Still, the environment selects: disentangling the Effect of Distance Decay on Soil Bacterial Communities

Niklas Neisse[1]*, Maria Kostakou[1], Sophia N. Meyer[2], Anja Linstädter[2], Antonis Chatzinotas[1,3,4], Stephanie D. Jurburg [1,3]*

[1] Department of Environmental Microbiology, Helmholtz Centre for Environmental Research - UFZ, Leipzig, Germany

[2] Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany

[3] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany

[4] Institute of Biology, Leipzig University, Leipzig, Germany

*Corresponding authors: neisse.n@protonmail.com; s.d.jurburg@gmail.com

Abstract

Soil bacterial communities are central to ecosystem functioning, yet the relative importance of dispersal limitation, environmental selection, and biotic interactions in shaping their spatial turnover remains unresolved. Distance–decay relationships (DDRs)—the decline in community similarity with geographic distance—are commonly observed for microbial communities, but their underlying drivers across ecologically relevant spatial scales remain unclear. We analyzed soil bacterial communities in temperate grasslands across two regions in Germany to quantify the contribution of geographic distance, soil physicochemical properties, plant community composition, and plant traits to bacterial β-diversity. Generalized linear modeling, variation partitioning, and commonality analysis revealed a clear DDR, with bacterial similarity declining by ~5% for each doubling of geographic distance. However, soil physicochemical heterogeneity accounted for over 50% of the explained variation in the bacterial DDR and nearly 30% of the total variation in community composition. Plant community composition independently explained additional variation, while plant functional traits had only marginal effects. Notably, fine-scale environmental heterogeneity within sites contributed to high turnover over short distances, indicating strong abiotic filtering even at the plot scale. Then, to further assess microbial distribution patterns, we examined the relationship between taxon abundance and spatial range. We found that rare taxa were both locally and broadly distributed, suggesting that

rarity does not necessarily constrain dispersal. Dominant taxa, particularly from Proteobacteria and Firmicutes, were consistently broadly distributed, in accordance with generalist lifestyles. In contrast, we found a small, but taxonomically diverse group of highly abundant taxa which were distributed only over intermediate ranges, suggesting that dispersal limitation does not constrain dominance. Our results demonstrate that soil bacterial DDRs emerge primarily from environmental filtering and plant–soil interactions, with a secondary role for spatial separation. These findings highlight the importance of integrating spatially explicit sampling with soil and vegetation data in microbial biogeography, and shed light on the complex patterns of dispersal limitation in microbial communities.

Introduction

Soil bacterial communities are essential to terrestrial ecosystem functioning, nutrient cycling, organic matter decomposition, and plant growth (Loreau et al., 2001), but knowledge of their spatial distribution is limited by their highly diversity, wide distributions, and considerable spatiotemporal turnover (Fierer & Jackson, 2006; Lemoine et al., 2023). Understanding the processes that structure bacterial communities (i.e., assembly) is crucial for predicting ecosystem responses to environmental change. Community assembly is shaped by four key mechanisms: selection, drift, speciation, and dispersal (Vellend, 2010). Among these, selection and dispersal are central to shaping bacterial diversity at local and regional scales (Nemergut et al., 2013), and while they have received considerable attention in microbial ecology, most work to date focuses on large scales (i.e., km), which likely exceed the dispersal capacity of most soil microbes and the local factors that drive their selection (Hanson et al., 2012; but see Richter-Heitmann et al., 2020).

The distance decay relationship (DDR) describes the observation that the similarity of community composition decreases with geographic distance, and has been well-documented for plants and animals (Graco-Roza et al., 2022; Nekola & McGill, 2014; Nekola & White, 1999), as well as microbes (Clark et al., 2021). Distance-decay relationships arise from the interaction between dispersal limitation, environmental gradients, and biotic interactions (Nekola & White, 1999). Dispersal limitation contributes to DDRs when a lack of connectivity between habitat patches prevents organisms from dispersing across patches, while environmental heterogeneity and biotic interactions, both forms of selection, contribute to DDRs because communities in close proximity are likely to encounter more similar environmental conditions and interact with similar organisms, respectively.

Existing research into DDRs in soil bacteria has found mixed evidence for this pattern (Fierer & Jackson, 2006; Rousk et al., 2010), and generally weaker DDRs, with differences in the strength of the relationship attributed to experimental designs (Barbour et al., 2022a), but the extent to which properties of soil microbial communities contribute to the observed DDRs is unclear. First, the extent to which microbes are limited by dispersal remains unresolved. On the one hand, passive dispersal mechanisms, high population densities, and the ability to enter dormancy suggest a high dispersal potential (Clark et al., 2021). On the other hand, accumulating evidence indicates that microbial dispersal can indeed be limited (Barbour et al., 2022b; Hanson et al., 2012), but the spatial extent of this dispersal limitation is unclear (Barbour et al., 2022a). Second, microbial community assembly in soils is influenced by plant–microbe associations, which play a significant role in structuring soil bacterial communities by exerting selective pressures on the local microbiota (Eisenhauer et al., 2010; Trivedi et al., 2020). However, since plant distributions themselves follow DDRs (Nekola & White, 1999), these interactions introduce further complexity into microbial assembly patterns. At the same time, a wide range of work has demonstrated the role of small-scale environmental differences, such as pH (Fierer & Jackson,

2006; Rousk et al., 2010; Zhou et al., 2024) and moisture (Delgado-Baquerizo et al., 2018), in shaping soil bacterial communities. Steep environmental gradients can occur over very short distances (Dumbrell et al., 2009; Vos et al., 2013), even over millimeters.

A more complete understanding of soil microbial community assembly requires disentangling the role of dispersal from the biotic and abiotic environment across spatial gradients that encompass expected bacterial dispersal ranges. Grasslands are ideal ecosystems to study soil bacterial DDRs due to their spatial continuity, microbial diversity, and variable environmental conditions. While stronger DDRs are expected in systems with low connectivity or limited dispersal (i.e., soil; Hanson et al., 2012), grassland soil microbial communities exhibit weaker DDRs than those in aquatic environments, suggesting that additional factors like complex dispersal patterns or uneven environmental gradients may influence microbial community turnover in these systems (Clark et al., 2021). We investigated the drivers of soil bacterial community assembly in grassland soils across two regions of Germany. We hypothesized that: (1) soil bacterial communities in grassland ecosystems exhibit distance-decay relationships from the local (m) to the regional (km) scales; (2) dispersal limitation plays a significant role in soil microbial community assembly; but (3) plant and environmental DDRs contribute to bacterial DDRs, accounting for a considerable portion of the observed variation in microbial community composition.

Methods

Study Area and Sampling Design

We sampled managed grasslands in 18 plots across the Hainich-Dün and Schorfheide-Chorin regions of Germany as part of the Biodiversity Exploratories (Fischer et al., 2010). These regions are separated by approximately 300 km. Schorfheide-Chorin is mainly shaped by sandy and loamy soils which originate from young glacial sediments, whereas Hainich-Dün is characterized by loamy and clayey soils from calcareous bedrocks (Fischer et al., 2010). All grassland plots lie on a land use intensity gradient and are managed by local farmers (Blüthgen et al., 2012).

Soil and vegetation data were collected in June 2023 in nine permanently installed research plots per region. In each plot, three 1 m^2 subplots were sampled along a south-to-north transect of 50 m. The first subplot had a distance of 13.5 m to the second subplot, while the second subplot had a distance of 18 m to the third subplot. Within each subplot, plant species inventories and cover estimations were conducted, and soil samples were taken. Furthermore, the cover of bare ground, litter, vegetation and moss were estimated and always added up to hundred percent. The total vegetation cover was further differentiated into the cover of grasses, herbs and legumes. Plant species occurring with a cover above one percent were sampled in each of the three subplots. The total dry biomass of a plant individual was estimated based on the dry mass of a portion of the sampled individual. The leaf area and dry mass of leaves from the collected individual were

measured to calculate the trait specific leaf area (SLA). Plant species inventory data is available in BExIS (http://doi.org/10.17616/R32P9Q; accession number 31976; Meyer, 2025).

Soil measurements

The top 10 cm of topsoil was collected from the center of each subplot with a 2.5 cm auger and mixed. A 5 g aliquot was stored at -20°C for community analyses. The rest of the samples were sieved (1-2 mm) and subsequently dried (at least 40°C for 24h) to assess soil carbon and nitrogen content using a CN analyser (Eurovector EA3100). Soil moisture was measured as (fresh weight [g] - dry weight [g]) / dry weight [g]. To measure soil pH, 10±0.2 g moist soil and 20 ml deionized water were shaken for a few seconds and an equilibrium period of 30 minutes was allowed before measuring pH with a pH meter.

DNA Extraction and Sequencing

DNA was extracted from 0.25 g of soil for each sample using the NucleoSpin Soil kit (Macherey Nagel), according to manufacturer instructions, and DNA quality and concentration were assessed with gel electrophoresis and Qubit, respectively. The 16S rRNA gene was amplified using primers targeting the V4 region (515F-806R), using the standard Earth Microbiome Project protocols (Thompson et al., 2017) and primers 16S_Illu_515F (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGYCAGCMGCCGCGGTAA-3') and 16S_Illu_806R (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGGACTACNVGGGTWTCTAAT-3'). Amplicons were sequenced with an Illumina MiSeq sequencer (Illumina Inc., San Diego, CA, United States) with the MiSeq Reagent Kit v3 (600 bp).

Bioinformatics

All bioinformatics and statistics analyses were conducted in R (version 4.3.2) (R. Core Team, 2025). R code for all analyses is available in GitHub (https://github.com/NeisseN/BEO_DDR). The *dada2* pipeline was used to infer amplicon sequence variants (ASVs) from the raw sequencing data according to standard protocols (Callahan et al., 2016). Reads were trimmed to exclude the first 10 nucleotides and truncated at 230 and 220 bp for the forward and reverse reads respectively, with a maximum of 3 expected errors per read. Taxonomic assignment was performed with the SILVA 16S rRNA reference database v138.1 (Quast et al., 2012). Raw sequences are publicly available in NCBI's Sequence Read Archives under accession number PRJNA1284051, and processed data and corresponding metadata are available in BExIS (accession numbers 32155 and 332156, respectively; Neisse, 2025). Prior to statistical analyses, all samples were standardized to 23,942 observations per sample and thereby excluding one sample using the *rarefy_even_depth* function of the *phyloseq* package (https://doi.org/10.1371/journal.pone.0061217).

Statistical analyses

To test the strength of the soil microbial DDR and account for the contributions plant and environmental DDRs to this pattern, we performed a commonality analysis based on a generalized linear model (GLM) with a Gamma error distribution and a log link function, using the *glmm.hp* function of the *glmm.hp* package (Lai et al., 2022). We considered four groups of explanatory variables, all measured at the sample level. First, geographic distance was represented by the pairwise Haversine distance on the natural based log scale between sample coordinates, calculated using the *distm* function from the *geosphere* package (version 1.5-20) (Hijmans et al., 2024). Second, abiotic environmental variation was captured as the Euclidean distance based on scaled soil carbon and nitrogen concentrations, moisture content, and pH, consistently identified as key drivers of microbial community composition across diverse ecosystems (Bell, 2010; Delgado-Baquerizo et al., 2018; Fierer & Jackson, 2006). Third, variation in plant community composition was assessed through Bray-Curtis dissimilarities with the *vegdist* function. Finally, plant traits were represented by the Euclidean distance of scaled total plant biomass, specific leaf area (SLA), and the percentage cover of four ground cover types: bare ground, litter, moss, and senescent litter.

To build the generalized linear model, we first addressed potential collinearity among the explanatory distance matrices by calculating pairwise Pearson correlations. We then performed bidirectional stepwise selection using the full model, including all interaction terms, as the upper scope. We identified only a significant interaction between physico-chemical Euclidean distance and plant compositional Bray-Curtis dissimilarity for model performance. Although plant trait similarity did not improve model fit, we retained it in the final model to ensure consistency with the variables used in the variation partitioning analysis. Model assumptions were checked with the *DHARMa* R package (Hartig et al., 2024). We used the commonality analysis with a lognormal approximation on our GLM to disentangle the unique and shared effects of each ecological factor on bacterial community composition. Individual and overlapping contributions were quantified as marginal R^2 values, calculated from commonality coefficients based on explained deviance. To assess statistical significance, we performed permutation tests with 10,000 iterations, where the response variable was randomly permuted while maintaining the structure of explanatory variables. The total model fit was assessed using a pseudo-R^2 calculated as the deviance explained by the final model relative to the null model. Mean values are reported as mean ± SD throughout.

To disentangle the interactions between drivers of soil microbial β-diversity we performed variation partitioning using the *varpart* function from the *vegan* R package on the Bray-Curtis dissimilarities of bacterial communities. We included four potential drivers of bacterial DDRs: 1) geographic location (spatial coordinates of the samples), 2) abiotic parameters, 3) the relative abundances of plants, and 4) plant traits. To select only relevant variables for each compartment, we applied the *ordiR2step* function on a distance based RDA (*dbRDA*) to perform stepwise

model selection with 999 permutations using the *vegan* package (version 2.6-10) (Oksanen et al., 2001). The significance of each fraction of the variation partitioning analysis was assessed with the *test_vp4* function from the *comecol* package in R (https://github.com/jgmv/comecol). This function performs permutation-based significance testing for each explanatory component using the *anova.cca* function from the *vegan* package.

Spatial range

To investigate how the spatial ranges of bacterial community members drive the observed DDRs, we investigated the relationship between the spatial distribution and relative abundance of ASVs by calculating the convex hull area and mean relative abundance for each ASV across all samples. The convex hull area (i.e., the smallest convex shape that encloses a set of observations), representing the geographic extent of each ASV, was calculated using the *CHullArea* function from the *GeoRange* R packages (Boyle, 2017), based on the longitude and latitude coordinates (WGS84; UTM Zone 33N, Berlin) of the samples in which the ASV was present. ASVs with less than 3 locations were assigned a convex hull area of 0, and all values were converted to the natural log scale (+1 m^2). ASVs were classified according to their relationship of occurrence (within plot, within site, and between site), and by phylum.

Results

Distance Decay in Bacterial Communities

We decomposed the DDR of soil bacterial communities to assess the contribution of geographic distance, the DDR in soil physicochemical properties, the DDR of plant communities, and the resulting gradient in plant community traits. The fitted GLM captured 53% of the variation in microbial community dissimilarities (Fadden's pseudo-R^2 = 0.53). As expected, geographic distance had a significant, negative effect on bacterial community similarity (β = -0.074, SE = 0.005, t(1269) = -14.03, $p < 0.001$; Fig. 1). Doubling the distance between sites was associated with a 5% decrease in bacterial community similarity, holding other variables constant (Supplemental Equation S1-S4).

The highest observed bacterial community similarity was 59.8%, occurring between two samples collected within the same plot, though not among the closest recorded distances. Nevertheless, despite originating from the same plot, some sample pairs exhibited very low similarity, with the lowest recorded value being 11.7%. The average similarity among sample pairs from the same plot was 40.4 ± 13.4%. At the minimum geographic distance of 12.96 m, a bacterial community similarity of 35.2 ± 3% was estimated (holding all other variables at their mean values). At average short range (20.13 m), mid range (5.7 km), and long range (309.4 km) distances, community similarities of 34.2 ± 2.8%, 22.6 ± 0.7%, and 16.9 ± 0.5% were observed. Notably, within-site comparisons included the 22 lowest similarity pairs (minimum similarity of 3%), and

exhibited high variation in physicochemical properties (Supplemental Fig. S1), which had a significant negative effect on bacterial community similarity (β = -0.279, SE = 0.011, t(1269) = -26.013, $p < 0.001$).

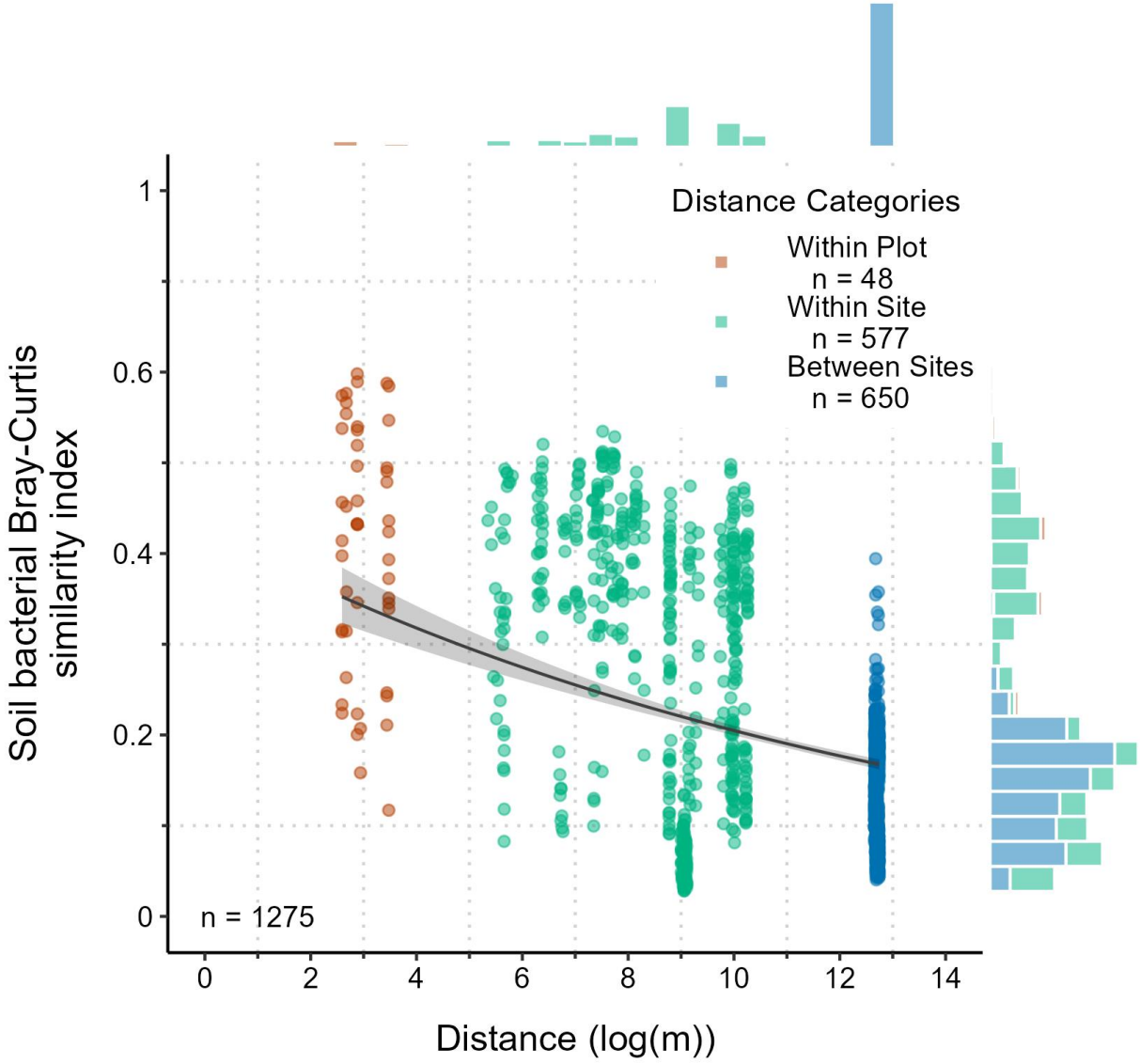

Fig. 1 Distance–decay relationship of soil bacterial communities across grasslands. Distance between samples (log-transformed Haversine distance in meters), modeled using a Gamma generalized linear model (GLM) with a log link and the difference in physicochemical properties, plant communities, plant community traits set to their mean ($p < 0.001$). Points represent raw pairwise similarity data; the line and shaded ribbon show the GLM-predicted relationship with 95% confidence intervals. Stacked bar charts along the X and Y axes show the distribution of pairwise distances, colored to match the comparison scale.

To evaluate the relative importance of DDRs of soil physicochemical parameters, plants, and their traits, we performed a commonality analysis on the GLM (total marginal adjusted R^2= 57.7%; Table 1). Geographic distance, and changes in physicochemical soil properties and plant communities, as well as their interaction were all highly significant ($p < 0.001$), while plant traits showed no significant contribution. Changes in physicochemical parameters accounted for the largest share of the explained variance, contributing 52.1% to the total R^2. This included the highest unique contribution (15.4%), along with a shared contribution of 13.5%, primarily with geographic distance and in three-way combinations involving either plant similarity or the interaction between physicochemical properties and geographic distance. Geographic distance was the second most influential predictor, explaining 38.7% of the total R^2, with a unique contribution of 9.6%. Changes in plant communities accounted for 6.4% of the total R^2, largely through shared variance with other predictors.

Table 1 Contribution shares of selective processes to bacterial community similarity in grasslands, based on commonality analysis using a lognormal approximation. The table presents the unique, average shared, and total (individual) contributions of each predictor to the model's adjusted R^2 (total $R^2 = 0.577$), along with their percentage contributions. The interaction term (E) reflects the conditional effect of physicochemical properties and geographic distance. Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$, ns = not significant (based on 1000 permutations).

		a. Unique		b. Average share		Individual (a+b)		Individual (%)
A.	Change in soil properties	0.154	***	0.135	***	0.289	***	52.1
B.	Geographic distance	0.096	***	0.119	***	0.215	***	38.7
C.	Changes in plant similarity	0.007	*	0.028	***	0.035	***	6.4
D.	Changes in plant traits	-0.001	ns	0.001	***	0.000	ns	0.1
E.	Interaction of A. and B.	0.012	***	0.004	***	0.016	***	2.8

Drivers of Bacterial Beta Diversity

To assess the direct role of a) location, b) physicochemical parameters, c) plants, and d) their traits to in modulating the soil bacterial community (i.e., rather than the effect of their DDRs), we first selected relevant drivers in for compartments b-d through stepwise selection. The variables selected for b-d were b) soil carbon concentration, pH, and the carbon-to-nitrogen (C:N) ratio; c) *Poa trivialis*, *Achillea millefolium*, *Lolium perenne*, *Anthoxanthum odoratum*, *Bromus hordeaceus*, *Medicago sativa*, *Geranium rotundifolium*, *Daucus carota*, *Crataegus*

spec., *Cynosurus cristatus*, *Phleum pratense*, and *Centaurea jacea*; and d) bare ground cover, total plant biomass, and ground cover of senescent litter.

Subsequently, we conducted a variance partitioning analysis, which explained 42% of the total variation in bacterial β-diversity (Fig. 2). All predictor groups contributed significantly to the explained variation (Supplemental Table S1). Plant community composition accounted for the largest proportion of the variance (~36%), followed by soil physicochemical properties (~30%), and location (~19%). Plant traits were the weakest drivers, explaining only about 10% of the variation. Among the unique fractions, plant community composition (7.4%) and location (1.8%) showed statistically significant effects. The unique contributions of soil physicochemical properties and plant traits were not significant (0.9%, and 0.8%, respectively). Notably, plant community composition, excluding the influence of soil physicochemical factors, still accounted for approximately 11% of the variance. Plant community composition and soil properties together accounted for about 41% of the total variance.

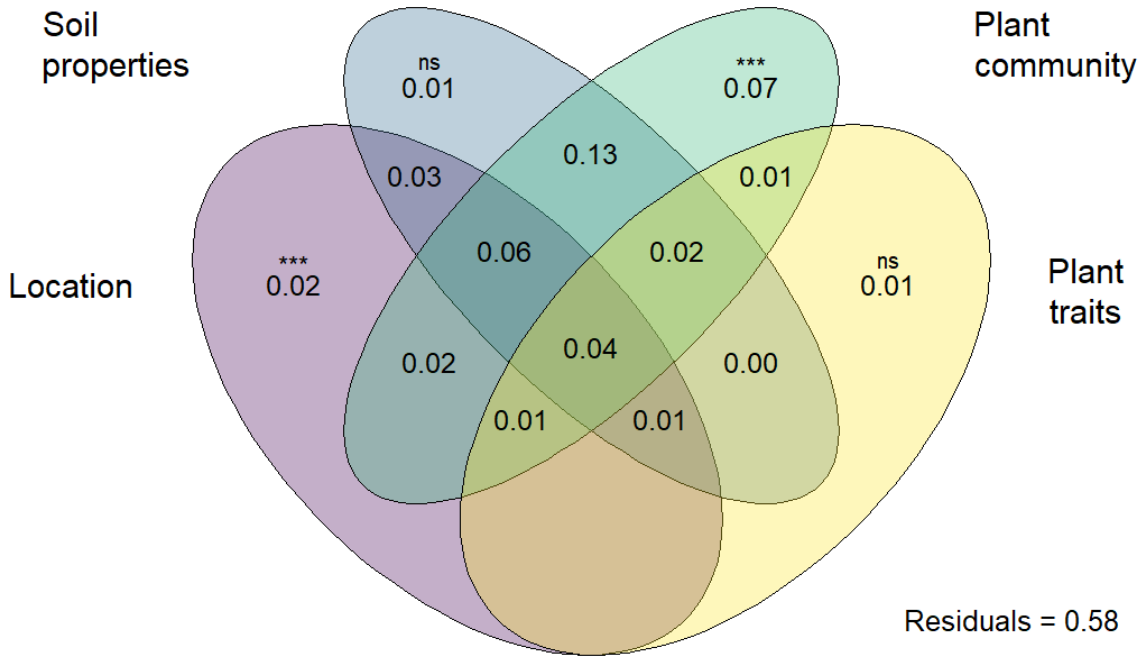

Fig. 2 Venn diagram summarizing the results of a variance partitioning analysing the effect of 1. Location, 2. Soil properties, 3. Plant community composition, 4. Plant traits on the Bray–Curtis similarity of bacterial communities. The values displayed represent the adjusted R^2 of both unique and overlapping contributions of each predictor group. Blank areas indicate no explained or shared variance. Significance codes: *** $p < 0.001$, ns = not significant, blank = not testable.

Spatial Range

We investigated the spatial distribution of bacterial community members by calculating the convex hull area as a proxy for range and determining the average relative abundance of each ASV across all samples (Fig. 3). Based on their spatial extent, we simultaneously classified ASVs into three categories: local, with a convex hull area of less than approximately 10 m^2; intermediate, ranging between roughly 150 m^2 and 170,000,000 m^2; and broad, exceeding around 5,000,000,000 m^2.

We observed a strong positive relationship between ASV relative abundance and range size (Spearman's $\rho = 0.73$, $p < 2.2 \times 10^{-16}$). Most ASVs (99%) exhibited average relative abundances below 0.1% (mean: 0.008%). The dominant bacteria (> 0.1% average relative abundance) primarily belonged to Proteobacteria (n = 57) and Actinobacteriota (n = 30). Acidobacteriota, Bacteroidota, Verrucomicrobiota, Firmicutes, Myxococcota, and Crenarchaeota included ≤ 10 dominant ASVs. Of these, 21 ASVs belonging to Actinobacteriota (8 ASVs), Proteobacteria (6 ASVs), Crenarchaeota (3 ASVs), Verrucomicrobiota (2 ASVs), Acidobacteriota and Myxococcota (1 ASV each) were restricted to intermediate spatial ranges (Table S1). All dominant Firmicutes belonged to Bacillales, including one unclassified *Planococcaceae*, and were detected at broad ranges. Notably, no dominant Firmicutes or Bacteroidota were detected at intermediate scales.

Highly dominant ASVs (>0.25% average relative abundance) were distributed among Proteobacteria (15 ASVs), Firmicutes (4 ASVs), Actinobacteriota (3 ASVs), and Crenarchaeota (3 ASVs). Only three ASVs belonging to Proteobacteria (unclassified *Sutterellaceae*), Actinobacteriota (unclassified *MB-A2-108*), and Crenarchaeota (unclassified *Nitrososphaeraceae*), occurred at intermediate scales. The rest occurred at the broadest spatial scale, including ASVs from Proteobacteria (14 ASVs), Firmicutes (3 ASVs), two ASVs each from Acidobacteriota and Actinobacteriota, and one each from Bacteroidota, Crenarchaeota, Myxococcota. The two most abundant taxa were a *Bradyrhizobium* and an unclassified *Bacillales*, each with a ~1% average relative abundance and spanning the broad spatial scale. No dominant ASVs occurred at local ranges (Fig. 3).

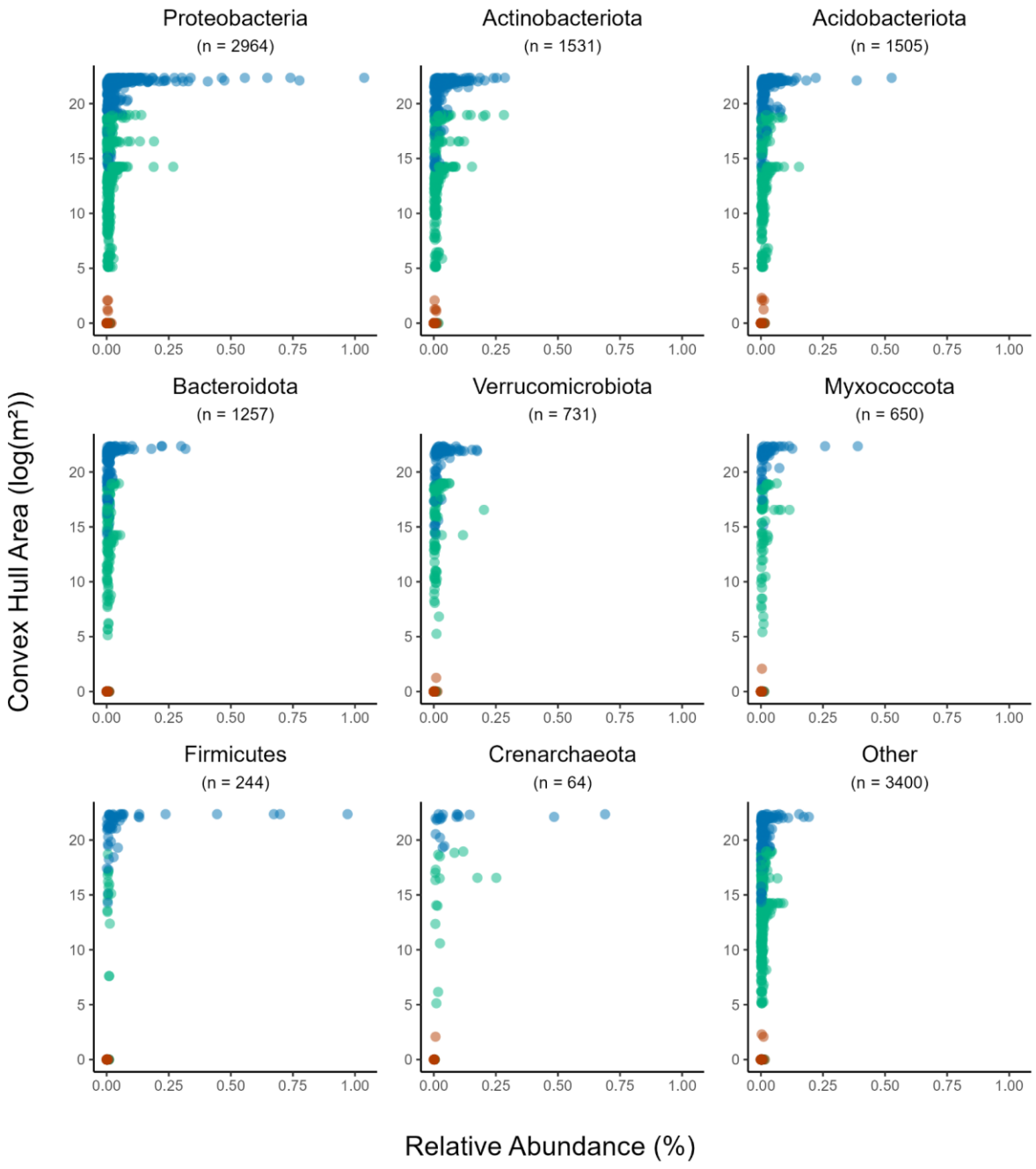

Fig. 3 Relationship between average relative abundance (%) and spatial range (convex hull area in ln(m²)) of soil bacterial ASVs in grasslands, faceted by phylum. Each point represents a single ASV. For faceting, the eight most abundant bacterial phyla, and the rest of the taxa are shown separately. Colors indicate the convex hull area relative to the sampling design: red is within plot; green is within sites, and blue is across sites.

Discussion

Understanding the factors governing the distribution of soil bacteria is essential to advance microbial biogeography and inform strategies for sustainable management of belowground diversity. While previous research has established that bacterial communities often exhibit distance–decay relationships (Martiny et al., 2011), the extent to which these patterns reflect spatial constraints, environmental filtering, biotic interactions or their combinations remains insufficiently resolved—particularly across multiple ecological scales (Barbour et al., 2022a). By integrating variation partitioning with commonality analysis, we disentangled the direct effects of geographic location and abiotic and biotic drivers on the bacterial community from the influence of changes in these abiotic and biotic compartments, which are though autocorrelation also spatially structured, on the bacterial DDR. Our findings shed light on the soil bacterial DDR, which emerges from the interaction between dispersal limitation and abiotic and biotic selective factors, and their own DDRs.

We found a significant, negative effect of geographic distance on bacterial community similarity, with a ~5% decrease in similarity with each doubling of distance while controlling for soil and plant-related variables. This aligns with existing research into microbial DDRs (Clark et al., 2021; Martiny et al., 2011), and underscores the role of spatial separation in microbial community assembly, likely through dispersal and historical contingencies. Indeed, geographic distance was the second most important driver of bacterial communities according to our commonality analysis. We found a small but significant individual effect of geographic location in our variation partitioning analysis, suggesting that while spatial processes do affect community composition, much of their influence is confounded with biotic and abiotic gradients driving gradual changes in bacterial communities. This aligns with findings from a synthesis by Hanson et al., (2012), which showed that environmental variables explained more variation in microbial composition (26.9%) than geographic distance (10.3%), with studies being more likely to detect significant effects from environmental factors than from spatial ones. Importantly, the lack of a role of geographic location may have partially resulted from our sampling design, which had a minimum distance between samples of 12.96 m, and while it boasted high spatial resolution relative to the existing literature (Griffiths et al., 2011), may have still been too large to capture dispersal ranges (Barbour et al., 2022a).

Soil physicochemical properties were the most important direct drivers of soil bacterial communities, explaining nearly 30% of the total variance in bacterial community composition in our variation partitioning analysis. Changes in soil properties accounted for over half of the variation in the bacterial DDR and contributed the largest unique share, underscoring their independent influence on microbial communities and highlighting the role of environmental filtering, consistent with niche-based theories of microbial biogeography (Chase & Myers, 2011; Fierer & Jackson, 2006). However, when considered independently of the contribution of soil properties, such as pH and inorganic carbon, had limited direct and marginally non-significant

effects on bacterial community structure (1%; $p = 0.08$). Nevertheless, their influence increased significantly when interactions with plant-driven biotic processes were considered, and their exclusive joint-share accounted for 13% of the total variation. This suggests that abiotic factors act in concert with biotic interactions to structure bacterial communities (Drenovsky et al., 2009) and points to synergistic interactions between abiotic and biotic filters. For instance, soil pH has long been recognized as a key determinant of bacterial diversity, but its effects may be amplified or mitigated by plant species composition (Fierer & Jackson, 2006; Rousk et al., 2010). Taken together, these findings support the notion that environmental conditions are important, but represent just one component of the broader set of selective factors influencing bacterial community assembly, including active biotic interactions (Vellend, 2010).

Bacterial communities varied considerably, even within the same plot over distances of several meters in response to changes in physicochemical parameters, highlighting the role of fine-scale environmental heterogeneity in driving soil microbial communities. The low similarity found between nearby samples, even within the relatively homogeneous grassland environment, further underscores the importance of revisiting standard sampling designs for soil microbiota, which often rely on composite sampling strategies, or pooling, to characterize microbial diversity at the plot scale, contributing further noise to the scales at which selection and microbial dispersal occur in soil ecosystems (Clark et al., 2021).

We found a strong contribution of plant community composition, but not plant traits, to the composition or change in microbial communities. The strong, unique contribution of plant composition to the variance in the soil bacterial communities likely reflects both direct interactions between plants and soil (e.g., through root exudates, litter quality, and rhizosphere dynamics; Philippot et al., 2013; Trivedi et al., 2020) while the shared contribution of plant composition and abiotic parameters likely highlights the role of plants in modulating the microbial environment (e.g., soil pH, nutrient availability, and organic matter content). The lack of influence of plant traits further suggests that trait-based filtering of soil bacteria may be weak or indirect, potentially mediated through plant effects on soil chemistry that were already captured by other measured variables.

Our dual analytical strategy allowed us to quantify the contribution of abiotic and biotic DDRs to soil bacterial DDRs. We found that the interaction between geographic location and physicochemical processes explained a small, but significant portion of the variance in bacterial DDRs, suggesting that spatial changes in soil properties may modulate the effects of distance in a multiplicative manner, adding further nuance to the role of the abiotic environment in structuring soil bacterial communities, both locally, and over spatial gradients (Legendre et al., 2005). Similarly, the influence of the plant DDR was primarily observed through shared effects with changes in geographic distance and soil physicochemical properties and not through unique contributions. This suggests that shifts in plant composition across sites do not consistently translate into corresponding microbial changes at the spatial scale examined, likely because their

effects are context-dependent and mediated through co-varying environmental conditions. Previous studies have shown that spatial heterogeneity in soil properties and stochastic dispersal processes can outweigh deterministic selection by plant communities in structuring microbial communities (K. Regan et al., 2017; K. M. Regan et al., 2014; Richter-Heitmann et al., 2020). In our study, abiotic filtering and dispersal limitation likely dominate microbial community assembly, consistent with findings from previous large-scale grassland surveys (Delgado-Baquerizo et al., 2018).

To further explore potential dispersal limitation and its relationship to dominance in soil microbial communities, we examined the relationship between bacterial ASV abundance and geographic range. We observed distinct biogeographic patterns shaped by taxonomic identity, with most taxa exhibiting low abundances, supporting the notion of a rare biosphere; however, we found similar long-tailed distributions of microbes with broad and local ranges, suggesting the existence of a large portion of rare generalists–low-abundance taxa that can disperse broadly or persist in diverse habitats. While the existence of rare microbes is well established (Lynch & Neufeld, 2015; Sogin et al., 2006), largely through dormancy (Lennon & Jones, 2011), less is known about the distribution of these taxa, and it is generally assumed that in general, rare taxa are locally distributed (Lindh et al., 2017; K. M. Meyer et al., 2018; Shade & Stopnisek, 2019; Thompson et al., 2017). Furthermore, Clark et al. (2021) emphasized that similarity analyses failing to capture rare taxa can overestimate compositional similarity across communities and thus weaken distance–decay relationships, under the assumption that rare taxa are also spatially restricted. Our observation that many rare ASVs are widespread intermediate and broad ranges (an area exceeding 5,000 km^2) challenges this assumption and highlights that rare taxa likely exhibit a range of dispersal ranges. The presence of both narrowly and broadly distributed rare taxa across all phyla suggests diverse dispersal strategies, potentially shaped by microbial life-history traits that facilitate survival and transport—such as dormancy in vegetative cells, cysts, or spores (Hanson et al., 2012; Locey et al., 2020). However such strategies, and their influence on community assembly, remain poorly characterized (Barbour et al., 2022a) and require further study.

In contrast, the accumulation of dominant ASVs (especially Proteobacteria) at the largest spatial scale reinforces the presence of successful generalists in the community, consistent with established ecological classifications (Fierer et al., 2007). These dominant phyla, including Proteobacteria, Actinobacteriota, and Bacteroidota, exhibited wide geographic distributions and often increased in abundance at broader spatial scales, suggesting high dispersal capacity and competitive success across heterogeneous soil environments (Shade & Gilbert, 2015). The consistent dominance of Proteobacteria across scales likely reflects their metabolic versatility, rapid growth, and ability to exploit diverse root-derived carbon sources (Fierer et al., 2007; Philippot et al., 2013; Spain et al., 2009).

Interestingly, we found a small and diverse group of taxa that were dominant at intermediate scales, likely indicating that dominant taxa can also be dispersal limited. These taxa were found across all phyla except for Firmicutes and Bacteroidota, further suggesting that these patterns emerge largely from dispersal limitation. The high abundance found among these taxa may reflect strong niche specialization or habitat patchiness, allowing certain taxa to thrive locally or regionally despite limited distribution. In contrast, the lack of abundant taxa at intermediate scales among Firmicutes and Bacteroidota likely reflects a greater overall dispersal capacity, e.g. due to their ability to form stress-resistant spores or cyst-like structures (Mandic-Mulec et al., 2015) that facilitate long-range dispersal and persistence under fluctuating conditions.

Conclusion

This study paints a more nuanced picture of ecological strategies among soil bacteria, their distribution within phyla, and the effect of the biotic and abiotic environment on the biogeographical distribution of these bacteria. We demonstrate that bacterial communities follow a distance decay relationship, but the strongest drivers of community assembly are soil physicochemical parameters and interactions with plants (i.e., selective factors). We find that dispersal plays an important role in modulating the community, and find large portions of the bacterial community that are dispersal limited; however relative to the contribution of selective factors, the role of dispersal limitation in structuring soil bacterial communities is small. Taken together, our findings suggest a modification to the hypothesis that “everything is everywhere, but the environment selects” (Baas Becking & Nicolai, 1934): “everything is *not* everywhere. Still, the environment selects”.

References

Baas Becking, L. G. M., & Nicolai, E. (1934). On the ecology of a Sphagnum bog. *Blumea: Biodiversity, Evolution and Biogeography of Plants*, *1*(1), 10–45.

Barbour, K. M., Barrón-Sandoval, A., Walters, K. E., & Martiny, J. B. H. (2022a). Towards quantifying microbial dispersal in the environment. *Environmental Microbiology*, *25*(1), 137–142. https://doi.org/10.1111/1462-2920.16270

Barbour, K. M., Barrón-Sandoval, A., Walters, K. E., & Martiny, J. B. H. (2022b). Towards quantifying microbial dispersal in the environment. *Environmental Microbiology*, *25*(1), 137–142. https://doi.org/10.1111/1462-2920.16270

Bell, T. (2010). Experimental tests of the bacterial distance–decay relationship. *The ISME Journal*, *4*(11), 1357–1365. https://doi.org/10.1038/ismej.2010.77

Blüthgen, N., Dormann, C. F., Prati, D., Klaus, V. H., Kleinebecker, T., Hölzel, N., Alt, F., Boch, S., Gockel, S., Hemp, A., Müller, J., Nieschulze, J., Renner, S. C., Schöning, I., Schumacher, U., Socher, S. A., Wells, K., Birkhofer, K., Buscot, F., … Weisser, W. W. (2012). A quantitative index of land-use intensity in grasslands: Integrating mowing, grazing and fertilization. *Basic and Applied Ecology*, *13*(3), 207–220. https://doi.org/10.1016/j.baae.2012.04.001

Boyle, J. (2017). *Package 'GeoRange': Calculating Geographic Range from Occurrence Data* (No. 0.1.0). https://cran.r-project.org/web/packages/GeoRange/GeoRange.pdf

Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J., & Holmes, S. P. (2016). Bioconductor Workflow for Microbiome Data Analysis: From raw reads to community analyses. *F1000Research*, *5*, 1492. https://doi.org/10.12688/f1000research.8986.2

Chase, J. M., & Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1576), 2351–2363. https://doi.org/10.1098/rstb.2011.0063

Clark, D. R., Underwood, G. J. C., McGenity, T. J., & Dumbrell, A. J. (2021). What drives study-dependent differences in distance–decay relationships of microbial communities? *Global Ecology and Biogeography*, *30*(4), 811–825. https://doi.org/10.1111/geb.13266

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-González, A., Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., & Fierer, N. (2018). A global atlas of the dominant bacteria found in soil. *Science*, *359*(6373), 320–325. https://doi.org/10.1126/science.aap9516

Drenovsky, R. E., Steenwerth, K. L., Jackson, L. E., & Scow, K. M. (2009). Land use and climatic factors structure regional patterns in soil microbial communities. *Global Ecology and Biogeography*, *19*(1), 27–39. https://doi.org/10.1111/j.1466-8238.2009.00486.x

Dumbrell, A. J., Nelson, M., Helgason, T., Dytham, C., & Fitter, A. H. (2009). Relative roles of niche and neutral processes in structuring a soil microbial community. *The ISME Journal*, *4*(3), 337–345. https://doi.org/10.1038/ismej.2009.122

Eisenhauer, N., Beßler, H., Engels, C., Gleixner, G., Habekost, M., Milcu, A., Partsch, S., Sabais, A. C. W., Scherber, C., Steinbeiss, S., Weigelt, A., Weisser, W. W., & Scheu, S. (2010). Plant diversity effects on soil microorganisms support the singular hypothesis. *Ecology*, *91*(2), 485–496. https://doi.org/10.1890/08-2338.1

Fierer, N., Bradford, M. A., & Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology*, *88*(6), 1354–1364. https://doi.org/10.1890/05-1839

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences*, *103*(3), 626–631. https://doi.org/10.1073/pnas.0507535103

Fischer, M., Bossdorf, O., Gockel, S., Hänsel, F., Hemp, A., Hessenmöller, D., Korte, G., Nieschulze, J., Pfeiffer, S., Prati, D., Renner, S., Schöning, I., Schumacher, U., Wells, K., Buscot, F., Kalko, E. K. V., Linsenmair, K. E., Schulze, E.-D., & Weisser, W. W. (2010). Implementing large-scale and long-term functional biodiversity research: The

Biodiversity Exploratories. *Basic and Applied Ecology*, *11*(6), 473–485. https://doi.org/10.1016/j.baae.2010.07.009

Graco-Roza, C., Aarnio, S., Abrego, N., Acosta, A. T. R., Alahuhta, J., Altman, J., Angiolini, C., Aroviita, J., Attorre, F., Baastrup-Spohr, L., Barrera-Alba, J. J., Belmaker, J., Biurrun, I., Bonari, G., Bruelheide, H., Burrascano, S., Carboni, M., Cardoso, P., Carvalho, J. C., … Soininen, J. (2022). Distance decay 2.0 – A global synthesis of taxonomic and functional turnover in ecological communities. *Global Ecology and Biogeography*, *31*(7), 1399–1421. https://doi.org/10.1111/geb.13513

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. https://doi.org/10.1111/j.1462-2920.2011.02480.x

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*(7), 497–506. https://doi.org/10.1038/nrmicro2795

Hartig, F., Lohse, L., & de Souza leite. (2024). *Package 'DHARMa': Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models* (No. 0.4.7). https://github.com/florianhartig/DHARMa

Hijmans, R. J., Karney, C., Williams, E., & Vennes, C. (2024). *Package 'geosphere': Spherical Trigonometry* (No. 1.5-20). https://github.com/rspatial/geosphere

Lai, J., Zou, Y., Zhang, S., Zhang, X., & Mao, L. (2022). glmm.hp: An R package for computing individual effect of predictors in generalized linear mixed models. *Journal of Plant Ecology*, *15*(6), 1302–1307. https://doi.org/10.1093/jpe/rtac096

Legendre, P., Borcard, D., & Peres-Neto, P. R. (2005). Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecological Monographs*, *75*(4), 435–450. https://doi.org/10.1890/05-0549

Lemoine, N. P., Adams, B. J., Diaz, M., Dragone, N. B., Franco, A. L. C., Fierer, N., Lyons, W.

B., Hogg, I. D., & Wall, D. H. (2023). Strong Dispersal Limitation of Microbial Communities at Shackleton Glacier, Antarctica. *mSystems*, *8*(1). https://doi.org/10.1128/msystems.01254-22

Lennon, J. T., & Jones, S. E. (2011). Microbial seed banks: The ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, *9*(2), 119–130. https://doi.org/10.1038/nrmicro2504

Lindh, M. V., Sjöstedt, J., Ekstam, B., Casini, M., Lundin, D., Hugerth, L. W., Hu, Y. O. O., Andersson, A. F., Andersson, A., Legrand, C., & Pinhassi, J. (2017). Metapopulation theory identifies biogeographical patterns among core and satellite marine bacteria scaling from tens to thousands of kilometers. *Environmental Microbiology*, *19*(3), 1222–1236. https://doi.org/10.1111/1462-2920.13650

Locey, K. J., Muscarella, M. E., Larsen, M. L., Bray, S. R., Jones, S. E., & Lennon, J. T. (2020). Dormancy dampens the microbial distance–decay relationship. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1798), 20190243. https://doi.org/10.1098/rstb.2019.0243

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J. P., Hector, A., Hooper, D. U., Huston, M. A., Raffaelli, D., Schmid, B., Tilman, D., & Wardle, D. A. (2001). Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges. *Science*, *294*(5543), 804–808. https://doi.org/10.1126/science.1064088

Lynch, M. D. J., & Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology*, *13*(4), 217–229. https://doi.org/10.1038/nrmicro3400

Mandic-Mulec, I., Stefanic, P., & van Elsas, J. D. (2015). Ecology of Bacillaceae. *Microbiology Spectrum*, *3*(2). https://doi.org/10.1128/microbiolspec.tbs-0017-2013

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*, *108*(19), 7850–7854. https://doi.org/10.1073/pnas.1016308108

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*(6), 1404–1413. https://doi.org/10.1038/s41396-018-0103-3

Meyer, S. (2025). *Plant species inventories in experimental plots (not all) of grasslands for trait measurements.* (No. 31976; Version 5) [Dataset]. Biodiversity Exploratories Information System. https://www.bexis.uni-jena.de

Nekola, J. C., & McGill, B. J. (2014). Scale dependency in the functional form of the distance decay relationship. *Ecography*, *37*(4), 309–320. https://doi.org/10.1111/j.1600-0587.2013.00407.x

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*(4), 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

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. https://doi.org/10.1128/mmbr.00051-12

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., … Borman, T. (2001). vegan: Community Ecology Package. In *CRAN: Contributed Packages* (No. 2.7-1). The R Foundation. https://doi.org/10.32614/cran.package.vegan

Philippot, L., Raaijmakers, J. M., Lemanceau, P., & van der Putten, W. H. (2013). Going back to the roots: The microbial ecology of the rhizosphere. *Nature Reviews Microbiology*, *11*(11), 789–799. https://doi.org/10.1038/nrmicro3109

Poghosyan, L., & Lehtovirta-Morley, L. E. (2024). Investigating microbial and environmental drivers of nitrification in alkaline forest soil. *ISME Communications*, *4*(1), ycae093.

https://doi.org/10.1093/ismeco/ycae093

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

R. Core Team. (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Regan, K. M., Nunan, N., Boeddinghaus, R. S., Baumgartner, V., Berner, D., Boch, S., Oelmann, Y., Overmann, J., Prati, D., Schloter, M., Schmitt, B., Sorkau, E., Steffens, M., Kandeler, E., & Marhan, S. (2014). Seasonal controls on grassland microbial biogeography: Are they governed by plants, abiotic properties or both? *Soil Biology and Biochemistry*, *71*, 21–30. https://doi.org/10.1016/j.soilbio.2013.12.024

Regan, K., Stempfhuber, B., Schloter, M., Rasche, F., Prati, D., Philippot, L., Boeddinghaus, R. S., Kandeler, E., & Marhan, S. (2017). Spatial and temporal dynamics of nitrogen fixing, nitrifying and denitrifying microbes in an unfertilized grassland soil. *Soil Biology and Biochemistry*, *109*, 214–226. https://doi.org/10.1016/j.soilbio.2016.11.011

Reji, L., Cardarelli, E. L., Boye, K., Bargar, J. R., & Francis, C. A. (2022). Diverse ecophysiological adaptations of subsurface Thaumarchaeota in floodplain sediments revealed through genome-resolved metagenomics. *The ISME Journal*, *16*(4), 1140–1152. https://doi.org/10.1038/s41396-021-01167-7

Richter-Heitmann, T., Hofner, B., Krah, F.-S., Sikorski, J., Wüst, P. K., Bunk, B., Huang, S., Regan, K. M., Berner, D., Boeddinghaus, R. S., Marhan, S., Prati, D., Kandeler, E., Overmann, J., & Friedrich, M. W. (2020). Stochastic Dispersal Rather Than Deterministic Selection Explains the Spatio-Temporal Distribution of Soil Bacteria in a Temperate Grassland. *Frontiers in Microbiology*, *11*. https://doi.org/10.3389/fmicb.2020.01391

Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R., &

Fierer, N. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME Journal*, *4*(10), 1340–1351. https://doi.org/10.1038/ismej.2010.58

Shade, A., & Gilbert, J. A. (2015). Temporal patterns of rarity provide a more complete view of microbial diversity. *Trends in Microbiology*, *23*(6), 335–340. https://doi.org/10.1016/j.tim.2015.01.007

Shade, A., & Stopnisek, N. (2019). Abundance-occupancy distributions to prioritize plant core microbiome membership. *Current Opinion in Microbiology*, *49*, 50–58. https://doi.org/10.1016/j.mib.2019.09.008

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of the National Academy of Sciences*, *103*(32), 12115–12120. https://doi.org/10.1073/pnas.0605127103

Spain, A. M., Krumholz, L. R., & Elshahed, M. S. (2009). Abundance, composition, diversity and novelty of soil Proteobacteria. *The ISME Journal*, *3*(8), 992–1000. https://doi.org/10.1038/ismej.2009.43

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., … Zhao, H. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463. https://doi.org/10.1038/nature24621

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., & Singh, B. K. (2020). Plant–microbiome interactions: From community assembly to plant health. *Nature Reviews Microbiology*, *18*(11), 607–621. https://doi.org/10.1038/s41579-020-0412-1

Vellend, M. (2010). Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, *85*(2), 183–206. https://doi.org/10.1086/652373

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of

bacterial diversity in soil. *FEMS Microbiology Reviews*, *37*(6), 936–954. https://doi.org/10.1111/1574-6976.12023

Zhang, B., Wu, X., Tai, X., Sun, L., Wu, M., Zhang, W., Chen, X., Zhang, G., Chen, T., Liu, G., & Dyson, P. (2019). Variation in Actinobacterial Community Composition and Potential Function in Different Soil Ecosystems Belonging to the Arid Heihe River Basin of Northwest China. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.02209

Zhou, X., Tahvanainen, T., Malard, L., Chen, L., Pérez-Pérez, J., & Berninger, F. (2024). Global analysis of soil bacterial genera and diversity in response to pH. *Soil Biology and Biochemistry*, *198*, 109552. https://doi.org/10.1016/j.soilbio.2024.109552

Supplementary materials

Formula to calculate the effect of doubling x in a generalized linear model with a log link.

$$\Delta = \beta * ln(2x) - \beta * ln(x) = \beta * [ln(x) + ln(2) - ln(x)] = \beta * ln(2) \quad (S1)$$

$$\beta * ln(2) = -0.736 * 0.6931 = -0.051 \quad (S2)$$

To back transform the log transformed geographic predictor, we use the exponential.

$$exp(-0.051) = 0.95 \quad (S3)$$

$$1 - 0.95 = 0.05 \quad (S4)$$

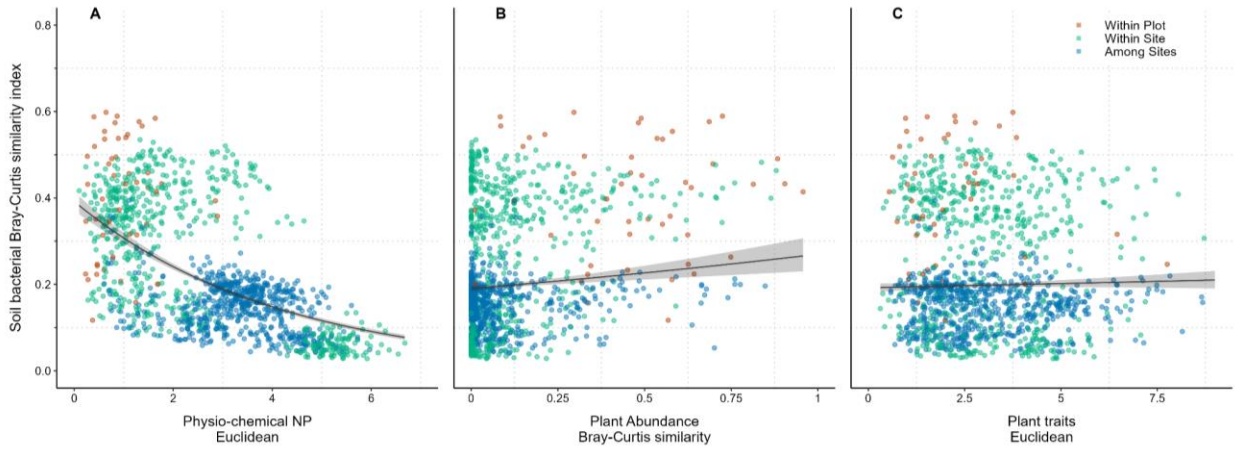

Fig. S1 Effect plots from a four-way Gamma generalized linear model (GLM) with a log link that predicted soil bacterial community similarity (Bray-Curtis index). Each panel displays the modeled relationship between Bray-Curtis similarity and one predictor: (A) physicochemical dissimilarity, (B) plant community dissimilarity, and (C) plant trait dissimilarity, with geographic distance and the remaining predictors held at their mean values. Physicochemical differences and plant community similarity significantly influenced bacterial similarity ($p < 0.001$), while plant trait dissimilarity showed no significant effect. Points represent observed pairwise similarity values; solid lines show model predictions, and shaded ribbons represent 95% confidence intervals.

Table S1 Taxonomic identification, spatial distribution, and relative abundance of dominant (> 0.1% average relative abundance) amplicon sequence variants (ASVs) that occur within the same sampling site, yet not within the same plot. Each ASV is listed with its corresponding phylum, family (representing the last known taxonomic identification), genus, log-transformed convex hull area (CH; m^2) as a proxy for spatial extent, and average relative abundance (%) across samples. Family depicts last known identification at that level. Unspecified genera indicate unresolved taxonomic classification at that level.

ID	Phylum	Family	Genus	Convex hull area	Abundance
ASV23	Actinobacteriota	MB-A2-108	unspecified	17.193	0.282
ASV28	Proteobacteria	*Sutterellaceae*	unspecified	15.441	0.268
ASV31	Crenarchaeota	*Nitrososphaeraceae*	unspecified	15.52	0.252
ASV40	Actinobacteriota	MB-A2-108	unspecified	17.193	0.208
ASV44	Verrucomicrobiota	*Xiphinematobacteraceae*	*Candidatus Xiphinematobacter*	15.52	0.202
ASV46	Proteobacteria	*Geminicoccaceae*	unspecified	15.52	0.19
ASV48	Proteobacteria	*Sutterellaceae*	unspecified	15.237	0.189
ASV55	Crenarchaeota	*Nitrososphaeraceae*	unspecified	15.52	0.176
ASV59	Actinobacteriota	*Nocardioidaceae*	Kribbella	15.298	0.199
ASV71	Actinobacteriota	*Actinomarinales*	unspecified	15.476	0.154
ASV73	Acidobacteriota	*Vicinamibacterales*	unspecified	15.376	0.153
ASV80	Proteobacteria	*Geminicoccaceae*	unspecified	15.52	0.134
ASV87	Proteobacteria	SC-I-84	unspecified	17.204	0.139
ASV90	Actinobacteriota	*Nocardioidaceae*	Kribbella	17.198	0.148
ASV94	Proteobacteria	TRA3-20	unspecified	17.188	0.116
ASV100	Actinobacteriota	*Gaiellales*	unspecified	17.204	0.134
ASV103	Actinobacteriota	*Gaiellaceae*	*Gaiella*	15.52	0.122
ASV104	Verrucomicrobiota	*Xiphinematobacteraceae*	*Candidatus Xiphinematobacter*	15.44	0.118