Seed-borne, endospheric and rhizospheric core microbiota as predictor for plant functional traits across rice cultivars are dominated by deterministic processes

Junjie Guo, Ning Ling, Yong Li, Kaisong Li, Huiling Ning, Qirong Shen, Shiwei Guo, Philippe Vandenkoornhuyse

DR NING LING (Orcid ID : 0000-0003-1250-4073)

DR YONG LI (Orcid ID : 0000-0002-3595-6673)

Article type : - Regular Manuscript

DR NING LING (ORCID ID: 0000-0003-1250-4073)

DR YONG LI (ORCID ID: 0000-0002-3595-6673)

Seed-borne, endospheric and rhizospheric core microbiota as predictor for plant functional traits across rice cultivars are dominated by deterministic processes

Brief heading: Core microbiota as predictor for plant functional traits

Junjie Guo[a], Ning Ling[a, b, *], Yong Li[c], Kaisong Li[a], Huiling Ning[a], Qirong Shen[a], Shiwei Guo[a], Philippe Vandenkoornhuyse[b]

a, Jiangsu Provincial Key Lab for Solid Organic Waste Utilization, Jiangsu Collaborative Innovation Center for Solid Organic Waste Resource Utilization, Nanjing Agricultural University, Nanjing 210095, Jiangsu, China

b, Université de Rennes 1, CNRS, UMR 6553 EcoBio, campus Beaulieu, Avenue du Général Leclerc, 35042 Rennes Cedex, France

c, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, Hubei, 430070 China

* Corresponding author: Ning Ling

E-mail: nling@njau.edu.cn

Tel.: (+86) 25 84396853

Received: 7 December 2020

Accepted: 17 February 2021

ORCID information:

Ning Ling: https://orcid.org/0000-0003-1250-4073

Yong Li: https://orcid.org/0000-0002-3595-6673

Summary

- A host-plant and its associated microbiota depend on one another. However, the assembly process and the functioning of host-associated microbiota are poorly understood.
- Herein, rice was used as model plant to investigate the assemblage of bacterial microbiota, including those in the seed, root endosphere and rhizosphere. We also assessed the degree to which endosphere and rhizosphere communities were influenced by vertical transmission through seed and identified the core microbes that potentially associate with plant phenotypic properties.
- Plant microhabitat, rather than subspecies type, was the major driver shaping plant-associated bacterial microbiota. Deterministic processes were primarily responsible for community assembly in all microhabitats. The influence of vertical transmission from seed to root-associated bacterial communities appeared to be quite weak (endosphere) or even absent (rhizosphere). A core microbial community composed of 15 generalist species persisted across different microhabitats and represented key connectors in networks. Host-plant functional traits were linked to the relative abundance of these generalist core microbes and could be predicted from them using machine learning algorithms.
- Overall, bacterial microbiota is assembled by host-plant interactions in deterministic-based manner. This study enhances our understanding of driving mechanism and associations of microbiota in various plant microhabitats and provides new perspectives to improve plant performance.

Keywords: Seed habitat; Root endosphere; Rhizosphere; Bacteria; Community assembly; Plant-microbiota association

Introduction

The current concept of holobiont supposes that a host-plant and its associated communities of microorganism form altogether the plant individual, whose ecology and evolution thus supposed inseparably woven together (Bordenstein & Theis, 2015; Vandenkoornhuyse *et al.*, 2015; Theis *et al.*, 2016). Virtually all tissues of a host-plant can be view as microhabitats for diverse microbiota that can determine plant health and productivity by providing additive ecological functions to help regulate the holobiont fitness and resistance to environmental changes (Turner *et al.*, 2013; Vandenkoornhuyse *et al.*, 2015). More specifically, the host-plant serves habitats and carbon sources supply for the microbiota (Sanchez-Canizares *et al.*, 2017; Shade *et al.*, 2017). Plant-associated microbiota provides its host different ecological services including nutrient and water acquisition, plant resistance to biotic and abiotic stresses (Turner *et al.*, 2013; Berg *et al.*, 2016; Compant *et al.*, 2019). Microbiota assemblages inhabiting these microhabitats may highly vary due to the different host-plant niche specific adaptations (Kroll *et al.*, 2017). Remarkably, the influence of microhabitats on plant-associated microbiota is suggested greater than host-plant cultivar and soil type in some cases (Edwards *et al.*, 2015; Coleman-Derr *et al.*, 2016; Hamonts *et al.*, 2018). Understanding the assembly rules of plant-associated microbiota is a current important topic, both to develop the mechanistic insight into holobiont formation and to provide important knowledge required for sustainable agriculture.

Understanding of fundamental eco-evolutionary processes that govern microbial community is certainly one of the most burning challenges in plant-associated microbiota assemblage research (Vandenkoornhuyse *et al.*, 2015; Sanchez-Canizares *et al.*, 2017). The conceptual framework of community ecology provides a testable backbone for unveiling how eco-evolutionary processes mediate the establishment of plant-associated microbiota (Dini-Andreote & Raaijmakers, 2018; Cordovez *et al.*, 2019). As the other communities, assembly in plant-associated microbiota is mediated by the interplay of four fundamental eco-evolutionary processes, namely, selection, dispersal, diversification (or speciation), and ecological drift (Zhou & Ning, 2017; Cordovez *et al.*, 2019; Fitzpatrick *et al.*, 2020). Selection and ecological drift are unambiguously deterministic and stochastic, respectively, while dispersal and diversification (or speciation) encompass both

deterministic and stochastic components (Zhou & Ning, 2017). In other words, both deterministic and stochastic components, which are embedded in these four fundamental processes, work in combination to control microbial community assembly (Zhou & Ning, 2017; Ning *et al.*, 2019). Several investigations have provided qualitative insight into how eco-evolutionary processes manipulate plant-associated microbial communities inhabiting different microhabitats (Lundberg *et al.*, 2012; Mendes *et al.*, 2014; Edwards *et al.*, 2015; Niu *et al.*, 2017). For instance, the microbiota assembly of emerging root system maybe prone to influenced by priority effects, while the microbiota that colonized in well-established vegetatively growing plant root and/or rhizosphere should be mainly structured by gradual changed environmental selection. Nevertheless, quantitative assessment regarding the relative importance of deterministic and stochastic processes that governing the assemblage of plant-associated microbiota is not well understood.

Traditionally, the members of plant-associated microbiota have been considered predominantly recruited horizontally during plant life from the surrounding environment, with soil being the main source and air only exerting a minute contribution (Vandenkoornhuyse *et al.*, 2015; Sanchez-Canizares *et al.*, 2017). Therefore, the rhizosphere between soil and roots is a highly dynamic below-ground environment with diverse soil-derived microbiota and represents a growth chamber or reservoir for root-endospheric microbiota recruitment (Hardoim *et al.*, 2015; Vandenkoornhuyse *et al.*, 2015). Emerging evidence has demonstrated that in addition to the horizontal transmission of microbiota from the surrounding environment, plant-associated microbiota is also transferred vertically directly from parent to offspring via stolon in clonal plants (Vannier *et al.*, 2018) and via seeds (Schardl, 2001; Truyens *et al.*, 2015; Frank *et al.*, 2017; Shade *et al.*, 2017). Microbiota inhabiting the seed may have the advantage of introducing itself as the initial microbiota to the plant root and rhizosphere, as they are already adapted to the plant tissue and readily access to the available space and nutrients in the host-plant during seed germination (Truyens *et al.*, 2015; Kong *et al.*, 2019). Thus, the seed could be expected to serve as a potential repository for the microbiota existing in the plant root or rhizosphere. Advances in understanding plant-associated microbiota assemblage patterns have been made considering microhabitats

represented by rhizosphere, root endosphere, leaf surface, flowers and seed habitat (Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012; Vorholt, 2012; Bringel & Couee, 2015; Edwards *et al.*, 2015; Beckers *et al.*, 2017; Cregger *et al.*, 2018; Fitzpatrick *et al.*, 2018; Kim *et al.*, 2020). However, the associations of microbiota between different plant microhabitats remain vague.

Both plant phenotypic properties and fitness can be mediated by the mutualistic microbial epiphytes and endophytes (Vandenkoornhuyse *et al.*, 2015), with plant-associated microbiota providing additional nutritional and defence capabilities or alter existing plant biochemical pathways (Pineda *et al.*, 2010; Friesen *et al.*, 2011; Cosme *et al.*, 2016; Ravanbakhsh *et al.*, 2020). Recent work show that plant functional traits (*i.e.*, morphological or physiological characteristics), such as specific leaf area, aboveground biomass, and specific root length, could be manipulated by the plant-associated microbiota composition (Perez-Jaramillo *et al.*, 2017; Fitzpatrick *et al.*, 2018; Ravanbakhsh *et al.*, 2020), especially by the root-associated microbiota (Vandenkoornhuyse *et al.*, 2015; Ravanbakhsh *et al.*, 2020). However, plant-associated microbiota is highly complex and diverse, with most community members are often supposed transient or opportunistic, which might contribute little to the host-plant biological function (Busby *et al.*, 2017). Current research suggests that transient associations should be filtered out to refine focus on the core microbiota, which are comprised of stable and persistent microbial taxa and have greater likelihood of influencing host-plant phenotype (Vandenkoornhuyse *et al.*, 2015; Busby *et al.*, 2017; Lemanceau *et al.*, 2017). From an ecological perspective, generalist species can be defined as taxa that are able to adapt to diverse habitats or environments (Sriswasdi *et al.*, 2017; Xu *et al.*, 2020). Thus, the wide and stable-distributed generalist species could be considered as members of the core microbiota, with high occupancy that persist across multiple hosts and microhabitats (Kokou *et al.*, 2019). Although the importance of generalist core microbiota members has been recognized, their connectivity in microbial interaction network and capacity of predicting host-plant phenotype are not yet understood.

Here, rice (*Oryza sativa* L.) was selected to be the model plant to narrow critical knowledge gap regarding host-microbiota systems. An original analysis of seed, root endosphere and rhizosphere bacterial microbiota was performed, characterizing the community assemblages and

quantifying the contribution of assembly processes with 99 rice varieties under controlled conditions. Meanwhile, the extent to which bacteria were transmitted from the seed to the root endosphere and/or rhizosphere communities as well as from the rhizosphere to the root endosphere communities were addressed. In addition, we identified generalist core microbial populations that are widespread across multiple plant microhabitats, and inferred the ecological role of core species in microbial interaction networks and their association with measured rice phenotypic characteristics. Our study sought to test the following hypotheses: (i) the bacterial microbiota assemblage of seed habitat, root endosphere and rhizosphere is dominated by deterministic processes of assembly; (ii) both vertical- and horizontal-transmission contribute to the observed differences in the root-associated bacterial communities; (iii) the plant phenotypes could be predicted from the core bacterial microbiota, which is represented by a small community but highly connected to other microbes. This work was developed to dive beyond the descriptive understanding of the plant-associated microbiota by developing predictive interpretations.

Materials and Methods

Rice seed collection and recovery of plant seedlings grown under axenic condition

Rice seeds from 99 cultivated varieties (belonged to different subspecies and originated from ten countries) were obtained from Huazhong Agricultural University, China (Figure S1). All rice accessions were stored in dry conditions in the dark until the following experiments. Given that most of the microbes inhabiting seeds are generally in a dormant stage (Truyens *et al.*, 2015; Shade *et al.*, 2017), the seed-associated microbiota was determined by recovering the activated microorganisms from plant seedlings that grown under axenic condition. In brief, rice seeds were surface sterilized by soaking in 15% NaOCl for 30 min and then washed several times with sterile distilled water. To verify the effectiveness of sterilization, 100 μL water from the final rinse was spread on Luria-Bertani (LB) plates and incubated at 37 °C. No bacterial colony was formed on LB plates. The surface-sterilized seeds were germinated at 25 °C in the dark for 48 h on sterile Murashige-Skoog (MS) medium containing 0.2% (w/v) phytagel (single seed per culture tube and three replicates per cultivated variety). After pre-germination, individual seed was continued to

incubated in each axenic tube under artificial light. Plant seedlings were collected at the same date after two weeks.

Rice cultivation, plant functional traits measurement and root-associated sample collection

After surface sterilization, the seeds were germinated at 25 °C in the dark for 48 h on filter paper wetted with sterile water. The pregerminated seeds continued to germinate in an artificial light growth chamber for seven days. The greenhouse experiment was conducted under non-sterile condition. Paddy soil used in this experiment was collected from the rice field on the outskirts of Nanjing, Jiangsu Province, China. The soil was air-dried, sieved through a 4-mm sieves and then homogenized thoroughly before adding to pots. For each rice variety, two seedlings were transplanted into a pot filled with soil (6 pots per variety). A total of 594 pots (99 cultivated varieties × 6 pots) were arranged randomly in a glasshouse and irrigated every day with distilled water to keep the soil submerged. Nutrient solution was supplied weekly to ensure plant growth staring six weeks after transplantation. Finally, half of these samples (3 pots) were used for plant analysis, and the other half were used for root-associated microbial community profiling.

Plant samples were harvested 9 weeks after seedling transplantation, at the tillering stage. To characterize the functional traits of the plants, several morphological and physiological traits were determined. Both the aboveground and underground dried-biomass were measured. Dried plant tissues were digested with H_2SO_4-H_2O_2 at 260-280 °C. Tissue nitrogen concentrations were determined with a continuous-flow auto-analyzer (AA3, SEAL Analytical, Norderstedt, Germany). Immediately before sampling, the net photosynthetic rate of newly expanded leaves of rice planted in each pot was measured by using an Li-Cor 6400 portable photosynthesis open system (LI-COR, Lincoln, NE, USA).

Root-associated samples from each plant were partitioned into two rhizocompartments (*i.e.,* root endosphere and rhizosphere compartments) following the procedure used previously (Leff *et al.*, 2017; Niu *et al.*, 2017), with modifications. Briefly, the entire root was vigorously shaken to remove loose soil and then transferred into a 50-ml sterile Falcon tube with 15 ml of sterile water. The rhizosphere compartment (root-adhering soil) was separated by vortexing the roots, centrifuging the soil suspension and collecting the resulting pellet. For collection of the

endosphere compartment, roots were thoroughly washed in sterile water to further discard any remaining soil. The water-washed roots were sterilized by dipping in 3% (v/v) sodium hypochlorite for 2 min, followed by dipping in 70% (v/v) ethanol for 2 min, and finally by rinsing 3 times with sterile distilled water to remove the rhizoplane microorganisms.

DNA extraction, 16S rRNA gene sequencing, and bioinformatic analysis

DNA extractions were performed from 250 mg of homogenized plant seedling sample (grown under axenic condition, 99 cultivated varieties × 3 replicates, n=297), root endosphere sample (99 cultivated varieties × 3 replicates, n=297) or rhizosphere sample (99 cultivated varieties × 3 replicates, n=297) according to the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) procedure. Isolated DNA was stored at −80 °C until further analyses. Before sequencing, three replicates of seed-derived, endospheric and rhizospheric DNA samples for each variety were pooled to one composite sample, respectively (Xu *et al.*, 2018).

The primer set 799F (5′- AAC MGG ATT AGA TAC CCK G -3′) and 1193R (5′- ACG TCA TCC CCA CCT TCC -3′), which exhibited low affinity for non-target DNA such as plastid (mostly chloroplast) DNA and mitochondrial DNA (Beckers *et al.*, 2016), was used to amplify the V5−V7 region of the bacterial 16S rRNA gene (Bulgarelli *et al.*, 2012). The 10-μl PCR mixture contained 1 μl of 10× TopTaq Buffer, 0.8 μl of dNTPs (2.5 mM), 1 μl of both forward and reverse primers (2 μM), 0.2 μl TopTaq DNA Polymerase (Qiagen) (5U/μl), 1 μl of DNA template, and 6 μl of ddH_2O. Amplification was performed with an initial denaturation of 2 min at 94 °C, followed by 28 cycles of 20 s at 94 °C (denaturation), 30 s at 55 °C (annealing), 1 min at 72 °C (extension), and a final extension at 72 °C for 2 min. Paired-end (2×250) sequencing was performed on the HiSeq platform (Illumina, USA) by Genesky Biotechnology Co., Ltd. (Shanghai, China).

Raw paired-end sequences were processed using USEARCH v.11.0 (Edgar, 2010). After forward and reverse reads were merged, low-quality sequences (length < 360 bp, total expected errors > 0.5) were filtered. The unique read sequences were clustered into operational taxonomic units (OTUs) by UPARSE algorithm with a 97% identity threshold. Finally, the taxonomic affiliation of representative sequence was performed using the RDP naïve Bayesian classifier (Wang *et al.*, 2007) with a minimum bootstrap threshold of 50%. The samples sequencing depth

was normalized to 16,204 sequences.

Bacterial community diversity and structure analyses

To assess bacterial α-diversity, both observed richness and Shannon indices were calculated on the normalized contingency table using MOTHUR (Schloss *et al.*, 2009). Kruskal-Wallis one-way analysis of variance was performed to compare α-diversity metrics between different subspecies within each plant microhabitat or between different plant microhabitats. Principal coordinate analysis (PCoA) was conducted to visualize the β-diversity pattern of bacterial communities between samples with different subspecies within each plant microhabitat based on the Bray-Curtis distance. Moreover, PCoA was also performed to visualize the β-diversity (both Bray-Curtis and Jaccard distances) of the bacterial communities between samples with different plant microhabitats. Both permutational multivariate analysis of variance (PERMANOVA) and analysis of similarity (ANOSIM) were performed to test the significance of bacterial community dissimilarity.

Association analysis of microbiota inhabiting in different microhabitats

The potential sources of microbiota inhabiting in root endosphere and rhizosphere was estimated using fast expectation-maximization microbial source tracking (FEAST) (Shenhav *et al.*, 2019). Procrustes analyses were performed using the distances plots (PCoA) as input based on the matrix of bacterial communities (Bray-Curtis) assess the pairwise concordance between bacterial microbiota inhabiting in different plant microhabitats. Moreover, the relationships between taxonomic composition dissimilarities (Bray-Curtis) among plant microhabitats were estimated based on Pearson correlations using Mantel test. Differential OTU abundance analysis was used to identify the enrichment of specific taxa in each plant microhabitat (R package DESeq2) (Love *et al.*, 2014). OTUs were considered enriched if they had a log2-fold change greater than 2 and an adjusted *P*-value less than 0.05.

Estimating the stochastic ratio of community assembly

To evaluate the relative importance of deterministic and stochastic processes to bacterial community assembly, the modified stochasticity ratio (MST), a metric to estimate ecological stochasticity according to a null-model-based statistical framework, was calculated as described

previously (Guo *et al.*, 2018). The MST index reflects the modified ratio of mean expected similarity in the null model to observed similarity (Guo *et al.*, 2018). The value of MST index developed with 50% as the boundary point to divide the deterministic-dominance (<50%) and stochastic-dominance (>50%) community assembly (Ning *et al.*, 2019). The MST analyses were performed based on both Bray-Curtis and Jaccard distance by using the "NST" package of R software (Ning *et al.*, 2019).

Sloan neutral community model fitting

To validate the potential contribution of neutral processes to bacterial community assembly, a Sloan neutral model was used to predict the relationship between occurrence frequency of taxa and their relative abundance in the metacommunity (sum of all samples for each microhabitat) (Sloan *et al.*, 2006; Burns *et al.*, 2016). The model predicts that abundant taxa in the metacommunity would disperse by chance and be randomly sampled by an individual, while rare taxa are more likely to be lost in different local communities due to stochastic loss and replacement of individuals. In this model, the parameters R^2 and m values represent the fit to the neutral model and migration rate, respectively. All the model analyses were computed using the R scripts as described previously (Burns *et al.*, 2016).

Definition of generalists in plant microhabitat

The generalists and specialists in each plant microhabitat were classified according to the method as previously described (Sriswasdi *et al.*, 2017; Xu *et al.*, 2020). Herein, each individual variety of rice was considered a unique environment. The enrichment of the number of OTUs that were assigned to a particular number of environments (i.e. observed distribution) was computed. The random background distribution (*i.e.* expected distribution) was obtained by 10,000 random permutations (the number of OTUs in each environment was preserved) of the OTU-environment association map. The OTUs were classified as generalists or specialists when the number of observed OTUs exceeds the expected distribution. Complementarily, the identities of generalists and specialists were verified by using niche breadth, as generalists are distributed across a wider range of environments and have a higher niche breadth value than specialists (Xu *et al.*, 2020).

The similarity between generalist composition and total community composition in each

plant microhabitat was tested through Procrustes rotation. Phylogenetic tree analyses of generalist species were performed using an in-house pipeline (http://mem.rcees.ac.cn:8080) with PyNAST alignment and FastTree program (Feng *et al.*, 2017). The phylogenetic tree was visualized and edited using iTOL (Letunic & Bork, 2016). The functional annotations of the generalist taxa were carried out using FAPROTAX v.1.1 (Louca *et al.*, 2016).

Identification of the generalist core microbiota

The generalist core microbes were identified based on the following criteria: shared generalist microbes that were present in all plant microhabitats. Representative sequences with generalist core microbes were selected for the following phylogenetic analyses, which were conducted with MEGA X through a neighbor-joining tree using the Kimura 2-parameter distance with 1000 bootstrap replicates (Kumar *et al.*, 2018). The phylogenetic tree was drawn and annotated using iTOL (Letunic & Bork, 2016).

Generalists network construction

Co-abundance networks were constructed for bacterial generalist communities based on sequencing data. Briefly, co-variations were determined across 99 samples at each plant microhabitat to construct the networks. Only OTUs that were classified as generalist were kept for network construction. Non-random network analyses were performed using the Python module "SparCC" follow default parameters (Friedman & Alm, 2012). Statistically significant ($p < 0.05$) inferred correlations with a magnitude of >0.3 or <−0.3 were used to construct networks. To describe node-level topological features, both the degree and betweenness centrality were calculated for each node by using "igraph" package of R software (Csardi & Nepusz, 2006). The nodes belonging to top 1% (top 3% in endosphere) of degree and betweenness centrality were defined as hubs of each network (Kim *et al.*, 2020). Significant difference between the nodes belonging to generalist core microbe or generalist non-core microbe in betweenness centrality was compared with the Wilcoxon rank sum test. Networks were visualized using Gephi 0.9.2-beta software (Bastian *et al.*, 2009).

Prediction of plant phenotypes by generalist core microbes

The mean value of each measured plant functional trait (Table S1) was calculated to match plant-associated bacterial microbiota data of each rice variety. Spearman analyses were performed to test the correlations between the plant functional traits and relative abundance of generalist core microbes inhabiting in each plant microhabitat. Random forest regression was used to evaluate the explain ability of generalist core microbes to plant functional phenotypes, with relative abundance of each microbes serving as predictors for the functional phenotypes. The significance of both the model and each predictor were assessed by using the "rfutilities" (Murphy *et al.*, 2010) and "rfpermute" (Archer, 2020) packages of R software, respectively.

Results

Assemblage of plant-associated bacterial microbiota

Neither observed richness nor Shannon diversity of bacterial communities were affected by the rice subspecies type in each plant microhabitat (Figure S2a-b). Both root endosphere and rhizosphere bacterial community structure did not differ significantly across the subspecies type (Figure S2c). Only the seed bacterial community structure exhibited significant but weak difference across the subspecies type (PERMANOVA: $R^2 = 0.06$, $P < 0.02$; ANOSIM: $R = 0.06$, $P = 0.046$, Figure S2c).

In contrast, α-diversity metrics revealed a significant difference among different plant microhabitats (Figure 1a-b). Both the greatest observed richness and highest Shannon diversity occurred in the rhizosphere, while the lowest were observed in the seed habitat (Figure 1a-b). Principal coordinate analysis (PCoA) based on both Bray-Curtis and Jaccard distance revealed that the bacterial microbiota formed three distinct clusters according to the microhabitat, which indicated a clear spatial compartmentalization of the bacterial microbiota (Bray-Curtis dissimilarity: PERMANOVA: $R^2 = 0.60$, $P < 0.001$; ANOSIM: $R = 0.97$, $P < 0.001$; Jaccard dissimilarity: PERMANOVA: $R^2 = 0.44$, $P < 0.001$; ANOSIM: $R = 0.97$, $P < 0.001$) (Figure 1c and Figure S3a). At the (sub) phylum level, *Proteobacteria* (mostly *α*-, *β*-, *γ*- and *δ-Proteobacteria*), *Actinobacteria* and, to a lesser extent, *Firmicutes*, *Bacteroidetes*, *Candidatus*

Saccharibacteria, *Acidobacteria*, *Verrucomicrobia, Gemmatimonadetes* and *Chloroflexi* dominated the plant-associated bacterial assemblages (Figure 1d).

The modified stochasticity ratio (MST) was calculated based on two taxonomic metrics (both Bray-Curtis and Jaccard dissimilarities), which indicated that the bacterial community in all the plant microhabitats was more strongly driven by deterministic assembly processes (MST < 50%), in which seed habitat exhibited the lowest stochasticity ratio (Figure 1e and Figure S3b). The neutral model was well fitted to the bacterial communities for all plant microhabitats, with the lowest R^2 value in seed habitat (0.73) (Figure S3c).

Linkages of bacterial microbiota inhabiting in different plant microhabitats

The fast expectation-maximization microbial source tracking (FEAST) analysis revealed that rhizosphere bacterial community contributed 79% on average to the root endosphere bacterial community (Figure S4a). In contrast, seed habitat exhibited a much smaller contribution to the root endosphere (4.2%) and rhizosphere (0.2%) bacteria (Figure S4a). By comparing the bacterial communities between different plant microhabitats in Procrustes analyses, remarkable congruence between seed microbiota and endosphere microbiota (M^2 = 0.49, r = 0.71, P = 0.047), as well as between rhizosphere microbiota and endosphere microbiota (M^2 = 0.38, r = 0.79, P = 0.001) were found (Figure 2a). Whereas a non-significant similarity existed between seed microbiota and rhizosphere microbiota (M^2 = 0.55, r = 0.67, P = 0.334) (Figure 2a). Mantel tests also corroborated a significant and positive relationship between seed community dissimilarity and endosphere community dissimilarity, as well as between rhizosphere community dissimilarity and endosphere community dissimilarity (Figure S4b). The enrichment of specific bacterial OTUs in the seed habitat, root endosphere or rhizosphere was identified using differential OTU abundance analysis (Figure 2b). By comparing the top 30 genera, it was found that *Bacillus*, *Paenibacillus*, *Sphingomonas*, *Rhizobium*, *Saccharibacteria*, *Nocardioides*, *Ferrovibrio* genera were enriched in multiple plant microhabitats (Figure 2c).

The identified generalist core microbiota of rice plant

By comparing the observed and expected OTU-environment distribution, the generalist and specialist bacteria in each plant microhabitat were identified (Figure S5a). Within each plant

microhabitat, OTUs observed in more than 35-, 25- and 54 samples were defined as the generalists in seed habitat, root endosphere and rhizosphere respectively (Figure 3a and Figure S5a). The distinction between the specialists and generalists was further validated by using niche breadth estimates confirming as expected that generalists had a wider niche breadth than specialists (Figure S5b). The generalist microbiota in each plant microhabitat consisted of approximately 17.2-35.6% of the OTUs, yet represented 95.5 to 98.4% of the sequences (Figure 3a and Figure S6a). The generalist microbiota was tightly associated with the overall microbiota, as reflected by high congruence between the beta diversity of the identified generalist microbiota and the overall microbiota across plant microhabitats (Figure S6b). As illustrated in Figure 3b-c, the taxonomic composition and potential ecological function of generalist microbiota in different microhabitats were also obviously different.

Only fifteen generalist OTUs overlapped among the seed, endosphere and rhizosphere microhabitat, and they were identified as the members of generalist core microbiota and accounted for 6.59% to 71.88% relative abundance of the bacterial community from each microhabitat (Figure 4a). These OTUs were mainly members of the *α-Proteobacteria*, *β-Proteobacteria*, *γ-Proteobacteria*, *Actinobacteria*, and *Firmicutes* (sub) phylum (Figure 4b). In the more detailed genera level of taxonomic resolution, these OTUs belonged to *Herbaspirillum, Acidovorax*, *Stenotrophomonas*, *Pseudoxanthomonas*, *Pantoea*, *Pseudomonas*, *Rhizobium*, *Sphingomonas*, *Microbacterium*, *Paenibacillus* and *Bacillus* (Figure 4b). The microbial networks inferred within individual microhabitats showed that the values of betweenness centrality were higher for generalist core taxa than for generalist non-core taxa, indicating that core taxa were located in central positions within the network more often than other taxa (Figure 5a and Figure S7). Spearman correlation analysis revealed that the relative abundance of most generalist core microbes that inhabiting in rhizosphere were positively related to one or numerous measured plant characteristics such as plant aboveground N content, aboveground N accumulation, and total N accumulation (Figure 5b). Additionally, the relative abundance of an OTU belonging to the genus *Pseudoxanthomonas* was found to be highly correlated with plant underground N content, and the relative abundance of one OTU of the genus *Sphingomonas* was significantly correlated with

aboveground biomass and total biomass (Figure 5b). However, there was a general lack of positive relationship between plant functional traits and relative abundance of generalist core microbes inhabiting in seed and root endosphere (Figure 5b). For instance, in root endosphere, only the relative abundance of five OTUs with the phylogenetic associations of genus *Pseudomonas*, *Pseudoxanthomonas*, *Pantoea*, and *Microbacterium* were positively correlated with plant functional traits (Figure 5b). Similarly, the random forest model results suggested that the generalist core microbes explained more variation for plant traits when inhabiting in rhizosphere than in seed habitat or root endosphere (Figure 5b).

Discussion

Plant-associated microbiota is influenced by multiple drivers including host signature, microbe-microbe interaction, and environmental factors (Cordovez *et al.*, 2019; Trivedi *et al.*, 2020). Although several work has corroborated that different host-plant subspecies/cultivars/genotypes can harbor distinct rhizosphere and root endosphere bacterial microbiota (Peiffer *et al.*, 2013; Edwards *et al.*, 2015; Zhang *et al.*, 2019), the magnitude of host subspecies/cultivar/genotype-dependent effect within individual species typically appear quite weak or even absent (Lundberg *et al.*, 2012; Leff *et al.*, 2017; Hamonts *et al.*, 2018), which is consistent with our results (Figure S2) from our experimental design. It is noteworthy that the seed bacterial microbiota is significantly dependent on subspecies type (Figure S2) as this suggests that convergent seed microbiota of different subspecies across various cultivars thus a fidelity of subspecies signature better express in the seed bacterial microbiota than the rhizosphere and root endosphere (Kim *et al.*, 2020). As expected from previous work on numerous plant species (Bulgarelli *et al.*, 2012; Lundberg *et al.*, 2012; Edwards *et al.*, 2015; Beckers *et al.*, 2017), large differences in community diversity and composition among different plant microhabitats (Figure 1a-d and Figure S3a), which are likely a general rule for the assembly of plant-associated bacterial microbiota. The bacterial variability among individuals of a given cultivar is usually negligible compared to that observed among host subspecies/cultivars/genotypes (Edwards *et al.,* 2015; Zhang *et al.,* 2019; Kim *et al.,* 2020). Therefore, the bacterial variability within host genotype

might not affect our main finding that plant microhabitat, rather than subspecies type, is the major driver shaping plant-associated bacterial microbiota. Given that plant-associated bacterial microbiota remarkably be shaped by the specific microhabitats, the community assembly rules of each microhabitats were further investigated in depth.

Deterministic processes dominated the assembly of plant-associated bacterial microbiota

Based on metacommunity theory, a given plant microbiome assemblage (*e.g.*, seed/endosphere/rhizosphere microbiome) was assumed to emerge as a result of multiple-scale processes, including abiotic constraints, biotic interactions, and dispersal (Cordovez *et al.*, 2019). In fact, both deterministic and stochastic components represent two complementary parts of assembly processes that shape the structure of host-associated microbial communities (Figure 1e and Figure S3b). In addition, the observation in the present study provided quantitative evidence that deterministic processes were more important than stochastic processes with respect to bacterial community assembly, regardless of in which plant microhabitat (Figure 1e and Figure S3b). The fact that rhizosphere microbiota mainly influenced by deterministic processes has been demonstrated in previous studies (Mendes *et al.*, 2014; Fan *et al.*, 2017). As generally admitted, the rhizosphere landscape is profoundly shaped by the secretion of rhizodeposition produced by plant metabolism (Bulgarelli *et al.*, 2013). The diverse and dynamic rhizodeposition results in a highly fluctuating rhizosphere environment (Sasse *et al.*, 2018), which causes a stronger influence of environmental selection in the assembly of the rhizosphere bacterial community. In addition, the heterogeneous distribution of root exudation in space (Sasse *et al.*, 2018) could have resulted in a shift in selective pressure for the rhizosphere arriving bacterial microbiota, which also contributed to the greater importance of selection in community assembly processes. In contrast to the environmental selection for rhizosphere microbiota, root endosphere-dwelling bacterial microbiota must specialize and coevolve with hosts (Kemen *et al.*, 2015), and their colonization involves a range of the bacterial colonization traits (*e.g.*, chemotaxis towards specific resources, flagellar motility and production of enzymes) and interplay with host-plant immune responses (Hardoim *et al.*, 2008; Zamioudis & Pieterse, 2012). Thus, the assembly of root endosphere microbiota was determined by the selection of plant dedicated genetic systems, which enable

microbe-plant interrelation and a deterministic endophytic colonization process (Hardoim *et al.*, 2008). It is worth noting that the ratio of ecological stochasticity for seed microbiota was clearly lower than rhizosphere and root endosphere microbiota (Figure 1e and Figure S3b), indicating that deterministic processes are relatively more important in seed habitat. Similar results were obtained based on the neutral model, which show that the distribution of microbiota deviated clearly from the neutral model in seed habitat (Figure S3c). Indeed, the seed represents a relatively stable and niche-specific microhabitat for bacteria with less freely available carbon sources and other substrates (Truyens *et al.*, 2015). By consequence, the conditions within the seed habitat may be more severe and exhibit stronger selective pressures when compared with other plant microhabitats (Shade *et al.*, 2017). The microbiota inhabiting seeds usually have conserved endophyte properties, such as tolerant to a high osmotic pressure, endospore formation and amylase activity, which were not found in microbiota inhabiting other microhabitats (Truyens *et al.*, 2015).

Horizontal transmission governed the assemblage of root endosphere and rhizosphere microbiota

Rhizosphere as the main origin of root endosphere microbiota have been reasonably corroborated in recent studies (Edwards *et al.*, 2015; He *et al.*, 2020). Similarly, most of the observed bacteria that inhabiting root endosphere originated from the rhizosphere (Figure S4a), consistent with the idea that the colonization of the root endosphere microbiota follows a two or three-step selective model in which microbe must first colonize the rhizosphere/rhizoplane from surrounding soil and then invade the roots (Bulgarelli *et al.*, 2013; Reinhold-Hurek *et al.*, 2015). Furthermore, the co-variation between the bacterial community found in rhizosphere and endosphere was clearly demonstrated herein, and the more dissimilar between the rhizosphere were, the more dissimilarities in the root endosphere community took place (Figure 2 and Figure S4b). The high congruence of community variation between rhizosphere and endosphere clarify and support the viewpoint of the rhizosphere microbiota served as important inoculum reservoir, thus structuring bacterial community found in the root (Vandenkoornhuyse *et al.*, 2015; Frank *et al.*, 2017). In addition, the seed microbiota also contributed to the community composition of root

endosphere (Figure S4a), even though the community co-variation in between seed habitat and endosphere was relatively weak (Figure 2 and Figure S4b). Consistent with our own results, the vertical transmission of bacterial taxa from seed to progeny was also supported by several studies that demonstrate overlap in microbiota taxa between seed and root (Yang *et al.*, 2017; Zheng & Gong, 2019). Microbiota inhabiting in seed could be regarded as either transient members of the seed habitat (seed-borne microbiota) or persistent members of the plant-associated microbiota (seed-transmitted microbiota) (Shade *et al.*, 2017). Only a small proportion of seed-associated microbiota would successfully remain in the next generation and travel within the plant, even colonize the rhizosphere (Frank *et al.*, 2017) among which well documented examples of vertically inherited fungi (Schardl, 2001). This seed vertical microbiota inheritance allows a continuity of partnership, and microbial alleles can be reliably associated with hosts across generations. In addition, a smaller proportion (~17%) of microbiota inhabiting in root endosphere originated from an "unknown source", indicating that other potential environmental sources (*e.g.,* bulk soil, water or air) might contribute to root endosphere microbiota (Sanchez-Canizares *et al.*, 2017). The potential contribution of seed microbiota as an origin on the assembly of the rhizosphere microbiota was negligible (Figure 2 and Figure S4). The composition of rhizosphere microbial communities may more appears to follow the pattern of soil microbial communities, that is, horizontal inheritance of soil microbiota (Edwards *et al.*, 2015; Leff *et al.*, 2017).

Generalist core microbes inhabiting in rhizosphere correlated to plant phenotypic characteristics

The core plant microbiota generally refers to those microbial taxa that is closely linked with a certain plant population/species and established through long lasting evolutionary mechanisms of selection and enrichment (Lemanceau *et al.*, 2017; Toju *et al.*, 2018). Herein, a group of rice core microbes consisting of generalist species sharing in various microhabitats were identified (Figure 3 and Figure 4). Most members of the rice generalist core microbiota (*e.g.*, *Acidovorax*, *Bacillus*, *Pseudomonas*, *Rhizobium*, *Sphingomonas*) had overlap with those identified in previous studies of other plant species such as *Arabidopsis thaliana* (Lundberg *et al.*, 2012; Thiergart *et al.*, 2020), sugarcane (Hamonts *et al.*, 2018), citrus (Xu *et al.*, 2018) and grapevine (Zarraonaindia *et al.*,

2015), suggesting that the presence of some core microbial taxa may be common across plant species (Trivedi *et al.*, 2020). Multiple members affiliated with these core species have been verified to exert different types of positive functions on plant health and growth (Tabassum *et al.,* 2017; Afzal *et al.*, 2019; Mhatre *et al.,* 2019; Singh *et al.,* 2019). The existence of common core microbiota members in various host-plants implied that a high conserved, co-evolutionary, host-independent core plant microbiota may exists that maintain plant holobiont fitness (Vandenkoornhuyse *et al.*, 2015; Muller *et al.*, 2016; Thiergart *et al.*, 2020). Moreover, the evolutionary selection mechanism might have favored the cooperative relationship between microbial taxa within a particular microhabitat (Hassani *et al.*, 2018), in which positive co-occurrence patterns were mostly observed (Figure 5a). The generalist core microbes trended to occupy central positions with high connectedness/connectivity within the co-occurrence network, in contrast to other generalist members of the microbiota, especially in root endosphere and rhizosphere (Figure 5a and Figure S7). The high interaction of generalist core microbes could raise possibilities for positively facilitating community diversity and stability, thereby promoting host-plant health and fitness (Agler *et al.*, 2016; Toju *et al.*, 2018). Interestingly, compared with inhabiting in seed habitat or root endosphere, generalist core microbes inhabiting in rhizosphere are more closely related to plant phenotypic characteristics and had stronger predictive ability for these traits (Figure 5b). These morphological and physiological traits are usually related to the strategies of plant resource acquisition, suggesting the ecological importance of microbial interactions and processes in the rhizosphere for plant resource consumption and turnover (Philippot *et al.*, 2013). By contrast, the presence of generalist core microbes inhabiting in seed habitat or root endosphere is likely to be the result of a long lasting co-evolutionary processes between host-plant and its associated microbiota, which highly adapted to host life without necessarily being direct beneficial (Hardoim *et al.*, 2015). Further analyses using gnotobiotic approaches and synthetic communities are expected to decipher the impact of these supposed important microorganisms in host plant growth survival and reproduction. Beside this important prospect, genomic analyses of these particular microorganisms could provide important predictive functional information.

Conclusions

Deterministic processes dominate the assembly of bacterial communities, whether in seed habitat, or in root endosphere and rhizosphere of well-established vegetatively growing plant. Among them, the horizontal transmission contributes far more than vertical transmission to root endosphere and rhizosphere community assemblage. Moreover, several important functional traits of the host-plant could be explained to an extent by the relative abundance of generalist core microbes inhabiting in rhizosphere. In conclusion, plant fitness can be manipulated potentially by altering the external bacterial microbiota sources, which could further participate in plant-microbe interactions in a deterministic manner. This idea opens new prospects on other important compartments of the plant microbiota (*i.e.* fungi, archaea, protists) and their interactions with host-plant. Our study goes beyond simply characterizing the assemblages of plant-associated bacterial communities to decipher the deterministic-based assembly mechanisms driving plant-associated bacterial communities from the perspective of community ecology, highlight the powerful role of horizontal transmission route for community assembly, as well as identify the presence of core microbes and their ability to predict host-plant phenotypes. This study may inform future efforts to transform the fundamental understanding of plant-microbiota associations into practical management strategies, which aim at engineering the desired plant-associated microbiome in diverse agricultural systems to enhance plant performance.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (41977080, 31902114, U2003210), the Natural Science Foundation of Jiangsu Province (BK20190543), the China Postdoctoral Science Foundation (2019M651861), the Young Elite Scientists Sponsorship Program by CAST (2019QNRC001), the Innovative Research Team Development Plan of the Ministry of Education of China (Grant No. IRT_17R56), and the Fundamental Research Funds for the Central Universities (Grant No. KYT201802). This work was also supported by a grant the French National program EC2CO (Ecosphère Continentale et Côtière) of the CNRS (HOLOBIONT – N°12669). We specially thank Dr. Matthieu Barret from Institut National de la Recherche Agronomique, France, for the valuable comments on the paper. We also thank American Journal Experts for polishing the language. Finally, we are grateful for the constructive comments of three anonymous reviewers. The authors declare that they have no conflict of interest.

Data availability

All the raw sequence data for 16S rRNA gene amplicons were submitted to the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/Traces/sra/) database under the accession number PRJNA660605.

Authors' contributions

N.L., Y.L. and S.G. conceived and designed the study. J.G., K.L., H.N. carried out the experiments. J.G., N.L. and P.V. performed the analyses and drafted the manuscript. S.G. and Q.S. made plentiful valuable comments for the manuscript. All the authors reviewed and approved the manuscript.

References

Afzal I, Shinwari ZK, Sikandar S, Shahzad S. 2019. Plant beneficial endophytic bacteria: mechanisms, diversity, host range and genetic determinants. *Microbiological Research* **221**: 36-49.

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim ST, Weigel D, Kemen EM. 2016. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLoS Biology* **14**: e1002352.

Archer E. 2020. rfPermute: Estimate permutation p-values for random forest importance metrics. R package version 2.1.81 [WWW document] URL https://cran.r-project.org/web/packages/rfPermute/fPermute.pdf [accessed 23 February 2020].

Bastian M, Heymann S, Jacomy M. 2009. Gephi: An open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media* **3**: 361-362.

Beckers B, Op De Beeck M, Thijs S, Truyens S, Weyens N, Boerjan W, Vangronsveld J. 2016. Performance of 16s rDNA primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies. *Frontiers in Microbiology* **7**: 650.

Beckers B, Op De Beeck M, Weyens N, Boerjan W, Vangronsveld J. 2017. Structural variability and niche differentiation in the rhizosphere and endosphere bacterial microbiome of field-grown poplar trees. *Microbiome* **5**: 25.

Berg G, Rybakova D, Grube M, Koberl M. 2016. The plant microbiome explored: implications for experimental botany. *Journal of Experimental Botany* **67**: 995-1002.

Bordenstein SR, Theis KR. 2015. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biology* **13**: e1002226.

Bringel F, Couee I. 2015. Pivotal roles of phyllosphere microorganisms at the interface between plant functioning and atmospheric trace gas dynamics. *Frontiers in Microbiology* **6**: 486.

Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, Rauf P, Huettel B, Reinhardt R, Schmelzer E, et al. 2012. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. *Nature* **488**: 91-95.

Bulgarelli D, Schlaeppi K, Spaepen S, Ver Loren van Themaat E, Schulze-Lefert P. 2013. Structure and functions of the bacterial microbiota of plants. *Annual Review of Plant Biology* **64**: 807-838.

Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, Bohannan BJ. 2016. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *The ISME Journal* **10**: 655-664.

Busby PE, Soman C, Wagner MR, Friesen ML, Kremer J, Bennett A, Morsy M, Eisen JA, Leach JE, Dangl JL.

2017. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biology* **15**: e2001793.

Coleman-Derr D, Desgarennes D, Fonseca-Garcia C, Gross S, Clingenpeel S, Woyke T, North G, Visel A, Partida-Martinez LP, Tringe SG. 2016. Plant compartment and biogeography affect microbiome composition in cultivated and native *Agave* species. *New Phytologist* **209**: 798-811.

Compant S, Samad A, Faist H, Sessitsch A. 2019. A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *Journal of Advanced Research* **19**: 29-37.

Cordovez V, Dini-Andreote F, Carrion VJ, Raaijmakers JM. 2019. Ecology and evolution of plant microbiomes. *Annual Review of Microbiology* **73**: 69-88.

Cosme M, Lu J, Erb M, Stout MJ, Franken P, Wurst S. 2016. A fungal endophyte helps plants to tolerate root herbivory through changes in gibberellin and jasmonate signaling. *New Phytologist* **211**: 1065-1076.

Cregger MA, Veach AM, Yang ZK, Crouch MJ, Vilgalys R, Tuskan GA, Schadt CW. 2018. The *Populus* holobiont: dissecting the effects of plant niches and genotype on the microbiome. *Microbiome* **6**: 31.

Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *Interjournal Complex Systems* **1695**: 1-9.

Dini-Andreote F, Raaijmakers JM. 2018. Embracing community ecology in plant microbiome research. *Trends in Plant Science* **23**: 467-469.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460-2461.

Edwards J, Johnson C, Santos-Medellin C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V. 2015. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences, USA* **112**: E911- E920.

Fan KK, Cardona C, Li YT, Shi Y, Xiang XJ, Shen CC, Wang HF, Gilbert JA, Chu HY. 2017. Rhizosphere-associated bacterial network structure and spatial distribution differ significantly from bulk soil in wheat crop fields. *Soil Biology and Biochemistry* **113**: 275-284.

Feng K, Zhang Z, Cai W, Liu W, Xu M, Yin H, Wang A, He Z, Deng Y. 2017. Biodiversity and species competition regulate the resilience of microbial biofilm community. *Molecular Ecology* **26**: 6170-6182.

Fitzpatrick CR, Copeland J, Wang PW, Guttman DS, Kotanen PM, Johnson MTJ. 2018. Assembly and ecological function of the root microbiome across angiosperm plant species. *Proceedings of the National*

Academy of Sciences, USA **115**: E1157-E1165.

Fitzpatrick CR, Salas-Gonzalez I, Conway JM, Finkel OM, Gilbert S, Russ D, Teixeira P, Dangl JL. 2020. The plant microbiome: from ecology to reductionism and beyond. *Annual Review of Microbiology* **74**: 81-100.

Frank AC, Saldierna Guzman JP, Shay JE. 2017. Transmission of bacterial endophytes. *Microorganisms* **5**: 70.

Friedman J, Alm EJ. 2012. Inferring correlation networks from genomic survey data. *PLoS Computational Biology* **8**: e1002687.

Friesen ML, Porter SS, Stark SC, von Wettberg EJ, Sachs JL, Martinez-Romero E. 2011. Microbially mediated plant functional traits. *Annual Review of Ecology, Evolution, and Systematics* **42**: 23-46.

Guo X, Feng JJ, Shi Z, Zhou XS, Yuan MT, Tao XY, Hale L, Yuan T, Wang JJ, Qin YJ, et al. 2018. Climate warming leads to divergent succession of grassland microbial communities. *Nature Climate Change* **8**: 813-818.

Hamonts K, Trivedi P, Garg A, Janitz C, Grinyer J, Holford P, Botha FC, Anderson IC, Singh BK. 2018. Field study reveals core plant microbiota and relative importance of their drivers. *Environmental Microbiology* **20**: 124-140.

Hardoim PR, van Overbeek LS, Berg G, Pirttila AM, Compant S, Campisano A, Doring M, Sessitsch A. 2015. The hidden world within plants: ecological and evolutionary considerations for defining functioning of microbial endophytes. *Microbiology and Molecular Biology Reviews* **79**: 293-320.

Hardoim PR, van Overbeek LS, Elsas JD. 2008. Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology* **16**: 463-471.

Hassani MA, Duran P, Hacquard S. 2018. Microbial interactions within the plant holobiont. *Microbiome* **6**: 58.

He R, Zeng J, Zhao D, Huang R, Yu Z, Wu QL. 2020. Contrasting patterns in diversity and community assembly of *Phragmites australis* root-associated bacterial communities from different seasons. *Applied and Environmental Microbiology* **86**: e00379-20.

Kemen AC, Agler MT, Kemen E. 2015. Host-microbe and microbe-microbe interactions in the evolution of obligate plant parasitism. *New Phytologist* **206**: 1207-1228.

Kim H, Lee KK, Jeon J, Harris WA, Lee YH. 2020. Domestication of *Oryza* species eco-evolutionarily shapes bacterial and fungal communities in rice seed. *Microbiome* **8**: 20.

Kokou F, Sasson G, Friedman J, Eyal S, Ovadia O, Harpaz S, Cnaani A, Mizrahi I. 2019. Core gut microbial

communities are maintained by beneficial interactions and strain variability in fish. *Nature Microbiology* **4**: 2456-2465.

Kong HG, Song GC, Ryu CM. 2019. Inheritance of seed and rhizosphere microbial communities through plant-soil feedback and soil memory. *Environmental Microbiology Reports* **11**: 479-486.

Kroll S, Agler MT, Kemen E. 2017. Genomic dissection of host-microbe and microbe-microbe interactions for advanced plant breeding. *Current Opinion in Plant Biology* **36**: 71-78.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**: 1547-1549.

Leff JW, Lynch RC, Kane NC, Fierer N. 2017. Plant domestication and the assembly of bacterial and fungal communities associated with strains of the common sunflower, *Helianthus annuus*. *New Phytologist* **214**: 412-423.

Lemanceau P, Blouin M, Muller D, Moenne-Loccoz Y. 2017. Let the core microbiota be functional. *Trends in Plant Science* **22**: 583-595.

Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research* **44**: W242-W245.

Louca S, Parfrey LW, Doebeli M. 2016. Decoupling function and taxonomy in the global ocean microbiome. *Science* **353**: 1272-1277.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**(12): 550.

Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J, Engelbrektson A, Kunin V, Del Rio TG, et al. 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**: 86-90.

Mendes LW, Kuramae EE, Navarrete AA, van Veen JA, Tsai SM. 2014. Taxonomical and functional microbial community selection in soybean rhizosphere. *The ISME Journal* **8**: 1577-1587.

Mhatre PH, Karthik C, Kadirvelu K, Divya KL, Venkatasalam EP, Srinivasan S, Ramkumar G, Saranya C, Shanmuganathan R. 2019. Plant growth promoting rhizobacteria (PGPR): A potential alternative tool for nematodes bio-control. *Biocatalysis and Agricultural Biotechnology* **17**: 119-128.

Muller DB, Vogel C, Bai Y, Vorholt JA. 2016. The plant microbiota: systems-level insights and perspectives.

Annual Review of Genetics **50**: 211-234.

Murphy MA, Evans JS, Storfer A. 2010. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. *Ecology* **91**: 252-261.

Ning D, Deng Y, Tiedje JM, Zhou J. 2019. A general framework for quantitatively assessing ecological stochasticity. *Proceedings of the National Academy of Sciences, USA* **116**: 16892-16898.

Niu B, Paulson JN, Zheng X, Kolter R. 2017. Simplified and representative bacterial community of maize roots. *Proceedings of the National Academy of Sciences, USA* **114**: E2450-E2459.

Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL, Buckler ES, Ley RE. 2013. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences, USA* **110**: 6548-6553.

Perez-Jaramillo JE, Carrion VJ, Bosse M, Ferrao LFV, de Hollander M, Garcia AAF, Ramirez CA, Mendes R, Raaijmakers JM. 2017. Linking rhizosphere microbiome composition of wild and domesticated *Phaseolus vulgaris* to genotypic and root phenotypic traits. *The ISME Journal* **11**: 2244-2257.

Philippot L, Raaijmakers JM, Lemanceau P, van der Putten WH. 2013. Going back to the roots: the microbial ecology of the rhizosphere. *Nature Reviews Microbiology* **11**: 789-799.

Pineda A, Zheng SJ, van Loon JJ, Pieterse CM, Dicke M. 2010. Helping plants to deal with insects: the role of beneficial soil-borne microbes. *Trends in Plant Science* **15**: 507-514.

Ravanbakhsh M, Kowalchuk GA, Jousset A. 2020. Targeted plant hologenome editing for plant trait enhancement. *New Phytologist* **229**: 1067-1077.

Reinhold-Hurek B, Bunger W, Burbano CS, Sabale M, Hurek T. 2015. Roots shaping their microbiome: global hotspots for microbial activity. *Annual Review of Phytopathology* **53**: 403-424.

Sanchez-Canizares C, Jorrin B, Poole PS, Tkacz A. 2017. Understanding the holobiont: the interdependence of plants and their microbiome. *Current Opinion in Microbiology* **38**: 188-196.

Sasse J, Martinoia E, Northen T. 2018. Feed your friends: do plant exudates shape the root microbiome? *Trends in Plant Science* **23**: 25-41.

Schardl CL. 2001. *Epichloe festucae* and related mutualistic symbionts of grasses. *Fungal Genetics and Biology* **33**: 69-82.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks

DH, Robinson CJ, et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**: 7537-7541.

Shade A, Jacques MA, Barret M. 2017. Ecological patterns of seed microbiome diversity, transmission, and assembly. *Current Opinion in Microbiology* **37**: 15-22.

Shenhav L, Thompson M, Joseph TA, Briscoe L, Furman O, Bogumil D, Mizrahi I, Pe'er I, Halperin E. 2019. FEAST: fast expectation-maximization for microbial source tracking. *Nature Methods* **16**: 627-632.

Singh P, Singh RK, Singh MP, Song QQ, Solanki MK, Yang LT, Li YR. 2019. Soil: microbial cell factory for assortment with beneficial role in agriculture. In: Singh DP, Gupta VK, Prabha R, eds. *Microbial Interventions in Agriculture and Environment*. Singapore: Springer, 63-92.

Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology* **8**: 732-740.

Sriswasdi S, Yang CC, Iwasaki W. 2017. Generalist species drive microbial dispersion and evolution. *Nature Communications* **8**: 1162.

Tabassum B, Khan A, Tariq M, Ramzan M, Khan MSI, Shahid N, Aaliya K. 2017. Bottlenecks in commercialisation and future prospects of PGPR. *Applied Soil Ecology* **121**: 102-117.

Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, Bosch TC, Cryan JF, Gilbert SF, Goodnight CJ, Lloyd EA, et al. 2016. Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems* **1**: e00028-16.

Thiergart T, Duran P, Ellis T, Vannier N, Garrido-Oter R, Kemen E, Roux F, Alonso-Blanco C, Agren J, Schulze-Lefert P, et al. 2020. Root microbiota assembly and adaptive differentiation among European *Arabidopsis* populations. *Nature Ecology and Evolution* **4**: 122-131.

Toju H, Peay KG, Yamamichi M, Narisawa K, Hiruma K, Naito K, Fukuda S, Ushio M, Nakaoka S, Onoda Y, et al. 2018. Core microbiomes for sustainable agroecosystems. *Nature Plants* **4**: 247-257.

Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. 2020. Plant-microbiome interactions: from community assembly to plant health. *Nature Reviews Microbiology* **18**: 607-621.

Truyens S, Weyens N, Cuypers A, Vangronsveld J. 2015. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environmental Microbiology Reports* **7**: 40-50.

Turner TR, James EK, Poole PS. 2013. The plant microbiome. *Genome Biology* **14**: 209.

Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A. 2015. The importance of the microbiome of the plant holobiont. *New Phytologist* **206**: 1196-1206.

Vannier N, Mony C, Bittebiere AK, Michon-Coudouel S, Biget M, Vandenkoornhuyse P. 2018. A microorganisms' journey between plant generations. *Microbiome* **6**: 79.

Vorholt JA. 2012. Microbial life in the phyllosphere. *Nature Reviews Microbiology* **10**: 828-840.

Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**: 5261-5267.

Xu J, Zhang Y, Zhang P, Trivedi P, Riera N, Wang Y, Liu X, Fan G, Tang J, Coletta-Filho HD, et al. 2018. The structure and function of the global citrus rhizosphere microbiome. *Nature Communications* **9**: 4894.

Xu Q, Ling N, Chen H, Duan Y, Wang S, Shen Q, Vandenkoornhuyse P. 2020. Long-term chemical-only fertilization induces a diversity decline and deep selection on the soil bacteria. *mSystems* **5**: e00337-20.

Yang L, Danzberger J, Scholer A, Schroder P, Schloter M, Radl V. 2017. Dominant groups of potentially active bacteria shared by barley seeds become less abundant in root associated microbiome. *Frontiers in Plant Science* **8**: 1005.

Zamioudis C, Pieterse CM. 2012. Modulation of host immunity by beneficial microbes. *Molecular Plant-Microbe Interactions* **25**: 139-150.

Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, Bokulich NA, Mills DA, Martin G, Taghavi S, et al. 2015. The soil microbiome influences grapevine-associated microbiota. *mBio* **6**: e02527-14.

Zhang J, Liu YX, Zhang N, Hu B, Jin T, Xu H, Qin Y, Yan P, Zhang X, Guo X, et al. 2019. *NRT1.1B* is associated with root microbiota composition and nitrogen use in field-grown rice. *Nature Biotechnology* **37**: 676-684.

Zheng Y, Gong X. 2019. Niche differentiation rather than biogeography shapes the diversity and composition of microbiome of *Cycas panzhihuaensis*. *Microbiome* **7**: 152.

Zhou J, Ning D. 2017. Stochastic community assembly: does it matter in microbial ecology? *Microbiology and Molecular Biology Reviews* **81**: e00002-17.

Supporting Information

Figure S1 Diagram of original collection countries of 99 rice varieties.

Figure S2 The α-diversity metrics and β-diversity patterns of bacterial communities in each plant microhabitat collected from different rice subspecies types.

Figure S3 The β-diversity, stochasticity and neutral model fitting of bacterial communities inhabiting in different plant microhabitats.

Figure S4 The relationship between bacterial community inhabiting in different plant microhabitats.

Figure S5 Classification of microbial generalists and specialists inhabiting in each plant microhabitat.

Figure S6 Cumulative percentage and relative abundance of generalist OTUs and associations between generalist microbiota and the total microbiota under different microhabitats.

Figure S7 Network hubs of seed habitat, root endosphere and rhizosphere.

Table S1 Plant functional traits of different rice cultivated varieties.

Figure legends:

Figure 1 The diversity, composition and stochasticity of bacterial communities inhabiting in seed habitat (Seed), root endosphere (Endo) and rhizosphere (Rhizo). (a) OTU richness and (b) Shannon indices of bacterial communities in seed habitat, root endosphere and rhizosphere. Horizontal lines within boxes denote medians. Tops and bottoms of boxes denote the 75th and 25th percentiles, respectively. Upper and lower whiskers extend to data no more than 1.5× the interquartile range from the upper edge and lower edge of the box, respectively. Different letters indicate significant differences among plant microhabitats ($P < 0.05$), based on Kruskal-Wallis one-way test. (c) PCoA plot depicting the β-diversity patterns of bacterial communities across different plant microhabitats based on Bray-Curtis distances. Significance tests of bacterial community dissimilarities among different plant microhabitats are based on PERMANOVA and ANOSIM tests. (d) Taxonomic composition of the bacterial communities under different plant microhabitats at the (sub) phylum level. (e) The modified stochasticity ratio (MST) of bacterial communities under different plant microhabitats was developed based on Bray-Curtis distances with 50% as the boundary point between more deterministic (<50%) and more stochastic (>50%) community assembly.

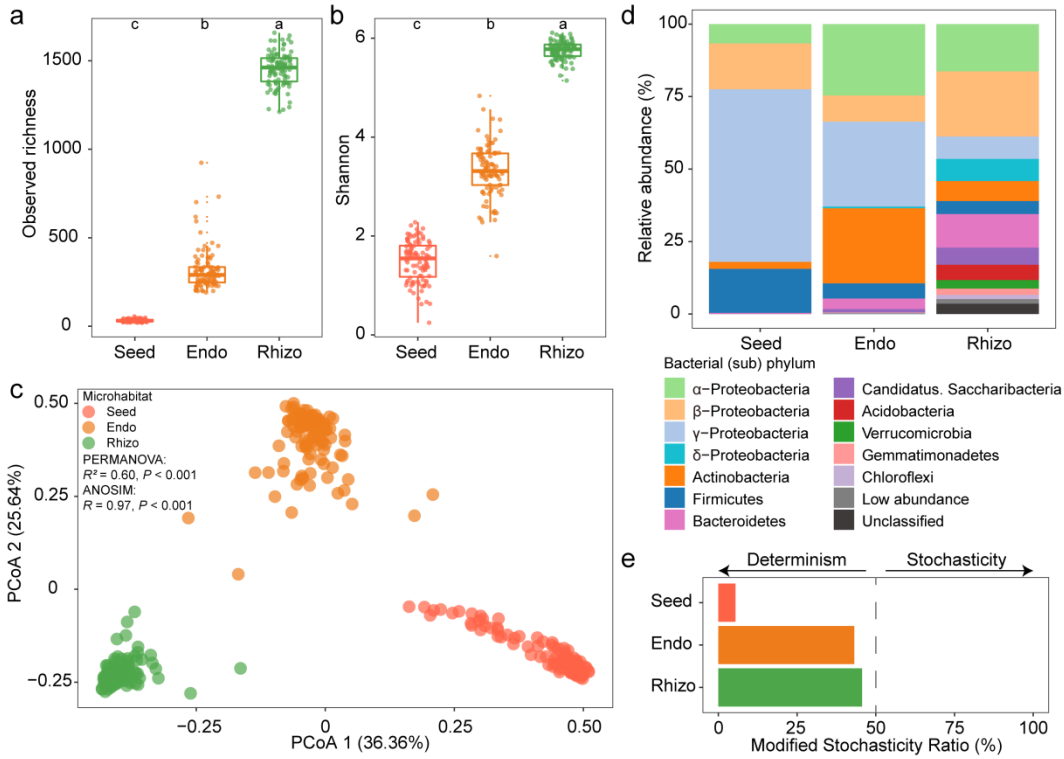
Figure 2 The association of microbiota inhabiting in different plant microhabitats. (a) Procrustes correlation between community compositions (Bray-Curtis metrics) of bacterial microbiota inhabiting in different plant microhabitats. M^2 represents the sum of squared distances between matched sample pairs. r indicates the correlation in a symmetric Procrustes rotation. P values were determined from 999 labelled permutations. Arrows indicate in which direction the ordination is stretched to fit the ordination of the bacterial compositions in one microhabitat to the ordination of bacterial compositions in another microhabitat. (b) Enrichment (positive) and depletion (negative) of OTUs between different plant microhabitats Each point represents an individual OTU. (c) Top 30 genera of enriched or depleted OTUs between different plant microhabitats. Colors of points indicate (sub) phylum classification.

Figure 3 Environmental distribution, phylogenetic classification, and potential ecological functions of generalist microbiota in various plant microhabitats. (a) The environmental distribution of the generalists and specialists. Each individual variety of rice was considered a unique environment. Relative abundance is shown on a logarithmic (log10) scale on the x axis. (b) Phylogenetic tree and taxonomic composition heat map of generalist microbiota. The scale bar represents 0.1 nucleic acid sequence divergence. (c) Potential ecological functions of generalist microbiota inhabiting in different plant microhabitats were based on FAPROTAX. Each bubble represents presence of potential ecological function of generalist microbiota in plant microhabitats.

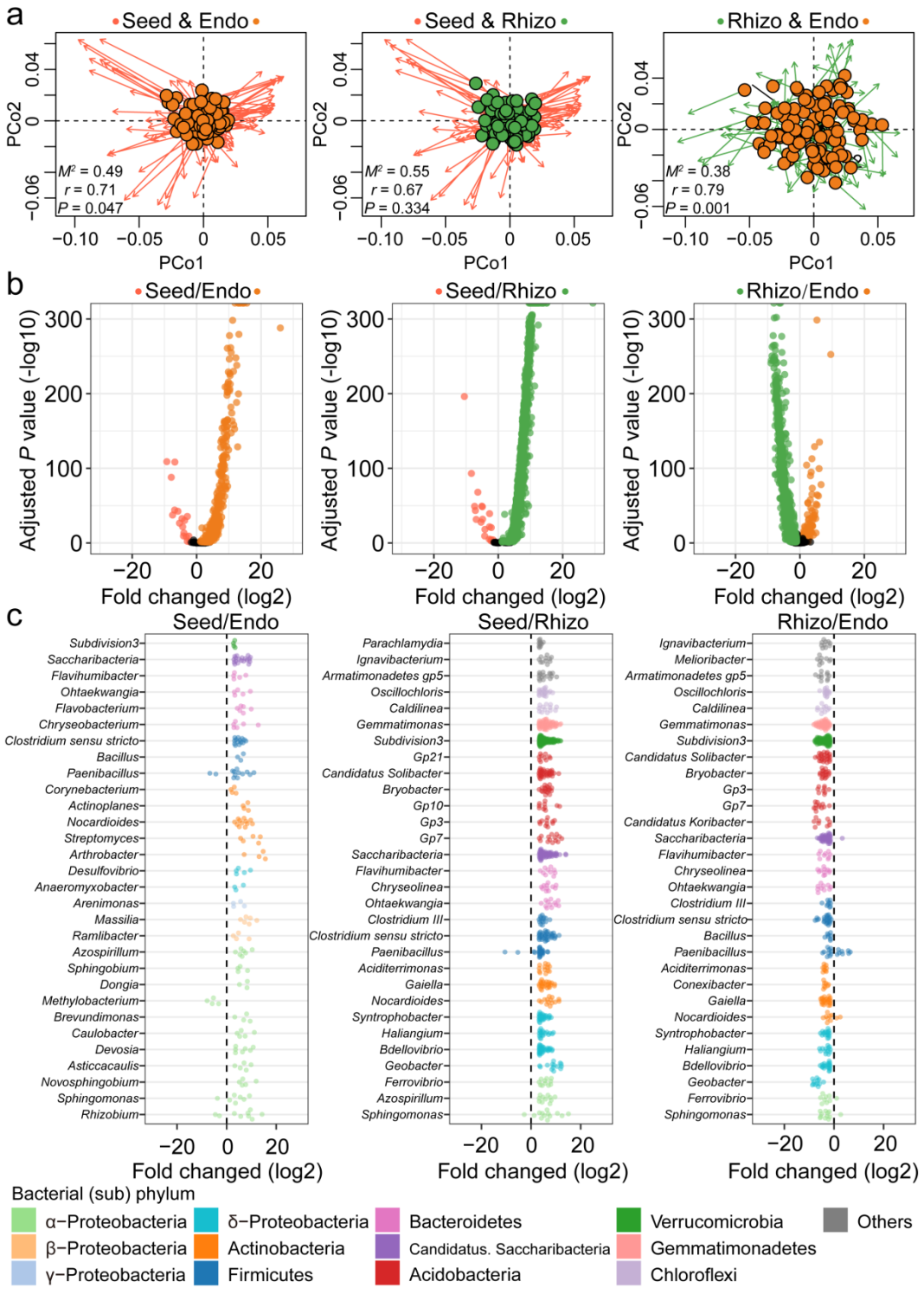
Figure 4 Characterization of the generalist core taxa of the plant-associated microbiota. (a) Relative abundance of generalist core taxa in each plant microhabitat. Error bars represent standard errors (n = 99). Venn diagram depicts the number of shared generalist microbes between each plant microhabitat. (b) Neighbour-joining tree for the generalist core taxa (shared generalist microbes that were present in all plant microhabitats). Bacterial OTU representative sequences from the present study are shown in bold. The scale bar represents 0.01 nucleic acid sequence divergence.

Figure 5 Microbial networks of generalist microbiota inhabiting in different microhabitats and associations between generalist core microbes and plant functional traits. (a) Co-abundance networks of generalist taxa in different microhabitats. Nodes represent the OTUs, edges indicate the significant correlation between the nodes. Horizontal lines within boxes denote medians. Tops and bottoms of boxes denote the 75th and 25th percentiles, respectively. Upper and lower whiskers extend to data no more than 1.5× the interquartile range from the upper edge and lower edge of the box, respectively. The difference of node-level betweenness centrality between generalist core and non-core microbes are based on the Wilcoxon rank sum test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant). (b) Associations between the plant functional traits and the relative abundance of generalist core microbes inhabiting in each plant microhabitat evaluated by correlation analysis and best random forest model. Only significant correlations ($P < 0.05$) are

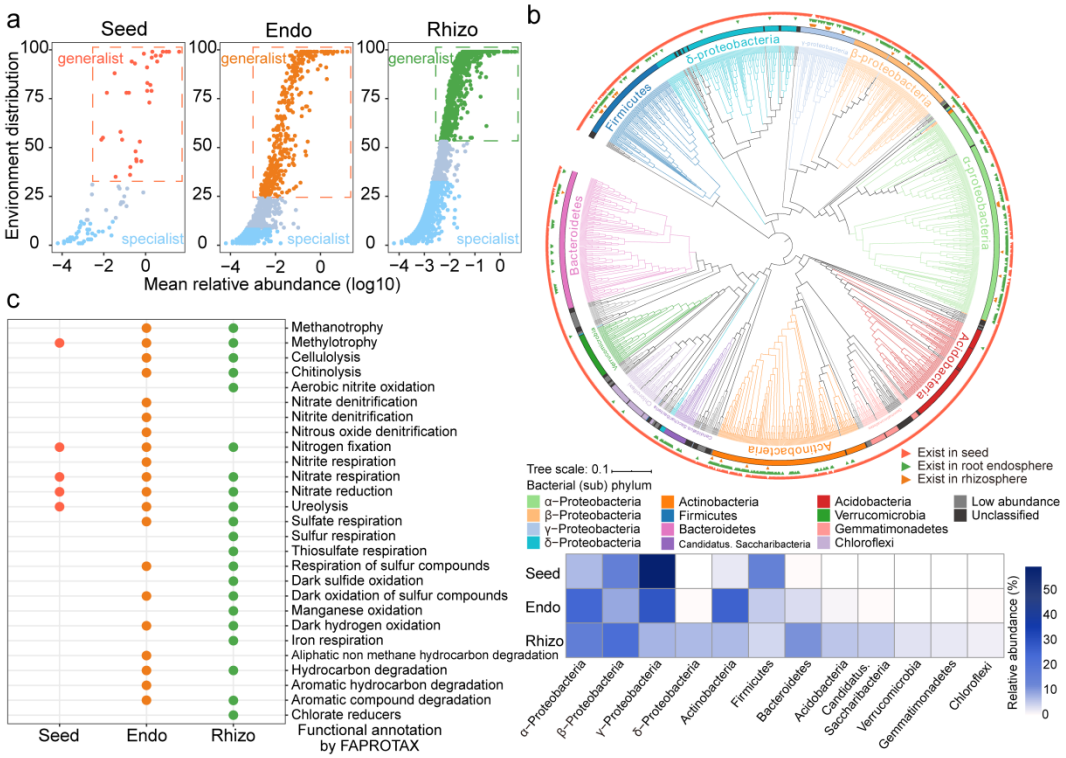
shown. Stronger colours (blue is positive, red is negative) in the heat maps represent stronger Spearman correlations. Circle size represents the generalist core microbe's importance (% of increased mean square error) predicted by random forest model. Significance levels of random forest model are as follows: *, $P < 0.05$; **, $P < 0.01$.



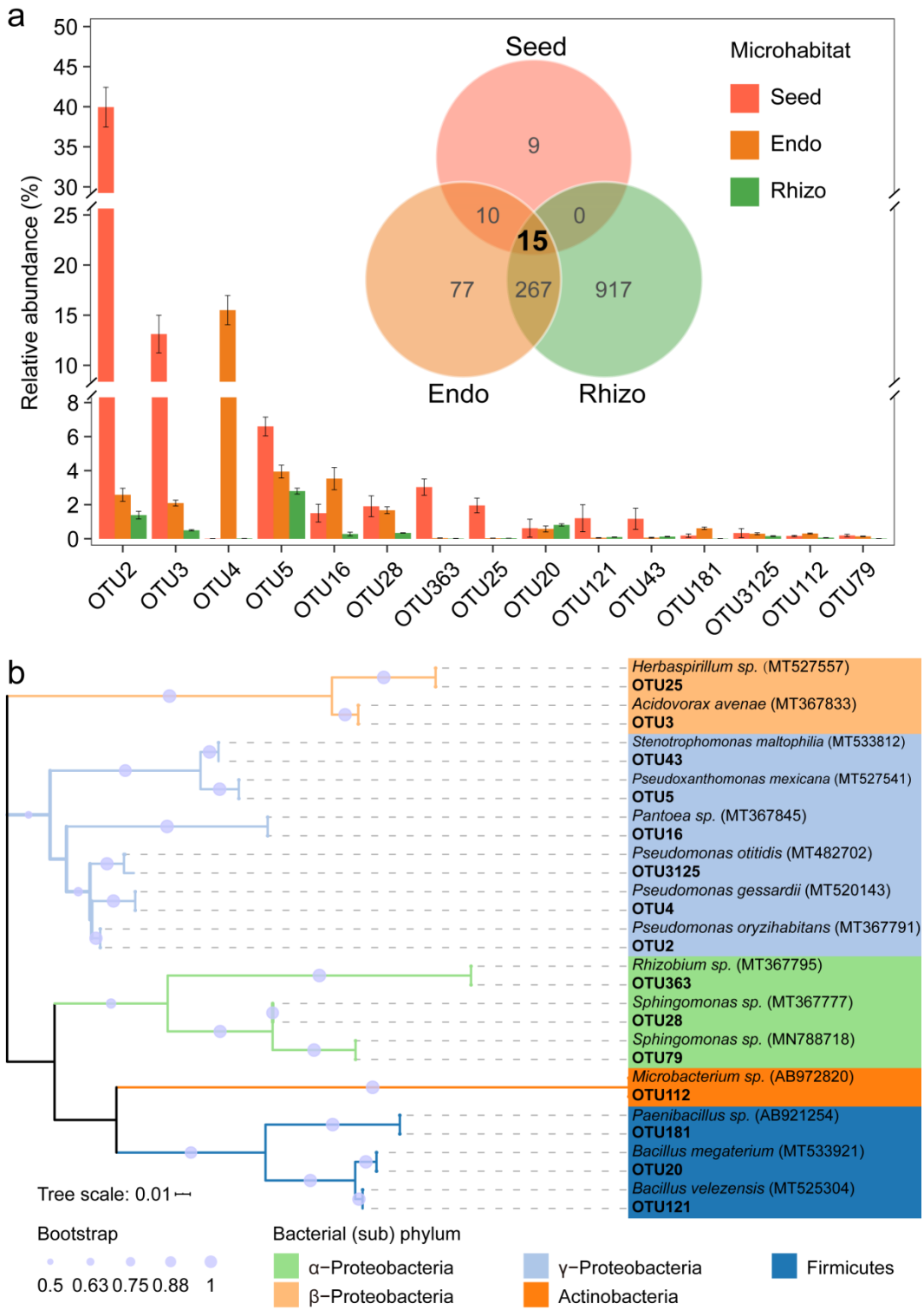
nph_17297_f1.tif



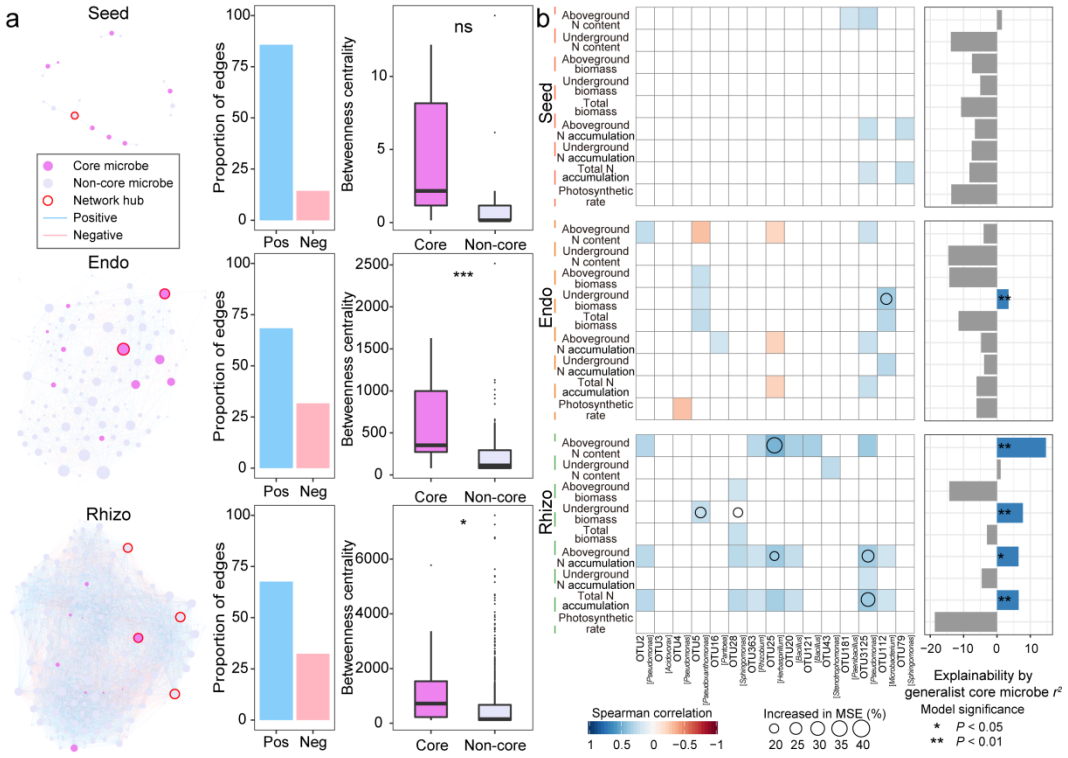
nph_17297_f2.tif



nph_17297_f3.tif



nph_17297_f4.tif



nph_17297_f5.tif