Science of the Total Environment 643 (2018) 569–578

Contents lists available at ScienceDirect

Science of the Total Environment

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

Distinct biogeographic patterns of rhizobia and non-rhizobial endophytes associated with soybean nodules across China

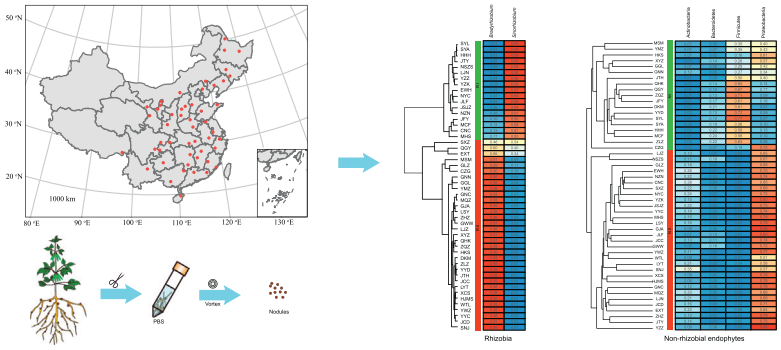
Baogang Zhang, Nini Du, Yajuan Li, Peng Shi *, Gehong Wei *

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

HIGHLIGHTS

- Proteobacteria and Firmicutes dominated non-rhizobial subcommunity.
- Rhizobia and non-rhizobial endophytes displayed distinct biogeographic patterns.
- Non-rhizobial endophytes had a lower dispersal probability than rhizobia.
- Rhizobia and non-rhizobial endophytes grouped separately in association network.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 17 April 2018
Received in revised form 19 June 2018
Accepted 19 June 2018
Available online xxxx

Editor: Frederic Coulon

Keywords:
Rhizobia
Non-rhizobial endophyte
Root nodule
Biogeography
Co-occurrence network

ABSTRACT

Both rhizobia and non-rhizobial endophytes (NRE) are inhabitants of legume nodules. The biogeography of rhizobia has been well investigated, but little is known about the spatial distribution and community assemblage of NRE. By using high-throughput sequencing, we compared biogeographic patterns of rhizobial and non-rhizobial subcommunities and investigated their bacterial co-occurrence patterns in nodules collected from 50 soybean fields across China. Dispersal probability was lower in NRE than in rhizobia, as revealed by a significant distance-decay relationship found in NRE, but not in rhizobia, in addition to a significant occupancy–abundance relationship in the entire community. Rhizobial and NRE subcommunities were significantly influenced by different environmental and spatial variables. Moreover, the rhizobial subcommunities were grouped into *Ensifer*- and *Bradyrhizobium*-dominated clusters that were significantly related to soil pH. The non-rhizobial subcommunities were grouped into Proteobacteria- and Firmicutes-dominated clusters that were more influenced by climatic than by edaphic factors. These results demonstrated that rhizobial and non-rhizobial subcommunities are characterized by distinct biogeographic patterns. Network analysis showed rhizobia and NRE as separately grouped and uncorrelated with each other, suggesting they did not share niche space in soybean nodules. In sum, these results broaden our knowledge of how bacteria are distributed and assemble as a community in root nodules.

1. Introduction

Many leguminous plants have the ability to establish binary symbiosis with some diazotrophic bacteria, collectively referred to as rhizobia. These rhizobia induce the formation of root nodules where biological

* Corresponding authors.
E-mail addresses: shipeng27@nwafu.edu.cn, (P. Shi), weigehong@nwsuaf.edu.cn (G. Wei).

https://doi.org/10.1016/j.scitotenv.2018.06.240
0048-9697/

nitrogen fixation occurs. Rhizobia have been found in Alphaproteobacterial genera (*Rhizobium*, *Ensifer*, *Bradyrhizobium*, *Mesorhizobium*, *Methylobacterium*, *Devosia*, *Azorhizobium*, *Allorhizobium*, and *Shinella*) and Betaproteobacterial genera (*Burkholderia* and *Cupriavidus*) (Peix et al., 2014). Together with rhizobia, a great diversity of endophytic bacteria, including those in the genera *Bacillus*, *Pseudomonas*, *Enterobacter*, *Chryseobacterium*, and *Sphingobacterium*, have since been detected inside legume nodules (De Meyer et al., 2015; Leite et al., 2016). Because these bacteria cannot induce nodules or perform biological nitrogen fixation they are called non-rhizobial endophytes (NRE) (Martínez-Hidalgo and Hirsch, 2017). NRE members can provide beneficial services to their host plants, such as plant growth promotion (Tariq et al., 2014), abiotic stress resistance, pathogen protection, as well as nodulation enhancement (Martinez-Hidalgo et al., 2014).

A fundamental goal in community ecology is to understand the factors that determine community distribution patterns. Many studies have explored the biogeographic patterns of rhizobia associated with several plant species, including the common bean (Cao et al., 2014; Verastegui-Valdes et al., 2014; Wang et al., 2016), soybean (Li et al., 2011; Yan et al., 2018; Zhang et al., 2011), cowpea (Chidebe et al., 2018), alfalfa (Donnarumma et al., 2014), and *Caragana* species (Lu et al., 2009). This work has revealed the influence of key environmental factors, such as precipitation, soil nutrient availability and soil pH, on the distribution of rhizobia. Although the literature is rich with studies of NRE's genetic diversity and potential roles (reviewed by Peix et al., 2014), just a few have investigated the spatial distributions of NRE associated with wild legumes, such as the genera *Sphaerophysa salsula* (Deng et al., 2011), *Caragana jubata* and *Oxytropis ochrocephala* (Xu et al., 2014), as well as the subfamily Faboideae (De Meyer et al., 2015), and all these studies relied on culture-dependent approaches and were conducted at regional scales. Hence, the full extent of NRE diversity remains unexamined and our knowledge of NRE biogeographic patterns at larger (i.e., continental) spatial scales is quite limited. Furthermore, biogeographic patterns of rhizobia and NRE have yet to be investigated in the same legume host species. Given the non-symbiotic roles of NRE, we hypothesized that rhizobia and NRE display distinct biogeographic patterns.

Microorganisms in natural ecosystems usually form complex ecological networks through direct (e.g., mutualism and competition) or indirect (e.g., environmental preferences) interactions (Faust and Raes, 2012). Characterizing the interactions among microorganisms—also called co-occurrence patterns—is crucial for better understanding their potential functions or ecological niches (Ju et al., 2014; Steele et al., 2011). Network analysis offers a powerful tool for studying the co-occurrence patterns of microbial communities, as demonstrated by its recent application to various complex environments, such as humans (Faust et al., 2012), oceans (Lima-Mendez et al., 2015), activated sludge (Ju et al., 2014), soils (Ma et al., 2016) and the plant rhizosphere (Fan et al., 2017; Shi et al., 2016). This work has revealed interesting co-occurrence patterns in microbial communities, such as non-random associations, highly connected modules (de Menezes et al., 2015), and relationships between functional groups (Bissett et al., 2013). However, co-occurrence patterns are poorly understood in nodule bacterial communities.

Soybean (*Glycine* max) is a major legume crop grown globally. It originated in China, where it widely cultivated (Li et al., 2008), which provides an excellent opportunity to study the biogeographic and co-occurrence patterns of its nodule bacterial communities at a continental scale. In this study, high-throughput sequencing technology was used to investigate the nodules' bacterial community composition in 50 soybean fields across China. We also used Molecular Ecological Network Analysis (Deng et al., 2012) to construct co-occurrence network for the nodule-dwelling bacteria. Our main objectives were (i) to determine and compare the biogeographic patterns of rhizobia and NRE in soybean nodules; and (ii) to investigate the co-occurrence patterns among bacterial taxa in these nodules.

2. Materials and methods

2.1. Sample collection and preparation

A total of 50 soybean fields across China (Fig. S1) were selected from which to collect soil and plant samples. All the fields were cultivated under conventional agricultural practices, in which chemical fertilizer and pesticide use is permitted yet organic, manure, or compost fertilizers were not used. Samples were collected in all fields at the flowering stage of soybean (May–August 2015). The methods we used for soil and root sampling, characterization of soil physicochemical properties, and climate data collection are described in detail by Zhang et al. (2018). Briefly, in each field, five topsoil samples (0–20 cm) were randomly from a ~100 m^2 plot and pooled as one bulk soil sample. A total of 15–20 randomly selected healthy plants were removed from the soil using a spade. Roots were gently shaken to remove loose soil, combined as one root sample per field. A subset of each soil sample was air-dried for an analysis of its physicochemical properties according to the standard protocols described by Bao (2000), namely soil pH (soil/water = 1:5, w/v), texture, organic matter (OC), total (TN) and available nitrogen (AN), and available phosphorus (P), potassium (K), calcium (Ca) and magnesium (Mg). Information on the fields' geographic coordinates, soybean cultivars, and soil and climate factors are provided in Supplementary Table S1.

Roots were placed in a sterile 50-mL tube containing 25 mL of a sterile phosphate-buffered saline solution (PBS, per litre: 6.33 g $NaH_2PO_4 \cdot H_2O$, 16.5 g $Na_2HPO_4 \cdot 7H_2O$, 200 μL Silwet L-77, pH 7.0), and vortexed at the maximum speed for 15 s to remove the rhizosphere soil from the root surfaces. Then the cleaned roots were transferred to a new sterile 50-mL tube with 25 mL of the sterile PBS buffer, and vortexed; this step was repeated until the PBS buffer appeared clear after vortexing. Next, the roots were moved into a new sterile tube and sonicated at low frequency for 5 min (five 30-s bursts, followed by five 30-s rests) to dislodge any attached microbes and to further clean the root exterior surface. Finally, the roots were removed and rinsed in a fresh volume of 25-mL PBS buffer. The efficacy of these procedures for removing microbes from soybean nodule surfaces had been confirmed in our recent study (Xiao et al., 2017).

2.2. DNA extraction, sequencing, and analysis

Nodules were taken from each root sample, and approximately 500 mg of healthy nodules were ground in liquid nitrogen using mortar and pestle. Their total DNA was then extracted using the FastDNA SPIN Kit for soil (MP Biomedicals, Santa Ana, USA) following the manufacturer's instructions. We used the primers 515F (5′-GTGCCAGCMGCCGCG GTAA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) (Caporaso et al., 2012) to amplify the V4 region of the bacterial 16S rRNA gene, by following the PCR protocols described in a recent study of ours (Zhang et al., 2018). Paired-end (250 bp) sequencing was conducted on an Illumina HiSeq 2500 instrument at Novogene Bioinformatics Technology Co., Ltd. (Beijing, China).

Raw sequence data were analyzed using QIIME (Caporaso et al., 2010), and the paired-end sequences merged using FLASH (Magoc and Salzberg, 2011). Those sequences with a length <200 bp, an average quality <25, or containing ambiguous bases were removed. Quality-filtered sequences were then clustered into operational taxonomic units (OTUs) based on 97% similarity using UPARSE (Edgar, 2013). Taxonomic annotation of each OTU was performed by the Ribosomal Database Project (RDP) Classifier (Wang et al., 2007) with the Greengenes database. All OTUs annotated as chloroplast or mitochondria were removed from the OTU matrix table, and singleton OTUs (containing only one sequence) were also removed to avoid possible biases. To correct for sequencing effort across samples, the OTU table was rarefied to 25,571 sequences per sample. Rhizobial OTUs were defined here as taxa that belonged to the genera *Bradyrhizobium*, *Ensifer*, *Rhizobium*, and

Mesorhizobium, since members of these four groups are reportedly soybean microsymbionts (Biate et al., 2014). The remaining OTUs were thus defined as NRE. OTUs occurring in at least 25 of 50 samples were defined as generalists, of which those with a mean relative abundance >0.1% (Ma et al., 2017) were defined as core OTUs. The obtained raw sequence data was deposited in the NCBI small-read archive database, under BioProject accession number PRJNA395393, with the run number SRR5859786-SRR5860102.

2.3. Network analysis

To identify and explore the co-occurrence patterns of bacteria in soybean nodules, network analysis was performed based on the random matrix theory (RMT)-based method in the Molecular Ecological Network Analyses Pipeline (http://ieg4.rccc.ou.edu/mena/main.cgi). More information on the theory, algorithms, and procedures of this pipeline is described elsewhere (Deng et al., 2012; Zhou et al., 2010). Network analysis was conducted as following steps. First, the rarefied OTU table was prepared in the correct format, as instructed by the pipeline, and submitted for network construction. Second, in the settings for MENA construction, the majority was set to 11, thus ensuring that only OTUs present in >20% (Ju et al., 2014) of the 50 sites were included in the network analysis. Using the default settings, the cutoff for the correlation coefficients between each pair of OTUs—that is, the similarity threshold—was determined as 0.68 for network construction. Third, a set of network properties was calculated: namely the R square of power-law, average clustering coefficient, and average path distance. Modules were densely connected node groups with more connections inside them than outside, detected by using the greedy modularity optimization algorithm. Definitions and interpretations of other key network properties are described in detail by Deng et al. (2012). Fourth, the network, node, and edge files were obtained by running the "Output for the Cytoscape software visualization" step. Network was visualized using Cytoscape v3.5.1 (Shannon et al., 2003). Fifth, the "Randomize the network structure and then calculate network properties" command was run to compare the topologies of the empirical and 100 randomly generated networks.

2.4. Statistical analyses

All statistical analyses were performed in R v3.2.3 (R Core Team, 2016) or PRIMER v7 (Clarke and Gorley, 2015). Mantel tests were used to calculate Spearman rank correlation coefficients among the Bray–Curtis dissimilarity matrices of the entire, rhizobial, and non-rhizobial bacterial communities. A cluster analysis of rhizobial and non-rhizobial subcommunities based on Bray–Curtis dissimilarity matrices was then performed with the vegan package. This generated two groups in each subcommunity, whose taxonomic distributions were compared using the Wilcoxon signed-rank test. To determine significant differences between groups in their rhizobial or non-rhizobial bacterial subcommunity compositions, permutational multivariate analysis of variance (PERMANOVA) with 999 permutations and non-metric multidimensional scaling (NMDS) ordination were performed using the vegan package. Relationships between the mean relative abundance of OTU and site occupancy (the number of samples in which an OTU occurred) were assessed with Spearman rank correlations.

Environmental factors (except pH) were log(x + 1)-transformed to reduce nonlinearity and to improve normality for the multivariate statistical analyses. Spatial variables were derived from the longitude and latitude coordinates of each field site, by using the principal coordinates of neighbor matrices (PCNM) analysis (Borcard et al., 2011). The geographic coordinates of each site were added to the list of spatial variables, because PCNM variables do not include linear trends associated with longitude and latitude. To select the significant edaphic, climatic, and spatial variables that influenced bacterial subcommunity composition, distance-based linear modelling (DistLM) analysis was carried out using a forward-selection protocol with an adjusted-R^2 selection criterion in PRIMER v7, from which the significant variables were used to plot distance-based redundancy ordinations (db-RDA). The relative contributions of edaphic, climatic, and spatial factors to bacterial subcommunity composition were determined with a variance partitioning analysis (vegan package), in which the Bray–Curtis dissimilarity matrix was the response variable, while the significant edaphic, climatic and spatial variables sets were explanatory variables. To investigate the effect of geographic distance on bacterial subcommunities, standard and partial Mantel tests with 999 permutations were carried out by calculating the Spearman correlations between the Bray–Curtis and geographical distances while controlling and not controlling for environmental distance. Spearman rank correlations were then used to examine the associations between environmental variables and core OTUs.

3. Results

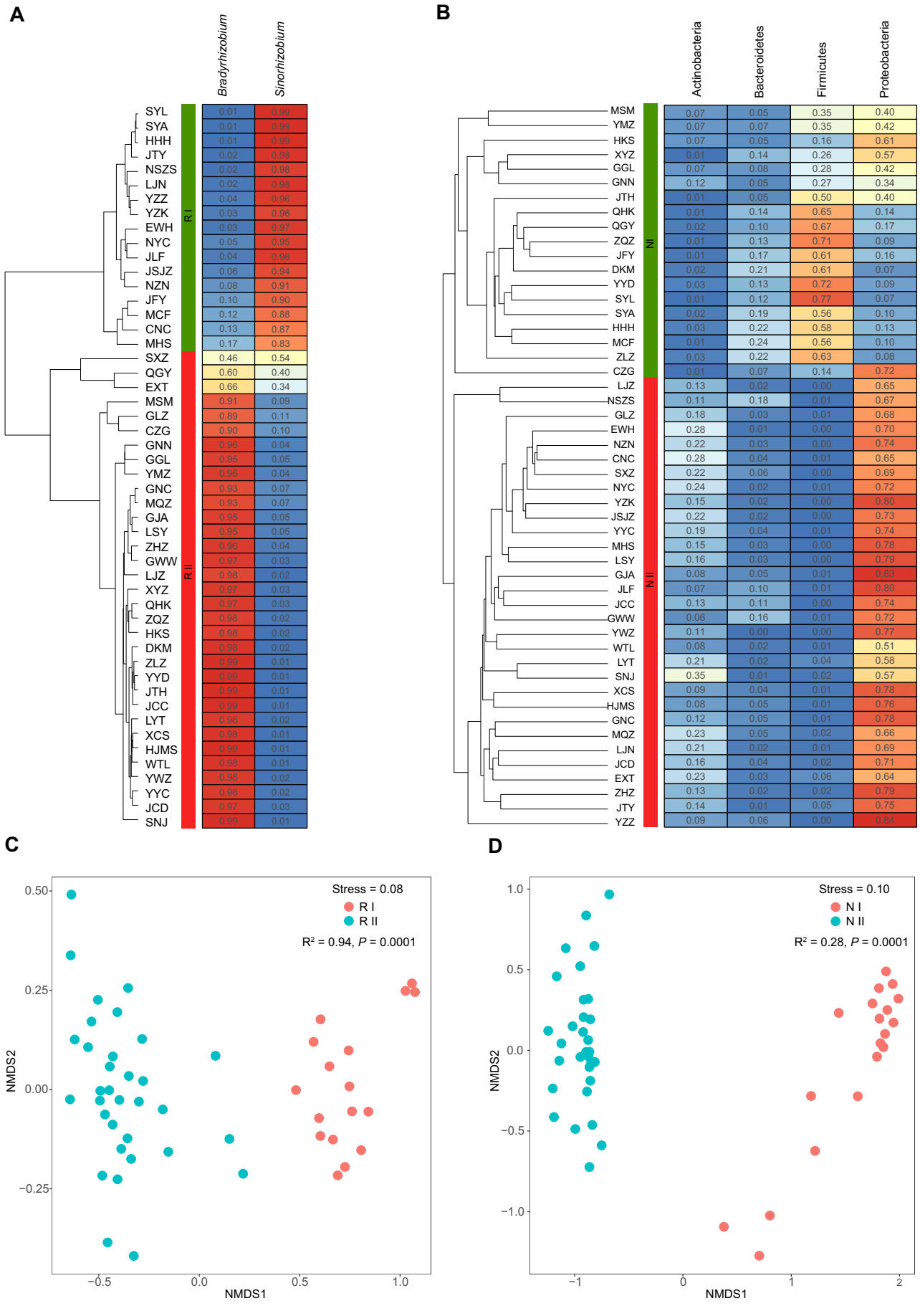
3.1. Rhizobial and non-rhizobial bacterial community composition

A total of 3435 bacterial OTUs (97% sequence identity) were clustered from 2,587,704 quality-filtered reads across the 50 sampled sites. Both the OTU rarefaction curves (Fig. S2) and Good's coverage values (mean ± SE: 0.996 ± 0.002) indicated that most of the bacterial taxa had been recovered from the nodule samples. After randomly selecting a subset of 25,571 reads per sample, 2450 OTUs that clustered from 1,278,550 reads were selected for further analysis. A total of 16 (0.65%) OTUs with 1,241,676 reads (97.1%) were classified as rhizobial OTUs, and likewise 2434 (99.35%) taxa and 29,824 (2.9%) reads were NRE OTUs (Table S2). As expected, *Bradyrhizobium* and *Ensifer* largely dominated the rhizobial subcommunity, accounting for 63.2% and 36.7% of the rhizobial reads, respectively (Fig. S3). In the NRE subcommunity, Proteobacteria, Firmicutes, and Actinobacteria were the prominent phyla, representing 53.7%, 19.5%, and 12.8% of the non-rhizobial reads, respectively (Fig. S3). At the taxonomic level of class, Alphaproteobacteria, Clostridia, and Gammaproteobacteria were the most abundant members in the NRE community, representing 27.2%, 18.0%, and 14.7% of the non-rhizobial reads, respectively.

3.2. Spatial distribution of bacterial taxa

As revealed by Mantel test, the NRE subcommunity had a biogeographic pattern that differed from both the rhizobial subcommunity ($r = 0.04$, $P > 0.05$) and the entire community ($r = 0.07$, $P > 0.05$). The rhizobial subcommunities of the 50 nodule samples clustered into two groups, R1 ($n = 17$) and R2 ($n = 33$) dominated by the genera *Ensifer* and *Bradyrhizobium*, respectively (Fig. 1A). Likewise, the NRE subcommunities also clustered into two groups, N1 ($n = 19$) and N2 ($n = 31$) dominated by the phyla Firmicutes and Proteobacteria, respectively (Fig. 1B). Additionally, samples in group N1 had a higher relative abundance of Actinobacteria but a lower relative one of Bacteroidetes when compared with group N2 (Fig. S4). These clustering results were confirmed by the MNDS ordination and PERMANOVA results (Fig. 1C–D), and revealed that both rhizobial and NRE subcommunities featured nonrandom spatial distributions. Specifically, group R1 samples were mainly from China's central region, while those of group R2 came from northeast and southern regions. For the NRE subcommunity, the group N1 samples were primarily from the southern region of China, while group N2 samples came from its northern and central regions (Fig. S5).

The number of samples each OTU occurred in significantly increased with the mean relative abundance of bacterial OTUs ($r = 0.907$, $P < 0.001$; Fig. 2). Notably, the rhizobia were more widely distributed than NRE. For example, 87.5% (14/16) of rhizobial OTUs occupied ≥50% of sites whereas only 1.8% (63/3434) of non-rhizobial OTUs did. These


A
Bradyrhizobium
Sinorhizobium
R I
R II
SYL 0.01 0.99
SYA 0.01 0.99
HHH 0.01 0.99
JTY 0.02 0.98
NSZS 0.02 0.98
LJN 0.02 0.98
YZZ 0.04 0.96
YZK 0.03 0.96
EWH 0.03 0.97
NYC 0.05 0.95
JLF 0.04 0.96
JSJZ 0.06 0.94
NZN 0.08 0.91
JFY 0.10 0.90
MCF 0.12 0.88
CNC 0.13 0.87
MHS 0.17 0.83
SXZ 0.46 0.54
QGY 0.60 0.40
EXT 0.66 0.34
MSM 0.91 0.09
GLZ 0.89 0.11
CZG 0.90 0.10
GNN 0.96 0.04
GGL 0.95 0.05
YMZ 0.96 0.04
GNC 0.93 0.07
MQZ 0.93 0.07
GJA 0.95 0.05
LSY 0.95 0.05
ZHZ 0.96 0.04
GWW 0.97 0.03
LJZ 0.98 0.02
XYZ 0.97 0.03
QHK 0.97 0.03
ZQZ 0.98 0.02
HKS 0.98 0.02
DKM 0.98 0.02
ZLZ 0.99 0.01
YYD 0.99 0.01
JTH 0.99 0.01
JCC 0.99 0.01
LYT 0.98 0.02
XCS 0.98 0.01
HJMS 0.99 0.01
WTL 0.98 0.01
YWZ 0.98 0.02
YYC 0.98 0.02
JCD 0.97 0.03
SNJ 0.99 0.01
B
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
N I
N II
MSM 0.07 0.05 0.35 0.40
YMZ 0.07 0.07 0.35 0.42
HKS 0.07 0.05 0.16 0.61
XYZ 0.01 0.14 0.26 0.57
GGL 0.07 0.08 0.28 0.42
GNN 0.12 0.05 0.27 0.34
JTH 0.01 0.05 0.50 0.40
QHK 0.01 0.14 0.65 0.14
QGY 0.02 0.10 0.67 0.17
ZQZ 0.01 0.13 0.71 0.09
JFY 0.01 0.17 0.61 0.16
DKM 0.02 0.21 0.61 0.07
YYD 0.03 0.13 0.72 0.09
SYL 0.01 0.12 0.77 0.07
SYA 0.02 0.19 0.56 0.10
HHH 0.03 0.22 0.58 0.13
MCF 0.01 0.24 0.56 0.10
ZLZ 0.03 0.22 0.63 0.08
CZG 0.01 0.07 0.14 0.72
LJZ 0.13 0.02 0.00 0.65
NSZS 0.11 0.18 0.01 0.67
GLZ 0.18 0.03 0.01 0.68
EWH 0.28 0.01 0.00 0.70
NZN 0.22 0.03 0.00 0.74
CNC 0.28 0.04 0.01 0.65
SXZ 0.22 0.06 0.00 0.69
NYC 0.24 0.02 0.01 0.72
YZK 0.15 0.02 0.00 0.80
JSJZ 0.22 0.02 0.00 0.73
YYC 0.19 0.04 0.01 0.74
MHS 0.15 0.03 0.00 0.78
LSY 0.16 0.03 0.00 0.79
GJA 0.08 0.05 0.01 0.83
JLF 0.07 0.10 0.01 0.80
JCC 0.13 0.11 0.00 0.74
GWW 0.06 0.16 0.01 0.72
YWZ 0.11 0.00 0.00 0.77
WTL 0.08 0.02 0.01 0.51
LYT 0.21 0.02 0.04 0.58
SNJ 0.35 0.01 0.02 0.57
XCS 0.09 0.04 0.01 0.78
HJMS 0.08 0.05 0.01 0.76
GNC 0.12 0.05 0.01 0.78
MQZ 0.23 0.05 0.02 0.66
LJN 0.21 0.02 0.01 0.69
JCD 0.16 0.04 0.02 0.71
EXT 0.23 0.03 0.06 0.64
ZHZ 0.13 0.02 0.02 0.79
JTY 0.14 0.01 0.05 0.75
YZZ 0.09 0.06 0.00 0.84
C
Stress = 0.08
R I
R II
R² = 0.94, P = 0.0001
NMDS1
NMDS2
D
Stress = 0.10
N I
N II
R² = 0.28, P = 0.0001
NMDS1
NMDS2

OTUs with a ≥50% site-occupancy were defined as generalists. Overall, rhizobial generalists were composed of *Ensifer* (7 OTUs), *Bradyrhizobium* (6 OTUs), and *Rhizobium* (1 OTU). Proteobacteria (52/63 OTUs) were prominent NRE generalists, followed by the Actinobacteria (9 OTUs), Bacteroidetes (1 OTU), and Cyanobacteria (1 OTU). At the genus level, some common genera were identified among the NRE generalists: namely *Streptomyces*, *Sphingomonas*, *Pseudomonas*, *Agrobacterium*, and *Burkholderia*.

3.3. Effects of ecological factors on bacterial community composition

The db-RDA ordination showed that rhizobial subcommunity composition was significantly influenced by three edaphic factors (soil pH, OC and TN), one climatic factor (MAP), and four spatial variables (PCNM 1, 2, 3, and 22; Fig. 3). Among the environmental variables detected, soil pH was the most important predictor, explaining 34.3% of the variation in rhizobial subcommunity composition (Table S3). Moreover, *Ensifer* dominated the R1 group derived from soybean nodules in alkaline soils, while *Bradyrhizobium* dominated the R2 group derived from neutral to acidic soils. (Fig. S6). However, the db-RDA for the NRE subcommunity showed a rather different pattern, being significantly influenced by two edaphic factors (soil Mg and Ca content), four climatic factors (MAP, MAT, MMT3, and TS), and five spatial variables (latitude, PCNM 2, 6, 26, and 28; Fig. 3).

Variation partitioning let us determine the relative contributions of the significant edaphic, climatic and spatial variables to bacterial subcommunity composition (Fig. 3). These three components together explained 62.1% and 35.4% of the variation in the rhizobial and NRE subcommunity compositions, respectively. For the rhizobial subcommunity, 46.1% of this variation arose from the joint effect of environmental (edaphic and climatic) factors and spatial variables. Pure spatial component amounted to ~16%, while pure edaphic component explained <1% of the variation (pure climatic component was negligible). For the NRE subcommunity, pure environmental component—edaphic, climatic, and their joint effect, excluding the spatial effect—accounted for 16.6% of the variation, while pure spatial component amounted to almost 12%. Climatic variables (11.3%) explained more variation in the NRE subcommunity composition than did edaphic factors (5.3%).

The Mantel tests also revealed that geographic distance had a negligible effect on the similarity among rhizobial subcommunities ($R = 0.05$, $P = 0.13$). However, geographic distance was correlated negatively with the similarity of the NRE subcommunities ($R = 0.13$, $P = 0.012$; Fig. S7). This distance–decay pattern remained significant when controlling for environmental distance in a partial Mantel test ($R = 0.10$, $P = 0.012$).

3.4. Associations between core OTUs and environmental variables

Since nodule samples were collected across a broad range of environmental gradients, we could identify the core bacteria in soybean nodules and study their spatial distribution patterns. Rhizobial core taxa included two *Bradyrhizobium* OTUs and three *Ensifer* OTUs: the former were negatively correlated with soil pH, organic carbon, C/N ratio, and Ca and Mg contents, while the latter were positively correlated with MAT and MAP and also showed contrasting environmental relationships (Fig. 4). All the non-rhizobial generalists (63 OTUs) were identified as core taxa, but almost half of them (33 OTUs) were not correlated with any measured environmental variable. Soil pH, Mg content, MMT3, and MAP were the major variables closely correlated with non-rhizobial core taxa. Furthermore, OTUs belonging to the order Rhizobiales (except one OTU belonging to Bradyrhizobiaceae) were positively correlated with soil pH and Mg content, but negatively correlated with MMT3 and MAP. OTUs belonging to the order Burkholderiales were also negatively correlated with MMT3 but not correlated with soil pH (Fig. 4).

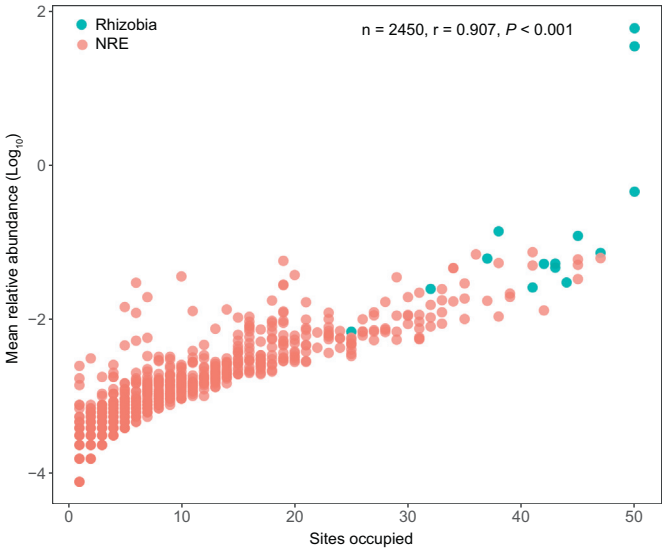


Fig. 2. Spearman rank correlation between the mean relative abundance of operational taxonomic units (OUTs) and number of sites occupied (NRE, non-rhizobial endophytes; n = number of OTUs).

3.5. Co-occurrence patterns in bacterial communities

The RMT-based network analysis let us explore the co-occurrence patterns of bacteria in the soybean nodules. The observed network consisted of 213 OTUs and 448 associations of 70.1% were positive (Fig. 5). The network degree followed a power-law distribution ($R^2 = 0.91$), thus indicating a scale-free characteristic of the network. The observed network had an average path length of 5.11 and an average clustering coefficient of 0.26, which were significantly greater than those of corresponding random networks (Table S4), thus indicating the small-world property of the observed network.

The constructed network had modular structure with a modularity value of 0.727, which exceeded that of random networks (mean ± SE: 0.459 ± 0.007). We therefore analyzed the phylogeny composition of major modules having at least 10 nodes. In the largest module (I), Proteobacteria and Firmicutes were dominant members and tended to co-exclude (i.e., negative associations) each other and with Actinobacteria members. Proteobacteria were dominant members of modules II, III, and IV, and co-occurred (i.e., positive associations) with Actinobacteria in modules II and IV. *Bradyrhizobium* and *Ensifer* co-excluded evenly in module V, but they did not correlate with any NRE members. Firmicutes dominated module VI, while Proteobacteria, Actinobacteria, and Bacteroidetes all co-occurred in module VII.

4. Discussion

4.1. Richness and community composition of nodule endophytes

Using high-throughput sequencing, we detected 3435 non-singleton OTUs in the soybean nodules collected from 50 fields across China. This was almost double the number reportedly associated with the same plant species in a greenhouse study, in which 1818 OTUs

Fig. 1. Bacterial community composition in soybean nodules collected from 50 fields across China. Hierarchical cluster diagrams (generate by using the average linkage) and heatmaps for rhizobial (A) and non-rhizobial (B) subcommunities. Heatmap values reflect relative abundances of dominant genera or phyla at each site. Non-metric multidimensional scaling (NMDS) ordination plots of rhizobial (C) and non-rhizobial (D) bacterial communities based on Bray-Curtis dissimilarities.

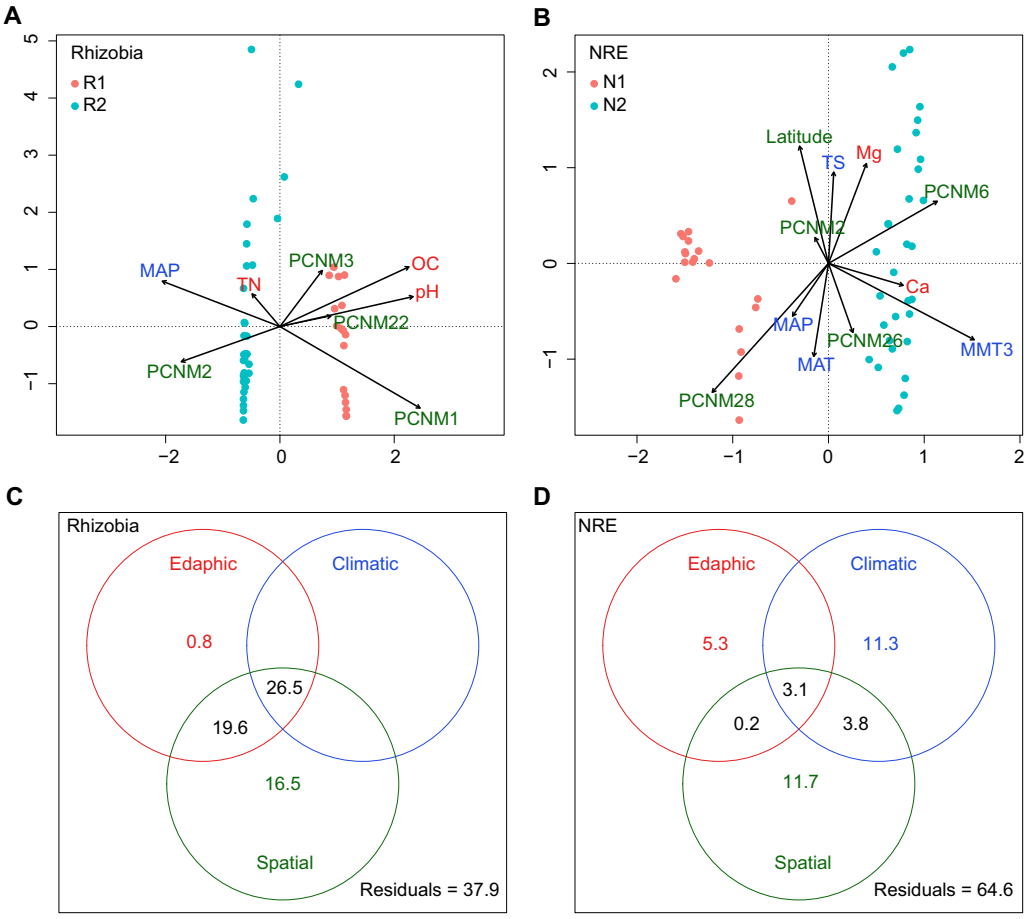


Fig. 3. Contribution of ecological factors to bacterial community composition in soybean nodules. Distance-based redundancy analysis (db-RDA) ordinations showing significant variables that influenced the composition of rhizobial (A) and non-rhizobial (B) sub-communities. Variation partitioning analysis illustrating the contributions of edaphic, climatic and spatial factors to the rhizobial (C) and non-rhizobial (D) sub-community compositions. NRE, non-rhizobial endophytes; OC, organic carbon; TN, total nitrogen; Mg, soil magnesium content; MAP, mean annual precipitation; MAT, mean annual temperature; MMT3, three-month mean temperature ranges; TS, temperature seasonality; and PCNM, principal coordinates of neighbor matrices.

were detected (Xiao et al., 2017). The higher bacterial richness of our study may be attributable to the deep sequencing used, in that bacteria were fully recovered from the nodules, as confirmed by Good's coverage values. The broad spatial gradients that our samples encompassed could also have contributed to high bacterial richness, given the restrictive distribution of many non-rhizobial taxa. To date, bacterial communities in nodules have rarely been investigated; hence, this study has greatly expanded the bacterial richness known in this widespread belowground habitat.

After filtering the rhizobial OTU, the NRE in soybean nodules were dominated by Proteobacteria, Firmicutes, Actinobacteria, and Bacteroidetes, not unlike in cowpea nodules (Leite et al., 2016). However, the dominant classes of Alphaproteobacteria, Clostridia and Gammaproteobacteria in the NRE subcommunities that we uncovered were not found in cowpea nodules, in which Flavobacteria and Actinobacteria classes were the dominant NRE members (Leite et al., 2016). These divergent findings suggest host lineage–dependent selection for bacteria occurs in legume nodules. The host plant might select lineage–dependent NRE members by modulating the quality and quantity of carbon sources delivered to them (Dudeja et al., 2012).

4.2. Distinct geographic patterns between rhizobial and non-rhizobial bacteria

In the present study, geographic distance significantly influenced the NRE but not rhizobial subcommunity composition in soybean nodules. Distance-decay patterns may result from spatially autocorrelated environmental factors or from species' dispersal limitations (Hanson et al., 2012). In our study, NRE subcommunity similarities were correlated with geographic distance regardless of whether environmental factors were controlled or not, indicating the NRE distance-decay pattern was driven by dispersal limitation. The absence of a distance-decay pattern for rhizobia highlights their strong dispersal ability, which could explain their ubiquitous distributions in soybean nodules. Compared with NRE, rhizobia represented a higher proportion of reads and so they can be considered as abundant taxa. Similar to findings reported for lakes (Liu et al., 2015) and soils (Nemergut et al., 2011), our result supports the view that more abundant microbial taxa are more apt realizing long-distance dispersal events (Martiny et al., 2006). Another plausible mechanism for the ubiquitous distribution of *Bradyrhizobium* and *Ensifer* is their large genome sizes (>6 Mb; Konstantinidis and Tiedje, 2004), which may foster diverse metabolism-related genes enabling these two taxa to grow on a wide spectrum of resources, and thereby enlarge their geographical distributions (Barberan et al., 2014).

Characterizing the key influential factors and their contributions to variation in bacterial community composition is critical to understanding bacterial biogeography. Our db-RDA analyses showed the rhizobial and NRE subcommunities were significantly related to different edaphic, climatic, and spatial variables, thus providing further evidence that these two bacterial subcommunities exhibit distinctive biogeographic patterns. Soil pH was the best predictor for the rhizobial subcommunity, as *Ensifer* and *Bradyrhizobium* were dominant members in alkaline and neutral to acidic soil pH samples, respectively (Fig. S6). This result is consistent with known genomic features of these two genera: genes involved in alkaline–saline adaptations and osmoprotection

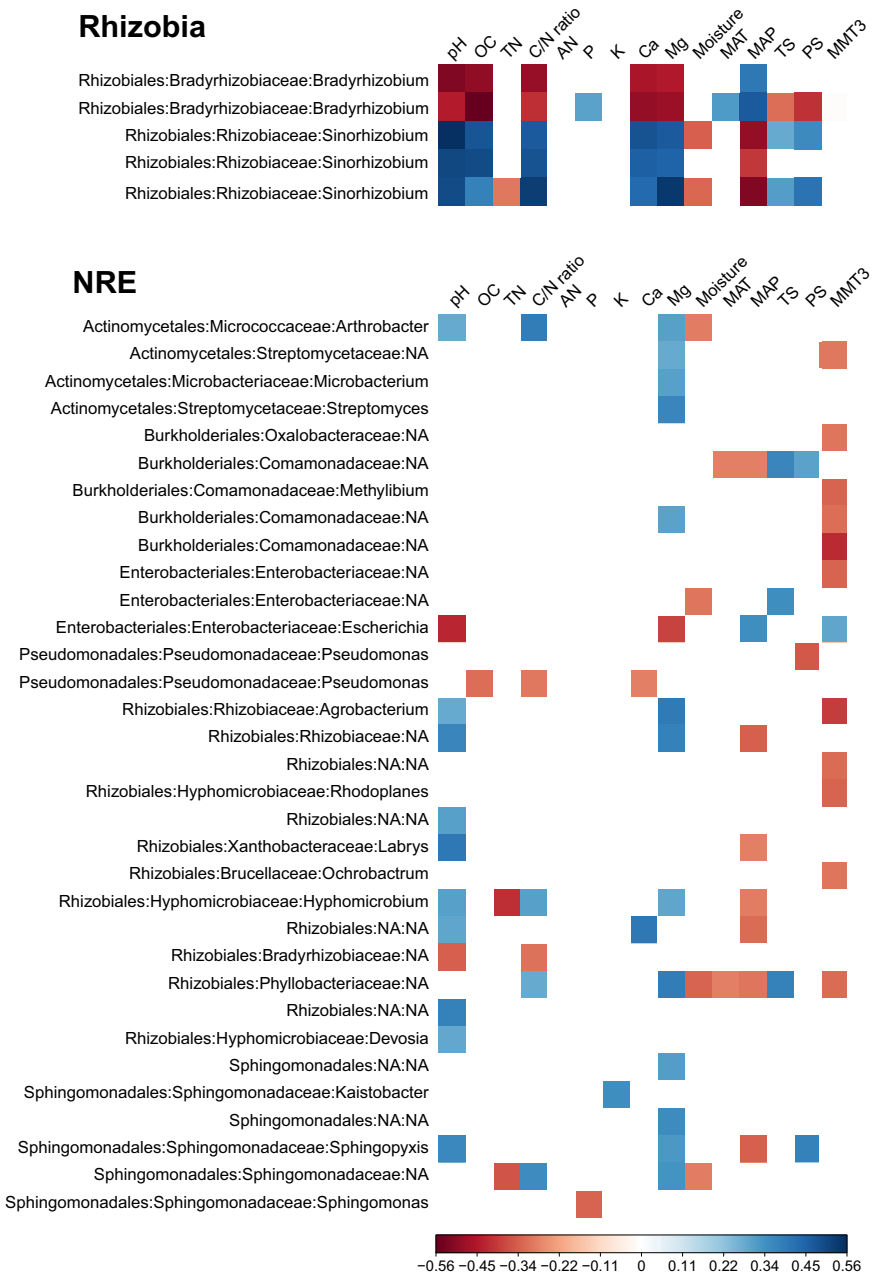


Fig. 4. Spearman rank correlations between the relative abundance of core OTUs and environmental variables in soybean fields. The coloring indicates the direction and strength of the correlation coefficients. Only those correlations with $P < 0.05$ are shown. Taxonomic information at order, family, and genus levels for each corresponding OTU is displayed to the left. "NA" indicates unclear classification at the corresponding taxonomic level. AN, available nitrogen; P, available phosphorus; K, available potassium; and PS, precipitation seasonality.

were enriched in *Ensifer* species, whereas genes involved in acid adaptations were enriched in *Bradyrhizobium* species (Tian et al., 2012). Environmental (edaphic, climatic) factors were more important than spatial variables in determining the NRE subcommunities, suggesting the paramount influence of deterministic processes over stochastic processes (Lindstrom and Langenheder, 2012).

Our study also revealed opposing niche preferences by the Proteobacteria and Firmicutes NRE members, which were driven more by climatic than by edaphic factors (Figs. 1 and 3). Generally, both precipitation and temperature are higher in China's Firmicutes-dominated southern region than in its Proteobacteria-dominated northern and central regions. The global biogeographic pattern of soil fungi suggests regional abiotic conditions probably stimulated evolutionary radiations in certain regions (Tedersoo et al., 2014). Therefore, the respective dominance of Proteobacteria and Firmicutes in different regions of China may have been stimulated by their climatic factors. Furthermore, the influence of significant environmental variables from the db-RDA plots was confirmed by the correlations between core taxa and environmental variables. However, our results also highlight that less influential environmental variables upon communities tend to be more important for specific species. For example, soil pH was positively correlated with many non-rhizobial core taxa, yet it did not influence whole NRE subcommunities.

That the rhizobial and NRE subcommunities were not influenced by the same factors also points to their different ecological niches and roles. Generally, rhizobial strains found in nodules function as mutualistic N2-fixers or as parasites (i.e., fixing little or no N2) (Denison and Kiers, 2004). *Bradyrhizobium* and *Ensifer* genera have inverse pH preferences. (Tian et al., 2012), and both had OTUs widely distributed among our sampling sites, with those found in low abundances perhaps occupying

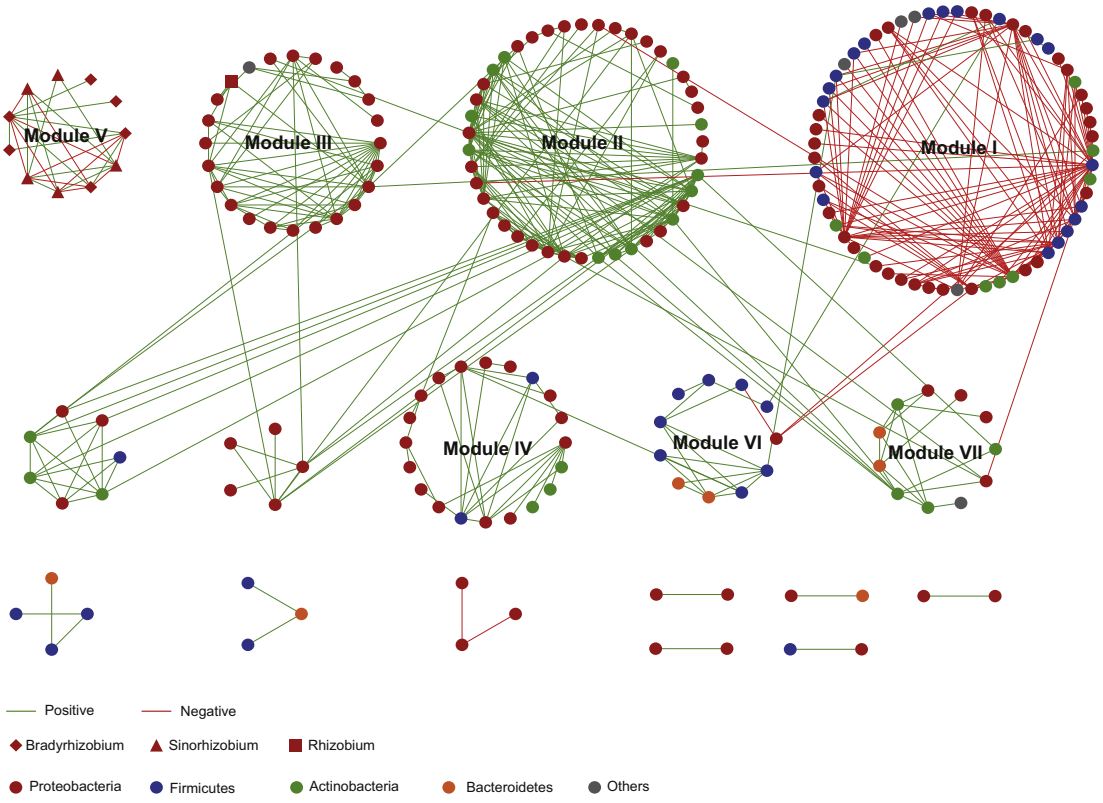


Fig. 5. The bacterial co-occurrence network for soybean nodules, based on the random matrix theory.

parasitic niches in soybean nodules in alkaline and neutral to acidic soils, respectively. Although the ecological roles of most NRE are not well known, some in the genera *Pseudomonas*, *Agrobacterium*, *Bacillus*, *Burkholderia*, *Microbacterium*, and *Streptomyces* reportedly exert plant growth-promoting activities in vivo (Egamberdieva et al., 2017; Palaniappan et al., 2010; Stajković et al., 2009; Sun et al., 2018; Tokala et al., 2002). Interestingly, the denitrifying *Hyphomicrobium* were also identified as core NRE in our study, suggesting that in addition to nitrogen fixation, other nitrogen cycle processes may operate in soybean nodules. Since the *Methylibium* are able to degrade a variety of carbon substrates (Imai et al., 2011; Joshi et al., 2015; Szabo et al., 2015), they may fulfill critical roles in microbial food web by transferring metabolic products to co-occurring NRE in nodules.

It should be noted that plant cultivars and agricultural practices (i.e., crop rotation, fertilizer and pesticide treatments) could also influence root-associated bacterial communities. In cowpea, pure host genotype did not influence the NRE communities (Leite et al., 2016). Recently, a two-decade study demonstrated differences in soil microbial diversity and community composition dependent on agricultural management, mainly owing to the fertilizer type quality used (organic or mineral) but not pesticides (Hartmann et al., 2015). Although all the soybean fields we sampled were cultivated under conventional management (i.e., mineral fertilizers used), among-site differences in chemical fertilizer type and quantity were not considered. Additional studies that explicitly include agriculture management practices should provide comprehensive insight into the biogeography of nodule endophytes.

4.3. Rhizobial and non-rhizobial taxa grouped separately in the co-occurrence network

The empirical network topologies differed from random networks, indicating nonrandom co-occurrence patterns for the nodule bacterial communities. This result also emphasizes the importance of deterministic processes, including bacterial interactions and niche differentiation, in structuring nodule bacterial communities (Ju et al., 2014). The negative associations between *Ensifer* and *Bradyrhizobium*, and between Proteobacteria and Firmicutes members were consistent with their environmental preferences discussed above. In particular, since it produces antimicrobial compounds (Barka et al., 2016), antagonistic interactions may underpin Actinobacteria's negative association with other members.

Highly interconnected modular structures are not unusual in nonrandom bacterial networks (Shi et al., 2016; Wu et al., 2016). Modules could reflect ecological niche overlap, resource partitioning, and phylogenetic relatedness (Olesen et al., 2007). Interestingly, the rhizobia (*Ensifer* and *Bradyrhizobium*) form a single module that was grouped separately from NRE. As we discussed earlier, rhizobia and NRE likely have different ecological niches and roles in root nodules. The separate grouping of rhizobia and NRE may also be attributed to resource partitioning. Rhizobia in nodules primarily use dicarboxylates delivered from roots as their carbon source (Udvardi and Poole, 2013). We presume that some NRE strains may be saprophytic bacteria that contribute to nodule decomposition, since 53.7% of the NRE we identified were Proteobacteria. This taxon uses a broad range of root-derived carbon substrates and has prominent members of putative saprophytes in roots (Bulgarelli et al., 2012; Philippot et al., 2013).

5. Conclusions

In summary, the rhizobial and NRE subcommunities displayed different biogeographic patterns in the soybean fields across China. The NRE have a lower probability of dispersal (i.e., are more dispersal limited) than do the rhizobia, and rhizobial and NRE subcommunities were significantly correlated with different environmental and spatial variables. Compositionally, the rhizobial subcommunities were grouped into *Ensifer*- and *Bradyrhizobium*-dominated clusters that were related to soil pH; the NRE subcommunities grouped into Proteobacteria- and

Firmicutes-dominated clusters that were more influenced by environmental than by spatial factors. Network analysis revealed that rhizobia and NRE grouped separately, showing no correlation with each other, which suggested niche sharing differed between these two subcommunities in soybean nodules. Taken together, these results provide a new perspective on bacterial distributions and community assembly in root nodules.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2018.06.240.

Acknowledgements

We thank Jun Zhang, Yao Liu, Yanqing Guo, and Lei Liu for their help with the field sampling. This study was funded by the National Key Research & Development Program (2016YFD0200308) and the National Natural Science Foundation of China (41671261 and 31672241). The authors declare no conflicts of interest.

References

Bao, S., 2000. Soil and Agricultural Chemistry Analysis. Agriculture Publication, Beijing, pp. 355–356.
Barberan, A., Ramirez, K.S., Leff, J.W., Bradford, M.A., Wall, D.H., Fierer, N., 2014. Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. Ecol. Lett. 17, 794–802.
Barka, E.A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Meier-Kolthoff, J.P., et al., 2016. Taxonomy, physiology, and natural products of Actinobacteria. Microbiol. Mol. Biol. Rev. 80, 1–43.
Biate, D.L., Kumar, L.V., Ramadoss, D., Kumari, A., Naik, S., Reddy, K.K., et al., 2014. Genetic Diversity of Soybean Root Nodulating Bacteria. Springer International Publishing.
Bissett, A., Brown, M.V., Siciliano, S.D., Thrall, P.H., 2013. Microbial community responses to anthropogenically induced environmental change: towards a systems approach. Ecol. Lett. 16 (Suppl 1), 128–139.
Borcard, D., Gillet, F., Legendre, P., 2011. Numerical Ecology with R. Springer Science & Business Media.
Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., et al., 2012. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. Nature 488, 91–95.
Cao, Y., Wang, E.-T., Zhao, L., Chen, W.-M., Wei, G.-H., 2014. Diversity and distribution of rhizobia nodulated with *Phaseolus vulgaris* in two ecoregions of China. Soil Biol. Biochem. 78, 128–137.
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.
Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., et al., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6, 1621–1624.
Chidebe, I.N., Jaiswal, S.K., Dakora, F.D., 2018. Distribution and phylogeny of microsymbionts associated with cowpea (*Vigna unguiculata*) nodulation in three Agroecological Regions of Mozambique. Appl. Environ. Microbiol. 84.
Clarke, K., Gorley, R., 2015. PRIMER v7: User Manual/Tutorial. PRIMER-EPlymouth.
de Menezes, A.B., Prendergast-Miller, M.T., Richardson, A.E., Toscas, P., Farrell, M., Macdonald, L.M., et al., 2015. Network analysis reveals that bacteria and fungi form modules that correlate independently with soil parameters. Environ. Microbiol. 17, 2677–2689.
De Meyer, S.E., De Beuf, K., Vekeman, B., Willems, A., 2015. A large diversity of non-rhizobial endophytes found in legume root nodules in Flanders (*Belgium*). Soil Biol. Biochem. 83, 1–11.
Deng, Z.S., Zhao, L.F., Kong, Z.Y., Yang, W.Q., Lindstrom, K., Wang, E.T., et al., 2011. Diversity of endophytic bacteria within nodules of the *Sphaerophysa salsula* in different regions of loess plateau in China. FEMS Microbiol. Ecol. 76, 463–475.
Deng, Y., Jiang, Y.H., Yang, Y., He, Z., Luo, F., Zhou, J., 2012. Molecular ecological network analyses. BMC Bioinf. 13, 113.
Denison, R.F., Kiers, E.T., 2004. Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. FEMS Microbiol. Lett. 237, 187–193.
Donnarumma, F., Bazzicalupo, M., Blazinkov, M., Mengoni, A., Sikora, S., Babic, K.H., 2014. Biogeography of *Sinorhizobium meliloti* nodulating alfalfa in different Croatian regions. Res. Microbiol. 165, 508–516.
Dudeja, S.S., Giri, R., Saini, R., Suneja-Madan, P., Kothe, E., 2012. Interaction of endophytic microbes with legumes. J. Basic Microbiol. 52, 248–260.
Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998.
Egamberdieva, D., Wirth, S.J., Shurigin, V.V., Hashem, A., Abd Allah, E.F., 2017. Endophytic bacteria improve plant growth, symbiotic performance of chickpea (*Cicer arietinum* L.) and induce suppression of root rot caused by *Fusarium solani* under salt stress. Front. Microbiol. 8, 1887.
Fan, K., Cardona, C., Li, Y., Shi, Y., Xiang, X., Shen, C., et al., 2017. Rhizosphere-associated bacterial network structure and spatial distribution differ significantly from bulk soil in wheat crop fields. Soil Biol. Biochem. 113, 275–284.
Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. Nat. Rev. Microbiol. 10, 538–550.
Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., et al., 2012. Microbial co-occurrence relationships in the human microbiome. PLoS Comput. Biol. 8, e1002606.
Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., Martiny, J.B., 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nat. Rev. Microbiol. 10, 497–506.
Hartmann, M., Frey, B., Mayer, J., Mader, P., Widmer, F., 2015. Distinct soil microbial diversity under long-term organic and conventional farming. ISME J. 9, 1177–1194.
Imai, S., Ichikawa, K., Muramatsu, Y., Kasai, D., Masai, E., Fukuda, M., 2011. Isolation and characterization of *Streptomyces*, *Actinoplanes*, and *Methylibium* strains that are involved in degradation of natural rubber and synthetic poly(cis-1,4-isoprene). Enzym. Microb. Technol. 49, 526–531.
Joshi, G., Schmidt, R., Scow, K.M., Denison, M.S., Hristova, K.R., 2015. Gene mdpC plays a regulatory role in the methyl-tert-butyl ether degradation pathway of *Methylibium petroleiphilum* strain PM1. FEMS Microbiol. Lett. 362.
Ju, F., Xia, Y., Guo, F., Wang, Z., Zhang, T., 2014. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. Environ. Microbiol. 16, 2421–2432.
Konstantinidis, K.T., Tiedje, J.M., 2004. Trends between gene content and genome size in prokaryotic species with larger genomes. Proc. Natl. Acad. Sci. U. S. A. 101, 3160–3165.
Leite, J., Fischer, D., Rouws, L.F., Fernandes-Junior, P.I., Hofmann, A., Kublik, S., et al., 2016. Cowpea nodules harbor non-rhizobial bacterial communities that are shaped by soil type rather than plant genotype. Front. Plant Sci. 7, 2064.
Li, Y., Guan, R., Liu, Z., Ma, Y., Wang, L., Li, L., et al., 2008. Genetic structure and diversity of cultivated soybean (*Glycine max* (L.) Merr.) landraces in China. Theor. Appl. Genet. 117, 857–871.
Li, Q.Q., Wang, E.T., Zhang, Y.Z., Zhang, Y.M., Tian, C.F., Sui, X.H., et al., 2011. Diversity and biogeography of rhizobia isolated from root nodules of *Glycine max* grown in Hebei Province, China. Microb. Ecol. 61, 917–931.
Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al., 2015. Ocean plankton. Determinants of community structure in the global plankton interactome. Science 348, 1262073.
Lindstrom, E.S., Langenheder, S., 2012. Local and regional factors influencing bacterial community assembly. Environ. Microbiol. Rep. 4, 1–9.
Liu, L., Yang, J., Yu, Z., Wilkinson, D.M., 2015. The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. ISME J. 9, 2068–2077.
Lu, Y.L., Chen, W.F., Wang, E.T., Guan, S.H., Yan, X.R., Chen, W.X., 2009. Genetic diversity and biogeography of rhizobia associated with Caragana species in three ecological regions of China. Syst. Appl. Microbiol. 32, 351–361.
Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., et al., 2016. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. ISME J. 10, 1891–1901.
Ma, B., Dai, Z., Wang, H., Dsouza, M., Liu, X., He, Y., et al., 2017. Distinct biogeographic patterns for archaea, bacteria, and fungi along the vegetation gradient at the continental scale in Eastern China. mSystems 2, e00174-16.
Magoc, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963.
Martínez-Hidalgo, P., Hirsch, A.M., 2017. The nodule microbiome: N2-fixing rhizobia do not live alone. Phytobiomes 1, 70–82.
Martinez-Hidalgo, P., Galindo-Villardon, P., Trujillo, M.E., Igual, J.M., Martinez-Molina, E., 2014. Micromonospora from nitrogen fixing nodules of alfalfa (*Medicago sativa* L.). A new promising plant probiotic bacteria. Sci. Rep. 4, 6389.
Martiny, J.B., Bohannan, B.J., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., et al., 2006. Microbial biogeography: putting microorganisms on the map. Nat. Rev. Microbiol. 4, 102–112.
Nemergut, D.R., Costello, E.K., Hamady, M., Lozupone, C., Jiang, L., Schmidt, S.K., et al., 2011. Global patterns in the biogeography of bacterial taxa. Environ. Microbiol. 13, 135–144.
Olesen, J.M., Bascompte, J., Dupont, Y.L., Jordano, P., 2007. The modularity of pollination networks. Proc. Natl. Acad. Sci. U. S. A. 104, 19891–19896.
Palaniappan, P., Chauhan, P.S., Saravanan, V.S., Anandham, R., Sa, T., 2010. Isolation and characterization of plant growth promoting endophytic bacterial isolates from root nodule of Lespedeza sp. Biol. Fertil. Soils 46, 807–816.
Peix, A., Ramírez-Bahena, M.H., Velázquez, E., Bedmar, E.J., 2014. Bacterial associations with legumes. Crit. Rev. Plant Sci. 34, 17–42.
Philippot, L., Raaijmakers, J.M., Lemanceau, P., van der Putten, W.H., 2013. Going back to the roots: the microbial ecology of the rhizosphere. Nat. Rev. Microbiol. 11, 789–799.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, p. 2016.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.
Shi, S., Nuccio, E.E., Shi, Z.J., He, Z., Zhou, J., Firestone, M.K., 2016. The interconnected rhizosphere: high network complexity dominates rhizosphere assemblages. Ecol. Lett. 19, 926–936.
Stajković, O., De Meyer, S., Miličić, B., Willems, A., 2009. Isolation and characterization of endophytic non-rhizobial bacteria from root nodules of alfalfa (*Medicago sativa* L.). Botanica Serbica 33, 107–114.
Steele, J.A., Countway, P.D., Xia, L., Vigil, P.D., Beman, J.M., Kim, D.Y., et al., 2011. Marine bacterial, archaeal and protistan association networks reveal ecological linkages. ISME J. 5, 1414–1425.
Sun, W., Xiong, Z., Chu, L., Li, W., Soares, M.A., White Jr., J.F., et al., 2018. Bacterial communities of three plant species from Pb-Zn contaminated sites and plant-growth

promotional benefits of endophytic *Microbacterium* sp. (strain BXGe71). J. Hazard. Mater. https://doi.org/10.1016/j.jhazmat.2018.02.003.
Szabo, Z., Gyula, P., Robotka, H., Bato, E., Galik, B., Pach, P., et al., 2015. Draft genome sequence of *Methylibium* sp. strain T29, a novel fuel oxygenate-degrading bacterial isolate from Hungary. Stand. Genomic Sci. 10, 39.
Tariq, M., Hameed, S., Yasmeen, T., Zahid, M., Zafar, M., 2014. Molecular characterization and identification of plant growth promoting endophytic bacteria isolated from the root nodules of pea (*Pisum sativum* L.). World J. Microbiol. Biotechnol. 30, 719–725.
Tedersoo, L., Bahram, M., Polme, S., Koljalg, U., Yorou, N.S., Wijesundera, R., et al., 2014. Fungal biogeography. Global diversity and geography of soil fungi. Science 346, 1256688.
Tian, C.F., Zhou, Y.J., Zhang, Y.M., Li, Q.Q., Zhang, Y.Z., Li, D.F., et al., 2012. Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. Proc. Natl. Acad. Sci. U. S. A. 109, 8629–8634.
Tokala, R.K., Strap, J.L., Jung, C.M., Crawford, D.L., Salove, M.H., Deobald, L.A., et al., 2002. Novel plant-microbe rhizosphere interaction involving *Streptomyces lydicus* WYEC108 and the pea plant (*Pisum sativum*). Appl. Environ. Microbiol. 68, 2161–2171.
Udvardi, M., Poole, P.S., 2013. Transport and metabolism in legume-rhizobia symbioses. Annu. Rev. Plant Biol. 64, 781–805.
Verastegui-Valdes, M.M., Zhang, Y.J., Rivera-Orduna, F.N., Cheng, H.P., Sui, X.H., Wang, E.T., 2014. Microsymbionts of *Phaseolus vulgaris* in acid and alkaline soils of Mexico. Syst. Appl. Microbiol. 37, 605–612.
Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73, 5261–5267.
Wang, L., Cao, Y., Wang, E.T., Qiao, Y.J., Jiao, S., Liu, Z.S., et al., 2016. Biodiversity and biogeography of rhizobia associated with common bean (*Phaseolus vulgaris* L.) in Shaanxi Province. Syst. Appl. Microbiol. 39, 211–219.
Wu, L., Yang, Y., Chen, S., Zhao, M., Zhu, Z., Yang, S., et al., 2016. Long-term successional dynamics of microbial association networks in anaerobic digestion processes. Water Res. 104, 1–10.
Xiao, X., Chen, W., Zong, L., Yang, J., Jiao, S., Lin, Y., et al., 2017. Two cultivated legume plants reveal the enrichment process of the microbiome in the rhizocompartments. Mol. Ecol. 26, 1641–1651.
Xu, L., Zhang, Y., Wang, L., Chen, W., Wei, G., 2014. Diversity of endophytic bacteria associated with nodules of two indigenous legumes at different altitudes of the Qilian Mountains in China. Syst. Appl. Microbiol. 37, 457–465.
Yan, J., Han, X.Z., Tian, C.F., Sui, X.H., Singh, R.P., Jiang, G.M., et al., 2018. Rhizobial biogeography and inoculation application to soybean in four regions across China. J. Appl. Microbiol. https://doi.org/10.1111/jam.13897.
Zhang, Y.M., Li Jr., Y., Chen, W.F., Wang, E.T., Tian, C.F., Li, Q.Q., et al., 2011. Biodiversity and biogeography of rhizobia associated with soybean plants grown in the North China Plain. Appl. Environ. Microbiol. 77, 6331–6342.
Zhang, B., Zhang, J., Liu, Y., Shi, P., Wei, G., 2018. Co-occurrence patterns of soybean rhizosphere microbiome at a continental scale. Soil Biol. Biochem. 118, 178–186.
Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., Zhi, X., 2010. Functional molecular ecological networks. MBio 1.