Community Assembly and Co-occurrence Patterns Underlying the Core and Satellite Bacterial Sub-communities in the Tibetan Lakes

Qi Yan1,2, Jianming Deng1,3, Feng Wang4,5, Yongqin Liu2,4 and Keshao Liu4*

1School of Life Sciences, Lanzhou University, Lanzhou, China, 2Center for the Pan-Third Pole Environment, Lanzhou University, Lanzhou, China, 3State Key Laboratory of Grassland and Agro-Ecosystems, School of Life Sciences, Lanzhou University, Lanzhou, China, 4State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment (TPESRE), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, 5College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China

Microbial communities normally comprise a few core species and large numbers of satellite species. These two sub-communities have different ecological and functional roles in natural environments, but knowledge on the assembly processes and co-occurrence patterns of the core and satellite species in Tibetan lakes is still sparse. Here, we investigated the ecological processes and co-occurrence relationships of the core and satellite bacterial sub-communities in the Tibetan lakes via 454 sequencing of 16S rRNA gene. Our studies indicated that the core and satellite bacterial sub-communities have similar dominant phyla (Proteobacteria, Bacteroidetes, and Actinobacteria). But the core sub-communities were less diverse and exhibited a stronger distance-decay relationship than the satellite sub-communities. In addition, topological properties of nodes in the network demonstrated that the core sub-communities had more complex and stable co-occurrence associations and were primarily driven by stochastic processes (58.19%). By contrast, the satellite sub-communities were mainly governed by deterministic processes (62.17%). Overall, this study demonstrated the differences in the core and satellite sub-community assembly and network stability, suggesting the importance of considering species traits to understand the biogeographic distribution of bacterial communities in high-altitude lakes.

Keywords: core and satellite sub-communities, biogeographic patterns, community assembly, co-occurrence patterns, Tibetan lakes

INTRODUCTION

In natural ecosystems, bacteria within a metacommunity could be partitioned into different ecological assemblages, such as abundant or rare sub-communities and core or satellite sub-communities in light of potential importance for the community function (Unterseher et al., 2011; Jeanbille et al., 2016; Lindh et al., 2017). Defining OTUs as abundant and rare taxa are often conducted on the relative abundance of each taxa (Campbell et al., 2011; Alonso-Sáez et al., 2015), while the division of the core and satellite taxa is based on occurrence in addition to abundance (Magurran and Henderson, 2003; Hu et al., 2017b).
The latter combines the positive feedback effect between abundance and occurrence, which could improve predictions and interpretations of patterns in biodiversity reacting to environmental change (Lindh et al., 2017). The core sub-communities are composed of the dominant species that are widely distributed and play a key role in the cycle of elements (Fuhrman, 2009; Pedrós-Alió, 2012), whereas the satellite sub-communities occur in low abundance and few locations and conduct specific metabolic functions, which constitute the seed bank of biodiversity (Pester et al., 2010; Van Der Gast et al., 2011; Lindh et al., 2017; Gendron et al., 2019). Up to now, this classification has proved to be a useful tool for understanding ecological principles of microorganisms, and has been applied in marine (Lindh et al., 2017) and rivers (Hu et al., 2017b) ecosystems, but has only infrequently been implemented in lake ecosystems.

Previous studies have reported that deterministic processes and stochastic processes play important roles in the regulation of spatial distribution of bacterial communities in natural environments (Sloan et al., 2006; Oñatue et al., 2010; Langenheder and Székely, 2011; Lindström and Östman, 2011; Lindström and Langenheder, 2012; Liao et al., 2016a,b). Deterministic processes refer to environmental filtering and biotic interactions influencing the fitness of microbial communities and determine the composition and abundance of microbes (Campbell et al., 2011; Gilbert et al., 2012; Zhang et al., 2014). Conversely, stochastic processes include dispersal limitation and random changes in species relative abundance, and therefore, changes in community composition are unpredictable (Hubbell, 2001; Dini-Andreote et al., 2015; Li et al., 2019). Recently, some studies have identified that different properties or traits of microbial sub-communities may assemble by different or same mechanisms (Pandit et al., 2009; Langenheder and Székely, 2011). For instance, the core and satellite sub-communities in a salinity-influenced watershed of China were mainly drove by deterministic processes (Hu et al., 2017b). The core sub-communities in arbuscular mycorrhizal fungi (AMF) were mainly influenced by deterministic processes related to soil properties, whereas the satellite sub-communities were considerably influenced by stochastic processes (Barnes et al., 2016). However, it remains unclear whether assembly processes of the core and satellite sub-communities in Tibetan lakes are similar or different when the range of distances over hundreds of kilometers? The ecological strategy can be elucidated by the contribution of deterministic and stochastic processes to microbial community assembly (Kraft et al., 2015; Jiao and Lu, 2020). Microorganisms with microscopic sizes and high dispersal capacity could display complex interaction webs within an ecological niche, which are also key to maintaining microbial community structure (Faust and Raes, 2012). Co-occurrence network analysis provides powerful support for revealing the complex microbial community structure and interactions among microorganisms, which could reflect shared niches among community members in the real world (Faust and Raes, 2012; Mikhailov et al., 2019; Mingkun et al., 2020). Hu et al. (2017b) demonstrated that due to different ecological niches, core and satellite sub-communities play different roles in the co-occurrence network and have different network topological characteristics.

In this study, we used 454 pyrosequencing of the bacterial 16S rRNA gene to investigate the diversity and composition of core and satellite bacterial sub-communities in 47 lake water samples of 30 lakes located on the Tibetan Plateau. The Tibetan Plateau has the largest number of plateau lakes group in the world (Zhang et al., 2011). A most recent study about the biogeography of microbial communities in Tibetan lakes reported that bacterial communities were mainly controlled by salinity-driven deterministic processes (Liu et al., 2020). Although the useful information gained from this study, the spatial distribution patterns, community assembly mechanisms, and the co-occurrence patterns may be different due to their different roles of the core and satellite bacterial sub-communities in the Tibetan lakes. Therefore, we sort to determine and compare the biogeographic patterns and underlying mechanisms for the core and satellite bacterial sub-communities at a regional scale. Specifically, we tested the following three hypotheses: (1) core and satellite taxa exhibit different biogeographic patterns in lakes on Tibetan Plateau; (2) core and satellite sub-communities assembly driven by divergent processes; and (3) compared to satellite, core sub-communities show a discrepant co-occurrence pattern.

**MATERIALS AND METHODS**

**Study Area and Sampling**

We investigated surface water from 30 Tibetan lakes in 2012, China (Supplementary Figure S1). These lakes are characterized by high-altitude location (above 3,900 meters), which covered an area from 79.81°E to 96.82°E longitudinally and 28.27°N to 34.58°N latitudinally. The mean annual air temperature of the lakes ranged from −9°C to +2°C, and the surface area ranged from 8 to 2,062 km² (Supplementary Table S1).

In total, 47 water samples were collected from 30 Tibetan lakes. The Schindler sampler was used to collect approximately 1L surface water samples (~0.5m depth) from twenty lakes, respectively. Duplicate samples were collected at the same time from same points from 10 lakes, AGC, BC, BGC, DZC, GZC, NMC, PE, PMYC, Yamdrok, and ZGTC (Supplementary Table S1). Water samples from each site for bacterial community analyses were pre-filtered through 20μm mesh (Millipore, United States) for removal of the large plankton and particles, and all filtrates were subsequently filtered through a 0.22μm polycarbonate membrane (Millipore, united states). Afterward, the membranes were put in sterile 2ml microcentrifuge tubes and were stored at −80°C for DNA extraction. Latitude, longitude, and altitude were measured using the Global Positioning System during the field work.

**DNA Extraction, Bacterial 16S rRNA Amplification, and 454 Sequencing**

Microbial DNA was extracted from filters using a FastDNA® Spin kit (MP Biomedicals, Santa Ana, CA) according to the
manufacturer's instructions. It was checked for concentration and purity using a NanoDrop Spectrophotometer (ND-1000 Thermo Fisher Scientific, Wilmington, DE, United States). The V4-V5 region of the bacterial 16S rRNA genes was amplified using the primer pair 515F (5'-GTGCCAGCMGCGGCTAA-3') and 907R (5'-CCGTCAATTCMTTTRAGTTT-3'; Christopher et al., 2011). An aliquot of 10 ng purified DNA template from each sample was amplified in triplicate in a 50 μl reaction system. The amplification conditions were as follows: 30 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 30 s, with a final extension at 72°C for 10 min (Liu et al., 2019b). Then, triplicate PCR products for each sample were pooled in equal quantity and purified using agarose gel DNA purification kits (TaKaRa, Japan). Finally, running on a Roche FLX 454 pyrosequencing machine (Roche Diagnostics Corporation, Branford, CT, United States; Liu et al., 2016). Raw sequence reads have been submitted to NCBI (BioProject ID PRJNA306720).

**Processing of Pyrosequencing Data**

Paired-end reads were quality trimmed using Trimmomatic v0.30 (Bolger et al., 2014) and combined using FLASH software (Magoč and Salzberg, 2011). The raw sequences data were subsequently analyzed using QHIME v1.9.0 (Caporaso et al., 2010). The reads which had ambiguous bases and mismatches to the barcode or primers or chimeric characteristics were discarded. Then, the sequences were clustered into OTUs using UPARSE algorithm in USEARCH v 11.0.667 with a 97% threshold of sequence similarity (Edgar, 2013). Representative sequences of each OTU were aligned using PyNAST (DeSantis et al., 2006). Taxonomic identity of each phylotype was determined using the SILVA 132 database (Quast et al., 2013) via the RDP classifier (Wang et al., 2007). Before tree construction, the filter_alignment.py script in qiime1 was used to remove highly variable regions, and then, a phylogenetic tree was constructed based on Neighbor-joining method (Saitou and Nei, 1987). All eukaryote, chloroplasts, mitochondria, and unknown sequences were culled before the OTU table was generated. To avoid biases generated by differences in sequencing depth and to make samples comparable, samples were randomly rarefied to the minimum number of retrieved sequences in the whole sample (2210). After taxonomies had been assigned, we deleted all archaea OTUs and obtained 5,233 OTUs and 103,870 sequences.

**Core and Satellite Sub-Community Classification**

The Poisson model of species abundance was examined by Krebs’ method, and the dispersion index was tested by Chi-square test to partition the bacteria into the core and satellite sub-communities (Van Der Gast et al., 2011; Hu et al., 2017b). Bacterial taxa that occurred only in a single sample were excluded from this analysis because their distributed in space would have no variance. Briefly, OTU occurrence plotted against the index of dispersion (the ratio of variance to the mean abundance) for each OTU, taking 2.5% of the χ² distribution as the confidence limit. OTUs that below the interval following a random distribution were considered as satellite sub-communities, whereas above were non-randomly distributed core sub-communities. Calculations were performed using the “vegan” and “plyr” R packages (Hu et al., 2017b).

**Distance Decay of the Community Dissimilarity**

To evaluate the distance decay of community similarity, the linear regression between In-transformed geographic distances and the Bray-Curtis dissimilarities was generated based on ordinary least squares. The relationships were evaluated using the Mantel test. The statistical significance of such comparisons was determined using 999 permutations and the analyses were performed using the “mantel” function of the “vegan” package in R (Jiao et al., 2020). Permutation test was used to test for significant differences between slopes in the “simba” R package (Nekola and White, 1999). Geographical distance between samples was calculated from the latitude and longitude coordinates using the “geosphere” packages (Hijmans, 2019).

Bray-Curtis dissimilarities were based on the core and satellite OTU tables using the vegdist function in the “vegan” package (Zhu et al., 2019).

**Phylogenetic Null Model Analysis**

Null model was used to quantify the contribution of different ecological processes (stochastic and deterministic; Stegen et al., 2013, 2015). This approach uses the beta mean nearest taxon distance (βMNTD) to represent the pairwise phylogenetic turnover between communities, and beta-nearest taxon index (βNTI) to represent the environmental impacts calculated by the standard deviation of the observed βMNTD from the βMNTD of the null model. When beta-nearest taxon index (βNTI) < 2 and ≥ 2 was identified as homogeneous selection and heterogeneous selection, respectively. Moreover, 999 random permutations of communities generate a null distribution of Bray-Curtis dissimilarity, and a Raup-Crick metric (RC Bray) is subsequently calculated by comparing empirically observed Bray-Curtis and simulated null distribution. The βNTI < 2 and RC Bray < −0.95 or the βNTI < 2 and RC Bray ≥ 0.95 were identified as homogenizing dispersal and dispersal limitation, respectively. When the βNTI < 2 and RC Bray < −0.95 were identified as “Undominated” (Dini-Andreote et al., 2015; Stegen et al., 2015; Isabwe et al., 2019). To demonstrate which process contributed more to the DDR slopes between the core and satellite sub-communities, samples controlled by dispersal limitation and heterogeneous selection were separately extracted from both sub-communities according to the results of Stegen’s null model. Then, the DDR slopes were calculated separately.

**Habitat Niche Breadth**

Niche breadth is often used to identify different levels of habitat specialization, which is a crucial trait that affects the relative importance of selection and dispersal limitation affecting...
communities (Pandit et al., 2009; Logares et al., 2013; Liao et al., 2016a). Niche breadth was calculated using Levins’ niche breadth (Levins, 1968) index ($B$):

$$
B_j = \frac{1}{\sum_{i=1}^{N} p_i^2}
$$

where $B_j$ represents the habitat niche breadth; $p_i$ is the mean relative abundance of OTU $j$ in lake $i$ (i.e., one of the 30 water samples); and $N$ is the total number of communities. A high B-values indicate a wide range of OTUs and even distribution, representing wide habitat niche breadth and more metabolic flexibility at the community level (Wu et al., 2018).

Network Construction
We used network analysis to examine co-occurrence networks of core and satellite sub-communities. To reduce noise and complexity of the datasets, we kept OTUs that appeared in ≥5 samples for network analysis. Spearman's rank coefficients ($\rho$) between those OTUs were calculated pairwise by the “Hmisc” package in an R environment. Only robust correlations with Spearman's correlation coefficients ($\rho$) > 0.6 and false discovery rate-corrected values of $p$ < 0.01 were used to construct networks (Hu et al., 2017a). Each node represents one OTU, and each edge represents a strong and significant correlation between two nodes. Network visualization was performed using the interactive platform Gephi (0.9.2). We use the “igraph” R package to calculate the node-level network topologies features (node degree, betweenness centrality, closeness centrality, and transitivity) and were examined by Kruskal-Wallis test to measure differences (Bastian et al., 2009; Mo et al., 2020). In addition, “igraph” package was used to calculate and compare the topology characteristics of the real networks and 10,000 Erdős-Rényi random networks, which had the same number of nodes and edges as the real networks (Jiao et al., 2020).

To understand the stability of the core and satellite networks, two indices were used to characterize the stability, including robustness and vulnerability. Natural connections were used to assess network stability by removing nodes in the network to evaluate the rate of robustness degradation (Peng and Wu, 2016). Network vulnerability is expressed as the maximal vulnerability of nodes in the network (Yuan et al., 2021).

Statistical Analyses
Diversity index was analyzed using “vegan” package in the R environment (R Core Team, 2013). Kruskal-Wallis test was performed with the PAST software to compare the α-diversity and niche differences of the core and satellite sub-communities and to identify the significantly and differentially abundant phyla/classes and genera between the core and satellite sub-communities. All the R analyses were performed in version 3.6.1.

RESULTS

OTUs Composition and Diversity of the Core and Satellite Sub-Communities
After removing low quality sequences, a total of 103,870 reads were obtained in this study and clustered into 5,233 OTUs (Table 1). Good’s coverage ranged from 86 to 96%, indicating that sequences identified in these samples represent the majority of bacterial sequences present in the collected water samples (Supplementary Table S2). A positive relationship between the mean abundance of OTUs and their occurrence was observed ($R^2 = 0.24$, $p < 0.001$; Figure 1A). The 1,276 OTUs fit a $\chi^2$ test were defined as the satellite sub-communities that with 4,500 (4.33%) reads. In contrast, the remaining 809 OTUs (93,493 reads), surpassing 2.5% confidence limit line of $\chi^2$ distribution, formed core sub-communities and accounted for 90.01% of the total reads (Table 1; Figure 1B).

In all taxa, Proteobacteria, Bacteroidetes, and Actinobacteria were the dominant phyla in the core and satellite sub-communities, together accounting for 71 and 78.62% of each sub-community sequences, respectively (Supplementary Table S3). Cyanobacteria was significantly abundant in the core sub-communities, while Betaproteobacteria, Alphaproteobacteria, Gammaproteobacteria, Gemmatimonadetes, Thermi, and TM7 were found to be significantly dominant in the satellite sub-communities (Kruskal-Wallis test, $p < 0.05$; Supplementary Figure S2). Among them, the genera Arthrobacter, B-42, Loktanella, and Rhodobacter harbored a higher abundance in the satellite sub-communities, while some genera, such as Planococcus, Psychroflexus, and Synechococcus exhibited significantly higher abundances in the core sub-communities. The α-diversity indices of the core and satellite sub-communities were compared based on Chao1 and Shannon indices (Figure 2). Both Chao1 and Shannon indices of the satellite sub-communities were significantly higher than those of the core sub-communities ($p < 0.001$).

Geographic Patterns of the Core and Satellite Sub-Communities
Distance-decay relationship (DDR) is a fundamental pattern in ecology, in which community similarity decreases as the geographic distance increases. In the current study, although the significant positive DDRs (Mantel $p < 0.05$; Figure 4) were observed, the fitness values were relatively low ($R^2 < 0.1$), indicating weak relationship of community dissimilarity with geographic distance for the core and satellite sub-communities. Meanwhile, the slope of DDRs was significant ($p < 0.01$) steeper for the core sub-communities (0.019) than that of the satellite sub-communities (0.004).

**TABLE 1** | The number of OTUs and sequences of the core and satellite bacteria sub-communities.

| Taxa             | OTU number | Sequence number |
|------------------|------------|-----------------|
| ALL OTUs         | 5,233      | 103,870         |
| Core OTUs        | 809 (15.46%) | 93,493 (90.01%) |
| Satellite OTUs   | 1,276 (24.38%) | 4,500 (4.33%)   |
Niche Breadth and Ecological Processes Underlying the Core and Satellite Sub-Communities

The niche breadth ($B$) analysis indicated that the average niche breadth for the core communities (4.11) was significantly wider than that of the satellite communities (2.65; $p<0.001$; Supplementary Figure S3).

The results of the null model quantify the relative contributions of major ecological processes of the core and satellite sub-communities in the Tibetan lakes (Figure 5). We found...
FIGURE 3 | Comparison of α-diversities of the core and satellite bacterial sub-communities. (A) Chao1 index and (B) Shannon index.

FIGURE 4 | The relationship between geographic distances and Bray-Curtis dissimilarities of the (A) core and (B) satellite bacterial sub-communities. The red line in each plot represents a linear regression model fit to Ln (geographic distance+1) vs. Ln (Bray-Curtis community dissimilarity). Gray band around the line indicates 95% confidence interval. Asterisks denote significant different between slopes ($p<0.01$).
that heterogeneous selection was the most important process structuring of the core and satellite bacterial sub-communities (41.26 and 55.32% of the overall community turnover, respectively). Dispersal limitation and undominated showed similar relative importance in shaping the core sub-communities (32.56% vs. 21.74% of the turnover; Figure 5A). In contrast, undominated process contributed about 27.38% to shaping the satellite sub-communities, while that of dispersal limitation process was less than 5.5% (Figure 5B). Generally, the results recommended that stochastic processes explained a higher proportion of the core sub-community variations than deterministic processes, while satellite sub-communities were primarily affected by deterministic processes. As shown in Supplementary Figure S4, core sub-community turnover that controlled by dispersal limitation process showed a negative distance-decay slope (−0.004), while satellite sub-community turnover showed a slight positive distance-decay slope (0.001). On the contrary, core sub-community turnover governed by heterogeneous selection process was significantly higher (p<0.05) than that of satellite sub-communities (Supplementary Figure S4).

**Co-occurrence Network of the Core and Satellite Sub-Communities**

The whole network included 5,145 associations among 518 microbial OTUs and exhibited scale-free characteristics (Power law: $R^2=0.71$). Meanwhile, the real network exhibited higher values of average clustering coefficient (0.58 vs. 0.04), average path length (5.99 vs. 2.41), and modularity (0.64 vs. 0.19) than those of the respective Erdős-Rényi random, suggesting the real network was non-random and modular structure (Table 2). We identified 423 and 95 core and satellite OTUs throughout the whole network, respectively (Figure 6A). In addition, the degree, betweenness, closeness, and eigenvector showed significantly higher values in the core sub-communities bacterial co-occurrence patterns than that in the satellite sub-communities in the Tibetan lakes (p<0.01; Figure 6B). The core sub-communities co-occurrence network exhibited higher robustness structure and lower network vulnerability compared to the satellite sub-communities (Supplementary Figure S5), indicating that the core sub-community network was more stable.

**DISCUSSION**

Community assembly mechanisms can predict community changes in space and time gradients, influence hydro-biogeochemical function, and have potential implications for ecosystem function and biodiversity conservation (Jiang and Patel, 2008; Hanson et al., 2012; Nemergut et al., 2013; Graham and Stegen, 2017). In this study, we used null model and network analysis to quantify the relative importance of ecological processes in shaping the core and satellite sub-communities and explore bacterial co-occurrence in the Tibetan lakes.

**Biogeographical Patterns of the Core and Satellite Communities**

In this study, our results showed that both of the core and satellite bacterial sub-communities displayed significant DDRs (Mantel p<0.05; Figure 4). This implies that the core and satellite bacterial sub-communities were not a random collection of taxa (Liu et al., 2015). This was consistent with previous studies on freshwater lakes, reservoirs, and marine environments (Galand et al., 2009; Liu et al., 2015; Liao et al., 2017) and provided further evidence from Tibetan lakes. However, within this general pattern, we also observed that the DDR slope of the core sub-communities was steeper than that of the satellite sub-communities, suggesting that the spatial turnover rate of the core sub-communities is higher than the satellite counterparts. This finding is consistent with the research results on bacterial communities in the reservoirs and rivers (Liu et al., 2015).

---

**FIGURE 5** | Delineation of the assembly processes underlying the core (A) and satellite (B) bacterial sub-communities. The percentage of turnovers governed by a process is used to represent its relative importance in community assembly. Low percentage contributions (<1.5%) are not shown.
However, our results are opposite to an earlier study which revealed that the satellite taxonomic communities had higher spatial turnover rates than core counterparts in Yongjiang river watershed of China (Hu et al., 2017b). This contrary conclusion might be ascribed to the different research zones and habitat types. In Hu et al. (2017b)’s study, 29 river surface water were considered. However, in the present study, 30 Tibetan lakes were studied, which exhibited larger geographic gradient.

Ecological Processes Underlying the Assembly of the Core and Satellite Communities

Studies have shown that environmental filtering or dispersal-related processes can generate the DDRs of bacterial communities (Lindström and Östman, 2011; Liu et al., 2015). The process of environmental filtering generally differentiates microbial composition among locations, which will tend to produce a distance-decay relationship. By contrast, high dispersal will weak or eliminate the distance-decay relationship by counteracting compositional differentiation and the distance-decay relationship should be stronger when dispersal is more limited (Hanson et al., 2012). To identify the main reason underpinning the different DDR between the core and satellite sub-communities, we used a null model that did not involve spatial and explanatory variables. Our results suggest that the heterogeneous selection was the most important process in structuring the core and satellite sub-communities (41.26% vs. 55.32%). In heterogeneous selection, the slope of the core sub-communities exhibited significantly higher ($p < 0.05$) than satellite sub-communities, while DDR controlled by dispersal limitation showed an opposite trend in the core and satellite sub-communities (Supplementary Figure S4). This could imply the important role of heterogeneous selection in shaping the different DDR slopes between the core and satellite sub-communities.

Table 2: Topological properties of co-occurrence networks of the Tibetan lake bacterial communities and their corresponding random networks.

| Network properties       | Value          |
|-------------------------|----------------|
| **Empirical networks**  |                |
| Nodes                   | 518            |
| Edges                   | 5,145          |
| Average clustering coeff| 0.58           |
| Diameter                | 9,73           |
| Average path length     | 5.99           |
| Average degree          | 19.86          |
| Modularity              | 0.64           |
| Power-law model         | 0.71           |
| **Random networks**     |                |
| Average clustering coeff| 0.04±0.001     |
| Average path length     | 2.41±0.001     |
| Modularity              | 0.19±0.003     |

of the core and satellite sub-communities may form different cell size communities, generating the discrepant assembly mechanisms. Cell size has often been regarded as an important factor in affecting the metabolic versatility (Farjalla et al., 2012) and dispersal potential (Liu et al., 2019c) of organisms. The metabolic activities and dispersal abilities due to the effect of cell size may affect stochasticity or deterministic adequacy for explaining their community assembly (Zinger et al., 2019; Gao et al., 2020). Finally, the 21.74 and 27.38% undominated processes that contributed to the assembly of the core and sub-communities, indicating that these sub-communities were shaped by a more complex assembly mechanism (Mo et al., 2018).

Bacterial sub-communities with wider niche breadth may have greater potential for dormancy (Wu et al., 2018; Mo et al., 2020). Thus, differences in niche breadth due to different species taxa and abundance in the core and satellite sub-communities (Supplementary Figure S2; Supplementary Table S3) can produce different dormancy strategies. The core sub-communities with wider niche breadth are more susceptible to enter dormancy of their cells than the satellite sub-communities, and reducing the active taxa affected by deterministic processes. This is an important metabolic strategy for microbial cells to manage with environmental stress and less vulnerable to deterministic processes (Lennon and Jones, 2011; Masanori et al., 2011; Nemergut et al., 2013).

Co-existence Patterns of the Core and Satellite Communities

Co-occurrence networks can partially reveal complex interactions within microbial communities and have been considered to be an important tools for investigating potential interactions within microbial food webs (Faust and Raes, 2012; Liu et al., 2019a; Du et al., 2020). Network topology features can reflect the complex interactions between microorganisms in the community. The present study showed that the core sub-communities have significantly higher network topology than satellite (Figure 6B). This suggests that there are stronger and more complex webs of interaction in the core than in the satellite sub-communities. Specific properties promoted community stability in co-occurrence networks, and competition could also increase the stability of the community structure (Ghoul and Mitri, 2016). More complex network structure indicates stronger connections between competitors and more efficient resource transfer (Morriën et al., 2017; Yao et al., 2019). The core sub-community network had higher connectivity than satellite networks (Supplementary Figure S5A), which suggests that it was more efficient at transferring information, energy, and resources. On the other hand, the simple network structure also reflects the fragility of the satellite bacterial sub-community structure in the case of ecosystem perturbations (Wang et al., 2018). In addition, our study also supports to Ghoul and Mitri’s (2016) argument that increasing the complexity of a co-occurrence network leads to more stable co-existence patterns.
CONCLUSION

In summary, this study has provided a better understanding of assembly mechanisms and co-occurrence patterns of the core and satellite bacterial sub-communities across multiple Tibetan lakes. Our results demonstrated that the core bacterial sub-communities exhibited similar biogeographic patterns to the satellite counterparts, but their patterns were generally shaped via different assembly mechanisms. For the core sub-communities, stochastic processes played important roles, while deterministic processes are of importance in shaping the satellite sub-community assembly. The co-occurrence pattern of the core sub-communities was more complex and more stable. Therefore, in future studies, bacterial community should be distinguished by traits of taxa in order to obtain comprehensive understanding of the biogeography and co-occurrence patterns of lake bacterial community.

Although the ecological model used can provide the in-depth results on the community assembly mechanisms, we acknowledge some limitations in the study. For example, the null model relies more on phylogenetic tree and lacks an explanation of the results through environmental factors. Therefore, it is necessary to use the null model and environment factors analysis at the same time in the subsequent research in order to obtain richer conclusions.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at NCBI (BioProject ID PRJNA306720).

AUTHOR CONTRIBUTIONS

This work was conceived by KL and YL. Field work was done by KL. Laboratory work was done by KL and FW. Analysis was carried out and the manuscript was written by QY and KL. Writing-reviewing and editing were done by JD. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (grant nos. 41771086 and 42006200), the Second Tibetan Plateau Scientific Expedition and Research Program (STEP; grant no. 2019QZKK0503), the National Key Research and Development Program Project (grant no. 2019YFC1509103), and the Strategic Priority Research Program (A) of the Chinese Academy of Sciences (grant no. XDA20050101).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.695465/full#supplementary-material
Zinger, L., Taberlet, P., Schimann, H., Bonin, A., Boyer, F., De Barba, M., et al. (2019). Body size determines soil community assembly in a tropical forest. *Mol. Ecol.* 28, 528–543. doi: 10.1111/mec.14919

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Yan, Deng, Wang, Liu and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.