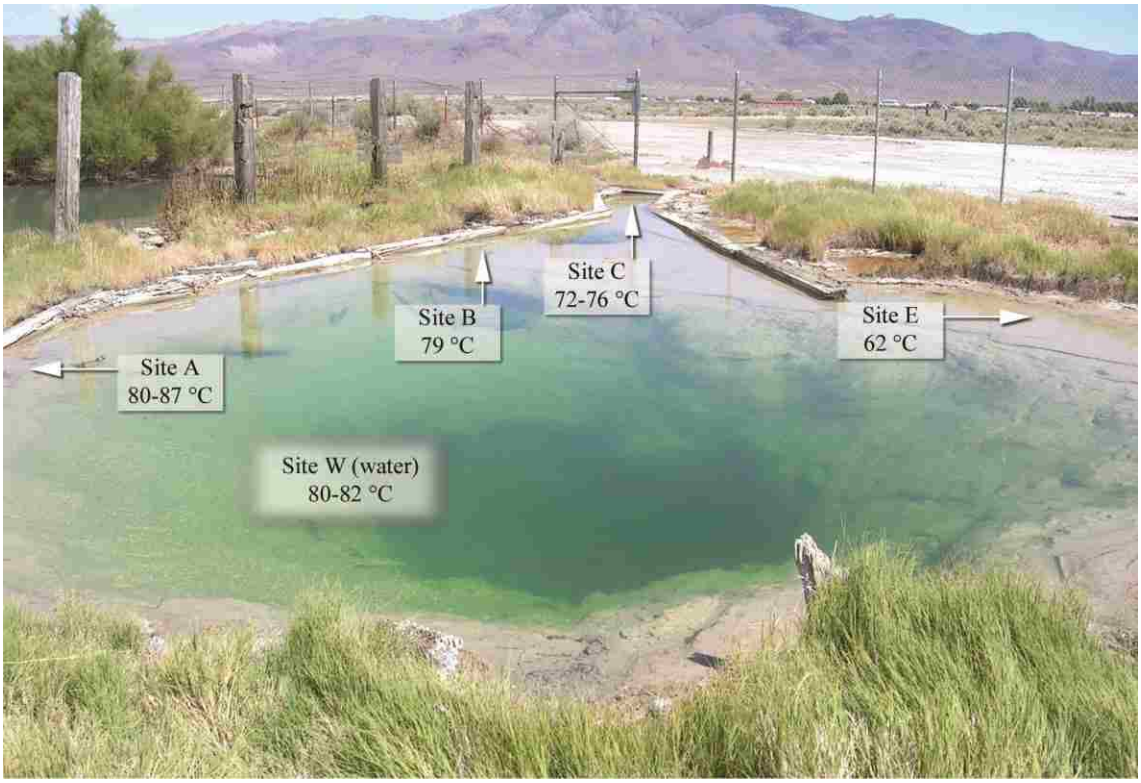


**Figure B.2.** Collector's (a,b) and rarefaction (c,d) curves of observed 97% OTUs in all samples.

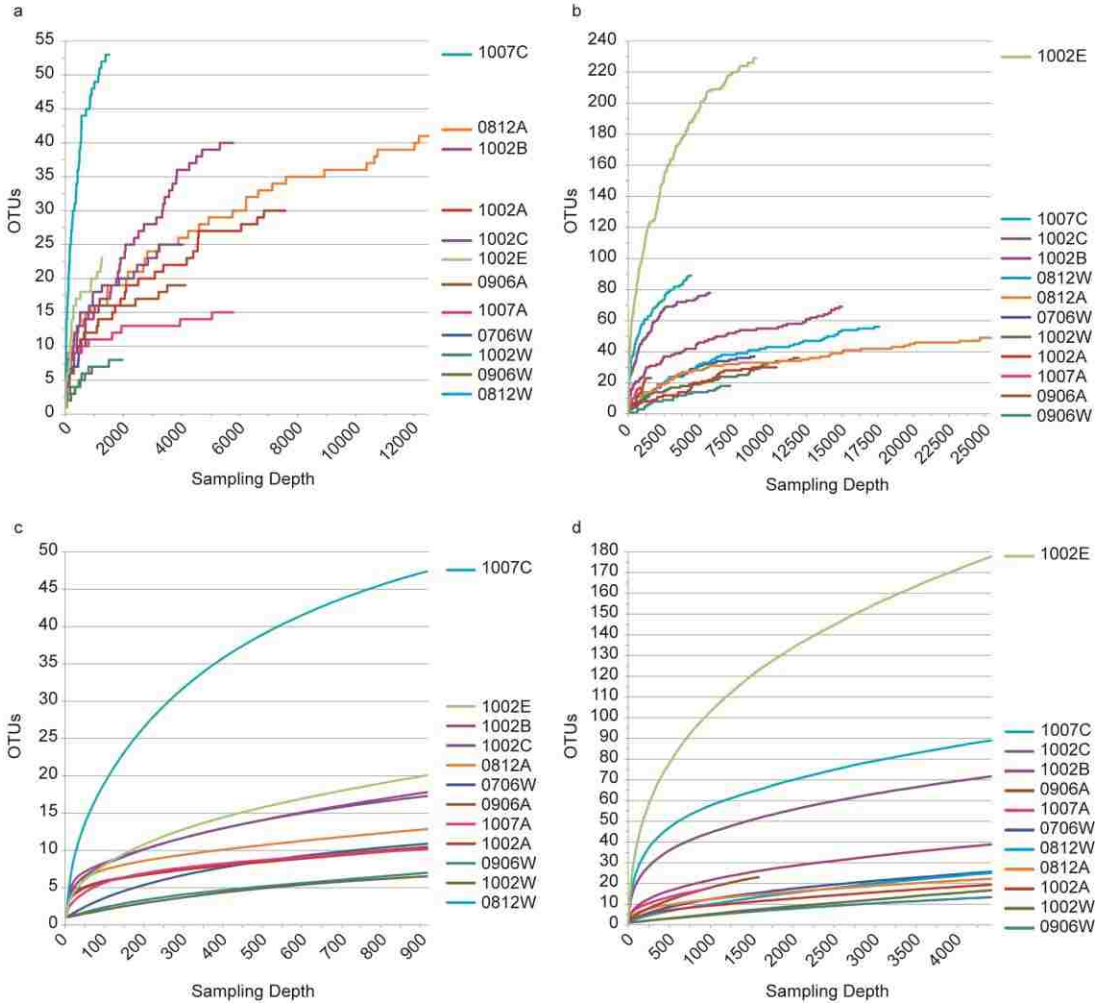
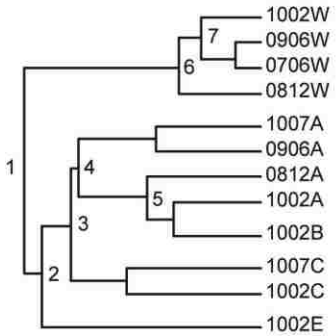


**Figure B.3.** ANOSIM analysis of sample divisions. (a) Cluster tree derived from 97% OTUs, with nodes labeled for reference. (b) ANOSIM results for all cluster tree nodes containing more than two samples. ANOSIM results also included for sample groups not represented by single nodes. ***, significant at α=0.001; **, significant at α=0.01; *, significant at α=0.05; ~, significant at α=0.1.

**a**




**b**

Results of ANOSIM analysis between groups of samples.

| Node ID | Groups[1] | ANOSIM R | ANOSIM p | Uncorrected Significance | Corrected Significance[2] |
|---|---|---|---|---|---|
| 1 | W-ABCE | 0.80 | 0.0022 | ** | * |
| 2 | ABC-E | 0.96 | 0.1259 | | |
| 3 | AB-C | 0.58 | 0.0978 | ~ | |
| 4 | AA-AAB | 0.75 | 0.0986 | ~ | |
| 5 | A-AB | 1 | 0.3258 | | |
| 6 | WWW-W | 1 | 0.246 | | |
| 7 | W-WW | 0 | 0.667 | | |
| N/A | AB-C-E | 0.75 | 0.011 | * | ~ |
| N/A | AB-W | 0.96 | 0.0085 | ** | ~ |

[1] Samples as grouped for ANOSIM calculation with groups separated by dash.
[2] Significance corrected for multiple tests.

**Figure B.4.** Scatter plots of sediment sample temperature versus alpha diversity measurement (richness or evenness) for the two domains surveyed (Archaea or Bacteria), with regression lines for each OTU level. (a) Archaeal richness vs. temperature, (b) Bacterial richness vs. temperature, (c) Archaeal evenness vs. temperature, (d) Bacterial evenness vs. temperature.

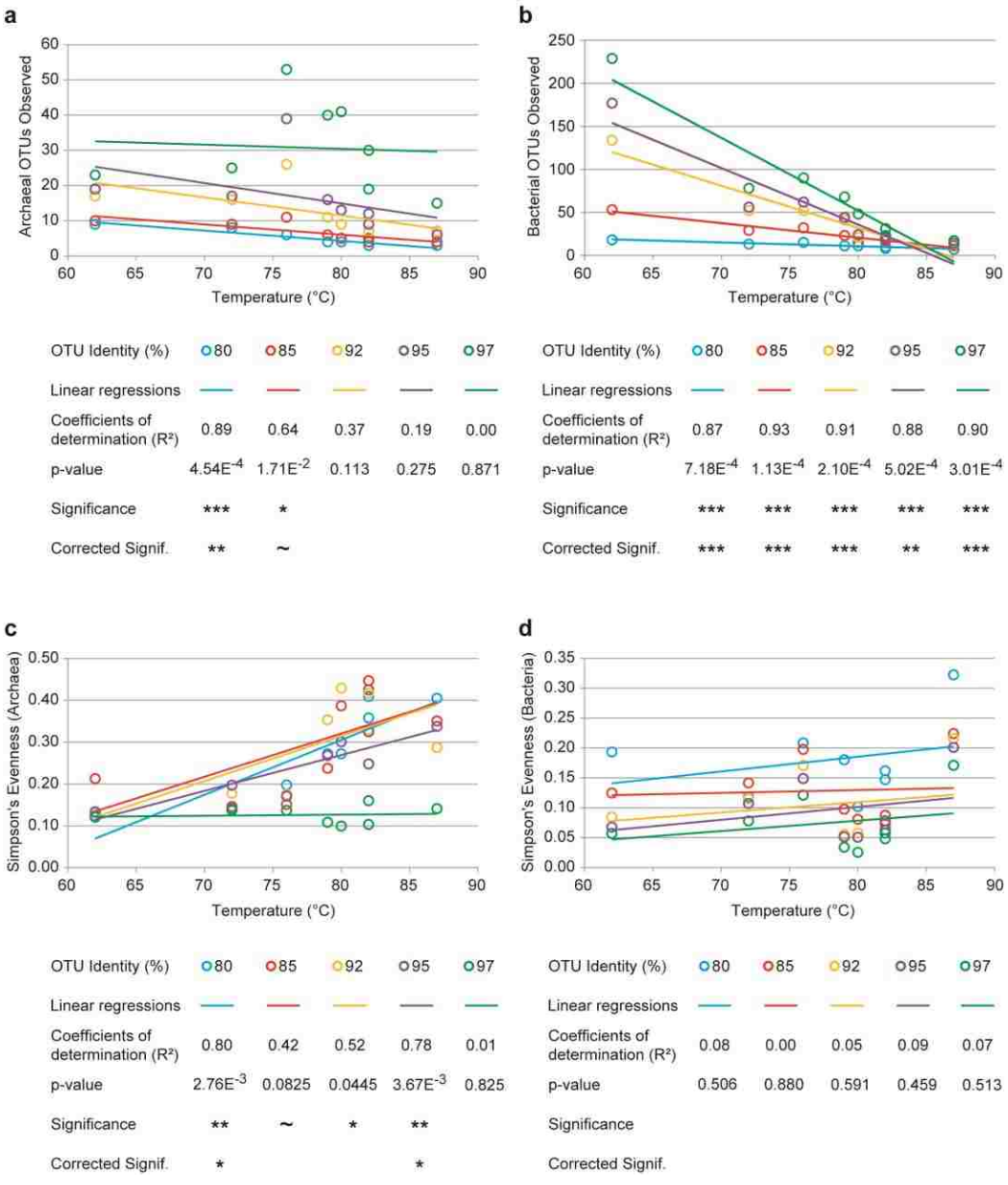


**Figure B.5.** Scatter plots with linear regressions of sediment sample temperature vs oxygen.

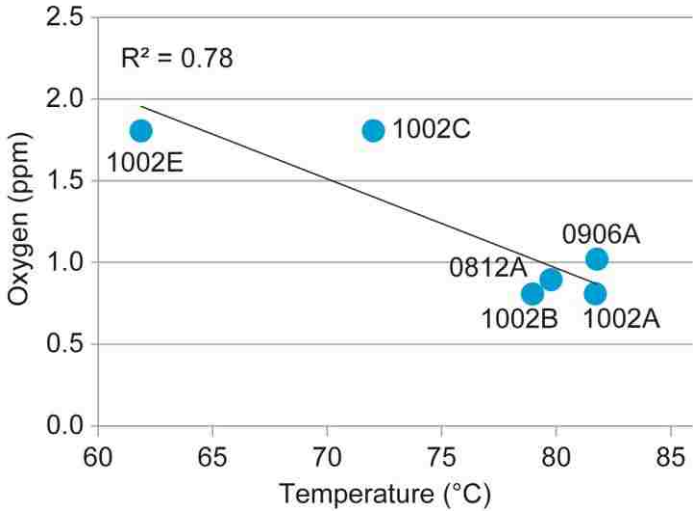


**Figure B.6.** Bar graphs highlighting relative abundance of significant (relative abundance of greater than or equal to 1% in at least one sample) 97% OTUs in selected taxa. (a) Distribution of three OTUs within the *Armatimonadetes*. (b) Distribution of three OTUs within the uncultured archaeal class, pSL4. (c) Distribution of three OTUs within *Thermoprotei.*

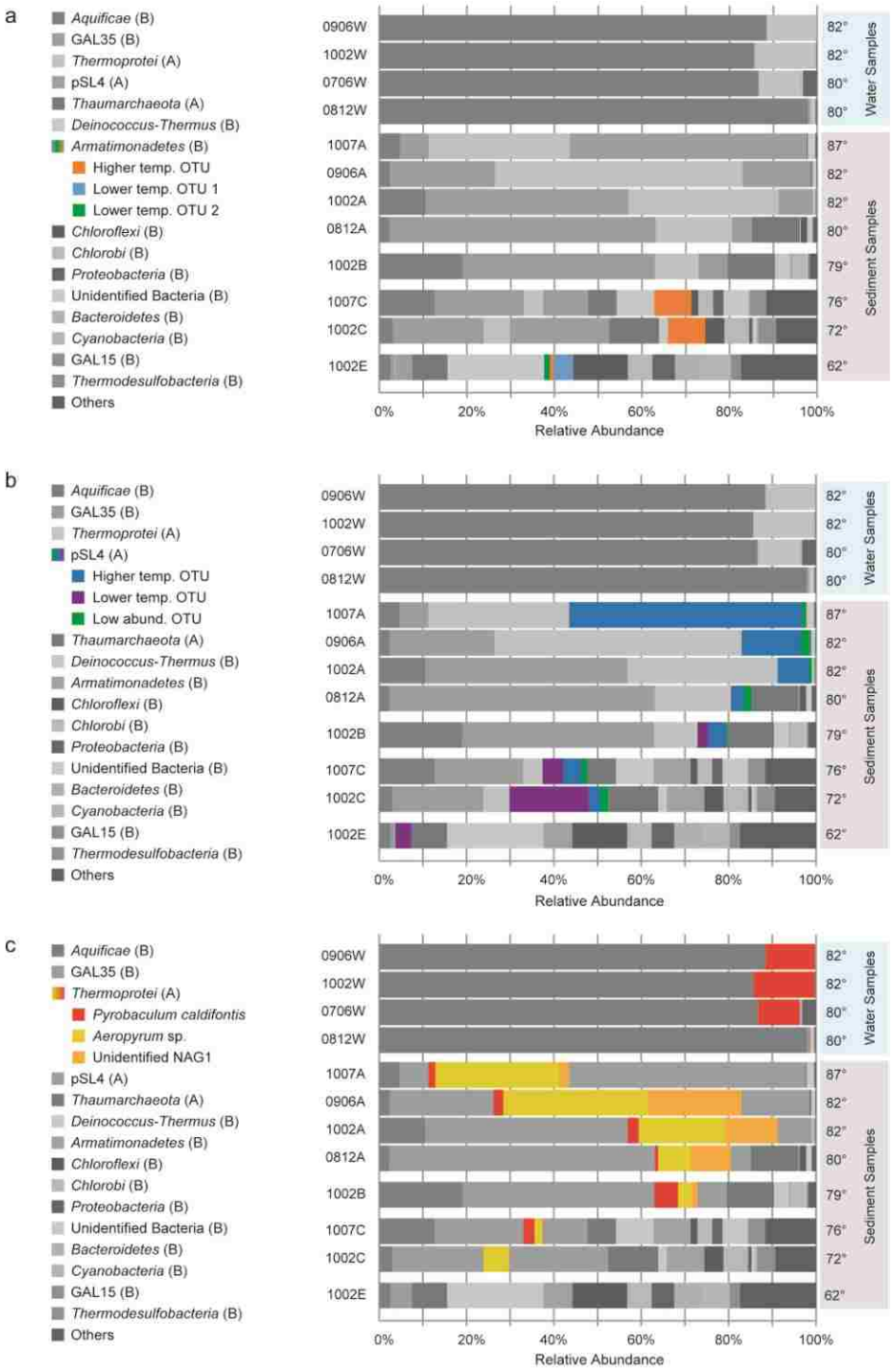


**Figure B.7.** Scatter plots of the relative abundances of the top ten OTUs identified by SIMPER analysis as contributing the most dissimilarity (in High vs Med vs Low temperature sediment samples) vs. sample temperature (See Table 2.3). Analysis was performed on the complete microbial community. Archaeal and bacterial OTUs are displayed separately for clarity. (a) OTUs in the domain Archaea, (b) OTUs in the domain Bacteria.

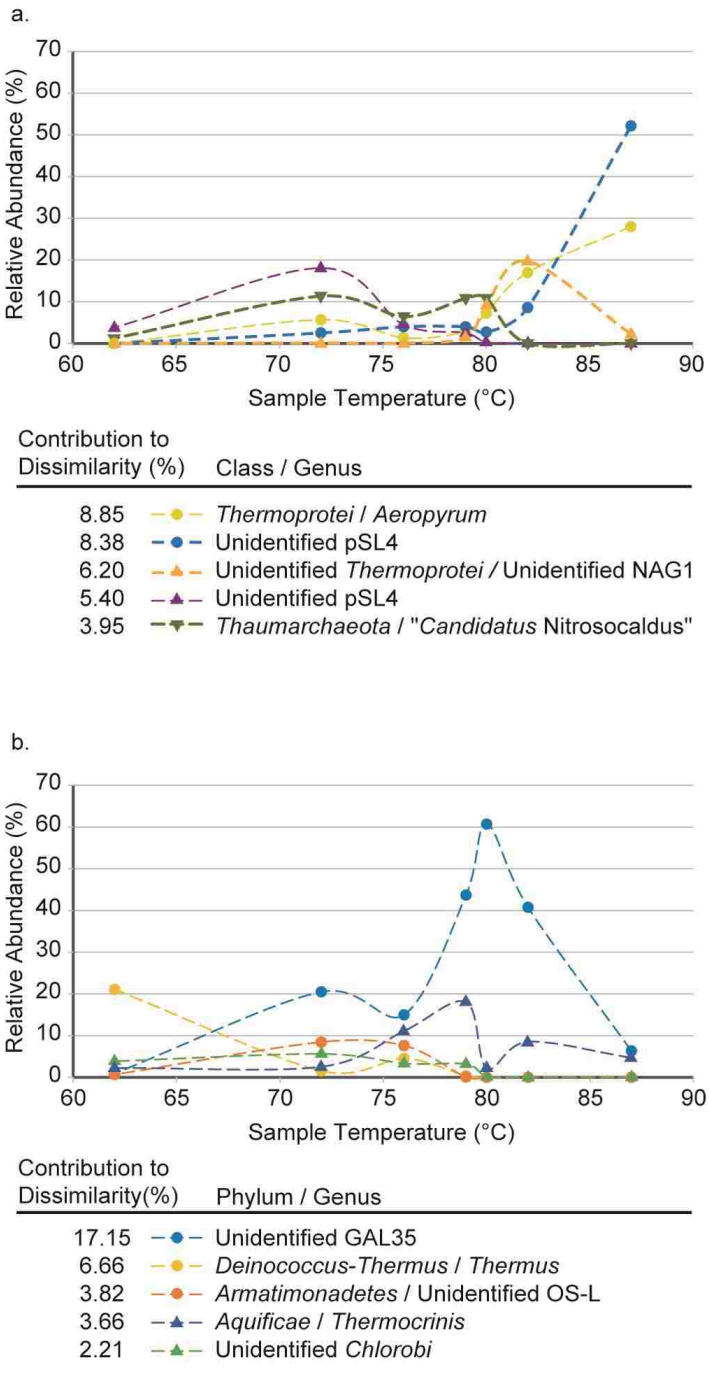
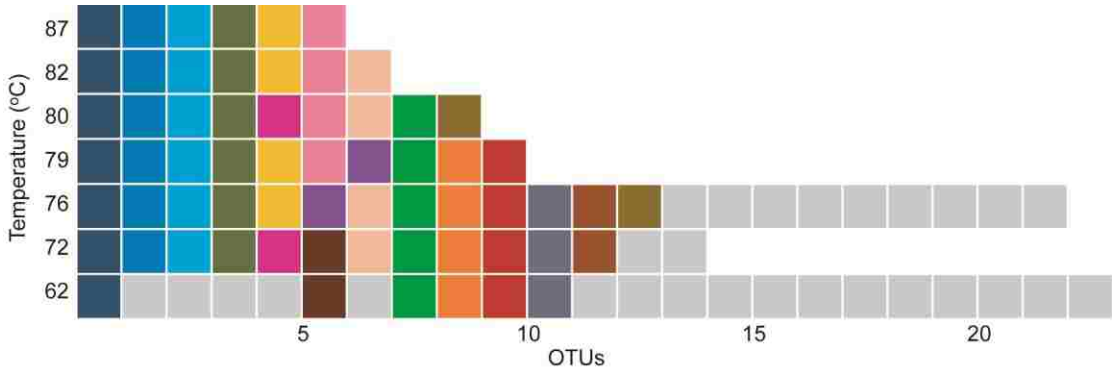


**Figure B.8.** Histogram displaying 97% OTUs present at ≥1% relative abundance in each sediment sample. Samples with identical temperature are combined, averaging relative abundances of each OTU. Grey blocks represent OTUs present above the 1% threshold in only one sample.




*Thermocrinis* sp. (B)
*Aeropyrum* sp. (A)
Unidentified GAL35 (B)
Unidentified pSL4 (A)
"*Candidatus* Nitrosocaldus yellowstonii" (A)
*Pyrobaculum calidifontis* (A)
Unidentified *Chlorobi* (B)
Unidentified pSL4 (A)
Unidentified pSL4 (A)
Unidentified NAG1 (A)
*Thermus scotoductus* (B)
*Thermus thermophilus* (B)
Unidentified Bacteria (B)
Unidentified GAL15 (B)
Unidentified OS-L (B)
Unidentified TK17 (B)

Any OTU present at only one temperature in >=1% relative abundance, including:

**76 °C**
*Archaeoglobus fulgidus* (A)
*Dictyoglomus turgidum* (B)
*Thermocrinis* sp. (B)
*Thermus* sp. (B)
Unidentified GAL35 (B)
Unidentified OP1 (B)
Unidentified *Syntrophobacteraceae* (B)
Unidentified *Thermodesulfobacteriaceae* (B)
Unidentified *Thermomicrobia* (B)

**72 °C**
Unidentified GN03 (B)
Unidentified O1aA90 (B)

**62 °C:**
*Caldilinea* sp. (B)
"*Candidatus* Nitrosocaldus" sp. (A)
*Chloroflexus* sp. (B)
LCP-6 sp. (B)
*Rhodothermus* sp. (B)
*Synechococcus* sp. (B)
*Synechococcus* sp. (B)
Unidentified *Anaerolineae* (B)
Unidentified CFB-26 (B)
Unidentified CH21 (B)
Unidentified *Gemmatimonadetes* (B)
Unidentified OPB92 (B)
Unidentified OS-K (B)
Unidentified OS-L (B)
Unidentified *Rhodospirillaceae* (B)
Unidentified SM1H02 (B)
Unidentified *Syntrophobacteraceae* (B)

**Figure B.9.** Bar chart displaying the community composition of each sample at the phylum/class level, including the 15 most abundant taxa of all samples with remaining taxa included as "Others". Novel taxa are in color and described taxa are in grayscale. "(A)" or "(B)" after taxon name in legend designates an archaeal or bacterial taxon, respectively.

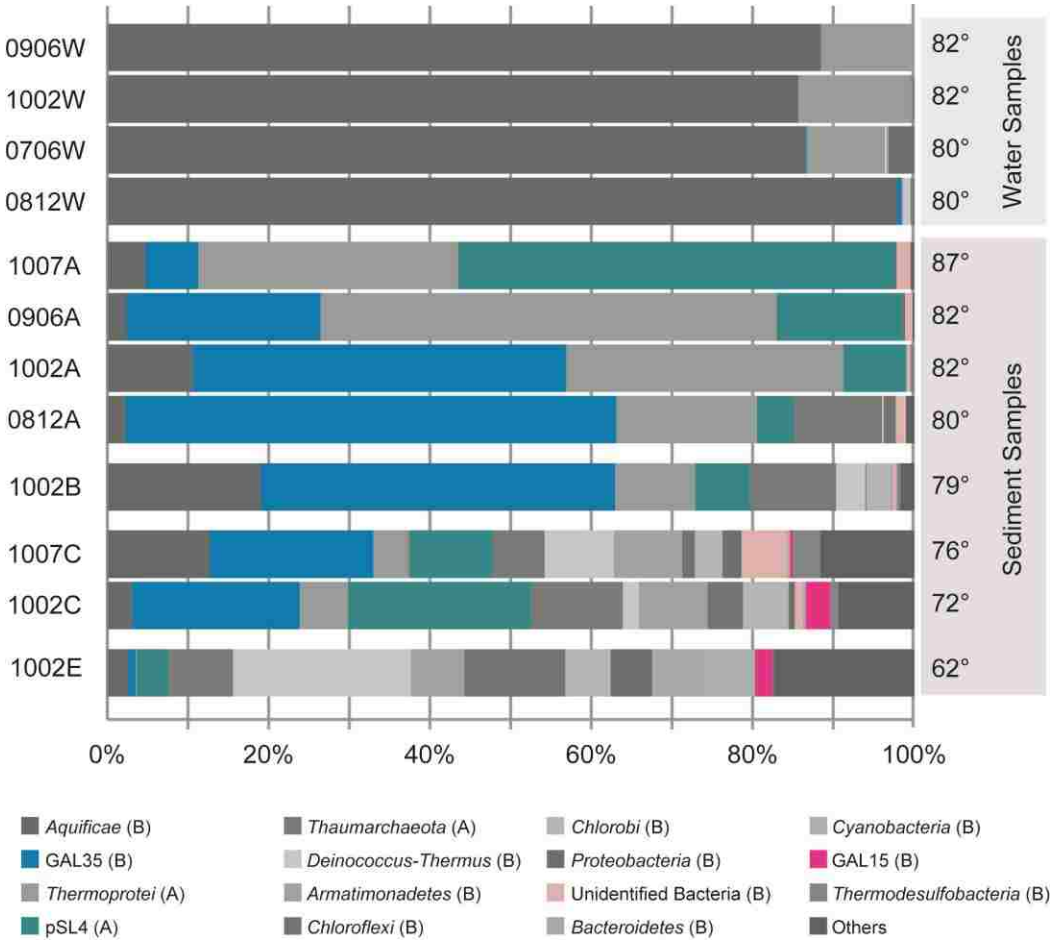


**Figure B.10.** Bar chart displaying the community composition of each sample at the genus level, including genera that represented at least 1% relative abundance of at least one sample, with remaining taxa included as "Others". Novel taxa are in color and described taxa are in grayscale. “(A)” or “(B)” after taxon name in legend designates an archaeal or bacterial taxon, respectively.

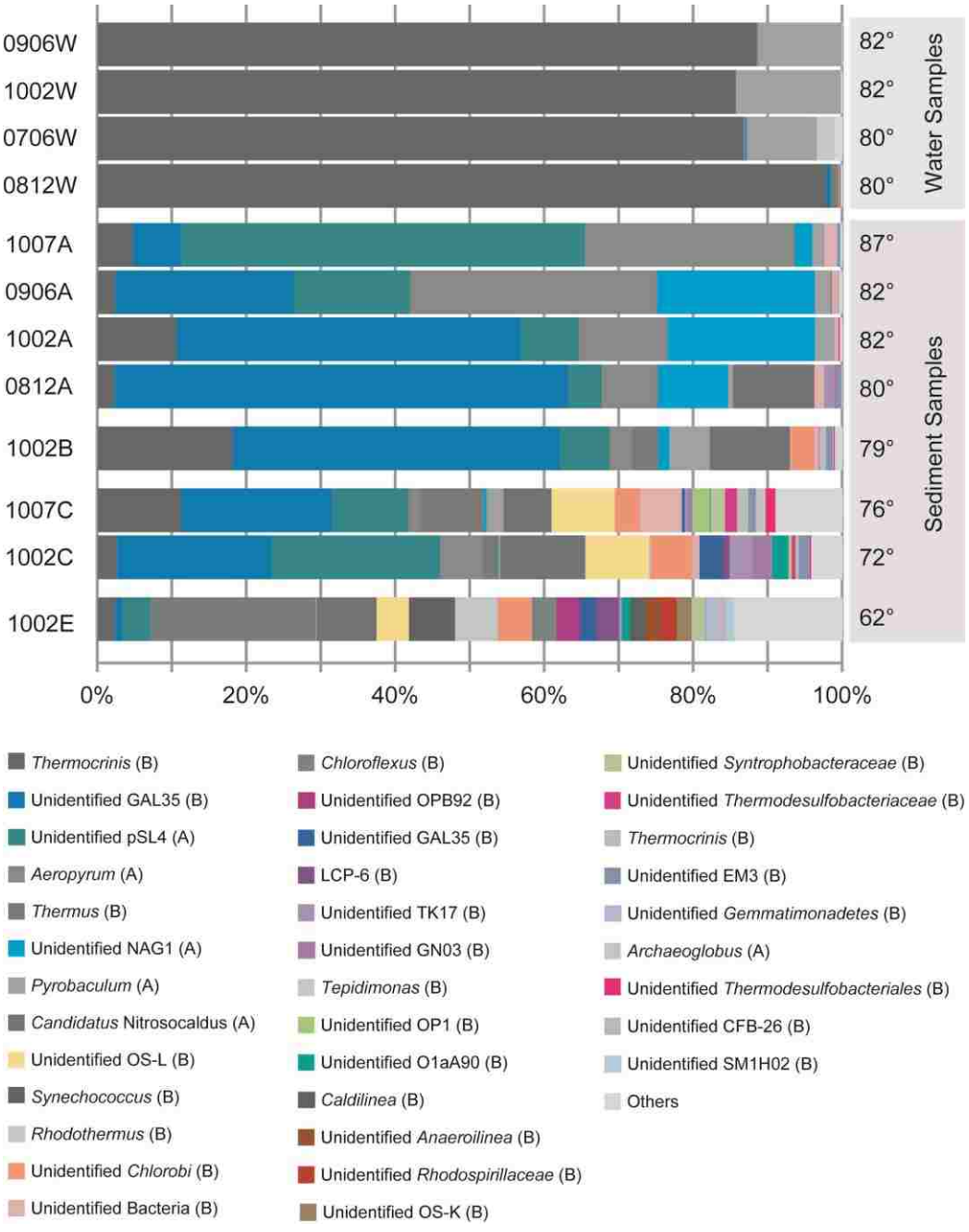


**Figure B.11.** Photograph of GBS with sample incubation sites, Site 85 and Site 77, indicated.

**Figure B.12.** Log of iButton temperature loggers at sample incubation sites. Incubation of substrates began on 29 August, 2009, with temperatures at each incubation site indicated by green dots. Temperature-logging iButtons were deployed in the water column at incubation sites on 15 September, 2009. Temperatures were recorded every two hours from this date until the end of substrate incubation at each site (red diamonds). Logged temperature readings at each site are graphed with solid, colored lines. Black lines indicate the average temperatures of each site over the time logged.

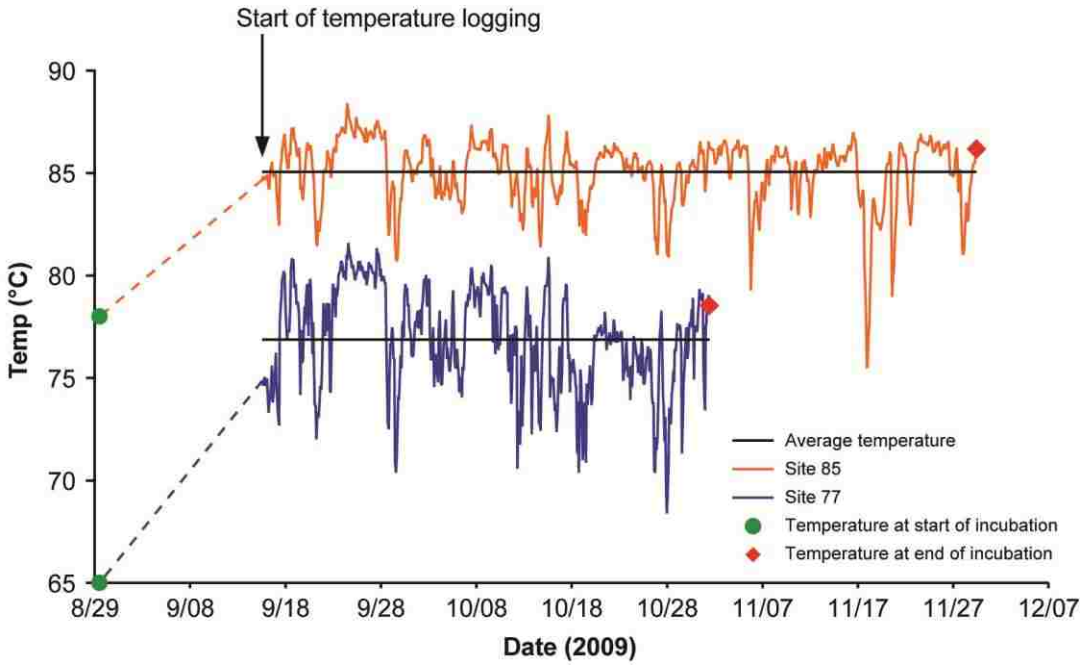


**Figure B.13.** Alpha diversity measurements, comparing natural and enrichment samples at 97%, 95%, 92%, 85%, and 80% OTU minimum identity levels. (a) OTUs observed in each sample. (b) Simpson's index of evenness in each sample. Note U85 and U77 points are overlapping (same evenness scores) for 97% and 95% OTU identities.

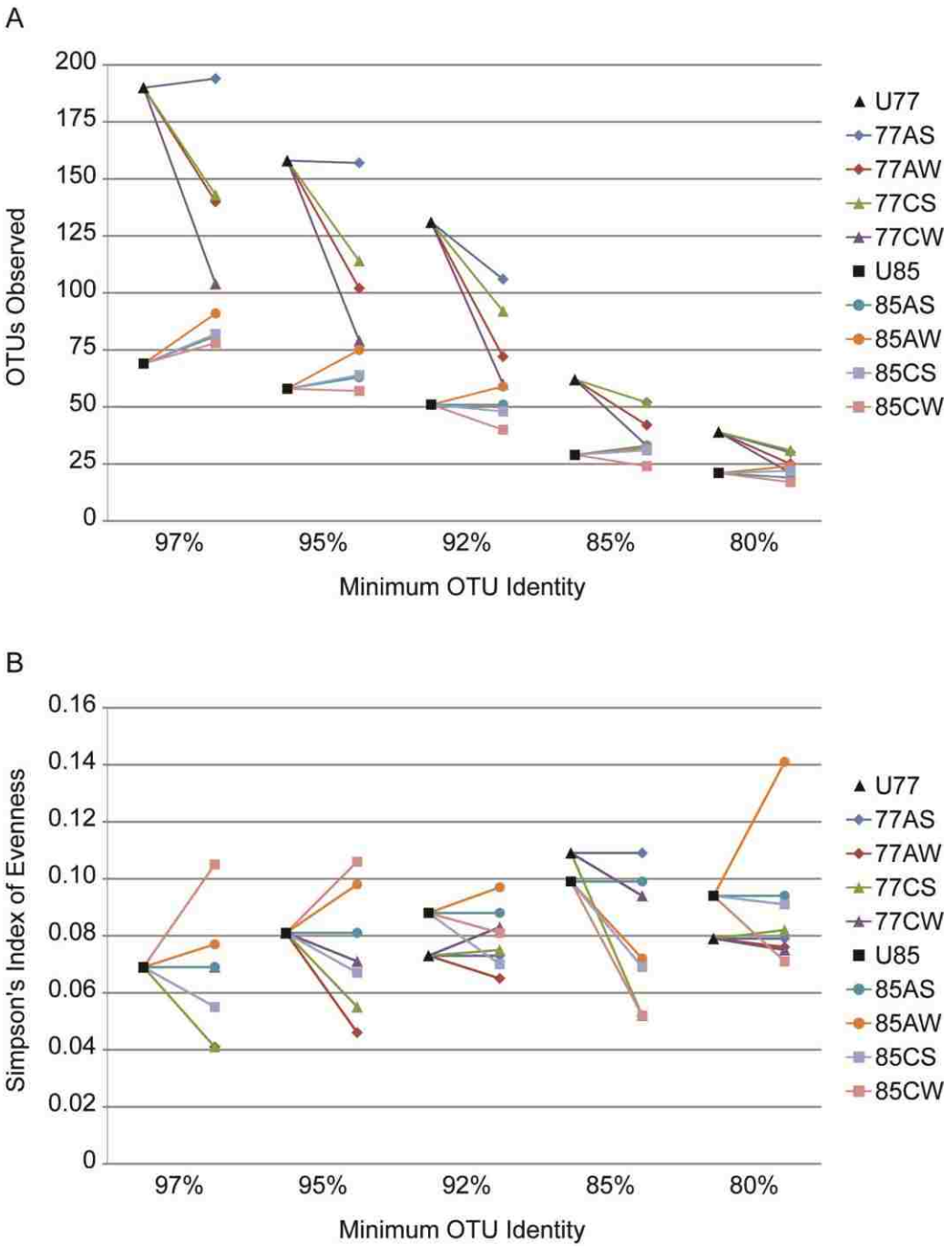
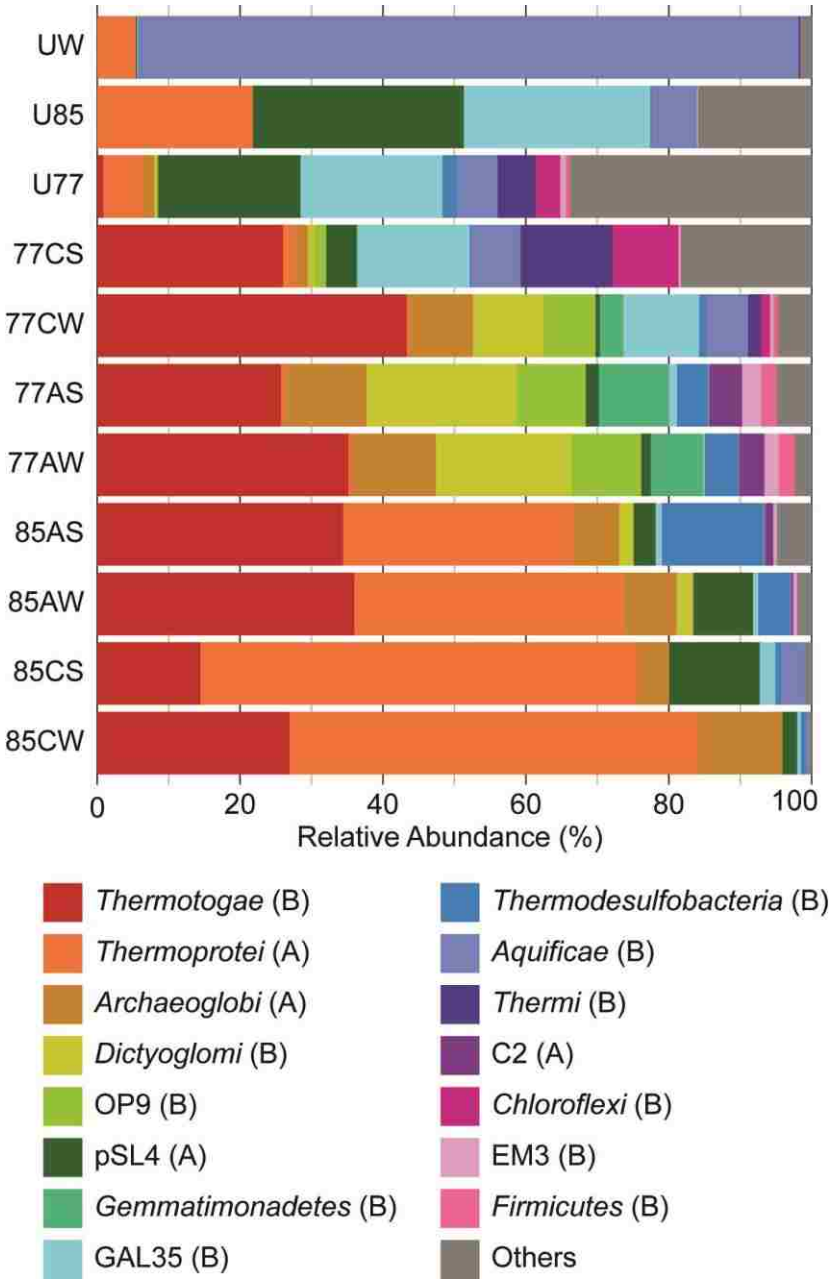


**Figure B.14.** Percent relative abundance of major bacterial phyla and archaeal classes in each sample. (A) Archaeal classes. (B) Bacterial phyla.



# APPENDIX C: EXCEL FILE CONTAINING ALPHA DIVERSITY METRICS PLOTTED AGAINST PHYSICOCHEMICAL MEASUREMENTS.

## References


**Alizadeh, H., Teymouri, F., Gilbert, T. & Dale, B. (2005).** Pretreatment of switchgrass by ammonia fiber explosion (AFEX). *Applied Biochemistry and Biotechnology* **124**, 1133–1141.

**Allgaier, M., Reddy, A., Park, J. I., Ivanova, N., D'haeseleer, P., Lowry, S., Sapra, R., Hazen, T. C., Simmons, B. A., & other authors**. **(2010).** Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. *PLoS ONE* **5**, e8812.

**Anderson, I., Rodriguez, J., Susanti, D., Porat, I., Reich, C., Ulrich, L. E., Elkins, J. G., Mavromatis, K., Lykidis, A., & other authors**. **(2008).** Genome sequence of *Thermofilum pendens* reveals an exceptional loss of biosynthetic pathways without genome reduction. *J Bacteriol* **190**, 2957–2965.

**Anderson, J. (1978).** A geochemical study of the southwest part of the Black Rock Desert and its geothermal areas; Washoe, Pershing, and Humboldt Counties, Nevada. *Colorado School of Mines Quarterly* **73**, 15–22.

**Anderson, M. J. (2001).** A new method for non-parametric multivariate analysis of variance. *Austral Ecol* **26**, 32–46.

**Ando, S., Ishida, H., Kosugi, Y. & Ishikawa, K. (2002).** Hyperthermostable endoglucanase from *Pyrococcus horikoshii*. *Appl Environ Microbiol* **68**, 430–433.

**Balch, W. E. & Wolfe, R. S. (1976).** New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM)-dependent growth of *Methanobacterium ruminantium* in a pressureized atmosphere. *Appl Environ Microbiol* **32**, 781–791.

**Barns, S. M., Fundyga, R. E., Jeffries, M. W. & Pace, N. R. (1994).** Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc Natl Acad Sci* **91**, 1609–1613.

**Bauer, M. W., Driskill, L. E., Callen, W., Snead, M. A., Mathur, E. J. & Kelly, R. M. (1999).** An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyzes β-1,4 bonds in mixed-linkage (1→3),(1→4)-β-d-glucans and cellulose. *J Bacteriol* **181**, 284–290.

**Blank, C. E., Cady, S. L. & Pace, N. R. (2002).** Microbial composition of near-boiling silica-depositing thermal springs throughout Yellowstone National Park. *Appl Environ Microbiol* **68**, 5123–5135.

**Blumer-Schuette, S. E., Kataeva, I., Westpheling, J., Adams, M. W. & Kelly, R. M. (2008).** Extremely thermophilic microorganisms for biomass conversion: status and prospects. *Curr Opin Biotechnol* **19**, 210–7.

**Blumer-Schuette, S. E., Lewis, D. L. & Kelly, R. M. (2010).** Phylogenetic, microbiological, and glycoside hydrolase diversities within the extremely thermophilic, plant biomass-degrading genus *Caldicellulosiruptor*. *Appl Environ Microbiol* **76**, 8084–8092.

**Bok, J.-D., Yernool, D. A. & Eveleigh, D. E. (1998).** Purification, characterization, and molecular analysis of thermostable cellulases CelA and CelB from *Thermotoga neapolitana*. *Appl Environ Microbiol* **64**, 4774–4781.

**Bray, J. R. & Curtis, J. T. (1957).** An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs* **27**, 325–349.

**Brock, T. D. (1967).** Micro-organisms adapted to high temperatures. *Nature* **214**, 882–885.

**Bronnenmeier, K., Kern, A., Liebl, W. & Staudenbauer, W. L. (1995).** Purification of *Thermotoga maritima* enzymes for the degradation of cellulosic materials. *Appl Environ Microbiol* **61**, 1399–1407.

**Brumm, P., Hermanson, S., Hochstein, B., Boyum, J., Hermersmann, N., Gowda, K. & Mead, D. (2011).** Mining *Dictyoglomus turgidum* for enzymatically active carbohydrases. *Applied Biochemistry and Biotechnology* **163**, 205–214.

**Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., & other authors**. **(2010).** QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335–6.

**Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., & other authors**. **(2010).** QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335–6.

**Carroll, A. & Somerville, C. (2009).** Cellulosic biofuels. *Annu Rev Plant Biol* **60**, 165–82.

**Cavaletti, L., Monciardini, P., Bamonte, R., Schumann, P., Rohde, M., Sosio, M. & Donadio, S. (2006).** New lineage of filamentous, spore-forming, Gram-positive bacteria from soil. *Appl Environ Microbiol* **72**, 4360–4369.

**Chao, A. (1984).** Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* **11**, 265–270.

**Charlier, D. & Droogmans, L. (2005).** Microbial life at high temperature, the challenges, the strategies. *Cellular and Molecular Life Sciences* **62**, 2974–2984.

**Chhabra, S. R., Shockley, K. R., Conners, S. B., Scott, K. L., Wolfinger, R. D. & Kelly, R. M. (2003).** Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. *J Biol Chem* **278**, 7540–7552.

**Chhabra, S. R., Shockley, K. R., Ward, D. E. & Kelly, R. M. (2002).** Regulation of endo-acting glycosyl hydrolases in the hyperthermophilic bacterium *Thermotoga maritima* grown on glucan- and mannan-based polysaccharides. *Appl Environ Microbiol* **68**, 545–554.

**Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A. S., McGarrell, D. M., Marsh, T., & other authors**. **(2009).** The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37**, D141–D145.

**Collins, M. D., Pirouz, T., Goodfellow, M. & Minnikin, D. E. (1977).** Distribution of menaquinones in *Actinomycetes* and *Corynebacteria*. *J Gen Microbiol* **100**, 221–230.

**Costa, K., Navarro, J., Shock, E., Zhang, C., Soukup, D. & Hedlund, B. (2009).** Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. *Extremophiles* **13**, 447–459.

**Cox, A., Shock, E. L. & Havig, J. R. (2011).** The transition to microbial photosynthesis in hot spring ecosystems. *Chemical Geology* **280**, 344–351.

**de la Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M. & Stahl, D. A. (2008).** Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol* **10**, 810–818.

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

**Dodsworth, J. A., Hungate, B. A. & Hedlund, B. P. (2011b).** Ammonia oxidation, denitrification and dissimilatory nitrate reduction to ammonium in two US Great Basin hot springs with abundant ammonia-oxidizing archaea. *Environ Microbiol* **13**, 2371–2386.

**Dodsworth, J. A., Hungate, B., de la Torre, J. R., Jiang, H. & Hedlund, B. P. (2011a).** Measuring nitrification, denitrification, and related biomarkers in terrestrial geothermal ecosystems. *Methods Enzymol* **486**, 171–203.

**Dodsworth, J. A., McDonald, A. I. & Hedlund, B. P. (2012).** Calculation of total free energy yield as an alternative approach for predicting the importance of potential chemolithotrophic reactions in geothermal springs. *FEMS Microbiology Ecology* **81**, 446-454.

**Dubinina, G. & Gorlenko, V. (1975).** New filamentous photosynthesizing green bacteria with gas vacuoles. *Mikrobiologiya* **44**, 452–458.

**Eder, W., Jahnke, L. L., Schmidt, M. & Huber, R. (2001).** Microbial diversity of the brine-seawater interface of the Kebrit Deep, Red Sea, studied via 16S rRNA gene sequences and cultivation methods. *Appl Environ Microbiol* **67**, 3077–3085.

**Edgar, R. C. (2010).** Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461.

**Engelbrektson, A., Kunin, V., Wrighton, K. C., Zvenigorodsky, N., Chen, F., Ochman, H. & Hugenholtz, P. (2010).** Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J* **4**, 642–647.

**Evans, B. R., Gilman, A. K., Cordray, K. & Woodward, J. (2000).** Mechanism of substrate hydrolysis by a thermophilic endoglucanase from *Thermotoga maritima*; *Biotechnology Letters* **22**, 735–740.

**Felsenstein, J. (2005).** PHYLIP (Phylogeny Inference Package) version 3.6.

**Fenchel, T., King, G. M. & Blackburn, T. H. (1998).** *Bacterial biogeochemistry: the ecophysiology of mineral cycling*, 2nd edn. Academic Press.

**Frock, A. D., Gray, S. R. & Kelly, R. M. (2012).** Hyperthermophilic *Thermotoga* species differ with respect to specific carbohydrate transporters and glycoside hydrolases. *Appl Environ Microbiol* **78**, 1978–1986.

**Garrity, G. M. & Holt, J. G. (2001).** Phylum BVI. *Chloroflexi* phy. nov. In *Bergey's Manual of Systematic Bacteriology: Volume One : The Archaea and the Deeply Branching and Phototrophic Bacteria*, 2nd edn., pp. 427–446. Springer.

**Gladden, J. M., Eichorst, S. A., Hazen, T. C., Simmons, B. A. & Singer, S. W. (2012).** Substrate perturbation alters the glycoside hydrolase activities and community composition of switchgrass-adapted bacterial consortia. *Biotechnology and Bioengineering* **109**, 1140–1145.

**Graham, J. E., Clark, M. E., Nadler, D. C., Huffer, S., Chokhawala, H. A., Rowland, S. E., Blanch, H. W., Clark, D. S. & Robb, F. T. (2011).** Identification and characterization of a multidomain hyperthermophilic cellulase from an archaeal enrichment. *Nature Communications* **2**, 375.

**Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S. K., & other authors**. **(2011).** Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res* **21**, 494–504.

**Haki, G. D. & Rakshit, S. K. (2003).** Developments in industrially important thermostable enzymes: a review. *Bioresource Technology* **89**, 17–34.

**Hall, T. (1999).** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp*.

**Hamilton-Brehm, S. D., Mosher, J. J., Vishnivetskaya, T., Podar, M., Carroll, S., Allman, S., Phelps, T. J., Keller, M. & Elkins, J. G. (2009).** *Caldicellulosiruptor obsidiansis* sp.

nov., an anaerobic, extremely thermophilic, cellulolytic bacterium isolated from Obsidian Pool, Yellowstone National Park. *Appl Environ Microbiol* **76**, 1014–1020.

**Hammer, O., Harper, D. A. T. & Ryan, P. D. (2001).** PAST: Paleontological Statistics software package for education and data analysis. *Paleontologia Electronica* **4**, 9.

**Hanada, S., Takaichi, S., Matsuura, K. & Nakamura, K. (2002).** *Roseiflexus castenholzii* gen. nov., sp. nov., a thermophilic, filamentous, photosynthetic bacterium that lacks chlorosomes. *Int J Syst Evol Microbiol* **52**, 187–193.

**Hedlund, B. P., McDonald, A. I., Lam, J., Dodsworth, J. A., Brown, J. R. & Hungate, B. A. (2011).** Potential role of *Thermus thermophilus* and *T. oshimai* in high rates of nitrous oxide (N2O) production in ~80 °C hot springs in the US Great Basin. *Geobiology* **9**, 471–480.

**Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W. & Foust, T. D. (2007).** Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807.

**Holm, S. (1979).** A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 65–70.

**Hreggvidsson, G. O., Kaiste, E., Holst, O., Eggertsson, G., Palsdottir, A. & Kristjansson, J. K. (1996).** An extremely thermostable cellulase from the thermophilic eubacterium *Rhodothermus marinus*. *Appl Environ Microbiol* **62**, 3047–3049.

**Huang, Z., Hedlund, B. P., Wiegel, J., Zhou, J. & Zhang, C. L. (2007).** Molecular phylogeny of uncultivated *Crenarchaeota* in Great Basin hot springs of moderately elevated temperature. *Geomicrobiology Journal* **24**, 535–542.

**Huber, R., Langworthy, T. A., König, H., Thomm, M., Woese, C. R., Sleytr, U. B. & Stetter, K. O. (1986).** *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Archives of Microbiology* **144**, 324–333.

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

**Huson, D. H., Richter, D. C., Rausch, C., Dezulian, M. F. & Rupp, R. (2007).** An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* **8**.

**International Energy Agency**. **(2008).** From 1st- to 2nd- generation biofuel technologies: An overview of current industry and RD&D activities. International Energy Agency.

**Jackson, C. R., Langner, H. W., Donahoe-Christiansen, J., Inskeep, W. P. & McDermott, T. R. (2001).** Molecular analysis of microbial community structure in an arsenite-oxidizing acidic thermal spring. *Environ Microbiol* **3**, 532–542.

**Jürgens, U. J., Meißner, J., Fischer, U., König, W. A. & Weckesser, J. (1987).** Ornithine as a constituent of the peptidoglycan of *Chloroflexus aurantiacus* diaminopimelic acid in that of *Chlorobium vibrioforme* f. thiosulfatophilum. *Archives of Microbiology* **148**, 72–76.

**Jürgens, U. J., Meissner, J., Reichenbach, H. & Weckesser, J. (1989).** L-ornithine containing peptidoglycan-polysaccharide complex from the cell wall of the gliding bacterium *Herpetosiphon aurantiacus*. *FEMS Microbiology Letters* **60**, 247–250.

**Keppen, O., Baulina, O. & Kondratieva, E. (1994).** *Oscillochloris trichoides* neotype strain DG-6. *Photosynthesis Research* **41**, 29–33.

**Kiss, H., Nett, M., Domin, N., Martin, K., Maresca, J. A., Copeland, A., Lapidus, A., Lucas, S., Berry, K. W., & other authors**. **(2011).** Complete genome sequence of the filamentous gliding predatory bacterium *Herpetosiphon aurantiacus* type strain (114-95T). *Stand Genomic Sci* **5**, 356–370.

**Kluskens, L. D., van Alebeek, G.-J. W. M., Voragen, A. G. J., de Vos, W. M. & van der Oost, J. (2003).** Molecular and biochemical characterization of the thermoactive family 1 pectate lyase from the hyperthermophilic bacterium *Thermotoga maritima*. *Biochem J* **370**, 651–659.

**Kluskens, L. D., van Alebeek, G.-J. W. M., Walther, J., Voragen, A. G. J., de Vos, W. M. & van der Oost, J. (2005).** Characterization and mode of action of an exopolygalacturonase from the hyperthermophilic bacterium Thermotoga maritima. *FEBS Journal* **272**, 5464–5473.

**Kozubal, M. A., Macur, R. E., Jay, Z. J., Beam, J. P., Malfatti, S. A., Tringe, S. G., Kocar, B. D., Borch, T. & Inskeep, W. P. (2012).** Microbial iron cycling in acidic geothermal springs of Yellowstone National Park: integrating molecular surveys, geochemical processes, and isolation of novel Fe-active microorganisms. *Front Microbiol* **3**, 109.

**Krebs, C. J. (1998).** *Ecological methodology*, 2nd edn. Benjamin Cummings.

**Kunin, V. & Hugenholtz, P. (2010).** PyroTagger : A fast , accurate pipeline for analysis of rRNA amplicon pyrosequence data. *The Open Journal* 1–8.

**Lane, D. J. (1991).** Nucleic acid techniques in bacterial systematics. In *Nucleic acid techniques in bacterial systematics*, pp. 115–174. Edited by E. Stackebrandt & M. Goodfellow. Wiley.

**Leboffe, M. J. & Pierce, B. E. (2006).** *Microbiology: Laboratory Theory and Application*, 2nd edn. Morton Publishing Company.

**Lee, K. C.-Y., Dunfield, P. F., Morgan, X. C., Crowe, M. A., Houghton, K. M., Vyssotski, M., Ryan, J. L. J., Lagutin, K., McDonald, I. R. & Stott, M. B. (2011).** *Chthonomonas calidirosea* gen. nov., sp. nov., an aerobic, pigmented, thermophilic micro-organism of a novel bacterial class, *Chthonomonadetes classis* nov., of the newly described phylum *Armatimonadetes* originally designated candidate division OP10. *Int J Syst Evol Microbiol* **61**, 2482–2490.

**Lee, N. & Reichenbach, H. (2006).** The Genus Herpetosiphon. In *The Prokaryotes*, pp. 854–877. Edited by M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer & E. Stackebrandt. Springer New York.

**Lefèvre, C. T., Abreu, F., Schmidt, M. L., Lins, U., Frankel, R. B., Hedlund, B. P. & Bazylinski, D. A. (2010).** Moderately thermophilic magnetotactic bacteria from hot springs in Nevada. *Appl Environ Microbiol* **76**, 3740–3743.

**Lewin, R. A. (1970).** New Herpetosiphon species (Flexibacterales). *Canadian Journal of Microbiology* **16**, 517–520.

**Lozupone, C. A. & Knight, R. (2007).** Global patterns in bacterial diversity. *Proc Natl Acad Sci USA* **104**, 11436–11440.

**Macur, R. E., Langner, H. W., Kocar, B. D. & Inskeep, W. P. (2004).** Linking geochemical processes with microbial community analysis: successional dynamics in an arsenic-rich, acid-sulphate-chloride geothermal spring. *Geobiology* **2**, 163–177.

**Marteinsson, V. T., Hauksdóttir, S., Hobel, C. F. V., Kristmannsdóttir, H., Hreggvidsson, G. O. & Kristjánsson, J. K. (2001).** Phylogenetic diversity analysis of subterranean hot springs in Iceland. *Appl Environ Microbiol* **67**, 4242–4248.

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

**Meyer-Dombard, D. R., Swingley, W., Raymond, J., Havig, J., Shock, E. L. & Summons, R. E. (2011).** Hydrothermal ecotones and streamer biofilm communities in the Lower Geyser Basin, Yellowstone National Park. *Environ Microbiol* **13**, 2216–2231.

**Miller, S. R., Strong, A. L., Jones, K. L. & Ungerer, M. C. (2009).** Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl Environ Microbiol* **75**, 4565–4572.

**Miller-Coleman, R. L., Dodsworth, J. A., Ross, C. A., Shock, E. L., Williams, A. J., Hartnett, H. E., McDonald, A. I., Havig, J. R. & Hedlund, B. P. (2012).** *Korarchaeota* diversity, biogeography, and abundance in Yellowstone and Great Basin hot springs and ecological niche modeling based on machine learning. *PLoS ONE* **7**, e35964.

**Minnikin, D. E., O'Donnell, A. G., Goodfellow, M., Alderson, G., Athalye, M., Schaal, A. & Parlett, J. H. (1984).** An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *Journal of Microbiological Methods* **2**, 233–241.

**Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y. Y., Holtzapple, M. & Ladisch, M. (2005).** Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology* **96**, 673–686.

**Mozhaev, V. V. (1993).** Mechanism-based strategies for protein thermostabilization. *Trends in Biotechnology* **11**, 88–95.

**Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., & other authors**. **(1999).** Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima. *Nature* **399**, 323–329.

**Nunoura, T., Hirayama, H., Takami, H., Oida, H., Nishi, S., Shimamura, S., Suzuki, Y., Inagaki, F., Takai, K., & other authors**. **(2005).** Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environ Microbiol* **7**, 1967–1984.

**Nunoura, T., Oida, H., Nakaseama, M., Kosaka, A., Ohkubo, S. B., Kikuchi, T., Kazama, H., Hosoi-Tanabe, S., Nakamura, K., & other authors**. **(2010).** Archaeal diversity and distribution along thermal and geochemical gradients in hydrothermal sediments at the Yonaguni Knoll IV Hydrothermal Field in the Southern Okinawa Trough. *Appl Environ Microbiol* **76**, 1198–1211.

**Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., Chee, G.-J., Hattori, M., Kanai, A., & other authors**. **(2011).** Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Research* **39**, 3204–3223.

**Oracle Corporation**. **(2012).** *MySQL Community Edition*. Oracle Corporation.

**Patel, B. K. C. (2010).** Phylum XX. Dictyoglomi phyl. nov. In *Bergey's Manual® of Systematic Bacteriology*, pp. 775–780. Edited by N. R. Krieg, J. T. Staley, D. R. Brown, B. P. Hedlund, B. J. Paster, N. L. Ward, W. Ludwig & W. B. Whitman. Springer New York.

**Perevalova, A. A., Svetlichny, V. A., Kublanov, I. V., Chernyh, N. A., Kostrikina, N. A., Tourova, T. P., Kuznetsov, B. B. & Bonch-Osmolovskaya, E. A. (2005).** *Desulfurococcus fermentans* sp. nov., a novel hyperthermophilic archaeon from a Kamchatka hot spring, and emended description of the genus *Desulfurococcus*. *Int J Syst Evol Microbiol* **55**, 995–999.

**Pierson, B. K. & Castenholz, R. W. (1974).** A phototrophic gliding filamentous bacterium of hot springs, *Chloroflexus aurantiacus*, gen. and sp. nov. *Archives of Microbiology* **100**, 5–24.

**Pierson, B. K., Giovannoni, S. J., Stahl, D. A. & Castenholz, R. W. (1985).** *Heliothrix oregonensis*, gen. nov., sp. nov., a phototrophic filamentous gliding bacterium containing bacteriochlorophyll. *Archives of Microbiology* **142**, 164–167.

**R Development Core Team**. **(2011).** *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

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

**Reysenbach, A. L., Ehringer, M. & Hershberger, K. (2000).** Microbial diversity at 83°C in Calcite Springs, Yellowstone National Park: another environment where the *Aquificales* and '*Korarchaeota'* coexist. *Extremophiles* **4**, 61–67.

**Reysenbach, A. L., Wickham, G. S. & Pace, N. R. (1994).** Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl Environ Microbiol* **60**, 2113–2119.

**Reysenbach, A.-L., Banta, A., Civello, S., Daly, J., Mitchell, K., Lalonde, S., Konhauser, K., Rodman, A., Rusterholtz, K. & Takacs-Vesbach, C. D. (2005).** *Aquificales* in Yellowstone National Park. In *Geothermal Biology and Geochemistry in Yellowstone National Park*. Edited by W. P. Inskeep & T. R. McDermott. Bozeman, MT: Montana State University Publications.

**Rubin, E. M. (2008).** Genomics of cellulosic biofuels. *Nature* **454**, 841–5.

**Saiki, T., Kobayashi, Y., Kawagoe, K. & Beppu, T. (1985).** *Dictyoglomus thermophilum* gen. nov., sp. nov., a chemoorganotrophic, anaerobic, thermophilic bacterium. *Int J Syst Bacteriol* **35**, 253–259.

**Sasser, M. (2006, July).** Technical note #101: Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME).

**Schleifer, K. H. (1985).** Analysis of the chemical composition and primary structure of murein. In *Methods in Microbiology*. Edited by G. Gottschalk. Academic Press.

**Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., & other authors**. **(2009).** Introducing

mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**, 7537–7541.

**Schönheit, P. & Schäfer, T. (1995).** Metabolism of hyperthermophiles. *World Journal of Microbiology and Biotechnology* **11**, 26–57.

**Simpson, E. (1949).** Measurement of diversity. *Nature* **163**, 688.

**Sims, R. E. H., Mabee, W., Saddler, J. N. & Taylor, M. (2010).** An overview of second generation biofuel technologies. *Bioresource Technology* **101**, 1570–1580.

**Sizova, M. V., Izquierdo, J. A., Panikov, N. S. & Lynd, L. R. (2011).** Cellulose- and xylan-degrading thermophilic anaerobic bacteria from biocompost. *Appl Environ Microbiol* **77**, 2282–2291.

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

**Stott, M. B., Crowe, M. A., Mountain, B. W., Smirnova, A. V., Hou, S., Alam, M. & Dunfield, P. F. (2008).** Isolation of novel bacteria, including a candidate division, from geothermal soils in New Zealand. *Environ Microbiol* **10**, 2030–2041.

**Svetlichnii, V. & Svetlichnaya, T. (1988).** *Dictyoglomus turgidus* sp. nov., a new extremely thermophilic eubacterium isolated from hot springs of the Uzon volcano caldera. *Mikrobiologiya* 364–369.

**Takahata, Y., Nishijima, M., Hoaki, T. & Maruyama, T. (2001).** *Thermotoga petrophila* sp. nov. and *Thermotoga naphthophila* sp. nov., two hyperthermophilic bacteria from the Kubiki oil reservoir in Niigata, Japan. *Int J Syst Evol Microbiol* **51**, 1901–1909.

**Tamaki, H., Tanaka, Y., Matsuzawa, H., Muramatsu, M., Meng, X.-Y., Hanada, S., Mori, K. & Kamagata, Y. (2011).** *Armatimonas rosea* gen. nov., sp. nov., of a novel bacterial phylum, *Armatimonadetes* phyl. nov., formally called the candidate phylum OP10. *Int J Syst Evol Microbiol* **61**, 1442–1447.

**Tamaoka, J., Katayama-Fujimura, Y. & Kuraishi, H. (1983).** Analysis of bacterial menaquinone mixtures by high performance liquid chromatography. *Journal of Applied Microbiology* **54**, 31–36.

**Tang, S.-K., Wang, Y., Chen, Y., Lou, K., Cao, L.-L., Xu, L.-H. & Li, W.-J. (2009).** *Zhihengliuella alba* sp. nov., and emended description of the genus *Zhihengliuella*. *Int J Syst Evol Microbiol* **59**, 2025–2032.

**Te'o, V. S. J., Saul, D. J. & Bergquist, P. L. (1995).** celA, another gene coding for a multidomain cellulase from the extreme thermophile *Caldocellum saccharolyticum*. *Applied Microbiology and Biotechnology* **43**, 291–296.

**USDA**. **(2012).** World agricultural supply and demand estimates. World Agricultural Outlook Board.

**Viikari, L., Alapuranen, M., Puranen, T., VehmaanperÃ¤, J., Siika-aho, M. & Olsson, L. (2007).** Thermostable enzymes in lignocellulose hydrolysis biofuels. In , Advances in Biochemical Engineering/Biotechnology, pp. 121–145. Springer Berlin / Heidelberg.

**Windberger, E., Huber, R., Trincone, A., Fricke, H. & Stetter, K. O. (1989).** *Thermotoga thermarum* sp. nov. and *Thermotoga neapolitana* occurring in African continental solfataric springs. *Archives of Microbiology* **151**, 506–512.

**Xue, Y., Wu, A., Zeng, H. & Shao, W. (2006).** High-level expression of an α-arabinofuranosidase from *Thermotoga maritima* in *Escherichia coli* for the production of xylobiose from xylan. *Biotechnology Letters* **28**, 351–356.

**Yabe, S., Aiba, Y., Sakai, Y., Hazaka, M. & Yokota, A. (2010).** A life cycle of branched aerial mycelium- and multiple budding spore-forming bacterium *Thermosporothrix hazakensis* belonging to the phylum *Chloroflexi*. *The Journal of General and Applied Microbiology* **56**, 137–141.

**Yabe, S., Aiba, Y., Sakai, Y., Hazaka, M. & Yokota, A. (2011).** *Thermogemmatispora onikobensis* gen. nov., sp. nov. and *Thermogemmatispora foliorum* sp. nov., isolated from fallen leaves on geothermal soils, and description of *Thermogemmatisporaceae* fam. nov. and *Thermogemmatisporales* ord. nov. within the class *Ktedonobacteria*. *Int J Syst Evol Microbiol* **61**, 903–910.

**Yang, S.-J., Kataeva, I., Wiegel, J., Yin, Y., Dam, P., Xu, Y., Westpheling, J. & Adams, M. W. W. (2010).** Classification of 'Anaerocellum thermophilum' strain DSM 6725 as *Caldicellulosiruptor bescii* sp. nov. *Int J Syst Evol Microbiol* **60**, 2011–2015.

**Youssef, N. H., Blainey, P. C., Quake, S. R. & Elshahed, M. S. (2011).** Partial genome assembly for a candidate division OP11 single cell from an anoxic spring (Zodletone Spring, Oklahoma). *Appl Environ Microbiol* **77**, 7804–7814.

**Youssef, N., Sheik, C. S., Krumholz, L. R., Najar, F. Z., Roe, B. A. & Elshahed, M. S. (2009).** Comparison of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl Environ Microbiol* **75**, 5227–5236.

**Zeikus, J. G., Dawson, M. A., Thompson, T. E., Ingvorsen, K. & Hatchikian, E. C. (1983).** Microbial Ecology of Volcanic Sulphidogenesis: Isolation and Characterization of *Thermodesulfobacterium commune* gen. nov. and sp. nov.,. *J Gen Microbiol* **129**, 1159–1169.

**Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. (2000).** A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**, 203–14.

**Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. (2000).** A greedy algorithm for aligning DNA sequences. *J Comput Biol* **7**, 203–14.

**Zhaxybayeva, O., Swithers, K. S., Lapierre, P., Fournier, G. P., Bickhart, D. M., DeBoy, R. T., Nelson, K. E., Nesbø, C. L., Doolittle, W. F., & other authors**. **(2009).** On the chimeric nature, thermophilic origin, and phylogenetic placement of the *Thermotogales*. *PNAS* **106**, 5865–5870.

**Zverlov, V., Mahr, S., Riedel, K. & Bronnenmeier, K. (1998).** Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile '*Anaerocellum thermophilum'* with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology* **144**, 457–465.

# VITA

Graduate College
University of Nevada, Las Vegas

Jessica K. Cole

Degree:
Bachelor of Science, Biology, 2010
University of Nevada, Las Vegas

Publications:

Hedlund B.P., **Cole J.K.**, Williams A.J., Hou W., Zhou E., Li W., and Dong H. 2012. A review of the microbiology of the Rehai geothermal field in Tengchong, Yunnan Province, China. Geoscience Frontiers 3:273-288.

Thesis Title: Natural and Lignocellulose-Enriched Microbial Communities in Great Boiling Spring, NV

Thesis Examination Committee:

Committee Chair, Brian Hedlund, Ph.D.
Committee Member, Eduardo Robleto, Ph.D.
Committee Member, Helen Wing, Ph.D.
Graduate Faculty Representative, Ernesto Abel-Santos, PhD.