Bacterial Community Succession during Soil and Ecosystem Development

Shankar Ganapathi Shanmugam

Follow this and additional works at: https://scholarsjunction.msstate.edu/td

Recommended Citation
https://scholarsjunction.msstate.edu/td/842

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.
Bacterial community succession during soil and ecosystem development

By

Shankar Ganapathi Shanmugam

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Agronomy
in the Department of Plant and Soil Sciences

Mississippi State, Mississippi

May 2013
Bacterial community succession during soil and ecosystem development

By

Shankar Ganapathi Shanmugam

Approved:

Daniel G. Peterson  
Professor of Plant and Soil Sciences  
(Committee Member)

Zenaida Magbanua  
Assistant Research Professor of Plant and Soil Sciences  
(Committee Member)

Frank B. Matta  
Professor of Plant and Soil Sciences  
(Committee Member)

Karen S. McNeal  
Associate Professor of Geosciences  
(Committee Member)

William L. Kingery  
Professor and Graduate Coordinator of Plant and Soil Sciences  
(Committee Member)

J. Mike Phillips  
Professor and Department Head of Plant and Soil Sciences

George Hopper  
Dean, College of Agriculture and Life Sciences
Organism succession during ecosystem development has been well studied for aboveground plant communities while the associated pattern of change in microbial communities remains largely unknown. A study was conducted along developmental sand-dune chronosequences bordering Lake Michigan at Wilderness State Park and Altamaha river valley of southeast Georgia with the hypothesis that soil bacterial communities will follow a pattern of change that is associated with soil, plant, and ecosystem development. This study site included 5 replicate sites along 14 dunes ranging in age from 105 to 500,000 years since deposition. The microbial composition and diversity in the soil was studied using bacterial tag-encoded FLX amplicon pyrosequencing of the 16S rRNA gene. As hypothesized, Bray-Curtis ordination indicated that bacterial community assembly changed along the developmental gradient at both sites. However, there was no seasonal effect at Michigan sites despite likely differences in plant carbon inputs. At the Michigan site, soil Ca, Mg levels and pH showed a significant log-linear correlation with soil development ($r = 0.83, 0.84$ and $0.81$, respectively). Bacterial diversity represented by Simpson’s reciprocal index (Simpson’s
I/D) showed a steady decline from the youngest to the oldest dunes with the largest decline (212 to 58) during the initial stages of soil development (105 to 450 years). The change in plant species abundance was higher in the youngest sites than the older sites. This change was significantly correlated with the change in microbial community distribution (p < 0.0001; r = 0.56). Similarly, at Georgia sites, soil development showed significant log-linear correlation with soil base cations (Ca and Mg) (r = 0.93 and 0.95). However, diversity indices and PLFA failed to show any particular change in trend across the developing chronosequences. When the results from both sites were used to study bacterial spatial patterns, local geochemical features were found to be a dominant factor in driving bacterial community structure, while geographic distance as a single factor could contribute to some community variation at a scale (50 – 1700 km). The results suggest that soil nutrients and plant community could be a strong driving force in shaping microbial community assembly across a developing soil ecosystem.

Key Words: Bacterial diversity, Ecosystem development, Pedogenesis, Soil nutrients and pH, Season, Vegetative succession.
DEDICATION

I would like to dedicate this piece of work to my family for their moral support and words of encouragements. I dedicate this work to my grandfather, freedom fighter, Mr.T.S.Arumugam Pillai. A special feeling of gratitude to my loving parents and brother.

I dedicate this work to my wife and my daughter for being there throughout the doctoral program and are very special.
ACKNOWLEDGEMENTS

I am heartily thankful to all my committee members: Dr. Daniel Peterson, Dr. Zenaida Magbanua, Dr. Frank Matta, Dr. Karen McNeal and Dr. William Kingery, for their insightful comments at different stages of my research.
### TABLE OF CONTENTS

**DEDICATION**........................................................................................................................................ ii  
**ACKNOWLEDGEMENTS**...................................................................................................................... iii  
**LIST OF TABLES**.................................................................................................................................. vii  
**LIST OF FIGURES** ................................................................................................................................ ix  
**CHAPTER**

I. **INTRODUCTION AND LITERATURE REVIEW** ............................................................. 1  
   - Microorganisms in soil ecosystem ...................................................................................... 1  
   - Soil chronosequences and nutrient dynamics ................................................................. 2  
     - Carbon dynamics in chronosequence studies ................................................................ 3  
     - Nitrogen dynamics in chronosequence studies ............................................................. 3  
     - Phosphorus dynamics in chronosequence studies ....................................................... 4  
   - Bacterial diversity in soil .................................................................................................... 4  
     - Abiotic factors .................................................................................................................. 4  
     - Biotic factors (Plant) ....................................................................................................... 5  
   - Bacterial species problem ............................................................................................... 6  
     - Defining bacterial species ............................................................................................. 7  
   - 16S rRNA gene ................................................................................................................. 8  
   - Genomics, metagenomics and next-generation sequencing ............................................ 10  
   - Relevance of research ...................................................................................................... 13  

II. **MATERIALS AND METHODS** ...................................................................................... 15  
   - Soil sampling and preparation ....................................................................................... 15  
   - Vegetative sampling ......................................................................................................... 15  
   - Soil characteristics .......................................................................................................... 16  
   - Phospholipid fatty acid analysis ..................................................................................... 17  
   - DNA extraction and pyrosequencing ............................................................................. 19  
   - Processing of 16S rRNA gene data ............................................................................... 20  
   - Statistical and multivariate analyses .............................................................................. 22  

III. **THE STUDY OF BACTERIAL COMMUNITY DIVERSITY AND SUCCESION ACROSS THE MICHIGAN CHRONOSEQUENCE** ........................................................................... 28
Site description........................................................................................................28
Soil and vegetative sampling..................................................................................29
Objectives and hypothesis......................................................................................30
  Objectives ...........................................................................................................30
  Hypothesis..........................................................................................................30
Results......................................................................................................................30
  Changes in soil characteristics during ecosystem development .......................30
  16S rRNA data...................................................................................................31
  Bacterial diversity indices....................................................................................32
  Microbial community structure and ecosystem properties...............................32
  Change in plant community composition across the chronosequences..............34
  Effect of seasonal variations on bacterial communities....................................35
  The change in microbial PLFA............................................................................36
Discussion...............................................................................................................37

IV. THE STUDY OF BACTERIAL COMMUNITY DIVERSITY AND SUCCESSION IN DEVELOPING DUNES OF SOUTHEASTERN GEORGIA .............................................................66

  Site description..................................................................................................66
  Objectives and hypothesis..................................................................................67
    Objectives .......................................................................................................67
    Hypothesis......................................................................................................67
  Results...............................................................................................................67
    Changes in soil characteristics during ecosystem development ....................67
    16S rRNA data...............................................................................................68
    Bacterial diversity indices..............................................................................68
    Microbial community structure and ecosystem properties..........................69
    Change in plant community composition across the chronosequences..........70
    Phylogenetic affiliation of the 16S rRNA genes............................................70
    The change in microbial PLFA.......................................................................71
  Discussion..........................................................................................................71

V. BIOGEOGRAPHY OF BACTERIAL COMMUNITY DIVERSITY AND SUCCESSION ..............................................................................................................................89

  Introduction.......................................................................................................89
  Objectives and hypothesis..................................................................................91
    Objectives ......................................................................................................91
    Hypothesis.....................................................................................................91
  Results...............................................................................................................91
    Microbial phylogenetic distribution.............................................................91
    Change in bacterial community structure and ecosystem properties............92
    Bacterial diversity indices.............................................................................93
Bacterial biogeography ................................................................. 94
Discussion ..................................................................................... 95

VI. GENERAL CONCLUSION AND SIGNIFICANCE OF RESEARCH ...... 108

Future directions .............................................................................. 111

REFERENCES .................................................................................. 113
LIST OF TABLES

3.1  Mehlich-3 extractable soil cations and selected soil properties from the mineral soil across the chronosequence. .................................................................47

3.2  Diversity indices of the 16S rRNA gene sequences........................................48

3.3  Correlations between the bacterial community distribution and selected environmental characteristics.............................................................49

3.4  Sum of the PLFAs of different taxonomic groups and ratios across the soil ages. ...........................................................................................................50

3.5  Diversity indices of the 16S rRNA gene sequences........................................51

4.1  Mehlich-3 extractable soil cations and selected soil properties from the mineral soil across the chronosequence. ....................................................75

4.2  Diversity indices of the 16S rRNA gene sequences........................................76

4.3  Total percent ground and canopy cover of different plant species across the chronosequences..................................................................................77

4.4  Sum of the PLFAs of different taxonomic groups and ratios across the soil ages ........................................................................................................78

4.5  Correlations between the bacterial community distribution and selected environmental characteristics across the Georgia chronosequences.................................................................79

5.1  Mehlich-3 extractable soil cations and selected soil properties from the mineral soil across the chronosequences....................................................98

5.2  Description of sampling sites ...........................................................................99

5.3  Mantel test correlations between the bacterial community distribution and selected environmental characteristics..........................................................99

5.4  Diversity indices of the 16S rRNA gene sequences........................................100
5.5 Pearson correlations between the bacterial community distribution and selected soil characteristics. .................................................................101
LIST OF FIGURES

2.1 Aerial photography of the beach-ridge complex at Wilderness State Park in Michigan. ........................................................................................................25

2.2 Aerial photography (University of Georgia Map Library) of the riverine dunes at Georgia showing the five sampling sites along the eastern coast of Altamaha and Ohooppe river basins. ...............................26

2.3 Overall steps involved in creating an emulsion amplified 16S rRNA gene library from the soil samples before pyrosequencing. ..................27

3.1 Bray –Curtis ordination plot showing the relationship between soil ecosystem development and bacterial community composition.................52

3.2 Bray –Curtis ordination plot showing (a) bacterial community composition and structure. (b) Change in plant species community. ..........................................................53

3.3 Relationship between percentage relative abundance of nine individual bacterial phyla across the chronosequence during ecosystem development ....................................................................................54

3.4 Relationship between the relative abundance of bacterial phyla across the chronosequences........................................................................55

3.5 Relationship between closest cultural representatives from the most abundant OTU based on RDP agent analysis........................................56

3.6 Rarefaction curves of the 16S rRNA gene libraries........................................57

3.7 A) Hierarchical Cluster Analysis of the most abundant OTU B) Hierarchical Cluster Analysis of the total percentage-cover of 13 different plant species.................................................................58

3.8 Plant succession shown as percentage change across soil development........59

3.9 Differences in the structure of microbial community associated with summer and winter samples from Michigan chronosequences using Bray-Curtis ordination analysis of the PLFA profiles..............60
3.10 Bray-Curtis ordination plot of PLFA profiles from different seasons across the soil ages. a) Summer PLFAs b) Winter PLFAs ..........................61
3.11 Sum of bacterial and fungal PLFA concentrations across soil development. ..................................................................................62
3.12 Change in fungal to bacterial PLFA ratios across soil development. ...............63
3.13 Change in Saturated/ Monosaturated PLFAs ratios across soil development. ..................................................................................64
3.14 Canonical correspondence analysis plot showing the relationship between the OTU distribution and environmental variable. ..............65
4.1 Bray–Curtis ordination plot showing relationship between soil ecosystem development and bacterial community composition and structure ...............................................................80
4.2 Bray-Curtis ordination showing relationship between the soil ecosystem development and change in percent cover of 20 plant species...........81
4.3 Relationship between percentage relative abundance of nine individual bacterial phyla across the chronosequence during ecosystem development ..................................................................................82
4.4 Relationship between the relative abundance of bacterial phyla across the chronosequences ..................................................................................83
4.5 Rarefaction curves of the 16S rRNA gene libraries ........................................84
4.6 Differences in the structure of microbial community from Georgia chronosequences using Bray-Curtis ordination analysis of the molar percentages of 39 PLFA profiles ..................................................................................85
4.7 Total PLFA and sum of bacterial and fungal PLFA concentrations across soil development. ..................................................................................86
4.8 Change in a) Fungal/Bacterial = fungal to bacterial PLFA ratios across the ages. b) S/M ratio= Saturated/ Monosaturated PLFAs. c) cyc/precursor ratio= changes in cyclopropyl and their precursor PLFA ratio. ..................................................................................87
4.9 Hierarchical Cluster Analysis of the most abundant OTU ...............................88
5.1 Bray–Curtis ordination plot showing relationship between soil ecosystem development and bacterial community composition and structure at both the sites (Michigan and Georgia). .........................102
5.2 Canonical correspondence analysis plot showing the relationship between the OTU distribution and environmental variable. ......................103

5.3 Hierarchical Cluster Analysis of the most abundant OTU.................................104

5.4 Relationship between the relative abundance of bacterial phyla across Michigan and Georgia chronosequences..........................................................105

5.5 Relationship between percentage relative abundance of nine individual bacterial phyla across the both the chronosequences during ecosystem development........................................................................106

5.6 Mantel correlogram for the spatial autocorrelation analysis of bacterial OTU distribution across all ages. .................................................................107
CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

Microorganisms in soil ecosystem

Microorganisms are key drivers of earth’s biogeochemical cycles and are considered the most abundant and diverse organisms on earth (Torsvik et al., 2002; Falkowski et al., 2008). Soil is considered to harbor the most diverse microbial communities than any other environment and a single gram of soil has been estimated to contain thousands of different bacterial species (Roesch et al., 2007). It has been estimated that only 1% of the bacteria are cultivable and at least 99% of bacteria observed under a microscope are not cultured by common laboratory techniques (Ovreas et al., 1998). The diversity of prokaryotes in soil is enormous and a vast majority of this diversity remains uncharacterized. Our knowledge of soil microbial diversity is limited in part by our inability to study various factors that regulate the diversity (Torsvik et al., 1996; Fierer et al., 2006). Recently cultivation-independent surveys of microbial communities in the environment using molecular techniques have greatly increased our understanding of prokaryotic diversity and have led to the discovery of considerable sequence diversity in soil (Zhou et al., 2002; Gans et al., 2005). Many novel taxonomic groups were identified using phylogenetic comparisons of the rRNA genes which were made possible only using culture independent surveys. In just a decade after introducing culture-independent surveys, at least 40 major bacterial divisions were identified in soil.
environmental samples which were previously not found in soil ecosystem (Konstantinidis et al., 2005; Konstantinidis et al., 2007). It also helped to reveal distribution patterns of the microbial diversity which were unknown before the advent of cultivation-independent methods. For example, *Archaea*, which were previously thought to be inhabitants of only extreme environments (hot springs, deep-thermal vents) were discovered from widespread habitats such as garden soils, forest soil and freshwater lakes (Xu, 2006).

**Soil chronosequences and nutrient dynamics**

Soil chronosequences are promising ecosystems to understand the trajectory of microbial community composition change in relation to sequentially changing soil factors. Soil chronosequences are excellent tools for evaluating soil genesis, nutrient availability and biologically significant elements and their implications for ecosystem functioning (Lichter, 1998b). Soil scientists categorize soil nutrients into two major categories based on their source as atmospherically-derived and rock-derived. Atmospherically- derived elements (carbon (C) and nitrogen (N)) should be biologically fixed or deposited directly from the atmosphere while rock-derived minerals (such as Ca, Mg, K and P) enter the ecosystem thorough chemical weathering (Chadwick et al., 1999). Walkers and Syers (1976) developed a conceptual model that explained the relationship between these two categories of nutrients from a soil development perspective. According to them the newly laid down (volcanic eruptions, eolian sand dunes) or newly exposed (glacial recessions) chronosequences are often rich in rock derived minerals and generally lack atmospherically derived elements.
**Carbon dynamics in chronosequence studies**

As mentioned above during initial soil development process there is a deficiency of the available carbon in the system. The system depends on the initial microbial colonizers (non-symbiotic nitrogen fixers) which use atmospheric CO$_2$ as carbon source and fix atmospheric nitrogen. This soil nitrogen would eventually help the initial plant colonizers to establish. These plants would decompose and provide carbon input to the system and as the soil develops there would be an accumulation of carbon with time (Allison et al., 2007). Primary succession along the chronosequences is characterized by initial colonization of communities of perennials and grasses, followed by softwood trees and shrubs, and finally hardwood trees and mixed pine forests (Lichter, 1998a). Thus, carbon inputs from plants combined with soil development help in accumulation of carbon for microbial growth (Tarlera et al., 2008). Lichter (1998b) studied 13 dune ridges along the sand-dune chronosequences bordering Lake Michigan and showed that total ecosystem carbon increased logarithmically along the chronosequence and attained steady-state once soil development attained equilibrium.

**Nitrogen dynamics in chronosequence studies**

Nitrogen is another element that is initially limited and becomes biologically available when the soil matures (Parfitt et al., 2005). Since nitrogen is absent in most exposed primary substrates, it should be obtained by biological nitrogen fixation or atmospheric deposition. As a result, during formative phases of soil development, soil fertility should be low with nitrogen as a limiting factor, a condition which favors N-fixing micro-organisms. Once the N-fixers inhabit the site, they would have a competitive advantage (Vitousek et al., 1997). Initially, with photoautotrophs (eg. *Cyanobacteria*) and
later with symbionts (e.g., *Rhizobium*) in association with N-fixing plants, nitrogen accumulates slowly as the soil ages. Eventually the biological available nitrogen will accumulate and will be in equilibrium with the rock-derived nutrients especially phosphorus (P) (Vitousek et al., 1997). Below this N-P equilibrium, nitrogen accumulates as the N-fixers dominate the system, while above this equilibrium, any excess nitrogen accumulated will get leached out of the system (Walker et al., 1976). Beyond this point non N-fixers tend to dominate the ecosystem and they obtain their soil nitrogen from heterotrophic nitrogen-mineralization of the organic nitrogen by the soil microbes. They liberate ammonium (NH$_4^+$) from soil organic matter decomposition.

**Phosphorus dynamics in chronosequence studies**

Phosphorus tends to become biologically available as the primary substrate is exposed due to the weathering process (Chadwick et al., 1999; Crews et al., 1995). As the soil ages the weatherable, rock-derived stock is depleted along the chronosequence and at later stages P becomes the limited nutrient (Walker et al., 1976). But unlike the atmospherically –derived nutrients, P becomes irreversibly depleted and as an ecosystem develops the available P gets converted to recalcitrant organic forms and the system attains a “terminal steady state” (Vitousek et al., 1997; Richardson et al., 2004).

**Bacterial diversity in soil**

**Abiotic factors**

Soil microorganisms are major players in regulating numerous biological processes in soil (Kuramae et al., 2011). Soils are highly heterogeneous both physico-chemically and biologically, thereby encompassing a wide range of niches available to
sustain microbial diversity (Reynolds et al., 2003). Numerous studies have shown the influence of both abiotic and biotic factors on microbial community structure (Brockett et al., 2011; Kuramae et al., 2011). Due to the fact that many factors are responsible for the composition of soil microbial community, differences in these factors may cause major differences in the distribution of microbial species. Soil type has been shown to be a primary factor determining microbial species composition (Bossio et al., 1998; Girvan et al., 2003). Bossio et al. (1998) found that soil type was the major determinant driving bacterial community structure in fields maintained under different management systems (organic and conventional systems). In a similar study, Girvan et al. (2003) showed that soil type was the key factor determining bacterial community composition in two cropping systems that had two major soil types. Other environmental factors, such as quantity and quality of available soil carbon (Blaalid et al., 2011), soil nitrogen (Eaton et al., 2011), soil water content (Schaaf et al., 2011), soil texture (Monroy et al., 2011) and land-use history (Lau et al., 2011), either alone or in combinations, have been shown to be significant contributors in shaping the soil microbial community structure. Soil pH has been often shown as significant contributor with significant correlations with soil microbial community structure, particularly bacteria. For example, recent studies across huge spatial scales have shown that soil pH could be a key driver that influences the bacterial community composition (Lauber et al., 2009; Rousk et al., 2010b).

**Biotic factors (Plant)**

Soil microbial communities play a crucial role in determining productivity, diversity and dynamics of above-ground plant communities (Reynolds et al., 2003). In a plant-soil subsystem, plants are major source of energy input in the form of litter deposits...
and root exudates that provide the carbon required for the heterotrophic microbial 
communities (Mitchell et al., 2010). Plant species established in nutrient-poor ecosystems 
have shown to have significant effect on the chemical properties of soil (Zhang et al., 
2011). This interaction and interdependence of plants with their biotic and abiotic soil 
environment is defined as plant-soil feedback (Ehrenfeld et al., 2005). However, little is 
known how successional changes in soil microbial communities affect vegetational 
changes over a longer timescale and how these feedbacks operate under long-term 
ecosystem change (Wardle et al., 2004). Studies have shown that plant species occupying 
different successional stages have different ecophysiological properties that might have a 
strong influence on soil biological properties (Porazinska et al., 2003). For example, fast-
growing plant species are common during early succession and they produce high-quality 
litter (nitrogen-rich) that favor bacterial dominated communities while slow-growers that 
dominate late successional stages produce low-quality litter (high in lignin) that tend to 
favor fungi which are efficient lignin decomposers (Kardol et al., 2006). Over the past 
decade biotic component of the ecosystems are considered as important determinants of 
the ecosystem functioning. The aboveground and belowground feedbacks (linkages and 
interactions between organisms aboveground and belowground) are regarded as the most 
important factors that drive processes at both community and ecosystem levels (Wardle et 
al., 2007).

**Bacterial species problem**

In spite of 3 billion years of evolution, only 5000 bacterial species have been 
described so far (Finlay et al., 1999). This is much less when compared with over a 
million animal species that have evolved in only 600 million years. This striking disparity
has been explained by various authors as inconsistencies that exist in the procedures to define species of bacteria (Oren, 2004). Several reasons have been attributed to this disparity. First, unlike higher organisms, it is difficult to identify and describe new bacterial species. For many organisms, it requires labor intensive and equipment intensive procedures to describe a species which might take several weeks to months of laboratory work. Secondly, for naming any bacterial species, it has to be isolated in pure culture and then described. There are many bacterial species that are yet to be described and it has been estimated that less than 1% of all bacteria have been isolated so far (Whitman et al., 1998). Thirdly, the species definition for bacteria is very broad when compared to plants and animals and this discrepancy is addressed in the following section. Microbial organisms share their genetic material often and in evolutionarily important ways. Lateral gene transfer usually refers to the transfer of individual genetic components across distant phylogenetic groups, while horizontal gene transfer refers to the interbreeding of genomes between species. Both these mechanisms pose a general problem in defining a species because most of the time the genetic lineages of the microbes do not generally match their expected phylogeny (Cohan, 2006).

**Defining bacterial species**

“Speciation is an evolutionary process in which a population containing distinct individuals diverges to form different clades that eventually become separate species” (Staley, 2006). A species concept is a set of prescribed rules that is based on the speciation process and that describes how to circumscribe a newly identified species. There is no convincing species concept (theory-based model) developed for microorganisms yet (Rosselló Mora et al., 2001). The microbial species today are defined by
pragmatic, arbitrary methods, the so called polyphasic approach (operational-based model), that is based on a set of criteria and cut-off levels for delineating strains. Unlike the species concepts, the operational species concept does not rely on the theory of speciation but on data-driven analyses (Ereshefsky, 2010). In early 70s, based on these cut-offs, a prokaryotic species was defined as a group of strains that showed 70% of DNA-DNA binding. Species recognized by this method clearly showed huge range of genetic variability. Thus, recently, the DNA-DNA hybridization method is replaced by the use of 16S rRNA gene sequences to identify closest relatives of an isolate and isolates that show less than 97% of 16S rRNA gene-sequence identity are considered to belong to different species (Riley et al., 2009). Thus, a cutoff of 3% divergence was recommended as a conservative criterion for demarcating species (Gevers et al., 2005; Ereshefsky, 2010). Based on these methods, defining bacterial species is considered very broad in comparison with that of animals and plants. For example, if we applied the 3% divergence to tree species, all oaks would be grouped as a single species and all pines would be grouped as a single species as well. This discrepancy is felt especially when the correlations between plant and microbial species are interpreted. However these methods are relatively quick and could be automated for screening of large number of isolates (Staley, 2006; Staley, 2009).

**16S rRNA gene**

In dealing with organisms undergoing profuse gene transfer both vertically and horizontally, a highly conserved gene is essential to study the phylogeny to avoid ambiguity in naming the taxons. Essential genes like those encoding for ribosomal RNA are unlikely to be transferred through the horizontal gene transfer mechanisms, as the
recipient organism would already have a functional orthologue (Fraser et al., 2007). In particular, the study of microbial evolution based on small ribosomal subunit RNA sequence analyses is gaining importance in the study of bacterial diversity. The resulting phylogenetic evidence has shown that the perceived lack of bacterial species is not simply due to lower levels of biodiversity relative to the multicellular eukaryotes but lack of proper tools to characterize large reservoir of genotypes from uncultured species (Fenchel et al., 2005; Fraser et al., 2007). The 16S rRNA is a 1500 bp nucleotide that is part of 30S small subunit (SSU) of prokaryotic ribosome. The conservative property of the 16S rRNA sequence was discovered a long time ago. Since then it has been widely used for bacterial identification and taxonomy. This gene is ubiquitous among prokaryotes and has relatively slow evolution rate. These properties make it highly conserved. It has highly conserved genes interspersed with hyper variable regions. The variable regions are used to discern phylogeny by comparing divergent bacteria, while the highly conserved regions are used as templates to design specific PCR primers. These PCR primers are used to amplify the conserved regions of a large population of prokaryotic cells in natural environment and the variability in these PCR amplicons are used to detect evolutionary shifts. It acts like a molecular chronometer that could act as a reliable marker (Woese et al., 1985). Although the absolute rate of evolution of the gene is unknown, the rate is approximately congruent with the speciation and hence these genes are described as “molecular clocks” (Woese et al., 1987). There is a huge database of the 16S rRNA gene sequence in GenBank that are deposited previously and these are useful to compare the sequence of unknown strain. These sequences are universal among bacteria and thus help in determining the relationship among them. They are useful in
differentiation of organism up to the genus level across major phyla of bacteria and also in classifying strains at multiple levels (Woese, 2000).

**Genomics, metagenomics and next-generation sequencing**

During 1960s, comprehensive studies on microbial diversity focused on culture-based studies and less attention was given to the possible habitat interaction. This limited our understanding of their behavior in natural ecosystem (Torsvik et al., 1996). This knowledge-gap could be filled through culture independent studies that would help us understand the importance of microbial phylogenetic groups and their interaction with the given habitat. This was achieved in 1980s, when microbiologist started to look beyond single organisms and started exploring density, diversity and composition of microbial populations. Introduction of genomics into microbial ecology paved way for exploring enormous biodiversity of uncultured microorganisms (Kirk et al., 2004). This is because most of the environments are dominated by uncultured organisms and culture-dependant techniques would severely underestimate the extent of microbial diversity. From 1990s, considerable molecular methods were developed to characterize microbial diversity by analyzing nucleic acids (DNA) extracted from the environmental samples. This led to exponential increase of several studies focusing on studying microbial communities in various environments, such as environments subjected to disturbances. This marked the era of “molecular ecology” that attempted to discover the unexplored microbial diversity through metagenomics (Xu, 2006).

Metagenome in this context is a set of microbes that make up the environment together and the goal of a metagenomic sequencing study is to sequence an entire microbial community without isolating or cultivating individual organisms. This is
accomplished by simultaneous sequencing of microbial genomes that comprise the
community (Maron et al., 2011). Polymerase chain reaction (PCR) techniques were used
to selectively amplify the target genes from all DNA sequences extracted from the
communities in a given environment. These amplicons (PCR amplified DNA segments)
were traditionally cloned into a vector and were used for sequencing the community
DNA. Many studies used Sanger sequencing of 16S rRNA gene libraries made by
cloning 16S rRNA amplicons obtained from environmental samples into E. coli vector
(Tiedje et al., 1999). Although sequencing by clones offered an initial high-resolution
snapshot of the ecological community, sequencing thousands of clones are cumbersome.
Sequencing depth for each sample is limited, usually to 90 clones (sequences) and was
costly and time consuming (Tiedje et al., 1999).

The advent of high-throughput, massively parallel sequencing technologies (next-
generation sequencing) in 2005 allowed for more environmental samples to be sampled at
a higher level of phylogenetic diversity, generating more robust methods of inferences
between environments (Roesch et al., 2007). Two platforms are used for sequencing
study, 1. The Roche 454 FLX platform uses “pyrosequencing” and 2. The Illumina
(Solexa) GAIIx, Hiseq platforms. Illumina platforms has much more sequencing depth
(up to 600 Gb) compared to Roche (400Mb). But Roche 454 produces longer reads (up to
1000 base pairs) compared to Illumina which is currently producing reads up to 200 bp
(Roesch et al., 2007; Lauber et al., 2009). The greater breadth of sequencing in
pyrosequencing was considered particularly advantageous for meaningful phylogenetic
inferences in soil ecology studies (Roesch et al., 2007)
Pyrosequencing has been utilized to describe and define community composition in multiple environments especially in under-sampled environments. Roesch et al. (2007) studied the bacterial diversity of forest and agricultural soils using pyrosequencing technique. They were able to produce up to 50,000 sequences by sequencing the 16S rRNA gene of the bacterial DNA and concluded that bacterial diversity in the forest soil was phylum rich while the agricultural soils were species rich with less number of phylum representatives. In another study, Acosta-Martinez et al. (2008) evaluated the influence of management and land use on bacterial diversity using modified FLX pyrosequencing approach. They compared the influence of agricultural systems with undisturbed pasture monoculture on bacterial diversity and pattern. They were able to reasonably predict the diversity and distribution of soil bacteria and concluded that it is a powerful method to characterize bacterial diversity in soil under various management and land use systems. Edwards et al. (2006), in one of the first attempts to use pyrosequencing technique to study environmental samples, identified crucial differences in metabolic potential of microbes in adjacent deep mine environments. They were able to extensively study microbes and their subsystems by integrating pyrosequencing technique with hydrogeology and chemistry. They also suggested that the technique would be widely used to study environmental samples in the future due to their speed and technical advantages.

Thus, a state of art pyrosequencing method would offer several advantages for metagenomics studies (Ronaghi, 2001). It will avoid cloning bias that is introduced in sample preparation which is often encountered in Sanger sequencing–based methods (Ronaghi, 2001). The Genome Sequencer FLX System is capable of producing 400,000
sequencing reads per instrument run. These unprecedented sequence numbers will offer more information and coverage to support research studies to answer environmental and ecological questions. These methods will also help to find functionally important but often poorly sampled microbial groups from environmental samples (Roesch et al., 2007). However, sequence output is enormous and they are almost 10 folds over the clone based sequences per sample. The depth of sequencing causes high burden on current computers and should therefore be done on parallel computing platforms. Although they are at high computational costs, they provide useful information to solve complex interactions.

**Relevance of research**

The majority of above mentioned research has focused on microbial community structure and diversity estimates on mature soil, specialty soils or system comparisons (Bardgett et al., 2005; Fierer et al., 2006; Upchurch et al., 2008). Only a few studies (Allison et al., 2007; Tarlera et al., 2008) have attempted to study the microbial diversity across a developing soil ecosystem that examined how the dynamics and complexity of the community structure would change over a progressive gradient. This would help understand the relative importance of various environmental factors that drive the microbial community structure in a developing ecosystem.

Studying microbial ecology from a temporal perspective is a recent development that involves studying microbial dynamics across soil chronosequences, receding glacier forefields and developing ecosystems (Fierer, 2008; Ranjard et al., 2010). Soil chronosequences are ideal environments for studying microbial diversity along a primary succession gradient. Because of the length of time required to observe soil development
in one place, we choose sites at different places that differed in soil development which are defined as soil chronosequences (Huggett, 1998). Soil chronosequences are promising ecosystems to understand the trajectory of microbial community composition change in relation to sequentially changing soil factors.

In this study, we investigated differences in bacterial community composition and diversity across a series of developmental sand-dune chronosequences. We used pyrosequencing-based approach to quantify phylogenetic differences in the bacterial communities with an objective to examine how changes in abundance and composition of soil bacterial communities across a gradient of ecosystem development correspond to changes in specific environmental characteristics. It is hypothesized that the bacterial communities will follow a pattern of change that would be associated with the dynamics of soil physico-chemical and vegetative characteristics.
CHAPTER II
MATERIALS AND METHODS

Soil sampling and preparation

Five replicate soil samples were taken at an interval of 20 feet between each sampling spot across a 100 feet transect ran along the crest of each dune. From each sampling spot 5-6 soil sub-sample cores were collected from zone of dominant root activity (A-horizon) at a depth of 0-15 cm using a stainless steel soil corer. After sampling, soil from other zones were carefully removed and immediately transferred into a Whirlpak® bags after homogenizing the sub-samples from each sampling spot. The sample bags were frozen immediately in cooler pack filled with dry ice. Upon arrival in the laboratory, soils were thawed for 25-30 minutes, homogenized through a sieve; extraneous roots and organic materials were removed and stored at -80°C. To study the seasonal effects on community composition, two sets of samples were collected in August and November 2008 at the Michigan Sites (Figure 2.1), representing the summer and winter samples respectively. One set of soil samples were collected in March 2008 at the Georgia sites (Figure 2.2).

Vegetative sampling

Plant sampling was done on each chronosequence by measuring the plant species composition, tree density, and percentage canopy cover across the sampling area.
covering a strip of 5 x 20 m (100 sq.m). The tree species composition was measured by counting the number of tree species within the sampling area and tree density was measured using the dbh (diameter breast height). Under-storey species cover was measured at five random spots within the sampling area using a 1 sq.m quadrat. Two quadrats were randomly placed on each sampling spot along each chronosequence transect. All the understorey species were identified and their percentage cover was estimated by visual observation as agreed by two observers. For the vegetation data, species that occurred in 3 or fewer plots were removed to avoid them influencing the ordination. The tree species canopy cover was estimated by fitting the dbh measurement into a conifer crown radius model (Dequiedt et al., 2011).

**Soil characteristics**

Soil organic matter content was measured by loss on ignition method (560 °C) by measuring the mass difference before and after ignition. The soil organic matter was calculated as the mass difference before and after ignition.

The respiration rates were determined by measuring CO₂ concentration from each serum bottles at regular intervals. The mineralizable carbon was estimated by measuring the cumulative CO₂-C from soil incubation experiment performed for a month using 1 L canning jar filled with 100 g of soil. CO₂ was sampled from each of the mason jars using 10 mL syringe and the gas samples were kept in vacuum vials before measurement on a gas chromatograph equipped with a TCD detector. (Varian 3600, Varian Instrument Group, Walnut Creek, CA).
Soil pH was measured on 1:2 soil and 0.01 M CaCl2 mixture. Soil extractable cations analysis was performed according to the Mehlich-3 extraction protocol (Maron et al., 2011).

**Phospholipid fatty acid analysis**

Phospholipid fatty acid analysis (PLFA) is a biochemical method that is used to study microbial community composition based on the grouping of fatty acids (Frostegård et al., 1996). PLFAs are integral parts of the cell wall of micro-organisms and are metabolized rapidly when the organism dies. Thus PLFA measurements have proved to be important indicators of viable biomass of a community (Frostegård et al., 1996; Moore-Kucera et al., 2008). Each microbe has a relatively constant PLFA composition that account for the total cell biomass. Thus the change in PLFA composition is often a good indicator of change in microbial community composition in an environment. There are also signature fatty acids (called markers) that could be used to differentiate major taxonomic groups among the microbial communities (Balser et al., 2005; Huang et al., 2011).

Phospholipid fatty acids (PLFAs) were extracted to estimate microbial biomass and community composition according to procedure proposed by White et al. (1979) and modified by Butler et al. (2003). All the glassware was cleaned with phosphorus free soap, rinsed well with deionized water and a final rinse with acetone. They were covered with aluminum foil, autoclaved and dried overnight. Ten grams of soil (dry weight) from each sample was measured into 160 ml serum bottles. Solvent solution containing a mixture of 50mM phosphate buffer (pH 7.1), chloroform and methanol (0.8:1:2) was added to the soil in the same order and extracted overnight. The samples were centrifuged
at 1000 rpm for 5 min and filtered using Whatman No1 filter paper and the soil was washed with methanol and chloroform and filtered twice to get the most of the lipids. The filtrate was added with 3 M NaCl solution and a pinch of Na₂SO₄ salt. The mixture was shaken well and the phases were allowed to separate overnight. The chloroform phase was removed using a pipette into a 10 ml glass tube. It was then dried under a stream of nitrogen in a warm water bath (< 35°C). The total lipids were fractionated into neutral, glycol and phospholipids using silicic acid bonded phase extraction columns (SPE) (Supelco, Cat.No. 505048). Neutral lipids were eluted out using chloroform and glycolipids using acetone. Finally, the phospholipids that were retained in the column were eluted out using methanol additions in a fresh glass tube and dried under a stream of nitrogen. The dried PLFAs were then saponified (methylated) under alkaline conditions using methanolic KOH, methanol and toluene mixture in a water bath for 15 minutes at 35°C. After neutralizing the mixture with 1 M acetic acid, the fatty acid methyl esters (FAME) were extracted using 1:4 chloroform and hexane mixture and transferred to fresh tubes. The mixture was completely evaporated under stream of ultra high purity nitrogen and the residue was re-suspended in 500 μl of hexane for GC analysis. Fatty acid methyl esters were separated, quantified and detected by an Agilent 6890 Series gas chromatograph (Santa Clara, CA) equipped with a flame ionization detector and controlled by a computer loaded with ChemStation and Sherlock software. Ultra high purity hydrogen was the carrier gas at a column head pressure of 20 KPa, septum purge of 5 ml min⁻¹, a split ratio of 40:1, injection temperature of 300 °C and injection volume of 2 μl. The oven temperature ramped from 170 °C to 288 °C at 28 °C min⁻¹ and the analysis time of each sample was 6 min. Peak identification was carried out by the
Microbial Identification System (MIDI, Inc.) following calibration with a standard mixture of 17 fatty acid methyl esters (1300A calibration mix). The signal was recorded along with time and then plotted to generate the chromatogram; the ChemStation software has an algorithm to interpret the chromatogram, detect peaks and calculate the area under the peak. A fatty acid name is then assigned to each peak based on the ECL (estimated chain length) of the fatty acids. Sherlock runs the calibration standard once for every 11 samples and updates the retention times used in these calculations each time.

**DNA extraction and pyrosequencing**

DNA was isolated from 0.5 g of soil from each soil sample using ZR Soil-Microbe DNA™ kit (Zymo Research). After extraction and purification, the DNA was inventoried and stored at -80 degrees. Small subunit bacterial rRNA gene fragments were amplified from an amount of DNA equivalent to that found in a 0.5 gram of soil to appropriate size using 515R-M (5’-CCGCGNGCKGCTGGCAC-3’) (Acosta-Martínez et al., 2008) and the sevenfold-degenerate primer 27F-YM+3 (Frank et al., 2008). The primers were synthesized in such a way that A and B sequencing adaptors (454 Life Science’s FLX) were immediately upstream of the 515R-M and 27F primer sequences, respectively. In addition to that an 8-nt sample–specific barcode tag was attached between the A-adaptor and primer 515R –M to identify and separate each sample sequence bioinformatically after sequencing. Each 25 µl PCRs consisted of 12.5 pmol of each forward and reverse primer, 1.25 µL of template DNA, and 22.5 µL of Platinum PCR SuperMix (Invitrogen). Samples were initially denatured at 95 °C for 3 min, then amplified by using 20 or 30 cycles of 94 °C for 30 s, annealing at 50 °C for 30 s, and extension at 72 °C for 1 min. A final extension of 4 min at 68 °C was added at the end of
the program to ensure complete amplification of the target region on a Veriti® 96-well Thermal Cycler (Applied Biosystems). The PCR products were run on a 1 % agarose gel and image quantified on a Typhoon Trio+ Variable mode imager (GE Healthcare) using Image Quant 5.2 (Molecular Dynamics). The PCR products from 5 replicates of each soil age were then pooled into equimolar concentrations and gel eluted using Zymoclean™ Gel DNA Recovery Kit (Zymo Research). The amplicons were then quantified on the Experion System (Bio-Rad) and a composite sample for pyrosequencing was prepared by pooling equal amounts of PCR amplicons from each soil age. The final mixed amplicon pool was further purified using the Agencourt AMPure XP system (Beckman Coulter Genomics) and submitted to the Environmental Genomics Core Facility at the University of South Carolina for pyrosequencing on a 454 Life sciences Genome Sequencer FLX (Roche) machine (Figure 2.3). Pyrosequencing generated 151,794 quality short-read bacterial 16S rRNA sequences from 125 samples across both the sites, generating an average of 1214 sequences per sample. The gene fragments averaged to approximately 530 bp (base pairs) in length.

**Processing of 16S rRNA gene data**

A two step pipeline was established to analyze the 16S rRNA gene sequence data. QIIME (Quantitative Insights into Microbial Ecology) was used to quality trim the raw sequences for primers, chimeras and to sort them based on the barcodes (Caporaso et al., 2010; Caporaso et al., 2011). QIIME was used to assign multiplex reads from sequence pools to specific samples using sequence barcodes, as well as to remove low-quality reads and to filter reads by length. The denoised data was then passed through the MOTHUR v1.22.0 (Schloss et al., 2009), a software for describing and comparing
microbial communities. To facilitate the downstream analysis of the large sequence datasets, identical sequences or artificially duplicated sequences, which can constitute a significant fraction of the dataset, were removed. The non-redundant sequence dataset was then aligned using Silva reference dataset (http://www.arb-silva.de/) (Quast et al., 2013). A second pass in MOTHUR, sequence reads were assigned (clustered) to OTUs (Operational Taxonomic Units) based on pairwise distances between all aligned sequences. A column formatted distance matrix was generated using average-neighborhood algorithm at an evolutionary distance D=0.03, which restricted the distance matrix to keep only sequence reads that had 97% sequence similarity. The alpha-diversity estimates such as rarefaction, richness, evenness, Shannon, Simpson’s reciprocal index and Chao1 estimates were done on OTUs at D=0.03 evolutionary distances (about 97% sequence similarity). This level of DNA sequence similarity is typically used to assign sequences to the same species (Schloss et al., 2009). The rarefaction curves (species accumulation curves) describe how comprehensively a microbial community has been assessed in the sample. It shows how richness or diversity changes as additional individuals from the community are sampled. It shows the relationship between the number of individuals sampled (x-axis) and number of genotypes actually observed (y-axis) in a given sample. The rarefaction curve rises as new genotype is added in each sampling effort and the curve reaches an asymptote (to x-axis) when no new species are found in the sample. Finally the phylogeny was assigned to representative sequences from OTUs using SILVA reference sequences.
Statistical and multivariate analyses

In order to visualize the differences in microbial community assemblage among the successional stages, 16S rRNA gene sequences and PLFA nonmetric multidimensional scaling (NMDS) ordination techniques were used. Ordination is a statistical tool used to simplify complex datasets by offering a graphical representation that makes it easy to discern trends. Ordination methods help to summarize a data set containing many variables into smaller number of variables. Ordination methods are very useful to iteratively search for the best way to represent the multivariate data into a reduced number of dimensions and to visually them (McCune et al., 2002). This will result in ordination diagrams in which the distances reflect the actual dissimilarities in the microbial community structure of the original samples. Distance between two points is proportional to the dissimilarity value for a given pair, such that points positioned close together are more similar than points plotted further apart. This method was used because it does not assume linear relationships between variables and is very effective method for ecological data (Clarke, 1993).

Bray-Curtis ordination analysis was used to visualize bacterial community structure and to investigate the components of the structure that has the most influence on the final ordination solution. This method was used to calculate dissimilarity matrix, which is a pair-wise comparison based on Sørensen distance measure. In this method, samples are randomly placed in 1 to 3 dimensions and Euclidean distance is calculated between samples. The elements of dissimilarity (Sørensen distance) are ranked in ascending order and plotted against the ordination distance (Euclidean distance). A stress value is calculated which measures the departure from monotonicity of the above
mentioned plot. Hence, a lower stress value means a more reliable relation. After multiple iterations the lowest stress value can be attained and samples containing common microbes will cluster together in ordination space (McCune et al., 2002). The final dimensionality of the dataset was determined based on the stress and stability measurements from the ordination analysis (McCune et al., 2002) using the PC-ORD software (MJM Software, Glendenen Beach, OR, USA). Each axis in ordination table represents a fraction of variance in the original data and the percentages in each axis (coefficients of determination) indicate the percentage of variance explained (represented) by that axis. A one-dimensional solution (1 axis) depicts the variability in dataset that could be interpreted by a single factor. In a two-dimensional solution (2 axes) the ordination diagrams could essentially show different factors controlling the variability in the dataset. The spread of data along the axis-2 (in 2-axes solution) indicates that a second factor (other than one that is represented in axis-1) might control the variability of the data points along that axis.

Mantel test was performed using the PC-ORD software (MJM Software, Glendenen Beach, OR, USA) to determine correlations between bacterial community composition, vegetation data and soil characteristics. For the Mantel test, Sørensen distance measure was used with a random starting configuration. Pearson and Kendall correlations ($r^2$ values) between the ordination axes and the environmental variables were calculated using the Sørensen distance measure. Log-linear correlations were performed using Excel 2007. Statistical comparisons to test for differences between ages and season effects on OTU distribution were conducted by using a two-way general linear model (GLM) and the statistical significance was measured using SAS version 9.2 (SAS.
Institute Inc., Cary, NC, USA). Multiple mean comparisons were done using Fisher’s least significant difference (LSD) in SAS version 9.2 (SAS Institute Inc., Cary, NC, USA). Hierarchical cluster analysis of the most abundant OTUs was done using the PC-ORD software (McCune et al., 1999). The relationship between microbial community composition and environmental parameters was analyzed by Canonical Correspondence Analysis (CCA) using the PC-ORD software (McCune et al., 1999). The significance of the correlations with each factor was evaluated through the Monte Carlo permutation test by applying 998 permutations.
Figure 2.1  Aerial photography of the beach-ridge complex at Wilderness State Park in Michigan.

Photograph showing the 9 sampling sites along each dune that is arranged parallel to the beach.
Notes: Map reproduced with permission (Lichter, 1995)
Figure 2.2  Aerial photography (University of Georgia Map Library) of the riverine dunes at Georgia showing the five sampling sites along the eastern coast of Altamaha and Ohoopee river basins.

Notes: Map reproduced with permission (Ivester et al., 2003).
Figure 2.3  Overall steps involved in creating an emulsion amplified 16S rRNA gene library from the soil samples before pyrosequencing.
CHAPTER III
THE STUDY OF BACTERIAL COMMUNITY DIVERSITY AND SUCCESSION ACROSS THE MICHIGAN CHRONOSEQUENCE

Site description
The study site consists of a series of beach-dune complex bordering Lake Michigan located in the Emmet County of northern Lower Michigan. Periods of unusual moist conditions altering with dry spells associated with swelling and receding of Lake Michigan formed a series of approximately 108 eolian deposited dune ridges running parallel to Lake Michigan with a depositional age from present day to approximately 4500 years (Lichter, 1995). The dune ridges have a parent material originating from glacial deposits and Paleozoic bedrock underlying the lake basin. The soil type is fine sand dominated by quartz but containing numerous other minerals in minor quantities. The chronology of the dunes was estimated using Accelerated Mass Spectroscopy (AMS) radiocarbon dating of the macrofossil remains from each dune (Lichter, 1995). The ridges are approximately 2.5 km long, 10–30 m wide, and vary between 3 and 5 m in height along the shore and reaching 15 m high parabolic dunes in the inland (Lichter, 1998b). The site has a uniform textured parent material which was predominantly sand transported by wind blowing from the lake towards the shore. Since Lake Michigan buffers the daily, seasonal and annual temperature fluctuations, the dune successional plant species are minimally affected by the climate change. Previous study
(Lichter, 1998) has shown patterns of primary succession with grasses and shrubs on younger dunes to mixed coniferous forests dominating the older dunes.

**Soil and vegetative sampling**

Five replicate soil samples were taken at an interval of 20 feet between each sampling spot across a 100 feet transect run along the crest of each dune. From each sampling spot 5-6 soil sub-sample cores were collected from zone of dominant root activity (A-horizon) at a depth of 0-15 cm using a stainless steel soil corer. Five replicate samples were also taken along the beach to simulate the material that might be a source of these eolian deposits (time zero). After sampling, soil from other zones were carefully removed and immediately transferred into a Whirlpak® bags after homogenizing the sub-samples from each sampling spot. The sample bags were frozen immediately in cooler pack filled with dry ice. Upon arrival in the laboratory, soils were thawed for 25-30 minutes, homogenized through a sieve; extraneous roots and organic materials were removed and stored at -80°C. To study the seasonal effects on community composition, two sets of samples were collected in August and November 2008, representing the summer and winter samples respectively.

Plant sampling was done on each chronosequence by measuring the plant species composition, tree density, and percentage canopy cover across the sampling area covering a strip of 5 x 20 m (100 sq.m). The tree species composition was measured by counting the number of tree species within the sampling area and tree density was measured using the dbh (diameter breast height). Under-storey species cover was measured at five random spots within the sampling area using a 1 sq.m quadrat. Two quadrats were randomly placed on each sampling spot along each chronosequence.
transect. All the understorey species were identified and their percentage cover was estimated by visual observation as agreed by two observers. For the vegetation data, species that occurred in 3 or fewer plots were removed to avoid them influencing the ordination. The tree species canopy cover was estimated by fitting the dbh measurement into a conifer crown radius model (Dequiedt et al., 2011).

**Objectives and hypothesis**

**Objectives**

- To study the composition and diversity of bacterial communities as a function of soil ecosystem age and season along a set of developing sand chronosequences
- To understand the relative importance of various environmental factors that drive the microbial community structure and diversity in a developing ecosystem

**Hypothesis**

It was hypothesized that the bacterial communities would follow a pattern of change that would be associated with the dynamics of vegetation change, soil physico-chemical characteristics and with season.

**Results**

**Changes in soil characteristics during ecosystem development**

Soil Ca and Mg levels showed a significant log-linear correlation with the soil development. They started leaching out of the system as the soil aged (Table 3.1) which was consistent with age related weathering. This change was rapid during the early stages
of soil development and started to stabilize during the later stages. This effect was reflected in the soil pH which started from near-neutral (7.6) and becoming increasingly acidic (3.5) as the soils got older due to the leaching of the carbonates. The soil pH also showed a similar trend in which the younger soils showed drastic change in pH and were significantly different from the relatively stable older chronosequences. Soil organic matter measured as percentage dry weight and total soil carbon showed a significant decrease along the chronosequence from younger to older soils. Potassium levels were relatively stable as the soil weathered. Soil Na and P levels did not show any significant correlation with soil development. Inherent low P pools (data not shown) would have caused the failure to show any significant trend in soil P levels.

16S rRNA data

For bacterial community analysis 5 replicates soil samples were obtained from each of the 9 developmental chronosequences and beach sand (0 years) in both the seasons (summer and winter). Thus a total of 50 different libraries were produced in each season from this study. Across all 100 samples, we obtained a total of 108,273 quality sequences, with a range of 1068 to 18,041 sequences from each soil. The average read length of the sequences was ~ 530 bp (base pairs). In order to verify whether the number of sequences analyzed affected the calculation of diversity indices, random sub-samples from each chronosequence library was used to build the distance matrix and to calculate the diversity indices. The results (Table 3.5) indicated that the sample size did not influence the diversity indices calculations. Thus to improve the coverage all sequences were used for the data analysis.
**Bacterial diversity indices**

As measured by the Shannon and Simpson’s reciprocal index, the bacterial diversity tended to decrease considerably across the chronosequence (Table 3.2) with largest decline during initial stages of soil development. The diversity declined from 212 to 58 (Simpson’s 1/D) during 105 to 450 years of soil development. The chao1 richness predictor values showed that only 20-37% of the OTU’s predicted by this estimator were actually observed. The rarefaction curves, which shows number of observed OTUs as a function of number of sequences sampled, also failed to plateau (Figure 3.6) indicating that the diversity is not completely sampled at evolutionary distance of 0.03. When grouped at the 97% similarity level, there were 19,893 OTUs in the complete data set, which had 15,013 singletons and 1900 doubletons, averaging to 75 % to 9.5 % of sequence representation. The decline in bacterial richness (Table 3.2) showed a similar change in trend associated with the soil pH, SOM and vegetation in which the biggest decline coincided with the youngest soils (105y, 155y and 210y) while the decline was smallest as the soils developed.

**Microbial community structure and ecosystem properties**

The distribution of 200 most abundant OTUs across the chronosequences and beach sand from the summer and winter samples showed strong relationship between the bacterial community change and ecosystem development (Figure 3.1). The bacterial community change at the younger sites was significantly different from the older sites as explained by the axis of maximum variability from Bray-Curtis ordination (Axis 1= 62 %). There was no significant difference in the pattern of bacterial OTU distribution between the summer and winter samples. But, the beach sand showed clear difference in
structure than the true chronosequences (Figure 3.1). Hence for further analysis the beach sand was excluded and the chronosequence samples were re-analyzed. Thus to depict the variability of the OTU distribution data across ages, the variability of the OTU distribution of the summer samples across the axis of maximum variability were plotted against the soil ages (Figure 3.2a). This pattern of variability was strongly related to the change in plant distribution (Figure 3.2a and b). There was also significant correlation between the Bray-Curtis ordination scores and pH (p<0.0001, r=0.97), vegetation (p<0.0001, r=0.87) and SOM (p<0.00001, r=0.87) (Table 3.3). Total carbon in the soil also showed significant log-linear relationship (r=0.76) with the ecosystem development, and in contrast, soil microbial mineralizable carbon showed no significant relationship (Table 3.1).

At a broad level of phylum-class classification, the most abundant bacterial phyla across the chronosequenes were Actinobacteria, α-Proteobacteria and Acidobacteria covering 71 % of all the sequences with 31%, 26% and 14% individual contribution, respectively (Figure 3.4). The other less abundant phyla were Cyanobacteria, β-Proteobacteria, Planctomycetes, Bacteroidetes, γ-Proteobacteria, Firmicutes and Chloroflexi that had an overall abundance of ~ 4% or less. α-Proteobacteria, β-proteobacteria, Bacteroidetes and Firmicutes showed a significant change in percentage of abundance across the soil and ecosystem development (Figure 3.3 and 3.4). Excluding the 105year chronosequence, α-Proteobacteria, β-proteobacteria, Bacteroidetes and Firmicutes showed a decreasing log-linear relationship across the ecosystem gradient with a higher abundance at younger soils and the abundance decreasing as the soil gets
older. Although not statistically significant *Acidobacteria* showed an increasing abundance as the soil developed and became increasingly acidic (Figure 3.3).

The distribution of bacterial phyla in the freshly deposited material (0 year) was very different from the true chronosequences (Figure 3.1). *Cyanobacteria* was the dominant phylogenetic group (25 %) followed by *α-Proteobacteria* (22 %) and *β-Proteobacteria* (20 %). Unlike the true chronosequences, *Actinobacteria* and *Acidobacteria* were among the less abundant phyla with only 9 and 4 % individual contributions, respectively. There were 23 OTU that were unique to the freshly deposited materials (0 years) which were not observed in true chronosequences. Thus, approximately 46% of the OTUs were unique to the freshly deposited beach sand and not found in chronosequence. Assessing the closest cultured representatives of 5 dominant OTU from these unique OTU group showed that three of the closest matches belonged to *Cyanobacteria* and the other two belonged to *Proteobacteria* and *Firmicutes* (Data not shown).

**Change in plant community composition across the chronosequences**

Changes in percentage cover of thirteen plant species (including herbs, shrubs and hardwood trees) that changed considerably across the chronosequences were investigated. The change in plant species abundance was higher in the youngest sites than the older sites (Figure 3.2b). Dune-building plant species were replaced by evergreen shrubs which were later replaced by mixed pine forests. This shift in early-succession species to late-succession species happened around 450 years of soil development (Figure 3.8). The early-succession species started disappearing and the mixed pine forest started establishing as soil ecosystem developed. Once the forest matured and stabilized, the
plant species input reduced and there was no significant change in the plant community across the older chronosequences (Figure 3. 2b). Mantel test of a relationship between vegetation and bacterial communities along the developmental gradient was done and was statistically significant (p<0.000001; r=0.56). This overall trend is comparable with the change in the most abundant bacterial OTUs wherein the biggest change in community was noticed during initial stages of soil development while the change got stabilized as the soil matured (Figure 3.2).

A relationship of closest representatives from each dominant OTU was studied to understand the organisms that were associated with the change in trend of the bacterial OTUs across the age gradient. In order to achieve that, 5 OTUs that showed the highest log-linear relationship with the changing ecosystem were chosen and plotted (Figure 3.5). The percentage abundance of each OTU across the ages was calculated based on the relative abundance of each OTU with the total number of sequences associated with that age. The percentage abundance varied from relatively non-detectable to ~ 4.5 % of sequence samples (Figure 3.5). Two of the closest matches belonged to phylum Acidobacteria, and the other three belonged to Proteobacteria, Planctomycetes and Actinobacteria.

**Effect of seasonal variations on bacterial communities**

There was no significant difference in the pattern of bacterial community structure across the chronosequence. Similarly, there was no significant season X age effect on the bacterial community (p =0.45) (Figure 3.1).
The change in microbial PLFA

The sum of multiple groups of PLFA biomarkers such as i15:0, a15:0, i16:0, i17:0, a17:0, cy17:0, cy19:0, 16:1ω7, 18:1ω7, and 17:1ω9 were considered for bacteria (Frostegård et al., 1996). Branched PLFAs (i15:0, a15:0, i16:0, i17:0, and a17:0) were used as biomarkers for gram-positive bacteria and cyclopropyl and mono PLFAs (cy17:0, 16:1ω7, 18:1ω7, and 17:1ω9) were associated with gram-negative bacteria (Bossio et al., 1998). For fungi 18:2ω6 was considered as marker (Frostegård et al., 1996). Fungi to bacteria ratio was calculated using 18:2ω6 and the sum of bacterial biomarkers (Frostegård et al., 1996). Different stress indicators were used to assess the physiological status of the bacteria. The ratio of saturated (SAT) to monounsaturated (MONO) PLFAs and the ratio of cyclopropyl PLFAs (cy17:0 and cy19:0) to their precursors (16:1ω7c and 18:1ω7c) were used as stress biomarkers.

The multivariate Bray-Curtis ordination plots showed significant difference between the summer and winter PLFAs (Figure 3.9) along the axis 1. There was a significant season x age effect observed along this axis (p-value = 0.01) (Data not shown). When the PLFA profiles from different seasons were analyzed separately, the PLFAs showed patterns similar to the bacterial OTU distribution. A one-dimensional solution of Bray-Curtis ordination technique better explained the distribution of PLFAs (Figure 3.10) which again showed that the differences were high during the initial stages of soil development while the older ages showed less difference in PLFA signatures. This trend was more evident in the winter samples than the summer samples (Figure 3.10). A significant difference between the summer and winter samples were noted in the sum of the PLFAs representing different microbial groups. Summer samples showed significant
higher levels of fungi, gram positive bacteria, gram negative bacteria and total bacteria
(Table 3.4) than the winter samples. But, the PLFA groups and ratios failed to show any
significant age related trend (Figure 3.11 and 3.12). The stress ratios failed to show
significant difference between the seasons (Figure 3.13).

Discussion

The goal of this study was to understand the patterns of bacterial community
change as a function of ecosystem development. We initiated a study along a
developmental sand-dune chronosequence bordering northern Lake Michigan at
Wilderness State Park with a hypothesis that the change in bacterial community
composition patterns will follow the soil and vegetative changes across the ecosystem
gradient. We also attempted to understand if season had an effect in the bacterial
community composition across the developmental gradient. There was a significant
change in bacterial community composition and diversity along the ecosystem
development showing strong correlations with vegetation and soil physico-chemical
gradients. The season had no effect on the bacterial community composition. The results
suggest that vegetation could be an indirect driver creating gradients in soil pH and soil
organic matter, which in turn, could have been a strong driving force in shaping microbial
community assembly across a developing soil ecosystem.

The chronosequences showed changes in soil properties depicting the classic
patterns of soil podzolization during ecosystem development (Huggett, 1998). During
450 years of soil development, carbonate mineral started weathering and large quantities
of Ca and Mg started leaching from the system as the pH of the upper mineral soil
decreased from 7.6 to 3.6. As the dunes got older, decomposing coniferous litter material
would have helped in hastening the mineral weathering process by production of organic and carbonic acids (Huggett, 1998; Lichter, 1998a). Soil organic matter (SOM), measured as the percentage loss on ignition, decreased along the gradient during initial phases of soil development and then leveled off to a steady state. Although the initial decrease in the SOM availability was not characteristic of the pedogenesis and ecosystem development (Huggett, 1998), the trend in accumulation rate was consistent with findings from other studies that involved longer chronosequences (Lichter, 1998a; Kara et al., 2008). SOM was strongly correlated (r=0.87, p <0.0001) with the ordination axes that explained most of the variation in bacterial community distribution (Table 3.3). Thus resource availability, probably the available carbon levels, could likely be one of the fundamental drivers of the bacterial community composition.

The results from this study show strong connection between pH and bacterial community composition, in broad agreement with other studies examining soil bacterial biogeography across pH gradients (Lauber et al., 2009; Rousk et al., 2010a). Of all of the soil characteristics examined, soil pH was the best predictor of soil bacterial community composition (Table 3.3), and there was a strong correlation between the primary axis of figure 3.2 and soil pH (r= 0.97 for Axis 1, p < 0.0001). We observed that the bacterial diversity positively track soil pH (range 3.5 – 7.6) corroborating the results from recent studies showing higher diversity at high pH in temperate and arctic biomes (Hartman et al., 2008; Chu et al., 2010; Rousk et al., 2010a).

At phylum level classification, the changes in relative abundances of specific taxonomic groups across the chronosequences pH gradient are similar to the pH responses observed in other studies. For instance, the relative abundance of
*Acidobacteria* has been shown to increase towards lower pH (Jones et al., 2009; Lauber et al., 2009; Chu et al., 2010). Consistent with those results, our results shows that the relative abundance of *Acidobacteria* changed from lower abundance in near-neutral condition to higher abundance in acidic condition. Soil pH followed this change, especially during the initial phases of soil development (up to 450 years) when there was a significant decline in soil pH (7.6 to 3.6). The relative abundance of *Actinobacteria*, the most abundant bacterial taxa across the Michigan chronosequences, showed increased abundance in response to increase in soil pH, corroborating the results from other studies (Lauber et al., 2009; Chu et al., 2010). Excluding the youngest soil (105 years), the relative abundance of *α*-Proteobacteria, *β*-Proteobacteria and Bacteriodetes groups tended to negatively correlate with soil pH, a pattern that was evident in other studies that included many soil types (Lauber et al., 2009; Chu et al., 2010).

Investigation of the closest cultural representatives of each of the five dominant OTUs that showed a log-linear increase in abundance across the soil development gradient showed that four (*Edaphobacter modestus, Singulisphaera acidiphila, Acidobacterium capsulatum* and *Methylovirgula ligni*) out of the five OTUs were acidophilic. Although this subset of bacteria had phylogenetic diversity (belonging to different taxonomic groups such as *Acidobacteria, Actinobacteria and Planctomycetes*), the members of those phyla tend to share common attributes that helped them survive the low pH conditions. It has been shown previously that soil bacteria have a narrow growth tolerance range to soil pH (Fernández-Calviño et al., 2011) and that soil pH would have imposed physiological constraints on soil bacteria, thereby favoring certain taxa that may tolerate low pH conditions (Lauber et al., 2009; Rousk et al., 2010b). While these results
may not explain the reduced diversity of other taxa along the soil development gradient, it could be inferred that soil pH would have selected these taxa that have a physiological advantage under low pH conditions.

However, the changes in the soil chemistry that we observed could have been indirectly driven by the plant distribution. Plant species established in nutrient-poor ecosystems, like our study system, have shown to have significant effect on the chemical properties of soil (Zhang et al., 2011). Particularly during primary succession, plants are considered key drivers of changes in physical and chemical properties of the soil along the successional gradient (Van Breemen et al., 1998; Kuramae et al., 2011) and regulators of weathering and soil formation through generation of weathering agents (Van Breemen et al., 1998). In our study, the distribution of the plant communities closely followed the microbial community composition as described by the hierarchical clustering analysis of the plant communities (Figure 3.7). The plants grouped broadly into two major clusters comprising of the younger dunes and older dunes, a pattern that is similar to the bacterial community grouping. Our results showed an increasing plant species richness between 105 and 450 years of soil development followed by a long-term decline in species richness between 450 and 4010 years. This was highly correlated to the pattern of change of bacterial communities across the Michigan chronosequence (e.g. Mantel test). This pattern was consistent with the observations of Reynolds et al (2003) who suggested that positive feedback shapes the early successional communities which slowly changes to negative feedback situation in the late successional communities. According to Lichter (1998a), who extensively studied the primary succession patterns of Michigan chronosequences, the increase in early successional plant species richness in younger
dunes could be attributed to the increasing soil organic matter accumulation, soil moisture and an increased opportunity for colonization. While the decrease in species richness of the late successional plant communities could be attributed to diminishing understory herbs due to competition for light and soil nutrients (Lichter, 1998b).

Studies have shown that plant distributions can have a significant impact on the chemical nature of the soil with regard to pH of the soil that control the nutrient availability and cycling (Ehrenfeld et al., 2005; Kara et al., 2008). The strong correlation observed between the soil pH and bacterial community composition in our study could be indirectly plant driven. The largest change in pH observed across the chronosequence was highly correlated with the plant species distribution (Table 3.3). As the conifers colonized the young dune ridges (around 200 yrs), soil changes were accelerated. Organic acids produced during decomposition of the coniferous litter rapidly acidified the upper mineral soil (Table 3.1) and the pH started declining as the soils got older. This plant induced acidity is well documented over decadal scale of plant succession and is often correlated to the litter chemistry (Kelly et al., 1998).

We also hypothesized that season would have an effect on bacterial community composition. Studies have hypothesized that the microbial activity will be high during summer compared to winter due to higher labile carbon pool becoming available as a result of thawing and new root growth (Bardgett et al., 1999; Griffiths et al., 2003). The general assumption in these studies has been that the soil microbial communities are predominantly inactive during winter (Bardgett et al., 1999). Since our study site underwent high temperature fluctuations between seasons, we expected to see differences in the bacterial community composition between seasons. However, the bacterial
community composition did not show much seasonal variations across the ages (Figure 3.1) and the interaction effect of the season across the ages was significantly lower ($p = 0.45$). This could be a reflection of the long term effect of the ages and climate on the soil bacterial community composition than the seasons. The season also failed to show significant difference in soil characteristics (extractable nutrients, pH, Total carbon and nitrogen) across the soil developmental gradient (data not shown). Thus the results suggest that the bacterial communities along these chronosequences were uniform and stable across seasons. Many of the previous studies that showed patterns of change in microbial community across the seasons were looking at the community dynamics across fertility gradients (Bardgett et al., 1999) or production systems with nutrient inputs (Steenwerth et al., 2006; Yao et al., 2006; Yao et al., 2011) or unvegetated glacier forefields (Lazzaro et al., 2012). Most of these studies only reported structural changes (PLFAs) and not compositional changes between seasons. Our study system was unique in studying the seasonal variation across a wide range of soil ages, under natural ecosystem with least impact of external nutrient input. Our study is corroborated by some of the recent findings that showed land-use types having significant effect on the microbial community composition than the seasons (Jangid et al., 2008; Drenovsky et al., 2010).

However, the PLFA analyses showed that the microbial communities were influenced by the season. The total microbial biomass was at least 1.2 (4010yrs) to 4 (105yrs) folds greater in summer than winter samples (Table 3.4). This is often associated with the higher activity of microbial communities in summer than winter due to higher nutrient availability (Moore-Kucera et al., 2008). But there were no significant
differences between the soil nutrient characteristics including soil organic matter and carbon availability between seasons (data not shown). Also this seasonal effect observed in microbial PLFA biomass was not reflected in the microbial community composition change between seasons (Figure 3.1). The variation in the abundance and composition of the soil bacterial communities were strongly influence by other environmental factors and not the season. Also the stress markers that are often associated with nutrient and water stress (Moore-Kucera et al., 2008) did not show any significant difference between the seasons. The levels of stress ratios S/M and cys/pre were not significantly different between the summer and winter samples (p-values = 0.10 and 0.23, respectively). These results suggest that the difference in PLFA profiles observed between the seasons would have reflected only a broad change in the structure of microbial communities and not the composition change of microbial communities across the soil development. This broad change in community structure indicates overall difference in microbial biomass with very little change in the dominant groups of microbes that control composition. Thus the difference in the microbial biomass (Total PLFA) would have been an indication of physiological adaptation (accumulation of phospholipids) rather than increase in cell number. The structural similarity between the PLFA distribution plots (Figure 3.11) of the summer and winter samples also corroborates our interpretation. A similar trend in PLFA was observed in a clone library study of grazing and poultry litter amended soils in Georgia (Jangid et al., 2008) and the authors explained that the change in PLFA associated with seasons could reflect the physiological adaptation of microbes than cellular abundance.
In order to understand the relative importance of the environmental variables that were shown to influence the bacterial community composition, a Canonical Correspondence Analysis (CCA) was performed. This technique has been shown to be useful to identify the best predictor that influences soil microbial communities in ecological studies (Mitchell et al., 2010). CCA was performed using plant (based on distribution scores from the Bray-Curtis ordination), pH and soil organic matter as categorical environmental variables. These variables were chosen because they were already shown to influence the soil microbial communities significantly (McCune et al., 2002). CCA significantly explained 27% of the OTU–environment relationship across the first two canonical axes. CCA grouped the ages similar to the other ordination techniques used in the study (Figure 3.14). All the three variables showed a significant correlation with axis 1 (Monte Carlo test of significance, p=0.01). In the joint plot (Figure 3.14) the length and the direction of the environmental lines indicate the strength of the environmental variables with the community. Our results showed that plant diversity decreased with soil development, in the same manner as pH (Figure 3.14). These results further confirm the close relationship between plant diversity and pH in shaping these soil bacterial communities along the development gradient.

Another interesting observation is that the biggest change in the community composition occurred during early stages of ecosystem development within a short time span of 450 years of soil development. However, the community change between 450 and 4010 years of soil development showed very little change. Cluster analysis of the bacterial community distribution showed that the most abundant OTUs of the youngest ages (105, 155 and 210 years) clustered together and were different from the older soil
age communities which clustered separately (Figure 3.7). Many studies on microbial community succession have shown that microbial diversity could be higher during early stages of succession, since these ecosystems are subjected to more biotic and abiotic stressors. The plant inputs in initial ecosystems (early succession) are less complex and could be effectively used by diverse heterotrophic organisms which can influence the microbial community structure. However, during late successional stages, the microbial community diversity attains a steady-state due to increase in complexity of plant inputs which favors more specialists than generalists (Schaaf et al., 2011). This trend is closely followed by the highest changes in the edaphic factors that occurred during early stage of soil development (Table 3.1). We observed many correlations in our study and the coefficients of these correlations will increase with increasing range of the variable; thus, whichever factor that has the greatest range could probably be considered as the best predictor, provided the factors are related (Prescott, 2005). For instance, pH as a variable had a highest range of values (7.6 to 3.5) across the chronosequences and they also showed highest correlation values with bacterial community composition. So pH could be considered one of the best predictors of soil bacterial community composition. However, in our study, plant and other soil factors were also significantly correlated with the change in soil microbial communities. However, we predict and it has been shown (Berendse, 1998) that the rapid change in the edaphic factors that occurred in these sandy (nutrient–poor) ecosystems could have been predominantly plant driven. We also found strong correlations to support this hypothesis.

In conclusion, we found consistent correlations between soil microbial communities, aboveground vegetation and soil chemistry. Thus direct links between plant
and microbial communities are still missing. It has been shown that soil chemistry and plant community could explain different parts of variation in the soil microbial community composition (Reynolds et al., 2003). For instance, plant species are known to directly impact the below-ground communities through the quality and quantity of the carbon supply through root exudates around the root interface (rhizosphere) which makes the species composition and the relative abundance of different species of rhizosphere communities differ from those in the bulk soil (Paterson, 2003). Thus by extending this study to sample the rhizosphere soils it is possible to identify specific plant-microbe interactions by further in-depth look into the root interface of the specific plant species.
<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Ca (µg/g)</th>
<th>K (µg/g)</th>
<th>Mg (µg/g)</th>
<th>pH</th>
<th>Total C (%)</th>
<th>SOM (%)</th>
<th>Mineralizable C (µg/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>1314 a</td>
<td>16 a</td>
<td>110 a</td>
<td>7.6 a</td>
<td>0.49 a</td>
<td>0.58 a</td>
<td>76 a</td>
</tr>
<tr>
<td>155</td>
<td>762 b</td>
<td>20 a</td>
<td>156 b</td>
<td>7.1 b</td>
<td>0.42 ab</td>
<td>0.45 ab</td>
<td>121 a</td>
</tr>
<tr>
<td>210</td>
<td>622 b</td>
<td>19 a</td>
<td>104 a</td>
<td>5.7 c</td>
<td>0.49 a</td>
<td>0.30 bc</td>
<td>169 a</td>
</tr>
<tr>
<td>450</td>
<td>141 c</td>
<td>22 a</td>
<td>20 c</td>
<td>3.6 d</td>
<td>0.39 ab</td>
<td>0.19 c</td>
<td>134 a</td>
</tr>
<tr>
<td>845</td>
<td>109 c</td>
<td>24 a</td>
<td>12 c</td>
<td>3.7 d</td>
<td>0.10 b</td>
<td>0.19 c</td>
<td>129 a</td>
</tr>
<tr>
<td>1475</td>
<td>116 c</td>
<td>24 a</td>
<td>11 c</td>
<td>3.6 d</td>
<td>0.20 ab</td>
<td>0.15 c</td>
<td>108 a</td>
</tr>
<tr>
<td>2385</td>
<td>137 c</td>
<td>25 a</td>
<td>13 c</td>
<td>3.6 d</td>
<td>0.13 b</td>
<td>0.18 c</td>
<td>127 a</td>
</tr>
<tr>
<td>3210</td>
<td>110 c</td>
<td>22 a</td>
<td>10 c</td>
<td>3.7 d</td>
<td>0.19 ab</td>
<td>0.14 c</td>
<td>85 a</td>
</tr>
<tr>
<td>4010</td>
<td>108 c</td>
<td>24 a</td>
<td>8 c</td>
<td>3.5 d</td>
<td>0.26 ab</td>
<td>0.15 c</td>
<td>153 a</td>
</tr>
</tbody>
</table>

r-value 0.83 0.79 0.84 0.81 0.76 0.85 0.36

*a* Soil properties with significant log-linear correlation to soil age (p<0.05). Na (~ 149 µg/g) and P (~ 4 µg/g) did not show a significant correlation with the soil ages. Means within columns followed by the same letter are not significantly different between them at P <0.05.

*b* Results from the summer samples alone.
Table 3.2  Diversity indices of the 16S rRNA gene sequences.

<table>
<thead>
<tr>
<th>Diversity Index</th>
<th>105y</th>
<th>155y</th>
<th>210y</th>
<th>450y</th>
<th>845y</th>
<th>1475y</th>
<th>2385y</th>
<th>3210y</th>
<th>4010y</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>8896</td>
<td>16921</td>
<td>7793</td>
<td>7221</td>
<td>7098</td>
<td>7869</td>
<td>6078</td>
<td>9412</td>
<td>4779</td>
</tr>
<tr>
<td>S</td>
<td>3156</td>
<td>6084</td>
<td>3326</td>
<td>1260</td>
<td>1675</td>
<td>1445</td>
<td>778</td>
<td>1537</td>
<td>822</td>
</tr>
<tr>
<td>Evenness</td>
<td>1.96</td>
<td>1.86</td>
<td>1.94</td>
<td>1.70</td>
<td>1.68</td>
<td>1.63</td>
<td>1.67</td>
<td>1.69</td>
<td>1.72</td>
</tr>
<tr>
<td>Richness (ace)</td>
<td>25814</td>
<td>63462</td>
<td>38723</td>
<td>9349</td>
<td>16737</td>
<td>12166</td>
<td>3909</td>
<td>12715</td>
<td>4399</td>
</tr>
<tr>
<td>Shannon</td>
<td>6.85</td>
<td>7.02</td>
<td>6.84</td>
<td>5.28</td>
<td>5.41</td>
<td>5.14</td>
<td>4.82</td>
<td>5.40</td>
<td>5.01</td>
</tr>
<tr>
<td>1/D</td>
<td>212</td>
<td>115</td>
<td>121</td>
<td>58</td>
<td>52</td>
<td>41</td>
<td>45</td>
<td>72</td>
<td>52</td>
</tr>
<tr>
<td>Chao 1</td>
<td>13419</td>
<td>30469</td>
<td>16576</td>
<td>4622</td>
<td>7151</td>
<td>5372</td>
<td>2112</td>
<td>6193</td>
<td>2445</td>
</tr>
</tbody>
</table>

\( ^a \) Calculations based on the Operational Taxonomic Units (OTU) formed at an evolutionary distance of <0.03. The results shown here are from the summer samples.  
\( ^b \) Number of sequences in the library.  
\( ^c \) Number of OTU  
\( ^d \) Simpson’s reciprocal index.  
\( ^+ \) Number of sequences in the library.
Table 3.3  Correlations between the bacterial community distribution and selected environmental characteristics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pearson Correlationᵃ</th>
<th>Mantel Testᵇ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>P-value</td>
</tr>
<tr>
<td>SOM (%)</td>
<td>0.87</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Total C (%)</td>
<td>0.43</td>
<td>0.003</td>
</tr>
<tr>
<td>Total N (%)</td>
<td>-0.10</td>
<td>0.494</td>
</tr>
<tr>
<td>pH</td>
<td>0.97</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Ca</td>
<td>0.88</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mg</td>
<td>0.84</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Vegetation</td>
<td>0.56</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

ᵃ Pearson correlations between the ordination scores of the axis explaining maximum variability (78 % from the original OTU distribution data from the summer samples) of the non-metric dimensional scaling ordination and selected environmental characteristics

ᵇ Mantel test of relationship between bacterial communities (summer) and environmental characteristics

c Pearson’s correlation coefficient

d Standardized Mantel statistic
Table 3.4  Sum of the PLFAs of different taxonomic groups and ratios across the soil ages.

<table>
<thead>
<tr>
<th>PLFA (nmol g^{-1} soil)</th>
<th>105y</th>
<th>155y</th>
<th>210y</th>
<th>450y</th>
<th>845y</th>
<th>1475y</th>
<th>2385y</th>
<th>3210y</th>
<th>4010y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total PLFA (Summer)</td>
<td>426</td>
<td>837</td>
<td>1089</td>
<td>768</td>
<td>432</td>
<td>394</td>
<td>722</td>
<td>551</td>
<td>767</td>
</tr>
<tr>
<td>Total PLFA (Winter)</td>
<td>95</td>
<td>323</td>
<td>348</td>
<td>254</td>
<td>89</td>
<td>371</td>
<td>171</td>
<td>150</td>
<td>639</td>
</tr>
<tr>
<td>Fungi (Summer)</td>
<td>73</td>
<td>127</td>
<td>187</td>
<td>122</td>
<td>35</td>
<td>65</td>
<td>120</td>
<td>101</td>
<td>156</td>
</tr>
<tr>
<td>Fungi (Winter)</td>
<td>13</td>
<td>31</td>
<td>40</td>
<td>23</td>
<td>25</td>
<td>83</td>
<td>35</td>
<td>42</td>
<td>168</td>
</tr>
<tr>
<td>Gram negative Bacteria (Summer)</td>
<td>81</td>
<td>182</td>
<td>246</td>
<td>151</td>
<td>68</td>
<td>83</td>
<td>168</td>
<td>131</td>
<td>171</td>
</tr>
<tr>
<td>Gram negative Bacteria (Winter)</td>
<td>10</td>
<td>58</td>
<td>78</td>
<td>49</td>
<td>29</td>
<td>132</td>
<td>63</td>
<td>44</td>
<td>185</td>
</tr>
<tr>
<td>Gram Positive Bacteria (Summer)</td>
<td>70</td>
<td>148</td>
<td>205</td>
<td>175</td>
<td>115</td>
<td>85</td>
<td>150</td>
<td>99</td>
<td>148</td>
</tr>
<tr>
<td>Gram Positive Bacteria (Winter)</td>
<td>24</td>
<td>77</td>
<td>96</td>
<td>74</td>
<td>13</td>
<td>57</td>
<td>29</td>
<td>21</td>
<td>99</td>
</tr>
<tr>
<td>Bacteria (Summer)</td>
<td>152</td>
<td>331</td>
<td>451</td>
<td>326</td>
<td>184</td>
<td>168</td>
<td>318</td>
<td>230</td>
<td>320</td>
</tr>
<tr>
<td>Bacteria (Winter)</td>
<td>35</td>
<td>135</td>
<td>174</td>
<td>123</td>
<td>41</td>
<td>189</td>
<td>92</td>
<td>65</td>
<td>283</td>
</tr>
<tr>
<td>Fungal/Bac Ratio (Summer)*</td>
<td>0.51</td>
<td>0.39</td>
<td>0.42</td>
<td>0.28</td>
<td>0.20</td>
<td>0.40</td>
<td>0.39</td>
<td>0.41</td>
<td>0.48</td>
</tr>
<tr>
<td>Fungal/Bac Ratio (Winter)*</td>
<td>1.00</td>
<td>0.20</td>
<td>0.25</td>
<td>0.27</td>
<td>0.62</td>
<td>0.44</td>
<td>0.41</td>
<td>0.65</td>
<td>0.60</td>
</tr>
<tr>
<td>SAT/MONO (Summer)*</td>
<td>0.82</td>
<td>1.02</td>
<td>0.97</td>
<td>13.17</td>
<td>27.19</td>
<td>12.22</td>
<td>1.22</td>
<td>1.70</td>
<td>0.89</td>
</tr>
<tr>
<td>SAT/MONO (Winter)*</td>
<td>8.09</td>
<td>7.05</td>
<td>3.13</td>
<td>7.13</td>
<td>0.43</td>
<td>0.49</td>
<td>0.47</td>
<td>0.64</td>
<td>0.59</td>
</tr>
<tr>
<td>cyc/precursor ratio (Summer)*</td>
<td>0.96</td>
<td>1.53</td>
<td>2.11</td>
<td>9.80</td>
<td>27.15</td>
<td>15.95</td>
<td>2.75</td>
<td>9.53</td>
<td>2.28</td>
</tr>
<tr>
<td>cyc/precursor ratio (Winter)*</td>
<td>0.00</td>
<td>0.79</td>
<td>18.60</td>
<td>10.72</td>
<td>3.28</td>
<td>2.64</td>
<td>2.87</td>
<td>2.71</td>
<td>2.99</td>
</tr>
</tbody>
</table>

(* denotes ratios)
Table 3.5  Diversity indices of the 16S rRNA gene sequences.

<table>
<thead>
<tr>
<th>Diversity Index</th>
<th>105y</th>
<th>155y</th>
<th>210y</th>
<th>450y</th>
<th>845y</th>
<th>1475y</th>
<th>2385y</th>
<th>3210y</th>
<th>4010y</th>
</tr>
</thead>
<tbody>
<tr>
<td>N(^b)</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
<td>4779</td>
</tr>
<tr>
<td>S(^c)</td>
<td>1969</td>
<td>2352</td>
<td>2259</td>
<td>950</td>
<td>1225</td>
<td>1027</td>
<td>674</td>
<td>977</td>
<td>815</td>
</tr>
<tr>
<td>Evenness</td>
<td>2.01</td>
<td>2.00</td>
<td>1.99</td>
<td>1.75</td>
<td>1.73</td>
<td>1.70</td>
<td>1.70</td>
<td>1.78</td>
<td>1.72</td>
</tr>
<tr>
<td>Richness (ace)</td>
<td>15557</td>
<td>27093</td>
<td>28193</td>
<td>6320</td>
<td>11929</td>
<td>7463</td>
<td>3042</td>
<td>7545</td>
<td>4433</td>
</tr>
<tr>
<td>Shannon</td>
<td>6.63</td>
<td>6.73</td>
<td>6.68</td>
<td>5.22</td>
<td>5.34</td>
<td>5.11</td>
<td>4.80</td>
<td>5.33</td>
<td>5.00</td>
</tr>
<tr>
<td>1/D(^d)</td>
<td>205</td>
<td>134</td>
<td>125</td>
<td>58</td>
<td>53</td>
<td>43</td>
<td>45</td>
<td>72</td>
<td>52</td>
</tr>
<tr>
<td>Chao 1</td>
<td>7572</td>
<td>12736</td>
<td>10863</td>
<td>3253</td>
<td>5045</td>
<td>3423</td>
<td>1813</td>
<td>3671</td>
<td>2380</td>
</tr>
</tbody>
</table>

(with normalized sequence numbers using random sub-samples)

\(^a\) Calculations based on the Operational Taxonomic Units (OTU) formed at an evolutionary distance of <0.03. The results shown here are from the summer samples.

\(^b\) Number of sequences in the library.

\(^c\) Number of OTU

\(^d\) Simpson’s reciprocal index.

+Number of sequences in the library.
Figure 3.1  Bray–Curtis ordination plot showing the relationship between soil ecosystem development and bacterial community composition

Notes: Freshly deposited beach sands were also sampled to assess the community composition of the source material that formed the eolian deposits of the dune soils. The 200 most abundant OTU were used for the ordination. OTU were formed using average neighborhood algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity). Error bars represent standard error (n=5). Percentages on each axis denote the amount of variability associated with each axis.

* Shows the p-value of the ANOVA done using the GLM procedure of SAS showing the interaction effect of the season and soil ages as factors with bacterial OTU distribution as dependant variable.
Figure 3.2  Bray –Curtis ordination plot showing (a) bacterial community composition and structure. (b) Change in plant species community.

Notes: For (a) Bray-Curtis ordination was used to analyze bacterial community structure based on relative proportion of 200 most abundant OTU. OTU were formed using average neighborhood algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity). For (b) grouping was done using Bray –Curtis ordination based on the presence-absence of 13 different plant species. Error bars represent standard error (n=5).
Figure 3.3  Relationship between percentage relative abundance of nine individual bacterial phyla across the chronosequence during ecosystem development (105 to 4010y).

Notes: Each point in the graph is the average (n=5) of the percentage abundance of each phyla (summer and winter) at each stage of development. Regression co-efficient and p-value for each phylum are shown.
Figure 3.4  Relationship between the relative abundance of bacterial phyla across the chronosequences

Notes: The data points on the graph are the average of percentage abundance from the summer and winter samples.
Figure 3.5  Relationship between closest cultural representatives from the most abundant OTU based on RDP agent analysis

Notes: 5 OTU that showed the highest log-linear relationship with the changing ecosystem were chosen and plotted. OTU were formed using average neighborhood algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity) and the abundance was calculated based on the number of sequences found in each OTU relative to the total number collected for that age.
Figure 3.6 Rarefaction curves of the 16S rRNA gene libraries.

Notes: OTU were formed using the average neighbor algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity).
Figure 3.7  

A) Hierarchical Cluster Analysis of the most abundant OTU  
B) Hierarchical Cluster Analysis of the total percentage-cover of 13 different plant species

Notes: Cluster analysis was done using Ward’s method calculating relative Euclidean distance. The distance axis represents the similarity index.
Figure 3.8  Plant succession shown as percentage change across soil development.
Figure 3.9 Differences in the structure of microbial community associated with summer and winter samples from Michigan chronosequences using Bray-Curtis ordination analysis of the PLFA profiles

Notes: Error bars represent standard error (n=5).
Figure 3.10  Bray-Curtis ordination plot of PLFA profiles from different seasons across the soil ages. a) Summer PLFAs b) Winter PLFAs

Notes: Error bars represent standard error (n=5).
Figure 3.11  Sum of bacterial and fungal PLFA concentrations across soil development.

a) Summer PLFAs.  b) Winter PLFAs
Notes: Letters denote significant difference (P< 0.05) between the years of development. Error bars represent standard error (n=5).
Figure 3.12  Change in fungal to bacterial PLFA ratios across soil development.

a) Summer PLFAs  b) Winter PLFAs
Notes: Letters denote significant difference (P< 0.05) between the years of development. Error bars represent standard error (n=5).
Figure 3.13  Change in Saturated/ Monosaturated PLFAs ratios across soil development.

a) Summer PLFAs  b) Winter PLFAs

Notes: Letters denote significant difference ($P< 0.05$) between the years of development. Error bars represent standard error ($n=5$).
Figure 3.14  Canonical correspondence analysis plot showing the relationship between the OTU distribution and environmental variable.

Notes: The entire environmental variables were significant (P< 0.05, Monte Carlo test of significance). Error bars represent standard error (n=5).
CHAPTER IV
THE STUDY OF BACTERIAL COMMUNITY DIVERSITY AND SUCCESSION IN DEVELOPING DUNES OF SOUTHEASTERN GEORGIA

Site description

These sand dunes are located along the coasts of Altamaha and Ohoopee river valley of southeast Georgia. They are chains of eolian deposited parabolic dunes ranging from 3.7 to 14.0 m high at the crests and are around 6 Km away from the modern river channel (Ivester et al., 2003). They are formed from a well-drained sandy parent material surrounded by coastal plain lowlands that are greater than 165,000 years old. The chronology of the dunes was estimated using optically stimulated luminescence (OSL) (Ivester et al., 2001). The Georgia dunes are basically fluvial deposits from the Altamaha river valley. They are formed during events of overbank flooding, shrinking and change of course of Altamaha River. The sand was entrained from the lateral sand bars, overbank flood deposits along the river valley to form the sand dunes. The dunes later migrated a shorter distance from their fluvial source due to westerly winds accumulating eolian deposits. Thus most of the dunes in these sites show a characteristic parabolic shape with their slip-face pointing the south-west. The dunes became stable once they were dominated by vegetation (Ivester, et al., 2001). A set of five differentially aged dunes (21K, 38K, 45K, 77K and 500K: K=1000 years) were chosen for the study. Thus, the
depositional age of the youngest dune was 21,000 years and the oldest dune was 500,000 years.

Objectives and hypothesis

Objectives

The objectives of this study were

1. To study the composition and diversity of bacterial communities as a function of soil ecosystem age along a set of developing sand chronosequences.

2. To understand the relative importance of various environmental factors that drive the microbial community structure and diversity in a developing ecosystem

Hypothesis

It is hypothesized that the bacterial communities will follow a pattern of change that would be associated with the dynamics of soil and ecosystem development.

Results

Changes in soil characteristics during ecosystem development

Soil extractable Ca, Mg and P showed significant log-linear correlation with the soil development (Table 4.1). Ca and Mg showed an accumulating trend across the developing soil chronosequences as the soil aged which was not characteristic of age related weathering. The oldest site (500,000 years) showed significant higher levels of Ca, Mg and P. However, total carbon and nitrogen levels showed an increasing trend (except for the youngest site) with the oldest site showing the highest accumulation levels
which was typical of a developing soil ecosystem. Soil organic matter measured as mass loss after ignition showed significant higher levels at older sites than the younger sites (Table 4.1). This was reflected in the mineralizable carbon levels accumulating as the soil develops with highest level at older site. Soil pH failed to show any trend despite the varying nutrient levels in the developing ecosystem. Na and K levels did not show any significant correlation with soil development. There was a significant difference in soil type between all the four riverine dunes (21K, 38K, 45K, and 77K) and the oldest site (500K). The younger site belonged to the Kershaw sand series (Sandy soils) while the 500K site belonged to Gilead series (loamy sand).

16S rRNA data

For bacterial community analysis 5 replicates soil samples were obtained from each of the 5 developmental chronosequences. Thus a total of 25 different libraries were produced from this study. Across all 25, we obtained a total of 43,521 quality sequences, with a range of 632 to 5357 sequences from each soil. The average read length of the gene sequences was approximately 262 bp (base pairs).

Bacterial diversity indices

Diversity indices failed to show any particular trend of change in diversity indices across the developing chronosequences (Table 4.2). However, excluding the 77K site, the Shannon diversity index showed increasing bacterial diversity with soil developmental age. It was also noted that the oldest soil (500K) showed highest bacterial richness (ace) and diversity (Shannon) than all the other sites (Table 4.2). This increased diversity was also supported by the rarefaction analysis. The rarefaction curves which predicts the
number of observed OTUs as a function of distance between sequences observed and the number of sequences sampled (Figure 4.5) showed that, except for the 77K site, the species richness was higher at the older sites than the younger sites. Rarefaction curves also failed to plateau (Figure 4.5) indicating that the diversity is not completely sampled at evolutionary distance of 0.03. The chao1 richness predictor values showed that only 48-56% of the OTU’s predicted by this estimator were actually observed (from Table 4.2). When grouped at the 97% similarity level, there were 2395 OTUs in the complete data set, which had 1296 singletons and 391 doubletons, averaging to 75 % to 9.5 % of sequence representation (Data not shown).

**Microbial community structure and ecosystem properties**

The distribution of 194 most abundant OTUs across the chronosequences showed significant relationship between the bacterial community change and ecosystem development (Figure 4.1). The bacterial community change at the oldest site (500K) was very different from the other sites as explained by the axis of maximum variability from Bray-Curtis ordination (Axis 1= 64 %). Although not all, some of the environmental factors that showed a significant log-linear correlation with age related development (Table 4.1) showed a significant correlation with the bacterial community distribution (Table 4.5). Mantel tests between bacterial community matrix and pair-wise edaphic factors matrices indicated significant correlations (P < 0.05) between bacterial community and Ca, Total carbon (%) and soil organic matter (r = 0.28, 0.19 and 0.48, respectively) (Table 4.5).
Change in plant community composition across the chronosequences

These dunes were dominated by turkey oaks, sand live oak, longleaf pine in canopy species and sparkleberry, lichen and mosses in understorey herbs. A vegetative succession along the chronosequences did not show a primary succession pattern as observed across the Michigan chronosequences (Table 4.3). However the oldest site (500K) had a different vegetation composition than the other ages dominated by oaks and holly (Table 3). A Bray-Curtis ordination plot showed that the older sites (77K and 500K) were structurally different in terms of the vegetative composition (Figure 4.2). Mantel test of a relationship between vegetation and bacterial communities along the developmental gradient was done and was statistically significant (p<0.009; r=0.94) (Table 4.5).

Phylogenetic affiliation of the 16S rRNA genes

Across all soils, the broad level of phylum classification (> 97% sequence similarity) showed that Actinobacteria were most abundant (35.9 %), followed by α-Proteobacteria (27.27 %) and Acidobacteria (12.2 %) covering 76 % of all the sequences (Figure 4.4). This trend was similar to what we observed in the Michigan chronosequences. The other less represented phyla were Cyanobacteria, β-Proteobacteria, Planctomycetes, Bacteroidetes, γ-Proteobacteria and Firmicutes with an overall abundance close to 4% or less. While looking at the relative abundance of individual phylum across the developing ecosystem (Figure 4.3), none of the phylogenetic groups, except Acidobacteria (P=0.02, r2 = 0.85), showed a significant trend of log-linear change across the chronosequences. However, Acidobacteria showed a decreased abundance across the soil developmental gradient (Figure 4.3). Although not
statistically significant, *Firmicutes* and *α-Proteobacteria* showed an increasing abundance across the soil ages (Figure 4.3).

**The change in microbial PLFA**

The molar percentage (mol%) distribution of the PLFA was influenced by the soil age and the older sites 77K and 500K were structurally different from the other sites along the axis that described 67% of variability in the dataset (Figure 4.6). Although not statistically significant, the older ages (77K and 500K) showed significant higher levels of total PLFA and bacterial PLFA (Table 4.4; Figure 4.7). The fungal bacterial ratio was not significantly different across the developing ecosystem (Figure 4.8). The stress markers measured as S/M (Saturated/ Monosaturated PLFAs) and cyc/precursor ratios (changes in cyclopropyl and their precursor PLFAs) were significantly lower at the older dunes (77K and 500K) indicating less nutrient stress in these sites (Table 4.4; Figure 4.8).

**Discussion**

The goal of this study was to understand the patterns of bacterial community change as a function of ecosystem development. We initiated a study along an eolian deposited sand-dune chronosequence near Altamaha river valley, Georgia with a hypothesis that the change in bacterial community composition patterns will follow the soil and vegetative changes across the ecosystem gradient. There was a significant change in bacterial community composition and diversity along the ecosystem development showing strong correlations with soil physico-chemical gradients and vegetation. Previous studies have shown that soil pH to be the dominant factor driving the bacterial community composition and a best predictor of soil bacterial biogeography.
across pH gradients (Lauber et al., 2009; Rousk et al., 2010a). This dominant effect of pH often masks the importance of other drivers shaping the microbial community structure. Our study site showed a relatively constant pH across the developmental gradient and offers an opportunity to understand other potentially important factors that shape soil microbial community structure. Soil pH was not significantly different between the youngest (21 K) and oldest site (500K) but the largest change in bacterial community diversity occurred between those sites (Table 4.2). This suggests that factors other than pH could have been an important factor in driving the microbial community structure.

Several soil factors showed strong log-linear relationship with the soil age along the development gradient. Soil extractable Ca, Total carbon and SOM also showed a consistent correlation with the bacterial community distribution. However, the highest variability exhibited by the 500K site from the other sites would have been the reason for this significant effect, especially of extractable Ca and SOM (Table 4.1). Cluster analyses of the bacterial community distribution showed that the most abundant OTU of the younger ages (21K, 38K, 45K and 77K) clustered together and were different from the oldest site (500K) which clustered separately (Figure 4.9). It had to be addressed that the 500 K site had a different parent material than the other sites and the younger sites belonged to the Kershaw sand series (Sandy soils) while the 500K site belonged to Gilead series (loamy sand). It has been previously shown that soil type which determines the soil physiochemical properties could be a primary determinant with an overriding effect on soil microbial community composition compared to other factors such as aboveground vegetation (Girvan et al., 2003).
The PLFA analysis failed to show a pattern in total bacterial, total fungal and total PLFA. The ratio of fungi to bacteria also did not show any trend. Previous studies have shown that the fungal to bacteria ratio is often related to productivity (Zeller et al., 2001; Moore-Kucera et al., 2008). Since the overall productivity at these sandy sites is low, little variation might be expected in the ratio of fungi to bacteria. But there was a significant lower levels of stress indicators such as S/M ratio (Saturated/Monounsaturated PLFAs) and cyc/precursor ratio (changes in cyclopropyl and their precursor PLFA ratio) along the older sites (77K and 500K). These stress markers are good indicators of nutrient and water stress in soil microbial communities (Moore-Kucera et al., 2008). Significant higher levels of mineralizable carbon and total carbon available for the microbes at these sites (Table 4.1) could explain the lower levels of stress markers. Carbon availability which is often viewed as a surrogate measure for energy and thus the total biomass a system can support would have influenced the microbial composition of different organisms that differ in their ability to tolerate or take advantage of the available carbon. The total fungal PLFA were higher at older sites (77K and 500K) than the younger sites (Table 4.4) indicating higher fungal abundance at the older sites possibly due to the greater ability of fungi to degrade recalcitrant materials such as lignin (Zeller et al., 2001). Although PLFA analysis only gives a broad level of community characterization, they are useful tool to detect structural alterations (Frostegård et al., 1996; Frostegård et al., 2011).

The results suggest that microbial community composition is most strongly controlled by soil carbon availability. The total carbon availability and SOM showed high correlations with the bacterial community distribution (Table 4.5). However, the low
levels of carbon and nitrogen availability across these oligotrophic ecosystems indicate that plant distribution could be an important driver of physical and chemical soil properties during succession (Kuramae et al., 2011). Lower levels of resources especially available nitrogen would expect to favor strong interactions between plant and bacteria along the developmental gradient. However, except for Wax myrtle that showed highest abundance across the oldest site (500 K), there were no other nitrogen fixers observed across these chronosequences (Table 4.3). Investigating the closest cultural representatives of each of the five dominant OTUs that showed the highest abundance across the soil development gradient showed that two out of the five OTUs were nitrogen fixers. But one of the nitrogen fixers is a free-living organism (*Azospirillum halopraeferens*). These results suggest that while the presence of plants is likely to have strong effects on microbial biomass, microbial community composition may be more strongly regulated by soil properties such as carbon availability.
Table 4.1  Mehlich-3 extractable soil cations and selected soil properties from the mineral soil across the chronosequence.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Ca</th>
<th>K</th>
<th>Mg</th>
<th>P</th>
<th>pH</th>
<th>Total C (%)</th>
<th>SOM (%)</th>
<th>Mineralizable C (µg/g)</th>
<th>Total N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21000</td>
<td>40.66 a</td>
<td>29.65 a</td>
<td>6.68 a</td>
<td>8.87 a</td>
<td>4.19 a</td>
<td>1.06 ab</td>
<td>0.23 a</td>
<td>109 a</td>
<td>0.04 a</td>
</tr>
<tr>
<td>38000</td>
<td>51.44 a</td>
<td>21.92 a</td>
<td>6.86 a</td>
<td>2.03 a</td>
<td>4.05 a</td>
<td>0.51 a</td>
<td>0.2 a</td>
<td>123 a</td>
<td>0.01 a</td>
</tr>
<tr>
<td>45000</td>
<td>51.04 a</td>
<td>21.80 a</td>
<td>8.97 a</td>
<td>3.81 a</td>
<td>3.97 ab</td>
<td>0.70 ab</td>
<td>0.2 a</td>
<td>149 a</td>
<td>0.03 a</td>
</tr>
<tr>
<td>77000</td>
<td>51.40 a</td>
<td>24.56 a</td>
<td>11.81 a</td>
<td>7.28 a</td>
<td>3.50 b</td>
<td>1.11 b</td>
<td>0.17 a</td>
<td>167 a</td>
<td>0.04 a</td>
</tr>
<tr>
<td>500,000</td>
<td>351.64 b</td>
<td>51.43 b</td>
<td>60.92 b</td>
<td>117.27 b</td>
<td>4.48 a</td>
<td>2.55 c</td>
<td>0.9 b</td>
<td>401 b</td>
<td>0.13 b</td>
</tr>
</tbody>
</table>

r-value | 0.93 | 0.84 | 0.95 | 0.92 | 0.39 | 0.88 | 0.9 | 0.98 | 0.9 |

^ Soil properties with significant log-linear correlation to soil age (p<0.05) are shown with bold r-values. Na did not show a significant correlation with the soil ages.

^Means within columns followed by the same letter are not significantly different between them at P <0.05
Table 4.2  Diversity indices of the 16S rRNA gene sequences.

<table>
<thead>
<tr>
<th>Diversity Index*</th>
<th>21 K</th>
<th>38 K</th>
<th>45 K</th>
<th>77 K</th>
<th>500 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>N+</td>
<td>6673</td>
<td>6827</td>
<td>8876</td>
<td>13123</td>
<td>8022</td>
</tr>
<tr>
<td>S#</td>
<td>629</td>
<td>645</td>
<td>790</td>
<td>845</td>
<td>867</td>
</tr>
<tr>
<td>Evenness</td>
<td>1.61</td>
<td>1.69</td>
<td>1.63</td>
<td>1.50</td>
<td>1.68</td>
</tr>
<tr>
<td>Goods Coverage</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.97</td>
<td>0.94</td>
</tr>
<tr>
<td>Richness (ace)</td>
<td>1854</td>
<td>1636</td>
<td>2391</td>
<td>2145</td>
<td>2474</td>
</tr>
<tr>
<td>Shannon</td>
<td>4.51</td>
<td>4.74</td>
<td>4.73</td>
<td>4.40</td>
<td>4.93</td>
</tr>
<tr>
<td>1/D&lt;sup&gt;+&lt;/sup&gt;</td>
<td>39</td>
<td>50</td>
<td>50</td>
<td>31</td>
<td>46</td>
</tr>
<tr>
<td>Chao 1</td>
<td>1216</td>
<td>1151</td>
<td>1633</td>
<td>1514</td>
<td>1536</td>
</tr>
</tbody>
</table>

*Calculations based on the Operational Taxonomic Units (OTUs) formed at an evolutionary distance of <0.03.
+Number of sequences in the library.
#Number of OTUs
K = 1000 years
Table 4.3  Total percent ground and canopy cover of different plant species across the chronosequences

<table>
<thead>
<tr>
<th>Common Name</th>
<th>Scientific name</th>
<th>21 K</th>
<th>38 K</th>
<th>45 K</th>
<th>77 K</th>
<th>500 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweetgum</td>
<td>Liquidambar styraciflua</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Waxmyrtle</td>
<td>Myrica cerifera</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Southern Red Oak</td>
<td>Quercus falcata</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Hickory</td>
<td>Carya sp.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Water Oak</td>
<td>Quercus nigra</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td>Laurel Oak</td>
<td>Quercus hemisphaerica</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>76</td>
<td>10</td>
</tr>
<tr>
<td>American Holly</td>
<td>Ilex opaca</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td>Sand Live Oak</td>
<td>Quercus virginiana var. geminata</td>
<td>81</td>
<td>81</td>
<td>48</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Slash Pine</td>
<td>Pinus elliotii</td>
<td>43</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Longleaf Pine</td>
<td>Pinus palustris</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Post Oak</td>
<td>Quercus stellata</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Turkey Oak</td>
<td>Quercus laevis</td>
<td>19</td>
<td>10</td>
<td>24</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Red Bay</td>
<td>Persea borbonia</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Sparkleberry</td>
<td>Vaccinium arboreum</td>
<td>14</td>
<td>5</td>
<td>5</td>
<td>57</td>
<td>19</td>
</tr>
<tr>
<td>Grass Sp</td>
<td>Panicum virgatum</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Catbrier</td>
<td>Smilax sp.</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Saw Palmetto</td>
<td>Serenoa repens</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Spreading Pricklypear</td>
<td>Opuntia humifusa</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reindeer Lichen</td>
<td>Cladina sp.</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Spanish Moss</td>
<td>Tillandsia usneoides</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>14</td>
</tr>
</tbody>
</table>

Notes: Each value represents the percentage cover recorded in a 500 sq. feet transect.
Table 4.4  Sum of the PLFAs of different taxonomic groups and ratios across the soil ages

<table>
<thead>
<tr>
<th>PLFA (nmol g⁻¹ soil)</th>
<th>21 K</th>
<th>38 K</th>
<th>45 K</th>
<th>77 K</th>
<th>500 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total PLFA</td>
<td>333</td>
<td>256</td>
<td>283</td>
<td>554</td>
<td>573</td>
</tr>
<tr>
<td>Total Fungi</td>
<td>11</td>
<td>23</td>
<td>20</td>
<td>31</td>
<td>28</td>
</tr>
<tr>
<td>Total G-ve</td>
<td>43</td>
<td>29</td>
<td>41</td>
<td>104</td>
<td>108</td>
</tr>
<tr>
<td>Total G +ve</td>
<td>146</td>
<td>70</td>
<td>105</td>
<td>201</td>
<td>208</td>
</tr>
<tr>
<td>Total Bacteria</td>
<td>188</td>
<td>99</td>
<td>146</td>
<td>305</td>
<td>316</td>
</tr>
<tr>
<td>Fungal/Bac Ratio*</td>
<td>0.08</td>
<td>0.25</td>
<td>0.18</td>
<td>0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>SAT/MONOSAT*</td>
<td>28.43</td>
<td>45.63</td>
<td>32.07</td>
<td>11.16</td>
<td>5.01</td>
</tr>
<tr>
<td>cyc/precursor ratio*</td>
<td>37.93</td>
<td>24.75</td>
<td>36.98</td>
<td>26.20</td>
<td>10.19</td>
</tr>
</tbody>
</table>

Notes: * denotes ratios
Table 4.5 Correlations between the bacterial community distribution and selected environmental characteristics across the Georgia chronosequences.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mantel Test $^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r^b$</td>
</tr>
<tr>
<td>Ca</td>
<td>0.28</td>
</tr>
<tr>
<td>K</td>
<td>0.02</td>
</tr>
<tr>
<td>Mg</td>
<td>0.20</td>
</tr>
<tr>
<td>P</td>
<td>0.21</td>
</tr>
<tr>
<td>pH</td>
<td>0.06</td>
</tr>
<tr>
<td>Total N (%)</td>
<td>0.14</td>
</tr>
<tr>
<td>Total C (%)</td>
<td>0.19</td>
</tr>
<tr>
<td>SOM (%)</td>
<td>0.48</td>
</tr>
<tr>
<td>Mineralizable C</td>
<td>0.20</td>
</tr>
<tr>
<td>Vegetation (% Cover)</td>
<td>0.94</td>
</tr>
</tbody>
</table>

$^a$ Mantel test of relationship between bacterial communities and environmental characteristics  
$^b$ Standardized Mantel statistic
Figure 4.1    Bray–Curtis ordination plot showing relationship between soil ecosystem development and bacterial community composition and structure

Notes: 194 most abundant OTUs were used for the ordination. OTUs were formed using average neighborhood algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity). Error bars represent standard error (n=5).
Figure 4.2  Bray-Curtis ordination showing relationship between the soil ecosystem development and change in percent cover of 20 plant species.
Figure 4.3  Relationship between percentage relative abundance of nine individual bacterial phyla across the chronosequence during ecosystem development

Notes: Each point in the graph is the average (n=5) of the percentage abundance of each phyla at each stage of development. Regression co-efficient and p-value for each phylum are shown.
Figure 4.4  Relationship between the relative abundance of bacterial phyla across the chronosequences.
Figure 4.5 Rarefaction curves of the 16S rRNA gene libraries.

Notes: OTU were formed using the average neighbor algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity).
Figure 4.6 Differences in the structure of microbial community from Georgia chronosequences using Bray-Curtis ordination analysis of the molar percentages of 39 PLFA profiles

Notes: Error bars represent standard error (n=5).
Figure 4.7  Total PLFA and sum of bacterial and fungal PLFA concentrations across soil development.

Notes: Each point denotes the average of five replications across each age. Letters denote significant difference (P < 0.05) between the years of development.
Figure 4.8  Change in a) Fungal/Bacterial = fungal to bacterial PLFA ratios across the ages. b) S/M ratio= Saturated/ Monosaturated PLFAs. c) cyc/precursor ratio= changes in cyclopropyl and their precursor PLFA ratio.

Notes: Each point denotes the average of five replications across each age. Letters denote significant difference ($P< 0.05$) between the years of development.
Figure 4.9   Hierarchical Cluster Analysis of the most abundant OTU

Notes: Cluster analysis was done using Ward’s method calculating relative Euclidean distance. The distance axis represents the similarity index.
CHAPTER V

BIOGEOGRAPHY OF BACTERIAL COMMUNITY DIVERSITY AND

SUCCESSION

Introduction

Microorganisms are considered the most diverse and abundant living organisms on earth (Gans et al., 2005). Despite their vast diversity, the distribution of microbes on small and large scales is poorly understood (Ranjard et al., 2010). Most of the studies on microbial diversity have focused on studying the diversity at a particular location, soil type or landscape and have ignored understanding the microbial community assembly across spatial scales (Martiny et al., 2006; Ranjard et al., 2010). Microbial biogeography is an emerging field of microbial ecology that studies the distribution of microbial diversity across space and time. It not only aims to reveal the relationship between the microorganisms and their habitat but also to determine the environmental factors that select or maintain these organisms in those habitats (Martiny et al., 2006; Fierer, 2008). Studying such relationships would provide more insights on relative influence of environmental and evolutionary changes that determine the structure of microbial communities (Green et al., 2006; Green et al., 2008).

There have been varying views about microbial biogeography over the past decade. The ubiquity model proposed by Beijerinck was considered a paradigm of microbial biogeography until the advent of modern molecular techniques. According to
this theory, due to their small size, abundance and high dispersal rate, micro-organisms have cosmopolitan distribution. In essence, any bacteria species can be found anywhere in the world owing to their small size and higher numbers. This view is supported even in a recent literature (Fenchel et al., 2004) where the authors argue that any organism less than 1 mm in size is likely to present everywhere due to their unlimited capability of long distance dispersal whereas larger organisms have constrained distribution. The underlying assumption in their argument is that due to the higher local abundance, microbes have the capability to successfully colonize any remote location by chance (Fenchel et al., 2004; Martiny et al., 2006). In contrast to the cosmopolitan theory, some researchers proposed the endemicty model. Although it has been mentioned in the literature (Martiny et al., 2006; Hanson et al., 2012) this model is least supported. Most of the studies supporting this theory rely on limited dispersal and local adaptation. For example extremophiles (microbes that inhabit extreme habitat) and obligate symbionts (eg. *Rhizobium*) have much lesser potential for universal dispersal than the non-extremophiles and non-symbionts (Whitaker et al., 2003). A conservative middle-ground model supports the idea that microbes can exist ubiquitously as long as the environment is suitable for their survival. The basic hypothesis of these researches that focused on this idea were based on the Bass-Becking’s statement “everything is everywhere; the environment selects” (Green et al., 2006). According to this idea habitats that have similar habitat and physical conditions would support similar microbial communities.

Recently, many theories that are established in macro-organism community assemblies are being tested on microbial community assemblies. The main aim of these studies are to find if there are specific factors that determine the abundance of certain
communities in certain environment and what regulates them (Martiny et al., 2006; Horner-Devine et al., 2007). These specific factors are divided into two major groups, environmental factors and geographic factors. According to these theories either environmental or geographical factors or factors from both groups determine the biogeography of bacteria (Martiny et al., 2006). So it is indispensable to understand which environmental factors most contribute to the structure and diversity of bacterial communities in soil on very broad geographic scales.

**Objectives and hypothesis**

**Objectives**

- To assess the biogeographic distribution and diversity of bacterial communities across two US soils and their relationship with soil characteristics.

**Hypothesis**

It was hypothesized that the geographic distribution of bacterial communities will be more closely related to soil physico-chemical characteristics than the physical distance between soils.

**Results**

**Microbial phylogenetic distribution**

Pyrosequencing generated 119,588 quality short-read bacterial rRNA sequences from 125 samples across both the sites, resulting on average 956 sequences per sample. At a broad level of phylum-class classification, the most abundant bacterial phyla across the chronosequences were *Actinobacteria*, *Proteobacteria* (Including α, β and γ) and *Acidobacteria* covering 93 % of all the sequences with 42%, 38% and 13% individual
contribution, respectively (Figure 5.4). The remaining phyla present within the libraries, such as *Bacteroidetes, Cyanobacteria, Firmicutes* and *Planctomycetes*, all comprised less than 7% of the clones. None of the bacterial phyla showed a log-linear trend across the developing gradient, except *Bacteroidetes*, that showed a decreasing log-linear relationship across the ecosystem gradient with a higher abundance at younger soils and the decreasing abundance as the soil gets older (Figure 5.5).

**Change in bacterial community structure and ecosystem properties**

The distribution of 100 most abundant OTU across the chronosequences from Michigan and Georgia samples showed strong relationship between the bacterial community change and ecosystem development (Figure 5.1). The ordination results showed that Axis 1 explained 78% of the variation in bacterial community composition while Axis 2 explained 6% of the variation (Figure 5.1). The spread of the data was dominated by axis 1 which indicated that the difference between the sites is much less than the difference within the sites. Soil extractable Ca, Mg and K showed significant log-linear correlation with the soil development (Table 5.1). Ca and Mg showed a decreasing trend across the developing soil chronosequences as the soil aged which was characteristic of age related weathering. This effect was reflected in the soil pH which started from near-neutral (7.6) and becoming increasingly acidic (3.5) as the soils got older due to the leaching of the carbonates. Carbon and Nitrogen levels showed an increasing trend with the oldest site showing the highest accumulation levels which was typical of developing soil ecosystem (Table 5.1).

In order to understand the relative importance of the environmental variables, a Canonical Correspondence Analysis (CCA) was performed on those variables that were
shown to influence the bacterial community composition. This technique has been shown to be useful to identify the best predictor that influences soil microbial communities in ecological studies (Mitchell et al., 2010). CCA analysis gives more weight to the community structure that is more related to the environmental variables (McCune et al., 2002). Thus CCA analysis was performed on the bacterial OTU distribution using the important underlying environmental variables that shapes the community structure that were measured using a log-linear relationship. Soil Ca, Mg, pH, Total Carbon (%) and total nitrogen (%) were chosen for the ordination. CCA significantly explained 34 % of the OTU–environment relationship across the first two canonical axes. The joint plot (Figure 5.2) shows that CCA grouped the ages similar to the other ordination techniques used in the study. In the joint plot the length and the direction of the environmental vectors indicate the strength of the environmental variables with the community. Variables Ca, Mg and pH showed a significant correlation along canonical axis 1 (Monte Carlo test of significance, p=0.001) while carbon and nitrogen percentages showed significant correlation along canonical axis 2 (Monte Carlo test of significance, p=0.001). Variables Ca, Mg and pH showed positive correlations ($r= 0.81$, $r=0.85$ and $r=0.98$, respectively) implying that they significantly decreased as the soil aged. Whereas, variables carbon and nitrogen (%) showed negative correlations ($r= -0.71$ and $r= -0.56$, respectively) implying that they significantly increased as the soil aged.

**Bacterial diversity indices**

In order to calculate diversity indices, OTUs were formed at $D = 0.03$ (about 97% sequence similarity). Based on the Shannon and Simpson’s reciprocal index, the bacterial diversity tended to decrease considerably across the chronosequence (Table 5.4). The
diversity declined from 78 to 43 (Simpson’s I/D) during 105 to 500,000 years of soil development. The chao1 richness predictor values showed that only 29-57% of the OTU’s predicted by this estimator were actually observed indicating that the diversity is not completely sampled at evolutionary distance of 0.03.

**Bacterial biogeography**

A simple Mantel test showed that there was no significant correlation between the bacterial community distribution and the geographic distance (Table 5.3). The standardized Mantel statistic ($r_m$) was not significant at 95 % confidence level using 999 permutations ($r_m = 0.13, P = 0.35$) indicating that the bacterial community assembly did not show a geographic pattern. In order to further understand this relationship a Mantel correlogram was developed to plot autocorrelation as a function of geographic distance (Cho et al., 2000; Sokal et al., 2008). To achieve this, geographic distances between the sites were divided into classes. This resulted in three classes: class 1 (0 – 1 Km), class 2 (1 – 50 Km) and class 3 (50 – 1700 Km). For each class, a n x n matrix was constructed containing zeroes for site pairs whose geographic distances fall within the class and ones for pairs that do not fall into that range class. Then for each distance class matrix, a simple mantel test was performed between the bacterial community distribution and the distance class matrices (Figure 5.6). For the first two distance classes there were no significant spatial autocorrelations observed for the distribution pattern of bacterial community structure ( p-value, class 1= 0.96 and class 2= 0.25) (Figure 5.6). However, for the third distance class (50- 1700 Km) there was weak negative correlation which was significant ($r_m = -0.19, < 0.001$).
A simple Mantel test also showed that there is significant correlation between bacterial community distribution and soil physico-chemical characteristics ($r_M = 0.59$, $p < 0.001$) (Table 5.3). However, controlling for variables is an important challenge in microbial biogeography studies when we try to make meaningful comparisons between the geographic distance and genetic distance, as they may make the comparisons more complex (Whittaker et al., 2001). So we performed a partial mantel statistic on our distance matrices. After controlling for geographical distances, a partial Mantel test showed that there is still a very high significant correlation between bacterial community distribution and soil physico-chemical characteristics ($r_M = 0.84$, $p = 0.001$) (Table 5.3). But, when the soil physico-chemical characteristics was included as a control matrix, there was no significant correlation between bacterial community distribution and the geographical distance ($r_M = 0.02$, $p = 0.41$) (Table 5.3).

**Discussion**

The broad goal of this experiment was to understand the biogeographic distribution and diversity of bacterial communities across two US soils and the factors that control it. We investigated whether the geographic distance or environmental conditions or both could be the important factors driving variation in bacterial communities at different spatial scales. We hypothesized that the geographic distribution of bacterial communities will be more closely related to local environmental variations (soil physico-chemical characteristics) than the physical distance between the soils.

As it turned out, the community similarities were not positively correlated with geographic distance ($r_M = 0.13$, $p = 0.35$). Thus the Mantel test shows there was no significant isolation by distance at a confidence level of 95%. However, when the
distance variance was partitioned into classes, there was a marginal correlation at a spatial scale of 50-1700 km. Our results showed that, at this scale, the community dissimilarity had a moderate negative correlation with the geographic distance (Figure 5.6). But, partial correlations of community distances and environmental conditions, keeping the geographic distance controlled, results in a stronger correlation ($r_M = 0.84$, $p = 0.001$). This was higher than the correlation observed due to the soil physico-chemical factors when the effect of the geographic distance was not controlled ($r_M = 0.59$, $p < 0.001$), indicating an increase genetic dissimilarities once the geographic isolation effect was removed.

Thus, the results suggest that local environmental conditions could have a stronger effect than the geographic distance, and could be a major contributor in shaping the bacterial community structure at smaller scales (Martiny et al., 2006). Previous studies at smaller geographic scales have shown similar effect on the microbial communities in which the environmental conditions was considered a major driving factor in shaping the variability in the community structure (Kuske et al., 2002; Horner-Devine et al., 2004). At intermediate geographic scale, studies have shown individual distance effect (Reche et al., 2005) and environmental effect (Green et al., 2006) on microbial communities. Other similar studies conducted at this intermediate scale have shown that both distance and environmental factors could be major determinants in shaping microbial communities (Ge et al., 2008). Similar to what we observed, Griffiths et al. (2011) showed that the correlation between community dissimilarity and the geographic distance disappeared as the soils got separated farther. Thus our results suggest that at a geographic scale of 50-1700 km, the geographic distance and
environmental conditions would have had different extents in shaping the bacterial community structure.

The results also showed that season had no effect on the bacterial community distribution (Figure 5.1). However, carbon and nitrogen seem to be important drivers of communities between the two sites along the axis 2 (that described 9% variation in the data among the sites) (Figure 5.2). The low variability between sites could be due to the similarity in soil characteristics. Both the sites were eolian deposited sand dunes which are nutrient poor with a sandy parent material (Table 5.2). However the 6% variation between the sites could be due to basic site difference that would be difficult to account for but could be important. A hierarchical cluster analysis of the microbial community composition broadly grouped the bacterial community (OTU) into two major clusters comprising of the younger dunes and older dunes. The older dunes were further grouped based on the location (Figure 5.3). This pattern was similar to the vegetation distribution across the older sites (at both Michigan and Georgia sites) that were dominated by mixed-pine forests (Table 5.2).

In conclusion, this study attempts to survey bacterial spatial patterns across two pristine US dune chronosequences which are approximately 1700 km apart. We show that local geochemical features could be a dominant factor in driving bacterial community structure, while geographic distance as a single factor could contribute to some community variation at a scale (50 – 1700 km). Thus, the results show that the bacterial abundance is spatially structured and could be more dependent on local filters such as soil characteristics than the global filters such as climatic factors or the presence of natural barriers. Hence supporting Bass-Becking’s idea that “everything is
everywhere; *the environment selects*” which implies that similar habitat and physical conditions would support similar microbial communities. However, for the results to be viewed as definitive we need more samples from more locations to increase the distance classes and to include more environmental variables.

Table 5.1  Mehlich-3 extractable soil cations and selected soil properties from the mineral soil across the chronosequences.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Ca</th>
<th>K</th>
<th>Mg</th>
<th>pH</th>
<th>Total C</th>
<th>Total N</th>
<th>Mineralizable &lt;sup&gt;b&lt;/sup&gt; C (µg/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>105</td>
<td>1314</td>
<td>16</td>
<td>110</td>
<td>7.6</td>
<td>0.49</td>
<td>0.01</td>
<td>76</td>
</tr>
<tr>
<td>155</td>
<td>762</td>
<td>20</td>
<td>156</td>
<td>7.1</td>
<td>0.42</td>
<td>0.02</td>
<td>121</td>
</tr>
<tr>
<td>210</td>
<td>622</td>
<td>19</td>
<td>104</td>
<td>5.7</td>
<td>0.49</td>
<td>0.02</td>
<td>169</td>
</tr>
<tr>
<td>450</td>
<td>141</td>
<td>22</td>
<td>20</td>
<td>3.6</td>
<td>0.39</td>
<td>0.02</td>
<td>134</td>
</tr>
<tr>
<td>845</td>
<td>109</td>
<td>24</td>
<td>12</td>
<td>3.7</td>
<td>0.10</td>
<td>0.01</td>
<td>129</td>
</tr>
<tr>
<td>1475</td>
<td>116</td>
<td>24</td>
<td>11</td>
<td>3.6</td>
<td>0.20</td>
<td>0.01</td>
<td>108</td>
</tr>
<tr>
<td>2385</td>
<td>137</td>
<td>25</td>
<td>13</td>
<td>3.6</td>
<td>0.13</td>
<td>0.02</td>
<td>127</td>
</tr>
<tr>
<td>3210</td>
<td>110</td>
<td>22</td>
<td>10</td>
<td>3.7</td>
<td>0.19</td>
<td>0.01</td>
<td>85</td>
</tr>
<tr>
<td>4010</td>
<td>108</td>
<td>24</td>
<td>8</td>
<td>3.5</td>
<td>0.26</td>
<td>0.02</td>
<td>153</td>
</tr>
<tr>
<td>21000</td>
<td>41</td>
<td>30</td>
<td>7</td>
<td>4.2</td>
<td>1.06</td>
<td>0.04</td>
<td>109</td>
</tr>
<tr>
<td>38000</td>
<td>51</td>
<td>22</td>
<td>7</td>
<td>4.0</td>
<td>0.51</td>
<td>0.01</td>
<td>123</td>
</tr>
<tr>
<td>45000</td>
<td>51</td>
<td>22</td>
<td>9</td>
<td>4.0</td>
<td>0.70</td>
<td>0.03</td>
<td>149</td>
</tr>
<tr>
<td>77000</td>
<td>51</td>
<td>25</td>
<td>12</td>
<td>3.5</td>
<td>1.11</td>
<td>0.04</td>
<td>167</td>
</tr>
<tr>
<td>500,000</td>
<td>352</td>
<td>51</td>
<td>61</td>
<td>4.5</td>
<td>2.55</td>
<td>0.13</td>
<td>401</td>
</tr>
<tr>
<td>r-value</td>
<td>0.59</td>
<td>0.70</td>
<td>0.53</td>
<td>0.54</td>
<td>0.69</td>
<td>0.66</td>
<td>0.60</td>
</tr>
</tbody>
</table>

<sup>a</sup> Soil properties with significant log-linear correlation to soil age (p<0.05). Na, P and soil organic matter (SOM) did not show a significant correlation with the soil ages. Means within columns followed by the same letter are not significantly different between them at P <0.05

<sup>b</sup> Results at the Michigan sites are from the summer samples alone
Table 5.2  Description of sampling sites

<table>
<thead>
<tr>
<th>Region</th>
<th>Site</th>
<th>Geographic Coordinates</th>
<th>Dominant Vegetation</th>
<th>Climate</th>
<th>Soil type and classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid-West</td>
<td>Wilderness State Park (MI)</td>
<td>45.724 N 84.940 W</td>
<td>Pine (<em>Pinus strobes, Pinus resinosa</em>)</td>
<td>Temperate</td>
<td>Sandy, frigid Mollic Endoaquents</td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South – East</td>
<td>Altamaha River Valley (GA)</td>
<td>31.691 N 81.787 W</td>
<td>Turkey oak (<em>Quercus laevis</em>) Pine (<em>Pinus palustrus</em>)</td>
<td>Sub-tropical</td>
<td>Sandy, Typic Quartzipsamments</td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3  Mantel test correlations between the bacterial community distribution and selected environmental characteristics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mantel Test a</th>
<th>Partial Mantel Test b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r_M</td>
<td>P-value</td>
</tr>
<tr>
<td>Soil physico-chemical characteristics</td>
<td>0.59</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Distance</td>
<td>0.13</td>
<td>0.35</td>
</tr>
</tbody>
</table>

a Mantel test of relationship between bacterial community similarity matrix, environmental characteristics and geographic distance. r_M Standardized Mantel statistic
b 1 Partial mantel test of relationship between bacterial community similarity matrix and environmental characteristics after controlling for geographical distance. 2 Partial mantel test of relationship between bacterial community similarity matrix and geographical distance after controlling for environmental characteristics
Table 5.4  Diversity indices of the 16S rRNA gene sequences.

<table>
<thead>
<tr>
<th>Diversity Index</th>
<th>105y</th>
<th>155y</th>
<th>210y</th>
<th>450y</th>
<th>845y</th>
<th>1475y</th>
<th>2385y</th>
<th>3210y</th>
<th>4010y</th>
<th>21 K</th>
<th>38 K</th>
<th>45 K</th>
<th>77 K</th>
<th>500K</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>8896</td>
<td>16921</td>
<td>7793</td>
<td>7221</td>
<td>7098</td>
<td>7869</td>
<td>6078</td>
<td>9412</td>
<td>4779</td>
<td>6673</td>
<td>6827</td>
<td>8876</td>
<td>13123</td>
<td>8022</td>
</tr>
<tr>
<td>S</td>
<td>2286</td>
<td>4360</td>
<td>2365</td>
<td>872</td>
<td>1180</td>
<td>1003</td>
<td>530</td>
<td>1063</td>
<td>566</td>
<td>634</td>
<td>634</td>
<td>790</td>
<td>848</td>
<td>855</td>
</tr>
<tr>
<td>Evenness</td>
<td>1.81</td>
<td>1.71</td>
<td>1.81</td>
<td>1.59</td>
<td>1.55</td>
<td>1.51</td>
<td>1.57</td>
<td>1.57</td>
<td>1.61</td>
<td>1.60</td>
<td>1.66</td>
<td>1.62</td>
<td>1.49</td>
<td>1.66</td>
</tr>
<tr>
<td>Richness (ace)</td>
<td>14399</td>
<td>34633</td>
<td>22001</td>
<td>6279</td>
<td>10571</td>
<td>7903</td>
<td>2467</td>
<td>6885</td>
<td>3286</td>
<td>1810</td>
<td>1589</td>
<td>2414</td>
<td>2247</td>
<td>2349</td>
</tr>
<tr>
<td>Shannon</td>
<td>6.1</td>
<td>6.2</td>
<td>6.1</td>
<td>4.7</td>
<td>4.7</td>
<td>4.5</td>
<td>4.3</td>
<td>4.8</td>
<td>4.4</td>
<td>4.5</td>
<td>4.7</td>
<td>4.7</td>
<td>4.4</td>
<td>4.9</td>
</tr>
<tr>
<td>1/D</td>
<td>78.1</td>
<td>74.3</td>
<td>79.8</td>
<td>39.1</td>
<td>33.5</td>
<td>28.8</td>
<td>31.0</td>
<td>45.9</td>
<td>34.9</td>
<td>37.4</td>
<td>44.0</td>
<td>46.7</td>
<td>29.8</td>
<td>43.6</td>
</tr>
<tr>
<td>Chao 1</td>
<td>7685</td>
<td>17920</td>
<td>9297</td>
<td>3050</td>
<td>4784</td>
<td>3534</td>
<td>1425</td>
<td>3323</td>
<td>1638</td>
<td>1251</td>
<td>1146</td>
<td>1604</td>
<td>1531</td>
<td>1477</td>
</tr>
</tbody>
</table>

\(^a\) Calculations based on the Operational Taxonomic Units (OTU) formed at an evolutionary distance of <0.03. (K = 1000 years)

\(^b\) Number of sequences in the library.

\(^c\) Number of OTU

\(^d\) Simpson’s reciprocal index.

K = 1000 years
Table 5.5  Pearson correlations between the bacterial community distribution and selected soil characteristics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pearson Correlation&lt;sup&gt;a&lt;/sup&gt;</th>
<th>r&lt;sup&gt;b&lt;/sup&gt;</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total C (%)</td>
<td>0.14</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Total N (%)</td>
<td>0.02</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>0.95</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Ca</td>
<td>0.84</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Mg</td>
<td>0.90</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Pearson correlations between the ordination scores of the axis explaining maximum variability (78 % from the original OTU distribution data) of the non-metric dimensional scaling ordination and selected environmental characteristics

<sup>b</sup> Pearson’s correlation coefficient
Figure 5.1 Bray–Curtis ordination plot showing relationship between soil ecosystem development and bacterial community composition and structure at both the sites (Michigan and Georgia).

Notes: 100 most abundant OTUs were used for the ordination. OTUs were formed using average neighborhood algorithm in MOTHUR at evolutionary distance of 0.03 (97% sequence similarity). Error bars represent standard error (n=5).
Figure 5.2  Canonical correspondence analysis plot showing the relationship between the OTU distribution and environmental variable.

Notes: All environmental variables were significant (P<0.05, Monte Carlo test of significance). Error bars represent standard error (n=5).
Figure 5.3  Hierarchical Cluster Analysis of the most abundant OTU

Notes: Cluster analysis done using Ward’s method and relative Euclidean distance. The distance axis represents the similarity index.
Figure 5.4  Relationship between the relative abundance of bacterial phyla across Michigan and Georgia chronosequences.
Figure 5.5  Relationship between percentage relative abundance of nine individual bacterial phyla across the both the chronosequences during ecosystem development (105 to 500,000y).

Notes: Each point in the graph is the average (n=5) of the percentage abundance of each phyla at each stage of development. Regression co-efficient and p-value for each phylum are shown.
Figure 5.6 Mantel correlogram for the spatial autocorrelation analysis of bacterial OTU distribution across all ages.

Notes: Standard Mantel statistics ($r_M$) are plotted against the distance classes. Closed symbols represent significant autocorrelation at 95% confidence level.

Distance classes:
- Class 1 (0 - 1 Km)
- Class 2 (1 – 50 Km)
- Class 3 (50 – 1700 Km)
CHAPTER VI

GENERAL CONCLUSION AND SIGNIFICANCE OF RESEARCH

This study was aimed at understanding the role of biotic and abiotic factors in structuring microbial community composition in a developing ecosystem. To fulfill our objectives we assessed the bacterial community composition and diversity from the environmental samples using the emerging state of art pyrosequencing technique. Given the ecological nature of our objectives, we needed a genomic technique that would help to deeply survey the diversity and composition of the bacterial communities. Pyrosequencing of the 16S rRNA genes allowed us to describe large number of sequences from two different locations, providing valuable assessment of overall bacterial diversity with coverage not yet reported in many soil microbial ecology studies.

At Michigan sties, both biotic and abiotic factors had significant influence on soil bacterial community. However, soil factors (pH), as opposed to plant species composition, seemed to be the main driver of microbial community structure. Although many studies have shown the influence of pH on soil microbial community, our study was unique in studying its effect along naturally developing ecosystems that are predominantly unperturbed and with minimal influence by recent human activity. However, this effect could be influenced by plant-driven pedogenesis which was shown by strong correlation observed between the plant distribution and bacterial community composition in our study. The largest change in pH observed across the chronosequence
was highly correlated with the plant species distribution. Another interesting observation is that the biggest change in the community composition occurred during early stages of ecosystem development. As the soil matured, stable soil bacterial community structure developed that mimicked the structural stability of aboveground plant communities. A similar effect was observed at Georgia sites, where plant community was relatively stable across the chronosequence that was highly correlated with stable bacterial communities associated with the climax community. This could be because of the extreme environmental conditions occurring in bare soils (young soils) compared to those in vegetated soils where there are fewer moisture and temperature fluctuations and more organic matter. Our findings suggest that bare soils provide a large range of microenvironments where different bacterial species can survive, resulting in more dynamic communities at younger sites as opposed to stable communities at older sites.

The overall dynamics of bacterial community composition was observed to be higher at Michigan site than the Georgia site. This again could be attributed to the stable pedogenesis processes at the Georgia sites. Both the chronosequences were dominated by wind-blown sediments dominated by sands. But higher temperature and rainfall amounts would have accelerated the weathering process at Georgia sites resulting in higher amount of weathered minerals. Michigan sites had greater abundance of weatherable minerals as evidenced by faster horizon development at these sites when compared to Georgia sites (Lichter, 1998). Thus, Georgia sites mimicked latter stages (climax) of ecosystem development before retrogression supporting stable communities (both plant and bacterial communities).
Our bacterial biogeography study showed the importance of local filters (environmental factors) in shaping up the soil bacterial community structure. Total bacterial abundance was relatively well structured spatially. However, the distribution of bacterial community structures was much less spatially autocorrelated (distance and community distribution) than the continuous variables (soil characteristics and community distribution). It has been shown that, while studying soil microbial biogeography, it is difficult to distinguish between the effects of dispersal limitation and the effects of environmental heterogeneity on community structure. Also designing studies that distinguish between the effect of dispersal limitation and environmental heterogeneity are difficult. Many previous studies that asserted the importance of dispersal limitation (geographic distance) on microbial biogeography failed to measure the direct influence of environmental heterogeneity (Green et al., 2006) or they had limitations in including the number of environmental factors that directly affected the bacterial community assembly (Reche et al., 2005). These discrepancies could be mitigated by studying microbial communities across identical habitats that are geographically isolated. Although it is difficult to find identical habitats or may be impossible, our study was one of its kinds and modestly aimed at understanding the impact of geographical isolation on microbial communities under similar habitats (nutrient poor developing systems showing succession patterns).

Our study also showed that there was evidence of nitrogen fixation (free-living nitrogen-fixation) in vegetated soils and there was low (but measurable) increase of total nitrogen as the soils matured and became dominated by vegetation (Figure 5.2). They would have had impacts on the microbial dynamics and this finding adds to the body of
evidence that reiterates the importance of asymbiotic nitrogen fixation in low-nutrient ecosystems. At Michigan sites, the sparsely vegetated young dunes supported more *Cyanobacteria* (autotrophs) than the vegetated old dunes. Approximately 30% of the beach sand bacteria (0 year) were derived from phylum *Cyanobacteria*. Although phylum *Cyanobacteria* constituted a relatively small portion of the whole bacterial community, they would have contribution to the initial nitrogen-fixation in these nutrient-poor ecosystems and helped in plant establishment. When the plant started colonizing the dunes, these autotrophs were quickly replaced by heterotrophs that would have efficiently used the increased availability of plant carbon inputs.

In summary, by using a high-throughput sequencing approach, our study reveals the trajectory of bacterial community change that was influenced by pedogenesis and distribution of aboveground plant communities. This study has a broader ecological impact on microbial function in relation to soil development, soil fertility, vegetation and seasonal changes. This study also highlights the importance of studying above-ground and below-ground feedbacks from a natural setting that is well-replicated.

**Future directions**

- This is the first study, to the best of our knowledge, providing a comprehensive overview of relative roles of environmental factors on microbial community assemblage in a natural setting with in-depth sequencing of microbial communities.
- Our results add to the growing body of evidence that reiterated the enormous metagenomic diversity of bacteria. We used a correlative approach to understand the above-ground and below-ground ecosystem
properties while the mechanisms controlling microbial distribution in these ecosystems still remains largely unknown. However, our study provides a framework to test a range of environmental gradients that might have a direct or indirect influence on soil microbial community assembly. They need to be tested over a wide range of ecosystems to better understand the mechanisms that structure microbial communities.

- Our study also concurs with recent studies that indentify the importance of plant community structure as important determinants of soil microbial dynamics. However, there is still lack of demonstrating direct links between plant and microbial communities. Studies focusing more on rhizosphere microbial communities and stable-isotope probing of plant inputs that enter microbial metabolism would shed more light on direct link between plant performance and microbial functionality.
REFERENCES


Chu, H., Fierer, N., Lauber, C.L., Caporaso, J., Knight, R., Grogan, P., 2010. Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. Environmental microbiology 12, 2998-3006.


