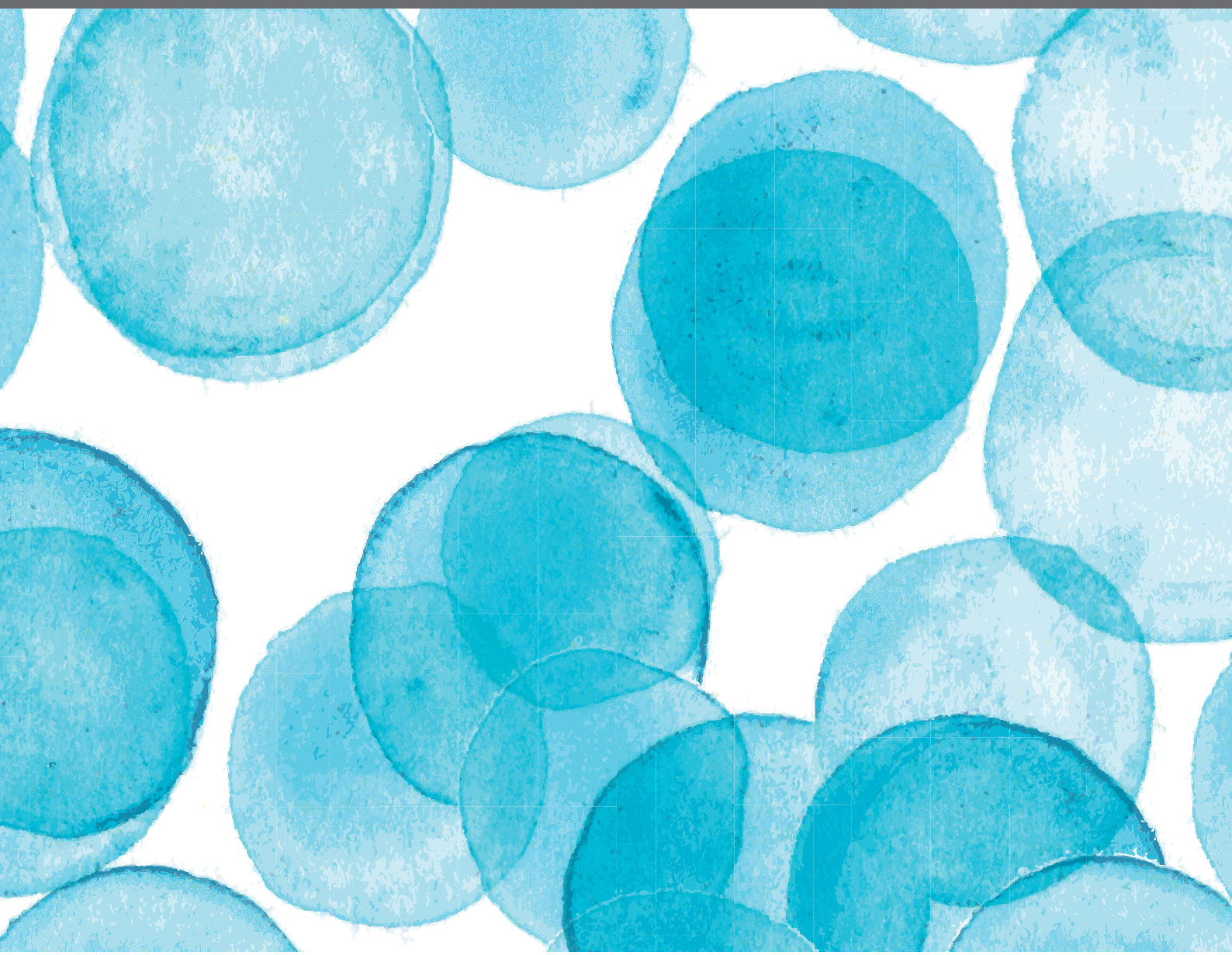
COMMUNITY ASSEMBLY MECHANISMS SHAPING MICROBIOME SPATIAL OR TEMPORAL DYNAMICS

EDITED BY: Daliang Ning, Stilianos Fodelianakis and Jianjun Wang
PUBLISHED IN: Frontiers in Microbiology



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88976-525-6
DOI 10.3389/978-2-88976-525-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

COMMUNITY ASSEMBLY MECHANISMS SHAPING MICROBIOME SPATIAL OR TEMPORAL DYNAMICS

Topic Editors:
Daliang Ning, University of Oklahoma, United States
Stilianos Fodelianakis, Swiss Federal Institute of Technology Lausanne, Switzerland
Jianjun Wang, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences (CAS), China

Citation: Ning, D., Fodelianakis, S., Wang, J., eds. (2022). Community Assembly Mechanisms Shaping Microbiome Spatial or Temporal Dynamics. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-525-6

Table of Contents

frontiers in Microbiology

ORIGINAL RESEARCH
published: 02 July 2021
doi: 10.3389/fmicb.2021.689762

Distinct Functions and Assembly Mechanisms of Soil Abundant and Rare Bacterial Taxa Under Increasing Pyrene Stresses

Yuzhu Dong[1,2†], Shanghua Wu[1,2†], Ye Deng[1,2], Shijie Wang[1,2], Haonan Fan[1,2], Xianglong Li[1,2], Zhihui Bai[1,2] and Xuliang Zhuang[1,2]*

[1] Key Laboratory of Environmental Biotechnology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China, [2] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, China

Elucidating the relative importance of species interactions and assembly mechanisms in regulating bacterial community structure and functions, especially the abundant and rare subcommunities, is crucial for understanding the influence of environmental disturbance in shaping ecological functions. However, little is known about how polycyclic aromatic hydrocarbon (PAH) stress alters the stability and functions of the abundant and rare taxa. Here, we performed soil microcosms with gradient pyrene stresses as a model ecosystem to explore the roles of community assembly in determining structures and functions of the abundant and rare subcommunities. The dose–effect of pyrene significantly altered compositions of abundant and rare subcommunities. With increasing pyrene stresses, diversity increased in abundant subcommunities, while it decreased in the rare. Importantly, the abundant taxa exhibited a much broader niche width and environmental adaptivity than the rare, contributing more to pyrene biodegradation, whereas rare taxa played a key role in improving subcommunity resistance to stress, potentially promoting community persistence and stability. Furthermore, subcommunity co-occurrence network analysis revealed that abundant taxa inclined to occupy the core and central position in adaptation to the pyrene stresses. Stochastic processes played key roles in the abundant subcommunity rather than the rare subcommunity. Overall, these findings extend our understanding of the ecological mechanisms and interactions of abundant and rare taxa in response to pollution stress, laying a leading theoretical basis that abundant taxa are core targets for biostimulation in soil remediation.

Keywords: assembly processes, abundant and rare taxa, biodegradation, environmental adaptability, polycyclic aromatic hydrocarbons

OPEN ACCESS

Edited by:
Jianjun Wang,
Nanjing Institute of Geography and Limnology (CAS), China

Reviewed by:
Jun Yang,
Institute of Urban Environment (CAS), China
Rafael Vazquez-Duhalt,
Universidad Nacional Autónoma de México, Mexico

***Correspondence:**
Xuliang Zhuang
xlzhuang@rcees.ac.cn

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *01 April 2021*
Accepted: *20 May 2021*
Published: *02 July 2021*

Citation:
Dong Y, Wu S, Deng Y, Wang S, Fan H, Li X, Bai Z and Zhuang X (2021) Distinct Functions and Assembly Mechanisms of Soil Abundant and Rare Bacterial Taxa Under Increasing Pyrene Stresses. Front. Microbiol. 12:689762. doi: 10.3389/fmicb.2021.689762

INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) have been extensively researched on account of their biotoxicity and high detection rate in the environment (Ma and Cao, 2010; Keith, 2015). Soil is the reservoir and transfer station of PAHs in the environment, which cause great effects on microorganisms and soil function (Crampon et al., 2018). The soil microbial community consists

of a few taxa with high abundance (defined as abundant bacterial taxa) and a great quantity of taxa with low abundance (defined as rare bacterial taxa) (Hanson et al., 2012; Jousset et al., 2017), and prior studies indicated that abundant and rare microbial taxa often exhibited different distribution patterns and functional traits (Xue et al., 2020; Zhao et al., 2020). To date, some studies have evaluated the influence of PAHs on the whole bacterial community (Gao et al., 2019) and found obvious changes in community compositions and functions of soil microecosystems (Lors et al., 2010; Shahsavari et al., 2019). However, it is not clear how the abundant and rare bacterial taxa respond and adapt to different levels of PAH disturbance in soils and, in turn, the dose–effect of PAHs on the functions of the two subcommunities.

Recently, a growing body of research has emphasized the ecological importance of rare taxa (Lynch and Neufeld, 2015; Jousset et al., 2017), and distinct succession patterns and functional characteristics were found in abundant and rare taxa (Jia et al., 2018; Jiao and Lu, 2020b; Rocca et al., 2020). The abundant taxa are usually perceived as the most active category in biogeochemical cycles, especially carbohydrate metabolism, and take up the coring niche (Jiao et al., 2017; Kurm et al., 2019; Liang et al., 2020). Besides, due to their high abundance, the abundant taxa might present greater survivability to environmental stresses (Jiao et al., 2019; Jiao and Lu, 2020b). As for the rare taxa, according to numerous studies, they serve as a reservoir of genetic and functional diversity for the whole community and contribute to the maintenance of microbial diversity (Jousset et al., 2017; Rocca et al., 2020). That is, when faced with environmental disturbance, the rare taxa could respond promptly to maintain community stability (Ji et al., 2020; Zhang et al., 2020). For example, some rare taxa may turn to be dominant taxa in the community to enrich ecosystem functions for the disturbance (Jiao et al., 2019; Du et al., 2020). The interaction and transformation of abundant and rare bacterial taxa are critical to maintaining soil functional redundancy and community stability (Li et al., 2019; Liang et al., 2020). Exploring the co-occurrence relationship and succession pattern between abundant and rare bacterial taxa is conducive to identify the keystone taxa in soil microecosystems, so as to better predict and restore soil functions under different levels of PAH pollution (Stegen et al., 2012; Jiao and Lu, 2020a). But so far, very few works have addressed the mutual relations and succession patterns of the abundant and rare bacteria taxa under different levels of PAH concentrations.

Community assembly, studying the processes that shape the traits and abundance of taxa in ecological communities, is a key issue in evaluating the influence of environmental pollutants on the soil bacterial community and in turn affects the transfer and biodegradation of the pollutants (Stegen et al., 2013; Jia et al., 2018; Guittar et al., 2019). Since there are distinct ecological responses to environmental changes, the abundant and rare subcommunities might be dominated by different assembly processes under disparate disturbances (Jia et al., 2018; Liang et al., 2020). Some researchers suggested that abundant and rare bacterial taxa have a similar community assembly mechanism (Liao et al., 2017), while others reached the opposite conclusion that there is a considerable discrepancy in the proportion of stochastic and deterministic processes in the assembly of abundant and rare subcommunities (Du et al., 2020; Ji et al., 2020). Obviously, we still lack a comprehensive understanding of the universality in community assembly of abundant and rare bacterial communities under specific environmental disturbances. Clarifying the essential mechanisms for microbial succession and assembly under PAH stress has vital importance on soil microbial remediation (Dua et al., 2002; Gavrilescu et al., 2015).

In this study, we chose pyrene, four-aromatic ring model compounds for PAHs, to establish a soil microcosm incubation experiment in order to simulate different pollution levels with the purpose of: (i) clarifying the uncertain succession patterns (such as diversity, distribution, function) of abundant and rare bacterial taxa under different levels of pyrene stresses; (ii) revealing the co-occurrence relationships of the rare and abundant bacterial taxa in response to different levels of pyrene stresses; (iii) assessing the major processes controlling the assembly of the abundant and rare bacterial subcommunities along with pyrene stresses. Based on different ecological functions and environmental adaptations of abundant and rare bacterial taxa, we hypothesized that soil abundant bacterial taxa may have broader adaptivity to serious PAH stresses than rare bacterial taxa and distinct mechanisms dominated the assembly of the abundant and rare subcommunities. Our study could help predict the responses of soil bacteria to environmental pollutions and understand the generation and maintenance of bacterial functions in further *in situ* bioremediation technologies.

MATERIALS AND METHODS

Soil Sampling and Experimental Setup

Surface soil samples were collected by thoroughly mixing several soil cores from a farmland (40°23′ E and 116°40′ N) in Changping District, Beijing, North China. All the samples were packed in sterile self-sealing bags and then sent to the laboratory immediately under dry ice conditions. The soil samples were passed through a 2-mm sieve after naturally air-dried under dark conditions to ensure homogeneity for subsequent experiments. The soil had a pH of 7.9, and the soil organic carbon and total nitrogen concentrations were 1.12 and 1.34 g kg^{-1}, respectively. The total concentration of 16 Environmental Protection Agency (EPA) priority PAHs in this soil was 66 $\mu g\ kg^{-1}$, which was far less than the standards of unpolluted soils (<200 $\mu g\ kg^{-1}$) according to the classification of PAH pollution levels indicated by Maliszewska-Kordybach (1996). Besides, the soil also was below the soil pollution risk screening values for agricultural and development land from the Soil Environmental Quality Risk Control standard for soil contamination of agricultural land (GB15618-2018) and development land (GB36600-2018) issued by the Ministry of Ecology and Environment of China.

Five treatments were carried out with six replicates: unpolluted soil (Control Treatment) and polluted soils under

four different pyrene concentrations (1, 10, 100, and 500 mg pyrene per kilogram dry soil, represented by PYR1, PYR10, PYR100, and PYR500, respectively). Pyrene-polluted soils were prepared in reference to the method of Brinch et al. (2002). Briefly, acetone stock solutions with different concentrations of pyrene were prepared first and then the same volume of stock solutions was spiked into 250-g homogenized soils. After the acetone completely evaporated, the spiked soils were thoroughly mixed with the remaining 750-g unpolluted soils. Meanwhile, an equal volume solution of pure acetone was added to the unpolluted soils in the same procedure to eliminate the effects of acetone.

For each treatment, 10-g pretreated soils were placed in a 120-ml serum bottle; then, sterile water was used to keep the soil moisture at about 60% every 2 days. All the microcosms were incubated for 35 days at 25°C in the dark. At the end of the 35-day incubation period, the pyrene concentration in Treatment PYR1 was lower than the detection limit, and all the soil CO_2 emission rates showed a downward trend. Therefore, soil samples were stored at −80°C for microbial analyses.

Residual pyrene in soils was obtained by accelerated solvent extraction (ASE 350, Dionex, Thermo Scientific, United States) and solid-phase extraction purification and quantified by gas chromatography coupled to mass spectrometry (GC-MS, Shimadzu, Kyoto, Japan). The detailed parameters of GC-MS were 50°C hold for 2 min, rise to 180°C at a rate of 20°C/min, increase at 10°C/min to 290°C, then hold for 10 min (Rombolà et al., 2015).

To further confirm the results that pyrene stresses influenced the abundance of abundant and rare taxa, and the abundant taxa played an important role in pyrene degradation, unpolluted soils from four different places (Qingdao, Changsha, Kunming, and Guangzhou, represented by QD, CS, KM, and GZ, respectively) in China were collected to establish the microcosm. The PYR500 and its controls (soils without pyrene) were harvested after 35 days of cultivation and subjected to DNA extraction, 16S rRNA gene sequencing, and the further analysis.

DNA Extraction and Illumina Sequencing

Extraction and purification of total bacterial genomic DNA from microcosms were performed using a FastDNA® SPIN Kit for Soil (MP Biochemicals, United States) following the manufacturer's instructions. The quantity and purity of soil DNA were determined by a Nanodrop 2000 UV-Vis Spectrophotometer (NanoDrop Technologies, Wilmington, DE, United States). DNA was quantified for *nidA* gene (pyrene dioxygenase gene) through the primer *nidA*-F (5′-TTC CCG AGT ACG AGG GAT AC-3′) and *nidA*-R (5′-TCA CGT TGA TGA ACG ACA AA-3′) (Ren et al., 2016). The V3–V4 region of the bacterial 16S rRNA gene was amplified using the primer 338F (5′-ACT CCT ACG GGA GGC AGC AG-3′) and 806R (5′-GGA CTA CHV GGG TWT CTA AT-3′) (Lee et al., 2017). The 20-μl PCR mixture contained 4 μl of 5 × FastPfu buffer, 2 μl of 2.5 mM dNTPs, 0.8 μl of 5 μM each primer (forward and reverse), 0.4 μl of TransStart Fastpfu DNA polymerase, 0.2 μl of bovine serum albumin, and 10 ng of sample DNA. The PCR reaction conditions were 95°C for 3 min, followed by 26 cycles of 30 s at 95°C, 55°C for 30 s, 72°C for 45 s, and a final extension at 72°C for 10 min. Each PCR product was sequenced on the Illumina MiSeq PE 300 × 2 sequencer at Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China).

Sequence Analysis

The Galaxy pipeline at the Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences[1], was used to process and analyze the high-throughput sequencing data (Feng et al., 2017). Briefly, paired-end reads were merged to a sequence of each sample according to the overlapped regions of reads (Zhang et al., 2017). Meanwhile, quality control and filtering were carried out to remove reads with lengths less than 50 bp, average score less than 20, and containing N bases. Effective sequences of each sample were obtained by distinguishing according to their unique barcodes and primer sequences with Fastp (Chen et al., 2018) and FLASH (Magoc and Salzberg, 2011). For the 30 soil samples, the mean average length was 418.87 ± 1.09 bp, and 1,374,419 quality-filtered clean sequences were obtained, ranging from 39,134 to 51,214 with a mean of 45,814 sequences per sample. Operational taxonomic units (OTUs) were picked using UPARSE at 97% similarity level, and sequences were then assigned to the SILVA reference using the RDP classifier (Quast et al., 2013). To compare different samples, we used a randomly selected subset of 39,134 sequences from each sample to normalize sequencing effort across samples. Finally, a resample OTU table was obtained for further statistical analysis.

Data Analyses

Generally, different cutoffs of the relative abundance are applied to distinguish abundant and rare bacterial taxa. In our study, (i) OTUs with a relative abundance ≥ 1% in a sample were defined as abundant bacterial taxa, (ii) OTUs with a relative abundance < 0.01% in a treatment were defined as rare bacterial taxa, and (iii) OTUs with a relative abundance < 1% in a treatment and ≥ 0.01% in a sample were defined as moderate taxa (Ji et al., 2020; Liang et al., 2020; Mo et al., 2021).

Here, α-diversity indices (Shannon-Wiener and Simpson) were calculated through the "vegan" package, and one-way ANOVA was used to determine whether there were any statistically significant differences among treatments, followed by the *post hoc* Tukey honestly significant difference (HSD) test. In addition, β-diversity was measured based on Bray–Curtis dissimilarity using metaMDS in "vegan" package, and the non-metric multidimensional scaling (NMDS) was visualized using "ggplot2" package (Oksanen et al., 2013). Furthermore, analysis of similarities (ANOSIM) was performed to test whether there were significant differences in the soil bacterial community structures at different treatments. Resistance of abundant and rare taxa in polluted soils was calculated based on the Shannon–Wiener index compared to the Control (Liang et al., 2020).

[1]http://mem.rcees.ac.cn:8080/

Co-occurrence network was carried out through Molecular Ecological Network Analysis Pipeline (MENA[2]) based on the Random Matrix Theory (RMT) and Spearman correlation (Deng et al., 2012, 2016). According to the RMT-based modeling, a cutoff was chosen to construct the network, and then global network properties were calculated including individual nodes' centrality, degree, betweenness, and clustering coefficient. Integrating all the results obtained from MENA, co-occurrence network was visualized by Gephi.

To predict the functional composition of the microbial community in the samples, the OTU abundance table was first standardized to remove the influence of the copy numbers of 16S marker gene in the species genome through Tax4Fun (Ashauer et al., 2015). Then, the corresponding relationship between SILVA classification and Kyoto Encyclopedia of Genes and Genomes (KEGG) database was established to predict microbial community function. Linear discriminant analysis Effect Size (LEfSe) was carried out to find the functions most likely to explain the differences among treatments (Segata et al., 2011).

Niche width was measured according to Levins' coefficient (Levins, 1968):

$$B_i = {}^{1}\Big/{\sum_{j=1}^{r} P_{ij}^2} \quad (1)$$

where B_i is the habitat niche width of OTU_i, and P_{ij} is the proportion of OTU_i in the total OTUs within a given resource state j. The average number of all OTUs' B was calculated to represent the niche width of the bacterial community.

Threshold indicator taxa analysis (Baker and King, 2010) was performed to analyze the threshold value of each abundant and rare bacterial taxa in response to variation of pollutant concentration. The z score of abundant and rare bacterial taxa was used to integrate taxon occurrence, abundance, and directivity.

In order to uncover the assembly mechanism of abundant and rare subcommunities, we performed the Sloan neutral community model (Sloan et al., 2010) and null model (Stegen et al., 2012) using the R scripts. The neutral community model was fitted by the nonlinear least-square fitting method, and the 95% confidence interval was predicted by the "Hmisc" package (Miller et al., 2016). In terms of the null model, Beta Taxon Index (βNTI) and Raup–Crick (RC_{Bray}) were calculated to represent phylogenetic and taxonomic diversity (Stegen et al., 2013; Zhou and Ning, 2017). $|\beta NTI| > 2$ indicates the dominance of deterministic processes, while $|\beta NTI| < 2$ indicates the dominance of stochastic processes. $|\beta NTI| < 2$ and $RC_{Bray} < -0.95$ represent homogenizing dispersal. $|\beta NTI| < 2$ and $RC_{Bray} > 0.95$ represent dispersal limitation. $|\beta NTI| < 2$ and $|RC_{Bray}| < 0.95$ represent "undominated" assembly (mainly consists of weak selection, weak dispersal, diversification, and/or drift). $\beta NTI < -2$ represents homogeneous selection. $\beta NTI > 2$ represents variable selection.

[2]http://ieg4.rccc.ou.edu/mena

Accession Numbers

All raw sequences of Illumina sequencing in this study have been submitted to the NCBI Sequence Read Archive (SRA) database, and the BioProject accession numbers for this research were PRJNA638003 and PRJNA728746.

RESULTS

Relative Abundance and Taxonomic Compositions of Abundant and Rare Bacterial Taxa

In order to investigate the succession of abundant and rare bacterial taxa under different pyrene concentrations, we classified each OTU based on the selected cutoff and calculated the total relative abundance of abundant and rare bacterial taxa. Overall, abundant bacterial taxa (18–37 OTUs) accounted for only a small part of the bacterial community but represented 17.1%–43.8% of the soil microbial community abundance. However, 6,085–7,851 OTUs were attached to rare bacterial taxa, only accounting for 5.3%–8.5% of all sequences. With the increasing pyrene concentrations in soils, the proportion of abundant bacterial taxa increased, while the relative abundance of rare bacterial taxa decreased (**Figures 1A,B**; **Supplementary Figure 1**). Compared with polluted soils containing pyrene, the abundant bacterial taxa in unpolluted soil (Control) were significantly lower regardless of pyrene concentrations. In contrast, the rare bacterial taxa in the Treatment Control showed an opposite trend, significantly higher than polluted soils. The majority (94.6%–100%) of abundant bacterial taxa in polluted soils were from unpolluted soils, with about 21.6% from initially abundant bacterial taxa and 8.04% from initially rare bacterial taxa. However, a large part (63.3%–82.4%) of rare bacterial taxa in polluted soils was not detected in the original unpolluted soils (**Supplementary Table 1**).

To better figure out the microorganisms' representative for abundant and rare bacterial taxa under different levels of pyrene stresses, we investigated the taxonomic compositions of the microbial community. The abundant bacterial taxa were comprised of nine phyla, dominated by Proteobacteria and Firmicutes, whereas up to 41 different phyla constituted the rare bacterial taxa (**Figures 1C,D**). Proteobacteria accounted for a high proportion in both abundant (28.5%–60.1%) and rare bacterial taxa (7.2%–35.6%). Interestingly, the relative abundance of Proteobacteria in abundant bacterial taxa increased with pyrene concentration, while a completely opposite trend was shown in rare bacterial taxa. As the pyrene concentration increased, the relative proportion of Firmicutes (19.1%–35.4%) in abundant bacterial taxa was on a downward path. Across abundant bacterial taxa, Chloroflexi and Deinococcus-Thermus were not involved in unpolluted soil but present in polluted soils (1.1%–3.8% and 2.0%–6.0% respectively), suggesting that the presence of pyrene activates specific bacteria taxa. Certain phyla occurred only in the rare bacterial taxa, such as Planctomycetes and Gemmatimonadetes. Compared with Control, the relative abundances of GAL15,

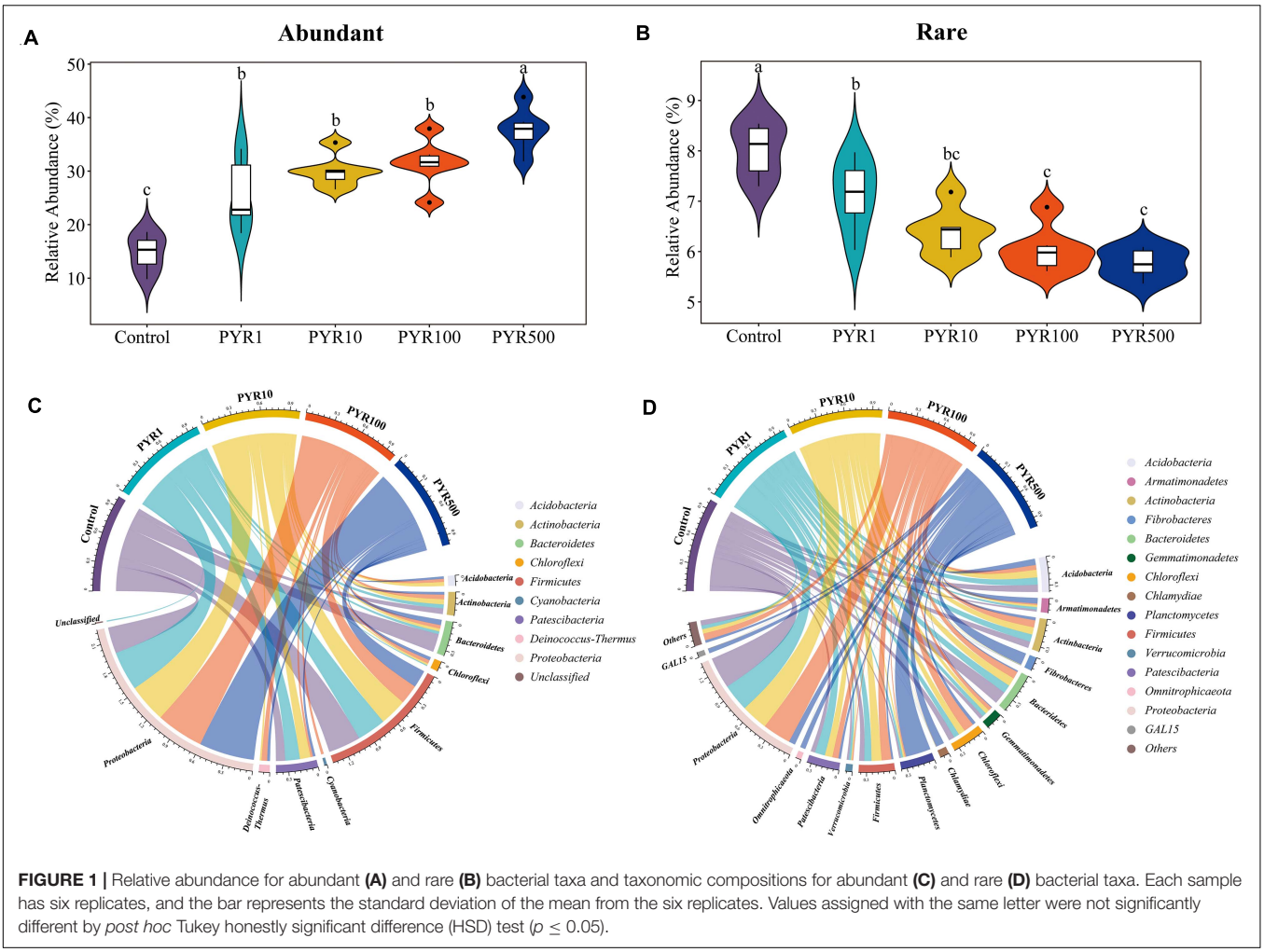


FIGURE 1 | Relative abundance for abundant **(A)** and rare **(B)** bacterial taxa and taxonomic compositions for abundant **(C)** and rare **(D)** bacterial taxa. Each sample has six replicates, and the bar represents the standard deviation of the mean from the six replicates. Values assigned with the same letter were not significantly different by *post hoc* Tukey honestly significant difference (HSD) test ($p \leq 0.05$).

Omnitrophicaeta, Planctomycetes, and Fibrobacteria in the rare subcommunity were significantly increased in PYR500. In particular, as for PYR500, the rare bacterial taxa were dominated by Planctomycetes (28.3%) and exhibited a higher relative abundance than other treatments (1.5%–2.4%).

Diversities and Functions of Abundant and Rare Bacterial Taxa Under Different Pollution Concentrations

Pollutants significantly changed the α-diversity of abundant and rare bacterial taxa (**Figure 2A**; **Supplementary Figure 2**). Compared with unpolluted soils, the Shannon–Wiener index of abundant bacterial taxa (2.16–3.96) in polluted soil increased, whereas rare bacterial taxa (7.07–7.58) decreased (**Figure 2A**; **Supplementary Figure 3**; ANOVA, $p < 0.05$). Moreover, we observed a significant influence of the pollution concentrations on both abundant and rare subcommunity structure based on Bray–Curtis dissimilarity (**Supplementary Figure 4**; ANOSIM, $R = 0.899$, $p = 0.001$ and $R = 0.975$, $p = 0.001$, respectively), and particularly, rare bacterial taxa were more sensitive than abundant bacterial taxa.

To further illustrate the differences between abundant and rare bacterial taxa under different pollution concentrations, community functions were analyzed by Tax4Fun based on KEGG Ortholog database. As for the six mainly functional categories, "Metabolism" always accounted for more than 60% across all samples in both abundant and rare bacterial taxa, followed by "genetic information processing" (17.4%–20.6%) (**Supplementary Figure 5A**). Rare bacterial taxa had five more functions than abundant bacterial taxa, mainly related to cellular processes (**Supplementary Figure 5B**). Interestingly, compared with Control, abundance of "Metabolism" showed an increasing trend along with increasing pollution concentrations no matter in abundant or rare bacterial taxa. However, there was a downturn in "genetic information processing" of polluted soils in both abundant and rare bacterial taxa (**Figure 2B**). Furthermore, based on LEfSe analysis, many functions significantly differed among treatments in both abundant and rare bacterial taxa with the LDA score (log 10) > 3.0 (**Supplementary Figure 6**). As for the abundant bacterial taxa, functions affiliated to the cellular processes were more abundant in unpolluted soils, while metabolism (related to degradation) showed a more important position in polluted soils. Moreover, functions related to bacterial

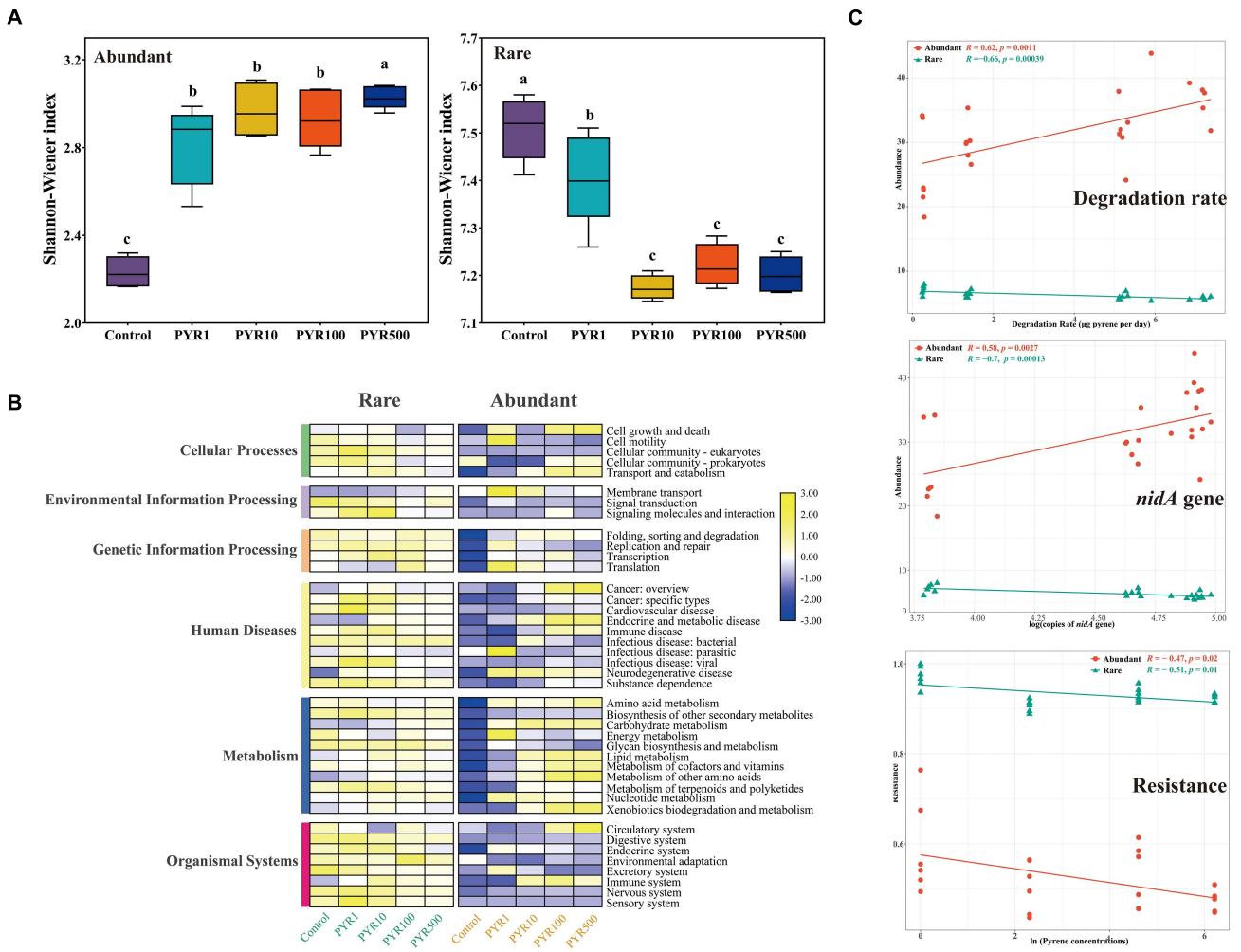


FIGURE 2 | Changes of the diversities and functions in abundant and rare subcommunities under different levels of pyrene stress. **(A)** The Shannon–Wiener index and **(B)** heatmap of potential community functions based on Tax4Fun in abundant and rare subcommunities. **(C)** The correlation analysis between the abundance of abundant and rare taxa and pyrene degradation rate or logarithm of *nidA* gene copies; the correlation analysis between resistance of abundant and rare taxa and natural logarithm of pyrene stress concentrations.

membrane transport, such as ABC transporters, were more prevalent in the rare bacterial taxa under pyrene stresses.

In order to further verify the metabolism functions of abundant and rare bacterial taxa, correlation analysis between the abundance of abundant and rare taxa and pyrene degradation rates under different levels of pyrene stresses was performed (**Figure 2C**). Obviously, the degradation rate was significantly positively correlated with the abundance of abundant taxa (Spearman R = 0.62, p = 0.0011), whereas it was negatively correlated with the rare (R = −0.66, p = 0.00039). Interestingly, the same conclusion was found in the correlation between the copies of *nidA* gene (pyrene dioxygenase gene) and the abundance of abundant (Spearman R = 0.58, p = 0.0027) and rare (Spearman R = −0.7, p = 0.00013) taxa. When focusing on the resistance ability of subcommunities under stress environments, we found that the resistance of rare subcommunity (0.8886–0.9995) was much higher than that of abundant subcommunity (0.4481–0.7638), and both were negatively correlated with pyrene stress concentrations (rare: Spearman R = −0.51, p = 0.01; abundant: Spearman R = −0.47, p = 0.02).

Co-occurrence Patterns of Abundant and Rare Subcommunities

To better examine the interaction between microorganisms with different abundances in the community, a subcommunity co-occurrence network was constructed at the OTU level based on the Spearman's correlation relationships. The co-occurrence network exhibited a scale-free character (**Supplementary Figure 7**, R square power-law: 0.917), indicating a nonrandom structure. In the whole network, there were about 14, 79, and 290 OTU nodes of abundant, moderate, and rare bacterial taxa, respectively. Most abundant nodes tended to have edges with rare nodes. Two important node-level topological features of

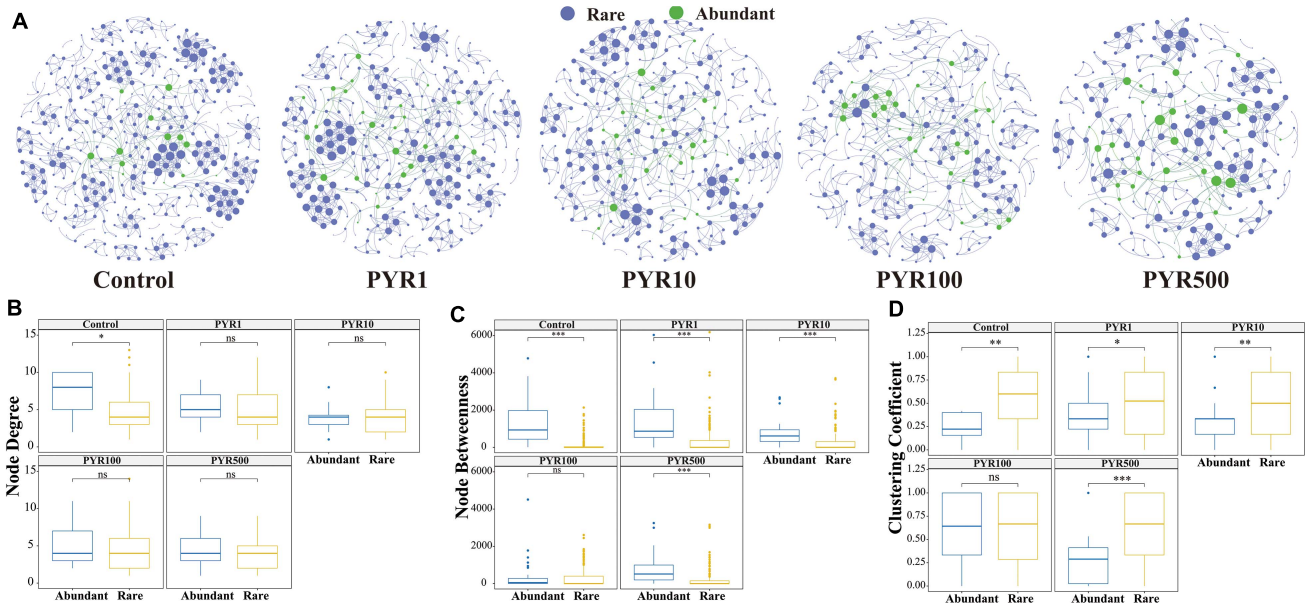


FIGURE 3 | Network analysis **(A)** and topological relations **(B–D)** between abundant and rare taxa under different levels of pyrene stress. The blue and green nodes represent abundant and rare operational taxonomic units (OTUs), respectively. All the networks were visualized by Gephi. The size of each node is proportional to its number of connections. Node degree **(B)**, betweenness **(C)**, and clustering coefficient **(D)** of abundant and rare networks under different levels of polycyclic aromatic hydrocarbon (PAH) stress. Asterisks indicate significance: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001 based on Tukey honestly significant difference (HSD) test.

different subcommunities including degree and betweenness were performed to further resolve the differences. The values of node degree and betweenness were significantly higher for rare and moderate taxa than abundant bacterial taxa (Tukey HSD, $p < 0.001$ and $p < 0.01$ respectively), but there was no significant difference between rare and moderate taxa in both topological features (Tukey HSD, $p > 0.05$).

In order to further reveal the direct relationship between abundant and rare bacterial taxa, co-occurrence networks were built through linking abundant OTUs to rare OTUs under different pollution concentrations. With the increase of pollutant concentrations, the co-occurrence networks tended to be simplified, among which nodes and links showed a decreasing trend gradually (**Figure 3**). In particular, more nodes related to abundant bacterial taxa showed high levels in PYR500 compared with rare bacterial taxa. We also calculated three important node-level topological features of different subcommunities including degree, betweenness, and clustering coefficient (**Figure 3**). Importantly, we only found that the node degree of abundant bacterial taxa was significantly higher than that of rare bacterial taxa in unpolluted soils. Besides PYR100, other treatments of betweenness in abundant bacterial taxa were significantly higher than those in rare bacterial taxa ($p < 0.05$), and nodes with high betweenness had more control over the network, indicating that more information may be passed through abundant bacterial taxa. However, except PYR100, the clustering coefficient in abundant bacterial taxa was significantly lower than that in rare, suggesting that the adjacency points of rare bacterial taxa had a higher interconnection degree.

Ecological Assembly Processes of the Abundant and Rare Subcommunities

To examine the adaptive capacity of microorganisms in specific PAH-polluted environments, we calculated the average niche width of the community. Notably, along the increasing PAH concentrations in soil, the average niche width (8.83–11.05) of the community showed a trend of gradual increase (**Figure 4A**; ANOVA, $p < 0.05$). When the pyrene concentrations exceeded 10 mg kg^{-1} soils, the bacterial community had a significantly broader niche width than Control and PYR1 (Tukey HSD, $p < 0.05$). In addition, we performed the niche width of subcommunities to further refine the response of abundant and rare bacterial taxa to pyrene stresses. There was a significant difference across the niche width of three subcommunities in all treatments, and the rare subcommunity stayed at the narrowest niche width (**Figure 4B**). Interestingly, the niche width of abundant subcommunity was significantly broader than that of the rare subcommunity (Tukey HSD, $p < 0.05$), whereas it was narrower than that of the moderate subcommunity (Tukey HSD, $p > 0.05$). Moreover, we also discovered the consistent trend for the threshold values of subcommunities to respond to pyrene stresses in soils using threshold indicator taxa analysis (TITAN2), which is calculated based on the sum z scores for each taxon in subcommunities (**Figure 4C**). That is, the abundant subcommunity exhibited a significantly broader range of threshold for pyrene stresses than the rare subcommunity (Tukey HSD, $p < 0.05$).

To address the effect of pyrene stresses on community and subcommunity assembly, we performed Sloan neutral

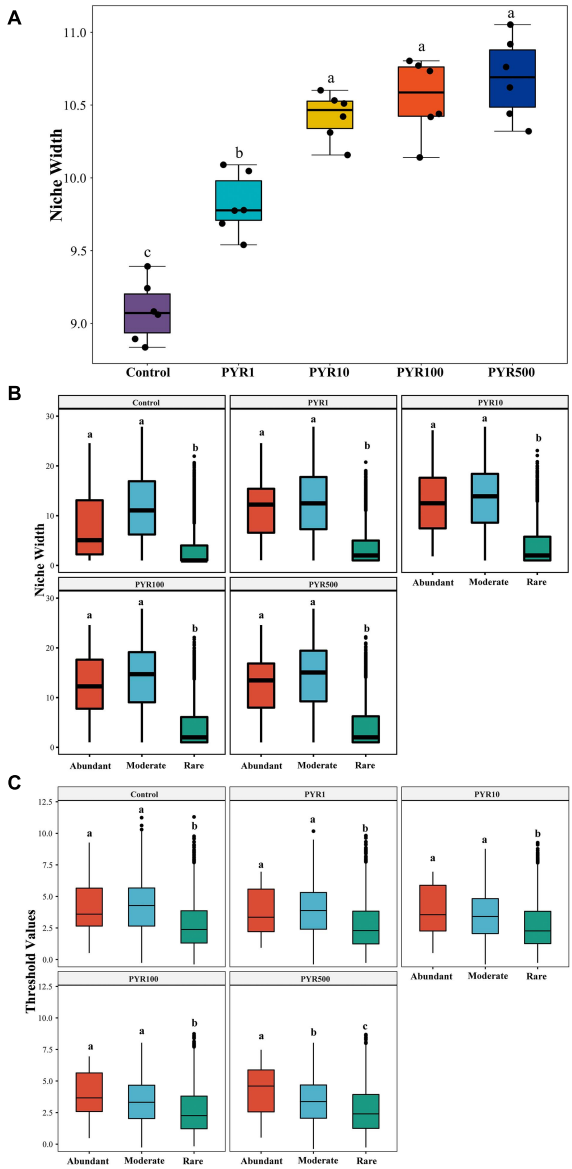


FIGURE 4 | The environmental adaptability of the entire bacterial community **(A)** and subcommunities **(B,C)** under different levels of pyrene stress. The niche width of the **(A)** entire community and **(B)** subcommunities, **(C)** the environmental threshold to pyrene stresses of abundant and rare subcommunities. Values assigned with the same letter were not significantly different by *post hoc* Tukey honestly significant difference (HSD) test ($p \leq 0.05$).

community model to determine the relative importance of the neutral processes. Along the increasing pyrene concentrations, the goodness of fit (R^2 = 0.599–0.657) of the neutral community model displayed a downward trend (**Figure 5**). We discovered that the neutral community model accounted for a large part of the community variance (84.4%–87.6%) between the occurrence frequency of OTUs and their mean relative abundance. There was a gradually decreasing trend in the Nm-value with the increased levels of PAHs pollution (Nm = 27,597–24,320). Almost all of the abundant bacterial taxa were distributed in the 95% confidence interval of the predicted neutral community model, whereas majority of the rare bacterial taxa were above the predicted occurrence frequency.

To better illustrate the differences of assembly processes between the abundant and rare subcommunities, the null model was performed based on the βNTI. For the entire community, the βNTI values in Control were mainly between −2 and 2 but gradually below −2 along the pyrene concentrations (**Supplementary Figure 9A**). We found that the maximum mean βNTI values for abundant and rare subcommunities were both in PYR500 (**Supplementary Figure 8**). Majority of the βNTI values for the abundant subcommunity dominated between −2 and +2 in all treatments, whereas the distributions of βNTI gradually shifted with increasing pyrene concentration in the rare subcommunity from stochastic community assembly ($-2 < \beta NTI < +2$) to deterministic community assembly ($|\beta NTI| > 2$). Given the Raup–Crick distance based on taxonomic dissimilarity index, along the pyrene concentration, the stochastic assembly (dispersal limitation, 86.7%–93.3%) occupied a large proportion in the abundant subcommunity (**Figure 6**). However, the trend in the fraction of dispersal limitation was a decrease with pyrene concentration (from 73.3% at Control to 40% at PYR500) in the rare subcommunity. Importantly, in polluted soils (except PYR10), deterministic community assembly (heterogeneous selection and homogeneous selection) contributed more variation than stochastic community assembly (dispersal limitation). Similarly, the proportion of deterministic community assembly (heterogeneous selection and homogeneous selection) increased with the pyrene concentrations (**Supplementary Figure 9B**).

Changes of Different Soils in Abundant and Rare Taxa Under Pyrene Stresses

To get more evidence, we performed microcosm incubation using four different soils and analyzed their abundant and rare taxa. Similarly, the relative abundance of abundant taxa in PYR500 (28.8%–76.7%) were significantly higher than those in Control (10.1%–49.2%) across all soil samples ($p < 0.001$). The rare taxa in PYR500 showed a significantly lower relative abundance than in Control (**Supplementary Figure 10A**; $p < 0.01$). Importantly, the pyrene dioxygenase gene copies (*nidA*) were also significantly positively correlated with abundance of abundant taxa (**Supplementary Figure 10B**; Spearman $R = 0.69$, $p = 0.013$), further suggesting that the abundant taxa played an important role in pyrene degradation. Moreover, the rare taxa presented a significantly higher resistance index compared with the abundant taxa (**Supplementary Figure 10C**; $p < 0.001$), implying that the rare taxa are critical for maintaining community stability.

DISCUSSION

Previous studies have revealed the impact of PAHs on microbial communities (Lors et al., 2010; Niepceron et al., 2013; Crampon et al., 2018; Li et al., 2020), but few have reported the succession

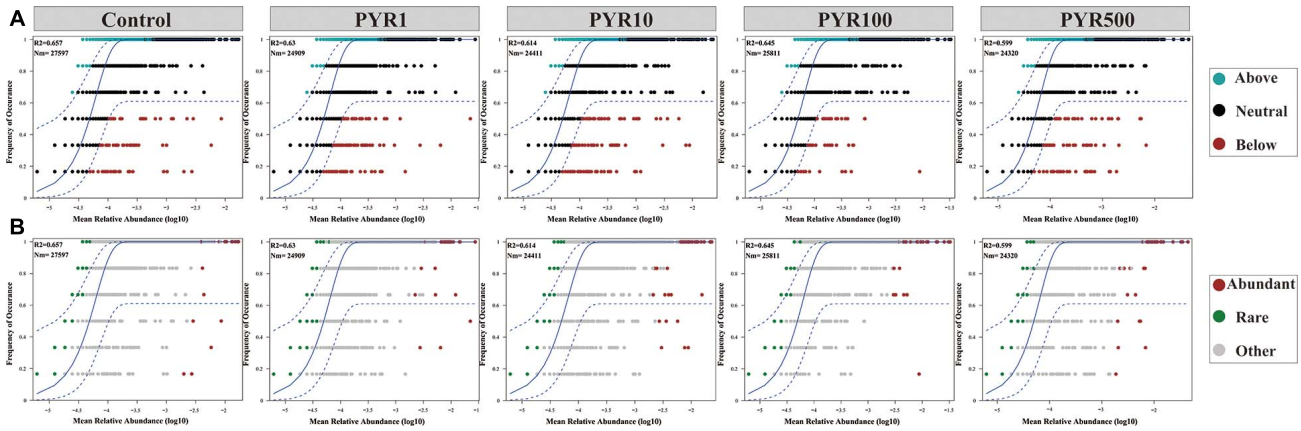


FIGURE 5 | Neutral model (relative abundance–frequency relationships) of the entire bacterial community **(A)** and the abundant and rare taxa **(B)** under different levels of pyrene stresses. The dashed blue line represents the 95% confidence interval above and below the prediction (the solid blue line). R2 indicates the coefficient of the neutral fit, and Nm indicates the metacommunity size times immigration.

patterns and community assembly mechanisms of abundant and rare bacterial subcommunities under various pyrene stresses. In this study, microcosm was constructed to explore the strategies of soil abundant and rare bacterial taxa in response to diverse pyrene concentrations. Pyrene stresses increased the proportion and diversity of abundant bacterial taxa, whereas rare bacterial taxa presented an opposite trend. Compared with the rare bacterial taxa, abundant bacterial taxa had better adaptive capacity and broader niche width to PAH-polluted environment. The abundant taxa are more likely to degrade pollutants, while the rare taxa contribute more to community resistance. We also reveal that distinct community assembly processes drove the abundant and rare subcommunities under different levels of pyrene stresses (**Figure 7**).

Broader Environmental Adaptations of Abundant Bacterial Taxa in Response to Pyrene Stresses

Increasing evidence has demonstrated the appreciable impact of PAHs on soil ecological functions (Wang et al., 2017; Liu et al., 2020). Numerous studies focused on the effects of PAH pollution in the natural environment on microbial community, but the dose–effects of pollutants are easily overlooked. Besides, although it has been well accepted that soils contaminated by PAHs have distinctive bacterial diversity and taxonomic compositions (Crampon et al., 2018; Ahmad et al., 2019), much less attention has been paid to the succession patterns (such as diversity, distribution, function) of abundant and rare bacterial taxa under different levels of pyrene stresses, not to mention their environmental adaptation.

Here in our study, results showed that there were markedly different succession patterns between the abundant and rare bacterial taxa under different levels of pyrene stresses. In terms of α diversity, the Shannon–Wiener index of rare bacterial taxa decreased with the increased pyrene concentrations, while the abundant bacterial taxa presented an opposite trend (**Figure 2A**; **Supplementary Figure 3**). Interestingly, the α diversity of the rare bacterial taxa was higher than the abundant bacterial taxa in all treatments (**Supplementary Figures 2**, **3**), suggesting that the rare bacterial taxa made more contributions to the composition of overall community (Xue et al., 2018; Du et al., 2020). Due to their high diversity, the rare bacterial taxa could increase the functional redundancy of the community, further providing wider ecological buffering space to strive against a changing environment (Hausmann et al., 2016; Jiao et al., 2017; Kurm et al., 2019).

As to the distribution, the abundant and rare bacterial taxa significantly separated along the pyrene concentration gradients, the abundant bacterial taxa had a gradual increase in the proportion (**Figure 1**). The new abundant bacterial taxa in polluted soils mainly came from the rare and moderate bacterial taxa in unpolluted soil, primarily belonging to Alphaproteobacteria, Gammaproteobacteria (Proteobacteria), and Bacilli (Firmicutes) (**Supplementary Table 1**). A wave of recent studies suggested that the rare bacterial taxa could act as the "seed bank," which plays an insurance effect on the whole community (Rocca et al., 2020). Our results supported prior studies reporting that the rare bacterial taxa have an opportunity to be activated to maintain the stability of the bacterial community under pollution stresses (Jousset et al., 2017). Intriguingly, four OTUs (OTU_2, OTU_27546, OTU_3, and OTU_8) belong to the rare species in uncontaminated soil before and turned into abundant under pyrene stresses (**Supplementary Table 1**). OTU_2 and OTU_27546 are aligned to Burkholderiaceae, while OTU_3 and OTU_8 belong to Paenibacillaceae and Bacillaceae, which all are known as typical PAH-degrading bacteria and carried several diverse ring-cleaving dioxygenase genes (e.g., 1,2 dioxygenase and 3,4-dioxygenase) (Ghosal et al., 2016; Li et al., 2018; Morya et al., 2020). However, most of the rare bacterial taxa in polluted soils were absent from the original unpolluted soil and have some unique phyla, such as Planctomycetes and Gemmatimonadetes, indicating that the rare bacterial

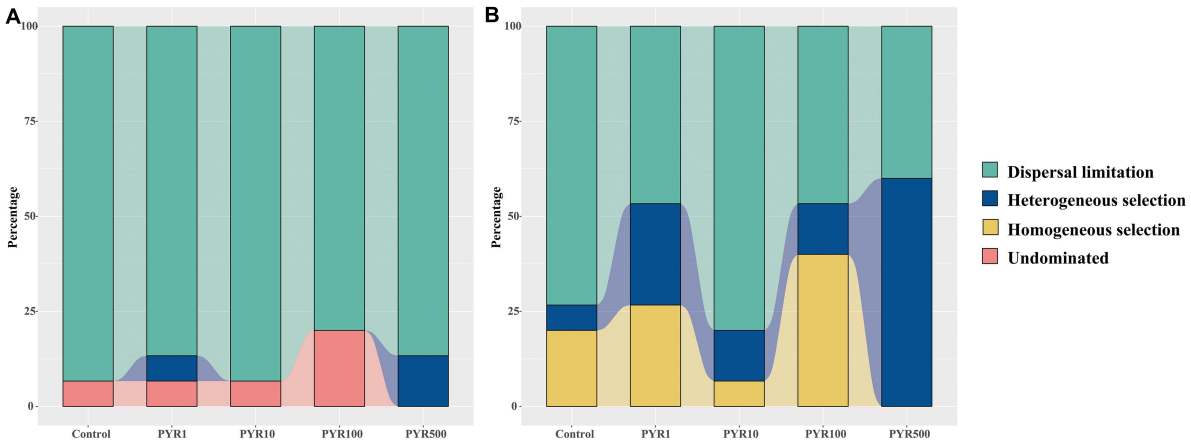


FIGURE 6 | The fraction of assembly mechanism in abundant **(A)** and rare **(B)** subcommunities based on the null model.

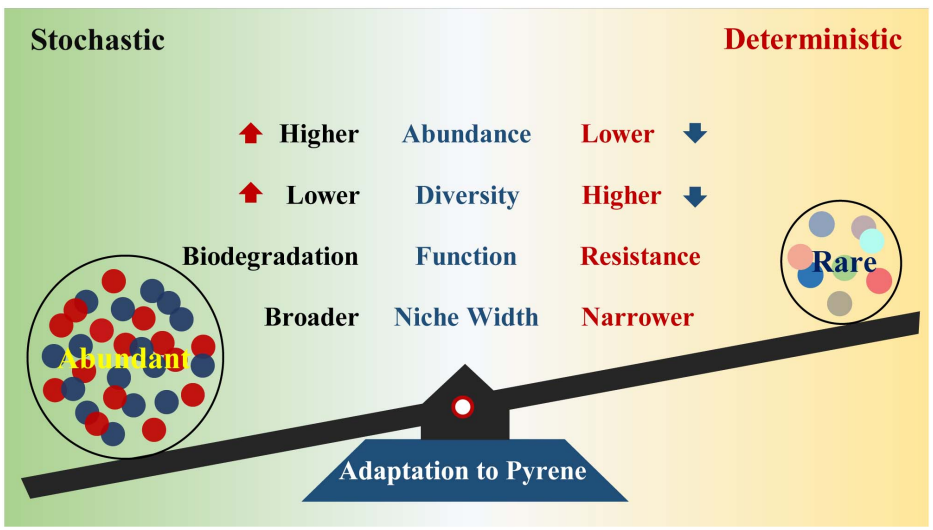


FIGURE 7 | Functions of the abundant and rare bacterial taxa to adapt to pyrene stresses.

taxa have environmental specificity under PAH selection (Kurm et al., 2019).

Historically, several studies have shown the occurrence of microbial adaptation in the environment after exposure to a xenobiotic chemical (Itrich et al., 2015; Poursat et al., 2019). PAH stresses altered the dynamic balance of abundant and rare bacterial subcommunities, further influencing environmental adaptation (Xue et al., 2018; Jiao and Lu, 2020b). Here, we discovered that the abundant bacterial taxa have broader adaptability to pollution stresses from two different aspects. First, the abundant bacterial taxa possess broader niche width than the rare, which was consistent with previous research (**Figure 4B**; Du et al., 2020; Jiao and Lu, 2020a). The niche width is an index of biodiversity of biological utilization resources, which means that the abundant bacterial taxa have the ability to make efficient use of various resources and survive in a more diverse environment compared with the rare bacterial taxa (Godoy et al., 2018). A prior study has demonstrated that many relatively abundant soil bacterial phylotypes could be found across a wide range of soils, suggesting the stronger environmental adaptability of abundant bacterial taxa (Delgado-Baquerizo et al., 2018). Moreover, the abundant bacterial taxa presented a higher response threshold to pyrene stresses, which proved that they can have a relatively durable survivability in polluted environment from another perspective (**Figure 4C**; Baker and King, 2010). Second, according to the correlation-based network analysis, we found that majority of the abundant bacterial taxa tend to be connected with the rare bacterial taxa, inferring that the two subcommunities were simultaneously affected by pyrene stresses (**Figure 3**). This is also demonstrated by the simplification of the network structures between abundant and rare bacterial taxa along the pyrene gradient. The topology of network could reflect the interaction among taxa, for example, node degree can indicate the links, and the node betweenness represents the impact on

connected nodes (Feng et al., 2017). Our results showed that the node degree and betweenness of the abundant subcommunity were higher than those of the rare subcommunity, indicating that the abundant bacterial taxa were inclined to occupy the core and central position in adaptation to the PAH stresses (Du et al., 2020).

Previous studies have shown that microorganisms can adapt to the environment by regulating their functions (Roller et al., 2013; Tikariha and Purohit, 2019). Based on the predicted subcommunity functions, the abundant bacterial taxa showed gradually enhanced metabolic functions along the PAH gradient and higher than that of the rare bacterial taxa. Importantly, functions related to xenobiotics biodegradation and metabolism (such as PAHs) and metabolism-specific simple carbohydrates (such as lipid, amino acids, and vitamins) presented higher relative abundance in the abundant subcommunities than the rare subcommunities (**Figure 2B**). This is similar to a previous work that concluded that high abundance taxa in polychlorinated biphenyl (PCB)-contaminated soils were relevant to PCB degradation (Xu et al., 2020). The enhanced xenobiotics metabolic capacity in abundant bacterial taxa was conducive to survival in different levels of pyrene stresses. Both pyrene degradation rate and *nidA* gene copies (pyrene dioxygenase gene) were significantly positively correlated with abundance of abundant taxa, negatively relative to the rare (**Figure 2C**), further proving abundant taxa rather than the rare taxa played an important role in PAH degradation. On the other hand, functions affiliated with membrane transport, particularly the ABC transporters, which are important for tolerance to many different kinds of pollutants, were more prevalent in the rare communities (**Figure 2B**). Meanwhile, the rare subcommunity displayed higher resistance than the abundant under different levels of pyrene stresses (**Figure 2C**). Together, these results indicated that the abundant bacterial taxa may play an important role in the degradation of pollutants, while the rare bacterial taxa play a key role in improving community tolerance (Jiao et al., 2017).

Distinct Assembly Mechanisms of Abundant and Rare Subcommunities Under Pyrene Stresses

The assembly process of microbial community can inevitably affect the diversity and composition of soil microbiome, thus influencing the functions of soil microecosystem (Leibold et al., 2017). Therefore, a deeper knowledge of assembly mechanisms of abundant and rare subcommunities under pyrene stresses will lead to a better understanding of the adaptability of bacterial abundance to environmental disturbances. The neutral community model is a prediction model based on neutral theory, which is an effective method to infer whether the stochastic process is dominant in the community assembly (Sloan et al., 2010). Our results clearly showed that almost all the abundant bacterial taxa were present in the predicted neutral region, while most of the rare bacterial taxa were above the predicted neutral region (**Figure 5**). The neutral community model could not explain 100% of the variation in the microbial community, suggesting that there may be other community assembly mechanisms that lead to non-neutral distribution (Dini-Andreote et al., 2015; Hou et al., 2020). We further performed the null model to elucidate the assembly mechanism of abundant and rare subcommunities (Stegen et al., 2013; Wang et al., 2021). In terms of unpolluted soils, stochastic processes (mainly dispersal limitation, more than 70% and 90%, respectively) participated in shaping both abundant and rare subcommunity assemblies (**Figure 6**). As for the abundant subcommunity, dispersal limitation was always dominant, independent of the PAH concentrations. However, the prominent role of deterministic processes in shaping the rare subcommunity assembly emerged along the pyrene stresses, except PYR10. As expected, due to their essentially low relative abundance and narrow niche width, the rare bacterial taxa are more sensitive to environmental filtering and less competitive than the abundant bacterial taxa (Lynch and Neufeld, 2015; Liang et al., 2020; Rocca et al., 2020). Importantly, the abundant bacterial taxa with a broader niche width are likely to make competitive use of various resources and adapt well to specific ecosystems through active growth and high abundance (Li et al., 2019; Jiao and Lu, 2020a). This indicates that PAH-induced taxa assembly has a great influence on the composition of bacterial subcommunities. As for the stochastic processes dominated in rare PYR10, we consider it as the turning point of steady state, with uncertainty, because the Shannon–Wiener index drops off a cliff in rare bacterial taxa of PYR10, whereas there was no similar phenomenon in abundant bacterial taxa. Considering the results of the neutral model and null model, we revealed that the stochastic process and deterministic process, respectively, dominate the assembly of abundant and rare bacterial subcommunities in PAH-polluted soils.

In summary, our study provides a better understanding of succession patterns and subcommunity assembly processes underlying the abundant and rare bacterial taxa under different levels of pyrene stresses and reveals the importance of the abundant bacterial taxa on the maintenance of community stability and adaptation to harsh environments. The abundance, diversity, and metabolism of specific carbohydrates in the abundant bacterial taxa rose across the increasing pyrene concentrations. Higher abundance and broader niche width are beneficial for the abundant bacterial taxa to cope with pyrene stresses. The rare bacterial taxa with higher phylogenetic diversity serve as a "seed bank" and play a crucial role in improving community stress resistance. Stochastic processes were dominant in driving the assembly of the abundant subcommunity, whereas the relative importance of deterministic processes progressively increased with pyrene stresses. The results based on the abundant and rare bacterial taxa may conduce to broaden our horizon about understanding the assembly and maintenance of bacterial diversity and function responses to pollution stresses.

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 in the article/ **Supplementary material**.

AUTHOR CONTRIBUTIONS

SWu designed the research. XZ, SWu, and ZB supervised the project. YDo, HF, XL, and SWa refined the experimental method. YDo performed the research and conducted the data analyses with assistance from SWu and YDe. YDo drafted the manuscript. XZ, SWu, and YDe reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Key R&D Program of China (No. 2019YFC1805803), the National Natural Science Foundation of China (Nos. 41907273, 91951108, and 31670507), the CAS International Partnership Program (No. 121311KYSB20200017), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA23010400), and the Science and Technology Service Network Initiative of the Chinese Academy of Sciences (Grant KFJ-STS-ZDTP-064).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.689762/full#supplementary-material

Supplementary Figures 1–10 | Correlation between natural logarithm of pyrene concentrations and richness of abundant and rare bacterial taxa. Simpson index of abundant and rare bacterial taxa under different levels of pyrene stresses. Correlation between natural logarithm of pyrene concentrations and Shannon-Wiener index of abundant and rare bacterial taxa. nMDS was calculated based on Bray-Curtis dissimilarity and 95% confidence ellipse was added for each treatment. Comparison of potential functions in abundant and rare subcommunities among treatments based on Tax4Fun. The linear discriminant analysis (LDA) value distribution histogram of functions in abundant and rare subcommunities. Network and topological relation in the whole communities. NTI of abundant and rare bacterial taxa under different levels of pyrene stresses based on the null model. The NTI value and fraction of assembly mechanism in entire community based on the null model. Changes of different soils in abundant and rare taxa under pyrene stresses.

Supplementary Table 1 | Information of abundant OTUs in treatments.

REFERENCES

Ahmad, M., Yang, Q., Zhang, Y., Ling, J., Sajjad, W., Qi, S., et al. (2019). The distinct response of phenanthrene enriched bacterial consortia to different PAHs and their degradation potential: a mangrove sediment microcosm study. *J. Hazardous Mater.* 380:120863. doi: 10.1016/j.jhazmat.2019.120863

Ashauer, K. P., Wemheuer, B., Daniel, R., and Meinicke, P. (2015). Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* 31, 2882–2884. doi: 10.1093/bioinformatics/btv287

Baker, M. E., and King, R. S. (2010). A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods Ecol. Evol.* 1, 25–37. doi: 10.1111/j.2041-210X.2009.00007.x

Brinch, U. C., Ekelund, F., and Jacobsen, C. S. (2002). Method for spiking soil samples with organic compounds. *Appl. Environ. Microbiol.* 68, 1808–1816. doi: 10.1128/AEM.68.4.1808-1816.2002

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, 884–890. doi: 10.1093/bioinformatics/bty560

Crampon, M., Bodilis, J., and Portet-Koltalo, F. (2018). Linking initial soil bacterial diversity and polycyclic aromatic hydrocarbons (PAHs) degradation potential. *J. Hazardous Mater.* 359, 500–509. doi: 10.1016/j.jhazmat.2018.07.088

Delgado-Baquerizo, M., Oliverio, A. M., Brewer, T. E., Benavent-Gonzalez, A., Eldridge, D. J., Bardgett, R. D., et al. (2018). A global atlas of the dominant bacteria found in soil. *Science* 359, 320–325. doi: 10.1126/science.aap9516

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., and Zhou, J. (2012). Molecular ecological network analyses. *BMC Bioinform.* 13:113. doi: 10.1186/1471-2105-13-113

Deng, Y., Zhang, P., Qin, Y., Tu, Q., Yang, Y., He, Z., et al. (2016). Network succession reveals the importance of competition in response to emulsified vegetable oil amendment for uranium bioremediation. *Environ. Microbiol.* 18, 205–218. doi: 10.1111/1462-2920.12981

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., and Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. U.S.U.* 112, E1326–E1332. doi: 10.1073/pnas.1414261112

Du, S., Dini-Andreote, F., Zhang, N., Liang, C., Yao, Z., Zhang, H., et al. (2020). Divergent co-occurrence patterns and assembly processes structure the abundant and rare bacterial communities in a salt marsh ecosystem. *Appl. Environ. Microbiol.* 86:e00322-20. doi: 10.1128/aem.00522-20

Dua, M., Singh, A., Sethunathan, N., and Johri, A. (2002). Biotechnology and bioremediation: successes and limitations. *Appl. Microbiol. Biotechnol.* 59, 143–152. doi: 10.1007/s00253-002-1024-6

Feng, K., Zhang, Z., Cai, W., Liu, W., Xu, M., Yin, H., et al. (2017). Biodiversity and species competition regulate the resilience of microbial biofilm community. *Mol. Ecol.* 26, 6170–6182. doi: 10.1111/mec.14356

Gao, P., da Silva, E. B., Townsend, T., Liu, X., and Ma, L. Q. (2019). Emerging PAHs in urban soils: concentrations, bioaccessibility, and spatial distribution. *Sci. Total Environ.* 670, 800–805. doi: 10.1016/j.scitotenv.2019.03.247

Gavrilescu, M., Demnerová, K. I, Aamand, J., Agathos, S., and Fava, F. (2015). Emerging pollutants in the environment: present and future challenges in biomonitoring, ecological risks and bioremediation. *New Biotechnol.* 32, 147–156. doi: 10.1016/j.nbt.2014.01.001

Ghosal, D., Ghosh, S., Dutta, T. K., and Ahn, Y. (2016). Current state of knowledge in microbial degradation of polycyclic aromatic hydrocarbons (PAHs): a review. *Front. Microbiol.* 7:1369. doi: 10.3389/fmicb.2016.01369

Godoy, O., Bartomeus, I., Rohr, R. P., and Saavedra, S. (2018). Towards the integration of niche and network theories. *Trends Ecol. Evol.* 33, 287–300. doi: 10.1016/j.tree.2018.01.007

Guittar, J., Shade, A., and Litchman, E. (2019). Trait-based community assembly and succession of the infant gut microbiome. *Nat. Commun.* 10:512. doi: 10.1038/s41467-019-08377-w

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* 10, 497–506. doi: 10.1038/nrmicro2795

Hausmann, B., Knorr, K.-H., Schreck, K., Tringe, S. G., del Rio, T. G., Loy, A., et al. (2016). Consortia of low-abundance bacteria drive sulfate reduction-dependent degradation of fermentation products in peat soil microcosms. *ISME J.* 10, 2365–2375. doi: 10.1038/ismej.2016.42

Hou, J., Wu, L., Liu, W., Ge, Y., Mu, T., Zhou, T., et al. (2020). Biogeography and diversity patterns of abundant and rare bacterial communities in rice paddy soils across China. *Sci. Total Environ.* 730:139116. doi: 10.1016/j.scitotenv.2020.139116

Itrich, N. R., McDonough, K. M., van Ginkel, C. G., Bisinger, E. C., LePage, J. N., Schaefer, E. C., et al. (2015). Widespread microbial adaptation to L-Glutamate-N,N-diacetate (L-GLDA) following its market introduction in a consumer cleaning product. *Environ. Sci. Technol.* 49, 13314–13321. doi: 10.1021/acs.est.5b03649

Ji, M., Kong, W., Stegen, J., Yue, L., Wang, F., Dong, X., et al. (2020). Distinct assembly mechanisms underlie similar biogeographical patterns of rare and abundant bacteria in tibetan plateau grassland soils. *Environ. Microbiol.* 22, 2261–2272. doi: 10.1111/1462-2920.14993

Jia, X., Dini-Andreote, F., and Salles, J. F. (2018). Community assembly processes of the microbial rare biosphere. *Trends Microbiol.* 26, 738–747. doi: 10.1016/j.tim.2018.02.011

Jiao, S., Chen, W., and Wei, G. (2017). Biogeography and ecological diversity patterns of rare and abundant bacteria in oil-contaminated soils. *Mol. Ecol.* 26, 5305–5317. doi: 10.1111/mec.14218

Jiao, S., and Lu, Y. (2020a). Abundant fungi adapt to broader environmental gradients than rare fungi in agricultural fields. *Global Change Biol.* 26, 4506–4520. doi: 10.1111/gcb.15130

Jiao, S., and Lu, Y. (2020b). Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ. Microbiol.* 22, 1052–1065. doi: 10.1111/1462-2920.14815

Jiao, S., Wang, J., Wei, G., Chen, W., and Lu, Y. (2019). Dominant role of abundant rather than rare bacterial taxa in maintaining agro-soil microbiomes under environmental disturbances. *Chemosphere* 235, 248–259. doi: 10.1016/j.chemosphere.2019.06.174

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., et al. (2017). Where less may be more: how the rare biosphere pulls ecosystems strings. *ISME J.* 11, 853–862. doi: 10.1038/ismej.2016.174

Keith, L. H. (2015). The source of US EPA's sixteen PAH priority pollutants. *Polycyclic Aromatic Compounds* 35, 147–160. doi: 10.1080/10406638.2014.892886

Kurm, V., Geisen, S., and Hol, W. H. G. (2019). A low proportion of rare bacterial taxa responds to abiotic changes compared with dominant taxa. *Environ. Microbiol.* 21, 750–758. doi: 10.1111/1462-2920.14492

Lee, S.-H., Sorensen, J. W., Grady, K. L., Tobin, T. C., and Shade, A. (2017). Divergent extremes but convergent recovery of bacterial and archaeal soil communities to an ongoing subterranean coal mine fire. *ISME J.* 11, 1447–1459. doi: 10.1038/ismej.2017.1

Leibold, M. A., Chase, J. M., and Ernest, S. K. M. (2017). Community assembly and the functioning of ecosystems: how metacommunity processes alter ecosystems attributes. *Ecology* 98, 909–919. doi: 10.1002/ecy.1697

Levins, R. (1968). *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton, NJ: Princeton University Press.

Li, J., Luo, C., Zhang, D., Song, M., Cai, X., Jiang, L., et al. (2018). Autochthonous bioaugmentation-modified bacterial diversity of phenanthrene degraders in PAH-contaminated wastewater as revealed by DNA-stable isotope probing. *Environ. Sci. Technol.* 52, 2934–2944. doi: 10.1021/acs.est.7b05646

Li, L., Lin, Q., Li, X., Li, T., He, X., Li, D., et al. (2019). Dynamics and potential roles of abundant and rare subcommunities in the bioremediation of cadmium-contaminated paddy soil by *Pseudomonas* chenduensis. *Appl. Microbiol. Biotechnol.* 103, 8203–8214. doi: 10.1007/s00253-019-10059-y

Li, X., Song, Y., Bian, Y., Gu, C., Yang, X., Wang, F., et al. (2020). Insights into the mechanisms underlying efficient Rhizodegradation of PAHs in biochar-amended soil: from microbial communities to soil metabolomics. *Environ. Int.* 144:105995. doi: 10.1016/j.envint.2020.105995

Liang, Y., Xiao, X., Nuccio, E. E., Yuan, M., Zhang, N., Xue, K., et al. (2020). Differentiation strategies of soil rare and abundant microbial taxa in response to changing climatic regimes. *Environ. Microbiol.* 22, 1327–1340. doi: 10.1111/1462-2920.14945

Liao, J., Cao, X., Wang, J., Zhao, L., Sun, J., Jiang, D., et al. (2017). Similar community assembly mechanisms underlie similar biogeography of rare and abundant bacteria in lakes on Yungui Plateau, China. *Limnol. Oceanogr.* 62, 723–735. doi: 10.1002/lno.10455

Liu, W., Wang, D., Wang, Y., Zeng, X., Ni, L., Tao, Y., et al. (2020). Improved comprehensive ecological risk assessment method and sensitivity analysis of polycyclic aromatic hydrocarbons (PAHs). *Environ. Res.* 187:109500. doi: 10.1016/j.envres.2020.109500

Lors, C., Ryngaert, A., Perie, F., Diels, L., and Damidot, D. (2010). Evolution of bacterial community during bioremediation of PAHs in a coal tar contaminated soil. *Chemosphere* 81, 1263–1271. doi: 10.1016/j.chemosphere.2010.09.021

Lynch, M. D. J., and Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13, 217–229. doi: 10.1038/nrmicro3400

Ma, J., and Cao, Z. (2010). Quantifying the perturbations of persistent organic pollutants induced by climate change. *Environ. Sci. Technol.* 44, 8567–8573. doi: 10.1021/es101771g

Magoc, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Maliszewska-Kordybach, B. (1996). Polycyclic aromatic hydrocarbons in agricultural soils in poland: preliminary proposals for criteria to evaluate the level of soil contamination. *Appl. Geochem.* 11, 121–127. doi: 10.1016/0883-2927(95)00076-3

Miller, C. A., Mcmichael, J., Dang, H. X., Maher, C. A., Ding, L., Ley, T. J., et al. (2016). Visualizing tumor evolution with the fishplot package for R. *BMC Genom.* 17:880. doi: 10.1186/s12864-016-3195-z

Mo, Y., Zhang, W., Wilkinson, D. M., Yu, Z., Xiao, P., and Yang, J. (2021). Biogeography and co-occurrence patterns of bacterial generalists and specialists in three subtropical marine bays. *Limnol. Oceanogr.* 66, 793–806. doi: 10.1002/lno.11643

Morya, R., Salvachua, D., and Thakur, I. S. (2020). Burkholderia: an untapoed but promising bacterial genus for the conversion of aromatic compounds. *Trends Biotechnol.* 38, 963–975. doi: 10.1016/j.tibtech.2020.02.008

Niepceron, M., Martin-Laurent, F., Crampon, M., Portet-Koltalo, F., Akpa-Vinceslas, M., Legras, M., et al. (2013). Gamma proteobacteria as a potential bioindicator of a multiple contamination by polycyclic aromatic hydrocarbons (PAHs) in agricultural soils. *Environ. Pollut.* 180, 199–205. doi: 10.1016/j.envpol.2013.05.040

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'hara, R., et al. (2013). *Community Ecology Package. R Package Version, 2.0-2*.

Poursat, B. A. J., van Spanning, R. J. M., de Voogt, P., and Parsons, J. R. (2019). Implications of microbial adaptation for the assessment of environmental persistence of chemicals. *Crit. Rev. Environ. Sci. Technol.* 49, 2220–2255. doi: 10.1080/10643389.2019.1607687

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl. Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Ren, G., Teng, Y., Ren, W., Dai, S., and Li, Z. (2016). Pyrene dissipation potential varies with soil type and associated bacterial community changes. *Soil Biol. Biochem.* 103, 71–85. doi: 10.1016/j.soilbio.2016.08.007

Rocca, J. D., Simonin, M., Bernhardt, E. S., Washburne, A. D., and Wright, J. P. (2020). Rare microbial taxa emerge when communities collide: freshwater and marine microbiome responses to experimental mixing. *Ecology* 101:e02956. doi: 10.1002/ecy.2956

Roller, M., Lucic, V., Nagy, I., Perica, T., and Vlahovicek, K. (2013). Environmental shaping of codon usage and functional adaptation across microbial communities. *Nucl. Acids Res.* 41, 8842–8852. doi: 10.1093/nar/gkt673

Rombolà, A. G., Meredith, W., Snape, C. E., Baronti, S., Genesio, L., Vaccari, F. P., et al. (2015). Fate of soil organic carbon and polycyclic aromatic hydrocarbons in a vineyard soil treated with biochar. *Environ. Sci. Technol.* 49, 11037–11044. doi: 10.1021/acs.est.5b02562

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12:R60. doi: 10.1186/gb-2011-12-6-r60

Shahsavari, E., Schwarz, A., Aburto-Medina, A., and Ball, A. S. (2019). Biological degradation of polycyclic aromatic compounds (PAHs) in soil: a current perspective. *Curr. Pollut. Rep.* 5, 84–92. doi: 10.1007/s40726-019-00113-8

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2010). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.* 8, 732–740. doi: 10.1111/j.1462-2920.2005.00956.x

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., and Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Tikariha, H., and Purohit, H. J. (2019). Different dimensions in microbial community adaptation and function. *Indian J. Microbiol.* 59, 387–390. doi: 10.1007/s12088-019-00813-1

Wang, J., Wang, J., Zhao, Z., Chen, J., Lu, H., Liu, G., et al. (2017). PAHs accelerate the propagation of antibiotic resistance genes in coastal water microbial community. *Environ. Pollut.* 231, 1145–1152. doi: 10.1016/j.envpol.2017.07.067

Wang, J., Wang, Y., Li, M., Xu, L., He, N., Yan, P., et al. (2021). Differential response of abundant and rare bacterial subcommunities to abiotic and biotic gradients across temperate deserts. *Sci. Total Environ.* 763, 142942–142942. doi: 10.1016/j.scitotenv.2020.142942

Xu, Y., Teng, Y., Wang, X., Li, R., and Christie, P. (2020). Exploring bacterial community structure and function associated with polychlorinated biphenyl biodegradation in two hydrogen-amended soils. *Sci. Total Environ.* 745:140839. doi: 10.1016/j.scitotenv.2020.140839

Xue, M., Guo, Z., Gu, X., Gao, H., Weng, S., Zhou, J., et al. (2020). Rare rather than abundant microbial communities drive the effects of long-term greenhouse cultivation on ecosystem functions in subtropical agricultural soils. *Sci. Total Environ.* 706:136004. doi: 10.1016/j.scitotenv.2019.136004

Xue, Y., Chen, H., Yang, J. R., Liu, M., Huang, B., and Yang, J. (2018). Distinct patterns and processes of abundant and rare eukaryotic plankton communities following a reservoir cyanobacterial bloom. *ISME J.* 12, 2263–2277. doi: 10.1038/s41396-018-0159-0

Zhang, H., Hou, F., Xie, W., Wang, K., Zhou, X., Zhang, D., et al. (2020). Interaction and assembly processes of abundant and rare microbial communities during a diatom bloom process. *Environ. Microbiol.* 22, 1707–1719. doi: 10.1111/1462-2920.14820

Zhang, X., Zhang, R., Gao, J., Wang, X., Fan, F., Ma, X., et al. (2017). Thirty-one years of rice-rice-green manure rotations shape the rhizosphere microbial community and enrich beneficial bacteria. *Soil Biol. Biochem.* 104, 208–217. doi: 10.1016/j.soilbio.2016.10.023

Zhao, L., Liu, Y.-W., Li, N., Fan, X.-Y., and Li, X. (2020). Response of bacterial regrowth, abundant and rare bacteria and potential pathogens to secondary chlorination in secondary water supply system. *Sci. Total Environ.* 719:137499. doi: 10.1016/j.scitotenv.2020.137499

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev. Mmbr.* 81, e00002–e00017. doi: 10.1128/MMBR.00002-17

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 20 August 2021
doi: 10.3389/fmicb.2021.699939

Evaluating the Assembly Dynamics in the Human Vaginal Microbiomes With Niche-Neutral Hybrid Modeling

Zhanshan (Sam) Ma[1,2]*

[1] *Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China,* [2] *Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China*

Using 2,733 longitudinal vaginal microbiome samples (representing local microbial communities) from 79 individuals (representing meta-communities) in the states of healthy, BV (bacterial vaginosis) and pregnancy, we assess and interpret the relative importance of stochastic forces (e.g., stochastic drifts in bacteria demography, and stochastic dispersal) vs. deterministic selection (e.g., host genome, and host physiology) in shaping the dynamics of human vaginal microbiome (HVM) diversity by an integrated analysis with multi-site neutral (MSN) and niche-neutral hybrid (NNH) modeling. It was found that, when the traditional "default" *P*-value = 0.05 was specified, the neutral drifts were predominant (≥50% metacommunities indistinguishable from the MSN prediction), while the niche differentiations were moderate (<20% from the NNH prediction). The study also analyzed two challenging uncertainties in testing the neutral and/or niche-neutral hybrid models, i.e., lack of full model specificity – non-unique fittings of same datasets to multiple models with potentially different mechanistic assumptions – and lack of definite rules for setting the *P*-value thresholds (also noted as P_t-value when referring to the threshold of *P*-value in this article) in testing null hypothesis (model). Indeed, the two uncertainties can be interdependent, which further complicates the statistical inferences. To deal with the uncertainties, the MSN/NNH test results under a series of *P*-values ranged from 0.05 to 0.95 were presented. Furthermore, the influence of *P*-value threshold-setting on the model specificity, and the effects of woman's health status on the neutrality level of HVM were examined. It was found that with the increase of *P*-value threshold from 0.05 to 0.95, the overlap (non-unique) fitting of MSN and NNH decreased from 29.1 to 1.3%, whereas the specificity (uniquely fitted to data) of MSN model was kept between 55.7 and 82.3%. Also with the rising *P*-value threshold, the difference between healthy and BV groups become significant. These findings suggested that traditional single *P*-value threshold (such as the *de facto* standard *P*-value = 0.05) might be insufficient for testing the neutral and/or niche neutral hybrid models.

Keywords: unified neutral theory of biodiversity and biogeography, multi-site neutral model, niche-Neutral hybrid model, human vaginal microbiome, bacterial vaginosis, hierarchical Dirichlet process

OPEN ACCESS

Edited by:
Daliang Ning,
University of Oklahoma, United States

Reviewed by:
Xiaoqian Yu,
University of Vienna, Austria
Hui Li,
Institute of Applied Ecology (CAS), China

***Correspondence:**
Zhanshan (Sam) Ma
ma@vandals.uidaho.edu

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *24 April 2021*
Accepted: *15 July 2021*
Published: *20 August 2021*

Citation:
Ma ZS (2021) Evaluating the Assembly Dynamics in the Human Vaginal Microbiomes With Niche-Neutral Hybrid Modeling. Front. Microbiol. 12:699939. doi: 10.3389/fmicb.2021.699939

INTRODUCTION

The relationship between human vaginal microbiome (HVM) and women's health has been investigated since the 1980s, when clinical microbiologists had postulated that the diversity and possibly stability of vaginal microbiome are involved in the occurrence/recurrence of bacterial vaginosis (BV) (e.g., Sobel, 1999; Fredricks et al., 2005; Fredricks, 2011; Ma et al., 2012). Those studies are among the earliest ecological approaches to diseases now often referred to as the human MADs (microbiome-associated diseases) with an ever more rapidly growing list including BV, IBD (inflammatory bowel disease), periodontitis, cystic fibrosis (CF), psoriasis and many others (Lynch and Pedersen, 2016; Knight et al., 2017; Young, 2017; Gilbert et al., 2018). The metagenomics technique and the launch of the human microbiome project (HMP) and MetaHIT (metagenome of human intestinal tract) have revolutionized the investigation of the human microbiome and associated diseases during the last decade or so. Nevertheless, many questions in the field are still open and new more complex questions are being raised. In the case of BV and vaginal microbiome, as described by Fredricks (2011) who borrowed Winston Churchill's words for a very different topic, "*BV remains a riddle, wrapped in a mystery, and inside an enigma*." A recent characterization "*that BV is not a single entity, but a syndrome linked to various community types that cause somewhat similar physiological symptoms*." by Ma et al. (2012) reflects the state-of-the-art understanding of BV etiology. Obviously, although the importance of vaginal microbiome ecology in BV etiology is repeatedly confirmed, the mechanistic relationship between BV and HVM is far from clear. A pair of questions of fundamental importance: what are the underlying mechanisms driving the dynamics of HVM and what are their implications to the occurrence/recurrence of BV, are still largely unanswered.

Addressing the question of community assembly and diversity maintenance, the essential ingredients of community structure and dynamics, has attracted extensive attention and also led to vigorous debate (Alonso et al., 2006; McGill et al., 2006; Chisholm and Pacala, 2010; Rosindell et al., 2011). Two leading and competing theories in this field have been the traditional niche theory with a history back to the 1910s (Grinnell, 1917; Hutchinson, 1957; Holt, 2009) and more recent neutral theory (Hubbell, 2001). Both theories were invented to explain a familiar phenomenon on the earth, which was described by Darwin (1859) in the last paragraph of his "*On the Origin of Species*" as "*It is interesting to contemplate a tangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us."* In modern ecological terminology, *entangled bank* is essentially the concept of ecological *community*. Darwin was wondering how diverse lives (species) could coexist and form a beautiful entangled bank, while his theory stipulated the universal struggle for life as a consequence to natural selection. The classic niche theory assumed that each species has its own niche in which its individuals are adapted to live and prosper, and the entangled bank consists of many different niches suitable for many different species. In terms of niche theory, deterministic traits a species possess or selective niche forces play critical roles in driving the assembly of an ecological community as well as the maintenance of diversity after the community is established.

In the late 1990s, Hubbell (2001) challenged the traditional niche view by proposing the unified neutral theory of biodiversity and biogeography (UNTB). Different from traditional niche theory, the UNTB was formulated as a probability distribution model, which can be fitted with the species abundance distribution data (the number of each species in a community), obtainable by sampling ecological communities, and rigorously tested statistically. The theory assumes that the individuals of all species in a community are demographically equivalent, but their birth/death rates are stochastic, which means birth-death, migration, and speciation are all random events. Consequently, random drift and dispersal play critical roles in driving community assembly and diversity maintenance. Some researchers argued that the concept of species equivalence is "flawed" given the existence of niche differences and competitive asymmetries among species. Nevertheless, the stochasticity in species demography (particularly of single-cell microbes) is also a biological reality and its role may not be ignored in many communities. In reality, both deterministic niche forces and stochastic neutral forces may be in effect in setting the rules of community assembly and diversity maintenance, and it may be the *hybrid* effects that shape the community dynamics. For this reason, in the last decade and so, several hybrid models that integrate neutral and niche effects have been developed (e.g., Tilman, 2004; Ofiteru et al., 2010; Stokes and Archer, 2010; Jeraldo et al., 2012; Pigolotti and Cencini, 2013; Tang and Zhou, 2013; Fisher and Mehta, 2014; Kalyuzhny et al., 2014a,b, 2015; Matthews and Whittaker, 2014; Noble and Fagan, 2015). As to the debates on the usefulness and validity of the UNTB, using an analogy, in modern statistics (especially in biostatistics), it has been widely recognized that many datasets do not follow the Gaussian distribution (the normal distribution); nevertheless, few statisticians would question the foundational role of the Gaussian distribution, not to mention its validity. Similarly, the merits and unique advantage of UNTB as a null model for testing the significance of stochastic drift and dispersal have been firmly established and widely applied in the community ecology of plants and animals.

In the present study, we use a pair of models, the first a multi-site neutral model (Harris et al., 2017) and the second, a niche-neutral hybrid model (Tang and Zhou, 2013), to evaluate the relative significance of neutral and niche effects in shaping the dynamics of HVM. We further investigate the difference in the neutral-niche *continuum* between BV patients and healthy women. Our approach is different from most existing applications of neutral or hybrid models in the following three aspects.

First, most existing neutral or niche-neutral hybrid models use spatially implicit/explicit community/metacommunity data, whereas we use longitudinal (time-series) sampling of the

community/metacommunity. In spatially explicit models, the metacommunity consists of multiple local communities, which are connected with each other through dispersal and migration. In temporal (time-series) models, the metacommunity consists of a series of "snapshots" of the same community at different time points, i.e., the time-series data obtained from sampling the vaginal microbiome of a subject at different time points in this study. Indeed, previously, Kalyuzhny et al. (2014a,b, 2015) used time-series data to perform dynamic analysis of the niches versus neutrality and they termed the analysis as a generalized neutral theory for explaining the static and dynamic properties of ecological communities. A reason we did not adopt their models is that the models we use in this study, as explained below, are truly multi-site mechanistically, which are mapped to the time-series points in our study.

Second, we use a *truly* multi-site neutral (MSN) model of UNTB, which was developed by Harris et al. (2017) to overcome the severe computational limitation of existing neutral theory models when the number of local communities is large and the migration rates among the local communities are different (Etienne, 2007, 2009a,b). The core technique Harris et al. (2017) developed was to approximate the multi-site UNTB model with the hierarchical Dirichlet process (HDP) and use an efficient Bayesian machine-learning algorithm. With their approach, fitting even the largest dataset can be performed in a reasonable amount of time. This important computational advance enables us to build a UNTB model for each subject by utilizing the time series sampling of her vaginal microbiome. This capability is of significant practical importance given the established connection between BV and the diversity of the vaginal microbiome, in particular, a long-standing puzzle in BV etiology – the rise of species diversity associated with BV (e.g., Sobel, 1999; Fredricks et al., 2005; Fredricks, 2011; Ma et al., 2012; Ma and Ellison, 2018, 2019).

Third, we also apply the niche-neutral hybrid (NNH) model by Tang and Zhou (2013) to further assess the neutral-niche hybrid effects in shaping the dynamics of HVM diversity. A major reason we prefer this hybrid model to other existing hybrid models (e.g., Tilman, 2004; Ofiteru et al., 2010; Stokes and Archer, 2010; Jeraldo et al., 2012; Pigolotti and Cencini, 2013; Tang and Zhou, 2013; Fisher and Mehta, 2014; Kalyuzhny et al., 2014a,b, 2015; Matthews and Whittaker, 2014; Noble and Fagan, 2015) is because both MSN and NNH use exactly the same data collection methods – either multi-site or multi-time-point sampling. The only essential difference between the neutral-niche hybrid model (NNH) and multi-site neutral model (MSN) is the assumption that niche differences exist among local communities in NNH, while the MSN assumes no niche differentiation. In a time-series setting, the NNH model can tell us whether deterministic forces (similar to habitat selection in a spatial setting) such as when changes in the host's physiology significantly influence the dynamics of vaginal microbiome diversity over time.

In summary, by building and testing the MSN and NNH models for each subject, we are able to evaluate the relative importance of stochastic forces (e.g., neutral dispersal, drift, and stochastic diversification) vs. deterministic forces (e.g., microbial interactions, host genome and physiology, menses, etc.) in shaping the dynamics of community diversity. Furthermore, if we treat BV or health status as part of the host physiology, testing the MSN/NNH models can reveal the impact of BV on the dynamics of the HVM diversity (assuming that diversity change is the consequence of BV), or reveal the diversity changes that induce BV (assuming that diversity change is the cause of BV). Regardless of the causal assumption, our approach offers a useful tool for evaluating the mechanisms (niche vs. neutral) of the dynamics of HVM diversity as well as the factors affecting the balances between different mechanisms. We demonstrate our approach (see **Figure 1**) by using the datasets (see **Table 1**) from three separate longitudinal studies on the HVM, including 79 subjects sampled at 2,733 time points (Gajer et al., 2012; Ravel et al., 2013; Romero et al., 2014).

MATERIALS AND METHODS

Human Vaginal Microbiome (HVM) Datasets and Analysis Strategy

Table 1 below listed the three published datasets (groups) of the HVM (human vaginal microbiome), which are reanalyzed in this study to perform the niche-neutral theoretic analysis. **Figure 1** is a diagram illustrating the background, objectives and the integrated niche-neutral approach to achieving the objectives of this study outlined previously. Two mathematical models: the multi-site neutral model (MSN) by Harris et al. (2017) and niche-neutral hybrid model (NNH) by Tang and Zhou (2013), are used to fit the same HVM datasets. Both the models are extensions or derived from Hubbell (2001) unified neutral theory of biodiversity and biogeography (UNTB). Brief description of the MSN and NNH models as well as their fittings is presented in **Table 1**. It is noted that these datasets are from three independent studies: minor difference in sequencing protocols may exist. For this reason, the samples from each of the 79 individuals are modeled independently. With the independent modeling, the influence from sequencing protocols should have minimized.

The Multi-Site Neutral Model (MSN) by Harris et al. (2017)

Hubbell's Unified Neutral Theory of Biodiversity (UNTB)

The UNTB conceptually distinguishes between metacommunity dynamics from local community dynamics coupled to metacommunity through migrations. The theory assumes that both the dynamics are driven by similar neutral processes, except that in metacommunity speciation, rather than migration are in operations (Hubbell, 2001, 2006). The neutral process or ecological equivalence between species implies that the demographic rates (birth/death) of all species are stochastic but equivalent on per capita basis (Harris et al., 2017). There are three key parameters (elements) with the UNTB, the immigration rate (I_i), which controls the coupling of a local community to the metacommunity. Another is the speciation rate, also known as the fundamental biodiversity number (θ), which can be interpreted as the rate at which new individuals are added to the

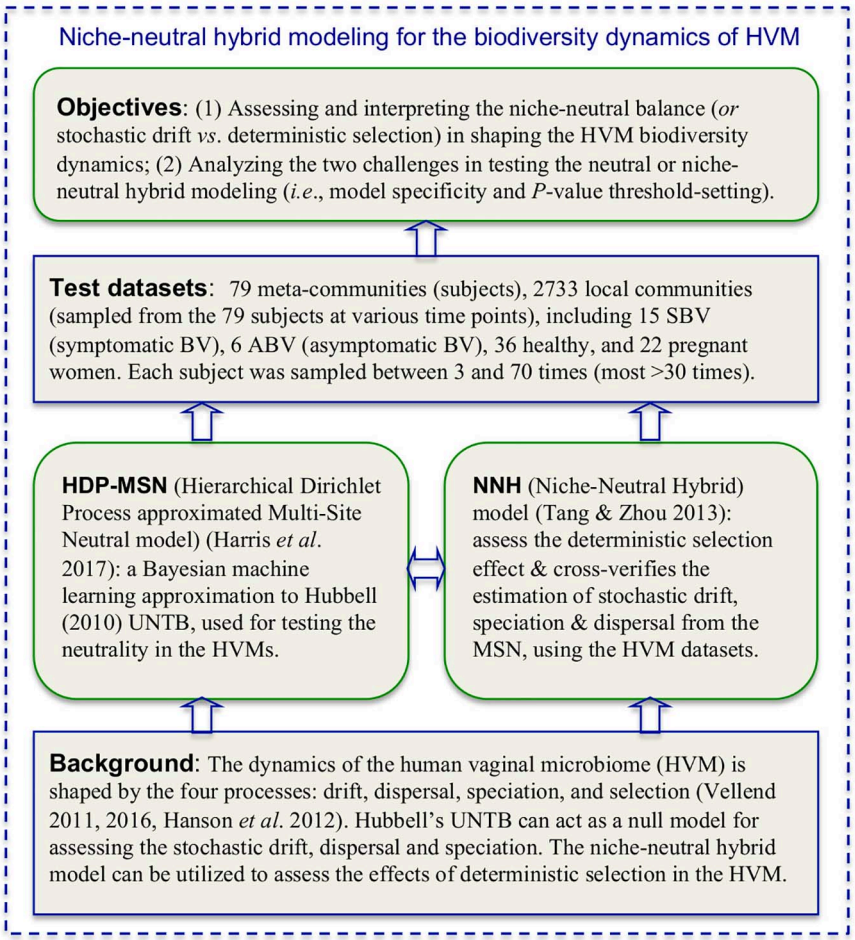


FIGURE 1 | A diagram illustrating the background, objectives and the integrated niche-neutral theoretic-approach to achieving the objectives of this study based on 2,733 longitudinal metagenomics (16S-rRNA) samples collected from 79 women (including BV patients).

TABLE 1 | The datasets of multi-site HVM (human vaginal microbiome) datasets utilized for testing the MSN (multi-site neutral) and NNH (niche-neutral hybrid) models.

Datasets	*N	**S	Sample description	Sources
ABV (Asymptomatic Bacterial Vaginosis)	6	66∼70	Ravel et al. (2013) sampled and DNA-pyrosequenced the vaginal microbiota of a cohort of 25 subjects over a 10-week period, consisting of 15 SBV, 6 ABV, and 4 healthy subjects (HEA-1). Total 16S-rRNA reads = 8,757,681, Average reads = 5285, The dataset is available from: https://doi.org/10.1186/2049-2618-1-2	Ravel et al. (2013) *Microbiome*
SBV (Symptomatic Bacterial Vaginosis)	15	59∼70		
HEA-1 (Healthy 1)	4	66∼69		
HEA-2 (Healthy 2): "32-healthy" cohort of HVMC study	32	25∼33	Gajer et al. (2012) sampled and DNA-pyrosequenced the vaginal microbiota of a cohort of 32 healthy individuals (HEA-2). Total 16S-rRNA reads = 2,522,080 from 937 samples; Average reads = 2692. The OTU table is available at: doi: 10.1126/scitranslmed.3003605	Gajer et al. (2012) *Science Translational Medicine*
PREG (Pregnancy)	22	3∼8	The vaginal microbiomes of a cohort of 22 normally pregnant women were sampled 6 times (for each individual) and DNA pyrosequenced. Total 16S-rRNA reads = 567,448; Average reads = 4082; The dataset is available at: https://doi.org/10.1186/2049-2618-2-4	Romero et al. (2014) *Microbiome*
Total or range	79	3∼70	A total of 79 meta-communities and 2,733 local communities (time-series samples) were sampled to conduct the tests.	

**N = the number of subjects (individuals) included in each dataset. **S = the approximate number of time points when samples of the HVMC were taken from each individual.*

metacommunity due to speciation. The third aspect of the UNTB is to assume that the SAD (species abundance distribution) of each community sample can be described by the multinomial (MN) distribution, formally:

$$\overline{X_i} \sim MN(N_i,\ \overline{\pi}_i) \tag{1}$$

where N_i is the size of i-th local community, $\overline{\pi}_i$ is a vector of the probability of observing a particularly species at i-th local community (Harris et al., 2017).

UNTB-HDP (Hierarchical Dirichlet Process) Limit to Metacommunities

A fully general case of fitting multiple sites (local communities) UNTB with potentially different immigration rates is computationally extremely challenging (actually intractable) even for small number of sites, and approximate algorithms must be utilized (Harris et al., 2017). Harris et al. (2017) developed an efficient Bayesian fitting framework by approximating the neutral models with the hierarchical Dirichlet process (HDP). The approximation was able to encapsulate the three essential elements of Hubbell (2001) UNTB, as stated above, but offers an efficient Bayesian fitting strategy for the multi-site UNTB.

Sloan et al. (2006, 2007) showed that for large local population sizes, assuming a fixed finite-dimensional metacommunity distribution with S species present, the local community distribution, $\overline{\pi}$, can be approximated by a *Dirichlet* distribution (Sloan et al., 2006, 2007). But it was Harris et al. (2017) who developed the general framework for approximating the UNTB computationally efficiently. Assuming there is a potentially infinite number of species that can be observed in the local community, then the stationary distribution of observing local population i is a Dirichlet process (DP), i.e.,

$$\overline{\pi_i}|\, I_i, \overline{\beta} \sim DP(I_i, \overline{\beta}) \tag{2}$$

where $\overline{\beta} = (\beta_1, ..., \beta_S)$ is the relative frequency of each species in the metacommunity, and I_i is the immigration rate.

At the metacommunity level, a Dirichlet process can still be used, but the base distribution is simply a uniform distribution over arbitrary species labels. The metacommunity distribution is then a purely stick breaking process, i.e.,

$$\overline{\beta} \sim Stick(\theta) \tag{3}$$

where θ is the fundamental biodiversity number. θ is a function of *speciation rate* (s) in the form of $\theta = (s/(1-s)(N-1)$, where N is the size of metacommunity (i.e., the fixed number of individuals in the metacommunity). The total number of species (S) in the metacommunity proportionally increases with θ. In addition, when θ increases, the SAD (species abundance distribution) is increasingly skewed to low abundance rare species (Harris et al., 2017). Note that speciation in the metacommunity is a counterpart of migration in a local community, except that the speciation is in operation on a longer timescale than migration. For this reason, both immigration rate (I_i) and biodiversity number θ have similar structure in their models. Specifically, $I_i = (m_i/(1-m_i)(N_i - 1)$, where m_i is the immigration probability to local community i, and N_i is the local community size. Obviously, when $I_i \to \propto$, the stationary distribution of local community should approach the metacommunity distribution since that means migration probability is equal to *1*, i.e., all members in the local community are immigrants. When $I_i \to\ 0$, local community can become dominated by a single species (Harris et al., 2017).

Given that both local community and metacommunity are approximated with Dirichlet processes, the problem can be formulated as a hierarchical Dirichlet process (HDP) (Teh et al., 2006; Harris et al., 2017). Alternatively, Dirichlet process (DP) can also be formulated as the so-called Chinese restaurant process, from which the Antoniak equation can be derived. The Antoniak equation represents the number species (S) observed following N draws from a Dirichlet process with biodiversity number θ, and is with the following form:

$$P(S\,|\theta, N) = s(N, S)\theta^S \frac{\Gamma(\theta)}{\Gamma(\theta + N)} \tag{4}$$

where $s(N, S)$ is the unsigned Stirling number of the first kind and $\Gamma(.)$ denotes the gamma function (Antoniak, 1974).

Gibbs Sampler (MCMC Algorithm) for the UNTB-HDP Model

The full UNTB-HDP model is obtained by combining previous equations (1–3) and also the distribution models of biodiversity number (θ) and immigration rate (I_i), both of which are assumed to follow Gamma distribution. Harris et al. (2017) developed an efficient Gibbs sampler for the UNTB-HDP approximation, which is a type of Bayesian Markov Chain Monte Carlo (MCMC) algorithm and can be summarized as the following four sampling steps, including sampling the biodiversity parameter, sampling the metacommunity distribution, sampling the immigration rate, and sampling the ancestral states. Harris et al. (2017) found through experiments that to ensure sampling was performed with the stationary distribution, 50,000 Gibb samples for each fitted dataset were necessary with the first 25,000 iterations removed as burn-in. The results are reported as the *median* values over the last 25,000 samples with upper and lower credible limits (Bayesian confidence) given by 2.5 and 97.5% quantiles of those samples.

Fitness Tests for the UNTB-HDP Multi-Site Neutral (MSN) Model

To determine whether an observed dataset fits the UNTB-HDP multi-site neutral (MSN) model (hereafter shortened as MSN model), Harris et al. (2017) proposed a similar Monte Carlo significance test to that used by Etienne (2007). Furthermore, Harris et al. (2017) also developed a procedure to test for the local neutral community assembly but with a fitted possibly non-neutral metacommunity because of the hierarchical nature of the MSN model. Specifically, with Harris et al. (2017) MSN model, two-level tests (local community and metacommunity levels) for neutrality can be performed. For both the tests, samples were generated from N = 2,500 sets of fitted MSN parameters, which were selected from every tenth iteration of the last 25,000 Gibbs samples (a total of 50,000 samples were simulated, and the first 25,000 samples were discarded as burn-in). N = 2,500

is chosen to compute the pseudo *P*-values for conducting the neutrality test (Harris et al., 2017). In addition, for each observed community sample, there is the *actual log-likelihood* L_0. Two additional parameters θ and M are particular worthy of mentioning: θ is the *median* of the *fundamental biodiversity par*ameters computed from 25,000 times of simulations, and M-value is the average of the medians of the *migration rates* of local communities in each metacommunity, also computed from 25,000 times of simulations.

To test the neutrality at the metacommunity level, P_M, which is "the proportion of the simulated neutral samples with their likelihoods *not* exceeding the observed data likelihood" (Harris et al., 2017). The computation of P_M is as follows: Assume L_M is the median of the log-likelihoods of the simulated neutral metacommunity samples, and N_M is the number of simulated neutral metacommunity samples, having their log-likelihoods satisfying $L \leq L_0$ (where L is the simulated likelihood and L_0 is the actual likelihood as mentioned previously), then the $P_M = N_{M/}N$ is a pseudo *P*-value for testing the neutrality at metacommunity level. If $P_M > 0.05$, the metacommunity appears to satisfy the MSN model, according to Harris et al. (2017).

To test the neutrality at the local community level, P_L, which is the proportion of the simulated locally neutral samples exceeding the observed data likelihood (Harris et al., 2017). It is computed as follows:

Assume L_L is the median of the log-likelihoods of the simulated local community samples, and N_L is the number of simulated local community samples, having their likelihoods not exceeding the L_0, then $P_L = N_{L/}N$, is the pseudo *P*-value for testing the neutrality at the local community level. If $P_L > 0.05$, the local community appears to satisfy the neutral model. Readers are referred to Harris et al. (2017) for the detailed algorithm and computational procedures (including the software in C language) for fitting the MSN model, which we used for analyzing HVM datasets in this study.

The Niche-Neutral Hybrid (NNH) Model by Tang and Zhou (2013)

Tang and Zhou (2013) proposed a hybrid niche-neutral model by revising Volkov et al. (2007) neutral model for multiple discrete communities. Volkov et al. (2007) assumed that the inter-species interactions in a steady-state community may be ignored, and all species in the community become functionally equivalent. They further assumed that birth and death probabilities of a species with n individuals are $b_n = b(n + \gamma)$ and $d_n = dn$, respectively, where b and d are the per-capita density-independent birth and death rates, and γ is a parameter for immigration. The migration was assumed to be species-independent, corresponding to immigration from a time-averaged metacommunity in a species-symmetric manner. This treatment of migration, in effect, ignored any immigration between local communities within the metacommunity, and also, the rates of immigration considered were small. By solving the master equation for the dynamics of a species, Volkov et al. (2007) obtained the probability that a species has n individuals, which follows the negative binomial distribution:

$$p(n) = \frac{(1-x)^{\gamma}}{\Gamma(\gamma)} \frac{x^n}{n!} \Gamma(n+\gamma) \tag{5}$$

where x is the ratio of the *per* capita birth to death rate (i.e., b/d, a measure of the lifetime reproductive success), and $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$, which is equal to $(z\text{-}1)!$ for integer z. They further obtained the mean number of species with abundance n:

$$< \varphi_n >= \theta \frac{x^n}{n!} \Gamma(n+\gamma) \tag{6}$$

where θ is the fundamental biodiversity parameter, and S is the number of observed species.

Tang and Zhou assumed that a semi-isolated local community consists of K non-overlapping niches. Within each niche, a number of species follow their own neutral rules independent of the other K-1 niches. By applying Volkov et al. (2007) neutral model for multiple discrete communities to a single niche of the community, Tang and Zhou (2013) derived the expected number of species with abundance n in niche i as:

$$< \varphi_{n,i} >= \theta_i \frac{x_i^n}{n!} \Gamma(n+\gamma_i) \tag{7}$$

where θ_i is the biodiversity parameter for niche i, x_i is the ratio of per capita birth to death rates of each species in niche i, and γ_i is a parameters for immigration of niche i. The total expected number of species with abundance n in the community consisting of K niches is represented by the following equation:

$$< \varphi_n; K >= \sum_{i=1}^{K} < \varphi_{n,i} > \tag{8}$$

Note that Eq. 8 is a summation of Eq. 7 across K niches, i.e., summing up all species with an abundance of n across all K niches. The following Chi-squared test statistic is utilized to determine the goodness-of-fitting for the niche-neutral hybrid model, i.e.,

$$\chi^2 = \sum_n \frac{(E_n - O_n)^2}{E_n} \tag{9}$$

where E_n is the expected number of species with abundance n, O_n is the observed number of species with abundance n.

To test the niche-neutral hybrid effects with Tang and Zhou (2013) NNH model, we computed the following items (listed in **Supplementary Table 2** of the online supplementary information (OSI) and partially in **Table 3**), including: the average number of individuals per niche (local community) in each metacommunity (J), the average species numbers per niche (local community) in each metacommunity (S), the average fundamental biodiversity parameter per niche (local community) in each metacommunity (θ), the average of the migration coefficients (m), the average of the birth to death ratio (x), the average of the migration rate (γ). To conduct the χ^2-test at the meta-community level, we computed χ^2-value [Eq. 9] and associated *P*-value. To test the neutrality at a local community level, Volkov et al. (2003, 2007) approach for fitting the relative species abundance (RSA)

distribution to their neutral model is adopted. Specifically, we computed and reported (see the last two columns in **Supplementary Table 2** and **Table 3**) the number and percentage of local communities (niches) that passed the local neutrality test.

The *P*-value of the Chi-squared test is then used to determine whether or not Tang and Zhou (2013) hybrid model is suitable for a series of microbial communities sampled from each individual. In the case of our time-series microbiome datasets, we treat each time point as a niche occupied by a local microbial community and fit the neutral model for each local community. Specifically, at the metacommunity level, if *P*-value > 0.05, then the metacommunity appears to satisfy the NNH, and the metacommunity assembly is co-driven by both niche and neutral processes, which also implies that the metacommunity itself does not satisfy the neutral theory, but within each niche, the local community is neutral. If *P*-value < 0.05, the metacommunity does not seem to satisfy the NNH, which also implies that within each niche, the local community is not neutral either, and the metacommunity assembly is solely influenced by the niche process. Readers are referred to Tang and Zhou (2013) for the detailed algorithm and computational procedures (including the software) for fitting the NNH model, which we used for analyzing HVM datasets in this study.

Model Specificity and *P*-Value Threshold Setting in Testing the Null Models

Two uncertainties have been well recognized in testing the neutral theory and niche-neutral hybrid models including the previous MSN and NNH models. One is the lack of full model specificity in fitting the neutral and/or niche-neutral hybrid models such as MSN/HHH models, and another is the lack of definite rules for setting the *P*-value thresholds in testing null models. What makes the statistical inferences more difficult is the potential interdependence between both uncertainties. There are no silver bullets to resolve them for various reasons including the complexity of the problem *per se* and limitations of the *P*-value setting in frequentist approaches to statistical inferences. In this article, no perfect solutions are offered, but we present two measures to relieve both issues. First, to evaluate the specificity of the MSN/NNH models, we classify the model-fittings as four possible categories: MSN-only, NNH only, both MSN & NNH, neither MSN nor MSN, and further observe the change of category proportions when *P*-value thresholds were specified differently. This allows us, at the minimum, to have an educated guess for the specificity of each model, particularly under different confidence levels (*P*-values). Second, besides testing the null models (MSN/NNH) under the traditional "default" *P*-value = 0.05, the results of model testing under a series of *P*-value thresholds are presented and analyzed. The variable *P*-value thresholds allow us to assess the goodness-of-fitting of the MSN/NNH models under various levels of confidence.

RESULTS

The Niche-Neutral Continuum in Shaping the HVM Dynamics Evaluated Under Traditional "Default" *P*-Value Threshold

Supplementary Table 1 in the OSI listed the *full* test results for the MSN with five HVM (human vaginal microbiome) datasets (groups) outlined in **Table 1**. **Table 2** below was excerpted from **Supplementary Table 1** to exhibit the results of 9 selected meta-communities. Similarly, **Table 3** below exhibited the results of

TABLE 2 | Fitting the HDP-MSN (hierarchical Dirichlet process approximated multi-site neutral model) (Harris et al., 2017) to the HVM (human vaginal microbiome) datasets for selected individuals, excerpted from **Supplementary Table 1** in the OSI*.

Datasets	Case No.	L_0	θ	M-Value	Meta-Community				Local Community			
					L_M	N_M	N	P_M	L_L	N_L	N	P_L
ABV	S12	−7164.578	17.818	11.849	−8855.097	2437	2500	0.975	−7387.665	2158	2500	0.863
SBV	S5	−5929.666	15.552	7.295	−7861.124	2469	2500	0.988	−6220.505	2339	2500	0.936
	S3	−1749.538	4.759	5.645	−1945.254	1715	2500	0.686	−1781.817	1672	2500	0.669
	S17	−10498.771	8.166	920.730	−6555.105	0	2500	0.000	−2900.912	0	2500	0.000
Healthy-1	S7	−7330.538	12.554	17.957	−10193.354	2475	2500	0.990	−7530.997	2096	2500	0.838
Healthy-2	#400	−1609.574	12.332	9.409	−3248.037	2500	2500	1.000	−1766.803	2330	2500	0.932
	#401	−3002.396	23.955	6.088	−4053.368	2495	2500	0.998	−3362.560	2492	2500	0.997
Pregnancy	N002	−325.419	12.392	2.623	−447.710	2384	2500	0.954	−403.615	2378	2500	0.951
	N003	−581.146	13.098	7.923	−687.729	2099	2500	0.840	−648.312	2244	2500	0.898

N = 2,500 is the number of Gibb samples selected from 25,000 simulated communities (i.e., every tenth iteration of the last 25,000 Gibbs samples, a total of 50,000 simulations were performed and with the first 25,000 discarded as burn-in), and the N is used to compute the pseudo P-value below for conducting the neutrality test. L_0 is the actual (observed) log-likelihood. θ is the median of biodiversity parameters computed from 25,000 times of simulations. M-value is the average medians of the migration rates of local communities in each metacommunity, also computed from 25,000 times of simulations. L_M is the median of the log-likelihoods of the simulated neutral metacommunity samples; and N_M is the number of simulated neutral metacommunity samples with their likelihoods satisfying the $L < L_0$ (L is the simulated likelihood and L_0 is the actually observed likelihood), $P_M = N_M/N$ is the pseudo P-value for testing the neutrality at metacommunity level; if $P_M > 0.05$, the metacommunity is indistinguishable from the prediction of the MSN model. L_L is the median of the log-likelihoods of the simulated local community samples, and N_L is the number of simulated local community samples with their likelihoods not exceeding the L_0. $P_L = N_L/N$, is the pseudo P-value for testing the neutrality at the local community level; if $P_L > 0.05$, the local community is indistinguishable from the neutral model. See **Figure 2 for an example of successfully fitting to the MSN model.*

TABLE 3 | Fitting the NNH (niche-neutral hybrid) model (Tang and Zhou, 2013) to the HVM (human vaginal microbiome) datasets for selected individuals, excerpted from **Supplementary Table 2** in the OSI*.

Datasets	ID	*J*	*S*	θ	*m*	*x*	γ	R^2	χ^2	*P*-value	N^{pass}	$\%^{(pass)}$
ABV	S12	5261.723	23.043	9277.776	0.000	0.691	0.488	0.996	52.253	0.000	22	46.8
SBV	S5	4647.896	18.042	466.265	0.001	0.643	0.557	0.983	154.291	0.000	18	37.5
	S3	164.000	8.455	4.032	0.007	0.803	1.236	0.981	0.925	0.996	11	100.0
	S17	5852.400	23.940	571.207	0.000	0.724	0.531	0.986	74.385	0.000	33	66.0
Healthy-1	S7	5902.193	20.421	650.047	0.000	0.687	0.462	0.986	87.923	0.000	13	37.1
Healthy-2	#400	2737.111	14.444	6.379	0.001	0.688	1.730	0.944	23.091	0.027	3	33.3
	#401	2614.200	22.933	7.292	0.000	0.760	1.242	0.983	12.745	0.310	13	86.7
Pregnancy	N002	4278.000	14.500	24.256	0.000	0.541	0.890	0.974	209327	0.000	1	50.0
	N003	4215.500	22.500	3.432	0.000	0.762	1.973	0.851	5.358	0.913	4	100.0

J: the average number of individuals per niche (local community) in each metacommunity, S: the average species numbers per niche (local community) in each metacommunity, θ: the average fundamental biodiversity parameter per niche (local community) in each metacommunity, m: the average of the migration coefficients, x: the average of the birth to death ratio, γ: the average of the migration rate, R^2: the goodness-of-fitting index, χ^2-value: the χ^2-value of chi-squared test for observed value against predicted value, P-value for the χ^2-test; when P-value > 0.05, the metacommunity satisfies the NNH model. The last two columns are the number and percentage of local communities (niches) that passed the local neutrality test. Note that R^2 = 1 resulted from approximation with four effective digits only (e.g., 0.99995, exact 1 is nearly impossible to achieve). See* ***Figure 3* *for an example of successfully fitting to the NNH model. Note that we use the P-value (FDR) after the FDR control was imposed to determine the outcome of testing the NNH model.*

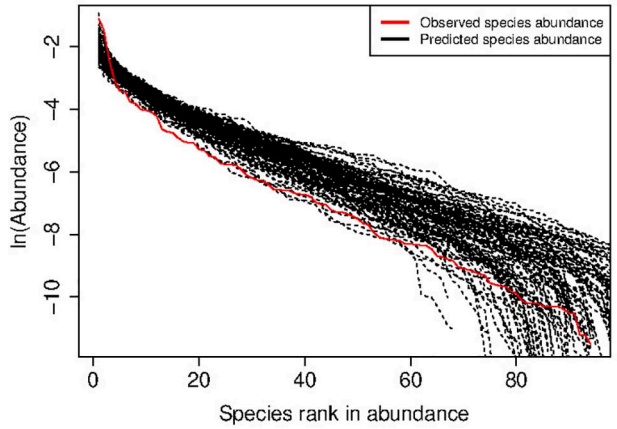


FIGURE 2 | An example of human vaginal microbial metacommunity (Subject#s35: disease status = SBV) showing successful fitting to the MSN (multi-site neutral) model: the observed *species (relative) abundance* is estimated from 68 sampling times of the HVM of the subject, and the predicted species abundance is from 25,000 times of MSN simulations.

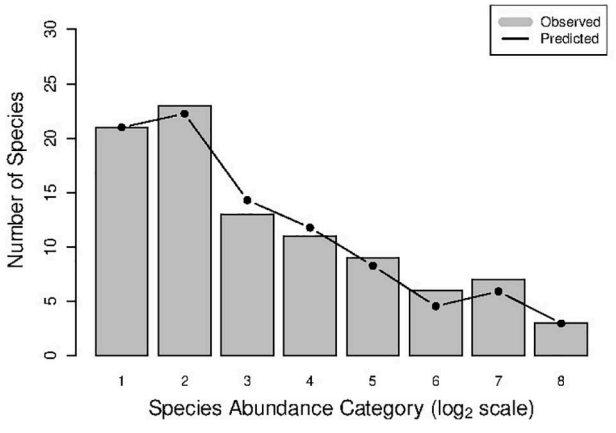


FIGURE 3 | An example of human vaginal microbial metacommunity (Subject#s3: disease status = SBV) showing successful fitting to the NNH (neutral-niche hybrid) model: the *observed number of species* is estimated from approximately 59 sampling times of the HVM of the subject, and the predicted species number is from the NNH simulations.

9 selected meta-communities from **Supplementary Table 2** in the OSI, where the full results for fitting the NNH model were listed. **Figures 2**, **3** illustrated two examples of fitting the MSN and NNH, respectively.

To better illustrate the full results in **Supplementary Tables 1**, **2** with **Tables 2**, **3**, we selected 4 meta-communities from each of the five HVM datasets, corresponding to the 4 possible outcomes of testing the MSN and NNH simultaneously (i.e., passing MSN or NNH alone, passing both or passing neither). With this scheme, a maximal number of 20 (4×5) samples could be selected, and it turned out that 11 of the combinations were missing from the results, leading to only 9 meta-communities being selected in **Tables 2**, **3**, respectively. The table legends were noted at the bottom sections of **Tables 2**, **3** below. **Tables 2**, **3**, therefore, offer windows to inspect the parameters and infer findings from fitting the MSN/NNH models. To inspect the complete test results of the 79 meta-communities and 2,733 local communities, readers are referred to **Supplementary Tables 1**, **2** in the OSI.

We now try to draw a big picture from the test results (**Supplementary Tables 1**, **2** and **Tables 2**, **3**) by computing the statistics of the passing rates from testing the MSN and NNH models. Recall that they use exactly the same data formats, i.e., with exactly the same specification for the local community and metacommunity. For example, with the dataset of "32-healthy" cohort, 32 subjects represented 32 meta-communities, and each metacommunity contained 25–33 local communities (or 25–33 *niches* in the case of NNH) given that each subject was sampled 25–33 times. **Table 4** (also see **Figure 4**) below exhibited the passing rates for both MSN

TABLE 4 | The passing percentages for testing the MSN (multi-site neutral) and NNH (niche-neutral hybrid) models with the HVM datasets, summarized from **Supplementary Tables 1, 2** and a series of the *P*-value thresholds (P_t) for testing MSN/NNH were set to P_t=0.05, 0.5, 0.9 or 0.95.

Microbiome *N		Meta community				Local community			
		0.05	0.5	0.9	0.95	0.05	0.5	0.9	0.95
The passing percentage (%) of MSN (Multi-site neutral) model									
ABV	6	100	67.7	33.3	16.7	100	50	0	0
SBV	15	86.7	80	26.7	13.3	86.7	66.7	13.3	6.7
HEA-1	4	100	100	75	75	100	100	50	25
HEA-2	32	100	100	90.6	75	100	100	75	65.6
Pregnancy	22	100	100	86.4	68.2	100	100	31.8	13.6
Overall	79	97.3	89.5	62.4	49.6	97.3	83.3	34.0	22.2
The passing percentage (%) of NNH (Niche-neutral hybrid) model									
ABV	6	0	0	0	0	81.4	57.9	19.9	14.1
SBV	15	6.7	6.7	6.7	6.7	74.0	55.2	23.9	14.8
HEA-1	4	0	0	0	0	59.2	39.8	21.0	10.6
HEA-2	32	53.1	15.6	3.1	3.1	78.8	68.1	32.5	19.8
Pregnancy	18	27.8	16.7	16.7	5.6	44.9	35.6	22.2	9.3
Overall	75	17.52	7.8	5.3	3.08	67.6	51.3	23.9	13.7

**N is the number of meta-communities (the number of individual subjects) investigated in each dataset.*

(the left) and NNH (the right) models; for each model, the passing rate at metacommunity level and local community level was listed separately. Note that in **Table 4**, the passing percentages for MSN/NNH corresponding to a series of *P*-value thresholds were presented, but here we only explain the result from the traditional "default" threshold (P = 0.05) and the results for other threshold values are explained in the following discussion section.

First, regarding the overall performance of the MSN model, 97.3% of meta-communities and local communities passed the neutrality test, respectively. The range of neutrality percentage was 86.7–100% across five datasets. Therefore, the stochastic neutral forces seem to play a dominant role in shaping the assembly of HVMs. At a local community level, the performance of NNH is significantly lower than that of MSN, with local neutrality passing the neutrality test at rate of 67.6%. However, at the metacommunity level, NNH also exhibited a moderate 17.5% of passing rate. Overall, niche differentiations appear to be moderate in the HVMs. In summary, the above findings indicate that both neutral and niche forces are in effect in shaping the community dynamics in the HVMs, but the neutral effects seem to play a dominant role.

The Effects of BV on the Neutral-Niche Continuum in the HVM

We further investigated the influence of BV (bacterial vaginosis) including both SBV (symptomatic BV) and ABV (asymptomatic BV) on the balance between neutral and niche forces in shaping the HVM dynamics by performing Fisher's exact test and Student's *t*-test. The Fisher's exact test was performed to evaluate the effect of BV on the rate of passing the neutrality test (MSN) or testing the niche-neutral hybrid effect (NNH) at the metacommunity level (the left side in **Table 5**), and Student's *t*-test on the passing rate of neutrality test at the local community level with either MSN or NNH model (the right side in **Table 5**). Similar to the previous sub-section, here we only

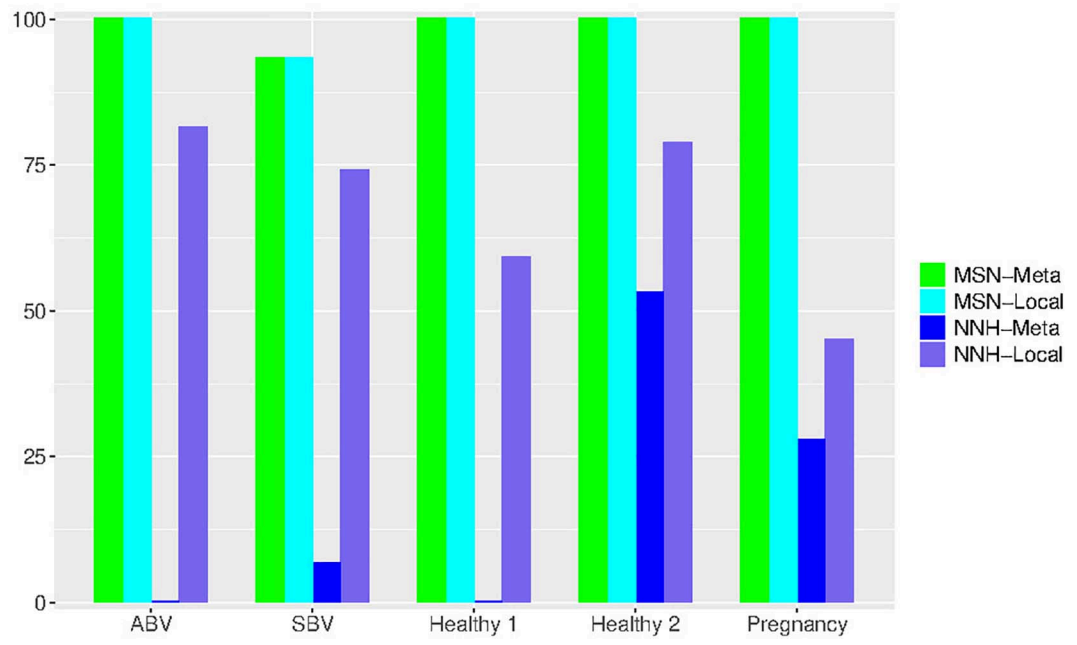


FIGURE 4 | Bar chart showing the passing percentages of testing the MSN and NNH models (when *P*-value = 0.05) for each human vaginal microbiome dataset (group), respectively: for each group, the passing percentages for both meta- and local-community of each model (MSN or NNH) were exhibited.

TABLE 5 | The *P*-values from testing the difference between various groups (ABV, SBV, HEA-1, HEA-2, and HEA) in their passing rates (from testing the MSN/NNH models) with *Fisher exact test* for the meta-community or *Student's t*-test for the local community ($^{\#,*}$).

Models	*P*-value from Metacommunity (Fisher Exact Test)			*P*-value from Local community (Student's *t*-Test)		
	Treatment	$^{\#}P_t$ = 0.05	$^{\#}P_t$ = 0.95	Treatment	$^{\#}P_t$ = 0.05	$^{\#}P_t$ = 0.95
MSN	ABV vs. HEA-1	1.000	0.559	ABV vs. HEA-1	1.000	1.000
	ABV vs. HEA-2	1.000	0.228	ABV vs. HEA-2	1.000	1.000
	SBV vs. HEA-1	1.000	0.127	SBV vs. HEA-1	1.000	1.000
	SBV vs. HEA-2	1.000	0.021	SBV vs. HEA-2	1.000	1.000
	ABV vs. SBV	1.000	1.000	ABV vs. SBV	1.000	1.000
	ABV vs. HEA (HEA-1++HEA-2)	1.000	0.230	ABV vs. HEA (HEA-1+HEA-2)	1.000	1.000
	SBV vs. HEA (HEA-1+HEA-2)	1.000	0.022	SBV vs. HEA (HEA-1+HEA-2)	1.000	1.000
	BV vs. HEA (HEA-1+HEA-2)	1.000	0.011	BV vs. HEA (HEA-1+HEA-2)	1.000	1.000
	HEA-1 vs. HEA-2	1.000	1.000	HEA-1 vs. HEA-2	1.000	1.000
NNH	ABV vs. HEA-1	1.000	1.000	ABV vs. HEA-1	0.017	0.476
	ABV vs. HEA-2	1.000	1.000	ABV vs. HEA-2	0.000	0.280
	SBV vs. HEA-1	1.000	1.000	SBV vs. HEA-1	0.299	0.548
	SBV vs. HEA-2	1.000	1.000	SBV vs. HEA-2	0.000	0.105
	ABV vs. SBV	1.000	1.000	ABV vs. SBV	0.044	0.677
	ABV vs. HEA (HEA-1+HEA-2)	1.000	1.000	ABV vs. HEA (HEA-1+HEA-2)	0.000	0.408
	SBV vs. HEA (HEA-1+HEA-2)	0.517	1.000	SBV vs. HEA (HEA-1+HEA-2)	0.000	0.182
	BV vs. HEA (HEA-1+HEA-2)	1.000	1.000	BV vs. HEA (HEA-1+HEA-2)	0.000	0.145
	HEA-1 vs. HEA-2	1.000	1.000	HEA-1 vs. HEA-2	0.017	0.113

*$^{\#}$The P-value thresholds for testing the MSN/NNH models were set to 0.05 and 0.95, corresponding to the two column heads "P = 0.05" and "P = 0.95"; be noted that they are column heads and are totally different the P-values in the table entries that are from Fisher or t-tests. *The HEA1 group has 4 subjects only, and we suggest following the inferences from the HEA2 (32 subjects) or HEA (=HEA1+HEA2) in case there were conflicting results between HEA1 with the other groups. Shaded entries are comparisons with significant differences ($P \leq 0.05$).*

analyze the BV effects under traditional *P*-value = 0.05 threshold and delay the analyses under alternative *P*-value thresholds to the discussion section.

Interestingly, both MSN and NNH exhibited slightly different results regarding the effects of BV status on the passing percentages of model tests. At the local community level, there appears to be significant differences in BV (SBV) and HEA (healthy groups) (*P*-value < 0.05, **Table 5**) according to the NNH model. However, at the metacommunity level, regardless of the MSN or NNH, the differences between various groups were statistically insignificant. The lack of difference between the HEA and pregnancy groups is also expected. Romero et al. (2014) defined a normal pregnancy as a woman with no obstetrical, medical or surgical complications, and delivered at term (38 to 42 weeks) without complications. The pregnancy group studied by Romero consisted of 22 normal pregnancies. Therefore, it appears that no statistically significant differences were detected between various groups in terms of MSN or NNH testing except for NNH at local community scale.

DISCUSSION

With traditional neutral theory of biodiversity, the spatially explicit or implicit model describing the metacommunity consisting of multiple local communities is the most frequently used metacommunity model for testing neutrality (Hubbell, 2001; Rosindell et al., 2011, 2012; Ma, 2020). The use of longitudinal community/meta-community samples to perform dynamic analysis of the niches vs. neutrality and to further assess and interpret the community static and dynamic properties by harnessing the neutral theory has been few but can be equally effective (Kalyuzhny et al., 2014a,b, 2015). The integrated modeling with MSN/NNH in previous sections demonstrated another approach to generalizing the neutral-theoretic analysis to temporal meta-communities. Furthermore, we take advantages of a recent advance in computational statistics made by Harris et al. (2017) HDP-MSN machine learning algorithm. The HDP-MSN overcomes a significant computational bottleneck that existed in estimating the migration rates (m) when the number of local communities is large, which prevented large-scale testing of the UNTB with truly multi-site datasets. Nevertheless, truly multi-site datasets are scarce, especially in the studies of the human microbiome, where community samples are usually taken from unrelated individuals, and therefore dispersal (migration) among individuals is unlikely to occur on ecological timescales. In this study, we use the time-series sampling data in place of spatial sampling data. That is, the vaginal microbial community of each subject was sampled in patients at varying numbers of time points (6–60, see **Table 1**). By using time-series data with the MSN and NNH models, one can effectively evaluate the levels of stochastic *neutral* forces and deterministic *niche* forces in driving the community dynamics. In the case of time-series data, stochastic neutral forces may include stochastic fluctuations in demography (in the birth-death processes of bacterial cell divisions and deaths), which is the analog to ecological drift in neutral theory. The deterministic forces in time series data can include diversity- or

TABLE 6 | Comparative summary of the performances of MSN and NNH models fitted to the human vaginal microbiome datasets of 79 subjects (meta-communities), summarized from **Supplementary Tables 1**, **2**, under different P_t-value thresholds for testing MSN/MMH models.

Microbiome	Meta-Community	MSN only		NNH only		Both MSN & NNH		**NOT** (MSN, NNH)	
		N	**%**	***N***	**%**	***N***	**%**	***N***	**%**
P_t **= 0.05**									
ABV	6	6	100.0	0	0.0	0	0.0	0	0.0
SBV	15	12	80.0	0	0.0	1	6.7	2	13.3
HEA-1	4	4	100.0	0	0.0	0	0.0	0	0.0
HEA-2 (32-Cohort)	32	15	46.9	0	0.0	17	53.1	0	0.0
Pregnancy	22	17	77.3	0	0.0	5	22.7	0	0.0
Overall	79	54	68.4	0	0.0	23	29.1	2	2.5
P_t **= 0.5**									
ABV	6	4	66.7	0	0.0	0	0.0	2	33.3
SBV	15	11	73.3	0	0.0	1	6.7	3	20.0
HEA-1	4	4	100.0	0	0.0	0	0.0	0	0.0
HEA-2 (32-Cohort)	32	27	84.4	0	0.0	5	15.6	0	0.0
Pregnancy	22	19	86.4	0	0.0	3	13.6	0	0.0
Overall	79	65	82.3	0	0.0	9	11.4	5	6.3
P_t **= 0.9**									
ABV	6	2	33.3	0	0.0	0	0.0	4	66.7
SBV	15	4	26.7	1	6.7	0	0.0	10	66.7
HEA-1	4	3	75.0	0	0.0	0	0.0	1	25.0
HEA-2 (32-Cohort)	32	28	87.5	0	0.0	1	3.1	3	9.4
Pregnancy	22	17	77.3	1	4.5	2	9.1	2	9.1
Overall	79	54	68.4	2	2.5	3	3.8	20	25.3
P_t **= 0.95**									
ABV	6	1	16.7	0	0.0	0	0.0	5	83.3
SBV	15	2	13.3	1	6.7	0	0.0	12	80.0
HEA-1	4	3	75.0	0	0.0	0	0.0	1	25.0
HEA-2 (32-Cohort)	32	24	75.0	1	3.1	0	0.0	7	21.9
Pregnancy	22	14	63.6	0	0.0	1	4.5	7	31.8
Overall	79	44	55.7	2	2.5	1	1.3	32	40.5

dominance-dependent regulatory forces for community stability (dynamics) (Ma and Ellison, 2018, 2019).

As explained in the previous section of results, the results from the integrated niche-neutral hybrid analyses under a default *P*-value = 0.05 with MSN and NNH models in this study seem to suggest that neutral drifts play a dominant role in driving the community dynamics of the HVM, while the deterministic niche differentiation is moderate (approximately 17% in terms of passing NNH test). As further elaborated below, the assessment of the relative significance of neutral vs. niche may be strongly influenced by the model-choice (MSN or NNH) and *P*-value thresholds. It should also be reiterated that the conclusions obtained from this study are from analyzing the temporal dynamics data of the HVM, rather than from analyzing spatial metacommunity samples as usually performed in community ecology. In other words, the local communities in our analyses are simply the snapshots of an individual woman's vaginal microbiome dynamics. Therefore, niche differentiations are also "temporal differentiations," which might be relatively weak due to the nature of longitudinal observations. Future studies with "orthodox" spatial metacommunity samples should shed a more comprehensive picture on the community assembly and diversity maintenance of the HVMs. In the remainder of this article, we discuss two uncertainties regarding the test of the neutral and/or niche-neutral hybrid models.

First, it is well known that a significant challenge in investigating the mechanisms of community assembly or distinguishing the neutral from niche effects is that multiple independents models with possibly different ecological assumptions about the mechanisms may produce similar goodness-of-fittings to the same datasets, which is termed the lack of full model specificity in previous sections. This can make the inferences of definite mechanisms from different models difficult since the mapping from assumptions to mechanisms may not be one-to-one. **Table 6** listed the breakups of successful fittings of the MSN/NNH models with *P*-value thresholds of 0.05–0.95, classified as four groups including successfully fitted to "MSN-only," "NNH-only," "both MSN & NNH," and "neither MSN nor NNH." Here, we first discuss the breakups when *P*-value threshold for testing the MSN/NNH model is set to 0.05, and the other thresholds are discussed shortly below. As exhibited in **Table 6**, overall, the "MSN-only group" (fitted to MSN uniquely) takes about 68% (ranged between 47 and 100%) of all cases and "both MSN & NNH" group (non-unique fittings)

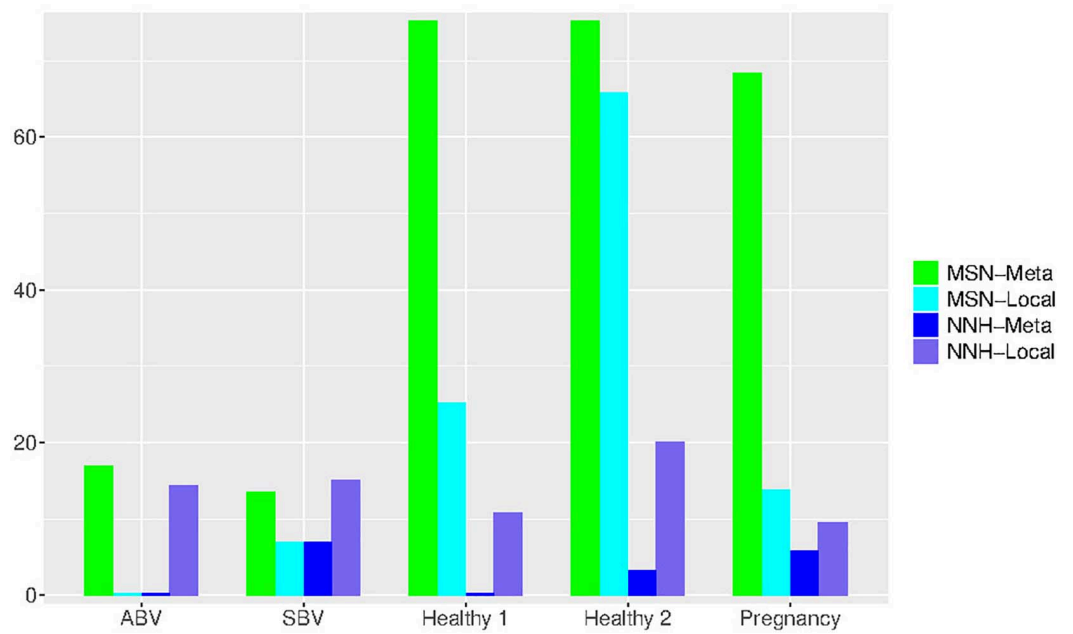


FIGURE 5 | Bar chart showing the passing percentages of testing the MSN and NNH models (when *P*-value = 0.95) for each human vaginal microbiome dataset (group), respectively: for each group, the passing percentages for both meta- and local-community of each model (MSN or NNH) were exhibited.

takes about 29%, (ranged between 0 and 53%) and in less than 3% cases (2 out of 79 individuals) neither MSN nor NNH model was fitted successfully. Therefore, in the majority (68% or 54 out of 79 individuals), the MSN model was able to uniquely interpret the neutral dynamics of the HVMs, when the *P*-value threshold for testing MSN/NNH was set to 0.05.

Second, another dilemma that may lead to uncertainty in testing the neutral or niche-neutral hybrid models is the choice of *P*-value threshold. Traditionally, the *P*-value was set to 0.05 in testing the neutral theory; when $P > 0.05$, the null hypothesis or model (satisfying the MSN or NNH model) cannot be rejected. In other words, when $P > 0.05$, the observed community is considered indistinguishable from what the theoretical model predicts. In previous sections, *P*-value = 0.05 was termed traditional "default." However, one may set *P*-value to other threshold values. The higher the *P*-value is, the more likely (the higher likelihood) that the community is consistent with the model prediction. That is, when the *P*-value is set to higher threshold values, it is more difficult to reject the null model. In terms of the neutrality test based on the MSN model, it implies that accepting neutral hypothesis is more reliable (conservative). In terms of the NNH model, it implies that accepting non-neutrality (niche differentiation) is more reliable (conservative). Consequently, when larger *P*-value thresholds are adopted, the confidence (reliability) to accepting the null model (MSN or NNH) is raised and the confidence to reject the null model (MSM or NNH) is lowered. **Table 4** listed the passing percentages (strictly, should be stated as percentage indistinguishable from model prediction) from testing MSN/NNH when *P*-value was set to 0.05, 0.5, 0.9, and 0.95, respectively. Obviously, as shown in **Table 4**, higher *P*-values correspond to a lower passing percentage of MSN-neutrality tests. When the *P*-value threshold was raised to 0.95, the passing percentage of MSN-neutrality test declined to approximately 50%, while the percentage was 97.3% when the *P*-value = 0.05. Raising the *P*-value threshold from 0.05 to 0.95 is a rather dramatic increase of the confidence level for not rejecting the null neutral model (or accepting the null model), still nearly half the metacommunities (microbiomes of individuals) passed the MSN model, suggesting that the neutral drifts indeed play a significant role in shaping the dynamics of the HVM.

An interesting observation is that, when the *P*-value threshold was set to default 0.05 (**Table 4**), the passing percentages of MSM testing were not significantly different between different treatments (see **Table 5** for Fisher's exact test). However, when the *P*-value threshold was set to 0.95 (**Figure 5**), the differences between the BV group (including ABV and SBV) and healthy groups (HEA1, HEA2, and pregnancy) were significant (13.3–16.7% vs. 68.2–75%). The healthy groups exhibited significantly more neutral communities than BV groups (**Table 5**). This result is actually puzzling. A traditional view has been that the vaginal microbiomes associated with BV usually have higher community diversity than the healthy counterparts, possibly due to the loss of dominant species such as *Lactobacillus* (e.g., Ma et al., 2012). Since communities with dominant species tend to contain more asymmetric interactions, they could be less likely neutral. Therefore, this puzzling result appears to contradict with the traditional view. Of course, the relationship between dominant species and non-neutrality is not necessarily positive; we hope that future more mechanistic studies will resolve the apparent inconsistency.

The adoption of different *P*-value thresholds may also affect the previously discussed specificity (uniqueness) of MSN/NNH

model fittings. **Table 6** shows the breakups of four categories (MSN-only, NNH-only, both MSN & NNH, neither MSN nor NNH) under different *P*-values ranging from *P*-value = 0.05 to 0.95. It appears that the model specificity of MSN seems to increase with the increase of *P*-value threshold adopted. This result should be expected since the increased *P*-value should raise the confidence level for accepting the null (MSN or NNH) model, and the passing percentages judged with higher confidence would decline accordingly.

Finally, the two previously discussed uncertainties associated with testing the neutral or niche-neutral hybrid models can be interdependent and proper resolving them requires both ecological science and statistical art. There may not be a perfect solution for resolving those issues due to both the ecological complexity and the limitation of frequentist statistical approaches. The art lies in balancing the trade-off between reliability (confidence) in hypothesis testing and model specificity. It is hoped that the demonstrative analysis and discussion included in this study with the HVMs will also be useful for other ecological and evolutionary modeling of biodiversity and biogeography.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

ZM designed and performed the study and wrote the manuscript.

FUNDING

This study received funding from the following sources: A National Natural Science Foundation (NSFC) Grant (No. 31970116) on "Medical Ecology of Human Microbiome", and the "Cloud-Ridge Industry Technology Leader" Award. An International Cooperation Grant (YNST) on Genomics and Metagenomics Big Data. The funders played no roles in interpreting the results.

ACKNOWLEDGMENTS

I deeply appreciate Hubbell for reviewing this manuscript and for his insightful comments and suggestions to improve my work. I am obliged to express my sincere appreciation to Rampal Etienne for his insightful advice and comments on testing neutral models based on the *P*-value. I thank Christopher Quince for his professional cooperation and support in updating their published work. I appreciate the computational support from LW Li of the Chinese Academy of Sciences.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.699939/full#supplementary-material

Supplementary Table 1 | Test results of fitting the MSN (multi-site neutral) model to the vaginal microbiome datasets.

Supplementary Table 2 | Fitting the NNH (Niche-Neutral Hybrid) model to the vaginal microbiome datasets.

REFERENCES

Alonso, D., Etienne, R., and Mckane, A. (2006). The merits of neutral theory. *Trends Ecol. Evol.* 21, 451–457. doi: 10.1016/j.tree.2006.03.019

Antoniak, C. E. (1974). Mixtures of dirichlet processes with applications to bayesian nonparametric problems. *Ann. Statist.* 2, 1152–1174.

Chisholm, R. A., and Pacala, S. W. (2010). Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proc. Natl Acad. Sci. U S A.* 107, 15821–15825. doi: 10.1073/pnas.1009387107

Darwin, C. (1859). *On the Origin of Spices*. London: John Murray.

Etienne, R. S. (2007). A neutral sampling formula for multiple samples and an 'exact' test of neutrality. *Ecol. Lett.* 10, 608–618. doi: 10.1111/j.1461-0248.2007.01052.x

Etienne, R. S. (2009a). Improved estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. *Ecology* 90, 847–852. doi: 10.1890/08-0750.1

Etienne, R. S. (2009b). Maximum likelihood estimation of neutral model parameters for multiple samples with different degrees of dispersal limitation. *J. Theor. Biol.* 257, 510–514. doi: 10.1016/j.jtbi.2008.12.016

Fisher, C. K., and Mehta, P. (2014). The transition between the niche and neutral regimes in ecology. *Proc. Natl. Acad. Sci. U S A.* 111, 13111–13116. doi: 10.1073/pnas.1405637111

Fredricks, D. N. (2011). Molecular methods to describe the spectrum and dynamics of the vaginal microbiota. *Anaerobe* 17, 191–195. doi: 10.1016/j.anaerobe.2011.01.001

Fredricks, D. N., Fiedler, T. L., and Marrazzo, J. M. (2005). Molecular identification of bacteria associated with bacterial vaginosis. *N. Engl. J. Med.* 353, 1899–1911. doi: 10.1056/nejmoa043802

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M., Zhong, X., et al. (2012). Temporal dynamics of the human vaginal microbiota. *Sci. Trans. Med.* 132:132ra52.

Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., and Knight, R. (2018). Current understanding of the human microbiome. *Nat. Med.* 24, 392–400. doi: 10.1038/nm.4517

Grinnell, J. (1917). The niche-relationships of the California Thrasher. *Auk* 34, 427–433. doi: 10.2307/4072271

Harris, K., Parsons, T. L., Ijaz, U. Z., Lahti, L., Holmes, I., Quince, C., et al. (2017). Linking statistical and ecological theory: hubbell's unified neutral theory of biodiversity as a hierarchical dirichlet process. *Proc. IEEE* 105, 516–529. doi: 10.1109/jproc.2015.2428213

Holt, R. D. (2009). Bringing the hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proc. Natl. Acad. Sci. U S A.* 106, 19659–19665. doi: 10.1073/pnas.0905137106

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, NJ: Princeton University Press.

Hubbell, S. P. (2006). Neutral theory and the evolution of ecological equivalence. *Ecology* 87, 1387–1398. doi: 10.1890/0012-9658(2006)87[1387:ntateo]2.0.co;2

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harb. Symp.* 22, 415–427.

Jeraldo, P., Sipos, M., Chia, N., Brulc, J. M., Dhillon, A. S., Konkel, M. E., et al. (2012). Quantification of the relative roles of niche and neutral processes in

structuring gastrointestinal microbiomes. *Proc. Natl. Acad. Sci. U S A.* 109, 9692–9698. doi: 10.1073/pnas.1206721109

Kalyuzhny, M., Kadmon, R., and Shnerb, N. M. (2015). A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecol. Lett.* 18, 572–580. doi: 10.1111/ele.12439

Kalyuzhny, M., Kadmon, R., and Shnerb, N. M. (2014a). A generalized neutral theory explains static and dynamic properties of biotic communities. *Quan. Biol.* 21, 62–74.

Kalyuzhny, M., Seri, E., Chocron, R., Flather, C. H., Kadmon, R., and Shnerb, N. M. (2014b). Niche versus neutrality: a dynamical analysis. *Am. Nat.* 184, 439–446. doi: 10.1086/677930

Knight, R., Callewaert, C., Marotz, C., Hyde, E. R., Debelius, J. W., McDonald, D., et al. (2017). The microbiome and human biology. *Ann. Rev. Genom. Hum. Genet.* 18, 65–86.

Lynch, S. V., and Pedersen, O. (2016). The human intestinal microbiome in health and disease. *N. Engl. J. Med.* 375, 2369–2379. doi: 10.1056/NEJMra1600266

Ma, B., Forney, L. J., and Ravel, J. (2012). The vaginal microbiome: rethinking health and disease. *Annu. Rev. Microbiol.* 66, 371–389. doi: 10.1146/annurev-micro-092611-150157

Ma, Z. S. (2020). Niche-neutral theoretic approach to mechanisms underlying the biodiversity and biogeography of human microbiomes. *Evol. Appl.* 2020, 1–13. doi: 10.1111/eva.13116

Ma, Z. S., and Ellison, A. M. (2018). A unified concept of dominance applicable at both community and species scale. *Ecosphere* 9:e02477. doi: 10.1002/ecs2.2477

Ma, Z. S., and Ellison, A. M. (2019). Dominance network analysis provides a new framework for studying the diversity-stability relationship. *Ecol. Monographs* 89:e01358. doi: 10.1002/ecm.1358

Matthews, T. J., and Whittaker, R. J. (2014). Neutral theory and the species abundance distribution: recent developments and prospects for unifying niche and neutral perspectives. *Ecol. Evol.* 4, 2263–2277.

McGill, B. J., Maurer, B. A., and Weiser, M. D. (2006). Empirical evaluation of neutral theory. *Ecology* 87, 1411–1423. doi: 10.1890/0012-9658(2006)87[1411:eeont]2.0.co;2

Noble, A. E., and Fagan, W. F. (2015). A niche remedy for the dynamical problems of neutral theory. *Theoret. Ecol.* 8, 149–161. doi: 10.1007/s12080-014-0240-x

Ofiteru, I. D., Lunn, M., Curtis, T. P., Wells, G. F., and Criddle, C. S. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *Proc. Natl. Acad. Sci. U S A.* 107, 15345–15350. doi: 10.1073/pnas.1000604107

Pigolotti, S., and Cencini, M. (2013). Species abundances and lifetimes: from neutral to niche-stabilized communities. *J. Theor. Biol.* 338, 1–8. doi: 10.1016/j.jtbi.2013.08.024

Ravel, J., Brotman, R. M., Gajer, P., Ma, B., Nandy, M., Fadrosh, D. W., et al. (2013). Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis. *Microbiome* 1:29.

Romero, R., Hassan, S. S., Gajer, P., Tarca, A. L., Fadrosh, D. W., Nikita, L., et al. (2014). The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women. *Microbiome* 2:4.

Rosindell, J., Hubbell, S. P., and Etienne, R. S. (2011). The unified neutral theory of biodiversity and biogeography at age ten. *Trends Ecol. Evol.* 26, 340–348. doi: 10.1016/j.tree.2011.03.024

Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., and Etienne, R. S. (2012). The case for ecological neutral theory. *Trends Ecol. Evol.* 27, 203–208. doi: 10.1016/j.tree.2012.01.004

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.* 8, 732–740. doi: 10.1111/j.1462-2920.2005.00956.x

Sloan, W. T., Woodcock, S., Lunn, M., Head, I. M., and Curtis, T. P. (2007). Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microb. Ecol.* 53, 443–455. doi: 10.1007/s00248-006-9141-x

Sobel, J. D. (1999). Is there a protective role for vaginal flora? *Curr. Infect. Dis. Rep.* 1, 379–383. doi: 10.1007/s11908-999-0045-z

Stokes, C. J., and Archer, S. R. (2010). Niche differentiation and neutral theory: an integrated perspective on shrub assemblages in a parkland savanna. *Ecology* 91, 1152–1162. doi: 10.1890/08-1105.1

Tang, J., and Zhou, S. (2013). Hybrid niche-neutral models outperform an otherwise equivalent neutral model for fitting coral reef data. *J. Theor. Biol.* 317, 212–218. doi: 10.1016/j.jtbi.2012.10.019

Teh, Y. W., Jordan, M., Beal, M., and Blei, D. (2006). Hierarchical dirichlet processes. *J. Am. Statist. Assoc.* 101, 1566–1581.

Tilman, D. (2004). Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proc. Natl. Acad. Sci. U S A.* 101, 10854–10861. doi: 10.1073/pnas.0403458101

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2003). Neutral theory and relative species abundance in ecology. *Nature* 424, 1035–1037. doi: 10.1038/nature01883

Volkov, I., Banavar, J. R., Hubbell, S. P., and Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature* 450, 45–49. doi: 10.1038/nature06197

Young, V. B. (2017). The role of the microbiome in human health and disease: an introduction for clinicians. *BMJ* 2017:356. doi: 10.1136/bmj.j831

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

frontiers in Microbiology

ORIGINAL RESEARCH
published: 17 September 2021
doi: 10.3389/fmicb.2021.695465

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

*Qi Yan[1,2], Jianming Deng[1,3], Feng Wang[4,5], Yongqin Liu[2,4] and Keshao Liu[4]**

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

OPEN ACCESS

Edited by:
Daliang Ning,
University of Oklahoma,
United States

Reviewed by:
Anyi Hu,
Institute of Urban Environment,
Chinese Academy of Sciences (CAS),
China
Jinbo Xiong,
Ningbo University, China

****Correspondence:***
Keshao Liu
liukeshao@itpcas.ac.cn

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *15 April 2021*
Accepted: *09 August 2021*
Published: *17 September 2021*

Citation:
Yan Q, Deng J, Wang F, Liu Y and Liu K (2021) Community Assembly and Co-occurrence Patterns Underlying the Core and Satellite Bacterial Sub-communities in the Tibetan Lakes. Front. Microbiol. 12:695465. doi: 10.3389/fmicb.2021.695465

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; Ofiţeru 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 droved 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 2062 km^2 (**Supplementary Table S1**).

In total, 47 water samples were collected from 30 Tibetan lakes. The Schindler sampler was used to collect approximately 1 L surface water samples (∼0.5 m 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 20 μm mesh (Millipore, United States) for removal of the large plankton and particles, and all filtrates were subsequently filtered through a 0.22 μm polycarbonate membrane (Millipore, united states). Afterward, the membranes were put in sterile 2 ml 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'-GTGCCAGCMGCCGCGGTAA-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 by using QIIME 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 qiime1was 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 χ^2 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 ln-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 ($\beta NTI) < -2$ 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 $|\beta NTI| < 2$ and $RC_{bray} < -0.95$ or the $|\beta NTI| < 2$ and $RC_{bray} \geq 0.95$ RCbray were identified as homogenizing dispersal and dispersal limitation, respectively. When the $|\beta NTI| < 2$ 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):

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

where B_j represents the habitat niche breadth; P_{ij} 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 (ρ) between those OTUs were calculated pairwise by the "Hmisc" package in an R environment. Only robust correlations with Spearman's correlation coefficients (ρ) > 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 χ^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 χ^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, $p < 0.05$; **Figure 2**). At the genus level, 43 genera showed significant differences between the two sub-communities ($p < 0.05$; **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 3**). 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%)

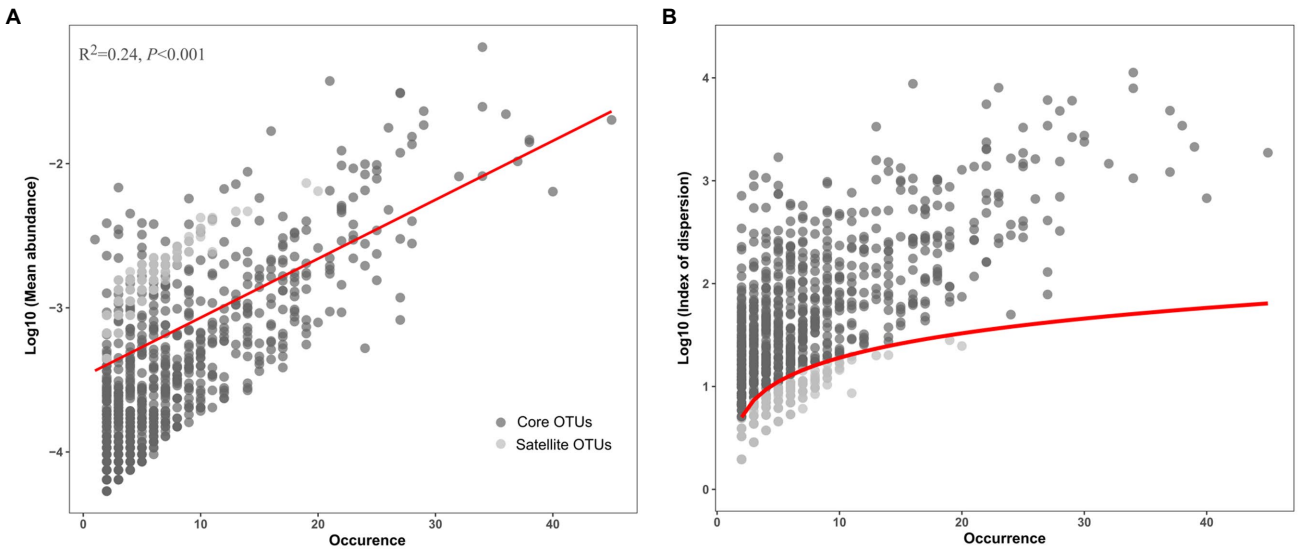


FIGURE 1 | Distribution **(A)** and dispersion **(B)** of the core and satellite OTUs in the Tibetan lakes. **(A)** The OTU occurrence (number of samples in which a given OTU was detected) plotted against the mean relative abundance of each OTU across samples. The red line represents the linear regression model fit to species abundance distribution. **(B)** The OTU occurrence plotted against the index of dispersion for each OTU calculated as variance to mean ratio of abundance for each OTU. The red line represents 2.5% confidence limit for the χ^2 distribution.

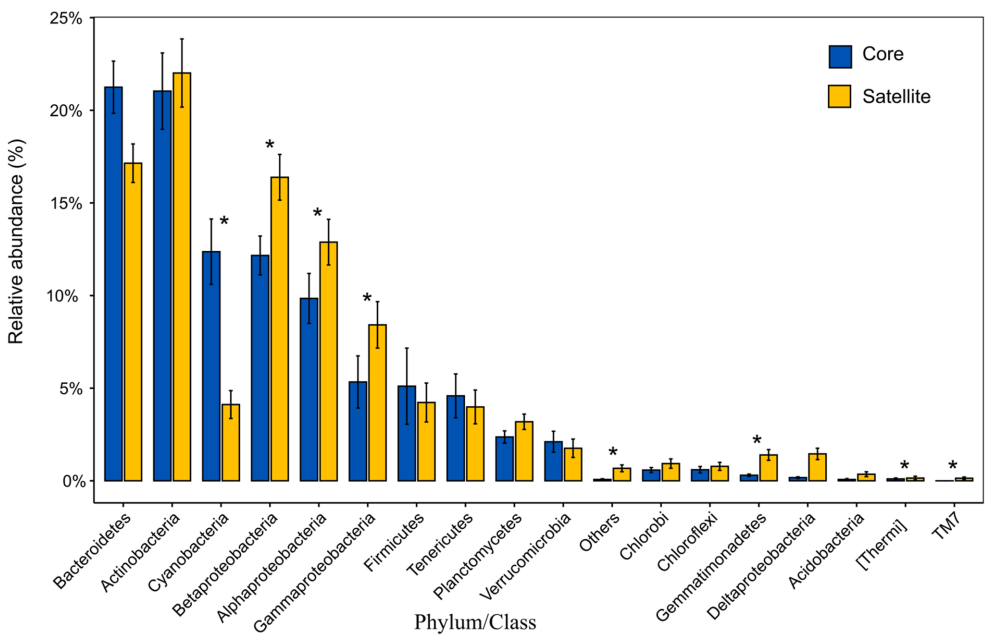


FIGURE 2 | The relative abundance of the 18 common phyla and the classes of Proteobacteria within the core and satellite communities. Asterisks indicate significant differentially abundant phyla/class between the core and satellite bacterial sub-communities ($p < 0.05$).

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

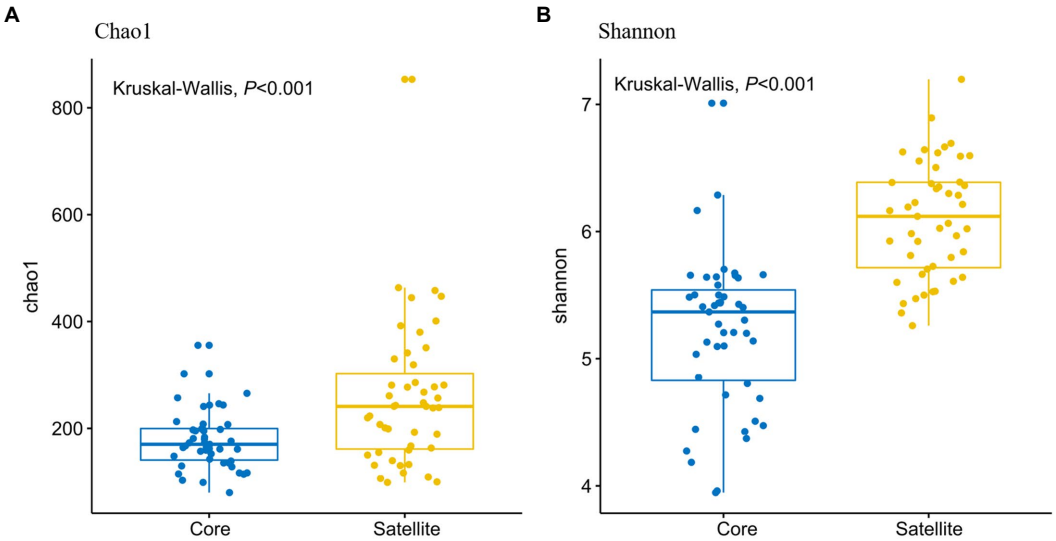


FIGURE 3 | Comparison of α-diversities of the core and satellite bacterial sub-communities. **(A)** Chao1 index and **(B)** Shannon index.

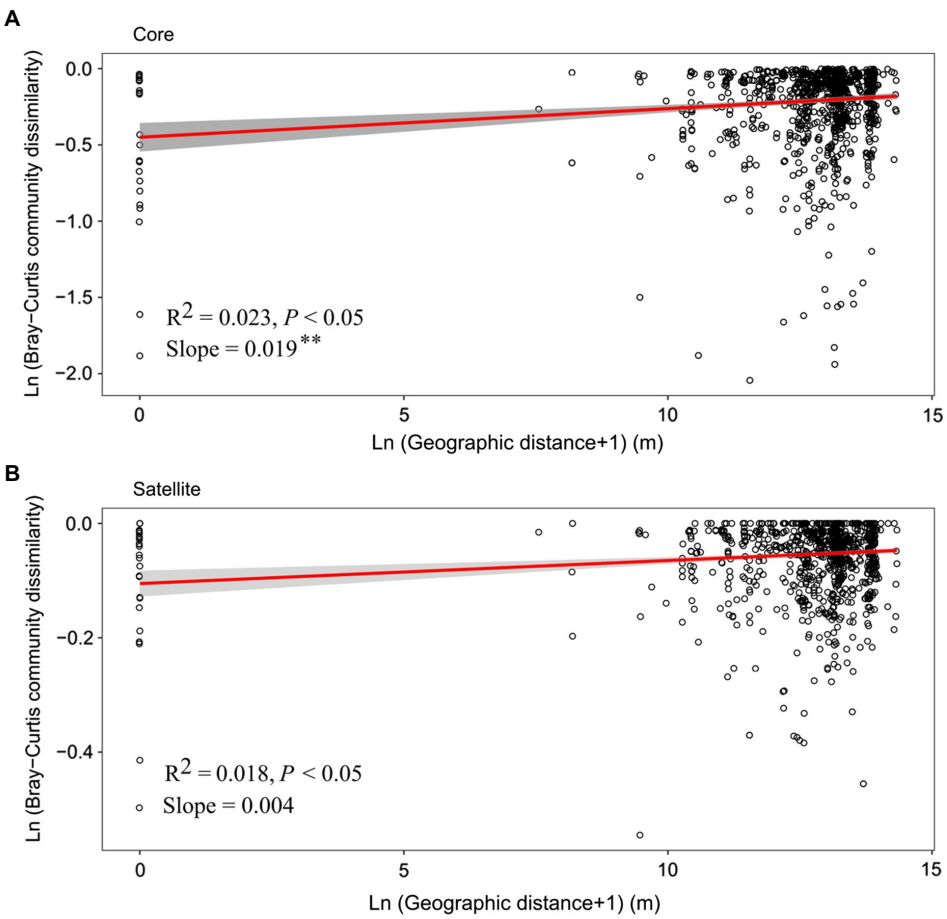


FIGURE 4 | The relationship between geographic distances and Bray-Curtis dissimilarities of the **(A)** core and **(B)** satellite bacterial sub-communities. The red line in each plot represents a linear regression model fit to Ln (geographic distance+1) vs. Ln (Bray-Curtis community dissimilarity). 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).

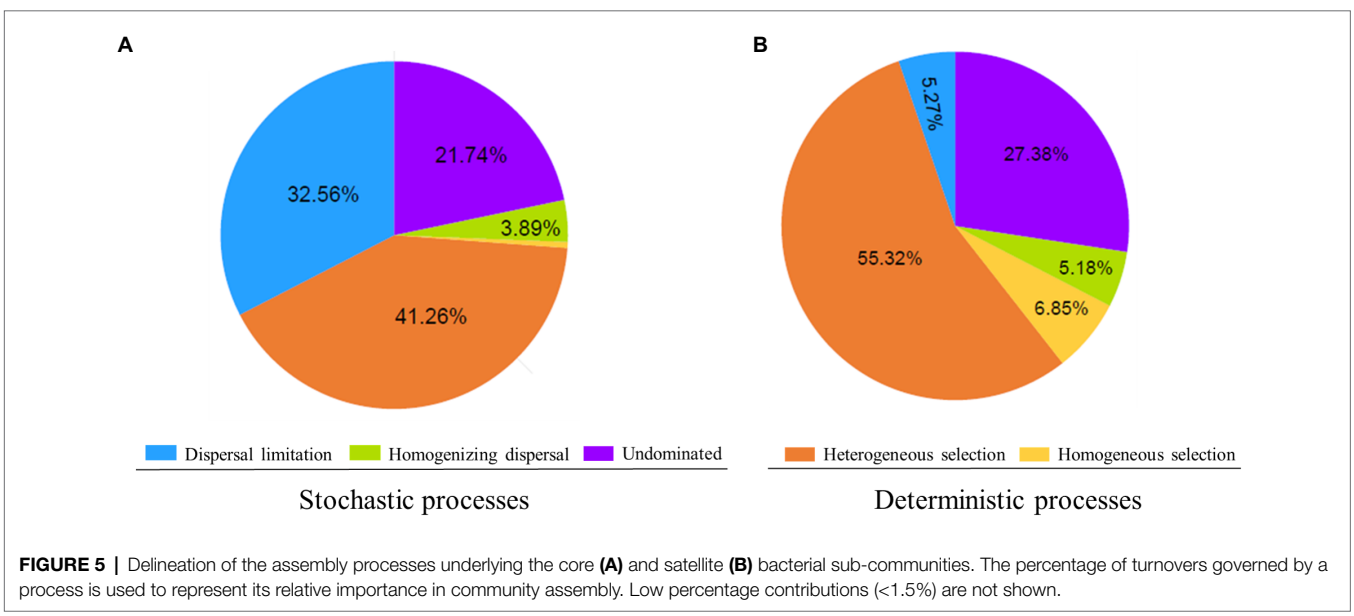


FIGURE 5 | Delineation of the assembly processes underlying the core **(A)** and satellite **(B)** bacterial sub-communities. The percentage of turnovers governed by a process is used to represent its relative importance in community assembly. Low percentage contributions (<1.5%) are not shown.

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

Network properties	Value
Empirical network	
Nodes	518
Edges	5,145
Average clustering coefficient	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 coefficient	0.04 ± 0.001
Average path length	2.41 ± 0.001
Modularity	0.19 ± 0.003

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 consideration. 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. A possible explanation for this might be due to environmental heterogeneity and the capability differences in the response to environmental change (Morrissey et al., 2019). Another possible explanation for this is that differences in species 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 shape 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.

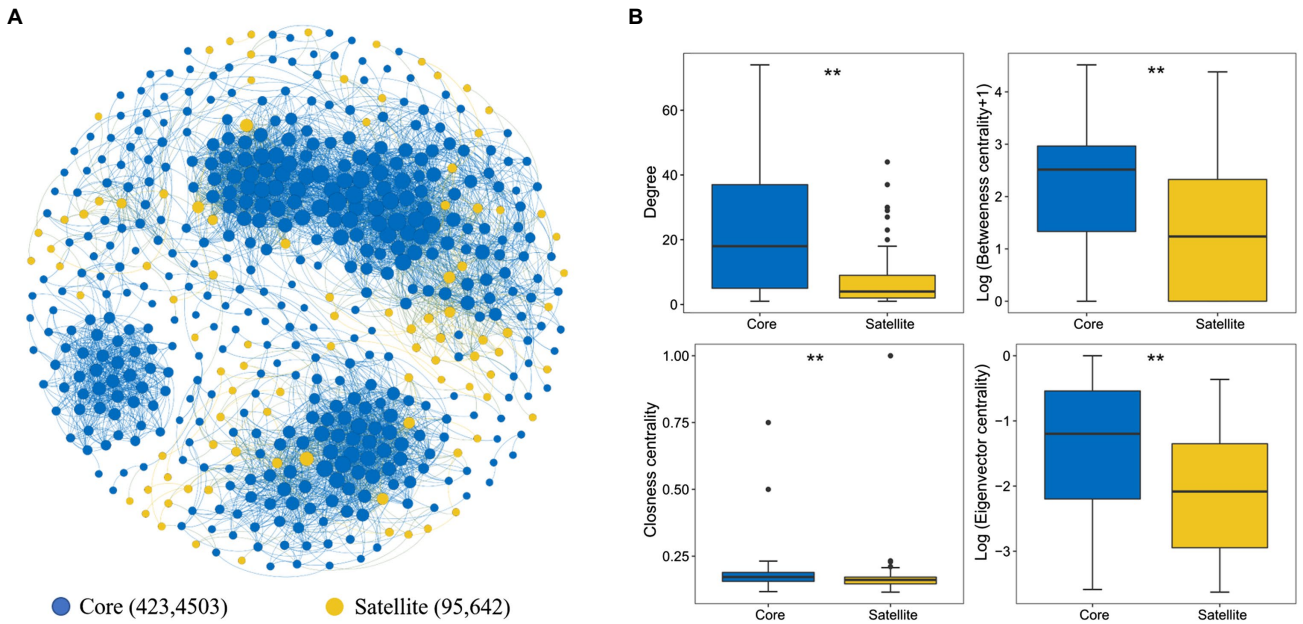


FIGURE 6 | Properties of the correlation-based network among the core and satellite sub-communities. **(A)** Network of inter-taxon associations within and between core and satellite sub-communities. A connection stands for a strong (Spearman's $\rho > 0.6$ or $\rho < -0.6$) and significant ($p < 0.01$) correlation. The size of each node is proportional to the degree of the OTU. Numbers represent the nodes and edges of the core and satellite sub-communities. **(B)**. Comparison of node-level topological features between two different sub-communities. "**" indicates significant differences ($p < 0.05$), determined by the Kruskal-Wallis test.

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

REFERENCES

Alonso-Sáez, L., Díaz-Pérez, L., and Morán, X. A. G. (2015). The hidden seasonality of the rare biosphere in coastal marine bacterioplankton. *Environ. Microbiol.* 17, 3766–3780. doi: 10.1111/1462-2920.12801

Barnes, C. J., Burns, C. A., Van, D. G. C. J., Mcnamara, N. P., and Bending, G. D. (2016). Spatio-temporal variation of core and satellite arbuscular mycorrhizal fungus communities in miscanthus giganteus. *Front. Microbiol.* 7:1278. doi: 10.3389/fmicb.2016.01278

Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *ICWSM* 8, 361–362. doi: 10.13140/2.1.1341.1520

Bolger, A. M., Marc, L., and Bjoern, U. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 15, 2114–2120. doi: 10.1093/bioinformatics/btu170

Campbell, B. J., Yu, L., Heidelberg, J. F., and Kirchman, D. L. (2011). Activity of abundant and rare bacteria in a coastal ocean. *Proc. Natl. Acad. Sci. U. S. A.* 108, 12776–12781. doi: 10.1073/pnas.1101405108

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Christopher, Q., Anders, L., Russell, D. P., et al. (2011). Removing noise from pyrosequenced amplicons. *BMC Bioinf.* 12:38. doi: 10.1186/1471-2105-12-38

DeSantis, T. Z., Hugenholtz, P., Keller, K., Brodie, E. L., Larsen, N., Piceno, Y. M., et al. (2006). NAST: a multiple sequence alignment server for comparative analysis of 16s rRNA genes. *Nucleic Acids Res.* 34, 394–399. doi: 10.1093/nar/gkl244

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., and Salles, J. F. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. U. S. A.* 112, E1326–E1332. doi: 10.1073/pnas.1414261112

Du, S., Dini-Andreote, F., Zhang, N., Liang, C., Yao, Z., Zhang, H., et al. (2020). Divergent co-occurrence patterns and assembly processes structure the abundant and rare bacterial communities in a salt marsh ecosystem. *Appl. Environ. Microbiol.* 86:e00322. doi: 10.1128/AEM.00322-20

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Farjalla, V. F., Srivastava, D. S., Marino, N. A., Azevedo, F. D., Dib, V., Lopes, P. M., et al. (2012). Ecological determinism increases with organism size. *Ecology* 93, 1752–1759. doi: 10.1890/11-1144.1

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550. doi: 10.1038/nrmicro2832

Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature* 459, 193–199. doi: 10.1038/nature08058

Galand, P. E., Casamayor, E. O., Kirchman, D. L., and Lovejoy, C. (2009). Ecology of the rare microbial biosphere of the Arctic Ocean. *Proc. Natl. Acad. Sci. U. S. A.* 106, 22427–22432. doi: 10.1073/pnas.0908284106

Gao, C., Montoya, L., Xu, L., Madera, M., Hollingsworth, J., Purdom, E., et al. (2020). Fungal community assembly in drought-stressed sorghum shows stochasticity, selection, and universal ecological dynamics. *Nat. Commun.* 11:34. doi: 10.1038/s41467-019-13913-9

Gendron, E. M. S., Darcy, J. L., Hell, K., and Schmidt, S. K. (2019). Structure of bacterial and eukaryote communities reflect in situ controls on community assembly in a high-alpine lake. *J. Microbiol.* 57, 852–864. doi: 10.1007/s12275-019-8668-8

Ghoul, M., and Mitri, S. (2016). The ecology and evolution of microbial competition. *Trends Microbiol.* 24, 833–845. doi: 10.1016/j.tim.2016.06.011

Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbrück, L., Reeder, J., Temperton, B., et al. (2012). Defining seasonal marine microbial community dynamics. *ISME J.* 6, 298–308. doi: 10.1038/ismej.2011.107

Graham, E. B., and Stegen, J. C. (2017). Dispersal-based microbial community assembly decreases biogeochemical function. *PRO* 5:65. doi: 10.3390/pr5040065

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* 10, 497–506. doi: 10.1038/nrmicro2795

Hijmans, R. J. (2019). Introduction to the "geosphere" package (Version 1.5-10).

Hu, A., Ju, F., Hou, L., Li, J., Yang, X., Wang, H., et al. (2017a). Strong impact of anthropogenic contamination on the co-occurrence patterns of a riverine microbial community. *Environ. Microbiol.* 19, 4993–5009. doi: 10.1111/1462-2920.13942

Hu, A., Wang, H., Yang, X., Hou, L., Li, J., Li, S., et al. (2017b). Seasonal and spatial variations of prokaryoplankton communities in a salinity-influenced watershed, China. *FEMS Microbiol. Ecol.* 93:fix093. doi: 10.1093/femsec/fix093

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton, New Jersey: Princeton University Press.

Isabwe, A., Ren, K., Wang, Y., Peng, F., Chen, H., and Yang, J. (2019). Community assembly mechanisms underlying the core and random bacterioplankton and microeukaryotes in a river–reservoir system. *Water* 11:1127. doi: 10.3390/w11061127

Jeanbille, M., Gury, J., Duran, R., Tronczynski, J., Agogué, H., Ben Saïd, O., et al. (2016). Response of core microbial consortia to chronic hydrocarbon contaminations in coastal sediment habitats. *Front. Microbiol.* 7:1637. doi: 10.3389/fmicb.2016.01637

Jiang, L., and Patel, S. N. (2008). Community assembly in the presence of disturbance: a microcosm experiment. *Ecology* 89, 1931–1940. doi: 10.1890/07-1263.1

Jiao, S., and Lu, Y. (2020). Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ. Microbiol.* 22, 1052–1065. doi: 10.1111/1462-2920.14815

Jiao, S., Yang, Y., Xu, Y., Zhang, J., and Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J.* 14, 202–216. doi: 10.1038/s41396-019-0522-9

Kraft, N. J., Adler, P. B., Godoy, O., James, E. C., Fuller, S., and Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Funct. Ecol.* 29, 592–599. doi: 10.1111/1365-2435.12345

Langenheder, S., and Székely, A. J. (2011). Species sorting and neutral processes are both important during the initial assembly of bacterial communities. *ISME J.* 5, 1086–1094. doi: 10.1038/ismej.2010.207

Lennon, J. T., and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* 9, 119–130. doi: 10.1038/nrmicro2504

Levins, R. (1968). *Evolution in Changing Environments: Some Theoretical Explorations (MPB-2)*. Princeton, New Jersey: Princeton University Press.

Li, Y., Gao, Y., Zhang, W., Wang, C., and Wu, H. (2019). Homogeneous selection dominates the microbial community assembly in the sediment of the three gorges reservoir. *Sci. Total Environ.* 690, 50–60. doi: 10.1016/j.scitotenv.2019.07.014

Liao, J., Cao, X., Wang, J., Zhao, L., Sun, J., Jiang, D., et al. (2017). Similar community assembly mechanisms underlie similar biogeography of rare and abundant bacteria in lakes on Yungui Plateau, China. *Limnol. Oceanogr.* 62, 723–735. doi: 10.1002/lno.10455

Liao, J., Cao, X., Zhao, L., Wang, J., Gao, Z., Wang, M. C., et al. (2016a). The importance of neutral and niche processes for bacterial community assembly differs between habitat generalists and specialists. *FEMS Microbiol. Ecol.* 92:fiw174. doi: 10.1093/femsec/fiw174

Liao, J., Zhao, L., Cao, X., Sun, J., Gao, Z., Wang, J., et al. (2016b). Cyanobacteria in lakes on Yungui Plateau, China are assembled via niche processes driven by water physicochemical property, lake morphology and watershed land-use. *Sci. Rep.* 6:36357. doi: 10.1038/srep36357

Lindh, M. V., Sjöstedt, J., Ekstam, B., Casini, M., Lundin, D., Hugerth, L. W., et al. (2017). Metapopulation theory identifies biogeographical patterns among core and satellite marine bacteria scaling from tens to thousands of kilometers. *Environ. Microbiol.* 19, 1222–1236. doi: 10.1111/1462-2920.13650

Lindström, E. S., and Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* 4, 1–9. doi: 10.1111/j.1758-2229.2011.00257.x

Lindström, E. S., and Östman, Ö. (2011). The importance of dispersal for bacterial community composition and functioning. *PLoS One* 6:e25883. doi: 10.1371/journal.pone.0025883

Liu, K., Hou, J., Liu, Y., Hu, A., Wang, M., Wang, F., et al. (2019c). Biogeography of the free-living and particle-attached bacteria in Tibetan lakes. *FEMS Microbiol. Ecol.* 95:fiz088. doi: 10.1093/femsec/fiz088

Liu, K., Liu, Y., Hu, A., Wang, F., Chen, Y., Gu, Z., et al. (2020). Different community assembly mechanisms underlie similar biogeography of bacteria and microeukaryotes in Tibetan lakes. *FEMS microbiol. ecol.* 96:fiaa071.

Liu, K., Liu, Y., Jiao, N., Zhu, L., Wang, J., Hu, A., et al. (2016). Vertical variation of bacterial community in Nam Co, a large stratified lake in central Tibetan Plateau. *Antonie Van Leeuwenhoek* 109, 1323–1335. doi: 10.1007/s10482-016-0731-4

Liu, J., Meng, Z., Liu, X., and Zhang, X.-H. (2019a). Microbial assembly, interaction, functioning, activity and diversification: a review derived from community compositional data. *Mar. Life Sci. Technol.* 1, 1–17. doi: 10.1007/s42995-019-00004-3

Liu, L., Yang, J., Yu, Z., and Wilkinson, D. M. (2015). The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *ISME J.* 9, 2068–2077. doi: 10.1038/ismej.2015.29

Liu, K., Yao, T., Liu, Y., Xu, B., Hu, A., and Chen, Y. (2019b). Elevational patterns of abundant and rare bacterial diversity and composition in mountain streams in the southeast of the Tibetan Plateau. *Sci. Earth Sci.* 62, 853–862. doi: 10.1007/s11430-018-9316-6

Logares, R., Lindström, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J., et al. (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.* 7, 937–948. doi: 10.1038/ismej.2012.168

Magoč, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Magurran, A. E., and Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714–716. doi: 10.1038/nature01547

Masanori, F., Hisaya, K., Tomoya, I., et al. (2011). Dissolved organic carbon as major environmental factor affecting bacterioplankton communities in mountain lakes of eastern Japan. *Microb. Ecol.* 63, 496–508. doi: 10.1007/s00248-011-9983-8

Mikhailov, I. S., Zakharova, Y. R., Bukin, Y. S., Galachyants, Y. P., Petrova, D. P., Sakirko, M. V., et al. (2019). Co-occurrence networks among bacteria and microbial eukaryotes of lake baikal during a spring phytoplankton bloom. *Microb. Ecol.* 77, 96–109. doi: 10.1007/s00248-018-1212-2

Mingkun, L., Xue, H., Jun, T., Huifeng, Z., and Xiaohui, B. (2020). Mutual environmental drivers of the community composition, functional attributes and co-occurrence patterns of bacterioplankton in the composite aquatic ecosystem of Taihu watershed in China. *FEMS Microbiol. Ecol.* 8:fiaa137. doi: 10.1093/femsec/fiaa137

Mo, Y., Zhang, W., Wilkinson, D. M., Yu, Z., Xiao, P., and Yang, J. (2020). Biogeography and co-occurrence patterns of bacterial generalists and specialists in three subtropical marine bays. *Limnol. Oceanogr.* 66, 793–508. doi: 10.1002/lno.11643

Mo, Y., Zhang, W., Yang, J., Lin, Y., Yu, Z., and Lin, S. (2018). Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. *ISME J.* 12, 2198–2210. doi: 10.1038/s41396-018-0153-6

Morriën, E., Hannula, S. E., Snoek, L. B., Helmsing, N. R., Zweers, H., De Hollander, M., et al. (2017). Soil networks become more connected and take up more carbon as nature restoration progresses. *Nat. Commun.* 8:14349. doi: 10.1038/ncomms14349

Morrissey, E. M., Mau, R. L., Hayer, M., Liu, X.-J., Schwartz, E., Dijkstra, P., et al. (2019). Evolutionary history constrains microbial traits across environmental variation. *Nat. Ecol. Evol.* 3, 1064–1069. doi: 10.1038/s41559-019-0918-y

Nekola, J. C., and White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *J. Biogeogr.* 26, 867–878. doi: 10.1046/j.1365-2699.1999.00305.x

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356. doi: 10.1128/MMBR.00051-12

Ofiţeru, I. D., Lunn, M., Curtis, T. P., Wells, G. F., Criddle, C. S., Francis, C. A., et al. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *Proc. Natl. Acad. Sci. U. S. A.* 107, 15345–15350. doi: 10.1073/pnas.1000604107

Pandit, S. N., Kolasa, J., and Cottenie, K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology* 90, 2253–2262. doi: 10.1890/08-0851.1

Pedrós-Alió, C. (2012). The rare bacterial biosphere. *Annu. Rev. Mar. Sci.* 4, 449–466. doi: 10.1146/annurev-marine-120710-100948

Peng, G. S., and Wu, J. (2016). Optimal network topology for structural robustness based on natural connectivity. *Physica A* 443, 212–220. doi: 10.1016/j.physa.2015.09.023

Pester, M., Bittner, N., Deevong, P., Wagner, M., and Loy, A. (2010). A 'rare biosphere'microorganism contributes to sulfate reduction in a peatland. *ISME J.* 4, 1591–1602. doi: 10.1038/ismej.2010.75

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.* 8, 732–740. doi: 10.1111/j.1462-2920.2005.00956.x

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Fredrickson, J. K., and Konopka, A. E. (2015). Estimating and mapping ecological processes influencing microbial community assembly. *Front. Microbiol.* 6:370. doi: 10.3389/fmicb.2015.00370

Unterseher, M., Jumpponen, A., Oepik, M., Tedersoo, L., Moora, M., Dormann, C. F., et al. (2011). Species abundance distributions and richness estimations in fungal metagenomics–lessons learned from community ecology. *Mol. Ecol.* 20, 275–285. doi: 10.1111/j.1365-294X.2010.04948.x

Van Der Gast, C. J., Walker, A. W., Stressmann, F. A., Rogers, G. B., Scott, P., Daniels, T. W., et al. (2011). Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J.* 5, 780–791. doi: 10.1038/ismej.2010.175

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wang, S., Wang, X., Han, X., and Deng, Y. (2018). Higher precipitation strengthens the microbial interactions in semi-arid grassland soils. *Global Ecol. Biogeogr.* 27, 570–580. doi: 10.1111/geb.12718

Wu, W., Lu, H.-P., Sastri, A., Yeh, Y.-C., Gong, G.-C., Chou, W.-C., et al. (2018). Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME J.* 12, 485–494. doi: 10.1038/ismej.2017.183

Yao, Z., Du, S., Liang, C., Zhao, Y., Dini-Andreote, F., Wang, K., et al. (2019). Bacterial community assembly in a typical estuarine marsh with multiple environmental gradients. *Appl. Environ. Microbiol.* 85:e02602. doi: 10.1128/AEM.02602-18

Yuan, M. M., Guo, X., Wu, L., Zhang, Y., Xiao, N., Ning, D., et al. (2021). Climate warming enhances microbial network complexity and stability. *Nat. Clim. Chang.* 11, 343–348. doi: 10.1038/s41558-021-00989-9

Zhang, G., Xie, H., Kang, S., Yi, D., and Ackley, S. F. (2011). Monitoring lake level changes on the Tibetan Plateau using ICESat altimetry data (2003–2009). *Remote Sens. Environ.* 115, 1733–1742. doi: 10.1016/j.rse.2011.03.005

Zhang, Y., Zhao, Z., Dai, M., Jiao, N., and Herndl, G. J. (2014). Drivers shaping the diversity and biogeography of total and active bacterial communities in the South China Sea. *Mol. Ecol.* 23, 2260–2274. doi: 10.1111/mec.12739

Zhu, H. Z., Zhang, Z. F., Zhou, N., Jiang, C. Y., Wang, B. J., Cai, L., et al. (2019). Diversity, distribution and co-occurrence patterns of bacterial communities in a karst cave system. *Front. Microbiol.* 10:1726. doi: 10.3389/fmicb.2019.01726

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 04 October 2021
doi: 10.3389/fmicb.2021.744834

Energy Availability Determines Strategy of Microbial Amino Acid Synthesis in Volatile Fatty Acid–Fed Anaerobic Methanogenic Chemostats

Jian Yao, Yan Zeng, Miaoxiao Wang and Yue-Qin Tang**

College of Architecture and Environment, Sichuan University, Chengdu, China

In natural communities, microbes exchange a variety of metabolites (public goods) with each other, which drives the evolution of auxotroph and shapes interdependent patterns at community-level. However, factors that determine the strategy of public goods synthesis for a given community member still remains to be elucidated. In anaerobic methanogenic communities, energy availability of different community members is largely varied. We hypothesized that this uneven energy availability contributed to the heterogeneity of public goods synthesis ability among the members in these communities. We tested this hypothesis by analyzing the synthetic strategy of amino acids of the bacterial and archaeal members involved in four previously enriched anaerobic methanogenic communities residing in thermophilic chemostats. Our analyses indicate that most of the members in the communities did not possess ability to synthesize all the essential amino acids, suggesting they exchanged these essential public goods to establish interdependent patterns for survival. Importantly, we found that the amino acid synthesis ability of a functional group was largely determined by how much energy it could obtain from its metabolism in the given environmental condition. Moreover, members within a functional group also possessed different amino acid synthesis abilities, which are related to their features of energy metabolism. Our study reveals that energy availability is a key driver of microbial evolution in presence of metabolic specialization at community level and suggests the feasibility of managing anaerobic methanogenic communities for better performance through controlling the metabolic interactions involved.

Keywords: thermophilic methanogenic community, functional groups, available energy, amino acids synthesis strategy, amino acids exchange

OPEN ACCESS

Edited by:
Jianjun Wang,
Nanjing Institute of Geography and Limnology, China

Reviewed by:
Bin Cao,
Nanyang Technological University, Singapore
Meng Li,
Shenzhen University, China
Yong Nie,
Peking University, China

****Correspondence:***
Miaoxiao Wang
wangmx2014@pku.edu.cn
Yue-Qin Tang
tangyq@scu.edu.cn

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *21 July 2021*
Accepted: *30 August 2021*
Published: *04 October 2021*

Citation:
Yao J, Zeng Y, Wang M and Tang Y-Q (2021) Energy Availability Determines Strategy of Microbial Amino Acid Synthesis in Volatile Fatty Acid–Fed Anaerobic Methanogenic Chemostats. Front. Microbiol. 12:744834. doi: 10.3389/fmicb.2021.744834

INTRODUCTION

In most natural environments, microbial individuals rarely live alone but co-colonize with other species to form complex communities. Members in these communities are connected through intricate interaction networks. These interactions not only influence the growth and survival of individual member, but also scale up to determine the assembly and functions of the whole community (Zengler and Zaramela, 2018). Among diverse modes of microbial interactions, exchange of public goods (PGs) among different members is one of the most pervasive

(Mitri and Foster, 2013), in which one member secretes metabolites to environment that benefit other members in the community. In turn, this community member can also obtain production from others. Previous studies suggest that many biologically essential metabolites can be shared as PGs, such as amino acids (AAs) (Mee et al., 2014; Embree et al., 2015; Hubalek et al., 2017), vitamins (Croft et al., 2005; Rodionova et al., 2015), siderophores (Kramer et al., 2020), and other cofactors (Jones, 1967). Associated with PG sharing, several genomic investigations indicate that many microorganisms in diverse environments [such as methanogenic chemostats (Hubalek et al., 2017), oil reservoir (Liu et al., 2018), and human gut (Soto-Martin et al., 2020)] contain only a small set of genes that encode these public functions, so they must survive by exchanging PGs with other members. This phenomenon reflects that PG sharing is a main force to drive the microbial genomic evolution (Morris et al., 2011) and play important roles in governing the assembly of the community (Zengler and Zaramela, 2018). Nevertheless, studies also suggest that the retained public functions of these auxotrophies were highly diverse (Hubalek et al., 2017; Liu et al., 2018; Soto-Martin et al., 2020). Given a strain residing in a community, we still lack knowledge to explain why it possesses the ability to execute the observed set of public functions (Zengler and Zaramela, 2018). The cause of this Gordian knot is that whether autonomously producing a specific PG benefits the community member itself is determined by many unknown biotic and abiotic factors present in complex natural communities. Uncovering these factors is crucial for understanding microbial evolution at community scale, as well as uncovering the assembly rule of microbial community with complex interaction networks.

The Black Queen Hypothesis (BQH) formulated by Morris et al. (2012) emphasizes that whether retaining (or loss of) a specific public function is selectively favored is determined by the traits of this function, including its energy cost, its essentiality, and its leakiness. A number of studies validated this prediction through theoretical models (Oliveira et al., 2014; Estrela et al., 2016; Mas et al., 2016; Zomorrodi and Segrè, 2017; Pacheco et al., 2019; Meijer et al., 2020; Wang et al., 2021), as well as experiments using synthetic microbial communities composed of engineered microbial auxotrophies (Mee et al., 2014; Pande et al., 2014; Cooper et al., 2019). These studies indicate that a leaky and essential public function was easier to be retained if the energy cost of performing this function is lower. Several studies extended the BQH framework to explain the scenario when multiple public functions can be shared, suggesting that the costs of all the function would profoundly affect the strategies of the strains involved (Mee et al., 2014; Embree et al., 2015). However, these studies simply assumed that the community members possessed similar metabolic functions except for their specific public functions. In addition, many studies explored the PG exchange strategies among community members in complex natural communities, such as anaerobic hydrocarbon-degrading community (Embree et al., 2015), freshwater mixed culture (Garcia et al., 2015), and kefir microbial community (Blasche et al., 2021). These studies showed that complex PG exchange relationships between community members were important causes for the stable coexistence of diverse species. Unfortunately, the factors that drive to these complex modes of PG exchange have not been fully revealed.

In complex natural community, the fitness of different members are highly varied and affected by complex factors, such as ecological niche (Neis et al., 2015), energy availability (Zengler and Zaramela, 2018), temporal shifts in community composition (Merchant and Helmann, 2012), and spatial structure (Allen et al., 2013). The fitness variation may also influence the strategies of PG production of the members involved. For example, if a community member possesses an ecological niche that could gain more energy to support PG production, it may possess ability to perform more public functions, even for those costly ones, because producing these PGs may only consume a relative smaller set of energy accounting for the overall energy it could obtain. Extending the framework of BQH, this hypothesis can be formulated as follows:

$$p \propto \frac{ea}{c} \quad (1)$$

This relationship depicts that, in a complex community, the probability of whether a member can perform a public function (p) is (1) negatively correlated with the energy cost of the function (c) and (2) positively correlated with the overall energy availability for a member (ea).

In this study, we set out to test this hypothesis in anaerobic methanogenic communities (AMCs). Compared with other communities, the available energy is relatively limited in the AMC system and thus plays fundamental roles in driving the evolution of the community members involved (Lyu et al., 2018), as well as being associated with the metabolic interactions among these members (Jackson and McInerney, 2002; Nobu et al., 2020). In our previous works, we enriched four AMCs in thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, or isovalerate as sole carbon source (accordingly, the four chemostats were simplified as ATL, PTL, BTL, and VTL thereafter) (Zheng et al., 2019; Chen et al., 2020a,b). Because AA is one type of important public functions that have been reported in many microbial communities [e.g., in ocean (Barbara and Mitchell, 2003) and in human gut (Perna et al., 2019), as well as other AMC systems (Embree et al., 2015; Hubalek et al., 2017)], we analyzed the abilities of AA synthesis of syntrophic acetate-oxidizing bacteria (SAOBs) involved in our enriched AMCs and found that AA synthesis ability was deficient in most of the SAOBs in our AMCs (Zeng et al., 2021). However, SAOBs represent only a subset of community members in AMCs. The PG synthesis strategies of other groups, as well as the underlying determinants, still remain poorly understood.

Several functional groups dominated our four AMCs. In ATL, *Methanosarcina* can directly utilize acetate for methanogenesis through the acetotrophic pathway. In addition, the H_2 and CO_2 generated from acetate oxidation by SAOBs can also be utilized by *Methanosarcina*, *Methanomassiliicoccus*, *Methanothermobacter*, *Methanobacterium*, and *Methanoculleus* to produce methane (**Supplementary Table 1**). In PTL, one additional group of bacteria, syntrophic propionate-degrading bacteria (SPOB), is active to degrade propionate to acetate (**Supplementary Table 1** and **Figure 1B**). Similarly, syntrophic

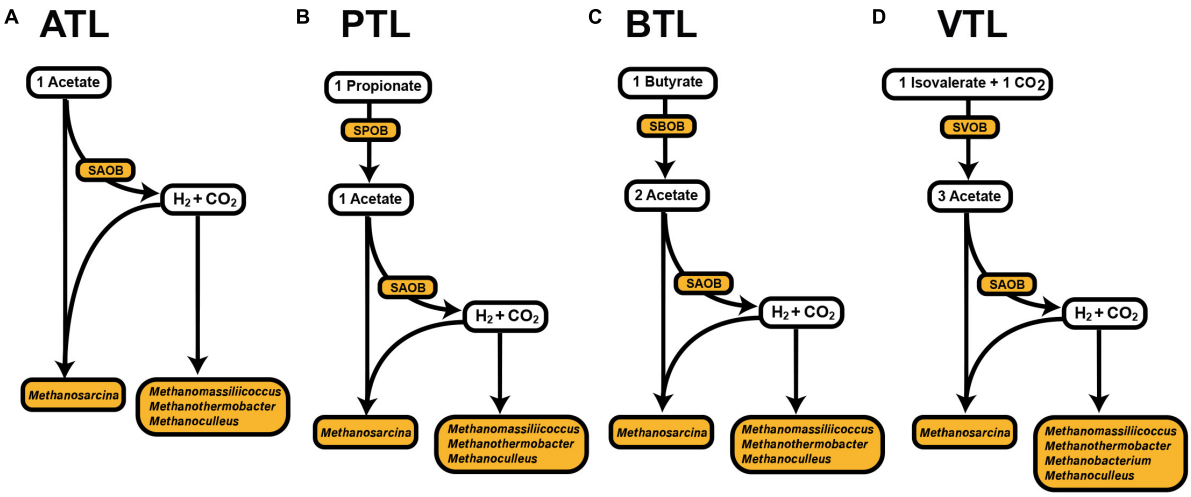


FIGURE 1 | Proposed main carbon flow in methanogenic process in the four anaerobic chemostats (Chen et al., 2020a, b; Zheng et al., 2019). The ATL **(A)**, PTL **(B)**, BTL **(C)**, and VTL **(D)** represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively.

butyrate-degrading bacteria (SBOB) degrade butyrate to acetate in BTL (**Supplementary Table 1** and **Figure 1C**), whereas syntrophic isovalerate-degrading bacteria (SVOB) degrade isovalerate to acetate in VTL (**Supplementary Table 1** and **Figure 1D**). In particular, microorganisms from different functional groups possess different energy availability, which is determined by the substrate availability (i.e., the fatty acids) in different environments, as well as their features of energy metabolisms. We thus expected to know whether the strategies of AA synthesis of the members in these functional groups are correlated with their energy availability, as well as the energy cost of the AA synthesis, which can be directly applied to test our hypothesis.

To this end, in this study, we analyzed the ability of AA synthesis of the communities in our methanogenic chemostats using our previously assembled metagenomic genomes, as well as the associated metatranscriptomic data.

MATERIALS AND METHODS

Settings of the Chemostats Containing the Anaerobic Methanogenic Communities

In our previous work, four thermophilic (55°C) methanogenic chemostats (ATL, PTL, BTL, and VTL, fed with acetate, propionate, butyrate, and isovalerate as sole carbon sources, respectively) were built and operated at a low dilution rate (0.025 day^{-1}, hydraulic retention time (HRT) = 40 days). The ATL and PTL were seeded with sludge from an anaerobic digester treating kitchen waste, whereas BTL and VTL were seeded with sludge from an anaerobic reactor treating swine manure. The total organic carbon (TOC) of synthetic wastewater fed to each chemostat was 8,000 mg L^{-1} (**Supplementary Figure 1**). These chemostats were stably operated over 200 days, and the biogas production was maintained at a constant level, and no fatty acids were accumulated in each chemostat. During the stable operation of these reactors, sludge samples were collected from ATL (days 306 and 307), PTL (day 223, 293, and 318), BTL (day 251 and 252), and VTL (day 295 and 296). The details about metagenomic and metatranscriptomic sample preparation, sequencing, assembly, and binning are available in the **Supplementary Methods**. More details about the operation and performance of chemostats are described in previous works (Chen et al., 2020a,b; Zeng et al., 2021).

Analysis of Amino Acid Synthesis Strategy in the Extracted Metagenome-Assembled Genomes

Two hundred twenty-seven high-quality metagenome-assembled genomes (MAGs, contamination < 10%, completeness > 70%) were recovered in our previous study (see **Supplementary Figure 2** for details about MAGs) (Chen et al., 2020a,b; Zeng et al., 2021), and raw sequence data were accessible at http://bigd.big.ac.cn/gsa (accession no. CRA004311). On the basis of these data, the gene sets related to AA synthesis were searched in these MAGs, by performing functional annotation on KEGG Automatic Annotation Server (Moriya et al., 2007). If an MAG had all the genes involved in the synthetic pathway of an AA (see **Supplementary Figure 3** for the complete synthetic pathway of all the AAs), it was defined to possess the ability to synthesize this AA at genomic level, denoted by the blue box in **Supplementary Figure 4**.

Metatranscriptomic sequencing of these communities was also performed in our previous study (raw sequence data are accessible at http://bigd.big.ac.cn/gsa, accession no. CRA004311). RPKM-NM (Zheng et al., 2019; Chen et al., 2020a,b) values of all the genes involved in AA synthesis were extracted from our previous datasets (**Supplementary Tables 2–5**). To quantify the

activity of synthesis of an AA in our MAGs at transcriptional level ($activity_{a.a.}$), RPKM-NM values of all genes involved in its synthetic pathway were summed and divided by the number of genes, formulated as follows:

$$Activity_{a.a.} = \frac{\sum_1^n RPKM - NM_{gene\ i}}{N(steps)} \quad (2)$$

In order to compare the overall AA synthesis activity among different MAGs, we defined the relative AA synthesis activity ($RA_{a.a.}$) as follows:

$$RA_{a.a.} = Activity_{a.a} \times \frac{MAG_i\ reads}{\sum_1^n MAG_i\ reads} \quad (3)$$

Here, $MAG_i\ reads$ represents the total number of reads that can be mapped to metatranscriptomic data for all genes in the *i*th MAG.

We quantified the relative contribution of AA synthesis of each functional group by defining $Contribution_{f.g.}$ based on $RA_{a.a.}$ as follows:

$$Contribution_{f.g.} = \frac{\sum_1^j RA_{a.a.}\ of\ MAG_i}{\sum_1^n RA_{a.a.}\ of\ MAG_i} \quad (4)$$

Here, j represents the number of MAGs within a functional group; n represents the number of all MAGs.

Statistical Analysis

The significance analysis of the difference in AA synthesis ability among different functional groups is performed using the "stats" package (4.0.2) of R (4.0.2) through analysis of variance. Correlation analyses were performed using Pearson correlation method using the cor.test() function of the stats package (4.0.2) in R (4.0.2).

RESULTS

Overall Strategy of Microbial Amino Acid Synthesis in the Four Anaerobic Methanogenic Communities

In our previous study, we recovered 57, 63, 52, and 55 high-quality MAGs from four enriched AMCs residing in ATL, PTL, BTL, and VTL, respectively. Nineteen MAGs were identified as methanogens at genus level, including four *Methanomassiliicoccus* (ATL89, PTL47, BTL15, and VTL23), four *Methanosarcina* (ATL14, PTL77, BTL39, and VTL53), one *Methanobacterium* (VTL28), six *Methanothermobacter* (ATL45, PTL2, PTL149, BTL70, VTL26, and VTL90), and four *Methanoculleus* (ATL103, PTL54, BTL76, and VTL77). Forty MAGs (out of 208 bacterial MAGs in total) were identified as potential SAOBs, because these MAGs contain genes encoding the WL or WL–glycine cleavage pathway, as well as complementary NADPH reoxidation and H2/formate-generating enzymes (Zeng et al., 2021). Similarly, one, two, and one MAGs were identified as SPOB, SBOB, and SVOB, respectively, including *Pelotomaculum* (PTL62) (Chen et al., 2020b), unclassified *Clostridiales* (VTL56) (Chen et al., 2020a), *Syntrophomonas* (BTL36) (McInerney et al., 1981), and *Syntrophothermus* (BTL6) (Luo et al., 2002). Meanwhile, we also found that many MAGs were highly abundant in the four AMCs, but did not contain the genes involved in the above core methanogenic pathways, which we defined as "noncore functional bacteria," including 45 MAGs in ATL, 55 MAGs in PTL, 40 MAGs in BTL, and 42 MAGs in VTL. Notably, in each AMC, none of the community members contains all the genes responsible for the synthesis of 20 essential AAs (**Supplementary Figure 4**). The prevalent lack of AA synthesis capacity among MAGs suggested that metabolic interdependency relying on PG sharing is essential for the survival of these MAGs in our four AMCs.

In order to depict the complex interaction modes of AA sharing in our AMCs, we evaluated the potential complementarity in AA synthesis between every two MAGs in each AMC (**Figure 2** and see **Supplementary Methods** for details about calculation). When an MAG contains all the genes encoding the enzymes used for the synthesis of an AA, it potentially "contributes" this AA to support the growth of other MAGs. In contrast, if an MAG does not contain these genes, they must "obtain" AAs produced by other MAGs. For each MAG, we counted how many AAs it could "contribute" to each MAG as the "contribute strength" (denoted as a link starting from the red patch in **Figure 2**). Similarly, we counted how many AAs it must "obtain" from one of the other MAGs as the "obtain strength" (denoted as a link ending in the blue patch in **Figure 2**). Then, the potential interactive connections between every pair of MAGs in each AMC were built. As shown in **Figure 2**, several MAGs tended to "contribute" more to others, whereas several MAGs "obtain" most of AAs from other members. To quantify this feature; we defined the contributing index (*ci*) of an MAG as the ratio of the summary of "contribute strength" to the summary of the "obtain strength" (see **Supplementary Figure 5** for details). When the *ci* value (**Supplementary Table 6**) of an MAG was over 1, we defined it as a "contributor." In contrast, MAGs with a *ci* less than 1 were defined as "beneficiaries." We found that the number of "contributors" MAGs was lower than that of "beneficiaries" MAGs in every AMC (17 against 40 in ATL, 23 against 40 in PTL, 22 against 30 in BTL, and 21 against 34 in VTL), and the overall abundance of "contributors" was also lower than that of "beneficiaries" in ATL (32 $<$ 59%), PTL (35 $<$ 52%), BTL (37 $<$ 59%), and VTL (42 $<$ 49%). However, at transcriptional level, the total activity of "contributors" was higher than that of "beneficiaries" in ATL (50 $>$ 42%), PTL (56 $>$ 31%), and VTL (62 $>$ 30%), except in BTL (27 $<$ 71%). These results suggest that in AMCs, a minority of MAGs contribute to producing the majority of AAs to support the growth of the whole community, and these MAGs exhibit higher metabolic activities, which may facilitate the production of AAs. This phenomenon is very similar to the findings in other natural communities (Embree et al., 2015; Liu et al., 2018).

Effect of Biosynthetic Cost on the Strategies of Amino Acid Synthesis

Previous studies showed that, in natural communities, the frequency of members that could synthesize an AA is negatively correlated with its energy cost (Mee et al., 2014). Thus, we

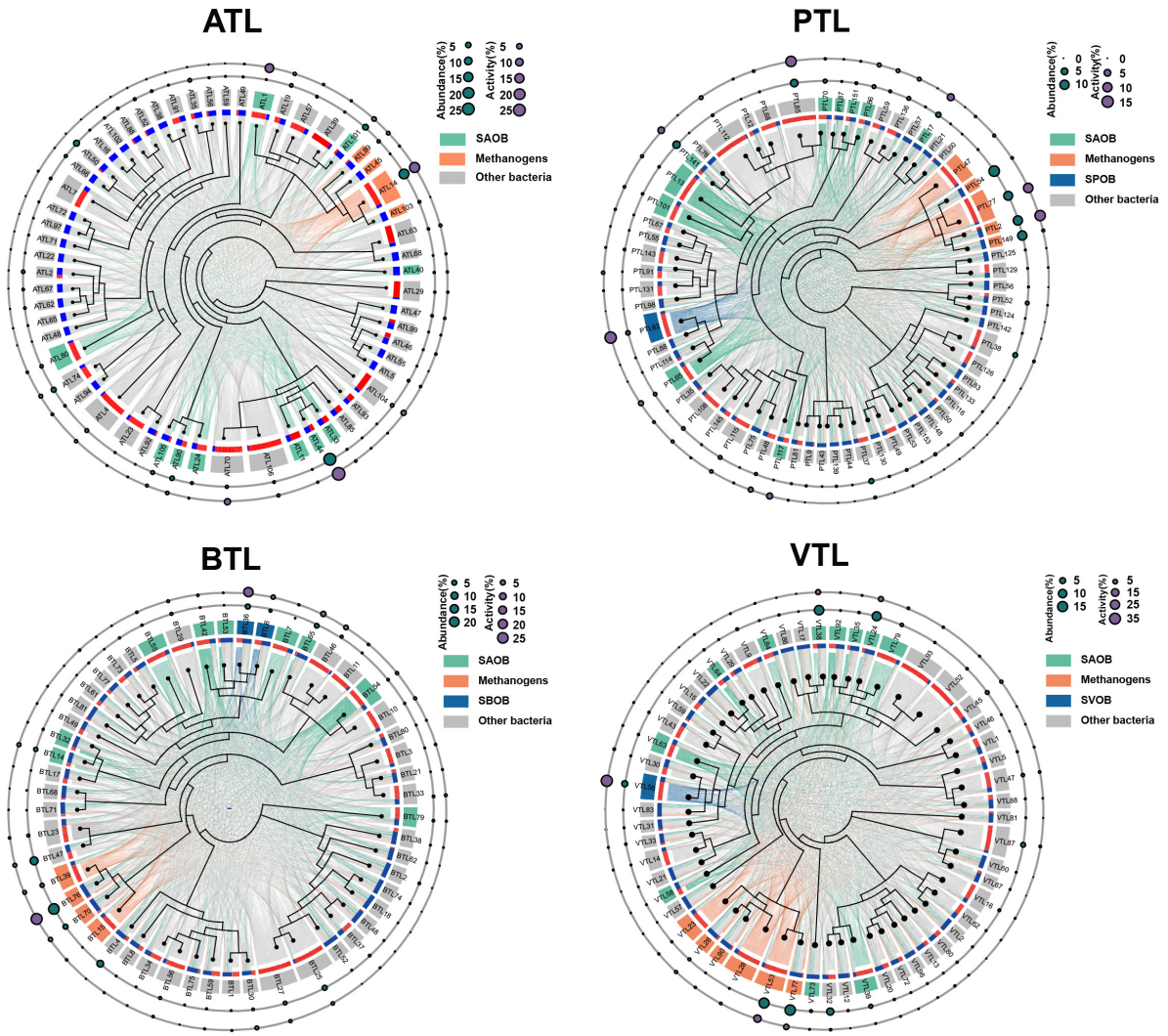


FIGURE 2 | Potential complementarity in amino acid (AA) synthesis among the community members in each chemostat. The ATL, PTL, BTL, and VTL represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively. The internal tree diagram depicts the phylogenetic relationship among the community members. The outer colored donuts indicate different functional groups, as shown in the left of each figure. All the members are connected by the lines reflecting their functional complementarity. Each link starts from an metagenome assembled genomes (MAG) that "contributes" AAs (labeled with the red patch) and ends with an MAG that "obtains" the AAs (labeled with the blue patch). The strength of the relationship is characterized by the thickness of the line. An additional diagram is given in **Supplementary Figure 5**. Accordingly, the width of red patch for a given MAG indicates its potential to "contribute" AAs to the community, whereas the width of blue patch for a given MAG indicates its potential to "cheat" AAs from the community. Dark green and purple bubbles in the outside concentric rings indicate the abundance and overall activity of community members, respectively.

next tested whether this observation was also the case in our AMCs. To evaluate the energy cost of AAs, we followed the framework proposed by Akashi and Gojobori (2002), which assessed the net ATP consumption in the synthetic pathway of each AA. Our results indicated that, generally, the number of MAGs that synthesize the expensive AAs (e.g., histidine and tryptophan) was less than that of MAGs synthesize the inexpensive ones at genome level and transcriptome level (e.g., glutamate and glutamine; **Figures 3A,B**), consistent with the findings in previous studies. Statistically, the number of MAGs containing synthesis genes of an AA was negatively correlated with the function cost of the AA (**Supplementary Figure 6A**). Nevertheless, 77 and 70 MAGs possessed genes to synthesize phenylalanine and tyrosine, respectively, which was higher than the number of MAGs that possessed genes to synthesize relatively inexpensive histidine (38 MAGs) and isoleucine (46 MAGs) (**Supplementary Figure 6A**), seemly contradictory to the previously proposed rule. Interestingly, we found that the genes encoding AA synthesis in these 77 and 70 MAGs were lowly transcribed (**Supplementary Figure 6B**), suggesting that MAGs in our AMCs adapted a regulatory strategy to decrease the synthetic activity of these two costly AAs and thus saved energy for their better survival. Combining metagenomic and metatranscriptomic data, we defined an "expression ratio" to

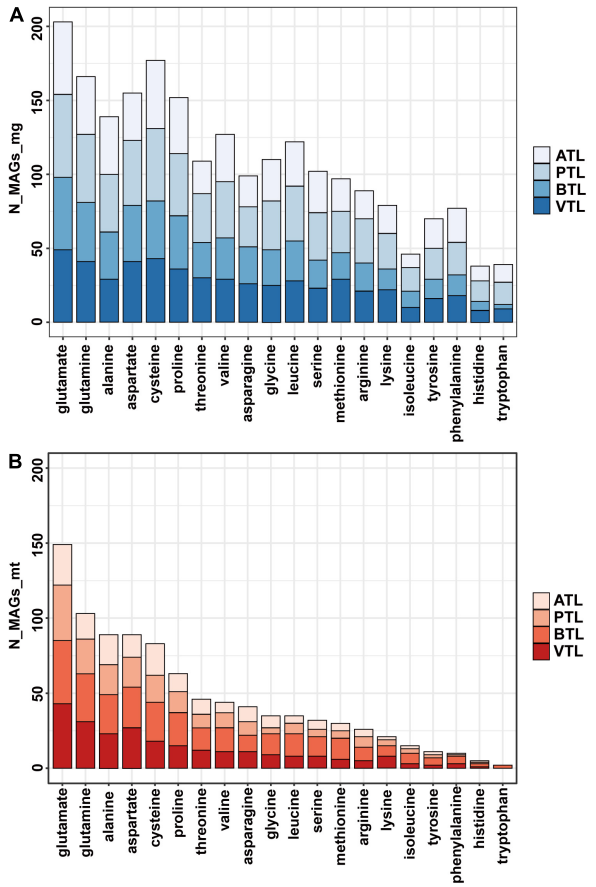


FIGURE 3 | The total number of MAGs which can synthesize each amino acid in different chemostats at genome level **(A)** and transcriptome level **(B)**. The ATL, PTL, BTL, and VTL represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively. "N_MAGs_mg" represents the number of MAGs that can synthesize each amino acid at genome level. "N_MAGs_mt" represents the number of MAGs that can synthesize each amino acid at transcriptomic level.

quantify the synthetic activity of synthesizing an AA in our communities (see legends to **Figure 4** for details). Pearson correlation test showed a significant negative correlation between the expression ratio of AAs and their biosynthesis cost in ATL (p = 0.03, r = −0.78), PTL (p = 0.0009, r = −0.68), BTL (p = 0.07, r = −0.40), and VTL (p = 0.0002, r = −0.75; **Figure 4**). These results indicate that our AMCs have been optimized at transcriptional level to reduce the metabolic burden. Therefore, function cost is a key driver that affects the strategy of PG production of the MAGs in our AMCs.

Strategies of Amino Acid Synthesis of Different Functional Groups

We next focused on the second point of our hypothesis, that is, to test whether the strategies of AA synthesis of the members in each function group were correlated with its energy availability. In general, the relative contributions of different functional groups to each AA were largely varied (**Figure 5** and **Supplementary Figure 7**). SAOBs and methanogens made major contributions to the synthesis of most AAs in ATL, but their contributions reduced in other chemostats, especially in PTL and BTL. As acetate is the main energy source for SAOBs and methanogens (Zheng et al., 2019; Chen et al., 2020a,b), this phenomenon might be related to the differences in the energy (acetate) availability in different chemostats. Under the same TOC fed to each chemostat, as acetate was supplied as the substrate in ATL, the energy (acetate) availability of the methanogenic group (SAOBs and methanogens) in ATL was higher than that in other chemostats, which enabled SAOBs and methanogens in ATL to maintain the highest activity of AA synthesis. In PTL, BTL, and VTL, acetate was generated as an intermediate of carbon metabolism. Theoretically, one-, two-, and three-molecule acetate can be produced from one-molecule propionate, butyrate, and isovalerate by SPOB, SBOB, and SVOB, respectively (**Figure 1**). Therefore, the carbon converting ratios of propionate to acetate is 2/3, that of butyrate to acetate is 4/4, and that for isovalerate is 6/5. As a result, the accessibility of acetate in VTL was higher than that in PTL and BTL, which enabled SAOBs and methanogens in VTL to obtain more energy from methanogenic metabolism to maintain the higher activity of AA synthesis than that in PTL and BTL. Similarly, the SPOB, SBOB, and SVOB made important contributions to synthesize several costly AAs (histidine, tyrosine, and phenylalanine), which may also be due to their higher energy availability deriving from the direct degradation of the supplying carbon sources (i.e., propionate, butyrate, or isovalerate). In summary, the AA synthesis ability of these functional groups was affected by their energy availability, consistent with the prediction of our hypothesis.

Moreover, we also found that the noncore functional group contributed to synthesizing every AA, especially playing a leading role in the synthesis of those costly AAs, such as histidine, tyrosine, phenylalanine, and tryptophane (**Figure 5**). This result suggests that those taxa that do not contribute to core metabolism of anaerobic methanogenesis are indispensable in PG synthesis to support other community members, which may explain why these groups always possess considerable abundance in AMCs (Rivière et al., 2009; Krakat et al., 2011; Tang et al., 2011).

Strategy of Amino Acid Synthesis Within Different Functional Groups

Despite common features of AA synthesis strategies of the community members in each functional group, our further analyses also indicate that, within a same functional group, strategies of AA synthesis are also considerably different among members (**Supplementary Figure 4**). In order to explore whether these intragroup differences were also associated with energy availability, we further compared the AA synthesis ability of MAGs within each functional group.

Amino Acid Synthesis in Syntrophic Acetate Oxidation Bacterias

To investigate the factors that affect the strategy of AA synthesis in SAOBs, we first arranged the strategy of AA synthesis of the MAGs of SAOBs according to their phylogenetic

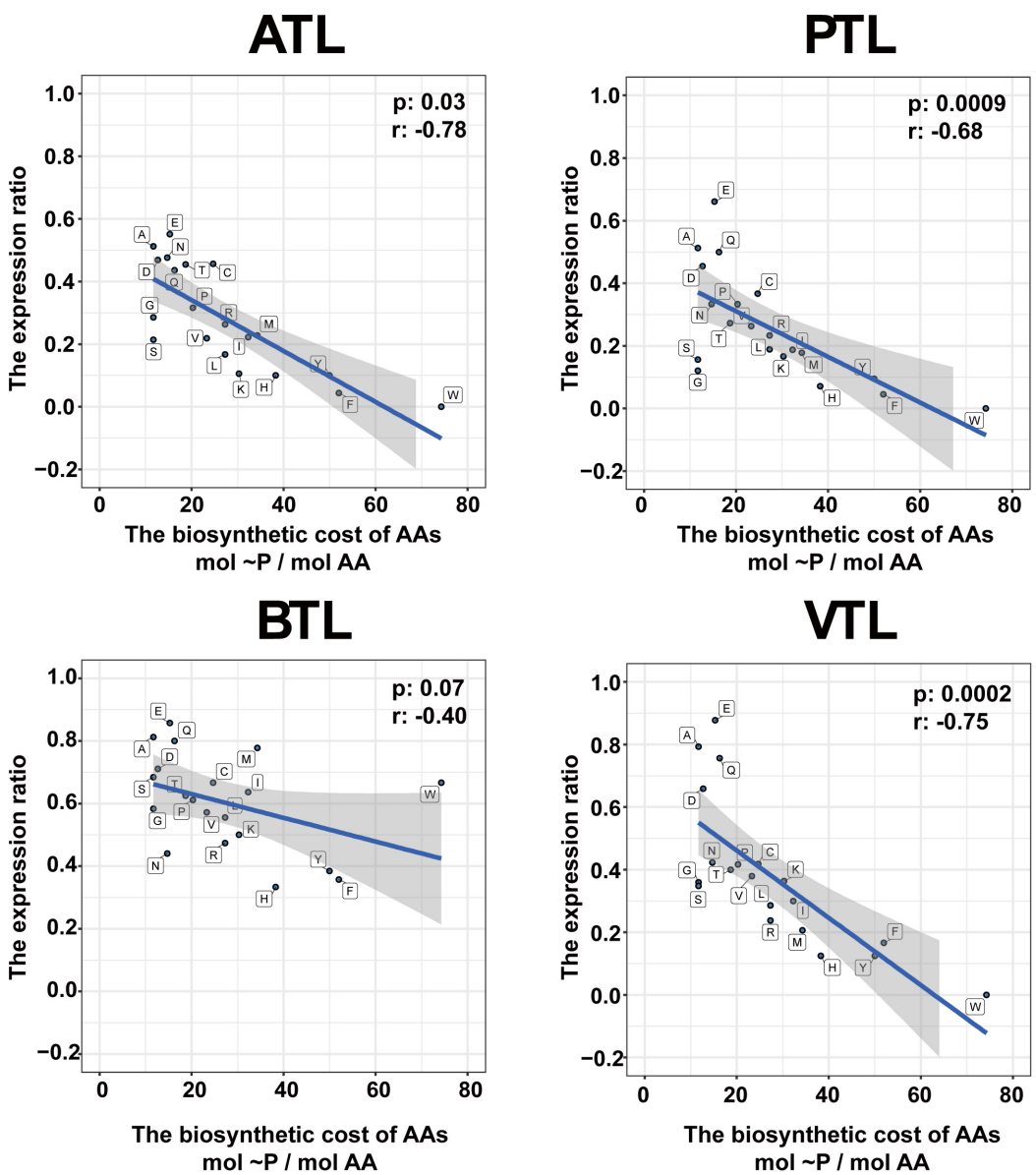


FIGURE 4 | Relationship between the amino acid expression ratio and biosynthetic cost of each amino acid. The ATL, PTL, BTL, and VTL represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively. The shaded area represents the 95% confidence region. The biosynthetic cost of each amino acid is represented by the number of high-energy phosphate bonds sacrificed in the biosynthetic pathway (Akashi and Gojobori, 2002). The *expression ratio* is defined as follows: $Expression\ ratio = \frac{N(NAG_{MT})}{N(MAG_{MG})}$. $N(MAG_{MT})$ represents the number of MAGs in which the genes responsible for the synthesis of the corresponding AA were actively transcribed. $N(MAG_{MG})$ is the number of MAGs that contain all the genes required for synthesizing the corresponding AA. Abbreviations of the amino acids: alanine (A), arginine (R), asparagine (N), aspartate (D), cysteine (C), glutamine (Q), glutamate (E), glycine (G), histidine (H), isoleucine (I), leucine (L), lysine (K), methionine (M), phenylalanine (F), proline (P), serine (S), threonine (T), tryptophan (W), tyrosine (Y), and valine (V).

relationships (**Figure 6**). We found that higher-cost AAs (e.g., histidine, isoleucine, tyrosine, and phenylalanine) were synthesized by a small number of MAGs (8 MAGs) that possessed lower relative abundance, whereas the lower-cost AAs could be synthesized by the majority of members with higher abundance. This result indicates that the strategy of AA synthesis within the group of SAOBs was also affected by the biosynthetic cost.

In addition, relative abundance of MAGs that possessed genes to synthesize a larger set of AAs was lower than that of MAGs that possess genes to synthesize only few AAs (**Figure 6**; left). Pearson correlation test further suggested that from a genomic perspective, the number of AAs that an MAG can synthesize was negatively correlated with its relative abundance in all the four AMCs (**Figure 7**; for ATL, $p = 0.57$, $r = -0.2$; for PTL, $p = 0.01$, $r = -0.76$; for BTL, $p = 0.0001$, $r = -0.96$; for VTL, $p = 0.14$,

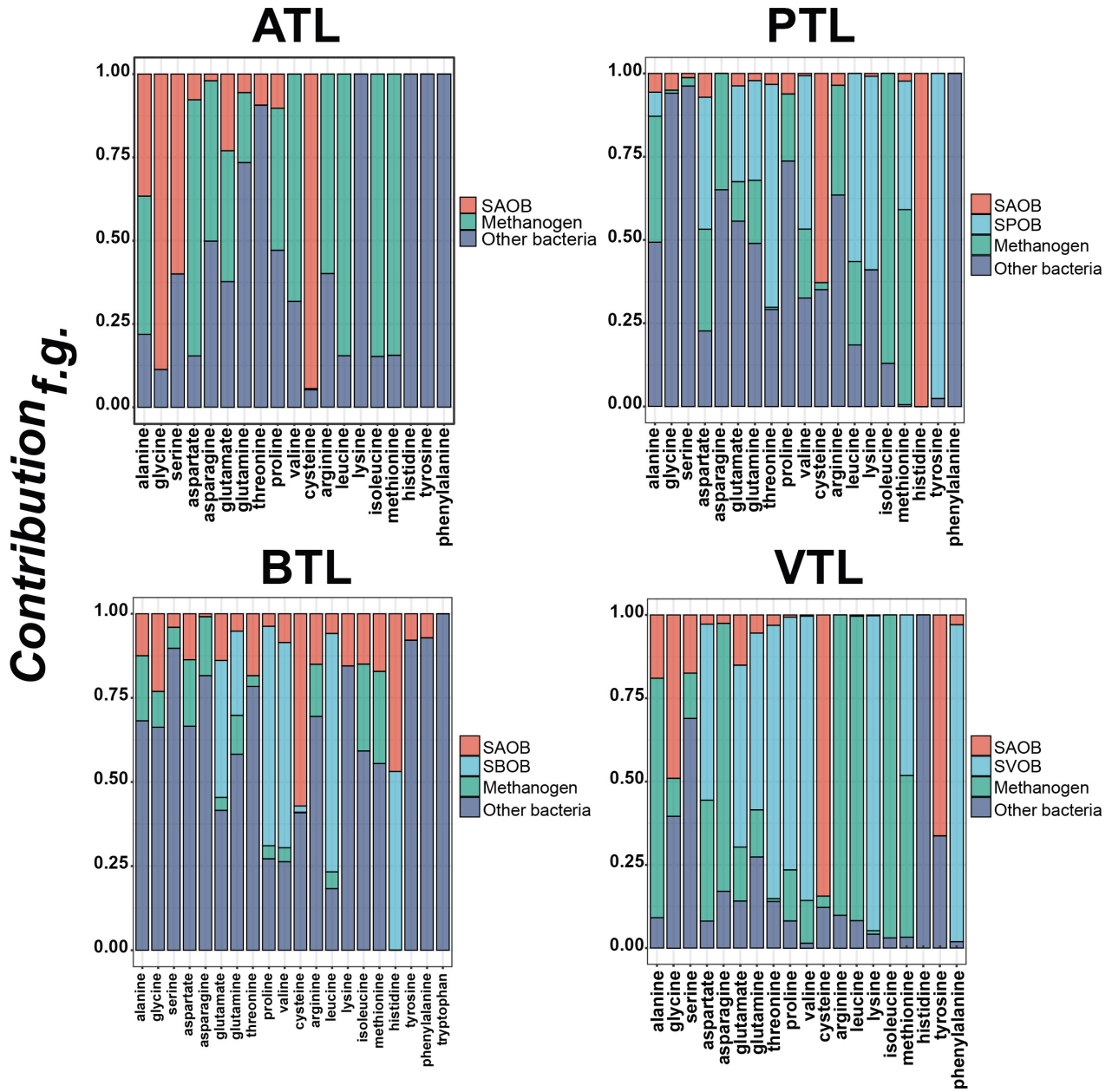


FIGURE 5 | The contributions of different functional groups to amino acid synthesis in each chemostat. The ATL, PTL, BTL, and VTL represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively. We quantified the relative contribution of AA synthesis of each functional group by defining $Contribution_{f.g.}$. Details on the definition are described in "Materials and Methods" section. More details about the contributions of different functional groups to the synthesis of each AA are shown in **Supplementary Figure 7**.

$r = -0.47$). Similarly, we also found this negative correlation at the transcriptomic level (**Supplementary Figure 8**). As we mentioned previously, acetate was supplied as the substrate in ATL, and thus, SAOBs in ATL possess higher energy availability. Higher energy availability causes SAOBs with different AA synthesis ability hold similar fitness, which might result in similar relative abundance of SAOBs in ATL (**Figure 7A**). Similarly, SAOBs in VTL possessed higher energy availability than those in PTL and BTL, which might cause the observed negative correlation being weaker in VTL than that in PTL and BTL (**Figures 7B–D**). Taken together, energy (acetate) availability determines whether an organism with a specific synthesis strategy (synthesize more AAs) is favored.

Amino Acid Synthesis in Methanogens

We then turned to investigate the factors that affect the strategy of AA synthesis in methanogens. As shown in **Figure 8**, all methanogens in our four AMCs did not contain complete

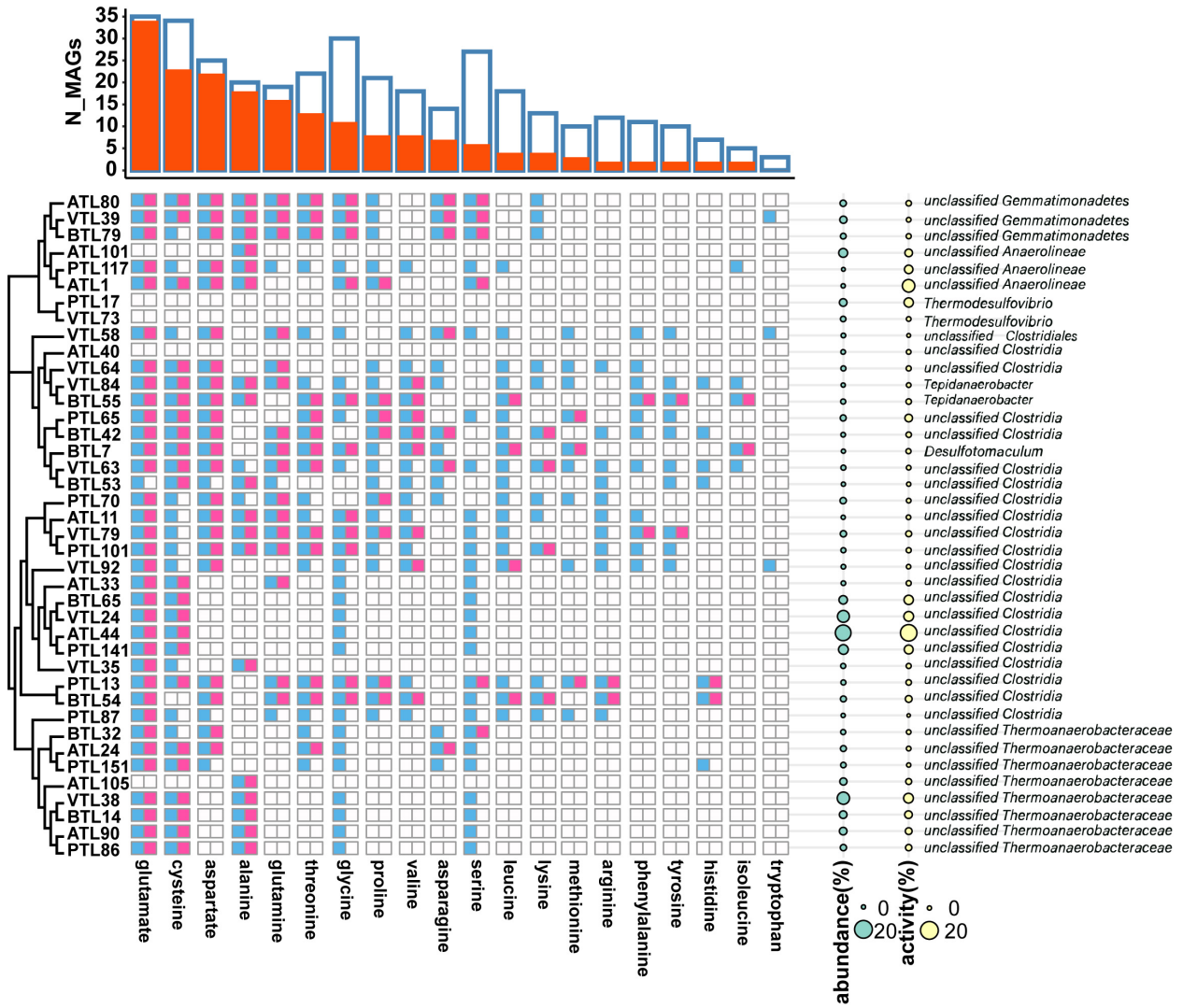


FIGURE 6 | The strategy of amino acid synthesis of syntrophic acetate oxidating bacteria (SAOBs) in different chemostats. The blue squares represent the presence of AA synthesis genes in the corresponding MAGs, whereas the red squares represent the activity of AA synthesis at transcriptomic level. Phylogenetic relationships among these MAGs are shown on the left side. On the right side, the cyan disks represent the abundance of the MAGs, whereas the yellow disks represent their transcriptional activities. The overall landscape of AA synthesis in SAOB MAGs is summarized on the top of the figure. Here, blue frames represent the number of MAGs containing the synthetic genes of the corresponding AA, and the red column represents the number of MAGs that actively transcribed the corresponding AA.

gene set encoding the enzyme of synthesizing those costly AAs (i.e., phenylalanine, tyrosine, histidine, and tryptophan), which indicate that the biosynthetic cost of AAs also had significant impacts on the AA synthesis strategies of methanogens.

Interestingly, MAGs belonging to *Methanosarcina* could synthesize more kinds of AAs and showed higher activity at transcriptomic level than other methanogens (**Figures 8B,C**). We previously found that, in our AMCs, whereas other methanogen taxa, such as *Methanoculleus* and *Methanothermobacter*, produced methane only by hydrogenotrophic pathway, *Methanosarcina* produced methane through both hydrogenotrophic and acetotrophic pathways (Zeng et al., 2021; **Figure 1**). In a previous study, we found that *Methanosarcina* mostly obtain energy through the acetic acid pathway (Zeng et al., 2021). As acetate was more available than hydrogen for the methanogens (because acetate was supplied as substrate or as a main intermediate), *Methanosarcina* could gain more energy from its methanogenic metabolism, possessing higher energy availability. Moreover, MAGs belonging to *Methanoculleus* in our thermophilic AMCs exhibited different AA synthesis strategies with those in mesophilic AMCs reported previously (Embree et al., 2015). While *Methanoculleus* in those AMCs could synthesize six AAs, including glutamate, glutamine, asparagine, aspartate, glycine, and serine, *Methanoculleus* in our AMCs could only synthesize alanine (**Figure 8A**). Through calculation of Gibbs free energy, we found that less energy could be generated in hydrogenotrophic pathway in thermophilic condition than that in mesophilic condition

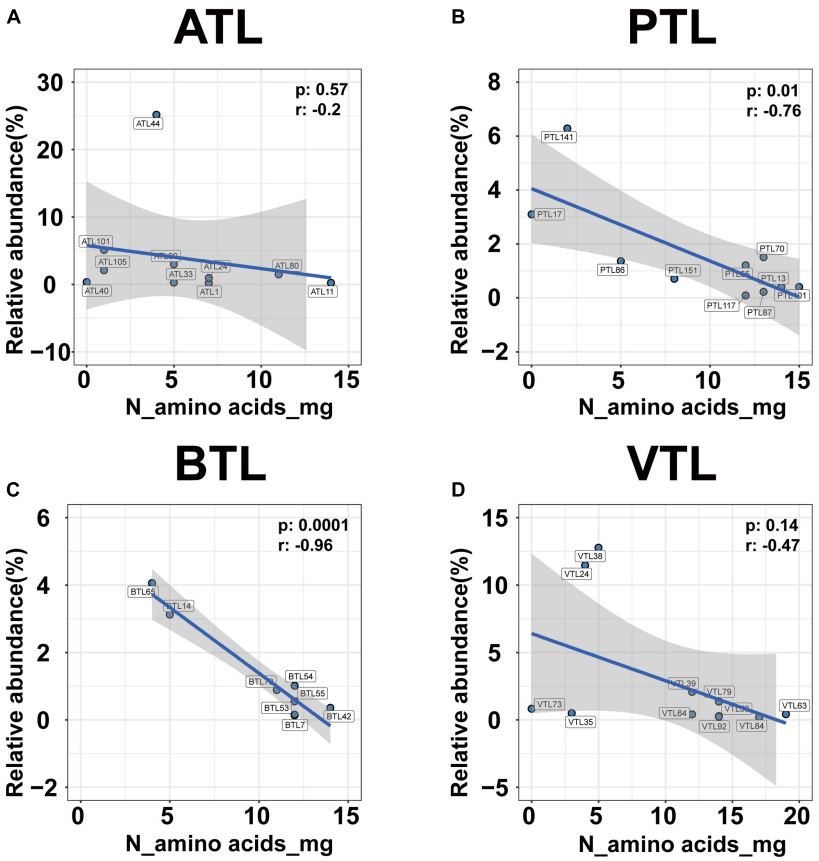


FIGURE 7 | The correlation between the relative abundance and the ability of amino acid synthesis of the MAGs belonging to SAOB at genomic level in ATL **(A)**, PTL **(B)**, BTL **(C)**, and VTL **(D)**. The ATL, PTL, BTL, and VTL represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively. The "N_amino acids_mg" represents the number of amino acids that can be synthesized by each MAG at genomic level. The shaded area represents the 95% confidence region.

(**Supplementary Table 7**). Therefore, *Methanoculleus* possessed lower energy availability in thermophilic condition, which might result in less capability of AA synthesis. In summary, these results suggest that strategies of the AA synthesis of methanogens are also largely affected by energy availability.

Furthermore, we observed a strong AA synthesis complementarity between MAGs belonging to *Methanosarcina* and those belonging to *Methanomassiliicoccus* (**Figure 8A**). *Methanosarcina* and *Methanomassiliicoccus* are frequently observed to co-occur in anaerobic methanogenesis systems (Wang et al., 2019; Wang F. et al., 2020), including our AMCs (Zheng et al., 2019; Chen et al., 2020a,b). Previous studies suggest that this co-occurrence is possibly due to the potential methyl supply of *Methanosarcina* to *Methanomassiliicoccus* (Dong et al., 2017). Here, in our studies, we proposed that the complementarity in PG synthesis and the associated metabolic interdependency might be another reason for their co-occurrence.

Amino Acid Synthesis in Noncore Functional Bacteria

As the noncore functional bacteria held a considerably high abundance and made essential contributions to the AA production in our AMCs (**Figure 5** and **Supplementary Figure 9**), we analyzed the factors that affect the strategy of AA synthesis in these MAGs. At phylum level, these MAGs were mainly classified into Proteobacteria, Planctomycetes, Chloroflexi, Bacteroidetes, Firmicutes, and Actinobacteria. The AA synthesis ability of MAGs within this group was largely varied. The MAGs with relatively higher AA synthesis ability were mainly from Firmicutes, Proteobacteria, and Chloroflexi (**Supplementary Figure 9**). We found that 20 of these MAGs contained genes encoding acetyl-coenzyme A synthetase (Acs), acetate kinase (Ack), and phosphate acyltransferase (Pta), and their encoding genes were also highly transcribed (**Supplementary Figure 9** and **Supplementary Table 8**). These enzymes are reported to catalyze the conversion of acetate to acetyl-CoA. Higher expression of these genes suggested that these MAGs could directly obtain energy from the metabolism of acetyl-CoA, resulting in higher energy availability of them than other noncore functional bacterial MAGs. Correlation analysis further showed that the AA synthesis ability of the "noncore" MAGs is significantly positively correlated with their metabolic activities of converting acetate to acetyl-CoA (**Supplementary Table 9**). We thus speculated that the strategies of AA synthesis

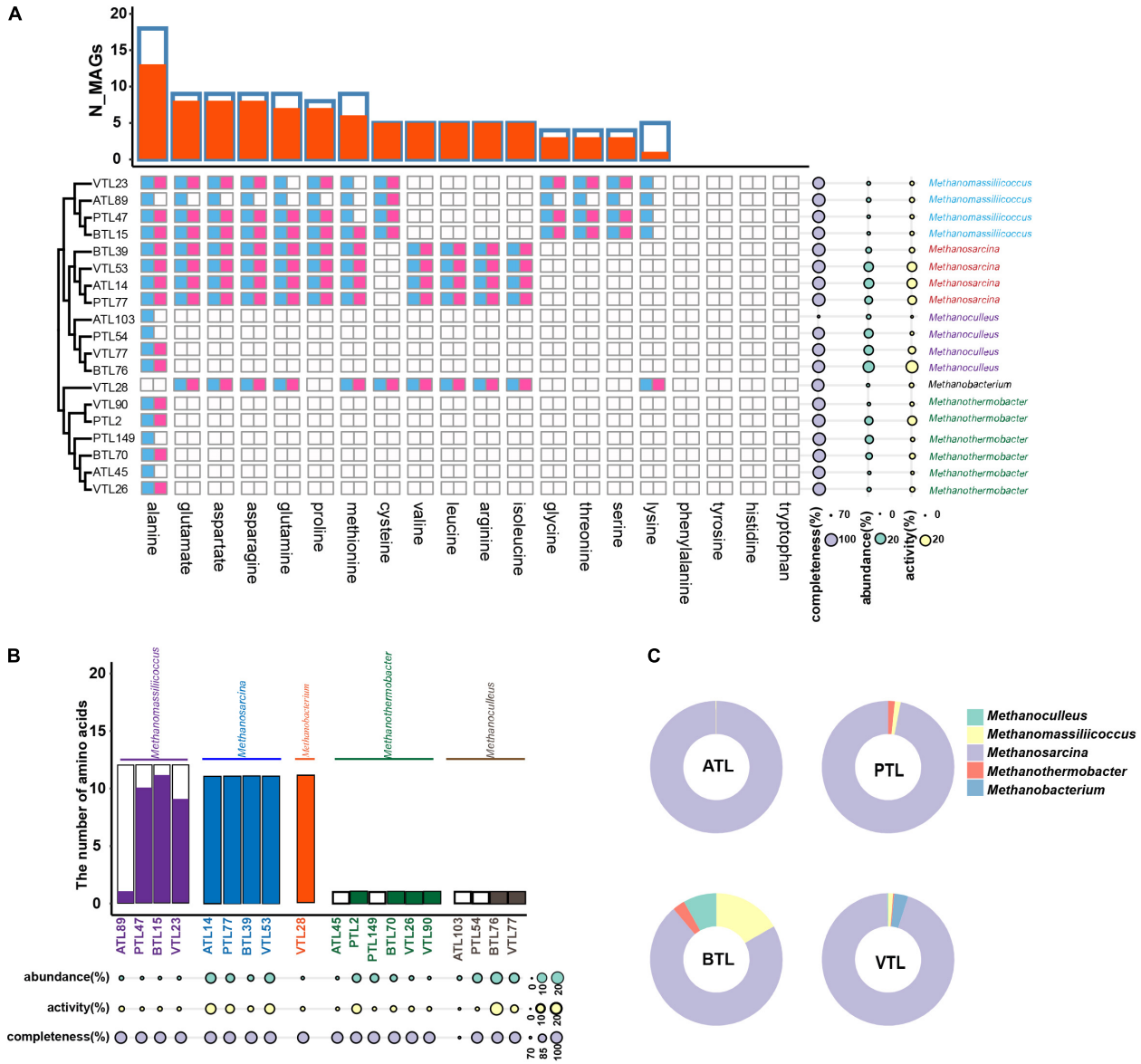


FIGURE 8 | The amino acid synthesis of methanogens in different reactors. **(A)** The blue squares represent the presence of AA synthesis genes in the corresponding MAGs, whereas the red squares represent the activity of AA synthesis at transcriptomic level. Phylogenetic relationships among these MAGs are shown on the left side. On the right side, the cyan disks represent the abundance of the MAGs, whereas the yellow disks represent their transcriptional activities. The overall landscape of AA synthesis in methanogen MAGs are summarized on the top of the figure. Here, blue frames represent the number of MAGs containing the synthetic genes of the corresponding AA, and the red column represents the number of MAGs that actively transcribed corresponding AA. **(B)** Summary of the AA synthesis ability of different methanogen taxa. The black frame represents the number of amino acids that the corresponding MAG can synthesize at genomic level, whereas the color-filled column represents the number of amino acids, of which the synthetic genes were actively transcribed in the corresponding MAG. **(C)** Comparison of relative AA synthesis activity ($RA_{a.a}$) of methanogens from different genera in each chemostat. The ATL, PTL, BTL, and VTL represent the thermophilic methanogenic chemostats supplemented with acetate, propionate, butyrate, and isovalerate as sole carbon source, respectively.

in these noncore functional bacteria were also affected by energy availability.

DISCUSSION

Here, we systematically probed the AA synthesis strategies of community members in four AMCs. We found that energy availability and energy cost of AA synthesis are key drivers to shape the abilities of AA synthesis in AMCs.

Extending the Framework of the Black Queen Hypothesis

The BQH provides an explanation for evolution of metabolic interdependency at community level. The basic framework of

BQH states that the benefit of not carrying a public function is determined by the saving energy that results from not performing the function. In the recent models and experiments, this benefit was formulated by roughly assessing the energy cost of the public function. For example, the cost of synthesizing an AA was assessed by the ATPs it would consume in the enzymatic reactions (Akashi and Gojobori, 2002). However, our results suggest that more factors should be included to evaluate this benefit in complex communities.

First, estimating cost by the amount of ATP consumption ignored many aspects involved in AA synthesis. For example, synthesis of glutamate consumes 15.3 ATPs, which is much higher than those of serine and glycine. According to the theory developed in previous studies (Estrela et al., 2016; Mas et al., 2016; Pacheco et al., 2019; Wang et al., 2021), the community members that can synthesize glutamate should be less than the ones that can synthesize serine and glycine. However, in our study, more MAGs synthesized glutamate rather than serine and glycine, which is also the case in other similar investigations (Mee et al., 2014). This phenomenon suggests that the cost of glutamate synthesis might be overestimated. In the estimation by Akashi et al., the ATP consumption of the synthesis of AA precursor (generated from the central carbon metabolism, such as Embden–Meyerhof–Parnas pathway, tricarboxylic acid, and the pentose phosphate cycles) was also taken into account. However, at community level, microorganisms could obtain precursor metabolites directly secreted by other community members (Zengler and Zaramela, 2018). Therefore, the energy cost in precursor metabolites should be prudently considered when evaluating the function cost. However, the accurate and quantitative assessment of the AA synthesis cost still faces obstacles. For example, it is difficult to quantify the exchange of precursor metabolites for AA synthesis among community members and identify all the AA biosynthesis pathways in a given member without comprehensive physiological investigation of it. More efforts should be made to overcome these obstacles in further studies.

Second, our results suggest that community members could also regulate the energy cost of a public function at transcriptomic level. This regulation occurred in the synthesis of two aromatic AAs, tyrosine and phenylalanine, whose encoding genes were considerably highly abundant but lowly transcribed in our systems (**Supplementary Figure 6**). Previous studies have revealed that the transcription of the related genes is regulated by a repressor, TrpR (Otwinowski et al., 1988). We searched our meta-omics data and found the identified trpR genes were highly active in the four MAGs (ATL5, ATL46, BTL5, and PTL65) that possessed genes to synthesize both tyrosine and phenylalanine (**Supplementary Table 10**). This result suggests that these MAGs could repress the transcription of these genes to decrease its energy cost. We proposed that this strategy may be, to some extent, better than directly losing the functional genes, by which microorganisms could actively change their metabolic states of PG production when facing fluctuated environmental conditions. Therefore, it is important to consider transcriptional physiology of community members when evaluating the benefit of not carrying a public function in complex communities.

Finally, our results indicate that if a community member could gain more energy in the habitat or niche it occupied, it possessed higher ability to produce more costly PGs. This phenomenon was also observed in other systems. For instance, in oil reservoir, aerobic microorganisms that possessed higher energy availability retained higher ability for synthesizing AAs and vitamins than anaerobic microorganisms (Liu et al., 2018). In the methanogenic system enriched by Zhu et al. (2020), *Coprothermobacter proteolyticus* DTU 632 lacked efficient pathway for electron disposal and the pathway of energy metabolism (acetate catabolism), so relied on other community members for those expensive AAs to reduce the energy burden. Therefore, estimating the benefit of function deficiency in a complex community should not only consider how much energy (ATP) is required for a member to perform a public function, but should also assess how much of the total energy this member can obtain from its overall metabolism. We thus propose that the ratio of essential function cost (*c*) to the overall energy availability of a member (*ea*) is a better indicator to quantify the benefit of function deficiency.

Implications From Phylogenetic Conservatism of Amino Acid Synthesis Strategy

A central goal in microbial ecology is to reveal the relationship between community composition and the community functioning (Zhou, 2009; Krause et al., 2014; Martiny et al., 2015). The key to achieving this goal is to understand whether the function traits of a microorganism are correlated with its phylogeny (Philippot et al., 2010; Martiny et al., 2013; Louca et al., 2018). Our results indicate that the strategies of AA synthesis of MAGs of SAOB and methanogens in our AMCs were clustered according to their phylogenetic relatedness (**Figures 6**, **8**), despite that these MAGs came from different chemostats initialized with different seedings (Chen et al., 2020a,b). For example, ATL90, BTL14, PTL86, and VTL38 (bacterial MAG) in ATL, BTL, PTL, and VTL adapted totally the same strategy of AA synthesis; that is, they only autonomously synthesized glutamate, cysteine, and alanine. Similarly, in all the four chemostats, *Methanothermobacter* and *Methanosarcina* MAGs (archaeal MAG) could only synthesize alanine, whereas all *Methanosarcina* could synthesize a variety of AAs, including alanine, glutamate, aspartate, glutamine, proline, methionine, valine, leucine, arginine, and isoleucine. This result suggests that the function of AA synthesis in microorganisms is highly phylogenetically conserved and may be mostly shaped by vertical inheritance driven by long-term nature selection, but not by rapid convergent evolution through random gene loss or horizontal gene transfer (Martiny et al., 2013). On the one hand, this result suggests that the evolution of PG (AA) exchange in AMCs is mostly driven by nature selection, such as the adaptive gene loss characterized by BQH (Morris et al., 2012). This observation is different from the cases in several other environments, for example, in the host-associated environments, which is driven by random genetic drift (Nilsson et al., 2005; Kuo et al., 2009). On the other hand, as previous studies

also suggest that the traits of energy metabolism (anaerobic methanogenesis) in AMCs are also phylogenetically conserved (Martiny et al., 2013; Evans et al., 2019), we hypothesized that AA synthesis ability of microorganisms in AMCs is coevolved with their core energy metabolism to achieve optimal survival strategy. For microorganisms that can only obtain limiting energy, they save energy from AA synthesis for their better growth/survival. For microorganisms that can obtain more energy from its core metabolism, they may retain more AA synthesis genes to increase its own autonomy, reducing its risk facing against environmental fluctuations. Nevertheless, this hypothesis requires further examination.

Metabolic Interdependency Affects the Performance of the Anaerobic Methanogenic Communities

Anaerobic methanogenic communities play important roles in treating our waste and driving global carbon cycles (Zhang et al., 2021). In AMCs, syntrophic fatty acid–oxidizing bacteria and methanogens cooperate to perform the conversion of fatty acids to methane, acting as the core functional groups (Nobu et al., 2015). However, it was confusing why the community members that are not involved in the core metabolism were present in the AMCs and how they contribute to the development and function of the AMCs (Ruiz-Sánchez et al., 2019; Wang P. et al., 2020). Our results suggest that those noncore functional bacteria are crucial for producing the costly AAs (tryptophan, tyrosine, histidine, and phenylalanine) that support the survival of the core functional groups. In addition to the AA exchange, several studies also indicate that methanogens also rely on those noncore members for vitamins (Hubalek et al., 2017), another type of PG widespread in natural microbial communities (Rodríguez-Fernández et al., 2018). Therefore, by contributing PGs, those previously thought "noncore" members actually plays a "core" role on the development of the AMCs, which further influences the metabolic efficiency of the AMCs. This finding suggests a novel insight on managing these AMCs for better output: manipulate those "noncore" members to contribute more public secretions; the methanogenic groups may perform better to treat our waste.

In addition, the mode of metabolic interdependency might be one of the causes for the vulnerability of AMCs (De Vrieze et al., 2012; Zhu et al., 2020). Our results suggest that those essential and costly AAs (tyrosine, histidine, tryptophan, and phenylalanine) were produced by only a few community members with lower abundance and activity. These low abundant members may be easily lost when the communities undergo environmental fluctuations. Without supplying the essential AAs, the fitness of the functional groups, especially the methanogens that possess low ability of AA synthesis, would rapidly decrease, resulting in the collapse of the community (Chen et al., 2008). In other words, the production of those costly AAs is highly limited in AMCs. Therefore, artificially feeding these AAs to the system is a potential strategy to increase the stability and robustness of AMCs.

Nevertheless, several studies indicate that community with such mode of metabolic interdependency can also benefit community stability. For example, computational simulation in one recent study found that the community with interdependent pattern has better resistance to nutrient disturbances than the community composed of only the autonomous population that can produce all PGs (Wang et al., 2021). Because the resources that were originally wastefully allocated to produce redundant PGs were saved to fight against the harsh environmental change, energy supply is even more limited in the AMCs, and thus, the energy saving by function loss may be selectively favored at the community level in anaerobic environment during long-term evolution.

Limitations of This Study

Despite these encouraging findings, we acknowledge four limitations of our study. First, we reconstructed the AA synthesis pathway of MAGs based on the KEGG database and determined the AA synthesis capacity of MAGs based on the expression of the pathway in MAGs. This approach is widely used in recent studies (Embree et al., 2015; Liu et al., 2018; Chen et al., 2020a; Zhu et al., 2020). However, it can be misleading due to the incomplete collection of metabolic pathways in the public databases and the limitations of our current understanding on biological metabolic pathways. Moreover, following classical knowledge, we assumed that the 20 AAs are essential to the growth of all the microorganisms. However, recent studies suggest that several bacteria and archaea grow without providing some of these "essential" AAs. For example, (Price et al., 2018) found that heterotrophic bacteria from 10 different genera annotated as AA auxotroph in IMG[1] could survive independently without extra AA supplement. Therefore, more culture-dependent experiments are required to advance our understanding on microbial metabolism of AAs and then guide our further analysis of AA exchange in complex microbial communities.

Second, we estimated the potential metabolic exchange of AAs between every two MAGs based on whether these MAGs carried out and expressed the related genes. However, this analysis is not enough to thoroughly understand their interactions. For example, the efficiency of PG exchange is largely affected by the transport of the PG across the cells (Kanzaki and Anraku, 1971), the diffusion rate of the PG (Julou et al., 2013), and the spatial positioning of different members (Allen et al., 2013). While the effects of these factors can hardly be achieved by bioinformatics investigations, culture-dependent studies could be adapted to quantify these effects. For example, combination of FISH and NanoSIMS technologies could be used to visualize the relative positioning of different community members in the AMCs, as well as measuring the distribution of PG *in situ*. We expect that these studies could further advance our understanding on PG exchange in AMCs.

Third, based on previous studies, we proposed a formula $\left(p \propto \frac{ea}{c}\right)$ to depict how the abilities of AA synthesis of a microorganism were determined. Unfortunately, because of the rapidity and complexity of the biological metabolic process,

[1] https://img.jgi.doe.gov/

we did not directly quantify energy availability of different microorganisms in our system and also lacked a precise way to measure function costs of the AA synthesis as described previously. Therefore, we did not test our hypothesis from a quantitative perspective. We expect further studies being set out to build the quantitative framework based on our proposed formula.

Finally, although the noncore functional bacteria play important roles in producing several costly AAs, the energy sources of these bacteria are still unclear. Our results suggest that some of these bacteria may acquire energy from metabolizing acetyl-CoA, but the conversion of acetate to acetyl-CoA consumes ATPs. The energy balance of the metabolism still required further investigation. Moreover, many noncore bacteria that did not contain these acetate acetylation genes were also considerably abundant in the AMCs. We hypothesized that these noncore bacteria might also scavenge by-products or the dead biomass of the core functional taxa. However, further experiments are still required to provide direct evidence for these hypotheses.

CONCLUSION

In summary, we revealed that metabolic interdependency based on AA exchange is prevalent in AMCs and is fundamental to connect different functional taxa through complex interaction webs. For the first time, we proposed that the strategy of public functions in a member residing in a community is largely influenced by how much energy it could acquire, which is determined by the niche it occupied and its metabolic strategy in specific environmental conditions. This finding shed light on how energy availability acts as a driving force of microbial evolution in complex microbial communities and also provides novel insights on how to manage AMCs and address grand challenges facing against environmental pollution.

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 below: http://bigd.big.ac.cn/gsa, CRA004311.

AUTHOR CONTRIBUTIONS

JY: formal analysis, visualization, and writing–reviewing and editing. YZ: resources, software, and methodology. MW: conceptualization and writing–review and editing. Y-QT: supervision, project administration, funding acquisition, and writing–reviewing and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by the Ministry of Science and Technology of China (Nos. 2018YFA0902102 and 2018YFA0902100) and National Natural Science Foundation of China (No. 51678378).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.744834/full#supplementary-material

REFERENCES

Akashi, H., and Gojobori, T. (2002). Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* 99, 3695–3700. doi: 10.1073/pnas.062526999

Allen, B., Gore, J., and Nowak, M. A. (2013). Spatial dilemmas of diffusible public goods. *Elife* 2:e01169. doi: 10.7554/eLife.01169

Barbara, G. M., and Mitchell, J. G. (2003). Marine bacterial organisation around point-like sources of amino acids. *FEMS Microbiol. Ecol.* 43, 99–109. doi: 10.1111/j.1574-6941.2003.tb01049.x

Blasche, S., Kim, Y., Mars, R. A. T., Machado, D., Mannsson, M., Kafkia, E., et al. (2021). Metabolic cooperation and spatiotemporal niche partitioning in a kefir microbial community. *Nat. Microbiol.* 2, 196–208. doi: 10.1038/s41564-020-00816-5

Chen, Y., Cheng, J. J., and Creamer, K. S. (2008). Inhibition of anaerobic digestion process: a review. *Bioresour. Technol.* 99, 4044–4064. doi: 10.1016/j.biortech.2007.01.057

Chen, Y. T., Zeng, Y., Li, J., Zhao, X. Y., Yi, Y., Gou, M., et al. (2020a). Novel syntrophic isovalerate-degrading bacteria and their energetic cooperation with methanogens in methanogenic chemostats. *Environ. Sci. Technol.* 54, 9618–9628. doi: 10.1021/acs.est.0c01840

Chen, Y. T., Zeng, Y., Wang, H. Z., Zheng, D., Kamagata, Y., Narihiro, T., et al. (2020b). Different interspecies electron transfer patterns during mesophilic and thermophilic syntrophic propionate degradation in chemostats. *Microb. Ecol.* 80, 120–132. doi: 10.1007/s00248-020-01485-x

Cooper, M. B., Kazamia, E., Helliwell, K. E., Kudahl, U. J., Sayer, A., Wheeler, G. L., et al. (2019). Cross-exchange of B-vitamins underpins a mutualistic interaction between *Ostreococcus tauri* and *Dinoroseobacter shibae*. *ISME J.* 13, 334–345. doi: 10.1038/s41396-018-0274-y

Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J., and Smith, A. G. (2005). Algae acquire vitamin B12 through a symbiotic relationship with bacteria. *Nature* 438, 90–93. doi: 10.1038/nature04056

De Vrieze, J., Hennebel, T., Boon, N., and Verstraete, W. (2012). *Methanosarcina*: the rediscovered methanogen for heavy duty biomethanation. *Bioresour. Technol.* 112, 1–9. doi: 10.1016/j.biortech.2012.02.079

Dong, M., Gonzalez, T. D., Klems, M. M., Steinberg, L. M., Chen, W., Papoutsakis, E. T., et al. (2017). In vitro methanol production from methyl coenzyme M using the *Methanosarcina barkeri* MtaABC protein complex. *Biotechnol. Prog.* 33, 1243–1249. doi: 10.1002/btpr.2503

Embree, M., Liu, J. K., Al-Bassam, M. M., and Zengler, K. (2015). Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15450–15455. doi: 10.1073/pnas.1506034112

Estrela, S., Morris, J. J., and Kerr, B. (2016). Private benefits and metabolic conflicts shape the emergence of microbial interdependencies. *Environ. Microbiol.* 18, 1415–1427. doi: 10.1111/1462-2920.13028

Evans, P. N., Boyd, J. A., Leu, A. O., Woodcroft, B. J., Parks, D. H., Hugenholtz, P., et al. (2019). An evolving view of methane metabolism in the Archaea. *Nat. Rev. Microbiol.* 17, 219–232. doi: 10.1038/s41579-018-0136-7

Garcia, S. L., Buck, M., McMahon, K. D., Grossart, H.-P., Eiler, A., and Warnecke, F. (2015). Auxotrophy and intrapopulation complementary in the 'interactome' of a cultivated freshwater model community. *Mol. Ecol.* 24, 4449–4459. doi: 10.1111/mec.13319

Hubalek, V., Buck, M., Tan, B., Foght, J., Wendeberg, A., Berry, D., et al. (2017). Vitamin and amino acid auxotrophy in anaerobic consortia operating under methanogenic conditions. *mSystems* 2, e00038–17. doi: 10.1128/mSystems.00038-17

Jackson, B. E., and McInerney, M. J. (2002). Anaerobic microbial metabolism can proceed close to thermodynamic limits. *Nature* 415, 454–456. doi: 10.1038/415454a

Jones, R. G. (1967). Ubiquinone deficiency in an auxotroph of *Escherichia coli* requiring 4-hydroxybenzoic acid. *Biochem. J.* 103, 714–719. doi: 10.1042/bj1030714

Julou, T., Mora, T., Guillon, L., Croquette, V., Schalk, I. J., Bensimon, D., et al. (2013). Cell–cell contacts confine public goods diffusion inside *Pseudomonas aeruginosa* clonal microcolonies. *Proc. Natl. Acad. Sci. U.S.A.* 110:12577. doi: 10.1073/pnas.1301428110

Kanzaki, S., and Anraku, Y. (1971). Transport of sugars and amino acids in bacteria. IV. Regulation of valine transport activity by valine and cysteine. *J. Biochem.* 70, 215–224. doi: 10.1093/oxfordjournals.jbchem.a129633

Krakat, N., Schmidt, S., and Scherer, P. (2011). Potential impact of process parameters upon the bacterial diversity in the mesophilic anaerobic digestion of beet silage. *Bioresour. Technol.* 102, 5692–5701. doi: 10.1016/j.biortech.2011.02.108

Kramer, J., Özkaya, Ö, and Kümmerli, R. (2020). Bacterial siderophores in community and host interactions. *Nat. Rev. Microbiol.* 18, 152–163. doi: 10.1038/s41579-019-0284-4

Krause, S., Le Roux, X., Niklaus, P. A., Van Bodegom, P. M., Lennon, J. T., Bertilsson, S., et al. (2014). Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front. Microbiol.* 5:251. doi: 10.3389/fmicb.2014.00251

Kuo, C.-H., Moran, N. A., and Ochman, H. (2009). The consequences of genetic drift for bacterial genome complexity. *Genome Res.* 19, 1450–1454. doi: 10.1101/gr.091785.109

Liu, Y. F., Galzerani, D. D., Mbadinga, S. M., Zaramela, L. S., Gu, J. D., Mu, B. Z., et al. (2018). Metabolic capability and in situ activity of microorganisms in an oil reservoir. *Microbiome* 6:5. doi: 10.1186/s40168-017-0392-1

Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O'Connor, M. I., et al. (2018). Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.* 2, 936–943. doi: 10.1038/s41559-018-0519-1

Luo, H. W., Zhang, H., Suzuki, T., Hattori, S., and Kamagata, Y. (2002). Differential expression of methanogenesis genes of *Methanothermobacter thermoautotrophicus* (formerly *Methanobacterium thermoautotrophicum*) in pure culture and in cocultures with fatty acid-oxidizing syntrophs. *Appl. Environ. Microbiol.* 68, 1173–1179. doi: 10.1128/AEM.68.3.1173-1179.2002

Lyu, Z., Shao, N., Akinyemi, T., and Whitman, W. B. (2018). Methanogenesis. *Curr. Biol.* 28, R727–R732. doi: 10.1016/j.cub.2018.05.021

Martiny, A. C., Treseder, K., and Pusch, G. (2013). Phylogenetic conservatism of functional traits in microorganisms. *ISME J.* 7, 830–838. doi: 10.1038/ismej.2012.160

Martiny, J. B. H., Jones, S. E., Lennon, J. T., and Martiny, A. C. (2015). Microbiomes in light of traits: a phylogenetic perspective. *Science* 350:aac9323. doi: 10.1126/science.aac9323

Mas, A., Jamshidi, S., Lagadeuc, Y., Eveillard, D., and Vandenkoornhuyse, P. (2016). Beyond the black queen hypothesis. *ISME J.* 10, 2085–2091. doi: 10.1038/ismej.2016.22

McInerney, M. J., Bryant, M. P., Hespell, R. B., and Costerton, J. W. (1981). *Syntrophomonas wolfei* gen. nov. sp. nov., an anaerobic, syntrophic, fatty acid-oxidizing bacterium. *Appl. Environ. Microbiol.* 41, 1029–1039. doi: 10.1128/aem.41.4.1029-1039.1981

Mee, M. T., Collins, J. J., Church, G. M., and Wang, H. H. (2014). Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2149–E2156. doi: 10.1073/pnas.1405641111

Meijer, J., van Dijk, B., and Hogeweg, P. (2020). Contingent evolution of alternative metabolic network topologies determines whether cross-feeding evolves. *Commun. Biol.* 3:401. doi: 10.1038/s42003-020-1107-x

Merchant, S. S., and Helmann, J. D. (2012). Elemental economy: microbial strategies for optimizing growth in the face of nutrient limitation. *Adv. Microb. Physiol.* 60, 91–210. doi: 10.1016/B978-0-12-398264-3.00002-4

Mitri, S., and Foster, K. R. (2013). The genotypic view of social interactions in microbial communities. *Annu. Rev. Genet.* 47, 247–273. doi: 10.1146/annurev-genet-111212-133307

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35, W182–W185. doi: 10.1093/nar

Morris, J. J., Johnson, Z. I., Szul, M. J., Keller, M., and Zinser, E. R. (2011). Dependence of the *Cyanobacterium Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the ocean's surface. *PLoS One* 6:e16805. doi: 10.1371/journal.pone.0016805

Morris, J. J., Lenski, R. E., and Zinser, E. R. (2012). The black queen hypothesis: evolution of dependencies through adaptive gene loss. *mBio* 3, e00036–12. doi: 10.1128/mBio.00036-12

Neis, E. P. J. G., Dejong, C. H. C., and Rensen, S. S. (2015). The role of microbial amino acid metabolism in host metabolism. *Nutrients* 7, 2930–2946. doi: 10.3390/nu7042930

Nilsson, A. I., Koskiniemi, S., Eriksson, S., Kugelberg, E., Hinton, J. C. D., and Andersson, D. I. (2005). Bacterial genome size reduction by experimental evolution. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12112–12116. doi: 10.1073/pnas.0503654102

Nobu, M. K., Narihiro, T., Mei, R., Kamagata, Y., Lee, P. K. H., Lee, P.-H., et al. (2020). Catabolism and interactions of uncultured organisms shaped by eco-thermodynamics in methanogenic bioprocesses. *Microbiome* 8:111. doi: 10.1186/s40168-020-00885-y

Nobu, M. K., Narihiro, T., Rinke, C., Kamagata, Y., Tringe, S. G., Woyke, T., et al. (2015). Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. *ISME J.* 9, 1710–1722. doi: 10.1038/ismej.2014.256

Oliveira, N. M., Niehus, R., and Foster, K. R. (2014). Evolutionary limits to cooperation in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17941–17946. doi: 10.1073/pnas.1412673111

Otwinowski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., et al. (1988). Crystal structure of trp represser/operator complex at atomic resolution. *Nature* 335, 321–329. doi: 10.1038/335321a0

Pacheco, A. R., Moel, M., and Segrè, D. (2019). Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat. Commun.* 10:103. doi: 10.1038/s41467-018-07946-9

Pande, S., Merker, H., Bohl, K., Reichelt, M., Schuster, S., de Figueiredo, L. F., et al. (2014). Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. *ISME J.* 8, 953–962. doi: 10.1038/ismej.2013.211

Perna, S., Alalwan, T. A., Alaali, Z., Alnashaba, T., Gasparri, C., Infantino, V., et al. (2019). The role of glutamine in the complex interaction between gut microbiota and health: a narrative review. *Int. J. Mol. Sci.* 20:5232. doi: 10.3390/ijms20205232

Philippot, L., Andersson, S. G. E., Battin, T. J., Prosser, J. I., Schimel, J. P., Whitman, W. B., et al. (2010). The ecological coherence of high bacterial taxonomic ranks. *Nat. Rev. Microbiol.* 8, 523–529. doi: 10.1038/nrmicro2367

Price, M. N., Zane, G. M., Kuehl, J. V., Melnyk, R. A., Wall, J. D., Deutschbauer, A. M., et al. (2018). Filling gaps in bacterial amino acid biosynthesis pathways with high-throughput genetics. *PLoS Genet.* 3:e1007147. doi: 10.1371/journal.pgen.1007147

Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., et al. (2009). Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME J.* 3, 700–714. doi: 10.1038/ismej.2009.2

Rodionova, I. A., Li, X., Plymale, A. E., Motamedchaboki, K., Konopka, A. E., Romine, M. F., et al. (2015). Genomic distribution of B-vitamin auxotrophy and uptake transporters in environmental bacteria from the chloroflexi phylum. *Environ. Microbiol. Rep.* 7, 204–210. doi: 10.1111/1758-2229.12227

Rodríguez-Fernández, D., Torrentó, C., Guivernau, M., Viñas, M., Hunkeler, D., Soler, A., et al. (2018). Vitamin B12 effects on chlorinated methanes-degrading

microcosms: dual isotope and metabolically active microbial populations assessment. *Sci. Total Environ*. 621, 1615–1625. doi: 10.1016/j.scitotenv.2017.10.067

Ruiz-Sánchez, J., Guivernau, M., Fernández, B., Vila, J., Viñas, M., Riau, V., et al. (2019). Functional biodiversity and plasticity of methanogenic biomass from a full-scale mesophilic anaerobic digester treating nitrogen-rich agricultural wastes. *Sci. Total Environ*. 649, 760–769. doi: 10.1016/j.scitotenv.2018.08.165

Soto-Martin, E. C., Warnke, I., Farquharson, F. M., Christodoulou, M., Horgan, G., Derrien, M., et al. (2020). Vitamin biosynthesis by human gut butyrate-producing bacteria and cross-feeding in synthetic microbial communities. *mBio* 11, e00886–20. doi: 10.1128/mBio.00886-20

Tang, Y.-Q., Ji, P., Hayashi, J., Koike, Y., Wu, X.-L., and Kida, K. (2011). Characteristic microbial community of a dry thermophilic methanogenic digester: its long-term stability and change with feeding. *Appl. Microbiol. Biotechnol*. 91, 1447–1461. doi: 10.1007/s00253-011-3479-9

Wang, B., Wang, Y., Cui, X., Zhang, Y., and Yu, Z. (2019). Bioconversion of coal to methane by microbial communities from soil and from an opencast mine in the Xilingol grassland of northeast China. *Biotechnol Biofuels* 12:236. doi: 10.1186/s13068-019-1572-y

Wang, F., Yi, W., Zhang, D., Liu, Y., Shen, X., and Li, Y. (2020). Anaerobic co-digestion of corn stover and wastewater from hydrothermal carbonation. *Bioresour. Technol*. 315:123788. doi: 10.1016/j.biortech.2020.123788

Wang, M., Liu, X., Nie, Y., and Wu, X. L. (2021). Selfishness driving reductive evolution shapes interdependent patterns in spatially structured microbial communities. *ISME J*. 15, 1387–1401. doi: 10.1038/s41396-020-00858-x

Wang, P., Qiao, Z., Li, X., Su, Y., and Xie, B. (2020). Functional characteristic of microbial communities in large-scale biotreatment systems of food waste. *Sci. Total Environ*. 746:141086. doi: 10.1016/j.scitotenv.2020.141086

Zeng, Y., Zheng, D., Gou, M., Xia, Z. Y., Chen, Y. T., Nobu, M. K., et al. (2021). Chasing the metabolism of novel syntrophic acetate-oxidizing bacteria in thermophilic methanogenic chemostats. *bioRxiv* [Preprint]. Available online at: https://connect.biorxiv.org/qr/2021.07.06.451242 (accessed July 6, 2021).

Zengler, K., and Zaramela, L. S. (2018). The social network of microorganisms–how auxotrophies shape complex communities. *Nat. Rev. Microbiol*. 16, 383–390. doi: 10.1038/s41579-018-0004-5

Zhang, M., Lu, G., Li, Z., Xu, F., Yang, N., Sun, X., et al. (2021). Effects of antimony on anaerobic methane oxidization and microbial community in an antimony-contaminated paddy soil: a microcosm study. *Sci. Total Environ*. 784:147239. doi: 10.1016/j.scitotenv.2021.147239

Zheng, D., Wang, H.-Z., Gou, M., Nobu, M. K., Narihiro, T., Hu, B., et al. (2019). Identification of novel potential acetate-oxidizing bacteria in thermophilic methanogenic chemostats by DNA stable isotope probing. *Appl. Microbiol. Biotechnol*. 103, 8631–8645. doi: 10.1007/s00253-019-10078-9

Zhou, J. (2009). Predictive microbial ecology. *Microb. Biotechnol*. 2, 154–156. doi: 10.1111/j.1751-7915.2009.00090_21.x

Zhu, X., Campanaro, S., Treu, L., Seshadri, R., Ivanova, N., Kougias, P. G., et al. (2020). Metabolic dependencies govern microbial syntrophies during methanogenesis in an anaerobic digestion ecosystem. *Microbiome* 8:22. doi: 10.1186/s40168-019-0780-9

Zomorrodi, A. R., and Segrè, D. (2017). Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nat. Commun*. 8:1563. doi: 10.1038/s41467-017-01407-5

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 October 2021
doi: 10.3389/fmicb.2021.752750

Quantifying the Importance of Abiotic and Biotic Factors Governing the Succession of Gut Microbiota Over Shrimp Ontogeny

*Wenqian Zhang[1,2], Zidong Zhu[3], Jiong Chen[1,2], Qiongfen Qiu[2] and Jinbo Xiong[1,2]**

[1] State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-Products, Ningbo University, Ningbo, China, [2] School of Marine Sciences, Ningbo University, Ningbo, China, [3] School of Biochemical Engineering, Jingzhou Institute of Technology, Jingzhou, China

OPEN ACCESS

Edited by:
Jianjun Wang, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences (CAS), China

Reviewed by:
Yuyi Yang, Wuhan Botanical Garden, Chinese Academy of Sciences (CAS), China
Adrian Ochoa-Leyva, National Autonomous University of Mexico, Mexico
Shan Gong Wu, Institute of Hydrobiology, Chinese Academy of Sciences (CAS), China

***Correspondence:**
Jinbo Xiong
xiongjinbo@nbu.edu.cn

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *03 August 2021*
Accepted: *31 August 2021*
Published: *08 October 2021*

Citation:
Zhang W, Zhu Z, Chen J, Qiu Q and Xiong J (2021) Quantifying the Importance of Abiotic and Biotic Factors Governing the Succession of Gut Microbiota Over Shrimp Ontogeny. Front. Microbiol. 12:752750. doi: 10.3389/fmicb.2021.752750

Intensive studies have evaluated abiotic factors in shaping host gut microbiota. In contrast, little is known on how and to what extent abiotic (geochemical variables) and biotic (i.e., surrounding microbes, younger shrimp, and age) factors assemble the gut microbiota over shrimp ontogeny. Considering the functional importance of gut microbiota in improving host fitness, this knowledge is fundamental to sustain a desirable gut microbiota for a healthy aquaculture. Here, we characterized the successional rules of both the shrimp gut and rearing water bacterial communities over the entire shrimp farming. Both the gut and rearing water bacterial communities exhibited the time decay of similarity relationship, with significantly lower temporal turnover rate for the gut microbiota, which were primarily governed by shrimp age (days postlarval inoculation) and water pH. Gut commensals were primary sourced (averaged 60.3%) from their younger host, rather than surrounding bacterioplankton (19.1%). A structural equation model revealed that water salinity, pH, total phosphorus, and dissolve oxygen directly governed bacterioplankton communities but not for the gut microbiota. In addition, shrimp gut microbiota did not simply mirror the rearing bacterioplankton communities. The gut microbiota tended to be governed by variable selection over shrimp ontogeny, while the rearing bacterioplankton community was shaped by homogeneous selection. However, the determinism of rare and stochasticity of abundant subcommunities were consistent between shrimp gut and rearing water. These findings highlight the importance of independently interpreting host-associated and free-living communities, as well as their rare and abundant subcommunities for a comprehensive understanding of the ecological processes that govern microbial successions.

Keywords: shrimp gut microbiota, bacterioplankton community, temporal succession, SourceTracker, ecological processes, rare and abundant sub-communities

INTRODUCTION

Litopenaeus vannamei is one of the most valuable shrimp species in aquaculture globally, while its production is being threatened by diverse diseases. It is now widely recognized that the gut microbiota contributes indispensable roles in sustaining host health (Xiong, 2018). For this reason, intensive studies have focused on factors shaping the gut microbiota, including life stage (i.e., larva, juvenile, or adult), disease (Lu et al., 2022), and surrounding environmental factors (Xiong et al., 2017; Hou et al., 2018). In contrast, we know little about how the biotic sources, e.g., younger host, affect the shrimp gut microbiota.

It is supposed that aquatic animals have a close association with their surrounding water microbiomes (De Schryver et al., 2014). However, survey studies show that the gut microbiota of shrimp is distinct from that in rearing water and/or sediment (Huang et al., 2018; Song et al., 2020). Our recent work evaluates to what extent rearing water and sediment bacterial communities affect the gut microbiota of shrimp, illustrating that shrimp acquire little of their gut commensals from rearing water (Xiong et al., 2019b). Instead, 66.7% gut commensals of the adult shrimp are derived from their juveniles (Zhang et al., 2021). In accordance, ample evidence has shown that the structures of gut microbiota are differed significantly along shrimp life stages, which are distinct from those in the rearing water (Burns et al., 2016; Yan et al., 2016; Zhang et al., 2018). In this context, temporal changes in the shrimp gut microbiota are not parallel with those in rearing bacterioplankton. However, it seems that the gut microbiota of larval shrimp is more similar with the rearing bacterioplankton community, compared with the adults (Burns et al., 2016; Xiong et al., 2018). A possible explanation is that the selection on external taxa is increased as host matured (Xiong et al., 2019a; Xiao et al., 2021). However, the deteriorated water quality imposes stress on shrimp, which in turn depresses their capability of filtering on external taxa (Xiong et al., 2017). For example, nutrient accumulation in rearing water significantly alters shrimp gut microbiota at later farming stage (Lucas et al., 2010; Zhang et al., 2014; Xiong et al., 2016). Accordingly, there is non-linear trend in the relative importance of deterministicity in governing the gut microbiota over shrimp development (Xiong et al., 2019b). It is now recognized that the gut microbiota is conjointly affected by rearing geochemical variables, bacterioplankton, and host development (Xiong et al., 2019b; Xiao et al., 2021), while little is known on the interplay among these variables. Theoretical evidence has proposed that the successional pattern of host-associated (e.g., gut microbiota) communities is distinct from that of free-living bacteria (e.g., bacterioplankton) (Baselga, 2010; Xiong et al., 2019b), whereas experimental evidence is lacking. For these reasons, it remains unclear how and to what extent the gut microbiota is affected by rearing bacterioplankton community as shrimp aged, whereas this knowledge is fundamental for sustaining a health aquaculture.

A microbial community is comprised by a large number of rare species and a few highly abundant taxa (Brown et al., 2014). It is becoming clear that rare biosphere is functionally and ecologically important in a given community (Lynch and Neufeld, 2015). For example, rare taxa serve a reservoir that can quickly respond to environmental changes, thereby promoting community stability in a wide variety of ecosystems (Shade et al., 2014). Additionally, rare subcommunity also contributes dispensable roles in nutrient cycling (Pester et al., 2010). Available studies have depicted that the freeing-living rare and abundant communities exhibit contrasting assembly processes (Mo et al., 2018). However, it remained uncertain whether the host-associated counterparts are ruled by the same ecological processes, as what has been observed for freeing-living community.

An ultimate goal of microbial ecology is to predict the responses of microbial communities to changing environments, yet this goal is difficult to achieve. One reason for this challenge is that there are two types of ecological processes, deterministicity and stochasticity, governing the microbial assembly (Van Der Gast et al., 2008). Deterministic processes include abiotic/biotic selection and biological interaction, while stochastic processes (also known as neutral processes) include dispersal-related processes and ecological drift (Venkataraman et al., 2015; Zhou and Ning, 2017). It has been perceived by ecologists that both deterministic and stochastic processes occur simultaneously in assembling local communities (Chase, 2010; Zhou et al., 2014), whereas no consensus has emerged regarding their relative importances. For a given community, if it is tailored by the dominance of deterministic processes, the temporally successional trend is predictable (Vanwonterghem et al., 2014). Intriguingly, it has been shown that the degrees of deviation in the gut microbiota from the successional trajectory as host aged are positively associated with the severity of the shrimp disease (Xiong et al., 2015). In this regard, it is essential to explore the underlying ecological processes governing the succession of gut microbiota over shrimp ontogeny.

Herein, we explored the successional rules of both the gut and the rearing water bacterial communities over the entire shrimp farming. The main purposes were (1) to evaluate the interplay among biotic (shrimp age), abiotic (water geochemical variables) factors, bacterioplankton community, and the shrimp gut microbiota; (2) to quantify the relative importances of external and internal sources to the gut microbiota over shrimp ontogeny; (3) to compare the underlying ecological processes governing the shrimp gut and rearing water bacterial communities, including total, abundant, and rare communities, by integrating multiple ecological approaches.

MATERIALS AND METHODS

Experimental Design and Sample Collection

Larval shrimp (*L. vannamei*) were introduced into 60 identical greenhouse ponds (concrete and rectangular, 30 m × 60 m, with a depth of 1.2 m) on April 8, at Zhanqi, Ningbo, eastern China (29°32′N, 121°31′E). One week later, both shrimp and rearing water samples were collected with an interval of 6–10 days from six selected ponds over the entire shrimp farming (from 15 April to 10 July). In order to remove

microorganisms and suspended particles, rearing seawater was disinfected with sodium hypochlorite and alum and then aerated in open reservoirs for 3 weeks before usage. To reduce the spatial variability, water samples were collected from four representative points (similar locations in all ponds) and then pooled to compose one biological sample for a given pond. Water samples were stored in an icebox and were transported to laboratory for further processing. In total, we collected 144 samples (6 replicates × 12 samplings × 2 habitats) for microbial community analysis.

Water temperature (WT), pH, salinity (SAL), and dissolved oxygen (DO) were recorded *in situ* using corresponding probes (Oxi 340i; WTW, Weilheim, Germany) at a depth of 50 cm (below water surface). The concentrations of water total phosphorus (TP) and total nitrogen (TN) were analyzed following seawater analysis standard of China (AQSIQ, 2007).

DNA Extraction, Amplification, and Sequencing

To collect microbial cells, 0.5 L of water sample was prefiltered through nylon mesh (100 μm pore size) and subsequently filtered onto a 0.22-μm membrane (Millipore, Boston, MA, United States) on the sampling day. To obtain high efficiency of DNA extracts, the pooled number of shrimp individuals was decided on the basis of their intestine size. Specifically, every three, two, or one intestine from larval, juvenile, or adult shrimp was pooled to compose one biological sample for each pond, respectively. The filters and shrimp intestines were placed into sterile tubes and were stored at −80°C.

Genomic DNA (gDNA) was extracted using a FastDNA Spin kit (MP Biomedicals, Carlsbad, CA, United States) following the manufacturer's protocols. The V3–V4 regions of bacterial 16S rRNA gene were amplified by primers: 341F (5′−CCTAYGGGRBGCA-SCAG−3′) and 806R (5′−GGACTACNNGGGTATCTAA-3′). For each sample, triplicate 50 μl PCRs were performed which contained 25 ng DNA extracts as template with the following conditions: 25 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 45 s, with a condition of 72°C for 10 min for the final elongation step. The triplicate amplicons for each sample were pooled and purified using a PCR fragment purification kit. Equimolar amounts of amplicons from each sample were pooled and then were sequenced in a single run using the Illumina MiSeq platform (Illumina, San Diego, CA, United States), resulting in 2 × 300 bp paired-end reads.

Processing of Illumina Sequencing Data

The paired-end reads were joined and assigned to samples based on barcode. The merged sequences were analyzed using the QIIME2 pipeline (Caporaso et al., 2010). Specifically, sequences at < 200 bp in length, showed ambiguous bases, or had a mean quality score < 20 were filtered. Then, sequences were binned into operational taxonomic unit (OTU) with 97% cutoff using UCLUST (Edgar, 2010). The most abundant sequence from each OTU was selected as representative and then was taxonomically assigned a closed reference (Greengenes Database, release 13.8) (DeSantis et al., 2006), which enables each identified OTU to have a close relative. After the taxonomy had been assigned, Archaea, Chloroplast, unclassified Bacteria, as well as singletons, were excluded from subsequent analysis.

Statistical Analysis

We defined OTUs with a mean relative abundance of $\geq 0.01\%$ across the samples as "abundant" OTUs, whereas OTUs with a mean relative abundance of $< 0.001\%$ as "rare" OTUs follow the criterion as described elsewhere (Logares et al., 2014; Liu et al., 2015).

All statistical analyses were performed in the R-environment[1] unless otherwise indicated. To improve normality and homoscedasticity, bacterial communities were Hellinger transformed, while environmental variables were normalized by using function decostand in package vegan. Heatmap was used to depict the abundance of the top 20 dominant bacterial genera in the shrimp gut microbiota and those in the bacterioplankton communities. Paired *t*-test (pond served a conditional factor) was used to evaluate the significance ($p < 0.05$ level) of diversity between gut and corresponding water bacterial communities at each sampling. A non-metric multidimensional scaling (NMDS) analysis was used to compare the differences in the structures of rearing water and shrimp gut bacterial communities based on Bray-Curtis distance. The significance between groups was tested using an analysis of similarity (ANOSIM) (Anderson, 2010). Permutational multivariate analysis of variance (perMANOVA) was conducted to quantify the relative contributions of habitat (gut or water), shrimp age (days postinoculation), and their interaction to the variations in bacterial community using the "adonis" function (Anderson, 2010).

The time decay of similarity relationship was used to compare the temporal turnover rate (the slope of the regression) between the gut and water bacterial communities over shrimp farming (Xiong et al., 2014). To account for zero similarity values, bacterial community similarity and lag of shrimp age were ln transformed (Talbot et al., 2014). Here, we treated pond as a conditional factor, thereby enabling us to compare the significance (paired *t*-test) in turnover rate between gut and rearing water communities. The multiple regression on distance matrices (MRM) was further used to determine variables that triggered the temporal turnover of bacterial communities. This approach offers advantages over the traditional partial Mantel test to investigate linear, non-linear or non-parametric relationships between a multivariate response distance matrix and any number of explanatory distance matrices (Legendre et al., 1994; Lichstein, 2007). To minimize the collinearity between environmental factors, we used variable clustering to assess the redundancy of variables by the "VARCLUS" procedure in package Hmisc before applying MRM. Then, a matrix randomization procedure with standardized predictor variables was implemented using package ecodist (Goslee and Urban, 2007). To reduce the effect of spurious relationships between variables, we ran the MRM test twice, after removal of insignificant variables by the first run. The results were reported from the second run. A structural equation

[1] http://www.r-project.org

model (SEM) was used to uncover the interplay among rearing water geochemical variables, bacterioplankton and gut bacterial communities, and shrimp age in AMOS 23.0 (IBM, Chicago, IL, United States) (Byrne, 2001).

SourceTracker analysis was employed to quantify the relative contributions of both external (rearing water bacterioplankton community) and internal (the gut microbiota of adjacent younger shrimp) resources to the shrimp gut microbiota (Knights et al., 2011). This approach analyzed the relative abundance of each OTU share in water or younger shrimp gut with older ones, to calculate the probability that each OTU detected in the shrimp gut was sourced from the rearing water or adjacent younger shrimp.

To evaluate the phylogenetic community assembly, the "standardized effect size" of phylogenetic community structure (ses.MNTD) was calculated for non-random phylogenetic relatedness (MNTD) by the difference between phylogenetic distances in the observed communities vs. null communities generated with 999 randomizations, divided by the standard deviation of phylogenetic distances in the distribution using the Picante package (Kembel et al., 2010; R Core Team, 2013). For a given community, ses.MNTD value less than –2 indicates that the community is more phylogenetically related than expected by chance (determinism), whereas ses.MNTD value greater than +2 indicates that a community is less closely related than expected by chance (stochasticity) (Webb et al., 2002; Stegen et al., 2012). Pairwise phylogenetic turnover between communities was calculated as the MNTD metric (βMNTD) using the "comdistnt" function (abundance.weighted = TRUE) in package picante (Kembel et al., 2010). The community assembly processes were further evaluated by βNTI using the "ses.mntd" function (Kembel et al., 2010) and a null modeling approach (Stegen et al., 2012), respectively. βNTI (the difference between the calculated βMNTD and the null-model estimation) values were quantified by either accounting for βNTI is the number of standard deviations that the observed βMNTD is from the mean of the null distribution. A value of βNTI of > 2 or < -2 indicates greater than or less than the expected phylogenetic turnover, respectively (Stegen et al., 2012).

RESULTS

Temporal Successions of Shrimp Gut and Rearing Water Bacterial Communities

In total, 3,842,244 high-quality sequences, with 17,385–36,579 sequences per sample (mean ± standard deviation, 26,868 ± 4,620) were collected across the enrolled 143 samples. After rarefaction, we obtained 32,811 OTUs in the analysis. Diversity of the gut microbiota was temporally stable over shrimp ontogeny, whereas bacterioplankton diversity linearly increased over the same timeframe, as supported by both Shannon and phylogenetic diversity indices (**Supplementary Figure 1**). The NMDS biplot depicted that the gut microbiota were distinct from bacterioplankton communities (**Figure 1**). There were sequential successions of both the gut microbiota

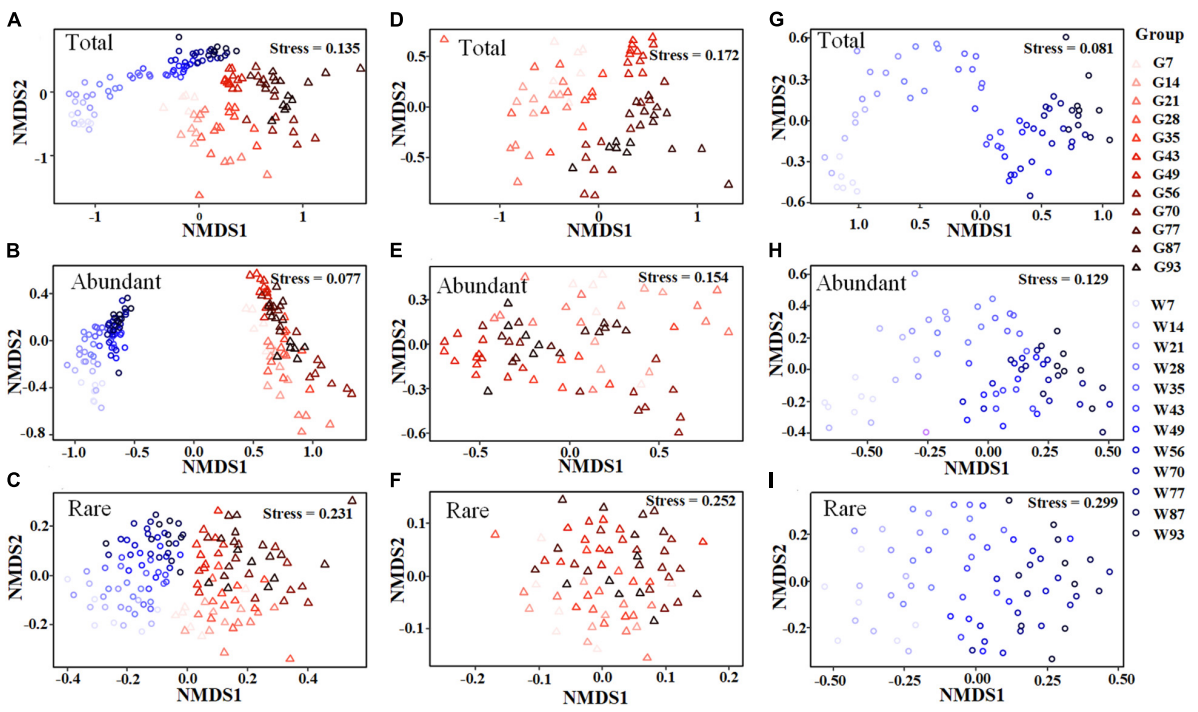


FIGURE 1 | Non-metric multidimensional scaling (NMDS) ordinations showing the structures of total, abundant and rare communities. NMDS ordination of the total, abundant, and rare bacterial communities in both the shrimp gut and rearing water **(A–C)**, shrimp gut **(D–F)**, rearing water **(G–I)** based on Bray-Curtis similarity, respectively. Samples were coded and colored by habitat and shrimp life stage, of which G and W indicate shrimp gut and water, respectively.

and bacterioplankton communities during shrimp farming, as evidenced by increased distances along NMDS axis 1 (**Figure 1**). These differences were confirmed by the ANOSIM, illustrating that shrimp gut bacterial communities were significantly distinct between every paired age (**Supplementary Table 1**). In contrast, there were no significant differences between some adjacent pairs of bacterioplankton communities, e.g., W28 vs. W21, W56 vs. W49, W77 vs. W70, and W93 vs. W87 (**Supplementary Table 2**). Furthermore, perMANOVA revealed that both habitat (shrimp gut or rearing water) and shrimp age significantly ($p < 0.001$ in both cases) contributed to the variations in bacterial community. Shrimp age exerted consistently higher importances than habitat in governing the total, abundant, and rare bacterial communities (**Table 1**). Both the gut microbiotas and bacterioplankton communities exhibited significant time decay of similarity relationship. The temporal turnover rate of bacterial community in shrimp gut (–0.290 ± 0.127) was significantly (paired *t*-test, $p = 0.001$) lower than that in rearing water (–0.827 ± 0.083) (**Figure 2**). MRM revealed that the temporal succession was primarily governed by shrimp age, water temperature, and pH (**Table 2**). Notably, each of the three variables exerted higher contributions in governing the bacterioplankton communities compared with the gut microbiota (**Table 2**), suggesting that bacterioplankton communities were more strongly affected by environmental factors. The MRM model explained 66% ($p < 0.001$) variation in bacterioplankton community, while only 21% ($p < 0.001$) in the gut microbiota. Additionally, shrimp age contributed larger partial regression coefficients in shaping abundant subcommunities than corresponding rare counterparts in both the gut and rearing water.

Factors Governing the Temporal Successions of Bacterial Community

A forward selection procedure identified four water variables (TP, DO, pH, salinity) and shrimp age that significantly contributed to the variations in bacterial community ($p < 0.01$) (**Supplementary Figure 2** and **Supplementary Table 3**). The four water variables were significantly associated with shrimp age, which were attributed to the temporal dynamics of water variables during shrimp farming. Bacterioplankton community was positively correlated with salinity ($\lambda = 0.28$, $p = 0.007$), pH ($\lambda = 0.23$, $p = 0.002$), and DO ($\lambda = 0.14$, $p = 0.048$), and was negatively affected by TP ($\lambda = -0.34$, $p < 0.001$) (**Figure 3** and **Supplementary Table 3**). Shrimp age was significantly associated with bacterioplankton community (0.50). The gut microbiota was affected by the combination of direct (0.34) and weak indirect (0.09) effects of shrimp age (**Figure 3** and **Supplementary Table 3**). Notably, bacterioplankton community contributed a weak and insignificant direct effect on the assembly of shrimp gut microbiota. As expected, abundant and rare subcommunities exhibited significant and positive contributions to corresponding total bacterial communities, with much higher contributions of the abundant subcommunities (**Figure 3** and **Supplementary Table 3**).

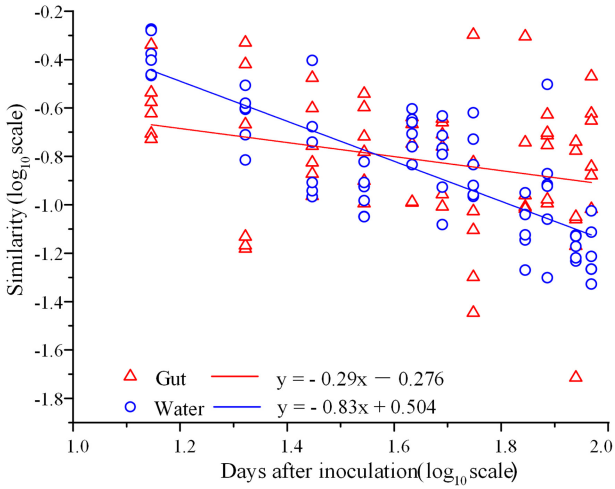


FIGURE 2 | Time-decay relationship for shrimp gut microbiota and bacterioplankton communities. The *x*-axis is log in days postlarval shrimp inoculation, and *y*-axis is log (similarity) calculated using the Bray-Curtis distance ($R = 0.407$, $p = 0.001$).

TABLE 1 | Quantitative effects of sampling time and habitats on variation in community composition using non-parametric permutational multivariate analysis of variance (perMANOVA) with adonis function.

	Age		Habitats		Age:habitats	
	R^2	*P*	R^2	*p*	R^2	*p*
Community structure						
Total	0.306	<0.001	0.150	<0.001	0.431	<0.001
Rare	0.116	<0.001	0.022	<0.001	0.138	<0.001
Abundant	0.315	<0.001	0.167	<0.001	0.455	<0.001

The R^2 values represent the proportion of the community variation constrained by each variable or their interaction.

Sources of Shrimp Gut Commensals Over Shrimp Ontogeny

The relative proportion of shared OTUs between shrimp gut and rearing water was negligible, ranged from 0.26 to 2.11% (**Supplementary Figure 3**). Thus, the rearing water bacterioplankton community contributed minor role in affecting the shrimp gut microbiota. To test whether gut microbiota parallelly changed with rearing water bacterial communities along shrimp farming, temporal dynamics of the top 20 dominant bacterial genera in the shrimp gut were compared with those in the rearing water (**Figure 4**). The relative abundances of *Vibrio*, *Salinivibrio*, and *Haloferula* genera were abundant in shrimp gut but were rare in rearing water. Only six dominant bacterial genera in shrimp gut, such as *Ruegeria*, *Marivita*, and *Flavobacterium*, were positively correlated with these in bacterioplankton communities, but not for the other 12 genera, including *Vibrio* and *Pseudoalteromonas* (**Figure 4A**). A similar pattern was observed for the most rare 20 bacterial genera in the shrimp gut, in which only genera of *Sedimentibacter*, *Cupriavidus*, and *Marinobacterium* were significantly associated with these

TABLE 2 | Results of the multiple regression on distance matrices (MRM) for shrimp gut, bacterioplankton, and their abundant and rare communities.

Habitat	Gut						Rearing water					
	Total		Rare		Abundant		Total		Rare		Abundant	
	R^2 = 0.21	p < 0.001	R^2 = 0.03	p < 0.001	R^2 = 0.15	p < 0.001	R^2 = 0.66	p < 0.001	R^2 = 0.27	p < 0.001	R^2 = 0.65	p < 0.001
Age	0.181	<0.001	0.007	0.001	0.118	<0.001	0.421	<0.001	0.229	<0.001	0.541	<0.001
WT	0.093	0.01	ND	ND	0.089	<0.001	0.301	<0.001	ND	ND	0.299	<0.001
pH	0.029	0.04	0.009	0.002	0.024	0.008	0.364	<0.001	ND	ND	0.358	<0.001
SAL	NS	NS	ND	ND	ND	ND	0.514	<0.001	0.198	<0.001	0.5	<0.001
BOD	NS	NS	ND	ND	ND	ND	0.216	0.002	ND	ND	0.216	<0.001
DO	ND	ND	ND	ND	ND	ND	0.258	<0.001	0.13	<0.001	0.256	<0.001
TP	ND	ND	0.014	<0.001	ND	ND	ND	ND	0.076	<0.001	ND	ND
TN	ND	ND	0.001	NS	ND	ND	0.254	<0.001	ND	ND	ND	ND

ND, not determined (removed by the VARCLUS results); NS, not significant.

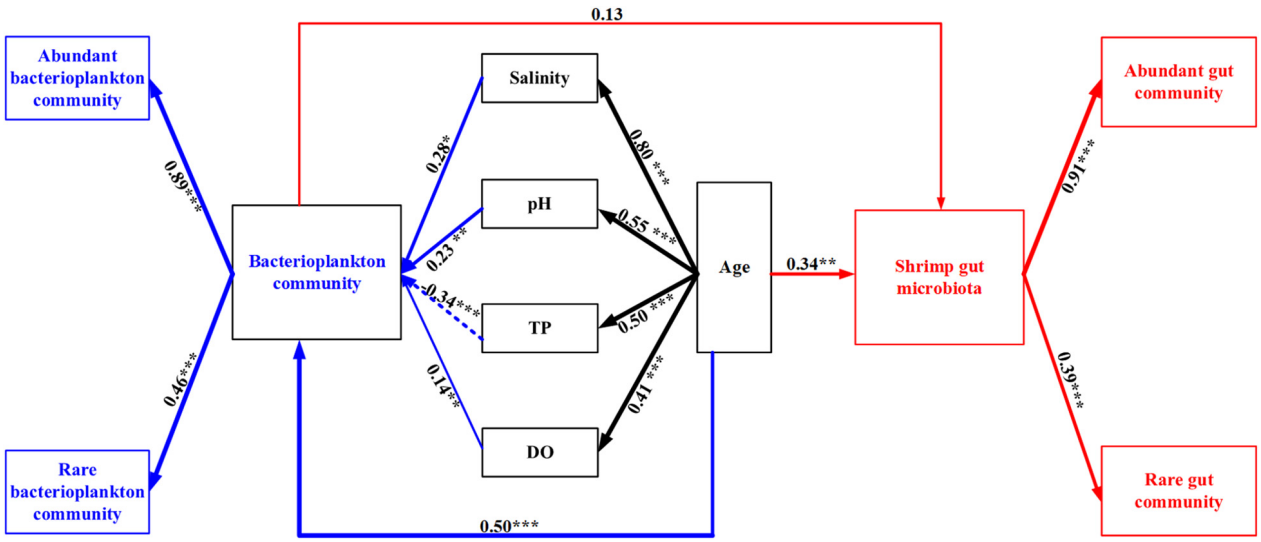


FIGURE 3 | Structural equation modeling (SEM) shows the effect of environmental factors and bacterioplankton community on shrimp gut community. The numbers on arrows indicate standardized path coefficients. Arrow widths show the strength of the causal relationship. Arrows in blue and red indicate the effects on the bacterioplankton community and the shrimp gut microbiota, respectively. Solid and dashed lines indicate positive and negative correlations, respectively. *p < 0.05, **p < 0.01, ***p < 0.001.

in rearing water (**Figure 4B**). Overall, the compositions and abundances of the bacteria in gut were insignificantly affected by the rearing bacterioplankton community over shrimp ontogeny.

SourceTracker analysis was used to quantify the relative contribution of external sources (rearing water) on the shrimp gut microbiota at each sampling (**Figure 5A**). In general, larval shrimp (breeding days less than 35) sourced little commensals from rearing water compared with juveniles and adults, with the exception on day 56. Bacterioplankton community contributed 43.3% (averaged contribution) of the species to shrimp gut microbiota, whereas most of the source was unknown (**Figure 5A**). When integrating the adjacent younger shrimp as an internal source for the gut microbiota in the model, the relative contribution of rearing water to gut microbiota sharply decreased to 19.1% (averaged proportion, ranged from 1.22 to 54.7%). Instead, gut commensals were primarily derived from the adjacent younger shrimp, with an averaged contribution of 60.3%. Accordingly, the proportion of unknown source of gut microbiota sharply decreased to 20.5% (**Figure 5B**). Taken together, the majority of gut commensals sourced little species from surrounding species pool, which were temporally sustained over shrimp ontogeny.

Ecological Processes Govern the Successions of Bacterial Community

The ses.MNTD values of gut and rearing water bacterial communities were significantly lower than zero, suggesting that the two communities tended to be phylogenetically clustering (**Supplementary Figure 4**). Additionally, most of the βNTI values of gut microbiotas and bacterioplankton communities were less than −2, indicating that the dominant role of deterministic processes assembled the gut microbiota and bacterioplankton

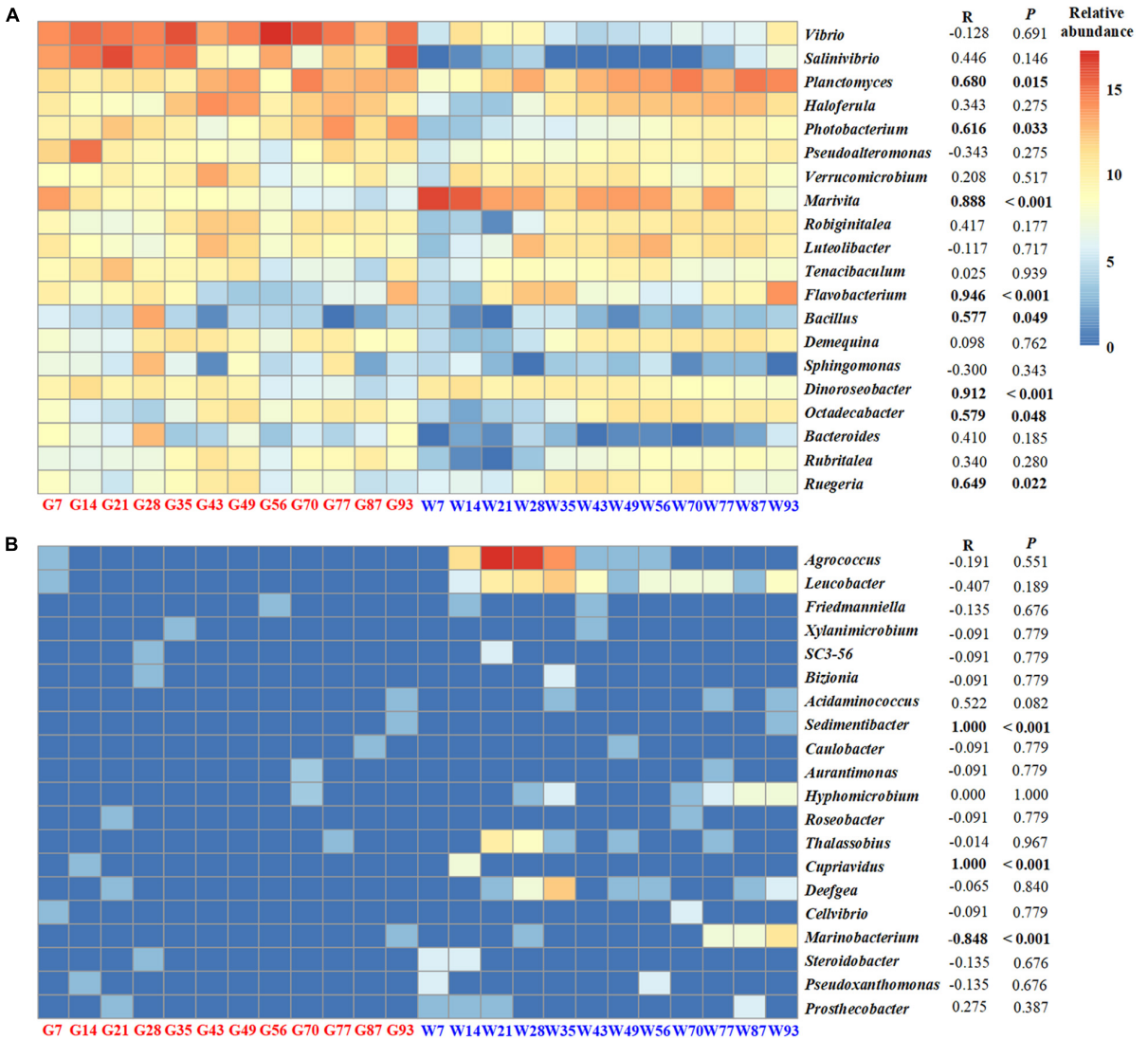


FIGURE 4 | The relative abundances of the top abundant 20 bacterial genera **(A)** and the most rare 20 bacterial genera **(B)** in the gut and those in bacterioplankton communities. Pearson's correlations and the significances of the genera between shrimp gut and rearing water are shown on the right.

community (**Supplementary Figure 5**). To evaluate the trends of ecological processes over shrimp ontogeny, βNTI values were regressed against the lag of shrimp age. The temporal trends of βNTI were different between shrimp gut and rearing water bacterial communities (**Figure 6**). Specifically, there was a significant and positive correlation (R = 0.16, p = 0.034) between βNTI values of the total gut microbiota as shrimp aged (**Figure 6A**), whereas those of total bacterioplankton community exhibited the opposing trend (R = –0.22, p = 0.003) (**Figure 6B**). There were no significant correlations between βNTI values of the abundant subcommunities in gut or rearing water during shrimp farming. In addition, most of the βNTI values of abundant subcommunity were between –2 and 2, indicating the dominant role of stochastic processes in assembling abundant subcommunity (**Figures 6C,D**). In contrast, the βNTI values of rare subcommunities in both the gut (R = –0.25, p = 0.001) and rearing water (R = –0.61, $p < 0.001$) significantly decreased over the same timeframe, which tended to be less than –2 (**Figures 6E,F**).

DISCUSSION

Despite recent progress, little is known about the underlying ecological processes governing the successional patterns of host-associated microbes, especially their abundant and rare counterparts. To address this pressing knowledge gap, we explored how and to what extent abiotic (water geochemical variables) and biotic (i.e., younger shrimp, host age, and rearing bacterioplankton community) factors affected the gut microbiota over shrimp ontogeny. In addition, we quantified the relative contributions of external (rearing water) and internal

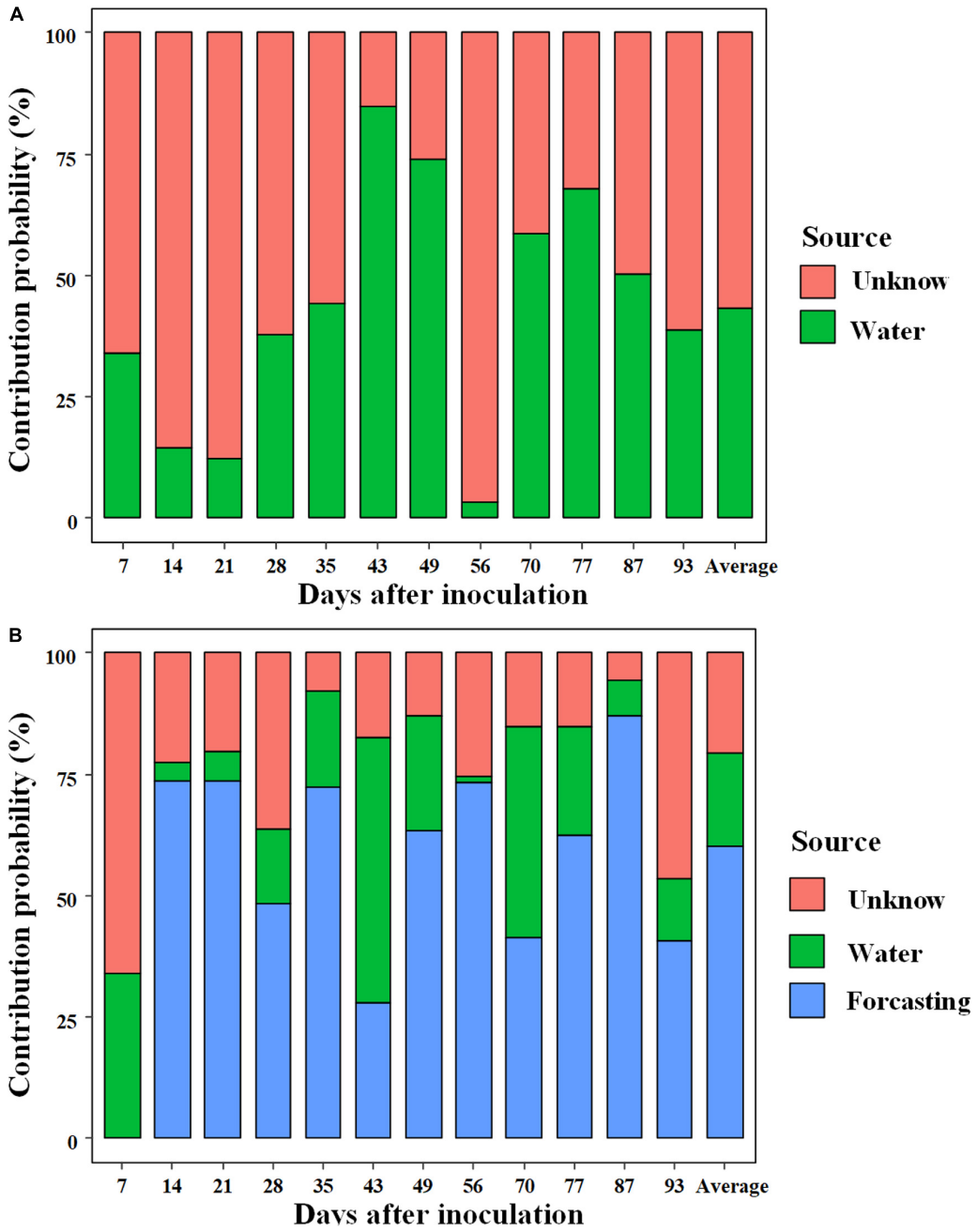


FIGURE 5 | SourceTracker analyzes the relative contribution of external (rearing water) **(A)** and internal (gut commensals of the adjacent younger shrimp) sources **(B)** to the shrimp gut commensals.

(adjacent younger shrimp gut microbiota) sources on the shrimp gut commensals. These findings yield novel insights into the assembly of gut microbiota over shrimp ontogeny from an ecological perspective. It is worthy to note that we used OTU clustering methods instead of the more recently developed amplicon sequence variants (ASVs). However, it has been shown that all α and β diversity metrics are highly positively correlated ($r > 0.90$) between samples analyzed with either ESVs or traditional OTUs. ESV or OTU methods often reveal similar ecological results, with indistinguishable statistical inferences (Glassman and Martiny, 2018). Similarly, a recent study depicts that OTUs and ASVs produce comparable shrimp microbiota (García-López et al., 2021). Thus, standard microbial community analyses are not overly sensitive to the particulars of binning approaches (Glassman and Martiny, 2018). In addition, we used a closed reference for taxonomical assignment, which has excluded spurious taxa. For these reasons, our findings are not biased by OTU clustering methods.

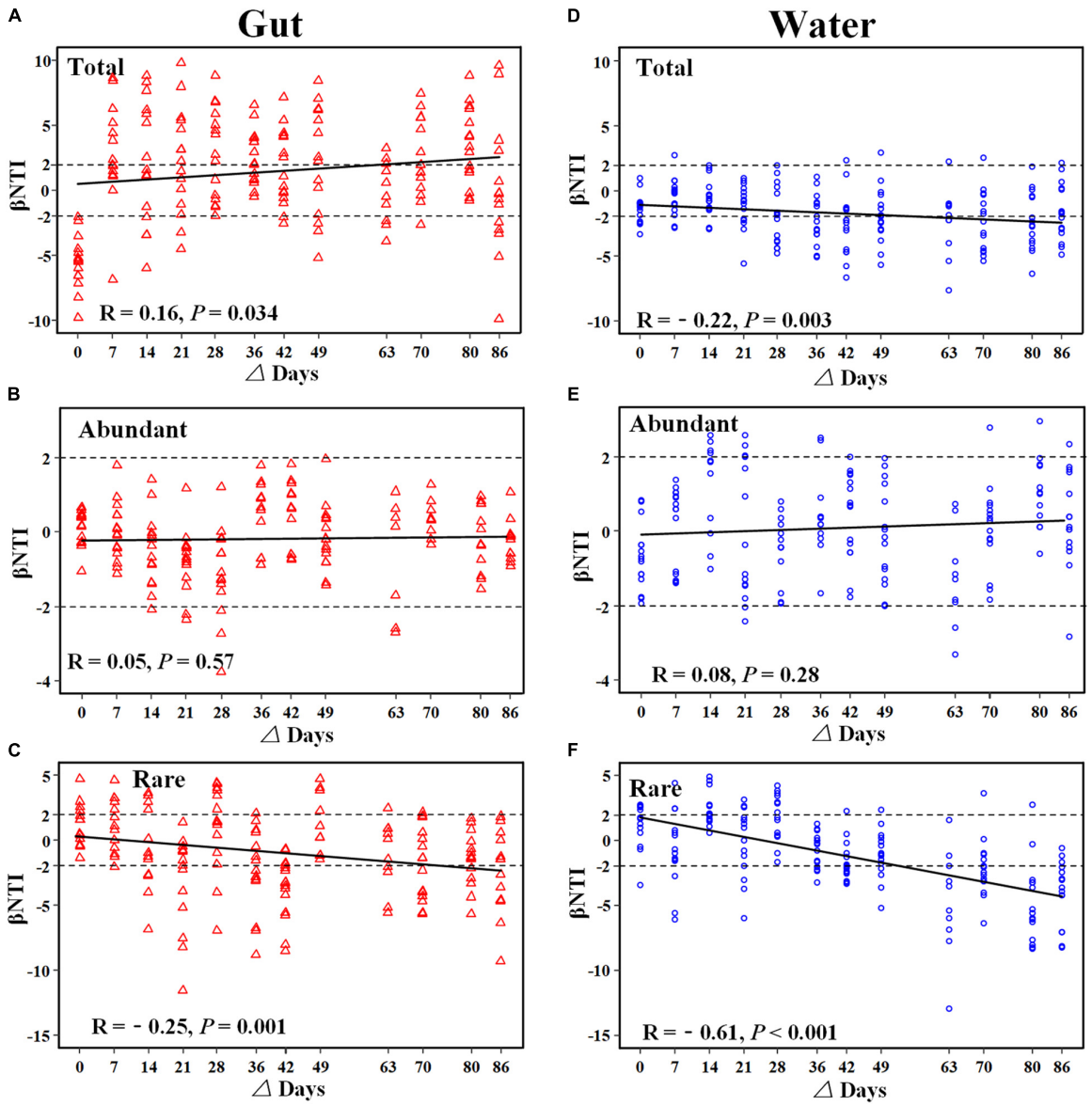


FIGURE 6 | Relationship between βNTI values of bacterial community over shrimp ontogeny. The horizontal dashed lines indicate the βNTI values of –2 and +2. An individual community below or above the two dashed lines indicates that determinism dominantly governs the community assembly, while between the two dashed lines indicates that stochasticity is dominant. The relationship between βNTI and differences in days gut microbiota **(A–C)** and bacterioplankton community **(D–F)** was fitted using linear regression.

Succession Pattern Between Bacterioplankton Community and Shrimp Gut Community

The diversity of shrimp gut microbiota was relatively stable, whereas the surrounding bacterioplankton diversity increased linearly during shrimp farming. Additionally, the diversity in gut microbiota was generally lower than that in corresponding bacterioplankton community (**Supplementary Figure 1**), in accordant with the notion that hosts select a subset of surrounding taxa that colonize into their gut (Zoqratt et al., 2018). In contrast, the linearly increased diversity of bacterioplankton could be attributed to the accumulation of nutritional sources along shrimp farming (**Supplementary Figure 2**), leading to the diversification of microbes. Both the shrimp gut microbiota and bacterioplankton community exhibited sequential changes over shrimp ontogeny (**Figure 1**). Consistently, there is ample evidence that the gut microbiota is primarily affected by host age in diverse aquatic animals (Burns et al., 2016; Yan et al., 2016). Given the sequential changes in both the gut microbiota and bacterioplankton communities (**Figure 1**), we compared temporal turnover rate of the two communities. Bacterioplankton community exhibited significantly steeper turnover (more rapid deviation from original

to new state) than gut microbiota (**Figure 2**), indicating that host-associated communities are more temporally stable than free-living counterparts. This finding was further supported by the MRM model, revealing that the succession of gut microbiota (R^2 = 0.21) were less explained by the shrimp age, compared with that of bacterioplankton community (R^2 = 0.66) (**Table 2**). It seems that host gut offers a relatively stable microenvironment for commensals. Consistent with this assertion, it has been shown that fish gut is a more suitable environment than external skin mucus (Sylvain et al., 2020). In other words, gut microbiota experiences regular incremental shifts such as the maturity of physiology and immunity over shrimp ontogeny. Additionally, the rapid succession of bacterioplankton community could be attributed to temporally varied water geochemical variables during shrimp farming (**Supplementary Figure 2**), instead of direct role of shrimp age itself (see details in **Figure 3**), although we have tried to minimize colinearity among environmental factors. Indeed, the important role of shrimp age in governing both communities did not completely rule out the effects of other factors. We also detected that water temperature and pH strongly affected the temporal successions of total and abundant bacterial communities in both habitats but not the rare subcommunities (**Table 2**). Water temperature and pH have been extensively observed as key factors in shaping bacterioplankton community (Luo et al., 2019; Nyirabuhoro et al., 2019). Interestingly, shrimp are poikilotherms, while their gut microbiotas are sensitive to changes in rearing water temperature (23.8–30.7°C) (**Supplementary Figure 2**). Changes in the environmental temperature of aquaculture water could affect the metabolic and physiological functions of shrimp, thus indirectly alter the gut microbiotas. In accordance, it has been shown that water temperature affects feeding, growth, and survival of *Litopenaeus vannaei*. Rearing water temperature was positively associated with the abundances of anaerobes and the anaerobic Bifidobacterium (Li et al., 2018). Considering the functional importance of gut microbiota in host health, this pattern may partially explain why a sudden change in water temperature generally causes shrimp disease (Estrada-Perez et al., 2020). Furthermore, water temperature could directly alter the bacterioplankton community (Yang et al., 2018), which in turn affects shrimp gut microbiota. It is worthy to emphasize that lifestyle is also key factor in shaping the gut microbiota of shrimp (Cornejo-Granados et al., 2018). More specifically, there are distinct gut microbiotas between wild and aquacultured shrimp (Cornejo-Granados et al., 2017), low- and high-salinity–cultured shrimp (Hou et al., 2020), and freshwater and marine conditions (Cornejo-Granados et al., 2018). In this regard, follow-up investigations following a spatial sampling strategy is needed to test whether the pattern observed here is shared between culture ecosystems or ecosystem dependent.

Interplay Among Geochemical Factor, Bacterioplankton Community, and Gut Microbiota Over Shrimp Ontogeny

SEM uncovered that rearing water salinity, pH, TP, and DO were the key determinants in driving the succession of bacterioplankton communities (**Figure 3**), corroborating recent studies obtained in shrimp, crab, and tilapia aquaculture conditions (Giatsis et al., 2015; Yang et al., 2018; Dai et al., 2020; Hou et al., 2020). Water DO, pH, and salinity directly affected the structures of bacterioplankton community (**Figure 3**), in accordance with the notion that bacterioplankton communities are extremely sensitive to subtle environmental changes (Or et al., 2012; Yang et al., 2018). In aquaculture ecosystem, the level of water phosphorus is usually low, thus bacterioplankton communities often experience P-unsaturation (Duhamel et al., 2021). Accordingly, the concentration of TP was significantly and positively correlated with the structures of bacterioplankton community (Liu et al., 2019). However, we found that a sharp increase in TP at the later farming stage exerted a negative effect on bacterioplankton communities (**Figure 3** and **Supplementary Figure 2**). In contrast, there were negligible and insignificant effects of water salinity, pH, TP, and DO on the gut microbiota (**Figure 3** and **Supplementary Table 3**). A possible explanation for this pattern is that host could buffer external environmental change. As a result, the gut microbiota is less affected by water geochemical factors. Bacterioplankton community only exerted a weak direct effect on the gut microbiota, compared with shrimp age (**Figure 3**). Consistently, ample evidence has shown that the gut microbiota in aquatic animals is distinct from surrounding environments (Zhang et al., 2018, 2021; Xiong et al., 2019b). It has been proposed that hosts selectively filter particular bacteria from the rearing environments, rather than randomly ingesting surrounding taxa (Stephens et al., 2016; Yan et al., 2016). In accordance, shrimp age (a proxy of gut maturity) is the main biological variable governing the succession of gut microbiota (**Figures 2**, **3** and **Table 2**). Together, changes in the geochemical variables strongly affect the structures of bacterioplankton community during shrimp farming, whereas the gut microbiota does not simply mirror the rearing bacterioplankton community.

Shrimp Gut Commensals Sourced From Their Larvae

Although shrimp live in rearing water, relative abundances of the top 20 bacterial genera in shrimp gut were insignificantly associated (12 in 20 cases) with those in bacterioplankton community (**Figure 4**). For example, *Vibrio* genus was predominant in shrimp gut but was rare in rearing water. This is consistent with previous studies showing that the dominant genera are distinct between shrimp gut and rearing water and sediment (Zhou et al., 2021). *Vibrio* and *Photobacterium* members are long known to be opportunistic pathogens in shrimp aquaculture (Manilal et al., 2010). Nevertheless, the majority of vibrios are not pathogenic, many are commensal or even beneficial, including the carbon cycle and osmoregulation (Johnson, 2013). Indeed, *Vibrio* species have been frequently detected as a dominant population in shrimp gut (Chaiyapechara et al., 2012). In accordance, a few vibrio strains, e.g., *Vibrio alginolyticus* UTM 102, have been applied as probiotics in shrimp aquaculture (Balcázar et al., 2007). Similarly, *Photobacterium* strains are common in the intestinal contents of marine animals (Chaiyapechara et al., 2012). In addition, *Ruegeria*, *Marivita*,

and *Flavobacterium* harbor the specific ability in degrading organic matter (Williams et al., 2012; Sun et al., 2019; Huang et al., 2020), which were enriched at the late stage of aquaculture water (**Figure 4**). The relative abundance of *Pseudoalteromonas* in shrimp gut was negatively associated with that in rearing water (**Figure 4**). Consistently, *Pseudoalteromonas* strains have been successfully used as probiotics in shrimp farming. In these regard, shrimp could select some beneficial commensals that improve their fitness.

Furthermore, we evaluated the contribution of internal sources (gut commensals of younger host) on the shrimp gut microbiota. Theoretically, aquatic animals are born without microorganisms, thus their gut commensals should source from the surrounding environments after birth (Yan et al., 2016). However, the majority of gut commensals of shrimp gut microbiota sourced from their younger host, rather than bacterioplankton communities (**Figure 5**), which reinforces the importance of the gut microbiome in younger host (Kerr et al., 2015). Similarly, it has been shown that shrimp acquires a small proportion of commensals from rearing water over development (Xiong et al., 2019b). The contribution of surrounding bacterioplankton communities on gut commensals markedly varied over shrimp ontogeny (**Figure 5**). We propose several explanations for this pattern. According to the co-evolution hypothesis (McFall-Ngai et al., 2013), it is mandatory for larva to recruit suitable taxa that expand the range of diet digestion due to incomplete digestive system. Thus, to improve hosts' fitness, the colonization of gut commensals is filtered from rearing species pool as a result of deterministic processes. However, as host matures, the initial "winners" could be reassembled, thereby resulting in host stage-specific gut microbiota (Stephens et al., 2016; Yan et al., 2016). Additionally, temporal dynamics of environmental variables directly alter the bacterioplankton communities (**Figure 4**), leading to non-linear contribution of bacterioplankton communities to gut microbiota along shrimp farming. The skewed source pattern on day 56 could be induced by sudden increase in TP content and the low level of DO (**Figure 2**). In accordance, the SEM uncovered that water TP and DO exerted indirect effects on the gut microbiota (**Supplementary Table 3**). However, there were still a high proportion of "unknown" sources (**Figure 5**), which could be attributed to the uncollected species pools, e.g., diet, air and farmer. Together, gut commensals primarily source from adjacent younger shrimp. In this regard, we propose the isolation of probiotics from larval gut, which could be persist over shrimp ontogeny.

Ecological Processes Governing the Assembly of Bacterial Community

Bacterioplankton communities are more closely phylogenetically clustered than the gut microbiotas, as supported by significantly lower mean value of ses.MNTD (**Supplementary Figure 4**), as observed in the present study and elsewhere (Xiong et al., 2019b). In addition, the βNTI values of gut microbiota and bacterioplankton community divergently changed during shrimp farming (**Figures 6A,B**), though both communities exhibited sequential shifts in the community structure (**Figures 1**, **2**). The gut microbiota tended to be governed by variable selection (βNTI values > 2), while the bacterioplankton community was affected by homogeneous selection (βNTI values < -2) (**Figures 6A,B**). The logic behind this may be that shrimp has not reached full maturity, though we collected samples over an entire shrimp farming (Lucas et al., 2010). Consistent with this assertion, the gut microbiota significantly changed between every paired sampling (**Supplementary Table 1**). Similarly, it has been shown that the succession of shrimp gut microbiota is more driven by species replacement than bacterioplankton community (Xiong et al., 2019b). In contrast, geochemical variables of rearing water were relatively stable at the later farming days (**Supplementary Figure 2**), thus bacterioplankton communities were governed by homogeneous selection. Accordingly, the structures of bacterioplankton community were comparable between some adjacent pairs (**Supplementary Table 2**). Thus, host-associated and free-living bacterial communities are governed by different ecological processes. Considering the functional importance of gut microbiota in host health, additional works are required to explore the underlying ecological processes in governing the overlooked host-associated microbes. Notably, rare subcommunities in both the gut and rearing water were affected by homogeneous selection (**Figures 6E,F**), whereas their abundant counterparts were shaped by random processes (**Figures 6C,D**). Rare members serve as "seed bank" in a given community, which could switch to abundant taxa in response to changing environments (Magurran and Henderson, 2003). That is, rare taxa adapt to specific conditions that are strongly selected by external factors. In accordance, rare subcommunities were governed by deterministic processes (**Figure 6**). Corroborating recent works, rare subcommunity shown to be dominated by deterministic processes, while abundant subcommunity is influenced largely by stochastic processes in agricultural soils (Jiao and Lu, 2020) and freshwater ecosystems (Liu et al., 2015). The broad fitness of abundant taxa facilitates their successive establishment across a wide range of environmental conditions (Wan et al., 2021), e.g., variations in host maturity and geochemical factors here. By this logic, the abundant subcommunities are less affected by local variables, leading to the predominance of stochasticity (**Figure 6**). That is, no phylogenetic signs were detected for abundant communities. Consistently, there is ample evidence that rare taxa exhibit greater sensitivity to environmental factors than abundant species (Mo et al., 2018). Under these scenarios, it seems that the rare subcommunities are governed by deterministic processes, while the assembly of their abundant counterparts was stochastic across habitats, such as host gut and rearing water here.

CONCLUSION

Host-associated bacterial community is more temporally stable than their free-living counterpart, as supported by the significant lower temporal turnover rate. In accordance, the gut microbiota is less affected by the rearing water geochemical variables, compared with bacterioplankton community. Intriguingly, the shrimp gut microbiota does not simply mirror the rearing bacterioplankton communities. Instead, gut commensals mainly

inherit from their younger shrimp, rather than the rearing water. It seems that host-associated and free-living microbes are assembled by divergently ecological processes. That is, the gut microbiota is governed by variable selection over shrimp ontogeny, while the rearing bacterioplankton community is shaped by homogeneous selection. However, the determinism of rare and stochasticity of abundant subcommunities are consistent between shrimp gut and rearing water. These findings greatly broaden our understanding on the underlying ecological processes governing the temporal successions of host-associated and free-living microbial communities.

DATA AVAILABILITY STATEMENT

Raw sequence data are available in the BIG Data Center, Chinese Academy of Sciences, under code CRA004710 at http://bigd.big.ac.cn/gsa.

ETHICS STATEMENT

The animal study was reviewed and approved by the Guide for the Care and Use of Laboratory Animals of Ningbo University.

AUTHOR CONTRIBUTIONS

JC and JX conceived and designed the research. WZ, ZZ, and QQ conducted the experiments. JX contributed analytical tools. JX and WZ analyzed the data. WZ wrote the manuscript with help from JX and ZZ. All authors read and approved the manuscript.

FUNDING

This work was financially supported by the Natural Science Fund for Distinguished Young Scholars of Zhejiang Province (LR19C030001), the National Natural Science Foundation of China (31872693 and 32071549), the Key Public Welfare Technology Application Research Project of Ningbo (202002N3032), and the K.C. Wong Magna Fund in Ningbo University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.752750/full#supplementary-material

REFERENCES

Anderson, M. J. (2010). A new method for non-parametric multivariate analysis of variance. *Aust. Ecol.* 26, 32–46. doi: 10.1046/j.1442-9993.2001.01070.x

AQSIQ (2007). The Specification for Marine Monitoring of China-Part 4: Seawater Analysis (GB 17378.4e2007). General Administration of Quality Supervision, Inspection and Quarantine (AQSIQ) of the People's Republic of China. Beijing: China Standard Press (in Chinese).

Balcázar, J. L., Rojas-Luna, T., and Cunningham, D. P. (2007). Effect of the addition of four potential probiotic strains on the survival of pacific white shrimp (*Litopenaeus vannamei*) following immersion challenge with *Vibrio parahaemolyticus*. *J. Invertebr. Pathol.* 96, 147–150. doi: 10.1016/j.jip.2007.04.008

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecol. Biogeogr.* 19, 134–143. doi: 10.1111/j.1466-8238.2009.00490.x

Brown, M. V., Ostrowski, M., Grzymski, J. J., and Lauro, F. M. (2014). A trait based perspective on the biogeography of common and abundant marine bacterioplankton clades. *Mar. Genomics* 15, 17–28. doi: 10.1016/j.margen.2014.03.002

Burns, A. R., Stephens, W. Z., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., et al. (2016). Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J.* 10, 655–664. doi: 10.1038/ismej.2015.142

Byrne, B. M. (2001). *Structural Equation Modeling with AMOS. Basic Concepts Applications & Programming*. Mahwah NJ: Lawrence Erlbaum Associates Publishers.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Chaiyapechara, S., Rungrassamee, W., Suriyachay, I., Kuncharin, Y., Klanchui, A., Karoonuthaisiri, N., et al. (2012). Bacterial community associated with the intestinal tract of P. Monodon in commercial farms. *Microb. Ecol.* 63, 938–953. doi: 10.1007/s00248-011-9936-2

Chase, J. M. (2010). Stochastic community assembly causes higher biodiversity in more productive environments. *Science* 328, 1388–1391. doi: 10.1126/science.1187820

Cornejo-Granados, F., Gallardo-Becerra, L., Leonardo-Reza, M., Ochoa-Romo, J. P., and Ochoa-Leyva, A. (2018). A meta-analysis reveals the environmental and host factors shaping the structure and function of the shrimp microbiota. *PeerJ* 6:e5382. doi: 10.7717/peerj.5382

Cornejo-Granados, F., Lopez-Zavala, A. A., Gallardo-Becerra, L., Mendoza-Vargas, A., Sánchez, F., Vichido, R., et al. (2017). Microbiome of Pacific Whiteleg shrimp reveals differential bacterial community composition between wild, aquacultured and AHPND/EMS outbreak conditions. *Sci. Rep.* 7:11783. doi: 10.1038/s41598-017-11805-w

Dai, W., Xiong, J., Zheng, H., Ni, S., Ye, Y., and Wang, C. (2020). Effect of *Rhizophora apiculata* plantation for improving water quality, growth, and health of mud crab. *Appl. Microbiol. Biotech.* 104, 6813–6824. doi: 10.1007/s00253-020-10716-7

De Schryver, P., Defoirdt, T., Sorgeloos, P., and Rall, G. F. (2014). Early mortality syndrome outbreaks: a microbial management issue in shrimp farming? *PLoS Pathog.* 10:e1003919. doi: 10.1371/journal.ppat.1003919

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072. doi: 10.1128/AEM.03006-05

Duhamel, S., Diaz, J. M., Adams, J. C., Djaoudi, K., Steck, V., and Waggoner, E. M. (2021). Phosphorus as an integral component of global marine biogeochemistry. *Nat. Geosci.* 14, 359–368. doi: 10.1038/s41561-021-00755-8

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Estrada-Perez, N., Ruiz-Velazco, J. M., Magallon-Barajas, F. J., Campa-Cordova, A. I., and Hernández-Llamas, A. (2020). Dynamic stock model for analysing semi-intensive production of whiteleg shrimp *Litopenaeus* (*Penaeus*) *vannamei* affected by the acute hepatopancreatic necrosis disease: assessment of disease severity indicators and relationships with pond water quality parameters. *Aquac. Res.* 51, 242–254. doi: 10.1111/are.14370

García-López, R., Cornejo-Granados, F., Lopez-Zavala, A. A., Cota-Huízar, A., and Ochoa-Leyva, A. (2021). OTUs and ASVs produce comparable taxonomic and

diversity from shrimp microbiota 16S profiles using tailored abundance filters. *Genes* 12:564. doi: 10.3390/genes12040564

Giatsis, C., Sipkema, D., Smidt, H., Heilig, H., Benvenuti, G., Verreth, J., et al. (2015). The impact of rearing environment on the development of gut microbiota in tilapia larvae. *Sci. Rep.* 5:18206. doi: 10.1038/srep18206

Glassman, S. S., and Martiny, J. B. H. (2018). Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. *mSphere* 3:e00148-18. doi: 10.1128/mSphere.00148-18

Goslee, S. C., and Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *J. Statal. Softw.* 22, 1–19. doi: 10.18637/jss.v022.i07

Hou, D., Huang, Z., Zeng, S., Liu, J., Wei, D., Deng, X., et al. (2018). Intestinal bacterial signatures of white feces syndrome in shrimp. *Appl. Microbiol. Biotech.* 102, 3701–3709. doi: 10.1007/s00253-018-8855-2

Hou, D., Zhou, R., Zeng, S., Wei, D., Deng, X., Xing, C., et al. (2020). Intestine bacterial community composition of shrimp varies under low-and high-salinity culture conditions. *Front. Microbiol.* 11:2765. doi: 10.3389/fmicb.2020.589164

Huang, F., Pan, L., Song, M., Tian, C., and Gao, S. (2018). Microbiota assemblages of water, sediment, and intestine and their associations with environmental factors and shrimp physiological health. *Appl. Microbiol. Biotech.* 102, 8585–8598. doi: 10.1007/s00253-018-9229-5

Huang, L., Guo, H., Chen, C., Huang, X., Chen, W., Bao, F., et al. (2020). The bacteria from large-sized bioflocs are more associated with the shrimp gut microbiota in culture system. *Aquaculture* 523:735159. doi: 10.1016/j.aquaculture.2020.735159

Jiao, S., and Lu, Y. (2020). Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ. Microbiol.* 22, 1052–1065. doi: 10.1111/1462-2920.14815

Johnson, C. N. (2013). Fitness factors in vibrios: a mini-review. *Microb. Ecol.* 65, 826–851. doi: 10.1007/s00248-012-0168-x

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kerr, C. A., Grice, D. M., Tran, C. D., Bauer, D. C., Li, D., Hendry, P., et al. (2015). Early life events influence whole-of-life metabolic health via gut microflora and gut permeability. *Crit. Rev. Microbiol.* 41, 326–340. doi: 10.3109/1040841X.2013.837863

Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., et al. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* 8, 761–763. doi: 10.1038/nmeth.1650

Legendre, P., Lapointe, F. J., and Casgrain, P. (1994). Modeling brain evolution from behavior: a permutational regression approach. *Evolution* 48, 1487–1499. doi: 10.1111/j.1558-5646.1994.tb02191.x

Li, E., Xu, C., Wang, X., Wang, S., Zhao, Q., Zhang, M., et al. (2018). Gut microbiota and its modulation for healthy farming of Pacific white shrimp *Litopenaeus vannamei*. *Rev. Fish. Sci. Aquac.* 26, 381–399. doi: 10.1080/23308249.2018.1440530

Lichstein, J. W. (2007). Multiple regression on distance matrices: a multivariate spatial analysis tool. *Plant Ecol.* 188, 117–131. doi: 10.1007/s11258-006-9126-3

Liu, L., Chen, H., Liu, M., Yang, J. R., Xiao, P., Wilkinson, D. M., et al. (2019). Response of the eukaryotic plankton community to the cyanobacterial biomass cycle over 6 years in two subtropical reservoirs. *ISME J.* 13, 2196–2208. doi: 10.1038/s41396-019-0417-9

Liu, L., Yang, J., Yu, Z., and Wilkinson, D. M. (2015). The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *ISME J.* 9, 2068–2077. doi: 10.1038/ismej.2015.29

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., et al. (2014). Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.* 24, 813–821. doi: 10.1016/j.cub.2014.02.050

Lu, J., Li, X., Qiu, Q., Chen, J., and Xiong, J. (2022). Gut interkingdom predator-prey interactions are key determinants of shrimp health. *Aquaculture* 546:737304. doi: 10.1016/j.aquaculture.2021.737304

Lucas, R., Courties, C., Herbland, A., Goulletquer, P., Marteau, A. L., and Lemonnier, H. (2010). Eutrophication in a tropical pond: understanding the bacterioplankton and phytoplankton dynamics during a vibriosis outbreak using flow cytometric analyses. *Aquaculture* 310, 112–121.

Luo, Z., Li, S., Hou, K., and Ji, G. (2019). Spatial and seasonal bacterioplankton community dynamics in the main channel of the Middle Route of South-to-North water diversion project. *Res. Microbiol.* 170, 24–34.

Lynch, M. D., and Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat. Rev. Microbiol.* 13, 217–229.

Magurran, A. E., and Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714–716.

Manilal, A., Sujith, S., Selvin, J., Shakir, C., Gandhimathi, R., and Kiran, G. S. (2010). Virulence of vibrios isolated from diseased black tiger shrimp, *Penaeus monodon*, Fabricius. *J. World Aquacult. Soc.* 41, 332–343.

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc. Nat. Acad. Sci. U.S.A.* 110, 3229–3236.

Mo, Y., Zhang, W., Yang, J., Lin, Y., Yu, Z., and Lin, S. (2018). Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. *ISME J.* 12, 2198–2210.

Nyirabuhoro, P., Liu, M., Xiao, P., Liu, L., Yu, Z., Wang, L., et al. (2019). Seasonal variability of conditionally rare taxa in the water column bacterioplankton community of subtropical reservoirs in China. *Microb. Ecol.* 80, 14–26.

Or, A., Shtrasler, L., and Gophna, U. (2012). Fine-scale temporal dynamics of a fragmented lotic microbial ecosystem. *Sci. Rep.* 2:207.

Pester, M., Bittner, N., Deevong, P., Wagner, M., and Loy, A. (2010). A 'rare biosphere'microorganism contributes to sulfate reduction in a peatland. *ISME J.* 4, 1591–1602. doi: 10.1038/ismej.2010.75

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna: The R Foundation for Statistical Computing.

Shade, A., Jones, S., Caporaso, J. G., Handelsman, J., Knight, R., Fierer, N., et al. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio* 5:e01371-14. doi: 10.1128/mBio.01371-14

Song, M., Pan, L., Zhang, M., Huang, F., Gao, S., and Tian, C. (2020). Changes of water, sediment, and intestinal bacterial communities in Penaeus japonicus cultivation and their impacts on shrimp physiological health. *Aquac. Int.* 28, 1847–1865. doi: 10.1007/s10499-020-00562-9

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Stephens, W. Z., Burns, A. R., Stagaman, K., Wong, S., Rawls, J. F., Guillemin, K., et al. (2016). The composition of the zebrafish intestinal microbial community varies across development. *ISME J.* 10, 644–654. doi: 10.1038/ismej.2015.140

Sun, F., Wang, Y., Wang, C., Zhang, L., Tu, K., and Zheng, Z. (2019). Insights into the intestinal microbiota of several aquatic organisms and association with the surrounding environment. *Aquaculture* 507, 196–202. doi: 10.1016/j.aquaculture.2019.04.026

Sylvain, F. -É, Holland, A., Bouslama, S., Audet-Gilbert, É, Lavoie, C., Val, A. L., et al. (2020). Fish skin and gut microbiomes show contrasting signatures of host species and habitat. *Appl. Environ. Microbiol.* 86:e00789-20. doi: 10.1128/AEM.00789-20

Talbot, J. M., Bruns, T. D., Taylor, J. W., Smith, D. P., Branco, S., Glassman, S. I., et al. (2014). Endemism and functional convergence across the North American soil mycobiome. *Proc. Nat. Acad. Sci. U.S.A.* 111:6341.

Van Der Gast, C. J., Ager, D., and Lilley, A. K. (2008). Temporal scaling of bacterial taxa is influenced by both stochastic and deterministic ecological factors. *Environ. Microbiol.* 10, 1411–1418. doi: 10.1111/j.1462-2920.2007.01550.x

Vanwonterghem, I., Jensen, P. D., Dennis, P. G., Hugenholtz, P., Rabaey, K., and Tyson, G. W. (2014). Deterministic processes guide long-term synchronised population dynamics in replicate anaerobic digesters. *ISME J.* 8, 2015–2028.

Venkataraman, A., Bassis, C. M., Beck, J. M., Young, V. B., Curtis, J. L., Huffnagle, G. B., et al. (2015). Application of a neutral community model to assess structuring of the human lung microbiome. *mBio* 6:e02284-14.

Wan, W., Gadd, G. M., Yang, Y., Yuan, W., Gu, J., Ye, L., et al. (2021). Environmental adaptation is stronger for abundant rather than rare microorganisms in wetland soils from the Qinghai-Tibet Plateau. *Mol. Ecol.* 30, 2390–2403.

Webb, C. O., Ackerly, D. D., and Donoghue, M. P. J. (2002). Phylogenies and community ecology. *Ann. Rev. Ecol. Syst.* 33, 475–505.

Williams, T. J., Long, E., Evans, F., DeMaere, M. Z., Lauro, F. M., Raftery, M. J., et al. (2012). A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J.* 6, 1883–1900.

Xiao, F., Zhu, W., Yu, Y., He, Z., Wu, B., Wang, C., et al. (2021). Host development overwhelms environmental dispersal in governing the ecological succession of zebrafish gut microbiota. *NPJ Biofilms Microbiomes* 7:5.

Xiong, J. (2018). Progress in the gut microbiota in exploring shrimp disease pathogenesis and incidence. *Appl. Microbiol. Biotech.* 102, 7343–7350.

Xiong, J., Dai, W., and Li, C. (2016). Advances, challenges, and directions in shrimp disease control: the guidelines from an ecological perspective. *Appl. Microbiol. Biotech*. 100, 6947–6954.

Xiong, J., Dai, W., Qiu, Q., Zhu, J., Yang, W., and Li, C. (2018). Response of host–bacterial colonization in shrimp to developmental stage, environment and disease. *Mol. Ecol.* 27, 3686–3699.

Xiong, J., Xuan, L., Yu, W., Zhu, J., Qiu, Q., and Chen, J. (2019b). Spatiotemporal successions of shrimp gut microbial colonization: high consistency despite distinct species pool. *Environ. Microbiol.* 21, 1383–1394.

Xiong, J., Nie, L., and Chen, J. (2019a). Current understanding on the roles of gut microbiota in fish disease and immunity. *Zool. Res.* 40, 70–76.

Xiong, J., Wang, K., Wu, J., Qiuqian, L., Yang, K., Qian, Y., et al. (2015). Changes in intestinal bacterial communities are closely associated with shrimp disease severity. *Appl. Microbiol. Biotech.* 99, 6911–6919.

Xiong, J., Zhu, J., Dai, W., Dong, C., Qiu, Q., and Li, C. (2017). Integrating gut microbiota immaturity and disease-discriminatory taxa to diagnose the initiation and severity of shrimp disease. *Environ. Microbiol.* 19, 1490–1501.

Xiong, J., Zhu, J., Wang, K., Wang, X., Ye, X., Liu, L., et al. (2014). The temporal scaling of bacterioplankton composition: high turnover and predictability during shrimp cultivation. *Microb. Ecol.* 67, 256–264.

Yan, Q., Li, J., Yu, Y., Wang, J., He, Z., Van Nostrand, J. D., et al. (2016). Environmental filtering decreases with fish development for the assembly of gut microbiota. *Environ. Microbiol.* 18, 4739–4754.

Yang, W., Zheng, C., Zheng, Z., Wei, Y., Lu, K., and Zhu, J. (2018). Nutrient enrichment during shrimp cultivation alters bacterioplankton assemblies and destroys community stability. *Ecotoxicol. Environ. Saf.* 156, 366–374.

Zhang, D., Wang, X., Xiong, J., Zhu, J., Wang, Y., Zhao, Q., et al. (2014). Bacterioplankton assemblages as biological indicators of shrimp health status. *Ecol. Indic.* 38, 218–224.

Zhang, X., Li, X., Lu, J., Qiu, Q., Chen, J., and Xiong, J. (2021). Quantifying the importance of external and internal sources to the gut microbiota in juvenile and adult shrimp. *Aquaculture* 531:735910.

Zhang, Z., Li, D., Refaey, M. M., Xu, W., Tang, R., and Li, L. (2018). Host age affects the development of southern catfish gut bacterial community divergent from that in the food and rearing water. *Front. Microbiol.* 9:495. doi: 10.3389/fmicb.2018.00495

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81:e00002-17.

Zhou, J., Deng, Y., Zhang, P., Xue, K., Liang, Y., Van Nostrand, J. D., et al. (2014). Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc. Nat. Acad. Sci. U.S.A.* 111, 836–845.

Zhou, L., Qu, Y., Qin, J. G., Chen, L., Han, F., and Li, E. (2021). Deep insight into bacterial community characterization and relationship in the pond water, sediment and the gut of shrimp (*Penaeus japonicus*). *Aquaculture* 539:736658. doi: 10.1016/j.aquaculture.2021.736658

Zoqratt, M. Z. H. M., Eng, W. W. H., Thai, B. T., Austin, C. M., and Gan, H. M. (2018). Microbiome analysis of Pacific white shrimp gut and rearing water from Malaysia and Vietnam: implications for aquaculture research and management. *PeerJ* 6:e5826.

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 02 December 2021
doi: 10.3389/fmicb.2021.783334

Biochar Addition Altered Bacterial Community and Improved Photosynthetic Rate of Seagrass: A Mesocosm Study of Seagrass *Thalassia hemprichii*

Jian Zhang[1,2,3,4,5†], Juan Ling[1,2,3,4,6†], Weiguo Zhou[1,2,3,4], Wenqian Zhang[1,2,3,4,5], Fangfang Yang[1,2,3,4], Zhangliang Wei[1,2,3,4], Qingsong Yang[1,2,3,4,6], Ying Zhang[1,2,3,4] and Junde Dong[1,2,3,4,6]*

[1]CAS Key Laboratory of Tropical Marine Bio-Resources and Ecology, Guangdong Provincial Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China, [2]Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China, [3]Key Laboratory of Tropical Marine Biotechnology of Hainan Province, Sanya Institute of Oceanology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Sanya, China, [4]Innovation Academy of South China Sea Ecology and Environmental Engineering, Chinese Academy of Sciences, Guangzhou, China, [5]College of Marine Science, University of Chinese Academy of Sciences, Beijing, China, [6]Sanya National Marine Ecosystem Research Station, Tropical Marine Biological Research Station in Hainan, Chinese Academy of Sciences, Sanya, China

Seagrass meadows, as typical "blue carbon" ecosystems, play critical ecological roles in the marine ecosystem and decline every year. The application of biochar in soil has been proposed as a potential soil amendment to improve soil quality and mitigate global climate change. The effects of biochar on soil bacterial activities are integrally linked to the potential of biochar in achieving these benefits. However, biochar has been rarely applied in marine ecosystems. Whether the application of biochar could work on the seagrass ecosystem remained unknown. In this study, we investigated the responses of sediment and rhizosphere bacterial communities of seagrass *Thalassia hemprichii* to the biochar addition derived from maize at ratios of 5% by dry weight in the soil during a one-month incubation. Results indicated that the biochar addition significantly changed the sedimental environment with increasing pH, total phosphorus, and total kalium while total nitrogen decreased. Biochar addition significantly altered both the rhizosphere and sediment bacterial community compositions. The significant changes in rhizosphere bacterial community composition occurred after 30 days of incubation, while the significant variations in sediment bacterial community composition distinctly delayed than in sediment occurred on the 14th day. Biochar application improved nitrification and denitrification, which may accelerate nitrogen cycling. As a stabilizer to communities, biochar addition decreased the importance of deterministic selection in sediment and changed the bacterial co-occurrence pattern. The biochar addition may promote seagrass photosynthesis and growth by altering the bacterial community compositions and improving nutrient circulation in the seagrass ecosystem, contributing to the seagrass health improvement. This study provided a theoretical basis for applying biochar to the seagrass ecosystem and shed light on the feasible application of biochar in the marine ecosystem.

Keywords: seagrass, biochar addition, nutrient cycling, bacterial community, community assembly

OPEN ACCESS

Edited by:
Daliang Ning,
University of Oklahoma,
United States

Reviewed by:
Kai Feng,
Chinese Academy of Sciences (CAS), China
Tianjiao Dai,
Tsinghua University, China
Xiaoli Zhang,
Chinese Academy of Sciences (CAS), China

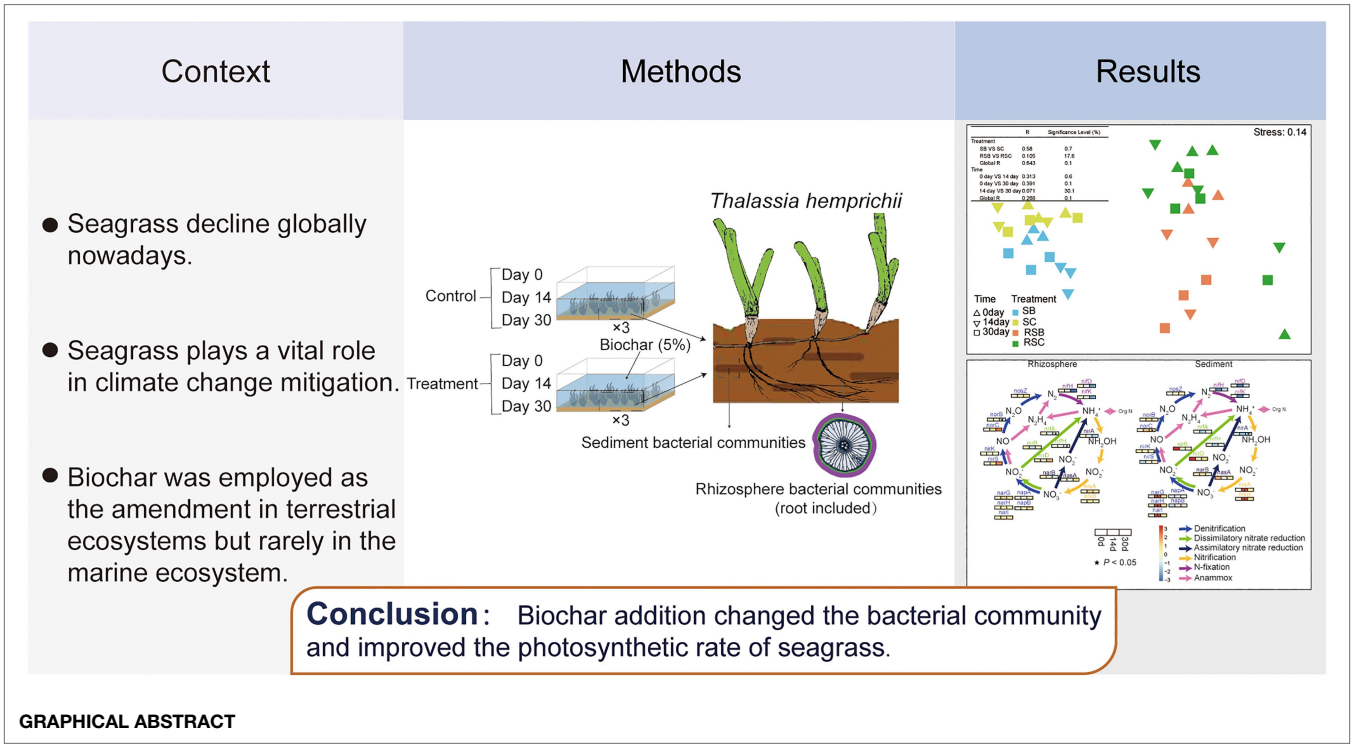
***Correspondence:**
Junde Dong
dongjd@scsio.ac.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *26 September 2021*
Accepted: *08 November 2021*
Published: *02 December 2021*

Citation:
Zhang J, Ling J, Zhou W, Zhang W, Yang F, Wei Z, Yang Q, Zhang Y and Dong J (2021) Biochar Addition Altered Bacterial Community and Improved Photosynthetic Rate of Seagrass: A Mesocosm Study of Seagrass Thalassia hemprichii. Front. Microbiol. 12:783334. doi: 10.3389/fmicb.2021.783334



GRAPHICAL ABSTRACT

INTRODUCTION

Seagrasses are the only marine flowering plants found along temperate and tropical coastlines worldwide, and seagrass meadows are one of the most widespread coastal habitats on earth (Cullen-Unsworth and Unsworth, 2018). Seagrass meadows play critical ecological roles in the marine ecosystem (Cullen-Unsworth and Unsworth, 2018). However, like many of the world's natural habitats, seagrass meadows are in decline, with estimated global losses of ~7% annually since 1990 (Waycott et al., 2009). Poor coastal water quality and coastal development are among the primary drivers of their loss (Waycott et al., 2009). Anthropogenic pollution and global climate change altered the sedimental environment and nutrient cycles of the seagrass ecosystem (Short and Neckles, 1999; Harris et al., 2021). Strategies need to be implemented to relieve the pressure on seagrass.

Biochar is a carbon-rich coproduct resulting from pyrolyzing biomass in oxygen-limited conditions (Lee et al., 2010). It is also a stable carbonaceous material with an extensive surface area and active functional groups (Spokas, 2010). The biochar application in the soil is evaluated globally to improve soil fertility (Novak et al., 2009). Recently, researchers have shown an increased interest in biochar amendments because it could promote additional photosynthetically fixed carbon into the soil, where it may contribute to longer-term carbon storage and thus mitigates increasing atmospheric CO_2 concentrations (Schmidt and Noack, 2000; Lehmann et al., 2006; Lehmann, 2007; Woolf et al., 2010). It is generally accepted that biochar is mainly unavailable to soil microbes, but it can induce changes in soil physicochemical properties and the introduction of metabolically available labile carbon compounds associated with the biochar, which may shift the soil microbial community composition and abundance (Grossman et al., 2010; Anderson et al., 2011). The variations mentioned above may well affect nutrient cycles or soil structure and indirectly affect plant growth (Yu et al., 2021).

Sediment, especially the rhizosphere of plants, is a complex and heterogeneous hotspot inhabited by various microorganisms, including bacteria, fungi, protists, nematodes, and viruses (Dodd et al., 1987; Huang et al., 2014; Wei et al., 2017; Pratama et al., 2020). Plants provide a multitude of niches for microorganisms' growth and proliferation. Lennon and Jones (2011) noted that the physicochemical properties of the soil, together with plant species, dominated where members of microorganisms can grow and thrive. Vice versa, soil microbial communities play central roles in most biogeochemical and ecological processes (Bardgett and van der Putten, 2014). They can form complex co-associations with plants and have essential roles in promoting the productivity and health of the plant in natural environments (Trivedi et al., 2020). Among the plant-associated microbiota, bacteria are the most dominant form. Thus, information on bacteria community composition, diversity, and their determinants is critical for understanding responses of plant microbial symbiont to environmental changes. Moreover, Martin et al. (2020) found that microbial indicators could detect the potential stress in the seagrass ecosystem while other seagrass health metrics failed to detect.

So far, biochar has been widely studied in terrestrial ecosystems (Laird et al., 2010; Song et al., 2020; Owsianiak et al., 2021), and biochar application in soil has been proven to be an effective method for enhancing nutrient cycling, and they could mediate biochar-plant root interactions and ultimately affected root

growth and overall plant performance (El-Naggar et al., 2019; Purakayastha et al., 2019). Zhang et al. (2018) found that the soil microbial activities increased and community structure changed under biochar amendment, which benefited the soil carbon sequestration and farmland systems stability and promoted soil nutrients cycling, thus improving crop yields. Plant stress is one of the major problems encountered in plant growth, and Kavitha et al. (2018) found biochar displayed great potential to mitigate plant stresses for both biotic and abiotic types of stresses. However, biochar application was mainly investigated in terrestrial agriculture and freshwater ecosystems, while for the marine environment, the information was rare. Given the functions of biochar to the terrestrial ecosystems, whether biochar could be applied to the marine seagrass ecosystem and get similar results. If biochar could promote the seagrass ecosystem, such as optimization of nutrient cycles and promotion of seagrass growth, it may be a solution to mitigate seagrass stresses under anthropogenic activities and global climate change.

In the present study, biochar was added as a soil amendment with the intention to improve the health condition of seagrass. The high-throughput DNA gene sequencing has been used to investigate the influence of biochar on the bacterial community, which indirectly explains biochar's effect on seagrass. The purpose of this study was to synthesize responses of seagrass sediment and rhizosphere bacterial community structure shifts and activities to biochar addition comprehensively. We have a hypothesis that if biochar addition could optimize the sediment and rhizosphere bacterial community structure of seagrass in the marine ecosystem and indirectly ameliorate the health status of seagrass? The result of this study could contribute to further biochar application in the seagrass ecosystem.

MATERIALS AND METHODS

Sample Collection, Microcosm Setup, and Experimental Design

Biochar was pyrolyzed from maize straw in this study. The maize straw was firstly pre-crushed, dried at 80°C, passed through a 2-mm sieve, and then pyrolyzed at 600°C for 1 h in the oven. Biochar pH was 9.0 approximately, which was measured with the standard procedure referred to ASTM (2017).

Seagrass *Thalassia hemprichii* were collected at Xincun Bay, Hainan province, China, (18°24′48″ N, 109°59′2″ E) on 13th June 2018. The *in situ* sediment was collected, extracting the surface layer (up to 10 cm deep) simultaneously. The samples were collected, stored in sterile sealing bags, and immediately transported to the laboratory.

The culture experiment was conducted indoors with constant room temperature at Tropical Marine Biological Research Station in Hainan, Chinese Academy of Sciences, from June 13, to July 19, 2018. Six independent microcosms manufactured by rectangular glass aquaria (24 L capacity, 30 cm height × 40 cm length × 20 cm width) were used for the experiment. Each microcosm contained about 10 cm of sediment (about 10 kg wet weight) and 10 L of artificial seawater, configured according to the ambient salinity (28.2 PSU) in the lab. Seagrasses were then transported into the glass vessels, where they were maintained for 1 week of indoor acclimation.

After the acclimation period, three aquaria with non-biochar-added soil were set as the blank control groups, while three aquaria with biochar-added soil with a final concentration of 5% were set as biochar addition groups. Each aquarium had an independent air pump providing proper aeration. The temperature was maintained at 29.0°C with a slight fluctuation (±0.5°C), close to the ambient temperature at the collection site (29.5°C). The lab allowed us to control incident light (270 μmol photons m^{-2} s^{-1}) above the saturation irradiance for these plants (Pérez and Romero, 1992) on a 12-h:12-h light: dark photoperiod. In order to better mimic the environmental conditions and eliminate artificial disturbances, no extra nutrients were added to the samples during the experiments, and the seawater overlying sediment was renewed every week with 0.2 μm membrane filtered seawater.

Samples used for bacterial and physicochemical analysis were collected simultaneously on the 1st, 14th, and 30th days. Sediment samples at each aquarium were collected from unvegetated areas. Rhizosphere sediment samples include the root of seagrass and soil that adheres to roots. After sampling, each sample was thoroughly homogenized using a sterilized spoon. All samples consisted of four types, including the rhizosphere sediment of blank control (RSC), the sediment of blank control (SC), the rhizosphere sediment of experiment group (RSB), and the sediment of experiment group (SB). "Soil" refers to both the sediment and rhizosphere sediment in afterward description.

All samples for DNA analysis were kept in sample protectors (TaKaRa, Dalian, China), frozen immediately, and stored at −80°C until further analysis. The temperature and salinity of the seawater adjacent to seagrass samples (within 3 cm) were measured using a YSI 6600V2 water quality sonde (YSI, Yellow Springs, OH, United States). Dissolved oxygen (DO) concentrations and pH values were measured using a portable pH/DO Meter (Thermo Fisher Scientific, MA, United States). Inorganic nutrients in seawater, including ammonium (oxidized by hypobromite), nitrate (diazotizing with sulfanilamide), nitrite [colored N-(1-naphthyl)-ethylenediamine-dihydrochloride], and phosphate (colored molybdophosphoric blue), were measured using standard methods as described previously with spectrophotometer (Huang et al., 2003). Chemical data [Total nitrogen (TN), total phosphorus (TP), total kalium (TK), available nitrogen, available phosphorus, available kalium, and nitrate-nitrogen (NO3-N)] of sediments were determined by using standard oceanographic methods with ultraviolet spectrophotometry method (General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, 2002). The rapid light curve (RLC) function of the Diving-PAM (Diving-PAM, Walz, Germany) was used to measure *in situ* photosynthetic performance (based on the effective quantum yield of PSII [Y] values) of intact seagrasses that were placed in small incubating chambers, and the rate of electron transport between photosystem II and photosystem I (ETR) was measured and used as a proxy for the photosynthetic rate.

DNA Extraction, PCR, and Sequencing

The bacterial 16S rRNA gene (total bacterial composition) was amplified using universal 16S rRNA gene (V4-V5) primers 515F-Y (5′-GTGYCAGCMGCCGCGGTAA) and 926R (5′-CCGYCAATTYMTTTRAGTTT). PCR cycling was performed in reaction mixtures consisting of 25 µl Ex Taq (2×; TaKaRa, Dalian, China), 1 µl of forward primer (10 µm), 1 µl of reverse primer (10 µm), and 1 µl of DNA in a 50-µl final volume. The PCR amplification program was as follows: initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 54°C for 45 s and extension at 72°C for 45 s, and final elongation at 72°C for 10 min. Libraries were constructed from the purified PCR products of each sample. The DNA was then purified with a Promega Wizard DNA Clean-Up System (Madison, WI, United States). Sequencing was performed on the Illumina MiSeq platform 2 × 250 bp.

Amplicon bioinformatic analysis was accomplished with EasyAmplicon v1.0 (Liu et al., 2021). Paired-end sequence data were merged, quality filtered, and dereplication using VSEARCH v2.15 subcommand –fastq_mergepairs, –fastx_filter and –derep_fulllength, respectively (Rognes et al., 2016). Then, the non-redundancy sequences are denoising into amplicon sequence variants (ASV) with USEARCH v10.0 (Edgar, 2010; *via* unoise3), and then, the singletons and chimeric sequences were removed. Chimera was removed by VSEARCH –uchime_ref against with SILVA database (Quast et al., 2013). Feature tables were created by vsearch –usearch_global. The USEARCH sintax algorithm classified the taxonomy of the features (ASVs) in RDP training set 16 (Cole et al., 2014). Samples were rarefied to 10,392 sequences per sample. The soil microbiome data set has been deposited in the NCBI Sequence Read Archive under accession number PRJNA750881.

PICRUSt2

Functional predictions of the microbial community were conducted using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) and the default analysis parameters (Douglas et al., 2020). PICRUSt2 uses the following tools and algorithms: HHMER, EPA-NG, GAPPA, and castor to align ASVs to reference sequences, place them into a reference tree, and perform hidden-state prediction functions (Eddy, 1998; Louca and Doebeli, 2018; Barbera et al., 2019; Czech et al., 2020; Douglas et al., 2020), respectively. Functional prediction analysis was performed at the gene-level (KEGG orthologs) and the pathway-level (Meta Cyc; Kanehisa and Goto, 2000; Karp et al., 2002). The nearest sequenced taxon index (NSTI) value is calculated to evaluate the prediction accuracy, and lower value means higher accuracy. In this study, NSTI values were 0.15 ± 0.002 (mean ± s.e., $n = 39$). The gene table was compared with KEGG pathways related to nitrogen metabolism (KO00910). As a result, a total of 18 nitrogen cycling genes (KOs) were chosen for subsequent analyses. The details of these genes (KOs) were shown in **Supplementary Table S2**. Furthermore, 30 genes involved in the carbon fixation, phosphorous, and sulfur metabolism were also selected for subsequent analyses.

Statistics Analysis

The phylogenetic diversity index (alpha diversity) and rarefaction curves were calculated based on the rarefied ASV table using the "vegan" R package in R software (version 4.0.4; Oksanen et al., 2015). All heatmap was generated using the "pheatmap" package in the R environment (R Core Team, 2018). The correlation between environmental variables and community composition was calculated using the "ggClusterNet"[1] R package in R software (version 4.0.4) with mantel test. Statistical analysis of metagenomic profiles (STAMP) was conducted to analyze the abundance profile. A two-sided Welch's *t* test carried by STAMP was used to identify distinct taxonomic compositions and metabolic pathways between blank control and experiment group (Parks et al., 2014; Li et al., 2020a).

To compare the β diversity of communities, non-metric multidimensional scaling ordination (nMDS) analyses were conducted based on Bray-Curtis similarity. Furthermore, an analysis of similarity (ANOSIM) was used to statistically test for significant differences in bacteria communities among groups, based on different times and treatments. In this analysis, complete separation is indicted by $R = 1$, whereas $R = 0$ suggests no separation. Both nMDS and ANOSIM were performed in PRIMER 7.0 (Clarke and Gorley, 2015).

Ecological Processes Influencing Bacterial Community Assembly

The null model (NM) was used to quantify the contributions of different ecological processes (stochastic vs. deterministic) to bacterial community structure (Stegen et al., 2013). The NM is pattern-generating model that deliberately exclude a mechanism of interest and allow for randomization tests of ecological and biogeographic data, a framework to quantitatively infer community assembly mechanisms by phylogenetic bin-based null model analysis (iCAMP) was used (Ning et al., 2020). We calculated the framework for bacterial community assembly in soil with the "iCAMP" R package[2], and the results showed the relative importance of different processes in the turnover of each bin within each group of samples.

Co-occurrence Network

A valid co-occurrence correlation was assigned between bacterial community composition if the spearman's correlation coefficient (r) was greater than 0.6 with an adjusted value of $p < 0.01$. Topological characteristics were calculated to describe the complexity of gene co-occurrence networks, including average degree (avgK, which is a key topological property to describe how well a node is connected to the others, higher avgK value means a more complex network), clustering coefficient (CC, which is used to measure the extent of module structure present in a network), characteristic path distance (CPD, which is the average value of the distances between every two nodes in a network, higher CPD value means a reduced coupling among nodes in a network), and network density (ND, which is closely related to the average degree).

[1]https://github.com/taowenmicro/ggClusterNet
[2]https://github.com/DaliangNing/iCAMP1

The topological role of each ASV was determined according to the Zi degree (how well a node is connected to other nodes in the same module) and Pi degree (how well a node is connected to the nodes in other modules; Xun et al., 2017). According to the suggested Zi and Pi degree thresholds (Olesen et al., 2007), all ASVs were categorized into four subcategories: peripherals ($Zi \leq 2.5$ and $0 \leq Pi \leq 0.62$), connectors ($Zi \leq 2.5$ and $Pi > 0.62$), module hubs ($Zi > 2.5$ and $Pi \leq 0.62$), and network hubs ($Zi > 2.5$ and $Pi > 0.62$). Overall, the correlations were calculated using the psych package (version 1.8.12; Revelle, 2017) in R software (version 4.0.4). The networks were visualized, and the topological characteristics were calculated using Gephi software (version 0.9.2; Bastian et al., 2009).

RESULTS

Responses of the Environment to Biochar Addition

The environmental parameters of the water, sediment, and seagrass samples were shown in **Supplementary Figure S1**. The addition of biochar significantly increased sediment pH from 7.94 ± 0.21 (mean ± s.e.) to 8.31 ± 0.08 (mean ± s.e.). Total phosphorus, available phosphate, total kalium, and available kalium of sediment were also significantly increased with biochar addition on day 30 ($p < 0.05$). Total nitrogen of the sediment was significantly decreased with biochar addition on day 30 ($p < 0.05$). There was no significant difference in the final NO_3-N concentration between control and biochar addition groups in sediment (**Supplementary Figure S1**). Biochar addition significantly increased the photosynthetic rate of seagrass from 0.76 ± 0.006 (mean ± s.e.) to 0.78 ± 0.003 (mean ± s.e.) on day 30.

Community Structure and Diversity

Bacterial community profiling of 18 sediments (three replications for biochar addition and three replications for control at three time points) and 21 rhizosphere samples (three replications for biochar addition and four replications for control with three time points) were conducted to investigate the effects of biochar on the structure of bacterial communities. The bacterial community profiling yielded 405,288 high-quality sequences. A total of 7,357 bacterial ASVs were identified across all samples (**Supplementary Table S1**). For α-diversity analyses, the communities were rarified to 10,392 sequences per sample, which captured most of the observed ASV richness (**Supplementary Figure S2**). The sediment bacterial community profiling yielded 6,206 ASVs with 187,056 sequences, while the rhizosphere bacterial community got 7,198 ASVs with 218,232 sequences (**Supplementary Table S1**).

The Shannon index, providing an estimate of alpha diversity in each sample, ranged from 6.72 to 7.45 with a mean of 7.17 ± 0.19 (mean ± s.e.) and did not differ significantly between the control and treatment group ($p > 0.05$; **Supplementary Figure S3**). Other alpha diversity indices (including ACE and Simpson) also did not show significant differences within and between groups (ANOVA, $p > 0.05$).

Seagrass rhizosphere sediment and sediment presented different microbial habitats with specific bacteria (**Supplementary Figure S4**). The differences between bacterial communities (β-diversity) were visualized and quantified using the dendrogram cluster for 12 subgroups from the control and biochar addition groups based on Bray-Curtis similarity (replicates were combined into one subgroup; **Figure 1A**). Bacterial communities of rhizosphere sediment and sediment were clearly separated into two clusters. It revealed that there were obvious differences between rhizosphere sediment and sediment bacterial communities. The result of nMDS also appeared to be two clearly differentiated plates (**Figure 1C**).

The result of nMDS (constrained by treatment) and ANOSIM highlighted the biochar effect on all bacteria communities (**Figure 1C**). Pairwise tests revealed significant differences ($p < 0.05$) between SB and SC, while there were no significant differences between RSB and RSC. Meanwhile, it also revealed that the community composition changed significantly with time.

Relative Abundance of the Different Classification Level

In general, most bacteria were gram-negative (91.49% ± 0.29 and 85.97% ± 1.16% for rhizosphere and sediment bacterial communities, respectively; **Figure 1B**). Phylum Bacteroidetes (6.74% ± 0.96 and 11.03% ± 1.37% for rhizosphere and sediment bacterial communities, respectively) and Proteobacteria (73.54% ± 3.11 and 67.26% ± 1.48% for rhizosphere sediment and sediment bacterial communities, respectively) were the most two abundant phyla for all the samples (**Figure 1B**). Phylum Cyanobacteria was much more abundant in the rhizosphere sediment (6.07% ± 2.22%) than sediment (2.25% ± 1.16%).

At the phylum level, Phylum Deferribacteres and Fusobacteria decreased significantly with biochar addition for rhizosphere sediment bacterial communities on day 30. While for sediment bacterial communities, Phylum Actinobacteria decreased, while Acidobacteria and Aminicenantes increased significantly with biochar addition on day 14 (**Figure 2A**).

Moreover, Class Deferribacteres, Deltaproteobacteria, and Fusobacteriia decreased significantly with biochar addition for rhizosphere sediment bacterial communities on day 30. Compared with the control group, the relative abundance of Class Alphaproteobacteria with biochar addition group was lower, while Class Gammaproteobacteria were higher on day 14 for sediment bacterial communities (**Figure 2B**).

At the ASV level, ASV7606 affiliated with the Phylum Actinobacteria showed significant differences between the control and biochar addition groups for both rhizosphere sediment and sediment bacterial communities. In addition, ASV7641 (Phylum Proteobacteria), ASV6824 (Phylum Fusobacteria), and ASV3977 (Phylum Bacteroidetes) presented significant differences between groups for rhizosphere sediment bacterial communities. In contrast, ASV7384 (Phylum Bacteroidetes) and ASV5617 (Phylum Proteobacteria) exhibited differences for sediment bacterial communities (**Figure 2C**).

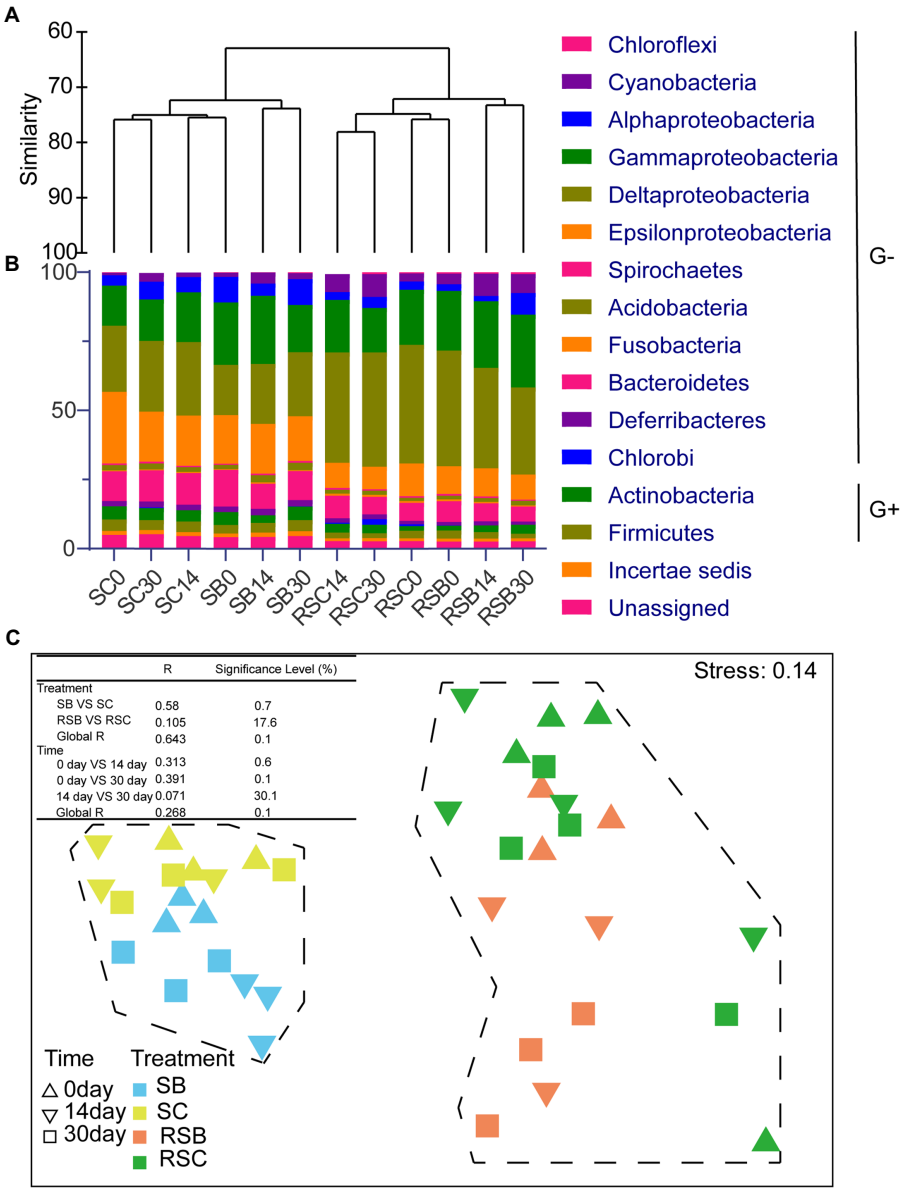


FIGURE 1 | (A) Dendrogram cluster for 12 combined samples based on Bray-Curtis similarity. **(B)** Histogram showing the relative abundance of different subgroups (Phylum and Class level), G+ means Gram-positive bacteria while G– means Gram-negative bacteria. **(C)** Non-metric multidimensional scaling ordinations (nMDS) for bacteria communities of 39 samples. [Table in the figure were analysis of similarities (ANOSIM) of bacterial communities. SB, Sediment bacteria of treatment groups; SC, sediment bacteria of control groups; RSB, rhizosphere bacteria of treatment groups; RSC, rhizosphere bacteria of control groups].

Relative Abundances of Genes Involved in the Nutrient Cycle

The relative abundances of genes involved in carbon fixation, phosphorus and sulfur metabolism, and nitrogen cycle were predicted by PICRUSt2 (**Supplementary Figure S5A**). The abundance of some function genes (e.g., *acsB* for carbon fixation; *ugpQ*, *glpA*, *glpD,* and *glpK* for phosphorus metabolism; *sat_met3* and *cysH* for sulfur metabolism; *nifD*, *nirK,* and *nosZ* for nitrogen cycle) was significantly different between the rhizosphere sediment and the sediment bacterial communities (**Supplementary Figure S5B**).

As illustrated in **Supplementary Figure S6**, both the rhizosphere sediment and sediment bacterial communities showed a significant positive correlation ($p < 0.05$) with NO_3-N. Therefore, in order to better identify the effect of biochar on nitrogen cycling genes, relative changes of nitrogen cycling genes between control and biochar addition groups were calculated in **Figure 3**. Effects of biochar addition on nitrogen cycling genes of the rhizosphere sediment and sediment were different. In sediment, these genes appeared significantly different on day 14 principally. For instance, the nitrogen fixation genes (*nifH*, *nifD,* and *nifK*) were restrained by biochar addition,

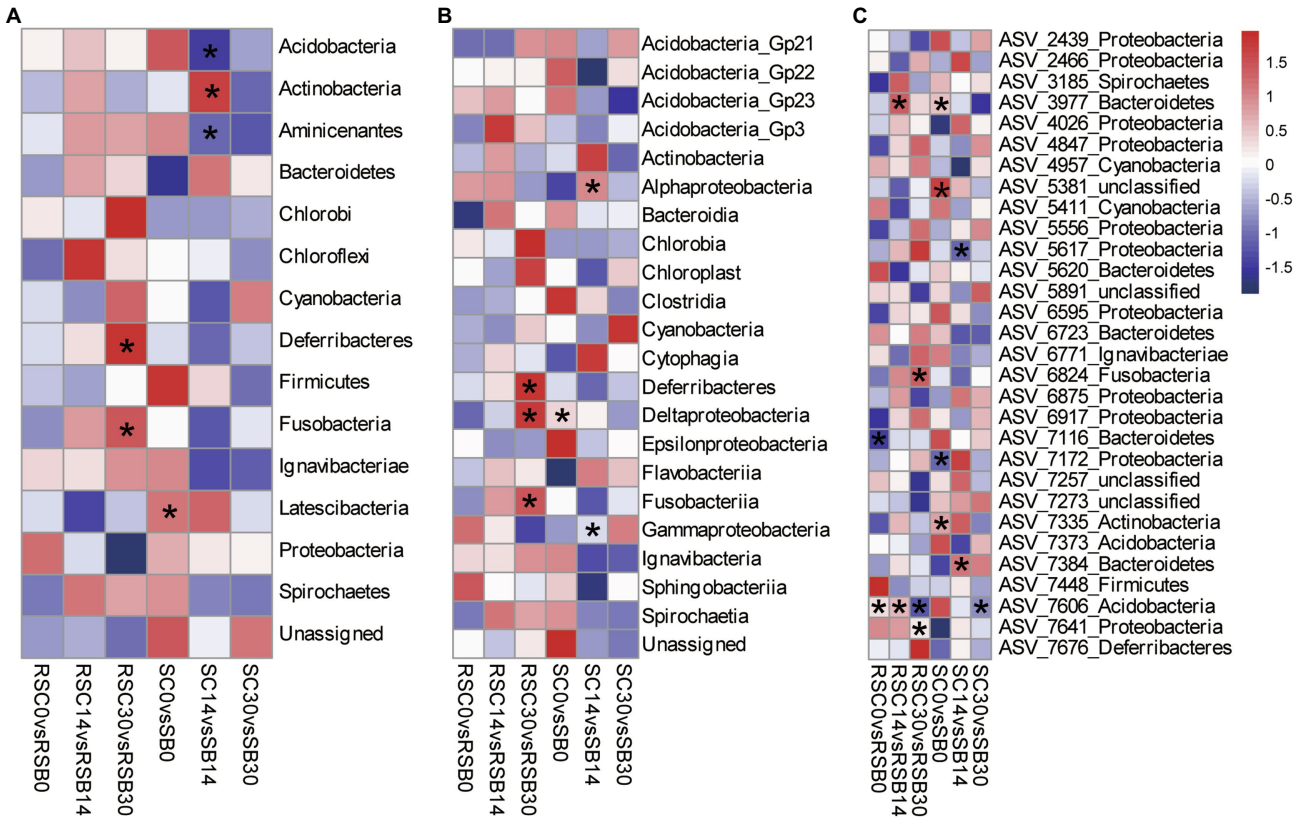


FIGURE 2 | Heatmap of differences between different groups at Phylum **(A)**, Class **(B)** and ASV (**C**: 30 ASVs with the highest relative abundance) level, respectively. The value represented by the color was calculated by Welch's *t* test in STAMP, value >0 (red) means control group has a higher abundance, while value <0 (blue) means biochar addition group has a higher abundance. (*$p < 0.05$).

and denitrification genes (*narG*, *narH,* and *narI*) and nitrification genes (*nrxA* and *nrxB*) were significant increased on day 14. In the rhizosphere sediment, the biochar exerted a primary effect on day 30, and the abundance of nitrogen fixation genes (*nifH*, *nifD*, and *nifK*) was decreased. Besides, biochar addition also significant decreased the abundance of dissimilatory nitrate reduction genes (*nrfA* and *nrfH*).

Ecological Processes Influencing Bacterial Community Assembly

It has been proved that both niche and habitat filtering have effects on bacterial community structure. Null model analysis was used to disentangle the relative importance of stochastic and deterministic processes (homogenous and heterogeneous selections) in the microbial assembly within biochar addition sediment of seagrass (Stegen et al., 2013). The determinism process showed a stronger impact on the community assembly for bacteria than stochasticity in all groups, and heterogeneous selection dominated the deterministic process (**Figure 4A**). Compared to sediment bacterial communities, the determinism process showed a stronger impact on rhizosphere bacterial communities. Biochar addition led to a higher relative importance of stochasticity in sediment on day 30, while there was barely any effect on rhizosphere bacterial communities (**Figure 4B**).

Co-occurrence Patterns of the Bacterial Communities

The bacterial community composition co-occurrence networks were constructed to identify the ecological interplay between co-occurrence taxa (**Figure 5**). Network topological features showed that the co-occurrence pattern in the rhizosphere sediment differed from the sediment network. There was a substantial change in the rhizosphere sediment bacterial community network with biochar addition, while the network topological features showed a minor fluctuation. After filtering most ASVs of low abundance, the final network had 180, 158, 187, and 174 nodes for RSC, RSB, SC, and SB, respectively. Network density, clustering coefficient, and average numbers of degrees were lower, while module numbers were higher for RSB than RSC (**Supplementary Table S3**).

The modular structure of the co-occurrence network was compared between RSC, RSB, SC, and SB. The RSC, RSB, SC, and SB networks parsed into three, six, five, and five major modules (modules with nodes number more than 10% of total nodes number), respectively, which accounted for 87.8, 77.2, 64.2, and 75.9% of their corresponding networks. The modules of RSC and RSB were primarily occupied by Class Deltaproteobacteria and Gammaproteobacteria, while the modules of SC and SB were primarily occupied by Class

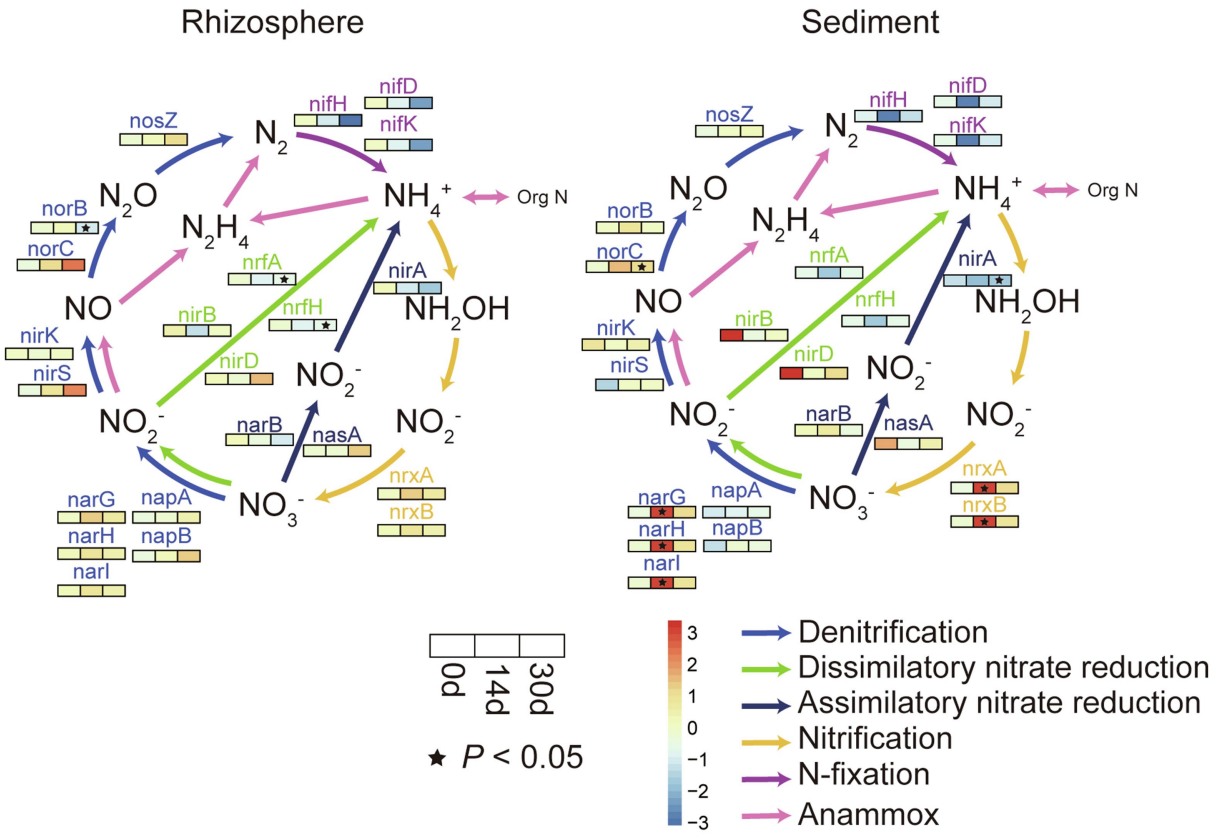


FIGURE 3 | Relative changes of N-cycling genes in rhizosphere and sediment. For each subfigure, the value represented by the color was calculated by Welch's *t* test in STAMP, colors indicate relative differences in gene abundance between the biochar addition groups and control groups, value >0 (red) means biochar addition group has a higher abundance, while value <0 (blue) means control group has a higher abundance. (*$p < 0.05$).

Deltaproteobacteria, Gammaproteobacteria, and Bacteroidetes (**Supplementary Figure S7**).

Topologically, the nodes represent distinct roles in the network. Ecologically, module hubs and connectors signify generalists, network hubs indicate supergeneralists, while peripherals represent specialists. Keystone taxa were identified and displayed *via* Zi-Pi plots (**Supplementary Figure S8**). There were no module hubs and network hubs for all networks (**Supplementary Figure S8**). About 16.7, 8.9, 16.6, and 14.4% of ASVs in the RSC, RSB, SC, and SB were connectors, respectively, and most of them were affiliated with the Phylum Proteobacteria (**Supplementary Table S4**).

DISCUSSION

Biochar Addition Changed the Environment and Improved the Photosynthetic Rate of Seagrass

Biochar addition changed some environment variables significantly in this study. There was a significant pH increase with biochar addition. Over time, the pH of biochar in the sediment may change and either decrease or increase depending on the biochar type. Nguyen and Lehmann (2009) observed a pH decrease with mineral-poor oak wood biochar from pH 4.9 to 4.7, but an increase with mineral-rich corn stover biochar from pH 6.7 to 8.1 during one-year incubation. The driving force behind a pH decrease is the oxidation of carbon to form acidic carboxyl groups (Cheng et al., 2006), whereas the increase in pH is likely related to the dissolution of alkaline minerals. Elevated pH caused by biochar addition might benefit bacteria over fungi (Rousk et al., 2009).

Biochar can affect the microbially-mediated transformation of nutrients significantly in the soil, and it could increase the adsorption of NO_3-N (Van Zwieten et al., 2010) and the soil contents of NO_3-N and TN (Li et al., 2020b). However, there was a minor decrease for NO_3-N while a significant decrease for TN with biochar addition in this study. In agricultural systems, the higher concentrations of available kalium would likely encourage plant uptake of NO_3-N (Chan et al., 2007), which led to a lower concentration of NO_3-N in the sediment of biochar addition groups. Van Zwieten et al. (2010) also found that biochar addition could increase kalium, which was consistent with our results.

Supplementary Figure S1 presents that the phosphorus in the sediment significantly increased in the biochar addition group. Moreover, Lu et al. (2020) found that the biochar addition increased the abundance of genes involved in inorganic phosphate solubilization and organic phosphorus mineralization,

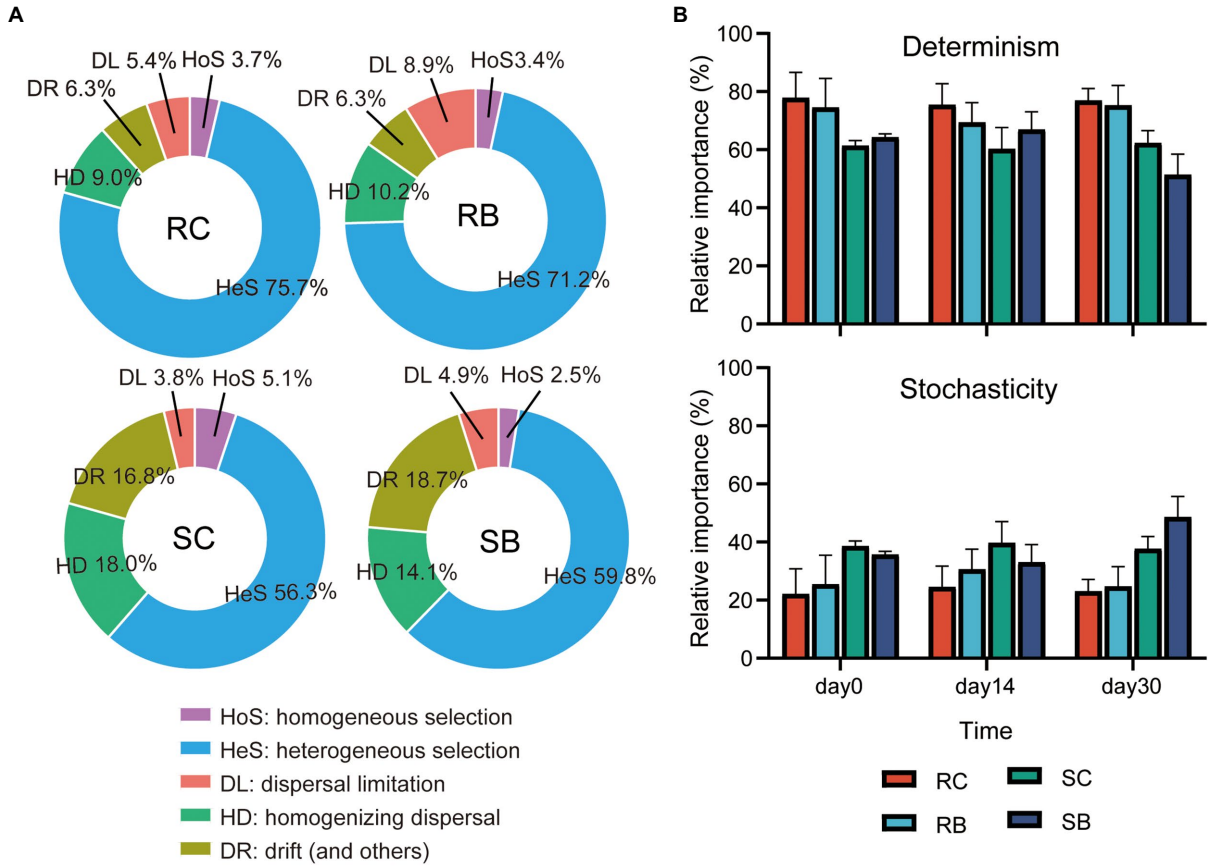


FIGURE 4 | Relative importance of different ecological processes in response to biochar addition. **(A)** Community assembly processes of bacterial community from different groups. **(B)** Changes of determinism and stochasticity; Data are presented as mean values ± SD. Error bars represented standard deviations; For RC, $n = 6$ comparisons among four biologically independent samples at each time point; For SC, RB and SB, $n = 3$ comparisons among three biologically independent samples at each time point. [Determinism: HoS + HeS; stochasticity: DL + HD + DR (Ning et al., 2020)].

but not those involved in phosphorus transportation of the phosphorus cycling. The microbial activities related to organic phosphorus mineralization were enhanced by biochar addition (Masto et al., 2013). Besides, biochar addition at a rate of 20 g kg^{-1} soil increased acid phosphatase activity (+32%) and alkaline phosphatase activity (+22.8%; Masto et al., 2013).

On day 30, there was a significant increase in the biochar addition group's *in situ* seagrass photosynthetic performance, of which a higher value means a strong photosynthetic performance. Higher photosynthetic performance could enhance seagrass growth and optimize CO_2 utilization. This result was consistent with our hypothesis that biochar addition would improve the health of the seagrass.

Biochar Addition Changed the Rhizosphere Sediment Bacterial Community Composition With Delayed Effect, but Not Alpha Diversity

Biochar addition could significantly change the bacterial community compositions (Lehmann et al., 2011; Xu et al., 2014; Zhang et al., 2018; Wei et al., 2020). The specific sets of microbes in the rhizosphere and sediment demonstrated clear separation by compartment in the nMDS plot. This distinctiveness of the plant rhizosphere microbiome was also found in previous studies (Bulgarelli et al., 2012; Lundberg et al., 2012; Peiffer et al., 2013; Yeoh et al., 2015; Zarraonaindia et al., 2015; Hartman et al., 2017). Plants recruit a rhizosphere sediment microbiome in their early life stages from a larger pool of sediment microbes. The initial composition of this sediment microbial pool is the most influential factor determining the composition of rhizosphere sediment microbial communities (Hartman et al., 2018). Therefore, the investigation on the response of both rhizosphere sediment microbiome and sediment microbes to the biochar addition was necessary. In this study, biochar exerted distinct effects on sediment bacterial communities than rhizosphere sediment bacterial communities. The differences between the control and biochar addition groups appeared on day 30 for rhizosphere sediment bacterial communities and day 14 for sediment bacterial communities. Taken together, biochar may have a direct and rapid effect on the sediment bacterial communities, and the change of sediment bacterial

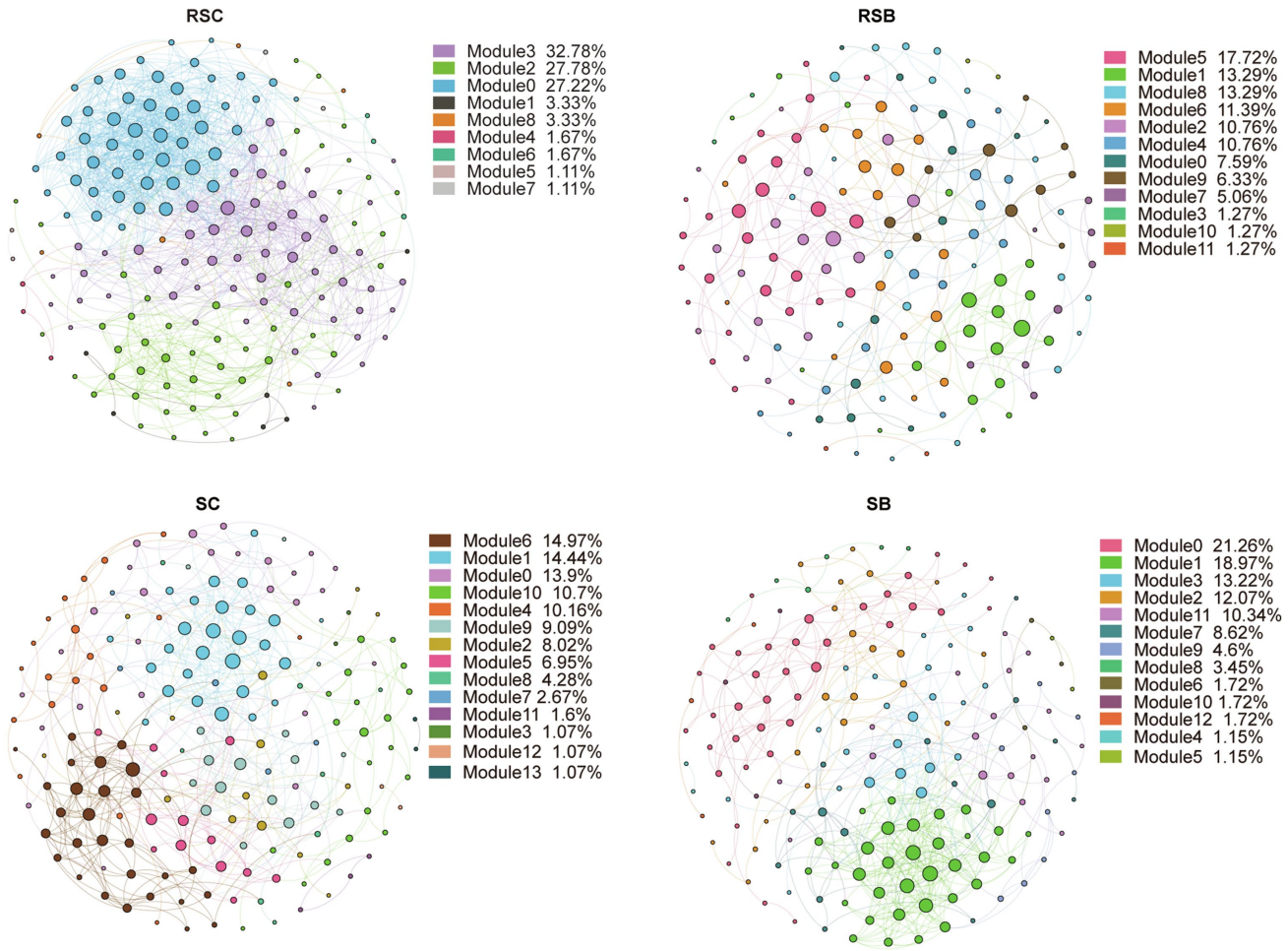


FIGURE 5 | Species-species and species-environment association network. A connection stands for a strong (Spearman's $|r|>0.6$) and significant (value of $p<0.01$) correlation.

pool led to rhizosphere sediment bacterial communities change with a "delayed effect."

Biochar addition could increase the sediment bacterial alpha diversity in constructed wetlands (Deng et al., 2019) and improve microbial activity in PAH-stressed soil (Li et al., 2020b). However, there were no significant differences in alpha diversity between control and biochar addition groups in this study. This may be explained by the different environments and different biochar properties. Effects of biochar application on soil bacterial community structure variations and activities remain controversial under different biochar characteristics, soil properties, and experiment conditions. The role of biochar in soil biological processes, therefore, represents a frontier in soil science research with many unexplained phenomena awaiting exploration.

Three Patterns of Bacterial Groups Induced by Biochar Addition

Phylum Bacteroidetes and Proteobacteria were the two most abundant phyla detected in this study. Their dominance was also observed in previous large-scale surveys of soil microorganisms (Fierer et al., 2007; Lauber et al., 2009). While this study reported the general characteristics of bacterial communities, it also revealed some specific patterns. For example, biochar addition groups with the oligotrophic environment had a higher relative abundance of Acidobacteria in sediment on day 14, which preferred oligotrophic soils (Fierer et al., 2007), and they did not seem to have outcompeted in soils of high CO_2 concentration plots despite with the increased flux of C (Austin et al., 2009).

In the present study, there were three changing patterns of all subgroups based on three sampling time points: Pattern 1: Rhizosphere sediment and sediment bacterial subgroups had consistent variation trend; Pattern 2: rhizosphere sediment and sediment bacteria subgroups had reverse variation trend; and Pattern 3: no correlation between changes of rhizosphere sediment and sediment bacteria subgroups (**Supplementary Figure S9**).

Pattern 1: The biochar addition may create a similar environment for some subgroups, and the relative abundance of these subgroups will increase or decrease depending on the

environment and their competitor in the same niche with a similar trend in rhizosphere sediment and sediment. For example, biochar addition caused the relative abundance of phylum Actinobacteria to decrease on day 14 while increasing on day 30, and that of phylum Cyanobacteria increased on day 14 while decreasing on day 30 in both rhizosphere and sediment.

Pattern 2: The biochar addition may create an uneven environment and have different effects on the sediment and rhizosphere sediment. Subsequently, some bacterial groups moved from the rhizosphere sediment to the sediment, while some bacterial groups moved from the sediment to the rhizosphere sediment driven by environmental factors. For example, biochar addition caused a decrease in the relative abundance of Phylum Fusobacteria and Aminicenantes in the rhizosphere sediment while an increase in sediment on day 30. Fusobacteria is a phylum of obligately anaerobic bacteria commonly found in marine sediment environments (Hofstad et al., 1991), and they have putative hydrocarbon-degrading qualities (Gutierrez et al., 2016). On the other hand, Phylum Aminicenantes exhibited the highest relative abundance in hydrocarbon-impacted environments, followed by marine habitats (especially hydrothermal vents and coral-associated microbiome samples; Farag et al., 2014). Taken together, a possible reason for this might be that biochar addition increased hydrocarbon concentration in the sediment or decreased hydrocarbon concentration in the rhizosphere sediment, which led to some hydrocarbon-related bacteria moving away from the rhizosphere.

Pattern 3: These subgroups may be governed by pattern 1 and pattern 2 with a combined effect, making their variation trends irregular. For example, Phylum Proteobacteria showed a similar variation trend before day 14, while had a reverse trend on day 30. In this study, only three sampling times were set. To further investigate the variation trends of different subgroups, more time points needed to be set.

Biochar Changed Nutrient Cycles, Especially the Nitrogen Cycle

Biochar addition significantly changed some bacterial groups involved in nutrient cycling. For example, the relative abundance of Phylum Deferribacteres decreased significantly with biochar addition in rhizosphere sediment bacterial communities on day 30. The previous study showed that Phylum Deferribacteres might have the *nifH* gene, which is important for nitrogen fixation (Zehr et al., 2003). Moreover, the relative abundance of Phylum Actinobacteria decreased significantly with biochar addition on day 14 for sediment bacterial communities. In the rhizosphere sediment, the enrichment of Actinobacteria could improve bacterial activity and nutrient cycling (Koranda et al., 2011; Zhang et al., 2019). Compared with the control group, the relative abundance of Class Alphaproteobacteria was lower on day 14 for sediment bacterial communities in the treatment group. Bacteria of Class Alphaproteobacteria frequently adopted an intracellular lifestyle as plant mutualists or plant or animal pathogens (Batut et al., 2004). A variety of metabolic strategies are found in this class, including photosynthesis, nitrogen fixation, ammonia oxidation, and methylotrophy (Williams et al., 2007).

The microbiota, mainly bacteria and archaea, drives the soil nitrogen cycle. Many investigations have been carried out on the effects of biochar application on soil microbiota (Kolton et al., 2011; Chen et al., 2012), and biochar addition improved the nitrogen cycle by changing the bacteria community composition (Chan et al., 2007, 2008; Major et al., 2009). Biochar addition restrained nitrogen fixation genes while promoting the transform between NO_3^- and NO_2^- in this study. Furthermore, Xiao et al. (2019) found that biochar addition significantly increased the abundance of ammonia-oxidizing archaea (AOA), *nirK*, *nirS,* and *nosZ* by an average of 25.3, 32.0, 14.6, and 17.0%, respectively. Biochar addition may improve both nitrification and denitrification and accelerate nitrogen cycling. The increased activity of nitrifying microorganisms in biochar may be due to the increase of ammonium nitrogen and DOC contents (Mierzwa-Hersztek et al., 2018). At the same time, DOC drives the turnover of C and N in microorganisms, which stimulates the growth of microorganisms and promotes the activity of denitrifying enzymes (Xiao et al., 2019). Furthermore, nitrification is an acidifying process (Bolan et al., 1991). The alkaline biochar may create much more favorable conditions for nitrifiers and then increase nitrification rates due to its liming effect (Prommer et al., 2014; Ulyett et al., 2014). Nishio (1996) and Rondon et al. (2007) found biochar addition to soil increased biological nitrogen fixation. In addition, Wu et al. (2020) found biochar decreased the diversity of the diazotrophic community and altered diazotroph community structure during composting. Biochar changed the community structure of nitrogen-fixing bacteria, but the effect on *nifH* gene abundance was not clearly determined.

NO_3-N but Not pH as the Key Driver of Both Rhizosphere Sediment and Sediment Bacteria Communities

Recent studies suggested that the richness and diversity of the soil bacterial communities were strongly related to soil pH (Nicol et al., 2008; Lauber et al., 2009; Li et al., 2020b). However, there was no significant correlation between the bacterial community's taxonomical composition and pH. The "size-plasticity" hypothesis argues that smaller individuals are less environment filtered than larger individuals because smaller individuals are more likely to have plasticity in metabolic abilities (Finlay, 2002; Langenheder et al., 2005). Therefore, bacteria may exist widely in such a narrow pH range, suggesting that the selection pressure of pH was invisible on the bacterial community.

In our result, both the rhizosphere sediment and sediment bacterial communities showed a significant positive correlation ($p < 0.05$) with NO_3-N concentration. One possible explanation was that biochar directly affected NO_3-N concentration (Sui et al., 2021); then, NO_3-N acted on bacterial communities. Another explanation was that biochar could act on nitrogen-related bacteria and seagrass (Lehmann et al., 2011; Yu et al., 2021), and the variations in bacteria and seagrass changed the concentration of NO_3-N afterward.

Biochar as a Stabilizer to the Original Environment

In this study, the distinct community assembly pattern of bacterial communities could be mainly explained by deterministic processes rather than stochastic processes, which supported the result found by Logares et al. (2013) that deterministic processes dominated the biogeography of bacterial communities which exposed to progressive long-term environmental change in coastal lakes. Previous studies also showed bacteria were predominantly structured by selection, while microeukaryotes were mainly structured by drift (Logares et al., 2018; Mo et al., 2020).

The biochar addition barely had any effect on the community assembly pattern of seagrass rhizosphere sediment bacterial communities, and it even led to a decreased importance of deterministic processes in sediment bacterial communities. Logares et al. (2018) found three phases of a community that changed after a disturbance. Phase 1: Stochastic processes initially governed by microbial community assembly. Phase 2: Changes in the local environment progressively increased the importance of deterministic selection. Phase 3: The emergence of stable environments led to stable levels of deterministic selection. Selection derived from the variations in the reproductive success across individuals and species caused by the biotic and abiotic pressures; the constant and reduced importance of deterministic selection meant biochar addition might act as a stabilizer to the original environment.

Biochar Addition Changed the Bacterial Co-occurrence Pattern

From the perspective of biotic factors, the relationships between microorganisms exert considerable influence and are also an important aspect of selection pressures. Network structure has important implications for the co-occurrence of species and their stability (Bascompte et al., 2003). The RSC network structure was more complex than RSB with more nodes and edges, while there were slight differences between SC and SB (**Figure 5, Supplementary Table S3**). In general, a more complex network structure may indicate more stable co-existence patterns, and a stable co-occurrence pattern mirrored fewer dynamic characteristics to some extent (Costa et al., 2006). Thébault and Fontaine (2010) demonstrated that high connectivity promoted community stability in mutualistic networks. SB had a lower network density and clustering coefficient with higher module numbers, which meant that the community was separated into more independent groups.

There were several keystones in our network, and all of them were connectors. The loss of these species may lead to the breakdown of the ecological networks and modules (Guimerà and Amaral, 2005). Therefore, these potential key species might be crucial in maintaining the stability of the bacterial communities. The identified connector taxa in the RSC/RSB and SC/SB were quite different, but they were primarily from Phylum Proteobacteria. Hence, they may have similar ecosystem functions.

However, these modules from all networks did not necessarily reflect their taxonomic classification. Most bacterial interactions were stronger between phyla/classes than within phylum/class, which provided evidence that the bacterial community structure is shaped by environmentally driven functional characteristics rather than phylogeny (Burke et al., 2011).

CONCLUSION

Our study investigated for the first time the influence of biochar addition on the bacterial relative abundance, composition, assembly, and co-occurrence network of bacterial communities in the seagrass ecosystem. Rhizosphere sediment and sediment bacterial communities responded differently to the biochar addition. The significant bacterial community composition changes in rhizosphere sediment occurred after incubation for 30 days with a delay effect than that of in sediment (14 days). Alteration of environmental factors and biotic interactions induced by biochar addition enhanced nitrification and denitrification, which may accelerate nitrogen cycling. More nitrogen absorption and photosynthetic performance of seagrass after biochar addition may lower the total nitrogen in sediment. Together, biochar addition could improve seagrass health, which has important implications for biochar application in the seagrass ecosystem.

DATA AVAILABILITY STATEMENT

The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found online at: https://www.ncbi.nlm.nih.gov/, PRJNA750881.

AUTHOR CONTRIBUTIONS

JL and JD: conceptualization, methodology, resources, and supervision. WZho: investigation. JZ and JL: formal analysis and data curation. JZ: writing – original draft and visualization. All authors: writing – review and editing.

FUNDING

The research was supported by the National Natural Science Foundation of China (41676163), Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou; GML2019ZD0402), National Key Research and Development Program of China (2018YFC1406505, 2018FY100105, and 2017YFC0506301), Innovation Academy of South China Sea Ecology and Environmental Engineering, Chinese Academy of Sciences (ISEE2021ZD03), and Science and Technology Planning Project of Guangdong Province, China (2020B1212060058).

ACKNOWLEDGMENTS

We thank all the members of Sanya National Marine Ecosystem Research Station in Hainan for their work in the field sample collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.783334/full#supplementary-material

REFERENCES

Anderson, C. R., Condron, L. M., Clough, T. J., Fiers, M., Stewart, A., Hill, R. A., et al. (2011). Biochar induced soil microbial community change: implications for biogeochemical cycling of carbon, nitrogen and phosphorus. *Pedobiologia* 54, 309–320. doi: 10.1016/j.pedobi.2011.07.005

ASTM (2017). *D3838-05. Standard Test Method for pH of Activated Carbon*. West Conshohocken: ASTM International.

Austin, E. E., Castro, H. F., Sides, K. E., Schadt, C. W., and Classen, A. T. (2009). Assessment of 10 years of CO_2 fumigation on soil microbial communities and function in a sweetgum plantation. *Soil Biol. Biochem.* 41, 514–520. doi: 10.1016/j.soilbio.2008.12.010

Barbera, P., Kozlov, A. M., Czech, L., Morel, B., Darriba, D., Flouri, T., et al. (2019). EPA-ng: massively parallel evolutionary placement of genetic sequences. *Syst. Biol.* 68, 365–369. doi: 10.1093/sysbio/syy054

Bardgett, R. D., and van der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature* 515, 505–511. doi: 10.1038/nature13855

Bascompte, J., Jordano, P., Melián, C. J., and Olesen, J. M. (2003). The nested assembly of plant-animal mutualistic networks. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9383–9387. doi: 10.1073/pnas.1633576100

Bastian, M., Heymann, S., and Jacomy, M. (2009). "Gephi: an open source software for exploring and manipulating networks." in *Proceedings of the 3rd International ICWSM Conference*; May 17–20, 2009; 361–362.

Batut, J., Andersson, S. G., and O'Callaghan, D. (2004). The evolution of chronic infection strategies in the alpha-proteobacteria. *Nat. Rev. Microbiol.* 2, 933–945. doi: 10.1038/nrmicro1044

Bolan, N. S., Hedley, M. J., and White, R. E. (1991). Processes of soil acidification during nitrogen cycling with emphasis on legume based pastures. *Plant Soil* 134, 53–63. doi: 10.1007/BF00010717

Bulgarelli, D., Rott, M., Schlaeppi, K., Loren, V., van Themaat, E., Ahmadinejad, N., et al. (2012). Revealing structure and assembly cues for Arabidopsis root inhabiting bacterial microbiota. *Nature* 488, 91–95. doi: 10.1038/nature11336

Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. U. S. A.* 108, 14288–14293. doi: 10.1073/pnas.1101591108

Chan, K. Y., Van Zwieten, L., Meszaros, I., Downie, A., and Joseph, S. (2007). Agronomic values of greenwaste biochar as a soil amendment. *Aust. J. Soil Res.* 45, 629–634. doi: 10.1071/SR07109

Chan, K. Y., Van Zwieten, L., Meszaros, I., Downie, A., and Joseph, S. (2008). Using poultry litter biochars as soil amendments. *Soil Res.* 46, 437–444. doi: 10.1071/SR08036

Chen, J., Qu, J., Liu, X., Zheng, J., Li, L., and Pan, G. (2012). "Inconsistent changes in microbial community structure and abundance with biochar amendment in rice paddy soils from South China." in *EGU General Assembly Conference 2012*; April 22–27, 2012; Vienna, Austria, 656.

Cheng, C. H., Lehmann, J., Thies, J. E., Burton, S. D., and Engelhard, M. H. (2006). Oxidation of black carbonby biotic and abiotic processes. *Org. Geochem.* 37, 1477–1488. doi: 10.1016/j.orggeochem.2006.06.022

Clarke, K. R., and Gorley, R. N. (2015). *Getting Started With PRIMER v7* (Plymouth: Plymouth Marine Laboratory, PRIMER-E), 20.

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244

Costa, R., Gotz, M., Mrotzek, N., Lottmann, J., Berg, G., and Smalla, K. (2006). Effects of site and plant species on rhizosphere community structure as revealed by molecular analysis of microbial guilds. *FEMS Microbiol. Ecol.* 56, 236–249. doi: 10.1111/j.1574-6941.2005.00026.x

Cullen-Unsworth, L. C., and Unsworth, R. (2018). A call for seagrass protection. *Science* 361, 446–448. doi: 10.1126/science.aat7318

Czech, L., Barbera, P., and Stamatakis, A. (2020). Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* 36, 3263–3265. doi: 10.1093/bioinformatics/btaa070

Deng, C., Huang, L., Liang, Y., Xiang, H., Jiang, J., Wang, Q., et al. (2019). Response of microbes to biochar strengthen nitrogen removal in subsurface flow constructed wetlands: microbial community structure and metabolite characteristics. *Sci. Total Environ.* 694:133687. doi: 10.1016/j.scitotenv.2019.133687

Dodd, J. C., Burton, C. C., Burns, R. G., and Jeffries, P. (1987). Phosphatase activity associated with the roots and the rhizosphere of plants infected with vesicular-arbuscular mycorrhizal fungi. *New Phytol.* 107, 163–172. doi: 10.1111/j.1469-8137.1987.tb04890.x

Douglas, G. M., Maffei, V. J., Zaneveld, J. R., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2020). PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* 38, 685–688. doi: 10.1038/s41587-020-0548-6

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

El-Naggar, A., El-Naggar, A. H., Shaheen, S. M., Sarkar, B., Chang, S. X., Tsang, D. C. W., et al. (2019). Biochar composition-dependent impacts on soil nutrient release, carbon mineralization, and potential environmental risk: a review. *J. Environ. Manage.* 241, 458–467. doi: 10.1016/j.jenvman.2019.02.044

Farag, I. F., Davis, J. P., Youssef, N. H., and Elshahed, M. S. (2014). Global patterns of abundance, diversity and community structure of the Aminicenantes (candidate phylum OP8). *PLoS One* 9:e92139. doi: 10.1371/journal.pone.0092139

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology* 88, 1354–1364. doi: 10.1890/05-1839

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *Science* 296, 1061–1063. doi: 10.1126/science.1070710

Grossman, J. M., O'Neill, B. E., Tsai, S. M., Liang, B., Neves, E., Lehmann, J., et al. (2010). Amazonian anthrosols support similar microbial communities that differ distinctly from those extant in adjacent, unmodified soils of the same mineralogy. *Microb. Ecol.* 60, 192–205. doi: 10.1007/s00248-010-9689-3

Guimerà, R., and Amaral, L. A. N. (2005). Functional cartography of complex metabolic networks. *Nature* 433, 895–900. doi: 10.1038/nature03288

Gutierrez, T., Berry, D., Teske, A., and Aitken, M. D. (2016). Enrichment of Fusobacteria in sea surface oil slicks from the deepwater horizon oil spill. *Microorganisms* 4:24. doi: 10.3390/microorganisms4030024

Harris, P. T., Westerveld, L., Nyberg, B., Maes, T., Macmillan-Lawler, M., and Appelquist, L. R. (2021). Exposure of coastal environments to river-sourced plastic pollution. *Sci. Total Environ.* 769:145222. doi: 10.1016/j.scitotenv.2021.145222

Hartman, K., van der Heijden, M. G., Roussely-Provent, V., Walser, J.-C., and Schlaeppi, K. (2017). Deciphering composition and function of the root microbiome of a legume plant. *Microbiome* 5:2. doi: 10.1186/s40168-016-0220-z

Hartman, K., van der Heijden, M. G. A., Wittwer, R. A., Banerjee, S., Walser, J. C., and Schlaeppi, K. (2018). Cropping practices manipulate abundance patterns of root and soil microbiome members paving the way to smart farming. *Microbiome* 6:14. doi: 10.1186/s40168-017-0389-9

Hofstad, T., Balows, A., Trüper, H. G., Dworkin, M., Harder, W., and Schleifer, K.-H. (1991). "The genus Fusobacterium," in *The Prokaryotes. 2nd Edn.* eds. A. Balows, H. G. Trüper, M. Dworkin, W. Harder and K.-H. Schleifer (New York, NY, USA: Springer-Verlag), 4114–4126.

Huang, X. F., Chaparro, J. M., Reardon, K. F., Zhang, R., Shen, Q., and Vivanco, J. M. (2014). Rhizosphere interactions: root exudates, microbes, and microbial communities. *Botany* 92, 267–275. doi: 10.1139/cjb-2013-0225

Huang, L., Tan, Y., Song, X., Huang, X., Wang, H., Zhang, S., et al. (2003). The status of the ecological environment and a proposed protection strategy

in Sanya Bay, Hainan Island, China. *Mar. Pollut. Bull.* 47, 180–186. doi: 10.1016/S0025-326X(03)00070-5

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Karp, P. D., Riley, M., Paley, S. M., and Pellegrini-Toole, A. (2002). The metacyc database. *Nucleic Acids Res.* 30, 59–61. doi: 10.1093/nar/30.1.59

Kavitha, B., Reddy, P. V. L., Kim, B., Lee, S. S., Pandey, S. K., and Kim, K.-H. (2018). Benefits and limitations of biochar amendment in agricultural soils: a review. *J. Environ. Manage.* 227, 146–154. doi: 10.1016/j.jenvman.2018.08.082

Kolton, M., Meller Harel, Y., Pasternak, Z., Graber, E. R., Elad, Y., and Cytryn, E. (2011). Impact of biochar application to soil on the root-associated bacterial community structure of fully developed greenhouse pepper plants. *Appl. Environ. Microbiol.* 77, 4924–4930. doi: 10.1128/AEM.00148-11

Koranda, M., Schnecker, J., Kaiser, C., Fuchslueger, L., Kitzler, B., Stange, C. F., et al. (2011). Microbial processes and community composition in the rhizosphere of European beech – the influence of plant C exudates. *Soil Biol. Biochem.* 43, 551–558. doi: 10.1016/j.soilbio.2010.11.022

Laird, D. A., Fleming, P., Davis, D. D., Horton, R., Wang, B., and Karlen, D. L. (2010). Impact of biochar amendments on the quality of a typical Midwestern agricultural soil. *Geoderma* 158, 443–449. doi: 10.1016/j.geoderma.2010.05.013

Langenheder, S., Lindström, E. S., and Tranvik, L. J. (2005). Weak coupling between community composition and functioning of aquatic bacteria. *Limnol. Oceanogr.* 50, 957–967. doi: 10.4319/lo.2005.50.3.0957

Lauber, C. L., Hamady, M., Knight, R., and Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.* 75, 5111–5120. doi: 10.1128/AEM.00335-09

Lee, J. W., Hawkins, B., Day, D. M., and Reicosky, D. C. (2010). Sustainability: the capacity of smokeless biomass pyrolysis for energy production, global carbon capture and sequestration. *Energy Environ. Sci.* 3, 1695–1705. doi: 10.1039/c004561f

Lehmann, J. (2007). A handful of carbon. *Nature* 447, 143–144. doi: 10.1038/447143a

Lehmann, J., Gaunt, J., and Rondon, M. (2006). Bio-char sequestration in terrestrial ecosystems–a review. *Mitig. Adapt. Strateg. Glob. Chang.* 11, 403–427. doi: 10.1007/s11027-005-9006-5

Lehmann, J., Rillig, M. C., Thies, J., Masiello, C. A., Hockaday, W. C., and Crowley, D. (2011). Biochar effects on soil biota–a review. *Soil Biol. Biochem.* 43, 1812–1836. doi: 10.1016/j.soilbio.2011.04.022

Lennon, J. T., and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* 9, 119–130. doi: 10.1038/nrmicro2504

Li, Y., Xu, C., Zhang, W., Lin, L., Wang, L., Niu, L., et al. (2020a). Response of bacterial community in composition and function to the various DOM at river confluences in the urban area. *Water Res.* 169:115293. doi: 10.1016/j.watres.2019.115293

Li, X., Yao, S., Bian, Y., Jiang, X., and Song, Y. (2020b). The combination of biochar and plant roots improves soil bacterial adaptation to PAH stress: insights from soil enzymes, microbiome, and metabolome. *J. Hazard. Mater.* 400:123227. doi: 10.1016/j.jhazmat.2020.123227

Liu, Y. X., Qin, Y., Chen, T., Lu, M., Qian, X., Guo, X., et al. (2021). A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* 12, 315–330. doi: 10.1007/s13238-020-00724-8

Logares, R., Lindström, E. S., Langenheder, S., Logue, J. B., Paterson, H., Laybourn-Parry, J., et al. (2013). Biogeography of bacterial communities exposed to progressive long-term environmental change. *ISME J.* 7, 937–948. doi: 10.1038/ismej.2012.168

Logares, R., Tesson, S. V., Canbäck, B., Pontarp, M., Hedlund, K., and Rengefors, K. (2018). Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environ. Microbiol.* 20, 2231–2240. doi: 10.1111/1462-2920.14265

Louca, S., and Doebeli, M. (2018). Efficient comparative phylogenetics on large trees. *Bioinformatics* 34, 1053–1055. doi: 10.1093/bioinformatics/btx701

Lu, H., Yan, M., Wong, M. H., Mo, W. Y., Wang, Y., Chen, X. W., et al. (2020). Effects of biochar on soil microbial community and functional genes of a landfill cover three years after ecological restoration. *Sci. Total Environ.* 717:137133. doi: 10.1016/j.scitotenv.2020.137133

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., et al. (2012). Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488, 86–90. doi: 10.1038/nature11237

Major, J., Steiner, C., Downie, A., Lehmann, J., and Joseph, S. (2009). "Biochar effects on nutrient leaching," in *Biochar for Environmental Management: Science and Technology*. eds. J. Lehmann and S. Joseph (London, UK: Routledge), 271.

Martin, B. C., Alarcon, M. S., Gleeson, D., Middleton, J. A., Fraser, M. W., Ryan, M. H., et al. (2020). Root microbiomes as indicators of seagrass health. *FEMS Microbiol. Ecol.* 96:fiz201. doi: 10.1093/femsec/fiz201

Masto, R. E., Kumar, S., Rout, T. K., Sarkar, P., George, J., and Ram, L. C. (2013). Biochar from water hyacinth (*Eichornia crassipes*) and its impact on soil biological activity. *Catena* 111, 64–71. doi: 10.1016/j.catena.2013.06.025

Mierzwa-Hersztek, M., Klimkowicz-Pawlas, A., and Gondek, K. (2018). Influence of poultry litter and poultry litter biochar on soil microbial respiration and nitrifying bacteria activity. *Waste Biomass Valorization* 9, 379–389. doi: 10.1007/s12649-017-0013-z

Mo, Y., Zhang, W., Wilkinson, D. M., Yu, Z., Xiao, P., and Yang, J. (2020). Biogeography and co-occurrence patterns of bacterial generalists and specialists in three subtropical marine bays. *Limnol. Oceanogr.* 66, 793–806. doi: 10.1002/lno.11643

Nguyen, B., and Lehmann, J. (2009). Black carbon decomposition under varying water regimes. *Org. Geochem.* 40, 846–853. doi: 10.1016/j.orggeochem.2009.05.004

Nicol, G. W., Leininger, S., Schleper, C., and Prosser, J. I. (2008). The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol.* 10, 2966–2978. doi: 10.1111/j.1462-2920.2008.01701.x

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., et al. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* 11:4717. doi: 10.1038/s41467-020-18560-z

Nishio, M. (1996). *Microbial Fertilizers in Japan. Extension Bulletin* (Taipei City: Food & Fertilizer Technology Center, FFTC), 1–13.

Novak, J. M., Busscher, W. J., Laird, D. L., Ahmedna, M., Watts, D. W., and Niandou, M. A. S. (2009). Impact of biochar amendment on fertility of a southeastern coastal plain soil. *Soil Sci.* 174, 105–112. doi: 10.1097/SS.0b013e3181981d9a

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., et al. (2015). Vegan community ecology package: ordination methods, diversity analysis and other functions for community and vegetation ecologists. R package version 2.3.

Olesen, J. M., Bascompte, J., Dupont, Y. L., and Jordano, P. (2007). The modularity of pollination networks. *Proc. Natl. Acad. Sci. U. S. A.* 104, 19891–19896. doi: 10.1073/pnas.0706375104

Owsianiak, M., Lindhjem, H., Cornelissen, G., Hale, S. E., Sørmo, E., and Sparrevik, M. (2021). Environmental and economic impacts of biochar production and agricultural use in six developing and middle-income countries. *Sci. Total Environ.* 755:142455. doi: 10.1016/j.scitotenv.2020.142455

Parks, D. H., Tyson, G. W., Hugenholtz, P., and Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30, 3123–3124. doi: 10.1093/bioinformatics/btu494

Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., et al. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc. Natl. Acad. Sci. U. S. A.* 110, 6548–6553. doi: 10.1073/pnas.1302837110

Pérez, M., and Romero, J. (1992). Photosynthetic response to light and temperature of the seagrass *Cymodocea nodosa* and the prediction of its seasonality. *Aquat. Bot.* 43, 51–62. doi: 10.1016/0304-3770(92)90013-9

Pratama, A. A., Terpstra, J., de Oliveria, A. L. M., and Salles, J. F. (2020). The role of rhizosphere bacteriophages in plant health. *Trends Microbiol.* 28, 709–718. doi: 10.1016/j.tim.2020.04.005

Prommer, J., Wanek, W., Hofhans, F., Trojan, D., Offre, P., Urich, T., et al. (2014). Biochar decelerates soil organic nitrogen cycling but stimulates soil nitrification in a temperate arable field trial. *PLoS One* 9:e86388. doi: 10.1371/journal.pone.0086388

Purakayastha, T. J., Bera, T., Bhaduri, D., Sarkar, B., Mandal, S., Wade, P., et al. (2019). A review on biochar modulated soil condition improvements and nutrient dynamics concerning crop yields: pathways to climate change

mitigation and global food security. *Chemosphere* 227, 345–365. doi: 10.1016/j.chemosphere.2019.03.170

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R Core Team (2018). R: a language and environment for statistical computing. Available at: https://www.r-project.org/about.html (Accessed March 15, 2021).

Revelle, W. R. (2017). psych: procedures for personality and psychological research. Version 1.8.12. Evanston: Northwestern University. Available at: https://CRANR-project.org/package=psych (Accessed March 17, 2021).

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Rondon, M., Lehmann, J., Ramírez, J., and Hurtado, M. (2007). Biological nitrogen fixation by common beans (*Phaseolus vulgaris* L.) increases with bio-char additions. *Biol. Fertil. Soils* 43, 699–708. doi: 10.1007/s00374-006-0152-z

Rousk, J., Brookes, P. C., and Bååth, E. (2009). Contrasting soil pH effects on fungal and bacterial growth suggest functional redundancy in carbon mineralization. *Appl. Environ. Microbiol.* 75, 1589–1596. doi: 10.1128/AEM.02775-08

Schmidt, M. W. I., and Noack, A. G. (2000). Black carbon in soils and sediments: analysis, distribution, implications, and current challenges. *Glob. Biogeochem. Cycles* 14, 777–793. doi: 10.1029/1999GB001208

Short, F. T., and Neckles, H. A. (1999). The effects of global climate change on seagrasses. *Aquat. Bot.* 63, 169–196. doi: 10.1016/S0304-3770(98)00117-X

Song, X., Razavi, B. S., Ludwig, B., Zamanian, K., Zang, H., Kuzyakov, Y., et al. (2020). Combined biochar and nitrogen application stimulates enzyme activity and root plasticity. *Sci. Total Environ.* 735:139393. doi: 10.1016/j.scitotenv.2020.139393

Spokas, K. A. (2010). Review of the stability of biochar in soils: predictability of O:C molar ratios. *Carbon Manage.* 1, 289–303. doi: 10.4155/cmt.10.32

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Sui, M., Li, Y., Jiang, Y., Wang, L., Zhang, W., Sathishkumar, K., et al. (2021). Sediment-based biochar facilitates highly efficient nitrate removal: physicochemical properties, biological responses and potential mechanism. *Chem. Eng. J.* 405:126645. doi: 10.1016/j.cej.2020.126645

Thébault, E., and Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* 329, 853–856. doi: 10.1126/science.1188321

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., and Singh, B. K. (2020). Plant–microbiome interactions: from community assembly to plant health. *Nat. Rev. Microbiol.* 18, 607–621. doi: 10.1038/s41579-020-0412-1

Ulyett, J., Sakrabani, R., Kibblewhite, M., and Hann, M. (2014). Impact of biochar addition on water retention, nitrification and carbon dioxide evolution from two sandy loam soils. *Eur. J. Soil Sci.* 65, 96–104. doi: 10.1111/ejss.12081

Van Zwieten, L., Kimber, S., Morris, S., Downie, A., Berger, E., Rust, J., et al. (2010). Influence of biochars on flux of N_2O and CO_2 from Ferrosol. *Aust. J. Soil Res.* 48, 555–568. doi: 10.1071/SR10004

Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., et al. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proc. Natl. Acad. Sci. U. S. A.* 106, 12377–12381. doi: 10.1073/pnas.0905620106

Wei, Z. W., Hu, X. L., Li, X. H., Zhang, Y. Z., Jiang, L. C., Jing, L., et al. (2017). The rhizospheric microbial community structure and diversity of deciduous and evergreen forests in Taihu lake area, China. *PLoS One* 12:e0174411. doi: 10.1371/journal.pone.0174411

Wei, Z., Wang, J. J., Fultz, L. M., White, P., and Jeong, C. (2020). Application of biochar in estrogen hormone-contaminated and manure-affected soils: impact on soil respiration, microbial community and enzyme activity. *Chemosphere* 270:128625. doi: 10.1016/j.chemosphere.2020.128625

Williams, K. P., Sobral, B. W., and Dickerman, A. W. (2007). A robust species tree for the alphaproteobacteria. *J. Bacteriol.* 189, 4578–4586. doi: 10.1128/JB.00269-07

Woolf, D., Amonette, J., Street-Perrott, F., Lehmann, J., and Joseph, S. (2010). Sustainable biochar to mitigate global climate change. *Nat. Commun.* 1:56. doi: 10.1038/ncomms1053

Wu, X., Sun, Y., Deng, L., Meng, Q., Jiang, X., Bello, A., et al. (2020). Insight to key diazotrophic community during composting of dairy manure with biochar and its role in nitrogen transformation. *Waste Manag.* 105, 190–197. doi: 10.1016/j.wasman.2020.02.010

Xiao, Z., Rasmann, S., Yue, L., Lian, F., Zou, H., and Wang, Z. (2019). The effect of biochar amendment on N-cycling genes in soils: a meta-analysis. *Sci. Total Environ.* 696:133984. doi: 10.1016/j.scitotenv.2019.133984

Xu, H. J., Wang, X. H., Li, H., Yao, H. Y., Su, J. Q., and Zhu, Y. G. (2014). Biochar impacts soil microbial community composition and nitrogen cycling in an acidic soil planted with rape. *Environ. Sci. Technol.* 48, 9391–9399. doi: 10.1021/es5021058

Xun, W., Huang, T., Li, W., Ren, Y., Xiong, W., Ran, W., et al. (2017). Alteration of soil bacterial interaction networks driven by different long-term fertilization management practices in the red soil of South China. *Appl. Soil Ecol.* 120, 128–134. doi: 10.1016/j.apsoil.2017.08.013

Yeoh, Y. K., Paungfoo-Lonhienne, C., Dennis, P. G., Robinson, N., Ragan, M. A., Schmidt, S., et al. (2015). The core root microbiome of sugarcanes cultivated under varying nitrogen fertiliser application. *Environ. Microbiol.* 18, 1338–1351. doi: 10.1111/1462-2920.12925

Yu, M., Liang, S., Dai, Z., Li, Y., Luo, Y., Tang, C., et al. (2021). Plant material and its biochar differ in their effects on nitrogen mineralization and nitrification in a subtropical forest soil. *Sci. Total Environ.* 763:143048. doi: 10.1016/j.scitotenv.2020.143048

Zarraonaindia, I., Owens, S. M., Weisenhorn, P., West, K., Hampton-Marcell, J., Lax, S., et al. (2015). The soil microbiome influences grapevine-associated microbiota. *mBio* 6:e02527-14. doi: 10.1128/mBio.02527-14

Zehr, J. P., Jenkins, B. D., Short, S. M., and Steward, G. F. (2003). Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ. Microbiol.* 5, 539–554. doi: 10.1046/j.1462-2920.2003.00451.x

Zhang, L., Jing, Y., Xiang, Y., Zhang, R., and Lu, H. (2018). Responses of soil microbial community structure changes and activities to biochar addition: a meta-analysis. *Sci. Total Environ.* 643, 926–935. doi: 10.1016/j.scitotenv.2018.06.231

Zhang, C., Tayyab, M., Abubakar, A. Y., Yang, Z., Pang, Z., Islam, W., et al. (2019). Bacteria with different assemblages in the soil profile drive the diverse nutrient cycles in the sugarcane straw retention ecosystem. *Diversity* 11:194. doi: 10.3390/d11100194

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 24 December 2021
doi: 10.3389/fmicb.2021.782621

The Distribution and Turnover of Bacterial Communities in the Root Zone of Seven *Stipa* Species Across an Arid and Semi-arid Steppe

Xiaodan Ma[1,2], *Lumeng Chao*[1,2], *Jingpeng Li*[1,2], *Zhiying Ding*[1,2], *Siyu Wang*[1,2], *Fansheng Li*[1,2] *and Yuying Bao*[1,2*]

[1] *Key Laboratory of Forage and Endemic Crop Biotechnology, Ministry of Education, School of Life Sciences, Inner Mongolia University, Hohhot, China,* [2] *State Key Laboratory of Reproductive Regulatory and Breeding of Grassland Livestock, Inner Mongolia University, Hohhot, China*

OPEN ACCESS

Edited by:
Stilianos Fodelianakis, Swiss Federal Institute of Technology Lausanne, Switzerland

Reviewed by:
Hannes Peter, Swiss Federal Institute of Technology Lausanne, Switzerland
Ramona Marasco, King Abdullah University of Science and Technology, Saudi Arabia
Tyler J. Kohler, Swiss Federal Institute of Technology Lausanne, Switzerland

***Correspondence:**
Yuying Bao
ndbyy@imu.edu.cn

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *24 September 2021*
Accepted: *26 November 2021*
Published: *24 December 2021*

Citation:
Ma X, Chao L, Li J, Ding Z, Wang S, Li F and Bao Y (2021) The Distribution and Turnover of Bacterial Communities in the Root Zone of Seven Stipa Species Across an Arid and Semi-arid Steppe. Front. Microbiol. 12:782621. doi: 10.3389/fmicb.2021.782621

The bacterial communities of the root-zone soil are capable of regulating vital biogeochemical cycles and the succession of plant growth. *Stipa* as grassland constructive species is restricted by the difference features of east–west humidity and north–south heat, which shows the population substituting distribution. The distribution, turnover, and potential driving factors and ecological significance of the root-zone bacterial community along broad spatial gradients of *Stipa* taxa transition remain unclear. This paper investigated seven *Stipa* species root-zone soils based on high-throughput sequencing combined with the measurements of multiple environmental parameters in arid and semi-arid steppe. The communities of soil bacteria in root zone had considerable turnover, and some regular variations in structure along the *Stipa* taxa transition are largely determined by climatic factors, vegetation coverage, and pH at a regional scale. Bacterial communities had a clear *Stipa* population specificity, but they were more strongly affected by the main annual precipitation, which resulted in a biogeographical distribution pattern along precipitation gradient, among which *Actinobacteria*, *Acidobacteria*, *Proteobacteria*, and *Chloroflexi* were the phyla that were most abundant. During the transformation of *Stipa* taxa from east to west, the trend of diversity shown by bacterial community in the root zone decreased first, and then increased sharply at *S. breviflora*, which was followed by continuous decreasing toward northwest afterwards. However, the richness and evenness showed an opposite trend, and α diversity had close association with altitude and pH. There would be specific and different bacterial taxa interactions in different *Stipa* species, in which *S. krylovii* had the simplest and most stable interaction network with the strongest resistance to the environment and *S. breviflora* had most complex and erratic. Moreover, the bacterial community was mainly affected by dispersal limitation at a certain period. These results are conducive to the prediction of sustainable ecosystem services and protection of microbial resources in a semi-arid grassland ecosystem.

Keywords: ***Stipa*** **taxa, root-zone, bacterial communities turnover, assembly processes, environmental factor**

INTRODUCTION

Soil microorganisms play an important role in regulating and maintaining biogeochemical cycles in a terrestrial ecosystem, and are sensitive to environmental changes (Chen et al., 2020). Arid and semi-arid steppe ecosystem, as one of the most important components of the terrestrial system (Kang et al., 2007; Wang et al., 2021), is becoming fragmented and degraded due to climate changes and anthropogenic activity (Gao et al., 2018; Zhang Y. et al., 2018). Climates, vegetation, and soil properties in such steppes vary greatly (Lv et al., 2016), and these changes will strongly affect the structure and function of the soil microbial community (Tu et al., 2017). Numerous studies have focused on soil microbial diversity and assembly processes in arid and semi-arid ecosystems, as well as on the variations and distribution of microorganisms along spatio-temporal, altitude, longitude, plant diversity, and drought gradients (Tu et al., 2017; Yao et al., 2017; Richter-Heitmann et al., 2020; Wang et al., 2021). However, plant root-zone soils have been relatively neglected in complex natural communities. To be specific, the rhizosphere acts as the region of interactions between soil microorganisms and the plant root system's own biological activities (Verma et al., 2019). Moreover, rhizosphere soil microbes are involved in the slow ecological process of vegetation succession and evolution to adapt to the regional environments (Song et al., 2019). This study systematically investigates the turnover and distribution of root-zone soil microbial communities and the correlation with plants along broad spatial gradients of the plant community succession distribution, which is critical to indicating and predicting ecosystem functions.

Abundant rhizosphere microbial resources create a highly evolved external functional environment for plants, thereby allowing organic matters to degrade, nutrients to be released from minerals, nitrogen to be fixed, and elemental forms to be transformed (Lau and Lennon, 2011; Lakshmanan et al., 2014; Na et al., 2018; Krishna et al., 2020). In turn, plants affect microorganisms by the decomposition of litter, turnover of roots, and the release of exudates (Philippot et al., 2013). The interactions help maintain the stability of ecosystem structures and functions, while positively affecting plant fitness and resistance, especially in extreme environments (Marasco et al., 2018). Bacteria comprise a large part of rhizosphere soil biodiversity and participate in most of the material transformation processes (Chu et al., 2020), hence greatly affecting plant growth and establishment. Studies have shown that the size, structure, and activity of the soil bacterial communities could be greatly affected by individual plant species (Liu et al., 2021). Thus, gaining insights into the variation and interactions of plant root-zone soil bacterial communities will enhance our understanding of the ecology and function of soil bacteria. Over the past few decades, spatial and temporal variation in bacterial communities and linkages with plant communities have been extensively investigated in a wide variety of habitats (Tian et al., 2017; Na et al., 2018; Zeng et al., 2019; Chu et al., 2020), whereas such studies have been rare for the root zone of natural forages.

As indicated from existing studies, the soil rhizosphere bacterial composition and relative abundance are highly affected by the biotic and abiotic environments (Fan et al., 2017; Zhang B. et al., 2018). Edaphic factors, particularly soil pH and nutrient availability, vegetation composition and species, regional climate, and altitude, have been shown to shape soil bacterial communities (Na et al., 2018; Zhalnina et al., 2018; Clairmont et al., 2019; Mohanram and Kumar, 2019). Moreover, bacterial evolutionary responses may be driven by edaphic and non-edaphic variables that function as selective pressures; for example, the relationship between the adaptive range of pH and biogeographical patterns (Na et al., 2018), or the greater sensitivity of bacteria in drier areas to environmental stimuli (Maestre et al., 2015). Recently, chemical element indices in soil (e.g., Fe, Cu, Ca, and Mg) have also been proposed to directly impact or predict bacterial diversity and composition (Corneo et al., 2013; Wang et al., 2020). Microbial community assembly (both deterministic and stochastic processes) is critical to maintaining the spatial distribution and composition of microbial communities from a local to a global scale (Stegen et al., 2013; Zhang et al., 2016). Microbial taxa interactions also describe the underlying ecological processes, which can affect the response of communities to environmental variations and may be more important than environmental variables in determining community structures (Zhang B. et al., 2018). For example, it is predicted that ecological networks that consist of weak interactions are more stable than those with strong interactions (Coyte et al., 2015). Yet little is known about how the mentioned ecological processes govern rhizosphere soil bacterial community turnover during the obviously geographic substitution of the plants.

The genus *Stipa* refers to a grazing tolerant, drought-resistant perennial bunchgrass species (Gao et al., 2018; Chen et al., 2020) that is an important ecological barrier and basically in the highly specialized stage of the grassland community succession and evolution (Lv et al., 2016; Liu et al., 2019). From eastern to western Inner Mongolia, across various combinations of hydro-thermal characteristic, different moisture ecotypes of *Stipa* occupy zonal habitats and form meadow steppe, typical steppe, and desert steppe. Accordingly, the *S. grandis* community is generally replaced by the mesophytic *S. baicalensis* community that connects the forest grassland subzone in the east under the progressively humid habitat conditions. As opposed to the mentioned, the *S. grandis* community is largely distributed in the eastern part of the typical steppe, and is often replaced by the *S. krylovii* community when the habitat conditions tend to be dry or frequently grazed. Also, the *S. krylovii* community in the western region of the typical steppe exhibits more xerophytic characteristics and is replaced by the warm-loving *S. breviflora* when the habitat is transmitted to the desert steppe to the west and southwest. Furthermore, *S. breviflora* crosses the southeastern edge of the desert steppe zone from east to west, and then the drier communities of *S. gobica*, *S. klemenzii*, and *S. glareosa* appear to the northwest. Moreover, *S. gobica* extends westward to the desert, and redundantly is combined with *S. klemenzii* and *S. glareosa* communities (Lv et al., 2016; Li et al., 2018; Nan et al., 2020). *Stipa* taxa show divergences

in morphological and physiological parameters (Han and Tian, 2016). The mentioned typical divergent evolution processes provide a unique opportunity and ideal study area to investigate the root-zone bacterial communities in natural ecosystems. The following hypotheses were verified here: (1) There is considerable turnover of root-zone soil bacterial communities along the *Stipa* taxa transition and largely determined by climatic factors, vegetation coverage, and pH at a regional scale. (2) Bacterial communities in root zone have a clear *Stipa* population specificity and the biogeographic distribution pattern would not be affected solely by climatic factors. (3) There would be specific and different bacterial taxa interactions and different resistance to environment in different *Stipa* species, and that the bacterial community assembly is dominated by dispersal limitation. This paper aimed to investigate the geographic distribution, turnover, and interaction of root-zone bacterial community across a transect of zonal *Stipa* species across broad spatial gradients. The identification of potential drivers is critical to elucidating between soil bacterial and plant populations, which can more effectively predict grassland ecosystem responses and functions under a changing climate, and lay a theoretical basis for gaining insights into the community structure and influencing factors of rhizosphere soil bacteria.

MATERIALS AND METHODS

Site Description

This study was conducted along a 1,700-km east–west transect across Inner Mongolia (41.52–50.12 N, 108.46–120.21 E) in northern *China*. A total of 32 sampling sites that had nearly natural plant communities with only light animal grazing were selected along the zonal distribution of *Stipa* spp.; six for *S. baicalensis* (S1–S6), 11 for *S. grandis* (S7–S17), eight for *S. krylovii* (S18–S25), three for *S. breviflora* (S28–S30), two for *S. gobica* (S31–S32), and one each for *S. glareosa* (S26) and *S. klemenzii* (S27). Samples were gradually collected from west to east from early August 2019 through the entire transect, to ensure that *Stipa* were in the same mature phenological stage (**Figure 1A**). Identification of *Stipa* species was according to FLORA INTRAMONGOLICA (EDITIO TERTIA Tomus 6: 205–217). Spatial geographic coordinates, climatic information, and the plant indexes (Plant coverage and Biomass of *Stipa*) for each sampling site are listed in **Supplementary Table S2**. The region has a temperate continental arid and semi-arid monsoonal climate with a mean annual average precipitation of 172.19–472.06 mm and a mean annual temperature of −0.12–6.56°C (https://disc.gsfc.nasa.gov/). The soil types in the study area include forest grassland black soil and chernozem, chestnut soil in the dry grassland, and brown calcareous soil and lime-calcium soil in the desert steppe.

Sampling of Root-Zone Soil and Plant Accessions

At the respective site ($n = 32$), three independent replicate locations (50 × 50 m, at least 5 km apart) were selected (**Figure 1**). In accordance with the double diagonal principle, five 1 × 1 m plots were evenly distributed on the center point of the respective independent location and the two diagonals of the four corners for a vegetation survey, which included plant diversity, number, vegetation coverage, and a calculation of the Shannon–Wiener index ($H' = -\sum_{i=1}^{r}[PiLnPi]$, where *Pi* denotes the ratio of the number of each species to the overall number of all species present). The *Stipa* species exhibits a caespitose growth form, which contains numerous tillers and culms emerging from a single rootstock. After establishment, soil, litter, roots, and live shoots accumulate to form tussocks (Bai et al., 1999). The active root system of *Stipa* has rhizosheath (a layer of adhering soil particle to the root surface) that does not easily fall off, creating sufficient water and inorganic salts for the plant and protecting the fibrous roots from mechanical damage (Basirat et al., 2019). The age of perennial bunch grasses is difficult to find through field observation. In the sampling process, plants with tussock fragmentation and self-thinning branches, or too small or excessively large plants should not be selected, as an attempt to collect representative root-zone soil samples and minimize the plant species-age effects on bacterial community diversity and composition (Bai et al., 1999; Na et al., 2018). Given the mentioned information, 25 individual *Stipa* were collected at the respective sampling site, and the excavation depth should exceed 30 cm (Chen, 1987) to ensure the root system to be as complete as possible. To be specific, 15 individual *Stipa* were applied for soil samples and 10 individual *Stipa* for vegetation samples. Several root segments were randomly sampled from each individual *Stipa* after shaking off the extremely loose root-attached soil. Next, the root segments from 15 individual *Stipa* were integrated as one site sample. Subsequently, root-zone soil (Edwards et al., 2015; Shi et al., 2019) was collected from the respective site sample with a disposable brush to brush out the soil around the rhizosheath and then divided equally into three parts. On the whole, 32 × 3 root-zone soil samples were collected, and immediately stored in liquid nitrogen until 16S rRNA high-throughput sequencing analysis could be performed. Brushes and gloves were sterile and disposable to avoid contamination during sampling. The mentioned 10 individual *Stipa* for vegetation samples fell into two equal parts—one part was used for weighing the average aboveground biomass of *Stipa* with the drying method at 80°C, and the other was mixed, crushed, and sieved through a sieve with 0.154 mm pore size into one sample to determine the nutritive value of the forage. The longitude and latitude of each location were recorded from a portable GPS. The collected soil samples fell into two parts: (1) stored at 4°C for soil physical and chemical analysis, and (2) sieved through pore size of 0.074 mm to analyze of soil chemical elements.

Analysis of Environmental Predictors

A total of 25 environmental predictors were measured. The pH was determined using an SKW 500 soil monitor (Palintest Co., Ltd, UK). Total nitrogen (TN), ammonium (NH_4^+-N), nitrate nitrogen (NO_3^--N), and plant crude protein (CP) were measured with an automatic Kjeldahl apparatus (K1100 Hanon; Haineng Future Technology Group Co., Ltd, China) after digestion in H_2SO_4 ($\rho = 1.8419$ g/L). Soil organic matter (SOM) and soil organic carbon (SOC) were determined by dichromate oxidation (Chen et al., 2015). Alkaline and neutral Olsen-phosphorus (AP)

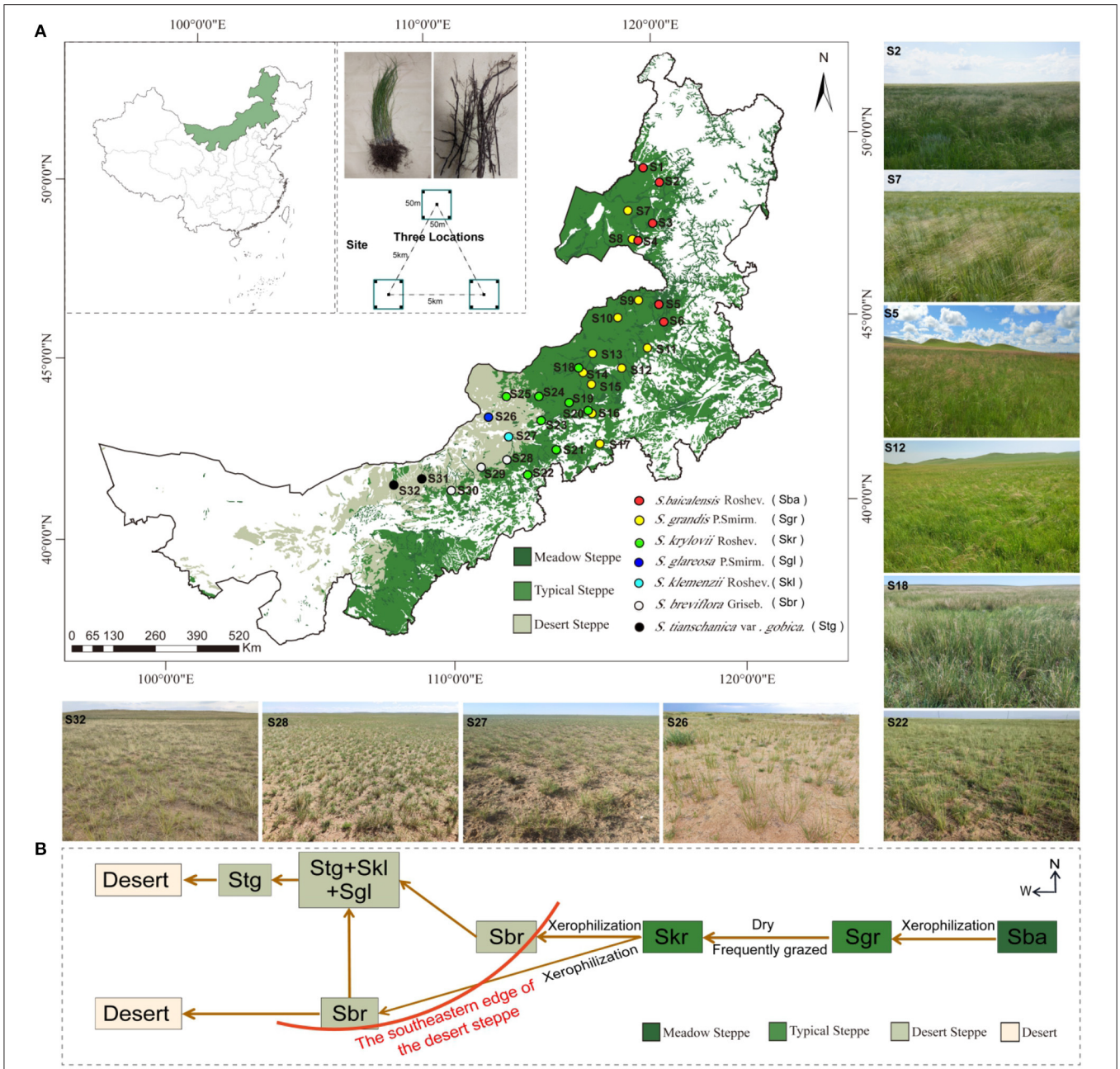


FIGURE 1 | (A) Map of sample locations, landscape, and schematic diagram in arid and semi-arid *Stipa* steppe. Each circle represents one sampling site, and each color represents a *Stipa* population; samples collected from different sites and species are referred to as *Stipa baicalensis* Roshev (Sba: S1–S6), *Stipa grandis* P. Smirn (Sgr: S7–S17), *Stipa krylovii* Roshev (Skr: S18–S25), *Stipa glareosa* P. Smirn (Sgl: S26), *Stipa klemenzii* Roshev (Skl: S27), *Stipa breviflora* Griseb (Sbr: S28–S30), and *Stipa tianschanica* var. *gobica* (Stg: S31–S32). **(B)** Schematic diagram of alternative distribution of *Stipa* population.

were determined based on the sodium bicarbonate extraction (0.05 mol·L^{-1}) molybdenum antimony anti-colorimetric method (Wang et al., 2020). Available potassium (AK) was determined with a kit according to the instructions (Suzhou Keming Biotechnology Co., Ltd., Suzhou, China). Root-zone soil pretreatment was carried out by microwave digestion system (MARS 6; CEM Co., Ltd, USA) according to a soil and sediment digestion of total metal elements–microwave-assisted acid digestion method (HJ 832−2017, Ministry of Environmental Protection MEP, *China*), and then soil chemical elements (S, Ca, Mg, Fe, Cu, Zn, and Mn) were determined by ICP-OES (PlasmaQuant PQ9000; Jena Analytical Instrument Co., Ltd, Germany). Plant crude fiber and crude fat were measured by an acid–base heating digestion method (GB/T5009.10-2003;

Ministry of Health, P. R. China) and a Soxhlet extraction method (GB/T6433-2006; General Administration of Quality Supervision GAQSIQ, China), respectively. Mean annual temperature and mean annual precipitation was obtained from the GES DISC of NASA (https://disc.gsfc.nasa.gov/).

DNA Extraction, PCR Amplification, and Illumina Sequencing

Triplicate genomic DNA samples were extracted from 0.5 g of the composite soil samples using an E.Z.N.A. soil DNA kit (Omega Bio-tek, Norcross, GA, USA) according to the manufacturer's instructions. The DNA extract was run on a 1% agarose gel, and the DNA concentration and purity were determined with a NanoDrop 2000 UV–vis spectrophotometer (Thermo Scientific, Wilmington, USA). To generate bacterial PCR amplicon libraries, universal 16S rRNA gene primers (338 forward: 5′-ACTCCTACGGGAGGCAGCAG-3′ and 806 reverse: 5′-GGACTACHVGGGTWTCTAAT-3′). The PCR amplification of the 16S rRNA gene was performed according to details in **Supplementary Materials**. Purified amplicons were pooled in equimolar quantities and paired-end sequencing (2 × 300 bp) on an Illumina MiSeq platform (Illumina, San Diego, USA) according to standard protocols by Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China). All of the sequence data generated for this study have been deposited in the NCBI Sequence Read Archive (SRR14415681–SRR14415776).

Bioinformatics Workflow

The raw 16S rRNA gene sequences obtained from the MiSeq platform were quality filtered, trimmed, de-noised, and merged using Mothur 1.32.2 (Schloss et al., 2009) where chimeric sequences were identified and removed using the UCHIME *de novo* algorithm (Perez-Jaramillo et al., 2019). Subsequently, the remaining high-quality sequences were clustered into operational taxonomic units (OTUs) using the UPARSE (version 7.0.1090, http://www.drive5.com/uparse/) algorithm, setting a distance limit of 0.03 using the open-reference OTU picking protocol. A representative sequence was aligned using the RDP Classifier (http://rdp.cme.msu.edu/) against the 16S rRNA database (Silva 132; https://www.arb-silva.de/) using a confidence threshold of 0.7. For downstream analysis, OTUs affiliated with chloroplasts and mitochondria were subsequently removed from the bacterial OTU table, and OTUs that were assigned to non-bacteria, including plants and protozoans, were removed from the bacterial OTU table (Perez-Jaramillo et al., 2019). The effects of sampling on diversity were corrected by rarifying the sequence numbers of each sample to that of the sample with the lowest number of reads (17,865 reads). The number of OTUs, community richness (Chao index), community diversity (Simpson index), community evenness (Shannon index), and a sequencing depth index (Good's coverage) were subsequently calculated using MOTHUR (v.1.30.1; www.mothur.org).

Statistical Analysis

All statistical analyses were conducted through a one-way ANOVA with IBM SPSS Statistics 26.0 ($p < 0.05$ was considered statistically significant). Non-metric multidimensional scaling [NMDS, stress < 0.20 was considered acceptable (Kuczynski et al., 2011)] and hierarchical clustering (unweighted pair group method with arithmetic means, UPGMA) that all used the Bray–Curtis distance to assess the *Stipa* population-specificity distribution of bacteria, as well as the similarity of bacterial community composition. To demonstrate the effect of *Stipa* population on the bacterial composition, permutational multivariate analysis of variance (PERMANOVA, $p < 0.05$) was conducted with the adonis function of the vegan package (Glasl et al., 2019) with 999 permutations. A linear discriminant analysis (LDA score threshold of 4.0) coupled with effect size (LEfSe) was adopted to search for statistically different biomarkers among groups of *Stipa* taxa root zone. In addition, Circos-0.67-7 was employed to determine the connection between the dominant phylum or genus and the sample group of *Stipa* species. The variance inflation factor (VIF) was exploited to remove the redundant variables (VIF > 20) (Fan and Xing, 2016) from the explanatory variables and avoid collinearity among a range of factors. The Mantel test and the Pearson correlation analysis were performed to examine the relationship between bacterial communities and environmental factors using the "Vegan" package (Oksanen et al., 2015) and the "ggcor" package (Wang et al., 2021) version 0.9 implemented in R, respectively. A redundancy analysis (RDA) was also performed to determine the most significant environmental variables shaping the microbial community composition.

Bacterial co-occurrence network analyses for *Stipa* species were conducted based on the Molecular Ecological Network Analyses Pipeline (MENAP, http://ieg4.rccc.ou.edu/MENA/main.cgi) at the OTU level (Deng et al., 2012). To simplify the networks for a more effective visualization and unified analysis conditions, only the 300 most abundant OTUs were analyzed and network topology parameters (i.e., network complexity and modularity) were determined to indicate the stability of the network and the resistance from environmental interference. Abundances of OTUs were log transformed. To compare networks under the identical conditions, a threshold of 0.82 (the recommended similarity threshold) was adopted to build the networks for the respective *Stipa* taxa in the study. Lastly, the networks were visualized with the Gephi 0.9.2 (Zhou et al., 2020). Furthermore, 100 random networks were generated and the properties were compared with the experience network to verify whether the constructed network is reasonable and effective. The different assembly processes of community composition were determined by phylogenetic null model analyses. All OTUs were adopted to build a maximum-likelihood tree in FastTree (Tang et al., 2021) to determine the weighted β-nearest taxon index (βNTI) using the picante package (Dini-Andreote et al., 2015). The βNTI values $> +2$ (variable selection) or < -2 (homogeneous selection) implies significantly more or less phylogenetic turnover than expected (Tripathi et al., 2018), respectively, thereby demonstrating the predominance of deterministic processes. If the $|\beta NTI| \leq 2$, stochastic processes are predominant. To more specifically differentiate between stochastic scenarios of assemblages of *Stipa* species root-zone soil bacteria, a Raup–Crick matrix (RCbray) based on the Bray–Curtis matrix was computed with the vegan package (Cheng

et al., 2021). |βNTI| < 2 and RCBray < −0.95, |βNTI| < 2 and RCBray > 0.95, and |βNTI| < 2 and |RCBray| <0.95 denote homogenizing dispersal, dispersal limitation, and undominated processes, respectively (e.g., weak selection, weak dispersal, diversification, and/or drift) (Stegen et al., 2013; Dini-Andreote et al., 2015). Furthermore, unless otherwise indicated, all analyses were conducted in *R* (version 3.5.1; *R* Development Core Team).

RESULTS

Vegetation and Soil Characteristics of the Study Area

In accordance with **Supplementary Tables S1, S2**, the MAP of the semi-arid grassland decreased gradually from east (472.06 mm) to west (152.08 mm). SOM, SOC, TN, NH_4^+-N, VC, and SB tended to increase with precipitation, while pH, altitude, and temperature tended to decrease. The soils were moderately alkaline and the average pH increased significantly from 6.06 to 8.74. Other test indices displayed significant differences among 32 sites, but no obvious rules to follow. There were significant differences in vegetation indexes among seven *Stipa* species, and the vegetation coverage (66.67 ± 14.38 to 27.50 ± 3.61) and *Stipa* biomass (3.00 ± 1.31 to 0.85 ± 0.32) decreased significantly under the process of Sbr replacing Sba distribution (**Supplementary Table S3**). Most soil parameters differed significantly between the various root-zone soils of *Stipa*, except for TK and NO_3^--N, in which macronutrients (SOM, SOC, TN, NH_4^+-N) in Sba were significantly higher than for other *Stipa* taxa (**Supplementary Table S3**). A Pearson correlation analysis (**Figure 6**) found that most variables (MAP, MAT, altitude, soil pH, SOC, TN, S, VS, and SB) had strong positive or negative correlations with each other. In addition, there was a significant positive correlation between soil chemical elements.

Turnover of Root-Zone Bacterial Community Composition of *Stipa* Taxa

We obtained a total of 4,665,705 raw reads by Illumina MiSeq high-throughput sequencing. After quality filtering, trimming, and assigning reads to the different samples, 3,150,865 high-quality reads were recovered in the dataset, representing 17,865 bacterial operational taxonomic units (OTUs) based on 97% sequence identity across all samples, with a median of 2,349 OTUs per sample (**Supplementary Table S4**). The observed OTUs and rarefaction curves of Shannon index, Chao richness, and Good's coverage (values were over 0.94) in all soil samples were plateaued (**Supplementary Figure S1**), suggesting that the sequencing depth was sufficient.

Circular visualization diagrams were adopted to display the overlapping and differentiating bacterial taxa among the seven *Stipa* taxa at the phylum and genus levels (**Figures 2A,B**). The bacterial community consisted of 37 different phyla, and the main bacterial phyla (top 11) collectively accounted for 96.98–98.76% of all taxon sequences (**Figure 2A**). Moreover, the five most abundant bacterial genera comprised 27.71% of the bacterial communities on average (**Figure 2B**). *Actinobacteria*, *Acidobacteria*, *Proteobacteria*, and *Chloroflexi* were the most abundant phyla observed in all of the *Stipa* taxa analyzed (**Figure 2A**). Except for Sgl, *Actinobacteria* (28.34–45.35%) was the dominant phylum in other *Stipa* species groups, especially in Sbr root zone (45.35%). *Acidobacteria* had a most significant advantage as dominant phylum in Sgl (32.81%) compared with other six *Stipa* taxa (10.73–24.10%). When Sba was replaced by Sbr from east to west along the transect, the relative abundance of *Chloroflexi* (phylum, from 8.17 to 14.80%) and *Rubrobacter* (genus, from 1.89 to 10.20%) increased significantly, and *Rubrobacter* was the most dominant genus in Sbr (10.20%); as opposed to the mentioned, the relative abundances of *Proteobacteria* (from 24.06 to 15.70%) decreased significantly (**Figures 2A,B** and **Supplementary Table S6**). Except Sbr, the relative abundances of genus *RB41* (7–15.27%) and *Subgroup_6* (5.85–13.04%) in the root zone of *Stipa* taxa were significantly higher than for other genera, especially in Sgl (**Figure 2B**). Taken together, these results revealed shifts in the bacterial community composition in root-zone soils of different *Stipa* taxa.

The LEfSe with LDA scores of 4 (**Figure 3**) revealed that, when compared with the Sba root-zone soil bacterial community (eight biomarkers at finer taxonomic level, one phylum, one class, three orders, two families, and one genera), the other *Stipa* species root-zone soil communities had less biomarkers (**Figure 3A**). However, Skr, as an intermediate transitional species, had no biomarkers. Furthermore, Stg and Sgl were distributed over a small area and had no biomarkers. To be specific, root-zone soil of Skl showed two significantly different taxa, *Bacteroidia* (class) and *Bacteroidetes* (phylum); Sgr were rich in *Rhizobiales* (order), *Xanthobacteraceae* (family), and *norank_f__Xanthobacteraceae* (genus); Sba were enriched with *Sphingomonadales* (order), *Chthoniobacterales* (order), *Betaproteobacteriales* (order), *Sphingomonadaceae* (family), *Chthoniobacteraceae* (family), and *Candidatus_Udaeobacter* (genus), while the root-zone soil of Sbr was enriched with *Thermomicrobiales* (order), *JG30_KF_CM45* (family and genus) (**Figure 3B**).

Bacterial Community Diversity Varied and Spatial Distribution of *Stipa* Taxa Root-Zone Soils

The bacterial biodiversity (Simpson index, **Figure 4A**) of the root-zone soil decreased when Sba was replaced by Skr from southeast to northwest along the transect, increased for Sbr, and then decreased sharply as Sbr transitioned to Stg. The trends of indices richness (Chao index **Figure 4C**) and evenness (Shannon index **Figure 4B**) of bacterial were opposite with the variation patterns of the mentioned bacterial diversity. We found that the α diversity of bacterial communities was different across 32 sampling sites with respect to the bacterial richness, diversity, and evenness for 96 soil samples, and a lower bacterial diversity was observed in more arid areas, as indicated by the changes in the Simpson index (**Supplementary Table S2**). Furthermore, bacterial evenness (Shannon index) and diversity (Simpson index) were more affected by altitude and pH, and bacterial evenness also affected crude protein; bacterial richness, as indicated by the Chao index, was affected mostly by the

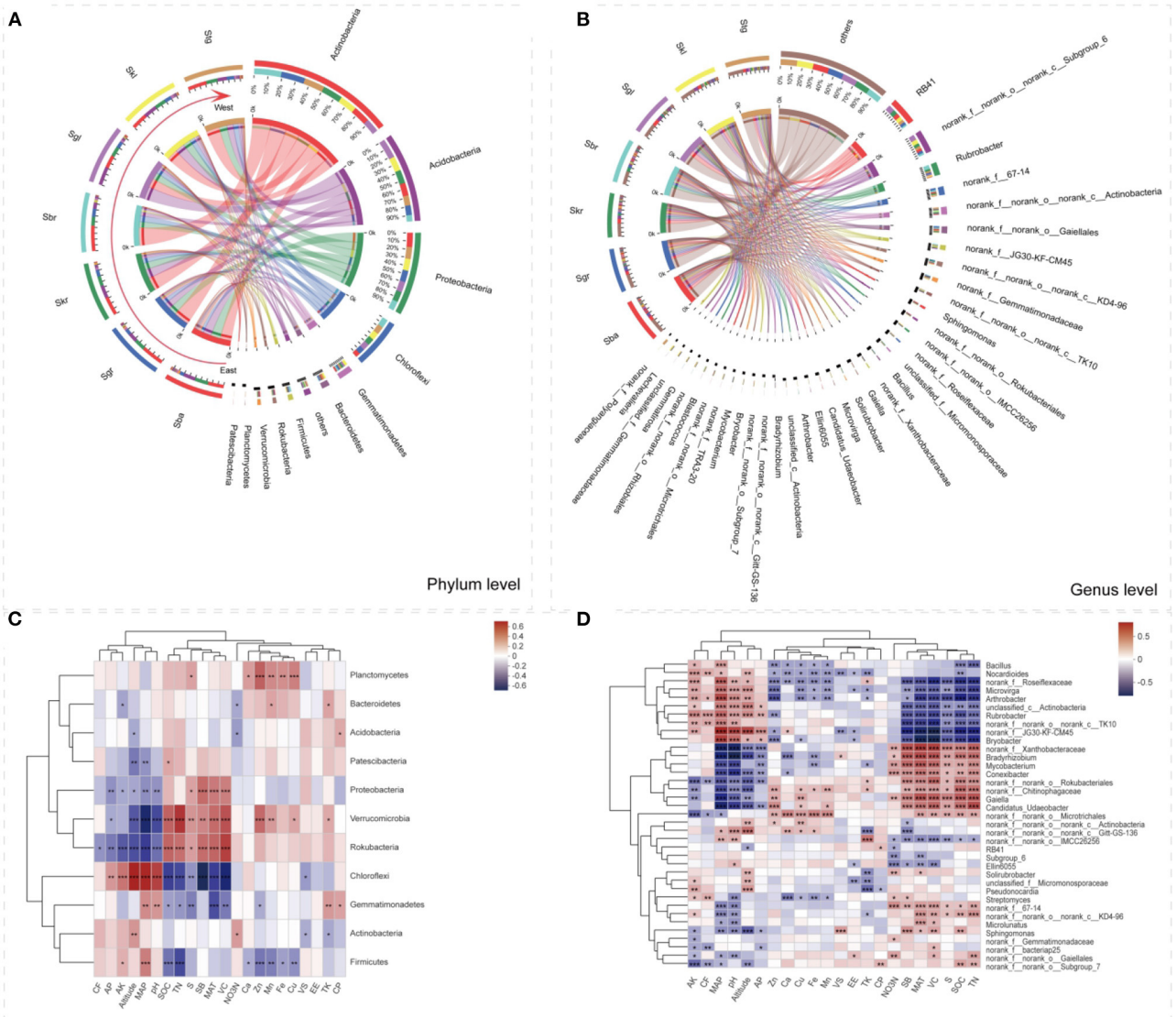


FIGURE 2 | The distributions of root-zone soil bacterial dominant taxa at the different *Stipa* species. **(A)** Represents the phylum level; **(B)** represents the genus level, and Spearman correlation analysis of relative abundance of bacterial keystone taxa with environmental factors: **(C)** phylum level, **(D)** genus level. Red represents positive correlation and blue represents negative correlation (significance levels are as follows: $^{*}0.01 < p \leq 0.05$, $^{**}0.001 < p \leq 0.01$, $^{***}p \leq 0.001$). Arrows refer to the geographic substitution order of *Stipa* species from east to west. SOC, soil organic carbon; TN, total nitrogen; TK, total kalium; NO_3^--N, nitrate nitrogen; AP, alkaline and neutral Olsen-phosphorus; AK, available potassium; Fe, iron element; Cu, copper element; Zn, zinc element; Mn, manganese element; Ca, calcium element; S, sulfur element; VS, Shannon–Wiener index of plot; VC, plant coverage; SB, biomass of *Stipa*; CP, crude protein; EE, crude fat; CF, crude fiber; MAT, mean annual temperature; MAP, mean annual precipitation. *S. baicalensis* (Sba), *S. grandis* (Sgr), *S. krylovii* (Skr), *S. glareosa* (Sgl), *S. klemenzii* (Skl), *S. breviflora* (Sbr), and *S. gobica* (Stg).

altitude and was also principally connected with crude fat (**Figure 6**).

NMDS analysis revealed a clear separation of the bacterial communities from different *Stipa* taxa (**Figure 5B**), and *Stipa* population specificity was further confirmed with a PERMANOVA ($p = 0.001$, **Supplementary Table S5**). In addition, the UPGMA clustering analysis also revealed similarity distribution in the bacterial community composition based on sample sites (**Figure 5A**). The samples formed two clear large clusters; moreover, six clear small clusters (I–VI) were generated, and two of these included only from Sba (VI) and Sgr (IV). The bacterial community was roughly distributed from southeast (VI) to northwest (I) along the reduced precipitation gradient and increased temperature gradient (**Figures 5C,D**). MAP ($R^2 = 0.78$, $p < 0.01$) was more dominant than the *Stipa* taxa ($R^2 = 0.71$, $p < 0.01$) for the distribution of bacterial community, which was indicated by the NMDS plot ordination analysis (**Supplementary Table S7**). In addition, soil pH ($R^2 = 0.63$, $p <$

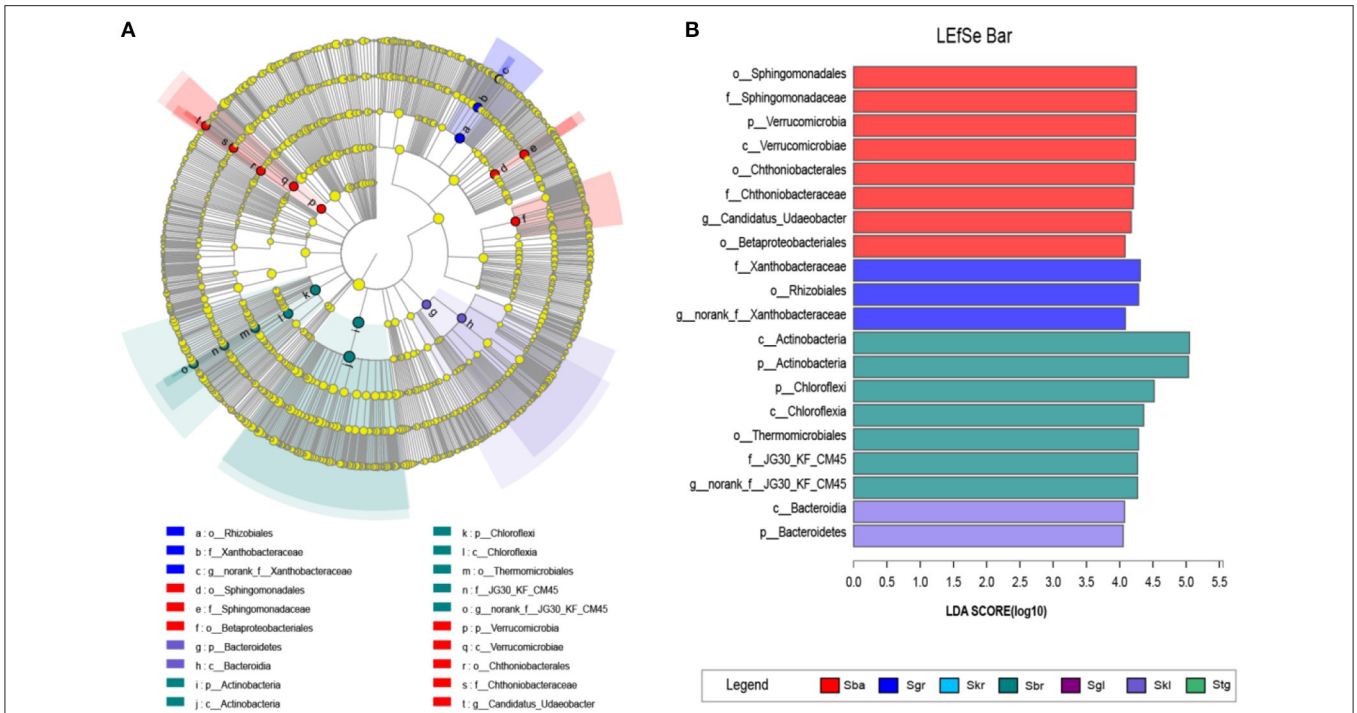


FIGURE 3 | LEfSe analysis of soil bacterial abundance in seven *Stipa* species root-zone soils. **(A)** Cladogram of microbial communities. Cladograms indicate the phylogenetic distribution of microbial lineages associated with the study *Stipa*; circles represent phylogenetic levels from kingdom to genus. **(B)** Indicator microbial groups with LDA scores > 4 and *p*-values < 0.05. *S. baicalensis* (Sba), *S. grandis* (Sgr), *S. krylovii* (Skr), *S. glareosa* (Sgl), *S. klemenzii* (Skl), *S. breviflora* (Sbr), and *S. gobica* (Stg).

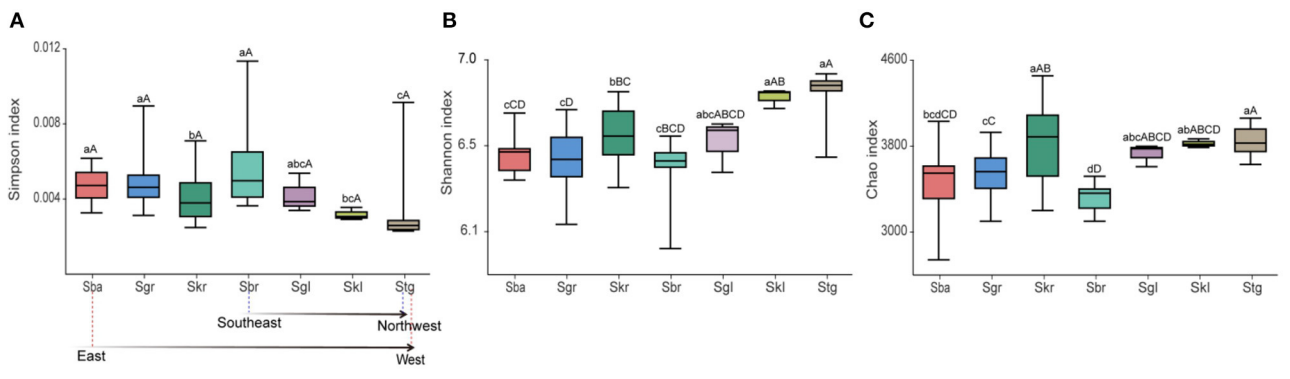


FIGURE 4 | Alpha-diversity indices of bacterial communities in root-zone soils of seven *Stipa* species along a transect from east to west. Different letters (a–g, $p < 0.05$/A–D, $p < 0.01$) above the bars show significant difference among groups based on one-way ANOVA. **(A)** Simpson index expresses microbial diversity. **(B)** Shannon index expresses microbial evenness. **(C)** Chao index expresses microbial richness. *S. baicalensis* (Sba), *S. grandis* (Sgr), *S. krylovii* (Skr), *S. glareosa* (Sgl), *S. klemenzii* (Skl), *S. breviflora* (Sbr), and *S. gobica* (Stg).

0.01), vegetation coverage ($R^2 = 0.64$, $p < 0.01$), and MAT ($R^2 = 0.68$, $p < 0.01$) also had a strong effect on the distribution of the bacterial communities across the geographical replacement gradient of the *Stipa* taxa (**Supplementary Table S7**). The mentioned results indicate that the composition of root-zone bacterial community not only had the *Stipa* population specificity, but also was strongly regulated by climate, and had a change pattern along the hydrothermal gradient. Especially in the process of transition from Sba to Sgr and then to Skr, root-zone bacteria distribution has a certain phenomenon of cross overlap.

Ecological Factors Effecting Root-Zone Bacterial Community Spatial Turnover

Three redundant variables (NH_4^+-N, SOM, and Mg) were eliminated based on VIF analysis to reveal the major

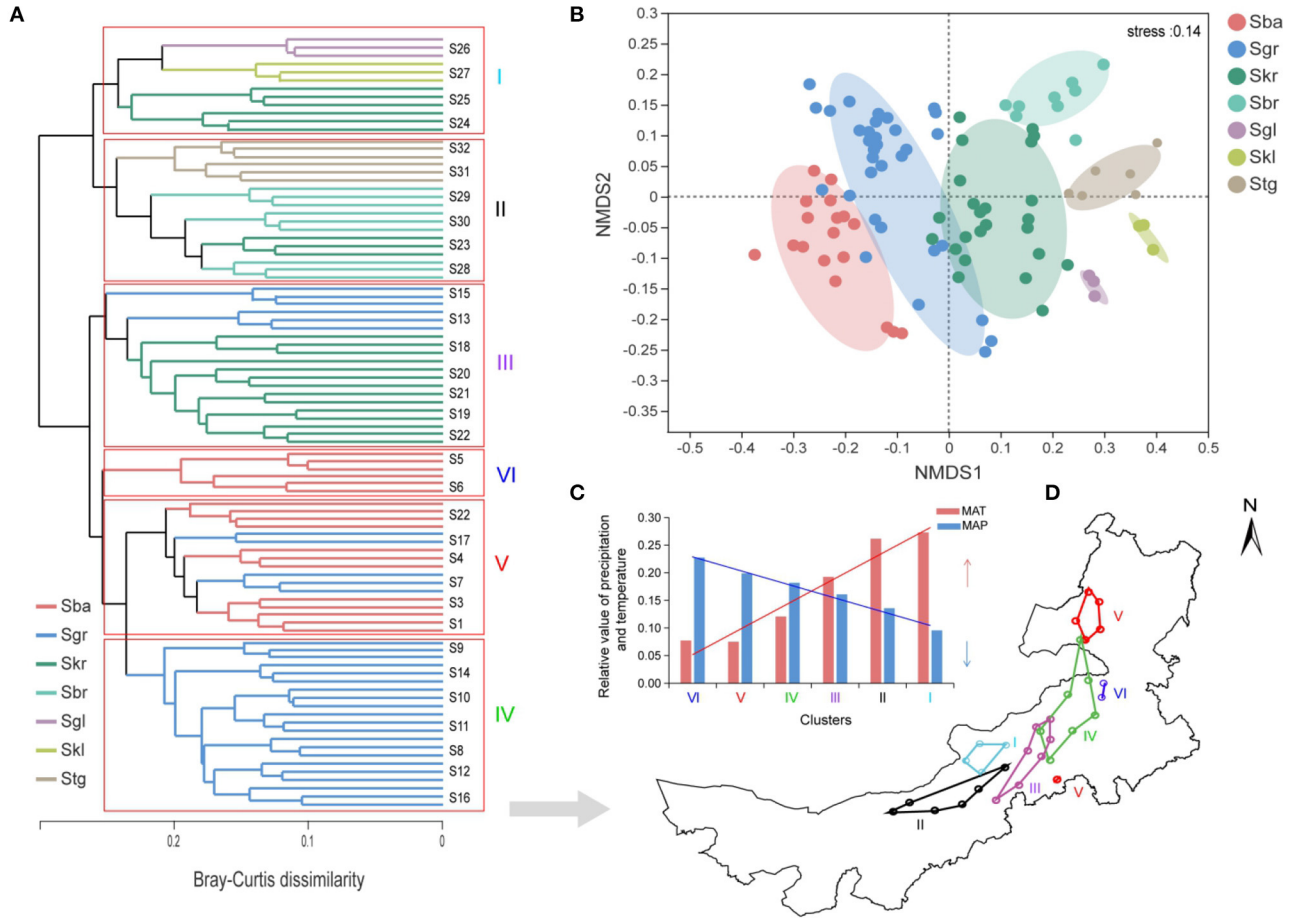


FIGURE 5 | Bacterial community differentiation in the root-zone soils from seven *Stipa* species sampled from different sites of arid and semi-arid steppe. **(A)** Similarity analysis of bacterial community in *Stipa* taxa root zone based on UPGMA clustering using Bray–Curtis distances, clustered into six clusters (I–VI), roman numeral color corresponds to **(C)**. **(B)** Population specificity of *Stipa* root-zone soil bacterial communities. **(C)** Columnar trend graphs of relative mean annual precipitation (MAP) and temperature (MAT) in different clusters; roman numerals correspond to **(A,D)**. **(D)** Biogeographical patterns of bacterial communities, with same color clustering. *S. baicalensis* (Sba), *S. grandis* (Sgr), *S. krylovii* (Skr), *S. glareosa* (Sgl), *S. klemenzii* (Skl), *S. breviflora* (Sbr), and *S. gobica* (Stg).

environmental variables shaping bacterial community composition and turnover. RDA and Mantel test were conducted to distinguish the impact of soil properties, climate, and vegetation factors on bacterial communities. RDA1 and RDA2 axis, respectively, explained 22.54 and 9.97% of the variance in soil bacterial community structures (**Supplementary Figure S2A**). The Mantel tests (**Figure 6**) denoted that the compositions of soil bacterial communities were significantly constructed by MAP (r = 0.58), MAT (r = 0.65), pH (r = 0.50), and VC (r = 0.53). These results were also supported by the RDA (**Supplementary Figure S2A**). The soil chemical elements (i.e., S, Fe, and Cu) also had an effect on the bacterial community structure, but not on the bacterial community distribution (**Figure 6**, **Supplementary Figure S2A**, and **Supplementary Table S7**). SOC, TN, and SB were significantly related to the spatial turnover and distribution of the root-zone soil bacterial community (**Figure 6**, **Supplementary Figure S2A**, and **Supplementary Table S7**). In accordance with the RDA examined, the bacteria phyla at seven *Stipa* root zones were mainly effected by VC (conditional effect = $R^2 = 0.26$, $p = 0.001$), SB ($R^2 = 0.22$, $p = 0.001$), and MAT ($R^2 = 0.19$, $p = 0.001$; **Supplementary Figure S2B**). At genus level, MAP (conditional effect = $R^2 = 0.60$, $p = 0.001$), MAT ($R^2 = 0.52$, $p = 0.001$), MAT ($R^2 = 0.51$, $p = 0.001$), pH ($R^2 = 0.41$, $p = 0.001$), SB ($R^2 = 0.39$, $p = 0.001$), SOC ($R^2 = 0.38$, $p = 0.001$), and TN ($R^2 = 0.38$, $p = 0.001$) affected bacteria, and all bacteria with relative abundance >1% could be explained by one soil property at least (**Supplementary Figure S2C**). The correlation heatmap denoted the relationship between relative abundance of bacterial keystone taxa with environmental factor (**Figures 2C,D**). It was found that the relative abundances of most dominant bacterial phyla and genus were significantly related to most soil physicochemical properties (i.e., soil pH, TN, AP, AK, S, Zn, Mn, Fe, Cu, and SOC), plant factors (i.e., SB and VC), and other indicators (i.e., MAT, MAP, and altitude). Furthermore, *Firmicutes* was significantly positively correlated with soil

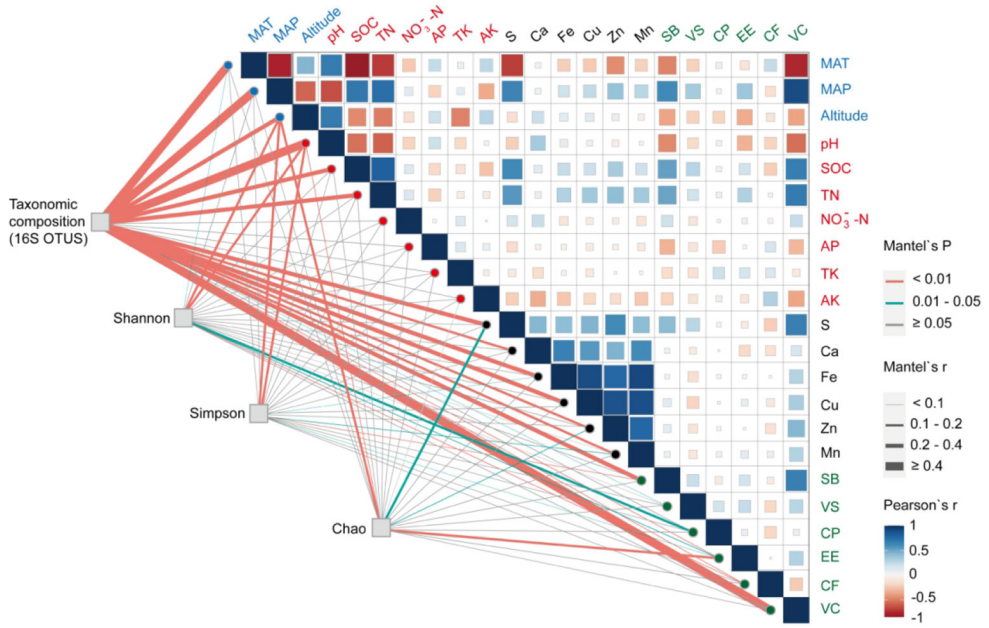


FIGURE 6 | Mantel correlations among bacterial community composition (OTU level), bacterial richness (Chao index), bacterial diversity (Simpson index), bacterial evenness (Shannon index), and environmental factors (climate, plant, and soil variables); Pearson correlation analysis was used among environmental factors. The width of edges represents the size of the correlation coefficient (Mantel's r), while edge color represents the statistical significance based on Mantel's *p*. Abbreviations are consistent with **Figure 2**.

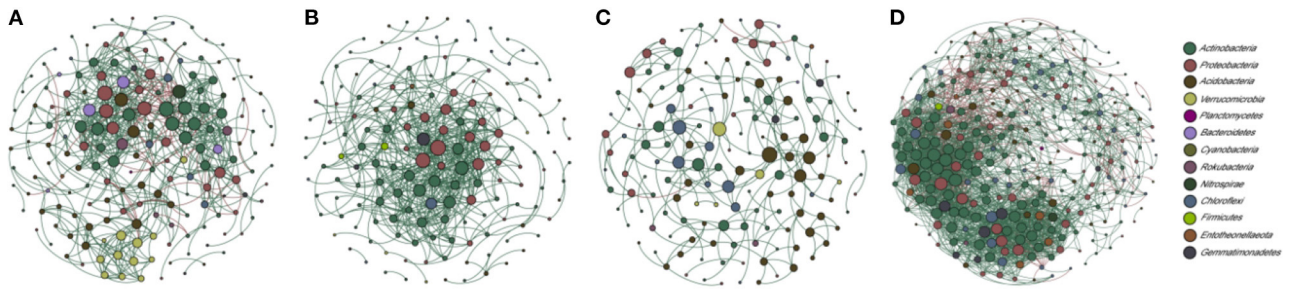


FIGURE 7 | Co-occurrence networks of the high-abundant bacteria from four *Stipa* species root-zone soil; in turn is *S. baicalensis* (Sba: **A**), *S. grandis* (Sgr: **B**), *S. krylovii* (Skr: **C**), and *S. breviflora* (Sbr: **D**). A connection stands for a robust (Spearman's r > |0.82|) and significant ($p < 0.05$) correlation. Green and red lines represent positive and negative associations between two nodes, respectively. Nodes are colored according to microbial phylum and the size of each node is proportional to the degree.

chemical elements (except S); on the contrary, *Planctomycetes* was significantly negatively correlated.

Bacterial Taxa Interactions and Community Assembly Processes in the Root Zone of *Stipa*

To explore the complex bacterial taxa interactions in different *Stipa* taxa root-zone soils and to identify the response to environmental change, a co-occurrence network analysis for abundant OTUs (top 300 OTUs) was performed. The Sgl, Stg, and Skl root-zone soils were excluded from the analyses because of limited samples. Based on the number of links (Sba = 648, Sgr = 527), avgK (Sba = 7.16, Sgr = 6.47), and the avgCC (Sba = 0.38, Sgr = 0.33), the network complexity was comparable between the Sba and Sgr root-zone soils, with nodes in the control network, respectively, grouped into 21 and 25 modules with respective quality values of 0.59 and 0.38 (**Figures 7A,B** and **Supplementary Table S8**). Moreover, the complexity (links = 226, avgK = 2.83, avgCC = 0.32) was less in Skr root-zone soils (**Figure 7C**), and the network grouped into 22 (quality value of 0.81) modules. However, the complexity (links = 2593, avgK = 18.13, avgCC = 0.49) greatly increased in Sbr root-zone soils (**Figure 7D**), which contained six modules, with a quality value of 0.49. There were absolutely more positive correlations in each network than negative correlations. *R* square of power-law values

of Sba, Sgr, and Skr networks were obtained for 0.87, 0.78, and 0.76, respectively, indicating the occurrence of scale-free network characteristics (**Supplementary Table S8**).

The βNTI values were determined, to report that the stochastic process was dominant in the bacterial communities of seven *Stipa* species root-zone soils (**Supplementary Figure S3A**; values >-2 and <2). We calculated the taxonomic-diversity metric RCBray to further quantify the relative effects of various ecological processes, including homogenizing dispersal, dispersal limitation and undominated processes, on the stochastic community assembly of every group (**Supplementary Figure S3B**). Consequently, the relative effects of dispersal limitation (24–49%) and undominated processes (27–41%) increased significantly across the transition from Sba to Skr, while the relative role of homogenizing dispersal (10–48%) was significantly reduced. In arid areas, the relative role of homogenizing dispersal (33–100%) was dominant in the bacterial community assembly in the root zone of *Stipa* species, except for Stg. Overall, the relative effects of dispersal limitation (73%) of *Stipa* genus root-zone soil bacteria were the most important in the grassland system.

DISCUSSION

Turnover of the Root-Zone Bacterial Communities Along the *Stipa* Taxa Transition

Diverse bacterial composition and structure were detected in root-zone soils when the *Stipa* taxa replace each other in space. According to the dominant bacteria taxon, *Actinobacteria* achieved the maximal relative abundance in most *Stipa* species root-zone soils (except for Sgl), which might be attributed to the climate and the soil characteristics of the research area. The members of the bacterial phylum *Actinobacteria* have been proven as an indicator of drought sensitivity and drought tolerance, and exhibit a strong metabolic capacity at low temperatures (Cheng et al., 2021). Given the Circos analysis, the top two *Stipa* taxa in *Acidobacteria* group were Sgl and Skl, which were located in the desert steppe that was characterized by low soil nutrition; in addition, *Acidobacteria* was the most dominant phylum in the Sgl root-zone soil; *Acidobacteria* was composed of various oligotrophs (Zeng et al., 2019), capable of tolerating harsh conditions (Yao et al., 2017). The phylum *Proteobacteria* comprises diverse groups of copiotrophic organisms in soil normally identified under nutrient-rich conditions (Janssen, 2006). The nutrient richness of root-zone soil of Sba is much higher than for other *Stipa* taxa (**Supplementary Table S3**), and the relative abundance of the phylum *Proteobacteria* decreased significantly in the process of Sbr replacing Sba. Moreover, the phylum *Proteobacteria* was negatively correlated with MAT, pH, and AP. The relative abundance of *Proteobacteria* in Sbr root zone was lower than that of several *Stipa* taxa (Sgl, Skl, and Stg) with higher content of AP. Root-zone bacterial phyla and genera of the *Stipa* varied significantly following the vegetation replacement distribution gradient (**Supplementary Table S6**). For instance, when Sba was replaced by Sbr from east to west following the transect, the relative abundance of the phylum *Chloroflexi* and the genus *Rubrobacter* was improved significantly, and *Rubrobacter* was reported as the most dominant genus in Sbr root-zone soil (**Supplementary Table S6** and **Figure 2**). The phylum *Chloroflexi* pertains to the group of heterotrophic oligotrophs in soil (Hug et al., 2013), and the genus *Rubrobacter* is significantly radiation resistant, halotolerant, thermotolerant, or even thermophilic (Kourilova et al., 2021). Most bacteria exhibited a relatively narrow growth tolerance, thereby causing individual taxa to be unable to adapt to soil pH variations (Fierer and Jackson, 2006), and then probably inducing variations of soil bacterial communities (Shen et al., 2013). pH (near neutral range) is suggested to exert a mediating effect in grassland system (Tripathi et al., 2018). This may explain why the relative abundance of *Actinobacteria* and *Acidobacteria* in this paper was considerable, whereas this has no significant correlation with neutral soil pH, while the areas exhibiting an acidic soil pH showed a significant correlation [e.g., the bacterial community in *Changbai* Mountain (Han et al., 2018)]. Together with the significant correlations between soil pH and SOC, TN and NH_4^+-N of root-zone process of grassland might explain why soil pH affected the root-zone bacterial community by affecting the availability of soil carbon and nitrogen or the number of biogeochemical processes involving carbon and nitrogen cycles (Jiao and Lu, 2020), thereby complying consistent with the research on *Caragana* spp. rhizosphere soil (Na et al., 2018). All the mentioned results collectively suggested that the variations of whole bacterial communities are strongly ecological-variable dependent and various phyla and genus may show different variation patterns.

Biomarkers of Soil Bacteria in the Root Zone of *Stipa* Taxa

Besides the differences in the relative abundance of the identical bacterial taxa, soils in different *Stipa* root zones also harbor different potential biomarkers (**Figure 3**), which have been characteristically isolated from soil samples or plant root systems. The Skl only had two significant taxa, *Bacteroidia* (class) and *Bacteroidetes* (phylum), that were involved in C and N metabolism and occurring extensively across various ecological niches (Wang et al., 2020). Several species of the family *Sphingomonadaceae* (*Sphingomonadales*) utilize a wide range of organic compounds for growth and survival under low-nutrient conditions (Glaeser and Kämpfer, 2014). The family *Chthoniobacteraceae* (*Chthoniobacterales*), represented by the genera *Chthoniobacter* and *Candidatus Xiphinematobacter*, can metabolize organic carbon (Victoria et al., 2012) and prevent nutrient leaching and erosion (Kant et al., 2011). The order *Betaproteobacteriales* exhibited a large relative abundance in these bacterial community studies. However, it is significantly enriched in the Sba than other *Stipa* taxa. Given the vulnerability of several species pertaining to *Betaproteobacteriales* to salinity, acidity, and grazing pressure, Sba achieves a better growth and living environment than other *Stipa* taxa. Moreover, Sba itself is easily replaced in heavily grazing and salinized areas. Several species of the order *Thermomicrobiales* (the family and genus

of *JG30_KF_CM45*) are characterized as thermophilic, neutral, and basophilic (*Bergey's Manual of Systematics of Archaea and Bacteria*). *Norank_f__JG30_KF_CM45* acts as a potential genus biomarker for the Sbr, with a significant positive correlation with MAT, pH, and AP, but a negative correlation with MAP, SOC, TN, and S (**Figure 2B**), while *Candidatus_Udaeobacter* (Sba) and *norank_f__Xanthobacteraceae* (Sgr) exhibit opposite behaviors (**Figure 2B**). The mentioned explains why the warm-loving, drought-resistant, alkali-resistant, and barren-resistant *norank_f__JG30_KF_CM45* is enriched, which may indicate available phosphorus transformation and utilization. The critical thing is that Sbr also likes warmth (Chen et al., 2020). *Candidatus_Udaeobacter* and *norank_f__Xanthobacteraceae* act as potential biomarkers for the Sba and Sgr, respectively, which might indicate vital ecosystem multifunctionality involved in SOC accumulation, S cycle, and N stock across the chronosequence (Li et al., 2021). The family *Xanthobacteraceae* (*Rhizobiales* and *norank_f__Xanthobacteraceae*) in the Sgr grows with aerobic chemoheterotrophs and fix nitrogen (Oren, 2014; Yu et al., 2019). The mentioned functional bacteria in root zone can directly indicate the soil nutrient transformation and the response of bacterial to environmental variations.

Climate Is the Main Driving Factor for the Turnover and Distribution of Bacterial Community in the Root Zone

Climate has been recognized as a vital driver of soil bacterial community structure in regional and large-scale studies (Zeng et al., 2019). Yao proved climate drives the differentiation of bulk soil bacterial communities in the eastern Inner Mongolia steppe (Yao et al., 2017). Likewise, temperature and precipitation are two major climatic factors affecting the *Stipa* species root-zone soil bacterial community structure and distribution in this paper. This is related to the influence of precipitation on plant community composition and soil nutrient availability. For instance, higher precipitation may upregulate soil organic matter decomposition rate, thereby reducing soil organic matter availability (Tian et al., 2017; Na et al., 2019), and increasing microbial metabolic rates and biochemical processes due to temperature (Zhou et al., 2016), which subsequently affects the soil bacterial community structure. Furthermore, the spatial pattern of root-zone soil bacterial communities was distributed along the hydrothermal combination gradient. The mentioned result was attributed to the long-term influence of the southeast ocean monsoon and the northeast–southwest arc mountain barriers (Greater *Khingan* Mountains and the *Yin* Mountains), thereby causing the climatic factors to form an arc belt distribution. Such a peculiar water–heat combination causes the zonation of soil and vegetation, which makes the differentiation of vegetation zones in this area roughly coincide with the distribution of the climatic zones (Han and Tian, 2016).

Plant and soil parameters were other crucial factors of the bacterial community structure and distribution. As indicated from existing studies, soil pH exerts a dominant effect on various ecosystems in shaping bacterial structures and large-scale spatial distributions (Fierer and Jackson, 2006; Fan et al., 2017), mainly in acidic and neutral soil environments (Yang et al., 2018). Different soil properties play a leading role in an alkaline environment (Wang et al., 2017); for example, the soil carbon content was the dominant factor of bacteria distribution in the *Ali* area (Chu et al., 2016). The result of this paper partially complies with the existing studies indicating that soil pH was the most crucial soil parameter, whereas SOC, SOM, TN, and NH_4^+-N significantly affected the structure and distribution of *Stipa* taxa root-zone soil bacterial communities. In this paper, chemical element (S, Fe, and Cu) also impacted the bacterial community structure, whereas it did not impact the bacterial community distribution except for S (**Figure 6** and **Supplementary Table S7**). Thus, the soil sulfur cycle plays a partial role in the ecological distribution of bacteria, whereas microbes in rhizosphere are critical to allowing plants to access soil organosulfur (Kertesz et al., 2007). This paper also reported that the plant coverage and biomass of *Stipa* had a significant impact on the distribution and structure of the bacterial community at the phylum and genus levels were particularly prominent (**Supplementary Figures S2, S6**); complying with Han's view (Han et al., 2018), while the species of *Stipa* significantly impacted the distribution of bacteria. This is because a higher plant biomass and differences in the plant cover are generally directly affecting the quality and quantity of organic matter input, with subsequent effects on bacterial communities (Fierer and Jackson, 2006). In arid systems, the changes in plant coverage are also especially important for nutrient and moisture maintenance (Yao et al., 2017). Furthermore, different vegetation types can determine litter composition, soil conditions, and resource acquisition strategies, thereby indirectly regulating the structure of the bacterial community through a superposition effect with soil factors (Yao et al., 2018). In contrast, the soil microbial community structure will also affect the natural selection pattern of plant traits, while regulating the response of plants to abiotic environmental pressure, thereby affecting the evolutionary process of the whole ecosystem (Berg and Smalla, 2009). Besides, high interspecific clustering of the bacterial root-zone communities was identified. Such an interspecific difference could be partially determined by the genetic diversity among host species (Yu and Hochholdinger, 2018).

On the whole, the effect of precipitation and *Stipa* species on the structure and distribution of the root-zone soil bacterial community exceeded that of soil pH on a large scale. It is noteworthy that the bacterial community composition in the root zone of *Stipa* taxa complied with the previous reports of bulk or rhizosphere soil in a semi-arid system (Wang et al., 2015; Nan et al., 2020), while the bacteria community structure was different. Many studies have confirmed that the rhizosphere microbiome consists of a subset of the bulk soil microbiome; besides, microbial community diversity decreased rapidly from bulk soil to rhizosphere (Reinhold-Hurek et al., 2015; Trivedi et al., 2020). During the assembly of microbial communities from bulk soil to rhizosphere, the bulk soil provides a microbial seed bank, the physical–chemical properties, the biogeography,

and climate conditions (Zhang et al., 2017). Based on reasons discussed previously, the differences observed in the root zone of *Stipa* also might be mediated by the bulk soil microbiota in their respective regions.

Soil pH and Altitude Have Strong Correlation With α Bacterial Community Diversity

The diversity, evenness, and richness of rhizosphere bacteria were inconsistent in the seven plant species. There was a sudden change in the root-zone soil bacterial diversity at Sbr in desert steppe. According to Dijkstra et al. (2012), bacterial diversity changed suddenly in arid areas, and soil microbes living in zones with scarce water availability might be easily activated by even small rainfall events, which did not reach a level that would satisfy the plants' needs. As reported from numerous studies, soil pH acts as a vital factor driving the diversity of rhizosphere soil bacterial community (Na et al., 2018; Wang et al., 2020). This paper further examined the previously found strong correlation between soil pH and bacterial root-zone community diversity (**Figure 6**), as the pH conditions affect the adaptation and selection of particular phylogenic groups (Wang et al., 2020). In turn, bacterial α diversity, as indicated by the Shannon and Chao indexes, significantly affects the nutrient composition of forage grasses, especially the accumulation of crude fat and crude protein, whereas the correlation coefficient is relatively small. In addition, the elevation is correlated with the variables (SOC, TN, TK, and pH) affecting the ecosystem and plays a key role in influencing root-zone bacterial diversity, evenness, and richness. Wang reported that microbial richness in the Qinghai–Tibet Plateau showed a significant negative correlation with altitude. The pattern of microbial diversity and altitude was inconsistent due to the differences in the primary factors regulating the bacterial diversity in soils from different regions (Wang et al., 2020). Furthermore, this paper reported that root-zone soil bacterial α diversity was positively correlated with altitude.

Bacterial Taxa Interaction and Community Assembly in *Stipa* Root-Zone Soil

Both deterministic and stochastic processes were governing the spatial distribution of microbial communities concurrently, and different processes were found to be dominant in various cases (Zhang et al., 2016). The root-zone bacterial community assembly was dominated by stochastic processes at *Stipa* taxa, which also complied with the results of several studies, i.e., the effects of stochastic processes were commonly dominant at larger geographic scales (Li et al., 2021). Previous research has shown that the spatio-temporal distribution of soil bacteria in a temperate grassland was dominated by stochastic dispersal (Richter-Heitmann et al., 2020). According to other studies, the variation of microbial community structure will be dominated by different assembly factors in different periods (Zhou et al., 2014; Liu et al., 2020). The results of this paper may only represent the assembly of *Stipa* rhizosphere microorganisms in such a period, which can be exploited to determine the stability of different *Stipa* rhizosphere bacterial communities of the period (Zhou and Ning, 2017). The significant variations of drought span and soil structure in the study area may impose the natural limitation on the undirected dispersal of microbes like others have reported (Li and Hu, 2021), which also explains why the relative effects of dispersal limitation of *Stipa* genus root-zone soil bacteria showed the absolute advantage in the study. As expected, the relative effects of dispersal limitation and undominated processes increased significantly across the transition from Sba to Skr, while the relative role of homogenizing dispersal decreased significantly (**Supplementary Figure S3B**), thereby demonstrating that the bacterial community structure tended to be increasingly stable, and environmental variations exerted fewer effects (Zhou and Ning, 2017). Homogeneous selection was the main process driving the assembly of Sbr, Sgl, and Skl bacterial communities, showing that community structure would change largely with environmental conditions (Cheng et al., 2021).

This paper reported that the potential correlation patterns of different *Stipa* taxa root-zone bacterial groups varied substantially for network complexity and modularity. Moreover, the M, GD, and avgCC of the empirical networks were clearly higher than those of the respective random networks (**Supplementary Table S8**), suggesting that empirical networks had a noticeable hierarchy and modularity in their topological properties. Some studies have reported that the interactions of root-associated microbes are more complex than those of microbes in bulk soil, which may be related to rhizosphere microbial diversity, and the main interactions are positive (>80%), which also indicates that rhizosphere may have greater potential for cooperative or mutualistic associations (Coyte et al., 2015; Fan et al., 2018), and so are the root zones of *Stipa*, because *four* networks had a large proportion of members that are connected through positive links (**Supplementary Table S8**). However, such bacterial community is considered unstable since bacterial taxa may make a coincident response to environmental fluctuations, thereby inducing positive feedback and co-oscillation (Coyte et al., 2015). According to the bacterial co-occurrence pattern characterized by a high average degree, an average clustering coefficient, and fewer modules and modularity at extremely arid areas indicated that the networks of Sbr root-zone soil exhibit a more complex bacterial community structure with closer and better-connected nodes, whereas such topological features are of low stability (De Vries et al., 2018). Nevertheless, the Skr network was simplest but most stable based on topological characteristics. This network exhibited stronger resistance to environmental variations. Skr is easy to replace other *Stipa* taxa when the growth environment is stressed, thereby also indicating that the population can strongly adapt to the environment. This also reveals that there may be potential co-evolution relationship between vegetation and their root-zone bacteria. In summary, network analysis and RCbray results exert a good indication and prediction effect on environmental interference. More network complexity and fewer modules were reported in the severe environment than in the mild environment, thereby demonstrating that network complexity of bacteria was facilitated by various

environmental factors, whereas the clustering of modules was not.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

XM: investigation, data curation, methodology, visualization, writing—original draft, writing—review, and editing. LC, SW, and JL: investigation. ZD and JL: data curation. FL: field investigation. YB: funding acquisition, supervision, resources, review, and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (31760005) and Science and Technology Major Project of Inner Mongolia Autonomous Region of China (Grant No. zdzx2018065).

ACKNOWLEDGMENTS

We thank the reviewers who helped improve the article with constructive comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.782621/full#supplementary-material

REFERENCES

Bai, Y. F., Xu, Z. X., Li, D. X., and Zhao, G. (1999). Study on age and bunch structure of four stipa species in inner mongolia plateau. *Acta Bot. Sin.* 41, 1125–1131.

Basirat, M., Mousavi, S. M., Abbaszadeh, S., Ebrahimi, M., and Zarebanadkouki, M. (2019). The rhizosheath: a potential root trait helping plants to tolerate drought stress. *Plant Soil* 445, 565–575. doi: 10.1007/s11104-019-04334-0

Berg, G., and Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiol. Ecol.* 68, 1–13. doi: 10.1111/j.1574-6941.2009.00654.x

Chen, L., Saixi, Y., Yi, R., and Baoyin, T. (2020). Characterization of soil microbes associated with a grazing-tolerant grass species, *Stipa breviflora*, in the Inner Mongolian desert steppe. *Ecol. Evol.* 10, 10607–10618. doi: 10.1002/ece3.6715

Chen, S. H. (1987). *Root System Types of Steppe Plants in Inner Mongolia.* Inner Mongolia: Inner Mongolia People's Publishing House. p. 42–50.

Chen, Z., Wu, W., Shao, X., Li, L., Guo, Y., and Ding, G. (2015). Shifts in abundance and diversity of soil ammonia-oxidizing bacteria and archaea associated with land restoration in a semi-arid ecosystem. *PLoS ONE* 10:e0132879. doi: 10.1371/journal.pone.0132879

Cheng, X., Yun, Y., Wang, H., Ma, L., Tian, W., Man, B., et al. (2021). Contrasting bacterial communities and their assembly processes in karst soils under different land use. *Sci. Total Environ.* 751:142263. doi: 10.1016/j.scitotenv.2020.142263

Chu, H., Gao, G. F., Ma, Y., Fan, K., and Delgado-Baquerizo, M. (2020). Soil microbial biogeography in a changing world: recent advances and future perspectives. *mSystems* 5:e00803-19.. doi: 10.1128/mSystems.00803-19

Chu, H., Sun, H., Tripathi, B. M., Adams, J. M., Huang, R., Zhang, Y., et al. (2016). Bacterial community dissimilarity between the surface and subsurface soils equals horizontal differences over several kilometers in the western Tibetan Plateau. *Environ. Microbiol.* 18, 1523–1533. doi: 10.1111/1462-2920.13236

Clairmont, L. K., Stevens, K. J., and Slawson, M. R. (2019). Site-specific differences in microbial community structure and function within the rhizosphere and rhizoplane of wetland plants is plant species dependent. *Rhizosphere* 9, 56–68. doi: 10.1016/j.rhisph.2018.11.006

Corneo, P. E., Pellegrini, A., Cappellin, L., Roncador, M., Chierici, M., Gessler, C., et al. (2013). Microbial community structure in vineyard soils across altitudinal gradients and in different seasons. *FEMS Microbiol. Ecol.* 84, 588–602. doi: 10.1111/1574-6941.12087

Coyte, K. Z., Schluter, J., and Foster, R. K. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* 350, 663–666. doi: 10.1126/science.aad2602

De Vries, F. T., Griffiths, R. I., Bailey, M., Craig, H., Girlanda, M., Gweon, H. S., et al. (2018). Soil bacterial networks are less stable under drought than fungal networks. *Nat. Commun.* 9:3033. doi: 10.1038/s41467-018-05516-7

Deng, Y., Jiang, Y. H., Yang, Y. F., He, Z. L., Luo, F., and Zhou, Z. J. (2012). Molecular ecological network analyses. *BioMed. Central* 13:113. doi: 10.1186/1471-2105-13-113

Dijkstra, F. A., Augustine, D. J., Brewer, P., and von Fischer, C. J. (2012). Nitrogen cycling and water pulses in semiarid grasslands: are microbial and plant processes temporally asynchronous? *Oecologia* 170, 799–808. doi: 10.1007/s00442-012-2336-6

Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., and Salles, F. J. (2015). Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc. Natl. Acad. Sci. U.S.A.* 112, E1326–E1332. doi: 10.1073/pnas.1414261112

Edwards, J., Johnson, C., Santos-Medellin, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112, E911–E920. doi: 10.1073/pnas.1414592112

Fan, K., Cardona, C., Li, Y., Shi, Y., Xiang, X., Shen, C., et al. (2017). Rhizosphere-associated bacterial network structure and spatial distribution differ significantly from bulk soil in wheat crop fields. *Soil Biol. Biochem.* 113, 275–284. doi: 10.1016/j.soilbio.2017.06.020

Fan, K., Weisenhorn, P., Gilbert, J. A., and Chu, H. (2018). Wheat rhizosphere harbors a less complex and more stable microbial co-occurrence pattern than bulk soil. *Soil Biol. Biochem.* 125, 251–260. doi: 10.1016/j.soilbio.2018.07.022

Fan, X., and Xing, P. (2016). The vertical distribution of sediment archaeal community in the "black bloom" disturbing Zhushan Bay of Lake Taihu. *Archaea* 2016:8232135. doi: 10.1155/2016/8232135

Fierer, N., and Jackson, B. R. (2006). The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U.S.A.* 103, 626–631. doi: 10.1073/pnas.0507535103

Gao, S. B., Mo, L. D., Zhang, L. H., Zhang, J. L., Wu, J. B., Wang, J. L., et al. (2018). Phenotypic plasticity vs. local adaptation in quantitative traits differences of Stipa grandis in semi-arid steppe, China. *Sci. Rep.* 8:3148. doi: 10.1038/s41598-018-21557-w

Glaeser, S. P., and Kämpfer, P. (2014). "The family sphingomonadaceae," in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria,* eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin: Springer Berlin Heidelberg), 641–707.

Glasl, B., Bourne, D. G., Frade, P. R., Thomas, T., Schaffelke, B., and Webster, S. N. (2019). Microbial indicators of environmental perturbations in coral reef ecosystems. *Microbiome* 7:94. doi: 10.1186/s40168-019-0705-7

Han, B., and Tian, S. Q. (2016). *Ecological Adaptability of Main Stipa Plants in Inner Mongolia*. Beijing: China Agricultural Science and Technology Press.

Han, D., Wang, N., Sun, X., Hu, Y., and Feng, F. (2018). Biogeographical distribution of bacterial communities in Changbai Mountain, Northeast China. *Microbiol. Open* 7:e00529. doi: 10.1002/mbo3.529

Hug, L. A., Castelle, C. J., Wrighton, K. C., Thomas, B. C., Sharon, I., Frischkorn, K. R., et al. (2013). Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome* 1:22. doi: 10.1186/2049-2618-1-22

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* 72, 1719–1728. doi: 10.1128/AEM.72.3.1719-1728.2006

Jiao, S., and Lu, Y. (2020). Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ. Microbiol.* 22, 1052–1065. doi: 10.1111/1462-2920.14815

Kang, L., Han, X., Zhang, Z., and Sun, J. O. (2007). Grassland ecosystems in China: review of current knowledge and research advancement. *Philos. Trans. R. Soc. Lond,. B,. Biol. Sci.* 362, 997–1008. doi: 10.1098/rstb.2007.2029

Kant, R., van Passel, M. W., Palva, A., Lucas, S., Lapidus, A., Glavina del Rio, T., et al. (2011). Genome sequence of Chthoniobacter flavus Ellin428, an aerobic heterotrophic soil bacterium. *J. Bacteriol.* 193, 2902–2903. doi: 10.1128/JB.00295-11

Kertesz, M. A., Fellows, E., and Schmalenberger, A. (2007). Rhizobacteria and plant sulfur supply. *Adv. Appl. Microbiol.* 62, 235–268. doi: 10.1016/S0065-2164(07)62008-5

Kourilova, X., Schwarzerova, J., Pernicova, I., Sedlar, K., Mrazova, K., Krzyzanek, V., et al. (2021). The first insight into polyhydroxyalkanoates accumulation in multi-extremophilic *Rubrobacter xylanophilus* and *Rubrobacter spartanus*. *Microorganisms* 9:909. doi: 10.3390/microorganisms9050909

Krishna, M., Gupta, S., Delgado-Baquerizo, M., Morrien, E., Garkoti, S. C., Chaturvedi, R., et al. (2020). Successional trajectory of bacterial communities in soil are shaped by plant-driven changes during secondary succession. *Sci. Rep.* 10:9864. doi: 10.1038/s41598-020-66638-x

Kuczynski, J., Stombaugh, J., Walters, W. A., Gonzalez, A., Caporaso, J. G., and Knight, R. (2011). Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Curr. Protoc. Bioinformatics* Chapter 10:Unit 107. doi: 10.1002/0471250953.bi1007s36

Lakshmanan, V., Selvaraj, G., and Bais, P. H. (2014). Functional soil microbiome: belowground solutions to an aboveground problem. *Plant Physiol.* 166, 689–700. doi: 10.1104/pp.114.245811

Lau, J. A., and Lennon, T. J. (2011). Evolutionary ecology of plant-microbe interactions: soil microbial structure alters selection on plant traits. *New Phytol.* 192, 215–224. doi: 10.1111/j.1469-8137.2011.03790.x

Li, M., Mi, T., He, H., Chen, Y., Zhen, Y., and Yu, Z. (2021). Active bacterial and archaeal communities in coastal sediments: biogeography pattern, assembly process and co-occurrence relationship. *Sci. Total Environ.* 750:142252. doi: 10.1016/j.scitotenv.2020.142252

Li, X., Wang, J., Zhang, S., Wang, H., Li, X., Li, X., et al. (2018). Distribution of fungal endophytes in roots of Stipa krylovii across six vegetation types in grassland of northern China. *Fungal Ecol.* 31, 47–53. doi: 10.1016/j.funeco.2017.11.001

Li, Y., and Hu, C. (2021). Biogeographical patterns and mechanisms of microbial community assembly that underlie successional biocrusts across northern China. *NPJ Biofilms Microbiomes* 7:15. doi: 10.1038/s41522-021-00188-6

Liu, C., Wentworth, T. R., Qiao, X., Guo, K., and Hou, D. (2019). Vegetation classification at the association level under the China Vegetation Classification System: an example of six Stipa steppe formations in China. *J. Plant Ecol.* 12, 1009–1024. doi: 10.1093/jpe/rtz028

Liu, M., Li, X., Zhu, R., Chen, N., Ding, L., and Chen, C. (2021). Vegetation richness, species identity and soil nutrients drive the shifts in soil bacterial communities during restoration process. *Environ. Microbiol. Rep.* 13, 411–424. doi: 10.1111/1758-2229.12913

Liu, W., Graham, E. B., Zhong, L., Zhang, J., Li, S., Lin, X., et al. (2020). Long-term stochasticity combines with short-term variability in assembly processes to underlie rice paddy sustainability. *Front. Microbiol.* 11:873. doi: 10.3389/fmicb.2020.00873

Lv, X., Zhou, G., Wang, Y., and Song, X. (2016). Sensitive indicators of zonal stipa species to changing temperature and precipitation in Inner Mongolia Grassland, China. *Front. Plant Sci.* 7:73. doi: 10.3389/fpls.2016.00073

Maestre, F. T., Delgado-Baquerizo, M., Jeffries, T. C., Eldridge, D. J., Ochoa, V., Gozalo, B., et al. (2015). Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15684–15689. doi: 10.1073/pnas.1516684112

Marasco, R., Mosqueira, M. J., Fusi, M., Ramond, J. B., Merlino, G., Booth, J. M., et al. (2018). Rhizosheath microbial community assembly of sympatric desert speargrasses is independent of the plant host. *Microbiome* 6:215. doi: 10.1186/s40168-018-0597-y

Mohanram, S., and Kumar, P. (2019). Rhizosphere microbiome: revisiting the synergy of plant-microbe interactions. *Ann. Microbiol.* 69, 307–320. doi: 10.1007/s13213-019-01448-9

Na, X., Xu, T., Li, M., Zhou, Z., Ma, S., Wang, J., et al. (2018). Variations of bacterial community diversity within the rhizosphere of three phylogenetically related perennial shrub plant species across environmental gradients. *Front. Microbiol.* 9:709. doi: 10.3389/fmicb.2018.00709

Na, X., Yu, H., Wang, P., Zhu, W., Niu, Y., and Huang, J. (2019). Vegetation biomass and soil moisture coregulate bacterial community succession under altered precipitation regimes in a desert steppe in northwestern China. *Soil Biol. Biochem.* 136:107520. doi: 10.1016/j.soilbio.2019.107520

Nan, J., Chao, L., Ma, X., Xu, D., Mo, L., Zhang, X., et al. (2020). Microbial diversity in the rhizosphere soils of three Stipa species from the eastern Inner Mongolian grasslands. *Global Ecol. Conserv.* 22:e00992. doi: 10.1016/j.gecco.2020.e00992

Oksanen, J., Blanchet, F. G., Kindt, R., and Minchin, R. P. (2015). *Vegan: Community Ecology Package*. R Package Version 22–1.

Oren, A. (2014). "The family xanthobacteraceae," in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria,* eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin: Springer Berlin Heidelberg), 709–726.

Perez-Jaramillo, J. E., de Hollander, M., Ramirez, C. A., Mendes, R., Raaijmakers, J. M., and Carrion, J. V. (2019). Deciphering rhizosphere microbiome assembly of wild and modern common bean (*Phaseolus vulgaris*) in native and agricultural soils from Colombia. *Microbiome* 7:114. doi: 10.1186/s40168-019-0727-1

Philippot, L., Raaijmakers, J. M., Lemanceau, P., and van der Putten, H. W. (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.* 11, 789–799. doi: 10.1038/nrmicro3109

Reinhold-Hurek, B., Bunger, W., Burbano, C. S., Sabale, M., and Hurek, T. (2015). Roots shaping their microbiome: global hotspots for microbial activity. *Annu. Rev. Phytopathol.* 53, 403–424. doi: 10.1146/annurev-phyto-082712-102342

Richter-Heitmann, T., Hofner, B., Krah, F.-S., Sikorski, J., Wüst, P. K., Bunk, B., et al. (2020). Stochastic dispersal rather than deterministic selection explains the spatio-temporal distribution of soil bacteria in a temperate grassland. *Front. Microbiol.* 11:1391. doi: 10.3389/fmicb.2020.01391

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Shen, C., Xiong, J., Zhang, H., Feng, Y., Lin, X., Li, X., et al. (2013). Soil pH drives the spatial distribution of bacterial communities along elevation on Changbai Mountain. *Soil Biol. Biochem.* 57, 204–211. doi: 10.1016/j.soilbio.2012.07.013

Shi, W., Li, M., Wei, G., Tian, R., Li, C., Wang, B., et al. (2019). The occurrence of potato common scab correlates with the community composition and function of the geocaulosphere soil microbiome. *Microbiome* 7:14. doi: 10.1186/s40168-019-0629-2

Song, Z., Liu, G., and Zhang, C. (2019). Response of rhizosphere microbial communities to plant succession along a grassland chronosequence in a semiarid area. *J. Soils Sediments* 19, 2496–2508. doi: 10.1007/s11368-019-02241-6

Stegen, J. C., Freestone, A. L., Crist, T. O., Anderson, M. J., Chase, J. M., Comita, L. S., et al. (2013). Stochastic and deterministic drivers of spatial and temporal turnover in breeding bird communities. *Global Ecol. Biogeogr.* 22, 202–212. doi: 10.1111/j.1466-8238.2012.00780.x

Tang, H., Zhang, N., Ni, H., Xu, X., Wang, X., Sui, Y., et al. (2021). Increasing environmental filtering of diazotrophic communities with a decade of latitudinal soil transplantation. *Soil Biol. Biochem.* 154:108119. doi: 10.1016/j.soilbio.2020.108119

Tian, J., He, N., Hale, L., Niu, S., Yu, G., Liu, Y., et al. (2017). Soil organic matter availability and climate drive latitudinal patterns in bacterial

diversity from tropical to cold temperate forests. *Funct. Ecol.* 32, 61–70. doi: 10.1111/1365-2435.12952

Tripathi, B. M., Stegen, J. C., Kim, M., Dong, K., Adams, J. M., and Lee, K. Y. (2018). Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *ISME J.* 12, 1072–1083. doi: 10.1038/s41396-018-0082-4

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., and Singh, K. B. (2020). Plant-microbiome interactions: from community assembly to plant health. *Nat. Rev. Microbiol.* 18, 607–621. doi: 10.1038/s41579-020-0412-1

Tu, B., Domene, X., Yao, M., Li, C., Zhang, S., Kou, Y., et al. (2017). Microbial diversity in Chinese temperate steppe: unveiling the most influential environmental drivers. *FEMS Microbiol. Ecol.* 93. doi: 10.1093/femsec/fix031

Verma, M., Mishra, J., and Arora, N. K. (2019). "Plant growth-promoting rhizobacteria: diversity and applications," in *Environmental Biotechnology: For Sustainable Future*, eds R. A. B. Khalid, R. Hakeem, and m. A. Dervash (Cham: Springer), 129–173.

Victoria, R., Banwart, S. A., Black, H., Ingram, J., Joosten, H., Milne, E., et al. (2012). *Benefits of Soil Carbon: Foresight Chapter in UNEP Yearbook 2012*. (Philadelphia, PA: United Nations Environment Programme), 19–33.

Wang, S. K., Zuo, X. A., Awada, T., Medima-Roldán, E., Feng, K. T., Yue, P., et al. (2021). Changes of soil bacterial and fungal community structure along a natural aridity gradient in desert grassland ecosystems, Inner Mongolia. *Catena* 205:105470. doi: 10.1016/j.catena.2021.105470

Wang, X., Van Nostrand, J. D., Deng, Y., Lu, X., Wang, C., Zhou, J., et al. (2015). Scale-dependent effects of climate and geographic distance on bacterial diversity patterns across northern China's grasslands. *FEMS Microbiol. Ecol.* 91:fiv133. doi: 10.1093/femsec/fiv133

Wang, X., Zhang, Z., Yu, Z., Shen, G., Cheng, H., and Tao, S. (2020). Composition and diversity of soil microbial communities in the alpine wetland and alpine forest ecosystems on the Tibetan Plateau. *Sci. Total Environ.* 747:141358. doi: 10.1016/j.scitotenv.2020.141358

Wang, X. B., Lu, X. T., Yao, J., Wang, Z. W., Deng, Y., Cheng, W. X., et al. (2017). Habitat-specific patterns and drivers of bacterial beta-diversity in China's drylands. *ISME J.* 11, 1345–1358. doi: 10.1038/ismej.2017.11

Yang, F., Wu, J., Zhang, D., Chen, Q., Zhang, Q., and Cheng, X. (2018). Soil bacterial community composition and diversity in relation to edaphic properties and plant traits in grasslands of southern China. *Appl. Soil Ecol.* 128, 43–53. doi: 10.1016/j.apsoil.2018.04.001

Yao, M., Rui, J., Niu, H., Heděnec, P., Li, J., He, Z., et al. (2017). The differentiation of soil bacterial communities along a precipitation and temperature gradient in the eastern Inner Mongolia steppe. *Catena* 152, 47–56. doi: 10.1016/j.catena.2017.01.007

Yao, X., Zhang, N., Zeng, H., and Wang, W. (2018). Effects of soil depth and plant-soil interaction on microbial community in temperate grasslands of northern China. *Sci. Total Environ.* 630, 96–102. doi: 10.1016/j.scitotenv.2018.02.155

Yu, P., and Hochholdinger, F. (2018). The role of host genetic signatures on root-microbe interactions in the rhizosphere and endosphere. *Front. Plant Sci.* 9:1896. doi: 10.3389/fpls.2018.01896

Yu, Z., Jin, J., Li, Y., Yang, Y., Zhao, Y., Liu, C., et al. (2019). Distinct effects of short-term reconstructed topsoil on soya bean and corn rhizosphere bacterial abundance and communities in Chinese Mollisol. *R. Soc. Open Sci.* 6:181054. doi: 10.1098/rsos.181054

Zeng, Q., An, S., Liu, Y., Wang, H., and Wang, Y. (2019). Biogeography and the driving factors affecting forest soil bacteria in an arid area. *Sci. Total Environ.* 680, 124–131. doi: 10.1016/j.scitotenv.2019.04.184

Zhalnina, K., Louie, K. B., Hao, Z., Mansoori, N., da Rocha, U. N., Shi, S., et al. (2018). Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. *Nat. Microbiol.* 3, 470–480. doi: 10.1038/s41564-018-0129-3

Zhang, B., Zhang, J., Liu, Y., Shi, P., and Wei, G. (2018). Co-occurrence patterns of soybean rhizosphere microbiome at a continental scale. *Soil Biol. Biochem.* 118, 178–186. doi: 10.1016/j.soilbio.2017.12.011

Zhang, R., Vivanco, J. M., and Shen, Q. (2017). The unseen rhizosphere root-soil-microbe interactions for crop production. *Curr. Opin. Microbiol.* 37, 8–14. doi: 10.1016/j.mib.2017.03.008

Zhang, X., Johnston, E. R., Liu, W., Li, L., and Han, X. (2016). Environmental changes affect the assembly of soil bacterial community primarily by mediating stochastic processes. *Glob. Chang. Biol.* 22, 198–207. doi: 10.1111/gcb.13080

Zhang, Y., Loreau, M., He, N., Wang, J., Pan, Q., Bai, Y., et al. (2018). Climate variability decreases species richness and community stability in a temperate grassland. *Oecologia* 188, 183–192. doi: 10.1007/s00442-018-4208-1

Zhou, H., Gao, Y., Jia, X., Wang, M., Ding, J., Cheng, L., et al. (2020). Network analysis reveals the strengthening of microbial interaction in biological soil crust development in the Mu Us Sandy Land, Northwestern China. *Soil Biol. Biochem.* 144:107782. doi: 10.1016/j.soilbio.2020.107782

Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D., et al. (2016). Temperature mediates continental-scale diversity of microbes in forest soils. *Nat. Commun.* 7:12083. doi: 10.1038/ncomms12083

Zhou, J., Deng, Y., Zhang, P., Xue, K., Liang, Y., Van Nostrand, J. D., et al. (2014). Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. *Proc. Natl. Acad. Sci. U.S.A.* 111, E836–E845. doi: 10.1073/pnas.1324044111

Zhou, J. Z., and Ning, L. D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81:4. doi: 10.1128/MMBR.00002-17

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 26 January 2022
doi: 10.3389/fmicb.2022.809074

Distinct Co-occurrence Relationships and Assembly Processes of Active Methane-Oxidizing Bacterial Communities Between Paddy and Natural Wetlands of Northeast China

*Xu Liu[1,2], Yu Shi[3], Teng Yang[1,2], Gui-Feng Gao[1,2], Liyan Zhang[4], Ruoyu Xu[5], Chenxin Li[5], Ruiyang Liu[5], Junjie Liu[6] and Haiyan Chu[1,2]**

[1]State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng, China, [4]Key Laboratory of Integrated Regulation and Resource Development on Shallow Lake of Ministry of Education, College of Environment, Hohai University, Nanjing, China, [5]High School Affiliated to Nanjing Normal University, Nanjing, China, [6]Key Laboratory of Mollisols Agroecology, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin, China

Studies of methane-oxidizing bacteria are updating our views of their composition and function in paddy and natural wetlands. However, few studies have characterized differences in the methane-oxidizing bacterial communities between paddy and natural wetlands. Here, we conducted a ^{13}C stable isotope-probing experiment and high-throughput sequencing to determine the structure profiling, co-occurrence relationships, and assembly processes of methanotrophic communities in four wetlands of Northeast China. There was a clear difference in community structure between paddy and natural wetlands. LEfSe analyses revealed that *Methylobacter*, *FWs*, and *Methylosinus* were enriched in natural wetlands, while *Methylosarcina* were prevailing in paddy, all identified as indicative methanotrophs. We observed distinct co-occurrence relationships between paddy and natural wetlands: more robust and complex connections in natural wetlands than paddy wetlands. Furthermore, the relative importance of stochastic processes was greater than that of deterministic processes, as stochastic processes explained >50% of the variation in communities. These results demonstrated that the co-occurrence relationships and assembly processes of active methanotrophic communities in paddy and natural wetlands were distinct. Overall, the results of this study enhance our understanding of the communities of methane-oxidizing bacteria in paddy and natural wetlands of Northeast China.

Keywords: CH_4 oxidation, DNA-SIP, methane-oxidizing bacterial communities, co-occurrence relationships, assembly processes

OPEN ACCESS

Edited by:
Jianjun Wang, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences (CAS), China

Reviewed by:
Qin Yao, Heilongjiang Bayi Agricultural University, China
Yu Luo, Zhejiang University, China
Jingjing Peng, China Agricultural University, China

****Correspondence:***
Haiyan Chu
hychu@issas.ac.cn

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *04 November 2021*
Accepted: *04 January 2022*
Published: *26 January 2022*

Citation:
Liu X, Shi Y, Yang T, Gao G-F, Zhang L, Xu R, Li C, Liu R, Liu J and Chu H (2022) Distinct Co-occurrence Relationships and Assembly Processes of Active Methane-Oxidizing Bacterial Communities Between Paddy and Natural Wetlands of Northeast China. Front. Microbiol. 13:809074. doi: 10.3389/fmicb.2022.809074

INTRODUCTION

Wetland ecosystem is one of the largest terrestrial sources of methane (CH_4) emissions. Paddy wetlands account for approximately 24 to 40 Tg CH_4 yr^{-1} and natural wetlands account for 100 to 183 Tg CH_4 yr^{-1} (Saunois et al., 2020), totally contributing about one-third of global CH_4 emissions (Bousquet et al., 2006). More than 80% of CH_4 produced is oxidized by methane-oxidizing bacteria (MOB or methanotrophs) before being released into the atmosphere (Mer and Roger, 2001). Wetlands soils are dynamic systems characterized by high methanotrophic activity at the aerobic–anaerobic interface; microbial guilds involved in CH_4 consumption in such soils in paddy and natural wetlands have received increased interest from researchers (Chowdhury and Dick, 2013; Kirschke et al., 2013). For example, Shiau et al. (2018) found that the *Methylocystis*-affiliated type II genotype is the predominant methanotroph in paddy fields at the regional scale of ~400 km. Yun et al. (2012) suggested that CH_4 consumption in Zoige natural wetland was driven by both type I and type II methanotrophs. Although several community studies of methanotrophs have been conducted, comparative studies of the structure of methanotrophic communities in paddy and natural wetlands are lacking.

Biotic interactions play a critical role in shaping the co-occurrence patterns of microbial guilds, and different co-occurrence patterns of methanotrophic communities are supposed to be correlated with essential ecological processes, such as polymer breakdown, syntrophy, and fermentation (Peng et al., 2018; Zhang et al., 2020). Microbial guilds compete with syntrophic partners to obtain carbon and ATP (Conrad, 2009). These partnerships are shaped by metabolic interactions as well as the degree of habitat niche overlap (Faust and Raes, 2012). However, detailed information on interspecific and intraspecific associations in this field based on empirical laboratory studies is difficult to obtain, especially for complex and diverse microbial communities. Inferring co-occurrence networks based on ecological data can provide key insights into co-occurrence relationships, the links among communities, and the functional potential of communities (Nemergut et al., 2013; Röttjers and Faust, 2018; Yang, 2021). For example, Li et al. (2021) explored the coexistence patterns of soil methanogens and revealed that the complex interactions are closely tied to CH_4 generation. Co-occurrence relationships are distinct from ecological interactions, and exploration of the former could provide new insights into the structure and functional potential of methanotrophic communities in paddy and natural wetlands.

The role of assembly mechanisms in determining the structure and function of methanotrophic communities has been a major focus of previous studies (McCalley et al., 2014; Hines, 2019). Comparison of intra- and inter-group phylogenetic distances with null models has provided new insights into assembly mechanisms, and this work has emphasized the importance of obtaining phylogenetic information for describing and predicting the structure and function of microbial communities (Stegen et al., 2013; Ning et al., 2020). Phylogenetic clustering and overdispersion compared with null models indicate the emphatic importance of deterministic processes in shaping community structure, representing variable and homogeneous selection, respectively (Emerson and Gillespie, 2008; Vellend, 2010; Weiher et al., 2011). Phylogenetic clustering could lead to the analogy of functional genotypes, whereas phylogenetic overdispersion indicates that diverse functional guilds are phylogenetically more distantly related to each other (Wang et al., 2012; Mittelbach and Schemske, 2015). For example, Fan et al. (2019) found that long-term fertilization shifts nitrogen fixers from being phylogenetically clustered to being phylogenetically overdispersed, which leads to the emergence of more phylogenetically diverse diazotrophic communities. In addition to deterministic processes, community assembly is mediated by stochastic processes, such as ecological drift and dispersal (Stegen et al., 2015). Stochastic processes play an important role in driving microbial community patterns in multiple ecosystems (Jiao et al., 2019). Generally, study of the assembly processes of methanotrophs can enhance our understanding of MOB communities and their functional potentials in paddy and natural wetlands.

Stable isotope probing (SIP) has been used in many ecological studies and is effective for studying the metabolic activities of methanotrophs (Dumont et al., 2011; Ho et al., 2013). A direct link between CH_4-uptake activity and methanotrophic taxa has been established using SIP and analysis of DNA markers specific to methanotrophs by feeding them with ^{13}C-labeled CH_4 (He et al., 2019, 2021; Sultana et al., 2019; Zhang et al., 2020). Here, we conducted a DNA-SIP experiment combined with high-throughput sequencing and multiple bioinformatics methods in four typical wetlands of Northeast China to study differences in methanotrophic communities between paddy and natural wetlands. Specifically, the aims of this study were to (i) identify differentially abundant and indicative methanotrophs in paddy and natural wetlands; (ii) characterize differences in the co-occurrence relationships and assembly processes of methanotrophic communities in paddy and natural wetlands; and (iii) explore the links between methanotrophic communities and functional potentials. We hypothesized that not only would the community compositions of active methanotrophs be distinct between paddy and natural wetlands, the co-occurrence relationships and assembly processes would be also different between paddy and natural wetlands. Overall, the results of our study provide new insights that could be used to aid the management of wetland ecosystems and enhance CH_4 consumption.

MATERIALS AND METHODS

Study Site and Sampling

Northeast China has a low mean annual temperature and is rich in soil carbon stocks; thus, the potential of carbon turnover is enormous (Tang et al., 2018). Paddy and natural wetlands are widespread in this region, accounting for ~16% of the total wetland area in China (Ding and Cai, 2007). Natural wetland soils were collected from two Chinese national wetland parks in Northeast China: Zhalong (ZL) and Xianghai (XH). Paddy

wetland soils were collected after the rice harvest in Changchun (CC) and Minzhu (MZ) in Northeast China. In each wetland, at least five soil cores were collected from depths of 0 to 20 cm. The cores were taken approximately 100 m apart and mixed to form one soil sample for each wetland. Visible roots and residues were eliminated from the soil samples. Samples were sieved through 2-mm mesh after natural withering and stored at 4°C before microcosm experiments. Detailed information on the study sites and sampling is provided in **Figure 1** (wetland locations and sample collection) and **Supplementary Table S1**.

DNA-SIP Microcosms

The four soil samples were treated with ^{12}C-labeled and ^{13}C-labeled CH_4 in triplicate in microcosms (4 sites × 2 treatments × 3 replicates = 24 total experimental units; **Figure 1**). For each microcosm, 5.0 g of soil (dry weight) was incubated at approximately 60% of the maximum water-holding capacity and 28°C in the dark in a 120-ml serum bottle sealed with a butyl stopper to determine the depletion of CH_4 in soils (Jia et al., 2019). For all incubations, 5.0 ml of the headspace air in the bottles was replaced with the same amount of $^{12}CH_4$ and $^{13}CH_4$ gas, which resulted in an initial CH_4 mixing ratio of approximately 40,000 ppmv in the headspace (Sultana et al., 2019). The CH_4 gas was >99% ^{13}C-atom pure. Gas chromatography was used to quantify the headspace CH_4 concentrations at two-day regular intervals (Shimadzu GC12-A, Japan). When more than 90% of CH_4 in the soil was consumed, soil samples were collected and separated into two subsamples for DNA

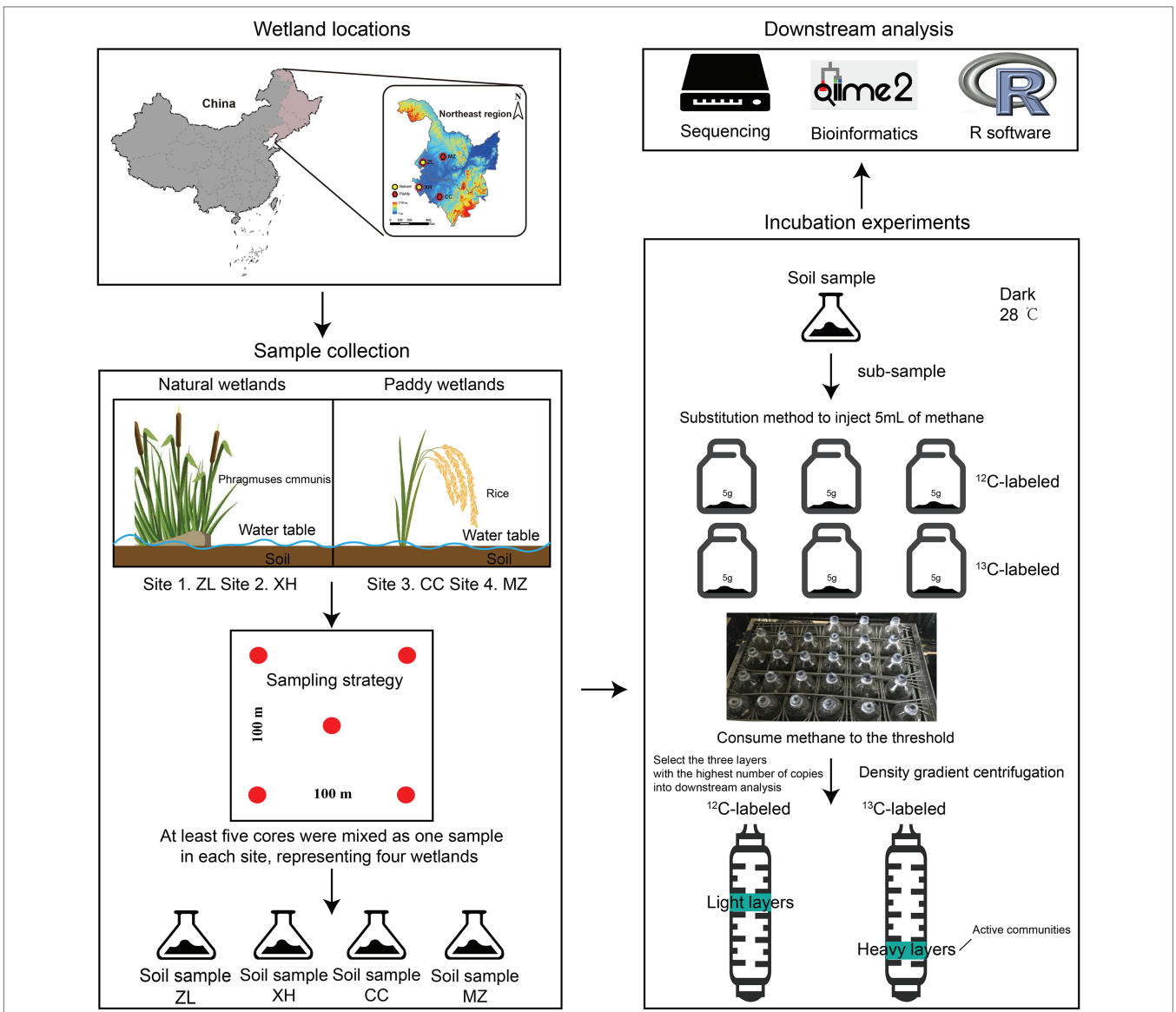


FIGURE 1 | A pipeline of the study on the differences in active methanotrophic communities between paddy and natural wetlands. CC on behalf of Changchun paddy wetland, MZ on behalf of Minzhu paddy wetland, XH on behalf of Xianghai natural wetland, and ZL on behalf of Zhalong natural wetland.

extraction and long-term preservation. Methane oxidation potential (MOP) was calculated using the formula $MOPs = \frac{(C1 - C2) * 120 * 0.001 * 16 * 301}{22.4 * 273 * T}$, where $C1$ and $C2$ represent the CH_4 concentration at the finishing time and starting time, respectively, and T stands for the temperature of the experiment.

DNA Extraction and SIP Gradient Fractionation

For the molecular survey of methanotrophic communities, 0.5 g of incubated soil of each microcosm was used for total DNA extraction with a FastDNA® Spin Kit (MP Biomedicals, Santa Ana, CA) per the manufacturer's instructions. The extracted DNA was dissolved in 70 μl of TE buffer. The quality and concentration of extracted DNA were quantified by NanoDrop1000 spectrophotometer (Thermo Fisher). When A260/A280 > 1.8 and A260/A230 > 2, the extracted DNA was considered of high quality and free from protein or RNA contamination. Besides, the samples were filtered with too low DNA concentration (<20 μg/ml). DNA samples that did not meet the above criteria were re-extracted. The qualified DNA was separated into two subsamples for high-speed buoyancy density centrifugation and long-term preservation.

To separate the different weights of labeled ($^{13}CH_4$) and unlabeled ($^{12}CH_4$) DNA, density gradient centrifugation of all extracted DNA was conducted. Soil DNA samples were collected from all CH_4-incubated microcosms (Lu and Jia, 2013). After adding CsCl solution into 5-ml sterile tubes with 2.0 μg of DNA, the mixed solution was adjusted with gradient buffer to a final density of 1.725 g/ml. The adjusted solution was then centrifuged in an ultracentrifuge (Beckman Coulter, Palo Alto, CA) at 177,000 g for 44 h at 20°C. DNA fractionations were carried out by the substitute method, wherein the gradient medium was displaced with sterile water from the top of the centrifuge tubes at a precisely controlled flow rate of 0.38 ml/min. Fifteen fractions per tube were collected and weighed on a 10,000 ppm balance. Polyethylene glycol 6000 (PEG6000) was used to precipitate the fractionated DNA from the calcium chloride medium, which was subsequently purified with 70% ethanol and dissolved in 30 μl of TE buffer for further amplification and sequencing.

Real-Time Quantitative PCR of *pmoA* Genes

Real-time quantitative PCR (qPCR) analysis was conducted using the CFX96 optical real-time detection system to determine the efficiency at which ^{13}C was incorporated into genomic DNA from methanotrophic communities (Bio-Rad, United States). The primers A189F (5′-GGNGACTGGGACTTCTGG-3′) and mb661r (5′-CCGGMGCAACGTCYTTACC-3′) were used to quantify the number of *pmoA* gene copies (Costello and Lidstrom, 1999). Reactions were performed in triplicate for each fractionated DNA sample. Reaction conditions were as follows: 95°C for 60 s, followed by 39 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s. Melting curve analysis was conducted by increasing the temperature from 65°C to 95°C in 0.1°C per second increments with continuous fluorescence acquisition. Individual standards were obtained from a 10-fold dilution series of plasmids containing a single *pmoA* gene fragment. Amplification efficiencies ranged from 91.6 to 96.7%, with R^2 values of 0.995–1.

Amplification and Sequencing of *pmoA* Gene Sequences

Three representative fractionations with a high copy number of the *pmoA* gene of each ^{13}C-labeled microcosm (heavy density of 1.7227–1.7422 g/ml) were selected according to the qPCR results. The A189F/mb661r primer set was used to amplify the *pmoA* genes, as it could retrieve the highest diversity of soil methanotrophs (Costello and Lidstrom, 1999). PCR conditions and procedures are commonly used to sequence *pmoA* functional genes (Zhang et al., 2020). Before sequencing, all PCR products were standardized to equimolar levels, and high-throughput sequencing was carried out using an Illumina MiSeq PE300 platform (Illumina, Inc., San Diego, CA, United States). All sequencing data in this study were deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession number PRJNA662020.

The raw data were analyzed using the QIIME pipeline (version 1.9.1; Caporaso et al., 2010). The raw reads were denoised to eliminate low-quality reads (length less than 200 bp or average quality score less than 29) to obtain high-quality sequences. The uchime3-denovo method in VSEARCH was used to detect chimeras (Rognes et al., 2016). Insertions and deletions causing the frameshifts in *pmoA* gene sequences were corrected using FrameBot[1] (Cole et al., 2014). The high-quality sequences were clustered into different operational taxonomic units (OTUs) by UCLUST (Edgar, 2010) based on a 93% similarity threshold of *pmoA* gene sequences (Degelmann et al., 2010). The PyNAST method was used to align representative sequences (DeSantis et al., 2006). The RDP-classifier (Cole et al., 2014) was used to determine the taxonomic identity of the OTUs, which was based on a database including 6,628 *pmoA* and *pmoA*-related sequences from pure culture methanotrophs and uncultured methanotrophic ecotypes (Dumont et al., 2014). After filtering low-quality reads and rarefying samples to equal sequencing depth, we obtained a total of approximately 115,754 quality-filtered, chimera-free, and frameshift-free sequences with an average of 1,172 methanotroph *pmoA* gene reads (the minimum number of sequences of all samples) for downstream analysis.

Co-occurrence Relationship Inference

Co-occurrence relationship inference was conducted at the OTU level *via* the SPIEC-EASI network method with the function *spiec.easi* in the spieceasi package (Kurtz et al., 2015). We focused on the dominant methanotrophs with filtered frequency (top 10% of relative abundance of active methanotrophs) and occurrence (>25% of all samples) across all wetland soils. Based on the filtered conditions, we constructed the total co-occurrence network involving 36 samples and then

[1]https://github.com/rdpstaff/fungene_pipeline

extracted the sub-network according to two types of wetlands. The generated network results were imported into Gephi software to determine the topological properties. Furthermore, we used within-module (z-score) and among-module (c-score) connectivity as a topological indicator to classify nodes as network hubs (z-score > 2.5; c-score > 0.6), module hubs (z-score > 2.5; c-score < 0.6), connectors (z-score < 2.5; c-score > 0.6), and peripherals (z-score < 2.5; c-score < 0.6; Shi et al., 2020). Module abundance in each network was calculated by averaging the standardized relative abundances (z-score) of the nodes that belonged to the specific modules. We used ForceAtlas 2 algorithms to display the visualization plot of the co-occurrence network in the Gephi platform.[2]

Quantification of Assembly Processes

Standardized mean pairwise distance (MPD) was calculated to determine the phylogenetic distance among methanotrophic communities using the picante package. This reveals the degree of dispersion of lineages within a community (Kembel et al., 2010). MPD values < −2 indicate phylogenetic clustering; MPD values > 2 indicate phylogenetic overdispersion; and MPD values between −2 and 2 indicate phylogenetic stochasticity.

The β-nearest taxon index (βNTI) and Bray–Curtis-based Raup–Crick metrics (RC_{bray}) were used to indicate the relative contributions of assembly processes, including deterministic processes (homogeneous selection and variable selection) and stochastic processes (homogenizing dispersal and dispersal limitation coupled with undominated processes). βNTI measures the deviation of observed β-mean nearest taxon distance (βMNTD) from the expected βMNTD in the null model. βNTI < −2 indicates homogeneous selection (HS), whereas βNTI > 2 indicates variable selection (VS). When $|\beta NTI| < 2$ and $RC_{bray} > 0.95$, dispersion limitation (DL) plays a dominant role. When $|\beta NTI| < 2$ and $RC_{bray} < -0.95$, homogenizing dispersal (HD) plays a dominant role. When $|\beta NTI| < 2$ and $|RC_{bray}| < 0.95$, no single process plays a dominant role (UD), and this is also known as drift (Stegen et al., 2013).

Statistical Analysis and Visualization

IBM SPSS Statistics 23.0 was used to perform analysis of variance, correlation analysis, and non-parametric difference tests (SPSS Inc., Cary, NC, United States). *Post hoc* Turkey's tests were conducted for multiple comparisons using the EasyStats package in R (version 3.6.1).[3] Differential abundance was determined using the DESeq2 package (Love et al., 2014). Principal coordinate analysis (PCoA) was conducted using the vegan package. Analysis of similarity (ANOSIM) in the vegan package was used to determine differences in community composition. Linear discriminant analysis effect size (LEfSe) analysis was used to identify the taxa that were differentially abundant and indicative in paddy and natural wetlands and was performed using the lefse package (Segata et al., 2011). Representative OTUs of methanotrophs were visualized using the iTOL website (Letunic and Bork, 2016).

Threshold indicator taxa analysis was conducted to identify indicator taxa using the TITAN2 package (Baker and King, 2010). This analysis uses standardized scores (Z-scores) to detect the MOPs exceeding thresholds based on their frequency and occurrence patterns. The z-scores were derived from normalizing indicator value scores (IndVals) with random permutations to determine the potential change. Responses of methanotrophs were standardized to the mean and standard deviation of permuted samples. Thus, the sum of the z-scores reflects the magnitude of potential change. TITAN differentiates taxa with positive (Z^+) and negative (Z^-) values; Z^+ taxa increase in frequency and abundance after the change point, and Z^- taxa show the opposite pattern.

RESULTS

Methane Oxidation Potential

DNA-SIP microcosm experiments revealed that all soil samples from paddy and natural wetlands displayed intense CH_4 consumption (**Figure 2A**). Assuming linear kinetics, MOP was highest in Changchun, followed by Xianghai, Minzhu, and Zhalong. No significant differences were observed in the MOP of all incubations between paddy and natural wetlands [$H(K) = 3.103$, $p = 0.078$], but significant differences in MOP were observed among the four wetland sites [$H(K) = 8.897$, $p = 0.031$].

Community Profiling of MOB Communities

Quantitative results of *pmoA* gene copy numbers across the entire buoyant density gradient of the fractionated DNA samples showed that the relative abundances of *pmoA* gene ^{13}C-labeled in heavy DNA fractions were significantly increased, compared to the background values from the same DNA fractions in the ^{12}C-labeled control treatment (**Supplementary Figure S1**). High-throughput sequencing of the *pmoA* genes was performed on C-labeled samples from SIP microcosms. The community structure of ^{12}C-labeled MOB is shown in **Supplementary Figure S2** as background data. The subsequent analysis focused on the ^{13}C-labeled MOB, which were considered active MOB communities.

Species richness and Faith's phylogenetic diversity did not significantly differ among wetland sites (**Supplementary Figure S3**). No correlations were observed between MOPs and alpha diversity (**Supplementary Table S2**). The only correlation detected was a significant relationship between the relative abundance of *Methylosarcina* and MOP at the genus level ($r = 0.734$, $p = 0.007$; **Supplementary Table S3**). The PCoA plot revealed large differences in the community composition of the four wetland sites (ANOSIM $r = 0.896$, $p = 0.001$) and two wetland types (ANOSIM $r = 0.842$, $p = 0.001$), and these findings were consistent with the other dissimilarity distances (**Figure 2B**; **Supplementary Table S4**). The type Ia methanotroph *Methylosarcina* was the most common taxon in each incubation (**Figures 2C**,**D**); it was the most common in the CC incubations, accounting for ~57.4% of total sequences

[2]https://gephi.org/
[3]https://www.r-project.org

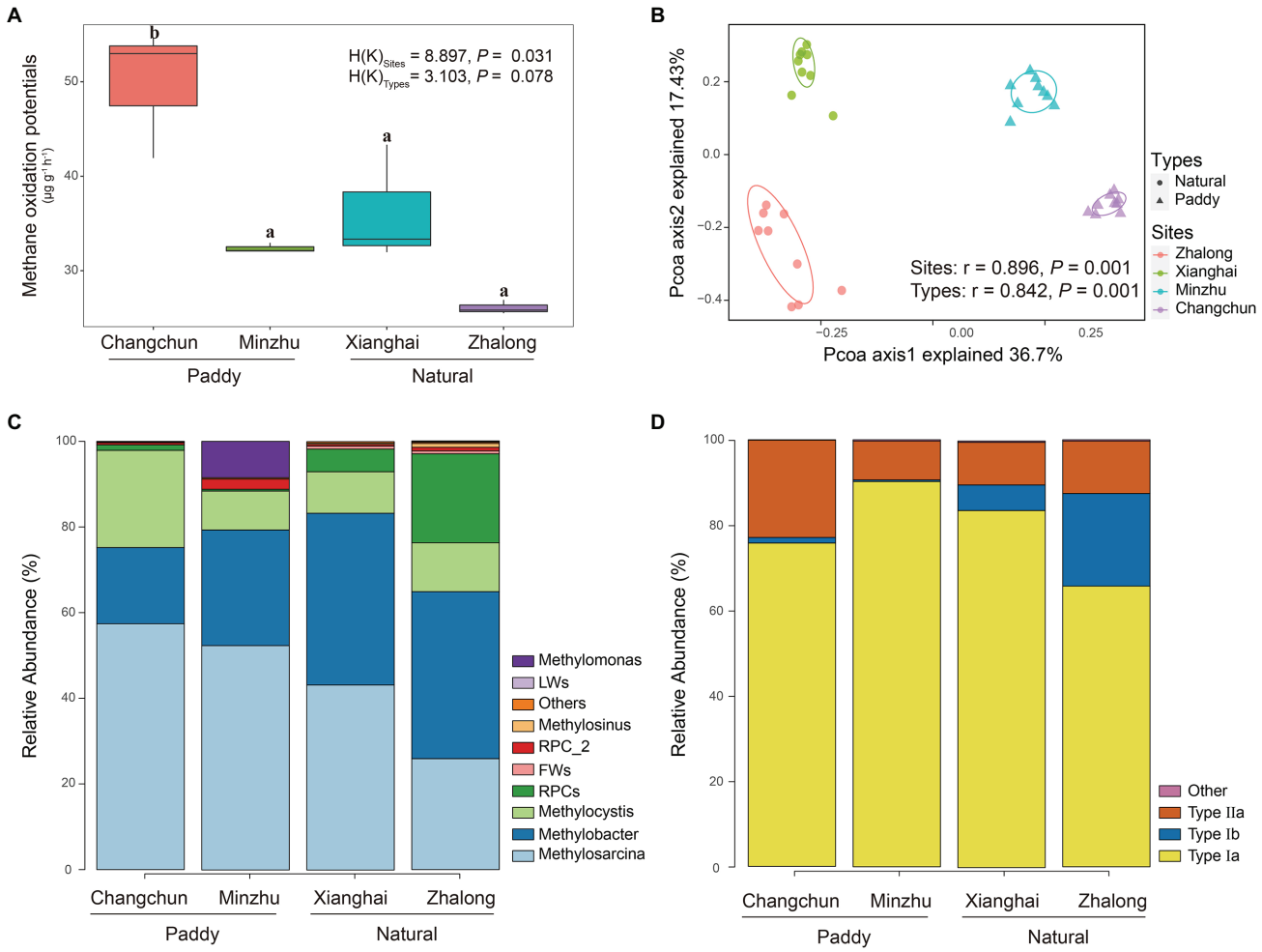


FIGURE 2 | (A) Methane oxidation potentials at the CH_4 incubations of four wetland soils. Boxes with different lowercase letters represent significant differences ($p < 0.05$). **(B)** Principal coordinates analysis (PCoA) biplots of Bray–Curtis distances for the active methanotrophic communities of four wetland soils. **(C,D)** Relative abundance of active methanotrophs on the genus levels and methanotrophic category levels identified with *pmoA* gene sequencing in four wetland soils.

(~25.9% in ZL; ~43.2% in XH; and ~52.4% in MZ). *Methylobacter*-affiliated *pmoA* genotypes were also common in all the incubations (~39.0% in ZL; ~40.1% in XH; ~27.0% in MZ; and ~17.8% in CC). Type Ib methanotroph *RPCs* were only common in ZL (~22.7% in ZL; ~4.2% in XH; ~0.5% in MZ; and ~1.6% in CC), and the type II methanotroph *Methylocystis* was only common in CC (~11.4% in ZL; ~9.7% in XH; ~9.1% in MZ; and ~22.7% in CC). The type II methanotroph *Methylomonas* was common in MZ (~8.5%). These methanotrophs accounted for more than 90% of the total microbes. LEfSe analysis revealed that *Methylobacter*, *Methylosinus*, and *FWs*-affiliated clusters were enriched in natural wetlands, and *Methylosarcina* was more prevalent in paddy than in natural wetlands, these methanotrophs used to distinguish the MOB communities between paddy and natural wetlands as indicators (**Figure 3**).

In the TITAN2 analysis, the peaks Z^- and Z^+ with the gradient of MOPs appeared around 13 and 30 $\mu g\, g^{-1}\ h^{-1}$, respectively. In paddy wetlands, type I *Methylosarcina* and type II *Methylocystis* were Z^+ taxa, which increased in frequency and occurrence as MOPs changed, whereas only type I genotypes (e.g., *Methylomonas*, *RPC_2*, *Methylosarcina*, and *Methylobacter*) were Z^- taxa (**Supplementary Figure S4**). Z^- and Z^+ taxa showed opposite patterns in natural wetlands: Type I and type II genotypes were generally Z^- taxa, and type I genotypes were generally Z^+ taxa (with the exception of 2 OTUs).

Co-occurrence Relationships of MOB Communities

Distinct patterns of co-occurrence relationships were observed between paddy and natural wetlands, which were determined based on the methanotrophic networks (**Figure 4**). A total of 1,416 positive and 124 negative links were observed in the network of paddy wetlands, whereas 3,127 positive and 366 negative links were detected in the network of natural wetlands, which involved a total of 464 and 576 OTUs, respectively (**Supplementary Table S5**). Degree scores defined

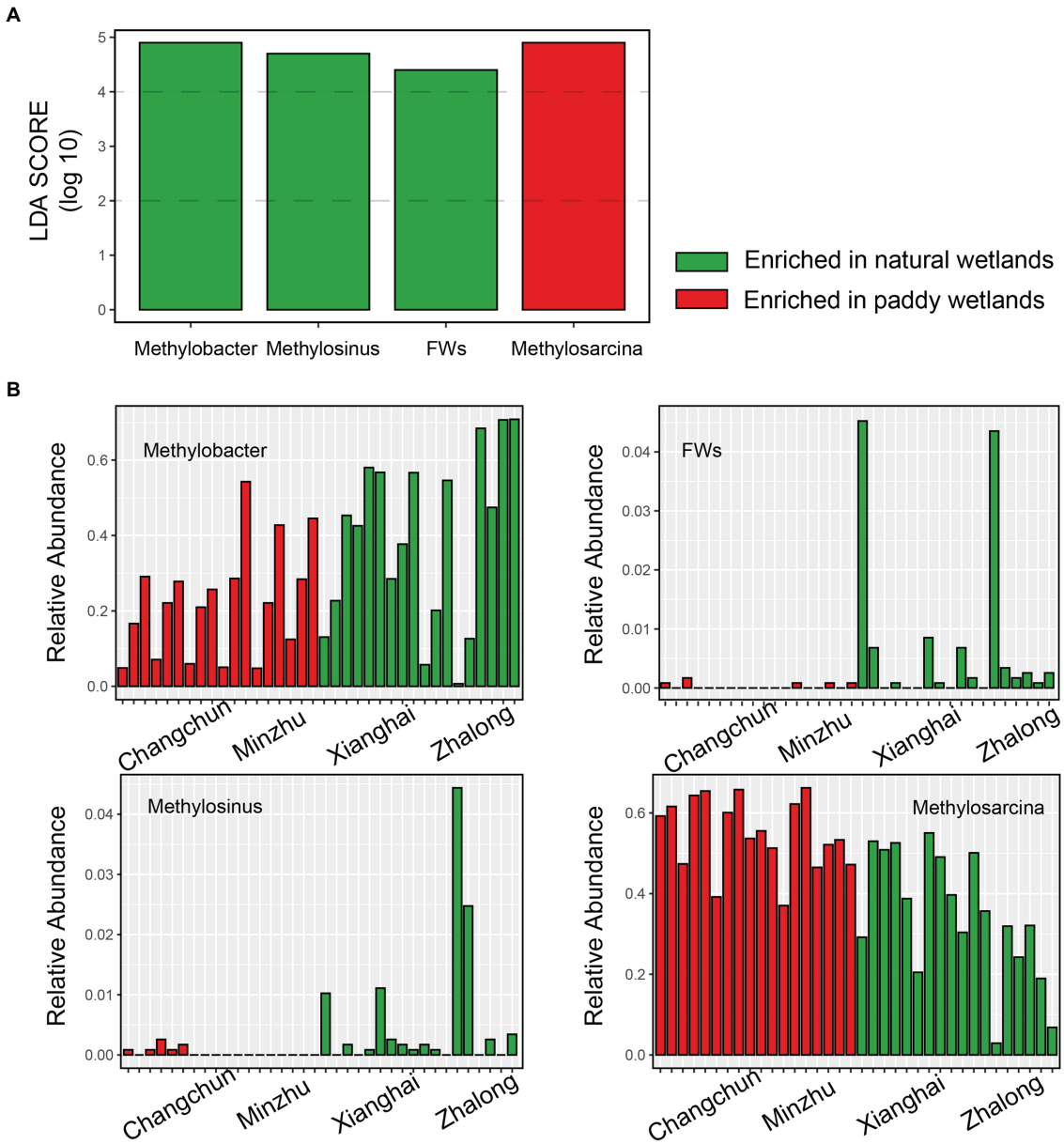


FIGURE 3 | The upper barplot showed LDA score of the active methanotrophic communities between paddy and natural wetlands identified by Lefse analysis **(A)**. The following four histograms showed the relative abundance of indicative methanotrophs that make differences **(B)**.

as the number of direct associations of the involved node and betweenness centralities indicating the degree of the involved node passed by network paths significantly differed between paddy and natural wetlands (Degree scores: $F = 26.68$, $p < 0.001$; Betweenness centralities: $F = 173.38$, $p < 0.001$; **Figure 4B**). The network analysis suggested that network stability was higher for the natural wetland network than the paddy wetland network, as the former had a greater proportion of removed nodes (**Figure 4C**). We also compared the abundance of methanotrophs in the networks and found that type I methanotrophs were mainly composed of differentially abundant taxa (**Figure 5A**). Furthermore, network module analysis divided the paddy wetland network into 10 modules and the natural wetland network into 5 modules. The community composition of each module is shown in **Figures 5B,C**. Linear regression showed that modules 1, 3, 4, 5, and 7 in the network for paddy wetlands were significantly related to MOPs; modules 1, 2, and 5 in the network for natural wetlands were significantly related to MOPs (**Supplementary Figure S5**). Other topological properties, such as the clustering coefficient, average neighbors, and centralization, are shown in **Supplementary Table S5**.

The nodes in the networks were divided into network hubs, module hubs, connectors, and peripherals by calculating Zi

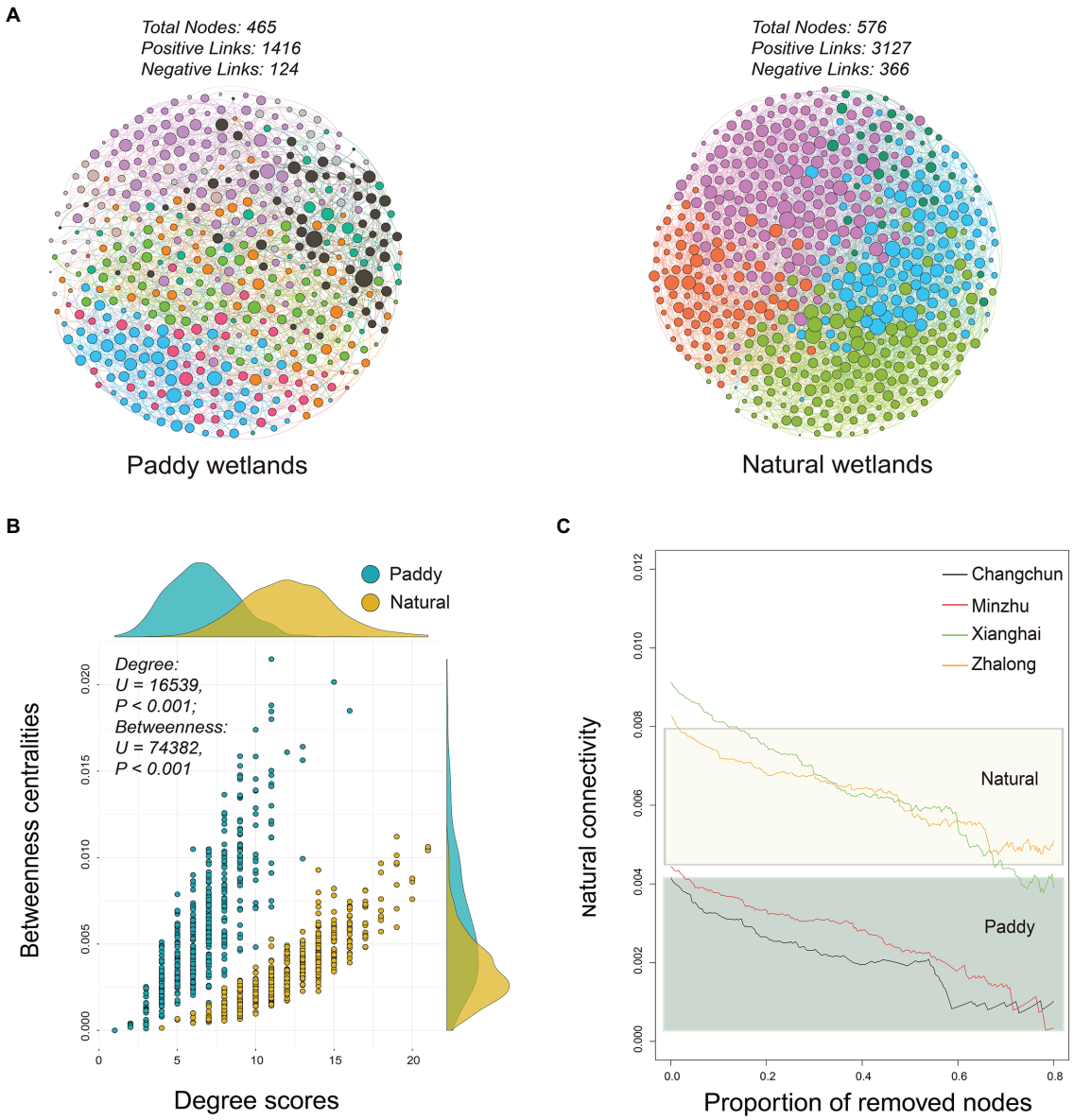


FIGURE 4 | (A) Networks visualizing co-occurring phylotypes of active methanotrophs in paddy and natural wetlands. All co-occurrence networks were analyzed by Spiec-Easi method. Colorful nodes represented different ecological module affiliations of active methanotrophs and edges on behalf of potential ecological co-occurring relationships. **(B)** Relationships between degree scores and betweenness centralities in active methanotrophic co-occurrence networks of four wetland soils. The difference test was adopted by Mann–Whitney U test. **(C)** Network robustness analysis of active methanotrophic co-occurrence networks in four wetland soils, measured as natural connectivities with remained proportion of nodes.

and Pi values. A less hub-based and more connected network structure in both paddy and natural wetlands was observed (**Supplementary Figure S6**). There were 5 OTUs that were regarded as module hubs in paddy wetlands, which were most closely affiliated to *RPC_2* (1 OTU), *Methylosarcina* (1 OTU), *Methylobacter* (1 OTU), and *Methylocystis* (2 OTU). Seven module hubs belonged to *Methylosarcina* (2 OTUs), *Methylobacter* (3 OTUs), *RPCs* (1 OTU), and *Methylocystis* (1 OTU), all of which were type I with the exception of *Methylocystis* in natural wetlands (**Supplementary Table S6**).

Assembly Processes of MOB Communities

A phylogenetic tree was constructed to characterize the taxonomic distribution and relative abundance of methanotrophs (**Figure 6A**); the tree revealed that type I methanotrophs were the most common. Standardized MPD was calculated to characterize patterns of phylogenetic distance among methanotrophic communities (**Figure 6B**; **Supplementary Table S7**), and these calculations revealed clear variation in phylogenetic distance (among sites: $F = 5.088$, $p = 0.005$; among types: $F = 5.458$, $p = 0.026$). Analysis of community assembly processes (**Figure 6C**;

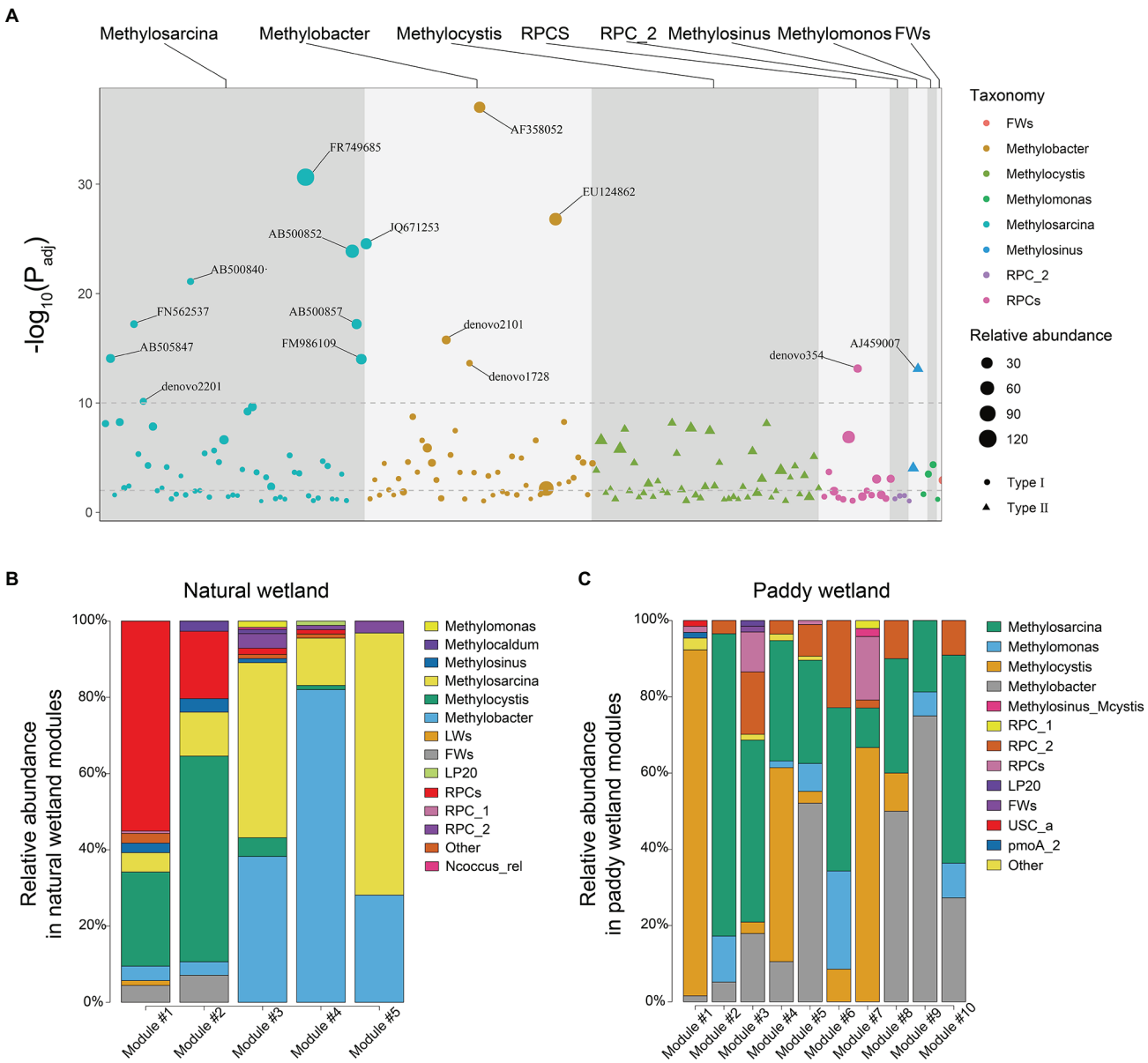


FIGURE 5 | (A) Manhattan plots showing differential operational taxonomic units (OTUs) in networks of paddy and natural wetlands. OTUs that belonged to Type I were depicted as circles shape where Type II OTUs as triangle shape. The dashed line corresponded to the adjusted value of *p* threshold of significance ($\alpha = 0.01/0.001$). The color of each dot represented the different taxonomic affiliation of the OTUs, and the size corresponded to their abundance in the respective samples. Gray backgrounds were used to denote the different taxonomic groups. **(B,C)** Relative abundance of the dominant genus in the respective main ecological modules of natural and paddy wetland networks, respectively.

Supplementary Table S8) revealed that the community turnover in CC was dominated by HD (~58.2%), followed by HS, UD, and VS. In MZ paddy wetlands, VS played an important role in community assembly (~31.6%), and UD, HD, and HS played equally important but secondary roles. The importance of stochastic processes in natural wetlands was highly variable compared with deterministic processes. Regression was used to evaluate the relationship between differences in MOPs and βNTI/RC_{bray} (**Supplementary Figure S7**). Differences in βNTI among paired sites were largely affected by VS (~28.2%) and negatively correlated with differences in MOPs. RC_{bray} was largely affected by DL (~3.37%) and UD (~26.3%) and positively correlated with differences in MOPs.

DISCUSSION

DNA-SIP experiments were conducted to study the structure profiling, co-occurrence relationships, and assembly processes of methanotrophic communities in paddy and natural wetlands.

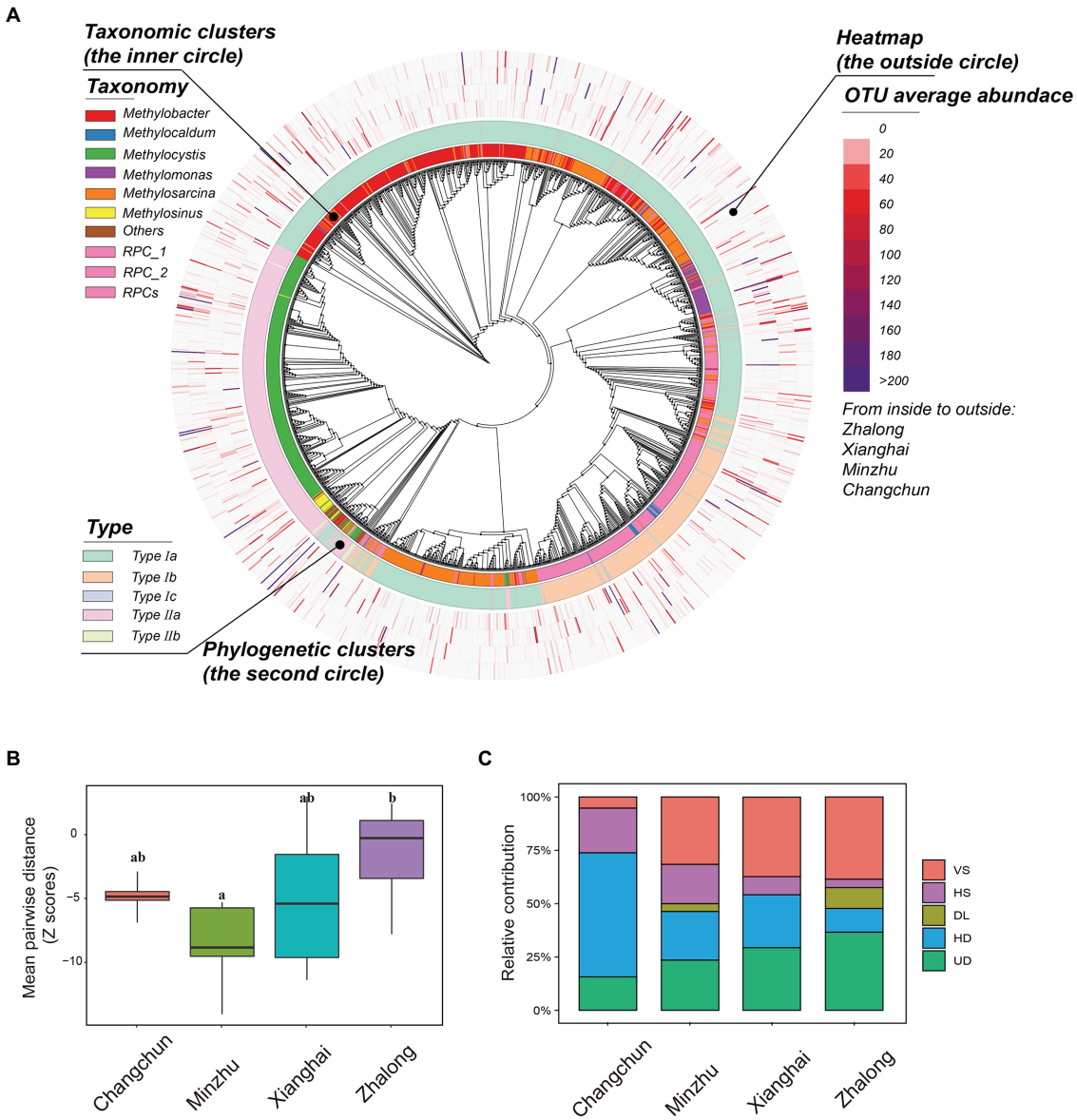


FIGURE 6 | (A) Phylogenetic tree displaying the taxonomic information, phylogenetic clusters, and OTU average abundance from inside layer to the outside layer of active methanotrophic phylotypes in four wetland soils. **(B)** The standardized difference of mean pairwise distance for active methanotrophic communities in four wetland soils using standard deviation (Z-score). Boxes with different lowercase letters represent significant differences ($p < 0.05$). **(C)** The contributions of different assembly processes of active methanotrophic communities in four wetland soils. The five processes indicated the relative importance in dominating the community dissimilarity between two samples. VS, variable selection; HS, homogeneous selection; DL, dispersal limitation; HD, homogenizing dispersal; UD, undominated.

Paddy and natural wetlands are two important types of wetlands that play an important role in CH_4 consumption driven by soil microorganisms. In our experiment, CH_4 oxidation activity was high in all incubations from the two types of wetlands, and methanotrophic communities were successfully labeled by ^{13}C; this permitted us to compare the structure and function of the active methanotroph community in both types of wetlands.

Community Compositions

Methanotrophic communities were highly diverse, and the composition of these communities in paddy and natural wetlands was distinct. Four differentially abundant and indicative methanotrophs were detected through LEfSe analysis. *Methylobacter* was enriched in natural wetlands, known to consist of psychrophilic methanotrophs with cold-adapted properties (Wartiainen et al., 2006; Dieser et al., 2014). *Methylosarcina*, which was prevailing in paddy wetlands, widely occurs in human-made ecosystems (McDonald et al., 2008; Martineau et al., 2010). *FWs*, an uncultured phylogenetical cluster, are usually present in natural plant-rich wetlands (Zhu et al., 2007). *Methylosinus* consists of acidophilic bacteria and MOB (Zhang et al., 2020). These

differentially abundant and indicative taxa explained a large amount of variation in community composition between paddy and natural wetlands.

Co-occurrence Relationships

One major finding of our study was that the co-occurrence relationships of methanotrophic communities in paddy and natural wetlands were distinct (**Figure 4**). One explanation for this pattern might be attributed to the more phylogenetically diverse phenotypes in natural wetlands compared with paddy wetlands, given that the survival of microbes critically depends on biotic interactions (Faust and Raes, 2012; Segar et al., 2020). Previous studies have explored the relationships between network interactions and phylogenetic distance (Goberna et al., 2019). Positive interactions (aggregative links) tend to occur between phylogenetically different species, whereas negative interactions (segregative links) either occur because of niche similarities or fitness differences. Our findings were consistent with this hypothesis: average positive links (positive links per node) were higher in natural wetlands than in paddy wetlands. This suggests that communities with larger phylogenetic distances may be characterized by robust network relationships through cooperation and facilitation interactions. Alternatively, this pattern might be explained by the fact that natural wetlands are more oligotrophic than paddy wetlands; the soil organic matter and total nitrogen content were markedly lower in XH and ZL (**Supplementary Table S1**). Thus, the links between methanotrophic taxa were strong, which enhanced the resistance of the natural wetland communities to disturbance.

The microbial composition of network modules was complex, and the structure of the network modules in paddy and natural wetlands differed. Various functions in different modules of entire biomes have been reported in previous studies (Fan et al., 2018; Shi et al., 2019; Ma et al., 2020). Presumably, this finding can be explained by differences in the efficiency of CH_4 consumption of methanotrophic guilds in different modules. Fan et al. (2019) showed that nitrogen fixation rates are significantly related to key modules of diazotrophic networks. The relative abundance of methanotrophs within different modules was correlated considerably with MOPs in this study (**Supplementary Figure S5**), indicating that module abundance and composition may modulate functional potentials (Bodelier et al., 2013). The niche partitioning of methanotrophs within a community can lead to the appearance of keystone species that dominate in a core function (e.g., module hubs) of CH_4 consumption (Toju et al., 2018; Ma et al., 2020). Both rare and highly abundant methanotrophs occupied critical positions in the networks (**Supplementary Figure S6**). For example, the module hub *RPC_2* was rare in paddy wetlands. However, methanotrophs were only identified to the genus level, and most of their ecological and metabolic functions remain poorly known. Future work focusing on uncultured methanotrophs, such as the provisional genus *RPC_2*, is important for identifying the roles of these keystone guilds in wetland ecosystems.

Assembly Processes

Community assembly processes that help us enhance the understanding of how communities were structured, also varied among different wetlands. Phylogenetic stochasticity was observed in ZL, and phylogenetic clustering was observed in XH, MZ, and CC (**Figure 6B**). Differences in functional potentials might be driven by phylogenetic stochasticity, but the efficiency of MOPs was limited in ZL. Saturated or flooding conditions (e.g., in paddy wetlands) were promoted by hydrologic mixing, and these conditions presumably enhanced the capacity of microorganisms to migrate across geographical regions (Liu et al., 2020), which might explain why the relative importance of HD and ecological drift was higher in paddy wetlands compared with other processes. Our findings were consistent with the results of Liu et al. (2020): UD contributed the most to shaping the communities in ZL natural wetlands, which enhanced the degree of phylogenetic stochasticity. By contrast, phylogenetic clustering can promote convergence in functional potentials and enhance the efficiency of MOPs depending on the methanotrophic guild. These results provide new insights into the CH_4 consumption potentials of methanotrophs, especially for the large quantity of unculturable or unknown methanotrophic taxa.

The dominant process shifted from HD at high MOPs to VS at low MOPs (**Figure 6C**; **Supplementary Table S8**). This result indicates that the observed processes can drive patterns consistent with the findings of previous studies (Zhang et al., 2019; Liu et al., 2020). HD resulted in greater community similarity under intensive farm management in paddy wetlands than in natural wetlands. These findings validated the assembly theory of Stegen et al. (2015), in which HD is dominant in wetland ecosystems with weak environmental selection, possibly because of fertilization (e.g., Changchun paddy wetland). In contrast, wetlands in which environmental selection is stronger were dominated by VS (e.g., Zhalong natural wetland). Consistent with the results of Feng et al. (2018), fertilization induced a shift in the dominant process affecting the diazotrophic community from deterministic processes to stochastic processes, suggesting that DL and drift became increasingly important factors shaping communities compared with selection. Based on the correlation between the assembly processes of methanotrophs and differences in the CH_4 oxidation potential, dispersal limitation and ecological drift were strongly associated with the potentials, explaining 9.0 and 13.7% of variations, respectively. These results indicated that the functional potential might be closely related to stochastic processes. Exploration of the associations between processes and functions is novel and adventurous, and this involves how to define and determine phylogenetic distance/community dissimilarity and the relative magnitude of different functional parameters in the future.

CONCLUSION

In sum, the co-occurrence relationships and assembly processes of methanotrophic communities in paddy and natural wetlands in Northeast China were distinct. DNA-SIP experiment and

high-throughput sequencing were used to compare the active methanotrophic communities between paddy and natural wetlands; network and phylogenetic analysis were applied to reveal co-occurrence relationships and assembly processes. Analysis of differences in the community composition revealed that *Methylosarcina*, *Methylobacter*, *Methylosinus*, and *FWs* were the most differentially indicative taxa between paddy and natural wetlands, and these taxa potentially contribute to differences in CH_4 consumption. The network structure of methanotrophic communities might be more stable in natural wetlands than in paddy wetlands. Degree scores and betweenness centralities differed between paddy and natural wetlands. Exploration of the relationships between module abundance and MOP revealed that the relative abundance of key sub-communities within modules accounted for the observed variation in functional potentials. Although stochastic processes contributed the most to community turnover, additional work is needed to credibly explore the potential links between community turnover and functional differences. The results of this study enhance our understanding of differences in community profiling, the co-occurrence relationships, and assembly processes of active methanotrophic communities between paddy and natural wetlands.

DATA AVAILABILITY STATEMENT

The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

All authors contributed the intellectual input and assistance to this study and manuscript preparation. HC, YS, JL, and XL provided initial concept and analysis. XL, LZ, RL, RX, and CL carried out field sampling, soil microbial, and chemical analyses. XL, TY, G-FG, and HC wrote manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA28020202), the National Natural Science Foundation of China (91951109), and the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, 2019QZKK0503).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.809074/full#supplementary-material

REFERENCES

Baker, M. E., and King, R. S. (2010). A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods Ecol. Evol.* 1, 25–37. doi: 10.1111/j.2041-210X.2009.00007.x

Bodelier, P. L., Meima-Franke, M., Hordijk, C. A., Steenbergh, A. K., Hefting, M. M., Bodrossy, L., et al. (2013). Microbial minorities modulate methane consumption through niche partitioning. *ISME J.* 7, 2214–2228. doi: 10.1038/ismej.2013.99

Bousquet, P., Ciais, P., Miller, J. B., Dlugokencky, E. J., Hauglustaine, D. A., Prigent, C., et al. (2006). Contribution of anthropogenic and natural sources to atmospheric methane variability. *Nature* 443, 439–443. doi: 10.1038/nature05132

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Chowdhury, T. R., and Dick, R. P. (2013). Ecology of aerobic methanotrophs in controlling methane fluxes from wetlands. *Appl. Soil Ecol.* 65, 8–22. doi: 10.1016/j.apsoil.2012.12.014

Cole, J. R., Wang, Q., Fish, J. A., Chai, B. L., McGarrell, D. M., Sun, Y. N., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244

Conrad, R. (2009). The global methane cycle: recent advances in understanding the microbial processes involved. *Environ. Microbiol. Rep.* 1, 285–292. doi: 10.1111/j.1758-2229.2009.00038.x

Costello, A. M., and Lidstrom, M. E. (1999). Molecular characterization of functional and phylogenetic genes from natural populations of methanotrophs in lake sediments. *Appl. Environ. Microbiol.* 65, 5066–5074. doi: 10.1128/AEM.65.11.5066-5074.1999

Degelmann, D. M., Borken, W., Drake, H. L., and Kolb, S. (2010). Different atmospheric methane-oxidizing communities in European beech and Norway spruce soils. *Appl. Environ. Microbiol.* 76, 3228–3235. doi: 10.1128/AEM.02730-09

DeSantis, J. T. Z., Hugenholtz, P., Keller, K., Brodie, E. L., Larsen, N., Piceno, Y. M., et al. (2006). NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Res.* 34(Suppl. 2), W394–W399. doi: 10.1093/nar/gkl244

Dieser, M., Broemsen, E. L. J. E., Cameron, K. A., King, G. M., Achberger, A., Choquette, K., et al. (2014). Molecular and biogeochemical evidence for methane cycling beneath the western margin of the Greenland ice sheet. *ISME J.* 8, 2305–2316. doi: 10.1038/ismej.2014.59

Ding, W.-X., and Cai, Z.-C. (2007). Methane emission from natural wetlands in China: summary of years 1995–2004 studies. *Pedosphere* 17, 475–486. doi: 10.1016/S1002-0160(07)60057-5

Dumont, M. G., Luke, C., Deng, Y. C., and Frenzel, P. (2014). Classification of pmoA amplicon pyrosequences using BLAST and the lowest common ancestor method in MEGAN. *Front. Microbiol.* 5:34. doi: 10.3389/fmicb.2014.00034

Dumont, M. G., Pommerenke, B., Casper, P., and Conrad, R. (2011). DNA-, rRNA- and mRNA-based stable isotope probing of aerobic methanotrophs in lake sediment. *Environ. Microbiol.* 13, 1153–1167. doi: 10.1111/j.1462-2920.2010.02415.x

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Emerson, B. C., and Gillespie, R. G. (2008). Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol. Evol.* 23, 619–630. doi: 10.1016/j.tree.2008.07.005

Fan, K., Delgado-Baquerizo, M., Guo, X., Wang, D., Wu, Y., Zhu, M., et al. (2019). Suppressed N fixation and diazotrophs after four decades of fertilization. *Microbiome* 7:143. doi: 10.1186/s40168-019-0757-8

Fan, K., Weisenhorn, P., Gilbert, J. A., and Chu, H. (2018). Wheat rhizosphere harbors a less complex and more stable microbial co-occurrence pattern than bulk soil. *Soil Biol. Biochem.* 125, 251–260. doi: 10.1016/j.soilbio.2018.07.022

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550. doi: 10.1038/nrmicro2832

Feng, M., Adams, J. M., Fan, K., Shi, Y., Sun, R., Wang, D., et al. (2018). Long-term fertilization influences community assembly processes of soil diazotrophs. *Soil Biol. Biochem.* 126, 151–158. doi: 10.1016/j.soilbio.2018.08.021

Goberna, M., Montesinos-Navarro, A., Valiente-Banuet, A., Colin, Y., Gomez-Fernandez, A., Donat, S., et al. (2019). Incorporating phylogenetic metrics to microbial co-occurrence networks based on amplicon sequences to discern community assembly processes. *Mol. Ecol. Resour.* 19, 1552–1564. doi: 10.1111/1755-0998.13079

He, R., Wang, J., Pohlman, J. W., Jia, Z., Chu, Y. X., Wooller, M. J., et al. (2021). Metabolic flexibility of aerobic methanotrophs under anoxic conditions in Arctic lake sediments. *ISME J.* 16, 78–90. doi: 10.1038/s41396-021-01049-y

He, D., Zhang, L. Y., Dumont, M. G., He, J. S., Ren, L. J., and Chu, H. Y. (2019). The response of methanotrophs to additions of either ammonium, nitrate or urea in alpine swamp meadow soil as revealed by stable isotope probing. *FEMS Microbiol. Ecol.* 95:fiz077. doi: 10.1093/femsec/fiz077

Hines, J. (2019). Ecosystem functioning: how much system is needed to explain function? *Curr. Biol.* 29, R1072–R1074. doi: 10.1016/j.cub.2019.09.037

Ho, A., Kerckhof, F.-M., Luke, C., Reim, A., Krause, S., Boon, N., et al. (2013). Conceptualizing functional traits and ecological characteristics of methane-oxidizing bacteria as life strategies. *Environ. Microbiol. Rep.* 5, 335–345. doi: 10.1111/j.1758-2229.2012.00370.x

Jia, Z., Cao, W., and Hernandez Garcia, M. (2019). DNA-based stable isotope probing. *Methods Mol. Biol.* 2046, 17–29. doi: 10.1007/978-1-4939-9721-3_2

Jiao, S., Xu, Y., Zhang, J., and Lu, Y. (2019). Environmental filtering drives distinct continental atlases of soil archaea between dryland and wetland agricultural ecosystems. *Microbiome* 7:15. doi: 10.1186/s40168-019-0630-9

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kirschke, S., Bousquet, P., Ciais, P., Saunois, M., Canadell, J. G., Dlugokencky, E. J., et al. (2013). Three decades of global methane sources and sinks. *Nat. Geosci.* 6, 813–823. doi: 10.1038/ngeo1955

Kurtz, Z. D., Müller, C. L., Miraldi, E. R., Dan, R. L., Blaser, M. J., and Bonneau, R. A. (2015). Sparse and compositionally robust inference of microbial ecological networks. *PLoS Comput. Biol.* 11:e1004226. doi: 10.1371/journal.pcbi.1004226

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Li, D., Ni, H., Jiao, S., Lu, Y., Zhou, J., Sun, B., et al. (2021). Coexistence patterns of soil methanogens are closely tied to methane generation and community assembly in rice paddies. *Microbiome* 9:20. doi: 10.1186/s40168-020-00978-8

Liu, W., Graham, E. B., Dong, Y., Zhong, L., Zhang, J., Qiu, C., et al. (2020). Balanced stochastic versus deterministic assembly processes benefit diverse yet uneven ecosystem functions in representative agroecosystems. *Environ. Microbiol.* 23, 391–404. doi: 10.1111/1462-2920.15326

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lu, L., and Jia, Z. J. (2013). Urease gene-containing archaea dominate autotrophic ammonia oxidation in two acid soils. *Environ. Microbiol.* 15, 1795–1809. doi: 10.1111/1462-2920.12071

Ma, B., Wang, Y., Ye, S., Liu, S., Stirling, E., Gilbert, J. A., et al. (2020). Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome* 8:82. doi: 10.1186/s40168-020-00857-2

Martineau, C., Whyte, L. G., and Greer, C. W. (2010). Stable isotope probing analysis of the diversity and activity of methanotrophic bacteria in soils from the Canadian high Arctic. *Appl. Environ. Microbiol.* 76, 5773–5784. doi: 10.1128/AEM.03094-09

McCalley, C. K., Woodcroft, B. J., Hodgkins, S. B., Wehr, R. A., Kim, E. H., Mondav, R., et al. (2014). Methane dynamics regulated by microbial community response to permafrost thaw. *Nature* 514, 478–481. doi: 10.1038/nature13798

McDonald, I. R., Bodrossy, L., Chen, Y., and Murrell, J. C. (2008). Molecular ecology techniques for the study of aerobic methanotrophs. *Appl. Environ. Microbiol.* 74, 1305–1315. doi: 10.1128/AEM.02233-07

Mer, J. L., and Roger, P. (2001). Production, oxidation, emission and consumption of methane by soils: a review. *Eur. J. Soil Biol.* 37, 25–50. doi: 10.1016/S1164-5563(01)01067-6

Mittelbach, G. G., and Schemske, D. W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends Ecol. Evol.* 30, 241–247. doi: 10.1016/j.tree.2015.02.008

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356. doi: 10.1128/MMBR.00051-12

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., et al. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* 11:4717. doi: 10.1038/s41467-020-18560-z

Peng, J., Wegner, C.-E., Bei, Q., Liu, P., and Liesack, W. (2018). Metatranscriptomics reveals a differential temperature effect on the structural and functional organization of the anaerobic food web in rice field soil. *Microbiome* 6:169. doi: 10.1186/s40168-018-0546-9

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Röttjers, L., and Faust, K. (2018). From hairballs to hypotheses–biological insights from microbial networks. *FEMS Microbiol. Rev.* 42, 761–780. doi: 10.1093/femsre/fuy030

Saunois, M., Stavert, A. R., Poulter, B., Bousquet, P., Canadell, J. G., Jackson, R. B., et al. (2020). The global methane budget 2000–2017. *Earth Syst. Sci. Data* 12, 1561–1623. doi: 10.5194/essd-12-1561-2020

Segar, S. T., Fayle, T. M., Srivastava, D. S., Lewinsohn, T. M., Lewis, O. T., Novotny, V., et al. (2020). The role of evolution in shaping ecological networks. *Trends Ecol. Evol.* 35, 454–466. doi: 10.1016/j.tree.2020.01.004

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12:R60. doi: 10.1186/gb-2011-12-6-r60

Shi, Y., Delgado-Baquerizo, M., Li, Y., Yang, Y., Zhu, Y.-G., Peñuelas, J., et al. (2020). Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems. *Environ. Int.* 142:105869. doi: 10.1016/j.envint.2020.105869

Shi, Y., Fan, K., Li, Y., Yang, T., He, J.-S., and Chu, H. (2019). Archaea enhance the robustness of microbial co-occurrence networks in Tibetan Plateau soils. *Soil Sci. Soc. Am. J.* 83, 1093–1099. doi: 10.2136/sssaj2018.11.0426

Shiau, Y. J., Cai, Y. F., Jia, Z. J., Chen, C. L., and Chiu, C. Y. (2018). Phylogenetically distinct methanotrophs modulate methane oxidation in rice paddies across Taiwan. *Soil Biol. Biochem.* 124, 59–69. doi: 10.1016/j.soilbio.2018.05.025

Stegen, J. C., Lin, X. J., Fredrickson, J. K., Chen, X. Y., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X. J., Fredrickson, J. K., and Konopka, A. E. (2015). Estimating and mapping ecological processes influencing microbial community assembly. *Front. Microbiol.* 6:370. doi: 10.3389/fmicb.2015.00370

Sultana, N., Zhao, J., Zheng, Y., Cai, Y. F., Faheem, M., Peng, X. L., et al. (2019). Stable isotope probing of active methane oxidizers in rice field soils from cold regions. *Biol. Fertil. Soils* 55, 243–250. doi: 10.1007/s00374-018-01334-7

Tang, X., Zhao, X., Bai, Y., Tang, Z., Wang, W., Zhao, Y., et al. (2018). Carbon pools in China's terrestrial ecosystems: new estimates based on an intensive field survey. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4021–4026. doi: 10.1073/pnas.1700291115

Toju, H., Tanabe, A. S., and Sato, H. (2018). Network hubs in root-associated fungal metacommunities. *Microbiome* 6:116. doi: 10.1186/s40168-018-0497-1

Vellend, M. (2010). Conceptual synthesis in community ecology. *Quarterly Review of Biology* 85, 183–206. doi: 10.1086/652373

Wang, J., Soininen, J., He, J., and Shen, J. (2012). Phylogenetic clustering increases with elevation for microbes. *Environ. Microbiol. Rep.* 4, 217–226. doi: 10.1111/j.1758-2229.2011.00324.x

Wartiainen, I., Hestnes, A. G., McDonald, I. R., and Svenning, M. M. (2006). *Methylobacter tundripaludum* sp. nov., a methane-oxidizing bacterium from Arctic wetland soil on the Svalbard islands, Norway (78 degrees N). *Int. J. Syst. Evol. Microbiol.* 56, 109–113. doi: 10.1099/ijs.0.63728-0

Weiher, E., Freund, D., Bunton, T., Stefanski, A., Lee, T., and Bentivenga, S. (2011). Advances, challenges and a developing synthesis of ecological community assembly theory. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 2403–2413. doi: 10.1098/rstb.2011.0056

Yang, Y. (2021). Emerging patterns of microbial functional traits. *Trends Microbiol.* 29, 874–882. doi: 10.1016/j.tim.2021.04.004

Yun, J., Zhuang, G., Ma, A., Guo, H., Wang, Y., and Zhang, H. (2012). Community structure, abundance, and activity of methanotrophs in the zoige wetland of the Tibetan Plateau. *Microb. Ecol.* 63, 835–843. doi: 10.1007/s00248-011-9981-x

Zhang, L. Y., Adams, J. M., Dumont, M. G., Li, Y. T., Shi, Y., He, D., et al. (2019). Distinct methanotrophic communities exist in habitats with different soil water contents. *Soil Biol. Biochem.* 132, 143–152. doi: 10.1016/j.soilbio.2019.02.007

Zhang, L. Y., Dumont, M. G., Bodelier, P. L. E., Adams, J. M., He, D., and Chu, H. (2020). DNA stable-isotope probing highlights the effects of temperature on functionally active methanotrophs in natural wetlands. *Soil Biol. Biochem.* 149:107954. doi: 10.1016/j.soilbio.2020.107954

Zhu, N., An, P., Krishnakumar, B., Zhao, L., Sun, L., Mizuochi, M., et al. (2007). Effect of plant harvest on methane emission from two constructed wetlands designed for the treatment of wastewater. *J. Environ. Manag.* 85, 936–943. doi: 10.1016/j.jenvman.2006.11.004

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 18 February 2022
doi: 10.3389/fmicb.2022.803420

Inferring the Contribution of Microbial Taxa and Organic Matter Molecular Formulas to Ecological Assembly

Robert E. Danczak[1], Aditi Sengupta[2], Sarah J. Fansler[1], Rosalie K. Chu[3], Vanessa A. Garayburu-Caruso[1], Lupita Renteria[1], Jason Toyoda[3], Jacqueline Wells[1] and James C. Stegen[1]*

[1]Ecosystem Sciences, Pacific Northwest National Laboratory, Richland, WA, United States, [2]Department of Biology, California Lutheran University, Thousand Oaks, CA, United States, [3]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, United States

OPEN ACCESS

Edited by:
Daliang Ning, University of Oklahoma, United States

Reviewed by:
Fangqiong Ling, Washington University in St. Louis, United States
Oskar Modin, Chalmers University of Technology, Sweden

***Correspondence:**
Robert E. Danczak
robert.danczak@pnnl.gov

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *28 October 2021*
Accepted: *14 January 2022*
Published: *18 February 2022*

Citation:
Danczak RE, Sengupta A, Fansler SJ, Chu RK, Garayburu-Caruso VA, Renteria L, Toyoda J, Wells J and Stegen JC (2022) Inferring the Contribution of Microbial Taxa and Organic Matter Molecular Formulas to Ecological Assembly. Front. Microbiol. 13:803420. doi: 10.3389/fmicb.2022.803420

Understanding the mechanisms underlying the assembly of communities has long been the goal of many ecological studies. While several studies have evaluated community wide ecological assembly, fewer have focused on investigating the impacts of individual members within a community or assemblage on ecological assembly. Here, we adapted a previous null model β-nearest taxon index (βNTI) to measure the contribution of individual features within an ecological community to overall assembly. This new metric, called feature-level βNTI (βNTI_{feat}), enables researchers to determine whether ecological features (e.g., individual microbial taxa) contribute to divergence, convergence, or have insignificant impacts across spatiotemporally resolved metacommunities or meta-assemblages. Using βNTI_{feat}, we revealed that unclassified microbial lineages often contributed to community divergence while diverse groups (e.g., Crenarchaeota, Alphaproteobacteria, and Gammaproteobacteria) contributed to convergence. We also demonstrate that βNTI_{feat} can be extended to other ecological assemblages such as organic molecules comprising organic matter (OM) pools. OM had more inconsistent trends compared to the microbial community though CHO-containing molecular formulas often contributed to convergence, while nitrogen and phosphorus-containing formulas contributed to both convergence and divergence. A network analysis was used to relate βNTI_{feat} values from the putatively active microbial community and the OM assemblage and examine potentially common contributions to ecological assembly across different communities/assemblages. This analysis revealed that P-containing formulas often contributed to convergence/divergence separately from other ecological features and N-containing formulas often contributed to assembly in coordination with microorganisms. Additionally, members of Family *Geobacteraceae* were often observed to contribute to convergence/divergence in conjunction with both N- and P-containing formulas, suggesting a coordinated ecological role for family members and the nitrogen/phosphorus cycle. Overall, we show that βNTI_{feat} offers opportunities to investigate the community or

assemblage members, which shape the phylogenetic or functional landscape, and demonstrate the potential to evaluate potential points of coordination across various community types.

Keywords: community assembly, β-nearest taxon index, null modeling, FTICR-MS, metacommunity ecology, meta-metabolome ecology

INTRODUCTION

Evaluating the processes which govern community diversity is often the goal of ecological studies across all ecosystems (Swenson et al., 2006; Kraft et al., 2007; Gilbert and Bennett, 2010; Smith and Lundholm, 2010; Chase and Myers, 2011; George et al., 2011; Stegen et al., 2013; Herren and McMahon, 2017; Zhou and Ning, 2017; Danczak et al., 2020b). Analogously, researchers have also focused on understanding the processes governing the composition of organic molecules or metabolites within organic matter (OM) assemblages (Danczak et al., 2020a, 2021). While methods might vary in how researchers investigate these processes (e.g., variation partitioning, trait-based analyses, and null modeling), each study attempts to determine when, where, and how various ecological assembly processes give rise to specific community/assemblage configurations. By better understanding the distribution of these processes and the circumstances under which they dominate, we will be able to better understand the fundamental principles governing community/assemblage structure. However, less attention has been paid to the impact of ecological processes on individual community/assemblage members or to the impact of individual members on ecological assembly. Hereafter, in order to limit confusion, both biological and chemical members are referred to as "features," while biological communities and OM assemblages are referred to as "communities" (**Table 1**).

Evaluating the impacts of individual features provides several benefits to ecological researchers. Firstly, feature-level metrics provide researchers the opportunity to investigate how specific features (e.g., a microbial taxon or metabolite) contribute to assembly under varied environmental regimes or across different spatiotemporal scales. As detailed by Ning et al. (2020), members of a given taxonomic level can respond differently to environmental stresses, with some experiencing variable selection and others being limited by dispersal. Feature-level metrics will allow researchers to disentangle the contributions (or lack thereof) of individual members and identify groups putatively most relevant to community assembly. The ability to observe individual contributions would allow researchers to evaluate the ecological roles of specific organisms and/or organic molecules in the absence of explicit physiological or biochemical information based on traits inferred from phylogeny/taxonomy. For example, Ning et al. (2020) observed that a group of Bacillales significantly experienced homogeneous selection in hot/dry environments potentially due to their enhanced survivability.

Secondly, feature-level metrics will allow ecologists to compare and relate the contributions to assembly dynamics across community types in order to observe potential ecological coordination. For example, these metrics can be directly compared across different community types to find groups of cross-community features, which exert coordinated control on the larger metacommunity/meta-assemblage, potentially highlighting a common ecological pressure or ecological interaction (e.g., a specific environmental condition driving the selection of features across community boundaries). Investigating how assembly compares across disparate groups has enabled a deeper understanding of the fundamental factors structuring ecological communities. Danczak et al. (2020b) revealed that viral and microbial communities experienced coordinated assembly processes despite facing separate environmental pressures in a fractured shale ecosystem. Jiao et al. (2020) demonstrated that the balance of cross-kingdom species interactions across Archaea-Bacteria-Fungi mediated community assembly in an agricultural soil ecosystem. These examples indicate that the ability to measure and relate feature-level contributions to ecological assembly will help identify components disproportionally impacting phylogenetic or functional community structure.

We propose that a new metric called feature-level β-nearest taxon index (βNTI_{feat}) based upon an existing null modeling framework (βNTI) will provide these benefits. βNTI is particularly capable in assessing the assembly dynamics associated with ecological metacommunities and OM meta-metabolomes/assemblages and we show that it can be adapted to feature-level analyses. βNTI has been used extensively to study assembly processes. For example, researchers have revealed relationships between microbial community development and organic matter degradation (Stegen et al., 2016, 2018), coordination of assembly

TABLE 1 | Table of terms used throughout the manuscript and their definitions.

Terms	Definitions
Feature	*Feature* is used liberally in this manuscript. This term is meant to describe anything that can be described in a relational dendrogram (e.g., microbial community members and FTICR-MS molecular formula).
Community	*Community* refers to a collection of ecological features. Here, that describes both microbial communities and organic matter assemblages.
Contribution	$\lvert\beta NTI_{feat}\rvert > 1$, with significant contributions inferred when $\lvert\beta NTI_{feat}\rvert > 2$; When a feature "contributes" to ecological assembly, it exerts some deterministic impact on the community.
Contribution to convergence	$\beta NTI_{feat} < -2$; Features that "contribute to convergence" are those which drive ecological/functional similarities across communities.
Contribution to divergence	$\beta NTI_{feat} > 2$; Features that "contribute to divergence" are those which drive ecological/functional differences across communities.

between viral and microbial communities (Danczak et al., 2020b), the balance of niche- and dispersal-based processes the soybean microbiome (Moroenyane et al., 2021), and emphasized the importance of salinity in the assembly of desert microbial communities (Zhang et al., 2019).

Recently, Ning et al. (2020) developed an iterative null model based on β-net relatedness index (βNRI) called iCAMP, which represents a potential route to identify these feature-level dynamics. By first identifying the minimum phylogenetic level at which a phylogenetic signal exists (e.g., a relationship between evolutionary history and niche occupancy; Blomberg et al., 2003; Stegen et al., 2012), iCAMP groups community members and measures the ecological pressures acting upon that level. Using these metrics calculated across the phylogenetic tree, iCAMP can then estimate the balance of assembly processes acting on the community. This method is an excellent way to account for phylogenetic groups experiencing varied assembly processes in whole community analyses and represents a novel way to follow sub-community assembly through time or space. However, while capable of identifying assembly processes at levels below the entire community, this approach still investigates processes impacting assembly at the subcommunity level rather than measuring the degree which an individual feature impacts or is impacted by assembly. Fodelianakis et al. (2021) instead took an approach, called "phyloscore analysis" that focuses instead on the ecological contributions of specific taxa within a microbial community. Likewise, βNTI_{feat} focuses on individual features to measure their ecological contribution to community dynamics and highlights a point of complementarity across community, subcommunity, and feature-level foci (**Table 1**).

Feature-level β-nearest taxon index provides insight into the degree to which each observed feature contributes to either ecological convergence or divergence (**Table 1**). Ecological convergence occurs when some feature drives similarities in the phylogenetic or functional landscape across samples or within a dataset. Such a feature would be more phylogenetically or functionally conserved than expected by random chance. In contrast, ecological divergence occurs when a feature drives phylogenetic or functional differences across samples or a dataset; these features are more divergent than expected by random chance. Based on these interpretations, βNTI_{feat} stands as the phylogenetic or dendrogram-based complement to taxonomic metrics like SIMPER (Clarke, 1993). Importantly, given that βNTI_{feat} does not rely on abundance or taxonomic-based distance metrics, it is able to overcome many of the limitations associated with SIMPER (Warton et al., 2012). When compared to the phyloscore metric described by Fodelianakis et al. (2021), it is much more similar though differs in some mathematical specifics (namely the null implementation) and in its application across different scales.

Here, we describe the theory behind the βNTI_{feat} calculations, apply it to microbial and environmental metabolomic data, and discuss how interpretations vary with scale and dataset. First, we reveal that βNTI_{feat} can identify microbial taxa (down to the amplicon sequence variant, or ASV, level) and environmental metabolites (down to the specific molecular formulas), which disproportionally contribute to the ecological structure of the respective community. Second, we demonstrate that we can track features with disproportionate contributions through time and relate them to each other to uncover ASVs and molecular formulas that have coordinated contributions to the biological and chemical composition of the study system.

MATERIALS AND METHODS

Sample Collection

Detailed sample collection is outlined in Sengupta et al. (2021a) but will be described briefly here. Sediments were collected from the hyporheic zone of the Columbia River shoreline in eastern Washington state on 14 January 2019 at 9 a.m. Pacific Standard Time. Five samples were collected from within a meter range, combined to make a composite sediment sample, and then sieved on site through a 2 mm sieve into a glass beaker. Sieved sediment was kept on blue ice for 30 min until transported back to the lab, where it was stored at 4°C until experimentation.

Experimental Design

A detailed experimental design is outlined in Sengupta et al. (2021a) but will be described briefly here. Sieved sediment was partitioned into two sets of vials, each vial containing 10 g of sediment: one set of vials were under inundated conditions for 23 days, and the other set were allowed to dry for 23 days. Once the vials were permitted to acclimate, they were subject treatment regimes designed around a series of wet/dry transition periods. While the details are important, these regimes largely translate to two bulk treatments based upon the number of days left dry: cumulatively dry (34, 31, and 27 days) and cumulatively inundated (0, 4, and 8 days). This resulted in a total of 20 vials per cumulative treatment.

16S rRNA Gene Sequencing and Processing

Detailed DNA and cDNA extraction methods can be found in Sengupta et al. (2021a). Briefly, incubated sediments were centrifuged and flash-frozen. gDNA was extracted following the protocol established by Bottos et al. (2018) RNA was extracted using the Qiagen PowerSoil RNA extraction kit (Qiagen, Germantown, MD), treated for contaminate DNA using DNase, quantified using a Qubit RNA kit (Thermo Fisher, Waltham, MA), and reverse transcribed into cDNA using the SuperScript™ IV First-Strand Synthesis System (Thermo Fisher Scientific, Waltham, MA). Amplicon sequencing for both gDNA and cDNA was performed following The Earth Microbiome Project. Sequences are accessible at NCBI's Sequence Read Archive using the accession number # PRJNA641165.

Sequence processing was performed using QIIME2 (Bolyen et al., 2019). Raw amplicon sequences were imported into the QIIME2 environment, denoised using DADA2 (*q2-dada2*), and assigned taxonomy using the SILVA v138 database (*q2-feature-classifier*; Quast et al., 2013; Callahan et al., 2016; Bokulich et al., 2018). A phylogenetic tree was generated by first aligning

amplicons using the MAFFT aligner (Katoh, 2002) and then generating a maximum-likelihood tree (*q2-phylogeny*). The 16S rRNA gene amplicon maximum-likelihood tree is stored as **Supplementary File 1**.

FTICR-MS Analysis

Detailed Fourier transform ion cyclotron resonance mass spectrometer (FTICR-MS) data acquisition is outlined in Danczak et al. (2020a) and Sengupta et al. (2021a) but will be described briefly here. In brief, a solid-phase extraction on a PPL cartridge (Bond Elut) was used to concentrate carbon and remove salt (Dittmar et al., 2008). Extracted samples were injected into a 12 Tesla (12T) Bruker SolariX FTICR-MS outfitted with a standard electrospray ionization source (ESI) configured to negative mode. Resulting spectra were processed using the BrukerDaltonik Data Analysis software (v4.2) to obtain a peak list; Formularity was then used to assign molecular formulas to detected peaks following the Compound Identification Algorithm (Kujawinski and Behn, 2006; Tolić et al., 2017). The report generated in Formularity was then processed using the ftmsRanalysis R package to calculate various molecular properties (e.g., double-bond equivalents, modified aromaticity index, nominal oxidation state of carbon, Kendrick's defect, etc.) and assign compound classes (Hughey et al., 2001; Kim et al., 2003; Koch and Dittmar, 2006; LaRowe and Van Cappellen, 2011; Bailey et al., 2017; Rivas-Ubach et al., 2018; Bramer et al., 2020). Using the methods outlined in Danczak et al. (2020a), we generated a molecular characteristics dendrogram (MCD) for use in dendrogram-informed analyses; the MCD is stored as **Supplementary File 2**.

Ecological Analyses

Multivariate differences across the microbial community and OM assemblage were detected using ordinations in combination with PERMANOVA statistics (*adonis*; vegan package v2.5-7; Oksanen et al., 2019). A principal coordinate analysis (PCoA; *pcoa*; ape package v5.5) was generated using both Bray–Curtis dissimilarity (*vegdist*; vegan package v2.5-7) and β-mean nearest taxon distance (βMNTD; *comdistnt*; picante package v1.8.2) were calculated for both the microbial community and OM assemblage (Kembel et al., 2010; Oksanen et al., 2019; Paradis and Schliep, 2019).

The βNTI_{feat} Calculation

The βNTI_{feat} is intrinsically linked to the βNTI calculation and helps us understand the relationship between observed dendrogram-based relationships of individual features and some null expectation (**Figure 1**). First, $\beta MNTD_{feat}$, the minimum relational distance of a feature in one community to the nearest feature in another, needs to be calculated for the observed community across the entire dataset:

$$\beta\text{MNTD}_{feat} = \frac{1}{n}\sum_{j=1}^{n} f_{a_i} \min\left(d_{a_i b_j}\right) \quad (1)$$

where f_{a_i} is the relative abundance of feature a in community i, n is the number of communities/samples in the dataset, and $\min\left(d_{a_i b_j}\right)$ is the average minimum relational distance (e.g., the distance between tips on a dendrogram—equivalent to phylogenetic distance) of the fixed feature a in the fixed community i to any feature b in all communities j. In these current analyses, we are allowing conspecifics in our calculations, but they can be excluded pending experimental design. In this case, a conspecific feature is one which is present across both halves of a pairwise comparison and the inclusion/exclusion implementation matches that of the *comdistnt* in the picante R package (Kembel et al., 2010). In practice, this metric measures the average minimum distance between a given feature in one community and all other features in other communities. The key departure from the standard βMNTD calculation is that this calculation occurs from a fixed perspective; only one community is compared to all other communities at a single time. As with the traditional βNTI calculation, $\beta MNTD_{feat}$ was also calculated for 999 randomized communities, which were generated by shuffling the tips of the provided dendrogram/phylogenetic tree using the function *taxaShuffle* from the picante R package (Kembel et al., 2010). Additionally, the null $\beta MNTD_{feat}$ calculation has small amounts of phylogenetic noise (e.g., 1×10^{-20}–5×10^{-20}) injected into them to allow for features present across both halves of the pairwise comparison to be included. By combining the null results with our observed results, we can calculate βNTI_{feat}:

$$\beta\text{NTI}_{feat} = \frac{\beta\text{MNTD}_{feat}^{obs} - \overline{\beta\text{MNTD}_{feat}^{null}}}{\beta\text{MNTD}_{feat}^{sd}} \quad (2)$$

where $\beta\text{MNTD}_{feat}^{obs}$ is the observed $\beta MNTD_{feat}$ measurement, $\overline{\beta\text{MNTD}_{feat}^{null}}$ is the average $\beta MNTD_{feat}$ for the null results, and $\beta\text{MNTD}_{feat}^{sd}$ is the SD of $\beta\text{MNTD}_{feat}^{null}$ values.

Estimating Feature-Level Ecological Dynamics

Following the same philosophy underlining the typical βNTI interpretations, βNTI_{feat} seeks to quantify the ecological processes occurring within and across communities. When describing the underlying theory of βNTI_{feat}, we use the term "community" as the general term to describe any assemblage of ecological data (**Table 1**). However, we want to stress that this method can be used on any relational set of data (e.g., microbial communities and DOM assemblages). Unlike βNTI, which evaluates whole communities through space or time, βNTI_{feat} is focused on identifying the contributions to community assembly by individual community members (**Figure 1**). For example, while βNTI is well-suited to identify variable selection as a process driving differences between two communities, it cannot measure which community members may be driving that variable selection. In contrast, βNTI_{feat} has been adapted from βNTI to track the influence of individual features, which significantly contribute either to convergence or divergence.

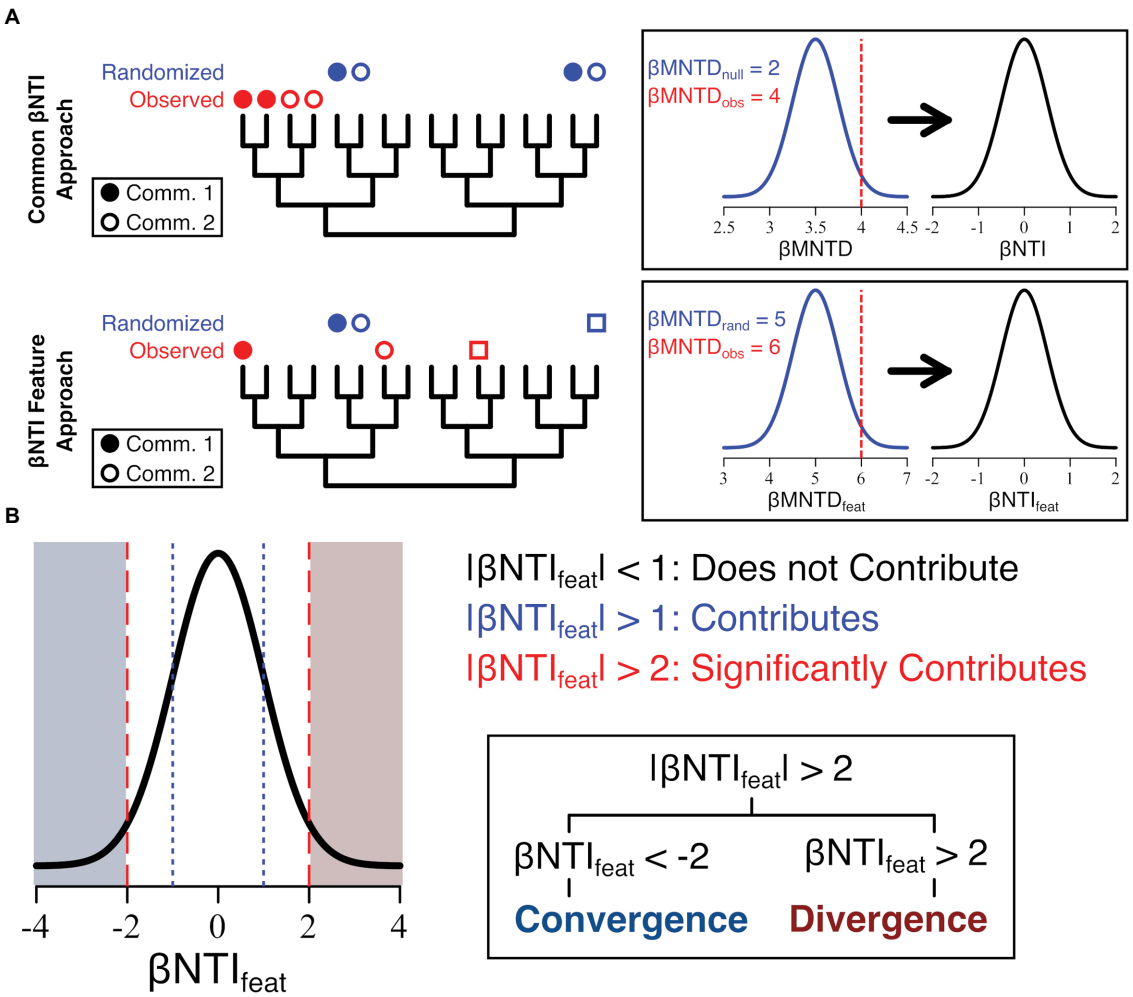


FIGURE 1 | Conceptual depiction of feature-level β-nearest taxon index (βNTI_{feat}) calculation as it compares to the common β-nearest taxon index (βNTI) calculation **(A)** and the subsequent interpretation **(B)**. Essentially, the βNTI_{feat} calculation mirrors βNTI though instead measures the distance between a given focal feature in one community/assemblage (closed circles) and the nearest member in another community/assemblage (open circles and squares).

Feature-level β-nearest taxon index can be calculated in three separate ways in order to account for variations in scale (e.g., a complete experiment vs. sub-groupings within the experiment). First, it can be calculated for an entire dataset yielding a single βNTI_{feat} value for each feature within a dataset. This approach is particularly useful in identifying which features in the whole dataset contribute to convergence or divergence. Second, it can be calculated within groups, for example, if you have a factor with two levels, you can calculate βNTI_{feat} for each of those levels. Results from this type of analysis will generate a βNTI_{feat} value for each feature within each level of the factor (i.e., if there are two factors and 100 features, 200 βNTI_{feat} values will be created) and is well-suited to compare differential feature contribution. Finally, βNTI_{feat} can be calculated truly pairwise (akin to the standard βNTI calculation) in order to provide complete spatial or temporal resolution at the feature level. By performing the βNTI_{feat} analysis on a fixed focal sample (a 0 day dry sample here), you can get a true temporal perspective. Importantly, all these calculations can be done with absolute abundance, relative abundance, or presence/absence data.

The interpretation of the βNTI_{feat} is directly analogous to the interpretation of βNTI but focuses on individual "features" instead. Here, a "feature" is any member of a community or assemblage that is subject to ecological pressures (a microorganism or environmental metabolite, for example; **Table 1**). $|\beta NTI_{feat}| < 1$ means that a feature has an insignificant contribution to ecological variation across the metacommunity or meta-assemblage, $1 < |\beta NTI_{feat}| < 2$ indicates that a feature somewhat contributes ecological variation, and $|\beta NTI_{feat}| > 2$ suggests that a given feature significantly contributes to ecological variation. These patterns can be further resolved based upon the sign of βNTI_{feat}. When βNTI_{feat} trends negative (e.g., <-1), we suggest that the feature contributes to convergence within the scale of analysis. Under this definition, these are features (or groups of related features) that significantly drive relational commonalities across a given

analytical scale. For example, these features could represent a phylogenetically conserved niche of microorganisms or consistent group of molecular formula that are disproportionally impacted by selective processes. When βNTI_{feat} instead trends positive (e.g., >1), we suggest that the feature contributes to divergence. Under this definition, these are features (or groups of related features) that drive relational differences across an analytical scale. These could be those microorganisms which arose to prominence under varied environmental conditions (e.g., they were selected for/against) or those organic matter constituents that were produced/consumed under specific conditions.

Network Analysis

Weighted gene co-expression network analysis (WGCNA) was used to relate βNTI_{feat} results and identify modules of related contributions to ecological assembly within community types (e.g., within the putative active community or molecular formula assemblage) or across community types (e.g., the putatively active community compared to the molecular formula assemblage; WGCNA package v1.70-3; Langfelder and Horvath, 2008). Networks were first generated in R and then visualized/analyzed using Cytoscape v3.8.2 (Shannon et al., 2003). Specifically, βNTI_{feat} values for the putatively active ASVs (as identified *via* cDNA amplicon data) were related to the βNTI_{feat} values for the molecular formulas. As part of this analysis, we identify modules of interconnected features (either ASVs or formula). These modules represent those features which have either positive or negative relationships in their contributions to ecological or functional convergence/divergence. Additional ecological metrics were calculated for modules containing both ASVs and molecular formulas including Shannon's diversity (H; *diversity*, vegan package v2.5-7), exp(H) per recommendation from Jost (2006), Pielou's evenness (J), number of taxonomic Orders within a module, and the number of ASVs in a module (e.g., richness) The relative abundance of elemental composition groups was also calculated for these modules. The Cytoscape network file is stored as **Supplementary File 3**.

Data and Code Availability

Fourier transform ion cyclotron resonance mass spectrometer data are available on the ESS-DIVE archive at https://data.ess-dive.lbl.gov/view/doi:10.15485/1807580, sequence data are available on the NCBI Sequence Read Archive PRJNA641165, and scripts used throughout this manuscript (including various versions of the βNTI_{feat} calculation) can be found on GitHub at https://github.com/danczakre/betaNTI-feature (Sengupta et al., 2021b).

RESULTS AND DISCUSSION

βNTI_{feat} Revealed That Uncultured Microbial Lineages Differentially Contribute to Ecological Processes Across Total and Putatively Active Communities

We first analyzed the overall microbial dynamics and observed that the total community (using 16S rRNA gene sequencing) and putatively active community (*via* transcribed 16S rRNA sequencing) were significantly divergent from each other (**Figures 2A,B**). A Jaccard-based PCoA of Family-level taxonomic assignments revealed that not only were the two community types divergent, but also the communities under cumulatively dry treatments were divergent from those under inundated treatments (DNA – PERMANOVA Pseudo-F: 1.4826, value of $p < .05$; RNA – Pseudo-F: 3.4025, value of $p < .001$; and Type – PERMANOVA Pseudo-F: 38.653, value of $p < .001$; **Figure 2A**). Similar cross treatment dynamics were also apparent once sample similarity was analyzed using phylogenetic relatedness (e.g., βMNTD; DNA – PERMANOVA Pseudo-F: 1.8191, value of $p < .05$; RNA – Pseudo-F: 3.7559, value of $p < .01$; and Type – PERMANOVA Pseudo-F: 69.54, value of $p < .001$; **Figure 2B**). Given that the largest differences existed between the total microbial community and the putatively active component (based upon Pseudo-F statistic for the Type comparison), this information was used as a baseline of differences to interpret downstream analyses.

When compared to existing taxonomic metrics (e.g., SIMPER), βNTI_{feat} allows researchers to identity microbial groups impacting ecological structure rather than composition providing insight into community development (Clarke, 1993). In turn, we can evaluate the degree to which specific microbes contribute to the breadth of ecological strategies contained within metacommunities. This goes far beyond information on variation in taxonomic composition or standard diversity metrics. Existing approaches can identity cross-community differences in the abundance of a given taxon, but cannot identify the assembly processes leading to those differences in abundance. For each taxon (or molecule), βNTI_{feat} quantifies the assembly processes it experiences, the degree to which it influences ecological structure, and elucidates the environmental conditions that lead to variation in these influences and contributions.

Feature-level β-nearest taxon index revealed that varied microbial lineages differentially impacted assembly of the total microbial community and the putatively active microbial community (**Figure 2C**). Looking at the βNTI_{feat} dynamics across both the total and putative active communities, we see that members of an unclassified/uncultured group of Bacteria, members of the Patescibacteria (intermingled with some Class Bacilli), and members of the Nanoarchaeota consistently contribute to ecological divergence (**Figure 2C**). Those taxa which contribute to ecological convergence; however, are less conserved and more widespread across the phylogenetic tree with ASVs appearing within the Crenarchaeota, Alphaproteobacteria, and Gammaproteobacteria (**Figure 2C**).

Given that the largest community differences existed between the total and putatively active communities (**Figures 2A,B**), we focused on evaluating which microbial groups differentially contributed across these groups. We see that those members of the uncultured bacterial group and the Patescibacteria are among the most notably differential. Specifically, while many members appear to contribute to either convergence or divergence across both communities, we see that fewer members of these lineages contribute to

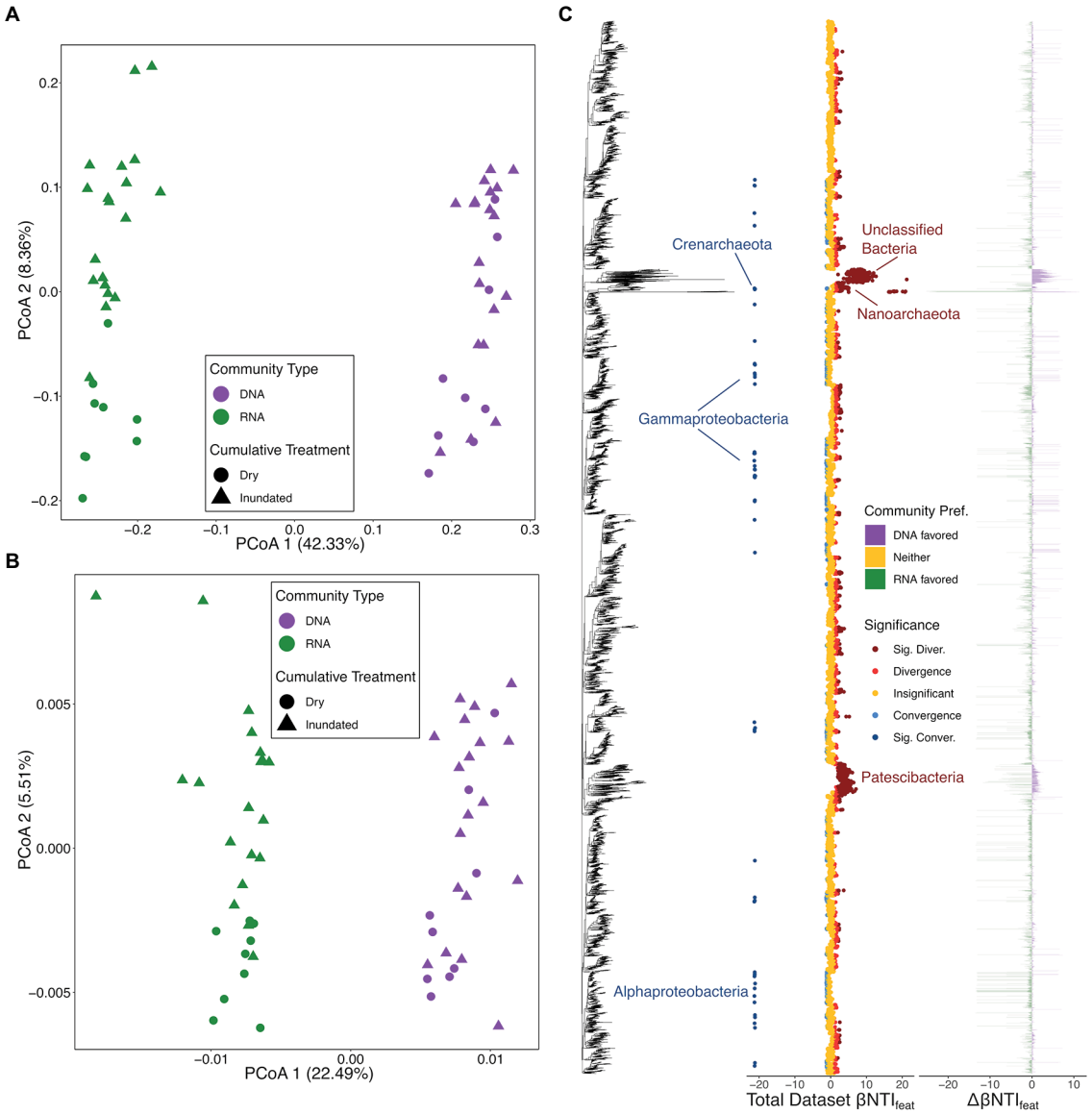


FIGURE 2 | Microbial ordinations and overview of microbial feature-level β-nearest taxon index (βNTI_{feat}) results. **(A)** Bray–Curtis dissimilarity-based principal coordinate analysis (PCoA) with colors representing community type, the total community defined by 16S rRNA gene amplicon results (DNA) and transcribed 16S rRNA gene amplicon results (RNA), and shapes distinguishing treatment type. **(B)** β-mean nearest taxon distance (βMNTD)-based PCoA with a legend matching panel **(B)**. **(C)** βNTI_{feat} results consisting of three panels: the 16S rRNA gene amplicon tree, βNTI_{feat} values for the whole dataset with colors indicating contribution to ecological assembly (Direction), and the difference in βNTI_{feat} values across the total microbial community and putatively active community (Community Pref.). Under the Direction legend, "Insignificant" represent $|\beta NTI_{feat}| < 1$, "Convergence" or "Divergence" indicate $1 < |\beta NTI_{feat}| < 2$, and "Sig. Conver." or "Sig. Diver." represent $|\beta NTI_{feat}| > 2$. Under the Community Pref. legend, "DNA favored" indicates that the given microbial group had a higher absolute βNTI_{feat} value in the total community, "RNA favored" indicates that the given microbial group had a higher absolute βNTI_{feat} value in the putative active community, and "Neither" indicates no difference between communities. Microbial groups discussed throughout the manuscript are called out where applicable.

assembly within the putatively active community. Out of 364 detected Patescibacteria across all microbial data, 305 contributed to assembly ($|\beta NTI_{feat}| > 1$) in the total community, while only 34 impacted the putatively active community (**Figure 2C**). This pattern demonstrates that the uncultured microbial lineages contribute more to the phylogenetic structure of the total community than the active community. The stronger role played by uncultured lineages in the total community may indicate that these taxa have a background role and are not as relevant to the ecology of the active community.

Breaking the data down based upon Family-level taxonomic groups that, on average, contributed to either convergence or divergence supported these broad phylogenetic patterns and we observed that contributions could be related to inferred functional potentials and spanned a range of relative abundances

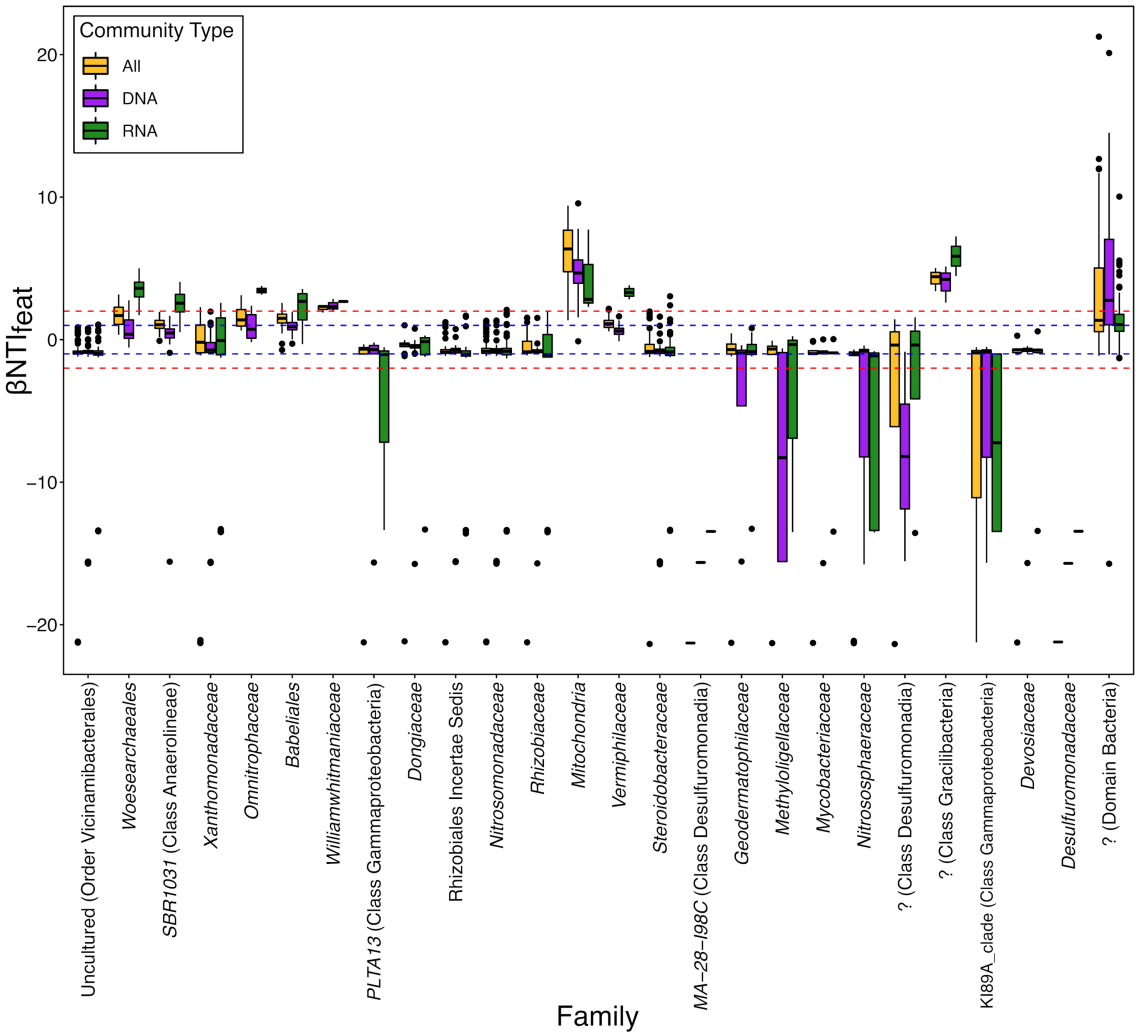


FIGURE 3 | Distribution of feature-level β-nearest taxon index (βNTI_{feat}) values for ASVs belonging to the Family-level taxonomic groups which contribute to either convergence or divergence on average. Some of these taxonomic groups only have a single ASV (e.g., Family MA-28-I98C), whereas other groups have many more (the top two were the unclassified Bacteria with 454 ASVs and the Nitrosomonadaceae with 147 ASVs). "DNA" represents the βNTI_{feat} values for the total community (defined by the 16S rRNA gene amplicon results), "RNA" represents the βNTI_{feat} values for the putatively active community (defined by the transcribed 16S rRNA gene amplicon results), and "All" represents βNTI_{feat} values calculated from the complete community (both DNA and RNA combined). The blue dashed line at +1 and −1 represents the "Contributes" threshold outlined in **Figure 1**, while the red dashed line at +2 and −2 represents the "Significantly Contributes" threshold.

(**Figure 3**; **Supplementary Figure 1**). For example, an unknown family from Class Desulfuromonadia significantly drives convergence ($\beta NTI_{feat} < -2$) in the total community (DNA) but not in either the overall dataset (All) or putatively active (RNA) communities (**Figure 3**). Conversely, the Family *Woesearchaeales* significantly drives ecological divergence ($\beta NTI_{feat} > 2$) within the putatively active community as compared to the total community or overall dataset. In the case of these example taxonomic groups, we hypothesize that this is likely tied to their respective inferred functional potentials. Class Desulfuromonadia tend to feature sulfate reducing bacteria, while described members of Family *Woesearchaeles* are fermentative (Castelle et al., 2015; Waite et al., 2020). Given that the experimental design allowed these sediments to maintain oxic conditions, this would act as a selective force against sulfate reducers due to thermodynamic constraints but not necessarily fermentative organisms (i.e., members could still form syntrophic relationships). This would translate to a diminished presence within the putative active community and thereby prevent them from significantly contributing to either convergence or divergence and suggests that these organisms might be detected in the DNA dataset as either dormant cells or relic DNA (Lennon et al., 2018).

Homologous Series Disproportionately Contribute to Convergence Regardless of Environment Type

The overall DOM patterns largely mirror those of the putatively active microbial community. Namely, both Jaccard- and βMNTD-based analyses revealed that significant differences existed

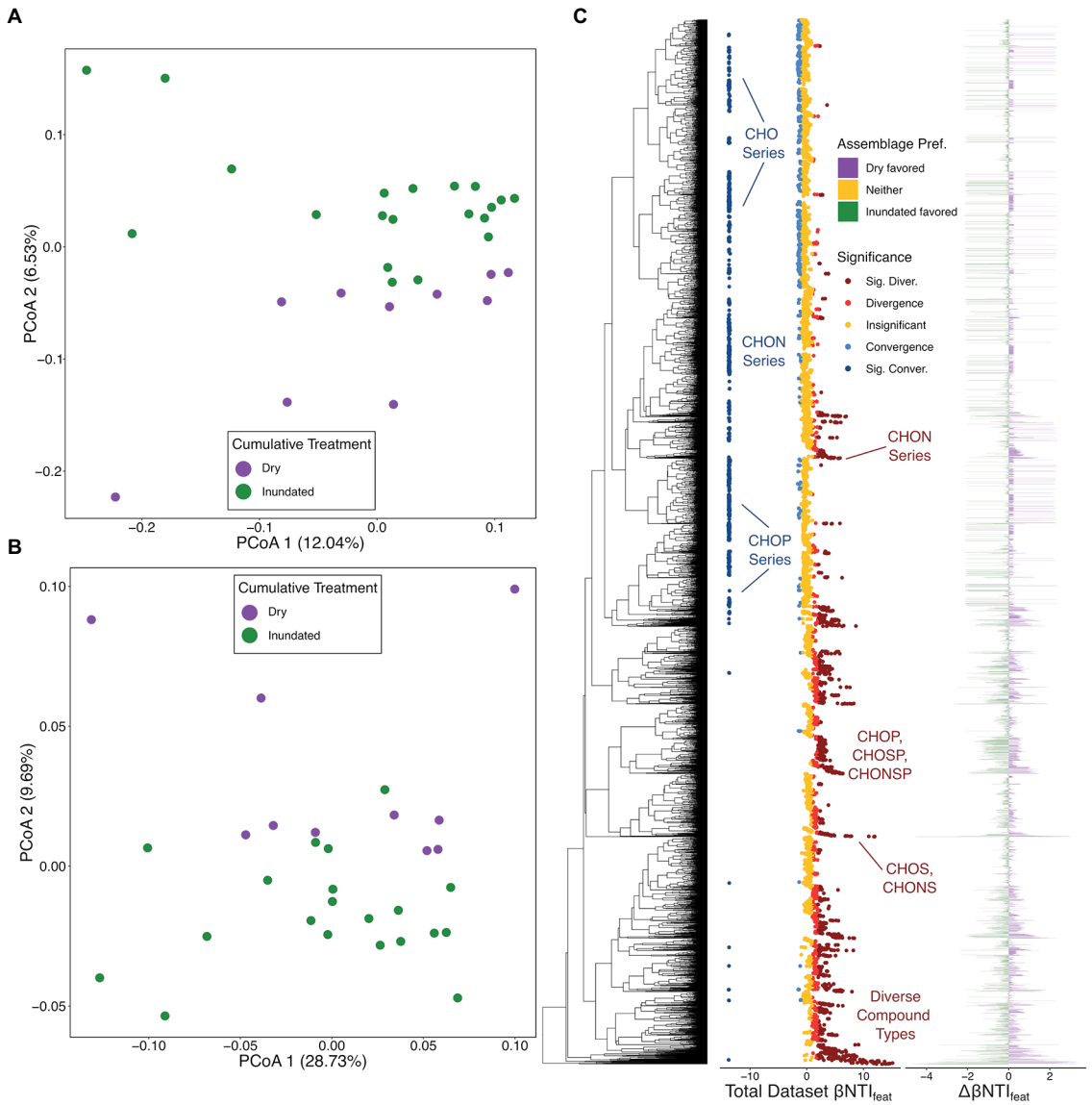


FIGURE 4 | Organic matter (OM) ordinations and overview of OM feature-level β-nearest taxon index (βNTI_{feat}) results. **(A)** Jaccard dissimilarity-based principal coordinate analysis (PCoA) with colors representing treatment type. **(B)** β-mean nearest taxon distance (βMNTD)-based PCoA with a legend matching panel **(B)**. **(C)** βNTI_{feat} results consisting of three panels: the molecular characteristics dendrogram (MCD), βNTI_{feat} values for the whole dataset with colors indicating contribution to ecological assembly (Direction), and the difference in βNTI_{feat} values across the dry-treatment OM assemblage and inundated-treatment OM assemblage (Assemblage Pref.). Under the Direction legend, "Insignificant" represent $|\beta NTI_{feat}| < 1$, "Convergence" or "Divergence" indicate $1 < |\beta NTI_{feat}| < 2$, and "Sig. Conver." or "Sig. Diver." represent $|\beta NTI_{feat}| > 2$. Under the Assemblage Pref. legend, "Dry favored" indicates that aa given molecular formula had a higher absolute βNTI_{feat} value within dry assemblages, "Inundated favored" indicates that a given molecular formula had a higher absolute βNTI_{feat} value in inundated assemblages, and "Neither" indicates no difference between assemblages. Organic matter groups discussed throughout the manuscript are called out where applicable.

between the organic matter from dry and inundated samples (PERMANOVA Pseudo-F: 1.5085, value of $p < .05$; **Figures 4A,B**). Using this information, we calculated βNTI_{feat} for the combined dataset, the dry-only dataset, and the inundated dataset.

Looking first at the complete dataset, we see a higher proportion of features significantly contributing to functional divergence or convergence within the organic matter dataset (45.08%) than in the ASV dataset (20.99%). This potentially points simultaneously to the more transitive nature of many environmental metabolites as compared to microbial community members (Schmidt et al., 2011; Graham et al., 2018; Coward et al., 2019), while also highlighting the consistent role that some conserved organic matter constituents play. Specifically, we see many larger molecular formula (e.g., C47H68O10, C34H63O4P, etc.) and some molecular formula with more complex compositions (e.g., C46H68NO4S2P, C42H52NO9S2P, etc.) significantly contribute to divergence across the entire dataset (**Figure 4C**). This might suggest that these more complex

compounds may play variable roles depending on the situation; they could be a hallmark of differential nutrient limitation (i.e., the functional needs of the meta-assemblage are highly variable; Graham et al., 2018; Garayburu-Caruso et al., 2020; Danczak et al., 2021) or highlight divergent degradation potential (e.g., common metabolites might be processed through different pathways).

Whereas divergence was driven by larger or more complex molecular formula compositions, we see that CHO-based homologous series (e.g., C23H30O14, C24H30O14, etc.), N-containing homologous series (e.g., C18H15NO10, C19H17NO10), and some P-containing homologous series (e.g., C17H25O9P, C18H27O9P) often contribute to convergence across the entire dataset. CHO-only molecular formulas are the most frequently observed types of organic matter from sedimentary and aquatic sources (Garayburu-Caruso et al., 2020) suggesting that these types of compounds could drive functional similarities across organic matter assemblages. As with those compounds which contribute to divergence, the contribution to convergence by the N-containing and P-containing formulas may also point to conserved functional needs or pathways (i.e., N/P-limitation drives selection for N/P-containing formulas, or ongoing metabolisms continually cycling these formulas).

Following the patterns established *via* the Jaccard- and βMNTD-based PCoAs, we evaluated the differences in environmental metabolites which contribute to convergence and divergence across the dry and inundated samples (**Figure 4C**). Unlike the ASV data, where groups contributed differently across the total and putatively active communities, we see a balanced set of contributions from formulas derived from dry and inundated organic matter with no clear discernable patterns.

Grouping molecular formula by either elemental composition categories or compound classes further revealed little consistent variation across wet-dry conditions suggesting that the OM meta-assemblage might be more impacted by ecosystem than treatment condition in this case (**Figure 5**). Similar patterns were observed in a headwater stream, where bulk environmental properties were conserved across ecosystem types despite divergent OM assemblages undergoing variable selection as a result of hypothesized thermodynamic redundancy (Danczak et al., 2021). Briefly, thermodynamic redundancy is like functional redundancy, where compositionally divergent OM assemblages have similar thermodynamic properties. Here, βNTI_{feat} helps us examine whether this hypothesized redundancy exists across groups contributing to community structure. For example, in contrast to the thermodynamic behavior observed in Danczak et al. (2021), we see that the nominal oxidation state of carbon (NOSC) significantly varies across those CHON and CHOP containing molecular formula which contribute to convergence/divergence (Mann–Whitney U test value of $p < .05$; **Supplementary Figure 2**). Specifically, we observe that CHOP formulas with less inferred structural complexity (e.g., lower aromaticity index and double-bond equivalents) and more inferred lability (e.g., lower NOSC) drove significant convergence; differences in the properties of CHON formulas appear to manifest more as variation in distributions (e.g., those formulas which contribute to convergence are more constrained). Formulas featuring other elemental compositions exhibit less consistent behavior. This pattern suggests that thermodynamic restrictions potentially help dictate the contributions of the N/P-containing homologous series to the overall OM assemblage and may indicate that thermodynamic redundancy is not at play in this particular system. In other words, the preference for certain N/P-containing formulas might be dictated by organisms targeting the most preferential carbon source rather than some other limitation.

These results also demonstrate the capability for βNTI_{feat} to uncover functional dynamics within OM assemblages and assign ecological importance in the absence of abundance information.

Network Analyses Reveal Groups of Molecular Formula and Microbes Which Contribute to Ecosystem Dynamics in a Coordinated Manner

While the previous sections focused on either analyzing complete datasets or subgroups within the overall dataset, βNTI_{feat} also provides sample-level information for each feature within a given dataset (**Supplementary Figure 3**). Using this approach, we can evaluate how the contribution of each feature to community assembly changes through either time or space by utilizing a single focal sample (here ECA_0Cyc_R2). Looking at the five most variable taxonomic groups from the putatively active community, we can see that different groups respond differently through treatment regimes (**Supplementary Figure 3**). For example, families DTB120, MVP-15, and order Myxococcocales rarely if ever cross the +2 βNTI_{feat} threshold for contributing to divergence. In contrast, the families *Williamwhitmaniaceae* and *Woesearchaeales* consistently cross that threshold. In contrast to these highly dynamic ASV taxa, the groups defined by a molecular formula's elemental composition are less variable across samples (**Supplementary Figure 3**).

Having established that βNTI_{feat} can track dynamics through time, we believe that this metric is well-suited in network analyses. By performing a WGCNA using βNTI_{feat} values for each feature—as opposed to abundances—we can obtain modules of putatively active ASVs and OM formulas. These modules should reflect features which have either positive or negative relationships among their contributions to ecological/functional convergence and divergence. Within-cluster relationships between ASVs and organic molecules do not indicate co-variance in abundances (as in a traditional network analysis). Instead, these relationships indicate linkages between the ecology of microbes and the functional properties of organic molecules. We propose that these putative ecology-function linkages are more likely to indicate causal connections between microbes and molecules than networks built on relative abundances. There are nonetheless multiple potential interpretations of within-cluster relationships, leading to additional questions. For example, could the ASVs be responsible for degrading the linked molecular formulas?

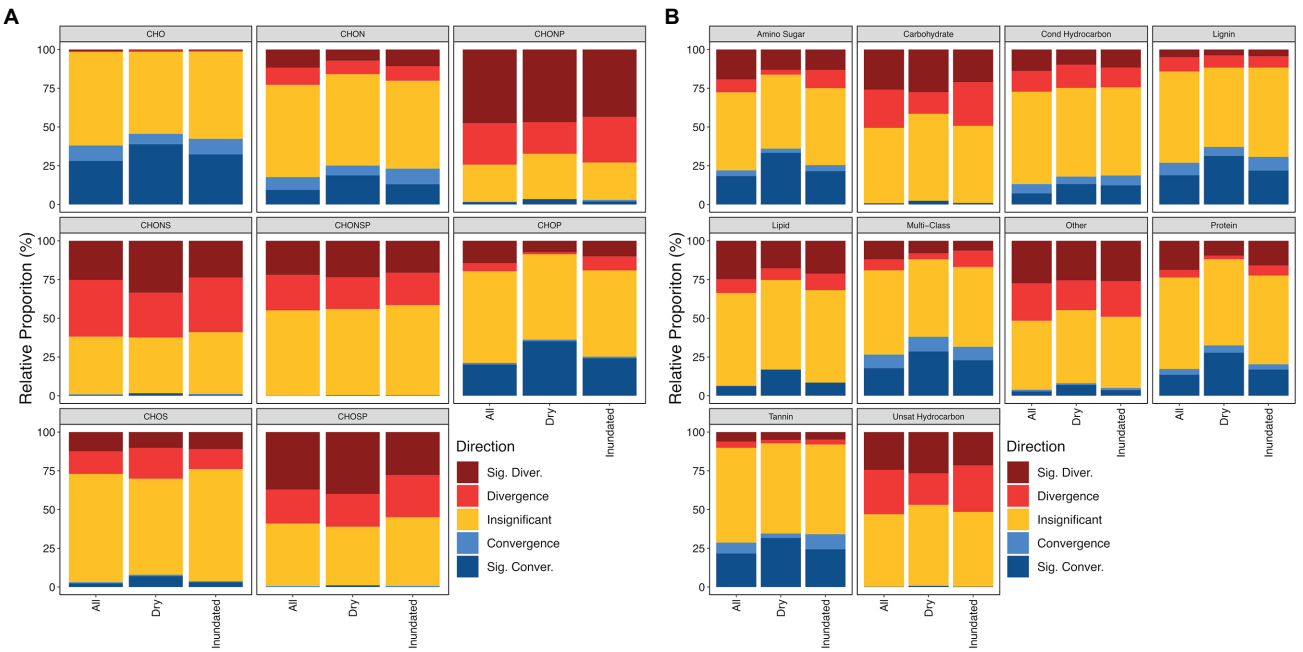


FIGURE 5 | Proportion of contributions across different elemental composition groups **(A)** and compound classes **(B)**. These graphs outline the frequency that we observe a given contribution type within a given grouping. For example, out of all CHO molecular formulas within dry assemblages, we see that they contribute to convergence more frequently than CHONSP molecular formulas.

If a module contains exclusively molecular formulas, does this point a single pathway or do they represent an ecological interaction among pathways? By revealing features that have coordinated impacts on the ecological or functional structure of their respective community, network analyses can highlight important organisms or molecules. A better understanding of the relationships among ecological pressures impacting members across communities within a metacommunity will enable researchers to develop transitive principles regarding the coordination of ecological assembly (e.g., the mechanisms impacting cross trophic level assembly, etc.).

The network analysis resulted in the creation of 54 modules, with 28 modules having some mixture of ASVs and organic matter, two featuring exclusively ASVs, and 24 containing only molecular formulas (**Figure 6**; **Supplementary Figure 4**). Of these 54 modules, two modules were removed because the consisted only of doublets (e.g., modules where one feature related to only one other single feature). Focusing on the 27 modules which have a mixture of ASVs and molecular formulas, we observe a range of microbial diversity: for example, we observe that the gray and tan modules have the greatest number of different taxonomic groups (**Supplementary Figure 4A**). These modules also exhibit a range of characteristics across their molecular formulas: we see that the gray module predominantly contains CHO-only formulas, while the tan module primarily contains CHON formulas (**Supplementary Figure 4B**). This approach now allows us to evaluate which factors and taxonomic groups are coordinated in driving ecological dynamics.

Given the importance of nitrogen and phosphorus homologous uncovered in the metabolite βNTI_{feat} analyses, we examined modules containing at least two N- or P-containing formulas and contained more than two nodes. There were 16 modules with two or more N-containing molecular formulas, 10 modules with two or more P-containing formulas, and nine modules with both two or more N- and P-containing formulas (**Figure 6**). Analyzing the membership of these modules first revealed that the P-containing modules only featured four ASVs across all 10 modules, with three modules comprised exclusively of P-containing modules. This pattern suggests that these P-containing molecular formulas potentially experience a set of selective pressures distinct from other metabolites and ASVs. This characteristic could explain the disproportionate effect that phosphorus homologous series had on the functional structure of the OM meta-assemblage. In contrast, N-containing modules featured broader microbial and molecular diversity indicating that N-containing molecular formulas experience a more common set of pressures than P-containing formulas. As hypothesized above, this might be driven by nutritional requirements or some sort of pathway regulation. Finally, five different ASVs from Family *Geobacteraceae* were detected in three out of the nine modules featuring both N- and P-containing molecular formulas. While we do not have direct functional information due to the limits of amplicon sequencing, members of this family have reported roles in both the nitrogen cycle (*via* nitrate reduction, nitrification, and nitrogen fixation) and the phosphorus cycle (*via* aggressive phosphate acquisition; Naik et al., 1993; N'Guessan et al., 2010; Ueki and Lovley, 2010;

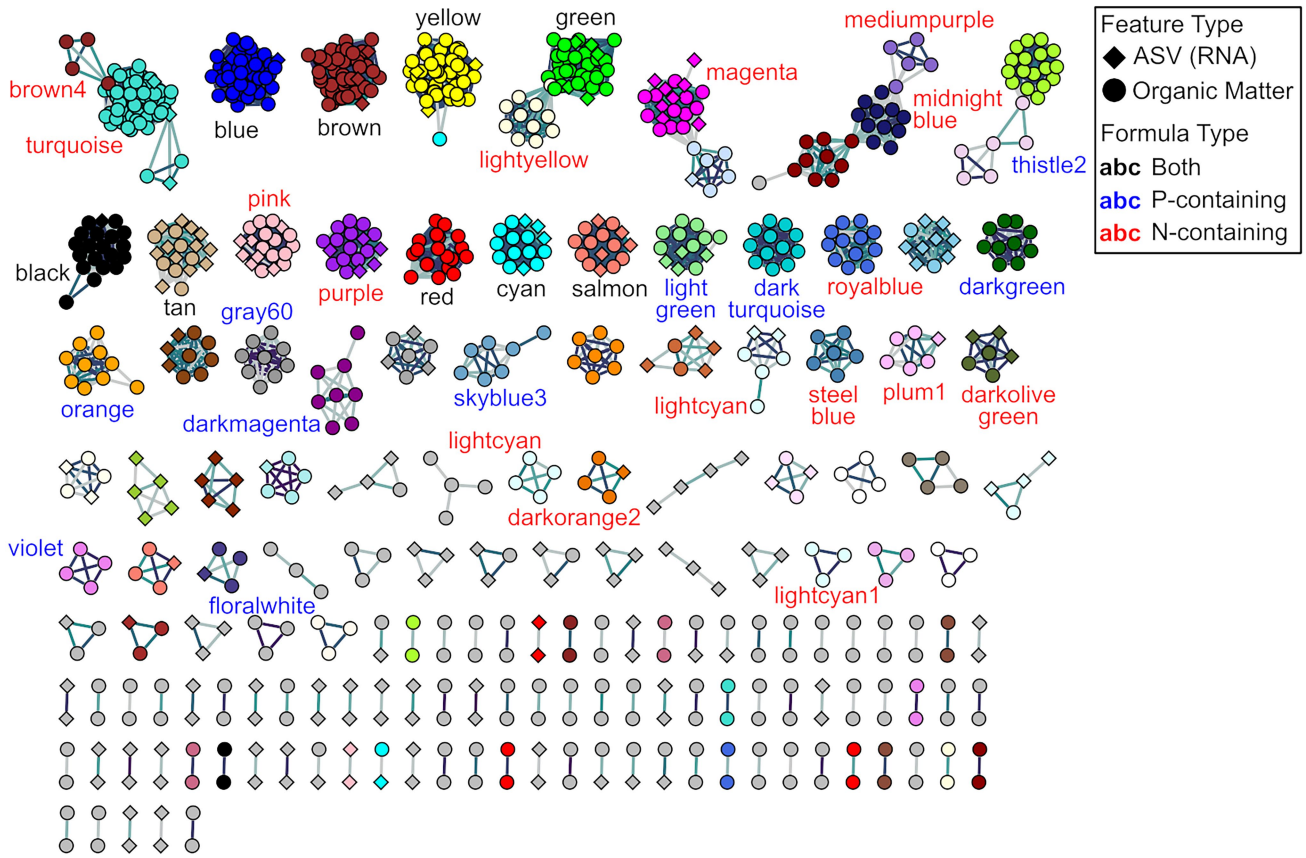


FIGURE 6 | Feature-level β-nearest taxon index (βNTI_{feat})-based weighted gene co-expression network analysis (WGCNA). Modules with labels represent those modules which contain N- and/or P-containing molecular formulas. Node colors indicate the different modules identified as part of the WGCNA pipeline, node shapes represent different feature types, and label colors indicate whether the modules contain either N- and/or P-containing molecular formulas.

Wang et al., 2018; Samaddar et al., 2019; Campeciño et al., 2020). Taken together, this potentially points to a combinatorial effect, whereby the functional development driven by N/P-containing formulas is tied to organismal selectivity (i.e., roles of *Geobacteraceae* in the N/P cycles).

CONCLUSION

Feature-level β-nearest taxon index is a novel metric that enables researchers to investigate the contributions to ecological convergence and divergence with a given metacommunity or meta-assemblage. Using βNTI_{feat}, we revealed that uncultured lineages contribute significantly to the ecological structure of microbial communities when assayed using 16S rDNA, but contributed little to the ecology of putatively actives communities assayed with 16S rRNA. These differences suggest that while these uncultured lineages represent a significant proportion of the microorganisms in some ecosystems (Brown et al., 2015; Anantharaman et al., 2016; Danczak et al., 2017; León-Zayas et al., 2017), they may be relatively minor contributors to the ecological composition of active microbes in the study system. This inference is complementary to traditional inferences based on relative abundances or taxonomic richness. That is, βNTI_{feat} quantifies the contribution of individual ASVs to the assembly of multi-dimensional ecological niche space occupied by communities, and how those contributions vary through space, time, and environmental conditions.

Applying βNTI_{feat} to OM provides analogous inferences, but instead quantifies the contribution of individual molecules to the assembly of the multi-dimensional functional space occupied by assemblages of organic molecules. In our study, we observed that OM homologous series primarily containing nitrogen or phosphorus differentially contributed to OM functional composition across the whole dataset. However, we do not see any noteworthy differences across wet-dry treatments indicating a stronger ecosystem-level control. Network analyses relating putatively active ASV βNTI_{feat} values to metabolite βNTI_{feat} values revealed formation of 54 distinct modules exhibiting coordinated contributions to their respective communities. Looking specifically at modules featuring N/P-containing molecular formulas, we observed that P-containing formulas appear to be largely disconnected from other metabolites and microorganisms, N-containing formulas are more integrated across feature types, and modules with both formula types exhibit a high incidence of *Geobacteraceae* members.

These dynamics potentially point to two larger scale hypotheses. First, the ecological contributions of P-containing formulas are often distinct from other metabolites indicating that they might occupy a unique functional space (at least compared to N-containing formulas). Second, certain microbial groups may offer contributions to convergence or divergence in concert with specific metabolite dynamics (here *Geobacteraceae* related to N/P-containing formulas). While these hypotheses still need to be independently verified, they point to specific conclusions that can be drawn from the use of βNTI_{feat}. As further research uses this metric, more broad scale, transferrable hypotheses can be generated and help us better understand community assembly.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: FTICR-MS data are available on the ESS-DIVE archive at: https://data.ess-dive.lbl.gov/view/doi:10.15485/1807580 and sequence data are available on the NCBI Sequence Read Archive PRJNA641165.

AUTHOR CONTRIBUTIONS

RD and JS conceptualized and designed the study. AS assisted in experimental design. JT, JW, LR, RC, SF, and VG-C performed the experiment. RD analyzed the data. RD and JS drafted the manuscript and all authors contributed to further writing. All authors contributed to the article and approved the submitted version.

FUNDING

The initial experimental stages of this work were supported by the PREMIS Initiative at the Pacific Northwest National Laboratory (PNNL) with funding from the Laboratory Directed Research and Development Program at PNNL, a multi-program national laboratory operated by Battelle for the United States Department of Energy under Contract DE-AC05-76RL01830. The later stages of this work (e.g., data analysis, conceptual interpretation manuscript development) were supported by the United States Department of Energy-BER program, as part of an Early Career Award to JS at PNNL. A portion of the research was performed using EMSL, a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.803420/full#supplementary-material

REFERENCES

Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., et al. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* 7:13219. doi: 10.1038/ncomms13219

Bailey, V. L., Smith, A. P., Tfaily, M., Fansler, S. J., and Bond-Lamberty, B. (2017). Differences in soluble organic carbon chemistry in pore waters sampled from different pore size domains. *Soil Biol. Biochem.* 107, 133–143. doi: 10.1016/j.soilbio.2016.11.025

Blomberg, S. P., Garland, T., and Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavorial traits are more labile. *Evolution* 57, 717–745. doi: 10.1111/j.0014-3820.2003.tb00285.x

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., et al. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:90. doi: 10.1186/s40168-018-0470-z

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bottos, E. M., Kennedy, D. W., Romero, E. B., Fansler, S. J., Brown, J. M., Bramer, L. M., et al. (2018). Dispersal limitation and thermodynamic constraints govern spatial structure of permafrost microbial communities. *FEMS Microbiol. Ecol.* 94:fiy110. doi: 10.1093/femsec/fiy110

Bramer, L. M., White, A. M., Stratton, K. G., Thompson, A. M., Claborne, D., Hofmockel, K., et al. (2020). FtmsRanalysis: an R package for exploratory data analysis and interactive visualization of FT-MS data. *PLoS Comput. Biol.* 16:e1007654. doi: 10.1371/journal.pcbi.1007654

Brown, C. T., Hug, L. A., Thomas, B. C., Sharon, I., Castelle, C. J., Singh, A., et al. (2015). Unusual biology across a group comprising more than 15% of domain bacteria. *Nature* 523, 208–211. doi: 10.1038/nature14486

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

Campeciño, J., Lagishetty, S., Wawrzak, Z., Alfaro, V. S., Lehnert, N., Reguera, G., et al. (2020). Cytochrome c nitrite reductase from the bacterium *Geobacter lovleyi* represents a new NrfA subclass. *J. Biol. Chem.* 295, 11455–11465. doi: 10.1074/jbc.RA120.013981

Castelle, C. J., Wrighton, K. C., Thomas, B. C., Hug, L. A., Brown, C. T., Wilkins, M. J., et al. (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr. Biol.* 25, 690–701. doi: 10.1016/j.cub.2015.01.014

Chase, J. M., and Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc. Lnod. B Biol. Sci.* 366, 2351–2363. doi: 10.1098/rstb.2011.0063

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral. Ecol.* 18, 117–143. doi: 10.1111/j.1442-9993.1993.tb00438.x

Coward, E. K., Ohno, T., and Sparks, D. L. (2019). Direct evidence for temporal molecular fractionation of dissolved organic matter at the iron oxyhydroxide interface. *Environ. Sci. Technol.* 53, 642–650. doi: 10.1021/acs.est.8b04687

Danczak, R. E., Chu, R. K., Fansler, S. J., Goldman, A. E., Graham, E. B., Tfaily, M. M., et al. (2020a). Using metacommunity ecology to understand environmental metabolomes. *Nat. Commun.* 11:6369. doi: 10.1038/s41467-020-19989-y

Danczak, R. E., Daly, R. A., Borton, M. A., Stegen, J. C., Roux, S., Wrighton, K. C., et al. (2020b). Ecological assembly processes are coordinated between bacterial and viral communities in fractured shale ecosystems. *mSystems* 5, e00098–e00120. doi: 10.1128/mSystems.00098-20

Danczak, R. E., Goldman, A. E., Chu, R. K., Toyoda, J. G., Garayburu-Caruso, V. A., Tolić, N., et al. (2021). Ecological theory applied to environmental metabolomes reveals compositional divergence despite conserved molecular properties. *Sci. Total Environ.* 788:147409. doi: 10.1016/j.scitotenv.2021.147409

Danczak, R. E., Johnston, M. D., Kenah, C., Slattery, M., Wrighton, K. C., and Wilkins, M. J. (2017). Members of the candidate phyla radiation are functionally differentiated by carbon- and nitrogen-cycling capabilities. *Microbiome* 5:112. doi: 10.1186/s40168-017-0331-1

Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G. (2008). A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods* 6, 230–235. doi: 10.4319/lom.2008.6.230

Fodelianakis, S., Washburne, A. D., Bourquin, M., Pramateftaki, P., Kohler, T. J., Styllas, M., et al. (2021). Microdiversity characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome. ISME J. doi:10.1038/s41396-021-01106-6 [Epub ahead of print].

Garayburu-Caruso, V. A., Danczak, R. E., Stegen, J. C., Renteria, L., Mccall, M., Goldman, A. E., et al. (2020). Using community science to reveal the global chemogeography of river metabolomes. *Metabolites* 10:518. doi: 10.3390/metabo10120518

George, D. B., Webb, C. T., Farnsworth, M. L., O'Shea, T. J., Bowen, R. A., Smith, D. L., et al. (2011). Host and viral ecology determine bat rabies seasonality and maintenance. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10208–10213. doi: 10.1073/pnas.1010875108

Gilbert, B., and Bennett, J. R. (2010). Partitioning variation in ecological communities: do the numbers add up? *J. Appl. Ecol.* 47, 1071–1082. doi: 10.1111/j.1365-2664.2010.01861.x

Graham, E. B., Crump, A. R., Kennedy, D. W., Arntzen, E., Fansler, S., Purvine, S. O., et al. (2018). Multi 'omics comparison reveals metabolome biochemistry, not microbiome composition or gene expression, corresponds to elevated biogeochemical function in the hyporheic zone. *Sci. Total Environ.* 642, 742–753. doi: 10.1016/j.scitotenv.2018.05.256

Herren, C. M., and McMahon, K. D. (2017). Cohesion: a method for quantifying the connectivity of microbial communities. *ISME J.* 11, 2426–2438. doi: 10.1038/ismej.2017.91

Hughey, C. A., Hendrickson, C. L., Rodgers, R. P., Marshall, A. G., and Qian, K. (2001). Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal. Chem.* 73, 4676–4681. doi: 10.1021/ac010560w

Jiao, S., Yang, Y., Xu, Y., Zhang, J., and Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J.* 14, 202–216. doi: 10.1038/s41396-019-0522-9

Jost, L. (2006). Entropy and diversity. *Oikos* 113, 363–375. doi: 10.1111/j.2006.0030-1299.14714.x

Katoh, K. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kim, S., Kramer, R. W., and Hatcher, P. G. (2003). Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van krevelen diagram. *Anal. Chem.* 75, 5336–5344. doi: 10.1021/ac034415p

Koch, B. P., and Dittmar, T. (2006). From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20, 926–932. doi: 10.1002/rcm.2386

Kraft, N. J. B., Cornwell, W. K., Webb, C. O., and Ackerly, D. D. (2007). Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *Am. Nat.* 170, 271–283. doi: 10.1086/519400

Kujawinski, E. B., and Behn, M. D. (2006). Automated analysis of electrospray ionization fourier transform ion cyclotron resonance mass spectra of natural organic matter. *Anal. Chem.* 78, 4363–4373. doi: 10.1021/ac0600306

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

LaRowe, D. E., and Van Cappellen, P. (2011). Degradation of natural organic matter: a thermodynamic analysis. *Geochim. Cosmochim. Acta* 75, 2030–2042. doi: 10.1016/j.gca.2011.01.020

Lennon, J. T., Muscarella, M. E., Placella, S. A., and Lehmkuhl, B. K. (2018). How, when, and where relic DNA affects microbial diversity. *mBio* 9, e00637–e00718. doi: 10.1128/mBio.00637-18

León-Zayas, R., Peoples, L., Biddle, J. F., Podell, S., Novotny, M., Cameron, J., et al. (2017). The metabolic potential of the single cell genomes obtained from the challenger deep, mariana trench within the candidate superphylum Parcubacteria (OD1). *Environ. Microbiol.* 19, 2769–2784. doi: 10.1111/1462-2920.13789

Moroenyane, I., Mendes, L., Tremblay, J., Tripathi, B., and Yergeau, E. (2021). Plant compartments and developmental stages modulate the balance between niche-based and neutral processes in soybean microbiome. *Microb. Ecol.* 82, 416–428. doi: 10.1007/s00248-021-01688-w

N'Guessan, A. L., Elifantz, H., Nevin, K. P., Mouser, P. J., Methé, B., Woodard, T. L., et al. (2010). Molecular analysis of phosphate limitation in Geobacteraceae during the bioremediation of a uranium-contaminated aquifer. *ISME J.* 4, 253–266. doi: 10.1038/ismej.2009.115

Naik, R. R., Francisco, M. M., and Stolz, J. F. (1993). Evidence for a novel nitrate reductase in the dissimilatory iron-reducing bacterium *Geobacter metallireducens*. *FEMS Microbiol. Lett.* 106, 53–58. doi: 10.1111/j.1574-6968.1993.tb05934.x

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., et al. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* 11:4717. doi: 10.1038/s41467-020-18560-z

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). Vegan: Community Ecology Package. Available at: https://cran.r-project.org/package=vegan (Accessed August 15, 2021).

Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528. doi: 10.1093/bioinformatics/bty633

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Rivas-Ubach, A., Liu, Y., Bianchi, T. S., Tolić, N., Jansson, C., and Paša-Tolić, L. (2018). Moving beyond the van krevelen diagram: a new stoichiometric approach for compound classification in organisms. *Anal. Chem.* 90, 6152–6160. doi: 10.1021/acs.analchem.8b00529

Samaddar, S., Chatterjee, P., Truu, J., Anandham, R., Kim, S., and Sa, T. (2019). Long-term phosphorus limitation changes the bacterial community structure and functioning in paddy soils. *Appl. Soil Ecol.* 134, 111–115. doi: 10.1016/j.apsoil.2018.10.016

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I. A., et al. (2011). Persistence of soil organic matter as an ecosystem property. *Nature* 478, 49–56. doi: 10.1038/nature10386

Sengupta, A., Fansler, S. J., Chu, R. K., Danczak, R. E., Garayburu-Caruso, V. A., Renteria, L., et al. (2021a). Disturbance triggers non-linear microbe–environment feedbacks. *Biogeosciences* 18, 4773–4789. doi: 10.5194/bg-18-4773-2021

Sengupta, A., Fansler, S. J., Toyoda, J. G., Chu, R. K., Garayburu-Caruso, V. A., Renteria, L., et al. (2021b). Respiration data, microbial community assembly data, and FTICR-MS data associated with: "Disturbance Triggers Non-Linear Microbe-Environment Feedbacks. Sengupta et al., 2021." *Biogeosciences* doi: 10.15485/1807580

Shannon, P., Markiel, A., Ozier, O., Baliga, N., Wang, J., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Smith, T. W., and Lundholm, J. T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography* 33, 648–655. doi: 10.1111/j.1600-0587.2009.06105.x

Stegen, J. C., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E., Nelson, W. C., Arntzen, E. V., et al. (2016). Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. *Nat. Commun.* 7:11237. doi: 10.1038/ncomms11237

Stegen, J. C., Johnson, T., Fredrickson, J. K., Wilkins, M. J., Konopka, A. E., Nelson, W. C., et al. (2018). Influences of organic carbon speciation on hyporheic corridor biogeochemistry and microbial ecology. *Nat. Commun.* 9:585. doi: 10.1038/s41467-018-02922-9

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Swenson, N. G., Enquist, B. J., Pither, J., Thompson, J., and Zimmerman, J. K. (2006). The problem and promise of scale dependency in community phylogenetics. *Ecology* 87, 2418–2424. doi: 10.1890/0012-9658(2006)87[2418:TPAPOS]2.0.CO;2

Tolić, N., Liu, Y., Liyu, A., Shen, Y., Tfaily, M. M., Kujawinski, E. B., et al. (2017). Formularity: software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra. *Anal. Chem.* 89, 12659–12665. doi: 10.1021/acs.analchem.7b03318

Ueki, T., and Lovley, D. R. (2010). Novel regulatory cascades controlling expression of nitrogen-fixation genes in *Geobacter sulfurreducens*. *Nucleic Acids Res.* 38, 7485–7499. doi: 10.1093/nar/gkq652

Waite, D. W., Chuvochina, M., Pelikan, C., Parks, D. H., Yilmaz, P., Wagner, M., et al. (2020). Proposal to reclassify the proteobacterial classes Deltaproteobacteria and Oligoflexia, and the phylum Thermodesulfobacteria into four phyla reflecting major functional capabilities. *Int. J. Syst. Evol. Microbiol.* 70, 5972–6016. doi: 10.1099/ijsem.0.004213

Wang, S., An, J., Wan, Y., Du, Q., Wang, X., Cheng, X., et al. (2018). Phosphorus competition in bioinduced vivianite recovery from wastewater. *Environ. Sci. Technol.* 52, 13863–13870. doi: 10.1021/acs.est.8b03022

Warton, D. I., Wright, S. T., and Wang, Y. (2012). Distance-based multivariate analyses confound location and dispersion effects. *Methods Ecol. Evol.* 3, 89–101. doi: 10.1111/j.2041-210X.2011.00127.x

Zhang, K., Shi, Y., Cui, X., Yue, P., Li, K., Liu, X., et al. (2019). Salinity is a key determinant for soil microbial communities in a desert ecosystem. *mSystems* 4, e00225–e00318. doi: 10.1128/msystems.00225-18

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81, e00002–e00017. doi: 10.1128/MMBR.00002-17

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.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 March 2022
doi: 10.3389/fmicb.2021.789983

Microbiome Transmission During Sexual Intercourse Appears Stochastic and Supports the Red Queen Hypothesis

Zhanshan (Sam) Ma[1,2]*

[1] Computational Biology and Medical Ecology Lab, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences (CAS), Kunming, China, [2] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences (CAS), Kunming, China

Microbes inhabit virtually everywhere on and/or in our bodies, including the seminal and vaginal fluids. They have significant importance in maintaining reproductive health and protecting hosts from diseases. The exchange of microbes during sexual intercourse is one of the most direct and significant microbial transmissions between men and women. Nevertheless, the mechanism of this microbial transmission was little known. Is the transmission mode stochastic, passive diffusion similar to the random walk of particles, or driven by some deterministic forces? What is the microbial transmission probability? What are the possible evolutionary implications, particularly from the perspective of sexual reproduction (selection)? We tackle these intriguing questions by leveraging the power of Hubbell's unified neutral theory of biodiversity, specifically implemented as the HDP-MSN (hierarchical Dirichlet process approximated multi-site neutral model), which allows for constructing truly multi-site metacommunity models, *simultaneously* including vaginal and semen microbiomes. By reanalyzing the microbiome datasets of seminal and vaginal fluids from 23 couples both before and after sexual intercourses originally reported by Mändar and colleagues, we found that the microbial transmission between seminal and vaginal fluids is a stochastic, passive diffusion similar to the random walk of particles in physics, rather than driven by deterministic forces. The transmission probability through sexual intercourse seems to be approximately 0.05. Inspired by the results from the HDP-MSN model, we further conjecture that the stochastic drifts of microbiome transmissions during sexual intercourses can be responsible for the homogeneity between semen and vaginal microbiomes first identified in a previous study, which should be helpful for sexual reproduction by facilitating the sperm movement/survival and/or egg fertilization. This inference seems to be consistent with the classic Red Queen hypothesis, which, when extended to the co-evolutionary interactions between humans and their symbiotic microbiomes, would predict that the reproductive system microbiomes should support sexual reproduction.

Keywords: microbiome transmission, semen microbiome, vaginal microbiome, neutral theory, multi-site neutral model (MSN), red queen hypothesis, coevolution

Abbreviations: BV, bacterial vaginosis; BVAB, BV associated anaerobic bacteria; CM, semen sample; CNA, vaginal sample before intercourse; CNB, vaginal sample after intercourse; HDP, hierarchical Dirichlet process; MSN, multi-site neutral model; HDP-MSN, hierarchical Dirichlet process approximation to multisite neutral model; SAD, species abundance distribution; UNTB, unified neutral theory of biodiversity.

OPEN ACCESS

Edited by:
Stilianos Fodelianakis, Swiss Federal Institute of Technology Lausanne, Switzerland

Reviewed by:
Osiris Gaona, National Autonomous University of Mexico, Mexico
Catherine Putonti, Loyola University Chicago, United States

***Correspondence:**
Zhanshan (Sam) Ma
ma@vandals.uidaho.edu

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *05 October 2021*
Accepted: *28 December 2021*
Published: *08 March 2022*

Citation:
Ma ZS (2022) Microbiome Transmission During Sexual Intercourse Appears Stochastic and Supports the Red Queen Hypothesis. Front. Microbiol. 12:789983. doi: 10.3389/fmicb.2021.789983

INTRODUCTION

Microbes inhabit virtually every corner of our body, including semen, male and female genital tracts. The genital microbiome have great importance in maintaining reproductive health and protecting hosts from disease (Ravel et al., 2011; Gajer et al., 2012; Macklaim et al., 2013). Studies show that the dysbiosis of vaginal microbiota is closely linked to an increased risk of certain diseases, such as bacterial vaginosis (BV) and sexually transmitted infections (e.g., Ma et al., 2012; Lewis et al., 2017; Smith and Ravel, 2017; van de Wijgert, 2017). Although the microbiome in the male genital tract exists primarily in the urethra and coronary sulcus, researchers typically use semen to study the microbiome of the male genital tract. Semen microbiome has been found to play a critical role in semen quality that is associated with male fecundity (Hou et al., 2013; Weng et al., 2014). In addition, semen can be a major vector for the sexual transmission of pathogens including HIV (Liu et al., 2015).

Multiple factors may influence the composition of genital-associated microbiota, including race, age, lifestyle, and sexual activity (Lewis et al., 2017; Noyes et al., 2018). During sexual intercourse, the genital microbiome can be exchanged between sexual partners, and the exchange may have significant influences on the vaginal microbiome, and to a less extent on the semen microbiome (Starnbach and Roan, 2008; Nelson et al., 2012; Liu et al., 2015; Zozaya et al., 2016; Plummer et al., 2018). Mändar et al. (2015) investigated the genital tract microbiota of 23 couples before and after intercourse, and postulated that there was association between semen and vaginal microbiomes. Their study revealed that the seminal microbiome caused the significant decrease in the relative abundance of *Lactobacillus crispatus* after intercourse, and *Gardnerella vaginalis* tend to dominate the vaginal communities of the women whose partners had leukocytospermia (Mändar et al., 2015). Vodstrcil et al.'s (2017) longitudinal sampling of the vaginal microbiome of 52 young women also revealed that penile-vaginal sex changed the vaginal communities into the *Gardnerella vaginalis* dominated microbiome. In spite of these apparently dramatic changes that occurred in vaginal microbiome after sexual intercourse, the relatively long term effects of the intercourse may be limited because of the resilience of normal vaginal microbiota (Borovkova et al., 2011). In addition, the evolutionary implications of the microbiome transmission *via* sexual intercourse are still little known (Ma and Taylor, 2020).

Existing literature on the influence of sexual intercourse on vaginal microbiome clearly highlights its healthy implications for woman (Starnbach and Roan, 2008; Ma et al., 2012; Nelson et al., 2012; Liu et al., 2015; Zozaya et al., 2016; Plummer et al., 2018). Nevertheless, existing studies ignored one important aspect, i.e., what is the microbial transmission (transfer) mechanism during the sexual intercourse? Is it stochastic, passive diffusion similar to the random walk of particles in physics, or driven by some deterministic forces? Is it possible to get rational estimation of the transmission probability and/or the portion of transmitted microbes? Indeed, it may not be practical to obtain such quantifications through experimental or observational studies. Fortunately, it is possible to get rational estimation for such important parameters through mathematical analysis based on the neutral theory of biodiversity (Hubbell, 2001; Li and Ma, 2016; Harris et al., 2017; Ma and Li, 2019; Ma, 2020, 2021a,b). In the present study, we apply Hubbell's (2001) unified neutral theory of biodiversity (UNTB), specifically the multi-site neutral model (MSN) implemented by Harris et al. (2017) to address the previously identified questions. The neutral theory enables us to determine whether or not the transmission of bacteria during the intercourse is a stochastic event similar to random walk of particles in physics or it is simply deterministic. It also allows for us to get rational estimation for the transmission probability and transmission level. We applied the neutral theory modeling by reanalyzing the microbiome (16s-rRNA) datasets of 23 couples originally reported by Mändar et al. (2015), which constitutes the first objective of the present study—investigating the mechanism of microbiome transmission during the sexual intercourse.

A secondary objective of the present study is to explore the evolutionary implications of the microbiome transmission during the sexual intercourse, which has been rarely addressed in existing literature. For example, one may wonder what are their potential evolutionary implications to the sexual reproduction? Specifically, would the microbiome exchange raise or lower the fitness of sexual reproduction? Indeed, one of the major mysteries of evolutionary biology is why costly sexual reproduction is evolved and maintained, whereas the apparently high efficiency of asexual reproduction is also compelling. That is, why and how would sexual reproduction still be evolved in organisms given the apparently compelling advantage of asexual reproduction, and the mystery has been known as sexual selection problem in literature. The Red Queen hypothesis (Van Valen, 1973; Žliobaitė et al., 2017; Scoville, 2019) has been one of the most favored theories to explain the evolution of sexual reproduction, i.e., a theory for the sexual selection problem. Multiple versions of Red Queen hypothesis have been developed in evolutionary biology. Arguably the most well-known version is the co-evolutionary or arms-race interactions between species (particularly the predator-prey system), in which both the predator and prey must continuously adapt to each other's innovative, and advantageous mutations to "out-compete" each other, such that neither go extinct and both survive and prosper. According to the Red Queen hypothesis, this arms race or back-and-forth co-evolution of the species is a continuous co-adaptation process over long evolutionary timelines. In the domain of sexual selection, according to the Red Queen theory, sexual reproduction, in which mate can be selected rather than undergoing "closed" and non-selective asexual reproduction, allows for selecting a partner with advantageous characteristics and is therefore more likely to produce offspring better fit for the environment (Scoville, 2019).

In the second mechanism described above for sexual selection (which is followed in this study), it was argued that the evolutionary advantages are particularly strong for one species in a symbiotic relationship if the other species can only undergoes asexual reproduction. For example, since most parasites are asexual, in a host-parasite interaction, if the host can freely select mates that seem immune to the parasite, then the host species would have an evolutionary advantage since its offspring

would be more resistant or immune to the parasite. Of course, this does not imply that the parasite could not co-evolve with hosts because it may still accumulate advantageous genes through other means such as simple DNA mutations (Scoville, 2019). We humans are typical sexual reproduction animals, although modern marriage systems may have exerted social limits to the degree of sexual selection. The Red Queen hypothesis has postulated that sexual selection in humans has played a critical role in shaking off some potentially dangerous microbial pathogens, although one may counter-argue that sexual activities *per se* provides venues for sexually transmitted pathogens. While threats of sexually transmitted pathogens are real, a consensus has been that the reproductive system microbiomes (mostly vaginal and semen microbiomes) are generally predominantly beneficial to human hosts such as suppressing/preventing invasions of opportunistic pathogens and maintaining healthy reproductive tract environment (e.g., right acidity in the human vaginal) (e.g., Ma et al., 2012). Nevertheless, comprehensive examinations of the roles of reproductive system microbiomes in human sexual reproduction from an evolutionary perspective are still missing to the best of knowledge.

In a recent study, Ma and Taylor (2020) suggested that co-evolutionary theories such as the Red Queen hypothesis (Van Valen, 1973; Žliobaitė et al., 2017) should be applicable for the co-evolution between human reproductive systems and their symbiotic microbiomes (mainly semen and vaginal microbiomes) due to the microbiome exchanges between both sexes. They argued that the long-term co-evolution should promote the dynamic homogeneity or stability of the microbiomes, possibly being beneficial for sexual reproduction (sexual selection) such as sperm movement and survival as well as egg fertilization. They further tested the hypothesis by analyzing the heterogeneity of the reproductive system (semen and vaginal microbiomes) based on Taylor's power law (TPL) (Taylor, 1961, 2019) and found no statistically significant differences between the semen and vaginal microbiomes, while both exhibiting significant differences with human gut microbiomes. That is, they demonstrated homogeneity between semen and vaginal microbiomes and therefore indirectly supported the extension of the Red Queen hypothesis to the human reproductive system microbiomes. Nevertheless, the microbiome datasets they used were from independent cohorts, which means that no apparent microbiome exchanges between the men and women in the cohorts actually occurred on ecological time scale. In other words, their results and inferences were on the evolutionary time scale, rather than on the ecological time scale (daily basis). Furthermore, their study only verified the homogeneity but without offering a mechanistic interpretation for the process maintaining the homogeneity at ecological time scale. In the present study, we aim to provide additional evidence to support the Red Queen hypothesis extension to the field of reproductive system microbiomes (Ma and Taylor, 2020) by leveraging the findings from pursuing the previously stated first objective. Specifically, we explore how the mechanism of microbiome transmission during the sexual intercourse influences the heterogeneity (the other side of homogeneity "coin") of the reproductive system microbiomes on ecological time scale. We conjecture that microbiome transmission during sexual intercourse should promote the homogeneity between semen and vaginal microbiomes on the ecological time scale, similar to what occurs on the evolutionary time scale as suggested by Ma and Taylor (2020). If the conjecture is confirmed, then one may infer that the microbiome transmissions between men and women either through sexual intercourse on ecological time scale or through other means on evolutionary time scale all support the Red Queen hypothesis, namely, that the co-evolution between reproductive system microbiomes and hosts facilitates the sexual reproduction (sexual selection). **Figure 1** below diagrammed the hypotheses (objectives) and supporting approaches of the present study. It should be noted that the secondary objective we pursue regarding the Red Queen hypothesis is of conjectural nature since our evidence is indirect and non-experimental. Future studies are required to cross-verify our conjecture.

MATERIALS AND METHODS

Datasets of Microbiome Transmission *via* Sexual Intercourse

Mändar et al. (2015)'s datasets in the form of OTU (operational taxonomic unit) tables, which are reanalyzed in this study, include 23 couples who sought consultation from a physician due to infertility of diverse etiologies. Semen samples were collected by masturbation, and each male was sampled only once. Each female participant was sampled twice, and the vaginal samples were collected in the evening before intercourse and next morning after intercourse. Seminal and vaginal samples were sequenced with Illumina HiSeq2000, and the obtained reads were processed with Mothur software pipeline. A total of 176,358 sequences were obtained, with an average 2,854 reads for each of the 46 vaginal fluid samples, and an average of 1,712 reads for each of the 23 semen samples. Those samples (3 samples per couple, and a total of 69 samples) were ideal materials for investigating microbiome transmission *via* sex, and we take advantage of the neutral theory of biodiversity for determining and estimating the transmission mode and level of the transmission during intercourse. For further information on the cohort information, readers are referred to Mändar et al. (2015). In this study, we used the OTU tables generously supplied to us by the original authors of Mändar et al. (2015).

As a side note, since no second semen samples were taken from the cohort, any discussion on microbiome transmission is primarily one way, from male to female, in this study. Nevertheless, for stable partners, the one-time semen samples cannot exclude the effect of female-to-male transmission obviously.

Multi-Site Neutral Model Approximated by Hierarchical Dirichlet Process

Hubbell's (2001) UNTB (unified neutral theory of biodiversity and biogeography) assumes that all individuals from different species are "neutral" in the sense that their differences, even if exist, would not translate into differences in their

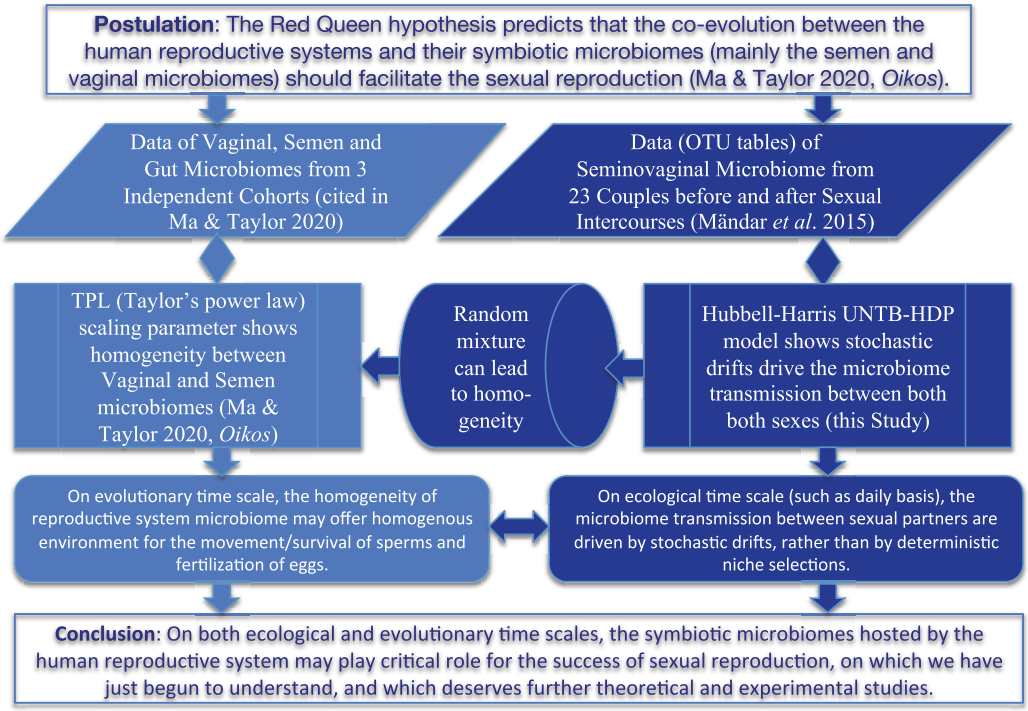


FIGURE 1 | A diagram showing the hypothesis and objectives of this study, including the comparative analysis with previous study (the left side) (Ma and Taylor, 2020).

probabilities of being, and persisting, in the present and future community (Alonso et al., 2006). The neutral theory diametrically contradicts the assumption of classic niche theory, which assumes that different species occupying different niches in their habitats are selected by natural selection to possess different characteristics.

Hubbell's (2001) UNTB conceptually distinguishes the local community dynamics from meta-community dynamics, but both are driven by similar neutral processes. Meta-community dynamics is controlled by two quantities: speciation probability and reproduction rate of an individual. The diversity of the local community is then maintained by immigration from the meta-community, but no speciation is assumed to occur unlike in the meta-community. With all these assumptions, Hubbell's neutral theory was formulated as a master equation (a stochastic differential equation), the solution of which is a probability distribution (sampling formula), which can be compared against the species abundance distribution obtained by sampling ecological communities, *via* rigorous statistical testing such as goodness-of-fitting test with χ^2 statistic.

A fully generalized case of fitting multiple sites UNTB with variable immigration rates among sites is computationally extremely challenging (actually intractable) even for small number of sites, and approximate algorithms must be utilized (Harris et al., 2017). Harris et al. (2017) developed an efficient Bayesian fitting framework by approximating the neutral models with the hierarchical Dirichlet process (HDP). Harris et al. (2017) approximation captures the essential elements of the UNTB, i.e., neutrality, finite populations, and multiple panmictic geographically isolated populations linked by relatively rare migration. With Harris et al.'s (2017) HDP-MSN model, i.e., multi-site neutral (MSN) model approximated by the HDP process, it is possible to simultaneously estimate the variable immigrations rates among a large number of sites within feasible computation timeframe, and therefore makes the UNTB truly multi-site. For this reason, we term Harris et al. (2017) implementation of Hubbell's UNTB as HDP-MSN model (hierarchical Dirichlet process approximation of multisite neutral model). Furthermore, the HDP-MSN model distinguishes between neutral local community (given a non-neutral meta-community) and the full UNTB (where the meta-community also assembles neutrally), and the neutrality tests can be performed at both meta-community level and local community level.

The Unified Neutral Theory of Biodiversity Model

As stated previously, a primary assumption in Hubbell's UNTB is that both local community dynamics and regional metacommunity dynamics are driven by similar neutral processes, although they are separated conceptually (Hubbell, 2001, 2006). Regarding the local community dynamics, assume there are M local communities indexed as $i = 1, 2... M$, each with N_i individuals and N_i is constant for each local community. At each time step, the local community dynamics for site i is driven by a random process—selecting an individual randomly and either replacing it by a randomly chosen individual immigrated

from the metacommunity with migration probability (m_i) or replacing it by an indigenous member randomly chosen from the local community (i) with probability (1-m_i). The UNTB further assumes that the local communities are at stationary state, and each site is assigned a vector $\bar{\pi}_i = (\pi_{i,1}, ..., \pi_{i,S})$, denoting the probability for observing a particular species at site i, which is simply the species abundance distribution (SAD) of site or local community i.

One parameter, *immigration rate* (I_i), controls the coupling of a local community to the meta-community by replacing the two parameters (m_i and N_i), i.e.,

$$I_i = (N_i - 1)[m_i/(1 - m_i)]. \quad (1)$$

Regarding the equivalent neutral dynamics of metacommunity, new species are generated through speciation with a probability ν. Similar to local community neutral dynamics, the speciation rate, also known as fundamental biodiversity number (θ), can be defined as:

$$\theta = (\nu/(1 - \nu))(N - 1), \quad (2)$$

where N is the fixed (total) number of individuals in the metacommunity. The parameter θ can be considered as the rate at which new individuals are added to the metacommunity as a result of speciation.

The third aspect of the UNTB is to treat the observed samples, i.e., the rows in the data matrix $\mathbf{X}_{MxS}$ with elements x_{ij} giving the abundance of species j is observed at site i, as a sample from the local community. As a side note, the matrix X is actually the OTU table of 16s-rRNA gene abundances in the case of test datasets we used in this study. Assume that the sample is taken with replacement, let $J_i = \sum_{j=1}^{S} x_{ij}$, and then the multinomial (MN) distribution describes the vector of observations at a given site i, i.e.,

$$\bar{X}_i \sim MN(J_i, \bar{\pi}_i). \quad (3)$$

In summary, the above three elements (the immigration rate, speciation rate, and multinomial distribution) constitute the building blocks of the neutral theory. These building blocks, together with the neutrality assumption—that all individuals from different species are "neutral" in the sense that their differences, even if exist, would not translate into differences in their probabilities of being, and persisting, in the present and future community (Alonso et al., 2006), may be implemented slightly differently in the following multi-site neutral (MSN) model by Harris et al. (2017). However, the fundamental ideas and elements of neutral theory are the same with classic neutral theory.

Hierarchical Dirichlet Process-Multi-Site Neutral Model

Neutral theory is one of the four paradigms of metacommunity theory. Since metacommunity consists of multiple local communities, it is essentially a multi-site model. It turned out that a fully general case of fitting multiple sites UNTB with different immigration rates is computationally extremely challenging (actually intractable) even for small number of sites, and approximate algorithms must be utilized (Harris et al., 2017). Harris et al. (2017) developed an efficient Bayesian fitting framework by approximating the neutral models with the hierarchical Dirichlet process (HDP). Harris et al. (2017) approximation captures the essential elements of the UNTB, i.e., neutrality, finite populations, and multiple panmictic geographically isolated populations linked by relatively rare migration—while little influenced by the specific details of the local community dynamics.

Sloan et al. (2006, 2007) showed that for large local population sizes, and assuming a fixed finite-dimensional metacommunity distribution with S species present, then the local community distribution, $\bar{\pi}_i$, can be approximated by a Dirichlet distribution (Sloan et al., 2006, 2007). But it was Harris et al. (2017) developed the general framework for approximating the UNTB computationally efficiently. Assuming there is a potentially infinite number of species that can be observed in the local community, then the stationary distribution of observing local population i is a Dirichlet process (DP), i.e.,

$$\bar{\pi}_i | I_i, \bar{\beta} \sim DP(I_i, \bar{\beta}) \quad (4)$$

where $\bar{\beta} = (\beta_1, ..., \beta_S)$ is the *relative frequency* of each species in the metacommunity.

At the metacommunity level, a Dirichlet process is still applicable, but then the base distribution is simply a uniform distribution over arbitrary species labels, and the concentration parameter is the biodiversity parameter (θ) (Harris et al., 2017). The metacommunity distribution follows the stick breaking process, i.e.,

$$\bar{\beta} \sim Stick(\theta). \quad (5)$$

Given that both local community and metacommunity are Dirichlet processes, it becomes a hierarchical Dirichlet process (HDP) in terms of the machine learning (Teh et al., 2006; Harris et al., 2017).

Alternatively, Dirichlet process can also be viewed as the so-termed Chinese restaurant process, from which the Antoniak equation can be derived. Antoniak equation represents the number of types (or species) (S) observed following N draws from a DP with concentration parameter θ and is in the following form:

$$P(S|\theta, N) = s(N, S)\theta^S \frac{\Gamma(\theta)}{\Gamma(\theta + N)} \quad (6)$$

where $s(N, S)$ is the unsigned Stirling number of the first kind and $\Gamma(.)$ denotes the gamma function (Antoniak, 1974).

Gibbs Sampler (MCMC Algorithm) for the Hierarchical Dirichlet Process-Multi-Site Neutral Model

The full HDP approximated neutral model (HDP-neutral) is formed by combining previous equations (4–6). Harris et al. (2017) devised an efficient Gibbs sampler for the HDP neutral approximation, which is a type of Bayesian Markov Chain Monte Carlo (MCMC) algorithm and can be summarized as the following four sampling steps:

(*a*) Sample the biodiversity parameter (θ) from the conditional probability:

$$P(\theta|S, T) \propto s(T, S)\theta^S \frac{\Gamma(\theta)}{\Gamma(\theta + T)} Gamma(\theta|\alpha, \zeta) \quad (7)$$

where θ is the biodiversity parameter. $T = \sum_{i=1}^{M} \sum_{j=1}^{S} T_{ij}$ is the number of ancestors, S is the number of species in metacommunity, $s(T, S)$ is the unsigned Stirling number of the first kind (Antoniak, 1974), and α and ζ are constants.

(*b*) Sample the metacommunity distribution:

$$\bar{\beta} = (\beta_1, \beta_2..., \beta_S, \beta_u) \sim DP(T_{\cdot 1}, T_{\cdot 2}, ..., T_{\cdot S}, \theta) \quad (8)$$

where $T_{\cdot j} = \sum_{i=1}^{M} T_{ij}$ is the number of ancestors of species j in metacommunity.

(*c*) Sample the immigration rates:

$$P(I_i|T_{ij}) \propto \frac{\Gamma(I_i)}{\Gamma(J_i + I_i)} I_i^{T_i} Gamma(I_i|\eta, \nu) \quad (9)$$

where both η and ν are constants.

TABLE 1 | Test results of fitting the HDP-MSN (hierarchical Dirichlet process, multi-site neutral) model of Harris et al. (2017) to the meta-communities consisting of 3-site semen-vaginal samples (CM = Semen Sample, CNA = vaginal sample before intercourse, and CNB = vaginal sample after intercourse) (*P*-value > 0.05 indicating significant or satisfactory fitting to the MSN)*,**.

ID	L_O	θ	*m*	*M-value*	Meta-community				Local Community			
					L_M	N_M	*N*	P_M	L_L	N_L	*N*	P_L
1	−382.709	40.978	0.151	73.153	−396.822	1,566	2,500	0.626	−389.420	1,417	2,500	0.567
2***	−1, 385.486	141.111	0.015	37.188	−1, 370.380	1,073	2,500	0.429	−1, 415.622	1,632	2,500	0.653
3	−708.010	91.888	0.013	35.028	−727.026	1,551	2,500	0.620	−744.482	1,775	2,500	0.710
6	−1, 032.545	100.840	0.020	39.051	−1, 051.866	1,499	2,500	0.600	−1, 070.271	1,803	2,500	0.721
7	−1, 040.569	97.987	0.010	32.693	−1, 069.604	1,594	2,500	0.638	−1, 091.907	1,930	2,500	0.772
8	−753.034	73.124	0.029	59.935	−837.017	2,162	2,500	0.865	−804.300	1,990	2,500	0.796
9	−820.546	106.933	0.012	24.899	−835.507	1,468	2,500	0.587	−855.005	1,776	2,500	0.710
12	−425.321	47.271	0.022	43.152	−460.434	1,819	2,500	0.728	−464.377	1,971	2,500	0.788
14	−717.459	93.516	0.011	33.766	−725.231	1,374	2,500	0.550	−742.193	1,638	2,500	0.655
15	−425.352	40.732	0.040	104.209	−476.039	2,024	2,500	0.810	−448.788	1,722	2,500	0.689
16	−692.768	95.069	0.010	27.169	−730.169	1,788	2,500	0.715	−732.498	1,800	2,500	0.720
17	−874.131	102.123	0.007	18.343	−917.696	1,813	2,500	0.725	−949.243	2,146	2,500	0.858
18	−768.402	97.810	0.011	26.005	−808.089	1,771	2,500	0.708	−819.856	1,986	2,500	0.794
21	−1, 381.209	129.878	0.021	40.338	−1, 392.766	1,381	2,500	0.552	−1, 431.878	1,936	2,500	0.774
22	−690.534	70.839	0.015	40.315	−739.996	1,910	2,500	0.764	−741.027	2,012	2,500	0.805
23	−1, 044.392	49.574	0.041	105.839	−1, 277.726	2,478	2,500	0.991	−1, 099.484	2,078	2,500	0.831
24	−740.746	47.534	0.072	151.599	−880.842	2,428	2,500	0.971	−762.288	1,649	2,500	0.660
25	−1, 262.028	79.649	0.024	76.674	−1, 386.549	2,225	2,500	0.890	−1, 301.471	1,725	2,500	0.690
26	−1, 178.871	96.351	0.016	38.733	−1, 290.164	2,253	2,500	0.901	−1, 256.236	2,161	2,500	0.864
27	−998.888	55.951	0.056	160.489	−1, 127.655	2,303	2,500	0.921	−1, 000.514	1,277	2,500	0.511
28	−1, 059.024	62.918	0.076	207.492	−1, 185.529	2,296	2„500	0.918	−1, 035.615	820	2,500	0.328
29	−1, 086.036	60.456	0.040	126.032	−1, 339.392	2,490	2,500	0.996	−1, 118.532	1,749	2,500	0.700
30	−981.639	66.499	0.013	35.685	−1, 081.354	2,207	2,500	0.883	−1, 055.134	2,176	2,500	0.870
Mean	−889.117	80.393	0.032	66.860	−961.211	1,890	2,500	0.756	−927.397	1,790	2,500	0.716
Passing rate (%)								100%				100%

**N = 2,500 is the number of Gibb samples selected from 25,000 simulated communities (i.e., every tenth iteration of the last 25,000 Gibbs samples), it is chosen to compute the pseudo P-value below for conducting the neutrality test.*

L_0 is the actual (observed) log-likelihood.

θ is the median of biodiversity numbers computed from 25,000 times of simulations.

m is the migration probability.

M-value is the average medians of the migration rates of local communities in each meta-community (i.e., the average median of the individuals migrated per generation), also computed from 25,000 times of simulations.

L_M is the median of the log-likelihoods of the simulated neutral meta-community samples; and N_M is the number of simulated neutral meta-community samples with their likelihoods satisfying $L_M \leq L_0$ (where L_M and L_0 are the simulated and actual likelihood respectively).

$P_M = N_M / N$ is the pseudo p-value for testing the neutrality at meta-community level; if $P_M > 0.05$, the meta-community is indistinguishable from the prediction of neutral model.

L_L is the median of the log-likelihoods of the simulated local community samples, and N_L is the number of simulated local community samples with their likelihoods satisfying $L_L \leq L_0$ (where L_L and L_0 are the simulated and actual likelihood respectively).

$P_L = N_L / N$, is the pseudo p-value for testing the neutrality at the local community level; if $P_L > 0.05$, the local community satisfies the neutral model.

Due to the typo/error in Harris et al. (2017), the P_M-values exhibited here are adjusted as ($P_M = 1 - P_{MS}$), where P_{MS} is output from their computational program. Similarly, the P_L-values are adjusted as ($P_L = 1 - P_{LS}$), where P_{LS} is output from their computational program.***Figure 2*** *displayed the fitting of the MSN to #2 sample by plotting the predicted and observed species abundance rank distribution.*

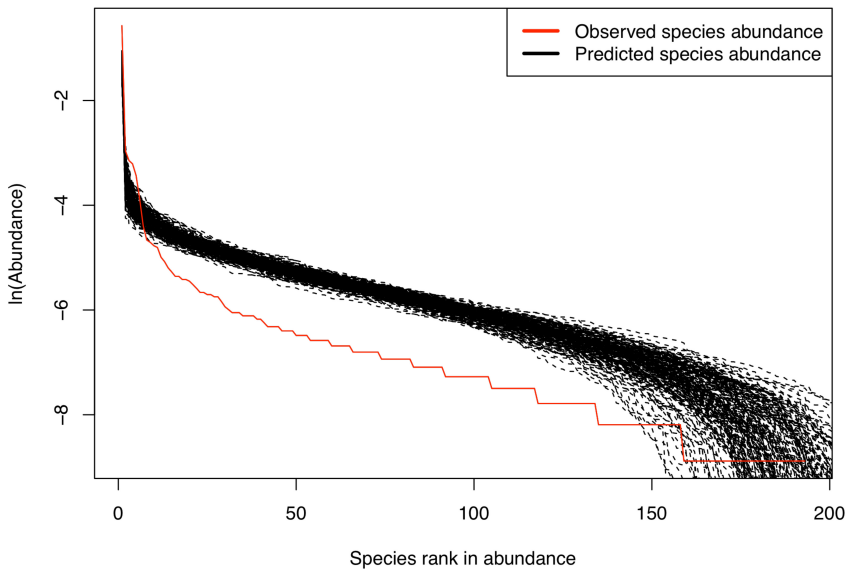


FIGURE 2 | Fitting Harris et al. (2017) MSN (multi-site neutral) model with the meta-community of (CM+CNA+CNB) samples from a randomly selected sample group (Couple#2).

(*d*) Sample the ancestral states:

$$P(T_{ij}|x_{ij}, I_i, \beta_i) = \frac{\Gamma(I_i\beta_j)}{\Gamma(x_{ij} + I_i\beta_j)} s(x_{ij}, T_{ij})(I_i\beta_j)^{T_{ij}} \quad (10)$$

where the various symbols have the same representations as previously defined.

Harris et al. (2017) discovered through experiments that to ensure sampling was from the stationary distribution, 50,000 Gibb samples for each fitted dataset were required with the first 25,000 iterations removed as burn-in. The results are reported as the *median* values over the last 25,000 samples with upper and lower credible limits (Bayesian confidence) given by 2.5% and 97.5% quantiles of those samples.

Goodness-of-Fitting Test for the Hierarchical Dirichlet Process-Multi-Site Neutral Model

To determine whether an observed dataset fits to the HDP-neutral model, Harris et al. (2017) proposed a similar Monte Carlo significance test to that used by Etienne (2007). For both the local and metacommunity level tests, samples were generated from 2,500 sets of fitted parameters, which were in turn sampled from every tenth iteration of the last 25,000 Gibbs samples (the first 25,000 were removed as burn-in as mentioned previously). The calculation and usage of the pseudo-*P* values for testing the goodness-of-fitting of the HDP-neutral model are explained in the footage for **Table 1** in the section of results, where actual model fittings to the datasets of seminovaginal microbiomes are presented. For the detailed computational procedures and computational program, readers are referred to Harris et al. (2017), which we used to perform the microbiome data analysis in this study. In addition, demonstration on the application of HDP-MSN model to the human microbiomes can be found in Ma (2020, 2021a,b).

RESULTS AND CONCLUSION

Table 1 listed the test results of fitting the MSN (multi-site neutral) model of Harris et al. (2017) to the semen-vaginal meta-community, which consists of the three samples from each couple (i.e., CM = semen sample, CNA = vaginal sample before intercourse, and CNB = vaginal sample after intercourse). A total of 23 meta-communities, one for each couple, were tested for their fitness to the MSN model, respectively. The neutrality-passing rate in the 23 couples is 100% (all 23) both at the local community and metacommunity level (**Figure 2**). **Table 2** listed the test results of fitting the MSN model, with pair of samples from a couple grouped as a meta-community. There are three possible pair-wise combinations, CM and CNA, CM and CNB, and CNA and CNB. The meta-communities from all pairs passed the neutrality test at both local and meta-community level (100% passing rates). These findings suggest that the transmission of microbes during sexual intercourse seems to be similar to a random "walk" (or random dispersal) and is driven by stochastic drifts. The 100% passing rate indicates that deterministic selection (forces) seems to play little role. **Table 2** also suggested that the transmission probability of microbiomes through sexual intercourse appears to be 0.05 approximately.

Hence, our analysis revealed that the microbiome transmission during the intercourse is primarily driven by stochastic neutral drift alone and should just be a random walk. The virtually universal neutrality among all the samples suggest that the neutrality is maintained within couples on daily basis, rather than only during the sexual intercourse. It should also be plausible to conjecture that the neutrality may possess both ecological and evolutionary implications, which we further elaborate in the discussion section.

TABLE 2 | Test results of fitting the HDP-MSN (hierarchical Dirichlet process, multi-site neutral) model of Harris et al. (2017) to the 2-sites meta-community (pair-wise combination of CM, CNA, CNB) ($p > 0.05$ indicating significant or satisfactory fitting to the MSN)*.

ID	L_O	θ	m	*M-value*	Meta-community				Local Community			
					L_M	N_M	N	P_M	L_L	N_L	N	P_L
Meta-community = CM (semen) and CNA (vaginal before): i.e., Semen-Vaginal before Sex												
1	−294.942	38.974	0.269	127.603	−281.606	891	2,500	0.356	−281.056	827	2,500	0.331
2	−1,249.588	167.832	0.019	48.708	−1,178.198	477	2,500	0.191	−1,234.428	1,016	2,500	0.406
3	−606.843	104.394	0.018	46.314	−586.250	930	2,500	0.372	−607.972	1,267	2,500	0.507
6	−941.873	117.855	0.025	48.872	−912.698	859	2,500	0.344	−950.423	1,395	2,500	0.558
7	−641.123	113.683	0.013	34.910	−613.257	792	2,500	0.317	−636.724	1,162	2,500	0.465
8	−519.786	56.211	0.145	185.879	−545.308	1,648	2,500	0.659	−521.949	1,292	2,500	0.517
9	−557.682	126.973	0.030	29.291	−523.500	701	2,500	0.280	−545.089	1,039	2,500	0.416
12	−318.186	47.318	0.242	67.048	−316.137	1,196	2,500	0.478	−323.589	1,367	2,500	0.547
14	−637.346	139.845	0.009	33.859	−591.837	599	2,500	0.240	−619.023	975	2,500	0.390
15	−445.940	80.545	0.044	35.486	−426.659	926	2,500	0.370	−432.225	1,072	2,500	0.429
16	−602.985	151.693	0.011	26.547	−564.504	752	2,500	0.301	−590.195	1,061	2,500	0.424
17	−722.901	154.230	0.007	20.553	−716.408	1,150	2,500	0.460	−750.418	1,676	2,500	0.670
18	−649.086	103.733	0.015	35.366	−639.784	1,090	2,500	0.436	−658.020	1,420	2,500	0.568
21	−1,041.511	162.229	0.019	41.977	−1,007.968	791	2,500	0.316	−1,050.771	1,406	2,500	0.562
22	−565.380	85.202	0.016	46.302	−571.889	1,353	2,500	0.541	−584.253	1,574	2,500	0.630
23	−687.356	49.269	0.066	139.642	−778.801	2,240	2,500	0.896	−682.369	1,149	2,500	0.460
24	−463.632	39.916	0.229	384.019	−514.220	2,094	2,500	0.838	−455.452	998	2,500	0.399
25	−1,050.043	89.466	0.023	92.515	−1,095.046	1,699	2,500	0.680	−1,053.375	1,268	2,500	0.507
26	−977.067	152.656	0.011	31.491	−973.298	1,189	2,500	0.476	−1,002.310	1,629	2,500	0.652
27	−886.104	75.389	0.029	99.172	−932.443	1,647	2,500	0.659	−886.128	1,251	2,500	0.500
28	−838.843	60.589	0.103	335.688	−875.010	1,775	2,500	0.710	−790.556	342	2,500	0.137
29	−1,029.231	114.178	0.014	51.628	−1,097.116	1,962	2,500	0.785	−1,068.638	1,509	2,500	0.604
30	−837.386	105.981	0.009	28.098	−857.412	1,513	2,500	0.605	−878.143	1,861	2,500	0.744
Mean	−720.210	101.659	0.059	86.564	−721.711	1,229	2,500	0.492	−721.874	1,242	2,500	0.497
Passing rate (%)								100%				100%
Meta-community = CM (semen) and CNB (vaginal after): i.e., Semen-Vaginal after Sex												
1	−289.030	43.371	0.204	105.245	−280.867	1,056	2,500	0.422	−282.879	1,056	2,500	0.426
2	−727.836	129.926	0.024	39.766	−688.822	677	2,500	0.271	−718.518	677	2,500	0.442
3	−511.738	66.266	0.054	99.539	−493.772	958	2,500	0.383	−498.098	958	2,500	0.418
6	−524.047	62.040	0.104	109.920	−516.595	1,127	2,500	0.451	−513.470	1,127	2,500	0.437
7	−922.580	120.417	0.016	41.932	−900.271	950	2,500	0.380	−936.888	950	2,500	0.585
8	−664.375	80.807	0.036	74.639	−683.925	1,547	2,500	0.619	−680.023	1,547	2,500	0.586
9	−734.546	127.632	0.013	31.013	−707.562	874	2,500	0.350	−740.545	874	2,500	0.537
12	−327.890	66.144	0.022	42.932	−314.591	994	2,500	0.398	−327.296	994	2,500	0.494
14	−619.570	92.639	0.036	51.576	−591.182	810	2,500	0.324	−610.084	810	2,500	0.445
15	−347.050	36.931	0.073	194.679	−366.486	1,617	2,500	0.647	−350.091	1,617	2,500	0.533
16	−657.279	148.840	0.012	26.854	−629.818	838	2,500	0.335	−658.247	838	2,500	0.506
17	−668.278	122.494	0.013	23.030	−666.608	1,219	2,500	0.488	−696.241	1,219	2,500	0.694
18	−660.140	139.453	0.015	27.505	−626.643	757	2,500	0.303	−659.396	757	2,500	0.495
21	−845.976	140.958	0.041	41.885	−799.162	567	2,500	0.227	−836.271	567	2,500	0.428
22	−600.325	102.615	0.022	37.687	−575.632	844	2,500	0.338	−595.463	844	2,500	0.465
23	−739.692	52.565	0.051	126.679	−836.581	2,237	2,500	0.895	−753.233	2,237	2,500	0.594
24	−673.891	54.907	0.055	140.013	−756.745	2,174	2,500	0.870	−684.364	2,174	2,500	0.554
25	−1,040.640	75.592	0.037	141.517	−1,102.943	1,898	2,500	0.759	−1,028.525	1,898	2,500	0.442
26	−1,041.126	127.860	0.013	41.347	−1,077.386	1,715	2,500	0.686	−1,091.201	1,715	2,500	0.759
27	−794.096	49.538	0.115	397.648	−863.984	2,040	2,500	0.816	−767.550	2,040	2,500	0.293
28	−957.051	62.429	0.084	326.120	−1,042.312	2,170	2,500	0.868	−920.118	2,170	2,500	0.235
29	−860.525	55.146	0.059	231.842	−986.236	2,364	2,500	0.946	−843.578	2,364	2,500	0.364
30	−768.696	64.007	0.020	54.922	−811.856	1,772	2,500	0.709	−791.770	1,772	2,500	0.611
Mean	−694.625	87.938	0.049	104.708	−709.564	1,357	2,500	0.543	−694.950	1,357	2,500	0.493
Passing rate (%)								100%				100%

(Continued)

TABLE 2 | (Continued)

ID	L_O	θ	*m*	*M-value*	Meta-community				Local Community			
					L_M	N_M	*N*	P_M	L_L	N_L	*N*	P_L
Meta-community = CNA + CNB: i.e., two vaginal samples												
1	−117.932	6.528	0.151	85.064	−146.579	2,029	2,500	0.812	−115.060	1,077	2,500	0.431
2	−687.812	52.987	0.026	80.201	−806.923	2,385	2,500	0.954	−704.006	1,556	2,500	0.622
3	−237.691	36.215	0.001	4.373	−285.973	2,096	2,500	0.838	−288.602	2,197	2,500	0.879
6	−558.172	98.290	0.006	15.714	−571.206	1,444	2,500	0.578	−591.573	1,820	2,500	0.728
7	−495.413	39.212	0.009	38.126	−615.493	2,410	2,500	0.964	−543.052	1,871	2,500	0.748
8	−274.276	51.506	0.002	4.737	−320.860	2,072	2,500	0.829	−327.750	2,184	2,500	0.874
9	−375.537	72.550	0.003	6.641	−388.970	1,514	2,500	0.606	−402.978	1,779	2,500	0.712
12	−211.853	24.140	0.001	3.883	−255.803	2,085	2,500	0.834	−253.221	2,145	2,500	0.858
14	−211.848	46.287	0.001	3.275	−243.275	1,846	2,500	0.738	−247.457	2,038	2,500	0.815
15	−120.950	23.085	0.000	1.592	−144.765	1,868	2,500	0.747	−145.086	1,899	2,500	0.760
16	−221.478	49.669	0.001	3.189	−253.301	1,905	2,500	0.762	−260.003	2,068	2,500	0.827
17	−377.262	34.993	0.003	11.786	−443.105	2,172	2,500	0.869	−418.600	1,968	2,500	0.787
18	−255.349	42.477	0.002	4.744	−300.117	2,090	2,500	0.836	−306.651	2,219	2,500	0.888
21	−940.462	130.900	0.012	29.270	−947.139	1,351	2,500	0.540	−978.608	1,844	2,500	0.738
22	−267.858	38.835	0.001	4.851	−306.574	1,996	2,500	0.798	−316.669	2,160	2,500	0.864
23	−612.796	32.891	0.041	133.344	−782.288	2,465	2,500	0.986	−636.483	1,830	2,500	0.732
24	−352.422	58.629	0.004	8.755	−405.049	2,085	2,500	0.834	−380.030	1,539	2,500	0.616
25	−342.730	19.430	0.049	100.793	−477.573	2,462	2,500	0.985	−356.722	1,688	2,500	0.675
26	−363.988	22.576	0.073	118.692	−488.442	2,457	2,500	0.983	−365.290	1,295	2,500	0.518
27	−270.576	14.159	0.050	94.893	−377.782	2,429	2,500	0.972	−274.642	1,391	2,500	0.556
28	−205.859	12.578	0.126	152.710	−290.304	2,433	2,500	0.973	−202.680	1,105	2,500	0.442
29	−298.507	15.997	0.071	137.058	−420.280	2,451	2,500	0.980	−306.712	1,524	2,500	0.610
30	−322.943	16.873	0.059	123.100	−434.570	2,425	2,500	0.970	−329.146	1,443	2,500	0.577
Mean	−353.205	40.905	0.030	50.730	−422.016	2,107	2,500	0.843	−380.479	1,767	2,500	0.707
Passing rate (%)								100%				100%

The interpretations of the symbols are the exactly the same as in **Table 1.*

Figure 3 shows the box chart for the fundamental biodiversity (θ) numbers estimated with the MSN models for the four different meta-community settings, as listed in **Tables 1**, **2**. It confirms the previous conclusions we draw from the neutrality tests with the MSN modeling reported in **Tables 1**, **2**. Specifically, θ is the lowest in the meta-community of the two vaginal samples setting (CNA and CNB), which simply indicates that the number of "novel" species (regional diversity) in the meta-community of two-sample vaginal microbiome is the lowest, compared with the other three meta-community settings, in which both semen and vaginal communities are included. This should be expected since the "range" of CNA and CNB metacommunity should be smaller than that of the others, and therefore hosts smaller microbiome diversity.

Figure 4 shows the box chart for the immigration probability (*m*) estimated with the MSN models for the four different meta-community settings, as listed in **Tables 1**, **2**. It confirms the previous conclusion we draw from the MSN neutrality tests reported in **Tables 1**, **2**. Specifically, *m* is the lowest in the meta-community of the two vaginal samples, which simply says that the dispersal (transmission) is the lowest between the two vaginal samples of a woman, compared with the other three meta-communities in which semen microbiome is included. This should be true obviously for the same reason as in the case of previously explained results of θ.

Table 3 listed the *p*-value from the Wilcoxon non-parametric test for the immigration probability (*m*) and the fundamental biodiversity number (θ). **Figure 5** further illustrated the same information as displayed in **Table 3**. In terms of the immigration probability (*m*), the meta-community of two vaginal samples (CNA and CNB) has significant differences (red links) with the meta-communities of CM and CNA or CM and CNB, and has no significant differences with all other meta-communities (green links). This should be expected, and it simply indicates that the transmission (dispersal) probability between man and woman after intercourse is significantly higher than the immigration probability naturally occurring within the vaginal microbiome.

In summary, previous results have shown that the microbiome transmission during the sexual intercourse appears to be driven by stochastic drifts of microbiome demography and dispersal, rather than by certain deterministic processes such as niche selections (e.g., the preferences of microbes to particular habitats). Further comparisons of the complementary seminovaginal microbiome samples before and after intercourse suggest that the level of stochastic drifts in the semen-vaginal metacommunity should be beyond the duration of sexual intercourse and be predominant on daily basis given that the neutrality passing rates were 100% in both before and after sexual intercourse. In other words, the microbiome exchanges between male and female, at least within couples, on ecological time scale

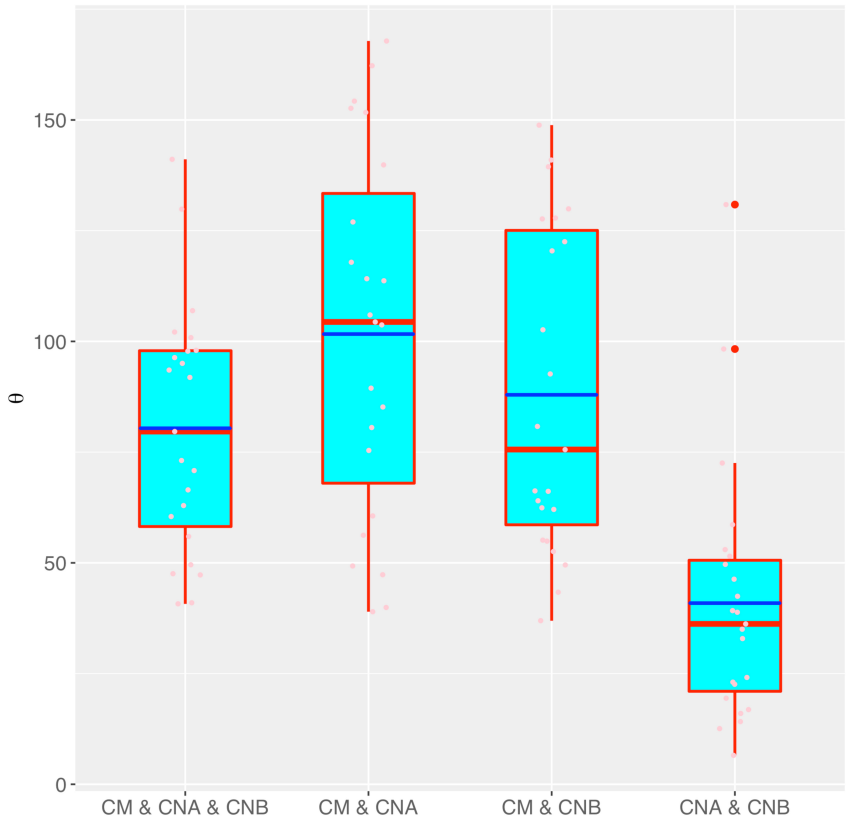


FIGURE 3 | The box chart for the fundamental biodiversity (θ) numbers estimated with the MSN models for the four different meta-community settings (also see **Table 3** for the *p*-values of the significance test for their differences in θ). Box red lines, blue lines, edges, whiskers, and bigger red points signify the median, mean, inter-quartile range (IQR), 1.5 × IQR, and > 1.5 × IQR, respectively. The smaller points in each box are the real values of θ of each sample.

are most likely driven by stochastic drifts and dispersal, rather than by certain deterministic forces.

The previous interpretations of the results are focused on the first or primary objective of this study, i.e., the underlying mechanisms of the microbial community/metacommunity assembly and diversity maintenance, including the possible transmission of microbes during sexual intercourse. As to the secondary objective of this study—the evolutionary implications of the findings of this study—is further discussed in the following section.

DISCUSSION

There are currently two major categories of hypotheses on the relationship between the evolutions of humans and their symbiotic microbiomes. The emerging theory of evolution considers the individual animal or plant as a community (or a holobiont) consisting of the host plus all of its symbiotic microbes. The collective genome of the holobiont is defined as the hologenome. The holobiont/hologenome theory maintains that the variations in the hologenome can be transmitted from generation to generation with reasonable fidelity, and are subject to evolutionary changes resulting from selection and drift (Rosenberg et al., 2009; Rosenberg and Zilber-Rosenberg, 2018). The theory further maintains that many factors including new acquisitions of microbes, horizontal gene transfers, and changes in microbial species abundance within hosts may cause genetic variation in the hologenome. Due to its mixture flavor of both Lamarckian and Darwinian, the theory stresses both cooperation and competition within and between holobionts (Rosenberg et al., 2009; Rosenberg and Zilber-Rosenberg, 2018), but the overall framework is still Darwinian evolution. The second category emphasizes the co-evolution between the hosts and microbiomes. For example, the classic Red Queen hypotheses (Van Valen, 1973; Žliobaitė et al., 2017) for explaining sexual selection and host/parasite co-evolutions have been applied to interpret the host/microbiome co-evolution (e.g., Papkou et al., 2018; Ma and Taylor, 2020). In reproductive biology, microbial symbionts were found to mediate reproductive isolation in *Drosophila*, but debates also exist (Leftwich et al., 2017; Shapiro, 2017; Schneider et al., 2019). Although, to the best of our knowledge, no experimental studies have been conducted with the human microbiomes, their roles in human reproductive biology cannot be excluded. Theoretically, Ma and Taylor (2020) postulated that the human semen and vaginal microbiomes, collectively termed human reproductive system microbiomes, may have coevolved with hosts to facilitate

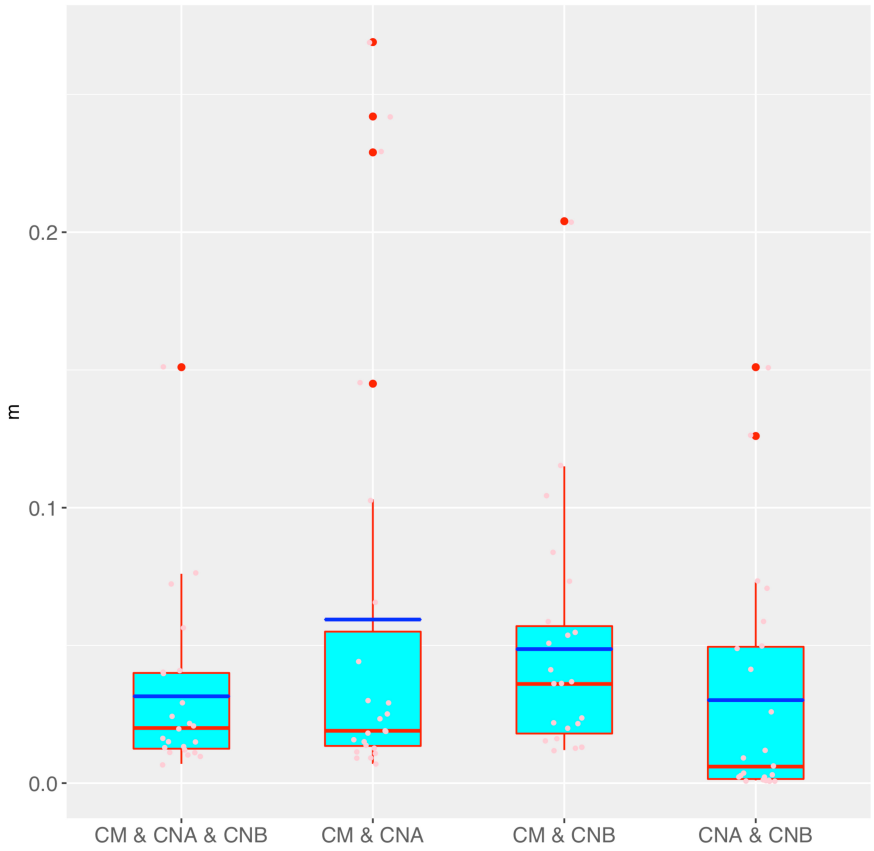


FIGURE 4 | The box chart for the immigration probability (*m*) estimated with the MSN models for the four different meta-community settings (also see **Table 3** for the *p*-values of the significance test for their differences in *m*). Box red lines, blue lines, edges, whiskers, and bigger red points signify the median, mean, inter-quartile range (IQR), 1.5 × IQR, and > 1.5 × IQR, respectively. The smaller points in each box are the real values of *m* of each sample.

the sexual reproduction such as offering beneficial environmental for the sperm movement/survival and/or egg fertilization.

While a hallmark of the Red Queen hypothesis is the antagonism or evolutionary conflicts, in which both species are locked in an "arms race" to maximize their relative fitness (Aleru and Barber, 2020), how would the mutualism or evolutionary cooperation between humans and their microbiomes fits to the picture of Red Queen dynamics? In the case of human gut microbiome, it has been found that our immune system is trained to discriminately treat pathogenic bacteria vs. beneficial ones that constitutes majority of the human gut microbiome. Positive selection—the rapid spread of new beneficial gene mutations in populations over time—has been observed in immune system related genes. Indeed, immune system components are among the most rapidly evolving genes in animal genomes. Commensal microbes are believed to be able to shift the balance of host-pathogen conflicts as described by the Red Queen dynamics (Aleru and Barber, 2020). In reproductive biology, microbial symbionts were found to mediate reproductive isolation in *Drosophila*, but debates also exist (Leftwich et al., 2017; Shapiro, 2017; Schneider et al., 2019). It should also be possible that the human and their microbiota have been coevolving with hosts through cooperation, competition (antagonism), and communication (signaling); consequently, the Red Queen type evolutionary dynamics should exist within and between holobiont(s), which are host plus all of its symbiotic microbes (Rosenberg et al., 2009; Rosenberg and Zilber-Rosenberg, 2018).

Ma and Taylor (2020) proposed that the co-evolution between human reproductive system and their symbiotic microbiomes (mainly the semen and vaginal microbiomes) should facilitate the sexual reproduction, as predicted by the classic Red Queen hypothesis. They further provided a piece of evidence to support this microbiome extension to the Red Queen theory by demonstrating that the heterogeneities of semen and vaginal microbiomes exhibited no significant differences, whereas both exhibiting significant differences with human gut microbiomes. Their logic was that the homogeneity or stability should be helpful for the success of sexual reproduction such as being beneficial for the sperm movement/survival and/or egg fertilization. However, Ma and Taylor (2020) study possessed two limitations, as mentioned in previous introduction section, one is the lack of a mechanistic interpretation for why the homogeneity between semen and vaginal microbiomes was the case, and the second is that the microbiome datasets they used were from independent cohorts of men and women (no apparent

TABLE 3 | The *p*-value from the Wilcoxon non-parametric test for the immigration probability (*m*) and the fundamental biodiversity number (θ) in **Tables 1**, **2**.

Meta-community I	Meta-community II	*M*	θ
CM and CNA and CNB	CM and CNA	0.575	0.068
CM and CNA and CNB	CM and CNB	0.054	0.617
CM and CNA and CNB	CNA and CNB	0.076	<0.001*
CM and CNA	CM and CNB	0.296	0.326
CM and CNA	CNA and CNB	0.026*	<0.001*
CM and CNB	CNA and CNB	0.012*	<0.001*

**Indicating the treatments with significant difference in the immigration probability at the significance level of P-value = 0.05.*

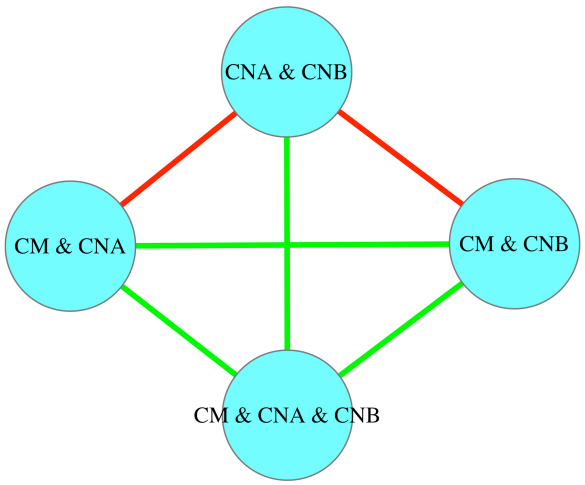


FIGURE 5 | The significance test for the immigration probability (*m*) between different meta-communities: In terms of the immigration probability (*m*), the meta-community of two vaginal samples ("CNA and CNB") has significant differences (red links) with the meta-communities of "CM and CNA" or "CM and CNB," and has no significant differences with all other meta-communities (green links).

microbiome exchanges on ecological time-scale such as daily basis), rather than from intimately connected couples as the datasets (Mändar et al., 2015) reanalyzed in this study.

The results from the present study actually fill the two gaps left by Ma and Taylor (2020) study. First, the stochastic drifts or random nature of microbiome exchanges explains the microbiome homogeneity within the reproductive system (i.e., semen and vaginal microbiomes). This is because random migration (mixture) is arguably the most effective mechanism (process) to achieve homogeneity in a fluid environment. Second, the time scale of the reproductive system microbiomes we used in this study is on ecological time scale (daily basis) given that the complementary seminovaginal microbiome samples were obtained both before and after sexual intercourses. Therefore, this study not only offers another piece of evidence to support the prediction of the Red Queen hypothesis on ecological time scale, but also presents a mechanistic interpretation for the process generating the microbiome homogeneity within the reproductive system as revealed by Ma and Taylor's (2020) previous study, which was postulated on the evolutionary time scale as explained previously. Combined together, both previous study by Ma and Taylor (2020) and the present one seem to confirm that the microbiome transmissions between men and women either through sexual intercourse on ecological time scale or through other means on evolutionary time scale all support the Red Queen hypothesis, namely, that the co-evolution between reproductive system microbiomes and hosts facilitates the sexual reproduction (sexual selection). However, we must reiterate the hypothetic nature of our discussion, that is, all assumptions are subject to further experimental and/or theoretical analyses (verifications).

In summary, this study, integrated with Ma and Taylor (2020), appears to cast relatively complete and reasonably strong evidence to support the extension of the classic Red Queen theory to the field of human reproductive system microbiome. That is, the co-evolution between human reproductive systems and their symbiotic microbiomes should facilitate the sexual reproduction. As the title of the classic monograph "*The Ecological Theater and the Evolutionary Play*" by G. E. Hutchinson (1965), implied, it is the ecology that sets theater (environment background) for evolution (adaptation) to act. We believe that the extension of the classic Red Queen hypothesis to the field of reproductive system microbiomes highlights the critical importance of symbiotic microbes to the success of sexual reproduction, on which our current understanding is still rather limited. Therefore, future theoretic and experimental studies from both ecological and evolutionary perspectives are dearly needed.

Finally, this study possesses several limitations that should be mentioned here. First, the discussion of microbiome transmission is primarily one way from male to female given that only one-time semen sample was taken from each couple in the reanalyzed datasets of Mändar et al. (2015). Second, Mändar et al. (2015) study was originally designed to investigate the relationships between infertility and microbiomes, but the implications of the infertility to the results presented in this reanalysis of their data are unknown due to lack of controls. Third, other factors such as multiple sexual partners, occurrences of diseases such as BV or HIV, are not considered in this study, and their implications are unknown. Fourth, no Type-II error analysis is performed in this study, which could detect false negatives in the neutrality tests or the potential non-neutral processes in those cases that have passed the neutrality test (Ma, 2021a,b). Fifth, as correctly pointed out by an anonymous expert reviewer, Mändar et al. (2015) study used the V6 region, which is not commonly used in vaginal or seminal microbiome studies given its limited ability to resolve gynecological taxa. Furthermore, the database used by Mändar et al. (2015) bioinformatics analysis, i.e., the Greengenes, is somewhat outdated and lacks representatives of understudied niches. Despite these unknown implications, we feel that the findings of this study are very likely robust against most of the additional factors. Part of the somewhat circular arguments comes from the prediction (expectation) of the Red Queen hypothesis. Given these limitations, it should be reiterated that findings in this study should be treated as postulations or evidence to support existing hypotheses (particularly the Red

Queen hypothesis). Sometimes, the evidence is indirect or even conjectural, and further experimental and/or theoretical studies are necessary to cross-verify our findings.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The source of the datasets is Mändar et al. (2015).

AUTHOR CONTRIBUTIONS

ZM designed the study, interpreted the results, and wrote the manuscript.

FUNDING

This study received funding from the following sources: A National Natural Science Foundation (NSFC) Grant (No. 31970116) on "Medical Ecology of Human Microbiome"; The Cloud-Ridge Industry Technology Leader Award.

ACKNOWLEDGMENTS

I am deeply indebted to R. Mändar for generous assistance in providing us the OTU tables from their original research (Mändar et al., 2015), which is used in this study. I appreciate the computational assistance from Lianwei Li of the Chinese Academy of Sciences (CAS).

REFERENCES

Aleru, O., and Barber, M. F. (2020). Battlefronts of evolutionary conflict between bacteria and animal hosts. *PLoS Pathog.* 16:e1008797. doi: 10.1371/journal.ppat.1008797

Alonso, D., Etienne, R., and Mckane, A. (2006). The merits of neutral theory. *Trends Ecol. Evol.* 21, 451–457. doi: 10.1016/j.tree.2006.03.019

Antoniak C. E. (1974) Mixtures of dirichlet processes with applications to bayesian nonparametric problems. *Ann. Statist*. 2, 1152–1174.

Borovkova, N., Korrovits, P., Ausmees, K., Türk, S., Jõers, K., Punab, M., et al., (2011). Influence of sexual intercourse on genital tract microbiota in infertile couples. *Anaerobe* 17, 414–418. doi: 10.1016/j.anaerobe.2011.04.015

Etienne, R. S. (2007). A neutral sampling formula for multiple samples and an 'exact' test of neutrality. *Ecol. Lett.* 10, 608–618. doi: 10.1111/j.1461-0248.2007.01052.x

Gajer, P., Brotman, R. M., Bai, G., Sakamoto, J., Schütte, U. M., Zhong, X., et al., (2012). Temporal dynamics of the human vaginal microbiota. *Sci. Transl. Med.* 4, 132–152.

Harris, K., Parsons, T. L., Ijaz, U. Z., Lahti, L., Holmes, I., and Quince, C. (2017). Linking statistical and ecological theory: hubbell's unified neutral theory of biodiversity as a hierarchical dirichlet process. *Proc. IEEE* 105, 516–529. doi: 10.1109/jproc.2015.2428213

Hou, D., Zhou, X., Zhong, X., Settles, M. L., Herring, J., Wang, L., et al., (2013). Microbiota of the seminal fluid from healthy and infertile men. *Fertil. Steril.* 100, 1261–1269. doi: 10.1016/j.fertnstert.2013.07.1991

Hubbell, S. P. (2001). *The Unified Neutral Theory Of Biodiversity And Biogeography*. Princeton, NJ: Princeton University Press.

Hubbell, S. P. (2006). Neutral theory and the evolution of ecological equivalence. *Ecology* 87, 1387–1398. doi: 10.1890/0012-9658(2006)87[1387:ntateo]2.0.co;2

Hutchinson, G. E. (1965). *The Ecological Theater and the Evolutionary Play*. London: Yale University Press.

Leftwich, P. T., Clarke, N. V. E., Hutchings, M. I., and Chapman, T. (2017). Gut microbiomes and reproductive isolation in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 12767–12772.

Lewis, F. M., Bernstein, K. T., and Aral, S. O. (2017). Vaginal microbiome and its relationship to behavior, sexual health, and sexually transmitted diseases. *Obstet Gynecol.* 129, 643–654. doi: 10.1097/AOG.0000000000001932

Li, L., and Ma, Z. S. (2016). Testing the neutral theory of biodiversity with human microbiome datasets. *Sci. Rep.* 6:31448. doi: 10.1038/srep31448

Liu, C. M., Hungate, B. A., Tobian, A. A., Ravel, J., Prodger, J. L., Serwadda, D., et al., (2015). Penile microbiota and female partner bacterial vaginosis in rakai, uganda. *MBio* 6:e00589. doi: 10.1128/mBio.00589-15

Ma, B., Forney, L. J., and Ravel, J. (2012). The vaginal microbiome: rethinking health and disease. *Ann. Rev. Microbiol.* 66, 371–389. doi: 10.1146/annurev-micro-092611-150157

Ma, Z. S. (2020). Niche-neutral theoretic approach to mechanisms underlying the biodiversity and biogeography of human microbiomes. *Evol. Appl.* 14, 322–334. doi: 10.1111/eva.13116

Ma, Z. S. (2021a). Cross-scale analyses of animal and human gut microbiome assemblies from metacommunity to global landscape. *mSystems* 6, e633–e621. doi: 10.1128/mSystems.00633-21

Ma, Z. S. (2021b). Philosophical skepticism concerning the neutral theory or randomness: misplaced or misconceived? A reply to Madison, "Stochasticity and randomness in community assembly: real or as-if?". *mSystems* 6:e0101421. doi: 10.1128/mSystems.01014-21

Ma, Z. S., and Li, W. (2019). How man and woman are different in their microbiome: ecological and network analyses of the microgenderome. *Adv. Sci.* 6:1902054.

Ma, Z. S., and Taylor, R. A. J. (2020). Human reproductive system microbiomes exhibited significantly different heterogeneity scaling with gut microbiome, but the intra-system scaling is invariant. *Oikos* 129, 903–911.

Macklaim, J. M., Fernandes, A. D., Di Bella, J. M., Hammond, J. A., Reid, G., and Gloor, G. B. (2013). Comparative meta-RNA-seq of the vaginal microbiota and differential expression by *Lactobacillus* iners in health and dysbiosis. *Microbiome* 1:12. doi: 10.1186/2049-2618-1-12

Mändar, R., Punab, M., Borovkova, N., Lapp, E., Kiiker, R., Korrovits, P., et al., (2015). Complementary seminovaginal microbiome in couples. *Res. Microbiol.* 166, 440–447. doi: 10.1016/j.resmic.2015.03.009

Nelson, D. E., Dong, Q., Van der Pol, B., Toh, E., Fan, B., Katz, B. P., et al., (2012). Bacterial communities of the coronal sulcus and distal urethra of adolescent males. *PLoS One* 7:e36298. doi: 10.1371/journal.pone.0036298

Noyes, N., Cho, K. C., Ravel, J., Forney, L. J., and Abdo, Z. (2018). Associations between sexual habits, menstrual hygiene practices, demographics and the vaginal microbiome as revealed by Bayesian network analysis. *PLoS One* 13:e0191625. doi: 10.1371/journal.pone.0191625

Papkou, A., Guzella, T., Yang, W., Koeppe, S., Pees, B., Schalkowski, R., et al., (2018). The genomic basis of red queen dynamics during rapid reciprocal host–pathogen coevolution. *Proc. Natl. Acad. Sci. U.S.A.* 116:201810402. doi: 10.1073/pnas.1810402116

Plummer, E. L., Vodstrcil, L. A., Danielewski, J. A., Murray, G. L., Fairley, C. K., Garland, S. M., et al., (2018). Combined oral and topical antimicrobial therapy for male partners of women with bacterial vaginosis: acceptability, tolerability and impact on the genital microbiota of couples A pilot study. *PLoS One* 13:e0190199. doi: 10.1371/journal.pone.0190199

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S., McCulle, S. L., et al., (2011). Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4680–4687.

Rosenberg, E., and Zilber-Rosenberg, I. (2018). The hologenome concept of evolution after 10 years. *Microbiome* 2018:78. doi: 10.1186/s40168-018-0457-9

Rosenberg, E., Sharon, G., and Zilber-Rosenberg, I. (2009). The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework. *Environ. Microbiol.* 11, 2959–2962. doi: 10.1111/j.1462-2920.2009.01995.x

Schneider, D. I., Ehrman, L., Engl, T., Kaltenpoth, M., Hua-Van, A., Le Rouzic, A., et al., (2019). Symbiont-driven male mating success in the neotropical drosophila paulistorum superspecies. *Behav. Genet.* 49, 83–98. doi: 10.1007/s10519-018-9937-8

Scoville, H. (2019). *What is the Red Queen Hypothesis?*. Available online at: https://www.thoughtco.com/red-queen-hypothesis-1224710 (Accessed October 1, 2020).

Shapiro, J. A. (2017). Biological action in Read–Write genome evolution. *Interface Focus* 7:20160115. doi: 10.1098/rsfs.2016.0115

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ. Microbiol.* 8, 732–740. doi: 10.1111/j.1462-2920.2005.00956.x

Sloan, W. T., Woodcock, S., Lunn, M., Head, I. M., and Curtis, T. P. (2007). Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microb. Ecol.* 53, 443–455. doi: 10.1007/s00248-006-9141-x

Smith, S. B., and Ravel, J. (2017). The vaginal microbiota, host defence and reproductive physiology. *J. Physiol.* 595, 451–463. doi: 10.1113/JP271694

Starnbach, M. N., and Roan, N. R. (2008). Conquering sexually transmitted diseases. *Nat. Rev. Immunol.* 8, 313–317. doi: 10.1038/nri2272

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature* 189, 732–735. doi: 10.1038/189732a0

Taylor, R. A. J. (2019). *Taylor's Power Law: Order and Pattern in Nature*. London: Academic Press, 657.

Teh, Y. W., Jordan, M. I., Beal, M. J., and Blei, D. M. (2006). Hierarchical dirichlet processes. *J. Am. Stat. Assoc.* 101, 1566–1581.

van de Wijgert, J. H. H. M. (2017). The vaginal microbiome and sexually transmitted infections are interlinked: consequences for treatment and prevention. *PLoS Med.* 14:e1002478. doi: 10.1371/journal.pmed.1002478

Van Valen, L. (1973). A new evolutionary law. *Evol. Theory* 1, 1–30. doi: 10.1016/j.jhevol.2008.09.007

Vodstrcil, L. A., Twin, J., Garland, S. M., Fairley, C. K., Hocking, J. S., Law, M. G., et al., (2017). The influence of sexual activity on the vaginal microbiota and *Gardnerella vaginalis* clade diversity in young women. *PLoS One* 2:e0171856. doi: 10.1371/journal.pone.0171856

Weng, S.-L., Chiu, C.-M., Lin, F.-M., Huang, W.-C., Liang, C., Yang, T., et al., (2014). Bacterial communities in semen from men of infertile couples: metagenomic sequencing reveals relationships of seminal microbiota to semen quality. *PLoS One* 9:e110152. doi: 10.1371/journal.pone.0110152

Žliobaitė, I., Fortelius, M., and Stenseth, N. C. (2017). Reconciling taxon senescence with the Red Queen's hypothesis. *Nature* 552, 92–95. doi: 10.1038/nature24656

Zozaya, M., Ferris, M. J., Siren, J. D., Lillis, R., Myers, L., Nsuami, M. J., et al., (2016). Bacterial communities in penile skin, male urethra, and vaginas of heterosexual couples with and without bacterial vaginosis. *Microbiome* 4:16. doi: 10.1186/s40168-016-0161-6

Conflict of Interest: The author declares 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.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 05 April 2022
doi: 10.3389/fmicb.2022.783923

Influence of Association Network Properties and Ecological Assembly of the Foliar Fugal Community on Crop Quality

Lei Xing[1,2], Qiqi Zhi[3], Xi Hu[2], Lulu Liu[2], Heng Xu[2], Ting Zhou[2], Huaqun Yin[3], Zhenxie Yi[1]* and Juan Li[1]**

[1] College of Agronomy, Hunan Agricultural University, Changsha, China, [2] Great Wall Cigar Factory, China Tobacco Sichuan Industrial Co., Ltd, Shifang, China, [3] School of Minerals Processing and Bioengineering, Central South University, Changsha, China

Revealing community assembly and their impacts on ecosystem service is a core issue in microbial ecology. However, what ecological factors play dominant roles in phyllosphere fungal community assembly and how they link to crop quality are largely unknown. Here, we applied internal transcriptional spacer high-throughput sequencing to investigate foliar fungal community assembly across three cultivars of a Solanaceae crop (tobacco) and two planting regions with different climatic conditions. Network analyses were used to reveal the pattern in foliar fungal co-occurrence, and phylogenetic null model analysis was used to elucidate the ecological assembly of foliar fungal communities. We found that the sensory quality of crop leaves and the composition of foliar fungal community varied significantly across planting regions and cultivars. In Guangcun (GC), a region with relatively high humidity and low precipitation, there was a higher diversity and more unique fungal species than the region of Wuzhishan (WZS). Further, we found that the association network of foliar fungal communities in GC was more complex than that in WZS, and the network properties were closely related to the sensory quality of crop. Finally, the results of the phylogenetic analyses show that the stochastic processes played important roles in the foliar fungal community assembly, and their relative importance was significantly correlated with the sensory quality of crop leaves, which implies that ecological assembly processes could affect crop quality. Taken together, our results highlight that climatic conditions, and plant cultivars play key roles in the assembly of foliar fungal communities and crop quality, which enhances our understanding of the connections between the phyllosphere microbiome and ecosystem services, especially in agricultural production.

Keywords: microbial community, phyllosphere, fungi, network, phylogenetic structure, null model, crop quality

OPEN ACCESS

Edited by:
Daliang Ning,
University of Oklahoma, United States

Reviewed by:
Jinbo Xiong,
Ningbo University, China
Xiao-Xia Zhang,
Institute of Agricultural Resources and Regional Planning (CAAS), China

****Correspondence:***
Huaqun Yin
yinhuaqun_cs@sina.com
Zhenxie Yi
yizhenxie@126.com
Juan Li
282617472@qq.com

Specialty section:
This article was submitted to Systems Microbiology, a section of the journal Frontiers in Microbiology

Received: *27 September 2021*
Accepted: *26 February 2022*
Published: *05 April 2022*

Citation:
Xing L, Zhi Q, Hu X, Liu L, Xu H, Zhou T, Yin H, Yi Z and Li J (2022) Influence of Association Network Properties and Ecological Assembly of the Foliar Fugal Community on Crop Quality. Front. Microbiol. 13:783923. doi: 10.3389/fmicb.2022.783923

INTRODUCTION

Assembly patterns within microbial communities are an important topic in microbial ecology and are closely related to the functioning of plant-associated ecosystems (Konopka, 2009). There are many different factors that could affect the structure of the plant-associated microbial community. For example, soil microbial communities were different across various plant species

(Ma et al., 2019), because of the range of nutrient content available in the leaf and root litter that alters decomposer abundance (Otsing et al., 2018). The foliar endophytic fungal community in *Cirsium arvense* could be associated with the soil nutrients and arbuscular mycorrhizal (AM) colonization (Eschen et al., 2010), and the latter would be further affected by root exudates, such as methyl salicylic acid and acibenzolar-*S*-methyl (Mannaa et al., 2020). Furthermore, climate change, such as warming, would decrease fungal species richness and change foliar fungal community composition, especially at the end of the growing season (Faticov et al., 2021). In the case of elevated atmospheric carbon dioxide, the growth of trees would also lead to the changes in the composition of microbial communities that colonize the fallen leaves (Kelly et al., 2010). Crop cultivar, tissue type, and climatic factors can all significantly influence phyllosphere fungal community structure; moreover, location-dependent climate conditions could contribute to the differences in abundance, diversity, and presence of genera containing pathogens, whereas the root communities were less affected by climatic factors (Latz et al., 2021). In addition, drought changes the composition of the root microbiome, where changes in the relative abundance of specific bacterial groups were associated with increased drought tolerance in plants (Fitzpatrick et al., 2018). Although bacterial abundance was negatively affected by O_3 stress, it was found that the fungal abundance was substantially stimulated (up to 12-fold compared with non-fumigated plants at 20°C). These changes were accompanied by modifications of the genetic structures and a relative increase in amino acids catabolism (Changey et al., 2018). The above findings advanced our understanding of the drivers in shaping plant-associated communities, but the mechanisms of assembly in the foliar fungal community remain largely unknown.

Exploring network assembly in microbial communities and their responses to environmental changes is fundamentally vital for the understanding of community organization (Zhou et al., 2010). Microbial community networks can provide a mechanistic association between species in a specific environment and information on the dynamics of community structure as a function of time or other external variables (Cardona et al., 2016). For example, climate change, such as warming, can significantly increase network complexity, including network size, connectivity, and number of keystone species (Faticov et al., 2021), whereas elevated CO_2 can increase modularity and hierarchy (Zhou et al., 2010). The community assembly of plant-associated microbes may have some differences, such as the rhizosphere microbial networks. In the rhizosphere, complexity of microbial ecological network increases with the growth of wild oat plants (Shi et al., 2016). *Artemisia annua* (sweet wormwood) promoted a specific root-associated microbial community assembly process, with increased abundance of plant growth–promoting microorganisms and building of interkingdom association networks, which may be beneficial for the fitness of plants in the natural environment (Shi et al., 2021). Together, these results revealed that network assembly of plant-associated microbial communities could be closely related to plant growth. However, how network assembly of the phyllosphere microbial community is affected by climatic conditions and crop cultivars have been less well studied.

Phylogenetic analyses based on null model provide a conceptual background for understanding the ecological processes of community assembly that determine which, and how many, species live in a particular environment (Campbell et al., 2011). Foliar fungi are of great importance to host plant growth and health and can also affect ecosystem functions. Most importantly, host environmental filtering caused by fungal infections outweighs competitive exclusion in driving foliar fungal community assembly in symptomatic leaves (Liu X. et al., 2021). Community co-occurrence theories can be explained mainly by niche-based theory and the null model (Gravel et al., 2006; Jiao et al., 2020). On the one hand, niche-based theory (Zhou and Ning, 2017) posits that deterministic processes play a key role in the community assembly process. Different species occupy different niches, and ecological selection can affect the community co-occurrence. On the other hand, the neutral model demonstrates that all species are equivalent on ecological function, and the community assembly is affected by stochastic processes but not their ecological abilities. The environment can play a vital role in stochastic processes that correlate with community assembly. However, the deterministic and stochastic processes that shape phyllosphere fungal community assembly have not been extensively explored.

The present study aims to reveal the assembly of phyllosphere fungal communities inhabiting a Solanaceae (tobacco) crop across climatic conditions and cultivars. We set up a large-scale field experiment with three crop cultivars and in areas with different climatic conditions in Hainan, China. We assessed the sensory quality of crop leaves and explored the phyllosphere fungal communities using internal transcriptional spacer (ITS) high-throughput sequencing technology. We hypothesized that (i) the sensory quality of crop leaves and the composition of the foliar fungal community are significantly affected by both the planting region and cultivar; (ii) the association network of foliar fungal communities and its characteristics are closely related to the sensory quality of the crop; (iii) the community assembly of all samples would be dominated by the ecological drift, which is community phylogeny structure with little effect on the sensory quality of the crop.

MATERIALS AND METHODS

Experimental Description

Three cultivars of the Solanaceae crop tobacco, including Haiyan 101 (HY101), Haiyan 201 (HY201), and Haiyan 109 (HY109), were used to conduct a field experiment in two regions of Hainan, China, at Guangcun (GC) town, Danzhou (19°49′ N, 109°28′ E) and Panyang town, Wuzhishan (WZS) City (18°87′ N, 109°40′ E). The three selected cultivars are core cultivars with consistent disease resistance and agronomic characteristics; moreover, each cultivar has different sensory quality traits. GC town is located in the northwest of Hainan, at an elevation of 51.6 m; it is a semihumid region. Mountain of five fingers (WZS), with an elevation of 154 m, is located on the central line of Hainan, which is the highest point in Hainan Province and is a humid mountain region. During the crop growth period in GC town, the average temperature was 23°C, average humidity

was 84%, total precipitation was 89 mm, and the average photosynthetically effective radiation was 306 μmol m^{-2} s^{-1}. In WZS, the average temperature was 23°C, average humidity was 78%, total precipitation was 132 mm, and the average photosynthetic effective radiation was 338 μmol m^{-2} s^{-1}.

The experiment was conducted from November 2018 to December 2019. The plants were transplanted in GC town on January 20 and in Panyang town on January 13. The experiment adopted a random block design with three replicates, the plot area was 90 m^2, and the row spacing was 40 × 100 cm. Other field management measures were carried out in accordance with local planting practices (Lei et al., 2021).

On April 17, 2019, the middle part of the fresh blade was collected at the maturity stage of the crop. A total of 18 plants were randomly selected from each plot, and the middle leaves of every sixth plant were taken as a sample, and each sample was divided into two parts. One part was kept at 4°C and brought back to the laboratory for subsequent indoor foliar microbial DNA extraction experiments. The other part was used to evaluate the sensory quality of the leaves.

Sensory Qualities of Crop Leaves

The evaluation in sensory quality of leaf aroma substances was based on the Sichuan China Tobacco "mellow and sweet" category to construct a leaf raw material evaluation table (9-point system), from which the average level was taken. The main evaluation contents were as follows: (i) coordination of flammability indicators: a sense of balance; (ii) combustion characteristics: gray, condensed gray, and combustibility; (iii) the mellowness of smoke: lingering, fullness, and smoothness; (iv) the mellowness of aroma: maturity, richness, and mellowness; (v) the mellow aftertaste: irritation, aftertaste, sweetness, and cleanliness; (vi) miscellaneous gas: woody gas, soil fishy gas, green mixed gas, burnt gas, and protein smell; (vii) fragrance: hay, floral, cellar, milk, woody, bean, sweet, honey, leather, baking, normal, resin, powder, and burnt sweet.

DNA Extraction and High-Throughput Sequencing

Fifteen grams of leaf samples obtained from various parts of the leaf surface (avoiding the main and branch veins) using a sterile puncher was added to 50 mL of 0.1% Tween-80 bacterial phosphate buffer (pH 7.0). The samples were then shaken for 30 min at 170 revolutions/min (rpm) and 28°C, the bacterial suspension was collected, and the leaf samples were washed twice more. The collected suspensions were centrifuged for 15 min (4°C, 10,000 rpm) to pellet the microorganisms. The pellet was suspended in sterile water and washed three times. Finally, the microorganisms were resuspended with 1 mL of sterile water for subsequent DNA extraction. After the above treatments, the leaves were rinsed three times with sterile water, with the last rinsing solution (1 mL) spread on LB plate medium and cultured in an incubator at 30°C for 2 days to determine whether the microorganisms on the surface of the leaves were completely eluted.

Genomic DNA extraction of foliar microorganisms was performed using the Plant Genomic DNA Kit (Plant Genomic DNA Kit) following the manufacturer's protocol. We used the primer pair ITS1-1F-F/ITS1-1F-R with barcodes to amplify the ITS. Amplicons were sequenced by Illumina NovaSeq platform with a Plant Genomic DNA Kit (2 × 250–base-pairs [bp] paired ends). The raw sequencing data were deposited in the NCBI Sequence Read Archive database according to accession number PRJNA778452.

Sequencing Processing and Statistical Analyses

Raw sequences were split into sample libraries with perfect matches to barcodes. Low-quality sequences with QC < 20 over a 5-bp window size were trimmed using Btrim, and sequences with a length of < 100 bp were removed (Kong, 2011). Then, the forward and reverse sequences were spliced together. Any sequences containing ambiguous bases or the incorrect length were removed, and remaining sequences were compared against the UNITE v8.2 database (Kõljalg et al., 2005) to remove possible chimeras. The length of the sequencing fragment was 200–400 bp. Then, UPARSE (Edgar, 2013) was used to cluster and produce operational taxonomic units (OTUs) at 97% similarity level. In order to ensure the authenticity of the data, we removed OTUs that were represented by only one sequence in overall data (global singletons). For comparability between different samples in subsequent data analysis, the ITS sequences were resampled to 10,000 per sample. Finally, RDP Classifier (Wang et al., 2007) was used to perform online comparison and systematic taxonomic annotation of the ITS sequences. The above analyses were conducted on a Galaxy Illumina sequencing analysis platform (Zhou et al., 2016) publicly available at http://zhoulab5.rccc.ou.edu:8080/.

Network Construction

To construct a microbial association network, correlations between pairwise OTUs that were present in more than a half of the samples were calculated using the SparCC method (Friedman and Alm, 2012; Preheim et al., 2013; Ju and Zhang, 2015; Watts et al., 2019) using the "microeco" v0.3.1 package (Liu C. et al., 2021) in R. Only significant correlations ($p < 0.01$) larger than 0.3 were retained for network construction. Network analyses were conducted in "igraph" v1.2.6 package (Ju et al., 2016). Some traits of association networks such as nodes, links, density, transitivity, modularity, centralization of degree, the average path distances, and diameter based on these methods were analyzed. The impact of ecological environment factors and crop cultivar on the structure of the foliar fungal community using permutational multivariate analysis of variance (PERMANOVA) using distance matrices (Alekseyenko, 2016).

Phylogenetic Analyses Based on Null Model

The fungal phylogenetic tree was constructed by FastTree v2.1.11 (Price et al., 2010). Considering the high degree of variation of ITS sequences, we performed constrained topology search based

on a guide tree, as reported previously (Nuccio et al., 2016). The guide tree was built from the full-length SSU sequences of 386 representative species with one representative for each fungal family. The SSU sequences were from Silva 138.1 Ref NR database (Quast et al., 2013). The nearest taxon index (NTI, i.e., negative of the standardized effect sizes of mean nearest taxon distance) and net relatedness index (NRI, i.e., negative of the standardized effect sizes of mean pairwise phylogenetic distance) of foliar fungal communities were calculated using "picante" v1.8.2 package in R (Kembel et al., 2010). The phylogenetic signals between closely species were estimated by Mantel correlogram (Borcard and Legendre, 2012). The ecological assembly process of the microbial community was assessed by using "microeco" v0.3.1 package (Stegen et al., 2013, 2015; Liu C. et al., 2021). The visualization of molecular ecological network was realized by Gephi 0.9.1 (Hernandez-Garcia et al., 2016).

Statistical Analysis

All packages discussed here for analyses were run in R v3.6 environment (Ihaka and Gentleman, 1996; R CoreTeam, 2013). The α diversity indices, including Shannon, Simpson, inverse Simpson, and Chao1, were calculated by using "diversity" function in the "vegan" v2.5-6 package (Dixon, 2003). The effects of crop cultivars and planting regions on microbial community structure were analyzed by principal components analysis (PCA) based on weighted UniFrac distance using the "rda" function. Samples were clustered based on sensory quality by using the "heatmap" function in R (Zhao et al., 2014). Venn plot analyses were conducted by using "VennDiagram" v1.6.20 (Chen and Boutros, 2011). For the correlations between sensory quality and diversity, phylogenetic indices were calculated by Pearson correlation. The differences in diversity and network properties were tested by least significant difference test.

RESULTS

Sensory Quality of Crop Leaves

The sensory quality of crop leaves was significantly affected by cultivar and ecological region. We estimated sensory quality indices using seven aspects, including coordination of flammability indicators, combustion characteristics, mellowness of smoke, mellowness of aroma, mellow aftertaste, miscellaneous gas, and fragrance. Our results showed that the combustibility of HY201 and HY109 displayed good performance in both planting regions, whereas HY101 had the lowest scores. Fragrance was a key indicator that determined the style and characteristics of crop leaves. On the whole, the samples from GC mainly displayed a sweet fragrance, whereas hay fragrance was more prominent in the samples from WZS. GC-HY201 had the highest scores in the mellowness of aroma, the mellowness, and the purity of aftertaste, and its corresponding sense of balance was also the highest, which showed obvious differences with WZS-HY201 and indicated that ecological environment had a great influence on the aroma of HY201.

The multiplicity in sensory quality across planting regions was differed with cultivar. The results of clustering analysis based on euclidean distance showed that the quality of WZS-HY201 was close to that of WZS-HY109, whereas GC-HY101 was close to that of WZS-HY101, which indicated that these two pairs of samples had, respectively, similar sensory quality (**Figure 1A**). The PCA results of sensory quality scores suggested that the sensory quality varied significantly across planting regions and cultivars, and the impact of planting regions was greater than that of cultivars (**Figure 1B**).

Associations Between Fungal Diversity and the Sensory Quality on Crop Leaf Surface

The results of non-metric multidimensional scaling of foliar fungal community showed that the composition of foliar fungal community varied significantly across planting regions and cultivars (**Figure 2A**). GC, with relative high humidity and low precipitation, had higher diversity, but more unique fungal species, than the region of WZS. The boxplots with stars showed the correlation between the phylogenetic diversity (PD) and the two regions (**Figure 2B**). As for the association between fungal diversity and sensory quality, the heatmap results indicated that the observed species number and Chao1 were significantly correlated ($p < 0.01$) with the mellowness of aftertaste, fragrance, and smoke, such as the cleanliness, sweetness, and mellowness (**Supplementary Figure 1**). Simpson diversity was significantly correlated ($p < 0.01$) with the hay fragrance, and inverse Simpson diversity was significantly correlated ($p < 0.01$) with maturity, aftertaste, and mellowness (**Supplementary Figure 1**). Linear regression analyzed the relationships between PD and differences in sensory quality of crop leaves (**Supplementary Figure 2**; the solid lines represent $p < 0.05$). In terms of the overall fragrance certain honey aroma, floral, cellar, and leather fragrances were observed in GC HY201, whereas hay and woody fragrances were more prominent among the samples of WZS HY201. In summary, the diversity of leaf fungi had a close relationship to its sensory quality.

Microbial Co-occurrence of Fungi on Leaf Surface

In order to explore the leaf fungal association networks, we applied the SparCC method (Watts et al., 2019) to construct molecular ecology networks of foliar fungi across all samples (**Figure 3**). Further the trendline with scatter plots suggested the correlation between the network properties and the sensory quality of crop (**Figure 4**). The mellowness of aroma: mellowness, richness and maturity, and the mellowness of smoke: fullness and lingering were significantly related with the nodes ($p < 0.01$). The miscellaneous gas: protein smell was significantly related with the transitivity ($p < 0.01$). The fragrance: honey and the mellowness of aroma: mellowness and maturity were significantly related with diameter ($p < 0.01$). A subnetwork for each sample was extracted based on the presence of nodes, and the associations between their properties and sensory quality were accessed by a linear model (**Table 1**). Our results showed that the number of nodes and connections of ecological networks

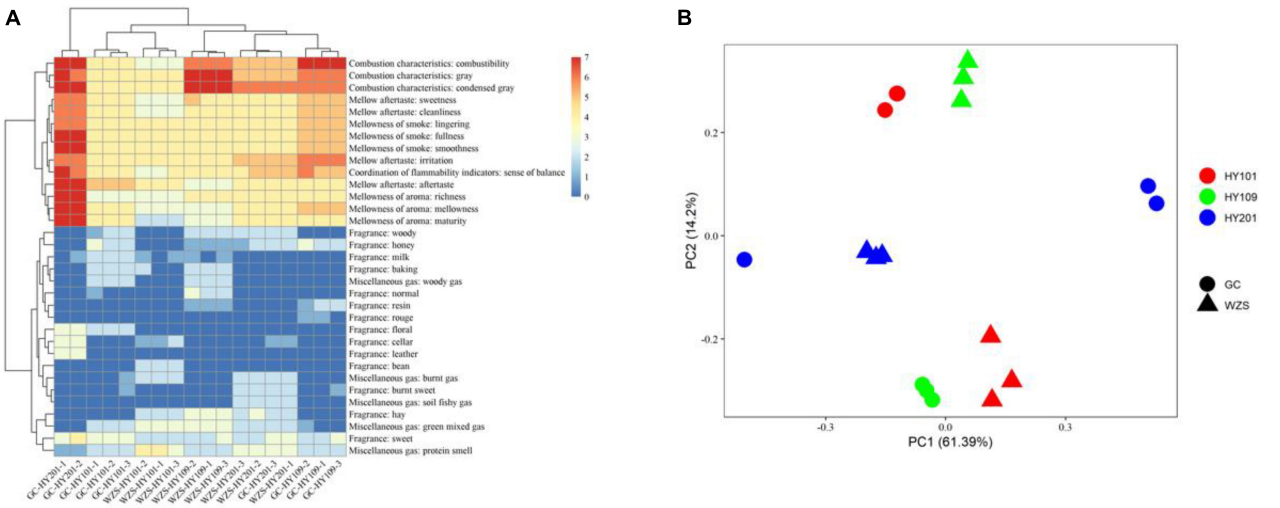


FIGURE 1 | (A) Heatmap for sensory quality of crop leaves. Color represents the scores of sensory qualities. Clustering analysis was based on Euclidean distance. **(B)** Principal components analysis for sensory quality of crop leaves across three crop cultivars and two planting regions. Color denotes different crop cultivars. Shape denotes different regions.

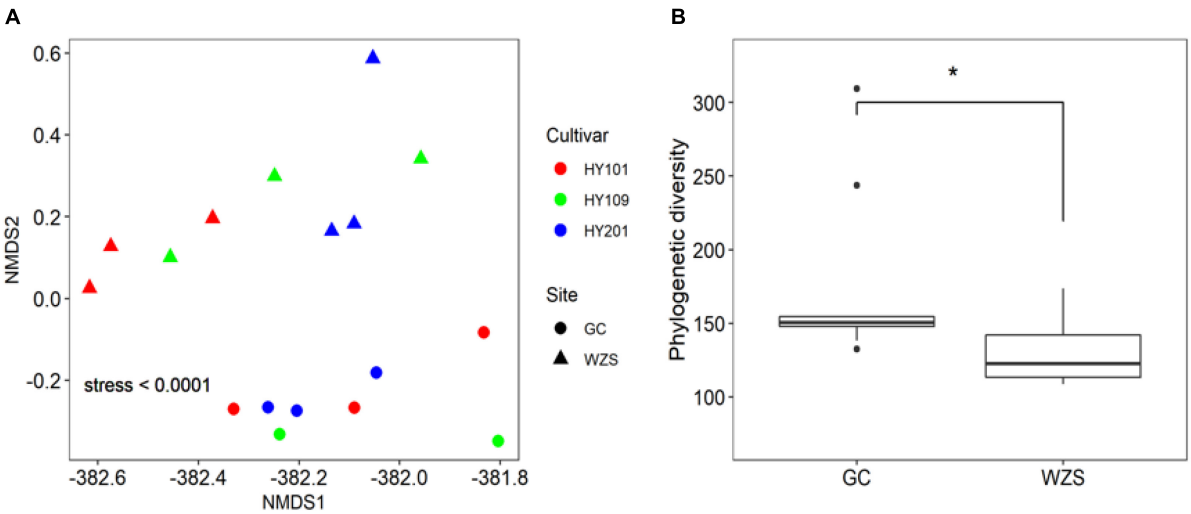


FIGURE 2 | (A) Non-metric multidimensional scaling (NMDS) based on Bray–Curtis similarity index of foliar fungal community. The composition of foliar fungal community varied significantly across planting regions and cultivars. **(B)** Difference of fungal phylogenetic diversity based on Kruskal–Wallis test between the two regions (GC and WZS, $^*p < 0.05$).

in GC was higher than those in WZS. The subnetwork of GC-HY109 had the greatest number of nodes (450), followed by GC-HY201 (233), WZS-HY109 (191), GC-HY101 (186), WZS-HY201 (173), and WZS-HY101 (148). In GC, the number of connections for HY101, HY109, and HY201 were 2,705, 90,087, and 3,076, respectively. While for WZS, the numbers of connections for HY101, HY109, and HY201 were 1,272, 3,517, and 2,739, respectively. The percentage of positive associations among treatments had no significant differences, whereas the modularity values of WZS-HY101 and GC-HY201 were 0.312 and 0.359, respectively, which were larger than those of GC-HY101 (0.124), GC-HY109 (0.17), WZS-HY109 (0.257), and WZS-HY201 (0.151). These results indicated that the size of subnetworks in GC was larger and more complex than those in WZS and that climatic conditions could affect the network characteristics. The average path length of HY201 was 3.518, followed by HY101 (2.948), and HY109 (2.591) in the GC planting region, which showed that the response speed of GC-HY201 leaf fungi to the external environment was lower than GC-HY101 and GC-HY109 and that the plant cultivar can affect the foliar fungal subnetwork structure of crop leaves. Furthermore, we found that 99.67% of the network nodes (2,093) were peripheral nodes, with only 0.33% of nodes (7) being module hubs, and no network hubs and connectors were detected.

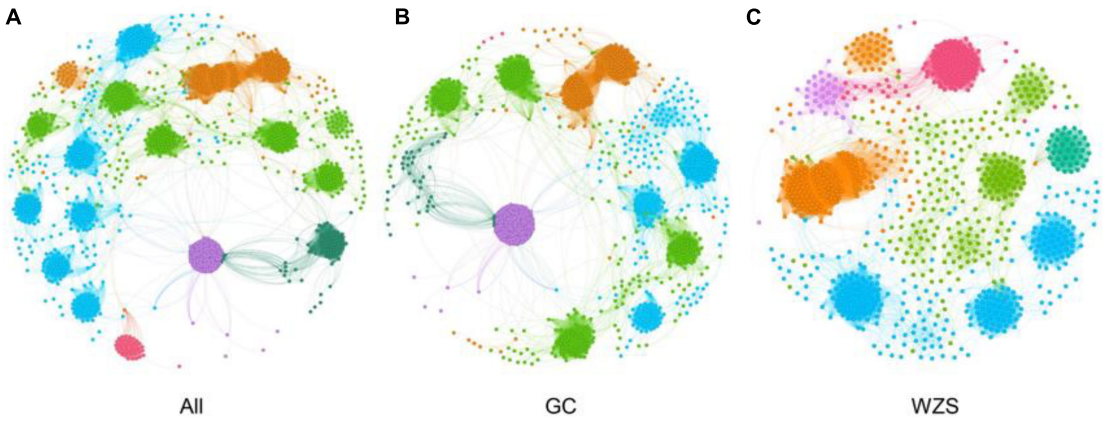


FIGURE 3 | The association network of foliar fungal community. Node color represents different modules. The network pattern was visualized by Gephi software. **(A)** The network of all samples in two regions. **(B)** The subnetwork of the samples in GC. **(C)** The subnetwork of the samples in WZS.

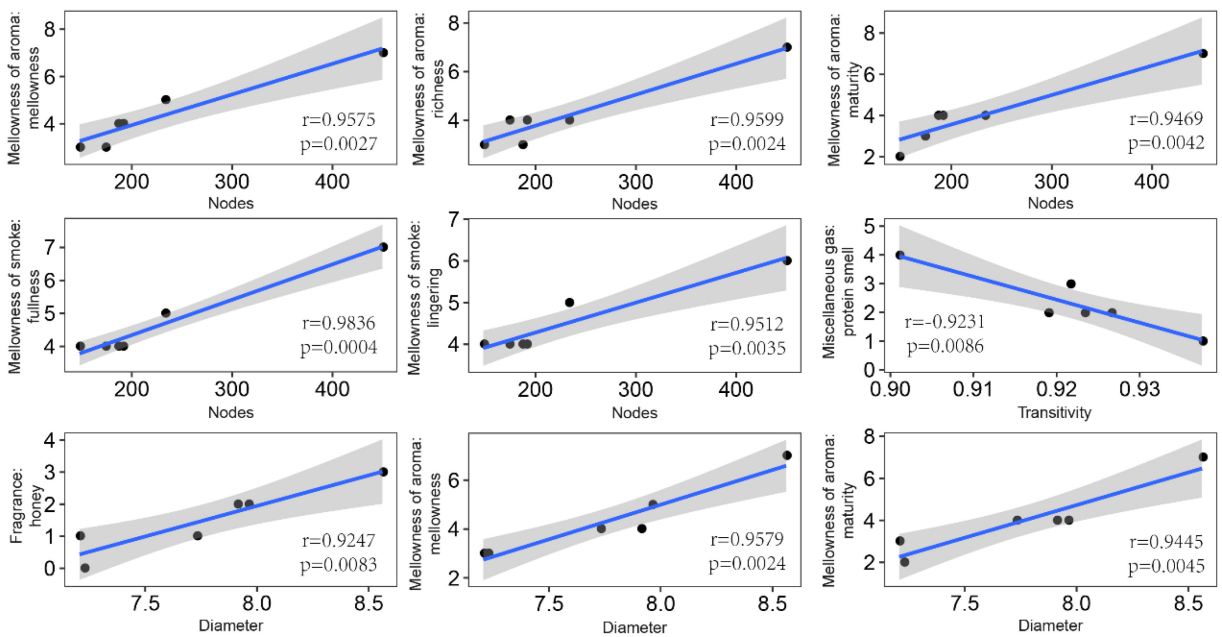


FIGURE 4 | Pearson correlations between the network properties and the sensory quality of crop leaves. The shadow indicates the 95% confidence interval. The miscellaneous gas: protein smell was significantly related with the transitivity ($p < 0.01$). The fragrance: honey, the mellowness of aroma: mellowness, and maturity were significantly related with diameter ($p < 0.01$).

Module hubs in the fungal networks included *Ilyonectria destructans* (OTU 298), *Penicillium brunneoconidiatum* (OTU 566), *Kodamaea ohmeri* (OTU 1026), *Solicoccozyma terricola* (OTU 1898), and *Penicillium maclennaniae* (OTU 1989).

Further, the network properties were significantly related to the sensory quality of the leaves. For example, we found that the number of nodes was significantly ($p < 0.05$) positively correlated with balance, mellowness of smoke, mellowness of aroma, and leather. The link number had a significantly positive correlation ($p < 0.05$) with fluency, richness, leather, and cellar. Density was positively correlated only with the leather ($p < 0.05$; **Supplementary Figure 3**). Furthermore, network transitivity had a significantly negative correlation with protein smell, whereas network diameter had a significantly positive correlation with maturity, alcohol, and sweetness ($p < 0.05$; **Supplementary Figure 3**). Overall, these results indicated that the properties of the leaf fungal microbial ecology networks can affect the leaf sensory quality to some extent.

Ecological Assembly Processes of the Fungal Community on Leaf Surfaces

In order to explore community assembly of the foliar fungal community, we calculated a set of indices based on null

TABLE 1 | Topological properties of molecular ecological subnetworks in foliar fungal communities on crop leaves across crop cultivars and plating regions.

Group	Nodes	Links	Density	Transitivity	Modularity	Centralization of degree	Average path distances	Diameter
GC-HY101	186 (18)[a]	2,705 (585)[a]	0.156 (0.015)[a]	0.919 (0.024)[a]	0.124 (0.02)[a]	0.242 (0.004)[a]	2.948 (0.248)[a]	7.911 (0.937)[a]
GC-HY109	450 (317)[a]	90,087 (145,523)[a]	0.367 (0.366)[a]	0.937 (0.057)[a]	0.17 (0.215)[a]	0.208 (0.092)[a]	2.591 (1.2)[a]	8.561 (2.976)[a]
GC-HY201	233 (24)[a]	3,076 (1,294)[a]	0.113 (0.046)[a]	0.923 (0.047)[a]	0.359 (0.191)[a]	0.179 (0.047)[a]	3.518 (0.64)[a]	7.96 (1.251)[a]
WZS-HY101	148 (22)[a]	1,272 (509)[a]	0.125 (0.075)[a]	0.901 (0.021)[a]	0.312 (0.199)[a]	0.195 (0.05)[a]	2.71 (0.569)[a]	7.226 (1.793)[a]
WZS-HY109	191 (99)[a]	3,517 (4,727)[a]	0.123 (0.062)[a]	0.922 (0.064)[a]	0.257 (0.038)[a]	0.215 (0.049)[a]	2.92 (0.517)[a]	7.73 (1.821)[a]
WZS-HY201	173 (14)[a]	2,739 (1,155)[a]	0.178 (0.055)[a]	0.927 (0.026)[a]	0.151 (0.016)[a]	0.246 (0.022)[a]	2.745 (0.337)[a]	7.206 (1.586)[a]

Values in parentheses are standard deviation (n = 3). The letter label means there are no statistical significances ($p < 0.05$) of network properties among groups, which is tested by multiple comparisons by means of least significant difference, and p-value is adjusted by the Bonferroni method.

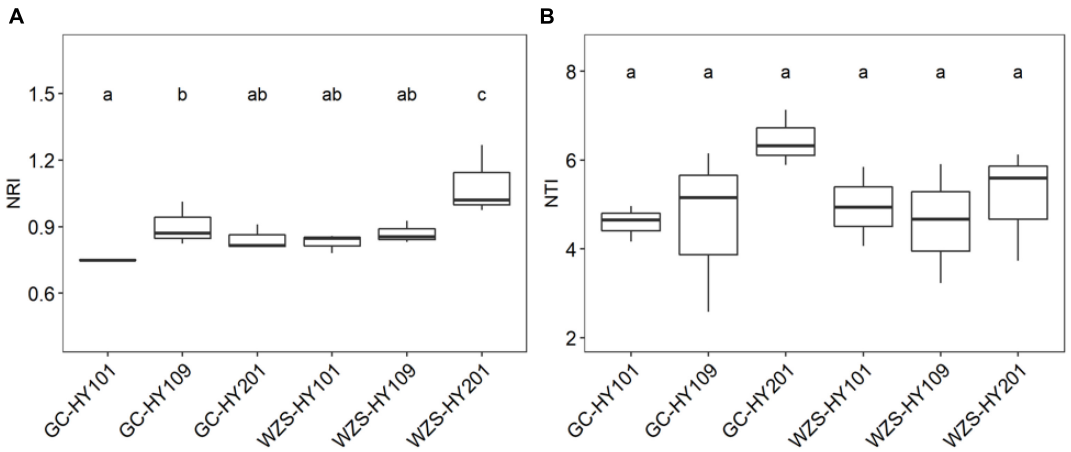


FIGURE 5 | (A) Boxplots for net relatedness index (NRI) for foliar fungal communities. **(B)** Boxplots for nearest taxon index (NTI) for foliar fungal communities. Different small letters indicate significant differences ($p < 0.05$).

model, including NRI, NTI, βNTI, and βNRI (**Figure 5** and **Supplementary Figure 5**; Stegen et al., 2012). We first detected the significant correlations between environmental distance and phylogenetic distances between phylogenetically close species, indicating phylogenetic signal in environment conditions (**Supplementary Figure 4**). Our results showed that NTI and NRI significantly varied among cultivars and planting regions (**Figure 5**). For example, the community NRI of fungi on leaf surfaces in GC region was between 0.74 and 1.01, whereas in WZS, it was between 0.78 and 1.26. The NRI values of the WZS fungal communities were larger than those of GC. The community NTI of fungi on leaf surfaces in GC was between 2.58 and 7.13, whereas in WZS, it was between 3.23 and 6.12. This meant that the community assembly of the foliar fungal community was affected by local α diversity. Pearson correlation analysis showed that there were significant relationships between NRI, NTI, PD, and sensory quality (**Supplementary Table 1**). The results showed that NRI was significantly related to the burnt sweet ($r = 0.525$, $p < 0.05$), the fragrance of milk ($r = -0.541$, $p < 0.05$), and the soil fishy gas ($r = 0.597$, $p < 0.05$); NTI was significantly correlated with the cellar fragrance ($r = 0.508$, $p < 0.05$), whereas the PD was significantly correlated with resin fragrance ($r = 0.663$, $p < 0.05$), rouge fragrance ($r = 0.907$, $p < 0.05$), burnt gas ($r = -0.500$, $p < 0.05$), and combustibility ($r = 0.514$, $p < 0.05$). This result showed that among the sensory quality of the crop, only the fragrance of milk and the soil fishy gas were greatly affected by NRI of the fungus on the leaf surface, and only the cellar fragrance was greatly affected by NTI of the fungus on the leaf surface, whereas the resin fragrance, rouge fragrance, burnt gas, and combustibility were greatly affected by PD of the fungus on the leaf surface. Furthermore, we found that the βNTIs of all foliar fungal communities were larger than 1 and less than 2, whereas the βNRIs of all foliar fungal communities were larger than 0.9 and less than 1.4, and the βNRIs of HY101 and HY109 were significantly different ($p < 0.05$), which indicated that spatial turnover was affected by the stochastic process of foliar fungal community assembly (**Supplementary Figure 5**; Stegen et al., 2012).

The assembly process of the foliar fungal community in GC-HY201 included drift (67%) and dispersal limitation (33%); moreover, in WZS-HY101 and WZS-HY109, it included drift (67%), limiting dispersal (33%, **Table 2**). This implied that stochastic processes could markedly influence crop fungal community assembly. The linear regression analysis showed that βNTI was linearly related to the sensory quality of crop leaves, indicating that the foliar fungal community βNTI had correlation with the differences in foliar sensory quality between pairwise samples (**Supplementary Figure 5**, the solid line represents $p < 0.05$).

DISCUSSION

In our study, we studied various plant cultivars and ecological regions with different climatic conditions in Hainan, China. Our results showed that the sensory quality and foliar fungal community structure showed significant differences across different planting regions and crop cultivars. There were significant associations between the sensory quality of crop leaves and foliar fungal diversity. Furthermore, ecological association networks of crop foliar fungi in GC were more complex than those in WZS, and the community assembly was dominated by ecological drift. Network assembly had significant correlations with sensory qualities, whereas ecological processes of community assembly had little effect on sensory qualities. These results highlighted how the phyllosphere fungal community is closely associated with crop quality.

Effects of Foliar Fungal Community on Crop Quality

Our results demonstrated that planting regions with different climate conditions and cultivars had significant effects on the sensory quality of a Solanaceae crop, which is likely to be led by the changes in foliar fungal community composition. Climate may affect the soil conditions and further alter the fungal community. Moreover, cultivation methods also will change the soil nutrients composition and the microbial network structure in the rhizosphere or foliar fungal communities. This is supported by previous studies of the cucumber rhizosphere, where AM fungi (AMF) significantly altered the nutrient composition of the branches of the host plant, with the strongest contrast observed between cucumber-irregular symbiotic plants and non-mycorrhizal cucumber plants (Ravnskov and Larsen, 2016). The composition of soil fungal communities changed with continuous cucumber cultivation, which may have been caused by the combined cultivation period of cucumber and excessive application of chemical fertilizers (Sun et al., 2021), such as nitrogen fertilizer and phosphate fertilizer.

Fungal community diversity and microbial interaction play key roles in plant growth and metabolism. Proteobacteria can directly inhibit Firmicutes from entering into the endophytic community and consequently modify the microbial community (Chen et al., 2020). Endophytes have been isolated from *Coffea canephora*, and the high biodiversity of fungal endophytes in coffee plants may help us understand the plant growth process (Vega et al., 2010). Soybean rhizosphere may act as allelochemicals in the interactions between root and soil microbial community in a long-term monocropped soybean field (Guo et al., 2011). AMF and plant growth–promoting bacteria are beneficial to horticultural crops, which could increase yield and enhance crop quality (Emmanuel and Babalola, 2020). Terpenoids are a group of structurally diverse natural products that are widely used in the flavor and fragrance industry. Furthermore, it was clarified that the fungal sesquiterpene synthase's function differs between the phyla Ascomycota and Basidiomycota (Zhang et al., 2020).

It was evident that there is interaction between the indigenous microbial community and grain metabolism even with good-quality, mature malting barley. In the malting ecosystem, the fungal community markedly contributed to the production of microbial β-glucanases and xylanases and was also involved in proteolysis (Laitila et al., 2007). Elevated temperature also increases aphid abundance but decreases AMF colonization rates of the wheat grain, which implies that climate may affect crop quality by altering plant-associated fungal communities (Tian et al., 2019).

Drivers in Shaping the Structure of Foliar Fungal Community

Many factors could affect the foliar fungal community structure, including plant cultivar (Martins et al., 2011), soil physical and chemical characteristics, and climate (Kauserud et al., 2013). Our results showed that planting regions with different climates and plant cultivars are key factors. The PCA results indicated that both cultivar and planting region climate could affect the community structure of foliar fungi, whereas through PERMANOVA using distance matrices (Alekseyenko, 2016), we found that the impact of ecological environment factors ($R^2 = 0.32$, $p < 0.05$) was more significant than crop cultivar ($R^2 = 0.42$, $p < 0.05$; **Table 3**).

The phyllosphere represents one of the most abundant habitats for microbiota colonization (Chen et al., 2020), and the role played by interactions between phyllosphere microorganisms in modifying the fungal community composition cannot be neglected. Related metastudies have identified climate as an important driving factor in different aspects of fungal biogeography, including the global distribution of common fungi and the composition and diversity of fungal communities (Vetrovsky et al., 2019). Climatic variability might modify trait selection in fungi, including spore size and dispersal characteristics. Changes in the composition and characteristics of fungal communities will have an important impact on

TABLE 2 | Ecological processes of community assembly for foliar fungi of crop leaves.

Sample	Dispersal limitation %	Even dispersal %	Drift %
GC-HY101	0	0	100
GC-HY201	33	0	67
GC-HY109	33	67	0
WZS-HY101	33	0	67
WZS-HY201	67	0	33
WZS-HY109	33	0	67

TABLE 3 | The impact of ecological environment factors and crop cultivar on the structure of the foliar fungal community using permutational multivariate analysis of variance using distance matrices based on Bray–Curtis similarity index.

	df	Sum of sequence	R^2	*F*	*p*
Planting region	1	72.83	0.32	2.61	0.0166
Cultivar	2	93.00	0.42	1.67	0.2333

interaction with plant communities and ecosystem functions (Andrew et al., 2016). Climate change may affect ecosystem functioning due to the narrow climatic tolerances of key fungal taxa. Mycorrhizal fungi appear to have narrower climatic tolerances than pathogenic fungi (Vetrovsky et al., 2019). Our results showed that the differences between tobacco cultivars play an important role in phyllosphere fungal community structure and affect the microbial co-occurrence pattern in phyllosphere fungal communities. Similarly, cucumber cultivars inoculated with different AMF had differential responses in terms of growth and branch nutrient composition, which revealed that plant cultivar could affect the microbial community functional diversity (Ravnskov and Larsen, 2016). Moreover, AMF also can enhance ecosystem resilience and reduce the negative impact of increased precipitation on nutrient losses (Martinez-Garcia et al., 2017). Also, mycorrhizal fungi can promote or hinder the successful spread of plants away from harsh environments (Bennett and Classen, 2020).

Community Assembly of Foliar Fungal Community on Leave Surface

The ecological assembly process is vital for the construction of microbial communities (Sloan et al., 2006). Spatial turnover in the composition of biological communities includes (ecological) drift, selection, and dispersal. Quantitatively estimating the influences of selection, dispersal, and drift is fundamental to our understanding of ecological systems (Stegen et al., 2013). In our study, the community NRI of fungi on leaf surfaces in GC region was between 0.74 and 1.01, whereas in WZS, it was between 0.78 and 1.26; furthermore, the NTI and NRI significantly varied among cultivars and planting regions (**Figure 5**), which implicates that the stochastic process plays a key role in the local species diversity and spatial turnover. Other studies have previously elucidated the importance of drift in community assembly process of the legume root nodules, including the core rhizobial communities (genus *Mesorhizobium*) that were driven by dispersal limitation in concert with drift (81.1% of *nodA* communities, Ramoneda et al., 2020). During the degradation of straw, ecological drift was important across all stages of decomposition (Bao et al., 2020). The βNTI was linearly related to the sensory quality of crop leaves, indicating that the ecosystem services may depend, to some extent, on the assembly process of microbial communities. This may have resulted from stochastic processes in the assembly of foliar fungal communities. Although these results have advanced our understanding of the relationships between microbial community assembly and crop quality, future work is still required to further reveal the connections between foliar fungal community and plant molecular metabolic mechanism.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available in NCBI under accession number PRJNA778452.

AUTHOR CONTRIBUTIONS

ZY and JL conceived and designed the works. LX, LL, and XH conducted the experiment. LX, QZ, and HY performed the bioinformatic analyses. LX and XH performed statistical analyses. LX wrote the original draft manuscript. HY, LX, QZ, and TZ reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study received funding from the China Tobacco Sichuan Industrial Co., Ltd (ctx201904). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication. All authors declare no other competing interests.

ACKNOWLEDGMENTS

We thank to the Hunan International Scientific and Technological Cooperation Base of Environmental Microbiology and Application (China) for bioinformatic analyses. This work was supported by the Development and application of mellow and sweet tobacco material for Chinese cigar in Hainan (KJSB201907010009).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.783923/full#supplementary-material

REFERENCES

Alekseyenko, A. V. (2016). Multivariate Welch t-test on distances. *Bioinform.* 32, 3552–3558. doi: 10.1093/bioinformatics/btw524

Andrew, C., Heegaard, E., Halvorsen, R., Martinez-Peña, F., Egli, S., Kirk, P. M., et al. (2016). Climate impacts on fungal community and trait dynamics. *Fung. Ecol.* 22, 17–25. doi: 10.1016/j.funeco.2016.03.005

Bao, Y., Feng, Y., Stegen, J. C., Wu, M., Chen, R., Liu, W., et al. (2020). Straw chemistry links the assembly of bacterial communities to decomposition in paddy soils. *Soil Biol. Biochem.* 148:107866. doi: 10.1016/j.soilbio.2020.107866

Bennett, A. E., and Classen, A. T. (2020). Climate change influences mycorrhizal fungal–plant interactions, but conclusions are limited by geographical study bias. *Ecol.* 101:e02978. doi: 10.1002/ecy.2978

Borcard, D., and Legendre, P. (2012). Is the Mantel correlogram powerful enough to be useful in ecological analysis? A simulation study. *Ecology* 93, 1473–1481. doi: 10.1890/11-1737.1

Campbell, C., Yang, S., Albert, R., and Shea, K. (2011). A network model for plant-pollinator community assembly. *Proc. Natl. Acad. Sci. U S A* 108, 197–202. doi: 10.1073/pnas.1008204108

Cardona, C., Weisenhorn, P., Henry, C., and Gilbert, J. A. (2016). Network-based metabolic analysis and microbial community modeling. *Curr. Opin. Microbiol.* 31, 124–131. doi: 10.1016/j.mib.2016.03.008

Changey, F., Bagard, M., Souleymane, M., and Lerch, T. Z. (2018). Cascading effects of elevated ozone on wheat rhizosphere microbial communities depend on temperature and cultivar sensitivity. *Environ. Pollut.* 242(Pt A), 113–125. doi: 10.1016/j.envpol.2018.06.073

Chen, H., and Boutros, P. C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinform.* 12:35. doi: 10.1186/1471-2105-12-35

Chen, T., Nomura, K., Wang, X., Sohrabi, R., Xu, J., Yao, L., et al. (2020). A plant genetic network for preventing dysbiosis in the phyllosphere. *Nature* 580, 653–657. doi: 10.1038/s41586-020-2185-0

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *J. Vegit. Sci.* 14, 927–930. doi: 10.1111/j.1654-1103.2003.tb02228.x

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Med.* 10:996. doi: 10.1038/NMETH.2604

Emmanuel, O. C., and Babalola, O. O. (2020). Productivity and quality of horticultural crops through co-inoculation of arbuscular mycorrhizal fungi and plant growth promoting bacteria. *Microbiol. Res.* 239:126569. doi: 10.1016/j.micres.2020.126569

Eschen, R., Hunt, S., Mykura, C., Gange, A. C., and Sutton, B. C. (2010). The foliar endophytic fungal community composition in Cirsium arvense is affected by mycorrhizal colonization and soil nutrient content. *Fungal. Biol.* 114, 991–998. doi: 10.1016/j.funbio.2010.09.009

Faticov, M., Abdelfattah, A., Roslin, T., Vacher, C., Hamback, P., Blanchet, F. G., et al. (2021). Climate warming dominates over plant genotype in shaping the seasonal trajectory of foliar fungal communities on oak. *New Phytol.* 231, 1770–1783. doi: 10.1111/nph.17434

Fitzpatrick, C. R., Copeland, J., Wang, P. W., Guttman, D. S., Kotanen, P. M., and Johnson, M. T. J. (2018). Assembly and ecological function of the root microbiome across angiosperm plant species. *Proc. Natl. Acad. Sci. U S A* 115, E1157–E1165. doi: 10.1073/pnas.1717617115

Friedman, J., and Alm, E. J. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLoS Comp. Biol.* 8:9. doi: 10.1371/journal.pcbi.1002687

Gravel, D., Canham, C. D., Beaudet, M., and Messier, C. (2006). Reconciling niche and neutrality: the continuum hypothesis. *Ecol. Lett.* 9, 399–409. doi: 10.1111/j.1461-0248.2006.00884.x

Guo, Z.-Y., Kong, C.-H., Wang, J.-G., and Wang, Y.-F. (2011). Rhizosphere isoflavones (daidzein and genistein) levels and their relation to the microbial community structure of mono-cropped soybean soil in field and controlled conditions. *Soil Biol. Biochem.* 43, 2257–2264. doi: 10.1016/j.soilbio.2011.07.022

Hernandez-Garcia, A., Gonzalez-Gonzalez, I., Zarco, A. I. J., and Chaparro-Pelaez, J. (2016). Visualizations of Online Course Interactions for Social Network Learning Analytics. *Intern. J. Emerg. Tech. Learn.* 11, 6–15. doi: 10.3991/ijet.v11i07.5889

Ihaka, R., and Gentleman, R. (1996). R: a language for data analysis and graphics. *J. Comp. Graph. Stat.* 5, 299–314.

Jiao, S., Yang, Y., Xu, Y., Zhang, J., and Lu, Y. (2020). Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J.* 14, 202–216. doi: 10.1038/s41396-019-0522-9

Ju, F., and Zhang, T. (2015). 16S rRNA gene high-throughput sequencing data mining of microbial diversity and interactions. *Appl. Microbiol. Biotech.* 99, 4119–4129. doi: 10.1007/s00253-015-6536-y

Ju, W. Y., Li, J. X., Yu, W. R., and Zhang, R. C. (2016). iGraph: an incremental data processing system for dynamic graph. *Front. Comp. Sci.* 10:462–476. doi: 10.1007/s11704-016-5485-7

Kauserud, H., Heegaard, E., Buntgen, U., Halvorsen, R., Egli, S., Senn-Irlet, B., et al. (2013). Reply to Gange et al.: Climate-driven changes in the fungal fruiting season in the United Kingdom. *Proc. Natl. Acad. Sci. U S A* 110:E335. doi: 10.1073/pnas.1221131110

Kelly, J. J., Bansal, A., Winkelman, J., Janus, L. R., Hell, S., Wencel, M., et al. (2010). Alteration of microbial communities colonizing leaf litter in a temperate woodland stream by growth of trees under conditions of elevated atmospheric CO2. *Appl. Environ. Microbiol.* 76, 4950–4959. doi: 10.1128/AEM.00221-10

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., et al. (2010). Picante: R tools for integrating phylogenies and ecology. *Bioinform.* 26, 1463–1464. doi: 10.1093/bioinformatics/btq166

Kõljalg, U., Larsson, K.-H., Abarenkov, K., Nilsson, R. H., Alexander, I. J., Eberhardt, U., et al. (2005). UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytologist.* 166, 1063–1068. doi: 10.1111/j.1469-8137.2005.01376.x

Kong, Y. (2011). Btrim: A fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. *Genomics* 98, 152–153. doi: 10.1016/j.ygeno.2011.05.009

Konopka, A. (2009). What is microbial community ecology? *ISME J.* 3, 1223–1230. doi: 10.1038/ismej.2009.88

Laitila, A., Kotaviita, E., Peltola, P., Home, S., and Wilhelmson, A. (2007). Indigenous Microbial Community of Barley Greatly Influences Grain Germination and Malt Quality. *J. Inst. Brew.* 113, 9–20. doi: 10.1002/j.2050-0416.2007.tb00250.x

Latz, M. A. C., Kerrn, M. H., Sorensen, H., Collinge, D. B., Jensen, B., Brown, J. K. M., et al. (2021). Succession of the fungal endophytic microbiome of wheat is dependent on tissue-specific interactions between host genotype and environment. *Sci. Total Environ.* 759, 143804. doi: 10.1016/j.scitotenv.2020.143804

Lei, X., Ge, L., Jie, L., Pinhe, L., Xi, H., Heng, X., et al. (2021). Effects of ecological environment and genotype on the fungal Community of cigar tobacco leaf in the mature period (in Chinese). *Acta Tabac. Sin.* 27, 119–126. doi: 10.16472/j.chinatobacco.2020.T0124

Liu, C., Cui, Y. M., Li, X. Z., and Yao, M. J. (2021). microeco: an R package for data mining in microbial community ecology. *FEMS Microbiol. Ecol.* 97:2. doi: 10.1093/femsec/fiaa255

Liu, X., Jia, P., Cadotte, M. W., Zhu, C., Si, X., Wang, Y., et al. (2021). Host plant environmental filtering drives foliar fungal community assembly in symptomatic leaves. *Oecologia* 195, 737–749. doi: 10.1007/s00442-021-04849-3

Ma, S., De Frenne, P., Vanhellemont, M., Wasof, S., Boeckx, P., Brunet, J., et al. (2019). Local soil characteristics determine the microbial communities under forest understorey plants along a latitudinal gradient. *Basic Appl. Ecol.* 36, 34–44. doi: 10.1016/j.baae.2019.03.001

Mannaa, M., Han, G., Jeon, H. W., Kim, J., Kim, N., Park, A. R., et al. (2020). Influence of Resistance-Inducing Chemical Elicitors against Pine Wilt Disease on the Rhizosphere Microbiome. *Microorganisms* 8:6. doi: 10.3390/microorganisms8060884

Martinez-Garcia, L. B., De Deyn, G. B., Pugnaire, F. I., Kothamasi, D., and van der Heijden, M. G. A. (2017). Symbiotic soil fungi enhance ecosystem resilience to climate change. *Glob Chang Biol* 23, 5228–5236. doi: 10.1111/gcb.13785

Martins, J. L., Jideani, I. A., Yusuf, I. Z., and Tahir, F. (2011). Rhizosphere mycology of three maize varieties. *J. Food Agricul. Environ.* 9, 706–710.

Nuccio, E. E., Anderson-Furgeson, J., Estera, K. Y., Pett-Ridge, J., de Valpine, P., Brodie, E. L., et al. (2016). Climate and edaphic controllers influence rhizosphere community assembly for a wild annual grass. *Ecology* 97, 1307–1318. doi: 10.1890/15-0882.1

Otsing, E., Barantal, S., Anslan, S., Koricheva, J., and Tedersoo, L. (2018). Litter species richness and composition effects on fungal richness and community structure in decomposing foliar and root litter. *Soil Biol. Biochem.* 125, 328–339. doi: 10.1016/j.soilbio.2018.08.006

Preheim, S. P., Perrotta, A. R., Martin-Platero, A. M., Gupta, A., and Alm, E. J. (2013). Distribution-Based Clustering: Using Ecology To Refine the Operational Taxonomic Unit. *Appl. Env. Microbiol.* 79, 6593–6603. doi: 10.1128/AEM.00342-13

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2-Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* 5:3. doi: 10.1371/journal.pone.0009490

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

R CoreTeam (2013). *R: A language and environment for statistical computing*. Vienna: R coreTeam.

Ramoneda, J., Le Roux, J. J., Frossard, E., Frey, B., and Gamper, H. A. (2020). Experimental assembly reveals ecological drift as a major driver of root nodule bacterial diversity in a woody legume crop. *FEMS Microbiol. Ecol.* 96:6. doi: 10.1093/femsec/fiaa083

Ravnskov, S., and Larsen, J. (2016). Functional compatibility in cucumber mycorrhizas in terms of plant growth performance and foliar nutrient composition. *Plant Biol.* 18, 816–823. doi: 10.1111/plb.12465

Shi, S. J., Nuccio, E. E., Shi, Z. J., He, Z. L., Zhou, J. Z., and Firestone, M. K. (2016). The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. *Ecol. Lett.* 19, 926–936. doi: 10.1111/ele.12630

Shi, Y., Pan, Y., Xiang, L., Zhu, Z., Fu, W., Hao, G., et al. (2021). Assembly of rhizosphere microbial communities in Artemisia annua: recruitment of plant growth-promoting microorganisms and inter-kingdom interactions between bacteria and fungi. *Plant Soil* 2021:9. doi: 10.1007/s11104-021-04829-9

Sloan, W. T., Lunn, M., Woodcock, S., Head, I. M., Nee, S., and Curtis, T. P. (2006). Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Env. Microb.* 8, 732–740. doi: 10.1111/j.1462-2920.2005.00956.x

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J.* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Fredrickson, J. K., and Konopka, A. E. (2015). Estimating and mapping ecological processes influencing microbial community assembly. *Front. Microb.* 6:370. doi: 10.3389/fmicb.2015.00370

Stegen, J. C., Lin, X. J., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J.* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Sun, K., Fu, L., Song, Y., Yuan, L., Zhang, H., Wen, D., et al. (2021). Effects of continuous cucumber cropping on crop quality and soil fungal community. *Environ. Monit. Assess.* 193:436. doi: 10.1007/s10661-021-09136-5

Tian, B., Yu, Z., Pei, Y., Zhang, Z., Siemann, E., Wan, S., et al. (2019). Elevated temperature reduces wheat grain yield by increasing pests and decreasing soil mutualists. *Pest Manag. Sci.* 75, 466–475. doi: 10.1002/ps.5140

Vega, F. E., Simpkins, A., Aime, M. C., Posada, F., Peterson, S. W., Rehner, S. A., et al. (2010). Fungal endophyte diversity in coffee plants from Colombia, Hawai'i, Mexico and Puerto Rico. *Fung. Ecol.* 3, 122–138. doi: 10.1016/j.funeco.2009.07.002

Vetrovsky, T., Kohout, P., Kopecky, M., Machac, A., Man, M., Bahnmann, B. D., et al. (2019). A meta-analysis of global fungal distribution reveals climate-driven patterns. *Nat. Comm.* 10:8. doi: 10.1038/s41467-019-13164-8

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Env. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Watts, S. C., Ritchie, S. C., Inouye, M., and Holt, K. E. (2019). FastSpar: rapid and scalable correlation estimation for compositional data. *Bioinformatics* 35, 1064–1066. doi: 10.1093/bioinformatics/bty734

Zhang, C. Q., Chen, X. X., Orban, A., Shukal, S., Birk, F., Too, H. P., et al. (2020). Agrocybe aegerita Serves As a Gateway for Identifying Sesquiterpene Biosynthetic Enzymes in Higher Fungi. *ACS Chem. Biol.* 15, 1268–1277. doi: 10.1021/acschembio.0c00155

Zhao, S., Guo, Y., Sheng, Q., and Shyr, Y. (2014). Heatmap3: an improved heatmap package with more powerful and convenient features. *BMC Bioinform.* 15, 1–2. doi: 10.1155/2014/986048

Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, X. (2010). Functional molecular ecological networks. *mBio* 1:4. doi: 10.1128/mBio.00169-10

Zhou, J., Deng, Y., Shen, L., Wen, C., Yan, Q., Ning, D., et al. (2016). Temperature mediates continental-scale diversity of microbes in forest soils. *Nat. Commun.* 7:12083. doi: 10.1038/ncomms12083

Zhou, J. Z., and Ning, D. L. (2017). Stochastic Community Assembly: does It Matter in Microbial Ecology? *Microbiol. Molecul. Biol. Rev.* 81:4. doi: 10.1128/MMBR.00002-17

Conflict of Interest: LX, XH, LL, HX, and TZ were employed by the China Tobacco Sichuan Industrial Co., Ltd.

The remaining 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.

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 23 May 2022
doi: 10.3389/fmicb.2022.783371

Rhizosphere Soil Microbial Community Under Ice in a High-Latitude Wetland: Different Community Assembly Processes Shape Patterns of Rare and Abundant Microbes

Jiaming Ma, Kang Ma, Jingling Liu* and Nannan Chen*

State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing, China

The rhizosphere soil microbial community under ice exhibits higher diversity and community turnover in the ice-covered stage. The mechanisms by which community assembly processes shape those patterns are poorly understood in high-latitude wetlands. Based on the 16S rRNA gene and ITS sequencing data, we determined the diversity patterns for the rhizosphere microbial community of two plant species in a seasonally ice-covered wetland, during the ice-covered and ice-free stages. The ecological processes of the community assembly were inferred using the null model at the phylogenetic bins (taxonomic groups divided according to phylogenetic relationships) level. Different effects of ecological processes on rare and abundant microbial sub-communities (defined by the relative abundance of bins) and bins were further analyzed. We found that bacterial and fungal communities had higher alpha and gamma diversity under the ice. During the ice-free stage, the dissimilarity of fungal communities decreased sharply, and the spatial variation disappeared. For the bacterial community, homogeneous selection, dispersal limitation, and ecological processes (undominated processes) were the main processes, and they remained relatively stable across all stages. For the fungal community, during the ice-covered stage, dispersal limitation was the dominant process. In contrast, during the ice-free stage, ecological drift processes were more important in the *Scirpus* rhizosphere, and ecological drift and homogeneous selection processes were more important in the *Phragmites* rhizosphere. Regarding the different effects of community assembly processes on abundant and rare microbes, abundant microbes were controlled more by homogeneous selection. In contrast, rare microbes were controlled more by ecological drift, dispersal limitation, and heterogeneous selection, especially bacteria. This is potentially caused by the low growth rates or the intermediate niche breadths of rare microbes under the ice. Our findings suggest the high diversity of microbial communities under the ice, which deepens our understanding of various ecological processes of community assembly across stages and reveals the distinct effects of community assembly processes on abundant and rare microbes at the bin level.

Keywords: bacterial and fungal communities, seasonally ice-covered, community assembly, temporal dynamic, rare taxa, iCAMP package

OPEN ACCESS

Edited by:
Stilianos Fodelianakis,
Swiss Federal Institute of Technology Lausanne, Switzerland

Reviewed by:
Guozhuang Zhang,
China Academy of Chinese Medical Sciences, China
Pengfa Li,
Nanjing Agricultural University, China
Xuanyu Tao,
University of Oklahoma, United States

****Correspondence:***
Jiaming Ma
201921180056@mail.bnu.edu.cn
Jingling Liu
jingling@bnu.edu.cn

Specialty section:
This article was submitted to Terrestrial Microbiology, a section of the journal Frontiers in Microbiology

Received: *26 September 2021*
Accepted: *23 March 2022*
Published: *23 May 2022*

Citation:
Ma J, Ma K, Liu J and Chen N (2022) Rhizosphere Soil Microbial Community Under Ice in a High-Latitude Wetland: Different Community Assembly Processes Shape Patterns of Rare and Abundant Microbes. Front. Microbiol. 13:783371. doi: 10.3389/fmicb.2022.783371

INTRODUCTION

It is increasingly recognized that microorganisms can exist and grow under the ice (Tran et al., 2018). For microbial communities in seasonally ice-covered habitats, the ice-covered stage is still a poorly studied period compared to the ice-free stage (Jansen et al., 2021). Some studies have found that bacterial diversity increases at sub-zero temperatures in controlled experiments (Juan et al., 2018), while other studies found that soil bacterial diversity decreased (Zhang et al., 2017) or increased at first and then decreased during freeze-thaw in the forest (Sang et al., 2021). In wetlands located in high-latitude regions where the soil is seasonally covered by ice or water rather than snow or air, the plant inputs can change dramatically, potentially differentiating wetland bacterial diversity from other habitats. The environment under the ice may have two opposing effects on the alpha diversity of bacteria. Although the more selective environment under the ice (low temperature, low light) may decrease the alpha diversity (Butler et al., 2019), the bacteria are believed to be mainly dormant under the ice, and dormancy could increase alpha diversity, thereby preventing bacterial species from going extinct (Butler et al., 2019). Thus, it remains uncertain whether bacterial communities have higher alpha diversity under the ice in wetlands. Fungi are another important component of the microbial community. Though some studies have shown that the fungal: bacterial biomass ratios may increase in the low-temperature environment (Robroek et al., 2013), studies on fungal diversity under the ice are even more sparse. Therefore, more evidence on the diversity patterns of microbial communities, especially the fungal community, across ice-covered seasons in wetlands needs to be investigated.

The diversity and biogeography patterns of microbial communities are shaped by many processes. Niche-based theory suggests that non-random and niche-based ecological processes govern diversity patterns, and these are known as deterministic processes (Chesson, 2000). On the other hand, neutral theory suggests that ecological processes, where all species have equivalent fitness, govern the diversity, and these processes are termed stochastic processes (Chave, 2004). It was gradually recognized that the deterministic and stochastic processes jointly govern the diversity patterns (Chase and Myers, 2011). To integrate these two processes that are based on different theories into a coherent framework, the community assembly processes can be divided into four fundamental ecological processes: selection (including environmental filtering, biotic interactions, and host filtering), dispersal, ecological drift (random birth and death), and diversification (speciation and extinction) (Vellend, 2016; Zhou and Ning, 2017; Ning et al., 2020).

The influence of environmental filtering (one of the selection factors) on bacterial communities under the ice has been studied, specifically relating to dissolved oxygen (Bertilsson et al., 2013), low light, and temperature (Cruaud et al., 2020). Host filtering of plants is also an important factor in selection. Microbiomes in the rhizosphere are considered highly affected by rhizo-deposits and are in higher abundance and activity than in bulk soil (Prashar et al., 2014). The species of plants play an important role in affecting rhizosphere microbial community assembly processes (Philippot et al., 2013; Fitzpatrick et al., 2018; Zhalnina et al., 2018; Matthews et al., 2019). The different development stages of plants could also influence the composition and the community assembly processes of the rhizosphere microbial community (Chaparro et al., 2014; Bell et al., 2015).

Most studies related to the rhizosphere microbial community have concentrated on the various stages of plant growth (Bell et al., 2015), including seedling, growth, and flowering (Chaparro et al., 2014), but similar studies during the non-growing season (e.g., the ice-covered stage) are limited. Some species of plants will stop growing and partially or entirely wither during the ice-covered season. A decrease in photosynthesis may dampen the secretory activity of the roots, thereby affecting the community assembly processes of the rhizosphere microbial community (He et al., 2020). Additionally, other ecological processes, such as dispersal and ecological drift of microbial communities, under the ice have also been less well-studied. To further understand how different ecological processes work together, it is necessary to quantify their relative importance. The null model analysis is a useful tool that uses randomization procedures to quantify these ecological processes and may provide an understanding of the relationship between changes in the environment and the microbial community (Hanson et al., 2012; Stegen et al., 2012, 2013; Zhou and Ning, 2017). This approach can be used to study and quantify the community assembly processes of the soil microbial communities that live in different plant rhizospheres under the ice.

The relative abundance of microbial taxa may play an important role in regulating different community assembly processes, for example, low abundance may enhance the relative importance of ecological drift (Nemergut et al., 2013). In natural ecosystems, the vast majority of microbial taxa have a low relative abundance of rare microbial taxa, which are driven by different processes than the abundant microbial taxa (Magurran and Henderson, 2003; Martinez et al., 2015; Jousset et al., 2017). The drivers of microbial community rarity include dispersal (Lee et al., 2021), narrow niche breadths (Jiao and Lu, 2020), low growth rates (Liao et al., 2017; Lee et al., 2021), and biotic interactions (Jousset et al., 2017). During the ice-covered season, those drivers change significantly, including the dispersal medium, growth ratio, and nutrient inputs, leading to changes in species rarity. Recent studies researched the relationship between community assembly processes and the relative abundance of microbial taxa by dividing the microbial community into abundant, intermediate, and rare microbial sub-communities according to an artificially selected threshold of relative abundance (Liao et al., 2017; Jiao and Lu, 2020; Lee et al., 2021; Wan et al., 2021; Zheng et al., 2021). However, it is important to note that the levels of division (sub-community) and the threshold of division (artificially selected) may influence the study results. A newly published phylogenetic bin-based null model (the "Icamp" model) provides an approach to solve this issue (Ning et al., 2020). It calculates ecological processes of community assembly at the taxonomic group level (hereafter, the term "bins" is used to represent these taxa groups) (Ning et al., 2020). Many recent studies use this framework (Ceja-Navarro et al., 2021; Dong et al., 2021; Le Roux et al., 2021; Sun C. et al.,

2021; Sun Y. et al., 2021). The influence of ecological processes on microbes with different levels of abundance (abundant, intermediate, and rare) could be further analyzed at the level of bins using this framework to explore the underlying mechanisms.

This research aimed to determine (1) the diversity patterns of microbial communities across ice-covered and ice-free stages; (2) the relative importance and the potential drivers of different community assembly processes across ice-covered and ice-free stages; (3) the effects of community assembly processes on different relative abundance microbes (abundant, intermediate, and rare microbes). The research was carried out in Momoge wetland, a seasonal ice-covered wetland located in Northeast China. Rhizosphere soil samples were collected throughout the ice-covered and ice-free stages. Microbial communities were characterized using the amplicon sequencing of the 16S rRNA gene (indicating bacterial communities) and ITS (indicating fungal communities). Diversity was calculated across different ice-covered stages, and the relative importance of different ecological processes was calculated using the "iCAMP" framework. The relationship between relative abundance and different ecological processes was also analyzed. We hypothesized that (1) the diversity of bacterial and fungal communities would be higher under the ice, (2) the selection processes would be influenced by the species of plants, and the dispersal limitation processes would dominate microbial communities during ice-covered stages, and (3) the rare microbial taxa would be more controlled by ecological drift than selection.

MATERIALS AND METHODS

Study Sites and Sampling

Our study was conducted at the Momoge National Nature Reserve (45°42′25″ to 46°18′ 0″ N, 123°27′0″ to 124°4′33.7″ E), located in northeastern China. As an important hydrology node in the Nenjiang River basin (Meng, 2020), the local average annual temperature is 4.4°C, and in January, it goes down to −17.4°C. For almost half of the year, water and soil in the Momoge wetlands are entirely frozen (mid-November to mid-March) or partially frozen (late October to mid-November and mid-March to early April) (Zheng et al., 2019). However, in recent years, the Momoge wetlands were disturbed by recession flows from farmlands (Meng et al., 2019) and exhibited an abnormally high water surface during autumn and winter. This may influence the microbial diversity pattern and the community assembly processes related to the rhizosphere soil microbial community, thus obstructing bio-geographical cycling processes and plant activities during the ice-cover period (Sun et al., 2020).

We chose three sites connected by surface water but over three kilometers apart as our sample sites (30 × 30 m) (**Supplementary Figure 1**, see details in **Supplementary Table 1**). Within those three sample sites, we designed sample transects at three water depths (0, 15, and 25 cm). Along these three sample transects, we chose two emergent aquatic plant species (*Phragmites australis* (Cav.) Trin. ex Steud. and *Scirpus mucronatus* Linn.) as sample plots (0.5 × 0.5 m). In each sample plot, we sampled plant rhizosphere soil (root around 1 cm); three similar cores of rhizosphere soils were collected and mixed as one sample. Each soil sample was divided into four parts: one was used for DNA extraction and sequencing (∼5 g), one for soil moisture measurements, one was freeze-dried for subsequent soil chemistry analysis, and one was stored at 4°C as a standby soil. Rhizosphere soil for DNA extraction and sequencing was saved at −20°C during transportation and stored at −80°C.

To reflect the dynamic change of the microbial communities, we set up three sample stages in the Momoge wetlands during the winter (See details in **Supplementary Table 2**). According to the duration of the freeze, we divided sample stages into the entirely ice-covered stage (December 2020), partially ice-covered stage (October 2020 and March 2021), and the ice-free stage (May 2021). Each stage was defined as follows: during the entirely ice-covered stage, the highest local air temperature is lower than 0°C, and water and soil are frozen throughout the whole day; during the partially ice-covered stage, the highest local air temperature is higher than 0°C while the lowest local air temperature is lower than 0°C, and the water and soil are frozen during the day; during the ice-free stage, the lowest local air temperature is higher than 0°C, and the water and soil are not frozen at any time (**Supplementary Figure 2**). During each sampling period, we collected 17 mixed rhizosphere soil samples, as described above (one mixed rhizosphere soil sample is missing from the expected 18, as we did not find any *Scirpus mucronatus* Linn. in Site 2 at 25 cm water depth), amounting to 68 mixed rhizosphere soil samples for all four stages.

DNA Extraction, Sequencing, and OTU Clustering

Total genome DNA from samples was extracted using the CTAB/SDS method. The V4 region of 16S rRNA genes was amplified using a specific primer (515F and 806R) for bacteria, and the ITS2 region was amplified using a specific primer (ITS3-2024F and ITS4-2409R) for fungi. All polymerase chain reactions (PCR) contained 15 μL of Phusion® High-Fidelity PCR Master Mix (New England Biolabs), 0.2 μM of forward and reverse primers, and ∼10 ng of template DNA. PCR conditions were 98°C for 1 min, followed by 30 cycles of denaturation at 98°C for 10 s, 50°C for 30 s, and 72°C for 30 s, and a final extension at 72°C for 5 min. The PCR products were mixed with the same volume of IX loading buffer (contained SYB green). The mixed PCR products underwent gel electrophoresis on 2% agarose gel for detection and were purified with a Qiagen Gel Extraction Kit (Qiagen, Germany). The library quality was assessed on the Qubit@2.0 Fluorometer (Thermo Scientific) and Agilent Bioanalyzer 2100 system. Finally, the library was sequenced on an Illumina NovaSeq PE250 platform, and 250 bp paired-end reads were generated. Paired-end reads were assigned to samples based on their unique barcode and truncated by cutting off the barcode and primer sequence. DNA extraction and sequencing were then performed at the Tianjin Sequencing Center and Clinical Lab (Beijing Novogene Technology Co., Ltd, Tianjin).

Paired-end reads were merged using FLASH (Magoč and Salzberg, 2011) to generate raw tags. Quality filtering of raw tags was performed according to Bokulich (Bokulich et al., 2013)

using the QIIME (V1.9.1) (Caporaso et al., 2010). To detect chimera sequences, the 16S RNA gene tags were compared with the Silva database (version 138.1) (Quast et al., 2013; Yilmaz et al., 2014) using the UCHIME algorithm (Edgar et al., 2011); the ITS tags were compared with the UNITE database (version 8.2) (Nilsson et al., 2019) using VSEARCH (version 1.3.0) (Rognes et al., 2016). The effective tags were obtained after removing the chimera sequences detected above (Haas et al., 2011). Based on these tags, operational taxonomic units (OTUs) were clustered using the Uparse algorithm (Edgar, 2013) (Uparse version 7.0.1001) at a 97% identity threshold. Each bacterial OTU was classified against the Silva database (version 138.1) using the Mothur algorithm (Schloss, 2009). Each fungal OTU was classified against the UNITE database (version 8.2) (Abarenkov et al., 2010) using the Blast algorithm (Altschul et al., 1990). To ensure comparability, each sample was homogenized to equal sequencing depth (see details in **Supplementary Table 3**).

Construction of Phylogenetic Trees

For the bacterial communities, the phylogenetic trees based on the 16S rRNA gene sequences were constructed in the FastTree software (Version 2.1.11) (Price et al., 2009) using the "maximum likelihood" method and in the Figtree software (Version 1.4.4) by setting the root using the "midpoint" method. For the fungal communities, the phylogenetic trees based on ITS sequences were constructed in the Ghost-tree software (Version 0.0.1 dev.) (Fouquier et al., 2016). This is because the ITS marker region has higher sequence variability, which not only makes ITS markers suitable for a more accurate taxonomic identification at the genus or species level but also makes the multiple sequence alignment in a long phylogenetic distance of ITS sequences highly unreliable (Fouquier et al., 2016; Tedersoo et al., 2018; Ning et al., 2020). To ensure the phylogenetic trees are robust enough for analysis based on phylogenetic distance, we applied the Ghost-tree method rather than other methods to construct the phylogenetic trees for fungi. As a construction method for hybrid-gene phylogenetic trees, the Ghost-tree constructed foundation trees according to aligned databases of the fungal 18S sequence and then grafted extension trees according to the aligned databases of the fungal ITS sequences (which have a more accurate taxonomic identification at the genus and species levels) (Fouquier et al., 2016) (**Supplementary Figure 3**). We chose the order level of foundation trees (equivalent to the family level of extension trees) to graft the extension trees because ITS markers are usually applied at the family or genus levels (Tedersoo et al., 2018). The extension trees at the family level contain more OTUs than the extension trees at the genus level in our Ghost-trees (**Supplementary Figure 4**, see details in **Supplementary Table 4**).

Statistical Analysis

The alpha diversity of each sample was indicated by the observed OTUs, and the gamma diversity of a region was estimated by the Chao algorithm (Chao, 1987) using the "adiv" package (version 2.0.1) (Pavoine, 2020). Wilcoxon signed-rank tests were used to assess whether alpha diversity was significantly different between different stages and plant rhizospheres.

The community dissimilarity was calculated using the Bray-Curtis distance in the "vegan" package (version 2.5-7) (Oksanen et al., 2019). To assess the influence of distance, samples were divided into two groups: the community dissimilarity of samples within the sites and the community dissimilarity of samples among the sites. The community dissimilarity of samples within the site indicated dissimilarity over short distances (<30 m) and was referred to as "intra-site"; the community dissimilarity of samples among sites indicated dissimilarity over long distances (>3,000 m) and was referred to as "inter-sites." The principal coordinates analysis (PCoA) based on Bray-Curtis distances was also performed on samples between two plant species. Adonis tests were undertaken to assess the difference in community dissimilarity between these two groups and plant rhizospheres. All statistical tests described were performed in the R environment (Version 4.1.1) (R Core Team, 2017).

Analysis of Community Assembly Processes by Null Model

To determine the relative importance of different ecological processes to community assembly, we employed the "iCAMP" package. In contrast to other frameworks based on the null model, the "iCAMP" package can calculate ecological processes based on individual taxonomic groups ("bins") rather than the entire community (Ning et al., 2020). The taxonomic groups were divided based on phylogenetic relationships between OTUs, usually containing 12∼48 OTUs, which could improve the accuracy of the results of community assembly processes that are necessary for subsequent analysis at the bin level (Ning et al., 2020). This framework uses the absolute abundance OTU table and rooted phylogenetic tree to calculate the relative importance of five community assembly processes: homogeneous selection, homogeneous dispersal, dispersal limitation, heterogeneous selection (variable selection), and undominated (including ecological drift, diversification, weak selection, and/or weak dispersal; hereafter, the term ecological drift is used to represent these processes) processes. The phylogenetic signal was based on pH to determine the parameters of the phylogenetic bins: bin size (24 for bacteria; 18 for fungi) and threshold of phylogenetic distance (0.05 for bacteria; 0.025 for fungi) (**Supplementary Figure 5**). Then, the beta net relatedness index (βNRI) was calculated based on 1,000-times randomization of the taxa across the tips of the phylogenetic tree. The Raup–Crick metric (RC) was similarly calculated, and both used the described threshold to divide community assembly processes (Ning et al., 2020). Finally, the relative importance of ecological processes to community assembly at the subcommunity and bin levels was ascertained.

Relationship Analysis of Microbial Rarity and Community Assembly Processes

To explore the relationship between microbial rarity and community assembly processes, bacterial and fungal communities were divided into rare, intermediate, and abundant sub-communities according to the relative abundance of phylogenetic bins. In previous studies, the sub-communities

were divided according to the relative of OTUs (Wan et al., 2021). Given that there was a lower number of bins compared to OTUs (bacteria: each bin contained 24 OTUs; fungi: each bin contained 18 OTUs) and lower numbers of fungal phylogenetic bins (from 24 to 66) compared to bacterial phylogenetic bins (from 142 to 194), the sub-communities were divided based on their abundance. The bacterial communities were categorized as "rare bacterial sub-communities" when including bins with relative abundances <0.5% of the total bins; as "abundant bacterial sub-communities" when including bins with relative abundances >1% of the total bins; and as "intermediate bacterial sub-communities" for the remaining bins. To ensure that the number of bins in each sub-community was >3, the fungal communities were categorized as "rare fungal sub-communities" when including bins with relative abundances <1% of the total bins; as "abundant fungal sub-communities" when including bins with relative abundances >5% of the total bins; and as "intermediate fungal sub-communities" for the remaining bins. We calculated the ratio of the ecological processes of the three microbial sub-communities to total ecological processes. The relative importance of each of the five ecological processes in these sub-communities, i.e., homogeneous selection, homogeneous dispersal, ecological drift, dispersal limitation, and heterogeneous selection (variable selection) processes, was determined using 'bin_contribution_to_each_process' ($BP_{\tau k}$) data. To calculate community assembly processes at the bin level, the relationship between the relative importance of different ecological processes in governing each bin ($P_{\tau k}$) and the relative abundance of each bin were analyzed by Spearman's rank correlation. The relative abundance of all bins (explanatory variables), the value of which was equal to zero, was deleted.

RESULTS

Alpha, Gamma, and Beta Diversity Analysis of Microbial Community

The alpha diversity of the bacterial communities during the ice-covered stage was greater than that during the ice-free stage (22.4–36.7%, $p < 0.05$) (**Figure 1**). Similarly, the alpha diversity of the fungal communities during the ice-covered stage was higher than that during the ice-free stage (45.9–341.5%, $p < 0.05$), and was more significant than the diversity pattern of the bacteria communities. The gamma diversity of the bacterial community during the ice-covered stage was also greater than that during the ice-free stage (31.1–35.9%). Similarly, the gamma diversity of the fungal community during the ice-covered stage was more than that during the ice-free stage (82.4–100.4%). In the other plant rhizosphere habitats, most of the alpha and gamma diversity of microbial communities did not show a significant difference during the same stage and showed a similar change between different stages. Only in the first partially ice-covered stage

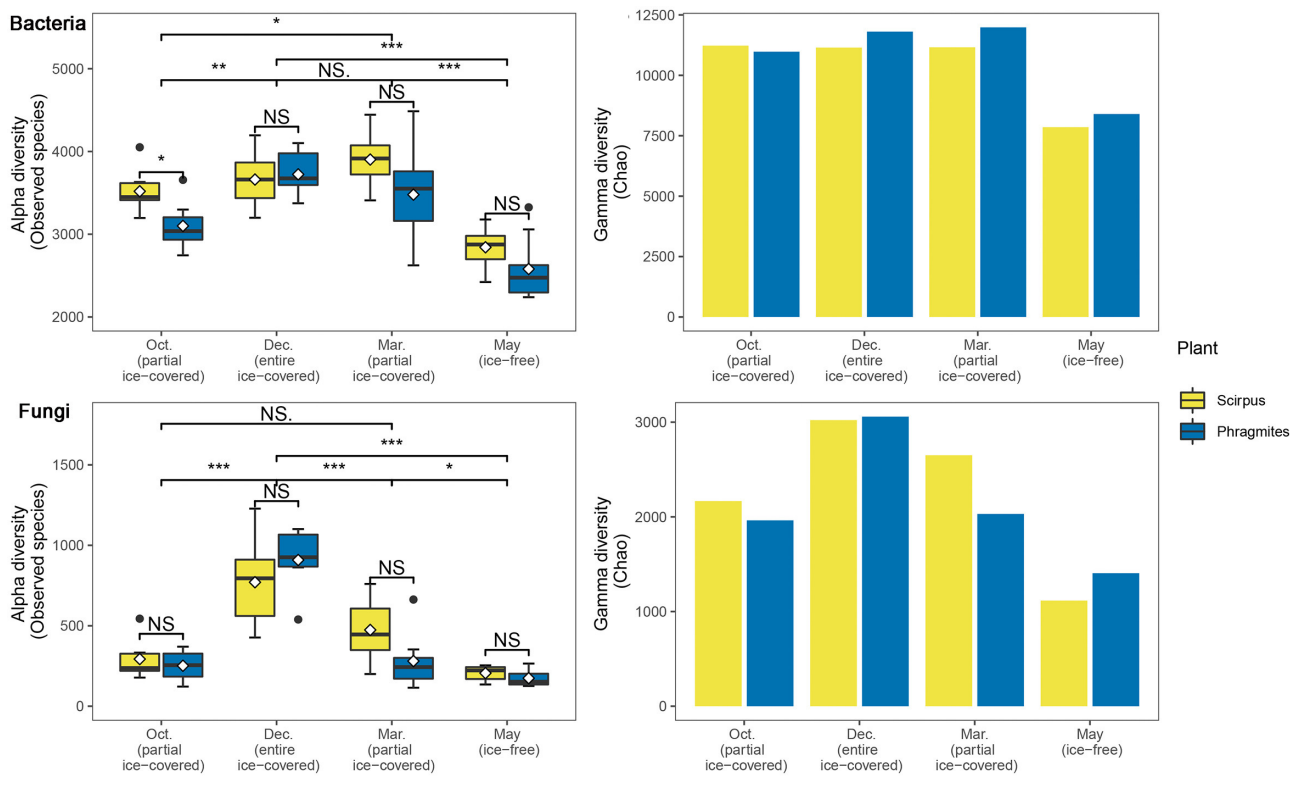


FIGURE 1 | Alpha and gamma diversities of bacterial and fungal communities during different ice-covered stages. Significance of the Wilcoxon signed-rank test: NS., not-significant; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

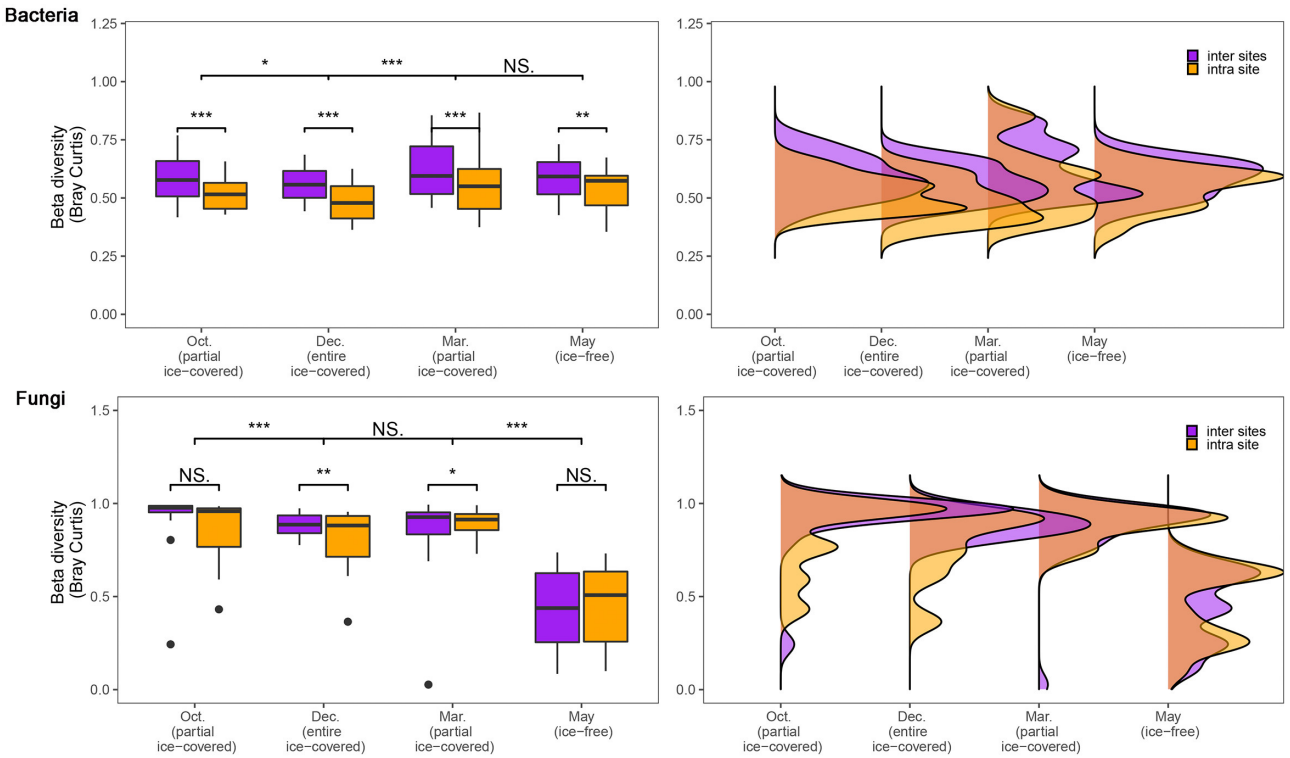


FIGURE 2 | Dissimilarity of bacterial and fungal communities during different ice-covered stages. Community dissimilarity was calculated using the Bray-Curtis distance. To assess the influence of geographical distances, the samples were divided into "inter sites" and "intra site" groups. The "inter site" group includes samples among different sites (>3,000 m); the "intra site" group includes samples within the same site (range of 30 × 30 m). Significance of the Adonis tests: NS., not-significant; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

(October 2020) did the alpha diversity of bacterial communities differ significantly between the two rhizosphere habitats.

To assess the influence of geographical distance on the microbial communities, the community dissimilarity (calculated by Bray-Curtis distance) between intra-sites and inter-sites was calculated and tested using the Wilcoxon signed-rank test (**Figure 2**). The dissimilarity of bacterial communities was significantly lower in intra-sites compared to inter-sites ($p < 0.05$) during all stages, with no exceptions. However, the dissimilarity of the fungal communities showed different patterns during the ice-free stage and other stages. During entirely and partially ice-covered stages, the dissimilarity of the fungal communities was significantly lower in intra-sites than in inter-sites ($p < 0.05$). During the ice-free stage, the dissimilarity of the fungal communities showed no significant difference either in intra-sites or in inter-sites. The dissimilarity of the fungal communities during the ice-free stage was significantly lower than at any other stage ($p < 0.05$).

Microbial Community Assembly Processes During Different Ice-Covered Stages

To reveal the community assembly mechanism during different ice-covered stages, ecological processes and their relative importance in the microbial communities were calculated (**Figure 3**). The results calculated by the "iCAMP" package showed that the main ecological processes of bacterial community assembly were homogeneous selection, dispersal limitation, and ecological drift. The main ecological processes of fungal community assembly were dispersal limitation, ecological drift, and homogeneous selection.

The turnover of different ecological processes during various ice-covered stages for the bacterial communities was minor. In contrast, for the fungal communities, there was a distinct shift between different stages. From entirely and partially ice-covered stages to the ice-free stage, the main ecological processes of the fungal community assembly changed from dispersal limitation (from 62.5–68.3 to 12.1 %) to homogeneous selection (from 10.8–18.6 to 28.5%) and ecological drift (from 14.0–20.5 to 58.3%). This means that during entirely and partially ice-covered stages, the fungal communities tended to be heterogeneous between different samples (dispersal limitation) and that during the ice-free stage, they tended to be homogeneous (homogeneous selection). Therefore, the community assembly processes of fungi were more influenced by the ice-covered environment or seasons than the community assembly processes of bacteria.

The turnover of different ecological processes was also different in the different plant rhizospheres. In the rhizosphere habitat of *Phragmites* (*Phragmites australis* (Cav.) Trin. ex Steud.), the proportion of homogeneous selection processes increased in both bacterial and fungal communities during the

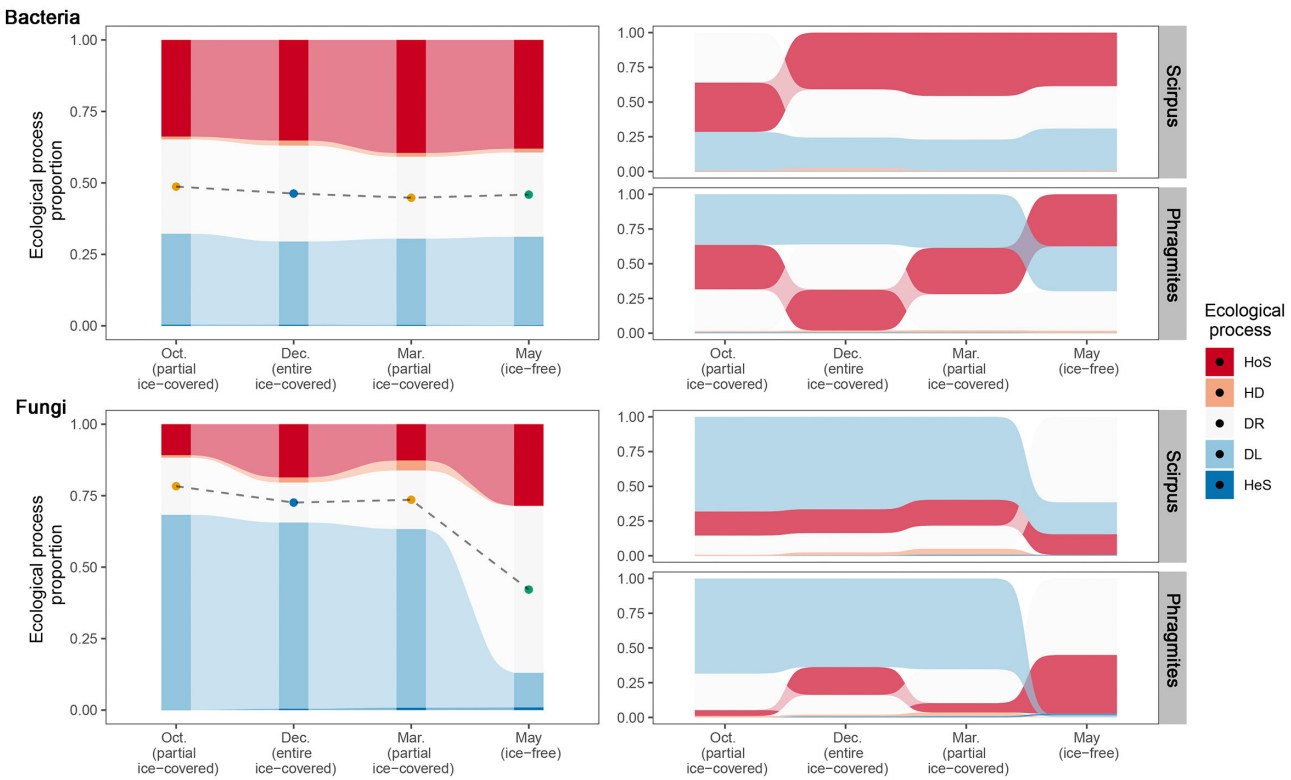


FIGURE 3 | Relative importance of different ecological processes to bacterial and fungal communities across different ice-covered stages and plant species. Different fill colors indicate different community assembly processes: HoS, homogeneous selection; HD, homogeneous dispersal; DR, ecological drift (i.e., undominated processes, including drift, diversification, weak selection, and/or weak dispersal); DL, dispersal limitation; HeS, heterogeneous selection. In the two left sub-figures, points depict the midpoint of the ecological drift process and reflect the deviation compared to the null model. In the two sub-figures on the right, processes were ranked according to the relative importance (proportion) from top to bottom to reflect the turnover of dominant processes across different ice-covered stages.

ice-free stage. In contrast, in the rhizosphere habitat of *Scirpus* (*Scirpus mucronatus Linn.*), the proportion of different ecological processes in the bacterial communities during various ice-covered stages was stable. However, the ecological drift process increased to the highest magnitude in the fungal communities during the ice-free stage.

Relationship Between Ecological Processes and Abundance of Microbial Subcommunities (Rare, Intermediate or Abundant) in Various Ice-Covered Stages

To explore the relationship between microbe rarity and the ecological processes of community assembly, the impact of the ecological processes on different microbial subcommunities (including rare, intermediate, and abundant microbial subcommunities) was calculated (**Figure 4**). During all stages in the bacterial communities, rare microbial subcommunities were more controlled by dispersal limitation than homogenous selection; however, the opposite was true for abundant subcommunities. In contrast, no such pattern was apparent for the fungal communities between rarity and ecological processes. During entirely and partially ice-covered stages, dispersal limitation was the dominant process for rare, intermediate, and abundant fungal microbial subcommunities. On the other hand, during the ice-free stage, abundant fungal subcommunities tended to be influenced by homogenous selection and ecological drift, while fungal intermediate and rare subcommunities tended to be influenced by dispersal limitation and ecological drift.

Relationship Between Ecological Processes and Abundance of Microbial Bins in Various Ice-Covered Stages

To verify the pattern of relative abundance and the ecological processes of microbial bins, Spearman's rank correlation was employed to test the relationship between the relative abundance of microbial bins and the relative importance of different ecological processes in governing each bin (**Figure 5**).

Interestingly, this result shows that ecological processes affect microbial bins of varying relative abundance in a different way for both bacterial and fungal communities. The relative importance of most processes for each bin was significantly correlated ($p < 0.05$) to the abundance of each bin. This means that if a microorganism belongs to an abundant microbial subcommunity (or a rare microbial subcommunity), it tends to be influenced by certain community assembly processes rather than others. For bacterial communities, the homogeneous selection tended to

FIGURE 4 | Relative importance of different ecological processes to abundant, intermediate, and rare bacterial and fungal subcommunity assemblies during different ice-covered stages. Different fill colors indicate different community assembly processes: HoS, Homogeneous Selection; HD, Homogeneous Dispersal; DR, ecological drift (i.e., undominated processes, including drift, diversification, weak selection, and/or weak dispersal); DL, Dispersal Limitation; HeS, Heterogeneous Selection.

exert control more on a relatively high abundance microbe rather than on a low abundance microbe (i.e., rare microbe). Moreover, the homogeneous selection, ecological drift, dispersal limitation, and heterogeneous selection processes tended to have more of an influence on low abundance than on high abundance bacteria. Compared to bacteria, the relationship between the fungal community assembly processes and the relative abundance had a similar direction but lower strength. Moreover, there was a significant correlation between the relative importance of ecological drift and heterogeneous selection processes for each OTU.

DISCUSSION

Higher Microbial Community Diversity Under the Ice

The diversity pattern of microbial communities under the ice, especially in fungal communities, is still unclear. In this study, the bacterial communities under the ice were characterized by a slightly higher alpha and gamma diversity (**Figure 1**). Furthermore, the alpha and gamma diversity for fungal communities under the ice exhibited patterns similar to the bacterial communities, and the pattern was even more significant than the bacterial diversity pattern (**Figure 1**). Previous studies in north wetlands suggested that soil contains more nitrogen and phosphorus during the ice-covered stage, which may be caused by the lower nutrient requirements and higher litter inputs of plants during the non-growth season (Gao, 2021). This could provide numerous niches and increase the diversity of microorganisms. Moreover, compared to bacteria, fungi are considered to have a higher resistance to low temperatures. A controlled experiment using the phospholipid fatty acid analysis reveals that fungi grow faster at sub-zero temperatures just as bacteria grow faster at above-zero temperatures (Haei et al., 2011). Other studies also support that fungi have higher diversity in frozen soil without any cover (Chen et al., 2021; Sang et al., 2021). Therefore, the higher resistance of fungi to low-temperature environments could improve the diversity of fungi under the ice in this study.

Different Temporal Dynamics of Bacterial and Fungal Community Assembly Processes

Community assembly processes shape the diversity and biogeography of microorganisms. Quantifying the relative importance of different processes could help us identify the main drivers of microbial communities across ice-covered and ice-free wetlands.

In this study, we used the "iCAMP" model to infer the relative importance of bacterial and fungal communities across ice-covered and ice-free stages. In this study, bacterial communities were mainly influenced by homogeneous selection during both

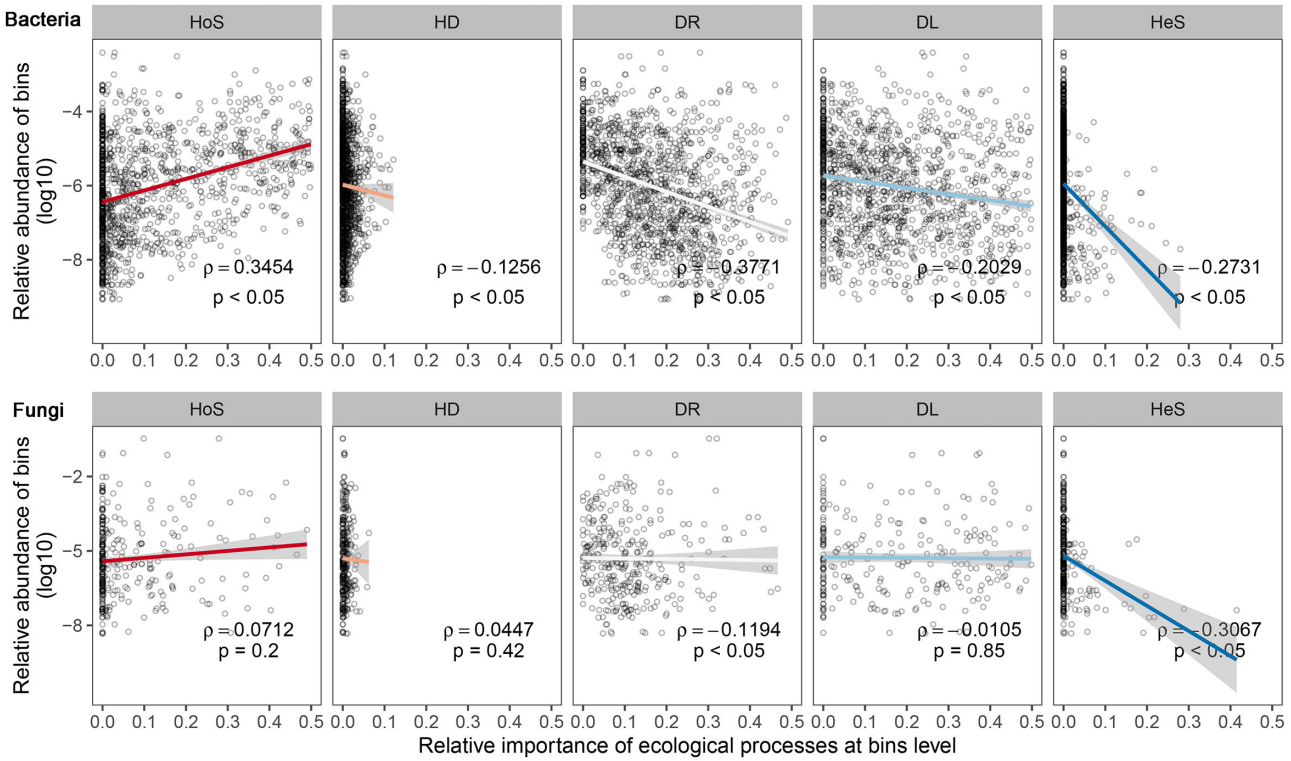


FIGURE 5 | Relationship of relative abundance of bacterial and fungal community taxa bins and their ecological processes of assembly at the bin level. Different sub-captions indicate different community assembly processes: HoS, homogeneous selection; HD, homogeneous dispersal; DR, ecological drift (i.e., undominated processes, including drift, diversification, weak selection, and/or weak dispersal); DL, dispersal limitation; HeS, heterogeneous selection. The horizontal axis depicts assembly processes' relative importance per bins' relative abundance; and the vertical axis is the bins' relative abundance, which has been transformed by logarithmic conversion. For each sub-figure, we add a tendency line (ggplot2 package, "lm" method); 95% confidence interval of the tendency line is depicted by shading. Significance of Spearman's rank correlation: p, coefficient of Spearman's rank correlation: ρ.

ice-covered and ice-free stages, while fungal communities were mainly influenced only during the ice-free stage rather than during the ice-covered stage (**Figure 3**). Homogeneous selection includes host filtering. Host filtering of plants in the rhizosphere habitat may be one driver for the turnover of fungal communities between stages. May is the growing season of *Phragmites* and *Scirpus* when the photosynthesis of these two plants is at its peak. This may lead to a stronger filter of roots on fungi than would be possible during the non-growing season (He et al., 2020). Therefore, homogeneous selection controlled more fungi in the ice-free stage compared to the ice-covered stage. Previous studies suggest that plant roots could filter the microbes in specific taxa and decrease the community dissimilarity (Edwards et al., 2015; Trivedi et al., 2020). The dissimilarity of the fungal communities sharply decreased from the ice-covered stage to the ice-free stage, which also provides indirect evidence that the increase in plant host filtering shapes the fungal communities in the ice-free stage (**Figure 2**; **Supplementary Figure 8**). Previous studies showed that *Phragmites* have stronger host activities in the ice-free stage (Fang et al., 2021) and milder host effects in the ice-covered stage (Gao, 2021) compared to other general wetland plants. These are consistent with our result that fungal and bacterial communities in *Phragmites* were more controlled by homogeneous selection in ice-free stages (**Figure 3**). Our study also showed that compared to *Phragmites*, *Scirpus* had a similar host effect during the ice-covered stage (**Supplementary Figure 7**) but milder host effects on fungi and bacteria in the growth stage (**Figure 3**; **Supplementary Figure 7**).

Our "iCAMP" result also showed that the dispersal limitation process of the fungal communities, but not of the bacterial communities, had a sharp decrease from the ice-covered to the ice-free stage (**Figure 3**). Water, in liquid form, is considered to be an ideal medium for the dispersal of microorganisms (Lindström and Langenheder, 2012). The main propagules of fungi are spores or sporocarps, which could disperse in water for a long distance because water, in its liquid form, could prevent them from spore desiccation and ultraviolet light damage (Golan and Pringle, 2017). In this study, the dispersal medium across ice-covered and ice-free stages changed from ice to water, and this may have facilitated the dispersal of fungi *via* spores and therefore decreased the dispersal limitation. The disappeared difference between intra-site and inter-site dissimilarity of fungal communities in the ice-free stage also implies that water, in its liquid form, may have facilitated the dispersal of fungi (**Figure 2**). Since the propagules of most bacteria are themselves, compared to other microorganisms dispersal through spores,

these bacteria are more easily limited by harsh abiotic conditions in long-distance dispersal (Yang and van Elsas, 2018). Moreover, bacteria could also disperse *via* fungal hyphae (Yang and van Elsas, 2018), which could have kept the dispersal of bacteria immune to the change of the abiotic medium in our research (**Figure 3**). This study provides indirect evidence that the change of the dispersal medium drives the spatial variation pattern of the fungal communities between ice-covered (water in solid form) and ice-free (water in liquid form) stages. In future, the underlying mechanisms need to be further verified by more direct evidence, including observations in natural ecosystems and controlled experiments.

Different Effects of Community Assembly Processes on Various Abundance Levels of Microorganisms (Both at the Subcommunity Level and Bin Level)

Previous research has generally studied the relationship between microbial abundance and community assembly processes at the subcommunity level, which could only provide qualitative information. Moreover, this relationship has not been researched in ice-covered wetlands. In this study, we used the "iCAMP" model to investigate this relationship in ice-covered wetlands, both at subcommunity and bin levels.

Our result at subcommunities levels suggested that, during the ice-free stage, abundant fungal subcommunities were more controlled by homogenous selection and the ecological drift process, while intermediate and rare fungal subcommunities were more controlled by dispersal limitation and the ecological drift process (**Figure 4**). Homogeneous ecological processes (including homogeneous dispersal and homogeneous selection) tended to dampen the community dissimilarity (Jia et al., 2018). The sharp decrease in the relative importance of dispersal limitation to abundant fungal subcommunities may have caused the dampened dissimilarity such that the intra-site and inter-site fungal communities were more similar. This result implies that the abundant fungal subcommunities, rather than the rare fungal subcommunities, dampened the dissimilarity of fungal communities during the ice-free stage in this study.

Our result at bin levels demonstrates that the dispersal limitation process showed a slightly higher tendency to control rare microbes, while homogeneous selection highly influenced abundant microbes. Compared to research in other ecological zones, research in bays (Mo et al., 2018) and wastewater (Lee et al., 2021) found similar patterns. Research in terrestrial ecosystems found that homogeneous selection greatly influenced rare microbes (Jiao and Lu, 2020; Zheng et al., 2021). The different niche breadths of rare microbes may be the reason for the different conclusions in different studies. Arguably, the narrow niche breadths of microbial taxa allow those taxa to be influenced by homogeneous selection (Jiao and Lu, 2020). In our study, rare microbes tended to have intermediate niche breadths rather than narrow ones, and abundant microbes tended to have both narrow and wide niche breadths (**Supplementary Figure 6**). Consequently, abundant microbes, rather than rare microbes, tended to be greatly influenced by homogeneous selection. In some studies where rare microbes had narrow niches, they were greatly controlled by homogeneous selection. Another explanation is the dormancy of rare microbes. Under the ice-covered surface, the environmental temperature was low such that some of the microorganisms were in dormancy or exhibiting low growth rates. Dormancy dampens the strength of homogeneous selection and leads to rare microbial taxa but will not cause extinction. It will also cause the homogeneous selection to influence abundant microbes rather than rare microbes. Those tendencies exist in bacterial and fungal communities equally, albeit they are stronger in bacterial communities. The possible explanation is that, during the ice-covered stage, the growth rate of abundant and rare microbes differed to a larger extent. That is, some microbes were active, had faster growth rates, and were more likely to be abundant, but were also more likely to be influenced by homogeneous environmental selection. However, some microbes exhibited very low activity (e.g., microorganisms in a dormant state), had lower growth rates, and were more likely to be rare but less influenced by homogeneous environmental selection. In our study, community assembly processes such as homogeneous selection, ecological drift, or dispersal limitation had a specific relationship with the relative abundance of microbial OTUs. Although diversity and assembly processes of the microbial community were influenced by external environmental factors, such as periods of ice-coverage, the presence of plants, and the dispersal medium, the pattern seemed to be steady and not influenced by these factors (**Supplementary Figure 9**). It might be valuable to verify whether this pattern is widespread in microbial communities and to further investigate the underlying mechanisms of this pattern.

CONCLUSION

Our study suggests that fungal communities have higher alpha diversity under the ice. Homogeneous selection, dispersal limitation, and ecological drift dominated the assembly of bacterial communities during the ice-covered stages and ice-free stages. Dispersal limitation dominated the assembly of fungal communities during the ice-covered stages. Host filtering mainly affected the assembly of fungal communities during the ice-free stage in the Phragmites rhizosphere. In our research, we found that abundant microbes were controlled more by homogeneous selection, while rare microbes were controlled more by ecological drift, dispersal limitation, and heterogeneous selection, especially for bacteria. In future, more directed observation and controlled experiments are required to comprehensively assess these conclusions.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the National Center for Biotechnology Information (NCBI) repository, accession number PRJNA813903.

AUTHOR CONTRIBUTIONS

JM provided concept, designed methodology, performed the field sampling, conducted the data analyses and wrote the manuscript. KM performed the field sampling and reviewed this manuscript. JL supervised this research and reviewed this manuscript. KM, JL, and NC assisted to methodology designing. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Key Research and Development Program of China (No. 2016YFC0500402).

ACKNOWLEDGMENTS

We are grateful to the reviewers for their valuable suggestions on this paper and also we thank the Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Mr. Zhongwei Zhang, and local villagers for their help with sampling.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.783371/full#supplementary-material

REFERENCES

Abarenkov, K., Henrik Nilsson, R., Larsson, K.-H., Alexander, I. J., Eberhardt, U., Erland, S., et al. (2010). The UNITE database for molecular identification of fungi–recent updates and future perspectives. *N. Phytol.* 186, 281–285. doi: 10.1111/j.1469-8137.2009.03160.x

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Bell, C. W., Asao, S., Calderon, F., Wolk, B., and Wallenstein, M. D. (2015). Plant nitrogen uptake drives rhizosphere bacterial community assembly during plant growth. *Soil Biol. Biochem.* 85, 170–182. doi: 10.1016/j.soilbio.2015.03.006

Bertilsson, S., Burgin, A., Carey, C. C., Fey, S. B., Grossart, H.-P., Grubisic, L. M., et al. (2013). The under-ice microbiome of seasonally frozen lakes. *Limnol. Oceanogr.* 58, 1998–2012. doi: 10.4319/lo.2013.58.6.1998

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., et al. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* 10, 57–59. doi: 10.1038/nmeth.2276

Butler, T. M., Wilhelm, A.-C., Dwyer, A. C., Webb, P. N., Baldwin, A. L., and Techtmann, S. M. (2019). Microbial community dynamics during lake ice freezing. *Sci. Rep.* 9:6231. doi: 10.1038/s41598-019-42609-9

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Ceja-Navarro, J. A., Wang, Y., Ning, D., Arellano, A., Ramanculova, L., Yuan, M. M., et al. (2021). Protist diversity and community complexity in the rhizosphere of switchgrass are dynamic as plants develop. *Microbiome* 9:96. doi: 10.1186/s40168-021-01042-9

Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43, 783–791. doi: 10.2307/2531532

Chaparro, J. M., Badri, D. V., and Vivanco, J. M. (2014). Rhizosphere microbiome assemblage is affected by plant development. *ISME J* 8, 790–803. doi: 10.1038/ismej.2013.196

Chase, J. M., and Myers, J. A. (2011). Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 2351–2363. doi: 10.1098/rstb.2011.0063

Chave, J. (2004). Neutral theory and community ecology. *Ecol. Lett.* 7, 241–253. doi: 10.1111/j.1461-0248.2003.00566.x

Chen, X., Gong, L., Zhao, J., and Zhu, H. (2021). Litter decomposition, microbial community dynamics and their relationships under seasonal snow cover. *Ecol. Eng.* 159:106089. doi: 10.1016/j.ecoleng.2020.106089

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* 31, 343–366. doi: 10.1146/annurev.ecolsys.31.1.343

Cruaud, P., Vigneron, A., Fradette, M.-S., Dorea, C. C., Culley, A. I., Rodriguez, M. J., et al. (2020). Annual bacterial community cycle in a seasonally ice-covered river reflects environmental and climatic conditions. *Limnol. Oceanogr.* 65:S21–S37. doi: 10.1002/lno.11130

Dong, Y., Sanford, R. A., Connor, L., Chee-Sanford, J., Wimmer, B. T., Iranmanesh, A., et al. (2021). Differential structure and functional gene response to geochemistry associated with the suspended and attached shallow aquifer microbiomes from the Illinois Basin, IL. *Water Res.* 202:117431. doi: 10.1016/j.watres.2021.117431

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996–998. doi: 10.1038/nmeth.2604

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., et al. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112, E911–E920. doi: 10.1073/pnas.1414592112

Fang, J., Dong, J., Li, C., Chen, H., Wang, L., Lyu, T., et al. (2021). Response of microbial community composition and function to emergent plant rhizosphere of a constructed wetland in northern China. *Appl. Soil Ecol.* 168:104141. doi: 10.1016/j.apsoil.2021.104141

Fitzpatrick, C. R., Copeland, J., Wang, P. W., Guttman, D. S., Kotanen, P. M., and Johnson, M. T. J. (2018). Assembly and ecological function of the root microbiome across angiosperm plant species. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1157–E1165. doi: 10.1073/pnas.1717617115

Fouquier, J., Rideout, J. R., Bolyen, E., Chase, J., Shiffer, A., McDonald, D., et al. (2016). Ghost-tree: creating hybrid-gene phylogenetic trees for diversity analyses. *Microbiome* 4:11. doi: 10.1186/s40168-016-0153-6

Gao, G. (2021). *The Effect of Freezing on the Substrate Environment and Microorganisms Around Plant Roots in Agricultural Drainage Wetlands*. Changchun, Northeast Normal University.

Golan, J. J., and Pringle, A. (2017). Long-distance dispersal of fungi. *Microbiol. Spectr.* 5. doi: 10.1128/9781555819583.ch14

Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., et al. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* 21, 494–504. doi: 10.1101/gr.112730.110

Haei, M., Rousk, J., Ilstedt, U., Öquist, M., Bååth, E., and Laudon, H. (2011). Effects of soil frost on growth, composition and respiration of the soil microbial decomposer community. *Soil Biol. Biochem.* 43, 2069–2077. doi: 10.1016/j.soilbio.2011.06.005

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nat. Rev. Microbiol.* 10, 497–506. doi: 10.1038/nrmicro2795

He, R., Zeng, J., Zhao, D., Huang, R., Yu, Z., and Wu, Q. L. (2020). Contrasting patterns in diversity and community assembly of phragmites australis root-associated bacterial communities from different seasons. *Appl. Environ. Microbiol.* 86:e00379–20. doi: 10.1128/AEM.00379-20

Jansen, J., MacIntyre, S., Barrett, D. C., Chin, Y.-P., Cortés, A., Forrest, A. L., et al. (2021). Winter limnology: how do hydrodynamics and biogeochemistry shape ecosystems under ice? *J. Geophys. Res. Biogeosci* 126:e2020JG006237. doi: 10.1029/2020JG006237

Jia, X., Dini-Andreote, F., and Falcão Salles, J. (2018). Community assembly processes of the microbial rare biosphere. *Trends Microbiol.* 26, 738–747. doi: 10.1016/j.tim.2018.02.011

Jiao, S., and Lu, Y. (2020). Abundant fungi adapt to broader environmental gradients than rare fungi in agricultural fields. *Glob. Change Biol.* 26, 4506–4520. doi: 10.1111/gcb.15130

Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., et al. (2017). Where less may be more: how the rare biosphere pulls ecosystems strings. *ISME J* 11, 853–862. doi: 10.1038/ismej.2016.174

Juan, Y., Jiang, N., Tian, L., Chen, X., Sun, W., and Chen, L. (2018). Effect of freeze-thaw on a midtemperate soil bacterial community and the correlation network of its members. *Biomed. Res. Int.* 2018:8412429. doi: 10.1155/2018/8412429

Le Roux, J. J., Crous, P. W., Kamutando, C. N., Richardson, D. M., Strasberg, D., Wingfield, M. J., et al. (2021). A core of rhizosphere bacterial taxa associates with two of the world's most isolated plant congeners. *Plant Soil* 468:1–18. doi: 10.1007/s11104-021-05049-x

Lee, S.-H., Kim, T.-S., and Park, H.-D. (2021). Transient-rare bacterial taxa are assembled neutrally across temporal scales. *Microbes Environ.* 36:ME20110. doi: 10.1264/jsme2.ME20110

Liao, J., Cao, X., Wang, J., Zhao, L., Sun, J., Jiang, D., et al. (2017). Similar community assembly mechanisms underlie similar biogeography of rare and abundant bacteria in lakes on Yungui Plateau, China. *Limnol. Oceanogr.* 62, 723–735. doi: 10.1002/lno.10455

Lindström, E. S., and Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* 4, 1–9. doi: 10.1111/j.1758-2229.2011.00257.x

Magoč, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Magurran, A. E., and Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714–716. doi: 10.1038/nature01547

Martinez, K. A., Gibson, D. J., and Middleton, B. A. (2015). Core-satellite species hypothesis and native versus exotic species in secondary succession. *Plant. Ecol.* 216, 419–427. doi: 10.1007/s11258-015-0446-z

Matthews, A., Pierce, S., Hipperson, H., and Raymond, B. (2019). Rhizobacterial community assembly patterns vary between crop species. *Front. Microbiol.* 10:581. doi: 10.3389/fmicb.2019.00581

Meng (2020). *Analysis of Wetland Agglomeration System Network Evolution and Hydrological Relationship in Lower Nenjiang River Basin*. Beijing, China, Beijing Normal University.

Meng, B., Liu, J., Kun, B., and Bin, S. (2019). Water fluxes of Nenjiang River Basin with ecological network analysis: Conflict and coordination between agricultural development and wetland restoration. *J. Clean. Prod.* 213, 933–943. doi: 10.1016/j.jclepro.2018.12.243

Mo, Y., Zhang, W., Yang, J., Lin, Y., Yu, Z., and Lin, S. (2018). Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. *ISME J* 12, 2198–2210. doi: 10.1038/s41396-018-0153-6

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., et al. (2013). Patterns and processes of microbial community assembly. *Microbiol. Mol. Biol. Rev.* 77, 342–356. doi: 10.1128/MMBR.00051-12

Nilsson, R. H., Larsson, K.-H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. (2019). The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47, D259–D264. doi: 10.1093/nar/gky1022

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., et al. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat. Commun.* 11:4717. doi: 10.1038/s41467-020-18560-z

Oksanen, J., Blanchet, G. F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019). *vegan: Community Ecology Package.* R package version 2.5-7. Available online at: https://CRAN.R-project.org/package=vegan (accessed: November 22, 2021)

Pavoine, S. (2020). adiv: An r package to analyse biodiversity in ecology. *Methods Ecol. Evol.* 11, 1106–1112. doi: 10.1111/2041-210X.13430

Philippot, L., Raaijmakers, J. M., Lemanceau, P., and van der Putten, W. H. (2013). Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.* 11, 789–799. doi: 10.1038/nrmicro3109

Prashar, P., Kapoor, N., and Sachdeva, S. (2014). Rhizosphere: its structure, bacterial diversity and significance. *Rev. Environ. Sci. Biotechnol.*13, 63–77. doi: 10.1007/s11157-013-9317-z

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks,1219

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Robroek, B. J. M., Heijboer, A., Jassey, V. E. J., Hefting, M. M., Rouwenhorst, T. G., Buttler, A., et al. (2013). Snow cover manipulation effects on microbial community structure and soil chemistry in a mountain bog. *Plant Soil* 369, 151–164. doi: 10.1007/s11104-012-1547-2

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584

Sang, C., Xia, Z., Sun, L., Sun, H., Jiang, P., Wang, C., et al. (2021). Responses of soil microbial communities to freeze–thaw cycles in a Chinese temperate forest. *Ecol. Process.* 10:66. doi: 10.1186/s13717-021-00337-x

Schloss, P. D. (2009). A high-throughput DNA sequence aligner for microbial ecology studies. *PLoS ONE* 4:e8230. doi: 10.1371/journal.pone.0008230

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., et al. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J* 7, 2069–2079. doi: 10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., and Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* 6, 1653–1664. doi: 10.1038/ismej.2012.22

Sun, B., Liu, J., Meng, B., and Bao, K. (2020). Structural variability and co-occurrence pattern differentiation in rhizosphere microbiomes of the native invasive plant Echinochloa caudate in Momoge National Nature Reserve, China. *Wetlands* 40, 587–597. doi: 10.1007/s13157-019-01209-z,

Sun, C., Zhang, B., Ning, D., Zhang, Y., Dai, T., Wu L., et al. (2021). Seasonal dynamics of the microbial community in two full-scale wastewater treatment plants: diversity, composition, phylogenetic group based assembly and co-occurrence pattern. *Water Res.* 200:117295. doi: 10.1016/j.watres.2021.117295

Sun, Y., Zhang, M., Duan, C., Cao, N., Jia, W., Zhao Z., et al. (2021). Contribution of stochastic processes to the microbial community assembly on field-collected microplastics. *Environm. Microbiol.* 23, 6707–6720. doi: 10.1111/1462-2920.15713

Tedersoo, L., Sánchez-Ramírez, S., Kõljalg, U., Bahram, M., Döring, M., Schigel, D., et al. (2018). High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal Divers.* 90, 135–159. doi: 10.1007/s13225-018-0401-0

Tran, P., Ramachandran, A., Khawasik, O., Beisner, B. E., Rautio, M., Huot, Y., et al. (2018). Microbial life under ice: Metagenome diversity and in situ activity of Verrucomicrobia in seasonally ice-covered Lakes. *Environ. Microbiol.* 20, 2568–2584. doi: 10.1111/1462-2920.14283

Trivedi, P., Leach, J. E., Tringe, S. G., Sa, T., and Singh, B. K. (2020). Plant-microbiome interactions: from community assembly to plant health. *Nat. Rev. Microbiol.* 18, 607–621. doi: 10.1038/s41579-020-0412-1

Vellend, M. (2016). The Theory of Ecological Communities (MPB-57). *Princeton, NJ: Princeton University Press.* doi: 10.1515/97814008 83790

Wan, W., Liu, S., Li, X., Xing, Y., Chen, W., and Huang, Q. (2021). bridging rare and abundant bacteria with ecosystem multifunctionality in salinized agricultural soils: from community diversity to environmental adaptation. *mSystems* 6:e01221–20. doi: 10.1128/mSystems.01 221-20

Yang, P., and van Elsas, J. D. (2018). Mechanisms and ecological implications of the movement of bacteria in soil. *Appl. Soil Ecol.* 129, 112–120. doi: 10.1016/j.apsoil.2018.04.014

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* 42, D643–D648. doi: 10.1093/nar/gkt1209

Zhalnina, K., Louie, K. B., Hao, Z., Mansoori, N., da Rocha, U. N., Shi, S., et al. (2018). Dynamic root exudate chemistry and microbial substrate preferences

drive patterns in rhizosphere microbial community assembly. *Nat. Microbiol.* 3, 470–480. doi: 10.1038/s41564-018-0129-3

Zhang, L., Wang, A., Yang, W., Xu, Z., Wu, F., Tan, B., et al. (2017). Soil microbial abundance and community structure vary with altitude and season in the coniferous forests, China. *J. Soils Sediments* 17, 2318–2328. doi: 10.1007/s11368-016-1593-0

Zheng, S., Liu, W., Ding, S., Yu, X., Zou, Y., and Luan, J. (2019). Hydrothermal regime of the soil during freezing and thawing period under different underlying surfaces in Momoge Wetlands. *Wetl. Sci.* 17, 68–73. doi: 10.13248/j.cnki.wetlandsci.2019.01.009

Zheng, W., Zhao, Z., Lv, F., Wang, R., Wang, Z., Zhao, Z., et al. (2021). Assembly of abundant and rare bacterial and fungal subcommunities in different soil aggregate sizes in an apple orchard treated with cover crop and fertilizer. *Soil Biol. Biochem.* 156:108222. doi: 10.1016/j.soilbio.2021.108222

Zhou, J., and Ning, D. (2017). Stochastic community assembly: does it matter in microbial ecology? *Microbiol. Mol. Biol. Rev.* 81, e00002–e00017. doi: 10.1128/MMBR.00002-17

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership