Soil Biology and Biochemistry 198 (2024) 109573

Contents lists available at ScienceDirect

Soil Biology and Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

# Bridging ecological processes to diversity formation and functional profiles in belowground bacterial communities

Beibei Chen[a], Ziheng Peng[a], Shi Chen[a], Yu Liu[a], Jiejun Qi[a], Haibo Pan[a], Hang Gao[a], Jiamin Gao[a], Chunling Liang[a], Jiai Liu[a], Xun Qian[a], Xiao Zhang[a], Sanfeng Chen[c], Jizhong Zhou[b], Gehong Wei[a,*], Shuo Jiao[a,**]

[a] State Key Laboratory of Crop Stress Biology in Arid Areas, Shaanxi Key Laboratory of Agricultural and Environmental Microbiology, College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, 712100, PR China
[b] Institute for Environmental Genomics and Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA
[c] State Key Laboratory for Agrobiotechnology and College of Biological Sciences, China Agricultural University, Beijing, 100193, PR China

ARTICLE INFO

*Keywords:*
Taxon pool size
Dispersal vs. selection
Diversity formation
Functional genes
Taxon coexistence

ABSTRACT

Revealing the generation and maintenance of biodiversity is a central goal in ecology, but how dispersal, selection, and regional taxon pool size shape soil microbial communities is not well understood. Here, we examined how dispersal and environmental selection affected soil bacterial diversity and their related metabolic functions by leveraging large-scale cross-biome soil surveys of ~1400 samples from diverse ecosystems across China, including agricultural, forest, grassland, and wetland soils. Our results showed that high dispersal increased $\alpha$-diversity and decreased $\beta$-diversity, whereas strong selection generated the opposite pattern in various ecosystems. This is likely due to dispersal enabling species access to otherwise unreachable habitats, and environmental selection excluding non-adapted species from communities. The $\alpha$-diversity increased with $\gamma$-diversity, whereas $\beta$-diversity did not covary. We also showed that bacterial phylotypes positively associated with dispersal and selection exhibited distinct metabolic diversity. Dispersal-induced phylotypes, which were abundant in agricultural soils, exhibited more metabolic diversity in fructose and mannose, starch and sucrose, and nitrogen metabolism. Conversely, selection-induced phylotypes, dominated in wetland soils, were primarily associated with sulfur and methane metabolism. In addition, the complexity of taxon associations increased when communities had higher selection increasing $\beta$-diversity. Our study establishes the predictive links of ecological processes to microbial diversity, metabolic functions, and taxon coexistence, thus facilitating a better understanding of the mechanisms underlying biodiversity generation and conservation.

## 1. Introduction

Soil microbial diversity plays a critical role in maintaining ecosystem function and services, such as soil genesis, decomposition, nutrient cycling, climate regulation and water detoxification (Huang et al., 2018; Benkwitt et al., 2020; Loreau et al., 2021). It is crucial to understand how soil microbial diversity is generated and maintained, which could improve our ability to predict how soil biodiversity loss might affect human wellbeing and ecosystem sustainability (Myers and Harms, 2009; Morrison-Whittle and Goddard, 2015). Although various ecological theories have been proposed (Vellend, 2016), the metacommunity theory provides an adequate framework to study the generation and maintenance mechanisms of soil microbial diversity (Leibold et al., 2004b). Under this theory, local communities are usually not isolated entities but are connected to each other by dispersal, and this process interacts with local selection and regional species pools (Wilson, 1992).

Disentangling the influence of dispersal, selection, and regional taxon pool size on soil microbial diversity is crucial for metacommunities (Leibold et al., 2004a). Variation in diversity may be caused by dispersal effects, as dispersal, which is defined as the

---

* Corresponding author. State Key Laboratory of Crop Stress Biology in Arid Areas, College of Life Sciences, Northwest A&F University, Yangling, Shaanxi 712100, PR China.

** Corresponding author. State Key Laboratory of Crop Stress Biology in Arid Areas, College of Life Sciences, Northwest A&F University, Yangling, Shaanxi, 712100, PR China.

*E-mail addresses:* weigehong@nwsuaf.edu.cn (G. Wei), shuojiao@nwsuaf.edu.cn (S. Jiao).

https://doi.org/10.1016/j.soilbio.2024.109573
Received 12 September 2023; Received in revised form 30 July 2024; Accepted 28 August 2024

movement of species, can promote species to reach habitats that would otherwise be inaccessible (Albright and Martiny, 2018). Selection, including environmental filtering (e.g., pH, temperature, moisture, and salinity) and biological interactions (e.g., competition, facilitation, mutualisms, and predation), could largely control the change in biodiversity (Chesson, 2000; Chave, 2004). For instance, when selection is variable, distinct taxa occupy different niches along environmental gradients, thereby creating diversity differences (Sunagawa et al., 2015; Thompson et al., 2017). Moreover, variation in diversity may also depend on changes in the size of regional species pools (Wang et al., 2021). In addition, the impacts of ecological processes on microbial communities may lead to variation in their functional traits (Walters et al., 2022). Numerous previous studies devote to exploring the biogeographic patterns of soil microbial diversity (Fierer et al., 2009; Delgado-Baquerizo et al., 2018; Oliverio et al., 2020), yet few studies have established the predictive links of ecological processes to soil microbial diversity and their related functional genes across relatively large spatial scales, which limits our ability to determine the mechanisms underpinning community formation and dynamics.

Exploring taxon coexistence can help integrate understanding of community structure and ecological rules (Vályi et al., 2016; Jiao et al., 2020). Ecological network analysis is used to explore the potential correlation among organisms coexisting in complex environments, which provides new insights into the structure and assembly of complex microbial communities (Barberán et al., 2012). Recently, the exploration of the factors influencing microbial co-occurrence associations is rapidly increasing. For example, dispersal may increase species coexistence because dispersal may facilitate more species to arrive and colonize a habitat (Baguette et al., 2013; Jiao et al., 2020), while an excessively high dispersal rate can be detrimental to species coexistence (Cadotte, 2006). Selection may be a key factor driving changes in microbial co-occurrence patterns, as selective pressure can not only directly affect potential associations between taxa, but also indirectly by stimulating changes in the relevant environment (Wang et al., 2018; Chen et al., 2021; Yuan et al., 2021). Alternatively, variation in regional species pool size can also lead to changes of taxon coexistence (Zobel, 1997; Cornell and Harrison, 2014; Karger et al., 2015). However, the combined influence of dispersal, selection, and regional species pool size on microbial co-occurrence associations across relatively large spatial scales remains scarce.

The aim of this study was to understand how dispersal, selection, and regional taxon pool size affect soil bacterial diversity, functional genes and taxon coexistence within the following framework (Fig. 1). A large-scale soil survey was conducted for bacterial community on 1379 soil samples across China that contain diverse ecosystems, including agricultural, forest, grassland, and wetland soils, which could help to detect generality of the findings across complex terrestrial ecosystems (e.g., natural and human-managed). A neutral model and an outlying mean index (OMI) analysis were employed to assess dispersal rate and the strength of environmental selection in bacterial communities, respectively. In this study, all references to selection specifically denote environmental selection. Given that dispersal effects promote more species to reach suitable habitats, selection effects filter out species that are not well-suited to the environment, and variation in diversity may also depend on regional species pool size (Zhou and Ning, 2017; Albright and Martiny, 2018; Wang et al., 2021), we hypothesized that (1) bacterial $\alpha$-diversity would increase with increasing dispersal/taxon pool size and decreasing selection, whereas $\beta$-diversity would decrease with increasing dispersal and decreasing selection/taxon pool size; (2) the complexity of taxon associations would increase when the strength of selection increases, whereas high dispersal would lead to less interconnected networks. With statistical evidence to directly examine the impacts of ecological processes on soil bacterial diversity, functional genes, and taxon coexistence, our results provide insights into the formation mechanisms and conservation of soil microbial diversity.

## 2. Methods and materials

### 2.1. Site and sampling

To elucidate the general rules of the biodiversity maintenance mechanisms, this study was conducted in four typical ecosystems (agricultural, forest, grassland, and wetland soils) in China, which extended from 19°59′N to 47°79′N and 86°30′E to 129°49′E with mean annual temperature from −2 to 24.3 °C and mean annual precipitation from 72 to 1853 mm. The plant community is dominated by *Zea mays* (agriculture), *Populus* spp., *Cunninghamia lanceolata*, *Casuarina equisetifolia* L. and *Phyllostachys heterocycla*, and shrubs (forest), *Stipa* spp., *Leymus* spp., *Festuca ovina*, and *Poa poophagorum* (grassland), and *Phragmites australis, Carex lasiocarpa*, and *Suaeda salsa* (wetland).

In total, 1379 soil samples were collected from 177 regions (50 agricultural fields, 44 forests, 43 grasslands, and 40 wetlands) across 60 study sites in China during July–August 2019 (Fig. S1). Among the 60 study sites, 28 sites include agricultural fields, forests, grasslands, and wetlands, 4 sites include agricultural fields, forests, and grasslands, 5 sites include agricultural fields, forests, and wetlands, 2 sites include forests, grasslands, and wetlands, 2 sites include agricultural fields, grasslands, and wetlands, 3 sites include agricultural fields and forests, 3 sites include agricultural fields and grasslands, 1 sites include forest and grasslands, 5 sites only include agricultural fields, 1 sites only include forests, 3 sites only include grasslands, and 3 sites only include wetlands. In each region, 5–12 plots (each of 10 × 10 m$^2$, adjacent plots: 100 m) were randomly established as described in previous study (Zhang et al., 2020), which allowed us to investigate the influence of dispersal, selection, and taxon pool size on bacterial diversity at a regional scale. In each plot, three soil cores (0–20 cm) were collected and mixed to form

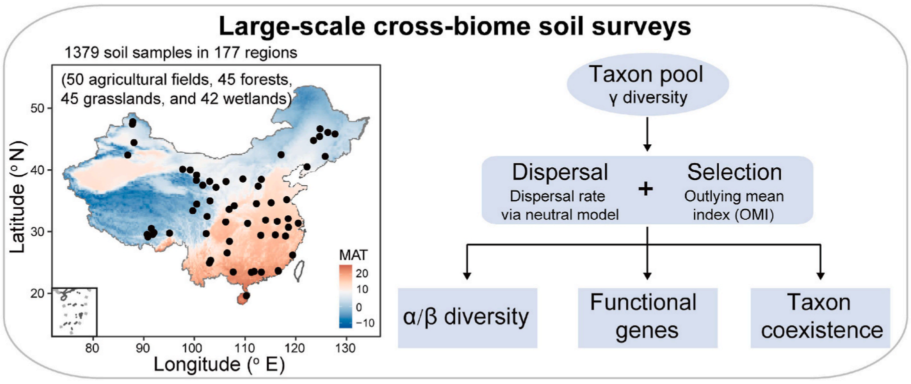

**Fig. 1.** A framework for studying how taxon pool size, dispersal, and selection affect bacterial diversity, functional genes, and taxon coexistence. The influence of dispersal, selection, and taxon pool size on bacterial diversity, functional genes, and taxon coexistence were identified by large-scale cross-biome soil surveys.

one composite sample and then transported to the laboratory in coolers.

### 2.2. Environmental factors measurement

Soil pH, moisture (SM), organic matter (SOM), nitrate-nitrogen ($NO_3^-$), ammonium-nitrogen ($NH_4^+$), and available phosphorus (AP) were measured as previously described (Ma et al., 2016; Shi et al., 2018). The mean annual temperature (MAT), mean annual precipitation (MAP), mean diurnal range (MDR), temperature seasonality (TS), temperature annual range (TAR), and precipitation seasonality (PS) of each site were extracted from the WorldClim database (www.worldclim.org). The aridity index (AI) of each site was extracted from the Global Potential Evapotranspiration database (Zomer et al., 2008).

### 2.3. DNA extraction and sequencing

DNA was extracted from 0.5 g of well homogenized soil using the MP FastDNA spin kit (MP Biomedicals, Solon, OH, USA). The V4 – V5 regions of bacterial 16S rRNA genes were amplified by primer pairs 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 907R (5′-CCGTCAATTCMTTTRAGTTT-3′) (Jiao et al., 2018). Sequencing was conducted on the Illumina NovaSeq PE 250 platform (Illumina Inc., San Diego, CA, USA). DADA2 pipeline was used to filter out low quality sequences, denoising, and removing chimeras (Callahan et al., 2016; Maver et al., 2021). Taxonomic assignment was conducted with a naïve Bayesian classifier using the Silva 138 database (Quast et al., 2012; Yilmaz et al., 2014). The sequences matching the mitochondria and chloroplast were also removed, and the remaining effective sequences were used to generate amplicon sequence variants (ASVs). In the end, 15,082 ASVs were eventually obtained. The representative sequences for 16S rRNA gene were aligned using Muscle (Edgar, 2022) and phylogenetic trees were constructed with FastTree (Price et al., 2009). Furthermore, samples were subsampled to an equal sampling depth, and the rarefaction curve reached saturation (Fig. S2).

To examine the influence of dispersal and selection-induced bacterial phylotypes on their related functional genes, we selected 10 regions covering the national country in each ecosystem (including agriculture, forest, grassland and wetland soils), and then mixed all soil samples from each region into one composite sample. 40 DNA samples (4 ecosystems × 10 regions) were used for shotgun metagenomic sequencing on an Illumina Novaseq 6000 platform (150-bp paired-end reads) at Guangdong Magigene Biotechnology Co., Ltd, but in total, only 39 DNA samples were used. Raw data were quality checked with FastQC (v0.11.9) and processed using Trimmomatic v.0.39 (leading: 3, trailing: 3, slidingwindow: 4:15, minlen:36) to trim adapters and discard bases with a quality score $<$15 and length $<$36 bp. After that, 12.2 Gbp clean data per sample were obtained. Clean reads were annotated for functional analysis of the microbiome using HUMAnN v3.7 (based on DIAMOND (version 2.1.6) 81 and Bowtie2 (version 2.5.1) 82) with ChocoPhlAn database (version “mpa_vJan21_CHOCOPhlAnSGB_202103”) and UniRef90 (version“uniref90_201901b”) protein database to quantify relative abundance of functional genes and metabolic pathways (Beghini et al., 2021). The annotation results were organized according to Kyoto Encyclopedia of Genes and Genomes (KEGG) database. According to the KEGG database, we identified amino acid metabolism, lipid metabolism, amino acid metabolism, lipid metabolism, nitrogen metabolism, phosphonate and phosphinate metabolism, sulfur metabolism, methane metabolism, xenobiotics biodegradation/metabolism and terpenoids polyketides metabolism, and carbohydrates metabolism (i.e., glycolysis/gluconeogenesis, TCA cycle, pentose phosphate pathway, fructose and Mannose metabolism, galactose metabolism, starch and sucrose metabolism, pyruvate metabolism, glyoxylate and dicarboxylate metabolism, butanoate metabolism, propanoate metabolism) (Xun et al., 2021).

### 2.4. Assessment of community dispersal rate

A neutral model was used to determine the importance of passive dispersal on community assembly via predicting the relationship between the frequency with which taxa occur in a set of local communities and their abundance across the wider metacommunity (Sloan et al., 2006). In the model, the dispersal rate is a parameter used to estimate the probability of a random loss of an individual in a local community to be displaced by an immigrant from the metacommunity (Jiao et al., 2020). It is calculated as follows: $\mathrm{Freq}_i = 1\text{-}\ I(1 \div N|N \times m \times p_i, N \times m \times (1\text{-}\ p_i))$ where $\mathrm{Freq}_i$ is the occurrence frequency of taxon $i$ across communities; $N$ is the number of individuals per community; $m$ is the estimated dispersal rate; $pi$ is the average relative abundance of taxon $i$ across communities; and $I$ is the probability density function of beta distribution. This analysis was conducted as nonlinear least squares fitting in the “minpack.lm” package (Elzhov et al., 2010). The overall fit of the model to observed data by comparing the sum of squares of residuals, $SS_{\mathrm{err}}$, with the total sum of squares, $SS_{\mathrm{total}}$: model fit $=$ $1{-}SS_{\mathrm{err}}/SS_{\mathrm{total}}$ (generalized $R$-squared) (Jiao et al., 2020). To determine whether the model was based on only the random sampling of the source metacommunity, the fit of the neutral model with the fit of a binomial distribution model was compared (Burns et al., 2016). Sampling from a binomial distribution represents the case where local communities are random subsets of the metacommunity in the absence of drift and dispersal limitation processes (Burns et al., 2016). The Akaike information criterion of each model was examined based on 1000 bootstrap replicates. Our result showed the neutral models were superior to the binomial distribution model in various ecosystems, indicating that dispersal had an influence beyond the random sampling of the source community (Fig. S3). In addition, the number of the shared taxon number across plots in each region was calculated, and the proportion of the number of the shared taxon divided by regional taxon pool size was used to estimate dispersal. A higher ratio indicates stronger dispersal (Mei and Liu, 2019).

### 2.5. Assessment of environmental selection

Outlying mean index (OMI) analysis was used to estimate the effects of environmental selection on bacterial communities (Luan et al., 2020). The OMI could determine phylotypes’ niche positions via measuring the distance between the mean environmental conditions used by each phylotype and the mean environmental conditions of the study area (Dolédec et al., 2000; Thuiller et al., 2005). The OMI makes no hypothesis on the shape of species response curves to the environment and gives equal weight to species-poor and species-rich sites (Luan et al., 2020). This analysis was ran using “niche” function in the “ade4” package (Dolédec et al., 2000; Thuiller et al., 2005). Specifically, we conducted principal component analysis to reduce the dimensions of all environmental variables. Each phylotype was correlated with the row profiles of the resulting environmental data, and the average position of each phylotype was then calculated along the ordination axes in each region. A high OMI value indicates that each phylotype has a narrow niche breadth, thus suggesting that each phylotype is subjected to higher environmental selection (Luan et al., 2020). In addition, environmental heterogeneity was estimated by the average environmental dissimilarity among sites using the Gower distance based on the combination of 13 variables related to soil properties (pH, SM, SOM, $NO_3^-$, $NH_4^+$, and AP) and climatic conditions (MAT, MAP, MDR, TS, TAR, PS, and AI) (Albrecht et al., 2021). A high environmental dissimilarity value indicates high environmental heterogeneity.

### 2.6. Statistical analyses

We calculated the variation in ($\gamma$, $\alpha$, and $\beta$) diversity in different ecosystems. Here, $\alpha$-diversity was measured as ASV richness of a single plot, and the $\gamma$-diversity (taxon pool size) was measured as the total ASVs richness in each sampling region (Wang et al., 2021; Xu et al., 2021). $\beta$-diversity was measured based on the Bray–Curtis dissimilarity in bacterial composition for each region. To visualize the sample relationships of bacterial communities among different ecosystems, NMDS analysis was performed based on Bray-Curtis dissimilarity by using the "metaMDS" function of the "vegan" package (Oksanen et al., 2015). To determine the significance in bacterial community composition based on Bray–Curtis dissimilarity matrices in different ecosystems, similarity analysis (Anosim) and permutational multivariate analysis of variance (Adonis) were performed using the anosim and adonis function, respectively, in "vegan" package in R (Oksanen et al., 2015). To assess the distance-decay relationship of community similarity, we calculated the slopes of ordinary least-squares regressions for the relationships between geographic distances and community similarities (1—dissimilarity of the Bray–Curtis metric) (Martiny et al., 2006). The geographical distances among the sampling sites were calculated from the sampling coordinates. Differences in the diversity ($\alpha$, $\beta$, and $\gamma$), dispersal rate, outlying mean index (OMI), and environmental heterogeneity were examined using Kruskal–Wallis tests in the "agricolae" package (De Mendiburu, 2014). Differences in the number of metabolic pathways were examined between dispersal and OMI-induced bacterial phylotypes using Wilcoxon rank sum test. All statistical analyses were conducted using R (v3.4.1) for Windows (http://www.r-project.org/).

**Dispersal and selection-induced phylotypes**. To examine the effect of dispersal and selection-induced bacterial phylotypes on their related metabolic diversity, we first determined the ASVs that occurred in more than 3% of all soil samples when all four ecosystems were considered together. Then, 349 ASVs and 222 ASVs with significant positive responses to dispersal or outlying mean index (OMI) were identified, respectively. Three different phylogenetic trees for these 571 ASVs with clear annotation at phylum/class-level were constructed using MEGA6 by maximum likelihood method, neighbor joining method, and unweighted pair group method using arithmetic average, respectively, and visualized using "ggtree" package in R (Yu et al., 2017).

**Random forest (RF) analysis**. We performed RF model to predict the relative importance of dispersal, selection, and taxa pool size for bacterial diversity in the "rfPermute" package (Bento et al., 2002). The cross-validated $R^2$ and significance were calculated in the "A3" package (Fortmann-Roe, 2015).

**Co-occurrence networks.** To estimate taxon coexistence in different ecosystems, co-occurrence networks were constructed in each ecosystem, respectively. To reduce rare ASVs in each ecosystem dataset, we removed ASVs with a relative abundance <0.01% (Jiao et al., 2020). Therefore, the correlation networks were constructed using 2151 ASVs in 390 agricultural samples, 2174 ASVs in 337 forest samples, 2190 ASVs in 347 grassland samples and 2214 ASVs in 305 wetland samples, respectively. Robust correlations with Spearman's correlation coefficients ($r$) > 0.6 and false discovery rate-corrected $p$-values <0.01 were used to construct networks in package "fdrtool" (Klaus and Strimmer, 2012). Each region networks (5–12 samples) were also generated in different ecosystem by the same method as above. Networks were visualized using the interactive Gephi platform (Bastian et al., 2009). To describe the complexity of taxon associations, five network-level topological features were calculated in the package "igraph" (Csardi and Nepusz, 2006), including the number of nodes and edges, average degree (the average connections of each node with another particular node in the network), clustering coefficient (the degree to which the nodes tend to cluster together) and network density (the intensity of connections among nodes) (Newman, 2003, 2006). Average degree was used to represent the complexity of taxon associations because these topological features were tightly associated (Xiong et al., 2021).

## 3. Results

### 3.1. Effects of dispersal, selection, and taxon pool size on soil bacterial diversity

The NMDS results showed that soil samples formed clear groupings according to different ecosystems in the ordination space (Fig. S4). Significant differences in bacterial community composition were found among different ecosystems (Adonis $R^2 = 0.079$, $p < 0.001$; and Anosim tests, $R = 0.252$, $p < 0.001$, Fig. S4). Distance-decay curves for all four typical ecosystems were significant. The slope was the steepest in wetland soils (slope = - 0.297, $R^2 = 0.102$) and the flattest in agricultural soils (slope = −0.192, $R^2 = 0.022$; Fig. 2a). The ecological processes behind the distance-decay curves were then estimated, and the variations of taxon pool size ($\gamma$-diversity), dispersal rate (DR - m value of Sloan neutral models), and outlying mean index (OMI - selection) were compared in agricultural, forest, grassland, and wetland soils. The $\gamma$-diversity was the highest in forest, grassland, and wetland soils, and lowest in agricultural soils (Fig. 2b). Effect of DR was significantly higher in agricultural than other soils (Fig. 2c), while OMI was the highest in wetlands and lowest in agricultural fields (Fig. 2d). This result was further supported by the proportion of shared taxon (potential dispersal) and the environmental heterogeneity (Figs. S5a and b). Moreover, bacterial $\alpha$-diversity was the highest in agricultural fields and lowest in wetlands, while $\beta$-diversity had a reversed pattern (Fig. 2e and f).

To further disentangle how taxon pool size and ecosystems processes affect bacterial diversity, we examined the relationship between bacterial ($\alpha$ and $\beta$) diversity and $\gamma$-diversity, dispersal rate, and OMI. Bacterial $\alpha$-diversity increased with increasing $\gamma$-diversity in agricultural, forest, and wetland soils and when all four ecosystems were considered together (Fig. 2g). No significant correlations were found between $\gamma$-diversity and bacterial $\beta$-diversity in various ecosystems, although $\gamma$-diversity slightly promoted bacterial $\beta$-diversity when all four ecosystems were considered together (Fig. 2g). In addition, the DR and proportion of the shared taxon had a significant positive influence on bacterial $\alpha$-diversity in all four ecosystems, but a negative effect on bacterial $\beta$-diversity (Figs. 2h and S5c). In contrast, bacterial $\alpha$-diversity decreased with increasing OMI or environmental heterogeneity, but $\beta$-diversity increased (Figs. 2i and S5d).

To disentangle the main predictors of soil bacterial diversity in various ecosystems, we evaluated the relative importance of $\gamma$-diversity, dispersal rate, and OMI for bacterial ($\alpha$ and $\beta$) diversity by random forest analysis. Dispersal rate was the primary factor affecting bacterial $\alpha$-diversity in agricultural fields, forests, grasslands, wetlands, and when all four ecosystems were considered together (Fig. S6). Moreover, dispersal rate was the most important variable for predicting bacterial $\beta$-diversity in agricultural, forest, and grassland soils, while the $\beta$-diversity was highly structured by OMI in wetland soils and when all four ecosystems were considered together (Fig. S6).

### 3.2. Effects of dispersal and selection-induced bacterial phylotypes on their related functional genes

To further examine the relative importance of dispersal and selection for the finer individual taxa levels, we first identified the bacterial phylotypes that were positively affected by dispersal and OMI, respectively. A total of 349 ASVs and 222 ASVs were mainly structured by dispersal and selection, respectively. Specifically, the relative abundance of Bacteroidota, Chloroflexi, Actinobacteriota, Acidobacteriota, Gemmatimonadota, Planctomycetota, Methylomirabilota, and Alphaproteobacteria was primarily increased by dispersal, whereas the relative abundance of Firmicutes and Desulfobacterota was mainly increased by selection (Fig. 3a and Fig. S7). Then, we compared the functional genes of the dispersal and selection-induced bacterial phylotypes via metagenomic sequencing. Our results showed that dispersal-induced

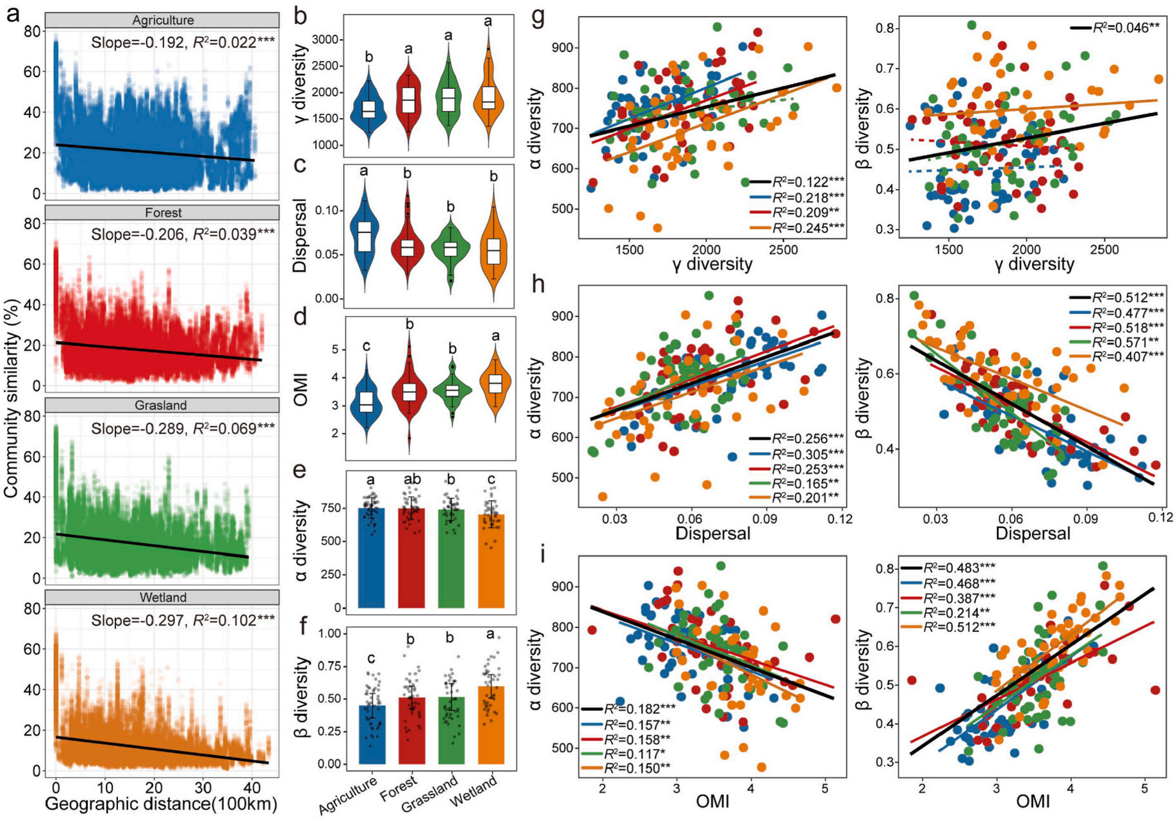

**Fig. 2. Relation between bacterial diversity and taxon pool size, dispersal, and selection.** Distance–decay relationships between geographic distances and bacterial similarities (1—dissimilarity of the Bray–Curtis metric, **a**). Changes in $\gamma$-diversity (**b**), dispersal rate (**c**), outlying mean index (OMI, **d**), $\alpha$-diversity (**e**), $\beta$-diversity (**f**) across different ecosystems. The lowercase suggests significant difference between ecosystems by Kruskal-Wallis tests. Linear correlation between bacterial diversity and $\gamma$-diversity (**g**), dispersal rate (**h**), and outlying mean index (OMI, **i**) was estimated by regression analysis in agricultural fields, forests, grasslands, wetlands, and when all four ecosystems were considered together, respectively. Solid and dashed lines represent the significant and nonsignificant correlation, respectively. Significance levels were as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

bacterial phylotypes, which had the highest relative abundance in agricultural soils (Fig. S8a), exhibited more metabolic diversity, particularly in the metabolism of fructose and mannose, starch and sucrose, nitrogen, terpenoids and polyketides. Conversely, OMI-induced bacterial phylotypes, prevailed in wetland soils (Fig. S8b), primarily associated with the sulfur, methane, and xenobiotics biodegradation/metabolism (Fig. 3b). However, there was no significant difference of the metabolic diversity between dispersal and selection-induced bacterial phylotypes in butanoate metabolism, galactose metabolism, glycolysis/gluconeogenesis, glyoxylate and dicarboxylate metabolism, propanoate metabolism, TCA cycle, lipid metabolism, and phosphonate and phosphinate metabolism (Fig. S9).

### *3.3. Relationships between ecological processes and taxon associations*

To examine how the ecological processes affect soil bacterial co-occurrence patterns, we established co-occurrence networks among bacterial taxa (Fig. 4). The differences of network-level topological features between agricultural, forest, grassland, and wetland soils were compared. The number of nodes and edges, average degree, clustering coefficient, and density were the highest in wetland, followed by grassland, forest and agricultural ecosystems (Fig. 4a and b), which indicated wetland network was more complex. In addition, OMI and $\beta$-diversity was primarily and positively correlated with the complexity of taxon associations in all four ecosystems (Fig. **4c-g**), indicating increasing selection promoted more potential associations between taxa increasing $\beta$-diversity. However, DR was negatively correlated with network complexity except for grassland ecosystems (Fig. 4c, d, f, and g). Negative relationship were observed between network complexity and $\gamma$-diversity in grassland soils and when all four ecosystems were considered together (Fig. 4e and g). There was no significant correlation between network complexity and $\alpha$-diversity in diverse ecosystems (Fig. 4c–f).

## 4. Discussion

Understanding the mechanisms underpinning the generation and maintenance of soil microbial diversity is a key focus of ecology (Zhou and Ning, 2017). Here, based on a national-scale soil survey, we demonstrated that: (i) dispersal significantly increased $\alpha$-diversity and decreased $\beta$-diversity, whereas strong selection generated the opposite pattern; (ii) dispersal and selection structured distinct metabolic diversity by shaping specific bacterial phylotypes; and (iii) the complexity of taxon associations increased when communities had higher selection increasing $\beta$-diversity.

In the present study, we observed that bacterial $\alpha$-diversity significantly increased with increasing dispersal rate. This may be attributed to that increasing dispersal increases the opportunities for soil microorganisms reaching the habitats by different dispersal strategies (Walters et al., 2022), resulting in increased $\alpha$-diversity. This could also be that local communities are not necessarily saturated—current biomass was far below carrying capacity of the environment (Cadotte, 2006), causing local $\alpha$-diversity to increase as dispersal rate increase. Our results showed that bacterial $\beta$-diversity significantly decreased with increasing dispersal. This may be due to the fact that dispersal can homogenize communities directly via the immigration of individuals, leading to a decrease in $\beta$-diversity (Evans et al., 2017; Fodelianakis et al., 2019). We also found that bacterial $\alpha$-diversity significantly decreased with

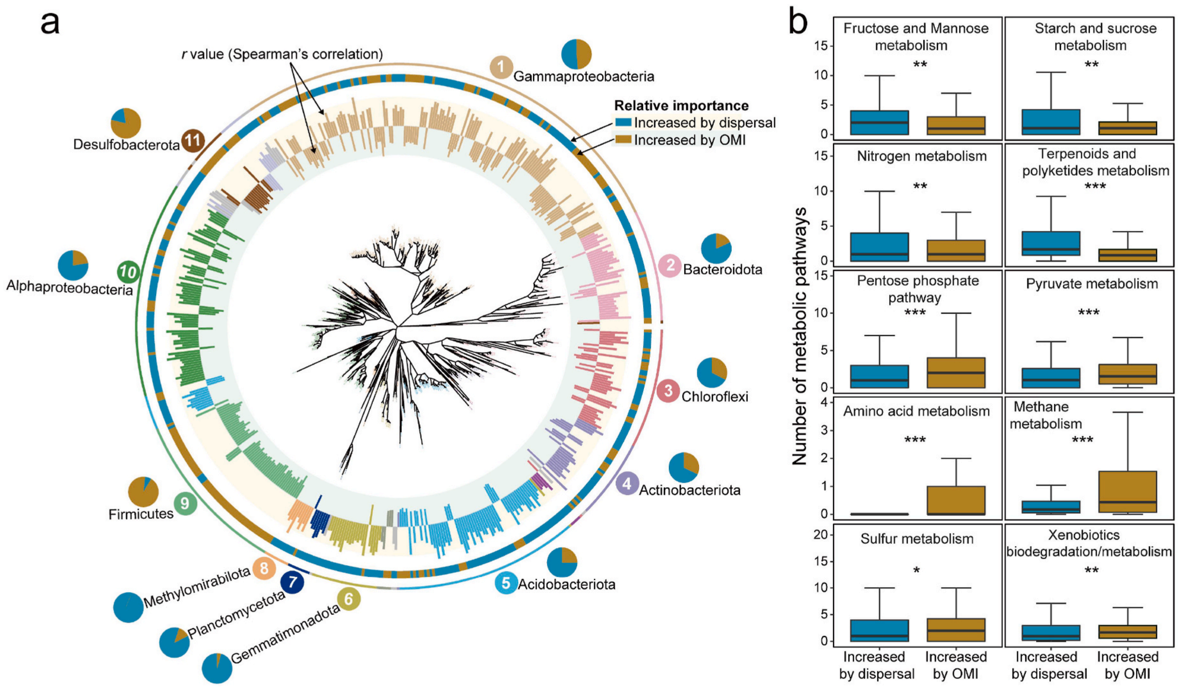

**Fig. 3.** Effects of dispersal and selection-induced bacterial phylotypes on their related functional genes. The phylogenetic tree was constructed with 571 species with clear taxonomic affiliations using maximum likelihood method (**a**). Only ASVs with significant positive response (Spearman's correlations, $p < 0.05$) to dispersal or outlying mean index (OMI) and occurring in more than 3% of all soil samples when all four ecosystems were considered together were included in the tree. The outside and inside bars of the second ring represent the positive effect ***r*** value (Spearman's correlation) of dispersal and outlying mean index on bacterial taxon relative abundances, respectively. Colors for both the branch and the outermost ring represent different Phyla or classes, and the gray color in the tree indicates unclassified or other minor phyla. Colors for the third ring represent ASVs primarily affected by dispersal (Brown) and outlying mean index (Blue), respectively. The pies represent the proportions of the primarily affected by dispersal (Brown) and outlying mean index (Blue) in different phyla/classes. Differences in the number of metabolic pathways were estimated between dispersal and OMI-induced bacterial phylotypes (**b**). The number of metabolic pathways was counted by Spearman correlation between dispersal/OMI-induced bacterial phylotypes and different metabolic pathways. Brown asterisks indicate significantly higher values in taxa increased by dispersal; Blue asterisks indicate significantly higher values in taxa increased by OMI (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns = not significant, Wilcoxon rank sum test).

increasing environmental selection. One possible explanation is that increased environmental selection allows only those microbes that can adapt to the environment to survive, while those unable to adapt are eliminated (Rillig et al., 2019; Yang et al., 2022). Moreover, our results showed that bacterial $\beta$-diversity significantly increased with rising environmental selection. This increase may be due to environmental selection filtering out some species from local communities, leading different species to occupy distinct and optimal niches along environmental gradients. Consequently, the variation in microbial $\beta$-diversity increases as species disappear and are replaced by others (Vellend, 2010; Dini-Andreote et al., 2015; Fodelianakis et al., 2019). In addition, we also identified positive effect of $\gamma$-diversity on $\alpha$-diversity, suggesting local species richness depends upon the regional species pool in a manner consistent with proportional sampling (Cornell, 1993). However, our results are not in line with previous findings, suggesting that $\alpha$-diversity begins to decline at intermediate dispersal rate (Loreau and Mouquet, 1999; Mouquet and Loreau, 2002), and $\beta$-diversity increases with increasing dispersal rate (Vannette and Fukami, 2017). Moreover, besides the positive relationships between ($\alpha$ and $\beta$) diversity selection, negative and unimodal relationships have also been reported (Ronen Kadmon and Omri Allouche, 2007; Bar-Massada and Wood, 2014; Barczyk et al., 2023). These inconsistent results may be attributed to that the effects of dispersal and selection on $\alpha$- and $\beta$-diversity are scale-dependent (Tamme et al., 2010; Barczyk et al., 2023). Overall, our findings provided statistical evidence on the predictable relationships between ecological processes and soil microbial diversity, suggesting that the relative variation in soil microbial diversity would count on which processes show a greater response at the larger spatial scales (Cadotte, 2006).

The second important finding of our study was that dispersal and selection had a significant differential influence on bacterial metabolic diversity. We found that dispersal-induced bacterial phylotypes, which dominated the agricultural soils, exhibited a pronounced metabolic diversity, particularly in the metabolism of fructose and mannose, starch and sucrose, nitrogen, terpenoids and polyketides. This could be that agricultural ecosystems are frequently subjected to intensive anthropogenic management, such as soil tillage and crop transplantation (Ponisio et al., 2016), which may facilitate microbial passive dispersal and allow more species to reach diverse habitats (Vannette and Fukami, 2017; Zhou and Ning, 2017). In turn, this enhances the fundamental metabolic functions of living microorganisms, including fructose and mannose metabolism, starch and sucrose metabolism, as well as nitrogen metabolism, which are all intracellular metabolic processes involving carbohydrates and nitrogen (Xun et al., 2021). Furthermore, bacterial phylotypes induced by strong selection were prevalent in wetland soils, and exhibited a higher metabolic diversity, particularly in the sulfur metabolism, methane metabolism, and xenobiotics biodegradation/metabolism. This could be that strong selection could screen the specialized species to perform specific metabolic functions. For example, our result showed that Desulfobacterota was mainly affected by selection. Previous studies have indicated that the sulfur and xenobiotics biodegradation/metabolism function were restricted to specialized microorganisms (Kumar et al., 2017; Hausmann et al., 2018). Moreover, several investigations have suggested that anaerobic environments (e.g., wetland ecosystems) might introduce specific species involved in methane metabolism, such as methanogens and methane-oxidizing species (Zhu et al., 2015; Lever, 2016). Our study illustrates the importance of ecological processes in driving bacterial

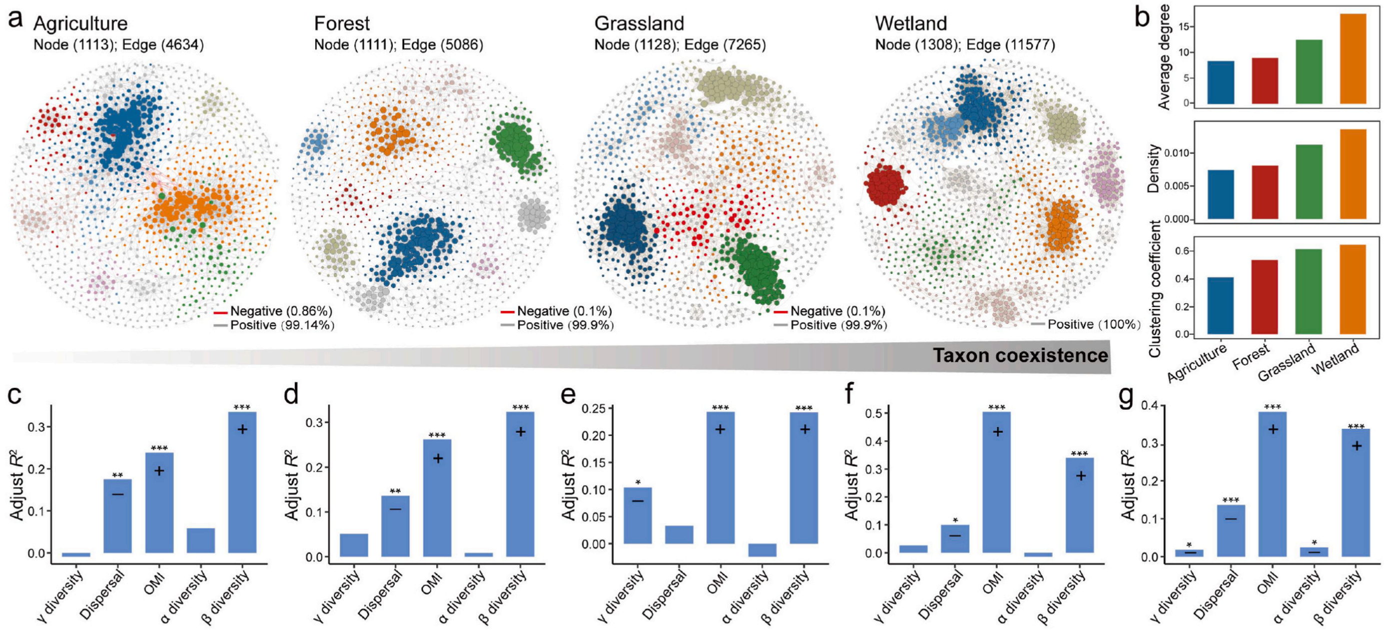

**Fig. 4.** Relationships of dispersal, selection, and taxon pool size, diversity on soil bacterial network complexity in different ecosystems. Co-occurrence networks were established in each ecosystem bases on Spearman's correlation analysis (**a**). Nodes colors indicate different network clusters; node size is positive to the degree of each ASV; edge thickness between two nodes is positively associated with correlation coefficient. Red edge refers to negative associations while gray edge to positive ones. A summary of node–edge statistics is in the top-left. Changes in unique network-level topological features in different ecosystems (**b**). The relation between $\gamma$-diversity, dispersal rate, outlying mean index (OMI), $\alpha$-diversity, $\beta$-diversity, and the complexity of soil bacterial correlation network was evaluated via linear least-squares regression analysis (adjust $R^2$; slope: positive (+) or negative (−)) in agricultural fields (**c**), forests (**d**), grasslands (**e**), wetlands (**f**), and when all four ecosystems (**g**) were considered together. Significance levels were as follows: $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$.

communities to perform different soil functions and metabolic functions across terrestrial ecosystems.

The third notable finding from our study showed that as environmental selection intensified across diverse ecosystems, the complexity of soil bacterial correlation networks correspondingly increased. Our finding is consistent with the notion that intensified environmental selection, such as climate warming and increasing precipitation, has the potential to enhance the complexity of species associations (Yuan et al., 2021; Wang et al., 2018). Rising temperature might stimulate various biotic potential association owing to more active individual metabolic processes and faster growth at higher temperature (Yuan et al., 2021). Higher network associations under a higher level of precipitation might be explained, in part, by the increasing biomass stimulated by a greater supply of water and nutrients, providing more opportunities for different species to associations with each other (Wang et al., 2018). Moreover, previous studies show that environmental heterogeneity facilitates more interconnected networks with increasing in $\beta$-diversity (Barberán et al., 2012; Huber et al., 2020) because higher heterogeneity lead to taxa occupying multiple different niches (Banerjee et al., 2019). In line with these previous researches (Barberán et al., 2012; Huber et al., 2020), our results showed that there was more complexity of bacterial correlation network in wetland soils, which might be associated with their higher environmental heterogeneity and $\beta$-diversity. In addition, when environmental selection is strong, microbes are able to resist harsh environmental stress by increasing huge number of connection with other taxa (Chen et al., 2021; Yuan et al., 2021). We also found that the complexity of bacterial association network decreased with increasing dispersal. This observation was supported by the notion that dispersal might decrease microbial interconnected networks with the declining in $\beta$-diversity (Huber et al., 2020), and too high dispersal could reduce species coexistence (Kneitel and Miller, 2003). Additionally, since dispersal homogenizes communities, species would compete for similar sources, reducing local species coexistence rate (Chesson and Neuhauser, 2002; De Cáceres et al., 2012). Overall, our findings facilitated the understanding of the maintaining mechanism of soil bacterial coexistence.

## 5. Conclusions

We proposed a conceptual framework describing how dispersal and selection affect soil bacterial diversity, functional genes, and taxon coexistence (Fig. 5). First, soil bacterial $\alpha$-diversity increased with increasing dispersal, but decreased with increasing selection. In contrast, soil bacterial $\beta$-diversity decreased when dispersal was high, but increased when selection was strong. Second, dispersal and selection-induced bacterial phylotypes exhibited significantly different metabolic diversity. Third, selection increased the number of interspecies interactions increasing $\beta$-diversity. Given the importance of biodiversity for maintaining ecosystem functions, understanding the mechanisms that maintain diversity is vital. By revealing the linkages between ecological processes and bacterial diversity, functional genes, taxon coexistence, our study provides an insightful understanding of the mechanisms underlying diversity formation and functional genes of soil bacterial communities.

## CRediT authorship contribution statement

**Beibei Chen:** Writing – review & editing, Writing – original draft, Data curation. **Ziheng Peng:** Data curation. **Shi Chen:** Data curation. **Yu Liu:** Data curation. **Jiejun Qi:** Data curation. **Haibo Pan:** Data curation. **Hang Gao:** Data curation. **Jiamin Gao:** Data curation. **Chunling Liang:** Data curation. **Jiai Liu:** Data curation. **Xun Qian:** Methodology. **Xiao Zhang:** Methodology. **Sanfeng Chen:** Writing – review & editing. **Jizhong Zhou:** Methodology. **Gehong Wei:** Data curation. **Shuo Jiao:** Funding acquisition.

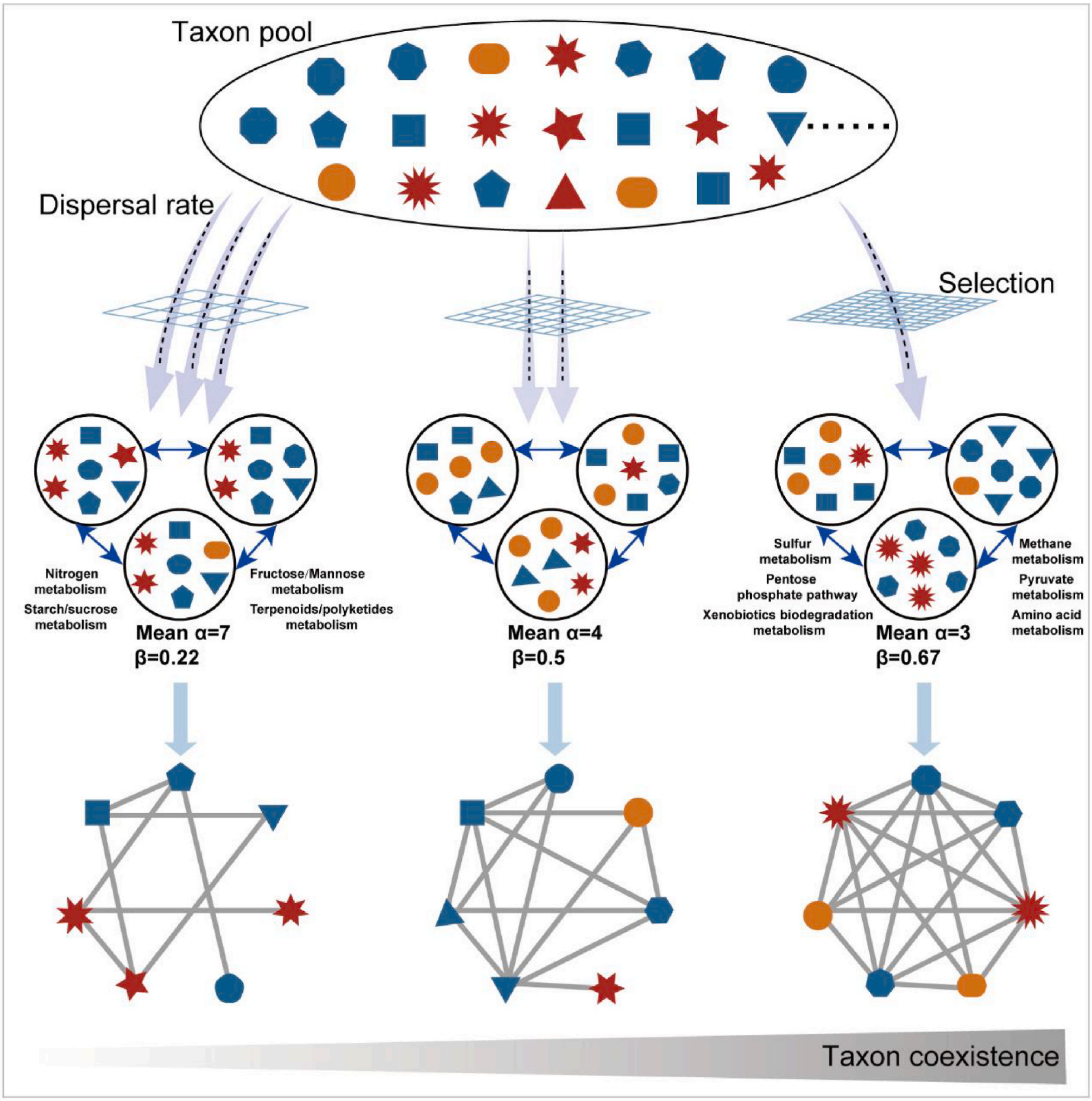

**Fig. 5.** A conceptual framework showing the influence of dispersal rate and selection on soil bacterial diversity, functional genes, and taxon coexistence. Throughout the figure, the geometric symbols refer to bacterial taxon. Large ovals refer to taxon pools and small circles to local communities consisting of a subset of taxon from taxon pool. The arrows with dotted line going from taxon pool to the local communities represent dispersal rate, and the grid size represent the strength of selection.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The data reported in this paper are available at https://doi.org/10.6084/m9.figshare.22276720.

## Acknowledgements

We thank Prof. Josh D. Neufeld from University of Waterloo for their constructive comments on the manuscript. This work was supported by the National Science Foundation for Excellent Young Scholars of China (grant No.: 42122050), National Science Foundation of China (grant No.: 42077222, 41830755, and 41807030), and the Fundamental Research Funds for the Central Universities (2022HHZX005).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.soilbio.2024.109573.

## References

Albrecht, J., Peters, M.K., Becker, J.N., Behler, C., Classen, A., Ensslin, A., Ferger, S.W., Gebert, F., Gerschlauer, F., Helbig-Bonitz, M., Kindeketa, W.J., Kühnel, A., Mayr, A. V., Njovu, H.K., Pabst, H., Pommer, U., Röder, J., Rutten, G., Schellenberger Costa, D., Sierra-Cornejo, N., Vogeler, A., Vollstädt, M.G.R., Dulle, H.I., Eardley, C. D., Howell, K.M., Keller, A., Peters, R.S., Kakengi, V., Hemp, C., Zhang, J., Manning, P., Mueller, T., Bogner, C., Böhning-Gaese, K., Brandl, R., Hertel, D., Huwe, B., Kiese, R., Kleyer, M., Leuschner, C., Kuzyakov, Y., Nauss, T., Tschapka, M., Fischer, M., Hemp, A., Steffan-Dewenter, I., Schleuning, M., 2021. Species richness is more important for ecosystem functioning than species turnover along an elevational gradient. Nature Ecology & Evolution 5, 1582–1593.

Albright, M.B.N., Martiny, J.B.H., 2018. Dispersal alters bacterial diversity and composition in a natural community. The ISME Journal 12, 296–299.

Baguette, M., Blanchet, S., Legrand, D., Stevens, V.M., Turlure, C., 2013. Individual dispersal, landscape connectivity and ecological networks. Biological Reviews 88, 310–326.

Banerjee, S., Walder, F., Büchi, L., Meyer, M., Held, A.Y., Gattinger, A., Keller, T., Charles, R., van der Heijden, M.G., 2019. Agricultural intensification reduces microbial network complexity and the abundance of keystone taxa in roots. The ISME Journal 13, 1722–1736.

Bar-Massada, A., Wood, E.M., 2014. The richness–heterogeneity relationship differs between heterogeneity measures within and among habitats. Ecography 37, 528–535.

Barberán, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. The ISME Journal 6, 343–351.

Barczyk, M.K., Acosta-Rojas, D.C., Espinosa, C.I., Schleuning, M., Neuschulz, E.L., 2023. Biotic pressures and environmental heterogeneity shape beta-diversity of seedling communities in tropical montane forests. Ecography, e06538, 2023.

Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. Proceedings of the International AAAI Conference on Web and Social Media, pp. 361–362.

Beghini, F., McIver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A.M., Valles-Colomer, M., Weingart, G., Zhang, Y., Zolfo, M., Huttenhower, C., Franzosa, E.A., Segata, N., 2021. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife 10, e65088.

Benkwitt, C.E., Wilson, S.K., Graham, N.A.J., 2020. Biodiversity increases ecosystem functions despite multiple stressors on coral reefs. Nature Ecology & Evolution 4, 919–926.

Bento, A.P., Gaulton, A., Hersey, A., Bellis, L.J., Chambers, J., Davies, M., Kruger, F.A., Light, Y., Mak, L., Mcglinchey, S., 2002. Classification and regression by randomForest. R News 23.

Burns, A.R., Stephens, W.Z., Stagaman, K., Wong, S., Rawls, J.F., Guillemin, K., Bohannan, B.J., 2016. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. The ISME Journal 10, 655–664.

Cadotte, M.W., 2006. Dispersal and species diversity: a meta-analysis. The American Naturalist 167, 913–924.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nature Methods 13, 581–583.

Chave, J., 2004. Neutral theory and community ecology. Ecology Letters 7, 241–253.

Chen, B., Jiao, S., Luo, S., Ma, B., Qi, W., Cao, C., Zhao, Z., Du, G., Ma, X., 2021. High soil pH enhances the network interactions among bacterial and archaeal microbiota in alpine grasslands of the Tibetan Plateau. Environmental Microbiology 23, 464–477.

Chesson, P., 2000. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics 31, 343–366.

Chesson, P., Neuhauser, C., 2002. Intraspecific aggregation and species coexistence. Trends in Ecology & Evolution 17, 210–211.

Cornell, H.V., 1993. Unsaturated patterns in species assemblages: the role of regional processes in setting local species richness. Species Diversity in Ecological Communities. University of Chicago Press, Chicago, Illinois, USA, pp. 243–252.

Cornell, H.V., Harrison, S.P., 2014. What are species pools and when are they important? Annual review of ecology. Evolution, and Systematics 45, 45–67.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. InterJournal, complex systems 1695, 1–9.

De Cáceres, M., Legendre, P., Valencia, R., Cao, M., Chang, L.-W., Chuyong, G., Condit, R., Hao, Z., Hsieh, C.-F., Hubbell, S., Kenfack, D., Ma, K., Mi, X., Supardi Noor, M.N., Kassim, A.R., Ren, H., Su, S.-H., Sun, I.-F., Thomas, D., Ye, W., He, F., 2012. The variation of tree beta diversity across a global network of forest plots. Global Ecology and Biogeography 21, 1191–1202.

De Mendiburu, F., 2014. Agricolae: statistical procedures for agricultural research. R package version 1, 1–4.

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, 320–325.

Dini-Andreote, F., Stegen, J.C., van Elsas, J.D., Salles, J.F., 2015. Disentangling Mechanisms that Mediate the Balance between Stochastic and Deterministic Processes in Microbial Succession, vol. 112. Proceedings of the National Academy of Sciences, pp. E1326–E1332.

Dolédec, S., Chessel, D., Gimaret-Carpentier, C., 2000. Niche separation in community analysis: a new method. Ecology 81, 2914–2927.

Edgar, R.C., 2022. High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. bioRxiv 2021 (2006), 449169, 2020.

Elzhov, T.V., Mullen, K.M., Spiess, A., Bolker, B., 2010. R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK. Plus support for bounds 1, 2, 1.

Evans, S., Martiny, J.B., Allison, S.D., 2017. Effects of dispersal and selection on stochastic assembly in microbial communities. The ISME Journal 11, 176–185.

Fierer, N., Strickland, M.S., Liptzin, D., Bradford, M.A., Cleveland, C.C., 2009. Global patterns in belowground communities. Ecology Letters 12, 1238–1249.

Fodelianakis, S., Lorz, A., Valenzuela-Cuevas, A., Barozzi, A., Booth, J.M., Daffonchio, D., 2019. Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. Nature Communications 10, 1–12.

Fortmann-Roe, S., 2015. Consistent and clear reporting of results from diverse modeling techniques: the A3 method. Journal of Statistical Software 66, 1–23.

Hausmann, B., Pelikan, C., Herbold, C.W., Köstlbacher, S., Albertsen, M., Eichorst, S.A., Glavina del Rio, T., Huemer, M., Nielsen, P.H., Rattei, T., Stingl, U., Tringe, S.G., Trojan, D., Wentrup, C., Woebken, D., Pester, M., Loy, A., 2018. Peatland Acidobacteria with a dissimilatory sulfur metabolism. The ISME Journal 12, 1729–1742.

Huang, Y., Chen, Y., Castro-Izaguirre, N., Baruffol, M., Brezzi, M., Lang, A., Li, Y., Härdtle, W., von Oheimb, G., Yang, X., Liu, X., Pei, K., Both, S., Yang, B., Eichenberg, D., Assmann, T., Bauhus, J., Behrens, T., Buscot, F., Chen, X.-Y., Chesters, D., Ding, B.-Y., Durka, W., Erfmeier, A., Fang, J., Fischer, M., Guo, L.-D., Guo, D., Gutknecht, J.L.M., He, J.-S., He, C.-L., Hector, A., Hönig, L., Hu, R.-Y., Klein, A.-M., Kühn, P., Liang, Y., Li, S., Michalski, S., Scherer-Lorenzen, M., Schmidt, K., Scholten, T., Schuldt, A., Shi, X., Tan, M.-Z., Tang, Z., Trogisch, S., Wang, Z., Welk, E., Wirth, C., Wubet, T., Xiang, W., Yu, M., Yu, X.-D., Zhang, J., Zhang, S., Zhang, N., Zhou, H.-Z., Zhu, C.-D., Zhu, L., Bruelheide, H., Ma, K., Niklaus, P.A., Schmid, B., 2018. Impacts of species richness on productivity in a large-scale subtropical forest experiment. Science 362, 80–83.

Huber, P., Metz, S., Unrein, F., Mayora, G., Sarmento, H., Devercelli, M., 2020. Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. The ISME Journal 14, 2951–2966.

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.

Kadmon, Ronen, Allouche, Omri, 2007. Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. The American Naturalist 170, 443–454.

Karger, D.N., Tuomisto, H., Amoroso, V.B., Darnaedi, D., Hidayat, A., Abrahamczyk, S., Kluge, J., Lehnert, M., Kessler, M., 2015. The importance of species pool size for community composition. Ecography 38, 1243–1253.

Klaus, B., Strimmer, K., 2012. Estimation of (local) false discovery rates and higher criticism. fdrtool. R package version 1, 15.

Kneitel, J.M., Miller, T.E., 2003. Dispersal rates affect species composition in metacommunities of Sarracenia purpurea inquilines. The American Naturalist 162, 165–171.

Kumar, M., Prasad, R., Goyal, P., Teotia, P., Tuteja, N., Varma, A., Kumar, V., 2017. Environmental biodegradation of xenobiotics: role of potential microflora. In: Hashmi, M.Z., Kumar, V., Varma, A. (Eds.), Xenobiotics in the Soil Environment: Monitoring, Toxicity and Management. Springer International Publishing, Cham, pp. 319–334.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., 2004a. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7, 601–613.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F., Holt, R.D., Shurin, J.B., Law, R., Tilman, D., Loreau, M., Gonzalez, A., 2004b. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7, 601–613.

Lever, M.A., 2016. A new era of methanogenesis research. Trends in Microbiology 24, 84–86.

Loreau, M., Mouquet, N., 1999. Immigration and the maintenance of local species diversity. The American Naturalist 154, 427–440.

Loreau, M., Barbier, M., Filotas, E., Gravel, D., Isbell, F., Miller, S.J., Montoya, J.M., Wang, S., Aussenac, R., Germain, R., Thompson, P.L., Gonzalez, A., Dee, L.E., 2021. Biodiversity as insurance: from concept to measurement and application. Biological Reviews 96, 2333–2354.

Luan, L., Jiang, Y., Cheng, M., Dini-Andreote, F., Sui, Y., Xu, Q., Geisen, S., Sun, B., 2020. Organism body size structures the soil microbial and nematode community assembly at a continental and global scale. Nature Communications 11, 6406.

Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., Brookes, P.C., Xu, J., Gilbert, J.A., 2016. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. The ISME Journal 10, 1891–1901.

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., Øvreås, L., Reysenbach, A.-L., Smith, V.H., Staley, J.T., 2006. Microbial biogeography: putting microorganisms on the map. Nature Reviews Microbiology 4, 102–112.

Maver, M., Escudero-Martinez, C., Abbott, J., Morris, J., Hedley, P.E., Mimmo, T., Bulgarelli, D., 2021. Applications of the indole-alkaloid gramine modulate the assembly of individual members of the barley rhizosphere microbiota. PeerJ 9, e12498.

Mei, R., Liu, W.-T., 2019. Quantifying the contribution of microbial immigration in engineered water systems. Microbiome 7, 1–8.

Morrison-Whittle, P., Goddard, M.R., 2015. Quantifying the relative roles of selective and neutral processes in defining eukaryotic microbial communities. The ISME Journal 9, 2003–2011.

Mouquet, N., Loreau, M., 2002. Coexistence in metacommunities: the regional similarity hypothesis. The American Naturalist 159, 420–426.

Myers, J.A., Harms, K.E., 2009. Seed arrival, ecological filters, and plant species richness: a meta-analysis. Ecology Letters 12, 1250–1260.

Newman, M.E., 2003. The structure and function of complex networks. SIAM Review 45, 167–256.

Newman, M.E., 2006. Modularity and Community Structure in Networks, vol. 103. Proceedings of the National Academy of Sciences, pp. 8577–8582.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'hara, R., Simpson, G., Solymos, P., Stevens, M., Wagner, H., 2015. Vegan: community ecology package. R package vegan, vers. 2.2-1. Worl. Agro. At the Centre 3, 7–81.

Oliverio, A.M., Geisen, S., Delgado-Baquerizo, M., Maestre, F.T., Turner, B.L., Fierer, N., 2020. The global-scale distributions of soil protists and their contributions to belowground systems. Science Advances 6, eaax8787.

Ponisio, L.C., M'Gonigle, L.K., Kremen, C., 2016. On-farm habitat restoration counters biotic homogenization in intensively managed agriculture. Global Change Biology 22, 704–715.

Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular Biology and Evolution 26, 1641–1650.

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, D590–D596.

Rillig, M.C., Ryo, M., Lehmann, A., Aguilar-Trigueros, C.A., Buchert, S., Wulf, A., Iwasaki, A., Roy, J., Yang, G., 2019. The role of multiple global change factors in driving soil functions and microbial biodiversity. Science 366, 886–890.

Shi, Y., Li, Y., Xiang, X., Sun, R., Yang, T., He, D., Zhang, K., Ni, Y., Zhu, Y.-G., Adams, J. M., Chu, H., 2018. Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. Microbiome 6, 27.

Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., Curtis, T.P., 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environmental Microbiology 8, 732–740.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D.R., Alberti, A., 2015. Structure and function of the global ocean microbiome. Science 348, 1261359.

Tamme, R., Hiiesalu, I., Laanisto, L., Szava-Kovats, R., Pärtel, M., 2010. Environmental heterogeneity, species diversity and co-existence at different spatial scales. Journal of Vegetation Science 21, 796–801.

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., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551, 457–463.

Thuiller, W., Lavorel, S., Araújo, M.B., 2005. Niche properties and geographical extent as predictors of species sensitivity to climate change. Global Ecology and Biogeography 14, 347–357.

Vályi, K., Mardhiah, U., Rillig, M.C., Hempel, S., 2016. Community assembly and coexistence in communities of arbuscular mycorrhizal fungi. The ISME Journal 10, 2341–2351.

Vannette, R.L., Fukami, T., 2017. Dispersal enhances beta diversity in nectar microbes. Ecology Letters 20, 901–910.

Vellend, M., 2010. Conceptual synthesis in community ecology. The Quarterly review of biology 85, 183–206.

Vellend, M., 2016. Princeton University Press, Princeton.

Walters, K.E., Capocchi, J.K., Albright, M.B.N., Hao, Z., Brodie, E.L., Martiny, J.B.H., 2022. Routes and rates of bacterial dispersal impact surface soil microbiome composition and functioning. The ISME Journal 16, 2295–2304.

Wang, S., Wang, X., Han, X., Deng, Y., 2018. Higher precipitation strengthens the microbial interactions in semi-arid grassland soils. Global Ecology and Biogeography 27, 570–580.

Wang, Y., Li, C., Tu, B., Kou, Y., Li, X., 2021. Species pool and local ecological assembly processes shape the β-diversity of diazotrophs in grassland soils. Soil Biology and Biochemistry 160, 108338.

Wilson, D.S., 1992. Complex interactions in metacommunities, with implications for biodiversity and higher levels of selection. Ecology 73, 1984–2000.

Xiong, C., Zhu, Y.-G., Wang, J.-T., Singh, B., Han, L.-L., Shen, J.-P., Li, P.-P., Wang, G.-B., Wu, C.-F., Ge, A.-H., Zhang, L.-M., He, J.-Z., 2021. Host selection shapes crop microbiome assembly and network complexity. New Phytologist 229, 1091–1104.

Xu, L., He, N., Li, X., Cao, H., Li, C., Wang, R., Wang, C., Yao, M., Zhou, S., Wang, J., 2021. Local community assembly processes shape β-diversity of soil phoD-harbouring communities in the Northern Hemisphere steppes. Global Ecology and Biogeography 30, 2273–2285.

Xun, W., Liu, Y., Li, W., Ren, Y., Xiong, W., Xu, Z., Zhang, N., Miao, Y., Shen, Q., Zhang, R., 2021. Specialized metabolic functions of keystone taxa sustain soil microbiome stability. Microbiome 9, 35.

Yang, G., Ryo, M., Roy, J., Lammel, D.R., Ballhausen, M.-B., Jing, X., Zhu, X., Rillig, M.C., 2022. Multiple anthropogenic pressures eliminate the effects of soil microbial diversity on ecosystem functions in experimental microcosms. Nature Communications 13, 4260.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., Glöckner, F.O., 2014. The SILVA and "all-species living tree project (LTP)" taxonomic frameworks. Nucleic Acids Research 42, D643–D648.

Yu, G., Smith, D.K., Zhu, H., Guan, Y., Lam, T.T.-Y., 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8, 28–36.

Yuan, M.M., Guo, X., Wu, L., Zhang, Y., Xiao, N., Ning, D., Shi, Z., Zhou, X., Wu, L., Yang, Y., Tiedje, J.M., Zhou, J., 2021. Climate warming enhances microbial network complexity and stability. Nature Climate Change 11, 343–348.

Zhang, X., Liu, S., Wang, J., Huang, Y., Freedman, Z., Fu, S., Liu, K., Wang, H., Li, X., Yao, M., Liu, X., Schuler, J., 2020. Local community assembly mechanisms shape soil bacterial β diversity patterns along a latitudinal gradient. Nature Communications 11, 5428.

Zhou, J., Ning, D., 2017. Stochastic community assembly: does it matter in microbial ecology? Microbiology and Molecular Biology Reviews 81, e00002, 00017.

Zhu, G., Zhou, L., Wang, Y., Wang, S., Guo, J., Long, X.-E., Sun, X., Jiang, B., Hou, Q., Jetten, M.S.M., Yin, C., 2015. Biogeographical distribution of denitrifying anaerobic methane oxidizing bacteria in Chinese wetland ecosystems. Environmental Microbiology Reports 7, 128–138.

Zobel, M., 1997. The relative of species pools in determining plant species richness: an alternative explanation of species coexistence? Trends in Ecology & Evolution 12, 266–269.

Zomer, R.J., Trabucco, A., Bossio, D.A., Verchot, L.V., 2008. Climate change mitigation: a spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. Agriculture, Ecosystems & Environment 126, 67–80.