Temporal Trends in Cyanobacteria through Paleo-Genetic Analyses

William Dodsworth

BSc. University of Guelph 2015

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies

University of Ottawa

Partial fulfillment of the requirements for:

Masters of Science – Biology General degree

Department of Biology

Faculty of Science

University of Ottawa

uOttawa

Abstract

With increasing eutrophication and climate change, temperate lakes are experiencing conditions favoring cyanobacterial dominance, leading to toxic blooms. However, the relative importance of these two factors remains unclear, due to a lack of historical records. This thesis analyzed sediment DNA from four lakes in Central Ontario to quantify trends over the past ~ 200 years in bacteria, cyanobacteria, and microcystin toxins through ddPCR of target genes. Climate related variables explained a small amount of variation. However, lakes with more development exhibited significant increases in cyanobacterial dominance, indicating regime shifts not occurring in less developed lakes. These shifts were likely driven by nutrient loading in concert with climate change. Sediment cores were also compared within a single multi-basin lake to determine the effect of depth and morphometry on DNA preservation. Sites of greater depth (20 m +) were more conducive to long-term DNA preservation and sheltered morphometry increased sediment DNA deposition.

Résumé

Avec l'eutrophisation soutenue et le changement climatique, les lacs tempérés connaissent des conditions favorables à la croissance des cyanobactéries, entraînant des proliférations potentiellement toxiques. Cependant, l'importance relative de ces deux facteurs reste incertaine, en raison du manque de données historiques. Cette thèse a analysé l'ADN des sédiments de quatre lacs du centre de l'Ontario afin de quantifier les tendances observées au cours des quelques 200 dernières années en ce qui concerne les bactéries, les cyanobactéries et les cyanotoxines (microcystines) par le biais de la ddPCR des gènes cibles. Les variables liées au climat ont expliqué une faible quantité de la variation. Cependant, la dominance des cyanobactéries dans les lacs plus développés a néanmoins fortement augmenté, et un tel changement de régime écosystémique ne s'est pas produit dans les lacs moins développés. Cette tendance a probablement été provoqué par une augmentation de la charge en éléments nutritifs associée au changement climatique. Les carottes de sédiments ont également été comparées dans un même lac comportant plusieurs bassins afin de déterminer l'effet de la profondeur et de la morphométrie sur la préservation de l'ADN. Les sites de plus grande profondeur (20 m +) étaient plus propices à la conservation à long terme de l'ADN et une morphométrie plus protégée augmentait le dépôt d'ADN dans les sédiments.

Acknowledgements:

My first acknowledgement is to my supervisor, Dr. Frances Pick. You took a chance on me, offered me guidance and a fascinating introduction to the world of research. I would like to thank you for the opportunities, wisdom and support you have offered me in my time here.

To Sofia Perrin and Phil Pelletier. The technologists of our core molecular floor. With answers and solutions for every problem I could stumble into I would like to thank you for your patience and all the help you have provided me throughout my project.

To my lab mates. Mary Ann Perron, Hebah Mejbel and Amber Dyck. As well as the summer students and technicians: Bella Richmond, Chloe Van de Panne, Meaghan MacIntyre-Newell and Jesse Fortier. Thank you all for the good memories, good times and for the role you each played in the completion of my thesis.

I would like to extend my thanks to my committee members Alexandre Poulain and Jesse Vermaire, and their respective labs, as well as my funding sources: NSERC CREATE-ABATE, Ontario Ministry of Environment, Conservation and Parks.

Thank you to my family and friends. For supporting my dreams, pushing me to face new challenges, listening to me ramble about my project, complain about setbacks and help me appreciate and celebrate the small (and big) victories in life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

apcA	**Allophycocyanin Alpha – Chain gene**
BP	**Breakpoint Analysis**
CAR	**Continuous time autoregressive model, lag 1**
CDD	**Cooling Degree Days (Degrees Celsius 18>)**
Chl a	**Chlorophyll a**
Cs^{137}	**Cesium 137 – Radio-Isotope**
CYA	**Cyanobacterial 16s rRNA gene (rRNA)**
ddPCR	**Droplet digital Polymerase Chain Reaction**
ELISA	**Enzyme linked Immuno-adsorbent assay**
GAM	**Generalized Additive Model**
GCN	**Gene Copy Number**
GE	**Gel - Electrophoresis**
***gln*A**	**Glutamine Synthetase A gene (Housekeeping)**
HAB	**Harmful Algal Bloom**
HDD	**Heating Degree Days (Degrees Celsius > 18)**
MC	**Microcystin**
***mcy*E**	**Microcystin E gene (protein sub-unit – ADDA)**
MICR	***Microcystis*-specific 16s rRNA gene (rRNA)**
NTC	**No Template Control**
P	**Phosphorus**
Pb^{210}	**Lead, Elemental - 210 isotope**
qPCR	**Quantitative Polymerase Chain Reaction**
TAE	**Tris-Acetic Acid-EDTA Buffer**
TE	**Tris EDTA Buffer**
TP	**Total Phosphorus**
TN	**Total Nitrogen**

Chapter 1: General Introductions

1.1 Cyanobacteria – Evolutionary History, Physiological and Ecological Traits

Cyanobacteria, previously known as blue-green algae, are photosynthetic prokaryotes occupying a primary production position at the base of aquatic ecosystems (Graham & Wilcox 2000; Whitton & Potts 2012; Pentecoaste & Whitton 2013; Hamilton *et al.* 2016). Species come in a variety of cellular and colonial shapes with sizes ranging from half a micron all the way up to several millimeters in diameter (Whitton & Potts 2012; Srivastava *et al.* 2012; Zhang *et al.* 2018). Similar to modern day eukaryotic plants, cyanobacteria make use of photosynthetic pigments and carbon dioxide to capture light and create complex organic compounds while releasing oxygen to the atmosphere (Jeltsch 2013; Hamilton *et al.* 2016). This self-sufficiency in energy capture and production has been a factor in the ubiquitous and global distribution of cyanobacteria *(*Garcia-Pichel *et al.* 2003*;* Rindi 2007*;* Wood *et al.* 2008).

There is evidence of cyanobacterial proliferation extending back beyond fossilized stromatolites that have an estimated age greater than 2.5 billion years (Bekker *et al.* 1980; Bekker *et al.* 2004; Jeltsch, 2013; Cardona *et al*. 2019). Cyanobacteria are also thought to be the driving force behind "The Great Oxygenation Event" as they were among the first organisms to produce oxygen as a by-product of photosynthetic processes (Kopp *et al.* 2005; Hamilton *et al.* 2016; Garcia-Pichel *et al*. 2019). This event was a turning point in the composition of the atmosphere. This release of oxygen saturated oxidative agents, such as iron, eventually leading to a build-up of atmospheric oxygen (Blankenship 2002; Kopp *et al.* 2005). With this came a shift in global climate, from the warmer conditions that cyanobacteria favor and evolved alongside, to

a more moderate state as greenhouse gases became oxidized and removed from the atmosphere (Kopp *et al.* 2005; Allen & Martin 2007; Garcia-Pichel *et al.* 2019). This proved to be the catalyst for the evolution of larger eukaryotic organisms, which make use of oxygen to power more complex life forms (Stamati *et al.* 2011; van der Giezen, 2011; Rockwell & Lagarias 2014).

Many planktonic cyanobacteria have an optimal growth temperature that is slightly greater than other phytoplankton that inhabit similar parts of the water column (Kosten *et al* 2012; Jankowiak *et al.* 2019). This trait was likely developed during the period before oxygenation when the earth was significantly warmer. This allows survival under many climates particularly tropical, sub-tropical freshwater and marine environments (Whitton & Potts 2012; Paerl & Paul 2012; Jankowiak *et al.* 2019). This also presents an interesting, potential, pre-adaptation to the current trend of increasing global temperature. Some cyanobacteria carry key genetic components that allow them to fix atmospheric nitrogen (N_2) during periods of nutrient deficiency (Berman-Frank *et al.* 2003., Carey *et al.* 2012., O'Neil *et al.* 2012). However, cyanobacteria have also been found to be highly responsive to nutrient enrichment from phosphorus, nitrogen and even iron (Stolte *et al*. 1994; Stolte & Riegman 1995; Roelke *et al.* 1999; Carey *et al*. 2012). Thus, cyanobacteria can survive in conditions of both low and high nutrient availability (Dolman *et al*. 2012; Bonilla & Pick 2017; Jankowiak *et al.* 2019). Some cyanobacteria are also capable of rudimentary forms of movement. This is not in the form of flagella or other motor extrusions but rather comes in the form of small gas filled vesicles produced within the cells (Reynolds *et al.* 1987; Whitton & Potts 2012). These are hollow, rigid, cylindrical structures that are created and deconstructed throughout the day, allowing

cyanobacteria to adjust their position within the water column (Bormans & Sherman 1999; Whitton & Potts, 2012). This allows cyanobacteria to maximize light exposure for photosynthesis (Bormans & Sherman 1999; Whitton & Potts 2012). This, in combination with a high tolerance to saline environments, allows cyanobacteria to inhabit fresh, brackish and even saltwater aquatic conditions (Paerl & Paul 2012). Many cyanobacteria also employ a variety of unique resting stages that confer resistance to adverse environmental conditions too harsh for eukaryotes (Paerl & Paul, 2012; Carey *et al.* 2012).

1.2 Harmful Algal Blooms (HAB's) – Cyanobacterial Proliferations

An algal bloom refers to an increase in abundance of certain species with biomass accumulation within the water column. Frequently, population increases coincide with increased availability of key nutrients (Stolte *et al*. 1994; Stolte & Riegman 1995; Roelke *et al*. 1999; Carey *et al.* 2012) as well as shifts in weather conditions (Wagner & Adrian 2009; Markensten *et al.* 2010; O'Neil *et al.* 2012; Paerl & Paul 2012; Elliot 2012). Cyanobacteria are particularly responsive to increased availability of nitrogen and phosphorus as well as increased temperature (Konopka & Brock 1978; Paerl & Huisman 2008; Paerl & Paul 2012; Dolman *et al.* 2012). This, in combination with their mobility within the water column, allows them to concentrate at the surface of the water column as a highly visible phenomenon (Paerl 2014). Although these blooms are a natural occurrence within aquatic ecosystems (Heisler *et al.* 2008; Paerl & Huisman 2008; Paerl, 2008; O'Neil *et al.* 2012), they have become more pronounced in recent years with more frequent and larger blooms capable of producing toxins (Paerl & Paul 2012; Paerl, 2014). The scientific community by consensus in the early 2000's has given these blooms the moniker Harmful Algal Blooms (HAB's) to designate excessive accumulations that have negative effects on the environment and may include the production of a variety of toxins. Among the range of

negative effects, the most prominent are the depletion of dissolved oxygen levels in the water column, fish kills, reduction of phytoplankton diversity in the water column through competition and the production of a variety of toxic compounds (Carmichael 1994; Carmichael 2001; Apeldoorn & Egmond 2007; Catherine *et al.* 2013). These have led to increased health risks to wildlife as well as human populations in freshwater and in marine coastal areas (Carmichael 1994; Carmichael 2001; Apeldoorn *et al.* 2007; Catherine *et al.* 2013).

1.3 Cyanotoxins – Toxin Production in Cyanobacteria

Cyanobacteria, across many genera (including: *Microcystis, Anabaena, Plankothrix, Nodularia, Dolichospermum* and *Aphanizomenonaceae),* can produce a variety of toxic compounds either, within their cytoplasm, intracellularly or by extracellular release directly into the surrounding environment (Francis 1878; Sivonen 2009). These toxins are collectively referred to as cyanotoxins and can be produced in large enough concentrations that they can produce symptoms of acute toxicity to wildlife (Stewart *et al.* 2008; Miller *et al.* 2010) as well as humans (Pouria *et al.* 1998, Codd *et al.* 2005).

Cyanotoxin classification depends on the organs or systems targeted and distinguished into three main groups: hepatotoxins, neurotoxins and endotoxins (Sivonen 2009). Some of the most frequently encountered cyanotoxins in fresh and brackish waters lie within the hepatotoxic class in the form of microcystins and nodularin toxins (Sivonen 2009). These toxins were named after the species that produced them when they were first described *Microcystis* and *Nodularia* (Sivonen 2009). However, individual taxa can possess the genes required to produce multiple different toxic compounds and the same toxic compounds can be produced by a variety of taxa (Sivonen 2009). Thus, a single bloom can produce a myriad of toxic compounds, a single toxin

or even none at all. This production appears to depend mainly on environmental selection of a particular strains of cyanobacteria possessing the toxin genes. However, the true underlying biological reasons for why these toxin genes are maintained and why cyanotoxins are produced is not well understood (Neilan *et al.* 2013; Sivonen 2009).

Due to the potential threats to human health, health organizations and agencies such as the WHO and Health Canada have set a guideline for a concentration limit of 1.5 µg/L of detectable microcystins for drinking water to protect the general public, and mitigate chronic health effects (WHO, 2003). The framework and plan for testing remains a reactive rather than proactive process with few standard laboratory methods available. This often leads to sporadic testing with few standardized guidelines, often at the discretion of municipal leadership, as to when or how this testing should be carried out (Health Canada 2016).

1.4 Cyanobacteria of Temperate Lakes

Cyanobacteria are commonly encountered in the plankton of all surface waters including lakes in temperate regions. In the typical seasonal succession of phytoplankton in temperate lakes (Sommer 1985; Sommer *et al.* 2012), cyanobacteria tend to be present in very early spring but build up biomass and, eventually, can dominate phytoplankton total biomass in late summer. It is the large colonial genera such as *Microcystis*, *Dolichospermum, Aphanizomenon, Planktothrix* that are particularly prone to generating toxic blooms in temperate lakes (Bonilla & Pick 2017). The factors that control these bloom- forming toxigenic taxa have been the subject of increasing research.

Across temperate lakes, total cyanobacterial biomass has been found to be a direct function of lake nutrient concentrations, in particular those of nitrogen and phosphorus (Pick & Lean 1997; Downing *et al.* 2001; Beaulieu *et al.* 2013). As a result, the process of anthropogenic eutrophication of temperate lakes (increasing nutrient concentrations) is accompanied by a rise in cyanobacteria. More recently climate change has been implicated as an additional trigger for cyanobacterial blooms (Paerl & Huisman 2008). At increasingly northern latitudes, cyanobacterial populations appear to be responding to shifts in temperature, precipitation and wind patterns stimulating dense growth as well as the formation of visible surface blooms (Pick 2016). However, the importance of these climate factors in explaining cyanobacterial blooms relative to well-known effects of nutrient drivers has not been established and may in fact vary regionally (Pilon *et al.* 2019)

1.5 Climate Change – Effects & Cyanobacterial Abundance

With the advent of the industrial age, humans have been changing the environment on a global scale. Industrialization, rapid human expansion as well as a decided lack of environmental awareness have led to degradation of key environments within our direct influence via dumping, modification and activity, as well as indirectly via greenhouse gas wastes in air and soils as well as through water pollution. In particular, the increased use and combustion of fossil fuels, as well as deforestation have led to rising atmospheric carbon dioxide levels. This increased concentration of carbon dioxide has contributed to the "greenhouse gas effect". The "greenhouse gas effect" is the cumulative increase in density of the earth's atmosphere. With enough density, the absorbance of more solar energy results in global changes manifested as rising temperatures, massive disruptions in weather and air patterns including changes in precipitation, melting of polar ice caps and rises in sea level.

Climate change has begun to impact aquatic ecosystems and the cyanobacteria that inhabit them (Schindler 1997). Researchers have reported a correlation between increased air temperature and water temperature, reduced ice cover, as well as an increase in the duration of both the growing season (Magnuson *et al.* 1997) as well as the duration of water column stability in temperate lakes (Ford & Stefan 1980; Robertson & Ragotzkie 1990). Increases in temperature may be contributing directly and/or indirectly through changes in stratification regimes to increases in cyanobacterial blooms in temperate lakes (Pick 2016; Huisman *et al*. 2018).

1.6 Paleolimnology – History as Recorded Through Lake Sediment Analysis

Without a working history of lake conditions beyond those taken through instrument records it has been difficult to determine whether the behaviour of a lake system is normal or is a significant deviation from the historical condition before the human settlement. Paleolimnology – the study of lake sediments – can help resolve this question. It is now possible to reconstruct past events utilizing chemical, physical and biological proxies preserved in the lake's sediments (Smol 2008). As particulates within the water column descend and settle within the water column, they trap pollutants, chemicals, microorganisms and various other elements trapping them above the older layer of sediments (Smol 2008). This leads to a sequence of layers that are similar to the rings visible across the trunk of a tree, with the oldest sediments being the buried deeper and the newer layers deposited on top (Smol 2008). These trapped layers can be retrieved in the form of sediment cores and assessed qualitatively and quantitatively to reconstruct past environmental conditions (Smol 2008). Since some elements do not preserve well, analyses of fossilized organisms such as diatoms, chironomid heads and Cladocera have been utilized as proxies of past lake conditions (Smol 2008). These taxa are often selective of, and highly

sensitive to, specific aquatic conditions within the lakes (Smol 2008). For example, certain species of chironomids favor low oxygen levels such that a sudden increase in the proportion of these fossilized chironomids indicates an extended period in the environment of lower oxygen.

Once a sediment core from a region of interest has been harvested and sectioned, it can then be analyzed and each section correlated to certain time periods using radio-isotope dating. Utilizing elements with known degradation curves and background levels, such as lead (^{210}Pb) and Cesium (^{137}Cs), proxy indicators can then be associated with specific time periods throughout history (Appleby 1986). A depth-time profile can then be created allowing a temporal study through time and the opportunity to identify specific periods associated with biotic or chemical compositions that indicate significant environmental change (Smol 2008).

There are some limitations to these traditional paleolimnological methods. Each sediment core must be dated separately as sedimentation rates vary from lake to lake and even within a lake (Kirchner 2011). Lakes of high productivity generally have greater sedimentation rates than lakes of low productivity (Stein 1990; Weffer *et al.* 1990). A similar trend can be seen with temperature with cooler lakes (northern) having generally lower productivity and sedimentation rates (Cornwell 1985; Hermanson 1990). As each lake has its own unique sedimentation rate, climate, anthropogenic activity level and internal/external productivity, dating must be conducted on each core individually.

Furthermore, traditional methodologies usually require fossilized physical remnants of organisms. However, with the advent of molecular biology, researchers have begun making use of environmental DNA (eDNA). Analysis of eDNA has been utilized in the field of environmental science and paleolimnology to enable the tracking of specific populations, of rare and endangered species as well as invasive species without relying on physical examination or

capture of these organisms. Furthermore, normally non-fossil producing organisms such as soft-bodied microbes, can now be tracked and quantified in paleolimnological analyses allowing for study of a wider range of organisms. eDNA analyses have better detection limits than most chemical analyses and can measure both presence and quantity of target genes associated with specific organisms and is highly sensitive, capable of functioning based off of very small, preserved, fragments of DNA.

1.7 Application of Molecular Biology Techniques to Paleolimnology: The Study of Ancient DNA

Beginning as early as 1998, researchers have found that sediments not only archive physical fossilized structures, but can capture and store both recent and ancient DNA (sedDNA) material (Coolen & Overmann 1998). The use of ancient DNA and molecular techniques has opened a much wider array of organisms to act as historical proxies and includes most known eukaryotic DNA (ex: pollens, macrophytes, dinoflagellates, turtles etc.) and prokaryotic DNA (ex: cyanobacteria and bacteria) sources. This plethora of accessible organisms can be responsive to a much wider variety of environmental factors and shifts in their population can be used to more accurately determine past lake conditions. Several studies have now tracked cyanobacterial genes in sediments as well as tracked cyanobacterial community structures through the sequencing of specific target genes (Pal *et al.* 2015; Domaizon *et al*. 2017). Most of these studies have focused on deep stratified temperate lakes where bottom conditions are assumed to be most conducive to DNA preservation. These conditions include continuous darkness, cool year-round temperatures and anoxia. Whether shallower systems preserve DNA as well as those at deeper sites has not yet been determined or thoroughly examined.

1.8 Thesis Hypotheses and Objectives

Part one of this thesis (Chapter 2) examined the general hypothesis that changes in cyanobacterial dominances over time are a function of both eutrophication and climate change. This was tested by comparing lakes with significant human impact to lakes less developed within the same climatic region. Trends in cyanobacterial genes in sediment were analyzed in relation to nutrient and climate factors in both types of lakes.

The second part of this thesis examined the role of lake depth and morphometry in the preservation of sediment DNA. The hypothesis here was that deeper sites would have higher quality and quantity of sediment DNA than shallower sites because of sediment focussing (Blais & Kalff 1998) and more favourable *in situ* conditions for preservation. To test this, cores were taken from different depths of a multi-bay lake in order to compare both DNA quantity and quality.

The specific objectives were:

1) To extract and quantify genomic DNA from sediment cores with a focus on comparing paired lakes (1 a more nutrient rich lake and 1 a less nutrient rich) from two separate lake districts in Ontario, Canada

2) To quantify and compare the levels of four cyanobacterial genes (16s ribosomal sub-unit, *glutamine synthetase*, *mcy*E toxic gene and *Microcystis* specific 16s ribosomal sub-unit) in sediment depositions.

3) To determine whether cyanobacteria are increasing with time in relation to nutrients (eutrophication) and/or climate (regional changes beyond the catchment level).

4) To determine if cyanobacterial populations are becoming more toxic over time and whether this is associated with eutrophication and/or climate change

5) To determine if the overall stability (quality and quantity) of sediment DNA through time is dependent on lake depth, which is conflated with conditions favoring preservation.

6) To develop a primer set to quantify a gene associated with allophycocyanin production, a pigment protein present in many cyanobacteria, as a potential new proxy for cyanobacteria.

Chapter 2: Effect of Climate and Nutrients on Cyanobacteria of Temperate Lakes: An Analysis of Sediment DNA Through Time

2.1 Introduction:

Temperate regions are beginning to exhibit changes in weather indicative of climate change (Emanuel *et al.*1985; Davis & Botkin 1985; Houghton *et al*. 1991; Steffen *et al*. 2007). These changes include increases in temperature, changes in precipitation levels and more extreme events, which are affecting both terrestrial and aquatic ecosystems (Mooij *et al;* 2005; Wiedner *et al.* 2007; Mooij *et al;* 2007; IPCC 2018). Aquatic systems in particular are sensitive to changes in climate. Lakes are directly impacted by changes in atmospheric temperature, precipitation, ice coverage and wind speed (Schindler 2009; Adrian *et al.* 2009). These climate variables may also be implicated in secondary processes affecting ice-cover thickness and duration, the length and stability of thermal stratification of the water column, water levels (drought and/or increased precipitation) as well as affecting internal biogeochemical cycles (carbon, iron, nitrogen and phosphorus cycling) (Bischoff *et al*. 1997; Magnuson *et al*. 2000; Adrian *et al*. 2009). These secondary effects can lead to changes in the external loading of particulates and nutrients from surrounding catchments. Increases in nutrient loading, are considered the most important factor leading to increases in algal biomass in lakes and in particular in the abundance of cyanobacteria, (Paerl 2006; Dolman *et al.* 2007; Pick 2016)

Temperate lakes have generally experienced an increase in cyanobacteria over the past 50-70 years (Taranu *et al.* 2015). Cyanobacterial blooms appear to be occurring with increasing frequency and severity of their formations (Garcia-Pichel *et al.* 2003; LeBlanc *et al.* 2008; Paerl & Huisman 2008; Winter *et al.*2011; Paerl & Otten 2013). This trend has been mainly attributed to increased fertilizer use in agriculture, as well as nutrients coming from urbanization and

deforestation (Smith *et al.* 2016; Dodds *et al.* 2016). In addition, over the past decade, climate change has been invoked as an additional factor promoting the growth of cyanobacteria (Wiedner *et al.* 2007; Paerl & Huisman 2008). However, the lack of long-term instrument records and the few lakes sufficiently monitored through time have made it difficult to determine the actual impact of climate change on lakes and its relative importance in explaining the apparent rise in cyanobacteria in temperate lakes. The majority of data providing evidence for an impact of climate change stem from laboratory experiments or simulations and from extreme northern climes rather than more temperate regions (Arp *et al.* 2015; Smol 2016; Wood *et al*. 2017; Paterson *et al.* 2017; Pryzytulska *et al*. 2017).

As a group, cyanobacteria have several physiological traits that may provide a competitive advantage under conditions of increasing nutrient levels and climate change. Cyanobacteria as a group seem to have higher growth rates at higher temperatures compared to eukaryotic algae and are highly responsive to increased nutrient availability (high P), while being capable of fixing atmospheric nitrogen during periods of low N availability (Fernandez *et al*. 2015; Huisman *et al*. 2018). Although cyanobacteria are a natural component of aquatic ecosystems, bloom levels can threaten both wildlife as well as human populations that depend on water resources. Cyanobacterial blooms can lead to reductions in water clarity, macrophyte suppression, fish kills, de-oxygenation events (dead zones) and loss of economic value to properties and communities (Hoagland & Scatasta 2006; Paerl 2014). Cyanobacteria may also present a direct health risk: several taxa are capable of producing a variety of metabolites responsible for acute toxicity (e.g. microcystins, anatoxin) and chronic toxicity (Bell & Codd 1994; Jochimsen *et al.* 1998; Codd *et al*. 2005; Svirčev *et al.* 2010). Some of the more common

cyanobacterial toxins encountered in freshwater are the hepatotoxic microcystins, which can be produced by a wide range of cyanobacterial taxa (Sivonen & Jones 1999).

To determine whether these apparent changes in cyanobacterial populations deviate from the historical norm researchers have made use of paleolimnology, the study of markers buried in sediment to recreate past lake conditions (Smol 1992). Soft-bodied organisms, such as cyanobacteria have historically been difficult to study as they do not fossilize or preserve well for traditional microscopic and taxonomic evaluation. However, over the past two decades, researchers been successfully applying molecular methodologies on sediment bound DNA (sedDNA) (Ogram *et al*. 1982; Domaizon *et al*. 2013; Pal *et al*. 2015) to reconstruct past aquatic environments by utilizing specific gene proxies for prokaryotic communities. Pal *et al.* (2015) made use of quantitative PCR (qPCR) to quantify several microbial markers from several test lakes exhibiting unusual bloom formations, and showed that cyanobacterial genetic markers correlated well with more traditional carotenoid pigment markers previously used as proxies for cyanobacteria in sediments. Other studies have also been making use of other quantitative methods including ddPCR (Te *et al.* 2015) and high throughput sequencing technologies combined with advanced bioinformatics analyses to analyze complex community changes in cyanobacteria through time (Monchamp *et al.* 2016; Monchamp *et al*. 2017; Pilon *et al.* 2019). These studies have rarely considered the relative drivers of change (Yan *et al.* 2017; Burge *et al.* 2018; Pilon *et al.* 2019) and few have investigated temporal trends in the production of cyanotoxins given the challenges in distinguishing toxic from non-toxic taxa without molecular tools (Brasell *et al.* 2015, Wood *et al.* 2017).

The following study was designed to determine the relative importance of nutrients vs. climate change in lakes from a temperate region, by conducting paleoDNA analyses of cores

from paired lakes within the same lake district. One lake was designated "more affected" with a greater level of human impact in the catchment and the other lake was a "less affected" lake with a lesser human impact. Two separate lake districts were chosen for the study: both districts had continuous climate records going back over 100 years (almost back to pre-European settlement) and nutrient data from surface water monitoring.

The specific objectives of this study were: 1) to re-construct the historical cyanobacterial abundances in each set of paired lakes since the beginning of significant land use change, 2) to determine whether cyanobacteria are influenced by nutrient and/or climate factors in these same regions and 3) to quantify the abundance of microcystin genes in the paired lakes and determine whether climate factors or nutrients play a role in determining the potential toxin production in these temperate lakes.

In order to address these objectives, ddPCR was used to quantify bacterial, cyanobacterial as well as microcystin synthetase genes and thus recreate their historical abundances in four temperate lakes paired by region and by degree of anthropogenic impact. As a proxy of bacterial abundance, the glutamine synthetase (*glnA)*, a widespread, housekeeping gene required for nitrogen metabolism in prokaryotes, was targeted (Stoeva *et al.* 2014). Cyanobacterial total abundance was assessed using the 16s rRNA gene found in all cyanobacterial taxa (Rinta-Kanto *et al.* 2005; Pal *et al.* 2015). The microcystin gene chosen was *mcyE*, responsible for the synthesis of the ADDA moiety, the protein moiety common to all variants of microcystins (Vaitomaa *et al*. 2003; Ngwa *et al.* 2014). The gene target for the *Microcystis* specific 16s rRNA was also assessed as this genus is a common bloom-forming cyanobacterium in freshwater, that is frequently associated with microcystin production (Rinta-Kanto *et al*. 2005). These gene abundance profiles as well as relatively complete climate data

provided an opportunity to investigate the relationship between cyanobacterial abundance, toxicity and climate change.

2.2 Methods and Materials

2.2.1 Study Sites

Lake sites were selected in pairs from two distinct lake districts in Central Ontario, the Rideau Lakes of Eastern Ontario and the Muskoka Region of Central Ontario (Table 1). Within each region, one lake was designated as "more impacted" because of higher densities of housing or significant industrial and agricultural activities within the watershed. The other lake within the region was less impacted with comparatively less development within its watershed.

2.2.2.1 Muskoka Lakes District

In the Muskoka district the two lakes were selected following consultation with the Ontario Ministry of the Environment, Conservation and Parks (OMOECP, Dr. Andrew Paterson & Ron Ingram). The impacted lake chosen was Three Mile Lake, near Bracebridge, Watt Township. The lake has relatively dense housing development (> 611 cottages/homes), several agricultural fields abutting the lake, and small businesses including several resorts and campgrounds (Three Mile Lake Association, 2019). The less impacted lake chosen was Blue Chalk Lake, also located near Bracebridge, Ontario. It is relatively undisturbed with fewer than a dozen residential cottages (< 10) and no commercial or agriculture within the catchment. Blue Chalk is currently oligotrophic (Table 1).

2.2.2.2 Rideau Lakes District

After consulting with Jesse Vermaire and Michael Murphy (Carleton University) Big Rideau Lake was selected as the more affected lake in the Rideau/Portland region, east of Ottawa. It has significant housing development (>>700) along its shoreline, several ferries and marinas as well as commercial and recreational boat traffic. Big Rideau is part of the Rideau Canal system, which has historically experienced toxic cyanobacterial blooms (Pick 2016). The Rideau Lake district is generally quite developed and it was challenging to find a lake with little development. As a result, Otty Lake was chosen even though it still has shoreline and catchment development (>525 private cottages/houses and few businesses). Otty Lake lies close to the town of Perth, Ontario and connects to the Tay River which eventually flows into the Rideau River system (as does Big Rideau). Both Big Rideau Lake and Otty Lake were settled in the early 1800s and development accelerated when mineral deposits were discovered in the area and subsequently mined for mica and apatite (Otty Lake Association 2018; Big Rideau Lake Association 2018). The mining ended in the watershed of Otty in the early 1900's, but continued until the early-mid 1900's in Big Rideau. In addition, both lakes were affected by logging operations, which largely ceased in the early 1900's. The area surrounding Otty lake has since been claimed as private summer cottage residences as well conservation lands through acquisitions of the Rideau Valley Conservation Association.

The majority of the Rideau lakes now have the invasive zebra mussel (*Dreissena polymorpha*), a species originally native to the Caspian and Black Sea and introduced to the Great Lakes from ship ballast water releases during the mid-1980s (~1986) (Herbert *et al.* 1989). Zebra mussels appeared in the Rideau River system in the mid-1990s (Basu & Pick 1997) and have expanded to many of the Rideau lakes, including Big Rideau and Otty. Studies of US lakes

affected by zebra mussels have found that their presence is associated with higher concentrations of cyanobacterial toxins (specifically microcystins) (Guillaume *et al*. 2006; Sarnelle *et al.* 2012).

2.2.2 Lake Sediment Sampling

Sediment cores were obtained on each lake by boat. Coring sites were identified prior to travel from the bathymetric charts of the Canadian Hydrographic Service for the Smith Falls to Kingston region (Canadian Hydrogeographic Service Ottawa, Canada) as well as advice from the Vermaire Laboratory (Carleton University). The coring sites in Muskoka were chosen based on expert opinions from researchers at the Dorset Environmental Center (OMOECP) as well as with the use of bathymetric charts. The depths at each site were gauged using both Vexilar hand-held sonar as well as with a Hummingbird LCR 400 ID bow mounted sonar unit (Vexilar, USA; HumminBird, USA). Surface water was collected using 1.5 Liter Nalgene plastic containers (Nalgene Labware, USA; ThermoFisher, Canada) initially rinsed several times with lake water prior to collection.

All sediment cores were collected using a gravity corer, using 7.5cm inner diameter core tube acrylic coring tubes. The sediment cores were sectioned in lab or dockside shortly after their collection. Duplicate cores were taken where possible in case of accidental core loss and to utilize as test duplicates.

Sediment cores were sectioned (Douglas & Smol 2007; Pal *et al.* 2015) at 0.5 cm intervals, separated using a sliding metal plate mounted on a stage. Small changes to this method were made to reduce cross-contamination between sections. To minimize contact with the core tube the portions of the sediment core near the center were prioritized for retrieval. The central portion of each core section was taken and placed into a sterile 8 oz Whirl-Pak ® bags (Fisher Scientific, Canada). The spatula or slicer were then sanitized via a short (~5 second) dip in a

weak ethanol solution (~10%) and rinsed thoroughly with de-ionized (DI) water between slices. The stage was wiped down with Kim® wipes as best as possible avoiding disturbing or contaminating the core sediments. This process was repeated until the bottom stopper was revealed. Once the stopper was revealed that sediment section as well as the one preceding it were discarded due to potential mixing and contamination from contact with the rubber. Sediment samples were stored at -20° C until further processing.

2.2.3 Sediment Dating

Subsamples of selected sediment layers were processed for dating by gamma spectrophotometry by the LANSET (LANSET, 2018) facility at the University of Ottawa, using an Ortec Germanium Gamma Spectrophotometer (Oak Ridge, TN, USA) and certified reference materials obtained from International Atomic Energy Association (Vienna, Austria). Spectral data were acquired utilizing Maestro -32 (software, Ortec) and results analyzed using ScienTissiME (Barry's Bay, ON, Canada). The data retrieved from each core were then analyzed utilizing different models based on the presence of ^{210}Pb within the sediment sections. The three models tested were Constant Rate of Supply (CRS), Constant Initial Concentration (CIC) and Constant Flux Constant Sedimentation (CFCS, Binford 1990; Appleby 2008). The most appropriate model was then selected based on the best fit to the individual core ^{210}Pb and ^{137}Cs peaks within the core samples.

2.2.4 DNA Extraction

Sediment samples were defrosted overnight at 4°C for DNA extractions using the DNeasy® Powersoil® Kit from Qiagen (Germany). First, Power Bead® tubes were labeled and their contents removed and placed within a sterile 2 mL centrifuge tube for storage. The empty PowerBead® tubes were pre-weighed with no cap to obtain an initial tube weight without

sediment on a Mettler Toledo AB 204S/FACT (Mettler-Toledo, USA). Sediments were then added into each tube and re-weighed on a Mettler Toledo AB 204S/FACT (Mettler-Toledo, USA) scale to obtain initial wet weight of the sediment. Sediment sample wet weight ranged from 0.2 g to 0.5 g depending on moisture content, composition and malleability of the sediments.

Before carrying out the procedure outlined by the DNeasy® Powersoil® Kit, sediments were initially subjected to an additional 1 to 3 washes using Magic Wash Sediment buffered as outlined in Poulain *et al.* (2015) and originally adapted from Zhou *et al.* (1996). These additional washes can improve DNA recovery by removing humic compounds and divalent cations, which can interfere with nucleic acid recovery from sediments. The amount of washes varied depending on the source lake, the depth of the core section and the level of DNA content present in the sediment section after an initial exploratory test without the wash buffer. A 3:1 ratio of buffer to sample weight was utilized. This meant that ~600 µL of the wash was added to each sample. Samples were vortexed on Vortex-Genie 2 (Scientific Industries Inc., USA) at maximum setting for 30 seconds. The samples were then spun down at 3000 rpm for 3 minutes. The supernatant was discarded. These steps were repeated for each wash cycle (between 1-3 per sample). The tube and adjusted sediment weight were retaken using a Mettler Toledo AB-204S/FACT scale and recorded (Mettler-Toledo, USA). The true weight of the sediment was then calculated by subtracting the empty tube weight initially recorded from the final tube total weight (sediment and tube together post wash step). This was conducted as the initial sediment weight value would shift as liquid weight and minute sediment particles were lost.

Minor modifications were made to the original methodology outlined by Qiagen in the DNeasy® Powersoil® Kit (Qiagen, Germany). The changes were as follows: initially, 60 µL of

solution C1, a lysis buffer for breaking down cellular membranes, was added to each tube as well as the full contents of a PowerBead® tube (contents included: silica beads of various sizes and a pH neutral carrying buffer for extraction). The samples were then vortexed at maximum speed utilizing a Vortex-Genie 2 (Scientific Industries Inc., USA) and PowerBead 24-slot adaptor for between 10 minutes for <10 samples or 20 minutes for >10 samples in order to lyse cells. Then samples were centrifuged at 10,000 x *g* for 1 minute using a SIGMA 1-16 microcentrifuge (Sigma Zentrifugen, Germany) and 500 µL of supernatant was transferred to a sterile 2 mL microcentrifuge tube (although up to 800 µL could be transferred at this step without affecting final results). Then 250 µL of C2 inhibitor solution, a proprietary proteinaceous and humic substance removal solution, was added and vortexed for ~5 seconds. The samples were then placed at 4°C for 5 minutes. The samples were then centrifuged again at 10,000 x *g*. Avoiding the pellet, 600 µL of supernatant was transferred to a sterile 2 mL microcentrifuge tube. A second organic and inorganic substance removal solution was added at this step in the form 200 µL of Solution C3 in each sample. The samples were vortexed for ~5 seconds at maximum speed. The samples were then stored at 4°C for 5 minutes. The samples were then centrifuged at 10,000x *g* for 1 minute. Then 750uL of the supernatant was transferred to a sterile 2mL microcentrifuge tube, followed by the addition of 1200 µL of solution C4. Solution C4 contains ethanol with a highly concentrated guanidine salt content which facilitates charge binding to silica columns. Then 650 µL of each sample was added to the Qiagen Spin Column and the sample centrifuged at 10,000x *g* for 1 minute. Flow through was discarded and another 650 µL of solution added to the spin column. Solution was added until all of the sample was passed through the spin filter. Once completed 500 µL of an ethanol-based wash, solution C5, was added to the spin column to further remove substances other than nucleic acids. The column was

centrifuged at 10,000x *g* for 1 minute and the flow through discarded. The sample was immediately spun for a further 1 minute at 10,000x *g*. The bottom of the column was then discarded and the filter placed in a sterile 2 mL microcentrifuge tubes. Finally, 80 µL of elution buffer (TE based) solution C6 was placed directly into the center of the column filter. The sample was spun for 1 minute at 10,000x *g.* The final DNA extraction solutions were stored at -20℃ in 2 mL tubes until they were defrosted at 4 ℃ for subsequent analyses.

2.2.5 Quality and Integrity Analyses

The quality of DNA was determined using a NanoDrop™ 2000 spectrophotometer (Thermo Scientific, USA). An absorbance ratio of 260: 280 nm between 1.8 and 2.0 is considered good quality DNA under ideal laboratory extraction conditions (NanoDrop™ 2000/2000c User Manual). For the purposes of this study, and due to the complex environmental matrix, samples yielding 260: 280 ratios of 1.6-2.2 were considered to be of "useable purity" for use in further analyses. Values slightly outside of 1.6-2.2 were still considered for analysis if they provided positive results in the integrity analysis outlined below.

NanoDrop™ spectrophotometry detects raw quantity and purity. However, it is unable to detect whether or not the DNA is double or single stranded or has retained original sequences (i.e.: no depurination/depyrimidination). Prior to quantitative analysis, simple endpoint PCR reactions with a known product can indicate that the extracted sequences are still complete enough for PCR applications and downstream quantifications. To test for integrity, each sample was amplified via endpoint PCR using verified Glutamine Synthetase (*glnA*) as well as cyanobacterial 16s rRNA (CYA) targeting primers (Rinto-Kanto *et al*. 2005, Pal *et al.* 2015). Primer sequences can be found in Appendix A: Table 1a. Each extracted sample was individually amplified using a reaction mix composed of 12.5 µL EconoTaq® PLUS green 2x

master mix (Lucigen, USA), 2.5 µL forward primer, 2.5 µL reverse primer (of either *gln*A or *CYA* primer pairs) and 6.5 µL sterilized deionized nuclease-free water. After preparation of the master mix, 24 µL of the mixture was dispensed into a pre-labeled 0.2 mL well of an 8-tube strip. After which 1 µL of template was added for each respective sample to its own individual tube to create 25 µL final reaction volumes. Each set of reactions was accompanied by a positive control reaction. In the positive control the sample volume (1µL) was replaced with extracted genomic DNA from a pure culture of a microcystin producing strain of *Microcystis aeroginosa* (CPCC300, Canadian Phycological Culture Center, Canada). This sample was amplified, to ensure that there were no issues with respect to reagents or amplification conditions, which could return a false negative. Accompanying each positive control was a negative control of sterile water (1 µL). This negative control was to ensure that the reaction was not contaminated and that the primers were not amplifying false targets (returning false positives). Endpoint PCR was run on S1000™ Thermal Cycler (Bio-Rad Laboratories, USA) using the protocols as outlined in Appendix A: Tables 2a, 3a, 4a and 5a utilizing the primers outlined in Appendix A: Table 1a. The above was repeated for each primer for each sample.

Once PCR was completed the samples were then loaded and run on a 100 mL, 1.5% agarose gel (Wisent Bioproducts, Canada) impregnated with 3.3µL of GelRed® nucleic acid dye (the dye intercalates and fluoresces with double-stranded DNA) per 50mL of agarose gel (Biotium, USA) to separate nucleic acid fragments by size. Each set of reactions was accompanied by a well containing GeneRuler® 1kb DNA ladder (Thermo Fisher, USA) to act as a comparator and assess amplification product sizes. The gels were suspended in 1x TAE buffer solution on an OWL B2 gel dock (Thermo Fisher, USA) at 107 Volts for approximately 35 to 50 minutes. Gels were then visualized on an MBI Lab Equipment Fusion FX5 UV visualizer cabinet

(Montreal-Biotech, Canada) and photos were captured utilizing Fusion Molecular Imaging software and labeled utilizing Microsoft Paint 3D or Microsoft Paint.

2.2.6 Gene Copy Quantification – Quantitative PCR

Target genes from the sediment samples were quantified utilizing the Droplet Digital Polymerase Chain Reaction system from BioRad (ddPCR, Bio-Rad Laboratories, USA). The genes chosen for analysis (primers provided in Appendix A: Table 1a) included the glutamine synthetase gene (*glnA*), (found as one copy per cell in most bacterial taxa). This target gene was used as a proxy for total bacterial abundance in previous studies of lake sediments, as well as a bacterial reference gene (Poulain *et al*. 2015; Pal *et al.* 2015; Pilon *et al*. 2019). To infer total cyanobacterial abundance, degenerate primers targeting 16S rRNA were utilized (Nübel *et al.* 1997; Pal *et al.* 2015). Ribosomal RNA, particularly the 16S, is highly conserved, considered 'slow" to evolve and relatively unique to species (Snel *et al*. 1999; Fox *et al*. 1977). A bloom-forming cyanobacterium that can potentially produce microcystins was also quantified by estimating the abundance of *Microcystis* specific 16s rRNA genes (Pal *et al.* 2015; Pilon et al. 2019). Finally, to quantify the abundance of microcystin genes, primers targeting *mcyE* were utilized (Vaitoma *et al*. 2003; Rantala *et al*. 2006). The *mcyE* primer set can capture most microcystin producers whereas others may be more biased to particular genera (ex: *mcyD)* (Pilon *et al.* 2019; Zuo *et al.* 2018).

Primer sets for each gene target were tested and optimized on the ddPCR™ platform (gradient for optimizing conditions are summarized in Appendix A: Table 6a). Following optimization, the four gene targets were amplified and quantified in each sediment section from the four lake cores. PCR reactions were prepared by creating a mixture using 11.5 µL of QX200™ Evagreen™ Super Mix (Bio-Rad Laboratories, USA), 0.23µL of forward primer, 0.23

µL of reverse primer and 6.04µL of sterilized deionized water per reaction. Enough mix was created to dispense 18µL for each sample into PCR 8-well strip tubes. Wells were then loaded with 5 µL of the desired sample DNA and samples (of 20 µL) were then loaded onto the DG-8™ ddPCR™ cartridge (Bio-Rad Laboratories, USA) in the marked "sample" wells. After all sample wells had been filled, 60µL of Droplet Digital™ droplet generation oil (Bio-Rad Laboratories, USA) was then added to the wells marked "oil". Bubble additions to the oil or sample were avoided as this would halt the droplet generation process. Cartridges were clipped into droplet generation holders and a specialized ddPCR™ DG-8™ rubber gasket was stretched over each cartridge. After checking for a smooth gasket fit the cartridge in holder were placed into the QX200™ Droplet Generation machine. Droplet generation oil and sample were drawn in equal parts to create droplets. These droplets, an oil "shell" containing reaction mixture, were channeled into a third set of wells labeled "droplets". Once complete the gaskets were removed. Utilizing a F1-ClipTip 30-300 µL (Thermo Fisher, USA) brand multichannel pipette ~40 µL of each droplet reaction was then transferred from the droplet generation wells and dispensed slowly into the corresponding 8 well columns in a ddPCR™ 96-wellplate (Bio-Rad Laboratories, USA). This was done slowly to reduce shear forces on the droplets and retain high droplet counts as well as eliminate injection of air bubbles, which may also disrupt the integrity of the oil droplet reactions. The plates were then covered utilizing specialized foil with sealant provided by BioRad and sealed using the PX1™ Plate Sealer (Bio-Rad Laboratories, USA). Once sealed, the plates were placed into a C1000™ Touch Thermocycler (Bio-Rad Laboratories, USA). and subjected to the appropriate PCR conditions for the primers (Appendix A Table: 7a, 8a, 9a, 10a**)**.

Once the PCR reactions were completed the plates were transferred to a QX200™ Droplet Reader. Using a precision needle and capillary suction to take in each droplet from the

reaction, each droplet was recorded individually based on the fluorescence generated when excited during the analysis by the double-stranded DNA bound fluorescing dye from the QX200 Evagreen™ Supermix. This was interpreted as a binary positive for fluorescence, as more dye can bind and fluoresce as more of the target is replicated, and as interpreted as negative or null amplification for samples with no change in bound fluorescence, indicating no DNA replication and by extension no target was present. An output file is generated indicating positive and non-positive amplification peaks and provides the droplet results numerically, utilizing Poisson distribution statistics to generate a copy number per µL value. The results were analyzed utilizing QuantaSoft™ software (Bio-Rad Laboratories, USA. An example of data output is shown in Appendix A: Figure 1a.

2.2.7 Environmental Drivers: Climate and Nutrient Data

To determine the effect of climate on the study lakes, seven variables commonly associated with climate change (minimum, maximum and mean temperature, heating degree days, cooling degree days, total yearly precipitation and daily average precipitation) were chosen. The data were obtained from climate stations within each region in the Canadian Government Historical Climate Archives (climate.weather.gc.ca/historical_data/). Data were extracted utilizing the 'rclimateca' package in R version 3.5.2. The historical data records attached to station numbers were gathered by querying each station using its unique identification numbers to generate yearly sums for each recorded climate factor. These climate data files were then combined into useable climate data files for the time period encompassed by the lake cores (years ~ 1800-2018). Multiple stations were utilized to ensure full time-period data values and checked manually by auto-plotting the data and searching for years devoid of observations. The stations used were all within 100 km of the lake regions. Stations closer to the

sampling sites were preferred, however, the closer stations did not always yield higher quality data and often had few to no measurements, missing/incorrectly inputted data or significant gaps within their records. To ensure that climate variables between stations did not differ greatly trends and averages of each climate variable were checked by utilizing the 'autoplot' function of the 'ggplot2' package (R version 3.5.2) of the individual as well as the combined station data. Values within overlapping years were averaged between the two stations.

Stations used with corresponding accession numbers for Muskoka were: Gravenhurst Station ID: 4450, Muskoka Station ID: 4477, Muskoka Alternate Station ID: 48368, Beatrice Station ID: 4409 and Beatrice Climate station ID: 44183. Stations used with corresponding accession numbers for the Rideau region were: Kingston University Station ID: 4301, Lyndhurst station ID: 4308 and Brockville Station ID: 4235. Climate variables were analyzed and tested as both whole year averages, as well as selected growing season only (May-November) yearly averages.

Nutrient data were also evaluated for their effects on sediment gene copy numbers over time. Nutrient data were acquired through the Lake Partner Program of OMOECP, a citizen-based data collection program compiled by the Dorset Environmental Science Center (https://desc.ca/programs/LPP, Lake Partner Program of OMOECP 2019). Dorset Environmental Science Center provided additional nutrient data in the form of more detailed nitrogen and phosphorus level measurements for both Three Mile Lake as well as Blue Chalk Lake from their lake monitoring program (DESC, 2019).

2.2.8 Statistical Analyses

All statistical analyses were conducted in R (version 3.5.2) and RStudio software (version 1.0.153). Data were tested for normality using the Shapiro-Wilks test from the *stats* package and if required data were transformed accordingly (square-root or log transformation).

2.2.8.1 Temporal Trends and Correlation Analyses

In order to reduce the number of climate variables for further analyses, correlations among climate variables of normal or relatively normally distributed data were assessed using the Pearson Product-Moment correlation test (coefficient r). However, most climate variables did not fit a normal distribution (transformed or un-transformed) and were instead assessed using Kendall Rank correlations (Kendall's tau). Results were considered significant when the p-value was less than 0.05 or had a correlation strength exceeding +/- 0.2. Climate variables found to be highly inter-correlated were not included in further analyses together. Correlations were tested using the 'chart.Correlations' function of the 'performanceanalytics' package in R 3.5.2.

To identify changes in climate variables over the whole period of time, Mann-Kendall trend analyses were utilized to identify any monotonic upwards or downwards trends, and the strength of those trends. This was done via the 'Kendall' package in R 3.5.2. Breakpoint analyses from the 'segmented' package in R were used to estimate the years where the change in trend in the selected climate variables as well as gene targets appeared to occur.

Data acquired from the Lake Partners Program were examined for temporal trends in nutrients. Total phosphorus yearly averages were plotted against time and examined for monotonic trends via the Mann-Kendall analysis ('Kendall' R version 3.5.2). Phosphorus values were also averaged by month, separated by year, and plotted to examine trends through time. Total Kjeldahl nitrogen (TKN) was also examined but data were only available for Three Mile

Lake and Blue Chalk Lake. Because nitrate concentrations are at detection in surface waters during the growing season in these lakes TKN was considered equivalent to total nitrogen. Correlations were tested using Kendall Rank correlation matrices between TKN and TP values. This was carried out using the 'chart.Correlations' function of the 'performanceanalytics' package in R 3.5.2.

Gene copy numbers were plotted against time. Each gene target was then assessed for temporal autocorrelation, an issue inherent to paleolimnological analyses due to the nature of sediment deposition. To ascertain whether each gene was temporally autocorrelated, gene copy numbers were subjected to an ACF test utilizing the 'ACF' function (R version 3.5.2). Gene targets found to be autocorrelated required the inclusion of time as a factor within further statistical models. Each gene target in each lake system was then analyzed for the estimated dates of initial divergence in trends through breakpoint analysis utilizing the 'segmented' package in R 3.5.2.

Sediment gene copy numbers for each of the four targets were assessed for normality of their respective distributions via the Shapiro-Wilk normality test. Data that did not fit a normal distribution were log-transformed and reassessed. The relationships between the various gene targets were examined through correlation analyses when data distributions were normal (or relatively normal) using the Pearson Product-Moment correlation coefficient (r). Gene targets with copy number data that did not fit a normal distribution (transformed or un-transformed) were instead assessed using Kendall rank correlation coefficient. Results were considered significant when the p-value was less than 0.05 or had a correlation strength exceeding +/-0.2. Correlation analysis was carried out using the 'chart.Correlations' function of the 'performanceanalytics' package in R 3.5.2.

2.2.9.2 Partitioning of Variation

Climate variables were analyzed for their significance in explaining the variation in gene copy targets through redundancy analysis (RDA) for 999 permutations carried out using the 'vegan' package in R 3.5.2. Inflation of the *p* value from multiple testing was controlled through the 'p.adjust' Holms *p* correction function from the base 'stats' package of R 3.5.2. Due to the low number of years of available TP data for the majority of lake sites (observation number = <15), TP could unfortunately not be included in the partitioning of variation analysis.

To ensure that only the most significant explanatory variables were selected, each gene from each lake was subjected to forward selection modeling using the 'forward.sel' function of the 'adespatial' package in R 3.5.2. This forward selection was done to generate a model including only variables that significantly explained gene copy numbers. The forward selection proceeds by testing and adding the most significant explanatory variable; then testing the model again and either adding a further variable if it explains more variation or halts the selection process.

Climate variables identified as significant explanatory variables by permutation RDA and forward selection analysis were subjected to a partitioning of variation analysis, alongside time to account for temporal autocorrelation within the gene copy data sets. Partitioning of variation utilized the 'varpart' function of the 'vegan' package and was carried out utilizing R statistical software version 3.5.2.

2.3 Results

2.3.1 Dating and Sedimentation Models for Lake Cores

The cores were analyzed using gamma spectrophotometry for Lead 210 (^{210}Pb) and the trend validated with Cesium 137 (^{137}Cs) peak analysis. Dating through this methodology is limited in accuracy to several hundred years, which was sufficient for this study. Each lake was assigned a sedimentation model of best fit (Table 2). The dating was modeled backwards to generate a more complete time series for each section (raw dating profiles in Appendix A Figure 2a, 3a, 4a & 5a). The cores covered an estimated 163 years in Three Mile and over 700 years in Blue Chalk in the Muskoka lake district. The cores in the Rideau district covered an estimated 132 years in Big Rideau and an estimated 149 years in Otty. Therefore, the lake cores dated back prior to early European settlement for the Muskoka lake regions and encompassed the main period of change in the Rideau district. The estimated average sedimentation rate was lowest in Blue Chalk Lake and highest in Otty Lake (Table 3**)**. All cores appeared to be undisturbed with the exception of the Otty Lake core, which showed a slight mid-core mixing event (dating profile in Appendix A: Figure 5a).

2.3.2 Temporal Trends in Climate Variables

The temperature-based climate variables (maximum temperature, minimum temperature, mean temperature, heating degree days, cooling degree days) were found to be highly correlated to one another within each region after using the Kendall-Rank correlation. Precipitation variables (average daily precipitation and total yearly precipitation) were found to be highly correlated to one another but less significantly correlated to the temperature related climate variables.

In the Muskoka region, climate variables experienced changes over the time period of the study (~1880-2018). Maximum temperature and heating degree-days were the variables with the most significant change through time (largest tau and most significant *p* values). Maximum temperature (+ ~0.5°C over the time period) and mean annual temperature both showed significant upwards trends over time (Fig. 1 & 2). In contrast, the average minimum temperature showed no significant trend through time (Appendix A: Figure 6a). However, heating degree days has decreased significantly over time (degrees required to heat to 18°C) (Fig. 3), whereas no change occurred in cooling degree days (degrees required to "cool" to 18°C) (Appendix A: Figure 7a). Precipitation (both total yearly precipitation –Fig. 4 and daily average precipitation – Appendix A: Figure 8a) has decreased significantly. Breakpoint analysis of the Muskoka climate variables estimated that a break occurred in maximum temperature around 1994, whereas a break estimate was found in mean annual temperature (MAT) to have occurred earlier in 1977. Heating degree-days showed a break point estimated around 1992, a similar time frame to that of maximum temperature. Total yearly precipitation and daily average precipitation both had initial break points estimated (1995) close to this time period as well.

In the Rideau region, climate variables related to temperature as well as precipitation also experienced significant trends over the time period of the cores (~1880-2018). Maximum and mean temperature have both increased significantly (~ 0.75°C for both) (Fig 5 & 6), but the minimum temperature has not changed significantly (Appendix A: Figure 9a). Heating degree days have declined (Fig 7) while cooling degree days have increased (Appendix A: Figure 10a). In contrast to the Muskoka's, daily precipitation average and total yearly precipitation average both experienced significant increases over time (Fig 8, Appendix A: Figure 11a). Breakpoint analysis of Rideau climate variables estimated that a break in the temporal trend occurred in

maximum, mean and minimum temperature in approximately the year 1994 (Fig. 5 & 6, Appendix A: Figure 9a). Heating degree days had a break estimate identified around 1992 (Figure 7). In contrast, cooling degree days had a breakpoint estimate identified in the year 1944 (Appendix A: Figure 10a). Daily average precipitation has increased since the year 1888 (Fig 8) and total yearly precipitation appeared to be on a continuous upward trend and a breakpoint was not apparent (Appendix A: Figure 11a).

2.3.3 Temporal Trends in Nutrients

Water chemistry data were available from early 2000s to present with the exception of Blue Chalk where monitoring extended further back to the mid-1970s. Yearly averages were computed based on ~ 8-10 samplings per year during the ice –free season. For Three-Mile Lake in the Muskoka's, there has been a significant increase in yearly averages for total phosphorus (TP) based on Mann-Kendall trend analysis (Fig 9 A). Even when removing the high values of the most recent measurements, the trend remained positive. In contrast, Blue Chalk Lake in the same region has experienced a slight but significant decline (Fig 9 B).

For the Rideau district, little change was observed through time in Big Rideau with respect to total phosphorus, whereas a slight downward trend was apparent in Otty Lake (Fig. 10 A & B).

Total Kjeldahl Nitrogen (TKN) was also examined through time. Since nitrate concentrations in the surface waters of all these lakes is close to detection, TKN is essentially equivalent to total nitrogen. TKN data were only available for the Muskoka region for Three Mile and Blue Chalk lake. TKN and TP were highly correlated with one another in Three Mile

lake but not correlated in Blue Chalk Lake (Appendix A: Figure 12a). TKN did not increase through time in either Blue Chalk or Three Mile Lake (Appendix A: Figure 13a).

2.3.4 Sediment DNA Extraction Quality and Integrity

Quality and integrity of each DNA extraction was initially assessed by examining absorbance ratios of 260: 280 nm wavelengths on a NanoDrop™ spectrophotometer. All sediment DNA extractions yielded 260: 280 nm ratio values between 1.6-2.2 (1.8 being considered "pure" DNA) indicating that DNA was of useable purity. In Three Mile Lake, the variance in the ratio tended to be higher at the bottom of the core compared to the top of the core (Appendix A: Figure 14a). The ratios from Blue Chalk Lake and Big Rideau lake remained relatively stable and did not change significantly down the core (Appendix A: Figure 15a). In Otty Lake, the ratio of 260: 280 nm was more variable near the top of the core and stabilized throughout the rest. In all of the lakes the extracted DNA values peaked in the top 1 cm of sediments utilized and progressively declined towards the bottom of the core (Table 3)

After initial assessment via NanoDrop™, samples were amplified to ascertain the integrity of each extraction via endpoint PCR utilizing the *gln*A and *CYA* primer sets (Appendix A: Table 1a**)**. All samples in this study produced amplicons of appropriate size for the desired target indicating that the integrity of the DNA was high enough to allow for quantification (Fig. 11).

2.3.5 Changes in Gene Copy Numbers Over Time and Breakpoint analysis

Gene copy numbers were generated for each target for each sediment section. An example of this output can be seen in Appendix A: Figure 1a. The gene copies per gram of sediment were aligned with the dates acquired from the sediment dating process. Duplicates were made every 5 cm and were averaged for *glnA*, CYA and *mcyE*.

2.3.5.1 Muskoka District

In Three Mile Lake, gene copy numbers were relatively stable (or even on the decline) until a significant rise post 1850-1900s for all of the gene targets (Fig. 12). A breakpoint was identified for *glnA* gene copies in the year 1834 and slightly more recent breakpoint for CYA in the year 1908 (Appendix A: Figure 16a**)**. *McyE* and *MICR* both remained close to detection levels (10^3) until 1900s **(**Fig. 12**)**. *MICR* appeared to begin increasing sharply above detection in approximately the year 1937 (breakpoint analysis estimate), followed closely by *mcyE* in the year 1961 (breakpoint analysis estimate) (Appendix A**:** Figure 16a**)**.

In contrast to Three Mile, gene copy numbers in Blue Chalk increased only slightly over time for each target (Fig. 13). Initial breakpoint estimates were identified for *gln*A and CYA in the years 1586 and 1517 respectively (Appendix A: Figure 17a). A secondary breakpoint was identified for *gln*A in the year 1989 (Appendix A: Figure 17a). The *MICR* target remained close to detection with a steady increase over time (10^3) with a breakpoint estimated much later in the year 1866 (Appendix A Figure 17a). The increases for most targets appear to be steady over time (Fig. 13). The exception to this is *mcyE* (microcystin toxin gene target, Fig. 13) which remained at a relatively constant level around detection (~ 10^3 gene copy numbers per gram of sediment) throughout the core (Fig. 13). A shallow peak was observed in recent years (breakpoint estimated as 2002, Appendix A 17a) which may be an erroneous measurement of a contaminated interface sample or the beginnings of *mcyE'*s presence in Blue Chalk Lake.

2.3.5.2 Rideau district lakes

In Big Rideau, gene copy numbers increase gradually for all gene targets over time (Fig. 14). However, *MICR* and *mcyE* both remained near detection (~ 10^3 gene copy numbers/g) until very recently (~1970+), with higher values in the most recent years (2016 – present). Initial breakpoints could not be identified (Appendix A: Figure 18a). However, a secondary breakpoint estimate was identified for *glnA* and estimated to be in the year 1975 where the increase in copies per gram of sediment appears to accelerate **(**Appendix A**:** Figure 18a**)**. The cyanobacterial abundance genes (CYA*)* also appeared to be on an ever-upwards trend through time with no easily discernable initial breakpoint. However, a secondary breakpoint was identified slightly earlier than that of *glnA,* estimated to be in the year 1968, where the trend appears to accelerate (Appendix A: Figure 18a**)**. The *MICR* gene copies had an estimated breakpoint very similar to that of *glnA* in the year 1969 (Appendix A: Figure 18a**)**. Microcystin gene abundance, quantified through *mcyE,* was found to have an estimated break around 1977, slightly later than that of *MICR* (Appendix A: Figure 18a**)**.

In Otty lake, gene copy numbers increased slightly for all of gene targets over time with minor declines in recent years in *gln*A and an increase in CYA targets (Fig. 15). The *MICR* target abundance was only slightly above detection and experienced a steady increase since the 1950s. Gene copies of *mcyE* remained near detection then increased above detection in the past two decades (Fig 15**)**. Higher values creating "humps" in the gene copy numbers in some targets may be from a mid-core mixing event identified during sediment dating analyses (Appendix A: Figure 19a). Similar to Big Rideau lake, initial breakpoints were difficult to identify for *glnA* and CYA as they appeared to be on a gradual increasing trend; breakpoint estimates however were identified in 2001 and 1924 respectively (Appendix A: Figure 19a**)**. The break estimated for *gln*A was at the start of a sharp decline in copy numbers in the top sections of the sediments

beginning in the year 2001 (Appendix A: Figure 19a). The year 2001 also appears to be the beginning of an increase in copy number of the CYA target (Appendix A: Figure 19a). *Microcystis* specific abundance (*MICR*) was estimated to have a break in the year 1993 **(**Appendix A: Figure 19a**)**. The toxin gene *mcyE* had an estimated break in 1998 when gene copy numbers appeared to increase above detection **(**Appendix A: Figure 19a**).**

2.3.6 Relationships Among Target Genes

Many of the gene targets correlated significantly with one another (Appendix A: Figure 20a, 21a, 22a & 23a). The strongest correlations were typically those between the *glnA* and CYA 16S rRNA targets. Relationships that were either not significant or weakly significant were mainly correlations involving the *Microcystis* specific 16S rRNA or the *mcyE* gene subunit. (Appendix A: Figure 20a, 21a, 22a & 23a). In several Blue Chalk, Otty and Big Rideau lakes, the latter was at detection throughout much of the time series.

2.3.7 Effects of Climate Variables on Gene Copy Numbers

Based on the, corrected (Holms), significance values of the permuted regression analysis of the RDA of climate variables on gene target abundances, several factors appeared more frequently as significant than others (Table 4). For Three Mile, the heating degree-days was the only significant climate variable that emerged and it affected *glnA* and CYA. No climate variables were found to affect gene targets in Blue-Chalk. In Big Rideau more climate variables affected the gene copy numbers: maximum temperature, MAT, and heating degree days all affected *glnA* and CYA. Maximum temperature also had a significant effect on *mcyE*. Similarly, Otty lake showed significance of the same climate variables but cooling degree days also had an effect on *glnA*. In addition, average daily precipitation and yearly total precipitation were also found to be significant factors in Otty (Table 4).

The significant factors were further tested and confirmed by utilizing a step-wise, forward selection permutation (a form of stepwise regression model fitting). Forward selection tests the significance of each potential explanatory variable then retains the most significant explanatory variable, adding it to the model, it then tests all the variables again in addition to the retained variable if a factor increases the amount of variation explained it is also retained and the process repeated to generate a model of only significant explanatory variables.

2.3.8 Partitioning of Variation Between Climate Variables and Time

Nutrient data were excluded from partitioning of variation analysis due to the low number of years of monitoring (~12-years per lake excluding Blue Chalk), which was further reduced when matched to years associated with sediment sections. The sample size was not large enough to carry out partitioning of variation analysis (requiring>15 measurements).

2.3.8.1 Maximum Temperature

Due to the strong correlations between several temperature related variables, maximum temperature and heating degree-days were selected for the partitioning of variation analyses (results for the other climate variables can be found in Appendix A: Figure 24a – Figure 34a**).** Maximum temperature did not explain a large percentage in variation for any of the individual lakes (Fig. 16, maximum of ~ 6% of variation for *glnA* in Appendix A: Figure 24a, 28a & 31a). Time alone proved to explain a high degree of the variation in gene copy for every gene target in all lakes (~16-84%). However, the interaction between time and maximum temperature explained some of the variation in gene copies in Three Mile (0-5.6%). The Rideau district lakes had a higher percentage of gene copy variation explained by the interaction of time and maximum temperature compared to the Muskoka district (Fig. 16, Appendix A: Figure 24a – Figure 34a). Big Rideau maximum temperature alone explained as much as 4.3-6.3% of gene

copy variation and the combination between maximum temperature and time explained between 13-28% of gene copy variation in most gene targets. Maximum temperature alone explained almost none of the variation in Otty lake's gene targets (0-0.2%) but the interaction between time and maximum temperature explained between 9-14% of gene target variation (Fig. 16, Appendix A: Figure 24a – Figure 34a).

2.3.8.2 Heating Degree Days

Overall heating degree-days was a slightly more significant variable (Fig. 17, Appendix A: Figure 24a – Figure 34a) than maximum temperature in terms of explaining the variation in gene copy numbers through time**.** In Three Mile lake, heating degree days alone explained up to 3.3% of the variation in the gene targets; the interaction between time and heating degree-days increased the amount of variation explained to between 2.2-17.9%. In Blue Chalk, heating degree was not a significant explanatory variable. In Big Rideau lake heating degree days alone explained a 0.2-3.9% of gene target variation depending on the gene target**;** the interaction between time and heating degree days explained between 11.9-21.6% depending on the gene target (Fig. 17, Appendix A: Figure 24a – Figure 34a**)**. However, heating degree days did not explain any variation in *mcy*E or *MICR* (Appendix A: Figure 24a – Figure 34a**)**. For Otty lake heating degree days alone explained less variation than in Big Rideau, although the interaction cooling degree-days and time explained ~21% of variation in *gln*A gene copy number (Appendix A: Figure 24a – Figure 34a).

2.4 DISCUSSION

2.4.1 Cyanobacterial and Toxin Gene Abundances Through Time

The objectives of this study were to compare the historical abundances of cyanobacteria as well as toxin genes in pairs of lakes (one impacted and one less impacted by development), and to determine whether the temporal changes were related to climate and/or nutrient factors. With the exception of the low impact lake in the Muskoka district (Blue-Chalk), the trends suggest an increase in cyanobacterial dominance since European settlement: the abundance of cyanobacteria (based on the CYA gene copy numbers) has increased relative to the abundance of total bacteria (based on the *glnA* gene copy numbers) (Fig 12, Fig 13). This result is consistent with paleolimnological analyses of cyanobacterial and algal pigments in over 100 temperate lakes, which show a clear increase in cyanobacteria since the mid-20th century (Taranu *et al.* 2015).

In the present study, climate related variables appear to have played a small role, in the changes observed. The bloom-forming and potentially toxic cyanobacterium, *Microcystis* (*MICR)* appears to have increased in recent times, but the pattern of change was different between the two lake districts and was most pronounced in the impacted Muskoka lake, Three-Mile (Fig 13). With respect to trends in cyanotoxins, the *mcyE* gene copies numbers, which can be considered a proxy for microcystin production, were historically at detection (at the bottom of all the cores) but have increased in a few lakes in recent times, most notably in Three-Mile Lake, Muskoka. Three Mile has also seen an increase in cyanobacterial blooms in recent years, and the increase in bloom events has become a concern to residents and the local lake association (Muskoka Water Web 2005; Cottage Life 2018; SimcoeMuskokaHealth.org 2019). This suggests either a recent introduction or increased propagation of both cyanobacteria and toxin genes and,

by extension, an increase in the abundance of toxigenic cyanobacteria over time. This is similar to the trend reported for Lake of the Woods, whereby both *Microcystis* and *mycD* genes appeared first in the sediment record in the late ‘70s to early ‘80s and proceeded to increase. In this case regional warming was the likely explanation (Pilon *et al.* 2019; Smol 2019).

There are some assumptions and limitations inherent with the paleo-genetic reconstruction methodology that should be considered when interpreting gene proxy data. Absolute values of gene copy numbers cannot be compared between lakes as extraction efficiency of DNA and DNA preservation likely vary with lake and sediment geochemistry. Trends within a lake through time must also be interpreted with caution, including trends of increasing absolute copy numbers over time, as this could be the result of sediment diagenesis and simply reflect a pattern of DNA degradation. In this study, DNA extracted from lower portions (older) of the cores was found to return consistently lower absolute values as well as more variable quality ratios (NanoDrop™ 260:280nm ratio readings) when compared to sediment sections closer to the water interface (newer). Therefore, it is considered more reasonable to examine the relative changes of genes to one another (Pal *et al.* 2015; Pilon *et al.* 2019). This is further compounded as not all gene targets are present in similar copy numbers on a per cell basis. Targets such as *glnA* are present as a single copy in prokaryotic cells (Stoeva *et al.* 2014), whilst the cyanobacterial 16S rRNA gene (CYA) is present as one to four copies per cell, depending on the species (Kaneko *et al*. 2007). Although this could lead to an over-representation of cyanobacterial abundance, all gene targets, and particularly *glnA* and CYA, exhibited strong correlations with one another (Appendix A: Figure 20a- 23a). This suggests that the changes observed over time were not the result of species-related copy number changes and rather are more likely attributable to overall changes in relative abundance of cyanobacteria.

The effect of degradation processes on DNA preservation is an important factor to consider in paleo-ecological research (Corinaldesi *et al.* 2008; Fernandez-Carazo *et al*. 2013). Although there remains some uncertainty as to the strength and rate of degradation within aquatic systems (Hebsgaard *et al*. 2005), steps were taken in this study to ensure that the effect of degradation was minimized. Lakes chosen were deep enough to stratify; stratification ensures that sediments from deep sites were subjected to cold temperatures (<4°C) year-round, which would slow the rate of DNA degradation and DNase enzymatic activities (Coolen & Overmann 1998; Burger *et al*. 1999; Strickler *et al.* 2015). Increased depth would also eliminate exposures to ultraviolet/ionizing radiation, which has been shown to damage DNA both directly and indirectly (Baust 2008; Strickler *et al.* 2015), and result in sediment anoxia, which is also likely conducive for DNA preservation (Corinaldesi *et al*. 2011).

None of the primer pairs utilized in this study generated an amplicon greater than 300 base pairs in length. Past studies as well as simulations have shown that shorter, rather than longer, sequences may be more persistent and resistant to the effects of degradation, preserving for longer periods of time whilst still remaining representative (Allentoft *et al*. 2012; Coolen *et al.* 2013; Capo *et al* 2017; Domaizon *et al.* 2017). Simulations of degradation of terrestrial ancient DNA have estimated that environmental samples may take as little as days, but may be preserved under ideal conditions for potentially tens of thousands of years (Willerslev *et al*. 2005; Corinaldesi *et al*. 2008; Allentoft *et al*. 2012). The portions of DNA surviving for tens of thousands of years are estimated to be fragments of no more than several hundred base-pairs in length, which is within the size range of the gene targets utilized in this study, making it less likely that they would be fully degraded (Allentoft *et al*. 2012).

2.4.2 Trends in Climate and Nutrients in Muskoka and Rideau Lake Districts

Several climate variables showed temporal changes consistent with climate change (IPCC 2018). Breakpoint analyses revealed specific points in time where patterns and trends deviated from prior trends with the majority occurring in the past 20-40 years and few occurring in earlier periods of time (1800-1900). This is consistent with IPCC (2013) reports pointing to more pronounced temperature changes globally post 1980s. Pal *et al.* (2015) observed greater change in cyanobacterial gene copy numbers post the 1980s relative to the 1930s in lakes of western Quebec (both inside and outside a protected park). In this study, the climate trends differed slightly between the two lake districts.

In both the Muskoka and Rideau districts, maximum and mean temperature have experienced a significant increasing trend. The Muskoka district average temperature has increased by ~ 0.5°C whereas the temperature in the Rideau has increased by ~0.75 °C. This is consistent with other analyses of the instrumental climate data records for Central Ontario during the past 100 years (e.g. Favot *et al*. 2019). This is quite similar to the global average (~0.8°C, NASA 2019) but lower than changes at higher latitudes, including those found at the previously mentioned Lake of the Woods (Rühland *et al.* 2010). In addition, heating degree days, a less commonly studied and a cumulative variable (of the degrees when a given day's mean temperature is below 18 °C) has declined significantly in both districts, reflecting milder winters and, most likely, a longer open water season (and growing season for cyanobacteria).

Trends in precipitation differed between the two districts. The Muskoka District has experienced a significant decrease in volume over the past 100 years, which is consistent with findings in previous studies from the region (Persaud *et al.* 2014). Overall, declines in precipitation could lead to longer water residence times, which creates the potential for more

stagnant water conditions as well as potentially leading to internal nutrient loading, which may favor cyanobacterial populations. In contrast, the Rideau district has experienced an increase in total annual and daily precipitation. This change in precipitation could be increasing the amount of nutrient run-off from land into lake systems, thereby indirectly contributing to cyanobacterial dominance.

In terms of nutrient changes, unfortunately, the monitoring record only went back as far as ~2000 and the surface water data could not be statistically related to the sediment records. From the records available it was found that over the past ~ 18 years, very little change in phosphorus has occurred in the less impacted Blue Chalk where relatively low total phosphorus concentrations (5-10 μg/L) were recorded and appeared to be declining over time. The concentrations are consistent with levels seen in oligotrophic conditions (~ 1-12 μg/L of TP, OECD, 2019). This decline is also consistent with other declines in TP that have been reported for Central Ontario Shield Lakes more generally since the 1970s (Palmer *et al.* 2011); this might be related to the overall decline in precipitation as well as changes in land use practices to counter eutrophication of lakes or the gradual reforestation of areas previously deforested. In contrast, the impacted Three Mile has seen elevated total phosphorus values in the last few years for reasons likely linked to watershed development and land-use intensification. Total phosphorus concentrations within Three Mile have been within the mesotrophic range (10-35 $\mu g/$L, OECD 2019), however, concentrations appear currently to be approaching eutrophic status (Table 1**).** The Rideau lakes have experienced a slight decline in total phosphorus over the past one to two decades. Big Rideau and Otty are within the low-mesotrophic range with the former now approaching oligotrophic levels (< 10 ug/L). Otty Lake, despite less catchment development, has in fact higher TP concentrations (12-16 $\mu g/$L) than Big Rideau. It was difficult

to find a suitable reference lake for the Rideau district given its longer period of settlement and the extent of historical industrial activity in the region. Otty was not as "pristine" as Blue-Chalk Lake in the Muskoka district and, in hindsight, Otty was not as suitable a reference lake.

2.4.3 Relationships Between Cyanobacteria, Climate and Nutrients

Based on the variance partitioning analyses, climate variables individually explained a very small fraction of the variation in gene copy numbers through time (=< 5.7%). This was the case for all lakes including the more "pristine" reference lakes where change might be less affected by nutrient related factors. It is possible that climate change is too recent to be detected in the sediment record since many breakpoints were identified in only the top ~3-12 cm of cores. Furthermore, it is possible that the magnitude of climate change to date has been too small to be resolved in these lake cores, compared to the greater changes in temperature observed for example in Lake of the Woods (Rühland *et al.* 2010, Rühland *et al.* 2018).

When the shared contribution of climate and time is considered, a much larger amount of variation in copy numbers was explained. Maximum temperature and heating degree days became the most significant explanatory variables, when their shared contribution was evaluated with time, both factors were found to explain between 14-21 % of cyanobacterial abundance (based on gene copy numbers of CYA). Big Rideau showed higher shared contributions than Muskoka on average for heating degree days and maximum temperature. In contrast, the effect of shared contribution of all climate variables and time appeared to be less important for total bacterial abundance explaining only 2.2-2.5% of the total variation in the Muskoka region. However, when the shared contribution of climate and time was evaluated in Rideau it explained similar amounts of variation to those observed for cyanobacterial genes (~21-28%). Temperatures influence on cyanobacteria biomass is well documented (e.g. Beaulieu *et al.* 2013)

and may have both direct (on growth rates) or indirect (on length of growing season and stratification) effects on their growth (Dokulil & Teubner 2000). Although temporal trends in precipitation (both negative and positive) were significant in both districts, precipitation did not contribute significantly to gene copy variation, except in Otty lake, either alone or in shared contribution with time, suggesting that the effects of temperature may be more important to cyanobacterial ecology than associated changes in hydrology and nutrient loading, at least in these lake districts.

Despite examining several climate variables, there were potentially other variables, such as wind speed, which could also help in explaining the variations in cyanobacteria over time. Wind speed can be important in explaining year-to-year variations in bloom severity (Persaud *et al.* 2014) and in the distribution of microcystin congeners (Taranu *et al*. 2019). Instrument records for wind speed are not as extensive as those for temperature and precipitation, thus limiting their utility in paleo-ecological studies.

2.4.4 Toxigenic Cyanobacteria and Microcystins

The relatively unimpacted Blue Chalk in Muskoka had no microcystin genes (*mycE*) above detection throughout the entire sediment record. Only the very top of the core (water interface) provided a single sample with gene copy numbers above detection; this could be either a very recent change or the result of some surface water contamination. Furthermore, there has been no increase in the relative dominance of cyanobacteria in Blue-Chalk. This result is in keeping with its oligotrophic (almost ultra-oligotrophic by international standards) status and confirmation of its pristine relatively unchanged state. Although there has been a slight overall rise in CYA abundance this rise is less than an order of magnitude, especially when compared to the impacted lake Three Mile. Three Mile has experienced an increase on the scale of several

orders of magnitude for all gene targets. Bacterial total abundance and cyanobacterial total abundance have both been increasing since ~1834 and ~1908 respectively with absolute cyanobacterial total abundance trending towards the total bacterial abundance. Both *Microcystis* and microcystin genes also began increasing since approximately the 1930s'– trending towards total bacterial abundance values, which in combination with the trends observed for cyanobacterial total abundance is suggestive of a significant regime shift. This is coincident with the recent rise in total phosphorus, well past the guideline for the protection of aquatic life in Ontario (TP > 20 ug/L). Changes in nutrient loading and/or internal loading are the most likely explanations for the rise in toxigenic cyanobacteria in Three-Mile Lake. Rising temperatures in summer and shorter winters are likely acting in concert to further exacerbate cyanobacterial blooms as suggested for other temperate lake systems (Paerl & Huisman 2008).

In Big Rideau and Otty, *Microcystis* spp. copy number remained low, only slightly above the detection threshold experiencing an increase in recent years in the 1970s onwards. Breakpoint analysis points towards an increase in *mcyE* copy numbers above the detection threshold beginning in 1977. Interestingly, as observed for *glnA* and CYA, *MICR* and *mcyE,* gene copy numbers also appear to approach one another in recent years. This is indicative of a shift within Big Rideau lake towards cyanobacterial dominance and a greater than historical proportion of cyanobacteria carrying the *mcyE* gene (i.e.: being capable of the formation of microcystin production).

Despite some evidence of internal mixing within the Otty Lake core, the general results were similar to those observed from the Big Rideau core. Otty lake, as mentioned previously was similar in trophic state to Big Rideau and thus was not an ideal unimpacted lake. Interestingly, at the top of the core there is a decrease in both *glnA* and CYA copy numbers. This could be due to

the recent zebra mussel invasion that would lead to increased grazing pressures and clearing of the water column (Raikow *et al*. 2004). Conversely, *MICR* and *mcyE* both increased during this same time period which may also be a result of the selective feeding habits of zebra mussels that can increase microcystin concentrations (Juhel *et al*. 2006, Knoll *et al*. 2008). Otty was also found to have an earlier breakpoint for *mcyE* than *MICR* which may indicate that *Microcystis* spp. are not the only producers of microcystins and that other cyanobacterial species may be responsible for its presence. Furthermore, *mcyE* gene copy numbers do not appear to approach either CYA or *MICR* in abundance, indicating no shift towards dominance of toxigenic cyanobacteria.

2.5 Conclusion:

Overall, for these Central and Eastern Ontario lakes, the magnitude of temperature changes appears to be minor compared to changes observed at more extreme northern latitudes and that those climate changes may not be "powerful" enough, or be occurring for long enough, to be reflected in the sediment proxies of this study. Furthermore, factors not explored in this study, such as wind speed, or not adequately tested (TP), may be playing a role in driving community changes. Despite this apparent lack of direct effects of climate on gene proxies, there is enough evidence to indicate that the Rideau region may be experiencing more significant effects of climate change than those observed in the Muskoka region, illustrating how the effects of climate can differ at small regional scales and regional differences should be examined in future studies. This research illustrates that the continuous trend in relative abundances towards cyanobacterial dominance, seen in impacted lakes and the capacity to produce toxins (Three Mile, Big Rideau) may in fact be the result of a synergistic effect of both nutrient concentrations (primary driver) and climate (additive driver).

Table 1

Table 1: General characteristics of four lakes (Three Mile, Blue Chalk, Big Rideau and Otty) from the Rideau and Muskoka districts in Ontario, Canada.

Lake Name	Three Mile	Blue Chalk	Big Rideau	Otty
Lake Coordinates	45.19 -79.46	45.19 - 78.93	44.71 -76.21	44.84 - 76.22
Average Depth (m)	4	9	12	9
Maximum Depth (m)	13	22	110	27
Shoreline Usage (Excluding Natural/Regenerative Riparian Zones)	54.46% Altered	15% Altered	27% Altered	27.1% Altered
Shoreline Land Usage	Housing (~>300), Recreational Centers, Marinas, Camping, Agriculture/Pasturelands	Housing (<20), Fishing Destination	Housing (>>500), Industry, Mining, Logging (Defunct), Shipping Route, Wineries	Housing (~>100), Mining (Defunct), Logging (Defunct) Designated Conservation Areas
Size (Surface Area km^2)	8.66	0.5	48	6.9
Watershed Size (km^2)	135.16	1.2	407	49.2
Zebra Mussels	No	No	Yes, 1980's	Yes, 1980's
Trophic State	Oligo - Mesotrophic	Oligotrophic	Meso- Eutrophic	Meso- Eutrophic

Table 2

Table 2: Individual lake sedimentation models with average sedimentation rates as well as maximum estimated date from the dated section. CFCS is Constant Sedimentation model and CRS is Constant Rate of Supply Model.

District	Lake	Sedimentation Model	Sedimentation Rate (est. in $g/cm^2/year$)	Estimated Date of Bottom Portion of Core
Muskoka	Three Mile Lake	CRS	0.038	1856
	Blue Chalk Lake	CFCS	0.0021	1222
Rideau	Big Rideau Lake	CRS	0.021	1887
	Otty Lake	CRS	0.06	1870

Table 3

Table 3: Average nucleic acid (ng/µL) extracted from lake cores (n = 70-110). Top 10 cm and the bottom 10 cm (n = 20). As well as the average 260: 280 nm ratio assessing DNA quality of each core from each lake site (n=70-110).

Region	Lake	Average Nucleic Acids extracted of All Sections (ng/µL)	Average Extraction Yield Top 10 cm of Core (ng/µL)	Average Extraction Yield Bottom 10 cm of Core(ng/µL)	Average 260:280nm NanoDrop™ Ratio (Whole Core)
Muskoka	Three Mile Lake	11.83	36.74	4.09	1.82
	Blue Chalk Lake	6.96	11.07	4.55	1.8
Rideau	Big Rideau Lake	18.53	24.85	11.98	1.78
	Otty Lake	30.61	26.73	25.87	1.83

Table 4

Table 4*:* Significance, *p* values, for permuted regression (999 repetitions) of RDA between climate factors and gene copy values of gene targets in each lake. Significant values are designated by symbols with significance value increasing in proportion to the number of asterisks. *P*-values subjected to Holms corrections (. , *, ** or NS for not significant)

Three Mile lake	***glnA***	**CYA**	***MICR***	***McyE***
Maximum Temperature	0.162	0.168	0.077	0.084
Minimum Temperature	0.399	0.459	0.448	1
Mean Temperature	0.285	0.31	0.18	0.35
Heating Degree Days	0.007**	0.007**	0.077	0.35
Cooling Degree Days	NS	NS	NS	0.847
Average Daily Precipitation	0.324	0.376	NS	0.653
Yearly Average Precipitation	NS	NS	NS	0.807
Blue Chalk lake				
Maximum Temperature	NS	0.511	NS	NS
Minimum Temperature	NS	NS	NS	NS
Mean Temperature	NS	NS	NS	NS
Heating Degree Days	NS	NS	NS	NS
Cooling Degree Days	NS	NS	NS	NS
Average Daily Precipitation	NS	NS	NS	NS
Yearly Average Precipitation	NS	NS	NS	NS
Big Rideau lake				
Maximum Temperature	0.007**	0.014*	0.07 .	0.014 *
Minimum Temperature	0.078 .	0.056 .	0.156	0.075 .
Mean Temperature	0.025*	0.018*	0.072 .	0.30
Heating Degree Days	0.024*	0.025*	0.1	0.192
Cooling Degree Days	NS	NS	NS	0.396
Average Daily Precipitation	0.064 .	0.333	0.66	0.95
Yearly Average Precipitation	NS	NS	NS	0.782
Otty Lake				
Maximum Temperature	0.048 *	0.007 **	0.020 *	0.007 **
Minimum Temperature	0.416	0.08 .	0.345	0.076 .
Mean Temperature	0.177	0.007 **	0.096 .	0.007 **
Heating Degree Days	0.416	0.006 **	0.007 **	0.002 **
Cooling Degree Days	0.014 **	0.057 .	0.345	0.904
Average Daily Precipitation	0.168	0.035 *	0.007 **	0.904
Yearly Average Precipitation	0.025 *	0.463	0.345	0.285

Figure 1

Maximum Temperature (Muskoka):

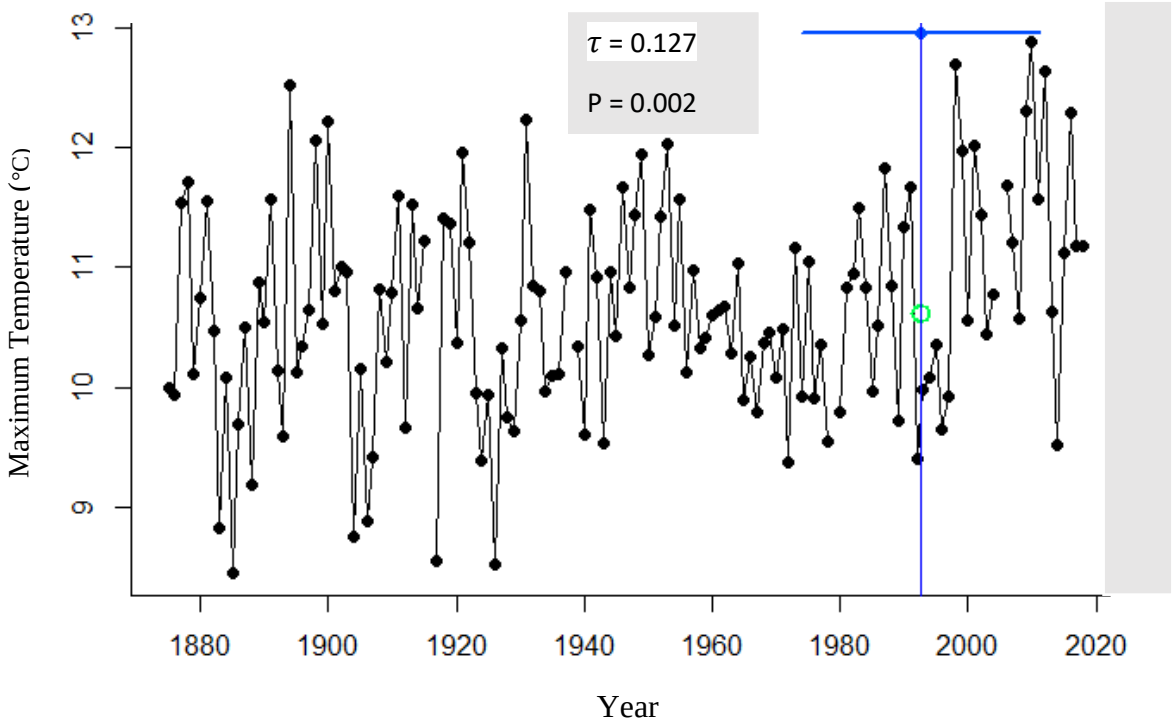


Figure 1: Maximum Temperature (°C) over Time (Year) in the Muskoka region. Blue line indicates the estimated "Breakpoint" (1994) where pattern begins to deviate. ταυ (τ) indicates positive trend and p (<0.05) indicates significance of the trend.

Figure 2

Mean Temperature (Muskoka):

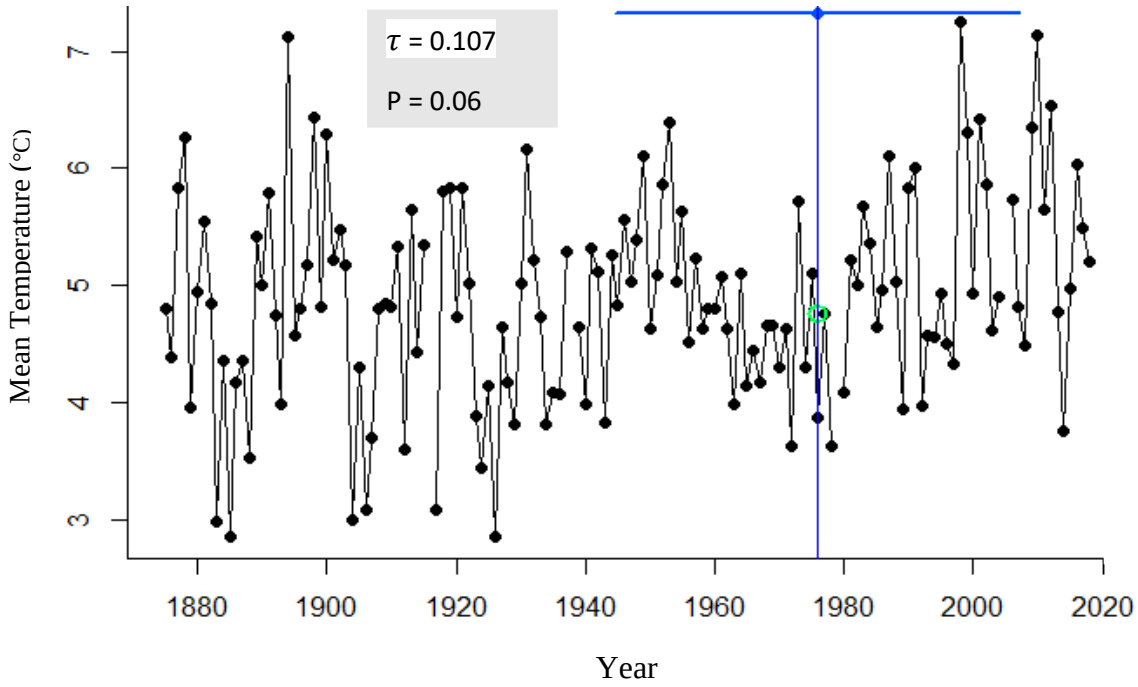


Figure 2: Mean Temperature (°C) over Time (Year Blue line indicates the estimated "Breakpoint" (1977) where pattern begins to deviate. tau (τ) indicates positive directionality and p value indicates a non-significance of the trend.

Figure 3

Heating Degree Days (Muskoka):

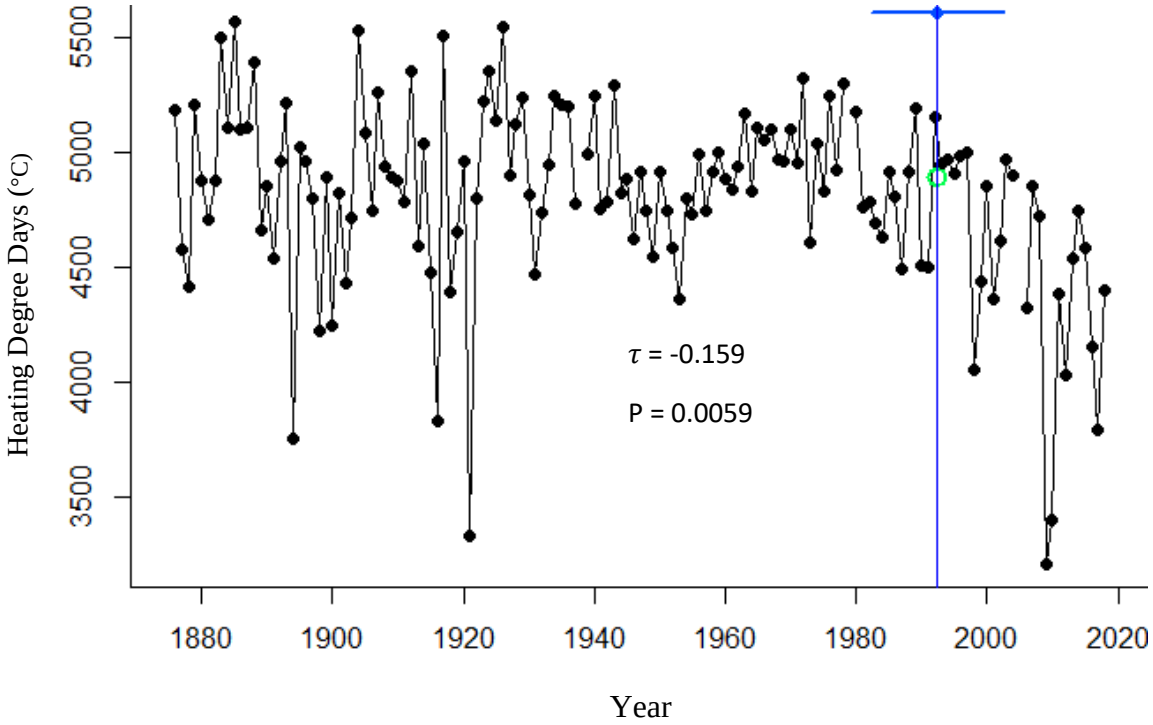


Figure 3: Heating degrees (°C) over Time (Year). Blue line indicates the estimated "Breakpoint" (1992 where pattern begins to deviate. ταυ (τ) indicates negative directionality and proportionality and p (<0.05) indicates significance of the trend.

Figure 4

Precipitation Total (Yearly, Muskoka):

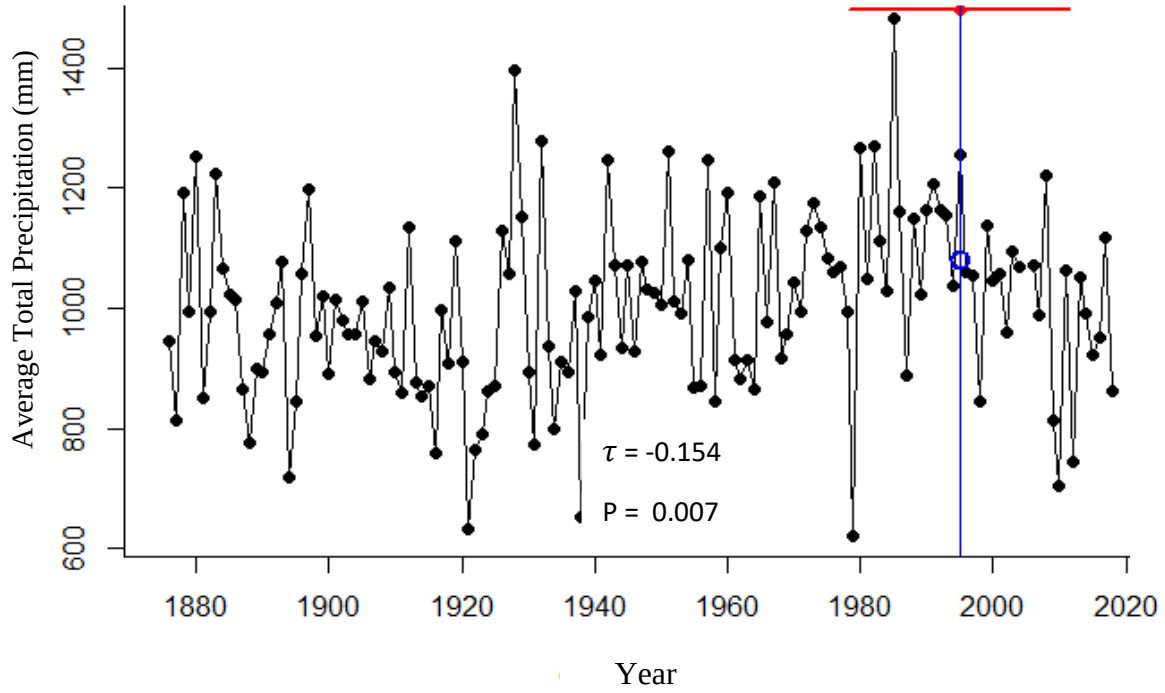


Figure 4: Precipitation total (mm) over Time (Year). Blue line indicates the estimated "Breakpoint" (1995) where pattern begins to deviate. tau (τ) indicates positive directionality and p value (<0.05) indicates significance.

Figure 5

Maximum Temperature (Rideau):

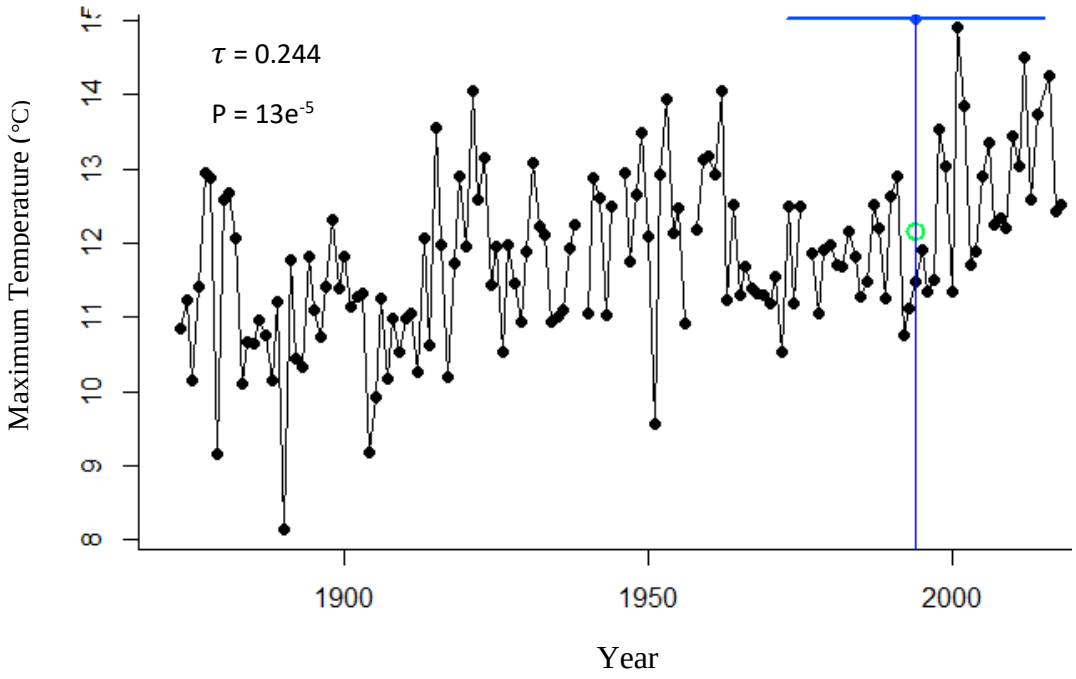


Figure 5: Maximum Temperature (°C) over Time (Year). Blue line indicates the estimated "Breakpoint" identified (1994) where pattern begins to deviate. tau (τ) indicates positive directionality and p value (<0.05) indicates significance.

Figure 6

Mean Temperature (Rideau):

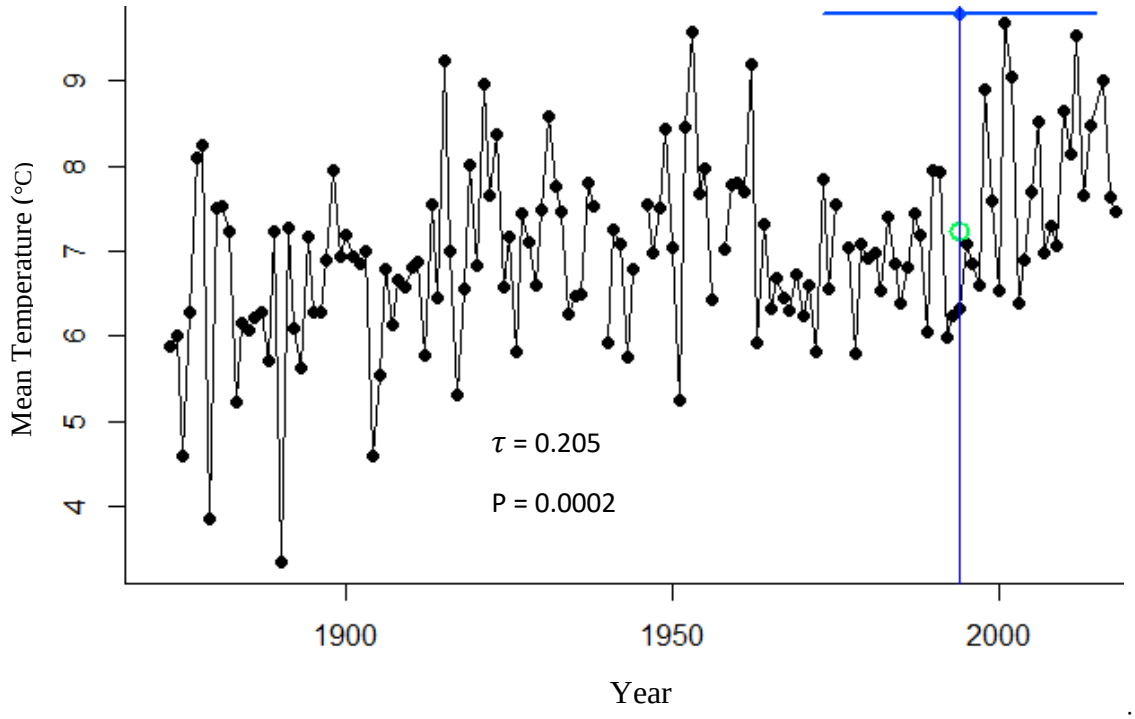


Figure 6: Mean Temperature (°C) over Time (Year). Blue line indicates the estimated "Breakpoint" (1994) where pattern begins to deviate. tau (τ) indicates positive directionality and p value (<0.05) indicates significance.

Figure 7

Heating Degree Days (Rideau):

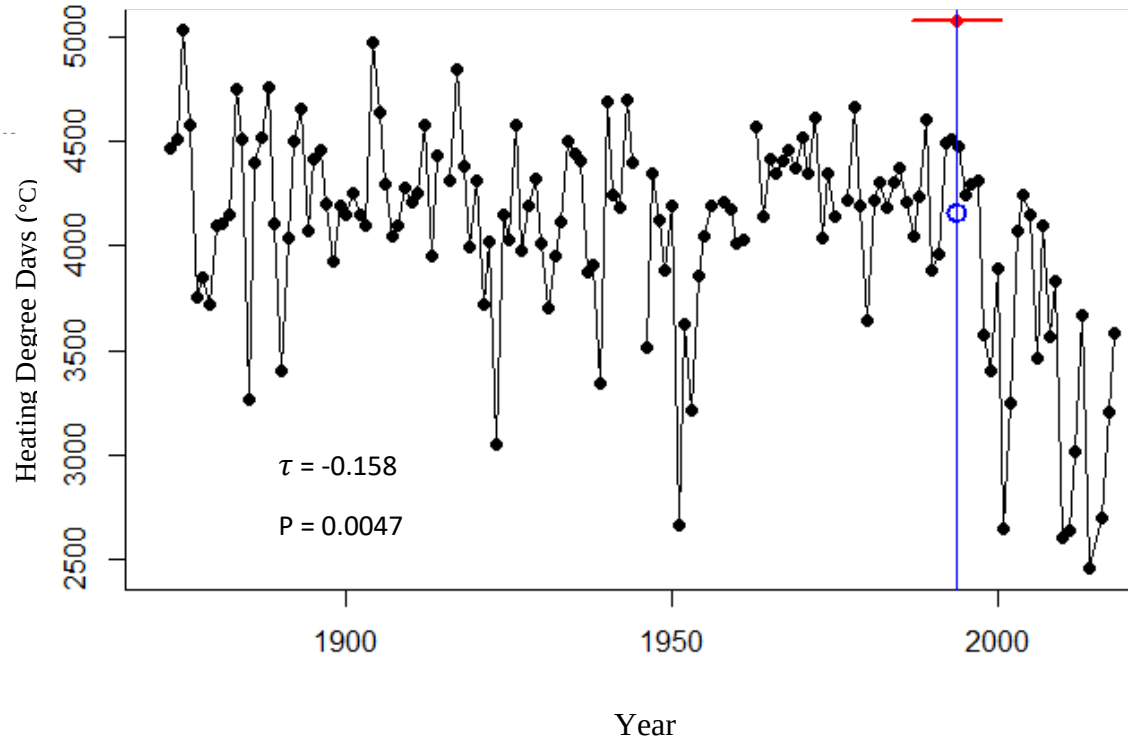


Figure 7: Heating Degree Days (°C) over Time (Year). Blue line indicates the estimated "Breakpoint" identified (1995) where pattern begins to deviate. Tau (τ) indicates positive directionality and p value (<0.05) indicates significance.

Figure 8

Precipitation Average (Daily)

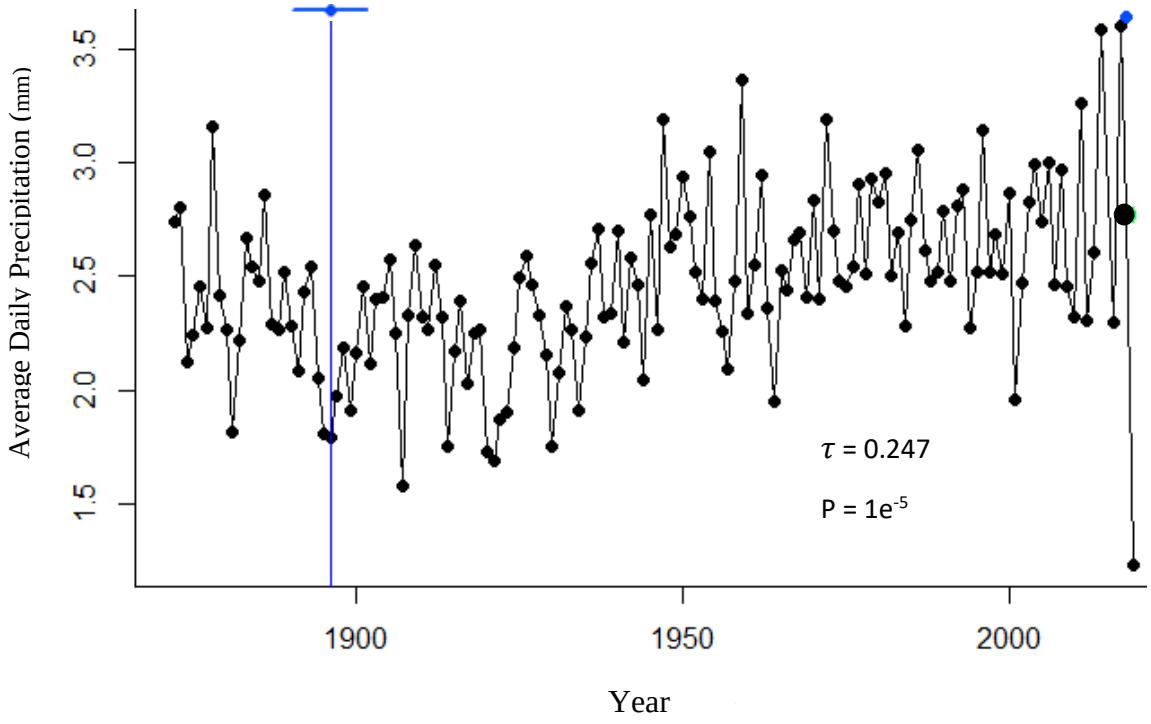


Figure 8: Precipitation Daily Avg (mm) over Time (Year). Blue line indicates the estimated "Breakpoint" identified (1888) where pattern begins to deviate. ταυ (τ) indicates positive directionality and p value (<0.05) indicates significance.

Figure 9

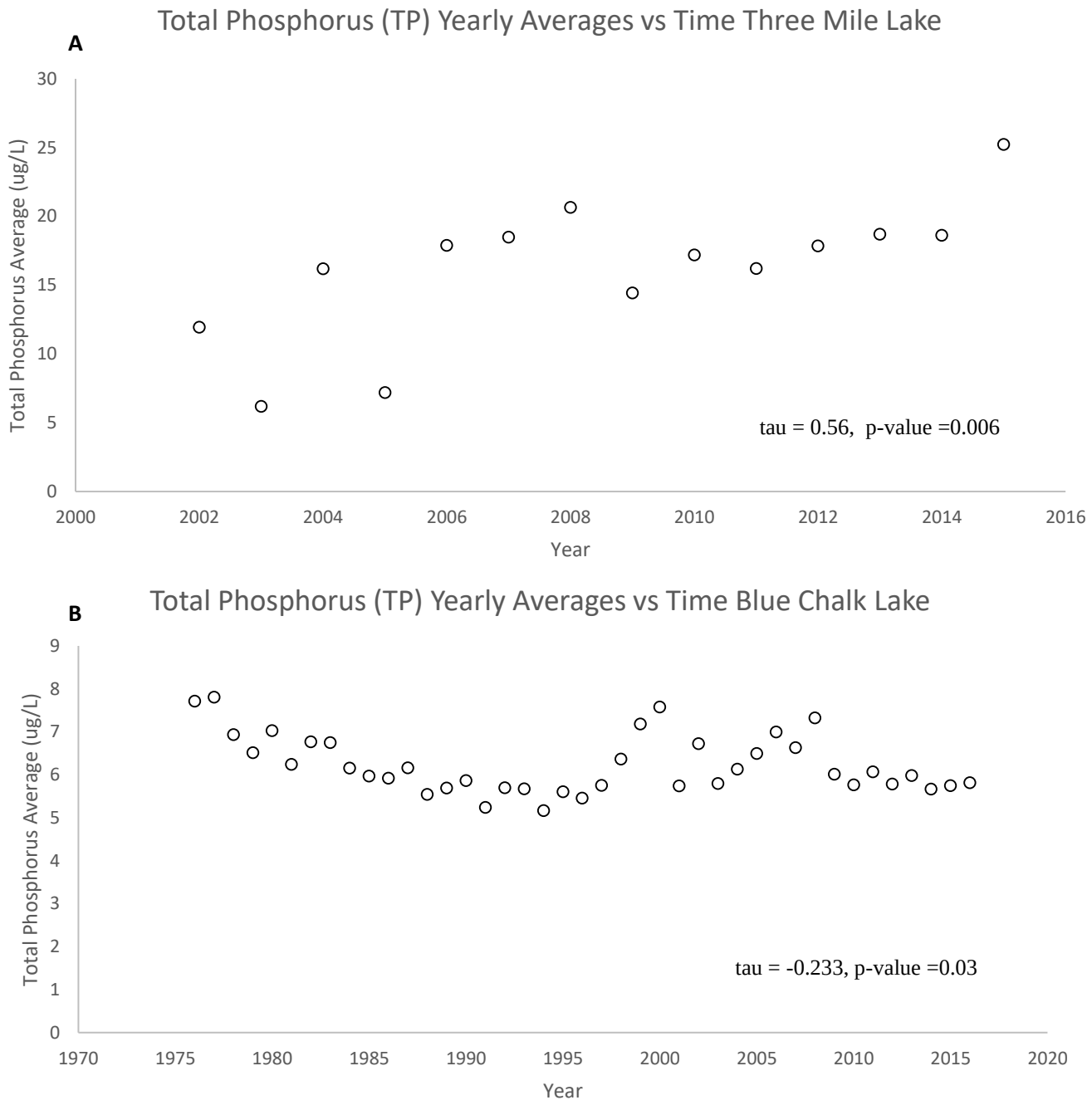
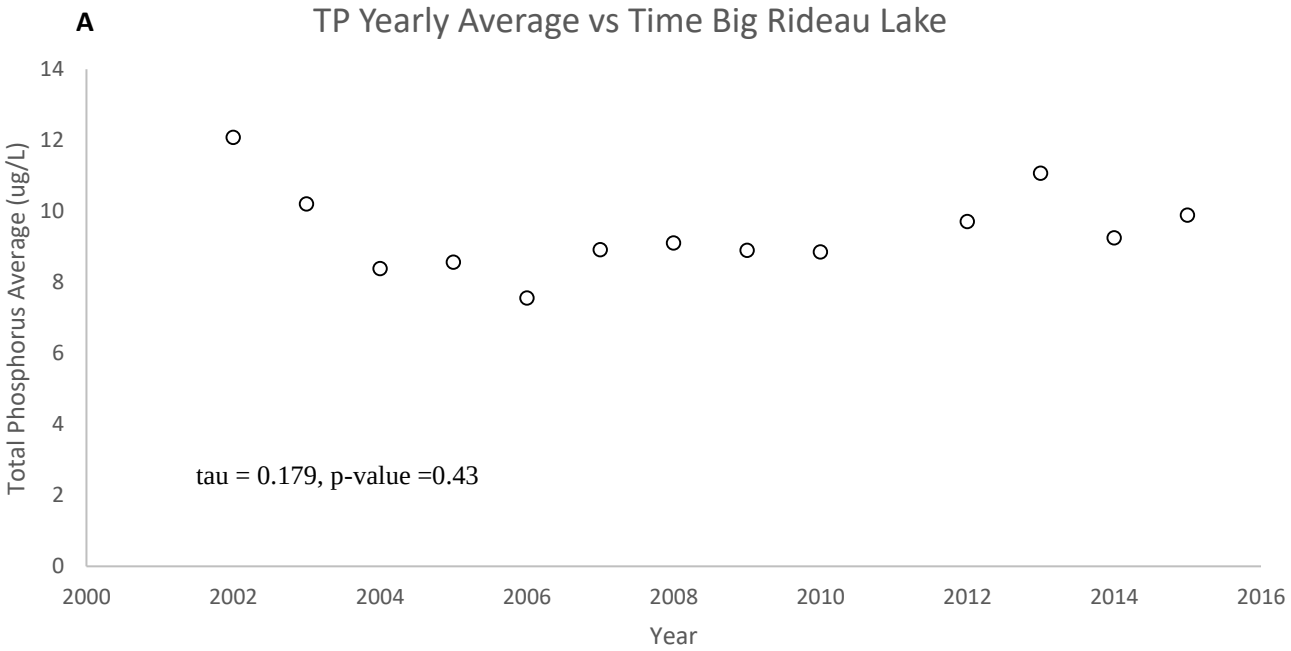


Figure 9 Average of (n = 8-10) yearly TP (µg/L) through (Year) in the Muskoka Region with Mann-Kendall trend test results, tau indicates directionality and strength. Results are considered significant when *p*= <0.05. **A):** Three Mile lake TP over time with significance of (*p*= 0.0062) and positive (tau= >0) trend over time. Confirmed with 500 permutation bootstrap analysis. **B):** Blue Chalk Lake TP over time which shows significance (*p*= 0.03), negative (tau=0>) trend over time. Confirmed with 500 permutation bootstrap analysis.

Figure 10



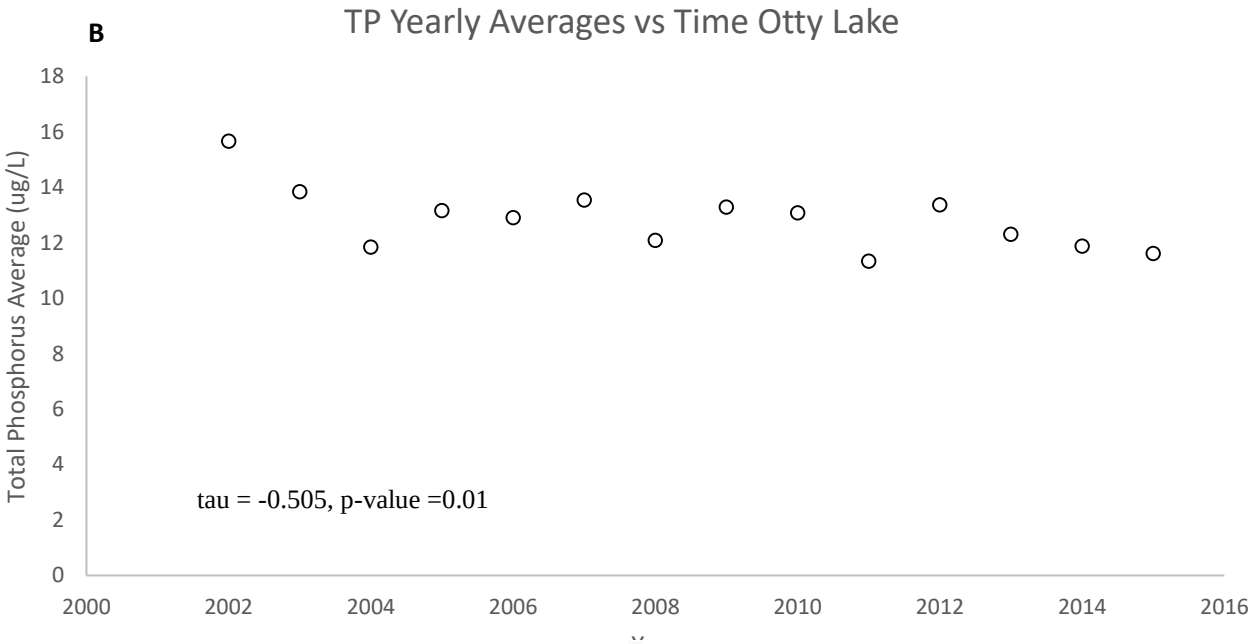


Figure 10: Average (n = 4-10) yearly TP measurements (µg/L) over time (Year) in the Rideau Region with Mann-Kendall trend test results. Tau value indicates directionality and strength. Results are considered significant when (p= <0.05). **A):** Big TP over time with no significant trend was found (p=>0.05). Confirmed through 500 permutation bootstrap analysis. **B)**: Otty Lake TP over time with significance of (p= 0.01) and negative (tau= 0>) trend over time. Confirmed with 500 permutation bootstrap analysis.

Figure 11

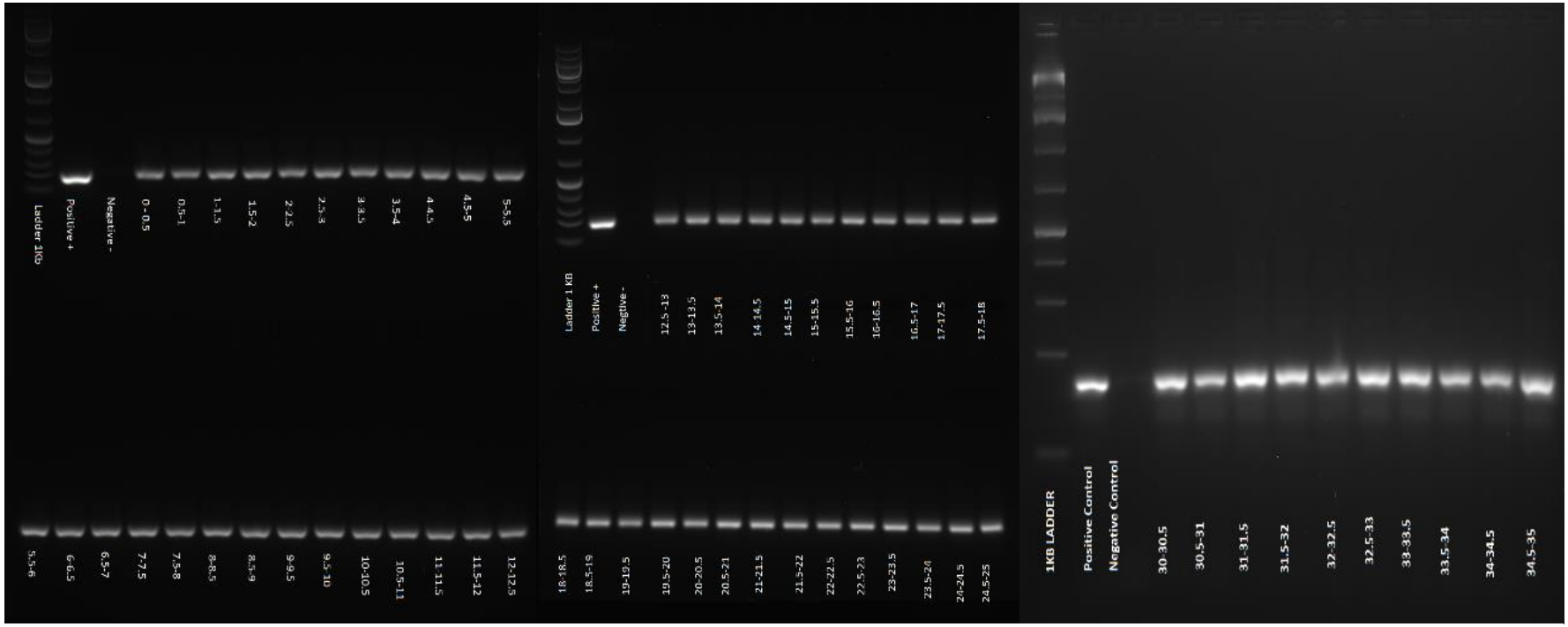


Figure 11. Examples of endpoint PCR results for glnA amplification from Blue Chalk Lake (Left), Three *M*ile Lake (Center) and Big Rideau Lake (Right) run on 1.5% agarose gels, impregnated with GelRed™ nucleic acid stain, suspended in 1xTAE Buffer at 120 Volts for approximately 40-60 minutes.

Figure 12

Three Mile

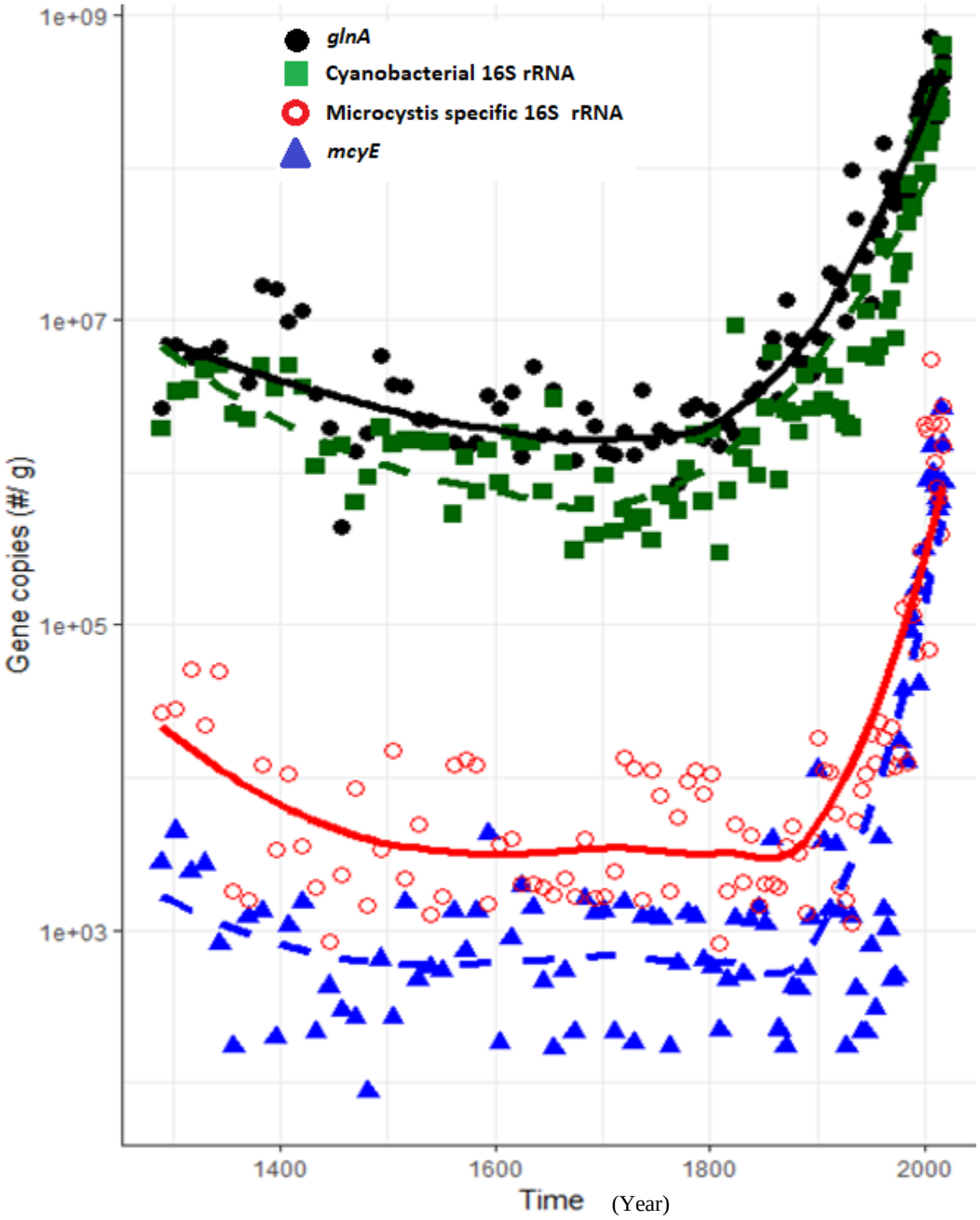


Figure 12: Gene copy numbers normalized per gram of sediment as a function of time for Three Mile Lake. Gene copies/g of sediment of *glnA* and CYA experienced generalized increases from the early 1800's. The *MICR* and *mcyE* gene copies remained near detection increasing rapidly in recent years. Note logarithmic scale. Lines fitted using "LOESS" method.

Figure 13

Blue Chalk

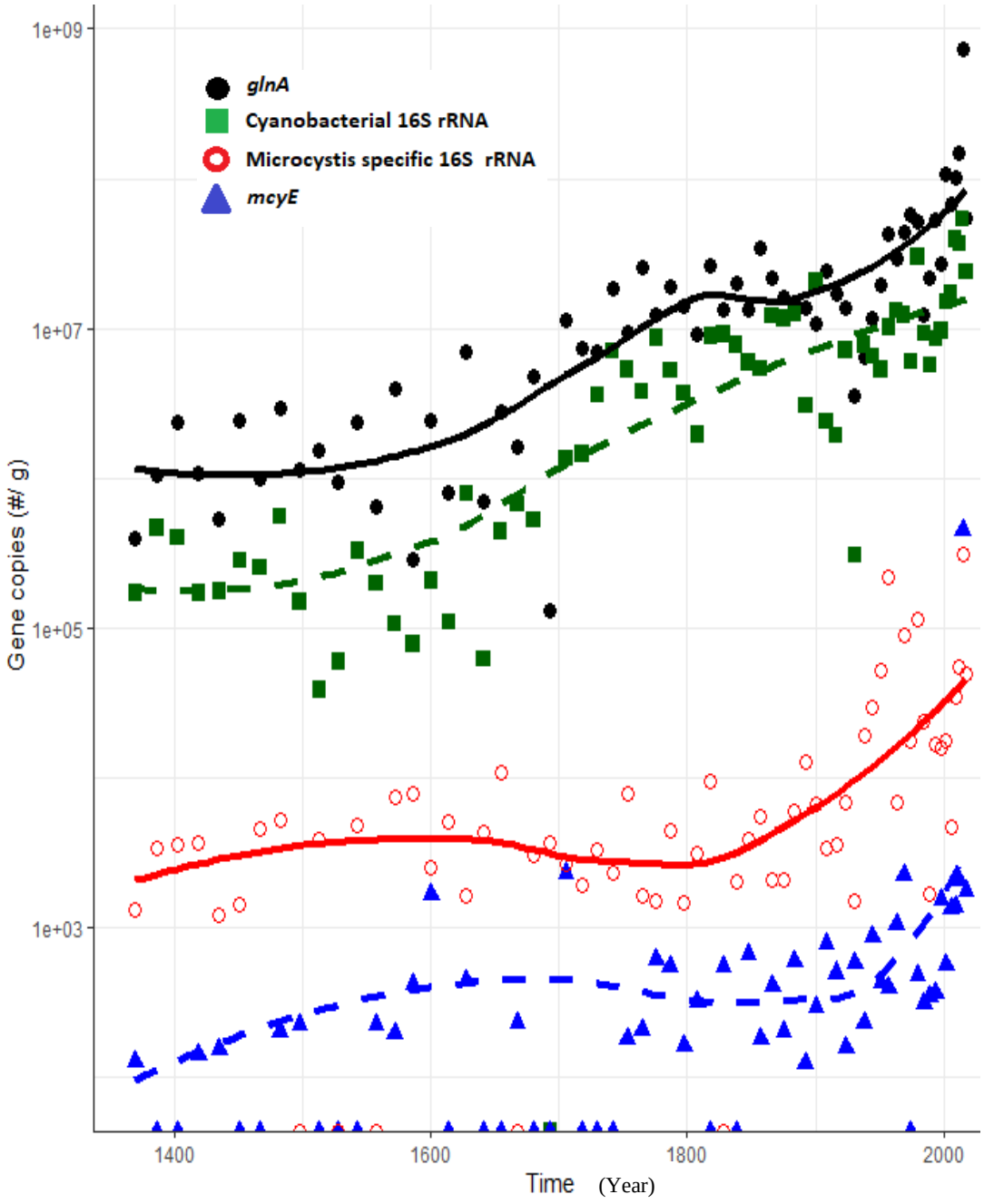


Figure 13: Gene copy numbers normalized per gram of as a function of time for Blue Chalk Lake. All gene targets showed generalized increases with time with the exception of *mcyE* which remained near detection throughout the entire time period. Note logarithmic scale. Lines fitted using "LOESS" method.

Figure 14

Big Rideau

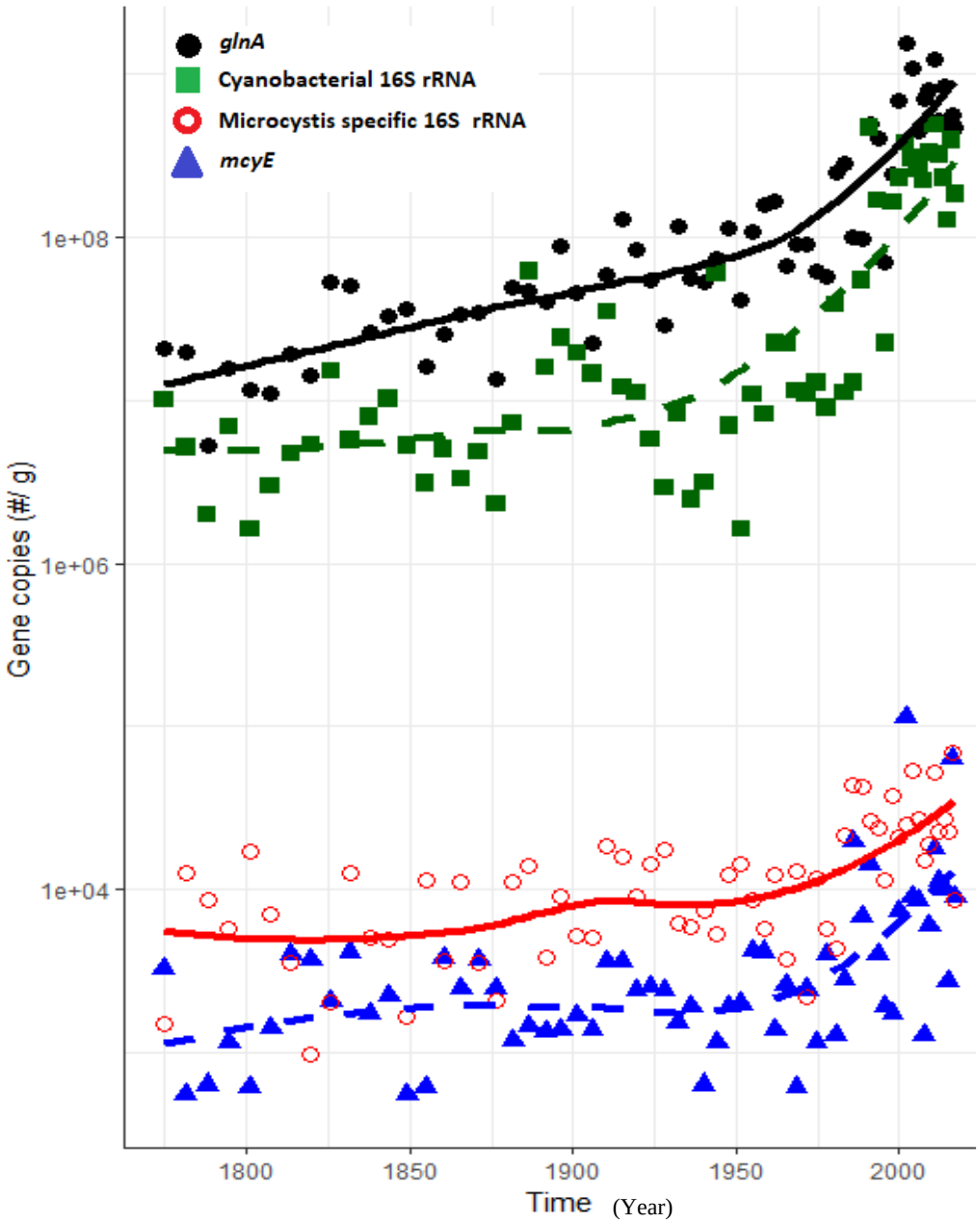


Figure 14: Gene copy numbers normalized per gram of sediment as a function of time for Big Rideau Lake. The *glnA* and CYA targets experienced increases in copy numbers per gram of sediment beginning in the early 1900's. The *MICR* and *mcyE* targets remained close to detection until the early 2000's Note logarithmic scale. Lines fitted using "LOESS" method.

Figure 15

Otty Lake

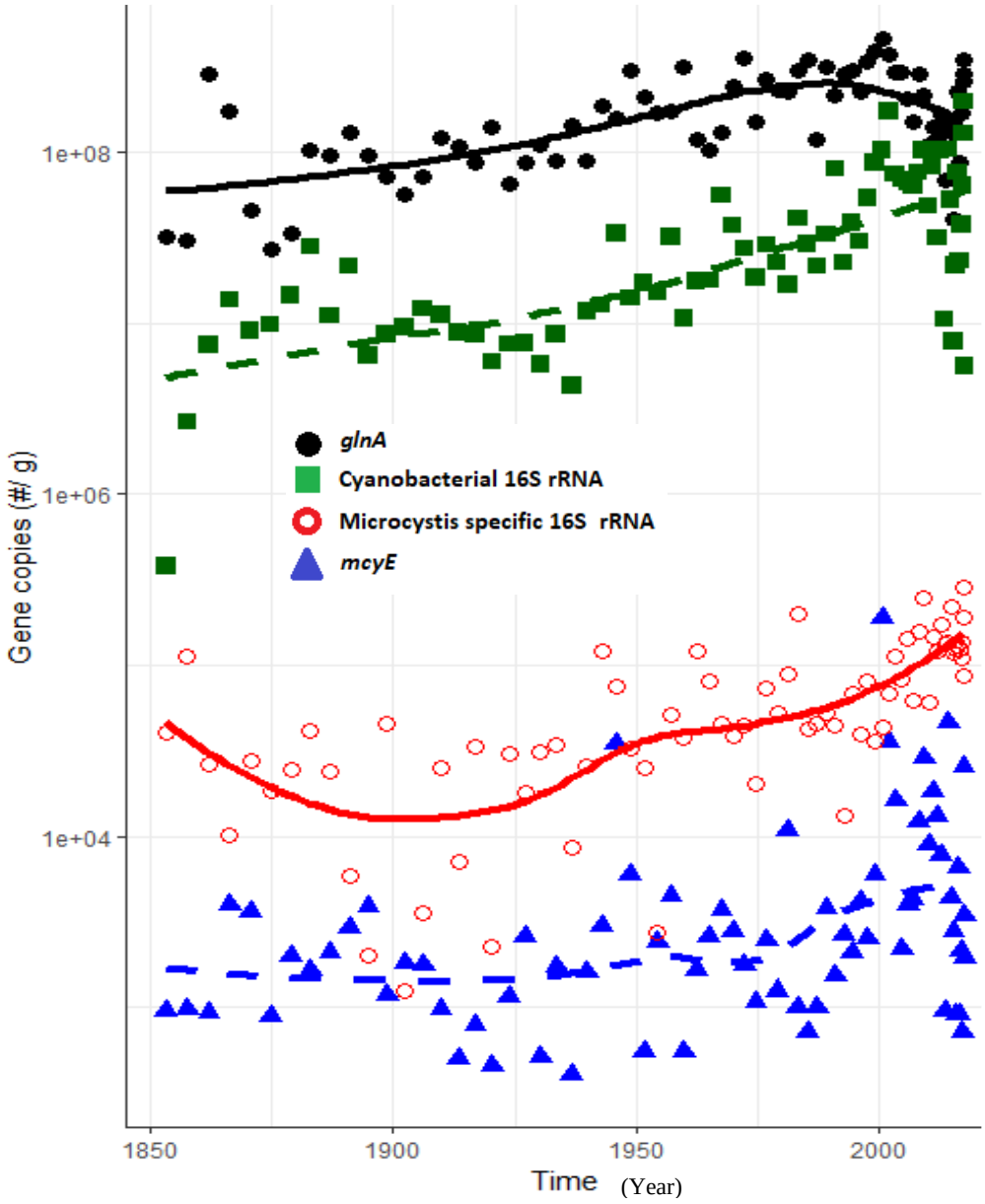


Figure 15: Gene copy numbers normalized per gram of as a function of time in Otty Lake. The *glnA* target experienced a gradual increase over time, decreasing in recent years (2000's). The CYA target continued to increase into the early 2000's. The *MICR* target increased at a similar rate to the more general CYA target. The toxin gene, *mcyE*, remained near detection only reaching detectable levels in most recent years. Note logarithmic scale. Lines fitted using "LOESS" method.

Figure 16

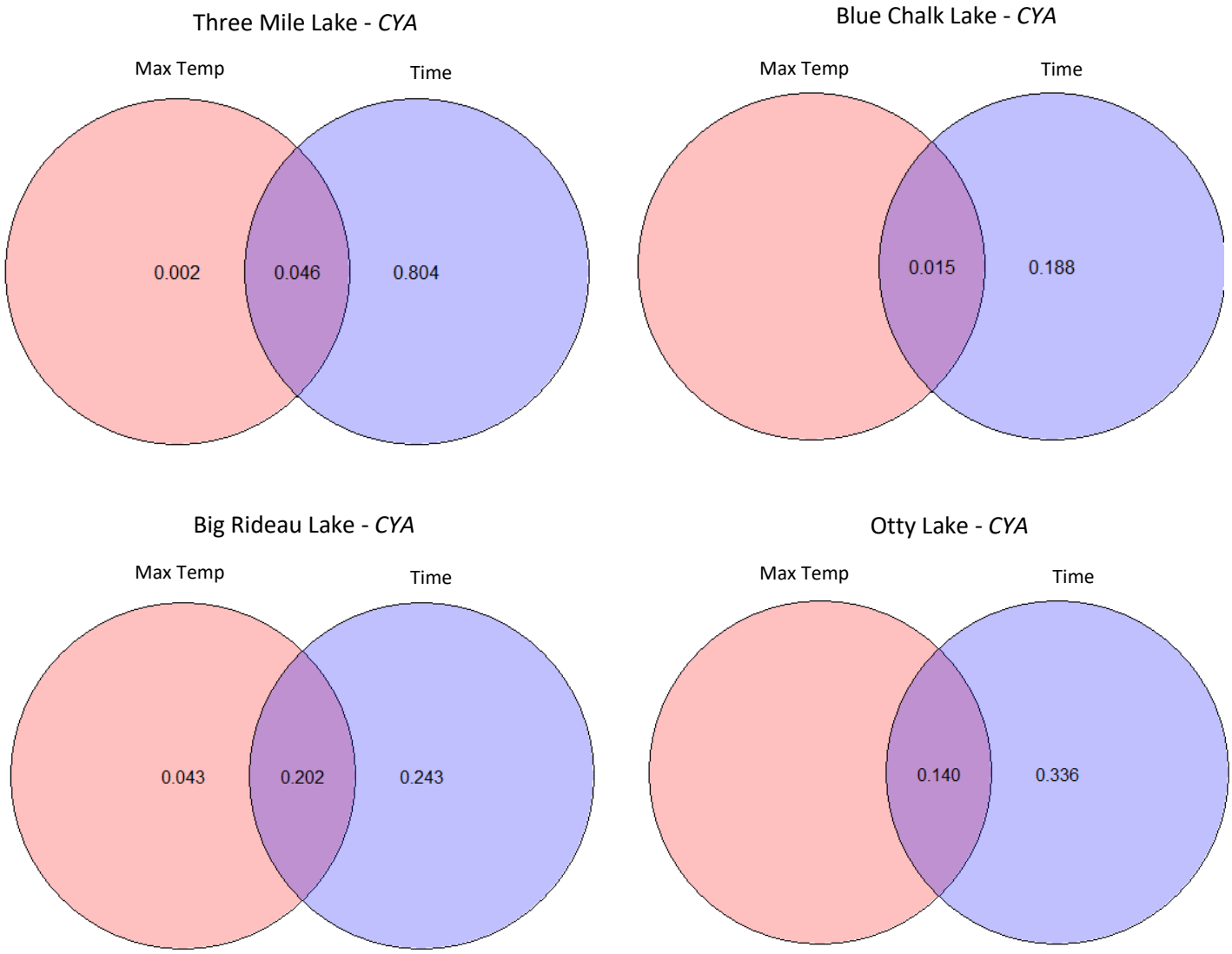


Figure 16 Example of partitioning of variation outputs showing the proportion of gene variation explained by maximum temperature (red) and time percent (blue) and the combination of the two factors (purple) of the *CYA* gene copy number in four lakes. **Top Left:** Output for Three Mile lake majority of variation explained by time (blue) with ~4.6% (purple) being explained by the combination of time and maximum temperature (red). **Top Right:** Blue Chalk lake minimal variation is explained by time (blue), maximum temperature (red) or even the combination of the factors (purple). **Bottom Left:** Big Rideau lake. The rideau region lakes had a much higher explanation of *CYA* variation by climate factors. In this case the combination (purple) between maximum temperature (red) and time (blue) explained 20.2% of gene variation. **Bottom Right:** Otty lake. Similar to Big Rideau, climate factors like maximum temperature explained a much higher percentage of the variation in selected gene targets. In this case the combination (purple) between maximum temperature (red) and time (blue) explained ~14% of the variation in *CYA* in Otty lake.

Figure 17

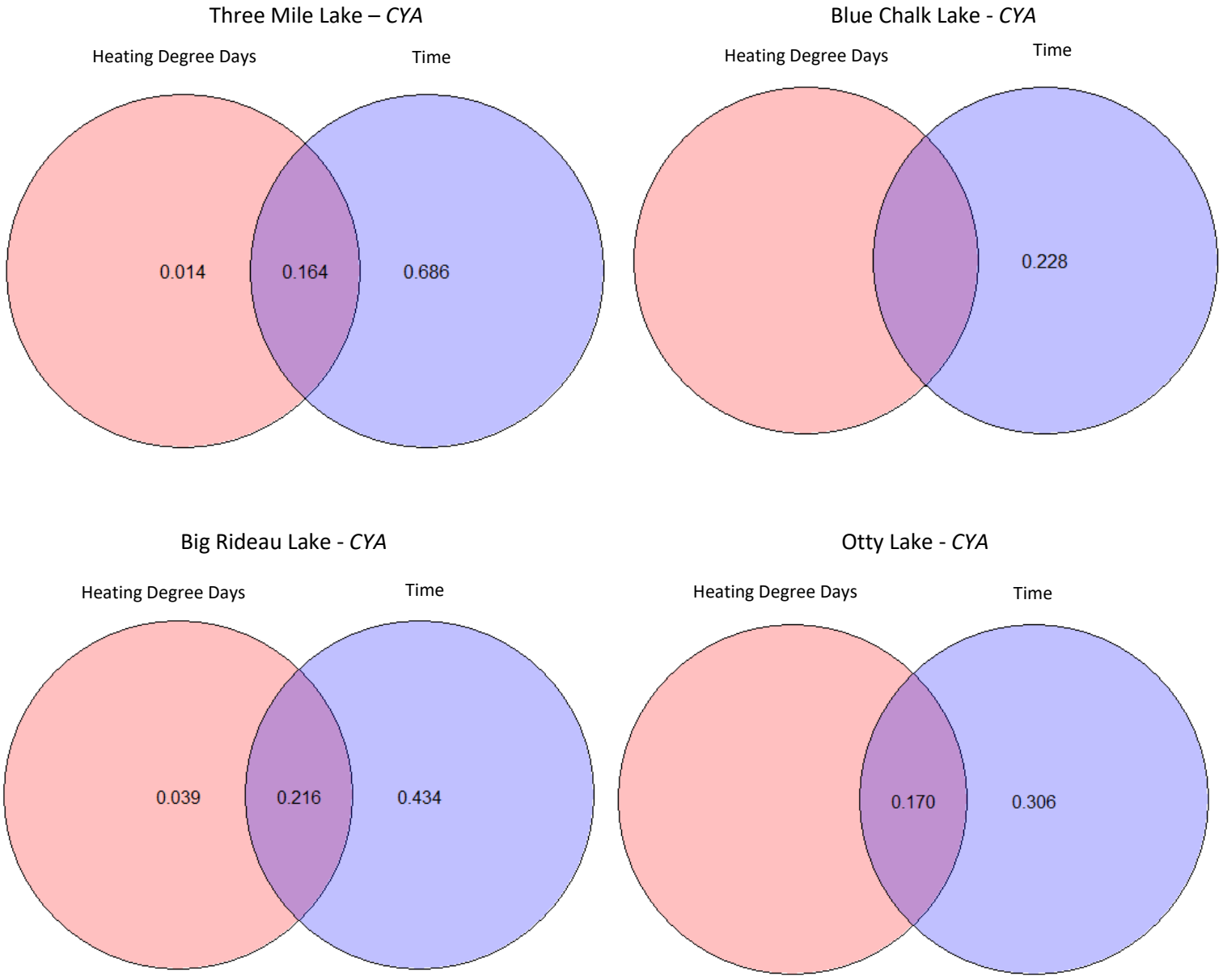


Figure 17: Example of partitioning of variation outputs using heating degree days (red) and time (blue) and the combination of the two factors (purple) in the explanation of variation (in % explained) of the *CYA* gene copy number in three lakes. **Top Left:** Output for Three Mile lake with ~16.4% of variation (purple) being explained by the combination of time (blue) and heating degree days (red). **Top Right:** Output for Blue Chalk Lake. Time (blue) was the only explanation of variation of *CYA* with 22.8% explained **Bottom Left:** Output for Big Rideau lake. The rideau region lakes had a much higher explanation of *CYA* variation by climate factors. In this case the combination (purple) between maximum temperature (red) and time (blue) explained 21.6% of gene variation. **Bottom Right:** Output for Otty lake. Similar to Big Rideau, climate factors like maximum temperature explained a much higher percentage of the variation in selected gene targets. In this case the combination (purple) between maximum temperature (red) and time (blue) explained ~17% of the variation in *CYA* in Otty lake.

Chapter 3: Effects of Lake Morphometry on the Preservation of Sediment DNA

3.1 Introduction:

Morphometry is a fundamental characteristic of lacustrine ecosystems that has significant direct and indirect effects on the physics, chemistry and biology of lakes (Kalff 2002). The surface area and shape of basins or bays, as well as their depth, can have a significant effect on turbulence and flow and by extension the rate at which particulates re-enter and exit the water column and the overall sedimentation of particles to the bottom. (Kalff 2002).

One of the most important effects of water column depth is on lake thermal regimes. Deep temperate lakes typically form, at depth, a thermocline during late spring through to early fall, leading to a separation of less dense surface waters from denser, colder, deep waters. As a result, the bottom waters of deep temperate lakes maintain a quasi-constant temperature of ~4°*C* Such low temperatures are favorable for the preservation of organic matter as well as inorganic material that has settled onto sediments. Basin formations and surface area also affect the length and depth of stratification as well as the specific temperatures found in deeper waters (Butcher *et al.* 2015). Depth may also play a secondary role in preserving organic matter by reducing ultraviolet radiation through its absorption by dissolved organic carbon and dispersion by the water itself (Fleischmann 1989; Kirk 1994; Rose *et al*. 2009). Stratification within highly productive lakes frequently leads to the development of low oxygen conditions within hypolimnetic waters and at the sediment water interface (Nürnberg 1995; Müller *et al*. 2005). As microbial communities carry out activities at the sediment interface, they rapidly deplete available oxygen ultimately causing complete anoxia (Nürnberg 1995; Müller *et al*. 2005), which may reduce oxidative stress on organic compounds within sediments.

Lake basin size, shape and slope pitch have a significant effect on how sediments and particulate matter exit the water column and deposit within the sediments. Utilizing ^{210}Lead isotopes, an atmospheric isotope marker, Blais *et al*. (1995) found that basin formations with steep walls caused a "sediment-focus" increasing the volume of particulate entering the deepest sites of lakes. These findings have implications with regards to how paleolimnological studies are conducted and results interpreted. For example, Moos *et al*. (2005) found that lake sites at depths greater than 9-14 m were the most ideal for reconstructing diatom assemblages from the sediment record. For most paleolimnological studies, sediment cores are usually taken from the deepest location in lakes because of sediment focusing and the best conditions for preservation of fossils and organic matter pertinent to the reconstruction of environmental change (Smol 2008).

Paleolimnological researchers making use of sediment preserved DNA (sedDNA) have been working under the assumption that DNA preserved in lacustrine environments are subjected to similar pressures to DNA preserved in more well studied terrestrial environments. Exposed DNA can degrade relatively quickly due to a variety of conditions present in many environments and include: enzymatic damage, direct and indirect ionizing radiation damage, oxidative stress as well as various mechanical processes. Temperature plays a key role, increasing the rates of damage accumulation and exacerbating enzymatic activities by other heterotrophic bacteria, leading to rapid degradation of environmental sequences (Zhu 2006; Strickler *et al*. 2015). Ultra-violet radiation exposure can directly produce damage in DNA segments (UV-B causing pyrimidine dimer formations) or form free radicals within cells that can cause direct physical damage to the DNA structure (UV-A) rendering them unusable relatively quickly, which has led to the widespread use of UV-B as a surface sterilization tool (Strickler *et al*. 2015). Abiotic damage imparted by extremes in pH or exposure to strong chemical agents, as well as

mechanical damage, such as the formation of ice crystals or shearing are also important sources of damage to environmental DNA. All of these damages can accumulate as the genetic material is not maintained by a living organism's repair system leading to permanent alterations and damage. It was long assumed that although DNA could be preserved for short periods of time its integrity would not persist for longer periods of time in exposed environments (Colson *et al.* 1997; Willerslev & Cooper 2004; Gilbert *et al*. 2005; Binladen *et al*. 2006; Poulakakis *et al.* 2006; Dabney *et al.* 2013). However, simulations and studies of DNA degradation kinetics in ancient, terrestrial, bone samples have revealed that DNA can retain stability, in smaller segment lengths, for longer periods (estimated half-life of 100bp lengths at -5 degrees Celsius being ~47 000 years) of time under certain conditions (Allentoft *et al.* 2012). These simulations have estimated that preservations rarely exceed 1-2 MYA (Allentoft *et al.* 2012). As a result, most paleoDNA studies of lake sediments have chosen sites for sediment core retrieval, which, in theory, would exhibit conditions that ensure better sedDNA preservation (Domaizon *et al*. 2017). Lake sites of extreme depth (to reduce oxygenic, UV as well as thermal pressures), have generally garnered positive results, yielding DNA capable of being not only amplified but also quantified and even sequenced for paleo-community analyses (Monchamp *et al.* 2016; Domaizon *et al.* 2017; Pilon *et al*. 2019). There are no studies to date that have specifically examined the effect of depth on sedDNA preservation and whether paleoDNA techniques can be applied to shallow and or tropical lake systems.

The objective of this study was to 1) examine the role of lake depth and morphometry in the preservation of sedDNA. Cores were gathered from four distinct areas of differing depth and degree of shelter from within a single lake (Big Rideau, Central Ontario). If depth is important, shallower sites were expected to contain lower quality DNA and lower amplification of target

genes at the bottom of cores whereas surficial sediments might be expected to have similar quality DNA and gene copy numbers (and possibly higher if benthic productivity is significant) compared to deep sites. The second objective was to develop novel primers targeting the allophycocyanin alpha chain (blue-green pigment) gene target in *Synechococcus*. The genus *Synechococcus* is ubiquitous in both marine and freshwater systems, with the exception of acidic environments and highly eutrophic lakes (Hauschild *et al*. 1991; Pick 1991). The genus also preserves in sediments as demonstrated using sediment DNA based reconstruction of population dynamics over ~ 200 years (Domaizon *et al*. 2013). The primers were implemented to assess the abundance and preservation of *Synechococcus* under various morphometric conditions and to compare trends with trends found in other gene targets (*glnA* & CYA).

3.2 Materials and Methods

3.2.1 Site Selection and Sediment Core Sampling

Big Rideau Lake is a large lake in the Rideau Lakes district of Eastern Ontario. It has multiple bays of varying depths and is thus suitable as a test site for the effect of lake depth on sediment DNA preservation. Locations for sediment coring were selected with the aid of hydrogeographic or fishing maps (Canadian Hydrogeographic Service Ottawa, Canada) as well as expert opinion. Cores were obtained from basins of 2.7 m (Green Lake area), 8.3 m (Little Lake area), 10.6 m (Cow Island area, and 21.3 m (Big Rideau Corridor area) (Table 5).

Coring was carried out as described in Chapter 2 Methods – Lake Sediment Sampling of this thesis. Sectioning and section storage were carried out using sterile techniques as described in Chapter 2 Methods – Lake Sediment Sampling.

3.2.2 DNA Extraction, Quality and Integrity

Extractions of nucleic acids from sediments were carried out utilizing a Qiagen Powersoil Kit™ with minor modifications (see: Chapter 2 – DNA Extraction). Samples of extracted DNA were stored at -20°C in sealed 2 mL tubes until further analysis.

DNA quality was first ascertained utilizing a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific, 2019) and analyzing the ratio of absorbance at 260 to 280 nm (see Chapter 2 – Quality and Integrity Analyses). Samples were further analyzed for nucleic acid concentration utilizing Quant-iT™ dsDNA High Sensitivity assay kit (Thermo Fisher, 2019). Samples were prepared on clear, Whatman® 96-well plates with flat-bottom wells (Whatman, 2019). Three sets of 8 wells were set aside for triplicate standards of known DNA concentrations. Then 1μL of each DNA extraction sample was placed into an individual well until each well on the plate contained a sample. The 24 wells set aside for known DNA standards of various concentrations were then inoculated such that 10 μL of one standard (0, 5, 10, 20, 40, 60, 80, 100 ng/μL) was in an individual well and each concentration was represented in triplicate. Each well in the plate was then given an additional volume of 200 μL of Quant-iT™ dsDNA HS reagent. Plates were covered fully with tin foil to reduce light exposure of the dye and incubated on an orbital plate shaker at low speed for 2 minutes. Samples were then placed into a Cytation™ 3 Fluorescent plate reader from BioTek (Biotek, 2019). The reader was set for an excitation maximum at 480 nm and an emission maximum at 530 nm, as outlined in the manual for the Quant-iT™ dsDNA High Sensitivity assay kit. Fluorescence values were then corrected by subtracting the values acquired from the 0 ng/μL standard's fluorescence value. The triplicate standards were then utilized to construct a standard curve and assigned a ng/μL DNA values to the remaining wells on the plate. This procedure was repeated until all extracted sediment sections were analyzed.

3.2.3 Primer Design

A gene targeting the photosynthetic pigment allophycocyanin (alpha chain) found in many cyanobacteria was first accessed through UniProtKB genetic and protein databases (UniProtKB 2019). Accession numbers are listed in Supplemental Information. Protein files were assessed for nucleic acid sequences utilizing R scripts Extracc_all.R and EMBL_Fetch.R (Lopez & McWilliam 2013; McWilliam *et al*. 2013) in combination with the "BioStrings" (Pagès 2019) and "BiocManager" built and added to the "BioConductor" repository for R version 3.5.2. After extracting and loading both scripts they were run using the following commands in a Linux (Linux, 2019) based setting, running Ubuntu 18.10 (Ubuntu, 2019), virtualized using VirtualBox (OracleOM, 2019) command terminal:

"/home/user/Desktop/EMBL_Fetch.R - -No-Save – f /home/user/Desktop/allophycocyaninAlpha.list"

A total of 615 associated nucleotide profiles between 40 and 480 base-pairs in length were pulled from the UniProtKB database as a comma-separated file and merged into a multi-fasta file format. Targets pulled by the search engine but not of the desired target itself were removed manually after visual curation of the list by entry. Sequences were then clustered for similarity utilizing VSEARCH software (Rognes *et al.* 2016) set to 0.99 similarity with centroids, and run on windows operating system using input command:

"/computer/Desktop/Vsearch.exe --clusterfast /computer/Desktop/filename.fasta --id 0.99 –centroids /computer/Desktop/output.fasta" (reference vsearch)

Clustering was done on Windows 10 operating system via the Console terminal and re-examined visually on UGene Pro (Unipro, 2019) visualization software. Once sequences were

clustered the output file was transferred back to Linux Virtualbox virtual generated console. The following scripts required a Linux Unix system to function and were run using the Terminal prompter. Sequences were initially aligned utilizing the TRANSLATORX (Abascal *et al.* 2010) PEARL-script program which was accessed and run through Linux command terminal with the following commands:

"perl ./translator_vLocal.pl -i /home/User/Desktop/allophycocyaninA.fasta -c 11 -o /home/User/Desktop/InitialAlignAllophycoA.fasta"

The alignment was then checked visually utilizing the UGene Pro software. The fasta files were then further aligned utilizing MUSCLE.exe command script (Edgar 2004) on virtual hosted Ubuntu 18.10. The MUSCLE program was executed in Terminal using input command:

"/home/computer/Desktop/muscle -in /home/computer/Desktop/inputfile.fasta -aa -maxiters 5 -out/home/computer/Desktop/alignment.fasta"

The MUSCLE software was set to change codon triplets to corresponding amino acids (-aa command) align them then convert the resulting alignment back to nucleic acids. This is done as several codon triplets can correspond to the same amino acid residue and this has been found to enhance the accuracy of the alignment. A maximum number of iterations (-maxiters 5) was set to 5 to reduce strain on CPU, GPU and time resources which can rapidly become overwhelmed with the number of sequences present. The alignment was then further refined in MUSCLE utilizing the '-refine' function:

". /MUSCLE – in /home/User/Desktop/alignment.fasta -out /home/User/Desktop/Refined.fasta -refine"

Sequences were then trimmed utilizing DegePrime software (Hugerth *et al*. 2014) DegePrime_Trim.exe using input command:

“. /Desktop$ perl /home/User/Desktop/TrimAlignment.ppl -i /home/User/Desktop/ApcARefine.fasta -o /home/User/Desktop/RefineTrim.fasta”

The alignment was then analyzed for best possible forward primers utilizing the Degeprime.exe application and the results were output as a .txt list file using the following command in Ubuntu terminal prompt:

“. /Desktop$ perl /home/User/Desktop/DegePrime.pl -i /home/User/Desktop/RefineTrim.fasta -d 27 -l 21 -skip 0 -o /home/User/Desktop/apcaForwards.txt”

The extra instructions -d 27 indicates a maximum degeneracy score of 27. The command -l indicates the maximum length of the primer at 21 base pairs. The command -skip 0 indicates that no sequences were to be skipped. Degenerate primer sequences were analyzed and compared to the initial alignment using the .txt primer output file as well as the trimmed alignment FASTA viewed in UGene pro.

The candidate primers were then examined using guidelines and oligo design tools available on the Integrated DNA Technologies (IDT) website (IDT, 2019) for sequences exhibiting ideal primer characteristics i.e.: a melting temperature (Tm) 55-65 °C, annealing temperatures (Ta) of ~5°C below the Tm,18-25 bp length and be composed of ~35-65% guanine and cysteine residues with few guanine repeats. Reverse primers were created by inputting the forward sequences into a reverse mirroring website program located on Harvard’s laboratory website: *http://arep.med.harvard.edu/labgc/adnan/projects/Utilities/revcomp.html/* (Harvard, 2019) and

subjected to the same analyses as outlined above. Pairs were chosen based on location within the sequence (~150-300bp apart) as well as for similarities in Tm's to the opposing primer.

Once a significant number of primer sequences (~10) were compiled, they were analyzed using IDT Oligo tools. The tools generated estimates for Tm's and Ta's as well as identify potential loop (hairpins etc.), self-annealing sites or cross-bound structures (between pairs). Once satisfied with the criteria the primer pairs that passed each test were ordered and tested physically through the use of endpoint PCR. Each PCR reaction required several positive controls, which should contain the targeted allophycocyanin alpha gene segment. The cultures utilized for positive controls were cyanobacterial strains CPCC300, *Microcystis aeroginosa* and CPCC663 a *Synechococcu*s sp., both acquired from the University of Waterloo based Canadian Phycological Culture Collection (CPCC, 2018). The primers were run along a temperature gradient to find the ideal annealing temperature on a Bio-Rad S1000 Thermal Cycler. Once an optimum temperature was determined, a standard endpoint PCR reaction were run with a positive control of CPCC663 genomic DNA and negative controls of deionized water (NTC) and Chlorella gDNA (reaction outline in Appendix B: Table 1b).

The reactions, once completed, were then run on a 1.5% agarose gel in 1x TAE at 107 volts for ~35 minutes and the results examined for single band amplification in the positive lane on an MBI Lab Equipment Fusion FX5 UV visualizer with photos captured and edited utilizing Fusion Molecular Imaging software (Montreal-Biotech, 2019) and Microsoft Paint 3D. A temperature gradient was repeated prior to ddPCR on a C1000 Touch Thermocycler (Bio-Rad, 2019) to ensure that the annealing temperature remained consistent for ddPCR quantification reactions utilizing the reaction parameter outline in Appendix A: Table 6a.

PCR products from the primer pairs F1 & R7 as well as pair F1 & R10 were sent to StemCore laboratories for sequencing (OHRI, 2019). Pure culture template was acquired via a DNA extraction of CPCC663 culture material. Template genomic DNA was amplified utilizing primer pairs F1-R7 and F1-R10 (sequences available in Appendix A: Table 1a) with the protocol for outlined in Appendix B: Table 2b. The products of each of these endpoint reactions were then separated on a 1.5% agarose gel, suspended in 1x TAE buffer at ~100 volts for ~45 minutes. Products were identified as bands in the gel utilizing a UV transillumination device and separated from the gel with a small scalpel. Gel pieces containing product were placed into separate sterile, pre-weighed 2 mL microcentrifuge tubes. The gel segments were then treated and extracted as outlined in the Bio Basic Inc. gel extraction kit instructions (Bio Basic Inc., 2019). Product was then analyzed utilizing NanoDrop™ spectrophotometry for both quality (via 260:280 nm ratio) as well as a rough estimate of quantity (ng/µL). Product was then diluted to 15uL of 1ng/µL concentration as per StemCore instructions. The tubes were then stored frozen in a sterile 2mL microcentrifuge tube. Product sequences were analyzed on UGene Pro software. Sequenced products were then evaluated through an NCBI Blast Nucleotide search to confirm the target specificity of the primers.

3.2.4 Gene Copy Quantification

Samples were quantified utilizing the Droplet Digital Polymerase Chain Reaction system from Bio Rad (Bio Rad, 2019). Reactions were prepared and assessed in the same manner as those outlined in Chapter 2 Methods – Quantitative PCR**.** Primer sequences and conditions for ddPCR can be located in Appendix A: Table 1a and Appendix A & B Table: 7a, 8a & 2b for *glnA*, CYA and *apcA*. Once completed, plates were transferred to a QX200™ Droplet Reader and the results were further analyzed utilizing QuantaSoft™ software (Bio Rad, 2019). An

example output for *apcA* analysis utilizing the F1-R7 primer set can be seen in Appendix B Figure 1b.

3.2.5 Statistical Analyses:

All statistical analyses were carried out in R (version 3.5.2) and RStudio software (version 1.0.153). All results were assessed for normality using the Shapiro.Test, histograms as well as the qqnorm functions of the 'stats' package. Result distributions not considered normally distributed were $\log_{10}$ or square-root transformed to render them more normal.

The relationship between raw extraction values (ng/µL) determined by both Nanodrop™ vs. Quant-iT® methodologies, was assessed using the 'cor.test' function with "method= pearson" of the 'stats' package. Results were visualized using the 'ggscatter' function from the 'ggpubr' package.

The cores from four different sites were divided into three groups: top (0-5 cm), middle (20-25cm) and bottom (33-40cm) for comparison of means using ANOVA testing through the 'aov' function from the 'stats' package. The results of the 'aov' were further assessed with a post-hoc test to identify which groups differed the most significantly using the 'TukeyHSD' function from the 'stats' package. Results were visualized using the 'boxplot' function from the 'graphics' package.

3.3 Results

3.3.1 Sediment DNA Quality and Integrity

Quality and integrity of each DNA extraction was initially assessed by examining absorbance ratios of 260: 280 nm wavelengths on a NanoDrop™ spectrophotometer. All sediment DNA extractions yielded 260: 280 nm ratio values between 1.60-2.2 (1.8 being considered "pure" DNA) indicating that DNA was of useable purity. After this initial

assessment, each sample was amplified with a primer set of a known target (*glnA*) utilizing endpoint PCR (Fig 11). All samples utilized in this study produced an amplicon of appropriate size for the desired target, indicating that the integrity of the DNA was high enough to allow for further quantification.

3.3.2 DNA Quantity: NanoDrop™ vs. Quant-iT® Methodologies

The concentrations of extracted DNA were generally similar between the four sites with DNA quantity decreasing steadily down the core from top (most recently deposited) to bottom (deposited the furthest back in time) (Appendix B: Figure 2b). The top sections yielded ~60 ng/µL on average. However, the most sheltered site, Little Lake, yielded significantly more DNA (~130 ng/µL). The bottom of the cores had the lowest DNA yields and results were more similar across all the cores with the average concentrations ranging from 7 to 12 ng/µL. The 260: 280 nm ratio (frequently used to quickly assess sample purity) increased in variability with depth (Appendix B: Figure 3b). The Quant-iT® fluorescent assay kit yielded similar results to those presented by NanoDrop™ spectrophotometry, however, it was generally more conservative (Appendix B: Figure 4b**)**. The measured nucleic acid quantities from the two methodologies were highly correlated (Fig. 18, rho= 0.97, $p=2.2x10^{-16}$).

3.3.3 Top, Middle and Bottom Comparisons Between Sites of Different Morphometry

No analyses were carried out comparing the top, middle and bottom portions down core as it was expected that values would decrease over time due to a variety of environmental factors and degradation of DNA materials over time. The top, middle and bottom of the cores were compared across sites using an ANOVA test (Tables 6, 7, 8, 9**)** and further analyzed using post-hoc Tukey's HSD analysis (Tables 6, 7, 8, 9**)** as well as graphically via boxplots (Fig 19, 20, 21, 22).

3.3.3.1 Total DNA Extracted:

Due to the high correlation between the NanoDrop™ and Quant-iT® kit evaluations of DNA content for the samples only NanoDrop™ extraction values were utilized in comparisons between sites. For the top portions of the core (represented by sections between 0-5 cm of each core) ANOVA analysis indicated significant differences among the top portions of cores (Table 6), the Little Lake site had significantly more DNA extracted (~130 ng/µL), while all other sites cored returned approximately the same values (~60 ng/µL).

The middle portions, comprising 20-25 cm of the core depth once again showed differences between the mean extraction values between sites differed (Table 6**)**. Graphical analysis (Fig 19**)** and post-hoc analysis indicated that Little Lake had significantly more DNA extracted per sediment sections (~ 60 ng/µL), that Corridor and Cow Island both returned similar DNA extraction values (~ 20 ng/µL) and that the shallowest location, the Green Lake site(depth ~2.7m), returned the least amount of DNA per section (~13 ng/µL) (Fig 19).

The bottom of the core was represented by extractions of sediment sections between 35-40 cm of each core. ANOVA analysis indicated that the mean extraction values differed between coring sites with both the deepest site (Corridor) and the most sheltered site (Little Lake) showing relatively similar DNA levels (~12 ng/µL) (Table 6**,** Fig 19). Meanwhile the unsheltered site of moderate depth (Cow Island), and the shallowest location, Green Lake, both returned similar, significantly lower, yields (~7 ng/µL) (Fig 19).

3.3.3.2 glnA Gene Copies per g of Sediment

When the tops of the cores were compared across sites as with the total DNA estimates, not all of the sites had similar mean *glnA* copies (Table 7). All sites except for Little Lake returned similar copies per gram of sediment of the *glnA* gene (2-4x10^8 gene copies/g) (Fig 20, Fig 21). The Little Lake coring site returned a significantly higher number of *glnA* gene copies (~6x10^8/g) (Fig 20, Fig 21). Comparison of Middle portions of the cores indicated that once again Little Lake had the highest *glnA* copies per gram of sediment (Fig 20, Fig 21). The Corridor and Cow island sites both gave similar copy yields (~3 x10^7/g) (Table 7, Fig 20, Fig 21) and the shallowest site, Green Lake, had the least *glnA* copies per gram of sediment (~8 x10^6/g) (Fig 19). The Little Lake site yielded the highest number of copies per gram of sediment (~2 x10^8/g) (Fig 20, Fig 21).

For the bottom portions of the core the Little Lake site similarly returned the highest copy numbers of glnA (~2 x10^7/g) with the deep Corridor site being slightly lower (~9.9 x10^6/g) (Fig 20). The shallowest site, Green Lake, and the unsheltered site of moderate depth, Cow Island, both had lower yet similar copy numbers per gram of sediment (< 3 x10^6/g) (Table 7, Fig 20, Fig 21).

3.3.3.3 CYA *Gene Copies per g of Sediment*

The top of the cores all yielded different average copy numbers for the cyanobacterial 16s rRNA gene (CYA). Little Lake yielded the highest number of copies per gram of sediment (~3 x10^8/g) and Corridor the lowest number of copies (~2x10^7/g) (Fig 22, Fig 23). Green Lake site and Cow Island both yielded moderate CYA copy numbers with Cow Island having slightly more per gram of sediment than Green Lake (Fig 22, Fig 23). Similarly, the middle portions of

the cores yielded significantly different gene copy numbers per gram of sediment across the four coring sites (Table 8). As with *glnA* results, Little Lake yielded the highest number of CYA copies per gram of sediment (~3-5x10^7/g) (Fig 22, Fig 23) and Green Lake the lowest (~9x10^5/g) with Corridor yielding only slightly more (Fig 22, Fig 23). The bottom portions of the cores were all found to be significantly different from one another with regards to CYA except the deepest, Corridor, and shallowest, Green Lake, sites, which had similar CYA copy number yields (~1x10^6 /g and ~2x10^6 /g average respectively) (Table 8, Fig 22, Fig 23). As found for glnA and the other portions of the cores, Little Lake yield the highest number of CYA copies (~4x10^6/g), and Cow Island the lowest at the bottom of the core (~4x10^5/g) (Fig 22, Fig 23).

3.3.3.4 apcA Gene Copies per g of Sediment

The primers designed for the allophycocyanin alpha-chain gene (*apcA*) amplified a single target (single band amplification) when utilized on extracted DNA materials from pure culture of desired targets (CPCC 300 & CPCC 663) (Appendix B: Figure 5b**)**. The primers showed a bias towards *Synechococcus spp.* producing much stronger amplification bands when used on *Synechococcus* pure culture extractions than with *Microcystis* pure culture extracts. This was expected as the degenerate primer was primarily constructed with sequences for *apcA* entered from *Synechoccocus* specimens. There was no visible amplification of off-targets when endpoint PCR was attempted with a green algal (*Chlorella* DNA extractions) or no template controls (deionized water). The product of the amplification was extracted from the gel and sequenced and the results assessed on NCBI BLAST; this confirmed that the primers amplified allophycocyanin alpha-chain (*apcA*) but further indicated a potential bias towards *Synechoccocus spp.*, as the resulting matches were primarily for *Synechoccocus* entries.

When the tops of the cores were compared, all of the sites except for Cow Island had similar mean yields of *apcA* copy numbers per gram of sediment (~1-2x10^6/g) (Fig 24, Fig 25). Cow Island had significantly more *apcA* copies (~3.2x10^6/g) (Table 9, Fig 24, Fig 25).

The middle portions of the cores were different among the sites except for Cow Island and the Corridor site with similar averages (~7x10^4/g) (Table 9, Fig 24, Fig 25). As found for the other gene targets, Little Lake had the highest *apcA* (~2x10^5/g) and the shallowest site, Green Lake, had the lowest *apcA* (~9x10^3/g).

The bottom portions of Corridor and Little Lake both yielded the highest number of *apcA* copies per gram of sediment with similar averages (~7x10^4/g) (Table 9, Fig 24, Fig 25**)**. Cow Island yielded less than either Corridor or Little Lake site (~1.5x10^4/g) (Table 9, Fig 24, Fig 25**)** and the Green lake site once again yielded the lowest number of *apcA* copies (~3x10^3/g) (Table 9, Fig 24, Fig 25).

3.4 Discussion

3.4.1 DNA Extraction Quantification: Quant-iT® vs NanoDrop™

Two methods of DNA quantification were employed to determine DNA concentrations from each sediment section. Quant-iT® kits have been considered a more accurate method for quantifying DNA due to their sensitivity and specificity for intact double-stranded DNA (dsDNA) (Shokere *et al.* 2009; Holden *et al.* 2009). Quant-iT® kits make use of a dsDNA specific binding dye which is considered to be much more specific in its function and also less likely to return false positive, or inflated, results. NanoDrop™ spectrophotometry is another common quantification methodology frequently used as a quick, relatively accurate method of quantifying DNA in a laboratory setting (Desjardins & Conklin 2010). However, shortcomings

have been identified with the methodology (Shokere *et al.* 2009; Holden *et al.* 2009). The NanoDrop™ platform makes use of the absorbance of light at specific wavelengths (i.e.: 260nm, 280nm, 230nm etc.) and the ratios generated between those wavelengths to quantify nucleic acids within solution. However, the dependence of this method of quantification on general wavelength absorption means that it cannot differentiate single stranded DNA (ssDNA) and dsDNA easily, many other contaminating substances may also absorb light at similar wavelengths skewing the results leading to over- or under-estimations of nucleic acid concentrations (Shokere *et al.* 2009; Holden *et al.* 2009).

In this study, the two methods were highly correlated (Fig. 18) returning very similar concentration results. However, overall the Quant-iT® kit was more conservative (i.e.: on average lower ng/µL concentrations per sediment section) and was on average 13% more conservative in its quantification than the NanoDrop™ method. This difference is likely due to contaminants from both the environment (e.g. organic molecules) or extraction process itself (guanidine salts etc.), which can absorb the wavelengths of light required for quantification and qualification thereby modifying the readings acquired via spectrophotometry. The more conservative readings acquired through the Quant-iT® kit are likely more accurate than those produced by the NanoDrop™ kit due to the specificity of the fluorescent dyes used to determine DNA content. The difference observed, at least in this study, was low enough that either quantification methodology could be used to obtain good estimates of extracted DNA concentrations. However, each methodology has its own advantages. NanoDrop™ provides an estimate of quality via the 260:280 nm wavelength ratio and is far more rapid than the Quant-iT® kit. Whereas the Quant-iT® kit likely remains the more accurate methodology as it should, by design, only bind with undegraded double-stranded DNA (dsDNA).

3.4.2 Comparison of Sediment DNA Across Total DNA Extracted:

Cores were not compared within themselves from top to bottom as it was expected that DNA concentrations would decrease with core depth, regardless of site, due to the degradative processes that occur under virtually all environmental conditions (Rees *et al.* 2014; Barnes *et al*. 2014; Strickler *et al.* 2015; Roussel *et al.* 2015). As expected, sediment DNA concentrations were highest within top layers of cores at all sites and decreased with depth within the cores (Appendix B: Figure 2b & 4b). Three out of the four sites presented similar concentrations of extracted materials in the top portions of the core, as predicted. This was likely due to the fact that the DNA bound within the sediments has not been subjected to significant stresses as yet and may be being replenished due to its proximity to the water interface (new sedimentation). In addition, cell replication and activity may still be occurring, particularly within the shallower sites because of solar radiation reaching the bottom and stimulating photosynthesis. Unexpectedly, the highest concentrations of DNA extracted within all portions (top, middle and bottom) of the core was from the Little Lake site. The Little Lake site was of moderate depth (~10m), and similar in depth to the Cow Island site, but Little Lake was located in a significantly more sheltered area within Big Rideau Lake. This may have amplified the sediment focusing effect noted for other paleolimnological signals (Blais *et al* 1995), thus increasing the amount of DNA accumulating in the sediments. The sheltered morphology may also be reducing water turbulence thus reducing resuspension of water-interface which could inhibit the entry of DNA to the sediment record.

Although the shallowest site (~ 2.7 m, Green Lake) was found to have similar DNA concentrations as most of the other sites at the top of the core (Fig 19)**, t**he total DNA content declined more rapidly with core depth, becoming the lowest in the middle and bottom portions of

any core examined in this study. Shallower sites are likely subjected to higher degrees of degradative pressure (from temperature, increased oxygen, UV and benthic microbial metabolism etc.) and would not present a good area to core for long-term studies. It may also be skewed towards benthic communities due to light penetration at this site. However, shallower sites may still provide a viable site to core for studies spanning a much shorter period of time as the top of the core returned similar DNA concentrations to those found elsewhere. Benthic vs. pelagic community biases were not examined in this study and, as researchers such as Moos *et al*. (2005) have noted, there may be a significant bias towards benthic or pelagic species of prokaryotes with depth. This may be important depending on the study question and should be considered during the site selection process if specific benthic vs. pelagic taxa were the desired targets of study.

The deepest site cored in this study (~21 m), Corridor, was predicted to have the highest concentrations of extracted DNA materials in the bottom portion of the core. The Corridor core had similar concentrations of extracted DNA material in the top portion of the core to most other sites in this study (Fig 19), with similar quantities at the top and middle portions of the core. Where it differed was that the bottom did appear to preserve the sediment DNA the best, leading to the highest concentrations of extracted total DNA content in the oldest sections of the core. (Fig 19). This indicates that deeper sites (20m +) likely afford significant protective benefits for sediment bound DNA from degradation, as predicted.

Thus, the best site for the extraction of sediment DNA in Big Rideau for long term studies was from the site of greatest depth (20m +). This was likely due to the natural conditions (cold temperatures, anoxia, absence of UV radiation) which arise at extreme depths and are conducive to long-term preservation of DNA as well as the alleviation of degradative pressures

on sedDNA. Shallow sites represented the least preferable locations to core for studies examining ancient DNA materials.

3.4.3 Comparison of Target Gene Copy Numbers Across Sites:

Similar trends were observed in the *glnA*, CYA and *apcA* copies per gram of sediment to those observed for the total sediment DNA (sedDNA) concentrations with few exceptions. Little Lake, which had the highest concentrations of extracted DNA per sediment layer as well as *glnA*, CYA and *apcA* copies per gram of all cores taken in most sections of the cores. This likely reflects the effect sheltered morphology has on particulate entering the sediment records as a larger amount of sedDNA deposition, and, by extension the number of copies of each target entering the sediment at this site.

The deepest site, Corridor, had the lowest number of *glnA*, CYA and *apcA* copies per gram of sediment in the top portion of the core. This is likely due to the significant depth at which this core was retrieved. The distance particulates would have to travel would be greatly increased and the time required to reach the sediment, as well as environmental effects such as turbulence, may play a role in the reduced copy numbers observed in the top portions of the Corridor core. As with the total DNA, the bottom of the core (i.e.: the oldest sediments) had higher gene copy numbers when compared to the other sites cored. This was likely due to the many conditions at depth favoring preservation (low temperature, anoxic etc.) allowing the sediments to retain higher *glnA*, CYA and *apcA* copy numbers than other sites. The greater copy numbers retrieved from the bottom portion suggests that although lower quantities may be entering the sediment greater preservation is occurring. That this trend was generally observed in all gene targets lends credence to the assumption that environmental conditions favorable for

sedDNA preservation are having an effect not only on the total extractable DNA material but also on copy numbers of individual targets.

The CYA and *apcA* gene both followed a similar trend, which was expected as they both targeted cyanobacteria, but some differences arose in comparison to *glnA*. First, lower overall numbers of *apcA* copies were quantified at all sites within the lake when compared to other gene targets. This was likely because the *apcA* primers were biased towards, *Synechoccocus* sp. over other cyanobacteria taxa, which possess the gene. Both *apcA* and CYA followed the same general trend observed with *glnA*, with the highest copy numbers in all portions of the core observed in the Little Lake region which was likely due to morphometry effects related to a more sheltered basin. The deepest site, Corridor, had the highest copy numbers of both *apcA* and CYA in the bottom portion of the core once again reinforcing the notion that environmental factors found at greater lake depth (> 20 m) play a significant role in sedDNA preservation. What was not consistent with the *glnA* results was that the shallowest site, Green Lake, actually had similar quantities of CYA as the deepest site (Corridor) in the bottom portion of the core. A possible explanation could be the presence of active benthic cyanobacteria, resulting from sufficient solar radiation reaching the bottom as well as more sediment mixing into the core from physical turbulence. Another possible explanation is historic bloom formations that were swept into the shallow bay area and deposited in large quantities during that time period. Differences in CYA and *apcA* were also observed at the Cow Island site, which had similar depth to the Little Lake site (~10.7m) but was not as sheltered. Although there were increased numbers of CYA and *apcA* copies in the top portion of the core from this site, all gene targets (*glnA*, CYA and *apcA*) rapidly declined becoming most similar to those observed in the shallowest site (Green Lake). This indicates that, at least in Big Rideau, depths of ~10 m or less likely do not retain the

environmental characteristics necessary for preservation. Furthermore, this indicates that the higher concentrations of extracted sedDNA as well as copies of all three targets observed at the Little Lake site are likely due to the deposition of large volumes of particulates rather than better conditions for sedDNA preservation.

3.5 Conclusions and Future Applications

Although this study was only conducted on one lake, it supported several assumptions about the preservation of sediment DNA. The deepest site in this study, both in terms of gene copies per gram of sediment as well as in raw extracted concentrations, was found to provide the highest concentrations/copy numbers at the bottom of the cores at (21.3m) indicating good DNA preservation over time. Morphometry of the lake was also important for the quantity of DNA accumulating in the sediments. Little Lake site was of similar depth to the Cow Island site; however, it was more sheltered and this sheltered status appeared to promote the entry of DNA into the sediments via either a focusing effect or the protective qualities provided by the sheltered basin. Shallow sites such as Green Lake (~2.7m depth) and unsheltered sites of moderate depth (~10m) presented the lowest average gene copy numbers and concentrations of DNA in most portions of the core, indicating that significant degradation or disruption of DNA entry into the sediment was occurring at these sites. The best sites for the retrieval of high quantities of high quality sedDNA for long-term paleo-genetic study are likely a combination of basins with significant depth (20m+) in highly sheltered locations.

Table 5

Table 5: Sediment collection information for cores retrieved from Big Rideau lake at four internal locations of varying depth and local morphometry. Green Lake was a shallow (~2.7m) unsheltered site. Corridor was the deepest (~21.3m) unsheltered site. Cow Island was a small basin located beside Cow Island and is unsheltered and of moderate depth (~10.6m). Little Lake was a sheltered site of moderate (~8.6m) depth (similar to Cow Island).

Lake/Region	Date Collected	Core Length (cm)	Depth Collected (m)	Coordinates	Collected By
Big Rideau - Green Bay	16/07/2018	40	2.7	44°42'12.9"N 76°12'52.8"W	William Dodsworth, Meaghan MacIntyre-Newell, Amber Dyke
Big Rideau - Corridor	16/07/2018	45.5	21.3	44°42'43.5"N 76°13'25.0"W	William Dodsworth, Meaghan MacIntyre-Newell, Amber Dyke
Big Rideau - Offshore Cow Island	16/07/2018	45	10.6	44°43'15.3"N 76°12'09.1"W	William Dodsworth, Meaghan MacIntyre-Newell, Amber Dyke
Big Rideau - Little Lake, Sheltered Bay	16/07/2018	43	8.3	44°43'41.4"N 76°10'38.4"W	William Dodsworth, Meaghan MacIntyre-Newell, Amber Dyke

Table 6

Table 6: ANOVA results when comparing Total DNA extraction values between Top, Middle and Bottom portions of 4 cores representing different morphometric conditions within Big Rideau Lake. Results were further refined through the use of Tukey's Post-Hoc test with adjusted p-values to identify the most different sites for each portion of the cores. For the Top, Little Lake, the most sheltered site of moderate depth (~10m) proved to be the only significantly different DNA yield (p=<0.05) while all other sites and conditions yielded similar amounts. For the Middle portions of the core, all sites proved to differ significantly from one another (p=<0.05) except for the Corridor and Cow island regions. In the Bottom of the core the unsheltered site of moderate depth (Cow Island) and the shallowest site (Green Lake) became similar (p=>0.05), while the most sheltered site (Little Lake) and the deepest site (Corridor) became similar (p=>0.05). However, Little Lake and Corridor remained significantly different from both Green Lake and Cow Island (p=<0.05).

Location	ANOVA Significance	Post-Hoc Tukey's Comparisons	Tukey's Adjusted Significance Values
Top	$1.03x10^{-11}$***	Cow Island-Corridor	0.99
		Green Lake-Corridor	0.97
		Little Lake -Corridor	$1x10^{-7}$***
		Green Lake-Cow Island	0.96
		Little Lake -Cow Island	$1x10^{-7}$***
		Little Lake -Green Lake	$1x10^{-7}$***
Middle	$1x10^{-16}$***	Cow Island-Corridor	0.59
		Green Lake-Corridor	$1.6x10^{-6}$***
		Little Lake -Corridor	$1x10^{-7}$***
		Green Lake-Cow Island	$1x10^{-7}$***
		Little Lake -Cow Island	$1x10^{-7}$***
		Little Lake -Green Lake	$1x10^{-7}$***
Bottom	$2.94x10^{-10}$***	Cow Island-Corridor	$2.1x10^{-6}$***
		Green Lake-Corridor	$4.4x10^{-6}$***
		Little Lake -Corridor	0.45
		Green Lake-Cow Island	0.99
		Little Lake -Cow Island	$1x10^{-7}$***
		Little Lake -Green Lake	$1x10^{-7}$***

α' *, **, *** = degree of significance

Table 7

Table 7: ANOVA results when comparing glnA copies per g of sediment between Top, Middle and Bottom portions of 4 cores representing different morphometric conditions within Big Rideau Lake. Results were further refined through the use of Tukey's Post-Hoc test with adjusted p-values to identify the most different sites for each portion of the cores. For the Top, Little Lake, the most sheltered site of moderate depth (~10m) proved to be the only significantly different glnA copy number ($p=<0.05$) while all other sites and conditions yielded similar copy numbers. For the Middle portions of the core, all sites proved to differ significantly from one another ($p=<0.05$) except for the Corridor and Cow island regions. In the Bottom of the core the all sites displayed different copy numbers per gram of sediment for glnA.

Location	ANOVA Significance	Post-Hoc Tukey's Comparisons	Tukey's Adjusted Significance Values
Top	$9.17x10^{-5}$***	Cow Island-Corridor	0.26
		Green Lake-Corridor	0.18
		Little Lake -Corridor	$3.9x10^{-5}$***
		Green Lake-Cow Island	0.99
		Little Lake -Cow Island	$8.5x10^{-3}$***
		Little Lake -Green Lake	$1.5x10^{-2}$***
Middle	$6.04x10^{-15}$***	Cow Island-Corridor	0.47
		Green Lake-Corridor	$2.3x10^{-5}$***
		Little Lake -Corridor	$1x10^{-7}$***
		Green Lake-Cow Island	$3x10^{-7}$***
		Little Lake -Cow Island	$1x10^{-7}$***
		Little Lake -Green Lake	$1x10^{-7}$***
Bottom	$2x10^{-16}$***	Cow Island-Corridor	$1x10^{-7}$***
		Green Lake-Corridor	$1x10^{-7}$***
		Little Lake -Corridor	$8x10^{-3}$
		Green Lake-Cow Island	$5x10^{-5}$***
		Little Lake -Cow Island	$1x10^{-7}$***
		Little Lake -Green Lake	$1x10^{-7}$***

α' *, **, *** = degree of significance

Table 8

Table 8: ANOVA results when comparing CYA copies per g of sediment between Top, Middle and Bottom portions of 4 cores representing different morphometric conditions within Big Rideau Lake. Results were further refined through the use of Tukey's Post-Hoc test with adjusted p-values to identify the most different sites for each portion of the cores. For the Top all sites and conditions yielded differing copy numbers (p=<0.05) with Cow Island yielding copy numbers between both Green Lake and Little Lake (p=>0.05). For the Middle portions of the core, all sites proved to differ significantly from one another (p=<0.05). In the Bottom of the core the all sites displayed different copy numbers per gram of sediment for CYA except Green Lake and Corridor region (p=>0.05).

Location	ANOVA Significance	Post-Hoc Tukey's Comparisons	Tukey's Adjusted Significance Values
Top	$9.6x10^{-13}$***	Cow Island-Corridor	$1x10^{-7}$***
		Green Lake-Corridor	$1x10^{-7}$***
		Little Lake -Corridor	$1x10^{-7}$***
		Green Lake-Cow Island	0.13
		Little Lake -Cow Island	0.5
		Little Lake -Green Lake	$4x10^{-3}$***
Middle	$2x10^{-16}$***	Cow Island-Corridor	$6x10^{-7}$***
		Green Lake-Corridor	$6x10^{-2}$***
		Little Lake -Corridor	$1x10^{-7}$***
		Green Lake-Cow Island	$1x10^{-7}$***
		Little Lake -Cow Island	$3x10^{-5}$***
		Little Lake -Green Lake	$1x10^{-7}$***
Bottom	$2x10^{-12}$***	Cow Island-Corridor	$3x10^{-7}$***
		Green Lake-Corridor	0.59
		Little Lake -Corridor	$3x10^{-4}$
		Green Lake-Cow Island	$1x10^{-7}$***
		Little Lake -Cow Island	$1x10^{-7}$***
		Little Lake -Green Lake	$1x10^{-2}$***

α' *, **, *** = degree of significance

Table 9

Table 9: ANOVA results when comparing *apcA* copies per g of sediment between Top, Middle and Bottom portions of 4 cores representing different morphometric conditions within Big Rideau Lake. Results were further refined through the use of Tukey's Post-Hoc test with adjusted p-values to identify the most different sites for each portion of the cores. For the Top all sites and conditions yielded similar copy numbers (p=>0.05) with Cow Island yielding significantly different copy number values (p=<0.05). For the Middle portions of the core, all sites proved to differ significantly from one another (p=<0.05) except for Cow Island and Corridor sites. In the Bottom of the core the all sites displayed different copy numbers per gram of sediment for *apcA* except Corridor and Little Lake region (p=>0.05).

Location	ANOVA Significance	Post-Hoc Tukey's Comparisons	Tukey's Adjusted Significance Values
Top	$5.5x10^{-7}$***	Cow Island-Corridor	$6.2x10^{-6}$***
		Green Lake-Corridor	0.43
		Little Lake -Corridor	0.95
		Green Lake-Cow Island	$6.5x10^{-4}$***
		Little Lake -Cow Island	$1.3x10^{-6}$***
		Little Lake -Green Lake	0.19
Middle	$3.25x10^{-11}$***	Cow Island-Corridor	0.99
		Green Lake-Corridor	$1x10^{-7}$***
		Little Lake -Corridor	$1x10^{-2}$***
		Green Lake-Cow Island	$2x10^{-7}$***
		Little Lake -Cow Island	$1x10^{-2}$***
		Little Lake -Green Lake	$1x10^{-7}$***
Bottom	$2x10^{-16}$***	Cow Island-Corridor	$1x10^{-6}$***
		Green Lake-Corridor	$1x10^{-7}$***
		Little Lake -Corridor	0.8
		Green Lake-Cow Island	$1x10^{-7}$***
		Little Lake -Cow Island	$2x10^{-5}$***
		Little Lake -Green Lake	$1x10^{-7}$***

α' *, **, *** = degree of significance

Figure 18

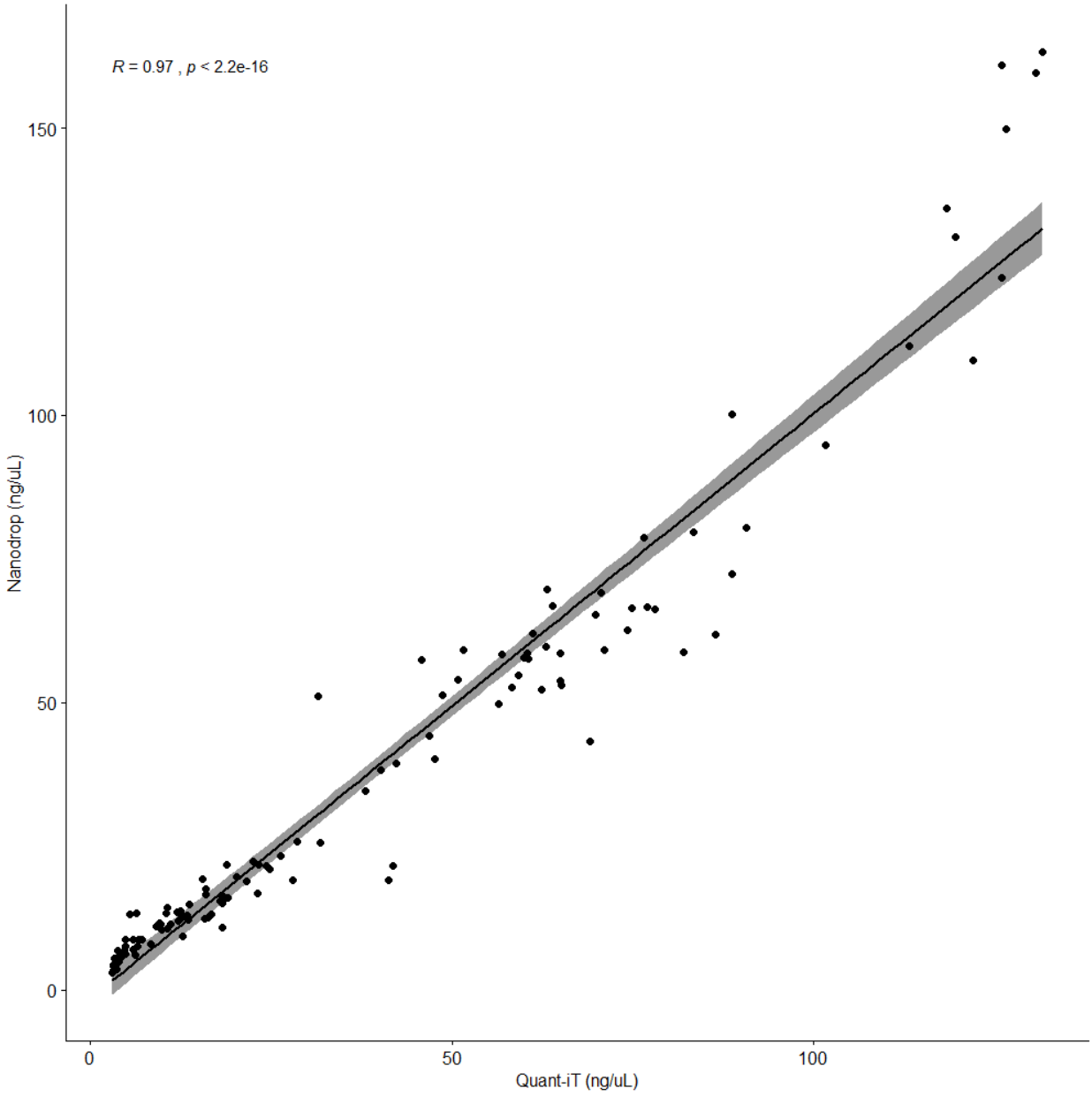


Figure 18: Correlation analysis of extracted nucleic acid quantities of 120 (n=120) sediment DNA extraction samples from two methodologies. The results obtained from both the NanoDrop™ and the Quant-iT® kit reveal a 97% correlation strength with a high significance level of $p=2.2x10^{-16}$. The two methodologies produce highly similar values when compared directly and can likely be used interchangeably for initial evaluation of extracted DNA for paleolimnological studies.

Figure 19

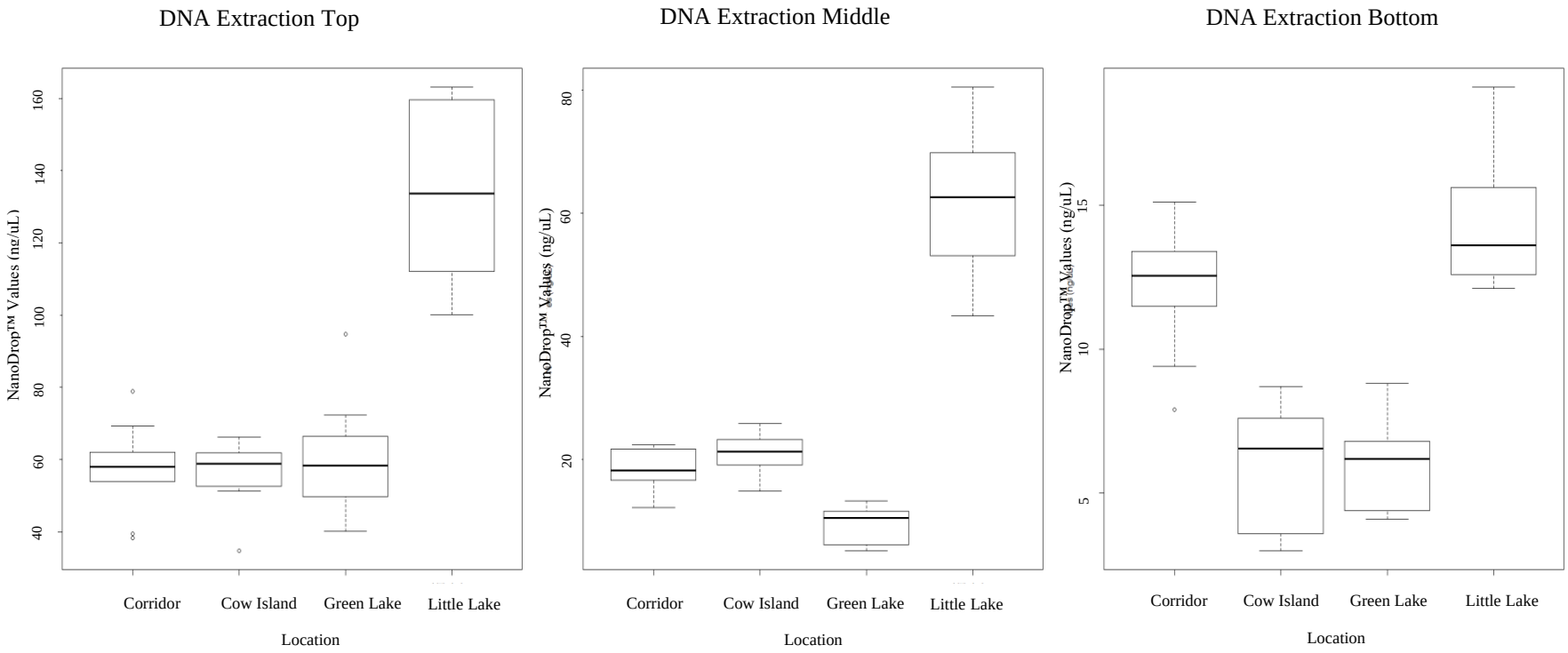


Figure 19: Plots of average DNA extraction value (in ng/µL) for sediment extractions from 4 sites (Corridor, Cow Island, Green Lake, Little Lake regions) within Big Rideau Lake. **Left**: Averages for top 10 sections (0-5cm) for each core. Corridor, Cow Island and Green Lake all display nearly identical average extraction values (~60 ng/µL). Little Lake, however, is a more sheltered site and has significantly higher average extraction values from the top portion of the core (~130 ng/µL average). **Middle:** All lake sites displayed differing average extraction values with Green Lake, the shallowest site, having the lowest average extraction values (~13 ng/µL average) and Little Lake, the most sheltered site, having the highest average extraction values (~60 ng/µL average extraction values). **Right**: At the bottom of the core, the average extraction values between the deepest site, Corridor region, and the most sheltered site, Little Lake, both had similar extraction values (~12 ng/µL average), whilst the shallowest site, Green Lake, and the unsheltered site of moderate depth, Cow Island, both had similar extraction values (~ 7 ng/µL average).

Figure 20

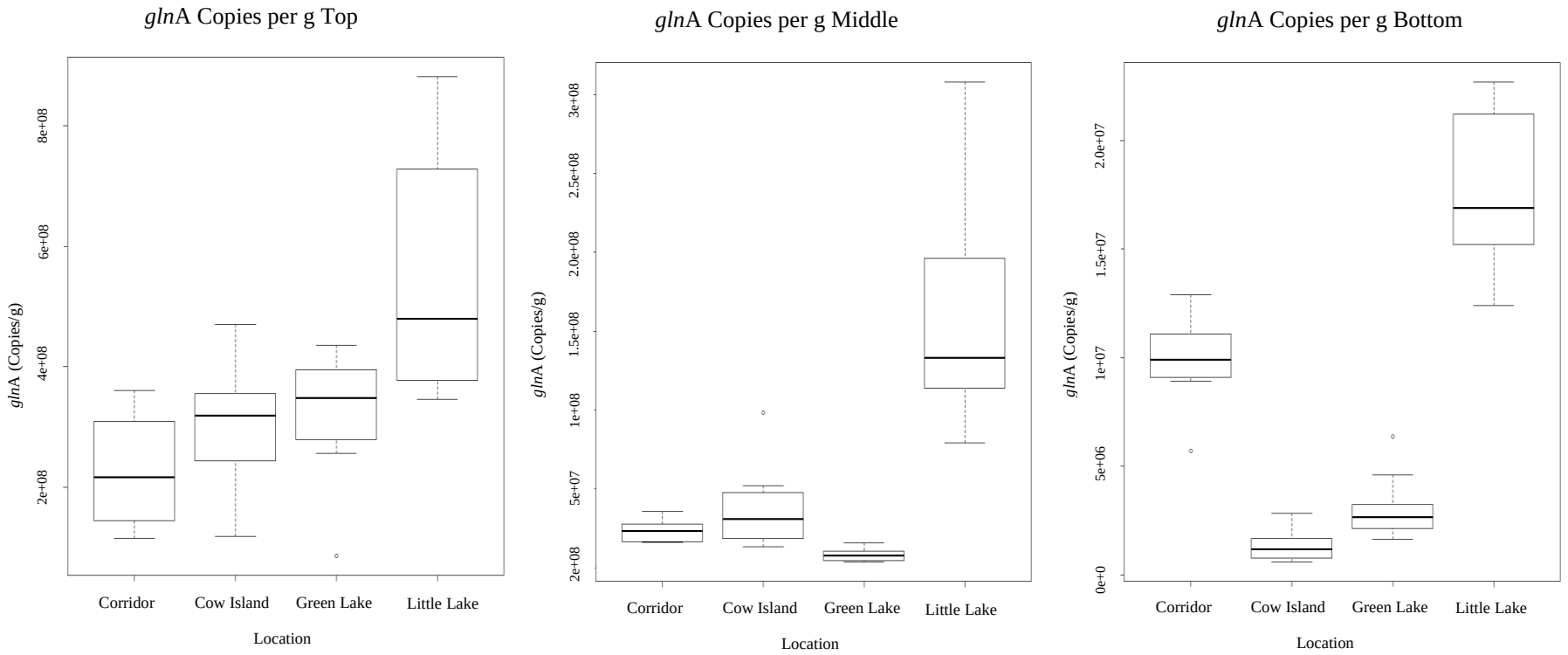


Figure 20: Plots of average glnA (in copies/g of sediment) for sediment extractions from 4 sites (Corridor, Cow Island, Green Lake, Little Lake regions) within Big Rideau Lake. **Left**: Averages for top 10 sections (0-5cm) for each core. Corridor, Cow Island and Green Lake all display similar glnA copy numbers(~2-4x10^8). Little Lake, is a more sheltered site and has significantly higher average extraction values and had a correspondingly higher copy number for glnA (~6x10^8 average). **Middle:** Averages for sections between (20-25cm) All lake sites displayed differing average extraction values except for Corridor and Cow Island sites which gave similar yields. Green Lake, the shallowest site, having the lowest glnA copy number (~8 x10^6copies/g) and Little Lake, the most sheltered site, having the highest average copy numbers (~2 x10^8) **Right**: At the bottom of the core (35-40cm), the copy number values between the deepest site, Corridor region, and the most sheltered site, Little Lake, became more similar (~9.9 x10^6 and ~2 x10^7 average respectively), whilst the shallowest site, Green Lake, and the unsheltered site of moderate depth, Cow Island, both decreased dramatically (< 3 x10^6 copies per g of sediment each).

Figure 21

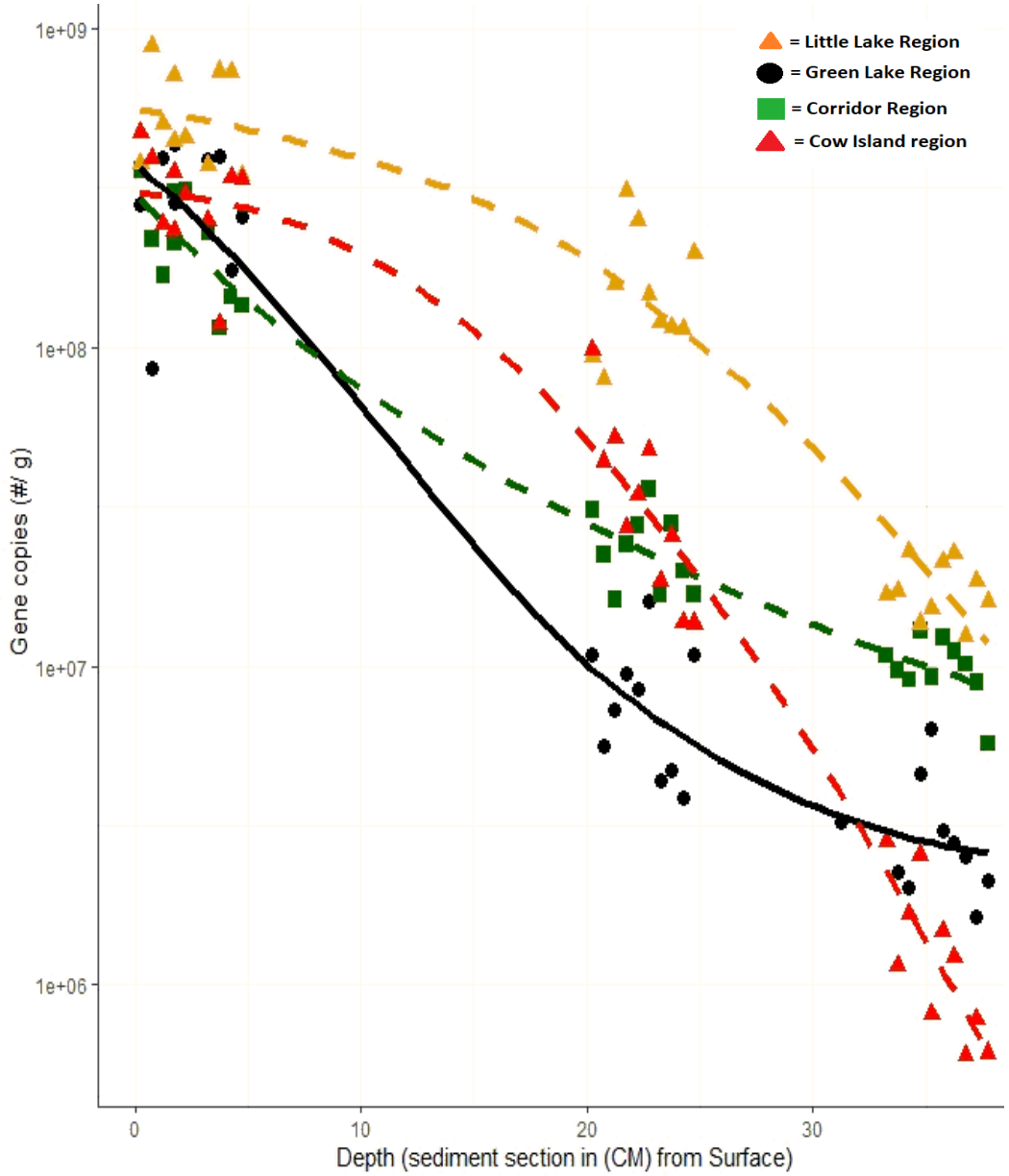


Figure 21: Copies per g of sediment of *glnA* target from each site plotted against depth of core section (cm from the surface). Little Lake (orange) site had the greatest return of *glnA* copies at all portions of the core (top, middle and bottom). Corridor (green), the deepest site had similar initial copy numbers but retained the highest copy numbers in the bottom portions of the core. Green Lake (black) the shallowest site and Cow island (Red) an unsheltered site of moderate depth both had similar initial *glnA* values but declined rapidly towards the bottom of the core. Copy numbers transformed using a logarithmic scale. Lines fitted using a "LOESS" method.

Figure 22

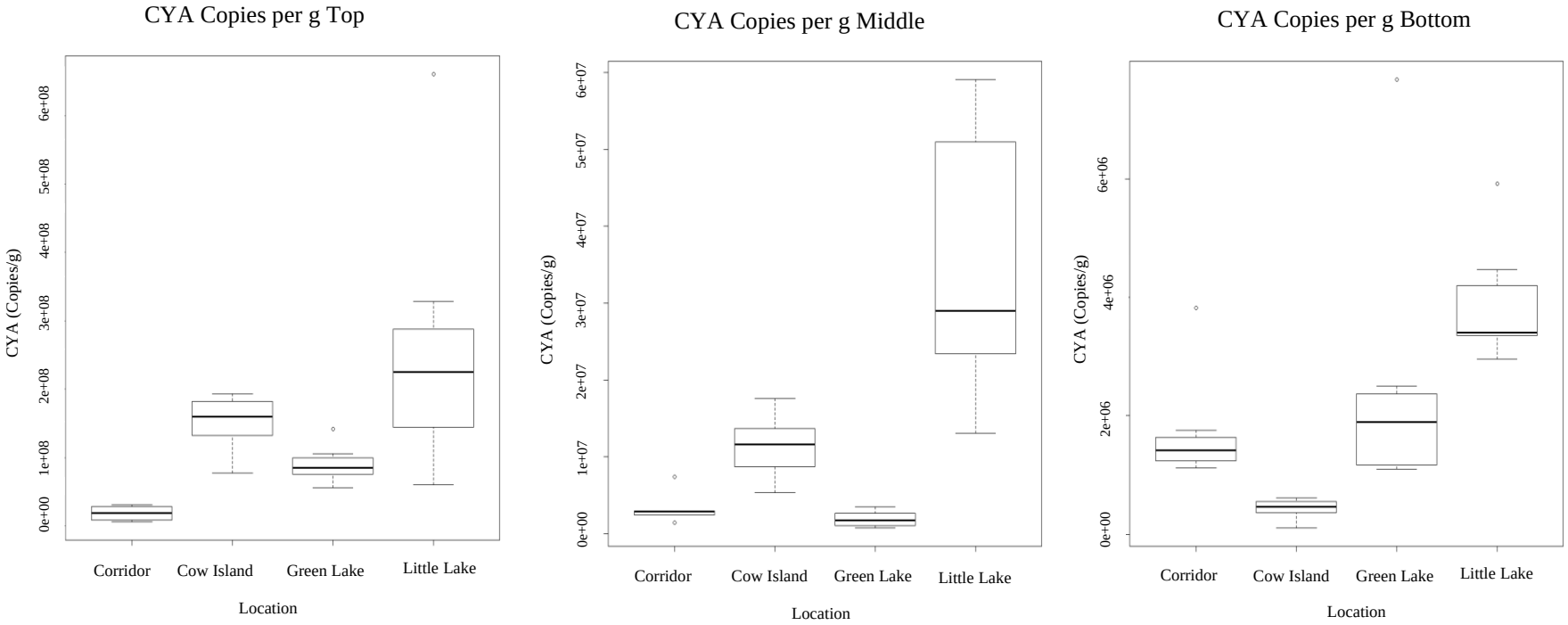


Figure 22: Plots of average CYA (in copies/g of sediment) for sediment extractions from 4 sites (Corridor, Cow Island, Green Lake, Little Lake regions) within Big Rideau Lake. **Left**: Averages for top 10 sections (0-5cm) for each core. Little Lake and Cow Island and Green lake all display similar CYA copy numbers(~2-3x10^8). The Corridor site, the deepest (20+m depth) had lower CYA copy numbers per g of sediment than any other site (~2x10^7). **Middle:** Average copy numbers for sections between (20-25cm) All lake sites displayed differing copy number value. Green lake, the shallowest site, displayed the lowest CYA copy number (~ 9x10^5copies/g) and Little Lake, the most sheltered site, having the highest copy number values (~3-5x10^7) **Right**: At the bottom of the core (35-40cm). Little Lake once more displayed the highest copy numbers for CYA in the bottom sections of the core (~4x10^6). The copy number values between the deepest site, Corridor region, and the shallowest site, Green Lake, became more similar (~1x10^6 and ~2x10^6 average respectively), whilst the unsheltered site of moderate depth, Cow Island, decreased dramatically with the lowest number of CYA copies recovered (~4 x10^5 copies per g of sediment each).

Figure 23

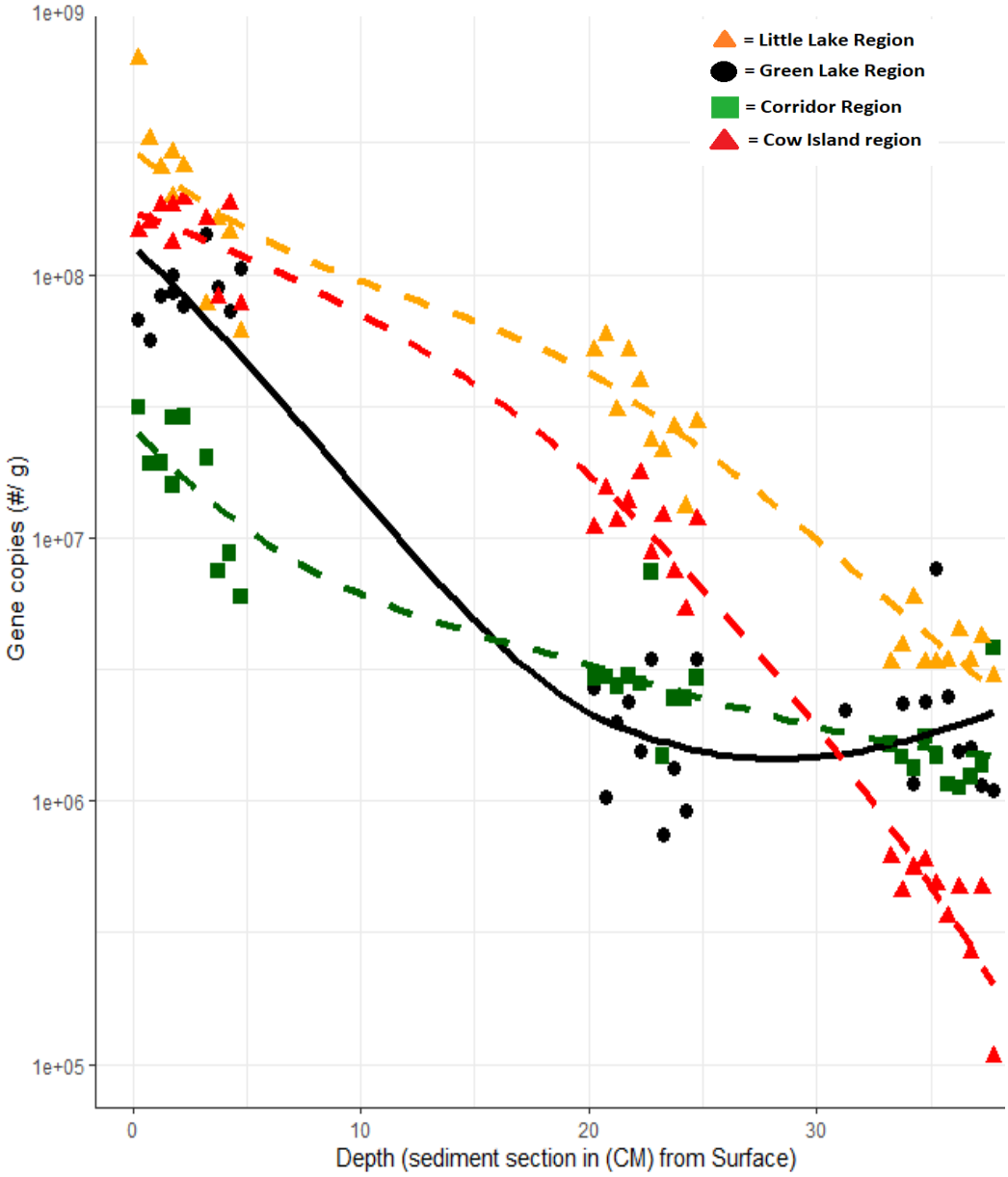


Figure 23: Copies per gram of sediment for the CYA gene target from 4 sites within Big Rideau Lake. Plotted against sediment section in cm from the water interface. Overall Little Lake Site (orange) had higher copies than most other sites and retained higher copies throughout all portions of the core (top, middle and bottom). Corridor (green) had lower initial CYA copies, however it retained slightly higher copy numbers in the bottom portions of the core. The shallowest site Green Lake (black) had similar initial values to other sites however it also retained similar CYA copy numbers to Corridor (green) and Little Lake (orange) in the deepest portions of the core. Cow Island (red) had higher CYA copies/g in the top and middle portions of the core, however, it rapidly declined near the bottom portion of the core with the fewest CYA copies returned. Copy numbers transformed using a logarithmic scale. Lines fitted using a "LOESS" method.

Figure 24

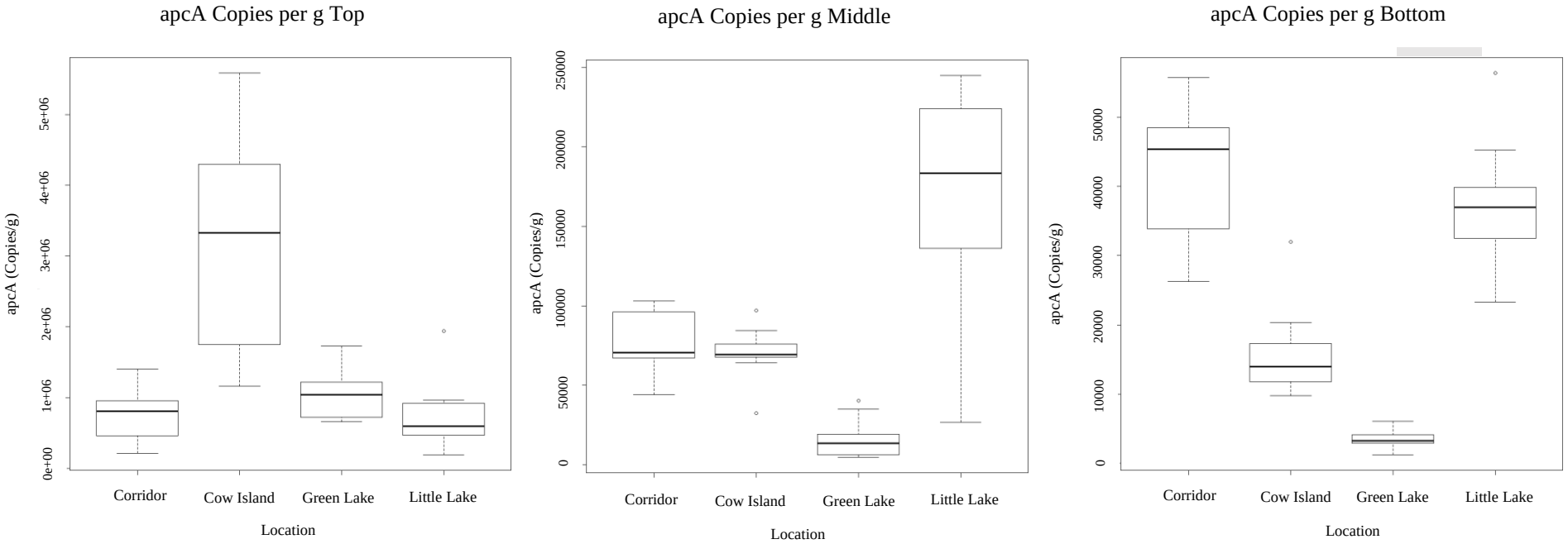


Figure 24: Plots of average apcA (in copies/g of sediment) for sediment extractions from 4 sites (Corridor, Cow Island, Green Lake, Little Lake regions) within Big Rideau Lake. **Left**: Averages for top 10 sections (0-5cm) for each core. Little Lake, Corridor and Green lake all display similar apcA copy numbers(~1-2x10^6). The Cow Island site had higher apcA copy numbers per gram of sediment than any other site (~3x10^6). **Middle:** Average copy numbers for sections between (20-25cm) All lake sites displayed differing copy number values except Cow Island and Corridor sites (~7x10^4) with Little Lake displaying the highest number of copies (~2x10^5) and Green Lake site displaying the lowest (~9x10^3). **Right**: At the bottom of the core (35-40cm). Little Lake and the Corridor region both became similar (~4x10^4) with Cow Island site yielding fewer copies (~1.5x10^4) the shallowest site, Green Lake displaying the fewest copies (~3x10^3).

Figure 25

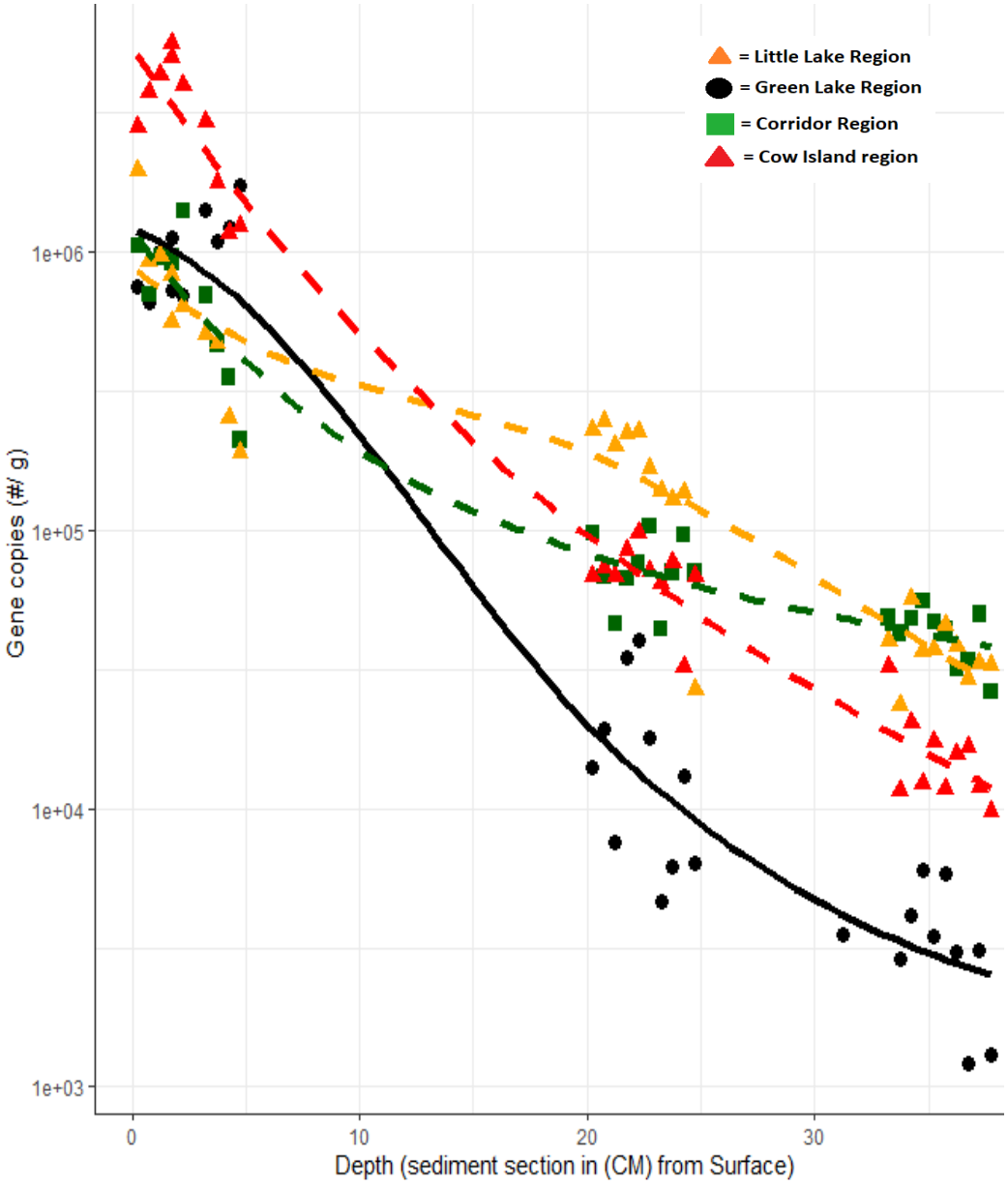


Figure 25: Copies per gram of sediment for the *apcA* gene target from 4 sites within Big Rideau Lake. Plotted against sediment section in cm from the water interface. All sites showed similar initial *apcA* copies/g values in the top of the core. The shallowest site, Green Lake (black) showed the most rapid decline through the core having the lowest copy numbers in the middle and bottom portions. Corridor (green) and Little Lake (orange) both had similar values throughout the core and the highest *apcA* copy numbers in the bottom portions of the core (30cm +). The Cow Island Site (red) had the highest initial *apcA* copy numbers but quickly declined to levels more similar to those of the Corridor (green) and Little Lake (orange) sites in the middle and bottom portions. Copy numbers transformed using a logarithmic scale. Lines fitted using a "LOESS" method.

Chapter 4 General Conclusions

The main objective of this thesis was to re-construct the history of cyanobacteria and microcystins (cyanobacterial toxins) in four Central Ontario lakes to determine whether nutrients and/or climate change have affected the relative abundance of cyanobacteria and toxigenic cyanobacteria from pre- to post European settlement (Chapter 2). In addition, four cores, from a variety of depths and morphometric conditions, were taken from a single lake to ascertain the role of site depth and morphometry in the preservation of sediment DNA over time (Chapter 3).

Although cyanobacteria are naturally abundant in aquatic systems, reports of blooms have been increasing in temperate regions (Winter *et al*. 2011). However, it is not known whether climate change has had a significant effect alone, or in concert, with nutrient additions in driving these blooms nor whether these blooms are becoming more or less toxic. These are important questions to guide future mitigation efforts.

In the two lake districts, aggregate climate data indicated that increases in maximum and mean temperature along with heating degree days have occurred in the past 20-40 years. Quantifications of gene copy numbers in all lakes found that proxies of bacterial and cyanobacterial abundances (*glnA* and CYA) have steadily increased over time. Although some of this may be attributed to sediment degradation over time, the relative abundance of cyanobacteria has clearly increased in the most developed lakes. Furthermore, these impacted lakes exhibited increased *mcyE* copy numbers when compared to unaffected lakes and also appeared to experience more dramatic increases in recent years. Examination of the effect of climate on gene target proxies indicated that climate had little to no effect on changes in gene abundance, nor on the increased toxicity observed in more anthropogenically affected lakes. Although climate did not appear to be having a significant effect, the changes in cyanobacteria or toxins may have

occurred too recently to be evaluated from the sediment record at the resolution of this study. Here sediment intervals of 0.5 cm represented an estimated 3-30+ years (depending on the lake or section). With respect to Three-Mile lake in Muskoka, it was more likely that the primary factor driving increases in cyanobacterial dominance and microcystins was due to an increase in nutrient loading from anthropogenic activities, however, monitoring data being very sparse made this difficult to support statistically. Recent chemical data from Three Mile Lake suggests an increase in trophic state from mesotrophic to eutrophic. In contrast, the reference lake in the same lake district of the Muskokas, Blue Chalk, remains oligotrophic and has experienced only a modest increase through time with regards to cyanobacterial abundance. Other methods of reconstructing phosphorus records, such as diatom inferred phosphorus approaches (Hall & Smol 1992; Rippey *et al*. 2011) could be used to characterize nutrient conditions through time. Climate factors could potentially be acting in a synergistic manner with nutrients in exacerbating cyanobacterial dominance within the lakes. However, continued research is required to fully discern and understand these complex interactions.

Little, if any, research has been carried out to determine whether site conditions have a significant effect on the deposition and preservation of sediment DNA for paleolimnological study in lacustrine or marine environments. Previous studies have typically selected sites with high rates of deposition as well as conditions considered to be conducive to DNA preservation. The results of Chapter 3 support prior assumptions with regards to site selection. Sites of great depth (20m +) provided the conditions required to preserve sedDNA as both quantity and copies per gram of sediment. Three of the gene targets were at greater concentrations in the deepest (oldest) sections of the core retrieved from the deepest site in Big Rideau Lake. Morphometry also appeared to play a significant role in how DNA enters the sediment, with a sheltered

morphometry significantly enhancing the amounts of material entering the sediment. Although the reason for this was not specifically determined in this study, it was highly likely sediment focusing combined with protection from wind and wave action were important factors in increased deposition at this site. What remained unclear, however, was whether morphometry played a role in DNA preservation. Whether the increased DNA content from the sheltered site (Little Lake) was due to morphometric factors preserving DNA, or, was merely a side-effect of the sediment focusing effect could not be determined.

This study could be expanded upon in future research by examining taxa distributions more closely. Previous studies have shown that microbial populations may be over-represented at certain depths and areas within a lake. Moos *et al.* (2005) found that diatom communities shifted between benthic and pelagic species between 0-30m and identified sites at the transition zone (8-13m) between these extremes as the best for representing the true community makeup of diatoms within a lake. The gene targets used in this study were not designed to differentiate between benthic and pelagic species thus a bias towards certain taxa may exist at sites of different depths. An analysis of community structure from high throughput sequencing of gene targets (e.g. via 16S rRNA sequencing as in Monchamp *et al.* 2016 Pilon *et al.* 2019) could help determine whether there were significant community or taxonomic differences present at different depths. What was conclusively determined was that highly sheltered basins of greater depth (20 m+) likely represent the best sites for the acquisition of higher quantities of extracted DNA that is well preserved for sustained periods of time.

The use of quantitative genetic techniques in paleolimnological studies of lake histories has proven valuable in filling knowledge gaps with respect to the rise in cyanobacterial dominance and cyanotoxin genes (Monchamp *et al.* 2016*;* Pilon *et al*. 2019). The present study

provides further support for a regime shift in temperate lakes. However, the specific influence of climate vs. nutrients requires additional research. With the risk of climate change continuing to increase globally (IPCC 2018), sediment DNA techniques including advanced sequencing technologies need to be further developed and utilized in parallel with traditional paleolimnological proxies. Given regional difference in climate change, such studies need to encompass different lake types and lake districts under different climates for a more comprehensive prediction of aquatic ecosystem changes.

References:

Abascal, F., Zardoya, R., & Telford, M. J. (2010). TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research*, *38*(SUPPL. 2), 7–13. https://doi.org/10.1093/nar/gkq291

Adrian, R., O'Reilly, C. M., Zagarese, H., Baines, S. B., Hessen, D. O., Keller, W., Livingston, D. M., Sommaruga, R., Straile, D., Van Donk, E., Weyhenmeyer, G.A., Winder, M. (2009). Lakes as sentinels of climate change. *Limnology and Oceanography*, *54*(6 PART 2), 2283–2297. https://doi.org/10.4319/lo.2009.54.6_part_2.2283

Allen, J. F., & Martin, W. (2007). Evolutionary biology: Out of thin air. *Nature*, *445*(7128), 610–612. https://doi.org/10.1038/445610a

Allentoft, M. E., Collins, M., Harker, D., Haile, J., Oskam, C. L., Hale, M. L., Campos, Paula F., Samaniego, Jose A., Gilbert, M. Thomas P., Willerslev, E., Zhang, G., Scofield, R Paul., Holdaway, Richard N., & Bunce, M. (2012). The half-life of DNA in bone: Measuring decay kinetics in 158 dated fossils. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1748), 4724–4733. https://doi.org/10.1098/rspb.2012.1745

Appleby, P. G. (2008). Three decades of dating recent sediments by fallout radionuclides: A review. *Holocene*, *18*(1), 83–93. https://doi.org/10.1177/0959683607085598

Appleby, P.G., Nolan, P.J., Gifford, D.W., Godfrey, M.J., Oldfield, F., Anderson, N.J., & Battarbee R.W. (1986). 210Pb dating by low background gamma counting. *Hydrobiologia*, *143*(1), 21–27. https://doi.org/10.1007/BF00026640

Arp, C. D., Jones, B. M., Grosse, G., Bondurant, A. C., Romanovsky, V. E., Hinkel, K. M., & Parsekian, A. D. (2016). Threshold sensitivity of shallow Arctic lakes and sublake permafrost to changing winter climate. *Geophysical Research Letters*, *43*(12), 6358–6365. https://doi.org/10.1002/2016GL068506

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

Basu, B. K., & Pick, F. R. (1997). Phytoplankton and zooplankton development in a lowland, temperate river. *Journal of Plankton Research*, *19*(2), 237–253. https://doi.org/10.1093/plankt/19.2.237

Baust, J. G. (2008). Strategies for the Storage of DNA. *Biopreservation and Biobanking*, *6*(4), 251–252. https://doi.org/10.1089/bio.2008.0604.lett

Beaulieu, M., Pick, F., & Gregory-Eaves, I. (2013). Nutrients and water temperature are significant predictors of cyanobacterial biomass in a 1147 lakes data set. *Limnology and Oceanography*, *58*(5), 1736–1746. https://doi.org/10.4319/lo.2013.58.5.1736

Beaulieu, M., Pick, F., Palmer, M., Watson, S., Winter, J., Zurawell, R., & Gregory-Eaves, I. (2014). Comparing predictive cyanobacterial models from temperate regions. *Canadian Journal of Fisheries and Aquatic Sciences*, *71*(12), 1830–1839. https://doi.org/10.1139/cjfas-2014-0168

Bekker, A., Holland, H. D., Wang, P. L., Rumble, D., Stein, H. J., Hannah, J. L., Coetzee, L.L., Beukes, N. J. (2004). Dating the rise of atmospheric oxygen. *Nature*, *427*(6970), 117–120. https://doi.org/10.1038/nature02260

Bell, S. G., & Codd, G. A. (1994). Cyanobacterial toxins and human health. *Reviews in Medical Microbiology*, *5*(4), 256–264.

Berman-Frank, I., Lundgren, P., & Falkowski, P. (2003). Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Research in Microbiology*, *154*(3), 157–164. https://doi.org/10.1016/S0923-2508(03)00029-9

Binford, M. W. (1990). Calculation and uncertainty analysis of 210Pb dates for PIRLA project lake sediment cores. *Journal of Paleolimnology*, *3*(3), 253–267. https://doi.org/10.1007/BF00219461

Binladen, J., Gilbert, M. T. P., & Willerslev, E. (2007). 800,000-year-old mammoth DNA, Modern elephant DNA or PCR artefact? *Biological Letters*, *3*, 55–56. https://doi.org/10.1038/362709a0

Bischoff, J. L., Menking, K. M., Fitts, J. P., & Fitzpatrick, J. A. (1997). Climatic Oscillations 10,000-155,000 yr B.P. at Owens Lake, California Reflected in Glacial Rock Flour Abundance and Lake Salinity in Core OL-92. *Quaternary Research*, *48*(3), 313–325. https://doi.org/10.1006/qres.1997.1933

Blais, J. M., & Kalff, J. (1995). The influence of lake morphometry on sediment focusing. *Limnology and Oceanography*, *40*(3), 582–588. https://doi.org/10.4319/lo.1995.40.3.0582

Blankenship, R. E. (2002). *Molecular Mechanisms of Photosynthesis*. (R. E. Blankenship, Ed.) (1st ed., Vol. 1). Oxford: Blackwell Publishing Ltd.

Bonilla, S. & F.R. Pick (2017). Freshwater bloom-forming cyanobacteria and anthropogenic change. Limnology & Oceanography. e-lecture series. https://doi.org/10.1002/loe2.10006

Bormans, M., Sherman, B. S., & Webster, I. T. (1999). Is buoyancy regulation in cyanobacteria an adaptation to exploit separation of light and nutrients? *Marine and Freshwater Research*, *50*(8), 897–906. https://doi.org/10.1071/MF99105

Brasell, K. A., Heath, M. W., Ryan, K. G., & Wood, S. A. (2014). Successional Change in Microbial Communities of Benthic Phormidium-Dominated Biofilms. *Microbial Ecology*, *69*(2), 254–266. https://doi.org/10.1007/s00248-014-0538-7

Burge, D. R. L., Edlund, M. B., & Frisch, D. (2018). Paleolimnology and resurrection ecology: The future of reconstructing the past. *Evolutionary Applications*, *11*(1), 42–59. https://doi.org/10.1111/eva.12556

Burger, J., Hummel, S., Herrmann, B., & Henke, W. (1999). DNA preservation: A microsatellite-DNA study on ancient skeletal remains. *Electrophoresis*, *20*(8), 1722–1728. https://doi.org/10.1002/(SICI)1522-2683(19990101)20:8<1722:AID-ELPS1722>3.0.CO;2-4

Butcher, J. B., Nover, D., Johnson, T. E., & Clark, C. M. (2015). Sensitivity of lake thermal and mixing dynamics to climate change. *Climatic Change*, *129*(1–2), 295–305. https://doi.org/10.1007/s10584-015-1326-1

Capo, E., Domaizon, I., Maier, D., Debroas, D., & Bigler, C. (2017). To what extent is the DNA of microbial eukaryotes modified during burying into lake sediments? A repeat-coring approach on annually laminated sediments. *Journal of Paleolimnology*, *58*(4), 479–495. https://doi.org/10.1007/s10933-017-0005-9

Cardona, T., Sánchez-Baracaldo, P., Rutherford, A. W., & Larkum, A. W. (2019). Early Archean origin of Photosystem II. *Geobiology*, *17*(2), 127–150. https://doi.org/10.1111/gbi.12322

Carey, C. C., Ewing, H. A., Cottingham, K. L., Weathers, K. C., Thomas, R. Q., & Haney, J. F. (2012). Occurrence and toxicity of the cyanobacterium Gloeotrichia echinulata in low-nutrient lakes in the northeastern United States. *Aquatic Ecology*, *46*(4), 395–409. https://doi.org/10.1007/s10452-012-9409-9

Carmichael, Wayne W. "The Toxins of Cyanobacteria." *Scientific American*, vol. 270, no. 1, 1994, pp. 78–86. *JSTOR*, www.jstor.org/stable/24942554.

Carmichael, W. W., Azevedo, S. M. F. O., An, J. S., Molica, R. J. R., Jochimsen, E. M., Lau, S., Rinheart, K.L., Shaw, G.R., Eaglesham, G. K. (2001). Human fatalities from cyanobacteria: Chemical and biological evidence for cyanotoxins. *Environmental Health Perspectives*, *109*(7), 663–668. https://doi.org/10.2307/3454781

Catherine, Q., Susanna, W., Isidora, E. S., Mark, H., Aurélie, V., & Jean-François, H. (2013). A review of current knowledge on toxic benthic freshwater cyanobacteria - Ecology, toxin production and risk management. *Water Research*, *47*(15), 5464–5479. https://doi.org/10.1016/j.watres.2013.06.042

Climate Change 1992, Edited by John Theodore Houghton, pp. 212. ISBN 0521438292. Cambridge, UK: Cambridge University Press, June 1992.

Codd, G. A., Morrison, L. F., & Metcalf, J. S. (2005). Cyanobacterial toxins: Risk management for health protection. *Toxicology and Applied Pharmacology*, *203*(3 SPEC. ISS.), 264–272. https://doi.org/10.1016/j.taap.2004.02.016

Colson, I. B., Bailey, J. F., Vercauteren, M., Sykes, B. C., & Hedges, R. E. M. (1996). The Preservation of Ancient DNA and Bone Diagenesis. *Ancient Biomolecules*, *1*(2), 109–117.

Coolen, M. J. L., & Overmann, J. (1998). Analysis of subfossil molecular remains of purple sulfur bacteria in a lake sediment. *Applied and Environmental Microbiology*, *64*(11), 4513–4521.

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

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

Cornwell, J. C. (1985). Sediment Accumulation Rates in an Alaskan Arctic Lake Using a Modified 210 Pb Technique. *Canadian Journal of Fisheries and Aquatic Sciences*, *42*(4), 809–814. https://doi.org/10.1139/f85-103

Dabney, J., Meyer, M., & Pääbo, S. (2013). Ancient DNA damage. *Cold Spring Harbor Perspectives in Biology*, *5*(7), 1–7. https://doi.org/10.1101/cshperspect.a012567

Davis, M. B., & Botkin, D. B. (1985). Sensitivity of cool-temperate forests and their fossil pollen record to rapid temperature change. *Quaternary Research*, *23*(3), 327–340. https://doi.org/10.1016/0033-5894(85)90039-0

DESC. (2019). Lake Partner Total Phosphorus (TP) Concentration Data, 2002-2017. Retrieved from https://desc.ca/web/sites/default/files/LPP/2017 TP data final.pdf

Desjardins, P., & Conklin, D. (2010). NanoDrop microvolume quantitation of nucleic acids. *Journal of Visualized Experiments*, (45), 1–5. https://doi.org/10.3791/2565

Dodds, W. K., & Smith, V. H. (2016). Nitrogen, phosphorus, and eutrophication in streams. *Inland Waters*, *6*(2), 155–164. https://doi.org/10.5268/IW-6.2.909

Dokulil, M. T., & Teubner, K. (2000). Cyanobacterial dominance in lakes. *Hydrobiologia*, *438*, 1–12. https://doi.org/10.1023/A:1004155810302

Dolman, A. M., Rücker, J., Pick, F. R., Fastner, J., Rohrlack, T., Mischke, U., & Wiedner, C. (2012). Cyanobacteria and cyanotoxins: The influence of nitrogen versus phosphorus. *PLoS ONE*, *7*(6). https://doi.org/10.1371/journal.pone.0038757

Domaizon, I., Savichtcheva, O., Debroas, D., Arnaud, F., Villar, C., Pignol, C., Alric, B., Perga, M. E. (2013). DNA from lake sediments reveals the long-term dynamics and diversity of <i>Synechococcus</i> assemblages. *Biogeosciences*, *10*(6), 3817–3838. https://doi.org/10.5194/bg-10-3817-2013

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

Downing, J. A., Watson, S. B., & Mccauley, E. (2001). RAPID COMMUNICATION / COMMUNICATION RAPIDE Predicting Cyanobacteria dominance in lakes, *1908*(Smith 1983), 1905–1908. https://doi.org/10.1139/cjfas-58-10-1905

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Elliott, J. A. (2012). Is the future blue-green? A review of the current model predictions of how climate change could affect pelagic freshwater cyanobacteria. *Water Research*, *46*(5), 1364–1371. https://doi.org/10.1016/j.watres.2011.12.018

Emanuel, W. R., Shugart, H. H., & Stevenson, M. P. (1985). Climatic change and the broad-scale distribution of terrestrial ecosystem complexes. *Climatic Change*, *7*(1), 29–43. https://doi.org/10.1007/BF00139439

Favot, E. J., Rühland, K. M., DeSellas, A. M., Ingram, R., Paterson, A. M., & Smol, J. P. (2019). Climate variability promotes unprecedented cyanobacterial blooms in a remote, oligotrophic Ontario lake: evidence from paleolimnology. *Journal of Paleolimnology*, *4*, 31–52. https://doi.org/10.1007/s10933-019-00074-4

Fernández, C., Estrada, V., & Parodi, E. R. (2015). Factors triggering cyanobacteria dominance and succession during blooms in a hypereutrophic drinking water supply reservoir. *Water, Air, and Soil Pollution*, *226*(3). https://doi.org/10.1007/s11270-014-2290-5

Fernandez-Carazo, R., Verleyen, E., Hodgson, D. A., Roberts, S. J., Waleron, K., Vyverman, W., & Wilmotte, A. (2013). Late Holocene changes in cyanobacterial community structure in maritime Antarctic lakes. *Journal of Paleolimnology*, *50*(1), 15–31. https://doi.org/10.1007/s10933-013-9700-3

Fleischmann, E. M. (1989). The measurement and penetration of ultraviolet radiation into tropical marine water. *Limnology and Oceanography*, *34*(8), 1623–1629. https://doi.org/10.4319/lo.1989.34.8.1623

Ford, D.E., Stefan, H. (1980) Stratification Variability in Three Morphometrically Different Lakes Under Identical Meteorological Forcing. *Journal of American Water Resources Association*. 16:243–247.

Fox, G.E., Magrum, L. J., Balch, W.E., Ralph, S., & Woese, C. R. (1977). Classification of methanogenic bacteria by 16S rRNA characterization. *Proc. Natl. Acad. Sci. USA*, *74*(10), 4537–4541.

Francis, G. (1878). Poisonous Australian Lake. *Nature*, *18*(444), 11–12. https://doi.org/10.1038/018011d0

Friedmann, E. I., & Ocampo-Friedmann, R. (1984). *The Antarctic Cryptoendolithic Ecosystem: Relevance to Exobiology*. *Origins of Life*.

Garcia-Pichel, F., Belnap, J., Neuer, S., & Schanz, F. (2009). Estimates of global cyanobacterial biomass and its distribution. *Algological Studies*, *109*(1), 213–227. https://doi.org/10.1127/1864-1318/2003/0109-0213

Garcia-Pichel, F., Lombard, J., Soule, T., Dunaj, S., Wu, S. H., & Wojciechowski, M. F. (2019). Timing the evolutionary advent of cyanobacteria and the later great oxidation event using gene phylogenies of a sunscreen. *MBio*, *10*(3), 1–14. https://doi.org/10.1128/mBio.00561-19

Gilbert, M. T. P., Bandelt, H. J., Hofreiter, M., & Barnes, I. (2005). Assessing ancient DNA studies. *Trends in Ecology and Evolution*, *20*(10), 541–544. https://doi.org/10.1016/j.tree.2005.07.005

Graham, L. E., & Wilcox, L. W. (2000). *Algae*. Prentice Hall. Retrieved from https://books.google.ca/books?id=sYXwAAAAMAAJ

Granéli, E., & Turner, J. T. (2006). *An Introduction to Harmful Algae*. *Ecology of Harmful Algae*. https://doi.org/10.1007/978-3-540-32210-8_1

Hamilton, T. L., Bryant, D. A., & Macalady, J. L. (2016). The role of biology in planetary evolution: Cyanobacterial primary production in low-oxygen Proterozoic oceans. *Environmental Microbiology*, *18*(2), 325–340. https://doi.org/10.1111/1462-2920.13118

Hauschild, C. A., McMurter, J. G., & Pick, F. R. (1991). j.0022-3646.1991.00698.x.pdf. *Journal of Phycology*, *27*, 698–702.

Hebert, P. D. N., Muncaster, B. W., & Mackie, G. L. (1989). Ecological and genetic studies on Dreissena polymorpha (Pallas): a new mollusc in the Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *46*(9), 1587–1591. https://doi.org/10.1139/f89-202

Health Canada. (2016). Cyanobacterial Toxins in Drinking Water. Retrieved from https://www.canada.ca/content/dam/canada/health-canada/migration/healthy-canadians/health-system-systeme-sante/consultations/cyanobacteria-cyanobacterie/alt/cyanobacteria-cyanobacterie-eng.pdf

Hebsgaard, M. B., Phillips, M. J., & Willerslev, E. (2005). Geologically ancient DNA: Fact or artefact? *Trends in Microbiology*, *13*(5), 212–220. https://doi.org/10.1016/j.tim.2005.03.010

Heisler, J., Glibert, P., Burkholder, J., Anderson, D., Cochlan, W., Dennison, W., Gobler, C., Dortch, Q., Heil, C., Humphries, E., Lewistus, A., Magnien, R., Marshall, H., Sellner, K., Stockwell, K., Stoecker, D., Suddleson, M.. (2008). Eutrophication and Harmful Algal Blooms: A Scientific Consensus. *Harmful Algae*, *8*(1), 3–13. https://doi.org/10.1016/j.hal.2008.08.006.Eutrophication

Hermanson, M. H. (1990). 210Pb and 137Cs chronology of sediments from small, shallow Arctic lakes. *Geochimica et Cosmochimica Acta*, *54*(5), 1443–1451. https://doi.org/10.1016/0016-7037(90)90167-J

Holden, M. J., Haynes, R. J., Rabb, S. A., Satija, N., Yang, K., & Blasic, J. R. (2009). Factors Affecting Quantification of Total DNA by UV Spectroscopy and PicoGreen Fluorescence. *Journal of Agricultural and Food Chemistry*, *57*(16), 7221–7226. https://doi.org/10.1021/jf901165h

Hudnell, H. K. (2008). *Cyanobacterial harmful algal blooms: State of the science and research needs*. *Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs*. https://doi.org/10.1007/978-0-387-75865-7

Hugerth, L. W., Wefer, H. A., Lundin, S., Jakobsson, H. E., Lindberg, M., Rodin, S., Engstrand, L., Andersson, A. F. (2014). DegePrime, a program for degenerate primer design for broad-taxonomic-range PCR in microbial ecology studies. *Applied and Environmental Microbiology*, *80*(16), 5116–5123. https://doi.org/10.1128/AEM.01403-14

Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M. H., & Visser, P. M. (2018). Cyanobacterial blooms. *Nature Reviews Microbiology*, *16*(8), 471–483. https://doi.org/10.1038/s41579-018-0040-1

IPCC. (2018). Global Warming of 1.5 °C Summary Report

Jankowiak, J., Hattenrath-Lehmann, T., Kramer, B. J., Ladds, M., & Gobler, C. J. (2019). Deciphering the effects of nitrogen, phosphorus, and temperature on cyanobacterial bloom intensification, diversity, and toxicity in western Lake Erie. *Limnology and Oceanography*, *64*(3), 1347–1370. https://doi.org/10.1002/lno.11120

Jeltsch, A. (2013). Oxygen, epigenetic signaling, and the evolution of early life. *Trends in Biochemical Sciences*, *38*(4), 172–176. https://doi.org/10.1016/j.tibs.2013.02.001

Juhel, G., Davenport, J., O'Halloran, J., Culloty, S., Ramsay, R., James, K., Furey, A., Allis, O. (2006). Pseudodiarrhoea in zebra mussels Dreissena polymorpha (Pallas) exposed to microcystins. *Journal of Experimental Biology*, *209*(5), 810–816. https://doi.org/10.1242/jeb.02081

Kalff J (2002) Limnology. Prentice Hall, Upper Saddle River

Kaneko, T., Nakajima, N., Okamoto, S., Suzuki, I., Tanabe, Y., Tamaoki, M., Nakamura, Y., Watanabe, A., Kawashima, K., Kishida, Y., Ono, A., Shimizu, Y., Takahashi, C., Minami, C., Fujishiro, T., Kohara, M., Nakazaki, N., Nakayama, S., Yamada, M., Tabata, S., Watanabe, M.M. (2007). Complete genomic structure of the bloom-forming toxic cyanobacterium microcystis aeruginosa NIES-843. DNA Research, 14(6), 247–256.

Kirchner, G. (2011). 210Pb as a tool for establishing sediment chronologies: Examples of potentials and limitations of conventional dating models. *Journal of Environmental Radioactivity*, *102*(5), 490–494. https://doi.org/10.1016/j.jenvrad.2010.11.010

Kirk, J.T.O. (2010). Light and Photosynthesis in Aquatic Systems. Light and Photosynthesis in Aquatic Ecosystems, third edition. Vol. VI. 1-651. 10.1017/CBO9781139168212.

Knoll, L. B., Sarnelle, O., Hamilton, S. K., Kissman, C. E. H., Wilson, A. E., Rose, J. B., & Morgan, M. R. (2008). Corrigendum: Invasive zebra mussels (Dreissena polymorpha) increase cyanobacterial toxin concentrations in low-nutrient lakes (Canadian Journal of Fisheries and Aquatic Sciences (2008) 65 (448-455)). *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(11), 2552. https://doi.org/10.1139/F08-182

Kopp, R. E., Kirschvink, J. L., Hilburn, I. A., & Nash, C. Z. (2005). The paleoproterozoic snowball Earth: A climate disaster triggered by the evolution of oxygenic photosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(32), 11131–11136. https://doi.org/10.1073/pnas.0504878102

Kosten, S., Huszar, V. L. M., Bécares, E., Costa, L. S., van Donk, E., Hansson, L. A., Jeppeson E., Kruk, C., Lacerot, G., Mazzeo, N., Meester, L.D., Moss, B., Lurling, M., Noges, T., Romo, S., Scheffer, M. (2012). Warmer climates boost cyanobacterial dominance in shallow lakes. *Global Change Biology*, *18*(1), 118–126. https://doi.org/10.1111/j.1365-2486.2011.02488.x

LeBlanc, S., Pick, F. R., & Hamilton, P. B. (2008). Fall cyanobacterial blooms in oligotrophic-to-mesotrophic temperate lakes and the role of climate change. *SIL Proceedings, 1922-2010*, *30*(1), 90–94. https://doi.org/10.1080/03680770.2008.11902091

Lopez R, Cowley A, Li W, McWilliam H.(2013) Using EMBL-EBI Services via Web Interface and Programmatically via Web Services. Curr Protoc Bioinformatics Volume 48 (2014) p.3.12.1-3.12.50. DOI: 10.1002/0471250953.bi0312s48

Magnuson, J. J., Robertson, D. M., Benson, B. J., Wynne, R. H., Livingstone, D. M., Arai, T., Assel, R.A., Barry, R.G., Card, V., Kuusisto, E., Granin, N.G., Prowse, T.D., Stewart, K.M., Vuglinski, V. S. (2000). Historical trends in lake and river ice cover in the Northern Hemisphere. *Science*, *289*(5485), 1743–1746. https://doi.org/10.1126/science.289.5485.1743

Magnuson, J. J., Webster, K. E., Assel, R. A., Bowser, C. J., Dillon, P. J., Eaton, J. G., Evans, H.E., Fee, E.J., Hall, R.I., Mortsch, L.R., Schindler, D.W., Quinn, F. H. (1997). Potential effects of climate changes on aquatic systems: Laurentian Great Lakes and Precambrian Shield region. *Hydrological Processes*, *11*(8), 825–871. https://doi.org/10.1002/(SICI)1099-

Marie-Eve, M. (2017). *Sedimentary DNA reveals centuries of hidden diversity in lake cyanobacterial communities*. *PhD Thesis*. Universite de Montreal. https://doi.org/https://doi.org/10.3929/ethz-b-000166606

Markensten, H., Moore, K., & Persson, I. (2010). Simulated lake phytoplankton composition shifts toward cyanobacteria dominance in a future warmer climate. *Ecological Applications*, *20*(3), 752–767. https://doi.org/10.1890/08-2109.1

McWilliam H., Li W., Uludag M., Squizzato S., Park Y.M., Buso N., Cowley A.P., Lopez R.(2013) Analysis Tool Web Services from the EMBL-EBI Nucleic Acids Research 41: W597-W600. PubMed Id: 23671338 Abstract DOI: 10.1093/nar/gkt376

Miller, M. A., Kudela, R. M., Mekebri, A., Crane, D., Oates, S. C., Tinker, M. T., Staedler, M., Miller, W.A., Toy-Choutka, S., Dominik, C., Hardin, D., Langlois, G., Murray, M., Ward, K., Jessup, D. A. (2010). Evidence for a novel marine harmful algal bloom: Cyanotoxin (microcystin) transfer from land to sea otters. *PLoS ONE*, *5*(9), 1–11. https://doi.org/10.1371/journal.pone.0012576

Monchamp, M. E., Walser, J. C., Pomati, F., & Spaak, P. (2016). Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two perialpine lakes. *Applied and Environmental Microbiology*, *82*(21), 6472–6482. https://doi.org/10.1128/AEM.02174-16

Monchamp, M. E., Walser, J. C., Pomati, F., & Spaak, P. (2016). Sedimentary DNA reveals cyanobacterial community diversity over 200 years in two perialpine lakes. *Applied and Environmental Microbiology*, *82*(21), 6472–6482. https://doi.org/10.1128/AEM.02174-16

Mooij, W. M., Janse, J. H., De Senerpont Domis, L. N., Hülsmann, S., & Ibelings, B. W. (2007). Predicting the effect of climate change on temperate shallow lakes with the ecosystem model PCLake. *Hydrobiologia*, *584*(1), 443–454. https://doi.org/10.1007/s10750-007-0600-2

Mooij, W. M., Hülsmann, S., De Senerpont Domis, L. N., Nolet, B. A., Bodelier, P. L. E., Boers, P. C. M., Dionisio Pires, M., Gons, H.J., Ibelings, Bas,W., Noordhuis, R., Portielje, R., Wolfstein, K., Lammens, E. H. R. R. (2005). The impact of climate change on lakes in the Netherlands: A review. *Aquatic Ecology*, *39*(4), 381–400. https://doi.org/10.1007/s10452-005-9008-0

Moos, M. T., Laird, K. R., & Cumming, B. F. (2005). Diatom assemblages and water depth in Lake 239 (Experimental Lakes Area, Ontario): Implications for paleoclimatic studies. *Journal of Paleolimnology*, *34*(2), 217–227. https://doi.org/10.1007/s10933-005-2382-8

Müller, B., Bryant, L. D., Matzinger, A., & Wüest, A. (2012). Hypolimnetic oxygen depletion in eutrophic lakes. *Environmental Science and Technology*, *46*(18), 9964–9971. https://doi.org/10.1021/es301422r

Muskoka, D. M. of. (2019). Muskoka Water Web. Retrieved from http://www.muskokawaterweb.ca/lake-data/muskoka-data/lake-data-sheets/three-mile-ml

NASA. (2019). NASA Global Climate Change: Vital Signs of the Planet. Retrieved from https://climate.nasa.gov/

Neilan, B. A., Pearson, L. A., Muenchhoff, J., Moffitt, M. C., & Dittmann, E. (2013). Environmental conditions that influence toxin biosynthesis in cyanobacteria. *Environmental Microbiology*, *15*(5), 1239–1253. https://doi.org/10.1111/j.1462-2920.2012.02729.x

Ngwa, F. F., Madramootoo, C. A., & Jabaji, S. (2014). Comparison of cyanobacterial microcystin synthetase (mcy) E gene transcript levels, mcy E gene copies, and biomass as indicators of microcystin risk under laboratory and field conditions. *MicrobiologyOpen*, *3*(4), 411–425. https://doi.org/10.1002/mbo3.173

Noro, M., Masuda, R., Dubrovo, I. A., Yoshida, M. C., & Kato, M. (1998). Molecular phylogenetic inference of the woolly mammoth Mammuthus primigenius, based on complete sequences of mitochondrial cytochrome b and 12S ribosomal RNA genes. *Journal of Molecular Evolution*, *46*(3), 314–326. https://doi.org/10.1007/PL00006308

Nübel, U., Garcia-Pichel, F., & Muyzer, G. (1997). PCR primers to amplify 16S rRNA genes from cyanobacteria. *Applied and Environmental Microbiology*, *63*(8), 3327–3332.

Nürnberg, G. K. (1995). Quantifying anoxia in lakes. *Limnology and Oceanography*, *40*(6), 1100–1111. https://doi.org/10.4319/lo.1995.40.6.1100

Ogram, A., Sayler, G. S., & Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. *Journal of Microbiological Methods*, *7*(2–3), 57–66. https://doi.org/10.1016/0167-7012(87)90025-X

O'Neil, J. M., Davis, T. W., Burford, M. A., & Gobler, C. J. (2012). The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae*, *14*, 313–334. https://doi.org/10.1016/j.hal.2011.10.027

Paerl, H. W. (2006). Assessing and managing nutrient-enhanced eutrophication in estuarine and coastal waters: Interactive effects of human and climatic perturbations. *Ecological Engineering*, *26*(1), 40–54. https://doi.org/10.1016/j.ecoleng.2005.09.006

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

Paerl, H. W. (2016). Impacts of Climate Change on Cyanobacteria in Aquatic Environments. In J. Marxsen (Ed.), *Climate Change and Microbial Ecology: Current Research and Future Trends* (pp. 1–22). Schlitz, Germany: Caister Academic Press.

Paerl, H. W., & Huisman, J. (2008). Bloom like it hot. *Science*, (April), 57–58. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.364.6826&rep=rep1&type=pdf

Paerl, H. W., & Otten, T. G. (2013). Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. *Microbial Ecology*, *65*(4), 995–1010. https://doi.org/10.1007/s00248-012-0159-y

Paerl, H. W., & Paul, V. J. (2012). Climate change: Links to global expansion of harmful cyanobacteria. *Water Research*, *46*(5), 1349–1363. https://doi.org/10.1016/j.watres.2011.08.002

Pagès H, Aboyoun P, Gentleman R, DebRoy S (2019). *Biostrings: Efficient manipulation of biological strings*. R package version 2.54.0.

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

Palmer, M. E., Yan, N. D., Paterson, A. M., & Girard, R. E. (2011). Water quality changes in south-central Ontario lakes and the role of local factors in regulating lake response to regional stressors. *Canadian Journal of Fisheries and Aquatic Sciences*, *68*(6), 1038–1050. https://doi.org/10.1139/f2011-041

Paterson, A. M., Rühland, K. M., Anstey, C. V., & Smol, J. P. (2017). Climate as a driver of increasing algal production in Lake of the Woods, Ontario, Canada. *Lake and Reservoir Management*, *33*(4), 403–414. https://doi.org/10.1080/10402381.2017.1379574

Persaud, A. D., Paterson, A. M., Dillon, P. J., Winter, J. G., Palmer, M., & Somers, K. M. (2015). Forecasting cyanobacteria dominance in Canadian temperate lakes. *Journal of Environmental Management*, *151*(January), 343–352. https://doi.org/10.1016/j.jenvman.2015.01.009

Persaud, A. D., Paterson, A. M., Ingram, R., Yao, H., & Dillon, P. J. (2014). Potential factors leading to the formation of cyanobacterial scums in a mesotrophic softwater lake in Ontario, Canada. *Lake and Reservoir Management*, *30*(4), 331–343. https://doi.org/10.1080/10402381.2014.937841

Pick, F. R. (1991). The abundance and composition of freshwater picocyanobacteria in relation to light penetration. *Limnology and Oceanography*, *36*(7), 1457–1462. https://doi.org/10.4319/lo.1991.36.7.1457

Pick, F. R. (2016). Blooming algae: A Canadian perspective on the rise of toxic cyanobacteria. *Canadian Journal of Fisheries and Aquatic Sciences*, *73*(7), 1149–1158. https://doi.org/10.1139/cjfas-2015-0470

Pick, F. R., & Lean, D. R. S. (1987). The role of macronutrients (C, N, P) in controlling cyanobacterial dominance in temperate lakes. *New Zealand Journal of Marine and Freshwater Research*, *21*(3), 425–434. https://doi.org/10.1080/00288330.1987.9516238

Pilon, S., Zastepa, A., Taranu, Z. E., Gregory-Eaves, I., Racine, M., Blais, J. M., Poulain A.J., Pick, F. R. (2019). Contrasting histories of microcystin-producing cyanobacteria in two temperate lakes as inferred from quantitative sediment DNA analyses. *Lake and Reservoir Management*, *35*(1), 102–117. https://doi.org/10.1080/10402381.2018.1549625

Poulain, A. J., Aris-Brosou, S., Blais, J. M., Brazeau, M., Keller, W. B., & Paterson, A. M. (2015). Microbial DNA records historical delivery of anthropogenic mercury. *ISME Journal*, *9*(12), 2541–2550. https://doi.org/10.1038/ismej.2015.86

Pouria, S., De Andrade, A., Barbosa, J., Cavalcanti, R. L., Barreto, V. T. S., Ward, C. J., Preiser, W., Poon, G.K., Neild, G.H., Codd, G. A. (1998). Fatal microcystin intoxication in haemodialysis unit in Caruaru, Brazil. *Lancet*, *352*(9121), 21–26. https://doi.org/10.1016/S0140-6736(97)12285-1

Przytulska, A., Bartosiewicz, M., & Vincent, W. F. (2017). Increased risk of cyanobacterial blooms in northern high-latitude lakes through climate warming and phosphorus enrichment. *Freshwater Biology*, *62*(12), 1986–1996. https://doi.org/10.1111/fwb.13043

Pugliese, G., & Favero, M. S. (1998). Liver Failure and Death After Exposure to Microcystins at Hemodialysis Center. *Infection Control & Hospital Epidemiology*, *19*(6), 465–465. https://doi.org/10.1017/s0195941700087890

Rantala, A., Rajaniemi-Wacklin, P., Lyra, C., Lepistö, L., Rintala, J., Mankiewicz-Boczek, J., & Sivonen, K. (2006). Detection of microcystin-producing cyanobacteria in Finnish lakes with genus-specific microcystin synthetase gene E (mcyE) PCR and associations with environmental factors. *Applied and Environmental Microbiology*, *72*(9), 6101–6110. https://doi.org/10.1128/AEM.01058-06

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, *51*(5), 1450–1459. https://doi.org/10.1111/1365-2664.12306

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., Kidd, K. A., MacCormack, T.J., Olden, J.D., Ormerod, S.J., Smol, J.P., Taylor, W. W., Tockner, K., Vermaire, J.C., Dudgeon, D., Cooke, S. J. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, *94*(3), 849–873. https://doi.org/10.1111/brv.12480

Reynolds, C. S., Oliver, R. L., & Walsby, A. E. (1987). Cyanobacterial dominance: The role of buoyancy regulation in dynamic lake environments. *New Zealand Journal of Marine and Freshwater Research*, *21*(3), 379–390. https://doi.org/10.1080/00288330.1987.9516234

Rindi F. (2007) Diversity, Distribution and Ecology of Green Algae and Cyanobacteria in Urban Habitats. In: Seckbach J. (eds) Algae and Cyanobacteria in Extreme Environments. Cellular Origin, Life in Extreme Habitats and Astrobiology, vol 11. Springer, Dordrecht

Rinta-Kanto, J. M., Ouellette, A. J. A., Boyer, G. L., Twiss, M. R., Bridgeman, T. B., & Wilhelm, S. W. (2005). Quantification of toxic Microcystis spp. during the 2003 and 2004 blooms in western Lake Erie using quantitative real-time PCR. *Environmental Science and Technology*, *39*(11), 4198–4205. https://doi.org/10.1021/es048249u

Rippey, B., John Anderson, N., & Foy, R. H. (1997). Accuracy of diatom-inferred total phosphorus concentrations and the accelerated eutrophication of a lake due to reduced flushing and increased internal loading. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(11), 2637–2646. https://doi.org/10.1139/f97-158

Robertson, D.M. & Ragotzkie, R.A. (1990). Changes in the thermal structure of moderate to large sized lakes in response to changes in air temperature *Aquatic Science*, *52*(4), 360-380. https://doi.org/10.1007/BF00879763

Rockwell, N. C., Lagarias, J. C., & Bhattacharya, D. (2014). Primary endosymbiosis and the evolution of light and oxygen sensing in photosynthetic eukaryotes. *Frontiers in Ecology and Evolution*, *2*(OCT), 1–13. https://doi.org/10.3389/fevo.2014.00066

Roelke, D. L., Eldridge, P. M., & Cifuentes, L. A. (1999). A model of phytoplankton competition for limiting and nonlimiting nutrients: Implications for development of estuarine and nearshore management schemes. *Estuaries*, *22*(1), 92–104. https://doi.org/10.2307/1352930

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *2016*(10), 1–22. https://doi.org/10.7717/peerj.2584

Rose, K. C., Williamson, C. E., Saros, J. E., Sommaruga, R., & Fischer, J. M. (2009). Differences in UV Transparency and thermal structure between alpine and subalpine lakes: Implications for organisms. *Photochem Photobiological Science*, *8*(9), 1244–1256. https://doi.org/10.1039/b905616e.Differences

Roussel, J. M., Paillisson, J. M., Tréguier, A., & Petit, E. (2015). The downside of eDNA as a survey tool in water bodies. *Journal of Applied Ecology*, *52*(4), 823–826. https://doi.org/10.1111/1365-2664.12428

Rühland, K. M., Paterson, A. M., & Smol, J. P. (2015). Lake diatom responses to warming: reviewing the evidence. *Journal of Paleolimnology*, *54*(1), 1–35. https://doi.org/10.1007/s10933-015-9837-3

Rühland, K. M., Rentz, K., Paterson, A. M., Teller, J. T., & Smol, J. P. (2018). The post-glacial history of northern Lake of the Woods: A multi-proxy perspective on climate variability and lake ontogeny. *Journal of Great Lakes Research*, *44*(3), 367–376. https://doi.org/10.1016/j.jglr.2018.04.002

Sarnelle, O., White, J. D., Horst, G. P., & Hamilton, S. K. (2012). Phosphorus addition reverses the positive effect of zebra mussels (Dreissena polymorpha) on the toxic cyanobacterium, Microcystis aeruginosa. *Water Research*, *46*(11), 3471–3478. https://doi.org/10.1016/j.watres.2012.03.050

Schindler, D. W. (2009). Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. *Limnology and Oceanography*, *54*(6 PART 2), 2349–2358. https://doi.org/10.4319/lo.2009.54.6_part_2.2349

Schindler, D. W. (1997). Widespread effects of climatic warming on freshwater ecosystems in North America. *Hydrological Processes*, *11*(8), 1043–1067. https://doi.org/10.1002/(SICI)1099-1085(19970630)11:8<1043::AID-HYP517>3.0.CO;2-5

Shokere, L. A., Holden, M. J., & Ronald Jenkins, G. (2009). Comparison of fluorometric and spectrophotometric DNA quantification for real-time quantitative PCR of degraded DNA. *Food Control*, *20*(4), 391–401. https://doi.org/10.1016/j.foodcont.2008.07.009

Simcoe Muskoka District Health Unit. (2019). BLUE-GREEN ALGAE. Retrieved from http://www.simcoemuskokahealth.org/Topics/SafeWater/bluegreenalgae_copy1.aspx

Sivonen, K. (2009). Cyanobacterial Toxins Defining Statement Cyanobacteria: General Description Mass Occurrences of Toxic Cyanobacteria Cyanobacterial Toxins Toxin Producers The Biogenesis of Cyanobacterial Toxins Detection Methods for Cyanotoxins Fact. Retrieved from https://s3.amazonaws.com/academia.edu.documents/45799963/Toxic_Cyanobacteria20160520-28017-63ytu3.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1538758306&Signature=b1cGvyLOMcTsG2%2FG49dFqENfyEA%3D&response-content-disposition=inline%3B filename%3DToxic

Smith, D. R., King, K. W., & Williams, M. R. (2015). What is causing the harmful algal blooms in Lake Erie? *Journal of Soil and Water Conservation*, *70*(2), 27A-29A. https://doi.org/10.2489/jswc.70.2.27A

Smol, J. P. (2016). Arctic and Sub-Arctic shallow lakes in a multiple-stressor world: a paleoecological perspective. *Hydrobiologia*, *778*(1), 253–272. https://doi.org/10.1007/s10750-015-2543-3

Smol, J. P. (2008). *Pollution of Lakes and Rivers*. *Library*. https://doi.org/10.1002/aqc.571

Smol, J. P. (2019). Under the radar: Long-term perspectives on ecological changes in lakes. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1906). https://doi.org/10.1098/rspb.2019.0834

Smol, J. P. (1992). Paleolimnology: an important tool for effective ecosystem management. *Journal of Aquatic Ecosystem Health*, *1*(1), 49–58. https://doi.org/10.1007/BF00044408

Smol, J. P., & Douglas, M. S. V. (2007). Crossing the final ecological threshold in high Arctic ponds. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12395–12397. https://doi.org/10.1073/pnas.0702777104

Snel, B., Bork, P., & Huynen, M. A. (1999). Genome phylogeny based on gene content. *Nature Genetics*, *21*(1), 108–110. https://doi.org/10.1038/5052

Sommer, U., Adrian, R., De Senerpont Domis, L., Elser, J. J., Gaedke, U., Ibelings, B., Jeppeson, R., Lurling, M., Molinero, J.C., Mooij, W.M., van Donk, E., Winder, M. (2012). Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. *Annual Review of Ecology, Evolution, and Systematics*, *43*(1), 429–448. https://doi.org/10.1146/annurev-ecolsys-110411-160251

Srivastava, A., Choi, G. G., Ahn, C. Y., Oh, H. M., Ravi, A. K., & Asthana, R. K. (2012). Dynamics of microcystin production and quantification of potentially toxigenic Microcystis sp. using real-time PCR. *Water Research*, *46*(3), 817–827. https://doi.org/10.1016/j.watres.2011.11.056

Stamati, K., Mudera, V., & Cheema, U. (2011). Evolution of oxygen utilization in multicellular organisms and implications for cell signalling in tissue engineering. *Journal of Tissue Engineering*, *2*(1), 1–12. https://doi.org/10.1177/2041731411432365

Steffen, W., Crutzen, P. J., & McNeill, J. R. (2007). The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature. *AMBIO: A Journal of the Human Environment*, *36*(8), 614–621. Retrieved from https://doi.org/10.1579/0044-7447(2007)36[614:TAAHNO]2.0.CO

Stein, R. (1990). Organic carbon content/sedimentation rate relationship and its paleoenvironmental significance for marine sediments. *Geo-Marine Letters*, *10*(1), 37–44. https://doi.org/10.1007/BF02431020

Stoeva, M. K., Aris-Brosou, S., Chételat, J., Hintelmann, H., Pelletier, P., & Poulain, A. J. (2014). Microbial community structure in lake and wetland sediments from a high arctic polar desert revealed by targeted transcriptomics. *PLoS ONE*, *9*(3), 1–12. https://doi.org/10.1371/journal.pone.0089531

Stolte, W., & Riegman, R. (1995). Effect of phytoplankton cell size on transient-state nitrate and ammonium uptake kinetics. *Microbiology*, *141*(5), 1221–1229. https://doi.org/10.1099/13500872-141-5-1221

Stolte, W., McCollin, T., Noordeloos, A. A. M., & Riegman, R. (1994). Effect of nitrogen source on the size distribution within marine phytoplankton populations. *Journal of Experimental Marine Biology and Ecology*, *184*(1), 83–97. https://doi.org/10.1016/0022-0981(94)90167-8

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

Svirčev, Z., Baltić, V., Gantar, M., Juković, M., Stojanović, D., & Baltić, M. (2010). Molecular Aspects of Microcystin-induced Hepatotoxicity and Hepatocarcinogenesis. *Journal of Environmental Science and Health, Part C*, *28*(1), 39–59. https://doi.org/10.1080/10590500903585382

Taranu, Z. E., Gregory-Eaves, I., Leavitt, P. R., Bunting, L., Buchaca, T., Catalan, J., Domaizon, I., Guilizzoni, P., Lami, A., McGowan, H., Morabito, G., Pick, F.R., Stevenson, M.A., Thompson, P.L., Vinebrooke, R. D. (2015). Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecology Letters*, *18*(4), 375–384. https://doi.org/10.1111/ele.12420

Taranu, Z. E., Pick, F. R., Creed, I. F., Zastepa, A., & Watson, S. B. (2019). Meteorological and Nutrient Conditions Influence Microcystin Congeners in Freshwaters. *Toxins*, *11*(11), 620–640.

Te, S. H., Chen, E. Y., & Gin, K. Y. H. (2015). Comparison of quantitative PCR and droplet digital PCR multiplex assays for two genera of bloom-forming cyanobacteria, Cylindrospermopsis and Microcystis. *Applied and Environmental Microbiology*, *81*(15), 5203–5211. https://doi.org/10.1128/AEM.00931-15

Ulrich Sommer, Seasonal Succession of Phytoplankton in Lake Constance, *BioScience*, Volume 35, Issue 6, June 1985, Pages 351–357, https://doi.org/10.2307/1309903

UniProt: a worldwide hub of protein knowledge.Nucleic Acids Res. 47: D506-515 (2019)

Vaitomaa, J., Rantala, A., Halinen, K., Rouhiainen, L., Tallberg, P., Mokelke, L., & Sivonen, K. (2003). Quantitative Real-Time PCR for Determination of Microcystin Synthetase E Copy Numbers for Microcystis and Anabaena in Lakes. *Applied and Environmental Microbiology*, *69*(12), 7289–7297. https://doi.org/10.1128/AEM.69.12.7289-7297.2003

Van Apeldoorn, M. E., Van Egmond, H. P., Speijers, G. J. A., & Bakker, G. J. I. (2007). Toxins of cyanobacteria. *Molecular Nutrition and Food Research*, *51*(1), 7–60. https://doi.org/10.1002/mnfr.200600185

van der Giezen, M. (2011). Mitochondria and the Rise of Eukaryotes. *BioScience*, *61*(8), 594–601. https://doi.org/10.1525/bio.2011.61.8.5

Vincent, W. F. (2007). *Algae and Cyanobacteria in Extreme Environments*. (J. Seckbach, Ed.) (11th ed.). Dordrecht, The Netherlands: Springer International Publishing.

Wagner, C., & Adrian, R. (2009). Cyanobacteria dominance: Quantifying the effects of climate change. *Limnology and Oceanography*, *54*(6 PART 2), 2460–2468. https://doi.org/10.4319/lo.2009.54.6_part_2.2460

Wang, S., Huang, W., & Xiang, M. fang. (2011). [Effect of temperature on the partition coefficient of isoflurane and sevoflurane in perflurocarbon emulsion (Oxygent(TM))]. *Nan Fang Yi Ke Da Xue Xue Bao = Journal of Southern Medical University*, *31*(10), 1718–1720.

Wefer G., Fischer G., Fütterer D.K., Gersonde R., Honjo S., Ostermann D. (1990) Particle Sedimentation and Productivity in Antarctic Waters of the Atlantic Sector. In: Bleil U., Thiede J. (eds) Geological History of the Polar Oceans: Arctic versus Antarctic. NATO ASI Series (Series C: Mathematical and Physical Sciences), vol 308. Springer, Dordrecht

Whitton, B. A. (2013). Preface. *Ecology of Cyanobacteria II: Their Diversity in Space and Time*, (January 2000), v–vi. https://doi.org/10.1007/978-94-007-3855-3

Whitton, B. A., and Potts, M. (2012). "Introduction to the Cyanobacteria," in Ecology of Cyanobacteria II: Their Diversity in Space and Time, ed. B. A. Whitton, 1–13

WHO. (2003). Cyanobacterial toxins: Microcystin-LR in Drinking-water Background document for development of WHO Guidelines for Drinking-water Quality. Retrieved from https://www.who.int/water_sanitation_health/dwq/chemicals/cyanobactoxins.pdf

Wiedner, C., Rücker, J., Brüggemann, R., & Nixdorf, B. (2007). Climate change affects timing and size of populations of an invasive cyanobacterium in temperate regions. *Oecologia*, *152*(3), 473–484. https://doi.org/10.1007/s00442-007-0683-5

Willerslev, E., & Cooper, A. (2005). Ancient DNA. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1558), 3–16. https://doi.org/10.1098/rspb.2004.2813

Willerslev, E., Hansen, A. J., Rønn, R., Brand, T. B., Barnes, I., Wiuf, C., Gilichinsky, D., Mitchell, D., Cooper, A. (2004). Long-term persistence of bacterial DNA. *Current Biology*, *14*(1), 13–14. https://doi.org/10.1016/j.cub.2003.12.012

Winter, J. G., Desellas, A. M., Fletcher, R., Heintsch, L., Morley, A., Nakamoto, L., & Utsumi, K. (2011). Algal blooms in Ontario, Canada: Increases in reports since 1994. *Lake and Reservoir Management*, *27*(2), 105–112. https://doi.org/10.1080/07438141.2011.557765

Wood, S. A., Borges, H., Puddick, J., Biessy, L., Atalah, J., Hawes, I., Dietrich, D. R., Hamilton, D. P. (2017). Contrasting cyanobacterial communities and microcystin concentrations in summers with extreme weather events: insights into potential effects of climate change. *Hydrobiologia*, *785*(1), 71–89. https://doi.org/10.1007/s10750-016-2904-6

Wood, S. A., Rhodes, L., Smith, K., Lengline, F., Ponikla, K., & Pochon, X. (2017). Phylogenetic characterisation of marine Chroococcus-like (Cyanobacteria) strains from the Pacific region. *New Zealand Journal of Botany*, *55*(1), 5–13. https://doi.org/10.1080/0028825X.2016.1205634

Yan, X., Xu, X., Wang, M., Wang, G., Wu, S., Li, Z., Sun, H., Shi, A., Yang, Y. (2017). Climate warming and cyanobacteria blooms: Looks at their relationships from a new perspective. *Water Research*, *125*, 449–457. https://doi.org/10.1016/j.watres.2017.09.008

Zhang, J. Y., Lin, G. M., Xing, W. Y., & Zhang, C. C. (2018). Diversity of growth patterns probed in live cyanobacterial cells using a fluorescent analog of a peptidoglycan precursor. *Frontiers in Microbiology*, *9*(APR), 1–10. https://doi.org/10.3389/fmicb.2018.00791

Zhou, J., Bruns, M. A., & Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology*, *62*(2), 316–322.

Zhu, B. (2006). Degradation of plasmid and plant DNA in water microcosms monitored by natural transformation and real-time polymerase chain reaction (PCR). *Water Research*, *40*(17), 3231–3238. https://doi.org/10.1016/j.watres.2006.06.040

Zuo, J., Chen, L., Shan, K., Hu, L., Song, L., & Gan, N. (2018). Assessment of different mcy genes for detecting the toxic to non-toxic Microcystis ratio in the field by multiplex qPCR. *Journal of Oceanology and Limnology*, *36*(4), 1132–1144. https://doi.org/10.1007/s00343-019-7186-1

Appendix A (Chapter 2)

Appendix: Table 1a

Table 1A: Primer targets sequences for *glnA*, CYA, *mcyE* and *MICR* including the sizes of amplicons and annealing temperatures and dilution factors utilized for ddPCR.

Primer		Gene Target	Primer Sequence	Size (bp)	Annealing Temperature (C)	Dilution Factor
*gln*A	F	Glutamine Synthetase	For 5'-GAT GCC GCC GAT GTA GTA-3'	156	60	1/256
	R		Rev 5'-AAG ACC GCG ACC TTP ATG CC-3'		60	
*mcy*E	F	*mcy* E Microcystin ADDA subunit	For 5'-TAA CTT TTT TGG GCA TAG TCC TG-3'	186	60	1 - 1/9
	R		Rev 5'-CGA ACW GCY GCC ATA ATC GC-3'		60	
*CYA*16s	108F	Cyanobacterial 16s rRNA	For 5'-ACG GGT GAG TAA CRC GTR A-3'	269	54	1/256
	377R		Rev 5'-CCA TGG CGG AAA ATT CCC C-3'		54	
MICR 16s	F	Microcystis specific 16s rRNA	For 5' - GCC GCR AGG TGA AAM CTA A - 3'	247	55	1 - 1/9
	R		Rev 5' - AAT CCA AAR ACC TTC CTC CC - 3'		54.5	
*apc*A	F1	Allophycocyanin Alpha Chain	For 5' -ATG AGC ATC GTC TCC AAC TCG - 3'	116	61.5	1 - 1/9
	R7		Rev 5'-CMC GRC TCT CGC TVA GAA CCT-3'		60.3	1 - 1/9
	R10		Rev 5'- YGC CGT CAT CTC YTC YCC GTA -3'	289	60.4	1 - 1/9

Appendix: Table 2a

Table 2A: Endpoint PCR protocol for *glnA*

***gln*A**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 25uL Samples	1	94	2:00
	2	94	0:20
	3	59	0:20
	4	72	0:20
	5	GOTO Step 2, 34 times	
	6	72	5:00
	7	10	HOLD

Appendix: Table 3a

Table 3a: Endpoint PCR protocol for CYA

***CYA* 16s rRNA**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 25uL Samples	1	95	2:00
	2	94	0:30
	3	Ramp, 2°C/s	
Increasing Gradient Cycle	4	53 to 60	1:00
	5	72	1:00
	6	GOTO Step 2, 34 times	
	7	72	5:00
	8	10	HOLD

Appendix: Table 4a

Table 4a: Endpoint PCR protocol for *mcyE*

***mcy*E (Microcystin marker)**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 25uL Samples	1	96	5:00
	2	94	1:00
	3	60	1:00
	4	72	1:00
	5		GOTO Step 2, 30 times
	6	72	10:00
	7	10	HOLD

Appendix: Table 5a

Table 5a: Endpoint PCR protocol for *MICR*

***MICR* specific 16s rRNA**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 25uL Samples	1	94	5:00
	2	94	1:00
	3	54.6	1:00
	4	72	5:00
	5	GOTO Step 2, 34 times	
	6	72	5:00
	7	10	HOLD

Appendix: Table 6a

Table 6a: ddPCR gradient test protocol

ddPCR Gradient	Step	Temperature (C°)	Time
Lid 105°C, 25uL Samples	1	95	5:00
	2	95	0:30
	3	51-63	1:00
	4	72	0:30
	5	GOTO Step 2, 49 times	
	6	4	5:00
	7	90	5:00
	8	10	HOLD

Appendix: Table 7a

Table 7a: ddPCR protocol for *glnA*

***gln*A (Glutamine Synthetase)**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 40uL Samples	1	95	5:00
		Set Ramp 2°C/s	
	2	95	0:30
		Set Ramp 2°C/s	
	3	59.5	1:00
		Set Ramp 2°C/s	
	4	72	0:30
		Set Ramp 2°C/s	
	5	GOTO Step 2, 49 times	
	6	4	5:00
		Set Ramp 2°C/s	
	7	10	HOLD

Appendix: Table 8a

Table 8a: ddPCR protocol for CYA

CYA 16s rRNA	Step	Temperature (C°)	Time
Lid 105°C, 40uL Samples	1	95	5:00
	2	95	0:30
		Set Ramp 2°C/s	
	3	53.2	1:00
	4	GOTO Step 2, 39 times	
	5	4	5:00
	6	90	5:00
	7	10	HOLD

Appendix: Table 9a

Table 9a: ddPCR protocol for *mcyE*

***mcy*E**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 40uL Samples	1	95	5:00
	2	95	0:30
		Set Ramp 2°C/s	
	3	59.5	1:00
	4	GOTO Step 2, 39 times	
	5	4	5:00
	6	90	5:00
	7	10	HOLD

Appendix: Table 10a

Table 10a: ddPCR protocol for *MICR*

***MICR* specific 16s rRNA**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 40uL Samples	1	94	5:00
	2	94	1:00
		Set Ramp 2°C/s	
	3	52.8	1:00
		Set Ramp 2°C/s	
	4	72	5:00
	5	GOTO Step 2, 34 times	
	6	72	5:00
		Set Ramp 2°C/s	
	7	10	HOLD

Appendix: Figure 1a

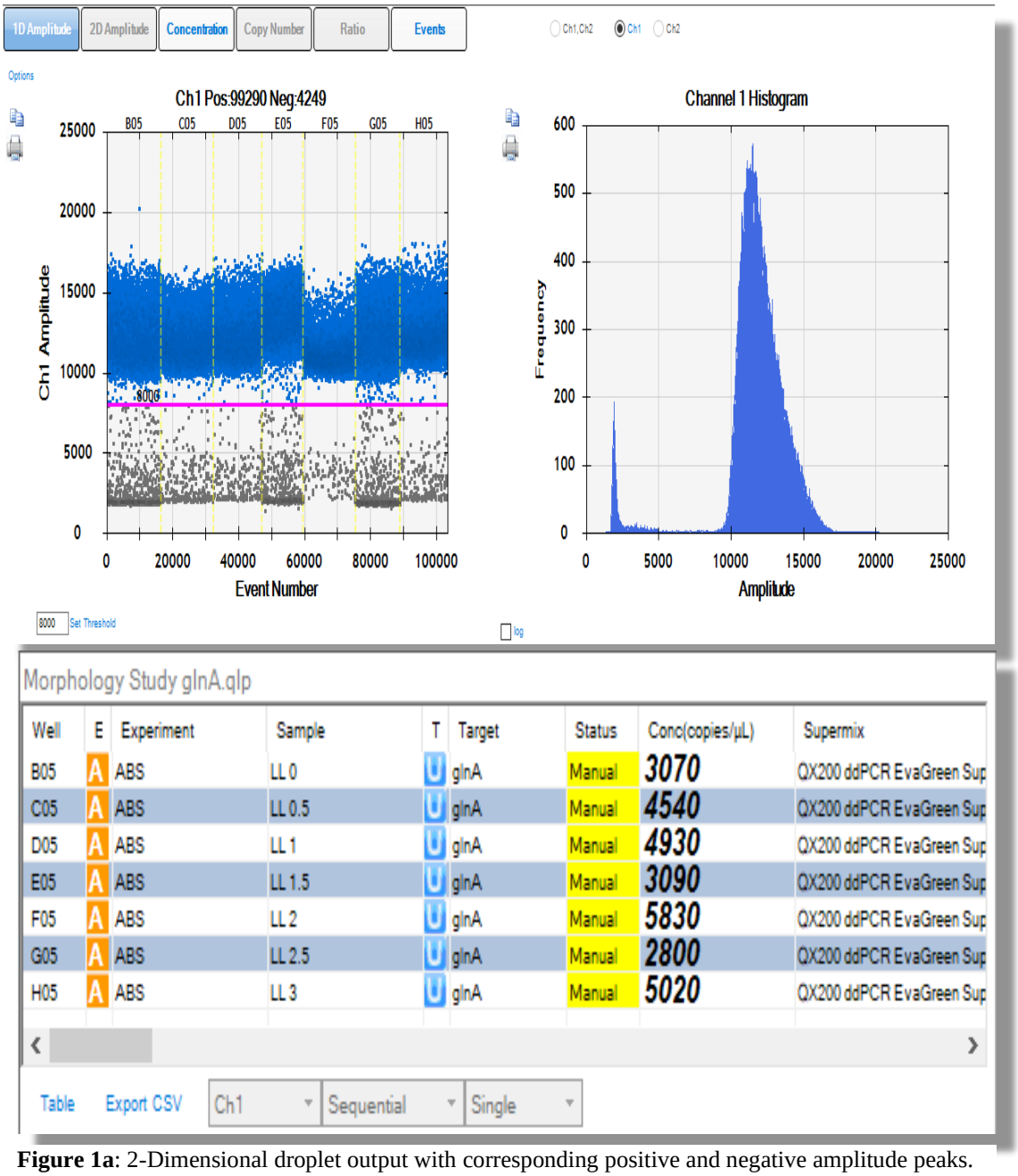
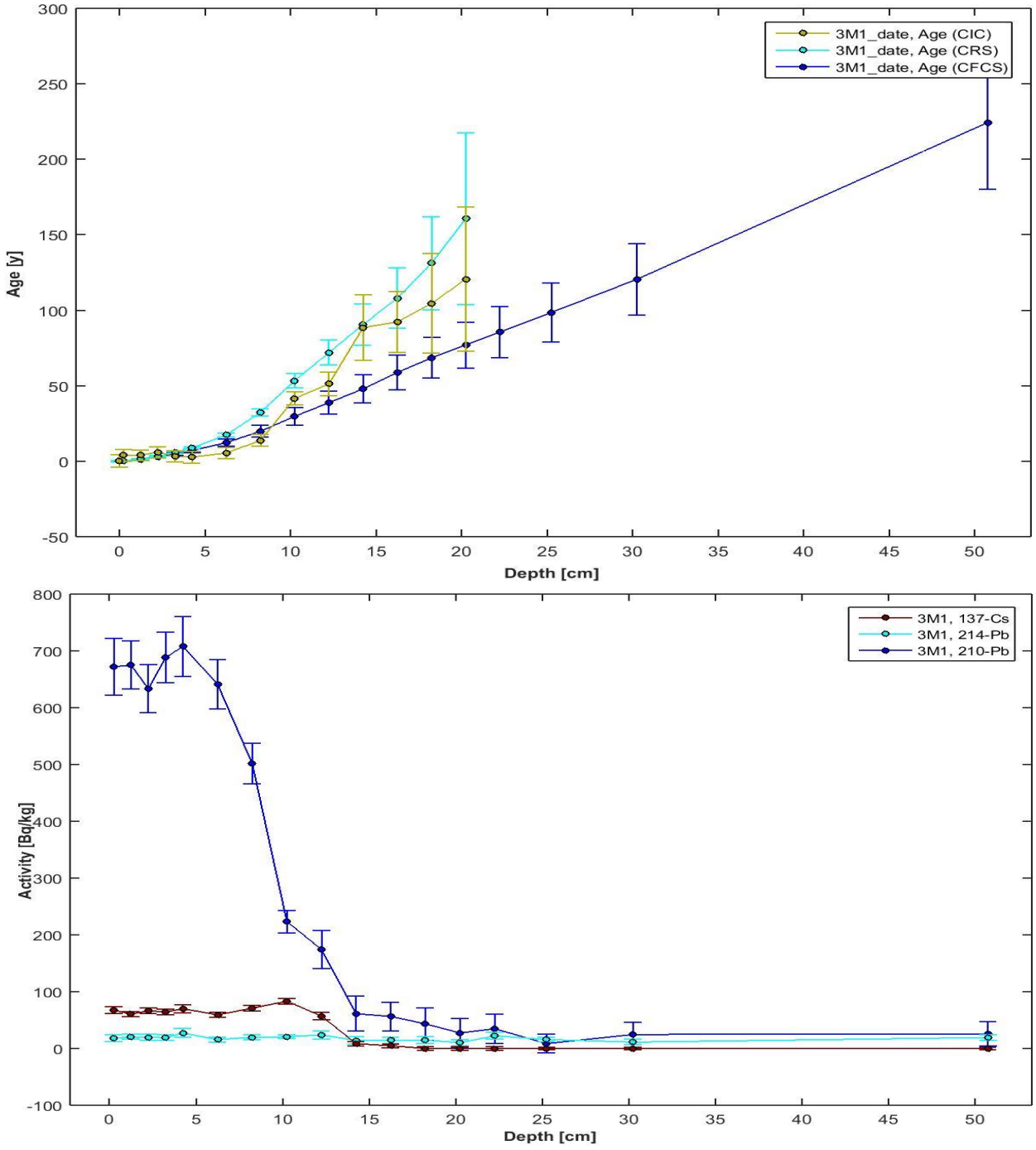


Figure 1a: 2-Dimensional droplet output with corresponding positive and negative amplitude peaks. A threshold (the purple line) Cut-off threshold to either restrict or be more liberal in accepting these values with more conservative thresholds often being preferred. Droplets floating between positive negative, referred to as "rain", may indicate partially successful amplification, inhibited amplification or off-target amplification. **Bottom:** Final copy per µL output for Glutamine Synthetase gene target in Big Rideau Lake samples. QuantaSoft™ has counted positive and negative droplets and generated a concentration of the gene target which can then be further manipulated.

Appendix: Figure 2a

Model and Isotope Profile Three Mile Lake Core



Appendix Figure 2a: Three Mile lake sediment dating profiles ***TOP:*** The fitting of three sediment dating models ***BOTTOM:*** The isotope peak profiles of ^{137}Cs and ^{210}Pb

Appendix: Figure 3a

Model and Isotope Profile Blue Chalk Lake Core

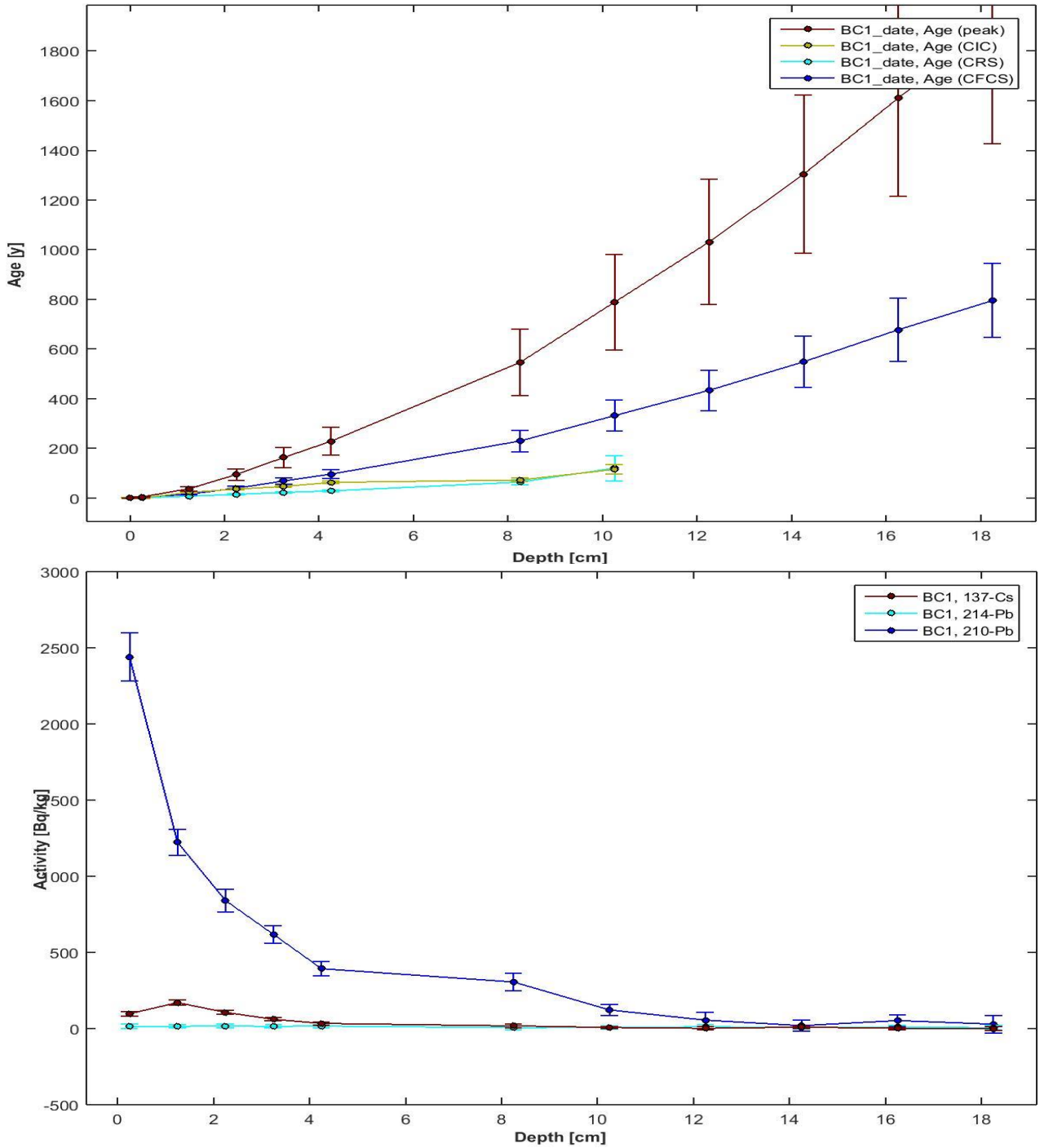


Figure 3a: Blue Chalk lake sediment dating profiles **TOP:** The fitting of three sediment dating models **BOTTOM:** The isotope peak profiles of ^{137}Cs and ^{210}Pb

Appendix: Figure 4a

Model and Isotope Profile Big Rideau Lake Core

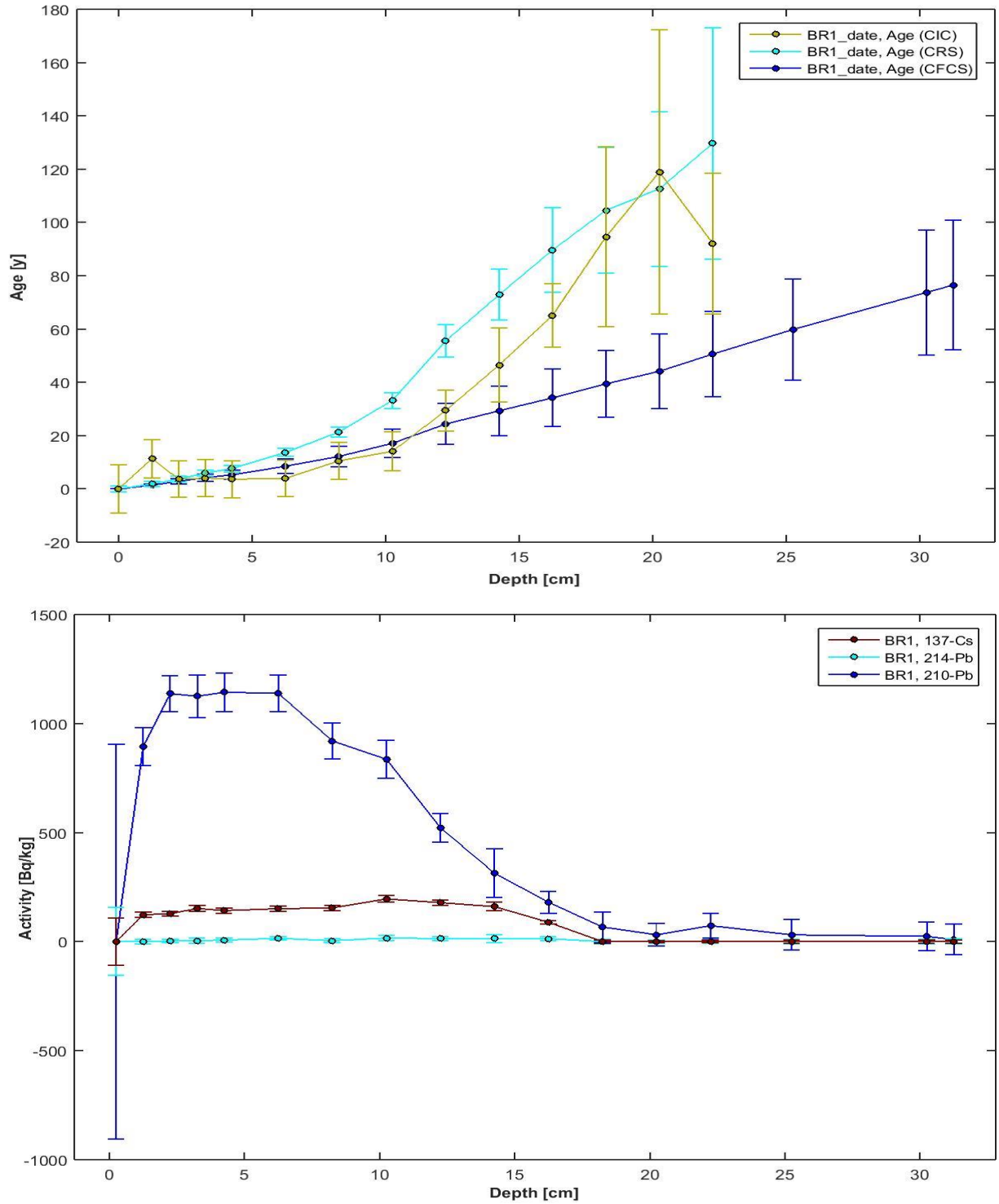


Figure 4a: Big Rideau lake sediment dating profiles **TOP:** The fitting of three sediment dating models **BOTTOM:** The isotope peak profiles of ^{137}Cs and ^{210}Pb

Appendix: Figure 5a

Model and Isotope Profile Otty Lake Core

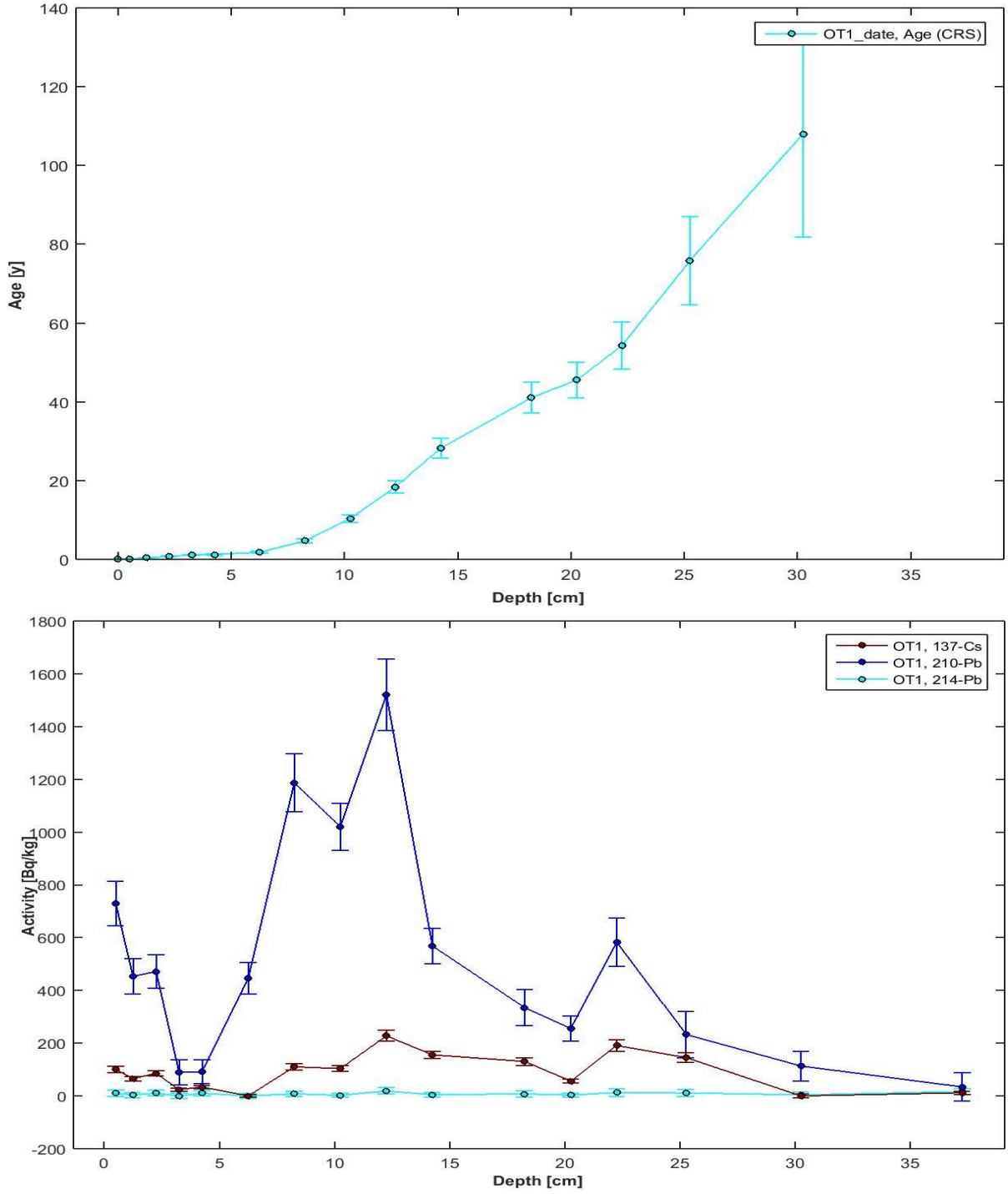


Figure 5a: Otty lake sediment dating profiles **TOP:** The fitting of three sediment dating models **BOTTOM:** The isotope peak profiles of ^{137}Cs and ^{210}Pb showing minor mid-core mixing,

Appendix: Figure 6a

Minimum Temperature (Muskoka):

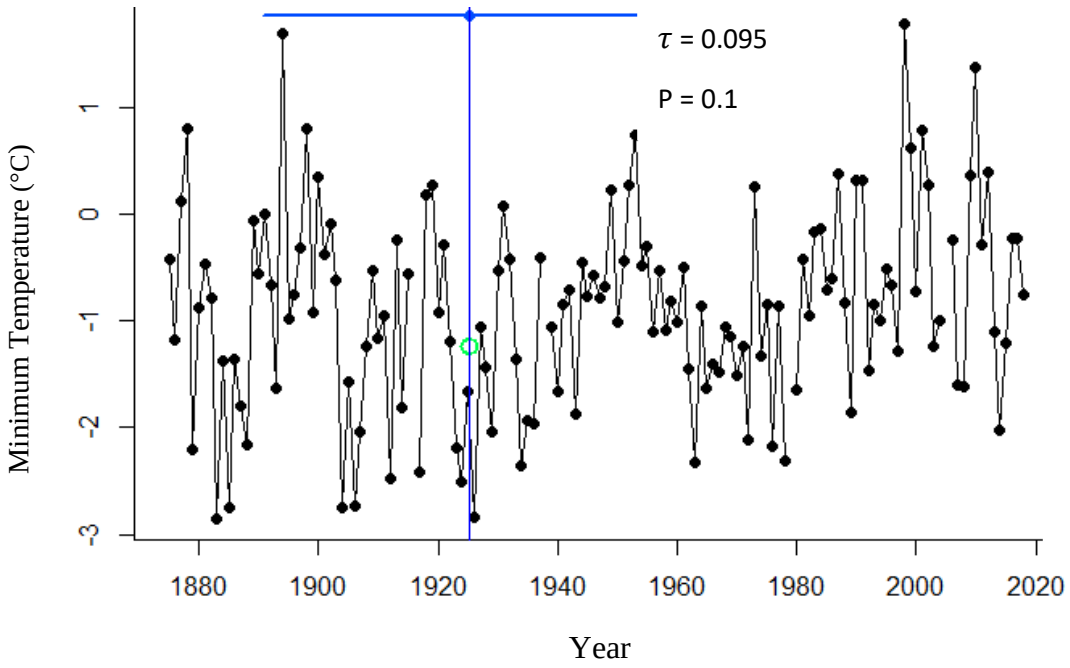


Figure 6a: Minimum Temperature (°C) over Time (Year). Blue line indicates the "Breakpoint" where pattern begins to deviate (1942). τau (τ) indicates positive trend and p indicates a non-significance of the trend.

Appendix: Figure 7a

Cooling Degree Days (Muskoka):

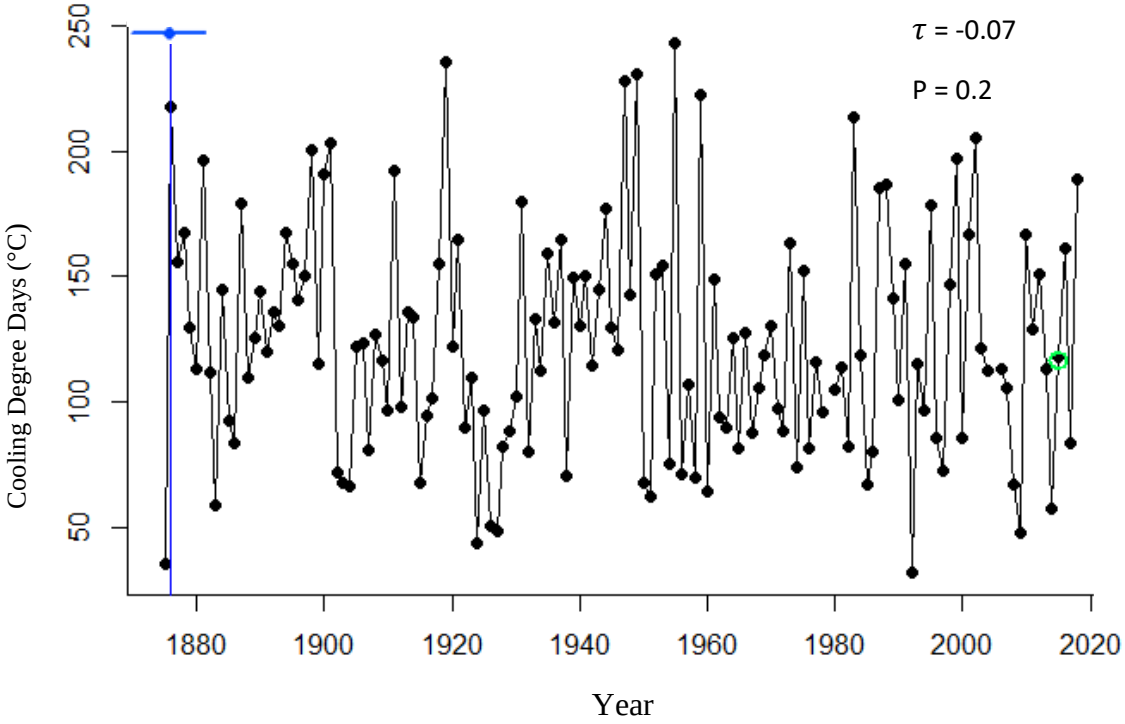


Figure 7a: Cooling degrees (°C) over Time (Year). Blue line indicates the "Breakpoint" (1875) where pattern begins to deviate. tau (τ) indicates negative directionality. p indicates trend is non-significant.

Appendix: Figure 8a

Precipitation Average (Daily Average, Muskoka)

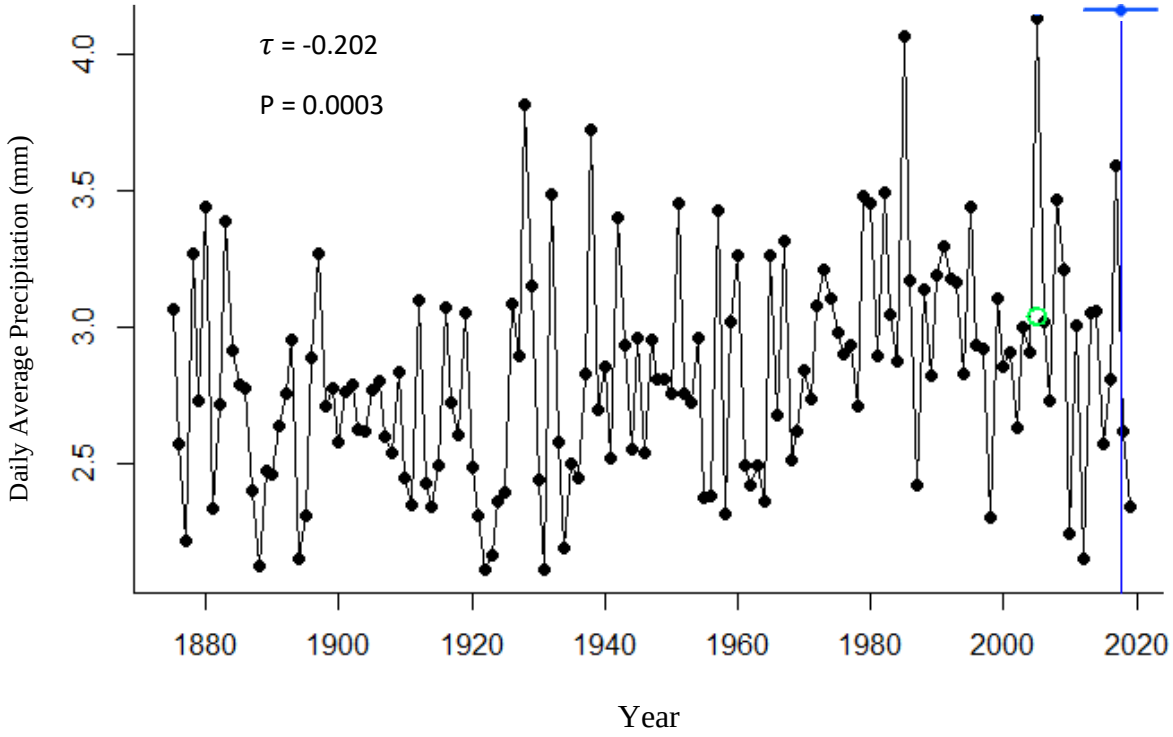


Figure 8a: Precipitation Daily Average (mm) over Time (Year). Blue line indicates the "Breakpoint" (2017) where pattern begins to deviate. ταυ (τ) indicates positive directionality and p value (<0.05) indicates significance of trend.

Appendix: Figure 9a

Minimum Temperature (Rideau):

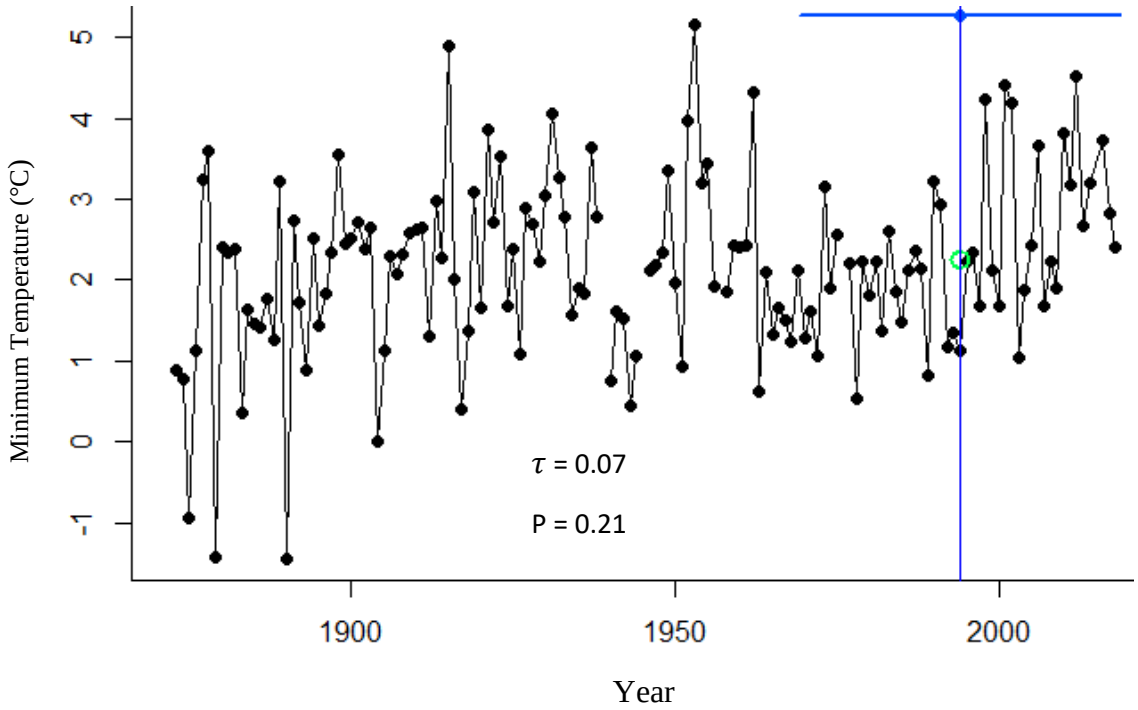


Figure 9a: Minimum Temperature (°C) over Time (Year). Blue line indicates the "Breakpoint" (1994) where pattern begins to deviate. Tau (τ) indicates positive directionality and p value indicates non-significance.

Appendix Figure 10a

Cooling Degree Days (Rideau):

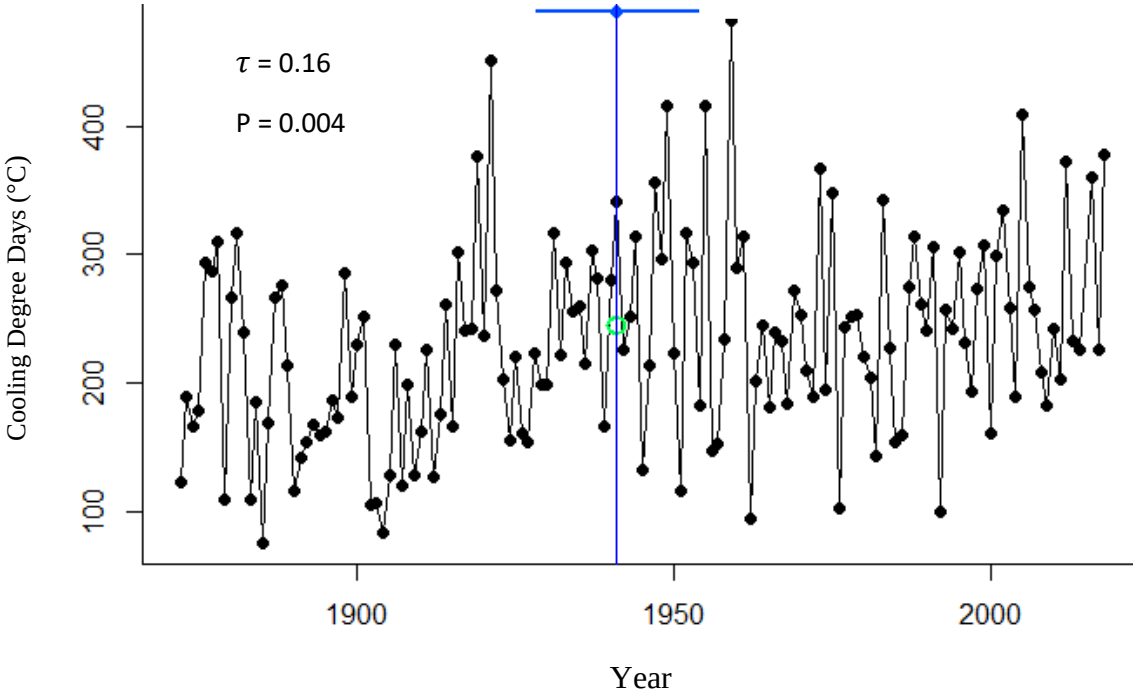


Figure 9a: Cooling Degree Days (°C) over Time (Year). Blue line indicates the "Breakpoint" (1944) where pattern begins to deviate. ταυ (τ) indicates positive directionality and p value (<0.05) indicates significance.

Appendix: Figure 11a

Precipitation Total (yearly average, Rideau):

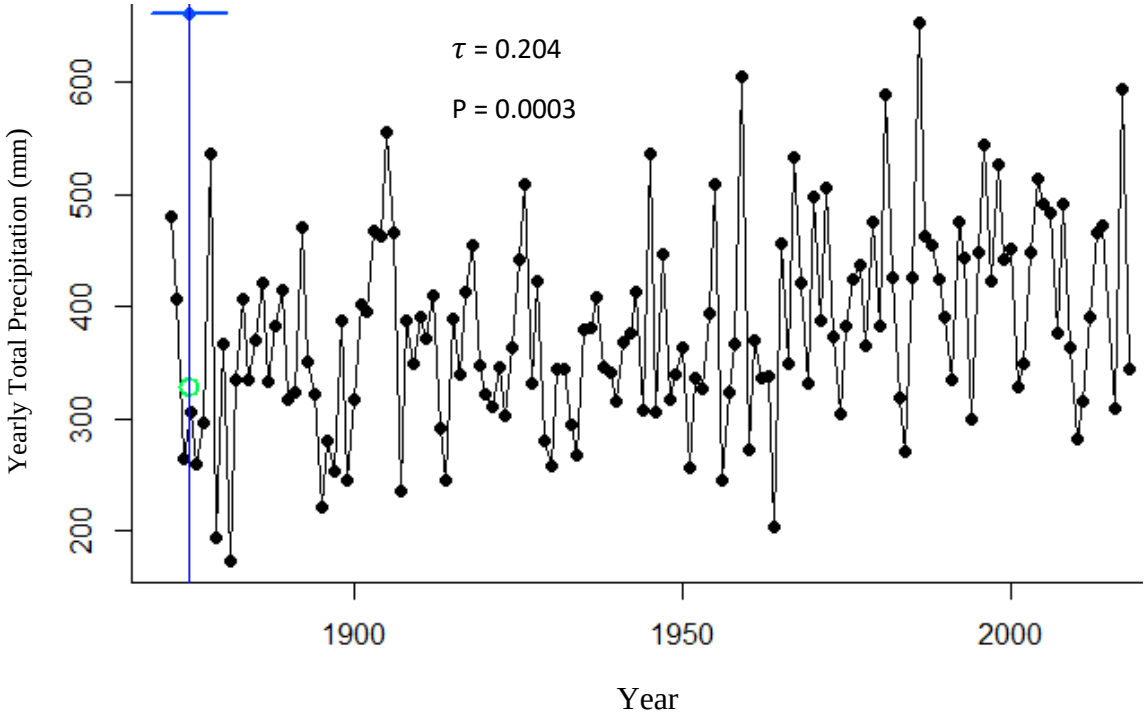


Figure 11a: Precipitation Daily Avg (mm) over Time (Year). Blue line indicates the "Breakpoint" (1888) where pattern begins to deviate. Tau (τ) indicates positive directionality and p value (<0.05) indicates significance.

Appendix: Figure 12a

Total Kjeldahl Nitrogen to Total Phosphorus Correlation Matrices for Three Mile lake and Blue Chalk lake

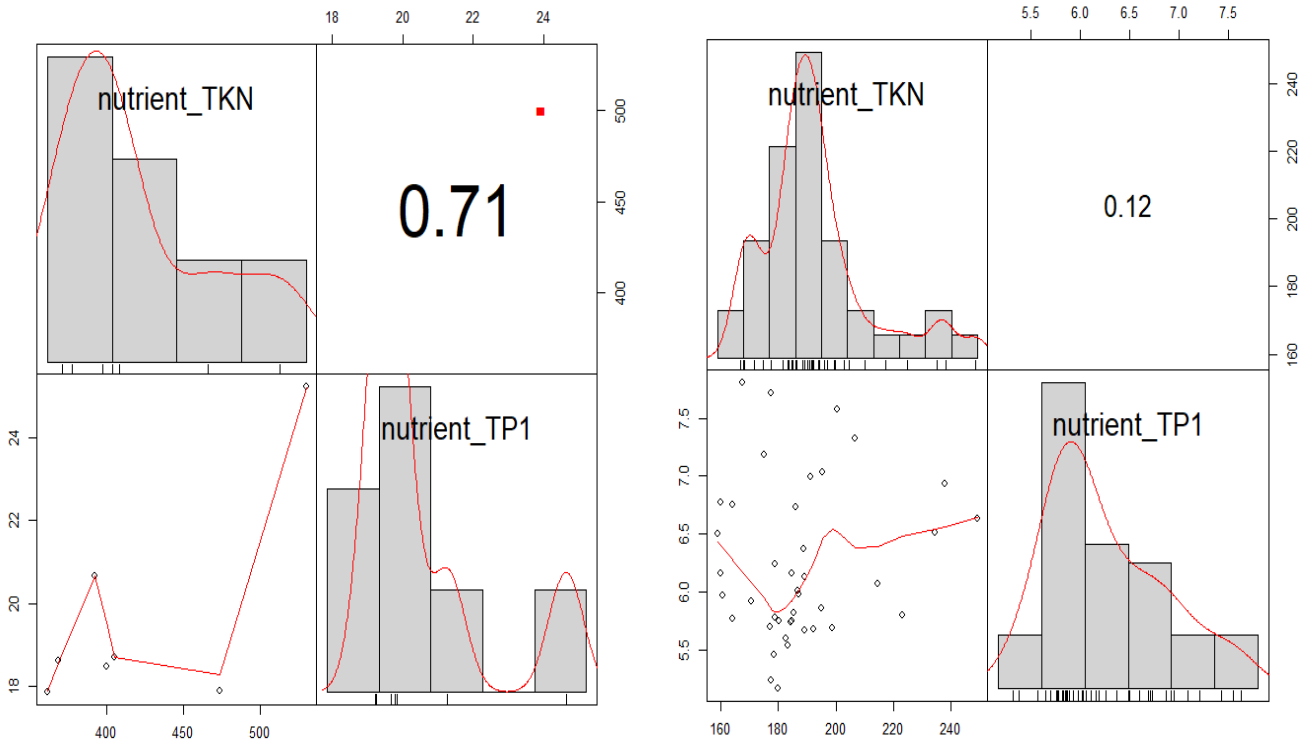


Figure 12a: Kendall Rank correlation plots of total phosphorus (nutrient_TP1 in ug/L) against total Kjeldahl nitrogen (TKN in ug/L). **LEFT:** Three Mile lake shows a moderately significant correlation between the two nutrient factors. **RIGHT:** Blue Chalk lake shows no significant correlation between the two nutrient factors.

Appendix: Figure 13a

Total Kjeldahl Nitrogen over time for Three Mile lake and Blue Chalk lake

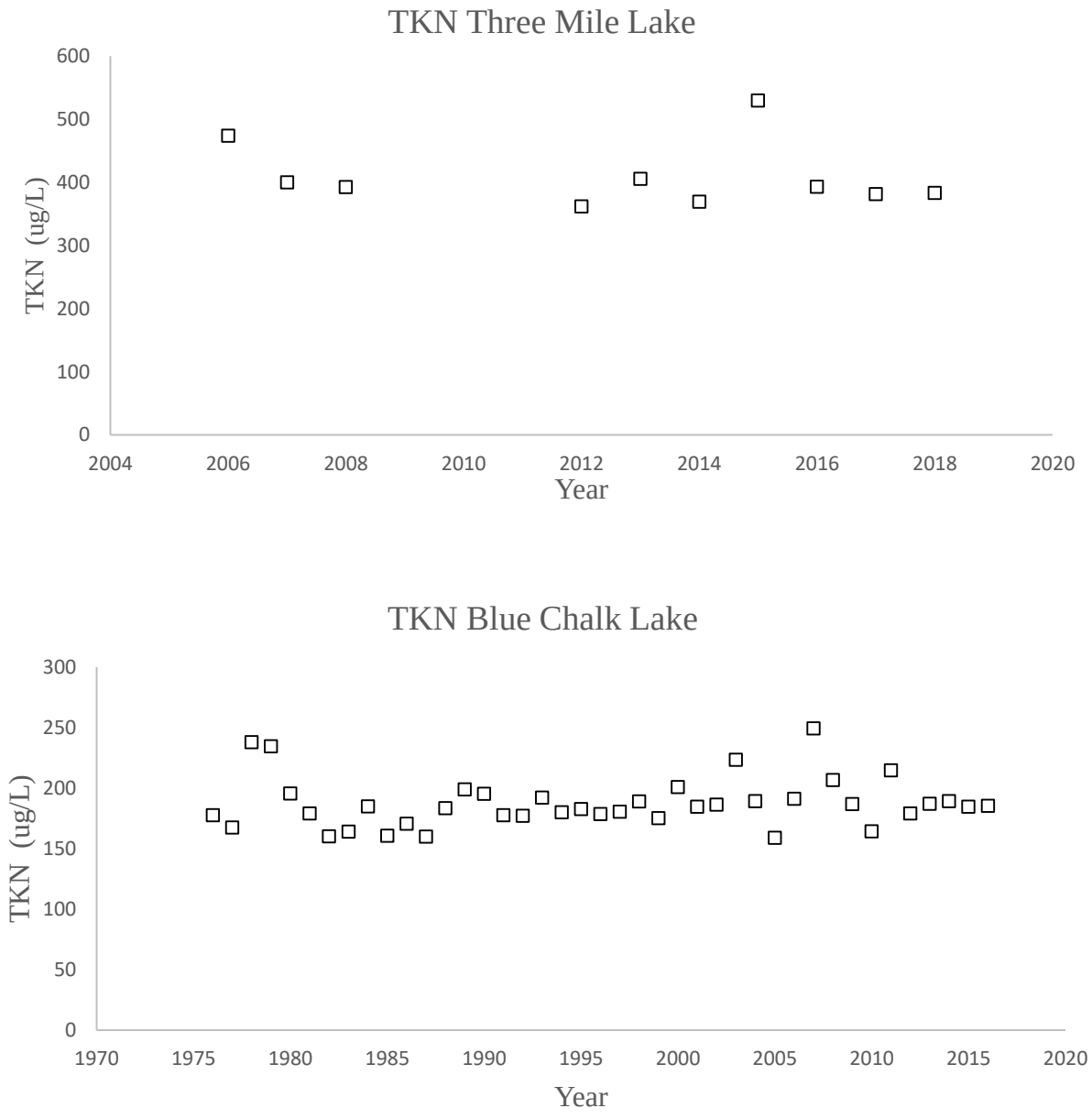


Figure 13a: Average yearly Total Kjeldahl nitrogen (μg/L) over time (Year) in the Muskoka region. **Top:** Three Mile TKN over time. No significant trend was found (p=>0.05). Confirmed through 500 permutation bootstrap analysis. **Bottom**: Blue Chalk lake TKN over time. TKN exhibited no significant upwards or downwards trend (p=>0.05). Confirmed with 500 permutation bootstrap analysis.

Appendix 14a

Nucleic Acid 260:280 Ratio's for Three Mile lake and Blue Chalk lake Sediment Extractions

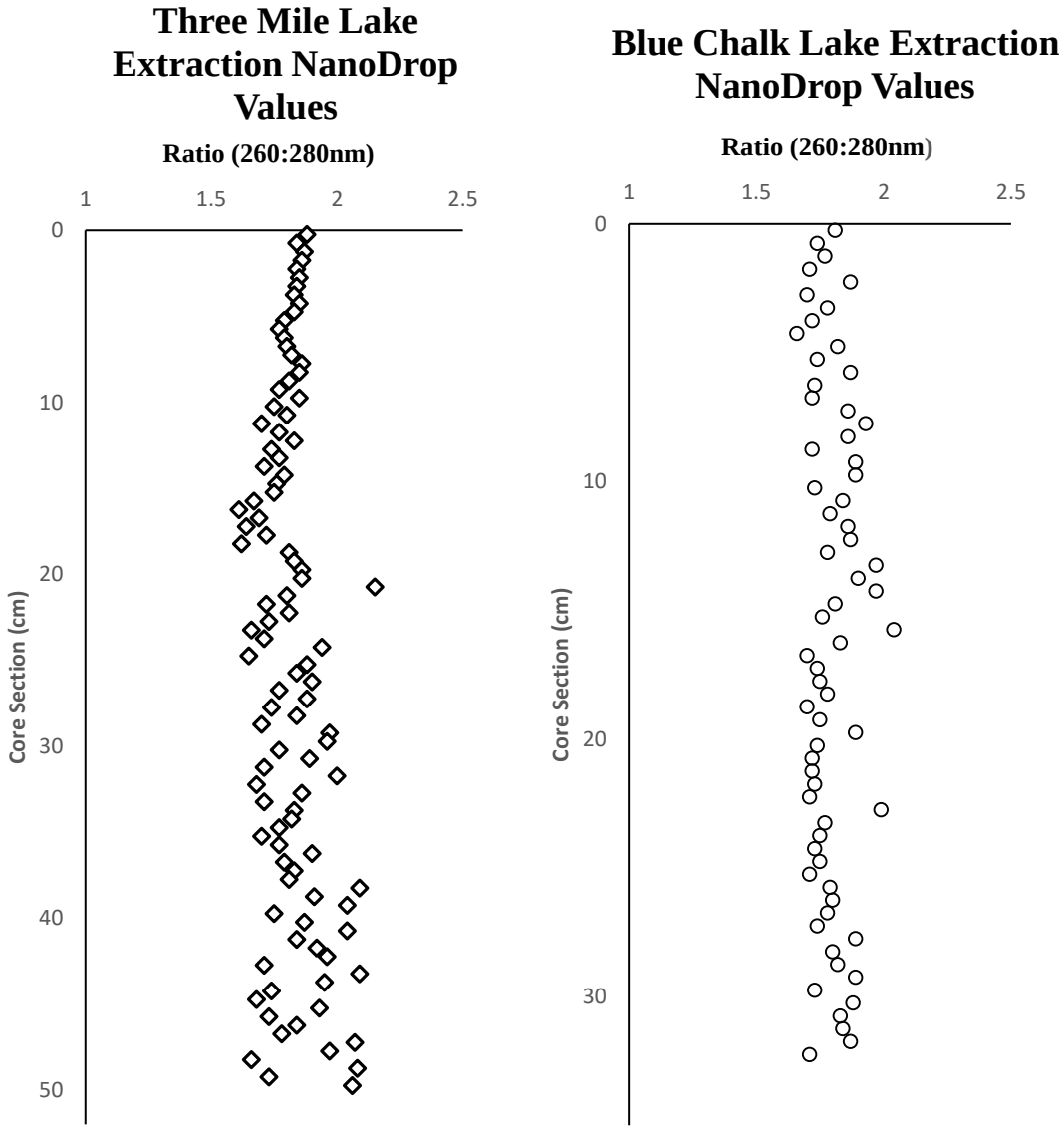
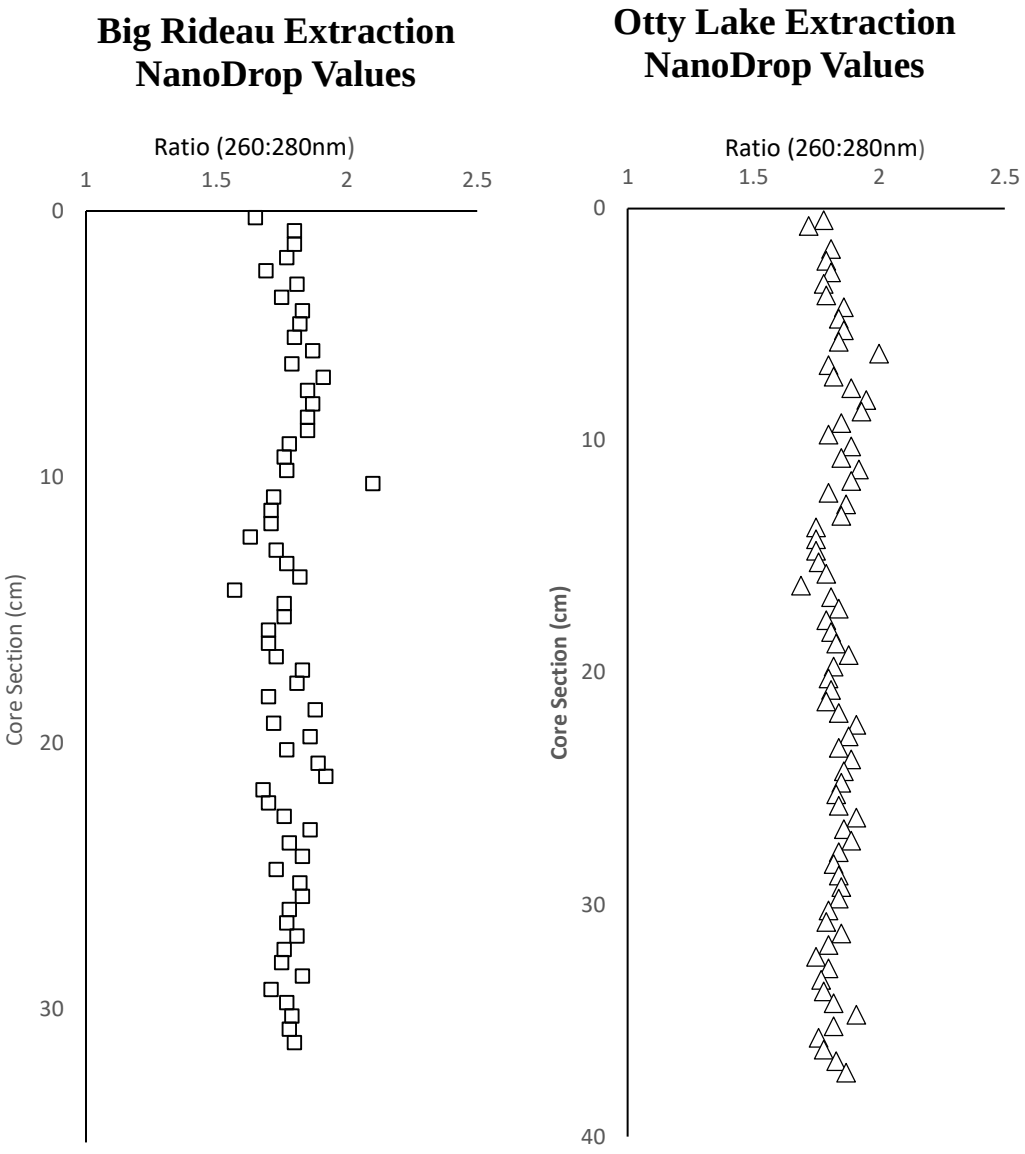


Figure 14a: NanoDrop™ nucleic acid extraction *quality* estimates from 260:280 ratio's in the Muskoka region lakes. **Left:** Three Mile Lake **Right:** Blue Chalk lake.

Appendix 15a

Nucleic Acid 260:280 Ratio's for Big Rideau lake and Otty Lake Sediment Extractions



Appendix 15a: NanoDrop™ nucleic acid extraction quality estimates from 260:280 ratio's in the Rideau region lakes **Left:** Big Rideau lake **Right:** Otty Lake

Appendix: Figure 16a

Breakpoint Analysis: Gene Targets in Three Mile Lake

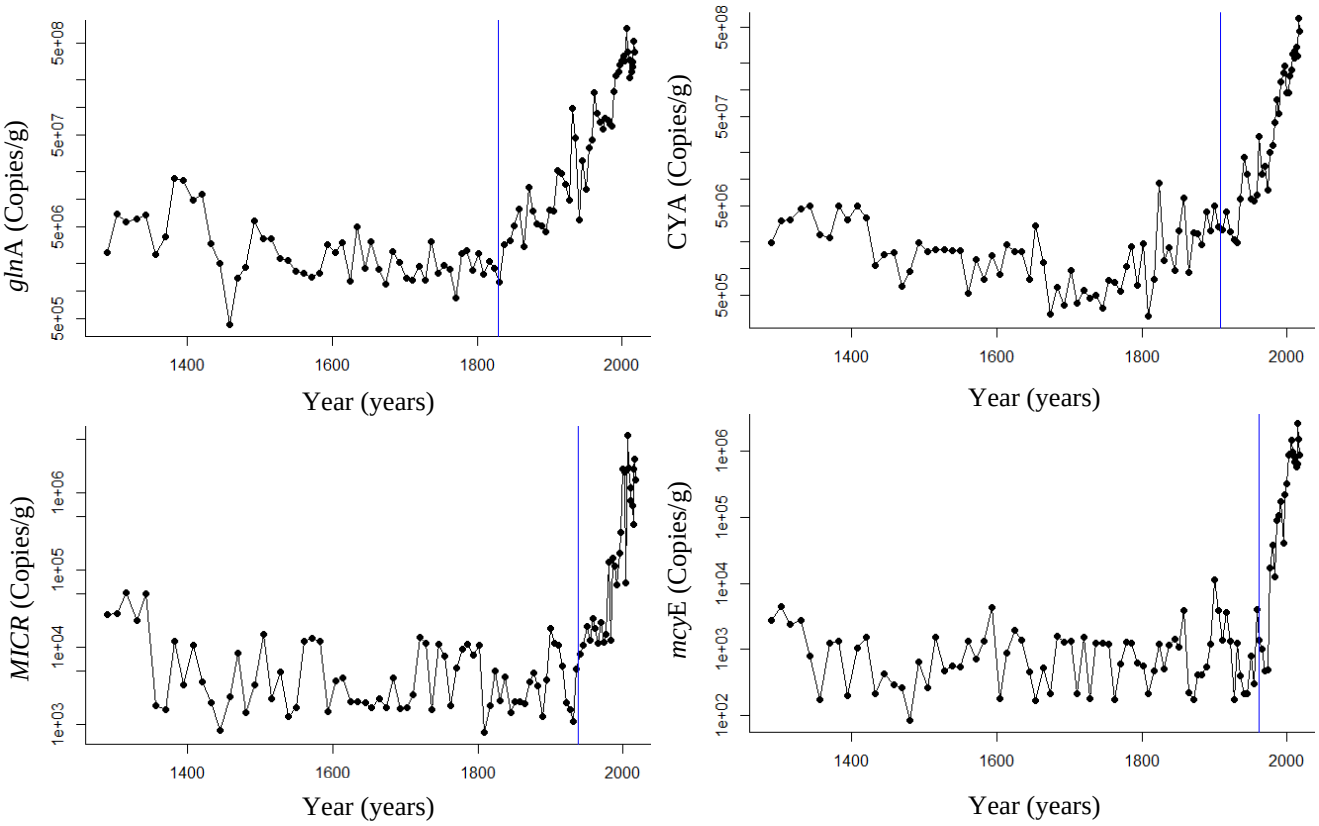


Figure 16a: Breakpoint analysis examining gene copy number of each gene target through time in Three Mile lake. Breakpoint estimates the year in which gene copy trend alters significantly. **TOP LEFT:** *gln*A gene copy increases through time with an initial breakpoint identified at 1834. **TOP RIGHT:** *CYA* gene copy increases through time with an initial breakpoint at 1908. **BOTTOM LEFT:** *MICR* gene copy number increasing through time with an initial breakpoint identified in 1937. **BOTTOM RIGHT:** *McyE* gene copy number increasing through time a much later breakpoint was identified in the year 1961.

Appendix: Figure 17a

Breakpoint Analysis: Gene Targets in Blue Chalk Lake

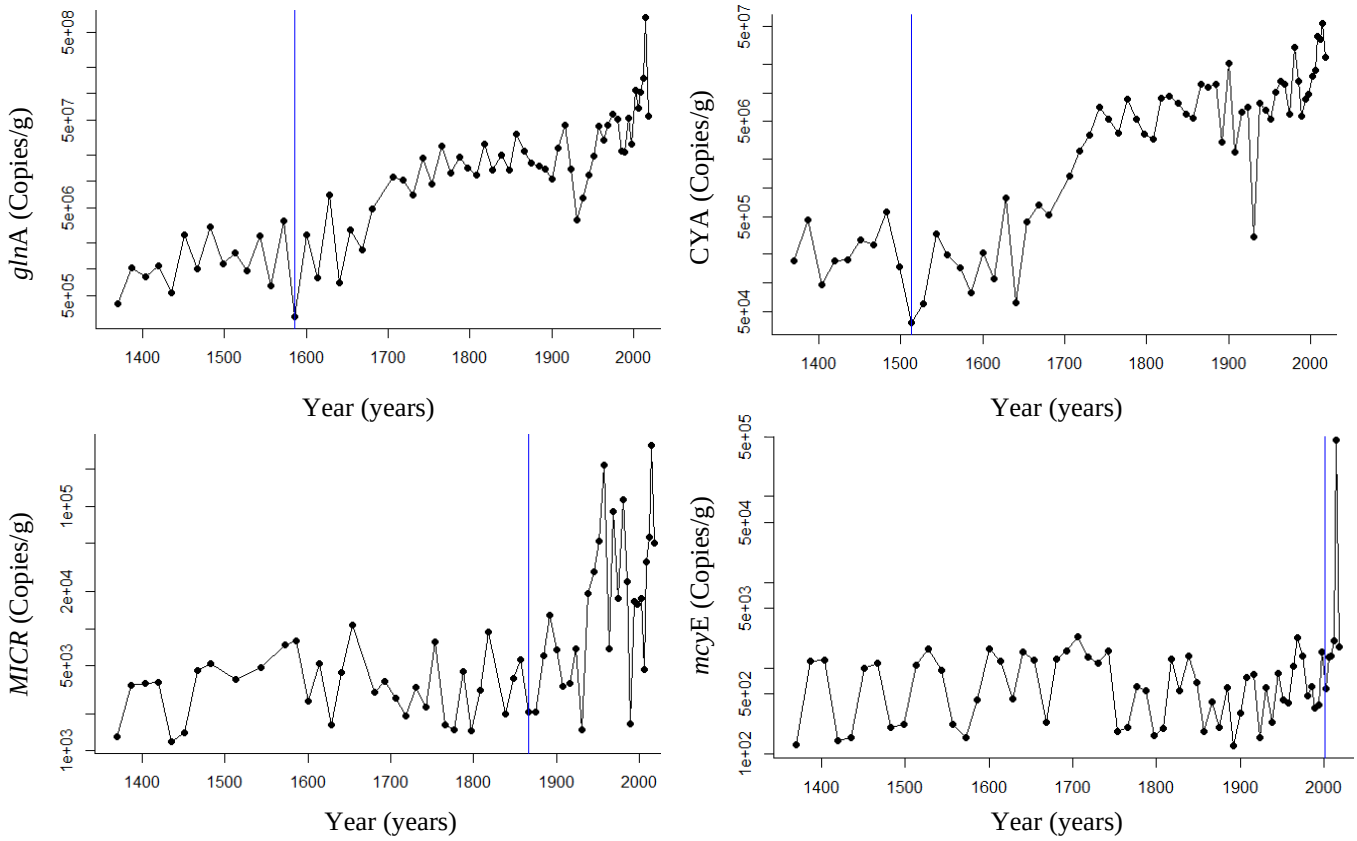


Figure 17a: Breakpoint analysis examining gene copy number of each gene target through time in Blue Chalk lake. Breakpoint estimates the year in which gene copy trend begins to differ significantly from established trends. **TOP LEFT:** *gln*A gene copy increases through time with an initial breakpoint identified at 1586. **TOP RIGHT:** *CYA* gene copy increases through time with an initial breakpoint at the year 1517. **BOTTOM LEFT:** *MICR* gene copy number increasing through time with an initial breakpoint identified in 1866, very similar to the breakpoint identified for *gln*A. **BOTTOM RIGHT:** *McyE* gene copy number remained basically unchanged at detection until a breakpoint was identified in the year 2002 based on two increased measurements.

Appendix: Figure 18a

Breakpoint Analysis: Gene Targets in Big Rideau Lake

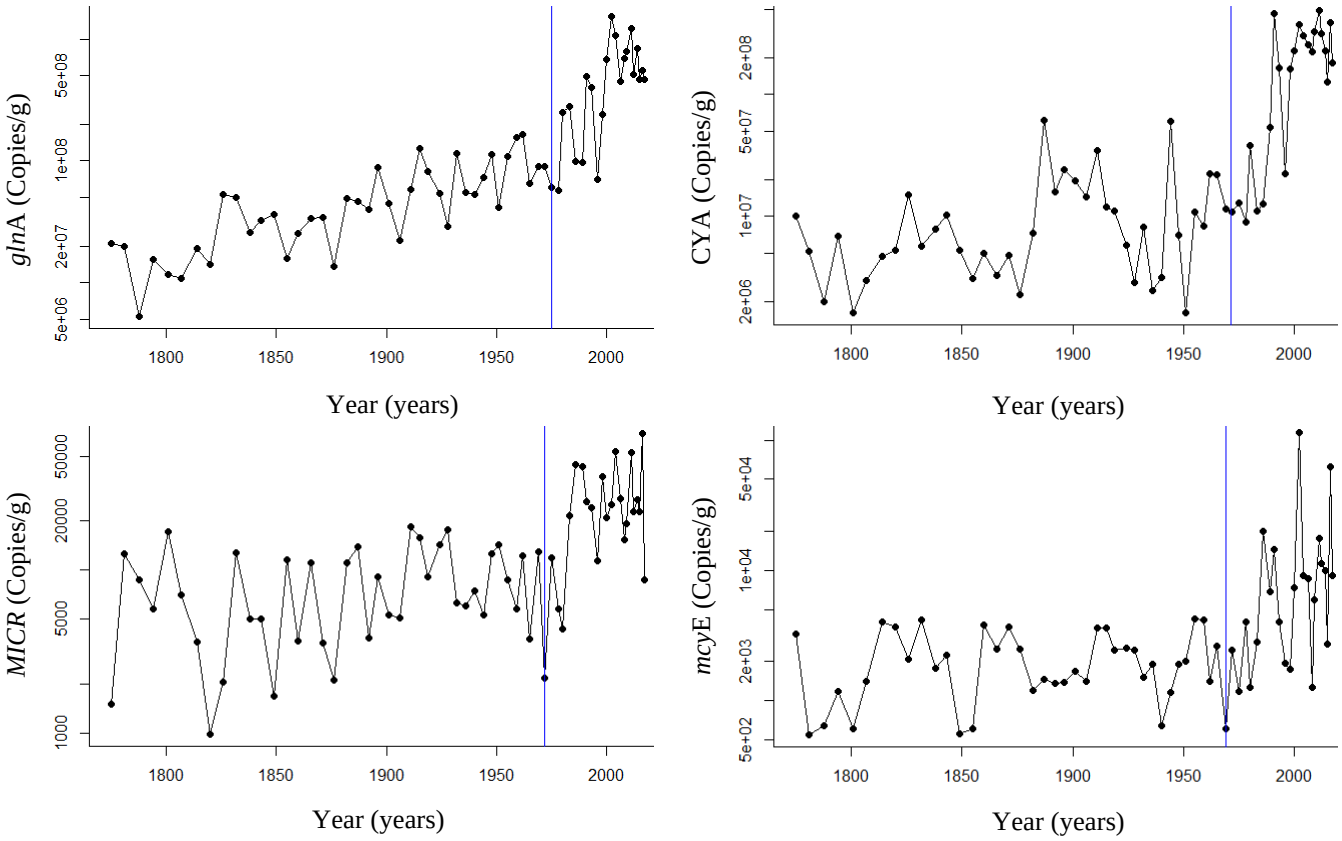


Figure 18a: Breakpoint analysis examining gene copy number of each gene target through time in Big Rideau lake. Breakpoint estimates the year in which gene copy trend begins to differ significantly from established trends. **TOP LEFT:** *gln*A gene copy increases through time with an initial breakpoint identified at 1975. **TOP RIGHT:** *CYA* gene copy increases through time with an initial breakpoint at the year 1968. **BOTTOM LEFT:** *MICR* gene copy number increasing through time with an initial breakpoint identified in 1969, very similar to the breakpoint identified for *gln*A. **BOTTOM RIGHT:** *McyE* gene copy number remained basically unchanged at detection until a breakpoint was identified in the year 1977.

Appendix: Figure 19a

Breakpoint Analysis: Gene Targets in Otty Lake

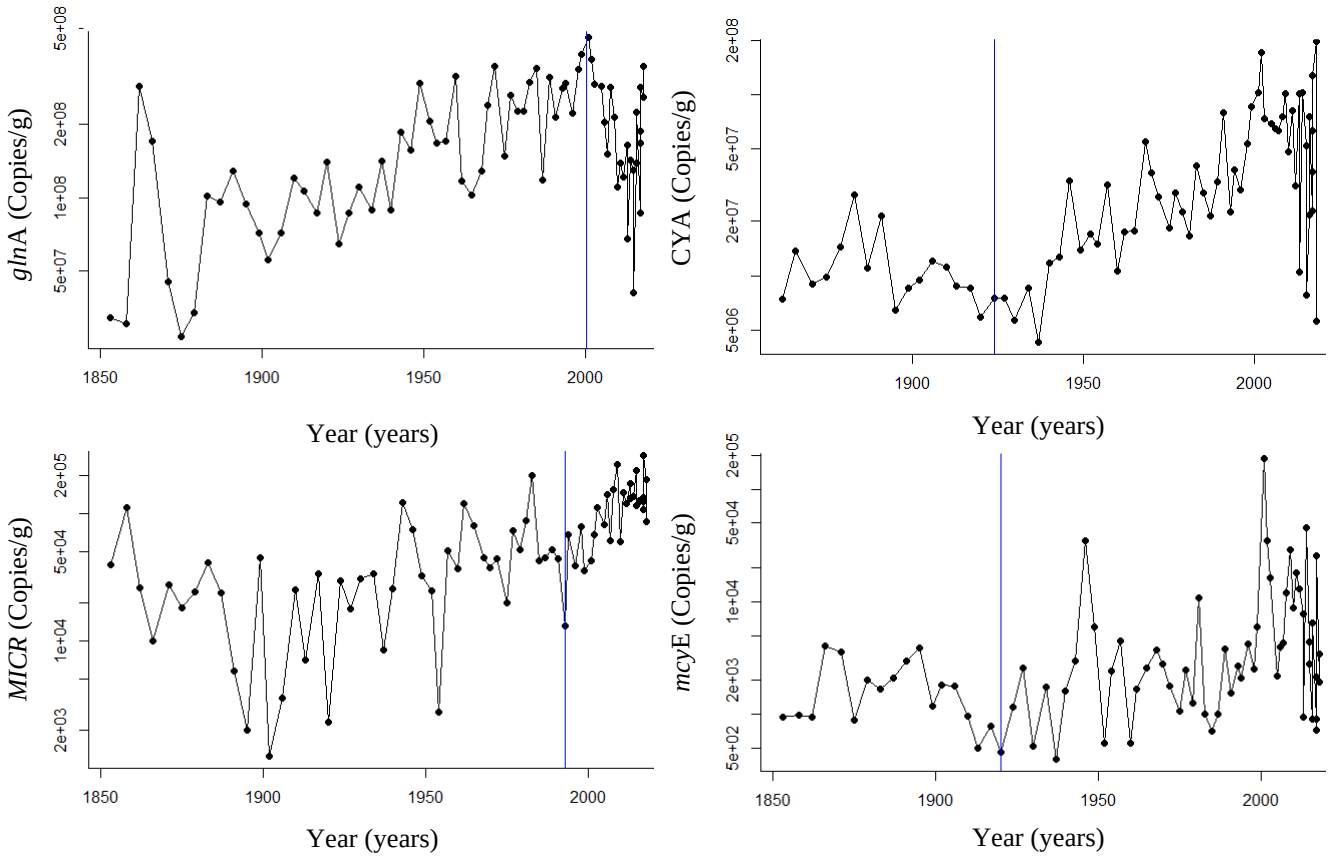


Figure 19a: Breakpoint analysis examining gene copy number of each gene target through time in Otty lake. Breakpoint estimates the year in which gene copy trend begins to differ significantly from established trends. **TOP LEFT:** *gln*A gene copy increases through time with an initial breakpoint identified at 2001. **TOP RIGHT:** *CYA* gene copy increases through time with an initial breakpoint at the year 1924. **BOTTOM LEFT:** *MICR* gene copy number increasing through time with an initial breakpoint identified in 1993. **BOTTOM RIGHT:** *McyE* gene copy number break was identified in 1917 earlier than the *MICR* breakpoint.

Appendix: Figure 20a

Three Mile Lake

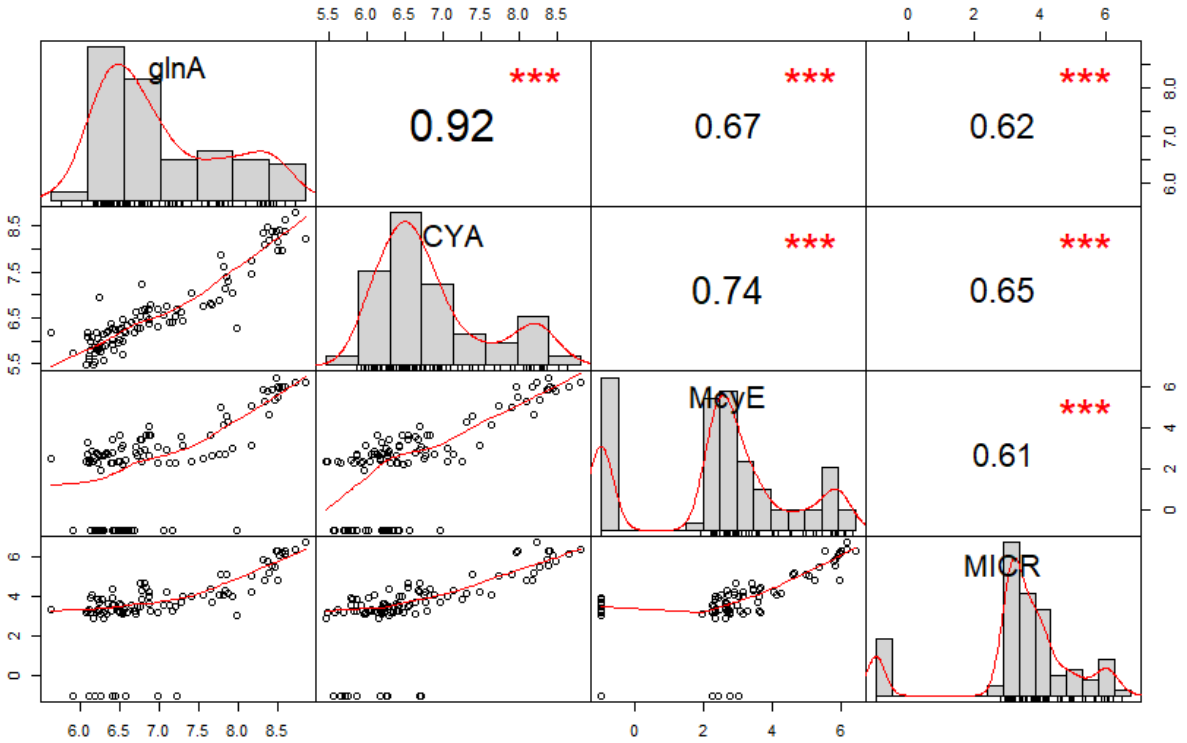


Figure 20a: Gene target correlations in Three Mile Lake using Kendall Rank correlation matrix. All gene targets were found to be highly correlated to one another.

Appendix: Figure 21a

Blue Chalk Lake

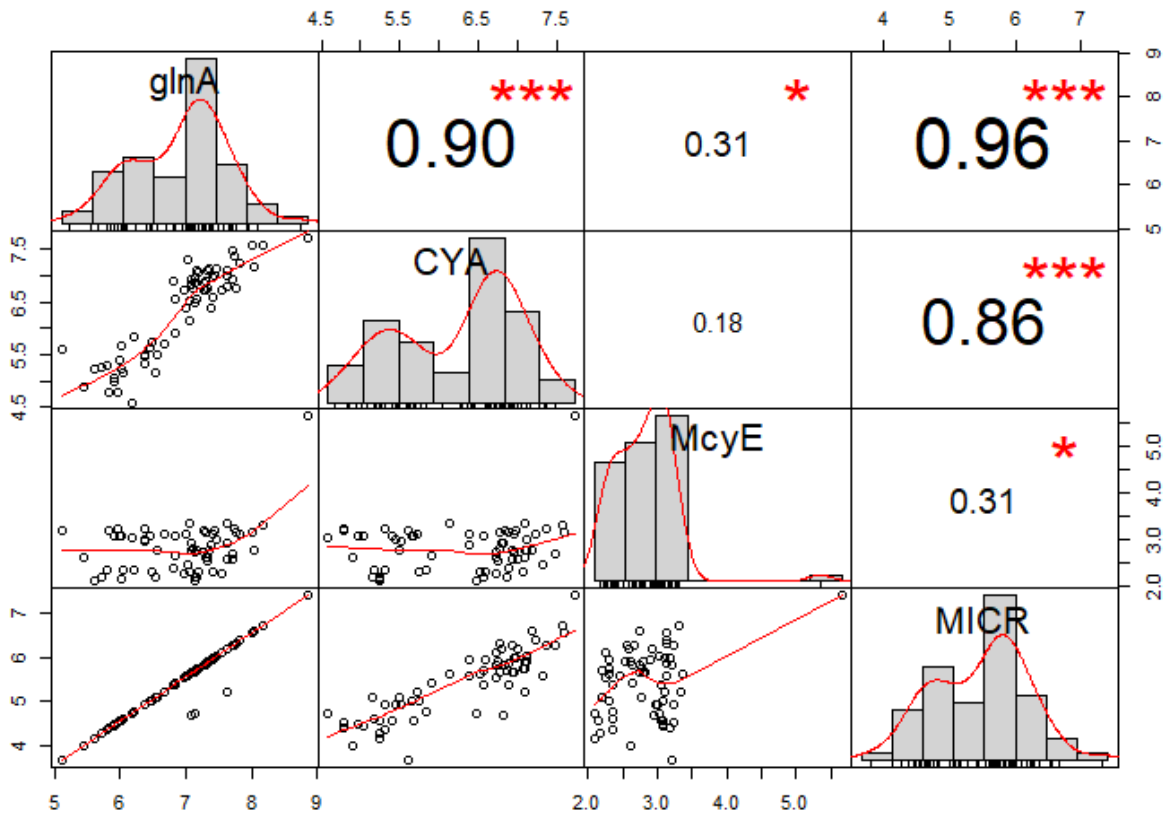


Figure 21a: Gene target correlations in Blue Chalk Lake using Kendall Rank correlation matrix. All gene targets were significantly correlated with the exception of *McyE* and CYA target (0.18).

Appendix: Figure 22a

Big Rideau Lake

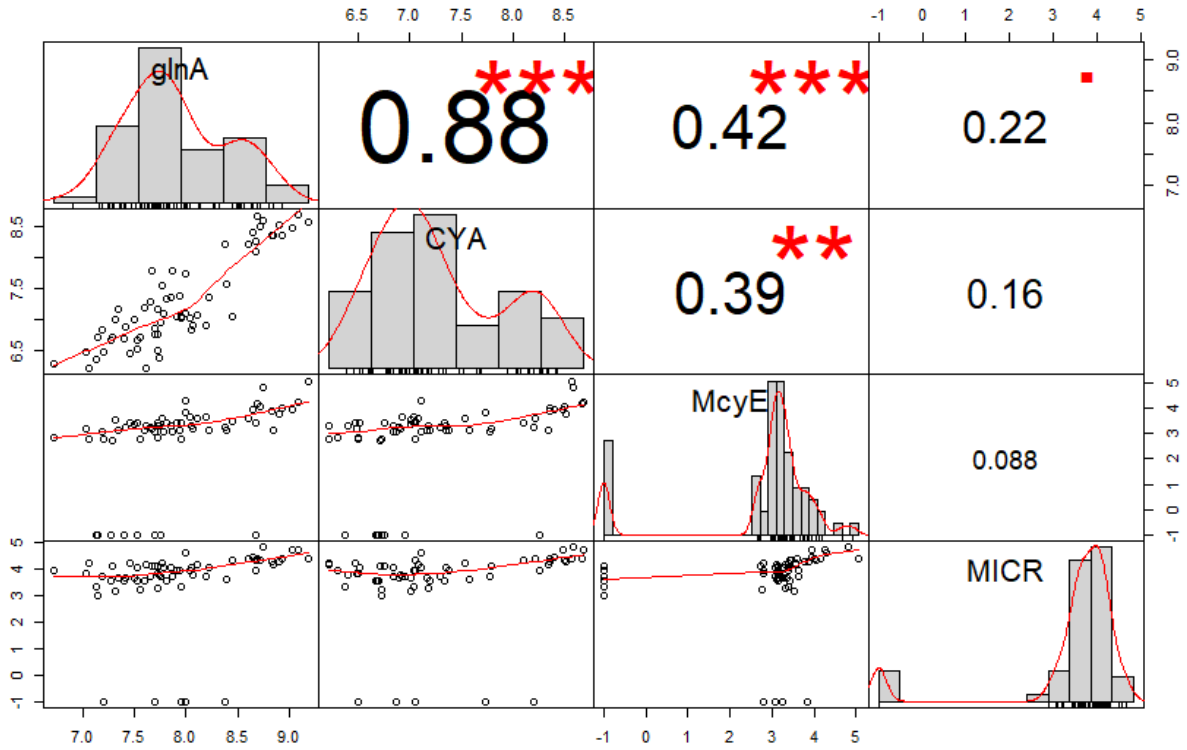


Figure 22a: Gene target correlations in Big Rideau Lake using Kendall Rank correlation matrix. All gene targets were significantly correlated with the exception of *MICR* which was found to not significantly correlate with any other gene targets in Big Rideau lake.

Appendix: Figure 23a

Otty Lake

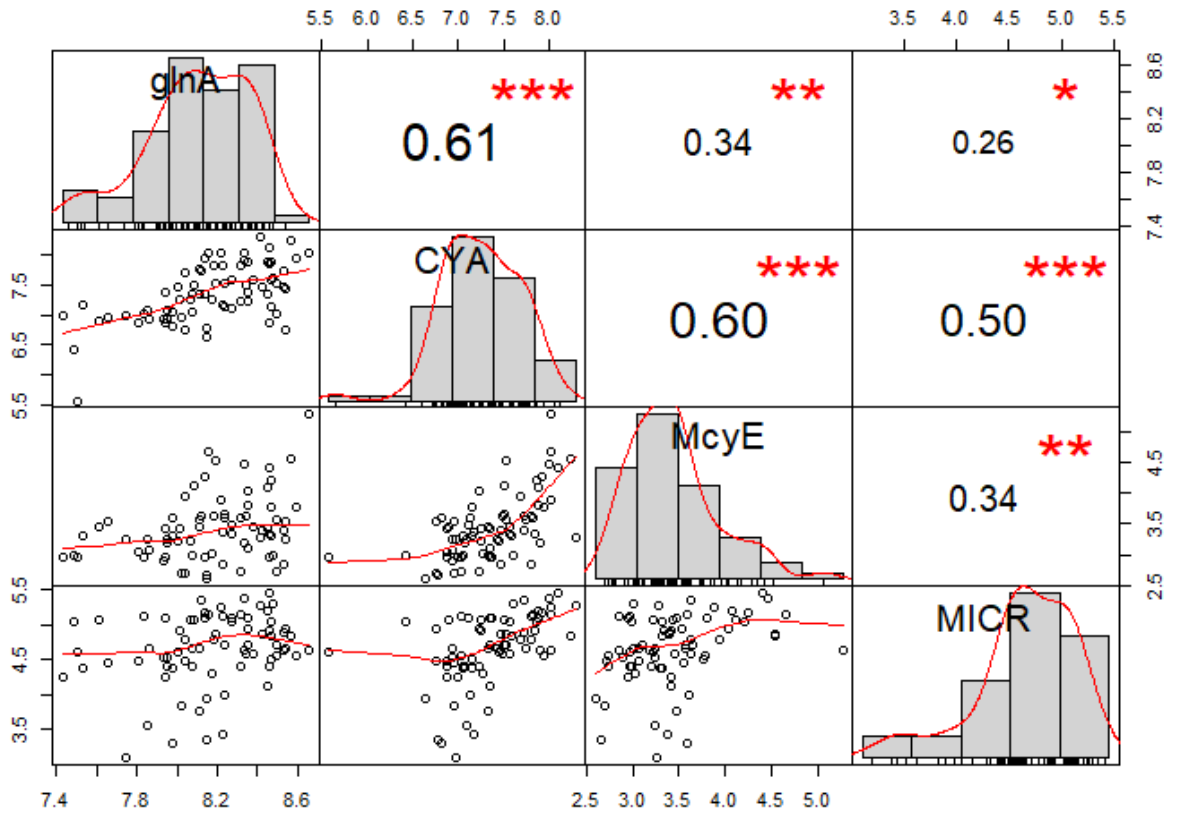


Figure 23a: Gene target correlations in Otty Lake using Kendall Rank correlation matrix. All gene targets were significantly correlated with one another.

Appendix: Figure 24a

Three Mile Lake, glnA - Partitioning of Variation, Time and Climate Factors

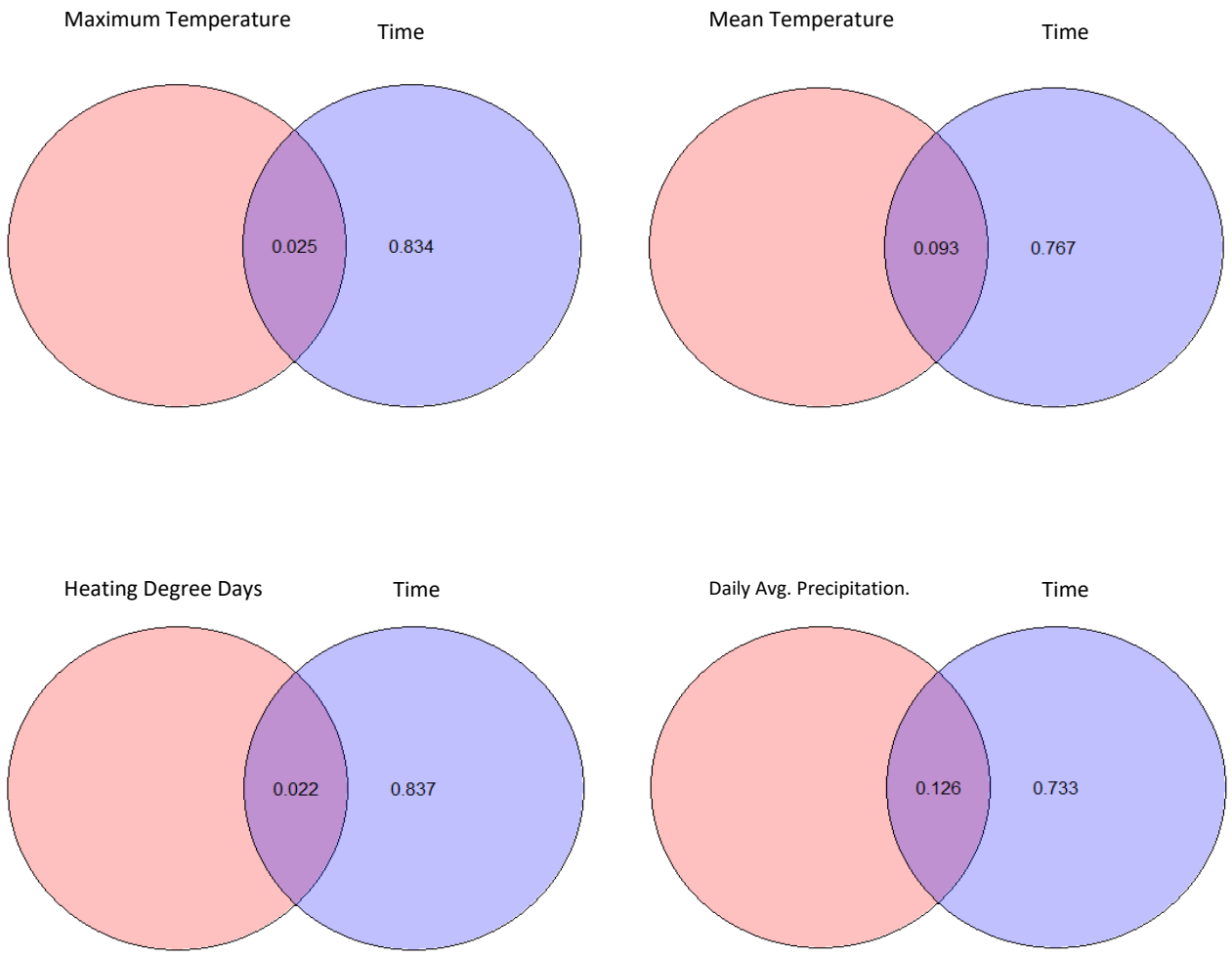


Figure 24a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on *gln*A in Three Mile Lake. **Top Left:** Time (blue) explained the majority of variation (83%) with maximum temperature explaining 2.5% of variation with time (purple). **Top Right:** Similar to maximum temperature time explained the majority of variation with mean temperature and time explaining ~9% (purple) **Bottom left**: Heating degree days and time explained ~2% of variation with time explaining the majority of variation. **Bottom Right:** Daily average precipitation and time explained 12.6% of variation of *gln*A in combination with time which still explains the majority of the variation in the *gln*A target.

Appendix: Figure 25a

Three Mile Lake, CYA - *Partitioning of Variation, Time and Climate Factors*

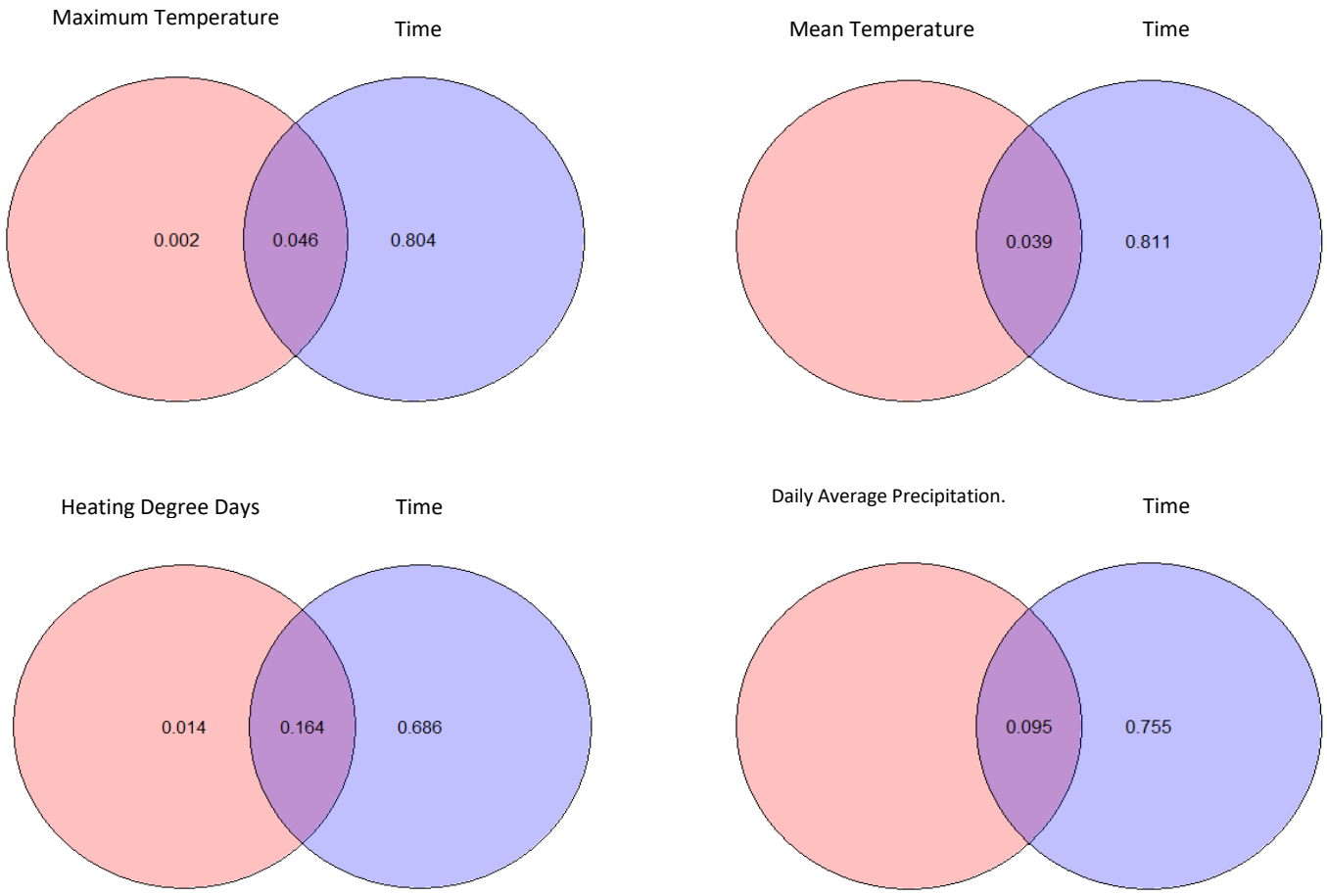


Figure 25a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on *CYA* in Three Mile Lake. **Top Left:** Time (blue) explained the majority of variation (80%) with maximum temperature explaining 4.6% of variation with time (purple). **Top Right:** Similar to maximum temperature time explained the majority of variation with mean temperature and time explaining ~3.9% (purple) **Bottom left**: Heating degree days was the most significant climate variable found in Three Milla lake and in combination with time (purple) it explained ~16.4% of variation in *CYA.* **Bottom Right:** Daily average precipitation and time explained 9.5% of variation of *gln*A in combination with time which still explains the majority of the variation in the *gln*A target.

Appendix: Figure 26a

Three Mile Lake, MICR - Partitioning of Variation, Time and Climate Factors

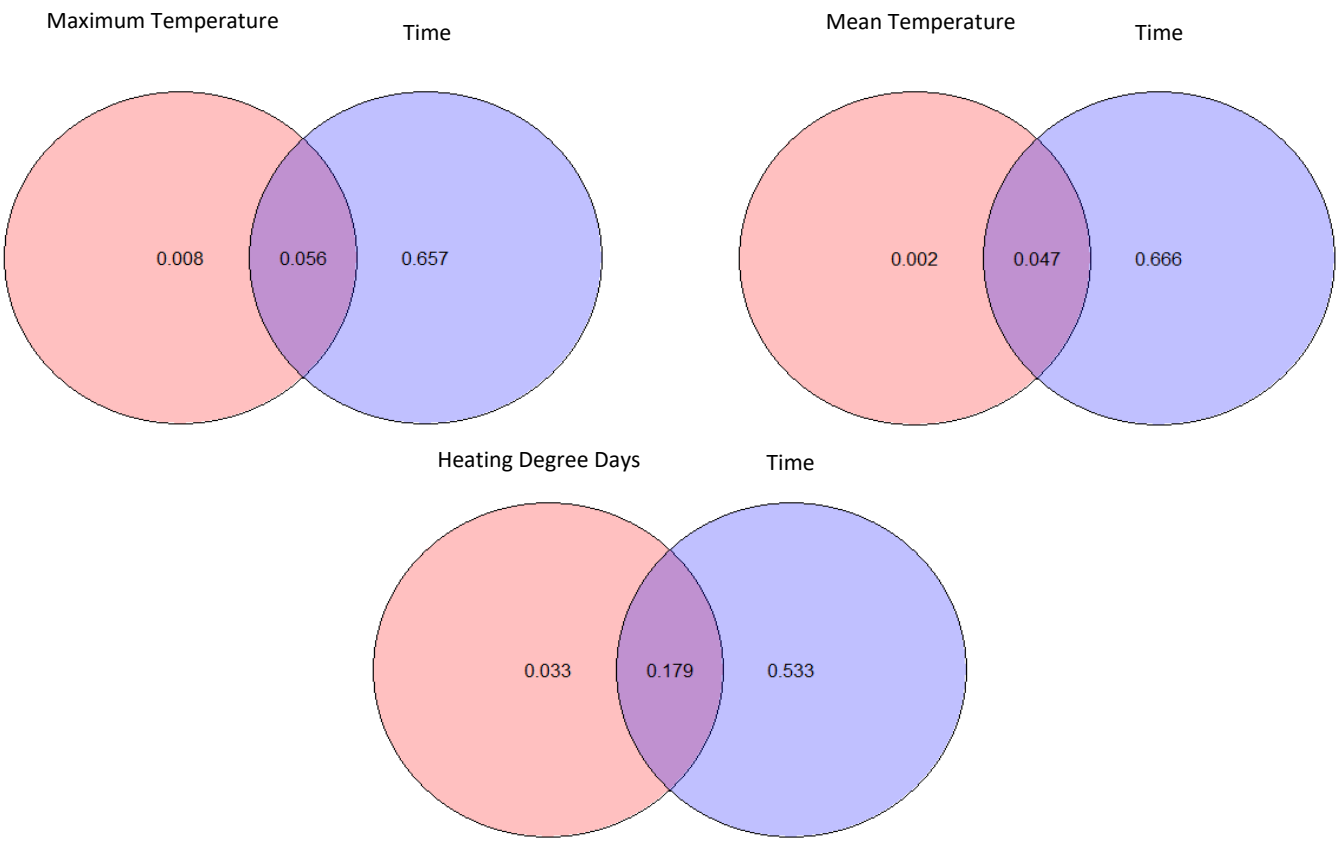


Figure 26a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on MICR in Three Mile Lake. **Top Left:** Time (blue) explained the majority of variation (80%) with maximum temperature explaining 5.6% of variation with time (purple). **Top Right:** Mean temperature was found to have a non-significant relationship to MICR and explained ~4.7% of variation with time (purple) **Bottom left**: Heating degree days was found to not explain a significant variation of MICR variation in Three Mile lake potentially explaining 17.9% of variation with time (purple).

Appendix: Figure 27a

Three Mile Lake, McyE - Partitioning of Variation, Time and Climate Factors

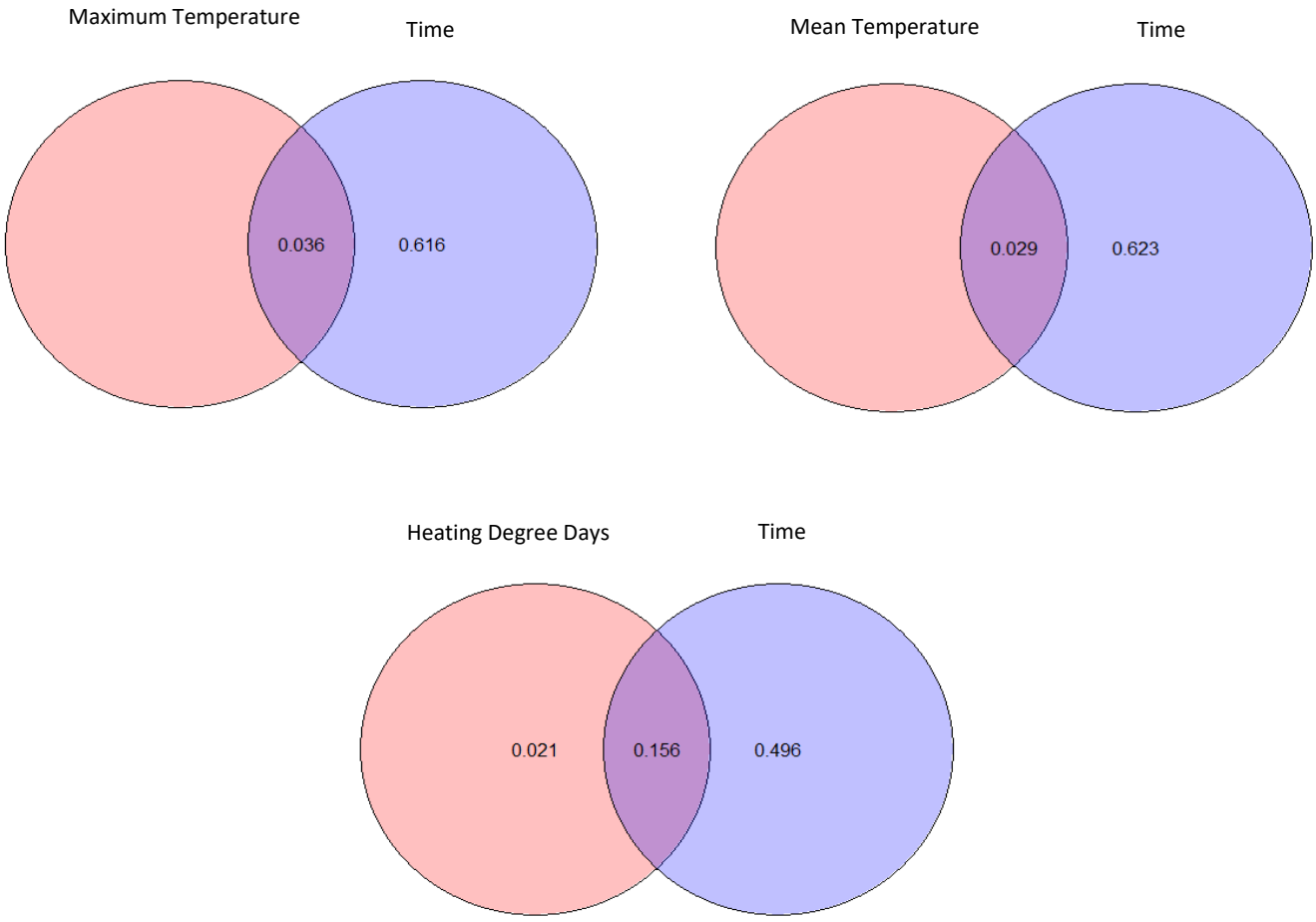


Figure 27a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on mcyE in Three Mile Lake. **Top Left:** Time (blue) explained the majority of variation (60%) with maximum temperature explaining 3.6% of variation with time (purple). **Top Right:** Mean temperature was found to have a non-significant relationship to McyE and explained ~2.9% of variation with time (purple) **Bottom**: Heating degree days was found to not explain a significant variation of McyE variation in Three Mile lake potentially explaining 15.6% of variation with time (purple).

Appendix: Figure 28a

Big Rideau, glnA Partitioning of Variation, Time and Climate Factors

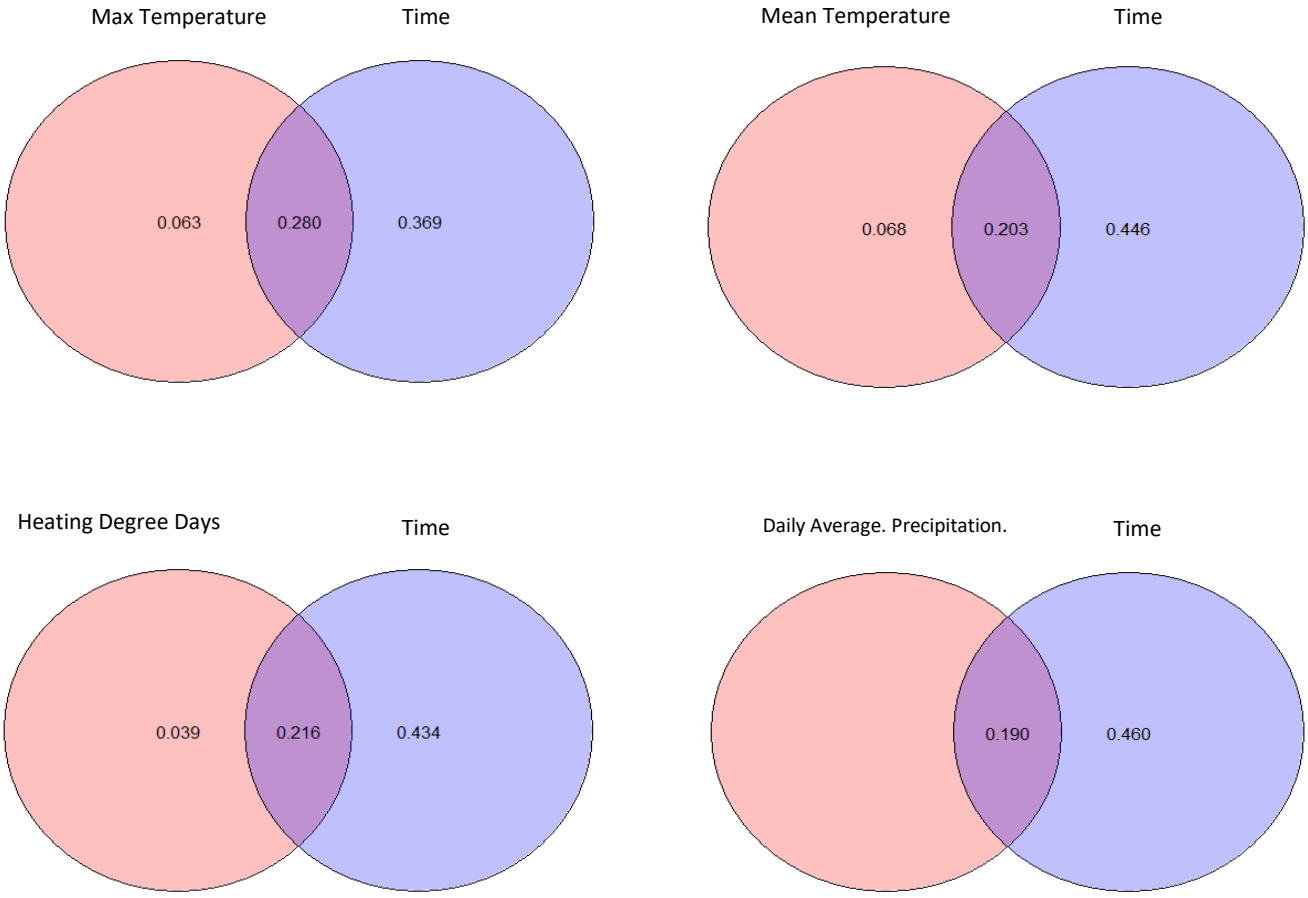


Figure 28a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on glnA in Big Rideau Lake. **Top Left:** Maximum temperature was highly significant in explaining variation of glnA, explaining as much as 28% of variation with time (purple). **Top Right:** Mean temperature is highly correlated to maximum temperature and was found to explain ~20% of variation with time (purple) **Bottom left**: Heating degree days was found to explain a significant variation of glnA variation in Three Mile lake potentially explaining 21.6% of variation with time (purple). **Bottom Right:** Daily average precipitation with time was found to explain 19% of variation with time.

Appendix: Figure 29a

Big Rideau, CYA *Partitioning of Variation, Time and Climate Factors*

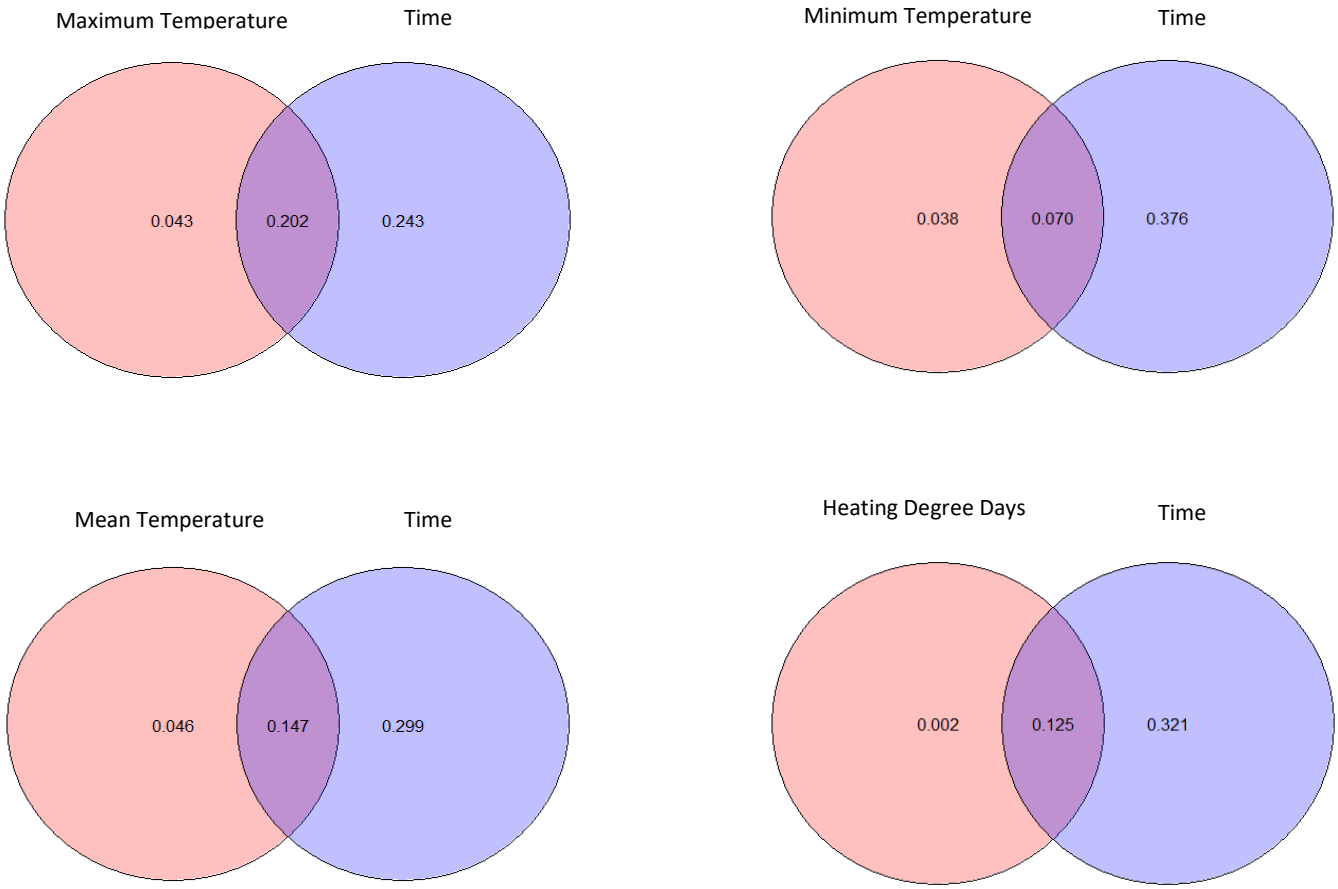


Figure 29a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on CYA in Big Rideau Lake. **Top Left:** Maximum temperature was highly significant in explaining variation of CYA, explaining as much as 20% of variation with time (purple). **Top Right:** Minimum temperature was found to explain only ~7% of variation with time (purple) **Bottom left**: Mean temperature was found to explain a significant variation of CYA variation in Big Rideau potentially explaining 14.7% of variation with time (purple). **Bottom Right:** Heating degree days with time was found to explain 12.5% of variation with time.

Appendix: Figure 30a

Big Rideau, McyE Partitioning of Variation, Time and Significant Climate Factors

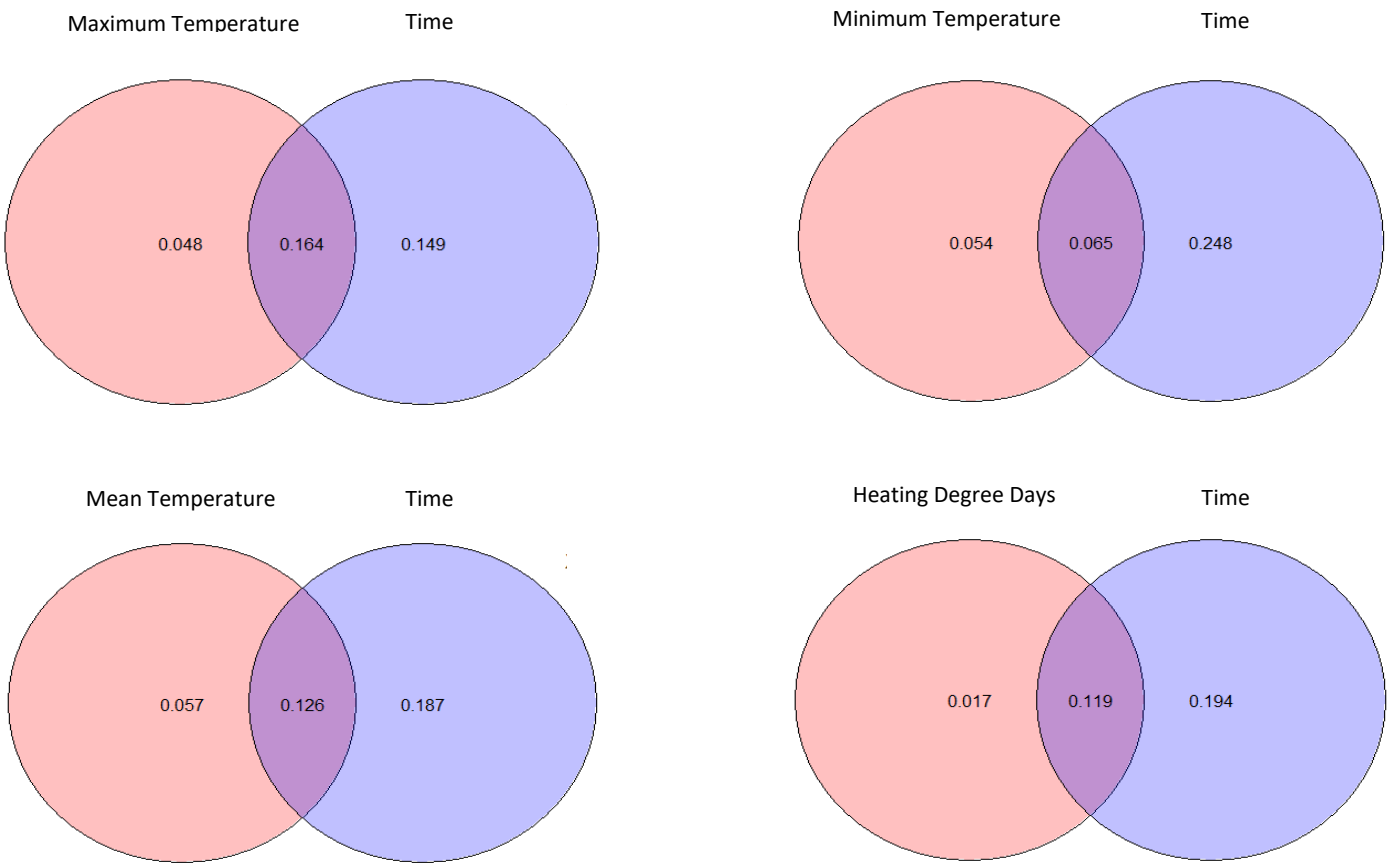


Figure 30a: Example of partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on McyE in Big Rideau Lake. **Top Left:** Maximum temperature was highly significant in explaining as much as 16% of variation with time (purple). **Top Right:** Minimum temperature was found to explain only ~6.5% of variation with time (purple) **Bottom left**: Mean temperature potentially explained 12.6.7% of variation with time (purple). **Bottom Right:** Heating degree days with time was found to explain 11.9% of variation with time.

Appendix: Figure 31a

Otty Lake, glnA Partitioning of Variation, Time and Climate Factors

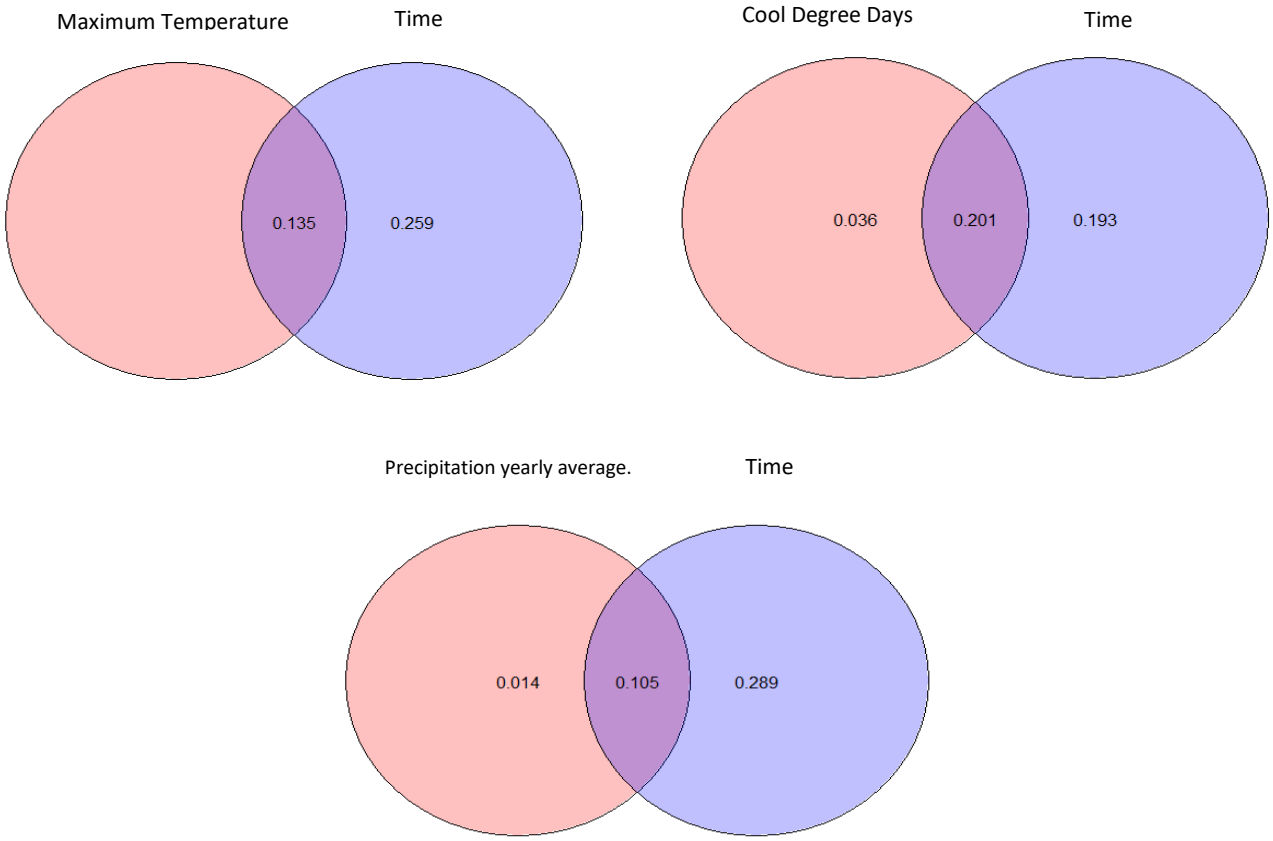


Figure 31a: Partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on glnA in Otty Lake **Top Left:** Maximum temperature and time explained 13.5% of variation with time (purple). **Top Right:** Cooling degree days was found to have a non-significant relationship to glnA and explained ~20% of variation with time (purple) **Bottom**: Yearly average precipitation was found to explain ~10.6% of variation with time (purple).

Appendix: Figure 32a

Otty Lake, CYA *Partitioning of Variation, Time and Climate Factors*

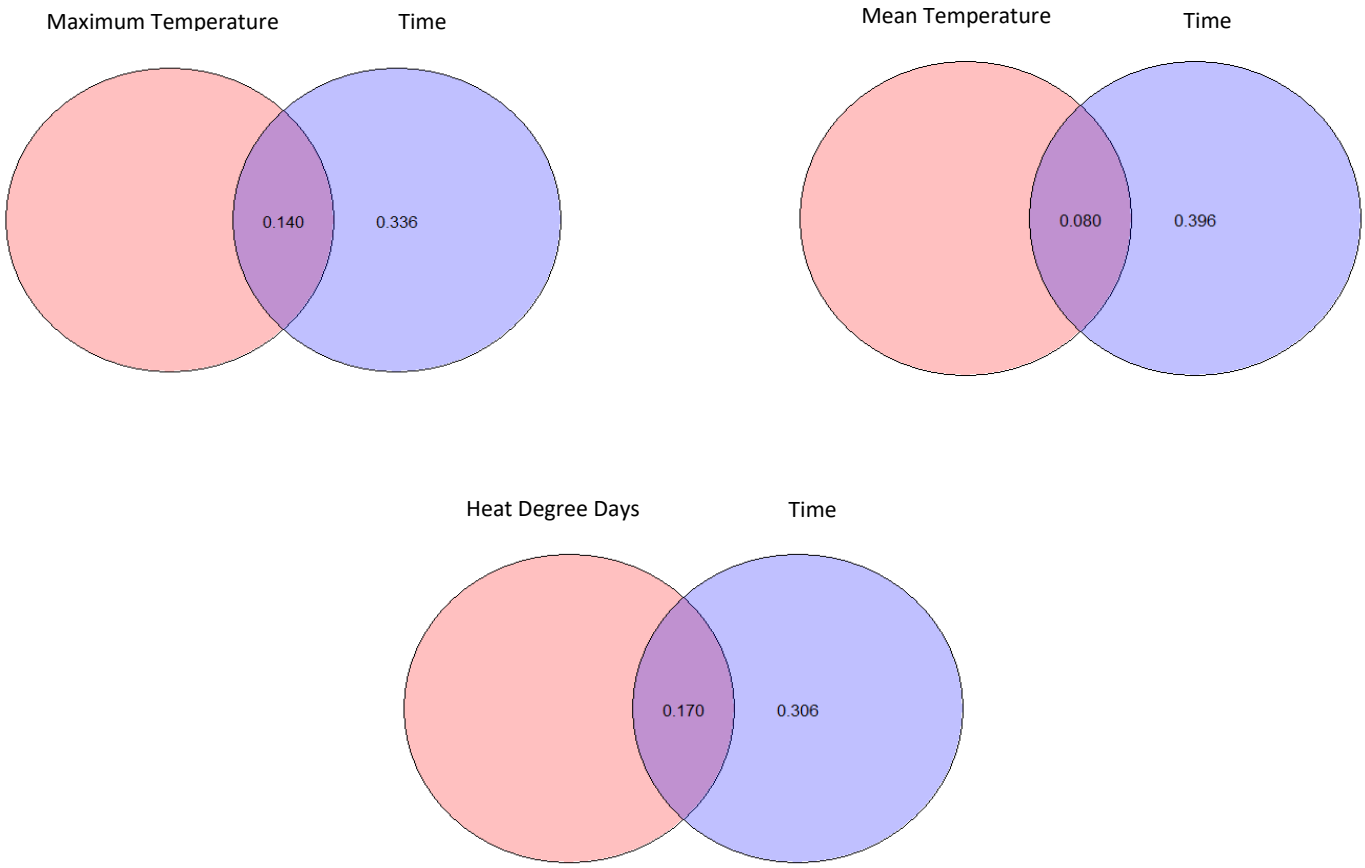


Figure 32a: Partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on CYA in Otty Lake **Top Left:** Maximum temperature and time explained 14% of variation with time (purple). **Top Right:** Mean temperature and time were found to explain ~6% of variation (purple) **Bottom**: Heating degree days was found to explain ~17% of variation with time (purple).

Appendix: Figure 33a

Otty Lake, MICR Partitioning of Variation, Time and Climate Factors

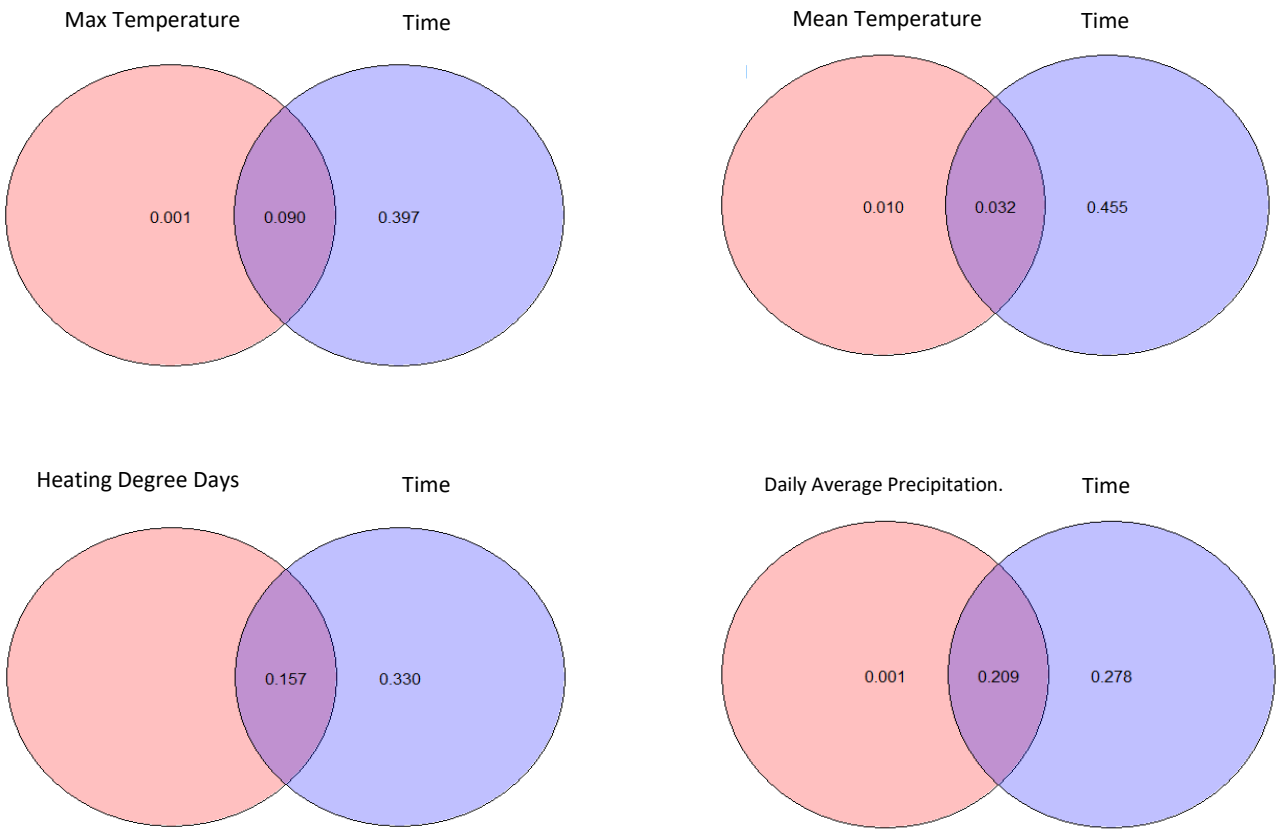


Figure 33a Partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on MICR in Otty Lake. **Top Left:** Maximum temperature was highly significant in explaining variation of CYA, explaining as much as 20% of variation with time (purple). **Top Right:** Minimum temperature was found to explain only ~7% of variation with time (purple) **Bottom left**: Mean temperature was found to explain a significant variation of CYA variation in Big Rideau potentially explaining 14.7% of variation with time (purple). **Bottom Right:** Heating degree days with time was found to explain 12.5% of variation with time.

Appendix: Figure 34a

Otty Lake, McyE Partitioning of Variation, Time and Climate Factors

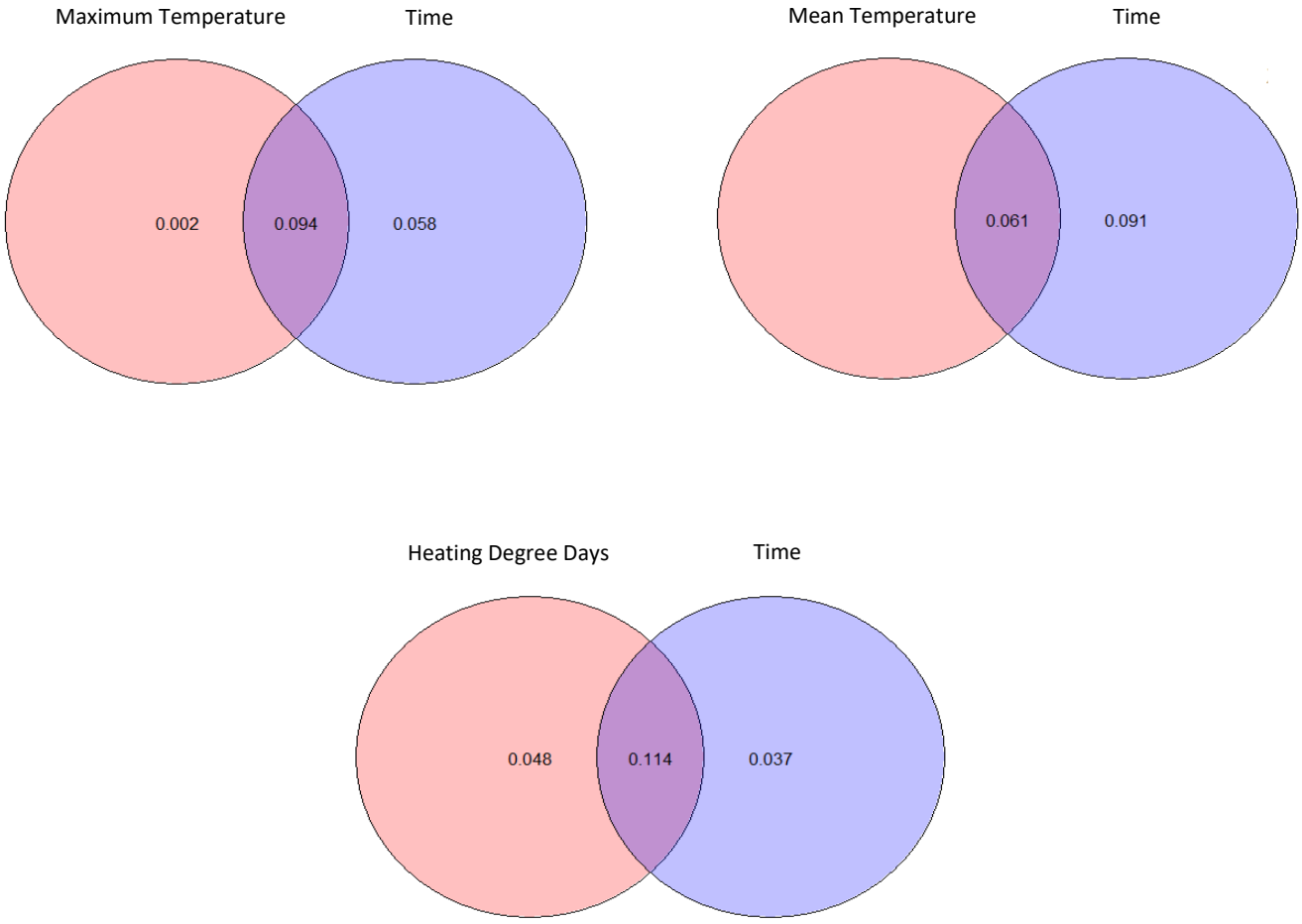


Figure 34a Partitioning of variation outputs examining climate variables (red) and time (blue) and the combination of the two factors (purple) on McyE in Otty Lake **Top Left:** Maximum temperature and time explained 9.4% of variation with time (purple). **Top Right:** Mean temperature and time were found to explain ~6% of variation (purple) **Bottom**: Heating degree days was found to explain ~11% of variation with time (purple).

Appendix B (Chapter 3)

Appendix: Table 1b

Table 1b: Endpoint PCR protocol for amplifying *apcA* using primers F1-R7 & F1-R10

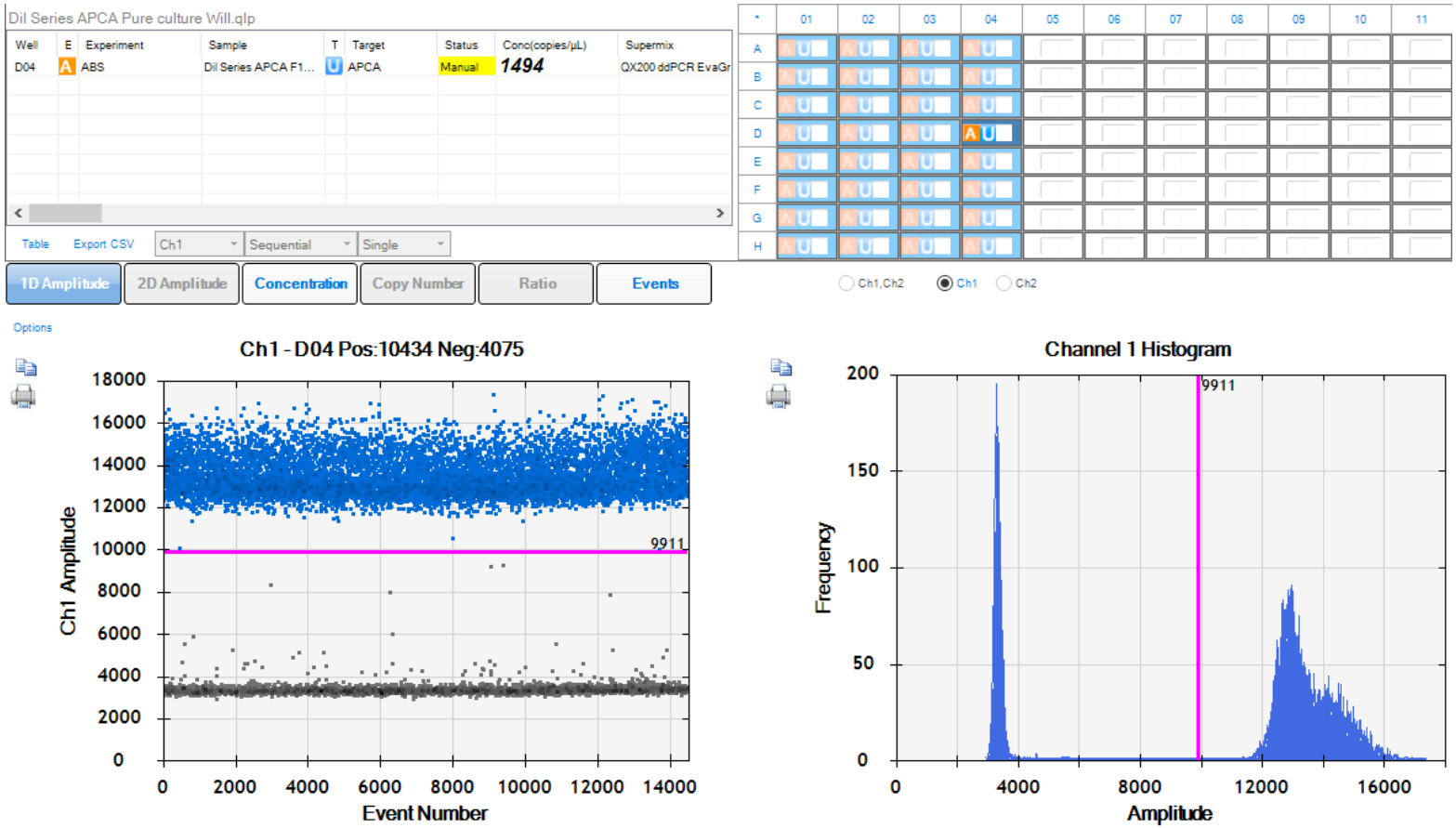
***apc*A**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 25uL Samples	1	94	2:00
	2	94	0:20
	3	60	1:00
	4	72	0:30
	5	GOTO Step 2, 34 times	
	6	72	10:00
	7	10	HOLD

Appendix: Table 2b

Table 2b: ddPCR protocol for amplifying *apcA* using primers F1-R7 & F1-R10

***apc*A (Allophycocyanin-Alpha Chain)**	**Step**	**Temperature (C°)**	**Time**
Lid 105°C, 40uL Samples	1	95	5:00
		Set Ramp 2°C/s	
	2	95	0:30
		Set Ramp 2°C/s	
	3	61.5	1:00
		Set Ramp 2°C/s	
	4	72	0:30
		Set Ramp 2°C/s	
	5	GOTO Step 2, 49 times	
	6	4	5:00
	7	90	5:00
	7	12	HOLD

Appendix: Figure 1b



Appendix 1b: Example ddPCR output for *apcA* primer set F1-R7 utilizing CPCC663 Synechoccocus pure culture extract

Appendix: Figure 2b

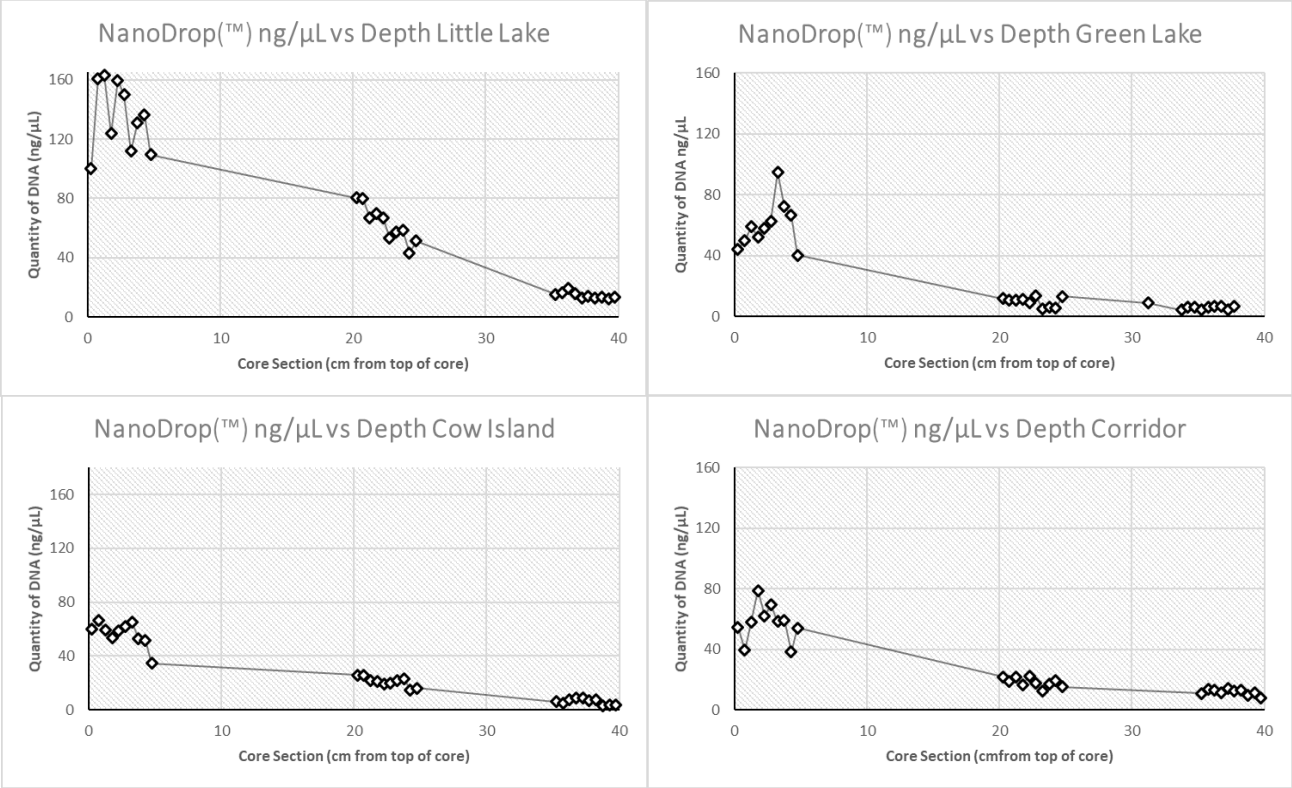


Figure 2b: Raw extracted DNA (In ng/uL) from sediment core sections against depth from four cores gathered from four different regions (Little Lake, Green Lake, Cow Island, Corridor) in Big Rideau Lake and measured through NanoDrop™ spectrophotometry. **Top Right**: Little Lake region is the most sheltered region cored in this study. It yielded the highest raw ng/uL of DNA material at both the top (average ~130 ng/uL) and middle portions (average ~70 ng/uL) of the cores gathered and yielded similar extraction value to the deepest (Corridor) core in the bottom sections (average ~14ng/uL). **Top Right:** Green Lake region was the shallowest region cored in this study (depth ~2.7m). It yielded similar DNA quantities in both the top and middle portion to all other sites except Little Lake and had similar yields in the bottom portion to the unsheltered Cow Island region. **Bottom Left:** Cow island region was of similar depth to the Little Lake region (~11m), however it yielded similar DNA quantities at both the top and middle portion of the core to all sections except the Little Lake region and had yields most similar to the Green Lake (shallowest site) region in the bottom portion. **Bottom Right:** The Corridor region was the deepest region cored in this study (~21.3m). The core yielded similar quantities of DNA to other cores taken at this site except for Little Lake in the top and middle portions of the core. However, it yielded similar DNA quantities to the Little Lake region and more DNA than from the bottom portion of the Green Lake or Cow Island bottom portions.

Appendix: Figure 3b

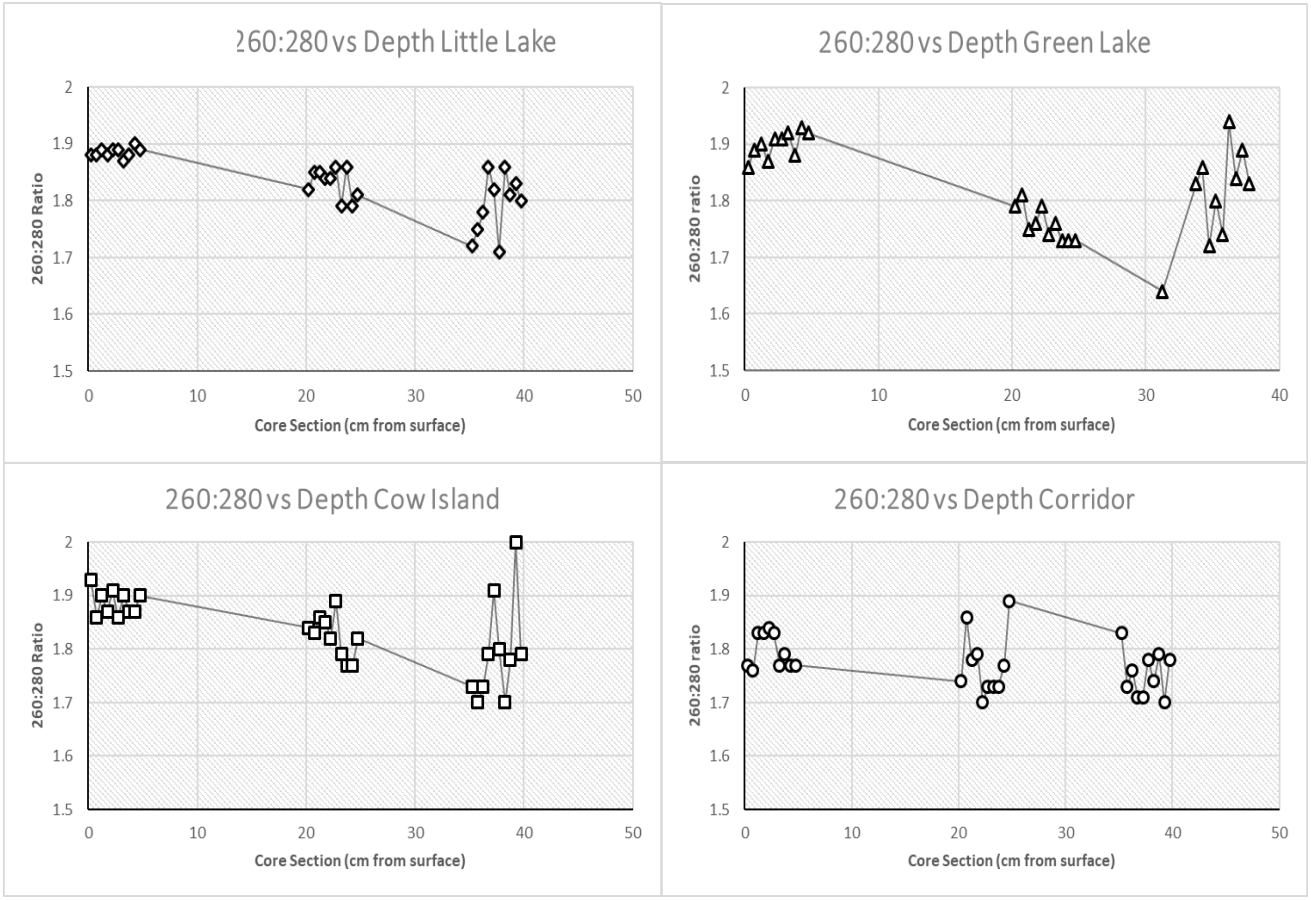


Figure 3b: 260:280 nm ratios of 30 (n=30) sediment section extractions from four different lake locations within Big Rideau lake representing a variety of depths and bathymetric conditions. **Top Left:** Little Lake region of Big Rideau lake. Little Lake was of moderate depth (~24m) but, represented the most sheltered of all locations cored. Ratio's became more variable with depth likely due to degradation and loss of nucleic acid materials over time. **Top Right:** Green Lake region of Big Rideau Lake. The shallowest location cored it displayed a more dramatic increase in variability in 260:280 nm ratio's likely due to higher detrivore activity within its sediments leading to increased degradation of nucleic acids. **Bottom Left**: Cow Island region of Big rideau lake. Cow Island was a less sheltered area of similar depth (~27m) to the Little Lake region. It shows similar variability in 260:280nm ratio variability to Little Lake until the deepest portion of the core where variability seemed increased. **Bottom Right**: The Corridor area of Big Rideau Lake. The deepest core taken (~70m) 260:280nm ratios remained relatively constant throughout the core. This is likely due to conditions present at extreme depths that have been long thought to slow the degradation of RNA and DNA materials (anoxic, low temperature, low light exposure etc.) and may represent those forces preserving DNA stably for longer periods of time.

Appendix: Figure 4b

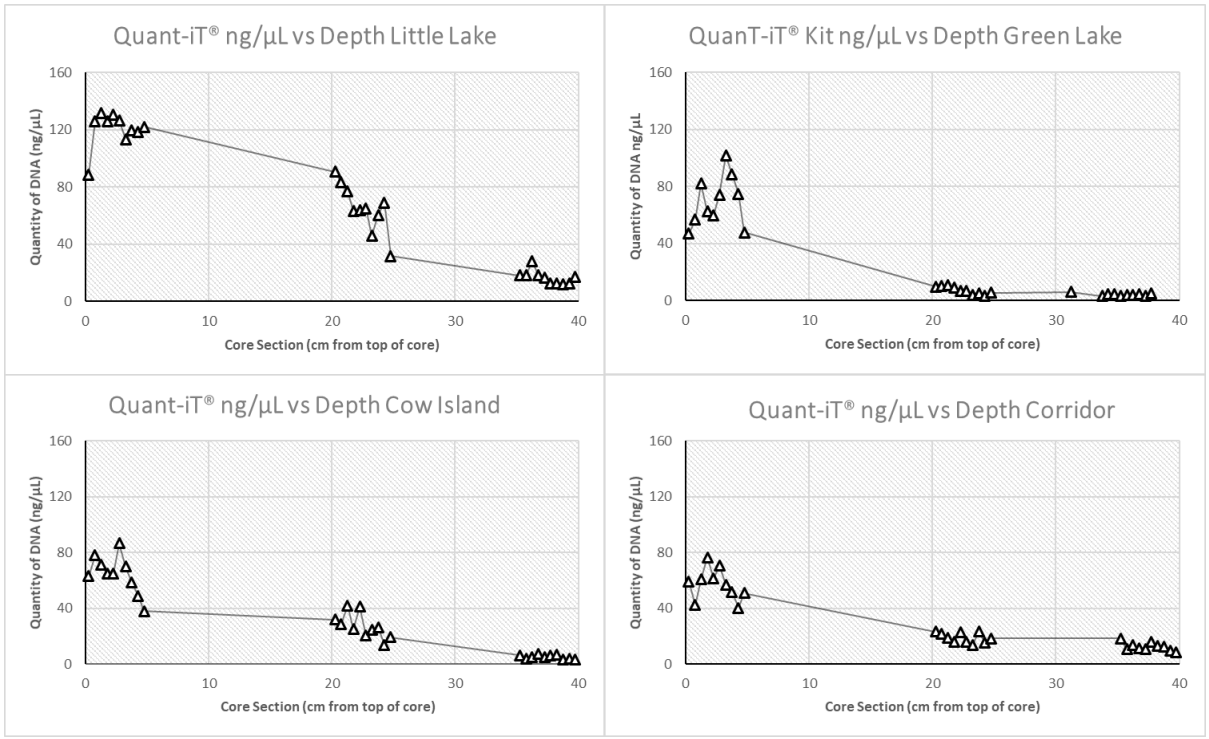


Figure 4b: Raw extracted DNA (In ng/uL) from sediment core sections against depth from four cores gathered from four different regions (Little Lake, Green Lake, Cow Island, Corridor) in Big Rideau Lake and measured through Quant-iT® fluorescent assay kit. **Top Right**: Little Lake region is the most sheltered region cored in this study. It yielded the highest raw ng/uL of DNA material at both the top (a more conservative average of ~120 ng/uL) and middle portions (average ~70 ng/uL) of the cores gathered and yielded similar extraction value to the deepest (Corridor) core in the bottom sections (average ~14ng/uL). **Top Right:** Green Lake region was the shallowest region cored in this study (depth ~2.7m). It yielded similar DNA quantities in both the top and middle portion to all other sites except Little Lake and had similar yields in the bottom portion to the unsheltered Cow Island region. **Bottom Left:** Cow island region was of similar depth to the Little Lake region (~11m), however it yielded similar DNA quantities at both the top and middle portion of the core to all sections except the Little Lake region and had yields most similar to the Green Lake (shallowest site) region in the bottom portion. **Bottom Right:** The Corridor region was the deepest region cored in this study (~21.3m). The core yielded similar quantities of DNA to other cores taken at this site except for Little Lake in the top and middle portions of the core. However, it yielded similar DNA quantities to the Little Lake region and more DNA than from the bottom portion of the Green Lake or Cow Island bottom portions.

Appendix: Figure 5b

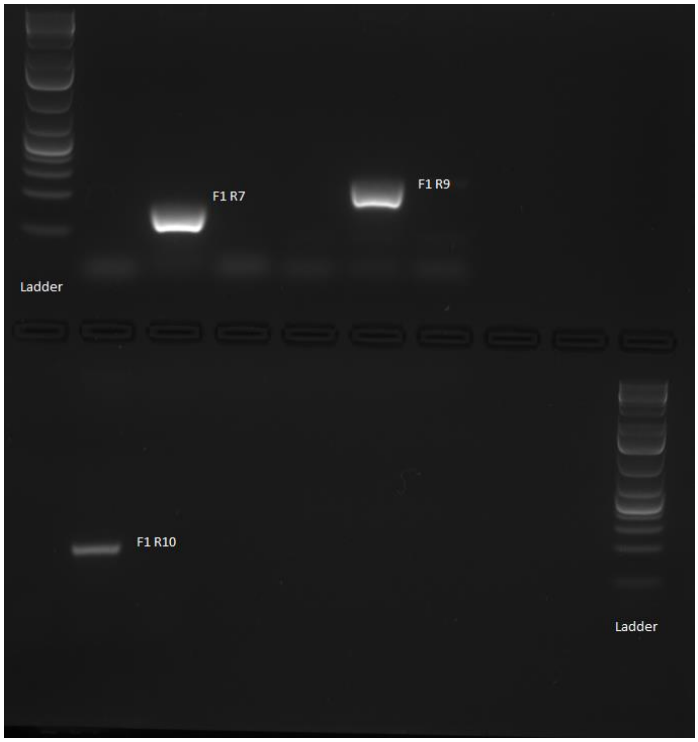


Figure 5b Endpoint PCR of 3 test primers for the morphometry study targeting the allophycocyanin alpha region of the allophycocyanin pigment gene. Single band amplification with no amplification in negative controls indicate functioning primers that are not self-annealing or forming homo/cross dimers. Gel was a 1.5% agarose impregnated with GelRed® nucleic acid dye. Gel was suspended in 1X TAE and run at 110V for ~40 minutes.

Supplemental Information

Organism List

ENA|A0A068N770|CP007542.1_Synechocystis_sp._PCC_6714,_complete_genome._:_Protein_AIE75927.1_CDS_(complement)

ENA|A0A078K1U6|LM993862.1_Phormidium_rubidum_A09DM_partial_apcA_gene_for_allophycocyanin,_strain_A09DM_:_Protein_CDY72720.1_CDS

ENA|A0A0H5PNJ2|LN847356.1_Synechococcus_sp._WH_8103_isolate_WH_8103_genome_assembly,_chromosome:_1_:_Protein_CRY91306.1_CDS_(complement)

ENA|A0A0X8WW16|AP017367.1_Leptolyngbya_sp._O-77_DNA,_complete_genome._:_Protein_BAU42719.1_CDS_(complement)

ENA|A0A182ARV3|LT578417.1_Cyanobium_sp._NIES-981_strain_NIES-981_genome_assembly,_chromosome:_I_:_Protein_SBO44547.1_CDS

ENA|A0A1J1JGW2|LO018304.1_Planktothrix_agardhii_str._7805_chromosome,_complete_genome._:_Protein_CUM60239.1_CDS

ENA|A0A1L6BXH8|KX758286.1_Synechococcus_sp._OH28_allophycocyanin_subunit_alpha_(apcA)_gene,_partial_cds._:_Protein_APQ41713.1_CDS

ENA|A0A1L6BXI5|KX758288.1_Synechococcus_sp._OH2_allophycocyanin_subunit_alpha_(apcA)_gene,_partial_cds._:_Protein_APQ41715.1_CDS

ENA|A0A1Q2U063|AP017959.1_Synechococcus_sp._NIES-970_DNA,_complete_genome._:_Protein_BAW97278.1_CDS_(complement)

ENA|A0A1Z3HPW6|CP021983.2_Halomicronema_hongdechloris_C2206_genome._:_Protein_ASC72351.1_CDS

ENA|A0A1Z4G798|AP018172.1_Calothrix_sp._NIES-2098_DNA,_complete_genome._:_Protein_BAY13390.1_CDS_(complement)

ENA|A0A1Z4GBI8|AP018174.1_Anabaenopsis_circularis_NIES-21_DNA,_nearly_complete_genome._:_Protein_BAY14874.1_CDS

ENA|A0A1Z4I9N8|AP018184.1_Nostoc_sp._NIES-2111_DNA,_nearly_complete_genome._:_Protein_BAY38762.1_CDS_(complement)

ENA|A0A1Z4IL84|AP018194.1_Scytonema_sp._HK-05_DNA,_nearly_complete_genome._:_Protein_BAY42804.1_CDS_(complement)

ENA|A0A1Z4J1M1|AP018194.1_Scytonema_sp._HK-05_DNA,_nearly_complete_genome._:_Protein_BAY47806.1_CDS

ENA|A0A1Z4KR01|AP018216.1_Anabaena_variabilis_NIES-23_DNA,_nearly_complete_genome._:_Protein_BAY71416.1_CDS_(complement)

ENA|A0A1Z4LMV7|AP018227.1_Calothrix_parasitica_NIES-267_DNA,_nearly_complete_genome._:_Protein_BAY82571.1_CDS_(complement)

ENA|A0A1Z4PR13|AP018268.1_Scytonema_sp._NIES-4073_DNA,_nearly_complete_genome._:_Protein_BAZ20430.1_CDS

ENA|A0A238HD08|LT852753.1_Nostoc_sp._R76DM_partial_apcA_gene_for_Allophycocyanin_alpha_subunit,_strain_R76DM_:_Protein_SMQ11640.1_CDS

ENA|A0A2H6L0U2|BFAC01000002.1_Microcystis_aeruginosa_NIES-87_DNA,_sequence002._:_Protein_GBE73647.1_CDS_(complement)

ENA|A0A2K8WSD8|CP024912.1_Cyanobacterium_stanieri_HL-69_chromosome,_complete_genome._:_Protein_AUC62014.1_CDS

ENA|A0A2L2P5Q1|MG014399.1_Synechococcus_sp._RS9907_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76490.1_CDS

ENA|A0A2L2P5Q3|MG014397.1_Synechococcus_sp._RS9915_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76488.1_CDS

ENA|A0A2L2P5Q4|MG014398.1_Synechococcus_sp._RS9909_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76489.1_CDS

ENA|A0A2L2P5Q5|MG014400.1_Synechococcus_sp._RS9902_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76491.1_CDS

ENA|A0A2L2P5Q7|MG014401.1_Synechococcus_sp._ROS8604_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76492.1_CDS

ENA|A0A2L2P5T9|MG014396.1_Synechococcus_sp._SYN20_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76487.1_CDS

ENA|A0A2L2P5U5|MG014403.1_Synechococcus_sp._PROS-9-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76494.1_CDS

ENA|A0A2L2P5U7|MG014395.1_Synechococcus_sp._TAK9802_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76486.1_CDS

ENA|A0A2L2P5U8|MG014394.1_Synechococcus_sp._WH_8101_isolate_WH8101_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76485.1_CDS

ENA|A0A2L2P5U9|MG014402.1_Synechococcus_sp._PROS-U-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76493.1_CDS

ENA|A0A2L2P5V0|MG014405.1_Synechococcus_sp._NOUM97013_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76496.1_CDS

ENA|A0A2L2P5V1|MG014412.1_Synechococcus_sp._BMK-MC-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76503.1_CDS

ENA|A0A2L2P5V2|MG014411.1_Synechococcus_sp._BOUM118_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76502.1_CDS

ENA|A0A2L2P5V3|MG014406.1_Synechococcus_sp._MVIR-18-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76497.1_CDS

ENA|A0A2L2P5V4|MG014413.1_Synechococcus_sp._BIOS-U3-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76504.1_CDS

ENA|A0A2L2P5V5|MG014404.1_Synechococcus_sp._PROS-7-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76495.1_CDS

ENA|A0A2L2P5V6|MG014407.1_Synechococcus_sp._MIT_S9220_isolate_MITS9220_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76498.1_CDS

ENA|A0A2L2P5V7|MG014409.1_Synechococcus_sp._MEDNS5_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76500.1_CDS

ENA|A0A2L2P5V8|MG014408.1_Synechococcus_sp._Minos11_isolate_MINOS11_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76499.1_CDS

ENA|A0A2L2P5V9|MG014410.1_Synechococcus_sp._M16.1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76501.1_CDS

ENA|A0A2L2P5W0|MG014420.1_Synechococcus_sp._A15-60_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76511.1_CDS

ENA|A0A2L2P5W1|MG014419.1_Synechococcus_sp._A15-62_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76510.1_CDS

ENA|A0A2L2P5W2|MG014421.1_Synechococcus_sp._A15-44_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76512.1_CDS

ENA|A0A2L2P5W3|MG014417.1_Synechococcus_sp._A18-25c_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76508.1_CDS

ENA|A0A2L2P5W4|MG014418.1_Synechococcus_sp._A15-127_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76509.1_CDS

ENA|A0A2L2P5W5|MG014414.1_Synechococcus_sp._BIOS-E4-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76505.1_CDS

ENA|A0A2L2P5W6|MG014423.1_Synechococcus_sp._A15-24_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76514.1_CDS

ENA|A0A2L2P5W7|MG014415.1_Synechococcus_sp._A18-46.1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76506.1_CDS

ENA|A0A2L2P5W8|MG014422.1_Synechococcus_sp._A15-28_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76513.1_CDS

ENA|A0A2L2P5W9|MG014416.1_Synechococcus_sp._A18-40_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76507.1_CDS

ENA|A0A2L2P5X2|MG014424.1_Cyanobium_sp._NS01_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76515.1_CDS

ENA|A0A2Z5X2N9|AP017560.1_Pseudanabaena_sp._ABRG5-3_DNA,_complete_genome._:_Protein_BBC25951.1_CDS_(complement)

ENA|A0A2Z6CCB6|AP017991.1_Planktothrix_agardhii_NIES-204_DNA,_complete_genome._:_Protein_BBD54171.1_CDS

ENA|A0A2Z6CX57|AP018318.1_Nostoc_sp._HK-01_DNA,_complete_genome._:_Protein_BBD60970.1_CDS_(complement)

ENA|A0A3G9JUA6|AP019314.1_Microcystis_viridis_NIES-102_DNA,_complete_genome._:_Protein_BBH41821.1_CDS_(complement)

ENA|A5GND7|CT971583.1_Synechococcus_WH7803_complete_genome_sequence._:_Protein_CAK24452.1_CDS

ENA|A5GV81|CT978603.1_Synechococcus_sp._RCC307_genomic_DNA_sequence._:_Protein_CAK28790.1_CDS

ENA|B0C0G1|CP000828.1_Acaryochloris_marina_MBIC11017,_complete_genome._:_Protein_ABW30754.1_CDS_(complement)

ENA|B0C947|CP000828.1_Acaryochloris_marina_MBIC11017,_complete_genome._:_Protein_ABW26582.1_CDS_(complement)

ENA|B0CFZ9|CP000828.1_Acaryochloris_marina_MBIC11017,_complete_genome._:_Protein_ABW29446.1_CDS

ENA|B0JRU9|AP009552.1_Microcystis_aeruginosa_NIES-843_DNA,_complete_genome._:_Protein_BAG00849.1_CDS_(complement)

ENA|B1WV59|CP000806.1_Cyanothece_sp._ATCC_51142_circular_chromosome,_complete_sequence._:_Protein_ACB52256.1_CDS

ENA|B1XQM2|CP000951.1_Synechococcus_sp._PCC_7002,_complete_genome._:_Protein_ACA99917.1_CDS_(complement)

ENA|B3VNK2|EU781500.1_Thermosynechococcus_vulcanus_str._Copeland_allophycocyanin_alpha_subunit_(apcA)_gene,_complete_cds._:_Protein_ACF18023.1_CDS_(complement)

ENA|D5A426|AP011615.1_Arthrospira_platensis_NIES-39_DNA,_nearly_complete_genome._:_Protein_BAI89320.1_CDS_(complement)

ENA|E9M3F3|HQ187783.1_Uncultured_Synechococcus_sp._clone_M60c096H4_allophycocyanin_subunit_B-like_(apcB)_gene,_partial_sequence;_and_allophycocyanin_subunit_A_(apcA)_gene,_partial_cds._:_Protein_ADU78748.1_CDS_(complement)

ENA|E9M3F5|HQ187853.1_Uncultured_Synechococcus_sp._clone_SynAculture_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78883.1_CDS_(complement)

ENA|E9M3F7|HQ187803.1_Uncultured_Synechococcus_sp._clone_M60c544G10_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78786.1_CDS_(complement)

ENA|E9M3G1|HQ187787.1_Uncultured_Synechococcus_sp._clone_M60c212H11_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78756.1_CDS_(complement)

ENA|E9M3H4|HQ187794.1_Uncultured_Synechococcus_sp._clone_M60c397F1_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78769.1_CDS_(complement)

ENA|E9M3H6|HQ187801.1_Uncultured_Synechococcus_sp._clone_M60c534H9_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78782.1_CDS_(complement)

ENA|E9M3H8|HQ187796.1_Uncultured_Synechococcus_sp._clone_M60c437F4_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78773.1_CDS_(complement)

ENA|E9M3I1|HQ187798.1_Uncultured_Synechococcus_sp._clone_M60c494H7_allophycocyanin_subunit_B-like_(apcB)_gene,_partial_sequence;_and_allophycocyanin_subunit_A_(apcA)_gene,_partial_cds._:_Protein_ADU78776.1_CDS_(complement)

ENA|E9M3J3|HQ187804.1_Uncultured_Synechococcus_sp._clone_M60c565B1_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78788.1_CDS_(complement)

ENA|E9M3J5|HQ187805.1_Uncultured_Synechococcus_sp._clone_M60c639B6_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78790.1_CDS_(complement)

ENA|E9M3L1|HQ187813.1_Uncultured_Synechococcus_sp._clone_M65c195A10_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78806.1_CDS_(complement)

ENA|E9M3M1|HQ187818.1_Uncultured_Synechococcus_sp._clone_M65c293C3_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78816.1_CDS_(complement)

ENA|E9M3M5|HQ187820.1_Uncultured_Synechococcus_sp._clone_M65c310B4_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78820.1_CDS_(complement)

ENA|E9M3M9|HQ187822.1_Uncultured_Synechococcus_sp._clone_M65c316F4_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78824.1_CDS_(complement)

ENA|E9M3P0|HQ187842.1_Uncultured_Synechococcus_sp._clone_M65c597G8_allophycocyanin_subunit_A_(apcA)_gene,_partial_cds._:_Protein_ADU78861.1_CDS_(complement)

ENA|E9M3Q8|HQ187837.1_Uncultured_Synechococcus_sp._clone_M65c548B6_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78853.1_CDS_(complement)

ENA|E9M3S6|HQ187847.1_Uncultured_Synechococcus_sp._clone_M65c647B11_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78871.1_CDS_(complement)

ENA|E9M3U0|HQ187997.1_Uncultured_Synechococcus_sp._clone_M60c032C2_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78884.1_CDS_(complement)

ENA|E9M3U2|HQ187998.1_Uncultured_Synechococcus_sp._clone_M60c051B3_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78887.1_CDS_(complement)

ENA|E9M3U4|HQ187999.1_Uncultured_Synechococcus_sp._clone_M60c055D3_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78889.1_CDS_(complement)

ENA|E9M3U6|HQ188054.1_Uncultured_Synechococcus_sp._clone_M60c667F8_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78994.1_CDS_(complement)

ENA|E9M3V2|HQ188003.1_Uncultured_Synechococcus_sp._clone_M60c144F7_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78897.1_CDS_(complement)

ENA|E9M3X8|HQ188016.1_Uncultured_Synechococcus_sp._clone_M60c626A5_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78923.1_CDS_(complement)

ENA|E9M3Z2|HQ188023.1_Uncultured_Synechococcus_sp._clone_M60c579D2_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78937.1_CDS_(complement)

ENA|E9M3Z5|HQ188025.1_Uncultured_Synechococcus_sp._clone_M60c658A8_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78940.1_CDS_(complement)

ENA|E9M401|HQ188028.1_Uncultured_Synechococcus_sp._clone_M60c146H7_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78946.1_CDS_(complement)

ENA|E9M403|HQ188029.1_Uncultured_Synechococcus_sp._clone_M60c211G11_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78948.1_CDS_(complement)

ENA|E9M405|HQ188030.1_Uncultured_Synechococcus_sp._clone_M60c657B8_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78950.1_CDS_(complement)

ENA|E9M408|HQ188032.1_Uncultured_Synechococcus_sp._clone_M60c045H2_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78953.1_CDS_(complement)

ENA|E9M414|HQ188035.1_Uncultured_Synechococcus_sp._clone_M60c084D4_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78959.1_CDS_(complement)

ENA|E9M415|HQ188036.1_Uncultured_Synechococcus_sp._clone_M60c124D6_allophycocyanin_subunit_B-like_(apcB)_gene,_partial_sequence;_and_allophycocyanin_subunit_A_(apcA)_gene,_partial_cds._:_Protein_ADU78960.1_CDS_(complement)

ENA|E9M419|HQ188038.1_Uncultured_Synechococcus_sp._clone_M60c129E6_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78964.1_CDS_(complement)

ENA|E9M423|HQ188040.1_Uncultured_Synechococcus_sp._clone_M60c194E10_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78968.1_CDS_(complement)

ENA|E9M425|HQ188041.1_Uncultured_Synechococcus_sp._clone_M60c206E11_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78970.1_CDS_(complement)

ENA|E9M437|HQ188047.1_Uncultured_Synechococcus_sp._clone_M60c395E1_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78982.1_CDS_(complement)

ENA|E9M445|HQ188052.1_Uncultured_Synechococcus_sp._clone_M60c486E7_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78990.1_CDS_(complement)

ENA|E9M451|HQ188055.1_Uncultured_Synechococcus_sp._clone_M60c679C9_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78996.1_CDS_(complement)

ENA|E9M454|HQ188057.1_Uncultured_Synechococcus_sp._clone_M60c011D1_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78999.1_CDS_(complement)

ENA|E9M456|HQ188062.1_Uncultured_Synechococcus_sp._clone_M60c643H6_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU79008.1_CDS_(complement)

ENA|E9M460|HQ188060.1_Uncultured_Synechococcus_sp._clone_M60c489B7_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU79005.1_CDS_(complement)

ENA|E9M461|HQ188061.1_Uncultured_Synechococcus_sp._clone_M60c577B2_allophycocyanin_subunit_B-like_(apcB)_gene,_partial_sequence;_and_allophycocyanin_subunit_A_(apcA)_gene,_partial_cds._:_Protein_ADU79006.1_CDS_(complement)

ENA|G9C3E4|HQ828096.1_Arthrospira_erdosensis_eb_allophycocyanin_alpha_chain_(apcA)_and_allophycocyanin_beta_chain_(apcB)_genes,_complete_cds._:_Protein_AEV40861.1_CDS

ENA|G9C3E6|HQ828097.1_Arthrospira_platensis_edb_allophycocyanin_alpha_chain_(apcA)_and_allophycocyanin_beta_chain_(apcB)_genes,_complete_cds._:_Protein_AEV40863.1_CDS

ENA|G9C3F2|HQ828100.1_Arthrospira_jenneri_fb_allophycocyanin_alpha_chain_(apcA)_and_allophycocyanin_beta_chain_(apcB)_genes,_complete_cds._:_Protein_AEV40869.1_CDS

ENA|P16572|M20807.1_F.diplosiphon_allophycocyanin_alpha_subunit_(apcA2)_gene,_complete_cds._:_Protein_AAA24877.1_CDS

ENA|P50030|BA000039.2_Thermosynechococcus_elongatus_BP-1_DNA,_complete_genome._:_Protein_BAC08509.1_CDS_(complement)

ENA|P72504|D86179.1_Spirulina_platensis_DNA_for_allophycocyanin_alpha_subunit,_allophycocyanin_beta_subunit,_complete_cds._:_Protein_BAA19985.1_CDS

ENA|P80555|BA000019.2_Nostoc_sp._PCC_7120_DNA,_complete_genome._:_Protein_BAB77545.1_CDS

ENA|Q01951|BA000022.2_Synechocystis_sp._PCC_6803_DNA,_complete_genome._:_Protein_BAA17874.1_CDS

ENA|Q02923|L02308.1_Synechocystis_sp._allophycocyanin_alpha_and_beta_subunit__(apcA_and_apcB)_genes,_and_core_linker_(apcC)_gene,_complete_cds._:_Protein_AAA69682.1_CDS

ENA|Q0I7Q2|CP000435.1_Synechococcus_sp._CC9311,_complete_genome._:_Protein_ABI45037.1_CDS

ENA|Q2JLK2|CP000240.1_Synechococcus_sp._JA-2-3B'a(2-13),_complete_genome._:_Protein_ABD02405.1_CDS

ENA|Q2JSL0|CP000239.1_Synechococcus_sp._JA-3-3Ab,_complete_genome._:_Protein_ABD00365.1_CDS_(complement)

ENA|Q55092|L76083.1_Synechocystis_sp._allophycocyanin_subunit_(apcA_and_apcB)_genes,_partial_cds's._:_Protein_AAC41530.1_CDS

ENA|Q7NL80|BA000045.2_Gloeobacter_violaceus_PCC_7421_DNA,_complete_genome._:_Protein_BAC89187.1_CDS

ENA|Q7U8X4|BX569690.1_Synechococcus_sp._WH8102_complete_genome;_segment_2/7_:_Protein_CAE07000.1_CDS_(complement)

ENA|Q8YZK7|BA000019.2_Nostoc_sp._PCC_7120_DNA,_complete_genome._:_Protein_BAB72408.1_CDS_(complement)

ENA|T1YXU4|KF052037.1_Arthrospira_platensis_qy3_allophycocyanin_alpha-subunit_(apcA)_and_allophycocyanin_beta-subunit_(apcB)_genes,_complete_cds._:_Protein_AGU69493.1_CDS

ENA|U5QFD5|CP003587.1_Gloeobacter_kilaueensis_JS1,_complete_genome._:_Protein_AGY57593.1_CDS_(complement)

ENA|V5V826|CP006735.1_Thermosynechococcus_sp._NK55_genome._:_Protein_AHB89511.1_CDS_(complement)

ENA|W0GWR0|CP006882.1_Synechococcus_sp._WH_8109,_complete_genome._:_Protein_AHF64605.1_CDS

ENA|A0A1L6BXH8|KX758286.1_Synechococcus_sp._OH28_allophycocyanin_subunit_alpha_(apcA)_gene,_partial_cds._:_Protein_APQ41713.1_CDS_0

ENA|A0A1L6BXI5|KX758288.1_Synechococcus_sp._OH2_allophycocyanin_subunit_alpha_(apcA)_gene,_partial_cds._:_Protein_APQ41715.1_CDS_0

ENA|A0A2L2P5Q3|MG014397.1_Synechococcus_sp._RS9915_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76488.1_CDS_0

ENA|A0A2L2P5U7|MG014395.1_Synechococcus_sp._TAK9802_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76486.1_CDS_0

ENA|A0A2L2P5U9|MG014402.1_Synechococcus_sp._PROS-U-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76493.1_CDS_0

ENA|A0A2L2P5V0|MG014405.1_Synechococcus_sp._NOUM97013_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76496.1_CDS_0

ENA|A0A2L2P5V1|MG014412.1_Synechococcus_sp._BMK-MC-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76503.1_CDS_0

ENA|A0A2L2P5V2|MG014411.1_Synechococcus_sp._BOUM118_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76502.1_CDS_0

ENA|A0A2L2P5V3|MG014406.1_Synechococcus_sp._MVIR-18-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76497.1_CDS_0

ENA|A0A2L2P5V4|MG014413.1_Synechococcus_sp._BIOS-U3-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76504.1_CDS_0

ENA|A0A2L2P5V5|MG014404.1_Synechococcus_sp._PROS-7-1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76495.1_CDS_0

ENA|A0A2L2P5V6|MG014407.1_Synechococcus_sp._MIT_S9220_isolate_MITS9220_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76498.1_CDS_0

ENA|A0A2L2P5W0|MG014420.1_Synechococcus_sp._A15-60_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76511.1_CDS_0

ENA|A0A2L2P5W1|MG014419.1_Synechococcus_sp._A15-62_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76510.1_CDS_0

ENA|A0A2L2P5W2|MG014421.1_Synechococcus_sp._A15-44_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76512.1_CDS_0

ENA|A0A2L2P5W6|MG014423.1_Synechococcus_sp._A15-24_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76514.1_CDS_0

ENA|A0A2L2P5W7|MG014415.1_Synechococcus_sp._A18-46.1_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76506.1_CDS_0

ENA|A0A2L2P5W8|MG014422.1_Synechococcus_sp._A15-28_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76513.1_CDS_0

ENA|A0A2L2P5W9|MG014416.1_Synechococcus_sp._A18-40_allophycocyanin_alpha_chain_(apcA)_gene,_complete_cds._:_Protein_AVH76507.1_CDS_0

ENA|A0A2Z5X2N9|AP017560.1_Pseudanabaena_sp._ABRG5-3_DNA,_complete_genome._:_Protein_BBC25951.1_CDS_(complement)_0

ENA|A0A3G9JUA6|AP019314.1_Microcystis_viridis_NIES-102_DNA,_complete_genome._:_Protein_BBH41821.1_CDS_(complement)_0

ENA|B0C0G1|CP000828.1_Acaryochloris_marina_MBIC11017,_complete_genome._:_Protein_ABW30754.1_CDS_(complement)_0

ENA|B0CFZ9|CP000828.1_Acaryochloris_marina_MBIC11017,_complete_genome._:_Protein_ABW29446.1_CDS_0

ENA|E9M3F5|HQ187853.1_Uncultured_Synechococcus_sp._clone_SynAculture_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78883.1_CDS_(complement)_0

ENA|E9M3F7|HQ187803.1_Uncultured_Synechococcus_sp._clone_M60c544G10_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78786.1_CDS_(complement)_0

ENA|E9M3H6|HQ187801.1_Uncultured_Synechococcus_sp._clone_M60c534H9_allophycocyanin_subunit_B_(apcB)_and_allophycocyanin_subunit_A_(apcA)_genes,_partial_cds._:_Protein_ADU78782.1_CDS_(complement)_0

ENA|E9M3P0|HQ187842.1_Uncultured_Synechococcus_sp._clone_M65c597G8_allophycocyanin_subunit_A_(apcA)_gene,_partial_cds._:_Protein_ADU78861.1_CDS_(complement)_0