Distinct co-occurrence patterns and driving forces of rare and abundant bacterial subcommunities following a glacial retreat in the eastern Tibetan Plateau

Yonglei Jiang¹ · Halfeng Song² · Yanbao Lei² · Helena Korpelainen³ · Chunyang Li¹

Abstract
Unraveling the dynamics and driving forces of abundant and rare bacteria in response to glacial retreat is essential for a deep understanding of their ecological and evolutionary processes. Here, we used Illumina sequencing datasets to investigate ecological abundance, successional dynamics, and the co-occurrence patterns of abundant and rare bacteria associated with different stages of soil development in the Hailuogou Glacier Chronosequence. Abundant taxa exhibited ubiquitous distribution and tight clustering, while rare taxa showed uneven distribution and loose clustering along the successional stages. Both abundant and rare subcommunities were driven by different factors during assembly: the interactions of biotic and edaphic factors were the main driving forces, although less important for rare taxa than for the abundant ones. In particular, the redundancy analysis and structural equation modeling showed that soil organic C, pH, and plant richness primarily affected abundant subcommunities, while soil N and pH were most influential for rare subcommunities. More importantly, variation partitioning showed that edaphic factors exhibited a slightly greater influence on both abundant (7.8%) and rare (4.5%) subcommunities compared to biotic factors. Both abundant and rare bacteria exhibited a more compact network topology at the middle than at the other chronosequence stages. The overlapping nodes mainly belonged to Proteobacteria and Acidobacteria in abundant taxa and Planctomycetia, Sphingobacteria, and Phycisphaerae in rare taxa. In addition, the network analysis showed that the abundant taxa exhibited closer relationships and more influence on other co-occurrences in the community when compared to rare taxa. These findings collectively reveal divergent co-occurrence patterns and driving forces for abundant and rare subcommunities along a glacier forefield chronosequence in the eastern Tibetan Plateau.

Keywords Rare and abundant bacteria · Driving forces · Network analysis · Microbial co-occurrence patterns · Hailuogou Glacier Chronosequence

Introduction
Glacial retreat caused by global warming can have profound and multiple impacts on ecosystems, including both above- and belowground community compositions (Miniaci et al. 2007; Schmidt et al. 2016; Jiang et al. 2018). Soil microorganisms are unique components of belowground ecosystems (Wang et al. 2016) and key players of biogeochemical processes (e.g., C, N, and P cycles) and ecological functions of soil (de Sosa et al. 2018; Schloter et al. 2018). Thus, changes in their spatial composition and distribution patterns in response to environmental fluctuations may directly influence ecosystem functioning. Along with the advancement of molecular tools, research on the ecological and evolutionary processes of microorganisms across space and time has received
increasing attention, especially in some harsh and pristine ecosystems, including tundra and glacier retreat areas (Bradley et al. 2016; Kazemi et al. 2016). For example, Zumsteg et al. (2012) claimed that abundant bacterial taxa are influenced by soil moisture, pH, and ion concentration in the Damma glacial foreland. Knebelman et al. (2012) found that many microbes are active (or potentially active) in recently deglaciated soils and able to quickly utilize available nutrients or aerially deposited carbon-rich substrates (e.g., pollen). These earlier studies mainly focused on the highly inclusive groups (kingdom or domain) and on the effects of environmental factors on soil community assemblages, while they largely neglected the lower and less inclusive groups and the interactions among microbial species. These groups are referred to as “rare biosphere” and have a lower abundance and extremely high diversity. Yan et al. (2017) suggest that to understand microbial successional dynamics, it is insufficient to analyze only the kingdom and/or domain groups. Both broad (kingdoms or domains) and less inclusive groups (e.g., rare taxa) should be investigated to better dissect community dynamics (Yan et al. 2017). Recent studies based on network analyses have proven the important roles of abundant and rare microbial interactions in determining community compositions (Wang et al. 2018; Xue et al. 2018). However, there is still limited knowledge about the interactions and assembly mechanisms of microbial communities in developing soils (<200 years), in particular after glacier retreat.

Microbial communities are comprised of a few highly abundant species and a large number of rare species, and the latter subcommunity is still largely unexplored (Mo et al. 2018). In most ecosystems, abundant microorganisms contribute to the bacterial biomass production and to the fluxes of organic matter (Jiao et al. 2017a; Liao et al. 2017), while the numerous rare species that are represented by only a few individuals provide a tremendous reservoir of genetic and functional diversity, which can enhance the functionality of abundant microbes and mediate ecosystem stability and function (Pedrós-Alió 2011). It is likely that abundant and rare community assemblages are subjected to different controlling factors. However, owing to previous methodological limitations, it has been difficult to directly identify both rare and abundant microbes (Logares et al. 2014) and explore microbial diversity in these subcommunities (Kimes et al. 2013; Wang et al. 2016). Moreover, it has been especially challenging to explore the dynamics of rare microbial communities in response to environmental changes; they could occupy multiple niches on a very fine scale. Currently, the rapid development of high-throughput sequencing technologies has substantially boosted our capacity to identify different types of microbes (Logares et al. 2014) and explore their diversity and dynamics in various ecosystems.

Although microbial distributions have been widely investigated in different ecosystems (Liu et al. 2015; Jiao et al. 2017a; Mo et al. 2018), there is still an ongoing debate about the factors that drive the distribution processes of microbial communities over time (Dini-Andreote et al. 2015; Xue et al. 2018). Recently, several studies have compared the relative influence of different environmental factors on abundant and rare taxa in different ecosystems (Yan et al. 2017; Xue et al. 2018). For example, Yan et al. (2017) revealed that rare bacterial species are less sensitive to changing conditions than abundant one at the early stages of deglaciation. Jiao et al. (2017a) found that the assemblages of rare and abundant bacteria are driven by different factors in oil-contaminated soils. Rare taxa are primarily affected by stochastic processes, while abundant taxa are mostly governed by edaphic factors (Jiao et al. 2017a). In contrast, Gobet et al. (2011) found that both rare and abundant bacterial taxa are influenced by similar environmental factors under nutrient stress in coastal sand. Thus, it is not clear whether rare microbial taxa show contrasting biodiversity patterns or differing ecological responses compared with abundant taxa when exposed to environmental changes, especially in developing ecosystems.

Despite the knowledge that the diversity and composition of microbial communities are driven by environmental conditions, our understanding of the complex interactions among microbial taxa is still limited (Barberán et al. 2014; Schloter et al. 2018). A solution may be to use network analysis-based approaches, which have been shown to be a powerful way to examine complex interactions within microbial communities (e.g., symbiosis, parasitism, competition, and predation) (Steele et al. 2011; Eiler et al. 2012). Different topological features of the network (e.g., degree, betweenness centrality, complexity, modularity, and clustering coefficient) can be used to identify the keystone species and estimate the stability of microbial communities (Banerjee et al. 2018). For example, a high betweenness centrality indicates the core and central location of a node in the network, whereas a low betweenness centrality indicates a peripheral location (Ma et al. 2016; Jiao et al. 2017b). Moreover, the presence or absence of certain interactive links might be strong indicators for a particular tightness of the interactive system (Schloter et al. 2018). Although some studies have been conducted to investigate microbial network interactions, which are affected by abiotic and/or biotic factors (Ma et al. 2016; Jiao et al. 2017b), very few studies have linked the co-occurrence network patterns of rare and abundant subcommunities with the variation in abiotic and/or biotic conditions, particularly in glacier forefields. Most studies have focused on abundant bacterial communities within “young” chronosequences (<200 years) along glacier forefront chronosequences via next-generation sequencing (Brown and Jumpponen 2014). However, the importance of rare biosphere and microbiota interactions has rarely been examined. In this study, we conducted network analyses for rare and abundant subcommunities in deglaciated soils to provide a deeper understanding of microbial interactions.
The habitats formed by Hailuogou Glacier Chronosequence provide an ideal natural place to study the diversity and trajectories of microbial succession, as the continuous succession stages allow detailed examinations through space that substitutes for time. Our previous studies have indicated that bacterial communities are influenced by deterministic processes, while fungal communities are mediated by stochastic processes along a glacier forefield chronosequence (Jiang et al. 2018). In the present study, we aimed to utilize our existing high-throughput sequencing data sets and to conduct a deeper investigation of abundant and rare microbial communities and of their subsequent trajectories and community assembly mechanisms in relation to biotic and abiotic factors along a glacier forefield chronosequence. Specifically, we hypothesized that: (1) distinct succession and assemblage patterns exist among abundant and rare taxa, and their relative contributions to community shifts become more heterogeneous as a result of soil development along the chronosequence; (2) the abundant taxa exhibit closer relationships and more influence on other co-occurrences in the community than the rare taxa, and their interactions are strengthened at middle succession stages due to the increasing abundance of tree species and soil nutrients; and (3) the driving forces for rare and abundant subcommunity assemblages differ among the successional stages of deglaciation. Taken together, this investigation allows us to uniquely characterize how the subcommunities of abundant and rare taxa change during primary succession on a local scale. The results could shed new light on the ecological mechanisms and microbial interactions in ecosystem development under future global changes.

Materials and methods

Study sites and sampling locations

The study sites and sampling locations used in this investigation have been published in our recent paper (Jiang et al. 2018). Detailed information about microbial high-throughput sequencing data sets and some other data used in this study are available in Jiang et al. (2018). Briefly, the Hailuogou Glacier is a monsoonal temperate glacier located at the Gongga Mountain (29° 30′ to 30° 20′ N, 101° 30′ to 102° 15′ E, 7556 m a.s.l.) on the southeastern fringe of the Tibetan Plateau. The climate is characterized by a mean annual precipitation of 2000 mm, most between June and October, and considerable seasonal temperature fluctuations, ranging from −4.3 °C in January to 11.9 °C in July, with an annual mean temperature of 3.8 °C. In August 2015, the soil samples were collected from seven sites (representing successional ages 3, 12, 30, 40, 52, 80, and 120 years) undergoing long-term primary succession from bare soil to pioneer communities and eventually to climax vegetation communities in the Hailuogou Glacier Chronosequence. The approximate age of each stage was calibrated according to tree rings and soil erosion rates assessed by 137Cs method (Jiang et al. 2018). At each site, three 5 × 5-m square experimental plots with a 10-m distance between the plots were established. Five soil cores to a depth of 20 cm were collected from the center and each corner of a 50 × 50-cm quadrat using a 5-cm-diameter soil corer after removing surface material by hand. The five soil cores were combined and roots were removed. Each composite soil sample was homogenized and passed through a 2-mm sieve. Soil properties, such as pH, density (SD), total N, total P, available P, organic C and plant richness, plant litter, plant leaf N and P, and litter C/N, have been published in Jiang et al. (2018).

Sequence analyses

The bacterial DNA was extracted from approximately 0.5 g of fresh soil using the PowerSoil® DNA Isolation Kit (MoBio Laboratories, Inc., Carlsbad, USA) including a negative control, as recent studies have highlighted that microbial DNA contaminations originating from DNA extraction kits can bias the obtained sequencing output (Schöler et al. 2017; Vestergaard et al. 2017). The V4–V5 hypervariable regions of the 16S rRNA gene were profiled using the primers 515F (5′-GTGCCAGCMGCGTCAG-3′) and 907R (5′-CCGTCAATTCMTTTRAGTTT-3′) on the Illumina MiSeq PE300 platform (Illumina Inc., San Diego, CA, USA) at Genesky Biotechnologies Inc., Shanghai, China. Detailed methods and data are available in Jiang et al. (2018). Bacterial sequence processing and quality control steps were performed using the pyrosequencing pipeline tools of QIIME (http://qiime.sourceforge.net) (Caporaso et al. 2010) and USEARCH (v 7.0.1090 for Windows 32 bit) according to UPARSE software package (http://drive5.com/uparse/) (Edgar 2013). The high-quality sequences were split into OTUs through UPARSE-OTU, which is a novel “greedy” algorithm that performs chimera filtering and OTU clustering simultaneously, based on a 3% dissimilarity level. We submitted all sequence data to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (http://trace.ncbi.nlm.nih.gov/Traces/sra/) with the accession number PRJNA354498.

Parameter calculations and statistical analyses

Abundant, rare, and intermediate species analyses For data analyses, we randomly selected the same number of sequences from each bacterial sample to avoid the effect of different sequencing depths. The definition of “abundant” or “rare” taxa used at local (one sample) and regional levels (across samples) referred to Logares et al. (2014) and Liu
et al. (2015), who investigated microbial eukaryotes and bacterioplankton, respectively. Previous studies have claimed that the relative abundances of “abundant” OTUs are above 1% in one sample (Logares et al. 2014; Liu et al. 2015). However, this definition is inappropriate in our study, because there are less than 20 “abundant” OTUs according to this definition. Therefore, we defined OTUs as “abundant” or “rare” when their relative abundance was above 0.5% or below 0.01% in a sample, and OTUs’ relative abundance between 0.01% and 0.5% was defined as “intermediate” following Jiao et al. (2017a). We defined specific OTUs as “abundant” or “rare” when their average relative abundances were above 0.05% and below 0.001% across all samples, respectively. Moreover, we defined the subcommunity as the core microbiome when a subcommunity consisting of OTUs present in >80% of all soil samples (Barberán et al. 2014).

Calculation of phylogenetic diversity Both bacterial α-diversity indices (e.g., the OTU richness and Shannon diversity (H)) and β-diversity based on Bray-Curtis distance between samples were estimated by QIIME using the minimum sequence numbers (44,455 reads per sample for bacteria). Wilcoxon rank sum tests were used to identify the taxonomic distributions between rare and abundant bacteria. In all statistical analyses, p values were corrected using the false discovery rate (FDR) correction. Furthermore, nearest taxon index (NTI) and net relatedness index (NRI) were calculated for the microbial phylogenetic diversity using the null model independent swap with 999 randomization runs with 1000 iterations in R utilizing the Picante package (Kembel et al. 2010; Brown and Jumpponen 2014).

Calculation of bacterial β-diversities of different subcommunities Beta-diversity was measured using Bray-Curtis dissimilarity. Microbial community similarity was calculated by 1 − (dissimilarity of the Bray-Curtis distance metric). To better understand the β-diversity of bacterial communities, we partitioned β-diversity into two dissimilarity fractions, the balanced variation (richness) and abundance gradient (turnover), based on the method of Carvalho et al. (2012), using the bray.ppart function in the betapart R package (Baselga and Orme 2012). The bacterial species replacement and abundance loss or gain from one site to another site were estimated by turnover and/or richness fractions (Qian et al. 2005; Ulrich et al. 2009).

Correlations of microbial community compositions with environmental factors To estimate differences between the whole, abundant, and rare bacterial communities, nonmetric multidimensional scaling (NMDS) ordinations were generated using the R vegan package based on Bray-Curtis dissimilarities. The relationships between the Bray-Curtis dissimilarity of a particular bacterial community and edaphic and biotic factors were calculated by the Mantel’s test, which is based on Pearson’s correlations. On the other hand, the R package “limma” was performed to identify differentially abundant OTUs enriched at three different successional stages (e.g., early, middle, and late) by fitting the linear models to the relative abundance values of OTUs (Bulgarelli et al. 2012). Ternary plots were produced using the methods reported by Bulgarelli et al. (2012). A redundancy discriminatory analysis (RDA) was performed to investigate the influence of environmental variables, including edaphic and biotic factors, on the bacterial communities. Before RDA, the environmental variables with high variance inflation factor (VIF) > 20 were eliminated to avoid colinearity among factors, and a forward selection was conducted to select those explanatory variables with significant explaining factors for further analyses (Blanchet et al. 2008). The contributions of environmental variables, including biotic and edaphic factors, to variation in bacterial communities were calculated by variance partitioning analysis (Oksanen and Blanchet 2015).

Topological features of microbial networks To better understand the co-occurrence patterns of different bacterial taxa, a network analysis was performed with the igraph R package (Jiao et al. 2017a; Zhang et al. 2018). Rare and abundant bacterial taxa were used to build individual co-occurrence networks based on an approach used by Barberán et al. (2014). In order to avoid biases that might arise from the count data itself, adequate appropriated correlation indices (e.g., Spearman, Pearson, SparCC) and permutation validation are required (Schöler et al. 2017; Schloter et al. 2018). Thus, the Spearman’s correlation between two OTUs was calculated before constructing networks. When the Spearman’s correlation coefficients were above 0.6 and FDR-corrected p values below 0.01, the two OTUs were considered statistically robust. The networks were formed by identifying the robust correlations from a pairwise abundance comparison of OTUs. Each node and edge in the network represented one OTU and a significant correlation between nodes, respectively. To describe the network topology, four unique node-level topological features (degree, betweenness, closeness, and eigenvector centrality) and a set of indexes (the average path length, network diameter, average degree, graph density, clustering coefficient, and modularity) were calculated using the igraph R package (Zhang et al. 2018).

Construction of phylogenetic molecular ecological networks The phylogenetic molecular ecological networks of species (pMENs) were calculated and constructed, as described by Deng et al. (2012) using the random matrix theory (RMT)-based network approach. The pMEN network was produced and statistically analyzed using the existing pipeline available at http://ieg2.ou.edu/MENA (Deng et al. 2012). Network properties, including the average geodesic distance (GD),
average connectivity (avgK), and average clustering coefficient (avgCC) of the pMEN, were used as indexes to test differences among bacterial subcommunities, as described by Deng et al. (2012). Moreover, the modularity index was calculated to estimate the network stability. The topological roles of nodes were characterized by their positions (within-module connectivity (Zi) and among-module connectivity (Pi)).

All data analyses were performed with R (version 3.3.3; http://www.r-project.org) in RStudio (version 1.0.44; http://rstudio.org), unless otherwise indicated.

Results

Microbial dynamics and phylogenetic diversity of abundant and rare bacterial taxa

After quality filtering and removing chimeras, the Illumina sequencing of the V4–V5 region of 16S rRNA yielded 1,282,898 high-quality sequences (44,455–80,714 per sample), which clustered into 5584 OTUs (2432±380 per sample) (Table S1; Jiang et al. 2018). In the whole data set, the rare taxa contained 1043 OTUs, which contributed to an average of 25.1% of total OTUs but only 0.4% of sequences in each sample. Only 6.7% of total OTUs belonged to the abundant microbiome, while they accounted for a substantial fraction of the reads (mean = 63.7%) in each sample (Table S1). At the local level, the rare subcommunities showed higher α-diversities (e.g., OTU richness and Shannon index) compared with the abundant subcommunities (Fig. S1). The distributions of rare and abundant taxa across soil samples were visualized with the percent changes for rare or abundant taxa in one sample to abundant, intermediate, rare, or not detected in other samples (Fig. 1). Among locally abundant OTUs (Fig. 1a, b), 32.25% were also abundant across samples, while 10.50% become rare in other samples. Among the locally rare OTUs (Fig. 1c, d), 52.38% were not detected elsewhere, 19.26% were intermediately abundant while almost none were abundant in other samples.

There were differences between the abundant and rare subcommunity components (Fig. S2): Alphaproteobacteria, Betaproteobacteria, Sphingobacteria, Acidobacteria_Gp6, Acidobacteria_Gp3, Acidobacteria_Gp2, Acidobacteria_Gp1, and Acidobacteria_Gp4 were the most abundant taxa, while Gammaproteobacteria, Deltaproteobacteria, Planctomycetia, Cytophagia, and Chlamydia were the rarest taxa. Microbial communities were phylogenetically distinct across the chronosequence (Fig. 2). The NRI value estimated for the whole community and abundant bacteria increased with soil age, whereas NTI decreased (Fig. 2a–e). However, both NTI and NRI showed significant decreases with soil age for rare bacteria (Fig. 2c, f).

Co-occurrence patterns and topological properties of different subcommunities

The meta-community co-occurrence network captured 293,352 associations (edges) among 3728 OTUs (nodes) (Fig. 3a). Moreover, 361 and 218 nodes were attributed to the abundant and rare taxa, respectively. Compared with rare taxa (82 inner associations), the inner associations of the abundant (111,014) and intermediate taxa (117,110) were more complex. For the external associations, we observed the highest numbers of edges (63411) between abundant and intermediate nodes and lowest numbers (23) between rare and abundant nodes. These observations showed that intermediate taxa frequently co-occurred with abundant taxa, while rare taxa seldom co-occurred with abundant taxa. In the network for the core and other taxa (Fig. 3b), there were 8126 and 131,664 inner associations, respectively. There were 53,562 external associations between core and other taxa (Fig. 3b). Furthermore, the values of four topological features (degree, betweenness, closeness, and eigenvector centrality) were higher for abundant subcommunities than for rare subcommunities and also higher for core taxa than for other taxa (Fig. S3). We also generated microbial co-occurrence networks at different successional stages and calculated the network topological features to explore the microbial responses to soil development (Fig. S4). At the middle stages of the chronosequence, the networks became most clustered and complex with highest numbers of nodes, edges, and inner and external associations (Fig. S4).

Module structure of different microbial subcommunities

Based on the modules calculated by phylogenetic molecular ecological networks, we found that the whole community and the abundant, rare, core, and intermediate subcommunities were clearly parsed into 31, 13, 9, 20, and 21 major modules (Table 1). Compared to the random networks, the empirical networks had higher values of avgCC, average GD, and modularity (Table 1). Additionally, the abundant sub-network possessed higher avgCC (0.14) and average connectivity (6.64) compared to the rare sub-network (0.09 and 4.49, respectively), which implied that the rare sub-network was more scattered and less connected than the abundant sub-network (Fig. S4 and Table 1). In addition, module hubs and connectors were distributed in various microbial taxa (Fig. S5). We detected 11 module hubs in the network for all bacteria, including three affiliated with Acidobacteria, one with Alphaproteobacteria, and three with Bacteroidetes, while others belonged to Alphaproteobacteria, Betaproteobacteria, Planctomycetia, Actinobacteria, and Omnitrophica. Additionally, we detected four connectors, which belonged to Chloroflexi, Gemmatimonadetes, Proteobacteria, and
Actinobacteria. Moreover, the ternary plot analysis indicated that most OTUs affiliated to Proteobacteria, Bacteroidetes, and Acidobacteria were frequent at all successional stages (Fig. 4). Actinobacteria and Gemmatimonadetes were mainly enriched at the early and mid-stage, respectively, while Planctomycetes and Verrucomicrobia were enriched at late stages (Fig. 4).

**Contrasting responses and driving forces of abundant vs. rare subcommunities**

The β-diversity of the five subcommunities exhibited significantly different responses across the chronosequence stages (Fig. 5a). The rare subcommunity showed clearly higher β-diversity than the other four subcommunities. The abundant and core subcommunities exhibited significantly lower β-diversity than the whole community (Fig. 5a). The division of microbial β-diversity into richness and turnover (species replacement) components showed that species turnover made a higher contribution to the β-diversity than did species richness in every subcommunity (Fig. 5b).

According to NMDS plots, the bacterial communities (including all, abundant and rare taxa) of different successional stages were separated into three clusters: S1 as cluster 1, S2–S5 as cluster 2, and S6–S7 as cluster 3 (Fig. 6a–c). Compared with the rare subcommunity, the pattern of abundant bacterial subcommunity was more similar to the pattern of the whole bacterial community (Fig. S6). Rare taxa showed striking scatter compared with all and abundant taxa, as indicated by the looser clustering among different successional stages and by the comparison of between-group distances among different successional stages (Fig. 6c, f). These observations were confirmed by the results of the RDA of the rare and abundant subcommunities, which showed that the bacterial communities clustered strongly according to successional stages. Furthermore, the total variation explained by the first two RDA components was much higher for the abundant subcommunity (73.9%) than for the rare subcommunity (19.43%) (Fig. 7a, b). In addition, the variation partitioning analysis showed that the interactions of biotic and edaphic factors are more important than edaphic and biotic factors alone in determining the composition of abundant and rare subcommunities (Fig. 7c, d). Furthermore, the SEM models showed that soil age, soil pH, SOC, and plant richness were important factors for the diversity of abundant taxa, while soil age, soil pH, total N, and SOC were important factors for the diversity of rare taxa (Fig. 7e, f).

**Discussion**

**Contrasting distribution patterns of abundant and rare taxa along the chronosequence**

Our results showed that abundant and rare microbes exhibited distinct responses to the soil age of the Hailuogou Glacier Chronosequence (Fig. 1). Compared with the abundant taxa,
the rare subcommunities showed higher richness and more diversity (Table S1 and Fig. S1). Such findings are consistent with many previous studies (e.g., in east Antarctica glacial foreland and in a coastal system), thus providing further evidence for the idea that rare microbes are also important contributors of the microbial diversity across the glacier chronosequence (Yan et al. 2017; Xue et al. 2018). Most abundant taxa were frequent across samples (Fig. 1a, b), while most rare taxa were not detected in other samples (Fig. 1c, d), which suggested that rare taxa are unevenly distributed across the chronosequence (Fig. 1c, d). This may be due to competitive exclusion or limited ability to detect rare species (Schöler et al. 2017; Vestergaard et al. 2017). In addition, a large fraction of soil microbial genomes is still uncharacterized due to the spatial heterogeneity of soil and the challenge of designing universal primers based on current database coverage (Schöler et al. 2017; Vestergaard et al. 2017). Our previous studies have showed that the characteristics of soil formed by deglaciation in Gongga Mountain are heterogeneous at early successional stages (Jiang et al. 2018). Compared with rare taxa, abundant species were more persistent (Fig. 1b), which may relate to their wider niche, higher resource competitiveness, and better adaptation to environmental fluctuations (Jiao et al. 2017a, b).

The phylogenetic composition of all microbes and abundant and rare subcommunities exhibited temporal differences across the chronosequence (Fig. 2). In all microbes and abundant subcommunities, NRI increased linearly and was higher than zero at late stages, providing evidence for phylogenetic clustering, especially at the later stages (Fig. 2b). In rare subcommunities, NRI decreased linearly and was lower than zero at middle and late stages, suggesting phylogenetic divergence at middle and late stages (Fig. 2c). Moreover, we observed large differences in the taxonomic distribution between the rare and abundant bacterial subcommunities (Fig. S2). These different shifts in the bacterial phylogenetic structure suggest that there might be different niches for abundant and rare bacterial consortia along the successional stages. A likely explanation for this pattern was that environmental factors may be especially complex and heterogeneous.
during habitats being formed by deglaciation (Sun et al. 2016; Yan et al. 2017; Jiang et al. 2018). However, the NRI and NTI results were not consistent, similarly as reported by Brown and Jumpponen (2014). The reason for this discrepancy is uncertain but it may be a result of common community reordering with new but closely related community members along the Hailuogou Glacier Chronosequence (Brown and Jumpponen 2014). Furthermore, the community similarity of abundant, core, and intermediate abundant subcommunities decreased with the successional age (Fig. S7), thus implying that these bacterial subcommunities were undergoing a directional change along the successional age. However, the similarity of rare subcommunities did not show great changes (Fig. S7). Our results corroborated the different responses of rare and abundant taxa to deglaciated soils and provided strong support for the hypothesis that the distributions patterns of abundant and rare taxa are distinct across the chronosequence.

**Different assemblages and co-occurrence patterns of rare and abundant subcommunities**

Microbial communities acted significantly differently at the three main successional stages (early, middle, and late) (Fig. 4). Among the bacterial communities studied across the chronosequence, the phyla Proteobacteria, Bacteroidetes, and Acidobacteria were dominant. However, Actinobacteria and Gemmatimonadetes were mainly enriched at early and middle stages, and Planctomycetes and Verrucomicrobia members were common at late stages (Fig. 4). Bao et al. (2019) found that Actinobacteria and Verrucomicrobia were the dominant bacterial

---

**Table 1** Topological properties of the empirical molecular ecological networks (MENs) among whole, abundant, rare, core, and intermediate communities and their associated random MENs along the glacier chronosequence

<table>
<thead>
<tr>
<th>Crop rotations</th>
<th>Empirical networks</th>
<th>Random networks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Modules</td>
<td>Network size</td>
</tr>
<tr>
<td>Whole</td>
<td>31</td>
<td>423</td>
</tr>
<tr>
<td>Abundant</td>
<td>13</td>
<td>173</td>
</tr>
<tr>
<td>Rare</td>
<td>9</td>
<td>156</td>
</tr>
<tr>
<td>Core</td>
<td>20</td>
<td>204</td>
</tr>
<tr>
<td>Intermediate</td>
<td>21</td>
<td>290</td>
</tr>
</tbody>
</table>

*Network size* the number of OTUs (e.g., nodes) in a network, *avgK* average connectivity, *GD* average geodesic distance, *avgCC* average clustering coefficient
Fig. 4 Ternary plot showing relative abundance of all OTUs (>5%) for early, middle, and late stages across three successional stages (3–120 years). Each point represents an OTU. The size of each circle represents its relative abundance. Its position represents its relative abundance with respect to each compartment. Differently colored circles represent OTUs significantly enriched in one compartment compared with the others (e.g., blue at late stages, orange at middle stages, and magenta at early stages), whereas gray circles represent OTUs that are not significantly enriched in a specific compartment.

Phyla degrading cellulose in soil, as revealed in a $^{13}$C-labeled cellulose degradation experiment. Furthermore, according to the connectivity within and among modules, module hubs showed that most of the keystone microbial species belonged to the phyla Proteobacteria, Bacteroidetes, and Acidobacteria (Fig. S5). Rime et al. (2016) claimed that these highly specialized organisms can strongly adapt to oligotrophic environments in glacial forelands.

Our present results of the co-occurrence network analysis conducted for rare and abundant bacteria are consistent with our previous studies (Jiang et al. 2018), both of which revealed that both rare and abundant networks are most complex at the middle stage of the chronosequence (Fig. S4). These results may be explained, in part, by the presence of more niches created by a greater plant diversity combined with a variety of organic substrates entering soil, thus providing more opportunities for different species to interact with each other (Steinberger et al. 1999; Sun et al. 2016). Although the network visualization could predict the complex interactions within microbial communities, it is challenging to precisely show specific niche observations (Schöler et al. 2017). This may have resulted from niche sharing between the new invaders and resident species. However, the topographic features of networks, such as keystone species, modules of highly co-occurring taxa, and even hubs of connected modules, can reinforce the idea of specific niche observations (Schöler et al. 2017; Schloter et al. 2018).

Compared with rare taxa, abundant taxa possessed more inner connections and closer clustering (Table 1 and Fig. 3), as

Fig. 5 The $\beta$-diversities of different bacterial subcommunities along the Hailusogou Glacier Chronosequence based on the Bray-Curtis dissimilarity (a) and the relative contributions of differences in species turnover and richness to the $\beta$-diversity of different subcommunities (b). Different letters show significant differences (p < 0.05; multiple comparisons with Kruskal-Wallis test).
indicated by their higher values of network-level topological features (Fig. S3). In addition, abundant and core subcommunities were more widely distributed and more often exhibited complex relationships within the network than rare and other taxa (Fig. 3a, b). Our results were consistent with Jiao et al. (2017a, b), who observed that the abundant taxa have closer relationships and a greater influence on other co-occurrences compared to rare microbes in oil-contaminated soil. However, contradictory results have been obtained by Yan et al. (2017), who found that low-abundance taxa can preserve more than 90% of the total bacterial OTUs and have important roles in maintaining the stability of microbial network structures in the East Antarctica glacier foreland. A likely explanation for this discrepancy is that environmental gradients and favorable conditions for rare and abundant taxa might be different across the chronosequence, as shown by our variation partitioning analysis and structural equation model (Fig. 7a–f). Brown and Junpponen (2014) also showed that when the forefront ecosystem develops over time, more and unique niches emerge promoting the co-occurrence of more dissimilar bacteria. On the other hand, it is possible that environmental effects, including strong UV irradiation, drastic diurnal temperature fluctuations, long-lasting snow cover, and extreme nutrient limitations on a glacier forefront, create oligotrophic surroundings, in which only highly specialized organisms can thrive (Rime et al. 2016; Sun et al. 2016). Admittedly, this study and most other studies on deglaciated primary succession involve detailed analyses of one glacier forefront and may thus be site-dependent (Brown and Junpponen 2014). In future studies, in order to obtain a clearer picture of the co-occurrence patterns of rare and abundant taxa during ecosystem succession, we should query multiple glacier forefronts to track the response strategies of bacteria to environmental changes.

**Divergent driving forces of abundant and rare subcommunities along the chronosequence**

The results of our study highlighted the presence of differing drivers for rare and abundant subcommunities across the chronosequence. Compared with the abundant subcommunity, the rare taxa showed a higher β-diversity (Fig. 5) and
Fig. 7 Redundancy ordinations (a, b), variation partitioning analysis (c, d), and structural equation modeling (e, f) of the selected environmental variables for abundant and rare subcommunity structures along the Halligou Glacier Chronosequence. AP, available P; SOC, organic C; TN, total N; TP, total P. In e and f, solid and dashed arrows represent positive and negative correlations, respectively. The thickness of the arrows reflects the magnitude of the standardized coefficients. GFI, goodness-of-fit index; RMSEA, root mean square error of approximation.

relative phylogenetic divergence, as supported by the looser clustering (Fig. 6c) and lower NRI values along the successional stages. The seven stages of succession ranging from a barren stage with some mosses to climax and a lush
forest stage (Sun et al. 2016; Jiang et al. 2018) did cause great variation in environmental conditions, thus providing a further explanation for the high variation in microbial β-diversity (Dini-Andreote et al. 2015). On the other hand, the Hailuogou glacial foreland provides the same environmental conditions for abundant and rare taxa; thus, the higher β-diversity of rare taxa might be related to new microbial inputs (e.g., immigration) (Dini-Andreote et al. 2015) or dormant local inputs (i.e., “seed bank”) (Jiao et al. 2017a). In support of this hypothesis, our study found that the proportion of turnover/richness partitioning was higher in the β-diversity of the rare subcommunity than in the abundant subcommunity (Fig. 5).

In general, the responses of soil microbial communities to environmental changes are closely related to their specific properties, such as physiological tolerance, dispersal capacity, and taxonomy, although they are subjected to the same environmental conditions (Jones and Lennon 2010; Liu et al. 2015). In our study, the amount of total variance explained by RDA was much higher for the abundant subcommunity (73.9%) than for the rare subcommunity (only 19.43%) (Fig. 7a, b). Previous studies have demonstrated that the large portion of unexplained variation in rare bacteria may be due to rare species being more often subjected to other biotic interactions (e.g., competition, trophic interactions, and root exudates) and/or occupying a greater variety of ecological niches (Jones and Lennon 2010; Liu et al. 2015). These observations were confirmed by the variation partitioning analysis and structural equation models (Fig. 7c–f), which further emphasized the different roles of edaphic and biotic factors in influencing the assembly and diversity of rare and abundant taxa. More importantly, we found that the purely edaphic variables are more influential than purely biotic factors in shaping the abundant and rare subcommunities. This was actually an expected result given that soil with low fluidity limits the microbial dispersal and directly provides the substrate for microbial communities (Chen et al. 2016; Jiao et al. 2017a).

Distinct edaphic and biotic factors impacted the assembly of the abundant subcommunity (primarily soil pH, plant richness, and SOC) and the rare subcommunity (primarily soil pH and total N) in the chronosequence (Fig. 7c, f). This result was similar to the findings obtained for marine organisms in northwestern Pacific Ocean (Wu et al. 2017) but contradicted the observations of Liao et al. (2017) in lakes on the Yungui Plateau, China. These differences possibly resulted from differing environmental gradients and distribution processes (e.g., stochastic and/or deterministic processes) among the studied areas. Moreover, soil pH and plant richness were negatively correlated with abundant taxa but positively correlated with rare taxa (Fig. 7e, f). Our findings could also indirectly explain why abundant and rare bacteria fill different niche spaces and play different roles in deglaciated soils.

Conclusions

In the present study, the rare and abundant bacteria exhibited dramatic differences in co-occurrence patterns and successional trajectories across the glacier chronosequence. Rare taxa showed an uneven distribution and loose clustering, while abundant taxa were ubiquitous and tightly clustering across the successional stages. Moreover, the interactions of biotic and edaphic factors were the primary agents shaping microbial structures and explained more community variation in abundant bacteria than that in rare taxa. In addition, the network topologies of both rare and abundant taxa were more complex at the middle stages than at the early and late stages of the chronosequence, implying that more complex assembly mechanisms may exist in the abundant and rare bacterial assemblages at the middle successional stages. However, abundant taxa showed closer relationships and exhibited a greater influence on other co-occurrences when compared to rare taxa. In all, our study provides important insights into microbial community patterns and driving forces in a glacier forefield. In future studies, we should investigate further the potential assembly mechanisms and functions of abundant and rare subcommunities during the microbial community succession using different space and time scales.

Acknowledgements The authors are grateful to the Gongga Mountain Alpine Ecosystem Observation Station, Chinese Academy of Sciences for logistic support. This work was supported by the Talent Program of the Hangzhou Normal University (2016QDL020).

References

Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, Raup P, Huettel B, Reinhardt R,


Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.