

# Dynamics and Heterogeneity of Microbial Communities in Urban Water Environments

Yanqing Lian (✉ [lianyq@ieecas.cn](mailto:lianyq@ieecas.cn))

Institute of Earth Environment Chinese Academy of Sciences <https://orcid.org/0000-0002-4093-4030>

Lisha Zhen

Shaanxi Provincial Institute of Microbiology

Xi Chen

Xi'an Polytechnic University

Yang Li

Institute of Earth Environment Chinese Academy of Sciences

Xiaona Li

Institute of Earth Environment Chinese Academy of Sciences

---

## Research

**Keywords:** Urban water environment, 16S rRNA gene sequencing, Dynamics of microbial community, heterogeneity of microbial community, LEfSe analysis

**Posted Date:** September 23rd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-75523/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Water samples for the 16S rRNA gene and water quality analyses were collected from around 155 kilometers of river segments surrounding the urban areas in Xi'an of China. Multiple statistical analyses showed the temporal dynamics of microbial communities and heterogeneity in their spatial distributions. The dynamic shifts of microbial communities in the Chan, Ba, and Feng Rivers from the Spring to the Summer seasons were apparent, but little in the Zao River. The heterogeneity of microbial distributions was more due to the influence of hydrologic conditions and various sources of inflows in the rivers. The LEfSe analysis showed the Chan and Zao Rivers, both were more impacted by the sewage effluents, were more differentially abundant with bacteria related to polluted water, but the Ba and Feng Rivers, both on the outer side of the city, were more abundant with microbial communities in soil and freshwater environments in August. Multiple statistical analyses indicated that environmental variables had a significant impact on microbial communities. The GIS-based spatial analysis not only showed heterogeneity of microbial community distributions along the rivers, more importantly, could help identify locations where pathogenic bacteria presented.

## Introduction

With the continuous development of biotechnology, methods for microbial research have significantly advanced. The high-throughput sequencing technology provides a considerable amount of biological information in studying microbial diversity and functionalities in the water environment. It allows for a more accurate analysis of the structures and functions of the microbial communities in a water environment to improve environmental quality and resilience. Many studies were carried out on the diversities of bacterial communities in the water environment in recent years. Researches in urban areas (Bai, Y. et al., 2018; Berglund, B. et al., 2015; Drury, B., et al., 2013) have shown that human activities, particularly sewage discharge, have had significant impacts to microbial communities in water and sediment.

Berglund, B. et al.'s (2015) study indicated that the treated effluent discharged from the wastewater treatment plant downstream would have contributed to the increase in antibiotic resistance genes (ARGs) in the river downstream. A study by Drury, B., et al. (2013) claimed that effluents from wastewater treatment plants would reduce the abundance and diversity of benthic bacterial communities in urban and suburban rivers. Vukovic, Z. et al. (2011) studied the microbial and heavy metal pollution of the Sava River and found that heavy metal levels in sediments correlated with those in the river water. The study by Peng, F. J., et al. (2019) showed that the community composition of the benthic bacteria was significantly correlated with galaxolide, total phosphorus, lead, and triclosan in urban rivers in Guangzhou of China, which suggested that sewages had altered the bacterial community compositions in urban streams.

Bower, P. A., et al. (2005) studied the spatial and temporal distributions of bacteria following major rain events and detected the human-specific *Bacteroides* as far as 2 km into Lake Michigan 1 to 9 days post the sewage overflow. Araya, R. et al.'s (2003) study revealed that beta-Proteobacteria and Cytophaga-

Flavobacterium clusters were the dominant bacterial populations in both stream water and biofilms in their study site. Zhang, H., et al. (2018)'s study indicated that Proteobacteria and Cyanobacteria were the most common phyla in the urban lakes in Xi'an of China, and the bacterial communities in the water among three urban lakes were dramatically different. There was a good correlation of fungal communities with water quality parameters such as CODMn and total nitrogen, and TN.

With the increasing use of pharmaceutical products, the antibiotic resistance genes (ARGs) have become a great concern to the health of humans and ecosystems in urban environments. Several studies have reported ARGs detected in water bodies and sediment in Xi'an. A survey by Guan, Y. et al. (2018) indicated that Proteobacteria and Bacteroidetes were the dominant phyla in the study area, and sediment had higher bacterial diversity and richness than water in the Ba River. Nutrients, heavy metals, and antibiotics might have influenced Bacterial communities. Pharmaceutical products, including sulphonamides, quinolones, macrolides, and tetracyclines were detected with high frequencies and Tetracyclines with high concentrations in water and sediment samples in the ba River (Jia, J. et al., 2018). Wang, J. et al. (2019) found that the Wei River's antibiotic residues were at a moderate contamination level. The primary sources were domestic sewage, livestock discharge, and aquaculture, and pharmaceutical wastewater.

Studies reviewed above have shown that human-induced environmental factors have a direct impact on microbial communities. This study was to conduct systematic research on the four major rivers in the urban environment of the Xi'an City, for a better understanding of the temporal dynamics and spatial heterogeneity of microbial communities and to identify biomarkers as indicators of environmental quality.

## **Materials And Methods**

### **Study area**

Xi'an was the birthplace of cultural civilization and the capitals for 13 dynasties in ancient China, with an annual rainfall of about 522.4 ~ 719. 5 mm and the annual average temperature around 13.0 °C ~ 13.7 °C. It is known for eight rivers surrounding the city, including the Wei, Jin, Feng, Lao, Yu, Hao, Chan, and Ba Rivers. The Chan and Ba Rivers on the east side and the Feng and Zao Rivers on the west side in the Xi'an City urban area (Fig. 1) are the focus of this study. The Chan, Ba, and Feng Rivers are originated from the Qin Mountains, with main stem lengths of 109-, 64-, and 84-kilometers, respectively. On the other hand, the Zao River starts from the alluvial plain with a reach length of about 30 kilometers, serving primarily as drainage for stormflow and sewage on the west side of the city.

### **Sample collection**

Sixty-two sets of water samples were collected in April, May, and August of 2019 from the four major rivers (Fig. 1). Sampling locations were selected to represent various sources of flows and water environments. Each set of samples was prepared for the microbial, physiochemical, environmental isotope, pharmaceutical products, and microplastics analysis. Microbial water samples were collected

using a battery charged pump to pump 1–5 liters of water through the 0.45 nanometers pore diameter filter paper. The filtered sample papers were properly stored before the 16S rRNA gene testing. Environmental parameters such as pH, temperature, dissolved oxygen, electric conductivity, hardness, Oxidation Reduction Potential (ORP) were measured on-site. Other environmental parameters such as cation, anion, nitrate, phosphorous, heavy metal, and environmental isotopes were analyzed in the laboratory.

Water samples were properly stored and shipped to the commercial laboratory for the 16S rRNA gene sequencing, and the sequenced data were stored on the cloud server for statistical analysis. The package UPARSE (Edgar, 2013) was used for high-accuracy and high-throughput OTU clustering. Sequences with greater than 97% similarity were assigned to the same OTUs. For each representative sequence, the Silva database was used based on the RDP classifier Bayesian algorithm to annotate taxonomic information the microbial communities at the domain, kingdom, phylum, class, order, family, genus, and species levels.

## Statistical analyses

After R. H. Whittaker (1960, 1972) introduced the alpha diversity ( $\alpha$ -diversity) together with the terms of beta diversity ( $\beta$ -diversity) and gamma diversity ( $\gamma$ -diversity), more indices have been subsequently developed to quantify the richness and diversity of biological communities. It was understood that the sampling process would lead to the overestimation of the beta-diversity of microbial communities (Zhou, J., et al., 2013). This paper selected to adopt the commonly used indices such as the sobs, Shannon, Simpson, Chao, Coverage, Shannoneven, and Simpstoneven.

The principal component analysis (PCA), the principal coordinates analysis (PCoA), and the non-metric multidimensional scaling (NMDS) are commonly used ordination methods for beta diversity analysis. According to Gauch (1982), the underlined linear assumption makes PCA unsuitable for most ecological data sets; thus, this study selected to use the principal coordinate analysis (PCoA), in which the distance metric was calculated using the Bray-Curtis algorithm. The Wilcoxon rank-sum test, which is particularly advantageous when comparing the averages of independent groups of samples for which a Gaussian type distribution cannot be assumed or unknown, was used to assess the microbial community's differences distributions between two or multiple groups. Spearman's rank correlation coefficient analysis could evaluate the relationship between environmental variables and microbial communities. The Linear discriminant analysis Effect Size (LEfSe) developed by Segata et al. (2011) was used to identify microbes that are abundantly different between two or multiple sample groups.

Software packages used in this study included: the Uparse 7.0.1090 for OTU classifications, Usearch 7.0 for OTU statistics, Mothur 1.30.2 for Alpha diversity, Qiime 1.9.1 for Beta diversity and sequence deduplication filtering, LEfSe for Linear discriminant analysis Effect Size analysis, and IQ-TREE 1.6.8 for construction of the phylogenetic tree.

## Results And Discussions

## Microbial community compositions

Microbiome data tend to be very noisy, and the total number of counts per sample is highly variable because of the experimental process and quality control filtering. For consistency in statistical analyses, all samples in this study were normalized by the minimum reads of 33,827. The taxonomy analysis has identified 55 phyla, 1502 genus, 3161 species, and 8042 OTUs from all samples. Out of the 55 phyla identified, 40 phyla are common in April, May, and August, 40 are common between April and May, 48 between April and August, 42 between May and August. April and August had two and three unique phyla, respectively, while May had none.

Major phyla with the relative percentage of abundance greater than 1% included Proteobacteria, Bacteroidetes, Firmicutes, Actinobacteria, Cyanobacteria, Deinococcus-Thermus, Patescibacteria, Epsilonbacteraeota, Chloroflexi, and Acidobacteria (Fig. 2a). Additionally, Nanoarchaeaeota, WS4, and Calditrichaeota were identified in a sample affected by the sewage discharge to the Chan River. Zixibacteria, WS4, and Calditrichaeota were also detected in samples from the Zao River. Diapherotrites, which was first identified by Youssef (2014) in the groundwater seepage, was found in the August samples from the Zao River.

Some pathogenic bacteria such as Lentisphaerae, Fusobacteria, Spirochaetes, Dependitiae, and Elusimicrobia were also found at various locations of the rivers with different abundance in different sampling periods. It indicated industrial and pharmaceutical wastes contaminated the urban water environments, and there existed a serious threat to the healthiness of eco-environments in the region.

## Microbial community differences

For the statistical analysis, samples were grouped by river segments, and each river segment was sub-grouped by sampling periods. Shown in Fig. 2c is the PCoA analysis of all 12 sample groups. Samples from the same river segment were marked with the same symbol but differentiated by the red, green, and blue colors for samples collected in April, May, and August. The PCoA plot showed that the differences in microbial communities from different sampling periods were far greater than the differences between river segments, with the variance explained on PC1 and PC2 as 26.65% and 10.24%. The separation of PC1 for the Chan, Ba, and Feng Rivers between the April/May and August samples were distinctive except for the Zao River.

The distance bar plot (Fig. 2b) showed that the dissimilarity of microbial communities between sample groups was higher than within each sample group. In a way, this test would also provide information for the similarity of microbial communities between sample groups. The ANOSIM (Analysis of Similarity) analysis had an R statistic of 0.65 with a p-value of 0.001, which indicates that the similarity between groups is greater than or equal to the similarity within the groups. ADONIS test results with  $R^2$  of 0.51 and a P-value of 0.001 implies that the classification of sampling groups by considering the geographical characteristics of rivers and the temporal variation of microbial communities is reasonable.

## Seasonal changes in microbial communities

As indicated by the Shannon indices (Table S1), the overall microbial diversity in August was higher than in April/ May. Samples from the Qin Mountains, the Chan, Ba, Chan\_Ba joined, and the Feng River segments all had higher diversity in August than in April or May. Shannon indices between river segments had larger variation in April and May than in August, which could be due to the dominance of recharge from the Qin Mountains areas in August but various sources of local inflows like the sewer outflows in April and May. The Zao River as a drainage channel for the municipal sewer discharge without the direct impact of flows from the Qin Mountains, its microbial communities in May and August did appear to have a significant change. Shannon indices varied between 1.33 and 1.54 in August but between 0.48 and 1.25 in April/May. Shannon evenness ranged from 0.37 to 0.44 in August but from 0.15 to 0.38 in April/May.

To further understand whether there were statistically significant differences between two or more sample groups, the Kruskal-Wallis H test was conducted on all sample groups. For the first ten major phyla, the difference among all groups for each phylum was significant, with p-values less than 0.05. It was clear that the proportions of Proteobacteria and Bacteroidetes were higher in August than in April/May for the Chan, Ba, Chan\_Ba joined, Feng, and Zao River segments. However, Firmicutes, Actinobacteria, Cyanobacteria, and Deinococcus-Thermus were higher in August than in April/May for all sample groups. While in the Zao River, Patescibacteria and Epsilonbacteria were notably higher than in other rivers, reflecting the microbial communities affected by stormwater and sewage discharges in an urban area.

Wilcoxon rank-sum test between two groups and the Fisher's Exact test between two samples was also conducted to understand microbial community diversities between locations and periods. A comparison of sample groups was to look at how microbial communities would be different on temporal scales, particularly their diversities at different sampling periods for the same river segment (Fig. 3). Shown in the figure are the first 15 major phyla in compared groups. The Proteobacteria was the dominant phylum in the Chan, Ba, Feng, and Zao rivers in April/May and August. It was also apparent that the microbial communities shifted from April/May to August. The Chan, Ba, and Zao Rivers had higher proportions of Proteobacteria and Bacteroidetes in April/May than in August. On the contrary, proportions of Firmicutes, Actinobacteria, Cyanobacteria, and Deinococcus-Thermus turned to be higher in August than in April/May. Shifts of proportions for the majority of the phyla in the Ba and Feng Rivers are statistically significant, with the p-values for the Wilcoxon rank-sum test less than 0.05, only Actinobacteria and Deinococcus-Thermus were statistically higher in August. For the Zao River, even though no statistically significant shift of proportion for any phylum, it was apparent that Firmicutes, Bacteroidetes, Deinococcus-Thermus, Cyanobacteria, and also the Epsilonbacteria had more proportions in August.

## Differentially abundant biomarkers between sampling periods

The Wilcoxon rank-sum test was conducted for the first 15 major phyla. Results showed three major phyla had statistically significant differences between April and August in the Feng River. The majority had significant differences in the Ba and Chan Rivers, but none had a significant difference from August

to May in the Zao River. While the Wilcoxon rank-sum test can check the statistical significance of the relative abundance difference of a selected class, the LEfSe method is advantageous in identifying clades with statistically and biologically differences between communities for a range of classes. The LEfSe analysis was from Phylum to Class in this study (Fig. 4 and Table S3), which has identified 12 abundantly differential clades in the Chan River, 44 in the Ba River, and 53 in the Feng River, but none in the Zao River. It worth noting no abundantly differential phylum from May to August in the Zao River further proved that its water environment did not have significant change between May and August. For the Chan River, which is also impacted by sewage effluent, Proteobacteria was more abundant in April, but Firmicutes, Actinobacteria, and Deinococcus\_Thermus were more abundant in August. The Ba River and Feng River, both on the outer side of the city, showed microbial communities that are more abundant in soil and freshwater environments. Both rivers had Proteobacteria and Bacteroidetes abundantly differential in April/May, but Firmicutes, Actinobacteria, Cyanobacteria, Chlamydiae, Planctomycetes, Chloroflexi, and Gemmatimonadetes in August. The existence of Gemmatimonadetes could imply the impact of soil erosion from the Chan River watershed. Chlamydiae, which is the most common bacterial STD in the United States with 2.86 million reported Chlamydiae infections annually, was found in the Ba River and Zao River in August. The differentially abundant microbial identified by the LEfSe analysis in the Feng River in August was also surprising. Proteobacteria and Bacteroidetes were more abundant in May, but the 18 other phyla more abundant in August included typical freshwater bacteria of Cyanobacteria, Armatimonadetes, Planctomycetes, Chloroflexi and Fibrobacteres, typical soil bacteria of Gemmatimonadetes, Kiritimatiellaeota, Acidobacteria and Patescibacteria, and Deinococcus\_Thermus and Dependistiaes often found in sewage polluted water. Other bacterias were Margulisbacteria, which was first found in the marine sample, Omnitrophicaeota, Latescibacteria, which is often related to algae, and Nitrospirae that can be found in a wide range of environment, including soils. The combination of bacteria indicated that in addition to the recharge from the Qin Mountains, the Feng River was also affected by soil erosion, non-point source pollution from horticulture and agriculture, and some degree of sewage pollution. It was quite alarming that Fusobacteria, Spirochaetes, and Chlamydiae, known for pathogenic diseases, were found in the Feng River.

### Microbial community differences along rivers

The composition of microbial communities in the study area was significantly impacted by environmental factors such as land-use changes and human activities such as sewage discharges. Shown in Figs. 5a and 5b are the composition of major phyla at each location, which shows the spatial distribution of microbial communities. Comparing samples from the same sampling locations in April/May and August can show the temporal dynamics of microbial communities. Figure 5a shows that Proteobacteria is a predominant phylum in all samples from April and May, with Proteobacteria percentage greater than 80% in samples from the Qin Mountains area. The percentages of Bacteroidetes, Firmicutes, and Actinobacteria were relatively high in the May samples from the Feng River, in samples from the upper portion of the Ba river. The percentage of Patescibacteria appeared to be relatively high in samples affected by sewage discharge, such as sample A18 and all samples from the Zao River. While in Fig. 5b, the composition of microbial communities in samples from August is different from April and

May. In addition to Proteobacteria, Bacteroidetes, Firmicutes, and Actinobacteria remaining to be dominant in most of the samples in August, the relative composition of Cyanobacteria were higher in August than in April/May samples in the Chan, Ba, and Feng Rivers. The percentage of Firmicutes was higher in August than in April in most of the samples from the Chan and Ba Rivers. Patescibacteria percentage was also high in sample B32, which was also due to the sewage discharge in August.

### **Microbial community differences between rivers**

Comparison analysis of microbial diversities through the Wilcoxon rank-sum test (Figure S1) showed there were no statistically significant shifts of proportions for the majority of the phyla in the Chan River and Ba River for the same sampling periods. It was only seen that Epsilonbacteraeota, Cyanobacteria, and Acidobacteria to be higher in the Chan River than in the Ba River in April and Spirochaetes to be higher in the Chan River in August. Differences between the Feng River and Zao River in May and August were statistically significant for a large portion of phyla because both rivers are not hydrologically well-connected. The Zao River had higher proportions of Patescibacteria, Epsilonbacteraeota, Chloroflexi, Fusobacteria, Chlamydiae, and Synergistetes in May and higher Patescibacteria, Epsilonbacteraeota, and Fusobacteria in August. In contrast, the Feng River had higher proportions of Verrucomicrobia in May and higher Cyanobacteria, Actinobacteria, Verrucomicrobia, Gemmatimonadetes, and Amatinomonadetes in August.

LEfSe analysis was also conducted to compare the abundance of microbial communities between two river segments for the same sampling period (Figure S2). In April, differentially abundant phyla in the Chan River included Epsilonbacteraeota, Cyanobacteria, Acidobacteria, Planctomycetes, Gemmatimonadetes, Chlamydiae, Kiritimatiellaeota, and BRC1, but none in the Ba River. In August, microbial communities in both the Chan and Ba Rivers were very similar due to the dominant inflows from the Qin Mountains areas. Spirochaetes were identified as abundant in the Chan River, which could be due to the effluent from the sewage discharge. Comparison of the Feng and Zao Rivers showed that the Zao River had Patescibacteria, Epsilonbacteraeota, Diapherotrites, Kiritimatiellaeota, Omnitrophicaeota, Elusimicrobia, Fusobacteria, Chlamydiae, WPS\_2, WS2, Synergistetes, and Dependitiae differentially abundant in both May and August.

Differences of major phyla between all four rivers were examined with the Kruskal-Wallis H test. The test results showed that Patescibacteria, Deinococcus\_Thermus, Epsilonbacterarota, Verrucomicrobia, and Chloroflexi were the phyla in the top 10 major phyla with statistically significant differences among the four rivers in April/May (Fig. 8a). Figure 8b-f showed Patescibacteria and Deinococcus\_Thermus were most abundant in the Zao River, Epsilonbacterarota in both the Chan and Zao Rivers. Verrucomicrobia was most abundant in the Feng River and the second most abundant in the Chan River. Chloroflexi was most abundant in the Zao River and the second most abundant in the Chan River.

The Kruskal-Wallis H test showed an increase in the typical freshwater bacteria of Cyanobacteria and Actinobacteria in the Feng, Chan, and Ba Rivers due to the recharge from the Qin Mountains streams.



Patescibacteria and Epsilonbacterarota remained to be the most abundant species only in the Zao River because of its primary sources of recharge from the domestic sewage effluents.

### **Differentially abundant biomarkers between rivers**

The LEfSe test was also conducted on the differential abundances at the phylum and class levels between the Chan, Ba, Feng, and Zao Rivers (Fig. 7, Table S4 and S5). For the April/May samples, Chan, Ba, Feng, and Zao Rivers had 13, 0, 5, and 51 differentially abundant clades, with differentially abundant phyla of Epsilonbacteraeota, Kiritimatiellaeota, Synergistetes, and Gemmatimonadetes in the Chan River, none in the Ba River, Verrucomicrobia and Planctomycetes in the Feng River, and Patescibacteria, Deinococcus\_Thermus, Chloroflexi, Omnitrophicaeota, Dependuntiae, Acidobacteria, Chlamydiae, WPS\_2, Elusimicrobia and Fibrobacteres in the Zao River. For the August samples, there were 1, 2, 14, and 51 differentially abundant clades of phylum and class in the Chan, Ba, Feng, and Zao Rivers, with differentially abundant phyla of Actinobacteria in the Chan River, none in the Ba River, Cyanobacteria, Verrucomicrobia, Gemmatimonadetes, Acidobacteria, Armatimonadetes, and Margulisbacteria in the Feng River, and Patescibacteria, Epsilonbacteraeota, Omnitrophicaeota, Lentisphaerae, Fusobacteria, Synergistetes, Kiritimatiellaeota, WPS\_2, Spirochaetes, Dependuntiae, Latescibacteria and Elusimicrobia in the Zao River. The existence of Epsilonbacteraeota and Synergistetes in the Chan and Zao Rivers indicated some degrees of pollution from domestic wastes, pathogenic bacteria such as Lentisphaerae, Fusobacteria, Spirochaetes, Dependuntiae, and Elusimicrobia was not only an indication of contamination from industrial and pharmaceutical wastes but also serious threat to the healthiness of eco-environment in the urban area.

Although both the Kruskal-Wallis H test and the LEfSe analysis are efficient in detecting species that are different among sample groups, it cannot show the spatial distribution of microbes. This study adopted the geographical information system (GIS) to map out the spatial distribution of selected bacteria with absolute abundance, as shown in Figure S4, which would help to identify the location of microbial such as the pathogenic bacteria that are harmful to human and the eco-environment health.

### **Microbial diversity driven by environmental factors**

Single water quality indices from the pH, Do, nitrogen, phosphorous, and COD were computed and combined into a composite water quality index (CWQI) for each river segment, with a smaller value indicating better water quality. Based on the index values (Fig. 8a), it was clear that water from the Qin Mountains areas has better quality. The composite water quality indices for the Chan River were high due to high nitrogen and phosphorous by pollution from the non-point source of stormflow and sewage discharge. The drainage area of the Ba River was predominantly agricultural land. Its water quality indices for April and August were both lower. It was understandable that the composite index for the joined river segment of the Chan and Ba Rivers fell between the Chan and Ba Rivers. Indices for the Feng River and the Zao River were lower than those of the Ba and Chan Rivers due to lower indices for nitrogen and phosphorous. For a seasonal change, the composite index for all river segments was higher in

August than in April and May due mostly to the increase of phosphorous and nitrogen. Influences of the water environment on microbial communities are analyzed in this section.

### **Correlation of microbial communities with environmental variables**

Pearson's correlation and Spearman's Rank Correlation are common algorithms for linear correlation analysis, but Pearson's correlation is sensitive to outliers; therefore, this study adopted the Spearman's Rank Correlation. The Spearman's rank correlation coefficients between environmental variables and microbial communities on the phylum level were presented in the heatmap shown in Fig. 8c. The strength of the correlation was indicated by color bands with a darker blue for negative correlation and red for positive correlation. The p-value is marked by an asterisk \* when  $0.01 < P \leq 0.05$ , \*\* for  $0.001 < P \leq 0.01$  and \*\*\* for  $P \leq 0.001$ . The average distance was used for clustering environmental variables and microbial communities.

There were three major clusters for environmental variables, i.e., cluster of pH, DO and Resistance, cluster of Turbidity, TOC, Nitrate and Total Nitrogen, and the cluster of Nitrite, EC, Soluble Phosphorus, Total Phosphorus, Hardness and CODMn. It can be seen in the heatmap that most of the major phyla were negatively correlated to the pH of the water sample; some were negatively related to the DO (dissolved oxygen) and resistance. Proteobacteria and Verrucomicrobia were positively correlated to all three variables. Proteobacteria had a stronger negative correlation with the second cluster of environmental variables. Other than Verrucomicrobia, Bacteroidetes, Defermentia, and WPS-2, other phyla had a positive correlation to the second cluster of environmental variables. Chlamydiae, Firmicutes, Deinococcus-Thermus, Epsilonbacteraeota, Kiritimatiellaeota, Synergistetes, Omnitrophicaeota, Fusobacteria, Elusimicrobia, Margulisbacteria, Tenericutes, Chloroflexi, and Spirochaetes, were positively correlated to the second and third clusters of environmental variables.

The impact of environmental variables was also tested using the redundancy analysis (RDA). Figure 8b showed that pH, DO, EC, Resistance, Nitrate, Total Nitrogen, Hardness, and TOC had more influence on microbial communities, all with a statistical p-value less than 0.05 (Table S2). To further examine the effect of environmental factors on microbial communities, the linear regression analysis was conducted for each environmental variable to April/May samples, August samples, and all April/May/August samples. Their respective coefficients of determination and p-values are summarized in Table S2. It can be seen that DO, Resistance, Nitrate, Total Nitrogen, Hardness, and TOC had a relatively better correlation with phyla, pH, Ammonia, Total Phosphorus, Soluble Phosphorus, and CODMn didn't show good linear correlation. The linear regression analysis showed that the environmental variable that had a major impact on microbial communities was consistent with the conclusions from the Spearman's Rank Correlation, the RDA analysis except for EC.

## **Conclusions**

Taxonomy analysis showed that the urban rivers in Xi'an were abundant with 55 phyla, 1502 genus, 3161 species, including typical freshwater bacteria, typical soil bacteria, and also bacteria often found in

sewage polluted water. The PCoA ordination analysis showed that the differences in microbial communities from different sampling periods were far greater than the differences between river segments, with the variance explained on PC1 and PC2 as 26.65% and 10.24%. Other than the Zao River, the separation of PC1 for the Chan, Ba, and Feng Rivers between the April/May and August samples was distinctive. ANOSIM analysis indicates that the similarity between groups is greater than or equal to the similarity within the groups. ADONIS implies that the classification of sampling groups by considering the geographical characteristics of rivers and the temporal variation of microbial communities is reasonable and also that the temporal or seasonal changes of microbial communities are more significant than the spatial variations.

Sources of recharge to the Chan, Ba, and Feng Rivers include the inflows from the Qin Mountains upstream areas, stormwater runoff from the agricultural areas, and urban areas, from treated sewer water or even some untreated sewer flow. The dynamic shifts of microbial communities in the Chan, Ba, and Feng Rivers from the Spring to the Summer seasons were apparent. Wilcoxon rank-sum test and the Fisher's Exact test showed statistically significant shifts of proportions between April/May and August for the majority of the phyla. None in the Zao River from May to August.

The LEfSe analysis identified the differentially abundant biomarkers between sampling periods in each river segment and between river segments in each sampling period. Proteobacteria was differentially abundant in April/May in the Chan, Ba, and Feng Rivers, but in August, the Chan River was differentially abundant with freshwater bacteria, the Ba River was more with the freshwater and typical soil bacteria, the Feng River was more differentially abundant with the freshwater bacteria, typical soil bacteria, bacteria from non-point source pollution, and bacteria related to polluted water. The Zao River did not have significantly abundant bacteria, which further proved that its water environment did not have a significant change between the two sampling periods.

Comparison of the Chan, Ba, Feng, and Zao Rivers for the same sampling period showed that the Chan River was differentially abundant with bacteria related to the polluted water environment and even the pathogenic bacterial Chlamydiae in April, but in August, both rivers did not appear to have differentially abundant bacteria due primarily to the dominant inflow from the Qin Mountains during the summer rainstorm season. For the Feng and Zao River, the Feng River was differentially abundant with soil bacteria Verrucomicrobia in May, but the Zao River was more with bacteria often found in the polluted water environment; but in August, the Feng River was more abundant with the typical freshwater and typical soil bacteria, whereas the Zao River was still more abundant with bacteria in the polluted water environment. Statistical analyses showed that environmental variables such as pH, DO, EC, Resistance, Nitrate, Total Nitrogen, Hardness, and TOC had a significant impact on microbial communities.

Overall, this study showed the temporal dynamics and spatial heterogeneity of microbial communities were due to the effects of natural environments such as the seasonal variation of hydrologic and hydraulic conditions and human-induced factors such as sewage discharges and non-point sources from urban stormflow and agriculture as well. It was alarming that pathogenic bacteria such as *Lentisphaerae*,

Fusobacteria, Spirochaetes, Dependuntiae, Chlamydiae, and Elusimicrobia were also found at various locations of the rivers. The greater variation of microbial communities in April/ May was likely due to local inflows such as the sewage discharges and tributaries along rivers. Microbial communities were relatively uniform in the Chan, Ba, and Feng Rivers in August due to the dominance of recharge from the Qin Mountains areas. Both the Chan River and Zao River were impacted more by sewage effluent. The Ba River and Feng River, both on the outer side of the city, showed microbial communities that are more abundant in soil and freshwater environments in August than in April/May. The GIS-based spatial analysis not only showed heterogeneity of microbial community distributions along the rivers, more importantly, could help identify locations where pathogenic bacteria presented. Further study will adopt the comprehensive functional gene array (GeoChip) technique to identify functional genes for effective in-situ bioremediation of water and sediments in wetlands for ecosystem restoration in the urban water environments.

## **Declarations**

## **Ethics approval and consent to participate**

Not applicable.

## **Consent for publication**

Not applicable.

## **Availability of data and material**

The datasets generated and/or analysed during the current study are submitted with this article.

## **Competing interests**

There was no competing interests.

## **Funding**

Funding for this research was provided by the Strategic Priority Research Program of the Chinese Academy of Sciences, Grant No. XDB40020100, and by the "Light of West China" Program of the Chinese Academy of Sciences.

## **Authors' contributions**

Dr. Lian directed the research, conducted data analysis, and prepared the tables, figures, and the manuscript. Dr. Lisha Zhen, Dr. Xi Chen, Ms. Yang Li and Ms Xiaona Li participated field campaign for water sampling. All authors read the manuscript.

## Acknowledgments

This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences, Grant No. XDB40020100, and by the "Light of West China" Program of the Chinese Academy of Sciences.

## Conflict of interest

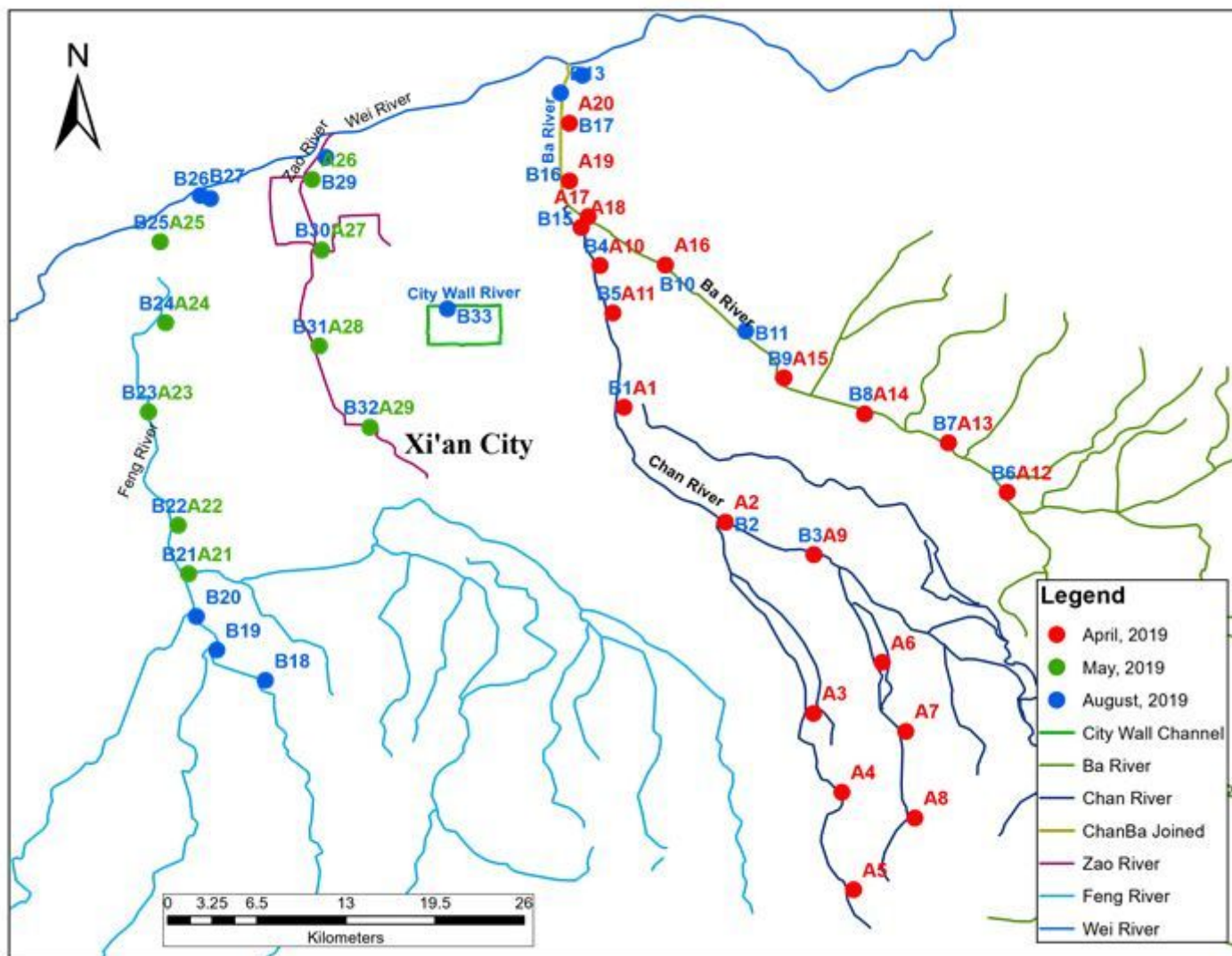
There is no conflict of interest.

## References

1. Araya R, Tani K, Takagi T, Yamaguchi N, Nasu M. Bacterial activity and community composition in stream water and biofilm from an urban river determined by fluorescent in situ hybridization and DGGE analysis. *FEMS Microbiol Ecol*. 2003;43(1):111–9.
2. Bai Y, Wang Q, Liao K, Jian Z, Zhao C, Qu J. Fungal Community as a Bioindicator to Reflect Anthropogenic Activities in a River Ecosystem. *Front Microbiol*. 2018;9:3152.
3. Berglund B, Fick J, Lindgren PE. Urban wastewater effluent increases antibiotic resistance gene concentrations in a receiving northern European river. *Environ Toxicol Chem*. 2015;34(1):192–6.
4. Bower PA, Scopel CO, Jensen ET, Depas MM, McLellan SL. Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. *Appl Environ Microbiol*. 2005;71(12):8305–13.
5. Drury B, Rosi-Marshall E, Kelly JJ. Wastewater treatment effluent reduces the abundance and diversity of benthic bacterial communities in urban and suburban rivers. *Appl Environ Microbiol*. 2013;79(6):1897–905.
6. Edgar RC, UPARSE. Highly accurate OTU sequences from microbial amplicon reads, *Nat Methods*. 2013; [PubMed:23955772, [dx.doi.org/10.1038/nmeth.2604](https://doi.org/10.1038/nmeth.2604)].
7. Gauch HG Jr. *Multivariate Analysis in Community Ecology*. Cambridge: Cambridge University Press; 1982. p. 298.
8. Guan Y, Jia J, Wu L, Xue X, Zhang G, Wang Z. Analysis of Bacterial Community Characteristics, Abundance of Antibiotics and Antibiotic Resistance Genes Along a Pollution Gradient of Ba River in Xi'an, China. *Front Microbiol*. 2018;9:3191.
9. Jia J, Guan Y, Cheng M, Chen H, He J, Wang S, Wang Z. Occurrence and distribution of antibiotics and antibiotic resistance genes in Ba River, China. *Sci Total Environ*. 2018;642:1136–44.

10. Peng FJ, Pan CG, Zhang NS, Ter Braak CJF, Salvito D, Selck H, Ying GG, Van den Brink PJ. Benthic invertebrate and microbial biodiversity in sub-tropical urban rivers: Correlations with environmental variables and emerging chemicals. *Sci Total Environ.* 2019;709:136281.
11. 10.1186/gb-2011-12-6-r60  
Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.*2011; 12(6):R60. Published 2011 Jun 24. doi:10.1186/gb-2011-12-6-r60.
12. Youssef NH, Rinke C, Stepanauskas R, Farag I, Woyke T, Elshahed MS. Insights into the metabolism, lifestyle and putative evolutionary history of the novel archaeal phylum 'Diapherotrites'. *ISME J.* 2015;9(2):447–60. doi:10.1038/ismej.2014.141.
13. Vukovic Z, Markovic L, Radenkovic M, Vukovic D, Stankovic S. Heavy metal and bacterial pollution of the Sava River in Serbia. *Arh Hig Rada Toksikol.* 2011;62(1):11–6.
14. Wang J, Wei H, Zhou X, Li K, Wu W, Guo M. Occurrence and risk assessment of antibiotics in the Xi'an section of the Weihe River, Northwestern China. *Mar Pollut Bull.* 2019;146:794–800.
15. Whittaker RH. Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol Monogr.* 1960;30:279–338. doi:10.2307/1943563.
16. Whittaker RH. Evolution and Measurement of Species Diversity. *Taxon.*1972. 21: 213–251. doi:10.2307/1218190.
17. Zhang H, Wang Y, Chen S, Zhao Z, Feng J, Zhang Z, Lu K, Jia J. Water Bacterial and Fungal Community Compositions Associated with Urban Lakes, Xi'an, China. *Int J Environ Res Public Health.*2018; 15(3).
18. Zhou J, et al. "Random sampling process leads to overestimation of beta-diversity of microbial communities." *mBio.*2013; 4(3): e00324-00313.

## Figures

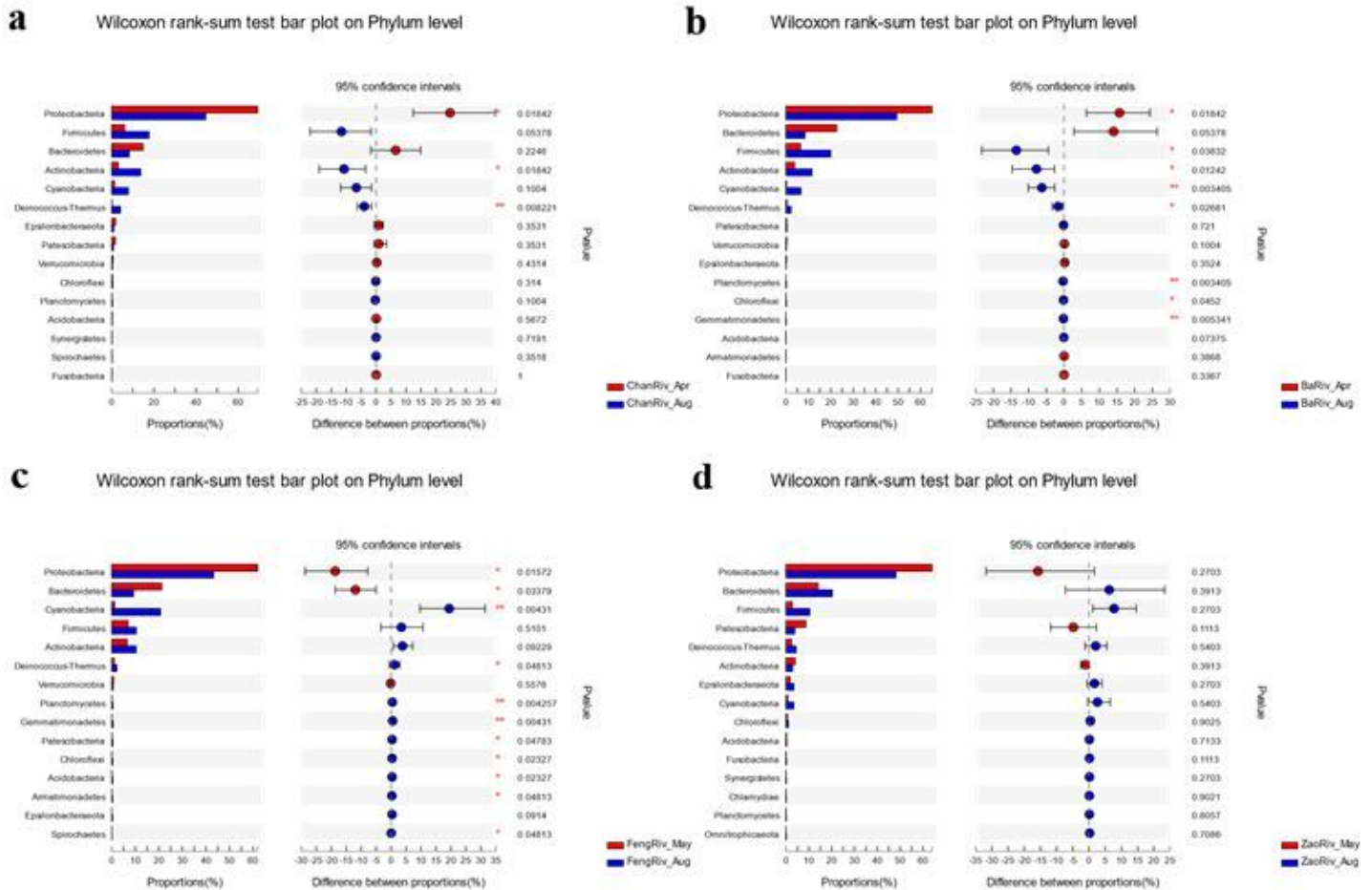


**Figure 1**

Map of sampling locations in Xi'an of China Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

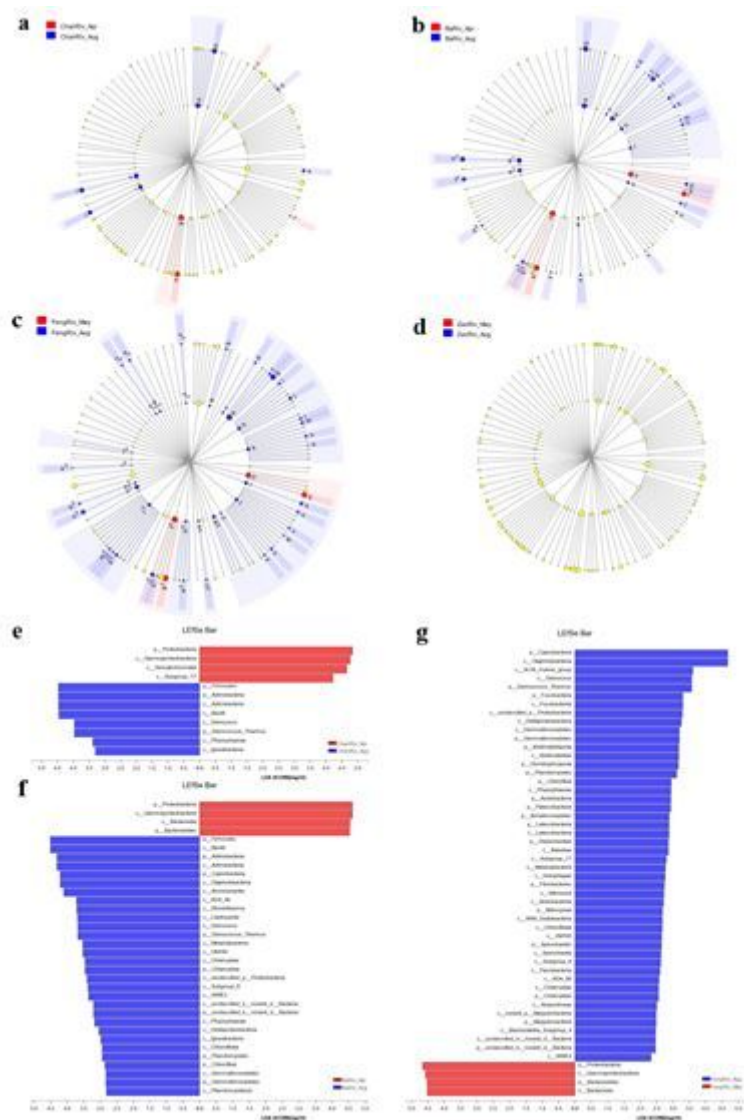






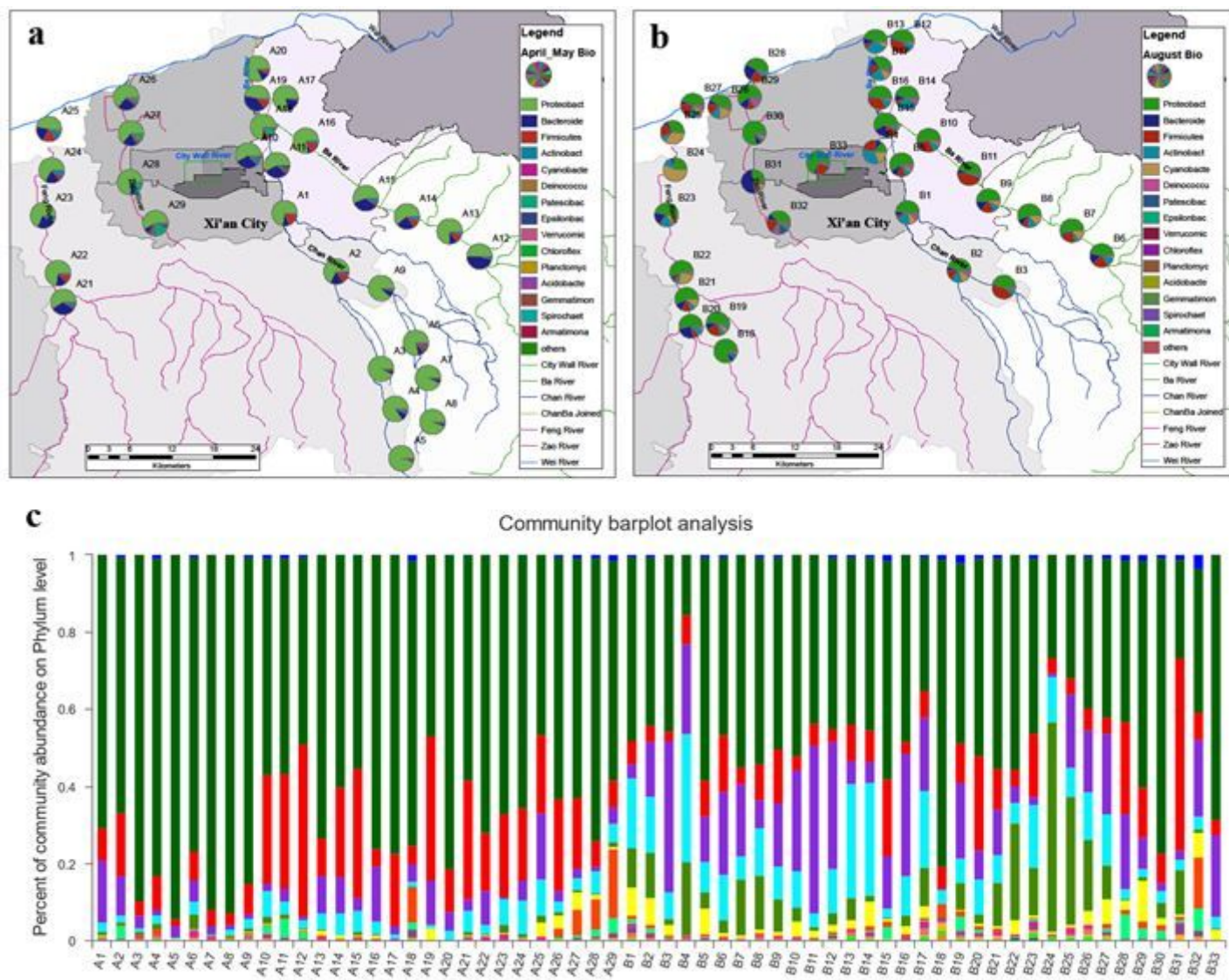
**Figure 3**

Comparison of the first 15 major phyla between groups of different sampling periods for the same river segment: a) and b) are samples from April and August of 2019 for the Chan and Ba Rivers, c) and d) are samples from May and August of 2019 for the Feng and Zao River.



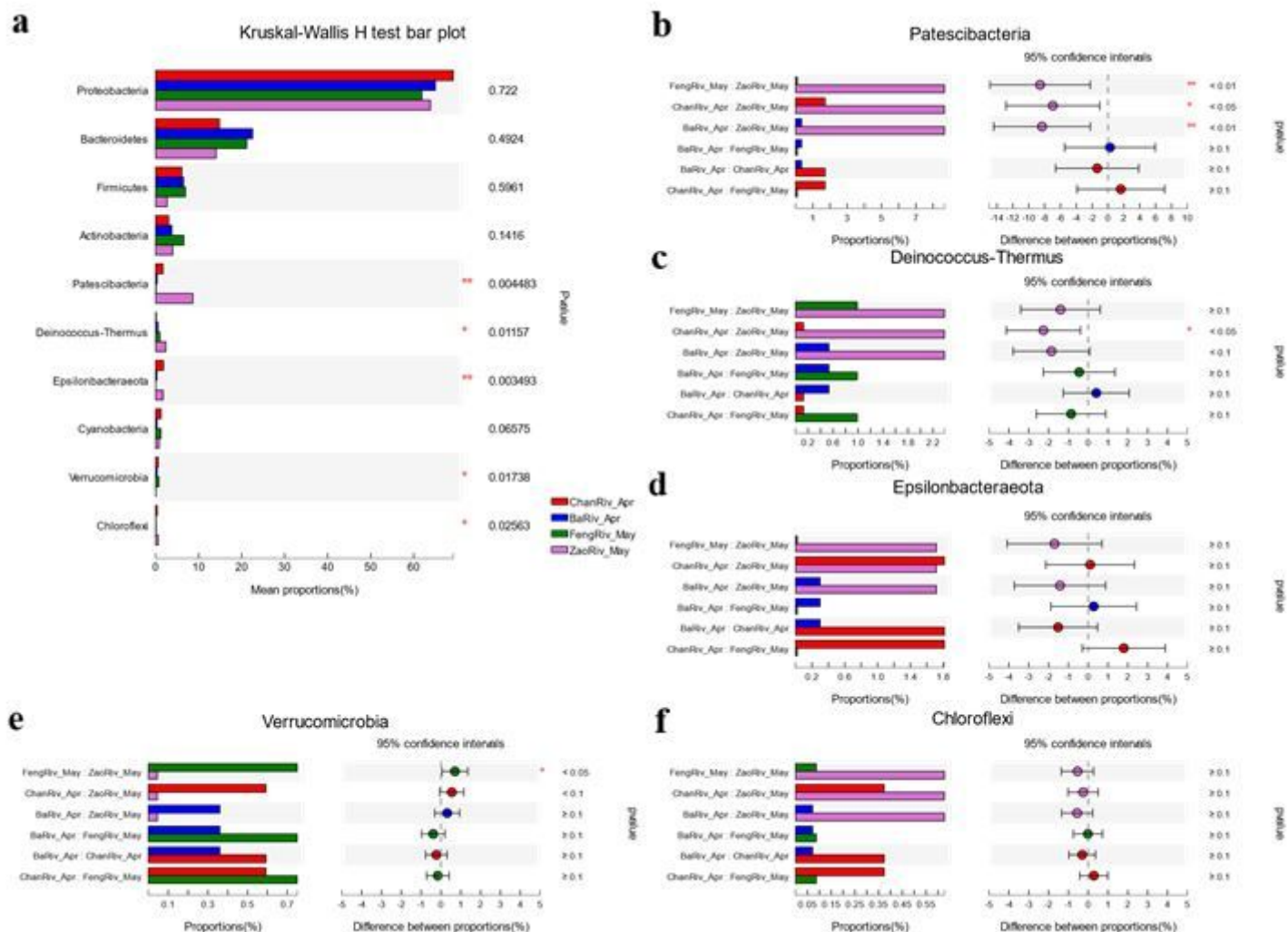
**Figure 4**

LefSe analysis for seasonal changes of microbial communities in each river segment. (a-d) are cladograms of taxonomic representation of statistically and biologically consistent differences between sample groups for the Chan, Ba, Feng, and Zao Rivers. Differences are represented in the color of the most abundant classes as shown in the legend of each figure. Each circle's diameter is proportional to the taxon's abundance. (e-g) Histograms of the LDA scores computed for phylum and class differentially abundant between April/May and August of 2019 samples for the Chan, Ba, and Feng Rivers, the LefSe analysis did not identify any statistically differential phylum or class between May and August samples in the Zao River. LefSe scores can be interpreted as the degree of consistent difference in relative abundance between features in the two classes of analyzed microbial communities. The histogram thus identifies which clades among all those detected as statistically and biologically differential explain the greatest differences between communities.



**Figure 5**

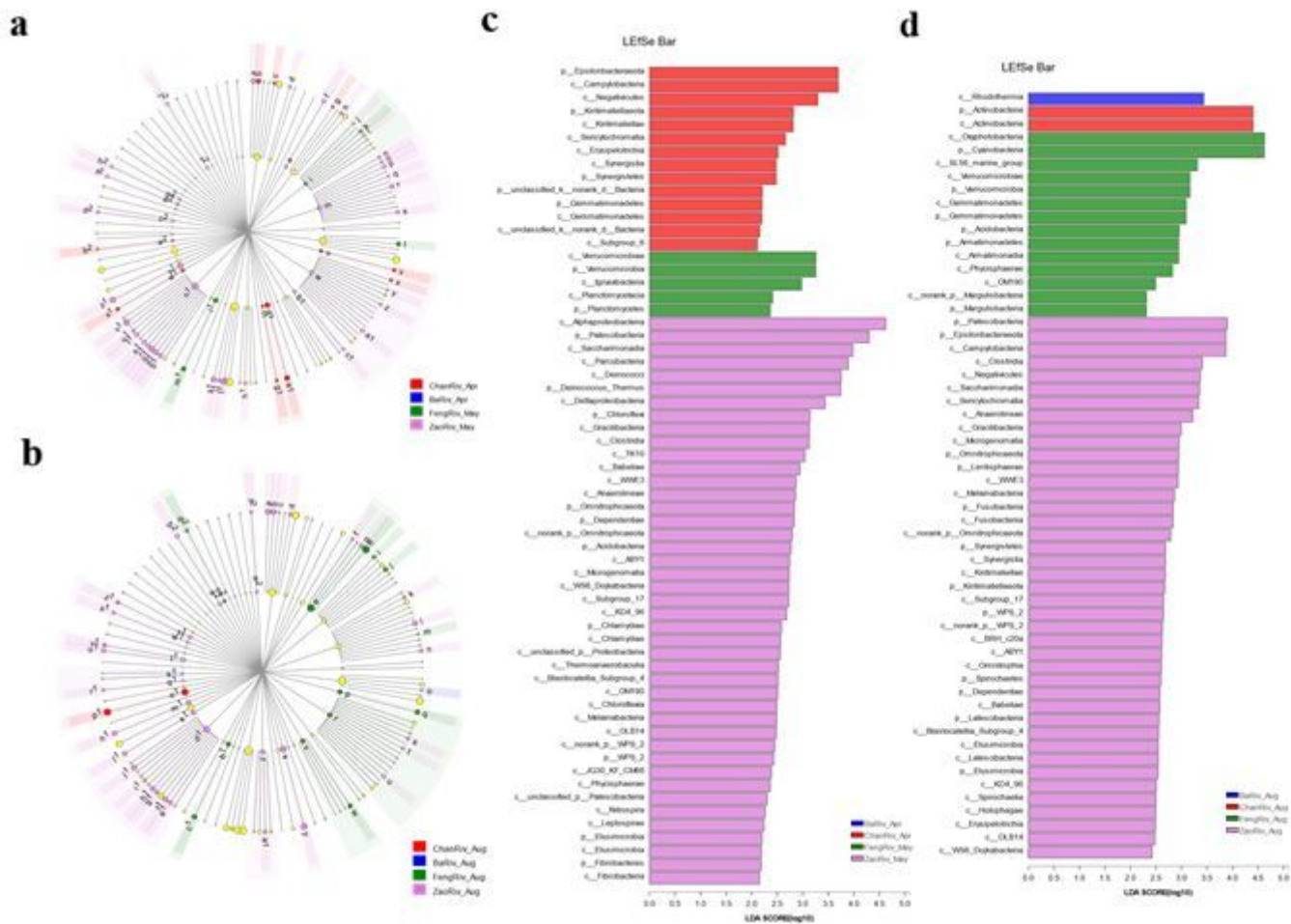
Relative abundances of the microbial community on phylum level for individual samples: 2a) microbial community abundance pie charts for all April/May samples, 2b) microbial community abundance pie charts for all August samples, 2c) barplot of microbial communities for major phyla with relative abundance greater than 1%.



**Figure 6**

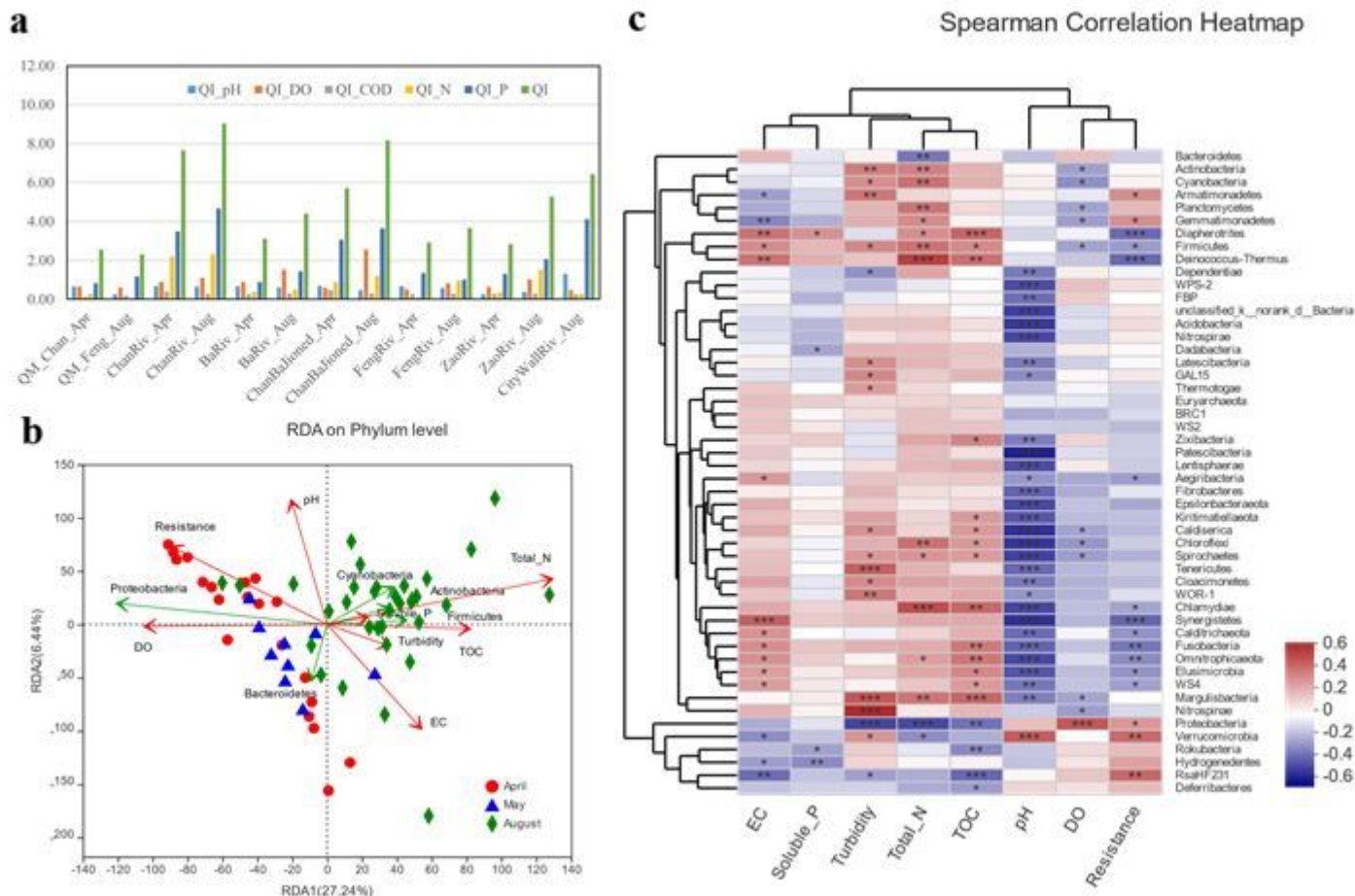
Kruskal-Wallis H test results for the statistical significances of major phyla between the Chan, Ba, Feng, and Zao Rivers in April/May: (a) Histogram of the mean proportions for the top ten major phyla, (b–f) are histograms of the mean proportions of phylum and difference between proportions for Patescibacteria, Deinococcus\_Thermus, Epsilonbacteraeota, Verrucomicrobia, and Chloroflexi, respectively, which had statistically significant differences between the compared rivers.





**Figure 7**

LEfSe analysis of the differential abundances of microbial phylum and class between the Chan, Ba, Feng, and Zao Rivers. a) and b) are cladograms of phylum and class for the April/May samples and August samples, respectively; c) and d) are phyla and classes of differentially abundant typically in each stream for the April/May samples and August samples, respectively.



**Figure 8**

Influences of environmental variables to microbial communities: a) water quality indices from the pH, Do, nitrogen, phosphorous, and COD, the composite index for all sample groups, b) Redundancy Analysis for correlation of environmental variables with microbial communities, c) Heatmap of Spearman's Rank Correlation between microbial community and environmental variables

# Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [WaterResearchXian08252020additionalfiles.docx](#)