Using sediment DNA archives for interpreting long-term cyanobacterial dynamics in the Anthropocene

Hebah Shaker Mejbel

Thesis submitted to the University of Ottawa in partial fulfillment of the requirements for the Doctorate in Philosophy (PhD) Biology

Department of Biology

Faculty of Science

University of Ottawa

DEDICATION

To my mom and dad, both the noor of my eyes.

ABSTRACT

Climate change and eutrophication, accelerated by anthropogenic activities, have impacted aquatic ecosystems worldwide. These impacts have stimulated the expansion of cyanobacterial blooms which pose severe threats to ecosystem functioning, environmental health, and the economy. However, the long-term effects of environmental change on bloom-forming cyanobacteria are not well understood as traditional paleolimnological approaches are of limited use in the reconstruction of cyanobacterial dynamics through time. Here, sediment DNA (sedDNA) was used to investigate long-term cyanobacterial trends using sediments from two experimental (fertilized L227 and acidified L223) and two reference (L224 and L442) lakes in the Experimental Lakes Area, Canada. First, to determine whether taxonomic bias might arise from the cyanobacterial sediment record, I performed a 1-year incubation experiment comparing the degradation rates of selected cyanobacterial genes under contrasting environmental conditions. Based on first-order linear decay models, *Synechococcus* sp. (Synechococcales) decayed the slowest under cold, anoxic conditions, followed by *Trichormus* (Nostocales), then *Microcystis* (Chroococcales), suggesting differential preservation of DNA. I then compared the quantitative performance of droplet digital polymerase chain reaction (ddPCR) and high-throughput sequencing (HTS) for the analysis of sedDNA and found that the ddPCR results were more consistent with the known history of the lakes. Furthermore, ddPCR showed that cyanobacterial abundance increased over the past century in all study lakes, but the greatest increase was observed in experimentally fertilized L227. HTS revealed shifts in the cyanobacterial community towards Nostocales dominance and a decrease in alpha diversity in response to phosphorus-only additions. An increase in abundance of the *mcyE* gene (indicative of microcystin producing taxa) was uniquely observed in L227 when nitrogen additions ceased.

Heating degree days were important in explaining variation in the cyanobacterial community composition in all lakes, but nutrients had a greater influence on the L227 community. When sediment data were compared to historical surface water phytoplankton records, moderate to strong correlations between the two archives were found, validating the use of sedDNA. This research demonstrated that sedDNA can elucidate cyanobacterial trends at the community, population, and species level over multidecadal timescales in response to environmental change.

RÉSUMÉ

Le changement climatique et l'eutrophisation, accélérés par les activités anthropiques, ont eu un impact sur les écosystèmes aquatiques du monde entier. Ces impacts ont stimulé la prolifération de cyanobactéries qui constituent de graves menaces pour le fonctionnement des écosystèmes, la santé de l'environnement, et l'économie. Cependant, les effets à long terme des changements environnementaux sur les efflorescences de cyanobactéries ne sont pas bien compris car les approches paléolimnologiques traditionnelles sont d'une utilité limitée dans la reconstruction de la dynamique des cyanobactéries au fil du temps. Ici, l'ADN sédimentaire (l'ADNséd) a été utilisé pour étudier les tendances à long terme des cyanobactéries à l'aide des sédiments de deux lacs expérimentaux (L227 fertilisé et L223 acidifié) et de deux lacs de référence (L224 et L442) dans la Région des Lacs Expérimentaux, Canada. Premièrement, pour déterminer si un biais taxonomique pouvait survenir avec l'enfouissement des sédiments, j'ai effectué une expérience d'incubation d'un an en comparant les taux de dégradation de gènes cyanobactériens sélectionnés dans des conditions environnementales contrastées. Basé sur des modèles de dégradation linéaire de premier ordre, *Synechococcus* sp. (Synechococcales) décroît le plus lentement dans des conditions froides et anoxiques, suivi par *Trichormus* (Nostocales), puis *Microcystis* (Chroococcales), suggérant une préservation différentielle de l'ADNséd. J'ai ensuite comparé les performances quantitatives de ddPCR et du HTS pour l'analyse de l'ADNséd et j'ai constaté que les résultats de la ddPCR étaient plus cohérents avec l'histoire connue des lacs. De plus, la ddPCR a montré que l'abondance des cyanobactéries a augmenté au cours du siècle dernier dans tous les lacs étudiés, mais la plus grande augmentation a été observée dans le L227 fertilisé expérimentalement. HTS a révélé des changements dans la communauté cyanobactérienne vers la dominance de Nostocales et une diminution de la diversité alpha en

réponse aux ajouts de phosphore uniquement. Une augmentation de l'abondance du gène *mcyE* (indiquant des taxons produisant des microcystines) a été observée uniquement dans L227 lorsque les ajouts d'azote ont cessé. Les degrés-jours de chauffage étaient importants pour expliquer la variation de la composition de la communauté cyanobactérienne dans tous les lacs, mais les nutriments ont eu une plus grande influence sur la communauté L227. Lorsque les données sur les sédiments ont été comparées aux enregistrements historiques du phytoplancton des eaux de surface, des corrélations modérées à fortes entre les deux archives ont été trouvées, validant l'utilisation de l'ADNséd. Cette recherche a démontré que l'ADNséd peut élucider les tendances des cyanobactéries au niveau de la communauté, de la population, et des espèces sur des échelles de temps multidécennale en réponse aux changements environnementaux.

ACKNOWLEDGEMENTS

I would like to convey my sincere gratitude and appreciation to my supervisor, Dr. Frances Pick for her continuous support, motivation, and guidance over the past few years. As someone who was coming in from a completely different research field, this project could not have been possible without her mentorship and unwavering patience and encouragement. Thank you for your advice, for challenging me, and for me providing me with countless opportunities—all of which have fostered enormous growth in me as a researcher and scientist.

I want to express my deep appreciation and thanks to my committee members, Dr. Danielle Fortin, Dr. Alexandre Poulain, and Dr. Myron Smith. Thank you for your comments and constructive feedback, for providing assistance whenever I needed it, and for your continuous guidance along the way. Thank you to all the past and present members of the Pick Lab. To Mary Ann Perron, thank you for your constant support, encouragement, and for being a listening ear whenever I needed to vent. To William Dodsworth, thank you for your friendship over the past few years and for helping me with my research so willingly. Amber Dyck, your passion and dedication inspire me to be the best version of myself. Thank you to the undergraduate students who have been integral in the completion of this thesis: Courtney Irwin, Daniel Pelletier, and Bryan Contreras Mercado. I could not have done this without your help.

To my fellow colleagues at CAREG and ARC, I would like to acknowledge and thank you for allowing me to use your equipment for my research, for sharing your expertise, and for the many insightful and engaging conversations we have had. I want to especially thank the members of the Poulain Lab for, not only their assistance with the bioinformatic processing, but also for allowing me to be an honourary member of their lab for the better part of 2020. I would

also like to thank Sofia Perin and Linda Kimpe for their technical expertise and for answering my many questions over the past few years.

There are many collaborators that I want to thank for their continuous support. I want to gratefully acknowledge the support from my collaborators at McGill University, especially Dr. Irene Gregory-Eaves and the members of the Gregory-Eaves Lab. My gratitude also goes to the researchers at the Experimental Lakes Area, particularly Michael Paterson, Scott Higgins, and Cyndy Desjardins. Thank you to Heather Roshon from the University of Waterloo for her guidance and for answering any question I had regarding cyanobacteria taxonomy. I want to thank the NSERC CREATE ABATE program for the financial support and for establishing a collaborative environment.

Finally, I want to express my most heartfelt gratitude to my family for which this journey would not have been possible. Thank you, Ayoub, Amjad, Mahmoud, Asil, Ala, Eman, Sosan, and Athari, for your continuous encouragement, for listening to me complain about things that may not have always made much sense, but most importantly, for teaching me to never settle for anything less than my best. To my mom and dad, thank you for your wisdom and unconditional love, for believing in me, and for supporting my dreams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANA	*Anabaena*-specific 16S rRNA gene
apcA	Allophycocyanin-alpha chain gene specific to *Synechococcus*
ASV	Amplicon Sequence Variant
bp	Base pair
C	Concentration
CDD	Cooling Degree Days
CPCC	Canadian Phycological Culture Centre
CRS	Constant Rate of Supply
CYA	Cyanobacteria-specific 16S rRNA gene
ddPCR	Droplet Digital Polymerase Chain Reaction
DGGE	Denaturing Gradient Gel Electrophoresis
DNA	Deoxyribonucleic Acid
eDNA	Environmental DNA
GCN	Gene Copy Number
glnA	Glutamine synthetase gene
HDD	Heating Degree Days
HTS	High-throughput sequencing
IISD-ELA	International Institute for Sustainable Development-Experimental Lakes Area
k	Decay rate
LOB	Limit of Blank
LOD	Limit of Detection
MAT	Mean Annual Temperature
mcyE	Microcystin synthetase E gene
MICR	*Microcystis*-specific 16S rRNA gene
n	Number of values
QIIME2	Quantitative Insights Into Microbial Ecology 2
qPCR	Quantitative Polymerase Chain Reaction
RDA	Redundancy Analysis
rRNA	Ribosomal Ribonucleic Acid
SD	Standard Deviation
SE	Standard Error
sedDNA	Sediment DNA
t	Time
TAP	Total Annual Precipitation
TN	Total Nitrogen
TP	Total Phosphorus
$\mathbf{T_{90}}$	Time it takes to reduce population by 90%

CHAPTER 1: GENERAL INTRODUCTION

As temperatures continue to rise globally, and as human activities accelerate the rate of eutrophication in freshwater ecosystems, a significant impact on water resources is expected. Together, the negative impacts of climate change and anthropogenic activities have led to drinking water advisories, wildlife mortalities, changes to food-web dynamics, economic losses, and the deterioration of water quality worldwide (Paerl and Otten, 2013; Huisman et al., 2018). This is particularly important in northern temperate regions, as mean annual temperatures in northern regions have increased at a faster rate than the global average (IPCC, 2018). The lack of sufficient long-term monitoring data, however, can make it difficult to assess historical trends and determine how and which anthropogenic factors have impacted aquatic ecosystems over decadal and centennial timescales (Smol, 2010). The analysis and reconstruction of environmental change through temporal and spatial scales using lake sediment records can lead to the successful management, restoration, and conservation of aquatic ecosystems that have been impacted by these drivers.

1.1 Cyanobacteria and harmful algal blooms

Climate change has been known to affect the physical, chemical, and biological characteristics of aquatic ecosystems (e.g., Barange and Perry, 2009; O'Neil et al., 2012). However, while the increased risk of flooding, droughts, and storms are of concern, it is perhaps the proliferation of cyanobacteria that creates the most alarm in freshwater and marine coastal systems. Understanding the impact of cyanobacteria taxa and climate change on water sources has been steadily growing, as more blooms are being reported around the world including Canada (for e.g., Ontario; Winter et al., 2011). The intensification of observed harmful cyanobacterial blooms worldwide has also been attributed to increased scientific and monitoring

efforts (Hallegraeff et al., 2021). In recent decades, the increasing intensity of cyanobacterial blooms has impacted aquatic ecosystem and human health, in addition to causing substantial economic loss (Bingham et al., 2015). Current monitoring programs are effective but are only typically initiated in cases of emerging blooms and are reactive rather than proactive.

Cyanobacteria are a diverse group of photosynthetic microbes, commonly known as blue-green algae due to the presence of blue-coloured phycocyanin as key photosynthetic pigments (Whitton, 2012). These prokaryotes play an important role in the cycling of nitrogen, carbon, and oxygen, were critical in creating the present oxygenic atmosphere and biosphere, and produce a large fraction of biomass on Earth (Falkowski et al., 2000). Taxa vary in size, with some being amongst the smallest species of phytoplankton (~0.6 µm) at the limit of the light microscope and others forming large macroscopic colonies (~1 – 2 mm) (Bonilla and Pick, 2017). They can be unicellular, colonial, or filamentous and found nearly worldwide, inhabiting different environments including deserts, lakes, wetlands, streams, hot springs, rivers, and salt lakes (Rippka et al., 1979). Some species can survive in a range of environments, which allows them to further proliferate and dominate over other species.

In addition to their physiological ability to survive at high temperatures, some cyanobacteria can persist when growth conditions are unfavourable through the production of heterocysts and specialized resting stages called akinetes (Vincent, 2009a; Kaplan-Levy et al., 2010). Heterocysts are differentiated vegetative cells responsible for converting nitrogen gas to ammonium, allowing the cyanobacteria to grow in conditions of limited nitrogen (Adams and Carr, 1981). Akinetes are thick specialized cells that are also produced by the Nostocales order and serve as survival structures when environmental conditions are unfavourable (Adams and Carr, 1981; Kaplan-Levy et al., 2010). When temperatures are low and nutrients are deficient,

akinetes store glycogen (for carbon), cyanophycin globules (for nitrogen), and nucleic acids, and remain dormant until conditions are favourable again (Kaplan-Levy et al., 2010; Sukenik et al., 2012).

When conditions are favourable and nutrient (primarily phosphorus and nitrogen) availability is high, planktonic cyanobacteria can form large blooms (Figure 1.1) where the increased biomass can cover the entire surface of water bodies (Reynolds and Walsby, 1975; Klemer and Konopka, 1989) and can be seen by satellite in large lakes (Becker et al., 2009; Sayers et al., 2019). Though blooms have been observed for centuries, ongoing global warming and eutrophication, have likely caused the frequency and intensity of blooms to rise. The increased biomass can cause anoxia, impact drinking water, and alter food webs (O'Neil et al., 2012; Paerl, 2014). Additionally, the rise of atmospheric carbon dioxide concentrations over the past decades has led to increased water temperatures, an increase in acidity of the oceans (lower pH), further water evaporation, vertical stratification, and water column stability (Barange and Perry, 2009; Paerl, 2014). The main causal driver, based on the sediment records, seems to be agricultural intensification (Taranu et al., 2015), triggering an increase in cyanobacterial and cyanotoxin concentrations (Zastepa et al., 2016).

1.2 Environmental drivers of cyanobacterial blooms

The direct causes of cyanobacterial dominance have been extensively examined for many years now and continue to be an area of scientific and societal interest. Most studies, however, are limited by examining cyanobacteria as one group rather than a diversity of taxa that each can contribute to bloom formation. Different species have varying rates of growth and biomass accumulation during the growing season that can be attributed to the level of nutrients, weather

conditions, and other factors. As such, the relative effects of eutrophication, climate change, and acidification on cyanobacterial diversity remain poorly understood.

As previously mentioned, high levels of nutrients in water often contribute to cyanobacterial dominance. Many studies have evaluated the impacts of nutrient loading and the increase in cyanobacterial biomass in Canadian lakes and worldwide (Schindler, 1977; Watson et al., 1997; Pick, 2016). When total phosphorus, the primary limiting nutrient in lakes, rises to 20 to 30 µg/L, the risk of a bloom also increases (Downing et al., 2001). Total nitrogen is also a strong contributor (Downing et al., 2001), and if inorganic nitrogen levels are low while phosphorus remains high, nitrogen-fixing cyanobacteria, primarily from the Nostocales order, typically become the dominant taxa. In lakes enriched with phosphorus, there is strong experimental and empirical evidence of the effect of nitrogen in explaining species composition differences between lakes (Dolman et al., 2012; Kolzau et al., 2018).

While nutrients have long been considered the key driver of blooms, climate change is considered to be the greatest factor currently impacting freshwater ecosystems, especially as temperatures globally have increased over the past 150 years. Global temperature increase is expected to exceed 2°C by the end of the 21st century (IPCC, 2021). A recent IPCC special report states that if global warming increases by 2°C or more, virtually all coral reefs would be lost by 2100 (IPCC, 2018). These temperature changes affect both the quantity of water and its quality. Temperature changes can elongate the growth season and may cause a shift in phytoplankton biomass towards cyanobacteria. For example, higher summer water temperatures in temperate regions during heatwaves have promoted nuisance cyanobacterial blooms (Jöhnk et al., 2007; Paerl and Huisman, 2009). The lack of long-term data studies, however, limits our

understanding of the combined effects of climate change and anthropogenic-induced eutrophication on cyanobacteria taxa and their dynamics over long periods of time.

While cyanobacteria are quite tolerant of high temperatures and extremes in nutrient concentrations, the effects of acidic environments on cyanobacterial dynamics are not completely understood. In controlled laboratory experiments, cyanobacterial growth ceased at pH levels below 4 (Brock, 1973). Other experiments have shown that some acidity may lead to an increase in cyanobacterial biomass. For example, in Lake 223 at the Experimental Lakes Area (ELA), H_2SO_4 was added into the lake between 1976-1983 and as the pH dropped from 6.7 to 5.1, cyanobacterial biomass increased 10-fold (Findlay and Kasian, 1996). Lake 302S at ELA was also acidified with H_2SO_4 between 1981 and 1989 (pH dropped down from 6.4 to 4.5) where fossil pigment record analysis (including pigments specific to filamentous, colonial, and nitrogen-fixing cyanobacteria) showed total algal biomass increased 200 to 400% upon acidification, but subsequently declined under severe acidification (Leavitt et al., 1999).

1.3 Paleolimnological approaches and sediment DNA

The long-term impacts of climate change and various anthropogenic activities on cyanobacterial dynamics can be analyzed using sediment archives. Sediments act as natural repositories of lake dynamics, containing both biotic and abiotic features of the lake. The biological and chemical materials that accumulate in lakes allows environmental change to be recorded over a long period of time (Smol, 2008). Sediment chronologies can be reconstructed using geochemical markers, radioisotope dating, and fossil records (Smol, 2008). This can be used to reconstruct how certain environmental conditions have impacted freshwater and other aquatic ecosystems.

In traditional paleolimnological methods, fossils, pollen, pigments, and other subfossil proxies are used to reconstruct environmental and evolutionary change. However, these approaches are limited as many organisms do not produce a diagnostic subfossil. Additionally, these techniques are often labor-intensive, somewhat unstandardized (as results can depend on the expertise and experience of the analyst), costly, and time consuming. Over the past decade, alternate approaches in tandem with traditional methods have been developed, including the use of sediment DNA (sedDNA).

The techniques and applications for environmental DNA have multiplied over the past decade (Taberlet et al., 2018; Ruppert et al., 2019). It is now well established that DNA is released from animals, plants, bacteria, viruses, fungi, and protists through passive or active secretion and/or from cell lysis. In aquatic ecosystems, intracellular DNA (DNA intact in cells) can persist in sediment for centuries (Ellegaard and Ribeiro, 2018). Extracellular DNA, too, can persist for long periods of time by binding and settling into the sediment through the formation of cationic bridges, in which the negatively charged phosphate groups on the DNA bind to the cations (such as sodium, calcium, and magnesium) in the sediment, making it stable for long-term analysis (Figure 1.2). Some of the greatest advantages of using sedDNA to track changes through temporal and spatial scales include the higher sensitivity and precision, the analysis of many taxa (down to the species level), and the standardized approach of molecular experiments. In general, molecular techniques are often performed in a consistent, repeatable manner, with only minor modifications, and do not solely rely on the expertise of the analyst.

While sedDNA has emerged as a promising tool for assessing diversity on a finer scale, and while rapid technological advancements have led to the successful reconstruction of complex communities, there are some difficulties associated with this approach. For instance, it is not

known how much DNA ends up in the sediment relative to the starting point or how long DNA preserves in sediment. In addition, a potentially high level of background contamination and low DNA quality and yield can interfere with successful DNA extractions. Optimized extraction procedures are required for downstream molecular biological applications. While the use of quantification methods including real-time PCR (qPCR), droplet digital PCR (ddPCR), and high-throughput sequencing (HTS) have been increasing in sedDNA studies over the past decade (Pal et al., 2015; Monchamp et al., 2016, 2018; Tse et al., 2018; Pilon et al., 2019; Dodsworth, 2020), a high level of potential PCR inhibitors (including humic and fulvic substances) can interfere with these molecular techniques. This indicates that the use of sedDNA requires further testing.

1.4 Analysis of cyanobacteria using sediment archives

Cyanobacteria have been observed to persist in lake sediments over millennia. For example, van Geel et al., (1994), found akinete fossils from sites going back to the Bronze Age (~3300 – 1200 BCE, approximately 5320 – 3220 years ago). Heterocysts have also been found to persist for long periods of time due to their thick cell walls. The use of akinetes and heterocysts for the reconstruction of temporal cyanobacterial trends is, however, somewhat limited as only a subset of cyanobacterial taxa produces these differentiated cells and the relationship between their numbers and population size is highly variable. Other studies have reconstructed long-term cyanobacterial trends using sedimentary algal pigments (Leavitt and Findlay, 1994; Deschpande et al., 2014; Pal et al., 2015; Tse et al., 2018) but the use of cyanobacterial toxins is still at the research stage (Klitzke et al., 2011; Waters, 2016; Zastepa et al., 2016; Henao et al., 2020). While the use of pigments and toxins have been successful in analyzing cyanobacterial dynamics as a group in response to environmental and anthropogenic change, cyanobacteria are a large and

ecologically (as well as functionally) diverse group of microbes, and taxa-specific data is limited. As such, the analysis of long-term cyanobacterial dynamics using sedDNA has been increasingly applied to the study of lakes over the past decade (Table 1.1). These analyses have provided both quantitative data (e.g., Pal et al., 2015) and qualitative data (community composition shifts) (e.g., Monchamp et al., 2016; Pilon et al., 2019).

Most sedDNA cyanobacterial studies have used the cyanobacterial-specific 16S rRNA gene (the gene that encodes for the small subunit ribosomal RNA molecules of ribosomes), to analyse long-term cyanobacterial dynamics (see Table 1.1). The primer set targeting cyanobacterial 16s rRNA, first designed by Nübel et al., (1997), has been modified over the past two decades for the use in sedDNA studies. Other sedDNA studies have assessed genus-specific 16S rRNA genes, such as those associated with the common cosmopolitan toxigenic bloom-forming genus *Microcystis* (Pilon et al., 2019), internal transcribed spacer regions (ITS) (Hou et al., 2014; Savichtcheva 2014), and genes specific for cyanobacterial metabolites (for e.g., *ociB* gene: Kyle et al., 2015). Pigment genes, including the alpha chain blue-green pigment specific to *Synechococcus* (*apcA*) (Dodsworth, 2020), have also been used to quantify cyanobacterial abundance.

The cyanobacterial DNA recovered from sediments is a reflection of both planktonic and benthic growth across the body of water that is then transported to the sediments (Figure 1.2). During the settling process some taxa may degrade more quickly than others, possibly as a result of differential grazing pressure or taxonomic bias, as seen in the case of marine dinoflagellates (Boere et al., 2011). However, the analysis of preservation differences between cyanobacterial taxa is lacking in the literature and studies typically assume equal preservation or that there is no taxonomic bias across the sediment record.

1.5 Study system

Often, choosing a study site for paleolimnological analysis of sediments can be challenging as many disturbances, whether natural or anthropogenic, can impact the sediment record. To overcome some of these challenges, I analyzed cyanobacterial dynamics using sediment freeze cores collected from the Experimental Lakes Area (ELA), Northwestern Ontario, Canada (Figure 1.3). These lakes are remote with next to no human development. Furthermore, ELA has been the site of several ecosystem-level experimental manipulations of anthropogenic stressors on a few selected lakes since the late 1960s. For example, some lakes have been manipulated by acidification (Lake 223 and Lake 302), eutrophication (for e.g., Lakes 226, 227, 304), or contaminants (Schindler, 2009). Five lakes (Lakes 114, 224, 239, 373, and 442) have had minimal human impact and can be used as reference lakes, or controls, for scientific research. The lakes at ELA have very long monitoring records, with some records going back more than 50 years, making it ideal for comparing the sediment archives with known ecosystem changes and validating sedDNA analyses.

One lake of interest is Lake 227, where two nutrients (nitrogen and phosphorus), have been added to the lake since 1969 (see Supplementary Table S3.1 in Appendix B) in a long-term eutrophication experiment that has led to cyanobacterial dominance (Schindler, 1974; Schindler et al., 1990; Findlay and Kasian, 1996). From 1990 onwards, only phosphorus was and continues to be added. This was performed in order to test the role of nitrogen, which is considered to not play an important role in eutrophication given the capacity of Nostocales to use atmospheric nitrogen gas (Schindler et al., 2008). Other studies conducted on Lake 227 have analyzed changes in phytoplankton biomass in response to reduced nitrogen levels (Higgins et al., 2018), the examination of radioisotope mobility (Crusius and Anderson, 1995), and the responses of

diatoms and chrysophyte assemblages in the sediment record (Zeeb et al., 1994). In both the experimentally eutrophied ELA lakes, 226 and 227, a several-fold increase in phytoplankton biomass and a change in the dominant species has been observed (Schindler et al., 1971; Paterson et al., 2011). For Lake 227, after the cessation of nitrogen loading in 1990, nitrogen-fixing species dominated the cyanobacterial community (Schindler et al., 2008). Lakes 224 and 442 are oligotrophic lakes and have been used as reference lakes in previous studies (Findlay and Kasian, 1996; Schiff et al., 2017; Lamothe et al., 2018).

Lake 223 is another lake at ELA that has been manipulated. To simulate the effects of acid precipitation, sulfuric acid was added from 1976 to 1983, and the pH gradually dropped from ~6.7 to 5.1 (Schindler et al., 1980; Findlay and Kasian, 1986). As a result of the acidification in Lake 223, the overall biomass and phytoplankton diversity increased (Findlay and Kasian, 1996). Fewer studies have assessed this lake compared to Lake 227. Although, there have been studies that have analyzed the impacts of Lake 223 acidification on fish populations (Mills et al., 1987), changes to sedimentary diatom assemblages (Findlay and Shearer, 1992), and the potential of pH recovery (Findlay and Kasian, 1996). Findlay and Kasian (1996) found that as the lake pH naturally recovered, the phytoplankton community also recovered. Lake 224 is an appropriate reference lake for Lake 223 as both lakes are relatively deep, oligotrophic, and have similar water renewal times (Findlay and Kasian, 1996; Lamothe et al., 2018) (see Supplementary Table S3.1 in Appendix B).

1.6 Thesis objectives

The purpose of this thesis was to analyze changes in cyanobacterial abundance and composition in response to experimental and environmental change using sediment DNA.

Specifically, I used a combination of molecular techniques to determine temporal trends in lakes impacted by eutrophication, acidification, and climate change. This thesis is comprised of five chapters.

Following the present introductory chapter, Chapter 2 addressed the environmental factors that influence cyanobacterial dominance and DNA preservation in sediments. My goal was to determine the impacts of oxygen, temperature, and time, on DNA preservation of target genes in sediments using different cyanobacterial taxa as the original source of DNA. I tested the hypothesis that long-term DNA preservation in sediment was a function of the environmental conditions and that cyanobacterial genes exposed to oxygen and warmer temperatures would experience much faster rates of degradation compared to cyanobacteria in anoxic and cold conditions. The findings can be used to guide the interpretation of population and community structure data based on sedDNA, be used to develop strategic sampling approaches, and help refine storage techniques that will limit DNA degradation.

In the third chapter, I analyzed and compared the coherence between ddPCR and HTS derived data using sediment DNA collected from four ELA lakes. The goal was to determine how well these two molecular approaches track one another by comparing absolute cyanobacteria abundance outputs and relative *Microcystis* abundances. The results provide insight into how both molecular techniques can be suitable for sedDNA analyses in the field of paleolimnology and explain how ddPCR may be more suitable for certain analyses.

The fourth chapter used both ddPCR and HTS to investigate cyanobacterial temporal dynamics over time in relation to environmental drivers. I quantified the absolute abundance of various cyanobacterial target genes and determined the temporal trends in cyanobacterial diversity through the sediment record of a eutrophied, an acidified, and two reference lakes at

ELA. The impacts of various environmental drivers on influencing the cyanobacterial community composition were also evaluated. I predicted an increase in cyanobacterial abundance in response to climate change and eutrophication, and that community composition will shift as a result of nutrient loading. I also expected that the combined effects of acidification and a warming climate will lead to a fluctuation of cyanobacterial abundance over time. The unique long-term phytoplankton records from ELA obtained through microscopy were used to compare with the sediment record and thus validate the molecular approaches used.

Overall, the approach and experimental results from this research can be used to analyze other freshwater ecosystems impacted by climate change and anthropogenic activities. This may allow for further inferences about the dominant drivers in a particular lake and identify what mitigation and management efforts should be implemented to reduce or prevent the negative impacts associated with these drivers.

Table 1.1 A summary of the paleolimnological studies from the past decade that investigated cyanobacterial sediment DNA (ordered by continent, then by year).

Continent	Location	Cyanobacterial target genes	Molecular approach	Analysis	Reference
Africa	Limpopo River, South Africa	• Cyanobacterial 16S rRNA • *mcyA*, *mcyE* • Nodularin synthetase • *Cylindrospermopsin* polypeptide synthetase	PCR, DGGE, HTS	Physical-chemical characteristics of river sediments on cyanobacterial growth and toxicity	Magonono et al., 2018
Asia	Lake Taihu, China	• Cyanobacterial 16S rRNA • *Microcystis* 16S rRNA • *Synechococcus* 16S-23S ITS	DGGE, qPCR	Effects of eutrophication	Ye et al., 2011
	Lake Chenghai, China	• Cyanobacterial 16S rRNA • *mcyA* • RNA polymerase gamma subunit *rpoC1*	HTS, qPCR	Effects of climate change and eutrophication	Cao et al., 2020
	Yunnan−Kweichow Plateau lakes, China	• Cyanobacterial 16S rRNA • *mcyA*	HTS, qPCR	Effects of climate change and eutrophication	Zhang et al., 2021
Australia (Oceania)	Lake Rotorua, New Zealand	• Cyanobacterial 16S rRNA • *Microcystis* 16S rRNA • *mcyE*	HTS, qPCR, ddPCR	If a single sediment core can capture dominant microbial community	Weisbrod et al., 2020
Europe	French Alps, France	• Cyanobacterial 16S rRNA • Cyanobacterial 16S rRNA-ITS • Phycocyanin Intergenic Spacer Region	qPCR	Validation of sediment DNA, effects of climate change and eutrophication	Savichtcheva et al., 2011, 2014
	Lac du Bourget, France	• Cyanobacterial 16S rRNA • Cyanobacterial 16S rRNA-ITS • *Synechococcus* ITS	qPCR	Effects of climate change and eutrophication	Domaizon et al., 2013
	Lakes in Southern Norway, Norway	• *Planktothrix* oligopeptide class cyanopeptolin (*ociB*)	qPCR	Identification and amplification of *Planktothrix* DNA in sediments	Kyle et al., 2015
	Perialpine lakes, Switzerland	• Cyanobacterial 16S rRNA	HTS	Effects of climate change and eutrophication	Monchamp et al., 2016, 2018

	Lake Aydat, France	• Cyanobacterial 16S rRNA • *mcyA*, *mcyB* • *anatoxin*-C	PCR	Cyanobacterial persistence, effects of eutrophication	Legrand et al., 2017, 2019
	Lake Tiefer, Germany	• Cyanobacterial 16S rRNA	HTS	Effects of climate change and eutrophication	Nwosu et al., 2021a
			HTS, qPCR	Comparing pelagic composition to sediment trap data	Nwosu et al., 2021b
	Gulf of Finland, Finland	• Cyanobacterial 16S rRNA • *mcyB, mcyE* • Nodularin synthetase genes (*ndaF*)	PCR, HTS	Fate of akinetes in brackish water sediment archives	Wood et al., 2021
North America	Gatineau Park Lakes, Québec, Canada	• Cyanobacterial 16S rRNA	PCR, qPCR	Effects of climate change and eutrophication	Pal et al., 2015
	Lake Diefenbaker, Saskatchewan, Canada	• Cyanobacterial 16S rRNA • *mcyA*	HTS, qPCR	Effects of climate change and eutrophication	Tse et al., 2018
	Lake of the Woods, Ontario/Manitoba, and Baptiste Lake, Alberta, Canada	• Cyanobacterial 16S rRNA • *Microcystis* 16S rRNA • *mcyD*	HTS, qPCR	Effects of climate change and eutrophication	Pilon et al., 2019
	Central Ontario Lakes, Ontario, Canada	• Cyanobacterial 16S rRNA • *Microcystis* 16S rRNA • *mcyE*	PCR, ddPCR	Effects of climate change and eutrophication	Dodsworth et al., 2020
	Experimental Lakes Area, Ontario, Canada	• Cyanobacterial 16S rRNA • *Microcystis* 16S rRNA	HTS, ddPCR	Comparison of molecular approaches on sediment DNA	This thesis; Mejbel et al., 2021b
South America	Laguna Blanca, Uruguay	• Saxitoxin-U • 16S-23S rRNA	DGGE, qPCR	Use of ancient DNA in subtropical lakes	Martínez de la Escalera et al., 2014

ddPCR: Droplet Digital Polymerase Chain Reaction; **DGGE**: Denaturing Gradient Gel Electrophoresis; **HTS**: High-throughput sequencing; **ITS**: Internal Transcribed Spacer Region; ***mcy***: Microcystin synthetase gene; **PCR**: Polymerase Chain Reaction; **qPCR**: Quantitative Polymerase Chain Reaction

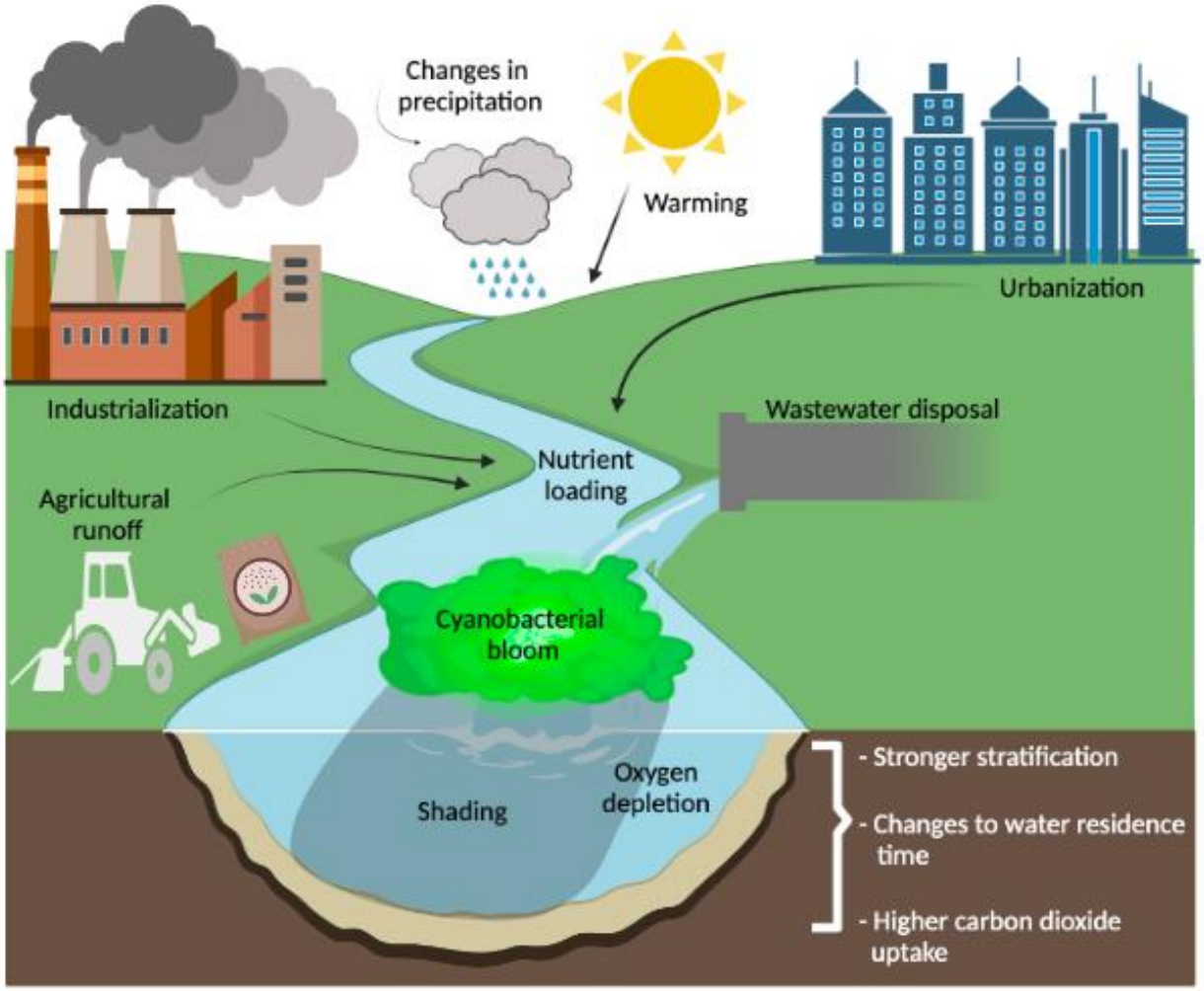

Figure 1.1 A conceptual figure illustrating the different factors that can contribute to the formation of a cyanobacterial bloom and the effects a bloom can have on a lake. Figure created through BioRender.

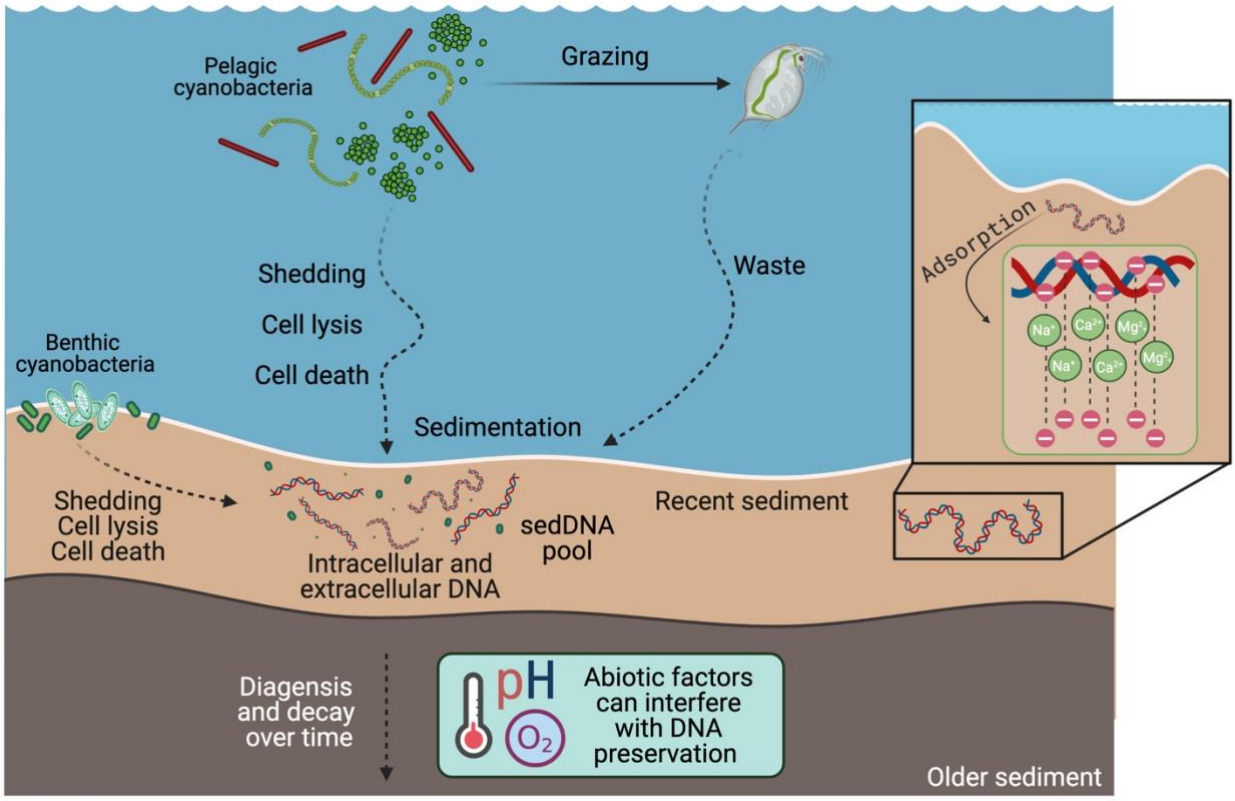

Figure 1.2 The release and deposition of DNA from pelagic and benthic cyanobacteria onto sediment. Figure from Mejbel et al., (2021a).

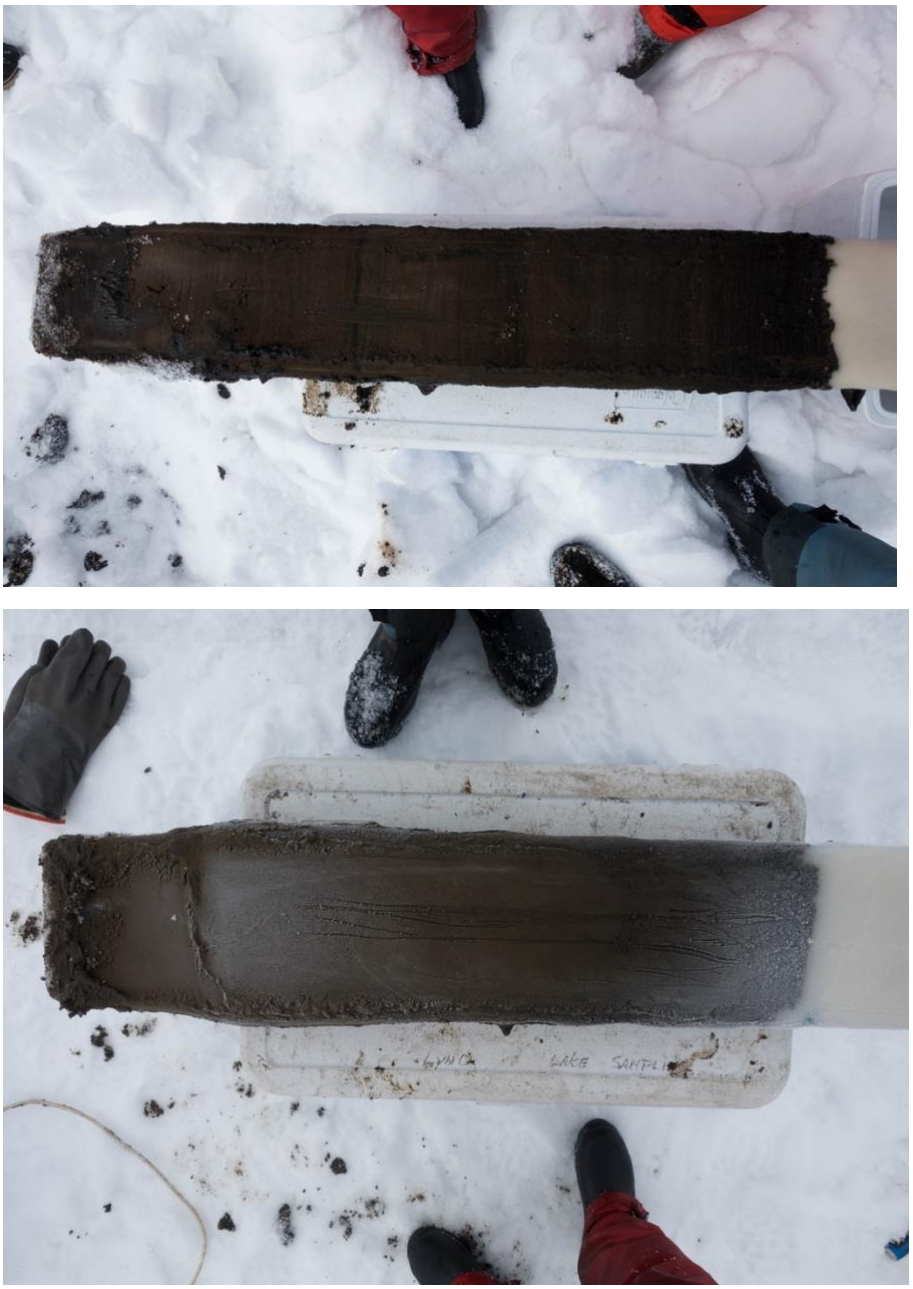

Figure 1.3 Lake 227 (top) and Lake 223 (bottom) freeze cores collected by IISD-ELA researchers in March 2018. Photos @ Michael Paterson.

CHAPTER 2: The effects of temperature and oxygen on cyanobacterial DNA preservation in sediments: a comparison study of major taxa

A version of this chapter has been published in
Environmental DNA

Mejbel, H.S., Dodsworth, W., and Pick, F.R (2022). The effects of temperature and oxygen on cyanobacterial DNA preservation in sediments: a comparison study of major taxa. *Environmental DNA.* 00, 1 – 15. doi: 10.1002/edn3.289

Centre for Advanced Research in Environmental Genomics, Department of Biology, University of Ottawa, Ottawa, ON, Canada

Author Contributions
HSM and FRP designed the study. HSM performed the molecular experiments, analyzed the data, and wrote the first draft of this article. WD and FRP revised the manuscript. All authors contributed to the article and approved the submitted version.

Keywords: anoxia, cyanobacteria, decay, droplet digital PCR, sediment DNA, temperature.

2. 1 Abstract

The analysis of sediment DNA has emerged as a promising alternative to classical paleolimnological proxies, which typically rely on hard fossils to reconstruct environmental change. In recent studies, the timing and causes of apparent increases in toxic cyanobacterial blooms in lakes around the world have been identified by high-throughput sequencing of cyanobacterial 16S rRNA genes from decadal- and centennial-scale sediment records. However, the reconstruction of ecological time series can be influenced by DNA degradation processes interfering with taxonomic resolution and quantification of target genes. To determine whether some cyanobacterial taxa degrade more rapidly than others, a one-year incubation experiment was performed to estimate degradation rates in sediments under contrasting environmental conditions of temperature (4°C vs. 25°C) and sediment redox (oxic vs. anoxic). Using sediments collected from an oligotrophic lake, serum bottles were inoculated with either *Microcystis aeruginosa*, *Synechococcus* sp., or *Trichormus variabilis*. These cyanobacteria were chosen because they are widespread yet exhibit contrasting morphologies. The absolute concentrations of the target genes were measured over time through droplet digital PCR and measured over time for one year. The three taxa exhibited log-linear declines with different rates of decay. Based on first-order linear decay models, *Microcystis* exhibited faster decay under oxic conditions, whereas *Synechococcu*s sp. and *Trichormus* were both temperature and anoxic dependent. Overall, cyanobacterial decay rates were lowest under 4°C anoxic conditions, corresponding to conditions in deep temperate lakes. Under 4°C, *Synechococcus* exhibited the lowest degradation rates followed by *Trichormus*, then *Microcystis*. These findings will be useful in the interpretation of population and community structure data based on sediment DNA and advances our current understanding of cyanobacterial preservation in sediment.

2.2 Introduction

The use of sediment DNA (sedDNA) has increased rapidly over the past several years to characterize and infer trends in aquatic systems over timescales of decades to millennia (Boere et al., 2011; Giguet-Covex et al., 2014; Sjögren et al., 2017; Pilon et al., 2019; Capo et al., 2021). This method has emerged as a valuable alternative to traditional paleolimnological methods which typically rely on either biological remains or non-genetic molecules and are often labour-intensive with limited taxonomic specificity (Coolen et al., 2004; Smol, 2008; Domaizon et al., 2017). Molecular-based techniques targeting sedDNA offer the potential to improve the analysis of complex microbial communities with higher specificity while also being more precise, cost-effective, and standardized (Pedersen et al., 2015; Zaiko et al., 2018). This is of particular importance to paleolimnological studies as some organisms do not leave fossil remains behind. However, limitations remain. Some of the major caveats in sedDNA analyses include the degradation of DNA in sediment and potential differential preservation of taxa along with their DNA (Boere et al., 2011; Capo et al., 2015). The reconstruction of ecological time series can therefore be influenced by DNA degradation processes interfering with taxonomic resolution and quantification of target genes.

Understanding the factors that affect DNA degradation rates in sediment is necessary for the accurate interpretation of data. Recent environmental DNA (eDNA) studies have shown the impacts of temperature, light (Barnes et al., 2014), pH (Strickler et al., 2015; Buxton et al., 2017), and dissolved organic carbon concentrations (Eichmiller et al., 2016) on DNA decay rates in water. Despite the fact that there is a need to consider the abiotic and biotic factors impacting eDNA (Stewart, 2019), very little work on the analysis of DNA degradation in sediment has been performed and the factors influencing decay rates of sedDNA remains largely undefined. In

general, previous studies have shown that adsorption onto clay minerals and other soil particulate material can slow down nuclease degradation (Blum et al., 1997; Cai et al., 2007), and the assumption is that taxa are preserved best under permanently anoxic and cold, dark conditions (Coolen et al., 2006; Domaizon et al., 2017). However, while some studies have demonstrated that DNA degradation can slow down under hypoxic or anoxic conditions, favouring long-term eDNA preservation (Corinaldesi et al., 2011; Jessen et al., 2017), others have found that the differences between oxic and anoxic degradation were often quite small (Arndt et al., 2013).

Temperature is often assumed to be the main variable affecting DNA degradation rates in sediment (Coolen et al., 2006; Coolen and Overmann, 2007), but the effects of other compounding variables may play a more important role than temperature alone. For example, DNA degraded faster in colder Antarctic sediments than in warmer Mediterranean Sea sediments possibly as a result of salinity, pH, and redox potential (Corinaldesi et al., 2008). Other abiotic factors, such as organic matter content and biotic factors such as DNase activity (Coolen and Overmann, 1998; Dell'Anno and Danovaro, 2005; Corinaldesi et al., 2008) have also been found to influence DNA preservation in sediments. To date sedaDNA studies have largely focused on deep, stratified lakes (Fernandez-Carazo et al., 2013; Hobbs et al., 2021; Kyle et al., 2015; Monchamp et al., 2016; Pal et al., 2015; Pilon et al., 2019; Yan et al. 2019) and seas (More et al., 2019; Orsi et al., 2017) where sediments are cold and/or anoxic. However, some studies have also focused on, or included, shallow lakes (Martínez De La Escalera et al., 2014; Moguel et al., 2021; Monchamp et al., 2017; Zhang et al., 2021), where sediments would reach higher temperatures and perhaps higher oxygen levels. There remains a need to analyze the combined impacts of environmental factors and differential preservation capacities between taxa to obtain a comprehensive understanding of DNA decay in sediments.

Differences in preservation capacity in sediment can lead to the over or underrepresentation of certain taxa, particularly for community-based analyses using metabarcoding techniques. Taxonomic biases in preservation have been reported in a few studies. For example, extracellular DNA from cryptophytes was found to be underrepresented in sediments of a deep Alpine lake when compared to the water column data (Capo et al., 2015). The low level of cryptophyte sedaDNA preservation could be explained by the absence of a cell wall in these microbes. Extracellular DNA is released and archived in sediment primarily through active or passive secretion or through cell lysis (Pietramellara et al., 2009). However, while the majority (> 90%) of sediment DNA is estimated to be extracellular (Dell'Anno and Danovaro, 2005), it is known that intracellular DNA (i.e., DNA intact in cells) can persist for longer periods of time (Ellegaard and Ribeiro, 2018). For example, in a study analyzing eukaryotic algae DNA from a 2700-year sediment core, dinoflagellate DNA degraded a thousand times more rapidly than DNA from diatoms, most likely due to diatom DNA preservation within intact cells (Boere et al., 2011). The analysis of differential preservation between cyanobacterial taxa has, thus far, been neglected in the literature and it is assumed in metabarcoding or community-based analyses that cyanobacteria taxa are equally preserved in sediments over time, indicating no bias against any particular taxonomic group within the cyanobacteria.

Reports of cyanobacterial blooms in fresh and marine waters have increased over the past few decades (Taranu et al., 2015), and more sedDNA analyses have been applied to identify the timing and drivers of these harmful cyanobacterial blooms (Pal et al., 2015; Monchamp et al., 2016; Pilon et al., 2019). However, more work is needed to clarify the potential compounding effects of species-specific preservation and the role of abiotic factors on DNA degradation in sediment. Here, I examined the sediment preservation of DNA in sediment from three different

lineages of cyanobacteria, two unicellular and one filamentous: *Microcystis aeruginosa* (from the Chroococcales order), *Synechococcus* sp. (Synechococcales), and *Trichormus variabilis* (also known as *Anabaena variabilis*; formerly *Anabaena flos-aquae*, from the Nostocales order) (Guiry and Guiry, 2021). All three taxa have been previously identified in sediment archives using metabarcoding and/or PCR amplification (Coolen et al., 2008; Savichtcheva et al., 2011; Domaizon et al., 2013; Monchamp et al., 2016, Pilon et al., 2019). *Microcystis* is a common bloom-forming genus, is colonial in nature, and can form large aggregations (Verspagen et al., 2006; Deng et al., 2015). In contrast *Synechococcus* sp. are typically unicellular and picoplanktonic in size (~ 1 μm in diameter); the *Synechococcus* genus is diverse and ubiquitous as a major primary producer in oceans and oligotrophic to mesotrophic freshwater ecosystems (Pick, 1991; Callieri et al., 2012). *Trichormus* species form filaments of spherical to ovoid cells and also produce differentiated cells, such as resting cysts (akinetes) and nitrogen-fixing cells (heterocysts), that are known to preserve well in sediments due to their larger size and thick cell walls (Adams and Duggan, 1999; Bradbury et al., 2004). There are many possible reasons as to why *Microcystis*, *Synechococcus*, and *Trichormus* might differ intrinsically in their rates of degradation in sediments. These include differences in colonial morphology, size, and cell structure (e.g., presence of exopolymeric substances in gelatinous sheaths or mucilage layers and chemistry of cell wall) (Callieri et al., 2012; Leak, 1967; Samuel et al., 2001; Skulberg et al., 1993; Šmarda et al., 2002; Whitton, 2012).

In the following study, I performed a one-year incubation experiment to estimate cyanobacterial DNA degradation rates in sediments under contrasting environmental conditions of temperature and sediment redox. Specifically, my goals were to 1) calculate decay rates of unicellular and filamentous cyanobacteria in sediment, and 2) compare the effects of oxygen

(oxic versus anoxic) and temperature (4°C versus 25°C) on three cyanobacterial taxa by targeting small gene fragments that collectively reflect the behaviour of the sedDNA. I tested the simple null hypothesis that degradation rates would not differ among taxa once live cells were mixed into sediments, and that preservation was instead a function of the sediment environmental conditions. Comparing degradation rates of cyanobacterial sedDNA under contrasting environmental conditions of temperature and sediment redox can advance our current understanding of cyanobacterial preservation in sediment and assist in the interpretation of abundance and community composition data. Findings from this study could also lead to the elucidation of strategic sampling approaches and improved storage techniques.

2.3 Materials and methods

2.3.1 Bacterial cultures used and experiment overview

Pure cultures and BG-11 medium obtained from the Canadian Phycological Culture Centre (CPCC, University of Waterloo, Canada) were used for the microcosm experiments; these were grown at room temperature in batch cultures and cells were harvested during the exponential phase. Serum bottles were filled with sediment and spiked with either *Microcystis aeruginosa* (CPCC 300), *Synechococcus* sp. (CPCC 663), or *Trichormus variabilis* (CPCC 67) (Figure 2.1). These microcosms were subsequently placed under contrasting environmental conditions (4°C oxic, 25°C oxic, 4°C anoxic, and 25°C anoxic) and target genes were quantified using droplet digital PCR (ddPCR) over the course of one year at seven specific timepoints. A schematic of the experimental set up is provided in Figure 2.2.

2.3.2 Sediment sampling

Sediments were collected in October 2019 from a small, soft water oligotrophic lake adjacent to Meilleurs Road, Laurentian Hills, Ontario (46°09’54.0”N 77°37’21.3”W). This lake was chosen as it does not experience cyanobacterial blooms and contains organic-rich sediments with low levels of dissolved silica, nitrate, and phosphate in the water column. Physical and chemical variables were measured *in situ* at different depths of the water column (Table S2.1). Sediments were collected using an Ekman grab from the deepest part of the lake (~3 m) and put into large 5-gallon screw top buckets for transportation. The buckets were then transferred back into the laboratory and placed in a 4°C refrigerator.

2.3.3 Experimental design

Prior to setting up the experiment, sediments were divided and placed into one litre beakers and thoroughly mixed using a sanitized scraper spatula to ensure a homogenous matrix. Large debris (macrophytes, twigs, etc.) were manually removed. To maintain the natural microbiological and chemical properties of the sediment such that the microcosms prepared for this study represent a typical environment, sediments were not autoclaved. Subsamples of 25 g of wet sediment were poured into 130 Weaton® 50-mL glass serum bottles to prepare the oxic microcosms. All cultures were thoroughly shaken prior to inoculation. Using sterile syringes, 2 mL of each culture was inoculated in triplicate for a total of 39 bottles for each target species (3 day 0 oxic bottles + 3 × 12 [six subsequent timepoints and two temperatures]). The bottles were then loosely covered with sterile aluminum foil. The remaining 13 serum bottles were not inoculated with a species but instead with 2mL of nuclease free water and used as experimental blanks. Inoculations were performed at room temperature in a stainless-steel laminar flow hood

(HEPA 100) that was treated with 70% ethanol and 10% bleach, then placed on a VWR OS-500 shaker at low speed for 30 seconds to ensure a homogenized mixture. Ten serum bottles (corresponding to three replicates per species plus one experimental blank) were "sacrificed" to estimate initial cyanobacterial gene abundance at day zero for all the oxic microcosms. To make the anoxic microcosms, 130 new bottles were prepared using the same methods. After inoculation, 120 bottles were placed in an anaerobic chamber (95% N_2; 5% H_2) and sealed with butyl rubber septa (Chemglass Life Sciences) and aluminum crimp caps. The inoculations were prepared in a laminar flow hood and not directly within the anaerobic chamber due to the high risk of potential cross-contamination and concerns with inadvertent oxygen entry (based on preliminary testing of inoculating within the chamber, which involved repeatedly opening and closing the air lock). The remaining ten bottles were used to estimate initial loading abundances at day zero for the anoxic microcosms. An additional 20 bottles (5 per environmental condition: 4°C oxic, 25°C oxic, 4°C anoxic, and 25°C anoxic) were set up to measure sediment mass over time and to determine moisture content at each timepoint under each condition.

To determine the effect of temperature on DNA preservation, half of the oxic and half of the anoxic microcosms were transferred and randomly positioned in dark sealed-off cabinets at 25°C. This temperature was chosen to mimic the warm summer temperatures of temperate shallow lakes. The other half were randomly positioned in a 4°C refrigerator to mimic year-round temperatures typically found at the bottom of deep temperate lakes. To limit the number of times the microcosms were exposed to light, temperatures were monitored only weekly for the first three months, then monthly for the subsequent months.

Replicates of the treatments were processed at specific timepoints to obtain DNA degradation estimates. The timepoints used for this experiment included zero days, one day (24

hours), 14 days (two weeks), 30 days (one month), 90 days (three months), 182 days (six months), and 365 days (one year).

2.3.4 DNA extraction and ddPCR quantification of target genes

DNA was extracted from the day zero bottles approximately two hours after inoculation. At each following time point, ten bottles (three from each culture and one experimental blank) were sacrificed from each environmental condition (oxic at 4°C, oxic at 25°C, anoxic at 4°C, and anoxic at 25°C, n=40) for DNA extraction and target gene quantification. DNA extractions were performed using ~0.7 g of subsample from each bottle. At the one-year time point, DNA extractions from the 25°C oxic microcosms were performed using ~0.4 g of subsample due to drying in those bottles. Samples were washed once with a sterile buffer (10 mM EDTA, 50 mM Tris-HCl, 50 mM Na_2HPO_4 at pH 8.0) (Zhou et al., 1996; Poulain et al., 2015) to remove potential PCR inhibitors and divalent cations. DNA was extracted using the DNeasy PowerSoil Kit ® according to the manufacturer's instructions (Qiagen, Germany) but centrifuge times were increased to one minute. Extractions were stored in a -20°C freezer for the first month, then -80°C subsequently. DNA purity was measured using a NanoDrop™ 2000 Spectrophotometer and DNA quantity was measured using a Quant-iT dsDNA assay kit (Invitrogen).

I estimated the abundance of specific cyanobacterial genes using the BioRad QX200 ddPCR system. Small DNA fragments preserve better than longer fragments and can be quantified using PCR technology (Jo et al., 2017; Wei et al., 2018; Xia et al., 2018). As such, the primer sets used in this study targeted gene fragments less than 300 bp in length. The ddPCR technology was used to quantify *Microcystis*-specific 16S rRNA (MICR), the allophycocyanin alpha chain blue-green pigment specific to *Synechococcus* (*apcA*), and *Anabeana*-specific 16S

rRNA (ANA) (primers are presented in Table 2.1). The later gene fragment was chosen to quantify the abundance of *Trichormus variabilis*, which is also commonly known as *Anabaena variabilis* (valid taxonomic synonym). Cyanobacteria 16S rRNA (CYA) was used as a proxy of the total amount of cyanobacteria in each microcosm. Estimates of CYA gene copy numbers included both the background native cyanobacterial populations as well as the spiked inoculated culture strains. The MICR and *apcA* target genes have been quantified in sediments going back decades (MICR: Pilon et al. 2019; Weisbrod et al., 2020, *apcA*: Dodsworth, 2020) and the CYA gene has been found in sediment layers dating back centuries (Pilon et al., 2019). A search through the literature yielded no results of any sedDNA study quantifying the ANA gene used in this experiment.

Each ddPCR reaction was prepared with 5 µL of extracted DNA, 6.04 µL of nuclease-free water, 0.23 µL of each primer, and 11.5 µL of the EvaGreen ddPCR Supermix. Droplets were generated using the manufacturer's guidelines except droplet generation oil volume was reduced to 65 µL. Thermocycling conditions included a 5 second step of 95°C, followed by 40 cycles of 95°C for 30 seconds with a 2°C/s ramp rate and a one minute annealing and extension step, followed by 5 minutes of 4°C and a final stabilization step of 90°C for 5 minutes. Annealing temperatures are outlined in Table 4.1. Absolute quantification of gene copy number per microliter of reaction was processed with the QuantaSoft Software (BioRad, Hercules, USA). Two negative controls using nuclease-free water and two positive controls using DNA extracted from the corresponding pure culture were included in each run. The copy numbers for all genes were normalized to copies per gram of dry sediment weight using: *(dd*4*EV*DF)/dw*, with *dd* being the number of copies per µL obtained from the ddPCR system, **4* since 5 µL of DNA were used in the transferred sample volume (20 µL) to generate the droplets, **EV* for the elution

volume, *DF* being the dilution factor used, and *dw* being the dry weight volume of the sediment. Gene copy numbers were normalized per gram of dry weight because water content differed between the microcosms throughout the course of the experiment. The average water content (%) in the microcosms under each condition at the different timepoints is presented in supplementary Table S2.2.

Estimates of the limit of detection (LOD) for each target gene were first calculated using the limit of the blank (LOB) (Armbruster and Pry, 2008), as similarly outlined in Taly et al. (2013) and Dong et al. (2018). The LOB was calculated by determining the mean and standard deviation (SD) of 20 no-template control samples and calculating the LOD using: $mean_{blanks} + 10(SD_{blanks})$. The LOD was established as approximately 1 copy/µL for each gene. This was confirmed using: $LOB + 1.645(SD_{low\ analyte\ samples})$ (as suggested by Armbruster and Pry; 2008). For this method, the LOB was calculated using: $mean_{blanks} + 1.645(SD_{blanks})$. Nine pooled DNA samples, diluted 100x for MICR and 1000x for *apcA*, ANA, and CYA, were used to calculate the SD of low analyte samples. Using the normalization equation described above, the LOD per gram of dry sediment for MICR, *apcA*, ANA, CYA was calculated as ~1.2×10^5, ~2.4×10^4, ~6.1×10^4, and ~3.1×10^6, respectively.

2.3.5 Data analysis

The degradation rate based on each target gene was estimated under the contrasting temperature/redox conditions from the decline in normalized gene copy numbers (GCN/g) over the entirety of the incubation experiment (one year) or until the last timepoint before the gene copy number average was below detection limits at a specific timepoint (n=9 for CYA, n=3 for MICR, *apcA*, and ANA). The use of exponential decay models to examine decay rates has been

proven successful in recent eDNA studies (Eichmiller et al. 2016; Kasai et al., 2020; Allan et al., 2021). Here, estimates of temporal change in sedDNA concentrations were quantified using a first-order exponential decay model:

$$C = C_0 e^{-kt}$$

where *k* is the decay rate constant, *C* is the concentration of the gene of interest (copies/g) at a specific timepoint (*t*), and C_0 refers to the concentration of the target gene at the start of the experiment (day 0). Because $C/C_0 = 1$ at $t = 0$, the y-intercept of the fitted line was 0. Two-tailed *t-tests* were performed to assess statistical differences between the decay rate constants of the oxic and anoxic microcosms at 4°C and at 25°C. Decay rate constants were subsequently plotted as a function of temperature, but because only two temperatures were used in this study, the relationship between temperature and decay rate constants cannot be statistically assessed.

Some treatments exhibited better fit when using a biphasic exponential decay model. As such, if there was enough data to determine a significant breakpoint, data were fitted instead using a biphasic model:

$$C = C_0 e^{-k_1 t'} e^{-k_2 t' - t'}$$

where k_1 refers to the first period of faster decay until the breakpoint occurred at time *t',* where then the second phase of slower decay (k_2) begins. The *segmented* package in R (R open-source Version 4.0.0 and RStudio software version 1.2.5042) was used to determine the breakpoints. The significance of the breakpoint was determined using the *davies.test* and confirmed using the *pscore.test.* Decay rate constants were again calculated either until the last day of the experiment or until the last timepoint before the average gene copy number was below detection limits (*n*=9 for CYA, *n*=3 for MICR, *apcA*, and ANA).

The number of days it would take for gene copy numbers to decline by 90% (T_{90}) was calculated using $T_{90} = -ln(1-0.9)/k$, where k refers to the calculated decay rate constant obtained from the one phase models. The estimated persistence time (in days) of each gene target in sediment was calculated using $t = -ln([C]/[C]_0)/k$, where k is the calculated decay rate, $[C]$ is the limit of detection for the gene of interest, and $[C]_0$ is the concentration of the sample that had the highest abundance (as outlined by Eichmiller et al., 2016).

2.4 Results

Total DNA content was tracked throughout the experiment. All microcosms (*Microcystis*-inoculated, *Synechococcus*-inoculated, *Trichormus*-inoculated, and experimental blanks) exhibited similar trends in declining DNA quantity (Figure 2.3), indicating that the addition of the different cell cultures did not affect the pattern of total DNA change. In each condition, DNA quantity declined rapidly in the first 0 to 14-days, but this was followed by a period of slower decay. When comparing DNA change overtime between the four environmental conditions, the extracted DNA quantity in the microcosms exposed to anoxic conditions decreased the least (Figure 2.3, bottom) (7.1×10^4 to ~1.5×10^4 ng/g), whereas microcosms under oxic 25°C conditions had the greatest decrease in DNA quantity throughout the experiment (Figure 2.3, top right) (7×10^4 to ~3×10^2 ng/g).

In the day zero experimental blank serum bottles (supplementary Figure S2.1), background CYA gene copy numbers were slightly greater than one order of magnitude above the corresponding LOD (~3.1×10^6 GCN/g) in both the oxic (1.6×10^8 GCN/g) and anoxic (5.5×10^7 GCN/g) microcosms. CYA copy numbers in the bottles were subsequently either at or below the LOD at each following time point. This was also observed with the *apcA* gene, where

gene copy numbers were ~10X greater than the LOD (~2.4 × 10^4 GCN/g) at time zero, but below the LOD subsequently. Some *Synechococcus* was expected in the blank bottles as they are commonly found in oligotrophic lakes (Pick, 1991). Bloom-forming taxa *Microcystis* and *Anabeana* were not detected in the blank microcosms, as all experimental blanks were consistently at or below detection limits for MICR (LOD: ~1.2 × 10^5 GCN/g) and ANA (LOD: ~6.1 × 10^4 GCN/g).

In the bottles inoculated with cyanobacterial strains, total CYA exhibited a pattern of log-linear decay with gene copy numbers consistently remaining above the corresponding LOD (~3.1 × 10^6 GCN/g) under all conditions except for those in the 25°C oxic environment (Figure S2.2, top). CYA gene copy numbers under 25°C oxic conditions were below detection limits after the 90-day timepoint. For the bottles inoculated with *Microcystis*, the MICR gene persisted for the entire experiment only under 25°C anoxic conditions, but gene copy numbers there were only slightly above the LOD (~1.2 × 10^5 GCN/g) after 90 days, reaching an average of ~3 ×10^5 GCN/g at 182 days and ~4 × 10^5 GCN/g at 365 days (Figure S2.2, bottom left). Although the abundance of MICR was highest under 4°C anoxic conditions for the first six months, gene copy numbers under this condition dropped below the corresponding LOD after that timepoint. Under 4°C oxic conditions, the MICR was below the LOD after 90 days. The calculated decay rate of MICR under 25°C oxic conditions may be unreliable as the average MICR gene copy numbers were below detection limits after the first timepoint (i.e., only the average at that one timepoint was used to calculate decay, since subsequent samples were removed from the model). In the microcosms inoculated with *Synechococcus* sp., *apcA* gene copy numbers in the 4°C anoxic bottles exhibited only a small level of degradation over the entire experiment (~3.5 × 10^6 down to ~4 × 10^5 GCN/g) and concentrations remained high by the 365-day timepoint (Figure S2.2,

bottom middle). In the 4°C oxic bottles, the average gene copy number declined from ~3.1 × 10^6 down to ~6.5 × 10^4 GCN/g. In the 25°C anoxic bottles, *apcA* concentrations declined two orders of magnitude over time (from ~3.5 × 10^6 down to ~3.2 × 10^4 GCN/g) and were only slightly above the corresponding LOD (~2.4 × 10^4 GCN/g) by the 365-day timepoint. Average *apcA* concentrations in the 25°C oxic bottles were below the LOD after the 90-day timepoint. For the microcosms inoculated with *Trichormus*, average ANA gene copy concentrations persisted for the entire experiment only under 4°C anoxic conditions. However, gene copy numbers there were only slightly above the corresponding LOD (~6.1 × 10^4 GCN/g) by the 365-day timepoint (~8.7 × 10^4 GCN/g) (Figure S2.2, bottom right). ANA gene copy numbers in the 25°C anoxic bottles had reached detection limits after the 182-day timepoint, after 90 days in the 4°C oxic bottles, and after 30 days in the 25°C oxic bottles.

Using a one-phase exponential decay model, decay rate constants (*k*) were calculated for each gene at each environmental condition. The rate of decline in gene copy numbers was different between the taxa (Table 2.2). For MICR, the *k* values were lower under anoxic conditions, and for *apcA*, under 4°C. For ANA, the one-phase decay model for the 25°C anoxic bottles had poor fit, and thus the *k* value there may not be reliable. However, based on the calculated *k* values for the 4°C oxic, 4°C anoxic, and 25°C oxic bottles, *Trichormus* decay rates also appeared to be temperature dependent. Overall, when comparing the three taxa, *Synechococcus* exhibited the lowest degradation rates followed by *Trichormus* then *Microcystis* (Figure 2.4 and Figure S2.3). Based on the results of the two-tailed *t*-tests (supplementary Table S2.3), *k* values between the oxic and anoxic microcosms were significantly different at both 4°C and at 25°C for all target genes, except for the CYA *k* value at 4°C. When *k* values were plotted as a function of temperature, positive trends with temperature (i.e., decay rate constants were

lower at 4°C and higher at 25°C temperatures) were observed only in the oxic microcosms (Figure S2.4).

Some data exhibited better fit when using a biphasic exponential decay model (Table 2.2). Except for those inoculated with *Trichormus*, all the 4°C microcosms had significant breakpoints and exhibited biphasic decay. The 25°C microcosms exhibiting significant breakpoints included the anoxic *Microcystis* bottles, as well as both the anoxic and oxic *Trichormus* and *Synechococcus* bottles. For all genes, the later the breakpoint was detected, the lower the k_1 value. When using these biphasic exponential models, *k* values again differed between the taxa. For CYA, the 4°C anoxic bottles had lower k_1 and k_2 than the 4°C oxic bottles. For MICR, the 25°C anoxic bottles had a much higher k_1 value than the 4°C microcosms. For *apcA*, the 25°C anoxic bottles had the lowest k_1 value, followed by the 4°C bottles, and then the 25°C oxic ones. For ANA, the calculated k_2-values were negative, indicating some growth during the early stages of the experiment. However, ANA were below detection limits after 30 days for the 25°C oxic bottles and after 182 days for the 25°C anoxic bottles (see Figure S2.5).

T_{90} and persistence times (Table 2.3) were reflective of the *k values* obtained from the one-phase exponential decay models. For CYA, *apcA*, and ANA, the calculated T_{90} and estimated persistence times were highest under 4°C anoxic conditions, and lowest under 25°C oxic conditions. For MICR, T_{90} and persistence were highest under 25°C anoxic and lowest at 25°C oxic conditions. However, rates were relatively similar between the 25°C anoxic and 4°C anoxic conditions.

2.5 Discussion

While the use of sedDNA for the reconstruction of historical cyanobacterial dynamics is increasing, there remains a need to examine the factors that can affect DNA degradation rates in sediment. This will be necessary for the accurate interpretation of population and community structure data from sedaDNA archives. Here, I investigated the effects of temperature and oxygen on gene fragment preservation of major cyanobacterial taxa in sediment using a microcosm approach and the results indicate variability in decay rate characteristics between *Microcystis*, *Synechococcus*, and *Trichormus*. Therefore, the null hypothesis that cyanobacterial taxa degrade in the same manner is rejected and differences at the cellular level among cyanobacteria taxa need to be considered when interpreting sediment archives.

2.5.1 The effects of oxygen and temperature on cyanobacterial sedDNA preservation

Both temperature and oxygen impacted cyanobacterial sedDNA degradation rates, but interestingly the effects differed between taxa. When comparing the three taxa under 4°C, *Synechococcus* had the lowest degradation rates followed by *Trichormus*, then *Microcystis*. While it is generally assumed that colder temperatures and the absence of oxygen are best for DNA preservation (Coolen and Overmann, 1998; Coolen et al., 2006; Corinaldesi et al., 2008; Domaizon et al., 2017), this was only true for *Synechococcus* and *Trichormus*. For *Microcystis*, the decay rate constants between the 4°C and 25°C anoxic microcosms were similar. This suggests that my hypothesis of cyanobacterial decay being only a function of the environmental conditions was partially incorrect, and that instead, decay is dependent on both environmental conditions and taxa preservation capacity.

For all four target genes, gene degradation was consistently slower under anoxic conditions in comparison to the oxic microcosms when exposed to the same temperature. This is consistent with other studies that show DNA degradation rates are lower in hypoxic or anoxic conditions (Corinaldesi et al., 2011; Jessen et al., 2017), and those that indicate preservation is enhanced in anoxic environments (Shi et al., 2003; Arndt et al., 2013). Oxygen may have contributed to higher decay rates by acting as a terminal electron acceptor during the oxidation of organic matter, leading to cell damage. Although only the decay of relatively small gene fragments was assessed in this study, fragmentation of cyanobacterial cells in response to oxidation has been observed in previous studies. Wert et al. (2013) used flow cytometry and chlorophyll-*a* analysis to compare cell degradation during the oxidation processes of drinking water treatment between *Microcystis aeruginosa* and filamentous *Oscillatoria* and *Lyngbya*. There, they found that *Microcystis aeruginosa* cells were more susceptible to oxidation degradation in comparison to the filamentous species (Wert et al., 2013).

The effects of temperature also differed between the taxa. This is consistent with a recent eDNA study analyzing the decay rates from diverse animal forms, where they found that not all eDNA decay was temperature dependent (Allan et al., 2021). While two temperatures points are not enough to statistically measure the effect of temperature, I found that decay rates were generally lower at 4°C, indicating that studies quantifying cyanobacterial gene concentrations from shallower or warmer sites, should be interpreted with caution. Although some studies have found that, in the dark, *Microcystis* cells have a greater tolerance to warm, oxic conditions in comparison to eukaryotic algae (Zhang et al., 2007), I found the MICR gene to be the least tolerant to warm, oxic environments in comparison to the other two cyanobacterial taxa assessed in this experiment. This suggests that *Microcystis* would not preserve well in shallow littoral

sediments, or warm shallow lakes where light reaches the bottom, and that estimates of *Microcystis* abundance from these environments might be underrepresented. However, *Microcystis* is colonial in nature and forms large aggregates (Deng et al., 2015; Verspagen et al., 2006); the *Microcystis* used in this study was not a colonial strain. More work is needed to evaluate how the mucilaginous colonial nature of *Microcystis* could influence sedaDNA degradation under these conditions.

Others have found that some cyanobacterial species can even grow under warm, dark conditions (Mannan and Pakrasi, 1993). *Trichormus* and other filamentous cyanobacteria are capable of heterotrophic growth in the dark when exposed to warm temperatures and are able to fix nitrogen aerobically when confined to differentiated cells called heterocysts (Fay, 1992; Gallon, 1992). Both potential heterotrophic growth and the transport of fixed nitrogen to other vegetative cells in the filament could slow down decay under warm temperatures. Although, here, when the cell cultures used for the experiments were observed under a light microscope, no akinetes or heterocysts were found, likely because of using cells in exponential growth with complete media (including nitrate and ammonium as a source of nitrogen). Under unfavourable conditions (absence of light and oxygen in the anoxic microcosms), it is, however, possible that in the microcosms inoculated with *T. variabilis*, some akinetes formed. Only under 4°C anoxic conditions did the ANA gene target persist for the entire 1-year experiment. Therefore, it is possible that for the microcosms in the other conditions (4°C oxic, 25°C anoxic, and 25°C oxic), the calculated decay rates were a result of rapid cell lysis, rapid degradation of DNA released from vegetative cells, and less DNA extracted from the thick cell-walled akinetes.

It is possible that the rates of decay were impacted by the target gene amplicon size. However, this is not likely to have affected results presented here. In this study, *apcA* had the

smallest fragment size (116 bp), followed by ANA (210 bp), then MICR (247 bp), and finally CYA (269 bp). If the rate of decay was a function of amplicon size, then *apcA* should have consistently had the lowest calculated decay rate and CYA would have had the greatest. Here, while *apcA* did generally have lower decay rates compared to the other target genes, ANA had the lowest rate of decay under 25°C anoxic conditions, and only under these conditions did CYA have the greatest calculated decay rate (Table 2.2). In addition, the fragment lengths of MICR, ANA, and CYA were relatively similar, yet in three of the four environmental conditions, MICR had much greater calculated decay rates compared to ANA and CYA (Figure S2.3). A recent study comparing different amplicon sizes found that calculated decay rates did not differ significantly between the different lengths assessed (96 bp, 285 bp, and 515 bp) (Bylemans et al., 2018). Although fragment length may have affected the rate of decay, differences in cell structure are probably what mainly explain the results found here.

2.5.2 Potential impacts of cyanobacterial taxa cell structure on sedDNA preservation

Microcystis, Synechococcus, and Trichormus differ in cell morphology and architecture. Overall, the lower *Synechococcus* decay rates suggest good preservation relative to *Microcystis* and *Trichormus.* The relative success of *Synechococcus* could be a combined result of their tiny size, their thick peptidoglycan layer, and efficient multifunctional envelope that allows for significant adaptation to changing environmental conditions (Callieri et al., 2012). While both *Microcystis* and *Synechococcus* have thick peptidoglycan layers (~10 – 20 nm in thickness) (Samuel et al., 2001; Skulberg et al., 1993) and have a hexagonal S-layer in the cell wall (Šmarda et al., 2002), *Synechococcus* is as distantly related to *Microcystis* as it is to *Nostocales* genera (Lyra et al., 2001). Some strains of *Synechococcus* have very thick cell envelopes

(consisting of a surface layer, outer membrane, peptidoglycan layer, and cell membrane) that are ~70 nm in diameter (Samuel et al., 2001). *Synechococcus* cells are also typically smaller (< 2 µm) than *Microcystis* cells (~2 – 7 µm) and are like small ball bearings and require tremendous pressure to physically lyse (> 16,000 PSI) (McMurter, 1991). For *Microcystis*, the cell wall ranges between 10 – 20 nm and cells are coated with thick proteinaceous material (Skulberg et al., 1993). For *Trichormus,* while no distinct S-layer is present in the cell wall (Šmarda et al., 2002), cells are surrounded by sheath material that is composed of pectic acids and acid mucopolysaccharides (Leak, 1967) that can slow down DNA degradation. In non-heterocyst forming filamentous bacteria, this outer membrane embeds a plasma membrane and a thick peptidoglycan layer surrounding the individual cells that ranges between 15 – 35 nm in thickness (Hoiczyk and Hansel, 2000; Wilk et al., 2011). Studies have also shown that the outer membrane curvature at the septum border between vegetative cells is further stabilized by intercellular connecting bridges (Wilk et al., 2011). This stabilization could have further slowed down DNA degradation.

2.5.3 The analysis and interpretation of cyanobacterial sedDNA archives

While there have been an increasing number of studies using sedDNA to track cyanobacterial dynamics, there are uncertainties regarding the interpretation of some results, especially since the use of sedDNA has not yet been fully validated. In a study analyzing cyanobacterial community changes from sedimentary archives collected from Swiss lakes, Monchamp et al. (2016) validated the approach of using sedDNA by comparing genetic data to long-term phytoplankton records collected over the past five decades. There, they found that reconstruction using DNA from sediment tracked the historical water data well overall in terms

of richness and taxonomic composition, implying little bias. However, when analyzing the water-sediment richness relationship at the level of order, weak but significant relationships were observed for mostly unicellular cyanobacteria (including Chroococcales and Synechococcales), but no relationship was observed with the Nostocales order (Monchamp et al., 2016). This indicates that preservation biases in sediment could be interfering with interpretation and that the estimates of the Nostocales order may be over or underrepresented in these sediment cores. However, further comparisons between community-based molecular data and surface water phytoplankton records are needed in order to verify this.

It should be noted that differences in settling rates and sediment deposition between cyanobacterial taxa would also strongly impact the interpretation of community data, but this too has been rarely addressed. A recent study found that the reconstruction of cyanobacterial dynamics using sedaDNA could be skewed and explained by differences in the deposition of freshwater cyanobacteria (Nwosu et al., 2021b). When comparing water column and sediment trap-based data to surficial sedaDNA data, Nwosu et al. (2021b) found that species of *Planktothrix* were more impacted by taphonomic processes within the water column (e.g., grazing by zooplankton) and post-depositional processes in comparison to aggregate- and colony-forming cyanobacterial taxa such as *Aphanizomenon* and *Snowella*.

Some studies analyze only one taxon, such as *Synechococcus*, and attempt to overcome the issue of potential differential preservation (and/or deposition) within sediment intervals by using relative abundance data (the ratio of *Synechococcus* abundance to the total cyanobacterial abundance) instead of absolute data (Domaizon et al., 2013). The use of relative abundance data, however, does not correct for preservation biases between cyanobacterial taxa. In addition, there is still some uncertainty in using relative abundance data using sedDNA (Mejbel et al., 2021b).

Interestingly, the high preservation capacity of *Synechococcus* has been noted in previous studies. In a study evaluating Holocene biological sources of bacteriohopanepolyols in the cold and euxinic (both anoxic and sulfidic) sediments of Ace Lake (Antarctica), stratigraphic analysis of the 16S rDNA showed that *Synechococcus* was present approximately 9400 years ago and the genus is still present today in the lake (Coolen et al., 2008). A *Synechococcus* strain found in the deepest sediment subsections of the 150-cm-long core showed a 99.8% sequence similarity (one ambiguous nucleotide position) to the 16S rDNA sequence of a pelagic *Synechococcus* strain. The presence of both *Synechococcus* sedaDNA and lipid biomarkers through the sediment archive demonstrated that at least the 16S rDNA of the identified *Synechococcus* phylotype was highly preserved under the cold and euxinic sediment conditions. DNA can be better preserved under euxinic conditions as the presence and accumulation of hydrogen sulfide reduces microbial activity and the bioavailability of organic matter, potentially reducing sedaDNA degradation (Coolen and Overmann, 1998; Coolen and Overmann, 2007; Coolen et al., 2008). More work, however, is needed to understand how other cyanobacterial taxa preserve under anoxic and sulfidic conditions and if there are differences in preservation capacity between taxa under these conditions.

2.5.4 The use of exponential decay models

Overall, 4°C anoxic conditions resulted in higher T_{90} and persistence times, but results were dependent on the calculated decay rate constants from the one-phase exponential decay model. The use of one-phase exponential models, however, oversimplifies the overall dynamics in natural lake environments and might not be the best model for analyzing sedDNA decay rates. In this study, those that exhibited significant breakpoints were also fitted using a biphasic

exponential decay model, which was found to better fit the data. Other recent studies have shown that decay experiments were better fitted using biphasic modelling and other models (for e.g., Weibull model) (Eichmiller et al., 2016; Bylemans et al., 2018; Allan et al., 2021). While sedDNA decay rates have rarely been reported in previous studies, the k_2 values obtained from the biphasic models were somewhat comparable to a recent study comparing sediment and environmental DNA decay rates of fish species (Sakata et al., 2020). However, both studies were performed under controlled settings, indicating that the decay rates observed in natural environments may not be fully represented.

In this study, the biphasic decay exhibited in some microcosms could have been the result of active or passive DNA extrusion and cell lysis, followed by some initial DNA fragmentation, then DNA stabilization, and subsequently DNA degradation in the sediment extracellular matrix. In other words, cell lysis and initial DNA fragmentation may have contributed to the first phase of rapid decay. It is expected that once the DNA was liberated from the cells and bound to the sediment, negatively charged phosphate groups on the DNA bound to the positively charged ions in the sediment, and formed cationic bridges, making the DNA stable for long-term analysis (Pietramellara et al., 2009). The second phase of slower decay in the sediment matrix then followed. Future work should investigate other filamentous and unicellular cyanobacteria and investigate different models of sedDNA decay rates to obtain a comprehensive understanding of the compounding effects of temperature, oxygen, and other abiotic factors.

2.5.5 Study limitations

The microcosm approach presented here offers new evidence on the variability in preservation capacity between major cyanobacterial groups in sediment and further elucidates

the impacts of temperature and oxygen on DNA decay rates. However, some limitations remain. The microcosms used for this experiment were not entirely representative of natural environments. Of perhaps greatest concern, was the level of evaporation in some microcosms. While I corrected for differences in moisture content by normalizing gene copy numbers per gram of dry weight, sediments are continuously wet in natural environments and the rates reported here (particularly for the 25°C microcosms) may be overestimating decay rate constants due to evaporation. The taxa used in this study have been recorded in sediments going back centuries (Savichtcheva et al., 2011; Monchamp et al., 2016; Pilon et al., 2019), yet the abundances of some genes were already below or near detection limits by the end of the 1-year experiment performed here. Larger microcosms with a greater amount of wet sediment and water volume may be necessary for future studies aiming to analyze sedDNA decay rates. Alternatively, gas-tight syringes and a valve system can be applied to add oxygen into the oxic microcosms, such that the bottles can retain moisture in sealed bottles (similar to the anoxic microcosm set-up). Nonetheless, the microcosms used here can be used to approximate the natural environment of some weather events or disturbances. For instance, data from the 25°C oxic microcosms can be used to represent an environment experiencing a drought and results can help further clarify how rapidly sedDNA from these three cyanobacterial species degrades under these conditions. Additionally, the culture media used to prepare the cell cultures was rich in nutrients (phosphorus and nitrogen) and the inoculation of these cultures within the microcosms may have increased the decay rates. Increased microbial activity is likely to lead to rapid ell lysis and a greater level of DNase production. Studies aiming to further analyze cyanobacterial decay rates in sediment should rinse cells with nitrogen- and phosphorus- free media or use a buffer without nutrients to analyse long-term trends.

Although alterations to the microcosm setup may further illuminate the current understanding of cyanobacterial sedDNA decay rates, reproducing the behaviour of any particular natural system is difficult. Natural lake environments differ greatly from one another and change over time. Biological (e.g., migration or shedding differences) and physical (e.g., sedimentation rate or climate) processes occur naturally and can interfere with sedDNA distribution and preservation in sediment. For instance, a shallow eutrophic lake may not be anoxic at the bottom of the water column, but a high sedimentation rate could lead to the quick burial of sediment DNA, meaning only minimal exposure to oxygen. Additionally, sediment chemistry and the level of surface-reactive particles (such as clay or humic substances) could impact DNA degradation rates. The sediments used in this study were highly organic and the clay fraction was small. A different study using sediments with a higher number of surface-reactive particles to adsorb nucleic acids will likely produce lower DNA decay rates. Other factors not evaluated in this study include processes within the water column, such as predation, cell lysis from viruses, and heterotrophic microbial decomposition (Pederson et al., 2015; Ellegaard et al., 2018; Nwosu et al., 2021b), which could affect the sediment deposition of cyanobacteria and hence their sedaDNA. Overall, microcosms used here were not meant to reproduce any particular lake system, but rather to gain a better understanding of potential preservation differences in sediment among cyanobacterial taxa.

2.5.6 Implications and conclusions

I found that differential preservation of DNA in sediment exists between major cyanobacterial taxa and that the effects of temperature and oxygen differ may also differ between these groups. In addition, the results demonstrated that data from studies where sediment DNA is

collected from shallow or warm lakes, should be interpreted with caution. These findings can help in the interpretation of cyanobacterial population and community structure data based on sedDNA. This is of particular importance as more studies are using sedDNA to understand the factors leading to harmful cyanobacterial bloom formation. This field is still in its infancy and more work is needed to analyse sedDNA decay rates between species to validate the use of sedDNA interpretation analyses. The influence of biological and physical processes, both in the water column and in the sediments, on sedaDNA archives should be evaluated in future studies.

Supplementary material is supplied in Appendix A.

Acknowledgements: Thank you to Sheelagh Ogilvie and Ogilvie's Cottages, Laurentian Hills, Ontario, for kindly allowing us to collect sediment from their lake and for the use of their boat. I also thank Heather Roshan from the Canadian Phycological Culture Centre (University of Waterloo, Ontario) for her technical expertise with the cell cultures, as well as Danielle Fortin and Tarek Najem (University of Ottawa) for advice and the use of their anaerobic chamber. This research was financially supported by an NSERC Discovery grant (RGPIN-2018-06356) and by the NSERC CREATE program.

Table 2.1 PCR primers used for gene quantification with droplet digital PCR for cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA).

Target	Primer	Primer sequence (5'-3')	Amplicon length (bp)	Annealing T (°C)	Dilution factor	Source
CYA	108F	ACG GGT GAG TAA CRC GTR A	269	56	256x	Rinta-Kanto et al., 2005
	377R	CCA TGG CGG AAA ATT CCC C				
MICR	184F	GCC GCR AGG TGA AAM CTA A	247	58	10x	Nübel et al., 1997
	431R	AAT CCA AAR ACC TTC CTC CC				
apcA	F1	ATG AGC ATC GTC TCC AAC TCG	116	60	2x	Dodsworth, 2020
	R7	CMC GRC TCT CGC TVA GAA CCT				
ANA	573F	AGT GGA AAC TAC AAA GCT AGA GTT	210	52	5x	Doblin et al., 2007
	780R	CTT GGG TCG ATA CGA GCT				

Table 2.2 Decay rate characteristics using one-phase and biphasic exponential decay models for cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at each condition. Decay rates were calculated using the average of 9 samples for CYA, and 3 samples each for MICR, *apcA*, and ANA.

		One-phase exponential decay model			**Biphasic exponential decay model**			
Target	Condition	*K* (day-$^{-1}$)	*SE*	R^2	k_1 ± SE (day^{-1})	k_2 ± SE (day^{-1})	Breakpoint ± SE (days)	R^2
CYA	25°C anoxic	0.020	± 0.003	0.86	-	-	-	-
	25°C oxic	0.059	± 0.01	0.87	-	-	-	-
	4°C anoxic	0.016	± 0.004	0.72	0.068 (± 0.02)	0.002 (± 0.003)	58.1 (± 20)	0.95
	4°C oxic	0.020	± 0.003	0.88	0.11 (± 0.05)	0.013 (± 0.003)	20.1 (± 11)	0.96
MICR	25°C anoxic	0.017	± 0.006	0.48	2.2 (± 0.6)	0.002 (± 0.002)	1.7 (± 0.5)	0.97
	25°C oxic	2.7	± 0[a]	1.00	-	-	-	-
	4°C anoxic	0.019	± 0.004	0.80	0.54 (± 0.3)	0.01 (± 0.003)	2.3 (± 2)	0.96
	4°C oxic	0.058	± 0.02	0.62	0.28 (± 0.2)	0.006 (± 0.003)	13.2 (± 9)	0.99
apcA	25°C anoxic	0.015	± 0.002	0.85	0.028 (± 0.005)	0.001 (± 0.004)	151.2 (± 38)	0.96
	25°C oxic	0.049	± 0.02	0.63	1.7 (± 0.3)	0.01 (± 0.005)	1.4 (± 0.3)	0.99
	4°C anoxic	0.0066	± 0.001	0.81	0.12 (± 0.1)	0.004 (± 0.0005)	6.6 (± 8)	0.98
	4°C oxic	0.012	± 0.002	0.83	0.089 (± 0.007)	0.007 (± 0.0004)	18.5 (± 2)	0.99
ANA	25°C anoxic	0.013	± 0.006	0.35	0.46 (± 0.49)	-0.003 (± 0.004)	4.7 (± 5)	0.91
	25°C oxic	0.092	± 0.03	0.73	0.90 (± 0.0001)	-0.02 (± 0.0001)	2.9 (± 0.001)	0.99
	4°C anoxic	0.0083	± 0.0008	0.94	-	-	-	-
	4°C oxic	0.029	± 0.005	0.84	-	-	-	-

[a]Reached limit of detection after the first timepoint (one day). As such, the decay rate was calculated based on only the average at the one-day timepoint.
"-" Denotes no significant breakpoint

Table 2.3 Time taken for gene copy numbers to decay by 90 % (T_{90} [days]) and the persistence time (days) for cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at each environmental condition given the calculated decay rate constants obtained from the one-phase models.

Target	Condition	T_{90} (days)	Estimated persistence time (days)
CYA	25°C anoxic	115	437
	25°C oxic	39	147
	4°C anoxic	144	546
	4°C oxic	121	432
MICR	25°C anoxic	135	303
	25°C oxic	0.85	1.8
	4°C anoxic	121	271
	4°C oxic	40	84
apcA	25°C anoxic	154	332
	25°C oxic	47	99
	4°C anoxic	349	754
	4°C oxic	192	403
ANA	25°C anoxic	177	242
	25°C oxic	25	32
	4°C anoxic	277	379
	4°C oxic	79	100

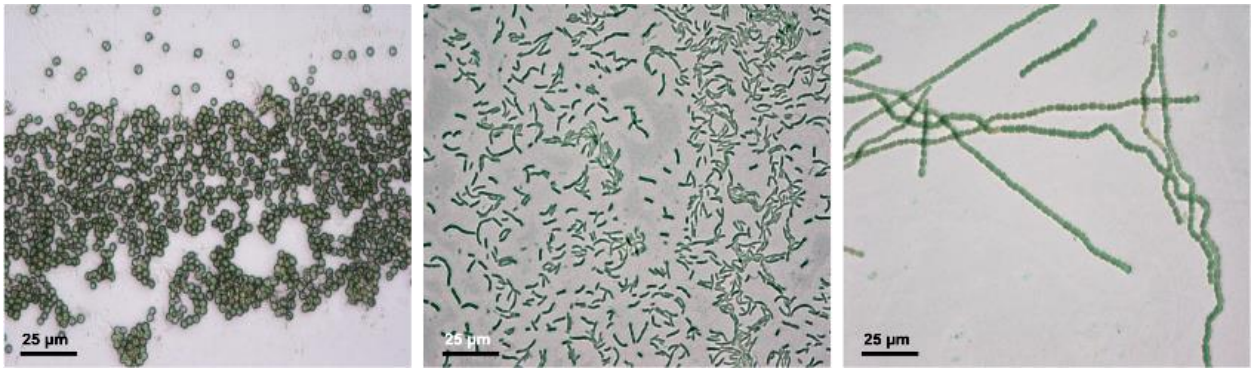

Figure 2.1 *Microcystis aeruginosa* (left), *Synechococcus* sp. (centre), and *Trichormus variabilis* (also known as *Anabaena variabilis*; formerly *Anabaena flos-aquae*) (right) under 400X magnification. Pictures were taken using an Olympus CX41 Compound Microscope and Infinity Capture software.

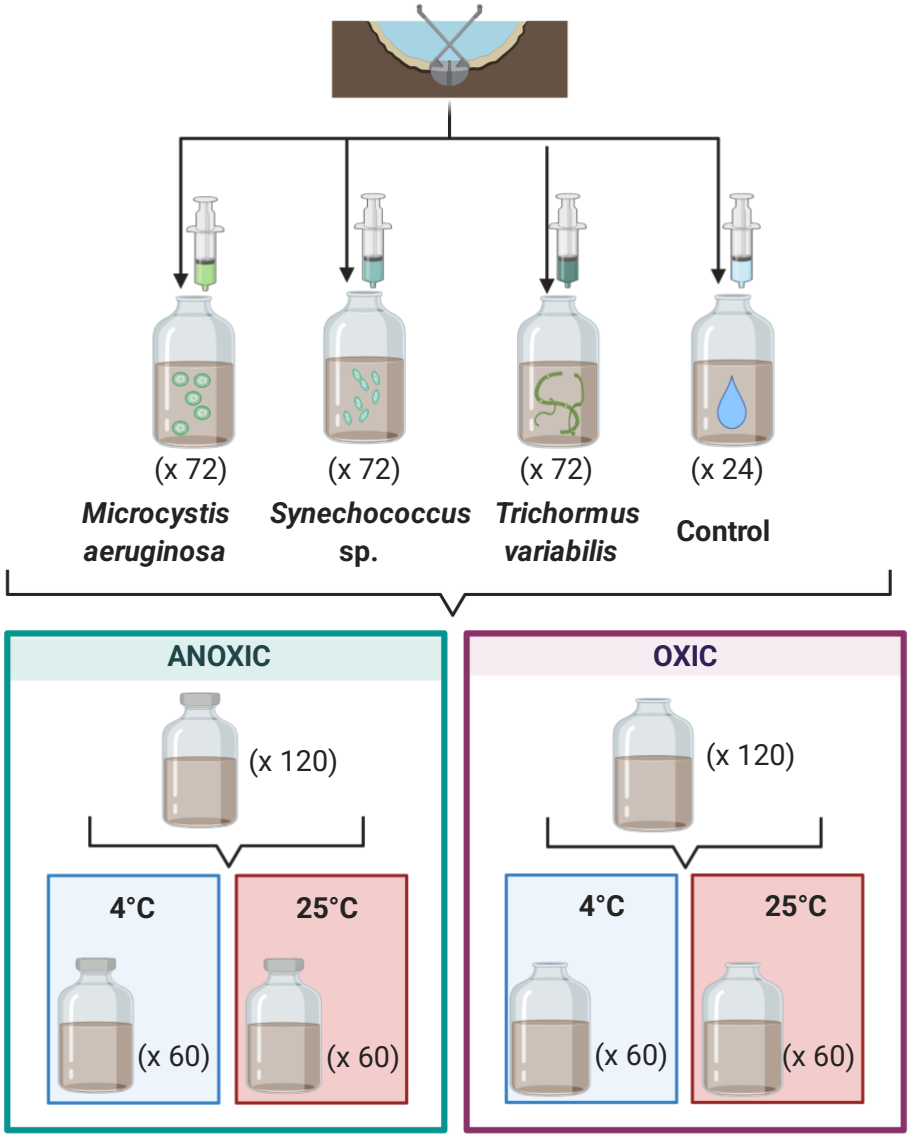

Figure 2.2 Schematic of the experimental design where the number in the brackets indicates the associated number of serum bottles. An additional 40 serum bottles (not shown here) were prepared: 20 to determine the absolute day 0 loading abundances in the oxic and anoxic bottles (3 of each species and 1 experimental blank per redox condition) and 20 to measure changes in sediment mass and water content. Pure cultures were obtained from the Canadian Phycological Culture Centre: *Microcystis aeruginosa* (CPCC 300), *Synechococcus* sp. (CPCC 663), and *Trichormus variabilis* (CPCC 67). Figure created through BioRender.

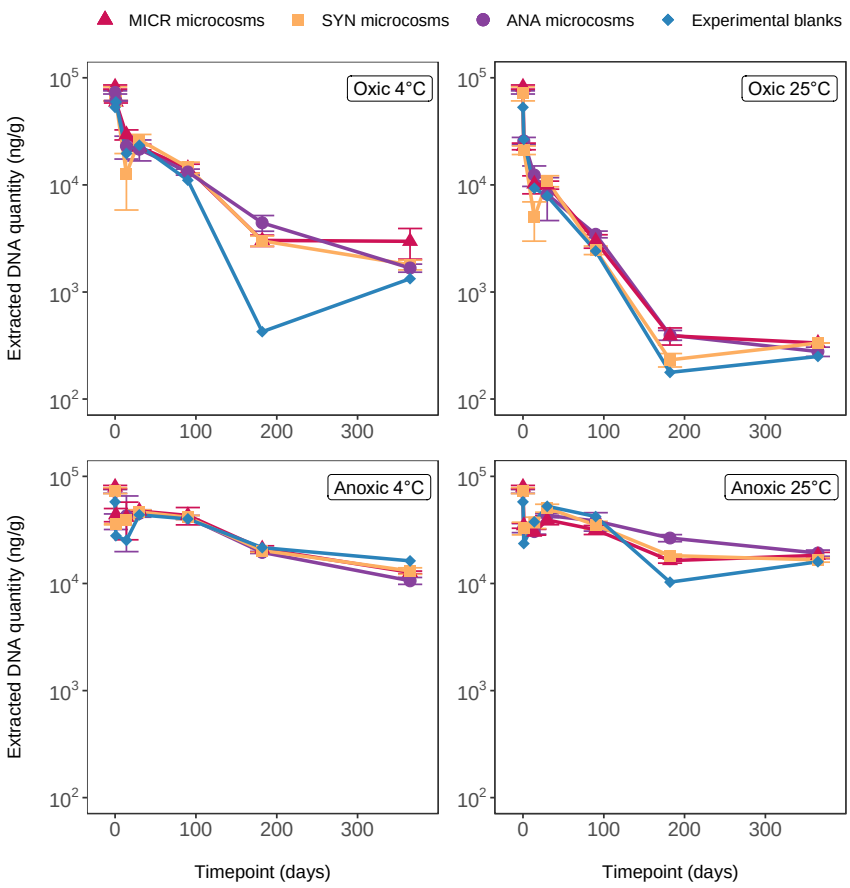

Figure 2.3 Mean (± standard error) DNA concentrations expressed per gram of dry sediment in the microcosms inoculated with *Microcystis aeruginosa* (MICR), *Synechococcus* sp. (SYN), *Trichormus variabilis* (ANA) (*n*=3 per each group), and in the experimental blanks under the four environmental conditions tested in this study; 4°C oxic (top left), 25°C oxic (top right), 4°C anoxic (bottom left), and 25°C anoxic (bottom right).

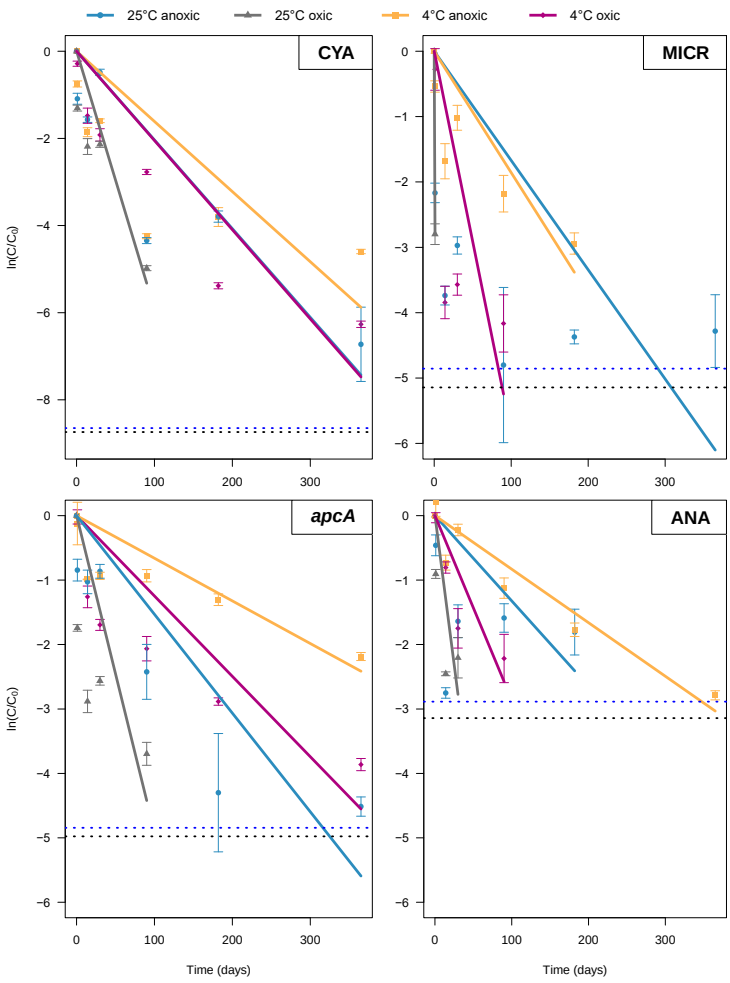

Figure 2.4 One-phase exponential decay curves for cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at each environmental condition. The solid lines show the decay curves, and the blue and black horizontal dotted lines are the detection limits for each target gene in the oxic and anoxic microcosms, respectively. The error bars represent the standard error of the mean (n=9 for CYA, n=3 for MICR, *apcA*, and ANA). Data where the average gene concentration were below the corresponding detection limits are not shown. Note the CYA plot (top left) has a different y-axis scale.

CHAPTER 3: Comparing quantitative methods for analyzing sediment DNA records of cyanobacteria in experimental and reference lakes

A version of this chapter has been published in
Frontiers in Microbiology

Mejbel, H.S.[1], Dodsworth, W.[1], Baud, A.[2,3], Gregory-Eaves, I.[2,3], and Pick, F.R[1]. (2021). Comparing quantitative methods for analyzing sediment DNA records of cyanobacteria in experimental and reference lakes. *Front. Microbiol.* 12:669910. doi: 10.3389/fmicb.2021.669910

[1]Centre for Advanced Research in Environmental Genomics, Department of Biology, University of Ottawa, Ottawa, ON, Canada

[2]Department of Biology, McGill University, Montréal, QC, Canada

[3]Groupe de Recherche Interuniversitaire en Limnologie

Author Contributions
HSM and FRP designed the study. HSM and AB sectioned the sediment cores. IG-E provided sterile space and assistance with sediment-sample preparation. AB performed the varve counting. HSM performed the molecular experiments, analyzed the data, and wrote the manuscript. WD, AB, IG-E, and FRP revised the manuscript. All authors contributed to the article and approved the submitted version.

Keywords*:* droplet digital PCR, high-throughput sequencing, metabarcoding, paleolimnology, sedDNA, targeted PCR

3.1 Abstract

Sediment DNA analyses are rapidly emerging as powerful tools for the reconstruction of environmental and evolutionary change. While there are an increasing number of studies using molecular genetic approaches to track changes over time, few studies have compared the coherence between quantitative PCR methods and metabarcoding techniques. Primer specificity, bioinformatic analyses, and PCR inhibitors in sediments could affect the quantitative data obtained from these approaches. I compared the performance of droplet digital PCR (ddPCR) and high-throughput sequencing (HTS) for the quantification of target genes of cyanobacteria in lake sediments and tested whether the two techniques similarly reveal expected patterns through time. Absolute concentrations of the cyanobacterial 16S rRNA gene were compared between ddPCR and HTS using dated sediment cores collected from two experimental (Lake 227, fertilized since 1969 and Lake 223, acidified from 1976–1983) and two reference lakes (Lakes 224 and 442) in the Experimental Lakes Area, Canada. Relative abundances of *Microcystis* 16S rRNA genes were also compared between the two methods. Moderate to strong positive correlations were found between the molecular approaches among all four cores but results from ddPCR were more consistent with the known history of lake manipulations. A 100-fold increase in ddPCR estimates of cyanobacterial gene abundance beginning in ~1968 occurred in Lake 227, in keeping with experimental addition of nutrients and increase in planktonic cyanobacteria. In contrast, no significant rise in cyanobacterial abundance associated with lake fertilization was observed with HTS. Relative abundances of *Microcystis* between the two techniques showed moderate to strong levels of coherence in top intervals of the sediment cores. Both ddPCR and HTS approaches are suitable for sediment DNA analysis, but studies aiming to quantify absolute abundances from complex environments should consider using ddPCR due to its high tolerance to PCR inhibitors.

3.2 Introduction

Lake sediments act as natural archives through the accumulation and preservation of organic and inorganic matter arising from the lake as well as its watershed and airshed. Changes within lake ecosystems, whether natural or anthropogenic, are stored in sediment profiles and can provide long-term historical records of environmental change (Smol, 2008). Because of the general lack of long-term monitoring records, lake sediments can be used to help identify the drivers of environmental change, such as climate change or nutrient loading. Typically, paleolimnological studies rely on subfossils or distinct chemicals preserved in sediment as proxies of past environments (Smol, 2008). However, not all organisms preserve well in sediment or leave behind a distinct biomarker, making it difficult to understand and track long-term responses to environmental change. As an alternate and complementary method to classical paleolimnological proxies, various DNA-based methods have emerged (Domaizon et al., 2017).

The recent advances and application of molecular approaches using DNA archived in sediment provide for considerable expansion in this field, particularly for analyses of microbial communities. Quantitative PCR techniques have shown to be successful in quantifying specific targets from sediment DNA (sedDNA) including bacteria, viruses, and other microorganisms that lack morphological features (Coolen et al., 2006; Savichtcheva et al., 2011; Brazeau et al., 2013; Pal et al., 2015; Pilon et al., 2019). Other, community-wide analyses have also been proven to track and describe community dynamics through time (Monchamp et al., 2016, 2018; Legrand et al., 2017; Pal et al., 2019). While there have been a few studies that have used a combination of molecular techniques on sedDNA, predominantly PCR-directed methods (Singh et al., 2017; Hamaguchi et al., 2018), there remains a lack of research directly examining the coherence and correlation between quantitative methods. This is of particular concern in the

analysis of sediment samples where PCR inhibitors (such as humic acids) can be found at high levels, which can then interfere with the sensitivity of the technology used and/or produce false results (Rochelle et al., 1992; Savichtcheva et al., 2011; Albers et al., 2013). Although inhibition is an issue for all PCR methods, diluting samples can sometimes overcome these limitations, yet this is not always possible or suitable when calculating concentrations of low-level test samples. Alternate methods should be applied as some PCR techniques have been found to be more tolerant to inhibitors than others (Taylor et al., 2017, 2019; Wood et al., 2019).

Quantitative studies have mainly used quantitative real-time polymerase chain reaction (qPCR) to obtain relative abundances of target genes from environmental samples (Nathan et al., 2014; Pal et al., 2015; Pilon et al., 2019). Although qPCR does have advantages, the need to run samples in triplicate along with ensuring the assay shows reliable amplification can make it challenging to quantify targets from complex environments (Taylor et al., 2019) such as sediments. With the development of droplet digital polymerase chain reaction (ddPCR), these limitations are reduced. ddPCR technology uses water-oil emulsion chemistry to split the PCR reaction into 20,000 individual vessels, or droplets, before amplification, reducing the normalization and calibration issues of qPCR (Hindson et al., 2011; Taylor et al., 2017). Studies that have described the relationship between qPCR and ddPCR techniques using sediment samples found that ddPCR significantly improved both the sensitivity and the detection of low concentrations of target genes (Singh et al., 2017; Hamaguchi et al., 2018).

In addition to targeted molecular approaches, there has been a recent emergence of paleolimnological studies using metabarcoding techniques on sediment samples for community-wide analyses (Monchamp et al., 2016; Legrand et al., 2017; Pilon et al., 2019). High-throughput sequencing (HTS) provides for a cost-effective alternative to traditional paleolimnological

methods by enabling simultaneous taxonomic identifications that probe sediment diversity, at times with increased specificity (Parducci et al., 2017; Tse et al., 2018). While there is still some uncertainty in interpretating the relative abundance data generated using metabarcoding approaches (Taberlet et al., 2018), some investigators are also exploring whether absolute read counts provide useful information (Lamb et al., 2018; Bylemans et al., 2019). Quantitative data using read counts obtained from metabarcoding has been suggested as an approach to monitor species interactions and distributions in support of management efforts (Hänfling et al., 2016; Ushio et al., 2018), but additional calibration is required to evaluate the quantitative performance of metabarcoding techniques. Some factors that can affect quantification data from HTS include the choice of reference database, primer biases that lead to unequal amplification (e.g., Hong et al., 2009; Shelton et al., 2016), and the bioinformatic pipeline chosen for analysis.

The performance between targeted and metabarcoding approaches using read counts has been evaluated in recent contemporary environmental DNA (eDNA) literature (Harper et al., 2018; Lamb et al., 2018; Bylemans et al., 2019). These studies found targeted PCR approaches were more sensitive with greater detection than eDNA metabarcoding methods. Although, others have found strong positive relationships between qPCR, ddPCR, and HTS technologies when using eDNA from seawater and biofouling samples (Wood et al., 2019). Quantitative outputs using absolute abundances from targeted PCR approaches (such as ddPCR) and read counts from metabarcoding approaches (such as HTS) has, to my knowledge, never been compared over timescales of decades using sedDNA. Furthermore, a greater amount of humic substances are often found in recent (or top) sediments (Calace et al., 2006; Hou et al., 2014) and these materials can affect the sensitivity of different genetic methods, and consequently the agreement

among them. The physical and chemical changes of sediments through time (i.e., diagenesis) could also affect differentially how well the two methods amplify target genes.

In this study, I evaluated the coherence between ddPCR and HTS methods using DNA extracted from lake sediments at the Experimental lakes area (ELA). ELA, located in Central Canada, has been the site of several long-term, ecosystem-level experimental manipulations (e.g., Schindler et al., 1990, 1996) and presents a unique opportunity to compare paleolimnological records across experimental and reference lakes. The latter have not been manipulated experimentally nor impacted by watershed land-use change. However, reference lakes may still be subject to climatic and atmospheric changes or the occasional forest fire. Furthermore, long-term monitoring records are available for many ELA lakes which enables validation of sediment proxies (e.g., Leavitt and Findlay, 1994). Cyanobacteria were selected as the ideal test organisms for this study as lake manipulation at several ELA sites has led to significant changes in cyanobacterial abundance and composition. I chose Lake 227 as it has experienced continued experimental fertilization since 1969 with a well-documented and concomitant increase in the abundance and dominance of cyanobacteria (Schindler, 1974; Anderson et al., 1987; Schindler et al., 1990; Findlay and Kasian, 1996). Cyanobacterial dynamics have also been recorded in a sediment core from Lake 227 through pigment analysis (Leavitt and Findlay, 1994). Lake 223 is another experimental lake where sulfuric acid was applied to the lake between 1976 until 1983 to simulate the effects of acid precipitation (Schindler et al., 1980; Findlay and Kasian, 1986). In Lake 223, phytoplankton biomass increased under acidification, but since ~1994, the lake has largely recovered to pre-manipulation conditions (Findlay and Kasian, 1996). Here, my objectives were to (1) measure the level of agreement between absolute cyanobacteria counts derived from HTS against absolute

concentrations derived from ddPCR, (2) compare relative abundances of *Microcystis* (a common genus of bloom-forming cyanobacteria) between the two methods, and (3) analyse the relationship between ddPCR and HTS from top and bottom sediments. I tested the hypothesis that ddPCR and HTS reflect similar patterns within each time series but that the level of coherence between the two methods would be weaker in top sediments as a result of potential organic matter interference.

3.3 Materials and methods

3.3.1 Study sites

Four lakes at ELA were used to determine the coherence between ddPCR analyses of gene copy numbers and HTS analyses of cyanobacterial gene targets. Lake 227 (L227) is an experimentally manipulated lake where nitrogen and phosphorus were added to the lake between 1969 and 1989. From 1990 onwards, only phosphorus has been continuously added (Schindler et al., 2008). Experimentation on Lake 223 (L223) began in 1976 when H_2SO_4 was added into the lake to drop the pH from 6.7 to 5.1 (Schindler et al., 1980; Cruikshank, 1984; Findlay and Kasian, 1986). Experimentation ended in 1983 to allow pH levels to naturally recover (Findlay and Kasian, 1990). Lakes 224 (L224) and 442 (L442) have been not been subjected to eutrophication nor acidification experiments and have served as reference lakes to manipulated lakes in previous studies (Findlay and Kasian, 1996; Higgins et al., 2018; Lamothe et al., 2018). L224 was subject to short radioisotope experiments in 1976 (Hesslein et al., 1980a,b; Quay et al., 1980; Crusius and Anderson, 1995). A map of ELA, lake coordinates, morphology, and lake manipulation details are presented in supplementary data (Figure S3.1, Table S3.1).

3.3.2 Sediment collection and sectioning

In March 2018, four sediment cores were collected from the maximum depths of lakes 227, 223, 224, and 442 using a 1.2 x 0.1 x 0.2-meter aluminum freeze corer (collected by the International Institute for Sustainable Development–Experimental Lakes Area [IISD–ELA], Table S3.2). Because of the flocculent nature of sediments in L227, freeze cores are a preferred way to obtain sediment cores from this lake (Michael Paterson, IISD-ELA, personal communication) and the same device was therefore used for all the lakes. Freeze coring is advantageous in obtaining undisturbed sediment, unaffected by the degassing of gas bubbles occurring from a change in hydrostatic pressure (Verschuren, 2000). The corer was filled with crushed dry ice and methanol and lowered through a hole in the ice. After approximately 20 minutes, the freeze cores were brought to the surface and, using a putty scraper, the excess unfrozen sediment was removed. A hammer and chisel were used to free the edges of the frozen sediment core from the freeze corer. The cores were then wrapped in cellophane, packed with dry ice and shipped to McGill University (Montréal, Québec) for subsectioning in the laboratory of I. Gregory-Eaves (Department of Biology).

In a cool sterile room, a rotary tool was used to subsample the sediment slab into 3 cm working intervals. Sterilized ceramic knives were used to scratch off a very thin surface layer from the face of the 3 cm intervals, moving horizontally across the core surface. Using the rotary tool, and following the line indicated by a laser and plastic meterstick, 0.5 cm subsections were created by making a shallow incision in the sediment. A ceramic knife was then inserted horizontally across the incision and tapped on using a sterilized hockey puck to cleave the slice. The workspace, ceramic knives and rotary tool blades were sterilized with 10% bleach, rinsed with deionized water, and dried with KimWipes® between each subsection. This process was

continued until each 3 cm interval block along the core was subsectioned. Every other 0.5-cm subsection was collected for molecular analyses while remaining samples were collected for other applications. Samples were stored in 5-mL sterile screwcap centrifuge tubes (VWR) then frozen at -20°C. All subsequent molecular analysis steps (extractions, optimization, and PCRs) were conducted in a separate and sterile part of the Core Molecular Biology and Genomics Laboratory at the University of Ottawa (Ottawa, Ontario).

3.3.3 DNA extraction

Sediment DNA was extracted using the DNeasy PowerSoil Kit according to the manufacturer's instructions (Qiagen, Germany). Minor modifications were made to the protocol. Lake sediment samples, ranging from 0.5 g to 0.7 g, were added to emptied PowerBead tubes. If DNA yield and quality did not meet standards, wet sediment weight was increased to 1 g. Subsection intervals from L224 and L442 were washed twice, and intervals from L227 and L223 were washed three times using 250 µL of a wash buffer (pH 8.0) found to improve DNA quality and quantity yield by removing humic compounds and divalent cations (Zhou et al., 1996; Poulain et al. 2015). Samples were vortexed for three minutes at maximum speed on a Vortex-Genie 2 (Scientific Industries Inc., USA), followed by a centrifuging step at 3000 r.p.m. for three minutes. After discarding the supernatant from the final washing step, the original PowerBead tube contents (glass beads, salts, and buffers) were re-added to the tubes. The extraction kit protocol was then carried out, but centrifuge time was increased to 1 minute at each centrifuge step. The purity of the DNA was measured using a NanoDrop^{TM} 2000 Spectrophotometer (Thermo Scientific, USA) and estimated based on the ratios of absorbance at 260 and 280 nm, and at 260 and 230 nm. DNA was considered pure if the ratios were at or around 1.8. Lower

ratios indicate some protein contamination (if 260:280 ratios are low) or organic contamination (if 260:230 ratios are low) (Yeates et al., 1997). DNA quantity was measured through a Quant-iT dsDNA assay kit (Invitrogen). Extractions were repeated with a decreased elution volume of 50 µL if low DNA yield and purity persisted. Final DNA extraction solutions were stored at -20°C for downstream applications and then -80°C for long-term storage.

3.3.4 Sediment dating

For the creation of the dating profiles, sediment samples (n=12–15) were homogenized, freeze-dried at -50°C, and their water content measured. Samples were then placed in an Ortec High Purity Germanium Gamma Spectrophotometer (Model GWL-120230, Oak Ridge, TN, USA) at the Laboratory for the Analysis of Natural and Synthetic Environmental Toxins (LANSET) core facility to measure ^{210}Pb, ^{137}Cs, and ^{226}Ra activity. Constant Rate of Supply (CRS) models were chosen for all cores and cross-calibrated against the ^{137}Cs peak (Blais et al., 1995). Sediment profiles of ^{210}Pb and ^{137}Cs data are presented in Figure S3.2. The peak from the ^{137}Cs was used as an independent marker and estimated to occur in 1963, when maximum nuclear weapon testing was in effect before atmospheric testing became prohibited.

Additional measures were taken with the L227 record to ensure the robustness of the chronology because ^{226}Ra was added to L227 in 1970 (Emerson and Hesslein, 1973), which is mobile in sediment and degrades to ^{210}Pb, potentially hindering accurate dating profiles. As a first step, ^{210}Pb excess activity was calculated through point-by-point subtractions of the supported activity from the total activity. To corroborate and confirm the ^{210}Pb dating, annual lamination (i.e., varve) counting was also conducted based on a high-resolution photograph taken by the Itrax XRF cores scanner housed at L'Institut National de la Recherche Scientifique in Québec.

3.3.5 Pre-droplet digital PCR optimization

Prior to running ddPCR, primers for each target gene were validated through qPCR (as suggested by Taylor et al., 2017). The cyanobacterial target gene chosen for analysis included cyanobacterial 16S rRNA (CYA; universal gene essential for DNA translation), as a proxy for the total abundance of cyanobacteria. The primer set chosen amplifies the V1-V3 regions of the 16S rRNA gene conserved in cyanobacterial species (Nübel et al., 1997; Rinta-Kanto et al., 2005). In addition, *Microcystis*-specific 16S rRNA (MICR) was also targeted to allow for testing the limits of the methods. *Microcystis* is a widely distributed and well-studied freshwater genus (Harke et al., 2016) that occurs in ELA lakes at low concentrations. Primers and associated sequences are listed in Table 3.1. All reactions were cycled in a Bio-Rad CFX96 Real-Time PCR Detection module and analyzed using Bio-Rad CFX Manager Software version 3.0. Each PCR mix contained 4 µL of 10x diluted pooled DNA (separate pools were made for each sediment core), 0.2 µL of nuclease-free water, 0.4 µL of each primer (forward and reverse), and 5 µL of the 2x SsoFast/Sso Advanced EvaGreen Supermix, brought to a final volume of 10 µL. The samples were then loaded in 0.2 mL 8-strip tubes in triplicate, run alongside one positive and one negative control sample. For the positive controls, 4 µL of DNA extracted from pure cultures of the microcystin-producing *Microcystis aeruginosa* strain CPCC 300 was used (CPCC300, Canadian Phycological Culture Center, Canada). Nuclease-free water was used for the negative controls. The annealing temperature for each primer set was first optimized using a thermal gradient. The qPCR cycle program consisted of an initial preheating step of 95°C for 3 minutes, followed by 40 cycles of denaturation at 95°C for 10 seconds, annealing for 20 seconds, and extension at 72°C for 20 seconds. The melt curve protocol followed with 10 seconds at 95°C and then 5 seconds each at 0.5°C increments between 60°C to 95°C.

The qPCR amplification products were then separated through 1.5% agarose gels at 100V in 1x TAE buffer for approximately 30 to 45 minutes. Products were then visualized using an MBI Lab Equipment Fusion UV visualizer cabinet (Montreal-Biotech, Canada) and captured using Fusion Molecular Imaging Software.

3.3.6 ddPCR

ddPCR was used to amplify and quantify CYA and MICR. Prior to running ddPCR on the unknown samples, pooled DNA (containing 5 µL from each extraction) was used to optimize the reaction conditions through temperature and dilution gradients. Temperature gradients were performed to establish the optimal annealing temperature and dilution gradients were run to determine the optimal DNA dilution and avoid saturation of the ddPCR reaction. Annealing temperatures and dilution factors are listed in Table 3.1.

Each PCR reaction was prepared with 5 µL of extracted DNA template, 6.04 µL of DNase-free water, 0.23 µL of each primer, and 11.5 µL of EvaGreen ddPCR Supermix. The samples were transferred according to the manufacturer's instructions into a DG8 cartridge with 65 µL of the QX200Droplet Generation Oil for EvaGreen (BioRad: 186-4005) and converted into droplets using a BioRad QX200 Droplet Generator. After the droplets were generated, 40 µL of the samples were then transferred into a ddPCR 96-well plate. Plates were sealed with specialized foil using a PX1 Plate Sealer then placed into a C1000 Touch Thermocycler and amplified under the following conditions: 95°C for 5 minutes, followed by 40 cycles of 95°C for 30 seconds, annealing for 1 minute, and 4°C for 5 minutes with a 2°C/second ramp rate. This was followed by a final step of 90°C for 5 minutes for stabilization. After PCR amplification, the 96-well plate was transferred into the QX200 Droplet reader for quantification. Quantification of

gene copy number per µL was determined using the QuantaSoft Software from Bio-Rad. Samples were accepted if at least 13 000 droplets were created and amplified (based on general guidelines from Bio-Rad).

To determine the absolute abundance of each target gene per gram of sediment, concentrations were corrected for DNA template volume, sample volume (volume of transferred PCR reaction mix), elution volume, and dilution factor before normalizing to the mass of dry weight (equation 1). Gene copy numbers were normalized per gram of dry weight because water content was higher in top intervals of the sediment cores. The range in water content (%) in top and bottom sections of each core is presented in supplementary Table S3.3.

$$Absolute\ abundance = \frac{ddPCR\ concentration * \left(\frac{sample\ volume}{template\ volume}\right) * elution\ volume * dilution\ factor}{mass\ of\ dry\ sediment} \quad (1)$$

Relative abundances of MICR derived from ddPCR were calculated using: (MICR abundance/CYA abundance) × 100. The limit of detection (LOD) was determined by calculating the limit of the blank (LOB) and adding it to 1.645 times the standard deviation (SD) of low analyte samples (Armbruster and Pry, 2008): LOD = LOB +1.645($SD_{low\ analyte\ samples}$). The LOB was calculated using: $mean_{blanks}$ + 1.645(SD_{blanks}). Eight pooled DNA samples, diluted 100x for MICR and 1000x for CYA, were used to calculate the SD of the low analyte samples. The LOD was determined as 1.55 copies/µL (copies per microlitre of reaction) for MICR and 1.43 copies/µL for CYA. After carrying through with normalization, the LOD for the MICR target was calculated as ~1.4×10^5 copies/gram of dry sediment for L227 and ~2.8×10^5 copies/g for

lakes 223, 224, and 442. The LOD for CYA was ~4.1×10^5/g for L227 and ~8.2×10^5 copies/g for the other three lakes. These LODs are conservative estimates and values obtained using ddPCR were, at times, below the average LODs for this method.

3.3.7 High-throughput sequencing

Library preparation and paired-end sequencing was completed at Genome Quebec Innovation Centre (Montréal, Quebec). Specific CYA primers (Table 3.1), with attached overhang adaptors, were used to amplify a 450-nt-long fragment of the V3-V4 regions of the 16S rRNA gene for metabarcoding (Nübel et al., 1997; Jungblut et al., 2005; Pilon et al., 2019). PCR amplification was prepared with 1 µL of DNA template diluted 10x (few samples required a 50x dilution), 5.512 µL of DNase-free water, 0.048 µL of each primer, 0.032 µL of Qiagen HotStarTaq DNA polymerase, 0.16 µL of dNTP mix, 0.4 µL of Roche DMSO, and 0.80 µL of Qiagen 10x Buffer with 15mM $MgCl_2$. The PCR cycle program consisted of heating at 96°C for 15 minutes, followed by 35 cycles of 96°C for 30 seconds, 52°C for 30 seconds, and 72°C for 60 seconds. This was followed by a final step of 72°C for 10 minutes. Barcodes were added to each sample through a second round of PCR. Amplicons were then quantified using Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies). The library was generated by pooling equal volumes of each sample, subsequently cleaned up with sparQ PureMag beads (Quantabio) and quantified using Kapa Illumina GA with Revised Primers-SYBR Fast Universal Kit (Kapa Biosystems). The amplicon pool was loaded at a final concentration of 7 pM with a 15% PhiX spike to improve the unbalanced base composition. Sequencing was performed on an Illumina MiSeq PE300 platform using the Miseq Reagent Kit. Two nuclease-free water samples were included as negative controls.

3.3.8 HTS data processing

Raw FASTQ files obtained from Genome Quebec were demultiplexed from within the MiSeq Illumina workflow. The 16S rRNA gene sequence data were then processed using the QIIME2 pipeline (v 2019.7; Bolyen et al., 2019). The cutadapt command line tool was used to trim the forward and reverse primer site of the merged reads. DADA2 denoise-single plugin was used to truncate and remove sequences with an average Phred quality score below 20, create the amplicon sequence variant (ASV) table with representative sequences and their frequency of occurrence, perform quality filtering, and remove chimeric sequences. The sequence variants were assigned using the q2-feature-classifier Naïve Bayes classifier after extracting the region of the 16S sequences most appropriate for the primers (359F/809R) used during sequencing. The classifier was trained on the SILVA 99% identity Full Seq OTU reference database (v 138) using the classify-sklearn command on the trained classifier. An additional classifier was produced and trained on the Greengenes 99% OTU reference database (v gg_13_8_99) to compare. ASVs assigned to chloroplasts or non-cyanobacterial species were removed from further analysis using the taxa filter-table plugin. Cyanobacterial and *Microcystis* ASV counts were normalized to counts per gram of dry sediment similarly to ddPCR (equation 1). Absolute cyanobacterial ASV counts were compared against absolute gene copy numbers derived from ddPCR to determine how well ASV counts track changes in abundance at the phylum level. Due to the compositional nature of the data, the coherence between ddPCR and HTS at the level of genus was only compared using relative abundances of *Microcystis*. Proportions were exported using the feature-table relative-frequency plugin and converted into percentages by multiplying by 100. Absolute ASV counts for *Microcystis* were compared between SILVA and Greengenes to assess the consistency of these databases at the genus level. A summary of the bioinformatics workflow is

presented in Figure S3.3. The negative controls had less than 0.1% of the average number of reads and were removed from analysis.

Here, the HTS LOD was estimated as one read (in expectation) per sample. For the absolute abundances, the LOD for both CYA and MICR was calculated as approximately 8.8×10^5 copies/g for L227 and 1.8×10^6 copies/g for lakes 223, 224, and 442.

3.3.9 Statistical analyses

All statistical analyses were conducted using R open source (Version 4.0.0) and RStudio software (version 1.2.5042). Absolute gene copy data were first verified for normal distribution through Shapiro-Wilk tests. Data were log or square root transformed to approach normal distribution and minimize heteroscedasticity of variance. Absolute abundances of CYA and MICR from ddPCR and HTS were plotted as a function of sediment depth. Proportions of MICR were also plotted as a function of sediment depth. Pearson's r and Kendall's τ (for L227; due to the lack of normality, even after log-transformation) correlation coefficients were calculated to evaluate the relationship between the two methods across the entire sediment core, top sediments, and bottom sediments. All values that were below the corresponding LOD were handled by using the common approach of substituting those values with the LOD/2 (in line with guidance provided by the United States Environmental Protection Agency, 2006). The use of LOD/2 has been found to be reliable in studies where much of the data are below the detection limit or where data are highly skewed (Glass and Gray, 2001). Relative abundances of MICR (as determined from ddPCR, SILVA, and Greengenes) were transformed by employing arcsine square root transformations. Temporal autocorrelation is an issue common in analyses of sediment records and environmental data in general and can potentially lead to the lack of

independence among time points in the collected samples (Legendre, 1993; Taranu et al., 2018). To test for the effect of temporal autocorrelation on *p*-values, gene copy numbers were subjected to the autocorrelation function (ACF) in R. If both variables within the correlation were found to be temporally autocorrelated, the *p*-value was corrected by permuting by blocks using the *permute* packages in R. If either variable was not found temporally autocorrelated, the *p*-value was not corrected (Legendre and Legendre, 2012).

Using ^{210}Pb profiles (see supplementary data Figure S3.2), top sediment samples for experimental lakes 227 and 223 were selected to correspond to and include the intervals associated with the approximate period of lake manipulation and all subsequent sediment deposition (1969 – 2018 for L227 and 1976 – 2018 for L223). Bottom sections chosen for reference lakes 224 and 442 were chosen to roughly parallel L223 bottom sections because sediment intervals corresponded to relatively similar timeframes in the pre-manipulation periods (Table S3.4). Paired *t* tests were also performed to assess statistical differences in DNA content between recent samples (top sediments) and older samples (bottom sediments).

3.4 Results

DNA was successfully extracted from all sections sampled from each lake core. The experimental lakes had more DNA in top sediments compared to the bottom sediments relative to the reference lakes (Table 3.2): for L227 the mean DNA concentration was 1.6X greater (p-value < 0.01) and for L223, 1.26X greater (p-value < 0.05). In contrast, there was no statistical difference in DNA content between top and bottom sections for L224 and L442 ($p > 0.05$ for both lakes). For all four cores, the mean DNA quality was high (260:280 ratio > 1.7). The bottom intervals of L227 and L223 contained DNA of higher purity (260:280 ratio > 1.8) in comparison

to the top subsection intervals. Only the top sections from L227 had a mean ratio less than 1.7. All lakes showed very low 260:230 ratios.

In L227, ddPCR results showed very little change in CYA abundance below 15 cm, with gene copy numbers consistently at ~10^8 copies/g of dry weight, between two to three orders of magnitude above the corresponding LOD (Figure 3.1, top left). Subsequently, above 15 cm, total CYA gene copy number increased two orders of magnitude (~100-fold from 10^8 to 10^{10} copies/g). In contrast, only a gradual ~10-fold increase from 10^{10} to 10^{11} was observed in CYA abundance as determined from HTS throughout the core, and no large increase in abundance was seen around 15 cm. Based on the sediment profile from L227 (Figure S3.2), 15.25 cm corresponds to ~1968 (± 2.8 years), the approximate time lake fertilization began. Absolute MICR abundances determined from ddPCR were consistently below the corresponding LOD (1.4×10^5) prior to lake fertilization but subsequently rose ~two orders of magnitude above detection limits. MICR abundances derived from HTS also rose ~two orders of magnitude above the corresponding LOD for this method (8.8×10^5) at 15 cm. Similar trends were observed when gene copy numbers were normalized per gram of wet sediment (Figure S3.4).

In L223, ddPCR results showed a slight increase in CYA abundance (10-fold from ~10^8 to 10^9 copies/g) over time from the bottom of the core to the top (Figure 3.1, top right). A slight increase in total cyanobacterial abundance was also observed with HTS from ~10^{10} to 10^{11} over time. Experimental acidification of L223 occurred over an eight-year period, between 1976 and 1983, corresponding to ~18 – 20 cm sediment depth based on the sediment profile (Figure S3.2). A small three-fold increase in CYA abundance, determined from ddPCR, was observed during this period, whereas HTS results did not show any increase during this time. ddPCR-derived MICR concentrations were consistently around the LOD but a ~three-fold increase above

detection limits occurred at approximately 16 cm. HTS-derived MICR abundances were near or below the corresponding LOD in bottom sediments but rose 100-fold from ~10^7 to 10^9 over time.

In L224, a ~10-fold increase in total CYA was observed over time from the bottom of the core to the top in both ddPCR and HTS (Figure 3.1, bottom left). In L442, a 100-fold increase in CYA was observed over time in both methods (Figure 3.1, bottom right). MICR abundances in L224, as determined from ddPCR, increased 10x above the LOD in top sediments, whereas ddPCR-derived MICR abundances in L442 were consistently at or near the LOD and this remained relatively unchanged over time. MICR abundances from HTS in both L224 and L442 increased ~100-fold (10^8 to 10^{10}), but in L442, the profile was skewed towards higher relative abundances in bottom sediments.

Relative abundances of MICR derived from ddPCR and from HTS using either the SILVA or Greengenes database showed similar patterns and tracked one another well only in top sediments of L227 (Figure 3.2). In L227, a rapid increase in MICR was observed with all three approaches at ~15 cm, consistent with the beginning of lake fertilization. All three methods also showed a decline in MICR proportions starting at 11.25 (~1988 ± 2 years— when nitrogen loading stopped). Below 15 cm, the ddPCR-derived MICR proportions appeared to be higher than the HTS results, but the ddPCR absolute abundances were below the LOD in these intervals (see Figure 3.1). In L223, SILVA and Greengenes appeared to track one another somewhat in top sediments, but relative abundances were much higher than the MICR levels obtained from ddPCR. While ddPCR and Greengenes-derived MICR abundances appeared to not change much throughout the sediment cores of lakes 224 and 442, a large increase (~5% to 40%) from the bottom of the core to the top in L224 (Figure 3.2, bottom left) and a large decrease (~50% to 15%) in L442 (Figure 3.2, bottom right) was found when using the SILVA database. The high

relative abundances of MICR in the two deepest sediment sections of L442 was not found when using Greengenes.

Total gene target copy abundance data were normally distributed after log or square root transformation (except for those from L227). Because SILVA and Greengenes were strongly correlated for the number of ASVs assigned to the cyanobacteria phylum in all four lakes (Figure S3.5), only the SILVA database was used to compare ddPCR CYA gene copy numbers to HTS-derived CYA counts. In L223, the ACF function showed CYA gene copy numbers derived from ddPCR were temporally autocorrelated, as were the CYA counts from HTS. All correlations were still significant after blockwise permutation. Calculated correlation coefficients showed strong positive relationships between ddPCR and HTS abundance for CYA in all lakes (Figure 3.3). When cores were split between top and bottom sediments, all four cores showed very strong positive correlations between ddPCR and HTS in bottom sediments (Figure 3.3, right). While there was a modest correlation in top sediments of L223 (Figure 3.3, middle), there were no significant correlations between the two methods in top sediments of lakes 227, 224, or 442. Correlation analysis using wet sediments also showed moderate to strong levels of coherence between ddPCR and HTS across the lake sediment cores and in bottom sediments (Figure S3.6).

SILVA and Greengenes taxonomic assignments for the *Microcystis* genus differed (Figure S3.7) and thus both SILVA and Greengenes MICR relative abundances were used to compare with ddPCR-derived MICR relative abundances (Table 3.3). When using SILVA-based taxonomy, only L227 and L442 showed a significant correlation between ddPCR and HTS for transformed relative MICR abundances. When using the Greengenes database, a correlation between ddPCR- and HTS-derived MICR relative abundances was only found in L227. The SILVA and Greengenes databases were significantly correlated with one another for relative

abundances of MICR in lakes 227 and 223. MICR abundances derived from ddPCR, SILVA, and Greengenes, were temporally autocorrelated in L227. Correlations were still significant after blockwise permutation. SILVA relative abundance outputs were used to determine the degree of correlation between ddPCR and HTS from top and bottom sediments (Figure 3.4). Calculated correlation coefficients between top and bottom sediments using Greengenes MICR proportions are presented in Figure S2.8. When using SILVA, moderate to strong correlations were observed in top sediments of lakes 227, 223, and 442, but no significant relationship was found in top sediments of L224. A strong correlation between ddPCR and SILVA was also found in bottom sediments of L442 but not in bottom sediments of the other three lakes (Figure 3.4).

3.5 Discussion

I compared the sensitivity between ddPCR and HTS by taking advantage of a unique set of experimental and reference lakes and studied the long-term temporal dynamics of cyanobacteria based on the examination of DNA preserved in sediment. These findings suggest that, while the two methods were significantly correlated with one another, the targeted ddPCR approach reflected more accurately the significant rise in cyanobacteria associated with experimental fertilization.

3.5.1 Comparison of absolute abundances between ddPCR and HTS at the level of cyanobacteria

The results showed quantification of the cyanobacterial 16S rRNA gene from ddPCR was more consistent with the known history of lake experimentation. In the case of experimental L227, while the ddPCR profile showed a dramatic increase in CYA abundance at 15.25 cm (~1968 ± 2.8 years), HTS did not show any change during lake fertilization. These trends were

also observed when gene copy numbers were normalized per gram of wet sediment (Figure S3.4). Eutrophication of L227 began in 1969 and led to the formation of large planktonic cyanobacterial blooms in response to phosphorus and nitrogen loading to the lake (Schindler, 1974, 1977). Phytoplankton records obtained from IISD-ELA show that L227 cyanobacterial biomass increased rapidly in 1969 (Paterson et al., 2011) to over 20 times above that seen in reference lakes at that time. In experimental L223, again only ddPCR illustrated an increase in CYA abundance during the period of experimentation. This is consistent with the surface water phytoplankton records. Findlay and Kasian (1986) found that as pH decreased during the lake acidification experiment, planktonic cyanobacteria increased in biomass. Cyanobacteria mean biomass increased tenfold from ~0.02 g/m^3 to 0.2 g/m^3 around 1978, before declining to pre-acidification levels in 1993-1994 (Findlay and Kasian, 1996). The ddPCR results for L223 reflected similar patterns and CYA abundance levels declined to pre-acidification levels at approximately 11 cm (early 2000s). HTS-derived results did not demonstrate any increase or decrease during these years.

The lack of an increase in CYA abundance from HTS in L227 and L223 during lake manipulation could be attributed to methodological factors. Primer biases could have impacted the detection and quantification of the cyanobacterial 16S rRNA target gene. The primer sequences used to target CYA in ddPCR have more degenerate bases and amplify a smaller amplicon size than the primers used for HTS. Although both primer sets have been successfully used in other studies (HTS: Nübel et al., 1997; Jungblut et al., 2005; ddPCR: Rinta-Kanto et al., 2005; Pal et al., 2015; Monchamp et al., 2016; Pilon et al., 2019), the choice of specific primers can impact the detection of selected target genes in complex environments (Lee et al., 2012). The amplicon size for the CYA primer set used for HTS was much longer (450 bp) than what was

used for ddPCR (269 bp). Cyanobacterial 16S rRNA gene copy number is known to vary between one to four copies per cell among certain cyanobacterial taxa (Schirrmeister et al., 2012). As such, there is likely a greater chance of detecting a taxon that has multiple copies of this 16S rRNA gene if one of those copies has a longer fragment preserved. This may have contributed to the greater overall cyanobacterial abundances observed when using HTS. A study comparing varying PCR amplicon sizes on microbial community analysis found that with shorter amplicons, the PCR reaction proceeded more efficiently and that artifact formation was less likely in comparison to longer fragments (Huber et al., 2009). While the data presented here point to some potential primer bias, it is also likely that the lack of an observable increase in CYA abundance using HTS can be attributed to a high number of false negative errors.

Humic substances have been found to represent potential biases in metabarcoding techniques and other PCR-based approaches on eDNA, generating false-negative results (Rochelle et al., 1992; Savichtcheva et al., 2011; Albers et al., 2013; Thomsen and Willerslev, 2015). Here, most of the samples were taken from the top (recent sediments) such that the impacts of lake manipulation on cyanobacterial dynamics in L223 and L227 could be assessed and compared to the reference lakes which have experienced little to no human impact. While strong correlations were found between the two methods in bottom sediments of both the experimental and reference lakes, no significant relationships between the two estimates of cyanobacterial abundance were observed in top sediments of lakes 227, 224, and 442. Higher levels of humic and fulvic acid substances within surface sediments (~5 cm) (Calace et al., 2006; Hou et al., 2014) may be the reason why top sections of these cores showed no significant relationship between ddPCR and HTS results for CYA gene copy numbers. Humic acids, co-extracted with DNA, can inhibit PCR amplification by suppressing the Taq polymerase,

preventing gene targets from amplifying. While the same DNA extractions were used for both ddPCR and HTS, tolerance to inhibitors can differ between the two methods. ddPCR has been found to be highly resistant to PCR inhibitors (Taylor et al., 2017, 2019). DNA templates are randomly partitioned into thousands of droplets of uniform size and volume, improving the performance of the individual endpoint PCR reactions within each droplet (Hindson et al., 2011), reducing the ability of inhibitors to impact amplification, and allowing for reliable detection of low-abundance targets (supplementary Figure S3.9). Low 260:230 ratios from sediment samples have been attributed to high organic contaminants and humic acid levels in other studies (Antony-Babu et al., 2013; Ramírez et al., 2018). The quality of the DNA was in fact particularly low in top sediments of L227; this was expected since the production and deposition of organic matter increased tremendously as a result of the lake fertilization (O'Connell et al., 2020). In this study, only a few samples were selected from the bottom of the reference lake sediment cores because cyanobacterial levels were expected to be low with little variability. Future studies evaluating differences between top and bottom studies (for e.g., studies comparing trends between contemporary and preindustrial sediments) should, however, use larger sample sizes to draw definitive conclusions. In addition to potential humic acid interference, the pooled samples used for sequencing could have also contributed to higher rates of false negatives (Manter et al., 2010).

A major limitation with sedDNA analyses in paleolimnological research includes incomplete DNA extraction or DNA extraction bias. Regardless of adjustments made during the DNA extraction protocol, low DNA yields persisted in a few samples from L227, L223, and L442 sediments, which could have ultimately impacted the number of ASVs passed through the bioinformatic pipeline. A recent study comparing ASV counts across a series of DNA dilutions

found that at low DNA concentrations, a disproportionate number of contaminants (from extraction reagents or laboratory environment) were amplified as a result of low signal-to-noise ratio (Caruso et al., 2019). Here, when experimenting with generous truncation parameters within the pipeline (for e.g., using lower quality reads, using only single end reads etc.), the number of ASVs passed through increased overall, but the CYA trends after taxonomy assignment did not differ to the results presented in this study. This indicates that the absolute ASV counts used for analysis were more impacted by the quality and quantity of the starting DNA material, rather than the specific filtering parameters chosen within the pipeline, and that contaminant sequences in samples with low DNA concentrations could have disproportionally been amplified, subsequently reducing the number of ASVs assigned to cyanobacteria. The combined effects of amplified contaminant sequences from some sediment sections and high amounts of false negatives, are potential reasons as to the slightly weaker correlations found between ddPCR and HTS when comparing absolute abundances in L227. Based on the results presented here, ddPCR minimizes PCR-inhibiting substances and should be used to quantify absolute concentrations of target genes from sedDNA records.

3.5.2 Comparison between ddPCR and HTS using relative abundances of the Microcystis genus

Microcystis occurs in these lakes (based on IISD-ELA phytoplankton records), but abundance levels were expected to be low. This allowed for testing the limits of both methods and for low gene copy numbers to be compared between ddPCR and HTS. While strong positive correlations were observed between ddPCR and HTS for absolute cyanobacterial abundances, the combination of overall low *Microcystis* copy numbers derived from ddPCR and little variation in MICR trends, could explain the weaker tracking between the two molecular

approaches when using proportion data. The correlation analyses showed moderate to strong levels of agreement between ddPCR and HTS in top sediments of lakes 227, 223, and 442, but results differed based on the use of either the SILVA or Greengenes reference taxonomy database. When using Greengenes, no correlation was observed in top sediments of L223 (Figure S3.8). In addition, the correlations across the core of L442 and across bottom sediments were significant only when using SILVA. This indicates that this more-frequently updated reference database could have picked up more *Microcystis* sequences than Greengenes, which was last updated in 2013. SILVA-derived relative abundances were much higher when compared to Greengenes, and it is possible that misannotations arose, particularly at the genus level in these publicly available DNA reference databases (Sierra et al., 2020). Additionally, SILVA relative *Microcystis* abundances in the reference lakes were approximately ~10 – 100x higher than in the experimental lakes. Picocyanobacteria (cyanobacteria < 2 μm in diameter) and small unicellular coccoid species are abundant in the reference lakes and range from 10^4 – 10^5/mL in ELA lakes during the summer (Pick, unpublished data). Both a high level of picocyanobacteria and coccoid species, and annotations to invalid names (in combination with the compositional nature of the sequencing data) could have contributed to the higher SILVA relative MICR abundances within the reference lakes.

Very low absolute MICR abundances and low downcore variation in gene copy numbers (as determined from ddPCR) could have contributed to the insignificant relationship between ddPCR and HTS across the sediment cores of lakes 223 and 224. In addition, the low sample size in L224 could have contributed to the poor statistical power for detecting correlations between the two methods in that core. No firm conclusion can be made regarding the correlation between ddPCR and HTS relative abundances across bottom sediments of the experimental lakes as

abundances were found below detection limits in many of those intervals (Figure 3.1). However, the very high consistency between these two methods in top sediments of L227 likely explains the moderate but significant correlation found across that sediment core. As previously mentioned, it is also possible that contaminants in samples with low DNA concentration reduced the total number of cyanobacterial ASVs in bottom sediments. This could have seriously skewed MICR proportions. Future studies assessing ddPCR and HTS relative abundance data from bottom sediments should use greater sample sizes or interpret results with caution if gene copy numbers are relatively low.

Another potential reason for the discrepancy between ddPCR and HTS for relative MICR abundances could be attributed to the lack of primer specificity. *Microcystis*-specific 16S rRNA primers were used to target the *Microcystis* genus from ddPCR, whereas data filtering through the bioinformatic pipeline was used to obtain *Microcystis* counts from HTS. Both Greengenes and SILVA had relatively higher MICR proportions when compared to ddPCR. It is possible that the MICR primers used in ddPCR were not the most universal and failed to pick up all *Microcystis* present in the samples. While no other *Microcystis* primer sets were tested in this study, future studies should sequence and/or add degenerate bases in the sequence primers to test this hypothesis and confirm the universality of the primers. Rigorous parameters for filtering in the bioinformatics pipeline and the choice of primers selected for analysis should be carefully considered to obtain reliable proportional data for the accurate inference of the composition.

3.5.3 Recommendations

The method of choice for sediment DNA research should depend on the ultimate goals of the study. Here, I found ddPCR to have fewer limitations than HTS, and ddPCR results were

more consistent with the history of cyanobacteria in these lakes. The high-throughput capacity of ddPCR and its ability to tolerate inhibitors makes this molecular method more suitable for quantifying targets from complex environments like sediments. Absolute abundance data from HTS, abundance measures derived from top sediments, or samples with very low gene copy numbers, should be interpreted with caution. However, metabarcoding approaches provide unparalleled resolution of microbial community diversity and expand taxonomic scope at unprecedented scales. If the main goal of a study is track abundance through time in an environment with a *priori* knowledge that low levels of inhibitors are present in the samples (based on nanodrop quality scores, for example), then either ddPCR or HTS are appropriate methods to consider. However, if the complexity of the environment is unknown, or if high levels of PCR inhibitors such as organic contaminants or humic acids are likely to be found, ddPCR is likely the more appropriate choice. If only relative abundance counts are necessary, then HTS alone is sufficient (although the choice of database for taxonomic assignment should be carefully considered). On the other hand, if overall community dynamics are assessed, and both abundance and community composition are required, I recommend to first analyse results through HTS and subsequently quantify change through time with ddPCR. As both ddPCR and HTS methods continue to improve and as costs continue to decline, the benefits of using both approaches in tandem can increase the validity of using sedDNA to reconstruct historical dynamics.

Supplementary material is provided in Appendix B.

Data Availability

Sequence data used for analysis are available in the National Center for Biotechnology Information Sequence Read Archive database under study accession number PRJNA679845.

Acknowledgements: This research would not have been possible without the long history of lake monitoring and whole ecosystem manipulations conducted at the Experimental Lakes Area. I am deeply grateful for the IISD-ELA and its researchers for their field work and continuous support. I thank Linda Kimpe (LANSET, University of Ottawa) for her expertise in ^{210}Pb dating of the sediment samples and Rebecca Garner for assisting with core subsectioning. This research was enabled in part by support provided by Compute Canada (www.computecanada.ca).

Table 3.1 The primer sequences used in droplet digital PCR (ddPCR) and High-throughput sequencing (HTS), amplicon sizes, annealing temperatures, and dilution factors of target genes cyanobacterial 16S rRNA and *Microcystis* 16S rRNA.

Gene	Primers	Sequence (5'-3')	Amplicon size (bp)	Annealing T (°C)	Dilution factor
Microcystis 16S rRNA (ddPCR)	MICR F[1]	GCC GCR AGG TGA AAM CTA A	247	56	10x
	MICR R[1]	AAT CCA AAR ACC TTC CTC CC			
Cyanobacteria 16S rRNA (ddPCR)	CYA 108F[1,2,]	ACG GGT GAG TAA CRC GTR A	269	56	32x
	CYA 377R[1.3]	CCA TGG CGG AAA ATT CCC C			
Cyanobacteria 16S rRNA (HTS)	CYA 359F[3,4]	GGG GAA TYT TCC GCA ATG GG	450	52	10x, 50x
	CYA 809R[3,4]	GCT TCG GCA CGG CTC GGG TCG ATA			

[1](Rinta-Kanto et al., 2005), [2](Urbach et al., 1992), [3](Nubel et al., 1997), [4](Jungblut et al., 2005)

Table 3.2 Comparison of mean DNA concentration (± SE) and purity between top and bottom subsection intervals.

Lake	Core depth (cm)	Number of subsections	DNA (ng/g of dry sediment)	260:280 ratio	260:230 ratio
227	Top (1-15.25)	11	4102 (± 355)	1.62 (± 0.02)	1.15 (± 0.08)
	Bottom (16.25-51.75)	12	2577 (± 255)	1.81 (± 0.02)	1.39 (± 0.08)
	Total core	23	3306 (± 266)	1.72 (± 0.03)	1.28 (± 0.06)
223	Top (1-20.25)	15	2870 (± 176)	1.74 (± 0.03)	1.41 (± 0.07)
	Bottom (21.75-41.75)	9	2311 (± 152)	1.82 (± 0.07)	1.39 (± 0.08)
	Total core	24	2660 (± 134)	1.77 (± 0.03)	1.40 (± 0.05)
224	Top (1-7.25)	9	4662 (± 508)	1.86 (± 0.01)	1.28 (± 0.07)
	Bottom (16.25-32.25)	4	4324 (± 414)	1.86 (± 0.02)	1.35 (± 0.05)
	Total core	13	4558 (± 367)	1.86 (± 0.01)	1.29 (± 0.05)
442	Top (1-12.25)	8	4183 (± 780)	1.90 (± 0.03)	1.33 (± 0.13)
	Bottom (13.25-39.75)	6	4098 (± 754)	1.99 (± 0.08)	1.30 (± 0.10)
	Total core	14	4146 (± 530)	1.93 (± 0.04)	1.32 (± 0.08)

Table 3.3 Kendall's (L227) and Pearson's (L223, L224, and L442) correlation coefficients of arcsine square root transformed relative *Microcystis* abundances derived from droplet digital PCR (ddPCR) and from High-throughput sequencing (HTS), using the SILVA and Greengenes databases.

			HTS	
Lake		ddPCR	SILVA	Greengenes
227	ddPCR	1.00	0.48*	0.57**
	SILVA		1.00	0.72**
	Greengenes			1.00
223	ddPCR	1.00	0.19	0.15
	SILVA		1.00	0.81**
	Greengenes			1.00
224	ddPCR	1.00	0.15	-0.34
	SILVA		1.00	0.23
	Greengenes			1.00
442	ddPCR	1.00	0.85**	-0.31
	SILVA		1.00	-0.02
	Greengenes			1.00

*$p < 0.05$, **$p < 0.001$

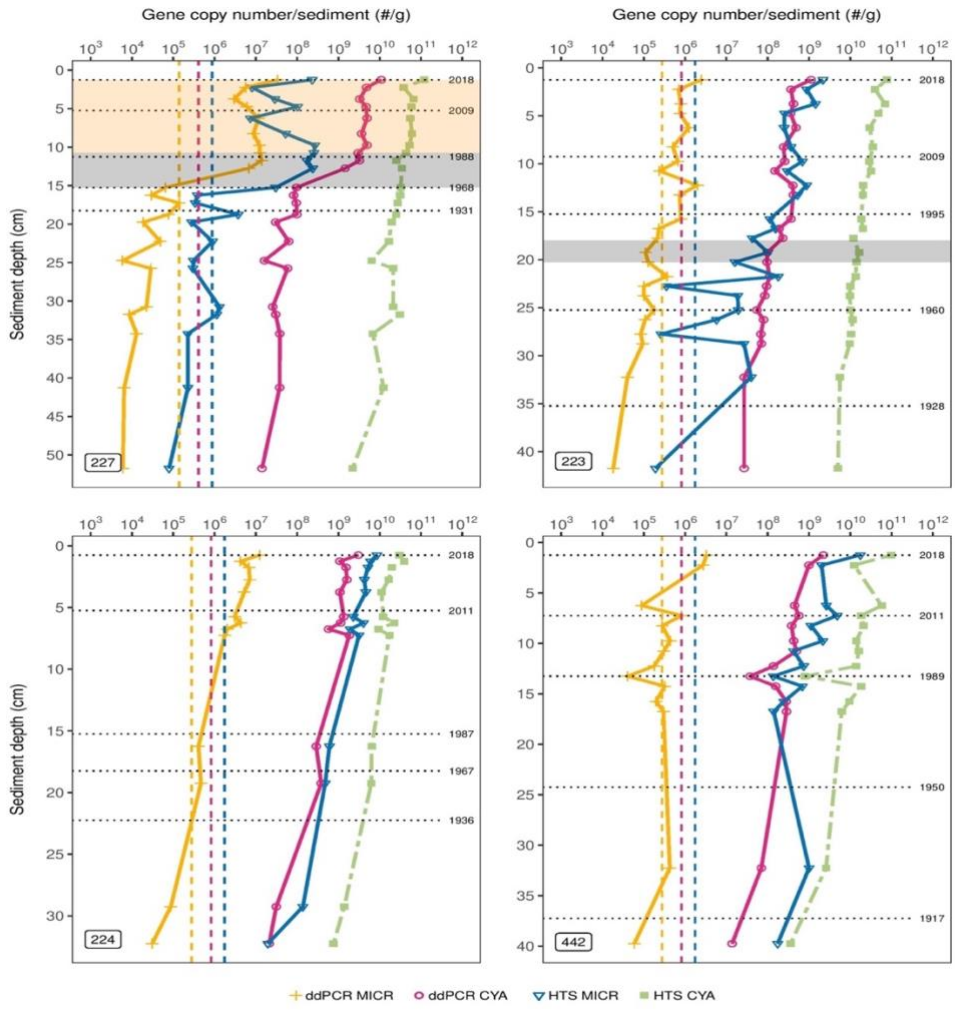

Figure 3.1 Target gene concentrations in sediment cores of the Experimental Lakes Area, Canada from manipulated lakes 227 and 223 and reference lakes 224 and 442. Absolute gene copy numbers of *Microcystis* 16S rRNA (MICR) and cyanobacterial 16S rRNA (CYA) as determined through droplet digital PCR (ddPCR) are shown, along with the amplicon sequence variant counts of cyanobacterial 16S rRNA and of the *Microcystis* genus from high-throughput sequencing (HTS) (using SILVA). Results are normalized per gram of dry sediment and presented on a logarithmic scale. The orange and pink dashed vertical lines are the detection limits (LODs) for ddPCR-derived *Microcystis* 16s rRNA and cyanobacterial 16S rRNA abundances, respectively. The blue dashed vertical line is the LOD of HTS-derived abundances for both cyanobacterial 16S rRNA and *Microcystis*. The grey shading in the top left plot corresponds to the period of phosphorus and nitrogen loading in Lake 227 (1969 – 1989), while the beige shading corresponds to the period of phosphorus loading only (1990 – 2018). The grey shading in the top right plot corresponds to the period of sulfuric acid loading in Lake 223 (1976 – 1983).

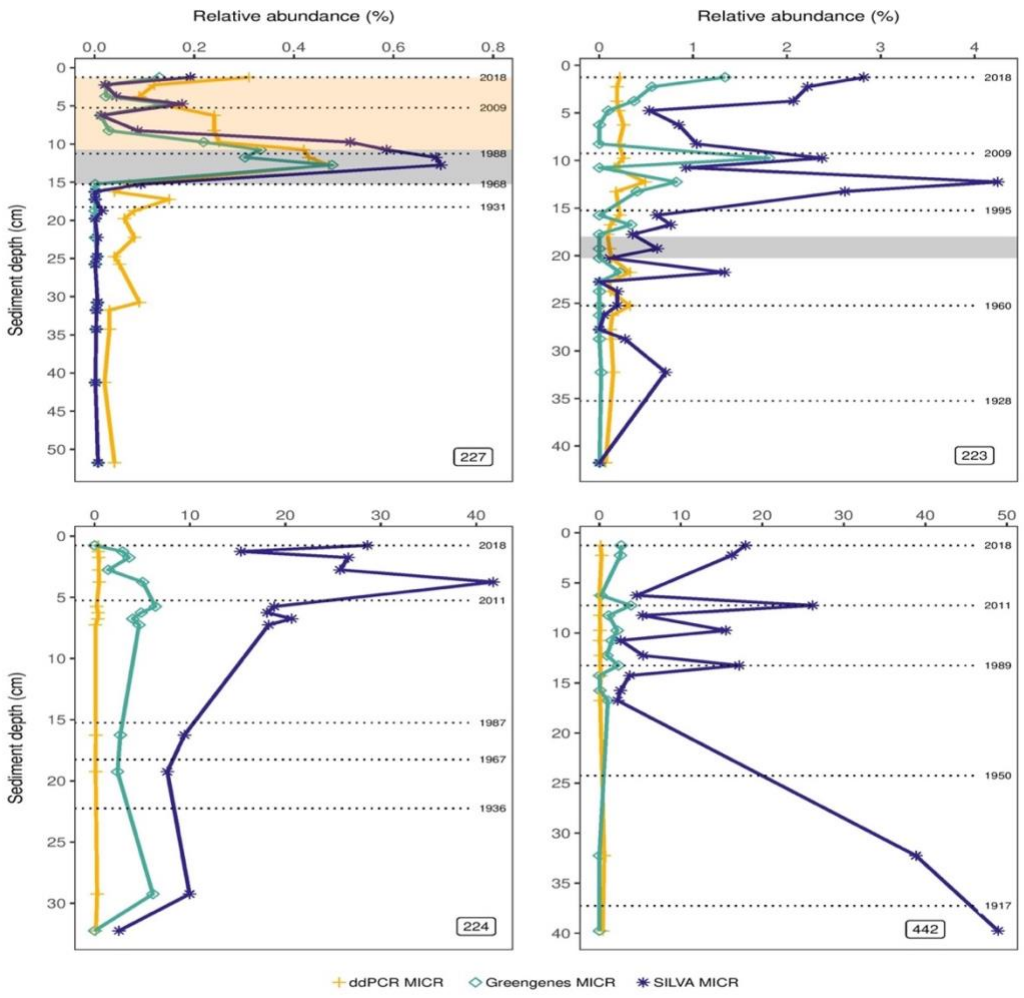

Figure 3.2 Percent abundance of the *Microcystis* genus (MICR) in lake sediment cores of the Experimental Lakes Area, Canada, from manipulated lakes 227 and 223, along with reference lakes 224 and 442, as determined through droplet digital PCR (ddPCR) and through high-throughput sequencing (HTS) using the SILVA and Greengenes databases. The grey shading in the top left plot corresponds to the period of phosphorus and nitrogen loading in Lake 227 (1969 – 1989), while the beige shading corresponds to the period of phosphorus loading only (1990 – 2018). The grey shading in the top right plot corresponds to the period of sulfuric acid loading in Lake 223 (1976 – 1983).

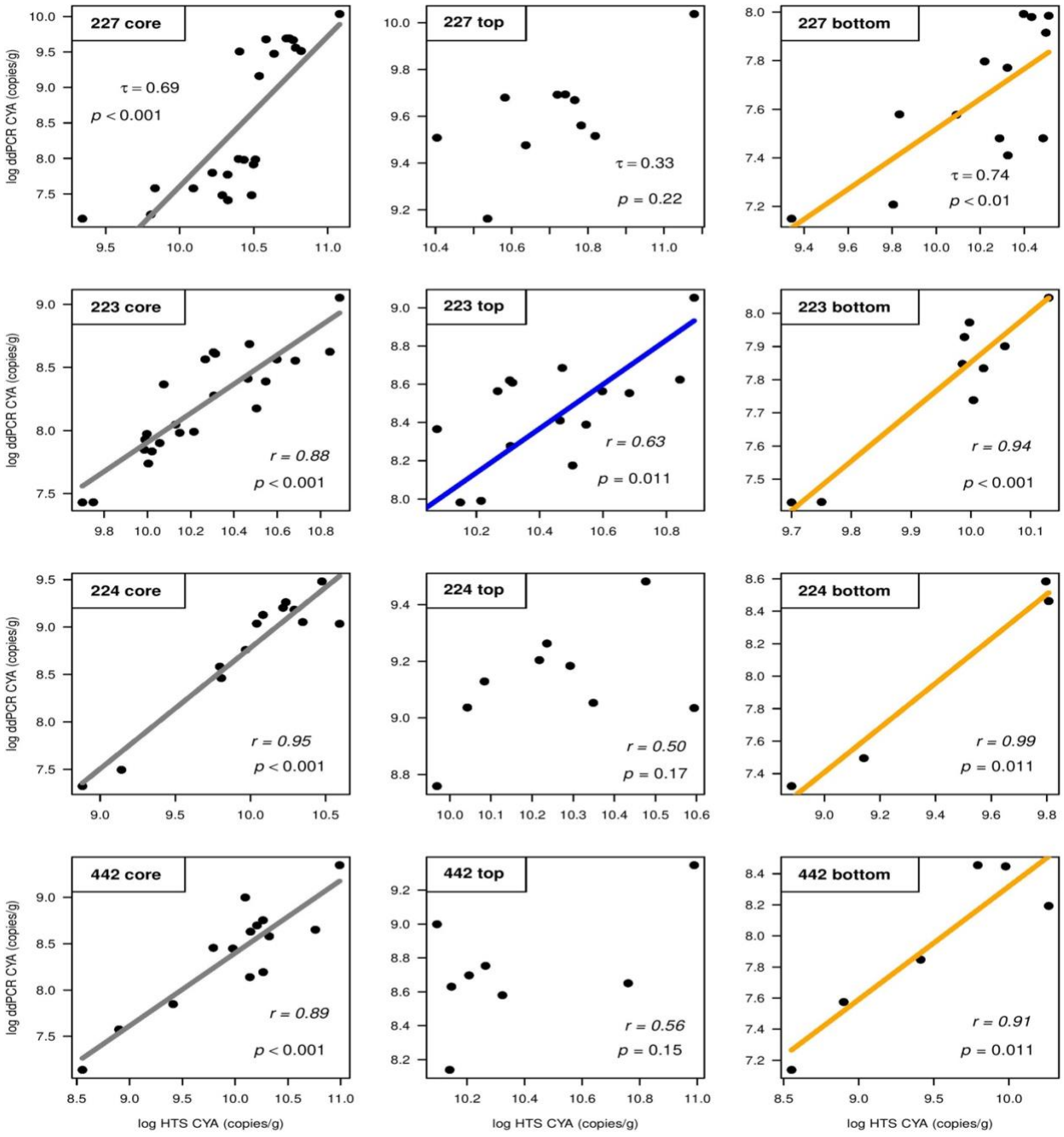

Figure 3.3 Correlations between log-transformed droplet digital PCR (ddPCR) gene copy numbers and high-throughput sequencing (HTS) amplicon sequence variant counts (using SILVA) for cyanobacterial 16S rRNA (CYA), normalized per gram of dry sediment across the whole core (left), top sediments (middle), and bottom sediments (right) of study lakes 227, 223, 224, and 442. Kendall's τ (Lake 227) and Pearson's r (lakes 223, 224, and 442) correlation coefficient values are shown.

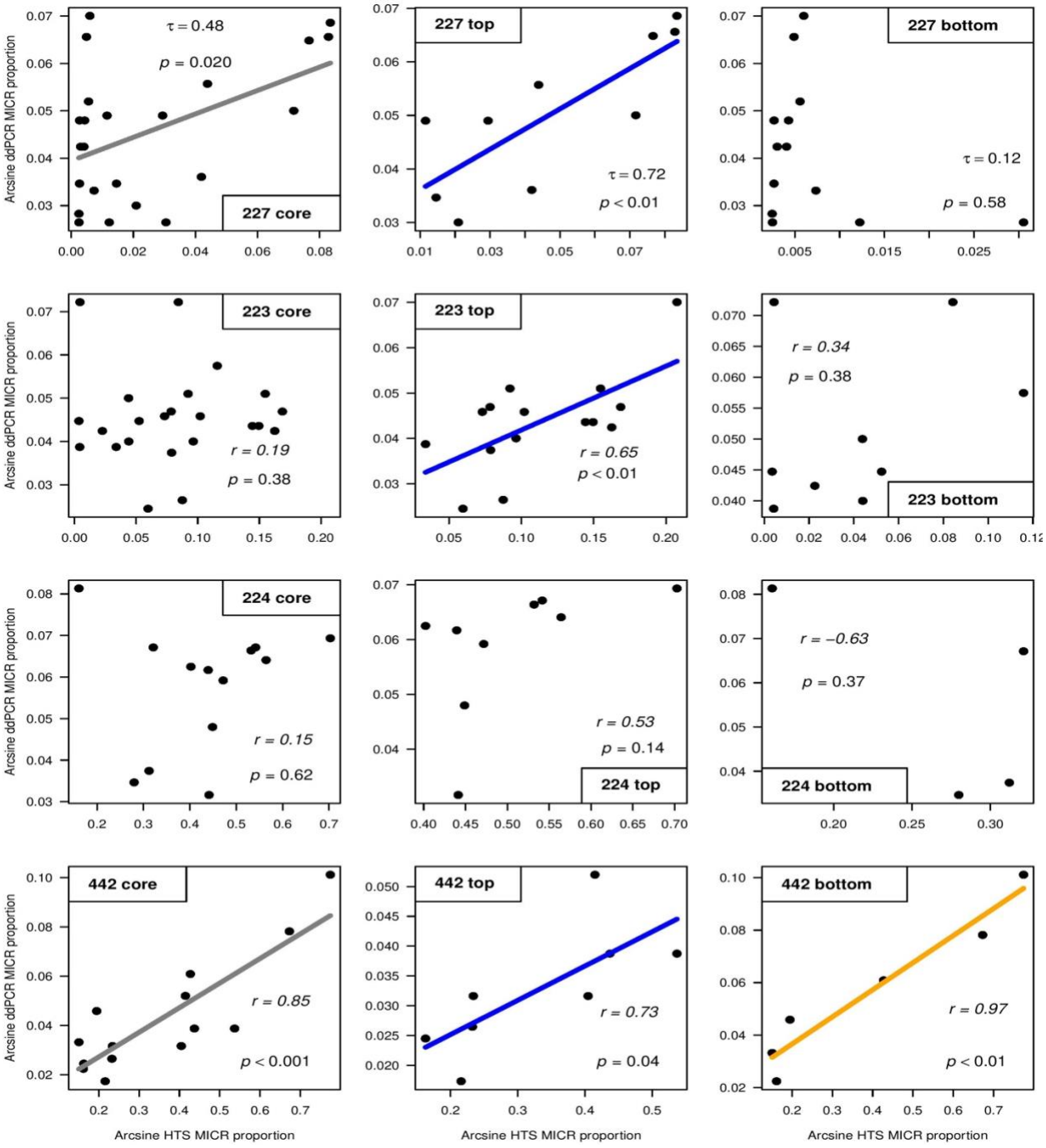

Figure 3.4 Correlations between arcsine square root transformed droplet digital PCR (ddPCR) and high-throughput sequencing (HTS) (using SILVA) relative abundance outputs for *Microcystis* (MICR) across the whole core (left), top sediments (middle), and bottom sediments (right) of study lakes 227, 223, 224, and 442. Kendall's τ (Lake 227) and Pearson's r (lakes 223, 224, and 442) correlation coefficient values are shown.

CHAPTER 4: Long-term cyanobacterial dynamics from lake sediment DNA in relation to experimental eutrophication and acidification at the Experimental Lakes Area

Hebah S. Mejbel[1], Courtney Irwin[1,2], William Dodsworth[1], Frances R. Pick[1]

[1]Centre for Advanced Research and Environmental Genomics, University of Ottawa

[2]Department of Laboratory Medicine and Pathobiology, University of Toronto

Author contributions

HSM and FRP designed the experiment. HSM performed the experiments and statistical analyses and wrote the first draft of this manuscript. CI assisted in the molecular analyses. CI, WD and FRP revised the manuscript.

Keywords*:* acidification, climate change, cyanobacteria, droplet digital PCR, eutrophication, high-throughput sequencing, paleolimnology, sediment DNA

4.1 Abstract

Cyanobacterial blooms in aquatic environments have impacted ecosystem health, altered food-webs, and contributed to substantial regional economic losses. The relative impacts of climate change, eutrophication, and other environmental stressors on the formation of cyanobacterial blooms remain unclear due to the lack of long-term data. The analysis of lake sediment archives can help address such questions. The abundance of target cyanobacterial genes was quantified using droplet digital PCR from eutrophied (L227), acidified (L223), and reference (L224 and L442) lake sediments at the Experimental Lakes Area, where various anthropogenic drivers have been manipulated and phytoplankton changes tracked over the past 50 years. Based on sediment DNA (sedDNA) records, L227 has seen more than two full orders of magnitude rise in cyanobacterial abundance in response to phosphorus and nitrogen loading. L223 also experienced a small increase in cyanobacteria during the acidification period. The abundance of *mcyE*, a gene involved in microcystin production, was also quantified. Microcystins have not been previously recorded in these lakes, but sedDNA archives identified the presence of *mcyE* and an increase in its abundance in L227 after phosphorus-only loading continued. To analyze shifts in the cyanobacterial communities, the cyanobacterial-specific 16S rRNA gene was sequenced. SedDNA revealed a shift towards Nostocales dominance in response to eutrophication in L227; this was also seen in the surface sediments of the other lakes but to a lesser extent. Heating degree days were important in explaining variation in the cyanobacterial communities, but nutrients had the greatest influence on the community in L227. When comparing the sedDNA data with surface water records, moderate to strong correlations were found between the two archives. This research validated the use of sedDNA for the analysis of long-term trends and the results can be used to analyze other freshwater ecosystems impacted by anthropogenic activities and climate change.

4.2 Introduction

Over the last century, anthropogenic activities and rising annual temperatures have led to substantial changes to freshwater ecosystems. Eutrophication, in particular, (for e.g., from coastal farming or intensive aquaculture) has led to community composition shifts, changes in phytoplankton biomass, cyanobacterial growth, and the formation of harmful algal blooms (Schindler et al., 2008; Pal et al., 2015; Taranu et al., 2015). According to sediment archives (Taranu et al., 2015), cyanobacterial biomass has increased in temperate lakes over the past several decades. However, few studies have assessed cyanobacterial dynamics in abundance and composition through large spatial and temporal scales. The lack of long-term data studies limits our understanding of the synergistic effects of environmental stressors on cyanobacterial dynamics and in particular the role of climate change.

In the absence of monitoring records, traditional paleolimnological approaches using biological remains (for e.g., subfossil diatoms, cladocera, and pigments) and other material (for e.g., isotope signatures) stored in sediment, can be used as proxies to reconstruct environmental change, evolutionary change, and spatiotemporal variability. However, these proxies have some limitations as not all organisms leave behind some form of physical subfossil. The analysis of past cyanobacterial trends using only fossil remains is especially difficult as only a few taxa produce akinetes or resting cysts. Other studies have used carotenoid pigments to evaluate long-term trends of cyanobacteria in freshwater ecosystems (Leavitt and Findlay, 1994; Deshpande et al., 2014; Pal et al., 2015; Monchamp et al., 2021). However, cyanobacterial pigments (such as zeaxanthin) can be found in other phytoplankton and there is poor taxonomic differentiation between groups of cyanobacteria based on pigments. This could hinder accurate quantification of cyanobacteria. Over the past decade, more research on the feasibility and usefulness of other

methods, including molecular and bioinformatics techniques, has been performed (Legrand et al., 2017; Monchamp et al., 2016, 2018; Pal et al., 2015; Mejbel et al., 2021b).

There has been an increase in the use of sediment DNA (sedDNA) as an alternate approach for the analysis of complex microbial communities in aquatic systems over time (Capo et al., 2021). This more targeted method allows for the analysis of organisms that rarely leave behind morphological microfossils. Organisms can leave behind genetic information through shedding, decay, and excretion, which can then be used to reconstruct the dynamics of communities within these environments. For example, one study compared the use of quantitative PCR with that of pigment analysis through high-performance liquid chromatography to reconstruct cyanobacterial dynamics and track historical trends (Pal et al., 2015). Other molecular techniques, such as high-throughput sequencing, have also been applied in studies assessing phylogenetic diversities and whole cyanobacterial communities using sedimentary archives (Monchamp et al., 2016; Tse et al., 2018; Pilon et al., 2019). The comparison of different molecular approaches for the analysis of cyanobacterial sedDNA has also been recently performed (Mejbel et al., 2021b).

Despite these advances, many questions remain regarding long-term cyanobacterial trends and the relative importance of drivers is not known. For example, it is still not clear whether the effects of climate change, eutrophication, and other stressors such as acidification, act synergistically, are additive, or reductive. In addition, there is still some debate regarding the impact of nitrogen on cyanobacterial dynamics and the effectiveness of controlling for nitrogen to limit cyanobacterial proliferation (Scott and McCarthy, 2010; Wurtsbaugh et al., 2019). Some studies have shown that limiting nitrogen levels in lakes has the potential of slowing down cyanobacterial growth (Paerl, 2014). Other have found that there is an elevated risk of toxin

production at low nitrogen:phosphorus ratios (Orihel et al., 2012). Furthermore, differences in N:P stoichiometry can lead to dominance of different taxa of bloom-forming cyanobacteria with varying toxicity (Dolman et al., 2012).

Further validation of the use of sedDNA is required, however. To my knowledge, only one other study has validated the use of sedDNA by comparing cyanobacterial richness estimates from sediment archives to phytoplankton pelagic records (Monchamp et al., 2016). Monchamp et al. (2016), examined sediments collected from natural perialpine lakes (one eutrophic and one mesotrophic) in Switzerland that have undergone cultural eutrophication. However, to address some of the questions mentioned above, and to examine dynamics that are occurring as a result of natural changes vs. anthropogenic influences, different lake types with long-term phytoplankton monitoring records should be tested.

In this study, sediments from four lakes at the Experimental Lakes Area (ELA) were collected to analyze long-term cyanobacterial trends in relatively untouched vs. experimentally manipulated systems. ELA, located in Kenora, Northwestern Ontario, is a controlled site with 58 lakes and watersheds used for scientific research. ELA has been the site of several ecosystem-level experimental manipulations of anthropogenic stressors on selected lakes since the late 1960s and the monitoring records for some of the lakes go back almost 50 years (Schindler, 2009). Reference lakes in this area are remote, largely unimpacted by human development, and the only natural drivers are those typical of the boreal forest (for e.g., forest fires) and those related to climate change. Here, two experimentally manipulated lakes, lakes 227 and 223, were selected to assess cyanobacterial responses to eutrophication and acidification, respectively. Lake 227 (L227) was an experimentally fertilized lake with constant annual inputs of nitrogen and phosphorus beginning in 1969. As of 1990 onwards, the lake continues to be fertilized with

phosphorus alone. Nitrogen loading stopped after 1989 because it was found that nitrogen reduction did not decrease eutrophication nor substantially change phytoplankton biomass (Schindler et al., 2008). In Lake 223 (L223), sulfuric acid was added from 1976 until 1983 when the pH was gradually dropped from ~7 to 5 (Schindler et al., 1980; Findlay and Kasian, 1986). Two oligotrophic and unmanipulated reference lakes, Lake 224 (L224) and Lake 442 (L442), were used in this study to compare how lakes exposed to climate change compare to the eutrophic (L227) and acidified (L223) lakes. L224 is often chosen as a reference site because it is chemically typical of ELA (Armstrong and Schindler, 1971) and the phytoplankton community is representative of the community composition of other oligotrophic ELA lakes (Findlay and Kasian, 1990, 1991). Physiochemical properties of L442 are characteristic of natural lakes on the Boreal shield (Schiff et al., 2017). However, there were a few significant events at ELA that could also influence the cyanobacteria record. This includes a major forest fire near lakes 239 and 240 in 1980 and the collapse of a road south of L442 in 2007 (Schindler et al., 1996). Here, my goals were to reconstruct long-term cyanobacterial dynamics, including changes to cyanobacterial abundance and shifts in the community compositions, and compare these trends to known long-term environmental and experimental drivers using a combination of molecular tools. Because long-term phytoplankton records are available for each of these lakes, I also compared the sedDNA data to microscopic surface water data. This was performed to evaluate whether taxa-specific gene copy numbers, species richness, and the cyanobacterial community composition from sediments track the surface water record.

This study will help further elucidate long-term changes to cyanobacteria when exposed to both phosphorus and nitrogen, followed by phosphorus alone. Likewise, no other study has assessed the long-term impact of acidification on cyanobacterial dynamics using sedDNA. The

comparison of sedDNA archives between eutrophied, acidified, and reference lakes within the same landscape has also not been previously performed.

4.3 Materials and methods

4.3.1 Sediment collection and core chronology

The study sites, sediment collection and core chronology steps are described in Mejbel et al. (2021b). In brief, a freeze corer was used to collect sediment cores in March 2018 from the maximum depths of ELA lakes 227, 223, 224, and 442. Due to the limited amount of sediment, every other 0.5-cm interval from the cores of lakes 227, 223, and 442 was collected for molecular analyses while the remaining sediment intervals were collected for other downstream applications. Cores were dated by measuring ^{210}Pb, ^{137}Cs, and ^{226}Ra activity. Constant Rate of Supply (CRS) models were chosen to create the depth-age profiles (see Figure S3.2 in Appendix B).

4.3.2 DNA extractions, Droplet Digital PCR (ddPCR) and High-throughput sequencing (HTS)

The DNeasy PowerSoil Kit was used to extract DNA according to the manufacturer's manual but minor modifications were made to the protocol. All DNA extraction details are described in Mejbel et al. (2021b). Four target genes were quantified using ddPCR. This included, 1) the cyanobacterial 16S rRNA gene (CYA) which is a highly conserved small subunit rRNA gene found in all cyanobacteria, 2) *Microcystis* 16S rRNA (MICR), a gene fragment specific to *Microcystis*— a genus ubiquitously found in freshwater environments, and 3) microcystin E (*mcyE*), a gene that is part of the operon involved in the production of the microcystin toxin. Microcystins are produced by several groups of cyanobacteria including *Microcystis*, *Anabaena*, and *Synechococcus*, but it is unknown how much microcystins are

present in these lakes and no study has measured the absolute quantity of cyanobacterial toxins in these sites. Finally, as a proxy of total bacterial abundance, glutamine synthetase (*glnA*), a housekeeping and essential gene needed for nitrogen metabolism in all bacteria, was also quantified. While the absolute abundance levels of CYA and MICR were previously described in Chapter 2, results were revaluated here in order to compare trends with the other target genes and because a greater sample size was analyzed in this chapter. Also, a reduced sample size was used in Chapter 2 since HTS was not performed on all the sediment samples, while the full sample set is presented here.

Primer sequences are presented in Supplementary Table S4.1. Reaction conditions were optimized, and abundances were normalized to gene copy number per gram of dry sediment (copies/g) using the methods described in Mejbel et al. (2021b). In brief, all four primers were first validated using quantitative PCR and gel-electrophoresis, followed by temperature and dilution optimization using the ddPCR technology (as suggested by Taylor et al., 2017). The absolute abundance of each target gene was subsequently quantified using ddPCR according to the manufacturer's instructions with minor modifications made to the protocol (see Mejbel et al., 2021b).

The limits of detection (LOD) were calculated using the methods described in Armbruster and Pry (2008). First, the limit of blank (LOB) was determined by using 20 blank samples (i.e., nuclease-free water), calculating the mean concentration of those blanks, and adding it to 1.645 the standard deviation (SD): $LOB = mean_{blanks} + 1.645(SD_{blanks})$. The LOD for each gene was subsequently quantified using: LOD= LOB +1.645(SD of low analyte samples). The LODs for CYA and MICR were previously described in Mejbel et al., (2021b). To determine the LOD for *mcyE* and *glnA*, twenty-two pooled DNA samples, diluted 100x for *mcyE*

and 1000x for *glnA*, were used to determine the SD of the low analyte samples. The LODs (as copies/µl and as copies/g) are presented in supplementary Table S4.2.

The high-throughput sequencing quality control and data processing steps are described in Mejbel *et al* (2021b). The QIIME2 pipeline (v 2019.7; Bolyen et al., 2019) was used to process the data and the Greengenes reference gene database (v gg_13_8_99) was used. The Greengenes database was chosen to analyze the trends here as it has been found to have fewer misannotations compared to other, more recently updated databases (such as SILVA or RDP) when corroborated against the nomenclature used by cyanobacterial taxonomists. The nomenclature used in Greengenes is also more consistent with the nomenclature used for the surface water monitoring of phytoplankton communities at ELA (Findlay et al., 1999). However, another classifier was trained on the SILVA 99% identity Full Seq OTU SILVA database (v 138) in order to assess L227 shifts more thoroughly. This was because Greengenes yielded a high number of unassigned features at lower taxonomic levels (i.e., genus and species).

For the diversity analyses, all samples from each lake were rarefied to an equal sampling depth. Sampling depths were set based on the sample that had the lowest sequence count. This corresponded to 13 114 sequences for L227, 4802 sequences for L223, 11 698 sequences for L224, and 4613 sequences for L442. Diversity metrics, including the number of observed features (or observed amplicon sequence variants [ASVs]; as a qualitative measure of community richness), Shannon's Diversity Index (as a quantitative measure of community richness), and Faith's Phylogenetic Diversity (as a qualitative measure of community richness incorporating phylogenetic relationships between features), were computed across different depths of the cores. Diversity values were averaged from 10 iterations (the number of rarefied tables computed at each sampling depth). For L227, diversity metrics were computed for the

period of pre-lake manipulation (pre-1969), the period of both phosphorus and nitrogen loading (1969 – 1989), and the period of phosphorus loading only (1990 – 2018). For L223, diversity metrics were computed for the period of pre-acidification (pre-1976), the acidification period (1976 – 1983), and the post-acidification period (1984 – 2018). For reference L224 and L442, diversity metrics were computed based on core depths (top, middle, and bottom).

4.3.3 Climate and nutrient data

Historical meteorological data for this region were obtained from the National Climate Data Archive of Environment and Climate Change Canada from the Kenora A station (49.79°N, 94.37°W). The effects of common climate variables were tested. This included minimum, maximum, and mean temperature, total precipitation (rain and snow), and heating and cooling degree days.

Water chemistry data (obtained from IISD-ELA) from the past several decades were available for each lake but the range of available data varied (supplementary Table S4.3). Typically, lakes were sampled once or twice a month for water chemistry variables during the ice-free season (May to October). Mean annual eplimnetic concentrations of total phosphorus (TP) and total nitrogen (TN) were used to evaluate the effects of nutrients on the cyanobacterial community in each lake.

4.3.4 Comparison of phytoplankton surface water enumerations to molecular data

To evaluate and verify the use of sedDNA to reconstruct temporal dynamics, molecular data were compared to historical pelagic phytoplankton records (data obtained from Scott Higgins, IISD-ELA). Phytoplankton samples from the epilimnion and metalimnion stratums

were historically collected from the centre buoy of each lake by members of ELA for microscopic identification and counting (e.g., Findlay et al., 1999). Counting was and continues to be conducted by the same analyst. The range of years with available surface water phytoplankton data were about the same as the years of available water chemistry data (see supplementary Table S4.3). Samples for microscopy were generally collected once or twice a month during the ice-free season (May to October). About two to four samples were collected from under the ice, but these measurements were removed from analysis for this study as this process was not consistent throughout the years. Productivity is low below the ice and not likely to impact the sediment records. Averages were calculated for the months where biweekly measurements were collected. Due to somewhat uneven sampling efforts between the years, only the monthly averages of May to September were calculated. Previous comparisons between sediment pigment data and water pigment data (as in Leavitt and Findlay, 1994) found that using counts for the entire year versus using 'only summer' data did not differ (personal communication; Peter Leavitt, University of Regina). Monthly averages were subsequently summed to obtain an annual estimate of cyanobacterial cell abundance (cell count/mL). To compare water and sediment abundance data, the annual estimate obtained from the water records was calculated by averaging the value of year before, year of, and year after the approximate year of the sedDNA samples (as performed by Monchamp et al., 2016). This was performed in order to correct for any potential errors in the sediment dating. I compared annual pelagic cyanobacterial cell density (cell count/mL) with sedDNA gene copy numbers (copies/g) obtained from ddPCR, the relative abundance of the common major orders between the two datasets, and cyanobacterial species richness (i.e., number of different cyanobacterial taxa) from the water with the ASV richness (the number of unique sequence variants) data obtained from

diversity analysis. Common trends in cyanobacterial abundance and community composition shifts from the surface water records were compared with the trends obtained from sedDNA.

4.3.5 Statistical analyses

The absolute gene copy numbers (GCN) as copies/g obtained from ddPCR were checked for normality using Shapiro-Wilk tests. Data that did not fit a normal distribution were log or square root transformed. GCNs for CYA, *glnA*, MICR, and *mcyE* were plotted as a function of sediment depth. Pearson's and Kendall's (for those that did not fit normal distribution) correlation coefficients were calculated to evaluate the relationship between the target genes. GCNs that were below the corresponding LOD were replaced with half the LOD (as suggested by the United States Environmental Protection Agency for water quality data at detection, 2006). Temporal autocorrelation is an issue occasionally seen with environmental data in which, because internal conditions are similar between samples taken closely together in time, the datapoints lack true independence. As such, GCNs were assessed for temporal autocorrelation using the autocorrelation *ACF* function in R and, *p-values* were corrected (using the *permute* function) if both variables were temporally autocorrelated (Legendre and Legendre, 2012). The cell density (copies/mL) of cyanobacteria from the water and absolute CYA abundance from the sediment (copies/g) were plotted as a function of time to compare cyanobacterial trends between 1970 and 2017.

For cyanobacterial community dynamics, the *phyloseq* package was used to import the ASV table, the taxonomic assignment, project-specific metadata, Shannon Index data, and visualize and subset the data. Assigned ASVs that did not appear at least 3 times in 20% of the samples were removed to protect against rare taxa biases (McMurdie and Holmes, 2013).

Mann-Kendall trend analyses were performed to identify any upward or downward trends in temperature, total annual precipitation (TAP), heating degree days (HDD), cooling degree days (CDD), TP, and TN. Correlation coefficients between abiotic climate variables were calculated. Because minimum, mean, and maximum temperature were highly correlated ($r > 0.9$, $p < 0.01$), only mean annual temperatures (MAT) were used in the models. Similarly, HDD and CDD were also strongly correlated ($r = -0.85$, $p < 0.001$), and thus only HDD was used. MAT and HDD were weakly correlated ($r = -0.37$, $p < 0.01$) and both were used as climate-associated variables. Correlation coefficients between TP and TN showed that the two were weakly or moderately correlated ($r = 0.22 – 0.65$), and thus both TP and TN were used in the analyses. Annual measurements of meteorological and nutrient data were calculated based on the year before, year of, and year after around the approximate year of the sedDNA sample.

Redundancy analysis (RDA) was used to evaluate the environmental drivers that significantly explained the variation in the cyanobacterial community composition in each lake. The climate-associated data (MAT, TAP, and HDD) were scaled and centred to account for differences in units and the relative abundance of the cyanobacterial orders were square root transformed (Legendre and Gallagher, 2001). TP and TN were log-transformed to improve normality. Forward selection modelling (with 999 permutations) was first performed to pre-select significant variables for analysis from the two groups (climate and nutrients). The respective effects of the variables that were significant ($p < 0.05$) were further evaluated using variation partitioning (Peres-Neto et al., 2006). This was performed to distinguish the effects of one predictor from another and to compare how much co-influence the explanatory environmental factors had on the cyanobacterial community in the experimental and reference lakes. To test the significance, permutations tests on the models were performed (minimum 999

permutations) using *anova*. The *varpart* function of the *vegan* package in R was used to perform the partitioning. The RDAs and variance partitioning were performed on a reduced subset for which both environmental and nutrient data were also available.

For the comparison between sedDNA and pelagic microscopy data, correlations were performed between log-transformed ddPCR GCNs (copies/g) and cyanobacterial cell copies (cell count/mL). Correlations were also performed between the relative abundances of the three most common orders (based on the HTS of the sedDNA samples and on the surface water records). Proportions were transformed by employing arcsine square root transformations. The relationship between the sediment ASV richness and pelagic species richness was assessed through correlation analyses as well. All data analysis was conducted using R (open-source v.4.0.0 and RStudio v.1.2.5042).

4.4 Results

4.4.1 Trends in gene copy numbers over time

The sediment dating profiles for each lake are presented in Appendix A, Figure S2.2. In L227, prior to the phosphorus and nitrogen experimentation that began in 1969 (~15.25 cm in sediment depth based on the depth profile), the abundance of CYA and *glnA* showed very little change overtime (Figure 4.1, top left). Below 15.25 cm, CYA abundance increased from ~10^7 (approximately two orders above the corresponding LOD: ~4.1×10^5) to 10^8 copies/g of dry weight, whereas for *glnA*, abundance was consistently at ~10^8 copies/g (approximately two to three orders of magnitude above the corresponding LOD: ~8.9×10^5). The absolute abundance of *mcyE* was slightly above the corresponding LOD (~9.8×10^4 copies/g) at the bottom of the sediment core but increased gradually overtime to ~10^7 copies/g prior to the lake manipulation experiment. The abundance of MICR was consistently below the LOD (~1.4×10^5) for the

samples that were below 15 cm in sediment depth (prior to fertilization). When the eutrophication experiment began in 1969 (corresponding to ~15.25 cm), CYA, *glnA,* and MICR showed an increase in abundance. For CYA, abundance increased 100-fold, from ~10^8 to 10^{10} copies/g. For *glnA,* abundance increased slightly greater than 10-fold from ~10^8 to 10^9 copies/g, and for MICR, abundance increased 100-fold from ~10^5 to 10^7 copies/g overtime. The abundance of *mcyE* did not increase much in response to dual fertilization and remained stable at ~10^7 copies/g. However, abundance increased 10-fold from ~10^7 to ~10^8 copies/g, after nitrogen loading stopped (~11.25 cm corresponding to ~1988 ± 2 years).

In the acidified lake, L223, CYA abundance was at detection limits (~8.2×10^5) in the deepest sediment samples (Figure 4.1, top right). However, CYA subsequently rose to ~10^8 copies/g and increased over time to ~10^9 copies/g. The acidification experiment in that lake, which began in 1976 and ended in 1983, corresponded to ~18 – 20 cm in sediment depth based on the sediment dating profile. CYA gene copy numbers showed a slight 3-fold increase in response to this acidification, but decreased back to reference lake conditions at ~11 cm. For *glnA*, gene copy numbers gradually rose from ~10^7 to 10^9 copies/g in that lake. After the acidification experiment ended, *glnA* abundance fluctuated between 10^8 and 10^9 copies/g, then remained relatively consistent at 10^9 copies/g at top sediments. The abundance of MICR was consistently at or below detection limits (~2.8×10^5) but increased above detection limits after acidification and subsequently stayed only slightly above the corresponding LOD. Gene copy numbers of *mcyE* were consistently at or below detection limits throughout the entire sediment core.

In reference lakes 224 and 442, CYA and *glnA* gene copy numbers increased gradually overtime. In L224, CYA increased 100-fold from ~10^7 to 10^9 copies/g and *glnA* remained stable

between ~10^8 to 10^9 copies/g. In L442, gene copy numbers of CYA and *glnA* increased 10-fold from ~10^8 to 10^9 copies/g over time. Absolute abundances of MICR and *mcyE* were consistently at or below detection limits in both lakes, but abundances increased ~10-fold above the corresponding LODs in the surface sediments (top ~5 cm).

Almost all target genes were significantly correlated with one another (Table 4.1). In L227, the strongest correlation was between CYA and *glnA*. For both lakes 223 and 224, the strongest correlations were between CYA and MICR, although moderate to strong correlations were also found between CYA and *glnA*. For L442, CYA had strong correlations with all target genes. In lakes 227 and 224, gene copy numbers of all four genes showed some temporal autocorrelation. Temporal autocorrelation was also found for CYA and MICR gene copy numbers in L223. All correlations were still significant ($p < 0.05$) after correcting for autocorrelation.

4.4.2 Down core community composition shifts

Metabarcoding of the cyanobacterial 16S rRNA gene yielded 8 676 724 sequence counts distributed in 74 samples. After trimming, quality filtering and preprocessing, a total of 4 396 942 sequence reads were obtained. The number of sequence reads ranged from 33 157 – 112 003 in L227, from 37 627 – 68 650 in L224, from 34 125 – 76 732 in L224, and from 21 337 – 94 264 in L442. After removing all ASVs assigned to non-cyanobacterial taxa and chloroplasts, a total of 3 655 659 sequences remained.

Three orders dominated the overall community composition of all four lakes, Chroococcales, Nostocales, Pseudanabaenales (Figure 4.2). Less than 5% contributed to Synechococcales and unassigned ASVs. In L227, below 15.25 cm, composition was mostly

composed of either Pseudanabaenales or Chroococcales (Figure 4.2, top left). However, above 15.25 cm, there was a decrease in the relative abundance of Chroococcales (16% to 7%) and Pseudanabaenales (from 69% to 48%) in response to high nitrogen loading (1969 – 1974) (see Table S3.1 in Appendix B) whereas Nostocales increased (from 9% to 38%) during this period. A further increase to 73% was observed for this order when only phosphorus loading continued (1990 onwards). Species from the Nostocales order continued to have the greatest relative abundance, eventually making up more than 96% of the community in top sediment samples. Although the abundance of Synechococcales was low, there was a small gradual decrease in relative abundance after fertilization began (above 15.25 cm) (supplementary Figure S4.1).

Evaluations at the level of genus using the SILVA database showed that in L227, below 15.25 cm, the community was mostly dominated by *Leptolyngbya*, *Limnolyngbya*, *Dolichospermum*, and unknown species (Figure 4.3). Greengenes-generated genus profiles are presented in supplementary Figure S4.2. Based on the trends observed in Figure 4.2 and by manually exploring the generated taxonomy table, most of the unknown species in Figure 4.3 were identified coming from the Chroococcales order. Between core depths 11.75 and 15.25 cm, the relative abundances of *Leptolyngbya* (Pseudanabaenales) and unknown species decreased. Between the sediment depths of 8.25 and 11.75 cm, the cyanobacterial community was generally mostly dominated by *Dolichospermum* (Nostocales). Above 8.25 cm, *Aphanizomenon* (Nostocales) subsequently dominated the entire cyanobacterial community.

In L223, only slight shifts in the major orders were observed overtime. The community was primarily composed of taxa from the Chroococcales and Pseudanabaenales order (Figure 4.2, top right). However, in the top 3.75 cm, the relative abundance of Nostocales taxa increased, making up the majority of the cyanobacterial species in the surface sediments. Similar trends

were observed in reference lakes 224 (Figure 4.2, bottom left) and 442 (Figure 4.2, bottom right), where composition was mostly made up of Chroococcales and Pseudanabaenales species in deeper sediments with relative abundances of Nostocales increasing in surface sediments.

Shannon diversity and Faith phylogenetic measures were used to evaluate changes in alpha diversity. Rarefaction curves plateaued in all samples (supplementary Figure S4.3), indicating the sequencing reached depths that represented the whole cyanobacterial community. For L227, according to the Shannon Diversity and Faith phylogenetic diversity indices, diversity was greatest in the middle of the core (10.5 – 15.5 cm), roughly corresponding to ~1969 – 1990, the period of dual phosphorus and nitrogen loading (Table 4.2). Diversity decreased when only phosphorus loading continued (~1990 – 2018; 0 –10 cm) with a Shannon Index that was slightly lower than the index in deep sediments (pre-1850 – 1967; 16 – 52 cm). For L223, diversity seemed to increase over time, with the greatest Shannon Diversity and Faith Phylogenetic diversity indices from post-acidification sediments (~2000 – 2018; 0 – 13.5 cm) and the smallest indices in the pre-acidification sediments (pre-1900 – 1973: 21.5 – 42 cm) (Table 4.2). Similar results were also found for reference lakes 224 and 442, where diversity indices were highest in top and middle sediment samples (~0 – 17 cm) and lowest in the deepest sediment samples (below 20 cm).

4.4.3 Impact of climate and nutrients

The impacts of various climate and nutrient variables on the cyanobacterial communities were assessed. First, the long-term trends of these climate variables were evaluated using Mann-Kendall analysis. Based on the meteorological archives, the mean annual air temperature (MAT) at ELA has increased by approximately 2°C since 1950 (Kendall's $\tau = 0.33$, $p < 0.001$) (Figure

4.4a). The total annual precipitation (TAP) has also increased ~100 mm over the past five decades ($\tau = 0.30$, $p < 0.001$) (Figure 4.4b). There was a clear trend in decreasing heating degree days (HDD) ($\tau = -0.57$, $p < 0.001$) (Figure 4.4c) and increasing cooling degree days (CDD) ($\tau = 0.57$, $p < 0.001$) at ELA (Figure 4.4d).

Nutrient data for 2017 were removed because data were only partially available for that year. In L227, TP concentrations averaged ~39 μg/L in the epilimnion (Figure 4.5, top). In lakes 223, 224, and 442, annual TP concentrations averaged ~6 μg/L. Based on Mann-Kendall analysis, there were no significant upward or downward long-term trends for phosphorus ($p > 0.05$) between 1970 and 2016 in any of the lakes. TN concentrations averaged 930 μg/L for L227 overtime, 281 μg/L for L223, 236 μg/L for L224, and 333 μg/L for L442 (Figure 4.5, bottom). Mann-Kendall analysis showed a significant decrease in TN concentrations between 1970 and 2016 in L227 ($\tau = -0.378$, $p < 0.001$) but no significant upward or downtrend trend was found in the other lakes ($p > 0.05$ for L223, L224, and L442).

The significant variables to influence the cyanobacterial community in each lake was determined through forward stepwise regression (999 permutations) (Table 4.4). In eutrophied L227, the cyanobacterial community was significantly related to HDD and nutrients (TP and TN), but not to MAT or TAP. Variation partitioning was used to examine how much was explained by these variables. In total, HDD and nutrients contributed to 91.8% of the variation of the cyanobacterial community in L227 (Figure 4.6, $F=12.9$, $p=0.018$). HDD alone contributed to 10% of the variation while the nutrients contributed to 35.8%. Approximately 46% of the variation was jointly explained by HDD and nutrients. TN was positively related to Chroococcales and Pseudanabaenales and negatively related to Nostocales. HDD was weakly positively related to Pseudanabaenales and Synechococcales.

In acidified L223 and reference L224, HDD had a significant effect on the variation of the cyanobacterial community, but there were no significant effects of MAT, TAP, or nutrients in either lake (Table 4.4). Based on the RDA models, approximately 17% of the variation in the community was explained by HDD in L223. HDD was positively related to Pseudanabaenales and weakly positively related to Chroococcales in this lake. In L224, HDD explained 36% of the cyanobacteria community profile. HDD was positively related to Pseudanabaenales, Chroococcales, and Synechococcales. In reference L442, the community was significantly influenced by HDD and TAP (Table 4.4), but not MAT or nutrients. In total, about 54.5% of the variation in the cyanobacterial community was explained by HDD and TAP (Figure 4.7). HDD contributed to about 51% of the variation, while TAP explained 12.9% of the variation. There was no co-influence of HDD and TAP on the cyanobacterial community in L442. HDD was positively related to Pseudanabaenales and Chroococcales. TAP was weakly positively related to Nostocales.

4.4.4 Comparison between the sediment and water data

Cyanobacteria cell density trends from the water column monitoring archives were compared with absolute CYA abundance trends from the sediment (supplementary Figure S4.4). Water phytoplankton records for L227 were available for the years of 1970 until 2017. In both the water and sediment data of L227, cyanobacterial abundance increased over time. Based on the cell density of cyanobacteria in the water, abundance increased ~10^5 to 10^6 overtime, with a peak in abundance between ~1985 – 1990 to ~10^7 cells/mL (supplementary Figure S4.4, left). Based on the sediment data, and as previously mentioned, CYA also increased 100-fold in response to eutrophication from 10^8 to 10^{10} over time (supplementary Figure S4.4, right). While

CYA abundance (copies/g) and cyanobacterial cell density (counts/mL) both appeared to peak in the mid 1980s, the sediment data showed that CYA absolute abundance remained stable and did not decrease after this time, whereas the water column cell density data showed a subsequent decrease from ~10^7 cells/mL back down to ~10^5 to 10^6 cells/mL after 1990.

Phytoplankton data for L223 were consistently monitored between the years of 1974 to 2004. As such, only the sediment samples that roughly corresponded to that period were used in this comparison. Based on the water data, cyanobacteria cell density increased 10-fold from ~10^3 to 10^4 between the years of 1980 and 1990, and subsequently decreased back to ~10^3 levels after that period. This roughly corresponds to the period of acidification (1976 – 1983) and the natural recovery of the lake (post-acidification). From the sediment, CYA abundance also increased around 1980, but only ~3-fold, and declined in abundance in the mid-2000s. Water phytoplankton records for L224 were available for the years of 1974 to 2017, while for L442, data were only available for the years of 1990 to 2017. For reference lakes 224 and 442, the surface water monitoring records showed that cyanobacterial cell density remained consistent between ~10^2 – 10^3 overtime but peaked suddenly to ~10^5 for the years of 2013-2017. In comparison, absolute CYA abundance from the sediment showed gradual 10-fold increases overtime (~10^8 – 10^9) for both lakes 224 and 442.

Community composition shifts in the water column archives were also compared with the sedDNA archives. In L227, surface water records showed that after 1990 (with the exception of 1997 and 2011), the Nostocales order dominated the overall cyanobacterial community (supplementary Figure S4.5, top left). This was also observed in the sequenced data of the sedDNA (Figure 4.2, top left). The relative abundance of Nostocales as derived from HTS was strongly correlated ($r = 0.88$, $p < 0.001$ $n=11$) with the relative abundance of this order in the

surface water records (Table 4.3). However, there were no significant correlations in either direction for the other two major orders, Chroococcales and Pseudanabaenales. The relative abundances of Oscillatoriales obtained from the surface water records were compared with the relative abundances of Pseudanabaenales from the sequenced sedDNA (Table 4.3). This harmonization was performed because, based on the updated Algaebase (Guiry and Guiry 2021), the two orders share similar morphological features and thus microscopic differentiation between Pseudanabaenales and Oscillatoriales species can be difficult.

In L223, surface water records showed that between 1974 and 2004, the community was mostly composed of Chroococcales and other unicellular cyanobacteria (supplementary Figure S4.5, top right). This was consistent with the sequenced sedDNA samples (Figure 4.2, top right), except during this period, the sequenced data showed that ~20% of the relative abundance was composed of Pseudanabaenales which was not observed in the surface water records. In the two reference lakes, the historical water records show that the abundance was mostly dominated by Chroococcales and Oscillatoriales with the relative abundance of the Nostocales order increasing in recent years (post-2012) (supplementary Figure S4.5, bottom). These trends were generally the same as the trends observed in the sequenced data (Figure 4.2, bottom). In the surface water records of all four lakes, there was generally a higher amount of cyanobacteria labelled as other unicellular or other filamentous cyanobacteria in the first few years of monitoring. For lakes 223 and 224, there was no significant correlation found between the relative abundances of the major orders derived from the sedDNA with the relative abundances derived from surface water records (Table 4.3). There was a weakly significant negative relationship between Pseudanabaenales and Oscillatoriales in L442.

Correlations were performed to evaluate the relationship between log-transformed absolute CYA abundance (copies/g) from the sediment and annual cyanobacterial cell density (cell count/mL) from the water (Figure 4.8). The two archives were strongly correlated in L227 ($r = 0.87$, $p < 0.001$, $n=13$). There were moderate correlations between the two datasets in L224 ($r = 0.43$, $p < 0.05$, $n=25$) and in L442 ($r = 0.62$, $p < 0.05$, $n=11$). No correlation between CYA abundance in the sediment and cyanobacterial cell density from water was found in L223 ($r = 0.07$, $p > 0.05$, $n=11$).

Correlations were also performed between annual ASV richness from the sediment based on HTS of the cyanobacterial 16S rRNA gene and annual species richness from the water data (Figure 4.9). In L227, the number of unique ASVs varied between 77 and 584 from the sedDNA samples, whereas species richness varied between 4 and 15 based on the water microscopy data. There was no significant correlation between the two methods for L227 ($r = 0.26$, $p > 0.05$, $n=11$). In L223, the number of ASVs varied between 39 and 622, whereas species richness in the water ranged between 5 and 8. The correlation was significant for L223 ($r= 0.79$, $p < 0.05$, $n=8$). In L224, ASV quantity ranged between 392 and 703 from the sediment, and species richness ranged between 13 and 18. The correlation was also statistically significant in L224 ($r = 0.74$, $p < 0.05$, $n=8$). Finally, in L442, the number of ASVs ranged between 101 and 387 from the sediment samples, whereas the water species richness varied between 10 and 22. The correlation between the two approaches was significant in L442 ($r = 0.73$, $p < 0.05$, $n=9$).

4.5 Discussion

Through the analysis of sedDNA and through the comparison of a eutrophied, acidified, and two reference lakes, I found that nutrients and climate variables influenced long-term

cyanobacterial dynamics, but that the impacts differed between the lakes. Nutrients had an important influence on increasing the absolute abundance of cyanobacterial genes and explaining the variation of the community in L227. Acidification had lesser effects on cyanobacterial abundance and community. Heating degree days had the greatest influence in explaining the cyanobacterial community variation in the reference lakes. The results also showed that both ddPCR and HTS approaches collectively validated the use of sedDNA to reconstruct cyanobacterial dynamics far back through time. The use of sedDNA provided a more comprehensive description of the cyanobacterial abundance and community in these lakes and added new knowledge beyond that of the monitoring record.

4.5.1 Temporal changes in gene abundance in response to environmental drivers and their co-influence

It is well established that excessive phosphorus loading and eutrophication can lead to algal bloom formation (Carpenter, 2005; Schindler et al., 2008; Taranu et al., 2015; Molot et al., 2021) and the dominance of cyanobacteria (Pick and Lean, 1987; Downing et al., 2001; Beaulieu et al., 2013). Here, results from L227 showed clear increases in cyanobacterial abundance corresponding with the periods of lake fertilization. This is consistent with the known history of the lake, in which a significant increase in cyanobacteria and algal biomass was observed when phosphorus and nitrogen fertilization began in 1969 (Schindler et al., 2008; Paterson et al., 2011; Higgins et al., 2018), leading to a greater dominance of cyanobacteria in terms of biomass. These results also showed, however, that despite continuous phosphorus-only loading after 1990, the absolute abundance of all genes remained relatively stable and only appeared to increase slightly in the surface sediments. These results were generally consistent with other ELA lake studies where they found that surface water cyanobacterial abundance did not dramatically decrease

after nitrogen additions stopped (Schindler et al., 2008; Molot et al., 2021). The lack of a cyanobacterial abundance stimulation during the period of excess phosphorus-enrichment (1990 – present) was surprising but demonstrated that either both nitrogen and phosphorus potentially stimulate cyanobacterial growth more than just phosphorus fertilization alone, or that the community was already stimulated with eutrophication and subsequently reached a new baseline.

The sedDNA analysis revealed that levels of a gene marker for microcystins can increase as well in response to continued phosphorus loading with a reduction in nitrogen. Specifically, I found that levels of *mcyE* began to increase after ~1988 ± 2 years (11.25 cm sediment depth), corresponding to when nitrogen loading stopped in this lake. There was only a moderate correlation between MICR and *mcyE* in this lake ($r = 0.60$) suggesting other toxigenic genera aside from *Microcystis* may be present. Microcystins are cyclic, nitrogen-rich, heptapeptides that can be produced by different cyanobacteria genera including *Microcystis*, *Anabaena* (recently renamed *Dolichospermum*), *Aphanizomenon*, and *Planktothrix* (Tillet et al., 2000; Renaud, 2009; Sivonen, 2009; Lyon-Colbert et al., 2018). Here, I found that after nitrogen loading ceased, the abundance of toxin producing nitrogen-fixers increased. These results and the theory that total nitrogen availability impacts microcystin concentrations is supported by other field and laboratory experiments (Dolman et al., 2012; Monchamp et al., 2014). Although it is generally accepted that microcystin concentrations increase with lake trophic status (Kotak and Zurawell, 2007), the results were more consistent with studies showing a potentially elevated risk of microcystin production as the nitrogen-to-phosphorus ratio decreases (Dolman et al., 2012; Orihel et al., 2012). Orihel et al., (2012) found that the probability of microcystin concentrations exceeding toxin thresholds was greatest when the N:P ratio was < 20. In L227, the N:P ratio was > 20 during the period of dual phosphorus loading and nutrient loading, but < 20 when only

phosphorus loading continued (see Table 3.1 in Appendix B). I do not suspect that the increase in *mcyE* is due to DNA degradation with depth, as microcystin-synthesizing genes (including *mcyD*) have been found in high gene copy numbers going far back in time well before European settlement (Pilon et al., 2019). Altogether, these findings indicate that a lake management paradigm focused on the reduction of phosphorus only may be insufficient and suggest the need to explore dual nutrient management strategies to mitigate cyanotoxin accumulation.

When interpreting an increase in gene abundance, it is important to note that cyanobacterial 16S rRNA can vary between 1 – 4 copies per cell in some species (Schirrmeister et al., 2012), with heterocystous and unicellular cyanobacteria differing in copy numbers. *Microcystis* 16S rRNA can vary between 1 – 2 copies per cell (Kaneko et al., 2007; Kyoung-Hee et al., 2012), but the *mcy* gene cluster (Kaebernick et al., 2002; Tanabe et al., 2004; Kyoung-Hee et al., 2012) and *glnA* (Reyes and Florencio, 1995; Zhang et al., 2000) are found only once per cell. This variation could interfere with the quantitative performance of the approaches used in this study and lead to the overrepresentation of cyanobacterial abundance. However, based on both the significant correlations found between the genes and *glnA* and the increase in both microbial and cyanobacterial abundance overtime in all cores, it is expected that the patterns observed were not due to variation in the number of copies per cell but rather indicative of the changes to the overall abundance levels of cyanobacteria in response to changes within the lake. In addition, CYA generally increased by orders of magnitude overtime, not 4-fold. Although, it could be speculated that the small increase in CYA observed during the acidification period in L223 (~18 – 20 cm) could be attributed to this (i.e., a switch in dominance from taxa that had 1 copy of the cyanobacterial-specific 16S rRNA gene per cell to taxa that had 4 copies per cell). Findlay and Kasian (1986), however, found an increase in cyanobacterial biomass during the

lake acidification experimentation and as the sedDNA data reflected somewhat similar patterns, it is unlikely that the increase was due to variation in the copy number of the 16S rRNA gene.

4.5.2 Shifts in the cyanobacterial community profile

Major shifts in the cyanobacterial community composition were only observed in L227, indicating that significant shifts in the overall cyanobacteria community profile is dependent on nutrient loading. Schindler et al. (2008), found that when the nitrogen loading was reduced in 1974, nitrogen-fixing *Aphanizomenon* (from the Nostocales order) increased, and the duration and predictability of monospecific *Aphanizomenon* blooms further increased when nitrogen loading stopped in ~1990 (Findlay et al., 1994; Hendzel et al., 1994). These results were also observed in this study, in which, based on the sequencing of the sedDNA, the relative abundance of Nostocales quadrupled in response to dual-nutrient loading, and subsequently made up almost the entire cyanobacterial community when only phosphorus loading continued. As previously mentioned, the increase in *mcyE* abundance around 1990 could reflect more microcystin-producing nitrogen-fixing cyanobacteria.

A shift towards Nostocales dominance in surface sediments was found in the other lakes assessed in this study. It has been suggested that the increase in abundance and distribution of cyanobacterial species, particularly Nostocales, in temperate regions may be due to the rise in global annual temperatures and climate change (Carey et al. 2012; Sukenik et al. 2012). An analysis of cyanobacterial diversity using sediment cores collected from northern Canadian lakes also observed a switch from Chroococcales dominance towards Nostocales dominance in more recent sediment samples (Pal, 2015). At ELA, the warmer temperatures, longer growing seasons, and longer thermal stratification, are all conditions likely to favour Nostocales growth. In ocean

water, the combined effects of acidification and warming have also usually been found to promote the growth of nitrogen-fixers (Hutchins et al., 2007, 2019; Hutchins and Fu, 2017). While pH levels have since recovered in L223, it is possible that the combined effects of acidification and warming during the acidification experiment triggered a shift in the community that became evident as time progressed. These findings demonstrate that by controlling for nitrogen, and if temperatures continue to rise, a switch towards *Nostocales* dominance (including toxigenic strains) may be observed in temperate regions and lakes exposed to acidification.

Diversity metrics also showed that cyanobacterial diversity seemed to increase over time but only in lakes 223, 224, and 442. Phytoplankton diversity is largely controlled by productivity and grazing pressures (Proulx et al., 1996) but climate may play a role, particularly temperature (Ibarbalz et al., 2019; Righetti et al., 2019). Based on the more recent switch towards Nostocales dominance observed in the surface sediments, it is expected that if temperatures continue to rise, then a reshuffle in the community composition may be observed and this could lead to a decrease in diversity. In L227, while the alpha diversity metrics showed that diversity increased slightly during the period of dual phosphorus and nitrogen loading, diversity decreased when only phosphorus loading continued, suggesting that a reduction in phosphorus and nitrogen fertilization levels can lead to a shift in alpha diversity, if there has been prior eutrophication. The consistent loss of particular taxa in L227 in response to nutrient additions and the increase in Nostocales, as well as the concurrent loss of rare taxa are the most likely explanation for the decrease in diversity in this lake.

4.5.3 The influence of environmental drivers on cyanobacterial communities

The ELA lies within the Boreal Forest, where even slight changes to temperature and precipitation levels have caused significant changes to lakes and streams in that region (Schindler et al., 1996). Longer growing seasons and rapid warming can have both direct and indirect effects on phytoplankton, including stronger stratification in the water column and less water mixing. Meteorological records show that mean annual temperatures have increased by approximately 0.42°C per decade in Northwest Ontario, while the global average increase per decade is approximately 0.15 – 0.30°C. In addition, annual precipitation levels have increased by ~20 mm per decade since 1970. HDD have also decreased by approximately 500 days over the past century in Ontario and are expected to continue to decrease (Bush and Lemmen, 2019). Since 1970, the number of ice-free days at ELA has increased by approximately 16 days (Wiltse et al., 2016).

Defining the complex interactions between climate variables, nutrient loading, and lake hydrological processes is difficult. To overcome some of these challenges, redundancy analysis and variance partitioning were used to identify the drivers that explained variation in the cyanobacterial community. Although water chemistry data has been tracked for decades at ELA and DNA was sequenced from top, middle, and bottom intervals of the cores, matching the water chemistry data to the sediment samples from the same timeframe reduced the overall sample size (especially in L442). Despite the small sample sizes, some differences were found between the reference and experimental lakes using these reduced datasets. Precipitation influenced only the community composition of L442 (explaining ~12%). This may be due to differences in the watershed area and lake volume, the species composition of surrounding trees, or the sample size. In all four lakes though, HDD influenced the variation of the community profiles. While

HDD explained 36% and 55% of the variation in L224 and L442 respectively, it had a lesser effect in the two experimental lakes (227 and 223). In L223, HDD explained about 17% of the variation, and none of the other variables evaluated had a significant effect, indicating that the variation might be explained by other confounding variables. In L227, only 10% of the variation was explained by HDD, and most of the variation was explained by either nutrients alone or the shared co-influence of nutrients and HDD. These results suggest that the HDD or other climate-related variables could be more important in oligotrophic lakes (TP < 10 μg/L in lakes 223, 224, and 442) than in eutrophic lakes (TP = 30 – 100 μg/L in L227) whereas nutrients (whether individually or synergistically with HDD) have greater impacts on explaining community variation in eutrophic lakes. While previous work evaluating the interaction of climate warming and eutrophication have indicated that both nutrients and temperature are important in explaining cyanobacteria biomass in eutrophic lakes, they found that the individual impact of nutrients alone was more important in oligotrophic lakes (Rigosi et al., 2014). To further evaluate the findings presented here, a larger study and a higher sample size would be necessary. Other physical lake properties such as water-column stability and stratification (which are affected by climate-change) could influence variation in the cyanobacterial community composition and should also be assessed (Steinberg and Hatmann, 1988; Taranu et al., 2012).

In addition to eutrophication, increased acidification has also been shown to increase cyanobacteria and phytoplankton biomass at least within the pH range of ~5 – 7 (Findlay and Kasian, 1986; Hutchins et al., 2007). The pH monitoring record was limited, and the synergistic impacts of pH and warming could not be evaluated in this study. In acidified L223, I found that CYA and MICR absolute abundance levels increased in response to the acidification experiment (1976 – 1983), indicating that both moderate acidification and warming in this region support

cyanobacterial growth. This was consistent with previous studies in L223, where it was found that phytoplankton biomass increased as pH levels decreased (Findlay and Kasian, 1986). The impacts of acidification on microcystin spread have been rarely evaluated. Overall, the results suggest that while acidity and a slight decrease in the pH level do support cyanobacterial growth, pH levels between 5 – 7 will not necessarily contribute to an increase in *mcyE*. Further work is required to evaluate the impacts of low pH levels on *mcyE* gene expression and microcystin production.

4.5.4 Congruency between the sediment data and surface water records

There have been a few environmental DNA studies that have assessed the coherence between molecular analyses and microscopy (Xiao et al., 2014; Monchamp et al., 2016; Kim et al., 2019). However, as the use of sedDNA to reconstruct historical aquatic ecosystems is increasing, there remains a need to evaluate these methods from different lake types. Here, when comparing the sedDNA absolute abundance data with the cell counts obtained from surface water records, there was only a strong correlation between the datasets for L227. Moderate correlations were found for the reference lakes, but no correlation was found in L223. While this may suggest that changes in lake pH can interfere with the accurate quantification of cyanobacteria, it is more likely that a combination of both methodological and biological reasons explains the discrepancy.

One concern with traditional paleolimnological approaches is that plankton identifications and cell counts can vary depending on the expertise of the researcher and the counting approach. At ELA, one analyst has performed the phytoplankton counts continuously over the past five decades. However, there may have been slight changes to the counting

approach overtime with improved taxonomic resolution (personal communication Cyndy Desjardin, IISD-ELA). These improvements over time, recent changes to the recording database, and the switch to a new system (due to the ELA transitioning from being a government facility to one run by a non-governmental organization) may have all contributed to the sudden rise in cyanobacterial cell density observed in reference lakes 224 and 442 in the recent monitoring years (2013 – 2017). On the other hand, the molecular experiments were generally performed in a consistent and repeatable manner and within a relatively short period of time. It is also important to note that in sedDNA studies, both planktonic (floating) and benthic (bottom-dwelling) taxa are collected in lake sediments through sediment focusing and therefore sedDNA integrates both communities over space and time. In contrast, with the surface-water monitoring at ELA, the benthic taxa are excluded from the record as they were not regularly monitored. As a result, the water column and sediment relationship would not be a perfect match. However, given that the pelagic volume is far greater than that of the littoral area in these deep lakes, the contribution of pelagic taxa to the sediment record would be far greater than that of the benthic littoral zone.

DNA degradation issues may have also contributed to the variability observed between the two methods. Larger gene fragments are more likely to undergo DNA degradation. In this case, the target genes used for ddPCR have small fragment sizes (< 300 bp) and good quality DNA was extracted from even the deepest sections of the sediment cores (see Chapter 3). The fate of the DNA could also depend on whether the conditions at the bottom of the water column are suitable for DNA preservation. It is generally accepted that cold, dark, anoxic conditions are most favourable for long-term DNA preservation. The lakes assessed in this study are relatively deep and stratified. L227 is the most shallow (maximum depth ~10.6 m), yet due to the high

phytoplankton biomass as a result of the continuous fertilization, the hypolimnion has become anoxic and the conditions have become favourable for long-term preservation. This suggests that DNA degradation issues may not be a particular concern in this study. Although, preservation biases between cyanobacterial groups may exist.

Differences in preservation capacity could interfere with the interpretation of abundance and community data in sedDNA studies. Under controlled conditions, I found that the *Anabeana*-specific 16S rRNA gene (used to quantify the abundance of *Trichormus variabilis*, also known as *Anabaena variabilis*; formerly *Anabaena flos-aquae*, from the Nostocales order) decayed at a slower rate in sediment compared to the *Microcystis*-specific 16S rRNA gene (used to quantify the abundance of *Microcystis aeruginosa*, from the Chroococcales order) (Mejbel et al., 2022). As such, the observed increase in Nostocales abundance in L223, L224 and L442 sediment cores could be, in part, attributed to better preservation by Nostocales species in comparison to Chroococcales species. On the other hand, if the abundance of Chroococcales decreased in recent sediments, then the observed trend is indeed true (and even more so) because, because the Chroococcales DNA in older sediments would have decayed considerably more. If Nostocales species overall do preserve better than Chroococcales species or even species from other genera, it is possible that the increase in alpha diversity observed in these lakes may also be the result of differences in preservation and taxonomic bias.

In the comparison between species richness from water and ASV richness from sediment, strong significant correlations were only observed for lakes 223, 224, and 442, and overall, relative abundance levels were not correlated between the two records (other than in L227 for the Nostocales order). This indicates that the use of relative abundance data using sedDNA should be interpreted with caution as also indicated by Monchamp et al. (2016). The differences

observed between HTS and surface water data could be due to the reference database used in this study. Here, both Greengenes and SILVA were used for taxonomic assignment, but Greengenes has not been updated since 2013, and taxonomy naming and designations in SILVA have been found to be somewhat erroneous (Komárek, 2010; MacKeigan et al., 2022). The lack of updates to Greengenes and the lack of exhaustive and well-curated reference databased may have contributed to the differences found between the sedDNA sequenced data and the surface water records.

In addition, species richness from the sedDNA data was measured based on the number of ASVs which distinguishes sequence variation by a single nucleotide. This avoids clustering units based on similarity (as in operational clustering units) and is a much more precise measurement of sequence variation, capable of quantifying and distinguishing between even cryptic species. This could be another explanation as to the discrepancy found between the molecular and microscopy data as it is sometimes difficult to differentiate between morphologically similar species using microscopy alone (Madigan et al., 2012). While the presence of picocyanobacteria (cyanobacteria < 2 μm in diameter) could also explain some of the discrepancy, the relative abundance of *Synechococcus*, (a widespread genus of picocyanobacteria in freshwater) was low in all lakes based on the sedDNA (using the Greengenes and SILVA databases), indicating that this may have had a small impact.

Other potential explanations that may have contributed to the moderate correlations between the sediment and pelagic data can be attributed to bioturbation of sediments, the presence of grazers, or buoyant cells. Additionally, in the case of lakes 223, 224, and 442, there was generally only a small and gradual level of change over time. As such, if there is little to no change (or movement in a specific direction), a relationship is less likely to be found. However,

while results from this study show that relative abundance data should be interpreted with caution, the moderate to strong correlations found in the comparisons of abundance and those found using species richness data, reinforce the validity of this reconstruction approach. Overall, the findings presented here open doors for future studies on long-term cyanobacterial dynamics.

4.6 Conclusions

Both ddPCR and HTS were applied to reconstruct historical cyanobacterial trends and community compositions in response to eutrophication in L227 and acidification in L223 and results were compared to dynamics in reference lakes 224 and 442 which have had very little human impact. This study demonstrated how highly sensitive molecular tools can be used for the reconstruction of long-term bacterial dynamics and showed the effectiveness of ddPCR and HTS to reconstruct historical cyanobacterial trends. Here, the use of sedDNA provided new insight regarding long-term changes to cyanobacteria. For instance, there has been no microcystin monitoring at ELA, and the sedDNA results showed an increase in the microcystin E gene (and likely toxins) post fertilization. Changes in alpha diversity in response to environmental factors were also further clarified through the sediment record— an increase in cyanobacterial alpha diversity was found in the oligotrophic lakes (223, 224, and 442), but a decrease was observed in eutrophied L227 when only phosphorus loading continued. Differences were found between the eutrophied and oligotrophic lakes regarding the historical influences of environmental drivers on cyanobacterial community composition, suggesting that the long-term impacts of climate change, nutrient levels, and their co-influence can depend on trophic status. I compared the surface water records to the sedDNA archives and found a strong correlation between the two datasets for

absolute cyanobacterial abundance in the eutrophied lake and between the two datasets for species richness in lakes 223, 224, and 442. These findings can be applied in conservation and management efforts aiming to mitigate the negative impacts of harmful cyanobacterial blooms and suggest the imperative need to consider regional warming coupled with anthropogenic activities. To manage toxigenic cyanobacteria, management approaches using a model that limits phosphorus only should consider an approach that considers nitrogen and phosphorus stoichiometry.

Supplementary material is supplied in Appendix C.

Acknowledgements: Thank you to Scott Higgins and Cyndy Desjardins for providing the long-term water chemistry and phytoplankton records. Thank you to Michael Paterson, Peter Leavitt, and all IISD-ELA researchers for their continuous support. Thank you to Mary Ann Perron for the stimulating discussions that led to significant improvements of this chapter. This work was funded by a Natural Science and Engineering Research Council of Canada (NSERC) Discovery grant (RGPIN-2018-06356), the NSERC CREATE ABATE program, and the University of Ottawa, and was enabled in part by support provided by Compute Canada (www.computecanada.ca). This chapter is dedicated to Dr. David William Schindler who passed away in 2021. David Schindler was the founding director of the Experimental Lakes Area and a leader in conservation science whose innovative research has shaped our understanding of inland waters. He leaves a lasting scientific legacy in limnology.

Table 4.1 Kendall's (lakes 227 and 223) and Pearson's (lakes 224, and 442) correlation coefficients between log-transformed target gene abundances (copies/g) of cyanobacterial 16S rRNA (CYA), glutamine synthetase (*glnA*), *Microcystis* 16S rRNA (MICR), and microcystin E (*mcyE*) abundances (copies/g) as obtained from droplet digital PCR on sediment core subsections.

Lake	**Gene**	CYA	*glnA*	MICR	*mcyE*
227	CYA	1.00	0.74*	0.62*	0.69*
	glnA		1.00	0.59*	0.56*
	MICR			1.00	0.60*
	mcyE				1.00
223	CYA	1.00	0.57**	0.70*	0.32*
	glnA		1.00	0.44**	0.27*
	MICR			1.00	0.25
	mcyE				1.00
224	CYA	1.00	0.75*	0.92*	0.60*
	glnA		1.00	0.58*	0.40*
	MICR			1.00	0.55*
	mcyE				1.00
442	CYA	1.00	0.70*	0.67**	0.70**
	glnA		1.00	0.57**	0.44**
	MICR			1.00	0.52**
	mcyE				1.00

After correcting for temporal autocorrelation. *$p < 0.05$, **$p < 0.001$

Table 4.2 Alpha diversity metrics through the sediment cores of Experimental Lakes Area, lakes 227, 223, 224, and 442. NA indicates not applicable.

Lake	Lake manipulation	Sediment depth (cm)	Estimated timeframe	# of samples	Observed ASVs[a]	Shannon Diversity	Faith phylogenetic diversity
227	Phosphorus loading	0 – 10	~1990 – 2018	7	234.51	2.69	6.87
	Phosphorus and nitrogen loading	10.5 – 15.5	~1968 – 1989	4	299.65	4.24	9.09
	Pre-lake manipulation	16 – 52	Pre-1850 – 1967	12	91.66	3.18	6.17
223	Post-acidification	0 – 13.5	~2000 – 2018	13	251.21	5.31	8.07
	Acidification with H_2SO_4	19 – 20.5	~1976 – 1983	2	77.75	3.84	5.67
	Pre-acidification	21.5 – 42	Pre-1900 – 1973	9	68.06	3.36	5.21
224	*NA*	0 – 4	~2013 – 2018	5	411.36	6.16	12.33
	NA	5.5 – 16.5	~1982 – 2011	5	603.98	7.03	15.14
	NA	19 – 32.5	Pre-1850 – 1961	3	158.47	4.33	9.08
442	*NA*	0 – 10	~2001 – 2010	6	199.63	5.57	8.48
	NA	12 – 17	~1977 – 1994	6	183.10	5.35	9.7
	NA	32 – 40	Pre 1920 – 1931	2	10.55	2.26	4.67

[a]Amplicon Sequence Variant average number for individual timeframes.

Table 4.3 Pearson's correlation coefficients between the relative abundances of the most common orders based on the sediment DNA sequencing data and those based on the Experimental Lakes Area surface water records of phytoplankton communities from microscopy.

Lake	Nostocales	Chroococcales	Pseudanabaenales and Oscillatoriales
L227	0.88**	0.27	0.35
L223	-0.23	-0.23	0.15
L224	-0.028	-0.26	-0.43
L442	0.31	0.37	-0.72*

*$p = 0.044$
**$p < 0.001$

Table 4.4 The *p*-values (significance level $p < 0.05$) obtained using forward selection with 999 permutations based on the redundancy analysis between selected environmental variables and the square root transformed relative abundance of the predominant cyanobacterial orders in lakes 227 (n=10), 223 (n=8), 224 (n=8), and 442 (n=6) of the Experimental Lakes Area, Canada. *ns* indicates no significant effect ($p > 0.05$).

Lake	Approximate time range	Mean annual temperature	Total annual precipitation	Heating degree days	Total phosphorus	Total nitrogen
227	1969 – 2016	*ns*	*ns*	0.034	0.036	0.019
223	1976 – 2016	*ns*	*ns*	0.02	*ns*	*ns*
224	1977 – 2016	*ns*	*ns*	0.005	*ns*	*ns*
442	1998 – 2016	*ns*	0.041	0.004	*ns*	*ns*

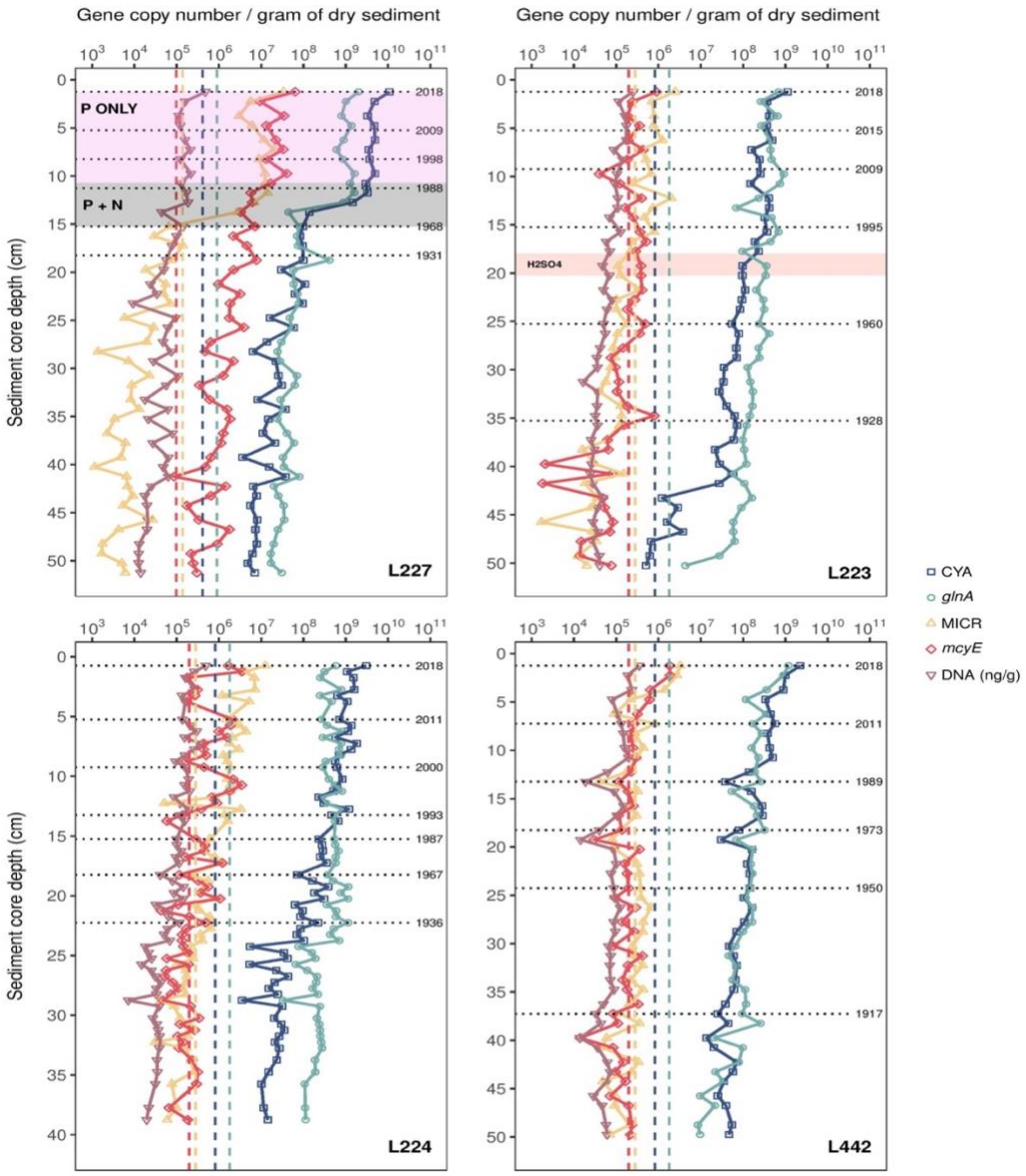

Figure 4.1 Normalized absolute gene copy numbers (copies/g) of cyanobacterial 16S rRNA (CYA), glutamine synthetase (*glnA*), *Microcystis* 16S rRNA (MICR), and microcystin E (*mcyE*), obtained from droplet digital PCR in sediment cores of the Experimental Lakes Area, Canada from Lake 227 (L227), Lake 223 (L223), Lake 224 (L224), and Lake 442 (L442). DNA extraction yields (ng/g) are also shown. Results are presented on a logarithmic scale. The dashed vertical lines correspond to the limit of detection of the gene with the same colour (blue: CYA, teal: *glnA*, yellow: MICR, and red: *mcyE*). The grey shading in the L227 plot (top left), labelled P + N, corresponds to the period of phosphorus and nitrogen loading in the lake between 1969 – 1989. The purple shading in the same plot, labelled P only, corresponds to the period of phosphorus loading only from 1990 – 2018. The pink shading in the L223 plot (top right), labelled H_2SO_4, corresponds to the period of sulfuric acid loading in the lake between 1976 – 1983.

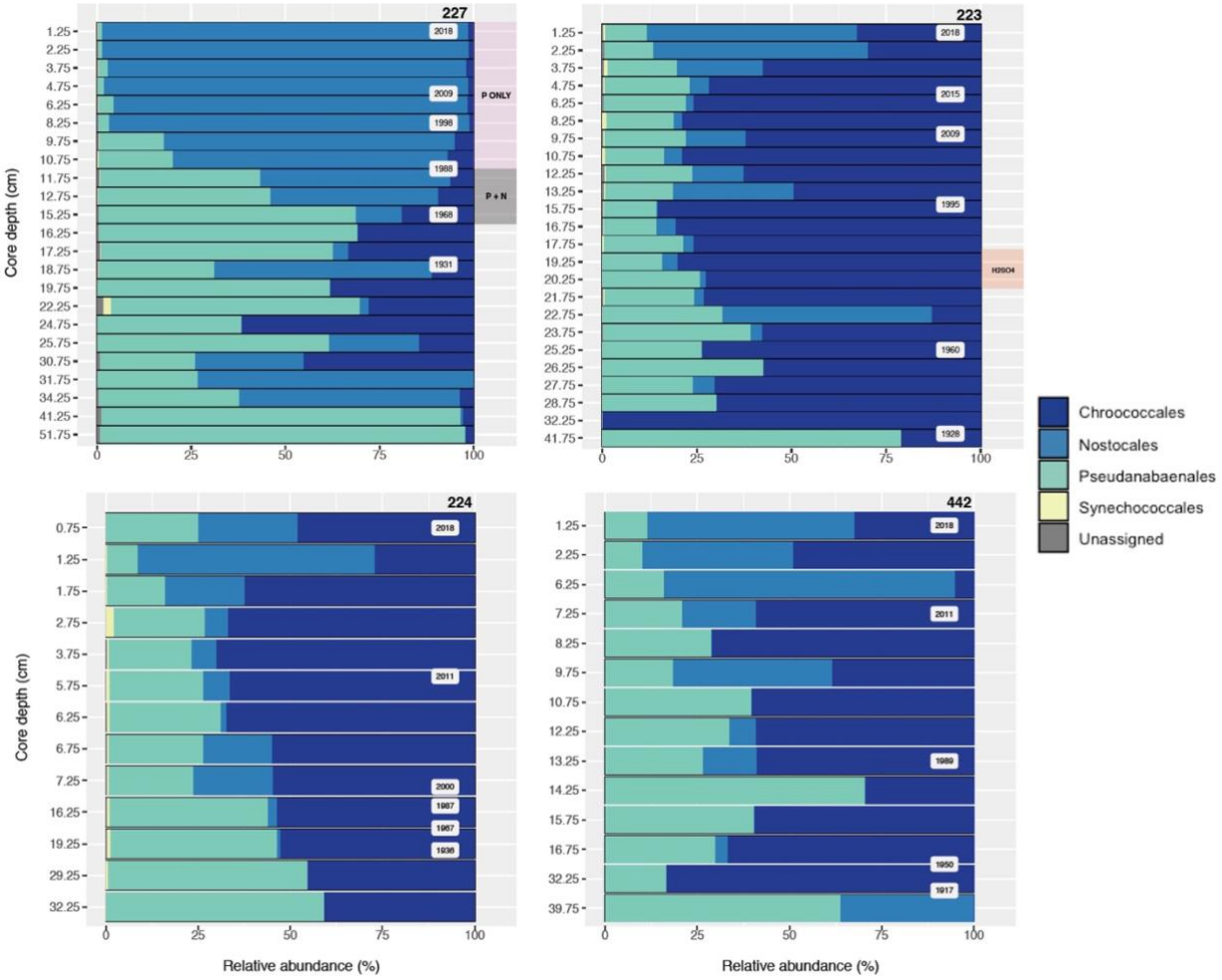

Figure 4.2 Cyanobacterial community composition in lake sediment cores of the Experimental Lakes Area, Canada, from manipulated lakes 227 (top left) and 223 (top right), along with reference lakes 224 (bottom left) and 442 (bottom right), as determined through high-throughput sequencing of the cyanobacterial 16S rRNA gene and using the Greengenes reference gene database. The relative abundance of each major taxa is presented at the level of order. The grey shading in the L227 plot (top left), labelled P + N, corresponds to the period of phosphorus and nitrogen loading in the lake between 1969 – 1989. The purple shading in the same plot, labelled P only, corresponds to the period of phosphorus loading only from 1990 – 2018. The pink shading in the L223 plot (top right), labelled H_2SO_4, corresponds to the period of sulfuric acid loading in the lake between 1976 – 1983.

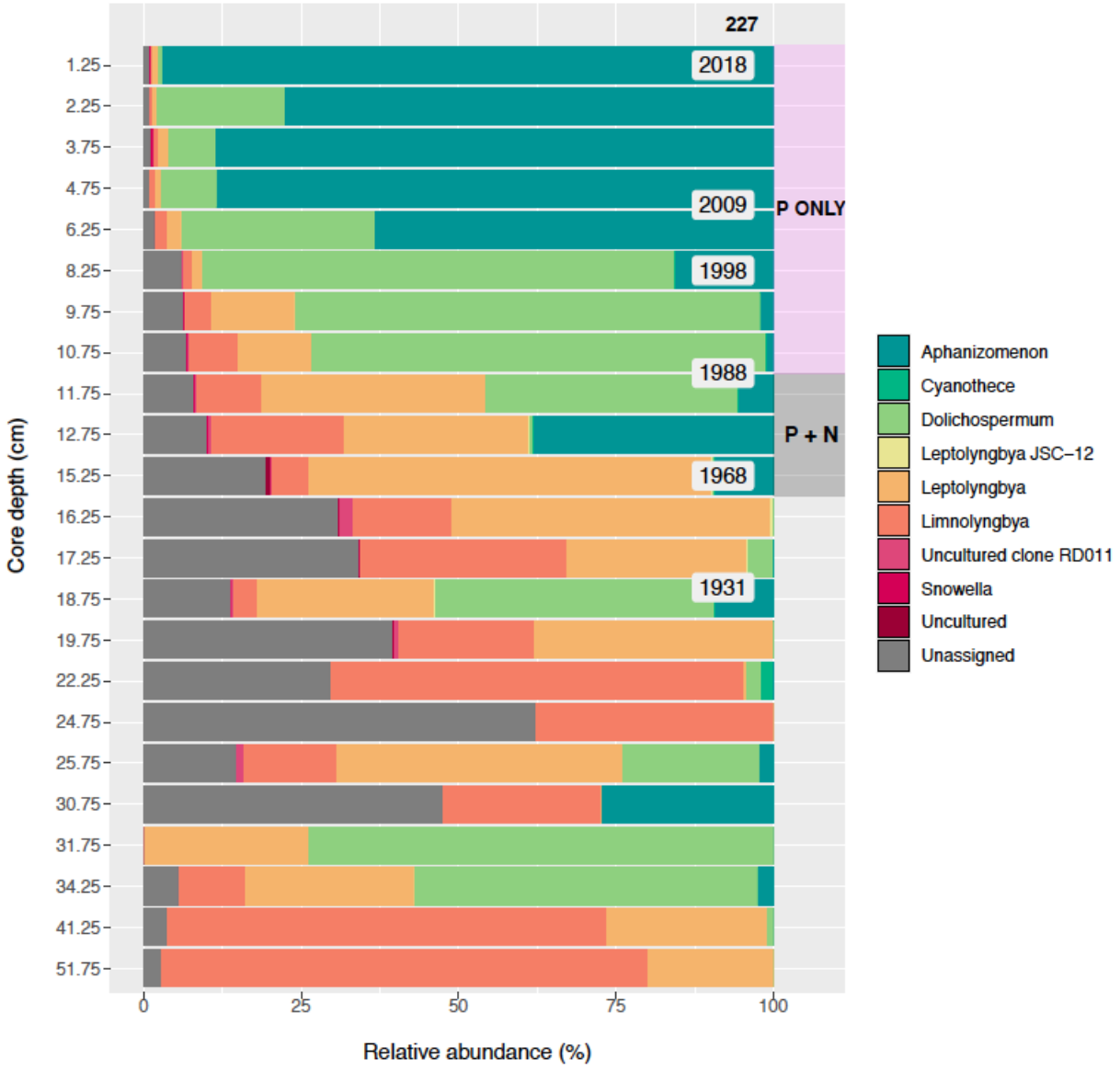

Figure 4.3 The cyanobacterial community composition in the sediment core of Lake 227 from the Experimental Lakes Area, Canada, as determined through high-throughput sequencing of the cyanobacterial 16S rRNA gene and using the SILVA reference gene database. The relative abundance of each major taxa is presented at the level of genus. The grey shading labelled P + N, corresponds to the period of phosphorus and nitrogen loading in Lake 227 between 1969 – 1989. The purple shading, labelled P only, corresponds to the period of phosphorus loading only from 1990 – 2018.

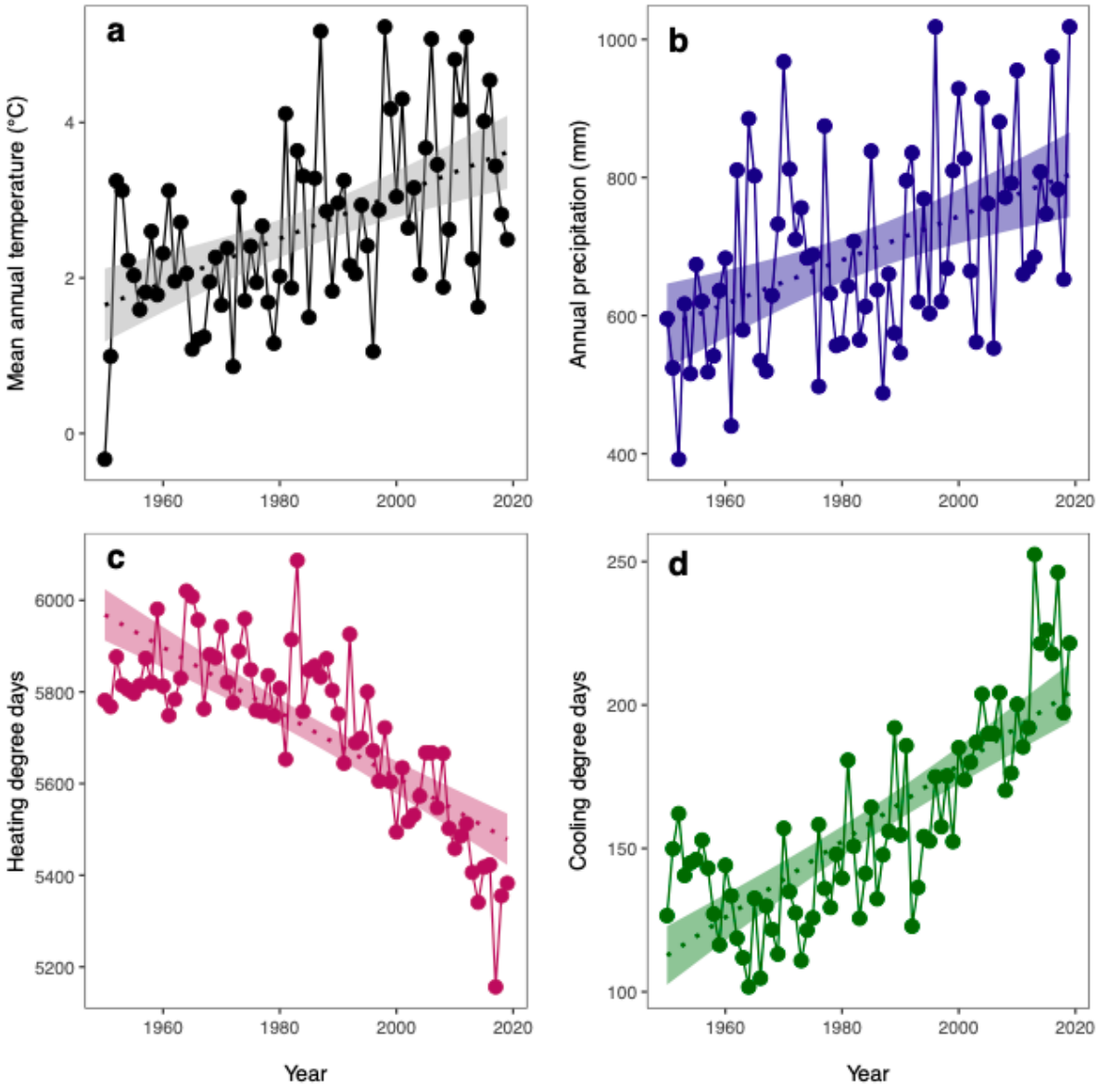

Figure 4.4 Changes to mean annual temperature (a), annual precipitation (b), heating degree days (c), and cooling degree days (d) at the Kenora meteorological station after 1950 using data obtained from the National Climate Data Archive of Environment and Climate Change Canada. The dotted line represents the linear regression, and the shading represents the confidence interval (95%).

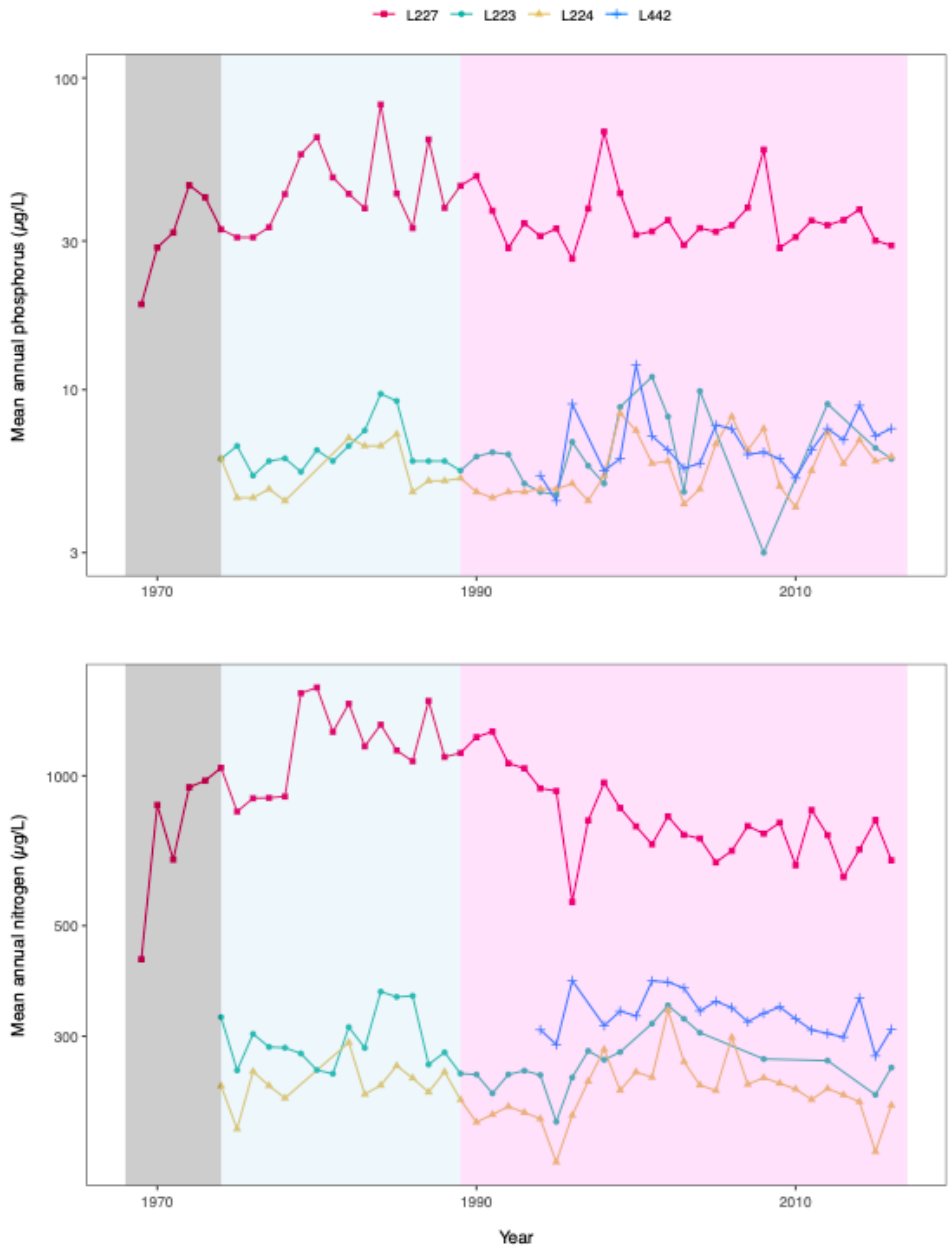

Figure 4.5 Mean annual eplimnetic nutrient concentrations in lakes 227 (L227), 223 (L223), 224 (L224), and 442 (L442) of the Experimental Lakes Area, Canada. The grey shading corresponds to the period of high nitrogen and phosphorus loading (~308 N and 24.8 P kg/year, 1969 – 1974) in L227. The blue shading corresponds to the period of fertilization with reduced nitrogen loading (~110 N and 22.3 P kg/year, 1975 – 1989) in L227, while the purple shading corresponds to the period where nitrogen was no longer added into L227 (~0 N and 26 P kg/year, 1989 – 2018). See Table S3.1 in Appendix B for further details. Data were obtained from Scott Higgins, IISD-ELA.

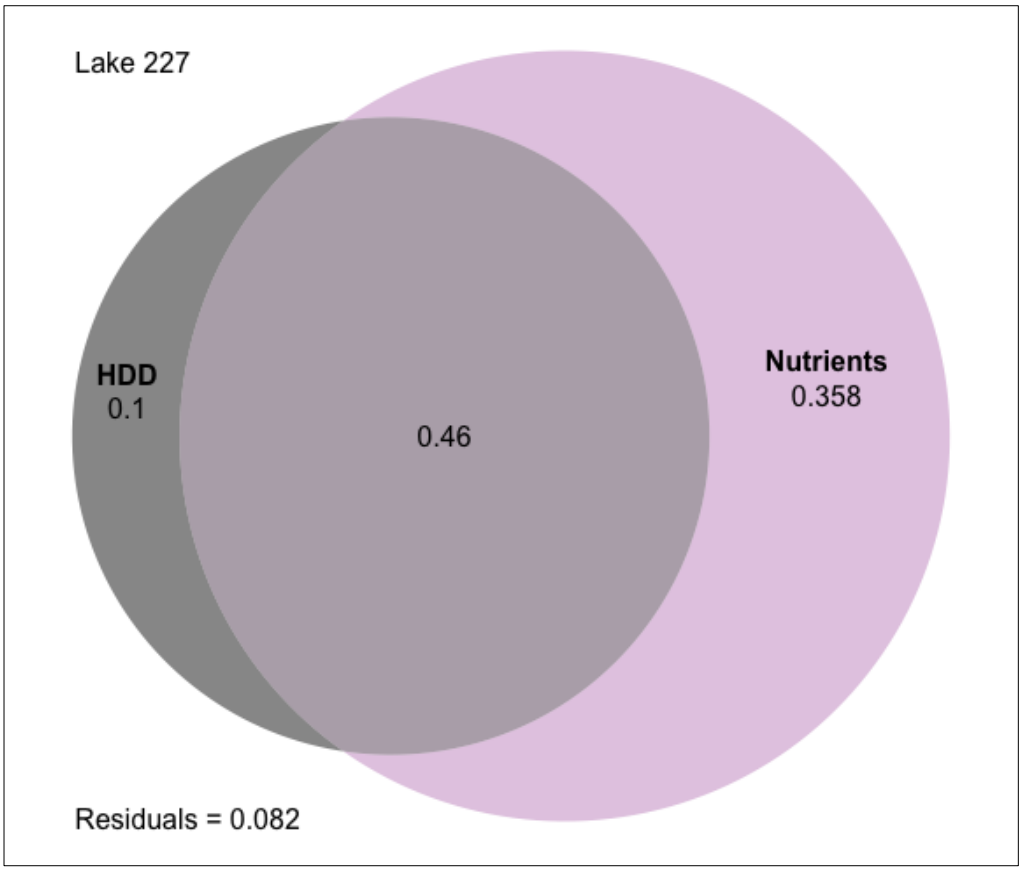

Figure 4.6 Variation partitioning showing the individual contributions of heating degree days (HDD) and nutrients, and their co-influence towards explaining the variation in the cyanobacterial community in Lake 227 (Experimental Lakes Area, Ontario, Canada). The residuals are elements of variation unexplained by these variables.

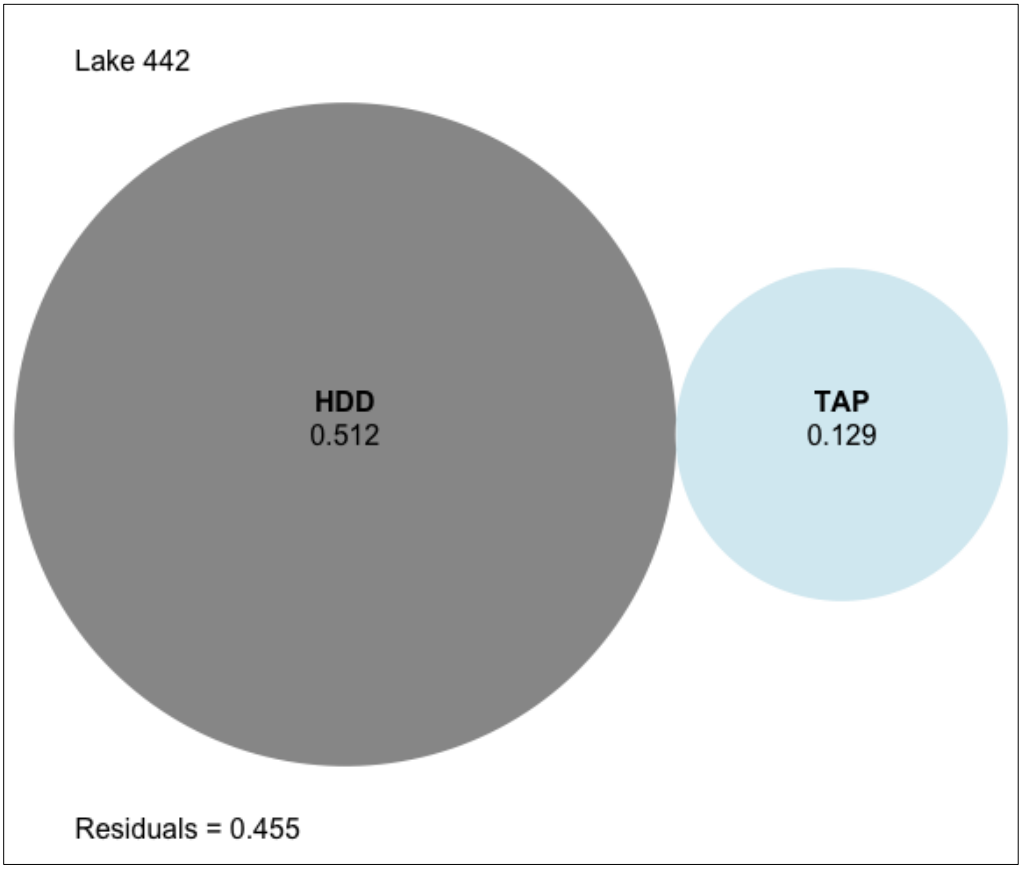

Figure 4.7 Variation partitioning showing the individual contributions of heating degree days (HDD) and total annual precipitation (TAP), and their co-influence towards explaining the variation in the cyanobacterial community in Lake 442 (Experimental Lakes Area, Ontario, Canada). The residuals are elements of variation unexplained by these variables.

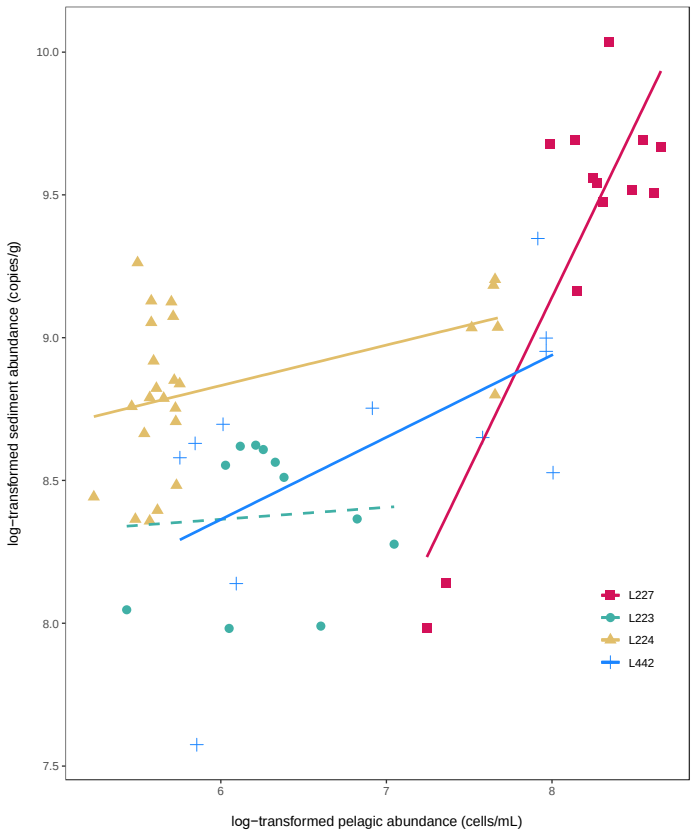

Figure 4.8 Correlations between log-transformed droplet digital PCR gene copy numbers of cyanobacterial 16S rRNA in sediment (copies/g) and log-transformed microscopic counts of cyanobacterial cells (cells/mL) in the water column of Experimental Lakes Area, lakes 227 (L227), 223 (L223), 224 (L224), and 442 (L442). Correlations were performed on a reduced dataset for which there was both sediment and water data available. For the water data, annual cyanobacterial abundance (cells/mL) was calculated using the sum of the monthly averages (May-September) of cyanobacterial cell density collected from the epilimnion and metalimnion layers. A solid line represents a significant correlation ($p < 0.05$) whereas a dashed line represents an insignificant correlation ($p > 0.05$).

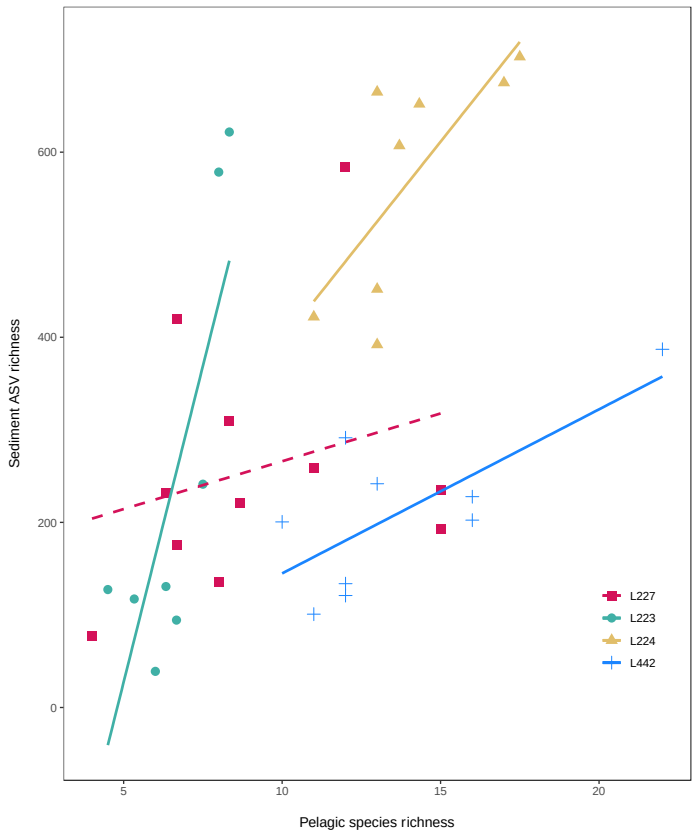

Figure 4.9 Correlations between annual ASV richness from high-throughput sequencing of the cyanobacterial 16S rRNA gene in sediments and cyanobacterial species richness in the water column of Experimental Lakes Area, lakes 227 (L227), 223 (L223), 224 (L224), and 442 (L442). Correlations were performed on a reduced dataset for which there was both sediment and water data available. For the water data, pelagic species richness was the measure of the number of different cyanobacterial species for that year. For the sediment data, sediment ASV richness was the measure of the number of unique ASVs found in that sediment sample. A solid line represents a significant correlation ($p < 0.05$) whereas a dashed line represents an insignificant correlation ($p > 0.05$).

CHAPTER 5: GENERAL CONCLUSIONS

5.1 Summary of research findings

Anthropogenic activities have produced changes on a global scale, leading to rising temperatures and excessive nutrient pollution (Galloway et al., 2008; IPCC, 2021). Climate change, too, has profoundly affected aquatic ecosystems. In reference to temperate lakes, water columns have become more stable during the ice-free season and mixing regimes have shifted (Barange and Perry, 2009; Vincent, 2009b). This has led to changes to the physical, biological, and chemical dynamics of freshwater and coastal areas that are reshaping natural communities, with conditions favouring cyanobacterial growth (Paerl and Huisman, 2009; O'Neil et al., 2012). Yet, much remains unknown about the combined effects of environmental drivers on cyanobacterial dynamics in lakes. It is imperative to disentangle the impact of these drivers to mitigate harmful cyanobacterial blooms. Paleolimnology has become an indispensable tool to reconstruct historical trends and is used to address questions regarding the impacts of environmental change over time and space. The general aim of this thesis was to evaluate how historical cyanobacterial dynamics have changed in temperate regions using sediment DNA (sedDNA) extracted from selected Experimental Lakes Area (ELA) lake sediment cores.

The application of molecular tools and the use of sedDNA have opened new areas of inquiry and provide for considerable expansion in the field of paleolimnology. However, there are concerns regarding DNA degradation and differences in preservation capacity between taxa, which could interfere with accurate interpretations of abundance and community data. To address some of these concerns, I performed a 1-year experiment that estimated and compared the degradation rates of selected cyanobacterial genes in sediments under contrasting environmental conditions (warm vs. cool and anoxic vs. oxic) (Chapter 2). Very few studies have assessed the factors impacting DNA decay rates in sediment. I found that cold, dark, anoxic

conditions slowed DNA decay the most and that *Synechococcus* sp. exhibited the lowest degradation rates followed by *Trichormus*, then *Microcystis*. These findings advanced the current understanding of cyanobacterial preservation in sediments and will be useful in the interpretation of population and community structure data from sedDNA. Results will also be useful in sediment sampling procedures, specifically in choosing sampling sites that are likely to best preserve sedDNA.

Next, I compared the quantitative performance between a targeted PCR approach and a metabarcoding approach for the analysis of sedDNA (Chapter 3). While the use of molecular tools in paleolimnology has been increasing over the past decade, it is not clear if the complexity of sediment samples (where there is potentially a high amount of PCR inhibitors) can impact the quantitative performance of these methods. The specific goal of Chapter 3 was to compare absolute and relative abundance data obtained from droplet digital PCR (ddPCR) with abundance data obtained from high-throughput sequencing (HTS). Using, sedDNA records of cyanobacteria, I found moderate to strong correlations between the two approaches (Mejbel et al., 2021b). However, based on previous studies evaluating phytoplankton and cyanobacterial dynamics at ELA, I found the ddPCR results to be generally more consistent with the known history of the study lakes (227, 223, 224, and 442). These results demonstrated that ddPCR quantification is a more appropriate choice for obtaining quantitative data from complex samples, possibly because the DNA sample is partitioned into thousands of droplets where each individual amplification event is separately analyzed. However, I highlight that if there is a *priori* knowledge that low levels of inhibitors are present in a sample, then either ddPCR or HTS are suitable for quantification of target genes in sediment. The insight from this chapter is important

for future paleolimnological DNA-based methods aiming to obtain reliable quantitative data from sediment samples.

Both ddPCR and HTS were used in Chapter 4 to reveal and interpret long-term cyanobacterial trends at ELA. There were two overarching goals in Chapter 4. First, I wanted to evaluate the long-term impacts of phosphorus and nitrogen loading, acidification, and climate change on cyanobacterial dynamics. Specifically, I analyzed how these drivers have affected the absolute abundance of selected cyanobacterial genes and the overall cyanobacterial community. Results showed that phosphorus and nitrogen loading in L227 led to a significant rise in the abundance of all genes (cyanobacterial 16S rRNA, *Microcystis* 16S rRNA, *mcyE*, and *glnA*), but abundances did not decrease when nitrogen loading stopped in 1989 or further increase with continuous loading of phosphorus after 1990. Interestingly, while microcystins have not been previously recorded in these lakes, the sedDNA archives identified both the presence of *mcyE* (a gene involved in microcystin production) and its increase after nitrogen loading stopped in 1989 in L227. Through sequencing, shifts towards Nostocales dominance were observed in the cyanobacterial community profiles. However, based on results found in Chapter 2, the switch in dominance in lakes 223, 224, and 442 may be the result of potential differences in preservation capacity between cyanobacterial taxa. In acidified L223, there was a small increase in cyanobacterial abundance during the acidification experiment, suggesting that slightly acidic environments (pH levels between 5 – 7) favour cyanobacterial growth. In the evaluation of the impacts of select climate- and nutrient-associated variables, I found that only heating degree days (rather than other metrics associated with climate warming) consistently influenced the cyanobacterial communities. In Ontario, the number of heating degree days have significantly declined and are expected to continue to decline (Bush and Lemmen, 2019). At ELA, the number

of heating degree days has decreased by ~500 days since 1950 (Figure 4.4, Chapter 4). Although heating degree days were important in the oligotrophic lakes (223, 224, and 442), nutrients had a far greater influence in explaining the variation in the cyanobacterial community of fertilized L227. Precipitation influenced the community composition of L442, but a greater sample size and an analysis of the watershed may be necessary to further understand the impacts of precipitation.

The second goal of Chapter 4 was to validate the use of sedDNA by comparing sediment data to surface water phytoplankton records. Here, I found moderate to strong correlations between the two archives for absolute abundance values and for species/ASV richness, substantiating the use of sedDNA in historical reconstructions of biological communities. Altogether, I showed that the investigation of cyanobacterial abundance and community composition using sedDNA, ddPCR, and HTS, can be proposed as an approach to track and determine the factors leading to cyanobacterial change.

5.2 Implications

5.2.1 The development of a standard analysis strategy and contributions to the field of paleolimnology

The approach described in these studies demonstrate the effectiveness of molecular tools for the reconstruction of historical microbial trends. Most traditional paleolimnological approaches depend on morphological remains (like cell walls and resting spores), and while not all taxa leave behind some form of hard fossil, they do leave DNA. When long term monitoring data are sparse and traditional methods are unable to fully reveal ecosystem changes, sedDNA can be used to address questions surrounding the evolution of ecosystems and the impacts of

human activities on aquatic ecosystems. Together, the use of ddPCR and HTS offer the potential to provide reliable, precise results even down to the species level and expedite the process for the analysis of complex sediment samples.

Concerns regarding DNA degradation issues with the use of sedDNA were tackled in Chapter 2. Previous investigations have described that cold, dark, anoxic conditions best preserve DNA (which results from Chapter 2 confirm), yet no comparative experimental study has been previously conducted. To my knowledge, the microcosm study performed in Chapter 2 was the first study to assess the impacts of oxygen and temperature on microbial sediment DNA and the first to compare degradation rates in sediments between taxa of different cell structure. Although there were limitations to the experimental design in terms of faithful reproduction of *in situ* conditions in the laboratory, the study presented new evidence on the variability in preservation capacity that exists between major cyanobacterial taxa. In addition, while there have been some successful studies that used sedDNA in shallow subtropical sites (Martínez De La Escalera et al., 2014; Zhang et al., 2021), I highlighted that cyanobacterial sedDNA data from shallow or warm lakes should be interpreted with caution, especially for the analysis of the *Microcystis* 16S rRNA gene. Understanding how cyanobacterial DNA is best preserved through sediments can be used to prevent issues with degradation biases in the development of future research studies.

By comparing the quantitative performance of common molecular tools (ddPCR: targeted approach, and HTS: community approach), as performed in Chapter 3, other concerns with sedDNA studies were addressed. This comparison attempted to address questions regarding the uncertainty of abundance data from HTS using complex samples. Although there are a few existing environmental DNA studies that have compared the quantitative performance between

targeted and metabarcoding approaches, comparing the quantitative outputs using absolute abundances from ddPCR and read counts from HTS had not been performed over timescales of decades using sedDNA. I provided recommendations about what method to choose for sedDNA research in the aim that a standard molecular paleolimnological analysis strategy can be developed (Mejbel et al., 2021b).

5.2.2 Contributions to our understanding of the role of climate and nutrients on cyanobacterial dynamics

There have been a multitude of research studies that have examined harmful bloom ecology, but many laboratory and field studies focus on a single environmental factor, or if assessing multiple stressors, do not consider long-term effects. Some studies have analyzed cyanobacterial dynamics through time in perialpine lakes, but while phosphorus discharge into the lakes was controlled, nitrogen was not, and it is uncertain what the nutrient inputs were over the studied time frame (Monchamp et al., 2016; Monchamp et al., 2018). In this study, by using sedDNA extracted from L227 core subsamples, I was able to further clarify 1) the interactions of climate change with dual-nutrient loading (1969-1989), 2) the effects of climate change with just phosphorus loading (1990 to present) and, 3) the effects of climate change alone with low nutrient levels and low human impact (reference lakes) on cyanobacterial abundance and community composition over long periods of time. I showed that, in regard to climate-associated variables, heating degree days, rather than temperature and precipitation, were important in influencing cyanobacterial communities. Although, greater sample sizes and analyzing the watershed area to lake volume, as well as the surrounding forestry, will be necessary to further elucidate the impacts of climate change. Altogether, these comparisons contributed to our

understanding of how climate change and nutrient fertilization impacts long-term cyanobacterial trends. This is especially important as there remains some dispute on whether controlling for nitrogen is substantiated in strategies aiming to prevent cyanobacterial bloom formation.

Schindler et al. (2008) found that despite extreme nitrogen limitations, the presence of nitrogen-fixing cyanobacteria allowed growth to continue and that controlling for nitrogen did not have an impact on the magnitude of cyanobacterial blooms. However, results from Chapter 4 showed that, while indeed the cyanobacteria abundance did not decrease when nitrogen loading stopped, it led to the dominance of the Nostocales order, a decrease in alpha diversity, and an increase in the level of *mcyE*, indicating that as nitrogen levels reduce, toxigenic nitrogen-fixing cyanobacteria may increase. Additionally, the ddPCR results showed that both phosphorus and nitrogen may have stimulated cyanobacterial growth more than just phosphorus alone. However, it is possible that growth plateaued at the end of dual-nutrient loading, then was simply maintained with phosphorus loading alone. Altogether, these results suggest that reducing nitrogen levels could competitively favor Nostocales spread, including those that form toxins and that an alternative approach of controlling these species is required. These results highlight the need to consider controlling the ratio of nitrogen to phosphorus in the development of mitigation strategies aiming to reduce harmful cyanobacterial blooms. However, this is a difficult task in itself.

5.2.3. Significance of research

The investigation of cyanobacterial genes in sediments both quantitatively and qualitatively could elucidate more effective management techniques in lakes affected by cyanobacterial blooms by defining reference conditions and restoration targets. The results

presented here can be used to analyze other freshwater ecosystems impacted by climate change and anthropogenic activities. Likewise, the results from L223 further clarified the impacts of acidification, a potential consequence of increased industrialization. The findings deepen the current understanding of the ecology of cyanobacteria and can be used to limit the consequences associated with cyanobacterial bloom formation. The work presented in this study can also be extended to other freshwater organisms and environments to help decision makers adopt the most successful and appropriate strategies, especially as it is also of relevance to ecosystem services management.

A key component for improving management efforts that create meaningful outcome is the communication and sharing of evidence. To accomplish this, and in pursuant of the scientific process, I presented this work at many conferences and workshops, through educational organizations and outreach initiatives (enabling elementary and high school students to learn more), wrote articles targeting lake associations and residents so that they can make informed decisions, and practiced in collaborative research by establishing partnerships across different sectors. Overall, as harmful cyanobacterial blooms continue to be of global concern, sharing this information will be important in improving scientific and monitoring activities and increasing observational efforts.

5.3 Moving forward

5.3.1 Immediate next steps

The techniques and work presented here provide information beyond that of surface water records but improving the reliability of results should be considered in future projects. Sediment samples are complex and PCR inhibition caused by the co-extraction of potential inhibitors (such as humic and fulvic acids) needs to be limited to prevent false negative results.

Modifications were performed to the DNA extraction protocol in these studies, but in order to improve DNA yield and quality, further improvements are needed and/or comparisons between DNA extraction kits should be performed to advance the reliability and use of molecular proxies. Additionally, the gene targets chosen for analysis had relatively short fragment sizes (< 300 bp for ddPCR but 450 bp for HTS), but there was some variability between the genes which may have interfered with the quantification. Target genes that have similar fragment sizes should be compared.

Molecular approaches do not only describe community composition, but also provide information about functional ecology, which the microscope record cannot provide. Here, only the cyanobacterial 16S rRNA gene was sequenced, but to provide a clearer interpretation of the microcystin record and to determine if microcystin concentrations are indeed increasing, it will be necessary to sequence other taxa-specific genes and the microcystin-synthesizing gene cluster. In addition, the presence of other toxins that Nostocales may produce (for e.g., anatoxins or saxitoxins) should also be evaluated. Conducting this using sediment subsections from L227 would be particularly important as the community composition is presently primarily composed of species from the Nostocales order.

5.3.2 Future research

The interactions of other abiotic factors should be considered to fully appreciate how these factors synergistically impact long-term cyanobacterial trends. In the studies performed for this thesis, the effects of nutrient loading, acidification, precipitation, temperature, and degree days were evaluated, but other factors that should be considered include partial pressure of carbon dioxide in lake water, light availability, changes in wind speed and stratification patterns,

storm events, and the intensity of heat waves. These interactions should be analyzed over large environmental gradients but also at regional scales where unique weather patterns develop. The approaches presented here should also be extended to lakes from other climate zones as most cyanobacterial sedDNA research has been performed on temperate lakes. More sedDNA analyses should be performed using samples collected from northern regions to further disentangle the relative impacts of climate and environmental change on cyanobacteria population and community composition dynamics.

Ultimately, sediment DNA can be used to address questions surrounding the evolution of ecosystems and the impacts of human activities on lakes, oceans, and coastal regions. As PCR and sequencing costs continue to decline and as more studies validate these approaches, the use of sedDNA to reconstruct historical dynamics will continue to increase. By analyzing how cyanobacterial dynamics have changed in response to environmental variables over decadal, centennial, and even millennial timescales, including those driven by anthropogenic activities, the association between environmental conditions and harmful cyanobacterial bloom formation can be resolved for regional and global spatial scales. This information will be critical for the management of surface waters and for the protection of ecosystem health.

REFERENCES

Adams, D.G., and Carr, N.G. (1981). The developmental biology of heterocyst and akinetes formation in cyanobacteria. *Crit. Rev. Microbiol.* 9(1), 45-100.

Adams, D. G., and Duggan, P. S. (1999). Heterocyst and akinete differentiation in cyanobacteria. *New Phytol.* 144(1), 3–33. doi: 10.1046/j.1469-8137.1999.00505.x

Albers, C.N., Jensen, A., Bælum, J., and Jacobsen, C.S. (2013). Inhibition of DNA polymerases used in Q-PCR by structurally different soil-derived humic substances. *Geomicrobiol. J.* 30, 675–681. doi: 10.1080/01490451.2012.758193

Allan, E.A., Zhang, W. G., Lavery, A., and Govindarajan, A. (2021). Environmental DNA shedding and decay rates from diverse animal forms and thermal regimes. *eDNA* 3(2), 492–514. doi: 10.1002/edn3.141

Anderson, R.F., Schiff, S.L. and Hesslein, R.H. (1987). Determining sediment accumulation and mixing rates using ^{210}Pb, ^{137}Cs and other tracers: problems due to post-depositional mobility or coring artifacts. *Can. J. Fish. Aquat. Sci.* 44, 231–250. doi:10.1139/f87-298

Antony-Babu, S., Murat, C., Deveau, A., Le Tacon, F., Frey-Klett, P., and Uroz, S. (2013). An improved method compatible with metagenomic analyses to extract genomic DNA from soils in *Tuber melanosporum* orchards. *J. Appl. Microbiol.* 115, 163–170. doi: 10.1111/jam.12205

Armbruster, D.A., and Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. *Clin. Biochem. Rev.* 29(Suppl 1), S49–52.

Armstrong, F.A.J., and Schindler, D.W. (1971). Preliminary chemical characterization of waters in the Experimental Lakes Area, northwestern Ontario. *J. Fish. Res. Board Can.* 28, 171–187.

Arndt, S., Jørgensen, B.B., LaRowe, D.E., Middleburg, J.J., Pancost, R.D., and Regnier, P. (2013). Quantifiying the degradation of organic matter in marine sediments: A review and synthesis. *Earth Sci. Rev.* 123, 53–86. doi: 10.1016/j.earscirev.2013.02.008

Barange, M., and Perry, R.I. (2009). Physical and ecological impacts of climate change relevant to marine and inland capture fisheries and aquaculture. Climate change implications for fisheries and aquaculture: overview of current scientific knowledge. *FAO Fish. Tech.* No. 530. Rome, FAO. 7-106.

Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L., and Lodge, D.M. (2014). Environmental conditions influence eDNA persistence in aquatic systems. *Environ. Sci. Technol.* 48(3), 1819–1827. doi: 10.1021/es404734p

Beaulieu, M., Pick, F.R., and Gregory-Eaves, I. (2013). Nutrients and water temperature are

significant predictors of cyanobacterial biomass in a 1147 lakes data set. *Limnol. Oceanogr*. 58, 1736–1746. doi: 10.4319/lo.2013.58.5.1736

Becker, R.H., Sultan, M.I., Boyer, G.L., Twiss, M.R., and Konopko, E. (2009). Mapping cyanobacterial blooms in the Great Lakes using MODIS. *J. Great Lakes Res*. 35, 447–453. doi: 10.1016/j.jglr.2009.05.007

Bingham, M., Sinha, S.K., and Lupi, F. (2015). Economic benfits of reducing harmful algal blooms in Lake Erie (Report No 67). Environmental Consulting & Technology Inc.

Blais, J.M., Kalff, J., Cornett, J.R., and Evans, D.R. (1995). Evaluation of ^{210}Pb dating in lake sediments using stable Pb, *Ambrosia* pollen, and ^{137}Cs. *J. Paleolimnol.* 13, 169–178. doi: 10.1007/BF00678105

Blum, S.A.E., Lorenz, M.G., and Wackernagel, W. (1997). Mechanism of retarded DNA degradation and prokaryotic origin of DNases in nonsterile soils. *Syst. Appl. Microbiol.* 20(4), 513–521. doi: 10.1016/S0723-2020(97)80021-5

Boere, A.C., Rijpstra, W.I.C., De Lange, G.J., Sinninghe Damsté, J.S., and Coolen, M.J.L. (2011). Preservation potential of ancient plankton DNA in Pleistocene marine sediments. *Geobiology* 9(5), 377–393. doi: 10.1111/j.1472-4669.2011.00290.x

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. doi: 10.1038/s41587-019-0209-9

Bonilla, S., and Pick, F.R. (2017). Freshwater bloom-forming cyanobacteria and anthropogenic change. *Limnol. Oceanogr. e-Lectures* 7(2), 1–62. doi: 10.1002/loe2.10006

Bradbury, J.P., Colman, S.M., and Reynolds, R.L. (2004). The history of recent limnological changes and human impact on Upper Klamath Lake, Oregon. *J. Paleolimnol.* 31, 151–165.

Brazeau, M.L., Blais, J.M., Paterson, A.M., Keller, W., and Poulain, A.J. (2013). Evidence for microbially mediated production of elemental mercury (Hg^0) in subarctic lake sediments. *J. Appl. Geochemistry* 37, 142–148. doi: 10.1016/j.apgeochem.2013.07.020

Brock TD. (1973). Lower pH limit for the existence of blue-green algae: evolutionary and ecological implications. *Science* 197: 480-483.

Bush, E. and Lemmen, D.S., editors (2019). Canada's Changing Climate Report; Government of Canada, Ottawa, ON.

Buxton, A.S., Groombridge, J.J., and Griffiths, R.A. (2017). Is the detection of aquatic environmental DNA influenced by substrate type? *PLoS ONE* 12(8): e0183371. doi: 10.1371/journal.pone.0183371

Bylemans, J., Furlan, E., Gleeson, D.M., and Hardy, C.M. (2018). Does Size Matter? An Experimental evaluation of the relative abundance and decay rates of aquatic eDNA. *Environ. Sci.Technol.* 52(11), 6408- 6416.

Bylemans, J., Gleeson, D.M., Duncan, R.P., Hardy, C.M., and Furlan, E.M. (2019). A performance evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes. *eDNA* 1, 402–414. doi: 1002.edn3.41

Cai, P., Huang, Q., Zhu, J., Jiang, D., Zhou, X., Rong, X., et al. (2007). Effects of low-molecular-weight organic ligands and phosphate on DNA adsorption by soil colloids and minerals. *Colloids Surf. B. Biointerfaces* 54(1), 53–59. doi: 10.1016/j.colsurfb.2006.07.013

Calace, N., Cardellicchio, N., Petronio, B.M., Pietrantonio, M., and Pietroletti, M. (2006). Sedimentary humic substances in the northern Adriatic sea (Mediterranean sea). *Mar. Environ. Res.* 61, 40–58. doi: 10.1016/j.marenvres.2005.05.002

Callieri, C., Cronberg, G., and Stockner, J.G. (2012). *Freshwater Picocyanobacteria: Single Cells, Microcolonies and Colonial Forms* in *Ecology of Cyanobacteria II: Their Diversity in Time and Space, 2nd Edn, ed.* Whitton B.A. Berlin: Springer. 229–269. doi: 10.1007/978-94-007-3855-3_8

Cao, X., Xu, X., Bian, R., Wang, Y., Yu, H., Xu, Y., et al. (2020). Sedimentary ancient DNA metabarcoding delineates the contrastingly temporal change of lake cyanobacterial communities. *Water res.* 183: 116077. doi: 10.1016/j.watres.2020.116077.

Capo, E., Debroas, D., Arnaud, F., and Domaizon, I. (2015). Is planktonic diversity well recorded in sedimentary DNA? Toward the reconstruction of past protistan diversity. *Microb. Ecol.* 70(4), 865–875. doi: 10.1007/s00248-015-0627-2

Capo, E., Giguet-Covex, C., Rouillard, A., Nota, K., Heintzman, P. D., Vuillemin, A., et al. (2021). Lake sedimentary dna research on past terrestrial and aquatic biodiversity: Overview and recommendations. *Quat.* 4(1):6. doi: 10.3390/quat4010006

Carey, C.C., Ibelings, B.W., Hoffmann, E.P., Hamilton, D.P., and Brookes, J.D. (2012). Eco-physiological adaptations that favour freshwater cyanobacteria in a changing climate. *Water Res.* 46, 1394–1407.

Carpenter, S.R. (2005). Eutrophication of aquatic ecosystems: Biostability and soil phosphorus. *Proc. Natl. Acad. Sci. USA* 102(29), 10002–10005.

Caruso, V., Song, X., Asquith, M., and Karstens, L. (2019). Performance of microbiome sequence inference methods in environments with varying biomass. *mSystems* 4:e00163–18. doi: 10.1128/mSystems.00163-18

Coolen, M.J.L., Boere, A., Abbas, B., Baas, M., Wakeham, S.G., and Sinninghe Damsté, J.S. (2006). Ancient DNA derived from alkenone-biosynthesizing haptophytes and other algae

in Holocene sediments from the Black Sea. *Paleoceanography* 21:PA1005. doi: 10.1029/2005PA001188

Coolen, M.J.L., Muyzer, G., Irene, W., Rijpstra, C., Schouten, S., Volkman, J. K., et al. (2004). Combined DNA and lipid analyses of sediments reveal changes in Holocene haptophyte and diatom populations in an Antarctic lake. *EPSL* 223(1-2), 225- 239. doi: 10.1016/j.epsl.2004.04.014

Coolen, M.J.L., and Overmann, J. (1998). Analysis of subfossil molecular remains of purple sulfur bacteria in a lake sediment. *Appl. Environ. Microbiol.* 64(11), 4513–4521. doi: 10.1128/aem.64.11.4513-4521.1998

Coolen, M.J.L., and Overmann, J. (2007). 217 000-year-old DNA sequences of green sulfur bacteria in Mediterranean sapropels and their implications for the reconstruction of the paleoenvironment. *Environ. Microbiol.* 9(1), 238–249. doi: 10.1111/j.1462-2920.2006.01134.x

Corinaldesi, C., Barucca, M., Luna, G. M., and Dell'Anno, A. (2011). Preservation, origin and genetic imprint of extracellular DNA in permanently anoxic deep-sea sediments. *Mol. Ecol.* 20(3), 642–654. doi: 10.1111/j.1365-294X.2010.04958.x

Corinaldesi, C., Beolchini, F., and Dell'Anno, A. (2008). Damage and degradation rates of extracellular DNA in marine sediments: Implications for the preservation of gene sequences. *Mol. Ecol.* 17(17), 3939–3951. doi: 10.1111/j.1365-294X.2008.03880.x

Cruikshank, D.R. (1984). W*hole lake chemical additions* in the Experimental Lakes Area, *1969–1983*. Canadian Data Report of Fisheries and Aquatic Sciences No. 449.

Crusius, J., and Anderson, R.F. (1995). Evaluating the mobility of ^{137}Cs, $^{239+240}Pu$ and ^{210}Pb from their distribution in laminated lake sediments. *J. Paleolimnol.* 13, 119–141. doi: 10.10007/BF000678102

Dell'Anno, A., and Danovaro, R. (2005). Ecology: Extracellular DNA plays a key role in deep-sea ecosystem functioning. *Science* 309(5744), 2179. doi: 10.1126/science.1117475

Deng, W., Monks, L., and Neuer, S. (2015). Effects of clay minerals on the aggregation and subsequent settling of marine Synechococcus. *Limnol. Oceanogr.* 60(3), 805–816. doi: 10.1002/lno.10059

Deshpande, B.N., Tremblay, R., Pienitz, R., and Vincent, W.F. (2014). Sedimentary pigments as indicators of cyanobacterial dynamics in a hypereutrophic lake. *J. paleolimnol.* 52, 171–184.

Doblin, M.A., Coyne. K.J., Rinta-Kanto, J.M., Wilhem, S.W., and Dobbs, F.C. (2007). Dynamics and short-term survival of toxic cyanobacteria species in ballast water from NOBOB vessels transiting the Great Lakes implications for HAB invasions. *Harmful Algae*

6(4), 519–530.

Dodsworth, W. (2020). Temporal trends in cyanobacteria through paleo-genetic analyses. Master's thesis, University of Ottawa. doi: 10.20381/ruor-24401

Dolman, A.M., Rüker, J., Pick, F.R., Fastner, J., Rohrlack, T., Mischke, U., et al. (2012). Cyanobacteria and cyanotoxins: The influence of nitrogen versus phosphorus. *PLoS ONE* 7(6): e38757.

Domaizon, I., Savichtcheva, O., Debroas, D., Arnaud, F., Villar, C., Pignol, C., et al. (2013). DNA from lake sediments reveals the long-term dynamics and diversity of *Synechococcus* assemblages. *Biogeosciences* 10(6), 3817–3838. doi: 10.5194/bg-10-3817-2013

Domaizon, I., Winegardner, A., Capo, E., Gauthier, J., and Gregory-Eaves, I. (2017). DNA-based methods in paleolimnology: new opportunities for investigating long-term dynamics of lacustrine biodiversity. *J. Paleolimnol.* 58, 1–21. doi: 10.1007/s10933-017-9958-y

Dong, L., Wang, S., Fu, B., and Wang, J. (2018). Evaluation of droplet digital PCR and next generation sequencing for characterizing DNA reference material for KRAS mutation detection. *Sci. Rep.* 8(1), 9650. doi: 10.1038/s41598-018-27368-3

Downing, J.A., Watson, S.B., and McCauley, E. (2001). Predicting Cyanobacteria dominance in lakes. *Can. J. Fish. Aquat. Sci.* 58, 1905–1908. doi: 10.1139/f01-143

Eichmiller, J.J., Best, S.E., and Sorensen, P.W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environ. Sci. Technol.* 50(4), 1859-1867. doi: 10.1021/acs.est.5b05672

Ellegaard, M., and Ribeiro, S. (2018). The long-term persistence of phytoplankton resting stages in aquatic 'seed banks'. *Biol. Rev.* 93(1), 166–183. doi: 10.1111/brv.12338

Emerson, S., and Hesslein, R. (1973). Distribution and uptake of artificially introduced radium-226 in a small lake. *J. Fish. Res. Board Canada* 30, 1485–1490. doi: 10.1139/f73-238

Falkowski, P., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., et al. (2000). The global carbon cycle: a test of our knowledge of earth as a system. *Science* 290(5490), 291–296. doi: 10.1126/science.290.5490.291

Fay, P. (1992). Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol. Rev.* 56(2), 340–373. doi: 10.1128/mr.56.2.340-373.1992

Fernandez-Carazo, R., Verleyen, E., Hodgson, D. A., Roberts, S. J., Waleron, K., Vyverman, et al. (2013). Late Holocene changes in cyanobacterial community structure in maritime Antarctic lakes. *J. Paleolimnol.* 50, 15– 31. doi: 10.1007/s10933-013-9700-3

Findlay, D.L., Hecky, R.E., Hendzel, L.L., Stainton, M.P., and Regehr, G.W. (1994). Relationship between N_2-fixation and heterocyst abundance and its relevance to the nitrogen budget of Lake 227. *Can. J. Fish Aquat. Sci.* 51, 2254–2266.

Findlay, D.L., and Kasian, S.E.M. (1986). Phytoplankton community responses to acidification of lake 223, experimental lakes area, northwestern Ontario. *Water Air Soil Pollut.* 30, 719–726. doi: 10.1007/BF00303337

Findlay, D.L., and Kasian, S.E.M. (1990). Phytoplankton communities of lakes experimentally acidified with sulfuric and nitric acids. *Can. J. Fish. Aquat. Sci.* 47, 1378–1386. doi: 10.1139/f90-157

Findlay, D.L., and Kasian, S.E.M. (1991). Response of a phytoplankton community to controlled partial recovery from experimental acidification. *Can. J. Fish. Aquat. Sci.* 48, 1022–1029.

Findlay, D.L., and Kasian, S.E.M. (1996). The effect of incremental pH recovery on the Lake 223 phytoplankton community. *Can. J. Fish. Aquat. Sci.* 53, 856–864. doi: 10.1139/cjfas-53-4-856

Findlay, D.L., Kasian, S.E.M. Turner, M.T., and Stainton, M.P. (1999). Responses of phytoplankton and epilithon during acidification and early recovery of a lake. *Freshw. Biol.* 42, 159 – 175.

Findlay, D.L., and Shearer, J.A. (1992). Relationships between sedimentary diatom assemblages and lakewater pH values in the experimental lakes area. *J. Paleolimnol.* 7, 145–156.

Gallon, J. R. (1992). Reconciling the incompatible: N2 fixation And O2. *New Phytol.* 122(4), 571–609. doi: 10.1111/j.1469-8137.1992.tb00087.x

Galloway, J.N., Townsend, A.R., Erisman, J.W., Bekunda, M., Cai, Z., Freney, J.R., et al. (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science*, 320, 889–892. doi: 10.1126/science.1136674

Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P. J., Griggo, C., Gielly, L., et al. (2014). Long livestock farming history and human landscape shaping revealed by lake sediment DNA. *Nat. Commun.* 5(1), 1–7. doi: 10.1038/ncomms4211

Glass, D.C., and Gray, C.N. (2001). Estimating mean exposure from censored data: exposure to benzene in the Australian petroleum industry. *Ann. Occup. Hyg.* 45, 275–282. doi: 10.1016/S0003-4878(01)00022-9

Guiry, M.D. and Guiry, G.M. (2021). *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org.

Hallegraeff, G.M., Anderson, D.M., Belin, C., Bottein, M-Y, D., Bresnan, E., Chinain, M., et al. (2021). Perceived global increases in algal blooms is attributable to intensified monitoring

and emerging bloom impacts. *Earth Environ. Commun.* 2:117. doi: 10.1038/s43247-021-00178-8

Hamaguchi, M., Shimabukuro, H., Hori, M., Yoshida, G., Terada, T., and Miyajima, T. (2018). Quantitative real-time polymerase chain reaction (PCR) and droplet digital PCR duplex assays for detecting *Zostera marina* DNA in coastal sediments. *Limnol. Oceanogr. Methods* 16, 253–264. doi: 10.1002/lom3.10242

Hänfling, B., Handley, L.L., Read, D.S., Hahn, C., Li, J., Nichols, P. et al. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Mol. Ecol.* 25, 3101–3119. doi: 10.1111/mec.13660

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

Harper, L.R., Handley, L.L., Hahn, C., Boonham, N., Rees, H.S., Gough, K.C., et al. (2018). Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). *Ecol. Evol.* 8, 6330–6341. doi: 10.1002/ece3.4013

Hendzel, L.L., Hecky, R.E., and Findlay, D.L. (1994). Recent change of N_2-fixation in Lake 227 in response to reduction of the N:P loading ratio. *Can. J. Fish Aquat. Sci.* 51, 2247–2253.

Henao, E., Rzymski, P., and Waters, M.N. (2020). A review on the study of cyanotoxins in paleolimnological research: Current knowledge and future needs. *Toxins* 12(1):6. doi: 10.3390/toxins12010006

Hesslein, R.H., Broecker, W.S., Quay, P.D., and Schindler, D.W. (1980a). Whole-lake radioarbon experiment in an oligotrophic lake at the Experimental Lakes Area, northwestern Ontario. *Can. J. Fish. Aquat. Sci.* 37, 454–463. doi: 10.1139/f80-059

Hesslein, R.H., Broecker, W.S., and Schindler, D.W. (1980b). Fates of metal radiotracers added to a whole lake: sediment-water interactions. *Can. J. Fish. Aquat. Sci.* 37, 378–386. doi: 10.1139/f80-052

Higgins, S.N., Paterson, M.J., Hecky, R.E., Schindler, D.W., Venkiteswaran, J.J., and Findlay, D.L. (2018). Biological nitrogen fixation prevents the response of a eutrophic lake to reduced loading of nitrogen: Evidence from a 46-year whole-lake experiment. *Ecosystems* 21, 1088–1100. doi: 10.1007/s10021-017-0204-2

Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., et al. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* 83, 8604–8610. doi: 10.1021/ac202028g

Hoiczyk, E., and Hansel, A. (2000). Cyanobacterial cell walls: News from an unusual

prokaryotic envelope. *J. Bacteriol.* 182(5), 1191–1199. doi: 10.1128/JB.182.5.1191-1199.2000

Hong, S., Bunge, J., Leslin, C., Jeon, S., and Epstein, S.S. (2009). Polymerase chain reaction primers miss half of rRNA microbial diversity. *ISME J.* 3, 1365–1373. doi: 10.1038/ismej.2009.89

Hou, D., He, J., Lü, C., Wang, W., and Zhang, F. (2014). Spatial distributions of humic substances and evaluation of sediment organic index on Lake Dalinouer, China. *J. Geochem.* 2014, 1–13. doi: 10.1155/2014/502597

Huber, J.A., Morrison, H.G., Huse, S.M., Neal, P.R., Sogin, M.L., and Welch, D.B.M. (2009). Effect of PCR amplicon size on assessments of clone library microbial diversity and community structure. *Environ. Microbiol.* 11, 1292–1302. doi: 10.1111/j.1462-2920.2008.01857.x

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W. Verspagen, J.M.H., and Visser, M. (2018). Cyanobacterial blooms. *Nat. Rev. Microbiol.* 16, 471–482. doi: 10.1038/s41579-018-0040-1

Hutchins, D.A., and Fu, F-X. (2017). Microorganisms and ocean global change. *Nat. Microbiol.* 2:17508. doi: 10.1038/nmicrobiol.2017.58

Hutchins, D., Fu, F-X., Zhang, Y., Warner, M.E., Feng, Y., Portune, K., et al. (2007). CO_2 control of *Trichodesmium* N_2 fixation, photosynthesis, growth rates, and elemental ratios: implications for past, present, and future ocean biogeochemistry. *Limnol. Oceanogr.* 52, 1293–1304. doi: 10.4319/lo.2007.52.4.1293

Hutchins, D. A., Jansson, J. K., Remais, J., Rich, V. I, Singh, B. K., and Trivedi, P. (2019). Climate change microbiology—problems and perspectives. *Nat. Rev. Microbiol.* 17, 391–396. doi: 10.1038/s41579-019-0178-5

Ibarbalz, F. M., Henry, N., Brandão, M.C., Martini, S., Busseni, G., Byrne, H., et al. (2019). Global trends in marine plankton diversity across kingdoms of life. *Cell* 179, 1084–1097.e21

IPCC, 2018: Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V., Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P.R., et al. (eds.)]. In Press.

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S. et al. (eds.)]. Cambridge University Press.

Jessen, G.L., Lichtschlag, A., Ramette, A., Pantoja, S., Rossel, P.E., Schubert, C. J., et al. (2017). Hypoxia causes preservation of labile organic matter and changes seafloor microbial community composition (Black Sea). *Sci. Adv.* 3(2): e1601897. doi: 10.1126/sciadv.1601897

Jo, T., Murakami, H., Masuda, R., Sakata, M. K., Yamamoto, S., and Minamoto, T. (2017). Rapid degradation of longer DNA fragments enables the improved estimation of distribution and biomass using environmental DNA. *Mol. Ecol. Resour.* 17(6), e25−e33. doi: 10.1111/1755-0998.12685.

Jöhnk, K.D., Huisman, J.E.F., Sharples, J., Sommeijer, B., Visser, P.M., and Stroom, J.M. (2007). Summer heatwaves promote blooms of harmful cyanobacteria. *Glob. Change Biol.* 14(3), 495−512. doi: 10.1111/j.1365-2486.2007.01510.x

Jungblut, A., Hawes, I., Mountfort, D., Hitzfeld, B., Dietrich, D., Burns, B., et al. (2005). Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environ. Microbiol.* 7, 519–529. doi: 10.1111/j.1462-2920.2005.00717.x

Kaebernick, M., Dittman, E., Börner, T., and Neilan, B.A. (2002). Multiple alternate transcripts direct the biosynthesis of microcystin, a cyanobacterial. *Appl. Environ. Microbiol.* 68, 449–455. doi: 10.1128/aem.68.2.449-455.2002

Kaneko, T., Nakajima, N., Okamoto, S., Suzuki, I., Tanabe, Y., Tamaoki, M., et al. (2007). Complete genomic structure of the bloom-forming toxic cyanobacteriam *Microcystis aeruginosa* NIES-843. *DNA Res.* 14, 247–256. doi: 10.1093/dnares/dsm026

Kaplan-Levy, R.N., Hadas, O., Summers, M.L., Rücker J., and Sukenik, A. (2010). Dormancy and Resistance in Harsh Environments. Chapter 2: Akinetes: dormant cells of cyanobacteria. New York, NY: Springer. p. 5–27.

Kasai, A., Takada, S, Yamazaki, A., Masuda, R., and Yamanaka, H. (2020). The effect of temperature on environmental DNA degradation of Japanese eel. *Fish. Sci.* 86, 465–471. doi: 10.1007/s12562-020-01409-1

Kim, D-K., Park, K., Jo, H., Kwak, I-S. (2019). Comparison of Water Sampling between Environmental DNA Metabarcoding and Conventional Microscopic Identification: A Case Study in Gwangyang Bay, South Korea. *Appl. Sci. 9*:3272. doi: 10.3390/app9163272

Klemer, A.R., and Konopka, A.E. (1989.) Causes and consequences of blue-green algal (cyanobacterial) blooms. *Lake Reserv. Manag.* 5(1), 9–19.

Klitzke, S., Beusch, C., and Fastner, J. (2011). Sorption of the cyanobacterial toxins cylindrospermopsin and anatoxin-a to sediments. *Water Res*. 45(3), 1338–1346. doi: 10.1016/J.WATRES.2010.10.019

Kolzau, S., Dolman, A.M., Voss, M., and Wiedner, C. (2018). The response of nitrogen fixing cyanobacteria to a reduction in nitrogen loading. *Hydrobiology* 103(1-2), 5–14.

Komárek, J. (2010). Recent changes (2008) in cyanobacteria taxonomy based on a combinatino of molecular background with phenotype and ecological consequences (genus and species concept). *Hydrobiologia*, 639(2010), 245–259. doi: 10.1007/s10750-009-0031-3

Kotak, B.G., and Zurawell, R.W. (2007). Cyanobacterial toxins in Canadian freshwaters: A review. *Lake Reserv. Manag.* 23, 109 –122. doi: 10.1080/07438140709353915

Kyle, M., Anderson, T., Haande, S., and Rohrlack, T. (2015). Historical *Planktothrix* diversity across seven Norwegian lakes implies environmentally driven niche differentiation. *Front. Ecol. Evol.* 15(3):104. doi: 10.3389/fevo.2015.00104

Kyle, M., Haande, S., Sønstebø, J., and Rohrlack, T. (2015). Amplification of DNA in sediment cores to detect historic *Planktothrix* occurrence in three Norwegian lakes. *J. Paleolimnol.* 53, 61– 72. doi: 10.1007%2Fs10933-014-9807-1

Kyoung-Hee, O., Jeong, D-H., and Cho, Y-C. (2012). Quantification of toxigenic *Microcystis* spp. in freshwaters by Quantitative Real-time PCR based on the microcystin synthetase A gene. *J. Microbiol.* 51, 18–24. doi: 10.1007/s12275-013-2354-z

Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G., and Taylor, M.I. (2018). How quantitative is metabarcoding: A meta-analytical approach. *Mol. Ecol.* 28, 420–430. doi: 10.1111/mec.14920

Lamothe, K.A., Jackson, D.A., and Somers, K.M. (2018). Long-term directional trajectories among lake crustacean zooplankton communities and water chemistry. *Can. J. Fish. Aquat. Sci.* 75, 1926–1939. doi: 10.1139/cjfas-2017-0518

Leak, L.V. (1967). Fine structure of the mucilaginous sheath of *Anabaena* sp. *J. Ultrastrcuct. Res.* 21(1-2), 61–74. doi: 10.1016/S0022-5320(67)80006-6

Leavitt, P.R., and Findlay, D.L. (1994). Comparison of fossil pigments with 20 years of phytoplankton data from eutrophic Lake 227, Experimental Lakes Area, Ontario. *Can. J. Fish. Aquat. Sci.* 51, 2286–2299. doi: 10.1139/f94-232

Leavitt, P.R., Findlay, D.L., Hall, R.I., Smol, J.P. (1999). Algal responses to dissolved organic carbon loss and pH decline during whole-lake acidification: Evidence from paleolimnology. *Limnol. Oceanogr.* 44, 757–774.

Lee, C.K., Herbold, C.W., Polson, S.W., Wommack, K.E., Williamson, S.J., McDonald, I.R., et al. (2012). Groundtruthing next-gen sequencing for microbial ecology-biases and errors in community structure estimates from PCR amplicon pyrosequencing. *PLoS ONE* 7:44224. doi: 10.1371/journal.pone.0044224

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm? *Ecology* 74, 1659–1673. doi: 10.2307/1939924

Legendre, P., and Gallagher, E.D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* 129, 271–280. doi: 10.1007/s004420100716

Legendre, P., and Legendre, L. (2012). *Numerical Ecology, 3rd Edn*. Amsterdam: Elsevier.

Legrand, B., Lamarque, A., Sabart, M., and Latour, D. (2017). Benthic archives reveal recurrence and dominance of toxigenic cyanobacteria in a eutrophic lake over the last 220 years. *Toxins* 9:271. doi: 10.3390/toxins9090271

Legrand, B., Miras, Y., Beauger, A., Dussauze, M., and Latour, D. (2019). Akinetes and ancient DNA reveal toxic cyanobacterial recurrences and their potential for resurrection in a 6700-year-old core from a eutrophic lake. *Sci. Total Environ.* 15(687), 1369–1380. doi: 10.1016/j.scitotenv.2019.07.100.

Lyon-Colbert, A., Su, S., and Cude, C. (2018). A systematic literature review for evidence of *Aphanizomenon flos-aquae* toxigenicity in recreational waters and toxicity of dietary supplements: 2000-2017. *Toxins* 10:254. doi: 10.3390/toxins10070254

Lyra, C., Suomalainen, S., Gugger, M., Vezie, C., Sundman, P., Paulin, L., et al. (2001). Molecular characterization of planktic cyanobacteria of Anabaena, Aphanizomenon, Microcystis and Planktothrix genera. *Int. J. Syst. Evol. Microbiol.* 51(2), 513–526. doi: 10.1099/00207713-51-2-513.

MacKeigan, P.W., Garner, R.E., Monchamp, M.È, Walsh, D.A., Onana, V.E., Kraemer, S.A. et al. (2022). Comparing microscopy and DNA metabarcoding techniques for identifying cyanobacteria assemblages across hundreds of lakes. *Harmful Algae,* 113:102187. doi: 10.1016/j.hal.2022.102187

Madigan, M.T., Martinko, J.M., Bender, K.S., Buckley, D.H., and Stahl, D.A. (2012). *Brock Biology of Microorganisms*. 13th edn. London, UK: Pearson.

Magonono, M., Oberholster, P.J., Shonhai, A., Makumire, S., and Gumbo, J.R. (2018). The presence of toxic and non-toxic cyanobacteria in the sediments of the Limpopo River Basin: Implications for human health. *Toxins* 10(7):269. doi: 10.3390/toxins10070269.

Mannan, R.M., and Pakrasi, H.B. (1993). Dark heterotrophic growth conditions result in an increase in the content of Photosystem II units in the filamentous cyanobacterium Anabaena variabilis ATCC 29413. *Plant Physiol.* 103(3), 971–977. doi: 10.1104/pp.103.3.971

Manter, D.K., Weir, T.L., and Vivanco, J.M. (2010). Negative effects of sample poooling on PCR-based estimates of soil microbial richness and community strucutre. *Appl. Environ. Microbial.* 76, 2086–2090. doi: 10.1128/AEM.03017-09

Martínez de la Escalera, G.M., Antoniades, D., Bonilla, S., and Piccini, C. (2014). Application of ancient DNA to the reconstruction of past microbial assemblages and for the detection of toxic cyanobacteria in subtropical freshwater ecosystems. *Mol. Ecol.* 23(23), 5791–5802. doi: 10.1111/mec.12979

McMurdie, P.J., and Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PloS ONE* 8(4):e61217. doi: 10.1371/journal.pone.0061217

McMurter, H.J.G. (1991). Flourescence characteristics of picocyanobacteria: Effects of light and nitrogen availability. Master's thesis, University of Ottawa. doi: 10.20381/ruor-15535.

Mejbel, H.S., Dodsworth, W., and Pick, F.R. (2021a). DNA from Muddy Time Capsules. *LakeLine,* Summer 2021: 16–20.

Mejbel, H.S., Dodsoworth, W., Baud, A., Gregory-Eaves, I., and Pick, F.R. (2021b). Comparing quantitative methods for analyzing sediment DNA records of cyanobacteria in experimental and reference lakes. *Front. Microbiol.* 12:669910. doi: 10.3389/fmicb.2021.669910

Mejbel, H.S., Dodsworth, W., and Pick, F.R. (2022). The effects of temperature and oxygen on cyanobacterial DNA preservation in sediments: a comparison study of major taxa. *Environmental DNA.* 00, 1 – 15. doi: 10.1002/edn3.289

Mills, K.H., Chalanchuk, S.M., Mohr, L.C., and Davies, I.J. (1987). Responses of fish populations in Lake 223 to 8 years of experimental acidification. *Can. J. Fish. Aquat. Sci.* 44: 114–125.

Molot, L.A., Higgins, S.N., Schiff, S.L., Venkiteswaran, J.J., Paterson, M.J., and Baulch, H.M. (2021). Phosphorus-only fertilization rapidly initiates large nitrogen-fixing cyanobacteria blooms in two oligotrophic lakes. *Environ. Res. Lett.* 16:064078.

Monchamp, M-È., Bruel, R., Frossard, V., McGowan, S., Lavrieux, M., Muschick, M., et al. (2021). Paleoecological evidence for multi-trophic regime shift in a perialpine lake (lake Joux, Switzerland). *Anthropocene* 35:100301. doi: 10.1016/j.ancene.2021.100301

Monchamp, M-È., Pick, F.R., Beisner, B.E., and Maranger, R. (2014). Nitrogen forms influence microcystin concentration and composition via change in cyanobacterial community structure. *PloS ONE* 9:e85573. doi: 10.1371/journal.pone.0085573

Monchamp, M-È., Spaak, P., Domaizon, I., Dubois, N., Bouffard, D., and Pomati, F. (2018). Homogenization of lake cyanobacterial communities over a century of climate change and eutrophication. *Nat. Ecol. Evol.* 2, 317–324. doi: 10.1038/s41559-017-0407-0

Monchamp, M-È., Walser, J-C., Pomati, F., and Spaak, P. (2016). Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two peri-alpine lakes. *Appl. Environ. Microbiol.* 82, 6472–6482. doi: 10.1128/AEM.02174-16

More, K. D., Giosan, L., Grice, K., and Coolen, M. J. L. (2019). Holocene paleodepositional changes reflected in the sedimentary microbiome of the Black Sea. *Geobiology*, 17, 436–448. doi: 10.1111/gbi.12338

Nathan, L.M., Simmons, M., Wegleitner, B.J., Jerde, C.L., and Mahon, A.R. (2014). Quantifying environmental DNA signals for aquatic invasive species across multiple detection platforms. *Environ. Sci. Technol.* 48, 12800–12806. doi: 10.1021/es5034052

Nwosu, E.C., Brauer, A., Kaiser, J., Horn, F. Wagner, D., and Liebner, S. (2021a). Evaluating sedimentary DNA for tracing changes in cyanobacteria dynamics from sediments spanning the last 350 years of Lake Tiefer, See, NE Germany. *J. Paleolimnol.* 66, 279–296. doi: 10.1007/s10933-021-00206-9.

Nwosu, E. C., Roeser, P., Yang, S., Ganzert, L., Dellwig, O., Pinkerneil, S., et al. (2021b). From Water into sediment—tracing freshwater cyanobacteria via DNA Analyses. *Microorganisms*, 2021, 1778. doi: 10.3390/microorganisms9081778

Nübel, U., Garcia-Pichel, F., and Muyzer, G. (1997). PCR primers to amplify 16S rRNA genes from cyanobacteria. *Appl. Environ. Microbiol.* 63, 3327–3332. doi: 10.1128/aem.63.8.3327-3332.1997

O'Connell, D.W., Ansems, N., Kukkadapu, R.K., Jaisi, D., Orihel, D.M., Cade-Menun, B.J., et al. (2020). Changes in sedimentary phosphorus burial following artificial eutrophication of Lake 227, Experimental Lakes Area, Ontario, Canada. *J. Geophys. Res.* 125:8. doi: 10.1029/2020jg005713

O'Neil, J.M., Davis, T.W., Burford, M.A., and Gobler, C.J. (2012). The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae* 14, 313-34.

Orihel, D.M., Bird, D.F., Brylinsky, M., Chen, H., Donald, D.B., Huang, D.Y., et al. (2012). High microcystin concentrations occur only at low nitrogen-to-phosphorus rations in nutrient-rich Canadian lakes. *Can. J. Fish. Aquat Sci.* 69, 1457–1462. doi: 10.1139/F2012-088

Orsi, W. D., Coolen, M. J. L., Wuchter, C., He, L., More, K. D., Irigoien, et al. (2017). Climate oscillations reflected within the microbiome of Arabian Sea sediment. *Sci. Rep.* 7, 6040. doi: 10.1038/s41598-017-05590-9

Paerl, H.W. (2014). Mitigating harmful cyanobacterial blooms in a human- and climatically-impacted world. *Life (Basel)* 4, 988–1012.

Paerl, H.W., and Huisman, J. (2009). Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environ. Microbiol. Rep.* 1(1), 27–37. doi: 10.1111/j.1758-2229.2008.00004.x.

Paerl, H.W., and Otten, T.G. (2013). Harmful cyanobacterial blooms: causes, consequences, and controls. *Microb. Ecol.* 65(4):995–1010. doi: 10.1007/s00248-012-0159-y

Pal, S. (2015). Temporal and spatial trends in toxic cyanobacteria as identified through lake sediment DNA. PhD thesis, University of Ottawa.

Pal, S., Gregory-Eaves, I., and Pick, F.R. (2015). Temporal trends in cyanobacteria revealed through DNA and pigment analyses of temperate lake sediment cores. *J. Paleolimnol.* 54, 87–101. doi: 10.1007/s10933-015-9839-1

Parducci, L., Bennett, K.D., Ficetola, G.F., Alsos, I.G., Suyama, Y., Wood, J.R., et al. (2017). Ancient plant DNA in lake sediments. *New Phytol.* 214, 924–942. doi: 10.1111/nph.14470

Paterson, M.J., Schindler, D.W., Hecky, R.E., Findlay, D.L., and Rondeau, K.J. (2011). Comment: Lake 227 shows clearly that controlling inputs of nitrogen will not reduce or prevent eutrophication of lakes. *Limnol. Oceanogr.* 56, 1545–1547. doi: 10.4319/lo.2011.56.4.1545

Pedersen, M.W., Overballe-Petersen, S., Ermini, L., Der Sarkissian, C., Haile, J., Hellstrom, M., et al. (2015). Ancient and modern environmental DNA. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 370(1660): 20130383. doi: 10.1098/rstb.2013.0383

Peres-Neto, P.R., Legendre, P., Dray, S., Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology* 87, 2614-2625.

Pick, F.R. (1991). The abundance and composition of freshwater picocyanobacteria in relation to light penetration. *Limnol. Oceanogr.* 36, 1457–1462. doi: 10.4319/lo.1991.36.7.1457

Pick, F.R. (2016). Blooming algae: a Canadian perspective on the rise of toxic cyanobacteria. *Can. J. Fish. Aquat. Sci.* 73(7), 1149–1158. doi: 10.1139/cjfas-2015-0470

Pick, F.R., and Lean, D.R.S. (1987). The role of macronutrients (C, N, P) in controlling cyanobacterial dominance in temperate lakes. *N. Z. J. Mar.* 21, 425–434. doi : 10.1080/00288330.1987.9516238

Pietramellara, G., Ascher, J., Borgogni, F., Ceccherini, M.T., Guerri, G., and Nannipieri, P. (2009). Extracellular DNA in soil and sediment: fate and ecological relevance. *Biol. Fertil.* 45(3), 219–235.

Pilon, S., Zastepa, A., Taranu, Z.E., Gregory-Eaves, I., Racine, M., Blais, J.M., et al. (2019). Contrasting histories of microcystin-producing cyanobacteria in two temperate lakes as inferred from quantitative sediment DNA analyses. *Lake Reserv. Manag.* 35, 102–117. doi: 10.1080/10402381.2018.1549625

Poulain, A.J., Aris-Brosou, S., Blais, J.M., Brazeau, M., Keller, W., and Paterson, A.M. (2015).

Microbial DNA records historical delivery of anthropogenic mercury. *ISME J.* 9, 2541–2550. doi: 10.1038/ismej.2015.86

Proulx, M., Pick, F.R., Mazumder, A., Hamilton, P.B., and Lean, D.R.S. (1996). Experimental evidence for interactive impacts of human lake algal species richness. *OIKOS* 76, 191 – 195.

Quay, P.D., Broecker, W.S., Hesslein, R.H., and Schindler, D.W. (1980). Vertical diffusion rates determined by tritium tracer experiments in the thermocline and hypolimnion of two lakes. *Limnol. Oceanogr.* 25, 201–218. doi: 10.4319/lo.1980.25.2.0201

Ramírez, G.A., Graham, D., and D'Hondt, S. (2018). Influence of commercial DNA extraction kit choice on prokaryotic community metrics in marine sediment. *Limnol. Oceanogr. Methods* 16, 525–536. doi: 10.1002/lom3.10264

Renaud, S.L. (2009). Microcystin production and the dominance of toxigenic strains of cyanobacteria in lake and culture studies. PhD thesis, University of Ottawa.

Reyes, J.C., and Florencio, F.J. (1995). Electron transport controls transcription of the glutamine synthetase gene (*glnA*) from the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Mol. Biol.* 27, 789–799.

Reynolds, C.S., and Walsby, A.E. (1975). Water-Blooms. *Biol.* 50, 437-481.

Righetti, D., Vogt, M., Gruber, N., Psomas, A., and Zimmermann, N.E. (2019). Global pattern of phytoplankton diversity driven by temperature and environmental variability. *Sci. Adv.* 5, eaau6253

Rigosi, A., Carey, C.C., Ibelings, B.W., and Brookes, J.D. (2014). The interaction between climate warming and eutrophication to promote cyanobacteria is dependent on trophic state and varies among taxa. *Limnol. Oceanogr.* 59, 99–114.

Rinta-Kanto, J.M., Ouellette, A.J.A., Boyer, G.L., Twiss, M.R., Bridgeman, T.B., and Wilhelm, S.W. (2005). Quantification of toxic *Microcystis* spp. during the 2003 and 2004 blooms in western Lake Erie using quantitative real-time PCR. *Environ. Sci. Technol.* 39, 4198–4205. doi: 10.1021/es048249u

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stainer, R.Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* 111(1), 1–61.

Rochelle, P.A., Fry, J.C., John Parkes, R., and Weightman, A.J. (1992). DNA extraction for 16S rRNA gene analysis to determine genetic diversity in deep sediment communities. *FEMS Microbiol. Lett.* 100, 59–65. doi: 10.1111/j.1574-6968.1992.tb14019.x

Ruppert, K.M., Kline, R.J., and Rahman, M.S. (2019). Past, present, and future perspectives of

environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Glob. Ecol. Conserv.* 17:e00547. doi: 10.1016/j.gecco.2019.e00547

Sakata, M.K., Yamamoto, S., Gotoh, R. O., Miya, M., Yamanaka, H., and Minamoto, T. (2020). Sedimentary eDNA provides different information on timescale and fish species composition compared with aqueous eDNA. *eDNA* 2(4), 505–518. doi: 10.1002/edn3.75

Samuel, A.D.T., Petersen, J.D., and Reese, T.S. (2001). Envelope structure of *Synechococcus* sp. WH8113, a nonflagellated swimming cyanobacterium. *BMC Microbiol.* 1:4. doi: 10.1186/1471-2180-1-4

Savichtcheva, O., Debroas, D., Kurmayer, R., Villar, C., Jenny, J.P., Arnaud, F., et al. (2011). Quantitative PCR enumeration of total/toxic *Planktothrix rubescens* and total cyanobacteria in preserved DNA isolated from lake sediments. *Appl. Environ. Microbiol.* 77, 8744–8753. doi: 10.1128/AEM.06106-11

Savichtcheva, O., Debroas, D., Perga, M. E., Arnaud, F., Villar, C., Lyautey, E., et al. (2014). Effects of nutrients and warming on *Planktothrix* dynamics and diversity: a palaeolimnological view based on sedimentary DNA and RNA. *Freshw. Biol.* 60, 31–49. doi: 10.1111/fwb.12465

Sayers, M.J., Grimm, A.G., Shuchman, R.A., Bosse, K.R., Fahnenstiel, G.L., Ruberg, S.A., et al. (2019). Satellite monitoring of harmful algal blooms in the Western Basin of Lake Erie: a 20-year time-series. *J. Great Lakes Res.* 45(3), 508–521. doi: 10.1016/j.jglr.2019.01.005

Schiff, S.L., Tsuji, J.M., Wu, L., Venkiteswaran, J.J., Molot, L.A., Elgood, R.J., et al. (2017). Millions of boreal shield lakes can be used to probe Archaean ocean biogeochemistry. *Sci. Rep.* 7:46708.

Schindler, D.W. (1974). Eutrophication and recovery in experimental lakes: Implications for lake management. *Science* 184, 897–899. doi: 10.1126/science.184.4139.897

Schindler, D.W. (1977). Evolution of phosphorus limitation in lakes. *Science* 195, 260–262. doi: 10.1126/science.195.4275.260

Schindler, D.W. (2009). A personal histroy of the Experimental Lakes Project. *Can. J. Fish. Aquat Sci.* 66, 1837–1847. doi: 10.1139/F09-134.

Schindler, D.W., Armstrong, F.A.J., Holmgren, S.K., and Brunskill, G.J. (1971). Eutrophication of Lake 227, Experimental Lakes Area, Northwestern Ontario, by addition of phosphare and nitrate. *J. FIsh. Res. Board Can.* 28, 1763–1782.

Schindler, D.W., Bayley, S.E., Parker, B.R., Beaty, K.G., Cruikshank, D.R., Fee, E.J., et al. (1996). The effects of climatic warming on the properties of boreal lakes and streams at the Experimental Lakes Area, northwestern Ontario. *Limnol. Oceanogr.* 41, 1004–1017. doi:

10.4319/lo.1996.41.5.1004

Schindler, D.W., Beaty, K.G., Fee, E.J., Cruikshank, D.R., DeBruyn, E.R., Findlay, D.L., et al. (1990). Effects of climatic warming on lakes of the central boreal forest. *Science* 250, 967–970. doi: 10.1126/science.250.4983.967

Schindler, D.W., Hecky, R.E., Findlay, D.L., Stainton, M.P., Parker, B.R., Paterson, M.J., et al. (2008). Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. *Proc. Natl. Acad. Sci. U.S.A.* 105, 11254–11258. doi: 10.1073/pnas.0805108105

Schindler, D.W., Wagemann, R., Cook, R.B., Ruszczynski, T., and Prokopowich, J. (1980). Experimental acidification of Lake 223, Experimental Lakes Area: Background data and the first three years of acidification. *Can. J. Fish. Aquat. Sci.* 37, 342–354. doi: 10.1139/f80-048

Schirrmeister, B.E., Dalquen, D.A., Anisimova, M., and Bagheri, H.C. (2012). Gene copy number variation and its significance in cyanobacterial phylogeny. *BMC Microbiol.* 12:177. doi: 10.1186/1471-2180-12-177

Scott, J. T., and McCarthy, M.J. (2010). Nitrogen fixation may not balance the nitrogen pool in lakes over timescales relevant to eutrophication management. *Limnol. Oceanogr*. 55(3), 1265–1270. doi: 10.4319/lo.2010.55.3.1265scot

Shelton, A.O., O'Donnell, J.L., Samhouri, J.F., Lowell, N., Williams, G.D., and Kelly, R.P. (2016). A framework for inferring biological communities from environmental DNA. *Ecol. Appl.* 26, 1645–1659. doi: 10.1890/15-1733.1

Shi, X., Yang, L., Niu, X., Xiao, L., Kong, Z., Qin, B., et al. (2003). Intracellular phosphorus metabolism of *Microcystis aeruginosa* under various redox potential in darkness. *Microbiol.* 158(4), 345–352. doi: 10.1078/0944-5013-00214

Sierra, M.A., Li, Q., Pushalkar, S., Paul, B., Sandoval, T.A., Kamer, A.R., et al. (2020). The influences of bioinformatics tools and reference databases in analyzing the human oral microbial community. *Genes* 11:878. doi: 10.3390/genes11080878

Singh, G., Sithebe, A., Enitan, A.M., Kumari, S., Bux, F., and Stenström, T.A. (2017). Comparison of droplet digital PCR and quantitative PCR for the detection of *Salmonella* and its application for river sediments. *J. Water Health* 15, 505–508. doi: 10.2166/wh.2017.259

Sivonen, K. (2009). *Cyanobacterial toxins* in *Encyclopedia of Microbiology (Third edition)*, 290–307. doi: 10.1016/B978-012373944-5.00005-5

Sjögren, P., Edwards, M. E., Gielly, L., Langdon, C. T., Croudace, I. W., Merkel, M. K. F., et al. (2017). Lake sedimentary DNA accurately records 20th Century introductions of exotic

conifers in Scotland. *New Phytol.* 213(2), 929–941. doi: 10.1111/nph.14199

Skulberg, O.M., Carmichael, W.W., Codd, G.A. and Skulberg, R. (1993). *Taxonomy of Toxic Cyanophyceae (cyanobacteria)* in *Algal toxins in seafood and drinking water.* I.R. Falconer (ed.). Academic Press, San Diego, CA. pp. 145–164.

Šmarda, J., Smajs, D., Komrska, J., and Krzyzanek, V. (2002). S-layers on cell walls of cyanobacteria. *Micron.* 33(3), 257–277. doi: 10.1016/S0968-4328(01)00031-2

Smol, J.P. (2008). *Pollution of Lakes and Rivers: A Paleoenvironmental Perspective. 2nd Edn.* Malden, MA: Wiley-Blackwell publishing.

Smol, J.P. (2010). The power of the past: using sediments to track the effects of multiple stressors on lake ecosystems. *Freshw. Biol.* 55(s1):43–59. doi: 10.1111/j.1365-2427.2009.02373.x

Steinberg, C.E.W., and Hartmann. H.M. (1988). Planktonic bloom-forming Cyanobacteria and the eutrophication of lakes and rivers. *Freshw. Biol.* 20, 279–287.

Stewart, K.A. (2019). Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. *Biodivers. Conserv.* 28, 983–1001. doi: 10.1007/s10531-019-01709-8

Strickler, K.M., Fremier, A.K., and Goldberg, C.S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biol. Conserv.* 183, 85–92. doi: 10.1016/j.biocon.2014.11.038

Sukenik, A., Hadas, O., Kaplan, A., and Quesada, A. (2012). Invasion of Nostocales (cyanobacteria) to subtropical and temperate freshwater lakes – physiological, regional, and global driving forces. *Fron. Microbiol.* 3:86.

Taberlet, P., Bonin, A., Zinger, L., and Coissac, E. (2018). *Evironmental DNA: For Biodiversity Research and Monitoring*. Oxford; New York, NY: Oxford University Press.

Taly, V., Pekin, D., Benhaim, L., Kotsopoulos, S.K., Le Corre, D., Li, X., et al. (2013). Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. *Clin. Chem.* 59(12), 1722–1731. doi: 10.1373/clinchem.2013.206359

Tanabe, Y., Kaya, K., and Watanabe, M.M. (2004). Evidence for recombination in the microcystin synthetase (*mcy*) genes of toxic cyanobacteria *Microcystis* spp. *J. Mol. Evol.* 58, 633–641

Taranu, Z.E., Carpenter, S.R., Frossard, V., Jenny, J-P., Thomas, Z., Vermaire, J.C., et al. (2018). Can we detect ecosystem critical transitions and signals of changing resilience from paleo-ecological records? *Ecosphere* 9:e02438. doi: 10.1002/ecs2.2438

Taranu, Z.E., Gregory-Eaves, I., Leavitt, P.R., Bunting, L., Buchaca, T., Catalan, J., et al. (2015). Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecol. Lett.* 18(4), 375–384. doi: 10.1111/ele.12420

Taranu, Z.E., Zurawell, R.W., Pick, F., and Gregory-Eaves, I. (2012). Predicting cyanobacterial dynamics in the face of global change: the importance of scale and environmental context. *Glob. Change Biol.* 18, 3477–3490. doi: 10.1111/gcb.12015

Taylor, S.C., Laperriere, G., and Germain, H. (2017). Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data. *Sci. Rep.* 7:2409. doi: 10.1038/s41598-017-02217-x

Taylor, S.C., Nadeau, K., Abbasi, M., Lachance, C., Nguyen, M., and Fenrich, J. (2019). The ultimate qPCR experiment: Producing publication quality, reproducible data the first time. *Trends Biotechnol.*, 37 761–774. doi: 10.1016/j.tibtech.2018.12.002

Thomsen, P.F., and Willerslev, E. (2015). Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. *Biol. Conserv.* 183, 4–18. doi: 10.1016/j.biocon.2014.11.019

Tillett, D., Dittmann, E., Erhard, M., Von Döhren, H., Börner, T., and Neilan, B.A. (2000). Structural organization of microcystin biosynthesis in Microcystis aeruginosa PCC7806: An integrated peptide-polyketide synthetase system. *Chem. Biol*. 7, 753-64.

Tse, T.J., Doig, L.E., Tang, S., Zhang, X., Sun, W., Wiseman, S.B., et al. (2018). Combining high-throughput sequencing of *seda*DNA and traditional paleolimnological techniques to infer historical trends in cyanobacterial communities. *Environ. Sci. Technol.* 52, 6842–6853. doi: 10.1021/acs.est.7b06386

United States Environmental Protection Agency. (2006). Data Quality Assessment: Statistical Methods for Practitioners. Washington, DC: Office of Environmental Information.

Urbach, E., Robertson, D.L., and Chisholm, S.W. (1992). Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. *Nature* 355, 267–270. doi: 10.1038/355267ao

Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., et al. (2018). Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. *MBMG* 2:e23297. doi: 10.1101/113472

van Geel, B., Mur, L.R., Ralska-Jasiewiczowa, M., and Goslan, T. (1994). Fossil akinete of *Aphanizomenon* and *Anabaena* as indicators for medieval phosphate-eutrophication of Lake Gosciaz (Central Poland). *Rev. Paleobot*. 83(1994), 97–105.

Verschuren, D. (2000). Freeze coring soft sediments in tropical lakes. *J. Paleolimnol.* 24, 361–

365. doi: 10.1023/A:1008191418497

Verspagen, J.M.H., Visser, P.M., and Huisman, J. (2006). Aggregation with clay causes sedimentation of the buoyant cyanobacteria Microcystis spp. *Aquat. Microb. Ecol.* 44(2), 165–174. doi: 10.3354/ame044165

Vincent, W.F. (2009a). *Cyanobacteria.* In Azim, S.M.E. & E.L. Gene (eds), Encyclopedia of Inland Waters. Academic Press, Oxford: 226–232. doi: 10.1016/B978-012370626-3.00127-7

Vincent, W.F. (2009b). *Effects of Climate Change on Lakes.* In Azim, S.M.E. & E.L. Gene (eds), Encyclopedia of Inland Waters. Academic Press, Oxford: 55–60. doi: 10.1016/B978-012370626-3.00233-7

Waters M.N. (2016). A 4700-year history of cyanobacteria toxin production in a shallow subtropical lake. *Ecosystems*, 19, 426–436. doi: 10.1007/s10021-015-9943-0

Watson, S.B., McCauley, E., and Downing, J.A. (1997). Patterns in phytoplankton taxonomic composition across temperature lakes of differing nutrient status. *Limnol. Oceanogr.* 42:487-495.

Wei, N., Nakajima, F., and Tobino, T. (2018). A microcosm study of surface sediment environmental DNA: decay observation, abundance estimation, and fragment length comparison. *Environ. Sci. Technol.* 52(21), 12428–12435. doi: 10.1021/acs.est.8b04956

Weisbrod, B., Wood, S.A., Steiner, K., Whyte-Wilding, R., Puddick, J., Laroche, O., et al. (2020). Is a central sediment sample sufficient? Exploring spatial and temporal microbial diversity in a small lake. *Toxins* 12(9), 580. doi: 10.3390/toxins12090580.

Wert, E.C., Dong, M.M., and Rosario-Ortiz, F.L. (2013). Using digital flow cytometry to assess the degradation of three cyanobacteria species after oxidation processes. *Water Res.* 47(11), 3752 – 3761. doi: 10.1016/J.WATRES.2013.04.038

Whitton, B. (2012). *Ecology of Cyanobacteria II: Their Diversity in Space and Time.* Springer Nature Switzerland.

Wilk, L., Strauss, M., Rudolf, M., Nicolaisen, K., Flores, E., Kühlbrandt, W., et al. (2011). Outer membrane continuity and septosome formation between vegetative cells in the filaments of Anabaena sp. PCC 7120. *Cell. Microbiol.* 13(11), 1744–1754. doi: 10.1111/j.1462-5822.2011.01655.

Wiltse, B., Paterson, A.W., Findlay D.L., and Cumming, B.F. (2016). Seasonal and decadal patterns in *Discostella* (Bacillariophyceae) species from bi-weekly records of two Boreal lakes (Experimental Lakes Area, Ontario, Canada). *J. Phycol.* 28, 817–826. doi: 10.1111/jpy.12443

Winter, J.G., DeSellas, A.M., Fletcher, R., Heintsch, L., Morley, A., Nakamoto, L., et al. (2011). Algal blooms in Ontario, Canada: Increases in reports since 1994. *Lake Reserv. Manag.* 27(2), 107–114. doi: 10.1080/07438141.2011.557765.

Wood, S.W., Kremp, A., Savela, H., Akter, S., Vartti, V-P., Saarni, S., et al. (2021). Cyanobacterial akinete distribution, viability, and cyanotoxin records in sediment archives from the Norther Baltic Sea. *Front. Microbiol.* 12:681881. doi: 10.3389/fmicb.2021.681881

Wood, S.A., Pochon, X., Laroche, O., von Ammon, U., Adamson, J., and Zaiko, A. (2019). A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for species-specific detection in environmental DNA. *Mol. Ecol. Resour.* 19, 1407–1419. doi: 10.1111/1755-0998.13055

Wurtsbaugh, W.A., Paerl, H.W., and Dodds, W.K. (2019). Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *WIREs Water* 6:e1372. doi: 10.1001/wat2.1373

Xia, Z., Zhan, A., Gao, Y., Zhang, L., Haffner, G.D., and MacIsaac, H.J. Early detection of a highly invasive bivalve based on environmental DNA (eDNA). (2018). *Bio. Invasions* 20, 437−447.

Xiao, X., Sogge, H., Lagesen, K., Tooming-Klunderud, A., Jakobsen, K.S., and Rohrlack, T. (2014). Use of high throughput sequencing and light microscopy show contrasting results in a study of phytoplankton occurrence in a freshwater environment. *PLoS ONE* 9:e106510. doi: 10.1371/journal.pone.0106510.

Yan, D., Xu, H., Yang, M., Lan, J., Hou, W., Wang, F. et al. (2019). Responses of cyanobacteria to climate and human activities at Lake Chenghai over the past 100 years. *Ecol. Indic.* 104, 755-- 763. doi: 10.1016/j.ecolind.2019.03.019

Ye, W., Tan, J., Liu, X., Lin, S., Pan, J., Li, D., et al. (2011). Temporal variability of cyanobacterial populations in the water and sediment samples of Lake Taihu as determined by DGGE and real-time PCR. *Harmful Algae* 10(5): 472−479. doi: 10.1016/j.hal.2011.03.002

Yeates, C., Gillings, M.R., Davison, A.D., Altavilla, N., and Veal, D.A. (1997). PCR amplification of crude microbial DNA extracted from soil. *Lett. Appl. Microbiol.* 25, 303–307. doi: 10.1046/j.1472-765x.1997.00232.x

Zaiko, A., Pochon, X., Garcia-Vazquez, E., Olenin, S., and Wood, S.A. (2018). Advantages and limitations of environmental DNA/RNA tools for marine biosecurity: Management and surveillance of non-indigenous species. *Front. Mar. Sci.* 5:322. doi: 10.3389/fmars.2018.00322

Zastepa, A., Taranu, Z.E., Kimpe, L.E., Blais, J.M., Gregory-Eaves, I., Zurawell, R.W., et al. (2016). Reconstructing a long-term record of microcystins from the analysis of lake

sediments. *Sci. Total Environ.* 579 (2017), 893–901.

Zeeb, B.A., Christie, C.E., Smol, J.P., Findlay, D.L., Kling, H.J., and Birks, H.J.B. (1994). Responses of diatom and chrysophyte assemblages in lake 227 sediments to experimental eutrophication. *Can. J. Fish. Aquat. Sci.* 51, 2300–2311

Zhang, H., Huo, S., Yeager, K.M., and Wu, F. (2021). Sedimentary DNA record of eukaryotic algal and cyanobacterial communities in a shallow Lake driven by human activities and climate change. *Sci. Total Environ.* 753, 141985. doi: 10.1016/j.scitotenv.2020.141985

Zhang, M., Kong, F., Shi, X., Xing, P., and Tan, X. (2007). Differences in responses to darkness between *Microcystis aeruginosa* and *Chlorella pyrenoidosa*. *J. Freshw. Ecol.* 22(1), 93–99. doi: 10.1080/02705060.2007.9664149

Zhang, Y., Pohlmann, E.L., Ludden, P.W., and Roberts, G.P. (2000). Mutageneis and functional characterization of the *glnB, glnA,* and *nifA* genes from the photosynthetic bacterium *Rhodospirillum rubrum. J. Bacteriol.* 182, 983–992 . doi: 10.1128/JB.182.4.983-992.2000

Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996). DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* 62, 316–322. doi: 10.1128/AEM.62.2.316-322.1996

APPENDIX A: Supplementary material for Chapter 2

Table S2.1 Surface water abiotic physical and chemical variables from the lake chosen for sediment sampling on the day of collection.

Variable	Depth (m)			
	0.5	1.0	2.0	3.0
Temperature (°C)	13.05	13.03	13.05	12.81
pH	7.28	7.19	7.11	6.36
Conductance/Specific conductance (SPC)	61.0	61.1	61.0	81.5
Oxidation-reduction potential (mV)	523	525	525	147
Dissolved oxygen saturation (%)	87.4	87.4	87.7	68.5
Dissolved oxygen concentration (mg/L)	9.26	9.28	9.31	7.3

Table S2.2 The average water content (%) in the microcosms at the timepoints used for the experiment.

	Water content (%)			
Timepoint	Anoxic 25	Oxic 25	Anoxic 4	Oxic 4
0 Days	95.2	95.2	95.2	95.2
24 hours	95.2	94.7	95.2	95.2
2 weeks	95.1	87.4	95.3	94.2
1 month	95.1	80.2	95.4	92.6
3 months	95.1	58.6	95.5	87.6
6 months	95.1	37.3	94.3	74.6
1 year	95.2	< 1	92.3	50.1

Table S2.3 The calculated *p*-values of two-tailed *t*-tests that assessed for statistical differences between the decay rate constants of the oxic and anoxic microcosms at 4°C and at 25°C. *n*=9 for cyanobacterial 16S rRNA (CYA) and *n*=3 for *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) for each condition.

Target gene	At 4°C	At 25°C
CYA	0.062	< 0.0001
MICR	0.016	< 0.01
apcA	< 0.01	< 0.01
ANA	0.033	< 0.0001

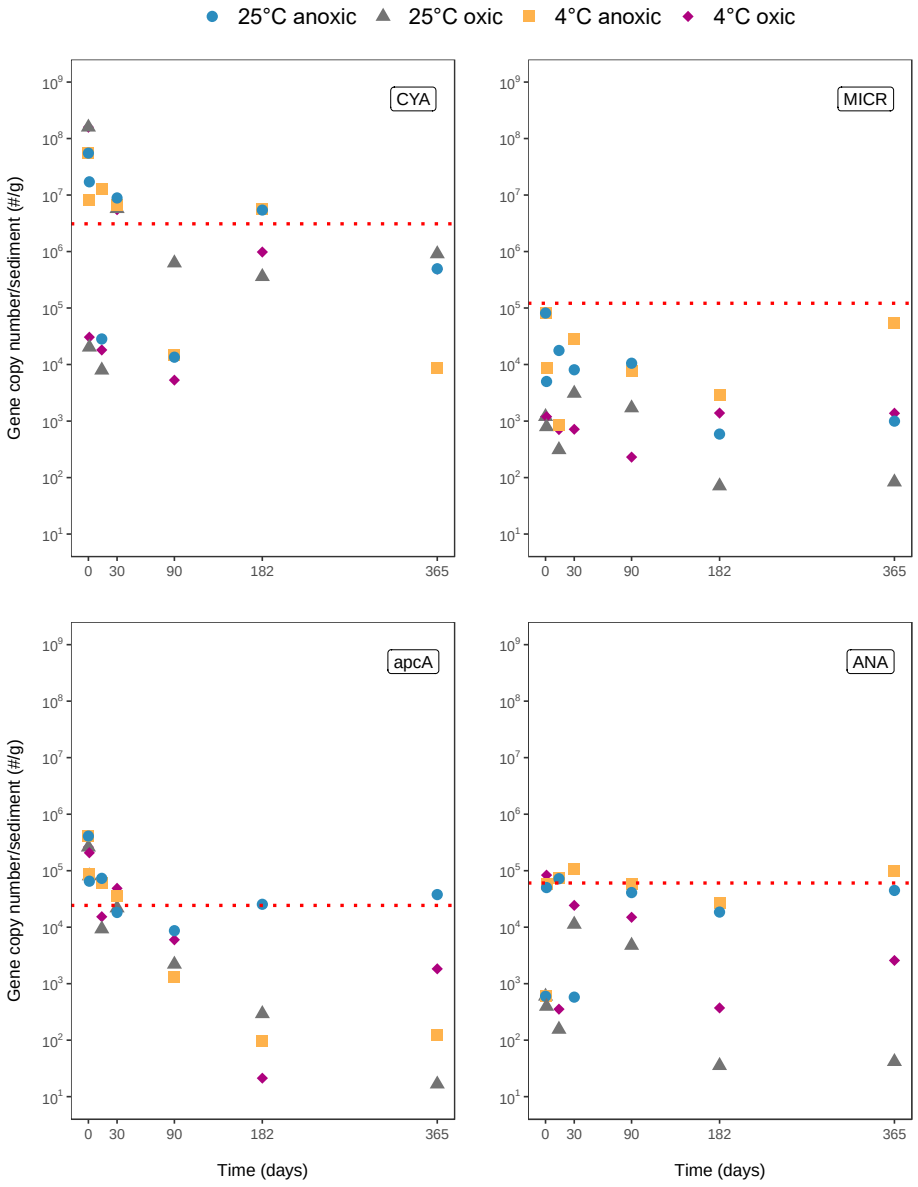

Figure S2.1 Absolute concentrations of cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at each environmental condition across 365 days in the experimental blank microcosms. Each point represents the droplet digital PCR-derived concentration, normalized per gram of dry sediment and presented on a logarithmic scale. The red dotted lines are the detection limits for each target gene. Note the CYA plot (top left) has a different y-axis scale.

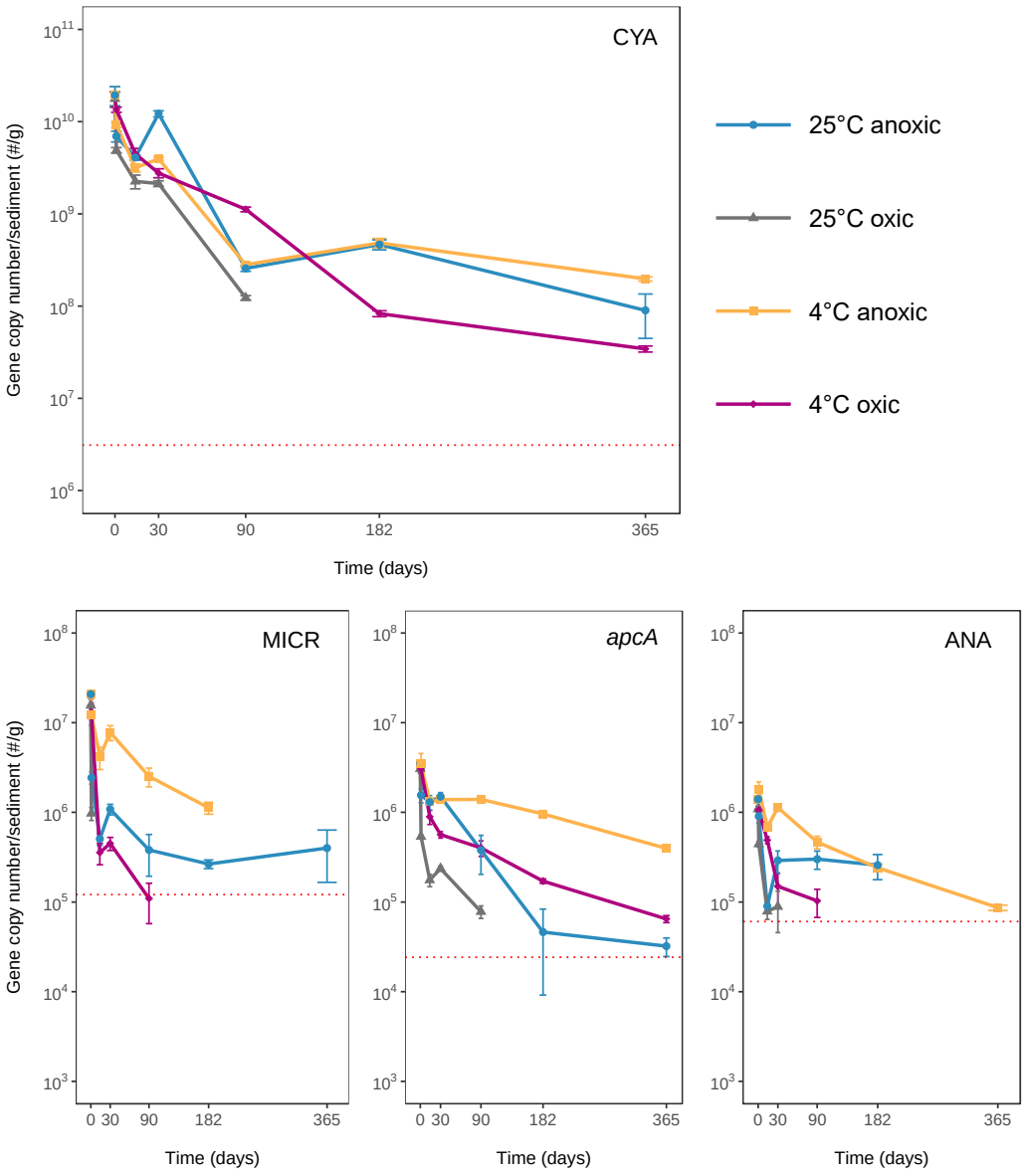

Figure S2.2 Absolute concentrations of cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at each environmental condition across 365 days. Gene copy numbers are normalized per gram of dry sediment and presented on a logarithmic scale. The error bars represent the standard error of the mean (*n*=9 for CYA, *n*=3 for MICR, *apcA*, and ANA). The red dotted lines are the detection limits for each target gene. Data where the average gene concentration were below the corresponding detection limits are not shown. Note the CYA plot (top) has a different y-axis scale.

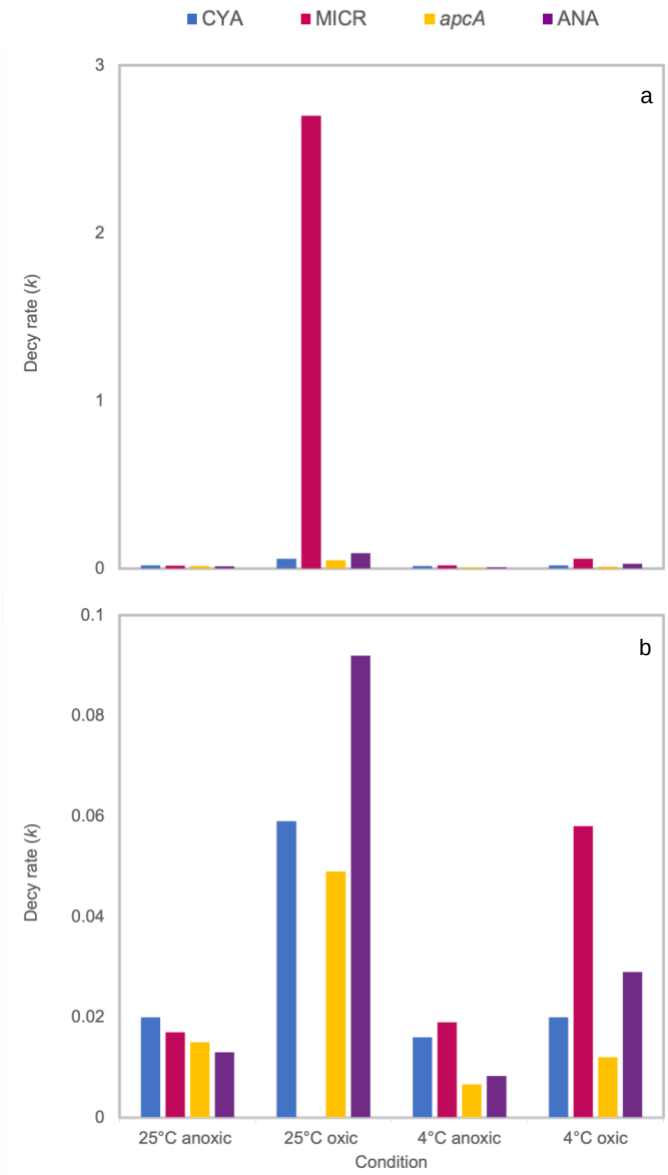

Figure S2.3 A comparison of the calculated decay rates of cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at each condition using the rates obtained from one-phase exponential decay models (a) and the same figure but with the decay rate for MICR under 25°C oxic conditions removed to better visualize and compare rates between the genes (b).

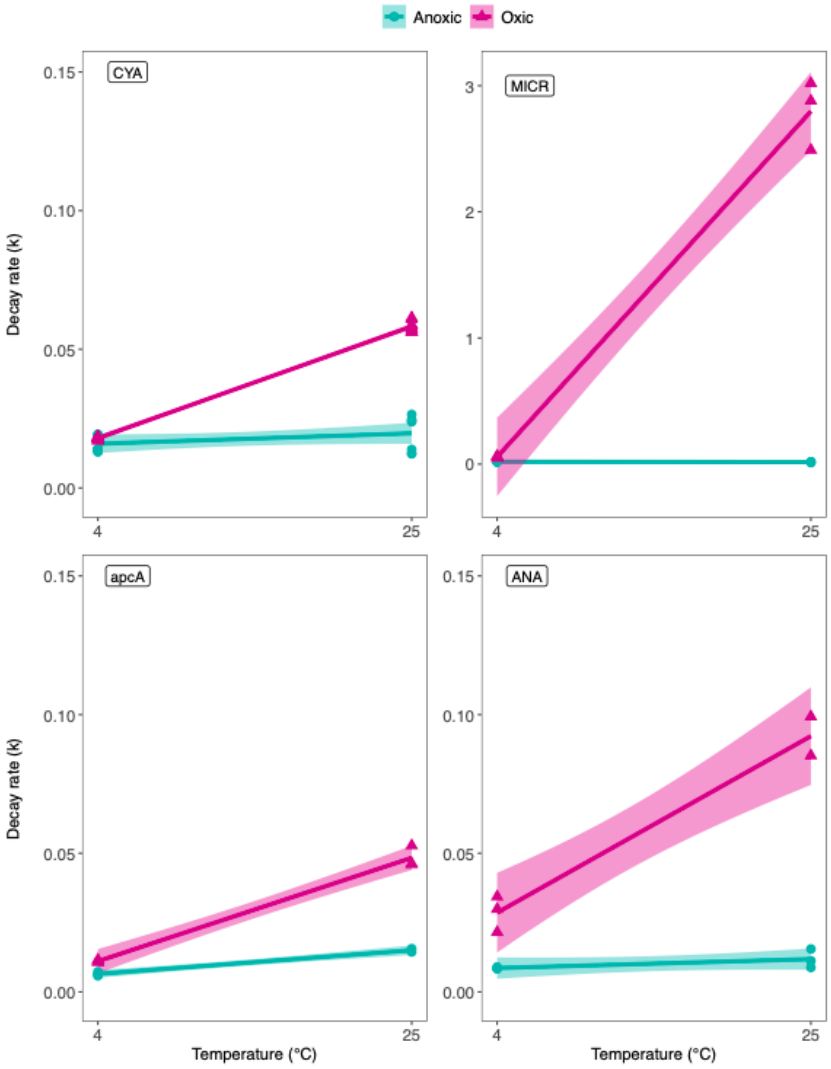

Figure S2.4 Relationship between temperature (°C) and the one phase exponential decay rate constants (*k*) for cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) under anoxic and oxic environments. The shading shows 95% confidence intervals of the mean (*n*=9 for CYA, *n*=3 for MICR, *apcA*, and ANA). Note the MICR plot (top right) has a different y-axis scale.

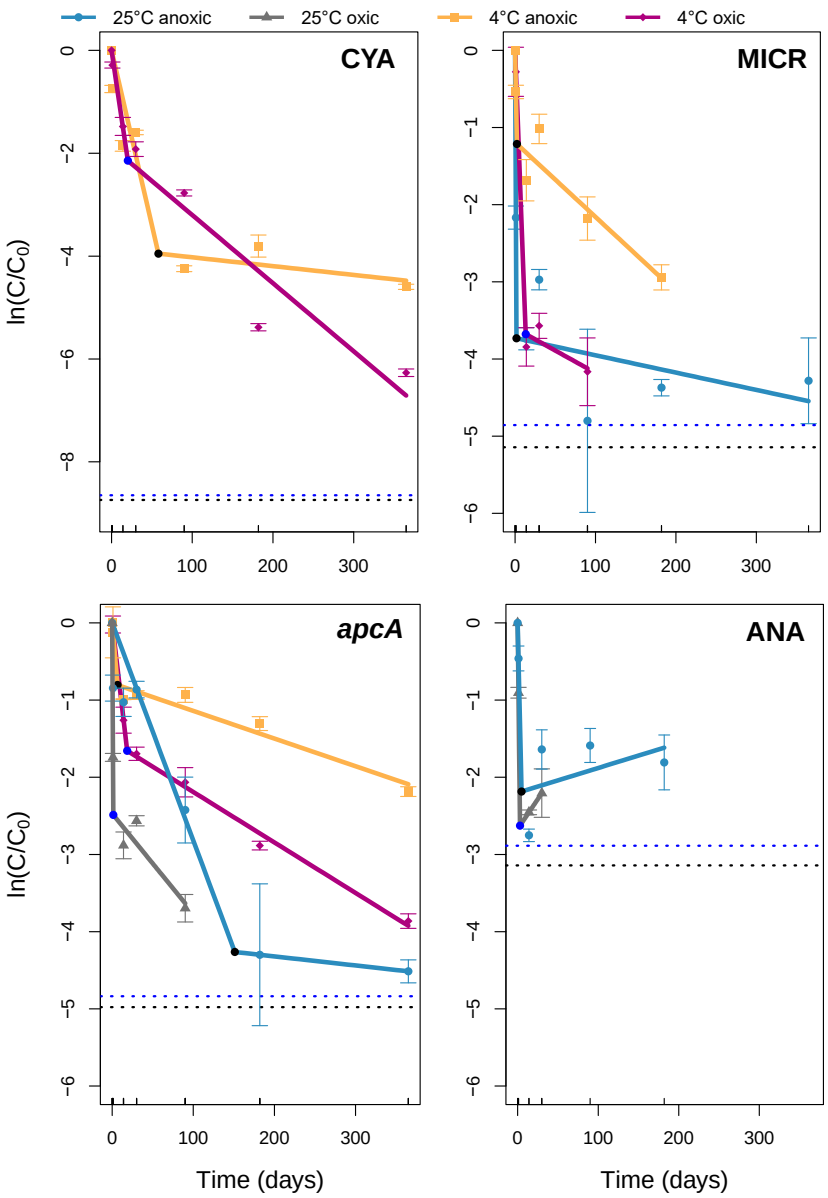

Figure S2.5 Biphasic decay curves for cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), *Synechococcus*-specific allophycocyanin alpha chain A (*apcA*), and *Anabaena* 16S rRNA (ANA) at the environmental conditions that exhibited a significant breakpoint. Each oxic breakpoint is highlighted with a blue circle and each anoxic breakpoint is highlighted with a black circle. The solid lines show the decay curves, and the blue and black horizontal dotted lines are the detection limits for each target gene in the oxic and anoxic microcosms, respectively. The error bars represent the standard error of the mean (*n*=9 for CYA, *n*=3 for MICR, *apcA*, and ANA). Data where the average gene concentration were below the corresponding detection limits are not shown. Note the CYA plot (top left) has a different y-axis scale.

APPENDIX B: Supplementary material for Chapter 3

Table S3.1 Experimental details and morphology* of IISD-ELA study lakes 227, 223, 224, and 442.

Lake	Coordinates	Lake type	Manipulation	Area (ha)	Max depth (m)	Water renewal time (yr)	Source
227	49°31' N 93°41' W	Experimental — Eutrophied	Years of experimentation: 1969-1974 N:P[a] loading ratio: 12:1 TN:TP[b] Concentration (µg/L): ~825:42 Years of experimentation: 1975-1989 N:P[a] loading ratio: 4:1 TN:TP[b] Concentration (µg/L): ~1200:42 Years of experimentation: 1990-present Nutrient loading: Phosphorus only TN:TP[b] Concentration (µg/L): ~800:42	5.0	10.6	4	(Anderson et al., 1987; Schindler et al., 2008)
223	49°42' N 93°42' W	Experimental — Acidified	Years of experimentation: 1976-1983 H_2SO_4 loading: ~3400 L/year[c] pH change: 6.7-5.13	27.3	14.4	8	(Schindler et al., 1980; Cruikshank, 1984; Findlay and Kasian, 1996)
224	49°41' N 93°43' W	Reference	Year of experimentation: 1976 Radioisotope addition: ^{14}C and ^{226}Ra Year of experimentation: 1976 Radioisotope addition: ^{75}Se, ^{203}Hg, ^{137}Cs, ^{59}Fe, ^{65}Zn, ^{60}Co Year of experimentation: 1976 Radioisotope addition: ^{3}H	25.9	27.4	13.5	(Hesslein et al., 1980a; 1980b; Quay et al., 1980; Crusius and Anderson, 1995)
442	49°46' N 93°49' W	Reference	NA	16.0	17.8	11.0	(Higgins et al., 2018)

*IISD-Experimental Lakes Area (https://www.iisd.org/ela/science-data/our-data/interactive-map/)

[a]Nitrogen:Phosphorus, [b]Total Nitrogen: Total Phosphorus average concentrations in epilimnion during ice-free season, [c]Average volume added during ice-free season.

Table S3.2 Sediment cores retrieved by IISD-ELA from lakes 227, 223, 224, and 442.

Lake	Date of collection	Core length (cm)	Depth collection (m)	Collected by
227	March 20, 2018	55	10.6	Cyndy Desjardins, Paul Faford, Michael Paterson, Stephen Paterson
223	March 21, 2018	55	14.60	Cyndy Desjardins, Stephen Paterson
224	March 21, 2018	37	26.2	Cyndy Desjardins, Stephen Paterson
442	March 22, 2018	55	17.2	Cyndy Desjardins, Paul Faford, Stephen Paterson

Table S3.3 The range in water content (%) in top and bottom subsection intervals of sediment cores collected from IISD-ELA Lake 227, 223, 224, and 442.

Lake	Core Depth (cm)	Range in water content (%)
227	Top (1-15.25)	96.95 – 98.74
	Bottom (16.25-51.75)	95.58 – 96.75
223	Top (1-20.25)	94.85 – 98.79
	Bottom (21.75-41.75)	93.48 – 96.04
224	Top (1-7.25)	97.55 – 98.76
	Bottom (16.25-32.25)	90.15 – 96.75
442	Top (1-12.25)	95.19 – 99.24
	Bottom (13.25-39.75)	77.41 – 96.02

Table S3.4 Core depths and associated years from top and bottom sediments from IISD-ELA lakes 227, 223, 224, and 442, based on the ^{210}Pb and ^{137}Cs radioisotope dating.

	Top sediments (post-manipulation*)		**Bottom sediments (pre-manipulation*)**	
Lake	Depth (cm)	Years	Depth (cm)	Years
L227	0 – 15.25	1968 – 2018	16.25 – 51.75	Pre-1850 – 1967
L223	0 – 20.25	1976 – 2018	21.75 – 41.75	Pre-1850 – 1973
L224	0 – 7.25	2004 – 2018	16.25 – 32.25	Pre-1850 – 1985
L442	0 – 12.25	1990 – 2018	13.25 – 39.75	Pre-1850 – 1989

*For lakes 227 and 223

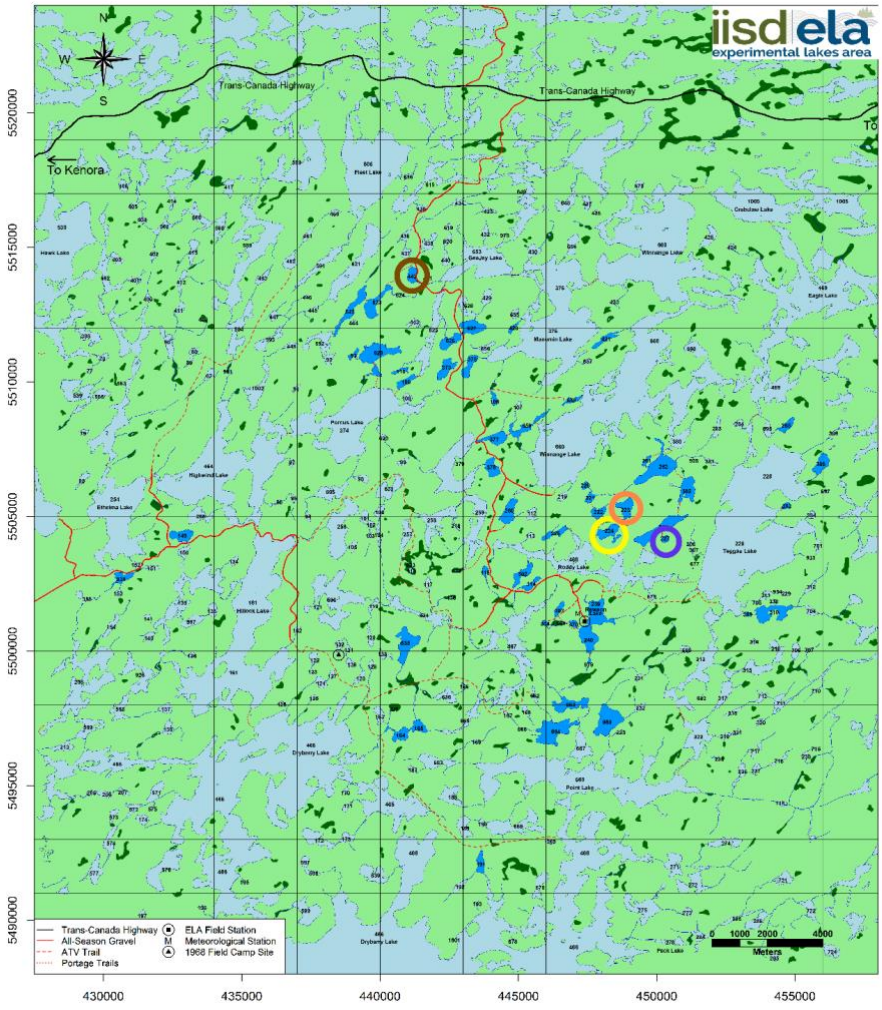

Figure S3.1 Map of the IISD Experimental Lakes Area. Circled sites correspond to study lakes 227 (purple), 223 (orange), 224 (yellow), and 442 (brown). Map from IISD Experimental Lakes Area (iisd.org/ela).

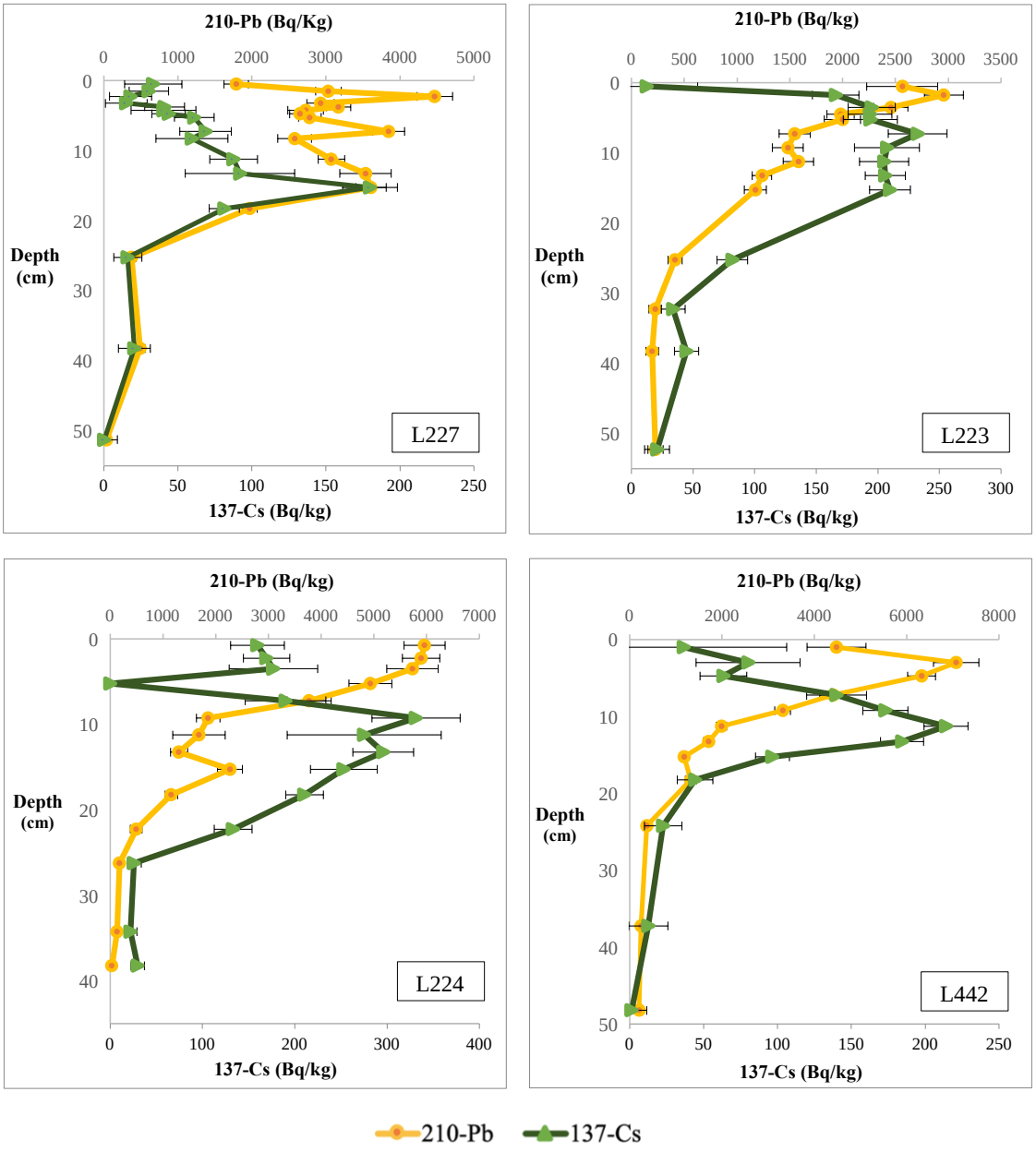

Figure S3.2 Depth profiles of ^{210}Pb and ^{137}Cs data in IISD-ELA lakes 227, 223, 224, and 442 sediment cores.

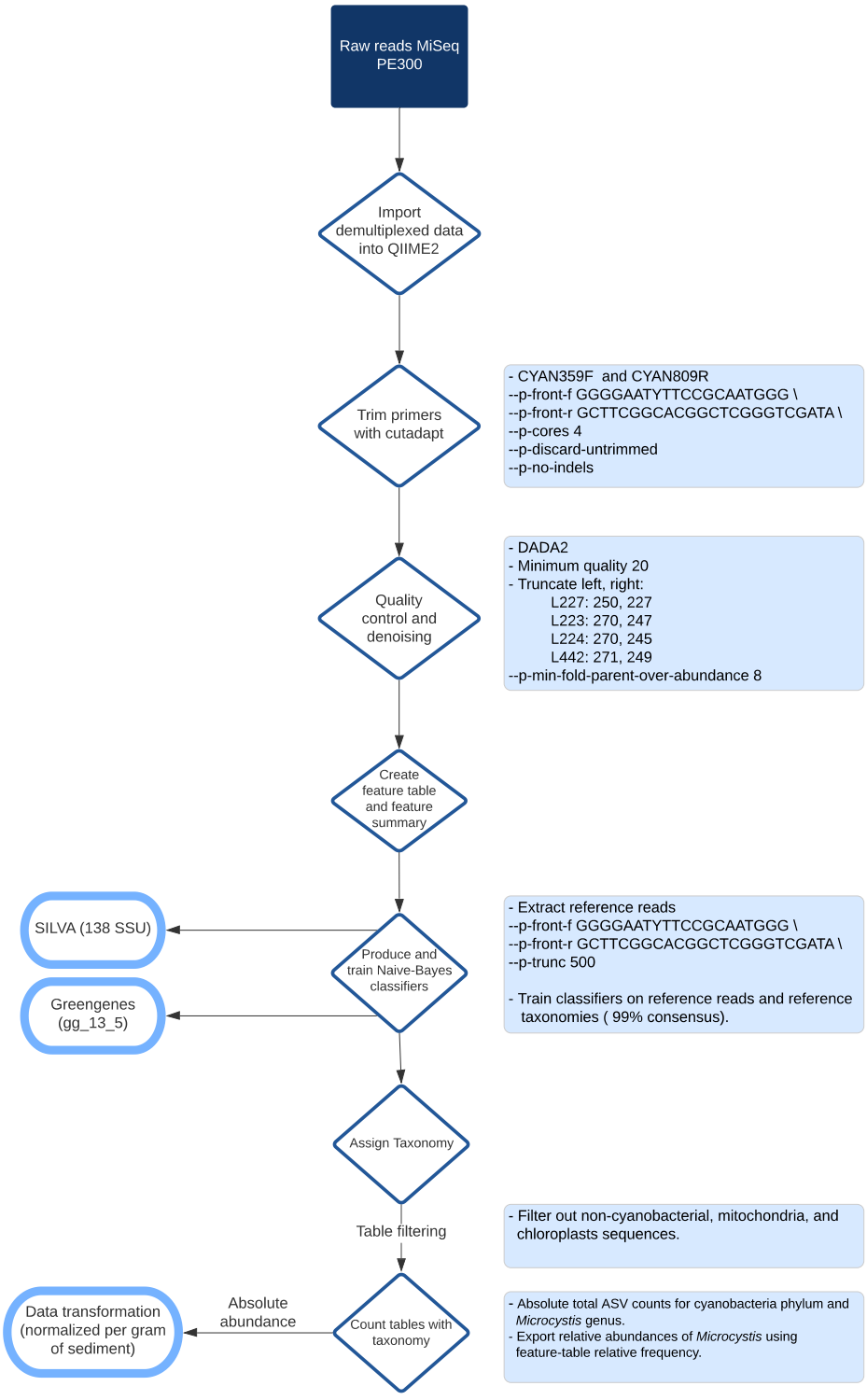

Figure S3.3 Summary of the main steps applied for the analysis of the sequencing data.

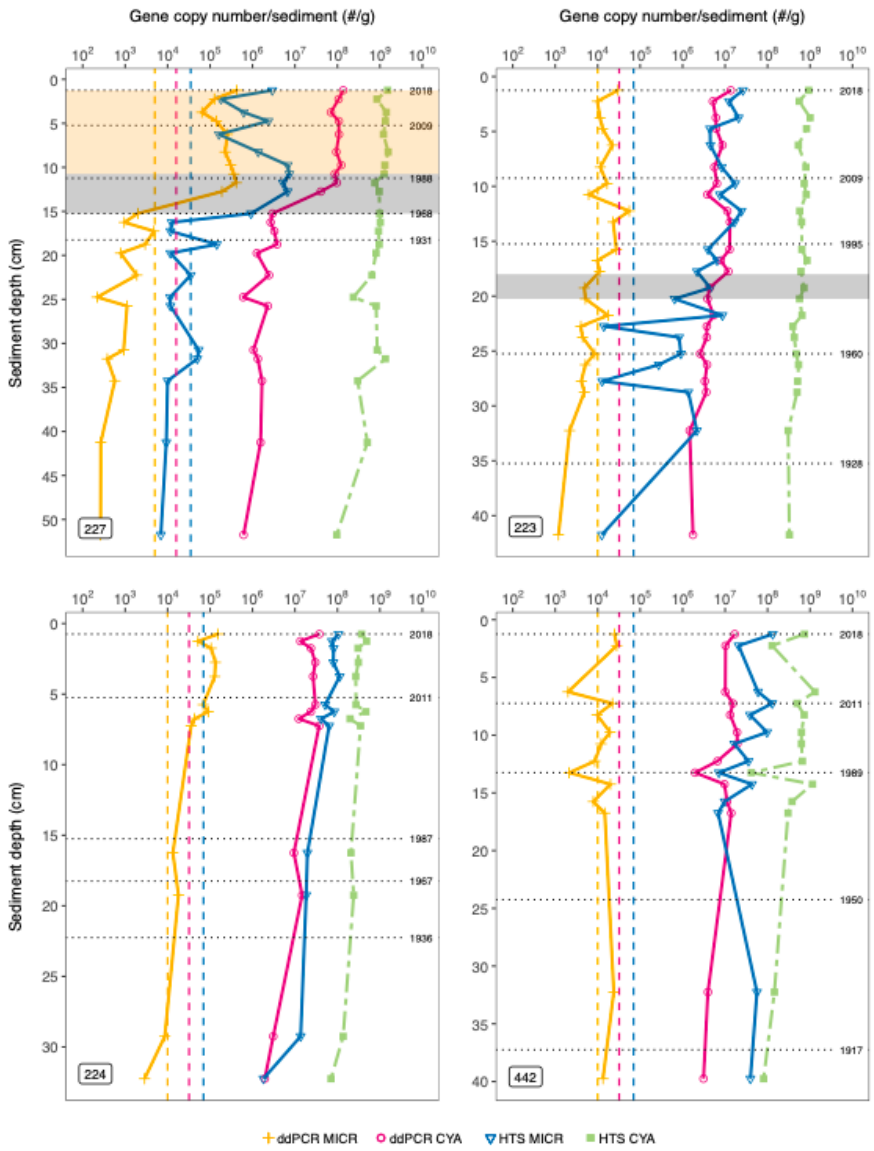

Figure S3.4 Target gene concentrations in sediment cores of the Experimental Lakes Area, Canada from manipulated lakes 227 and 223 and reference lakes 224 and 442. Absolute gene copy numbers of *Microcystis* 16S rRNA (MICR) and cyanobacterial 16S rRNA (CYA) as determined through droplet digital PCR (ddPCR) are shown, along with the amplicon sequence variant counts of cyanobacterial 16S rRNA and the *Microcystis* genus from high-throughput sequencing (HTS) (using SILVA). Results are normalized per gram of wet sediment and presented on a logarithmic scale. The orange and pink dashed vertical lines are the detection limits (LODs) for ddPCR-derived *Microcystis* 16s rRNA and cyanobacterial 16S rRNA abundances, respectively. The blue dashed vertical line is the LOD of HTS-derived abundances for both cyanobacterial 16S rRNA and *Microcystis*. The grey shading in the top left plot corresponds to the period of phosphorus and nitrogen loading in Lake 227 (1969-1989), while the beige shading corresponds to the period of phosphorus loading only (1990-2018). The grey shading in the top right plot corresponds to the period of sulfuric acid loading in Lake 223 (1976-1983).

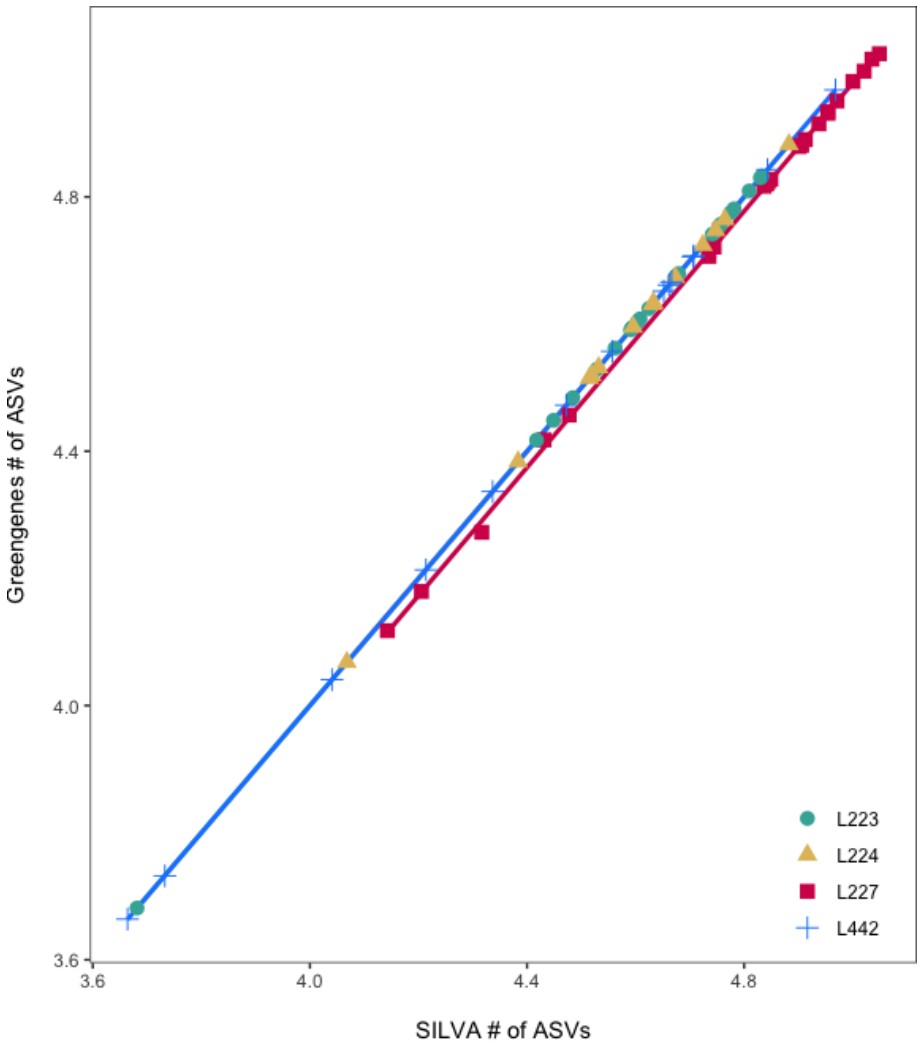

Figure S3.5 Correlations between SILVA and Greengenes log-transformed amplicon sequence variant (ASV) counts of the cyanobacteria phylum. Pearson's $r = 0.98 – 0.99$, $p < 0.001$, for IISD-ELA lakes 227 (L227), 223 (L223), 224 (L224), and 442 (L442).

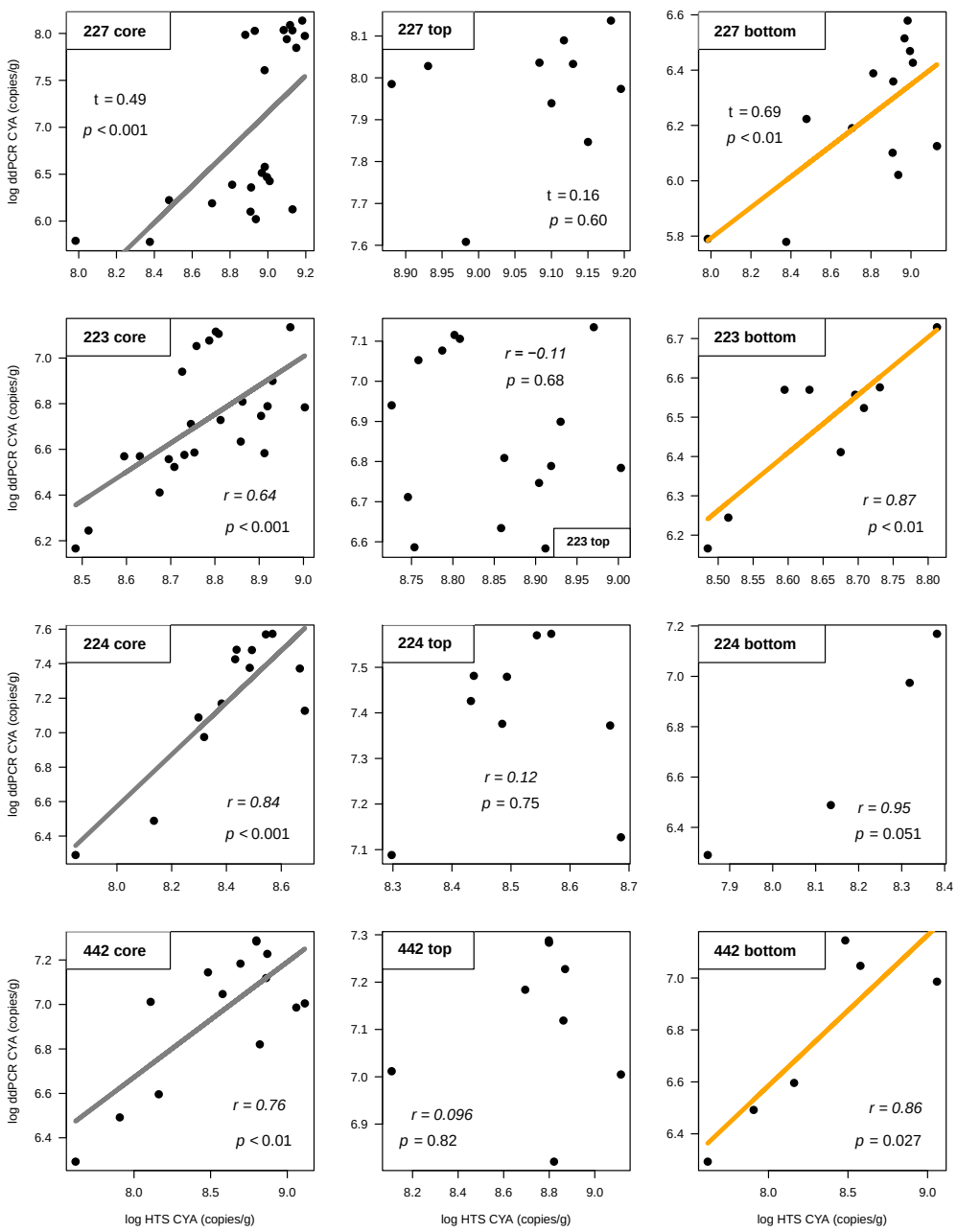

Figure S3.6 Correlations between log-transformed droplet digital PCR (ddPCR) gene copy numbers and high-throughput sequencing (HTS) amplicon sequence variant counts (using SILVA) for cyanobacterial 16S rRNA (CYA), normalized per gram of wet sediment across the whole core (left), top sediments (middle), and bottom sediments (right) of study lakes. Kendall's τ (Lake 227; top row) and Pearson's r (lakes 223, 224, and 442) correlation coefficient values are shown.

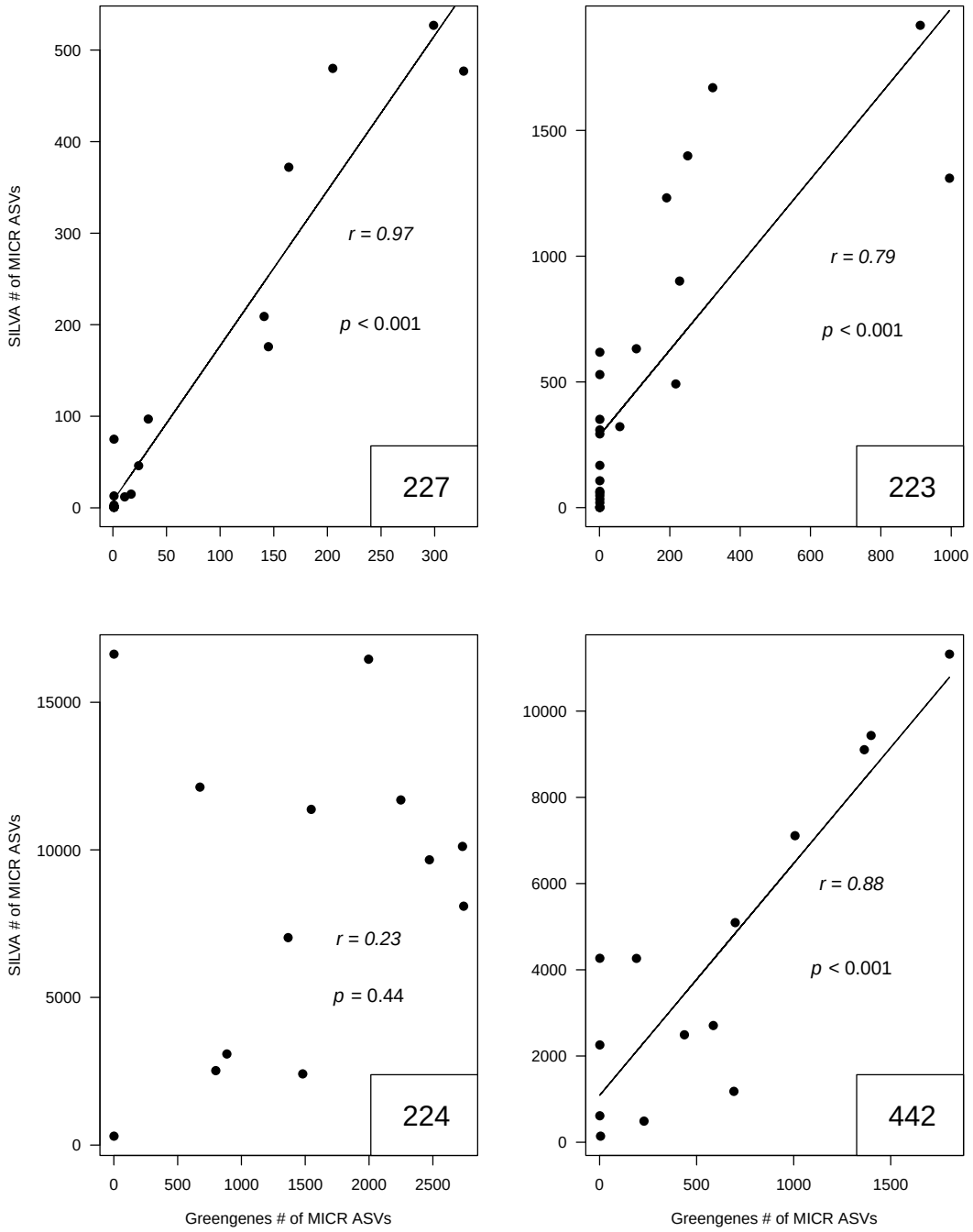

Figure S3.7 Correlations between SILVA and Greengenes amplicon sequence variant (ASV) counts for *Microcystis* (MICR). Pearson's *r* values are shown.

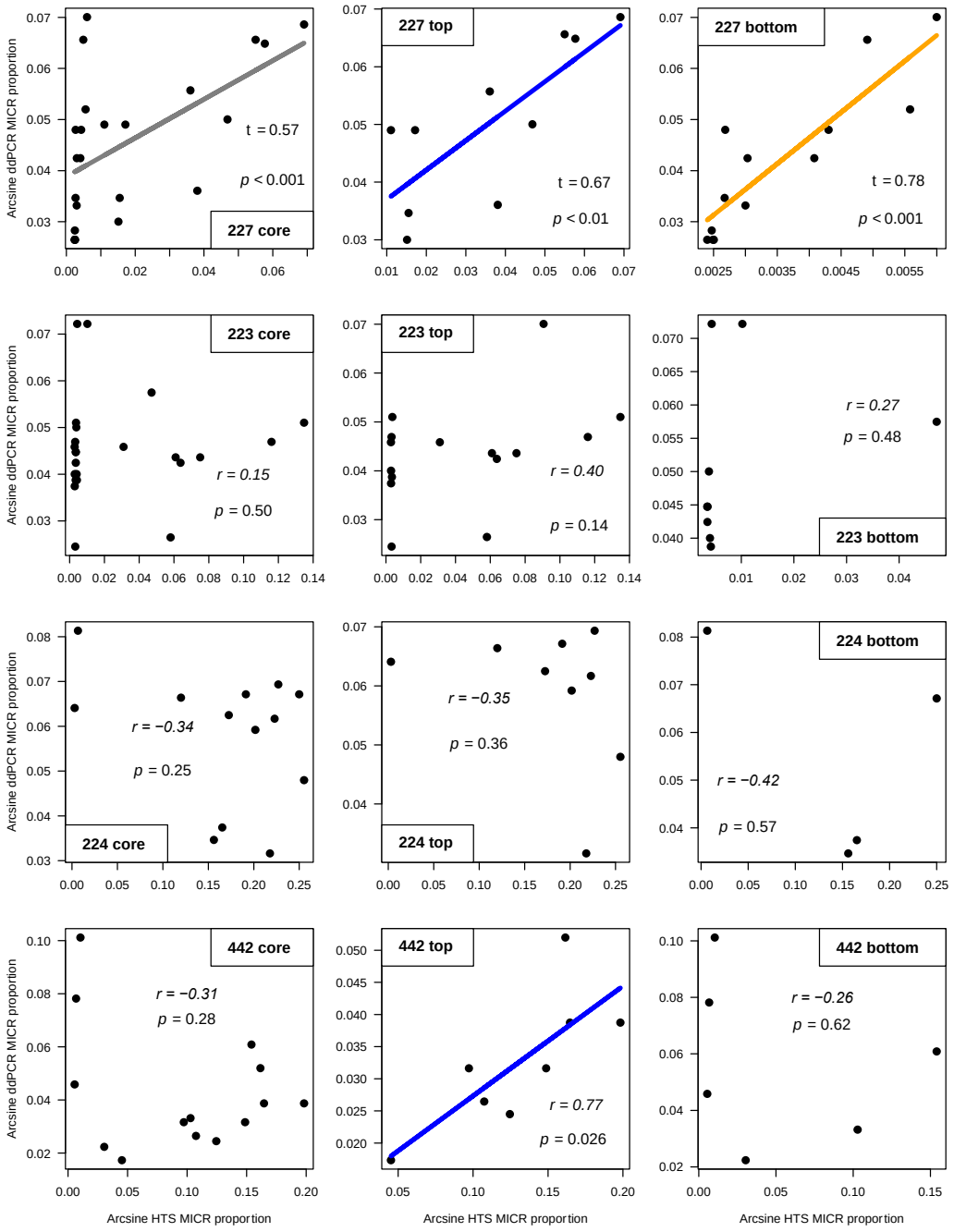

Figure S3.8 Correlations between arcsine square root transformed droplet digital PCR (ddPCR) and high-throughput sequencing (HTS) (using SILVA) relative abundance outputs for *Microcystis* (MICR) across the whole core (left), top sediments (middle), and bottom sediments (right) of study lakes. Kendall's τ (Lake 227; top row) and Pearson's r (lakes 223, 224, and 442) correlation coefficient values are shown.

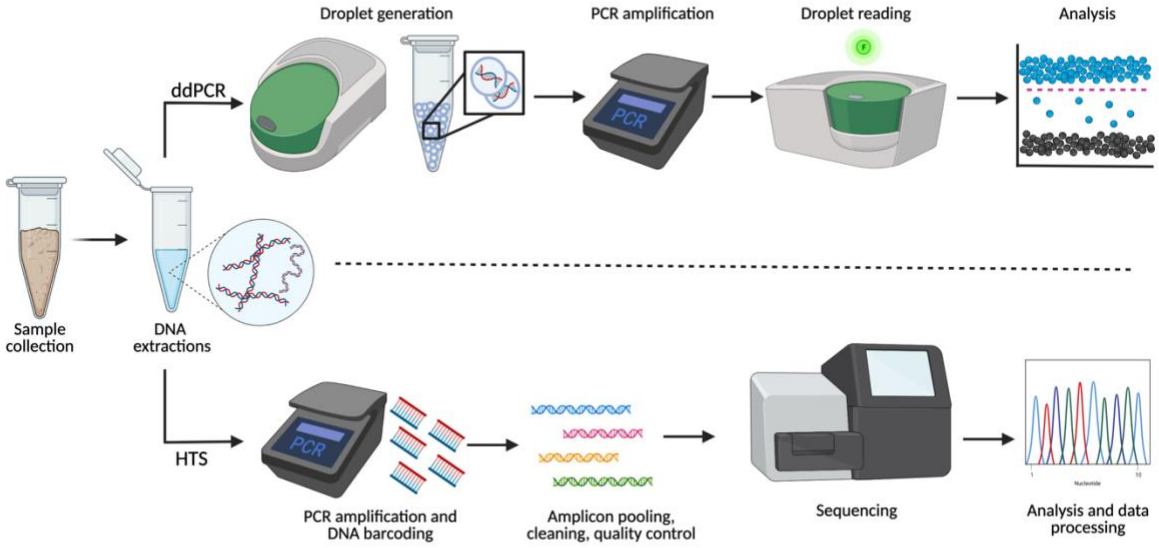

Figure S3.9 The droplet digital polymerase chain reaction (ddPCR) process in comparison to the general process of high-throughput sequencing (HTS). In ddPCR, samples are transferred into the droplet generation machine where the DNA samples are partitioned into thousands of droplets. Samples are then transferred into the PCR machine for target gene amplification. The plate is subsequently transferred into the droplet reader where the intensity of the fluorescence in each droplet is measured. Positive and negative droplets, based on the fluorescence signal and determined using QuantaSoft software, are then separated and concentrations are analyzed. In HTS, a targeted PCR is performed, followed by a second PCR for barcoding, which adds an index (or barcode) to each sample. After the amplicon pool (or library) has been cleaned, the pool is then sequenced (here, an Illumina MiSeq was used), and results are interpreted using bioinformatics analyses. Figure created through BioRender.

APPENDIX C: Supplementary material for Chapter 4

Table S4.1 The primer sequences used in droplet digital PCR (ddPCR) and high-throughput sequencing (HTS), as well as the amplicon sizes, annealing temperatures, and dilution factors of target genes cyanobacterial 16S rRNA (CYA), *Microcystis* 16S rRNA (MICR), microcystin E (*mcyE*), and glutamine synthetase (*glnA*).

Target gene	Primer	Primer sequence (5'-3')	Amplicon size (bp)	Annealing T (°C)	Dilution factor
Microcystis 16S rRNA (ddPCR)	MICR F[1]	GCC GCR AGG TGA AAM CTA A	247	56	10x
	MICR R[1]	AAT CCA AAR ACC TTC CTC CC			
Cyanobacteria 16S rRNA (ddPCR)	CYA 108F[1,2,]	ACG GGT GAG TAA CRC GTR A	269	56	32x
	CYA 377R[1.3]	CCA TGG CGG AAA ATT CCC C			
Microcystin E (*mcyE*)	*mcyE* F	TAA CTT TTT TGG GCA TAG TCC TG	186	59.5	10x
	mcyE R	CGA ACW GCY GCC ATA ATC GC			
Glutamine synthetase (*glnA*)	*glnA* F[4]	GAT GCC GCC GAT GTA GTA	156	59.5	100x
	glnA R[4]	AAG ACC GCG ACC TTY ATG CC			
Cyanobacteria 16S rRNA (HTS)	CYA 359F[3,5]	GGG GAA TYT TCC GCA ATG GG	450	52	10x, 50x
	CYA 809R[3,5]	GCT TCG GCA CGG CTC GGG TCG ATA			

[1](Rinta-Kanto et al., 2005), [2](Urbach et al., 1992), [3](Nubel et al., 1997), 4(Hurt et al., 2001), [5](Jungblut et al., 2005).

Table S4.2 The droplet digital PCR limits of detection (LOD) for each target gene as copies/μL and when normalized as copies/g of dry sediment.

	Initial LOD (copies/μL)	LOD after normalization (copies/g)	
Target gene	For all lakes	Lake 227*	Lakes 223, 224, and 442
Cyanobacterial 16S rRNA	1.43	~4.1×10^5	~8.2×10^5
Microcystis 16S rRNA	1.55	~1.4×10^5	~2.8×10^5
Microcystin E	1.10	~9.8×10^4	~1.9×10^5
Glutamine synthetase	1.00	~8.9×10^5	~1.8×10^6

*The limits of detection for the genes in this lake were slightly lower due to some alterations within the DNA extraction protocol when using Lake 227 sediments. For L227, a greater sediment weight and lower elution volume was used (compared to the average volumes used in the other lakes). See Mejbel et al., 2021b for details regarding DNA extraction protocol modifications.

Table S4.3 Years with available water chemistry (from the epilimnion) and phytoplankton data (from the epilimnion and metalimnion) from the Experimental Lakes Area, lakes 227, 223, 224, and 442. Records obtained from Scott Higgins, IISD-ELA.

Lake	Phytoplankton data	Water chemistry data
227	1970 – 2017	1969 – 2017
223	1974 – 2004, 2007 – 2008, 2015 – 2016	1974 – 2004, 2007 – 2008, 2012, 2015 – 2017
224	1974 – 2017	1974 – 2017
442	1990 – 2017	1994 – 2017

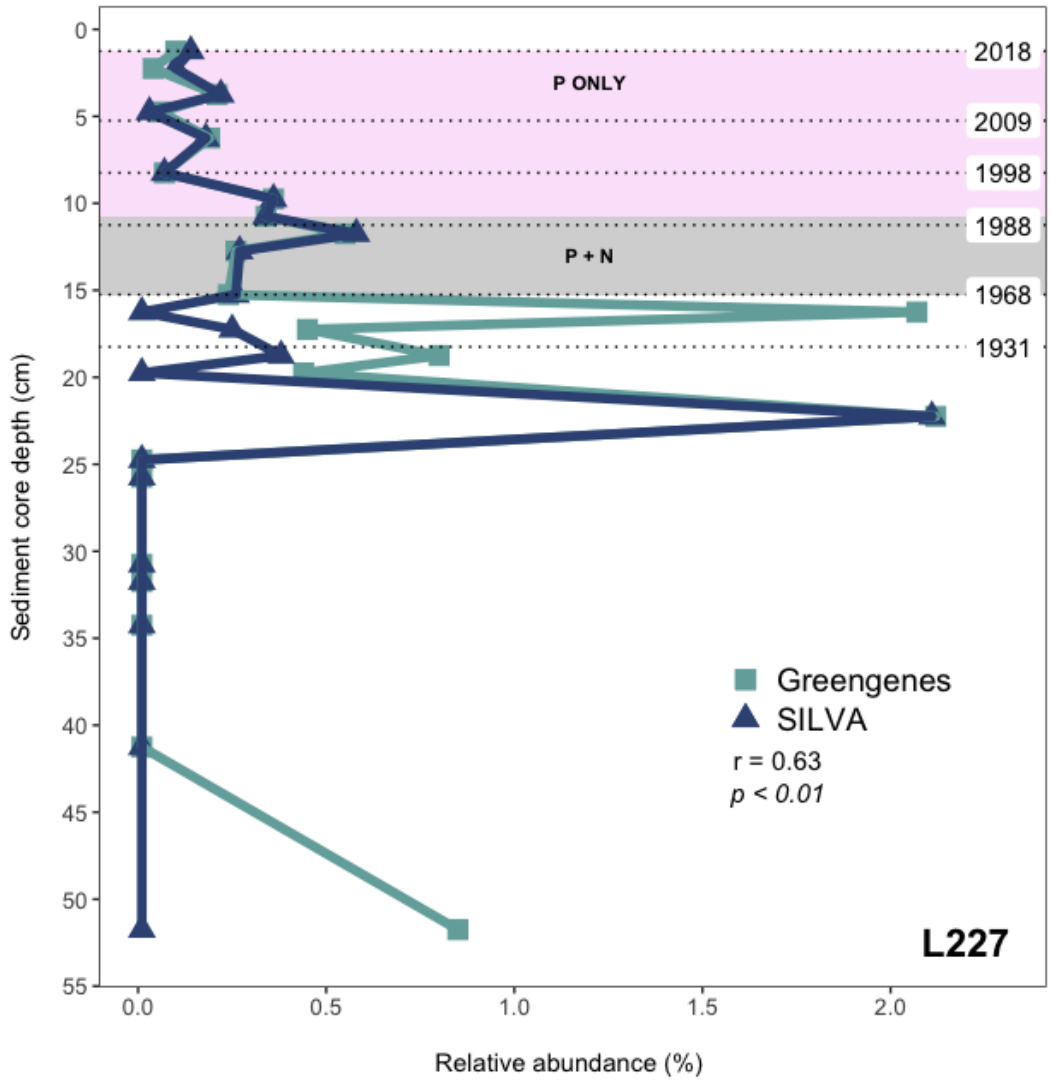

Figure S4.1 The relative abundance of the *Synechococcales* order in the sediment core of Lake 227 (L227), Experimental Lakes Area, Canada, as determined through high-throughput sequencing of the cyanobacterial 16S rRNA gene and using both the Greengenes and SILVA reference gene databases. The grey shading, labelled P + N, corresponds to the period of phosphorus and nitrogen loading in the lake between 1969 – 1989. The purple shading, labelled P only, corresponds to the period of phosphorus loading only from 1990 – 2018.

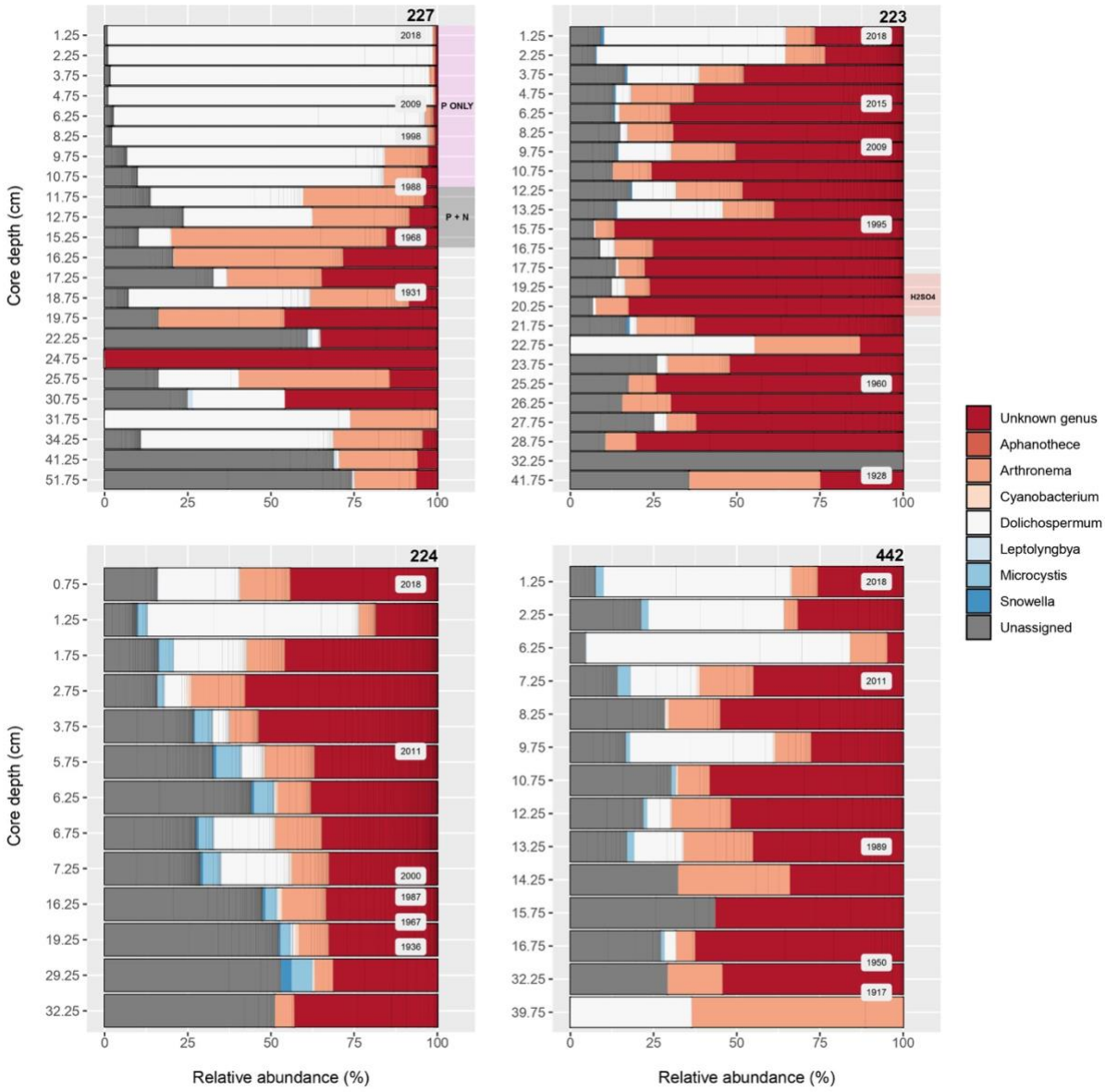

Figure S4.2 Cyanobacterial community composition in lake sediment cores of the Experimental Lakes Area, Canada, from manipulated lakes 227 (top left) and 223 (top right), along with reference lakes 224 (bottom left) and 442 (bottom right), as determined through high-throughput sequencing of the cyanobacterial 16S rRNA gene and using the Greengenes reference gene database. The relative abundance of each major taxa is presented at the level of genus. The grey shading on the right of the L227 plot (top left), labelled P + N, corresponds to the period of phosphorus and nitrogen loading in the lake between 1969 – 1989. The purple shading in the same plot, labelled P only, corresponds to the period of phosphorus loading only from 1990 – 2018. The pink shading on the right of the L223 plot (top right), labelled H_2SO_4, corresponds to the period of sulfuric acid loading in the lake between 1976 – 1983.

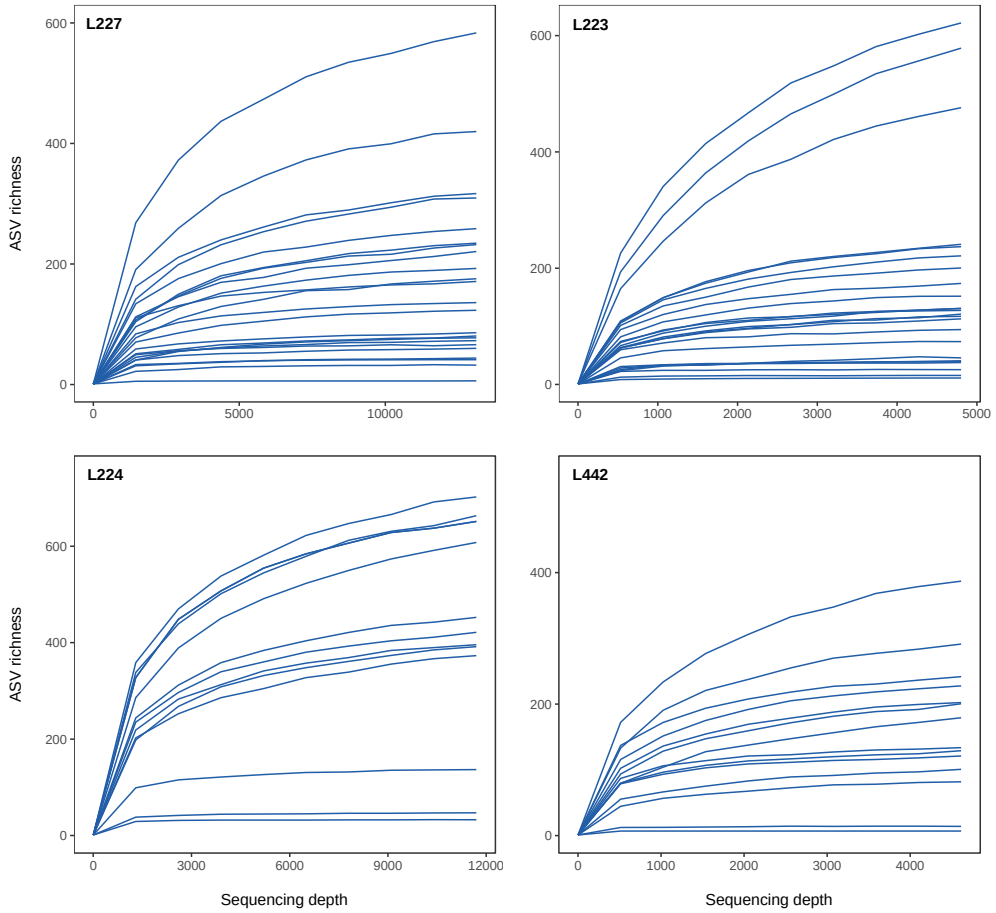

Figure S4.3 Rarefaction curves for all samples in Lake 227 (L227; top left), Lake 223 (L223; top right), Lake 224 (L224; bottom left), and Lake 442 (L442; bottom right) of the Experimental Lakes Area, Canada. The lines are representative of the average diversity values computed for a sample based on 10 rarefied tables generated at each sampling depth.

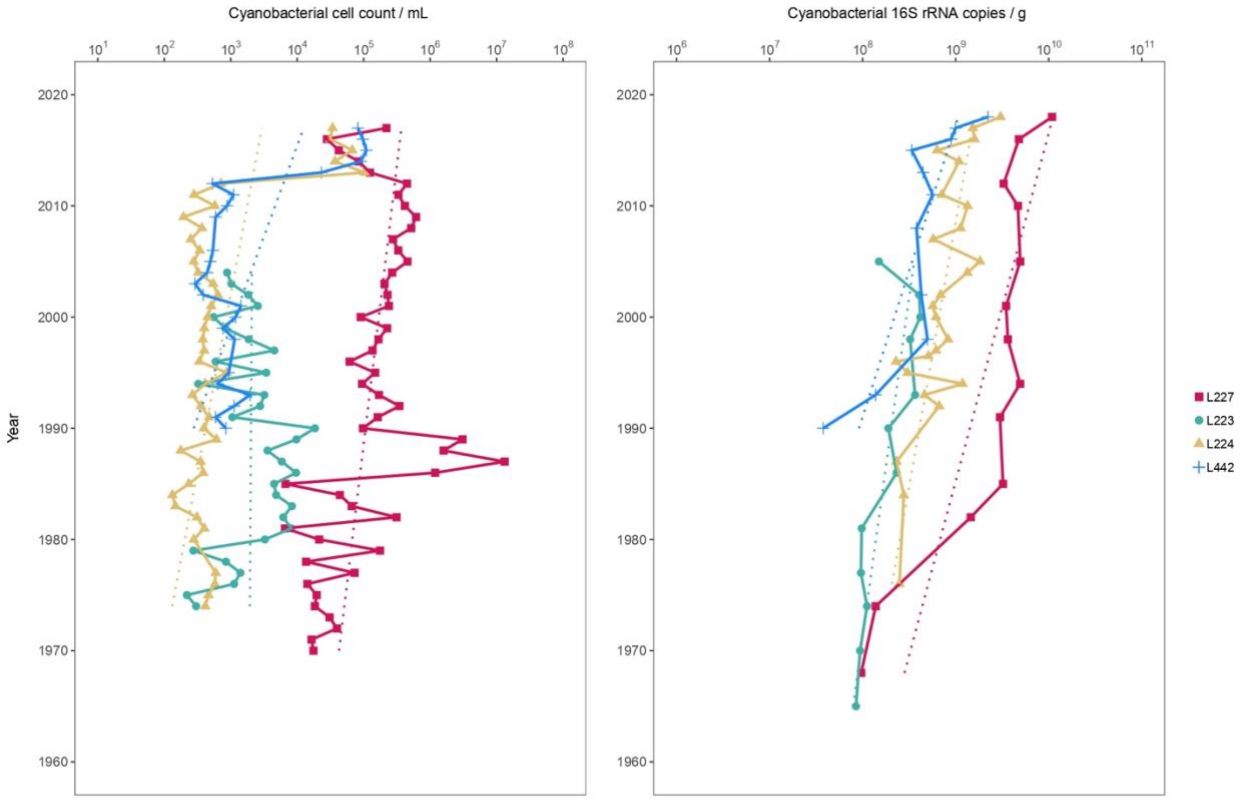

Figure S4.4 Comparison between the cyanobacterial cell density (cell count/mL) from the water (left) with the absolute abundance of the cyanobacterial 16S rRNA gene (copies/g) from the sediment (right) of Lake 227 (L227), Lake 223 (L223), Lake 224 (L224), and Lake 442 (Lake 442) from the Experimental Lakes Area. Annual cyanobacterial abundance from the water was calculated using the sum of the monthly averages (May-September) of cyanobacterial cell density in the epilimnion and metalimnion. Water data for L227 corresponded to the periods of 1969-2017, for L223: 1974-2004, for L224: 1974-2017, and L442: 1990-2017. The sediment plot on the right shows the samples that correspond to relatively similar timeframes. Note the x-axis scales at the top are not the same between the two plots.

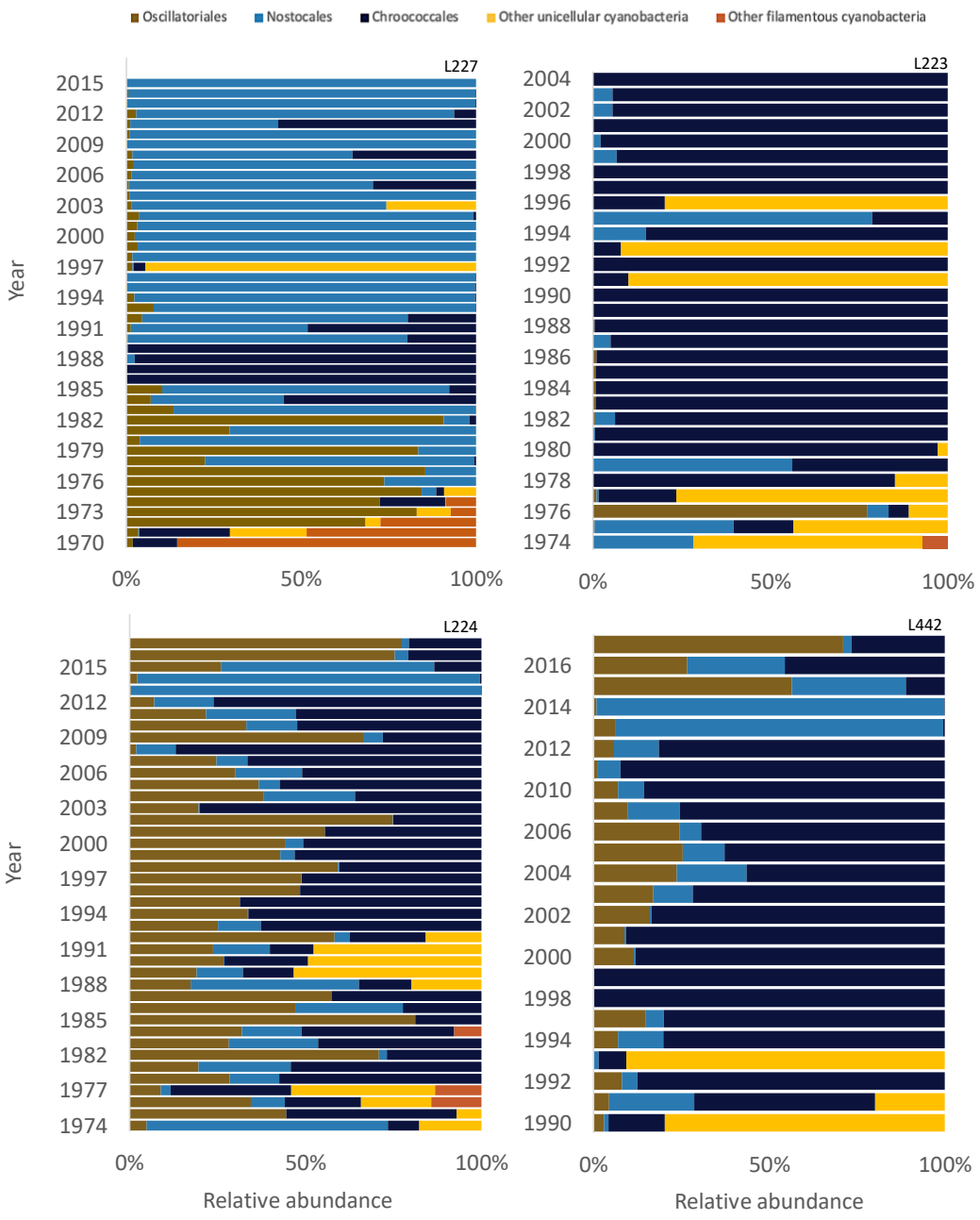

Figure S4.5 Cyanobacterial community composition from the epilimnion and metalimnion surface water records of manipulated lakes 227 (top left) and 223 (top right), along with reference lakes 224 (bottom left) and 442 (bottom right) from the Experimental Lakes Area, Ontario, Canada. The relative abundance of each major taxa at the level of order and of the unidentified unicellular and filamentous cyanobacteria is presented. Phytoplankton records provided by Scott Higgins, IISD-ELA.