CrossMark

ENVIRONMENTAL MICROBIOLOGY

Bacterial Diversity Patterns Differ in Soils Developing in Sub-tropical and Cool-Temperate Ecosystems

Shankar G. Shanmugam[1,2] · Zenaida V. Magbanua[1,2] · Mark A. Williams[5] ·
Kamlesh Jangid[4] · William B. Whitman[3] · Daniel G. Peterson[1,2] · William L. Kingery[1]

Abstract Microbial diversity patterns have been surveyed in many different soils and ecosystems, but we are unaware of studies comparing similar soils developing from similar parent materials in contrasting climates. In 2008, developmental chronosequences with ages ranging from 105 to 500,000 years across Georgia (GA) and Michigan (MI) were studied to investigate how bacterial community composition and diversity change as a result of local environmental gradients that develop during pedogenesis. Geographic factors were studied between and within locations spanning two scales: (1) regionally between 0.1 and 50 and (2) ~1700 km apart. The diversity was surveyed using high-throughput pyrosequencing, and variance partitioning was used to describe the effects of spatial, environmental, and spatio-environmental factors on bacterial community composition. At the local scale, variation in bacterial communities was most closely related to environmental factors ($r_M = 0.59$, $p = 0.0001$). There were differences in bacterial communities between the two locations, indicating spatial biogeography. Estimates of bacterial diversity were much greater in MI (numbers of OTU, ACE, and Chao1) and remained 2–3× greater in MI than GA after removing the effect of soil properties. The large differences in diversity between geographically separated bacterial communities in different climates need further investigation. It is not known if the rare members of the community, which contributed to greater bacterial diversity in GA relative to MI, play an important role in ecosystem function but has been hypothesized to play a role in ecosystem resiliency, resistance, and stability. Further research on the link between bacterial diversity and spatial variability related to climate needs further investigation.

Keywords Bacterial diversity · Ecosystem development · Biogeography · 16S rRNA sequencing · Microbial ecology · Soil nutrients

The original version of this article was revised: The article title was mistakenly amended to "Bacterial diversity patterns differed in two developing ecosystems".

✉ Shankar G. Shanmugam
  sg383@msstate.edu

1   Department of Plant and Soil Sciences, Mississippi State University, 117 Dorman Hall, Mississippi State, MS 39762, USA

2   Institute for Genomics, Biocomputing & Biotechnology, Mississippi State University, Mississippi State, MS 39762, USA

3   Department of Microbiology, University of Georgia, 527 Biological Sciences, Athens, GA 30602, USA

4   Microbial Culture Collection, National Centre for Cell Science, Pune 411007, Maharashtra, India

5   College of Agriculture and Life Sciences, Virginia Polytechnic and State University, Horticulture, 301 Latham Hall, Blacksburg, VA 24060, USA

Introduction

Microorganisms are considered the most diverse and abundant organisms on earth [1]. Progress has been made in understanding the distribution of bacteria across several biogeographical scales, but questions still remain about the factors important to community structure and diversity [2]. For example, soil properties such as pH, carbon, and nitrogen have been shown to have strong associations with bacterial community structure, but fewer studies have attempted to compare biogeographically separated and climatically different soils that have comparable parent materials [3, 4]. 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 [2, 5]. 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 [5, 6]. Studying such relationships would provide more insights on relative influence of environmental and evolutionary changes that determine the structure of microbial communities [7, 8].

There have been varying views about microbial biogeography over the past decade. The ubiquity model proposed by Beijerinck [6] 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, microorganisms have cosmopolitan distribution. In essence, any bacterial species can be found anywhere in the world owing to their small size and high numbers. This view is supported even in a recent literature [9] where the authors argue that any organism less than 1 mm is likely to be 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 [5, 9, 10]. In contrast to the cosmopolitan theory, some researchers proposed the endemicity model. According to this model, the endemic taxon is restricted to a particular region or habitat type, and thus, their distribution across the landscape exhibits a nonrandom relationship. Although it has been mentioned in the literature [5, 11], this model is least studied. 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 (ectomycorrhizal fungal association with tree species) have much lesser potential for universal dispersal than the non-extremophiles and non-symbionts [12, 13]. A conservative middle-ground model supports the idea that microbes can exist ubiquitously as long as the niche 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, but, the environment selects" [7]. According to this idea, habitats that have similar environmental and physical conditions would support similar microbial communities.

Theories that have been established in macro-ecology have recently been tested in microbial systems, with the aim of describing the factors that determine population abundance patterns [5, 14]. These factors are often divided into two major groups: environmental and spatial. Though these groups can co-vary and care must be taken during interpretation, these factors have been shown to effect bacterial community structure [4, 5, 15, 16].

In this study, bacterial community composition and diversity were described across a series of dune chronosequences developing with similar parent materials but varying climatic conditions. The goal was to compare the environmental soil gradients created by pedogenesis in two climatically different (subtropical and cool-temperate) locations in the continental US. Habitat filtering along the pedogenic gradients were hypothesized to be a major determinant of bacterial community change (using 16S rRNA gene sequences), but that differences between locations would help to explain some of the biogeographic differences related to factors such as climate and vegetation. 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.

Materials and Methods

Site Description

Two sites were chosen for this study, one located in Emmet County of the lower peninsula of Michigan in Wilderness Park. The GA dunes are located in the Altamaha and Ohoopee river valleys of southeast Georgia. The reason for selecting these sites was that they were sandy with similar parent materials in different climate zones (Table 1). At MI study site, a series of beach-dune ridges form a series of approximately 108 eolian deposits that run parallel to Lake Michigan. Depositional ages of the parent materials from present day to approximately 4500 years [17] were derived 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 [17]. 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 in the inland [18]. A set of nine differentially aged dunes (105, 155, 210, 450, 845, 1475, 2385, 3210, and 4010 years) were sampled. A previous study [18] has shown patterns of primary succession with grasses and shrubs on younger dunes to mixed coniferous forests dominating the older dunes.

The GA dunes are located in the Altamaha and Ohoopee river valleys of southeast Georgia. Chains of eolian deposited parabolic dunes ranging from 3.7 to 14.0 m high at the crests are found ~6 km away from the modern river channel [19]. 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 procedure (OSL) [20]. The GA dunes are eolion, blown into ridges from fluvial

Table 1 Description of sampling sites

Region	Site	Geographic coordinates	Climate	Soil type and classification
Mid-West USA	Wilderness State Park (MI)	45.724 N 84.940 W	Temperate	Sandy, frigid Mollic Endoaquents
South-East USA	Altamaha River Valley (GA)	31.691 N 81.787 W	Sub-tropical	Sandy, Typic Quartzipsamments

deposits of the Altamaha river valley. The dunes were stabilized once they were dominated by vegetation [20]. A set of four differentially aged dunes (21, 38, 45, and 77 K; K = 1000 years) were sampled. The oldest soils were deposited during the Pleistocene interglacial period and were assigned an age of 500 K based on their geology and proximity to the present coastline. These deposits are sedimentary but represent sandy marine deposits that are much older formations than the dunes.

Soil and vegetative sampling

Five replicate soil samples were taken at an interval of 20 ft between each sampling spot across a 100-ft transect along the crest of each dune. From each sampling spot, 5–6 cores (0–10 cm depth, 5-cm diameter) were collected from the zone of dominant root activity (A-horizon) using a stainless steel soil corer. After sampling, soil from other zones were carefully removed, and A-horizon soil from the same plot were homogenized with and quickly transferred to a Whirlpak® bag. The sample bags were frozen immediately in cooler pack filled with dry ice. Upon arrival in the laboratory, soils were thawed for 25–30 min and homogenized through a 4-mm sieve; extraneous roots and organic materials were removed and stored at −80 °C. Plant sampling was conducted on the selected dune ridges along the chronosequence by measuring the plant species composition, tree density, and percentage canopy cover across the sampling area covering a strip of 5×20 m (100 m^2). The tree species composition was measured by counting the number of tree species within the sampling area and understorey species cover was measured at five random spots within the sampling area using a 1 m^2 quadrat. The tree species canopy cover was estimated by fitting the dbh (diameter breast height) measurement into a conifer crown radius model [21].

Soil Characteristics

Soil organic matter content was estimated by measuring the mass difference before and after ignition (560 °C). The mineralizable C was estimated by measuring the cumulative CO_2-C produced from 100-g soil during a 1-month soil incubation (20 °C). Soil pH was measured on 1:2 soil and 0.01 M $CaCl_2$ mixture. Soil extractable cations were analyzed according to the Mehlich-3 extraction protocol [22].

DNA Extraction and Pyrosequencing

Total DNA was extracted 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°. Small subunit bacterial rRNA gene fragments were amplified from an amount of DNA equivalent to that found in a 0.5 g of soil (varied with each sample) to appropriate size using 515R-M (5′-CCGCNGCKGCTGGCAC-3′) [23] and the sevenfold-degenerate primer 27F-YM + 3 [24]. 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, an 8-nt sample-specific barcode tag was attached between the A-adaptor and primer 515R-M to allow multiplexing and to eventually 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 using standard protocol [25]. Pyrosequencing generated 151,794 quality short-read bacterial 16S rRNA sequences from 75 (14 dunes × 5 replications) samples across both the sites. The gene fragments averaged to approximately 530 bp (base pairs) in length. Sequences generated in this study have been deposited

in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under accession number SRP091769.

Processing of 16S rRNA Gene Data

A two-step pipeline was established to analyze the 16S rRNA gene sequence data. Quantitative Insights into Microbial Ecology (QIIME) was used to quality trim the raw sequences for primers, chimeras and to sort them based on the barcodes [26, 27]. The denoised data was then passed through MOTHUR v1.22.0 [28], 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 [29, 30]. The non-redundant sequence dataset was then aligned using SILVA reference dataset (http://www.arb-silva.de/) [31]. Sequence reads were assigned (clustered) to Operational Taxonomic Units (OTUs) based on pairwise distances between all aligned sequences in MOTHUR. A 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. This level of DNA sequence similarity is typically used to assign sequences to the same taxa [28]. Finally, the phylogeny was assigned to representative sequences from OTUs using SILVA reference taxonomy [28].

The 16S rRNA gene sequence possessed a 530-bp average length and formed 15,131 OTUs in total at 97% sequence similarity ($D = 0.03$). Each soil age was represented by between 4779 and 11,248 sequences forming 537 to 2352 OTUs per soil age. The most abundant OTU was represented by 19,258 sequences, accounting for ~13% of the entire sequence data set. The top 10 and 100 OTUs represented 48 and ~93 % of the entire sequence data set, respectively. Thus, top 100 dominant OTUs, those with at least 0.05 % average abundance across all samples, were used for the ordination techniques. This was done to minimize the stress (a measure of poorness of fit between the ordination and measured ecological distances) while using distance-based ordination methods when the entire dataset with all OTUs were used.

Statistical and Multivariate Analyses

In order to visualize the differences in microbial community assemblages related to environmental factors across the pedogenic gradient, 16S rRNA gene-derived OTUs were analyzed using Bray-Curtis analysis and mantel tests. PC-ORD (MJM Software, Gleneden Beach, OR, USA) software and the Sorensen's metric were used to obtain the n-dimensional ordination, with Monte Carlo tests (1000 randomized runs) to determine the occurrence of significance differences between groups [32]. Mantel tests were used to determine correlations between bacterial community composition, site vegetation, and soil variables. Partial mantel test were performed to examine correlation between the degree of similarity in bacterial community composition with environmental and spatial variables, correlating the community with one variable while holding the third group constant. [15]. Correlations between soil and site variables were performed by using log-linear correlations and Pearson coefficients for multiple comparisons. The effect of site pedogenic age on the distribution of OTU was conducted by using Bray-Curtis-derived ordination scores and testing the statistical significance using a two-way general linear model (GLM; 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 [33]. Factors that significantly influenced community composition were used to construct a soil variables matrix for Variance partitioning analysis in CANOCO (ver 4.5) [34]. The significance of the correlations with each factor was evaluated through the Monte Carlo permutation test by applying 998 permutations.

Results

Plant Community Succession Across the Chronosequences

At MI site, changes in percentage cover of 13 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 (Table 2). 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 (Table 2). The early-succession species started disappearing and the mixed pine forest started establishing as soil ecosystem developed.

GA dunes were dominated by turkey oaks, sand live oak, and longleaf pine covering the canopy while sparkleberry, lichen, and mosses covered the dune ground. A vegetative succession along the chronosequences did not show a primary succession pattern as observed across the MI chronosequences. However, the oldest site (500 K) had a different vegetation composition than the other ages dominated by oaks and holly (Table 3).

Table 2 Total percent ground and canopy cover of different plant species across the Michigan chronosequences

Common name	Scientific name	105 years	155 years	210 years	450 years	845 years	1475 years	2385 years	3210 years	4010 years
Bear berry	*Arctostaphylos uva-ursi*	26	35	0	0	0	0	0	0	0
Little blue stem	*Schizachyrium scoparium*	9	15	0	0	0	0	0	0	0
Braken fern	*Pteridium aquilinum*	0	0	0	40	19	4	5	8	6
Winter green	*Gaultheria procumbens*	0	0	2	10	6	6	7	0	7
Fir sapling	*Abies balsamea*	0	0	22	0	0	0	0	0	0
Moss		0	0	27	0	0	1	0	0	1
Lowbush blueberry	*Vaccinium angustifolium*	0	0	4	5	6	0	10	6	0
Huckle berry	*Vaccinium myrtilloides*	0	0	0	0	22	11	37	31	0
Sassafrasas	*Sassafras albidum)*	0	0	0	0	1	9	0	0	0
Lichen		0	0	0	0	0	6	0	0	30
Fir	*Abies balsamea*	0	52	35	17	0	35	0	0	0
White pine	*Pinus strobus*	0	9	18	27	9	0	35	0	9
Red pine	*Pinus resinosa*	0	0	9	9	27	9	18	35	9

Change in Bacterial Community Structure and Ecosystem Properties

The distribution of the 100 most abundant OTU, indicative of bacterial community change, and showed patterns related to ecosystem development and site characteristics likely reflecting conditions in MI and GA (Fig. 1). Axis 1 ordinates explained 78% and Axis 2 at 6% of the variation in bacterial community composition (Fig. 1). The spread of the data was dominated by axis 1 which indicated that variation associated with age and pedogenesis within each site (MI and GA) explains a greater proportion of the variance than differences between locations. Several soil characteristics showed significant log-linear correlation with the soil development (Table 4). Except for the oldest site (500 K), Ca and Mg accumulation showed a decreasing trend across the developing soil chronosequences which was characteristic of age related weathering. This effect was also reflected in soil pH, which declined from near-neutral (7.6) to acidic (3.5) as the soils aged and weathered. Carbon and nitrogen levels showed an increasing trend with the oldest site showing the highest accumulation levels, an occurrence typical of developing soil ecosystem (Table 4). 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 (Fig. 2). This pattern was similar to the vegetation distribution across the older sites (at both MI and GA sites) that were dominated by mixed-pine forests (Tables 2 and 3).

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 of soil microbial communities [35]. Soil Ca, Mg, pH, total carbon (%) and total nitrogen (%) were chosen for the ordination based on the factors that had a significant Pearson correlation with the bacterial community distribution (see Table S1 for details). CCA significantly explained 70% of the OTU–environment relationship across the first two canonical axes. Axis 1 (62 %) explained more of the age-related pattern while axis 2 (8 %) structured the community based on the location (Fig. 3). The joint plot shows that CCA grouped the ages similar to the other ordination techniques used in the study. Variables Ca, Mg, and pH showed a significant correlation along canonical axis 1 (Monte Carlo test of significance, $p = 0.001$) while % C and N showed significant correlation along canonical axis 2 (Monte Carlo test of significance, $p = 0.001$). Variables Ca, Mg, and pH showed positive correlations along axis 1 ($r = 0.81$, $r = 0.85$, and $r = 0.98$, respectively), implying that they significantly decreased as the soil aged. Whereas % C and N were negatively correlated along axis 2 ($r = -0.71$ and $r = -0.56$, respectively), increasing significantly as the soil aged.

Table 3 Total percent ground and canopy cover of different plant species across the Georgia chronosequences

Common name	Scientific name	21 K	38 K	45 K	77 K	500 K
Hickory	*Carya sp.*	0	0	0	0	24
Water Oak	*Quercus nigra*	0	0	0	0	43
Laurel Oak	*Quercus hemisphaerica*	0	0	0	76	10
American Holly	*Ilex opaca*	0	0	0	5	43
Sand Live Oak	*Quercus virginiana var. geminata*	81	81	48	0	0
Slash Pine	Pinus elliotii	43	0	0	0	0
Longleaf Pine	*Pinus palustris*	0	0	5	14	0
Post Oak	*Quercus stellata*	0	19	0	5	0
Turkey Oak	*Quercus laevis*	19	10	24	0	0
Red Bay	*Persea borbonia*	0	0	0	14	14
Sparkleberry	*Vaccinium arboreum*	14	5	5	57	19
Grass Sp	*Panicum virgatum*	10	5	5	0	0
Catbrier	*Smilax sp.*	5	0	5	5	5
Saw Palmetto	*Serenoa repens*	1	5	0	0	0
Spreading Pricklypear	*Opuntia humifusa*	1	0	0	0	0
Reindeer Lichen	*Cladina sp.*	1	0	0	10	0
Spanish Moss	*Tillandsia usneoides*	0	0	5	0	14

Phylogenetic Affiliation of the 16S rRNA Genes

At a broad level of phylum-class classification, the most abundant bacterial phyla across the chronosequenes were *Actinobacteria*, *Proteobacteria* (including α, β, and γ), and *Acidobacteria* covering 93 % of all the sequences with 42, 38, and 13% individual contribution, respectively (Fig. 4). The remaining phyla present within the libraries, such as *Bacteroidetes, Cyanobacteria, Firmicutes*, and *Planctomycetes*, all comprised less than 7% of the sequences. None of the bacterial phyla showed a log-linear trend across the age gradient, except *Bacteroidetes*, that showed a decreasing log-linear relationship with a higher abundance in younger compared to older soils (Fig. 5). Although not statistically significant, *Acidobacteria* showed an increasing abundance in older soils (low pH) compared to young soils (high pH).

Fig. 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)
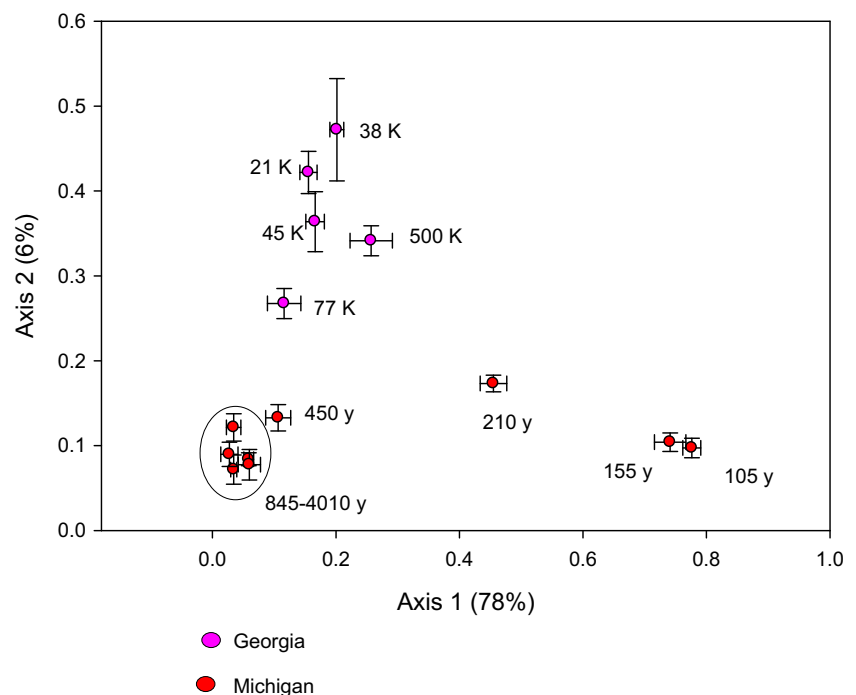
Table 4 Mehlich-3 extractable soil cations and selected soil properties from the mineral soil across the chronosequences

Location	Age (years)	Ca μg/g	K	Mg	pH	Total C (%)	Total N (%)	Mineralizable C (μg/g)
Michigan	105	1314	16	110	7.6	0.49	0.01	76
	155	762	20	156	7.1	0.42	0.02	121
	210	622	19	104	5.7	0.49	0.02	169
	450	141	22	20	3.6	0.39	0.02	134
	845	109	24	12	3.7	0.10	0.01	129
	1475	116	24	11	3.6	0.20	0.01	108
	2385	137	25	13	3.6	0.13	0.02	127
	3210	110	22	10	3.7	0.19	0.01	85
	4010	108	24	8	3.5	0.26	0.02	153
Georgia	21,000	41	30	7	4.2	1.06	0.04	109
	38,000	51	22	7	4.0	0.51	0.01	123
	45,000	51	22	9	4.0	0.70	0.03	149
	77,000	51	25	12	3.5	1.11	0.04	167
	500,000	352	51	61	4.5	2.55	0.13	401
	r value	0.59	0.70	0.53	0.54	0.69	0.66	0.60

Soil properties with significant log-linear correlation to soil age ($p < 0.05$)

Bacterial Diversity

To calculate diversity indices (using $D = 0.03$), the number of sequences per sample were normalized to 4779 by randomly subsampling a subset of sequences using QIIME scripts. This was done to avoid the possible influence of sample size on diversity estimates and normalized subset were used for further diversity measurements [36]. Based on the Shannon and Simpson's reciprocal index, the bacterial diversity tended to decrease considerably across the chronosequence (Table 5). The diversity declined from 205 to 47 (Simpson's 1/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 was not completely sampled at evolutionary distance of 0.03. Estimates of bacterial diversity were much
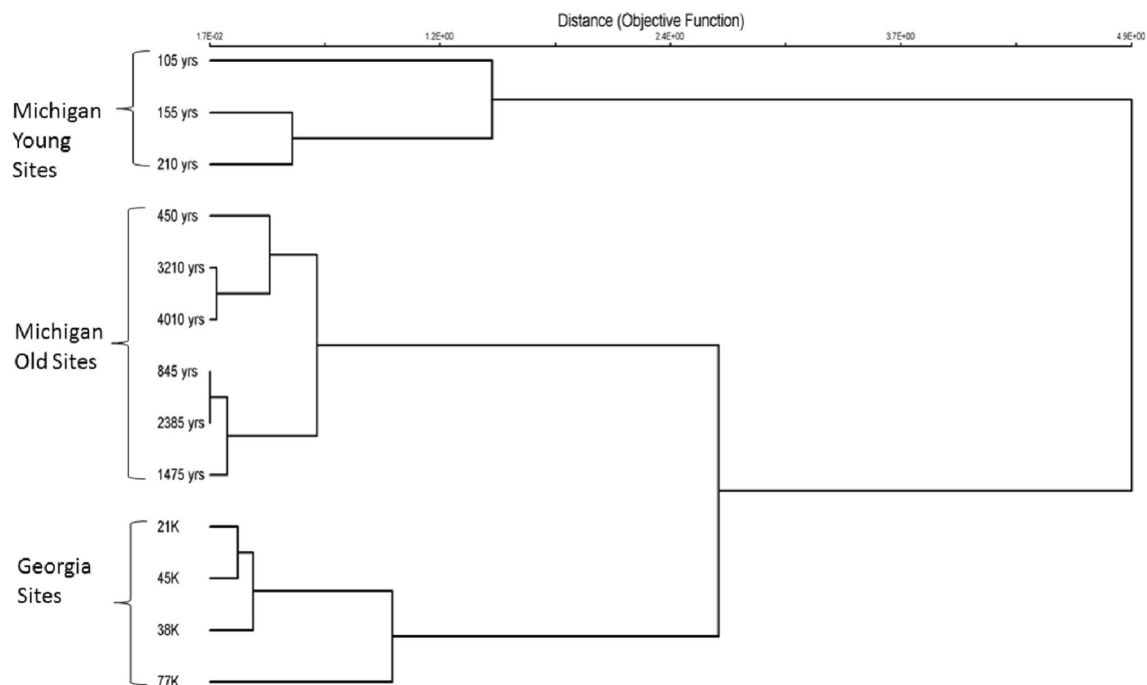

Fig. 2 Hierarchical Cluster Analysis of the 100 most abundant bacterial OTU. Notes: Cluster analysis done using Ward's method and relative Euclidean distance. The distance axis represents the similarity index

Fig. 3 Canonical correspondence analysis of the relationship between physical and chemical parameters and bacterial relative abundances (OTU intensities). The projected length of the vectors centered in the panel on the two axis represents the strength of the soil factors. Notes: All environmental variables were significant ($P < 0:05$, Monte Carlo test of significance). *Error bars* represent standard error ($n = 5$)
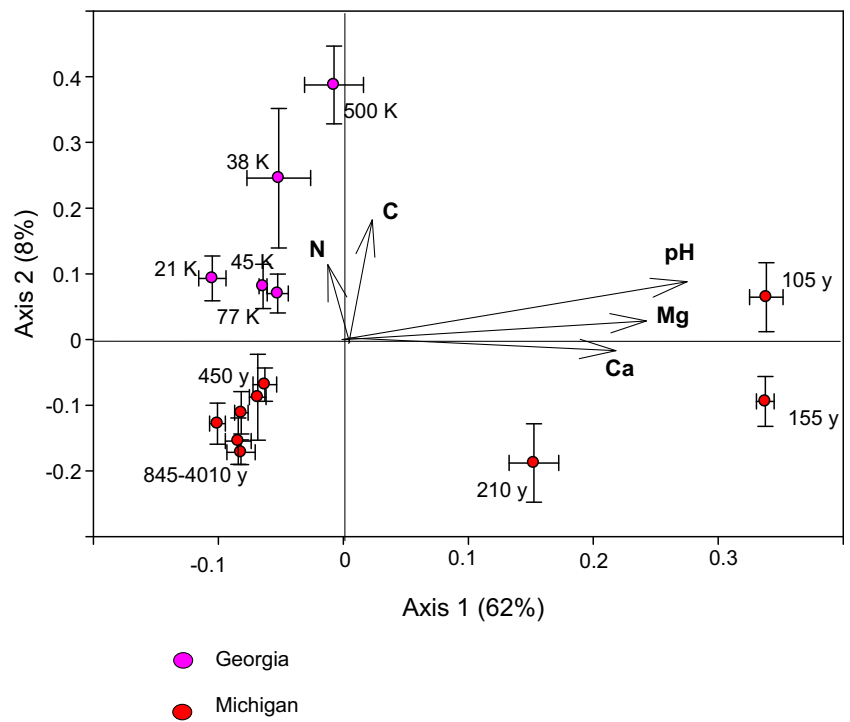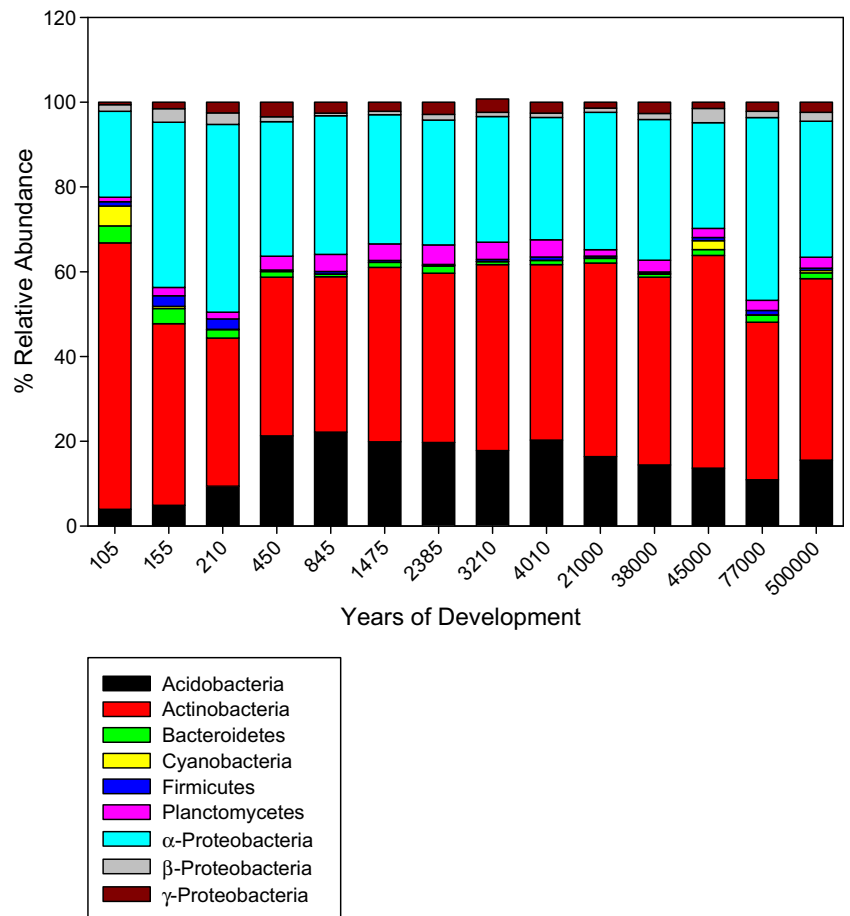
Fig. 4 Relationship between the relative abundance of bacterial phyla across Michigan and Georgia chronosequences
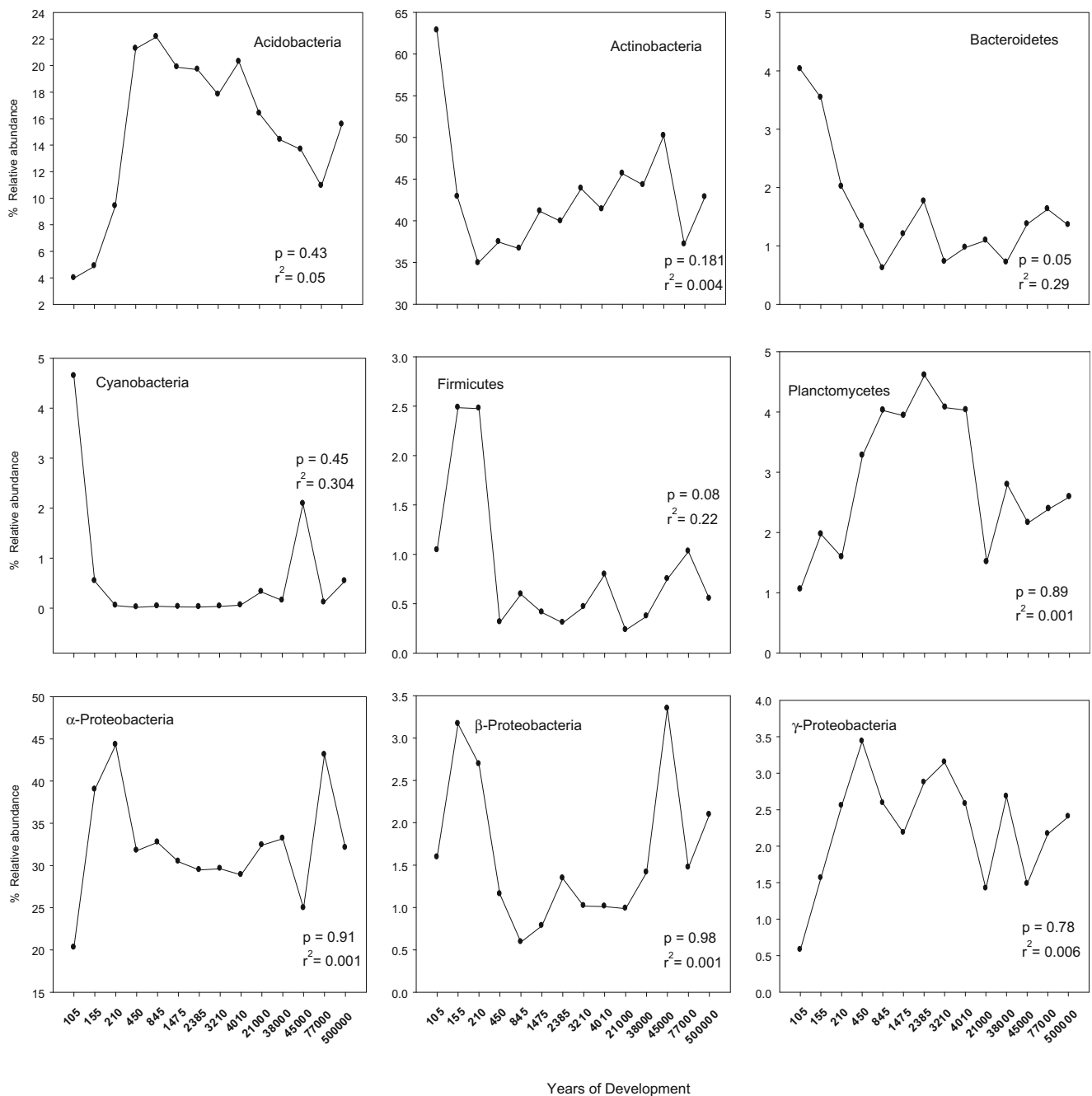
Fig. 5 Relationship between percentage relative abundance of nine individual bacterial phyla across both chronosequences during ecosystem development (105 to 500,000 years). Notes: Each point in the graph is the average (n = 5) of the percentage abundance of each phyla at each stage of development. Regression co-efficient and p value for each phylum are shown

greater in MI (numbers of OTU, ACE, and Chao1) and remained 2–3× greater in MI than GA after (Table 5).

Bacterial Biogeography

A Mantel test showed that there was no significant correlation between the bacterial community dissimilarity and the geographic distance (Table 6). The standardized Mantel statistic (r_M) was not significant at 95% confidence level using 999

permutations (r_M = 0.13, P = 0.35) and was unsupportive of a pattern of community change related to distance. Because the distance factor is strongly weighted by many values below 50 km and those above 1700 km, a Mantel correlogram was developed to plot autocorrelation as a function of geographic distance [37, 38] classes defined by: class 1 (0–1 km), class 2 (1–50 km), and class 3 (50–1700 km). For each class, a n × n matrix was constructed containing zeroes for site pairs whose geographic distances fall within the class and ones for pairs that

Table 5 Diversity indices of the 16S rRNA gene sequences

Diversity Index [a]	105 years	155 years	210 years	450 years	845 years	1475 years	2385 years	3210 years	4010 years	21 K	38 K	45 K	77 K	500 K
N[b]	4779	4779	4779	4779	4779	4779	4779	4779	4779	4779	4779	4779	4779	4779
S[c]	1969	2352	2259	950	1225	1027	674	977	815	541	537	590	555	660
Evenness	2.01	2.00	1.99	1.75	1.73	1.70	1.70	1.78	1.72	1.23	1.28	1.29	1.21	1.34
Richness (ace)	15,557	27,093	28,193	6320	11,929	7463	3042	7545	4433	1775	1823	2092	2322	2252
Shannon	6.63	6.73	6.68	5.22	5.34	5.11	4.80	5.33	5.00	4.5	4.7	4.7	4.5	4.9
1/D[d]	205	134	125	58	53	43	45	72	52	40	49	51	33	47
Chao 1	7572	12,736	10,863	3253	5045	3423	1813	3671	2380	1060	1030	1332	1194	1457

K = 1000 years

[a] Calculations based on the Operational Taxonomic Units (OTU) formed at an evolutionary distance of <0.03

[b] Number of sequences in the library

[c] Number of OTU

[d] Simpson's reciprocal index

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 (Fig. 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 $2 = 0.25$) (Fig. 6). However, for the third distance class (50–1700 km), there was weak negative correlation which was significant ($r_M = -0.19$, $p < 0.001$). A simple Mantel test also showed that there was significant correlation between bacterial community distribution and soil physico-chemical characteristics ($r_M = 0.59$, $p < 0.001$) (Table 6). However, controlling for variables is an important challenge in microbial biogeography studies when one tries to make meaningful comparisons between the geographic distance and genetic differences, as they may make the comparisons more complex [39]. 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 6). However, when the soil physico-chemical characteristics were included as a control matrix, there was no significant correlation between bacterial community distribution and geographical distance ($r_M = 0.02$, $p = 0.41$) (Table 6).

Relative Contribution of Temporal and Environmental Filters on Bacterial Community Structure

Variance partitioning analysis (VPA) was performed to quantify the relative contributions of geographic distance and soil parameters to the taxonomic structure of the bacterial communities. A subset of environmental parameters (Ca, Mg, pH, total carbon (%), and total nitrogen (%)) that had the highest Pearson correlation with the bacterial communities were selected by the CCA procedure. The combination of selected soil characteristics and geographic distance showed a significant ($p = 0.002$) correlation with the bacterial community structure. These factors explained 50.1% of the observed variation, leaving 49.9% of the variation unexplained. The soil factors explained 38.2% ($p = 0.003$), and geographic distance alone explained 11.9% ($p = 0.029$) variations, and no interaction effect was detected (Fig. 7). Although the sediment properties together explained more of the variation, geographic distance by itself explained 11.9% of the variation observed, more than any of the other 5 of the individual soil variables (Fig. 7).

Discussion

The broad goal of this experiment was to characterize patterns of bacterial community composition and diversity during soil pedogenesis in two different climate zones. Parent materials in

Table 6 Mantel test correlations between the bacterial community distribution and selected environmental characteristics

Variable	Mantel test [a]		Partial mantel test	
	r $_M$	P value	r $_M$	P value
Soil physico-chemical characteristics	0.59	<0.0001	0.84[b]	0.001
Geographic Distance	0.13	0.35	0.02[c]	0.41

[a] Mantel test of relationship between bacterial community similarity matrix ,environmental characteristics and geographic distance. r $_M$ Standardized Mantel statistic

[b] Partial Mantel test of relationship between bacterial community similarity matrix and environmental characteristics after controlling for geographical distance

[c] Partial mantel test of relationship between bacterial community similarity matrix and geographical distance after controlling for environmental characteristics

the two locations where similar, but the stages (and age) of ecosystem development differed, with the site in MI representing an aggrading system and that in GA representing a mature mid-succession and possibly retrogressive system. Aggrading ecosystem by definition is gradual accumulation of biomass in the form of living biomass, dead wood, soil fertility, and physical and biological complexity. Essentially, the ecosystem is building biomass over time during the developmental period through primary succession [40, 41]. A retrogressive ecosystem on the other hand is characterized by reductions in ecosystem productivity and plant biomass which is accompanied by shifts in aboveground and belowground communities dominated by stress-tolerant species. These ecosystems are often associated with plant communities reaching a stable and self-replacing climax [41, 42]. One of the important aspects of our study is that it represents first large-scale molecular survey of soil bacterial 16S rRNA gene β diversity across naturally occurring sand chronosequences. Although β diversity estimates (between sample variations) are considered important for the overall understanding of soil microbial community dynamics [8, 43], they are traditionally less studied than α diversity (diversity estimates within a single sample). Pedogenesis is the driver of soil change and is affected by parent material, climate, organisms such as plants and microbes [44], and the length of time that these processes have occurred. As shown previously, bacterial community change is sensitive to environmental gradients [45, 46], especially soil pH [47]. There are, however, a number of other insights with community change reflecting differences between the two locations.

The absolute diversity of taxa is shown to be underestimates; however, comparison of similar sized libraries provides confidence when making comparisons between soils. The soils in MI were estimated to contain up to 2500 different OTU, while those in GA were much lower and ranged between 500 and 600 OTU. Even when pH and soil properties were considered, the

number of OTU remained ~2× greater (~1000) in the soils of MI. Greater evenness and 3× greater richness (ACE, Chao1) support the conclusion that there is greater diversity in the bacterial communities in the soils of MI.

Bacterial species loss and turnover could also affect the diversity of bacterial taxa in a soil. Weathering rates and temperature related to climate are considerably greater in southern GA than northern-lower MI. Turnover and decomposition of organic matter is greater in GA and therefore likely supports a lower standing microbial biomass. This was supported by the CCA analysis (Fig. 3) which showed that carbon and nitrogen were important drivers of community differences between the two locations [48] . In GA dunes, high summer temperature, low organic matter at the surface to modulate temperature swings, and the feast and famine of rainfall might create some of the most extreme environmental swings for bacteria existing in the top 10 cm of soil. Many studies have shown that alternate wetting and drying of soil in sub-tropical conditions may have a direct influence on microbial community structure and diversity [49, 50]. In MI, low temperatures might be the biologically most difficult variable to adapt; however, large swings are modulated to some extent by the regional climatic effects of Lake Michigan. We also attribute pH to be another important filter that shaped microbial community diversity. The chronosequences showed changes in soil properties, depicting the classic patterns of soil podzolization during ecosystem development [51] . 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 [51, 52]. 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 [3, 53, 54]. 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 (Fig. 5). Dunes that were near neutral showed higher bacterial community diversity when compared to the older dunes that were more acidic and supported less bacterial diversity. Therefore, it appears that the established pH–diversity relationship [47, 54] observed here and elsewhere could be driven by the dominance of a few taxonomic groups in low pH soils.

While studying the effect of dispersal limitation on microbial community structure, the community similarities were not positively correlated with geographic distance (r $_M$ = 0.13, p = 0.35) when the distance matrix included all the distances. Thus the Mantel test shows there was no significant isolation

Fig. 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)
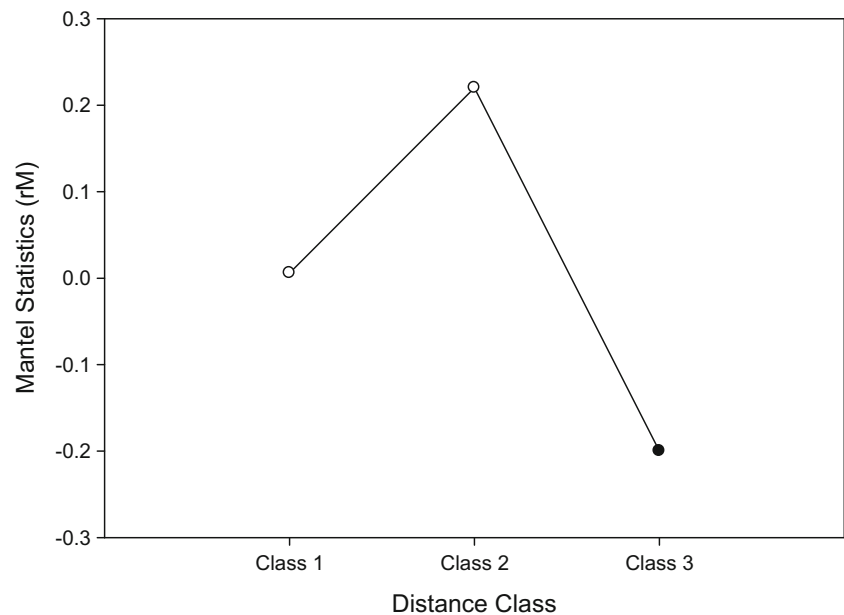
by distance at a confidence level of 95 %. However, when the distance variance was partitioned into classes, there was a marginal correlation at a spatial scale of 50–1700 km. Our results showed that, at this scale, the community dissimilarity had a moderate negative correlation with the geographic distance (Fig. 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 soil physico-chemical factors when the effect of the geographic distance was not controlled ($r_M = 0.59$, $p < 0.001$), indicating an increase in genetic dissimilarities once the geographic isolation effect was removed. Thus the result show
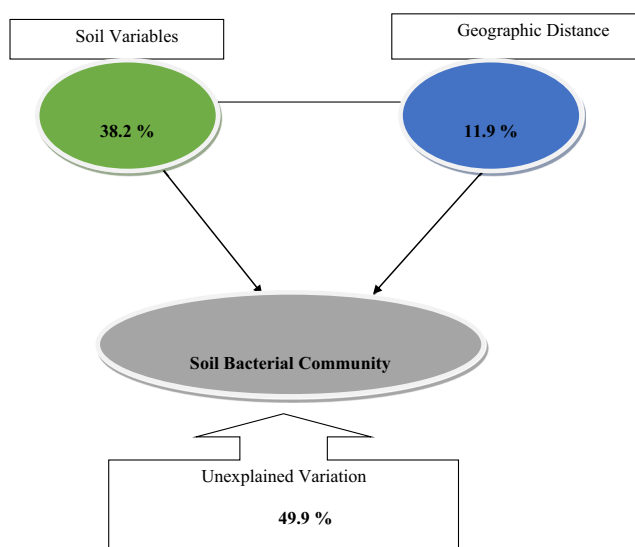

Fig. 7 Variance partition analysis of the effects of geographic distance and soil variables on the bacterial community structure

some evidence of dispersal limitation in shaping the community patterns but they are difficult to distinguish from the effects of environmental heterogeneity [37, 55].

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 [5, 56]. 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 [14, 57]. At intermediate geographic scale, studies have shown individual distance effect [58] and environmental effect [7] 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 [59]. Similar to what we observed, studies have shown that the correlation between community dissimilarity and the geographic distance disappeared as the soils got separated farther [60, 61]. 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.

In conclusion, this study attempts to survey bacterial spatial patterns across two pristine US dune chronosequences which are approximately 1700 km apart. Despite the large effect of environmental factors on community distances, we also found some evidence for residual spatial autocorrelation at closer spatial scales. 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 specific 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, but, *the environment selects*" which implies that similar habitat and physical conditions would support similar microbial communities. Soil pedogenesis is both a result and driver of ecosystem development and change, and thus understanding how microbial communities change during this natural process will help develop and describe fundamental ecological theory.

References

1. Gans J, Wolinsky M, Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309:1387

2. Ranjard L, Dequiedt S, Jolivet C, Saby NPA, Thioulouse J, Harmand J, Loisel P, Rapaport A, Fall S, Simonet P, Joffre R, Bouré NCP, Maron PA, Mougel C, Martin MP, Toutain B, Arrouays D, Lemanceau P (2010) Biogeography of soil microbial communities: a review and a description of the ongoing French national initiative. Agron Sustain Dev 30:359–365

3. Xiong J, Liu Y, Lin X, Zhang H, Zeng J, Hou J, Yang Y, Yao T, Knight R, Chu H (2012) Geographic distance and pH drive bacterial distribution in alkaline lake sediments across Tibetan Plateau. Environmental microbiology 14:2457-2466

4. Sun B, Wang X, Wang F, Jiang Y, Zhang X-X (2013) Assessing the relative effects of geographic location and soil type on microbial communities associated with straw decomposition. Appl Environ Microbiol 79:3327–3335

5. Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR (2006) Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 4:102–112

6. Fierer N (2008) Microbial biogeography: patterns in microbial diversity across space and time. ASM Press, Washington DC

7. Green J, Bohannan BJM (2006) Spatial scaling of microbial biodiversity. Trends Ecol Evol 21:501–507

8. Green JL, Bohannan BJM, Whitaker RJ (2008) Microbial biogeography: from taxonomy to traits. Science 320:1039

9. Fenchel T, Finlay BJ (2004) The ubiquity of small species: patterns of local and global diversity. Bioscience 54:777–784

10. Pasternak Z, Al-Ashhab A, Gatica J, Gafny R, Avraham S, Minz D, Gillor O, Jurkevitch E (2013) Spatial and temporal biogeography of soil microbial communities in arid and semiarid regions. PLoS One 8, e69705

11. Hanson CA, Fuhrman JA, Horner-Devine MC, Martiny JBH (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. Nat Rev Microbiol

12. Whitaker RJ, Grogan DW, Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science 301:976–978

13. Peay KG, Garbelotto M, Bruns TD (2010) Evidence of dispersal limitation in soil microorganisms: isolation reduces species richness on mycorrhizal tree islands. Ecology 91:3631–3640

14. Horner-Devine MC, Silver JM, Leibold MA, Bohannan BJM, Colwell RK, Fuhrman JA, Green JL, Kuske CR, Martiny JBH, Muyzer G (2007) A comparison of taxon co-occurrence patterns for macro-and microorganisms. Ecology 88:1345–1353

15. Winter C, Matthews B, Suttle CA (2013) Effects of environmental variation and spatial distance on Bacteria, Archaea and viruses in sub-polar and arctic waters. ISME J

16. Zheng Y-M, Cao P, Fu B, Hughes JM, He J-Z (2013) Ecological drivers of biogeographic patterns of soil Archaeal community. PLoS One 8, e63375

17. Lichter J (1995) Lake Michigan beach-ridge and dune development, lake level, and variability in regional water balance. Quat Res 44:181–189

18. Lichter J (1998) Rates of weathering and chemical depletion in soils across a chronosequence of Lake Michigan sand dunes. Geoderma 85:255–282

19. Ivester AH, Leigh DS (2003) Riverine dunes on the Coastal Plain of Georgia, USA. Geomorphology 51:289–311

20. Ivester AH, Leigh DS, Godfrey-Smith D (2001) Chronology of inland eolian dunes on the coastal plain of Georgia, USA. Quat Res 55:293–302

21. Dequiedt S, Saby N, Lelievre M, Jolivet C, Thioulouse J, Toutain B, Arrouays D, Bispo A, Lemanceau P, Ranjard L (2011) Biogeographical patterns of soil molecular microbial biomass as influenced by soil characteristics and management. Glob Ecol Biogeogr

22. Maron PA, Mougel C, Ranjard L (2011) Soil microbial diversity: methodological strategy, spatial overview and functional interest. C R Biol

23. Acosta-Martinez V, Dowd S, Sun Y, Allen V (2008) Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. Soil Biol Biochem 40: 2762–2770

24. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ (2008) Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol 74:2461–2470

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

26. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336

27. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci 108:4516

28. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537

29. Quince C, Lanzen A, Davenport RJ, Turnbaugh PJ (2011) Removing noise from pyrosequenced amplicons. BMC Bioinf 12:38

30. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6, e27310

31. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596

32. McCune B, Grace JB (2002) Analysis of ecological communities. MjM Software Design

33. McCune B, Mefford MJ (1999) PC-ORD: multivariate analysis of ecological data. MjM Software Design
34. ter Braak CJ, Verdonschot PF (1995) Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquat Sci 57:255–289
35. Mitchell RJ, Hester AJ, Campbell CD, Chapman SJ, Cameron CM, Hewison RL, Potts JM (2010) Is vegetation composition or soil chemistry the best predictor of the soil microbial community? Plant Soil 333:417–430
36. Gihring TM, Green SJ, Schadt CW (2012) Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. Environ Microbiol 14:285–290
37. Cho JC, Tiedje JM (2000) Biogeography and degree of endemicity of fluorescent Pseudomonas strains in soil. Appl Environ Microbiol 66:5448–5456
38. Sokal RR, Oden NL (2008) Spatial autocorrelation in biology: 1. Methodology. Biol J Linn Soc 10:199–228
39. Whittaker RJ, Willis KJ, Field R (2001) Scale and species richness: towards a general, hierarchical theory of species diversity. J Biogeogr 28:453–470
40. Whittaker R, Bush M, Richards K (1989) Plant recolonization and vegetation succession on the Krakatau Islands, Indonesia. Ecol Monogr 59:59–123
41. Walker LR, Wardle DA, Bardgett RD, Clarkson BD (2010) The use of chronosequences in studies of ecological succession and soil development. J Ecol 98:725–736
42. Wardle DA (2002) Communities and ecosystems: linking the aboveground and belowground components. Princeton University Press
43. Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, He J, Liu X, Zhang L, Zhang E (2013) Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. ISME J
44. Minasny B, McBratney AB, Salvador-Blanes S (2008) Quantitative models for pedogenesis—a review. Geoderma 144:140–157
45. Landesman WJ, Nelson DM, Fitzpatrick MC (2014) Soil properties and tree species drive ß-diversity of soil bacterial communities. Soil Biol Biochem
46. Hu Y, Xiang D, Veresoglou SD, Chen F, Chen Y, Hao Z, Zhang X, Chen B (2014) Soil organic carbon and soil structure are driving microbial abundance and community composition across the arid and semi-arid grasslands in northern China. Soil Biol Biochem 77:51–57
47. Lauber CL, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. Appl Environ Microbiol 75:5111
48. Beilke A, Bockheim J (2013) Carbon and nitrogen trends in soil chronosequences of the Transantarctic Mountains. Geoderma 197:117–125
49. Lundquist E, Scow K, Jackson L, Uesugi S, Johnson C (1999) Rapid response of soil microbial communities from conventional, low input, and organic farming systems to a wet/dry cycle. Soil Biol Biochem 31:1661–1675
50. Fierer N, Schimel J, Holden P (2003) Influence of drying–rewetting frequency on soil bacterial community structure. Microb Ecol 45:63–71
51. Huggett R (1998) Soil chronosequences, soil development, and soil evolution: a critical review. Catena 32:155–172
52. Lichter J (1998) Primary succession and forest development oncoastal Lake Michigan sand dunes. Ecol Monogr 68:487–510
53. Chu H, Fierer N, Lauber CL, Caporaso J, Knight R, Grogan P (2010) Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. Environ Microbiol 12:2998–3006
54. Rousk J, Brookes PC, Bååth E (2009) Contrasting soil pH effects on fungal and bacterial growth suggest functional redundancy in carbon mineralization. Appl Environ Microbiol 75:1589–1596
55. Ramette A, Tiedje JM (2007) Biogeography: an emerging cornerstone for understanding prokaryotic diversity, ecology, and evolution. Microb Ecol 53:197–207
56. Griffiths RI, Thomson BC, James P, Bell T, Bailey M, Whiteley AS (2011) The bacterial biogeography of British soils. Environ Microbiol 13:1642–1654
57. Kuske CR, Ticknor LO, Miller ME, Dunbar JM, Davis JA, Barns SM, Belnap J (2002) Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. Appl Environ Microbiol 68:1854–1863
58. Reche I, Pulido-Villena E, Morales-Baquero R, Casamayor EO (2005) Does ecosystem size determine aquatic bacterial richness? Ecology 86:1715–1722
59. Ge Y, He J, Zhu Y, Zhang J, Xu Z, Zhang L, Zheng Y (2008) Differences in soil bacterial diversity: driven by contemporary disturbances or historical contingencies? ISME J 2:254–264
60. Tripathi BM, Lee-Cruz L, Kim M, Singh D, Go R, Shukor NAA, Husni MHA, Chun J, Adams JM (2014) Spatial scaling effects on soil bacterial communities in Malaysian tropical forests. Microb Ecol 1–12
61. Pereira e Silva MC, Dias ACF, van Elsas JD, Salles JF (2012) Spatial and temporal variation of archaeal, bacterial and fungal communities in agricultural soils. PLoS One 7, e51554