Distinct ecological mechanisms drive the spatial scaling patterns of abundant and rare microbial communities in an ocean sediment

Qichao Tu[1], Xia Liu[1], Hongjun Li[2], and Wen Song[1]

[1]Shandong University Institute of Marine Science and Technology
[2]National Marine Environmental Monitoring Center

June 15, 2022

Abstract

Revealing the ecological mechanisms driving the diversity patterns followed by microbial communities across space and through time is an essential issue in microbial community ecology. In this study, two typical spatial scaling patterns, including diversity-area and distance-decay relationships, were investigated for microbial communities in an ocean sediment ecosystem. Strong spatial scaling patterns were observed at the whole community level and for the rare subcommunities, but hardly for the abundant subcommunities. Rare subcommunities were mainly responsible for the observed spatial scaling patterns, as also confirmed by extending spatial scaling diversity metrics to Hill numbers. Distinct ecological mechanisms underlay the differed spatial scaling patterns followed by abundant and rare subcommunities. Both environmental heterogeneity and local community assembly mechanisms drove the microbial spatial scaling patterns. Environmental heterogeneity was significantly associated with the spatial scaling metrics of rare but not abundant subcommunities. Strong ecological drift and dispersal limitation underlay the spatial scaling patterns of rare subcommunities, whereas high homogeneous selection weakened the spatial scaling patterns of abundant subcommunities. Such differed mechanisms driving the spatial scaling patterns of abundant and rare subcommunities were also experimentally confirmed by deep sequencing experiments. This study links microbial spatial scaling patterns with ecological mechanisms, providing novel mechanistic insights into the diversity patterns followed by different types of microbes.

1. Introduction

Microbes are ubiquitous in the Earth's biosphere, executing essential ecosystem functions and maintaining ecosystem stability(Bardgett & van der Putten, 2014; Fuhrman, Cram, & Needham, 2015; B. Gilbert & Lechowicz, 2004). The diversity of soil microbial communities is in general positively associated with ecosystem multifunctioning (Bradford et al., 2014; Delgado-Baquerizo et al., 2016; Lefcheck et al., 2015). As a major component of microbial communities, revealing the diversity patterns across space and through time is of essential importance for better understanding the underlying ecological mechanisms governing the distribution and assembly of microbial communities. However, the tiny size of microorganisms and immense complexity of microbial communities make this issue more challenging than macrobial communities.

Taxa-area relationship (TAR) and distance-decay relationship (DDR) are two typical and perhaps universal spatial scaling patterns followed by both macrobial and microbial communities(Green et al., 2004; M. C. Horner-Devine, M. Lage, J. B. Hughes, & B. J. Bohannan, 2004; Tu et al., 2016; Zinger, Boetius, & Ramette, 2014). Of these, TAR describes the pattern of continuously increasing species richness with increasing sampling area(Connor & Mccoy, 1979; Rosenzweig, 1995), whereas DDR describes the pattern that the composition of biological communities becomes more dissimilar with increasing geographic distance(Nekola & White, 1999). Although different in concept, both TAR and DDR are assumed to be the result of a set of common processes, including environmental heterogeneity and local community assembly processes (e.g., speciation, drift, and dispersal limitation) across sampling area and distance(Connor & Mccoy, 1979;

Hubbell, 2001). Specifically, higher environmental heterogeneity is associated with more ecological niche space and habitat types, allowing more microbial taxa to coexist(Allouche, Kalyuzhny, Moreno-Rueda, Pizarro, & Kadmon, 2012; Huber et al., 2020; Yang et al., 2015). The larger sampling area it is, the higher environmental heterogeneity and more coexisted microbial taxa are expected, resulting in TAR patterns. Environmental heterogeneity also contributes to DDR patterns for its being strongly correlated with geographic distance(Tilman, 1983). Local community assembly processes may also result in differed community structure and composition(Stegen et al., 2013; X. Zhang et al., 2020). leading to TAR and DDR patterns. However, TAR and DDR may not be directly derived from each other, and may be subjected to influences by different ecological factors(Zinger et al., 2014).

Microbial communities in natural ecosystems are typically composed by a small number of abundant taxa and an extremely long tail of rare taxa(M. D. Lynch & Neufeld, 2015; Sogin et al., 2006). The abundant taxa usually occupy $< 20\%$ of the total richness, but $> 80\%$ in relative abundance(Sogin et al., 2006). Although low in relative abundance, recent studies suggest that the rare microbial taxa execute nonnegligible ecosystem functions in the environment(Q.-L. Chen et al., 2020; Lyons & Schwartz, 2001; Mouillot et al., 2013; Xiong et al., 2021). Recent studies suggested that the abundant and rare subcommunities are structured by different community assembly mechanisms and environmental parameters(Jiao & Lu, 2020; Mo et al., 2018; W. Zhang et al., 2018). However, it remains not clear whether and how abundant and rare subcommunities differ in the spatial scaling patterns they may follow and how such patterns are linked to environmental heterogeneity and local community assembly processes.

In this study, we investigated the spatial scaling patterns followed by abundant and rare subcommunities of microbes in an ocean sediment ecosystem, aiming to address the following ecological questions: (1) Do abundant and rare subcommunities differ in following spatial scaling patterns? (2) How do environmental heterogeneity and local community assembly mechanisms respectively contribute to the spatial scaling patterns? We expected that abundant and rare subcommunities may differ in the spatial scaling patterns they follow, mainly due to their different life strategies (e.g., different adaptability to environmental conditions)(He et al., 2022; Wan et al., 2021). Specifically, abundant subcommunities may follow weak spatial scaling patterns, especially TAR, as they are more broadly distributed across the sampling space. As previously reported, abundant and rare subcommunities differ dramatically in local community assembly mechanisms(Jiao & Lu, 2020; Mo et al., 2018; W. Zhang et al., 2018). We therefore expected strong links between local community assembly and spatial scaling patterns. The results confirmed our expectation that spatial scaling patterns were rarely observed for abundant subcommunities, whereas rare subcommunities were mainly responsible for the observed microbial spatial scaling patterns. Distinct ecological mechanisms underlay the spatial scaling patterns followed by abundant and rare subcommunities. The study provided novel mechanistic insights into the spatial scaling patterns followed by different types of microbes.

2. Materials and methods

2.1 Experimental design and sample collection

A total of 29 sedimental samples were collected in the Beibu Gulf, a semi-enclosed oceanic bay located in the southern coast of China (Supplementary Figure 1). Samples were collected from a near rectangle area that was 149.25 km long and 128.73 km wide, covering approximately 19212.57 km^2. The sampling area were then divided into 9 blocks, with each block covering 3-4 samples. By merging neighborhood blocks, a series of multi-block sampling areas were generated, enabling TAR analysis in the experiment.

2.2 Environmental variables

A total of 19 environmental variables were measured at each sampling station, including temperature, salinity, pH, total organic carbon (TOC) i, total nitrogen (TN) , total phosphorus (TP) , ammonium nitrogen ($NH4^+$-N) , nitrate nitrogen (NO^{-3}-N) , nitrite nitrogen (NO^{-2} -N) , Organic matter , Sulfate (SO^{4-}) , Petroleum and sulfide. In addition, seven heavy metals, including Hg, Cd, Pb, Cr, As, Cu, and Zn. Of these, the pH value was measured by adding 5 ml of distilled water to 2 g of the precipitate and recording the pH value using a pH electrode (STARTER 300, OHAUS, Beijing, China). Total nitrogen (TN) content was analyzed by Kjeldahl

method(J. M. Lynch & Barbano, 1999). Total phosphorus (TP) content was determined by molybdenum blue colorimetry at 660 nm after hydrofluoric and perchloric acid digestion. Ammonium nitrogen (NH+4 -N), nitrate nitrogen (NO-3 - N), nitrite nitrogen (NO-2 -N) were quantified by automated discrete analysis (CleverChem 380, Germany)(Islam, Sarker, Yamamoto, Wahab, & Tanaka, 2004). Heavy metal element content was determined by inductively coupled plasma mass spectrometry (ICP-MS, Optima, 2000 DV, Perkin Elmer, USA).

2.3 DNA extraction, PCR amplification, and sequencing

Microbial DNA was extracted from 0.5 g soil samples using Soil DNA Mini kit (Omega Bio-Tek, Norcross, GA, USA) according to the manufacturer's protocol. The DNA preparation and sequencing library preparation were performed following the procedures described by Scholer and Vestergaard(Griffiths, Whiteley, O'Donnell, & Bailey, 2000). The V3-V4 region of the bacterial 16S rRNA gene were PCR amplified (95 °C for 2 min, followed by 27 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 45 s, with a final extension at 72 °C for 10 min) using the primer set 341F (5'-CCTACGGRRBGCASCAGKVRVGAAT-3') and 806R (5'-GGACTACNVGGGTWTCTAATCC-3'). A six-base barcode was added to each library for further demultiplexing samples. All samples were subject to paired-end high throughput sequencing at regular sequencing depth. The BBW11 sample was also subject to deep sequencing to achieve at least 1 million reads per sample, in addition to regular depth sequencing. The Illumnina HiSeq 2500 (Illumina, Inc., San Diego, CA, USA) was used for sequencing.

2.4 Sequence data processing

The raw FASTQ files were subject to a series of standard processing including demultiplexing, read merging, quality filtering and chimeric removal using DADA2 (version 1.20.0)(Callahan et al., 2016). A relative abundance microbial profile was generated at the level of amplicon sequence variant (ASV). The taxonomic information for each ASV was determined using the Ribosomal Database Project (RDP) Classifier (*http://rdp.cme.msu.edu*)(Cole et al., 2009) with a confidence interval of 80%. The "Rarefy" function in the "GuniFrac" package is used to rarefy the microbial profile to a same sequencing depth before further statistical analyses were carried out. All analyses were completed in R v 4.1.2.

Microbial ASVs were classified into two different categories, including abundant and the rare taxa, according to their relative abundance and/or frequency(Bickel & Or, 2021). Here, ASVs with an average relative abundance of <0.01% across all samples were defined as rare taxa, whereas the remaining ones were defined as abundant ASVs. Notably, different criteria were used to define abundant and rare taxa in different studies(Jiao, Chen, & Wei, 2017; Y. Xue et al., 2018). Since only 72 ASVs were not rare, we classified all of them as abundant, without considering more precise concept such as occasional taxa.

For all ASVs, indices including niche breadth and niche overlap were calculated to see how well they may adapt to the environment. The niche breadth was evaluated using the Levins' standardized niche breadth index(Feinsinger, Spears, & Poole, 1981). The niche overlap was calculated using Pianka's niche overlap index equation, with the value of Pianka's index between 0 and 1(Pianka, 1974). The R package "spaa" was employed to calculate niche breadth and niche overlap indices.

2.5 Spatial scaling pattern analyses

Two typical spatial scaling patterns were analyzed, including TAR and DDR. The community richness (alpha-diversity) and Bray-Curtis community similarity (beta-diversity) were respectively used as diversity metrics for TAR and DDR analyses. The slope coefficients between log-transformed diversity indices and log-transformed geographic area/distance were calculated for TAR and DDR. Both the diversity indices of TAR and DDR were extended to different diversity orders using Hill numbers, which are a parametric family of diversity indices differentiated by the parameter q (Chao, Chiu, & Jost, 2014). The extended TAR and DDR were respectively termed as DAR_q and DDR_q. The R package "hillR"(Chiu & Chao, 2014) (*https://github.com/dajiang/hillR*) was used for both alpha- and beta-diversity indices calculation under different diversity order q .

In addition, we also correlated the diversity metrics with environmental heterogeneity to investigate how microbial spatial scaling patterns were affected by environmental conditions. Here, the Euclidean distance based on normalized environmental variables was calculated to represent the environmental heterogeneity between different samples. All 19 environmental variables were included for environmental heterogeneity calculation. The analysis, as well as TAR and DDR, were carried out for the whole community, the abundant and the rare subcommunities.

2.6 Inferring local community assembly mechanisms

To quantify the relative importance of ecological processes in structuring the bacterial and fungal metacommunity, the iCAMP approach was employed(Ning et al., 2020), which is a more sophisticated development of the approach proposed by Stegen et al. (Stegen et al., 2013; Stegen, Lin, Konopka, & Fredrickson, 2012). The iCAMP approach uses a quantitative framework to infer community assembly mechanisms through a phylogenetic-bin-based null model analysis. Within this framework, ecological processes are divided into five processed, including homogeneous selections (HoS), heterogeneous selections (HeS), dispersal limitation (DL), homogenizing dispersal (HD) and drift (DR) processes. In the approach, multiple phylogenetic bins were generated based on the phylogenetic tree. The null model analysis within each bin is calculated by beta Net Relatedness Index (βNRI) and modified Raup–Crick metric (RC). The fraction of pairwise comparisons with βNRI < - 1.96 is considered as the percentages of homogeneous selection, whereas those with βNRI > +1.96 as the percentages of heterogeneous selection. Next, taxonomic diversity metric RC is used to partition the remaining pairwise comparison with |βNRI| [?] 1.96. The fraction of pairwise comparisons with RC < - 0.95 is treated as the percentages of homogenizing dispersal, while those with RC > + 0.95 as dispersal limitation. The remaining ones with |βNRI| [?] 2 and | RC | [?] 0.95 represent the percentages of drift. The βNRI and RC are calculated using the "picante" package(Kembel et al., 2010) and "iCAMP"(Ning et al., 2020) package in R.

3. Results

3.1 Overall diversity of the sedimental microbial communities

A total of 29 sedimental samples in the Beibu Gulf were collected and subjected to total DNA extraction. For each sample, the V3-V4 region of the 16S rRNA gene were amplified and sequenced, targeting the bacterial communities in the Beibu Gulf sediment. After data processing including quality filtering, chimera removal and rarefaction, 8,007 merged sequences per sample were retained. These sequences were then clustered into 13,073 bacterial ASVs (Supplementary Table 1). By applying 0.01% relative abundance as the cutoff, 13,001 bacterial ASVs (51.19% in relative abundance) were classified as rare taxa, and 72 ASVs (48.81% in relative abundance) as abundant taxa.

Taxonomically, the bacterial communities were dominated by *Gamma-Proteobacteria* (31.39%), *Delta-Proteobacteria*(20.01%), *Acidobacteria* (10.57%), *Bacteroidetes*(9.62%), and *Actinobacteria* (3.69%) (Fig. 1a). On average, each sediment sample was found with 299 ASVs, of which 65 were abundant and 235 were rare (Fig. 1b). Further analysis suggested that abundant and rare subcommunities differed dramatically in Pielou's evenness and Shannon-Wiener diversity indices (Supplementary Figure 2), as well as within community similarity (Fig. 1c). Such results suggested that the abundant and rare subcommunities tended to differ in spatial scaling patterns as well as local community assembly mechanisms.

3.2 Rare taxa were mainly responsible for microbial spatial scaling patterns

We first investigated whether abundant and rare subcommunities followed similar spatial scaling patterns, such as TAR and DDR (Fig. 2a and b). The typical community richness and Bray-Curtis similarity were respectively used to quantify TAR and DDR. As a result, clear TAR (z = 0.494, $P < 0.001$) and DDR (d = -0.242, $P < 0.001$) patterns were observed for the bacterial communities (Fig. 2a and b). As expected, stronger TAR pattern was found for rare subcommunities (z = 0.517, $P< 0.001$) than abundant subcommunities (z = 0.029,P = 0.034) (Fig. 2a). Similarly, rare subcommunities (d = -0.447, $P < 0.001$) harbored stronger DDR patterns than abundant subcommunities (d = -0.12, $P < 0.001$) (Fig. 2b). For

both TAR and DDR, rare subcommunities even showed stronger spatial scaling patterns than the whole community. The results suggested that rare subcommunities were mainly responsible for the spatial scaling patterns followed by microbial communities.

To verify the major contribution of rare subcommunities to the spatial scaling patterns of microbial communities, we extended both alpha- and beta-diversity to Hill numbers to analyze TAR and DDR patterns at different diversity orders. As such, the ambiguous definition of abundant and rare taxa can be well resolved by giving continuously decreasing weight on rare taxa. By setting the diversity order $q$to different values (here 0 [?] q [?] 2), different weight is given to microbial taxa with different relative abundance. The higher diversity order q is, the lower weight is given to rare taxa. Taking alpha diversity for example, the Hill numbers equal to community richness when $q = 0$, indicating all microbial taxa are equally treated. When the order q is set to 1 and 2, the Hill numbers respectively equals to the Shannon-Wiener and Simpson diversity index (Supplementary Figure 4). In the case rare subcommunities were mainly responsible for the microbial spatial scaling patterns, decreased DAR_q and DDR_q slope coefficients with increasing q values were expected. As a result, sharply decreased slope coefficients of DAR_q and DDR_q were observed when the diversity order q increased from 0 to 2 (Fig. 2c). Such results confirmed that rare subcommunities were mainly responsible for the observed spatial scaling patterns followed by microbial communities.

3.3 Linking environmental heterogeneity with microbial spatial scaling patterns

Environmental heterogeneity could be an important factor responsible for the spatial scaling patterns of biological communities. Therefore, we first investigated whether and how environmental heterogeneity was associated with microbial spatial scaling patterns. For each sample pair, the Euclidean distance was calculated based on a set of 19 environmental factors and used as environmental heterogeneity. For bacterial communities, significant associations were observed between environmental heterogeneity and microbial spatial scaling metrics, except for abundant subcommunities (Fig. 3a and b). Rare subcommunities were found with stronger association with environmental heterogeneity than abundant subcommunities (Fig. 3a and b), suggesting that environmental heterogeneity more influenced rare subcommunities. Additionally, weak association was observed between environmental heterogeneity and geographic distance (Supplementary Figure 3). The results suggested that environmental heterogeneity played important roles in driving the spatial scaling patterns of sedimental microbial communities via rare subcommunities.

3.4 Local community assembly mechanisms also drove the spatial scaling patterns of microbial communities

To further disentangle the underlying mechanisms driving the spatial scaling patterns of microbial communities, especially the different patterns between abundant and rare subcommunities, the following experimental and statistical investigations were carried out.

First, deep sequencing of a randomly selected sample (BBW11) was performed to investigate the dispersal potential of microbial communities. Here, microbial ASVs were mapped to the deep sequencing dataset at the levels of ASV and read (Fig. 4 a and b). As a result, as high as 72.62% microbial ASVs and 88.57% reads could be mapped to the deep sequenced dataset. Dramatically differed mapping ratios were observed between abundant and rare subcommunities. Abundant ASVs (100%) and reads (100%) can be completely mapped to the deep sequencing datasets. In contrast, rare ASVs (65.22%) and reads (74.46%) were mapped to the deep sequencing datasets at much lower ratios. The results suggested that abundant taxa had higher dispersal rate and better adaptability to the environment than rare taxa.

Second, the niche breadth and niche overlap were also calculated using the Levins' standardized niche breadth index and the Pianka's niche overlap index (Fig. 4c and d). In general, rare ASVs were similar with the whole community regarding the niche breadth and niche overlap. In contrast, abundant ASVs had much higher niche breadth and overlaps than rare ASVs. Such results were consistent with the deep sequencing experiment and suggested that the abundant taxa can better adapt to the environment and coexist with each other than rare taxa.

Third, null model analysis was carried out to investigate the links between local community assembly mech-

anisms and microbial spatial scaling patterns. According to βNRI and RC_{bray} values, the contribution of five different processes to the compositional variations of microbial communities were quantified, including homogeneous selection, heterogeneous selection, dispersal limitation, homogeneous dispersal, and drift (Fig. 5). At the whole community level, drift (62.32%) is mainly responsible for the compositional variations of microbial communities, followed by dispersal limitation (24.45%) and homogeneous selection (10.56%) (Fig. 5a). Distinct community assembly processes were observed for abundant and rare subcommunities. Abundant subcommunities were mainly structured by homogeneous selection (38.69%) and drift (36.55%), whereas rare subcommunities were mainly structured by drift (58.8%) and dispersal limitation (24.04%) (Fig. 5b and c). Such differed contributions of homogeneous selection and dispersal limitation to abundant and rare subcommunities were consistent with the deep sequencing experiment results. The results here demonstrated that distinct community assembly mechanisms shaped the compositional variations of abundant and rare subcommunities, resulting in differed spatial scaling patterns.

4. Discussion

Revealing the underlying mechanisms driving the spatial scaling patterns of the complex soil microbial communities is an essential issue in microbial ecology and community ecology(Jiang et al., 2018; O'Brien et al., 2016). In this study, we focused on the ecological mechanisms structuring the spatial scaling patterns of abundant and rare subcommunities in an ocean sediment ecosystem. Rare subcommunities were mainly responsible for the spatial scaling patterns followed by microbes. Environmental heterogeneity was significantly associated with the spatial scaling metrics of whole community and rare subcommunities, but not abundant subcommunities. Further experimental and statistical analysis suggested that distinct ecological mechanisms underlay the spatial scaling patterns followed by abundant and rare subcommunities.

We found that rare subcommunities were mainly responsible for the spatial scaling patterns followed by microbes. The slope coefficients of rare subcommunities even exceeded the values of the whole community. The role of rare microbial subcommunities has been ambiguous and awkward in microbial ecology, and were frequently ignored due to their low relative abundance and frequency in microbial profiles(Sogin et al., 2006). However, recent studies demonstrated that rare microbial taxa execute important ecosystem functions in various ecosystems(Q.-L. Chen et al., 2020; Xiong et al., 2021; M. Xue et al., 2020). Recent studies in macrobial ecology also show that common species contribute little to the spatial scaling pattern of functional diversity(van Schalkwyk, Pryke, & Samways, 2019; White, Pakeman, & Buckley, 2022). This suggests that the major contribution of rare subcommunities to the spatial scaling patterns might be a common rule in both macrobial and microbial community ecology.

Traditional microbial TAR and DDR analyses generally employ community richness and Bray-Curtis similarity indices and do not distinguish abundant and rare subcommunities(Barreto, Conrad, Klose, Claus, & Enrich-Prast, 2014; Feinstein & Blackwood, 2012; M. C. Horner-Devine, M. Lage, J. B. Hughes, & B. J. M. Bohannan, 2004). A recent study extended TAR to DAR using Hill numbers, incorporating species abundance in spatial scaling analysis(Ma, 2018). Consensus has been achieved that Hill numbers, also knowns as the effective number of species, are the best choice to quantify abundance-based species diversity(Ellison, 2010). In this study, Hill numbers were also employed to quantify both alpha- and beta-diversity, showing decreasing spatial scaling patterns when lower weight was given to rare taxa (i.e., increasing q value). The employment of Hill numbers confirmed the major contribution of rare subcommunities to microbial spatial scaling patterns, bypassing the ambiguous definition of abundant and rare taxa.

Multiple ecological mechanisms may drive the plain and well recognized spatial scaling patterns of biological communities. Environmental heterogeneity (niche theory) and dispersal limitation (neutral theory) are generally considered as the most important factors responsible for the spatial scaling patterns(B. Gilbert & Lechowicz, 2004; Stein, Gerstner, & Kreft, 2014). The effects of environmental conditions on microbial communities may differ for abundant and rare subcommunities, as revealed by previous studies(Jiao et al., 2017; Jiao & Lu, 2020; Mo et al., 2018). In this study, significant associations between environmental heterogeneity and the spatial scaling metrics of whole and rare subcommunities were observed, but not with that of abundant subcommunities. This suggested that environmental heterogeneity was an important factor

responsible for the spatial scaling patterns of microbial communities, via affecting rare subcommunities.

In addition to environmental heterogeneity, local community assembly mechanisms may also contribute to the spatial scaling patterns of biological communities, such as processes like dispersal limitation(W. Chen, Jiao, Li, Du, & Yang, 2019). The distinct local community assembly mechanisms for abundant and rare subcommunities well explained their differed spatial scaling patterns. Specifically, high homogeneous selection result in highly similar communities(Vellend, 2010; Zhou & Ning, 2017), leading to flat spatial scaling patterns for abundant subcommunities. In contrast, dispersal limitation and drift result in highly dissimilar communities(Vellend, 2010; Zhou & Ning, 2017), leading to strong spatial scaling patterns for rare subcommunities. Noteworthily, the deep sequencing experiment, which was previously used to uncover the ultimate microbial diversity in the environment(Gibbons et al., 2013; J. A. Gilbert et al., 2012), also demonstrated the distinct ecological mechanisms underlying abundant and rare subcommunities. A recent study suggests that the compositional variations of different type of microbes were structured by different mechanisms due to different organismal body size(Luan et al., 2020). Here, the differed spatial scaling patterns and community assembly mechanisms between abundant and rare subcommunities should be due to their differed life strategies (e.g., adaptability to the environment).

In conclusion, this study investigated the ecological mechanisms driving spatial scaling patterns of sedimental microbial communities in a coastal sediment. Rare subcommunities were mainly responsible for the spatial scaling patterns followed by microbes. Distinct ecological mechanisms shaped the spatial scaling patterns of abundant and rare subcommunities. The results in this study are heuristic that different mechanisms may underlie the spatial/temporal patterns of microbes with different relative abundance. The study provided novel mechanistic insights into the spatial scaling patterns followed by different microbes.

Acknowledgment

This study was supported by the National Natural Science Foundation of China (92051110, 31971446), by National Key Research and Development Program of China (2019YFA0606700, 2020YFA0607600), by the Natural Science Foundation of Shandong Province (ZR2020YQ21), and by the Qilu Young Scholarship of Shandong University. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Conflict of interest: none declared.

Supplementary information is available via the publisher's website.

Data availability: sequencing data generated in this study is deposited at NCBI SRA portal under project ID PRJNA836637.

References

Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M., & Kadmon, R. (2012). Area-heterogeneity tradeoff and the diversity of ecological communities. *Proc Natl Acad Sci U S A, 109* (43), 17495-17500. doi:10.1073/pnas.1208652109

Bardgett, R. D., & van der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning.*Nature, 515* (7528), 505-511. doi:10.1038/nature13855

Barreto, D. P., Conrad, R., Klose, M., Claus, P., & Enrich-Prast, A. (2014). Distance-decay and taxa-area relationships for bacteria, archaea and methanogenic archaea in a tropical lake sediment. *PLoS One, 9* (10), e110128. doi:10.1371/journal.pone.0110128

Bickel, S., & Or, D. (2021). The chosen few-variations in common and rare soil bacteria across biomes.*ISME J, 15* (11), 3315-3325. doi:10.1038/s41396-021-00981-3

Bradford, M. A., Wood, S. A., Bardgett, R. D., Black, H. I., Bonkowski, M., Eggers, T., . . . Jones, T. H. (2014). Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *Proc Natl Acad Sci U S A, 111* (40), 14478-14483. doi:10.1073/pnas.1413707111

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data.*Nat Methods, 13* (7), 581-583. doi:10.1038/nmeth.3869

Chao, A., Chiu, C.-H., & Jost, L. (2014). Unifying Species Diversity, Phylogenetic Diversity, Functional Diversity, and Related Similarity and Differentiation Measures Through Hill Numbers. *Annual Review of Ecology, Evolution, and Systematics, 45* (1), 297-324. doi:10.1146/annurev-ecolsys-120213-091540

Chen, Q.-L., Ding, J., Zhu, D., Hu, H.-W., Delgado-Baquerizo, M., Ma, Y.-B., . . . Zhu, Y.-G. (2020). Rare microbial taxa as the major drivers of ecosystem multifunctionality in long-term fertilized soils. *Soil Biology and Biochemistry, 141* . doi:10.1016/j.soilbio.2019.107686

Chen, W., Jiao, S., Li, Q., Du, N., & Yang, G. (2019). Dispersal limitation relative to environmental filtering governs the vertical small-scale assembly of soil microbiomes during restoration. *Journal of Applied Ecology, 57* (2), 402-412. doi:10.1111/1365-2664.13533

Chiu, C. H., & Chao, A. (2014). Distance-based functional diversity measures and their decomposition: a framework based on Hill numbers. *PLoS One, 9* (7), e100014. doi:10.1371/journal.pone.0100014

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., . . . Tiedje, J. M. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res, 37* (Database issue), D141-145. doi:10.1093/nar/gkn879

Connor, E. F., & Mccoy, E. D. (1979). Statistics and Biology of the Species-Area Relationship.*American Naturalist, 113* (6), 791-833. doi:Doi 10.1086/283438

Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar, D., . . . Singh, B. K. (2016). Microbial diversity drives multifunctionality in terrestrial ecosystems. *Nat Commun, 7* , 10541. doi:10.1038/ncomms10541

Ellison, A. M. (2010). Partitioning diversity. *Ecology, 91* (7), 1962-1963. doi:10.1890/09-1692.1

Feinsinger, P., Spears, E. E., & Poole, R. W. (1981). A Simple Measure of Niche Breadth. *Ecology, 62* (1), 27-32. doi:Doi 10.2307/1936664

Feinstein, L. M., & Blackwood, C. B. (2012). Taxa-area relationship and neutral dynamics influence the diversity of fungal communities on senesced tree leaves. *Environ Microbiol, 14* (6), 1488-1499. doi:10.1111/j.1462-2920.2012.02737.x

Fuhrman, J. A., Cram, J. A., & Needham, D. M. (2015). Marine microbial community dynamics and their ecological interpretation. *Nat Rev Microbiol, 13* (3), 133-146. doi:10.1038/nrmicro3417

Gibbons, S. M., Caporaso, J. G., Pirrung, M., Field, D., Knight, R., & Gilbert, J. A. (2013). Evidence for a persistent microbial seed bank throughout the global ocean.*Proc Natl Acad Sci U S A, 110* (12), 4651-4655. doi:10.1073/pnas.1217767110

Gilbert, B., & Lechowicz, M. J. (2004). Neutrality, niches, and dispersal in a temperate forest understory. *Proc Natl Acad Sci U S A, 101* (20), 7651-7656. doi:10.1073/pnas.0400814101

Gilbert, J. A., Steele, J. A., Caporaso, J. G., Steinbruck, L., Reeder, J., Temperton, B., . . . Field, D. (2012). Defining seasonal marine microbial community dynamics.*ISME J, 6* (2), 298-308. doi:10.1038/ismej.2011.107

Green, J. L., Holmes, A. J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., . . . Beattie, A. J. (2004). Spatial scaling of microbial eukaryote diversity. *Nature, 432* (7018), 747-750. doi:10.1038/nature03034

Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G., & Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Applied and Environmental Microbiology, 66* (12), 5488-5491. doi:Doi 10.1128/Aem.66.12.5488-5491.2000

He, Z., Liu, D., Shi, Y., Wu, X., Dai, Y., Shang, Y., . . . Cui, Z. (2022). Broader environmental adaptation of rare rather than abundant bacteria in reforestation succession soil. *Sci Total Environ, 828* , 154364. doi:10.1016/j.scitotenv.2022.154364

Horner-Devine, M. C., Lage, M., Hughes, J. B., & Bohannan, B. J. (2004). A taxa-area relationship for bacteria. *Nature, 432* (7018), 750-753. doi:10.1038/nature03073

Horner-Devine, M. C., Lage, M., Hughes, J. B., & Bohannan, B. J. M. (2004). A taxa-area relationship for bacteria. *Nature, 432* (7018), 750-753. doi:10.1038/nature03073

Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography* . Princeton: Princeton University Press.

Huber, P., Metz, S., Unrein, F., Mayora, G., Sarmento, H., & Devercelli, M. (2020). Environmental heterogeneity determines the ecological processes that govern bacterial metacommunity assembly in a floodplain river system. *ISME J, 14* (12), 2951-2966. doi:10.1038/s41396-020-0723-2

Islam, M. S., Sarker, M. J., Yamamoto, T., Wahab, M. A., & Tanaka, M. (2004). Water and sediment quality, partial mass budget and effluent N loading in coastal brackishwater shrimp farms in Bangladesh. *Marine Pollution Bulletin, 48* (5-6), 471-485. doi:10.1016/j.marpolbul.2003.08.025

Jiang, Y., Lei, Y., Yang, Y., Korpelainen, H., Niinemets, Ü., & Li, C. (2018). Divergent assemblage patterns and driving forces for bacterial and fungal communities along a glacier forefield chronosequence. *Soil Biology and Biochemistry, 118* , 207-216. doi:10.1016/j.soilbio.2017.12.019

Jiao, S., Chen, W., & Wei, G. (2017). Biogeography and ecological diversity patterns of rare and abundant bacteria in oil-contaminated soils. *Mol Ecol, 26* (19), 5305-5317. doi:10.1111/mec.14218

Jiao, S., & Lu, Y. (2020). Abundant fungi adapt to broader environmental gradients than rare fungi in agricultural fields. *Glob Chang Biol, 26* (8), 4506-4520. doi:10.1111/gcb.15130

Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., . . . Webb, C. O. (2010). Picante: R tools for integrating phylogenies and ecology.*Bioinformatics, 26* (11), 1463-1464. doi:10.1093/bioinformatics/btq166

Lefcheck, J. S., Byrnes, J. E. K., Isbell, F., Gamfeldt, L., Griffin, J. N., Eisenhauer, N., . . . Duffy, J. E. (2015). Biodiversity enhances ecosystem multifunctionality across trophic levels and habitats. *Nat Commun, 6* , 6936. doi:10.1038/ncomms7936

Luan, L., Jiang, Y., Cheng, M., Dini-Andreote, F., Sui, Y., Xu, Q., . . . Sun, B. (2020). Organism body size structures the soil microbial and nematode community assembly at a continental and global scale. *Nat Commun, 11* (1), 6406. doi:10.1038/s41467-020-20271-4

Lynch, J. M., & Barbano, D. M. (1999). Kjeldahl nitrogen analysis as a reference method for protein determination in dairy products. *Journal of Aoac International, 82* (6), 1389-1398.

Lynch, M. D., & Neufeld, J. D. (2015). Ecology and exploration of the rare biosphere. *Nat Rev Microbiol, 13* (4), 217-229. doi:10.1038/nrmicro3400

Lyons, K. G., & Schwartz, M. W. (2001). Rare species loss alters ecosystem function - invasion resistance. *Ecology Letters, 4* (4), 358-365. doi:DOI 10.1046/j.1461-0248.2001.00235.x

Ma, Z. S. (2018). DAR (diversity-area relationship): Extending classic SAR (species-area relationship) for biodiversity and biogeography analyses. *Ecol Evol, 8* (20), 10023-10038. doi:10.1002/ece3.4425

Mo, Y., Zhang, W., Yang, J., Lin, Y., Yu, Z., & Lin, S. (2018). Biogeographic patterns of abundant and rare bacterioplankton in three subtropical bays resulting from selective and neutral processes. *ISME J, 12* (9), 2198-2210. doi:10.1038/s41396-018-0153-6

Mouillot, D., Bellwood, D. R., Baraloto, C., Chave, J., Galzin, R., Harmelin-Vivien, M., . . . Thuiller, W. (2013). Rare species support vulnerable functions in high-diversity ecosystems. *PLoS Biol, 11* (5), e1001569. doi:10.1371/journal.pbio.1001569

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography, 26* (4), 867-878. doi:DOI 10.1046/j.1365-2699.1999.00305.x

Ning, D., Yuan, M., Wu, L., Zhang, Y., Guo, X., Zhou, X., . . . Zhou, J. (2020). A quantitative framework reveals ecological drivers of grassland microbial community assembly in response to warming. *Nat Commun, 11* (1), 4717. doi:10.1038/s41467-020-18560-z

O'Brien, S. L., Gibbons, S. M., Owens, S. M., Hampton-Marcell, J., Johnston, E. R., Jastrow, J. D., . . . Antonopoulos, D. A. (2016). Spatial scale drives patterns in soil bacterial diversity. *Environ Microbiol, 18* (6), 2039-2051. doi:10.1111/1462-2920.13231

Pianka, E. R. (1974). Niche overlap and diffuse competition. *Proc Natl Acad Sci U S A, 71* (5), 2141-2145. doi:10.1073/pnas.71.5.2141

Rosenzweig, M. L. (1995). *Species diversity in space and time* . Cambridge ; New York: Cambridge University Press.

Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R., . . . Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci U S A, 103* (32), 12115-12120. doi:10.1073/pnas.0605127103

Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., . . . Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *ISME J, 7* (11), 2069-2079. doi:10.1038/ismej.2013.93

Stegen, J. C., Lin, X., Konopka, A. E., & Fredrickson, J. K. (2012). Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J, 6* (9), 1653-1664. doi:10.1038/ismej.2012.22

Stein, A., Gerstner, K., & Kreft, H. (2014). Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecol Lett, 17* (7), 866-880. doi:10.1111/ele.12277

Tilman, D. (1983). Some Thoughts on Resource Competition and Diversity in Plant-Communities. *Ecological Studies, 43* , 322-336.

Tu, Q., Deng, Y., Yan, Q., Shen, L., Lin, L., He, Z., . . . Zhou, J. (2016). Biogeographic patterns of soil diazotrophic communities across six forests in North America. *Mol Ecol, 25* (12), 2937-2948. doi:10.1111/mec.13651

van Schalkwyk, J., Pryke, J. S., & Samways, M. J. (2019). Contribution of common vs. rare species to species diversity patterns in conservation corridors. *Ecological Indicators, 104* , 279-288. doi:10.1016/j.ecolind.2019.05.014

Vellend, M. (2010). Conceptual synthesis in community ecology. *Q Rev Biol, 85* (2), 183-206. doi:10.1086/652373

Wan, W., Gadd, G. M., Yang, Y., Yuan, W., Gu, J., Ye, L., & Liu, W. (2021). Environmental adaptation is stronger for abundant rather than rare microorganisms in wetland soils from the Qinghai-Tibet Plateau. *Mol Ecol, 30* (10), 2390-2403. doi:10.1111/mec.15882

White, H. J., Pakeman, R. J., & Buckley, Y. M. (2022). Common species contribute little to spatial patterns of functional diversity across scales in coastal grasslands. *Journal of Ecology* . doi:10.1111/1365-2745.13858

Xiong, C., He, J. Z., Singh, B. K., Zhu, Y. G., Wang, J. T., Li, P. P., . . . Zhang, L. M. (2021). Rare taxa maintain the stability of crop mycobiomes and ecosystem functions. *Environ Microbiol, 23* (4), 1907-1924. doi:10.1111/1462-2920.15262

Xue, M., Guo, Z., Gu, X., Gao, H., Weng, S., Zhou, J., . . . Zhou, X. (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., & Yang, J. (2018). Distinct patterns and processes of abundant and rare eukaryotic plankton communities following a reservoir cyanobacterial bloom. *ISME J, 12* (9), 2263-2277. doi:10.1038/s41396-018-0159-0

Yang, Z., Liu, X., Zhou, M., Ai, D., Wang, G., Wang, Y., . . . Lundholm, J. T. (2015). The effect of environmental heterogeneity on species richness depends on community position along the environmental gradient. *Sci Rep, 5* , 15723. doi:10.1038/srep15723

Zhang, W., Pan, Y., Yang, J., Chen, H., Holohan, B., Vaudrey, J., . . . McManus, G. B. (2018). The diversity and biogeography of abundant and rare intertidal marine microeukaryotes explained by environment and dispersal limitation. *Environ Microbiol, 20* (2), 462-476. doi:10.1111/1462-2920.13916

Zhang, X., Liu, S., Wang, J., Huang, Y., Freedman, Z., Fu, S., . . . Schuler, J. (2020). Local community assembly mechanisms shape soil bacterial beta diversity patterns along a latitudinal gradient. *Nat Commun, 11* (1), 5428. doi:10.1038/s41467-020-19228-4

Zhou, J., & Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology?*Microbiol Mol Biol Rev, 81* (4). doi:10.1128/MMBR.00002-17

Zinger, L., Boetius, A., & Ramette, A. (2014). Bacterial taxa-area and distance-decay relationships in marine environments. *Mol Ecol, 23* (4), 954-964. doi:10.1111/mec.12640

Figure legends

Fig. 1 The composition and diversity metrics of the bacterial communities in the Beibu Gulf sediment. a Community composition at the phylum level; **b** Community richness (log transformed) for the whole community, the abundant and the rare subcommunities; **c** Community dissimilarity for the whole community, abundant and rare subcommunities. The Bray-Curtis dissimilarity was calculated and used for community dissimilarity.

Fig. 2 Spatial scaling patterns followed by sedimental microbes in Beibu Gulf. a Taxa-area relationship (TAR) for the whole community, the abundant and the rare subcommunitiesm, by investigating the relationship between log-transformed richness and area; **b**Distance decay relationship (DDR) for the whole community, the abundant and the rare subcommunities, by investigating the relationship between log-transformed community similarity (1-Bray-Curthis dissimilarity) and geographic distance; **c** The slope coefficient (z-value) of the diversity-area relationships (DAR_q) with different diversity orders (q) by extending TAR to DAR_q and the slope coefficient (d -value) of the distance decay relationship (DDR_q) with diversity order (q) by extending DDR to DDR_q.

Fig. 3 Linking microbial spatial scaling diversity metrics with environmental heterogeneity. a The association between environmental heterogeneity and differed species richness of the whole community, the abundant subcommunities and the rare subcommunities. The number of unique taxa in two sample pairs were calculated for differed richness. **b** Association between community similarity and environmental heterogeneity. The community similarity was calculated as 1-Bray-Curtis dissimilarity. The Euclidean distance based on normalized environmental variables was calculated to represent the environmental heterogeneity between different samples.

Fig. 4 Properties of microbial taxa at the levels of whole community, abundant subcommunities, and rare subcommunities. a Percentage of microbial ASVs mapped to the deep sequencing dataset;**b** Percentage of reads mapped to the deep sequencing dataset;**c** Niche breadth of microbial ASVs; **d** Niche overlap of microbial ASVs. The Levin's and Pianka's indices were respectively calculated for niche breadth and niche overlap.

Fig. 5 Local community mechanisms driving the compositional variations of sedimental microbial communities in the Beibu Gulf. Local community mechanisms were quantified for the whole community (**a**), the abundant subcommunities (**b**), and the rare subcommunities (**c**). The contribution of five different ecological processes, including homogeneous selection (HoS), heterogeneous selection (HeS), drift (DF), dispersal limitation (DL), and homogeneous dispersal (HD), were quantified.