3-27-2023

Soil depth and geographic distance modulate bacterial β-diversity in deep soil profiles throughout the U.S. Corn Belt

Lucas Dantas Lopes

Stephanie L. Futrell

Emily E. Wright

Gerasimos J. Danalatos

Michael J. Castellano

See next page for additional authors

Authors

Lucas Dantas Lopes, Stephanie L. Futrell, Emily E. Wright, Gerasimos J. Danalatos, Michael J. Castellano, Tony J. Vyn, Sotirios V. Archontoulis, and Daniel P. Schachtman

Received: 23 May 2022 | Revised: 3 March 2023 | Accepted: 27 March 2023

DOI: 10.1111/mec.16945

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Soil depth and geographic distance modulate bacterial β-diversity in deep soil profiles throughout the U.S. Corn Belt

Lucas Dantas Lopes[1] | Stephanie L. Futrell[1] | Emily E. Wright[2] | Gerasimos J. Danalatos[1] | Michael J. Castellano[2] | Tony J. Vyn[3] | Sotirios V. Archontoulis[2] | Daniel P. Schachtman[1]

[1]Department of Agronomy and Horticulture, Centre for Plant Science Innovation, University of Nebraska–Lincoln, Lincoln, Nebraska, USA

[2]Iowa State University, Department of Agronomy, Ames, Iowa, USA

[3]Purdue University, Department of Agronomy, West Lafayette, Indiana, USA

Correspondence
Daniel P. Schachtman, Department of Agronomy and Horticulture, Centre for Plant Science Innovation, Nebraska Center for Biotechnology, University of Nebraska–Lincoln, 1901 Vine Street, Lincoln, NE 68588, USA.
Email: daniel.schachtman@unl.edu

Funding information
Foundation for Food and Agriculture Research; National Institute of Food and Agriculture, Grant/Award Number: 2020-67019-31796

Handling Editor: Simon Creer

Abstract

Understanding how microbial communities are shaped across spatial dimensions is of fundamental importance in microbial ecology. However, most studies on soil biogeography have focused on the topsoil microbiome, while the factors driving the subsoil microbiome distribution are largely unknown. Here we used 16S rRNA amplicon sequencing to analyse the factors underlying the bacterial β-diversity along vertical (0–240 cm of soil depth) and horizontal spatial dimensions (~500,000 km^2) in the U.S. Corn Belt. With these data we tested whether the horizontal or vertical spatial variation had stronger impacts on the taxonomic (Bray-Curtis) and phylogenetic (weighted Unifrac) β-diversity. Additionally, we assessed whether the distance-decay (horizontal dimension) was greater in the topsoil (0–30 cm) or subsoil (in each 30 cm layer from 30–240 cm) using Mantel tests. The influence of geographic distance versus edaphic variables on the bacterial communities from the different soil layers was also compared. Results indicated that the phylogenetic β-diversity was impacted more by soil depth, while the taxonomic β-diversity changed more between geographic locations. The distance-decay was lower in the topsoil than in all subsoil layers analysed. Moreover, some subsoil layers were influenced more by geographic distance than any edaphic variable, including pH. Although different factors affected the topsoil and subsoil biogeography, niche-based models explained the community assembly of all soil layers. This comprehensive study contributed to elucidating important aspects of soil bacterial biogeography including the major impact of soil depth on the phylogenetic β-diversity, and the greater influence of geographic distance on subsoil than on topsoil bacterial communities in agroecosystems.

KEYWORDS
distance decay, microbial biogeography, soil bacteria, subsoil ecology

1 | INTRODUCTION

An important and still unresolved topic in terrestrial microbial ecology is how bacterial communities are distributed across the space and what are the key factors driving their biogeographic patterns (Chu et al., 2020; Hanson et al., 2012; Martiny et al., 2006; Meyer et al., 2018). In order to shed light on these questions research has been conducted to analyse samples on continental and even global scales, concluding that bacteria have weak biogeographic patterns that are largely independent of geographic distance and strongly driven by edaphic variables, particularly by soil pH (Bahram et al., 2018; Fierer et al., 2012; Fierer & Jackson, 2006; Karimi et al., 2018; Lauber et al., 2009; Thompson et al., 2017). A few studies on soil bacterial biogeography have also been performed exclusively in agroecosystems such as rice, maize, soybean, and sugarcane plantations (Durrer et al., 2017; Jiao et al., 2020; Li et al., 2021; Liu et al., 2014). Agroecosystems that comprise a large swath of terrestrial landscape may differ from natural ecosystems because there are direct anthropogenic impacts that may influence the spatial distribution of bacterial communities. However, most studies on both natural and agroecosystems analysed only the most superficial soil layer, that is, topsoil and therefore the ecological processes governing the bacterial community distribution in the subsoil are largely unexplored.

Despite the focus of most soil biogeography studies was on understanding the drivers shaping the topsoil bacterial communities along the horizontal spatial dimension, previous studies showed that the bacterial communities also change in the vertical spatial dimension, that is, across soil depth (Eilers et al., 2012; Hao et al., 2021; Jiao et al., 2018; Mundra et al., 2021; Stone et al., 2014). The soil physicochemical conditions change dramatically from the topsoil to the subsoil, for example, soil texture, carbon, and oxygen availability (Hartemink et al., 2020; Sun et al., 2020; Wordell-Dietrich et al., 2017). Therefore, the changes in bacterial community composition along the vertical spatial dimension may be as large as the changes in the horizontal spatial dimension. However, it is unknown which of the two spatial dimensions has the largest impact on the structure and diversity of bacterial communities. In addition to answer basic ecological questions, knowledge about the factors shaping the subsoil microbiome will be important for agriculture, since sustainable agricultural practices aim to enhance the root exploration of the subsoil resources – including the benefits of growth promotion by rhizobacteria (Gocke et al., 2017; Lynch, 2021; Singh et al., 2020; Zhang & Peng, 2021). Moreover, this knowledge may also impact fields of environmental concern such as carbon sequestration and climate change, where subsoil microbes play a key role (Jia et al., 2017; Sun et al., 2021; Xu et al., 2021).

Previous studies indicated the existence of distance-decay relationships in the topsoil bacterial communities (Hanson et al., 2012; Jiao et al., 2020; Li et al., 2021; Meyer et al., 2018), that is, the greater the geographic distance, the larger the dissimilarity between the bacterial communities (Bell, 2010). However, the distribution of the topsoil bacterial communities has in general been shown to be more strongly associated with edaphic variables such as pH than geographic distance (Fierer & Jackson, 2006; Lauber et al., 2009; Griffiths et al., 2011; Bahram et al., 2018; Karimi et al., 2018; Li et al., 2021). The few bacterial biogeography studies available that analysed multiple soil depths indicated different biogeographic patterns in the topsoil compared to the subsoil communities in natural ecosystems (Du et al., 2021; Li et al., 2022; Liu et al., 2019; Seuradge et al., 2017). However, the drivers of bacterial β-diversity depend on the environment and spatial scale (Martiny et al., 2011). Therefore, the existence of distance-decay relationships in subsoil bacterial communities inhabiting deeper soil layers across larger horizontal spatial scales is unclear, as well as the factors driving the biogeography of subsoil bacteria in agroecosystems.

Here we analysed the soil bacterial communities in eight soil layers to a depth of 240 cm from 26 on-farm production fields geographically distributed across eight states in the midwestern United States covering a region of approximately 500,000 km^2. Our hypotheses were: (i) The bacterial taxonomic and phylogenetic β-diversity change more in the vertical (along soil depth) than in the horizontal (between geographical locations) spatial dimension due to the drastic changes in soil physicochemical properties along soil depth; (ii) The bacterial communities in the subsoil layers of agroecosystems also show a distance-decay relationship in the horizontal spatial dimension, which is larger than in the topsoil due to stronger selection and dispersal limitation. Our study found a greater impact of soil depth than geographic location on the phylogenetic β-diversity, and a stronger influence of geographic distance on the structure of subsoil bacterial communities compared to those in the topsoil. These results highlight that the variation in the vertical dimension is as important as in the horizontal dimension to the spatial distribution of soil bacterial communities, and indicate that the biogeographic patterns of subsoil and topsoil bacteria are different in agroecosystems.

2 | MATERIALS AND METHODS

2.1 | Sampling locations and soil analyses

Soil samples were collected from a total of 26 locations across the midwest of the United States of America, including the states of Illinois (4 locations), Indiana (3 locations), Iowa (7 locations), Minnesota (2 locations), Missouri (1 location), Nebraska (7 locations), South Dakota (1 location) and Wisconsin (1 location) (Figure 1a). The nearest cities to each of the 26 agricultural fields and the states where they are located are provided in Table S1. The fields were in a maize-soybean rotation. The 26 locations had a maximum longitudinal variation of 1170 km and latitudinal variation of 573 km, comprising an approximate area of 500,000 km^2. Soil samples were collected from 5 April to 6 May 2021, using Giddings Hydraulic probe. At least two different soil cores ($n = 2$) were collected at each location, and a total of 54 individual soil cores were collected among the 26 locations (Table S1). The soil samples were taken before the fields were

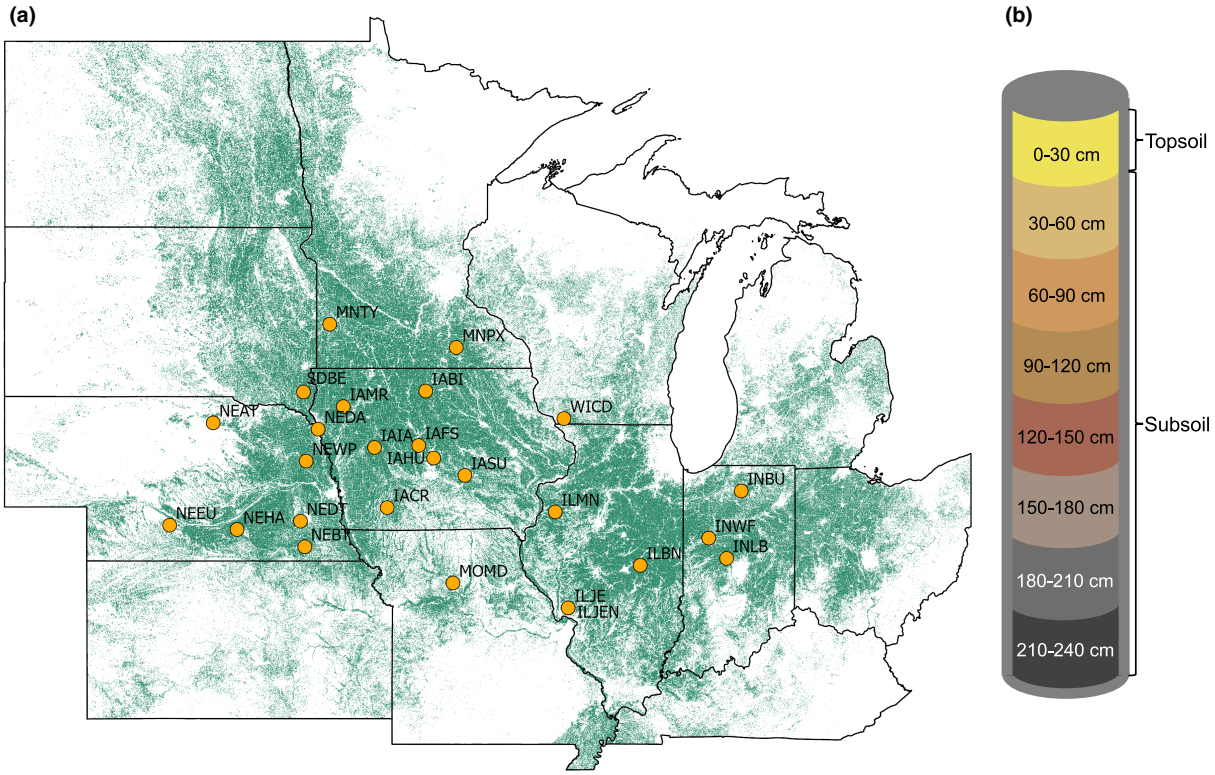

FIGURE 1 Geographic locations where soil cores were sampled and the different soil layers analysed. (a) Map showing the midwestern region of the United States of America (USA) where soil cores were collected. The yellow dots indicate the specific locations of sample collection and their respective codes. Samples were collected in 26 locations across eight states of the USA midwest where the Corn Belt is located. At each location at least two independent soil cores were collected and used for 16S rRNA gene sequencing. The background green colour represents the region planted with maize and/or soybean. The eight states of the 26 locations, their nearest cities, all the samples used for sequencing (total of 385), the core replicate of each sample at each location, and the number of sequencing reads before and after sequence analysis are provided in Table S1. The map shows the variation in the horizontal spatial dimension in our study, comprising an approximate area of 500,000 km^2. (b) Schematic overview of the soil core showing the whole soil profile sampled containing the eight soil layers analysed, including the topsoil (0–30 cm) and subsoil (30–60, 60–90, 90–120, 120–150, 150–180, 180–210, and 210–240 cm). The different soil layers sampled indicate the variation in the vertical spatial dimension in our study, comprising 0–240 cm of soil depth.

planted with maize. In the previous year, the fields were all planted with soybean. Each soil core was split into 30 cm increments up to 240 cm of soil depth, generating samples from the following soil layers: 0–30 cm, 30–60 cm, 60–90 cm, 90–120 cm, 120–150 cm, 150–180 cm, 180–210 cm, and 210–240 cm (Figure 1b). At some sites below a certain depth soil samples could not be collected due to soil physical factors or shallow water tables. The soil samples from each layer were placed in plastic bags, kept cool, and transported to the lab. Part of each soil sample was stored at −20°C for molecular analysis.

Two replicate samples (each from a different soil core) from the eight soil layers comprising a total of 385 individual samples were collected in the 26 geographic locations and used for DNA extraction and 16S rRNA gene sequencing (Table S1). Thirty-eight soil cores had samples from all eight soil layers to a soil depth of 240 cm (304 samples), five soil cores had samples from seven soil layers to a soil depth of 210 cm (35 samples), one soil core had samples from six soil layers to a soil depth of 180 cm (6 samples), two soil cores had samples from five soil layers to a soil depth of 150 cm (10 samples), six soil cores had samples from four soil layers to a soil depth of 120 cm (24 samples), and two soil cores had samples from three soil layers to a soil depth of 90 cm (6 samples), resulting in a total of 385 samples.

Each soil sample was analysed for nitrate (NO_3^-, ppm) and ammonium (NH_4^+, ppm) using methods described by Hood-Nowotny et al. (2010). Soil samples were also analysed for soil organic matter (SOM, %) by measuring the soil weight (dried for 24 h at 105°C) before and after removing the organic matter using a furnace at 360°C. Soil phosphorus analysis (total P, ppm) was performed using the Bray method (Bray & Kurtz, 1945). In soils above pH 7.3 the P was validated using the Olsen bicarbonate method (Olsen & Sommers, 1982). Soil potassium (K, ppm), calcium (Ca, ppm) and magnesium (Mg, ppm) were extracted with ammonium acetate and measured with inductively coupled plasma (ICP). Soil pH was measured in a 1:1 soil: water slurry using a pH meter. Cation exchange capacity (CEC) was calculated as the sum of cations which were normalized by their equivalent weights. Sand, clay and silt content (%) were analysed with a hydrometer (Gee & Bauder, 1979). Gravimetric soil moisture was measured by weighing soil samples before and after drying for 24 h at 105°C.

2.2 | 16S rRNA gene amplification and Illumina Miseq sequencing

The PowerSoil-htp 96 well Soil DNA Isolation kit (MoBio) was used to extract the DNA from the soil samples, which was quantified using QuantiFluor ONE dsDNA Dye (Promega Corporation). The V4 hypervariable region of the 16S rRNA gene was amplified in the samples using the primer pair V4_515F/V4_806R (Caporaso et al., 2013). The PCR thermal cycling conditions were initial denaturing at 95°C for 5 min, 30 cycles of 95°C for 40 s (denaturing), 65°C for 30 s (annealing), and 72°C for 1 min (extension), followed by final extension at 72°C for 10 min. The 385 samples were divided into two sequencing libraries, each containing samples from locations of 5/8 states (Table S1). Samples from three states (Wisconsin, South Dakota and Missouri) were only collected in one location and therefore were present in only one of the two libraries. A total of 288 samples were pooled in each library, including the samples used in this study, and other samples not used in this study, H_2O blanks and mocking communities. Among the samples used in this study, 232 were pooled in Library 1 and 153 were pooled in Library 2 (Table S1). The sequencing libraries were prepared using a dual-index primer system containing the Illumina adapter, an eight-nucleotide index sequence, a 10-nucleotide pad-sequence, a two-nucleotide linker sequence and the primer sequence (Kozich et al., 2013). Amplicons were purified and normalized using SequalPrep normalization plates (Invitrogen), pooled in an equimolar mixture, concentrated using a SpeedVac, and size-selected within a range of 200–700 bp using the SPRIselect beads (Beckman Coulter). The amplicon libraries were quantified and quality-checked using a high-sensitivity DNA kit on an Agilent 2100 Bioanalyser (Agilent Technologies), spiked with 20% PhiX and loaded in the Illumina flowcell. Sequencing was performed on an Illumina MiSeq platform using the MiSeq 600 cycles version 3 kit (Gohl et al., 2016).

2.3 | Bioinformatic analysis of reads

The sequencing reads from the two libraries were pooled together and analysed with the QIIME 2.0 software (Bolyen et al., 2019). The DADA2 algorithm was used for denoising and generating the amplicon sequence variants (ASV) table (Callahan et al., 2017). The SILVA database version 138 was used for taxonomic classification of reads (Quast et al., 2013). Mitochondria and chloroplast sequences were filtered from the ASV table. Archaeal sequences were not removed, but we used the term "bacterial community/diversity" in the manuscript because bacterial taxa dominated the communities of all soil depth layers studied. Initially, samples from the eight soil layers and 26 locations were analysed together in order to compare the influence of soil depth (vertical variation) versus geographic location (horizontal variation) on the bacterial community structure. For that, the β-diversity was analysed using the unrarefied data which included all the 385 samples from the 26 locations and eight soil depth layers, except those without any reads after sequence processing (Figure 2). We also analysed the results with rarefied data (Figure S1). The samples contained a range of reads so that 184/385 (48%) of the samples contained >50,000 reads, 65/385 (17%) of the samples contained between 50,000–20,000 reads, 26/385 (7%) of the samples contained 20,000–10,000 reads, 44/385 (11%) of the samples contained 10,000–1000 reads and 66/385 (17%) of the samples contained between 1000–0 reads. We used the Bray-Curtis and weighted Unifrac indices to assess the taxonomic and phylogenetic β-diversity, respectively (Fierer et al., 2012; Fukuyama, 2019; Lozupone et al., 2007, 2011).

Since we observed a variable number of reads among the samples from the different soil layers, which decreased along soil depth (Table S1), we also assessed the impact of soil depth versus geographic location on the bacterial β-diversity using a similar number of reads among samples. For that, data were rarified to 1000 reads due to the low number of sequences in the samples from the deepest soil layers. Using this rarefaction cutoff, a total of 319 samples were used in the analysis. For α-diversity analyses, the Shannon index, Faith's PD and observed ASVs were calculated to determine the species diversity, species phylogenetic diversity and species richness, respectively. The data were also split for each soil layer in order to analyse the bacterial communities inhabiting the different soil depths separately. The goal here was not to compare the influence of soil depth versus geographic location, but to investigate the factors contributing to the β-diversity variation in the horizontal dimension (i.e., across geographical locations) at each specific soil depth. Therefore, samples from each soil depth were rarefied to different values, according to the number of reads available among them (Table S1): 43,000 (0–30 cm), 14,400 (30–60 cm), 10,900 (60–90 cm), 1000 (90–120 cm), 1900 (120–150 cm), 4000 (150–180 cm), 410 (180–210 cm) and 110 (210–240 cm) reads. All the Bray-Curtis and weighted Unifrac matrices were exported for statistical analyses.

2.4 | Statistical analysis

The Bray-Curtis and weighted Unifrac matrices were exported to the R software where all multivariate analyses were performed using the vegan package version 2.5-6 (Dixon, 2003). Constrained principal coordinate analyses (CAP) were generated using the "capscale" function and the ordination of samples was visualized with the "ggplot" function in the ggplot package version 3.3.0 (Wickham, 2011). Permutative multivariate analysis of variance (PERMANOVA, 999 permutations) was performed to test if the bacterial communities were significantly affected by soil depth, geographic location, latitude, longitude and their interaction using the "adonis" function. Both soil depth and geographic location were used as categorical factors in the PERMANOVA, in order to compare the influence of vertical and horizontal spatial variation, respectively. Data from each soil layer were also individually analysed with PERMANOVA to compare the influence of geographic location on the bacterial communities in the different layers (Bell, 2010). The ASV table was used for a canonical correspondence analysis (CCA) using the "cca"

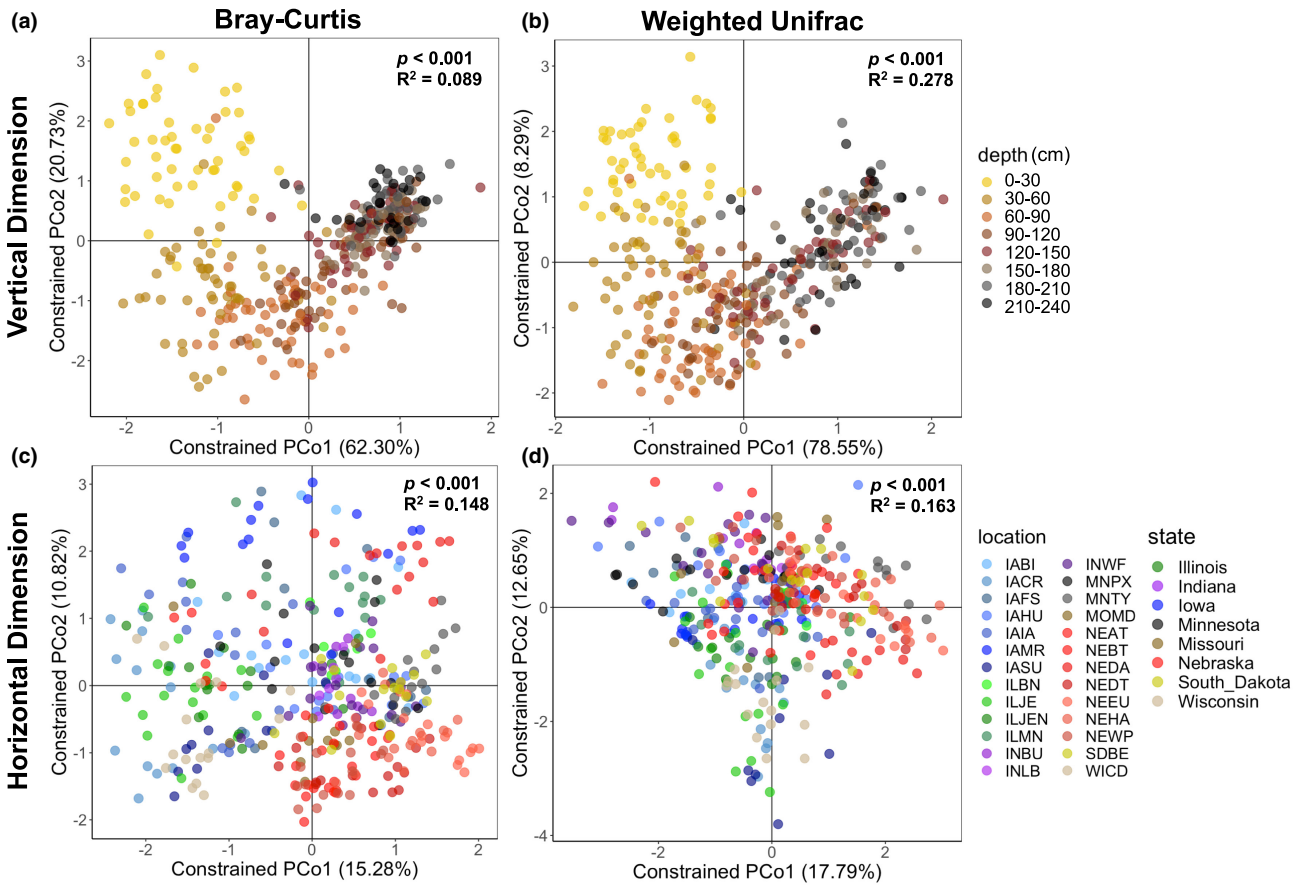

FIGURE 2 Bacterial taxonomic and phylogenetic β-diversity according to soil depth or geographic location. Constrained principal coordinate analysis (CAP) showing the separation of samples from the different soil layers (a, b) or geographic locations (c, d) based on the Bray-Curtis (taxonomic β-diversity) or weighted Unifrac (phylogenetic β-diversity) indices using the unrarefied data. All samples were used in this analysis (Table S1). The sample ordination was constrained for the factors soil depth (a, b) or location (c, d). The different soil layers and the geographic locations represent the vertical (a, b) and horizontal (c, d) spatial variation in our study, respectively. Results of PERMANOVA (p-value and R^2) using the "adonis" function are shown and corresponds to the factor tested and index used in each graph. In the graphs showing the geographic locations, the legend codes reflect each specific place where the samples were collected. Different locations from the same state are shown in distinct shades of the same colour.

function with backward selection to detect the bacterial genera that significantly affected the changes in bacterial community structure between soil depth layers. In this case, the abundance of all specific genera was used as the second table. The same process was used to detect soil physicochemical variables significantly affecting the bacterial communities along soil depth, but the values of physicochemical variables of the samples were used as the second table. The vectors of significant variables associated with the samples of the different soil depth layers were identified in the CCA plots. CCA was also used to identify the soil variables affecting the bacterial communities between locations (horizontal variation) in each individual soil layer.

A matrix containing the geographic distances in kilometres between each of the 26 locations was constructed. Mantel tests were performed to assess the existence of significant correlations between the geographic distance matrix and the Bray-Curtis matrix of each soil layer. In addition, the Spearman's correlation coefficient Rho (ρ) resulting from the Mantel tests was used as a measure of the degree of the bacterial community distance-decay in each soil layer (Bowman & Arnold, 2021). The higher the ρ value, the greater is the distance-decay. Mantel tests were also used to test for correlations between the bacterial communities and the soil physicochemical data in each layer of the soil profile. The resulting ρ-values were used to compare the influence of edaphic variables and geographic distance on the β-diversity of each soil layer. Ecological models were used to infer the species rank abundance distribution (SAD) in the samples in order to verify whether the bacterial communities from the different soil depths were assembled due to neutral (stochastic) or niche-based (deterministic) processes. For that, we used the "radfit" function in the vegan package version 2.5-6 (Dixon, 2003). The models tested were the Broken-stick (null) which represents neutral processes, and the Pre-emption, Log-normal, Zipf and Zipf-Mandelbrot which represent niche-based processes. The resulting Akaike information criterion (AIC) of the five models were compared in each sample. The lowest AIC value indicated the best-fit model to explain the rank abundance distribution of each sample (Bozdogan, 1987; Dumbrell et al., 2010). Differences in relative abundance of bacterial phyla and genera along the soil depth profile were

analysed with the Welch's *t*-test and the Benjamini-Hochberg FDR *p*-value correction using the STAMP software (Parks et al., 2014).

3 | RESULTS

3.1 | Bacterial phylogenetic β-diversity was impacted more by soil depth than geographic location

This study sought to characterize the spatial changes and the factors that impact the bacterial community distribution in top and subsoils across a large section of midwestern USA, which is one of the most productive agricultural regions in the world (Figure 1). The 16S rRNA amplicon sequencing resulted in 23,134,871 reads, from which 18,849,502 passed all quality filters comprising a total of 75,140 individual ASVs.

Here we compared the analysis of unrarefied to rarefied data and used two contrasting methods of analysis to capture different aspects of the β-diversity. Taxonomic β-diversity was determined using Bray Curtis that uses counts (not actual phylogenetic relationships) to represent specific taxa and compares the difference in the number of different taxa in each sample. In contrast, weighted Unifrac (phylogenetic β-diversity) considers the phylogenetic relationships or distances of the taxa in the communities. The rarefied and unrarefied data multivariate analyses indicated that soil depth down to 240 cm (Figure 2a,b) and geographic location across >1170 Km of distance (Figure 2c,d) significantly affected the bacterial community structure (Table S2). However, the two different β-diversity indices (Bray-Curtis and weighted Unifrac) revealed distinct results regarding the impact of each factor on the bacterial community structure. The unrarefied analysis using the Bray-Curtis index showed a larger influence of geographic location on the β-diversity (14.8% of the variation) as compared to soil depth (8.9% of the variation) (Figure 2; Table S2). On the other hand, the phylogenetic β-diversity as assessed with the weighted Unifrac index was less impacted by geographic location (16.3% of variation) than by soil depth (27.8% of variation) (Figure 2; Table S2). The results of the analysis using the rarefied data (1000 reads per sample) also showed that the taxonomic β-diversity (Bray-Curtis) was impacted more by geographical location (16.7% of the variation) than soil depth (9.9% of the variation), while the phylogenetic β-diversity (weighted UniFrac) was less impacted by geographic distance (18.2% of the variation) and more by soil depth (28.3% of the variation; Figure S1). In sum, these results showed that the changes in bacterial community structure of these agricultural soils were greater due to soil depth (vertical variation along 240 cm) than geographic location (horizontal variation > 1170 km) when the phylogenetic information (weighted UniFrac) was used as compared to the analysis with the Bray-Curtis index. However, the influence of location was virtually the same using weighted UniFrac and Bray-Curtis (Figure 2; Figure S1). The influence of location on the β-diversity was more clearly observed in the ordination using the rarefied data than using the unrarefied data according to both indices (Figure 2; Figure S1), probably because the rarefaction reduced the differences between soil layers associated with variable number of reads (Table S1), and therefore optimized the clustering of samples by geographic location (Figure S1).

Latitude and longitude also significantly affected the bacterial β-diversity according to both indices, with a slightly greater influence of longitude than latitude regardless of the β-diversity index used (Figure S2, Table S2). The larger influence of longitude on the bacterial communities reflected the greater variation in longitude (up to 1170 km) than latitude (up to 573 km) in our study. However, latitude and longitude explained only 1.0%–2.6% of the sample variation, while location explained 14.8%–16.3% of the sample variation (Table S2), suggesting that other factors in addition to the latitude and longitude of the geographic locations contributed to shape the bacterial communities in the horizontal spatial dimension. Soil depth showed significant interactions with location, latitude and longitude for both β-diversity indices (Table S2). The α-diversity analyses indicated that the species richness, species diversity and species phylogenetic diversity decreased in the deeper soil layers (Figure S3). The species richness based on the observed ASVs decreased significantly from the topsoil to the following soil depth layers in the order: 0–30 > 30–60 > 60–90 > 90–120 >120–150 cm = 150–180 > 180–210 = 210–240 cm (Figure S3a). The species diversity based on the Shannon index decreased significantly from the topsoil to the following layers in the order: 0–30 > 30–60 > 60–90 = 90–120 >120–150 = 150–180 > 180–210 = 210–240 cm (Figure S3b). The species phylogenetic diversity based on the Faith's PD index decreased significantly from the topsoil to the following layers in the order: 0–30 > 30–60 = 60–90 > 90–120 >120–150 = 150–180 > 180–210 = 210–240 cm (Figure S3c).

3.2 | Bacterial phylotypes and soil properties changed along the soil profile with *Pseudomonas* dominating the subsoil

There was a large shift in the bacterial community composition along the soil profile, observed at both phylum and genus levels (Figure 3a,b). At the phylum level, there was a substantial enrichment of Pseudomonadota (previously Proteobacteria) and depletion of Acidobacteriota (previously Acidobacteria), Chloroflexota (previously Chloroflexi), Verrucomicrobiota (previously Verrucomicrobia), Actinomycetota (previously Actinobacteria) and Planctomycetota (previously Planctomycetes) in the deepest soil layers (Figure 3c). The significant shifts in the relative abundance of many dominant bacterial phyla (Figure 3a,c) explain the large changes in phylogenetic β-diversity along soil depth (Figure 1b), since phylum is the highest hierarchical level in taxonomy after domain. The large changes at the phylum level indicate that the bacterial taxa in deep soils have a very different phylogenetic makeup compared to those inhabiting the topsoil. Significant changes in relative abundance were also observed for less abundant phyla such as Latescibacterota, Methylomirabilota, Armatimonadota, Nitrospirota, Sumerlaeota and GAL15 (Figure 3c).

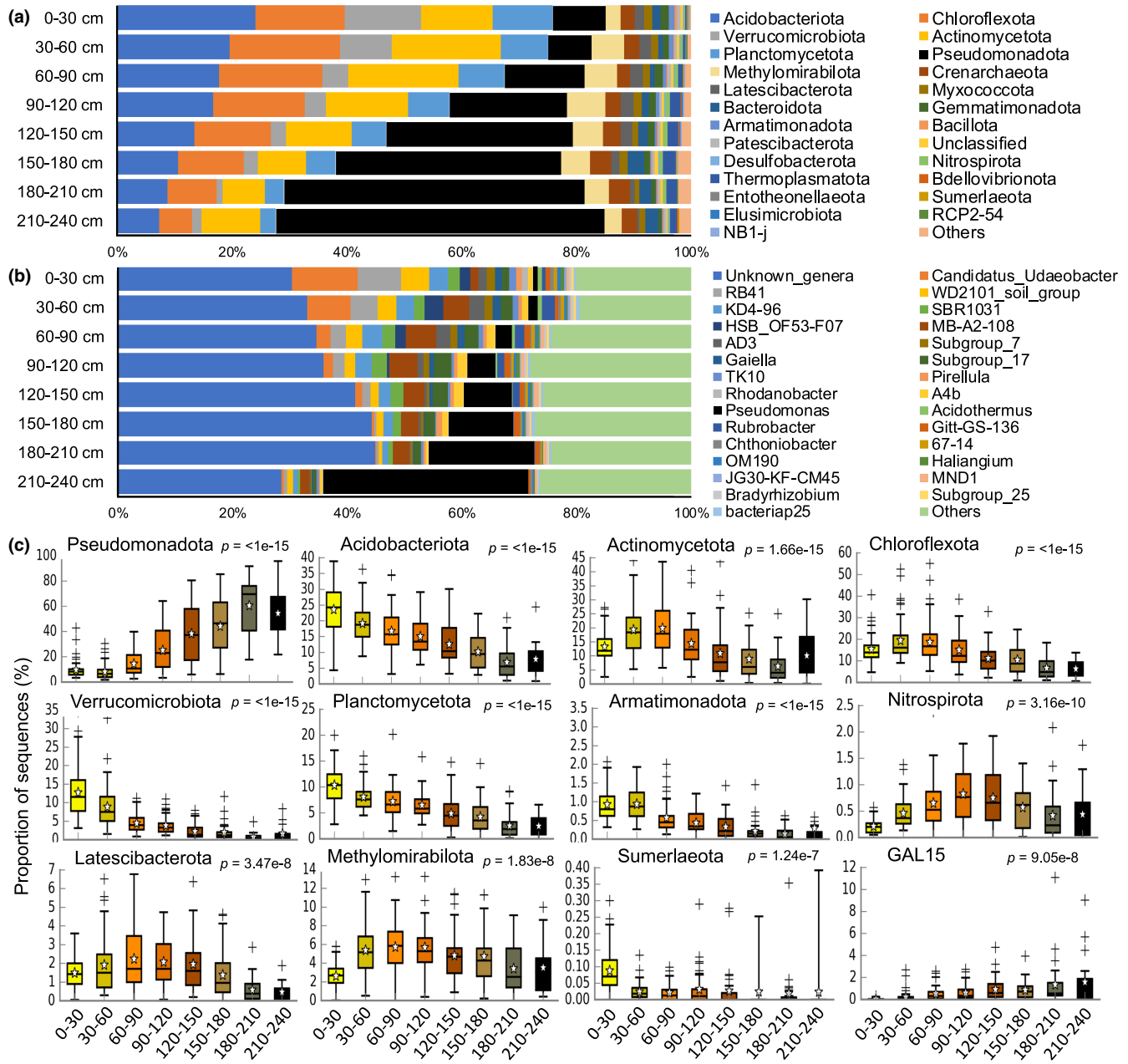

FIGURE 3 Shifts in bacterial community composition along soil depth. (a–b) Horizontal bar charts showing the average relative abundance of the main phyla (a) and genera (b) in each soil layer analysed considering all sampling locations. Others indicate phyla (a) or genera (b) with average relative abundance below 0.1% and 0.4%, respectively. Unknown genera indicate the ASVs that were unclassified at the genus level. (c) Box plots showing the main phyla significantly changing in relative abundance between the soil depth layers according to the Welch's *t*-test and the Bonferroni *p*-value correction. The stars in the boxes indicate the average relative abundance of each phylum in each soil layer considering all locations sampled.

The largest shift in relative abundance was observed for the phylum Pseudomonadota, which progressively dominated the deep soils (Figure 3a,c). An investigation at the genus level indicated that this Pseudomonadota dominance was mainly due to *Pseudomonas*, which was the genus with the largest relative abundance below 120 cm of soil depth (Figure 3b). In the deepest soil layer analysed (210–240 cm) *Pseudomonas* average relative abundance was 27.5% (Table S3). In addition to *Pseudomonas*, 86 other genera changed in relative abundance along the soil profile (Table S3). Most of them decreased in relative abundance along soil depth, for example, *Candidatus Udaeobacter* (Verrucomicrobiota), RB41 (Acidobacteriota), WD2101 soil group (Planctomycetota) and KD4-96 (Chloroflexota) (Figure 3b). Canonical correspondence analysis (CCA) confirmed that *Pseudomonas* was the genus with the largest impact on the bacterial community changes along soil depth (Figure 4a; Table S4). *Pseudomonas* was clearly associated with samples from the deepest soil layers in the CCA ordination (Figure 4a). Other genera that significantly contributed to the differences in bacterial

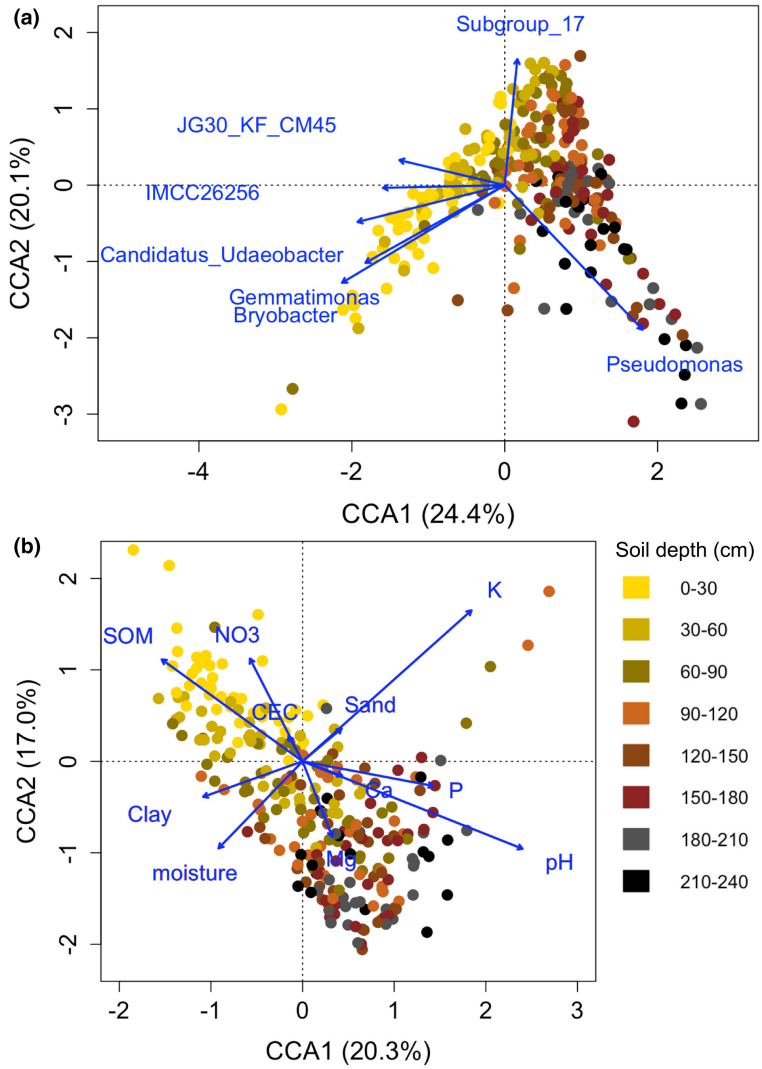

FIGURE 4 Canonical correspondence analysis showing the main bacterial genera and soil physicochemical properties affecting the bacterial community structure along soil depth. (a) CCA showing the seven bacterial genera with the largest influences on the bacterial community changes along soil depth. The ASV table was used for the analysis and the second table contained the abundance of all specific genera. A backward selection model was used to identify the genera significantly ($p < .05$) affecting the bacterial communities. The seven genera with largest effects (chi-square values) were used in the CCA plot. The direction of the vectors indicates where the relative abundance of each genera increased. (b) CCA showing the soil physicochemical variables significantly affecting the bacterial community structure and associated with the different soil depth layers. The ASV table was used for the analysis and the second table contained the values of soil physicochemical variables for each sample. A backward selection model was used to identify the soil variables significantly ($p < .05$) affecting the bacterial communities. The direction of the vectors indicates where the values of each variable were greater. Soil organic matter (SOM) and nitrate (NO3) were associated with the bacterial communities in the uppermost soil layers, while pH, magnesium (Mg) and phosphorus (P) were associated with the bacterial communities in the deepest soil layers. The other variables were not clearly associated with changes along soil depth.

community structure between soil depth layers were *Bryobacter* (Acidobacteriota), *Gemmatimonas* (Gemmatimonadota), *Candidatus Udaeobacter* (Verrucomicrobiota), *IMCC26256* (Actinomycetota), *JG30-KF-CM45* (Thermomicrobiota) which were associated with shallow/surface soil samples, and *Subgroup17* (Acidobacteriota) which was associated with samples from intermediate soil depth layers (Figure 4a).

We also performed a CCA with the physicochemical data, which indicated that variation in many soil physicochemical variables significantly affected the bacterial communities studied, including SOM, NO_3^-, CEC, sand and clay content, soil moisture, Mg, Ca, P and pH (Figure 4b, Table S4). Among the soil physicochemical variables analysed, only silt content and NH_4^+ did not affect the bacterial communities (Table S4). The two soil variables that showed the largest association with the changes in the bacterial community composition down the soil profile were pH which significantly increased and SOM that significantly decreased in deeper soil layers (Figure 4b; Figure S4). In addition to SOM, the soil variable NO_3^- was also correlated with the bacterial communities in the topsoil. On the other hand, the variables P and Mg were correlated with the

bacterial communities in the deepest soil layers, as observed for pH (Figure 4b). Variation in the other soil variables were not clearly correlated with soil depth, but possibly with other factors such as location.

3.3 | Bacterial community distance-decay in the horizontal dimension was stronger in the subsoil than in the topsoil and niche-based processes were dominant in all soil layers

We next separated the samples from each of the eight soil layers to verify in which soil layers the bacterial communities were more impacted by the horizontal spatial variation. PERMANOVA showed that the factor location significantly affected the structure of the bacterial communities in all soil layers, but it explained a lower percentage of the sample variation in the topsoil (64%) than in the subsoil layers (70%–74%) (Figure S5). Furthermore, we generated a matrix containing the pairwise distances in kilometres between each of the 26 locations sampled to check if geographic distance was correlated with the bacterial community structure of each soil layer using Mantel tests. Results indicated that the bacterial communities from all soil layers were significantly correlated with geographic distance (Figure 5). As the geographic distance between the locations increased, there was a trend of increasing dissimilarity between the soil bacterial communities. However, the lowest distance-decay (ρ value) was observed in the topsoil bacterial communities (Figure 5). The lowest slope and R^2 of the regressions were also observed in the topsoil (Figure 5). These results indicated that the topsoil harbours the bacterial communities with the weakest distance-decay relationships.

In addition, we aimed to assess whether edaphic variables or geographic distance had a greater influence on the structure of bacterial communities in the different soil depths. Mantel tests indicated significant correlations between the soil physicochemical data and the bacterial communities at all soil depths, but the influence of edaphic variables was larger than geographic distance only in the topsoil according to ρ values (Table 1). However, specific key soil variables may have had a greater effect than the entire set of soil physicochemical data. Therefore, CCAs were conducted to identify the soil variables significantly affecting the bacterial community composition at each soil depth layer. The bacterial communities were affected by a different set of soil variables at each soil layer, but pH and K were present virtually at all depths (Figure S6). In addition, pH was the best predictor of the bacterial community changes in the four uppermost soil layers (0–120 cm), while K was the best predictor in the four deepest layers studied (120–240 cm) according to chi-square values (Table S5).

Based on the CCA results, we used the set of soil variables as well as the best predictor variable significantly affecting the bacterial communities in each soil layer for Mantel tests. Results indicated a larger impact of selected soil variables and pH than geographic

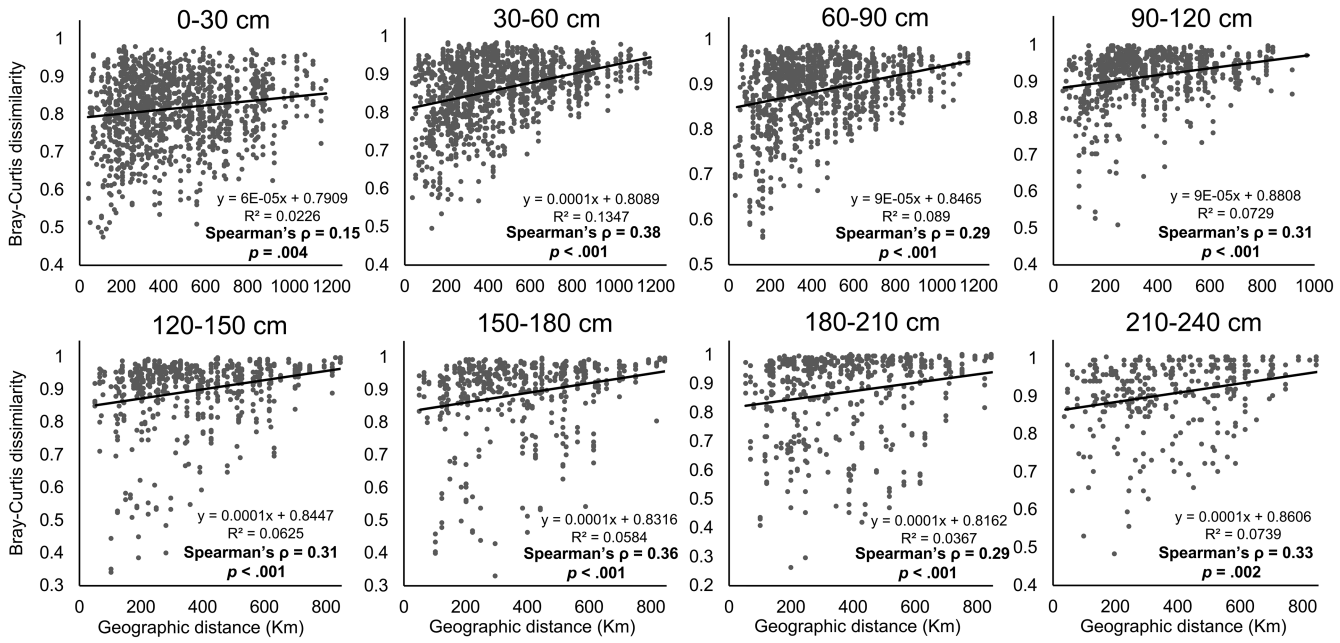

FIGURE 5 Dissimilarity of the bacterial communities as a function of geographic distance between locations in each soil depth layer. The dots in the graph represents the pairwise distances (Bray-Curtis index and geographic distance in km) between all samples. The fitted line shows the distance-decay relationship of the bacterial communities in each soil depth layer. The steeper the slope of the line, the greater the bacterial community distance-decay. In addition, the results of Mantel tests are shown in the graphs and indicate if there was a significant multivariate correlation ($p < .05$) between the bacterial communities of each soil layer and geographic distance. The Spearman's correlation coefficient rho (ρ) is also shown in the graphs and indicates the degree of the bacterial community distance-decay in each soil layer. The greater the ρ value the larger the distance-decay. The slope and R^2 of the regressions are also shown.

TABLE 1 Factors influencing the bacterial β-diversity at each soil depth layer.

	Geographic distance		All soil variables		Selected soil variables			Best predictor soil variable		
Soil layer	*p*-value	Rho (ρ)	*p*-value	Rho (ρ)		*p*-value	Rho (ρ)		*p*-value	Rho (ρ)
0–30 cm	.004	0.147	.004	0.211	pH, Mg, NO_3, K	<.001	0.195	pH	<.001	**.463**
30–60 cm	<.001	0.378	<.001	0.248	pH, Ca, P, K, Silt	<.001	0.298	pH	<.001	**.429**
60–90 cm	<.001	0.291	.005	0.196	pH, Sand, NH_4, SOM, moisture, Mg, Clay, K	<.001	0.235	pH	<.001	**.543**
90–120 cm	<.001	**0.308**	.004	0.235	pH, K, Sand, NO_3, P, K, Mg	.006	0.216	pH	<.001	.286
120–150 cm	<.001	**0.306**	.007	0.236	K, SOM, moisture, NO_3, Silt, pH, Ca, CEC, NH_4	.007	0.221	K	<.001	.236
150–180 cm	<.001	**0.364**	.015	0.211	K, NO_3, Silt, Clay, Mg	<.001	0.297	K	<.001	.315
180–210 cm	<.001	0.287	.011	0.276	K, NO_3, pH, Sand, Ca	<.001	**0.439**	K	<.001	.420
210–240 cm	.002	0.330	.021	0.244	K, pH	<.001	**0.363**	K	.004	.359

	Percentage of samples best fitting to each model according to the Akaike information criterion (AIC)				
Soil layer	Broken-stick (null), %	Pre-emption, %	Log-normal, %	Zipf, %	Zipf-Mandelbrot, %
0–30 cm	0.00	0.00	**65.91**	0.00	34.09
30–60 cm	0.00	0.00	**83.72**	0.00	16.28
60–90 cm	0.00	0.00	**74.42**	6.98	18.60
90–120 cm	0.00	0.00	**63.89**	13.89	22.22
120–150 cm	0.00	0.00	**47.06**	23.53	29.41
150–180 cm	0.00	0.00	**50.00**	31.25	18.75
180–210 cm	0.00	3.44	27.59	**41.38**	27.59
210–240 cm	5.56	0.00	22.22	**44.44**	27.78

Note: Results of Mantel tests (999 permutations) between the Bray-Curtis distance matrices of each soil layer and (1) the geographic distance matrix, (2) the matrices with all soil physicochemical variables in each soil layer, (3) the matrices with the set of selected soil physicochemical variables that significantly affected the bacterial communities in each soil layer (Table S5), and (4) the soil physicochemical variable best predicting the changes in the bacterial communities in each soil layer (Table S5). The *p*-value and the Spearman's correlation coefficient Rho (ρ) is shown for each Mantel test. Higher ρ values indicate stronger correlations. The highest ρ found in the correlation tests for each soil depth layer is highlighted in bold. Below are the ecological models explaining the species rank abundance distribution (SAD) among the samples of each soil depth layer. The percentage of samples with the lowest Akaike Information Criterion (AIC) for each model is shown at each soil depth. The model with lowest AIC fits best to explain the SAD of each sample. The ecological models assessed were Broken-stick (neutral model), Pre-emption, Log-normal, Zipf and Zipf-Mandelbrot (niche-based models). The model showing the highest percentage of samples with lower AIC values for each soil depth layer is highlighted in bold.

distance on the bacterial communities for the depth 0–30 cm, a larger effect of pH than geographic distance on the bacterial communities for the depths 30–90 cm, a larger effect of geographic distance than any edaphic variable at depths 90–180 cm, and a larger influence of selected soil variables and K than geographic distance in the two deepest layers (180–240 cm) according to ρ-values (Table 1). Therefore, our results suggest that the influence of geographic distance can be greater than the influence of edaphic variables on the bacterial biogeography of some subsoil layers (90–180 cm). Finally, we assessed whether the assembly of the bacterial communities inhabiting the different soil depth layers was due to neutral or niche-based mechanisms. The ecological models suggested that niche-based mechanisms explained the species abundance distribution (SAD) of the bacterial communities in all soil layers studied (Table 1, Table S6). The log-normal, Zipf and Zipf-Mandelbrot models which represent deterministic processes explained the bacterial community assembly of samples from all soil layers (Table 1). Most samples (47.06%–83.72%) from the 0–180 cm fit the log-normal model best, while most samples (41.38%–44.44%) from the 180–240 cm fit the Zipf model best (Table 1).

4 | DISCUSSION

This is the first comprehensive soil bacterial biogeography study in the U.S. Corn Belt, which is one of the most productive agricultural regions for food, feed and energy (Riccetto et al., 2020). Our exploration of deeper soil profiles across a vast geographic region in agroecosystems allowed new insights into multiple questions in microbial biogeography. We found a larger influence of geographic location on the bacterial taxonomic β-diversity, but a greater impact of soil depth on the phylogenetic β-diversity – even considering the much larger horizontal than vertical spatial variation in our study (>1170 km vs. only 240 cm, respectively). Phylogenetic β-diversity

accounts not only for the different bacterial ASVs and their abundance (community structure), but also their evolutionary relatedness (Graham & Fine, 2008; Lozupone et al., 2007; Wang et al., 2013). Therefore, our results indicate that the bacterial taxa inhabiting the topsoil are in general different and unrelated to those inhabiting the deep soils, while the changes across the horizontal space are more associated with community organization and not as much with taxa phylotype. These results agree with a previous study showing that several phyla that were more abundant in the subsoil were rarer in the topsoil (Brewer et al., 2019). The large phylogenetic shifts in the bacterial communities along the soil profile were probably associated with the differences in the soil environmental conditions. For example, most bacteria in topsoil are heterotrophic and aerobic but carbon and oxygen availability decreases with depth (Fierer, 2017; Paul, 2014; Sun et al., 2020; Wordell-Dietrich et al., 2017), which may select for bacterial taxa that possess alternative types of metabolism belonging to different phylogenetic groups. SOM was found to be one of the main soil properties associated with the bacterial community changes along soil depth (Figure 4b), which was also observed in previous studies (Eilers et al., 2012; Hao et al., 2021). Therefore, the lower carbon availability was probably a strong selective factor for the phylotypes thriving in the subsoil.

Pseudomonadota (previously Proteobacteria) dominated the deepest soil layers, more specifically the genus *Pseudomonas*. Both Pseudomonadota and *Pseudomonas* are known to comprise taxa with many different types of metabolic capabilities and functions (Kersters et al., 2006; Silby et al., 2011). The genomes of *Pseudomonas* are especially adaptable to many different environmental conditions (Silby et al., 2011). Moreover, they have been found to possess a great metabolic plasticity such as in situations with different oxygen availability – several *Pseudomonas* spp. are facultative anaerobes, despite aerobic respiration being their primary form of metabolism (Kampers et al., 2021; Tribelli et al., 2019). Therefore, the enrichment of these taxa in the subsoil was probably enabled by their great metabolic diversity and adaptability. In a recent study, Pseudomonadota and Pseudomonadaceae were enriched in very alkaline as compared to neutral soils in the Nebraska Sandhills (Lopes et al., 2021). Therefore, the dominance of those taxa in the subsoil may also be associated with its higher pH, since it was correlated with the bacterial community changes along soil depth. It is unclear if *Pseudomonas* also dominates the subsoils of other geographic regions, since they show a strong degree of endemicity, that is, strains with distinct genetic background inhabit the different places (Cho & Tiedje, 2000; Lopes et al., 2018).

Another interesting taxon associated with the bacterial community changes along soil depth was the ubiquitous genus *Candidatus Udaeobacter* (Verrucomicrobiota), which was depleted in deeper soils (Brewer et al., 2017; Willms et al., 2021). A previous study revealed their aerobic and heterotrophic lifestyle and a low metabolic versatility due to a reduced genome size, which may limit their survival in subsoils (Brewer et al., 2017). In addition, bacteria from this genus have preference for acidic soils (Willms et al., 2021). Other taxa enriched in topsoils and depleted in subsoils were *Bryobacter* (Acidobacteriota) and *Gemmatimonas* (Gemmatimonadota), but less information is available in the literature regarding these genera. The few studies available suggest that *Bryobacter* is also an aerobic heterotroph that prefers acidic soils (Kulichevskaya et al., 2010), while *Gemmatimonas* has obligate aerobic species performing N_2O reduction or heterotrophic photosynthesis (Koblizek et al., 2020; Park et al., 2017; Zeng et al., 2021). Therefore, the physicochemical variation from top to subsoils including decreased O_2 and light availability, increased pH, and lower N and SOM concentrations were directly associated with the main indicator genera of the bacterial community changes along the soil profile. In addition to the impact of the physicochemical variation, the large change in phylogenetic β-diversity along soil depth was probably associated with the reduced bacterial α-diversity in the subsoil that may limit the establishment of many phylotypes present in the topsoil (Brewer et al., 2019; Eilers et al., 2012; Hao et al., 2021). Previous studies showed that microbial biomass and bacterial abundance decrease exponentially with soil depth (Hao et al., 2021; Sun et al., 2020), which may contribute to the lower α-diversity in the subsoil.

Our results also indicated that the topsoil bacterial communities showed a lower distance-decay than those in the subsoil. The large anthropogenic interference in the topsoil (e.g., soil tillage, fertilization, plant carbon inputs) may to some extent homogenize the topsoil microbiome of agricultural fields and reduce the influence of local environmental factors (Rodrigues et al., 2013; Wang et al., 2022). In contrast, the bacterial communities inhabiting the subsoil may be subjected to fewer anthropogenic influences and therefore are more prone to selection by local factors associated with pedogenic processes. In addition, the subsoil bacterial communities probably have more influence of dispersal limitation, since the dispersion of bacteria in the subsoil is constrained by the physical barrier of topsoil. Despite soil bacteria being horizontally dispersed by water movement, the wind is a main factor for long distance dispersion which requires contact with the atmosphere (Smith et al., 2013; Tecon & Or, 2017). Both selection by local factors and dispersal limitation are processes that contribute to the increase in the bacterial community distance-decay (Bell, 2010; Hanson et al., 2012). Therefore, the greater selection and dispersal limitation in the subsoil probably explains the larger distance-decay of the bacterial communities in deeper soil layers. Similar results were found in undisturbed grassland soils in China and in the Qinghai-Tibet Plateau (alpine steppes and alpine meadow), where the effect of geographic distance also increased in the subsoil compared to the topsoil (Du et al., 2021; Li et al., 2022). Therefore, both natural and agricultural ecosystems have a lower distance-decay in the topsoil than in the subsoil, which may be maximized in agroecosystems due to the human activities in the topsoil.

Previous studies showed that the topsoil bacterial communities are more influenced by edaphic variables than geographic distance, with pH as the best predictor of changes in community structure (Fierer & Jackson, 2006; Lauber et al., 2009; Griffiths et al., 2011; Karimi et al., 2018; Bahram et al., 2018). In our study, the topsoil bacterial communities were also more influenced by edaphic variables

than geographic distance and this effect was observed down to 90 cm of soil depth, with pH being the best predictor of the bacterial community changes. However, beyond that depth, from 90–180 cm, geographic distance had a greater impact on bacterial community structure than any soil variable according to ρ values (Table 1). The greater influence of distance than edaphic variables suggests that below 90 cm of soil depth dispersal limitation was stronger, probably because the bacterial communities are so deep in the soil profile that dispersion is severely impaired (Hanson et al., 2012). A previous subsoil biogeography study also detected increased dispersal limitation in deeper soil horizons (Li et al., 2022). However, the two deepest soil layers analysed in this study (180–240 cm) showed again a larger influence of edaphic variables than geographic distance, particularly due to K concentrations. A previous study analysing shifts in topsoil bacterial communities between soils of different pedogenic types also found a large influence of K in addition to pH (Pershina et al., 2018). Therefore, the strong impact of K on the β-diversity of the deep soils might be associated with a larger pedogenic influence on the subsoil than on the topsoil bacterial communities in the studied agricultural fields.

The reduced bacterial abundance and α-diversity in the deepest subsoil layers (Brewer et al., 2019; Eilers et al., 2012; Hao et al., 2021; Sun et al., 2020) probably increased the influence of ecological drift as well, since drift is more impactful in small and less diverse populations (Nemergut et al., 2013; Siqueira et al., 2020). In addition, a previous study suggested that the importance of drift on bacterial communities increases under high selection and dispersal limitation (Fodelianakis et al., 2021), which seem to be both stronger in the subsoil. As noted above, *Pseudomonas* relative abundance tripled from the 90–180 cm to the 180–240 cm soil layer (Table S3), which was probably maximized by the effect of ecological drift in the deepest subsoil layers that may have eliminated other phylotypes, contributing to *Pseudomonas* dominance. Although dispersal limitation and ecological drift are probably more effective in deeper soils, our results showed that the bacterial community assembly in all soil layers fit best to niche-based models, suggesting that selection was always stronger compared to neutral processes, which agrees with previous subsoil biogeography studies in undisturbed grasslands (Du et al., 2021; Li et al., 2022). Therefore, this and previous studies indicate that the assembly of topsoil and subsoil bacterial communities are influenced more by deterministic than stochastic processes in both natural and agroecosystems. Here, most samples from 0–180 cm fit to the log-normal model best, which is the most commonly observed SAD in communities of macro and microorganisms (Shoemaker et al., 2017). On the other hand, most samples from 180–240 cm fit to Zipf, which is also a deterministic model but implies a different community organization. The Zipf is a power-law model that describes one of the most uneven forms of SAD, with a greater dominance of the most abundant taxa (Shoemaker et al., 2017). Therefore, we speculate that the bacterial communities in the deepest soils fit best to the Zipf model due to the dominance of *Pseudomonas*. Despite being less common, the Zipf model also explained the microbial community assembly in a few studies (Ashby et al., 2007; Cutler et al., 2015; Leon-Palmero et al., 2018; Yan et al., 2020).

Overall, our study showed that bacterial phylogenetic β-diversity changes more in the vertical than in the horizontal spatial dimension in a highly productive agricultural region in the midwestern USA. However, this might not be true for larger horizontal spatial scales comprising different continents and biomes, which requires further investigation. We also showed that the topsoil bacterial communities of agricultural fields have a lower distance-decay than those in the subsoil, which is similar to what was observed in undisturbed grassland ecosystems. Moreover, in our study geographic distance had a larger influence than edaphic variables in some subsoil layers, which is fundamentally different to our knowledge of the bacterial biogeography in the topsoil. These results suggest that different factors drive the spatial distribution of bacterial communities in the subsoil, probably associated with stronger selection, dispersal limitation and ecological drift than in the topsoil. The species rank abundance distribution also changed along the soil profile, with a different community organization at the deepest soil layers. However, deterministic mechanisms were prevalent in all soil layers analysed, indicating that selection was the most important ecological process in topsoils and subsoils. In sum, we showed that studying the processes shaping the microbiome not only in shallow soil layers, but also in the deepest soils will be essential for a comprehensive knowledge of microbial biogeography. Understanding how microbiomes are shaped across the space can impact other important research areas such as sustainable agriculture, carbon sequestration and climate change.

AUTHOR CONTRIBUTIONS

Lucas Dantas Lopes analysed the sequencing data, performed multivariate statistical analyses and wrote the manuscript. Stephanie L. Futrell prepared the sequencing libraries. Emily E. Wright collected the soil cores from the fields and organized the samples for analysis. Gerasimos J. Danalatos synthesized soil databases and performed soil analysis. Michael J. Castellano and Tony J. Vyn edited the manuscript. Sotirios V. Archontoulis coordinated field data collection, acquired funding and edited the manuscript. Daniel P. Schachtman coordinated the study, acquired funding and edited the manuscript.

ACKNOWLEDGEMENTS

We thank Mari Butler from University of Nebraska at Lincoln for her help with DNA extractions, Garrett Verhagen from Purdue University and Mickalla Stallman from Iowa State University for their assistance with soil data collection from the fields, Bayer Crop Science Testing Team for providing meta-data and identifying fields for experimentation, Clarice Mensah and Slobodan Trifunoivic for valuable comments.

FUNDING INFORMATION

This work was supported by USDA-NIFA award 2020-67019-31796 to Daniel Schachtman, and by FFAR to Sotirios Archontoulis and Bayer Crop Science on the project entitled “Evaluating the relative influence of maize breeding, field management, and environmental setting on crop production and sustainability targets”.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The sequences used in this manuscript have been submitted to NCBI-SRA under the BioProject accession no. PRJNA824344.

ORCID

Lucas Dantas Lopes https://orcid.org/0000-0002-6355-3599

REFERENCES

Ashby, M. N., Rine, J., Mongodin, E. F., Nelson, K. E., & Dimster-Denk, D. (2007). Serial analysis of rRNA genes and the unexpected dominance of rare members of microbial communities. *Applied and Environmental Microbiology*, *73*(14), 4532–4542.

Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., Bengtsson-Palme, J., Anslan, S., Coelho, L. P., Harend, H., Huerta-Cepas, J., Medema, M. H., Maltz, M. R., Mundra, S., Olsson, P. A., Pent, M., Polme, S., Sunagawa, S., Ryberg, M., … Bork, P. (2018). Structure and function of the global topsoil microbiome. *Nature*, *560*(7717), 233–237.

Bell, T. (2010). Experimental tests of the bacterial distance-decay relationship. *The ISME Journal*, *4*, 1357–1365.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensive microbiome data science using QIIME2. *Nature Biotechnology*, *37*, 852–857.

Bowman, E. A., & Arnold, A. E. (2021). Drivers and implications of distance decay differ for ectomycorrhizal and foliar endophytic fungi across an anciently fragmented landscape. *The ISME Journal*, *15*, 3437–3454.

Bozdogan, H. (1987). Model selection and Akaike's information criterion: The general theory and its analytical extensions. *Psychometrika*, *52*, 345–370.

Bray, R. H., & Kurtz, L. T. (1945). Determination of total, organic, and available forms of phosphorus in soils. *Soil Science*, *59*, 39–45.

Brewer, T. E., Aronson, E. L., Arogyaswamy, K., Billings, S. A., Botthoff, J. K., Campbell, A. N., Dove, N. C., Fairbanks, D., Gallery, R. E., Hart, S. C., Kaye, J., King, G., Logan, G., Lohse, K. A., Maltz, M. R., Mayorga, E., O'Neill, C., Owens, S. M., Packman, A., … Fierer, N. (2019). Ecological and genomic attributes of novel bacterial taxa that thrive in subsurface soil horizons. *mBio*, *10*, e01318–e01319.

Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., & Fierer, N. (2017). Genome reduction in an abundant and ubiquitous soil bacterium '*Candidatus Udaeobacter copiosus*'. *Nature Microbiology*, *2*, 16198.

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*, 2639–2643.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2013). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 4516–4522.

Cho, J. C., & Tiedje, J. M. (2000). Biogeography and degree of endemicity of fluorescent *pseudomonas* strains in soil. *Applied and Environmental Microbiology*, *66*(12), 5448–5446.

Chu, H., Gao, G.-F., Ma, Y., Fan, K., & Delgado-Baquerizo, M. (2020). Soil microbial biogeography in a changing world: Recent advances and future perspectives. *mSystems*, *5*(2), e00803-19.

Cutler, N. A., Chaput, D. L., Oliver, A. E., & Viles, H. A. (2015). The spatial organization and microbial community structure of an epilithic biofilm. *FEMS Microbiology Ecology*, *91*, fiu027.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*, 927–930.

Du, X., Deng, Y., Li, S., Escalas, A., Feng, K., He, Q., Wang, Z., Wu, Y., Wang, D., Peng, X., & Wang, S. (2021). Steeper spatial scaling patterns of subsoil microbiota are shaped by deterministic assembly process. *Molecular Ecology*, *30*(4), 1072–1085.

Dumbrell, A. J., Nelson, M., Helgason, T., Dytham, C., & Fitter, A. H. (2010). Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME Journal*, *4*, 337–345.

Durrer, A., Gumiere, T., Taketani, R. G., Costa, D. P., Silva, M. C. P., & Andreote, F. D. (2017). The drivers underlying biogeographical patterns of bacterial communities in soils under sugarcane cultivation. *Applied Soil Ecology*, *110*, 12–20.

Eilers, K. G., Debenport, S., Anderson, S., & Fierer, N. (2012). Digging deeper to find unique microbial communities: The strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biology and Biochemistry*, *50*, 58–65.

Fierer, N. (2017). Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, *15*, 579–590.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626–631.

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., Owens, S., Gilbert, J. A., Wall, D. H., & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), 21390–21395.

Fodelianakis, S., Valenzuela-Cuevas, A., Barozzi, A., & Daffonchio, D. (2021). Direct quantification of ecological drift at the population level in synthetic bacterial communities. *The ISME Journal*, *15*, 55–66.

Fukuyama, J. (2019). Emphasis on the deep or shallow parts of the tree provides a new characterization of phylogenetic distance. *Genome Biology*, *20*, 131.

Gee, G. W., & Bauder, J. W. (1979). Particle size analysis by hydrometer: A simplified method for a routine textural analysis and a sensitivity test of measured parameters (Bouyoucos method). *Soil Science Society of America Journal*, *43*, 1004–1007.

Gocke, M. I., Huguet, A., Derenne, S., Kolb, S., Dippold, M. A., & Wiesenberg, G. L. B. (2017). Disentangling interactions between microbial communities and roots in deep subsoil. *Science of the Total Environment*, *575*, 135–145.

Gohl, D. M., Vangay, P., Garbe, J., MacLean, A., Hauge, A., Becker, A., Gould, T. J., Clayton, J. B., Johnson, T. J., Hunter, R., & Knights, D. (2016). Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nature Biotechnology*, *34*(9), 942–949.

Graham, C. H., & Fine, P. V. A. (2008). Phylogenetic beta diversity: Linking ecological and evolutionary processes across space and time. *Ecology Letters*, *11*, 1265–1277.

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of british soils. *Environmental Microbiology*, *13*(6), 1642–1654.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, *10*, 497–506.

Hao, J., Chai, Y. N., Lopes, L. D., Ordonez, R. A., Wright, E. E., Archontoulis, S., & Schachtman, D. P. (2021). The effects of soil depth on the structure of microbial communities in agricultural soils in Iowa (United States). *Applied and Environmental Microbiology*, *87*(4), e02673-20.

Hartemink, A. E., Zhang, Y., Bockheim, J. G., Curi, N., Silva, S. H. G., Grauer-Gray, J., Lowe, D. J., & Krasilnikov, P. (2020). Soil horizon variation: A review. *Advances in Agronomy*, *160*(1), 125–185.

Hood-Nowotny, R., Umana, N. H.-N., Inselbacher, E., Oswald-Lachouani, P., & Wanek, W. (2010). Alternative methods for measuring inorganic, organic, and total dissolved nitrogen in soil. *Soil Science Society of America Journal*, *74*, 1018–1027.

Jia, J., Feng, X., He, J.-S., He, H., Lin, L., & Liu, Z. (2017). Comparing microbial carbon sequestration and priming in the subsoil versus topsoil of a Qinghai-Tibetan alpine grassland. *Soil Biology and Biochemistry*, *104*, 141–151.

Jiao, S., Chen, W., Wang, J., Du, N., Li, Q., & Wei, G. (2018). Soil microbiomes with distinct assemblies through vertical soil profiles drive the cycling of multiple nutrients in reforested ecosystems. *Microbiome*, *6*, 146.

Jiao, S., Yang, Y., Xu, Y., Zhang, J., & Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *The ISME Journal*, *14*, 202–216.

Kampers, L. F. C., Koehorst, J. J., van Heck, R. J. A., Suarez-Diez, M., Stams, A. J. M., & Schaap, P. J. (2021). A metabolic and physiological design study of *pseudomonas putida* KT2440 capable of anaerobic respiration. *BMC Microbiology*, *21*(1), 1–15.

Karimi, B., Terrat, S., Dequiedt, S., Saby, N. P. A., Horrigue, W., Lelievre, M., Nowak, V., Jolivet, C., Arrouays, D., Wincker, P., Cruaud, C., Bispo, A., Maron, P.-A., Prevost-Boure, N. C., & Ranjard, L. (2018). Biogeography of soil bacteria and archaea across France. *Science Advances*, *4*, eaat1808.

Kersters, K., de Vos, P., Gillis, M., Swings, J., Vandamme, P., & Stackebrandt, E. (2006). Introduction to the proteobacteria. In M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, & E. Stackebrandt (Eds.), *The prokaryotes: A handbook on the biology of bacteria* (pp. 3–37). Springer.

Koblizek, M., Dachev, M., Bina, D., Nupur, Piwosz, K., & Kaftan, D. (2020). Utilization of light energy in phototrophic gemmatimonadetes. *Journal of Photochemistry and Photobiology B: Biology*, *213*, 112085.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, *79*, 5112–5120.

Kulichevskaya, I. S., Suzina, N. E., Liesack, W., & Dedysh, S. N. (2010). *Bryobacter aggregatus* gen. Nov., sp. nov., a peat-inhabiting, aerobic chemo-organotroph from subdivision 3 of the Acidobacteria. *International Journal of Systematic and Evolutionary Microbiology*, *60*(2), 301–306.

Lauber, C. L., 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. *Applied and Environmental Microbiology*, *75*(15), 5111–5120.

Leon-Palmero, E., Joglar, V., Alvarez, P. A., Martin-Platero, A., Llamas, I., & Reche, I. (2018). Diversity and antimicrobial potential in sea anemone and holothurian microbiomes. *PLoS One*, *13*(5), e0196178.

Li, B., Shen, C., Wu, H.-Y., Zhang, L.-M., Wang, J., Liu, S., Jing, Z., & Ge, Y. (2022). Environmental selection dominates over dispersal limitation in shaping bacterial biogeographical patterns across different soil horizons of the Qinghai-Tibet plateau. *Science of the Total Environment*, *838*(2), 156177.

Li, S., Li, Y., Hu, C., Zheng, X., Zhang, H., Bai, N., Zhang, H., Tian, M., Ban, S., Wu, G., Pan, A., Li, P., & Lv, W. (2021). Stochastic processes drive bacterial and fungal community assembly in sustainable intensive agricultural soils of Shanghai, China. *Science of the Total Environment*, *778*, 146021.

Liu, J., Sui, Y., Yu, Z., Shi, Y., Chu, H., Jin, J., Liu, X., & Wang, G. (2014). High throughput sequencing analysis of biogrographical distribution of bacterial communities in the black soils of Northeast China. *Soil Biology and Biochemistry*, *70*, 113–122.

Liu, N., Hu, H., Deng, Y., Liu, Y., Hao, B., Zhang, X., Dimitrov, D., Feng, X., & Wang, Z. (2019). Contrasting biogeographic patterns of bacterial and archaeal diversity in the top- and subsoils of temperate grasslands. *mSystems*, *4*(5), e00566-19.

Lopes, L. D., Davis, E. W., II, Silva, M. C. P., Weisberg, A. J., Bresciani, L., Chang, J. H., Loper, J. E., & Andreote, F. D. (2018). Tropical soils are a reservoir for fluorescent *pseudomonas* spp. biodiversity. *Environmental Microbiology*, *20*(1), 62–74.

Lopes, L. D., Hao, J., & Schachtman, D. P. (2021). Alkaline soil pH affects bulk soil, rhizosphere and root endosphere microbiomes of plants growing in a sandhills ecosystem. *FEMS Microbiology Ecology*, *97*, fiab028.

Lozupone, C. A., Hamady, M., Kelley, S. T., & Knight, R. (2007). Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Applied and Environmental Microbiology*, *73*(5), 1576–1585.

Lozupone, C. A., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *The ISME Journal*, *5*, 169–172.

Lynch, J. P. (2021). Harnessing root architecture to address global challenges. *Plant Journal*, *109*(2), 415–431.

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Ovreas, L., Reysenbach, A.-L., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*, *4*, 102–112.

Martiny, J. B. H., Eisen, J. A., Penn, K., Allison, S. D., & Horner-Devine, M. C. (2011). Drivers of bacterial β-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(19), 7850–7854.

Meyer, K. M., Memiaghe, H., Korte, L., Kenfack, D., Alonso, A., & Bohannan, B. J. M. (2018). Why do microbes exhibit weak biogeographic patterns? *The ISME Journal*, *12*, 1404–1413.

Mundra, S., Kjonaas, O. J., Morgado, L. N., Krabberod, A. K., Ransedokken, Y., & Kauserud, H. (2021). Soil depth matters: Shift in composition and inter-kingdom co-occurrence patterns of microorganisms in forest soils. *FEMS Microbiology Ecology*, *97*(3), fiab022.

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews*, *77*(3), 342–356.

Olsen, S. R., & Sommers, L. E. (1982). Phosphorus. In A. L. Page, et al. (Eds.), *Methods of soil analysis: Part 2. Chemical and microbiological properties. Agronomy mongraphs* (Vol. *9*, 2nd ed., pp. 403–430). ASA and SSSA.

Park, D., Kim, H., & Yoon, S. (2017). Nitrous oxide reduction by an obligate aerobic bacterium, *Gemmatimonas aurantiaca*. *Applied and Environmental Microbiology*, *83*(12), e00502-17.

Parks, D. H., Tyson, G. W., Hugenholtz, P., & Beiko, R. G. (2014). STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics*, *30*, 3123–3124.

Paul, E. (2014). *Soil microbiology, ecology and biochemistry*. Academic press.

Pershina, E. V., Ivanova, E. A., Korvigo, I. O., Chirak, E. L., Sergaliev, N. H., Abakumov, E. V., Provorov, N. A., & Andronov, E. E. (2018). Investigation of the core microbiome in main soil types from the east European plain. *Science of the Total Environment*, *631-632*, 1421–1430.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glockner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *D1*, D590–D596.

Riccetto, S., Davis, A. S., Guan, K., & Pittelkow, C. M. (2020). Integrated assessment of crop production and resource use efficiency indicators for the U.S. Corn Belt. *Global Food Security*, *24*, 100339.

Rodrigues, J. L. M., Pellizari, V. H., Mueller, R., Baek, K., Jesus, E. C., Paula, F. S., Mirza, B., Hamaoui, G. S., Tsai, S. M., Feigl, B., Tiedje, J. M., Bohannan, B. J. M., & Nüsslein, K. (2013). Conversion of the amazon rainforest to agriculture results in biotic homogeneization of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3), 988–993.

Seuradge, B. J., Oelbermann, M., & Neufeld, J. D. (2017). Depth-dependent influence of different land-use systems on bacterial biogeography. *FEMS Microbiology Ecology*, *93*, fiw239.

Shoemaker, W. R., Locey, K. J., & Lennon, J. T. (2017). A macroecological theory of microbial biodiversity. *Nature Ecology and Evolution*, *1*(5), 107.

Silby, M. W., Winstanley, C., Godfrey, S. A. C., Levy, S. B., & Jackson, R. W. (2011). *Pseudomonas* genomes: Diverse and adaptable. *FEMS Microbiology Reviews*, *35*, 652–680.

Singh, B. K., Trivedi, P., Egidi, E., Macdonald, C. A., & Delgado-Baquerizo, M. (2020). Crop microbiome and sustainable agriculture. *Nature Reviews Microbiology*, *18*, 601–602.

Siqueira, T., Saito, V. S., Bini, L. M., Melo, A. S., Petsch, D. K., Landeiro, V. L., Tolonen, K. T., Jyrkankallio-Mikkola, J., Soininen, J., & Heino, J. (2020). Community size can affect the signals of ecological drift and niche selection on biodiversity. *Ecology*, *101*(6), e03014.

Smith, D. J., Timonen, H. J., Jaffe, D. A., Griffin, D. W., Birmele, M. N., Perry, K. D., Ward, P. D., & Roberts, M. S. (2013). Intercontinental dispersal of bacteria and archaea by transpacific winds. *Applied and Environmental Microbiology*, *79*(4), 1134–1139.

Stone, M. M., DeForest, J. L., & Plante, A. F. (2014). Changes in extracellular enzyme activity and microbial community structure with soil depth at the Luquillo critical zone observatory. *Soil Biology and Biochemistry*, *75*, 237–247.

Sun, Q., Guo, S., Wang, R., & Song, J. (2021). Responses of bacterial communities and their carbon dynamics to subsoil exposure on the loess plateau. *Science of the Total Environment*, *756*, 144146.

Sun, T., Wang, Y., Hui, D., Jing, X., & Feng, W. (2020). Soil properties rather than climate and ecosystem type control the vertical variations of soil organic carbon, microbial carbon, and microbial quotient. *Soil Biology and Biochemistry*, *148*, 107905.

Tecon, R., & Or, D. (2017). Biophysical processes supporting the diversity of microbial life in soil. *FEMS Microbiology Reviews*, *41*, 599–623.

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Zech Xu, Z., Jiang, L., … Knight, R. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463.

Tribelli, P. M., Lujan, A. M., Pardo, A., Ibarra, J. G., Do Porto, D. F., Smania, A., & Lopez, N. I. (2019). Core regulon of the global anaerobic regulator Anr targets central metabolism functions in *pseudomonas* species. *Scientific Reports*, *9*, 9065.

Wang, H., Liu, S., Li, H., Tao, X., Wang, H., Qi, J., & Zhang, Z. (2022). Large-scale homogenization of soil bacterial communities in response to agricultural practices in paddy fields, China. *Soil Biology and Biochemistry*, *164*, 108490.

Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J. C., He, J., Liu, X., Zhang, L., & Zhang, E. (2013). Phylogenetic beta diversity in bacterial assemblages across ecosystems: Deterministic versus stochastic processes. *The ISME Journal*, *7*, 1310–1321.

Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*, 180–185.

Willms, I. M., Bolz, S. H., Yuan, J., Krafft, L., Schneider, D., Schoning, I., Schrumpf, M., & Nacke, H. (2021). The ubiquitous soil verrucomicrobial clade '*Candidatus Udaeobacter*' shows preferences for acidic pH. *Environmental Microbiology Reports*, *13*(6), 878–883.

Wordell-Dietrich, P., Don, A., & Helfrich, M. (2017). Controlling factors for the stability of subsoil carbon in a dystric Cambisol. *Geoderma*, *304*, 40–48.

Xu, M., Li, X., Kuyper, T. W., Xu, M., Li, X., & Zhang, J. (2021). High microbial diversity stabilizes the responses of soil organic carbon decomposition to warming in the subsoil on the Tibetan plateau. *Global Change Biology*, *27*(10), 2061–2075.

Yan, L., Herrman, M., Kampe, B., Lehmann, R., Totsche, K. W., & Kusel, K. (2020). Environmental selection shapes the formation of near-surface groundwater microbiomes. *Water Research*, *170*, 115341.

Zeng, Y., Nupur, Wu, N., Madsen, A. M., Chen, X., Gardiner, A. T., & Koblizek, M. (2021). Gemmatimonas groenlandica sp. nov. is an aerobic anoxygenic phototroph in the phylum Gemmatimonadetes. *Frontiers in Microbiology*, *11*, 606612.

Zhang, Z., & Peng, X. (2021). Bio-tillage: A new perspective for sustainable agriculture. *Soil and Tillage Research*, *206*, 104844.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lopes, L. D., Futrell, S. L., Wright, E. E., Danalatos, G. J., Castellano, M. J., Vyn, T. J., Archontoulis, S. V., & Schachtman, D. P. (2023). Soil depth and geographic distance modulate bacterial β-diversity in deep soil profiles throughout the U.S. Corn Belt. *Molecular Ecology*, *00*, 1–15. https://doi.org/10.1111/mec.16945