Title: Sporadic Diurnal Fluctuations of Cyanobacterial Populations in Oligotrophic Temperate Systems Can Prevent Accurate Characterization of Change and Risk in Aquatic Systems

Authors: Ellen S. Cameron[a], Anjali Krishna[a], Monica B. Emelko[b], Kirsten M. Müller[a,*]

[a] Department of Biology, University of Waterloo, 200 University Ave. W, Waterloo, Ontario, Canada, N2L 3G1

[b] Department of Civil & Environmental Engineering, University of Waterloo, 200 University Ave. W, Waterloo, Ontario, Canada, N2L 3G1

*** Corresponding author. Tel.: +1 519 888 4567x32224**

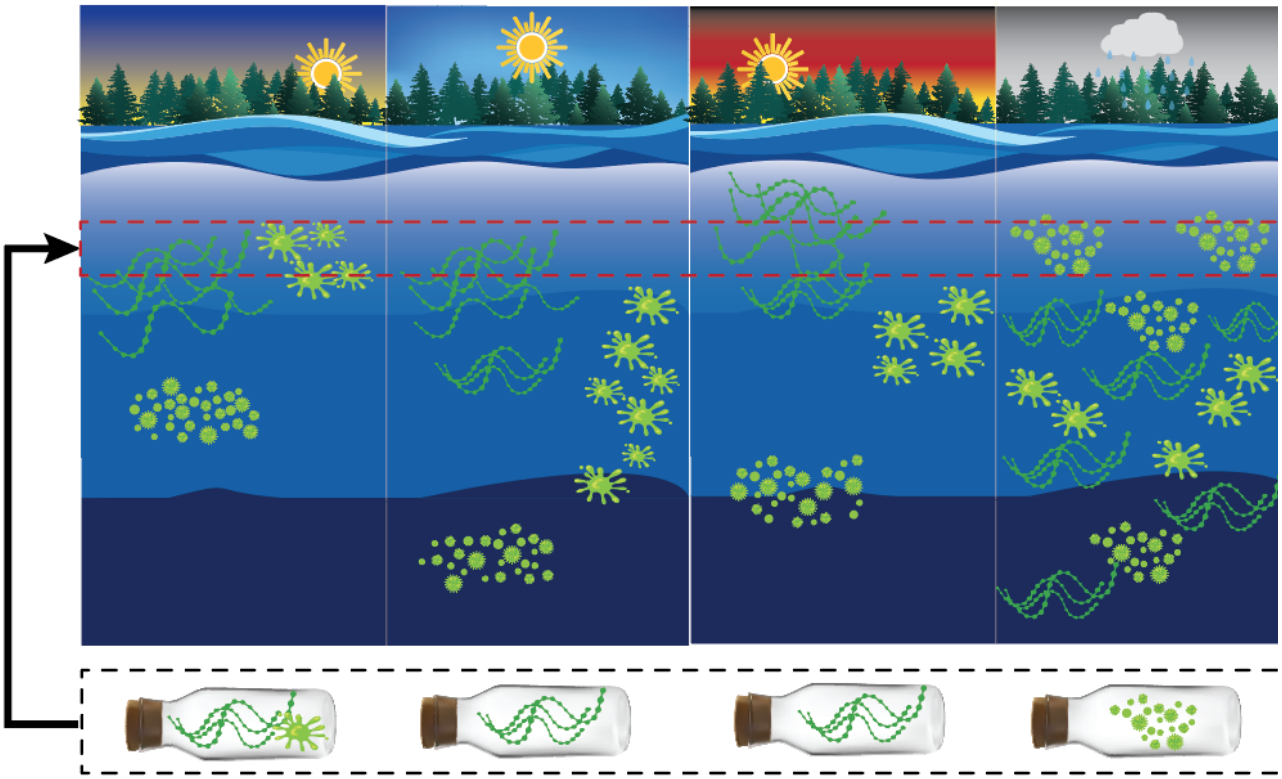
E-mail address: kirsten.muller@uwaterloo.ca (K.M. Müller).

Abstract

Continental-scale increases in aquatic system eutrophication are linked with increased cyanobacteria threats to recreational water use and drinking water resources globally. Increasing evidence suggests that diurnal vertical migration of cyanobacteria are key factors that must be considered in cyanobacterial bloom risk management. While this has been discussed in marine and eutrophic freshwater contexts, reports of diurnal vertical migration of cyanobacteria in oligotrophic freshwater lakes are scant**.** Typical monitoring protocols do not reflect these dynamics and frequently focus only on surface water sampling approaches, and either ignore sampling time or recommend large midday timeframes (e.g., 10AM-3PM), thereby preventing accurate characterization of cyanobacterial community dynamics. To evaluate the impact of diurnal migrations and water column stratification on cyanobacterial abundance and composition, communities were characterized in a shallow well-mixed lake interconnected to a thermally stratified lake in the Turkey Lakes Watershed (Ontario, Canada) using amplicon sequencing of the 16S rRNA gene across a multi-time point sampling series in 2018 and 2022. This work showed that cyanobacteria are present in oligotrophic lakes and their community structure varies (i) diurnally, (ii) across the depth of the water column, (iii) interannually within the same lake and (iv) between different lakes that are closely interconnected within the same watershed. It underscored the need for integrating multi-timepoint, multi-depth discrete sampling guidance into lake and reservoir monitoring programs to describe cyanobacteria community dynamics and signal change to inform risk management associated with the potential for cyanotoxin production. Ignoring variability in cyanobacterial community dynamics (such as that reported herein) and reducing sample numbers can lead to a false

sense of security and missed opportunities to identify and mitigate changes in trophic status and associated risks such as toxin or taste and odor production, especially in sensitive, oligotrophic systems.

Graphical Abstract



Keywords

Amplicon sequencing, diel cycles, vertical migration, buoyancy, algae, eutrophication

Highlights

- Cyanobacterial populations fluctuate sporadically across diurnal cycles
- Cyanobacterial communities can vary significantly between interconnected lakes
- Significant annual shifts in communities signal higher risk and need for monitoring
- Cyanobacteria monitoring for risk management should incorporate time and depth

1. Introduction

Cyanobacteria are recognized as a threat to surface water quality as they can form dense blooms and produce secondary metabolites including potent cyanotoxins (Harke et al., 2016; Huisman et al., 2018; Paerl, 2014; Vu et al., 2020) and, compounds associated with taste and odor problems (e.g., geosmin, 2-methyl isoborneol). As a result of anthropogenic activities, cyanobacterial bloom frequency and intensity have been increasing globally (Huisman et al., 2018) and climate change-exacerbated landscape disturbances (e.g., wildfires) further promote their proliferation (Emelko et al., 2016; Silins et al., 2014). These organisms can challenge water treatment processes beyond their response capacities (Emelko et al., 2011) and exacerbate disinfection by-product formation potential after landscape disturbances such as wildfires (Tsai et al., 2017). Accordingly, continued increases in toxic bloom occurrence globally (Huisman et al., 2018) underscore the need to advance source water protection and drinking water safety plan development guidance. Improved accuracy in characterization of cyanobacterial communities and their distributions which are underpinned by diurnal migrations will accelerate that advancement and associated risk management.

The diurnal migration of cyanobacterial populations is driven by cellular characteristics (Naselli-Flores et al., 2021) and water column stability (Walsby et al., 1997), as summarized in Table S1. Diurnal migration rates (e.g., flotation and/or sinking rates) and vertical distribution of populations may be distinctive to individual taxa arising from characteristic differences in cell sizes or the presence of specialized cellular structures impacting cellular density (Naselli-Flores et al., 2021; Reynolds et al., 1987). However, these characteristic distributions are further impacted by water column stability

(Walsby et al., 1997) with the potential for oscillatory diurnal variation in the distribution of cyanobacterial populations within systems of differing water column stability (Hunter et al., 2008). Rapid changes caused by oscillatory diurnal variation in the distribution of cyanobacterial populations can challenge detection if community structure and system dynamics are not well characterized.

Climate variability and landscape change-associated impacts on freshwater systems are key drivers of the need to better describe cyanobacterial growth and behavior in freshwater systems. Warming temperatures alter thermal stratification and mixing regimes in lakes (Woolway and Merchant, 2019) and increases in extreme precipitation and associated runoff increase nutrient delivery and eutrophication in receiving waters (Scholz and Brahney, 2022; Stoddard et al., 2016) and contribute to increases in cyanobacterial blooms, even in oligotrophic systems (Reinl et al., 2021; Sterner et al., 2020). Improved community characterization that reflects contemporary understanding of the diversity of cyanobacterial populations and their adaptations is therefore essential to managing risks attributable to the presence of these potentially toxic organisms in water supplies and better informing pre-emptive mitigation of potential health impacts and drinking water treatment challenges.

Key factors that must be considered in cyanobacterial bloom risk management include diurnal vertical migration of cyanobacteria, which has been reported in marine (Olli, 1999) and eutrophic freshwater (Hunter et al., 2008; von Orgies-Rutenberg et al., 2018) systems. More recently, advanced technologies including remote sensing techniques have been used to monitor cyanobacterial blooms evolution in eutrophic systems and has enabled high-resolution exploration of diurnal variability of visible

bloom biomass (Li et al., 2023; Wang et al., 2023; Xue et al., 2023). In contrast, reports of diurnal vertical migration of cyanobacteria in oligotrophic freshwater lakes are scant and have been limited to spectrophotometric analysis of chlorophyll-*a* and microscopic cell enumeration over periods of 24 hours or less (Frempong, 1981; Ganf, 1974; Gilbert et al., 2010; Ibelings et al., 1991; Olli, 1999; Shahraki et al., 2020). Despite next-generation sequencing and metagenomic approaches being used to characterize cyanobacterial communities in freshwater lakes (Valadez-Cano et al., 2022) and in the context of water treatment (Jalili et al., 2021; Le et al., 2022). Studies using next-generation sequencing technology to characterize diurnal variation of cyanobacteria remain limited in number and scope, however (Gilbert et al., 2010; Shahraki et al., 2020). While the use of multi-time point sampling series provides insights into the variability observed in community composition over periods of hours, investigations that use them do not typically encompass the natural dynamic variability that aquatic systems may experience day-to-day.

Here, we provide guidance for the development of cyanobacterial monitoring programs for risk management in oligotrophic freshwater systems, the impact of sampling time and depth on cyanobacterial community composition were evaluated using multi-time point sampling over consecutive days and amplicon sequencing of the V4 region in the 16S rRNA gene. Taxonomic composition and community diversity analyses were used to examine diurnal trends in cyanobacterial distribution and the associated impact of sampling time. Fluctuations in cyanobacterial community composition were evaluated (i) temporally over a 48-hour multi-time point sampling period and (ii) spatially within the water column of a shallow well-mixed lake and an interconnected,

thermally stratified lake. Temporal differences were further characterized in the thermally stratified lake during a second aquatic growing season.

2. Methods

2.1 Study Site: Turkey Lakes Watershed

The Turkey Lakes Watershed (TLW) Study was established in 1980 to investigate ecosystem effects of acidic atmospheric deposition; Jeffries et al. (1988) provide a comprehensive description of the physical characteristics of the watershed. In brief, it is approximately 50 km north of Sault Ste Marie, Ontario on the Canadian Shield in an uneven-aged tolerant hardwood and mixed conifer forest (Jeffries et al., 1988). This watershed consists of four interconnected lakes fed by both first order streams and groundwater: Batchawana Lake, Wishart Lake, Little Turkey Lake, and Big Turkey Lake (Jeffries et al., 1988; Figure S1). The shallow depth of Wishart Lake often results in complete wind-induced mixing that prevents water column stratification whereas Little Turkey Lake, Big Turkey Lake and Batchawana Lake are deeper and regularly undergo thermal stratification (Figure S2). These lakes are classified as oligotrophic to mesotrophic, and cyanobacteria have been reported as the dominant members of phytoplankton communities and previously reported as major contributors to peaks in primary production during summer periods (Jeffries et al., 1988).

2.2 Sample Collection

Water samples were collected from the deepest point in Little Turkey (47°02’37.2”N 84°24’24.4”W) and Wishart (47°03’00.0”N 84°23’58.3”W) Lakes during a two-day period across three time points on both days (2018a/b, 2022a/b): morning (8-9 a.m.), midday (12-1 p.m.) and afternoon (4-5 p.m.) during two sampling events (August

2018 and August 2022). Samples were collected from Secchi depth, the water depth at which light penetration is approximately 1% of surface illumination and is considered the maximum depth at which there is sufficient light for photosynthesis (Bukata et al., 1988). In Wishart Lake, these samples were collected from near the bottom (4 m) due to high water clarity. Secchi depths ranged from 4 to 5.25 m in Little Turkey Lake. Water samples were also collected at the surface (0 m) in Little Turkey Lake on the first sampling day in 2018 (2018a) and both sampling days in 2022 (2022a, 2022b) to evaluate potential temporal correlations in cyanobacterial communities between depths.

Water samples collected using a Masterflex E/S portable sampler peristaltic pump were vacuum filtered through a 47 mm GF/C filter (Whatman plc, Buckinghamshire, United Kingdom). For samples collected in August 2018, 250 mL of vacuum filtered water were filtered a second time through a 0.22 µm Sterivex™ filter to collect additional microbes. All filters were stored at -20°C prior to DNA extraction. Sampling details are provided in Table S2.

2.3 DNA Extraction, V4 16S rRNA Gene Amplicon Sequencing

DNA was extracted using the DNeasy PowerSoil Kit (QIAGEN Inc., Venlo, Netherlands) following the manufacturer's protocol. Following this, elution buffer was added to spin columns for 15 minutes prior to elution. DNA was quantified using a NanoDrop spectrophotometer (Table S2); absolute values were only accurate at DNA concentrations of more than 10ng/µl. The DNA extracts were submitted to a commercial laboratory (Metagenom Bio Inc.,Waterloo, ON) for amplicon sequencing. Primers designed to target the 16S rRNA gene V4 region [515FB (GTGYCAGCMGCCGCGGTAA) and 806RB (GGACTACNVGGGTWTCTAAT)]

(Walters et al., 2015) were used for PCR amplification. Polymerase chain reactions (PCR) were conducted in triplicate for each sample (25 µL). Reaction mixtures were composed of 2.5 µL of 10X *Taq* buffer, 0.5µL of 10mM dNTP, 0.25 µL of BSA (20mg/mL), 5.0 µL of 1 µM forward primer, 5.0 µL of 1 µM of reverse primer, 5.0 µL of sample DNA, 0.2 µL *Taq* DNA polymerase (5u/µL) and 6.55 µL of PCR water. DNA was initially denatured at 95°C for 5 minutes, followed by 35 cycles of: 1) 95°C for 30 seconds, 2) 30°C for 30 seconds, 3) 72°C for 50 seconds. A final extension was performed at 72°C for 10 minutes. Triplicate PCR products were pooled, gel purified and quantified using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific Inc.). Library DNA was prepared and sequenced with MiSeq Reagent Kit v2 (2x 250 cycles) and sequenced with the Illumina MiSeq platform (Illumina Inc., San Diego, United States).

2.4 Sequence Processing & Library Size Normalization

The program QIIME2 (v. 2019.10; Bolyen et al., 2019) was used for bioinformatic processing. Demultiplexed paired-end sequences were trimmed and denoised, including the removal of chimeric sequences and singleton sequence variants, using *DADA2* (Callahan et al., 2016). Amplicon sequence variants across the sample sets were clustered at 99% identity to account for single nucleotide discrepancies in sequenced amplicons using *vsearch* (Rognes et al., 2016) to generate a representative set of operational taxonomic units (OTUs). Taxonomic classification was performed using a Naïve-Bayes taxonomic classifier trained using the SILVA138 database (Quast et al., 2013; Yilmaz et al., 2014). Sequences classified as cyanobacteria at the phylum level were aligned to reference 16S rRNA cyanobacterial sequences in the Cyanoseq database (Lefler et al., 2023) using *MAFFT* (v7.490; Katoh and Standley, 2013). A phylogenetic

tree was generated using *iqtree* (v. 2.2.0; Nguyen et al., 2015) using an unrestricted model (UNREST+FO+R10) selected by the ModelFinder Bayesian Information Criterion and visualized with Interactive Tree of Life (iTOL; Letunic and Bork, 2021; Figure S3). Confirmed taxonomic classifications for OTUs classified as cyanobacteria at the phylum level were manually curated to reflect taxonomic levels above the genus level according to taxonomic labels in the Cyanoseq database based on placements in the phylogenetic tree.

Files from QIIME2 were imported into R (v. 4.0.1; R Core Team, 2020) for downstream analyses using *qiime2R* (v. 0.99.23; Bisanz, 2018). Initial sequence libraries were filtered to exclude OTUs that were taxonomically classified as mitochondria or chloroplast sequences using *phyloseq* (v. 1.32.0; McMurdie and Holmes, 2012). For cyanobacterial community analysis, OTUs classified as cyanobacteria at the phylum level were selected to create libraries consisting of only cyanobacterial sequences.

2.5 Cyanobacterial Communities - Taxonomic Composition & Diversity Analyses

Bacterial community composition was assessed at the phylum level and relative abundances were visualized using a stacked bar plot produced with *mirlyn* (Cameron and Tremblay, 2020). Relative abundances were randomized across bacterial phyla within samples to identify significantly dominant groups within the bacterial community in relation to sampling time and adjusted with a Bonferroni correction using the *randomseqsig* function in *mirlyn*. Cyanobacterial community composition was assessed through the selection of a subset of sequences classified as “cyanobacteria” at the phylum level and relative abundances of cyanobacterial families were also visualized using a stacked bar plot. For bacterial and cyanobacterial community composition bar plots,

groups with an average relative abundance lower than 1% across samples were collapsed into a low abundance category for visualization purposes.

As amplicon sequencing only indirectly and partially represents source diversity and clustering, amplification, and the lack of certainty about how many zeros should be included in the data compromise statistical inference about source diversity (Schmidt et al., 2021), community diversity analyses were performed using relative abundances instead of rarefying. While rarefying can be utilized to compare libraries in a statistically meaningful manner (Cameron et al., 2021), library sizes between samples collected in 2018 and 2022 differed by orders of magnitude creating extreme data loss of sequences or entire samples. OTUs present at >1% in the bacterial community were utilized for diversity analyses to reduce the risk of accidental non-detection. Alpha diversity was assessed using direct counts of unique OTUs and the Shannon Index (Shannon, 1948) to identify trends in sample diversity as a function of time. Bray-Curtis dissimilarity metric was calculated on compositional community data using *vegdist* in *vegan* (Oksanen et al., 2022) and used for permutational analysis of variance with *adonis2*. Bray-Curtis distances were visualized using ComplexHeatMap (Gu, 2022). To identify individual OTUs driving compositional changes, highly abundant (>10% of cyanobacterial community) were selected and normalized using Z-scores to identify enrichment or under-enrichment of individual abundant OTUs as a function of time.

2.6 Data Availability

Sequence data analyzed in this study is available in the European Nucleotide Archive (ENA) under study accession ERP134980.

3. Results

3.1 Diurnal Conservation of Bacterial & Cyanobacterial Communities

Cyanobacteria classified sequences were found to comprise a variable proportion of the bacterial community composition across sampling events in 2018 and 2022 (Figure 1A). Relative abundances of cyanobacteria were more similar during sampling events in Little Turkey Lake than Wishart Lake in 2018 (Figure 1A; Figure 2A). Cyanobacteria did not contribute to a significant portion of the bacterial community in all sampling events of 2018 ($p = 1$; Table S4). In contrast, cyanobacterial sequences comprised a significant portion of the bacterial community in Little Turkey Lake in 2022. Specifically, cyanobacteria comprised a significant portion of the bacterial community (i) at all sampling times at Secchi depth ($p < 0.005$) on both sampling days (2022a; 2022b) and (ii) in morning and midday surface samples on the second sampling day (2022b; $p < 0.003$).

Cyanobacterial communities were frequently dominated by sequences classified to the family Prochlorococcaceae with minimal diurnal fluctuations (Figure 1B; Figure 2B). Cyanobacterial communities observed in Little Turkey Lake in 2018 included large proportions of taxonomic families not observed in Wishart Lake or during the 2022 sampling events. A higher abundance of Cyanothrichaceae, Microcystaceae, and Trichocoleusaceae classified sequences were observed in the 2018 samples from Little Turkey Lake. Relative abundances of Cyanothrichaceae decreased from morning (11.1%) to afternoon (1.4%) at surface depths, possibly due to downward vertical migration. At Secchi depth on the first sampling day (2018a), Trichocoleusaceae were found in similar relative abundances (48.6%) in the morning sampling event but experienced a drastic

decrease by midday (10.1%) before a second increase at the final sampling time (19.8%). On the second sampling day (2018b) in Little Turkey Lake, Trichocoleusaceae relative abundances were relatively consistent from morning (19.0%) to midday (19.1%) and then increased in the afternoon (25%). Notably, the relative abundances observed in the morning sampling point at Secchi depth of the 2018a sampling event had higher similarity to those at the water surface than those at Secchi depth during subsequent sampling events. Fluctuations in Prochlorococcaceae relative abundances were negatively correlated with Trichocoleusaceae and Microcystaceae abundances due to the compositional constraint of amplicon sequencing data.

Similarity between cyanobacterial communities across sampling events was analyzed using the Bray-Curtis dissimilarity metric (Figure 3). Sampling time did not significantly impact cyanobacterial community composition in either of the study lakes ($p = 0.97$; Table S5) suggesting diurnal conservation of cyanobacterial community composition. However, significant compositional differences were observed between the two lakes sites ($p = 0.002$). The insignificance of sampling time and significance of sampling site were also observed for bacterial community composition ($p = 0.99$; $p = 0.001$; Table S6).

Within Little Turkey Lake, interannual variability between sampling events had a significant impact on cyanobacterial communities ($p = 0.001$). Sampling depth of the stratified water column in Little Turkey did not significantly impact the community structure ($p = 0.154$) and high similarity was still observed between samples collected at the same sampling depth. Additionally, similarity was observed within lakes across the sampling days of each year, except in the unique case of the morning of the first sampling

day of 2018 (2018a). Cyanobacterial communities at Secchi depth in the morning of the 2018a sampling event were highly similar to surface communities observed in Little Turkey Lake, a phenomenon that was not observed during other sampling events (Figure 3).

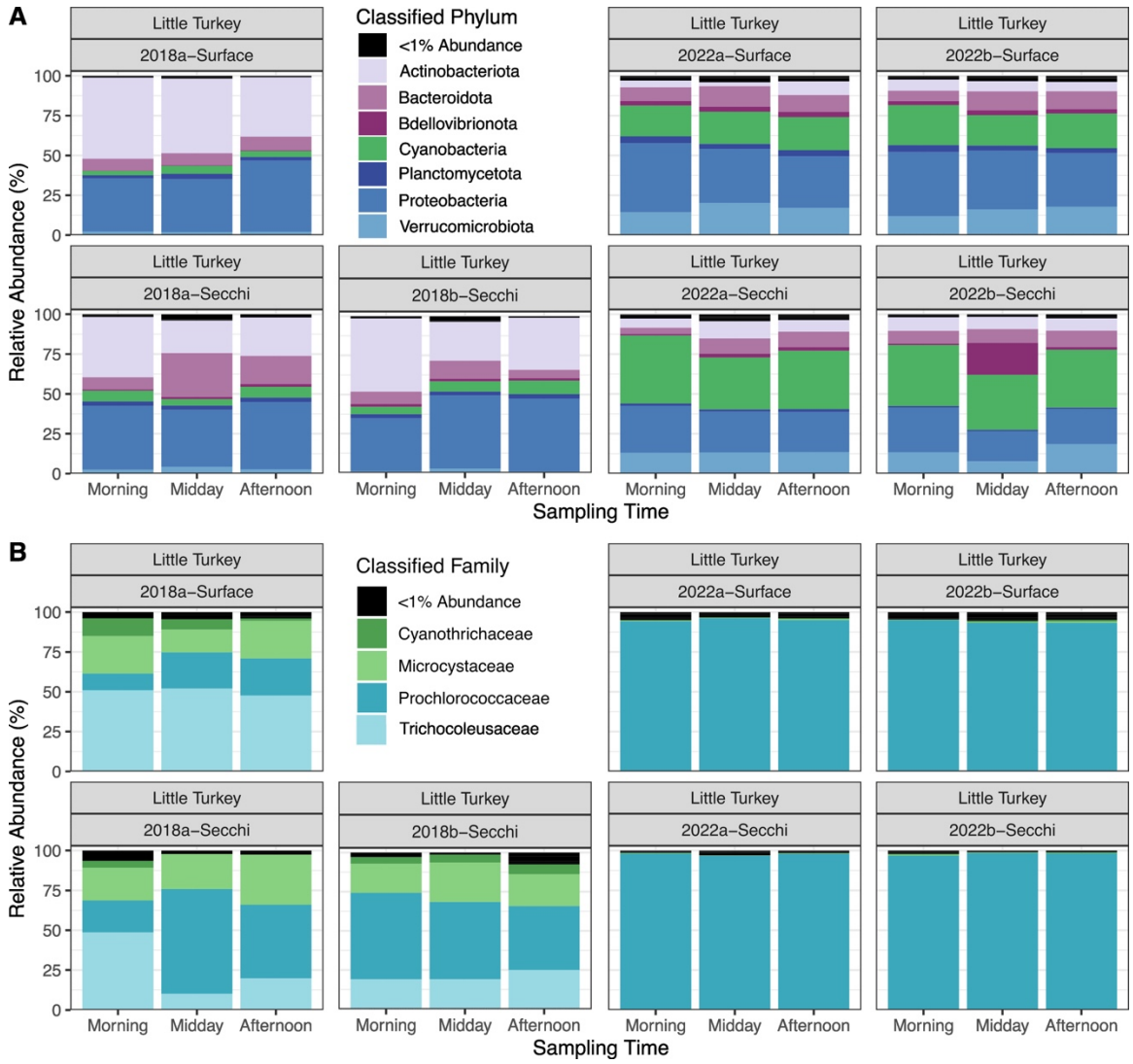


Figure 1: Stacked bar plots depicting the taxonomic composition of bacterial communities in terms of the relative abundance of major bacterial phyla (A) and cyanobacterial communities in terms of relative abundance of major cyanobacterial families (B) measured across a multi-time point sampling series over two consecutive

sampling days (2018a/2018b, 2022a/2022b) in a thermally stratified lake (Little Turkey Lake).

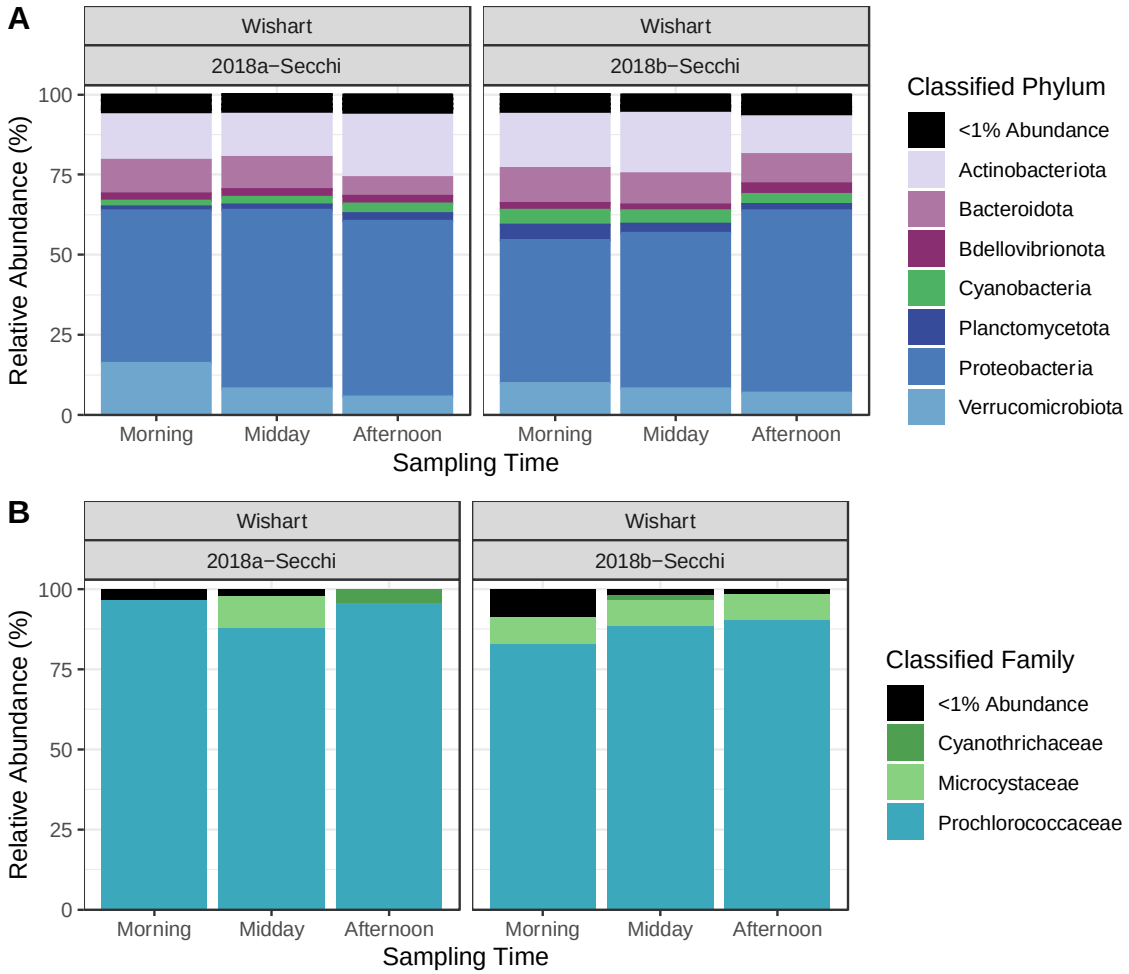


Figure 2: Stacked bar plots depicting the taxonomic composition of bacterial communities in terms of the relative abundance of major bacterial phyla (A) and cyanobacterial communities in terms of relative abundance of major cyanobacterial families (B) measured across a multi-time point sampling series over two consecutive sampling days (2018a/2018b) in a shallow, non-stratified lake (Wishart Lake).

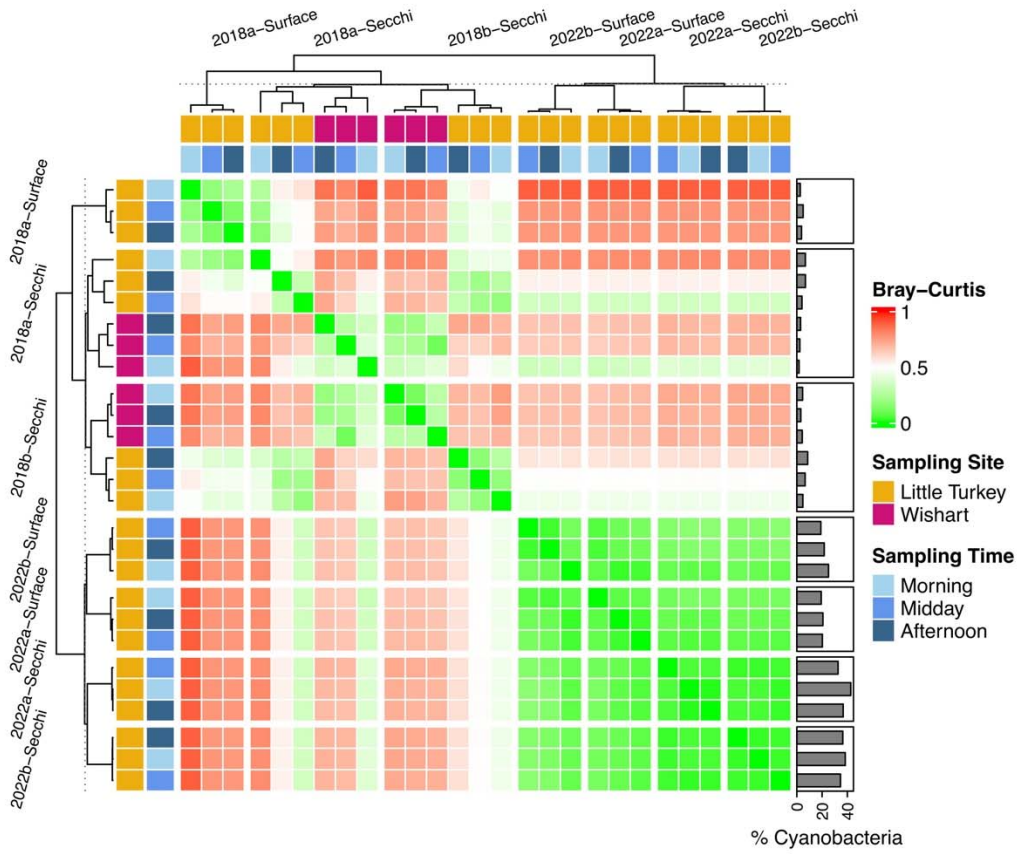


Figure 3: Compositional similarity of cyanobacterial communities across a multi-time point sampling series conducted over two consecutive sampling days (2018a/2018b, 2022a/2022b). Bray-Curtis dissimilarity metric (0 = completely identical, 1 = completely dissimilar) characterizes the differences present in the OTU composition of cyanobacterial communities. Overall proportion of cyanobacteria in bacterial communities displayed on the right-hand side highlights and contextualizes the interannual variability observed within Little Turkey Lake bacterial communities.

3.2 Sporadic Diurnal Variability in Relative Abundances of OTUs

Cyanobacterial community diversity across sampling times was assessed using (i) counts of unique OTUs (>1% abundance) and (ii) the Shannon Index (Figure 3A). Little Turkey Lake exhibited different trends in sample diversity dependent on the sampling event. Within samples collected during the 2018 sampling events, a higher number of cyanobacterial OTUs were detected in comparison to the 2022 sampling series. However, there was no consistent trend of diurnal fluctuations, revealing the sporadic fluctuations which may occur throughout a diurnal period. For example, 2018 sampling events at Secchi depth revealed opposing trends at midday where diversity scores decrease during the 2018a sampling event but increase during the 2018b sampling event. Similar results are also observed during the 2022 sampling events. The 2022b sampling event also reveals a consistent OTU count across sampling times, but the Shannon Index suggests changes in proportionality of these detected OTUs across time. Although results from Wishart Lake are limited to the 2018 sampling events, contrary to the sporadic diurnal fluctuations of diversity observed in Little Turkey Lake, Wishart Lake was shown to have corresponding spikes in diversity at the midday sampling point on both sampling days.

Diurnal fluctuations of highly abundant cyanobacterial OTUs (>10%) were identified using Z-scores; time points associated with a higher Z-score reflect greater deviation from the mean (Figure 3B). Additionally, temporal distribution of cyanobacterial OTUs was visualized to identify ubiquitous OTUs within cyanobacterial communities (Figure 4). Specifically, OTU5755 (Prochlorococcaceae) was found in all samples independent of sampling site, sampling year or sampling time. Diurnal

fluctuations in the occurrence of OTU5755 could not be discerned based on the Z scores; both enrichment and de-enrichment were observed. Absence of diurnal linked fluctuations was also observed for other highly independent OTUs including those from families which may be associated with gas vacuolate taxa (e.g., Trichocoleusaceae – OTU5691). Distribution of OTUs across sampling events indicated that the predominant populations varied temporally (Figure 4). While OTU5691 and OTU5695, both classified to the family Trichocoleusaceae and comprised a significant portion of cyanobacterial communities in 2018, they were rarely detected and detected at substantially lower relative abundances in the 2022 sampling events. 2022 sampling events were found to have a higher count of unique Prochlorococcaceae OTUs than those detected in 2018.

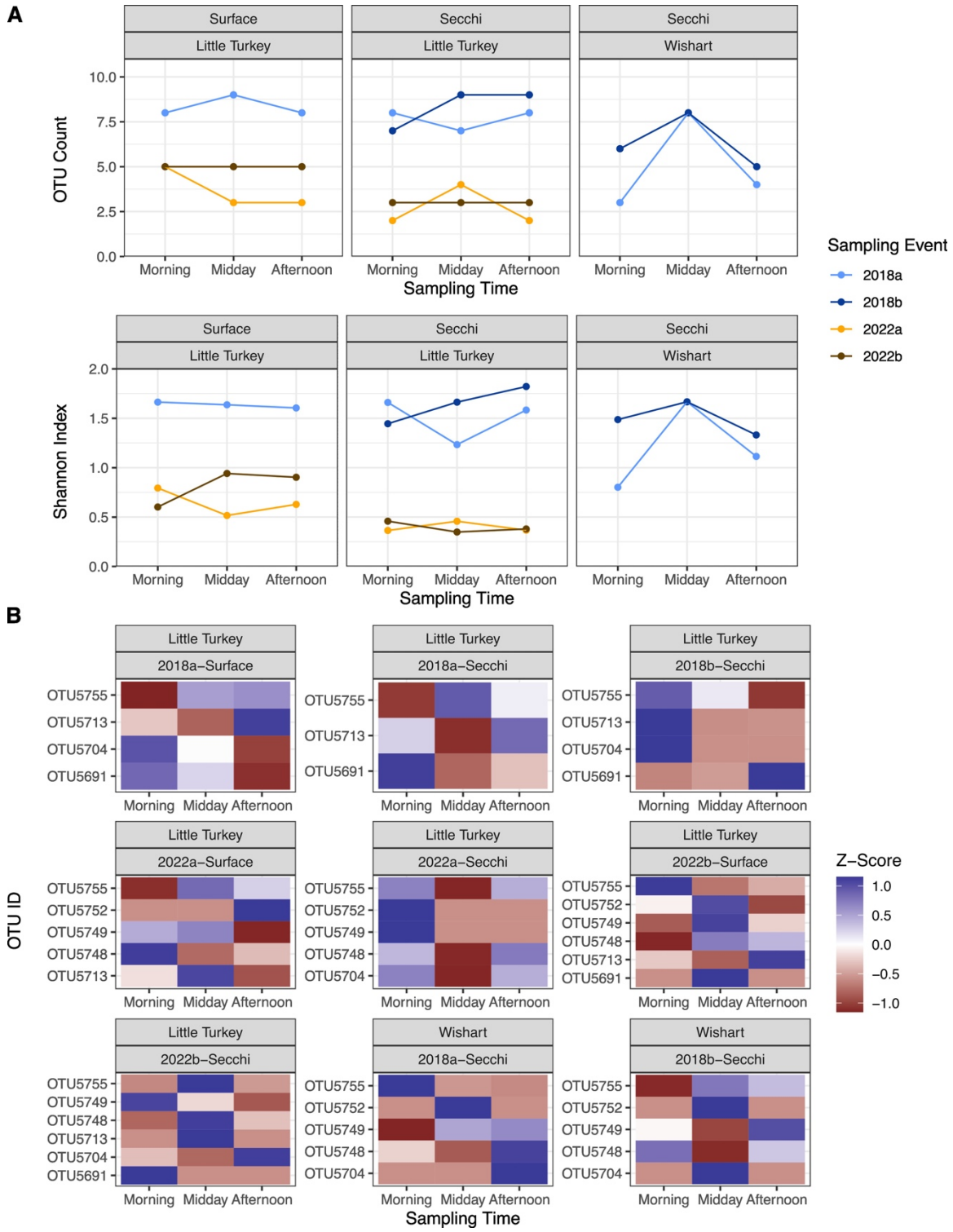
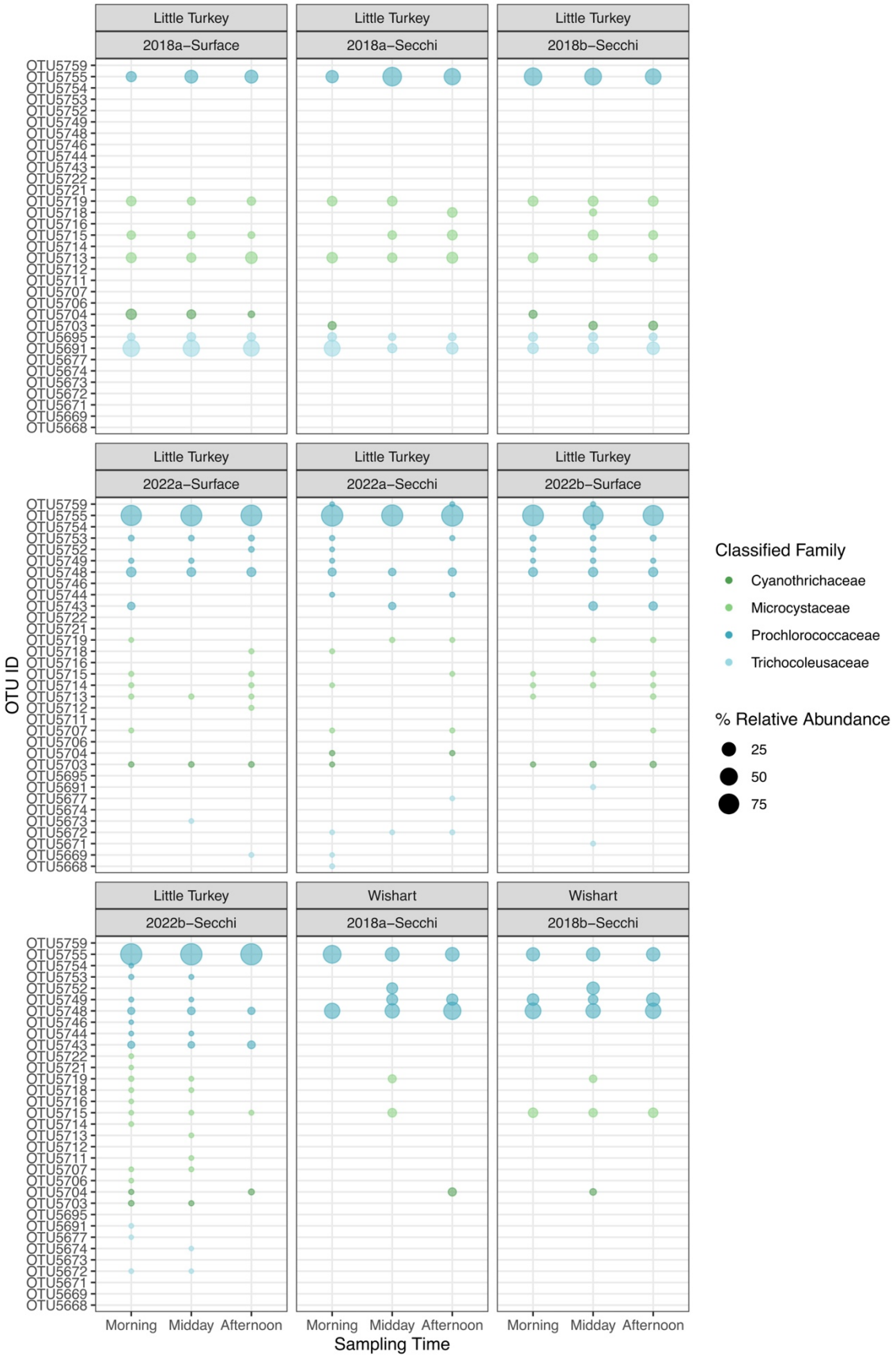


Figure 3: Temporal Fluctuations in alpha-diversity metrics of cyanobacterial communities driven by changes in OTU detection (OTU counts; A) and composition

(Shannon Index; B). Z-score normalized relative abundances for highly abundant operational taxonomic units (>10% of cyanobacterial community) for sampling event to demonstrate diurnal trends of specific cyanobacterial populations (C).


Little Turkey
2018a–Surface
Little Turkey
2018a–Secchi
Little Turkey
2018b–Secchi
Little Turkey
2022a–Surface
Little Turkey
2022a–Secchi
Little Turkey
2022b–Surface
Little Turkey
2022b–Secchi
Wishart
2018a–Secchi
Wishart
2018b–Secchi
OTU ID
OTU5759
OTU5755
OTU5754
OTU5753
OTU5752
OTU5749
OTU5748
OTU5746
OTU5744
OTU5743
OTU5722
OTU5721
OTU5719
OTU5718
OTU5716
OTU5715
OTU5714
OTU5713
OTU5712
OTU5711
OTU5707
OTU5706
OTU5704
OTU5703
OTU5695
OTU5691
OTU5677
OTU5674
OTU5673
OTU5672
OTU5671
OTU5669
OTU5668
Morning
Midday
Afternoon
Sampling Time
Classified Family
Cyanothrichaceae
Microcystaceae
Prochlorococcaceae
Trichocoleusaceae
% Relative Abundance
25
50
75

340

Figure 4: Temporal distribution of individual OTUs of cyanobacterial families present in >1% average relative abundances coloured based on family level taxonomic classification. Sizes of points represent relative abundance of OTU in cyanobacterial community.

4. Discussion

Diurnal migrations of cyanobacterial populations have been observed in marine and eutrophic freshwater systems and attributed to both cellular characteristics (Naselli-Flores et al., 2021) and water column stability (Walsby et al., 1997). Oscillatory diurnal variation in the distribution of cyanobacterial populations would thus be expected across depth in the water column. In contrast to trends frequently observed in marine and eutrophic freshwater systems, we demonstrate that in highly oligotrophic systems, diurnal fluctuations of cyanobacterial populations are sporadic, but overall bacterial community composition is conserved temporally. These observations contradict the patterns commonly described for diurnal migrations of gas vacuolate taxa. Consequently, these contrasting results—especially the unpredictable nature of cyanobacterial population distributions—have significant implications for the development of water quality monitoring programs and associated risk management strategies, including source water protection and drinking water safety plans.

Evidence of diurnal migrations of colonial and filamentous cyanobacteria has been reported previously; larger colonies have been shown to migrate across greater depths more rapidly than smaller ones (Reynolds et al., 1987; Westwood and Ganf, 2004); they also experience faster sinking rates (Ganf, 1974). In both Little Turkey and Wishart Lakes, cyanobacterial communities were predominantly composed of OTUs

classified to the family Prochlorococcaceae. Notably, the family Prochlorococcaceae includes small unicellular and colonial taxa that may facilitate their ability to maintain water column position (Śliwińska-Wilczewska et al., 2018). Furthermore, smaller colonies also experience more random spatial movement (Chien et al., 2013). The dominance of small unicellular and colonial taxa of the Prochlorococcaceae family concurrent with the absence of visible cyanobacterial biomass may thus help to explain the absence of discernible diurnal trends in the studied cyanobacterial OTUs.

Previous studies have demonstrated that thermal stratification of the water column plays a significant role in shaping the distribution of cyanobacteria and other phytoplankton through the creation of discrete zones in the water column (*i.e.*, nutrient depleted, light rich surface waters *vs.* nutrient rich, light limited deep waters; Reynolds et al., 1987; Spigel and Imberger, 1987). To investigate the impact of thermal stratification on cyanobacterial communities in an oligotrophic lake (*i.e*., Little Turkey Lake), samples were collected in both the epilimnion (water surface) and metalimnion (Secchi depth) zones (Figure S2). During both sampling events this analysis revealed common OTUs over time and across depth, thereby suggesting a broader distribution of cyanobacterial species across the water column rather than restriction to specific zones. In 2022, the composition of the cyanobacterial communities was highly similar regardless of spatio-temporal differences. In contrast, in 2018, cyanobacterial communities at the water surface and Secchi depth were distinguishable, consistent with previous spatial distribution research conducted in Little Turkey Lake (Cameron et al., 2022). On both occasions, sampling time did not significantly impact the cyanobacterial community

composition, further supporting the hypothesis that small unicellular and colonial cyanobacterial populations will exhibit random fluctuations.

Notably, cyanobacterial communities of Little Turkey Lake in 2018 and 2022 did not exhibit similarity. This result is consistent with seasonal studies conducted in consecutive years (*i.e.*, 2018-2020) in Little Turkey Lake and that reported (i) similar cyanobacterial community composition between sampling years, and (ii) variability in the relative abundance of cyanobacteria within the broader bacterial community (Cameron et al., 2022). Here, cyanobacteria were a statistically significant portion of the bacterial community in 2022, but not 2018, underscoring the potential for significant interannual variability in bacterial community structure and the need for further investigation into environmental drivers of these differences in oligotrophic systems.

Critically, the increased availability and rapidly decreasing costs of next-generation sequencing tools have eliminated most of the barriers to their wider use in resource management and the drinking water industry. The insights enabled by these approaches allow for proactive rather than reactive management of drinking water supplies (Chapman, 2010). Although amplicon sequencing has revolutionized our ability to study microbial communities, these data do not quantify abundance of community members. Rather, they only indicate community composition and relative abundance (Gloor et al., 2017, 2016) which introduces challenges in the interpretation of changes in community structure. For example, an observed decrease in relative abundances of cyanobacteria may only be an artifact of increased relative abundances of sequences classified to other taxonomic groups and not reflective of an actual decrease in cyanobacterial populations. Thus, to best inform cyanobacterial community dynamics for

the purposes of resource and risk management, traditional cell enumeration techniques (e.g., flow cytometry; Patel et al., 2019) or other molecular approaches (e.g., qPCR; Chiu et al., 2017) that provide community density information are needed in conjunction with amplicon sequencing.

The sporadic diurnal variations in cyanobacterial populations and the differences between interconnected oligotrophic lakes that were reported herein emphasize the challenge in developing lake and reservoir monitoring protocols in absence of sufficient baseline knowledge of ecological community structures in the specific systems of interest. Here, cyanobacterial relative abundances fluctuated substantially between the hours of 9 a.m. and 4 p.m., demonstrating that large, generalized time frames are insufficient for characterizing cyanobacteria occurrence and dynamics in aquatic systems and associated threats to water quality and treatability. This work specifically demonstrated that following typical guidance that limits sample collection to a discrete depth in stratified lakes or a single timepoint may result in vast underestimation of cyanobacterial relative abundances. While cyanobacteria sampling protocols and guidelines do occasionally mention that time and depth of sample collection can affect outcomes, they are often vague and refer to collection of “surface waters” at times “later in the day” or from “10 a.m. to 3 p.m.” (Table S7), and do not address the potential for or implications of random non-diurnal migrations of smaller taxa. The present investigation fills a critical knowledge gap by reporting these insights for oligotrophic systems.

The design of sample collection schemes is closely linked to the utility of the information that is delivered using these approaches. The analysis provided herein has delivered clear evidence to demonstrate the need for integrating multi-timepoint, multi-depth discrete sampling guidance into lake and reservoir monitoring programs to describe cyanobacterial community dynamics and signal change to inform risk management associated with the potential for cyanotoxin production. This work further emphasized that cyanobacteria are present in oligotrophic lakes and their community structure varies (i) diurnally, (ii) across the depth of the water column, (iii) interannually within the same lake and (iv) between different lakes that are closely interconnected within the same watershed. Ignoring this variability and reducing sample numbers can lead to a false sense of security and missed opportunities to identify and mitigate changes in trophic status and associated risks such as toxin or taste and odor production, especially in sensitive, oligotrophic systems.

5. Conclusions

- The potential for diurnal migration should be reflected in cyanobacterial monitoring programs through the inclusion of multiple-sampling times or conducting sampling at an ecologically significant time of day.
- Sampling of lakes should not be restricted to the surface and should include discrete sampling and multiple depths to reflect spatial heterogeneity of cyanobacterial communities present in the water column.
- Cyanobacterial communities are highly dynamic and may vary significantly in composition, prevalence and occurrence times between sampling years at the same sampling site.
- Cyanobacterial monitoring may be enhanced through the incorporation of system characteristics (e.g., thermal stratification) and characterization of communities for design of sampling protocols that are system specific.

Acknowledgements

We acknowledge the support of the *for*Water NSERC Network for Forested Drinking Water Source Protection Technologies [NETGP-494312-16]. We are also grateful for the continued support of Natural Resources Canada and Environment and Climate Change Canada in sample collection at Turkey Lakes Watershed Research Station.

6, References

Bisanz, J.E., 2018. qiime2R: Importing QIIME2 artifacts and associated data into R sessions.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K. bin, Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A. v., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R.,

Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37, 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bukata, R.P., Jerome, J.H., Bruton, J.E., 1988. Relationships Among Secchi Disk Depth, Beam Attenuation Coefficient, and Irradiance Attenuation Coefficient for Great Lakes Waters. J Great Lakes Res 14, 347–355. https://doi.org/10.1016/S0380-1330(88)71564-6

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13, 581–583. https://doi.org/10.1038/nmeth.3869

Cameron, E.S., Müller, K.M., Stone, M., Buttle, J., Leach, J., Webster, K., Emelko, M.B., 2022. Early Seasonal Increases and Persistence in Relative Abundance of Potentially Toxic Cyanobacteria: Concerning Impacts of Extended Ice-Free Periods in Northern Temperate Lakes. bioRxiv 2022.12.20.521158. https://doi.org/10.1101/2022.12.20.521158

Cameron, E.S., Schmidt, P.J., Tremblay, B.J.M., Emelko, M.B., Müller, K.M., 2021. Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. Sci Rep 11. https://doi.org/10.1038/s41598-021-01636-1

Chapman, A.D., 2010. Cyanobacteria, in: Algae Source to Treatment - Manual of Water Supply Practices M57. American Water Works Association, pp. 125–146.

Chien, Y.C., Wu, S.C., Chen, W.C., Chou, C.C., 2013. Model simulation of diurnal vertical migration patterns of different-sized colonies of microcystis employing a particle trajectory approach. Environ Eng Sci 30, 179–186. https://doi.org/10.1089/ees.2012.0318

Chiu, Y.T., Chen, Y.H., Wang, T.S., Yen, H.K., Lin, T.F., 2017. A qPCR-based tool to diagnose the presence of harmful cyanobacteria and cyanotoxins in drinking water sources. Int J Environ Res Public Health 14. https://doi.org/10.3390/ijerph14050547

Emelko, M.B., Silins, U., Bladon, K.D., Stone, M., 2011. Implications of land disturbance on drinking water treatability in a changing climate: Demonstrating the need for " source water supply and protection" strategies. Water Res 45, 461–472. https://doi.org/10.1016/j.watres.2010.08.051

Emelko, M.B., Stone, M., Silins, U., Allin, D., Collins, A.L., Williams, C.H.S., Martens, A.M., Bladon, K.D., 2016. Sediment-phosphorus dynamics can shift aquatic ecology and cause downstream legacy effects after wildfire in large river systems. Glob Chang Biol 22, 1168–1184. https://doi.org/10.1111/gcb.13073

Frempong, E., 1981. Diel Variation in the Abundance, Vertical Distribution, and Species Composition of Phytoplankton in a Eutrophic English Lake. J Ecol 69, 919. https://doi.org/10.2307/2259645

Ganf, G.G., 1974. Diurnal Mixing and the Vertical Distribution of Phytoplankton in a Shallow Equatorial Lake (Lake George , Uganda). Journal of Ecology 62, 611–629.

Gilbert, J.A., Field, D., Swift, P., Thomas, S., Cummings, D., Temperton, B., Weynberg, K., Huse, S., Hughes, M., Joint, I., Somerfield, P.J., Mühling, M., 2010. The taxonomic and functional diversity of microbes at a temperate coastal site: A “multi-omic” study of seasonal and diel temporal variation. PLoS One 5. https://doi.org/10.1371/journal.pone.0015545

Gloor, G.B., Macklaim, J.M., Pawlowsky-Glahn, V., Egozcue, J.J., 2017. Microbiome datasets are compositional: And this is not optional. Front Microbiol 8, 1–6. https://doi.org/10.3389/fmicb.2017.02224

Gloor, G.B., Macklaim, J.M., Vu, M., Fernandes, A.D., 2016. Compositional uncertainty should not be ignored in high-throughput sequencing data analysis. Austrian Journal of Statistics 45, 73–87. https://doi.org/10.17713/ajs.v45i4.122

Gu, Z., 2022. Complex heatmap visualization. iMeta 1. https://doi.org/10.1002/imt2.43

Harke, M.J., Steffen, M.M., Gobler, C.J., Otten, T.G., Wilhelm, S.W., Wood, S.A., Paerl, H.W., 2016. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, Microcystis spp. Harmful Algae 54, 4–20. https://doi.org/10.1016/j.hal.2015.12.007

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M.H., Visser, P.M., 2018. Cyanobacterial blooms. Nat Rev Microbiol 16, 471–483.

Hunter, P.D., Tyler, A.N., Willby, N.J., Gilvear, D.J., 2008. The spatial dynamics of vertical migration by Microcystis aeruginosa in a eutrophic shallow lake: A case study using high spatial resolution time-series airborne remote sensing. Limnol Oceanogr 53, 2391–2406. https://doi.org/10.4319/lo.2008.53.6.2391

Ibelings, B.W., Mur, L.R., Walsby, A.E., 1991. Diurnal changes in buoyancy and vertical distribution in populations of Microcystisin two shallow lakes. J Plankton Res 13, 419–436. https://doi.org/10.1093/plankt/13.2.419

Jalili, F., Trigui, H., Guerra Maldonado, J.F., Dorner, S., Zamyadi, A., Shapiro, B.J., Terrat, Y., Fortin, N., Sauvé, S., Prévost, M., 2021. Can Cyanobacterial Diversity in the Source Predict the Diversity in Sludge and the Risk of Toxin Release in a Drinking Water Treatment Plant? Toxins (Basel) 13. https://doi.org/10.3390/TOXINS13010025

Jeffries, D.S., Kelso, J.R.M., Morrison, I.K., 1988. Physical, Chemical and Biological Characteristics of the Turkey Lakes Watershed, Central Ontario, Canada. Canadian Journal of Fisheries and Aquatic Sciences 45, 3–13.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30, 772–780. https://doi.org/10.1093/molbev/mst010

Le, K.T.N., Maldonado, J.F.G., Goitom, E., Trigui, H., Terrat, Y., Nguyen, T.L., Husk, B., Shapiro, B.J., Sauvé, S., Prévost, M., Dorner, S., 2022. Shotgun Metagenomic Sequencing to Assess Cyanobacterial Community Composition

following Coagulation of Cyanobacterial Blooms. Toxins (Basel) 14. https://doi.org/10.3390/toxins14100688

Lefler, F.W., Berthold, D.E., Laughinghouse, H.D., 2023. Cyanoseq: A database of cyanobacterial 16S rRNA gene sequences with curated taxonomy. J Phycol. https://doi.org/10.1111/jpy.13335

Li, H., Qin, C., He, W., Sun, F., Du, P., 2023. Learning and inferring the diurnal variability of cyanobacterial blooms from high-frequency time-series satellite-based observations. Harmful Algae 123. https://doi.org/10.1016/j.hal.2023.102383

McMurdie, P.J., Holmes, S., 2012. Phyloseq: A bioconductor package for handling and analysis of high-throughput phylogenetic sequence data, in: Pacific Symposium on Biocomputing.

Naselli-Flores, L., Zohary, T., Padisák, J., 2021. Life in suspension and its impact on phytoplankton morphology: an homage to Colin S. Reynolds. Hydrobiologia 848, 7–30. https://doi.org/10.1007/s10750-020-04217-x

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32, 268–274. https://doi.org/10.1093/molbev/msu300

Olli, K., 1999. Diel vertical migration of phytoplankton and heterotrophic flagellates in the Gulf of Riga. Journal of Marine Systems 23, 145–163. https://doi.org/10.1016/S0924-7963(99)00055-X

Paerl, H.W., 2014. Mitigating harmful cyanobacterial blooms in a human- and climatically-impacted world. Life 4, 988–1012. https://doi.org/10.3390/life4040988

Patel, R., de Oliveira, A., Newby, R., Chu, T., 2019. Flow cytometric analysis of freshwater cyanobacteria: A case study. Water (Switzerland) 11. https://doi.org/10.3390/w11071422

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41, 590–596. https://doi.org/10.1093/nar/gks1219

Reinl, K.L., Brookes, J.D., Carey, C.C., Harris, T.D., Ibelings, B.W., Morales-Williams, A.M., De Senerpont Domis, L.N., Atkins, K.S., Isles, P.D.F., Mesman, J.P., North, R.L., Rudstam, L.G., Stelzer, J.A.A., Venkiteswaran, J.J., Yokota, K., Zhan, Q., 2021. Cyanobacterial blooms in oligotrophic lakes: Shifting the high-nutrient paradigm. Freshw Biol. https://doi.org/10.1111/fwb.13791

Reynolds, C.S., Oliver, R.L., Walsby, A.E., 1987. Cyanobacterial dominance: The role of buoyancy regulation in dynamic lake environments. N Z J Mar Freshwater Res 21, 379–390. https://doi.org/10.1080/00288330.1987.9516234

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: A versatile open source tool for metagenomics. PeerJ 2016. https://doi.org/10.7717/peerj.2584

Schmidt, P.J., Cameron, E.S., Müller, K.M., Emelko, M.B., 2021. Ensuring that fundamentals of quantitative microbiology are reflected in microbial diversity analyses based on next-generation sequencing. bioRxiv. https://doi.org/https://doi.org/10.1101/2021.06.19.449110

Scholz, J., Brahney, J., 2022. Evidence for multiple potential drivers of increased phosphorus in high-elevation lakes. Science of the Total Environment 825. https://doi.org/10.1016/j.scitotenv.2022.153939

Shahraki, A.H., Chaganti, S.R., Heath, D.D., 2020. Diel Dynamics of Freshwater Bacterial Communities at Beaches in Lake Erie and Lake St. Clair, Windsor, Ontario. Microb Ecol 81, 1–13.

Shannon, C.E., 1948. A Mathematical Theory of Communication. The Bell System Technical Journal 27, 623–656.

Silins, U., Bladon, K.D., Kelly, E.N., Esch, E., Spence, J.R., Stone, M., Emelko, M.B., Boon, S., Wagner, M.J., Williams, C.H.S., Tichkowsky, I., 2014. Five-year legacy of wildfire and salvage logging impacts on nutrient runoff and aquatic plant, invertebrate, and fish productivity. Ecohydrology 7, 1508–1523. https://doi.org/10.1002/eco.1474

Śliwińska-Wilczewska, S., Maculewicz, J., Felpeto, A.B., Latała, A., 2018. Allelopathic and bloom-forming picocyanobacteria in a changing world. Toxins (Basel) 10, 1–20. https://doi.org/10.3390/toxins10010048

Spigel, R.H., Imberger, J., 1987. Mixing processes relevant to phytoplankton dynamics in lakes. N Z J Mar Freshwater Res 21, 361–377. https://doi.org/10.1080/00288330.1987.9516233

Sterner, R.W., Reinl, K.L., Lafrancois, B.M., Brovold, S., Miller, T.R., 2020. A first assessment of cyanobacterial blooms in oligotrophic Lake Superior. Limnol Oceanogr 65, 2984–2998. https://doi.org/10.1002/lno.11569

Stoddard, J.L., van Sickle, J., Herlihy, A.T., Brahney, J., Paulsen, S., Peck, D. v., Mitchell, R., Pollard, A.I., 2016. Continental-Scale Increase in Lake and Stream Phosphorus: Are Oligotrophic Systems Disappearing in the United States? Environ Sci Technol 50, 3409–3415. https://doi.org/10.1021/acs.est.5b05950

Tsai, K.P., Uzun, H., Karanfil, T., Chow, A.T., 2017. Dynamic Changes of Disinfection Byproduct Precursors following Exposures of Microcystis aeruginosa to Wildfire Ash Solutions. Environ Sci Technol 51, 8272–8282. https://doi.org/10.1021/acs.est.7b01541

Valadez-Cano, C., Hawkes, K., Calvaruso, R., Reyes-Prieto, A., Lawrence, J., 2022. Amplicon-based and metagenomic approaches provide insights into toxigenic potential in understudied Atlantic Canadian lakes. Facets 7, 194–214. https://doi.org/10.1139/facets-2021-0109

von Orgies-Rutenberg, M., Rolfes, C., Eckel, T., Quiroz, A., Skalbeck, J., Riley, D., Sander, H., 2018. Diurnal vertical migration of cyanobacteria and chlorophyta

in eutrophied shallow freshwater lakes. Fundamental and Applied Limnology 191. https://doi.org/10.1127/fal/2017/1021

Vu, H.P., Nguyen, L.N., Zdarta, J., Nga, T.T.V., Nghiem, L.D., 2020. Blue-Green Algae in Surface Water: Problems and Opportunities. Curr Pollut Rep 6, 105–122. https://doi.org/10.1007/s40726-020-00140-w

Walsby, A.E., Hayes, P.K., Boje, R., Stal, L.J., 1997. The selective advantage of buoyancy provided by gas vesicles for planktonic cyanobacteria in the Baltic Sea. New Phytologist 136, 407–417. https://doi.org/10.1046/j.1469-8137.1997.00754.x

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., Jansson, J.K., Caporaso, J.G., Fuhrman, J.A., Apprill, A., Knight, R., 2015. Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems 1, e0009-15. https://doi.org/10.1128/mSystems.00009-15.Editor

Wang, S., Zhang, X., Wang, C., Chen, N., 2023. Temporal continuous monitoring of cyanobacterial blooms in Lake Taihu at an hourly scale using machine learning. Science of the Total Environment 857. https://doi.org/10.1016/j.scitotenv.2022.159480

Westwood, K.J., Ganf, G.G., 2004. Effect of mixing patterns and light dose on growth of Anabaena circinalis in a turbid, lowland river. River Res Appl 20, 115–126. https://doi.org/10.1002/rra.725

Woolway, R.I., Merchant, C.J., 2019. Worldwide alteration of lake mixing regimes in response to climate change. Nat Geosci 12, 271–276. https://doi.org/10.1038/s41561-019-0322-x

Xue, K., Ma, R., Shen, M., Wu, J., Hu, M., Guo, Y., Cao, Z., Xiong, J., 2023. Horizontal and vertical migration of cyanobacterial blooms in two eutrophic lakes observed from the GOCI satellite. Water Res 240, 120099. https://doi.org/10.1016/j.watres.2023.120099

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., Glöckner, F.O., 2014. The SILVA and “all-species Living Tree Project (LTP)” taxonomic frameworks. Nucleic Acids Res 42, D643–D648. https://doi.org/10.1093/nar/gkt1209