Microbial phylogenetic relatedness links to distinct successional patterns of bacterial and fungal communities

Qiang Lin[a,e]*, Francisco Dini-Andreote[b,c], Travis B. Meador[a], Roey Angel[a], Lenka Meszárošová[d], Petr Heděnec[f,g], Lingjuan Li[a], Petr Baldrian[d], Jan Frouz[a,e]*

Running title: Ecological succession of bacteria and fungi

[a] Biology Centre of the Czech Academy of Sciences, Institute of Soil Biology & SoWa Research Infrastructure, Na Sádkách 7, CZ, 37005, České Budějovice, Czech Republic

[b] Department of Plant Science, The Pennsylvania State University, University Park, PA, USA

[c] Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA

[d] Laboratory of Environmental Microbiology, Institute of Microbiology of the Czech Academy of Sciences, Vídeňská 1083, 14220 Praha 4, Czech Republic

[e] Institute for environmental studies, Faculty of Science, Charles University, Benátská 2, 12800, Praha 2, Czech Republic

[f] Department of Geosciences and Natural Resource Management, Faculty of Science, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg C, Denmark

[g] Engineering Research Center of Soil Remediation of Fujian Province University, College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China

21 ***Corresponding author:** Qiang Lin, E-mail: qiangl2019@gmail.com and Jan Frouz, E-mail:

22 jan.frouz@natur.cuni.cz

Abstract

Development of soil microbial communities along ecological succession is crucial for ecosystem recovery and maintenance. However, mechanisms mediating microbial community dynamics and co-occurrence patterns along ecological succession remain unclear. Here, we explored community dynamics and taxa co-occurrence patterns in bacterial and fungal communities across a well-established chronosequence of post-mining lands spanning 54 years of recovery. Meanwhile, by synthesizing previous studies and ecological theories, we devised two conceptual models that integrate microbial phylogeny with patterns in community dynamics and in taxa co-occurrence. We further tested these models by using empirical data. At early successional stages, bacterial community structures became increasingly phylogenetically clustered with soil age, which was co-determined by the environmental selection from soil vegetation cover and by heterogeneous responses of less phylogenetically similar bacteria to the increasing resource availability along succession. At later successional stages, bacterial community phylogenetic structures displayed progressively lower variability. The fungal community phylogenetic structures varied relatively less and were independent of soil age, soil properties and vegetation cover, which was attributed to the dominance of stochastic processes in community turnover along succession. Network analysis revealed a decrease in bacterial co-occurrence complexity along succession, which aligned with a decrease in average pairwise phylogenetic distances between co-occurring bacteria. These patterns together implied a decrease in potential bacterial cooperation that was probably mediated by increasing resource availability along succession. The increased complexity of fungal co-occurrence along succession was independent of phylogenetic distances between co-occurring

45 fungi. This study provides new sights into ecological mechanisms underlying bacterial and
46 fungal community succession.
47
48 **Keywords**: soil bacteria and fungi, ecological succession, community assembly, phylogenetic
49 relatedness, co-occurrence, post-mining lands

1. Introduction

Elucidating the co-development of soil microbiomes and the ecosystem that hosts them is crucial for understanding ecosystem functioning, recovery and maintenance (Bardgett and van der Putten, 2014; Dini-Andreote et al., 2014; Bahram et al., 2018). Therefore, studying natural development of post-mining lands provides a unique opportunity to advance knowledge on the mechanisms underlying soil microbiome development along with *de novo* ecosystem recovery, as strip-mining activities often degrade existing ecosystems to an almost primordial stage.

Elucidating soil microbial community dynamics and microbial taxa co-occurrence patterns along ecological succession could provide important clues to understand successional patterns of microbial communities (Brown and Jumpponen, 2014; Dini-Andreote et al., 2014; Harantová et al., 2017; Morriën et al., 2017), whereas addressing either of them remains challenging (Additional file 1: Fig. S1). Besides environmental selection effects, the niche characteristics of the initial community could also affect community dynamics (i.e., the transition from an initial community to a later community), as community niche characteristics determine how the community respond to environmental selection/disturbance. For example, a few studies have revealed "legacy effects" based on the finding that initial microbial communities differing in niche characteristics, subsequently subjected to the identical environmental selection, finally transitioned to different community assemblages (Evans and Wallenstein, 2014; Banerjee et al., 2016a). However, it remains challenging to understand how the niche characteristics of the initial community would affect community dynamics (Additional file 1: Fig. S1a).

Due to the high complexity of soil microbiomes, direct evaluation of their niche characteristics is not technically feasible. Instead, community phylogenetic structures could

provide clues for community niche characteristics (Webb et al., 2002). As such, a phylogenetically clustered community (species in this community are more closely related than null model expectation (Stegen et al., 2012)) consists of species with similar niches; an overdispersed community (species in this community are more distantly related than null model expectation (Stegen et al., 2012)) consists of species with distinct niches (Webb et al., 2002). Here, to address the challenge in elucidating community dynamics (Additional file 1: Fig. S1a), we devised a conceptual model (Fig. 1a) to describe how the phylogenetic structure of the initial community will affect the later community structure, namely community dynamics. In our model (Fig. 1a), if the initial community is phylogenetically overdispersed (nearest taxon index (NTI) < -2 (Stegen et al., 2012)), this community is predicted to show heterogeneous responses to the environmental selection. In this case, the later community is expected to become more phylogenetically clustered than the initial community. For example, Tripathi et al. found that under pH selection, the phylogenetically overdispersed (or less phylogenetically clustered) community became more phylogenetically clustered (Tripathi et al., 2018). This happens because taxa performed heterogeneous responses that well-adapted taxa became predominant while maladapted taxa became rare or were excluded in the later community. These predominant taxa are probably closely related, as they show similar adaptations to pH selection (Webb et al., 2002), thus resulting in a more phylogenetic clustered community. In contrast, if the initial community is phylogenetically clustered (nearest taxon index (NTI) > 2 (Stegen et al., 2012)), this community is predicted to show a homogeneous response to the environmental selection (Fig. 1a), so that the phylogenetic structure of the later community is expected to keep relatively stable as the initial community. This happens because taxa perform heterogeneous

94 responses that taxa in the initial community are closely related and thus are presumed to develop
95 similarly in response to the environmental selection (Webb et al., 2002). Taken together, our
96 model (Fig. 1a) conceptualizes another factor (i.e., the phylogenetic structure of the initial
97 community) underlying community dynamics along ecological succession.

98 Although microbial taxa co-occurrence patterns are important for understanding microbial
99 community assembly along ecological succession (Dini-Andreote et al., 2014; Morriën et al.,
100 2017), ecologically explaining taxa co-occurrence patterns is challenging (Additional file 1:
101 Fig. S1b). For example, taxa co-presence (positive correlations between taxa in the network)
102 could imply two distinct ecological inferences: 1) potential cooperation for
103 survival/development across taxa with distinct niches, or 2) similar environmental preferences
104 across taxa with similar niches (Barberan et al., 2012; Deng et al., 2012). Taxa co-exclusion
105 (negative correlations between taxa) could also imply two distinct ecological inferences: 1)
106 potential competition across taxa with similar/overlapped niches, or 2) distinct environmental
107 preferences across taxa with distinct niches (Barberan et al., 2012; Deng et al., 2012)
108 (Additional file 1: Fig. S1b). Therefore, it is challenging to differentiate the ecological
109 inferences derived from taxa co-occurrence patterns and thus hinders the understanding of
110 ecological relationships among taxa (Additional file 1: Fig. S1b).

111 Remarkably, all ecological inferences derived from microbial co-occurrence patterns are
112 tightly linked to species niches, so the phylogenetic relatedness among co-occurring species
113 can help to differentiate the ecological inferences derived from a specific co-occurrence pattern,
114 based on the assumption of phylogenetic niche conservatism (Webb et al., 2002). Here, to
115 address the challenge in ecologically elucidating taxa co-occurrence patterns (Additional file 1:

Fig. S1b), we devised a conceptual model (Fig. 1b) that integrates phylogenetic relatedness between the co-occurring taxa with taxa co-occurrence patterns to differentiate ecological inferences from microbial co-occurrence. In the model (Fig. 1b), in scenario (1): taxa co-presence (positive correlations between taxa in the network) implies similar environmental preferences between the co-present taxa, if the co-present taxa are closely related (evaluated by the phylogenetic distance between the co-present taxa), because similar environmental preferences are usually found among closely related taxa (Webb et al., 2002); in scenario (2): taxa co-presence implies potential cooperation between the co-present taxa, if the co-present taxa are distantly related, because functionally complementary cooperation tends to establish between distantly related taxa (Morris et al., 2013; Zelezniak et al., 2015); in scenario (3): taxa co-exclusion (negative correlations between taxa in the network) implies potential competition between the co-excluding taxa, if the co-excluding taxa are closely related, because competition tends to occur between closely related taxa that occupy similar/overlapped niches (Hibbing et al., 2010); in scenario (4): taxa co-exclusion implies distinct environmental preferences between the co-excluding taxa, if the co-excluding taxa are distantly related, because distinct environmental preferences are usually found among distantly related taxa (Webb et al., 2002).

In this study, we focused on two major components of soil microbiomes (bacterial and fungal communities) and utilized phylogenetic information to disentangle the interplay of ecological processes underlying microbial community dynamics and co-occurrence patterns along ecological succession on post-mining lands. We addressed three main questions: (*i*) To what extent do the phylogenetic structures of bacterial and fungal communities change along succession? (*ii*) What are the relative influences of distinct ecological processes that mediate

bacterial and fungal community dynamics? (*iii*) How do bacterial and fungal co-occurrence patterns change along ecological succession, and why?

2. Materials and methods

2.1. Data collection

A well-established soil chronosequence from the lignite mining district near Sokolov in the Czech Republic (Frouz et al., 2001; Frouz and Nováková, 2005; Mudrak et al., 2016) was used to investigate the patterns of soil succession and recovery. The chronosequence and other background information in this area were provided by Sokolovská Uhelná mining company. The chronosequence was further validated using historical aerial photography and additional independent methods (Frouz, 2013). In this area, the mean annual temperature is 6.8 °C, the mean annual precipitation is 650 mm and the altitude is ~550 m a.s.l (Frouz et al., 2001). This chronosequence spans 54 years of ecosystem development from bare overburden of alkaline Miocene clay via herbs, grasses, goat willow shrubs to young forests dominated by birch and aspen (Mudrak et al., 2016; Harantová et al., 2017). A total of four distinct chronosequence sites (i.e., Sites I, II, III and IV) were sampled in triplicate at three time-points (in the end of May in 2006 (2008 only for the Site I), 2010 and 2015) (Additional file 1: Table S1) (see (Harantová et al., 2017) for additional details). The aboveground vegetation was surveyed as previously described (Harantová et al., 2017). Soil samples were subjected to total DNA extraction (Additional file 2: S1), quantification of microbial biomass, and characterization of physicochemical properties (i.e., total nitrogen, organic carbon and pH), as previously described (Sagova-Mareckova et al., 2008; Harantová et al., 2017). The primer sets 515F/806R

(Caporaso et al., 2012) and nu-SSU-0817/nu-SSU-1196 (Borneman and Hartin, 2000) were used to amplify the bacterial 16S rRNA gene and the fungal 18S rRNA gene, respectively, as previously described (Zifcakova et al., 2016; Navrátilová et al., 2019). PCR products were purified and sequenced on an Illumina MiSeq platform in the Laboratory of Environmental Microbiology, Institute of Microbiology of the Czech Academy of Sciences.

2.2. Data analysis

Amplicon sequences were sorted according to their unique barcodes and trimmed for quality (sequences containing ambiguous characters, with read length < 200bp, or with quality score < 15 were removed) (Edgar and Flyvbjerg, 2015), followed by chimera removal (Edgar et al., 2011). Sequences were clustered at 97% and 98.5% similarities, into operational taxonomic units (OTUs), for bacteria and fungi, respectively (Li and Godzik, 2006). Taxonomic assignments of bacterial sequences were performed using the Greengenes database (version 13_8) (DeSantis et al., 2006). Each fungal OTU was assigned to its closest genus using the Genbank database. Singletons, non-bacterial and non-fungal sequences were removed, and the OTU tables were rarefied to 7942 sequences per sample (bacteria) and 1431 sequences per sample (fungi).

Representative sequences of fungal and bacterial OTUs were aligned with PyNAST (Caporaso et al., 2010), referring to the SILVA alignment version 108 (https://www.arb-silva.de/download/archive/qiime/), and the Greengenes core set alignment (http://greengenes.lbl.gov/Download/Sequence_Data/Fasta_data_files/). The high-quality alignments of bacterial and fungal representative sequences were used to construct

phylogenetic trees with FastTree, respectively (Price et al., 2009). To evaluate the phylogenetic relatedness between OTUs, the phylogenetic distance of pairwise OTUs was determined using the function *cophenetic* in the R (v. 4.0.2) package 'stats' (v. 4.0.2) based on the phylogenetic tree. To evaluate the niche differences between OTUs, the niche distance of pairwise OTUs for all measured environmental variables (aboveground vegetation and soil properties) was determined as previously described (Stegen et al., 2012). To test the phylogenetic signal, the function *mantel.correlog* (permutations = 999) in the R package 'vegan' (v. 2.5-6) was used to measure the Pearson's correlation between OTU phylogenetic distances and OTU niche distances (Stegen et al., 2012). To evaluate the community phylogenetic structure, nearest taxon index (NTI) of microbial communities was calculated in the R package 'picante' (v.1.8.2) with the function *ses.mntd* (abundance.weighted = TRUE, null.model="taxa.labels", iterations = 1000) with 999 randomization across all samples. NTI value > +2 indicates phylogenetic clustering (species in a local community are more closely related than null model expectation), whereas NTI value < -2 indicates phylogenetic overdispersion (species in a local community are more distantly related than null model expectation) (Table 1) (Stegen et al., 2012; Stegen et al., 2013). Mantel test was used to evaluate correlations between Euclidean distances of NTI values and vegetation cover, using the R package 'vegan'. To estimate pairwise phylogenetic turnover between communities, Beta mean nearest taxon distance (βMNTD) was calculated using the function *comdistnt* (abundance.weighted = TRUE) in the R package 'picante'. β-nearest taxon index (βNTI) was used to estimate the degree to which observed βMNTD deviates from the mean of the null distribution βMNTD, normalized by its standard deviation. βNTI was calculated based the phylogenetic null model (Additional file 2: S2) (Stegen et al., 2013; Stegen

et al., 2015). βNTI value < -2 indicates that phylogenetic turnover is driven by homogeneous selection (significantly less turnover than null model expectation), whereas βNTI value > +2 indicates variable selection (significantly greater turnover than null model expectation) (Table 1) (Stegen et al., 2013; Stegen et al., 2015). |βNTI| < 2 indicates that phylogenetic turnover is not driven by deterministic processes (no significant differences between observed turnover and null model expectation) but by dispersal limitation, homogenizing dispersal, or undominated process (including ecological drift and other stochastic processes excluding dispersal) (Table 1) (Stegen et al., 2013; Stegen et al., 2015). In the scenario of |βNTI| < 2, to disentangle these ecological processes, the Bray-Curtis-based Raup-Crick metric (RC_{bray}) was calculated with the modified approach by Stegen et al. (Stegen et al., 2013) from Chase et al. (Chase et al., 2011). Taken together, the relative importance of ecological processes to the phylogenetic turnover of microbial communities was evaluated following the previous method (Table 1) (Stegen et al., 2013; Stegen et al., 2015). In particular, in the scenario of |βNTI| < 2, the percentages of RC_{bray} > +0.95, RC_{bray} <−0.95 and $|RC_{bray}|$ < 0.95 were used to quantify the relative importance of dispersal limitation, homogenizing dispersal and undominated process, respectively. The percentages of βNTI values > +2 and < -2 were used to quantify the relative importance of variable and homogeneous selection, respectively.

Based on these above analyses, we found large differences in vegetation cover, soil properties and microbial community turnover between the early successional stages (ES, including Site I and II) and the later successional stages (LS, including Site III and IV) (see the Result section). In this context, co-occurrence network analysis was performed separately on the data from ES and LS, to reveal microbial co-occurrence patterns along the succession. Two

network groups were constructed based on bacterial and fungal communities, respectively. OTUs with occurrence in more than three samples and average relative abundance > 0.1% were selected in early (158 OTUs) and later (168 OTUs) successional stages for bacteria, and in early (83 OTUs) and later (78 OTUs) successional stages for fungi. Spearman correlations between two taxa with adjusted *P*-value (Benjamini and Hochberg, 1995) < 0.01 and rho coefficient > 0.7 or < −0.7 were considered statistically robust to indicate positive and negative co-occurrences, respectively (Barberan et al., 2012). FDR correction (Benjamini and Hochberg, 1995) of *P*-values was conducted in R using the package 'vegan'. The taxa involved in the network were designated as co-occurring taxa. A total of 999 Erdős-Rényi random networks (Erdős and Rényi, 1960) with the same number of nodes and edges as each correspondingly observed network were generated using the function *erdos.renyi.game* in the R package 'igraph' (v. 1.2.6). The topological properties and visualizations of networks were performed using Gephi (https://gephi.github.io/). To evaluate the significance of differences in comparisons, Wilcoxon rank sum test and permutational multivariate analysis of variance (PERMANOVA) were conducted in R using the packages 'stats' and 'vegan'. The regression analysis with the method of "loess" (Cleveland and Grosse, 1991) was conducted in R using the packages "ggplot2" (v. 3.3.2).

3. Results

3.1. Changes in environmental variables along the chronosequence

Vegetation cover and soil properties significantly changed across the post-mining soil chronosequence, especially between the early successional stages (ES, including Site I and II)

and the later successional stages (LS, including Site III and IV), but no significant differences were found within ES (Site I *VS* Site II) or LS (Site III *VS* Site IV) (Additional file 1: Table S2 and S3). In ES, the increase of vegetation cover with soil age was accompanied by the significant increase in biomass of bacteria and fungi (Additional file 1: Table S4). In LS, the sites were gradually covered by trees. The soil organic carbon and total nitrogen were significantly higher (ca. two-fold change, $p < 0.01$, by Wilcoxon rank sum test) in LS than those in ES, which was likely attributed to the greater vegetation cover at these mature soil stages (Additional file 1: Table S2). Soil pH showed an opposite pattern, gradually decreasing with soil age from 7.4 to 6.7.

3.2. Phylogenetic structures and turnover of microbial communities across the chronosequence

All NTI values of bacterial communities across the chronosequence were greater than + 2 (Fig. 2), indicating that bacterial community structures were more phylogenetically clustered than null model expectation (Stegen et al., 2012). Bacterial NTI values increased with soil age ($p < 0.05$) in ES, thereafter showing less variability in LS (Fig. 2). Bacterial NTI values were significantly positively correlated with vegetation cover but not with soil properties (Additional file 1: Table S5 and S6). Interestingly, bacterial NTI values showed significant positive correlations with bacterial biomass (based on both qPCR and PLFA results) only in ES (Additional file 1: Table S6). NTI values of most fungal communities were greater than + 2 (Fig. 2), indicating that fungal community structures were phylogenetically clustered. Fungal NTI values did not significantly correlate with vegetation cover or fungal biomass ($p > 0.05$),

but were only significantly positively correlated with soil pH in LS (Additional file 1: Table S5 and S6).

Across short phylogenetic distances for both bacterial and fungal communities, significant positive correlations were found between OTU phylogenetic distances and OTU niche distances (Additional file 1: Fig. S2), indicating a significant phylogenetic signal. Thus, bacterial and fungal communities exhibited phylogenetic niche conservatism, supporting the use of phylogenetic information to derive ecological inferences (Losos, 2008; Stegen et al., 2013). With the combination of RC_{bray} and βNTI, we quantified the ecological processes governing the turnover of microbial community structures (Fig. 3), based on the criteria shown in Table 1. For the turnover of bacterial community structures, variable selection was stronger in ES (63.9 % at Site I and 69.4 % at Site II) than in LS (2.8 % at Site III and 38.9 % at Site IV), whereas homogeneous selection was stronger in LS (8.3 % at Site III and 30.6 % at Site IV) than in ES (0 % at Site I and 5.6 % at Site II). Across all successional stages (including the turnover between sites and within sites), the turnover of bacterial community structures was mainly driven by dispersal limitation and environmental selection (variable and homogeneous selection) (Fig. 3a). For the turnover of fungal community structures, variable selection was more important in LS (44.4 % at Site III and 52.8 % at Site IV) than in ES (25 % at Site I and 2.7 % at Site II). Across all successional stages, the turnover of fungal community structures was largely driven by dispersal limitation, undominated processes and environmental selection (Fig. 3b).

3.3. Microbial co-occurrence patterns along succession

292 Owing to the large differences in vegetation cover, soil properties and microbial community
293 turnover between ES and LS, we constructed networks separately based on data in ES and LS,
294 to reveal the differences in microbial co-occurrence patterns across successional stages. Among
295 the network topological properties, the observed "modularity" was greater than that in the
296 corresponding Erdős-Rényi random networks (Additional file 1: Table S7), suggesting that
297 observed networks had modular structures (Newman, 2006). Other topological properties such
298 as 'network diameter' and 'average path length' in each observed network were different from
299 those in the corresponding Erdős-Rényi random networks (Additional file 1: Table S7). These
300 implied that all the observed networks were distinguishable from random networks.

301 All four networks were dominated by co-presence (positive connections) (Fig. 4 and
302 Additional file 1: Table S7). In the bacterial networks, the nodes (OTUs) were mainly from the
303 phyla Proteobacteria, Actinobacteria, Gemmatimonadetes and Bacteroidetes, and node number
304 declined from ES to LS (Fig. 4). The node degree (number of connections to a node) that
305 represents the connectedness of the network significantly declined from ES to LS (Fig. 5a). The
306 network average degree (representing network complexity (Deng et al., 2012; Dini-Andreote
307 et al., 2014)) and the number of total connections also declined from ES to LS (Additional file
308 1: Fig. S3 and Table S7). Taken together, the signals of network average degree, node degree,
309 node number, and connection number indicated a decrease in the network complexity from ES
310 to LS. Notably, the decrease in the complexity was more obviously reflected in co-presence
311 than in co-exclusion (negative connections) (Fig. 5a and Additional file 1: Fig. S3 and Table
312 S7). However, in the fungal networks, an opposite pattern was observed, such that the network
313 complexity increased from ES to LS (Fig. 4 and 5a, and Additional file 1: Fig. S3 and Table

S7).

Microbial co-occurrence is influenced by microbial niches and thus is usually tightly related to phylogenetic relatedness between co-occurring taxa (Fig. 1b) (Losos, 2008; Stegen et al., 2012). Therefore, the relationship between microbial co-occurrence patterns and corresponding pairwise phylogenetic distances (between co-occurring taxa) was evaluated (Fig. 5b). Greater pairwise phylogenetic distances between two taxa correspond to greater phylogenetical difference between taxa. In bacterial networks, the average pairwise phylogenetic distances were greater ($p < 0.01$) in ES than in LS (Fig. 5b). The differences in average pairwise phylogenetic distances between ES and LS were greater and had higher significance in bacterial co-presence than co-exclusion. For fungal communities, the average pairwise phylogenetic distances were not significantly different ($p > 0.05$) between in ES and in LS, no matter in co-presence or co-exclusion (Fig. 5b). Collectively, the decreasing network complexity corresponded to the decreasing average pairwise phylogenetic distances in bacterial communities across the chronosequence, while there was no such corresponding relationship observed in fungal communities.

There were great variations in microbial co-occurrence patterns across the chronosequence, which was probably related to the turnover of these co-occurring microorganisms (that were involved in networks). Thus, we further evaluated different ecological processes determining the turnover of these co-occurring microorganisms. We found that dispersal limitation played a dominant role in the turnover of co-occurring bacterial communities, and dispersal limitation together with undominated process mainly governed the turnover of co-occurring fungal communities (Additional file 1: Fig. S4). Variable selection exerted a greater role in the

turnover of co-occurring bacterial communities in ES than in LS, whereas an opposite pattern was observed in fungi.

4. Discussion

4.1. Ecological processes driving microbial community succession

The extent of phylogenetic clustering of bacterial communities significantly increased with soil age in ES (Fig. 2), indicating the increasing convergence of bacterial taxa niches. Shifts in microbial community phylogenetic structures are usually related to environmental variations (Stegen et al., 2012; Brown and Jumpponen, 2014). Changes in environmental properties in ES (Additional file 1: Table S2) yielded high environmental heterogeneity, which was further validated by major contributions of variable selection to the turnover of bacterial community structures in ES (Fig. 3a), because variable selection could indicate environmental heterogeneity (Dini-Andreote et al., 2015; Stegen et al., 2015). The high environmental heterogeneity in ES was probably the cause of great dynamics of bacterial community phylogenetic structures. Additionally, significantly positive correlations between bacterial NTI values with vegetation cover in ES (Additional file 1: Table S5) further suggest that environmental selection determined phylogenetic structures of bacterial communities in ES.

Low soil nutrient availability has been reported to enhance phylogenetic clustering of bacterial communities (Feng et al., 2017), which is not consistent with the finding in this study. In ES, the gradually increasing vegetation cover from barren soil via grassland to forest (Additional file 1: Table S2) likely account for the increasing amounts of rhizodeposits and litter, thereby enhancing soil nutrient availability (Harantová et al., 2017). This notion is

supported by the increased biomass of bacteria and fungi in ES. Remarkably, the substantial increase of bacterial biomass was accompanied by the increase of bacterial community phylogenetic clustering in ES (Additional file 1: Table S6), which can be explained by taxa heterogeneous responses in our model (Fig. 1a). Specifically, the initial bacterial communities in ES were less phylogenetically clustered (Fig. 2), so the nutrient acquisition strategies among bacterial taxa were presumably less similar (Webb et al., 2002). With progressive increments in nutrient availability along the chronosequence, bacterial taxa performed heterogeneous responses that they substantially expanded their populations over time, with different nutrient acquisition strategies. This notion is supported by the fact that relative abundances of different bacterial genera increased at different rates with soil age in ES (Harantová et al., 2017). Therefore, the substantial but uneven increase of biomass across bacterial taxa probably resulted in subsequent predominance of some taxa and thereby increased phylogenetic clustering of bacterial communities in ES. In prior studies, the increasing phylogenetic clustering of microbial communities was also found to align with the increasing predominance of some taxa in the community subjected to environmental selection (Yan et al., 2016; Wang et al., 2017). This can be explained by that predominant taxa under the similar environmental selection are likely closely-related (Webb et al., 2002), thus resulting in a phylogenetically clustered community. This explanation is supported by evidences that the predominance of well-adapted taxa to pH selection made the community phylogenetically clustered (Tripathi et al., 2018). These examples further prove our inference that less phylogenetically clustered communities caused taxa heterogeneous responses to the increment in nutrient availability and thus co-determined the increasing phylogenetic clustering of bacterial communities with soil

age in ES, thereby supporting our conceptual model (Fig. 1a). As succession proceeded, environmental stability and the buffering capacity of soil progressively increased (Dini-Andreote et al., 2014), which probably contributed to the relative stability (less variability) of bacterial community phylogenetic structures in LS. Meanwhile, the significant correlations between bacterial NTI values and vegetation cover in LS indicated that high forest cover probably resulted in homogeneous selection and directly or indirectly contributed to the relative stability of bacterial community phylogenetic structures. Additionally, bacterial homogeneous responses indicated by highly phylogenetic clustering of the bacterial community probably (at least partially) contributed to the relative stability of bacterial community phylogenetic structures in LS.

In ES, the extent of phylogenetic clustering of fungal communities remained relatively stable (less variable) with soil age, despite the increase in fungal biomass. This is consistent with few changes in the relative abundances of most fungal genera across soil ages in ES (Harantová et al., 2017), thus resulting in the relatively low level of community phylogenetic turnover. Besides, the dynamics of fungal community phylogenetic structures in ES were also attributed to the dominance of stochastic processes (e.g., undominated process and dispersal limitation) in community phylogenetic turnover (Fig. 3b). Additionally, neither vegetation cover nor soil properties significantly correlated with fungal NTI values in ES, which further endorses the dominance of stochastic processes in the turnover of fungal community phylogenetic structures.

The relative importance of variable selection in determining fungal community dynamics increased from ES to LS (Fig. 3b), which is not in agreement with other studies where

deterministic processes progressively decreased (Tian et al., 2017) or remained unchanged (Brown and Jumpponen, 2015) along successional gradients on glacier forefields. This disagreement was likely due to distinct initial conditions and environmental backgrounds of succession. Soil pH significantly correlated with fungal NTI values in LS. Therefore, the stronger variable selection on fungal community dynamics in LS was probably attributed to soil pH as well as the increased vegetation cover, which is known to at least partially influence fungal communities (Urbanova et al., 2015). Despite the stronger variable selection in LS, fungal community phylogenetic structures did not significantly change with soil age (Fig. 2). This might be explained by important roles of undominated process and dispersal limitation in determining fungal community dynamics (Fig. 3b). Interestingly, in comparison to the ecological processes that determined bacterial community dynamics, we found higher proportions of undominated process driving fungal community dynamics within each site as well as across all sites (Fig. 3). This difference probably explains the distinct dynamics of bacterial and fungal community phylogenetic structures along succession. Dispersal limitation refers to the restriction of taxa movement to and/or establishment in a new location (Martiny et al., 2011; Hanson et al., 2012). In this study, the sampling within a site was conducted in the same geocoordinate with different years (Additional file 1: Table S1). In this context, we assumed that the effect of geographic distance within a site can be neglected, so dispersal limitation (in this case) within each site (Fig. 3) may rather be attributed to the restriction of the continuous establishment of taxa across the chronosequence. Although here we categorized dispersal limitation into stochastic processes, dispersal limitation could be partly caused by deterministic factors (e.g., habitat features that probably changed across the chronosequence in

424 this study) (Hanson et al., 2012). Therefore, the relatively higher proportional influence of
425 dispersal limitation on the dynamics of the bacterial versus fungal communities across all sites
426 (Fig. 3) suggested that bacterial communities were more responsive to the ecological succession
427 than fungal communities.
428
429 **4.2. Mechanisms underlying microbial co-occurrence patterns along succession**
430 Bacterial co-occurrence exhibited decreasing complexity along succession, while the
431 complexity of fungal co-occurrence showed the opposite pattern (Fig. 4 and 5a, and Additional
432 file 1: Fig. S3). A similar development of bacterial co-occurrence was reported in a salt marsh
433 chronosequence, where the higher co-occurrence complexity at initial successional stages was
434 attributed to the higher temporal turnover of bacterial communities (Dini-Andreote et al., 2014).
435 Analogously, in this study, high variability of bacterial community phylogenetic structures (Fig.
436 2) and high community temporal turnover (Additional file 1: Fig. S5) probably underlay the
437 high complexity of bacterial co-occurrence in ES. Moreover, the importance of variable
438 selection in the turnover of bacterial co-occurring communities and whole communities (Fig. 3
439 and Additional file 1: Fig. S4) decreased together with bacterial co-occurrence complexity
440 along succession. Because variable selection probably indicates environmental heterogeneity
441 (Dini-Andreote et al., 2015; Stegen et al., 2015), we speculate that the greater environmental
442 heterogeneity contributed to the higher complexity of bacterial co-occurrence in ES.
443 When linking bacterial co-occurrence patterns to their phylogeny, we found that the
444 decreasing complexity of bacterial co-presence from ES to LS aligned with the general increase
445 of phylogenetic clustering of bacterial communities (Fig. 2), and remarkably with the decrease

of average pairwise phylogenetic distances between co-present taxa from ES to LS (Fig. 5b). That is, bacterial co-presence complexity decreased as the co-present taxa became more phylogenetically similar along succession. Based on our model (Fig. 1b), in scenario (1) where taxa co-presence bases on the similar environmental preference between the taxa: if the decrease of bacterial co-presence complexity along succession was attributed to the decreased similarity in environmental preference between the co-present bacterial taxa, the corresponding average pairwise phylogenetic distances between co-present taxa were expected to increase. This would occur because similar environmental preference favors phylogenetically similar (closely related) taxa (Webb et al., 2002; Tripathi et al., 2018). In scenario (2) where taxa co-presence bases on potential cooperation among the taxa: if the decrease of bacterial co-presence complexity along succession was attributed to the decreasing potential cooperation between distantly related bacteria, the corresponding average pairwise phylogenetic distances between co-present taxa were expected to decrease. This would occur because cooperation tends to be established based on metabolic dependencies between distantly related species whose niches are less overlapped (Morris et al., 2013; Zelezniak et al., 2015). In light of these concepts, only the scenario 2 was in agreement with our results, thus indicating that bacterial co-presence was likely an indication of potential cooperation across taxa. Interestingly, this inference, in turn, suggests the tight relationship between potential bacterial interactions and their corresponding phylogenetic relatedness in the natural environment. Most microbial interactions, regardless of cooperation or competition, are mostly established by nutrient demand (Hibbing et al., 2010; Morris et al., 2013). Nutrient addition was reported to substantially alter potential microbial interactions (Banerjee et al., 2016b), and high bacterial co-occurrence complexity has been

468 observed in barren soils (Dini-Andreote et al., 2014; Feng et al., 2017). Low nutrient availability
469 has been suggested to push the exchange of metabolites and nutrients between species for
470 survival (Zelezniak et al., 2015; Morriën et al., 2017). Therefore, in this study, low nutrient
471 availability in ES probably strengthened the potential cooperation between functionally distinct
472 bacteria, which resulted in the relatively higher complexity of bacterial co-presence.

473 Bacterial co-exclusion can be attributed to potential competition or distinct environmental
474 preferences between co-excluding taxa (Fig. 1b; scenarios 3 and 4). If the decreasing
475 complexity of bacterial co-exclusion from ES to LS (Additional file 1: Fig. S3 and Table S7)
476 indicated the decreasing potential in bacterial competition, the corresponding average pairwise
477 phylogenetic distances between co-excluding taxa were expected to increase, because
478 competition is more prone to occur among phylogenetically similar taxa that occupy a similar
479 niche (Webb et al., 2002; Violle et al., 2011; Stegen et al., 2012). However, this scenario
480 conflicted with our observations (Fig. 5b). If the decreasing complexity of bacterial co-
481 exclusion from ES to LS indicated the decreasing differences in environmental preferences
482 between co-excluding taxa, the corresponding average pairwise phylogenetic distances
483 between co-excluding taxa were expected to decrease, which coincided with our observations
484 (Fig. 5b). This inference was further supported by the increase of phylogenetic clustering of
485 bacterial communities from ES to LS (Fig. 2).

486 In contrast to bacterial community development, high turnover rates of fungal
487 communities (Additional file 1: Fig. S5) did not result in high network complexity in ES.
488 However, the higher environmental heterogeneity inferred by the higher proportional influence
489 of variable selection on fungal community turnover (Fig. 3b), coincided with the higher

490 complexity of fungal co-occurrence in LS than in ES (Fig.4 and 5a and Additional file 1: Fig.
491 S2). This indicates that the higher environmental heterogeneity may contribute to the higher
492 complexity for fungal co-occurrence in LS. The higher environmental heterogeneity of fungal
493 communities might be due to the increased vegetation cover in LS, as discussed in section 4.1.
494 Interestingly, although fungal co-occurrence complexity increased along succession, the
495 corresponding average pairwise phylogenetic distances between taxa in co-occurrence were not
496 significantly different between in ES and in LS (Fig. 5). Thus, we speculate that fungal co-
497 occurrence patterns along succession were unlikely related to their phylogeny. The less
498 difference in average phylogenetic distances of co-occurring fungi between in ES and in LS
499 was likely a result of the highly similar phylogeny of regional species, which was reflected in
500 the flat dynamics of fungal phylogenetic structures along succession.
501
502 **5. Conclusions and implications**
503 This study shows that bacterial community phylogenetic structures were more responsive than
504 fungal community phylogenetic structures to environmental gradients along the soil primary
505 succession, which was attributed to stronger deterministic effects on bacterial than fungal
506 community phylogenetic turnover. Significant changes in bacterial phylogenetic structures only
507 occurred at early successional stages, and aligned with substantial increase of bacterial biomass.
508 This implies that bacterial taxa niches and bacteria-dependent functions in the ecosystem
509 changed mainly at early successional stages, as the expansion of communities. Both bacterial
510 and fungal co-occurrence patterns significantly varied along succession, but only the former
511 aligned with phylogenetic relatedness between co-occurring taxa, thereby implying potential

512 bacterial cooperation based on our conceptual models. Thus, bacterial co-occurrence patterns
513 along soil primary succession were phylogeny-associated. Taken together, our results boost
514 understanding of ecological processes underlying microbial community development along soil
515 primary succession. Our conceptual models help to address two key fundamental challenges in
516 microbial community assembly and have broad application, as the two challenges exist not only
517 in soil ecosystems but also in other various ecosystems.
518

519 **Acknowledgements**

520 We thank Sokolovská Uhelná mining company for providing research permits and background
521 data about the sites.
522

523 **Author's contributions**

524 QL analyzed the data and wrote the manuscript. QL, FDA, TBM, RA, PH, LJL and LM revised
525 the manuscript. LM performed experimental works. JF and PB conceived the study and revised
526 the manuscript. All authors read the manuscript and approved the final draft.
527

528 **Funding**

529 This study was supported by grant of Ministry of Education Youth and Sport of the Czech
530 Republic LM2015075 and EF16_013/0001782.
531

532 **Data Accessibility**

533 The original sequencing data are available at public database (http://metagenomics.anl.gov/)

with dataset number 4741652.3 for bacteria and 4827823.3 for fungi.

Ethics approval and consent to participate

Not applicable

Competing interest

Authors declare that they have no competing interests.

Appendices

Supplementary information is available at ##.

References

Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., Huerta-Cepas, J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M., Polme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., Bork, P. (2018). Structure and function of the global topsoil microbiome. Nature, 560(7717), 233-237.

Banerjee, S., Helgason, B., Wang, L., Winsley, T., Ferrari, B.C., Siciliano, S.D. (2016a). Legacy effects of soil moisture on microbial community structure and N2O emissions. Soil Biology and Biochemistry, 95(40-50.

Banerjee, S., Kirkby, C.A., Schmutter, D., Bissett, A., Kirkegaard, J.A., Richardson, A.E. (2016b). Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. Soil Biology and Biochemistry, 97(188-198.

Barberan, A., Bates, S.T., Casamayor, E.O., Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. The ISME Journal, 6(2), 343-351.

Bardgett, R.D., van der Putten, W.H. (2014). Belowground biodiversity and ecosystem functioning. Nature, 515(7528), 505-511.

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B, 57(1), 289-300.

565 Borneman, J., Hartin, R.J. (2000). PCR primers that amplify fungal rRNA genes from environmental
566 samples. Applied Environmental Microbiology, 66(10), 4356-4360.
567 Brown, S.P., Jumpponen, A. (2014). Contrasting primary successional trajectories of fungi and bacteria
568 in retreating glacier soils. Molecular Ecology, 23(2), 481-497.
569 Brown, S.P., Jumpponen, A. (2015). Phylogenetic diversity analyses reveal disparity between fungal and
570 bacterial communities during microbial primary succession. Soil Biology and Biochemistry,
571 89(52-60.
572 Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., Knight, R. (2010). PyNAST: a
573 flexible tool for aligning sequences to a template alignment. Bioinformatics, 26(2), 266-267.
574 Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J.,
575 Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R. (2012). Ultra-high-
576 throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The
577 ISME journal, 6(8), 1621-1624.
578 Chase, J.M., Kraft, N.J., Smith, K.G., Vellend, M., Inouye, B.D. (2011). Using null models to disentangle
579 variation in community dissimilarity from variation in α - diversity. Ecosphere, 2(2), 1-11.
580 Cleveland, W.S., Grosse, E. (1991). Computational methods for local regression. Statistics and computing,
581 1(1), 47-62.
582 Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., Zhou, J. (2012). Molecular ecological network analyses. BMC
583 Bioinformatics, 13(1), 113.
584 DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P.,
585 Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and
586 workbench compatible with ARB. Applied Environmental Microbiology, 72(7), 5069-5072.
587 Dini-Andreote, F., de Cassia Pereira e Silva, M., Triado-Margarit, X., Casamayor, E.O., van Elsas, J.D., Salles,
588 J.F. (2014). Dynamics of bacterial community succession in a salt marsh chronosequence:
589 evidences for temporal niche partitioning. The ISME Journal, 8(10), 1989-2001.
590 Dini-Andreote, F., Stegen, J.C., van Elsas, J.D., Salles, J.F. (2015). Disentangling mechanisms that mediate
591 the balance between stochastic and deterministic processes in microbial succession.
592 Proceedings of the National Academy of Sciences of the United States of America, 112(11),
593 E1326-E1332.
594 Edgar, R.C., Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation
595 sequencing reads. Bioinformatics, 31(21), 3476-3482.
596 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. (2011). UCHIME improves sensitivity and
597 speed of chimera detection. Bioinformatics, 27(16), 2194-2200.
598 Erdős, P., Rényi, A. (1960). On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci, 5(1),
599 17-60.
600 Evans, S.E., Wallenstein, M.D. (2014). Climate change alters ecological strategies of soil bacteria. Ecol
601 Lett, 17(2), 155-164.
602 Feng, Y.Z., Guo, Z.Y., Zhong, L.H., Zhao, F., Zhang, J.B., Lin, X.G. (2017). Balanced fertilization decreases
603 environmental filtering on soil bacterial community assemblage in north China. Front microbiol,
604 8(12.
605 Frouz, J., 2013. Soil biota and ecosystem development in post mining sites. CRC Press.
606 Frouz, J., Keplin, B., Pižl, V., Tajovský, K., Starý, J., Lukešová, A., Nováková, A., Balík, V.r., Háněl, L., Materna,
607 J. (2001). Soil biota and upper soil layer development in two contrasting post-mining
608 chronosequences. Ecological Engineering, 17(2-3), 275-284.

609 Frouz, J., Nováková, A. (2005). Development of soil microbial properties in topsoil layer during
610 spontaneous succession in heaps after brown coal mining in relation to humus microstructure
611 development. Geoderma, 129(1-2), 54-64.
612 Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., Martiny, J.B. (2012). Beyond biogeographic patterns:
613 processes shaping the microbial landscape. Nature Reviews Microbiology, 10(7), 497.
614 Harantová, L., Mudrák, O., Kohout, P., Elhottová, D., Frouz, J., Baldrian, P. (2017). Development of
615 microbial community during primary succession in areas degraded by mining activities. Land
616 Degradation & Development, 28(8), 2574-2584.
617 Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B. (2010). Bacterial competition: surviving and
618 thriving in the microbial jungle. Nature Reviews. Microbiology, 8(1), 15.
619 Li, W., Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or
620 nucleotide sequences. Bioinformatics, 22(13), 1658-1659.
621 Losos, J.B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the relationship between
622 phylogenetic relatedness and ecological similarity among species. Ecology letters, 11(10), 995-
623 1003.
624 Martiny, J.B., Eisen, J.A., Penn, K., Allison, S.D., Horner-Devine, M.C. (2011). Drivers of bacterial β-
625 diversity depend on spatial scale. Proceedings of the National Academy of Sciences, 108(19),
626 7850-7854.
627 Morriën, E., Hannula, S.E., Snoek, L.B., Helmsing, N.R., Zweers, H., De Hollander, M., Soto, R.L., Bouffaud,
628 M.-L., Buée, M., Dimmers, W. (2017). Soil networks become more connected and take up more
629 carbon as nature restoration progresses. Nature Communications, 8(14349.
630 Morris, B.E., Henneberger, R., Huber, H., Moissl-Eichinger, C. (2013). Microbial syntrophy: interaction for
631 the common good. FEMS Microbiology Reviews, 37(3), 384-406.
632 Mudrak, O., Dolezal, J., Frouz, J. (2016). Initial species composition predicts the progress in the
633 spontaneous succession on post-mining sites. Ecological Engineering, 95(665-670.
634 Navrátilová, D., Tláskalová, P., Kohout, P., Dřevojan, P., Fajmon, K., Chytrý, M., Baldrian, P. (2019).
635 Diversity of fungi and bacteria in species-rich grasslands increases with plant diversity in shoots
636 but not in roots and soil. Fems Microbiology Ecology, 95(1), fiy208-fiy208.
637 Newman, M.E. (2006). Modularity and community structure in networks. P Natl Acad Sic USA, 103(23),
638 8577-8582.
639 Price, M.N., Dehal, P.S., Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with
640 profiles instead of a distance matrix. Molecular Biology and Evolution, 26(7), 1641-1650.
641 Sagova-Mareckova, M., Cermak, L., Novotna, J., Plhackova, K., Forstova, J., Kopecky, J. (2008). Innovative
642 methods for soil DNA purification tested in soils with widely differing characteristics. Applied
643 and Environmental Microbiology, 74(9), 2902-2907.
644 Stegen, J.C., Lin, X., Fredrickson, J.K., Chen, X., Kennedy, D.W., Murray, C.J., Rockhold, M.L., Konopka, A.
645 (2013). Quantifying community assembly processes and identifying features that impose them.
646 The ISME journal, 7(11), 2069.
647 Stegen, J.C., Lin, X., Fredrickson, J.K., Konopka, A.E. (2015). Estimating and mapping ecological processes
648 influencing microbial community assembly. Frontiers in microbiology, 6(370.
649 Stegen, J.C., Lin, X., Konopka, A.E., Fredrickson, J.K. (2012). Stochastic and deterministic assembly
650 processes in subsurface microbial communities. ISME Journal, 6(9), 1653.
651 Tian, J.Q., Qiao, Y.C., Wu, B., Chen, H., Li, W., Jiang, N., Zhang, X.L., Liu, X.Z. (2017). Ecological succession
652 pattern of fungal community in soil along a retreating glacier. Frontiers in microbiology, 8(13.

653 Tripathi, B.M., Stegen, J.C., Kim, M., Dong, K., Adams, J.M., Lee, Y.K. (2018). Soil pH mediates the balance
654 between stochastic and deterministic assembly of bacteria. ISME Journal, 12(4), 1072-1083.
655 Urbanova, M., Snajdr, J., Baldrian, P. (2015). Composition of fungal and bacterial communities in forest
656 litter and soil is largely determined by dominant trees. Soil Biology and Biochemistry, 84(53-64.
657 Violle, C., Nemergut, D.R., Pu, Z., Jiang, L. (2011). Phylogenetic limiting similarity and competitive
658 exclusion. Ecology letters, 14(8), 782-787.
659 Wang, Y., Li, C., Kou, Y., Wang, J., Tu, B., Li, H., Li, X., Wang, C., Yao, M. (2017). Soil pH is a major driver
660 of soil diazotrophic community assembly in Qinghai-Tibet alpine meadows. Soil Biology and
661 Biochemistry, 115(547-555.
662 Webb, C.O., Ackerly, D.D., McPeek, M.A., Donoghue, M.J. (2002). Phylogenies and community ecology.
663 Annu Rev Ecol Syst, 33(1), 475-505.
664 Yan, Q., Li, J., Yu, Y., Wang, J., He, Z., Van Nostrand, J.D., Kempher, M.L., Wu, L., Wang, Y., Liao, L., Li, X.,
665 Wu, S., Ni, J., Wang, C., Zhou, J. (2016). Environmental filtering decreases with fish
666 development for the assembly of gut microbiota. Environmental Microbiology, 18(12), 4739-
667 4754.
668 Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D.R., Bork, P., Patil, K.R. (2015). Metabolic
669 dependencies drive species co-occurrence in diverse microbial communities. P Natl Acad Sic
670 USA, 112(20), 6449-6454.
671 Zifcakova, L., Vetrovsky, T., Howe, A., Baldrian, P. (2016). Microbial activity in forest soil reflects the
672 changes in ecosystem properties between summer and winter. Environmental Microbiology,
673 18(1), 288-301.

674

675

676

677

678

679

680

681

682

683

684

685

686 **Table 1.** Criteria and definitions for patterns and ecological processes underlying
687 community phylogenetic structures and turnover, respectively (Webb et al., 2002;
688 Stegen et al., 2012; Stegen et al., 2013; Stegen et al., 2015).

Phylogenetic structure for a local (single) community	**Criteria**	**Assembly patterns**	**Definitions**
	NTI > +2	Phylogenetic clustering	Species in a local community are more closely related than null model expectation.
	|NTI| < 2	Stochastic processes	Phylogenetic relatedness among species in a local community do not differ significantly from null model expectation.
	NTI < -2	Phylogenetic overdispersion	Species in a local community are more distantly related than null model expectation.
Phylogenetic turnover between regional (pairwise) communities	**Criteria**	**Ecological processes**	**Definitions**
	βNTI > +2	Variable selection	Caused by heterogeneous environmental conditions, deterministically leading regional communities toward more dissimilarity.
	|βNTI| < 2 and RC_{bray} > +0.95	Dispersal limitation	Caused by the restriction of taxa movement to or/and establishment in a new location, usually acting with ecological drift to result in more dissimilarity between regional communities.
	|βNTI| < 2 and |RC_{bray}| < 0.95	Undominated	Caused by ecological drift (e.g. random changes in species relative abundance) or other stochastic processes, signifying a scenario that neither selection nor dispersal primarily determine the community turnover.
	|βNTI| < 2 and RC_{bray} < –0.95	Homogenizing dispersal	Caused by very high dispersal rates, overwhelming other processes and leading toward very few variations between regional communities.
	βNTI < -2	Homogeneous selection	Caused by homogeneous environmental conditions, deterministically leading regional communities toward more similarity.

689
690
691
692
693
694

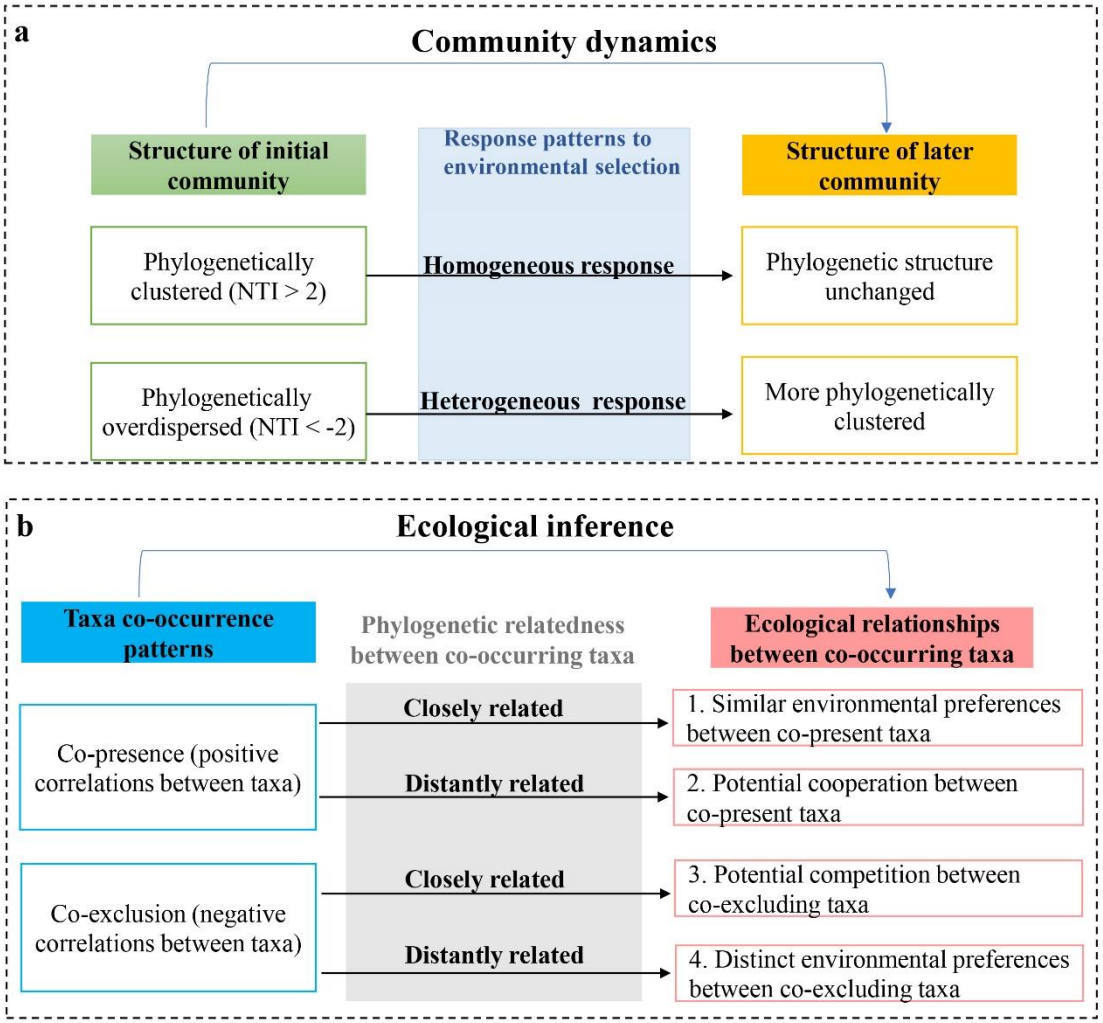

Fig. 1 Conceptual models in elucidating microbial community dynamics (a) and taxa co-occurrence (b) separately. The model (a) uses the phylogenetic structure of the initial community to predict the structure of the later community, namely community dynamics. If the initial community is phylogenetically overdispersed (nearest taxon index (NTI) < -2), this initial community is predicted to show heterogeneous responses to the environmental selection, so that the later community is expected to become more phylogenetically clustered than the initial community. If the initial community is phylogenetically clustered (NTI > 2), this initial community is predicted to show a homogeneous response to the environmental selection, so that the phylogenetic structure of the later community is expected to keep relatively stable as the initial community. The model (b) integrates taxa co-occurrence patterns and phylogenetic relatedness between the co-occurring taxa to differentiate ecological inferences from taxa co-occurrence. Taxa co-presence and co-exclusion are separately based on positive and negative correlations between taxa in a network. Taxa phylogenetic relatedness is evaluated by phylogenetic distance between co-occurring taxa.

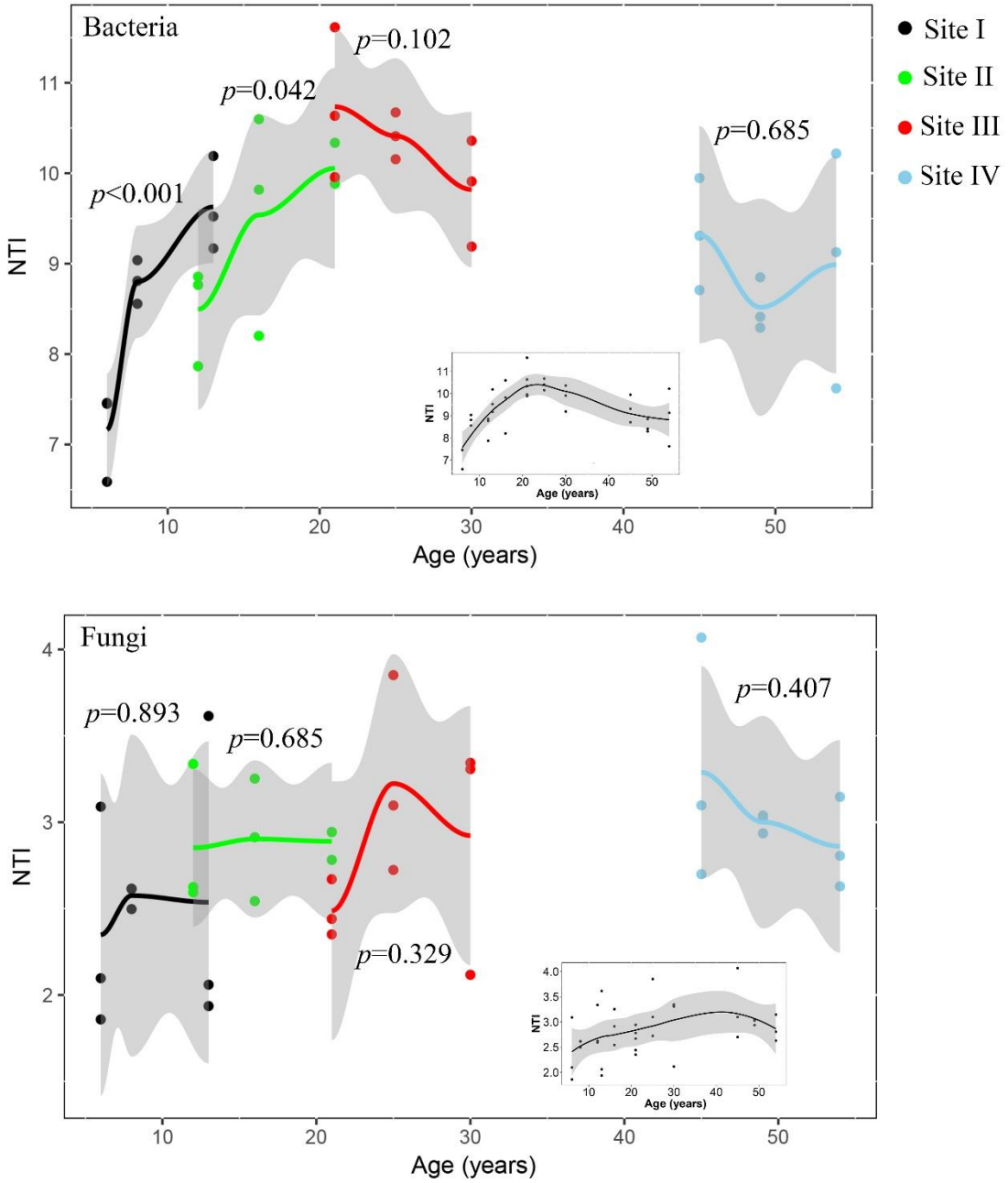

Fig. 2 Distribution of NTI values at different sites across the chronosequence, and all sites (embedded). *P*-values of spearman's correlations between ages and NTI values at different sites are shown. The method of the regression is "loess". Grey shadow represents 95% confidence intervals.

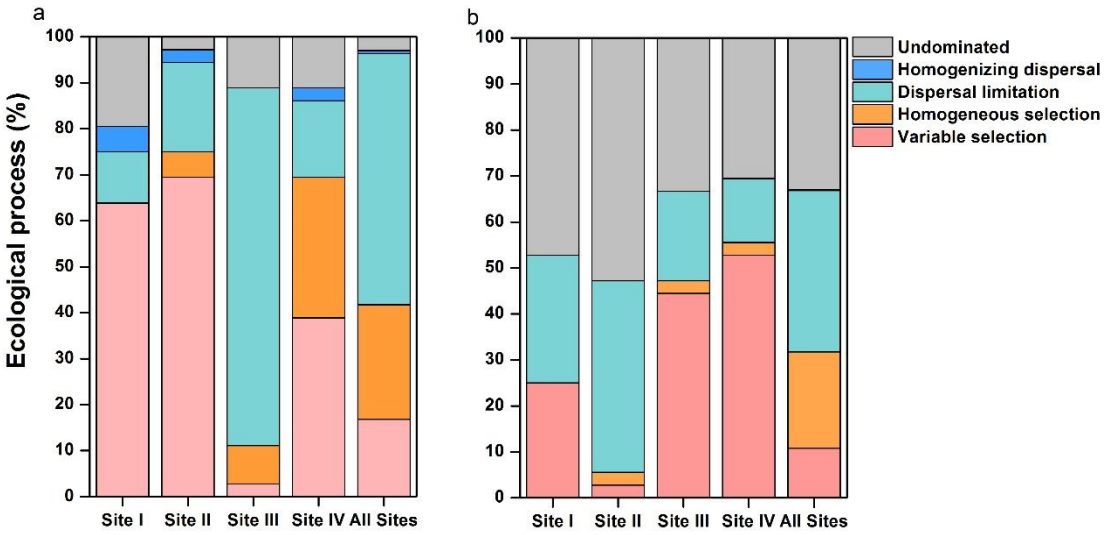

Fig. 3 Relative contributions of different ecological processes to bacterial (a) and fungal (b) community turnover at different sites across the soil chronosequence. All Sites: including the community turnover between sites and within sites.

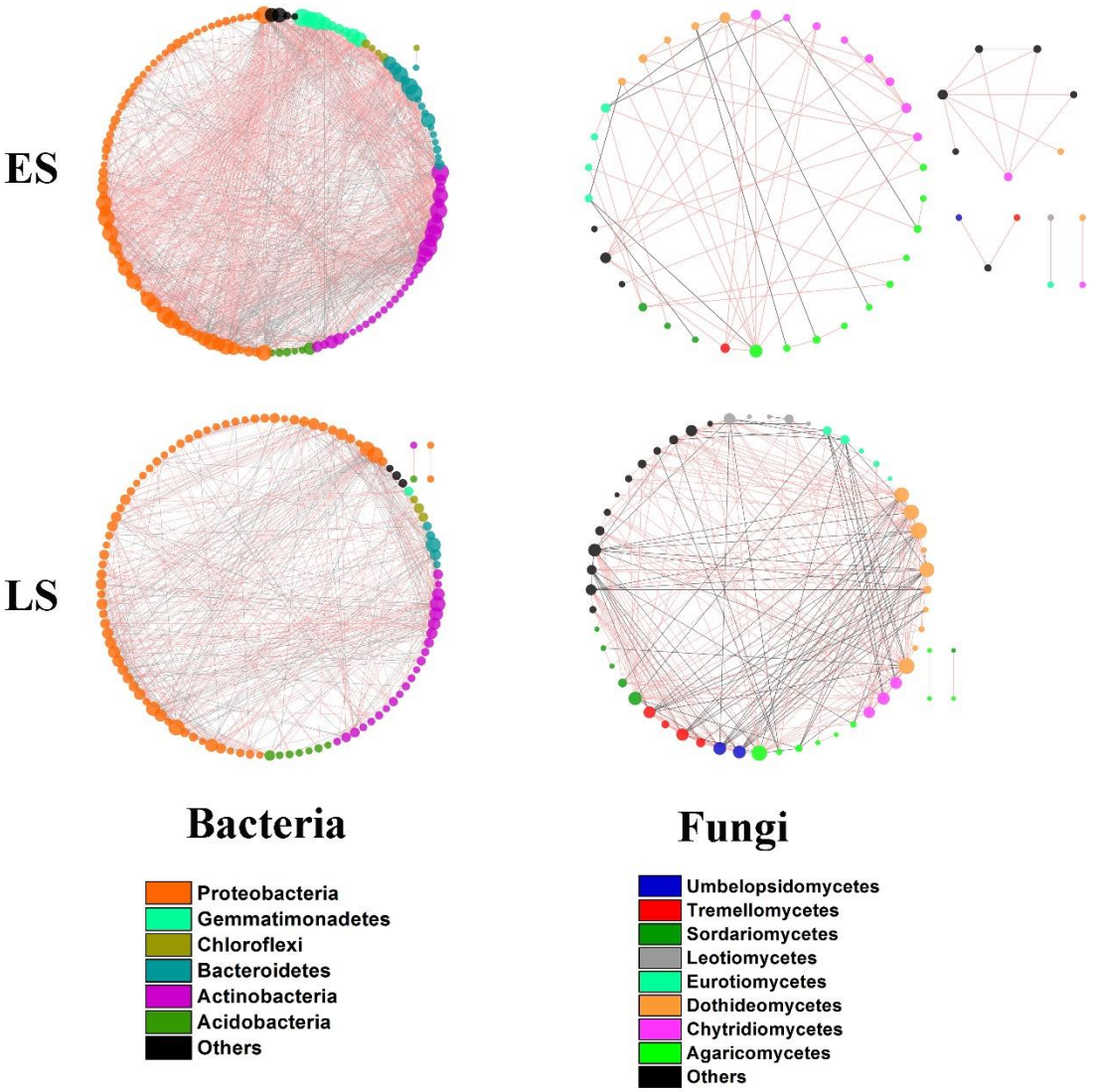

Fig. 4 Co-occurrence networks of bacterial and fungal taxa, respectively. The nodes are sized by node degree (number of connections to a node), and colored by bacterial phyla and fungal classes, respectively. Positive and negative connections are colored by red and grey, respectively. ES: early successional stages including Site I and II; LS: later successional stages including Site III and IV.

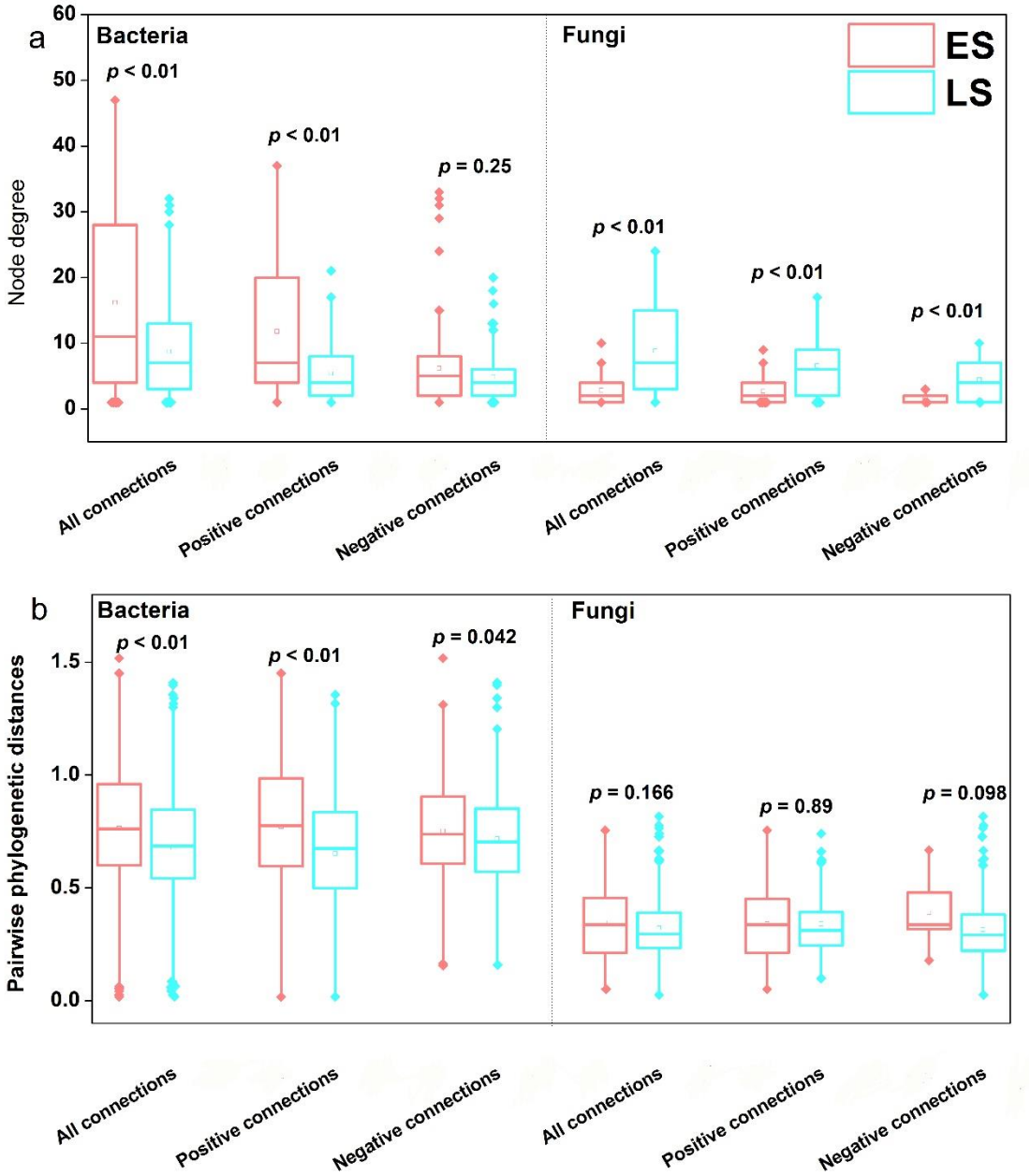

Fig. 5 Node degree (a) and pairwise phylogenetic distances between connecting nodes (b) in bacterial and fungal networks. All connections: the network consisting of all connections (edges); positive connections: the sub-network consisting of only positive connections; negative connections: the sub-network consisting of only negative connections. The *P*-value is shown for the statistical significance in each pairwise comparison based on Wilcoxon rank sum test. The square and line inside the box represent the mean and median, respectively. ES: early successional stages including Site I and II; LS: later successional stages including Site III and IV.