1-2018

Bioassessment of a Drinking Water Reservoir Using Plankton: High Throughput Sequencing vs. Traditional Morphological Method

Wanli Gao
Nanyang Normal University

Zhaojin Chen
Nanyang Normal University

Yuying Li
Nanyang Normal University

Yangdong Pan
Portland State University, pany@pdx.edu

Jingya Zhu
Nanyang Normal University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/esm_fac

Part of the Environmental Microbiology and Microbial Ecology Commons, Environmental Sciences Commons, and the Hydrology Commons

Let us know how access to this document benefits you.

Citation Details

Authors

Wanli Gao, Zhaojin Chen, Yuying Li, Yangdong Pan, Jingya Zhu, Shijun Guo, Lanqun Hu, and Jin Huang

Bioassessment of a Drinking Water Reservoir Using Plankton: High Throughput Sequencing vs. Traditional Morphological Method

Wanli Gao [1,2,†], Zhaojin Chen [1,2,†], Yuying Li [1,2,*], Yangdong Pan [2,3], Jingya Zhu [1,2], Shijun Guo [1,2], Lanqun Hu [2,4] and Jin Huang [2,4]

[1] Key Laboratory of Ecological Security for Water Source Region of Middle Route Project of South-North Water Diversion of Henan Province, College of Agricultural Engineering, Nanyang Normal University, Nanyang 473061, Henan, China; highly110@163.com (W.G.); zhaojin_chen@163.com (Z.C.); zhujy15@126.com (J.Z.); 152256558851@163.com (S.G.)

[2] Collaborative Innovation Center of Water Security for Water Source Region of Middle Route Project of South-North Water Diversion of Henan Province, Nanyang Normal University, Nanyang 473061, Henan, China; pany@pdx.edu (Y.P.); hlq65@sina.com (L.H.); nyhbhj@163.com (J.H.)

[3] Department of Environmental Science and Management, Portland State University, Portland, OR 97207, USA

[4] Emergency Centre for Environmental Monitoring of the Canal Head of Middle Route Project of South-North Water Division, Xichuan 474475, Henan, China

* Correspondence: lyying200508@163.com; Tel.: +86-377-6352-5027

† These authors contributed equally to this work.

Abstract: Drinking water safety is increasingly perceived as one of the top global environmental issues. Plankton has been commonly used as a bioindicator for water quality in lakes and reservoirs. Recently, DNA sequencing technology has been applied to bioassessment. In this study, we compared the effectiveness of the 16S and 18S rRNA high throughput sequencing method (HTS) and the traditional optical microscopy method (TOM) in the bioassessment of drinking water quality. Five stations reflecting different habitats and hydrological conditions in Danjiangkou Reservoir, one of the largest drinking water reservoirs in Asia, were sampled May 2016. Non-metric multi-dimensional scaling (NMDS) analysis showed that plankton assemblages varied among the stations and the spatial patterns revealed by the two methods were consistent. The correlation between TOM and HTS in a symmetric Procrustes analysis was 0.61, revealing overall good concordance between the two methods. Procrustes analysis also showed that site-specific differences between the two methods varied among the stations. Station Heijizui (H), a site heavily influenced by two tributaries, had the largest difference while station Qushou (Q), a confluence site close to the outlet dam, had the smallest difference between the two methods. Our results show that DNA sequencing has the potential to provide consistent identification of taxa, and reliable bioassessment in a long-term biomonitoring and assessment program for drinking water reservoirs.

Keywords: Danjiangkou Reservoir; plankton; high throughput sequencing; generalized procrustes analysis; bioassessment

1. Introduction

Clean freshwater resources are becoming increasingly scarce globally [1–4]. Due to climate change, economic development and population growth, approximately four billion persons of the world's population are facing severe water scarcity, with nearly half of them living in India and China [1,5]. Drinking water security may substantially hamper the sustainable development of humanity, particularly in developing countries. To better protect and manage diminishing freshwater

resources, researchers have developed and tested multiple biological indices for assessing water quality and the ecological integrity of aquatic ecosystems since biological assemblages provide a direct measure of the aquatic ecosystems' conditions [6]. Multimetric indices of fish, macroinvertebrates and periphyton have been effectively used to assess surface water quality in the USA [7]. For example, in an integrated biosurvey as a tool for evaluation of aquatic life use attainment and impairment in Ohio surface waters, Yoder (1991) [8] reported that biological monitoring and assessment indicated that approximately 50% of 645 Ohio stream/river segments were impaired while chemical monitoring and assessment showed no signs of impairment. Biological assessment has become an integral part of water resource protection and management.

One of the key components for a successful bioassessment program is to accurately characterize the composition of biological assemblages. Traditional bioassessment methods rely on specialized analysts to identify taxonomic groups, a time-consuming process. The quality of taxonomic analysis largely depends on analysts training, experience, and interpretation of the taxonomic literature. For instance, several studies on inter-analysts comparison of the diatom identification showed that inconsistency among independent analysts can contribute uncertainty to the bioassessment [9,10]. The inconsistency among the analysts may be more problematic for long-term monitoring programs, particularly with climate change. With the development of DNA sequencing technology [11–14], DNA sequencing has been applied to bioassessment, particularly in freshwater [15–21]. High throughput sequencing (HTS) was developed to characterize biological assemblages in environmental samples. The method is faster in terms of sampling processing, and may become cheaper as the technology improves, and more importantly with the ability to provide more reliable and richer biological information than the traditional morphology-based method [22–24]. Different plankton assemblages as bioindicators were characterized using DNA sequencing including bacterioplankton [18,24–26], and phytoplankton [27–33]. Baird and Hajibabaei (2012) [17] envisioned a new paradigm (i.e., Biomonitoring 2.0) in ecosystem assessment based on a HTS platform, though a complete paradigm shift may require more research.

Several researchers have assessed the effectiveness of water quality assessment using the traditional optic microscopic method and the DNA sequencing method [21,23,34,35]. Due to several limitations, such as incomplete taxonomic coverage in DNA reference libraries and biases related to molecular procedures, the two methods may not generate identical compositional data, particularly at the species level. Despite the difference, biotic indices based on the data generated by the two methods were highly consistent. Multiple researchers suggested that DNA sequencing generates a richer amount of information on biotic diversity with consistent and increased taxonomic resolution and thus has a great potential to improve the effectiveness of current bioassessment. Not surprisingly, most of these studies focused on benthic diatoms in streams and rivers, since diatoms are commonly used as bioindicators in lentic ecosystems. The studies that systematically compare the plankton assemblages characterized by traditional optical microscopy (TOM) and HTS sequencing methods in drinking water reservoirs are still limited.

China is facing challenges in both water quality and quantity. With uneven water distribution, the North China Plain is highly water-stressed, while water resources are relatively more abundant in the south [5,36,37]. Based on the strategic demands for China's regional sustainable development, the Chinese government has launched the South–North Water Diversion (SNWD) Project to transfer water from the Yangtze River and its tributaries to the more arid and industrialized North China Plain [36–38]. As the largest drinking water source in China [38], affecting more than 53 million people in Beijing and other receiving-water regions, the water quality in the Danjiangkou Reservoir, source water for the middle route of the SNWD, is required to be in good quality, be stable long-term, be continuously improved, and be able to adapt to further changes such as climate, and hence the establishment of a continuous ecological monitoring database on the reservoir is particularly important. Such a database can provide scientific data support for best and adaptive management practices.

In this study, we compared the effectiveness of the TOM and HTS methods in assessing water quality in the Danjiangkou Reservoir.

2. Materials and Methods

2.1. Study Area

As one of the largest river impoundments in the Yangtze River basin, the Danjiangkou Reservoir (32°36′~33°48′ N; 110°59′~111°49′ E), with a maximum depth of about 80 m, is located at the juncture of Hubei and Henan provinces of central China. Its drainage area includes the upper Hanjiang and Danjiang rivers, with a total area of 95,000 km² (Figure 1). The Danjiangkou Reservoir has a variety of functions, such as flood control, electricity generation, irrigation, shipping and drinking water supply [36–38].
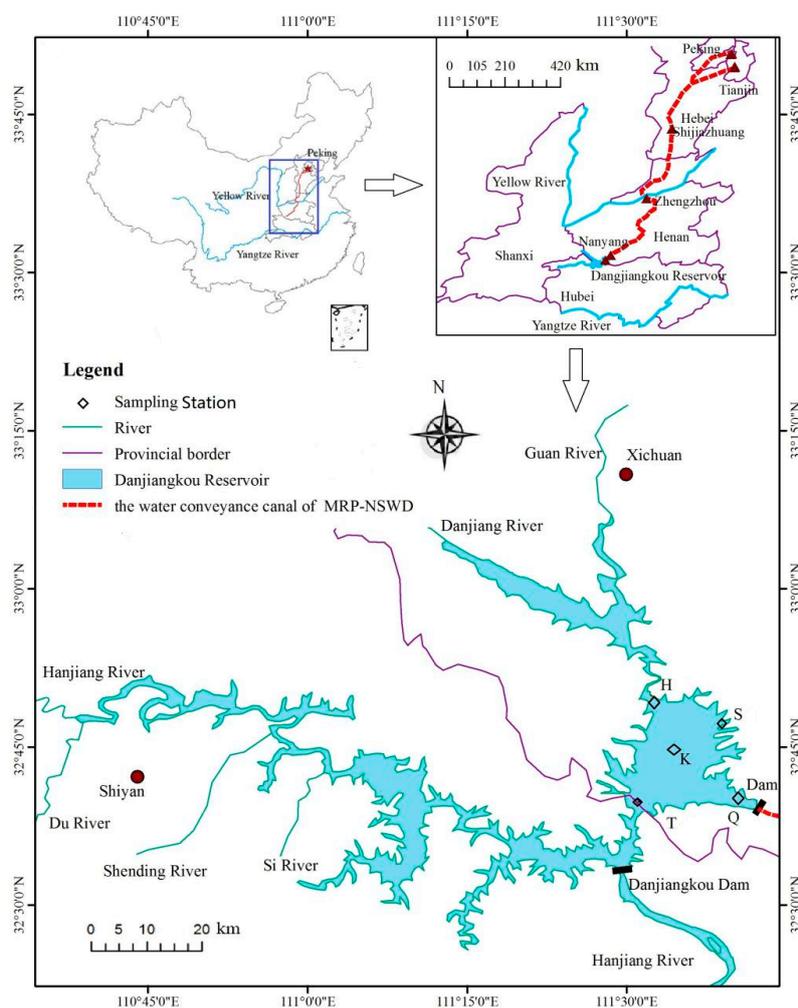
Figure 1. Locations of the five sampling stations in the Danjiangkou Reservoir and the water conveyance canal of the Middle Route Project of South-North Water Division in China. Station codes represent the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishan.

The reservoir is located in the north sub-tropic monsoon climate region with an annual mean temperature of 15~16 °C and annual precipitation 800~1000 mm, of which 80% is concentrated in the period from May to October. The monthly maximum precipitation is 193.7 mm and the minimum is 31.0 mm. Soil types include alluvial soil, lime concretion black soil, yellow brown loam and purple soil. The watershed around the reservoir is dominated by mountainous areas (85%) and forest coverage (76%) [39].

2.2. Field Sampling

Based on the previous studies [40,41], we selected the five most representative sampling stations in the reservoir. In each station, water samples were collected at the 0~50 cm below the water surface for HTS analysis of planktonic bacterial and eukaryotic assemblages and TOM analysis of the phytoplankton and physico-chemical variables in May 2016 (rainy season). Qushou (Q) station is 100 m upstream of the water outlet dam of the water conveyance canal of the Middle Route Project of SNWD. Heijizuo (H) station is close to the confluence of the Danjiang and Guan rivers, two major tributaries to the reservoir in Henan province. Songgang (S) station is located in a reservoir bay which was influenced by a shipping dock and tourists. Kuxin (K) station is located in the middle of the reservoir while Taizishan (T) station is in the confluence of the two sub-basins of the reservoir (Figure 1). There were three replicates for each sample.

2.3. Morphological Identification of Phytoplankton Assemblages

Phytoplankton samples were preserved immediately with 1% Lugol's solution. The water samples for phytoplankton analysis were stored in a 2 L glass sedimentation utensil. After 48 h sedimentation, each sample was condensed to about 30 mL, and then stored in darkness at 4 °C until analysis. Phytoplankton were identified under a microscope (Nike E200) according to Hu and Wei (2006) [42]. Phytoplankton analysis followed the standard method [43]. Before counting, each sample was gently mixed. Phytoplankton cells were counted in a Fuchs-Rosental counting chamber at $400\times$ in Z line microscope fields. Algal biovolumes were calculated from measured morphometric characteristics (diameter, length and width). Conversion to biomass was based on 1 mm^3 of algal volume being equivalent to 1 mg of fresh weight biomass. A minimum of 500 individual units were counted with a counting error of less than 10%. Three subsamples were analyzed for each sample.

The Shannon–Wiener index [44] (H') of phytoplankton was used to evaluate water quality. The water quality was classified as severe ($H' = 0\sim1$), moderately β ($H' = 1\sim2$), moderately α ($H' = 2\sim3$) trophic condition and clean condition ($H' > 3$) [45].

2.4. DNA Extraction and Sequencing

Samples were first filtered through a 0.45 μm diameter filter for planktonic eukaryote sequencing and the filtrates were then collected through a 0.22 μm diameter filter for bacterioplankton sequencing. Plankton genomic DNA in the stored filters was extracted using the E.Z.N.A.® Water DNA Kit (OMEGA bio-tek, Norcross, GA, USA), following the manufacturer's instructions. Electrophoresis and Nano Drop ND 2000 (Thermo Scientific, Pittsburgh, PA, USA) were used to estimate the quantity of extracted DNA. The V3-V4 region of bacterial 16S rRNA gene was amplified with 338F and 806R [18,46] and the V5-V7 region of eukaryote 18S rRNA gene was amplified with 0817F and 1196R with sample-identifying barcodes [47]. High throughput sequencing (Illumina MiSeq PE300 platform, Illumina, Inc., CA, USA) was performed by Shanghai Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China).

In both cases, the MiSeq sequencing data were processed using the QIIME Pipeline [11]. The operational taxonomic units (OTUs) were determined (at 97% similarity level) using USEARCH. The OTU number of each sample was used to represent taxa richness. Rarefaction curves and a Shannon–Wiener index (H') were generated, and the ACE, Shannon, Simpson, and Chao1 estimators were calculated to compare the richness and diversity of plankton. Taxonomic classification at the phylum and genus levels was performed using the ribosome database project (RDP) algorithm.

2.5. Physico-Chemical Variables

Physico-chemical variables were measured using standard methods [43]. Water temperature (T), pH, electric conductivity (EC), and dissolved oxygen (DO) were measured in situ using a YSI 6920 (YSI Inc., Yellow Springs, OH, USA). Secchi depth (SD) was determined with a 30 cm diameter

Secchi disk. Water samples for chemical analysis were transported to the laboratory within 24 h, stored in a refrigerator at 4 °C, and analyzed within one week after sample collection.

Permanganate index (COD_{Mn}) was calculated using the potassium permanganate index method and chemical oxygen demand (COD) was measured by the potassium dichromate method. Total phosphorus (TP) was determined with acidified molybdate to form reduced phosphor-molybdenum blue and measured spectro-photometrically. Total nitrogen (TN) was assayed with alkaline persulfate digestion and UV spectro-photometry, and ammonia nitrogen (NH_4^+-N) was measured with Nessler's reagent spectro-photometric method. Chlorophyll a (Chla) concentration was estimated spectro-photometrically after extraction in 90% ethanol [43].

The evaluation for the water trophic state was based on the five variables of TN, TP, COD_{Mn}, SD and Chla [48]. Using the linear interpolation method, we converted the value of each environmental variable into the assigned value (E), and then used the arithmetic mean of each assigned value as the trophic state index (EI). EI values between 60~100 indicate hyper-eutrophic conditions, between 50~60 mildly eutrophic conditions and between 20~50 moderately eutrophic conditions [48].

2.6. Data Analysis

Datasets between HTS and TOM were compared at the phylum/genus level. All genera belonging to the concept of phytoplankton, including Cyanobacteria (from the 16S sequencing dataset), Bacillariophyta, Chlorophyta, and other groups of eukaryotic phytoplankton (from the 18S sequencing dataset) were determined from the cleaned HTS datasets. The numbers of OTUs were compared with their corresponding number of phytoplankton species (or genera) found by TOM.

The per-mutational multivariate analysis of variance (PERMANOVA) and analysis of similarity (ANOSIM) were performed using R software with the "vegan" package to assess the significant differences in assemblage structure among the sampling stations. Non-metric multi-dimensional scaling (NMDS) was also conducted with the R function "metaMDS" in the same package, and the first two NMDS axis scores (NMDS I and NMDS II) for each station were used as reduced multi-dimensional data for the plankton assemblages. The relative abundance of biomass data was used for the NMDS analysis. Generalized procrustes analysis (GPA) was used to compare the ordination based on phytoplankton assemblages and DNA barcoding [49]. To assess the differences in the environmental variables among stations, one-way analysis of variance (ANOVA) was applied with statistical significance set prior at $p < 0.05$. All data analyses were performed in R (Ver. 3.4.0, R Development Core Team, 2017).

3. Results

3.1. Evaluation of Water Quality and Trophic Status

The physico-chemical variables indicated that the water quality of the reservoir was overall good as drinking water according to the environmental quality standards of surface water of China (GB38382-2002). However, TN was higher than the Class III surface water standard. TN at H station was 54.08% higher than that of other stations. Both COD_{Mn} and TP met Class II surface water standards while other indicators met Class I surface water standards (Table 1). There were significant differences in the environmental variables among the stations except TP ($p < 0.05$). According to TN, the five stations could be divided into three different groups (i.e., H station, K and T stations, Q and S stations, Table 1).

The mean E value of TN was 60.05. The E values in the H and T stations were higher, over 60.00. The E values of the other four environmental variables were between 20.00~49.05. The mean EI values, ranged between 38.78~41.16, were significantly different among the five stations ($p < 0.05$).

Table 1. Main physico-chemical characteristics of water samples. Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang.

Station	pH	SD (m)	DO (mg/L)	TN (mg/L)	NH_4^+-N (mg/L)	TP (mg/L)	COD (mg/L)	COD_{Mn} (mg/L)	Chla (mg/m^3)	EI
Q	8.70 ± 0.009e	3.00 ± 0.000b	7.50 ± 0.035a	0.91 ± 0.003a	0.041 ± 0.005bc	0.02 ± 0a	13.97 ± 0.006d	3.81 ± 0.006e	2.42 ± 0.035bc	41.16 ± 0.003b
K	8.50 ± 0.006d	5.00 ± 0.003e	9.17 ± 0.015d	0.96 ± 0.007b	0.069 ± 0.003d	0.02 ± 0a	11.60 ± 0.145a	3.46 ± 0.006a	2.13 ± 0.015a	38.78 ± 0.002a
S	8.09 ± 0.003a	2.90 ± 0.003a	8.42 ± 0.029c	0.92 ± 0.023ab	0.048 ± 0.001c	0.02 ± 0a	14.37 ± 0.088e	3.74 ± 0.006d	2.40 ± 0.023b	41.13 ± 0.001b
H	8.35 ± 0.003b	4.20 ± 0.003c	8.34 ± 0.020b	1.46 ± 0.020d	0.031 ± 0.001ab	0.02 ± 0a	11.97 ± 0.088b	3.52 ± 0.003b	2.49 ± 0.019cd	41.04 ± 0.003b
T	8.44 ± 0.003c	4.30 ± 0.000d	9.20 ± 0.012d	1.03 ± 0.006c	0.027 ± 0.006a	0.02 ± 0a	12.33 ± 0.033c	3.56 ± 0.015c	2.53 ± 0.003d	40.25 ± 0.000b

Notes: (1) Mean ± standard error; (2) Different lowercase letters in the same column showed that the indicator was significant among stations at $p < 0.05$ level; (3) According to China SL395-2007 evaluation of the water trophic state [48], we converted the value of each environmental variable into the assigned value (E) using the linear interpolation method, and then used the arithmetic mean of each assigned value as the trophic state index (EI).

3.2. Phytoplankton Assemblages Characterized by TOM

A total of 39 taxa were recorded, belonging to five divisions, 17 families and 26 genera. Bacillariophyta, Chlorophyta and Cyanophyta accounted for 51.28%, 23.08% and 15.38% of the total genus/species and 72.53%, 22.15% and 3.78% of the total biomass, respectively (Figure 2). *H'* values varied from the highest of 5.66 at Q station to the lowest of 0.67 at T station (Table A1), suggesting that the reservoir was under a moderate trophic condition.
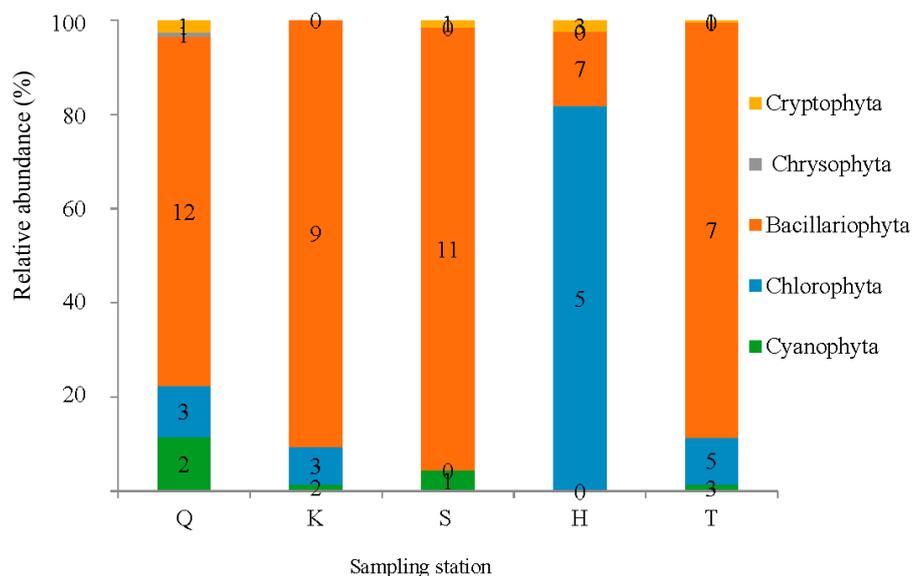
Figure 2. Relative abundance (%) of biomass and taxa number of phytoplankton assemblages by traditional optical microscopy method (TOM). Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang.

The NMDS analysis showed that phytoplankton assemblages substantially varied among the five stations (Figure 3a). The NMDS axis I mainly reflected the change of Chlorophyta and Bacillariophyta in the riverine and reservoir areas. The biomass of Chlorophyta, dominated by *Pandorina* sp. (56.34% of the total biomass) at H station, was much higher than that of the other four stations. In contrast, the biomass of Bacillariophyta at H station was the lowest. The NMDS axis II was negatively correlated with the total phytoplankton biomass. *Aulacoseira granulata* (Ehr.) Simonsen accounted for 53.70% and 36.81% of the total biomass in K and T stations, respectively, while *Navicula* sp. accounted for 35.25% and 30.84% in Q and S stations, respectively.
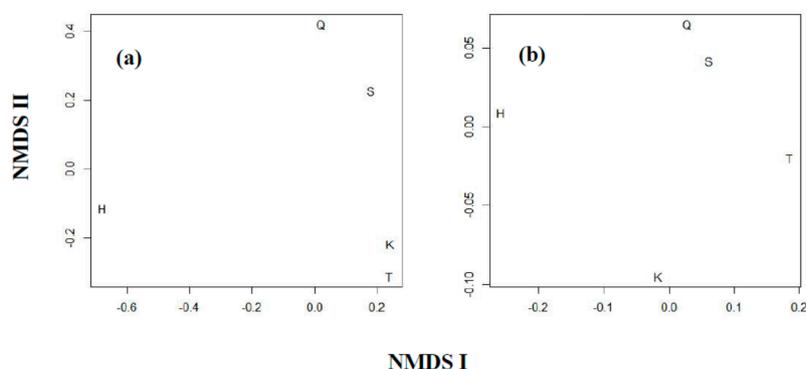
Figure 3. Comparison of the NMDS plots based on phytoplankton (**a**) and DNA barcoding (**b**). Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang.

3.3. Plankton Assemblages Characterized by HTS

The HTS results of 16S rRNA and 18S rRNA gene showed that the library coverage and Simpson index of 18S rRNA planktonic eukaryotic assemblages (18S assemblages) were higher than those of 16S rRNA bacterioplankton assemblages (16S assemblages) (Tables 2 and A1, and Figure 4). In contrast, the ACE index, Chao1 index and *H*′ values of the 18S assemblages were lower than those of the 16S assemblages. The OTUs number, ACE index, Chao1 index, *H*′ and Simpson index of the 18S assemblages were significantly different among the stations ($p < 0.05$).

The HTS results showed that the mean bands and OTUs of 16S assemblages of 15 samples were 29,468 and 343 (Table 2 and Figure 4a). Out of 25 phyla of bacteria, Proteobacteria, Bacteroidetes, Actinobacteria, Cyanobacteria and Firmicutes were dominant. The first dominant Proteobacteria was 36.71%, 32.70%, 34.87%, 29.57% and 34.24% of the total OTUs of Q, K, S, H and T stations, and the fourth dominant Cyanobacteria accounted for 7.46%, 10.97%, 11.47%, 9.86% and 8.68%, respectively.
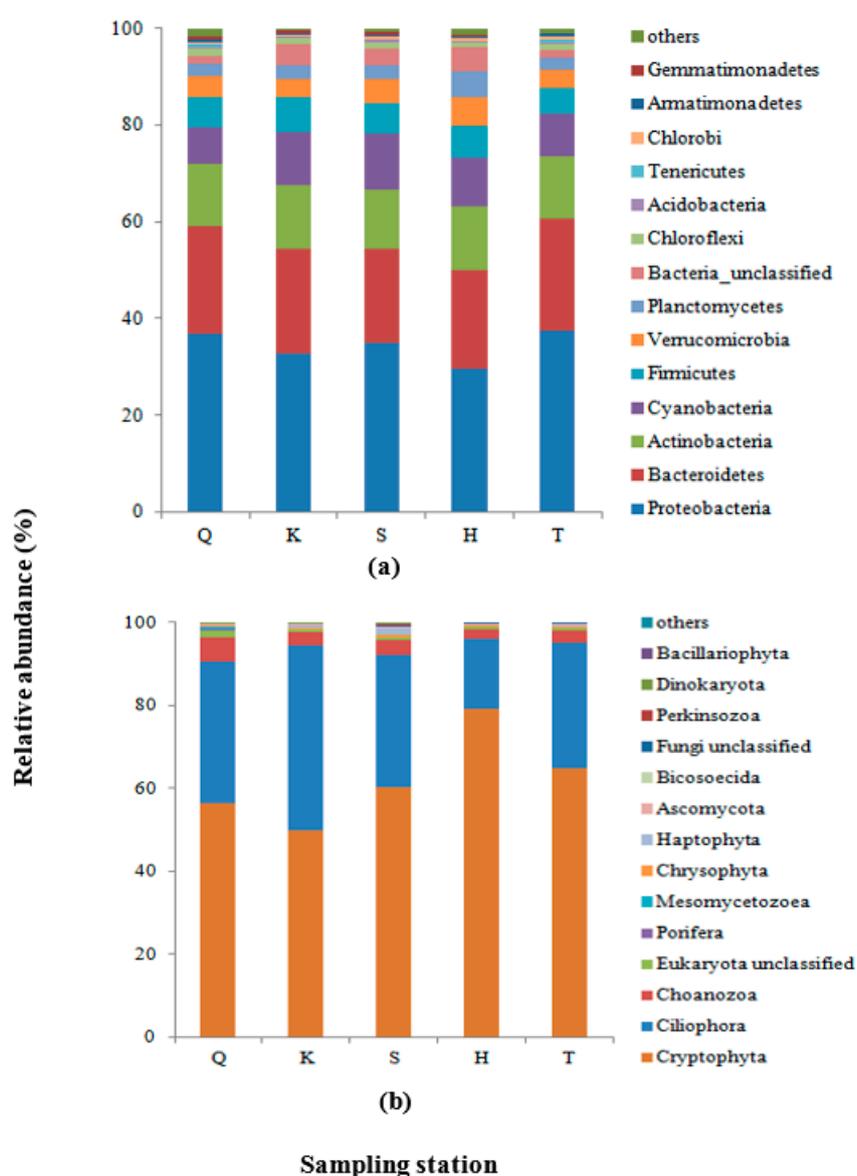
Figure 4. Relative abundance (%) of 16S (**a**) and 18S (**b**) assemblage structures at phylum levels in different stations. Only taxa with 0.05% abundance or more in the total dataset were identified in the legend. Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang.

Table 2. Estimation of molecular diversity of 16S and 18S assemblages. Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang.

Station	Sequencing Band No.	Clustering at 97% Similarity Threshold					
		OTU No.	ACE Index	Chao 1 Index	Library Coverage (%)	Shannon-Wiener Index	Simpson Index
				16S assemblages			
Q	32465 ± 1731a	345 ± 32a	421 ± 29a	421 ± 28a	99.75 ± 0.01a	3.85 ± 0.139a	0.0395 ± 0.007a
K	31068 ± 4516a	334 ± 17a	396 ± 10a	412 ± 10a	99.76 ± 0.05a	4.07 ± 0.070a	0.0293 ± 0.002a
S	27449 ± 3454a	353 ± 37a	420 ± 34a	417 ± 31a	99.70 ± 0.03a	3.97 ± 0.111a	0.0373 ± 0.003a
H	28242 ± 1686a	351 ± 15a	418 ± 12a	414 ± 14a	99.74 ± 0.03a	4.08 ± 0.133a	0.0321 ± 0.005a
T	28116 ± 3473a	332 ± 34a	387 ± 29a	389 ± 24a	99.76 ± 0.03a	4.02 ± 0.135a	0.0332 ± 0.004a
				18S assemblages			
Q	34240 ± 3685a	153 ± 5.49c	168 ± 7.81b	164 ± 5.67b	99.93 ± 0.021a	2.57 ± 0.05c	0.1362 ± 0.005a
K	36042 ± 2763a	104 ± 4.37a	119 ± 5.69a	124 ± 1.15a	99.95 ± 0.004a	2.20 ± 0.02b	0.1785 ± 0.004b
S	36651 ± 1579a	125 ± 4.35b	137 ± 4.70a	137 ± 3.48a	99.95 ± 0.003a	2.47 ± 0.05c	0.1300 ± 0.004a
H	36931 ± 4520a	108 ± 7.31ab	133 ± 13.17a	127 ± 7.21a	99.94 ± 0.014a	2.03 ± 0.01a	0.2355 ± 0.002c
T	33678 ± 3723a	115 ± 3.79ab	128 ± 4.33a	126 ± 2.85a	99.94 ± 0.007a	2.15 ± 0.07ab	0.2041 ± 0.017b

Note: (1) Mean ± standard error; (2) Different lowercase letters in the same column indicated that the indicator was significant among stations in $p < 0.05$.

The HTS result of 18S assemblages showed that the mean bands and OTUs were 35,508 and 121 (Table 2). The OTUs of Bikonta, Unikonta and the unclassified accounted for 59.57%, 38.30% and 2.13% of the total OTUs, respectively. There were 20 phyla of planktonic eukaryotes, mainly including Cryptophyta, Ciliophora, Choanomonada, Haptophyta and other groups (Figure 4b). Cryptophyta, Ciliophora and Choanozoa accounted for 97.21% of the total OTUs. Phytoplankton and zooplankton accounted for 63.25% and 35.48% of the total OTUs, respectively. The first dominant Cryptophyta accounted for 56.15%, 49.85%, 60.20%, 79.11% and 64.65% of the total OTUs in Q, K, S, H and T stations, and the second dominant Ciliophora was 34.17%, 44.37%, 32.00%, 16.77% and 30.45%, respectively. Picked phytoplankton included Cryptophyta, Chrysophyta, Haptophyta, Perkinsozoa, Dinokaryota, Bolidomonas, Bacillariophyta, Chlorophyta, Charophyta and Euglenophyta, the first assemblages of which were of 62.00% of the total OTUs.

The NMDS analysis of the DNA sequencing data showed that 16S and 18S assemblages varied among the stations (Figures 3b, 5, 6 and A1). The spatial variation of plankton assemblages among the stations was greater than the measurement error among the replicates of each station (Figures 3b and 5). The difference in bacterioplankton assemblages among the stations was lower than that of eukaryote assemblages (Figures 4 and A1). The bacterioplankton assemblages along the NMDS axis I from left to right may reflect different habitat conditions, and along the NMDS axis II, the differences among bacterioplankton assemblages were less evident (Figure A1a). The OTU percentages of Cyanobacteria_norank and SubsectionI_Family I_norank in H station were higher than those of the other four stations while the OTUs of Synechococcus were much lower. The relative abundance of dominant Cryptomonadales_uncultured along the NMDS axis I from left to right decreased from 43.86% to 14.56%, and along the NMDS axis II from bottom to top, the OTUs of Choanozoa_incertae_sedis gradually dropped from 3.76% to 1.53% (Figure A1b).
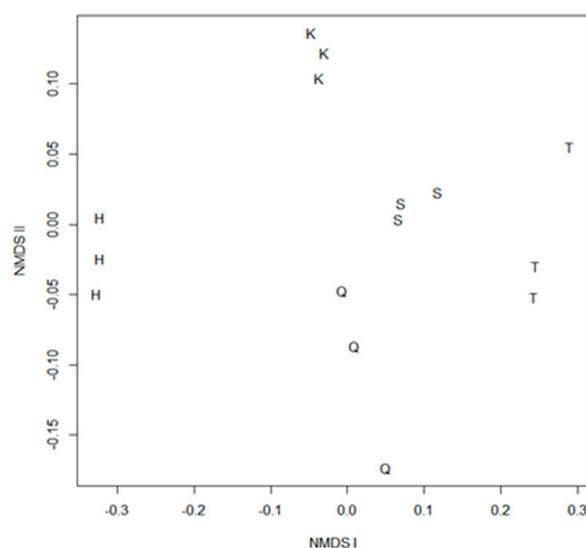
Figure 5. NMDS plot showing both spatial variation among five stations and measurement errors for each station based on the relative abundance of 16S and 18S sequencing barcoding at the genus level. Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang. Three identical letters represent three repetitions of each station. There were three repetitions for each sample.
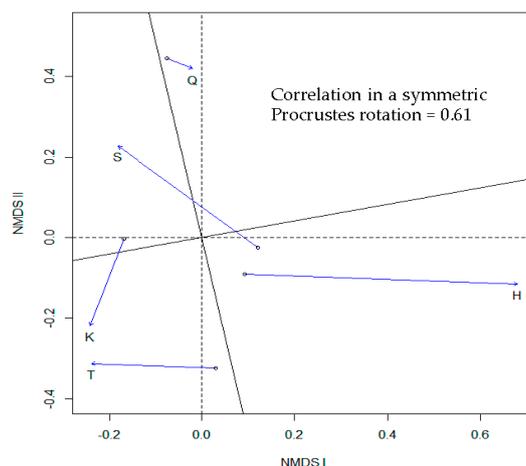
Figure 6. Procrustes plot comparing phytoplankton assemblages and DNA barcoding. Solid lines represented Procrustes residuals from both methods. Station codes represented the first letter of their names: Q: Qushou, K: Kuxin, S: Songgan, H: Heijizuo, T: Taizishang.

3.4. Comparison of Phytoplankton Assemblages by TOM and HTS

The plankton assemblages identified by the two methods were substantially different (Tables 2 and A1; Figures 2–6 and Figure A2). At the phylum level, the number of the phyla identified by HTS was more than that by TOM. However, the genera number with almost the same taxa was similar (Figures 2 and 4). Cyanophyta, Chlorophyta, Bacillariophyta, Chrysophyta and Cryptophyta were detected by both methods with less of the same genera, followed by Microcystis, Anabaena, Navicula and Cryptomonas. Bacillariophyta was dominant with a mean of 51.28% of the total genera/species by TOM while Cryptophyta was dominant with mean 98.27% of the total OTUs by HTS. Dinokaryota, Charophyta, Euglenophyta and the other rare groups were only identified by HTS.

The NMDS analysis showed the substantial spatial difference of the plankton assemblages among the stations using either method (Figures 3, 5, 6 and A1). The plankton assemblages of H station were completely different from those of the other four stations (Figures 3, 5 and 6). The stations could be divided into three group types such as H station, Q and S stations, K and T stations (Figure 3).

Based on the two datasets, GPA was used to compare the difference in assemblages by the two methods. The correlation between TOM and HTS in a symmetric Procrustes rotation was 0.61 (Figure 6), revealing good concordance between the two datasets. The residuals varied among the sampling stations. Q station, located just above the outlet dam, had the smallest residual while H station, located in the confluence of two tributaries with strong river impact, had the largest residual (Figure 6).

4. Discussion

Our study showed that the plankton spatial pattern in the reservoir was consistent using both the TOM and HTS datasets (Figure 3). The reservoir in general has good water quality and thus the spatial differences among the sampled stations may largely reflect the variation of habitat and hydrological conditions. For instance, station H, a strong riverine site located at the confluence of two major tributaries, was consistently different from the rest of the stations in the NMDS plots. Our results are consistent with findings by other researchers [23,50,51]. For instance, both DNA sequencing and TOM analyses generally captured frequency shifts of abundant taxa over the seasonal samples [50]. Vasselon et al. (2017) [23] compared HTS and morphological water quality index values in 33 river sites in Island Mayotte and found that the water quality status was congruent between the two methods. Eiler et al. (2013) [29] reported that DNA-sequencing-derived phytoplankton composition differed

significantly among lakes with different trophic status, showing that DNA sequencing could resolve phytoplankton assemblages at a level relevant for ecosystem management.

Procrustes analysis showed that consistency between the two methods varied among the stations in the reservoir (Figure 6). The two methods generated the best agreement at station Q. The station, located just above the outlet dam, may best integrate spatial variability in the reservoir. Both TOM and 18S sRNA results indicated that the sample from this station had the highest taxa richness. In contrast, the least agreement between the two methods was at station H. Located in the downstream of the two tributaries, the station may be heavily influenced by the rivers and consequently its plankton composition with more benthic algae and less true planktonic algae was substantially different from the rest of the stations. The site-specific discrepancy between the two methods may also be due to the incomplete taxonomic coverage in the reference library, particularly for benthic diatoms. For instance, Vasselon et al. (2017) [23] reported that only 13% of the benthic diatom species was shared by the two methods in 33 river sites in Island Mayotte.

It was expected that the plankton assemblages identified by TOM and HTS would be substantially different in terms of their composition. Compared to TOM, HTS can detect a wide variety of plankton, including nanoplankton and rare taxa. Other studies also found similar results in freshwater ecosystems [32,34,50,52]. Xiao et al. (2014) [32] found that the species compositions detected by TOM and 454 HTS did not always match at the taxa level after analyzing 300 weekly water samples over 20 years in Lake Gjersjøen, Norway, a drinking water source. Studies in Lake Tegel, Germany, showed that because the 480 used diatom sequences of the 18S region were generated from world-wide occurrences, only a small number of individuals precisely matched on the species level [52]. A deep-branching taxonomically unclassified cluster was frequently detected by DNA sequencing, but could not be linked to any group identified by microscopy. DNA sequencing allows approximately three orders of magnitude larger SSU rDNA sequencing [16]. Deleting rare species can affect the sensitivity of biotic indices to detect environmental degradation [53]. In the absence of other nuclear markers less susceptible to copy number variation, rDNA-based diversity studies need to be adjusted for confounding effects of copy number variation [50]. Evans et al. (2007) [54] assessed the effectiveness of several genes (cox1, rbcL, 18S and ITS rDNA) to distinguish cryptic species within the model "morphospecies", *Sellaphora pupula agg.*, and found that tree topologies were very similar although support values were generally lower for cox1. To assess the proportional biomass of diatoms and dinoflagellates along the Swedish west marine coast, Godhe et al. (2008) [27] designed two real-time PCR assays with special primers to find that the linear regression of the proportion of SSU rDNA copies of dinoflagellate and diatom origin versus the proportion of dinoflagellate and diatom biovolumes or biomass per liter was significant. Thus, for diatoms, linear regression of the number of small subunit ribosomal DNA (SSU rDNA) copies versus biovolume or biomass per liter was significant, but no such significant correlation was detected in the field samples for dinoflagellates. Exploring the alternative markers (e.g., ~1400 bp of rbcL; 748 bp at the 3_end of rbcL (rbcL-3P); large ribosomal subunit (LSU D2/D3) and UPA) to diatom barcoding (e.g., COI-5P, rbcL-3P) should be used as the primary marker for diatom barcoding, while LSU D2/D3 should be sequenced as a secondary marker to facilitate bioassessment [55].

The high reproducibility and potential for standardization and parallelization makes the high throughput sequencing approach an excellent candidate for the simultaneous monitoring of plankton assemblages in drinking water quality, mainly including both phytoplankton and bacterioplankton. For a long-term biomonitoring and assessment program in the Danjiangkou Reservoir, a critical drinking water source for northern China, our study shows that the DNA barcoding method has great potential. DNA sequencing can provide a rapid and consistent identification of taxa. The next step is to design a specific assay with a specific DNA extraction method and a specific primer for the local flora in the reservoir.

5. Conclusions

In summary, the plankton spatial pattern among the stations in the Danjiangkou Reservoir was consistently detected by both TOM and HTS methods, reflecting the variation of habitat and hydrological conditions in the reservoir. To develop a consistent and accurate long-term ecological monitoring database for drinking water quality, DNA sequencing may serve as a promising alternative.

Author Contributions: All the authors contributed equally to this work. Specifically, Wanli Gao and Zhaojin Chen performed the experiments and contributed to writing the paper; Yuying Li and Yangdong Pan conceived and designed the experiments; Jingya Zhu and Shijun Guo did field sampling and analyzed the data. Jin Huang measured the water quality. Lanqun Hu provided the laboratory space for measuring water quality.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Shannon-Wiener index (*H'*) of phytoplankton, 16S and 18S assemblages. *H'* values were the mean of the three repetitions of sequencing data. Station codes represented the first letter of their names: H: Heijizuo, K: Kuxin, Q: Qushou, S: Songgan, T: Taizishang.

Station	Phytoplankton Assemblages	16S Assemblages	18S Assemblages
Q	5.66	3.85	2.57
K	1.55	4.07	2.20
S	1.93	3.97	2.47
H	3.85	4.08	2.03
T	0.67	4.02	2.15
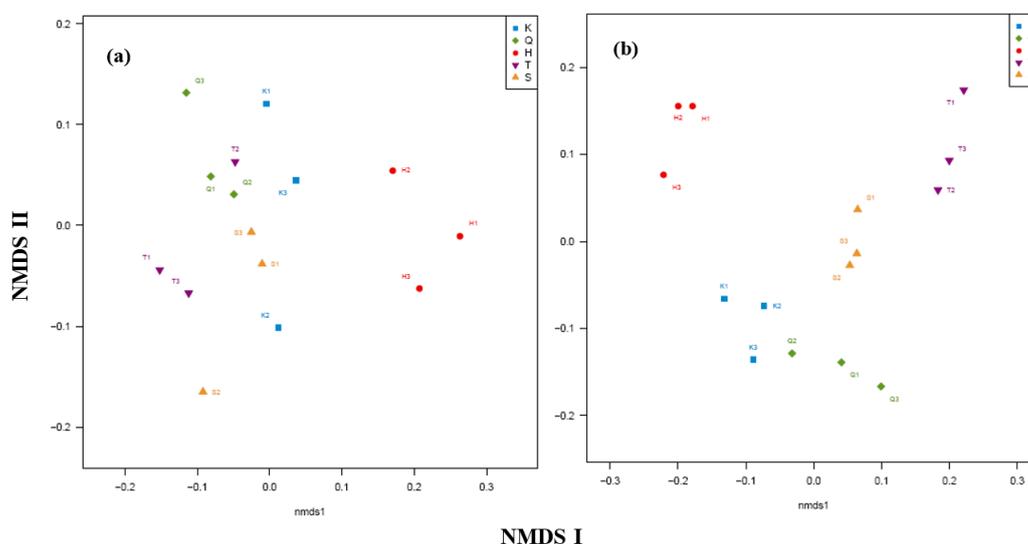
Figure A1. NMDS analysis of 16S (**a**) and 18S (**b**) assemblages structure. Station codes represented the first letter of their names: H: Heijizuo, K: Kuxin, Q: Qushou, S: Songgan, T: Taizishang. The numbers after the station codes indicated the repeating serial numbers.
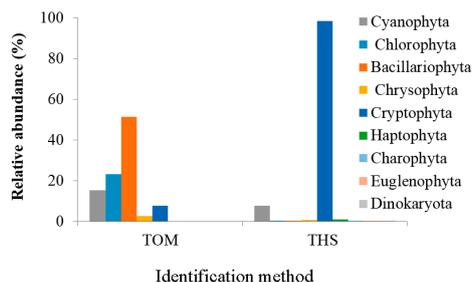
Figure A2. Relative abundance (%) of taxa and OTUs of phytoplankton using traditional optical microscopy method (TOM) and high throughput sequencing method (HTS). OTUs percentage was only that the number of OTUs of one assemblages accounted for the total OTUs number of phytoplankton identified from the 16S and 18S sequencing datasets.

References

1. Vörösmarty, C.J.; Green, P.; Salisbury, J.; Lammers, R.B. Global water resources: Vulnerability from climate change and population growth. *Science* **2000**, *289*, 284–288. [CrossRef] [PubMed]

2. Simonovic, S.P. World water dynamics: Global modeling of water resources. *J. Environ. Manag.* **2002**, *66*, 249–267. [CrossRef]

3. Oki, T.; Kanae, S. Global hydrological cycles and world water resources. *Science* **2006**, *313*, 1068–1072. [CrossRef] [PubMed]

4. Guitton, M.J. The water challenges: Alternative paths to trigger large-scale behavioural shifts. *Lancet Planet. Health* **2017**, *1*, 46–47. [CrossRef]

5. Mekonnen, M.M.; Hoekstra, A.Y. Four billion people facing severe water scarcity. *Sci. Adv.* **2016**, *2*, 1–6. [CrossRef] [PubMed]

6. Karr, J.R. Biological integrity: A long-neglected aspect of water resource management. *Ecol. Appl.* **1991**, *1*, 66–84. [CrossRef] [PubMed]

7. Barbour, M.T.; Gerritsen, J.; Snyder, B.D.; Stribling, J.B. *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish*, 2nd ed.; Environmental Protection Agency, Office of Water: Washington, DC, USA, 1999; pp. 23–178.

8. Yoder, C. The integrated biosurvey as a tool for evaluation of aquatic life use attainment in Ohio surface waters. In *Biological Criteria: Research and Regulation*; Environmental Protection Agency: Washington, DC, USA, 1991; pp. 110–122.

9. Prygiel, J.; Carpentier, P.; Almeida, S.; Coste, M.; Druart, J.C.; Ector, L.; Guillard, D.; Honoré, M.A.; Iserentant, R.; Ledeganck, P.; et al. Determination of the biological diatom index (IBD NF T 90–354): Results of an intercomparison exercise. *J. Appl. Phycol.* **2002**, *14*, 27–39. [CrossRef]

10. Kelly, M.; Bennett, C.; Coste, M.; Delgado, C.; Delmas, F.; Denys, L.; Ector, L.; Fauville, C.; Ferréol, M.; Golub, M.; et al. A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: Results of an intercalibration exercise. *Hydrobiologia* **2009**, *621*, 169–182. [CrossRef]

11. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335. [CrossRef] [PubMed]

12. van Dijk, E.L.; Auger, H.; Jaszczyszyn, Y.; Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **2014**, *30*, 418–426. [CrossRef] [PubMed]

13. Esling, P.; Lejzerowicz, F.; Pawlowski, J. Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Res.* **2015**, *43*, 2513–2524. [CrossRef] [PubMed]

14. Evans, K.M.; Mann, D.G. A Proposed Protocol for Nomenclaturally Effective DNA Barcoding of Microalgae. *Phycologia* **2009**, *48*, 70–74. [CrossRef]

15. Dorigo, U.; Volatier, L.; Humbert, J.F. Molecular approaches to the assessment of biodiversity in aquatic microbial communities. *Water Res.* **2005**, *39*, 2207–2218. [CrossRef] [PubMed]

16. Creer, S. Second-generation sequencing derived insights into the temporal biodiversity dynamics of freshwater protists. *Mol. Ecol.* **2010**, *19*, 2829–2831. [CrossRef] [PubMed]

17. Baird, D.J.; Hajibabaei, M. Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Mol. Ecol.* **2012**, *21*, 2039–2044. [CrossRef] [PubMed]

18. Zhang, Y. Microbial Community Dynamics and Assembly: Drinking Water Treatment and Distribution. Ph.D. Thesis, University of Tennessee, Knoxville, TN, USA, 2012.

19. Stein, E.D.; Martinez, M.C.; Stiles, S.; Miller, P.E.; Zakharov, E.V. Is DNA Barcoding Actually Cheaper and Faster than Traditional Morphological Methods: Results from a Survey of Freshwater Bioassessment Efforts in the United States? *PLoS ONE* **2014**, *9*, e95525. [CrossRef] [PubMed]

20. Stein, E.D.; White, B.P.; Mazor, R.D.; Jackson, J.K.; Battle, J.M.; Miller, P.E.; Pilgrim, E.M.; Sweeney, B.W. Does DNA Barcoding Improve Performance of Traditional Stream Bioassessment Metrics? *Freshw. Sci.* **2013**, *33*, 302–311. [CrossRef]

21. Visco, J.A.; Apothéloz-Perret-Gentil, L.; Cordonier, A.; Esling, P.; Pillet, L.; Pawlowski, J. Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. *Environ. Sci. Technol.* **2015**, *49*, 7597–7605. [CrossRef] [PubMed]

22. Keck, F.; Vasselon, V.; Tapolczai, K.; Rimet, F.; Bouchez, A. Freshwater biomonitoring in the Information Age. *Front. Ecol. Environ.* **2017**, *15*, 266–274. [CrossRef]

23. Vasselon, V.; Rimet, F.; Tapolczai, K.; Bouchez, A. Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). *Ecol. Ind.* **2017**, *82*, 1–12. [CrossRef]

24. Zhang, J.; Zhu, C.; Guan, R.; Xiong, Z.; Zhang, W.; Shi, J.; Sheng, Y.; Zhu, B.; Tu, J.; Ge, Q.; et al. Microbial profiles of a drinking water resource based on different 16S rRNA V regions during a heavy cyanobacterial bloom in Lake Taihu, China. *Environ. Sci. Pollut. Res.* **2017**, *24*, 12796–12808. [CrossRef] [PubMed]

25. Glöckner, F.O.; Zaichikov, E.; Belkova, N.; Denissova, L.; Pernthaler, J.; Pernthaler, A.; Amann, R. Comparative 16S rRNA analysis of lake bacterioplankton reveals globally distributed phylogenetic clusters including an abundant group of Actinobacteria. *Appl. Environ. Microbiol.* **2000**, *66*, 5053–5065. [CrossRef] [PubMed]

26. Zhao, D.; Shen, F.; Zeng, J.; Huang, R.; Yu, Z.; Wu, Q.L. Network analysis reveals seasonal variation of co-occurrence correlations between Cyanobacteria and other bacterioplankton. *Sci. Total Environ.* **2016**, *573*, 817–825. [CrossRef] [PubMed]

27. Godhe, A.; Asplund, M.E.; Härnström, K.; Saravanan, V.; Tyagi, A.; Karunasagar, I. Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. *Appl. Environ. Microbiol.* **2008**, *74*, 7174–7182. [CrossRef] [PubMed]

28. Monchy, S.; Grattepanche, J.D.; Breton, E.; Meloni, D.; Sanciu, G.; Chabé, M.; Delhaes, L.; Viscogliosi, E.; Sime-Ngando, T.; Christaki, U. Microplanktonic community structure in a coastal system relative to a phaeocystis bloom inferred from morphological and tag pyrosequencing methods. *PLoS ONE* **2012**, *7*, e39924. [CrossRef] [PubMed]

29. Eiler, A.; Drakare, S.; Bertilsson, S.; Pernthaler, J.; Peura, S.; Rofner, C.; Simek, K.; Yang, Y.; Znachor, P.; Lindström, E.S. Unveiling Distribution Patterns of Freshwater Phytoplankton by a Next Generation Sequencing Based Approach. *PLoS ONE* **2013**, *8*, e53516. [CrossRef] [PubMed]

30. Alemzadeh, E.; Haddad, R.; Ahmadi, A.R. Phytoplanktons and DNA barcoding: Characterization and molecular analysis of phytoplanktons on the Persian Gulf. *Iran. J. Microbiol.* **2014**, *6*, 296–302. [PubMed]

31. Bazin, P.; Jouenne, F.; Friedl, T.; Deton-Cabanillas, A.F.; Le Roy, B.; Véron, B. Phytoplankton diversity and community composition along the estuarine gradient of a temperate macrotidal ecosystem: Combined morphological and molecular approaches. *PLoS ONE* **2014**, *9*, e94110.

32. Xiao, X.; Sogge, H.; Lagesen, K.; Tooming-Klunderud, A.; Jakobsen, K.S.; Rohrlack, T. Use of high throughput sequencing and light microscopy show contrasting results in a study of phytoplankton occurrence in a freshwater environment. *PLoS ONE* **2014**, *9*, e101560. [CrossRef] [PubMed]

33. Taylor, J.D.; Cottingham, S.D.; Billinge, J.; Cunliffe, M. Seasonal microbial community dynamics correlate with phytoplankton-derived polysaccharides in surface coastal waters. *ISME J.* **2013**, *8*, 245. [CrossRef] [PubMed]

34. Zimmermann, J.; Abarca, N.; Enk, N.; Skibbe, O.; Kusber, W.H.; Jahn, R. Taxonomic reference libraries for environmental barcoding: A best practice example from diatom research. *PLoS ONE* **2014**, *9*, e108793. [CrossRef] [PubMed]

35. Kermarrec, L.; Franc, A.; Rimet, F.; Chaumeil, P.; Frigerio, J.M.; Humbert, J.F.; Bouchez, A. A next-generation sequencing approach to river biomonitoring using benthic diatoms. *Freshw. Sci.* **2014**, *33*, 349–363. [CrossRef]

36. Berkoff, J. China: The south-north water transfer project-is it justified? *Water Policy* **2003**, *5*, 1–28.

37. Shao, X.; Wang, H.; Wang, Z. Interbasin transfer projects and their implications: A China case study. *Int. J. River Basin Manag.* **2003**, *1*, 5–14. [CrossRef]

38. Zhang, Q. The south-to-north water transfer project of China: Environmental implications and monitoring strategy. *J. Am. Water Resour. Assoc.* **2009**, *45*, 1238–1247. [CrossRef]

39. Shen, H.; Cai, Q.; Zhang, M. Spatial gradient and seasonal variation of trophic status in a large water supply reservoir for the south-to-north water diversion project, China. *J. Freshw. Ecol.* **2015**, *30*, 249–261. [CrossRef]

40. Li, Y.Y.; Gao, W.L.; Li, J.F.; Wen, Z.Z.; Liu, H.; Hu, L.Q.; Zhang, N.Q.; Cheng, X. Spatio-temporal distribution of phytoplankton and trophic status in the water resource area of the middle route of China's south to north water transfer project. *Chin. J. Ecol.* **2008**, *27*, 14–22. (In Chinese with English Abstract)

41. Chen, Z.; Ding, C.; Zhu, J.; Li, B.; Huang, J.; Du, Z.; Wang, Y.; Li, Y. Community structure and influencing factors of bacterioplankton during low water periods in Danjiangkou Reservoir. *China Environ. Sci.* **2017**, *37*, 336–344, (In Chinese with English Abstract). [CrossRef] [PubMed]

42. Hu, H.J.; Wei, Y.X. *The Freshwater Algae of China: Systematics, Taxonomy and Ecology*; Science and Technology Press: Beijing, China, 2006. (In Chinese)

43. Chinese State Environment Protection Bureau (CSEPB). *Water and Wastewater Monitoring Analysis Methods*, 4th ed.; Chinese Environment Science Press: Beijing, China, 2006; pp. 20–345. (In Chinese)

44. Shannon, C.E.; Wiener, W. *The Mathematical Theory of Communication*; University of Illinois Press: Urbana, IL, USA, 1949; p. 117.

45. Jin, X.C.; Tu, Q.Y. *The Standard Methods for Observation and Analysis of Lake Eutrophication*, 2rd ed.; China Environmental Science Press: Beijing, China, 1990; pp. 10–167. (In Chinese)

46. Lee, C.K.; Barbier, B.A.; Bottos, E.M.; McDonald, I.R.; Cary, S.C. The Inter-Valley Soil Comparative Survey: The ecology of Dry Valley edaphic microbial communities. *ISME J.* **2011**, *6*, 1046. [CrossRef] [PubMed]

47. Wang, Y.; Zheng, Y.; Wang, X.; Wei, X.; Wei, J. Lichen-associated fungal community in Hypogymnia hypotrypa (Parmeliaceae, Ascomycota) affected by geographic distribution and altitude. *Front. Microbiol.* **2016**, *7*, 1231. [CrossRef] [PubMed]

48. Water Resources Department of the People's Republic of China. *SL395-2007 Technological Regulations for Quality Assessment of Surface Water Resources in China*; China Water Conservancy and Hydropower Press: Beijing, China, 2007; pp. 1–5. (In Chinese)

49. Peres-Neto, P.R.; Jackson, D.A. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia* **2001**, *129*, 169–178. [CrossRef] [PubMed]

50. Medinger, R.; Nolte, V.; Pandey, R.V.; Jost, S.; OttenwÄLder, B.; SchlÖTterer, C.; Boenigk, J. Diversity in a hidden world: Potential and limitation of next-generation sequencing for surveys of molecular diversity of eukaryotic microorganisms. *Mol. Ecol.* **2010**, *19*, 32–40. [CrossRef] [PubMed]

51. Nolte, V.; Pandey, R.V.; Jost, S.; Medinger, R.; OttenwÄLder, B.; Boenigk, J.; SchlÖTterer, C. Contrasting seasonal niche separation between rare and abundant taxa conceals the extent of protist diversity. *Mol. Ecol.* **2010**, *19*, 2908–2915. [CrossRef] [PubMed]

52. Jahn, R.; Zetzsche, H.; Reinhardt, R.; Gemeinholzer, B. *Diatoms and DNA Barcoding: A Pilot Study on an Environmental Sample*; Kusber, W.H., Jahn, R., Eds.; Botanic Garden and Botanical Museum Berlin: Dahlem, Germany; Freie Universität Berlin: Berlin, Germany, 2007; pp. 63–68.

53. Cao, Y.; Williams, D.D.; Williams, N.E. How important are rare species in aquatic community ecology and bioassessment? *Limnol. Oceanogr.* **1998**, *43*, 1403–1409. [CrossRef]

54. Evans, K.M.; Wortley, A.H.; Mann, D.G. An assessment of potential diatom "barcode" genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in Sellaphora (Bacillariophyta). *Protist* **2007**, *158*, 349–364. [CrossRef] [PubMed]

55. Hamsher, S.E.; Evans, K.M.; Mann, D.G.; Poulíčková, A.; Saunders, G.W. Barcoding diatoms: Exploring alternatives to COI-5P. *Protist* **2011**, *162*, 405–422. [CrossRef] [PubMed]