



OPEN

Seasonal and spatial patterns of eukaryotic phytoplankton communities in an urban river based on marker gene

Jing Yang¹, Junping Lv¹, Qi Liu¹, Fangru Nan¹, Bo Li², Shulian Xie¹ & Jia Feng¹✉

The seasonal and spatial eukaryotic phytoplankton composition in the Fenhe River was investigated based on the 18S rDNA V4 region. The relationship between phytoplankton functional groups and environmental factors was explored to effectively capture the responses of these taxa to environmental gradients and their effects on ecosystem function. Our results indicated that the Chlorophyta and Bacillariophyta had higher relative abundance than other taxa, and their diversity and richness indices in spring were higher than those in other seasons. The linear discriminant analysis effect size (LEfSe) analyses detected that the potential seasonal biomarkers included *Desmodesmus*, *Cyclotella*, *Pseudoschroederia*, *Discostella*, *Scenedesmus*, *Monoraphidium*, and *Nannochloropsis*; the spatial biomarkers included *Amphora*, *Neochloris*, *Hindakia*, *Pseudomuriella*, *Coccomyxa*, *Chloroidium*, *Scherffelia*, *Chromochloris*, and *Scotinosphaera*. The systemic evolution and distribution characteristics of the first 50 representative sequences showed that the dominant genus included *Desmodesmus* in spring, *Pseudopediastrum* in summer, *Mychonastes* in autumn, and *Monoraphidium* in winter. Main seasonal variation of phytoplankton functional groups was as follows: spring (J + F + C + X1) → summer (J + F + X1 + X2) → autumn (J + F + X1 + C) → winter (X1 + J + B + X2). Pearson correlation, redundancy analysis, and variance partitioning analysis showed temperature and phosphate were the determining factors causing the changes of phytoplankton functional groups and community composition in the Fenhe River.

The process of urbanization can affect water bodies both qualitatively and quantitatively and can lead to a series of processes that are collectively referred to as “urban river syndrome”¹. The increase of population density and the changes of land cover can threaten basic ecosystem services², causing unprecedented loss of habitat and biodiversity, affecting public health, and posing challenges to water management³. In order to deal with these impacts, new monitoring methods and public participation are needed to restore the ecology and integrate environments with society and economy⁴. The Fenhe River is the second largest tributary of the Yellow River in China; it is located in Shanxi Province, which has a semi-arid climate. Over the past few decades, streamflow has been continuously decreasing because of the impact of human activities and climate change, and the river ecosystems are experiencing a rapid loss of biodiversity and ecosystem integrity, thus have become one of the subjects that have been extensively studied^{5,6}. However, our understanding of the geographical distribution pattern of microbial communities and the ecological factors affecting the diversity in urban river is still very limited.

Phytoplankton is an important organismal group that can indicate the changes in aquatic environment because of its sensitivity to the changes in physical, biological, and chemical structure of aquatic ecosystems^{7,8}, they mainly involve in the degradation of natural and man-made residues and nutrient cycling, and play an important role in many biogeochemical processes⁹. Previous studies have shown that phytoplankton does not usually show strong diffusion restrictions^{10,11}, indicating that its distribution is more likely influenced by local environmental variables. These variables can include nutrient concentrations, light conditions¹², water temperatures¹³, pH, and predation by grazers and zooplankton¹⁴, all of which are the main factors affecting the growth and reproduction of algae in water bodies. Environmentally driven changes in phytoplankton diversity can also have a significant impact on the stability and function of food webs. Therefore, the understanding of factors determining the phytoplankton communities has become one of the biggest challenges in ecology.

¹School of Life Science, Shanxi University, Taiyuan 030006, China. ²School of Geographical Science, Taiyuan Normal University, Jinzhong 030619, China. ✉email: fengj@sxu.edu.cn

Environmental variable	Season	$\bar{x} \pm s$ (s.d.)	F-value	Anova P-value
Water temperature (°C)	Spring	20.283 ± 1.053	1573.605	0.001
	Summer	22.983 ± 0.256		
	Autumn	21.150 ± 0.259		
	Winter	3.167 ± 0.234		
Nitrate (mg/L)	Spring	1.572 ± 0.264	9.172	0.001
	Summer	1.268 ± 0.551		
	Autumn	0.735 ± 0.620		
	Winter	0.341 ± 0.169		
Nitrite (mg/L)	Spring	0.035 ± 0.006	37.247	0.000
	Summer	0.145 ± 0.034		
	Autumn	0.043 ± 0.024		
	Winter	0.035 ± 0.096		
Phosphate (mg/L)	Spring	0.095 ± 0.009	18.223	0.000
	Summer	0.160 ± 0.033		
	Autumn	0.062 ± 0.027		
	Winter	0.059 ± 0.031		
DOC (mg/L)	Spring	6.770 ± 0.220	176.402	0.000
	Summer	6.198 ± 0.457		
	Autumn	6.282 ± 0.293		
	Winter	1.181 ± 0.773		

Table 1. One-way ANOVA analysis of physiochemical parameters in seasonally. $\bar{x} \pm s$ (s.d.) is expressed as mean \pm standard deviation, and *P*-value less than 0.05 is statistically significant.

The development of molecular biotechnological technology based on amplicon pyrosequencing has provided a new way to study phytoplankton communities in environments. The small subunits of the ribosome RNA gene (18S rRNA gene) are present in a wide range of living cells and are highly conserved. These subunits contain conserved regions and variable regions¹⁵, making the 18S rRNA gene one of the most commonly used genes. The structure of the 18S rRNA gene has been described as containing nine highly variable regions, V1 to V9, of which the V4 and V9 regions are commonly used in environmental studies¹⁶. A study has suggested that species richness and community composition estimated based on the V4 region are similar to those estimated using full-length ribosomal RNA genes¹⁷; therefore, the V4 region is suitable for environmental studies.

Given that phytoplankton is a highly diverse group of photosynthetic organisms¹⁸, functional groups are commonly used to classify phytoplankton species to avoid losing important characteristics and responses^{19–21}. The definition of functional groups can bring together different species with common functional characteristics that respond to environmental changes in a similar way and have the same impact on ecosystem function²². Besides, the functional groups can more validly assess the response of phytoplankton communities to the environmental changes²³. Therefore, in biodiversity research, the combination of species diversity and its functions is particularly useful for the understanding of the patterns of diversity and their underlying mechanisms.

In this study, we investigated the phytoplankton diversity and functional groups using 18S rDNA V4 region, with related physical and chemical variables (including water temperature, carbon, nitrogen, and phosphorus) in the Fenhe River. We combined the functional group theory with the phytoplankton diversity succession and attempted to reveal the biogeographical distribution and seasonal migration of phytoplankton and identify possible driving mechanisms based on environmental factors and phytoplankton functional groups in the Fenhe River.

Results

Effects of seasonality on physiochemical properties of water. The characteristics of the physical and chemical factors of the river are shown in Table 1. Variance analysis showed that the environmental conditions during the four seasons were significantly different ($P < 0.01$), indicating the seasonal changes can affect the water quality in the Fenhe River. Although the spatial changes were not significantly different, the samples from downstream sites were separated from those from the upstream and midstream sites during the four seasons (Fig. 1). In addition, the levels of nitrite, phosphate, and DOC in the upstream sites were lower than those in the downstream sites, suggesting that the water quality in the former was better than that in the latter.

Phytoplankton diversity and richness analysis. The total raw reads of all samples ranged from 782,712 (autumn) to 1,547,327 (summer). After removing chimera, the number of reads in each sample ranged from 41,526 to 85,363, with a mean value of 61,629. The temporal and spatial richness estimates and diversity indices are shown in Fig. 2. The Simpson diversity index ranged between 0.08 and 0.35 with an average value of 0.15 ± 0.06 , and the Shannon diversity index varied from 1.86 to 3.17 with a mean value of 2.65 ± 0.37 . The Shannon and Simpson diversity indices of winter were significantly different ($P < 0.001$) from those of other seasons, whereas the Chao and ACE indices of all the four seasons were significantly different ($P < 0.001$). Over-

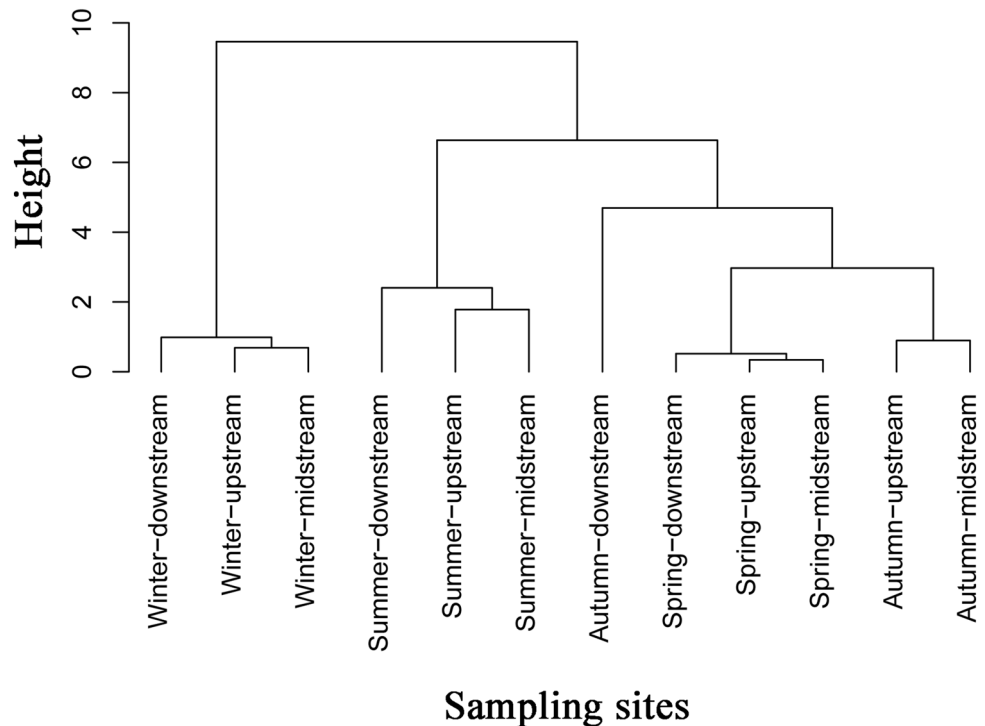


Figure 1. Dendrogram of environmental variables data with hierarchical cluster analysis based on the Euclidean distance and Ward linkage. Each season is divided into upstream, midstream, and downstream in spatially.

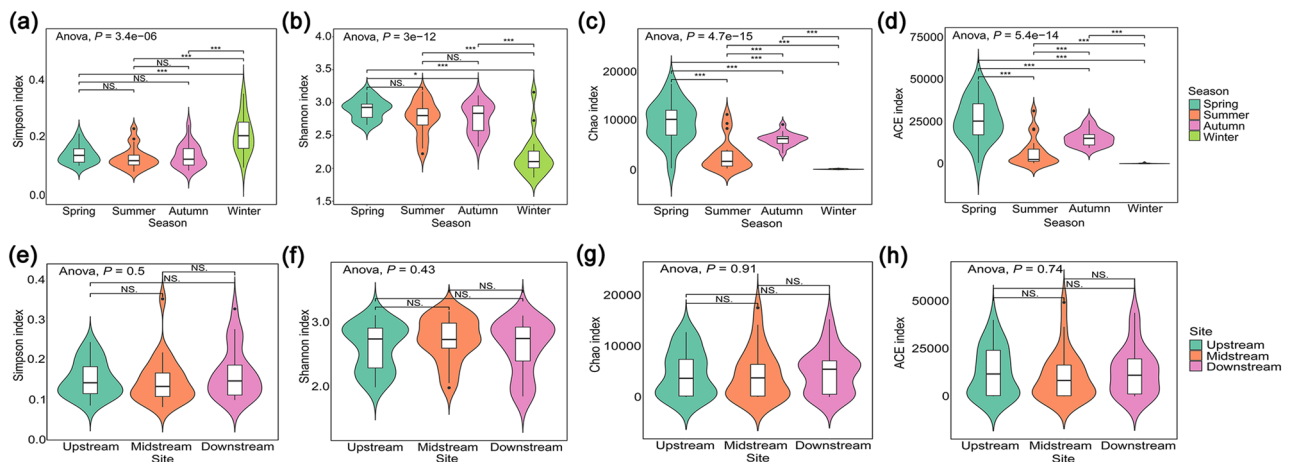


Figure 2. Boxplot of diversity index and richness based on one-way ANOVA. Upper row is the season variation, lower row is the spatial variation. Significant P -values in post-hoc test are designated with star notation: *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, and NS not significant.

all, the diversity value in summer and autumn was equivalent, while that in winter was lowest due to its lower temperature, which is not suitable for the growth of algae. Although there were no obvious differences between the sampling sites, the upstream sites had high Shannon index but low Simpson index in spring, indicating that phytoplankton communities in these sites had high diversity.

Differences in phytoplankton communities analyzed based on LEfSe. The seasonal and spatial differentially abundant taxa (i.e., the biomarkers) from the phylum to the genus level identified by LEfSe analysis are shown in Fig. 3. The linear discriminant analysis (LDA) score describes the degree to which the relative abundance of various microbial groups in given microbial communities consistently changes between seasons²⁴. Our results emphasized that family and genus were main the potential biomarkers in different seasons were identified and found four to seven phytoplankton taxa unique to each season. The phytoplankton groups that

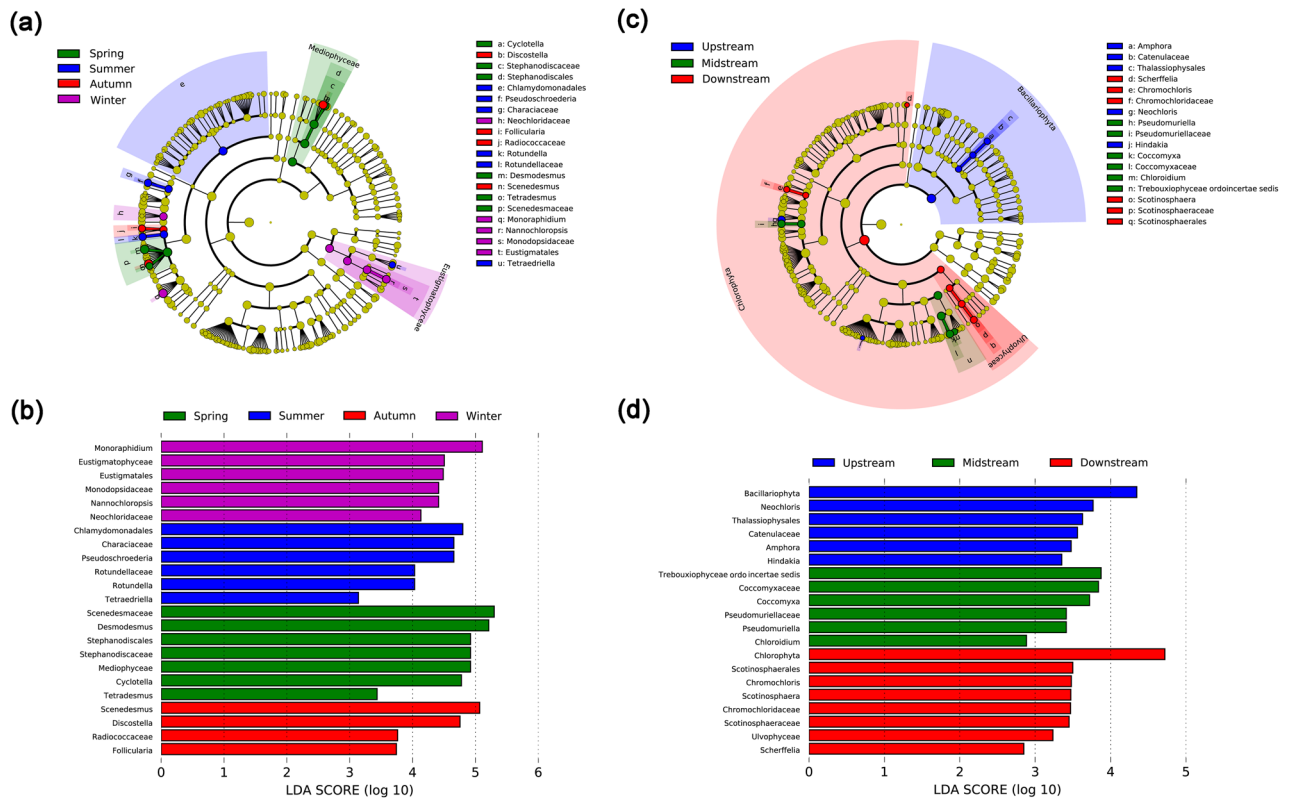


Figure 3. Linear discriminant analysis effect size (LEfSe) of the eukaryotic phytoplankton communities with an LDA score higher than 2.0 and *P* values less than 0.05. Each successive circle represents a phylogenetic level. Colour regions indicate taxa enriched in the different plant compartments. Bar graph shows LDA scores for phytoplankton taxa. Only taxa meeting an LDA significant threshold >2.0 are shown. (a,b) reveal the season, (c,d) reveal the sampling sites.

were enriched in spring included Mediophyceae, Stephanodiscales, Scenedesmaceae, *Desmodesmus*, and *Cyclotella*, and their LDA scores were higher than 4.5. The order Chlamydomonadales, the family Characiaceae, and the genus *Pseudoschroederia* were the biomarkers that preferred summer; their LDA scores were >4.5. Fewer phytoplankton groups were enriched in autumn, *Discostella* and *Scenedesmus*; their LDA scores were higher than 4.5. The taxa enriched in winter samples mainly included Eustigmatales at the order level, Neochloridaceae and Monodopsidaceae at the family level, *Monoraphidium* and *Nannochloropsis* at the genus level. In spatial analysis, the biomarkers enriched at the upstream sites included the Thalassiosphaerales at the order level, Catenuclaceae at the family level, *Amphora*, *Neochloris*, and *Hindakia* at the genus level. Trebouxiophyceae ordincertae sedis (order), Coccomyxaceae and Pseudomuriellaceae (families), *Pseudomuriella*, *Coccomyxa*, and *Chlorodinium* (genus) were enriched in the midstream sites. The downstream sites were mainly enriched with the members of the phyla Chlorophyta, the order Scotinosphaerales, the families Chromochloridaceae and Scotinosphaeraeae, and the genera *Scherffelia*, *Chromochloris*, and *Scotinosphaera*.

Phylogenetic analysis and distribution characteristics of phytoplankton communities. In this study, the first 50 representative sequences of samples collected in winter were different from those of samples collected in other seasons. Therefore, two phylogenetic evolution trees were constructed (Figs. 4 & 5). Firstly, we used the 50 most abundant OTUs to blast against the NCBI database, downloaded each of the most relevant sequences with the clearest taxonomic annotations, and then clustered them with our 50 OTUs. As shown in Fig. 4, 36 OTUs in Chlorophyta formed a monophyletic clade with sequences belonging to 27 genera. In Bacillariophyta, 7 OTUs (OTU1268, OTU708, OTU938, OTU941, OTU1252, OTU1170, and OTU1096) formed a monophyletic clade with sequences belonging to 2 genera, *Cyclotella* and *Discostella*, indicating that the species in these two genera were highly abundant. At the same time, we found that OTU1170 had the closest relationship with the *Peridiniopsis jiulongensis* H.Gu (now is accepted taxonomically as *Unruhdinium jiulongense* (H.Gu) Gottschling) diatom endosymbiont, which suggests that OTU1170 may be a *Peridiniopsis* diatom endosymbiont sequence. Besides, OTU1170 was clustered with OTU1252 and OTU1096, which belonged to *Discostella*. Previous studies have shown that the endosymbionts of *Peridiniopsis* may originate from the *Discostella*-like species²⁵, indicating the two are closely related. For the phylogenetic analysis of Ochrophyta, 7 OTUs were closely related to the genera such as *Nannochloropsis*, *Eustigmatos*, *Tetraedriella*, and *Goniochloris*, which have not or hardly been reported to be present in the Fenhe River, reflecting that there are substantially more species of Ochrophyta in this river.

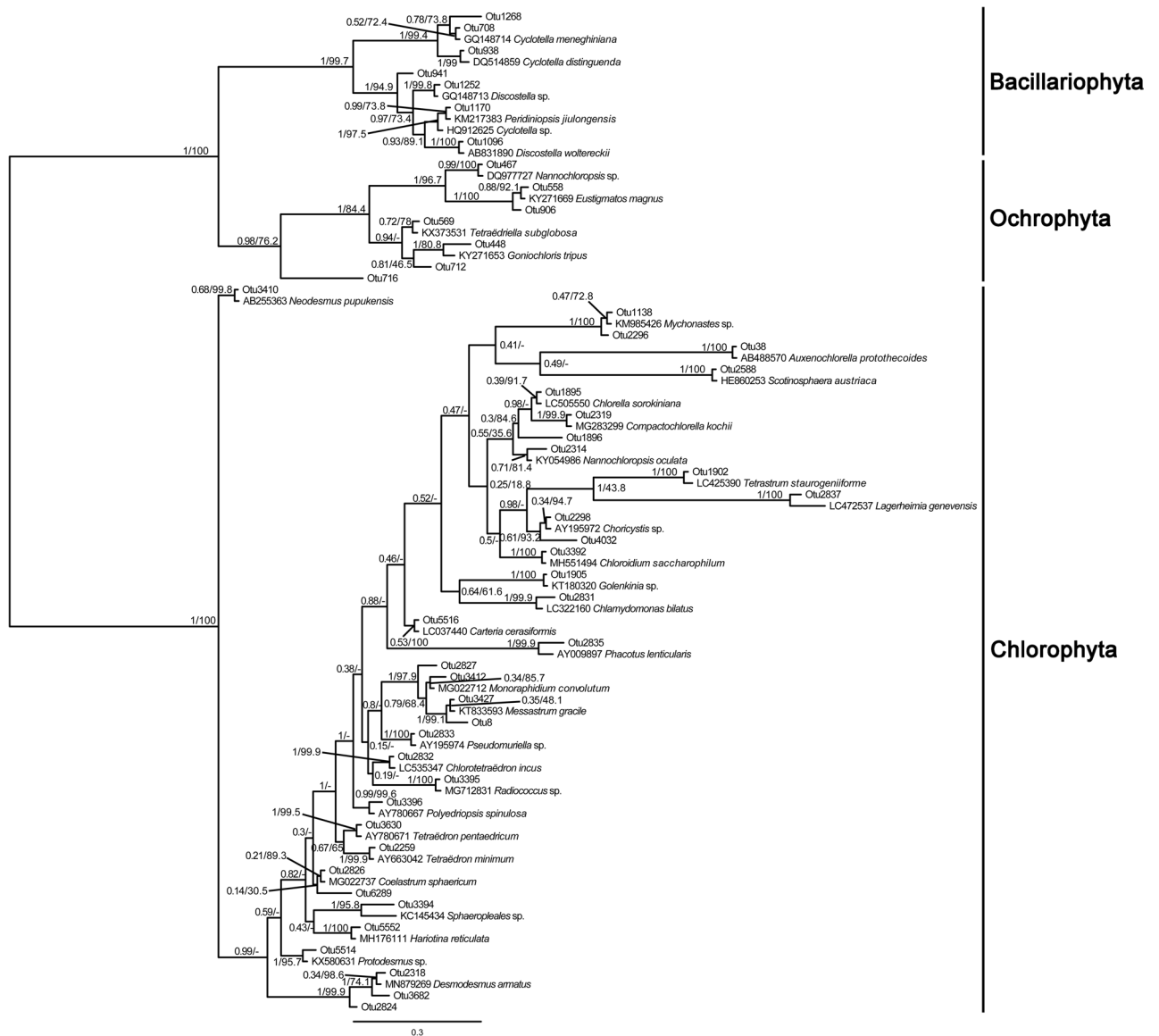


Figure 4. Phylogenetic analysis of the 50 most abundant OTUs in spring, summer, and autumn. Numbers on the left side at the branches represent Bayesian test support values and right side at the branches represent maximum-likelihood bootstrap values.

As illustrated in Fig. 5, the abundance of some genera such as *Neochloris*, *Botryosphaerella*, *Schroederia*, and *Chromochloris* was higher in winter than in other seasons. For Ochrophyta, 6 OTUs formed the monophyletic clade with sequences belonging to 4 genera, including *Vischeria*, *Nannochloropsis*, *Pseudotetraëdriella*, and *Vacuoliviride*. Two OTUs from Cryptophyta were clustered with *Cryptomonas obovoidea* Pascher and *Hemiselmis cryptochromatica* C.E.Lane & J.M.Archibald, respectively, with a support rate of 100%.

To find the most widely distributed genera in the Fenhe River, a heatmap showing the relative abundance of 50 genera was plotted (Fig. 6). The heatmap clearly showed that the genus *Desmodesmus*, *Cyclotella*, and *Fragilaria* had the highest relative abundance and were most widely distributed in spring, which is in agreement with the occupancy-abundance relationship. In summer, the average proportions of *Pseudomuriella*, *Pseudopediastrum*, *Rotundella*, and *Oocystis* were higher than those of other genera at the same sites. *Mychonastes*, *Scenedesmus*, *Neodesmus*, *Follicularia*, and *Schroederia* were widely distributed in autumn. The *Cryptomonas* genus at site S2 reached its peak during winter.

Relationships between phytoplankton communities and physicochemical parameters. The dynamic pattern of phytoplankton communities along the Fenhe River can be influenced by the fluctuations of physical and chemical variables. In general, specific phytoplankton at the different taxonomic levels (mainly phylum, class, order, family, and genus) were related to certain environmental variables (see Supplementary Table S1 online). For instance, at the phylum level, the variations of most members of the Bacillariophyta and Ochrophyta were significantly associated with both physical and chemical properties ($P < 0.01$ and $P < 0.05$). Although the Chlorophyta had no significant correlation with these parameters, at the class level, Chlorophyceae

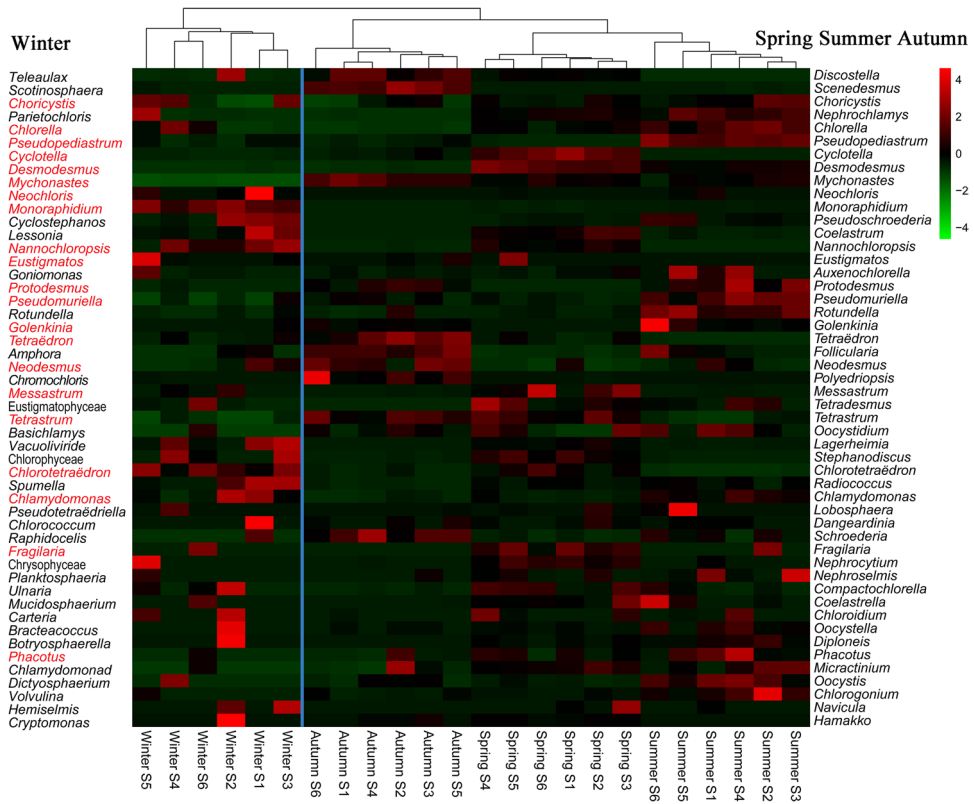


Figure 6. Heatmap of the 50 most abundant genera in all sampling sites based on the relative abundance. One column represents one sample, from left-most column to right-most column is the sites S1–S6 in winter, sites S1–S6 in autumn, sites S1–S6 in spring, sites S1–S6 in summer. The red font indicates the same algal genera in winter as in other seasons.

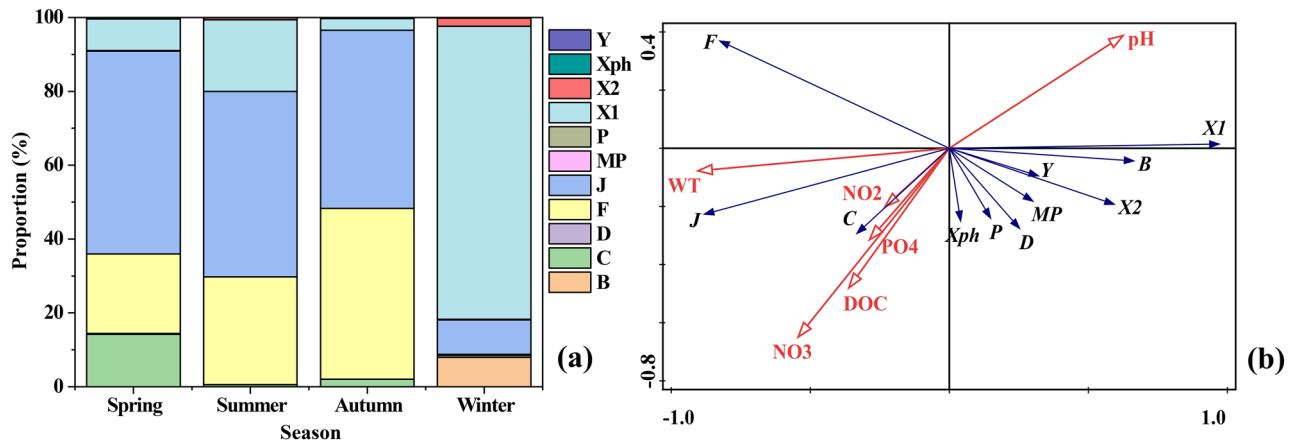


Figure 7. The proportion of the phytoplankton functional groups divided according to the reference^{19,36} during the four seasons in Fenhe River (a); RDA ordination diagram of the phytoplankton functional groups and environmental variables (b).

Phytoplankton functional groups and driving factors. The proportion of phytoplankton functional groups divided according to the references^{19,36} is shown in Fig. 7. Group J was dominant in spring, summer, and autumn, accounting for 54.97%, 50.25%, and 48.29%, respectively (Fig. 7a). Besides, group F also presented high proportions, its average proportion accounted for 21.54% in spring, 29.18% in summer, and 46.24% in autumn. Group X1 occurred during all seasons, however, the proportion was particularly high in winter. Groups B and X2 were also dominant in winter compared to other seasons.

The RDA ordination diagram of environmental variables and phytoplankton functional groups is shown in Fig. 7b. The first two axes explained 82.89% and 3.87% of the cumulative variance of the relationship of

Statistic	Axis 1	Axis 2	Axis 3	Axis 4
Eigenvalues	0.8289	0.0387	0.0098	0.0012
Cumulative % variance of species data	82.89	86.76	87.74	87.86
Species–environmental correlation	0.9807	0.6262	0.6230	0.2993
Cumulative % variance of species–environmental relationship	94.34	98.75	99.86	100

Table 2. Summary statistics for the first four axes of RDA.

Upstream and midstream data			Downstream data		
Variables	Explained %	<i>P</i>	Variables	Explained %	<i>P</i>
pH	41.8	0.008	Water temperature	39.4	0.047
Phosphate	15.6	0.036	Phosphate	36.5	0.012
Nitrite	10.1	0.024	pH	12.2	0.040
Water temperature	6.4	0.216	Nitrate	3.8	0.302
DOC	6.3	0.070	Nitrite	3.3	0.374
Nitrate	0.6	0.786	DOC	0.4	0.836
Unexplained	19.2		Unexplained	4.4	

Table 3. Spatial variance partitioning analysis between phytoplankton functional groups and environmental variables. Bold indicates a difference that is statistically significant ($P < 0.05$).

species–environmental variables, and the eigenvalues were 0.8289 and 0.0387, respectively (Table 2). The species–environment correlation values of axis 1 and axis 2 were 0.9807 and 0.6262, respectively. Monte Carlo permutation tests showed that environmental factors such as water temperature ($F = 46.3$, $P = 0.002$), nitrite ($F = 13.5$, $P = 0.002$), and phosphate ($F = 3.8$, $P = 0.014$) were the main factors regulating the phytoplankton functional groups. In Fig. 7b, axis 1 was positively correlated to pH while negatively correlated to water temperature. Axis 2 was negatively correlated with nitrite, phosphate, and DOC. Moreover, groups X1 were greatly affected by pH, and group J showed a strong positive correlation with temperature. Groups C and Xph showed different degrees of positive correlation with phosphate.

Cluster analysis based on environmental variables (Fig. 1) showed that the upstream and midstream samples were clustered into a branch, however, they were separated from the downstream samples, indicating that the upstream and midstream environments were similar, and showed certain environmental differences from the downstream. The variance partitioning analysis also showed the same spatial difference (Table 3). The pH could explain 41.8% ($P < 0.01$) of the changes in phytoplankton functional groups of upstream and midstream, followed by phosphate (15.6%) and nitrite (10.1%). Water temperature had the largest explanation rate (39.4%) ($P < 0.05$) to phytoplankton functional groups at the downstream of river, followed by phosphate (36.5%), and that the explanation rate of pH was relatively low (12.2%).

Discussion

High-throughput sequencing not only can rapidly and accurately reveal the information of a number of unknown taxa, but also can provide a better understanding of ecological factors that influence their distribution patterns²⁶. The 18S rDNA was selected as the target gene because it contains the variable regions that allow detailed classification and highly conserved regions that are required for primer annealing. In addition, it is easy to amplify because of its high copy numbers^{27,28}. Thus far, most research has focused on investigating marine environments, but freshwater ecosystems, which also have great biodiversity, are rarely studied. As a result, most sequences in the database are derived from marine microorganisms, causing many freshwater organisms to lack the corresponding OTUs in the SILVA 18S database, thus cannot be classified or assigned annotations. This emphasizes the importance of further research on freshwater environments. In our study, the eukaryotic phytoplankton composition and diversity in the Fenhe River were evaluated based on 18S rDNA, and the effects of seasonal physiochemical factors on the formation of the communities at different taxonomic levels were explored. The results showed the Fenhe River was environmentally heterogeneous, and mainly autotrophic organisms. We also discovered its high diversity value and temporal distribution pattern of phytoplankton. Although the spatial filters had no effect on the diversity, which is in contrary to our initial expectation, the samples from downstream sites were separated from those from the upstream and midstream sites during all the four seasons.

The Fenhe River runs across the urban, agricultural, and industrial areas; thus, the sources of pollutants in surface water at different sampling sites may be different^{6,29}. There is less human disturbance and more surrounding aquatic plants (such as reeds) in the upstream area, which have a purifying effect on water quality. Phytoplankton in this area is evenly distributed and has the highest diversity. The downstream area is highly urbanized, densely populated, and is closely related to the good economic development nearby²⁹. Domestic sewage is often a potential pollution source in this area, as a result, it had relatively low diversity. In the area in which we studied, the annual precipitation is 538.6 mm, which is concentrated in summer and autumn, but

there is less rainfall in spring and winter³⁰. In summer, due to the increase of temperature and rainfall, there are more tourists, irrigation and drainage from the surrounding farmlands, causing the water ecological environment to become greatly disturbed, and green algae are prone to outbreaks. After entering autumn, which is a dry season, the water flow velocity decreases; therefore, it is easy to cause individual phytoplankton species to become the dominant species, causing the diversity of phytoplankton and its community structure stability to decline. By contrast, the temperature drops sharply in winter, which is no longer suitable for the growth of most phytoplankton. Less rainfall and lower surface non-point source pollution in spring, and the temperature are suitable for the growth of phytoplankton, which resulting in the increase of its diversity.

Phytoplankton taxa that were significantly different (enriched) were identified by LEfSe analysis, and the frequency and relative abundance of the representative genera were presented by heatmap. Causing the change of phytoplankton communities might be closely associated with local environments. The phytoplankton functional groups have recently been considered as the effective indicator for the understanding of seasonal fluctuations of phytoplankton communities, as they can weaken the individual role of species and explain the response of the phytoplankton communities to environmental changes^{31,32}. It is more important and comprehensive than the response of the corresponding individual and population, thus can better reflect the characteristics of the habitat³³. Factors influencing phytoplankton functional groups usually include water temperature, nutrient availability, and light^{34,35}. The changes in water temperature not only can affect the metabolic rate of phytoplankton, but also are related to the degree of dissolution of various nutrients required for growth and reproduction of phytoplankton in water body.

The dominant species in the Fenhe River have both crossover and succession with the seasonal changes. The seasonal succession of phytoplankton functional groups was as follows: spring (J + F + C + X1) → summer (J + F + X1 + X2) → autumn (J + F + X1 + C) → winter (X1 + J + B + X2). Among the main dominant genus (see Supplementary Table S2 online), *Fragilaria* belonging to P group has the highest frequency throughout the year, the P group tends to be strong light and rich nutrients, thus is generally enriched in well-ventilated, nutrient-rich lakes, as well as in the epilimnion of stratified shallow lakes³⁶. Group C contained central diatoms, for example, *Cyclotella* sp. (mainly *Cyclotella meneghiniana* Kützing), which was a representative genus with relatively high abundance in spring, which can adjust to small and constantly mixed eutrophication to mesotrophic habitats^{19,36}. The growth of most diatoms depends on the availability of nutrients, including carbon, nitrate, silicon, and phosphate, while the increase of TN/TP ratio is related to a higher biomass of group C³⁷. The significant increase of carbon and nitrogen concentrations in spring may explain the changes in phytoplankton communities observed here. An increase of nitrogen is usually related to a sudden input of wastewater. As a complex mixture of compounds, DOC is a critical component of carbon transfer from terrestrial to freshwater ecosystems, as well as from sources to marine areas³⁸. It has been determined that the subtle changes of DOC concentrations can cause shifts in the structure of aquatic microorganisms³⁹. One of the main terrestrial inputs of DOC comes from vegetation, such as the decomposition of organic litter and the direct input of submerged plants⁴⁰, which can release a large amount of DOC into rivers.

It is well known that the trophic status and nutrient concentrations of lakes have different effects on the types of microorganisms⁴¹. Nitrogen and/or phosphorus are usually the factors that limit the growth of phytoplankton⁴². Chlamydomonadales mainly participate in the remineralization of nitrogen and phosphate. In our study, the family Chlamydomonadaceae and the genus *Chlamydomonas* belonging to X2 group showed significant negative correlation with temperature and phosphate, and had high abundance in winter, which was able to tolerate stratification, but was sensitive to mixing and filter-feeding grazers. Group X1 included small and slender single-celled green algae (*Monoraphidium* sp. and *Chlorella* sp.), they could adjust to low light and were sensitive to nutrient deficiency¹⁹. Correlation analysis and RDA analysis showed significant negative relationships between group X1 and nutrients (nitrogen and phosphorus) concentrations. The widespread distribution pattern of *Cryptomonas* belonging to Y group can be related to its growth in a variety of environments, as it has been found existed in rivers, reservoirs, and lakes. At the same time, the distribution of this genus does not have any seasonal specificity, because it could occur in the cool winter or bloom in summer⁴³. Besides, it is mainly distributed in moderately enriched systems with a high surface, volume ratio, relatively fast growth, and rapid phosphorus uptake rates⁴⁴. According to the literature⁴⁵, the Y group is tolerant to high light attenuation coefficient values, suggesting that it can adjust to the environments with insufficient light. In our study, this group was found in low light conditions during winter. Other studies have also shown that *Cryptomonas* has the ability to enhance nutrient absorption through mixotrophy, the mechanism that enables it to grow and reproduce under a light-limited environment⁴⁶. In addition, it has strengthened competitiveness due to its flagella, which allows it to migrate vertically between water layer with optimal light conditions and nutrient concentrations.

Organisms often have an indicating effect on temperature changes, especially for those residing in aquatic environments, where the range of temperature changes is smaller than that on land, which causing aquatic organisms lack a temperature compensation mechanism, therefore, there are more thermophilic organisms. Correlation analysis demonstrated that diatoms, especially for Bacillariophyta and *Discostella*, were sensitive to temperature changes. The class Chlorophyceae, the order Sphaeropleales, the family Mychonastaceae, and the genus *Mychonastes* and *Tetrastrum* have specific requirements for environmental factors (high temperature, high contents of carbon, nitrogen, and phosphorus), all of which belonged to a widespread species group. The Sphaeropleales (“Chlorococcales”) is a diverse taxon that responds positively to increasing thermal stability by adopting the inclusive range of morphological adaptations to contrast large subsidence losses, for example, small size and mucus formation^{47,48}. However, other organisms residing in a cold environment, *Chlorotetraëdron*, *Neochloris*, and *Nannochloropsis* has lower requirements for nutrients and carbon, and are mainly distributed in areas with low temperature.

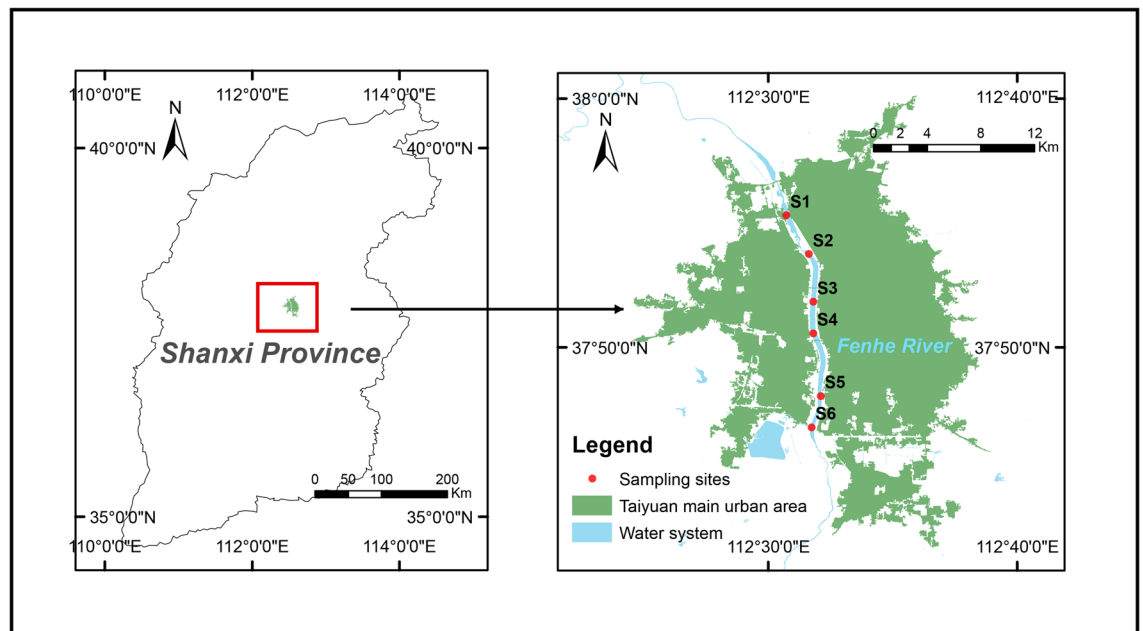


Figure 8. The location of the sampling sites in Fenhe River. The map is produced using GIS software ArcMap (version 10.2) (<https://developers.arcgis.com/>).

Conclusion

This study combined the diversity with functional groups to identify factors driving the seasonal composition of phytoplankton communities in the Fenhe River. The 18S rDNA V4 region revealed that the diversity and composition of phytoplankton varied with seasons, and the diversity and number of species in spring were significantly higher than that in other seasons. LEfSe analysis identified the potential seasonal and spatial biomarkers and detected four to seven phytoplankton taxa each season. Besides, analysis of the systemic evolution and distribution characteristics of the first 50 representative sequences showed that the dominant genus in spring was *Desmodesmus*, whereas that in summer, autumn, and winter was *Pseudopediastrum*, *Mychonastes*, and *Monoraphidium*, respectively. The phytoplankton functional groups have both crossover and succession with the seasonal changes. RDA analysis demonstrated that water temperature, nitrite, and phosphate were considered as the critical factors affecting this fluctuation. The variance partitioning result showed that pH had the largest explanation rate to phytoplankton functional groups of upstream and midstream, water temperature had the largest explanation rate to phytoplankton functional groups at the downstream of river. Pearson correlation analysis indicated that carbon and phosphorus were more likely the determinants than nitrogen in the process of shaping the phytoplankton communities in the Fenhe River. Taken together, these results greatly enhance our knowledge on the spatiotemporal dynamics and diversity of this ecologically significant but little-known urban river.

Methods

The Fenhe River is the largest river in Shanxi Province, with a catchment area of 39 471 km² and a total length of about 710 km. The study area is located in Taiyuan (latitude: 37°27'–38°25', longitude: 111°30'–113°09'), Shanxi Province, northern China (Fig. 8). The Taiyuan section of the Fenhe River is a length of 188 km, flows through Taiyuan from north to south, and receives domestic sewage and industrial wastewater discharged along the bank. At the same time, its upstream is also an important supply area for Taiyuan's drinking water source. The urban section of river starts from Chaicun Bridge in the north and reaches at Xiangyun Bridge in the south, with a total length of 20 km; and includes many factories and enterprises, such as coal mines, thermal power stations, coal preparation plants, etc. Therefore, a large amount of untreated industrial sewage is discharged into the river, which is often a potential source of pollution for water bodies in the area. According to statistics, the industrial production and residential water consumption in the Fenhe River is about 2.32×10^9 m³/a, which accounts for 46% of the total water consumption in Shanxi Province. And the city is densely populated and highly urbanized, with a population of about 4.5 million. In addition, the region has a semi-humid continental monsoon climate with an average annual precipitation of 538.6 mm. According to the Shanxi Provincial Hydrology and Water Resources Survey Bureau, the water flows of the Fenhe River range from 2.12 to 9.44 m³/s, and the water velocities range between 0.35 to 0.58 m/s. The six sampling sites selected in this study are mainly located in the upstream and downstream of the Taiyuan section of the river, as well as the industrial or urban wastewater discharge outlets. The map of Fenhe River is produced using GIS software ArcMap (version 10.2) (<https://developers.arcgis.com/>).

To investigate the influence of seasonal factors, water samples were taken in March (spring), July (summer), September (autumn), and December (winter) of 2019. The distribution of sampling sites is shown in Fig. 7. The phytoplankton community diversity of the water samples collected from each site was analyzed based on the V4

region of 18S rDNA, and the environmental parameters of the site were measured within 24 h. Water temperature was measured in situ using a multi-parameter water quality monitor, and the concentrations of nitrate (NO_3^-), nitrite (NO_2^-), and phosphate (PO_4^{3-}) were determined according to the APHA standard methods⁴⁹. The dissolved organic carbon (DOC) concentration was determined according to the method described by Bolan et al.⁵⁰. All experiments were performed in triplicate.

For the analysis of marker gene, 250 mL of water samples were concentrated using an isopore membrane filter (0.22 μm) and then stored at -80°C until subsequent DNA extraction. Genomic DNA was extracted using an E.Z.A.N. Mag-Bind DNA Kit according to the manufacturer's instructions. The V4 region of 18S rDNA was amplified using primers V4F (5'-GGCAAGTCTGGTGCCAG-3') and V4R (5'-ACGGTATCTRATCRTCCTT CG-3'), and the PCR amplification was performed according to Sun et al.⁵¹. The PCR amplification cycle was: the first round of amplification was 94°C for 3 min, followed by 5 cycles at 94°C for 30 s, 45°C for 20 s, 65°C for 30 s, then followed by 20 cycles at 94°C for 20 s, 55°C for 20 s, 72°C for 30 s, and a final extension at 72°C for 5 min; the second round of amplification was 95°C for 3 min, followed by 5 cycles at 94°C for 20 s, 55°C for 20 s, 72°C for 30 s, and a final extension at 72°C for 5 min. Two round PCR reactions were performed in a 30 μL reaction, containing 15 μL 2 \times Taq master Mix, 10 μM of forward and reverse primers, and 20 ng of genomic DNA or PCR products of first-round amplification, respectively.

The quality of raw data was controlled using Qiime 1.7.0. The barcode and adapter sequences were removed by Cutadapt⁵². The chimera and non-specific amplified sequences were removed by Usearch to obtain the effective tags. All the effective tags from each sample were clustered using Usearch software, and the sequences with 97% identity were clustered as Operational Taxonomic Units (OTUs). Selected the top 50 most abundant OTUs as the representative sequences and blast against the nt database in NCBI. Then download the most related sequence with clear taxonomy annotations, and cluster with our 50 OTUs. PHYML software and MrBayes version 3.1.2 were executed to construct the Maximum Likelihood and Bayesian trees, respectively.

The linear discriminant analysis (LDA) effect size (LEfSe) analysis was carried out using a normalized relative abundance matrix (<http://huttenhower.sph.harvard.edu/lefse/>) to detect the potential biomarkers. The LEfSe method based on the Kruskal–Wallis test was employed to identify features based on seasons and sites that were significant differences, and the LDA was used to assess the effect size of each feature. The significant P -value less than 0.05 and the LDA threshold score of 2.0 were used as criteria for identifying biomarkers. The diversity index and richness were analyzed using one-way analysis of variance (ANOVA). Post hoc comparisons were made using Tukey's method, a value of $P < 0.05$ was considered to be significant. Hierarchical cluster analysis of environmental variables based on the Euclidean distance and Ward linkage was conducted by R version 3.5.1 software. Heatmap generated using R version 3.5.1 software was used to detect the distribution of phytoplankton. SPSS 26.0 software was used to perform ANOVA variance analysis to identify the differences between seasonal variables and determine the Pearson correlations between the dominant groups at the different taxonomic levels and the physicochemical characteristics of water.

In this study, the taxa were classified into phytoplankton functional groups according to Reynolds et al.¹⁹ and Padišák et al.³⁶. The relationship between the phytoplankton functional groups and environmental factors was determined by multivariate analysis (Canoco 5.0). Before the analysis, all the phytoplankton functional groups and environmental factors (except for pH) were converted into $\log_{10}(x + 1)$. The detrended correspondence analysis (DCA) was employed to decide whether a linear or unimodal ordination method should be applied. The results demonstrated that the maximum value of the four-axis length gradient was 2.1, and linear model RDA was performed to identify the effects of environmental variables on phytoplankton functional groups. Prior to conducting the RDA analysis, Monte Carlo tests were used to screen out those environmental variables that had significant impacts on the phytoplankton functional groups. On the basis of the RDA analysis results, variance partitioning analysis (Canoco 5.0) was carried out to evaluate the spatial effects of the environmental parameters on the changes in phytoplankton functional groups.

Data availability

The datasets generated during and/or analysed during the current study are available in the NCBI repository (accession: PRJNA648254).

Received: 6 August 2021; Accepted: 10 November 2021

Published online: 30 November 2021

References

- Meyer, J. L., Paul, M. J. & Taulbee, W. K. Stream ecosystem function in urbanizing landscapes. *J. North. Am. Benthol. Soc.* **24**, 602–612 (2005).
- Garcia, X. et al. Placing ecosystem services at the heart of urban water systems management. *Sci. Total Environ.* **563**, 1078–1085 (2016).
- Varis, O., Biswas, A. K., Tortajada, C. & Lundqvist, J. Megacities and water management. *Int. J. Water Resour. D.* **22**, 377–394 (2006).
- Palmer, M. A., Filoso, S. & Fanelli, R. M. From ecosystems to ecosystem services: Stream restoration as ecological engineering. *Ecol. Eng.* **65**, 62–70 (2014).
- Liang, W. et al. Quantifying the impacts of climate change and ecological restoration on streamflow changes based on a Budyko hydrological model in China's Loess Plateau. *Water Resour. Res.* **51**, 6500–6519 (2015).
- Hua, K., Xiao, J., Li, S. J. & Li, Z. Analysis of hydrochemical characteristics and their controlling factors in the Fen River of China. *Sustain. Cities Soc.* **52**, 101827 (2020).
- Reynolds, C. S. *The Ecology of Phytoplankton* (Cambridge University Press, 2006).
- Wojciechowski, J., Heino, J., Bini, L. M. & Padial, A. A. Temporal variation in phytoplankton beta diversity patterns and metacommunity structures across subtropical reservoirs. *Freshw. Biol.* **62**, 751–766 (2017).
- Machado, K. B. et al. Diversity patterns of planktonic microeukaryote communities in tropical floodplain lakes based on 18S rDNA gene sequences. *J. Plankton Res.* **41**, 241–256 (2019).

10. Soininen, J., McDonald, R. & Hillebrand, H. The distance decay of similarity in ecological communities. *Ecography* **30**, 3–12 (2007).
11. De Bie, T. *et al.* Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecol. Lett.* **15**, 740–747 (2012).
12. Siegel, D. *et al.* Regional to global assessments of phytoplankton dynamics from the SeaWiFS mission. *Remote Sens. Environ.* **135**, 77–91 (2013).
13. Deng, Y. B. *et al.* Temporal and spatial dynamics of phytoplankton primary production in Lake Taihu derived from MODIS data. *Remote Sens.* **9**, 195 (2017).
14. Jia, J. *et al.* Identifying the main drivers of change of phytoplankton community structure and gross primary productivity in a river-lake system. *J. Hydrol.* **583**, 124633 (2020).
15. Hugerth, L. W. & Andersson, A. F. Analysing microbial community composition through amplicon sequencing: from sampling to hypothesis testing. *Front. Microbiol.* **8**, 1561 (2017).
16. Pawlowski, J. *et al.* Eukaryotic richness in the abyss: insights from pyrotag sequencing. *PLoS ONE* **6**, e18169 (2011).
17. Dunthorn, M., Klier, J., Bunge, J. & Stoeck, T. Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. *J. Eukaryot. Microbiol.* **59**, 185–187 (2012).
18. Borics, G., Abonyi, A., Salmaso, N. & Ptasnik, R. Freshwater phytoplankton diversity: models, drivers and implications for ecosystem properties. *Hydrobiologia* **848**, 53–75 (2021).
19. Reynolds, C. S. *et al.* Towards a functional classification of the freshwater phytoplankton. *J. Plankton Res.* **24**, 417–428 (2002).
20. Salmaso, N. & Padišák, J. Morpho-functional groups and phytoplankton development in two deep lakes (Lake Garda, Italy and Lake Stechlin, Germany). *Hydrobiologia* **578**, 97–112 (2007).
21. Kruk, C. *et al.* A morphological classification capturing functional variation in phytoplankton. *Freshw. Biol.* **55**, 614–627 (2010).
22. Kruk, C., Devercelli, M. & Huszar, V. L. Reynolds functional groups: A trait-based pathway from patterns to predictions. *Hydrobiologia* **848**, 113–129 (2021).
23. Kruk, C., Mazzeo, N., Lacerot, G. & Reynolds, C. Classification schemes for phytoplankton: A local validation of a functional approach to the analysis of species temporal replacement. *J. Plankton Res.* **24**, 901–912 (2002).
24. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, 1–18 (2011).
25. Zhang, Q., Liu, G. X. & Hu, Z. Y. Morphological differences and molecular phylogeny of freshwater blooming species, *Peridiniopsis* spp. (Dinophyceae) from China. *Eur. J. Protistol.* **47**, 149–160 (2011).
26. Leray, M. & Knowlton, N. Censusing marine eukaryotic diversity in the twenty-first century. *Philos. T. R. Soc. B.* **371**, 20150331 (2016).
27. Pawlowski, J. *et al.* CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *Plos Biol.* **10**, e1001419 (2012).
28. Guillou, L. *et al.* The protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* **41**, D597–D604 (2013).
29. Meng, Z. L., Yang, Y. G., Qin, Z. D. & Huang, L. Evaluating temporal and spatial variation in nitrogen sources along the lower reach of Fenhe River (Shanxi Province, China) using stable isotope and hydrochemical tracers. *Water* **10**, 231 (2018).
30. Yang, Y. G., Guo, T. T. & Jiao, W. T. Destruction processes of mining on water environment in the mining area combining isotopic and hydrochemical tracer. *Environ. Pollut.* **237**, 356–365 (2018).
31. Cao, J. *et al.* Succession of phytoplankton functional groups and their driving factors in a subtropical plateau lake. *Sci. Total Environ.* **631**, 1127–1137 (2018).
32. Long, S. X. *et al.* Responses of phytoplankton functional groups to environmental factors in the Pearl River. *South China. Environ. Sci. Pollut. Res.* **2**, 1–12 (2020).
33. Kruk, C. & Segura, A. M. The habitat template of phytoplankton morphology-based functional groups. *Hydrobiologia* **698**, 191–202 (2012).
34. Marce, R. *et al.* Interaction between wind-induced seiches and convective cooling governs algal distribution in a canyon-shaped reservoir. *Freshw. Biol.* **52**, 1336–1352 (2007).
35. Bohnenberger, J. E., Rodrigues, L. R., Da Motta-Marques, D. & Crossetti, L. O. Environmental dissimilarity over time in a large subtropical shallow lake is differently represented by phytoplankton functional approaches. *Mar. Freshw. Res.* **69**, 95–104 (2018).
36. Padišák, J., Crossetti, L. O. & Naselli-Flores, L. Use and misuse in the application of the phytoplankton functional classification: A critical review with updates. *Hydrobiologia* **621**, 1–19 (2009).
37. Shen, H. L. *et al.* Phytoplankton functional groups in a high spatial heterogeneity subtropical reservoir in China. *J. Gt. Lakes Res.* **40**, 859–869 (2014).
38. Jones, T. G. *et al.* Transformations in DOC along a source to sea continuum; impacts of photo-degradation, biological processes and mixing. *Aquat. Sci.* **78**, 433–446 (2016).
39. Eiler, A., Langenheder, S., Bertilsson, S. & Tranvik, L. J. Heterotrophic bacterial growth efficiency and community structure at different natural organic carbon concentrations. *Appl. Environ. Microbiol.* **69**, 3701–3709 (2003).
40. Tian, Y. Q., Wang, D., Chen, R. F. & Huang, W. Using modeled runoff to study DOC dynamics in stream and river flow: A case study of an urban watershed southeast of Boston Massachusetts. *Ecol. Eng.* **42**, 212–222 (2012).
41. Llíros, M. *et al.* Bacterial community composition in three freshwater reservoirs of different alkalinity and trophic status. *PLoS ONE* **9**, e116145 (2014).
42. Abell, J. M., Özkundakci, D. & Hamilton, D. P. Nitrogen and phosphorus limitation of phytoplankton growth in New Zealand lakes: Implications for eutrophication control. *Ecosystems* **13**, 966–977 (2010).
43. Wehr, J. D., Sheath, R. G. & Kociolek, J. P. *Freshwater Algae of North America: Ecology and Classification* (Academic Press, 2005).
44. Bovo-Scomparin, V. M. & Train, S. Long-term variability of the phytoplankton community in an isolated floodplain lake of the Ivinhema River State Park. *Brazil. Hydrobiologia* **610**, 331–344 (2008).
45. Wang, L. *et al.* Weekly dynamics of phytoplankton functional groups under high water level fluctuations in a subtropical reservoir bay. *Aquat. Ecol.* **45**, 197–212 (2011).
46. Jones, R. I. Mixotrophy in planktonic protists: An overview. *Freshw. Biol.* **45**, 219–226 (2000).
47. Reynolds, C. S. Variability in the provision and function of mucilage in phytoplankton: facultative responses to the environment. *Hydrobiologia* **578**, 37–45 (2007).
48. Salmaso, N., Boscaini, A. & Pindo, M. Unraveling the diversity of eukaryotic microplankton in a large and deep perialpine lake using a high throughput sequencing approach. *Front. Microbiol.* **11**, 789 (2020).
49. Association, A. P. H. *APHA Standard Methods for the Examination of Water and Wastewater* (American Public Association, 2005).
50. Bolan, N., Baskaran, S. & Thiagarajan, S. An evaluation of the methods of measurement of dissolved organic carbon in soils, manures, sludges, and stream water. *Commun. Soil Sci. Plant Anal.* **27**, 2723–2737 (1996).
51. Sun, Z. *et al.* Community dynamics of prokaryotic and eukaryotic microbes in an estuary reservoir. *Sci. Rep.* **4**, 6966 (2015).
52. Matejić, J. *et al.* Total phenolic content, flavonoid concentration, antioxidant and antimicrobial activity of methanol extracts from three *Seseli* L. taxa. *Open Life Sci.* **7**, 1116–1122 (2012).

Acknowledgements

This study was funded by the National Natural Science Foundation of China (No. 31770223), the Excellent Achievement Cultivation Project of Higher Education in Shanxi (No. 2020KJ029), Graduate Education Innovation Project in Shanxi (No. 2019SY030), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (No. 2019L0778) and the Shanxi “1331 Project”. We are grateful to the Dr. Tac A.L. (University of British Columbia) for the editorial assistance with the English.

Author contributions

J.Y., J.P.L., Q.L., F.R.N., and B.L. collected, transported and processed the water samples used in this study. J.Y., J.F., and S.L.X. analyzed and interpreted the data. J.Y. wrote the manuscript. All authors revised, read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-02183-5>.

Correspondence and requests for materials should be addressed to J.F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021