COMPARISON OF BACTERIAL COMMUNITIES IN LIVING EASTERN HEMLOCK (*TSUGA CANADENSIS)* RHIZOSPHERES VERSUS PRESUMABLY DEAD EASTERN HEMLOCK RHIZOSPHERES

A thesis submitted to the faculty of the Graduate School of Western Carolina University in partial fulfillment of the requirements for the degree of Master of Science

By

Kendall Leigh Fuller

Director: Dr. Seán O'Connell
Associate Professor and Department Head
Department of Biology

Committee Members: Dr. Thomas Martin, Biology,
Dr. Ron Michaelis, Biology

March 2012

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Seán O'Connell, for all the assistance, guidance and support. This work would not have been possible without him. I would also like to thank my committee members, Dr. Thomas Martin, Dr. Ron Michaelis and Dr. Sabine Rundle for their guidance and assistance. I would like to thank Emily York for the use of data from her previous work and Carter Dillow for the use of his samples. Also, I would like to thank the Biology Department of Western Carolina University for the use of equipment and funding, along with the Graduate School of Western Carolina University for funding. Finally, I would like to give special thanks to my friends and parents, Kenneth and Sherry Fuller for their encouragement, support and dog sitting efforts that allowed me to complete this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

COMPARISON OF BACTERIAL COMMUNITIES IN LIVING EASTERN HEMLOCK (*TSUGA CANADENSIS)* RHIZOSPHERES VERSUS PRESUMABLY DEAD EASTERN HEMLOCK RHIZOSPHERES

Kendall Leigh Fuller, M.S.

Western Carolina University

Director: Dr. Seán O'Connell

The purpose of this study was to identify bacterial species associated with the rhizosphere of living and dead hemlocks (those that have been likely killed by adelgid infestation). Study of the bacterial diversity associated with the rhizosphere of hemlocks, could lead to elucidation of partnerships between bacteria and hemlocks. Samples were collected from Albright Grove, Great Smoky Mountains National Park (GSMNP), in early February 2011 from the soil attached to hemlock roots from six trees, consisting of paired live/dead trees that were found side by side. Species richness and evenness within the samples were evaluated. Variation based on time of year and over time was assessed for hemlock rhizospheres. Some of the bacteria detected in this work were unidentified at the phylum level and are likely new taxa to science. Overall, *Acidobacteria* was the dominant phylum making up 63% of all samples, followed by *Proteobacteria* at 23%, other phyla were represented at levels ≤ 6%. Results indicated significant differences in the composition at the phylum level of Live 2011 and Dead 2011 samples. Other significant differences were found at lower levels of classification between all six comparisons of the four sample sets. This research has built upon the research of two former M.S. thesis students by using clone libraries prepared in the summer of 2006 and DNA extracts from samples collected during the winter of 2008 from the same field site.

The association of microbial communities with living hemlocks is important. If hemlocks cease to exist in the GSMNP because of the infestation of the Hemlock Woolly Adelgid (HWA), then unidentified microorganisms that may be specific to hemlock may become extinct as well. If reforestation efforts were ever to take place for Eastern Hemlock in GSMNP, microbial communities associated with healthy trees could be vital in the success of this effort.

INTRODUCTION

Development of Microbiology and Classification of Microorganisms

Development of microscopes led to the first observation of bacteria by Antonie van Leeuwenhoek in 1676 and a series of findings in the 17th and 18th century led to the founding of bacteriology. Ferdinand Cohn is credited with this founding, based on his work that led him to the discovery of the genus *Bacillus*; he described the entire life cycle (vegetative cell→endospore→vegetative cell) in 1875 (Brock et al. 2006, Baker et al. 2007). Cohn discovered the process of spore formation, and he also found that vegetative cells but not endospores were killed by boiling (Brock et al. 2006). Cohn's early discoveries helped the first medical microbiologist, Robert Koch, form what is now termed "Koch's Postulates," demonstrating that a specific microorganism causes a specific disease (Brock et al., 2006). By the start of the 20th century, after a series of many discoveries, the first journal in microbiology, *Journal of Bacteriology* was published by the American Society for Microbiology (Baker et al. 2007). Much work has been completed since, and microbiology continues to produce revolutionary findings in science.

The development of molecular techniques, including ribosomal RNA sequencing has led to the discovery of three phylogenetically distinct domains of organisms: *Bacteria, Archaea* and *Eukarya* (Figure 1). The domains *Bacteria* and *Archaea* are made up exclusively of prokaryotes, though interestingly molecular and biochemical classification reveals that *Archaea* are more closely related to eukaryotic species than are species of *Bacteria* (Brock et. al 2006). Phylogenies based on RNA sequencing have allowed microbiologists to identify and discover organisms without culturing them,

which is important because not all bacteria can be cultured in a laboratory setting and culture-independent analysis allows for much more rapid classification (Brock et al. 2006).

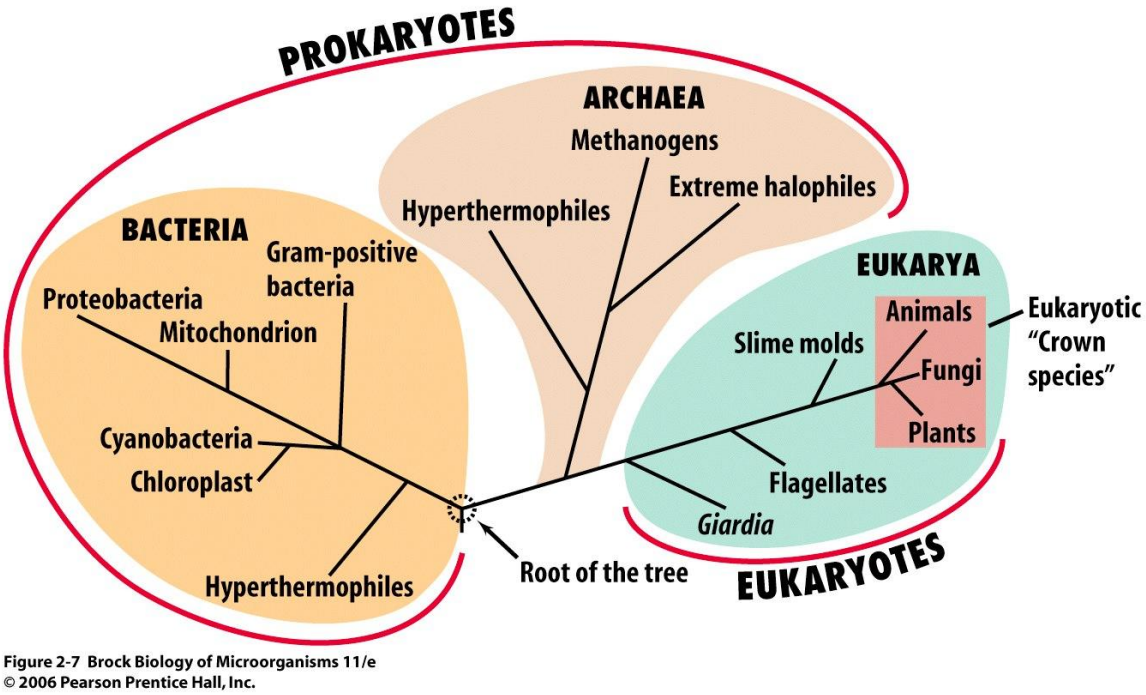

Figure 1. Phylogentic tree of life defined by rRNA gene sequencing. The three domains are *Bacteria*, *Archaea* and *Eukarya* (Brock et al. 2006).

The microbial world is immensely diverse and conceptually large with more than 10^{30} individual prokaryotes on Earth (Sloan et al. 2008). Culture independent analysis of the small subunit of the 16S rRNA genes is the most frequently used molecular method for bacterial identification due to the highly conserved nature of the gene (Cardenas et al. 2009). When determining biodiversity among bacterial communities based on differences in the 16S rDNA, the term species may not be the most appropriate term; instead the term operational taxonomic unit (OTU) may be more appropriate (Sloan et al. 2008). Sequences determined to be 97% identical or higher are classified into the same

OTU, which is the typical classification definition for bacterial species (Cardenas et al. 2009). The substantial heterogeneity and evolutionary divergence within prokaryotes exemplifies the problem of defining species, because what is suitable for one phylogenetic group might not be valid for another (Rossellò-Mora and Lòpez-Lòpez, 2008). The acceptance of microbial ecologists that the straightforward idea of "species" may not be applicable to bacterial diversity has led to the use of a number of terms including OTU to describe diversity and the term OTU is used in this study for those sequences found to share ≥ 97% similarity (Rossellò-Mora and Lòpez-Lòpez, 2008).

As of 2004, the domain *Bacteria* had been divided into 24 phyla (containing cultured representatives) and three times that with uncultured species, as shown in Figure 2 (Schloss and Handelsman 2004). Given the high level of microbial diversity, identification of microorganisms is essential and one of the most valuable and authoritative methods for identification is 16S rRNA gene analysis (Cardenas et al. 2009). Highly conserved regions in the rRNA gene allow for the use of "universal" primers and for amplification of the gene from DNA extracted straight from the environment; other regions of the gene have increasing variation which is reflective of evolutionary distance (Cardenas et al. 2009). The first step in microbial identification using the 16S rRNA gene is DNA extraction, which can be completed using commercial kits available for samples from water, soil or most any other type of environment (Cardenas et al. 2009). "Universal" primers are then used to amplify the extracted DNA, which is then inserted into a vector and inserted into a host cell such as *Escherichia coli* (Cardenas et al. 2009). Hosts are then grown on selective media for host cells carrying the PCR product in the vector, thus creating a clone library (Cardenas et al. 2009).

Vectors can then be extracted from the clones and sequenced. Sequences can be classified and compared to other bacterial rRNA sequences using the Ribosomal Database Project (RDP; Cole et al. 2008) Classifier and Library Compare tools as well as other analyses (Cardenas et al. 2009).

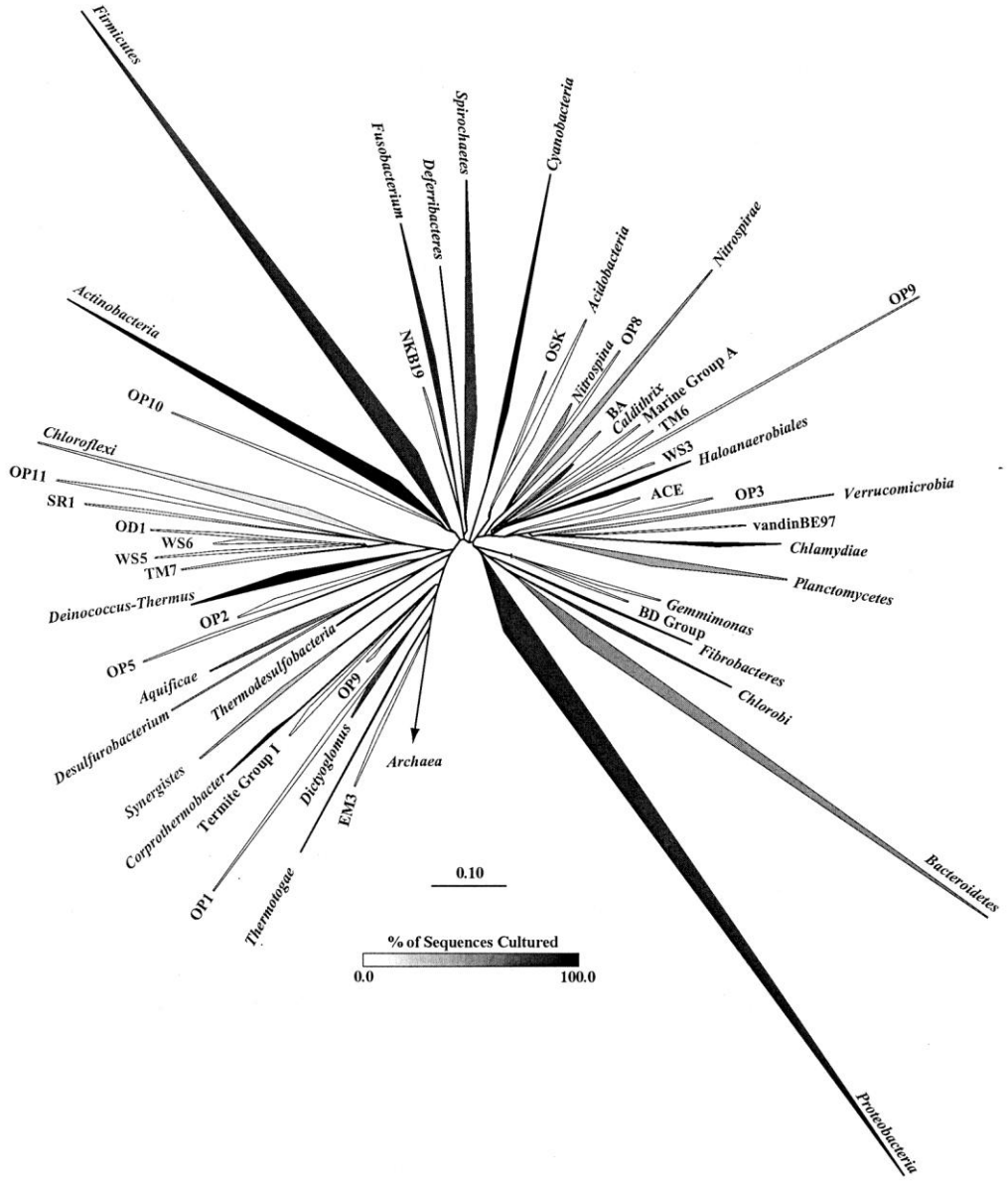

Figure 2. Phylogenetic tree of the *Bacteria* showing established phyla (italicized Latinized names) and candidate phyla. The width of each branch corresponds to the number of sequences in each phylum and shading represents the fraction of sequences that were acquired from cultured representatives (darker branches have more cultured representatives). Figure taken from Schloss and Handelsman 2004.

Taxonomists are using molecular techniques like rRNA and other gene sequencing to identify species, whereas before biochemical tests and phenotypic characteristics that tested the ability of bacteria to metabolize different resources were solely used (Fraser et. al. 2009). The best method for identifying and classifying bacteria is likely a combination of genetic identification and culture based techniques. However, only a minute amount of the vast level of microbial diversity can be cultured in the laboratory and it is important to recognize that not all parameters, such as interactions with other organisms can be sufficiently simulated in the laboratory (Zengler 2008). And the real issue with classifying and identifying bacteria seems to be defining a bacterial species in general, whether using molecular techniques or phenotypic characteristics.

Importance of Prokaryotes

Prokaryotes are vital components of the earth's biosphere and are important for a variety of reasons. Prokaryotes inhabit every location where life exists, ranging in conditions considered ideal for growth to the most extreme environments (Schlegel and Jannasch 2006). Classification of prokaryotes began in 1872 when Ferdinand Cohn grouped bacteria by a morphological approach, which soon proved to be insufficient (Truper and Schleifer, 2006). Prokaryotes are classified into two domains, *Bacteria* and *Archaea,* both of which are incredibly genetically diverse and techniques for classifying these organisms are more recently becoming widespread (Fraser et al. 2009).

Bacteria are fundamental players in the biogeochemical pathways in every ecosystem in the biosphere (Nehl and Knox 2006). Nitrogen fixation, for example, is one of the most important processes essential to all life that is exclusively carried out by prokaryotes (Zinder and Dworkin 2006). Some nitrogen fixation takes place in free living

bacteria, but perhaps more importantly to humans much occurs in symbiotic bacteria in association with plants. Nitrogen fixation by prokaryotes happens when N_2 is reduced to ammonia, which can be converted to organic form (Manoharachary and Mukerji 2006). The enzyme involved in reducing the nitrogen gas is nitrogenase. Many bacteria are capable of fixing nitrogen including *Rhizobium spp.*, which infect the roots of plants and produce root nodules (White 2007). The root nodules are where bacteria fix nitrogen that is then provided to the plant. In this symbiotic relationship the plant provides the bacteria with organic nutrients from photosynthesis, while bacteria provide nitrogen in a usable form (White 2007). To date, no eukaryotes have been found to fix nitrogen and given that nitrogen is a limiting factor in plant growth, the entire world's food chain is dependent upon the ability of prokaryotes to fix nitrogen (White 2007).

Nitrogen is not the only element where microorganisms occupy a large role in terms of elemental cycling. Microorganisms, through decomposition, play a large part in the carbon cycle, ultimately converting a range of carbon compounds (many made through photosynthesis) to methane and carbon dioxide (Baker et al. 2007) while also releasing sulfur, iron, phosphorous, and other nutrients. Methanogens and chemoorganotrophs break down the carbon compounds via fermentation, anaerobic respiration or aerobic respiration processes (Brock et al. 2006).

The Rhizosphere

The rhizosphere was first defined in 1904 by German agronomist Hiltner "as the 'effect' of the roots of legumes on the surrounding soil, in terms of higher microbial activity because of the organic matter released by the roots" (Barriuso et al. 2008). The rhizosphere is a "very complex environment in which the effects of the plant on soil

microorganisms and the effects of the microorganisms on the plant are interacting and are interdependent" (Manoharachary and Mukerji 2006). The effect of the rhizosphere makes this zone its own ecosystem (one of the largest on earth), with its impacts being crucial to the functioning of the biosphere (Barriuso et al. 2008). Some estimate plants release between 20-50% of photosynthates through their roots, making the energy exchange in the rhizosphere massive (Barriuso et al. 2008). Microbial growth stimulated by the release of organic compounds by plants allows for a diverse microbial community in the rhizosphere. The rhizosphere is important for a variety of reasons, including plant nutrition and health as well as microbial driven carbon sequestration and nutrient cycling essential for ecosystem functioning (Berg and Smalla 2009). Advances in molecular tools have allowed for a wealth of information to be discovered regarding microbial effects in the rhizosphere.

The unique microenvironment that is the rhizosphere, creates a complex microbial network that impacts all food chains, and plant species influence on the microbial communities present in the rhizosphere has been shown in many studies (Berg and Smalla 2009). The finding that unique plant species harbor unique microbial communities in the rhizosphere has implications for agriculture, biocontrol and nature conservation (including reforestation), as well as many others (Berg and Smalla 2009).

Eastern Hemlock and Hemlock Woolly Adelgid Infestation

Great Smoky Mountains National Park (GSMNP) has the largest widespread distribution of old-growth Eastern Hemlocks (*Tsuga canadensis)* of any other National Park in the United States, with individual trees exceeding 150 feet tall and six feet in diameter. The Hemlock Woolly Adelgid (*Adelges tsugae*; HWA), an invasive species,

was identified in GSMNP in 2002 and has since been found throughout the park, having the potential to eradicate Eastern Hemlocks from the area (GSMNP: HWA 2006). HWA is a small piercing/sucking insect that is native to Asia, and was originally introduced into North America in the 1920s, colonizing the eastern United States in the 1950s. HWA have a complex polymorphic life cycle in which the nymphs feed on fluids from the hemlock after inserting their stylets into the xylem, while most likely releasing toxins that cause desiccation and loss of needles on the hemlocks (McClure and Cheah 1999). HWA have few natural predators in the eastern United States and winters with consistently low temperatures seem to be the only natural occurrence that can lower the abundance of the HWA population (McClure and Cheah 1999). The decline and possible loss of hemlocks due to the HWA infestation will have complex effects on the ecosystem in which the infestations occur. Effects include changes in energy and nutrient flux, including nitrogen and in species composition (Stadler et al. 2006). Bacteria have been shown to have significantly larger colony forming units (CFUs) on HWA-infested trees and the surrounding soil; filamentous fungi have also been shown to be more numerous on HWA infested trees (Stadler et al. 2006).

GSMNP managers have made extensive efforts to control the HWA population, by spraying insecticides on trees, applying systemic insecticides into the soil or trunks of trees and by releasing 65,000+ predatory beetles (*Sasajiscymnus tsugae* and *Laricobius nigrinus*) as a biological control. Treatments have produced some promising results as trees have produced new growth and foliage has regained its color in some plots (GSMNP: HWA 2006). Others have also had hopeful results in treating the HWA infestation. A study completed in Frederick County, Maryland in 2003 showed hemlocks

treated with a systemic insecticide, imidacloprid, can control the HWA population and that hemlocks can recover by generating new growth on many branches (Webb et al. 2003). However, these treatments are only feasible where trees can be easily accessed and much of the hemlock forest in GSMNP, which are not ready accessible have witnessed extensive die back.

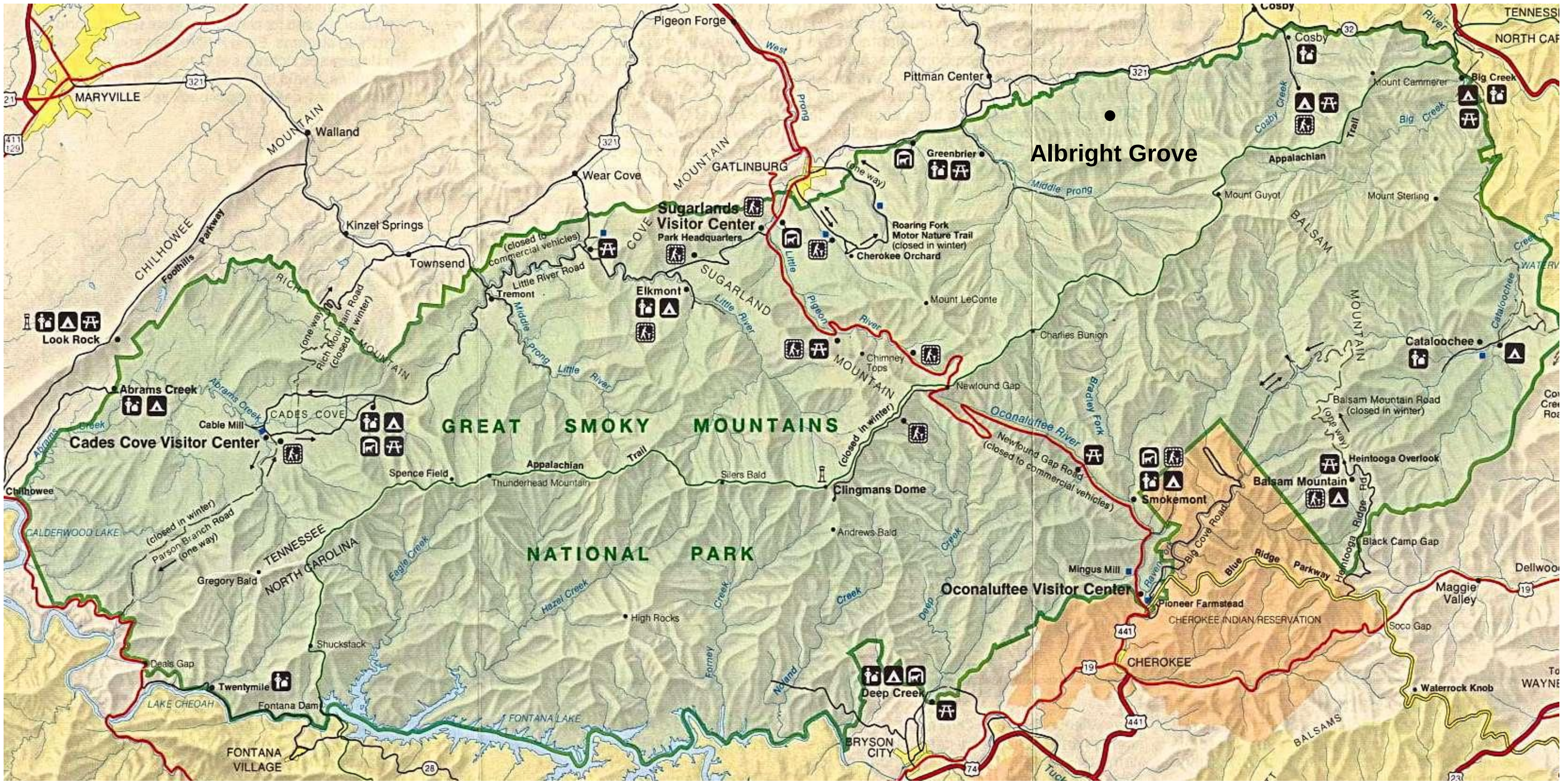

Figure 3. Map of Great Smoky Mountains National Park (GSMNP) showing the location of Albright Grove in Tennessee, where samples used in this study were collected from 2006-2011.

Purpose

The purpose of this study was to identify bacterial species associated with the rhizosphere of live and dead hemlocks (those that have presumably been killed by HWA infestation). Study of the bacterial diversity associated with the rhizosphere of healthy versus dead hemlocks could lead to elucidation of partnerships between bacteria and hemlocks. If hemlocks are eradicated in the GSMNP because of the infestation of the HWA, then unidentified microorganisms that may be specific to hemlock may become extinct as well. This research has built upon the research of two former M.S. thesis students by using clone libraries prepared in the summer of 2006 and DNA extracts from samples collected during the winter of 2008 from the same field site. Variation based on time of year and over time was assessed for hemlock rhizospheres. Samples were collected from Albirght Grove, GSMNP (Figure 3).

Hypothesis One

Bacterial diversity (richness and evenness) will be higher in the rhizospheres of living Eastern Hemlock versus dead hemlock in the same forest plots. Previous research has shown that bacterial diversity in living hemlock rhizospheres is greater than in nearby bulk soil, possibly due to the nutrients provided by the trees. Dead trees would provide fewer such photosynthates and would be expected to provide substrates for general decomposers only.

Hypothesis Two

Bacterial diversity (richness and evenness) will be similar in living hemlock rhizospheres sampled in the same forest plot independent of season and years between samples taken. Microbial diversity is high in the soil environment and community composition can

change temporally; however the magnitude of diversity is expected to stay the same as different species assemblages respond to temperature, moisture and other physiochemical factors.

METHODS AND MATERIALS

Sample collection

Samples were collected in February 2011 from Albright Grove, an old growth forest, in GSMNP. Aspetic techniques, including the use of flame sterilized tools, were used throughout collection of the samples. Soil attached to hemlock roots from six trees was collected. Trees were paired as living and dead trees growing side by side and were of approximately the same size. Sample one was from paired trees approximately 10-20 ft tall with a 10 in diameter, sample two was from paired trees approximately 20-30 ft tall with a 20 in diameter and sample 3 was from paired trees approximately 10-20 ft tall with a 15 in diameter. Living trees were characterized as containing many green needles while dead trees were those with no or few green needles. Samples were taken by exposing primary roots (1-4 in below the surface) and sampling the secondary root; loose soil was shaken off the roots and roots were collected in a 50mL sterile tube. The samples were placed on ice and transported to Western Carolina University where they were stored at -70°C until processed. Samples from previous studies were collected in a similar manner in June 2006 and December 2008. These samples were used as a comparison to the 2011 samples.

Soil preparation for DNA extraction

Soil samples were used in a culture-independent method, molecular cloning of 16S rDNA, to identify prevalent species from hemlock rhizospheres. Sample weights were between 1.67-3.75 g without roots. Depending upon weight of the sample 16.7 mL-37.5 mL of sterile deionized water was added to the sample (i.e., a 1:10 dilution) in a 125 mL KiMax flask with a magnetic stir bar. Samples were stirred for five min at a setting of

4-5 on a Corning stirrer/hot plate (Model PC-620, Corning, NY). A volume of 2.0 mL of each sample was pipetted into a microcentrifuge tube and centrifuged at 10,000 x g for 10 min. The supernatant was discarded and samples were weighed. More sample slurry was added and repeatedly centrifuged with the supernatant being discarded until ~0.3g of sample was obtained.

DNA Extraction

DNA was extracted from the samples using the PowerLyzer, PowerSoil DNA Isolation kit using the method provided (MoBio® Laboratories Inc., Solana Beach, CA). A bead-beater was used for one minute at 2,500 RPM instead of a vortex for lysing the cells in the soil.

Emily York's samples were collected in June 2006 and DNA from her samples were then extracted from the hemlock rhizosphere soil samples using a Power Soil DNA Extraction kit employing the alternative lysis method (MoBio® Laboratories Inc., Solana Beach, CA). Samples were processed for molecular cloning using the same method as stated below in 2006/2007 (York 2008). Carter Dillow's samples were collected in December 2008 and he used the same DNA extraction procedure as below (Dillow 2009); his DNA extracts were processed along with samples from the living and dead trees in my study.

PCR Amplification of Bacterial 16S rDNA

The full bacterial 16S rRNA gene from the DNA extracts were amplified by polymerase chain reaction (PCR) using forward primer 27F and reverse primer 1492R (Corinaldes et al. 2005). PCR reactions were performed in a 50 µL reaction mix containing nuclease free water, Promega Master Mix 2x (Promega, Inc., Madison, WI),

1.0 µL of DNA, and 1.0 µL extra Mg++ for each sample (final concentration 3mM). Eppendorf Mastercycler personal thermocycler (Eppendorf Corp.,Westbury, NY) parameters were denaturation for 1 min at 94°C, annealing for 1 min at 55°C and elongation for two minutes at 72°C. Thirty cycles of this PCR regime were completed with a final elongation of ten min at 72°C. PCR products were visualized on a 1% agarose gel run at 75 V for 45 min and compared to a 100 bp ladder. PCR products were stored at 4ºC until further analysis.

Molecular Cloning

Molecular cloning of the amplified PCR products was performed using the Promega pGEM-T Easy kit (Promega Inc., Madison, WI). The Promega pGEM-T Easy kit utilizes a linearized vector with a single 3' terminal thymidine at both ends to facilitate ligation of PCR products. Successful ligation of an insert into the pGEM-T Easy Vector, interrupts the coding sequence of β-glactosidase, disrupting the ability of the transformed *E. coli* cells to use X-gal. Therefore recombinant clones can be identified by color screening, with white colonies indicating transformed recombinant clones and blue colonies indicating unsuccessful transformation/recombination. After blue/white colony screening 100 white colonies per sample were grown in LB media with 15% glycerol, then transferred to 96-well plates or PCR tubes and stored at -70°C until the inserts could be sequenced.

DNA Sequencing

Clones were grown for 24 hours at 37.5°C and plasmid DNA was isolated from *E. coli* DNA using the Zymo Zyppy -96 Plasmid Mini-Prep kit (Zymo Research Corp., Irvine, CA). Instructions were followed as provided with the exception of centrifugation

during step four was processed at 2250g for 12 min. Using PCR with 341F and 907R primers each sample was amplified in a 50µL reaction using 1.0 µL of DNA. The 341F and 907R primers target a small section of the 16S rDNA gene as opposed to the full gene. Each sample reaction contained 23µL nuclease free water, 0.5µL of each primer and 25µL of Promega Master Mix 2x. Thermal cycler conditions consisted of 30 cycles of denaturation for 1 min at 94°C, annealing for 1 min at 55°C and elongation for 2 min at 72°C with a final elongation of 10 min at 72°C. PCR products were purified using PSI Ψ Clone PCR 96 kits (Princeton Separations, Adelphia, NJ). A 10µL sequencing PCR reaction was performed using 0.75µL of DNA. Each sample reaction contained 6.25µL nuclease free water, 1.0µL of 341F primer (3.2pmol) and 2.0µL of Big Dye Terminator Version 3.1 Cycle Sequencing Mix (Applied Biosystems, Foster City, CA). Thermal cycler condition consisted of 25 cycles of denaturation for 10 sec at 96°C, annealing for 5 sec at 50°C and elongation for 4 min at 60°C. Sequencing PCR products were purified using Centri Sep 96 kit (Princeton Separations, Adelphia, NJ) and dried down using a Vacufuge Concentrator 5301 system (Eppendorf Corp., Westbury, NY). The dried samples were resuspended in 10µL of HiDi formamide (Applied Biosystems, Foster City, CA) and denatured for 5 min at 95°C and sequenced using a Genetic Analyzer 3130xl capillary sequencer (Applied Biosystems Inc., Foster City, CA).

The four clones per sample that were not processed in 96-well plates were processed using whole cell 50µL PCR reactions containing 14µL nuclease free water, 25µL Promega MasterMix 2X (Promega, Inc., Madison, WI) and 0.05µL M13F and M13R primers. These primers were used for sequencing inserts cloned into *Lac*Z-containing plasmids. A reaction was initiated at 99°C for 15 min with cells in water and

then Promega MasterMix 2X and other PCR reagents were added at a 80°C hold within 5 min (i.e., a "hot start" PCR was used to lyse *E.coli* cells). The thermal cycler condition then consisted of 35 cycles of denaturation for 1 min at 94°C, annealing for 1 min at 55°C and elongation for one min at 72°C, with a final elongation of four min at 72°C. PCR products for these samples were purified using Millipore Centrifugal Filter Units (Millipore Corporation, Billerica, MA). The cleaned PCR products were quantified using agarose gel electrophoresis and a Lambda Hind III molecular weight marker. PCR reactions with 341F and 907R primers as well as sequencing PCR reactions were then performed as stated above. PCR products from the sequencing PCR reaction were purified using Illustra Autoseq G-50 Dye Terminator Removal Kit (GE Healthcare, Buckinghamshire, UK). The dried samples were resuspended in 10µL of HiDi formamide (Applied Biosystems, Foster City, CA) and transferred to a 96 well plate. Samples were denatured for 5 min at 95°C and sequenced using a 3130xl capillary sequencer (Applied Biosystems Inc., Foster City, CA).

Data Analysis

Potential chimeric sequences were removed using a Database Enabled Code of Ideal Probe Hybridization Employing R (DECIPHER) Find Chimeras web tool and any sequences identified as potential chimeras were not analyzed further (Wright et al. 2012). Non-chimeric clone sequences were classified from the domain to the genus level using the Ribosomal Database Project (RDP) Classifier based on the small subunit of the 16S rDNA sequence (Cole et al. 2008, Wang et al. 2007). Sample libraries were compared using the LibCompare tool from the Ribosomal Database Project (Cole et al. 2008). Bacterial richness and evenness patterns were assessed using species rarefaction curves where a species was defined as sequences sharing ≥97% similarity (Schloss and

Handelsman 2005). A computer program for Defining Operational Taxonomic Units and Estimating Species Richness (DOTUR; Schloss and Handelsman 2005), was used to assign sequences to operational taxonomic units (OTUs) based on a distance matrix created using Green Genes alignment tool and distance matrix calculator (DeSantis et al. 2006a, 2006b). Shannon-Weaver diversity index values and Chao richness estimates with 95% confidence valves were calculated using DOTUR. Pair-wise log likelihood ratio tests (G-test) of independence were performed for sequences assigned at the phylum level using "R" (The R Foundation for Statistical Computing, Vienna, Austria) (R Development Core Team. 2008). An experiment-wise error rate of 0.10 using the Holm sequentially rejective procedure was maintained while determining significant P-values (Holm 1979).

RESULTS

Molecular cloning methods produced 937 DNA sequences from the 16S rRNA gene between all samples of live, dead, those collected in December, 2008 and those processed by Emily York from samples collected in June, 2006 (Table 1). Samples collected from living and dead Eastern Hemlock (*Tsuga canadensis*) in February, 2011 will be referred to as "Live 2011" and "Dead 2011," respectively. While those samples collected in December, 2008 and June, 2006 will be referred to as "Winter 2008" and "Summer 2006", respectively. Diversity among samples was analyzed from the phylum level down to genus, and species defined here as sequences sharing ≥ 97% homology for species. *Acidobacteria* dominated the collection of all sequences at 63.0%, followed by *Proteobacteria* at 23.6% (Table 1). Other phyla represented in the collection were *Actinobacteria, Armatimonadetes, Bacteroidetes, Firmicutes, Planctomycetes* and *Verrucomicrobia* at levels equal to or less than 6.0% each. The collection also contained bacterial sequences that were unclassified at the phylum level, consisting of 2.0% of all samples (Table 1). *Bacteroidetes* were found most often in samples from Winter 2008, while *Firmicutes* were only found in samples collected before 2011 (Table 1).

Table 1. Ribosomal Database Project Classifier results for phylum level diversity for bacterial 16S rDNA sequences cloned from Eastern Hemlock (*Tsuga canadensis*) rhizospheres from Albright Grove, Great Smoky Mountains National Park, showing the number of sequences assigned to each phylum.

	Live 2011	**Dead 2011**	**Winter 2008**	**Summer 2006**	**All**
Acidobacteria	166	150	145	129	590
Proteobacteria	47	64	73	37	222
Planctomycetes	13	11	15	17	56
Bacteroidetes	6	15	5	2	28
Verrucomicrobia	1	5	5	2	13
Actinobacteria	0	4	1	0	5
Firmicutes	0	0	3	1	4
Armatimonadetes	0	0	0	1	1
Unclassified	6	7	4	2	19
Total	239	256	251	191	937

Table 2. Summary of pair-wise log likelihood ratio tests (G-tests) of independence for sequences assigned to phyla. Note that P-values denoted by an asterisk are significant while maintaining an experiment-wise error rate of 0.10 using the Holm sequentially rejective procedure (Holm 1979). There were six degrees of freedom for all pair-wise comparisons; *Actinobacteria*, *Firmicutes*, and *Armatimonadetes* were pooled into a "rare-phyla" category due to a lower number of sequences classified in those phyla. (Table constructed by Thomas Martin, 2012)

	Dead 2011	Winter 2008	Summer 2006
Live 2011	G=15.5, P=0.017*	G=15.9, P=0.014*	G=8.3, P=0.216
Dead 2011		G=7.3, P=0.294	G=16.9, P=0.010*
Winter 2008			G=8.9 P=0.179

Pair-wise log likelihood ratio tests (G-test) of independence were performed for sequences assigned to phyla between sample sets using the data as shown in Table 1 (Table 2). Comparison of Live 2011 and Dead 2011 samples yielded a significant p-value of 0.017. Comparison of Live 2011 and Winter 2008 samples yielded a significant p-

value of 0.014. The other significant difference was found between Summer 2006 and Dead 2011 samples with a p-value of 0.010. Significant differences were not found between Summer 2006 and Live 2011 samples, Winter 2008 and Dead 2011 samples, and Winter 2008 and Summer 2006 samples (Table 2).

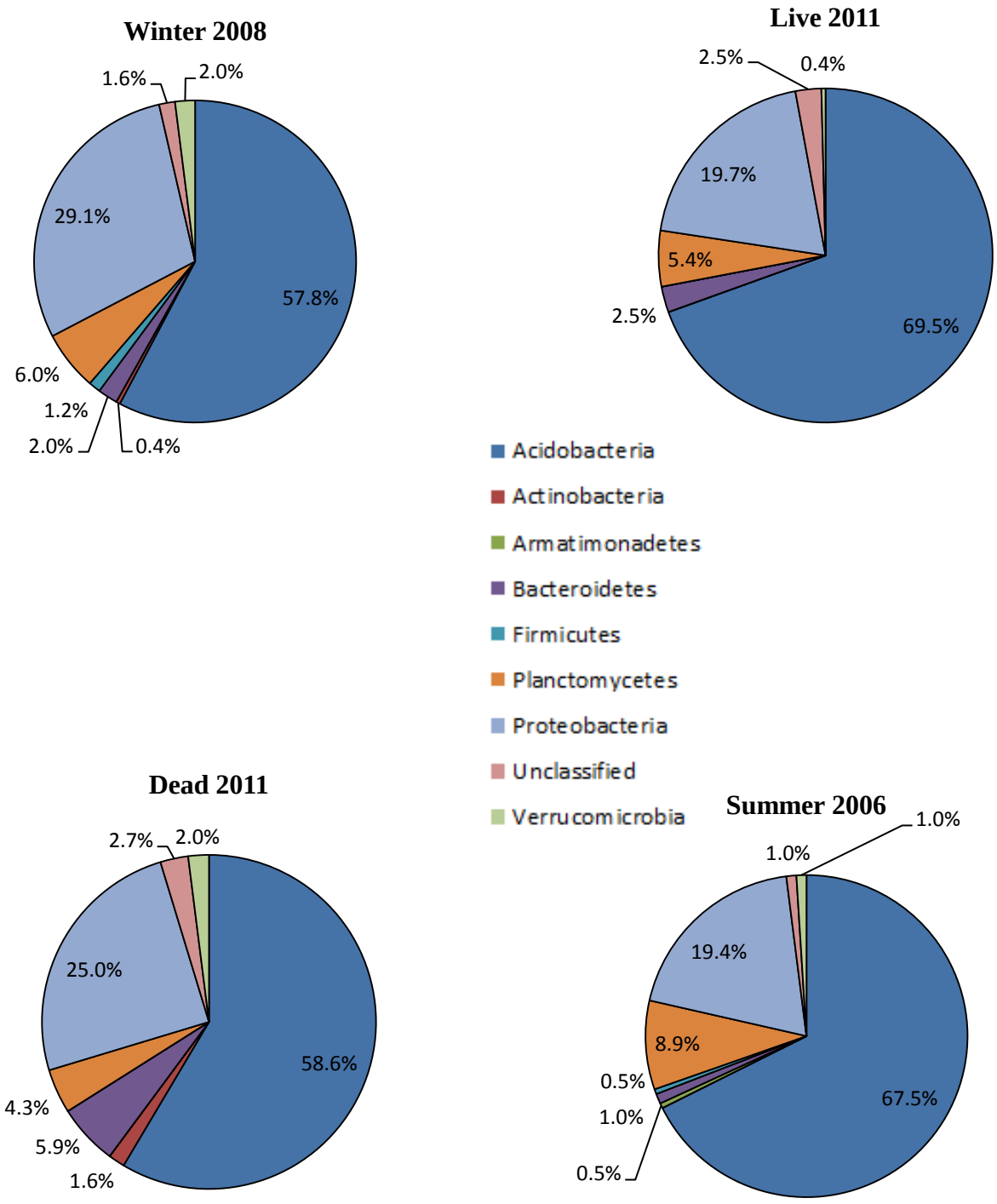

Figure 4. Ribosomal Database Project Classifier results for phylum level diversity of clones from Eastern Hemlock (*Tsuga canadensis*) rhizospheres from Albright Grove, Great Smoky Mountains National Park, showing percent of sequences assigned to each phylum for each sample.

Acidobacteria dominated in all samples, but made up almost 70% of sequences from the Live 2011 samples while consisting of ~58% of those from Winter 2008 (Figure 4). *Proteobacteria* made up ~29% of sequences from Winter 2008 and only ~20% from Live 2011 (Figure 4). Overall a total of eight phyla were found among all samples (Tables 1 and 3), as well as some sequences that were unclassified at the phylum level. The number of orders ranged from 14 in the Dead 2011 sample to nine from the Summer 2006 sample. Also, 27 families and 37 genera were found amongst all samples, with a fairly even distribution of number of different taxa among the various samples (Table 3). Live 2011 samples had a high of 20 genera, while the low was from Winter 2008 and Summer 2006 samples with 18 genera each.

All sequences were analyzed to the genus level using the RDP Classifier and the results are shown in Tables 4-8. *Acidobacteria Gp1* and *Gp2* were found 194 and 190 times, respectively, among all samples. *Acidobacteria Gp14* and *Gp5* were the least common *Acidobacteria* sequences found only occurring one and three times, respectively (Table 4). Within *Proteobacteria* the *Alphaprotebacteria* were most numerous among all samples and the *Deltaproteobacteria* were found the least. Unclassified *Rhizobiales* occurred 36 times overall, most frequently in Winter 2008 samples (Table 5). Within the *Betaproteobacteria*, sequences classified as *Burkholderia* occurred 16 times overall, but no sequences were found in the Dead 2011 samples. Sequences classified as *Dyella* (*Gammaproteobacteria*) occur 14 times, but none were found in the Summer 2006 sample (Table 5).

Sequences classified in phyla other than *Proteobacteria* and *Acidobacteria*, consisted of less than 9.0% in each sample. *Planctomycetes* were the third most dominant

phylum in all samples except Dead 2011, where *Bacteriodetes* was the third most dominant (Figure 4). Sequences classified as *Singulisphaera,* a genus within the phylum *Planctomycetes*, were numerous among all samples (36), but most common in Winter 2008 samples occurring 13 times. Sequences classified within the phylum *Firmicutes* were most rare among all sequences occurring only four times and never from samples in 2011. There were 19 sequences that could not be classified at the phylum level (Table 7). *Actinobacteria* were only found in Dead 2011 and Winter 2008 samples, occurring four times and one time, respectively (Table 8).

Table 3. Ribosomal Database Project Classifier results for taxonomic diversity for bacterial 16S rDNA sequences cloned from soil surrounding Eastern Hemlock (*Tsuga canadensis*) rhizosphere from Albright Grove, Great Smoky Mountain National Park, showing the number of sequences assigned to each classification.

	All	**Live 2011**	**Dead 2011**	**Winter 2008**	**Summer 2006**
Phylum	8	5	6	7	7
Class	20	15	15	16	15
Order	18	11	14	11	9
Family	27	15	16	16	12
Genus	37	20	19	18	18

Table 4. Ribosomal Database Project Classifier results for sequences classified as *Acidobacteria* for bacterial 16S rDNA sequences cloned from Eastern Hemlock (*Tsuga canadensis*) rhizosphere from Albright Grove, Great Smoky Mountains National Park, showing the number of sequences assigned to each classification down to the genus level. (The number range in parentheses (XX-XXX%) represents the confidence values assigned by the RDP Classifier and the whole number after the ":" represents the number of total sequences.)

Phylum: *Acidobacteria* (99-100%): 590						
Class	**Genus**	**Live 2011**	**Dead 2011**	**Winter 2008**	**Summer 2006**	**Total**
Acidobacteria_Gp1 (99-100%):256	*Gp1* (90-100%)	57	48	51	38	194
Acidobacteria_Gp1 (99-100%):256	*Granulicella* (95-100%)	4	3	4	3	14
Acidobacteria_Gp1 (99-100%):256	*Terriglobus* (99-100%)	0	0	1	3	4
Acidobacteria_Gp1 (99-100%):256	*Edaphobacter* (87-96%)	1	0	1	0	2
Acidobacteria_Gp1 (99-100%):256	Unclassified_ Gp1 (<80%)	19	9	10	4	42
Acidobacteria_Gp2 (96-100%):190	*Gp2* (96-100%)	51	52	48	39	190
Acidobacteria_Gp3 (100%):119	*Gp3* (82-100%)	28	32	24	35	119
Acidobacteria_Gp6 (100%):13	*Gp6* (100%)	2	4	4	3	13
Acidobacteria_Gp13 (100%):8	*Gp13* (100%)	2	1	2	3	8
Acidobacteria_Gp5 (100%):3	*Gp5* (100%)	2	1	0	0	3
Acidobacteria_Gp14 (97%):1	*Gp14* (97%)	0	0	0	1	1

Table 5. Ribosomal Database Project Classifier results for sequences classified as *Proteobacteria* for bacterial 16S rDNA sequences cloned from Eastern Hemlock (*Tsuga canadensis*) rhizospheres from Albright Grove, Great Smoky Mountains National Park, showing the number of sequences assigned to each classification down to the genus level. Sequences from the live and dead samples are labeled "L" and "D" respectively, while those sampled in the winter 2008 and summer 2006 are labeled "W" and "S," respectively. "T" is the total for all samples. (The number range in parentheses (XX-XXX%) represents the confidence values assigned by the RDP Classifier and the whole number after the ":" represents the number of total sequences.)

Phylum: *Proteobacteria* (80-100%):221							
Class: *Alphaproteobacteria* (98-100%):99							
Order	**Family**	**Genus**	**L**	**D**	**W**	**S**	**T**
Rhizobiales (91-100%):62	*Bradyrhizobiaceae* (100%):13	*Bradyrhizobium* (93-100%)	1	2	6	4	13
Rhizobiales (91-100%):62	*Hyphomicrobiaceae* (90-100%):6	*Devosia* (100%)	0	5	0	0	5
Rhizobiales (91-100%):62	*Hyphomicrobiaceae* (90-100%):6	*Rhodoplanes* (90%)	0	1	0	0	1
Rhizobiales (91-100%):62	*Methylocystaceae* (86-96%):4	Unclassified_Methylocystaceae (<80%)	2	0	0	2	4
Rhizobiales (91-100%):62	*Beijerinckiaceae* (81-100%):3	Unclassified_Beijerinckiaceae (<80%)	0	0	3	0	3
Rhizobiales (91-100%):62	Unclassified_Rhizobiales (<80%):36		1	8	21	6	36
Rhodospirillales (82-100%):29	*Acetobacteraceae* (100%):6	*Acidocella* (100%)	0	0	0	1	1
Rhodospirillales (82-100%):29	*Acetobacteraceae* (100%):6	Unclassified_Acetobacteraceae (<80%)	0	2	0	3	5
Rhodospirillales (82-100%):29	*Rhodospirillaceaen* (88-96%):2	Unclassified_Rhodospirillaceae (<80%)	1	0	1	0	2
Rhodospirillales (82-100%):29	Unclassified_Rhodospirillales (<80%):21		1	5	9	6	21
Caulobacterales (100%):2	*Caulobacteraceae* (100%):2	*Phenylobacterium* (97-100%)	2	0	0	0	2
Alphaproteobacteria_incertae_sedis (97%):1		*Rhizomicrobium* (97%)	0	0	0	1	1
Unclassified_Alphaproteobacteria (<80%):5			2	1	2	0	5

Class: *Gammaproteobacteria* (82-100%):73							
Order	**Family**	**Genus**	**L**	**D**	**W**	**S**	**T**
Xanthomonadales (81-100%):30	*Xanthomonadaceae* (100%):19	*Dyella* (92-100%)	5	5	4	0	14
Xanthomonadales (81-100%):30	*Xanthomonadaceae* (100%):19	*Rhodanobacter* (95%)	1	0	0	0	1
Xanthomonadales (81-100%):30	*Xanthomonadaceae* (100%):19	Unclassified_Xanthomonadaceae (<80%)	1	1	1	1	4
Xanthomonadales (81-100%):30	*Sinobacteraceae* (81-96%):11	*Steroidobacter* (80-95%)	3	4	3	0	10
Xanthomonadales (81-100%):30	*Sinobacteraceae* (81-96%):11	Unclassified_Sinobacteraceae (<80%)	0	1	0	0	1
Enterobacteriales (100%):7	*Enterobacteriaceae* (100%) :7	*Escherichia/Shigella* (99-100%)	1	0	5	0	6
Enterobacteriales (100%):7	*Enterobacteriaceae* (100%) :7	Unclassfied_Enterobacteriaceae (<80%)	0	0	1	0	1
Legionellales (100%):2	*Legionellaceae* (100%):1	*Legionella* (100%)	1	0	0	0	1
Legionellales (100%):2	*Coxiellaceae* (100%):1	*Coxiella* (100%)	0	1	0	0	1
Unclassified_Gammaproteobacteria (<80%):34			10	14	5	5	34
Class: *Betaproteobacteria* (96-100%):33							
Order	**Family**	**Genus**	**L**	**D**	**W**	**S**	**T**
***Burkholderiales* (83-100%):25**	*Burkholderiaceae* (96-100%):19	*Burkholderia* (96-100%)	8	0	5	3	16
Burkholderiales (83-100%):25	*Burkholderiaceae* (96-100%):19	Unclassified_Burkholderiaceae (<80%)	3	0	0	0	3
Burkholderiales (83-100%):25	*Oxalobacteraceae* (95-100%):4	Unclassified_Oxalobacteraceae (<80%)	1	2	1	0	4
Burkholderiales (83-100%):25	*Burkholderiales_incertae_sedis* (84%):1	Unclassified_Burkholderiales_incertae_sedis (<80%)	0	1	0	0	1
Burkholderiales (83-100%):25	Unclassified_Burkholderiales (<80%):1		0	1	0	0	1
Rhodocyclales (100%):1	*Rhodocyclaceae* (100%):1	Unclassified_Rhodocyclaceae (<80%)	0	1	0	0	1
Unclassified_Betaproteobacteria (<80%):7			1	4	2	0	7

Table 6. Ribosomal Database Project Classifier results for sequences classified as *Deltaproteobacteria* for bacterial 16S rDNA sequences cloned from Eastern Hemlock (*Tsuga canadensis*) rhizospheres from Albright Grove, Great Smoky Mountains National Park, showing the number of sequences assigned to each classification down to the genus level. Sequences from the live and dead samples are labeled "L" and "D" respectively, while those sampled in the winter 2008 and summer 2006 are labeled "W" and "S," respectively. "T" is the total for all samples. for bacterial 16S rDNA sequences cloned from Eastern Hemlock (*Tsuga canadensis*) rhizosphere from Albright Grove, Great Smoky Mountains National Park, showing the number of sequences assigned to each classification down to the genus level. (The number range in parentheses (XX-XXX%) represents the confidence values assigned by the RDP Classifier and the whole number after the ":" represents the number of total sequences.) Those classified in this table show the suborder classification.

Phylum: *Proteobacteria* (80-100%):221								
Class: *Deltaproteobacteria* (84-100%):14								
Order	**Suborder**	**Family**	**Genus**	**L**	**D**	**W**	**S**	**T**
Myxococcales (94-100%):9	*Sorangiineae* (97-100%):4	*Polyangiaceae* (100%):1	*Byssovorax* (92%)	1	0	0	0	1
Myxococcales (94-100%):9	*Sorangiineae* (97-100%):4	Unclassified_Sorangiineae (<80%):3		0	1	0	2	3
Myxococcales (94-100%):9	*Cystobacterineae* (89-97%):3	*Cystobacteraceae* (84%):1	Unclassified_Cystobacteraceae (<80%)	1	0	0	0	1
Myxococcales (94-100%):9	*Cystobacterineae* (89-97%):3	Unclassified_Cystobacterineae (<80%):2		0	1	1	0	2
Myxococcales (94-100%):9	*Nannocystineae* (93%):1	*Haliangiaceae* (82%):1	*Haliangium* (82%)	0	0	0	1	1
Myxococcales (94-100%):9	Unclassified_Myxococcales (<80%):1			0	1	0	0	1
Bdellovibrionales (84%):1		*Bdellovibrionaceae* (84%):1	*Bdellovibrio* (84%)	0	1	0	0	1
Unclassified_Deltaproteobacteria (<80%):4				0	0	2	2	4

Table 7. Ribosomal Database Project Classifier results for sequences classified in phyla other than *Acidobacteria*, *Proteobacteria* or *Actinobacteria* for bacterial 16S rDNA sequences cloned from soil surrounding Eastern Hemlock (*Tsuga canadensis*) rhizosphere from Albright Grove, Great Smoky Mountain National Park, showing the number of sequences assigned to each classification down to the genus level. Sequences from the live and dead samples are labeled "L" and "D" respectively, while those sampled in the winter 2008 and summer 2006 are labeled "W" and "S," respectively. "T" is the total for all samples. (The number range in parentheses (XX-XXX%) represents the confidence values assigned by the RDP Classifier and the whole number after the ":" represents the number of total sequences.)

Phylum: *Planctomycetes* (99-100%):56								
Class	**Order**	**Family**	**Genus**	**L**	**D**	**W**	**S**	**T**
Planctomycetacia (99-100%):56	*Planctomycetales* (99-100%):56	*Planctomycetaceae* (99-100%):56	*Singulisphaera* (93-100%)	5	5	13	8	31
***Planctomycetacia* (99-100%):56**	*Planctomycetales* (99-100%):56	*Planctomycetaceae* (99-100%):56	*Gemmata* (82-99%)	0	1	0	1	2
***Planctomycetacia* (99-100%):56**	*Planctomycetales* (99-100%):56	*Planctomycetaceae* (99-100%):56	*Planctomyces* (94-100%)	1	0	1	0	2
***Planctomycetacia* (99-100%):56**	*Planctomycetales* (99-100%):56	*Planctomycetaceae* (99-100%):56	Unclassified_Planctomycetaceae (<80%)	7	5	1	8	21
Phylum: *Bacteroidetes* (100%):28								
Sphingobacteria (96-100%):26	*Sphingobacteriales* (96-100%):26	*Chitinophagaceae* (91-100%):17	*Ferruginibacter* (100%)	0	1	0	0	1
Sphingobacteria (96-100%):26	*Sphingobacteriales* (96-100%):26	*Chitinophagaceae* (91-100%):17	Unclassified_Chitophagaceae (<80%)	4	6	4	2	16
Sphingobacteria (96-100%):26	*Sphingobacteriales* (96-100%):26	*Sphingobacteriaceae* (100%):9	*Mucilaginbacter* (100%)	0	8	1	0	9
Flavobacteria (100%):1	*Flavobacteriales* (100%):1	*Flavobacteriaceae* (100%):1	*Flavobacterium* (100%)	1	0	0	0	1
Unclassified_Bacteroidetes (60%):1				1	0	0	0	1
Phylum: *Verrucomicrobia* (98-100%):13								
*Subdivision3 (*100%):10		*Subdivision3_genera_incertae_sedis* (100%)		1	4	4	1	10

Spartobacteria (85-100%):2			*Spartobacteria_genera_incertae_sedis* (84-100%)	0	0	1	1	2
Opitutae (100%):1	*Opitutae* (100%):1	*Opitutaceae* (100%):1	*Opitutus* (100%)	0	1	0	0	1
Phylum: *Firmicutes* (81-100%):4								
Clostridia (80-87%):3	Unclassified_Clostridia (56-73%):3			0	0	2	1	3
Bacilli (100%):1	*Bacillales* (100%):1	*Bacillaceae* (100%):1	*Bacillus* (99%)	0	0	1	0	1
Bacilli (100%):1	*Bacillales* (100%):1	*Bacillaceae* (100%):1	*Armatimonadetes_gp5* (99%)	0	0	0	1	1
Unclassified Bacteria (<80%):19				6	7	4	2	19

Table 8. Ribosomal Database Project Classifier results for sequences classified in the phylum *Actinobacteria* within the class *Acidimicobidae* for bacterial 16S rDNA sequences cloned from Eastern Hemlock (*Tsuga canadensis*) rhizosphere from Albright Grove, Great Smoky Mountains National Park, showing the number of sequences assigned to each classification down to the genus level at the cutoff of 95%. Sequences from the live and dead samples are labeled "L" and "D" respectively, while those sampled in the winter 2008 and summer 2006 are labeled "W" and "S," respectively. "T" is the total for all samples. Those classified in this table show the subclass and suborder classification, which are commonly used for *Actinobacteria*. (The number range in parentheses (XX-XXX%) represents the confidence values assigned by the RDP Classifier and the whole number after the ":" represents the number of total sequences.)

Phylum: *Actinobacteria* (100%):5 **Class: *Actinobacteria* (100%):5**									
Subclass	**Order**	**Suborder**	**Family**	**Genus**	**L**	**D**	**W**	**S**	**T**
Acidimicrobidae (82-99%):4		*Acidimicrobineae* (82-99%):4			0	0	1	0	1
Acidimicrobidae (82-99%):4			Unclassified_Acidimicrobineae (<80%):4		0	3	0	0	3
Actinobacteridae (100%):1	*Actinomycetales* (100%):1	*Streptosporangineae* (96%):1	*Thermomonosporacea e* (96%):1	Unclassfied_Thermomonosporaceae (<80%)	0	1	0	0	1

Based on distributional patterns of sequences, *Acidobacteria* was significantly different among Live 2011 versus Dead 2011, Winter 2008 versus Summer 2006 and Winter 2008 versus Live 2011 (Table 9). Significant differences also occurred between samples within *Proteobacteria*. These occurred in *Alphaproteobacteria* between four of the six comparisons including Live 2011 versus Dead 2011 and in *Rhizobiales* between all comparisons except Summer 2006 versus Dead 2011 (Table 9). *Betaproteobacteria* including the genus *Burkholderia* and *Gammaproteobacteria,* including the order *Xanthomonadales* and family *Enterobacteriaceae,* also significantly differed among samples (Table 9). Another group differing significantly among samples was the genus *Mucilaginibacter* within the *Bacteroidetes*; no sequences were found within this group among Live 2011 and Summer 2006 samples (Table 9).

Table 9. Ribosomal Database Project Library Compare results showing distributed patterns of taxa from clone libraries obtained from Eastern Hemlock (*Tsuga canadensis*) rhizospheres from Albright Grove, Great Smoky Mountains National Park, that were significantly different between microbial communities. (Comparisons are labeled as follows: a-Live 2011 vs Dead 2011, b- Summer 2006 vs Dead 2011, c-Summer 2006 vs Live 2011, d- Summer 2006 vs Winter 2008, e-Winter 2008vs Live 2011, f-Winter 2008 vs Dead 2011)

Samples	**Live 2011 (239)**	**Dead 2011 (256)**	**Winter 2008 (251)**	**Summer 2006 (191)**	**Significance**
Taxon					
Acidobacteria	166	150	145	129	a,d,e
Acidobacteria_Gp1	81	60	67	48	a,c
Acidobacteria_Gp3	28	32	24	35	d
Gp3	28	32	24	35	d
Bacteroidetes	6	15	5	2	b,f
Sphingobacteria	4	15	5	2	a,f
Sphingobacteriales	4	15	5	2	a,b,f
	0	8	1	0	a,f
Sphingobacteriaceae					
Mucilaginibacter	0	8	1	0	a,b,f
Planctomycetes	13	11	15	17	b
Planctomycetacia	13	11	15	17	b
Planctomycetales	13	11	15	17	b
Planctomycetaceae	13	11	15	17	b
Proteobacteria	47	64	73	37	d,e
Alphaproteobacteria	10	24	42	23	a,c,e,f
Rhizobiales	4	16	30	12	a,c,d,e,f
Hyphomicrobiaceae	0	6	0	0	a,b,f
Devosia	0	5	0	0	a,f
Rhodospirillales	2	7	10	10	c,e
Acetobacteraceae	0	2	0	4	c,d
Betaproteobacteria	13	9	8	3	c
Burkholderiales	12	4	6	3	a
Burkholderiaceae	11	0	5	3	a,f
Burkholderia	8	0	5	3	a,f
Gammaproteobacteria	22	26	19	6	c,d
Enterobacteriales	1	0	6	0	d,f
Enterobacteriaceae	1	0	6	0	d,f
	1	0	5	0	f
Esherichia/Shigella					
Xanthomonadales	10	11	5	1	b,c

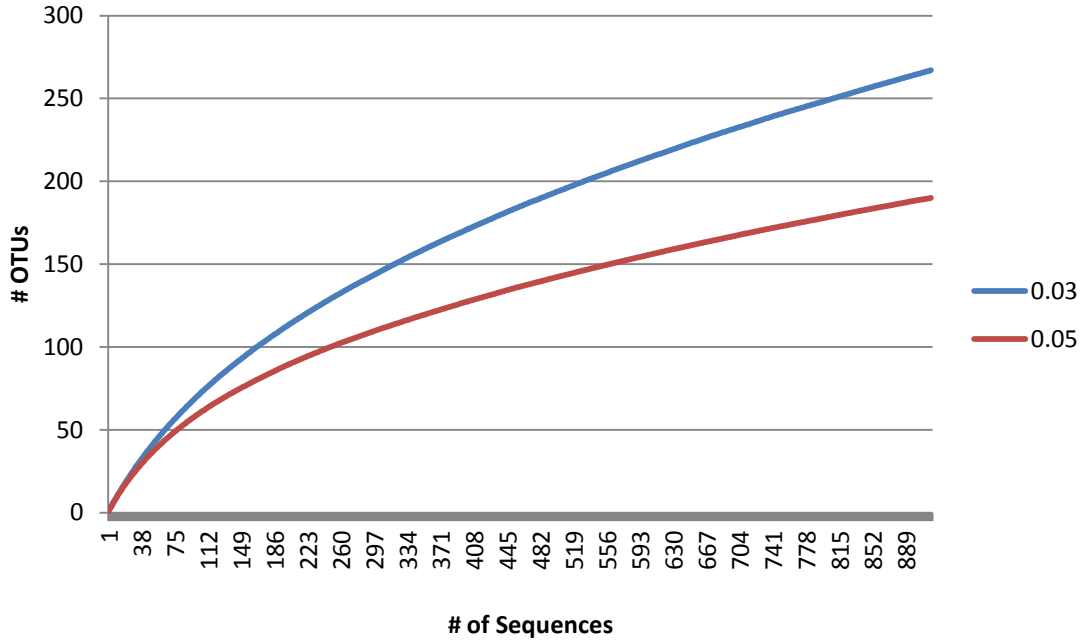

Figure 5. Rarefaction curve showing all bacterial sequences from Eastern Hemlock rhizospheres based on number of operational taxonomic units (OTUs) at the genus (0.05) and species (0.03) level as computed using DOTUR.

Species richness among all samples produced 267 unique operational taxonomic units (OTU), defined as those sequences sharing ≥ 97% similarity (Figure 5). Rarefaction curves of all sequences shows species richness is likely to continue to increase with more samples. Species richness is lowest from samples from Summer 2006 and highest from Winter 2008 (Figure 6). There were 115 unique OTUs in the sequences from the Winter 2008 sample, 110 unique OTUs from the Dead 2011 sample, 99 unique OTUs from the Live 2011 and 71 unique OTUs from the Summer 2006 sample.

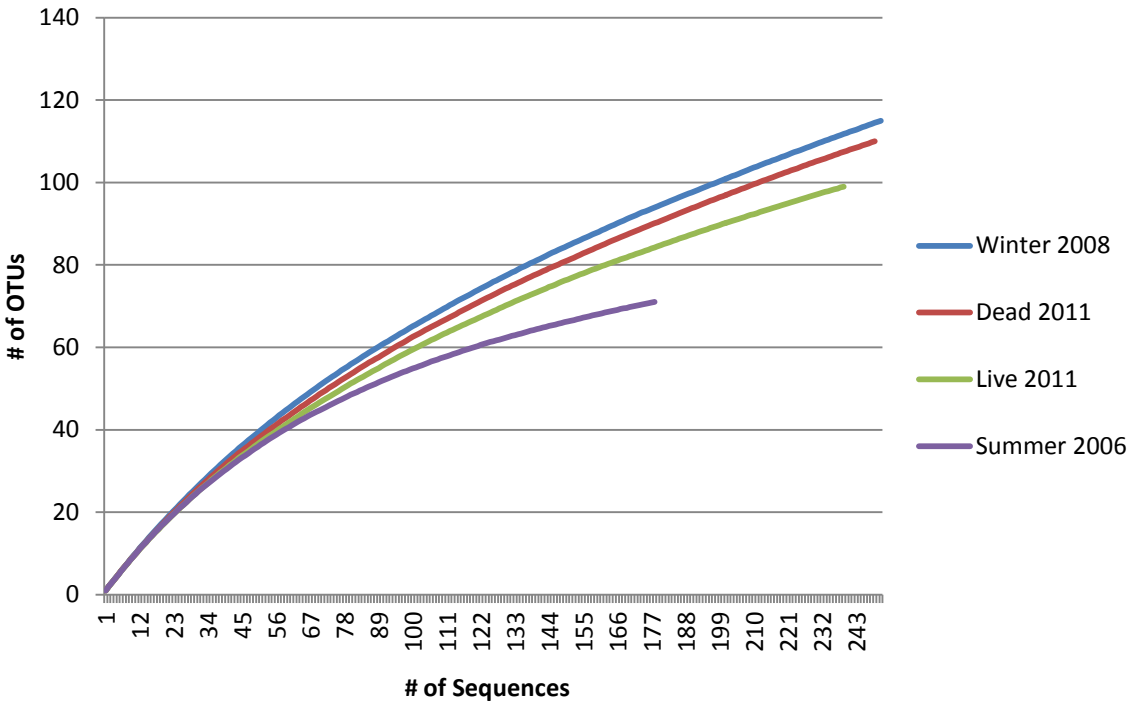

Figure 6. Rarefaction curve showing the number of OTUs based on collection effort for Eastern Hemlock rhizosphere for bacterial sequences in each sample at the 97% cutoff as computed using DOTUR.

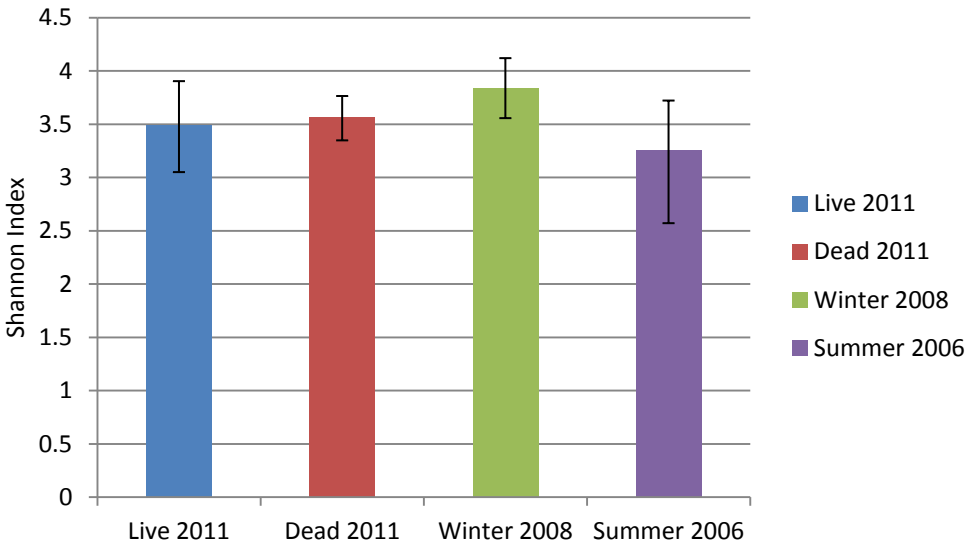

Figure 7. Bar graph showing the average Shannon Index (H) with 95% confidence intervals as calculated using DOTUR from replicates within each sample of bacterial sequences cloned from Eastern Hemlock rhizospheres (n=3).

Shannon Index (H) values were calculated using DOTUR for each replicate (3 per sample) then averaged for each sample set. The H values ranged from 3.25 to 3.83. Average values were highest for samples from Winter 2008 and lowest for those from Summer 2006 (Figure 7).

Chao values for species richness estimates for each sample were calculated using DOTUR for each replicate (n=3) and averaged for each sample (Figure 8). Average Chao values were 130.42 for Winter 2008, 84.58 for Live 2011, 72.79 for Dead 2011 and 46.21 for Summer 2006. Confidence values at the 95% level are displayed in Figure 8 and are as high as 213 for Live 2011 and as low as 38 for Summer 2006.

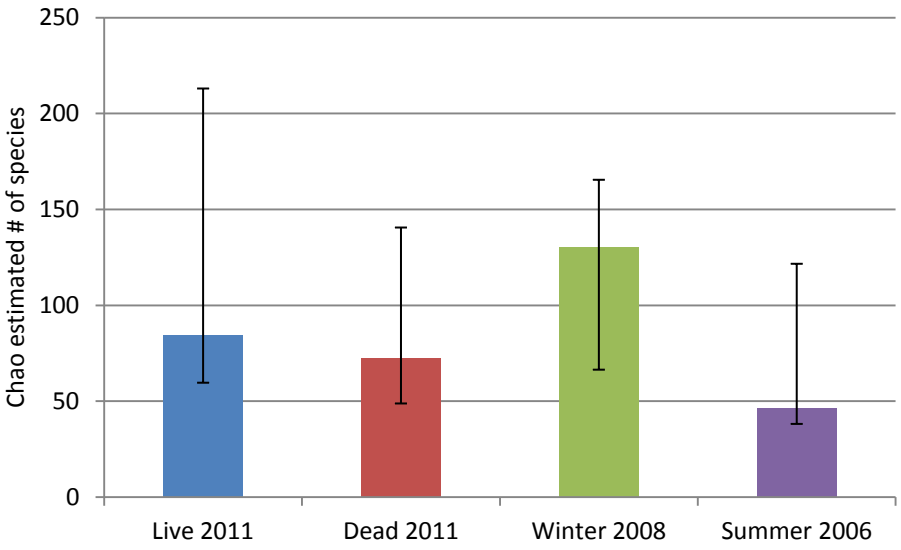

Figure 8. Bar graph showing the average Chao values with 95% confidences intervals as calculated using DOTUR from replicates within each sample of bacterial sequences cloned from Eastern Hemlock rhizospheres (n=3).

Rank abundance data calculated using DOTUR shows that within the 267 unique sequences, 27 sequences occurred in dominant OTUs, which occurred twice (Table 10). The dominant OTUs are named for convenience, *Acidobacteria Gp2*_A and *Acidobacteria Gp3*_A. Of the 27 sequences in the dominant OTU *Acidobacteria Gp3*_A, no sequences were from samples taken in Summer 2006, while the other OTU was more evenly distributed for all samples. The next dominant OTU with 26 sequences was found and named *Acidobacteria Gp2*_B. There were 137 singletons among all sequences representing the extreme of rare OTUs detected. Thirty-eight OTUs occurred as doublets and twenty-six as triplets, while others were observed between four and nine times. The OTU containing 18 sequences was determined to be unclassified within the order *Rhizobiales* and no sequences within this OTU were from the Live 2011 sample. One of the OTUs containing 15 sequences was determined to be *Burkholderia* and no sequences were from the Dead 2011 sample. The OTU containing 11 sequences was determined to

be *Dyella* with no sequences from the Summer 2006 sample and only one was from the Winter 2008 sample (Table 10). Overall, both *Acidobacteria Gp1* and *Gp2* were represented in eight OTUs each, accounting for 16 of the 26 OTUs containing 10 or more sequences. *Acidobacteria Gp3* was represented by two OTUs and *Acidobacteria Gp6, Terriglobus* and *Granulicella* by one OTU each.

Table 10. Dominant OTUs displaying the number of times each sequence occurred per sample, as determined by DOTUR. Classification is based on Ribosomal Database Project Classifier, with "_A, _B, etc." representing multiple OTUs within a particular taxonomic group. ("*****" indicates that classification is based on sequences that Classifier returned that genera ≥80% match, but not all sequences within that OTU were classified to that level.)

OTU/ # of Sequences	**Genus/Taxon**	**Live 2011**	**Dead 2011**	**Winter 2008**	**Summer 2006**
27	*Acidobacteria Gp2*_A	7	8	7	5
27	*Acidobacteria Gp3*_A	15	8	4	0
26	*Acidobacteria Gp2*_B	15	3	6	2
20	*Acidobacteria Gp2*_C	6	7	6	1
19	*Acidobacteria Gp1*_A	1	9	3	6
18	*Unclassified_Rhizobiales (Proteobacteria)*	0	5	10	3
18	*Acidobacteria Gp2*_D	3	2	8	5
16	*Acidobacteria Gp3*_B	1	7	4	4
15	*Terriglobus* (Acidobacteria)*	3	0	4	8
15	*Burkholderia (Proteobacteria)*	7	0	5	3
15	*Acidobacteria Gp1*_B	5	5	1	4
15	*Acidobacteria Gp1*_C	4	5	2	4
15	*Acidobacteria Gp2*_E	4	7	2	2
14	*Acidobacteria Gp1*_D	5	5	3	1
13	*Bradyrhizobium* (Proteobacteria)*	1	2	6	4
13	*Acidobacteria Gp1*_E	3	1	5	4
12	*Acidobacteria Gp1*_F	4	4	3	1
12	*Acidobacteria Gp2*_F	3	4	1	4
12	*Unclassified_Gammaproteobacteria*	3	5	1	3
11	*Granulicella* (Acidobacteria)*	2	3	4	2
11	*Dyella (Proteobacteria)*	5	5	1	0
11	*Acidobacteria Gp1*_G	1	3	6	1
11	*Acidobacteria Gp1*_H	5	2	4	0
11	*Acidobacteria Gp2*_G	3	3	0	5
10	*Acidobacteria Gp6*	1	4	2	3
10	*Acidobacteria Gp2_H*	2	2	2	4

DISCUSSION

Hypothesis one stated that bacterial diversity would be higher in the rhizospheres of living Eastern Hemlock versus dead hemlock in the same plots. This was not supported by the data in this study. Rarefaction curves show that bacterial diversity is likely to continue to increase in samples from both living and dead rhizospheres (Figure6). The Dead 2011 samples actually contained the second most unique OTUs in this study with 110. Estimated numbers of species using the Chao equation were higher in living trees sampled in 2011, however, confidence intervals overlapped among all samples. Shannon index values were higher in the Dead 2011 samples than in the Live 2011 samples, showing more species richness within the Dead 2011 samples, but no significant differences were found. However, bacterial communities associated with paired living and dead Eastern Hemlock rhizospheres sampled in 2011 were significantly different in composition of phyla indicated by Pair-wise log likelihood ratio tests (G-test) of independence (Table 2). Upon review of the data it seems that diversity among rhizospheres of living and dead Eastern Hemlock have similar levels of bacterial species richness, but the composition of the microbial communities associated with these rhizospheres will be different, when sampled in the same general location and at the same time. This is consistent with the second hypothesis in my study.

Hypothesis two stated that bacterial diversity in the rhizospheres of living Eastern Hemlock would be equal independent of season and years sampled, this was supported by this study. Average values for Chao and the Shannon index were slightly lower in samples collected from Summer 2006, but no significant differences were found. There were 71 unique OTUs found in the Summer 2006 sample, 110 from the Winter 2008

sample and 99 from the Live 2011 sample. The lower number of unique OTUs in the Summer 2006 sample could be due to a smaller clone library with 191 clones, whereas the Live 2011 sample had 239 and Winter 2008 had 251 clones, respectively. Rarefaction curves show that diversity is likely to increase among all living samples, as additional samples are generated.

In a study that was analogous to this one, Filion et al. (2004) profiled rhizosphere microbial communities associated with healthy and diseased Black Spruce (*Picea mariana*) seedlings grown in a nursery, using molecular culture-independent approaches. They found the majority of 16S rDNA clones from healthy seedlings were from *Proteobacteria* (27%) and *Acidobacteria* (25%); in the diseased seedling the *Proteobacteria* still dominated making up 22%, but *Acidobacteria* only consisted of 16% of the clone library (Filion et al. 2004). Shannon index values in this study were higher in healthy seedlings (4.036) than diseased seedlings (3.566), suggesting a higher species richness in healthy seedlings. Filion et al. found 84 different OTUs among the 300 cloned inserts, 34 OTUs (of the 84) were associated only with the rhizosphere of healthy seedlings, 24 OTUs were only associated with the rhizosphere of diseased seedlings, while 24 OTUs overlapped between both samples.

In another study Chow et al. (2002) characterized bacterial communities associated with the rhizosphere of Lodgepole Pine (*Pinus contorta*) in forest soils differing in disturbance (whole-tree harvesting without soil compaction and with heavy soil compaction) and geographic location, using much of the same methods as in this study. Their results show no significant differences between disturbance areas (Chow et. al 2002). Overall *Proteobacteria* and *Acidobacteria* were dominant making up 55% and

19% of the clone library, respectively (Chow et al. 2002). Within the *Proteobacteria, α-Proteobacteria* and *β-Proteobacteria* made up 24% and 19% of the libraries, respectively (Chow et al. 2002). *Actinobacteria* comprised 3% of the clone library while unclassified bacteria were prevalent making up 15% of all clones (Chow et. al 2002).

In the studies completed by Filion et al. (2004) and Chow et al. (2002) on the rhizospheres of other conifers, *Proteobacteria* was the dominant phylum, followed by *Acidobacteria.* In a review of the literature, Janssen (2006) showed that these two phyla show global trends as being the most dominant taxa in soils. These two phyla were the two most dominant found in this study as well, however *Acidobacteria* was the most dominant with sequences occurring nearly three times for every one time *Proteobacteria* was found. The differences between the three studies could be due to a number of factors including plant species type or other influences within the rhizosphere of these conifers. In the study of Black Spruce, trees were grown in a greenhouse setting in containers, thus limiting the potential pool of bacteria to colonize the rhizosphere. *Acidobacteria* at the time of these studies may not have been as numerous in databases such as GenBank and RDP, therefore yielding more unclassified bacteria that could be classified as *Acidobacteria*. Plant species and soil type are the main influences on structure and function of microbial communities in the rhizosphere, but which is the dominant control is unclear (Berg and Smalla, 2008), and one of these could account for differences between studies. Many other factors that influence microbial communities in the rhizosphere are shown in Figure 9.

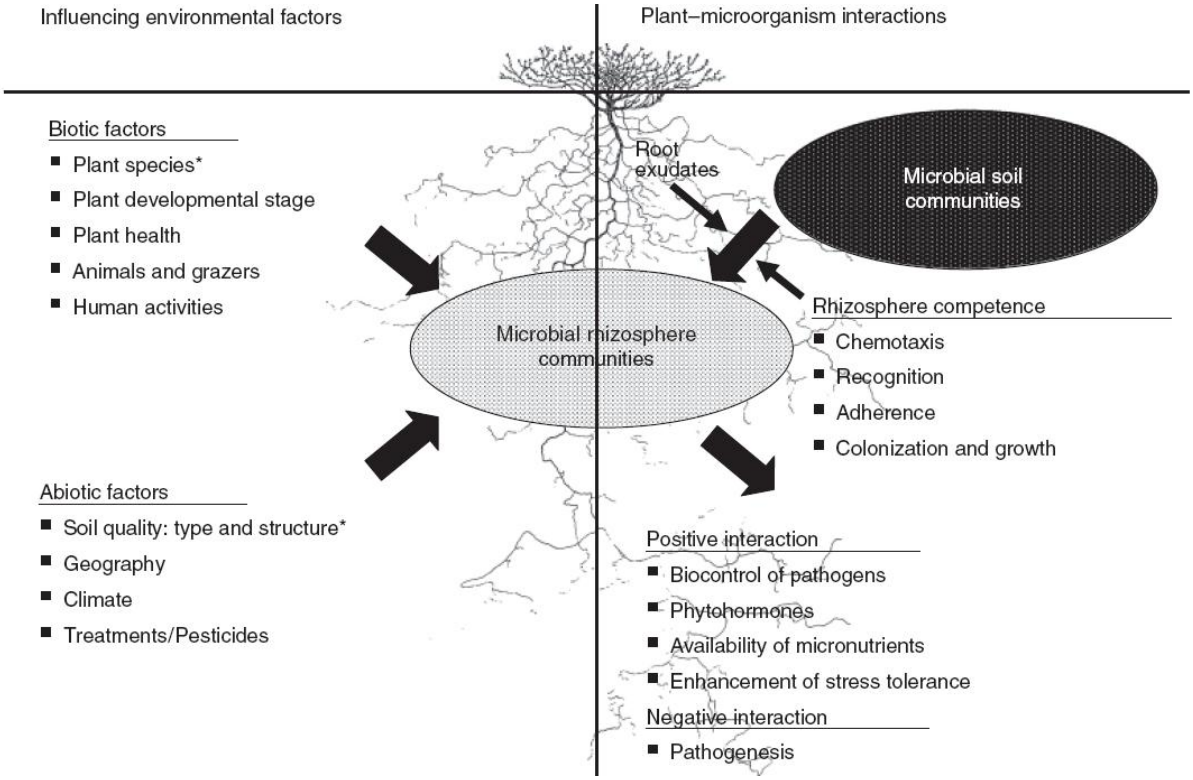

Figure 9. Multitude of factors influencing rhizosphere microbial communities (Figure taken from Berg and Smalla, 2008).

Acidobacteria dominated the clone libraries in this study, making up 69.5% of the Live 2011 clones, 58.6% of Dead 2011 clones, 57.8% of Winter 2008 clones and 67.5% of Summer 2006 clones (Figure 4). *Acidobacteria* were identified based on analysis of 16S rRNA gene sequences and 'currently contain two classes, three orders, three families and six described genera' (Thrash and Coates 2011). There are more than 3,000 sequences making up 26 groups in public databases, though most are uncultivated and little is known about their physiological and ecological roles (Barns et al. 2007). *Acidobacteria* have been found in a multitude of habitats around the world such as freshwater habitats, arid Arizona soils, hot spring microbial mats and sewage sludge (Quaiser et al. 2003). Diversity within the phylum and the abundance of *Acidobacteria*

species mainly in soil habitats suggests an important ecological niche (Quaiser et al. 2003).

Acidobacteria have also been shown to be dominant in rhizosphere of a chestnut tree (*Castanea crenata*), accounting for 65% of the sequences in clone libraries (Lee et al. 2008). *Acidobacteria* groups 1, 2 and 3 accounted for 22.5%, 36.7% and 6.1% of clone libraries, respectively (Lee et al. 2008). Further analysis by Lee et al. revealed that *Acidobacteria* were not only dominant in overall numbers, but also metabolically active within the rhizosphere of the chestnut trees (Lee et al. 2008). They constructed a complementary ribosomal RNA (crDNA) library, in which they found *Acidobacteria* to be present in similar composition to the 16S rRNA clone library, suggesting high metabolic activity (because of a high number of ribosomes from which the rRNA was cloned) (Lee et al. 2008).

Within *Acidobacteria, Gp1* and *Gp2* were most prevalent across all samples from this study of Eastern Hemlock rhizospheres, occurring 194 and 190 times, respectively. There were also 42 unclassified *Acidobacteria_Gp1* (Table 4). The two most dominant OTUs were *Gp2*_A and *Gp3*_A containing 27 sequences each from all samples (Table 10). The numbers of sequences from each of the four samples were similarly spread across samples in the OTU classified as *Gp2.* However, of the 27 sequences in the OTU classified as *Gp3* none were from Summer 2006 samples and only four were from Winter 2008 samples. There were 15 sequences from Live 2011 and eight from Dead 2011 (Table 10). This could suggest a seasonal trend or a trend in a changing structure in the bacterial community associated with the HWA infestation.

Groups within *Acidobacteria* account for 19 of the 26 OTUs containing more than ten sequences. *Gp1,* which contained the most sequences within *Acidobacteria* in this study were grouped into eight different OTUs (of the 26 OTUs with more than ten sequences) suggesting there is high diversity within that group. *Gp2* was also grouped into eight different OTUs, with fewer sequences (n=190), this could suggest even higher diversity (Table 10). *Terriglobus* within *Acidobacteria_Gp1* was found as an OTU with 15 sequences, with no sequences found in Dead 2011 samples. RDP Classifier only identified four total sequences as *Terriglobus* of $\geq$ 97% similarity, but the OTU contained 15 total sequences of $\geq$ 97% sequence similarity (Table 10). This is an example of the uncertain classification within *Acidobacteria* and the need for further work on taxonomic classification within this group. *Acidobacteria* made up 590 of 937 total sequences in this study, but the lack of culture based or other information on these organisms makes it difficult to speculate on the roles they are playing for Eastern Hemlock within GSMNP. Sheer numbers lead to conjecture that *Acidobacteria* have major ecological roles, but exactly what they may be, is undetermined. Through metagenomic analysis in another study, one uncultured *Acidobacterium* fosmid clone was shown to have homology to genes from members of *Rhizobiales* (an order shown to have nitrogen-fixing members; Quaiser et al. 2003). While this could be the result of horizontal gene transfer, it suggests that perhaps some acidobacterial species are nitrogen fixers and could be assisting hemlocks in their nutritional needs (Quaiser et al. 2003, Andreote et al. 2009).

Proteobacteria were the next most dominant phyla found among all samples and ranged from 19.4% to 29.1% found in Summer 2006 and Winter 2008 samples, respectively (Figure 4). The phylum *Proteobacteria* contains the largest number of

known species, with many recognizable species such as *Escherichia coli*, *Salmonella enteritidis* and *Rhizobium leguminosarum* (Baker et al. 2007). *Proteobacteria* are divided into the Alpha, Beta, Gamma, Delta and Epsilon classes. In this study, *Alphaproteobacteria* were found most often among the *Proteobacteria* occurring 99 times out of the total 222 sequences, while *Deltaproteobacteria* occurred least often with only 14 total sequences throughout all samples (Table 5, 6). LibCompare found significant differences between samples among several different classes and genera within the *Proteobacteria. Rhizobiales* within the *Alphaproteobacteria* was significantly different between all six sample pairings except Summer 2006 versus Dead 2011 (Table 9, Appendix A). *Devosia,* a genus within the *Rhizobiales,* was found five times and only among the dead samples. *Burkholderia* within the class *Betaproteobacteria* was not found in the Dead 2011 samples, but occurred in all others. Within the *Gammaproteobacteria, Escherichia/Shigella* was not found in Dead 2011 or Summer 2006 samples, but was found five times in Winter 2008 samples and once in Live 2011 samples. *Xanthomonadales* within the *Gammaproteobacteria* also significantly differed among groups (Appendix A, Table 9).

Some of the more dominant OTUs also came from the *Proteobacteria,* including unclassified *Rhizobiales, Burkholderia*, *Bradyrhizobium*, unclassified *Gammaproteobacteria* and *Dyella* (Table 10). One of the OTUs containing 13 sequences was classified as *Bradyrhizobium (Alphaproteobacteria),* which is a slow growing gram negative bacteria that has host-specific nitrogen fixation and nodulation genes; nitrogenase activity has been shown to occur in free living cells as well (Kuykendall

2011). Of the 13 times *Bradyrhizobium* was found throughout all samples, six were from Winter 2008 and four from Summer 2006.

Another dominant OTU, containing 18 sequences was determined to be unclassified bacteria within the order *Rhizobiales (Alphaproteobacteria)* (Table 10). *Rhizobiales* is an order containing nitrogen-fixing groups such as *Rhizobium* and other groups like the *Methylobacterium* that are known to be in association with plants and contain nitrogenase genes (Andreote et al. 2009). Given that 18 of the 36 unclassified *Rhizobiales* are in the same OTU it could represent a novel taxon within this group (Table 10). The sequences within this OTU occurred in two of three replicates for the Dead 2011 and Summer 2006 samples and in all replicates from the Winter 2008 samples. No sequences within this OTU were from the Live 2011 samples, and only one other sequence from the Live 2011 samples was an unclassified *Rhizobiales* (Table 5, 10).

Burkholderia is one of the only genera of *Betaproteobacteria* that was shown to be significantly different between Live 2011/Dead 2011 and Winter 2008/Dead 2011 (Table 9, Appendix A). One of the more dominant OTUs containing 15 sequences was classified as *Burkholderia*, which leaves one unique sequence within this genus from the Live 2011 samples (Table 10). Members within this genus have been shown to promote plant growth and have been useful in bioremediation; *Burkholderia cepacia* is a natural colonizer of the rhizosphere and has anti-fungal and anti-nematodal properties (Woods and Sokol, 2006). Richardson et al. illustrated in 2002 that isolates of *B. cepacia* are "genetically diverse but that woodland rhizospheric isolates are related to each other and unrelated to plant or human pathogenic strains." *Burkholderia* occurred in all samples

except Dead 2011 (Table 5), and possible reasoning could be that these organisms rely on photosynthates released in the rhizosphere.

Dyella is a novel genus within *Gammaproteobacteria* described by Xie and Yokota in 2005, in which they isolated three novel strains from garden soil in Tokyo, Japan. An OTU with 11 sequences was classified as *Dyella* in this study and no sequences from this genus were found in the samples from Summer 2006 (Table 10). Sequences within this genus were found in all replicates from the Winter 2008 and Live 2011 samples and two of the three replicates from the Dead 2011 samples. The lack of information available for *Dyella* and the common distribution across samples from this study exemplifies the need for further research before any speculation of an ecological role of these organisms can be made. However, like many of the other proteobacteria menitoned above, *Dyella spp.* have been cultured on heterotophic growth media.

Mucilaginibacter within the phylum *Bacteroidetes* showed an interesting distribution across samples occurring eight times in the Dead 2011 samples, once in Winter 2008 and it did not occur in the Summer 2006 or Live 2011 samples (Table 9). *Mucilaginibacter* was found in two of three replicates in the Dead 2011 samples and it was significantly different (as indicated by RDP Library Compare Tool) between Live and Dead 2011, tree replicate number 2 (Appendix B). Trees sampled in 2006 and 2008 did not have visible signs of HWA infestation so the organisms within this genus found in the Dead 2011 samples could represent a changing bacterial community after the presumed death of Eastern Hemlock caused by HWA infestation. *Mucilaginibacter* is a genus proposed in 2007 by Pankratov et al. (2007) and it has been amended several times. *Mucilaginibacter* currently, contains 10 species, in which one strain has been

isolated from the rhizosphere of *Platycodon gradiflorum,* also known as Doraji, and has been shown to utilize carbon sources through the hydrolysis of cellulose and starch (Kim et al. 2010).

The phylum *Planctomycetes* was the third most common phylum found within all samples (Table 7). Specifically sequences in the genus *Singulisphaera* occurred 31 times, but none of the OTUs with more than ten sequences were classified as *Planctomycetes.* This suggests that there is a high level of species richness within this genus in this study, most likely species that have yet to be described. *Singulisphaera* was evenly distributed across all samples and replicates, as it was found in all replicates, in all samples except for one replicate (tree 3) from the Live 2011 samples. *Singulisphaera* is a novel genus described in 2008 by Kulichevskaya et al. (2008) based on four isolates from acidic wetlands in northern Russia. Zaicnikova et al. (2010) have also described a novel species, *Singulisphaera mucilagenosa*, isolated from low mineral water formed by xylotrophic fungi (those that have the ability to consume wood as an energy source), on decaying spruce wood. It is interesting that, other than *Acidobacteria* groups that are not well classified, *Singulisphaera* is the genus found most commonly among all samples. As a novel genus little is known about their ecological role, but with the frequency found in this study across all samples it does suggest they are likely filling some type of role in the microbial community of the rhizosphere of Eastern Hemlock.

G-tests showed significant differences in comparisons of sequences assigned at the phylum level between samples from Live 2011 and Dead 2011, Live 2011 and Winter 2008, and Summer 2006 and Dead 2011 (Table 2). The comparisons could suggest microbial communities associated with the rhizospheres of living Eastern Hemlock do

not vary by season, as the only significant difference found with Summer 2006 samples was with Dead 2011 samples. Significant differences between Live and Dead 2011 samples in bacterial community composition is likely due to groups such as *Mucilaginabacter* (not present in Live 2011 samples), unclassified *Rhizobiales* (only one sequence in Live 2011 samples) and *Burkholderia* (not present in Dead 2011 samples). Significant differences were found between Dead 2011 samples and two of the three living samples, suggesting that bacterial community composition in rhizospheres of living and dead Eastern Hemlock was different more often than not. Summer 2006 samples were processed using a different DNA extraction kit (see methods), but this should not have affected any of the significant differences found. The PowerSoil kit used in 2006 had larger beads for shearing cells, but resulted in overall similar recoveries of DNA sequences as found in 2008 and 2011 samples, which used the PowerLyser PowerSoil kit. Paired replicates of Live 2011 and Dead 2011 samples were compared using the RDP Library Compare tool and no pairs were found to have significant differences between the same groups (Appendix B)

Bacterial identification and assessment of microbial diversity is important for the understanding of biochemical cycling and use of biocontrols to promote plant growth, among many other reasons, and methods and techniques to describe microbial communities can range from microscopy, culturing, and DNA-DNA hybridization to metagenomics (Fuhrman 2008). Sequencing 16S rRNA genes recovered from the environment, like those used in this study, is the most frequently used molecular method for bacterial identification, and it allows for the study of the structure of the whole microbial community (Cardenas et al. 2009). It also provides the information needed to

attempt to isolate target taxa not yet cultivated or potentially find biogeochemical functions performed by these organisms (Röling and Head 2005).

However there are limitations with sequence based approaches to assess microbial diversity (Fierer and Lennon 2011). One should be aware that universal primers used in PCR to amplify the whole 16S rRNA gene do have some mismatches with certain known groups, but they are "nearly" universal maximizing the available phylogenetic information for bacterial identification (Fuhrman et al. 2008). Some bias can also be created by interference of the clone product with the clone host (if using molecular cloning), by differences in amplicon size and by artifacts like chimeras being created because of the hybridization of incomplete PCR products, which then serve as templates in further cycles of PCR (Cardenas et al. 2009). Bias in this study was controlled by checking for chimeras using the DECHIPER Chimera Check web tool and not including sequences identified as potential chimeras in further analysis. All amplicon sizes analyzed were $\geq$ 400 base pairs. With the widespread use of sequence based approaches, the number of bacterial sequences in rRNA databases is increasing rapidly from less than 50,000 sequences in 2002 to over 2,000,000 in 2012, but we are still far from surveying microbial diversity in most environments (Fierer and Lennon 2011, Cole et. al 2009).

Isolation of the organisms via culturing or a combination of sequencing techniques and methods that can identify functions are ways organisms identified by 16S rRNA gene sequencing can be linked to an ecological role in their environment (Röling and Head 2005). "Metagenomics describes the functional and sequenced-based analysis of collective microbial genomes contained in environmental samples" and it allows for the study of genetic material without first having to culture the organisms (Riesenfeld et

al. 2004, Nelson, 2008). Metagenomics studies have tried to link function with phylogeny, which requires the association of a phenotype to its DNA sequence (i.e., finding the genome of one species in a community and searching for gene coding regions and presumptive functions) (Riesenfeld et al. 2004). Three metagenomic approaches to link phylogeny and function have been identified by Riesenfeld et al. in a 2004 review of metagenomics, they include: screening a metagenomic library for a phenotype then trying to uncover the phylogenetic origin of the cloned DNA, screening clones for a specific phylogenetic anchor like 16S rRNA or another gene and then sequencing then whole clone and looking for genes neighboring the anchor or sequencing the complete metagenome and identifying genes of interest in the recreated genomes.

PCR primers or hybridization probes can be designed to find target genes, but the downfall of this sequence based approach is that the target genes must be from already known protein families (Vieites et al. 2008). There are approaches that do not require *a priori* familiarity of a gene sequence such as differential display (DD), cDNA-amplified fragment-length polymorphism (cDNA-AFLP), suppression subtractive hybridization (SSH), and serial analysis of gene expression (SAGE; Vieites et al. 2008). To highlight one of these approaches SSH allows for the differentiation between two very similar samples (Vieities et al. 2008). The SSH technique involves the separation of larger DNA fragments by inhibiting PCR amplification of smaller fragments within the sample, by binding adaptors to them (Nelson 2008). Subtracted cDNA or genomic DNA libraries can be created and differences between closely related DNA samples can be detected (Vieites et al. 2008). While it is ideal to have a pure culture in hand to completely classify a bacterial species, it is estimated that >99% of bacteria resist being grown in the lab.

Therefore, these novel culture-independent methods may be our best chance to learn more about environmental species and perhaps learn enough to bring some to cultivation.

CONCLUSIONS AND POSSIBLE FUTURE WORK

Bacterial diversity of the rhizospheres of living and dead Eastern Hemlock were assessed in this study. Variation based on time of year and over time was also assessed for hemlock rhizospheres. Results indicated significant differences in the composition at the phylum level of Live 2011 and Dead 2011 samples. Other significant differences were found at lower levels of classification between the four sample sets, including *Acidobacteria_Gp1, Acidobacteria_Gp3, Sphingobacteriales, Rhizobiales,* and *Burkholderia,* among others (Table 9). Overall, *Acidobacteria* was the dominant phylum making up 63% of all samples, followed by *Proteobacteria* at 23%, while other phyla were represented at levels $\leq$ 6%. New taxa to science could be represented with 19 of 937 total sequences that were unclassified at the phylum level (Table 1). Of the 26 OTUs containing at least 10 sequences, 21 were *Acidobacteria* and all others were *Proteobacteria* (Table 10).

The association of microbial communities within the rhizosphere of living hemlocks is important. If hemlocks cease to exist in the GSMNP because of the infestation of the HWA, then unidentified microorganisms that may be specific to hemlock may become extinct as well. If reforestation efforts were ever to take place for Eastern Hemlock in GSMNP, microbial communities associated with healthy trees could be vital in the success of this effort. Further work is needed to assess functions of the bacteria associated with the rhizosphere of the living Eastern Hemlock samples, especially *Acidobacteria*, where little is known about their physiological and ecological roles (Barns et al. 2007). Since most *Acidobacteria* are uncultivated, this could be accomplished through metagenomic analysis, with sequencing of DNA from millions of

cells or by activity-based screening approaches such as differential display, cDNA-amplified fragment-length polymorphism or DNA microarrays (Vieites et al. 2008). Further work could also be completed on the many unclassified bacteria found in this study; sequencing of the full 16S rRNA gene would provide more specific phylogenetic information. It would also be useful to construct 16S rRNA clone libraries to determine if the active bacteria in the rhizosphere reflect the results of this study. Of particular interest are the *Rhizobiales, Bradyrhizobium,* and *Acidobacteria* clones that might represent nitrogen-fixing bacteria whose potential role in hemlock rhizospheres is yet to be deduced. These are especially interesting groups in that they were very commonly observed OTUs in this work

LITERATURE CITED

Andreote, F. D., R. T. Carneiro, J. F. Salles, J. Marcon, C. A. Labate, J. L. Azevedo, W. L. Araujo. 2009. Culture-independent assessment of Rhizobiales-related Alphaproteobacteria and the diversity of *Methylobacterium* in the rhizospere and rhizoplane of transgenic eucalyptus. Microbiology Ecology. Vol. 57: 82-93.

Baker, S., J. Nicklin, N. Khan, R. Killington. 2007. Microbiology. 3rd Ed. Taylor and Francis Group. New York.

Barriuso J., Solano B.R., Lucas J.A., Lobo A.P., Garcia-Villaraco A. and Manero, F.J.G. 2008. Ecology, genetic diversity and screening strategies of plant growth promoting rhizobacteria (PGPR). *In* Plant-Bacteria Interactions. Ahmad I., Pichtel J. and Hayat S. (Ed.) Wiley-VCH. Weinheim, Germany. p. 1-13.

Barns, S. M., E.C. Cain, L. Sommerville, and C. Kuske. 2007. *Acidobacteria* phylum sequences in uranium-contaminated subsurface sediments greatly expand the known diversity within the phylum. Applied and Environmental Microbiology. Vol. 73 (9): 3133-3116.

Berg, G. and K. Smalla. 2009. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. Federation of European Microbiological Societies: Microbiology Ecology. Vol. 68: 1-13

Brock, T. D., M.T. Madigan and J.M. Martinko. 2006. *Brock Biology of Microorganisms*. 11th Ed. Prentice Hall. New Jersey.

Cardenas, E., J.R.Cole, J.M. Tiedje and J. H. Park. 2009. Microbial community analysis using RDP II (Ribosomal Database Project II): Methods, tools and new advances. Environmental Engineering Research Vol.14 (1): 3-9.

Chow, M. L., C. C. Radomski, J. M. McDermott, J. Davies, P. E. Axelrood. 2002. Molecular characterization of bacterial diversity in Lodgepole pine (*Pinus contorta*) rhizosphere soils from British Columbia forest soils differing in disturbance and geographic source. Federation of European Microbiological Societies: Microbiology Ecology. Vol. 42: 347-57.

Cole, J. R., Q. Wang, E. Cardenas, J. Fish, B. Chai, R. J. Farris, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, T. Marsh, G. M. Garrity, and J. M. Tiedje. 2008. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research.* Vol. 37 (Database issue): D141-145; doi:10.1093/nar/gkn879.

Corinaldes C, R Danovaro, and A Dell'Ano. 2005. Simultaneous recovery of extracellular and intracellular DNA suitable for molecular studies or studies from marine sediments. Applied and Environmental Microbiology. 71: 46-50.

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

DeSantis, T. Z., P. Hugenholtz, K. Keller, E. L. Brodie, N. Larsen, Y. M. Piceno, R. Phan, and G. L. Andersen. 2006b. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. Nucleic Acids Research Vol. 34:394-9

Dillow, C. 2009. Diversity Pattern of 16S rDNA of Bacteria and *amoA* of Archaea and Bacteria from Eastern Hemlock (*Tsuga canadensis*). M.S. Thesis. Western Carolina University.

Fraser, C., E. J. Alm, M. F. Polz, B. G. Spratt and W.P. Hanage. 2009. The bacterial species challenge: Making sense of genetic and ecological diversity. Science. Vol. 323. no. 5915, p. 741 – 746.

Fierer, N. and J. T. Lennon. 2011. The generation and maintenance of diversity in microbial communities. American Journal of Botany. Vol. 93 (3): 439-448.

Filion M., R. C. Hamelin, L. Bernier and M. St-Arnaud. 2004. Molecular profiling of rhizosphere microbial communities associated with healthy and diseased black spruce (*Picea mariana*) seedlings grown in a nursery. Applied and Environmental Microbiology. Vol. 70. 6:3541-51.

Fuhrman, J.A.. 2008. Measuring Diversity. *In* Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back. Ed. K. Zengler. ASM Press, Washington, DC.

Great Smoky Mountains National Park: Hemlock Woolly Adelgid. 2006. National Park Service. U.S. Department of the Interior. ParkNet. http://www.nps.gov/grsm/naturescience/hemlock-woolly-adelgid.htm.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. Vol. 6: 65-70.

Janssen, P.H.. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Applied and Environmental Microbiology. Vol. 73. No. 3:1719-28.

Kim, B. C., K. H. Lee, M. N. Kim, J. Lee and K. S. Shin. 2010. *Mucilaginibacter dorajii* sp. nov., isolated from the rhizosphere of *Platycodon grandiflorum*. Federation of European Microbiological Societies Microbiology Letters. Vol. 309: 130-135.

Kulichevskaya, I. S., A. O. Ivanova, O. I. Baulina, P. L. E. Bodelier, J. S. Sinninghe and S. N. Dedysh. 2008. *Singulisphaera acidiphila* gen. nov., sp. nov., a non-filamentous, *Isosphaera*-like planctomycete from acidic northern wetlands. International Journal of Systematic and Evolutionary Microbiology. Vol. 58: 1186-1193.

Kuykendall L.D. 2005 "Genus I. Bradyrhizobium" *In* Bergey's Manual of Systematic Bacteriology 2nd ed. Vol. 2. Eds. G.M. Garrity, D.J. Brenner, N. R. Krieg, J. T. Staley. Springer, New York.

Lee, S.H., J.O. Ka, and J.C. Cho. 2008. Members of the phylum *Acidobacteria* are dominant and metabolically active in rhizosphere soil. Federation of European Microbiological Societies: Vol. 285: 263-269.

Manoharachary C. and Mukerji, K.G. 2006. Rhizosphere biology-an overview. *In* Microbial Activity in the Rhizosphere. Mukereji K. G., Manoharachasry C. and Singh J. (Eds.). Springer. Leipzig, Germany p.1-11.

McClure M.S. and Cheah C.A.S.J. 1999. Reshaping the ecology of invading population of hemlock woolly adelgid *Adelges tsugae* (Homoptera: Adelgidae), in eastern North America. Biological Invasions Vol. 1:247-254.

Nehl D.B. and Knox O.G.G. 2006. Significance of bacteria in the rhizosphere. *In* Microbial Activity in the Rhizosphere. Mukereji K. G., Manoharachasry C. and Singh J. (Eds.). Springer. Leipzig, Germany p.1-11.

Nelson, K.E. 2008. Metagenomics as a tool to study biodiversity. *In* Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back. Ed. K. Zengler. ASM Press, Washington, DC.

Pankratov, T.A., B.J. Tindall, W. Liesack and S.N. Dedysh. 2007. *Mucilaginibacter paludis* gen. nov., sp. nov. and *Mucilaginibacter gracilis* sp. nov., pectin-, xylan- and laminarin-degrading members of the family *Sphingobacteriaceae* from acidic sphagnum peat bog. International Journal Systematic and Evolutionary Microbiology. Vol. 57: 2349-2354.

Quaiser, A., T. Ochsenreiter, C. Lanz, S. C. Schuster, A. H. Treusch, J. Eck and C. Schleper. 2003. Acidobacteria form a coherent but highly diverse group within the bacterial domain: evidence from environmental genomics. Molecular Microbiology. Vol 50. (2): 563-575.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Richardson, J., D.E. Stead, J.G. Ephinstone and R.H.A. Coutts. 2002. Diversity of *Burkholderia* isolates from woodland rhizosphere environments. Journal of Applied Microbiology.Vol. 93: 616-603

Riesenfeld, C.S., P.D. Schloss, and J. Handelsman. 2004. Metagenomics: Genomic analysis of microbial communities. Annual Reviews of Genetics. Vol. 38:525-52

Röling W.F.M., Head I. M. 2005. Prokaryotic systematics: PCR and sequence analysis of amplified 16S rRNA genes. *In* Molecular Microbial Ecology. Eds. A. M. Osborn and C. J. Smith. Taylor & Francis Group. New York.

Rossellò-Mora R. and A. Lòpez-Lòpez. 2008. The Least Common Denominator: Species or Operational Taxonomic Units? *In* Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back. Ed. K. Zengler. ASM Press, Washington, DC.

Schlegel, H. G., and Jannasch, H.W. 2006. Prokaryotes and Their Habitats. *In* The Prokaryotes: A Handbook on the Biology of Bacteria. Springer. New York, NY. 1:137-84.

Schloss, P.D. and J. Handelsman. 2004. Status of the Microbial Census. Microbiology and Molecular Biology Reviews. Vol. 68 (4). 686

Schloss, P.D. and Handelsman, J. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Applied and Environmental Microbiology. 71:1501-1506.

Sloan, W.T., C. Quince and T. P. Curtis. 2008. The Uncountables. *In* Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back. Ed. K. Zengler. ASM Press, Washington, DC.

Stadler B., Muller T. and Orwig D. 2006. The ecology of energy and nutrient fluxes in hemlock forests invaded by hemlock woolly adelgid. Ecology 87:1792-1804.

Thrash, J.C. and J.D. Coates. 2010. "Phylum XVII. Acidobacteria phyl. nov." *In* Bergey's Manual of Systematic Bacteriology 2nd ed. Vol. 4. Eds. N.R. Krieg, W. Ludwig, W.B. Whitman, B.P. Hedlund, B.J. Paster, J.T. Staley, N. Ward, D. Brown, A. Parte. Springer, New York.

Truper H.G. and Schleifer K. H. 2006. Prokaryote characterization and identification. *In* The Prokaryotes. A Handbook on the Biology of Bacteria. 3rd Ed. Vol 1. Eds. M. Dworkin, S. Falkow, E. Rosenberg, K.H. Schleifer and E.Stackebrandt, New York, Springer.

Vieites, J.M., M.E. Guazzaroni, A. Beloqui, P. N. Golyshin and M. Ferrer. 2008. Metagenomics approaches in systems microbiology. FEMS Microbiology Reviews. Vol. 33:236-55.

Wang, Q, G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl Environ Microbiol. Vol. 73(16):5261-7.

Webb R.E., Frank J. R. and Raupp M. J. 2003. Eastern hemlock recovery from hemlock woolly adelgid damage following imidacloprid therapy. Journal of Arboriculture Vol. 29:298-302.

White, David. 2007. *The Physiology and Biochemistry of Prokaryotes*. 3rd Ed. Oxford University Press.

Woods, D. E., and P.A. Sokol. 2006. "The Genus *Burkholderia.*" *In* The Prokaryotes. Vol. 5. Eds. M. Dworkin, S. Falkow, E. Rosenberg, K. H. Schleifer, E. Stackebrandt. Springer. New York.

Wright, E.S., L.S. Yilmaz, D.R. Noguera. 2012. DECIPHER: A Search-Based Approach to Chimera Identification for 16S rRNA Sequences. Applied and Environmental Microbiology. Vol. 78 (3):717-25

Xie, C.H. and A.Yokota. 2005. *Dyella japonica* gen. nov., sp. nov., a γ-proteobacterium isolated from soil. International Journal of Systematic and Evolutionary Microbiology. Vol. 55: 753-756.

York, E. 2008. Characterization of Bacterial Assemblages from the Rhizosphere of Easten Hemlock (*Tsuga canadensis*) using Culturing Methods and Molecular Cloning. M.S. Thesis. Western Carolina University.

Zaicnikova, M.V., Y. Y. Berestovaskaya, V. N. Akimov, N. A. Kostrikina and L. V. Vasilieva. 2011. *Singulispaera mucilagenosa* sp. nov., a novel acid-tolerant representative of the order *Planctomycetales*. Microbiology. Vol. 80 (1): 101-107.

Zengler, K. 2008. Does Cultivation Still Matter? *In* Accessing Uncultivated Microorganisms: From the Environment to Organisms and Genomes and Back. Ed. K. Zengler. ASM Press, Washington, DC.

Zinder, S. H. and M. Dworkin. 2006. Morphological and Physiological Diversity. *In* The Prokaryotes: A Handbook on the Biology of Bacteria. Springer. New York, NY: 1:185-220.

APPENDIX A

Ribosomal Database Project Library Compare results showing distributed patterns of taxa from clone libraries obtained from Eastern Hemlock (*Tsuga canadensis*) rhizospheres from Albright Grove, Great Smoky Mountains National Park, that were significantly different between microbial communities (Comparisons are labeled as follows: a-Live 2011 vs Dead 2011, b-Summer 2006 vs Dead 2011, c-Summer 2006 vs Live 2011, d- Summer 2006 vs Winter 2008, e-Winter 2008 vs Live 2011, f-Winter 2008 vs Dead 2011).

Comparison	**a**	**b**	**c**	**d**	**e**	**f**
Taxon						
Acidobacteria	1.21E-02			3.57E-02	7.14E-03	
Acidobacteria_Gp1	9.88E-03		4.88E-02			
Acidobacteria_Gp3				7.36E-03		
Gp3				7.36E-03		
Bacteroidetes		8.75E-03				2.94E-02
Sphingobacteria	1.80E-02	8.75E-03				2.98E-02
Sphingobacteriales	1.80E-02	8.75E-03				2.98E-02
Sphingobacteriaceae	5.29E-03	1.33E-02				2.33E-02
Mucilaginibacter	5.29E-03	1.33E-02				2.33E-02
Planctomycetes		4.66E-02				
Planctomycetacia		4.66E-02				
Planctomycetales		4.66E-02				
Planctomycetaceae		4.66E-02				

Proteobacteria				1.93E-02	1.55E-02	
Alphaproteobacteria	2.26E-02		2.36E-03		6.80E-06	1.39E-02
Rhizobiales	1.13E-02		1.51E-02	4.44E-02	6.69E-06	2.57E-02
Hyphomicrobiaceae	1.98E-02	4.04E-02				1.67E-02
Devosia	3.83E-02					3.31E-02
Rhodospirillales			7.31E-03		2.81E-02	
Acetobacteraceae			3.46E-02	3.01E-02		
Betaproteobacteria			4.14E-02			
Burkholderiales	3.49E-02					
Burkholderiaceae	3.21E-04					2.94E-02
Burkholderia	2.85E-03					2.94E-02
Gammaproteobacteria		4.38E-03	1.14E-02	4.55E-02		
Enterobacteriales				3.81E-02		1.46E-02
Enterobacteriaceae				3.81E-02		1.46E-02
Esherichia/Shigella						2.94E-02
Xanthomonadales		1.53E-02	1.84E-02			

APPENDIX B

Ribosomal Database Project Library Comparison results showing taxa that were significantly different between microbial community replicates of paired living and dead Eastern Hemlock (*Tsuga canadensis*) trees.

Taxon	Name	L1 (84)	D1 (82)	Significance
Phylum	*Proteobacteria*	4	15	1.02E-02
Class	*Alphaproteobacteria*	0	7	7.09E-03
Order	*Rhizobiales*	0	5	2.91E-02
Phylum	*Acidobacteria*	71	59	4.88E-02
		L2 (87)	**D2 (82)**	
Family	*Sphingobacteriaceae*	0	7	6.14E-03
Genus	*Mucilaginibacter*	0	7	6.14E-03
		L3 (68)	**D3 (92)**	
Class	*Acidobacteria_Gp1*	29	23	1.83E-02
Family	*Burkholderiaceae*	8	0	9.05E-04
Genus	*Burkholderia*	5	0	1.18E-02
Family	*Hyphomicrobiaceae*	0	6	4.16E-02