Published in final edited form as:
Sci Total Environ. 2022 July 15; 830: 154568. doi:10.1016/j.scitotenv.2022.154568.

Cyanotoxin-encoding genes as powerful predictors of cyanotoxin production during harmful cyanobacterial blooms in an inland freshwater lake: Evaluating a novel early-warning system

Xiaodi Duan[1], **Chiqian Zhang**[1], **Ian Struewing**[2], **Xiang Li**[3], **Joel Allen**[2], **Jingrang Lu**[4]

[1]Pegasus Technical Services, Inc., Cincinnati, OH 45219, USA.

[2]Office of Research and Development, United States Environmental Protection Agency, Cincinnati, OH 45268, USA.

[3]Oak Ridge Institute for Science and Education at the United States Environmental Protection Agency, Cincinnati, OH 45268, USA.

[4]Office of Research and Development, United States Environmental Protection Agency, Cincinnati, OH 45268, USA.

Abstract

Freshwater harmful cyanobacterial blooms (HCBs) potentially produce excessive cyanotoxins, mainly microcystins (MCs), significantly threatening aquatic ecosystems and public health. Accurately predicting HCBs is thus essential to developing effective HCB mitigation and prevention strategies. We previously developed a novel early-warning system that uses cyanotoxin-encoding genes to predict cyanotoxin production in Harsha Lake, Ohio, USA, in 2015. In this study, we evaluated the efficacy of the early-warning system in forecasting the 2016 HCB in the same lake. We also examined potential HCB drivers and cyanobacterial community composition. Our results revealed that the cyanobacterial community was stable at the phylum level but changed dynamically at the genus level over time. Microcystis and Planktothrix were the major MC-producing genera that thrived in June and July and produced high concentrations of MCs (peak level 10.22 μg·L^{-1}). The abundances of the MC-encoding gene cluster mcy and its transcript levels significantly correlated with total MC concentrations (before the MC concentrations peaked) and accurately predicted MC production as revealed by logistic equations. When the Microcystis-specific gene mcyG reached approximately 1.5×10^3 copies·mL^{-1} or when its transcript level reached approximately 2.4 copies·mL^{-1}, total MC level exceeded 0.3 μg L^{-1} (a health advisory limit) approximately one week later (weekly sampling scheme). This study

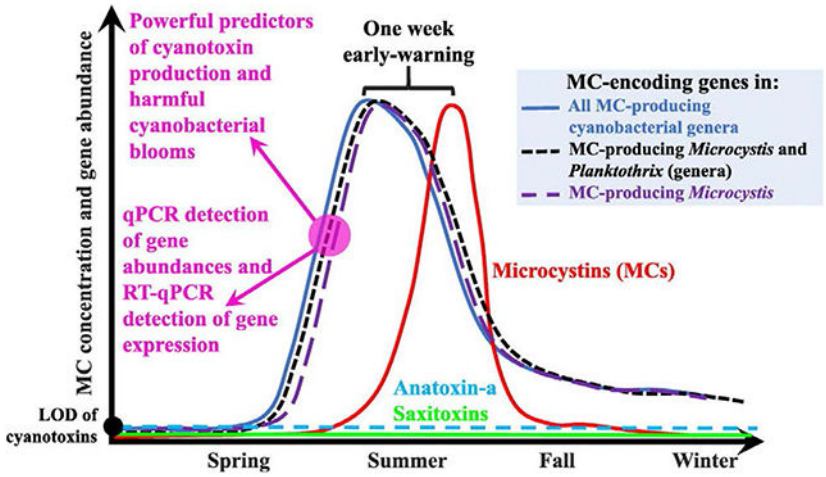
lu.jingrang@epa.gov.

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

Credit authorship contribution statement
XD and CZ analyzed the data, drafted and revised the manuscript, and contributed equally to this work. IS conducted the microbiological experiments. XL performed the bioinformatic analysis. JA determined the physiochemical water quality parameters. IS, XL, JA, and JL revised and polished the manuscript. JL was the principal investigator who designed the study and finalized the article. All authors significantly contributed to this study, proofread the manuscript, and approved the submitted version for publication in *Science of the Total Environment*.

suggested that cyanotoxin-encoding genes are promising predictors of MC production in inland freshwater lakes, such as Harsha Lake. The evaluated early-warning system can be a useful tool to assist lake managers in predicting, mitigating, and/or preventing HCBs.

Graphical Abstract



Keywords

Cyanobacteria; Microcystins; Microcystis; Planktothrix; Public health; qPCR/RT-qPCR

1. Introduction

Harmful cyanobacterial blooms (HCBs; also known as cyanobacterial harmful algal blooms even though cyanobacteria are not eukaryotic algae) frequently occur in inland freshwater bodies during warm seasons. In the current study, we define an HCB event (or HCB in short) to be the significant growth of cyanobacteria with cyanotoxin production during warm seasons. HCBs potentially produce a large number of cyanotoxins with high concentrations such as microcystins (MCs, the dominant freshwater cyanotoxins), cylindrospermopsin, anatoxins, nodularins, and saxitoxins (Brooks et al., 2016; Munoz et al., 2019; Pearson et al., 2010; Schmale et al., 2019; Watson et al., 2014). Cyanotoxins significantly threaten aquatic ecosystems (e.g., causing fish poisoning and fish/waterfowl/mammalian mortality) and potentially contaminate municipal drinking water, posing serious public health risks (Backer and Miller, 2016; Brooks et al., 2017; Ferrão-Filho and Kozlowsky-Suzuki, 2011; Grattan et al., 2016; Landsberg, 2002). Ingestion of cyanotoxins can cause acute and chronic toxicity effects on the human liver, kidney, lung, nervous tissues, gastrointestinal tract, and immune systems (Buratti et al., 2017; Funari and Testai, 2008; Kubickova et al., 2019; Vu et al., 2020; Zanchett and Oliveira-Filho, 2013). Therefore, authorities shut down municipal water supplies when serious HCBs occur in drinking water sources (e.g., the 2014 drinking water crisis in Toledo, Ohio, USA, caused by a *Microcystis* bloom in Lake Erie) (Pelley, 2016; Steffen et al., 2017; Wynne and Stumpf, 2015). The United States Environmental Protection Agency (US EPA) has issued ten-day non-regulatory health advisory limits for MCs (i.e., MC concentrations at which adverse health effects are unlikely to occur over a ten-day

exposure) in drinking water (0.3 μg·L^{-1} for bottle-fed infants and pre-school children; 1.6 μg·L^{-1} for school-age children through adults) (US EPA, 2015) and recreational water (8.0 μg·L^{-1}) (US EPA, 2019). An early-warning system predicting HCBs is critical for timely or early actions to protect drinking water sources and ecosystems and to limit recreational exposure. Once massive cyanobacterial growth has occurred, effectively alleviating HCBs and cyanotoxin production is difficult or even impossible.

Monitoring the changes in abundances and transcript levels of cyanotoxin-encoding genes is a promising approach to predict cyanotoxin production (Lu et al., 2020). MCs have more than 279 variants and are encoded solely by the gene cluster *mcy* (Abdallah et al., 2021; Bouaïcha et al., 2019; Cheng et al., 2021; Overlingė et al., 2021; Pearson et al., 2007; Pearson et al., 2019; Rouhiainen et al., 2004; Tanabe et al., 2009; Tillett et al., 2000; Tooming-Klunderud et al., 2008). The abundances and transcript levels of *mcy* have direct, positive correlations with MC concentrations (Chen et al., 2016a; Christiansen et al., 2008; Pacheco et al., 2016). For instance, *mcy* abundances positively correlated with MC levels in two freshwater ponds in India (Singh et al., 2015), one river in the US (Otten et al., 2015), and two inland lakes in China (Li et al., 2014). Therefore, *mcy* is a potential predictor of MC production. We developed a novel early-warning system to forecast MC production in William H. Harsha Lake (Ohio, USA) with a history of HCBs (Lu et al., 2020). The early-warning system monitored the abundances and expression (i.e., the transcript levels) of *mcy* using quantitative polymerase chain reactions (qPCR) and quantitative reverse transcription PCR (RT-qPCR), respectively. We established and used thresholds of qPCR and RT-qPCR signals of *mcy* to predict whether and when total MC concentrations would exceed health advisory limits during the 2015 HCB.

Our novel early-warning system is different from and has clear advantages over others (Brown et al., 2020; Davidson et al., 2016; Zohdi and Abbaspour, 2019). First, our approach relies on the direct, significant correlations between cyanotoxin-encoding genes and cyanotoxin production and is thus more straightforward. Previous early-warning systems rely on correlating cyanobacterial growth with environmental factors and cyanobacterial physiological parameters. Those factors and parameters affect cyanobacterial growth and can be potential indicators of cyanotoxin production. However, those factors and parameters control HCBs in indirect, complex, and often unpredictable ways. Therefore, early-warning systems based on those factors and parameters are less reliable. Second, our early-warning system is straightforward, rapid, and cost-effective. Previous early-warning systems, which generally target cyanobacterial blooms/growth rather than cyanotoxin production, are mainly based on complex mathematical (i.e., neural network) models (Henrichs et al., 2021; Lee and Lee, 2018; McGillicuddy, 2010; Ralston and Moore, 2020), remote sensing (Karki et al., 2018; Seltenrich, 2014; Tian and Huang, 2019; Wynne et al., 2013), optical and/or fluorescence detection (Ezenarro et al., 2021; Wang et al., 2018), and underwater camera imaging (Song et al., 2020). These complicated systems are difficult to use and can be cost-prohibitive.

The current study evaluating the early-warning system is novel and important. First, to implement an early-warning system in the field, we must fully evaluate the system against HCBs in different years and waterbodies. If the early-warning system is only accurate or

useful for one single bloom, it would have limited scientific and practical implications. Therefore, we have conducted comprehensive, systematic, serial studies to evaluate the early-warning system. The first work in the serial studies proposed the concept of the early-warning system and was based on the 2015 HCB in Harsha Lake (Lu et al., 2020). In the current work, we focused on evaluating the efficacy of the early-warning system in forecasting cyanotoxin production in 2016 in the same lake. We have also tested the performance of the early-warning system in Harsha Lake in 2017 and 2018 and lakes in different geographical regions in the US (Smithville Lake, Missouri, in 2017; Detroit Lake, Oregon, in 2019; and Spanaway Lake, Washington, in 2019) (Table S1). Along with the fieldwork, we have evaluated the early-warning system in more detail in laboratory experiments using a toxic *Microcystis aeruginosa* strain and tested the efficacies of various approaches in preventing or mitigating cyanotoxin production (Table S1). Our serial studies form a complete story from the initial proposal of the early-warning concept to the application of engineering strategies to prevent cyanobacterial growth. Second, while the proposal of a novel early-warning system forecasting HCBs is significant, follow-up work evaluating the system is even more important than initial development and testing (Lu et al., 2020). Without appropriate evaluation, the early-warning system cannot be applied in the field. Third, a unique, clear contribution of the current work to the field is that it established thresholds of qPCR and RT-qPCR signals. When those thresholds are reached, massive cyanotoxin production will likely happen after a short period. The period is approximately one week in Harsha Lake but could be shorter or longer in other freshwater bodies. Water treatment plants and recreational water services need those thresholds to predict massive cyanotoxin production so that timely or early actions can be taken.

This study aimed to evaluate the novel early-warning system that we developed to forecast the 2015 HCB in Harsha Lake (Lu et al., 2020). We first examined whether an HCB event reoccurred in Harsha Lake in 2016 by characterizing cyanobacterial community dynamics. We then tested whether the early-warning system could accurately predict cyanotoxin production during the 2016 HCB event. In addition, to understand how environmental factors shaped the cyanobacterial community and affected cyanotoxin production, we measured numerous physicochemical parameters (including but not limited to nutrient-metabolism-related parameters). We hypothesized that the abundances and expression of cyanotoxin-encoding genes (mainly the *mcy* cluster) are promising predictors of cyanotoxin production during HCBs in freshwater bodies.

2. Materials and methods

2.1. Study sites and lake sample collection

Harsha Lake is a multi-use reservoir in the interior low plateau ecoregion in central Clermont County, Ohio, USA. The lake has a surface area of 8.7 km^2 and an average depth of 13.1 m with a maximum depth of 34.0 m (Zhu et al., 2019). The lake provides for flood prevention, drinking water, and recreational and wildlife habitats. As a eutrophic inland freshwater lake heavily impacted by nutrient loading from urban and agricultural land, Harsha Lake experiences HCBs with diverse cyanobacterial taxa (Lu et al., 2020; Lu et al., 2019; Zhu et al., 2019).

We sampled lake water from two sites, Harsha Buoy (BUOY; latitude 39.0325, longitude −84.1377) and East Fork Lake Surface (EFLS; near the intake of a municipal water utility; latitude 39.0367, longitude −84.1381) with a prolonged (March to September) and more frequent sampling scheme. Water samples were taken from the lake surface (0 to 0.5 m depth) three times a week in June; weekly in May, July, and August; and biweekly in April and September (32 sampling events for BUOY and 31 sampling events for EFLS) from March 25th to September 28th, 2016. Chen et al. (2017) describes the sampling sites in detail. We used this sampling scheme because low-frequency, short-term sampling is an important but neglected reason for the poor performance of *mcy* and qPCR as predictors of MC production during HCBs (Beversdorf et al., 2015; Pacheco et al., 2016; Sabart et al., 2015). We sampled from only the lake surface for two reasons. First, while *Planktothrix* could live in deeper water than *Microcystis* (Reynolds and Rogers, 1976; Su et al., 2015; Yu et al., 2018), our previous study (Lu et al., 2020) and the current study showed that compared to *Microcystis*, *Planktothrix* was not a major MC producer at the two sampling sites in Harsha Lake in 2015 and 2016 as determined by the abundances and expression of the *mcy* genes. Second, in our 2015 study (Lu et al., 2020), MC concentrations at a depth of 8 m were extremely low (data not shown).

We collected a 5 L water sample at each sampling event. An aliquot (approximately 9 mL) of each sample was transferred to a glass container and frozen at −20 °C for total MC analysis. A second aliquot (50 mL) was transferred into a 50 mL conical tube (Thermo Fisher Scientific; Waltham, Massachusetts, USA) for nutrient analyses. The second aliquot was split into two sub-samples. The first sub-sample was transferred to a screw-top culture tube for total nitrogen (TN) and total phosphorus (TP) concentration determination. The second sub-sample was transferred to a disposable culture tube for total reactive phosphorus (TRP), nitrite (NO_2^-), nitrate (NO_3^-), and ammonium (NH_4^+) concentration analysis. A third aliquot (40 mL) of each sample was transferred to a scintillation vial containing one drop of concentrated sulfuric acid for total organic carbon (TOC) determination. Duplicates of a fourth aliquot (100 to 300 mL per duplicate) of each sample were filtered for DNA and RNA isolation.

2.2. Assessment of physicochemical water quality

We placed an EXO Sonde device (YSI Inc.; Yellow Springs, Ohio, USA) at BUOY to determine the following physicochemical water quality parameters: optical dissolved oxygen (ODO), water temperature, pH, specific conductivity, chlorophyll relative fluorescence, turbidity, blue-green algae phycocyanin (BGA-PC), oxidation-reduction potential (ORP), and water depth (the waves and wind constantly changed the relative locations of the Sonde device to the water surface). Some environmental parameters provided by the EXO Sonde device (i.e., salinity, conductivity, total dissolved/suspended solids, and water pressure) were derivative and thus were omitted from data analysis. Following the *Standard Methods* (APHA, 2017), we analyzed the concentrations of dissolved nutrients (NH_4^+, NO_2^-, NO_3^-, TN, TRP, and TP) using a Lachat Quickchem 8000 Flow Injection Analysis System (Hach Company; Loveland, Colorado, USA) and TOC using a Phoenix 8000 TOC Analyzer (Teledyne Tekmar; Mason, Ohio, USA). In addition, we freeze-thawed the first aliquots for three cycles and centrifuged them for 15 min (3000 ×*g*, 4 °C) to release intracellular MCs.

We then used an Enzyme-Linked Immunosorbent Assay (ELISA) (Gaget et al., 2017) to determine total MC concentrations in raw water using a Microcystins/Nodularins (ADDA) SAES ELISA Kit (Eurofins Abraxis, Inc.; Warminster, Pennsylvania, USA) on a Chromate 4300 microplate reader (Awareness Technologies; Palm City, Florida, USA). The limit of detection (LOD) for total MC concentrations in lake water was 0.15 $\mu g \cdot L^{-1}$.

2.3. DNA and RNA extraction, DNA sequencing, and qPCR and RT-qPCR

We filtered the fourth aliquot (100 to 300 mL per duplicate) of each sample using an EMD Millipore Durapore® membrane filter (pore size 0.45 μm; MilliPore; Foster City, California, USA). Each filter with captured biomass was transferred to a 1.5 mL microcentrifuge tube containing 600 μL of a cell lysis Buffer RLT Plus with RNase inhibitor (QIAGEN; Valencia, California, USA) and stored at −80 °C before total genomic DNA and RNA extraction. We lysed the stored cells by shaking and beating the tubes with a Mini-Beadbeater-16 (BioSpec Products, Inc.; Bartlesville, Oklahoma, USA). We then centrifuged the tubes (10,000 ×*g*, 3 min, room temperature) and transferred the supernatant to new microcentrifuge tubes. We used an AllPrep® DNA/RNA Mini Kit (QIAGEN) to purify total genomic DNA and RNA from the supernatant. DNA contamination was removed from the RNA extracts (TURBO DNA-free™ Kit; Life Technologies Co.; Carlsbad, California, USA). We reverse-transcribed the cleaned RNA to cDNA using a High Capacity cDNA Reverse Transcription Kit (Life Technologies Co.). DNA and RNA concentrations were determined using a Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies, Inc.; Wilmington, Delaware, USA). The DNA and RNA extracts were stored at −80 °C before use.

We assessed the temporal variations of microbial communities using high-throughput 16S rRNA gene sequencing to determine whether the microbial community structures during the 2015 and 2016 HCBs were similar. The sequencing also allowed us to determine how the relative abundance (RA) of cyanobacteria against total bacteria changed over time and how the RA of dominant cyanobacterial genera against cyanobacteria changed over time. The sequencing was conducted following an established procedure (Lamendella et al., 2018; Zhang et al., 2021a). Briefly, we prepared the sequencing libraries by amplifying the V4 region using Illumina iTag PCR Mixtures (Illumina, Inc.; San Diego, California, USA) on an MJ Research PTC-200 Thermocycler (Bio-Rad; Hercules, California, USA). Each PCR (25 μL) contained (final concentration or mass) 5 to 10 ng of template DNA, 1× PCR buffer, 0.8 mM dNTPs (each), 0.625 U of Taq polymerase, 0.2 μM barcoded forward primer (515F), and 0.2 μM universal Illumina reverse primer (806R) (Caporaso et al., 2012; Walters et al., 2016). Pooled PCR amplicons were gel-purified using a QIAquick Gel Extraction Kit (QIAGEN) and quantified using a Qubit 2.0 fluorometer (Life Technologies Co.). We then checked the quality of the libraries using a 2100 Bioanalyzer DNA 1000 Chip (Agilent Technologies; Santa Clara, California, USA). Library pools were size-verified using a Fragment Analyzer™ Automated CE System (Advanced Analytical Technologies, Inc.; Ames, Iowa, USA) and quantified using a Qubit High Sensitivity Double-Stranded DNA Kit (Invitrogen™, Thermo Fisher Scientific). Laragen, Inc. (Culver City, California, USA) sequenced the purified DNA libraries using an Illumina MiSeq Reagent Kit v2 (500-cycles; Illumina, Inc.) with a 16S rRNA gene library sequencing primer set for 300 base pair (bp) paired-end reads. The raw sequencing data are available at the National Center

for Biotechnology Information's Sequence Read Archive (SRA) website (ncbi.nlm.nih.gov/sra) under the BioProject ID PRJNA752140 (Accession numbers SAMN20584775 to SAMN20584893).

The raw sequencing reads were merged against a minimum overlap length of 190 bp. We conducted quality filtering for the merged reads against a maximum error rate of 1.0% and truncated the filtered reads at a length of 370 bp using USEARCH (version 7) (Edgar, 2010). We subsequently analyzed the reads using QIIME (version 1.9.0) (Caporaso et al., 2010; Caporaso et al., 2011). We identified and removed chimeras using USEARCH61 (Edgar, 2013). We identified the open reference operational taxonomic units (OTUs) using the USEARCH61 algorithm and performed taxonomy assignment using the Greengenes 16S rRNA gene database (the May 2013 release of gg_13_5_99) with a minimum similarity of 97% (DeSantis et al., 2006). Any water sample with less than 1000 sequences was removed from the OTU table before cumulative sum scaling (CSS) normalization.

We determined the abundances of the cyanotoxin-encoding and nutrient metabolism (such as N_2-fixation) genes in the water samples via SYBR® Green qPCR and their expression (i.e., transcript levels) via SYBR® Green RT-qPCR (Table S2) (Chen et al., 2017; Lu et al., 2020; Lu et al., 2019; Zhang et al., 2021b). Specifically, we quantified the abundances and expression of *mcyA* and *mcyE* to target cyanobacterial MC producers, *mcyG* to target MC producers in *Microcystis*, *anaC* to target cyanobacterial anatoxin-a producers, and *sxtA* to target cyanobacterial saxitoxin producers. To monitor N_2-fixation, we quantified the abundances and expression of the N_2-fixation gene *nif* in genera *Dolichospermum* (née *Anabaena*) (Li et al., 2016) and *Nostoc*. We also quantified the abundances and expression of the inorganic phosphate scavenging gene *pstS* in *Dolichospermum*. To facilitate reporting the hosts of these functional genes, we named the qPCR/RT-qPCR targets *mcyA-Cya* (*Cya*: cyanobacteria), *mcyE-Cya*, *mcyG-Mic* (*Mic*: *Microcystis*), *anaC-Cya*, *sxtA-Cya*, *nif-Ana* (*Ana*: *Anabaena*, a genus with a current name *Dolichospermum*), *nif-Nos* (*Nos*: *Nostoc*), and *pstS-Ana*.

qPCR and RT-qPCR were performed on a QuantStudio™ 6 Flex System (Life Technologies Co.). Each reaction (20 μL) contained (final concentration or volume) 10 μL of 2× qPCR SYBR® Green Master Mix (Life Technologies Co.), 0.2 μM primers (each; Integrated DNA Technologies, Inc.; Coralville, Iowa, USA), and 2 μL of template DNA or cDNA. The thermal cycling conditions were 40 cycles of (95 °C for 15 s; the annealing temperature of 54, 56, or 60 °C for 30 s; and 72 °C for 30 s) and a final hold step at 72 °C for 5 min (Table S2).

The DNA or cDNA was quantified against a standard series constructed in-house. The standard series for *mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic* were constructed from the genomic DNA of *M. aeruginosa*. The standard series for *anaC-Cya* was constructed from the genomic DNA of *Aphanizomenon flos-aquae* FACHB-1039 isolated from Dianchi Lake, Kunming City, Yunnan Province, China. We also identified *Aph. flos-aquae* from Hasha Lake (Zhu et al., 2019). The standard series of *nif-Nos, nif-Ana, pstS-Ana*, and *sxtA-Cya* were conventional PCR products generated from DNA isolated from water samples from Harsha Lake. Each quantification was run in triplicate, and each qPCR plate contained a

triplicate, six-point standard curve with target gene concentrations ranging from 10^1 to 10^6 copies per microliter (tenfold serial dilution). We checked PCR inhibition by analyzing 10-fold diluted DNA extracts using qPCR and removed qPCR datapoints where significant PCR inhibition was detected following an established protocol (Zhang et al., 2021b).

2.4. Data processing and statistical analysis

qPCR and RT-qPCR data were transformed to $\log_{10}(GCN{\cdot}mL^{-1})$ (GCN: genome or gene copy number). We performed statistical analysis using R 3.6.1 (R Core Team, 2018). Specifically, we analyzed the diversity and structure of microbial communities using the phyloseq package (McMurdie and Holmes, 2013). We calculated the Pearson correlation coefficients and the associated *p*-values among various functional genes using the corrplot package (Wei and Simko, 2021). We conducted a TOBIT analysis with left-censored values using the AER package (Kleiber and Zeileis, 2008) to determine the relationship between total MC concentrations and *mcy* abundances and expression. By generating receiver operating characteristic (ROC) curves and calculating the areas under the ROC curves (AUC) using the pROC package, we made 7- and 14-d lag correlations (95% confidence interval) between total MC concentrations and the abundances and transcript levels of *mcy*. We then used the ROC curves to predict total MC concentrations and compared the predicted concentrations with health advisory limits of 0.3, 1.6, and 8.0 (and/or 4) $\mu g{\cdot}L^{-1}$ (Lu et al., 2020; Robin et al., 2011). To correlate physicochemical water quality parameters with the abundances and expression of functional genes and total MC concentrations, we performed a canonical correspondence analysis (CCA) using the vegan package (Dixon, 2003; Ji et al., 2019). We analyzed the effects of nutrients and other physicochemical parameters on MC concentrations via the MASS package based on Akaike Information Criterion (AIC) with a backward elimination procedure (Zhang, 2016) to rule out unrelated factors and optimize the multivariance linear regression model. The significance level was 0.05.

3. Results and discussion

3.1. Dynamic cyanobacterial community structures and potential cyanotoxin producers

The microbial community structures were monitored from March to September 2016 at BUOY and EFLS in Harsha Lake. We visualized the similarity among the bacterial community structures in different water samples with an Unweighted Pair Group Method with Arithmetic Mean (UPGMA) dendrogram (Fig. S1). Samples clustered by sampling time for both sites, suggesting significant temporal variations in the cyanobacterial community structures. We speculated that the significant variations were due to environmental variables such as pH, water temperature, light intensity, and nutrient levels. Other reasons or factors could also contribute to the significant temporal changes in the cyanobacterial community structures.

At the phylum level, the bacterial community structures were stable throughout the sampling period with limited variations in RA (Fig. S2). Proteobacteria, Cyanobacteria, Bacteroidetes, Actinobacteria, Planctomycetes, Verrucomicrobia, and Chloroflexi (approximate descending order of RA) were the dominant phyla at both sampling sites. The RA of cyanobacteria was

greater between late May and late September than before late May and after late September, indicating significant growth of cyanobacteria during the warm season, especially from early June to late July.

The cyanobacterial community structures at the genus level changed substantially from early June to late July (Fig. S3). *Dolichospermum*, *Synechococcus*, *Pseudanabaena*, *Prochlorothrix*, *Planktothrix*, and *Microcystis* were the dominant genera and potential cyanotoxin producers at both sampling sites. Similar to 2015 (Lu et al., 2020), *Microcystis* and *Planktothrix* were the major MC producers in 2016 and more abundant in June and/or July than in other months, indicating the reoccurrence of an HCB event in 2016. In addition, the RA of *Microcystis* and *Planktothrix* significantly and positively correlated with each other at both sampling sites (Pearson correlation coefficients ≥ 0.80, p-values < 0.05) (Fig. S4). Therefore, the two major MC producers co-occurred and had a close relationship with each other. However, from early July to mid-November, the RA of *Planktothrix* was high and much greater than *Microcystis* (Fig. S3). We speculated that the greater RA of *Planktothrix* was for two reasons. First, *Planktothrix* is a better competitor for light (a major driver for phytoplankton growth) than *Microcystis* (de Araujo Torres et al., 2016; Oberhaus et al., 2007). Second, *Microcystis* prefers a warmer environment, while *Planktothrix* is adapted to a broader temperature range (Jöhnk et al., 2008; Paerl and Huisman, 2008). Water temperature in Harsha Lake typically starts decreasing significantly in late August.

The RA of other cyanobacterial genera also changed dynamically over time. For instance, the RA of *Dolichospermum* increased from early April to early June (peak RA approximately 40%) but subsequently decreased and stabilized starting in early July (RA approximately 10%). The RA of *Synechococcus* was generally greater after June than in earlier months. From late June to early November, *Synechococcus* was more abundant than other identified cyanobacterial genera. Previous studies similarly found that the abundances of *Dolichospermum* and *Synechococcus* had significant temporal variations in lakes (Becker et al., 2007; Zhang et al., 2020).

The dominant cyanobacteria, such as *Microcystis* and *Synechococcus*, grow well in eutrophic freshwater bodies (Akins et al., 2018; Beaulieu et al., 2018; Liu et al., 2019). The relatively high abundances of the dominant cyanobacteria (Fig. S3) were thus a result of the eutrophication in Harsha Lake in 2016 as they were in 2015. Taken together, the dynamic cyanobacterial community structures preliminarily evidenced a reoccurred HCB event and the associated microbial communities in 2016 because of consistent eutrophication in Harsha Lake (Figs. S1 to S3).

3.2. Temporal variations in cyanotoxin-encoding genes and cyanotoxin production

We monitored the abundances and expression (i.e., transcript levels) of multiple cyanotoxin-encoding genes (Table S3) while focusing on three MC-encoding genes (*mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*). The qPCR/RT-qPCR target *mcyE-Cya* covers all major MC-producing cyanobacterial genera (such as *Dolichospermum*, *Microcystis*, *Planktothrix*, and *Nostoc*) and certain nodularin-producing genera (such as *Nodularia*) (Jungblut and Neilan, 2006; Lu et al., 2020). The qPCR/RT-qPCR target *mcyA-Cya* covers fewer genera (mainly MC-producing *Microcystis* and *Planktothrix* and possibly other MC-producing

genera) (Hisbergues et al., 2003). The qPCR/RT-qPCR target *mcyG-Mic* is specific to MC-producing *Microcystis* (Ngwa et al., 2014). In agreement with previous studies in Harsha Lake (Lu et al., 2020) and the Macau storage reservoir (Zhang et al., 2014a), the abundances of these three qPCR/RT-qPCR targets were comparable at each sampling time. In addition, temporal variations in the abundances of these targets evidenced that an HCB occurred from late May to early August 2016 in Harsha Lake (Figs. 1A and 1C), similar to the 2015 HCB in the same lake. Since *mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic* had comparable abundances and transcript levels during the 2016 HCB (Fig. 1), *Microcystis* and *Planktothrix* were the dominant MC producers, while other MC-producing cyanobacterial genera such as *Dolichospermum* were in low abundances or potentially absent during the 2016 HCB. In addition, we found that *Microcystis* produced significantly more MCs than *Planktothrix* because both the abundances and expression of *mcyE-Cya* and *mcyG-Mic* were comparable.

The abundances of the three MC-encoding genes (*mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*) (Figs. 1A and 1C) varied significantly from late March to late September. Between late March and late May, the abundances of each MC-encoding gene were at or below approximately 1 $\log_{10}(\text{GCN}\cdot\text{mL}^{-1})$ at both sampling sites. Starting in late May, the abundances of each MC-encoding gene dramatically increased and peaked in mid- to late June at approximately 5.5 $\log_{10}(\text{GCN}\cdot\text{mL}^{-1})$. The gene abundances then gradually decreased until early August, whereafter the abundances for each MC-encoding gene were stable at approximately 2.5 $\log_{10}(\text{GCN}\cdot\text{mL}^{-1})$. These significant temporal changes in the abundances of the three MC-encoding genes implied that an HCB reoccurred in 2016 (Lu et al., 2020). To further understand the occurrence of this HCB, we measured the transcript levels of these three MC-encoding genes. The transcript levels followed a similar temporal trend to that of the gene abundances. For instance, the transcript levels also peaked in late June to early July at approximately 2.5 $\log_{10}(\text{GCN}\cdot\text{mL}^{-1})$ at both sampling sites (Figs. 1B and 1D). Subsequently, the transcript levels decreased and were minimal after early August.

The temporal variations in the abundances and transcript levels of these three MC-encoding genes (*mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*) suggested that MC-producing cyanobacteria (mainly *Microcystis* and *Planktothrix*, particularly *Microcystis*) significantly grew between late May and late July. However, the abundances and expression of these MC-encoding genes indicated only the production of total MCs by cyanobacteria but might not indicate the concentrations of MC-producing cyanobacteria. For instance, from late May to early July, the RA of *Microcystis* and *Planktothrix* was comparable, while the RA of *Planktothrix* was high and much greater than that of *Microcystis* from early July to mid-November (Fig. S3). Even though the RA of *Planktothrix* was comparable to or even greater than that of *Microcystis* during the HCB, *Microcystis* was still the dominant MC producer at the two sampling sites in Harsha Lake in 2016 as determined by the abundances and expression of the *mcy* genes (Fig. 1). In other freshwater bodies, however, *Planktothrix* could be a major or dominant MC producer. For instance, in 1995 and 1996 in 55 German freshwater bodies, the concentrations of MCs per unit dry cyanobacterial biomass were highest when *P. rubescens* was dominant, followed by *P. agardhii* and *Microcystis* spp. (Fastner et al., 1999; Tonk et al., 2005).

The temporal trend in MC production was similar to those of the abundances and expression of the three MC-encoding genes (Fig. 1). The significant MC production confirmed the reoccurrence of an HCB event in 2016. The MC concentrations were below the LOD (0.15 $\mu g \cdot L^{-1}$) before early June at both sampling sites. Starting in early June, total MC concentrations dramatically increased, peaking at approximately 10 $\mu g \cdot L^{-1}$ in mid-June at both sites, which is greater than the health advisory limit for recreational water of 8 $\mu g \cdot L^{-1}$ (US EPA, 2019). From mid-June to late July, total MC concentrations dramatically decreased. After early August, total MC levels were below the LOD even though the abundances of the MC-encoding genes were relatively high and stable at approximately 2.5 $\log_{10}(GCN \cdot mL^{-1})$. The undetected total MCs when the MC-encoding genes had relatively high abundances were presumably because these genes were not highly expressed. Indeed, the transcript levels of these MC-encoding genes were minimal after early August. Overall, before total MC concentrations peaked, temporal changes in the abundances of MC-encoding genes, mRNA production by these genes, and total MC concentrations displayed similar patterns. The close correlations among the MC-encoding genes, their expression, and MC concentrations before total MC concentrations peaked confirmed that MC-encoding genes are powerful molecular markers to assess and predict MC production during HCBs (Lu et al., 2020). In addition, total MC levels in three (4.8%) of the 63 water samples (32 for BUOY and 31 for EFLS) exceeded the health advisory limit of 8 $\mu g \cdot L^{-1}$ for recreational water (US EPA, 2019).

This study focused on the dominant cyanotoxins (i.e., MCs) and the MC-encoding genes (i.e., the *mcy* genes) in Harsha Lake. In addition, we assessed the concentrations of anatoxin-a and saxitoxin and genes involved in their production (i.e., *anaC* and *sxtA*, respectively) (Chia et al., 2018; Sabart et al., 2015) for two reasons. First, low levels of anatoxin-a (but not saxitoxin) were detected in 2015 from Harsha Lake (data not shown). Second, we used qPCR and RT-qPCR as an early-warning tool to target not only MCs but also other common cyanotoxins in freshwater bodies such as anatoxin-a and saxitoxin.

anaC-Cya had high abundances and was highly expressed from late May to late September (Fig. 2), but it did not produce anatoxin-a to a level greater than the LOD (0.05 $\mu g\ L^{-1}$) at both sampling sites throughout the HCB event. Although *sxtA-Cya* was detected throughout the HCB, its expression was low or even minimal (Fig. 2), resulting in undetectable saxitoxin concentrations (i.e., < 0.05 $\mu g\ L^{-1}$). The undetected saxitoxin could be presumably because some non-saxitoxin-producing species contain the *sxtA* gene (Ballot et al., 2010; Le Tortorec et al., 2016; Murray et al., 2011; Stüken et al., 2011; Zhang et al., 2014b), where *sxtA* might not be transcribed into mRNA (Akbar et al., 2020) or the function of *sxtA* is not limited to saxitoxin production (Geffroy et al., 2021; Hackett et al., 2013). Additionally, *Microcystis* as the dominant MC producer in the current study might have suppressed the production of anatoxin-a and saxitoxin (Al-Tebrineh et al., 2010; Chia et al., 2018).

3.3. Temporal variations in nutrient metabolism

We tracked the temporal changes in the abundances and expression of two nutrient metabolism genes (the N_2-fixation gene *nif* and the inorganic phosphate scavenging gene *pstS*) using three qPCR/RT-qPCR assays (*nif-Ana* and *pstS-Ana* specific to

Dolichospermum; *nif-Nos* specific to *Nostoc*). The abundances of *nif-Ana* and *nif-Nos* increased dramatically starting in mid-April and peaked in early June (Figs. 1A and 1C). Afterward, the abundances of *nif* gradually decreased until mid-August. From mid-August to late August (the end of the monitoring), the abundances of *nif* significantly increased again. The transcript levels of *nif* followed a similar pattern. The transcript levels of *nif-Ana* and *nif-Nos* significantly increased starting in mid-May and peaked in early June at approximately 5 $\log_{10}(GCN{\cdot}mL^{-1})$ (Figs. 1B and 1D). The transcript levels of *nif* then dramatically decreased to approximately 2.5 $\log_{10}(GCN{\cdot}mL^{-1})$ in mid-June and subsequently rapidly increased again to approximately 4.5 $\log_{10}(GCN{\cdot}mL^{-1})$ in late June. The transcript levels then gradually decreased until early to mid-August and significantly increased again until the end of the monitoring (late August).

The temporal changes in *nif* abundances and dissolved nitrogen concentrations were closely associated. For instance, the abundances of *nif* significantly increased starting in mid-April, and the transcript levels of *nif* significantly increased starting in mid-May (Fig. 1). NH_4^+-N concentrations formed significant peaks between late April and late May (Fig. S5), suggesting active N_2-fixation (Bothe et al., 2010; Zehr, 2011). The significant increase in NH_4^+-N concentrations could have promoted the growth of nitrifiers (phylogenetically not closely associated with cyanobacteria) (Chen et al., 2016b; Hampel et al., 2018; Purkhold et al., 2000), which consume NH_4^+-N and generate NO_2^--N and NO_3^--N. Consequently, peaks for the concentrations of NO_2^--N and NO_3^--N occurred approximately one week after the peaks for the concentrations of NH_4^+-N. Starting in early June, the concentrations of dissolved NH_4^+-N, NO_2^--N, NO_3^--N, and TN gradually decreased, indicating that cyanobacteria (and other bacteria) significantly consumed dissolved inorganic nitrogen (mainly NH_4^+-N) for growth (Rittmann and McCarty, 2020), decreasing inorganic nitrogen concentrations in this lake.

The abundances of the phosphate transporter gene *pstS* significantly increased starting in mid-April, peaked in early June, and subsequently gradually decreased until mid-August (Fig. 3). From mid-August to the end of the monitoring (late August), the abundances of *pstS* significantly increased again. Similarly, the transcript levels of *pstS* significantly increased starting in mid-May, peaked in early June, subsequently decreased until mid-August, and then significantly increased again. Therefore, *pstS* was active during the 2016 HCB.

The concentrations of dissolved TRP and TP changed significantly over time. The concentrations of dissolved TRP and TP significantly increased starting in late April, peaked in mid-May, and subsequently gradually decreased. The abundances of *pstS-Ana* were low before mid-April, dramatically increased since mid-April, peaked in early June, and then gradually decreased (Fig. 3). Therefore, *pstS-Ana* was well regulated by extracellular phosphorus concentrations (Harke et al., 2012; Pereira et al., 2019). For instance, when the concentrations of dissolved TRP and TP peaked in mid-May, the expression of *pstS-Ana* was suppressed. Afterward, the transcript levels of *pstS-Ana* increased when the concentrations of dissolved TP and TRP decreased. Therefore, *pstS-Ana* was more active when external phosphorus concentrations were low but was downregulated when external phosphorus levels were high (Dyhrman and Haley, 2006).

3.4. The effect of nutrient metabolism on cyanotoxin production

Nutrient metabolism is an important cyanobacterial activity that catalyzes and sustains cyanotoxin production and HCBs (Glibert et al., 2018; Lu et al., 2019). Therefore, closely monitoring cyanobacterial nutrient metabolism can help explain the mechanisms, dynamics, and kinetics of HCBs.

This study suggested that active nutrient metabolism, especially N_2-fixation, potentially promoted total MC production for two reasons. First, cyanobacterial nutrient metabolism and cyanotoxin production were closely correlated. The abundances of the MC-encoding gene cluster *mcy* (i.e., *mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*) and the N_2-fixation gene *nif* (*nif-Ana* and *nif-Nos*) were positively correlated (R_{Pearson} 0.75 to 0.82, p-values < 0.05) (Fig. 4). The transcript levels of *mcy* and *nif* also had significant positive correlations (R_{Pearson} 0.49 to 0.65, p-values < 0.05). Moreover, the abundances (R_{Pearson} 0.74 to 0.81, p-values < 0.05) and transcript levels (R_{Pearson} 0.59 to 0.65, p-values < 0.05) of *mcy* and *pstS-Ana* had significant positive correlations. In addition to *mcy*, the genes involved in the synthesis of anatoxin-a (*anaC*) and saxitoxin (*sxtA*) had positive correlations with *nif* (R_{Pearson} 0.59 to 0.78 for gene abundances, 0.16 to 0.61 for transcript levels, p-values generally less than 0.05) and *pstS-Ana* (R_{Pearson} 0.56 and 0.66 for gene abundances, 0.23 and 0.55 for transcript levels, p-values generally less than 0.05) (Fig. 4). A significant positive correlation between the concentrations of two potential cyanotoxin producers (*Aphanizomenon* and *Dolichospermum*) and *nif* abundances existed in Utah Lake (a freshwater lake in Salt Lake City, Utah, USA) (Li et al., 2020). We also identified multiple potential cyanotoxin-producing species in the genera *Aphanizomenon* and *Dolichospermum* such as *Aph. flos-aquae*, *D. circinale*, *D. crassum*, and *D. ellipsoides* from Harsha Lake in 2015 (Zhu et al., 2019). However, a cyanobacterial species could include both cyanotoxin-producing and non-producing strains. The identification and confirmation of cyanotoxin-producing strains should be included in future studies. The significant correlations between nutrient metabolism genes and cyanotoxin-encoding genes suggested that nutrient-related factors (such as nutrient circulation, loading, and concentrations) and cyanotoxin (mainly total MC) production could have an inherent connection in freshwater bodies (Bartoli et al., 2018; Huisman et al., 2018; Lu et al., 2019; Paerl et al., 2011b; Park et al., 2017; Xu et al., 2017).

Second, nutrient metabolism genes became active before increases in the abundances and transcript levels of cyanotoxin-producing genes and total MC concentrations. The initial significant increases in the transcript levels of *nif* (on May 11th) were approximately three weeks before those of *mcy*, *anaC-Ana*, and *sxtA-Cya* (all on June 01st) at both sampling sites (Figs. 1B, 1D, and 2). Significant growth of N_2-fixing cyanobacteria and increases in the transcript levels of N_2-fixing genes before significant growth of cyanotoxin producers such as *Microcystis* and production of cyanotoxins are common in eutrophic lakes (Tanvir et al., 2021). The initial increases in the transcript levels of *pstS* on May 18th (BUOY) and May 11th (EFLS) (Fig. 3) were also earlier than those of *mcy*, *anaC-Cya*, and *sxtA-Cya*. In addition, during the HCB, the abundances and transcript levels of *nif-Ana*, *nif-Nos*, and *pstS-Ana* were generally much greater than those of *mcy*, *anaC-Cya*, and *sxtA-Cya* (Fig. 1, Fig. 2, Fig. 3). Thus, heterocystous cyanobacteria outgrew cyanotoxin (mainly

MC) producers, and nutrient metabolism was more active than cyanotoxin production. The more active nutrient metabolism genes compared with cyanotoxin-encoding genes further suggested that nutrient metabolism could trigger cyanotoxin production and stimulate HCBs in freshwater bodies (Lu et al., 2019; Paerl and Otten, 2013). Therefore, controlling nutrient loadings and suppressing cyanobacterial nutrient metabolism, specifically N_2-fixation, are useful strategies for mitigating and preventing HCBs in freshwater aquatic ecosystems (Hamilton et al., 2016; Nwankwegu et al., 2019; Paerl et al., 2016; Paerl et al., 2018).

Even though nutrient-metabolism-related genes may promote cyanotoxin production, these genes are not good predictors of cyanotoxin production for multiple reasons. First, nutrient metabolism (N_2-fixation and inorganic phosphate scavenging) and cyanotoxin production are separate pathways. Second, certain non-N_2-fixing cyanobacterial species can produce cyanotoxins, and certain N_2-fixing cyanobacterial species do not produce cyanotoxins. Third, in the current study, anatoxin-a and saxitoxins remained below the LOD or were even absent after significant increases in the transcript levels of the N_2-fixing genes and the inorganic phosphate scavenging gene. Similarly, even though the nutrient metabolism genes and the genes involved in the synthesis of anatoxin-a and saxitoxins had significant positive correlations (Fig. 4), anatoxin-a and saxitoxins were undetected. Fourth, significant increases in the transcript levels of *nif* genes might indicate that cyanotoxin (such as MC) production could occur subsequently but might not predict cyanotoxin concentrations (i.e., no clear quantitative correlation). Indeed, because total MC production could be affected by various factors and those factors could affect one another (i.e., intercorrelations), a simple correlation between total MC concentrations and the abundances/transcript levels of *nif* was not found (data not shown). These factors include but are not limited to light intensity, sulfur concentration, nutrient (i.e., nitrogen and phosphorus) levels, nitrogen-to-phosphorus ratio, iron concentration, water temperature, pH, salinity, turbidity, the presence of xenobiotics, and the interacts between MC producers and their predators and competitors (Jähnichen et al., 2011; Kotak et al., 2000; Pineda-Mendoza et al., 2016; Sevilla et al., 2008). Therefore, nutrient-metabolism-related genes are not useful predictors of MC production.

3.5. mcy *as a promising predictor of MC production: evaluating the early-warning system*

This study focused on evaluating our recently developed early-warning system (Lu et al., 2020), which uses MC-encoding genes as molecular markers to predict MC production in Harsha Lake. At both sampling sites, the abundances of the three *mcy* genes (*mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*) increased before the initial significant increases in total MC concentrations (Figs. 1A and 1C). Furthermore, the temporal variations of MC concentrations and the abundances/expression of the MC-encoding genes followed similar patterns before total MC concentrations peaked (Fig. 1). Therefore, the MC-encoding genes are promising molecular markers to predict and assess MC production during HCBs in Harsha Lake.

To evaluate the early-warning system, we correlated total MC concentrations with the same-day abundances and transcript levels of the *mcy* genes using an association model. We also correlated total MC concentrations with the past (7 d earlier) abundances and transcript levels of the *mcy* genes (i.e., the early-warning system). Specifically, we used

TOBIT regression with left-censored values to determine *mcy* abundances and transcript levels that would indicate the same-day or predict 7-d later total MC concentrations to reach the three health advisory limits. These limits are 0.3 $\mu g \cdot L^{-1}$ in drinking water for bottle-fed infants and pre-school children, 1.6 $\mu g \cdot L^{-1}$ in drinking water for school-age children through adults (US EPA, 2015), and 8.0 $\mu g \cdot L^{-1}$ in recreational water (US EPA, 2019).

For the same-day association model, the correlation coefficients between predicted and observed total MC concentrations when the abundances/transcript levels of *mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic* were used as the predictors were 0.80/0.82, 0.77/0.81, and 0.87/0.87, respectively (Fig. S6). Therefore, the abundances (determined via qPCR) and transcript levels (determined via RT-qPCR) of *mcy* are comparably powerful in indicating the same-day total MC concentrations. The same-day association model (confidence level 95%) indicated that total MC concentrations would exceed the health advisory limits of 0.3, 1.6, and 8.0 $\mu g \cdot L^{-1}$ if the mean abundances and mean transcript levels of the three *mcy* genes are greater than 3.70 and 0.94, 4.59 and 1.81, and 5.44 and 2.64 $\log_{10}(GCN \cdot mL^{-1})$, respectively (Fig. S7A). These thresholds for the MC-encoding genes are comparable to those for the 2015 HCB (Fig. S7B) (Lu et al., 2020). Therefore, the abundances and transcript levels of *mcy* in two consecutive years both significantly correlated with the same-day total MC concentrations, supporting the feasibility of using the *mcy* genes to assess MC production in Harsha Lake and probably other freshwater bodies. In addition, the World Health Organization (WHO) assumes that each MC-producing cyanobacterial cell contains 0.2 pg of total MCs (US EPA, 2019; World Health Organization, 2003). If all cells with the MC-encoding genes produce MCs, total MC levels would exceed 0.3, 1.6, and 8.0 $\mu g \cdot L^{-1}$ when *mcy* abundances are greater than 3.18, 3.90, and 4.60 $\log_{10}(GCN \cdot mL^{-1})$, respectively. These gene abundances of 3.18, 3.90, and 4.60 $\log_{10}(GCN \cdot mL^{-1})$ are comparable to the ones determined by the same-day association model [3.70, 4.59, and 5.44 $\log_{10}(GCN \cdot mL^{-1})$, respectively]. Therefore, the same-day association model can accurately indicate total MC levels on the basis of the abundances and transcript levels of MC-encoding genes.

Weekly and biweekly samplings are both common strategies for monitoring HCBs and cyanotoxin production in waterbodies. However, it is unknown which sampling frequency is more appropriate in Harsha Lake. We compared the abilities of our novel early-warning system based on 7- and 14-d sampling frequencies in predicting total MC production in Harsha Lake. We developed ROC curves against health advisory limits of total MC of 0.3, 1.6, and 8.0 $\mu g \cdot L^{-1}$ for the 7-d (Fig. S8) and 14-d frequencies. The AUC for the 7-d frequency were generally greater than those for the 14-d frequency (Fig. S9). Therefore, monitoring *mcy* weekly (i.e., a 7-d sampling frequency) through qPCR and RT-qPCR will more accurately indicate MC production during HCBs, especially when HCBs and the production of cyanotoxin are significant (i.e., during the peak stage). However, weekly sampling is more costly and time-consuming, while biweekly sampling could still well predict, assess, and/or indicate cyanotoxin production and the propagation of HCBs. As a result, when HCBs and the production of cyanotoxin are less pronounced (e.g., when water is cooler), biweekly sampling can reduce the cost of monitoring. In other lakes with different depths, volumes, nutrient levels, and water temperatures, the adequate sampling frequencies should be individually determined on the basis of the early-warning system.

We evaluated the early-warning system by predicting total MC concentrations during the 2016 HCB against a 7-d sampling frequency (Fig. 5). The correlation coefficients between simulated and observed total MC concentrations when the abundances/transcript levels of *mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic* were used as the predictors were 0.70/0.68, 0.65/0.58, and 0.80/0.70, respectively. These relatively large correlation coefficients (≥ 0.65) suggested that the abundances of the cyanotoxin-encoding genes (determined via qPCR) can predict MC production. In addition, RT-qPCR results not only predict cyanotoxin production (correlation coefficients ≥ 0.58) but also the dynamics of an HCB event from its beginning, to its peak, and to its end. Therefore, both gene abundances and transcript levels of *mcy* are good predictors of total MC production during HCBs in Harsha Lake. This early-warning system indicated that when the mean abundances and mean transcript levels for the three *mcy* genes reached 3.05 and 0.28, 4.45 and 1.62, and 5.80 and 2.92 $\log_{10}(GCN \cdot mL^{-1})$, total MC concentrations would have a 95% probability to reach health advisory limits of 0.3, 1.6, and 8.0 $\mu g \cdot L^{-1}$, respectively, after approximately one week in Harsha Lake (Fig. 6). On the other hand, the MC-encoding genes are not the sole factors determining or affecting total MC concentrations. Future studies might need to consider the impacts of other factors or parameters on MC production to more accurately predict HCBs.

We developed this early-warning system to reduce the frequency of (rather than to replace) direct MC measurements and thus to decrease MC analytical costs. qPCR and RT-qPCR are more cost-effective and easier to perform than direct MC determination assays such as ELISA and liquid chromatography/mass spectrometry/mass spectrometry (Lu et al., 2020). Before HCBs and significant MC production, measuring the abundances and expression of MC-encoding genes via qPCR and RT-qPCR, respectively, is more economically feasible and easier than directly determining total MC concentrations. Therefore, before significant MC production, the frequency of direct measurements of MC concentrations can be reduced, and qPCR/RT-qPCR targeting *mcy* can be used as the major HCB monitoring approach. Once the abundances and transcript levels of *mcy* reach the thresholds established by the early-warning system, a series of preventative measures could be implemented to predict and prevent HCBs and minimize the risks of cyanotoxins (mainly MCs). Similar to direct MC analysis methods, which are unable to provide on-site data, qPCR and RT-qPCR also normally require a laboratory. However, the qPCR/RT-qPCR caveat may no longer be an issue with the fast development of portable PCR devices for rapid on-site sample processing and gene abundance/expression determination (Ahrberg et al., 2016; Gou et al., 2018; Kuske et al., 1998; Marx, 2015; Zhu et al., 2020).

Once an HCB event is predicted by our novel early-warning system, actions minimizing the risks of cyanotoxins should be taken such as direct measurements of MC concentrations, cyanobacteria and cyanotoxin removal from waterbodies, advisory notices on recreational uses (e.g., beach bathing and fishing), and implementing enhanced treatment protocols at drinking water utilities. In addition, bank filtration, coagulation, and/or the floc-and-sink assays are promising approaches that can remove cyanobacteria and cyanotoxins from raw surface water (de Lucena-Silva et al., 2019; Romero et al., 2014; Sukenik et al., 2017).

3.6. Effects of physicochemical water quality parameters on MC production

Physicochemical water parameters, especially water temperature and nutrient levels or loading, significantly affect cyanobacterial growth and MC production (Jing et al., 2013; Kaebernick and Neilan, 2001; Paerl et al., 2016). This study supported that warm water could stimulate cyanotoxin production during HCBs. For instance, total MCs were detectable (LOD 0.15 $\mu g \cdot L^{-1}$) only when water temperatures reached approximately 20 °C (Figs. 1A, 1C, and S10). In addition, warm water promoted cyanobacterial nutrient metabolism. For example, increases in *nif* and *pstS* abundances (Fig. 1, Fig. 3) followed an approximate 10 °C elevation of water temperatures starting in mid-April (10 °C), to late April (18 °C), and to early June (27 °C) (Fig. S10). Previous studies also indicated that warm water (indexing heat resource in a region) significantly promotes and sustains HCBs (Jing et al., 2013; Paerl and Otten, 2013; Paul, 2008). This is presumably because cyanobacteria are well adapted to warm water and their maximal growth occurs at a relatively high temperature (Paerl et al., 2011a; Paerl and Paul, 2012; Xu et al., 2021).

We used multiple linear regression to assess the relationship between total MC concentrations and water quality parameters (Table S4 and Fig. S10), such as water temperature, pH, NH_4^+, NO_2^-, TN, and TP. However, the model was insignificant (*p*-value > 0.05). The insignificant linear correlation can also be concluded from the divergent concentration profiles of total MCs (Figs. 1A and 1C), nitrogen compounds (Fig. S5), and phosphorus compounds (Fig. 3). Other environmental parameters not monitored in this study, possibly including light intensity, wind intensity, flow rate, precipitation, stormwater runoff, nutrient loading, metal ions, and biological agents (predators and competitors), might have more strongly affected total MC production.

We also used CCA to assess the linear correlation between various environmental parameters and MC- and nutrient-metabolism-related factors. These factors include total MC concentrations, the abundances and expression of cyanotoxin-encoding genes (*mcy*, *anaC*, and *sxtA*), the abundances and expression of the N_2-fixation gene *nif* (*nif-Ana* and *nif-Nos*), and the abundances and expression of the inorganic phosphate scavenging gene *pstS* (*pstS-Ana*). The constrained CCA explained less than 50% of the total inertia, confirming that some environmental parameters not monitored in this study more strongly affected total MC production and nutrient metabolism. Taken together, except for the *mcy* genes, commonly measured water quality parameters (such as nutrient-metabolism-related parameters) are not useful predictors of MC production.

The CCA map (Fig. 7) showed that the abundances of *mcyA-Cya* and *mcyE-Cya* were positively associated with ORP and water temperature but negatively correlated with the concentrations of dissolved TRP, NH_4^+, NO_2^-, and NO_3^-. The negative correlation between *mcyA-Cya*/*mcyE-Cya* abundances and nitrogen compound concentrations could possibly be because these two genes were upregulated under nitrogen stress conditions (Pineda-Mendoza et al., 2016; Zhou et al., 2020). *mcyG-Mic* abundances and total MC concentrations positively correlated with turbidity, BGA-PC, TOC, and TN, confirming that *Microcystis* as a non-N_2 fixing genus produces more MCs at higher levels of dissolved nitrogen (Harke and Gobler, 2013; Kaebernick and Neilan, 2001).

4. Conclusions

We previously proposed a novel early-warning system that uses cyanotoxin-encoding genes to predict cyanotoxin (mainly MC) production during HCBs in Harsha Lake. To evaluate the system, we examined the HCB event in Harsha Lake in 2016. During the warm season, *Microcystis* and *Planktothrix* were the dominant MC producers, the MC-encoding genes had high abundances and were highly expressed, and total MC concentrations were relatively high (e.g., exceeding the US EPA health advisory limit in recreational water of 8 $\mu g \cdot L^{-1}$ in 4.5% of the 63 water samples), suggesting the recurrence of an HCB event. Before total MC concentrations peaked, temporal variations in the abundances of the three MC-encoding genes (*mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*), transcript levels of these genes, and total MC concentrations displayed similar patterns. In addition, total MC concentrations (time-lag) and the abundances and expression of the MC-encoding genes had statistically significant correlations. Moreover, approximately one week after the qPCR/RT-qPCR signals reached the “actionable” thresholds established by the early-warning system, total MC concentrations would exceed health advisory limits. For instance, when the mean abundances and mean transcript levels of the three *mcy* genes reached 3.05 and 0.28, 4.45 and 1.62, and 5.80 and 2.92 $\log_{10}(GCN \cdot mL^{-1})$ in Harsha Lake in 2016, total MC concentrations would reach health advisory limits of 0.3, 1.6, and 8.0 $\mu g \cdot L^{-1}$, respectively, after approximately one week. These findings indicated that cyanotoxin-encoding genes are powerful molecular markers to predict MC production during HCBs in Harsha Lake, further supporting our qPCR/RT-qPCR-based early-warning system. Other than the MC-encoding genes, commonly measured water quality parameters (such as nutrient-related parameters), are not useful predictors of cyanotoxin production. Direct and frequent (i.e., weekly or even daily) determinations of MC concentrations along with additional measures to control, mitigate, and/or prevent HCBs should be implemented once the early-warning system indicates that the qPCR/RT-qPCR signals reach these thresholds.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments and disclaimer

This research was supported by the United States Environmental Protection Agency (US EPA), Office of Research and Development' (ORD') research programs: Safe and Sustainable Water Resources (SSWR 4.3.1 and 4.3.3). We thank Drs. Eric Villegas and Mark Bagley (ORD, US EPA) for their critical review and valuable comments and suggestions. The views expressed in this manuscript are those of the authors and do not necessarily reflect on the funding agency. It has been subjected to Agency review and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Abdallah MF, Van Hassel WH, Andjelkovic M, Wilmotte A, Rajkovic A, 2021. Cyanotoxins and food contamination in developing countries: review of their types, toxicity, analysis, occurrence and mitigation strategies. Toxins 13 (11), 786. [PubMed: 34822570]

Ahrberg CD, Ilic BR, Manz A, Neužil P, 2016. Handheld real-time PCR device. Lab Chip 16 (3), 586–592. [PubMed: 26753557]

Akbar MA, Mohd Yusof NY, Tahir NI, Ahmad A, Usup G, Sahrani FK, Bunawan H, 2020. Biosynthesis of saxitoxin in marine dinoflagellates: an omics perspective. Mar. Drugs 18 (2), 103. [PubMed: 32033403]

Akins LN, Ayayee P, Leff LG, 2018. Composition and diversity of cyanobacteriaassociated and free-living bacterial communities during cyanobacterial blooms. Ann. Microbiol 68 (8), 493–503.

Al-Tebrineh J, Mihali TK, Pomati F, Neilan BA, 2010. Detection of saxitoxin-producing cyanobacteria and Anabaena circinalis in environmental water blooms by quantitative PCR. Appl. Environ. Microbiol 76 (23), 7836–7842. [PubMed: 20935128]

APHA, 2017. Standard Methods for the Examination of Water and Wastewater. 23rd edn. American Public Health Association (APHA), American Water Works Association (AWWA), and Water Environment Federation (WEF), Washington, DC, USA.

de Araujo Torres C, Lürling M, Marinho MM, 2016. Assessment of the effects of light availability on growth and competition between strains of Planktothrix agardhii and Microcystis aeruginosa. Microb. Ecol 71 (4), 802–813. [PubMed: 26691315]

Backer LC, Miller M, 2016. Sentinel animals in a one health approach to harmful cyanobacterial and algal blooms. Vet. Sci 3 (2), 8. [PubMed: 27152315]

Ballot A, Fastner J, Wiedner C, 2010. Paralytic shellfish poisoning toxin-producing cyanobacterium Aphanizomenon gracile in Northeast Germany. Appl. Environ. Microbiol 76 (4), 1173–1180. [PubMed: 20048055]

Bartoli M, Zilius M, Bresciani M, Vaiciute D, Vybernaite-Lubiene I, Petkuviene J, Giordani G, Daunys D, Ruginis T, Benelli S, 2018. Drivers of cyanobacterial blooms in a hypertrophic lagoon. Front. Mar. Sci 5, 434.

Beaulieu JJ, Balz DA, Birchfield MK, Harrison JA, Nietch CT, Platz MC, Squier WC, Waldo S, Walker JT, White KM, 2018. Effects of an experimental water-level drawdown on methane emissions from a eutrophic reservoir. Ecosystems 21 (4), 657–674. [PubMed: 31007569]

Becker S, Richl P, Ernst A, 2007. Seasonal and habitat-related distribution pattern of synechococcus genotypes in Lake Constance. FEMS Microbiol. Ecol 62 (1), 64–77. [PubMed: 17825073]

Beversdorf LJ, Chaston SD, Miller TR, McMahon KD, 2015. Microcystin mcyA and mcyE gene abundances are not appropriate indicators of microcystin concentrations in lakes. PLOS ONE 10 (5), e0125353. [PubMed: 25945933]

Bothe H, Schmitz O, Yates MG, Newton WE, 2010. Nitrogen fixation and hydrogen metabolism in cyanobacteria. Microbiol. Mol. Biol. Rev 74 (4), 529. [PubMed: 21119016]

Bouaïcha N, Miles CO, Beach DG, Labidi Z, Djabri A, Benayache NY, Nguyen-Quang T, 2019. Structural diversity, characterization and toxicology of microcystins. Toxins 11 (12), 714. [PubMed: 31817927]

Brooks BW, Lazorchak JM, Howard MD, Johnson MVV, Morton SL, Perkins DA, Reavie ED, Scott GI, Smith SA, Steevens JA, 2016. Are harmful algal blooms becoming the greatest inland water quality threat to public health and aquatic ecosystems? Environ. Toxicol. Chem 35 (1), 6–13. [PubMed: 26771345]

Brooks BW, Lazorchak JM, Howard MDA, Johnson M-VV, Morton SL, Perkins DAK, Reavie ED, Scott GI, Smith SA, Steevens JA, 2017. In some places, in some cases, and at some times, harmful algal blooms are the greatest threat to inland water quality. Environ. Toxicol. Chem 36 (5), 1125–1127. [PubMed: 28423202]

Brown AR, Lilley M, Shutler J, Lowe C, Artioli Y, Torres R, Berdalet E, Tyler CR, 2020. Assessing risks and mitigating impacts of harmful algal blooms on mariculture and marine fisheries. Rev. Aquac 12 (3), 1663–1688.

Buratti FM, Manganelli M, Vichi S, Stefanelli M, Scardala S, Testai E, Funari E, 2017. Cyanotoxins: producing organisms, occurrence, toxicity,mechanism of action and human health toxicological risk evaluation. Arch. Toxicol 91 (3), 1049–1130. [PubMed: 28110405]

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R, 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7 (5), 335–336. [PubMed: 20383131]

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R, 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. U.S.A 108 (Supplement 1), 4516–4522. [PubMed: 20534432]

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, 2012. Ultra-high-throughput microbial community analysis on the illumina HiSeq and MiSeq platforms. ISME J. 6 (8), 1621–1624. [PubMed: 22402401]

Chen K, Allen J, Lu J, 2017. Community structures of phytoplankton with emphasis on toxic cyanobacteria in an Ohio inland lake during bloom season. J. Water Resour. Prot 9 (11), 1299–1318.

Chen Q, Christiansen G, Deng L, Kurmayer R, 2016a. Emergence of nontoxic mutants as revealed by single filament analysis in bloom-forming cyanobacteria of the genus Planktothrix. BMC Microbiol. 16, 1–23. [PubMed: 26728027]

Chen X, Jiang H, Sun X, Zhu Y, Yang L, 2016b. Nitrification and denitrification by algae attached and free-living microorganisms during a cyanobacterial bloom in Lake taihu, a shallow eutrophic Lake in China. Biogeochemistry 131, 135–146.

Cheng R, Zhu H, Shutes B, Yan B, 2021. Treatment of microcystin (MC-LR) and nutrients in eutrophic water by constructed wetlands: performance and microbial community. Chemosphere 263, 128139. [PubMed: 33297127]

Chia MA, Jankowiak JG, Kramer BJ, Goleski JA, Huang IS, Zimba PV, do Carmo Bittencourt-Oliveira M, Gobler CJ, 2018. Succession and toxicity of Microcystis and Anabaena (Dolichospermum) blooms are controlled by nutrient-dependent allelopathic interactions. Harmful Algae 74, 67–77. [PubMed: 29724344]

Christiansen G, Molitor C, Philmus B, Kurmayer R, 2008. Nontoxic strains of cyanobacteria are the result of major gene deletion events induced by a transposable element. Mol. Biol. Evol 25 (8), 1695–1704. [PubMed: 18502770]

Davidson K, Anderson DM, Mateus M, Reguera B, Silke J, Sourisseau M, Maguire J, 2016. Forecasting the risk of harmful algal blooms. Harmful Algae 53, 1–7. [PubMed: 28073436]

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL, 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ.Microbiol 72 (7), 5069–5072. [PubMed: 16820507]

Dixon P, 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci 14 (6), 927–930.

Dyhrman ST, Haley ST, 2006. Phosphorus scavenging in the unicellular marine diazotroph crocosphaera watsonii. Appl. Environ. Microbiol 72 (2), 1452–1458. [PubMed: 16461699]

Edgar RC, 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26 (19), 2460–2461. [PubMed: 20709691]

Edgar RC, 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10 (10), 996–998. [PubMed: 23955772]

Ezenarro JJ, Ackerman TN, Pelissier P, Combot D, Labbé L, Muñoz-Berbel X, Mas J, Del Campo FJ, Uria N, 2021. Integrated photonic system for early warning of cyanobacterial blooms in aquaponics. Anal. Chem 93 (2), 722–730. [PubMed: 33305581]

Fastner J, Neumann U, Wirsing B,Weckesser J,Wiedner C, Nixdorf B, Chorus I, 1999. Microcystins (hepatotoxic heptapeptides) in german fresh water bodies. Environ. Toxicol 14 (1), 13–22.

Ferrão-Filho ADS, Kozlowsky-Suzuki B, 2011. Cyanotoxins: bioaccumulation and effects on aquatic animals. Marine Drugs 9 (12), 2729–2772. [PubMed: 22363248]

Funari E, Testai E, 2008. Human health risk assessment related to cyanotoxins exposure. Crit. Rev. Toxicol 38 (2), 97–125. [PubMed: 18259982]

Gaget V, Lau M, Sendall B, Froscio S, Humpage AR, 2017. Cyanotoxins: which detection technique for an optimum risk assessment? Water Res. 118, 227–238. [PubMed: 28433693]

Geffroy S, Lechat M-M, Le Gac M, Rovillon G-A, Marie D, Bigeard E, Malo F, Amzil Z, Guillou L, Caruana A, 2021. From the sxtA4 gene to saxitoxin production: what controls the variability among Alexandrium minutum and Alexandrium pacificum strains? Front. Microbiol 12, 341.

Glibert PM, Heil CA, Wilkerson FP, Dugdale RC, 2018. Global ecology and oceanography of harmful algal blooms. In: Glibert PM, Berdalet E, Burford MA, Pitcher GC, Zhou M (Eds.), Global Ecology and Oceanography of Harmful Algal Blooms. Springer, pp. 93–112.

Gou T, Hu J, Wu W, Ding X, Zhou S, Fang W, Mu Y, 2018. Smartphone-based mobile digital PCR device for DNA quantitative analysis with high accuracy. Biosens. Bioelectron 120, 144–152. [PubMed: 30173010]

Grattan LM, Holobaugh S, Morris JG Jr., 2016. Harmful algal blooms and public health. Harmful Algae 57, 2–8.

Hackett JD, Wisecaver JH, Brosnahan ML, Kulis DM, Anderson DM, Bhattacharya D, Plumley FG, Erdner DL, 2013. Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. Mol. Biol. Evol 30 (1), 70–78. [PubMed: 22628533]

Hamilton DP, Salmaso N, Paerl HW, 2016. Mitigating harmful cyanobacterial blooms: strategies for control of nitrogen and phosphorus loads. Aquat. Ecol 50 (3), 351–366.

Hampel JJ, McCarthy MJ, Gardner WS, Zhang L, Xu H, Zhu G, Newell SE, 2018. Nitrification and ammonium dynamics in taihu Lake, China: seasonal competition for ammonium between nitrifiers and cyanobacteria. Biogeosciences 15 (3), 733–748.

Harke MJ, Gobler CJ, 2013. Global transcriptional responses of the toxic cyanobacterium, Microcystis aeruginosa, to nitrogen stress, phosphorus stress, and growth on organic matter. PLOS ONE 8 (7), e69834. [PubMed: 23894552]

Harke MJ, Berry DL, Ammerman JW, Gobler CJ, 2012. Molecular response of the bloom-forming cyanobacterium, Microcystis aeruginosa, to phosphorus limitation. Microb. Ecol 63 (1), 188–198. [PubMed: 21720829]

Henrichs DW, Anglès S, Gaonkar CC, Campbell L, 2021. Application of a convolutional neural network to improve automated early warning of harmful algal blooms. Environ. Sci. Pollut. Res 28, 28544–28555.

Hisbergues M, Christiansen G, Rouhiainen L, Sivonen K, Börner T, 2003. PCR-based identification of microcystin-producing genotypes of different cyanobacterial genera. Arch. Microbiol 180 (6), 402–410. [PubMed: 14551674]

Huisman J, Codd GA, Paerl HW, Ibelings BW, Verspagen JM, Visser PM, 2018. Cyanobacterial blooms. Nat. Rev. Microbiol 16 (8), 471–483. [PubMed: 29946124]

Jähnichen S, Long BM, Petzoldt T, 2011. Microcystin production byMicrocystis aeruginosa: direct regulation by multiple environmental factors. Harmful Algae 12, 95–104.

Ji B, Liang J, Ma Y, Zhu L, Liu Y, 2019. Bacterial community and eutrophic index analysis of the East Lake. Environ. Pollut 252, 682–688. [PubMed: 31185357]

Jing YS, Jing ZH, Hu JY, Chen F, 2013. Meteorological conditions influences on the variability of algae bloom in Taihu Lake and its risk prediction. Appl. Mech. Mater 253–255, 935–938.

Jöhnk KD, Huisman J, Sharples J, Sommeijer B, Visser PM, Stroom JM, 2008. Summer heatwaves promote blooms of harmful cyanobacteria. Glob. Chang. Biol 14 (3), 495–512.

Jungblut A-D, Neilan BA, 2006. Molecular identification and evolution of the cyclic peptide hepatotoxins, microcystin and nodularin, synthetase genes in three orders of cyanobacteria. Arch. Microbiol 185 (2), 107–114. [PubMed: 16402223]

Kaebernick M, Neilan BA, 2001. Ecological and molecular investigations of cyanotoxin production. FEMS Microbiol. Ecol 35 (1), 1–9. [PubMed: 11248384]

Karki S, Sultan M, Elkadiri R, Elbayoumi T, 2018. Mapping and forecasting onsets of harmful algal blooms using MODIS data over coastal waters surrounding Charlotte County, Florida. Remote Sensing 10 (10), 1656.

Kleiber C, Zeileis A, 2008. Applied Econometrics with R. Springer,NewYork,NewYork,USA.

Kotak BG, Lam AK, Prepas EE, Hrudey SE, 2000. Role of chemical and physical variables in regulating microcystin-LR concentration in phytoplankton of eutrophic lakes. Can. J. Fish. Aquat. Sci 57 (8), 1584–1593.

Kubickova B, Babica P, Hilscherová K, Šindlerová L, 2019. Effects of cyanobacterial toxins on the human gastrointestinal tract and the mucosal innate immune system. Environ. Sci. Eur 31, 31.

Kuske CR, Banton KL, Adorada DL, Stark PC, Hill KK, Jackson PJ, 1998. Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. Appl. Environ. Microbiol 64 (7), 2463–2472. [PubMed: 9647816]

Lamendella R, Wright JR, Hackman J, McLimans C, Toole DR, Bernard Rubio W, Drucker R,Wong HT, Sabey K, Hegarty JP, Stewart DB Sr., 2018. Antibiotic treatments for clostridium difficile

infection are associated with distinct bacterial and fungal community structures. mSphere 3 (1), e00572–17. [PubMed: 29359185]

Landsberg JH, 2002. The effects of harmful algal blooms on aquatic organisms. Rev. Fish. Sci 10 (2), 113–390.

Le Tortorec AH, Tahvanainen P, Kremp A, Simis SG, 2016. Diversity of luciferase sequences and bioluminescence production in Baltic Sea Alexandrium ostenfeldii. Eur. J. Phycol 51 (3), 317–327.

Lee S, Lee D, 2018. Improved prediction of harmful algal blooms in four Major South Korea's Rivers using deep learning models. Int. J. Environ. Res. Public Health 15 (7), 1322. [PubMed: 29937531]

Li D, Gu AZ, He M, 2014. Quantification and genetic diversity of total and microcystinproducing Microcystis during blooming season in Tai and Yang-cheng lakes, China. J. Appl. Microbiol 116 (6), 1482–1494. [PubMed: 24471490]

Li H, Barber M, Lu J, Goel R, 2020.Microbial community successions and their dynamic functions during harmful cyanobacterial blooms in a freshwater lake. Water Res. 185, 116292. [PubMed: 33086464]

Li X, Dreher TW, Li R, 2016. An overview of diversity, occurrence, genetics and toxin production of bloom-forming Dolichospermum (Anabaena) species. Harmful Algae 54, 54–68. [PubMed: 28073482]

Liu C, Shi X, Fan F,Wu F, Lei J, 2019. N: P ratio influences the competition of Microcystis with its picophytoplankton counterparts, mychonastes and synechococcus, under nutrient enrichment conditions. J. Freshw. Ecol 34 (1), 445–454.

Lu JR, Zhu B, Struewing I, Xu N, Duan SS, 2019. Nitrogen-phosphorus-associated metabolic activities during the development of a cyanobacterial bloom revealed by metatranscriptomics. Sci. Rep 9, 2480. [PubMed: 30792397]

Lu JR, Struewing I, Wymer L, Tettenhorst DR, Shoemaker J, 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 Res. 170, 115262. [PubMed: 31785564]

de Lucena-Silva D, Molozzi J, dos Santos Severiano J, Becker V, de Lucena Barbosa JE, 2019. Removal efficiency of phosphorus, cyanobacteria and cyanotoxins by the "flock & sink" mitigation technique in semi-arid eutrophic waters. Water Res. 159, 262–273. [PubMed: 31102855]

Marx V, 2015. PCR heads into the field. Nat. Methods 12 (5), 393–397. [PubMed: 25924072]

McGillicuddy D Jr., 2010. Models of harmful algal blooms: conceptual, empirical, and numerical approaches. J. Mar. Syst 83 (3–4), 105. [PubMed: 28366998]

McMurdie PJ, Holmes S, 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE 8 (4), e61217. [PubMed: 23630581]

Munoz M, Nieto-Sandoval J, Cirés S, de Pedro ZM, Quesada A, Casas JA, 2019. Degradation of widespread cyanotoxins with high impact in drinking water (microcystins, cylindrospermopsin, anatoxin-a and saxitoxin) by CWPO. Water Res. 163, 114853. [PubMed: 31310856]

Murray SA, Wiese M, Stüken A, Brett S, Kellmann R, Hallegraeff G, Neilan BA, 2011. sxtA-based quantitative molecular assay to identify saxitoxin-producing harmful algal blooms in marine waters. Appl. Environ. Microbiol 77 (19), 7050–7057. [PubMed: 21841034]

Ngwa F, Madramootoo C, Jabaji S, 2014. Monitoring toxigenic Microcystis strains in the Missisquoi bay, Quebec, by PCR targeting multiple toxic gene loci. Environ. Toxicol 29 (4), 440–451. [PubMed: 22431468]

Nwankwegu AS, Li Y, Huang Y, Wei J, Norgbey E, Sarpong L, Lai Q, Wang K, 2019. Harmful algal blooms under changing climate and constantly increasing anthropogenic actions: the review of management implications. 3 Biotech 9 (12), 449.

Oberhaus L, Briand J-F, Leboulanger C, Jacquet S, Humbert JF, 2007. Comparative effects of the quality and quantity of light and temperature on the growth of Planktothrix agardhii and P. Rubescens. J. Phycol 43 (6), 1191–1199.

Otten TG, Crosswell JR, Mackey S, Dreher TW, 2015. Application ofmolecular 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.

Overlingė D, Toruńska-Sitarz A, Kataržytė M, Pilkaitytė R, Gyraitė G, Mazur-Marzec H, 2021. Characterization and diversity of microcystins produced by cyanobacteria from the curonian lagoon (SE Baltic Sea). Toxins 13 (12), 838. [PubMed: 34941676]

Pacheco ABF, Guedes IA, Azevedo S, 2016. Is qPCR a reliable indicator of cyanotoxin risk in freshwater? Toxins 8 (6), 172. [PubMed: 27338471]

Paerl HW, Huisman J, 2008. Blooms like it hot. Science 320 (5872), 57–58. [PubMed: 18388279]

Paerl HW, Otten TG, 2013. Harmful cyanobacterial blooms: causes, consequences, and controls. Microb. Ecol 65 (4), 995–1010. [PubMed: 23314096]

Paerl HW, Paul VJ, 2012. Climate change: links to global expansion of harmful cyanobacteria. Water Res. 46 (5), 1349–1363. [PubMed: 21893330]

Paerl HW, Hall NS, Calandrino ES, 2011a. Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. Sci. Total Environ 409 (10), 1739–1745. [PubMed: 21345482]

Paerl HW, Xu H, McCarthy MJ, Zhu G, Qin B, Li Y, Gardner WS, 2011b. Controlling harmful cyanobacterial blooms in a hyper-eutrophic lake (Lake Taihu, China): the need for a dual nutrient (N & P) management strategy. Water Res. 45 (5), 1973–1983. [PubMed: 20934736]

Paerl HW, Gardner WS, Havens KE, Joyner AR, McCarthy MJ, Newell SE, Qin B, Scott JT, 2016. Mitigating cyanobacterial harmful algal blooms in aquatic ecosystems impacted by climate change and anthropogenic nutrients. Harmful Algae 54, 213–222. [PubMed: 28073478]

Paerl HW, Otten TG, Kudela R, 2018. Mitigating the expansion of harmful algal blooms across the freshwater-to-marine continuum. Environ. Sci. Technol 52 (10), 5519–5529. [PubMed: 29656639]

Park Y, Pyo J, Kwon YS, Cha Y, Lee H, Kang T, Cho KH, 2017. Evaluating physicochemical influences on cyanobacterial blooms using hyperspectral images in inland water, Korea. Water Res. 126, 319–328. [PubMed: 28965034]

Paul VJ, 2008. Global warming and cyanobacterial harmful algal blooms. In: Hudnell HK (Ed.), Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs (Advances in Experimental Medicine and Biology). Springer, New York, New York, USA, pp. 239–257.

Pearson L, Mihali T, Moffitt M, Kellmann R, Neilan B, 2010. On the chemistry, toxicology and genetics of the cyanobacterial toxins, microcystin, nodularin, saxitoxin and cylindrospermopsin. Mar. Drugs 8 (5), 1650–1680. [PubMed: 20559491]

Pearson LA, Barrow KD, Neilan BA, 2007. Characterization of the 2-hydroxy-acid dehydrogenase McyI, encoded within the microcystin biosynthesis gene cluster of Microcystis aeruginosa PCC7806. J. Biol. Chem 282 (7), 4681–4692. [PubMed: 17142460]

Pearson LA, Crosbie ND, Neilan BA, 2019. Distribution and conservation of known secondary metabolite biosynthesis gene clusters in the genomes of geographically diverse Microcystis aeruginosa strains. Mar. Freshw. Res 71 (5), 701–716.

Pelley J, 2016. Taming toxic algae blooms. ACS Cent. Sci 2 (5), 270–273. [PubMed: 27280160]

Pereira N, Shilova IN, Zehr JP, 2019. Use of the high-affinity phosphate transporter gene, pstS, as an indicator for phosphorus stress in the marine diazotroph Crocosphaera watsonii (Chroococcales, Cyanobacteria). J. Phycol 55 (4), 752–761. [PubMed: 30929262]

Pineda-Mendoza RM, Zúñiga G, Martínez-Jerónimo F, 2016. Microcystin production in Microcystis aeruginosa: effect of type of strain, environmental factors, nutrient concentrations, and N: P ratio on mcyA gene expression. Aquat. Ecol 50 (1), 103–119.

Purkhold U, Pommerening-Röser A, Juretschko S, Schmid MC, Koops H-P, Wagner M, 2000. Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA and amoA sequence analysis: implications for molecular diversity surveys. Appl. Environ. Microbiol 66 (12), 5368–5382. [PubMed: 11097916]

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ralston DK,Moore SK, 2020. Modeling harmful algal blooms in a changing climate. Harmful Algae 91, 101729. [PubMed: 32057346]

Reynolds C, Rogers D, 1976. Seasonal variations in the vertical distribution and buoyancy of Microcystis aeruginosa Kütz. Emend. Elenkin in Rostherne Mere, England. Hydrobiologia 48, 17–23.

Rittmann BE, McCarty PL, 2020. Environmental Biotechnology: Principles and Applications. 2nd end. McGraw-Hill, New York, New York, USA.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M, 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12, 77. [PubMed: 21414208]

Romero L, Mondardo R, Sens M, Grischek T, 2014. Removal of cyanobacteria and cyanotoxins during lake bank filtration at Lagoa do Peri, Brazil. Clean Technol. Environ. Policy 16 (6), 1133–1143.

Rouhiainen L, Vakkilainen T, Siemer BL, Buikema W, Haselkorn R, Sivonen K, 2004. Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium Anabaena strain 90. Appl. Environ. Microbiol 70 (2), 686–692. [PubMed: 14766543]

Sabart M, Crenn K, Perrière F, Abila A, Leremboure M, Colombet J, Jousse C, Latour D, 2015. Co-occurrence of microcystin and anatoxin-a in the freshwater lake Aydat (France): analytical and molecular approaches during a three-year survey. Harmful Algae 48, 12–20. [PubMed: 29724471]

Schmale DG III, Ault AP, Saad W, Scott DT,Westrick JA, 2019. Perspectives on harmful algal blooms (HABs) and the cyberbiosecurity of freshwater systems. Front. Bioeng. Biotechnol 7, 128. [PubMed: 31231642]

Seltenrich N, 2014. Keeping tabs on HABs: new tools for detecting, monitoring, and preventing harmful algal blooms. Environ. Health Perspect 122 (8), A206–A213. [PubMed: 25229077]

Sevilla E, Martin-Luna B, Vela L, Bes MT, Fillat MF, Peleato ML, 2008. Iron availability affects mcyD expression and microcystin-LR synthesis in Microcystis aeruginosa PCC7806. Environ. Microbiol 10 (10), 2476–2483. [PubMed: 18647335]

Singh S, Rai PK, Chau R, Ravi AK, Neilan BA, Asthana RK, 2015. Temporal variations in microcystin-producing cells and microcystin concentrations in two fresh water ponds. Water Res. 69, 131–142. [PubMed: 25463934]

Song J, Bi H, Cai Z, Cheng X, He Y, Benfield MC, Fan C, 2020. Early warning of Noctiluca scintillans blooms using in-situ plankton imaging system: an example from Dapeng Bay, P.R. China. Ecol. Indic 112, 106123.

Steffen MM, Davis TW, McKay RML, Bullerjahn GS, Krausfeldt LE, Stough JM, Neitzey ML, Gilbert NE, Boyer GL, Johengen TH, 2017. Ecophysiological examination of the Lake Erie Microcystis bloom in 2014: linkages between biology and the water supply shutdown of Toledo, OH. Environ. Sci. Technol 51 (12), 6745–6755. [PubMed: 28535339]

Stüken A, Orr RJ, Kellmann R, Murray SA, Neilan BA, Jakobsen KS, 2011. Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. PLOS ONE 6 (5), e20096. [PubMed: 21625593]

Su M, Yu J, Zhang J, Chen H, An W, Vogt RD, Andersen T, Jia D, Wang J, Yang M, 2015. MIB-producing cyanobacteria (Planktothrix sp.) in a drinking water reservoir: distribution and odor producing potential. Water Res. 68, 444–453. [PubMed: 25462751]

Sukenik A, Viner-Mozzini Y, Tavassi M, Nir S, 2017. Removal of cyanobacteria and cyanotoxins from lake water by composites of bentonite with micelles of the cation octadecyltrimethyl ammonium (ODTMA). Water Res. 120, 165–173. [PubMed: 28486167]

Tanabe Y, Sano T, Kasai F,Watanabe MM, 2009. Recombination, cryptic clades and neutral molecular divergence of the microcystin synthetase (mcy) genes of toxic cyanobacterium Microcystis aeruginosa. BMC Evol. Biol 9, 115. [PubMed: 19463155]

Tanvir RU, Hu Z, Zhang Y, Lu J, 2021. Cyanobacterial community succession and associated cyanotoxin production in hypereutrophic and eutrophic freshwaters. Environ. Pollut 290, 118056. [PubMed: 34488165]

Tian Y, Huang M, 2019. An integrated web-based systemfor themonitoring and forecasting of coastal harmful algae blooms: application to Shenzhen city, China. J. Mar. Sci. Eng 7 (9), 314.

Tillett D, Dittmann E, Erhard M, Von Döhren H, Börner T, Neilan BA, 2000. Structural organization of microcystin biosynthesis inMicrocystis aeruginosa PCC7806: an integrated peptide–polyketide synthetase system. Chem. Biol 7 (10), 753–764. [PubMed: 11033079]

Tonk L, Visser PM, Christiansen G, Dittmann E, Snelder EO, Wiedner C, Mur LR, Huisman J, 2005. The microcystin composition of the cyanobacterium Planktothrix agardhii changes toward a

more toxic variant with increasing light intensity. Appl. Environ. Microbiol 71 (9), 5177–5181. [PubMed: 16151102]

Tooming-Klunderud A, Fewer DP, Rohrlack T, Jokela J, Rouhiainen L, Sivonen K, Kristensen T, Jakobsen KS, 2008. Evidence for positive selection acting on microcystin synthetase adenylation domains in three cyanobacterial genera. BMC Evol. Biol 8, 256. [PubMed: 18808704]

US EPA, 2015. Drinking Water Health Advisory for the Cyanobacterial Microcystin Toxins (EPA Document Number: 820R15100), Washington, DC, USA.

US EPA, 2019. Recommended Human Health Recreational AmbientWater Quality Criteria or Swimming Advisories for Microcystins and Cylindrospermopsin (EPA Document Number: 822-R-19-001), Washington, DC, USA.

Vu HP, Nguyen LN, Zdarta J, Nga TT, Nghiem LD, 2020. Blue-green algae in surface water: problems and opportunities. Curr. Pollut. Rep 6 (2), 105–122.

Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JA, Jansson JK, Caporaso JG, Fuhrman JA, 2016. Improved bacterial 16S rRNA gene (V4 and V4–5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. mSystems 1 (1), e00009–15.

Wang H, Zhu R, Zhang J, Ni L, Shen H, Xie P, 2018. A novel and convenient method for early warning of algal cell density by chlorophyll fluorescence parameters and its application in a highland lake. Front. Plant Sci 9, 869. [PubMed: 30002664]

Watson SB, Whitton BA, Higgins SN, Paerl HW, Brooks BW,Wehr JD, 2014. Harmful algal blooms. In: Wehr JD, Sheath RG, Kociolek JP (Eds.), Freshwater Algae of North America: Ecology and Classification (Aquatic Ecology), 2nd edn. Elsevier Inc. (Academic Press), San Diego, California, USA, pp. 873–920.

Wei T, Simko V, 2021. R Package ‘corrplot’: Visualization of a Correlation Matrix Version 0.88.

World Health Organization, 2003. Guidelines for Safe Recreational Water Environments. Coastal and Fresh Waters. 1. World Health Organization Geneva, Switzerland.

Wynne TT, Stumpf RP, 2015. Spatial and temporal patterns in the seasonal distribution of toxic cyanobacteria in Western Lake Erie from 2002–2014. Toxins (Basel) 7 (5), 1649–1663. [PubMed: 25985390]

Wynne TT, Stumpf RP, Tomlinson MC, Fahnenstiel GL, Dyble J, Schwab DJ, Joshi SJ, 2013. Evolution of a cyanobacterial bloom forecast system in western Lake Erie: development and initial evaluation. J. Great Lakes Res. 39, 90–99.

Xu H, Paerl HW, Zhu G, Qin B, Hall NS, Zhu M, 2017. Long-term nutrient trends and harmful cyanobacterial bloom potential in hypertrophic Lake Taihu, China. Hydrobiologia 787 (1), 229–242.

Xu H, Qin B, Paerl HW, Peng K, Zhang Q, Zhu G, Zhang Y, 2021. Environmental controls of harmful cyanobacterial blooms in Chinese inland waters. Harmful Algae 110, 102127. [PubMed: 34887007]

Yu Q, Liu Z, Chen Y, Zhu D, Li N, 2018. Modelling the impact of hydrodynamic turbulence on the competition between Microcystis and Chlorella for light. Ecol. Model 370, 50–58.

Zanchett G, Oliveira-Filho EC, 2013. Cyanobacteria and cyanotoxins: from impacts on aquatic ecosystems and human health to anticarcinogenic effects. Toxins 5 (10), 1896–1917. [PubMed: 24152991]

Zehr JP, 2011. Nitrogen fixation by marine cyanobacteria. Trends Microbiol. 19 (4), 162–173. [PubMed: 21227699]

Zhang C, Qin K, Struewing I, Buse HY, Santo Domingo J, Lytle D, Lu J, 2021a. The bacterial community diversity of bathroom hot tap water was significantly lower than that of cold tap and shower water. Front. Microbiol 12, 625324. [PubMed: 33967975]

Zhang C, Struewing I, Mistry JH, Wahman DG, Pressman J, Lu J, 2021b. Legionella and other opportunistic pathogens in full-scale chloraminated municipal drinking water distribution systems. Water Res. 205, 117571. [PubMed: 34628111]

Zhang M, Yang Z, Yu Y, Shi X, 2020. Interannual and seasonal shift between Microcystis and Dolichospermum: a 7-year investigation in Lake Chaohu, China. Water 12 (7), 1978.

Zhang W, Lou I, Ung WK, Kong Y, Mok KM, 2014a. Analysis of cylindrospermopsin and microcystin-producing genotypes and cyanotoxin concentrations in the Macau storage reservoir. Hydrobiologia 741, 51–68.

Zhang Y, Zhang S-F, Lin L, Wang D-Z, 2014b. Comparative transcriptome analysis of a toxin-producing dinoflagellate Alexandrium catenella and its non-toxicmutant. Mar. Drugs 12 (11), 5698–5718. [PubMed: 25421324]

Zhang Z, 2016. Variable selection with stepwise and best subset approaches. Ann. Transl. Med 4 (7), 136. [PubMed: 27162786]

Zhou Y, Li X, Xia Q, Dai R, 2020. Transcriptomic survey on the microcystins production and growth of Microcystis aeruginosa under nitrogen starvation. Sci. Total Environ 700, 134501. [PubMed: 31689655]

Zhu B, Cao H, Li G, Du W, Xu G, Santo Domingo J, Gu H, Xu N, Duan S, Lu J, 2019. Biodiversity and dynamics of cyanobacterial communities during blooms in temperate lake (Harsha Lake, Ohio, USA). Harmful Algae 82, 9–18. [PubMed: 30928013]

Zhu H, Zhang H, Xu Y, Laššáková S, Korabečná M, Neužil P, 2020. PCR past, present and future. BioTechniques 69 (4), 317–325. [PubMed: 32815744]

Zohdi E, Abbaspour M, 2019. Harmful algal blooms (red tide): a review of causes, impacts and approaches to monitoring and prediction. Int. J. Environ. Sci. Technol 16 (3), 1789–1806.

Duan X et al. Science of the Total Environment 830 (2022) 154568 [PubMed: 35302035]

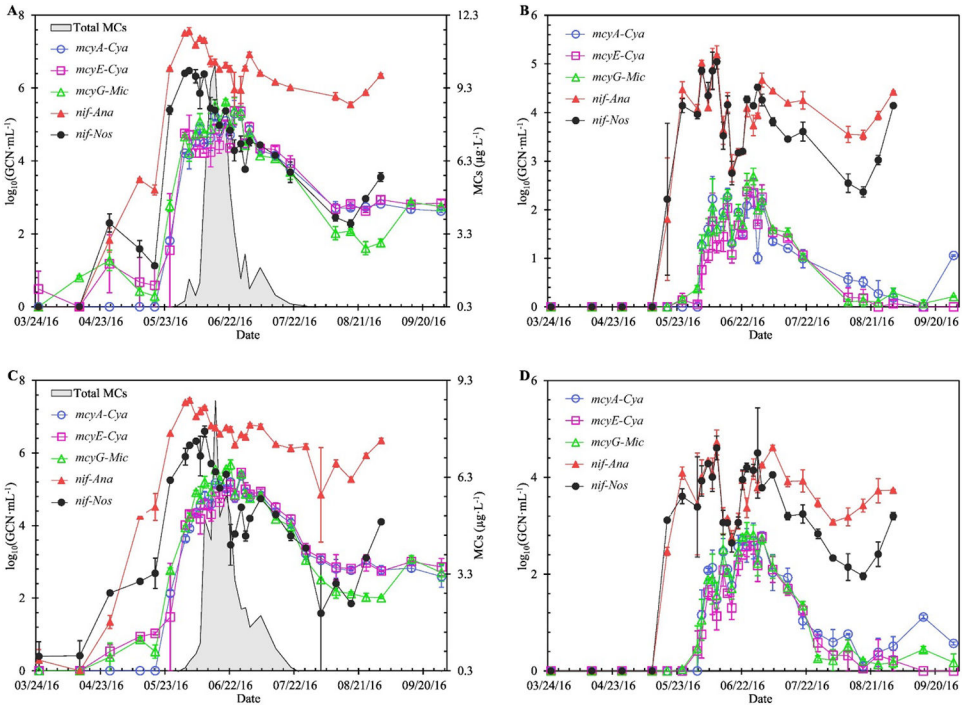


Fig. 1.

Abundances (**A** and **C**) and transcript levels (**B** and **D**) of *mcy* and *nif* genes compared with total microcystin concentrations (**A** and **C**) at BUOY (**A** and **B**) and EFLS (**C** and **D**) in Harsha Lake in 2016. **GCN**: Genome or gene copy number. The error bars represent the standard errors of the means for duplicate measurements. Date format: Month/Day/Year.

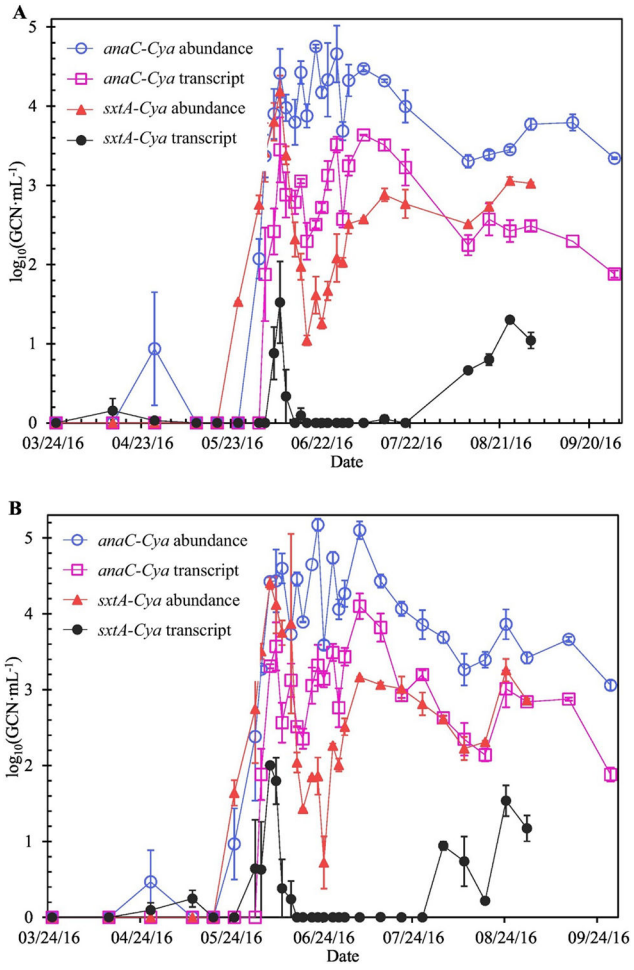


Fig. 2.

Abundances and transcript levels of *anaC* and *sxtA* at BUOY (**A**) and EFLS (**B**) in Harsha Lake in 2016. The error bars represent the standard errors of the means for duplicate measurements. **GC**N: Genome or gene copy number. Date format: Month/Day/Year.

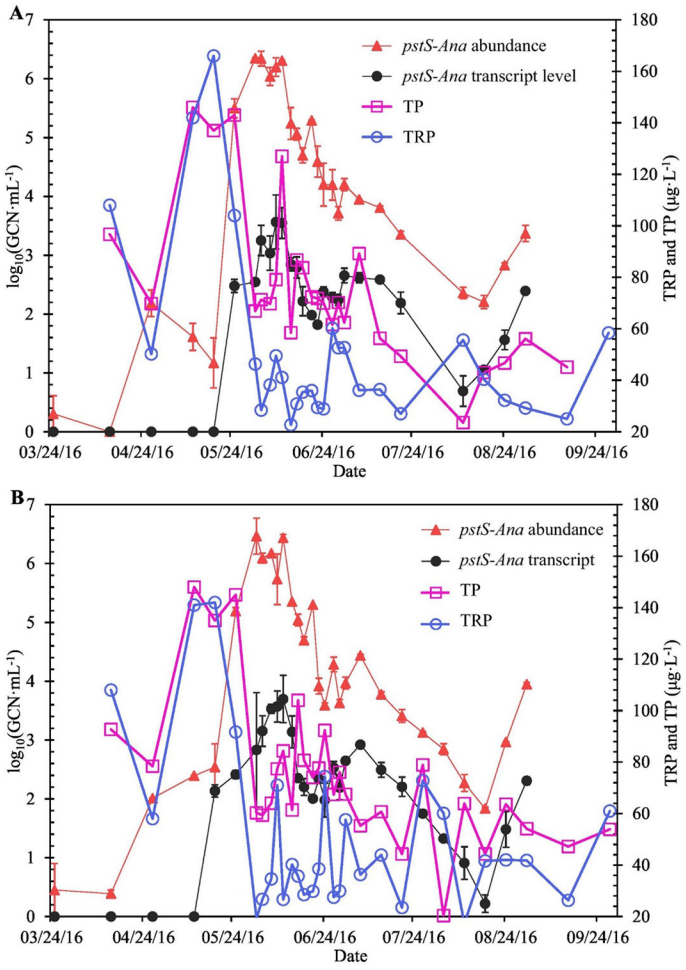


Fig. 3.

Dissolved total phosphorus (TP) concentrations, dissolved total reactive phosphorus (TRP) concentrations, *pstS-Ana* abundances, and *pstS-Ana* transcript levels at BUOY (**A**) and EFLS (**B**) in Harsha Lake in 2016. **GCN**: Genome or gene copy number. The error bars (for the abundances and transcript levels of *pstS-Ana*) represent the standard errors of the means for duplicate measurements. Date format: Month/Day/Year.

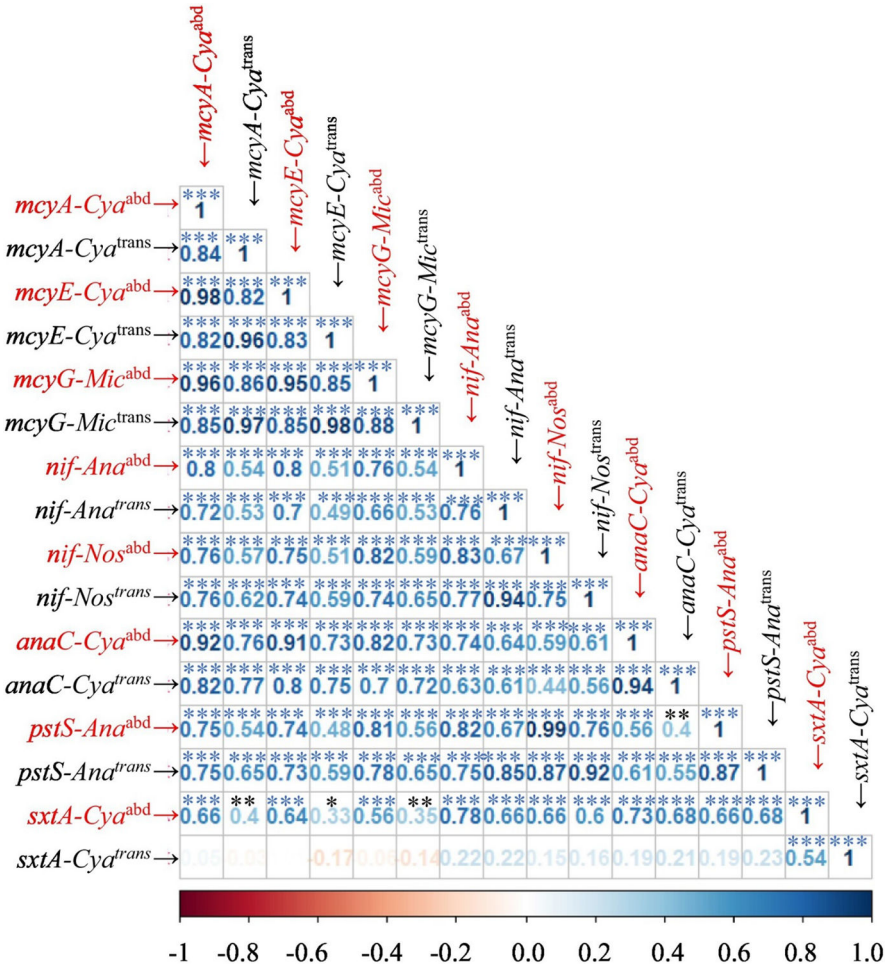


Fig. 4.

Pearson correlation coefficients between pairwise variables (i.e., gene abundances and transcript levels) for the 2016 harmful cyanobacterial bloom in Harsha Lake. The superscript ^{abd} indicates gene abundances [$\log_{10}$(GCN·mL^{-1})]. The superscript ^{trans} indicates transcript levels [$\log_{10}$(GCN·mL^{-1})]. **GCN**: Genome or gene copy number. The *, **, and *** on top of the Pearson correlation coefficients indicate that the *p*-values are in the ranges from 0.01 to 0.05, 0.001 to 0.01, and 0 to 0.001, respectively.

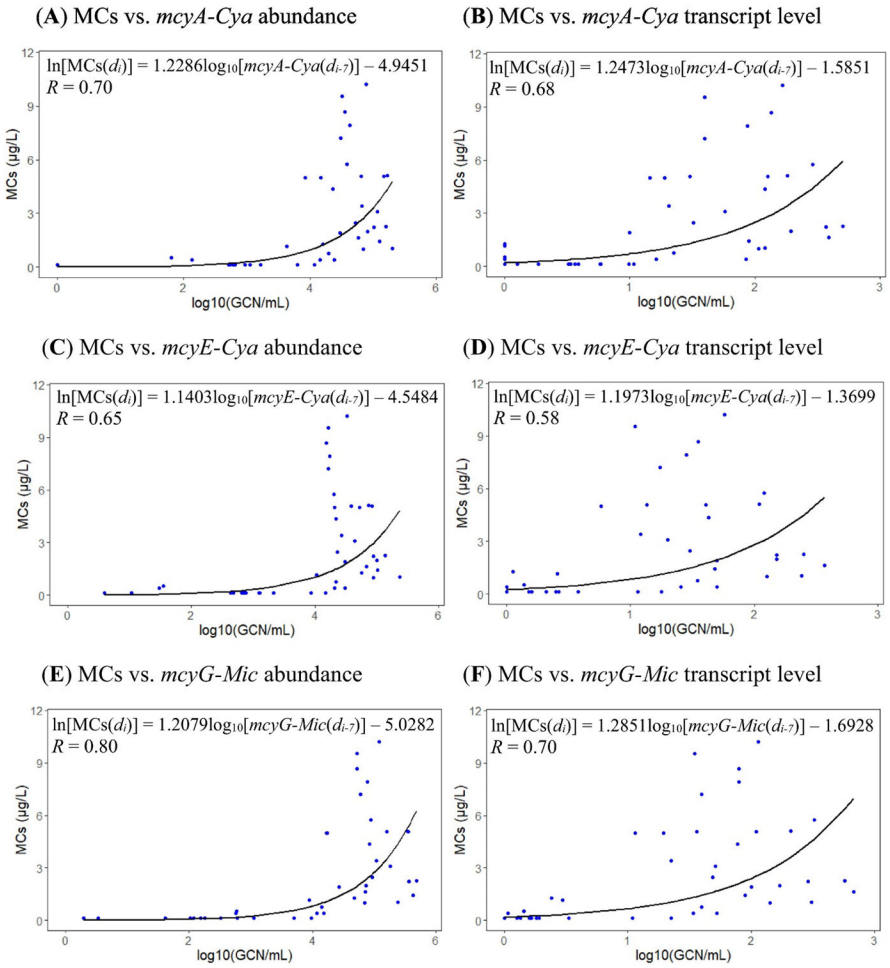


Fig. 5.

Prediction of 7-d lagged (d_i) total microcystin (MC) concentrations (left censored) using the current (d_{i-7}) abundances (**A**, **C**, and **E**) and transcript levels (**B**, **D**, and **F**) of *mcyA-Cya* (**A** and **B**), *mcyE-Cya* (**C** and **D**), and *mcyG-Mic* (**E** and **F**) in Harsa Lake in 2016. The *R*'s indicate the correlation coefficients between predicted and observed total MC concentrations. For each sub-figure, the sample size is 42.

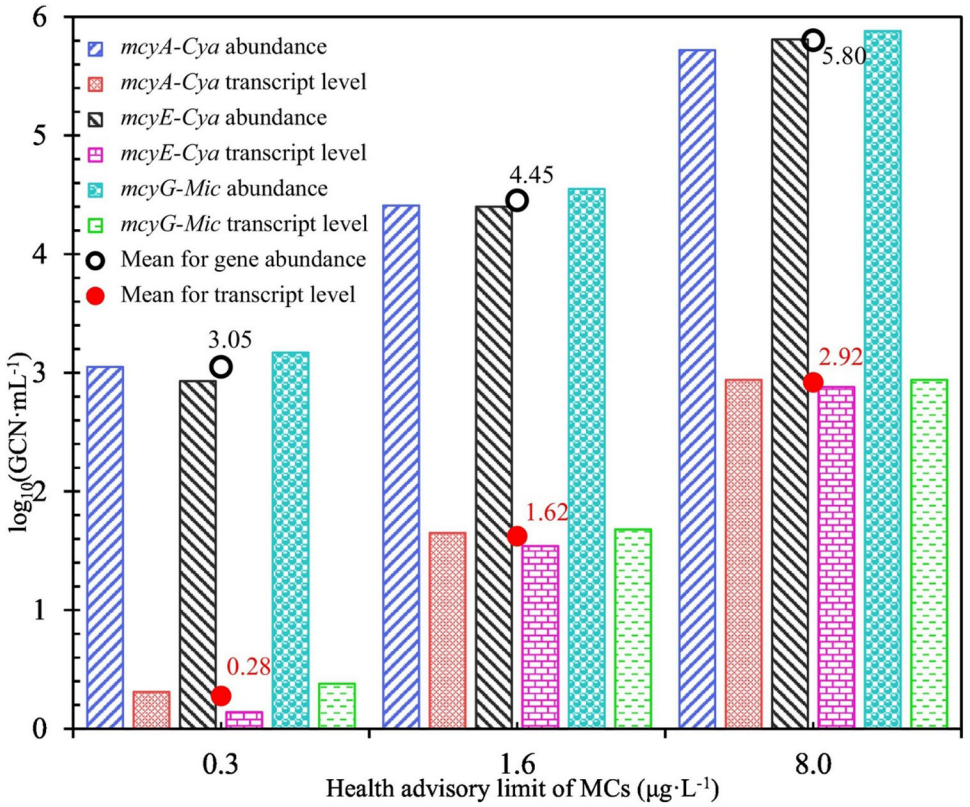


Fig. 6.

Abundances and transcript levels of *mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic* that predicted 7-d lagged total microcystin (MC) concentrations at health advisory limits of 0.3, 1.6, and 8.0 μg L^{-1} in Harsa Lake in 2016. **GCN**: Genome or gene copy number. The black open and red closed circles indicate the arithmetic means for the abundances and transcript levels, respectively, of the three *mcy* genes (*mcyA-Cya*, *mcyE-Cya*, and *mcyG-Mic*).

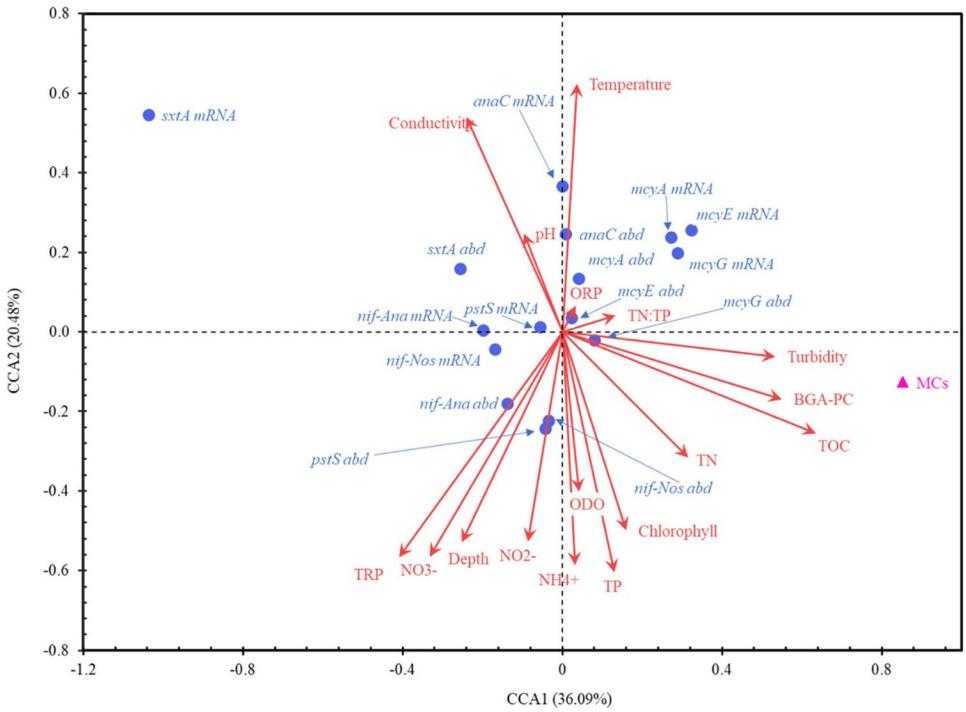


Fig. 7.

A Canonical correspondence analysis (CCA) map for the effects of physicochemical water quality parameters on the abundances and transcript levels of functional genes (*mcy*, *anaC*, *sxtA*, *nif*, and *pstS*) and total microcystin (MC) concentrations at BUOY in Harsha Lake in 2016.