Analysis of bacterial diversity and biogeography at the Central Arizona-Phoenix Long Term Ecological Research (CAP LTER) site

Brian Anthony Rash

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations

Recommended Citation

https://digitalcommons.lsu.edu/gradschool_dissertations/2455

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
ANALYSIS OF BACTERIAL DIVERSITY AND BIOGEOGRAPHY AT THE CENTRAL ARIZONA-PHOENIX LONG TERM ECOLOGICAL RESEARCH (CAP LTER) SITE

A Dissertation

Submitted to the Graduate Faculty of the Louisiana State University and Agricultural and Mechanical College
In partial fulfillment of the Requirements for the degree of Doctor of Philosophy

In

The Department of Biological Sciences

by

Brian Anthony Rash
B.S., Bowling Green State University, 2000
December 2004
ACKNOWLEDGEMENTS

I dedicate this work to my parents Curtis Anthony and Regina Sue Rash and my wife Tammy Waguespack Rash whose unconditional love and continual emotional support this work would not have been possible.

I thank my advisor, Fred A. Rainey, who assisted in this work and provided continual guidance and financial support. I would also like to thank Drs. Bruce Williamson, Meredith Blackwell, Gregg Pettis, Kenneth Damann and Kevin Carman for serving on my advisory committee and for lending their scientific advice.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS......................................................................................................................ii

LIST OF TABLES..............................................................................................................................v

LIST OF FIGURES............................................................................................................................vi

ABSTRACT........................................................................................................................................ix

CHAPTER

1 INTRODUCTION............................................................................................................................1
   Culture-Dependent Techniques in Prokaryotic Diversity Studies...............................................2
   Using Culture-Independent Techniques in the Study of Prokaryotic Diversity…………………..4
   Ribosomal DNA Gene Sequences...............................................................................................10
   Diversity Indices and Statistics in Bacterial Ecology.................................................................14
   The CAP LTER Site: a Suitable Model for Urban Ecology Studies...........................................17

2 THE INFLUENCE OF LAND USE ON BACTERIAL DIVERSITY......................................23
   Introduction.................................................................................................................................24
   Materials and Methods..............................................................................................................29
   Results.......................................................................................................................................35
   Discussion.................................................................................................................................51

3 BACTERIAL BIOGEOGRAPHY AT THE CAP LTER..............................................................65
   Introduction.................................................................................................................................66
   Materials and Methods..............................................................................................................68
   Results.......................................................................................................................................68
   Discussion.................................................................................................................................108

4 DETECTION OF UBIQUITOUS TAXA ACROSS THE CAP LTER.................................111
   Introduction.................................................................................................................................112
   Materials and Methods..............................................................................................................116
   Results.......................................................................................................................................118
   Discussion.................................................................................................................................121

5 CONCLUSIONS..........................................................................................................................125

BIBLIOGRAPHY...............................................................................................................................129

APPENDIX A – VARIABLES USED FOR DIVERSITY ANALYSES............................................166
LIST OF TABLES

2.1 Description of CAP LTER soils used in this study........................................28
2.2 Distribution of clone OTUs in phylogenetic groups.................................37
2.3 Summary of MANOVA results...................................................................40
2.4 Comparison of observed and estimated total species richness in CAP LTER soils.................................................................44
2.5 Effects of land use on CAP LTER bacterial diversity..............................45
2.6 Effects of watering type on CAP LTER bacterial diversity......................47
2.7 LIBSHUFF comparisons of specific phyla contained in CAP LTER soil clone libraries.................................................................50
3.1 16S rRNA phylotype distribution of alpha and beta proteobacteria within CAP LTER soils.................................................................70
3.2 16S rRNA phylotype distribution of delta proteobacteria, gamma proteobacteria, and Acidobacteria within CAP LTER soils........................................81
3.3 16S rRNA phylotype distribution of Actinobacteria, Bacteriodetes, Chrloroflexi, and Cyanobacteria within CAP LTER soils.........................................93
3.4 16S rRNA phylotype distribution of Firmicutes, Gemmatimonadetes, Planctomycetales, and Verrucomicrobia within CAP LTER soils.................99
4.1 CAP LTER phylotypes chosen for 16S rRNA gene PCR primer design........115
4.2 Forward 16S rRNA gene primers used in the survey of environmental DNA samples.................................................................117
LIST OF FIGURES

2.1 CAP LTER site and sample collection locations........................................29
2.2 Bacterial group clone abundance separated by land use type.......................39
2.3 Correspondence analysis of community structure across CAP LTER land use
types.............................................................................................................41
2.4 Accumulation curves derived from 16SrDNA clone library data...................42
2.5 Results of a statistical analysis for the effect of ecological factors on bacterial
diversity.........................................................................................................46
2.6 $\text{C}_{\text{MH}}$ values for comparisons of ARDRA fingerprint assemblages across CAP
LTER samples as a function of distance......................................................48
2.7 LIBSHUFF comparison of replicate X13A and X13B clones.....................49
2.8 Results of a complete-linkage cluster analysis............................................52
3.1 16S rRNA gene tree showing positions of mostly heterotrophic members of the
$\text{Sphingomonadaceae}$ family found in CAP LTER samples................................71
3.2 16S rRNA gene tree showing positions of the $\text{Methyobacteraceae}$,
$\text{Caulobacteraceae}$, $\text{Alpha-3}$, $\text{Stappia-Rosebium}$, and $\text{Bradyrhizobiaceae}$ groups
found in CAP LTER samples........................................................................74
3.3 16S rRNA gene tree showing positions of the $\text{Rhizobiaceae}$, $\text{Bartonellaceae}$,
$\text{Hyphomicrobiaceae}$, $\text{Alpha-4}$, $\text{Rhodospirillaceae}$, and $\text{MND-8}$ groups found in
CAP LTER samples.......................................................................................75
3.4 16S rRNA gene tree showing positions of the $\text{Alpha-5}$, $\text{Rohodobacteraceae}$, and
$\text{Acetobacteraceae}$ found in CAP LTER samples........................................76
3.5 16S rRNA gene tree showing positions of the $\text{Oxalobacteraceae}$,
$\text{Commonadaceae}$, $\text{Neisseriaceae}$, and 646-2 groups found in CAP LTER
samples...........................................................................................................78
3.6 16S rRNA gene tree showing positions of the $\text{Polyangiaceae}$, Q3, $\text{Myxococceae}$,
$\text{Syntrophonobacteriodetes}$, 6G20, and $\text{Bdellovibrio}$ groups found in CAP LTER
samples...........................................................................................................80
3.7 16S rRNA gene tree showing positions of the $\text{Ectothiorhodospiraceae}$,
$\text{Thiotrichaceae}$, $\text{Pseudomonadaceae}$, $\text{Moracellaceae}$, $\text{Xanthomonadaceae}$, and
$\text{Coxiallaceae}$ groups found in CAP LTER samples.....................................83
3.8 16S rRNA gene tree showing positions of the *Acidobacteria* groups I and II, and phylotypic groups 384-2, iii-8, and Sva0725 found in CAP LTER samples………84

3.9 16S rRNA gene tree showing positions of the *Acidobacteria* group III found in CAP LTER samples.................................................................85

3.10 16S rRNA gene tree showing positions of the *Acidobacteria* groups V and VI found in CAP LTER samples...........................................87

3.11 16S rRNA gene tree showing positions of the *Geodermatophilaceae, Pseudonocardiaceae, Kineosporaceae, Frankiaceae*, and *Acidotherrnaceae* groups found in CAP LTER samples.........................................................88

3.12 16S rRNA gene tree showing positions of the *Rubrobacteraceae* group found in CAP LTER samples........................................................86

3.13 16S rRNA gene tree showing positions of the *Acidomicrobiaceae* group found in CAP LTER samples..............................................91

3.14 16S rRNA gene tree showing positions of the *Microbiaceae, Cellomonadaceae, Intrasporangiaceae, Micrococcaeae, Actinosynnemataceae, Streptomycetaceae, Nocardiodaceae, Propionibacteraceae, Micronomonosporaceae*, and 7L14 groups found in CAP LTER samples.................................................................92

3.15 16S rRNA gene tree showing positions of the *Flexibacteraceae, Flavobacteraceae, BD-7 and KL* groups found in CAP LTER samples..........95

3.16 16S rRNA gene tree showing positions of the *Chloroflexi* group found in CAP LTER samples..............................................................96

3.17 16S rRNA gene tree showing positions of the *Cyanobacteria*, phylotypes found in CAP LTER samples.........................................................96

3.18 16S rRNA gene tree showing positions of the *Balliceae* group found in CAP LTER samples.................................................................99

3.19 16S rRNA gene tree showing positions of the *Alicyclobacillaceae, Clostridiaceae, Peptococceae, Paenibacillaceae, Staphylococceae, Carnobacteraceae*, and *Sytophonomonadaceae*, groups found in CAP LTER samples..............100

3.20 16S rRNA gene tree showing positions of the *Gemmatimonadetes* group I found in CAP LTER samples.................................................102

3.21 16S rRNA gene tree showing positions of the *Gemmatimonadetes* groups II, III, and IV found in CAP LTER samples..................................103
3.22 16S rRNA gene tree showing positions of the *Gemmata, Isophaera, Planctomyces, Pirellula, and 96-2* groups found in CAP LTER samples.........104

3.23 16S rRNA gene tree showing positions of the *Spartobacteria, Verrucomicrobiae*, and *Verrumicrobia* III and IV groups found in CAP LTER samples.................106

3.24 Unrooted radial 16S rRNA tree showing positions of the *Deinococcus-Thermus* phylum, and candidate phyla OP3, OP10, TM 7, and WS 6 in the CAPLTER...107
ABSTRACT

The limiting factor involved in past assessments of soil bacterial diversity when using culture-independent techniques has often been the lack of sampling and replication. As a result, analyses of community structural shifts across soil environments have lacked statistical power. In this study, 23 16S rRNA gene clone libraries consisting of over 11,000 clones were constructed from soils at the Central Arizona Phoenix Long Term Ecological Research (CAP LTER) site. Subsequent ARDRA fingerprinting and partial 16S rRNA gene sequencing allowed for a more robust investigation of various components that may explain any observed variations in bacterial species composition. The designated land use type of the soils best explained the overall diversity gradient. Based on Simpson’s reciprocal index, diversity was found to significantly increase when comparing urbanized and agricultural soils to open desert samples located outside the metropolitan region. Land use type appears to be a powerful indicator of overall diversity due to irrigation methods that differ greatly across land use types. Experiment-wise comparisons of complete CAP LTER clone libraries via the LIBSHUFF method yielded no statistical similarity in sequence libraries, except for two replicate libraries constructed from one urban soil. However, inter-phylum LIBSHUFF analysis of the clones also shows degrees of phylogenetic partitioning between land use categories and that open desert remnant patches located within the city limits more closely resemble those urban soils than the open deserts outside of Phoenix. Examination of constructed 16S rRNA phylogenetic trees that include CAP LTER phylotypes indicate some distinct clustering of sequences appears to be driven by land use type rather than geography, and
that most of these groups may be endemic to the region. However, some ubiquitous phylotype groups were discovered and were used as templates for specific PCR primer design, allowing for the detection of ten different groups in all soil samples analyzed. Overall, this study suggests that anthropogenic factors have altered soil bacterial communities, the biogeography of many species is controlled in some manner by land use type, and that a small subset of taxa is ecologically tolerant despite the heterogeneity of habitats within the site.
CHAPTER 1:

INTRODUCTION
Culture-Dependent Techniques in Prokaryotic Diversity Studies

Culture-dependent techniques are defined as methods that involve the culturing or enrichment of prokaryotes using various forms of artificial media either in a laboratory or *in situ* setting. This approach may include the direct isolation of individual colonies or the enrichment of a prokaryotic consortium. Up until the 1980s, most prokaryotic ecology studies described in the scientific literature employed culture-dependent methods that produced culturable strains that were held in culture and lab strain collections. These organisms can be phenotypically and genetically analyzed and given a valid taxonomic designation based on presently accepted guidelines. However, a significant proportion of these strains, especially those that have never been formally described or extensively studied and are held in lab strain collections, are unknown entities in terms of their phenotypic and phylogenetic diversity. It is only since extensive phylogenetic analysis and identification of strains by gene sequencing became routine that the species richness of these collections could be assessed (Woese 1994; Rainey and Stackebrandt 1996).

Despite recent progress, there remains a significant lack of understanding of prokaryotic diversity in nature and this is compounded by the belief that approximately only 1-5% of prokaryotes can be readily cultured (Ward et al. 1992; Amann et al. 1995, Head et al. 1998). The large discrepancy between the number of colonies that form on solid media and the total number of cells observed microscopically is a consequence of the present inability of artificial media and culture conditions to facilitate the growth of the majority of prokaryotes. There are multiple explanations for this noted inconsistency. Most culture conditions cannot mimic the specific microhabitats that many prokaryotes thrive within, a complex niche that requires strict concentrations of various compounds.
required for sustained cell growth. The fact that many of these same essential compounds can become inhibitory or lethal to oligotrophic prokaryotes at high concentrations further complicates culture-dependent approaches (Buck 1974; Hattori and Hattori 1980; Hattori 1981). This knowledge has allowed for the design of new formulas of artificial media which often contain diluted amounts of metabolic substrates and substitute artificial compounds in the place of agar as solidifying agents (Connon and Giovannoni 2002; Jannsen et al. 2002). Difficulty in culturing also arises from the state of dormancy that many bacteria exhibit in situ, and the resulting culture conditions which are unable to trigger these cells into reproductive cycles (Demming and Baross 2000, Colwell 2000). Moreover, there are many examples of bacterial species that can in fact be cultured using traditional culture media but are not given sufficient time to produce colonies (Mitsui et al. 1997; Connon and Giovannoni 2002; Janssen et al. 2002; Kaeberlein et al. 2002; Sait et al. 2002; Joseph et al. 2003; Rappe and Giovannoni 2003). Many forms of media will be dominated initially by colonies associated with r-strategist bacteria, a group associated with the rapid response to emerging nutrient pools by the rapid transcription of rRNA and subsequent synthesis of ribosomes (Klappenbach et al. 2000). Therefore, most media protocols select for this r-type of ecological strategy, and in the process restricts the colonization of those species with lower rRNA gene copy numbers that divide at a slower pace. Janssen et al. indicates that a 12 week incubation period can successfully recover members of slow-growing, previously uncultured prokaryotic groups (2002). These taxa generally have lower rRNA gene copy numbers, a variation of a k-strategy that is described as the delayed response to resource input and sensitivity to disturbance. Given the limitations of the culture-dependent approach, any
ecological survey using only these methods will be impacted by these biases, and therefore deemed inadequate in capturing a highly resolved picture of the prokaryotic components of any ecosystem.

**Using Culture-Independent Techniques in the Study of Prokaryotic Diversity**

Due to the inability of microbiologists to exhaustively sample any given environment using currently available culturing methods, the culture-independent approach has been used to assess bacterial diversity across many sample sites. Culture-independent methodologies focus primarily on the capture of nucleic acids from cells present within an environmental sample. Once extracted from cells, genetic material can be analyzed in various ways to identify the phylogenetic and physiological makeup of communities that are present in the environmental sample. These procedures have been found to be most beneficial in analyzing the prokaryotic diversity of soils based on the increased number of species found in soil coupled with an overall deficiency in sampling when compared to other ecosystems (Rashit and Bazin 1987, Zhou et al. 2002). Based upon DNA reassociation kinetics data, it is estimated that one gram of soil contains between 2,000 and 18,000 genomes (Torsvik et al. 1990, Torsvik et al. 1996; Sandaa et al. 1999). Under a similar diversity estimation model, an estimated 2,000 to 4,000 genomes were present in the Sargasso sea using a genome shotgun sequencing approach (Venter et al. 2004), indicating a decrease in overall species richness when comparing soil communities to sediment counterparts. Furthermore, implementation of an assumed lognormal species distribution model combined with rarefaction analysis provides total bacterial species estimates of roughly 16,000 to 50,000 species of bacteria per gram of soil (Dunbar et al. 2001). Although these estimates have not been tested using a
complete sample of a soil habitat, it is clear that there is a large number of unsampled bacterial species rendering many previous studies on changes in bacterial species composition obsolete. Thus, culture-independent approaches are essential in beginning to understand the extent of prokaryotic diversity in any ecosystem, and have exposed the relative ineffectiveness of culturing approaches to reveal the full breadth of diversity present in such environments. As previously stated, culture-independent techniques often initially involve extracting nucleic acids in the form of DNA or RNA from an environmental sample. A complete collection of nucleic acids is crucial in obtaining all the genomic information contained within the sample for subsequent analysis. There are various means of extracting nucleic acids that are dependent on the type of environment from which the sample was collected. Soils require the most rigorous nucleic acid extraction methods due to the attachment of prokaryotic cells to soil particles (Priemé et al. 1996). Most extraction methods employ both chemical and physical means for bacterial cell wall disruption. Compounds such as lysozyme and sodium dodecyl (lauryl) sulfate (SDS) act in disrupting the outer membranes of bacteria, resulting in cell lysis and the dispersal of contaminating proteins respectively. Physical means include the use of bead-beating and freeze/thaw methods that will disrupt the more rigorous cell walls, including those of bacterial spores and phyla such as the Actinobacteria which are surrounded by a thick peptidoglycan layer. Once the nucleic acids have been extracted, they are purified by the addition of proteinase K followed by phenol/chloroform washes. The nucleic acids are then captured onto a positively charged material and washed in order to remove other contaminating compounds, thus purifying the remaining genetic material.
At this point, a variety of culture-independent approaches using DNA extracts can be employed. One approach, known as Denaturing or Temperature Gradient Gel Electrophoresis (DGGE or TGGE), was first developed by Myers et al. (1987) to detect small changes in cloned and genomic DNA sequences. DGGE is used to obtain a profile of the amplified environmental DNA consortium that can be visualized in 2-D polyacrylamide gels (Muyzer et al. 1993; Duineveld et al. 1998; Eichner et al. 1999; van Hennen et al. 1999). The generated profiles (also known as fingerprints) are the result of the fact that different DNA sequence helices melt or denature at different rates in accordance with small changes in DNA sequence (Gotoh 1983), allowing for separation across a gradient gel. The resulting DGGE/TGGE fingerprints can then be compared across environmental samples via band matching algorithms. Bands of interest can also be excised from the gel and sequenced for further analysis. While this approach is relatively inexpensive and efficient for comparing banding patterns of various samples, as the complexity of the community increases it becomes more difficult to excise individual bands for sequence analysis and quantification of the richness and evenness of individual bands, which in turn can restrict the measurement of diversity.

Another widely used culture-independent technique is gene clone library construction. The steps involved in this method are the amplification of a specific sequence located on the various genomes present within the total environmental DNA sample using the Polymerase Chain Reaction (PCR), insertion of the individual amplicons into an engineered plasmid, and the transformation of these plasmids into competent *E.coli* host cells, potentially generating thousands of clones per environmental DNA sample. Once the individual clones are collected, several different fingerprinting
techniques can be implemented to separate and identify potential species without the time
and budget constraints of extensive sequencing. One such method known as Restriction
Fragment Length Polymorphism (RFLP) involves the use of restriction enzymes that cut
DNA at specific sites, resulting in a collection of DNA fragments (Lee et al. 1993;
Gundersen et al. 1994; Muyzer et al. 1995). These fragments are then separated using an
ultra-pure agarose that can resolve fragments that differ by less than 25 bp during
electrophoresis. The gels are then stained with ethidium bromide, and visualized using
UV illumination. The various DNA fragments migrate at different rates based upon
length and thus distinct banding patterns are observed for those sequences that are
different and contain different restriction sites. The extent of variation between
individual RFLP fingerprints depends on the restriction enzymes used and sequence
composition. Using multiple restriction enzymes in most cases will allow for enhanced
resolution by increasing the number of different sized fragments. A more specific RFLP
method involving ribosomal genes, usually the 16S rRNA gene, is known as Amplified
Ribosomal DNA Restriction Analysis (ARDRA). This method produces identical
ARDRA fingerprint groups for 16S rRNA gene clones with highly similar sequences
which then can be considered as an Operational Taxonomic Units (OTU), and therefore
assumed to be both species-like and a distinct ecological entity (Wayne et al. 1987;
Stackebrandt et al. 2002). The number of these OTUs can be determined for individual
samples, allowing for the estimation of diversity within an individual sample or the
calculation of the number of shared OTUs across samples for direct comparisons of
different environments.
The first ecological study of a soil sample using 16S rRNA gene clone libraries was performed by Liesack and Stackebrandt and published in the August 1992 issue of *Journal of Bacteriology* (1993). In this study, 113 clones were constructed from an Australian soil sample and phylogenetically identified, revealing a substantial amount of diversity never before observed in nature. Since this classic work, hundreds of subsequent studies have applied 16S rRNA gene clone library construction for answering a variety of questions pertaining to the ecology of prokaryotes and eukaryotes. However, although the number of 16S rRNA gene clones generated for analysis has increased over time, the overall numbers of clones typically range between 50-400 clones per sample studied. These small 16S rRNA gene clone libraries represent a significant undersampling of the prokaryotes in most environments (Torsvik 1990; Dunbar et al. 2001). The reason for the low number of clones generated in the past is a consequence of the high cost and labor-intensiveness of the various steps involved in the cloning and sequencing processes. Constructing larger 16S rRNA gene clone libraries (>400 clones) allows for a larger numerical and genetic dataset that is crucial in determining the degree of diversity in terms of species evenness and richness, as well as obtaining the statistical power to make significant conclusions about the environments sampled. Because these sequences can be incorporated into a phylogenetic context and are known to originate from the environmental sample under investigation, this paper will refer to the source organisms of the sequences as “phylotypes” (Reysenbach et al. 1994; Delong et al. 1994), although they are commonly referenced elsewhere as “environtaxa” and “ribotypes” (Snipes et al. 1989). The data collected using 16S rRNA gene clone libraries can be analyzed in multiple ways, most notably by quantitative measurement of species richness
and evenness values. Significant changes in these values across separate environmental samples may indicate selective pressures that influence changes in bacterial diversity.

There are recognized caveats associated with these culture-independent techniques. First, the efficiency of DNA extraction techniques vary widely. This variation is a consequence of each procedure’s ability to lyse even the most rigorous of cell walls, or the shearing of the DNA due to over robustness of the physical methods (Martin-Laurent et al. 2001). Lysis treatments can be ineffective in lysing certain organisms due to cell membrane resistance or, in the case of terrestrial sites, protection of the cells by soil structures (Frostegard et al. 1999). Due to the observed biases in DNA extraction protocols in the literature, many datasets across sample sites cannot be directly compared. Furthermore, biases associated with the PCR reaction have been widely studied, most dealing with the ability of certain template strands to be polymerized more efficiently (Suzuki and Giovannoni 1996; Polz and Cavanaugh 1998), and the production of sequence artifacts including chimeras and heteroduplexes (Wang and Wang 1996; Speksnijder et al. 2001). In soils, there is the added problem the presence of humic acids, which are high molecular weight, degradation-resistant compounds within organic-rich environments that interfere with PCR amplification (Aiken et al. 1985; Steffan et al. 1988; Rochelle et al. 1992; Tsai and Olsen 1992). There also are recorded biases involving varying \( rrn \) copy numbers across bacterial genomes (Linton et al. 1994; Rainey et al. 1996; Crosby and Criddle 2003). Since 16S rRNA gene function is essential to the bacterial cell, the number of functioning copies located on the genome ranges from one (Cole and Girons 1994) to as many as fourteen copies (Young and Cole 1993). Because the copy number varies drastically amongst bacterial species, and those with higher copy
numbers can be amplified to a greater extent in PCR reactions, the quantification of
species numbers can be biased (Farrelly et al. 1995).

The ability to detect novel taxa using culture-independent techniques and
subsequent phylogenetic assignment alongside cultured taxa is the key to understanding
overall prokaryotic diversity and phylogeny. This approach has provided for an
increased understanding of the prokaryotic domains Bacteria and Archaea, which can
now be divided into over 40 phyla, many of which are absent of any cultured
representation (Hugenholtz et al. 1998). Many studies that use 16S rRNA gene clone
library technology report phylotypes for which no phylogenetically-similar cultured
isolates are available (Rheims et al. 1996; Bornemann et al. 1996; Bornemann and
1998; Hugenholtz et al. 1998; McCaig et al. 2001). The conclusions of these molecular
ecology studies are that the recovered 16S rRNA gene sequences represent new higher-
order phylogenetic groups. Overall, 16S rRNA gene clone library construction has
allowed researchers to finally begin to answer fundamental ecological questions
involving the relationships amongst bacteria and investigate total species distribution
(Dunbar et al. 1997), total species richness estimation (Hughes et al. 2001), and
prokaryotic biogeography (Bowman et al. 2003).

Ribosomal RNA Gene Sequences

Ribosomes are involved with the translation of RNA to proteins, a process that is
essential for all forms of life. Ribosomes are composed of both a small and large
subunits. The small subunit of the prokaryotic ribosome contains one 16S rRNA
molecule (~1,500 nucleotides) and 20 proteins and is labeled as the 30S subunit of the
ribosome. The function of the 30S subunit is to attach to a mRNA sequence located upstream of the initiation codon, which in most cases is a conserved region known as the Shine-Delgarno sequence. The Shine-Delgarno sequence provides a substrate for the remaining parts of the ribosome to merge and move along the mRNA fragment. The larger 50S subunit consists of one 23S rRNA (2,900 nucleotides), one 5S rRNA (120 nucleotides) and 30 proteins. The function of the 50S subunit is to provide a complex where tRNAs carrying the amino acid can be attached to each other via the peptidyltransferase reaction. Due to its importance in cellular function, ribosomes compose approximately 40 percent of an active cell’s dry weight (Pace et al. 1986).

Of the three ribosomal RNAs, the 16S rRNA has been the molecule of choice for the majority of prokaryotic phylogenetic analyses. The 5S rRNA gene sequences were initially used, but were quickly deemed limited in terms of sequence length and data content to differentiate between taxa and provide for a robust phylogenetic analysis. The 16S rRNA gene sequence was selected over the larger 23S rRNA gene sequence due to the ability to retrieve sufficient evolutionary information with the smaller gene sequences that require less sequencing time (Pace et al. 1986). In addition, the limitations of PCR and sequencing technology at the point when such studies were initiated lead to the use of the 16S rRNA gene sequence and as the database increased in the number of reference sequences, its use in phylogenetic characterization quickly became the standard.

Despite advances in sequencing and the study of molecular evolution using alternative genes, 16S rRNA gene sequences remain the most commonly used molecular marker in prokaryotic phylogenetic studies for a number of reasons. It is one of the slowest evolving sequences found in all living organisms, which makes it useful for
examining ancient evolutionary events (Hillis and Dixon 1991). All 16S rRNA gene sequences are similar in length (~1500 nucleotides) in prokaryotes and contain for the most part ultra-conserved regions, however the variable regions usually diverge and provide differentiation according to taxonomic rank (Rochelle et al. 1992). Due to the conformational complexity of binding with twenty-one proteins to form the 30S ribosomal subunit, and the essential housekeeping functions they provide, the rRNA genes are extremely resistant to selective pressures that affect other genes and evolve at an estimated rate of only 1-2% per 50 million years (Moran et al. 1993). The stability of the 16S rRNA gene allows it to be easily aligned and compared across all taxa (Pace et al 1986; Woese 1987). Currently, two readily available 16S rRNA gene databases exist, most notably Genbank (http://www.ncbi.nlm.nih.gov/blast), and RDP (http://rdp.cme.msu.edu/html), each containing over 140,000 and 86,000 partial and full sequences, respectively. In addition, the ARB software package (http://www.arb-home.de), designed exclusively for 16S rRNA gene alignment and analysis, has allowed for a user-friendly interface for the identification of environmental sequences and construction of phylogenetic trees (Strunk et al. 2004). The sequences in these databases represent valid prokaryotic species, as well as environmental isolates and phylotypes from a diverse array of ecosystems and geographical locations. This extensive database of sequences is many magnitudes greater than what is available for other prokaryotic genes, and therefore allows researchers the best opportunities to correctly affiliate sequences at the genus-species level as well as to construct higher-order phylogenetic relationships amongst a large number of taxa (Hills and Dixon 1991). The importance of the rRNA gene in the study of evolution is highlighted by its use as the molecular marker of choice
in the proposal for the three domain structure of organismal phylogeny, differentiating *Archaea, Bacteria, and Eucarya* (Woese et al. 1990).

16S rRNA phylogenetic trees also act as the principal mechanism in determining taxonomic levels within the prokaryotic domains as described by the taxonomic outline of *Bergey's Manual of Systematic Systematic Bacteriology* (http://www.cme.msu.edu/Bergeys; http://rdp.cme.msu.edu/hierarchy/hb_intro.jsp). This designation of species boundaries has been chosen because of the positive correlation between 16S rRNA gene and whole genome sequence divergence. A dissimilarity score of greater than 3 percent in the 16S rRNA gene can be reasonably correlated to at least a 30 percent dissimilarity in the entire DNA between two genomes, which designates separate species status and correlates with the observed phenotypic differences found amongst bacteria (Wayne et al. 1987; Stackebrandt and Geobel 1994; Stackebrandt et al. 2002). Although this standard is applicable in most known cases, there are major phylogenetic groups such as the *Actinobacteria* and the *Firmicutes* that display accelerated rates of evolution, where a sequence of divergence of less than 1 percent can constitute separate species, causing bacterial species boundaries to be less genetically definitive domain-wide (Ochman et al. 1999). This phenomenon is hypothesized to be caused by quantum evolution of 16S rRNA genes, the substantial increase in substitution rates in certain phylogenetic groups when protons remain attached to the tautomeric nitrogen atom of a nucleotide, resulting in a definitive increase in mutations during DNA replication in the daughter strands (Cavalier-Smith 2002).

In spite of its practicality, the use of 16S rRNA ribosomal gene sequence for phylogenetic analysis in bacteria does have potential problems. Certain studies suggest
that short segments of the 16S rRNA gene (100-200 bp) have been exchanged between species via horizontal gene transfer (HGT) in the past (Mylvaganam and Dennis 1992; Wang and Zhang 2000; Gogarten 2002). HGT events could alter vertical evolutionary descent and therefore create inaccuracies in phylogenetic inferences based solely on the 16S rRNA gene sequence. However, such HGT events are most likely to occur between two species or strains of a species with low levels of phylogenetic distance, and therefore still allow relationships above the species level to remain accurate. These higher order relationships are more valued in environmental studies in which one is usually comparing overall diversity in less resolved terms. Moreover, in eukaryotic studies there have been multiple cases in which the construction of nuclear gene trees are incongruent with that of mitochondrial 16S rRNA-based trees (Cao et al. 1994; Cummings et al. 1995; Zardoya and Meyer 1996; Naylor and Brown 1997; Springer et al. 2001). These artifacts in tree-building have been accredited to fluctuating rates of mutation across the 16S rRNA gene, where certain slowly evolving fragments of the gene are less prone to problematic factors such as multiple substitutions, and therefore retain a better ancient phylogenetic signal than those gene positions that evolve faster (Brochier and Phillipe 2000). The effect of this phenomenon on prokaryotic-based phylogenetic analysis has not been studied.

**Diversity Indices and Statistics in Bacterial Ecology**

Over time, an increasing number of more comprehensive prokaryotic diversity studies have led to the application of ecological diversity indices that previously have reserved for macrobial research (Dunbar et al. 2001; Zhou et al. 2002; Hill et al. 2003; Cummings et al. 2003). This estimation of diversity is essential to provide insight into
the participation of prokaryotes in various ecological processes as well as biogeography and community assembly (Curtis et al. 2002).

There are many apparent difficulties in assessing overall bacterial diversity. One of the utmost concerns is the difference between three sets of elements: functional, structural, and compositional diversity (Noss 1990). Compositional diversity, the measure of species abundance, is more widely used, notably in prokaryotic ecological studies because of the inability to culture the majority of these organisms. This has led to the discovery of 16S rRNA gene phylotypes for which no functional data are available. This lack of phenotypic data is compounded by the understanding that in some instances, closely-related taxa based upon 16S rRNA similarity demonstrate significant phenotypic divergence (Fox et al. 1992; Palys et al. 1997). Examples of this discrepancy in 16S rRNA gene evolutionary rates include *Bacillus globisporus* and *Bacillus psychrophilus*, two species that share 99.8% similarity in the 16S rRNA gene sequence but differ in terms of optimal temperature and nutrients required for growth (Larkin and Stokes 1967; Gordon et al. 1973; Nakamura 1984) as well as protein-coding gene divergence (Palys et al. 2000). However, Nübel et al. (1999) showed an overall positive correlation between genetic and ecological divergence in a microbial mat ecosystem.

Another problem in assessing diversity is how the rare species in a sample are treated in the calculations of diversity. Due to the undersampling biases involved in culture-independent techniques, many studies collect only the more abundant species. Thus, undetected rare species are neglected in these surveys. This approach may be problematic when studying prokaryotic communities since little is known about how prokaryotic cells behave *in situ*, and so it would seem dangerous to assume that all rare
species have little effect on the system. Also, the undersampling of species using culture-independent methods makes the quantification of rare members in theoretical models error-prone; therefore, it is important to include rare species in diversity calculations (Gaston 1994).

Even with these numerous shortcomings, molecular based prokaryotic diversity indicators are useful in underscoring shifts in overall community structure. The measure of community structure and/or richness are indicators of selective pressures that influence diversity within environments. By using various treatment effects as predictor variables for changes in diversity values we can potentially explain how certain biotic and abiotic parameters are affecting prokaryotic community composition. These relative measures of species diversity can also be used as evidence for the functional diversity of a community, as it has been shown to be positively correlated to eukaryotes and some prokaryotes (Martin 2002).

Until recently, prokaryotic ecology studies lacked the application of statistical machinery to permit a robust analysis of environmental data. However, mass sequencing of 16S rRNA phylotypes has permitted the use a number of statistics amended specifically for use of sampled 16S rRNA gene sequences rather than directly-sampled organisms. The most commonly applied statistical analysis is LIBSHUFF analysis, which compares and analyzes genetic similarity of two clone libraries at different degrees of evolutionary distance to test for the degree of overlap (Singleton et al. 2001). Another currently used statistical approach involves a comparison of two clone libraries by calculating overall genetic diversity coupled with quantifying the amount of phylogenetic covariance (Martin 2002). The ability to avoid OTU definitions in these methods can
complement DNA fingerprinting results to further validate observed changes in prokaryotic communities.

The Central Arizona Phoenix Long Term Ecological Research (CAP LTER) Site: a Suitable Model for Urban Ecology Studies

Studying the urban environment has become a top priority for many in the field of ecology because of its potentially rapid and pronounced effects on the global ecosystem. Currently, around half of the world’s population now resides in cities, but this number is expected to grow to around 61% by the year 2025 (United Nations 1997; Sadik 1999). Due to this urban population explosion, sprawl has increased the disturbance of surrounding environments (Burchell et al. 2002). This trend is quite ominous given that various anthropological studies have concluded that rapid urbanization in the ancient civilizations ultimately led to abandonment of entire cities and the destruction of civilization (Redman 1992; Rice 1996; O’Hara et al. 1993). More recently, there is mounting evidence that increased urbanization is positively correlated with increases in the transmission of Lyme disease in the Northeastern US (Schmidt and Ostfeld 2001), Nipah virus in Malaysia (Chua et al. 1999; Lam and Chua 2002), as well as other pathogenic agents worldwide (Palumbi 2001; Rose et al. 2001). Another disturbing prediction by E. O. Wilson is that loss in biodiversity is will significantly accelerated by increasing human influences (2002). Decreasing biodiversity is expected to negatively affect basic life-support functions for the human race, including atmospheric gas regulation, climate regulation, water purification, soil formation, waste removal, biological control of populations, food, and raw materials (McDaniel and Borton 2002). Based on these numerous observations and predictions, urban areas are now considered to be complex ecological entities, affecting the local ecosystem directly and remotely.
through land conversion (Machilis et al. 1997; Alberti et al. 2003). To avoid possible irreversible damage to the ecology of urban ecosystems in the future, the study of urban ecology has become imperative in the hopes that it will allow for alterations in environmental management policy that will maintain a reasonable quality of life for humankind (Patz et al. 2004).

One specific way to study how ecological patterns differ in urbanized settings is to look for changes in communities along gradients of land use type. Land use type equates to a pattern of human activity, and ultimately displays how a given tract of land is influenced by human decisions. Changes in land use are predicted to cause significant changes in biodiversity in various environments worldwide (Wolters et al. 2000; Lake et al. 2000; Smith et al. 2000). The mechanisms involved in such an extreme alteration within ecosystems are habitat loss, increased disturbance, and the immigration of new chemical inputs (Adams and Wall 2000). Such activities have already been shown to be directly involved in the degradation of water quality in oligotrophic lake and marine ecosystems, creating substantial economic losses (Tikkanen et al. 1997; Lithrop et al. 1999; Carpenter et al. 1999). Similar change in terrestrial biomes has resulted in the conversion of pristine habitats and overall loss in community complexity and diversity (Rapport et al. 1985). Therefore, the study of bacterial assemblages along land use type gradients may reveal overall diversity loss in urban bacterial populations that would indicate the decreased capacity of the resident soils to sustain biological productivity that promotes plant and animal health (Girvan et al. 2002). Such knowledge could lead to amended conservation policies for urban ecosystems as required.
The CAP LTER provides for a plausible environment to test for potential changes in bacterial diversity across land use gradients. It incorporates a 6400 km² area within the Southwest United States. At the center of this site lies the city of Phoenix with primarily rural environments dominating the outer boundaries. This site provides a wide range of land use types including but not limited to agricultural, commercial, parks, residential, and natural desert environment sites. These sites are classified into five major land use categorizations based on observational data as defined by the Maricopa Association of Governments (1997). These land use types include agricultural, urban, open desert, desert remnant, and mixed varieties. The extent of available land use types allows for the study of potential ecological effects of land use and how the urban landscape can alter community composition of various forms of life in a compact, yet geographically diverse setting (Grimm et al. 2000; Hope et al. 2003; Cousins et al. 2003).

The CAP LTER site is also attractive as a model study site because of its location within the Sonoran desert, a region that has been studied for over fifty years, most notably by plant ecologist Forrest Shreve (1951), who noted the great variation amongst native vegetation. The Sonoran desert encompasses over 310,000 km² of land that includes significant portions of Arizona, California, and the upper coast line of the Gulf of California located in Mexico. This area has the highest temperatures and lowest precipitation in North America (Schmidt 1989; Turner 1990). Due to the extreme nature of the physical environment, it has been considered that these arid systems would be more highly disturbed through urbanization and agriculture than in more temperate counterparts (Hodge 1970; Jackson and Comus 1999). The fragility of the desert
ecosystem therefore allows for an increased opportunity to observe subtle long-term ecological changes in this region.

The LTER program was established because of the importance of historical change to the understanding of present and future ecological issues (Turner et al. 2003). The subtle, complex progression of many environmental processes operate in large time frames and thus are unidentifiable in short-term projects. LTERs have been developed in an attempt to combat this problem, and facilitate the documentation of the range of long-term changes in many locations (Hobbie et al. 2003). At the CAP LTER region, the land use history of the site can be reconstructed for over 100 years, and this information allows temporal changes in ecological communities due to land use change to be evaluated. Irrespective of current land use, it is considered that 100 years ago the majority of the area encompassing the CAP LTER existed in the form of undeveloped natural desert. The considerable amount of land use change has occurred recently and rapidly. The proportion of the CAP LTER region converted to urban land type has increased from 0% in 1912 to over 18% in 1995 (Knowles-Yànez et al. 1999). Yet, even within areas of greatest change, i.e. the city center, remnant patches of natural desert environment still remain. Direct comparison of prokaryotic community composition of samples categorized by one land use type to another can be investigated for significant changes in community structure. Certain taxonomic groups within the Bacteria have been identified as major components of natural desert ecosystems in terms of abundance (Barns et al. 1999; Dunbar et al. 2002). The ability of community members to maintain a presence within the ecosystem despite these land use changes can now be investigated.
Overall, the CAP LTER site provides a unique opportunity to study the possible effects of anthropogenic changes on soil bacterial community assembly as well as the biogeography of individual phylotype groups. Potential shifts in the bacterial diversity of agricultural soils can also be assessed to supplement other studies that reveal increased species richness amongst cultivated lands (Tørsvik and Ovreas 1998; Buckley and Schmidt 2001; Buckley and Schmidt 2003; Girvan et al. 2003). Furthermore, comparison of the bacterial diversity within desert remnant soils to both open desert and urban samples can provide insight on the degree of disturbance that these sites sustain due to their proximity to the urban sector.

First, using the CAP LTER soils as a model system, combined with various geochemical and ecological data surveyed, we tested whether bacterial composition of any sample is related to environmental parameters such as land use, as well as temporal parameters based on land history. The experimental approach involved using DNA extraction from soil samples, 16S rRNA gene clone library construction, ARDRA fingerprinting, gene sequencing, and phylogenetic analysis. The data collected was then used in various statistical analyses to test for relationships between bacterial diversity and land use type. The findings suggest that urbanization and cultivation of soil in any manner has increased the total soil bacterial diversity, although it appears to play a bigger role in shaping some bacterial phyla than others. This correlation cannot be explained by the biogeochemical data collected, except for those variables involved with watering/irrigation type. The increased water content associated with the urbanization and agriculture soils coupled with the increased availability of various metabolic substrates appear to drive this trend in CAP LTER soils. The results presented here can
be used to better understand how urban ecology is affecting the capacity for soils to function as a viable ecological system.

Second, a biogeographical study of phylotypes within the 16S rRNA gene clone libraries was also assembled through phylogenetic tree construction. This analysis revealed the presence of many phylotype groups in soils of an assortment of land use types, indicating a widespread distribution of these taxa across the site that is likely the result of a wide variety of dispersal and survival stages. The majority of phylogenetic clades that appeared to be endemic to individual land use types were highly diverged from other phylotypes in the Genbank database. The overall dissimilarity with other environmental 16S rRNA sequences in the databases is evidence for endemism of many bacterial species within the CAP LTER site.

Third, using the sequence data collected from the 16S rRNA gene clone libraries, we designed PCR primers specific for phylotypes that appear in 75% of the clone libraries constructed. PCR-based assays were then performed to determine if certain bacterial phylotypes were found across all 200 CAP LTER soils from which genomic DNA was extracted. We estimate that ten phylotypes appear to be ubiquitous to CAP LTER soils. These cosmopolitan phylotypes are of particular interest because of an apparent ability to survive in a variety of habitats, and further genomic analysis may lead to increased awareness of the physiological robustness of these taxa.
CHAPTER 2:

THE INFLUENCE OF LAND USE ON BACTERIAL DIVERSITY
Introduction

Knowledge of the diversity of complex bacterial assemblages in the environment has increased with the development of more advanced culture-independent methodologies (Ward et al. 1990; Amann et al. 1995; Hugenholtz 1998b). An assemblage is defined here as a collection of bacterial phylotypes in a given space. Assemblages can include phylotypes in a cryptobiotic state that are not interacting with other populations. They differ from communities, where the populations are interacting with one another either directly or indirectly. In the early stages of culture-independent studies, single 16S rRNA gene clone libraries consisting of less than one hundred clones were constructed from one environmental sample and were primarily aimed at understanding the extent of uncultured prokaryotic diversity (Liesack and Stackebrandt 1992; Stackebrandt et al. 1993; Moyer et al. 1994; Schuppler et al. 1995; Haddad et al. 1995). Recently, studies coupling larger 16S rRNA gene clone libraries (>250 clones) with one of many fingerprinting techniques have been implemented repeatedly to provide quantitative estimates of the richness and evenness of prokaryotic assemblages across multiple environments. (Martinez-Murcia et al. 1995; Smit et al. 1997; McCaig et al. 2001; Ramirez-Monero et al. 2003; Hackl et al. 2004). Progress has also been accelerated by the use of bacterial diversity estimates that are imperative in understanding assembly of bacterial communities (Hughes et al. 2001; Bohannon and Hughes 2003). Furthermore, ecological statistics involving phylogenetic analyses have been converted for use in comparing overall microbial sequence diversity between samples (Singleton et al. 2001; Martin 2002; Schloss et al. 2004). The cumulative application of these methods not only increases understanding of the parameters that influence overall bacterial species
composition and diversity, but also aids in the determination of the ecological roles certain bacterial groups fulfill (Staley and Gosnik 1999).

The employment of culture-independent techniques may be crucial in the study of the prokaryotic diversity in soils, where empirical and reassociation kinetics studies suggest the highest extent of bacterial diversity exists (Torsvik et al. 1990; Dunbar et al. 2002). Soils are complex consortia of aggregated microhabitats, allowing for the colonization of a diverse assemblage of prokaryotic species across a very small distance (Metting 1992; Zhou et al. 2002). Recent studies have attempted to explain observed changes in prokaryotic diversity in soil based on geochemical gradients caused by a variety of abiotic and biotic factors (Tørsvik et al. 1996; Furlong et al. 2002; Zhou et al. 2002; Humanyoun et al. 2003; Borneman and Triplett 2003). However, due to the inefficiency of culture-independent methods involving the construction of small 16S rRNA gene, clone libraries have not completely revealed the entire diversity of soil environments. This undersampling effect coupled with the absence of replication diminishes the statistical power of many culture-independent studies (Valinsky et al. 2002).

One factor potentially involved in the observed bacterial species heterogeneity in soil may be a result of anthropogenic changes in land use type. It is evident that mass extinction events involving certain eukaryotic species have occurred as a result of the disruption of pristine habitats caused by urbanization (Vale and Vale 1976; Luniak 1994; Kowarik 1995; Marzluff 2001; McKinney 2002) and agricultural development (Collins 1984; Bennington et al. 2002) resulting in a loss of biodiversity. Moreover, a decrease in biodiversity relative to birds and mammals has also been reported for protected sites in
Yellowstone National Park, presumably due to the increasing human population surrounding it (Hansen et al. 2002). Loss of biodiversity can be detrimental to ecosystem function, and can have adverse effects on the quality of life, in terms of food, medicine, and pollution (Pimentel et al. 1997).

However, little is known about how land use changes affect microbial diversity. The few studies that have been carried out involving land use variation within agricultural soils have shown significant shifts in microbial species composition. Boddington and Dodd (2000) observed increases in arbuscular mycorrhizal fungi within cultivated lands due to increased plant populations. Girvan et al. (2003) found distinct changes in prokaryotic communities within arable farm soils with different farm land management practices. Furthermore, the findings of Buckley and Schmidt (2001; 2003) indicate that throughout the world, agro-ecosystems display similar distributions of higher-order taxonomic groups that differentiate them from uncultivated soils. Changes in community DNA composition were also observed between geographically distinct grasslands in the United Kingdom categorized by vegetation type (Clegg et al. 2000). Models involving soil types that have been subjected to other forms of anthropogenic disruption have not been tested in this manner. We hypothesize that urbanized soils, like their agricultural counterparts, have been disturbed to a point in which bacterial assemblages as well as overall bacterial diversity have been significantly altered.

The Central Arizona-Phoenix Long Term Ecological Research (CAP LTER) site is located in the Sonoran desert within the southwestern United States. It represents an arid landscape containing a diverse assemblage of land use types. In the past 100 years portions of the land encompassing the CAP LTER site have been converted from
undisturbed desert soil with rugose crusts to disturbed agricultural as well as urbanized soils, ranging from residential to industrial settings (McIntyre et al. 2001). Additionally, patches of desert surrounded by the urban landscape, known as “desert remnants” lie near or within the boundaries of the Phoenix metropolitan area. The study of these soils could provide an estimate of how recent urbanization events affect nearby undisturbed desert soil communities. Despite the relative proximity of desert remnant soils to the high population areas, these soils often display higher clay content, lower pH, and lower C/N ratios than urban soils, suggesting that these soils represent a potentially novel habitat (Green and Oleksyszyn 2002). The confinement of several diverse land use types within a small geographical region (~6400 km²) make this urban LTER site a suitable model for examining the possible changes in biodiversity across land use gradients (Grimm et al. 2001). Land use has been shown to be a determinant in changes in plant diversity, mainly due to the increased introduction of non-native plant species in urban areas (Hope et al. 2002). Furthermore, land use was found to be important in explaining system wide diversity of arbuscular mycorrhizal fungi (AMF) community structure in the region with increased numbers of AMF species located in the urbanized soils (Cousins et al. 2003). This study adds a prokaryotic component or more specifically a bacterial component to the study of the CAP LTER site to determine if these apparent ecological trends also apply to bacterial species diversity.

16S rRNA gene clone libraries of 22 CAP LTER soils of varying land use types were constructed, generating as many as 522 clones per sample (Table 2.1). Amplified Ribosomal DNA Restriction Analysis (ARDRA) was used as a DNA fingerprinting technique for the efficient resolution of the phylogenetic identity of 16S rRNA gene
Table 2.1- Description of CAP LTER soils used in this study

<table>
<thead>
<tr>
<th>Soil Type</th>
<th>Soil ID</th>
<th>Land Description</th>
<th>No. of clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural</td>
<td>Q12</td>
<td>Agricultural Land; Cropland; Active</td>
<td>475</td>
</tr>
<tr>
<td>Agricultural</td>
<td>R18</td>
<td>Agricultural Land; Cropland; Fallow</td>
<td>460</td>
</tr>
<tr>
<td>Agricultural</td>
<td>V18</td>
<td>Agricultural Land; Cropland; Fallow</td>
<td>458</td>
</tr>
<tr>
<td>Mixed</td>
<td>L14</td>
<td>mixed between agricultural field and open desert</td>
<td>500</td>
</tr>
<tr>
<td>Open</td>
<td>AF14</td>
<td>Open; Desert</td>
<td>491</td>
</tr>
<tr>
<td>Open</td>
<td>AG20</td>
<td>Open; Desert</td>
<td>442</td>
</tr>
<tr>
<td>Open</td>
<td>F12</td>
<td>Undisturbed desert; Riparian Zone</td>
<td>459</td>
</tr>
<tr>
<td>Open</td>
<td>I11</td>
<td>Undisturbed desert</td>
<td>508</td>
</tr>
<tr>
<td>Open</td>
<td>I17</td>
<td>Undisturbed desert</td>
<td>489</td>
</tr>
<tr>
<td>Open</td>
<td>J7</td>
<td>Undisturbed desert</td>
<td>409</td>
</tr>
<tr>
<td>Open</td>
<td>K9</td>
<td>Undisturbed desert</td>
<td>522</td>
</tr>
<tr>
<td>Open</td>
<td>U20</td>
<td>Undisturbed desert</td>
<td>458</td>
</tr>
<tr>
<td>Open</td>
<td>V7</td>
<td>Undisturbed desert</td>
<td>480</td>
</tr>
<tr>
<td>Remnant</td>
<td>U21</td>
<td>Desert park remnant</td>
<td>496</td>
</tr>
<tr>
<td>Remnant</td>
<td>V11</td>
<td>Desert park remnant</td>
<td>480</td>
</tr>
<tr>
<td>Remnant</td>
<td>X14</td>
<td>Residential lot remnant</td>
<td>465</td>
</tr>
<tr>
<td>Remnant</td>
<td>X15</td>
<td>Residential lot remnant</td>
<td>509</td>
</tr>
<tr>
<td>Urban</td>
<td>AC17</td>
<td>Urban; Residential; Mesic</td>
<td>478</td>
</tr>
<tr>
<td>Urban</td>
<td>M9</td>
<td>Urban; Open; Vacant</td>
<td>458</td>
</tr>
<tr>
<td>Urban</td>
<td>V13</td>
<td>Urban; Non-Residential; Institutional</td>
<td>501</td>
</tr>
<tr>
<td>Urban</td>
<td>W13</td>
<td>Urban; Transportation; Highway</td>
<td>495</td>
</tr>
<tr>
<td>Urban</td>
<td>X13A</td>
<td>Urban; Residential; Medium Lot; Single family; Mixed</td>
<td>341</td>
</tr>
<tr>
<td>Urban</td>
<td>X13B</td>
<td>Urban; Residential; Medium Lot; Single family; Mixed</td>
<td>462</td>
</tr>
</tbody>
</table>

clones, and unique ARDRA patterns were sequenced to obtain overall phylogenetic composition of the entire clone library. The relative abundances of the 16S rRNA genes were collected to analyze differences in phylum-level diversity among the land use types. Using this large set of data (over 11,000 ARDRA fingerprints), various statistical analyses were performed to compare 16S rRNA gene clone libraries in order to observe shifts in patterns of diversity. The results suggest that bacterial
diversity has increased in CAP LTER soils disturbed by agricultural or urbanization, and that bacterial communities in desert remnant patches are being influenced by bordering urban areas.

Materials and Methods

Sample Collection

The samples used in the study were part of the 204 soil samples collected aseptically from various recorded sites of the CAP LTER using a dual-density, randomized, tessellation-stratified design to obtain an unbiased, spatially-dispersed sample that allowed for maximum postdesign stratification (Fig. 2.1). This sampling approach ensures representative and unbiased characterization of ecological resources.
The sampling unit was a 30 X 30m plot. One hundred gram soil samples were dug from the upper 5 cm of soil crust using a sterile scoop. The sampling area within the plot was chosen based on the absence of vegetation and leaf litter. Classification of land use type at each of the surveyed sites was according to the Maricopa Association of Governments land-use classification scheme (1997). The five main land use categories included urban (n=91), desert (n=73), agricultural (n=23), transportation (n=6), and a “mixed” class (n=11). The soil classifications differed from the Maricopa Association of Governments scheme in two ways: the addition of a mixed category for sites at which more than one land use type was within an individual survey plot, and a desert remnant subcategory within the desert samples that were defined as desert patches completely surrounded by an urban landscape. Once collected, soils were stored in a dark room for a period of three months prior to DNA extraction.

**DNA Extraction and Purification**

Bulk environmental DNA was extracted by methods modified from the protocol of Herrick et al. (1993). Two grams of soil was initially mixed with lysozyme and 10% SDS to chemically lyse cell membranes. Three freeze/thaw cycles were then performed as a physical lysing method. Proteinase K was added to inhibit or degrade proteins. Polyethylene glycol (30% w/v) was added to precipitate the DNA at 4°C. DNA was further purified using a phenol/chloroform extraction and a Prep-A-Gene kit (BioRad Industries).

**PCR and Cloning**

Community 16S rRNAs were amplified by PCR from 20 to 200 ng of DNA in reaction mixtures totaling 50-µl and containing (as final concentrations) 1X PCR Buffer
(Applied Biosystems), 200 nM each deoxynucleotide triphosphate, 50 nM each forward and reverse primer, and 1 U of Amp-Taq polymerase (Roche). Reaction mixtures for universal primers were incubated in a Geneamp PCR system (Perkin-Elmer) at 98\(^0\)C for 6 minutes (for initial denaturation); after cooling to 90\(^0\)C 1U of Amp-Taq polymerase was added followed by 28 cycles at 52\(^0\)C for 30 s, 54\(^0\)C for 1 min, and 94\(^0\)C for 30 s; followed by a final extension period of 7 min at 72\(^0\)C. For all clone libraries, rRNAs were amplified with the primers 27F (specific for Bacteria) (5’-AGAGTTTGATCCTGGCTCAG-3’) and 1492R (5’-GGTTACCTTGTTACGACTT-3’). PCR products were cloned with a TOPO XL cloning kit in accordance with the manufacturer’s instructions (Invitrogen Corporation).

**ARDRA and Sequencing**

Initially, a 0.5 µL aliquot of each cloned PCR product mixture was amplified from generated *E. coli* host colonies using the phage primers M-13F (5’-TGTAAGACGACGCAGGTG-3’) and M-13R (5’-GTACGACGACGTTATG-3’) in a reaction mixture containing 1X Buffer, 200 nM of each deoxynucleotide triphosphate, and 50nM each primer. A 10 mL aliquot of amplified M13 PCR products was then digested with TaqI restriction endonuclease for 1 h at 65\(^0\)C. Those TaqI fingerprints that were identical for two or more clones were digested separately with RsaI restriction endonuclease for 1 h at 37\(^0\)C. The digestion fragments were separated using NuSeive agarose (BioWhittaker Molecular Applications) (3% wt/vol) and subjected to electrophoresis in 1X TBE buffer for 6 h. The gel was stained with 10mL of ethidium bromide (0.5 µg/mL) and visualized by UV transillumination. Restriction fragment profiles were subjected to complete-linkage cluster analysis by the Pearson coefficient to
identify similarity using Bionumerics software (Applied Maths). Unique ARDRA fingerprint clusters were designated as operational taxonomic units (OTUs). PCR products were purified and concentrated by using a Prep-A-Gene kit (Bio-Rad). DNA was eluted in 20 µL of distilled H$_2$O, and 1.0 µL of the resulting preparation was used in sequencing reactions. The sequencing reactions were performed with a Big Dye terminator cycle sequencing kit and a BioRad iCycler 170 thermal cycler according to the thermal profile that included one initial denaturation step at 94ºC followed by 25 cycles at 96ºC for 30 s, 50ºC for 5 s, and 60ºC for 4 min. Sequence reaction mixtures were purified by ethanol precipitation. Inserts were sequenced on an ABI 3100 automated DNA sequencer (Applied Biosystems International) using the 27-forward primer. Sequences were then manually aligned and analyzed using BioEdit (version 5.09) (Hall 1999). Representative fingerprints were sequenced and imported into the BLAST search utility of the National Center for Biotechnology Information nucleotide database for determination of similar to 16S rRNA gene sequences in the database (Altschul et al. 1990).

**Analysis of 16S rRNA Gene Sequence Data**

Based on the BLAST search results and phylogenetic placement using ARB software package (Strunk et al. 2004), clones were designated into phylogenetic groups. Sequence data corresponding to bases of 101 to 578 in terms of *Escherichia coli* numbering (Brosius et al. 1978) were used to generate phylogenetic trees using the neighbor joining method (Saitou and Nei 1987) for phylogenetic identification as well as to calculate evolutionary distance matrices using the Jukes-Cantor algorithm (Jukes and Cantor 1969) in the PHYLIP software package (Felsenstein 1989).
**Chimera Checking**

All clones exhibiting less than 90% sequence similarity to existing GenBank sequences and/or containing secondary structure anomalies were analyzed using the CHIMERA_CHECK program (version 2.7) of the Ribosomal Database Project (RDP) (Maidak et al. 2001; Cole et al. 2003). A portion of the sequences were also submitted to the BELLEPHRON program (http://foo.maths.uq.edu.au/~huber/bellerophon.pl), which is designed to check for chimeric sequences from 16S rRNA gene clone libraries. The proportions of overall chimeras found in each individual sample was below 3%.

**Sample Coverage and Species Richness Estimates**

The reciprocal of Simpson’s index (1/D) was used as the diversity estimate due to its significant discriminating ability and frequent use in ecological studies (Magurran 1988). Species richness and sample coverage calculations were performed with the program EstimateS (version 5.0.1; R. K. Colwell, University of Connecticut [http://vicerory.eeb.unconn.edu/estimates]). One hundred randomizations were calculated for all tests. The estimated total species richness was calculated for each 16S rRNA gene clone library sample using the nonparametric estimators ACE (abundance-based coverage estimator) and Chao1 (Chao 1987; Chao 1992; Hughes et al. 2001). Since the distribution of estimates is not normal, Burnham’s log transformation is used to determine 95% confidence intervals of the total bacterial species: $S + [(N - S)/C], S +[(N - S)C]$, where $C = \exp(1.96\{\log[1 + \sigma^2 (N - S)^2]\})^{0.5}$ (Chao 1987, Chao and Lee 1992). Accumulation curves were produced by using software available online at http://www.uga.edu/~starta/software.html. See Appendix A for individual diversity values.
CAP LTER Variables

A total of thirteen variables were chosen to represent the main geochemical, geographic, and socioeconomic characteristics of the study site (Appendix B), while having minimal colinearity values. These include: years since land use changed, elevation, distance from urban center, population density, income per capita, vegetation cover, percentage clay, watering type, percentage total nitrogen, NH$_4$ concentration, NO$_3$ concentration, percentage organic carbon, and percentage total carbon. Distances between sites were obtained from GPS coordinates of sample sites.

Statistical Analyses

Data used in all tests followed a normal distribution according to the Shapiro-Wilkes test unless noted otherwise. Percent abundance was arcsine transformed when statistical tests were performed. The statistical significance of differences in the compositions of pairs of sequence libraries of approximately 500 base pairs in length was tested with the LIBSHUFF program using genetic distance matrices generated as described above (Singleton et al. 2001). Statistical analysis using linear and multiple regression models were performed using SAS statistical software version 9. When applicable, Generalized Linear Modeling (GLM) was selected for fitting datasets, and when overall treatments mean variances were large (F*>1), the least squared distance correction was used for pairwise comparisons. Complete-linkage cluster analysis was also performed in SAS using PROC CLUSTER for fitting Sørenson’s similarity data and $\Delta$C values representing a Cramér-von Mises-type statistic, a measure of the area between homologous and heterologous curve data in the LIBSHUFF analysis. Sørenson’s similarity index (CC$_s$) was used to estimate the presence and absense of shared ARDRA
fingerprints between two samples, were calculated by using the equation \( CC_s = \frac{2c}{s_1 + s_2} \), where \( s_1 \) and \( s_2 \) represent the number of species in assemblages 1 and 2 respectively (i.e., two given sites) and \( c \) is the number of species common to both communities (Sørenson 1948). Morisita-Horn similarity coefficients (\( C_{MH} \)), which were used to compare the richness of common ARDRA fingerprints found at pairs of sampling locations, were calculated by using the equation 
\[
C_{MH} = \frac{2 \times \sum (A_i \times B_i)}{N_A \times N_B \times \left( \frac{\sum A_i^2}{N_A^2} + \frac{\sum B_i^2}{N_B^2} \right)},
\]
where \( A_i \) and \( B_i \) are the numbers of isolates of genotype \( i \) in samples \( A \) and \( B \), respectively, and \( N_A \) and \( N_B \) are the total numbers of isolates in samples \( A \) and \( B \), respectively (all \( \Sigma \) are summed from \( i = 1 \) to \( S \), the total number of species) (Wolda 1983). Ordination scores and correlation data were gathered using Principal Components Analysis (PCA) performed using MiniTab version 12. Variance inflation factors between the independent variables used in these analyses never exceeded ten, and most were substantially less, indicating that collinearity is not a significant problem. For increased statistical power, we assume a priori a 10% probability of Type I error instead of the more common assumption of 5% (as discussed by Gill 1978). Thus, we report those correlations as “significant” where \( P < 0.05 \) and “suggestive” where \( P < 0.1 \).

Results

OTU Definition

The use of multiple restriction enzymes in this study allowed for the presence of ARDRA fingerprint groups with high levels of relatedness, which permitted the designation of each unique ARDRA fingerprint as an individual OTU. To quantify the genetic relatedness for three fingerprint groups that differ by one band, the nearly complete 16S rRNA gene insert was sequenced (~1400 nucleotides). These three groups
had sequence similarities of 96.8± 2.3%. Differences in 16S rRNA gene sequences of more than 3% are typically compatible to <70% DNA-DNA hybridization, the current standard in the definition of bacterial species (Wayne et al. 1987; Stackebrandt and Geobel 1994). Therefore, it was assumed in this study that each unique ARDRA fingerprint represents a different taxon at the species level or above (genus, family, order, etc.) under the current guidelines set forth.

**Phylogeny of 16S rRNA Gene Clones**

Twenty two CAP LTER soils were chosen for 16S rRNA gene clone library construction, including one soil (X13) from which two separate clone libraries were constructed from two separate environmental DNA extracts taken from the same sample in order to test the reproducibility of the results generated. From the total of 11,096 clones analyzed, 1,465 OTUs were categorized into 17 existing or candidate phyla (Table 2.2).

Representatives of the alpha proteobacteria were the most commonly sampled group present with an overall mean abundance of 46.6% within all 16S rRNA gene clone libraries constructed. Other bacterial groups present in large numbers included the *Acidobacteria* (11.1%), *Gemmatimonadetes* (8.3%), *Actinobacteria* (7.8%), *Firmicutes* (5.9%), and beta proteobacteria (5.4%). The phyla *Bacteriodetes* (2.9%), *Verrucomicrobia* (2.2%), and *Planctomycetes* (1.6%) were also observed, but in all cases these groups were nearly undetectable in at least one sample (<0.6% abundance).

Members of *Cyanobacteria* were only found in high numbers (19.8%) in the soil L14 that was categorized as mixed because of its location at the border between land use categories, open and agricultural, within the sample plot. All other land use categories
Table 2.2 - Distribution of clone OTUs in phylogenetic groups

<table>
<thead>
<tr>
<th>Phylogenetic group</th>
<th>Open Desert (%)</th>
<th>Desert Remnant (%)</th>
<th>Urban (%)</th>
<th>Agricultural (%)</th>
<th>Mixed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J7</td>
<td>44.3</td>
<td>64.0</td>
<td>45.4</td>
<td>43.9</td>
<td>41.8</td>
</tr>
<tr>
<td>F12</td>
<td>66.8</td>
<td>47.4</td>
<td>51.9</td>
<td>32.5</td>
<td>33.9</td>
</tr>
<tr>
<td>AF14</td>
<td>73.2</td>
<td>65.2</td>
<td>58.8</td>
<td>66.5</td>
<td>40.6</td>
</tr>
<tr>
<td>K9</td>
<td>76.9</td>
<td>42.7</td>
<td>59.7</td>
<td>40.0</td>
<td>47.4</td>
</tr>
<tr>
<td>I7</td>
<td>87.9</td>
<td>42.7</td>
<td>40.4</td>
<td>31.9</td>
<td>75.8</td>
</tr>
<tr>
<td>V7</td>
<td>50.0</td>
<td>35.6</td>
<td>57.2</td>
<td>58.6</td>
<td>43.9</td>
</tr>
<tr>
<td>AG20</td>
<td>18.2</td>
<td>34.5</td>
<td>25.2</td>
<td>50.0</td>
<td>61.0</td>
</tr>
<tr>
<td>I11</td>
<td>5.5</td>
<td>7.5</td>
<td>6.0</td>
<td>48.1</td>
<td>30.0</td>
</tr>
<tr>
<td>U20</td>
<td>4.0</td>
<td>7.5</td>
<td>4.0</td>
<td>66.5</td>
<td>31.9</td>
</tr>
<tr>
<td>V11</td>
<td>1.9</td>
<td>4.8</td>
<td>4.0</td>
<td>58.5</td>
<td>48.1</td>
</tr>
<tr>
<td>U21</td>
<td>0.8</td>
<td>4.9</td>
<td>0.4</td>
<td>61.0</td>
<td>47.4</td>
</tr>
<tr>
<td>X14</td>
<td>2.3</td>
<td>2.2</td>
<td>2.4</td>
<td>58.6</td>
<td>50.0</td>
</tr>
<tr>
<td>X15</td>
<td>1.9</td>
<td>3.8</td>
<td>2.2</td>
<td>41.8</td>
<td>40.6</td>
</tr>
<tr>
<td>X13A</td>
<td>45.2</td>
<td>3.8</td>
<td>3.8</td>
<td>47.4</td>
<td>47.4</td>
</tr>
<tr>
<td>X13B</td>
<td>56.0</td>
<td>3.8</td>
<td>3.8</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>M9</td>
<td>40.4</td>
<td>3.8</td>
<td>3.8</td>
<td>58.5</td>
<td>58.5</td>
</tr>
<tr>
<td>AC17</td>
<td>14.1</td>
<td>3.8</td>
<td>3.8</td>
<td>43.9</td>
<td>43.9</td>
</tr>
<tr>
<td>W13</td>
<td>12.6</td>
<td>3.8</td>
<td>3.8</td>
<td>32.5</td>
<td>32.5</td>
</tr>
<tr>
<td>V13</td>
<td>2.4</td>
<td>3.8</td>
<td>3.8</td>
<td>33.9</td>
<td>33.9</td>
</tr>
<tr>
<td>R18</td>
<td>1.9</td>
<td>3.8</td>
<td>3.8</td>
<td>41.8</td>
<td>41.8</td>
</tr>
<tr>
<td>V18</td>
<td>0.0</td>
<td>3.8</td>
<td>3.8</td>
<td>40.6</td>
<td>40.6</td>
</tr>
<tr>
<td>Q12</td>
<td>0.0</td>
<td>3.8</td>
<td>3.8</td>
<td>41.8</td>
<td>41.8</td>
</tr>
<tr>
<td>L14</td>
<td>0.0</td>
<td>3.8</td>
<td>3.8</td>
<td>40.6</td>
<td>40.6</td>
</tr>
</tbody>
</table>

- Proteobacteria: 38.4%
- Acidobacteria: 5.9%
- Actinobacteria: 27.9%
- Firmicutes: 2.0%
- Bacteroidetes: 3.2%
- Gemmatimonadetes: 5.5%
- Planctomycetes: 0.5%
- Verrucomicrobia: 0.5%
- Cyanobacteria: 1.0%
- Chloroflexi: 0.0%
- Thermus-Deinococcus: 0.0%
- Dehalococcoides: 0.0%

<table>
<thead>
<tr>
<th>Phylum</th>
<th>OTU ID</th>
<th>Open Desert (%)</th>
<th>Desert Remnant (%)</th>
<th>Urban (%)</th>
<th>Agricultural (%)</th>
<th>Mixed (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidobacteria</td>
<td>59</td>
<td>8.7</td>
<td>4.2</td>
<td>10.0</td>
<td>6.8</td>
<td>4.0</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>27.9</td>
<td>6.8</td>
<td>7.3</td>
<td>2.3</td>
<td>2.9</td>
<td>2.5</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>2.0</td>
<td>2.2</td>
<td>1.8</td>
<td>1.9</td>
<td>2.0</td>
<td>2.3</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>3.2</td>
<td>2.8</td>
<td>0.6</td>
<td>3.6</td>
<td>3.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Gemmatimonadetes</td>
<td>5.5</td>
<td>7.4</td>
<td>8.8</td>
<td>7.9</td>
<td>7.4</td>
<td>8.1</td>
</tr>
<tr>
<td>Planctomycetes</td>
<td>0.5</td>
<td>1.7</td>
<td>1.6</td>
<td>0.8</td>
<td>0.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Verrucomicrobia</td>
<td>1.7</td>
<td>0.4</td>
<td>0.4</td>
<td>1.9</td>
<td>3.3</td>
<td>1.9</td>
</tr>
<tr>
<td>Cyanobacteria</td>
<td>0.5</td>
<td>2.8</td>
<td>1.0</td>
<td>0.4</td>
<td>2.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Chloroflexi</td>
<td>1.0</td>
<td>0.4</td>
<td>0.7</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Thermus-Deinococcus</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Dehalococcoides</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>OP10</td>
<td>1.5</td>
<td>0.6</td>
<td>0.6</td>
<td>1.1</td>
<td>1.8</td>
<td>0.6</td>
</tr>
<tr>
<td>OP8</td>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>OP9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>OP1</td>
<td>0.2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>OP11</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Unknowna</td>
<td>3.4</td>
<td>1.5</td>
<td>1.4</td>
<td>1.9</td>
<td>1.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

*Values represent percentage of clones belonging to a phylogenetic group

*Sum of alpha, beta, gamma, and delta proteobacteria

*Clones which did not correlate with a known phylum.
were typically low in *Cyanobacteria* abundance on average (1.1%). In the soils
examined, there were also 5 other distinct phylogenetic groups detected at low levels
(<1%), the gamma and delta proteobacteria, *Chloroflexi, Dehalococcoides*, and
*Deinococcus-Thermus*. Other rare phylotypes collected were identified as representatives
of OP3, OP8, OP9, and OP11 candidate divisions. The OP series designation is from
Obsidium Pool, a hot spring in Yellowstone National Park, in which they were initially
discovered (Hugenholtz et al. 1998a). These phylogenetic groups at present have no
cultured representatives, and therefore little is known of their physiological and
ecological capabilities. Due to a high probability of sampling error, the presence or
absence of these groups cannot be measured with significant statistical power, and
therefore have been removed from the following comparisons (Dunbar et al. 1999). The
remaining clones which are categorized as unknown were less than 85% similar in their
16S rRNA gene sequence to any sequence within the Genbank database, and therefore
were unable to be phylogenetically classified with high confidence under the methods
used in this study.

**Land Use Influences on Bacterial Group Abundance**

Overall bacterial group abundance based on 16S rRNA gene clone library data
was compiled. Percent abundance means of the most prevalent phylogenetic groups
were separated by land use type, and subjected to one-way ANOVA (Fig. 2.2). Land use
type was shown to be a strong predictor of the abundance of alpha proteobacteria,
*Firmicutes* and *Acidobacteria*. Pairwise differences are strong in the alpha
proteobacteria, in which the open deserts contain a significantly higher percentage of 16
rRNA gene clones than the other three land use types, while agricultural soils are
Figure 2.2 - Bacterial group clone abundance separated by land use type. F statistics shown from ANOVA for the effects of land use type when the alpha error values are suggestive (P<0.1) or significant (P>0.05). Bars topped by the same letter do not differ significantly according to the least significant difference (LSD) test. Error bars represent standard error.

Comprised of a distinctly smaller proportion. A significantly higher abundance of Acidobacteria was found in remnant and agricultural soils as compared to open desert sites. Members of the Firmicutes phylum are detected in agricultural settings to a much higher degree than in all other land use types, while the proportions of this group found in open desert samples was significantly smaller. The statistical analysis also suggests slight changes in the abundances of beta proteobacteria and Gemmatimonadetes groups across land use types (P<0.1), particularly between both remnant and urban soils versus the open desert samples (Fig 2.2D).

Multiple analysis of variance (MANOVA) based upon all the most abundant bacterial group abundance data across the four main land use types showed a significant effect (Table 2.3). Pairwise contrasts revealed distinct differences only between urban and agricultural soils, but suggestive variation between open desert and urban/remnant samples. Correspondence analysis using multivariate data was used to visualize patterns
Table 2.3 - Summary of MANOVA results.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Pillai’s Trace</th>
<th>d.f.</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landtype</td>
<td>1.46</td>
<td>21,42</td>
<td>1.91</td>
<td>0.0364</td>
</tr>
</tbody>
</table>

Pairwise tests

<table>
<thead>
<tr>
<th>Test</th>
<th>Pillai’s Trace</th>
<th>d.f.</th>
<th>F-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag v Open</td>
<td>0.72</td>
<td>7,12</td>
<td>4.36</td>
<td>0.0127</td>
</tr>
<tr>
<td>Ag v Rem</td>
<td>0.45</td>
<td>7,12</td>
<td>1.41</td>
<td>ns</td>
</tr>
<tr>
<td>Ag v Urb</td>
<td>0.45</td>
<td>7,12</td>
<td>1.41</td>
<td>ns</td>
</tr>
<tr>
<td>Open v Rem</td>
<td>0.57</td>
<td>7,12</td>
<td>2.31</td>
<td>0.0965</td>
</tr>
<tr>
<td>Open v Urb</td>
<td>0.59</td>
<td>7,12</td>
<td>2.43</td>
<td>0.0842</td>
</tr>
<tr>
<td>Rem v Urb</td>
<td>0.41</td>
<td>7,12</td>
<td>1.19</td>
<td>ns</td>
</tr>
</tbody>
</table>

in overall land use community structure and the influence of bacterial group abundance on particular land use types (Fig. 2.3). Inspection of the correspondence analysis reveals cluster zones between urban, agricultural, and open desert samples. Desert remnant datapoints are noticeably present within both urban and open desert zones. Alpha proteobacteria abundances appear to heavily influence the scaling of the open desert libraries while the low G+C Gram positive members of Firmicutes are heavily influencing the agricultural cluster zone. Group abundance means for Actinobacteria vary widely across the soils and therefore are not influencing any land use type disproportionately. All other bacterial groups within the data set appear to be stable in abundance across various land use types with ordination scores below 0.1. The
Figure 2.3 - Correspondence analysis of community structure across CAP LTER land use types. Datapoints represented by open desert (black squares), desert remnants (blue triangles), urban (red circles), and agricultural (green diamonds). Ellipses are used to visually aid differences in bacterial diversity structure. Community is represented by alpha proteobacteria (a), beta proteobacteria (b), Acidobacteria (c), Bacteriodetes (d) Firmicutes (f), Gemmatimonadetes (g), and Actinobacteria (t). The two datapoints located within the open square represent two replicates from CAP LTER sample X13.

Euclidean distance between ordination scores of the two replicate X13 datapoints differ marginally ($3.8 \times 10^{-4}$).

**Total Species Richness Estimates**

Based upon the OTUs collected through ARDRA analysis within each 16S rRNA gene clone library, accumulation curves for the CAP LTER clone libraries were constructed to determine overall sampling effort (Fig. 2.4). Accumulation curves that are linear signify that the collected sample is extremely undersampled to point in which any estimation of total species richness is invalid (Hughes et al. 2001). The accumulation curves generated are non-linear, but have not reached an asymptote, indicating that total
Figure 2.4 - Accumulation curves derived from 16SrDNA clone library data. Datapoints represent agricultural (blue triangles), open desert (yellow circles), desert remnant (gray diamonds), and urban (black square) soils.
diversity of the CAP LTER soils are undersampled in the constructed 16S rRNA gene clone libraries. When categorized by land use type, urban accumulation curves are the most linear, especially when compared to their open desert counterparts, indicating a divergence in total species richness when comparing these two land use types. Since the curves continued increasing even at the endpoint, it is impossible to calculate the amount of sampling needed to successfully differentiate any of the CAP LTER communities with confidence. However, when nonparametric Chao1 and ACE richness estimators of total species richness are calculated, the 95% CIs overlap across all estimates (Table 2.4), ranging from 4,189 to 23,761 species. Neither these soils nor the entire land use types they belong to vary significantly in terms of total species richness (α=0.05).

**Statistical Analysis Using the Reciprocal of Simpson’s Index**

The reciprocal of Simpson’s index has been used to assess the level of dominance within a community and indicates that land use type was a powerful predictor for overall bacterial diversity in CAP LTER soils (Table 2.5). Subsequent pairwise comparisons revealed significant increases in urban and agricultural diversity as opposed to undisturbed open desert soils. On average, the open and desert remnant samples have 1/D values below 50, which suggest that these communities demonstrate dominance profile, while urban and agricultural soils as a group resemble a more uniform diversity pattern with 1/D values over 50. (Zhou et al. 2002). Inter-phyla diversity estimates involving the five most abundant phyla vary widely across land use type samples with no significant stratification being observed, which is contrary to the distinct changes observed in overall bacterial diversity values. This indicates that the significant changes
### Table 2.4 - Comparison of total OTUs and estimated total species richness in CAP LTER soils.

<table>
<thead>
<tr>
<th></th>
<th>Agricultural</th>
<th></th>
<th></th>
<th>Open</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q12</td>
<td>R18</td>
<td>V18</td>
<td>AF14</td>
<td>AG20</td>
<td>F12</td>
<td>I11</td>
<td>I17</td>
<td>J7</td>
<td>K9</td>
<td>U20</td>
<td>V7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTUs</td>
<td>183</td>
<td>225</td>
<td>210</td>
<td>174</td>
<td>177</td>
<td>192</td>
<td>237</td>
<td>196</td>
<td>196</td>
<td>159</td>
<td>176</td>
<td>161</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE</td>
<td>424</td>
<td>757</td>
<td>509</td>
<td>537</td>
<td>505</td>
<td>567</td>
<td>595</td>
<td>458</td>
<td>656</td>
<td>500</td>
<td>454</td>
<td>586</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LL CI&lt;sup&gt;a&lt;/sup&gt;</td>
<td>192</td>
<td>237</td>
<td>220</td>
<td>195</td>
<td>190</td>
<td>217</td>
<td>259</td>
<td>208</td>
<td>219</td>
<td>171</td>
<td>189</td>
<td>174</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UL CI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6,398</td>
<td>23,761</td>
<td>9,019</td>
<td>6,592</td>
<td>8,459</td>
<td>5,922</td>
<td>6,072</td>
<td>6,130</td>
<td>9,951</td>
<td>9,468</td>
<td>5,998</td>
<td>13,166</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chao1</td>
<td>437</td>
<td>697</td>
<td>520</td>
<td>463</td>
<td>477</td>
<td>407</td>
<td>485</td>
<td>447</td>
<td>452</td>
<td>443</td>
<td>384</td>
<td>493</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LL CI</td>
<td>193</td>
<td>235</td>
<td>228</td>
<td>187</td>
<td>188</td>
<td>201</td>
<td>251</td>
<td>207</td>
<td>206</td>
<td>169</td>
<td>186</td>
<td>171</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UL CI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6,339</td>
<td>23,145</td>
<td>5,688</td>
<td>6,694</td>
<td>8,169</td>
<td>5,607</td>
<td>5,051</td>
<td>5,881</td>
<td>6,971</td>
<td>8,398</td>
<td>4,531</td>
<td>11,034</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Urban</th>
<th>Remnant</th>
<th>Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AC17</td>
<td>M9</td>
<td>V13</td>
</tr>
<tr>
<td>OTUs</td>
<td>188</td>
<td>184</td>
<td>196</td>
</tr>
<tr>
<td>ACE</td>
<td>438</td>
<td>626</td>
<td>651</td>
</tr>
<tr>
<td>LL CI</td>
<td>202</td>
<td>208</td>
<td>216</td>
</tr>
<tr>
<td>UL CI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4,513</td>
<td>8,405</td>
<td>10,461</td>
</tr>
<tr>
<td>Chao1</td>
<td>391</td>
<td>463</td>
<td>553</td>
</tr>
<tr>
<td>LL CI</td>
<td>198</td>
<td>196</td>
<td>209</td>
</tr>
<tr>
<td>UL CI&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4,215</td>
<td>6,866</td>
<td>9,674</td>
</tr>
</tbody>
</table>

<sup>a</sup> Lower limit value of generated 95% confidence intervals

<sup>b</sup> Upper limit value of generated 95% confidence intervals
Table 2.5- Effects of land use on CAP LTER bacterial diversity

<table>
<thead>
<tr>
<th>Phylogenetic group</th>
<th>1/D values</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Agricultural</td>
<td>Open</td>
<td>Urban</td>
<td>Remnant</td>
<td></td>
</tr>
<tr>
<td>A Proteobacteria</td>
<td>6.79</td>
<td>9.87</td>
<td>11.03</td>
<td>8.49</td>
<td>0.29 (ns)</td>
</tr>
<tr>
<td>B Acidobacteria</td>
<td>38.20</td>
<td>28.03</td>
<td>33.07</td>
<td>21.23</td>
<td>0.85 (ns)</td>
</tr>
<tr>
<td>C Actinobacteria</td>
<td>32.50</td>
<td>38.08</td>
<td>41.42</td>
<td>50.05</td>
<td>0.28 (ns)</td>
</tr>
<tr>
<td>D Firmicutes</td>
<td>5.67</td>
<td>18.01</td>
<td>5.67</td>
<td>11.78</td>
<td>0.85 (ns)</td>
</tr>
<tr>
<td>E Gemmatimonadetes</td>
<td>38.10</td>
<td>38.91</td>
<td>39.10</td>
<td>22.64</td>
<td>0.85 (ns)</td>
</tr>
<tr>
<td>F Bacteria</td>
<td>61.93</td>
<td>24.79</td>
<td>51.03</td>
<td>42.18</td>
<td>3.24 (0.0479)</td>
</tr>
<tr>
<td>Urban-Open</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.74 (0.0284)</td>
</tr>
<tr>
<td>Agric.-Open</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.02 (0.0187)</td>
</tr>
</tbody>
</table>

*a*Effects of land use on domain and phylum-level diversity. All *F* and *P* values are derived from analysis of variance (ANOVA), with *P* values determined on the basis of *F*(3,17) for the experimental study.

*b*Pairwise comparisons of domain level diversity using Least Squared Distance

overall diversity values are not explained by the more abundant groups but the rarer taxa that were not included in the analysis.

**Other Potential Effects on Diversity**

Simpson’s reciprocal index values were also used as a dependent variable in Principal Components and simple linear regression analysis with various biogeochemical, geographic, and urban numerical data collected from the same Survey 200 sites (Fig. 2.5). After the normalization of the variable set, multiple regression analysis was performed, in which no combination of these variables allowed for cumulative explanation of the variance. Furthermore, all simple linear regression models using Simpson’s reciprocal index values as a response variable to individual continuous
Figure 2.5 - Results of a statistical analysis for the effect of ecological factors on bacterial diversity. All tested relationships are indicated by arrows. Measurements of variables are included in methods. Bacterial diversity refers to values of Simpson’s reciprocal index. Numbers at arrows represent partial correlation coefficients in multiple regression analysis. Numbers within boxes equals percentage of explained variation ($r^2$ adjusted) in one-way ANOVA. Asterisks represent overall significance when testing the variable against land use type in a 1-way ANOVA, *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

variables failed to adequately explain the observed variability across the samples ($r^2<0.32$). However, many of the variable means, in particular those that are correlated with urbanization (population density, distance from urban center, vegetation cover) are significantly different across land use type, and suggest that land use types are distinct ecological systems, particularly within the anthropogenically-disturbed soils.
Table 2.6 - Effects of watering type on CAP LTER bacterial diversity.

<table>
<thead>
<tr>
<th>Watering Type</th>
<th>Open</th>
<th>Remnant</th>
<th>Urban</th>
<th>Agricultural</th>
<th>1/D</th>
<th>F value (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flood</td>
<td>3</td>
<td>62.0</td>
<td></td>
<td></td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Hand</td>
<td>4</td>
<td>61.1</td>
<td></td>
<td></td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>32.0</td>
<td>5.42 (0.0132)</td>
</tr>
<tr>
<td>Overall</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>6.38</td>
<td>(0.0081)</td>
</tr>
</tbody>
</table>

aEffects of watering treatment on bacterial diversity. Overall experimental F and P values are derived from analysis of variance (ANOVA), with P values determined on the basis of F(2,21) for the experimental study. The single mixed soil was not subjected to analysis.

bNumber of soils categorized by corresponding watering type.

cFlood is defined as soils which are subjected to a flood irrigation system. Hand watering is defined as soils that are irrigated by hand twice a week. “None” watering type is defined as soil that receive no artificial watering.

dSince only land use category contains all watering type observations, ANOVA is not applicable because there is zero degrees of freedom in the model.

eSignificant difference between Flood (P=0.0256) and Hand (P=0.0160) watering types according to LSD pairwise comparisons.

However, when another form of categorical data, “watering type”, was tested as the lone treatment effect, it demonstrated a highly significant relationship with diversity (Table 2.6). The only dichotomous land use type using watering method data was that of the urban soils, in which four designated under “hand watering” and two were classified as “none”. Surprisingly, the diversity of the non-hand watered soils are below the uniform distribution threshold (1/D=50) as opposed to the other urban soils that are significantly higher, indicating that watering may be the prime factor for the observed differences in overall diversity within land use types.

Morisita-Horn’s index values, which indicate the amounts of shared OTUs between two sites, were derived from the ARDRA fingerprint database. Similarity scores were plotted against the respective distances between pairs of Survey 200 soils (Fig. 2.6). A simple linear regression analysis of these two variables indicates that distance is a
Figure 2.6 - \( C_{\text{MH}} \) values for comparisons of ARDRA fingerprint assemblages across CAP LTER samples as a function of distance.

weak indicator of two samples similarity to one another (\( r^2=0.02 \)), and a slight increase in similarity as a function of distance is observed from the best-fit line produced.

**LIBSHUFF Comparisons of the Clones**

As expected from the wide range of diversity values mentioned above, as well as the large numbers of 16S rRNA gene clones analyzed, 325 direct pairwise comparisons of CAP LTER samples revealed no comparisons in which samples differ significantly using the Bonferroni correction \( (P>0.0001) \) (data not shown). However, direct comparisons of the replicate 16S rRNA gene sequence libraries (X13A and B) were highly similar (Fig. 2.7).

Because diversity estimates are often more indicative of community change within specific groups (Hill et al. 2003), several phyla were also chosen as individual
templates for LIBSHUFF analysis (Table 2.7). *Acidobacteria*, *Firmicutes*, and *Actinobacteria* were also chosen due to their relative abundance in 16S rRNA gene clone libraries and previous studies demonstrating differences in species composition across samples (Kuske et al. 2002; Stach et al. 2003). The phylum *Gemmatimonadetes* was chosen because of the lack of knowledge of this newly-described phylogenetic group (Zhang et al. 2003). For these calculations, sequence libraries were pooled based upon land use type, and not as individual libraries as was done previously with total sequencelibraries. Comparisons of *Acidobacteria* and *Actinobacteria* differed in the majority of land use comparisons analyzed. This result indicates that these phyla vastly differ in soils of varying land use type. The *Firmicutes* and *Gemmatimonadetes* phyla are not significantly different in six of the pairwise LIBSHUFF comparisons. The

---

**Figure 2.7 - LIBSHUFF comparison of replicate X13A and X13B clones.** Homologous coverage curves are shown with black circles and heterologous coverage curves are shown with orange circles. Solid lines indicate the difference between the homologous and heterologous coverage curves at each value of D, and broken lines indicate the 95% value of the random shufflings.
Table 2.7 - LIBSHUFF comparisons of specific phyla contained in CAP LTER soil clone libraries

<table>
<thead>
<tr>
<th>Comparison no.</th>
<th>Library 1 (X) data</th>
<th>Library 2 (Y) data</th>
<th>Acidobacteria</th>
<th>Actinobacteria</th>
<th>Firmicutes</th>
<th>Gemmatimonadetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 A</td>
<td>Agricultural</td>
<td>Open</td>
<td>213</td>
<td>87</td>
<td>0.087</td>
<td>261</td>
</tr>
<tr>
<td>B</td>
<td>Open</td>
<td>Agricultural</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 A</td>
<td>Agricultural</td>
<td>Remnant</td>
<td>213</td>
<td>87</td>
<td>0.001</td>
<td>261</td>
</tr>
<tr>
<td>B</td>
<td>Remnant</td>
<td>Agricultural</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 A</td>
<td>Agricultural</td>
<td>Urban</td>
<td>213</td>
<td>87</td>
<td>0.004</td>
<td>261</td>
</tr>
<tr>
<td>B</td>
<td>Urban</td>
<td>Agricultural</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 A</td>
<td>Open</td>
<td>Remnant</td>
<td>344</td>
<td>304</td>
<td>0.001</td>
<td>93</td>
</tr>
<tr>
<td>B</td>
<td>Remnant</td>
<td>Open</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 A</td>
<td>Open</td>
<td>Urban</td>
<td>347</td>
<td>304</td>
<td>0.002</td>
<td>93</td>
</tr>
<tr>
<td>B</td>
<td>Urban</td>
<td>Open</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 A</td>
<td>Remnant</td>
<td>Urban</td>
<td>299</td>
<td>85</td>
<td>0.076</td>
<td>64</td>
</tr>
<tr>
<td>B</td>
<td>Urban</td>
<td>Remnant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P*-values in red indicate a non-rejection of the null hypothesis that the two libraries are equal using the Bonferroni correction, indicating that the two libraries do not differ significantly.
agricultural-urban desert sequence libraries appeared to diverge at the highest degree, as these libraries only produced similarities in the heterologous coverage of the *Gemmatimonadetes* (Table 2.7, Comp. 3A). However, desert remnant 16S rRNA gene sequence libraries were found to be similar to urban sequences across all four phyla tested (Table 2.7, Comp. 6A).

**Cluster Analysis of CAP LTER Soils**

Cluster analyses were performed based on differences between \( \Delta C \) values generated through LIBSHUFF analysis and the degree of shared OTUs across CAP LTER samples based upon Sørenson’s similarity index values (Fig. 2.8). Within the dendogram in Figure 2.8A, two primary groups are present. Group 1 includes three open desert samples (K9, Ill, J7), two desert remnant soils (X15, U21), and the single mixed soil sampled (L14). Group 2 is distinctly divided into two more homogeneous groups. Five of the six soils found in cluster A are open desert samples, while cluster B contains all three agricultural soils, five of six urban soils, and two desert remnant soils. The replicate samples (X13A and B) form the strongest cluster.

The Sørenson’s index dendogram in Figure 2.8B is derived from shared ARDRA fingerprints as opposed to phylogenetic relatedness. This dendogram contains more heterogeneous groupings with respect to land use. The two initial groupings both contain soils belonging to all four major land use categories. Clusters with higher similarity values (>50%) are highly variable as well. Overall Sørenson’s index values were extremely low, ranging from 17% to 60% within clusters.

**Discussion**

**Inter-Phylum Abundance**

In this study, there were multiple instances in which significant differences in
Figure 2.8 - Results of a complete-linkage cluster analysis. Clusters based on (A) ΔC values based on the difference in homologous (X) and heterologous (XY) coverage generated from the LIBSHUFF analysis and (B) similarity in bacterial species composition based on Sorenson’s index. Letters in parentheses indicate the land-use categories (O open; U urban; A; agricultural; R remnant; M mixed) Numerical labels 1 and 2 indicate the first two clusters formed from the method and a, b indicate clusters of particularly high amount of genetic coverage.

overall abundances of broad phylogenetic groups were observed across land use types.

The most highly variable phylum in terms of abundance appears to be the low G+C Gram phylum *Firmicutes*, members of which are found most frequently in agricultural soils.

This result may be correlated with the decreased presence of *Actinobacteria* and the subsequent higher concentrations of antimicrobial agents produced by this group, which are most effective in lysing the thick peptidoglycan membranes of Gram positive organisms (Moré et al. 1994). The high abundance of alpha proteobacteria has been observed in other bacterial community diversity studies (Nüsslein and Tjede 1999;
Sessitch et al. 2001; Buckley and Schmidt 2003). This high abundance is most likely due to the tendency of this group to employ an r-strategy as a reproductive approach, permitting these opportunist taxa to grow quickly within an environment during the rare instances when limiting resources become readily available and then maintain dominance during long periods of substrate starvation (MacArthur and Wilson 1967; Gadgil and Solbrig 1972). The physiological diversity within the alpha proteobacteria, including both the free-living and legume symbionts with N-fixing capabilities, may also permit this group to sustain a presence in a variety of habitats. However, the differentiation in alpha proteobacteria abundance between open desert and all other land use types is a surprising one, suggesting that there may be anthropogenic factors that inhibit the accelerated reproductive stages that many of the proteobacteria exploit. For instance, open desert samples are composed mainly of larger sand particles, while urbanized samples have increased concentrations of smaller silt and clay particles (Green and Oleksyszyn 2003). The alpha proteobacteria appear to thrive on sand particles where they are protected from predatory interactions, while other phyla dominate soils with smaller particles (Sesitsch et al. 2001). Surprisingly, the variability of the abundances of Proteobacteria are explained reasonably well by the abundance of Acidobacteria in an inverse relationship (r²=0.61, two-tailed t test). The reason for this observed relationship is unclear, but as previously suggested it may represent competition-based distribution and potential metabolic similarity of members of these two taxa, considering no other significant relationship was found between other phyla (Quaiser et al. 2003). It could also be indicative of the resource heterogeneity of the soils, where the Proteobacteria/Acidobacteria ratio increases with improved nutrient status (Smit et al.
A more stable nutrient supply in these soils would further encourage the growth of bacterial populations with rapid reproductive strategies.

The relatively low overall abundance of *Acidobacteria* species were also unexpected because much higher abundances have been previously reported in southwestern USA soils (Dunbar et al. 2002). However, increased presence of *Acidobacteria* and beta proteobacteria observed in agricultural and urban soils may be indicative of their importance within the rhizosphere, the area surrounding the absorbing root-soil interface. The increased presence of rhizosphere habitats may best explain the increased presence of these groups due to increasing plant diversity observed in these soils via non-native plant introduction (Hope et al. 2003). The presence of a diverse assemblage of legume roots in urbanized soils requires an increased number of nitrogen-fixing symbionts and free-living species often filled by alpha and/or beta proteobacterial rhizobia (Moulin et al. 2001; Chen et al. 2004). It is also been considered that some members of *Acidobacteria* have specific functional roles with specific plant roots in arid grasslands (Kuske et al. 2002). The result of a diverse root assemblage coupled with the sensitivity of plant-bacterial symbiosis suggests that the increased variety of roots within these soils leads to increased abundance of these groups. The significant presence of the *Gemmatimonadetes* phylum, a newly described and relatively unknown taxonomic group initially named the “BD group” (Hugenholtz et al. 2001), suggests that members of this group are physiologically robust and may fill a valuable, yet undetermined ecological role in these soils irrespective of land use type. Significant abundances of these taxa are supported by the work of Mummey and Stahl, who reported a *Gemmatimonadetes*
incidence rate of 8% in Wyoming grassland soils (2003), which is comparable to the results generated in this study.

The single mixed soil that was tested showed dramatic changes in species composition, most notably the levels of *Cyanobacteria*. The reasons for this shift remain unclear, but suggest that geographic locations bordering two land use types contain more dynamic microhabitats that allow for the survival of vastly different bacterial taxa. These photosynthetic bacteria are notably dominant in nearby mature Chihuahuan desert crusts, but are less prominent in recently disturbed soils (Yeager et al. 2004). This observation would suggest that the lack of *Cyanobacteria* in urbanized sites is a function of disturbance effects, but similar low abundances in open desert samples with presumably intact crusts is confounding. One alternative explanation for the low abundance of *Cyanobacteria* is the mixed soil sampling protocol, where both the thin crust and the upper subsurface layers were collected. The low numbers of *Cyanobacteria* reported here are more representative of those subsurface layers that light cannot penetrate in which phyotosynthesis cannot occur.

In previous studies, suggested changes in the phylum-level diversity may be dependent on a mutual colonization history of local bacterial populations, and therefore may not be relied upon as strong evidence for demonstrating shifts in bacterial diversity due to an absence of an initial dispersal event (Felske et al. 1998; Findlay et al. 1999; Dunbar et al. 2001). However, due to the geographic compactness of many of the sampled CAP LTER sites it is reasonable to disregard dispersal as a factor in changes in abundance. The Morisita-Horn similarity values were used in pairwise comparisons of CAP LTER samples to test the effects of distance on sample similarity. The low amount
of explained variability coupled with the positive slope provide further evidence of the heterogeneity of soil microhabitats (Grundmann and Debouzie 2000; Oda et al. 2003). The weak relationship between distance and shared OTUs among CAP LTER soils also suggests that the observed changes in overall diversity are not explained by dispersal and that other factors must be responsible for these shifts. There is also an issue regarding the effects of temporal shifts in bacterial diversity in soils, a phenomenon that has been shown in other soil systems (Torsvik et al. 1996; Smit et al. 2001; Ramirez et al. 2001; Girvan et al. 2003). In this case there are no data on short-term temporal changes in bacterial diversity, but instead there is an underlying assumption that the desert remnant and urban soils were all undisturbed open desert samples before the urbanization event occurred, which signifies a measure of long-term disturbance on the CAP LTER soils. Therefore, it can be considered that the open desert samples can act as a control against other land use types for measuring the long-term temporal changes in diversity.

**Diversity Estimates and Statistical Comparisons**

The culmination of the wide variety of methodologies used in this study allows for comparison of coverage, richness, and phylogenetic diversity of multiple bacterial 16S rRNA gene clone libraries. As previously mentioned here and in studies of a similar nature, many of these analyses are dependent on an OTU definition for ecological calculations (Zhou et al. 2002; Stach et al. 2003; Humayoun et. al. 2003; Bornemann et al. 2003). Although the methods in this study allow for a less conservative OTU definition, it is imperative to recognize the potential caveats of using OTUs in diversity statistics. For example, it has been shown that in the case of the Actinobacteria, an OTU definition as high as 99% may not be enough to distinguish two species from two
different environments (Stach et al. 2003). Furthermore, OTU designation based on 16S rRNA gene sequence relatedness does not necessarily discriminate between distinct ecological entities (Vandamme et al. 1996). It has been shown previously that environmental bacterial isolates with very high 16S rRNA gene sequence similarity values (>99%) function differently in an ecological sense (Boddinghaus et al. 1990; Ash et al. 1991; Williams et al. 1991). The observed variable evolutionary rates across bacterial taxa are compounded by the fact that there is no strictly-enforced delineation pertaining to the acceptable definition of an OTU. Thus, many culture-independent studies define taxonomic entities differently, making cross-study comparisons difficult. However, in this more comprehensive study with large OTU datasets available for the analyses, we define an OTU based upon the accepted 97% benchmark, providing for a higher probability that these OTU groups are generated from the same bacterial species, and therefore are confident that relative richness estimates are still applicable in comparisons of CAP LTER soils.

Likewise, it is imperative to note that a relatively large amount of soil (2g) was used for bulk DNA extraction, presumably homogenizing vast numbers of microhabitats that are spatially isolated. However, the nearly identical phylogenetic and group abundance results observed in the two separate X13 16S rRNA gene clone libraries suggests that the construction of larger soil clone libraries provides improved resolution of the overall diversity bacterial assemblages in soil. In the future, further 16S rRNA clone library replication of soil samples for ecological research will allow for the utilization of dynamic statistical analyses to more accurately investigate observed shifts in bacterial diversity across environmental sites.
Although this is one of the most exhaustive studies of its kind, accumulation curves and nonparametric richness estimators suggest that these soil bacterial communities are still undersampled, and that predictions of total species richness based upon the generated CAP LTER clone databases will be imprecise. This lack of complete sampling highlights the need for advances in technology and high-throughput methods in order to efficiently sample entire prokaryotic soil communities in environmental samples. However, the sample size was large enough for accumulation curves to become non-linear, permitting more robust analysis of the attained OTUs via statistical analysis once only permissible in macrobial organism datasets.

The implementation of the LIBSHUFF method is essential for quantifying the overall evolutionary distance of environmental bacterial sequences within a given clone library. The ability to observe significant difference in two samples is dependent on the size of data sets involved. Therefore, it is not surprising that the analysis of the 16S rRNA clone libraries of 330 to 522 clones yielded significantly different results in nearly all of the pairwise comparisons performed, providing further evidence for the vast extent of 16S rRNA gene diversity in soil habitats. This observed heterogeneity is due mostly to cases where phylotypes from rare phyla that are present in only a few of the soils are detected. This causes a great amount of higher-order divide across the homologous and heterologous coverage curves and increases the magnitude of dissimilarity. However, when using specific phylogenetic group data, more detailed information can be obtained from the analysis. The Acidobacteria sequences appear to be phylogenetically distinct across changing soil regimes, indicating the diverse physiological capabilities of members of this relatively unknown phylum. Most sequence libraries comparisons of
*Firmicutes* and *Gemmatimonadetes* in the majority of land use comparisons do not significantly differ, revealing that closely related phylogenetic groups (genus-level) are present, presumably filling similar ecological roles within these environments, although in the case of low G+C Gram positive bacteria this may be a consequence of the decreased recombination rates resulting in a lesser degree of 16S rRNA gene sequence divergence (Palys et. al 1997). In all four phyla studied, nearly all desert remnant sequences were also found in their urban counterparts. In contrast, these remnant sequences are not related to those found in open desert soils in three of the four phyla. Coupled with the notion that these remnant soils were comparable to natural open desert bacterial populations before the rapid population growth of the area, the LIBSHUFF data provides evidence of urbanization of not only disturbed community structure within the urban sector, but also within nearby natural settings. This conclusion appears to be more suggestive of a significant urban influence on remnants than the multivariate abundance data, which is a further indication of the greater precision obtained via phylogenetic-based results in assessing the ecological similarity of communities compared to other ecological diversity indices which are based solely upon frequency of phylotypes or fingerprints (Martin 2002).

The two dendograms based upon cluster analysis of genetic and OTU data demonstrate a high degree of topological incongruence (Fig. 2.8A, B). While the dendogram based upon genetic coverage generally indicates land use gradients amongst CAP LTER soils, the Sørenson’s index dendogram varies among clusters, suggesting that taxa are not shared at a higher extent within land use types. The discrepancy in these two analyses is likely due to the omission of species richness in the calculation of
Sørenson’s index values. The bias in richness quantification is more prevalent in comparing microbial communities in soil environments where complete sampling cannot be achieved and many rare species within a sample are not captured. For this particular study, it is assumed that large abundances of shared sequences, represented by the ΔC values, are more reliable for assessing trends in bacterial community diversity among environmental sites as opposed to shared OTUs represented by Sørenson’s index values.

While a suitable model reasonably explaining bacterial diversity estimates using other ecological variables was not found, there is substantial evidence of habitat transition within anthropogenically disturbed samples as indicated by the significant changes in imperviousness, total carbon, vegetation cover, nitrate concentration, and clay content across the different land use types (Fig. 2.5). These variables do not explain the observed discrepancy in diversity estimates because the amount of diversity within land use types can be extensive, so the ability to find more concrete patterns of observed changes can be diminished. Therefore, even larger 16S rRNA gene clone library datasets are needed for increased power in multivariate analysis, so that more rigid variables such as those mentioned previously can be re-examined.

**Mechanisms Responsible for Diversity**

The mechanisms for which a dominant, competitive diversity profile in soils with low carbon have been previously proposed as: (i) superabundant resources, (ii) resource heterogeneity, (iii) spatial isolation, and (iv) nonequilibrium conditions (Zhou et al. 2002). Spatial isolation appears to be the most crucial mechanism in surface soils. Isolation of soil particles caused by the absence of water regimes result in a lack of resource migration and therefore enhance competitive interactions among microbes,
resulting in dominance by the most aggressive species. (Treves et al. 2003; Tjiede et al. 2003). A uniform diversity profile would be discovered with the absence of one of these four proposed mechanisms, where the lack of competition allows for the survival of nearly equal amounts of genotypes in the soil. This model appears does not fit the low diversity values observed when comparing land use effects on these CAP LTER soils, which are low in organic carbon concentration (0.78-5.09 g·mg/L). The lack of watering events at the surface of the open and remnant desert sites results in spatially isolated soil particles representing unique microhabitats that are almost completely void of resource or species immigration. In the rare events when moisture in the form of precipitation appears, the fast growing r-strategist bacteria, presumably the alpha proteobacteria, likely out compete the other species in utilizing the newly-introduced nutrients and increase in numbers. In desert soils, these rare rain events may also be considered as extreme disturbance events, which would lead to the competitive exclusion of slow-growing bacterial by their r-strategist counterparts under the intermediate disturbance hypothesis (Connel 1978). Under the spatial isolation model, the abrupt rise in the abundance of fast growers would be observed only shortly after water activity and would quickly decrease through time in correlation with nutrient availability, resulting in similar proportions of species. However, Dunbar et al. (2001) demonstrated that in the four Arizona soils they sampled, a uniform distribution did not fit the 16S rRNA clone dataset, while a log-normal based model appears to apply. In the desert soils studied here, only six ARDRA fingerprints were observed on the average of ten times are more, but the average abundance of these same fingerprints was disproportionately high amongst all 16S rRNA gene clones sampled (20.9). However, the majority of ARDRA fingerprints are
singletons (100.2 per sample on average). This survey therefore also suggests a 
lognormal-like distribution for the open desert CAP LTER soils, and that the spatial 
isolation theory may not be applicable to soils within this arid environment.

The apparent competitive pattern observed in urbanized and agricultural soils is 
most likely the result of artificial watering and subsequent resource heterogeneity in these 
systems. Watering type is a strong predictor of bacterial diversity. This added moisture 
in urban and agricultural soils would eliminate the spatial isolation component of the 
diversity model by allowing for a means for frequent species and substrate immigration 
to occur. Although the soils in this study are not significantly different in carbon 
abundance across land use types (data not shown), there is likely to be more 
heterogeneity of the carbon substrates types in the agricultural and urbanized soils, as 
indicate by the increase in functional diversity based upon BIOLOG data (Enticknap and 
Rainey 2001). Furthermore, it reasonable to assume that the amount of disturbance 
(flooding, traffic, vegetation, fertilization, etc.) is much greater in urbanized and 
agricultural soils and the higher diversity statistics indicate the ability of a bacterial 
community to recover from disturbance and utilize resources effectively (Chapin et al. 
1997). Agricultural soils appear to contain the most phylogenetically diverse bacteria of 
all land use types. The key component shared by both agricultural and urban sites is the 
introduction of water and nutrients via irrigation and/or fertilization. Ultimately, the 
degree of increasing diversity among urbanized and agricultural land use types suggests a 
lesser degree of limiting nutrient availability for bacterial species in soil, thereby relaxing 
strong competitive interactions that are present in the open desert. The same increases in 
diversity have been observed in other organisms within the same CAP LTER survey sites
(Hope et al. 2002, Cousins et al. 2003). Furthermore, increasing diversity in urban settings appears to contradict the theory that the influence of humans across the world is leading to rapid loss in overall biodiversity (Wilson 2004). Based on the results generated in this study, perhaps this mass extinction event does not apply to the prokaryotic world.

The data presented here builds upon the understanding the prokaryotic ecology of various environments and reveals the necessity of these studies to search for key factors that may control diversity. It demonstrates that the implementation of high-throughput methods in analyzing 16S rRNA gene data can yield a foundation to answering topics in prokaryotic ecology by using a variety of ecological variables. Ultimately, the anthropogenic influence on microbial communities appears to be related to the land use management via irrigation, dramatically altering soil water content. Findings of subtle shifts in bacterial diversity due to urbanization, as well as the identification of key species or phylogenetic groups controlling the biogeochemical processes, are important to assess the long-term “soil health,” the capacity of soil to function as a living system in order to sustain biological productivity and environmental quality (Doran and Zeiss 2002; McDonnel and Pickett 1993). Microbial processes have long been known to influence agricultural output (Killiam 1994; O’Donnell et al. 1994) yet how these changes in bacterial populations are affecting the urban soil processes is presently unclear. However, observed abrupt ecological shifts are predicted to be the result of biogeochemical changes brought about by changes in microbial species composition (Lockeretz et al. 1981; Drinkwater et al. 1998; Knops and Tilman 2000; Paul et al. 1999; Anderson 2003). If such events also occur in the CAP LTER soils, the health of the soil
within a region could be adversely affected. Further studies of this nature will allow for
the beginning of a means to address and ultimately amend or maintain current
environmental management of urbanized areas if needed.
CHAPTER 3:

BACTERIAL BIOGEOGRAPHY AT THE CAP LTER
Introduction

Biogeography is defined as the field of biology that attempts to document and explain spatial and/or temporal patterns in biodiversity. The processes involved in producing biogeographical patterns are evolution, extinction, and dispersal. Gloger’s rule (1833) and Bergmann’s rule (1847), both of which dealt with the distribution of species as a function of body appearance, were among the first attempts to describe spatial changes in biodiversity. Since then, numerous biogeographical studies involving eukaryotic organisms have explained the manner in which species are distributed in nature (see Brown and Gibson 1983 for review).

Due to the vast heterogeneity within soil microhabitats, knowledge of the biogeography of bacteria in soil is relatively limited. In the beginning of microbial ecology theory, findings based on culture-dependent results supported the Baas-Becking “(1934) everything is everywhere; the environment selects” theory for explaining bacterial biogeography. Under this dogma, any observed spatial shift in microbial diversity must be shaped in large part by ecological parameters because of an assumed event of global dispersal of microorganisms. The theory also indicates that if there has been a global dispersal event involving soil bacterial species, their immense range of adaptability may allow for an immense consortium of bacterial species residing in heterogeneous environments. This degree of adaptability comes from accelerated rates of evolution via mutation and generation time (Arbor 2000; Ochman et al. 2000; Spiers et al. 2000), large effective populations, inter-population genetic exchange (Shapiro 1985), and subsequent periodic selection (Levin 1981). Ultimately, these unique properties belonging to microbial species can limit the probability of local extinctions, erasing
geographic barriers and leading to pandemic presence of bacterial species. When differences in the spatial distribution of species across environmental gradients are observed, it is assumed that there are disruptive components within these habitats of which ecological and evolutionary forces have resulted in the absence of those taxa which are physiologically and/or genetically inflexible.

The “everything is everywhere theory” had been unchallenged until the rise of molecular-based biogeographical studies in which actual prokaryotic genotypes can be compared across geographic locations. It has been discovered that within strains of the thermophilic arcahaeon Sulfolobus, there is a significant correlation between genetic divergence of multiple protein-coding loci and geographic distance, even in strains that demonstrate high relatedness in terms of the 16S rRNA gene (Whitaker et al. 2003). This study suggests that gene flow in these populations is limited and that genetic drift has created vastly different genetic composition within local populations. The work of Papke et al. (2003) also described geographic distinction between thermophilic populations of members of Cyanobacteria, in which geographical patterns do arise across genotypes in different continents, and that these genotypes do not always correlate with microhabitat structure or properties (temperature, pH).

Increased bacterial community structure analyzes through culture-independent methods have increased understanding of biogeographical relationships and lead to better comprehension of terrestrial ecosystem processes and the ecological roles that specific taxa fulfill (Staley and Gosnik 1999). Due to the diverse collection of landscapes in the CAP LTER created by anthropogenic disturbances via agriculture and urbanization and the observed changes in phylum-level diversity within the entire region, we hypothesize
that the distribution of lower-level bacterial groups would be spatially restricted based upon their ability to survive within a limited number of land use types. Using phylotypic data collected from construction of 16S rRNA gene libraries, we generated phylogenetic trees for all major bacterial taxa found in the CAP LTER to catalog the phylogenetic diversity of bacterial phylotypes, and recorded the presence and abundance of significant phylotypic groups across land use types. Overall, we found evidence for biogeographic limitation of phylotypes across the site, indicating that specific bacterial populations are endemic to not only specific land use types but the entire CAP LTER as well.

Materials and Methods

Phylogenetic Tree Construction

Phylogenetic analyses based 16S rRNA sequence data were carried out using the ARB-software package (Strunk et al. 2003) along with a modified sequence base containing approximately 5,000 bacterial 16S rRNA sequences available through the Genbank interface (http://www.ncbi.nlm.nih.gov). The 16S rRNA sequence lengths used in phylogenetic tree reconstructions were in the range of 490-600 nucleotide positions. The 16S rRNA alignments of selected clones representing unique ARDRA fingerprints were made using the ARB edit tool. The construction of phylogenetic trees was based on the neighbor joining analysis with the Felsenstein correction. Bootstrapping analyses were performed to test the robustness of the trees using PHYLIP 6.0 in a distance and neighbor joining analyzes with 100 replicates (Felsenstein 1988).

Results

Biogeographic Analysis and Distribution of CAP LTER Clones

Phylogenetic trees representing ten higher-order bacterial taxa were constructed. Robust
phylotypic clusters with three or more phylotypes, or single phylotypes that group within validly described taxonomic groups in accordance with Bergey’s taxonomic outline (http://rdp.cme.msu.edu/index.jsp) were designated as significant and categorized by their distribution in the 16S rRNA clone libraries. From the phylogenetic analyses, eighty-five significant 16S rRNA groups were found (Tables 3.1-3.4). Since many clusters do not contain sequences from cultured organisms, they were designated arbitrarily either with the names of corresponding 16S rRNA clones from previous studies or the higher-order phylogenetic group they represent.

**Alpha Proteobacteria**

Alpha proteobacteria was the most observed sampled group, consisting of fourteen major phylotypic groups (Table 3.1). The majority of phylotypes (3,466) are affiliated with the *Sphingomonadaceae* family (Fig. 3.1). *Sphingomonas* strains are usually identified as aerobic, heterotrophic rod-shaped bacteria and assumed to be ubiquitous in nature. *Sphingomonas* isolates have been collected from a diverse collection of anthropogenically-disturbed environments (Mueller et al. 1990; Mueller et al. 1990; Feng et al. 1997; Lloyd-Jones and Lau 1997; Atkins 1999; Meyer et al. 1999; Momma et al. 1999; Bastiaens et al. 2000; Cassidy et al. 1999; Pinyakong et al. 2000; Sørenson et al. 2001; Leys et al. 2004). One selective advantage associated with *Sphingomonas* species that may explain their significant presence in these habitats is likely a result of the ability to degrade a variety of artificial pollutants, including chlorinated phenols (Cassidy et al. 1999; Crawford and Ederer 1999), polycyclic aromatic hydrocarbons (PAHs), insecticides (Nagata et al. 1998, Kim et al. 2004), and herbicides (Adkins 1999; Kohler 1999).
Table 3.1 - 16S rRNA phylotype distribution of alpha and beta proteobacteria within CAP LTER soils

<table>
<thead>
<tr>
<th>Taxonomic group (α)</th>
<th>Phylotypic or phylogenetic group</th>
<th>φ</th>
<th>No. of phylotypes (no. of clones) in the following land use type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Agricultural</td>
</tr>
<tr>
<td>Alpha Proteobacteria (26.7)</td>
<td>Sphingomonadaceae</td>
<td>7.1</td>
<td>35 (292)</td>
</tr>
<tr>
<td>Alpha-1</td>
<td></td>
<td>13.2</td>
<td>5 (9)</td>
</tr>
<tr>
<td>Alpha-2</td>
<td></td>
<td>10.1</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Methylbacteriaceae</td>
<td></td>
<td>12.5</td>
<td>6 (16)</td>
</tr>
<tr>
<td>Caulobacteraceae</td>
<td></td>
<td>12.6</td>
<td>5 (14)</td>
</tr>
<tr>
<td>Alpha-3</td>
<td></td>
<td>5.6</td>
<td>3 (24)</td>
</tr>
<tr>
<td>Stappia-Rosebium</td>
<td></td>
<td>12.2</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Bradyrhizobiaceae</td>
<td></td>
<td>*</td>
<td>1 (19)</td>
</tr>
<tr>
<td>Rhizobiaceae</td>
<td></td>
<td>15.4</td>
<td>3 (7)</td>
</tr>
<tr>
<td>Bartonellaceae</td>
<td></td>
<td>*</td>
<td>1 (13)</td>
</tr>
<tr>
<td>Hyphomicrobiaceae</td>
<td></td>
<td>6.2</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Alpha-4</td>
<td></td>
<td>9.4</td>
<td>3 (24)</td>
</tr>
<tr>
<td>Rhodospirillaceae</td>
<td></td>
<td>12.1</td>
<td>5 (14)</td>
</tr>
<tr>
<td>MND8</td>
<td></td>
<td>13.2</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Alpha-5</td>
<td></td>
<td>10.1</td>
<td>5 (18)</td>
</tr>
<tr>
<td>Rhodobacteraceae</td>
<td></td>
<td>8.5</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Acetobacteraceae</td>
<td></td>
<td>17.6</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Beta Proteobacteria (22.7)</td>
<td>Oxalobacteraceae</td>
<td>16.1</td>
<td>12 (35)</td>
</tr>
<tr>
<td>Comamonadaceae</td>
<td></td>
<td>19.1</td>
<td>8 (40)</td>
</tr>
<tr>
<td>Neisseriaceae</td>
<td></td>
<td>6.0</td>
<td>2 (2)</td>
</tr>
<tr>
<td>646-2</td>
<td></td>
<td>18.1</td>
<td>5 (19)</td>
</tr>
</tbody>
</table>

a Percent 16S rRNA dissimilarity throughout phylotypes in the group
b Phylotypes and phylogenetic groups correspond to cluster of similar 16S rRNA gene sequences in the phylogenetic trees in Fig. 3-1 to 3-5.

Forty-two phylotypes (2,656 total clones) grouped within the unknown putative species II group previously described (Ley et al. 2004). The two isolate sequences found within this group are *Sphingomonas* sp. SIA181, which was cultured in low nutrient from...
Figure 3.1 - 16S rRNA gene tree showing positions of mostly heterotrophic members of the *Sphingomonadaceae* family found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

accretion ice at Lake Vostok, Antarctica (Christner et al. 2001), and a German agricultural soil clone SC-I-5 (Lukow 1999). Other phylotypes (clone 376-2, clone 768-2, clone 565-2) found in a single Californian agricultural soil (Valinsky et al. 2002) are also found in this group.
A smaller collection of phylotypes resemble the *Sphingomonas* subgroup I. The described species within this group include *Sphingomonas mali* and *Sphingomonas asaccharolytica*, both of which were isolated from plant roots in Japan (Takeuchi et al. 1995), as well as *Sphingomonas aquatilis*, isolated from a Korean mineral water source (Lee et al. 2001). Ten phylotypes are grouped near known polycyclic aromatic hydrocarbon (PAH)-degrading species *Sphingomonas capsulata* and *Sphingomonas subarctica*, both reclassified to be members of the genus *Novosphingubium* (Yabuuchi et al. 2002). Also present in this group are *Sphingomonas* sp. B28161 isolated from a paper-processing machine (Vaisanen et al. 1998) and two freshwater lake phylotypes (Pernthaler et al. 1999; Mitsutani et al. 2000). *S. capsulata* was isolated from a clinical sample (Yabuuchi et al. 1990) and *S. subartica* was isolated from an Antarctic lake (Nohynek et al. 1996). There also appears to be no significant changes in phylotype richness or distribution across land use types (Table 3.1).

Another cluster of five phylotypes appear to be most similar (95.0-96.7% 16S rRNA gene similarity) to type species *Porphyrobacter tepidarius*, a thermophilic aerobic bacterium isolated from a brackish hot spring in Japan (Hanada et al. 1997). Chemoheterotrophy appears to be a primary metabolism of *Porphyrobacter* sp., however the cells also contain bacteriophyll *a* and can perform photosynthesis. There were also two phylotypic groups (alpha-1 and alpha-2) which had no less than 93% similarity to any members of *Sphingomonadaceae*, likely indicating new genera. These two clusters were detected mostly in urban and desert remnant land use type, indicating a restricted habitat range for these taxa (Table 3.1).
CAP LTER phylotypes also grouped within other families of the alpha proteobacteria (Figs. 3.2-3.4). Thirty-seven phylotypes grouped within the *Methylobacteraceae, Caulobacteraceae, and Bradyrhizobiaceae* families (Fig. 3.2). Seventeen phylotypes were closely related to the genus *Methylobacterium*, composed of facultative methylotrophs that typically grows on highly reduced carbon sources (Green 1992). Members of this group are believed to be ubiquitous in nature, and are often observed within artificial environments, most notably chlorine-contaminated water sources (Garf and Bauer 1973; Hiraishi and Kaneko 1994; Hirainshi et al. 1995; Castro et al. 2004). *Methylobacterium* strains have also been shown to degrade PAHs, compounds known to contaminate natural environments (Andereoni 2004). Most of the phylotypes in this group demonstrate 5-7% divergence from *Methylobacterium* species, indicating that this group represents a new genus. Thirteen phylotypes group within the *Caulobacteraceae*, more specifically near the genera *Phenylobacterium* and *Brevundimonas* which have also been shown to degrade aromatic compounds (Lingens 1989). The rest of the phylotypes within this family group near *Brevundimonas*, composed of species that have been shown to different ecophysiological behavior across different environments, indicating an ability to quickly exploit selective advantages to become superior competitors (Jasper and Overmann 2004). The members of the alpha-3 group are phylogenetically similar (5.8% dissimilarity across 12 phylotypes), but show over 8% divergence from all other 16S rRNA sequences in the database, indicating that the phylotypes represent a potentially novel genus. There are also three phylotypes grouped within the *Stappia-Roseibium* group, and one phylotype located within the
Figure 3.2 - 16S rRNA gene tree showing positions of the *Methylobacteraceae*, *Caulobacteraceae*, Alpha-3, *Stappia-Rosebium*, and *Bradyrhizobiaceae* groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%. 
Figure 3.3 - 16S rRNA gene tree showing positions of the Rhizobiaceae, Bartonellaceae, Hyphomicrobiaceae, Alpha-4, Rhodospirillaceae, and MND-8 groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

family Bradyrhizobiaceae, a group associated with symbiotic activities with legumes.

Members of these groups are represented in nearly land use types (Table 3.1).

Aside from Bradyrhizobaceae, CAP LTER phylotypes grouped within the four other families within the order Rhizobiales (Fig. 3.3). Phylotypes grouped within the nitrogen-fixing members of the Rhizobaceae and Rhodospirillaceae that form nodules on
the roots of a diverse assemblage of plants (Zahran 2001) as well as in the *Bartonellaceae* and *Hyphomicrobaceae*. Two other distinct cluster of phylotypes branch near these groups as well (alpha-2 and MND-8 groups). The MND-8 group is also composed of two phylotypes collected from a Wisconsin freshwater lake (Stein et al. 2001) and from an industrial water treatment plant (Juretschko et al. 2002). All of these groups appear to be more prevalent in open desert and urban samples compared to remnant and agricultural soils (Table 3.1)

A diverse assemblage of CAP LTER phylotypes were also associated with the families *Rhodobacteraceae* and *Acetobacteraceae*, more specifically the genera

*Paracoccus*, *Amaricoccus*, and *Paracraurococcus* (Fig. 3.4). Members of these genera
are consistently found in 16S rRNA gene clone libraries constructed from contaminated environments around the world (Levantesi et al. 2002; Smith et al. 2003; Deb et al. 2004). Members of *Paracoccus* in particular have been found to degrade PAHs in soil (Zhange et al. 2004). The distribution of *Rhodobacteraceae* is skewed towards urban soils, while the *Acetobacteraceae* are more dominant in open soils (Table 3.1).

**Beta Proteobacteria**

CAP LTER phylotypes grouped with three families and one phylotypic group within the beta proteobacteria (Fig. 3.5). No obvious land use zoning in phylotypic or phylogenetic group abundance was observed (Table 3.1). Twenty-four phylotypes were associated with the *Oxalobacteraceae* family. More specifically, the phylotypes formed three clusters that were most similar to the genera *Massilia*, *Zooglea*, and *Herbaspirillum*. The only validly described species of the genus *Massilia* is *M. timonae*, a non-fermentative aerobe isolated from human blood samples (La Scole et al. 1998; Lindquist et al. 2003). However, other phylotypes corresponding to this species have been found in a variety of soil habitats (Wery et al. 2003; Padmanabhan et al. 2003). *Massilia*-like phylotypes have also been associated with the degradation of PAHs in soil (Bodour et al. 2003). Bacteria associated with *Zooglea* have also shown the ability to degrade PAHs and other pollutants (Williams et al. 1983; Wang et al. 2002) and play a key role in wastewater treatment (Rossello-Mora 1995). Members of the *Herbaspirillum* genus are mostly nitrogen-fixing diazotrophs and are commonly found in cropland (Chelius et al. 2001) and forest (Widmer et al. 1999) soils.

Ten phylotypes were associated with the *Comamonadaceae* are not closely related to any described species (<5% 16S rRNA gene similarity). Within this family, a group of
Figure 3.5 - 16S rRNA gene tree showing positions of the Oxalobacteraceae, Commonadaceae, Neisseriaceae, and 646-2 groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

Six phylotypes clustered near a marine clone collected from the coast of Japan (Mitsutani et al. 2000). Another phylotype grouped near Polaromonas, a heterotrophic bacterium isolated from Antarctic sea ice (Irgens et al. 1995). The final three phylotypes were over 8% dissimilar to all other 16S rRNA gene sequences, indicating a novel genus within the Comamonadaceae.

Two other phylotypic groups were observed in the phylogenetic analyzes, the
646-2 and beta-2 groups. The 646-2 group included four CAP LTER phylotypes as well as three additional phylotypes collected from Californian agricultural soils (Valinsky et al. 2000) and activated sludge (Juretschko et al. 2002). The 16S rRNA phylotypes of other group (beta-2) are more than 10% dissimilar to the other members of the database, providing no clues as to the characteristics of these phylotypes.

**Delta Proteobacteria**

The majority of CAP LTER phylotypes associated with this group are myxobacteria, a group that includes the Polyangiaceae and Myxococcaeae families (Fig. 3.6). The myxobacteria are described as a mostly aerobic group that has the ability to produce fruiting bodies under starvation conditions (Green 1992). Seven of these phylotypes were only found in urban samples, indicating a more diverse presence of these bacteria in these regions. Two agricultural and urban phylotypes were phylogenetically related to the Syntrophobacteraceae and Syntrophaceae, bacteria that are typically syntrophic propionate oxidizers that can also degrade a variety of man-made pollutants (Bakermans and Madsen 2002; Fang et al. 2004). The phylotypes in this group are more closely related to other phylotypes from wastewater treatment plants (Layton et al. 2000) and batch reactors (Dubert et al. 2001).

Two phylotypic groups are also present in this cluster that include the 6G20 group that is most closely related to an Australian desert soil phylotype (Holmes et al. 2000) and the Q3 group that contains a phylotype from a potato rhizosphere in Germany (Heuer et al. 2002). Within open and remnant land use types, there are five phylotypes that are related to the genus Bdellovibrio, a group consisting of predatory, interperiplasmic bacteria that are known to frequent desert and other sandy-type soil habitats in order to
Figure 3.6 - 16S rRNA gene tree showing positions of the Polyangiaceae, Q3, Myxococceae, Syntrophonobacteriodetes, 6G20, and Bdellovibrio groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
Table 3.2 - 16S rRNA phylotype distribution of delta proteobacteria, gamma proteobacteria, and Acidobacteria within CAP LTER soils

<table>
<thead>
<tr>
<th>Taxonomic group (Φ)</th>
<th>Phylotypic or phylogenetic group</th>
<th>Φ*</th>
<th>No. of phylotypes (no. of clones) in the following land use type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Agricultural</td>
</tr>
<tr>
<td>Delta proteobacteria (27.1)</td>
<td>Polyangiaceae</td>
<td>20.0</td>
<td>2 (2)</td>
</tr>
<tr>
<td></td>
<td>Q3</td>
<td>3.4</td>
<td>2 (2)</td>
</tr>
<tr>
<td></td>
<td>Myxococceae</td>
<td>8.2</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>Syntrophonobacteroides</td>
<td>11.5</td>
<td>3 (5)</td>
</tr>
<tr>
<td></td>
<td>6G20</td>
<td>12.0</td>
<td>2 (11)</td>
</tr>
<tr>
<td></td>
<td>Bdeellovibrio</td>
<td>8.2</td>
<td>5 (5)</td>
</tr>
<tr>
<td>Gamma proteobacteria (26.1)</td>
<td>Ectothiorhodospirae</td>
<td>17.1</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>Thiotrichaceae</td>
<td>*</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>Hydrocarboniphaga</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pseudomonadaceae</td>
<td>15.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moraxellaceae</td>
<td>17.1</td>
<td>2 (3)</td>
</tr>
<tr>
<td></td>
<td>Xanthomonadaceae</td>
<td>12.3</td>
<td>3 (3)</td>
</tr>
<tr>
<td></td>
<td>Coxiallaceae</td>
<td>*</td>
<td>1 (16)</td>
</tr>
<tr>
<td>Acidobacteria (27.8)</td>
<td>Group I</td>
<td>*</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>Group II</td>
<td>15.3</td>
<td>13 (28)</td>
</tr>
<tr>
<td></td>
<td>Group III</td>
<td>21.1</td>
<td>14 (40)</td>
</tr>
<tr>
<td></td>
<td>Group V</td>
<td>15.9</td>
<td>33 (107)</td>
</tr>
<tr>
<td></td>
<td>384-2</td>
<td>23.2</td>
<td>3 (3)</td>
</tr>
<tr>
<td></td>
<td>Iii-8</td>
<td>14.0</td>
<td>6 (14)</td>
</tr>
<tr>
<td></td>
<td>Sva0725</td>
<td>11.2</td>
<td>2 (3)</td>
</tr>
</tbody>
</table>

*a Percent 16S rRNA dissimilarity throughout phylotypes in the group

*b Phylotypes and phylogenetic groups correspond to cluster of similar 16S rRNA gene sequences in the phylogenetic trees in Fig. 3-6 to 3-9.

*c Asterisks represent groups with one phylotype. Therefore, no genetic variability can be measured.

prey on Gram negative bacteria (Jurkevitch 2000). The increase of alpha proteobacterial prey abundance in open and desert remnants (Table 3.2), may be correlated to this distribution pattern within the CAP LTER soils.
**Gamma Proteobacteria**

The gamma proteobacteria phylotypes detected were diverse throughout the CAP LTER region (Fig. 3.7), corresponding to chemoheterotrophic bacterial groups *Ectothiorhodospiraceae* and *Hydrocarboniphaga*, as well as a phylotypic group (gamma-1) with no similar representation in the database. Four phylotypes found only in agricultural and/or urban sites were grouped within the *Pseudomadaceae* and *Moraxellaceae*, two families that are composed of aerobic mesophiles found in many soil environments (Moore et al. 1996; Mergaert et al. 2003). Five phylotypes were associated with the genera *Stenotrophomonas* and *Lysobacter* of the family *Xanthomonadaceae*. *Stenotrophomonas* species have been found to degrade man-made pesticides and have been used as a bioremediators (Chen et al. 2004), which may be indicative of their ability to survive in urban and agricultural soils that are assumed to contain higher amounts of man-made compounds. *Lysobacter* species are found in a diverse range of water and soil environments and have the ability to lyse eukaryotic cell walls (Sullivan et al. 2003). When comparing phylotype richness of the gamma proteobacteria across land use types, no changes in diversity were observed (Table 3.2).

**Acidobacteria and Relatives**

The presence of *Acidobacteria* in environmental 16S rRNA gene clone libraries has been well documented (Liesack and Stackebrandt 1992; Stackebrandt et al. 1993; Udea et al. 1995; Bornemann et al. 1996; Furlong et al. 2002). Due to the lack of cultured *Acidobacteria* isolate 16S rRNA gene sequences, phylogenetic groups in this study were designated in the same manner as Barns et al. (1999). Figure 3.8 represents Groups I and II, as well three other phylotypic groups. Group I includes the type species
**Figure 3.7 - 16S rRNA gene tree showing positions of the Ectothiorhodospiraceae, Thiotrichaceae, Pseudomonadaceae, Moraxellaceae, Xanthomonadaceae, and Coxiallaceae groups found in CAP LTER samples.** Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

*Acidobacterium capsulatum*, which is a chemoorganotrophic bacterium isolated from an acidic mineral environment in Japan (Kishimoto et al. 1991). Only one phylotype from an agricultural soil was found within this group, which is surprising because of the detection of this group in two northern Arizona soils (Barns et al. 1999). Group II contains twenty CAP LTER phylotypes, the majority of which branch near Arizona desert soil clones (Kuske et al. 1997) or an agricultural clone from Germany (Lukow 2000). The three phylotypic groups are related to phylotypes from other environments including clone 384-2 from a Californian agricultural soil (Valinsky et al. 2000), a clone iii-8 from a German field (Ludwig et al. 1997), and a marine sediment collected from the
Figure 3.8 - 16S rRNA gene tree showing positions of the *Acidobacteria* groups I and II, and phylotypic groups 384-2, iii-8, and Sva0725 found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

Arctic ocean (Ravenschlag et al. 1999). Group III (Fig. 3.9) contains fifty-four phylotypes representing a large degree of genetic dissimilarity (21.1% total) but grouping near the Arizona soil clones S027 and C105 (Kuske et al. 1997). Group IV is
Figure 3.9 - 16S rRNA gene tree showing positions of the Acidobacteria group III found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
represented by a small number of forest and agricultural soil clones in Genbank database and is not represented by any phylotypes in this study. Group V is represented most frequently, with fifty-eight CAP LTER phylotypes grouping here (Fig. 3.10). These phylotypes are also phylogenetically diverse (15.9%) and display similarity to a collection of agricultural, rhizosphere, and grassland soil clones. Group VI is composed of the iron-reducing *Geothrix* (Coates et al. 1999) and homoacetogenic *Holphaga* (Liesack et al. 1994) genera. There are no CAP LTER phylotypes that are within this group, which is not surprising considering phylotypes from other sources are from tundra (Zhou et al. 1997) and a uranium-contaminated aquifer (Anderson et al. 2003) environments. These results indicate that group VI is limited to certain specialized environments, unlike the other groups within *Acidobacteria*. Group VII represents an additional undescribed group that is associated with one phylotype from a field containing genetically modified plants in Germany (Ludwig et al. 1997).

Additionally, two phylogenetic groups within the *Acidobacteria* group appear more frequently in certain land use types (Table 3.2). Group III phylotype richness increases in open and remnant desert land use types. This trend is supported by the large number of phylotypes in the Genbank database belonging to this group coming from arid desert soils. In contrast, group V phylotypes are more prevalent in the more moisture-rich urban and agricultural environments.

**Actinobacteria and Relatives**

Phylogenotypes belonging to the *Actinobacteria* are present in over twenty phylotypic and phylogenetic groups. Eight phylotypes are grouped within the *Geodermatophilaceae* (Fig. 3.11), a family, with only three described species that typically colonize stone and
Figure 3.10 - 16S rRNA gene tree showing positions of the *Acidobacteria* groups V and VI found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

monument surfaces as well as arid soils (Eppard et al. 1996; Urze et al. 2001; Urzi et al. 2004; Rainey et al. 1997). Groups of 5 phylotypes were identified as being members of the families *Pseudonocardiaceae* and *Kineosporaceae*. One phylotype each was grouped with *Frankiaceae* species, a group of nitrogen-fixing symbionts (Benson and Sylvester...
Figure 3.11 - 16S rRNA gene tree showing positions of the Geodermatophilaceae, Pseudonocardiaceae, Kineosporaceae, Frankiaceae, and Acidothermaceae groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

1993; Normand et al. 1996), and Acidothermaceae, a group of cellulytic bacteria originally isolated from hot springs (Mohagheghi et al. 1986).

The Rubrobacteraceae were found to be diverse within all the CAP LTER soils (Fig. 3.12). This result is not surprising because of previous studies that have shown this group to be widely distributed in nature (Rheims et al. 1996; Holmes et al. 2000). In this study, the majority of Rubrobacteraceae phylotypes were found in open desert samples, possibly indicating a selective advantage of members of this group in undisturbed desert environments. More specifically, the phylotypes are designated into three phylogenetic groups and one phylotypic group with intra-group sequence divergence in the range of 10.7-16.1%. The Rubrobacter group includes two described thermophilic species
Figure 3.12 - 16S rRNA gene tree showing positions of the *Rubrobacteraceae* group found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

(*R. radiotolerans* and *R. xylanophilus*), both of which have been isolated from thermal sources (Suzuki et al. 1993; Ferreria et al. 1999). The phylotypes are not closely related to the two *Rubrobacter* species (>6.8% sequence divergence), and more closely related to a Australian desert soil phylotype (Holmes et al. 2000). A phylotypic group described as 1G9 is composed of eleven phylotypes and is only related to another desert clone from arid soil in Australia (Holmes et al. 2000). Ten phylotype groups were shown to fall...
within the *Solirubrobacter/Conexibacter* clade, two recently described species isolated from a US agricultural soil (Singleton et al. 2003) and forest soil in Italy (Monciardini et al. 2003), and German peat bog clones (Rheims et al. 1996). The *Thermoleophilium* group, assigned for the species *Thermoleophilium album*, an aklane-degrading bacterium isolated from thermal and non-thermal regions of the USA (Yakimov et al. 2003). Within this group, the majority of CAP LTER phylotypes group near an Australian desert soil clone (Holmes et al. 2000) and two agricultural soil clones (Valinsky et al. 2002).

The remaining phylogenetic group within *Actinobacteria* found in the CAP LTER soils is the *Acidimicrobiaceae* family (Fig. 3.13), another group which is lacking sufficient numbers of cultured representatives and is composed mainly of phylotypes from a diverse collection of environments (Rheims et al. 1996; Bond et al. 2002; Johnson et al. 2003). Thirty-one phylotypes are subdivided into four groups. The *Microthrix* group contains seven CAP LTER phylotypes. *Microthix parvicella* and other environmental phylotypes have been shown to be filamentous bacteria typically present in wastewater treatment plants (Blackall et al. 1996; Rosetti et al. 1997; Krhutkova et al. 2002; de los Reyes et al. 2002), but they have also been discovered in rice paddy soils (Ludemann and Conrad 2000). Other phylotypic groups within the *Acidimicrobiaceae* are most closely related to environmental sequences from a wide range of habitats, including clones from marine sediments (Reed et al. 2002) and subtropical forest (Stackebrandt et al. 1993) locations.

A small number of phylotypes also grouped within the families

*Microbacteriaceae, Cellumondaceae, Friedmanellaceae, Intrasporangiaceae,*

*Microccaceae, Actinosynnemataceae, Streptomycetaceae, Nocardioidaceae,*
Figure 3.13 - 16S rRNA gene tree showing positions of the *Acidimicrobiaceae* group found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

*Propionibacteraceae*, and *Micromonsporaceae* (Fig. 3.14). Most of the members of these families are heterotrophic aerobes commonly found in soils. There are two other phylotypic groups labeled 7L14 and AY10 that are strongly diverged from other actinobacterial groups (>13%).

Most of the *Actinobacteria* groups show no land use type specificity (Table 3.3). However, the members of the *Rubrobacteraceae* family, are clearly more phylotypically
Fig 3.14 - 16S rRNA gene tree showing positions of the Microbiaceae, Cellumonadaceae, Intrasporangiaceae, Micrococcaeae, Actinosynnemataceae, Streptomyctaceae, Nocardiodaceae, Propionibacteraceae, Micrononosporaceae, and 7L14 groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

rich in open desert environments. This may be a result of radiation resistance of some Rubrobacter species (Suzuki et al. 1988; Chen et al. 2004). Radiation resistance has been shown to positively correlate with desiccation resistance (Mattimore and Battista 1996), which may be a strong selective advantage for bacterial populations residing in open desert environments that receive little water.
Table 3.3 - 16S rRNA phylotype distribution of *Actinobacteria*, *Bacteriodetes*, *Chloroflexi*, and *Cyanobacteria* within CAP LTER soils

<table>
<thead>
<tr>
<th>Taxonomic group (a)</th>
<th>Phyotypic or phylogenetic group (b)</th>
<th>No. of phylotypes (no. of clones) in the following land use type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Agricultural</td>
</tr>
<tr>
<td>Actinobacteria (25.5)</td>
<td>Geodermatophilaceae</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>Pseudonocardoidaceae</td>
<td>15.4</td>
</tr>
<tr>
<td></td>
<td>Kinetoplastidae</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>Frankiaceae</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Acidothermaceae</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Acidimicrobiaceae</td>
<td>15.1</td>
</tr>
<tr>
<td></td>
<td>Rubrobacteriae</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>Microbacteriaceae</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>Cellulosimicrobiaceae</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>Nocardiaceae</td>
<td>11.3</td>
</tr>
<tr>
<td></td>
<td>Intrasporangiaceae</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Microccaceae</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>Streptomycetaceae</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>Propionibacteraceae</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Micromonosporaceae</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>TL14 group</td>
<td>17.0</td>
</tr>
<tr>
<td>Bacteriodetes (26.7)</td>
<td>Flexibacteraceae</td>
<td>21.1</td>
</tr>
<tr>
<td></td>
<td>Flavobacteraceae</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>BD7 group</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>KL group</td>
<td>11.6</td>
</tr>
<tr>
<td>Chloroflexi (25.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyanobacteria (17.2)</td>
<td>Microcoleus</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>Leptolyngbya</td>
<td>*</td>
</tr>
</tbody>
</table>

a Percent 16S rRNA dissimilarity throughout phylotypes in the group
b Phylotypes and phylogenetic groups correspond to cluster of similar 16S rRNA gene sequences in the phylogenetic trees in Figs. 3-10 to 3-16.

c Asterisks represent groups with one phylotype. Therefore, no genetic variability can be measured.

**Bacteriodetes and Relatives**

Two of the four previously described families in the *Bacteriodetes* and two new phylotypic group are represented within the CAP LTER phylotypes (Fig. 3.15). Thirty-
six phylotypes group within the *Flexibacteraceae*. Most of the members of this family are found in soil, chemoorganotrophic and possess gliding motility. However, most of these phylotypes are highly divergent (>7%) from other sequences in the database, and therefore extrapolating potential physiological or ecological characteristics is difficult. Twenty-four phylotypes grouped within the family *Flavobacteriaceae*, a family that encompasses closely related soil and clinical organisms (Bernardet et al. 2002). The two phylotypic groups most closely resemble the deep-sea sediment strain BD-7 (Li et al. 1999) and a clone taken from a clean room facility (Venkateswaran et al. 2003). Within the *Bacteriodetes*, the BD-7 phylotypic group was the only cluster in which differentiation in phylotypic richness was observed in urban soil types (Table 3.3).

**Chloroflexi**

The seven phylotypes that grouped within the *Chloroflexi* phylum were also highly diverged from other 16S rRNA gene sequences (<9.5%) located in the Genbank database (Fig. 3.16). Six phylotypes are most similar to an Australian desert clone (Holmes et al. 2000). The majority of the phylotypes were found within only one land use type, indicating a higher degree of habitat specificity than other groups (Table 3.3).

**Cyanobacteria**

Nearly all CAP LTER cyanobacterial phylotypes formed a group that included the *Anabaena* and *Microcoleus* genera within the phototrophic *Cyanobacteria* phylum (Fig. 3.17). The phylogenetic position of the phylotypes was expected because of the multiple studies involving soil crusts in which these taxa are present in large numbers due to nitrogen-fixing capabilities, which is beneficial for organisms in nitrogen-limited
Figure 3.15 - 16S rRNA gene tree showing positions of the *Flexibacteraceae*, *Flavobacteraceae*, BD-7 and KL groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
Figure 3.16 - 16S rRNA gene tree showing positions of the *Chloroflexi* group found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

Figure 3.17 - 16S rRNA gene tree showing positions of the *Cyanobacteria*, phylotypes found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
environments such as arid soil crusts (Anderson and Rushforth 1976; Belnap and Gardner 1993; Belnap 2002; Redfield et al. 2002; Yeager et al. 2003). Most phylotypes are present in all five land use types, indicating the ability to maintain populations within soil crusts despite urbanization and agricultural disturbances (Table 3.3).

**Firmicutes and Relatives**

Fifty-seven percent of the phylotypes within the *Firmicutes* fall within the family *Bacillaceae* (Fig. 3.18). Although present throughout the CAP LTER, the *Bacillaceae* are dominated by phylotypes from agricultural land use types (Table 3.4). Eight of these phylotypes are less than 3.5% dissimilar to *Bacillus* sp. LMG-1945, which was isolated from a wall painting in Europe (Gurtner et al. 2000). The majority of phylotypes are less than 5% dissimilar from well documented species *B. fastidiosus*, *B. niacini*, *B. benzoeverans*, *B. sirlis*, *B. silvestris*, and *B. cereus*. All of these species are commonly found in soils and have spore-forming capabilities. Phylotypes resembling the previously described *Bacillus* species have been shown to demonstrate a high degree of genomic variability across similar environments, indicating the capability for *Bacillus* populations in soil to fill separate ecological roles based upon habitat (Felske et al. 2004). Most members of the *Bacillaceae* also possess the ability to form spores, a means of maintaining a presence in soils when environmental conditions are potentially lethal. This ecological robustness may explain the presence of *Bacillaceae* in the various CAP LTER soils. Smaller amounts of phylotypes grouped within the spore-forming *Alicyclobacillaceae*, *Paenibacillaceae*, and *Clostridiaceae* families, as well as the non-spore-forming *Peptococceae*, *Staphylococceae*, *Carnobacteraceae* and *Syntrophonomaceae* groups (Fig. 3.19).
Table 3.4 - 16S rRNA phytype distribution of *Firmictues*, *Gemmatinonadetes*, *Planctomycetales*, and *Verrucomicrobia* within CAP LTER soils

| Taxonomic group (α) | Phylotypic or phylogenetic groupβ | No. of phylotypes (no. of clones) in the following land use type |  |  |  |  |
|---------------------|----------------------------------|---------------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Firmicutes (25.5)   | *Bacillus*                       | θ                           | *Bacillus*       | 10.1            | 19 (236)         | 5 (13)          | 12 (60)         | 10 (60)         | 19 (236)        |
|                     | *Alicyclobacillus*               |                               | 7.6              | 2 (3)           | 3 (5)            | 1 (1)           | 4 (12)          |
|                     | *Clostridiaceae-1 group*         |                               | 9.7              | 2 (3)           | 2 (4)            |                  | 2 (26)          |
|                     | *Thermoterrabacterium*           |                               | *               |                  |                  |                  | 1 (1)           |
|                     | *Paenibacillus*                  |                               | 11.1             | 3 (11)          | 2 (3)            | 2 (3)           | 3 (48)          |
|                     | *Staphylococcus*                 |                               | *               | 1 (1)           | 1 (3)            | 1 (3)           | 1 (1)           |
|                     | *Dolosigranulum*                 |                               | *               | 1 (1)           | 1 (1)            |                  |                 |
|                     | *Clostridiaceae-2 group*         |                               | 8.0              | 1 (1)           | 1 (1)            | 1 (1)           | 2 (1)           |
|                     | *Syntrophomonas*                 |                               |                  | 1 (1)           | 1 (1)            | 1 (1)           | 1 (1)           |
| Gemmatimonadetes (26.7) | Group I                          |                               | 14.1             | 23 (42)         | 5 (9)            | 35 (108)        | 31 (93)         | 27 (59)         |
|                     |                                  |                               |                  |                 |                 |                  |                 |                 |
|                     |                                  |                               | 15.4             | 17 (25)         | 3 (3)            | 14 (42)         | 20 (25)         | 20 (33)         |
|                     |                                  |                               |                  |                 |                 |                  |                 |                 |
|                     |                                  |                               | 12.3             | 14 (55)         | 9 (13)           | 25 (89)         | 14 (48)         | 23 (84)         |
|                     |                                  |                               |                  |                 |                 |                  |                 |                 |
|                     |                                  |                               | 25.7             | 1 (4)           | 1 (1)            |                  |                 | 2 (2)           |
| Planctomycetales (28.3) | *Gemmata*                         |                               | 20.6             | 6 (6)           | 1 (1)            | 9 (13)          | 8 (8)           | 3 (5)           |
|                     |                                  |                               | 8.0              | 1 (1)           | 1 (1)            | 1 (14)          | 2 (6)           | 1 (3)           |
|                     |                                  |                               |                  |                 |                  |                  |                 |                 |
|                     |                                  |                               | 16.6             | 2 (2)           | 3 (3)            | 1 (1)           | 3 (4)           |
|                     |                                  |                               |                  |                 |                  |                  |                 |                 |
|                     |                                  |                               | 20.1             | 6 (14)          | 2 (2)            | 8 (22)          | 4 (11)          | 5 (23)          |
| Verrucomicrobia (28.9) | *Spartobacteria*                 |                               | 19.3             | 13 (18)         | 3 (4)            | 29 (48)         | 15 (35)         | 22 (42)         |
|                     |                                  |                               |                  |                 |                 |                  |                 |                 |
|                     |                                  |                               | 22.1             | 4 (4)           | 2 (2)            |                  | 4 (5)           |
|                     | *Verrucomicrobia group III*      |                               |                  |                 |                  |                  |                 |                 |
|                     | *Verrucomicrobia*                |                               | 10.0             | 1 (2)           | 1 (3)            |                  | 2 (3)           |
|                     | *Verrucomicrobia group IV*       |                               |                  |                 |                  |                  |                 |                 |

a Percent 16S rRNA dissimilarity throughout phylotypes in the group

b Phylotypes and phylogenetic groups correspond to cluster of similar 16S rRNA gene sequences in the phylogenetic trees in Fig. 3-17 to 3-22.

c Asterisks represent groups with one phylotype. Therefore, no genetic variability can be measured.
Figure 3.18 - 16S rRNA gene tree showing positions of the Balliceae group found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
Figure 3.19 - 16S rRNA gene tree showing positions of the Alicyclobacillaceae, Clostridiaceae, Peptococceae, Paenibacillaceae, Staphylococceae, Carnobacter Aceae, and Syntrophonomonadaceae, groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

Gemmatimonadetes and Relatives

Gemmatimonadetes is represented by a single cultured species, G. aurantiaca, a heterotrophic aerobe isolate from a wastewater treatment system, as well as three environmental phylotypic groups composed of environmental sequences collected from a wide range of habitats (Zhang et al. 2003). Group I is well represented in the CAP LTER, with 91 phylotypes that diverge by 13.4% (Fig. 3.20). Six phylotypes form a tight group with G. aurantiaca, while the majority of the other phylotypic groups most closely
resemble Australian soil clones (Holmes et al. 2000), Wyoming grassland clones
(Mummey and Stahl 2003), an Arizona desert soil clone (Kuske et al. 1997), a recently
disturbed Canadian forest soil clone (Axelrood et al. 2002), and a Georgia forest wetland
soil clone (Broff et al. 2002). Phylotypes representing Gemmatimonadetes groups II-IV
are also present, but at lower levels than Group 1 (Fig. 3.21). One clade of ten
phylotypes show lower-order similarity to Arizona soil clone S0134 (>92%). However,
the remaining phylotypes that group within group II are highly divergent from other
phylotypes in the database (>10%). This is also the case in group III, in which clusters of
three and five phylotypes demonstrate high similarity with clones from a Gulf of Mexico
hydrate (Lanoil et al. 2001) and a Wisconsin agricultural soil (Bintrim et al. 1997)
respectively. Group IV has previously contained only phylotypes of marine origin
(Zhang et al. 2003), but here we find two urban phylotypes that may be distantly related
to this group (>10%). Overall, no changes in phylotypic richness were observed in any
of the phylotypic groups (Table 3.4).

**Planctomycetes and Relatives**

The only described family within the phylum Planctomycetes is Planctomycetaceae,
which contains four validly described genera that demonstrate a diverse array of
physiological characteristics (Staley et al. 1992). All of the described species have been
isolated or observed in aquatic environments, but 16S rRNA gene clone libraries have
uncovered the presence of Planctomycetes in soil habitats (Neef et al. 1997). Twenty-
three phylotypes are grouped within the Planctomycetaceae, the majority of which were
associated with the genus Gemmata (Fig. 3.22).
Figure 3.20 - 16S rRNA gene tree showing positions of the Gemmatimonadetes group I found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
Figure 3.21 - 16S rRNA gene tree showing positions of the Gemmatimonadetes groups II, III, and IV found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.
Figure 3.22 - 16S rRNA gene tree showing positions of the Gemmata, Isophaera, Planctomyces, Pirellula, and 96-2 groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

There was one divergent phylotypic group of eleven CAP LTER phylotypes that appears to be related to an agricultural soil clone from California (Valinsky et al. 2000). No other phylotypes from the Genbank database were more than 90% similar to the 96-2 group, thus indicating that this group represents a novel lineage within this family. No land use type stratification in phylotypes was observed (Table 3.4).

**Verrucomicrobia and Relatives**

The phylum Verrucomicrobia contains only a few validly described species
(Hedlund et al. 1997; Hugenholtz et al. 1998b; Ward-Rainey et al. 1995), but is detected regularly in soils from around the world. *Verrucomicrobia* has been divided into five monophyletic subdivisions by Hugenholtz et al. (1998b). The majority of CAP LTER phylotypes fall within the radiation of subdivision 2 group (Fig. 3.23), also known as the *Spartobacteria*, and includes the aerobic nitrogen-fixing species *Chthoniobacter flavus* (Sangwan et al. 2004) and the endosymbiotic genus *Ximenthobacter* (Vanderkerckhove et al. 2000). Many of these phylotypes are found in the four major land use types, but are more dominant in the urbanized soils. Most of the CAP LTER phylotypes are similar an Arizona desert clone collected by Kuske et al. (1997) and to *C. flavus*, indicating that *Spartobacteria* may fill an important ecological role in nitrogen cycling within desert environments. Eight phylotypes are grouped within subdivision 3, a group that includes a Yellowstone hot spring phylotype (Hugenholtz et al. 1998a) and industrial sludge (Juretschko et al. 2002) clones. Only one agricultural clone was grouped within the *Verrucomicrobia* group, which is not surprising considering both phylotypes and isolates are all of aquatic origin. Only two phylotypes fall into the *Opitutus* group, which comprises a collection of phylotypes and one valid species, *Opititus terrae* (Chin et al. 2001), all of which were found in soil environments (Janssen et al. 1997; Holmes et al. 2000).

**Other Groups**

Phylotypes that were affiliated with other phyla and candidate phyla were also found in this study (Fig. 3.24). Four phylotypes were related to members of the *Deinococcus-Thermus* phylum, more specifically the genus *Deinococcus*, which comprises isolates that exhibit radiation and desiccation resistance through advanced
Figure 3.23 - 16S rRNA gene tree showing positions of the **Spartobacteria**, **Verrucomicrobia**, and **Verrumicrobia III and IV** groups found in CAP LTER samples. Colored blocks adjacent to phylotypes represent presence in an agricultural (green), mixed (black), open (orange), remnant (yellow), and urban (blue) land use types. Grey circles represent bootstrap values of >50%. Black circles represent bootstrap values of >75%.

DNA repair mechanisms (Minton 1994, Minton 1996). Much like the desiccation-resistant *Rubrobacter* species, the majority of CAP LTER *Deinococcus* phylotypes are found only in open desert soils.
Figure 3.24 - Unrooted radial 16S rRNA tree showing positions of the *Deinococcus-Thermus* phylum, and candidate phyla OP3, OP10, TM7, and WS6 in the CAP LTER.
CAP LTER phylotypes were also affiliated with at least four candidate phyla that lack cultured representation. This includes four phylotypes that group with WS6, a group first described by phylogenetic analysis of bacterial communities in contaminated aquifers (Dojka et al. 1998). Six phylotypes are distantly related (<90%) to members of the OP3 group, including phylotypes collected from hydrothermal deep-sea sediments (Li et al. 1999; Teske et al. 2002) and heavy-metal contaminated soils (Marsh et al. 1999). Twenty-one phylotypes were affiliated with the OP10 group. This group includes other phylotypes from a Yellowstone hot spring (Hugenholtz et al. 2001), Australian desert soils (Holmes et al. 2000), and a trichloroethylene-contaminated soil (Lowe et al. 2002). Although some phylotypes are only found in individual land use types, across the OP10 group there appears to be no significant changes in distribution across land use types. Finally, the candidate phylum TM7 is represented in a small number the CAP LTER soils. This group also includes phylotypes from deep-sea hydrothermal sediments (Li et al. 1999), Yellowstone hot springs (Hugenholtz et al. 2001), and trichlorobenzene-contaminated soils (von Wintzingerode et al. 1999). The three TM7 group CAP LTER phylotypes are found only in open desert or remnant samples, the two land use types that receive the least amount of precipitation. Since no members of these groups have been isolated through culture, little is known about the physiology and ecology of these groups, which limits the ability to explain the limited distribution of these groups at the CAP LTER site.

**Discussion**

As a result of the extensive numbers of 16S rRNA gene clones generated and analyzed in this study, the observed phylogenetic diversity within the CAP LTER region
is remarkable. Due to this high degree of species heterogeneity, we constructed phyllogenetic trees to serve as a useful tool for identifying instances of habitat specificity with phylotypic and/or phylogenetic groups. Most groups sampled are distributed evenly across land use types, indicating a wide range of adaptability for members of these groups. Many of these phylotypes are closely related to strains and or phylotypes that were collected from contaminated environments. The ability to degrade environmental pollutants may provide a selective advantage for many bacterial species throughout the CAP LTER site. This might enable such species to sustain a presence in urbanized environments that are likely to contain higher levels of pollutants. Most of the phylogenetic groups, in which CAP LTER phylotypes are clustered closely together, are highly dissimilar when compared to other environmental sequences within the available database of 16S rRNA gene sequences. This tight grouping of phylotypes, although defined as separate OTUs based upon ARDRA fingerprint data, may be filling similar ecological roles within given habitats. However, the slight deviation in 16S rRNA gene sequences may be accounted for in the increasingly higher degrees of niche specialization. These genetic polymorphisms are analogous to Wallace’s “bed of nails” model (1968), where each genomic polymorphism represents an adaptive peak achieved by individuals at a particular microhabitat. In certain taxonomic groups, genomic divergence due to ecological adaptation has been shown not to be directly correlated to significant changes in 16S rRNA gene sequences (Jaspers and Overmann 2004). Phylotypes found in multiple land use types may undergo genomic polymorphisms to gain selective advantages over competing species based on the physical properties of the different soils. Many of the CAP LTER phylotypes analyzed were most similar to other
phylotypes collected from the western United States (California, Arizona, Wyoming), indicating a broad geographic range of some bacterial species that may fill an important role in soil ecology.

The distribution of phylotypic/phylogenetic groups across many land use types may also be due to the heterogeneity within individual soil samples, especially agricultural and urban soils. Agricultural lands are subcategorized into pastures and croplands despite significant differences in the management of these soils in terms of irrigation and fertilization. Likewise, urban samples include residential lawn, transportation, and industrial sites, where the physical properties of soils can vary widely. The CAP LTER phylotype dataset is not large enough to derive comparisons based on subcategories, limiting the validity of inferences on habitat constriction. However, the results obtained provide insight into possible habitat restrictions on the biogeography of certain bacterial groups across land use types. In these cases, the availability of niches within these land use types is lacking, presumably due to changes in soils which ultimately limit habitat range. Research focusing on these groups is needed to determine the mechanisms responsible for these observed restrictions in habitats.
CHAPTER 4:

DETECTION OF UBIQUITOUS TAXA ACROSS THE CAP LTER
Introduction

Among the prokaryotic taxa detected in different environments throughout geographically distant locations, there is some observed degree of phylogenetic overlap, which provides further evidence of the “everything is everywhere” theory pertaining to global microbial distribution (Bass-Becking 1934). Several epidemiological studies have also shown that several spore-forming Bacillus and various other pathogenic genotypes of species including Salmonella enterica, Haemophilus influenzae, and Neisseria meningitidis are detectable at a global scale (Smith et al. 1990; Musser et al. 1990; Maynard-Smith et al. 1993; Roberts and Cohan 1995). These observed distributions of pathogenic prokaryotes appear fit the “everything is everywhere” model as a consequence of the ability to be disseminated through plant and animal vectors. Widespread distribution of other free-living bacterial groups has also been observed. 16S rRNA genes sequenced from bacterial cultures as well as phylotypes showing high gene sequence similarity are being detected repeatedly, often from geographically isolated locations. Strains of the cyanobacterium Microcoleus chthonoplastes containing identical 16S rRNA gene sequences were isolated across both the North American and European continents (Garcia-Pichel et al. 1996). Similar results have been documented in Gram positive, Gram negative, and symbiotic bacterial species (Grayson et al. 1999; Cho et al. 2000; Guevara et al. 2002). There is even a documented case in which known thermophilic bacteria have also been successfully isolated from cold seawater (Isaken et al. 1994). Multiple studies of archaeal species have shown high degrees of similarity within widely disparate locations (Stetter et al. 1993; Maggot 1996).

Many of these cases of phylogenetic overlap across environments involve prokaryotic taxa with little or no cultured representation, making any distinctive
explanation of ecophysiology difficult. Analysis of phylotypes present in a peat bog in Germany highlighted the presence of members of the deep branching actinomycete lineages represented in culture by the genera *Rubrobacter* and *Acidimicrobium* in a wide range of environments at geographically isolated locations (Rheims et al. 1996). Studies using 16S rRNA gene primers specific for the *Acidobacterium* phylum have shown there to be representatives of this taxa in 27 environmental samples that were diverse in terms of soil type, pH, and temperature (Barns et al. 1999). In a compilation of results of numerous molecular studies it was noted that 16S rRNA gene sequences representing certain bacterial divisions (referred to as “cosmopolitan” divisions) have been found in diverse habitats, and that the majority of these phylotypes lack cultured representation (Hugenholtz et al. 1998).

Global dispersion of prokaryotic species may best explain widespread distribution. This phenomenon is supported by the ability of various microbial species to be transported via air currents across continents in a span of approximately six days (Jaffe et al. 1999; Brown and Hovmeller 2002). It has been calculated that this method of dispersal permits $10^{18}$ viable soil bacterial cells to be transported across continents annually (Griffin 2002). However, even when assuming global dispersal events involving soil bacterial species occur, endemic populations of some bacterial taxa have been observed (Papke et al. 2003; Whitaker et al. 2003). Therefore, those taxa that are constantly detected across habitats must possess an immense range of adaptability and may allow for an immense consortium of bacterial species residing in heterogeneous environments. This degree of adaptability comes from accelerated rates of evolution via mutation and generation time (Arbor 2000; Ochman et al. 2000; Spiers et al. 2000), large
effective populations, inter-population genetic exchange (Shapiro 1985), and subsequent periodic selection (Levin 1981). Ultimately, these unique properties belonging to prokaryotic species can limit the probability of local extinctions, erasing geographic barriers and leading to pandemic distribution.

The ubiquitous presence of microbial species in these same habitats indicates that these taxa are ecological opportunists with the capability to survive in a diverse array of niches. Ecological opportunist microbes were initially observed in the laboratory of Gause, where different Paramecium species competed directly against each other, leading to the foundation of the “niche-exclusion” principle (Gause 1934). Since these classic experiments were performed, numerous studies have attempted to explain what components within a prokaryotic genome are essential for increased physiological suppleness. When studying housekeeping loci, no substantial distinction between ecologically dominant species has been found (Sprangenberg 1998; Kiewtiz and Tummler 2000). However, within these genomes, increased assortments of protein kinases have been detected which allow for mutated conformations of enzymatic targets or altered activation timing (Jablanka 1998). Increased levels of variation in physiological genes within prokaryotic phylotypes have also been observed, permitting diverse niche specialization within one defined species (Maynard Smith 1991; Futuyma 1998; Oda et al. 2004).

We have demonstrated that there are abrupt changes in bacterial diversity based upon land use type at the CAP LTER region that is a result of disjunctive patterns of distribution among many taxa. Using the 16S rDNA clone database that we established
The bacterial communities were analyzed by ARDRA fingerprints to identify phylotypes that are common in the CAP LTER region. These phylotypes were used as templates for designing phylotype-specific PCR reactions.

### Table 4.1 - CAP LTER phylotypes chosen for 16S rRNA gene PCR primer design.

<table>
<thead>
<tr>
<th>Taxonomic Group (banding pattern number)</th>
<th>Closest relative (accession no.)</th>
<th>Similarity (%)</th>
<th>Overall abundance in all 16S rRNA gene clone libraries (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha Proteobacteria (392)</td>
<td>Sphingomonas sp. SIA181-1A1 (AF395032)</td>
<td>95.8</td>
<td>11.6</td>
</tr>
<tr>
<td>(374)</td>
<td>Sphingomonas sp. SIA181-1A1 (AF395032)</td>
<td>97.9</td>
<td>8.6</td>
</tr>
<tr>
<td>(373)</td>
<td>Sphingomonas Clone 768-2 (AF423293)</td>
<td>97.0</td>
<td>2.9</td>
</tr>
<tr>
<td>(375)</td>
<td>Sphingomonas Clone 768-2 (AF423293)</td>
<td>96.3</td>
<td>2.2</td>
</tr>
<tr>
<td>(376)</td>
<td>Sphingomonas asaccharolytica (Y09639)</td>
<td>96.4</td>
<td>0.5</td>
</tr>
<tr>
<td>(359)</td>
<td>Rhizosphere isolate RS1-21 (AJ252588)</td>
<td>97.5</td>
<td>0.5</td>
</tr>
<tr>
<td>(398)</td>
<td>Bosea thiostrepsis (AJ250799)</td>
<td>94.8</td>
<td>2.4</td>
</tr>
<tr>
<td>(444)</td>
<td>Methylbacterium organophilum (D32226)</td>
<td>95.6</td>
<td>3.4</td>
</tr>
<tr>
<td>(445)</td>
<td>Environmental clone SME102 (AF445680)</td>
<td>95.6</td>
<td>1.2</td>
</tr>
<tr>
<td>(423)</td>
<td>Rhodobacter gluconicum (AB1077986)</td>
<td>91.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Beta Proteobacteria (358)</td>
<td>clone F10201 (AF236005)</td>
<td>92.3</td>
<td>0.9</td>
</tr>
<tr>
<td>(377)</td>
<td>Methylkivibacteriaceae (AY157761)</td>
<td>97.2</td>
<td>0.7</td>
</tr>
<tr>
<td>(63)</td>
<td>Herbaspirillum sp. (Y101-46)</td>
<td>98.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Acidobacteria (493)</td>
<td>Clone S027 (AF013554)</td>
<td>96.2</td>
<td>0.7</td>
</tr>
<tr>
<td>(435)</td>
<td>Clone #0319-7B4 (AF234087)</td>
<td>96.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Gemmatimonadetes (360)</td>
<td>Clone FW125 (AF353020)</td>
<td>90.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Firmicutes (424)</td>
<td>Bacillus sp. MK03 (AB006267)</td>
<td>97.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Bacteroidetes (435)</td>
<td>Flexibacter aggregans (AB078038)</td>
<td>90.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Actinobacteria (411)</td>
<td>Geodermatophilus obscurus (Y92356)</td>
<td>96.8</td>
<td>0.9</td>
</tr>
<tr>
<td>(71)</td>
<td>Arthrobacter sp. (AF197053)</td>
<td>98.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Planctomycetes (487)</td>
<td>Clone WD2101 (AJ292687)</td>
<td>91.4</td>
<td>0.3</td>
</tr>
</tbody>
</table>

and that encompasses twenty-two sites and over 11,000 clones collected from the CAP LTER region, various analyses were performed to determine if overall biogeography of resident bacterial populations can be explained by recently-formed anthropogenic barriers. In this chapter, ARDRA fingerprints that were detected in a clear majority (>75%) of 16S rDNA clone libraries and within the four major land use types (agricultural, open, remnant, urban) were used as models for the design of phylotype-specific PCR reactions.
to test for the presence of these taxa in the remaining 167 CAP LTER soil samples collected (Table 4.1).

**Materials and Methods**

**Specific PCR Amplification of 16S rRNA Genes from Bulk Environmental DNA**

To test for the presence of members of specific bacterial groups, a variety of forward oligonucleotide primers were designed according to ARDRA fingerprint data using the PRIMROSE software package, which generates oligonucleotide primers against the Ribosomal Database Project-II 16S rRNA gene database (Ashhelford et al. 2002). PCR reactions were carried out using the designed group specific forward primer with the universal bacterial 1492r reverse primer. Amplification was performed in a 25 µL (total volume) reaction mixture using approximately 2, 20, and 200 ng of environmental DNA, 1U Taq polymerase, each deoxynucleotide phosphate at a concentration of 200 µM, and each primer at a concentration of 1 µM. The cycling conditions for the specific F-1492R reactions follow those described previously in Chapter 2 except for altered annealing temperatures based upon forward primer sequence and a prolonged final extension period of 15 minutes (Table 4.2). For reamplification reactions, 0.5 mL from the reactions with the respective primer set were used as template verify the presence/absence of a band. 3 mL of each reaction mixture was analyzed on 1% agarose gels, and DNA was visualized by ethidium bromide and UV illumination. A visible band of the anticipated size (depending on primer set used) was deemed a positive result. No visible product or products of the wrong size were counted as negative results. All samples were tested at least twice with the primer set to confirm results. Control reactions using the 27F and 1492R primers were carried out as described in Chapter 2.
Table 4.2 - Forward 16S rRNA gene primers used in the survey of environmental DNA samples.

| Phylogenetic Group | Primer | Nucleotide sequences (5’→3’)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Target region(^b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha Proteobacteria</td>
<td>Sphing1</td>
<td>GCCGCCGCTTTCCACCCCTTA</td>
</tr>
<tr>
<td></td>
<td>Sphing2</td>
<td>ATCATCCGGGTTAAAAGGAC</td>
</tr>
<tr>
<td></td>
<td>Sphing3</td>
<td>CGACTTATCCRACGCGGGC</td>
</tr>
<tr>
<td></td>
<td>Sphing4</td>
<td>GCCGAAGTTTCCRTCCGTA</td>
</tr>
<tr>
<td></td>
<td>Sphing5</td>
<td>TTACGCCGGCTCACCCCTAG</td>
</tr>
<tr>
<td></td>
<td>Rhizo1</td>
<td>TCTTAAAGGCAATTCCGAGG</td>
</tr>
<tr>
<td></td>
<td>Bose1</td>
<td>CGCCGAAATCTCTTCCGG</td>
</tr>
<tr>
<td></td>
<td>Rhod1</td>
<td>CGGCGATAAATCTTTCYCCY</td>
</tr>
<tr>
<td></td>
<td>SME1</td>
<td>GTTCCACTMACCTCTCCGG</td>
</tr>
<tr>
<td></td>
<td>A061</td>
<td>AATGCGCCGAGGGYCTT</td>
</tr>
<tr>
<td></td>
<td>FTL1</td>
<td>CGCTCCARRAGCAGTGGGCC</td>
</tr>
<tr>
<td></td>
<td>Herb1</td>
<td>MAACTCCCTCTCCCTGACA</td>
</tr>
<tr>
<td>Beta Proteobacteria</td>
<td>Acidobacteria</td>
<td>Acido1</td>
</tr>
<tr>
<td></td>
<td>Acidobacteria</td>
<td>Acid2</td>
</tr>
<tr>
<td></td>
<td>Gemmatimonadetes</td>
<td>Gemm1</td>
</tr>
<tr>
<td></td>
<td>Firm1</td>
<td>CCGCGCTAATCAAAGRAGAG</td>
</tr>
<tr>
<td></td>
<td>Bacteroidetes</td>
<td>CFB1</td>
</tr>
<tr>
<td></td>
<td>Actinobacteria</td>
<td>Actino1</td>
</tr>
<tr>
<td></td>
<td>Planctomyceales</td>
<td>Actino2</td>
</tr>
<tr>
<td></td>
<td>Planctomyceales</td>
<td>Plancto1</td>
</tr>
</tbody>
</table>

\(^a\)Y, mixture of C and T; W, mixture of A and T; Y, mixture of C and T; R, mixture of G and A; M, mixture of A and C; N, mixture of A, C, G, and T.

\(^b\)Positions correspond to *E. coli* nucleotide numbering (Gutell 1983).

\(^c\)Proportions in red indicate a 95% confidence interval at 100% based on standard error of proportion.
Results

Selection of ARDRA Fingerprint Groups for PCR Primer Design

The initial design of phylotype-specific PCR primers was directed by the partial sequencing of between five and twenty 16S rRNA gene clones that shared the same ARDRA fingerprint. The selected phylotypes were represented in at least 75% of CAP LTER 16S rRNA clone libraries constructed (Table 1). The abundance of fingerprints within the ARDRA pattern type ranged from 0.4 to 11.6% of all 16S rRNA gene clones analyzed.

Ten CAP LTER phylotypes that met these selection requirements were associated with the alpha proteobacteria. Two fingerprint groups that were most dominant in the CAP LTER 16S rRNA gene clone libraries were closely related to *Sphingomonas* strain SIA181, which was isolated from sea ice in Lake Vostok, Antarctica (Christner et al. 2001). Two phylotypes were most similar to a California agricultural soil clone sequence (Valinsky et al. 2000). In terms of isolate sequences, the fingerprint groups were most similar to *Sphingomonas mali*, a species isolated from apple tree roots in Japan (Takuchi et al. 1995). When compared to each other, the two fingerprint groups were ~ 96.8% identical, just below the 97% benchmark for delineation of prokaryotic species (Stackebrandt et al. 1994; Palys et al. 1997). Another fingerprint group was associated to *Sphingomonas asachharolytica*, a species also isolated in the same soil as *S. mali* (Takuchi et al. 1995). One other selected fingerprint group was associated with an unidentified *Sphingomonas* strain from a soil rhizosphere in Germany (Lukow 1999).

Other groups affiliated with the alpha proteobacteria were also present in the majority of all CAP LTER 16SrRNA gene clone libraries. One fingerprint group was
~95% identical to *Bosea thiooxidans*, a chemolithotrophic bacterium that is well known for its thiosulfate oxidizing abilities (Das et al. 1999). A second fingerprint group was associated with *Methylobacterium organiphilum*, a type II methylotroph in the alpha-2 subclass that has been shown to be resistant to various normally toxic compounds (Hirashi et al. 1995; Mostafa and Helling 2003). While another selected group was similar to an *Aminobacter*-type clone collected from Yellowstone National Park hot springs (Bonheyo et al. 1999). A fingerprint group that was related to the *Rhodobacter* genus, a group composed of purple nonsulfur bacteria that demonstrate a diverse array of metabolisms, including photoautotrophy with carbon dioxide, chemoheterotrophy under aerobic conditions in the dark, and nitrogen fixation was selected for detection (Gest 1972; Kiley and Kaplan 1988).

Three fingerprint groups associated with the beta proteobacteria also met the primer design requirements. One group was affiliated with a soil clone collected from deep-sea sediments off Japan (Mitsutani et al. 1999). Another fingerprint group was closely associated with the genus *Massilia*, a group that contains phylotypes commonly found in soils (Wery et al. 2003; Padmanabhan et al. 2003). The final beta proteobacteria group that fits the criteria was closely related to *Herbaspirillum seropedicae*, an endosymbiotic diazotroph commonly isolated from crop plant root nodules (Baldani et al. 1986).

Within the high G+C Gram positive bacteria, there were three fingerprint groups to be analyzed. One group was composed of phylotypes that shared similar sequences with *Bacillus* sp. MK03, an agar-degrading soil bacterium that is 95% identical to *Bacillus batavienesis* (Suzuki and and Sawai 2001). Two other fingerprint groups were
associated with the genera *Geodermatophilus* and *Arthrobacter*, both of which are commonly found in soils around the world (Hagedorn and Holt 1975; Keddie et al. 1986; Urzi et al. 2003).

The remaining fingerprint groups that meet the chosen selection requirements were affiliated with phyla that are poorly characterized due to the lack of cultured species. These groups had relatively low abundances compared to the other selected fingerprint groups (<1%). One ARDRA fingerprint group that meets the requirements was distantly related to a *Planctomyceales* clone collected from a polychlorinated biphenyl-contaminated soil (Nogales et al. 2001). Two fingerprint groups were associated with *Acidobacteria*, one group with a northern Arizona desert soil clone (Kuske et al. 1997) and a second group, with an Australian desert clone (Holmes et al. 2000). A fingerprint group affiliated with a contaminated forest soil clone with the newly described *Gemmatimonadetes* group was also selected (Broft et al. 2002).

**Design and Specificity of 16S rRNA Gene Primers Specific for CAP LTER Phylogenotypes**

The specificity of candidate primers was validated through the PRIMROSE software (Ashhelford et al. 2000). Amplification of 16S rRNA gene sequences outside the target range was unlikely due to the presence of at least two mismatches with the primers. Primers were then tested against twenty-three 16S rRNA gene clones that represent the known diversity within the *Bacteria*. The results indicate that the primers were specific for only the CAP LTER phylogenotypes (not shown). Specific primer sequences and optimal annealing temperature are provided in Table 4.2.

**Distribution of Potentially Ubiquitous CAP LTER Phylogenotypes**

All extracted environmental DNA samples were initially tested for their ability to
support PCR using universal bacterial primers. Fourteen CAP LTER samples yielded no visible DNA band and were removed from the analysis. After PCR verification, DNA samples were subjected to PCR using the designed primer sets. Amplification of 167 CAP LTER environmental DNA samples were recorded according to the primer set used (Table 4.2). The proportion of samples with positive results was recorded for all samples tested as well as by land use type, and 95% confidence intervals were calculated to determine what fingerprint groups had upper bound error rates at 100%.

Ten fingerprint groups had presence proportions that reached 100%. Six of those groups were associated with *Sphingomonas* species. The other phylotypes included here belong to the *Methylobacterium*, *Bosea*, *Bacillus*, and *Geodermatophilus* genera. Across land use types, 17 of the 21 selected fingerprint groups (81%) were statistically ubiquitous in agricultural soils, more than any other type.

**Discussion**

Based upon on the results, ten phylotypes are ubiquitous to the Survey 200 soils and likely the entire CAP LTER site. In order for the observed ubiquitous distribution to occur, the taxa that are represented by these phylotypes must invoke one or more survival strategies that provide strong selective advantages for a variety of heterogeneous habitats. For the *Sphingomonas* species, this strategy is likely explained by the diverse metabolic capabilities towards natural and man-made organic compounds and resistance to antimicrobial agents (Mueller et al. 1990; Mueller et al. 1990; Feng et al. 1997; Lloyd-Jones and Lau 1997; Atkins 1999; Meyer et al. 1999; Momma et al. 1999; Bastiaens et al. 2000; Cassidy et al. 1999; Pinyakong et al. 2000; Sørenson et al. 2001). *Sphingomonas* species also demonstrate a presence at high cellular concentrations of $10^5$ to $10^6$ cells per
gram of soil (Leys et al. 2004), an indication of an *r*-specific reproductive strategy that may permit *Sphingomonas* populations to colonize soils quickly and aids in overall dispersal across the site. Similar physiological characteristics may explain the ubiquity of the *Methylobacterium* phylotype, including the ability to survive within anthropogenically disturbed environments (Garf and Bauer 1973; Hiraishi and Kaneko 1994; Hirainshi et al. 1995).

The fingerprint group that is affiliated with the genus *Bosea* employs selective strategies that likely differ from *Sphingomonas*. Members of *Bosea* are aerobic inorganic sulfur oxidizers that in some cases can grow autotrophically with carbon dioxide (Das et al. 1996; Stubner et al. 1998). Like *Sphingomonas*, *Bosea* species have also been isolated from wastewater facilities, indicating an ability to maintain a presence in a wide range of habitats (Ouattara et al. 2003; Chen et al. 2004). They have also been found to be resistant to predation by protists such as amoeba species, which may indicate a key selective advantage against other species (Greub and Raolut 2004). Furthermore, in rhizosphere environments, introduced *Bosea* species have been shown to maintain a selective advantage over indigenous communities because of the ability to degrade opines, condensates which include the pairing of an amino acid and a sugar that are produced by some wild plant species (Oger et al. 2004). The ability to degrade opine may serve as an alternative metabolic mechanism for *Bosea* species in agricultural or urban soils, which often have increased rhizosphere diversity. These characteristics combined may explain how this phylotype can be present in such diverse habitats at the CAP LTER.

The survival strategies of the *Bacillus* and *Geodermatophilus* phylotypes may involve spore-forming abilities. Spore formation is a convenient tool to provide spatial
and temporal escape from unfavorable conditions (Nicholson et al. 2000). *Bacillus* spores can germinate into viable cells after periods of tens to thousands of years or longer (Kennedy et al. 1994) and have been shown to be readily dispersed via wind currents (Robets and Cohan 1995). *Bacillus* species have also been found in rhizosphere communities (Seldin et al. 1984; Padney and Palini 1997), which may explain their presence in urban and agricultural soils with increased plant diversity. The ecological role that the *Geodermatophilus* group fills is relatively unknown other than the biodeteriation of stone surfaces and ability to survive in extremely arid environments (Eppard et al. 1996; Urzi et al. 2001; Navarro-Gonzales et al. 2003). The type species *Geodermatophilus obscurus* is characterized as employing a two-form life cycle and possessing the ability to produce motile spores, which may improve dispersion (Ludemann 1968).

It is also important to note that the 16S rRNA gene sequences of the individual fingerprint groups were not identical. Dissimilarity across the sequences was common and range from a one to fifteen base pair changes. Although some of this dissimilarity can be attributed to *Taq* polymerase errors, similar results in phylotype dissimilarity have been observed and verified in another molecular ecology study of an estuary sample (Acinas et al. 2004). The amount of 16S rRNA gene sequence heterogeneity within ARDRA fingerprint types therefore indicates that these representative populations within the CAP LTER are not clonal, but in a phylogenetic analysis would appear “bushy” when compared to each other. This is further evidence of the presence of ‘microdiverse clusters’ within bacterial populations that reflect the true diversity with certain species-level phylotypes. Whether not the slight variation in 16S rRNA genes actually
corresponds to changes in niche structure is unclear (Giovannoni 2003). If the “micorodiverse cluster” theory is true, this observed microdiversity may also explain the ubiquitous biogeography of these taxa at the CAP LTER as a result of their ability to efficiently evolve to fill previously unfilled, slightly dissimilar niches across spatial gradients in the environment.

The continued study of the ubiquitous taxa is of particular interest in the study of microbial ecology for several reasons. Ubiquitous taxa most likely could be engineered and used in bioremediation based upon their ecological resiliency. Additionally, primers designed specifically for the ubiquitous phylotypes observed in this study could be used on other soils outside of the CAP LTER to indicate whether these specific phylotype group sequences are endemic to the region. The ability to culture and isolate ubiquitous taxa and subsequent genome sequencing may unlock the true reasons behind their wide distribution, and may lead to the detection of novel natural compounds that could be produced to benefit mankind through genetic engineering.
CHAPTER 5:
CONCLUSIONS
Culture-independent comparisons of multiple environments have indicated that various factors can have a dramatic impact on prokaryotic communities in soils. This study examines the influence of anthropogenic factors on diversity, assemblage structural shifts, and biogeography of bacteria within the CAP LTER. However, certain bacterial phyotypes from divergent higher order phylogenetic groups are shown to be ubiquitous within the CAP LTER, suggesting a high degree of ecological tolerance within these taxa.

Chapter 2 introduced a comprehensive assay of the bacterial diversity of CAP LTER soils that are categorized by land use type. This analysis included the construction of twenty-three 16S rRNA gene clone libraries and generation of over 11,000 clones to permit robust statistical analyses. Subsequent ARDRA fingerprinting and partial 16S rRNA gene sequencing allowed for both individual site and land use type group comparisons using OTU as well as phylogenetic standards. Across land use types, significant changes in major bacterial group abundances are observed. Alpha proteobacteria dominate the undisturbed desert soils, while Firmicutes abundance is significantly higher in the agricultural soils. Members of Acidobacteria phylum are definitely present in higher proportions in urban and desert remnant comparisons. These fluxes in group abundance are an indication of the specialization of certain bacterial groups within land use types. Total species richness estimates ranged from $4.3 \times 10^3$ to $2.4 \times 10^5$ within the samples, providing further evidence of the vast species heterogeneity within soil environments. Overall bacterial diversity estimates were shown to significantly increase in agricultural and urban samples. Conversely, other continuous geochemical and physical variables did not sufficiently explain the variance in diversity estimates across the samples. The use of artificial irrigation methods helped explained
the diversity values within and across land use types. Collected 16S rRNA gene sequence libraries assembled according to land use type revealed that in major bacterial groups, desert remnant soils resemble nearby urban soils more than the undisturbed open desert samples outside the metropolitan area. Thus, land use change appears to not only affect bacterial communities directly within these sample areas, but also those undisturbed patches of natural desert in close proximity to the city.

Chapter 3 represents a concise phylogenetic census of the CAP LTER phylotypes enabling the examination of the total phylogenetic diversity of bacteria at the sites studied and the overall biogeography of individual phylotype groups. The majority of CAP LTER phylotype groups exhibited little biogeographical intolerance to land use type. This result is best explained by the physiological and evolutionary robustness among these phylotypes coupled with the heterogeneity of habitats within the bounds of samples that are categorized under the same land use type. However, there were isolated phylotype groups that were present in one or two land use types exclusively. The 16S rRNA gene sequences of these groups were highly diverged from existing phylotypes in the Genbank database, making any inference of the observed ecological intolerance difficult. In general, the majority of the CAP LTER phylotypes demonstrated high dissimilarity with Genbank sequences, indicating a potential endemicity of bacteria to the western United States region.

Chapter 4 details those CAP LTER phylotype groups that were well represented in the 16S rRNA gene clone libraries and the subsequent design of specific PCR primers to determine their presence in the other soil samples collected. Of the 21 phylotype groups selected for analysis, 10 were found to be present within all of the 167 CAP
LTER soils studied irrespective of land use type. Examination of the most closely related taxa to these phylotype groups revealed potential selective advantages to explain the observed continuous distribution.

Overall, the results presented here indicate that anthropogenic factors such as urbanization and agriculture are disrupting the diversity and spatial distribution of bacterial communities within CAP LTER soils. Considering the sprawl of urban environments throughout the world, the next step is to determine whether these ecological consequences are translating into deteriorating soil health, and the productivity of soils which are clearly dependent on microbial communities. The observed changes in bacterial diversity estimates in urban, remnant, and agricultural soils suggests that the increasing immigration and evolution of bacteria to fill new niches that arise from the altered chemical inputs, physical disturbance, and altered vegetation composition associated with land use change. However, the ability for many of the more dominant types to remain in these soils regardless of land use type designation is a testament to the immense buffering capacity that soils exhibit. Ultimately, the sustained presence of these more dominant members is beneficial because they likely represent the key components of maintaining a higher degree of soil health. However, the longer-term impacts of land use change may be much more significant, and their effects upon the bacterial community were not examined in this study; therefore further attention is warranted to this subject.


Bass-Becking LGM (1934) Geobiologie of inteiding tot de mileukunde. (van Stockum WP, Zoon NV) The Hague, Netherlands


Gause GF (1934) The struggle for existence. Baltimore (MD), Williams & Wilkins


Linton D, Dewhirst FE, Clewley JP, Owen RJ, Burnens AP, Stanley J (1994) Two types of 16S rRNA gene are found in *Campylobacter helveticus*: analysis, applications and characterization of the intervening sequence found in some strains. Microbiology 140:847–855


Mylvaganam S, Dennis PP (1992) Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaeabacterium Haloarcula marismortui. Genetics 130:399-410


enzymogenes based on phylogenetic analysis of 16S rRNA, fatty acid composition and phenotypic characteristics. J Appl Microbiol 94:1079-1086


Wang, Y, Zhang Z. (2000) Comparative sequence analyses reveal frequent occurrence of short segments containing an unusually high number of non-random base variation in bacterial rRNA genes. Microbiology (UK) 146:2845-2854


Williams AM, Rodriguez UM, Collins MD (1991) Intragenic relationships of Enterococci as determined by reverse transcritase sequencing of small-subunit rRNA. Res Microbiol 142:67-74


APPENDIX A – VARIABLES USED FOR DIVERSITY ANALYSES.

<table>
<thead>
<tr>
<th>SITE ID</th>
<th>land use type</th>
<th>watering type</th>
<th># of positive BLOG wells</th>
<th>Simpson</th>
<th>Shannon</th>
<th>Alpha</th>
<th>Simpson/ Proteins</th>
<th>Simpson/ Acids</th>
<th>Simpson/ Actinomycetes</th>
<th>Simpson/ Grams</th>
<th>Simpson/ Firm</th>
<th>Chao1</th>
<th>95% Chao1 CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC17</td>
<td>urban</td>
<td>hand/drip/overhead</td>
<td>21</td>
<td>40.62</td>
<td>4.6</td>
<td>122.75</td>
<td>14.16</td>
<td>50</td>
<td>15.89</td>
<td>26.05</td>
<td>5.06</td>
<td>385.27</td>
<td>197.94, 11567.41</td>
</tr>
<tr>
<td>AF14</td>
<td>open</td>
<td>none</td>
<td>16</td>
<td>20.56</td>
<td>4.08</td>
<td>96.36</td>
<td>22.13</td>
<td>21.17</td>
<td>24.79</td>
<td>26</td>
<td>11.37</td>
<td>408.1</td>
<td>184.37, 13306.18</td>
</tr>
<tr>
<td>AG20</td>
<td>open</td>
<td>none</td>
<td>11</td>
<td>32.72</td>
<td>4.36</td>
<td>110.69</td>
<td>14.35</td>
<td>39</td>
<td>17.14</td>
<td>51.66</td>
<td>10.02</td>
<td>472.34</td>
<td>188.09, 8044.86</td>
</tr>
<tr>
<td>F12</td>
<td>open</td>
<td>none</td>
<td>16</td>
<td>27.91</td>
<td>4.39</td>
<td>126.1</td>
<td>11.04</td>
<td>3.14</td>
<td>27.78</td>
<td>66.11</td>
<td>18.33</td>
<td>390.28</td>
<td>190.87, 14859.63</td>
</tr>
<tr>
<td>I11</td>
<td>open</td>
<td>none</td>
<td>4</td>
<td>38.15</td>
<td>4.62</td>
<td>172.91</td>
<td>8.22</td>
<td>25.3</td>
<td>36.12</td>
<td>21.25</td>
<td>42.85</td>
<td>502.03</td>
<td>250.65, 14581.57</td>
</tr>
<tr>
<td>I17</td>
<td>open</td>
<td>none</td>
<td>4</td>
<td>22.04</td>
<td>4.34</td>
<td>118.04</td>
<td>7.64</td>
<td>25</td>
<td>27.55</td>
<td>33.68</td>
<td>15</td>
<td>454.06</td>
<td>207.53, 14987.81</td>
</tr>
<tr>
<td>J7</td>
<td>open</td>
<td>none</td>
<td>7</td>
<td>63.05</td>
<td>4.72</td>
<td>145.99</td>
<td>14.06</td>
<td>70</td>
<td>19.72</td>
<td>39</td>
<td>29</td>
<td>478.37</td>
<td>206.66, 18060.06</td>
</tr>
<tr>
<td>R9</td>
<td>open</td>
<td>none</td>
<td>10</td>
<td>10.65</td>
<td>3.64</td>
<td>78</td>
<td>2.59</td>
<td>25.3</td>
<td>*</td>
<td>8.88</td>
<td>25.5</td>
<td>460.64</td>
<td>169.4, 81134.76</td>
</tr>
<tr>
<td>L14</td>
<td>mixed</td>
<td>none</td>
<td>6</td>
<td>41.57</td>
<td>4.47</td>
<td>116.38</td>
<td>7.84</td>
<td>4.55</td>
<td>15.5</td>
<td>16.6</td>
<td>16.6</td>
<td>455.4</td>
<td>200.12, 15456.45</td>
</tr>
<tr>
<td>M9</td>
<td>urban</td>
<td>none</td>
<td>16</td>
<td>18.21</td>
<td>4.18</td>
<td>100.56</td>
<td>2.88</td>
<td>3.1</td>
<td>*</td>
<td>72.5</td>
<td>5.37</td>
<td>459.27</td>
<td>195.46, 15365.35</td>
</tr>
<tr>
<td>Q12</td>
<td>ag</td>
<td>flood</td>
<td>20</td>
<td>31.68</td>
<td>4.35</td>
<td>110.58</td>
<td>13.78</td>
<td>47.35</td>
<td>16.71</td>
<td>38.5</td>
<td>5</td>
<td>454.63</td>
<td>194.21, 15630.15</td>
</tr>
<tr>
<td>R18</td>
<td>ag</td>
<td>flood</td>
<td>9</td>
<td>68.94</td>
<td>4.85</td>
<td>171.58</td>
<td>3.41</td>
<td>38.04</td>
<td>60</td>
<td>57.27</td>
<td>6.97</td>
<td>654.42</td>
<td>233.84, 49290.64</td>
</tr>
<tr>
<td>U20</td>
<td>open</td>
<td>none</td>
<td>5</td>
<td>23.9</td>
<td>4.23</td>
<td>104.41</td>
<td>2.81</td>
<td>50.32</td>
<td>42.16</td>
<td>46.2</td>
<td>14.09</td>
<td>388.81</td>
<td>186.16, 12031.12</td>
</tr>
<tr>
<td>U21</td>
<td>remnant</td>
<td>none</td>
<td>7</td>
<td>21.18</td>
<td>4.22</td>
<td>115.51</td>
<td>3.07</td>
<td>27.04</td>
<td>84.33</td>
<td>22.5</td>
<td>9.73</td>
<td>515.32</td>
<td>182.71, 18069.18</td>
</tr>
<tr>
<td>V7</td>
<td>open</td>
<td>none</td>
<td>18</td>
<td>18.3</td>
<td>3.89</td>
<td>84.29</td>
<td>10.4</td>
<td>35</td>
<td>91</td>
<td>23.8</td>
<td>8.57</td>
<td>530.69</td>
<td>172.29, 22813.85</td>
</tr>
<tr>
<td>V11</td>
<td>remnant</td>
<td>none</td>
<td>22</td>
<td>55.37</td>
<td>4.74</td>
<td>153.29</td>
<td>13.34</td>
<td>14.56</td>
<td>42.75</td>
<td>22.09</td>
<td>6.14</td>
<td>579.94</td>
<td>230, 20865.88</td>
</tr>
<tr>
<td>V13</td>
<td>urban</td>
<td>none</td>
<td>27</td>
<td>39.97</td>
<td>4.48</td>
<td>119.42</td>
<td>19.68</td>
<td>52.5</td>
<td>91</td>
<td>7.45</td>
<td>6.14</td>
<td>528.8</td>
<td>206.53, 14939.44</td>
</tr>
<tr>
<td>V18</td>
<td>ag</td>
<td>flood</td>
<td>15</td>
<td>91.59</td>
<td>4.88</td>
<td>157.08</td>
<td>3.18</td>
<td>29.22</td>
<td>20.78</td>
<td>21.52</td>
<td>5.05</td>
<td>526.9</td>
<td>227.9, 15220.24</td>
</tr>
<tr>
<td>W13</td>
<td>urban</td>
<td>hand/drip/overhead</td>
<td>5</td>
<td>51.75</td>
<td>4.57</td>
<td>117.78</td>
<td>3.42</td>
<td>27.02</td>
<td>8.25</td>
<td>14.8</td>
<td>13</td>
<td>420.05</td>
<td>205.78, 13031.09</td>
</tr>
<tr>
<td>X13A</td>
<td>urban</td>
<td>hand/drip/overhead</td>
<td>23</td>
<td>91.59</td>
<td>4.92</td>
<td>205.98</td>
<td>22.59</td>
<td>13.26</td>
<td>45.11</td>
<td>74.37</td>
<td>15</td>
<td>574.23</td>
<td>208.94, 18268.32</td>
</tr>
<tr>
<td>X13B</td>
<td>urban</td>
<td>hand/drip/overhead</td>
<td>23</td>
<td>63.79</td>
<td>4.82</td>
<td>180.83</td>
<td>3.42</td>
<td>52.56</td>
<td>46.87</td>
<td>75.6</td>
<td>13.14</td>
<td>690.13</td>
<td>243.83, 27019.03</td>
</tr>
<tr>
<td>X14</td>
<td>remnant</td>
<td>none</td>
<td>17</td>
<td>73.29</td>
<td>4.92</td>
<td>187.9</td>
<td>10.7</td>
<td>19.32</td>
<td>46.86</td>
<td>22.09</td>
<td>6.14</td>
<td>630.01</td>
<td>247.21, 22728.18</td>
</tr>
</tbody>
</table>

*a Asterisks represent samples with only singleton OTUs based upon ARDRA fingerprint data and therefore were excluded from analysis*
### APPENDIX B – OTHER VARIABLES USED FOR STATISTICAL ANALYSES.

<table>
<thead>
<tr>
<th>SITE ID</th>
<th>lat/long type</th>
<th>Elevation (Meters)</th>
<th><em>d</em></th>
<th>slope</th>
<th><em>p</em></th>
<th>Imp*</th>
<th>% Vegetation Cover</th>
<th>%clay</th>
<th>%totalN</th>
<th>%NO3-N (mg N per kg)</th>
<th>%NHI3-N (mg N per kg)</th>
<th>%totalC</th>
<th>%organicC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC17</td>
<td>urban</td>
<td>20</td>
<td>396</td>
<td>30579.74547</td>
<td>0</td>
<td>1706.41337</td>
<td>33.9</td>
<td>7.12571167</td>
<td>16.5826935</td>
<td>0.106</td>
<td>1.53476599</td>
<td>2.70213626</td>
<td>1.702</td>
</tr>
<tr>
<td>AF14</td>
<td>open</td>
<td>0</td>
<td>482</td>
<td>45335.9745</td>
<td>0</td>
<td>0.04515903</td>
<td>0</td>
<td>65</td>
<td>4.9</td>
<td>1.800</td>
<td>2.026</td>
<td>0.055</td>
<td>0.605</td>
</tr>
<tr>
<td>AG20</td>
<td>open</td>
<td>0</td>
<td>458</td>
<td>47331.81008</td>
<td>0</td>
<td>90.70827463</td>
<td>0</td>
<td>100</td>
<td>9.6</td>
<td>2.548</td>
<td>11.473</td>
<td>0.066</td>
<td>0.349</td>
</tr>
<tr>
<td>F12</td>
<td>open</td>
<td>0</td>
<td>401</td>
<td>61996.1073</td>
<td>0</td>
<td>3.96469</td>
<td>0</td>
<td>33.2464384</td>
<td>5.940428245</td>
<td>0.062</td>
<td>1.96429379</td>
<td>18.14386589</td>
<td>0.698</td>
</tr>
<tr>
<td>H1</td>
<td>open</td>
<td>0</td>
<td>517</td>
<td>55420.13073</td>
<td>0</td>
<td>3.22796</td>
<td>0</td>
<td>49.60174894</td>
<td>8.872753223</td>
<td>0.287</td>
<td>3.671395718</td>
<td>0.4566291</td>
<td>0.2</td>
</tr>
<tr>
<td>H7</td>
<td>open</td>
<td>0</td>
<td>335</td>
<td>46740.88635</td>
<td>0</td>
<td>3.96469</td>
<td>0</td>
<td>37.57306826</td>
<td>9.991022025</td>
<td>0.045</td>
<td>2.306747396</td>
<td>3.062544294</td>
<td>0.625</td>
</tr>
<tr>
<td>J7</td>
<td>open</td>
<td>0</td>
<td>547</td>
<td>60041.94638</td>
<td>0</td>
<td>20.96659</td>
<td>0</td>
<td>68.59813373</td>
<td>17.65247472</td>
<td>0.122</td>
<td>4.1190233</td>
<td>6.518358945</td>
<td>0.25</td>
</tr>
<tr>
<td>K9</td>
<td>open</td>
<td>0</td>
<td>480</td>
<td>50794.0104</td>
<td>0</td>
<td>20.96659</td>
<td>0</td>
<td>8.498876525</td>
<td>12.83900335</td>
<td>0.073</td>
<td>1.74745993</td>
<td>5.566546271</td>
<td>0.446</td>
</tr>
<tr>
<td>L14</td>
<td>mixed</td>
<td>0</td>
<td>375</td>
<td>38540.83664</td>
<td>16</td>
<td>47.94878</td>
<td>0</td>
<td>16.94858475</td>
<td>4.739319074</td>
<td>0.029</td>
<td>2.976173258</td>
<td>2.204757112</td>
<td>0.363</td>
</tr>
<tr>
<td>M9</td>
<td>urban</td>
<td>5</td>
<td>440</td>
<td>45493.86038</td>
<td>0</td>
<td>11.41012</td>
<td>0</td>
<td>7.733794086</td>
<td>11.6671132</td>
<td>0.039</td>
<td>1.4575919</td>
<td>1.833419689</td>
<td>0.28</td>
</tr>
<tr>
<td>Q12</td>
<td>agricultural</td>
<td>88</td>
<td>357</td>
<td>23437.9491</td>
<td>0</td>
<td>200.30233</td>
<td>0</td>
<td>0</td>
<td>20.79437792</td>
<td>0.085</td>
<td>3.37744342</td>
<td>19.45108365</td>
<td>1.068</td>
</tr>
<tr>
<td>R1B</td>
<td>agricultural</td>
<td>100</td>
<td>309</td>
<td>12109.34387</td>
<td>0</td>
<td>52.01702</td>
<td>0</td>
<td>0</td>
<td>13.7775529</td>
<td>0.076</td>
<td>2.429165529</td>
<td>102.8736661</td>
<td>0.956</td>
</tr>
<tr>
<td>U20</td>
<td>open</td>
<td>0</td>
<td>740</td>
<td>12581.64381</td>
<td>10</td>
<td>43.95941</td>
<td>0</td>
<td>51.766292</td>
<td>10.1269916</td>
<td>0.098</td>
<td>2.806180332</td>
<td>33.2124099</td>
<td>0.964</td>
</tr>
<tr>
<td>U21</td>
<td>remnant</td>
<td>0</td>
<td>399</td>
<td>15726.23135</td>
<td>13</td>
<td>710.69014</td>
<td>0</td>
<td>25.82799955</td>
<td>9.671157997</td>
<td>0.021</td>
<td>1.224842333</td>
<td>1.519804364</td>
<td>0.427</td>
</tr>
<tr>
<td>V7</td>
<td>open</td>
<td>0</td>
<td>547</td>
<td>38721.26264</td>
<td>0</td>
<td>8.83609</td>
<td>0</td>
<td>46.01808262</td>
<td>10.5583187</td>
<td>0.029</td>
<td>1.336917493</td>
<td>9.512132031</td>
<td>0.302</td>
</tr>
<tr>
<td>V11</td>
<td>remnant</td>
<td>0</td>
<td>475</td>
<td>23540.01124</td>
<td>18</td>
<td>973.841</td>
<td>0</td>
<td>41.78640037</td>
<td>0</td>
<td>0.04</td>
<td>1.171124753</td>
<td>2.059741341</td>
<td>0.38</td>
</tr>
<tr>
<td>V13</td>
<td>urban</td>
<td>30</td>
<td>441</td>
<td>15652.36369</td>
<td>0</td>
<td>1613.43284</td>
<td>100</td>
<td>0</td>
<td>12.81286477</td>
<td>0.086</td>
<td>4.344382817</td>
<td>0.872609653</td>
<td>1.479</td>
</tr>
<tr>
<td>V14</td>
<td>agricultural</td>
<td>100</td>
<td>347</td>
<td>7921.98134</td>
<td>0</td>
<td>272.32716</td>
<td>0</td>
<td>0</td>
<td>12.96405136</td>
<td>0.066</td>
<td>2.600610077</td>
<td>42.8104607</td>
<td>0.64</td>
</tr>
<tr>
<td>W15</td>
<td>urban</td>
<td>10</td>
<td>438</td>
<td>15391.27933</td>
<td>17</td>
<td>1046.69088</td>
<td>16.0111111</td>
<td>42.91717043</td>
<td>12.86372692</td>
<td>0.018</td>
<td>3.188363075</td>
<td>1.775508202</td>
<td>1.733</td>
</tr>
<tr>
<td>X15A</td>
<td>urban</td>
<td>45</td>
<td>412</td>
<td>19158.36593</td>
<td>0</td>
<td>1396.4616</td>
<td>32.76666667</td>
<td>16.02591016</td>
<td>12.94184841</td>
<td>0.046</td>
<td>1.572771497</td>
<td>8.117485428</td>
<td>1.315</td>
</tr>
<tr>
<td>X15B</td>
<td>urban</td>
<td>45</td>
<td>412</td>
<td>19158.36593</td>
<td>0</td>
<td>1396.4616</td>
<td>32.76666667</td>
<td>16.02591016</td>
<td>12.94184841</td>
<td>0.046</td>
<td>1.572771497</td>
<td>8.117485428</td>
<td>1.315</td>
</tr>
<tr>
<td>X14</td>
<td>remnant</td>
<td>0</td>
<td>645</td>
<td>15500.38466</td>
<td>30</td>
<td>220.94085</td>
<td>0</td>
<td>18.86009704</td>
<td>3.66305997</td>
<td>0.25</td>
<td>1.268174285</td>
<td>4.835211318</td>
<td>0.739</td>
</tr>
<tr>
<td>X15</td>
<td>remnant</td>
<td>0</td>
<td>650</td>
<td>12675.94738</td>
<td>27</td>
<td>239.03904</td>
<td>0</td>
<td>62.48007339</td>
<td>5.081526425</td>
<td>0.10225</td>
<td>2.517402643</td>
<td>15.28685337</td>
<td>0.96725</td>
</tr>
</tbody>
</table>

*Years since land use change
*Distance from urban center (m)
*Population density (people per km²)
*Percent of sample impervious to soil
VITA

Brian Anthony Rash was born to Curtis Anthony and Regina Sue Rash on July 26th, 1978 in Dayton, Ohio. In 1996, he graduated with honors from Northmont High School in Clayton, Ohio, after which time he moved to Bowling Green State University to pursue undergraduate study. Brian received a Bachelor of Science degree in microbiology in 2000, and relocated to Baton Rouge, Louisiana to begin graduate school at Louisiana State University in the laboratory of Dr. Fred A. Rainey. Brian will receive the degree of Doctor of Philosophy at the December 2004 commencement.