Environment International 142 (2020) 105869

Contents lists available at ScienceDirect

Environment International

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

Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems

Yu Shi[a], Manuel Delgado-Baquerizo[b], Yuntao Li[a], Yunfeng Yang[c], Yong-Guan Zhu[d,e,*], Josep Peñuelas[f,g], Haiyan Chu[a,i,*]

[a] State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, 71 East Beijing Road, Nanjing 210008, China
[b] Departamento de Sistemas Físicos, Químicos y Naturales, Universidad Pablo de Olavide, 41013 Sevilla, Spain
[c] State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China
[d] Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China
[e] State Key Lab of Urban and Regional Ecology, Research Centre for Eco-environmental Sciences, Chinese Academy of Sciences, Beijing 10085, China
[f] CSIC, Global Ecology Unit CREAF-CSIC-UAB, Bellaterra, Catalonia E-08193, Spain
[g] CREAF, Cerdanyola del Vallès, Catalonia E-08193, Spain
[h] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Handling Editor: Guo-Ping Sheng

Keywords:
Microbial network
Kinless hub
Functional gene
Fungi
Bacteria

ABSTRACT

Microbial taxa within complex ecological networks can be classified by their universal roles based on their level of connectivity with other taxa. Highly connected taxa within an ecological network (kinless hubs) are theoretically expected to support higher levels of ecosystem functions than less connected taxa (peripherals). Empirical evidence of the role of kinless hubs in regulating the functional potential of soil microbial communities, however, is largely unexplored and poorly understood in agricultural ecosystems. Here, we built a correlation network of fungal and bacterial taxa using a large-scale survey consisting of 243 soil samples across functionally and economically important agricultural ecosystems (wheat and maize); and found that the relative abundance of taxa classified as kinless hubs within the ecological network are positively and significantly correlated with the abundance of functional genes including genes for C fixation, C degradation, C methanol, N cycling, P cycling and S cycling. Structural equation modeling of multiple soil properties further indicated that kinless hubs, but not provincial, connector or peripheral taxa, had direct significant and positive relationships with the abundance of multiple functional genes. Our findings provide novel evidence that the relative abundance of soil taxa classified as kinless hubs within microbial networks are associated with high functional potential, with implications for understanding and managing (through manipulating microbial key species) agricultural ecosystems at a large spatial scale.

1. Introduction

Soil microbial taxa live together within ecological networks, forming ecological clusters of taxa strongly co-occurring with each other (Toju et al., 2018a). Taxa within these ecological clusters are important ecological units that share environmental preferences (e.g. Faust et al., 2015) but may differ with other taxa in their level of connectivity within and across ecological clusters. To put it simply, some taxa are highly connected within an ecological cluster (provincial hubs) or within and between ecological clusters (kinless hubs). Other taxa, however, support a much lower number of connections within the network (e.g. peripheral and connector nodes (Guimera and Amaral, 2005; Poudel et al., 2016). Kinless hubs, unlike other nodes, strongly support the structure of ecological networks (Faust et al., 2015; Shade and Handelsman, 2012; Toju et al., 2018b), and are expected to create niche for other taxa (Toju et al., 2018a). The importance of ecological networks for investigating the role of microbial communities have been recently highlighted in natural ecosystems (Wagg et al., 2019; Delgado-Baquerizo et al., 2020), however, very little is known about whether ecological networks can also provide useful information to maintain ecosystem functions in agricultural environments. This knowledge is essential to increase plant production to support a continuously growing human population worldwide.

The importance of the relative abundance of kinless hubs for

* Corresponding authors at: Key Laboratory of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen 361021, China (Y.-G. Zhu).
E-mail address: hychu@issas.ac.cn (H. Chu).

https://doi.org/10.1016/j.envint.2020.105869
Received 12 February 2020; Received in revised form 22 April 2020; Accepted 16 May 2020

ecosystem functioning has been reported for plant communities (Rey et al., 2016; Soliveres et al., 2012; Cavieres et al., 2014). The relative abundance of microbial kinless hubs may similarly comprise a higher abundance of important functional attributes of nutrient cycling and carbon (C) and nitrogen (N) fixation. These hubs have been previously defined as keystone species for plant communities, where they play important roles in the fluxes of energy and materials and provide some amount of production in terrestrial ecosystems (Toju et al., 2018b; Ellison et al., 2005; Rooney et al., 2006). Probably, these species also represent organisms with limited inherent function which rely upon functions performed by ecological clusters with which they associate (Kokou et al., 2019). Empirical evidence supporting the importance of the relative abundance of kinless hubs within microbial networks, however, is lacking. Identifying the environmental preferences and functional importance of microbial taxa in kinless hubs is essential to better understand the soil processes supported by microbial communities in terrestrial ecosystems. Similar to what has been reported for plants (Rey et al., 2016; Cavieres et al., 2014), we hypothesize that highly connected microbial taxa within and across ecological clusters (kinless hubs) might also be positively associated with abundances of functional genes involved multiple ecosystem processes including cycling of C, N, phosphorus (P) and sulfur (S). All these functions are fundamental for maintaining plant production. Because of this, we hypothesized that the abundance of kinless hubs might be of paramount importance to maintain productivity in human-managed ecosystems.

To test this hypothesis, we analyzed 243 soil samples on a large scale (> 1000 km) across the North China Plain, which is probably the most intensive agricultural ecosystem worldwide (Han et al., 2018). This region has been intensively managed over the last century under functionally and economically important wheat-maize double cropping rotations, accounting for > 50% of the total cereal production in China (Jeong et al., 2014). Most recent studies suggest that agricultural productivity and sustainability could be largely improved by managing soil microbes (Toju et al., 2018a), because multiple soil ecological processes (e.g. nutrition cycling) critical for cropping are driven by the functional traits of soil microbes (Martiny et al., 2015). The importance of microbial taxa within ecological networks for maintaining potential functioning in these ecosystems, however, remains unclear. We performed qPCR analyses to estimate the abundances of the functional genes involved in C, N. P and S cycling and sequenced amplicons of 16Sr RNA and ITS to obtain information on the diversity and community composition of bacteria and fungi.

2. Methods and materials

2.1. Site description, DNA extraction and soil variability

We collected 243 soil samples from 27 sites from 20 November to 30 November 2014 (Fig. S1A) to survey the patterns of soil microbial communities across the North China Plain. The rotation system is winter wheat and summer corn in this region (Shi et al., 2019a,b). All the sampling soils in our studying area were identified as Ochric Aquic Cambosols (Chinese soil taxonomy). Briefly, nine plots were selected at each site (100 km^2), with 3.3 km between any two plots (Fig. S1B). We collected 12 cores in each plot to a depth of 15 cm, which were mixed to form one sample per plot, and all 243 samples were stored in ice boxes and immediately shipped to the laboratory. Soil DNA was extracted and soil variability was measured as described by Shi et al. (2018). The information of sampling sites and soil characteristics were also described in Tables S1 and S2

2.2. Data analysis

High-throughput bacterial DNA analysis has been well described (Shi et al., 2018; Caporaso et al., 2012). Briefly, using a common primer set (515F, 5′-GTGCCAGCMGCCGCGGTAA-3′; 806R, 5′-GGACTACHVGGGTWTCTAAT-3′) combined with adapter sequences and barcode sequences, the bacterial 16Sr RNA V4 hyper-variable regions were amplified (Caporaso et al., 2011). The fungal ITS1 region was amplified using a common primer set (ITS1-F: CTTGGTCATTTAG AGGAAGTAA, ITS2-R: GCTGCGTTCTTCATCGATGC) combined with adapter and barcode sequences (Caporaso et al., 2011). Triplicates of each sample were amplified in a 50-μl reaction under the following conditions: 30 cycles of denaturation at 94 ˚C for 30 s, annealing at 54 ˚C for 30 s and extension at 72 ˚C for 1 min with a final extension at 72 ˚C for 7 min. The triplicate PCR products for each sample were pooled for both the bacteria and fungi, purified using a QIA quick PCR purification kit (Qiagen, Hilden, Germany) and then quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, USA). Equimolar ratios of PCR products were mixed in a single tube and sequenced using an Illumina MiSeq platform (Caporaso et al., 2012).

A total of 243 DNA samples were analyzed by high-throughput qPCR for 71 functional genes (e.g. C, N and P cycling) on a Wafergen Smart Chip Real-Time qPCR platform as described by Wang et al. (2014) and Zhu et al. (2016). Each DNA sample was amplified by PCR in triplicate at a threshold cycle of 31 as the detection limit. The same gene was regarded as only a single unique functional gene when targeted by multiple primer sets. The relative abundances (gene copies normalized to the 16S rRNA gene) of functional genes for each sample were calculated as described by Pfaffl (2001). The high-throughput fungal data were processed and analyzed using QIIME 1.9.1, with the workflow available at http://nbviewer.ipython.org/github/biocore/qiime/blob/1.9.1/examples/ipynb/illumina_overview_tutorial.ipynb, following the method described by Caporaso et al. (2012). Each bacterial phylotype was taxonomically identified using the Ribosomal Database Project classifier (Wang et al., 2007) trained on the Greengenes 13_8 16S rRNA database (McDonald et al., 2012). Each fungal OTU was taxonomically identified using the UNITE/QIIME ITS reference OTU database (Koljalg et al., 2013). For resampling, 20 005 bacterial sequences and 11 076 fungal sequences were randomly selected to ensure that all data sets were equally sampled.

2.3. Co-occurrence analyses

Co-occurrence network analyses were conducted based on a SparCC correlation matrix using the *WGCNA* package in R (Langfelder and Horvath, 2012). The OTU table was filtered with two constraints to improve the reliability of the network: only OTUs in > 30% of all samples and only OTUs with average relative abundances > 0.01% were retained. A total of 4473 bacterial OTUs and 685 fungal OTUs were thereby retained and used to construct the co-occurrence network. False discovery rates were corrected to adjust the *P* value of the correlation network. Before constructing the network, the reliability of the association was tested using the method described by Cougoul et al (2019). Network properties were calculated using the R igraph package (Csardi and Nepusz, 2006). Network images were generated using Gephi (http://gephi.github.io/). Random networks with the same numbers of nodes and edges as the empirical network were generated, and topological indices were summarized by 999 iterations based on the Erdös-Réyni model (Erdős 1960) to determine if the network properties were prone to errors. The entire network contained 1239 nodes and 37 203 edges (Table S3). Our correlation network contained eight modules (Fig. S2). The microbial community which is structured into microbial modules indicate the different groups of organisms which had distinct associations (de Menezes et al., 2015).

In our study, the calculations of z-score and c-score were according to the methods described by Guimera and Amaral (2005). Here, in order to vividly describe the role of hubs, kinless hubs (z-score > 2.5; c-score > 0.62), provincial hubs (z-score > 2.5; c-score ≤ 0.62), connectors (z-score ≤ 2.5; c-score > 0.62) and peripherals (z-score ≤ 2.5; c-score ≤ 0.62) were defined according to their within-module degree (z-score) and participation coefficient (c-score)

threshold value (Poudel et al., 2016). The kinless hubs mean the nodes were highly connected both within and between modules, these taxa are expected to mediate interactions species within and between modules and might be very important for functioning by allowing the exchange of energy and matter within and between modules; the provincial hubs mean the nodes were highly connected within a module, these taxa are expected to mediate interactions species within modules and might be very important for functioning by allowing the exchange of energy and matter within modules; the connectors were highly connected between modules, these taxa are expected to mediate interactions species between modules and might be very important for functioning by allowing the exchange of energy and matter between modules; and the peripherals mean the node had few or no links with other nodes (Guimera and Amaral, 2005; Toju et al., 2018b).

2.4. Statistical analysis

We conducted reduced major axis regression to determine the relationships between network hubs, soil properties and the functional genes and used a structural equation model (SEM) based on AMOS to infer the direct and indirect effects of soil variables, kinless hubs, connectors, provincial hubs and peripherals on the soil functions. Functional potential in the SEM was calculated using the average standardized abundance of all functional genes. Kinless hubs were identified using the average standardized abundance of the kinless -hub OTUs. Similarly, connector hubs, provincial hubs and peripherals were identified as above. The families with relative abundance > 0.5% were chosen to perform the heat maps. Z-score (*(i-m)/s*, i is the abundance of an OTU in each site, *m* is the mean value of the abundance of the OUT across the whole sites, *s* is the standard deviation of the abundance of the OUT across the whole sites) was calculated for each family in order to eliminate the disequilibrium of sequences among families. Package *pheatmap* (https://CRAN.R-project.org/package=pheatmap) was used to generate the heat maps.

2.5. Availability of data and materials

The data sets were deposited in the NCBI Sequence Read Archive with accession numbers SRP100578 (bacteria) and SRP116613 (fungi). Functional-gene data are shown in Table S4. More data supporting our findings are available in the supplementary information and from the corresponding authors upon reasonable request.

3. Results

3.1. General information of fungal community and functional genes

Soil bacterial diversity and community composition were described by Shi et al. (2018).

Briefly, in total 75 179 bacterial operational taxonomic units (97% similarity) were identified from 15 184 073 high-quality sequences across 243 soil samples. Actinobacteria, Alphaproteobacteria, Acidobacteria, Gammaproteobacteria, Betaproteobacteria, and Chloroflexi were the dominant at the phyla or class level, which were accounting for > 75% of total sequences. In this study, the dominant bacterial families (relative abundance > 0.5%) were shown in the heatmap and Acidobacteriaceae had the highest relative abundance in LB (Fig. S3). For the bacterial community composition, it showed strong correlation with soil pH. Using Illumina next-generation sequencing technologies, we obtained 11 076 fungal quality sequences for 243 soils and identified 5 068 operational taxonomic units (97% similarity). At the phyla level, Ascomycota (~85.3%), Basidiomycota (~6.0%) and Zygomycota (~5.1%) were dominant, accounting for > 96% of total sequences (Fig. S4, Table S5). According to the heatmap, Sordariaceae, Herportrichiellaceae and Microascaceae belonging to Ascomycota had the highest relative abundance in SZ, LB and ZX (Fig. S5). Using Mantel test, we found that the soil fungal community composition significantly correlated with soil pH, soil moisture, TN and TP (Table S6), with soil pH having the strongest influence on fungal community. Fungal alpha diversity indexes such as Simpson, Shannon, Chao1 and observed species, significantly decreased with increasing soil DOC, TN and AP (Table S7). From the 27 North China Plain sites that were sampled in nine plots, we detected a total of 71 different functional genes involved in Carbon (fixation, methanol, degradation), nitrogen, phosphorus and sulfur cycling. The relative abundance of these genes across the site were well described in Table S4. The heatmap showed that most of the functional genes had high abundance in site LQ, YJ, MJ and DM (Fig S6.). Using Mantel test, we found that the soil functional community significantly correlated with soil pH, TP and TK (Table S8). Most of functional genes showed significantly positive correlations with soil pH and TK and negative correlations with DTN and DON (Table S9).

3.2. General information of the network

We identified eight ecological clusters within our correlation network. Most microbial species (nodes) within our ecological network (94%) were classified as peripherals, 3% as connector hubs, 3% as provincial hubs and only 0.2% as kinless hubs (Fig. 1A). These network roles were not evenly distributed over the identified ecological clusters. All kinless (two phylotypes) hubs were in module 7 (Fig. 1B) and were classified as bacterial taxa. These two relatively rare kinless hubs accounted for 0.17% of all reads for soil bacteria. Kinless hubs included phylotypes from the classes Gemmatimonadetes and Phycisphaerae. Most connectors were in modules 6 and 8. Connector hubs included the phyla Proteobacteria, Actinobacteria, Acidobacteria, Bacteroidetes, Verrucomicrobia, Zygomycota and Ascomycota. Provincial hubs were in modules 11, 2 and 4. The community composition of each ecological cluster is shown in Fig. S7 and S8, and Table S10. The taxonomy for the provincial and connector hubs is presented in Table S11.

3.3. Links between the relative abundances of hubs and functional genes

Our study provided empirical evidence that the relative abundance of taxa classified as kinless hubs within the ecological network were positively and significantly correlated with the abundance of functional genes involved in C, N, P and S cycling (Fig. 2, Table 1). The relative abundances of taxa classified as connector hubs were weakly positively correlated with the abundances of the soil functional genes (Fig. S9, Table S12). In contrast, the relative abundances of taxa classified as module hubs and peripherals were negatively correlated with the abundances of these functional genes (Fig. S10 and S11, Table S12).

In order to confirm the links between the relative abundances of hubs and functional genes, we fitted structure equation models (SEM) after accounting for multiple soil variables (Table 1, Figs. S12-S15, Table S13). The SEM results confirmed the positive association between the relative abundance of kinless hubs and the abundances of multiple functional genes ($\beta = 0.32$, standardized regression weights) (Fig. 3A, Table S14), and negative association between provincial hubs, peripherals and functional genes ($\beta = -0.27$ and -0.26, standardized regression weights, respectively) (Fig. 3B and D). The relative abundance of connector hubs was weakly correlated with soil functions ($\beta = 0.01$, standardized regression weights) (Fig. 3C). The relative abundance of taxa within different network roles was associated with different soil properties. For example, pH, TP, TK, Ca and Cd positively and directly affected the kinless hubs (Fig. 3A). Soil AP and Mg strongly and positively affected the provincial hubs (Fig. 3B), and the abundance of taxa within the connector hubs was positively correlated with soil pH and Cu (Fig. 3C). Almost none of the factors positively affected the peripherals (Fig. 3D, Table S15). The soil variables had different effects on the hubs and functional genes, but pH directly affected all hubs and functional genes, e.g. positive effects on kinless and connector hubs and functional genes, and negative effects on provincial and peripheral hubs.

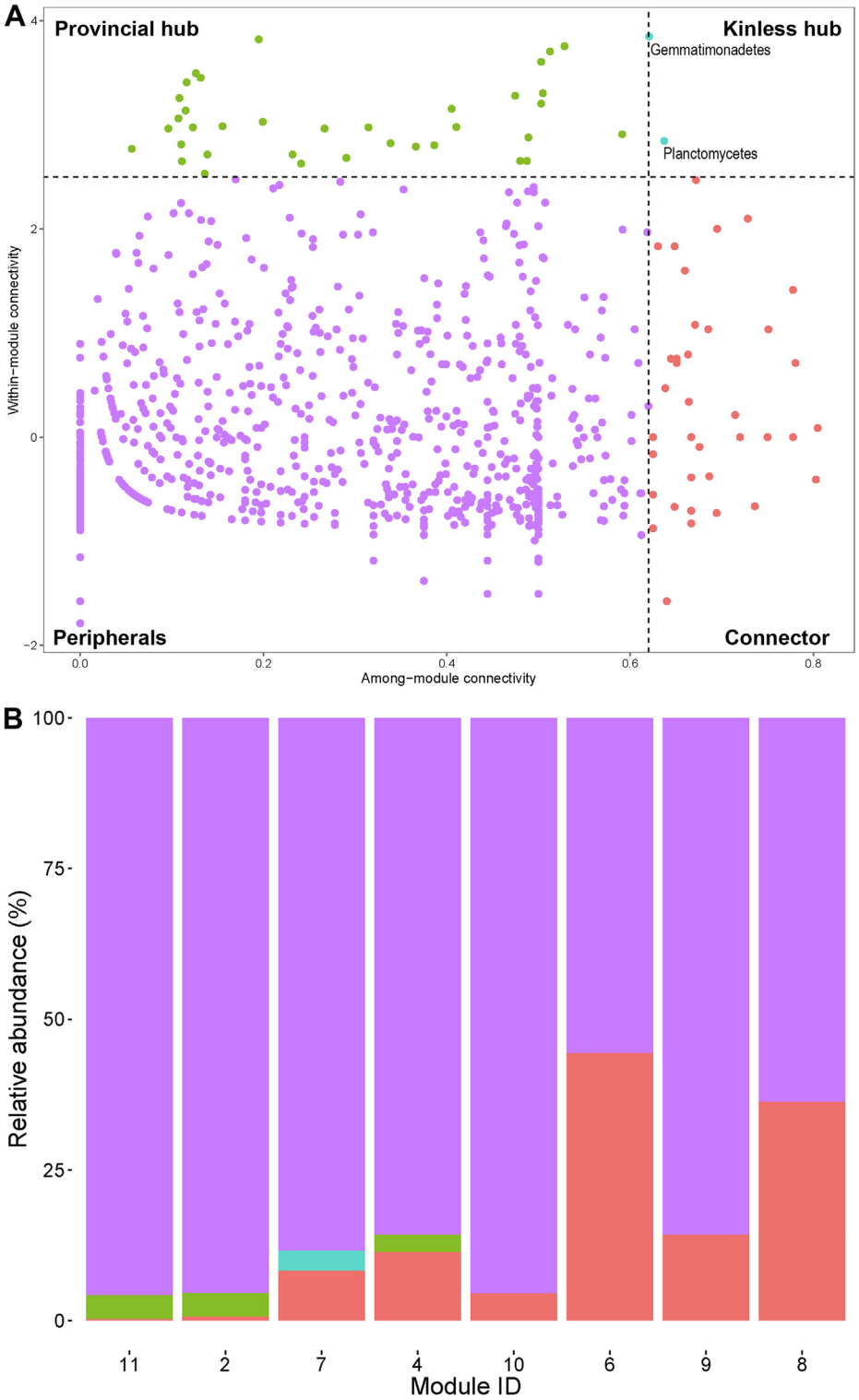

Fig. 1. A, Distribution of bacterial and fungal OTUs based on their network roles. Nodes in the network are classified as peripherals, modular hubs, network hubs or connectors depending on their role in the network. B, Relative abundance of peripherals (purple), provincial hubs (green), kinless hubs (blue) and connectors (red) among the modules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

Our study provides evidence that the relative abundance of taxa classified as kinless-hub microbes within ecological networks are associated with high levels of functional potential, similar to plant communities. The relative abundance of taxa classified as connector hubs was also important for maintaining multiple functional genes within the network, but the relative abundances of taxa classified as peripheral and provincial were negatively correlated with the abundance of functional genes involved in C, N, P and S cycling. These results remained valid even after accounting for multiple soil properties. Our study indicated that relatively rare but highly connected taxa within a correlation network played an important role in driving the abundance of functionally important genes. These results remained valid after

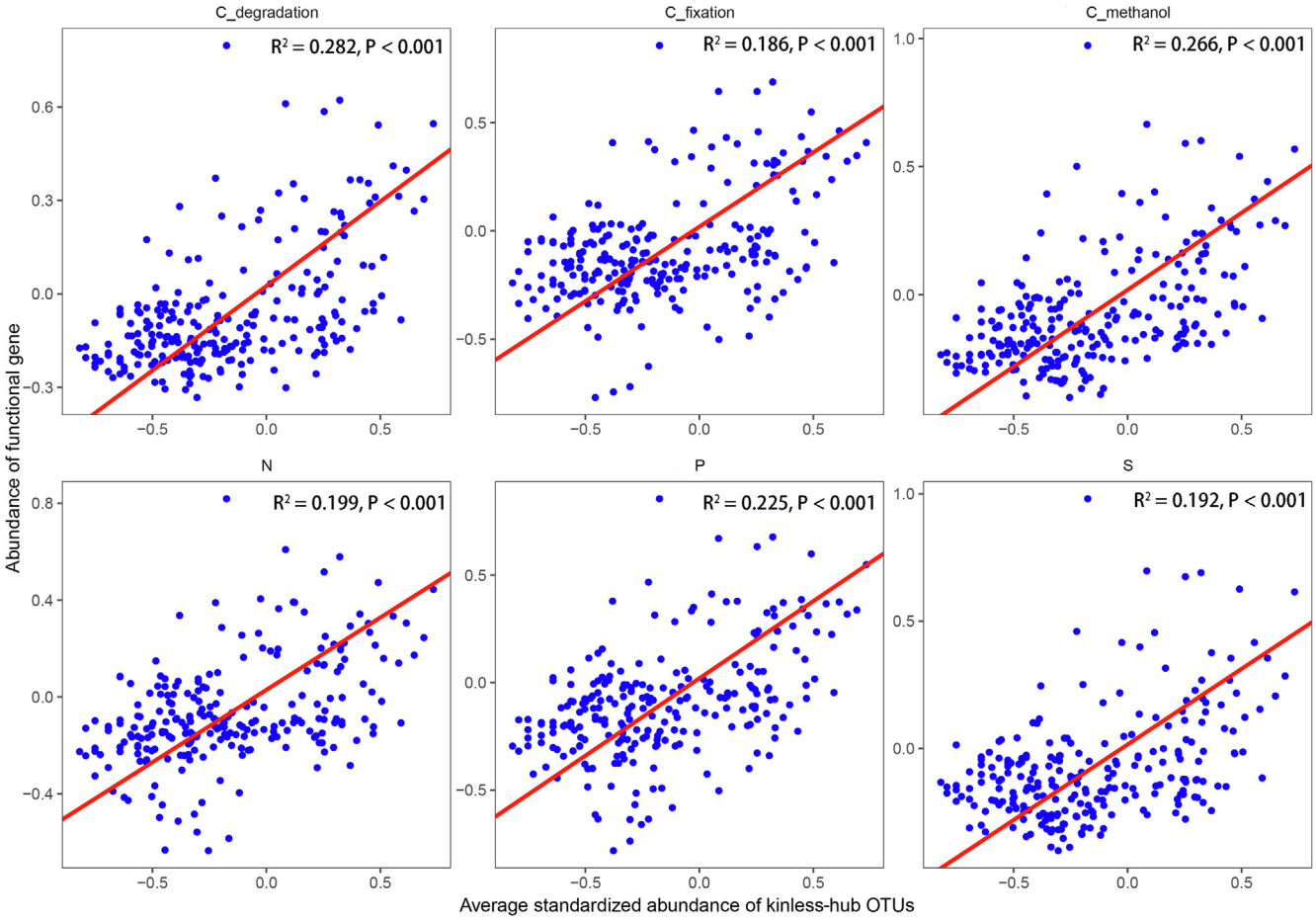

Fig. 2. Relationships between the average standardized abundance of the kinless-hub OTUs (OTU237047, OTU575940) and the abundances of the functional genes (both OTUs and genes were lg transformed), which are grouped by C_degradation (carbon degradation), C_fixation (carbon fixation), C_methanol (carbon methanol), N (nitrogen cycling), P (phosphorus cycling) and S (sulfur cycling). Reduced major axis regression analysis can be found in Table S12.

considering the influence of multiple soil properties simultaneously using an SEM, i.e. locations with a high abundance of kinless taxa were positively associated with high levels of functional potential. Our results provide evidence that taxa playing different roles within the ecological network differentially influence the abundance of functional genes across environmental gradients.

4.1. The relative abundance of kinless taxa was positively associated with the abundance of functional genes

We quantified the abundance of 71 functional genes involved in C, N, S and P cycling (Table S4) across 27 sites on the North China Plain. More than 90% of the functional genes were positively correlated with the relative abundance of two identified kinless hubs, and about 70% of these genes were positively correlated with the relative abundance of taxa classified as connector hubs (Table S16). The importance of kinless and connector hubs as regulators of functional potential may thus be associated with their topological role within the network. Kinless and connector hubs were correlated with other taxa within and between ecological clusters (C scores > 0.62), suggesting that these taxa can play multiple roles in maintaining multiple ecosystem processes by promoting nutrient exchange and resource availability. Our results highlight the importance of these taxa for multiple functional genes involved in the regulation of multiple ecosystem functions, including C and N fixation, P mineralization and litter decomposition. The relative abundance of taxa classified as provincial and peripheral was negatively associated with the abundance of functional genes, unlike kinless and connector taxa, supporting the results of a long-term field-scale experiment where provincial hubs were negatively correlated with C-cycling genes (Deng et al., 2015). Provincial hubs were highly correlated with taxa within a module, and they had almost no links with other modules (Deng et al., 2015). Therefore, high abundance of them probably displayed fierce competition with members in other modules. This may be not helpful in forming the ecological networks entirely. Because the whole ecological network contains multiple modules and carry out various functions, some high abundance of phyla in a module may be harmful to the stability of ecological network. This might be the reason why we found that some high abundance of phyla in the module had negative correlation with the functional genes. Analogically, high abundance of peripherals may also influence the whole ecological network performing multifunction.

4.2. The relative abundance of rare but highly connected kinless and connector hubs which dominate the soil functioning are independent of their abundance

The strong positive and significant associations between the abundance of key functional genes and the relative abundance of rare but highly connected kinless and connector hubs suggest that these species depend on its position in the microbial network and are independent of their abundance (Banerjee et al., 2018). This means they may play fundamental roles in monitoring the functioning of multiple ecosystem processes in economically important agricultural ecosystems, including wheat and maize cropping. For example, the strongest positive correlation between exoPG (pectinase), pmoA (methane/ammonia monooxygenase subunit A) and kinless hubs may indicate a strong ability to

Table 1
Heatmap of the strength of the correlations between kinless hubs, connectors, provincial hubs, peripherals and soil factors. Darker colors (blue is negative, red is positive) represent stronger correlations; the correlation coefficient (r) is indicated only for significant correlations ($P < 0.05$).

Soil property	Kinless_hub	Connector hub	Provincial_hub	Peripherals
pH	0.344	0.476	-0.685	-0.232
SM	0.207	0.244		-0.188
OC				
DOC	-0.213	-0.228	0.18	
TN				
DTN	-0.323	-0.218		0.201
DON	-0.423	-0.202		0.285
DIN		-0.155		
TP		0.195	-0.437	-0.152
TK	0.358	0.237	-0.324	-0.217
AP	-0.293	-0.355	0.324	0.174
AK	0.368	0.249		-0.183
EC			-0.215	
Mg	0.43	0.366	-0.55	-0.268
Ca	0.435	0.252	-0.512	-0.199
K	0.16		-0.263	-0.217
Fe	0.367	0.216	-0.177	-0.332
Cr	0.36	0.273	-0.131	-0.178
Mn	0.196			-0.316
Cu	0.481	0.314	-0.292	-0.333
Zn	0.262	0.214	-0.186	-0.22
Cd	0.471	0.252	-0.397	-0.232
Pb	0.213	0.222		
As	0.464	0.172	-0.229	
Functional gene				
C_degradation	0.545	0.17	-0.358	-0.348
C_fixation	0.5	0.18	-0.346	-0.313
C_methanol	0.549	0.304	-0.477	-0.388
N_cycling	0.481	0.288	-0.496	-0.382
P_cycling	0.517	0.267	-0.49	-0.342
S_cycling	0.464	0.146	-0.3	-0.291

Cd, cadmium; K, potassium; Ca, calcium; TK, total potassium; TP, total phosphorus; Mg, magnesium; Cu, copper; AP, available phosphorus; DOC, dissolved organic carbon; Fe, iron; Mn, manganese; SM, soil moisture; OC, organic carbon; TN, total nitrogen; DTN, dissolved total nitrogen; DON, dissolved organic nitrogen; DIN, dissolved inorganic nitrogen; AK, available potassium; EC, Conductivity; Pb, lead; Zn, zinc; Cr, chromium; C_degradation, carbon degradation; C_fixation, carbon fixation; C_methanol, carbon methanol function; N_cycling, nitrogen cycling; P_cycling, phosphorus cycling; S_cycling, sulfur cycling.

accumulate N (Kuypers et al., 2018). Zeng et al., found that Gemmatimonadetes contained chlorophyll-based phototrophic species (Zeng et al., 2014), suggesting a strong ability to support fundamental biological processes (Hohmann-Marriott and Blankenship, 2011). These studies support the idea that although the species are rare, they are likely to provide complementary or unique metabolic pathways to service the ecosystem (Jousset et al., 2017). In addition, taxa as keystone species might be context dependent, and they have large influence on the ecosystem under certain conditions (Jiao et al., 2017; Banerjee et al., 2018). Taxa within the Gemmatimonadetes phylum can adapt to dry conditions (Drees et al., 2006). The North China Plain is generally arid (Jeong et al., 2014), further suggesting that this taxon can be functionally important in globally dominant ecosystems, such as those in drylands (Girvan et al., 2005). For example, Gemmatimonadetes were identified as keystone species in tobacco soils in North China (Chen et al., 2019). The other identified kinless hub belonged to Phycisphaerae within Planctomycetes which had been found in various habitats (Spring et al., 2018). For example, some species have ability in

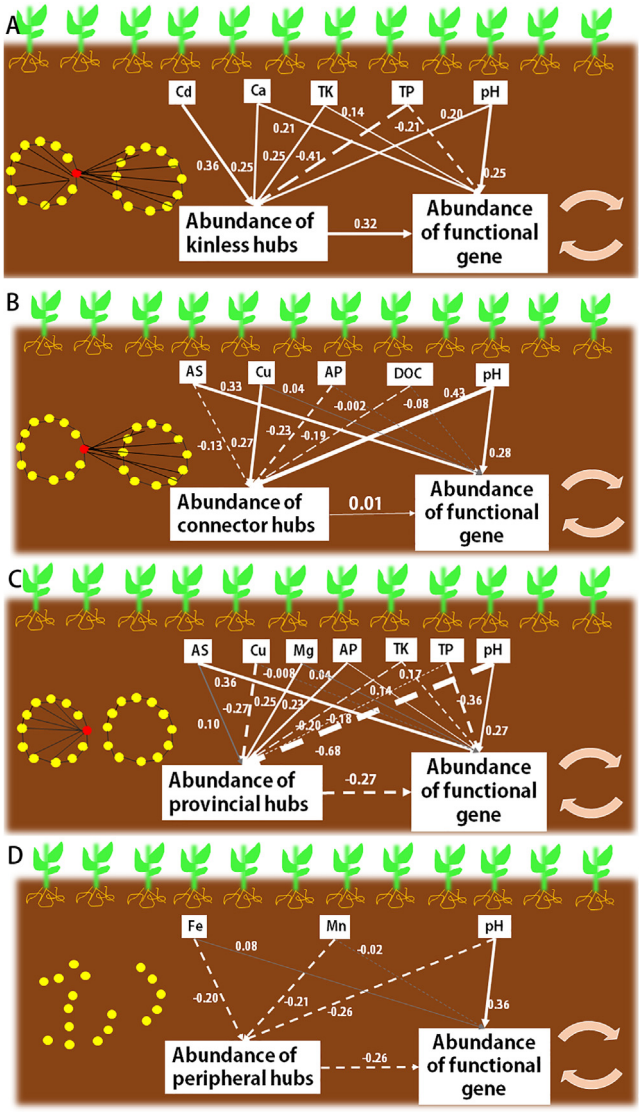

Fig. 3. Structural equation model of the relationships of the soil variables, and abundance of the functional genes with kinless hubs (A), provincial hubs (B), connector hubs (C) and peripheral hubs (D). Solid white arrows represent significantly positive paths ($P < 0.05$), dashed white arrows represent significantly negative paths ($P < 0.05$), solid gray arrows represent non-significantly positive paths ($P > 0.05$) and dashed gray arrows represent non-significantly negative paths ($P > 0.05$). The path coefficients (standardized effect sizes) are reported. The details of model fitness are presented in Table S3. Cd, cadmium; K, potassium; Ca, calcium; TK, total potassium; TP, total phosphorus; Mg, magnesium; Cu, copper; AP, available phosphorus; DOC, dissolved organic carbon; Fe, iron and Mn, manganese; AS, asenic. Here, the yellow and red dots are the nodes in the network. Yellow dots are the common nodes in the modules, and the red dots are the hubs in the modules. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

oxidizing ammonium under anaerobic condition (Jetten et al., 2001), and with this process, these species could live in hot springs (Tekere et al., 2011) or Antarctic water (Piquet et al., 2010). Moreover, Planctomycetes have often been associated with the rhizospheres of plants (Mendes et al., 2013), such as cotton (Qiao et al., 2017), wild beet (Zachow et al., 2014) and winter wheat (Mahoney et al., 2017). Rhizospheres are loci of nutrient exchange and elemental cycling (Kwak et al., 2018) and harbor millions of microbes (Xu et al., 2018) that improve nutrient cycling or suppress pathogens most likely to be recruited (Berendsen et al., 2012). Planctomycetes may therefore indirectly promote plant growth by being involved in multiple ecological processes. Future studies should aim to culture these two microbial taxa and test their potential applicability to promote ecosystem function in cropland ecosystems under controlled conditions.

We also observed some significant and positive associations between the relative abundance of taxa classified as connector hubs and the absolute abundance of functional genes across soil samples. Connector taxa included both bacterial and fungal taxa, unlike kinless taxa. A connector belonging to Sordariomycetes, a class of Ascomycota, were positively correlated with most of the genes (Table S17). Sordariomycetes is prevalent in various soil habitats and has been widely reported as a decomposer of leaf litter and wood (Lutzoni et al., 2004; Spatafora and Blackwell, 1993). Another Verrucomicrobia connector was significantly positively correlated with most of the genes (Table S17, the representative sequences can be found in Table S18). The abundance of Verrucomicrobia has been positively correlated with a variety of genes associated with C cycling (Fierer et al., 2013). It's clear that the untested species in our study such as archaea (Shi et al., 2019a,b) or eukaryotic microbes (Xiong et al., 2020) could play critical role in the soil functions, and they may have keystone species in agricultural soils (Banerjee et al., 2018). In addition, the role of plant in keystone species was not considered, because only buck soils were collected in our study. Our results provide a direct link between the abundance of highly connected taxa within a microbial network and soil functional genes, suggesting that studying the abundance of these hub-related species would provide baseline information on monitoring or improving ecosystem functions in agricultural ecosystem.

5. Conclusions

Taken together, our findings suggest that the relative abundance of a reduced number of highly connected taxa is fundamental for maintaining high functional potential, with implications for understanding the functioning of intensive agricultural ecosystems at a large spatial scale. Aiming to culture the thousands of microbial species in our soils is not feasible, however, targeting the reduced subset of keystone species identified in this study could potentially have important implications for monitoring ecosystem functions under human-managed ecosystems.

Declaration of Competing Interest

The authors declare that they have no competing interests.

CRediT authorship contribution statement

Yu Shi: Methodology. **:** Data curation, Visualization, Investigation, Validation, Writing - review & editing. **Manuel Delgado-Baquerizo:** Writing - review & editing, Validation. **Yuntao Li:** Investigation, Methodology. **: .** **Yunfeng Yang:** Writing - review & editing. **Yong-Guan Zhu:** Conceptualization, Supervision, Writing - review & editing. **Josep Peñuelas:** Writing - review & editing. **Haiyan Chu:** Conceptualization, Supervision, Writing - review & editing, Validation.

Acknowledgements

We thank J. Q. Su, B. X. Zheng, T. Yang, X. J. Xiang, C. C. Shen, R. B. Sun, D. He, K. P. Zhang, Y. Y. Ni and K. K. Fan for their assistance in soil sampling and laboratory analyses. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB15010101), the National Key Research and Development Program of China (2017YFD0200604) and the China Biodiversity Observation Networks (Sino BON). JP was financially supported by the European Research Council Synergy grant (ERC-SyG-2013-610028 IMBALANCE-P). M.D-B. is supported by a Ramón y Cajal grant from the Spanish Government (agreement n° RYC2018-025483-I) and by the BES grant agreement n° LRA17\1193 (MUSGONET).

Author contributions

HC and YZ designed the experiments. YS and YL collected the samples and determined the soil variables. YS and YL performed bioinformatics. YS, MD-B, JP, YY, YZ and HC wrote the manuscript.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2020.105869.

References

Banerjee, S., Schlaeppi, K., van der Heijden, M.G.A., 2018. Keystone taxa as drivers of microbiome structure and functioning. Nat. Rev. Microbiol. 16, 567–576.

Berendsen, R.L., Pieterse, C.M.J., Bakker, P.A.H.M., 2012. The rhizosphere microbiome and plant health. Trends Plant Sci. 17, 478–486.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. PNAS 108, 4516–4522.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 6, 1621–1624.

Cavieres, L.A., Brooker, R.W., Butterfield, B.J., Cook, B.J., Kikvidze, Z., Lortie, C.J., Michalet, R., Pugnaire, F.I., Schob, C., Xiao, S., Anthelme, F., Bjork, R.G., Dickinson, K.J.M., Cranston, B.H., Gavilan, R., Gutierrez-Giron, A., Kanka, R., Maalouf, J.P., Mark, A.F., Noroozi, J., Parajuli, R., Phoenix, G.K., Reid, A.M., Ridenour, W.M., Rixen, C., Wipf, S., Zhao, L., Escudero, A., Zaitchik, B.F., Lingua, E., Aschehoug, E.T., Callaway, R.M., 2014. Facilitative plant interactions and climate simultaneously drive alpine plant diversity. Ecol. Lett. 17, 193–202.

Chen, L., Jiang, Y., Liang, C., Luo, Y., Xu, Q., Han, C., Zhao, Q., Sun, B., 2019. Competitive interaction with keystone taxa induced negative priming under biochar amendments. Microbiome 7, 77.

Cougoul, A., Bailly, X., Vourc'h, G., Gasqui, P., 2019. Rarity of microbial species: in search of reliable associations. PLoS ONE 14, 1–15.

Csardi, G., Nepusz, T., 2006. The igraph software package for complex network research. Int. J. Comp. Sys.

de Menezes, A.B., Prendergast-Miller, M.T., Richardson, A.E., Toscas, P., Farrell, M., Macdonald, L.M., Baker, G., Wark, T., Thrall, P.H., 2015. Network analysis reveals that bacteria and fungi form modules that correlate independently with soil parameters. Environ. Microbiol. 17, 2677–2689.

Delgado-Baquerizo, M., Reich, P.B., Trivedi, C., Eldridge, D.J., Abades, S., Alfaro, F.D., Bastida, F., Berhe, A.A., Cutler, N.A., Gallardo, A., Garcia-Velazquez, L., Hart, S.C., Hayes, P.E., He, J.Z., Hseu, Z.Y., Hu, H.W., Kirchmair, M., Neuhauser, S., Perez, C.A., Reed, S.C., Santos, F., Sullivan, B.W., Trivedi, P., Wang, J.T., Weber-Grullon, L., Williams, M.A., Singh, B.K., 2020. Multiple elements of soil biodiversity drive ecosystem functions across biomes. Nat. Ecol. Evol. 4, 210–220.

Deng, Y., Zhang, P., Qin, Y., Tu, Q., Yang, Y., He, Z., Schadt, C.W., Zhou, J., 2015. Network succession reveals the importance of competition in response to emulsified vegetable oil amendment for uranium bioremediation. Environ. Microbiol. 18, 205–218.

Drees, K.P., Neilson, J.W., Betancourt, J.L., Quade, J., Henderson, D.A., Pryor, B.M., Maier, R.M., 2006. Bacterial community structure in the hyperarid core of the Atacama Desert. Chile. Appl. Environ. Microbiol. 72, 7902–7908.

Ellison, A.M., Bank, M.S., Clinton, B.D., Colburn, E.A., Elliott, K., Ford, C.R., Foster, D.R., Kloeppel, B.D., Knoepp, J.D., Lovett, G.M., Mohan, J., Orwig, D.A., Rodenhouse, N.L., Sobczak, W.V., Stinson, K.A., Stone, J.K., Swan, C.M., Thompson, J., Von Holle, B., Webster, J.R., 2005. Loss of foundation species: consequences for the structure and dynamics of forested ecosystems. Front. Ecol. Environ. 3, 479–486.

Erdős, P., 1960. On sets of distances of n points in Euclidean space. Magyar Tudom. Akad. Matem. Kut. Int. Közl. Publ. Math. Inst. Hung. Acad. Sci. 5, 165–169.

Faust, K., Lima-Mendez, G., Lerat, J.S., Sathirapongsasuti, J.F., Knight, R., Huttenhower, C., Lenaerts, T., Raes, J., 2015. Cross-biome comparison of microbial association networks. Front. Microbiol. 6, 1200.

Fierer, N., Ladau, J., Clemente, J.C., Leff, J.W., Owens, S.M., Pollard, K.S., Knight, R., Gilbert, J.A., McCulley, R.L., 2013. Reconstructing the microbial diversity and function of pre-agricultural tallgrass prairie soils in the United States. Science 342, 621–624.

Girvan, M.S., Campbell, C.D., Killham, K., Prosser, J.I., Glover, L.A., 2005. Bacterial diversity promotes community stability and functional resilience after perturbation. Environ. Microbiol. 7, 301–313.

Guimera, R., Amaral, L.A.N., 2005. Functional cartography of complex metabolic networks. Nature 433, 895–900.

Han, D., Wiesmeier, M., Conant, R.T., Kuehnel, A., Sun, Z.G., Koegel-Knabner, I., Hou, R.X., Cong, P.F., Liang, R.B., Ouyang, Z., 2018. Large soil organic carbon increase due to improved agronomic management in the North China Plain from 1980s to 2010s. Glob. Change Biol. 24, 987–1000.

Hohmann-Marriott, M.F., Blankenship, R.E., 2011. Evolution of photosynthesis. Annu. Rev. Plant Biol. 62, 515–548.

Jetten, M.S.M., Wagner, M., Fuerst, J., van Loosdrecht, M., Kuenen, G., Strous, M., 2001. Microbiology and application of the anaerobic ammonium oxidation ('anammox') process. Curr. Opin. Biotech. 12, 283–288.

Jeong, S.J., Ho, C.H., Piao, S.L., Kim, J., Ciais, P., Lee, Y.B., Jhun, J.G., Park, S.K., 2014. Effects of double cropping on summer climate of the North China Plain and neighbouring regions. Nat. Clim. Change 4, 615–619.

Koljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F.S., Bahram, M., Bates, S.T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., Douglas, B., Drenkhan, T., Eberhardt, U., Duenas, M., Grebenc, T., Griffith, G.W., Hartmann, M., Kirk, P.M., Kohout, P., Larsson, E., Lindahl, B.D., Luecking, R., Martin, M.P., Matheny, P.B., Nguyen, N.H., Niskanen, T., Oja, J., Peay, K.G., Peintner, U., Peterson, M., Poldmaa, K., Saag, L., Saar, I., Schuessler, A., Scott, J.A., Senes, C., Smith, M.E., Suija, A., Taylor, D.L., Telleria, M.T., Weiss, M., Larsson, K.H., 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol. Ecol. 22, 5271–5277.

Jiao, S, Chen, W.M., Wei, G.H., 2017. Biogeography and ecological diversity patterns of rare and abundant bacteria in oil-contaminated soils. Mol. Ecol. 26, 5305–5317.

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., Kusel, K., Rillig, M.C., Rivett, D.W., Salles, J.F., van der Heijden, M.G.A., Youssef, N.H., Zhang, X.W., Wei, Z., Hol, W.H.G., 2017. Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J. 11, 853–862.

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. Nat. Microbiol. https://doi.org/10.1038/s41564-019-0560-0.

Kwak, M.J., Kong, H.G., Choi, K., Kwon, S.K., Song, J.Y., Lee, J., Lee, P.A., Choi, S.Y., Seo, M., Lee, H.J., Jung, E.J., Park, H., Roy, N., Kim, H., Lee, M.M., Rubin, E.M., Lee, S.W., Kim, J.F., 2018. Rhizosphere microbiome structure alters to enable wilt resistance in tomato. Nat. Biotechnol. 2018. https://doi.org/10.1038/nbt.4232.

Kuypers, M.M.M., Marchant, H.K., Kartal, B., 2018. The microbial nitrogen-cycling network. Nat. Rev. Microbiol. 16, 263–276.

Langfelder, P., Horvath, S., 2012. Fast R functions for robust correlations and hierarchical clustering. J. Stat. Softw. 46, 1–17.

Lutzoni, F., Kauff, F., Cox, C.J., McLaughlin, D., Celio, G., Dentinger, B., Padamsee, M., Hibbett, D., James, T.Y., Baloch, E., Grube, M., Reeb, V., Hofstetter, V., Schoch, C., Arnold, A.E., Miadlikowska, J., Spatafora, J., Johnson, D., Hambleton, S., Crockett, M., Shoemaker, R., Hambleton, S., Crockett, M., Shoemaker, R., Sung, G.H., Lucking, R., Lumbsch, T., O'Donnell, K., Binder, M., Diederich, P., Ertz, D., Gueidan, C., Hansen, K., Harris, R.C., Hosaka, K., Lim, Y.W., Matheny, B., Nishida, H., Pfister, D., Rogers, J., Rossman, A., Schmitt, I., Sipman, H., Stone, J., Sugiyama, J., Yahr, R., Vilgalys, R., 2004. Assembling the fungal tree of life: progress, classification and evolution of subcellular traits. Am. J. Bot. 91, 1446–1480.

Mahoney, A.K., Yin, C., Hulbert, S.H., 2017. Community structure, species variation, and potential functions of rhizosphere-associated bacteria of different winter wheat (Triticum aestivum) cultivars. Front. Plant Sci. 8, 132.

Martiny, J.B., Jones, S.E., Lennon, J.T., Martiny, A.C., 2015. Microbiomes in light of traits: a phylogenetic perspective. Science 350, aac9323.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., Hugenholtz, P., 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 6, 610–618.

Mendes, R., Garbeva, P., Raaijmakers, J.M., 2013. The rhizosphere microbiome: significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. FEMS Microbiol. Rev. 37, 634–663.

Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 9, e25.

Piquet, A.M.T., Bolhuis, H., Davidson, A.T., Buma, A.G.J., 2010. Seasonal succession and UV sensitivity of marine bacterioplanktonatan Antarctic coastal site. FEMS Microbiol. Ecol. 73, 68–82.

Poudel, R., Jumpponen, A., Schlatter, D., Paulitz, T., Gardener, B.M., Kinkel, L., Garrett, K., 2016. Microbiome networks: a systems framework for identifying candidate microbial assemblages for disease management. Phytopathology 106, 1083–1096.

Qiao, Q.H., Wang, F.R., Zhang, J.X., Chen, Y., Zhang, C.Y., Liu, G.D., Zhang, H., Ma, C.L., Zhang, J., 2017. The variation in the rhizosphere microbiome of cotton with soil type, genotype and developmental stage. Sci. Rep. 7, 3940.

Rey, P.J., Alcantara, J.M., Manzaneda, A.J., Sanchez-Lafuente, A.M., 2016. Facilitation contributes to Mediterranean woody plant diversity but does not shape the diversity-productivity relationship along aridity gradients. New Phytol. 211, 464–476.

Rooney, N., McCann, K., Gellner, G., Moore, J.C., 2006. Structural asymmetry and the stability of diverse food webs. Nature 442, 265–269.

Shade, A., Handelsman, J., 2012. Beyond the Venn diagram: the hunt for a core microbiome. Environ. Microbiol. 14, 4–12.

Soliveres, S., Eldridge, D.J., Hemmings, F., Maestre, F.T., 2012. Nurse plant effects on plant species richness in drylands: the role of grazing, rainfall and species specificity. Perspect. Plant Ecol. 14, 402–410.

Shi, Y., Fan, K., Li, Y., Yang, T., He, J.-S., Chu, H., 2019a. Archaea enhance the robustness of microbial co-occurrence networks in Tibetan Plateau soils. Soil Sci. Soc. Am. J. 83, 1093–1099.

Shi, Y., Li, Y.T., Yuan, M.Q., Adams, J.M., Pan, X.Z., Yang, Y.F., Chu, H.Y., 2019b. A biogeographic map of soil bacterial communities in wheats field of the North China Plain. Soil Ecol. Lett. 1, 50–58.

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

Spatafora, J.W., Blackwell, M., 1993. Molecular systematics of unitunicate perithecial ascomycetes – the clavicipitales-hypocreales connection. Mycologia 85, 912–922.

Spring, S., Bunk, B., Sproer, C., Rohde, M., Klenk, H.P., 2018. Genome biology of a novel lineage of planctomycetes widespread in anoxic aquatic environments. Environ.

Microbiol. 20, 2438–2455.
Tekere, M., Lotter, A., Olivier, J., Jonker, N., Venter, S., 2011. Metagenomic analysis of bacterial diversity of Siloam hot water spring, Limpopo, South Africa. Afr. J. Biotechnol. 10, 18005–18012.
Toju, H., Peay, K.G., Yamamichi, M., Narisawa, K., Hiruma, K., Naito, K., Fukuda, S., Ushio, M., Nakaoka, S., Onoda, Y., Yoshida, K., Schlaeppi, K., Bai, Y., Sugiura, R., Ichihashi, Y., Minamisawa, K., Kiers, E.T., 2018a. Core microbiomes for sustainable agroecosystems. Nat. Plants 4, 247–257.
Toju, H., Tanabe, A.S., Sato, H., 2018b. Network hubs in root-associated fungal metacommunities. Microbiome 6, 116.
Wagg, C., Schlaeppi, K., Banerjee, S., Kuramae, E.E., van der Heijden, M.G.A., 2019. Fungal-bacterial diversity and microbiome complexity predict ecosystem functioning. Nat. Commun. 10, 4841.
Wang, F.H., Qiao, M., Su, J.Q., Chen, Z., Zhou, X., Zhu, Y.G., 2014. High throughput profiling of antibiotic resistance genes in urban park soils with reclaimed water irrigation. Environ. Sci. Technol. 48, 9079–9085.
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.
Xiong, W., Song, Y., Yang, K., Gu, Y., Wei, Z., Kowalchuk, G.A., Xu, Y., Jousset, A., Shen, Q., Geisen, S., 2020. Rhizosphere protists are key determinants of plant health. Microbiome. 8, 27.
Xu, J., Zhang, Y., Zhang, P., Trivedi, P., Riera, N., Wang, Y., Liu, X., Fan, G., Tang, J., Coletta-Filho, H.D., Cubero, J., Deng, X., Ancona, V., Lu, Z., Zhong, B., Roper, M.C., Capote, N., Catara, V., Pietersen, G., Vernière, C., Al-Sadi, A.M., Li, L., Yang, F., Xu, X., Wang, J., Yang, H., Jin, T., Wang, N., 2018. The structure and function of the global citrus rhizosphere microbiome. Nat. Commun. 9, 4894.
Zachow, C., Muller, H., Tilcher, R., Berg, G., 2014. Differences between the rhizosphere microbiome of Beta vulgaris ssp maritima – ancestor of all beet crops – and modern sugar beets. Front. Microbiol. 5, 415.
Zeng, Y.H., Feng, F.Y., Medova, H., Dean, J., Koblizek, M., 2014. Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. Proc. Natl. Acad. Sci. USA 111, 7795–7800.
Zhu, Y.G., Su, J.Q., Cao, Z.H., Xue, K., Quensen, J., Guo, G.X., Yang, Y.F., Zhou, J.Z., Chu, H.Y., Tiedje, J.M., 2016. A buried Neolithic paddy soil reveals loss of microbial functional diversity after modern rice cultivation. Sci. Bull. 61, 1052–1060.