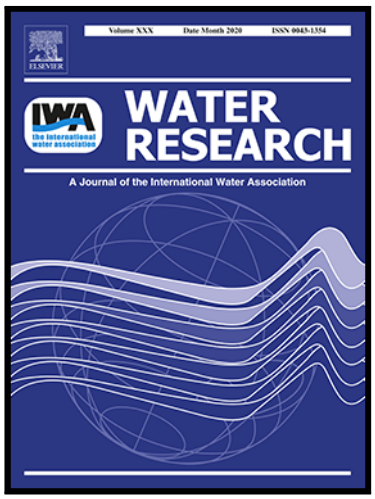


Using Cyanobacteria and Other Phytoplankton to Assess Trophic Conditions: A qPCR-Based, Multi-Year Study in Twelve Large Rivers across the United States

Chiqian Zhang , Kyle D. McIntosh , Nathan Sienkiewicz , Erin A. Stelzer , Jennifer L. Graham , Jingrang Lu

PII: S0043-1354(23)00114-8
DOI: https://doi.org/10.1016/j.watres.2023.119679
Reference: WR 119679

To appear in: *Water Research*

Received date: 26 October 2022
Revised date: 24 January 2023
Accepted date: 27 January 2023

Please cite this article as: Chiqian Zhang , Kyle D. McIntosh , Nathan Sienkiewicz , Erin A. Stelzer , Jennifer L. Graham , Jingrang Lu , Using Cyanobacteria and Other Phytoplankton to Assess Trophic Conditions: A qPCR-Based, Multi-Year Study in Twelve Large Rivers across the United States, *Water Research* (2023), doi: https://doi.org/10.1016/j.watres.2023.119679

Tables

Using Cyanobacteria and Other Phytoplankton to Assess Trophic Conditions: A qPCR-Based, Multi-Year Study in Twelve Large Rivers across the United States

Authors:

Chiqian Zhang[1], Kyle D. McIntosh[2], Nathan Sienkiewicz[3], Erin A. Stelzer[4], Jennifer L. Graham[5], and Jingrang Lu[3*]

Affiliations:

[1]Department of Civil and Environmental Engineering, College of Sciences and Engineering, Southern University and A&M College, Baton Rouge, Louisiana 70813, United States

[2]Oak Ridge Institute for Science and Education at the United States Environmental Protection Agency's Office of Research and Development, Oakridge 37830, TN

[3]Office of Research and Development, United States Environmental Protection Agency, Cincinnati, Ohio 45268, United States

[4]United States Geological Survey, Columbus, Ohio 43229, United States

[5]United States Geological Survey, Troy, New York 12180, United States

[*]Corresponding author. E-mail address: lu.jingrang@epa.gov (J. Lu)

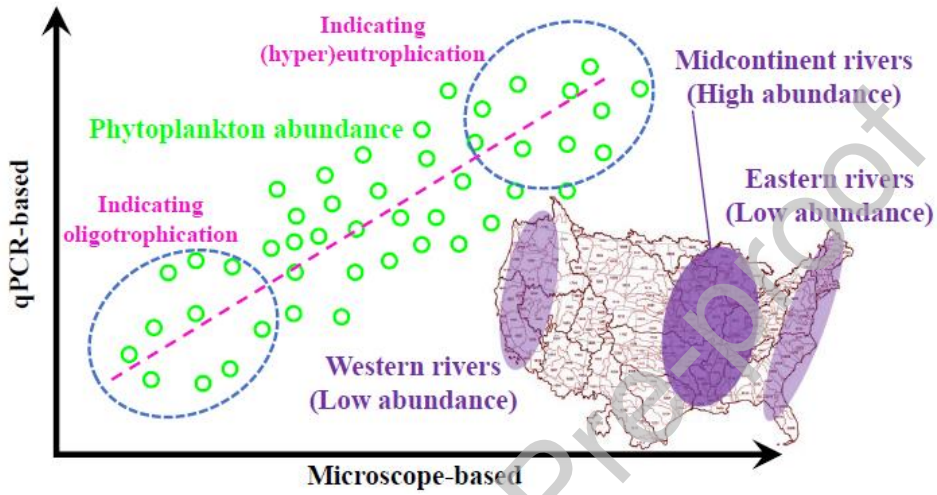
Tables

Highlights

- Assessed phytoplankton in 12 freshwater rivers across the U.S. using qPCR and microscopy
- qPCR- and microscope-based phytoplankton abundance had a significant positive correlation
- qPCR-based phytoplankton abundance was indicative of river trophic conditions
- qPCR is a promising numerical tool to quantify phytoplankton and assess river trophic status
- Higher phytoplankton abundance indicated the midcontinent river sites were more eutrophic

Graphical abstract

Tables



Abstract: Phytoplankton is the essential primary producer in fresh surface water ecosystems. However, excessive phytoplankton growth due to eutrophication significantly threatens ecologic, economic, and public health. Therefore, phytoplankton identification and quantification are essential to understanding the productivity and health of freshwater ecosystems, as well as the impacts of phytoplankton overgrowth (such as cyanobacterial blooms) on public health. Microscopy is the gold standard for phytoplankton assessment but is time-consuming, has low throughput, and requires rich experience in phytoplankton morphology. Quantitative polymerase chain reaction (qPCR) is accurate and straightforward with high throughput. In addition, qPCR does not require expertise in phytoplankton morphology. Therefore, qPCR can be a useful alternative tool for molecular identification and enumeration of

phytoplankton. Nonetheless, a comprehensive study is missing which evaluates and compares the feasibility of using qPCR and microscopy to assess phytoplankton in freshwater. This study focused on 1) comparing the performance of qPCR and microscopy in identifying and quantifying phytoplankton and 2) evaluating qPCR as a molecular tool to assess phytoplankton and indicate eutrophication. We assessed phytoplankton using both qPCR and microscopy in twelve large, freshwater rivers across the United States from early summer to late fall in 2017, 2018, and 2019. qPCR- and microscope-based phytoplankton abundance had a significant positive linear correlation (adjusted R^2 = 0.836, p-value < 0.001). Phytoplankton abundance had limited temporal variation within each sampling season and over the three years studied. The sampling sites in the midcontinent rivers had higher phytoplankton abundance than those in the eastern and western rivers. For instance, the concentration (geometric mean) of Bacillariophyta, Cyanobacteria, Chlorophyta, and Dinoflagellates at the sampling sites in the midcontinent rivers was approximately three times that at the sampling sites in the western rivers and approximately 18 times that at the sampling sites in the eastern rivers. Welch's analysis of variance indicated that phytoplankton abundance at the sampling sites in the midcontinent rivers was significantly higher than that at the sampling sites in the eastern rivers (p-value = 0.013) but was comparable to that at the sampling sites in the western rivers (p-value = 0.095). The higher phytoplankton abundance at the sampling sites in the midcontinent rivers was presumably because these rivers were more eutrophic. Indeed, low phytoplankton abundance occurred in oligotrophic or low trophic sites, whereas more eutrophic sites had greater phytoplankton abundance. This study demonstrated that qPCR-based phytoplankton abundance can be a useful numerical indicator of the trophic conditions and water quality in freshwater rivers.

Tables

Keywords: Microscopy; Freshwater bodies; Eutrophication; Spatiotemporal variation; Nutrients; Indicator

1. Introduction

Fresh surface water represents approximately 0.02% of the Earth's water but is the most important physical resource for society (Stets et al., 2020; Wilhelm et al., 2004). In fresh surface water, phytoplankton is a vital part of biogeochemical cycling and is the most critical primary producer. Phytoplankton photosynthesizes and provides food and energy to other organisms, ensuring normal and healthy ecosystem functioning and services (Dokulil and Qian, 2021; Li et al., 2022; Naselli-Flores and Padisák, 2022; Winder and Sommer, 2012). Eutrophication, the accumulation of nutrients (primarily nitrogen and phosphorus) and minerals in natural surface waterbodies due to natural and anthropogenic activities, threatens over 80% of rivers and lakes globally (Dodds, 2006; Smith et al., 2006; Sutcliffe and Jones, 1992; Wilhelm et al., 2004). In the United States (U.S.), total nitrogen (TN) and total phosphorus (TP) concentrations exceeded the reference median concentrations in 76% to 100% and 53% to 96%, respectively, of rivers in 14 ecoregions (Dodds et al., 2009). Over 90% of rivers in 12 of the 14 ecoregions had TN and TP levels higher than the reference median values. Overgrowth of phytoplankton, especially harmful or toxic Cyanobacteria, is the biological consequence of freshwater eutrophication (Duan et al., 2022; Lu et al., 2020; Wilhelm et al., 2004). Phytoplankton overgrowth deteriorates water quality, threatens ecosystems (e.g., aquatic organism mortality and biodiversity reduction), and can limit recreational opportunities. The overgrowth also poses public health risks and raises other issues by releasing harmful toxins (mainly cyanotoxins) and taste-and-odor compounds, causing the water to be unsafe and/or unpleasant for consumption and recreation (Chislock et al., 2013; Graham et al., 2010; Steffen et al., 2017; Wilhelm et al., 2004; Zhang et al., 2021b). In addition, the annual economic losses due to freshwater eutrophication are over 2.2 billion dollars in the U.S. (Dodds et al., 2009).

Tables

Phytoplankton overgrowth (e.g., algal and cyanobacterial blooms) is a consequence of eutrophication and poses substantial risks to the environment and public health. Therefore, evaluating phytoplankton is critical to understanding freshwater bodies' ecological conditions and trophic status (Aboim et al., 2020; Dembowska et al., 2018; Li et al., 2020; Sayers et al., 2021; Zhang et al., 2021c). Microscopy is the gold standard for assessing phytoplankton (Clementson et al., 2021; Ding et al., 2022; Dugdale et al., 2012; Hoang et al., 2018; Liu et al., 2019); however, microscopy is time-consuming, labor-intensive, and less efficient (Dunker, 2019; Malkassian et al., 2011; Rivas-Villar et al., 2021; Yu et al., 2021). Microscopy also requires rich experience in phytoplankton morphology and is difficult to operate for high-frequency (spatial and/or temporal) sample analysis. Moreover, microscopy is unable to identify cells with significant morphology changes (e.g., broken and shrunk cells), cannot discriminate cryptic species, cannot determine if a strain can produce toxins, and may be confounded by fixatives (Jeffrey and Vesk, 1997; Xiao et al., 2014). Furthermore, microscopic analysis of phytoplankton is prone to human error and bias (e.g., misidentification of species and missing picoplanktonic taxa) (Culverhouse, 2007; Luo et al., 2006; Palińska and Surosz, 2008).

To overcome the limitations of microscopy, researchers have developed alternative approaches to identify and quantify phytoplankton and study phytoplankton community structure. Those alternative methods include pigment-based approaches (often via high-performance liquid chromatography or spectrofluorometry) (Lee et al., 2020; Pan et al., 2020; Srichandan et al., 2020; Stoyneva-Gärtner et al., 2020; Yu et al., 2021), fatty-acid-based procedures (Cañavate et al., 2019; Canavate, 2019; Dijkman and Kromkamp,

2006; Galloway and Winder, 2015), analytical and imaging flow cytometry (Dunker, 2020; Latasa et al., 2021; Liu et al., 2021; Moorhouse et al., 2018; Read et al., 2014), and high-throughput sequencing or next-generation sequencing (dos Santos et al., 2021; Gong et al., 2020; Malashenkov et al., 2021; Yang et al., 2021b). These alternatives have advantages but also limitations. For instance, pigments (e.g., chlorophyll) are a widely used proxy for algal biomass and an indicator of trophic status (Chen and Chen, 2022; Dodds et al., 1998; Mineeva and Makarova, 2018). However, pigment-based approaches are insensitive and inaccurate (i.e., unable to differentiate phytoplankton species accurately), thereby providing little information on phytoplankton community composition (Catherine et al., 2012; Tian et al., 2020; Yu et al., 2021). Fatty-acid-based methods have similar limitations to the pigment-based approaches because they also lack sufficient reference libraries (Cañavate et al., 2019). Flow cytometry is less accurate, requires complicated staining, is bottlenecked by small sample volumes processed, and has a low taxonomic resolution (Dashkova et al., 2017; Dubelaar and Jonker, 2000; Markina, 2019). Even though high-throughput sequencing is much better at evaluating phytoplankton community diversity compared to quantitative polymerase chain reaction (qPCR), it is not as good at quantifying abundances, is expensive, could be disrupted by sequencing errors, and requires extensive bioinformatic expertise to interpret the data (Poretsky et al., 2014; Zhou et al., 2015).

Because of the shortcomings of microscopy and the alternative assays, we propose using qPCR to screen phytoplankton in surface waterbodies. qPCR is a rapid, accurate, sensitive, robust, reproducible, high-throughput, and cost-effective methodology (Galazzo et al., 2020; Smith and Osborn, 2009). qPCR instruments are more accessible than flow cytometers and high-throughput sequencing

platforms (Azat, 2021; Raso and Biassoni, 2014). qPCR also does not require knowledge of phytoplankton taxonomy or morphology but can provide information about both abundance and taxonomic identification of phytoplankton. In addition, unlike pigment-based approaches and flow cytometry, qPCR can be coupled with high throughput sequencing (Jian et al., 2020) to assess phytoplankton community with high accuracy, large dynamic range, and high sample processing capability (i.e., sequencing of millions of DNA molecules in parallel) (Churko et al., 2013; Reuter et al., 2015). Furthermore, because of the rapid innovation in qPCR instrumentation and protocols, we can upgrade and scale up qPCR to a high throughput assay (Pearson et al., 2021; Porter and Hajibabaei, 2018; Wilcox et al., 2020).

Most studies use qPCR to assess genes encoding cyanotoxins (e.g., microcystins, cylindrospermopsin, anatoxin-a, and saxitoxin) and Cyanobacteria (especially cyanotoxin-producing Cyanobacteria) (Almuhtaram et al., 2021; Christensen et al., 2019; Eldridge et al., 2017; Kulabhusan and Campbell, 2021; Otten et al., 2015; Pacheco et al., 2016; Sagova-Mareckova et al., 2021; Schweitzer-Natan et al., 2019; Tavakoli et al., 2021; Zupančič et al., 2021). However, the feasibility of qPCR in identifying and quantifying phytoplankton, especially non-Cyanobacterial phytoplankton, in freshwater rivers has not been fully evaluated in a comparative study with microscopy.

This study aimed to evaluate the feasibility of qPCR in assessing phytoplankton by monitoring phytoplankton in twelve large, freshwater rivers across the U.S. from early summer to late fall in 2017, 2018, and 2019. Because freshwater phytoplankton

community dynamics are driven by local environmental factors (Stomp et al., 2011), we assessed how environmental parameters, such as nutrients, shaped the phytoplankton community. We hypothesized that qPCR- and microscopy-based phytoplankton abundance would be positively correlated, supporting the use of qPCR as a numerical tool for phytoplankton assessment.

2. Materials and methods

2.1. Study sites and river water sample collection

This study included twelve large, inland and coastal freshwater rivers in the western (two), midcontinent (five), and eastern (five) regions across the U.S. (Table 1 and Figure 1). In this study, the western region refers to the region in the west of the Sierra Nevada and Cascade ranges, and the eastern region refers to the region in the east of the Appalachian range. The midcontinent region refers to the region between the western and eastern regions. We collected river water samples in 2017, 2018, and 2019 from one sampling site in each river according to the United States Geological Survey (USGS)'s National Water Quality Assessment (NAWQA) protocol (USGS, 2022a). We used isokinetic sampling techniques to ensure that the river samples are representative. The sampling spanned early summer through late fall each year. Since phytoplankton does not grow well and has low abundance during cold seasons (i.e., winter and spring), especially in large streams (Savoy et al., 2019), we did not sample during cold seasons. Two sets of river water samples were collected in parallel during each sampling event, defined as river water collection from a site on a sampling day. The first set was an isokinetic sample for physicochemical water quality parameter determination and microscopic analysis of phytoplankton. The second set was a near-surface grab sample collected at the centroid of a river for qPCR. In this study, we collected

the samples for qPCR from the centroids of the streams to minimize microbial contamination of the sampling equipment. In addition, when a single sample location must be selected, sampling from the centroid of a flow is generally representative of the cross-sections of the streams (Graham et al., 2012). In total, we collected river water during 166 sampling events for qPCR (Table S1) and a similar number of sampling events for physicochemical water quality parameter determination and microscopic analysis. Additional details on sample collection are described elsewhere (Graham et al., 2020; USGS, 2022a; Zuellig et al., 2021).

2.2. Microscopic assessment of phytoplankton and physicochemical water quality determination

The microscopic assay for phytoplankton identification and enumeration is described somewhere else (Graham et al., 2020; Graham et al., 2021; Graham et al., 2022; King et al., 2020a; King et al., 2020b; Zuellig et al., 2021). Briefly, we collected river water samples in 500 mL high-density polyethylene (HDPE) bottles to shoulder-full using Teflon churns. Each river water sample was preserved with approximately 5 mL of acidified Lugol's iodine in the same bottle. In 2017 and 2018, we concentrated phytoplankton by settling and identified/enumerated phytoplankton using a light microscope (Komárek and Anagnostidis, 2007; Komárek, 2008; Komárek and Anagnostidis, 2008; Komárek, 2013; Komárek et al., 2014; Komárek, 2016; Rosen et al., 2017; Rosen et al., 2018; Wehr et al., 2015). In 2019, we quantified phytoplankton using a permanent *N*-(2-hydroxypropyl)methacrylamide mounting technique according to established protocols (phycotech.com). For all river samples, we identified and enumerated phytoplankton using an Olympus BX51 research-grade microscope (Olympus Corporation, Tokyo, Japan) and variable magnifications depending on identified taxa. We enumerated phytoplankton to the lowest possible taxonomic level (typically genus or species) via natural units (all samples) or cell

counts (2018 and 2019 samples). For the 2017 samples, we did not enumerate phytoplankton by cell count. A "natural unit" is the natural growth form of phytoplankton, such as an individual cell, a colony, or a filament. The "natural unit" is commonly used for microscopic phytoplankton enumeration (Bellinger and Sigee, 2015; Graham et al., 2004b; Lehman et al., 2021; Onyema, 2018; Vuorio et al., 2020). In this work, the microscope-based phytoplankton abundance is expressed in "$NU \cdot L^{-1}$," in which "NU" stands for "natural units." At least, 400 natural units were counted per sample (Rosen et al., 2018). Phytoplankton identification and enumeration methods and data for the 2017, 2018, and 2019 samples are available in King et al. (2020a) and King et al. (2020b), Graham et al. (2021), and Graham et al. (2022), respectively.

We also determined multiple physicochemical water quality parameters (Graham et al., 2020; USGS, 2022a; Zuellig et al., 2021). We analyzed how 21 critical physicochemical parameters affected the phytoplankton community: water temperature, river discharge, dissolved oxygen concentration, water pH, particulate nitrogen concentration, dissolved nitrogen concentration, TN concentration, total organic nitrogen concentration, dissolved organic nitrogen concentration, ammonia and ammonium nitrogen concentration, nitrite nitrogen concentration, nitrate nitrogen concentration, dissolved Kjeldahl nitrogen concentration, total Kjeldahl nitrogen concentration, TP concentration, dissolved phosphorus concentration, orthophosphate phosphorus concentration, dissolved organic carbon concentration, total carbon dioxide concentration, dissolved carbonate ion concentration, and dissolved bicarbonate ion concentration. Water quality data are available through the USGS National Water Information System database (USGS, 2022b).

2.3. qPCR assessment of phytoplankton

We quantified nine critical phytoplankton groups via qPCR following established protocols (Lu et al., 2016; Lu et al., 2017; Zhang et al., 2021a). The nine qPCR targets are (1) Bacillariophyta (division), (2) Cyanobacteria (division), (3) Chlorophyta (specific to three classes in the division Chlorophyta: Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae), (4) Dinoflagellates (class), (5) *Dolichospermum* (genus), (6) *Microcystis* (genus), (7) *Planktothrix* (genus), (8) *Synechococcus* (genus), and (9) *Cylindrospermopsis raciborskii* (species; updated name: *Raphidiopsis raciborskii*). *Dolichospermum*, *Microcystis*, *Planktothrix*, *Synechococcus*, and *C. raciborskii* are Cyanobacteria. We selected these nine targets because they are common and dominant phytoplankton in freshwater bodies (Duan et al., 2022; Lu et al., 2020; Lueangthuwapranit et al., 2011; Marshall, 2009; Wilhelm et al., 2004). The goal of this study was not to assess all phytoplankton in the twelve freshwater rivers using qPCR. Therefore, the analysis of certain phytoplankton groups common in fresh water, such as Prochlorophyta and Xanthophyceae (which are two classes in the division Ochrophyta), was excluded from this work. We included 50, 50, and 66 sampling events in 2017, 2018, and 2019, respectively (166 in total) for qPCR (Table S1). We included biological duplicates for 25 out of the 166 sampling events. We collected river water samples for qPCR in autoclaved and bleached 500 mL polypropylene bottles and filtered the samples after collection using Nuclepore polycarbonate filters (Whatman/GE Healthcare, Piscataway, New Jersey, U.S.A.) (Graham et al., 2020). The filters with captured biomass were stored in tubes containing acid-washed glass beads (i.e., bead tubes) and shipped to our laboratory in Cincinnati, Ohio, for qPCR analysis.

We added 400 µL of 1× Tissue & Cell Lysis Solution (Epicentre Technologies Corp.; Madison, Wisconsin, U.S.A.) into each bead tube and stored the bead tubes at -20 °C before genomic DNA extraction. We lysed the stored cells by shaking the bead tubes for 1 min with a Mini-Beadbeater-16 (BioSpec Products, Inc.; Bartlesville, Oklahoma, U.S.A.). We then centrifuged the tubes at 12,000 × *g* for 3 min at room temperature and transferred the supernatant containing DNA to sterile 1.5 mL microcentrifuge tubes. One microliter of Proteinase K (50 µg·µL^{-1}, Epicentre Technologies Corp.) was mixed with the supernatant of each sample, and the mixture was incubated at 65 ˚C for 15 min. Subsequently, 1 µL of RNase A (5 ng·µL^{-1}; Epicentre Technologies Corp.) was added to each supernatant sample, and the mixture was incubated at 37 ˚C for 30 min before 200 µL of MPC protein (Epicentre Technologies Corp.) was added. The final mixture was incubated on ice for 5 min before being centrifuged at 15,000 × *g* for 10 min at room temperature. We then purified total genomic DNA from the supernatant using a ZymoBIOMICS™ DNA Miniprep Kit (Zymo Research; Irvine, California, U.S.A.). The DNA samples were quantified with a Qubit Fluorometer (Thermo Fisher Scientific Inc.; Waltham, Massachusetts, U.S.A.) and stored at -20 ˚C until use.

We determined the abundance of the nine phytoplankton groups via SYBR® Green qPCR assays on a QuantStudio™ 6 Flex Real-Time PCR System (Applied Biosystems™; Waltham, Massachusetts, U.S.A.) and a QuantStudio™ 5 Flex Real-Time PCR System (Applied Biosystems™) (Table S2). Each qPCR master mix (20 µL) contained (final concentration or volume) 10 µL of 2× qPCR SYBR® Green Master Mix (Life Technologies Co.; Carlsbad, California, U.S.A.), 0.5 µL of primers (each; 250 nM; Integrated DNA Technologies, Inc.; Coralville, Iowa, U.S.A.), 2 µL of DNA template, 2 µL of Bovine Serum Albumin solution (0.1 ng·µL^{-1}, Life

Technologies Co.), and 5 μL of nuclease-free water. We diluted the original DNA extract for each sample ten times with molecular-grade water. Both the undiluted and the ten-fold diluted DNA extracts were used as qPCR templates. Assessing the DNA of the phytoplankton in both the undiluted and the ten-fold diluted DNA samples helped detect qPCR inhibition and operation errors (Zhang et al., 2021a).

The target genes were quantified relative to in-house constructed standard series. All qPCR standards were recombinant plasmid DNA with the same backbone (Invitrogen pCR™4-TOPO® TA vector, Thermo Fisher Scientific Inc.). The insert for each qPCR standard was a conventional PCR product amplified with the same primer set for the corresponding qPCR assay (Table S2). Each qPCR plate contained a triplicate standard series (ten-fold serial dilution) with plasmid concentration ranging from 10^1 to 10^6 copies per microliter (2017 and 2018 samples) or from 10^0 to 10^6 copies per microliter (2019 samples). The phytoplankton abundance in the river water samples is expressed in gene or genome copy number per liter ($GCN \cdot L^{-1}$), where one "GCN" is equal to one copy of the qPCR amplicon. In most cases, each qPCR amplicon has one copy in the genome of each qPCR target (nine targets in total). However, each genome of *Microcystis* has one or two copies of the qPCR amplicon and each genome of *Synechococcus* has one, two, or three copies of the qPCR amplicon. The limit of detection (LOD) in the DNA extracts was 3 $GCN \cdot \mu L^{-1}$ for each qPCR assay.

2.4. Data filtering, analysis, and reporting

Tables

We processed the data for physicochemical water quality and microscope-based phytoplankton abundance according to Graham et al. (2020), Zuellig et al. (2021), and USGS (2022a). We filtered qPCR results (e.g., removal of data points with significant PCR inhibition, incorrect melt curves, and other issues) following Zhang et al. (2021a). If the concentrations of a qPCR target in the undiluted and the ten-fold diluted DNA extracts for a sampling event were both below 1 $GCN \cdot \mu L^{-1}$ or "undetermined" (i.e., no significant fluorescent signal after 40 qPCR cycles), we assumed that the concentration of that target was 1 $GCN \cdot L^{-1}$. The assumption was made because a zero value (i.e., 0 $GCN \cdot L^{-1}$) would complicate downstream data processing (especially the calculation of geometric means or GMs). If the concentration of a qPCR target in a DNA extract was below the LOD (3 $GCN \cdot \mu L^{-1}$) but higher than 1 $GCN \cdot \mu L^{-1}$, we assumed that the concentration was accurate (Klymus et al., 2020; Kralik and Ricchi, 2017).

We used Microsoft Office's Excel® (Professional Plus 2021; version 2112; Microsoft Corporation, Redmond, Washington, U.S.A.) for data management and figure plotting. We used SPSS® Statistics (version 26.0; The International Business Machines Corporation, Armonk, New York, U.S.A.) for statistical analysis, such as analysis of variance (ANOVA). The major assumptions for ANOVA, such as normal distributions, were also checked. A constrained canonical correspondence analysis (CCA) map was generated with an Excel® add-in XLSTAT (version 2019.2.2) (Addinsoft, 2022). The significance level was set at 0.05.

3. Results and discussion

3.1. Water quality conditions of the twelve rivers across the U.S.

Rivers are important, representative freshwater ecosystems, acting as a conveyor belt transporting nutrients and pollutants from the land to the oceans (Wang et al., 2018). Therefore, we utilized freshwater rivers to represent freshwater bodies. We strategically selected the twelve large freshwater rivers to include various geographic regions, stream drainage areas, and streamflow rates so that the rivers could be suggestive of the streams in the regions (Graham et al., 2020; Zuellig et al., 2021). Those rivers are in three major regions in the U.S.: the western region (two rivers), the midcontinent region (five rivers), and the eastern region (five rivers) (Table 1 and Figure 1).

The twelve rivers included both inland and coastal rivers. The drainage areas [arithmetic mean (AM) ± standard deviation: 175,953 ± 361,458 km^2] of the rivers ranged from 6,294 km^2 (the Chattahoochee River) to 1,353,269 km^2 (the Missouri River). From the water years of 2017 to 2019, the annual average volumetric flow rates (1,230 ± 1,510 $m^3 \cdot s^{-1}$) at the sampling sites ranged from 54 $m^3 \cdot s^{-1}$ (the Trinity River, 2017) to 5,870 $m^3 \cdot s^{-1}$ (the Ohio River, 2019). The USGS defines a water year as 12 months from October 1st for any given calendar year through September 30th of the following year.

The sampling sites in the twelve large rivers had substantially different water quality conditions from water years 2017 to 2019 (Table 1). The annual average suspended solids (SS) concentrations (95.2 ± 118.5 $mg \cdot L^{-1}$) ranged from 14.5 $mg \cdot L^{-1}$ (the Willamette River, 2017) to 402 $mg \cdot L^{-1}$ (the Missouri River, 2017). The annual average TN concentrations (2,202 ± 2,016 $\mu g \cdot L^{-1}$) ranged from 315 $\mu g \cdot L^{-1}$ (the Sacramento River, 2019) to 8,040 $\mu g \cdot L^{-1}$ (the Trinity River, 2017). From the water years of 2017 to 2019, the annual

average TP concentrations (201 ± 221 μg·L^{-1}) ranged from 46 μg·L^{-1} (the Connecticut River, 2019) to 833 μg·L^{-1} (the Trinity River, 2017). The summer median concentrations of TN and TP for rivers in 14 ecoregions across the U.S. ranged from 248 to 3,372 μg·L^{-1} (1,173 ± 839 μg·L^{-1}) and from 21 to 184 μg·L^{-1} (89 ± 52 μg·L^{-1}), respectively (Dodds et al., 2009). Therefore, the TN and TP concentrations at the sampling sites in the twelve rivers encompassed ranges of nutrient conditions of rivers in various ecoregions across the U.S.

The sampling sites in the three regions had significantly different TN and TP concentrations (one-way ANOVA, *p*-values < 0.001). Tukey *post hoc* tests revealed that TN and TP concentrations were significantly higher at the sampling sites in the midcontinent rivers than those in the eastern and western rivers (*p*-values ≤ 0.001). However, both TN and TP concentrations were comparable at the sampling sites in the western and eastern rivers (*p*-values > 0.5). Nutrient concentration is a commonly used indicator for the trophic status of natural surface waterbodies (Dai et al., 2016; Huo et al., 2013; Knoll et al., 2015; Yao et al., 2019). Therefore, on the basis of nutrients, the sampling sites in the midcontinent rivers were more eutrophic than those in the eastern and western rivers. By contract, the sampling sites in the eastern and western rivers had similar trophic statuses. In addition, the sampling sites in the three regions had significantly different SS concentrations (one-way ANOVA, *p*-value = 0.001). The sampling sites in the midcontinent rivers had significantly higher SS concentrations than those in the western and eastern rivers (Turkey *post hoc* tests, *p*-values < 0.011). By contrast, the sampling sites in the western and eastern rivers had comparable SS concentrations (Turkey *post hoc* test, *p*-value > 0.05). Previous studies similarly found that rivers in the midcontinent region or the midwest region in the U.S. typically had higher nutrient

concentrations and poorer water quality than rivers in the western and eastern regions (Graham et al., 2020; Smith et al., 1997; Stets et al., 2015; Zuellig et al., 2021). The higher nutrient concentrations in midcontinent rivers could be due to various reasons, both natural and anthropogenic (Khatri and Tyagi, 2015). For instance, the more intensive agricultural activities (higher use of fertilizers with nitrogen and phosphorus) in the midcontinent region might have contributed to the higher river water nutrient concentrations in that region (Najwah and Philip, 2021; Stets et al., 2015). Soils in the midcontinent region are also naturally nutrient rich (Omernik, 1987; Omernik and Griffith, 2014; U.S. EPA, 2022). In addition, land use and climate changes are stronger in the midcontinent region (Chapin et al., 1997). Furthermore, the sampling site in the Trinity River in the midcontinent region had the highest nutrient concentrations in the current study, which could be because the sampling site was downstream of a municipal wastewater treatment plant. This site was built by the USGS and has been sampled for decades. In addition, no USGS sampling site upstream of the treatment plant exists. Other sampling sites were not strongly affected by wastewater. Future studies need to explore further the reasons for the higher nutrient concentrations in rivers in the midcontinent region.

This study used nutrient (primarily TN and TP) levels as the surrogate of eutrophication (de Jonge et al., 2002; Harper, 1992; Richardson and Jørgensen, 1996). TN and TP are widely used indicators of eutrophication in waterbodies (Dai et al., 2016; Huo et al., 2013; Knoll et al., 2015; Yao et al., 2019). Widely accepted TN concentration boundaries for delineating oligotrophic from mesotrophic streams and mesotrophic from eutrophic streams are 700 and 1,500 $\mu g \cdot L^{-1}$, respectively (Dodds et al., 1998). TP concentration boundaries for delineating oligotrophic from mesotrophic streams and mesotrophic from eutrophic streams are 25 and

75 $\mu g \cdot L^{-1}$, respectively. On the basis of these TN and TP concentration boundaries, trophic conditions varied among the sampling sites in the twelve rivers (Table 1). The sampling site in the Chattahoochee River in the eastern region and all sampling sites in the midcontinent region were eutrophic sites. The sampling site in the Susquehanna River in the eastern region and the sampling site in the Willamette River in the western region were mesotrophic sites. In the eastern region, the Delaware River sampling site and the Hudson River sampling site were mesotrophic (indicated by TN concentrations) to eutrophic (indicated by TP concentrations). The sampling site in the Connecticut River in the eastern region and the sampling site in the Sacramento River in the western region were oligotrophic (indicated by TN concentrations) to mesotrophic (indicated by TP concentrations).

Because the sampling sites in the twelve rivers covered a wide range of geographical regions, stream drainage areas, streamflow rates, nutrient concentrations, and trophic conditions, our study locations suggested a wide diversity of freshwater river conditions across the U.S. Therefore, the findings of the current work can suggest the general conditions of freshwater rivers in the U.S.

3.2. qPCR as a useful numerical tool to assess phytoplankton

We compared the qPCR- and microscope-based abundance (Graham et al., 2020; Graham et al., 2021; Graham et al., 2022; King et al., 2020a) for eight phytoplankton groups at the sampling sites in the twelve rivers in 2017, 2018, and 2019 (Figure 2): Bacillariophyta, Cyanobacteria, Chlorophyta, Dinoflagellates, *Dolichospermum*, *Microcystis*, *Planktothrix*, and *Synechococcus*. *C. raciborskii* was excluded from this analysis because it was not identified by microscopy. In this comparison, the microscope-based phytoplankton

abundance is in natural units because the abundance in cell counts is available for only the 2018 and 2019 samples. Overall, the qPCR- and microscope-based phytoplankton abundance had a significant positive linear correlation (adjusted R^2 = 0.836, p-value < 0.001, n = 815) (Equation 1).

$$A_{qPCR} = 1.176 \times A_{Micriscopy} + 0.803 \quad \text{(Equation 1)}$$

where A_{qPCR} is the qPCR-based phytoplankton abundance [$\log_{10}(GCN \cdot L^{-1})$] and $A_{Microscopy}$ is the microscope-based phytoplankton abundance [$\log_{10}(NU \cdot L^{-1})$].

For each of the eight phytoplankton groups, the linear correlation between qPCR- and microscope-based abundance was also statistically significant with a relatively large R-squared value (average R-squared value 0.556) (Table 2). These results suggest that qPCR is a feasible method of assessing phytoplankton abundance and community composition in the twelve rivers and presumably general natural surface waterbodies. Future studies need to evaluate and confirm the feasibility of qPCR for phytoplankton assessment in other waterbodies.

For most sampling events, the qPCR-based phytoplankton abundance was equal to or higher than the microscope-based abundance (Table 2 and Figure 2). Indeed, the overall linear regression between qPCR- and microscope-based phytoplankton abundance showed higher qPCR-based abundance (Equation 1). Four reasons can explain the higher qPCR-based abundance. First, qPCR could

overestimate the abundance of phytoplankton by detecting extracellular DNA in the river water samples (Bairoliya et al., 2021; Liang and Keeley, 2013; Nagler et al., 2018; Nielsen et al., 2007). Second, phytoplankton cells with significant morphology changes are challenging to pick up during microscopic analysis, whereas qPCR accurately enumerates those cells. Similarly, small phytoplankton cells (sizes less than 3 μm) and rare phytoplankton taxa can escape the detection of microscopy but can be easily quantified by qPCR (Brewin et al., 2014; Graham et al., 2020; Malashenkov et al., 2021; Xiao et al., 2014). For instance, *Synechococcus* cells are small, and the linear correlation between qPCR- and microscope-based *Synechococcus* abundance is the weakest (i.e., having the smallest adjusted *R*-squared value among the eight phytoplankton groups). Therefore, microscopy might underestimate phytoplankton abundance. For example, we generated 1,304 pairs of data points for the qPCR- and microscope-based phytoplankton abundance from 163 sampling events in 2017, 2018, and 2019. For 477 out of the 1,304 pairs of data points, qPCR detected phytoplankton, while microscopy failed to detect phytoplankton (Figure 1). By contrast, in only 12 of the 1,304 pairs of data points, microscopy detected phytoplankton, while qPCR failed to recover phytoplankton. Interestingly, 11 of those 12 pairs of data points were for *Planktothrix*, where the microscope-based *Planktothrix* abundance was in the range from 2.31 to 3.91 $\log_{10}(NU \cdot L^{-1})$. This finding indicates that the qPCR assay targeting *Planktothrix* (Table S2) often underestimates the abundance, and an updated, more accurate qPCR assay for *Planktothrix* is needed. Third, we determined phytoplankton abundance using microscopy and qPCR in single-composite samples and near-surface grab samples, respectively. The different samples for microscopic and qPCR analyses could be another potential reason for the higher qPCR-based phytoplankton abundance. Lastly, in the current study, we quantified phytoplankton in natural units using microscopy (Graham et al., 2020). In the cases of colonies and filaments, multiple cells comprise a natural unit because a natural unit

does not include the enumeration of individual cells (BS EN 15204:2006, 2006; Graham et al., 2020). Therefore, for colonies and filaments, the number of natural units determined by microscopy would be lower than the number of genome copies determined by qPCR (Figure 2, Table 2, and Equation 1). For the 2019 samples, we estimated the number of cells in colonies and filaments and calculated the total number of cells in each natural unit of phytoplankton. We then compared the qPCR-based phytoplankton abundance ($GCN \cdot L^{-1}$) and microscope-based phytoplankton abundance ($cells \cdot L^{-1}$). The abundance showed a strong positive linear correlation with comparable unstandardized coefficient and constant values to the ones in Equation 1 (data not shown). As a result, we believe the first three explanations would be more important reasons for the higher qPCR-based phytoplankton abundance.

Microscopy and qPCR both have strengths and shortcomings (Pacheco et al., 2016; Xiao et al., 2014). For instance, a microscope is easy to operate (i.e., a low technology threshold) and directly describes phytoplankton morphology. However, microscopic analysis is time-consuming, has low throughput, and requires taxonomic or morphology experience. qPCR has high throughput and requires no taxonomic knowledge. However, qPCR fails to describe phytoplankton morphology and viability and does not cover all phytoplankton taxa. In addition, the sensitivity, efficiency, and accuracy of qPCR are compromised by DNA isolation bias, DNA amplification bias, and inhibitors (such as humic acid in environmental samples) (Acharya et al., 2017; Dechesne et al., 2016; Lima et al., 2022; Opel et al., 2010). Furthermore, a specific qPCR assay (i.e., a specific primer set and a specific thermocycling condition) can detect only one phytoplankton group (or closely related groups). To detect multiple phytoplankton groups, one needs to use multiple qPCR assays. Therefore, abandoning microscopy and adopting qPCR as the sole method for phytoplankton assessment is inadequate. Instead,

combining use of qPCR and microscopy would more comprehensively reveal phytoplankton abundance and community structure and provide deeper insights into phytoplankton dynamics in freshwater bodies (Malashenkov et al., 2021).

3.3. Low temporal variation in phytoplankton abundance

Phytoplankton abundance had limited temporal variation at the sampling sites in the twelve rivers within each sampling season. At each sampling site, the qPCR-based phytoplankton abundance fluctuated within each sampling season (Figures S1 and S2). No universal trend in the temporal variation of phytoplankton abundance across the twelve sampling sites is observed. In addition, at some sampling sites, the patterns of the temporal variation of phytoplankton abundance changed substantially across in the three years studied. For instance, the phytoplankton abundance at the sampling site in the Kansas River increased during the 2017 sampling season, decreased during the 2018 sampling season, and was relatively stable during the 2019 sampling season (Figures S1D and S2D). Future studies need to explore the underlying reasons for the distinct temporal variation of phytoplankton abundance both within and among sampling sites across years.

Even though phytoplankton abundance fluctuated within each sampling season, it was generally stable (Figures S1 and S2). The relatively stable phytoplankton abundance over time within each sampling season was presumably because the sampling period was from early summer to late fall, covering primarily the warm seasons when phytoplankton abundance is typically high. Future studies should extend the sampling period to assess the seasonal variation in phytoplankton abundance in these twelve rivers.

Phytoplankton abundance also had limited temporal variation throughout the three years studied (2017, 2018, and 2019). The qPCR-based phytoplankton abundance was comparable over the three years (Figure 3). Indeed, repeated measures ANOVA indicated that the annual GM concentrations were comparable for most phytoplankton groups across all three years at the sampling sites in eleven rivers (Table S3). The sampling site in the Trinity River was excluded from the analysis because the microscope-based phytoplankton abundance for that site was available for only 2019. Therefore, phytoplankton at the sampling sites in the eleven rivers occurred from early summer to late fall at comparable annual abundance. The comparable annual phytoplankton abundance indicated that the environmental conditions (i.e., drivers of phytoplankton community dynamics) for these sampling sites did not vary substantially from 2017 to 2019 (Stomp et al., 2011). Other factors might also contribute to the comparable annual phytoplankton abundance over the three years and should be explored in future studies. In other freshwater bodies, clear temporal (interannual) trends in phytoplankton abundance and community dynamics existed, which could be due to the substantial changes in local and global environmental conditions such as nutrient input, water regime, and global warming (Alvarez-Cobelas et al., 2019; Chen et al., 2003; Hardenbicker et al., 2014; Verasztó et al., 2010; Znachor et al., 2020). To assess the sustained responses of phytoplankton community dynamics to changes in local and global environmental conditions such as eutrophication management (e.g., nutrient loading reduction efforts) and global climate change, we need to conduct decadal or even multidecadal studies (Duarte et al., 2009; Gallegos et al., 2010). Studying the phytoplankton communities in a large number of waterbodies at the same time with varying local environmental conditions, such as different nutrient inputs, is also needed.

3.4. Moderately higher phytoplankton abundance in the midcontinent rivers

The sampling sites in the midcontinent rivers generally had higher phytoplankton abundance than the sites in the eastern and western rivers (Figures 1, 4, and S3 to S6). For instance, the GM concentration of Bacillariophyta, Cyanobacteria, Chlorophyta, and Dinoflagellates at the sampling sites in the midcontinent rivers [7.31 $\log_{10}(GCN \cdot L^{-1})$] was approximately three times that at the sampling sites in the western rivers [6.81 $\log_{10}(GCN \cdot L^{-1})$] and approximately 18 times that at the sampling sites in the eastern rivers [6.06 $\log_{10}(GCN \cdot L^{-1})$]. Welch's ANOVA indicated that the $\log_{10}$-transformed GM abundance of phytoplankton at the sampling sites in the midcontinent rivers was significantly higher than that at the sampling sites in the eastern rivers (p-value = 0.013) but was comparable to that at the sampling sites in the western rivers (p-value = 0.095) (Table S4). Additionally, the $\log_{10}$-transformed GM abundance of phytoplankton at the sampling sites in the eastern and western rivers was comparable (p-value = 0.949). In conclusion, the sampling sites in the midcontinent rivers had higher phytoplankton abundance than the sampling sites in the western rivers and especially the sampling sites in the eastern rivers. A study similarly found that midcontinent lakes in the U.S. had higher phytoplankton species richness than other regions (Stomp et al., 2011). The higher phytoplankton abundance at the sampling sites in the midcontinent rivers was presumably due to the higher TN and TP concentrations at those sites (Table 1). Rivers in the midcontinent region typically have higher nutrient levels because soils in that region are naturally rich in nutrients and that region has more intensive agricultural activities (Omernik, 1987; Omernik and Griffith, 2014; Stets et al., 2015; U.S. EPA, 2022).

3.5. The positive correlation between nutrient concentrations and phytoplankton abundance

We investigated how critical physicochemical parameters (mainly nutrients) shaped the phytoplankton community at the sampling sites in the twelve large, freshwater rivers across the U.S. (Figure S7). Phytoplankton abundance was often significantly and linearly correlated with nutrient concentrations, water temperature, and water pH (*p*-values < 0.05) (Table S5). In addition, a constrained CCA map showed that phytoplankton abundance (except for *Planktothrix*) was associated with intermediate or high nutrient concentrations and water temperatures (Figure 5). For instance, Bacillariophyta, Cyanobacteria, Chlorophyta, and Dinoflagellates were associated with intermediate levels of ammonium/ammonia, nitrite, nitrate, dissolved organic nitrogen, and orthophosphate phosphorus. Similarly, phytoplankton abundance in 540 freshwater lakes and reservoirs across the U.S. had significant positive linear correlations with both TN and TP concentrations (*p*-values < 0.001). By contrast, water temperature had a significant positive linear correlation with phytoplankton richness (*p*-value < 0.001) (Stomp et al., 2011). Additionally, the CCA map shows that the stream flow rate at the sampling sites was a driver for *Planktothrix*. Therefore, the phytoplankton community at the sampling sites in the twelve rivers was shaped by the critical physicochemical parameters (mainly nutrients), indicating local environmental conditions are the main drivers for phytoplankton community dynamics in rivers.

The correlations between phytoplankton abundance and the tested physicochemical parameters were generally weak, as indicated by the small *R*-squared values (Table S5). The low degrees of correlations indicated that the tested physicochemical parameters explained a small portion of the total variation in phytoplankton abundance. Indeed, the first and second axes in the constrained CCA map

explained only 23.52% and 5.96% of the total inertia, respectively (Figure 5). Therefore, some environmental parameters not tested in the current study might have more significantly affected phytoplankton abundances, such as light intensity, wind intensity, precipitation, stormwater runoff, and biological agents (predators and competitors) (Affronti Jr and Duquette, 2007; Cao et al., 2018; Elser and Hassett, 1994; Feng et al., 2021; Hunt and Matveev, 2005; Jones and Redfield, 1984; Kiefer and Cullen, 1991; Reichwaldt et al., 2004; Rhee and Gotham, 1981; Schelske et al., 1995; Silva et al., 2019; Striebel et al., 2008; Webster, 1990). Future studies should include more physicochemical parameters when determining how environmental conditions shape phytoplankton communities. The low degrees of correlations could also be because those critical physicochemical parameters were determined only several times during each sampling season (Table S1) and thus might not reflect the overall water quality (e.g., trophic conditions) at the sampling sites. The annual average TN and TP concentrations (numerous sampling events included) at the sampling sites could more accurately indicate the water quality at those sites. Indeed, phytoplankton abundance had positive linear correlations with the twelve sampling sites' annual average TN and TP concentrations (Figure 6). Compared with TN, TP explained more variation in phytoplankton abundance. Specifically, TN and TP individually explained 26.1% and 38.9% of the total variation in the $\log_{10}$-transformed GM abundance of phytoplankton, respectively. In a multiple linear regression model, TN and TP significantly predicted the $\log_{10}$-transformed GM abundance of phytoplankton (adjusted $R^2 = 0.385$, p-value < 0.001, $n = 34$) at the sampling sites (Equation 2).

$$\log_{10} A_{GM} = 0.266 \times \log_{10} C_{TN} + 0.766 \times \log_{10} C_{TP} + 4.268 \quad \text{(Equation 2)}$$

where A_{GM} is the annual GM of the qPCR-based abundance of four major phytoplankton groups (Bacillariophyta, Cyanobacteria, Chlorophyta, and Dinoflagellates) for individual sampling sites ($GCN \cdot L^{-1}$), C_{TN} is the annual AM of TN concentrations for individual sampling sites ($\mu g \cdot L^{-1}$), and C_{TP} is the annual AM of TP concentrations for individual sampling sites ($\mu g \cdot L^{-1}$).

The large multiple correlation coefficient (R) of 0.650 indicated a good level of prediction. The significant positive correlations between phytoplankton abundance and nutrient concentrations further suggested that qPCR-based phytoplankton abundance can indicate freshwater river trophic conditions. Even though nutrients explain a substantial amount of the variation in phytoplankton abundance, they are not the only factors affecting phytoplankton growth. For instance, the CCA map (Figure 5) indicates that phytoplankton abundance was affected by multiple other environmental variables, such as stream flow rate, water temperature, and water pH.

3.6. qPCR as a numerical tool for assessing phytoplankton abundance and trophic conditions of freshwater rivers

qPCR is a useful numerical tool for assessing phytoplankton and indicating trophic conditions of freshwater rivers. First, as stated above, microscopy and other approaches, such as pigment-based assays, for assessing phytoplankton have multiple shortcomings. qPCR, a widely used molecular technique, has many advantages over conventional microscopy and other phytoplankton evaluation assays. This study demonstrated that qPCR- and microscope-based phytoplankton abundance at the sampling sites in the rivers had a

strong positive linear correlation (Figure 2, Table 2, and Equation 1). Even though qPCR cannot and should not replace microscopy for phytoplankton assessment, it is a proper alternative technique to microscopy for identifying and quantifying phytoplankton.

Second, phytoplankton abundance indicates river trophic conditions. Trophic state index (TSI) is a common and widely accepted indicator for the trophic status of natural surface waterbodies (Bilgin, 2020; Hu et al., 2021; Kiersztyn et al., 2002). TSI can be based on transparency (i.e., Carlson's TSI) (Carlson, 1977), chlorophyll *a* concentration (i.e., the modified Carlson's TSI) (Aizaki et al., 1981), and other biological, chemical, and physical indicators such as COD, TN concentration, and TP concentration (Lopes et al., 2019; Primpas et al., 2010; Xu et al., 2001). No matter how TSI is determined or calculated, it is the numerical indicator for surface water trophic status (Lopes et al., 2019; Neverova-Dziopak et al., 2023; Yang et al., 2016). Overgrowth of phytoplankton (such as Cyanobacteria and algae) is the hallmark symptom or distinctive biological consequence of eutrophication (González and Roldán, 2019; Qin et al., 2013; Sari et al., 2022; Xu et al., 2010; Yang et al., 2021a). Therefore, phytoplankton abundance and TSI have an inherent positive correlation (Adamovich et al., 2016; Caputo et al., 2008; Dembowska et al., 2015). Altogether, phytoplankton abundance, which can be feasibly determined via qPCR, is a useful indicator of the trophic conditions of natural surface waterbodies.

Third, qPCR is a promising molecular tool for assessing phytoplankton and, thus, the trophic conditions of natural surface waterbodies. The growth or overgrowth of phytoplankton, especially Cyanobacteria, is the most direct and significant biological consequence of eutrophication in freshwater bodies (Wilhelm et al., 2004). Multiple assays are available to assess phytoplankton in fresh water. These

assays mainly include chlorophyll *a* concentration determination. Chlorophyll *a* is a photosynthetic pigment in all phytoplankton (Chew and Bryant, 2007; Gregor and Maršálek, 2004; Patel, 2011) and a good proxy for phytoplankton biomass, water quality, and eutrophication (Boyer et al., 2009; Chen and Chen, 2021; Filazzola et al., 2020; Guo et al., 2018; Xia and Zeng, 2021). Nonetheless, chlorophyll *a* assessment as a first-tier tool fails to provide information on phytoplankton community composition (Almuhtaram et al., 2021; Garmendia et al., 2013). Other commonly used approaches for phytoplankton assessment such as microscopy, fatty-acid-based procedures, flow cytometry, and next-generation sequencing, also have limitations. Therefore, a more feasible approach to assessing phytoplankton and indicating freshwater trophic status is highly needed. qPCR determines not only the abundance but also the community structure of phytoplankton. qPCR is also powerful in monitoring potential cyanotoxin producers and predicting cyanotoxin production (Duan et al., 2022; Lu et al., 2020; Pacheco et al., 2016). Moreover, the qPCR-based phytoplankton abundance was significantly correlated with nutrient (i.e., TN and TP) levels (Figures 5 and 6 and Equation 2). As a result, qPCR-based phytoplankton abundance is a useful indicator of the trophic conditions and health of freshwater bodies. However, caution should be taken for waterbodies with very high nutrient concentrations because hyper-eutrophication could inhibit phytoplankton's growth and primary production (Filstrup and Downing, 2017; Graham et al., 2004a; Parker et al., 2012; Yoshiyama and Sharp, 2006). In addition, nutrient concentration is a commonly used surrogate for freshwater trophic conditions, but the correlation between phytoplankton abundance and nutrient concentrations is weak (Figure 6). The weak correlation was mainly because other environmental parameters besides nutrients play an essential role in shaping the phytoplankton community (Figure 5). Therefore, the indicative role of qPCR for the trophic conditions and health of freshwater bodies should be more comprehensively assessed in future studies.

Phytoplankton community structure is complex, and quantifying all phytoplankton taxa using qPCR is labor-intensive, time-consuming, and perhaps impossible. Supposedly, a selected phytoplankton taxon can serve as an indicator of other phytoplankton. In this case, one may use qPCR to assess the selected taxon to represent the phytoplankton community and to indicate freshwater trophic status, reducing the workload and increase the feasibility of qPCR. Cyanobacteria are a common and frequently dominant phytoplankton division in many freshwater bodies and often form (harmful) cyanobacterial blooms during warm seasons (Duan et al., 2022; Lu et al., 2020; O'Neil et al., 2012; Paerl et al., 2001). The relatively high abundance of Cyanobacteria makes their detection, either through microscopy or qPCR, feasible. Cyanobacterial abundance can also potentially indicate cyanotoxin production in waterbodies. We thus hypothesized that Cyanobacteria can be an indicator of the phytoplankton community. To test this hypothesis, we first tested whether the abundance of Cyanobacteria was comparable to that of other phytoplankton. Statistical analysis indicated that the qPCR-based Cyanobacteria abundance was significantly higher than the other three major phytoplankton groups (Bacillariophyta, Chlorophyta, and Dinoflagellates) (Table S6). Therefore, Cyanobacteria were a dominant phytoplankton group at the sampling sites in the twelve freshwater rivers (Figures 3A and 4A). In addition, Cyanobacterial abundance had a significant positive linear correlation with other tested phytoplankton groups (except for *C. raciborskii*, which might have a different growth pattern than other common Cyanobacterial species) (Table S7). Therefore, a high or low abundance of Cyanobacteria suggests that the abundance of other phytoplankton is high or low, respectively. As a result, as a representative and common phytoplankton division, Cyanobacteria can predict and indicate the occurrence and abundance of other important phytoplankton. Using qPCR to detect and

quantify Cyanobacteria could be a simple, fast, and effective way to assess the phytoplankton community and trophic conditions of freshwater rivers (or riverine systems). However, it is worth noting that Cyanobacterial assessment is only a rough method to indicate the whole phytoplankton community. With time and resources permitting, other phytoplankton groups should also be assessed.

3.7. Limitations and future work

We sampled twelve large, freshwater rivers throughout the U.S. to capture a wide range of streamflow, water quality, and phytoplankton community conditions. Our study demonstrated that qPCR works well to assess phytoplankton abundance and to indicate river trophic status across the wide range of conditions. Our study design leveraged an ongoing sampling effort within the USGS NWQN. While this study design was adequate for method (i.e., qPCR) validation, it would not be valid to determine reach-level or ecosystem-level trophic status. Future studies designed to evaluate river trophic status need to compare traditional metrics with qPCR. qPCR cannot be incorporated into trophic status evaluations until further research and comparisons have been completed.

Three limitations of this study should be addressed by future research. First, we sampled from two, five, and five rivers from the western, midcontinent, and eastern regions of the U.S., respectively (Table 1 and Figure 1). Results based on the low number of rivers, especially in the western region, might need to be more representative. Future studies should include more rivers in each region to represent the stream conditions better. Second, we sampled from one site in each river. The ecosystem of each river is highly dynamic longitudinally. Water quality parameters, environmental conditions, and phytoplankton community composition vary significantly

along each river. Therefore, one sampling site cannot represent the entire ecosystem of a river. Phytoplankton abundance and community composition determined in the current work thus represent only the sampling sites rather than the entire rivers. Future studies aiming to assess the spatial variation of phytoplankton in these rivers should include multiple sampling sites along each river. Third, we leveraged an ongoing USGS NWQN sampling effort (USGS, 2022a) to conduct this study and added additional analyses to routinely collected samples within the NWQN. In the NWQN, the sampling effort differed for the regions or rivers. Therefore, the number of samples in the current work was unequal across the different regions or rivers. Future studies comparing phytoplankton abundance and composition in different regions and/or rivers need to ensure an equal sampling effort.

4. Conclusions

This work is the first qPCR-based, large-scale, multi-year, comprehensive study on the occurrence and abundance of phytoplankton in freshwater rivers across the U.S. This study demonstrated that qPCR is a useful molecular tool to quantify phytoplankton and assess eutrophication in freshwater rivers. Using qPCR and microscopy, we assessed dominant phytoplankton at the sampling sites in twelve large, freshwater rivers across the U.S. from early summer to late fall in 2017, 2018, and 2019. Phytoplankton abundance varied little within each sampling season and throughout the three years studied. The sampling sites in the midcontinent rivers had moderately higher phytoplankton abundance than the sampling sites in the eastern and western rivers primarily because the sampling sites in the midcontinent rivers were more eutrophic (i.e., higher TN and TP concentrations). Phytoplankton abundance was comparable at the sampling sites in the eastern and western rivers studied. The qPCR- and microscope-based phytoplankton abundance had a significant

positive linear correlation. In addition, qPCR-based phytoplankton abundance had significant positive correlations with nutrient levels, a surrogate measure of eutrophication. Therefore, qPCR is a useful molecular tool to determine phytoplankton abundance and indicate freshwater rivers' trophic conditions and quality. Finally, Cyanobacteria had a much higher abundance and significant positive linear correlations with other phytoplankton. Therefore, Cyanobacteria as a dominant phytoplankton group can be potentially used as an indicator representing the phytoplankton community and inferring trophic conditions in freshwater rivers.

Acknowledgments

This research was supported by the United States Environmental Protection Agency (U.S. EPA)'s Office of Research and Development's research program Safe and Sustainable Water Resources (SSWR 5.2.1) and the United States Geological Survey's (USGS's) National Water Quality Program's National Water Quality Assessment Project. The New York State Department of Environmental Conservation supported sampling from the Hudson River. We sincerely thank Ian Struewing, Eric Villegas, and Mark Bagley (Office of Research and Development, U.S. EPA) for their technical support, constructive comments, and helpful suggestions. The views expressed in this manuscript do not necessarily represent the views or policies of the U.S. EPA. The manuscript has been subjected to the U.S. EPA, and the USGS has reviewed and approved for publication. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Declaration of Interest Statement

All authors claim that no actual or potential conflict of interest exists in relation to this study.

References

Aboim, I.L., Gomes, D.F., and Mafalda Junior, P.O. 2020. Phytoplankton response to water quality seasonality in a Brazilian neotropical river. *Environmental Monitoring and Assessment* 192, 70.

Acharya, K.R., Dhand, N.K., Whittington, R.J., and Plain, K.M. 2017. PCR inhibition of a quantitative PCR for detection of *Mycobacterium avium* subspecies *Paratuberculosis* DNA in feces: diagnostic implications and potential solutions. *Frontiers in Microbiology* 8, 115.

Adamovich, B., Zhukova, T., Mikheeva, T., Kovalevskaya, R., and Luk'yanova, E. 2016. Long-term variations of the trophic state index in the Narochanskie Lakes and its relation with the major hydroecological parameters. *Water Resources* 43(5), 809-817.

Affronti Jr, L.F., and Duquette, B.T. 2007. Stormwater influence on phytoplankton composition and dynamics in Lake Joyce, Virginia. *Virginia Journal of Science* 58(4), 2.

Aizaki, M., Otsuki, A., Fukushima, T., Hosomi, M., and Muraoka, K. 1981. Application of Carlson's trophic state index to Japanese lakes and relationships between the index and other parameters: with 2 figures and 4 tables in the text. *SIL Proceedings, 1922-2010 (Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen)* 21(1), 675-681.

Almuhtaram, H., Kibuye, F.A., Ajjampur, S., Glover, C.M., Hofmann, R., Gaget, V., Owen, C., Wert, E.C., and Zamyadi, A. 2021. State of knowledge on early warning tools for cyanobacteria detection. *Ecological Indicators* 133, 108442.

Alvarez-Cobelas, M., Rojo, C., and Benavent-Corai, J. 2019. Long-term phytoplankton dynamics in a complex temporal realm. *Scientific Reports* 9, 15967.

Azat, C. 2021. Not just a pathogen: the importance of recognizing genetic variability to mitigate a wildlife pandemic. *Molecular Ecology Resources* 21(5), 1410-1412.

Bairoliya, S., Koh Zhi Xiang, J., and Cao, B. 2021. Extracellular DNA in environmental samples: occurrence, extraction, quantification, and impact on microbial biodiversity assessment. *Applied and Environmental Microbiology* 88(3), e0184521.

Bilgin, A. 2020. Trophic state and limiting nutrient evaluations using trophic state/level index methods: a case study of Borçka Dam Lake. *Environmental Monitoring and Assessment* 192, 794.

Boyer, J.N., Kelble, C.R., Ortner, P.B., and Rudnick, D.T. 2009. Phytoplankton bloom status: Chlorophyll a biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. *Ecological Indicators* 9(6), S56-S67.

Brewin, R.J., Sathyendranath, S., Lange, P.K., and Tilstone, G. 2014. Comparison of two methods to derive the size-structure of natural populations of phytoplankton. *Deep Sea Research Part I: Oceanographic Research Papers* 85, 72-79.

Cañavate, J.-P., van Bergeijk, S., Giráldez, I., González-Ortegón, E., and Vílas, C. 2019. Fatty acids to quantify phytoplankton functional groups and their spatiotemporal dynamics in a highly turbid estuary. *Estuaries and Coasts* 42(8), 1971-1990.

Canavate, J.P. 2019. Advancing assessment of marine phytoplankton community structure and nutritional value from fatty acid profiles of cultured microalgae. *Reviews in Aquaculture* 11(3), 527-549.

Cao, J., Hou, Z., Li, Z., Chu, Z., Yang, P., and Zheng, B. 2018. Succession of phytoplankton functional groups and their driving factors in a subtropical plateau lake. *Science of the Total Environment* 631, 1127-1137.

Caputo, L., NASELLI‐FLORES, L., Ordonez, J., and Armengol, J. 2008. Phytoplankton distribution along trophic gradients within and among reservoirs in Catalonia (Spain). *Freshwater Biology* 53(12), 2543-2556.

Carlson, R.E. 1977. A trophic state index for lakes. *Limnology and Oceanography* 22(2), 361-369.

Catherine, A., Escoffier, N., Belhocine, A., Nasri, A., Hamlaoui, S., Yéprémian, C., Bernard, C., and Troussellier, M. 2012. On the use of the FluoroProbe®, a phytoplankton quantification method based on fluorescence excitation spectra for large-scale surveys of lakes and reservoirs. *Water Research* 46(6), 1771-1784.

Chapin, F.S., III, Walker, B.H., Hobbs, R.J., Hooper, D.U., Lawton, J.H., Sala, O.E., and Tilman, D. 1997. Biotic control over the functioning of ecosystems. *Science* 277(5325), 500-504.

Chen, C.-K., and Chen, Y.-C. 2021. Detection of chlorophyll fluorescence as a rapid alert of eutrophic water. *Water Supply* 22(3), 3508–3518.

Chen, C.-K., and Chen, Y.-C. 2022. Detection of chlorophyll fluorescence as a rapid alert of eutrophic water. *Water Supply* 22(3), 3508-3518.

Chen, Y., Qin, B., Teubner, K., and Dokulil, M.T. 2003. Long-term dynamics of phytoplankton assemblages: *Microcystis*-domination in Lake Taihu, a large shallow lake in China. *Journal of Plankton Research* 25(4), 445-453.

Chew, A.G.M., and Bryant, D.A. 2007. Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. *Annual Review of Microbiology* 61, 113-129.

Chislock, M.F., Doster, E., Zitomer, R.A., and Wilson, A.E. 2013. Eutrophication: causes, consequences, and controls in aquatic ecosystems. *Nature Education Knowledge* 4(4), 10.

Christensen, V.G., Maki, R.P., Stelzer, E.A., Norland, J.E., and Khan, E. 2019. Phytoplankton community and algal toxicity at a recurring bloom in Sullivan Bay, Kabetogama Lake, Minnesota, USA. *Scientific Reports* 9, 16129.

Churko, J.M., Mantalas, G.L., Snyder, M.P., and Wu, J.C. 2013. Overview of high throughput sequencing technologies to elucidate molecular pathways in cardiovascular diseases. *Circulation Research* 112(12), 1613-1623.

Clementson, L.A., Richardson, A.J., Rochester, W.A., Oubelkheir, K., Liu, B., D'Sa, E.J., Gusmão, L.F.M., Ajani, P., Schroeder, T., and Ford, P.W. 2021. Effect of a once in 100-year flood on a subtropical coastal phytoplankton community. *Frontiers in Marine Science* 8, 163.

Culverhouse, P.F. 2007. Human and machine factors in algae monitoring performance. *Ecological Informatics* 2(4), 361-366.

Dai, C., Tan, Q., Lu, W., Liu, Y., and Guo, H. 2016. Identification of optimal water transfer schemes for restoration of a eutrophic lake: an integrated simulation-optimization method. *Ecological Engineering* 95, 409-421.

Dashkova, V., Malashenkov, D., Poulton, N., Vorobjev, I., and Barteneva, N.S. 2017. Imaging flow cytometry for phytoplankton analysis. *Methods* 112, 188-200.

de Jonge, V.N., Elliott, M., and Orive, E. 2002. Causes, historical development, effects and future challenges of a common environmental problem: eutrophication. *Hydrobiologia* 475-476, 1-19.

Dechesne, A., Musovic, S., Palomo, A., Diwan, V., and Smets, B.F. 2016. Underestimation of ammonia‐oxidizing bacteria abundance by amplification bias in *amoA*‐targeted qPCR. *Microbial Biotechnology* 9(4), 519-524.

Dembowska, E.A., Napiórkowski, P., Mieszczankin, T., and Józefowicz, S. 2015. Planktonic indices in the evaluation of the ecological status and the trophic state of the longest lake in Poland. *Ecological Indicators* 56, 15-22.

Dembowska, E.A., Mieszczankin, T., and Napiórkowski, P. 2018. Changes of the phytoplankton community as symptoms of deterioration of water quality in a shallow lake. *Environmental Monitoring and Assessment* 190(2), 95.

Dijkman, N.A., and Kromkamp, J.C. 2006. Phospholipid-derived fatty acids as chemotaxonomic markers for phytoplankton: application for inferring phytoplankton composition. *Marine Ecology Progress Series* 324, 113-125.

Ding, Y., Li, M., Pan, B., Zhao, G., and Gao, L. 2022. Disentangling the drivers of phytoplankton community composition in a heavily sediment-laden transcontinental river. *Journal of Environmental Management* 302, 113939.

Dodds, W.K., Jones, J.R., and Welch, E.B. 1998. Suggested classification of stream trophic state: distributions of temperate stream types by chlorophyll, total nitrogen, and phosphorus. *Water Research* 32(5), 1455-1462.

Dodds, W.K. 2006. Eutrophication and trophic state in rivers and streams. *Limnology and Oceanography* 51(1 Part 2), 671-680.

Dodds, W.K., Bouska, W.W., Eitzmann, J.L., Pilger, T.J., Pitts, K.L., Riley, A.J., Schloesser, J.T., and Thornbrugh, D.J. 2009. Eutrophication of US freshwaters: analysis of potential economic damages. *Environmental Science & Technology* 43(1), 12–19.

Dokulil, M.T., and Qian, K. 2021. Photosynthesis, carbon acquisition and primary productivity of phytoplankton: a review dedicated to Colin Reynolds. *Hydrobiologia* 848(1), 77-94.

dos Santos, A.L., Ribeiro, C.G., Ong, D., Garczarek, L., Shi, X.L., Nodder, S.D., Vaulot, D., and Gutiérrez-Rodríguez, A. (2021) Phytoplankton diversity and ecology through the lens of high throughput sequencing technologies in *Advances in Phytoplankton Ecology: Applications of Emerging Technologies*. Clementson, L.A., Eriksen, R.S. and Willis, A. (eds), pp. 353-413, Elsevier Inc., Cambridge, Massachusetts, U.S.A.

Duan, X., Zhang, C., Struewing, I., Li, X., Allen, J., and Lu, J. 2022. Cyanotoxin-encoding genes as powerful predictors of cyanotoxin production during harmful cyanobacterial blooms in an inland freshwater lake: evaluating a novel early-warning system. *Science of the Total Environment* 830, 154568.

Duarte, C.M., Conley, D.J., Carstensen, J., and Sánchez-Camacho, M. 2009. Return to *Neverland*: shifting baselines affect eutrophication restoration targets. *Estuaries and Coasts* 32(1), 29-36.

Dubelaar, G.B., and Jonker, R.R. 2000. Flow cytometry as a tool for the study of phytoplankton. *Scientia Marina* 64(2), 135-156.

Dugdale, R., Wilkerson, F., Parker, A.E., Marchi, A., and Taberski, K. 2012. River flow and ammonium discharge determine spring phytoplankton blooms in an urbanized estuary. *Estuarine, Coastal and Shelf Science* 115, 187-199.

Dunker, S. 2019. Hidden secrets behind dots: improved phytoplankton taxonomic resolution using high ‐ throughput imaging flow cytometry. *Cytometry Part A* 95(8), 854-868.

Dunker, S. 2020. Imaging flow cytometry for phylogenetic and morphologically based functional group clustering of a natural phytoplankton community over 1 year in an urban pond. *Cytometry Part A* 97(7), 727-736.

Eldridge, S.L.C., Driscoll, C., and Dreher, T.W. 2017. *Using High-Throughput DNA Sequencing, Genetic Fingerprinting, and Quantitative PCR as Tools for Monitoring Bloom-Forming and Toxigenic Cyanobacteria in Upper Klamath Lake, Oregon, 2013 and 2014 (Report 2017–5026).* U.S. Geological Survey, doi.org/10.3133/sir20175026.

Elser, J.J., and Hassett, R.P. 1994. A stoichiometric analysis of the zooplankton–phytoplankton interaction in marine and freshwater ecosystems. *Nature* 370(6486), 211-213.

Feng, F., Li, Y., Latimer, B., Zhang, C., Nair, S.S., and Hu, Z. 2021. Prediction of maximum algal productivity in membrane bioreactors with a light-dependent growth model. *Science of the Total Environment* 753, 141922.

Filazzola, A., Mahdiyan, O., Shuvo, A., Ewins, C., Moslenko, L., Sadid, T., Blagrave, K., Imrit, M.A., Gray, D.K., and Quinlan, R. 2020. A database of chlorophyll and water chemistry in freshwater lakes. *Scientific Data* 7, 310.

Filstrup, C.T., and Downing, J.A. 2017. Relationship of chlorophyll to phosphorus and nitrogen in nutrient-rich lakes. *Inland Waters* 7(4), 385-400.

Galazzo, G., van Best, N., Benedikter, B.J., Janssen, K., Bervoets, L., Driessen, C., Oomen, M., Lucchesi, M., Van Eijck, P.H., and Becker, H.E. 2020. How to count our microbes? The effect of different quantitative microbiome profiling approaches. *Frontiers in Cellular and Infection Microbiology* 10, 403.

Gallegos, C.L., Jordan, T.E., and Hedrick, S.S. 2010. Long-term dynamics of phytoplankton in the Rhode River, Maryland (USA). *Estuaries and Coasts* 33(2), 471-484.

Galloway, A.W., and Winder, M. 2015. Partitioning the relative importance of phylogeny and environmental conditions on phytoplankton fatty acids. *PLOS ONE* 10(6), e0130053.

Garmendia, M., Borja, Á., Franco, J., and Revilla, M. 2013. Phytoplankton composition indicators for the assessment of eutrophication in marine waters: present state and challenges within the European directives. *Marine Pollution Bulletin* 66(1-2), 7-16.

Gong, W., Hall, N., Paerl, H., and Marchetti, A. 2020. Phytoplankton composition in a eutrophic estuary: comparison of multiple taxonomic approaches and influence of environmental factors. *Environmental Microbiology* 22(11), 4718-4731.

González, E.J., and Roldán, G. (2019) Eutrophication and phytoplankton: some generalities from lakes and reservoirs of the Americas in *Microalgae—From Physiology to Application*. Vítová, M. (ed), pp. 27-45, IntechOpen, London, UK.

Graham, J.L., Jones, J.R., Jones, S.B., Downing, J.A., and Clevenger, T.E. 2004a. Environmental factors influencing microcystin distribution and concentration in the Midwestern United States. *Water Research* 38(20), 4395-4404.

Graham, J.L., Loftin, K.A., Meyer, M.T., and Ziegler, A.C. 2010. Cyanotoxin mixtures and taste-and-odor compounds in cyanobacterial blooms from the Midwestern United States. *Environmental Science & Technology* 44(19), 7361-7368.

Graham, J.L., Ziegler, A.C., Loving, B.L., and Loftin, K.A. 2012. *Fate and Transport of Cyanobacteria and Associated Toxins and Taste-and-Odor Compounds from Upstream Reservoir Releases in the Kansas River, Kansas, September and October 2011 (Report 2012–5129)*. U.S. Geological Survey, pubs.usgs.gov/sir/2012/5129/.

Graham, J.L., Dubrovsky, N.M., Foster, G.M., King, L.R., Loftin, K.A., Rosen, B.H., and Stelzer, E.A. 2020. Cyanotoxin occurrence in large rivers of the United States. *Inland Waters* 10(1), 109-117.

Graham, J.L., Rosen, B.H., Dubrovsky, N.M., Loftin, K.A., and Stelzer, E.A. 2021. *Phytoplankton Data for Samples Collected at Eleven Large River Sites throughout the United States, June through October 2018*. U.S. Geological Survey Data Release, doi.org/10.5066/P9N4Q9HG.

Graham, J.L., Rosen, B.H., Dubrovsky, N.M., Loftin, K.A., Stelzer, E.A., St. Amand, A.L., and Welk, R.J. 2022. *Phytoplankton data for samples collected at twelve large river sites throughout the United States, June through October 2019*. U.S. Geological Survey Data Release, doi.org/10.5066/P9CDG5QI.

Graham, J.M., Kent, A.D., Lauster, G., Yannarell, A., Graham, L., and Triplett, E.W. 2004b. Seasonal dynamics of phytoplankton and planktonic protozoan communities in a northern temperate humic lake: diversity in a dinoflagellate dominated system. *Microbial Ecology* 48(4), 528-540.

Gregor, J., and Maršálek, B. 2004. Freshwater phytoplankton quantification by chlorophyll a: a comparative study of *in vitro*, *in vivo* and *in situ* methods. *Water Research* 38(3), 517-522.

Guo, J., Zhang, C., Zheng, G., Xue, J., and Zhang, L. 2018. The establishment of season-specific eutrophication assessment standards for a water-supply reservoir located in Northeast China based on chlorophyll-a levels. *Ecological Indicators* 85, 11-20.

Hardenbicker, P., Rolinski, S., Weitere, M., and Fischer, H. 2014. Contrasting long‐term trends and shifts in phytoplankton dynamics in two large rivers. *International Review of Hydrobiology* 99(4), 287-299.

Hoang, H.T.T., Duong, T.T., Nguyen, K.T., Le, Q.T.P., Luu, M.T.N., Trinh, D.A., Le, A.H., Ho, C.T., Dang, K.D., and Némery, J. 2018. Impact of anthropogenic activities on water quality and plankton communities in the Day River (Red River Delta, Vietnam). *Environmental Monitoring and Assessment* 190(2), 67.

Hu, M., Ma, R., Cao, Z., Xiong, J., and Xue, K. 2021. Remote estimation of trophic state index for inland waters using Landsat-8 OLI imagery. *Remote Sensing* 13(10), 1988.

Hunt, R.J., and Matveev, V.F. 2005. The effects of nutrients and zooplankton community structure on phytoplankton growth in a subtropical Australian reservoir: an enclosure study. *Limnologica* 35(1-2), 90-101.

Huo, S., He, Z., Su, J., Xi, B., and Zhu, C. 2013. Using artificial neural network models for eutrophication prediction. *Procedia Environmental Sciences* 18, 310-316.

Jeffrey, S.W., and Vesk, M. (1997) Introduction to marine phytoplankton and their pigment signatures in *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*. Jeffrey, S.W., Mantoura, R.F.C. and Wright, S.W. (eds), pp. 33–84, UNESCO Publishing, Paris, France.

Jian, C., Luukkonen, P., Yki-Järvinen, H., Salonen, A., and Korpela, K. 2020. Quantitative PCR provides a simple and accessible method for quantitative microbiota profiling. *PLOS ONE* 15(1), e0227285.

Jones, R.C., and Redfield, G.W. 1984. Effects of urban stormwater runoff on reservoir phytoplankton: With 3 figures and 3 tables in the text. *SIL Proceedings, 1922-2010 (Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen)* 22(3), 1486-1492.

Khatri, N., and Tyagi, S. 2015. Influences of natural and anthropogenic factors on surface and groundwater quality in rural and urban areas. *Frontiers in Life Science* 8(1), 23-39.

Kiefer, D.A., and Cullen, J.J. 1991. Phytoplankton growth and light absorption as regulated by light, temperature, and nutrients. *Polar research* 10(1 Part 1), 163-172.

Kiersztyn, B., Siuda, W., and Chróst, R. 2002. Microbial ectoenzyme activity: useful parameters for characterizing the trophic conditions of lakes. *Polish Journal of Environmental Studies* 11(4), 367-374.

King, L.R., Rosen, B.H., Graham, J.L., Dubrovsky, N.M., Foster, G.M., Loftin, K.A., and Stelzer, E.A. 2020a. *Phytoplankton Data for Samples Collected at Eleven Large River Sites throughout the United States, June through September 2017*. U.S. Geological Survey Data Release, doi.org/10.5066/P9EYP85Z.

King, L.R., Rosen, B.H., Graham, J.L., Dubrovsky, N.M., Foster, G.M., Loftin, K.A., and Stelzer, E.A. 2020b. *Phytoplankton Tally Sheet, Including Photomicrographs, for Samples Collected at Eleven Large River Sites throughout the United States, June through September 2017*. U.S. Geological Survey, 10.5066/P9KKN921.

Klymus, K.E., Merkes, C.M., Allison, M.J., Goldberg, C.S., Helbing, C.C., Hunter, M.E., Jackson, C.A., Lance, R.F., Mangan, A.M., and Monroe, E.M. 2020. Reporting the limits of detection and quantification for environmental DNA assays. *Environmental DNA* 2(3), 271-282.

Knoll, L.B., Hagenbuch, E.J., Stevens, M.H., Vanni, M.J., Renwick, W.H., Denlinger, J.C., Hale, S.R., and González, M.J. 2015. Predicting eutrophication status in reservoirs at large spatial scales using landscape and morphometric variables. *Inland Waters* 5(3), 203-214.

Komárek, J. 2008. The cyanobacterial genus *Macrospermum*. *Fottea* 8(1), 79-86.

Komárek, J., Kaštovský, J., Mareš, J., and Johansen, J.R. 2014. Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. *Preslia* 86, 295–335.

Komárek, J. 2016. A polyphasic approach for the taxonomy of cyanobacteria: principles and applications. *European Journal of Phycology* 51(3), 346-353.

Kralik, P., and Ricchi, M. 2017. A basic guide to real time PCR in microbial diagnostics: definitions, parameters, and everything. *Frontiers in Microbiology* 8, 108.

Kulabhusan, P.K., and Campbell, K. 2021. Recent trends in the detection of freshwater cyanotoxins with a critical note on their occurrence in Asia. *Trends in Environmental Analytical Chemistry* 32, e00150.

Latasa, M., Scharek, R., Morán, X.A.G., Gutiérrez-Rodríguez, A., Emelianov, M., Salat, J., Vidal, M., and Estrada, M. 2021. Dynamics of phytoplankton groups in three contrasting situations of the open NW Mediterranean Sea revealed by pigment, microscopy, and flow cytometry analyses. *Progress in Oceanography* 201, 102737.

Lee, M., Won, N.-I., and Baek, S.H. 2020. Comparison of HPLC pigment analysis and microscopy in phytoplankton assessment in the Seomjin River estuary, Korea. *Sustainability* 12(4), 1675.

Lehman, P.W., Kurobe, T., Huynh, K., Lesmeister, S., and Teh, S.J. 2021. Covariance of phytoplankton, bacteria, and zooplankton communities within *Microcystis* blooms in San Francisco estuary. *Frontiers in Microbiology* 12, 1184.

Li, Y., Meng, J., Zhang, C., Ji, S., Kong, Q., Wang, R., and Liu, J. 2020. Bottom-up and top-down effects on phytoplankton communities in two freshwater lakes. *PLOS ONE* 15(4), e0231357.

Li, Y., Zhang, C., He, X., and Hu, Z. 2022. Solids retention time dependent, tunable diatom hierarchical micro/nanostructures and their effect on nutrient removal. *Water Research* 216, 118346.

Liang, Z., and Keeley, A. 2013. Filtration recovery of extracellular DNA from environmental water samples. *Environmental Science & Technology* 47(16), 9324-9331.

Lima, A., França, A., Muzny, C.A., Taylor, C.M., and Cerca, N. 2022. DNA extraction leads to bias in bacterial quantification by qPCR. *Applied Microbiology and Biotechnology* 106, 7993–8006.

Liu, L., Fu, M., Sun, K., Xu, Q., Xu, Z., Zhang, X., and Wang, Z. 2021. The distribution of phytoplankton size and major influencing factors in the surface waters near the northern end of the Antarctic Peninsula. *Acta Oceanologica Sinica* 40(6), 92-99.

Liu, Y., Xu, X., Wang, T., and Ni, J. 2019. Microscopic view of phytoplankton along the Yangtze River. *Science China Technological Sciences* 62(11), 1873-1884.

Lopes, O.F., Rocha, F.A., de Sousa, L.F., da Silva, D.M.L., Amorim, A.F., Gomes, R.L., da Silva Junior, A.L.S., and de Jesus, R.M. 2019. Influence of land use on trophic state indexes in northeast Brazilian river basins. *Environmental Monitoring and Assessment* 191(2), 1-14.

Lu, J., Struewing, I., Vereen, E., Kirby, A., Levy, K., Moe, C., and Ashbolt, N. 2016. Molecular detection of *Legionella* spp. and their associations with *Mycobacterium* spp., *Pseudomonas aeruginosa* and amoeba hosts in a drinking water distribution system. *Journal of Applied Microbiology* 120(2), 509-521.

Lu, J., Buse, H., Struewing, I., Zhao, A., Lytle, D., and Ashbolt, N. 2017. Annual variations and effects of temperature on *Legionella* spp. and other potential opportunistic pathogens in a bathroom. *Environmental Science and Pollution Research* 24(3), 2326-2336.

Lu, J.R., Struewing, I., Wymer, L., Tettenhorst, D.R., Shoemaker, J., and Allen, J. 2020. Use of qPCR and RT-qPCR for monitoring variations of microcystin producers and as an early warning system to predict toxin production in an Ohio inland lake. *Water Research* 170, 12.

Lueangthuwapranit, C., Sampantarak, U., and Wongsai, S. 2011. Distribution and abundance of phytoplankton: influence of salinity and turbidity gradients in the Na Thap River, Songkhla Province, Thailand. *Journal of Coastal Research* 27(3), 585-594.

Luo, W., Pflugmacher, S., Pröschold, T., Walz, N., and Krienitz, L. 2006. Genotype versus phenotype variability in *Chlorella* and *Micractinium* (Chlorophyta, Trebouxiophyceae). *Protist* 157(3), 315-333.

Malashenkov, D.V., Dashkova, V., Zhakupova, K., Vorobjev, I.A., and Barteneva, N.S. 2021. Comparative analysis of freshwater phytoplankton communities in two lakes of Burabay National Park using morphological and molecular approaches. *Scientific Reports* 11, 16130.

Malkassian, A., Nerini, D., van Dijk, M.A., Thyssen, M., Mante, C., and Gregori, G. 2011. Functional analysis and classification of phytoplankton based on data from an automated flow cytometer. *Cytometry Part A* 79(4), 263-275.

Markina, Z.V. 2019. Flow cytometry as a method to study marine unicellular algae: development, problems, and prospects. *Russian Journal of Marine Biology* 45(5), 333-340.

Marshall, H.G. 2009. Phytoplankton of the York river. *Journal of Coastal Research* 10057, 59-65.

Mineeva, N., and Makarova, O. 2018. Chlorophyll content as an indicator of the modern (2015–2016) trophic state of Volga River reservoirs. *Inland Water Biology* 11(3), 367-370.

Moorhouse, H., Read, D., McGowan, S., Wagner, M., Roberts, C., Armstrong, L., Nicholls, D., Wickham, H., Hutchins, M., and Bowes, M. 2018. Characterisation of a major phytoplankton bloom in the River Thames (UK) using flow cytometry and high performance liquid chromatography. *Science of the Total Environment* 624, 366-376.

Nagler, M., Insam, H., Pietramellara, G., and Ascher-Jenull, J. 2018. Extracellular DNA in natural environments: features, relevance and applications. *Applied Microbiology and Biotechnology* 102(15), 6343-6356.

Najwah, A.A., and Philip, G. 2021. The comparison of three environmental metrics for Cr, Pb, and Zn in the agricultural region of the mid-continent of USA. *Journal of Geoscience and Environment Protection* 9(4), 147-165.

Naselli-Flores, L., and Padisák, J. 2022. Ecosystem services provided by marine and freshwater phytoplankton (In press). *Hydrobiologia*, doi.org/10.1007/s10750-10022-04795-y.

Neverova-Dziopak, E., Kowalewski, Z., and Preisner, M. 2023. The universal trophic index: new methodological approach to eutrophication monitoring and control. *Aquatic Sciences* 85, 6.

Nielsen, K.M., Johnsen, P.J., Bensasson, D., and Daffonchio, D. 2007. Release and persistence of extracellular DNA in the environment. *Environmental Biosafety Research* 6(1-2), 37-53.

O'Neil, J.M., Davis, T.W., Burford, M.A., and Gobler, C.J. 2012. The rise of harmful cyanobacteria blooms: the potential roles of eutrophication and climate change. *Harmful Algae* 14, 313-334.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77(1), 118-125.

Omernik, J.M., and Griffith, G.E. 2014. Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. *Environmental Management* 54(6), 1249-1266.

Onyema, I. 2018. Water chemistry, microscopy and algal pigment concentration analyses of phytoplankton in the western and eastern parts of the Lagos lagoon. *Egyptian Academic Journal of Biological Sciences, H. Botany* 9(1), 75-85.

Opel, K.L., Chung, D., and McCord, B.R. 2010. A study of PCR inhibition mechanisms using real time PCR. *Journal of Forensic Sciences* 55(1), 25-33.

Otten, T.G., Crosswell, J.R., Mackey, S., and Dreher, T.W. 2015. Application of molecular tools for microbial source tracking and public health risk assessment of a *Microcystis* bloom traversing 300 km of the Klamath River. *Harmful Algae* 46, 71-81.

Pacheco, A.B.F., Guedes, I.A., and Azevedo, S.M. 2016. Is qPCR a reliable indicator of cyanotoxin risk in freshwater? *Toxins* 8(6), 172.

Paerl, H.W., Fulton, R.S., Moisander, P.H., and Dyble, J. 2001. Harmful freshwater algal blooms, with an emphasis on cyanobacteria. *The Scientific World Journal* 1, 76-113.

Palińska, K.A., and Surosz, W. 2008. Population of *Aphanizomenon* from the Gulf of Gdańsk (Southern Baltic Sea): differences in phenotypic and genotypic characteristics. *Hydrobiologia* 607, 163-173.

Pan, H., Li, A., Cui, Z., Ding, D., Qu, K., Zheng, Y., Lu, L., Jiang, T., and Jiang, T. 2020. A comparative study of phytoplankton community structure and biomass determined by HPLC-CHEMTAX and microscopic methods during summer and autumn in the central Bohai Sea, China. *Marine Pollution Bulletin* 155, 111172.

Parker, A.E., Dugdale, R.C., and Wilkerson, F.P. 2012. Elevated ammonium concentrations from wastewater discharge depress primary productivity in the Sacramento River and the Northern San Francisco Estuary. *Marine Pollution Bulletin* 64(3), 574-586.

Patel, B.H. (2011) Natural dyes in *Handbook of Textile and Industrial Dyeing (Volume 1): Principles, Processes and Types of Dyes*. Clark, M. (ed), pp. 395-424, Woodhead Publishing Limited, Philadelphia, Pennsylvania, U.S.A.

Pearson, L.A., D'Agostino, P.M., and Neilan, B.A. 2021. Recent developments in quantitative PCR for monitoring harmful marine microalgae. *Harmful Algae* 108, 102096.

Poretsky, R., Rodriguez-R, L.M., Luo, C., Tsementzi, D., and Konstantinidis, K.T. 2014. Strengths and limitations of 16S rRNA gene amplicon sequencing in revealing temporal microbial community dynamics. *PLOS ONE* 9(4), e93827.

Porter, T.M., and Hajibabaei, M. 2018. Scaling up: a guide to high‐throughput genomic approaches for biodiversity analysis. *Molecular Ecology* 27(2), 313-338.

Primpas, I., Tsirtsis, G., Karydis, M., and Kokkoris, G.D. 2010. Principal component analysis: development of a multivariate index for assessing eutrophication according to the European water framework directive. *Ecological Indicators* 10(2), 178-183.

Qin, B., Gao, G., Zhu, G., Zhang, Y., Song, Y., Tang, X., Xu, H., and Deng, J. 2013. Lake eutrophication and its ecosystem response. *Chinese Science Bulletin* 58(9), 961-970.

Raso, A., and Biassoni, R. (2014) Twenty years of qPCR: a mature technology? in *Quantitative Real-Time PCR: Methods and Protocols, Methods in Molecular Biology, Volume 1160*. pp. 1-3, Springer Science+Business Media, New York, U.S.A.

Read, D.S., Bowes, M.J., Newbold, L.K., and Whiteley, A.S. 2014. Weekly flow cytometric analysis of riverine phytoplankton to determine seasonal bloom dynamics. *Environmental Science: Processes & Impacts* 16(3), 594-603.

Reichwaldt, E.S., Wolf, I.D., and Stibor, H. 2004. The effect of different zooplankton grazing patterns resulting from diel vertical migration on phytoplankton growth and composition: a laboratory experiment. *Oecologia* 141(3), 411-419.

Reuter, J.A., Spacek, D.V., and Snyder, M.P. 2015. High-throughput sequencing technologies. *Molecular Cell* 58(4), 586-597.

Rhee, G.-u., and Gotham, I.J. 1981. The effect of environmental factors on phytoplankton growth: light and the interactions of light with nitrate limitation 1. *Limnology and Oceanography* 26(4), 649-659.

Richardson, K., and Jørgensen, B.B. (1996) Eutrophication: definition, history and effects in *Eutrophication in Coastal Marine Ecosystems, Volume 52*. Jørgensen, B.B. and Richardson, K. (eds), pp. 1-19, The American Geophysical Union, Washington, DC, U.S.A.

Rivas-Villar, D., Rouco, J., Carballeira, R., Penedo, M.G., and Novo, J. 2021. Fully automatic detection and classification of phytoplankton specimens in digital microscopy images. *Computer Methods and Programs in Biomedicine* 200, 105923.

Rosen, B.H., Davis, T.W., Gobler, C.J., Kramer, B.J., and Loftin, K.A. 2017. *Cyanobacteria of the 2016 Lake Okeechobee and Okeechobee Waterway Harmful Algal Bloom (Report 2017-1054).* U.S. Geological Survey, pubs.er.usgs.gov/publication/ofr20171054.

Rosen, B.H., Loftin, K.A., Graham, J.L., Stahlhut, K.N., Riley, J.M., Johnston, B.D., and Senegal, S. 2018. *Understanding the effect of salinity tolerance on cyanobacteria associated with a harmful algal bloom in Lake Okeechobee, Florida (Report 2018–5092).* U.S. Geological Survey, doi.org/10.3133/sir20185092.

Sagova-Mareckova, M., Boenigk, J., Bouchez, A., Cermakova, K., Chonova, T., Cordier, T., Eisendle, U., Elersek, T., Fazi, S., and Fleituch, T. 2021. Expanding ecological assessment by integrating microorganisms into routine freshwater biomonitoring. *Water Research* 191, 116767.

Sari, S.R., Tsushida, M., Sato, T., and Tominaga, M. 2022. Highly sensitive detection of phosphate using well-ordered crystalline cobalt oxide nanoparticles supported by multi-walled carbon nanotubes. *Materials Advances* 3(4), 2018-2025.

Savoy, P., Appling, A.P., Heffernan, J.B., Stets, E.G., Read, J.S., Harvey, J.W., and Bernhardt, E.S. 2019. Metabolic rhythms in flowing waters: An approach for classifying river productivity regimes. *Limnology and Oceanography* 64(5), 1835-1851.

Sayers, M.J., Fahnenstiel, G.L., Shuchman, R.A., and Bosse, K.R. 2021. A new method to estimate global freshwater phytoplankton carbon fixation using satellite remote sensing: initial results. *International Journal of Remote Sensing* 42(10), 3708-3730.

Schelske, C.L., Carrick, H.J., and Aldridge, F.J. 1995. Can wind-induced resuspension of meroplankton affect phytoplankton dynamics? *Journal of the North American Benthological Society* 14(4), 616-630.

Schweitzer-Natan, O., Ofek-Lalzar, M., Sher, D., and Sukenik, A. 2019. Particle-associated microbial community in a subtropical lake during thermal mixing and phytoplankton succession. *Frontiers in Microbiology* 10, 2142.

Silva, T.F., Vinçon-Leite, B., Lemaire, B.J., Petrucci, G., Giani, A., Figueredo, C.C., and Nascimento, N.d.O. 2019. Impact of urban stormwater runoff on cyanobacteria dynamics in a tropical urban lake. *Water* 11(5), 946.

Smith, C.J., and Osborn, A.M. 2009. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiology Ecology* 67(1), 6-20.

Smith, R.A., Schwarz, G.E., and Alexander, R.B. 1997. Regional interpretation of water‐quality monitoring data. *Water Resources Research* 33(12), 2781-2798.

Smith, V.H., Joye, S.B., and Howarth, R.W. 2006. Eutrophication of freshwater and marine ecosystems. *Limnology and Oceanography* 51(1 Part 2), 351-355.

Srichandan, S., Baliarsingh, S.K., Lotliker, A.A., Prakash, S., Samanta, A., and Sahu, K.C. 2020. A baseline investigation of phytoplankton pigment composition in contrasting coastal ecosystems of north-western Bay of Bengal. *Marine Pollution Bulletin* 160, 111708.

Steffen, M.M., Davis, T.W., McKay, R.M.L., Bullerjahn, G.S., Krausfeldt, L.E., Stough, J.M., Neitzey, M.L., Gilbert, N.E., Boyer, G.L., and Johengen, T.H. 2017. Ecophysiological examination of the Lake Erie *Microcystis* bloom in 2014: linkages between biology and the water supply shutdown of Toledo, OH. *Environmental Science & Technology* 51(12), 6745-6755.

Stets, E.G., Kelly, V.J., and Crawford, C.G. 2015. Regional and temporal differences in nitrate trends discerned from long‐term water quality monitoring data. *Journal of the American Water Resources Association* 51(5), 1394-1407.

Stets, E.G., Sprague, L.A., Oelsner, G.P., Johnson, H.M., Murphy, J.C., Ryberg, K., Vecchia, A.V., Zuellig, R.E., Falcone, J.A., and Riskin, M.L. 2020. Landscape drivers of dynamic change in water quality of US rivers. *Environmental Science & Technology* 54(7), 4336-4343.

Stomp, M., Huisman, J., Mittelbach, G.G., Litchman, E., and Klausmeier, C.A. 2011. Large‐scale biodiversity patterns in freshwater phytoplankton. *Ecology* 92(11), 2096-2107.

Stoyneva-Gärtner, M.P., Morana, C., Borges, A.V., Okello, W., Bouillon, S., Deirmendjian, L., Lambert, T., Roland, F., Nankabirwa, A., and Nabafu, E. 2020. Diversity and ecology of phytoplankton in Lake Edward (East Africa): present status and long-term changes. *Journal of Great Lakes Research* 46(4), 741-751.

Striebel, M., Spörl, G., and Stibor, H. 2008. Light‐induced changes of plankton growth and stoichiometry: experiments with natural phytoplankton communities. *Limnology and Oceanography* 53(2), 513-522.

Tavakoli, Y., Mohammadipanah, F., Te, S.H., You, L., and Gin, K.Y.-H. 2021. Biodiversity, phylogeny and toxin production profile of cyanobacterial strains isolated from lake Latyan in Iran. *Harmful Algae* 106, 102054.

Tian, Y., Gao, L., Deng, J., and Li, M. 2020. Characterization of phytoplankton community in a river ecosystem using pigment composition: a feasibility study. *Environmental Science and Pollution Research* 27(34), 42210-42220.

U.S. EPA 2022. *Ecoregions*. U.S. EPA, epa.gov/eco-research/ecoregions.

USGS 2022a. *Tracking Water Quality in U.S. Streams and Rivers: USGS National Water Quality Network Data, Water-Quality Loads, and Trends*. U.S. Geological Survey, nrtwq.usgs.gov/nwqn/#/.

USGS 2022b. *National Water Information System: USGS Water Data for the Nation*. U.S. Geological Survey, doi.org/10.5066/F7P55KJN.

Verasztó, C., Kiss, K.T., Sipkay, C., Gimesi, L., Vadadi Fülöp, C., Türei, D., and Hufnagel, L. 2010. Long-term dynamic patterns and diversity of phytoplankton communities in a large eutrophic river (the case of River Danube, Hungary). *Applied Ecology and Environmental Research* 8(4), 329-349.

Vuorio, K., Mäki, A., Salmi, P., Aalto, S.L., and Tiirola, M. 2020. Consistency of targeted metatranscriptomics and morphological characterization of phytoplankton communities. *Frontiers in Microbiology* 11, 96.

Wang, C., Liu, Y., Zhan, Q., Yang, W., and Wu, N. 2018. Global trends in phytoplankton research of river ecosystems during 1991–2016: a bibliometric analysis. *Fundamental and Applied Limnology* 191(1), 25-36.

Webster, I. 1990. Effect of wind on the distribution of phytoplankton cells in lakes. *Limnology and Oceanography* 35(5), 989-1001.

Wilcox, T.M., McKelvey, K.S., Young, M.K., Engkjer, C., Lance, R.F., Lahr, A., Eby, L.A., and Schwartz, M.K. 2020. Parallel, targeted analysis of environmental samples via high-throughput quantitative PCR. *Environmental DNA* 2(4), 544-553.

Wilhelm, C., Becker, A., Toepel, J., Vieler, A., and Rautenberger, R. 2004. Photophysiology and primary production of phytoplankton in freshwater. *Physiologia Plantarum* 120(3), 347-357.

Winder, M., and Sommer, U. 2012. Phytoplankton response to a changing climate. *Hydrobiologia* 698(1), 5-16.

Xia, J., and Zeng, J. 2021. Environmental factor assisted chlorophyll-*a* prediction and water quality eutrophication grade classification: a comparative analysis of multiple hybrid models based on a SVM. *Environmental Science: Water Research & Technology* 7(6), 1040-1049.

Xiao, X., Sogge, H., Lagesen, K., Tooming-Klunderud, A., Jakobsen, K.S., and Rohrlack, T. 2014. Use of high throughput sequencing and light microscopy show contrasting results in a study of phytoplankton occurrence in a freshwater environment. *PLOS ONE* 9(8), e106510.

Xu, F.-L., Tao, S., Dawson, R., and Li, B.-G. 2001. A GIS-based method of lake eutrophication assessment. *Ecological Modelling* 144(2-3), 231-244.

Xu, H., Paerl, H.W., Qin, B., Zhu, G., and Gaoa, G. 2010. Nitrogen and phosphorus inputs control phytoplankton growth in eutrophic Lake Taihu, China. *Limnology and Oceanography* 55(1), 420-432.

Yang, B., Jiang, Y.-J., He, W., Liu, W.-X., Kong, X.-Z., Jørgensen, S.E., and Xu, F.-L. 2016. The tempo-spatial variations of phytoplankton diversities and their correlation with trophic state levels in a large eutrophic Chinese lake. *Ecological Indicators* 66, 153-162.

Yang, J., Strokal, M., Kroeze, C., Chen, X., Bai, Z., Li, H., Wu, Y., and Ma, L. 2021a. Seasonal River Export of Nitrogen to Guanting and Baiyangdian Lakes in the Hai He Basin. *Journal of Geophysical Research: Biogeosciences* 126(10), e2020JG005689.

Yang, J., Lv, J., Liu, Q., Nan, F., Li, B., Xie, S., and Feng, J. 2021b. Seasonal and spatial patterns of eukaryotic phytoplankton communities in an urban river based on marker gene. *Scientific Reports* 11, 23147.

Yao, J., Wang, G., Xue, B., Wang, P., Hao, F., Xie, G., and Peng, Y. 2019. Assessment of lake eutrophication using a novel multidimensional similarity cloud model. *Journal of Environmental Management* 248, 109259.

Yoshiyama, K., and Sharp, J.H. 2006. Phytoplankton response to nutrient enrichment in an urbanized estuary: apparent inhibition of primary production by overeutrophication. *Limnology and Oceanography* 51(1 Part 2), 424-434.

Yu, X., Yang, J.R., Chen, J., Isabwe, A., and Yang, J. 2021. On the use of chemotaxonomy, a phytoplankton identification and quantification method based on pigment for quick surveys of subtropical reservoirs. *Environmental Science and Pollution Research* 28(3), 3544-3555.

Zhang, C., Ian, S., Mistry, J.H., Wahman, D., Pressman, J., and Lu, J. 2021a. *Legionella* and other opportunistic pathogens in full-scale chloraminated municipal drinking water distribution systems. *Water Research* 205, 117571.

Zhang, Y., Li, M., Dong, J., Yang, H., Van Zwieten, L., Lu, H., Alshameri, A., Zhan, Z., Chen, X., and Jiang, X. 2021b. A critical review of methods for analyzing freshwater eutrophication. *Water* 13(2), 225.

Zhang, Y., Gao, W., Li, Y., Jiang, Y., Chen, X., Yao, Y., Messyasz, B., Yin, K., He, W., and Chen, Y. 2021c. Characteristics of the phytoplankton community structure and water quality evaluation in autumn in the Huaihe River (China). *International Journal of Environmental Research and Public Health* 18(22), 12092.

Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S.G., and Alvarez-Cohen, L. 2015. High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. *mBio* 6(1), e02288-14.

Znachor, P., Nedoma, J., Hejzlar, J., Seďa, J., Komárková, J., Kolář, V., Mrkvička, T., and Boukal, D.S. 2020. Changing environmental conditions underpin long-term patterns of phytoplankton in a freshwater reservoir. *Science of the Total Environment* 710, 135626.

Zuellig, R.E., Graham, J.L., Stelzer, E.A., Loftin, K.A., and Rosen, B.H. 2021. *Cyanobacteria, Cyanotoxin Synthetase Gene, and Cyanotoxin Occurrence among Selected Large River Sites of the Conterminous United States, 2017–18 (Report 2021–5121)*. U.S. Geological Survey, doi.org/10.3133/sir20215121.

Zupančič, M., Kogovšek, P., Šter, T., Remec Rekar, Š., Cerasino, L., Baebler, Š., Krivograd Klemenčič, A., and Eleršek, T. 2021. Potentially toxic planktic and benthic cyanobacteria in Slovenian freshwater bodies: detection by quantitative PCR. *Toxins* 13(2), 133.

Table 1. The twelve large, freshwater rivers across the United States

River (region)	Sampling site (USGS station number)	Drainage area (km^2)[a]	Water temperature (°C)[c]	Flow rate ($m^3 \cdot s^{-1}$)[c]	SS ($mg \cdot L^{-1}$)[e]	TN ($\mu g\ N \cdot L^{-1}$)[e]	TP ($\mu g\ P \cdot L^{-1}$)[e]
Sacramento (W)	Freeport, CA (11447650)	61,445[b]	15.2 | 16.0 | 15.3	1,128 | 493 | 897	35.6 | 34.5 | 33.4	318 | 317 | 315	61 | 60 | 59
Willamette (W)	Portland, OR (14211720)	29,008	13.0 | 13.5 | 13.5	1,243 | 792 | 665	14.5 | 14.6 | 14.6	767 | 778 | 790	62 | 62 | 62
Trinity (M)	Dallas, TX (08057410)	16,260	22.6 | 22.1 | 26.3	54 | 66 | 206	85.1 | 80.9 | 76.8	8,040 | 7,800 | 7,570	833 | 777 | 722
Kansas (M)	De Soto, KS (06892350)	154,767	15.9 | 17.4 | 14.5	226 | 83 | 619	284 | 266 | 249	1,880 | 1,800 | 1,730	466 | 459 | 452
Missouri (M)	Hermann, MO (06934500)	1,353,269	16.7 | 15.8 | 13.6	2,589 | 2,284 | 5,505	402 | 401 | 399	2,710 | 2,790 | 2,880	425 | 425 | 425
Ohio (M)	Cannelton, IN (03303280)	251,229	NA | NA | NA	3,769 | 5,004 | 5,870	80.4 | 81.0 | 81.7	1,640 | 1,640 | 1,640	171 | 172 | 174
Mississippi (M)	Hastings, MN (05331580)	96,089	NA | NA | NA	708 | 756 | 1,053	40.6 | 38.4 | 36.4	4,290 | 4,380 | 4,480	112 | 109 | 106
Chattahoochee (E)	Whitesburg, GA (02338000)	6,294	NA | NA | NA	70 | 107 | 172	74.5 | 73.3 | 72.0	2,710 | 2,750 | 2,800	89 | 90 | 90
Susquehanna (E)	Conowingo, MD (01578310)	70,189	NA | NA | NA	1,122 | 1,561 | 1,757	17.1 | 17.0 | 17.0	1,470 | 1,460 | 1,450	50 | 49 | 48
Delaware (E)	Trenton, NJ (01463500)	17,560	13.6 | 13.3 | 13.3	287 | 406 | 567	17.4 | 17.9 | 18.4	1,210 | 1,220 | 1,230	79 | 81 | 84
Hudson (E)	Poughkeepsie, NY (01372043)	30,303	NA | NA | NA[d]	692 | 630 | NA[d]	NA | NA | NA	942 | 928 | 859	93 | 88 | 76
Connecticut (E)	Thompsonville, CT (01184000)	25,019	NA | NA | NA	490 | 506 | 675	22.6 | 22.5 | 22.9	569 | 567 | 564	47 | 47 | 46

W: Western region. **M**: Midcontinent region. **E**: Eastern region. **USGS**: The United States Geological Survey. **CA**: California. **OR**: Oregon. **TX**: Texas. **KS**: Kansas. **MO**: Missouri. **IN**: Indiana. **MN**: Minnesota. **GA**: Georgia. **MD**: Maryland. **NJ**: New Jersey. **NY**:

New York. **CT**: Connecticut. **SS**: Suspended solids. **TN**: Total nitrogen. **TP**: Total phosphorus. Except for the case of the sampling site in Poughkeepsie, NY, the concentrations of SS, TN, and TP are annual flow-normalized concentrations [computing assay: the Weighted Regressions on Time, Discharge, and Season (WRTDS) generalized flow-normalization method]. For the case of the sampling site in Poughkeepsie, NY, we calculated the concentrations of SS, TN, and TP manually using the discrete water-quality results. **NA**: Not available. [a]Source: https://nrtwq.usgs.gov/nwqn/#/NET except for the sampling site in Freeport, CA. [b]Source: Graham et al. (2020). [c]Source: https://nwis.waterdata.usgs.gov/nwis except for the sampling site in Poughkeepsie, NY. [d and e]Source: https://nrtwq.usgs.gov/nwqn/#/NET. [c, d, and e]In each cell, the values for 2017, 2018, and 2019 are listed from left to right and pipe delimited. The "2017," "2018," and "2019" are water years. The USGS defines a water year to be a 12-month period from October 1st in any given calendar year through September 30th in the following calendar year. The water year is designated by the calendar year in which it ends. [c, d, and e]Water quality parameters at the sampling sites.

Table 2. The linear correlations between qPCR- and microscope-based phytoplankton abundance at the sampling sites in the twelve large, freshwater rivers

	Bacill.	**Cyanobacteria**	**Dinoflagellates**	***Dolichospermum***	***Microcystis***	***Planktothrix***	***Synechococcus***	**Chlorophyta**
R^2; *p*	0.452; < 0.001*	0.359; < 0.001*	0.602; < 0.001*	0.836; < 0.001*	0.901; < 0.001*	0.963; < 0.001*	0.107; 0.005*	0.224; < 0.001*
B*; *p	0.989; < 0.001*	0.427; < 0.001*	1.049; < 0.001*	1.509; < 0.001*	1.337; < 0.001*	1.309; < 0.001*	0.393; 0.005*	0.430; < 0.001*
C*; *p	1.388; 0.004*	5.475; < 0.001*	1.759; < 0.001*	0.383; 0.217	0.107; 0.480	0.040; 0.549	4.832; < 0.001*	4.780; < 0.001*

n	163	141	60	62	55	115	64	155

Statistical analysis: SPSS. Test: Simple linear regression. Sampling time: Early summer to late fall in 2017, 2018, and 2019. Independent variable: Microscope-based phytoplankton abundance [unit: $\log_{10}(NU \cdot L^{-1})$] at the sampling sites. Dependent variable: qPCR-based phytoplankton abundance [unit: $\log_{10}(GCN \cdot L^{-1})$] at the sampling sites. One sampling site was selected from each river. Before the test, data points meeting the following two criteria were deleted (pairwise deletion): 1) For a target phytoplankton group at any sampling event, the qPCR-based phytoplankton abundance is 0 $\log_{10}(GCN \cdot L^{-1})$, while the microscope-based abundance is higher than 0 $\log_{10}(NU \cdot L^{-1})$; and 2) For a target phytoplankton group at any sampling event, the microscope-based phytoplankton abundance is 0 $\log_{10}(NU \cdot L^{-1})$, while the qPCR-based abundance is higher than 0 $\log_{10}(GCN \cdot L^{-1})$. For a target phytoplankton group at a sampling event, if the microscope-based abundance is 0 $\log_{10}(NU \cdot L^{-1})$ and the qPCR-based abundance is 0 $\log_{10}(GCN \cdot L^{-1})$, the data points were both kept. ***R^2*; *p***: The adjusted *R*-squared and *p*-value for the overall regression model, respectively. ***B*; *p***: The unstandardized coefficient and *p*-value for the independent variable, respectively. ***C*; *p***: The constant and the associated *p*-value, respectively. ***n***: Sample size. *: Significance (*p*-value < 0.05). **GCN**: Gene or genome copy number. **NU**: Natural units (cells, colonies, or filaments). **Bacill.**: Bacillariophyta (division). **Cyanobacteria**: Division. **Dinoflagellates**: Class. **Chlorophyta**: Specific to three classes in the division Chlorophyta (Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae). ***Dolichospermum***, ***Microcystis***, ***Planktothrix***, and ***Synechococcus***: Genera.

Figures

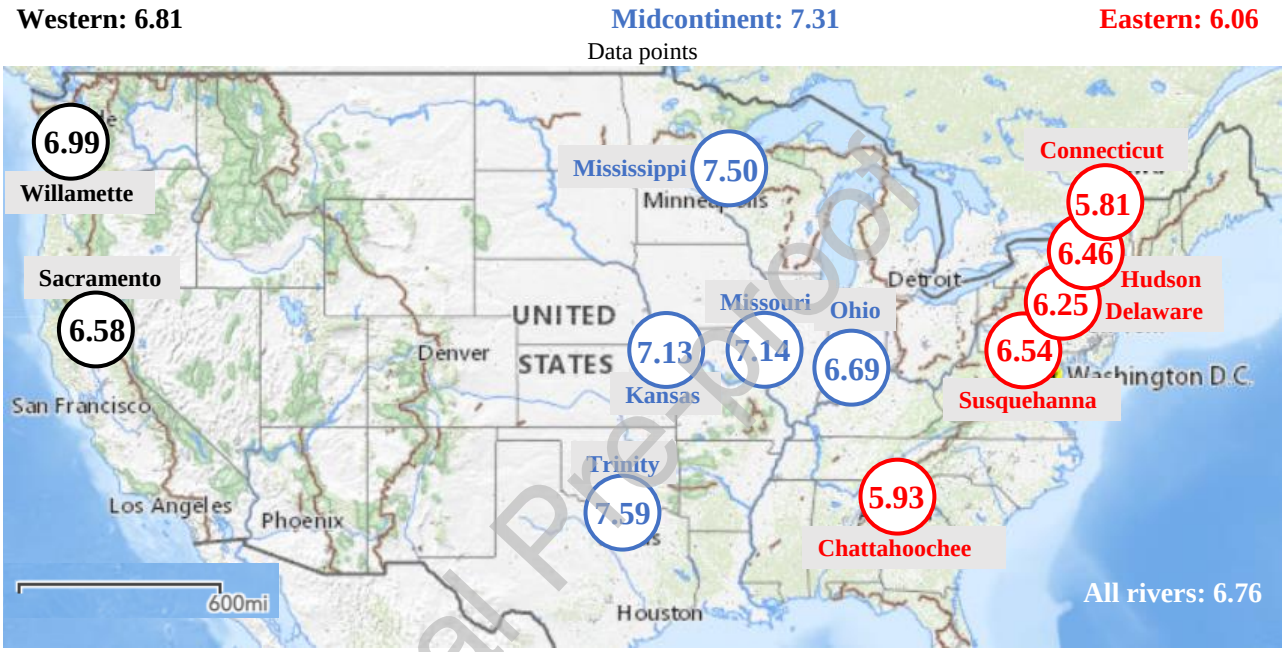


Figure 1. Concentrations of phytoplankton at the sampling sites (approximate locations) in twelve large, freshwater rivers across the United States. Source of the map: apps.nationalmap.gov/viewer/. One sampling site was selected from each river. The number within each circle: The geometric mean [unit: $\log_{10}(GCN \cdot L^{-1})$] of the qPCR-based concentration of four major phytoplankton groups (Bacillariophyta, Cyanobacteria, Dinoflagellates, and Chlorophyta) at each sampling site from 2017 to 2019. The geometric mean [unit: $\log_{10}(GCN \cdot L^{-1})$] of the qPCR-based concentration of the four major phytoplankton groups for each geographic region of the United States (western, midcontinent, and eastern) from 2017 to 2019 is on the top of the map. The geometric mean [unit:

$log_{10}(GCN \cdot L^{-1})$] of the qPCR-based concentration of the four major phytoplankton groups for all the sampling sites from 2017 to 2019 is on the lower right corner of the map. **GCN**: Gene or genome copy number. **Sacramento**: The Sacramento River (sampling site: Freeport, California). **Willamette**: The Willamette River (Sampling site: Portland, Oregon). **Trinity**: The Trinity River (sampling site: Dallas, Texas). **Kansas**: The Kansas River (sampling site: De Soto, Kansas). **Missouri**: The Missouri River (sampling site: Hermann, Missouri). **Ohio**: The Ohio River (sampling site: Cannelton Locks and Dam close to Cannelton, Indiana). **Mississippi**: The Mississippi River (sampling site: Hastings, Minnesota). **Chattahoochee**: The Chattahoochee River (sampling site: Whitesburg, Georgia). **Susquehanna**: The Susquehanna River (sampling site: Conowingo, Maryland). **Delaware**: The Delaware River (Trenton, New Jersey). **Hudson**: The Hudson River (sampling site: Poughkeepsie, New York). **Connecticut**: The Connecticut River (sampling site: Thompsonville, Connecticut).

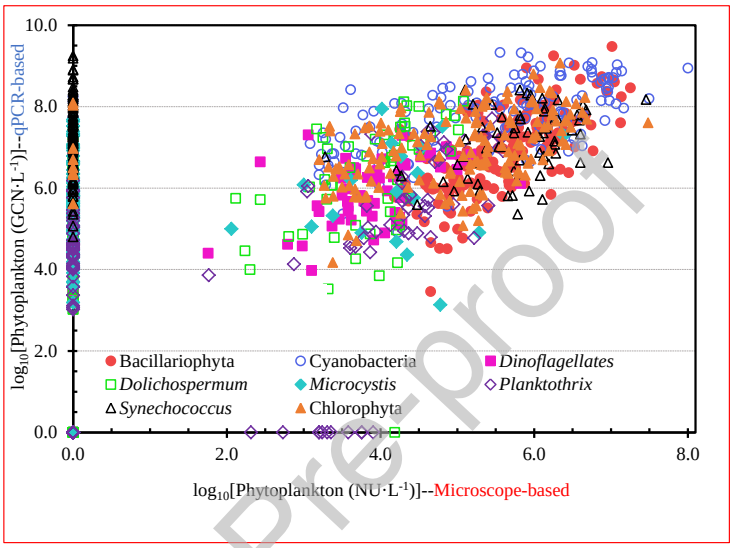
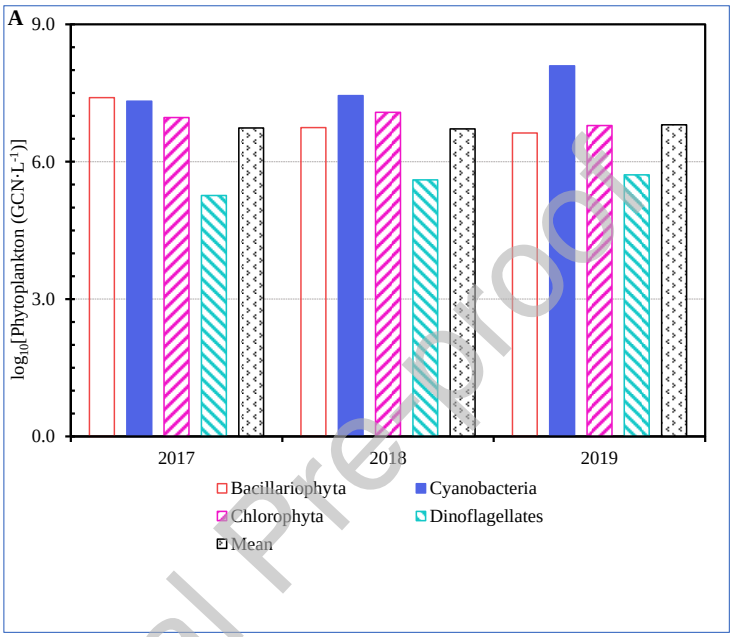


Figure 2. A comparison of qPCR-based and microscope-based phytoplankton abundance at the sampling sites in the twelve large, freshwater rivers. One sampling site was selected from each river. All data points are included in the plot. **Bacillariophyta** and **Cyanobacteria**: Divisions. **Dinoflagellates**: Class. **Chlorophyta**: Covers three classes within the division Chlorophyta (Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae). ***Dolichospermum***, ***Microcystis***, ***Planktothrix***, and ***Synechococcus***: Genera. Sampling time: Early summer to late fall in 2017, 2018, and 2019. **GCN**: Gene or genome copy number. **NU**: Natural units (cells, colonies, or filaments).

Figures



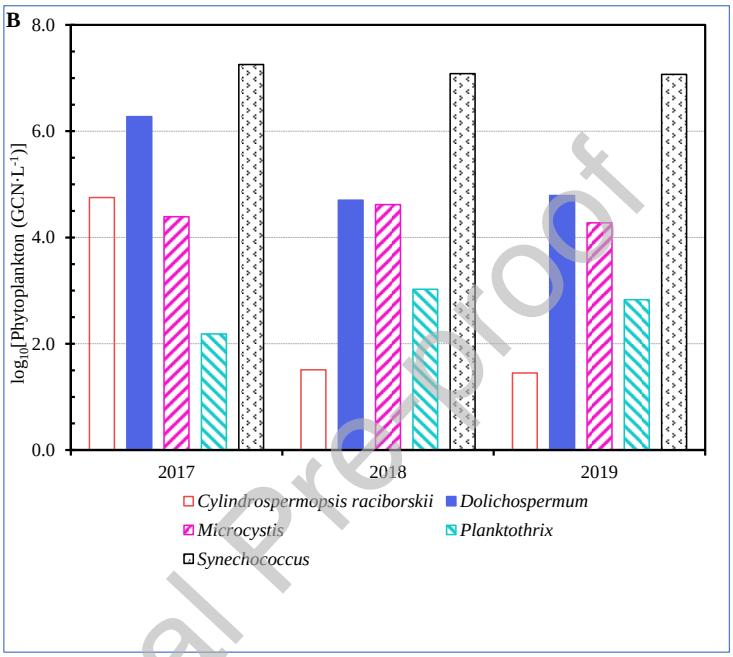
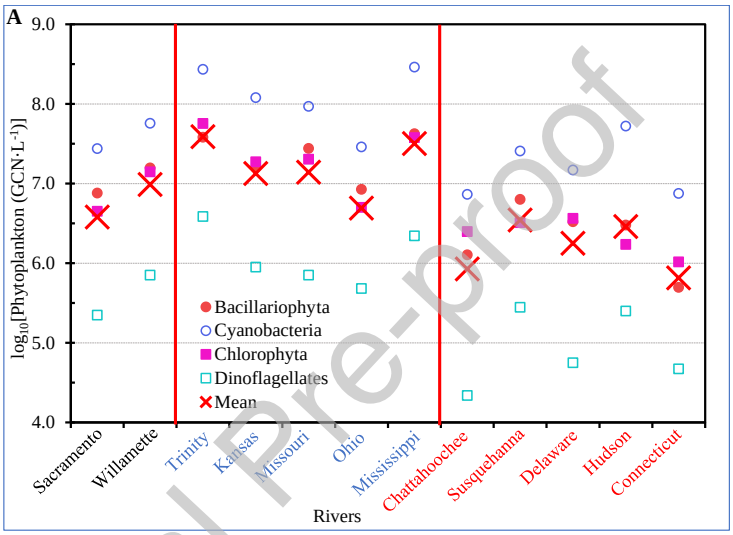


Figure 3. The geometric mean abundance (qPCR-based) of phytoplankton in 2017, 2018, and 2019 (sampling season: early summer to late fall) for all the sampling sites in the twelve rivers. One sampling site was selected from each river. **GCN**: Gene or genome copy number. (**A**) Four major phytoplankton groups. **Bacillariophyta** and **Cyanobacteria**: Divisions. **Chlorophyta**: Specific to three classes within the division Chlorophyta (Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae). **Dinoflagellates**: Class. **Mean**: The geometric mean for the four major phytoplankton groups. (**B**) Four phytoplankton genera (***Dolichospermum***, ***Microcystis***,

Planktothrix, and ***Synechococcus***) and a species (***Cylindrospermopsis raciborskii***) in Cyanobacteria.



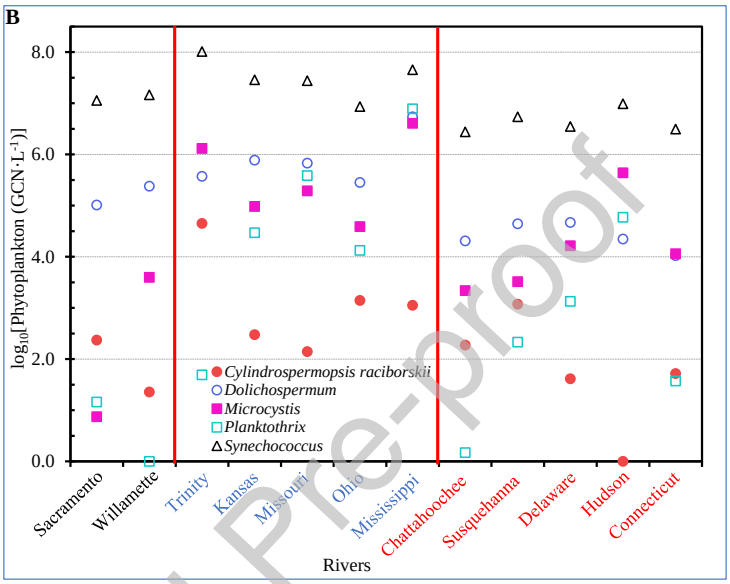


Figure 4. The geometric mean abundance (qPCR-based) of phytoplankton at the sampling sites in the twelve rivers (2017 to 2019, sampling season: early summer to late fall). One sampling site was selected from each river. **GCN**: Gene or genome copy number. (**A**) Major phytoplankton groups. **Bacillariophyta** and **Cyanobacteria**: Divisions. **Chlorophyta**: Specific to three classes within the division Chlorophyta (Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae). **Dinoflagellates**: Class. **Mean**: The geometric mean for the four major phytoplankton groups. (**B**) Four genera (***Dolichospermum***, ***Microcystis***, ***Planktothrix***, and ***Synechococcus***) and a species (***Cylindrospermopsis raciborskii***) in Cyanobacteria.

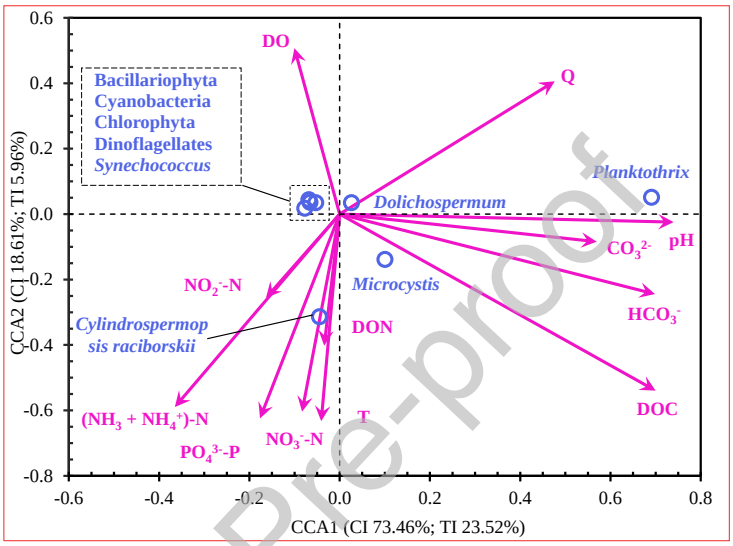
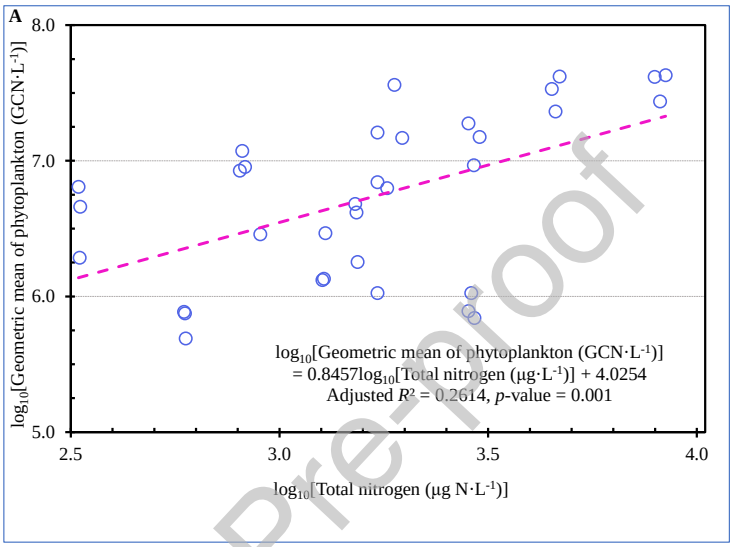


Figure 5. A constrained Canonical correspondence analysis (CCA) map (generated with XLSTAT) showing the linear associations among the qPCR-based abundance of eight phytoplankton groups and twelve critical water quality physicochemical parameters all measured at the sampling sites in the twelve rivers. One sampling site was selected from each river. The length of an arrow indicates the significance of the corresponding environmental variable, the direction of an arrow indicates the correlation between the corresponding environmental variable and an axis or a phytoplankton group, the angle between two arrows indicates the correlation between the two corresponding environmental variables, and the specific location of a phytoplankton group relative to an arrow indicates the environmental preference of the group (Palmer, 1993; Teer Braak, 1986). **Bacillariophyta** and **Cyanobacteria**: Divisions.

Chlorophyta: Specific to three classes within the division Chlorophyta (Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae). **Dinoflagellates**: Class. ***Dolichospermum***, ***Microcystis***, ***Planktothrix***, and ***Synechococcus***: Genera. ***Cylindrospermopsis raciborskii***: Species. Unit for phytoplankton abundance: $\log_{10}(GCN \cdot L^{-1})$. **GCN**: Gene or genome copy number. **T**: Water temperature (°C). **Q**: Discharge ($m^3 \cdot s^{-1}$). **DO**: Dissolved oxygen ($mg \cdot L^{-1}$). **DON**: Dissolved organic nitrogen (mg $N \cdot L^{-1}$). **(NH_3 + NH_4^+)-N**: Ammonia and ammonium nitrogen (mg $N \cdot L^{-1}$). **NO_2^--N**: Nitrite nitrogen (mg $N \cdot L^{-1}$). **NO_3^--N**: Nitrate nitrogen (mg $N \cdot L^{-1}$). **PO_4^{3-}-P**: Orthophosphate phosphorus (mg $P \cdot L^{-1}$). **DOC**: Dissolved organic carbon ($mg \cdot L^{-1}$). **CO_3^{2-}**: Dissolved carbonate ion ($mg \cdot L^{-1}$). **HCO_3^-**: Dissolved bicarbonate ion ($mg \cdot L^{-1}$). **CI**: Constrained inertia. **TI**: Total inertia.

Figures



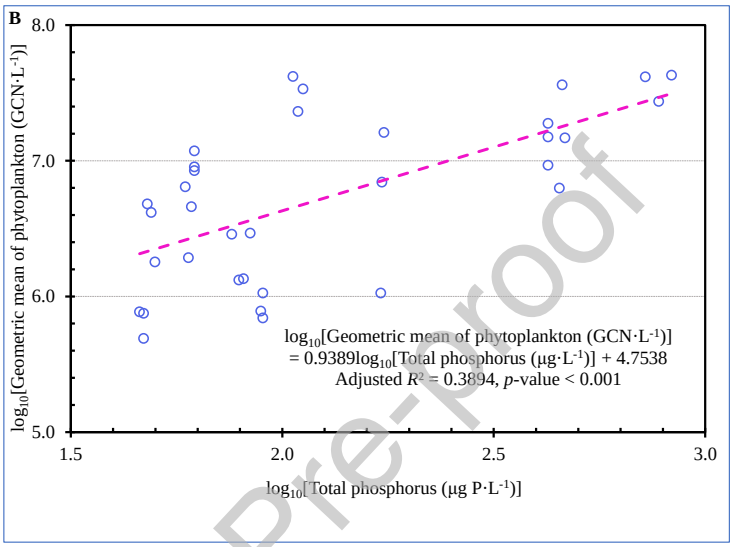


Figure 6. The linear correlations between the geometric mean abundance (qPCR-based) of four major phytoplankton groups (Bacillariophyta, Cyanobacteria, Chlorophyta, and Dinoflagellates) and (**A**) total nitrogen and (**B**) total phosphorus concentrations. The dashed lines: Simple linear regression lines. **Bacillariophyta** and **Cyanobacteria**: Divisions. **Chlorophyta**: Specific to three classes within the division Chlorophyta (Chlorophyceae, Trebouxiophyceae, and Klebsormidiophyceae). **Dinoflagellates**: Class. Adjusted *R*-squared's and *p*-values: SPSS-generated.