

Soil acidification mediates changes in soil bacterial community assembly processes in response to agricultural intensification

Xingang Zhou,^{1,2} Muhammad Khushi u Rahman,²
Junjie Liu^{3*} and Fengzhi Wu^{1,2*}

¹Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (Northeast Region), Ministry of Agriculture and Rural Affairs, Northeast Agricultural University, Harbin, 150030, China.

²Department of Horticulture, Northeast Agricultural University, Harbin, 150030, China.

³Key Laboratory of Mollisols Agroecology, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin, 150081, China.

Summary

Agricultural intensification is known to alter the assembly of soil microbial communities, which regulate several critical ecosystem processes. However, the underlying ecological processes driving changes in microbial community assembly, particularly at the regional scale, remain poorly understood. Using 16S rDNA sequencing, we characterized soil bacterial community assembly in three land-use types with increasing land-use intensity: open fields cultivated with main crops (CF) or vegetables (VF), and greenhouses cultivated with vegetables (VG). Compared with CF, VF and VG altered bacterial community composition and decreased spatial turnover rates of edaphic variables and bacterial communities. Bacterial community assembly was primarily governed by deterministic processes; however, bacterial communities in VF and VG were phylogenetically less clustered and more influenced by variable selection and less by dispersal limitation. Soil pH was the most important edaphic variable mediating the changes in bacterial community assembly processes induced by agricultural intensification. Specifically, decreasing soil pH led to stochastic assembly of bacterial community. Soil pH was lower in more intensively

managed lands, especially in case of VG (pH range: 5.86–7.42). Overall, agricultural intensification altered soil bacterial community assembly processes, which was associated with soil acidification. These findings may have implications for improving soil quality and agroecosystem sustainability.

Introduction

Soils harbour considerably diverse microbial communities that are essential players in terrestrial ecosystem functioning, such as organic material decomposition, nutrient cycling and plant growth (Bardgett and van der Putten, 2014). Therefore, maintaining the diversity of soil microbial communities is critical for crop production and agroecosystem sustainability (Geisen *et al.*, 2019). The activity and diversity of soil microbial communities are sensitive to changes in soil characteristics, such as soil pH, moisture, carbon content and nutrient availability (Chen *et al.*, 2017; Feng *et al.*, 2017; Barnett *et al.*, 2020). Land-use change due to agricultural intensification is one of the most significant anthropogenic activities that greatly affect soil microbial communities by altering edaphic variables (Geisen *et al.*, 2019). Studies have shown that agricultural intensification can be detrimental to the diversity and functioning of soil microbial communities (Tsiafouli *et al.*, 2015; de Graaff *et al.*, 2019). However, the ecological processes driving these changes in microbial community assembly, especially at the regional scale, are largely unexplored.

Deterministic and stochastic processes are two fundamental types of processes affecting the assembly of species into a local community (Chase, 2007; Zhou and Ning, 2017). On the one hand, the niche theory emphasizes that deterministic processes, such as environmental filtering and biotic interactions (e.g. mutualisms, facilitation, competition and predation) control the local community composition. In contrast, the neutral theory stresses that stochastic processes, such as pervasive dispersal limitation and ecological drift (random changes in the relative abundance of species), can explain a wide variety of biodiversity patterns in nature (Chase, 2007;

Received 24 April, 2021; revised 14 July, 2021; accepted 16 July, 2021. *For correspondence. E-mail liujunjie@iga.ac.cn; Tel. +86 451 86602737; Fax +86 451 86602737. E-mail fzwunewau@yahoo.com; Tel. +86 451 55190215; Fax +86 451 55190443.

Zhou and Ning, 2017). Soil microbial communities are reported to exhibit spatial biogeographic patterns, which is a greater dissimilarity in soil microbial community composition with increasing geographic distance (referred to as a distance-decay relationship) (Hanson *et al.*, 2012; Powell *et al.*, 2015; Jiao and Lu, 2020). This microbial biogeographic pattern can be driven by both stochastic and deterministic processes (Hanson *et al.*, 2012). For example, selection and drift can accelerate the spatial turnover rate while high dispersal rate can decelerate it (Hanson *et al.*, 2012). Anthropogenic activities, such as deforestation, fertilization and agricultural management regimes (organic versus conventional managements) have been reported to alter the soil microbial biogeography (Rodrigues *et al.*, 2013). Possibly, agricultural intensification may decrease the soil environmental heterogeneity along the spatial gradient, thereby, decreasing the spatial turnover rate of soil microbial communities.

The balance between stochastic and deterministic assembly processes is mediated by various environmental factors, such as soil pH (Tripathi *et al.*, 2018; Jiao and Lu, 2020), soil organic matter content (Zhou *et al.*, 2014; Feng *et al.*, 2018) and soil nutrient content (Feng *et al.*, 2017). For example, resource supply can increase the importance of stochastic processes (Zhou *et al.*, 2014; Feng *et al.*, 2018). Recently, soil pH has been shown to serve as a major mediator of the balance between stochastic and deterministic assembly processes of soil bacterial communities (Tripathi *et al.*, 2018; Jiao and Lu, 2020). However, contrasting results have been reported in the literature. For example, Tripathi *et al.* (2018) showed that stochastic assembly processes are predominant in soils with pH close to neutral, and a shift in pH towards relatively extreme conditions (acidic or alkaline) leads to a more deterministic assembly. However, other studies have shown an increased role of stochasticity in more acidic environments (Ren *et al.*, 2015; Barnett *et al.*, 2020; Jiao and Lu, 2020). Intensified agroecosystems are usually subjected to high levels of fertilizer application, and thus leading to soil acidification (Guo *et al.*, 2010; Hu *et al.*, 2017). However, it is unclear whether soil acidification caused by agricultural intensification is linked to the changes in the balance between stochastic and deterministic assembly processes.

Over the past half-century, large areas of open fields for main crop production (such as legumes and cereals) have been converted to greenhouses to meet the demand for vegetables year-round. For example, in China, there are 3.86 million hectares of greenhouses for vegetable production, accounting for approximately 2.86% of the farmland area (Shen *et al.*, 2021). Compared with main crop cultivation, vegetable cultivation

usually receives higher quantities of manure and fertilizer (Hu *et al.*, 2017). Since more cropping of vegetables are cultivated in the greenhouse per year, greenhouse production is also characterized by more intensive agricultural management practices, such as irrigation and tillage operations (Boulard *et al.*, 2011). Therefore, among the three land-use types of open fields cultivated with main crops (CF), open fields cultivated with vegetables (VF) and greenhouses cultivated with vegetables (VG), the land-use intensity is the highest for VG and lowest for CF. Since highly intensified agricultural systems are considered to be unsustainable in the long term (Tsiafouli *et al.*, 2015; de Graaff *et al.*, 2019; Shen *et al.*, 2021), identification of the factors that govern changes in the microbial assembly processes induced by agricultural intensification can help to enhance agroecosystem sustainability.

In this study, we aimed to address the following questions: (i) Whether and how agricultural intensification affect the spatial turnover rate and assembly processes of soil bacterial communities? (ii) Is soil acidification an important factor regulating the assembly processes of soil bacterial communities in response to agricultural intensification? To answer these questions, we evaluated the abundance, composition, geographic pattern and assembly processes of soil bacterial communities in agricultural fields with three different land-use types. We analysed 288 soil samples collected from 32 sites across China. Each site comprised adjacent triplets of three land-use types (i.e. CF, VF and VG). We hypothesized that vegetable production in greenhouses could (i) decrease the level of spatially structured environmental gradients, and therefore lead to a weaker distance-decay pattern of soil bacterial communities, and (ii) alter the relative importance of stochastic and deterministic processes in driving soil bacterial community assembly by altering edaphic variables, especially soil pH.

Experimental procedures

Soil sampling and edaphic variable analysis

Soil samples were collected in July 2017 from 32 sites located in 12 provinces of China (Fig. S1a). The maximum distance between the sampling sites was 3120 km. Each site comprised agricultural fields with three different land-use types (i.e. CF, VF and VG) (Fig. S1b–e). Within each site, fields with different land-use types were adjacent to each other (<1 km apart). For CF and VF, the fields had been cultivated with main crops or vegetables respectively, for more than 20 years. For VG, open fields had been converted to greenhouses and cultivated with vegetables for more than 3 years (Table S1). Of note, all greenhouses were used for soil-based vegetable

production; in other words, vegetables were cultivated in the soil, and not via soilless production. As different crops were cultivated at our sampling sites (Table S1) and as plant roots can strongly affect the soil microbial community composition (Bardgett and van der Putten, 2014), bulk soils were sampled in this study. For each land-use type at each site, three plots, each with an area of approximately 50 m², were randomly selected, and 10 soil cores per plot were taken from the upper soil layer (0–15 cm) and pooled. Therefore, there were three composite samples for each land-use type at each site. In total, 288 soil samples were collected. Soils were sieved (2 mm), and large stones and plant debris were removed. A part of these sampled soils was used for edaphic variable analyses (Supplementary Methods) and the other part was stored at –80 °C for DNA extraction.

Soil DNA extraction

Soil DNA was extracted from 0.25 g of soil with the Power Soil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, USA) according to the manufacturer's instructions. The quality of extracted DNA was evaluated using electrophoresis in a 1.2% (wt./vol.) agarose gel and a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). Each composite soil sample was extracted in triplicate, and the extracted DNA solutions were pooled.

Quantitative PCR analysis

SYBR Green quantitative PCR assays were performed to estimate soil bacterial abundance using the primer set 338F/518R, which targets the V3 region of the bacterial 16S rRNA gene (Muyzer *et al.*, 1993). Quantitative PCR assays were conducted using a qTOWER 3G touch real-time PCR system (Analytik Jena, Jena, Germany) (the detailed PCR conditions are provided in the Supplementary Methods). All amplifications were performed in triplicate.

High-throughput amplicon sequencing and data processing

High-throughput amplicon sequencing was performed to profile the diversity and composition of soil bacterial communities. Extracted soil DNA was amplified using the primer set F338/R806, targeting the V3–V4 regions of the bacterial 16S rRNA gene (Derakhshani *et al.*, 2016). Both forward and reverse primers were tagged with specific overhang Illumina adapters. The detailed PCR conditions and library preparation protocols are provided in the Supplementary Methods. The libraries were

sequenced (2 × 300) on an Illumina MiSeq platform (Illumina, San Diego, USA).

Raw sequencing data were analysed using the QIIME pipeline (Caporaso *et al.*, 2010). Briefly, adaptor sequences, barcodes and 30 low-quality bases at the end of each read were removed. Paired reads were joined (minimum overlapping read length of 20 base pairs) and quality filtered (Phred score of 20), and reads with less than 200 base pairs were removed. Chimaeras were screened and removed using USEARCH with the UCHIME algorithm (Edgar, 2013). Sequences with ≥97% similarity were assigned to the same operational taxonomic unit (OTU) using UPARSE with an agglomerative clustering algorithm (Edgar, 2013). Taxonomies were assigned to the representative sequence of each OTU using the SILVA database release 128 (Quast *et al.*, 2013). OTUs classified as chloroplasts and mitochondria, and singleton OTUs were eliminated. To avoid potential bias caused by the sequencing depth, sequence counts of all samples were normalized to the minimum number of sequence (10 763) per sample. The sequences data were deposited in the Sequence Read Archive at NCBI under the accession numbers PRJNA660593, PRJNA660598, PRJNA660802 and PRJNA660808.

Statistical analyses

Bacterial community α - and β -diversity analyses. Unless stated otherwise, all statistical analyses were conducted using 'R' (v3.6.2, <http://www.r-project.org/>). Taxon accumulation curves and Good's coverage were used to evaluate the sampling effort of the sequencing data. Rarefaction curves of the observed OTUs and Good's coverage indicated that a large proportion of the bacterial diversity was covered by sequencing (Fig. S2a,b). Bacterial community α -diversity indices, including Shannon index and Faith's phylogenetic diversity, were calculated using the 'vegan' (Oksanen *et al.*, 2019) and 'picante' (Kembel *et al.*, 2010) packages.

For the analysis of bacterial community β -diversity, taxonomy-based Bray–Curtis distance and phylogeny-based β -mean nearest taxon distance (β MNTD) were calculated using the *vegdist* function in the 'vegan' (Oksanen *et al.*, 2019) and the *comdistnt* function in the 'picante' packages (Kembel *et al.*, 2010) respectively. Variation in bacterial community structure was visualized using principal coordinates analysis (PCoA). The proportion of variation in the community structure explained by land-use type was quantified using constrained analysis of principal coordinates (CAP), and statistical significance was tested using the *permutest* function with 9999 permutations in the 'vegan' package (Oksanen *et al.*, 2019). To test the effects of sampling sites and land-use types

on community β -diversity, permutational analysis of variance (PERMANOVA) was performed using the *adonis* function with 9999 permutations in the 'vegan' package (Oksanen *et al.*, 2019).

Identification of land-use type sensitive OTUs. Land-use type sensitive OTUs were identified according to the method described by Hartman *et al.* (2018). First, indicator species analysis was performed to calculate the point-biserial correlation coefficient of an OTU's positive association with one or a combination of land-use types. This analysis was performed using the *multipatt* function with 9999 permutations in the 'indicspecies' package (De Caceres *et al.*, 2010). Next, the likelihood ratio test was performed to evaluate the differential OTU abundance among land-use types with the 'edgeR' package (Robinson *et al.*, 2010) and the *P*-values were adjusted with the Benjamini–Hochberg (BH) method. Lastly, sensitive OTUs were defined as these validated by both indicator species analysis and likelihood ratio test at $P < 0.05$. A maximum-likelihood tree was constructed based on representative sequences for each sensitive OTU and was visualized using the iTOL tool.

Correlation network and cohesion analyses. Correlation networks were constructed for all samples and for each land-use type based on the relative abundances of OTUs. Spearman's correlations between OTUs were calculated using the *rcorr* function in the 'Hmisc' package (Harrell Jr and Dupont, 2020). To reduce rare OTUs in each dataset, OTUs with relative abundance less than 0.5% were removed. A correlation coefficient was considered statistically robust if the Spearman's correlation coefficient (ρ) was >0.6 and the BH-adjusted *P*-value was <0.01 . The Gephi software was used to visualize the network graphs. The nodes in the networks represent bacterial OTUs, and the edges represent strong and significant correlations among the nodes. Some key topological features of the networks (including average connectivity, average path length, clustering coefficient, network density and modularity) and topological features of nodes (including degree and betweenness centrality) were calculated using the 'igraph' package (Csardi and Nepusz, 2006). The possible keystone OTUs were those that had low betweenness centrality and high degree values (within the lowest 5% of betweenness centrality and top 5% of node degree values) (Berry and Widder, 2014; Hartman *et al.*, 2018). To determine whether the constructed networks were nonrandom, 1000 Erdős–Rényi random networks with the same number of nodes and edges as that in the observed networks were constructed, and the topological features of the real and random networks were compared. Differences in the correlation network structure were determined using

PERMANOVA analysis of the Spearman's correlation distance matrix (Williams *et al.*, 2014).

Community connectivity was measured using the cohesion metric according to the protocol reported by Herren and McMahon (2017). Briefly, pairwise correlations were calculated between all OTUs in a community. Then, a null model was used to verify the strength of these correlations. Expected correlations generated using the null model were subtracted from the observed correlations, yielding positive and negative connectedness values for each OTU. Finally, cohesion metrics were determined for each community by calculating the sum of each OTU's connectedness value multiplied by its abundance in the community, yielding a positive and negative cohesion metric for each community.

Analyses of distance–decay relationship and factors affecting β -diversity. Geographical distances among sampling sites were calculated from the sampling coordinates. The spatial turnover rates (i.e. the distance decay of similarity) of bacterial community and edaphic variable were determined by regressing the pairwise community similarity (1-dissimilarity of the Bray–Curtis or β MNTD metrics) or edaphic variable similarity respectively, against the pairwise geographic distance using ordinary least-squares linear regressions. Edaphic variable similarity was calculated based on the edaphic variables (Powell *et al.*, 2015): $Ed = (1 - Euc_d/Euc_{max})$, where Euc_d is the Euclidean distance between two sites and Euc_{max} is the maximum Euclidean distance between sites in the distance matrix. For edaphic variables (except pH), the proportion data were arcsine(sqrt)-transformed, and others were log-transformed. The differences of slopes among land-use types were evaluated using the *diffslope* function with 9999 permutations in the 'simba' package (Jurasinski and Jurasinski, 2012).

The relationship between bacterial community similarity, geographic distance and environmental similarity was evaluated with the partial Mantel test using the *mantel* function in the 'ecodist' package (Goslee and Urban, 2007). The environmental distance matrix was created with non-redundant edaphic variables that best explaining bacterial community dissimilarities, which were identified using the *bioenv* function in the 'vegan' package (Oksanen *et al.*, 2019). To tease apart the relative impact of land-use type, crop type, and edaphic and geographic variables on bacterial community β -diversity, multiple regressions on matrices (MRM) was performed with Spearman's correlation and 9999 permutations using the 'ecodist' package (Goslee and Urban, 2007). Land-use type and crop type were transformed into categorical variables. Before applying MRM, variable clustering was performed to remove redundant environmental variables with the Spearman's correlation method (cutoff: $\rho < 0.6$)

using the *varclus* function in the ‘Hmisc’ package (Harrell Jr and Dupont, 2020). The partial regression coefficients (*b*) of the MRM model provided a measure of the per-unit effect of an independent variable on the dependent variable (bacterial community β -diversity) while controlling for the effects of the other independent variables. Moreover, a distance-based redundancy analysis with forward selection of edaphic variables was performed using the *capscale* and *ordR2step* functions in the ‘vegan’ package (Oksanen *et al.*, 2019).

Ecological modelling. Phylogenetic Mantel correlograms analysis was performed to evaluate the relationship between OTUs’ niche differences and phylogenetic distances. The distances between the OTUs’ environmental optima were calculated using the abundance-weighted mean approach (Stegen *et al.*, 2013). The phylogenetic distance matrix was generated using the *cophenetic.phylo* function in the ‘ape’ package (Paradis *et al.*, 2004). The correlation between phylogenetic distance and the difference in environmental optima was estimated using the *mantel.correlog* function with 999 permutations and BH correction in the ‘vegan’ package (Oksanen *et al.*, 2019). A phylogenetic signal was detected when the OTUs’ niche differences were significantly related to the between-OTU phylogenetic distances.

The standardized effect size measure of the mean nearest taxon distance (SES.MNTD) was calculated using the *ses.mntd* function with the ‘taxa.labels’ null model and 999 randomizations in the ‘picante’ package (Kembel *et al.*, 2010). SES.MNTD values greater or less than zero indicate that communities are more distantly or closely phylogenetic clustered respectively, than expected by chance (Tripathi *et al.*, 2018). To quantify the potential ecological processes (variable selection, homogeneous selection, dispersal limitation, homogenizing dispersal, or drift) that govern bacterial community assembly, the null model method proposed by Stegen *et al.* (2013) was used. The β -nearest taxon index (β NTI) and Bray–Curtis-based Raup–Crick (RCbray) metrics, which were based on the difference between the observed β MNTD and Bray–Curtis dissimilarities and their null distributions (999 community randomizations) respectively, were calculated using R scripts written by Stegen *et al.* (2013). For a given community, $|\beta$ NTI| >2 and <2 indicated the dominance of deterministic and stochastic processes respectively. Furthermore, the relative influences of variable or homogeneous selection were quantified as the fraction of pairwise comparisons with β NTI >+2 or β NTI <−2 respectively. The relative influences of dispersal limitation or homogenizing dispersal were quantified as the percentage of pairwise comparisons with $−2 < \beta$ NTI < 2 and RCbray >0.95 or <−0.95

respectively. The undominated fraction was quantified as the percentage of pairwise comparisons with $−2 < \beta$ NTI < 2 and $−0.95 < \text{RCbray} < 0.95$.

To identify the primary factors that influenced the bacterial community assembly processes, the MRM test was performed with the geographic distance and Euclidean distance matrices of each of the edaphic variables as independent variables and with the β NTI matrix among samples as the dependent variable. Partial Mantel tests were performed to evaluate the relationship between β NTI and either of geographic distance, soil pH and edaphic variables excluding soil pH. To further assess the role of soil pH in driving the bacterial community assembly processes, the β NTI values were regressed against the Euclidean distance matrix of soil pH. The statistical significance of the resulting comparisons was determined using Mantel tests with 999 permutations in the ‘ecodist’ package (Goslee and Urban, 2007).

The Sloan neutral model, which assumes that community assembly is driven solely by chance and dispersal and ignores microbial community phylogenetic turnover, was used to determine the potential importance of neutral processes in community assembly based on the community abundance data (Sloan *et al.*, 2006). This model was fitted to the frequency of detection of OTUs and their abundances by the parameter *m* (migration rate), and the fit of the model was assessed using a generalized R^2 .

Two-way ANOVA was conducted to analyse the effects of sampling sites and land-use types on bacterial abundance, α -diversity indices, soil physicochemical properties, SES.MNTD and β NTI. In cases where the data did not satisfy the normality and homogeneity of variance, the values were arcsine(sqrt)-transformed or log-transformed before analysis. Means were compared between land-use types based on Tukey’s HSD test. Spearman’s correlations among bacterial abundance, bacterial community α -diversity indices, relative abundances of sensitive OTUs and soil physicochemical properties were tested using the *rcorr* function in the ‘Hmisc’ package (Harrell Jr and Dupont, 2020) and the *P*-values were adjusted using the BH method.

Results

Edaphic properties and bacterial abundance

Across all samples, both sampling sites and land-use types exerted significant effects on soil edaphic properties (two-way ANOVA, $P < 0.05$) (Table S2). Soil pH was lower in VF and VG than in CF, with VG having the lowest value (Tukey’s HSD test, $P < 0.05$) (Fig. S3a). A contrasting pattern was observed for soil electric conductivity, organic matter, dissolved organic carbon, ammonium-nitrogen, nitrate-nitrogen, Olsen phosphorus,

available potassium, total nitrogen and total phosphorus (Tukey's HSD test, $P < 0.05$).

Both sampling sites and land-use types affected soil bacterial abundance (Two-way ANOVA, $P < 0.05$) (Table S2). VG had higher soil bacterial abundance than CF and VF (Tukey's HSD test, $P < 0.05$) (Fig. 1A). Soil bacterial abundance was negatively correlated with soil pH but positively correlated with soil electric conductivity, organic matter, dissolved organic carbon and soil nutrient contents (Spearman's correlation, BH-corrected $P < 0.05$) (Fig. S3b).

Bacterial community α - and β -diversities

High-throughput amplicon sequencing generated 9 417 461 quality sequences from 288 soil samples. The dataset was composed of 20 488 bacterial OTUs grouped at 97% similarity. Both sampling sites and land-use types altered soil bacterial community α -diversity (two-way ANOVA, $P < 0.05$) (Table S2). The number of OTUs, Shannon index and Faith's phylogenetic diversity were lower in the VG than in CF and VF (Tukey's HSD,

$P < 0.05$) (Fig. 1A). Soil pH was positively correlated with the number of OTUs and Shannon index, whereas soil electric conductivity, Olsen phosphorus, nitrate- and ammonium-nitrogen were negatively correlated with the α -diversity indices (Spearman's correlation, BH-corrected $P < 0.05$) (Fig. S3b).

Bacterial community taxonomic and phylogenetic β -diversities varied widely across the sampling sites (Fig. 1B). In addition, CAP analysis, which was constrained by the land-use type, revealed a distinct bacterial community β -diversity across different land-use types (Fig. 1C). Land-use type explained a small but significant fraction of the total variation in β -diversity as measured using Bray–Curtis (5.4%; 95% CI = 3.9%, 7.5%; $P < 0.001$) and β MNTD distance dissimilarities (9.8%; 95% CI = 5.0%, 17.3%; $P < 0.001$). Furthermore, PERMANOVA analysis showed that land-use type exerted a significant effect on bacterial community taxonomic and phylogenetic β -diversities ($P < 0.05$) (Table S3). Pairwise PERMANOVA comparisons indicated significant differences among the taxonomic β -diversities of the three land-use types and between the phylogenetic β diversities of CF or VF and VG ($P < 0.05$) (Table S3). Moreover,

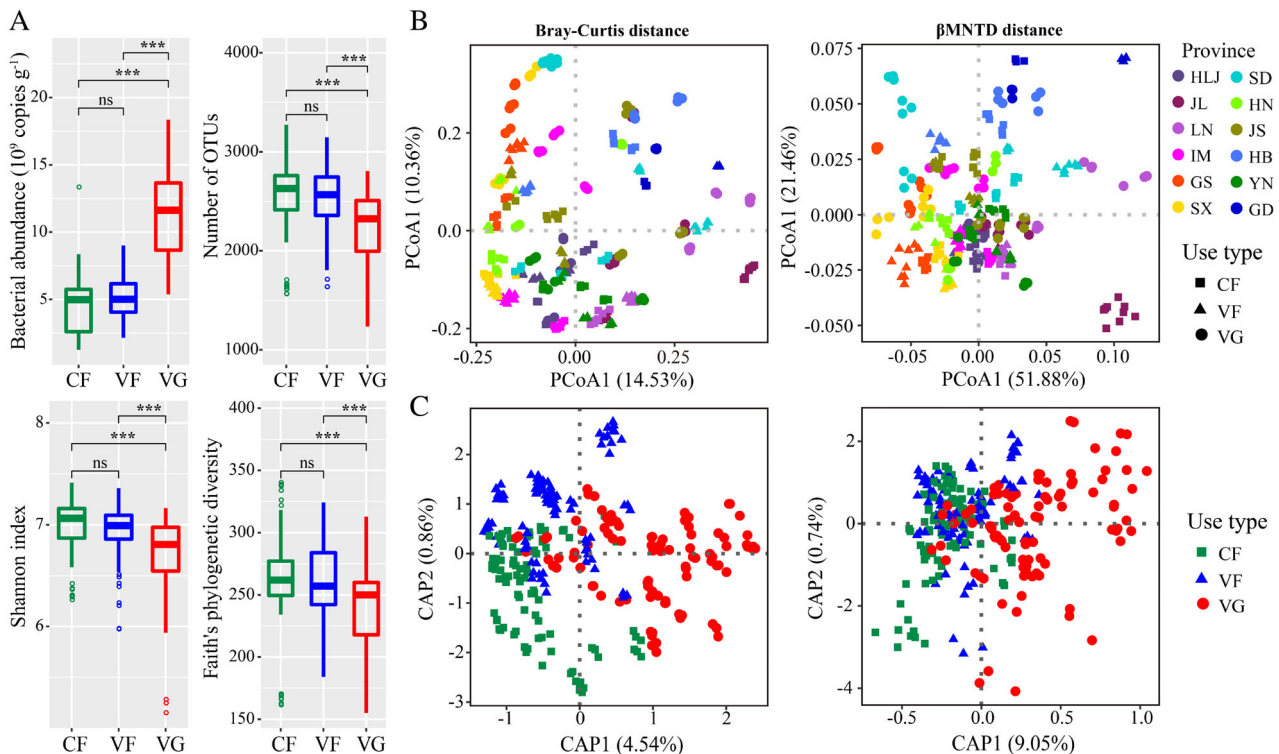


Fig. 1. Abundance, α - and β -diversities of soil bacterial communities.

A. Box plots showing differences in bacterial abundance (estimated using quantitative PCR assays) and α -diversity among land-use types. ns indicates no significant effect ($P > 0.05$); Triple asterisk indicates significant differences at $P < 0.001$ (Tukey's HSD test).

B. PCoA and (C) CAP analysis of bacterial community β -diversities based on the Bray–Curtis and β MNTD distance dissimilarities. HLJ, Heilongjiang; JL, Jilin; LN, Liaoning; IN, Inner Mongolia; GS, Gansu; SX, Shanxi; SD, Shandong; HN, Henan; JS, Jiangsu; HB, Hubei; YN, Yunnan; GD, Guangdong. CF, open field cultivated with main crops; VF, open field cultivated with vegetables; VG, greenhouse cultivated with vegetables. [Color figure can be viewed at wileyonlinelibrary.com]

the taxonomic and phylogenetic similarities were lower in VG than in other treatments (Fig. S2c).

Factors that contribute to variation in soil bacterial community β -diversity

Overall, both soil bacterial community taxonomic and phylogenetic similarities decreased significantly with increasing geographic distance for all samples (Fig. S4) and the samples in each treatment (Fig. 2A; Fig. S5a).

Partial Mantel test revealed that geographic distance was significantly related to bacterial community similarities ($P < 0.01$) (Table S4). Moreover, bacterial community similarities in VG had a weaker correlation with geographic distance than those in CF and VF. The slopes of the distance–decay curves of VF and VG were shallower than that of CF, with VG having the lowest value ($P < 0.001$). The edaphic variable similarity decreased with the increasing of the geographic distance ($P < 0.05$), with VF and VG having shallower slopes of distance–

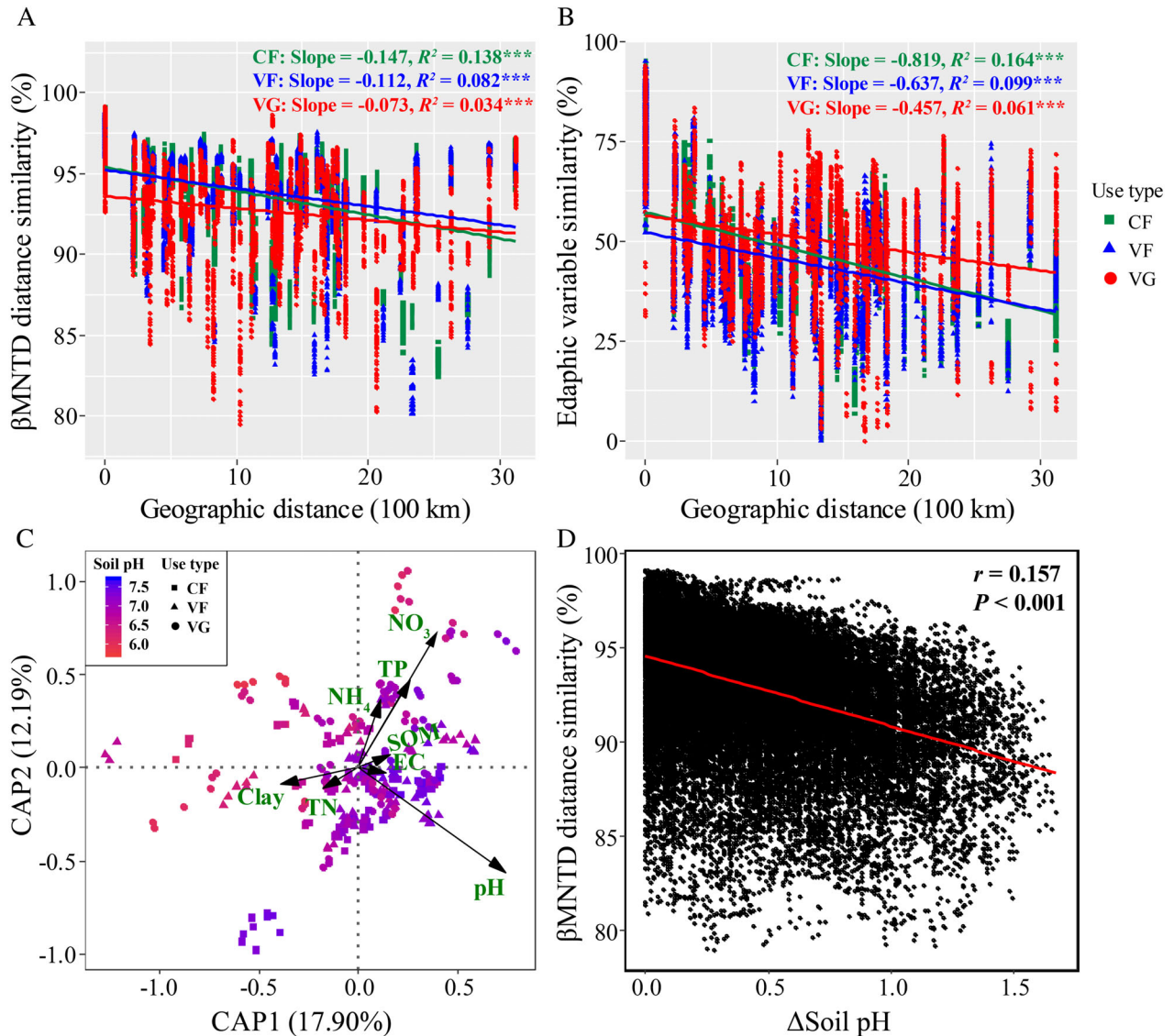


Fig. 2. Drivers of soil bacterial community phylogenetic β -diversity.

A. Distance–decay curves of bacterial community (based on the β MNTD distance) and (B) edaphic variable similarities. Solid lines represent the linear regression models. Triple asterisk indicates significant correlation at $P < 0.001$.

C. CAP analysis revealing edaphic variables influencing bacterial community phylogenetic β -diversity. Samples are coloured according to the soil pH. SOM, soil organic matter; TN, total nitrogen; EC, Soil electric conductivity; P, Olsen phosphorus; NH_4 , ammonium-nitrogen; NO_3 , nitrate-nitrogen; TP, total phosphorus.

D. Relationships between the β MNTD distance similarities and differences in soil pH. CF, open field cultivated with main crops; VF, open field cultivated with vegetables; VG, greenhouse cultivated with vegetables. [Color figure can be viewed at wileyonlinelibrary.com]

decay curves than CF ($P < 0.001$) (Fig. 2B). Moreover, the average between sample similarity of edaphic variables for VG was higher than that for CF and VF ($P < 0.01$) (Fig. S3c).

The MRM models, testing the combined effects of geographic distance, edaphic variables and land-use types, explained the majority of the variability in the similarities of the bacterial communities of all samples and each land-use type ($P < 0.001$) (Table S5). Land-use type had a weak yet significant influence on bacterial community taxonomic and phylogenetic β -diversities ($P < 0.05$). Geographic distance had smaller partial regression coefficients for VG than for CF and VF.

Edaphic variables contributed a larger proportion of variation than geographic distance to bacterial community β -diversity for all samples and each treatment (Table S5). Among all edaphic variables tested, soil pH had the greatest partial regression coefficient. The CAP analysis also revealed that soil pH was the most important edaphic factor influencing bacterial community taxonomic and phylogenetic β -diversities (Fig. 2C; Fig. S5b). Moreover, bacterial community taxonomic and phylogenetic similarities were negatively correlated with differences in soil pH ($P < 0.001$) (Fig. 2D; Fig. S5c).

Land-use type sensitive-OTUs

In total, 46 bacterial phyla were observed across all samples with *Proteobacteria* (32.78%), *Actinobacteria* (17.88%), *Chloroflexi* (12.80%), *Acidobacteria* (11.99%) and *Bacteroidetes* (7.36%) as the dominant phyla (relative abundance higher than 5%) (Fig. S6). Based on indicator species analysis and likelihood ratio test, 274 OTUs were identified to be sensitive to land-use type, with the relative abundances of these sensitive OTUs accounting for 6.48% of the total sequences (Fig. 3; Fig. S7). The taxonomic patterns of these sensitive OTUs are described in the Supplemental Material. The relationship patterns between edaphic variables and the relative abundances of land-use sensitive OTUs varied among different types of sensitive OTUs (Fig. 3C). For example, soil pH was negatively correlated with most OTUs enriched solely in the VG (VG-sensitive OTUs) but positively correlated with most OTUs enriched in both CF and VF (CF_VF-sensitive OTUs) (Spearman's correlation, BH-corrected $P < 0.05$). However, the soil nitrate–nitrogen content was negatively correlated with most CF- and CF_VF-sensitive OTUs but positively correlated with most VG- and VF_VG-sensitive OTUs (Spearman's correlation, BH-corrected $P < 0.05$).

Bacterial association patterns

We constructed a meta correlation network for the bacterial communities of all samples and separate correlation networks for each treatment (Fig. 4A; Fig. S8a). The bacterial association patterns differed among treatments (PERMANOVA, *Pseudo-F* = 275.25, $P < 0.001$). The VG network contained a larger proportion of negative correlations but fewer number of edges, lower average connectivity, clustering coefficient and modularity than CF and VF networks (Table S6). Moreover, VG had lower positive cohesion and higher negative cohesion values than CF and VF (Fig. 4B). The meta correlation network comprised 195 land-use type sensitive OTUs (71% of the total number of sensitive OTUs) (Table S7). CF_VF-sensitive OTUs had a higher average degree, whereas VG-sensitive OTUs had a higher proportion of negative edges than other types of sensitive OTUs in the meta correlation network. Land-use type sensitive OTUs had low to medium node degrees and betweenness centralities (Fig. 4C; Fig. S8b). Moreover, only a few land-use type sensitive OTUs were identified as keystone OTUs in the correlation network (Table S7). CF_VF-sensitive OTUs tended to correlate more frequently among themselves than with other sensitive OTUs (Fig. 4D; Fig. S8c). Meanwhile, VG- and VF_VG-sensitive OTUs formed more connections among themselves than with other sensitive OTUs.

Ecological processes governing bacterial community assembly

The Mantel correlograms analysis revealed significant phylogenetic signals at relatively short phylogenetic distances (Fig. S9a). Therefore, SES.MNTD and β NTI were used to analyse the bacterial community assembly processes because these two metrics emphasize phylogenetic relationships across short phylogenetic distances (Stegen *et al.*, 2013). The SES.MNTD values were less than zero (Fig. 5A), suggesting that bacterial communities were more closely phylogenetic clustered than expected by chance (Stegen *et al.*, 2013). The β NTI values were less than -2 (Fig. 5A; Fig. S9b), indicating that the deterministic process played a major role in governing bacterial community assembly (Stegen *et al.*, 2013). Both SES.MNTD and β NTI values were significantly higher for bacterial communities of VG than for those of the other two treatments (Tukey's HSD test, $P < 0.05$).

By combining β NTI and RCbray metrics, we found that homogeneous selection was the most important process governing bacterial community assembly (Fig. 5B). The proportion of variable selection was lower, whereas the proportion of dispersal limitation was higher in VF

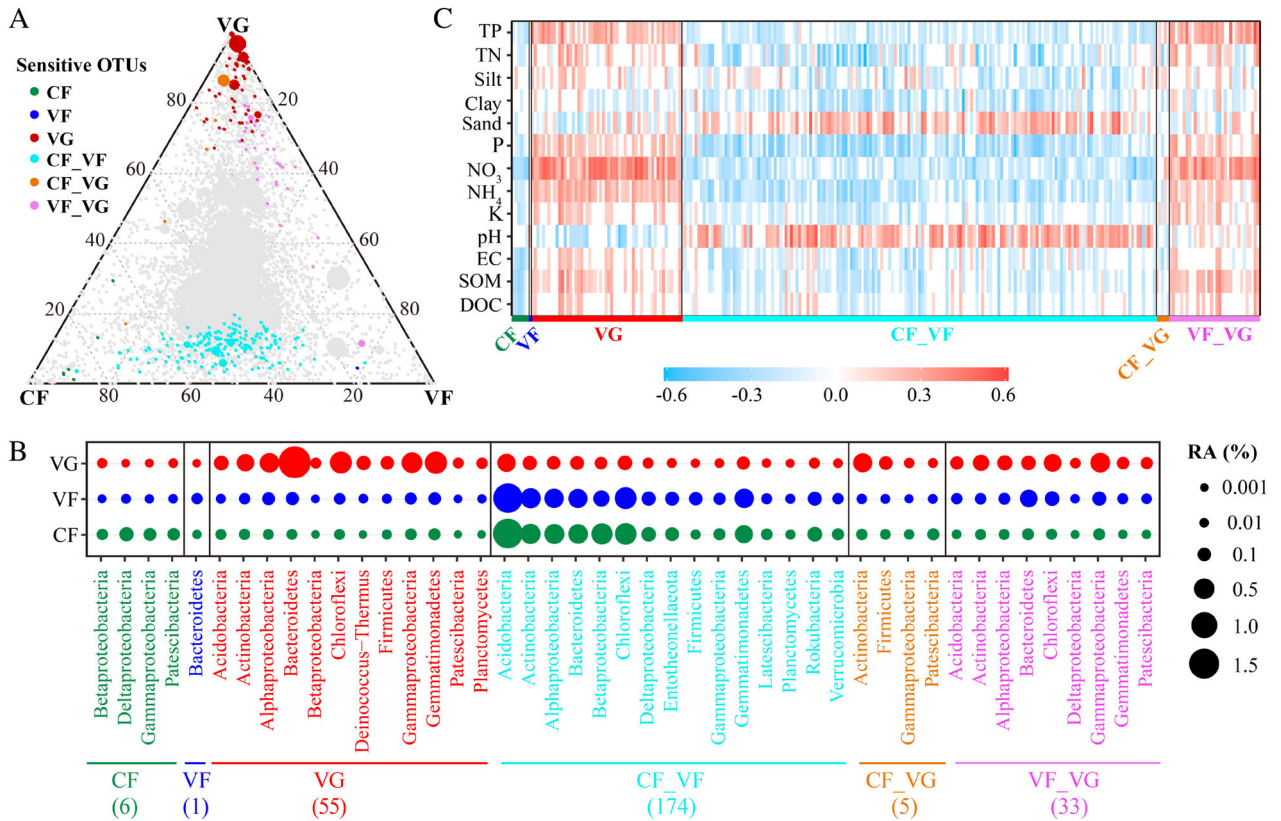


Fig. 3. Occurrence of land-use type sensitive OTUs and their relationship with edaphic variables.

A. Ternary plot showing relative abundances of land-use type sensitive OTUs. Each circle represents an OTU. Its position represents its relative abundance with respect to each land-use type, and its size is proportional to its mean relative abundance across all samples. Coloured circles represent sensitive OTUs and grey ones represent non-sensitive OTUs.

B. Relative abundances of sensitive OTUs summarized at the phylum/Proteobacteria class level. Numbers in the brackets indicate the number of sensitive OTUs.

C. Spearman's correlation between relative abundances of sensitive OTUs and edaphic variables. Only significant correlations (BH-corrected $P < 0.05$) are shown. DOC, dissolved organic carbon; SOM, soil organic matter; EC, soil electric conductivity; K, available potassium; NH₄, ammonium-nitrogen; NO₃, nitrate-nitrogen; P, Olsen phosphorus; TN, total nitrogen; TP, total phosphorus. CF, open field cultivated with main crops; VF, open field cultivated with vegetables; VG, greenhouse cultivated with vegetables. [Color figure can be viewed at wileyonlinelibrary.com]

and VG than in CF. The Sloan neutral model performed well and explained a large proportion of the variation in the metacommunity of all samples and the communities of each land-use type (Fig. S10). Moreover, the estimated migration rate (m) was relatively low in VF and VG than in CF, with VG having the lowest value.

Partial Mantel and MRM tests showed that among all measured edaphic variables and geographic distance, pH had the strongest association with β NTI ($P < 0.001$) (Table S4, S5). Moreover, SES.MNTD values were negatively correlated with soil pH ($P < 0.001$) (Fig. 6A; Fig. S11a). When soil samples were divided into different soil pH categories, the relative importance of homogeneous selection tended to increase with soil pH (Fig. 6B). Significant positive correlations were observed in the pairwise comparisons of β NTI values and differences in soil pH across all samples and within each treatment

($P < 0.001$) (Fig. 6C; Fig. S11b). When comparing between land-use types, differences in soil pH were also positively related to differences in β NTI ($P < 0.001$) (Fig. 6D).

Discussion

Agricultural intensification altered the spatial turnover rate and assembly processes of soil bacterial communities

Agricultural intensification is considered to pose a major threat to the global biodiversity (Tsiafouli *et al.*, 2015; Geisen *et al.*, 2019). We examined soil bacterial community assembly at a regional scale with respect to different land-use types and found that agricultural intensification altered the composition and decreased the spatial

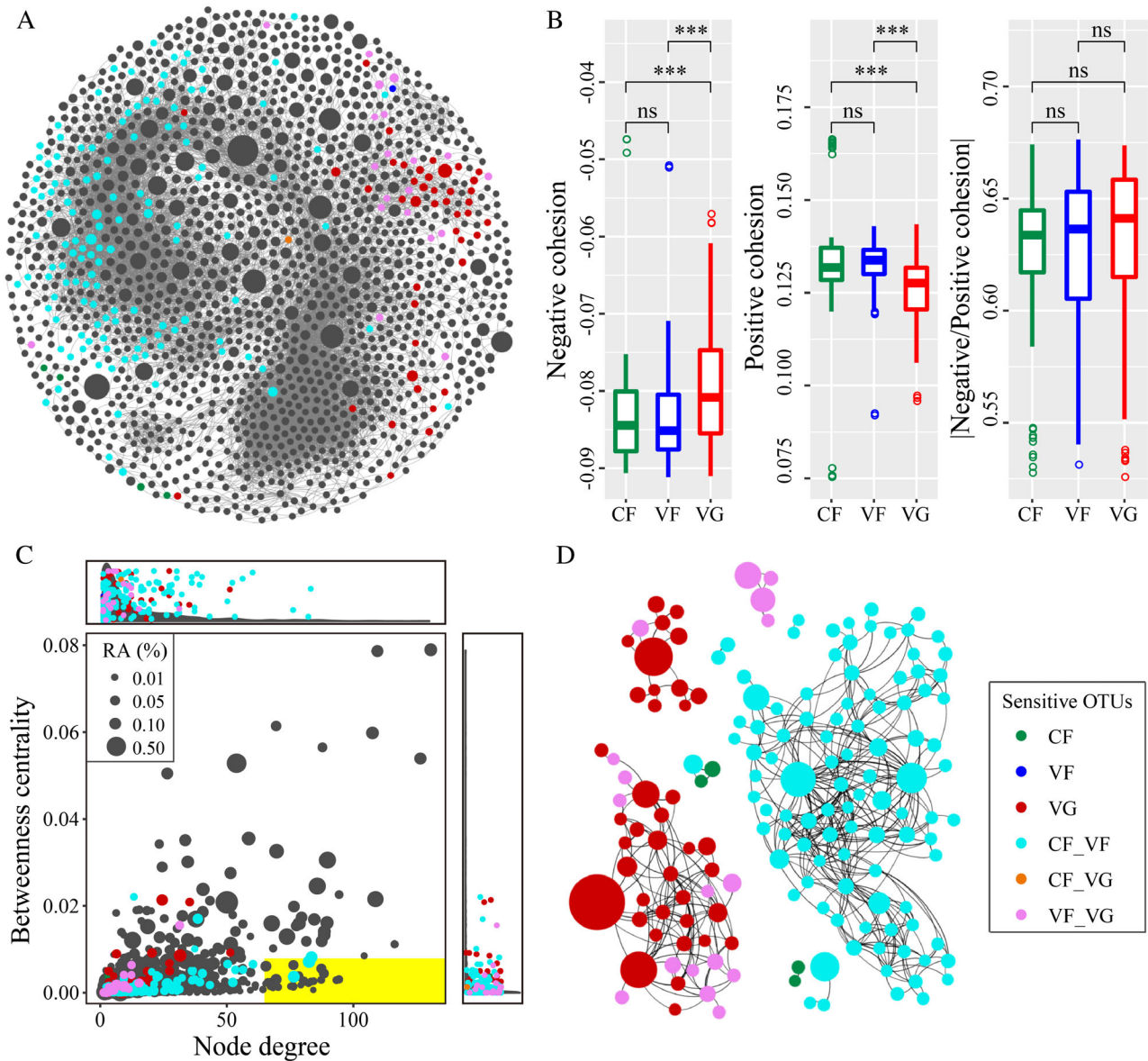


Fig. 4. Potential interspecies associations of soil bacterial communities.

A. The meta correlation network showing significant correlations ($\rho > 0.6$, BH-corrected $P < 0.001$) among OTUs with relative abundances above 0.5%. The size of each node is proportional to the relative abundance of the OTU; the edge thickness is proportional to the value of the Spearman's correlation coefficient.

B. Cohesion values for each land-use type. ns indicates no significant effect ($P > 0.05$); Triple asterisk indicates significant differences at $P < 0.001$ (Tukey's HSD test).

C. Degree-betweenness centrality plot of OTUs in the correlation networks. Keystone OTUs have yellow background. Side panels indicate the distributions of node degrees and betweenness centrality for sensitive OTUs compared to the density of all OTUs in the network.

D. The correlation network demonstrates correlations among sensitive OTUs, which was a sub-network of the meta correlation network. OTUs are coloured by their association to land-use types; the grey colour indicates the non-sensitive OTUs. CF, open field cultivated with main crops; VF, open field cultivated with vegetables; VG, greenhouse cultivated with vegetables. [Color figure can be viewed at wileyonlinelibrary.com]

turnover rate of soil bacterial communities. These results indicated that the bacterial community composition was more similar under more intensively managed conditions across spaces. A low spatial turnover rate of microbial communities can result from high dispersal rates and low environmental gradients (Hanson *et al.*, 2012). In this

study, the spatial turnover rate of edaphic variable similarity was lower in VF and VG than in CF, which confirmed our first hypothesis. A low spatial turnover rate of edaphic variable similarity can decrease the habitat preference and homogenize soil bacterial communities across spaces (Bahram *et al.*, 2016; Fodelianakis

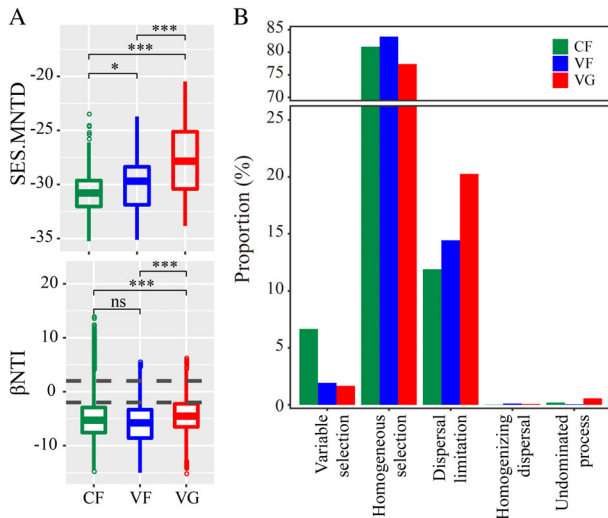


Fig. 5. Ecological processes governing bacterial community assembly within each land-use type.

A. Boxplots showing differences in SES.MNTD and β NTI values among land-use types. Single asterisk and triple asterisk indicate significant differences at $P < 0.05$ and $P < 0.001$ (Tukey's HSD test) respectively.

B. Relative contributions of ecological processes governing bacterial community assembly within each land-use type. CF, open field cultivated with main crops; VF, open field cultivated with vegetables; VG, greenhouse cultivated with vegetables. [Color figure can be viewed at wileyonlinelibrary.com]

et al., 2019). However, the bacterial communities of CF and VG exhibited a lower migration rate as revealed by the Sloan neutral model (Sloan *et al.*, 2006). Possibly, the effect of low spatial heterogeneity of edaphic variables overwhelmed that of low dispersal on the bacterial community distance-decay pattern for VF and VG.

The null model showed that soil bacterial community assembly was governed by both deterministic and stochastic processes, with deterministic processes exerting a stronger influence than stochastic processes. However, the relative importance of deterministic processes versus stochastic processes in bacterial community assembly varied among land-use types, which supported our second hypothesis. Specifically, agricultural intensification decreased the relative importance of variable selection and increased that of dispersal limitation on bacterial community assembly. These results support the proposition that, in systems with low habitat heterogeneity, which generally have lower habitat preference, the importance of stochastic processes may increase (Bahram *et al.*, 2016). The SES.MNTD value was higher in VF and VG than in CF, indicating that soil bacterial community was less closely phylogenetic clustered in VF and VG. The weak phylogenetic clustering for VF and VG in a more homogeneous environment might be associated with the decreased importance of deterministic

processes, since environmental filtering can generate phylogenetic clustering (Goberna *et al.*, 2014).

Although the soil bacterial community composition was more similar across spaces under greenhouse conditions, greenhouse production increased the variance of soil bacterial community β -diversity, that is, it decreased the soil bacterial community similarity, in contrast to the other two treatments. The increased β -diversity may be attributed to the increased stochasticity in VG under high land-use intensity, since stochastic processes and environmental perturbations (e.g. disturbance and fertilization) could promote variations in the community composition (Vellend, 2010; Zaneveld *et al.*, 2017; Wang *et al.*, 2020).

Soil acidification as an important mediator of changes in bacterial community assembly processes induced by agricultural intensification

In this study, edaphic variables, particularly soil pH, rather than geographical distance, better explained the patterns of bacterial community β -diversities. Most VG-sensitive OTUs were negatively correlated with soil pH, whereas most CF_VF-sensitive OTUs were positively correlated, which also indicated that soil pH was a main factor affecting the soil bacterial community composition (Fierer and Jackson, 2006; Tripathi *et al.*, 2018; Jiao and Lu, 2020). Most soil microorganisms prefer a neutral pH environment and soil bacterial communities generally exhibit lower diversity in both strongly alkaline and acidic soils (Fierer and Jackson, 2006). Greenhouse vegetable production systems usually receive high inorganic fertilizer inputs, which can cause significant soil acidification (Hu *et al.*, 2017). We observed that soil pH was positively correlated with soil bacterial taxonomic α -diversity. Therefore, the decrease in soil bacterial taxonomic α -diversity in VG may be associated with the decrease in soil pH caused by the high rate of inorganic fertilizer application.

In this study, edaphic variables were strongly associated with the β NTI value, with soil pH as the most influential factor. In particular, β NTI values were negatively correlated with the changes in soil pH, indicating that variations in edaphic variables, especially soil pH, exerted substantial effects on the bacterial community assembly process. Greenhouse vegetable production acidified the soil and promoted stochastic processes. Moreover, soil bacterial communities were less phylogenetically clustered and subjected to more stochastic assembly processes in acidic soils. The endocellular pH of most microorganisms is near neutral; therefore, it is generally assumed that soil pH deviant far from neutral would exert deterministic selective effects on soil microbial communities and cause phylogenetic clustering (Fierer and Jackson, 2006; Tripathi *et al.*, 2018). However, empirical

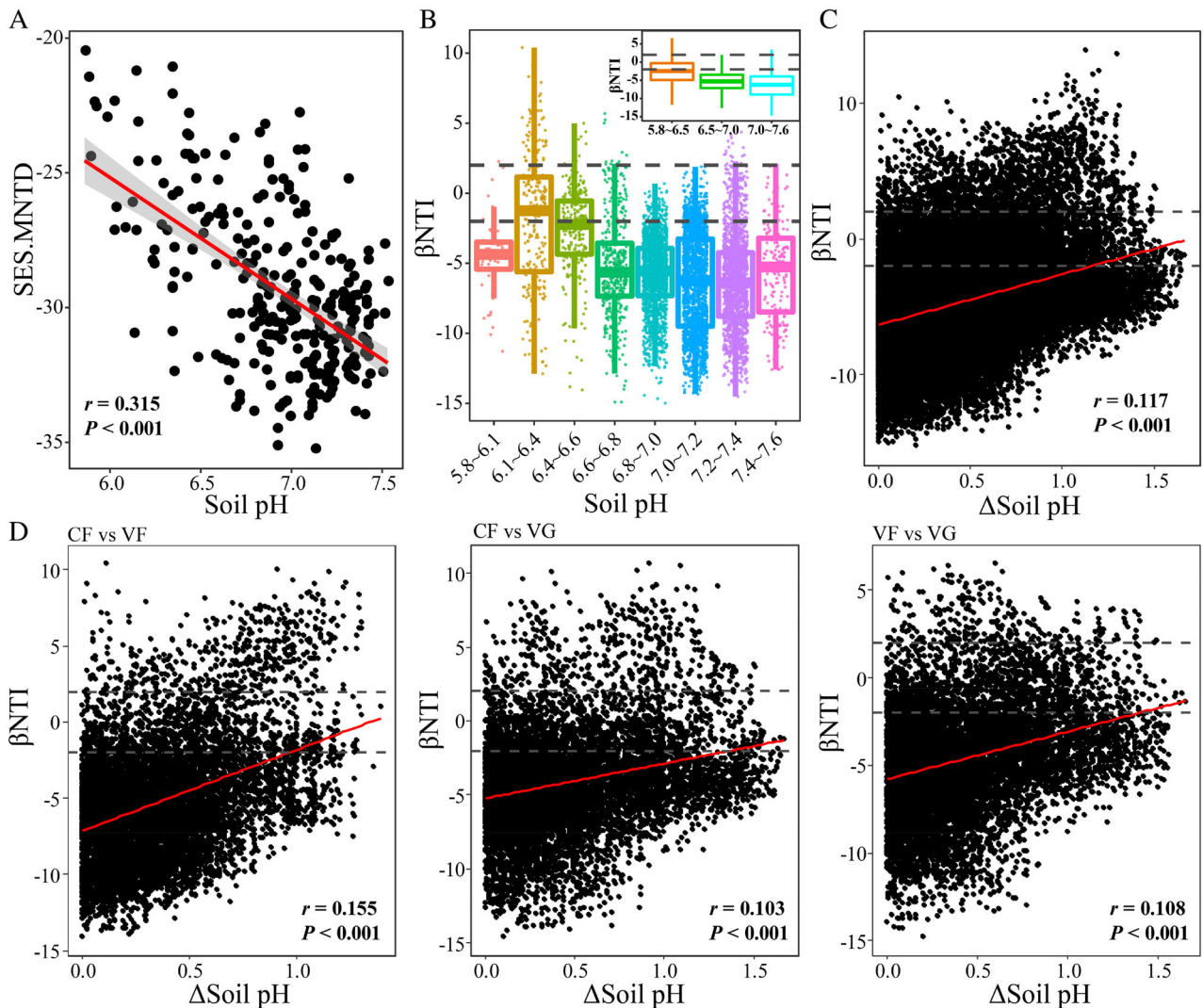


Fig. 6. Relationships between SES.MNTD or β NTI of soil bacterial communities and soil pH.

A. Relationships between SES.MNTD and soil pH across all samples.

B. Boxplots showing β NTI distributions across different categories of soil pH. Horizontal dashed grey lines indicate lower and upper significance thresholds at β NTI = -2 and +2 respectively.

C. Relationships between β NTI and differences in soil pH across all samples.

D. Relationships between β NTI and differences in soil pH between treatments. The solid red line represents the linear regression model. Shaded areas represent the estimated 95% confidence intervals. CF, open field cultivated with main crops; VF, open field cultivated with vegetables; VG, greenhouse cultivated with vegetables. [Color figure can be viewed at wileyonlinelibrary.com]

studies have reported contrasting results. For example, the study of bacterial communities in non-agricultural soils (e.g. glacier, tundra, forest and desert) supported this assumption and showed that homogeneous selection dominated in more acidic and alkaline soils, whereas stochastic assembly processes dominated in soils with close-to-neutral pH (Tripathi *et al.*, 2018). However, studies of bacterioplankton in freshwater lakes (Ren *et al.*, 2015) and bacterial communities in agricultural soils have shown an increased role of stochasticity in acidic environments (Barnett *et al.*, 2020; Jiao and Lu, 2020). This inconsistency among studies may arise

from differences in the attributes of these ecosystems. In addition to its direct effects, soil pH can indirectly affect the soil bacterial communities by altering the solubility of elements (e.g. phosphorus, aluminium and iron) and plant growth (Lammel *et al.*, 2018). Some bacterial taxa may respond differently to the direct and indirect effects of soil pH. As such, it is possible that these indirect effects can alter the direct effect of pH on the assembly of bacterial communities. Moreover, agroecosystems are generally subjected to high levels of disturbance, which can function in conjunction with environmental factors to influence the diversity and composition of ecological

communities (Houseman *et al.*, 2008). For example, Barnett *et al.* (2020) found that the land-use type and soil pH exerted interactive effects on bacterial community assembly and this interaction varied across pH classes (i.e. acidic and neutral soils).

In addition to soil acidification, changes in other edaphic variables were also associated with changes in the bacterial community assembly processes. Increasing soil fertility could increase the relative importance of stochastic processes, as increasing soil nutrients could weaken niche selection by reducing the competition for resources and enhancing stochastic processes by promoting the growth of soil microbial communities (Zhou *et al.*, 2014; Chen *et al.*, 2017; Feng *et al.*, 2018). Our quantitative PCR analysis revealed an increase in bacterial abundance in VG, indicating that an increase in soil nutrients owing to the high application rates of fertilizers may promote stochasticity in VG. Disturbance, which imposes various stresses on soil microbial communities, is an important ecological factor affecting microbial community assembly (Stegen *et al.*, 2013; Zhou and Ning, 2017). For example, disturbances induced by drought can promote deterministic processes (Chase, 2007; Valverde *et al.*, 2014; Lupatini *et al.*, 2019). Vegetables cultivated in greenhouses are frequently irrigated, and therefore, soils in greenhouses are exposed to minor variation in soil moisture, which might be associated with the decreased importance of deterministic processes in VG.

Soil bacterial communities were less closely associated under greenhouse conditions

The correlation network of VG had lower average connectivity and clustering coefficient than that of CF and VF, indicating that the bacterial taxa were less closely associated in VG (Newman, 2006). The lower complexity observed in VG may be related to the high-level nutrient contents, as the complexity of the bacterial community network could decrease in soils with higher nutrient levels (Dini-Andreote *et al.*, 2014; Feng *et al.*, 2017). The observation of a less closely association pattern and high contribution of stochastic processes in VG is consistent with the assumption that stochastic processes play a more important role in structuring an ecological community with more random species associations (Veech, 2013; Danczak *et al.*, 2018).

In this study, different types of land-use type sensitive OTUs exhibited different association patterns. For example, CF_VF-sensitive OTUs tended to correlate more among themselves than with other types of sensitive OTUs. A similar trend was observed for the VG- and VF_VG-sensitive OTUs. Moreover, different types of sensitive OTUs responded differently to the changes in

edaphic variables. For example, most VG-sensitive OTUs were negatively and positively correlated with soil pH and nitrate content respectively. However, a contrasting relationship was observed for CF_VF-sensitive OTUs. Correlated species pairs may share similar ecological characteristics (Williams *et al.*, 2014). Therefore, these different types of sensitive OTUs may share different ecological niches and play different roles in agroecosystems.

Implications and limitations of this study

As essential functional components of soil microbial communities, soil bacteria are key determinants of plant health and productivity in agricultural ecosystems (Bardgett and van der Putten, 2014). Previous studies have demonstrated that agricultural intensification can negatively affect the functioning of soil microbial communities (Tsiafouli *et al.*, 2015; de Graaff *et al.*, 2019; Jin *et al.*, 2019). Although soil function was not measured in this study, the decreases in the α -diversity in VG indicated that this production system may not be sustainable (Wagg *et al.*, 2019). The finding that soil acidification act as an important mediator of the changes in bacterial community assembly processes induced by agricultural intensification suggests that agricultural practices that avoid soil acidification, such as reducing the application rate of inorganic fertilizers, may help maintain the functioning of soil microbial communities. A shortcoming of this experiment is that we measured certain edaphic variables, but not other environmental factors, such as air temperature and precipitation regimes, which are reported to be important driving factors of soil microbial assembly (Lupatini *et al.*, 2019; Jiao and Lu, 2020). Soil microorganisms can disperse passively via air and water flow (Hanson *et al.*, 2012; Fodelianakis *et al.*, 2019). As greenhouses are covered by glasses or plastic films, the passive dispersal of soil microorganisms across spaces caused by wind and rainfall events would be minimal for VG, resulting in a lower migration rate. Therefore, in addition to high land-use intensity, the semi-closed environment in greenhouses also contributed to the increased importance of dispersal limitation in VG. Further experiments are warranted to assess the role of other environmental factors in mediating the observed changes in soil microbial assembly.

Conclusions

This study provides evidence that agricultural intensification alters the diversity and assembly processes of the soil bacterial community. Specifically, agricultural intensification led to the homogenization of the soil environment across spaces, particularly decreased soil pH, which reduced the importance of variable selection and

increased the importance of dispersal limitation. Moreover, increasing land-use intensity altered the potential interspecies associations of soil bacterial communities and reshaped the soil bacterial community composition by favouring some taxa over others. These findings may help improve our understanding of soil bacterial community ecology in agroecosystems and provide references for forecasting how soil bacterial communities respond to anthropogenically induced environmental changes.

Acknowledgements

This work was funded by the National Key Research and Development Program (2018YFD1000800) and the National Natural Science Foundation of China (32072655).

References

- Bahram, M., Kohout, P., Anslan, S., Harend, H., Abarenkov, K., and Tedersoo, L. (2016) Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *ISME J* **10**: 885–896.
- Bardgett, R.D., and van der Putten, W.H. (2014) Below-ground biodiversity and ecosystem functioning. *Nature* **515**: 505–511.
- Barnett, S.E., Youngblut, N.D., and Buckley, D.H. (2020) Soil characteristics and land-use drive bacterial community assembly patterns. *FEMS Microbiol Ecol* **96**: fiz194.
- Berry, D., and Widder, S. (2014) Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front Microbiol* **5**: 219.
- Boulard, T., Raepel, C., Brun, R., Lecompte, F., Hayer, F., Carmassi, G., and Gaillard, G. (2011) Environmental impact of greenhouse tomato production in France. *Agron Sustain Dev* **31**: 757–777.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.
- Chase, J.M. (2007) Drought mediates the importance of stochastic community assembly. *Proc Natl Acad Sci U S A* **104**: 17430–17434.
- Chen, R.R., Zhong, L.H., Jing, Z.W., Guo, Z.Y., Li, Z.P., Lin, X. G., and Feng, Y.Z. (2017) Fertilization decreases compositional variation of paddy bacterial community across geographical gradient. *Soil Biol Biochem* **114**: 181–188.
- Csardi, G., and Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal Complex Syst* **1695**: 1–9.
- Danczak, R.E., Johnston, M.D., Kenah, C., Slatery, M., and Wilkins, M.J. (2018) Microbial community cohesion mediates community turnover in unperturbed aquifers. *mSystems* **3**: e00066-00018.
- De Caceres, M., Legendre, P., and Moretti, M. (2010) Improving indicator species analysis by combining groups of sites. *Oikos* **119**: 1674–1684.
- de Graaff, M.A., Hornslein, N., Throop, H.L., Kardol, P., and van Diepen, L.T.A. (2019) Effects of agricultural intensification on soil biodiversity and implications for ecosystem functioning: a meta-analysis. *Adv Agron* **155**: 1–44.
- Derakhshani, H., Tun, H.M., and Khafipour, E. (2016) An extended single-index multiplexed 16S rRNA sequencing for microbial community analysis on MiSeq illumina platforms. *J Basic Microbiol* **56**: 321–326.
- Dini-Andreote, F., Silva, M., Triado-Margarit, X., Casamayor, E.O., van Elsas, J.D., and Salles, J.F. (2014) Dynamics of bacterial community succession in a salt marsh chronosequence: evidences for temporal niche partitioning. *ISME J* **8**: 1989–2001.
- Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996–998.
- Feng, Y., Chen, R., Stegen, J.C., Guo, Z., Zhang, J., Li, Z., and Lin, X. (2018) Two key features influencing community assembly processes at regional scale: initial state and degree of change in environmental conditions. *Mol Ecol* **27**: 5238–5251.
- Feng, Y., Guo, Z., Zhong, L., Zhao, F., Zhang, J., and Lin, X. (2017) Balanced fertilization decreases environmental filtering on soil bacterial community assemblage in North China. *Front Microbiol* **8**: 2376.
- Fierer, N., and Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* **103**: 626–631.
- Fodelianakis, S., Lorz, A., Valenzuela-Cuevas, A., Barozzi, A., Booth, J.M., and Daffonchio, D. (2019) Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. *Nat Commun* **10**: 1314.
- Geisen, S., Wall, D.H., and van der Putten, W.H. (2019) Challenges and opportunities for soil biodiversity in the anthropocene. *Curr Biol* **29**: R1036–R1044.
- Goberna, M., Garcia, C., and Verdu, M. (2014) A role for biotic filtering in driving phylogenetic clustering in soil bacterial communities. *Glob Ecol Biogeogr* **23**: 1346–1355.
- Goslee, S.C., and Urban, D.L. (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Softw* **22**: 1–19.
- Guo, J.H., Liu, X.J., Zhang, Y., Shen, J.L., Han, W.X., Zhang, W.F., et al. (2010) Significant acidification in major Chinese croplands. *Science* **327**: 1008–1010.
- Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat Rev Microbiol* **10**: 497–506.
- Harrell Jr, F.E. & Dupont, C. (2020) Hmisc: harrell miscellaneous. R Package Version 4.4-0.
- Hartman, K., van der Heijden, M.G.A., Wittwer, R.A., Banerjee, S., Walsler, J.C., and Schlaeppli, K. (2018) Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome* **6**: 14.
- Herren, C.M., and McMahon, K.D. (2017) Cohesion: a method for quantifying the connectivity of microbial communities. *ISME J* **11**: 2426–2438.
- Houseman, G.R., Mittelbach, G.G., Reynolds, H.L., and Gross, K.L. (2008) Perturbations alter community convergence, divergence, and formation of multiple community states. *Ecology* **89**: 2172–2180.

- Hu, W., Zhang, Y., Huang, B., and Teng, Y. (2017) Soil environmental quality in greenhouse vegetable production systems in eastern China: current status and management strategies. *Chemosphere* **170**: 183–195.
- Jiao, S., and Lu, Y.H. (2020) Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ Microbiol* **22**: 1052–1065.
- Jin, X., Zhang, J.H., Shi, Y.J., Wu, F.Z., and Zhou, X.G. (2019) Green manures of Indian mustard and wild rocket enhance cucumber resistance to Fusarium wilt through modulating rhizosphere bacterial community composition. *Plant Soil* **441**: 283–300.
- Jurasinski, G. & Jurasinski, M.G. (2012) simba: a collection of functions for similarity analysis of vegetation data. R package version 0.3-5.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., *et al.* (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.
- Lammel, D.R., Barth, G., Ovaskainen, O., Cruz, L.M., Zanatta, J.A., Ryo, M., *et al.* (2018) Direct and indirect effects of a pH gradient bring insights into the mechanisms driving prokaryotic community structures. *Microbiome* **6**: 106.
- Lupatini, M., Suleiman, A.K.A., Jacques, R.J.S., Lemos, L. N., Pylro, V.S., Van Veen, J.A., *et al.* (2019) Moisture is more important than temperature for assembly of both potentially active and whole prokaryotic communities in subtropical grassland. *Microb Ecol* **77**: 460–470.
- Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes encoding for 16S rRNA. *Appl Environ Microbiol* **59**: 695–700.
- Newman, M.E.J. (2006) Modularity and community structure in networks. *Proc Natl Acad Sci U S A* **103**: 8577–8582.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., *et al.* (2019) vegan: community ecology package. R package version 2.5-6.
- Paradis, E., Claude, J., and Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.
- Powell, J.R., Karunaratne, S., Campbell, C.D., Yao, H.Y., Robinson, L., and Singh, B.K. (2015) Deterministic processes vary during community assembly for ecologically dissimilar taxa. *Nat Commun* **6**: 8444.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.
- Ren, L., Jeppesen, E., He, D., Wang, J., Liboriussen, L., Xing, P., and Wu, Q.L. (2015) pH influences the importance of niche-related and neutral processes in lacustrine bacterioplankton assembly. *Appl Environ Microbiol* **81**: 3104–3114.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140.
- Rodrigues, J.L.M., Pellizari, V.H., Mueller, R., Baek, K., Jesus, E.D., Paula, F.S., *et al.* (2013) Conversion of the Amazon rainforest to agriculture results in biotic homogenization of soil bacterial communities. *Proc Natl Acad Sci U S A* **110**: 988–993.
- Shen, W.S., Hu, M.C., Qian, D., Xue, H.W., Gao, N., and Lin, X.G. (2021) Microbial deterioration and restoration in greenhouse-based intensive vegetable production systems. *Plant Soil* **463**: 1–18.
- Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., and Curtis, T.P. (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* **8**: 732–740.
- Stegen, J.C., Lin, X.J., Fredrickson, J.K., Chen, X.Y., Kennedy, D.W., Murray, C.J., *et al.* (2013) Quantifying community assembly processes and identifying features that impose them. *ISME J* **7**: 2069–2079.
- Tripathi, B.M., Stegen, J.C., Kim, M., Dong, K., Adams, J.M., and Lee, Y.K. (2018) Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J* **12**: 1072–1083.
- Tsiafouli, M.A., Thebault, E., Sgardelis, S.P., de Ruiter, P.C., van der Putten, W.H., Birkhofer, K., *et al.* (2015) Intensive agriculture reduces soil biodiversity across Europe. *Glob Change Biol* **21**: 973–985.
- Valverde, A., Makhalyane, T.P., and Cowan, D.A. (2014) Contrasting assembly processes in a bacterial metacommunity along a desiccation gradient. *Front Microbiol* **5**: 668.
- Veech, J.A. (2013) A probabilistic model for analysing species co-occurrence. *Glob Ecol Biogeogr* **22**: 252–260.
- Vellend, M. (2010) Conceptual synthesis in community ecology. *Q Rev Biol* **85**: 183–206.
- Wagg, C., Schlaeppli, K., Banerjee, S., Kuramae, E.E., and van der Heijden, M.G.A. (2019) Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. *Nat Commun* **10**: 4841.
- Wang, Z., Li, Y., Li, T., Zhao, D., and Liao, Y. (2020) Conservation tillage decreases selection pressure on community assembly in the rhizosphere of arbuscular mycorrhizal fungi. *Sci Total Environ* **710**: 136326.
- Williams, R.J., Howe, A., and Hofmockel, K.S. (2014) Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. *Front Microbiol* **5**: 358.
- Zaneveld, J.R., McMinds, R., and Thurber, R.V. (2017) Stress and stability: applying the Anna Karenina principle to animal microbiomes. *Nat Microbiol* **2**: 17121.
- Zhou, J.Z., Deng, Y., Zhang, P., Xue, K., Liang, Y.T., Van Nostrand, J.D., *et al.* (2014) Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc Natl Acad Sci U S A* **111**: E836–E845.
- Zhou, J.Z., and Ning, D.L. (2017) Stochastic community assembly: does it matter in microbial ecology? *Microbiol Mol Biol Rev* **81**: e00002-00017.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1: Supplementary Information