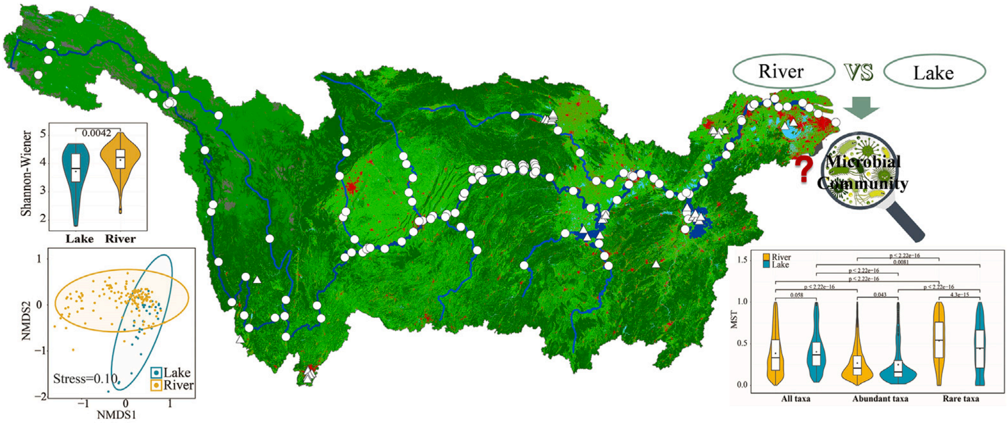
iScience

CellPress OPEN ACCESS

Article

Microbiome analysis in Asia's largest watershed reveals inconsistent biogeographic pattern and microbial assembly mechanisms in river and lake systems



Xi Liu, Lu Zhang, Yingcai Wang, ..., Yuxin Hu, Huaiying Yao, Zhi Wang

ssshyx@163.com (Y.H.)
zwang@apm.ac.cn (Z.W.)

Highlights

Deterministic process dominated microbial community assemble in watershed

Stochastic process contributed more to microbes in rivers than in lakes

Environmental variations mainly shaped the microbial biogeography in the lake

The effects of land use on microbes were stronger in lakes than in rivers

Liu et al., iScience 27, 110053
June 21, 2024 © 2024 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.isci.2024.110053

iScience

CellPress
OPEN ACCESS

Article

Microbiome analysis in Asia's largest watershed reveals inconsistent biogeographic pattern and microbial assembly mechanisms in river and lake systems

Xi Liu,[1,5] Lu Zhang,[2,4,5] Yingcai Wang,[1] Sheng Hu,[1] Jing Zhang,[1] Xiaolong Huang,[1] Ruiwen Li,[1] Yuxin Hu,[1,*]
Huaiying Yao,[3] and Zhi Wang[2,6,*]

SUMMARY

Microorganisms are critical to the stability of aquatic environments, and understanding the ecological mechanisms of microbial community is essential. However, the distinctions and linkages across biogeographic patterns, ecological processes, and formation mechanisms of microbes in rivers and lakes remain unknown. Accordingly, microbiome-centric analysis was conducted in rivers and lakes in the Yangtze River watershed. Results revealed significant differences in the structure and diversity of microbial communities between rivers and lakes, with rivers showing higher diversity. Lakes exhibited lower community stability, despite higher species interactions. Although deterministic processes dominated microbial community assembly both in rivers and lakes, higher stochastic processes of rare and abundant taxa exhibited in rivers. Spatial factors influenced river microbial community, while environmental factors drove differences in the lake bacterial community. This study deepened the understanding of microbial biogeography and formation mechanisms in large watershed rivers and lakes, highlighting distinct community aggregation patterns between river and lake microorganisms.

INTRODUCTION

Rivers and lakes are vital components of the hydrological cycle and essential sources of fresh water, providing resources for human industrial and agricultural production and maintaining ecosystem functions.[1,2] However, these aquatic ecosystems have been severely degraded by the increased disturbance from anthropogenic activities and global environmental changes.[2,3] As the functional basis of aquatic ecosystems, microbes participate in maintaining the stability of the structure and function of aquatic ecosystems.[4,5] The microbial communities in the river-lake ecosystem have been widely studied worldwide.[6] Many independent studies on aquatic microbes have been carried out in river watersheds, lakes, and freshwater reservoirs.[7–12] However, although some studies have been conducted on microbes in rivers and lakes, limited comparative studies have focused on a complete watershed, especially the river and lake systems of one of the largest watersheds in the world. It is unclear whether there are differences in the biogeographic patterns, mechanisms of community construction, and responses to anthropogenic activities of microbes between rivers and lakes across a wide range of environmental gradients and geographic spatial differences on a complete large watershed.

In response to different stresses of environmental changes or geographical patterns, microbes often form specific community structures in response to various influences. The cooperation and competition between microbes form a complex network of microbial ecological relationships to maintain the stability, balance, and biogeochemical function of the ecosystem.[13] Deterministic (niche based) and stochastic (neutral) processes are closely related yet antagonistic processes that account for microbial community variation that reflects their relative importance in shaping microbial community collections.[14] The understanding of microbial interactions and community assembly mechanisms in natural environments is crucial for understanding microbial community structure and functional stability. Although rivers and lakes are connected aquatic environments, river-lake ecosystems have distinct differences that are affected by

---

[1] Ecological Environment Monitoring and Scientific Research Center, Yangtze River Basin Ecological Environment Supervision and Administration Bureau, Ministry of Ecological Environment, Wuhan 430010, Hubei, China
[2] Key Laboratory for Environment and Disaster Monitoring and Evaluation of Hubei, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430077, Hubei, China
[3] Research Center for Environmental Ecology and Engineering, School of Environmental Ecology and Biological Engineering, Wuhan Institute of Technology, Wuhan 430205, Hubei, China
[4] University of Chinese Academy of Sciences, Beijing 100049, China
[5] These authors contributed equally
[6] Lead contact
*Correspondence: ssshyx@163.com (Y.H.), zwang@apm.ac.cn (Z.W.)
https://doi.org/10.1016/j.isci.2024.110053

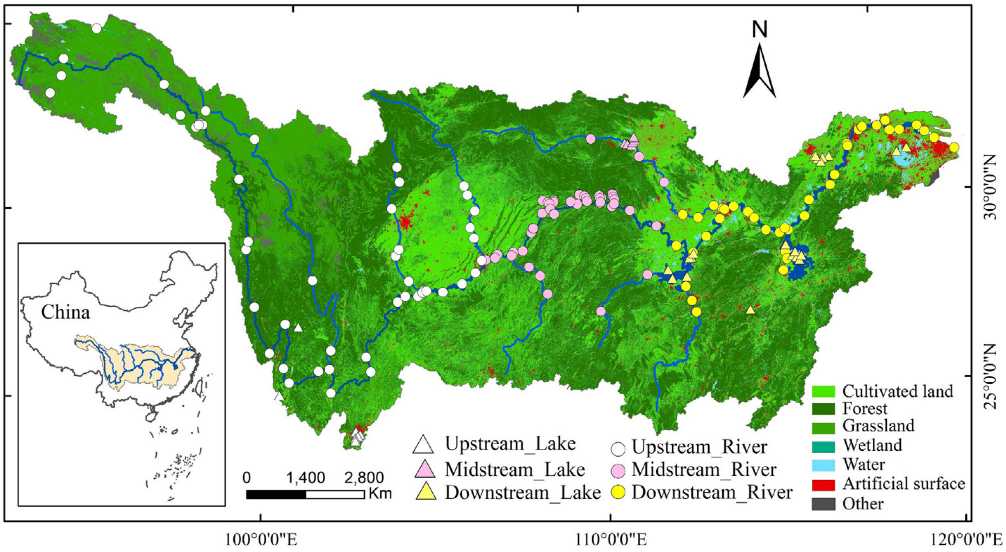


Figure 1. Location of water sampling sites and land-use types in Yangtze River Basin

hydrological conditions, environmental physicochemical indexes, and other factors. However, considering the limited comparative studies of rivers and lakes within the watershed, differences in microbial interactions and formation mechanisms have not been determined in these interconnected but distinct habitats. Moreover, land use, as a macroscopic manifestation of human activity on the earth's surface, directly or indirectly affect the microbial community in watershed.[15–17] Although land use plays an important role in the deterministic process of microbial community construction. However, it is unclear how land use affects the construction of microbial communities and maintains community stability.

Therefore, in the present study, the whole Yangtze River watershed served as the research framework, facilitating an enhanced comprehension of microbial characteristics and driving mechanisms within river-lake ecosystems, particularly in response to land use. The Yangtze River, with a length of 6,300 km, stands as the largest river in Asia and the third-longest globally. The Yangtze River basin is rich in river system and has many important lakes and reservoirs, such as Taihu Lake, Poyang Lake, and Danjiangkou Reservoir. The watershed, supporting a population exceeding 400 million, has experienced sustained human activities, influencing diverse land use patterns throughout history, rendering it an optimal study locale. We hypothesized that (1) microbial diversity would be higher in rivers than in lakes due to spatial heterogeneity and closer links with terrestrial ecosystems, (2) environmental factors and land use generated by human activities have a greater impact on microorganisms in lakes than that in rivers, and (3) the construction of river microbial communities is more stochastically due to the more stable hydrological environment of the lake. This study contributes a comprehensive characterization, geographic pattern, and interconnection of microbes in river and lake ecosystems, elucidating symbiotic patterns, ecological processes, and the driving mechanisms shaping microbial community structures under the influence of environmental factors and land use. These results could reveal the biogeographic pattern and formation mechanism of microbes in rivers and lakes of large catchments, which may aid in the effective management of watershed aquatic ecology during the intensification of anthropogenic activities.

RESULTS

Land use patterns and water quality parameters

The Yangtze River Basin has seven land use types, namely, cultivated land, forest, grassland, wetland, residence, water, and unused land. Based on the composition and distribution of land use (Figure 1), the Yangtze River Basin is dominated by forest (43.76%), followed by cultivated land (28.09%) and grassland (21.0%), in which wetland and water body account for 2.61%. Residence has higher proportion in the midstream and downstream of the watershed, totaling 3.06%. In this study, the effects of land use on microbial communities within different buffer scales (500 m, 1,000 m, 1,500 m, 2,000 m, 2,500 m, and 5,000 m) were explored, and the redundancy analysis (RDA) results showed that land use exerted the most substantial explanatory influence on the microbial community within the 2,500-m buffer zone, both in the river and within the lake (Table S1). For optimal analysis, land use parameters within the 2,500-m buffer zone, demonstrating the greatest impact on the bacterial community, were selected for subsequent analysis. Differences were observed in the water quality parameters between rivers and lakes in watershed (Table S2). The physicochemical parameters of NH_4^+-N, COD_{Mn}, WT, pH, and Chl.a were significantly higher in lake than that in river, while conductivity was significantly higher in river than that in lake ($p < 0.05$). In addition, a significant shift was observed in the physicochemical parameters along the upstream to downstream of the watershed (Table S3).

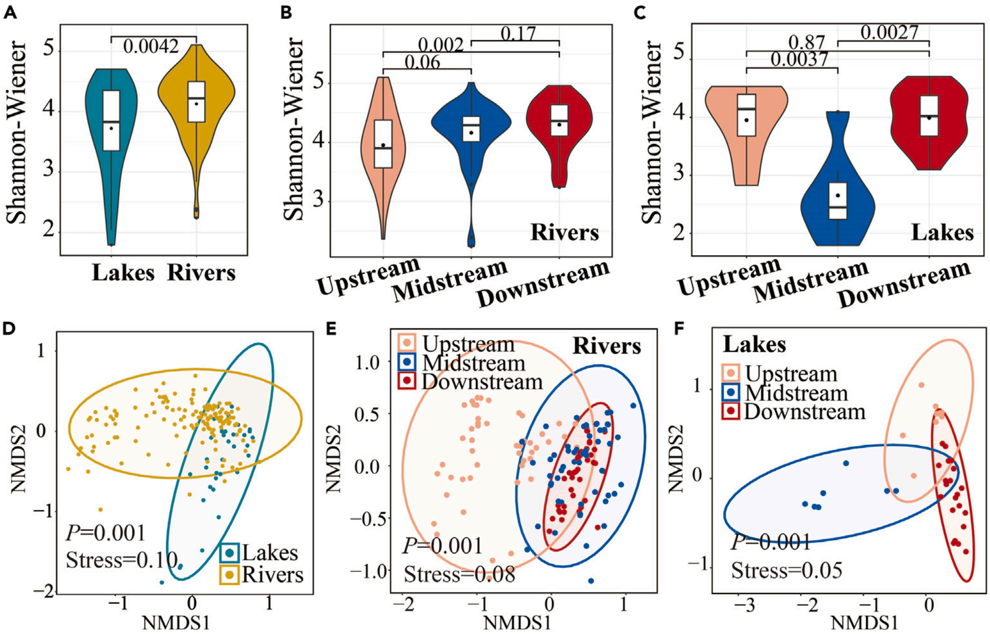


Figure 2. Characteristics of bacterial communities in Yangtze River Basin

(A) Shannon index of bacterial communities in rivers and lakes.

Comparison of Shannon indices among upstream, midstream, and downstream in rivers (B) and lakes (C).

(D) Differences in the bacterial OTUs of rivers and lakes. Differences in the bacterial OTUs among upstream, midstream, and downstream in rivers (E) and lakes (F).

Characterization of microbial community in rivers and lakes

At the phyla level, Proteobacteria, Actinobacteria, and Bacteroidetes were the dominant phyla both in rivers and lakes (Figure S1A). In total, 1,042 OTUs (82.7%) were shared in rivers and lakes, while the shared rare OTUs (170) were more than abundant OTUs (84) between rivers and lakes (Figure S1B). Although the dominant phyla and OTUs of rivers and lakes were similar, NMDS analysis showed obvious differences in microbial community between the rivers and lakes (Figure 2D). Moreover, the Shannon index of rivers (4.13) was significantly higher than that in the lakes (3.72) (Figure 2A, $p < 0.05$). A significant shift was observed in the bacterial community compositions along the upstream to downstream of the watershed (Figures 2B, 2C, 2E, 2F, and S1E–S1H). The NMDS showed a strong and significant difference of OTUs among the upstream, midstream, and downstream of the watershed both in rivers and lakes (Figures 2E, 2F, and S1E–S1H). The Shannon index was highest downstream of the watershed both in rivers and lakes (Figures 2B and 2C; $p < 0.05$). More shared microbes were observed on rare taxa than that of abundant taxa among the upstream, midstream, and downstream of rivers, while the opposite was true in lakes (Figures S1C and S1D).

Ecological processes of microbial community

The index of modified stochasticity ratio (MST) was used to evaluate the ecological process of bacterial community. Deterministic process dominated the whole bacterial community assembly in rivers (61%) and lakes (60%, Figure 3A). Although no significant difference was observed between the contribution of stochastic processes to the assembly of whole bacterial community in rivers (39%) and lakes (40%), the contributions of stochastic processes to the assembly of abundant and rare OTUs were significantly higher in river (28% and 54%) than that in lake (25% and 44%, $p < 0.05$, Figure 3A). Moreover, the contribution of deterministic processes to the assembly of abundant taxa were higher than that in whole bacterial community, while the contribution of stochastic processes to the assembly of rare taxa were higher than that in whole bacterial community both in river and lake ($p < 0.05$, Figure 3A).

Spatial differences were observed in the ecological processes of bacterial communities (Figures 3B–3G). The dominant ecological processes of whole bacterial community shifted from deterministic to stochastic along the upstream to downstream in the river, while the dominant ecological processes shifted from stochastic to deterministic along the upstream to downstream in the lake (Figures 3C, 3D, 3F, and 3H). Interestingly, the contribution of deterministic process contribution of abundant OTUs was the highest, while the stochastic process contribution of rare OTUs was highest both in downstream of rivers and lakes ($p < 0.05$, Figures 3D and 3E).

Patterns of co-occurrence among microbial communities in rivers and lakes

Potential microbial interactions were identified by constructing the co-occurrence networks (Figure 4). The multiple topological properties of networks indicated that microbial co-occurrence patterns differed in rivers and lakes, as well as from upstream to downstream both in rivers and lakes (Figure 4 and Table S4). The number of total nodes, total links, average degree, and modularity of lake networks were higher than

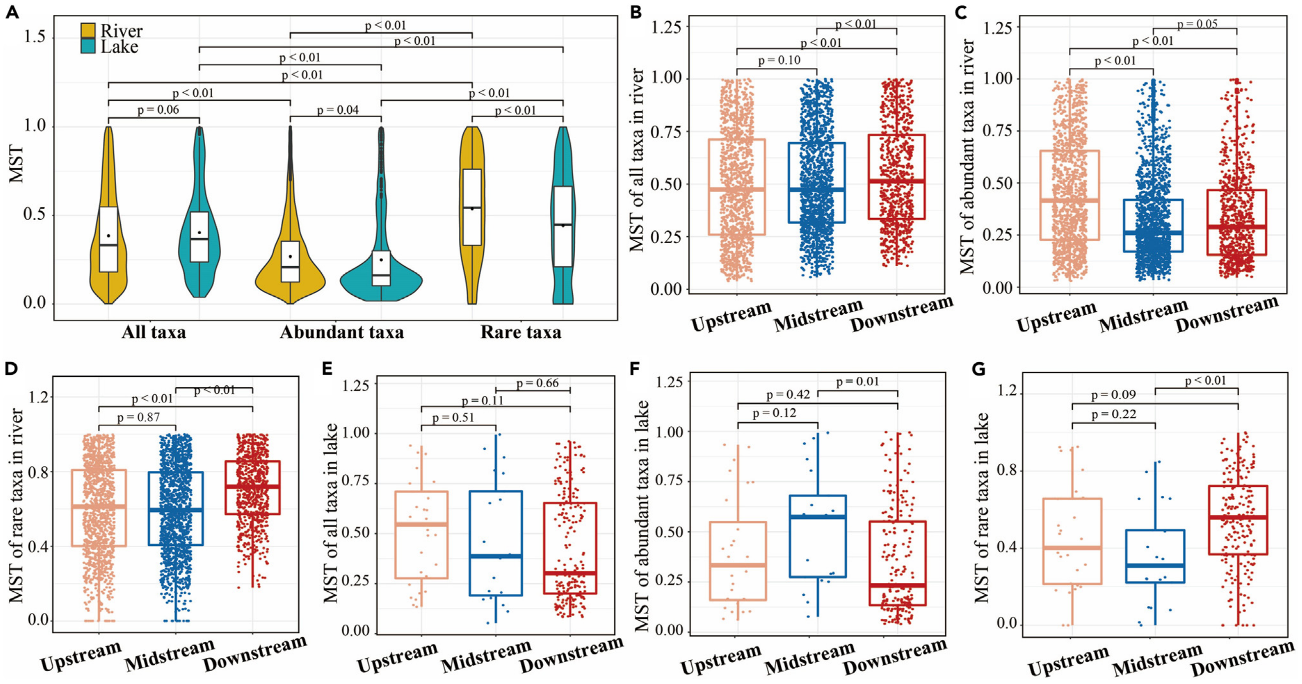


Figure 3. Ecological process of bacterial OTUs

(A) Stochastic process contribution in constructing the bacterial OTUs in rivers and lakes. Stochastic process contribution in constructing the bacterial OTUs (B), the abundant taxa (C) and rare taxa (D) in upstream, midstream, and downstream of rivers. Stochastic process contribution in constructing the bacterial OTUs (E), the abundant taxa (F), and rare taxa (G) in upstream, midstream, and downstream of lakes.

those in river (Table S4). The positive relationship between microbes was much higher than the negative relationship both in rivers and lakes (Figure 4). The keystone taxa in rivers and lakes were identified from networks based on their topological roles (Figure S2). Although the keystone taxa of rivers and lakes were different, most of them belonged to Proteobacteria. Cohesion indexes were calculated to evaluate community stability (Figure S3). The index of N:P cohesion was used to determine community stability, and the higher the value, the higher the community stability. The N:P cohesion value in the river network (0.52) was higher than that in the lake network (0.43) ($p < 0.05$), indicating that microbial community in rivers is more stable (Figure S3A). Moreover, the N:P cohesion value varied along the upstream to downstream of the watershed (Figure S3B). The N:P cohesion value suggested that the microbial communities were the most stable in downstream of rivers (0.65), while the highest stability in lakes was observed midstream (0.47).

Formation mechanism of microbial community

Effect of physiochemical factors and spatial factors on microbial community

For a better understanding of the relationships between geographic, elevation, and environmental distances and bacterial community structure, the linear regressions of similarities with those distances were analyzed (Figure 5). Elevation and geographical distances were more closely related to whole bacterial OTUs and abundant OTUs than rare OTUs both in rivers and lakes (Figures 5A and 5B). Notably, the relationship between elevation and geographical distance and bacterial community was higher in rivers than that in lakes (Figures 5A and 5B). However, the relationship between environmental distance and bacterial community was higher in lakes than that in rivers (Figures 5A and 5B). Linear regression analysis showed that elevation, geographic, and environmental distances were significantly correlated with MST in whole bacterial OTUs ($p < 0.05$, Figures 5C and 5D). The relationship between elevation distance and geographic distance and MST was higher in rivers (Figures 5C and 5D). Moreover, the relationship between community similarity of abundant taxa and elevation and geographical distance was higher than that of rare taxa both in rivers and lakes (Figures 5A and 5B). However, the relationship between the community similarity of rare taxa and environmental distance was stronger than that of abundant taxa in rivers, while the opposite was true in lakes (Figures 5A and 5B).

Coupling effects of environmental factors on microbial community

Partial least squares path modeling analysis (PLSPMA) was used to comprehensively understand the mechanism maintaining microbial community stability in river and lake ecosystems (Figure 6). In either river or lake ecosystem, land use played a direct and strong role in water nutrition and water property (Figures 6A and 6C). However, land use had a stronger effect on keystone taxa in river ecosystem, and it played a stronger role in microbial diversity in lake ecosystem (Figures 6A and 6C). Cultivated land, residence, and grass remarkably affected the

Figure 4. Co-occurrence network patterns of bacterial communities

Network of bacterial communities in rivers (A) and lakes (E). Network of bacterial communities in the upstream (B), midstream (C) and downstream (D) of rivers. Network of bacterial communities in the upstream (F), midstream (G) and downstream (H) of lakes. Edges represent interactions between nodes, pink represents positive correlation, and green represents negative correlation. Only OTUs with an average relative abundance >0.01% were used in network analysis.

microbes both in lakes and rivers (Figures 6A and 6C), and differences were observed in the bacterial community of samples dominated different land use types (Figure S4). *Limnohabitans, Acinetobacter*; *Polynucleobacter*, and *Pseudarcicella* were the dominant genera in the samples dominated by grass. However, the dominant genera changed to *CL500-29 marine group* (related to eutrophication), *Curvibacter* (related to denitrification), and *hgcI_clade* (related to nitrogen cycle) in the samples dominated by residence and cultivated land. In addition, the effects of land use on bacterial community in lakes (0.45) was stronger than that in rivers (−0.4) (Figures 6B and 6D). Water nutrient was likely to affect microbial diversity, and the increase in nutrient concentrations could promote microbial diversity in both rivers and lakes. However, water quality was likely to affect keystone taxa in river ecosystem, indicating that environmental indicators such as water temperature and pH play an important role in shaping microbial niches (Figures 6A and 6C). However, water nutrient and property strongly affected both microbial diversity and the keystone taxa of lake system (Figures 6A and 6C). Notably, the stability of microbial community in river ecosystem was more directly affected by keystone taxa, and it was more directly and heavily affected by biodiversity in lake ecosystem (Figures 6A and 6C).

DISCUSSION

The number of studies based on high-throughput sequencing technology has increased rapidly, revealing the great diversity of bacterial communities in aquatic environments. However, most studies that investigated the variability and diversity of bacterial communities in aquatic ecosystem have focused on small-scale spatial dimensions, such as short rivers or independent lakes.[1,3,11,12,18] The distinctions and linkage across bacterial profiles, biogeographic patterns, ecological processes, and formation mechanisms of microbes in connected rivers and lakes of a large watershed and the mechanism of their response to land use remain unknown. In the present study, the spatial patterns, ecological processes, and driving mechanisms of microorganism in lakes and rivers (∼6,300 km) were observed within a large watershed. Results show that changes in environmental factors dominated by anthropogenic drive the composition of lakes microbial community structure, while changes in spatial pattern directly caused differences in the rivers microbial community. This study provides a deep understanding of the biogeographic pattern and formation mechanism of microbes in the rivers and lakes of large watershed, lending ways for the effective management of watershed aquatic ecology.

Spatial variation and linkage of microbial communities between rivers and lakes

River–lake systems are closely connected aquatic ecosystems that are affected by both natural and anthropogenic conditions. Rivers are essential channels for the migration of terrigenous materials to lakes and play a fundamental role in controlling the health of lakes,

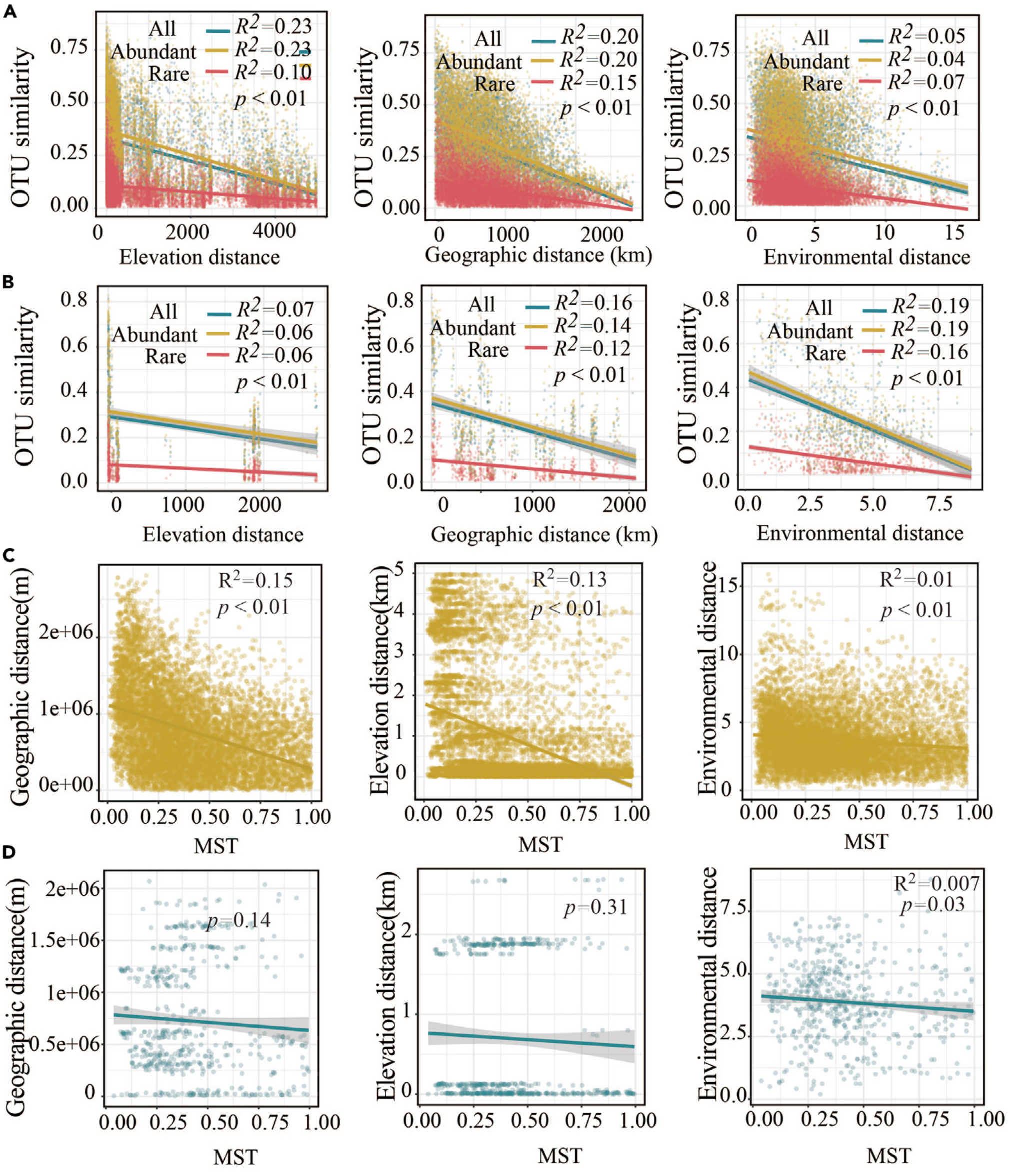


Figure 5. Relationship between microbial community and environmental and spatial factors

Patterns of correlations between OTUs similarity and the elevation, geographic, and environmental distance in rivers (A). Patterns of correlations between OTUs similarity and the elevation, geographic, and environmental distance in lakes (B). Patterns of correlations between MST and the geographic, elevation, and environmental distance in rivers (C). The patterns of correlations between MST and the geographic, elevation, and environmental distance in lakes (D). OTUs similarity was calculated using the Bray-Curtis distance between samples.

consequently, microbes are shared between rivers and lakes.[10] However, lakes and rivers are generally subject to different environmental conditions, such as velocity, water retention time, organic matter, and nutrient content, thus affecting the composition and function of microbial communities, resulting in differences in the structure of microbial communities between rivers and lakes.[19] Crevecoeur et al.[19] confirmed that lakes act as funnels, which selectively retain and exclude exotic taxa, and the environmental filters and other processes that form lake microbial communities operate on an extremely small spatial scale. Microbial diversity was significantly higher in rivers than that in lakes, thus supporting the patterns of microbial community diversity in the river-lake continuum in Bosten Lake watershed, northwest China.[20] This finding was obtained possibly because microbes in terrestrial soils (which contain highly diverse microbes) are likely to enter rivers.[21–23] Moreover, the coupling between the rivers microbes and the surrounding regional microbes is more frequent than that in lakes habitats.[20] The spatial heterogeneity of the rivers environment result in different ecological niches for different microbes to inhabit.[20] The results of PLSPMA analysis also showed that water temperature, pH, Chl.a, conductivity, COD_{Mn}, and other environmental factors are important driving factors of microbial diversity (Figure 6).

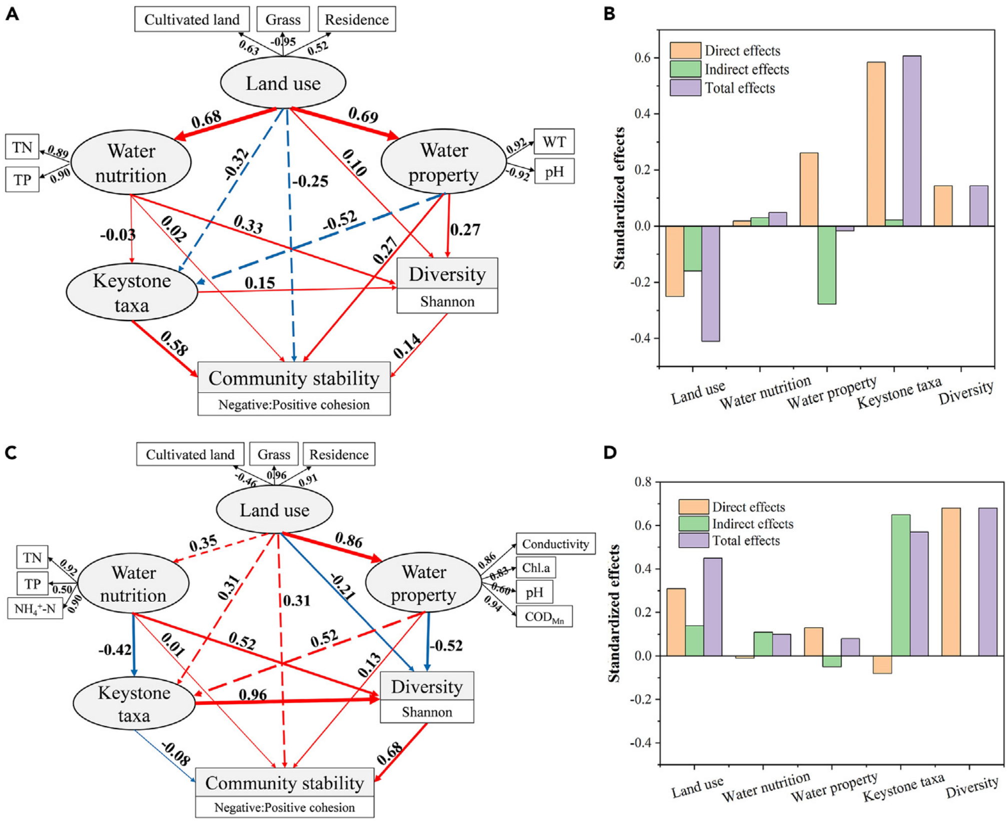


Figure 6. Driving factors of microbial community

The PLSPMA model was used to reveal the relationships among land use, water nutrition, and water property on microbial community in (A) rivers and (C) lakes. The bar chart shows the effects of the latent variables on the microbial community in (B) rivers and (D) lakes. The arrow indicated the direction of the effect between the variables, and the path coefficient indicates the strength of effect. The linewidth is proportional to the strength of effect.

Spatially, the diversity of microbes was higher in the midstream and downstream of rivers, which was consistent with the spatial changes of TN and TP concentrations. Nutrient concentration enhanced the diversity of microbial community, thus supporting that greater diversity and abundance are observed among bacterial communities at higher TN levels.[24] Increased biodiversity provides the tolerance of microbes to environmental stress.[25,26] Thus, high TN concentrations in the midstream and downstream promote bacterial diversity in rivers. Similarly, no significant difference was observed in nutrient concentration and microbial diversity between lakes in upstream and downstream of the watershed. In addition, the diversity of bacterial community is correlated with the hydraulic retention time affected by water transfer.[27] The flowing water of the torrent system apparently transfers microbials relatively quickly from the upstream habitat to the downstream habitat. Moreover, as a widely studied alpha diversity pattern,[28] species diversity decreases with increasing latitude.[29] However, in the present study, the microbial diversity of the lakes did not increase from the upstream to downstream of watershed. Although the lake was connected with the rivers, the independence between the lakes was strong, and the lakes in upstream can hardly interact hydrologically with the lakes downstream. Therefore, compared with rivers, the microbes of lakes in the upstream, midstream, and downstream of watershed had more obvious differences in community structure and shared fewer microbes.

Driving factors and ecological mechanism of microbial communities

Defining the ecological mechanisms underlying the formation of microbial community profiles is a central challenge in microbial ecology.[18,30,31] Neutral theory suggests that microbial communities are formed by random fluctuations and limited dispersal of species abundance (births and deaths), while niche theory suggests that microbial communities are formed by deterministic processes (such as habitat heterogeneity or species sorting) caused by differences in habitat preferences and species fitness.[30,32] Both deterministic and stochastic processes regulate microbial community assembly.[33] In the present study, the whole microbial communities were dominated by deterministic processes both in rivers and lakes considering the wide span of the watershed, which covers an area of 1.8 million square kilometers, complex and diverse human activities, and distinct geology, geomorphology, climate, and environment. These changes in the external environment severely affected the shaping of microbial community structure. Elevation, geographical and environmental distance were substantially correlated with MST. Spatially, the dominant ecological process of microbial community shifted from deterministic to stochastic in rivers from upstream to downstream, while the opposite was observed in the lakes along the upstream to the downstream of the watershed. The spatial variation of microbial ecological processes was consistent with the nutrient concentrations, presumably resulting from the tendency of microbes to form their communities by deterministic process in nutrient-poor conditions.[34,35]

Furthermore, the microbial community structure is distinguished by a markedly skewed rank-abundance distribution, featuring a limited number of dominant taxa and extensive tails comprising rare bacterial species.[19] The contribution of abundant and rare taxa to the structural stability, function, and assembly of microbial communities is a key issue in microbial ecology.[19] In the present study, the ecological processes of abundant taxa were dominated by deterministic processes, while the community construction of rare taxa showed strong stochastic process. Rare taxa are limited by habitat specificity and may thus exhibit stronger stochastic dispersal than abundant taxa, use a wide range of resources, and adapt effectively to the environment.[36,37] This finding is supported by the fact that more rare taxa than abundant taxa are shared among the upstream, midstream and downstream of rivers, that is, abundant taxa are more constrained by dispersal than rare taxa, which is consistent with previous studies in the Northwest Pacific.[38] However, some studies have found that dispersal remarkably affects rare taxa, but not on abundant taxa.[39,40] These differences can be attributed to differences in habitat and geography.[41–43] The flow is directed, and microbes in the upstream of the rivers can be transferred downstream by passive diffusion, while those in downstream are less likely to be transferred upstream. Thus, although dispersal can be deterministic, stochastic, or both,[14] the different dispersal limits of microbial communities in connected rivers in the present study reflect stochastic processes.

For both rare and abundant taxa, the stochastic process has a strong effect on microbes in rivers, thus supporting that microbes are strongly influenced by environmental filters in lakes (Figures 5 and 6). The river environment is relatively continuous, and the bacterial community in the water can spread from upstream to downstream easily. However, as relatively independent individuals, the heterogeneity of nutrient conditions in different lakes is obvious. Bacterial communities spread more strongly in rivers than in lakes. Moreover, considering the limitation caused by perturbation processes, microbes may be more strongly influenced by competition and selection processes in lakes.[19]

In addition, distance decay analysis results showed that geographical and environmental distances are important in constructing the spatial similarity of bacterial communities in rivers and lakes (Figures 5 and 6). Notably, spatial factors more remarkably affect the variation in fluvial microbial communities in rivers (including total, abundant and rare taxa) than that in lakes (Figures 5 and 6). In lakes, environmental factors are substantial factors in explaining the differences in bacterial community assemblages (Figures 5 and 6). In the present study, the bacterial community in the rivers is mainly affected by topography and elevation but is less affected by local environmental parameters. This phenomenon was observed possibly because rivers are large dynamic rapids systems. Microbes in the upstream of the rivers can easily spread downstream with the flow of water. Therefore, microbes can easily migrate through rivers as the water flows from upstream to downstream.[20] However, the effect of environmental selection on microbial community in lakes is more intense than that in rivers, mainly because lakes are relatively closed and stable aquatic systems.[20] When foreign microbes enter the lake environment, the arriving microbes become more strongly selected by physical and chemical conditions.[19] Some of the surviving alien microbes arrive in the lakes and are then selected based on the physical and chemical conditions of the lakes.[19] Thus, in this slowly changing environment, the selection of physicochemical factors attenuates the influence of spatial factors, such as geographic distance, on lake microbial community structure. Therefore, the spatial effect of microbial communities in rivers is considerably stronger than that in lakes, while microbial communities are more susceptible to local environmental factors.

Effects of land use on microbial communities

Land use is a spatial reflection of human activities and an important anthropogenic driving force that affects the change in physical and chemical parameters in the watershed.[44] Many studies have investigated the effects of land use on microbial production in aquatic environments.[1,45,46] In the present study, land use indirectly affected microbial community diversity and stability by influencing physicochemical parameters, and even directly and strongly affected microbial keystone species, especially cultivated land and residence that are strongly related to human activities (Figure 6). Pollutants from anthropogenic point source inputs (e.g., urban runoff and sewage pollutants) and agricultural non-point source pollution enter rivers and lakes through surface runoff, thus remarkably reducing surface water quality and affecting microbial community.[2] In addition, remarkable differences were observed in the microbial community structure in aquatic environments dominated by different land use types. For example, high levels of urban activity (measured as the percentage of landscape devoted to residence) may shift microbial communities toward greater denitrification and anammox activity.[47] Notably, the effect of land use on microbes in lakes is much greater than that in rivers, confirming that microbes in lake are more susceptible to local environmental influences than those in rivers. Considering that lakes act as "sink" of pollutants and microbes of the watershed, their hydrological environment is stable. Both native and foreign microbes are more susceptible to the physical and chemical properties of lakes in such a slowly changing and relatively stable environment.

Stability of microbial community in rivers and lakes

Many studies have generally focused on the abundant and rare taxa in microbial communities, while network analysis has shown that specific microbes play unique and important roles in community structure and function.[48,49] Keystone species plays an important role in microbial community stability in river and strongly affects microbial community diversity in lakes (Figure 6). Our study showed that only a few of these keystone species belong to rich species, and most of them belong to intermediate species. Keystone species are important for the stability of microbial communities, and the disappearance of these keystone species may lead to fragmentation of modules and networks, making intermediate species important for maintaining ecosystem stability.[13] The co-occurrence network results showed that the interactions between microbes are dominated by positive connections, mainly because the positive interaction is mainly considered as cooperation.[50] Microbial interactions can support ecosystem function and stability. Microbial cooperation may contribute to the resilience of plankton communities under changing environments, because microbial interaction networks can provide a buffer against environmental disturbances.[51]

Limitations of the study

It must be recognized that this study still has limitations. First, this study was based on one sampling, the seasonal variation of microbes of rivers and lakes in the large-scale watershed is not clear. However, sampling was conducted during the warm rainy season (June – October), and all areas within the watershed were in the local hot season during this period. Therefore, we assumed that the spatial differences in microbes in the watershed due to seasonal variations could be neglected. In addition, the Yangtze River is 6,300 km long and it flows through a basin with an area of 1.8 million km^2, and there is a great variation in microbial habitats within the watershed, with large differences in gradients of environmental and climatic factors (temperature, elevation, nutrients, hydrodynamics, etc.). Thus, we believe that the results based on a single sample can characterize the spatial distribution of microbes and identify the driver factors in a large watershed where environmental and climatic factors vary considerably. Second, the presence of many dams in the mainstem and tributaries of the Yangtze River affects the hydrodynamic conditions. The presence of a dam may indeed affect microbial communities. However, none of the sampling sites were in lakes/reservoirs formed by dams and all of the sampling sites are far away from dams. In addition, rivers still have hydrological connectivity and microbes in rivers still can spread with water flow. Therefore, as this study focuses on the spatial patterns of microbial distribution characteristics, our results can explain the spatial characteristics and differences of microbes in rivers and lakes in a large-scale watershed. This seasonal variation of microbes and the effects of dams on microbial communities within a large-scale watershed need further studies.

Conclusion

This study comprehensively investigated the distinctive patterns and assembly mechanisms governing microbial communities within river and lake ecosystems across a relatively large-scale watershed. The examination involved microbial community profiles, land use patterns, and environmental factors. Many microbes are shared in this closely related river-lake aquatic systems. However, considering the influence of hydrological conditions and environmental factors, the microbial community structure in rivers and lakes remarkably differed, and the microbial diversity in rivers was significantly higher than that in lakes. Land use strongly influenced microbial community structure and assemblage processes, either directly or indirectly (by physicochemical factors). The effect of land use on lake microbes was stronger than that in river microbes. In addition, spatial factors remarkably affected the changes in river microbial communities. However, environmental factors are important in explaining differences in bacterial community assemblages in lakes. Although deterministic processes dominated the microbial community assembly mechanism both in rivers and lakes, the higher stochastic processes of rare and abundant taxa were exhibited in rivers than in lakes. Microbial keystone species play an important role in the stability and diversity of microbial communities. This study deepens our understanding of the biogeographic pattern and formation mechanism of microbes in the rivers and lakes of large-scale watershed and provides a reliable basis for the effective management of watershed water ecology and land use planning under the influence of intensive anthropogenic activities. For example, microbes in lakes are more likely to be affected by environmental factors, especially land use that reflects human activities. Therefore, establishing ecological buffer zones around lakes to reduce the disturbance of human activities is of great significance for maintaining the microbial diversity and stability of lakes. In addition, the scale effects of land use on rivers and lakes were explored, and the results showed that land use has the greatest effect on microbes within a certain distance. Therefore, the optimal buffer zone can be considered in land use management planning and land use strategy. In other words, the concept of "minimum acceptable width" can be considered in the watershed protection to obtain ideal ecological benefits at the lowest cost via the appropriate width.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Study area and sample collection
  - Water physicochemical parameter
  - Land use parameter acquirement
  - DNA extraction, sequencing, and bioinformatic analysis
  - Statistical analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110053.

ACKNOWLEDGMENTS

This research was funded by National Key research and development program of China (2021YFC3201002, 2021YFC3200103 and 2021YFC3200105), Yangtze River science research joint fund of National Natural Science Foundation of China (2022U2240213).

AUTHOR CONTRIBUTIONS

Z.W.: Conceptualization, Supervision, Project administration, Funding acquisition, Resources, Writing – Review and Editing. **Y.H.**: Investigation, Supervision, Project administration, Funding acquisition, Resources, Writing – Review and Editing. **X.L.**: Investigation, Project administration, Funding acquisition, Resources, Writing – Review and Editing. **L.Z.**: Visualization, Methodology, Formal analysis, Writing – Original Draft, Writing – Review and Editing. **Y.W.**: Writing – Reviewing and Editing. **S.H.**: Writing – Reviewing and Editing. **J.Z.**: Writing – Reviewing and Editing. **X.H.**: Writing – Reviewing and Editing. **R.L.**: Writing – Reviewing and Editing. **H.Y.**: Writing – Reviewing and Editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 3, 2024
Revised: April 2, 2024
Accepted: May 17, 2024
Published: May 27, 2024

REFERENCES

1. Kraemer, S.A., Barbosa da Costa, N., Shapiro, B.J., Fradette, M., Huot, Y., and Walsh, D.A. (2020). A large-scale assessment of lakes reveals a pervasive signal of land use on bacterial communities. ISME J. *14*, 3011–3023. https://doi.org/10.1038/s41396-020-0733-0.
2. Hu, A., Yang, X., Chen, N., Hou, L., Ma, Y., and Yu, C.P. (2014). Response of bacterial communities to environmental changes in a mesoscale subtropical watershed, Southeast China. Sci. Total Environ. *472*, 746–756. https://doi.org/10.1016/j.scitotenv.2013.11.097.
3. Liao, H., Yen, J.Y., Guan, Y., Ke, D., and Liu, C. (2020). Differential responses of stream water and bed sediment microbial communities to watershed degradation. Environ. Int. *134*, 105198. https://doi.org/10.1016/j.envint.2019.105198.
4. Battin, T.J., Luyssaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A., and Tranvik, L.J. (2009). The boundless carbon cycle. Nat. Geosci. *2*, 598–600. https://doi.org/10.1038/ngeo618.
5. Madsen, E.L. (2011). Microorganisms and their roles in fundamental biogeochemical cycles. Curr. Opin. Biotechnol. *22*, 456–464. https://doi.org/10.1016/j.copbio.2011.01.008.
6. Zeglin, L.H. (2015). Stream microbial diversity in response to environmental changes: review and synthesis of existing research. Front. Microbiol. *6*, 454. https://doi.org/10.3389/fmicb.2015.00454.
7. Sanfilippo, G.E., Homola, J.J., Ross, J., Kanefsky, J., Kimmel, J., Marsh, T.L., and Scribner, K.T. (2021). Watershed-scale landuse is associated with temporal and spatial compositional variation in Lake Michigan tributary bacterial communities. J. Great Lake. Res. *47*, 862–874. https://doi.org/10.1016/j.jglr.2021.02.009.
8. 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. https://doi.org/10.1038/s41396-018-0159-0.
9. Wan, W., Grossart, H.P., He, D., Yuan, W., and Yang, Y. (2021). Stronger environmental adaptation of rare rather than abundant bacterioplankton in response to dredging in eutrophic Lake Nanhu (Wuhan, China). Water Res. *190*, 116751. https://doi.org/10.1016/j.watres.2020.116751.
10. Yuan, S., Zhang, W., Li, W., Li, Z., Wu, M., and Shan, B. (2023). Shifts in the bacterial community caused by combined pollutant loads in the North Canal River, China. J. Environ. Sci. *127*, 541–551. https://doi.org/10.1016/j.jes.2022.05.026.
11. Read, E.K., Patil, V.P., Oliver, S.K., Hetherington, A.L., Brentrup, J.A., Zwart, J.A., Winters, K.M., Corman, J.R., Nodine, E.R., Woolway, R.I., et al. (2015). The importance of lake-specific characteristics for water quality across the continental United States. Ecol. Appl. *25*, 943–955. https://doi.org/10.1890/14-0935.1.
12. Savio, D., Sinclair, L., Ijaz, U.Z., Parajka, J., Reischer, G.H., Stadler, P., Blaschke, A.P., Blöschl, G., Mach, R.L., Kirschner, A.K.T., et al. (2015). Bacterial diversity along a 2600 km river continuum. Environ. Microbiol. *17*, 4994–5007. https://doi.org/10.1111/1462-2920.12886.
13. Liu, S., Yu, H., Yu, Y., Huang, J., Zhou, Z., Zeng, J., Chen, P., Xiao, F., He, Z., and Yan, Q. (2022). Ecological stability of microbial communities in Lake Donghu regulated by keystone taxa. Ecol. Indicat. *136*, 108695. https://doi.org/10.1016/j.ecolind.2022.108695.
14. Zhou, J., and Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? Microbiol. Mol. Biol. Rev. *81*, e00002-17. https://doi.org/10.1128/MMBR.00002-17.
15. Liu, X., Wang, X., Zhang, L., Fan, W., Yang, C., Li, E., and Wang, Z. (2021). Impact of land use on shallow groundwater quality characteristics associated with human health risks in a typical agricultural area in Central China. Environ. Sci. Pollut. Res. Int. *28*, 1712–1724. https://doi.org/10.1007/s11356-020-10492-x.
16. Liu, X., Wang, Z., Zhang, L., Fan, W., Yang, C., Li, E., Du, Y., and Wang, X. (2021). Inconsistent seasonal variation of antibiotics between surface water and groundwater in the Jianghan Plain: Risks and linkage to land uses. J. Environ. Sci. *109*, 102–113. https://doi.org/10.1016/j.jes.2021.03.002.
17. Zhang, L., Liu, X., Wei, H., Yang, C., Li, E., and Wang, Z. (2022). Impact of land use on surface water quality: a case study of active agriculturally disturbed basin in China. Int. J. Environ. Sci. Technol. *19*, 4435–4446. https://doi.org/10.1007/s13762-021-03363-6.
18. Wan, W., Gadd, G.M., Gu, J.D., He, D., Liu, W., Yuan, W., Ye, L., and Yang, Y. (2021). Dredging alleviates cyanobacterial blooms by weakening diversity maintenance of bacterioplankton community. Water Res. *202*, 117449. https://doi.org/10.1016/j.watres.2021.117449.
19. Crevecoeur, S., Prairie, Y.T., and del Giorgio, P.A. (2022). Tracking the upstream history of aquatic microbes in a boreal lake yields new insights on microbial community assembly. PNAS Nexus *1*, pgac171. https://doi.org/10.1093/pnasnexus/pgac171.
20. Tang, X., Xie, G., Shao, K., Hu, Y., Cai, J., Bai, C., Gong, Y., and Gao, G. (2020). Contrast diversity patterns and processes of microbial community assembly in a river-lake continuum across a catchment scale in northwestern China. Environ. Microbiome *15*, 10. https://doi.org/10.1186/s40793-020-00356-9.

21. Hauptmann, A.L., Markussen, T.N., Stibal, M., Olsen, N.S., Elberling, B., Bælum, J., Sicheritz-Pontén, T., and Jacobsen, C.S. (2016). Upstream Fresh water and Terrestrial Sources Are Differentially Reflected in the Bacterial Community Structure along a Small Arctic River and Its Estuary. Front. Microbiol. *7*, 1474. https://doi.org/10.3389/fmicb.2016.01474.
22. Nino-Garcia, J., Ruiz-Gonzalez, C., and del Giorgio, P. (2016). Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. ISME J. *10*, 1755–1766. https://doi.org/10.1038/ismej.2015.226.
23. Ruiz-Gonzalez, C., Nino-Garcia, J., and del Giorgio, P. (2015). Terrestrial origin of bacterial communities in complex boreal freshwater networks. Ecol. Lett. *18*, 1198–1206. https://doi.org/10.1111/ele.12499.
24. Dai, Y., Yang, Y., Wu, Z., Feng, Q., Xie, S., and Liu, Y. (2016). Spatiotemporal variation of planktonic and sediment bacterial assemblages in two plateau freshwater lakes at different trophic status. Appl. Microbiol. Biotechnol. *100*, 4161–4175. https://doi.org/10.1007/s00253-015-7253-2.
25. Cottingham, K., Brown, B., and Lennon, J. (2001). Biodiversity may regulate the temporal variability of ecological systems. Ecol. Lett. *4*, 72–85. https://doi.org/10.1046/j.1461-0248.2001.00189.x.
26. Matias, M.G., Combe, M., Barbera, C., and Mouquet, N. (2012). Ecological strategies shape the insurance potential of biodiversity. Front. Microbiol. *3*, 432. https://doi.org/10.3389/fmicb.2012.00432.
27. Fornarelli, R., Antenucci, J.P., and Marti, C.L. (2013). Disturbance, diversity and phytoplankton production in a reservoir affected by inter-basin water transfers. Hydrobiologia *705*, 9–26. https://doi.org/10.1007/s10750-012-1351-2.
28. Ibarbalz, F.M., Henry, N., Brandão, M.C., Martini, S., Busseni, G., Byrne, H., Coelho, L.P., Endo, H., Gasol, J.M., Gregory, A.C., et al. (2019). Global Trends in Marine Plankton Diversity across Kingdoms of Life. Cell *179*, 1084–1097.e21. https://doi.org/10.1016/j.cell.2019.10.008.
29. Zuo, J., Liu, L., Xiao, P., Xu, Z., Wilkinson, D.M., Grossart, H., Chen, H., and Yang, J. (2023). Patterns of bacterial generalists and specialists in lakes and reservoirs along a latitudinal gradient. Glob. Ecol. Biogeogr. *32*, 2017–2032. https://doi.org/10.1111/geb.13751.
30. 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. https://doi.org/10.1111/1462-2920.14815.
31. Wang, K., Ye, X., Chen, H., Zhao, Q., Hu, C., He, J., Qian, Y., Xiong, J., Zhu, J., and Zhang, D. (2015). Bacterial biogeography in the coastal waters of northern Zhejiang, East China Sea is highly controlled by spatially structured environmental gradients. Environ. Microbiol. *17*, 3898–3913. https://doi.org/10.1111/1462-2920.12884.
32. Stegen, J.C., Fredrickson, J.K., Wilkins, M.J., Konopka, A.E., Nelson, W.C., Arntzen, E.V., Chrisler, W.B., Chu, R.K., Danczak, R.E., Fansler, S.J., et al. (2016). Groundwater–surface water mixing shifts ecological assembly processes and stimulates organic carbon turnover. Nat. Commun. *7*, 11237. https://doi.org/10.1038/ncomms11237.
33. Chen, W., Pan, Y., Yu, L., Yang, J., and Zhang, W. (2017). Patterns and Processes in Marine Microeukaryotic Community Biogeography from Xiamen Coastal Waters and Intertidal Sediments, Southeast China. Front. Microbiol. *8*, 1912. https://doi.org/10.3389/fmicb.2017.01912.
34. Zhou, J., Deng, Y., Zhang, P., Xue, K., Liang, Y., Van Nostrand, J.D., Yang, Y., He, Z., Wu, L., Stahl, D.A., et al. (2014). Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proc. Natl. Acad. Sci. USA *111*, E836–E845. https://doi.org/10.1073/pnas.1324044111.
35. Chase, J.M. (2010). Stochastic Community Assembly Causes Higher Biodiversity in More Productive Environments. Science *328*, 1388–1391. https://doi.org/10.1126/science.1187820.
36. Barberán, A., Ramirez, K.S., Leff, J.W., Bradford, M.A., Wall, D.H., and Fierer, N. (2014). Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. Ecol. Lett. *17*, 794–802. https://doi.org/10.1111/ele.12282.
37. Jousset, A., Bienhold, C., Chatzinotas, A., Gallien, L., Gobet, A., Kurm, V., Küsel, K., Rillig, M.C., Rivett, D.W., Salles, J.F., et al. (2017). Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J. *11*, 853–862. https://doi.org/10.1038/ismej.2016.174.
38. Wu, W., Logares, R., Huang, B., and Hsieh, C.H. (2017). Abundant and rare picoeukaryotic sub-communities present contrasting patterns in the epipelagic waters of marginal seas in the northwestern Pacific Ocean. Environ. Microbiol. *19*, 287–300. https://doi.org/10.1111/1462-2920.13606.
39. 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. https://doi.org/10.1111/mec.14218.
40. 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. https://doi.org/10.1038/s41396-018-0153-6.
41. Wang, J., Shen, J., Wu, Y., Tu, C., Soininen, J., Stegen, J.C., He, J., Liu, X., Zhang, L., and Zhang, E. (2013). Phylogenetic beta diversity in bacterial assemblages across ecosystems: deterministic versus stochastic processes. ISME J. *7*, 1310–1321. https://doi.org/10.1038/ismej.2013.30.
42. Wang, Z., Du, Y., Yang, C., Liu, X., Zhang, J., Li, E., Zhang, Q., and Wang, X. (2017). Occurrence and ecological hazard assessment of selected antibiotics in the surface waters in and around Lake Honghu, China. Sci. Total Environ. *609*, 1423–1432. https://doi.org/10.1016/j.scitotenv.2017.08.009.
43. Shi, Y., Li, Y., Xiang, X., Sun, R., Yang, T., He, D., Zhang, K., Ni, Y., Zhu, Y.G., Adams, J.M., and Chu, H. (2018). Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. Microbiome *6*, 27. https://doi.org/10.1186/s40168-018-0409-4.
44. Peng, F., Guo, Y., Isabwe, A., Chen, H., Wang, Y., Zhang, Y., Zhu, Z., and Yang, J. (2020). Urbanization drives riverine bacterial antibiotic resistome more than taxonomic community at watershed scale. Environ. Int. *137*, 105524. https://doi.org/10.1016/j.envint.2020.105524.
45. Ji, L., Wang, Z., Zhang, L., Zhu, X., and Ning, K. (2022). Determining the primary sources of groundwater bacterial communities in a large-scale plain area: Microbial source tracking and interpretation for different land use patterns. Agric. Ecosyst. Environ. *338*, 108092. https://doi.org/10.1016/j.agee.2022.108092.
46. Laperriere, S.M., Hilderbrand, R.H., Keller, S.R., Trott, R., and Santoro, A.E. (2020). Headwater Stream Microbial Diversity and Function across Agricultural and Urban Land Use Gradients. Appl. Environ. Microbiol. *86*, e00018-20. https://doi.org/10.1128/AEM.00018-20.
47. Caillon, F., and Schelker, J. (2020). Dynamic transfer of soil bacteria and dissolved organic carbon into small streams during hydrological events. Aquat. Sci. *82*, 41. https://doi.org/10.1007/s00027-020-0714-4.
48. Ma, Q., Liu, S., Li, S., Hu, J., Tang, M., and Sun, Y. (2020). Removal of malodorant skatole by two enriched microbial consortia: Performance, dynamic, function prediction and bacteria isolation. Sci. Total Environ. *725*, 138416. https://doi.org/10.1016/j.scitotenv.2020.138416.
49. Zhou, J., Deng, Y., Luo, F., He, Z., and Yang, Y. (2011). Phylogenetic Molecular Ecological Network of Soil Microbial Communities in Response to Elevated CO2. mBio *2*, e00122-11. https://doi.org/10.1128/mBio.00122-11.
50. Ju, F., Xia, Y., Guo, F., Wang, Z., and Zhang, T. (2014). Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. Environ. Microbiol. *16*, 2421–2432. https://doi.org/10.1111/1462-2920.12355.
51. Konopka, A., Lindemann, S., and Fredrickson, J. (2015). Dynamics in microbial communities: unraveling mechanisms to identify principles. ISME J. *9*, 1488–1495. https://doi.org/10.1038/ismej.2014.251.
52. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. j. *17*, 10. Next Generation Sequencing Data Analysis. https://doi.org/10.14806/ej.17.1.200.
53. Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460–2461. https://doi.org/10.1093/bioinformatics/btq461.
54. Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods *10*, 996–998. https://doi.org/10.1038/NMETH.2604.
55. 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. https://doi.org/10.1093/bioinformatics/btr381.
56. Gao, Y., Zhang, W., Li, Y., Wu, H., Yang, N., and Hui, C. (2021). Dams shift microbial community assembly and imprint nitrogen transformation along the Yangtze River. Water Res. *189*, 116579. https://doi.org/10.1016/j.watres.2020.116579.
57. Martin, M. (2011). CUTADAPT removes adapter sequences from high-throughput sequencing reads. EMBnet. j. *17*, 10. https://doi.org/10.14806/ej.17.1.200.

58. 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. https://doi.org/10.1128/AEM.00062-07.
59. McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. *6*, 610–618. https://doi.org/10.1038/ismej.2011.139.
60. R Core Team (2021). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
61. Sanchez, G., Trinchera, L., and Russolillo, G. (2015). plspm: Tools for Partial Least Squares Path Modeling (PLS-PM). Version R package version 0.4.9.
62. Oksanen, J., Simpson, G., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P., hara, R., Solymos, P., Stevens, H., Szöcs, E., et al. (2022). vegan community ecology package version 2.6-2 April 2022.
63. Wickham, H. (2016). ggplot2 (Springer International Publishing). https://doi.org/10.1007/978-3-319-24277-4.
64. Rottjers, L., and Faust, K. (2018). From hairballs to hypotheses-biological insights from microbial networks. FEMS Microbiol. Rev. *42*, 761–780. https://doi.org/10.1093/femsre/fuy030.
65. Olesen, J.M., Bascompte, J., Dupont, Y.L., and Jordano, P. (2007). The modularity of pollination networks. Proc. Natl. Acad. Sci. USA *104*, 19891–19896. https://doi.org/10.1073/pnas.0706375104.
66. Banerjee, S., Baah-Acheamfour, M., Carlyle, C.N., Bissett, A., Richardson, A.E., Siddique, T., Bork, E.W., and Chang, S.X. (2016). Determinants of bacterial communities in Canadian agroforestry systems. Environ. Microbiol. *18*, 1805–1816. https://doi.org/10.1111/1462-2920.12986.
67. Herren, C.M., and McMahon, K.D. (2017). Cohesion: a method for quantifying the connectivity of microbial communities. ISME J. *11*, 2426–2438. https://doi.org/10.1038/ismej.2017.91.
68. Ning, D., Deng, Y., Tiedje, J.M., and Zhou, J. (2019). A general framework for quantitatively assessing ecological stochasticity. Proc. Natl. Acad. Sci. USA *116*, 16892–16898. https://doi.org/10.1073/pnas.1904623116.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw data	This paper	NCBI: PRJNA892799
Software and algorithms		
Lima	https://github.com/pacificbiosciences/barcoding	V1.7.0
SMRT link	https://www.pacb.com/products-and-services/analytical-software/smrt-analysis/	V8
Cutadapt	Martin et al.[52]	V1.9.1
USEARCH	Edgar[53]	V10.0
UPARSE	Edgar[54]	V7.1
UCHIME	Edgar et al.[55]	V4.2

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact Zhi Wang (Email: zwang@apm.ac.cn).

Materials availability

This study did not generate new unique data.

Data and code availability

- The datasets analyzed for this study can be found in the Short Read Archive database at NCBI accession: PRJNA892799.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon reasonable request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

This study does not use experimental models in the life sciences.

METHOD DETAILS

Study area and sample collection

The Yangtze River, which is the largest river in Asia and the third-longest in the world, originates from the Qinghai-Tibet Plateau and flows eastward into the East China Sea. It flows through a basin with an area of 1.8 million km^2 and has a population of over 400 million.[56] The location of the water sampling sites in Yangtze River Basin is shown in Figure 1, and the samples were collected from June to October 2020. Sampling sites were evenly distributed throughout the basin, including the main and tributaries of the Yangtze River, as well as major lakes and reservoirs. The water samples obtained by the water collector at a depth of approximately 0.5m were placed into a portable refrigeration and immediately conveyed to a laboratory for subsequent analytical procedures. In addition, we divided the sampling sites into three parts, namely, upstream, midstream, and downstream, based on the spatial distribution of sampling sites, drainage basin topography, and land use pattern (Figure 1). A total of 143 and 36 sampling were collected in rivers and lakes, respectively. There were 50, 54 and 39 sampling in the upstream, midstream and downstream of rivers, and 8, 7 and 21 sampling in the upstream, midstream and downstream of lakes, respectively. None of the sampling sites were in lakes/reservoirs formed by dams and all of the sampling sites are far away from dams.

Water physicochemical parameter

Hydrochemical parameters were measured, encompassing total phosphorus (TP), total nitrogen (TN), ammonium nitrogen (NH_4^+-N), and chemical oxygen demand (COD_{Mn}), employing established laboratory techniques (APHA, 1998). Concurrently, conventional water physicochemical parameters such as dissolved oxygen (DO), water temperature (WT), pH, chlorophyll-a (Chl.a), and conductivity (SpCond) were ascertained *in situ* using a portable multi-parameter water quality analyzer (EXO2, YSI, USA). In addition, environmental distance represents the similarity of physicochemical parameters of water between two samples, measured by calculating the standardized Euclidean distance of water physicochemical parameters between two samples. The larger the environmental distance, the larger the difference in water physicochemical parameters between two samples, and the smaller the environmental distance, the more similar the water physicochemical parameters are between samples.

Land use parameter acquirement

The land use data of Yangtze River Basin were obtained from Landsat 8 imagery (Database: http://www.gscloud.cn) using ArcGIS 10.8. Seven distinct land use types were delineated based on the land cover characteristics of the watershed, encompassing, forest, grassland, residential areas, water body, wetland, cultivated land, and unused land. Unused land was excluded from the subsequent analyses, because it accounts for a very small portion of the watershed. Land use types within six buffer zones for each sampling point were extracted at intervals ranging from 500 m to 5,000 m, specifically at 500 m, 1,000 m, 1,500 m, 2,000 m, 2,500 m, and 5,000 m buffer distances.

DNA extraction, sequencing, and bioinformatic analysis

For each sampling site, 1,500 mL of water from each sample was subjected to concentration on isopore membrane filters (0.2 μm, Millipore, Ireland) and subsequently preserved at −80°C until DNA extraction. The DNA isolation process involved utilizing a DNA isolation kit (Omega, D5625-01, USA). Subsequently, the complete length of the 16S rRNA gene was amplified employing the primer set pA (AGAGTTTGAT CCTGGCTCAG) and B23S (CTTCGCCTCTGTGTGCCTAGGT) (Kurobe et al., 2013). The polymerase chain reaction (PCR) analyses were conducted in triplicate using 20 μL mixtures comprising 10 μL of 2× PCR buffer, 5 μL of template DNA, 4 μL of 2.5 mM dNTPs and 0.5 μL of each primer (10 μM). The PCR program consisted of an initial denaturation step at 95°C for 5 min, followed by 25 cycles at 95°C for 30 s, 50°C for 30 s, and 72°C for 3 min, with a final extension at 72°C for 7 min. Post-amplification, the PCR products underwent detection through 2% agarose gel electrophoresis in 1.0× TAE buffer, and subsequent purification was executed using the AxyPrep DNA gel extraction kit (Axygen, AP-GX-250, USA). Quantification of the purified PCR products was performed on a Microplate reader (BioTek, FLx800, USA) using a Quant-iT PicoGreen dsDNA assay kit (Invitrogen, P7589, USA), after which they were pooled together. The samples were normalized to equimolar amounts in the final mixture and subjected to sequencing on the PacBio Sequel platform (PacBio, USA).

The raw data from each sample were demultiplexed and assigned unique barcodes through lima v1.7.0 (https://github.com/pacificbiosciences/barcoding). Circular Consensus Sequencing data were acquired using SMRT Link v8 (https://www.pacb.com/products-and-services/analytical-software/smrt-analysis/). Evaluation of the raw read quality was conducted using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Subsequently, adapter and primer sequences were excised employing cutadapt v1.9.1.[57] The raw data underwent quality filtration using USEARCH V10.0,[53] applying an expected error filtering approach with a maximum expected error threshold of 1.0. Operational taxonomic units (OTUs) were delineated at a 97% similarity cutoff using UPARSE (version 7.1), and identification and removal of chimeric sequences were executed using UCHIME (version 4.2).[54,55] Taxonomic assignment of each 16S rRNA gene sequence was performed using the RDP Classifier[58] against the Greengenes 16S rRNA database.[59]

Statistical analysis

Partial least squares path modeling analysis (PLSPMA), used to analyze the effects of land use and water physicochemical parameters on microbes, was conducted in R 4.1.1[60] (http://www.r-project.org/) using "plspm" package.[61] Non-metric multidimensional scaling (NMDS), used to analyze the differences in microbial communities, was conducted in R with the "vegan" package[62] (version 2.6-2). Redundancy analysis (RDA), used to analyze the effects of land use at different scales on microbes, was conducted in R using "vegan" package[62] (version 2.6-2). Mann-Whitney U test and regression analysis, were also conducted in R using "ggplot2" package,[63] with significant difference was determined at $p < 0.05$. All default parameters were accepted. OTUs exhibiting relative abundances less than 0.01% of the total sequences were classified as "rare taxa", those exceeding 0.1% were categorized as "abundant taxa," and OTUs with relative abundances ranging from 0.01% to 0.1% were designated as "intermediate taxa".[30] Only OTUs with an average relative abundance > 0.01% were retained to obtain accurate and reliable results for network analysis.[64] Spearman correlation analysis was applied with thresholds set at $|r| > 0.60$ and $P < 0.05$ to identify significant correlations among OTUs. The networks of co-occurrence of the relationship in microbes were visualized using Gephi 0.9.2. Environmental distances were calculated as standardized Euclidean distances based on physiochemical characteristics observed at each sampling site. Specifically, environmental distance represents the similarity of physicochemical parameters of water between two samples. The larger the environmental distance, the larger the difference in water physicochemical parameters between two samples, and the smaller the environmental distance, the more similar the water physicochemical parameters are between samples.

The classification of network nodes into distinct topological roles was based on the calculation of within-module connectivity (Zi) and among-module connectivity (Pi) values. Specifically, nodes were categorized as module hubs ($Zi \geq 2.5$, $Pi < 0.62$), network hubs ($Zi \geq 2.5$, $Pi \geq 0.62$), connectors ($Zi < 2.5$, $Pi \geq 0.62$), and peripherals ($Zi < 2.5$, $Pi < 0.62$) according to established criteria.[65] The three identified categories, comprising module hubs, network hubs, and connectors, were collectively considered as potential keystone taxa.[66] In addition, rich species as those having relative abundances above 0.1%, rare species as those having relative abundances below 0.01% and intermediate species as those having relative abundances between 0.01% and 0.1%.[9]

Microbial community connectivity was quantified by calculating both positive and negative cohesion, serving as metrics to characterize associations among taxa attributable to positive and negative species interactions, as well as niche similarities and differences among microbial taxa.[67] Initially, an abundance-weighted matrix was generated based on pairwise correlations across taxa. Subsequently, the connectedness matrix, incorporating average positive and negative correlations, was calculated. Null model was used to correct bias in these correlations. The positive and negative cohesions were then computed as follows:[67]

$$\text{Cohesion} = \sum \text{abundance}_i \times \text{connectedness}_i,$$

Furthermore, the ratio of negative cohesions to positive cohesions (N:P cohesion) was computed as an index reflecting community stability, where a higher value indicated a greater degree of community stability.

The discernment of deterministic and stochastic mechanisms governing microbial community assembly processes involved the computation of the modified stochasticity ratio (MST), implemented in the R using the NST package.[68] This index served as a quantitative measure indicating the proportion of stochastic processes influencing community assembly. A higher MST value suggested a greater contribution of stochastic processes, wherein MST values exceeding 50% signified the dominance of stochastic processes in microbial community assembly, while MST values below 50% indicated the prevalence of deterministic processes in shaping the microbial community structure. All default parameters were accepted.

QUANTIFICATION AND STATISTICAL ANALYSIS

In this study, Mann-Whitney U test was used to detect the difference in the comparison results of microbes in river and lake, with significant difference was determined at $p < 0.05$.