Factors Regulating Phytoplankton Community Composition in a Large Lake

Jessica Owen
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
Owen, Jessica, "Factors Regulating Phytoplankton Community Composition in a Large Lake" (2019). Electronic Theses and Dissertations. 7651.
https://scholar.uwindsor.ca/etd/7651

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
Factors Regulating Phytoplankton Community Composition in a Large Lake

By

Jessica M. Owen

A Thesis
Submitted to the Faculty of Graduate Studies
through the Great Lakes Institute for Environmental Research
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2019

© 2019 Jessica M. Owen
Factors Regulating Phytoplankton Community Composition in a Large Lake

by

Jessica M. Owen

APPROVED BY:

______________________________________________
J. Ciborowski
Department of Biological Sciences

______________________________________________
C.G. Weisener
Great Lakes Institute for Environmental Research

______________________________________________
K. Stammler
Essex Region Conservation Authority

______________________________________________
S. Chaganti, Co-advisor
Great Lakes Institute for Environmental Research

______________________________________________
G.D. Haffner, Co-advisor
Great Lakes Institute for Environmental Research

January 17th, 2019
DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.
ABSTRACT

Aquatic environments have experienced long term anthropogenic stressors from habitat alteration, invasive species, and nutrient inputs which leads to increased eutrophication. Lake Erie is a large lake system that has historically seen highly eutrophic conditions. Nutrient loading targets were put in place in the 1970s and successfully reduced lake wide eutrophication by the 1980s. However, Lake Erie has recently seen an increase in eutrophication and harmful phytoplankton species, such as *Microcystis*. Therefore, studying community composition and factors regulating community composition is necessary. This increase in eutrophication and potential shifts to problematic species was studied on spatial and temporal scales, and through microscopic and genomic techniques. Microscopy was used to determine annual average biomass which ranged between 4-6 g/m$^3$ in 2016 and 2017. Microscopy and next generation sequencing (NGS) were used to determine phytoplankton community composition which was very diverse. Presence of oligotrophic species, such as *Dinobryon*, and eutrophic species, such as *Microcystis*, were seen in Lake Erie often during the same periods. Diatoms were a very dominant phytoplankton class especially in 2016 whereas cyanobacteria were slightly more dominant in 2017. Microscopic and genomic techniques were also used to determine factors regulating this composition through CCA plots. It was concluded that chemical factors such as SRP and NO$_2$-NO$_3^-$, and physical factors such as wind, water temperature, and euphotic depth were important for potential harmful phytoplankton growth and distribution. It also concluded that there was a large spatial and temporal distribution of phytoplankton communities in what has been considered a well-mixed western basin. Future research is needed throughout the western basin on both the Canadian and American side to determine if there are long term patterns and shifts in phytoplankton community composition.
DEDICATION

To my parents Geoff and Carol Owen, for your love and support.
ACKNOWLEDGEMENTS

I would like to thank everyone who has supported and contributed to my success throughout my MSc. Firstly, I would like to thank my supervisor Dr. Doug Haffner and Dr. Subbarao Chaganti for providing the opportunity to study with them. Without their guidance, I would not have the skills I have today. Secondly, I would like to thank my committee members Dr. Katie Stammler, Dr. Chris Weisener, and Dr. Jan Ciborowski for their input and direction. I would like to thank Essex Region Conservation Authority (ERCA), the township of Pelee Island, and the Ontario Ministry of Natural Resource and Forestry (OMNRF) Wheatley for providing equipment and passage for sample collection. I was especially grateful for their support during the summer sample season of 2017 when Lake Erie experienced the algal bloom.

Thank you to everyone in the Haffner and Heath lab who aided with lab and field work: Erin Hillis, Lei Zhang, Lucas Wilson, Kate Brown, Uyen Vo, Brenten Vercruysse, Arthur Adamczak, Cecilia Heuvel, Dylan Xia, Fu mei, Mariam Ageli, Danielle VanMensel, Abdolrazagh Hashemi, Mubarak Ziab, and Shelby Mackie. I would like to thank Todd Leadley for assisting with sampling and driving the boat during the sample season. I would like to thank Da-wei Qiu and Nargis Ismail for providing the facilities and guidance during nutrient analyses. I would also like to thank Yingming Zhao and Tyler Genereaux from the OMNRF for providing a boat to use while the Loon was in for repairs. I would like to thank Celine Lajoie for guidance with presentations and edits, as well as Erin Hillis and Anne MacLeod who provided assistance on presentations, edits, and statistical analyses. I would like to thank Mary-Lou Scratch, Christine Weisener, and Christina Semeniuk for their administrative support.

I would like to thank my family: Carol Owen, Geoff Owen, Kris Owen, Devin Owen, Sidney Owen, Margaret Owen, and dogs Zinzan and Rio their constant love and support throughout my MSc. Lastly, I would like to thank my friends: Nicole Jahn, Tristan (Angie) Desaulniers, Blake DiBiase, and Taylar Rock for all their love. Without their encouragement and support, I would have never pursued my MSc at GLIER.

Funding support was provided by Natural Sciences and Engineering Research Council of Canada (NSERC), the University of Windsor, ERCA and Mitacs.
# TABLE OF CONTENTS

DECLARATION OF ORIGINALITY ..................................................................................... iii

ABSTRACT ....................................................................................................................... iv

DEDICATION .................................................................................................................... v

ACKNOWLEDGEMENTS ................................................................................................. vi

LIST OF TABLES ........................................................................................................... ix

LIST OF FIGURES ......................................................................................................... x

LIST OF APPENDICES .................................................................................................. xiii

CHAPTER ONE: GENERAL INTRODUCTION ................................................................ 1

1.1 Introduction .............................................................................................................. 1

1.2 References ............................................................................................................... 6

CHAPTER TWO: FACTORS REGULATING PHYTOPLANKTON COMMUNITY COMPOSITION AND
ABUNDANCE IN THE WESTERN BASIN OF LAKE ERIE ........................................... 10

2.1 Introduction .............................................................................................................. 10

2.2 Methods .................................................................................................................. 13

2.2.1 Sample Collection .............................................................................................. 13

2.2.2 Water Column Profiles ...................................................................................... 14

2.2.3 Phytoplankton community composition and biomass ...................................... 14

2.2.4 Nutrients ........................................................................................................... 16

2.2.5 Chlorophyll-a .................................................................................................... 17

2.2.6 Wind Data ......................................................................................................... 18

2.2.7 Qualitative and Statistical Analyses .................................................................. 18

2.3 Results .................................................................................................................... 19

2.3.1 RBR *maestro* logger Data Collection .............................................................. 19

2.3.2 Community Composition in 2016 and 2017 ...................................................... 20

2.3.3 Cyanobacterial Genus Composition and Changes in % Cyanobacteria Genus Composition ........................................................................................................ 21

2.3.4 Environmental Factors and Phytoplankton Community .................................. 22

2.4 Discussion ............................................................................................................... 22

2.4.1 Community Composition .................................................................................. 22

2.4.2 Harmful Algal Blooms ..................................................................................... 26

2.4.3 Environmental Factors ..................................................................................... 27
## CHAPTER THREE: GENOMIC TECHNIQUES AS A TOOL TO TRACK AND UNDERSTAND PHYTOPLANKTON DYNAMICS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>51</td>
</tr>
<tr>
<td>3.2 Methods</td>
<td>54</td>
</tr>
<tr>
<td>3.2.1 Sampling Protocol</td>
<td>54</td>
</tr>
<tr>
<td>3.2.2 Water Column Profiles</td>
<td>54</td>
</tr>
<tr>
<td>3.2.3 Nutrients</td>
<td>55</td>
</tr>
<tr>
<td>3.2.4 Chlorophyll-α Extraction</td>
<td>55</td>
</tr>
<tr>
<td>3.2.5 DNA Extraction, PCRs and Next Generation Sequencing</td>
<td>55</td>
</tr>
<tr>
<td>3.2.6 16S rRNA gene and 18S rRNA gene Taxonomy and Bioinformatics</td>
<td>56</td>
</tr>
<tr>
<td>3.2.7 Qualitative and Statistical Analyses</td>
<td>57</td>
</tr>
<tr>
<td>3.3 Results</td>
<td>58</td>
</tr>
<tr>
<td>3.3.1 rRNA gene reads and OTU counts</td>
<td>58</td>
</tr>
<tr>
<td>3.3.2 Community Composition and Change in Harmful Cyanobacteria Frequency</td>
<td>59</td>
</tr>
<tr>
<td>3.3.3 Spatial-Temporal Factors and Clustering</td>
<td>61</td>
</tr>
<tr>
<td>3.3.4 Microbial Clustering</td>
<td>62</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>65</td>
</tr>
<tr>
<td>3.4.1 Harmful Algal Growth</td>
<td>65</td>
</tr>
<tr>
<td>3.4.2 Spatial and Temporal Variation</td>
<td>66</td>
</tr>
<tr>
<td>3.4.3 Environmental Factors influencing Microbial Community Composition</td>
<td>67</td>
</tr>
<tr>
<td>3.5 Conclusion</td>
<td>68</td>
</tr>
<tr>
<td>3.6 References:</td>
<td>70</td>
</tr>
</tbody>
</table>

## CHAPTER FOUR: GENERAL DISCUSSION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Discussion</td>
<td>90</td>
</tr>
<tr>
<td>4.2: Final Conclusion and Future Work</td>
<td>93</td>
</tr>
<tr>
<td>4.3: References</td>
<td>96</td>
</tr>
</tbody>
</table>

APPENDIX A ............................................................................. 98

VITA AUCTORIS ...................................................................... 107
LIST OF TABLES

**Table 2.1:** Sampling dates for the western basin a) 2016 sampling year b) 2017 sampling year

**Table 2.2:** One-way ANOVA Kruskal Wallis showing difference between % composition and biomass of each phytoplankton group throughout the sample season and bloom conditions in 2016 and 2017. ***= p<0.0001

**Table 2.3:** One-way ANOVA Kruskal Wallis showing difference between % composition and biomass of each cyanobacteria group throughout sample season and bloom conditions in 2017 and 2017. * p<0.05, *** p<0.0001

**Table 2.4:** PERMANOVA results showing significance of CCA data results for phytoplankton percent community composition. 2016 and 2017 were analyzed together. * p<0.05

**Table 3.1:** Primers and primer sequences used for DNA PCR amplification for 16S rRNA and 18S rRNA in 2016 and 2017

**Table 3.2:** PERMANOVA results showing significance of CCA data results for 16S and 18S rRNA gene data for 2016 and 2017. * p<0.05

**Table 3.3:** Shepard plot values of stress for nMDS analyses using Bray-Curtis similarity index
LIST OF FIGURES


Figure 2.2: Seasonal changes in phytoplankton biomass at Colchester Inshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average historical seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………37

Figure 2.3: Seasonal changes in phytoplankton biomass at Colchester Offshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………38

Figure 2.4: Seasonal changes in phytoplankton biomass at Pelee Inshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………39

Figure 2.5: Seasonal changes in phytoplankton biomass at Pelee Offshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………40

Figure 2.6: Seasonal changes in phytoplankton biomass at Colchester Inshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………41

Figure 2.7: Seasonal changes in phytoplankton biomass at Colchester Offshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………42

Figure 2.8: Seasonal changes in phytoplankton biomass at Pelee Inshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie………………………………………………………………………………………………………………………………43
Figure 2.9: Seasonal changes in phytoplankton biomass at Pelee Offshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie…………………………………………………………44

Figure 2.10: Pie charts showing yearly changes in % cyanobacteria from 2016 and 2017 at a) Colchester Inshore; b) Colchester Offshore; c) Pelee Inshore; and d) Pelee Offshore. Chryso=Chrysophytes, Crypto=Cryptophytes, Chloro=chlorophytes, and Cyanos=cyanophytes………..45

Figure 2.11: Pie charts showing biomass changes in cyanobacteria community composition from 2016 to 2017 at a) Colchester Inshore; b) Colchester Offshore; c) Pelee Inshore; and d) Pelee Offshore…………46

Figure 2.12: One-way ANOVA Kruskal Wallis showing significant changes in biomass for cyanobacteria genus types during bloom conditions in 2016 and 2017. ***p<0.0001………………47

Figure 2.13: One-way ANOVA Kruskal Wallis showing significant changes in biomass for whole community composition to group during bloom conditions in 2016 and 2017……………………………………..47

Figure 2.14: One-way ANOVA Kruskal Wallis showing significant changes in % community composition for whole community to phytoplankton group during bloom conditions in 2016 and 2017…………………………………………………………………………………………….48

Figure 2.15: Box and whisker plot showing the distribution of wind speeds 60 hours before the end of the sampling dates in 2016. Average speeds were 21-Jun= 11.3km/h, 19-Jul=14.9km/h, 26-Aug= 14.6km/h, 12-Sep= 13.5km/h, and 4-Oct= 7.1km/h……………………………………………………………………………………………………48

Figure 2.16: Box and whisker plot showing the distribution of wind speeds 60 hours before the end of the sampling dates in 2017. Average speeds were 8-Jun= 15.6km/h, 21-Jun=16.2km/h, 6-Jul= 8.7km/h, 26-Jul=12.3km/h, 8-Aug=9.4km/h, 30-Aug= 6.9km/h, 15-Sep=7.6km/h, 27-Sep=8.5km/h……………………………………..49

Figure 2.17: CCA plots for monthly distribution of phytoplankton communities in 2016 and 2017…….49

Figure 2.18: CCA plot for site location and distribution of phytoplankton communities in 2016 and 2017………………………………………………………………………………………………………………………………………………50

Figure 2.19: CCA plot for depth distribution of phytoplankton communities in 2016 and 2017…….50

Figure 3.1: Phytoplankton community composition determined by 16S rRNA reads in 2016 and 2017…………………………………………………………………………………………………………………………………………………………………………………………79

Figure 3.2: Phytoplankton community composition determined by 18S rRNA reads in 2016 and 2017…………………………………………………………………………………………………………………………………………………………………………………………79

Figure 3.3: Relative abundance of Microcystis and Anabaena through the sample season at all sample locations in 2016. Depths shown are 1m, 3m, 7m, and 10m. Microcystis is orange, and Anabaena is grey…………………………………………………………………………………………………………………..80
Figure 3.4: Relative abundance of *Microcystis* and *Anabaena* through the sample season at all sample locations in 2017. Depths shown are 1m, 3m, 7m, and 10m. *Microcystis* is orange, and *Anabaena* is grey………………………………………………………………………………...81

Figure 3.5: nMDS plot of 16S rRNA 2016 Month................................................................................82

Figure 3.6: nMDS plot of 16S rRNA 2017 Month................................................................................82

Figure 3.7: nMDS plots of 18S rRNA 2016 Month.............................................................................83

Figure 3.8: nMDS plots of 18S rRNA 2017 Month.............................................................................83

Figure 3.9: 16S rRNA CCA for depth in 2016.....................................................................................84

Figure 3.10: 16S rRNA CCA for site in 2016.....................................................................................84

Figure 3.11: 16S rRNA CCA for season in 2016................................................................................85

Figure 3.12: 16S rRNA CCA for depth in 2017................................................................................85

Figure 3.13: 16S rRNA CCA for site in 2017.....................................................................................86

Figure 3.14: 16S rRNA CCA for season in 2017................................................................................86

Figure 3.15: 18S rRNA CCA for depth in 2016................................................................................87

Figure 3.16: 18S rRNA CCA for site in 2016.....................................................................................87

Figure 3.17: 18S rRNA CCA for season in 2016................................................................................88

Figure 3.18: 18S rRNA CCA for depth in 2017................................................................................88

Figure 3.19: 18S rRNA CCA for site in 2017.....................................................................................89

Figure 3.20: 18S rRNA CCA for season in 2017................................................................................89
LIST OF APPENDICES

Appendix A.............................................................................................................................................94
CHAPTER ONE: GENERAL INTRODUCTION

1.1 Introduction

Aquatic environments have been under increasing stress including overharvesting, invasive species, habitat alteration, and cultural eutrophication (Wilson 2002). These issues have long been present leading to harmful algal blooms becoming more prevalent in both marine and freshwater systems. Harmful algal blooms (HABs), especially *Microcystis* dominant growths, are present all over the world except Antarctica (Harke et al 2016a, Zurawell et al 2005). These harmful algal blooms are caused by an increase in nutrients into an aquatic system increasing phytoplankton biomass (Davis et al 2015, IJC et al 2014, Fuller et al 2002, Munawar et al 2002). In 1974, Schindler concluded that phosphorus was the main contributor in regulating freshwater eutrophication. As well as nutrients, temperature is also important for harmful algal bloom growth. With increasing atmospheric and water temperatures, harmful algal blooms are predicted to become more common, more severe, and have a longer growing season especially in temperate lakes such as the Laurentian Great Lakes (Visser et al 2016).

Harmful algal blooms can be described as large colonies of algae that are able produce toxins or toxic effects on the local ecosystem (NOAA 2016). The algae present do not need to produce toxins for a system to be affected by a harmful algal bloom. However, harmful algal blooms are often dominated by toxin producing cyanobacteria (Paerl et al 2001) in freshwater systems. Cyanobacteria become dominate due to their ability to outcompete other phytoplankton species at higher temperatures (Paerl and Huisman 2008, Johnk et al 2008, Elliott et al 2006), extreme light condition, and low nutrient conditions (Paerl et al 2001). Buoyant cyanobacteria, such as *Microcystis* and *Anabaena*, can migrate throughout the water column due to their gas vesicles (Huisman and Hulot 2005, Walsby et al 1994). This feature allows many cyanobacteria species to exploit their most favoured part of the water column and grow in large numbers.
(Huisman and Hulot 2005). The surface accumulation of cyanobacteria might also cause water temperatures to increase allowing them to further exploit surface waters causing a positive feedback system where they are the dominant phytoplankton group at the surface (Paerl and Huisman 2009, Hense 2007).

The Laurentian Great Lakes are a large freshwater resource and hold approximately 20% of the world’s surface freshwater (Davis et al 2015). Eutrophication has long been an issue in the Great Lakes where harmful algal blooms were first observed in the 1960s especially in Lake Erie (Davis et al 2015, Bertram 1993, Makarewicz et al 1991, Rosa et al 1987). In 1972, the Great Lakes Water Quality Agreement (GLWQA) was established to help improve water quality in the Great Lakes (IJC 2012). The easiest method to reduce the severity of HABS at the time was to reduce phosphorus from point source discharges to reach the phosphorus loading target for Lake Erie (IJC 2015, Perri et al 2015). This loading target was set at 11,000 MT/yr and current loadings have been at or below the loading target since the mid-1980s (Scavia et al 2014).

Lake Erie is the most productive Laurentian Great Lake, providing many benefits to the ecosystem including profitable fisheries and a diverse ecosystem (Watson et al 2016, LEIA et al 2012). Since the 1990s, Lake Erie’s ecosystem has been increasingly plagued by eutrophication and this has caused the return of the harmful algal blooms (IJC 2015). These HABs are dominated by *Microcystis* spp and *Anabaena* spp which can produce hepatotoxins and neurotoxins (Harke et al 2016b, Bridgeman et al 2012, Stumpf et al 2012, Rinta-Kanto et al 2005). The reappearance of these *Microcystis* dominant HABs have impacted water quality including recreational water and drinking water use.

Guidelines have been set to prevent populations from encountering toxic waters. For drinking water, Microcystin concentrations need to be below 1.5 µg/L (WHO 1998) to prevent...
any issues with toxicity and in Windsor Essex county concentrations need to be below 10 µg/L (WECHU 2019) for recreational use. The government of Canada prevents recreational use above 20 µg/L of Microcystin (Health Canada 2012). If concentration of Microcystin exceeds these limits, water intake plants will stop distributing water and recreational beaches will be closed. One of the most well-known water quality issues caused by HABs in Lake Erie was the Toledo Water Crisis in 2014. This event affected nearly 500,000 people for two days and cost nearly $65 million (USD) in economic losses (Steffen et al 2016). Canada’s largest island in Lake Erie, Pelee Island, has not been able to drink water from local wells many summers due to issues with high Microcystin concentrations (CTV Windsor 2017; CTV Windsor 2014). Source water protection legislation in Essex County states that microcystin-LR is a toxin of concern and therefore a drinking water issue (ERCA 2015). Due to this drinking water issue, the Essex Region Source Protection Committee states that all water intake plants need to monitor microcystin concentrations through the summer season through the Drinking Water Surveillance Program (ERCA 2015).

There are many hypotheses to why HABs are becoming more prevalent in Lake Erie. While total phosphorus (TP) entering the lake has remained under 11,000 MT/yr since the 1980s, one hypothesis suggests that there is increased soluble reactive phosphorus (SRP) entering the western basin from the Maumee River (Kane et al 2014, Charlton and Milne 2004). As well, invasive zebra mussels (*Dreissena polymorpha*) have changed nutrient cycling within the basin increasing the amount of SRP available to phytoplankton (Lindim 2015, Conroy et al 2005, James et al 2001, Johengen et al 1995). Another nutrient hypothesis states that there is legacy phosphorus located in the lake sediments (Matisoff and Carson 2014). This may suggest that there are also cyanobacteria propagules located in the sediments, and that these are resuspended
into the above water column during mixing events such as heavy winds and springtime storm events (IJC 2014). This suggests that sediment is an internal loading source for both nutrients and cyanobacterial biomass. However, phosphorus is not the only nutrient of concern in Lake Erie. Studies have shown that both phosphorus and nitrogen together can cause a greater increase in phytoplankton biomass than phosphorus alone (Paerl et al 2016). Other hypotheses suggest that climate change will lead to cyanobacteria dominant HABs (cHABs) due to their large niche (Hallegraeff 2010) and high temperature tolerance (Davis et al 2009; Fujimoto et al 1997).

These many hypotheses reveal that there are many possible factors contributing to the prevalence and persistence of HABs in Lake Erie. While these hypotheses are important to determine HAB presence and mitigation, they do not explain the overall phytoplankton community response with respect to composition and relative abundance of species. It has been predicted that cyanobacteria, especially *Microcystis*, will become the dominant phytoplankton species (Hallegraff 2010, Davis et al 2009, Fujimoto et al 1997), especially in Lake Erie where the conditions are ideal for the cyanobacteria (Davis et al 2009, Paerl and Huisman 2008, Paul 2008, Paerl et al 2001, Paerl 1998).

Research in Lake Erie on factors regulating community composition and relative abundance has been rare, and mainly limited to phytoplankton community composition. Two Canadian studies examined how phytoplankton biomass changed with sampling years. Nicholls et al 1993 concluded that the P control actions of the GLWQA decreased total phosphorus entering the lake, as well as total biomass after the 1970s. However, dreissenid mussels also greatly decreased phytoplankton biomass across phytoplankton groups (Nicholls et al 1993). Makarewicz et al 1999 showed a very similar pattern noting that cyanobacterial concentrations
had increased in late spring and decreased in the late summer after the dreissenid mussel invasion in 1987.

The objective of this thesis was to determine the relative importance of factors regulating the seasonal, spatial, temporal, and vertical distribution of the phytoplankton composition and relative abundance throughout the western basin of Lake Erie. The last study that observed phytoplankton community composition was nearly 20 years ago where it was found that total phosphorus loads and the dreissenid mussels influenced community composition. Phytoplankton community composition and factors regulating this community composition may have changed since the 1990s. Chapter Two examines the phytoplankton community composition and relative abundance throughout the Western Basin of Lake Erie using traditional microscopic techniques. This approach added to the understanding of current community composition, as well as seasonal, temporal, spatial, and vertical changes in composition and relative abundance. Chapter Three addresses phytoplankton community composition using genomic techniques as a novel approach for tracking changes in community composition throughout the basin. In these two chapters nutrients (SRP, TP, NO$_2$-NO$_3$), chl-$\alpha$, and light attenuation were analyzed to determine if these factors influenced compositional changes. Chapter Four compares and integrates the knowledge of algal community assemblages using microscopic and genomic approaches. It is suggested that these two techniques can be used in combination to quantify changes in community composition especially with a focus on the increasing harmful algal blooms.
1.2 References


Bridgeman, T.B; Chaffin J.D; Kane, D.D; Conroy, J.D; Panek, S.E; Armenio, P.M. From River to Lake: Phosphorus partitioning and algal community compositional changes in Western Lake Erie. J. Great Lakes Res. 2012. 38: 90-97.


CHAPTER TWO: FACTORS REGULATING PHYTOPLANKTON COMMUNITY COMPOSITION AND ABUNDANCE IN THE WESTERN BASIN OF LAKE ERIE

2.1 Introduction

In recent years, there has been an increased number of reports of harmful algal blooms, especially cyanobacteria blooms (HAB/cHAB). (Beaulieu et al 2013; O’Neil et al 2012). It is predicted that with climate change, cyanobacteria will become the dominant phytoplankton group in many aquatic environments (Hallegraeff 2010; Davis et al 2009; Fujimoto et al 1997). Lake Erie will potentially shift to a more dominant cyanobacteria ecosystem as a result of its highly productive nature, shallow depth, and warm waters that provide ideal conditions for cyanobacteria (Levy 2017).

Phytoplankton community composition, however, has varied historically in Lake Erie especially in the western basin. Pre-1960, the phytoplankton community was generally dominated by diatoms (Bacillariophyta) especially filamentous centric taxa such as Melosira/Aulacoseira (Allinger and Reavie 2013), and only during the occasional autumn season would more eutrophic species, such as cyanobacteria species Anabaena and Microcystis, dominate the algal assemblage (Allinger and Reavie 2013; Davis et al 1962). During the 1960s, however, the summer phytoplankton community and total biomass were reported to be large and dominated by cyanobacteria such that 70% of the algal community was comprised of Anabaena, Microcystis and Aphanizomenon (Casper 1965). In the 1970s and 1980s, total algal biomass decreased because of phosphorus abatement programs implemented under the Great Lakes Water Quality Agreement in 1972 (Makarewicz and Bertram 1991). In addition to decreasing total algal biomass, decreased P loads also resulted in changes to algal community composition. Makarewicz (1993) concluded that about 54% of the algal biomass in the western basin was made up primarily of diatoms especially Fragilaria spp in the 1980s. Conroy et al (2005),
however, noted that the algal biomass has not been at its minimum since the early 1990s. During this period total spring biomass measured 1.27 mg/L (Makarewicz et al 1999) and total summer biomass measured 1.79 mg/L (Makarewicz et al 1999). Despite these variations in community composition over time, there have been few studies attempting to understand factors regulating the relative abundance of algal species in Lake Erie. Furthermore, data on the algal assemblage of Lake Erie are also rare, with most studies relying only on chlorophyll-α concentrations to measure algal biomass and production (Hillis 2017, unpublished).

It has been hypothesized that the invasion of the dreissenid mussel (*Dreissena polymorpha*) in 1987 and their feeding methods selectively consumed large quantities of the algal biomass after 1987 (Nicholls et al 1993). Two studies that examined the phytoplankton community composition and biomass in the 1990s in Lake Erie concluded that this invasion strongly modified both algal community composition and biomass (Nicholls et al 1993, Makarewicz et al 1999). Nicholls et al (1993) noted that although P controls decreased algal biomass, the mussel invasion further decreased biomass/density to extremely low levels. In the mid-1970s biomass was as high as 4,508 areal standard units (ASU/mL) and decreased to 73 ASU/mL in 1989 and 30 ASU/mL in 1990. Cyanophyte densities decreased from 120 ASU/mL to 4 ASU/mL and diatom density decreased from about 520 ASU/mL to 44 ASU/mL from 1989 to 1990. Makarewicz et al (1999) quantified community composition and biomass during the same period and observed that while cyanobacteria concentrations decreased during the autumn (0.08 g/m$^3$ to 0.01 g/m$^3$), concentrations increased in the spring (0.02 g/m$^3$ to 0.065 g/m$^3$). Total algal biomass decreased from about 2.5 g/m$^3$ pre-mussel invasion to about 1.8 g/m$^3$ (Makarewicz et al 1999) post mussel invasion. As well as decreased biomass, compositional makeup of the
community changed. Munawar et al 2002 noted that while biomass was higher in the 1970s, fewer species were identified.

During phytoplankton surveys in the 1990s, less common species made up a larger proportion of the community with many species from Chlorophyta and Bacillariophyta contributing significantly to the total algal biomass (Munawar et al 2002). Since the late 1990s there has been an increase in phytoplankton biomass especially cyanobacteria in the summer months. (Conroy et al 2005). Increases in soluble reactive phosphorus (SRP) (Scavia et al 2014), selective filter feeding by zebra mussels (Watson et al 2016; Hecky et al 2004), legacy phosphorus in sediments (Matisoff and Carson 2014), and climate change (O’Neil et al 2012; Paerl and Huisman 2008) have all been identified as possible causes of the recent increases in cyanobacteria biomass. Kruk et al (2011) suggested that while predicting future composition will be very difficult, understanding the morphology of phytoplankton species currently present in a Lake Erie will be the best indicator of how the algal community will react to changing environmental conditions and which species will become dominant.

Understanding changes in environmental conditions and how different species are responding to these changes is a critical challenge if we are to predict how lakes respond to remedial actions. Distinguishing phytoplankton biomass by classes (i.e potentially toxin producing species vs non-toxic species) helps to quantify human and environmental health hazards but does not provide insight as to how the assemblage is responding to multiple stressors in a spatially and temporally dynamic system like Lake Erie. More detailed knowledge on factors regulating species composition, relative abundance, seasonality and spatial distributions is essential. It is also essential because it could help predict future conditions that could exacerbate human and environmental health hazards.
The objective of this study was to determine the current phytoplankton community composition and relative abundance in the western basin of Lake Erie. We aimed to determine factors regulating seasonal, temporal, and vertical changes in community composition and abundance by examining the following three questions.

a) Are there seasonal and temporal community shifts in phytoplankton community composition and abundance, and is this pattern stable between different years?

b) Does the relative abundance of species remain constant during bloom periods (i.e. are all taxa responding in a similar manner)?

c) Are there strong differences in community composition and relative abundance among sites and depth in the well mixed western basin?

2.2 Methods

2.2.1 Sample Collection

Water column sampling occurred at two sites in the nearshore waters (< 7 m) (Colchester Inshore: N41°58.801', W82°56.183'; Pelee Inshore: N41°50.290', W82°40.268') and two sites in the offshore waters (>10 m) (Colchester Offshore: N41°51.396', W82°59.137'; Pelee Offshore: N41°50.290', W82°39.837') on the Canadian side of the western basin of Lake Erie. Colchester and Pelee Inshore were chosen due to their assumed influence from agricultural runoff. Colchester Offshore was chosen due to its assumed influence from Lake Huron via the Detroit River, and Pelee Offshore was chosen due to its assumed influence from the Maumee River nutrient and algal plume (Fig 2.1). Sampling was conducted monthly from June to October in 2016 and April to September bi-monthly in 2017 (Table 2.1). The water column samples were collected at four depths (1, 3, 7, and 10 m) and analyzed for chlorophyll-a, nutrients (SRP, TP, NO₂⁻ and NO₃⁻), and phytoplankton community composition and biomass using traditional
techniques using the inverted microscope method (Utermöhl 1958). Average maximum depth for Colchester Inshore was 7.2 m and was 11.3 m at all other sites.

2.2.2 Water Column Profiles

The RBR maestro logger (Brancker) was used to determine environmental parameters such as water temperature, pH, dissolved oxygen, and conductivity throughout the water column from surface to just above the sediments at each site. Water temperature was averaged for entire water column. Light measurements were taken at each metre until extinction (1 μmol/cm²/s) using a LI-193 spherical quantum light sensor. The light attenuation coefficient and euphotic depth were determined according to the Beer-Lambert Law such that

a) \( \varepsilon_{\text{par}} = (\ln I_o - \ln I_z)/z \), and therefore

b) \( 0.01 I_o = I_o e^{-\varepsilon_{\text{par}} Z_{\text{e}}} \) (assuming 1% of the surface light intensity represents the euphotic depth \( Z_{\text{e}} \)) (Zhang et al 2006). Therefore, euphotic depth was calculated as

c) \( Z_{\text{e}} = 4.6/\varepsilon_{\text{par}} \)

Where;

\( I_o = \) irradiance at the surface

\( I_z = \) irradiance at a given depth, \( z \).

\( \varepsilon_{\text{par}} = \) the light attenuation coefficient.

To minimize the effect of waves focusing surface light, the light attenuation coefficient was measured between 2 m (Wetzel 2001) and \( I_z \) (1 μmol/cm²/s)

2.2.3 Phytoplankton community composition and biomass

Water samples were collected at each depth and stored in 250 mL glass amber bottles with 5 mL of Lugol’s iodine to preserve the samples (Eaton et al 1995). To determine algal
composition and relative abundance, the bottles were shaken, and a sub-sample placed in 5 mL graduated cylinders overnight which allowed phytoplankton to settle onto 1 mL slides. The slides were viewed under 400x magnification on a Leica DM IRB inverted microscope (Utermöhl 1958) using the Leica Application Suite 4.5 software to measure cell size to estimate bio-volumes of all taxa observed. Ten random fields of view were examined, and photos taken for each slide. Total biomass (g/m$^3$) was determined by calculating the average cell bio-volume. This was done by assuming a specific gravity of unity (Strickland 1960) for each phytoplankton cell, such that phytoplankton with a biovolume of $10^9$ μm$^3$ have an assumed mass of 1 mg. Multiple equations (which are listed below) were used to correct for area examined to total slide area, and to convert slide area to (in mL) to biomass at depth (g/m$^3$). The total slide area was 0.7854 cm$^2$ and the field of view area at 400x magnification was 0.00050195 cm$^2$. After taxa specific bio-volumes were determined, total biovolume was obtained by summation to get total biomass of the algae sampled at each depth (g/m$^3$). The equations used are as follows:

a) Average biovolume for one 40x field of view (μm$^3$):

$$\text{avg} = (n1+n2+...+n10)/10$$

b) Average biovolume in total slide view (μm$^3$/5 mL):

$$\text{avgT}= (\text{avg} \times 0.7854)/0.00050195$$

c) Average biovolume per 1 mL (μm$^3$/mL) in 5 mL:

$$\text{AvgS}= (\text{avgT})/5$$

d) Biomass per depth (g/m$^3$): Sum of all bio-volumes

$$\text{biomass} = [\text{sum(avgS)}]/10^6$$

Phytoplankton species were identified using Freshwater Algae (Bellinger and Siege 2010). Biomass was determined by calculating the geometric biovolume (Hillebrand et al 1999)
of the cells observed. Some adaptations to biovolume calculations were made because some dimensions were not measurable. For example, centric diatoms such as *Stephanodiscus* or *Cyclotella*, were assumed to have a cell depth equal to width (Sun and Liu 2003) and therefore was measured as a sphere. Biomass was calculated and organized by each taxonomic group (Cyanobacteria, Chlorophyta, Chrysophyta, Cryptophyta, Bacillariophyta, Dinophyta) and added together to determine overall algal biomass of each sample. Percent composition of each group was also calculated, with a focus on percent cyanobacteria. This was done through the formula:

a)  \( \% \text{ Group} = \frac{\text{Biomass Group}}{\text{Total Biomass}} \times 100 \)

2.2.4 Nutrients

Soluble reactive phosphorus (SRP), total phosphorus (TP), and nitrite-nitrate (NO\(_2\)-NO\(_3\)) were analyzed on the Unity Scientific SmartChem 175-200 Discrete Chemistry Analyzer using SmartChem 200 operating software in the Nutrient Lab at Great Lakes Institute for Environmental Research at the University of Windsor. For TP, raw water was stored in glass amber bottles at 4 °C until analysis. SRP and NO\(_2\)-NO\(_3\) samples were filtered using 0.45 μm pore size EMD Millipore Nitrocellulose membrane filters in the lab immediately after field work and stored in glass amber bottles at 4°C until analysis. TP was processed using the acid digestion EPA method 365.4 1974. For this method, all samples underwent digestion where samples were treated with 1 mL of sulphuric acid: \( \text{H}_2\text{SO}_4 \) (30%) and 1 mL of persulphate solution (0.4g ammonium persulphate: \( \text{(NH}_4\text{)}_2\text{S}_2\text{O}_8 \)). These samples were put on hot plates and heated to 90-110 °C where they digested for about 2 hours. After digestion, samples were stored at 4 °C until analysis. For analysis calibration standards, field blanks, and method blanks were used to compare to sample results. SRP was processed using the EPA method 365.1, Revision 2.0 1993.
Calibration standards, field blanks, and method blanks were used to compare to sample results. NO$_2$-NO$_3^-$ was processed using EPA method 353.2, Revision 2.0 1993. For this method calibration standards, field blanks, and method blanks were used to compare to sample results.

### 2.2.5 Chlorophyll-α

Uncorrected chlorophyll-α (chl-α) concentrations were determined using the Chlorophyll extractive method (EPA method 445.0, Revision 1.2 1997). Between 100 mL and 1 L of water was filtered using 1.2 μm pore size Whatman GF/C filters depending on observable algal densities. The filters were then stored in the freezer at -20 °C in hexane cleaned aluminum foil until extraction and analysis. Chl-α was extracted in 30 mL of magnesium carbonate acetone solution (Eaton et al 1995) and chlorophyll concentrations measured on a Turner Designs Handheld Fluorometer AquaFlor probe using EPA method 445.0, Revision 1.2 1997. Uncorrected chlorophyll-α concentrations were calculated using the formula

$$C_{S,C} = C_{E,U} \times \text{extract volume (L)} \times DF$$

Where;

$C_{S,C}$= uncorrected chlorophyll-α concentration (ug/L)

$C_{E,U}$= uncorrected chlorophyll-α concentration (ug/L) in extract solution

Extract volume= volume (L) of extract prepared (0.03L)

Sample volume= volume (L) of water sample filtered

DF= dilution factor (1)
2.2.6 Wind Data

Wind speeds during and prior to sample collection (total of 60 hours) were obtained from the Environment Canada weather station at the Windsor Airport (Environment Canada 2016-2017). This station was chosen due to its proximity to the sample location and availability of historical data. Monthly data were available and downloaded where temperature, air pressure, and wind speeds were present. Only winds speeds 48 hours before sample date and 12 hours during the sample date were collected.

2.2.7 Qualitative and Statistical Analyses

Qualitative bar graphs and pie charts were used to summarize changes in algal biomass, cyanobacterial genus biomass, and % cyanobacteria composition. For algal biomass, spatial and temporal changes (Table 2.1) in algal groups, % cyanobacteria and cyanobacteria genera are presented as averages over all sample dates and depths. All are organized based on site of collection and sample year. Box and whisker plots were used to demonstrate average and distribution of wind speeds for each sample date.

All statistical tests were performed on PAleontological STatistics (PAST) software. For relationships between phytoplankton community composition and environmental parameters, canonical component analyses (CCAs) were used. They were analyzed spatial and temporally to determine if environmental parameters correlated to sites, depths, or months in both sample years. For data points relating to site they will be grouped together based on sample location. This will be the same for depths where they will be grouped based on depth sampled, and for month based on sample date. The environmental parameters tested were TP, chlorophyll-α, water temperature, $Z_{eu}$, wind speed, SRP, NO$_2$-NO$_3^-$, site, depth, and sample date. After CCAs were analyzed, PERMANOVAs using Bray-Curtis Index were used to determine if the relationships
between spatial-temporal variations and environmental parameters were significant. In total, 1,110 data points for each analysis will be analyzed in each CCA.

One-way analysis of variance (ANOVAs) Kruskal Wallis were performed for changes in abundance of cyanobacteria genera from each site in 2016 and 2017. One-way Kruskal Wallis ANOVAs were also performed for changes in biomass community composition, % community composition, cyanobacteria genus composition, and % cyanobacteria genus composition throughout sample season and during bloom conditions. All data between sites, sample dates, and depths were put together and used for Kruskal Wallis ANOVAs. After the ANOVAs were analyzed, Mann Whitney U pairwise was conducted to determine where the significant variations were. Mann Whitney U pairwise is used as a post hoc test for non-parametric data.

2.3 Results
2.3.1 RBR maestro logger Data Collection

Environmental parameters collected from the RBR maestro logger (brancker) were water temperature, pH, dissolved oxygen percent (DO2), and specific conductivity. In 2016, water temperature throughout the sample season ranged from 16.9 °C to 25.7 °C, with average seasonal temperature noted as 23.4 °C. The water column was mostly isothermal with weak temperature gradients observed during the June and October sampling period. The averages for these environmental parameters were 8.2 for pH, 95.2% for DO2, and 232.6 uS/cm for specific conductivity. In 2017, water temperature ranged from 18.8 °C to 23.9 °C, with an average 20.9 °C, DO2 had an average of 94%, specific conductivity had an average of 202.4 uS/cm, and pH had an average of 8.2.
2.3.2 Community Composition in 2016 and 2017

In 2016, the largest biomass was observed at offshore sites, Colchester Offshore and Pelee Offshore (Figure 2.3 and 2.5). For Colchester Offshore (Fig 2.3) biomass was at its peak on August 26, with densities of 12.7 g/m³, 10.6 g/m³, 15.9 g/m³, 17.8 g/m³ at 1, 3, 7 and 10 m depth respectively. Pelee Offshore (Fig 2.5) had comparable biomass during this period ranging from 10-20 g/m³. At both sites, diatoms and cyanobacteria groups were dominant with diatoms making up 60-80% of the biomass. The largest biomass at Colchester Inshore (Fig 2.2) was on June 21, with diatoms representing over 90% of the biomass. During this period, biomass was over 3 g/m³, with the largest biomass of 6.3 g/m³ observed at 7m. The rest of the sampling dates yielded algal biomasses under 3 g/m³. At the Pelee Inshore site (Fig 2.4), the largest biomass (10 g/m³) formed on September 12 at 3 m depth. In this case, 70% of the biomass was cyanobacteria. At all four-sampling sites, however, there was the common presence of oligotrophic taxa such as Chrysophyceae, Dinobryon spp. with Dinobryon at times reaching a maximum of 67% of total algal abundance.

In 2017, although a cooler year, cyanobacteria were much more dominant. The largest cyanobacteria biomass was present during the September 24-28th at Colchester Harbour. At this site, the largest biomass (18 g/m³) was observed at 3 m depth (Fig 2.6) with almost 100% of the biomass being cyanobacteria. The other sites reflected this bloom but to a lesser extent in the surface waters with total biomasses of 17.4 g/m³, 12.5 g/m³, 9.7 g/m³ and 1.1 g/m³ at Colchester Inshore, Colchester Offshore, Pelee Inshore and Pelee Offshore respectively. The largest biomass at the Colchester Offshore (Fig 2.7) was during this time period at 10 m with biomass at 16 g/m³. Pelee Inshore (Fig 2.8) had its largest biomass (12 g/m3) at 3 m and Pelee Offshore (Fig 2.9) had its largest biomass at 3 m at 12.5 g/m³. Similar to 2016, oligotrophic taxa such as Chrysophyta Dinobryon spp were commonly present. The largest biomass of Dinobryon was on June 21st.
with the majority of sites and depth having biomass over 10% with the largest coming from Pelee Offshore at 7 m where Dinobryon represented 74% of the relative abundance.

Cyanobacteria composed a greater proportion of the community composition at all stations in 2017 compared to 2016. Percent cyanobacteria was 17% higher at Colchester Inshore (Fig 2.10a), 18% higher at Colchester Offshore (Fig 2.10b) and Pelee Inshore (2.10c), and 13% higher at Pelee Offshore (2.10d) in 2017. However, there were no significant differences in composition and relative abundance between the two years (ANOVA Kruskal Wallis, p>0.05) (Table 2.2).

During bloom periods in 2016 and 2017, changes in both total biomass (Fig 2.13) and species specific relative abundance (Fig 2.14) were significant especially for diatoms (Fig 2.13 and Fig 2.14, ANOVA Kruskal Wallis p<0.0001) and cyanobacteria (Fig 2.13 and Fig 2.14, ANOVA Kruskal Wallis p<0.0001). Between these two years, average diatom biomass decreased from 4.1 g/m³ to only 0.1 g/m³, and average cyanobacteria biomass increased from 2.1 g/m³ to 10.1 g/m³. In summary, diatoms declined with respect to total algal biomass from 49.8% in 2016 to 2.4% 2017, whereas cyanophytes increased from 43.1% to 95.4%.

2.3.3 Cyanobacterial Genus Composition and Changes in % Cyanobacteria Genus Composition

In 2016 there was no cyanophyte bloom observed and diatoms tended to dominate the algal assemblage. In 2017, a cyanophyte bloom was observed during the period of September 24-28. During bloom conditions, changes in % cyanobacteria genus composition was considered insignificant based on ANOVA Kruskal Wallis test (Table 2.3 and Fig 2.11a-d). However, changes in biomass were significant (Table 2.3) especially for Microcystis spp (ANOVA Kruskal Wallis p<0.0001) and Anabaena spp (ANOVA Kruskal Wallis p<0.0001). Biomass for
Microcystis increased from 2 g/m³ in 2016 to 9.5 g/m³ in 2017 whereas Anabaena biomass increased from about 0.02 g/m³ in 2016 to 0.4 g/m³ in 2017 (Fig 2.12).

### 2.3.4 Environmental Factors and Phytoplankton Community

Site, month, and depth were analyzed for influence of environmental factors on phytoplankton community distribution. Significance was determined for each. (Table 2.4). Both sample month (Fig 2.17) and site location (Fig 2.18) were significant and highly influenced by environment factors whereas depth showed no significant environmental influence (Fig 2.19). For monthly distribution, April and June were correlated with TP, SRP, and NO₂-NO₃⁻. SRP AND NO₂-NO₃⁻ would have been highly influential chemical sources in 2017 because that is the only sample year they were collected (Fig 2.17). Physical environmental factors were shown to be more influential in the late summer where Zₑu, water temperature, and wave height were especially important in September/October distribution of the phytoplankton community (Fig 2.17). For site location, Pelee Inshore and Pelee Offshore were distinct and smaller subsections of the Colchester sites. Both Pelee sites seemed to be more influenced by TP and water temperature compared to the Colchester sites (Fig 2.18). Colchester sites seemed to also have a chemical environmental influence however it came from SRP and NO₂-NO₃⁻ (Fig 2.18). Although depth wasn’t significantly correlated to environmental parameters there is a slight difference in community. Communities with depth higher in water column seems to be slightly more distinct and smaller that communities present lower in the water column (Fig 2.19).

### 2.4 Discussion
#### 2.4.1 Community Composition

The Western Basin of Lake Erie supports temporally and spatially variable phytoplankton assemblages, making it very difficult to determine the overall health of the lake. While average
biomass was relatively similar in 2016 and 2017, there were important changes in community composition between the two years. In both years, Chrysophyta, Bacillariophyta, Cyanophytes, Chlorophyta, and Cryptophyta dominated the assemblage but with very high spatial and temporal heterogeneity. In 2016, cyanobacteria were commonly observed lower in the water column early in the sample season, and this was true at all four sample sites. This distribution supports the hypothesis that sediments are critical for the overwintering of cyanophyte propagules and that the seeding influence of algae from the sediments (Kitchens et al 2018) in the western basin can play an important role in the development of the final algal assemblage.

Other taxa commonly observed at all four sites was chrysophyte Dinobryon. This indicates an oligotrophic influence (Rawson 1956), potentially from Lake Huron via the Detroit River (Beeton 1965). Dinobryon was present early in the sample season in 2016 and was found as late as early July. Similar to 2016, Dinobryon was found early in the sample season. However, in 2017, Dinobryon cells were found later in the sample season (July 26th) at the Pelee sites, indicating that the Detroit River might have had a strong influence on the algal assemblage, and this effect lasted longer than the previous year. There appear to be many different source populations that have the potential to dominate the phytoplankton assemblage of the Western Basin of Lake Erie. Sediments can be a major source of both diatom and blue-green algal propagules whereas major tributaries are critical sources of other genera such as Dinobryon and possibly cyanophytes from the Maumee River.

Community composition stayed relatively constant during the sample season in 2016. Offshore sites (Colchester and Pelee) had larger concentrations of bacillariophytes compared to the nearshore sites. Cyanobacteria, however, maintained larger populations at the inshore sites than the offshore sites. However, in 2017, community composition showed changes throughout
the sampling season with development of a significant algal bloom September 24th to 28th. Lake Erie harmful algal blooms are generally dominated by *Microcystis* (Bridgeman et al 2013; Budd et al 2002; Brittain et al 2000) which have an optimal growth rate at temperatures between 25 and 30°C (Kruger and Eloff 1978). Although cyanophytes are commonly described as slow growing, the 2017 bloom developed very quickly. Such rapid changes in abundance can be related to growth, or to other factors such as advective input from tributaries or the ‘piling up’ of cells in specific locations by wind. Wind can determine the location of algal biomass in the western basin (Holland 1993) and may have been an influential factor determining the size and location of the 2017 bloom. It also possible, however, that this bloom may have originated from the Maumee River in Ohio, reaching the Ontario shoreline by the bloom date of September 27th.

As noted for Colchester Inshore, this bloom was confined to the upper surface layers, with little evidence of increased population densities at lower depths which would be predicted if the bloom was based on in-lake environmental conditions.

Average seasonal biomass stayed the same and was relatively consistent in both years; the average biomass was 4.53 g/m³ in 2016 and 5.63 g/m³ in 2017. This was consistent with studies such as Conroy et al 2005 which stated that average algal biomass in Lake Erie have ranged between 4-5 g/m³ during the summer months since 1996. Kane et al (2014) reported a similar biomass to Conroy et al 2005 with average of 4 g/m³. It should be noted that current management models are based more on total algal biomass and phosphorus loads, but there is little evidence the total algal biomass has changed in Lake Erie. What seems to be changing is that this biomass is becoming more and more dominated by cyanobacteria.

In 2016, the offshore sites had larger total algal biomass than nearshore sites and their assemblage was dominated by diatoms probably re-suspended from the sediments based on
many cells appearing moribund with shrinking cytoplasm. In 2017, the Pelee sites had larger algal biomass compared to the Colchester sites. However, the largest total algal biomass of the sample season came from the Colchester sites during the bloom period in September where over 90% of the biomass was cyanobacteria. Lake Erie has been described and managed as a homogenous and well mixed system (Conroy and Culver 2005) however spatial and temporal variation in biomass and community composition suggest that this may not true.

Vertical sampling was integral to this study. It is often assumed that the western basin of Lake Erie is a well-mixed system (Conroy and Culver 2005) and simple surface sampling strategies have been used to characterize the algal assemblage. While the lake is well mixed based on isothermal conditions being maintained throughout the year, the community composition and biomass of algae indicates the lake supports very high temporal and spatial heterogeneity. In 2016, for example lower depths supported larger biomasses, especially from *Dinobryon*, diatoms, and cyanobacteria, compared to surface depths especially at offshore sites. Makarewicz et al (1999) suggested that this stratification could be due to the dreissenid impact on the nearshore environment, although this has only been observed in the central and eastern basins of Lake Erie. A dreissenid effect would predict lower population at lower depths, not the elevated biomasses observed in this study. The largest biomass observed in 2017 was in the surface water at Colchester Inshore, but at the other three sites the larger concentrations of cyanobacteria were constrained lower in the water column. High vertical heterogeneity in the development of HABs in well mixed water columns is a function of many processes including differential growth rates, buoyance regulation, and source of propagules and the effects of strong mixing events. Many studies such as Bridgeman et al (2011) have used plankton tows (water column averages) to collect phytoplankton and determine algal community assemblage. While
this method might address water column heterogeneity, net tows are not effective for collecting small cell sizes. The data presented in this thesis suggests that only discrete sampling with depth can provide critical information on the factors regulating the distribution, composition and duration of HABs.

2.4.2 Harmful Algal Blooms

Community composition showed that there were two peaks in the algal assemblage during 2016. Offshore sites developed maximum biomass around August 26th while nearshore sites had a slightly delayed peak on September 12th (see also NOAA 2016). During these conditions, more than 50% of the bloom was represented by diatoms that were likely re-suspended from the sediment. In 2017, despite being a cooler year, a typical cyanophyte bloom developed. It was seen at all four sites on September 27th and the biomass was over 90% cyanobacteria characterized by potentially toxin producing species. These two sampling years clearly demonstrate that community composition and relative abundance of species can change from year to year. It also suggests that HABs can have different growth sources (IJC 2014) and changing environmental factors such as nutrient loading, wind, advective flow and temperature influencing distribution and size of HABs from year to year (IJC 2014).

Cyanobacterial biomass in the Western Basin of Lake Erie has increased since 1996 from about 1 g/m$^3$ to about 4 g/m$^3$ (Kane et al 2014). This increase in cyanophyte biomass did not result in a clear increase in the overall algal biomass of approximately which has stayed at about 6 g/m$^3$. Data from 2016 and 2017 sample years reveal dramatic differences in composition and relative biomass from a low bloom year to a large bloom year, with cyanobacteria biomass being up to seven times higher in 2017 than in 2016, where Colchester Inshore had a biomass of 2.42 g/m$^3$ September 12, 2016 and a biomass of 15.46 g/m$^3$ September 27, 2017 at 1m depth. It is
questionable if the current monitoring strategies with respect to the high temporal and spatial (vertical and horizontal) variability can reliably characterize a bloom in a manner that would resolve the relative importance of limiting factors. Even more challenging, will be demonstrating a change in the lake’s status having spent large sums on watershed remedial action programs.

2.4.3 Environmental Factors

Chlorophyll-α showed large distribution in 2016 which had a very small bloom whereas 2017 showed much lower concentrations of chlorophyll-α which showed a large bloom. As was shown in the CCA analyses both spatially and temporally, chl-α concentrations did not seem to have a large environmental influence. As suggested by other researchers, chlorophyll-α is not a good proxy for measuring and estimating phytoplankton biomass (Hillis 2017, unpublished; Holland 1993) and thus not a good indicator of lake status (or how lake status has changed as a result of remedial action efforts). Chemical factors, such as nutrients (TP, SRP, and NO₂-NO₃⁻) showed strong influence on the community especially during the spring months (April and June). SRP and NO₂-NO₃⁻ were especially influential in 2017 where there was a large HAB present. This suggests that concentrations of these two nutrients might be better predictors of cyanobacteria HABs in Lake Erie (Michalak et al 2013) than previously known. Physical factors, such as water temperature and euphotic depth, were also shown to be important especially for monthly distribution of phytoplankton communities.

2.5 Conclusion

While biomass has stayed the same since the mid-1990s with yearly averages around 4 g/m³, community composition has changed. Diatoms make up majority of the biomass in Lake Erie and are present even in bloom condition. Oligotrophic species such as Dinobryon are
present in Lake Erie well in the summer indicating that the Detroit River has an influence on the community composition. 2016 and 2017 only saw differences in composition during the HAB conditions. In 2016, the bloom was dominated by diatoms whereas the 2017 bloom was dominated by cyanobacteria, particularly *Microcystis*. This indicates that HABs can be influenced by different environmental factors and seeding/nutrients sources each year. Environmental parameters such wind patterns, water temperature, euphotic depth, and bioavailable nutrients (SRP, NO$_2$–NO$_3^-$) can influence blooms/HABs more than previously thought. While there was no difference in community composition through the water column, biomass did differ, with larger biomass at lower depths. Our current sample regimes may be neglecting composition and biomass in the lower depths and therefore neglecting important information about HABs. This study shows the need to look at whole community composition and biomass throughout the season, locations, and depths not just cyanobacteria communities in surface HABs. More frequent long-term research needs to be conducted at multiple locations and depths throughout the western basin of Lake Erie to understand if there is a pattern to changes in community composition and if this will affect the proliferation of larger and more toxic HABs.
2.6 References


Table 2.1: Sampling dates for the western basin a) 2016 sampling year b) 2017 sampling year

a) 2016

<table>
<thead>
<tr>
<th>Month</th>
<th>Sample Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>June 21, 2016</td>
</tr>
<tr>
<td>July</td>
<td>July 7, 2016</td>
</tr>
<tr>
<td></td>
<td>July 19, 2016</td>
</tr>
<tr>
<td>August</td>
<td>August 10, 2016</td>
</tr>
<tr>
<td></td>
<td>August 26, 2016</td>
</tr>
<tr>
<td>September</td>
<td>September 12, 2016</td>
</tr>
<tr>
<td>October</td>
<td>October 4, 2016</td>
</tr>
</tbody>
</table>

b) 2017

<table>
<thead>
<tr>
<th>Month</th>
<th>Sample Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>April</td>
<td>April 28, 2017</td>
</tr>
<tr>
<td>May</td>
<td>N/a</td>
</tr>
<tr>
<td>June</td>
<td>June 8, 2017</td>
</tr>
<tr>
<td></td>
<td>June 21, 2017</td>
</tr>
<tr>
<td>July</td>
<td>July 6, 2017</td>
</tr>
<tr>
<td></td>
<td>July 26, 2017</td>
</tr>
<tr>
<td>August</td>
<td>August 9, 2017</td>
</tr>
<tr>
<td></td>
<td>August 30, 2017</td>
</tr>
<tr>
<td>September</td>
<td>Sept 15, 2017</td>
</tr>
<tr>
<td></td>
<td>Sept 27, 2017</td>
</tr>
</tbody>
</table>
Table 2.2: One-way ANOVA Kruskal Wallis showing difference between % composition and biomass of each phytoplankton group throughout the sample season and bloom conditions in 2016 and 2017. *** = p<0.0001 \( n_1=1,110 \)  \( n_2=92 \)

<table>
<thead>
<tr>
<th>Comparison type</th>
<th>Difference between 2016 and 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>% composition (( n_1 ))</td>
<td>No</td>
</tr>
<tr>
<td>Biomass of all algal species (( n_1 ))</td>
<td>No</td>
</tr>
<tr>
<td>% composition during bloom (( n_2 ))</td>
<td>Yes diatoms*** cyanobacteria***</td>
</tr>
<tr>
<td>Biomass of all algal species during bloom (( n_2 ))</td>
<td>Yes diatoms*** cyanobacteria***</td>
</tr>
</tbody>
</table>

Table 2.3: One-way ANOVA Kruskal Wallis showing difference between % composition and biomass of each cyanobacteria group throughout sample season and bloom conditions in 2016 and 2017. * p<0.05, *** p<0.0001 \( n_1 = 222, n_2=92 \)

<table>
<thead>
<tr>
<th>Comparison Type</th>
<th>Difference between 2016 and 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>% cyanobacteria genus composition (( n_1 ))</td>
<td>No</td>
</tr>
<tr>
<td>Biomass of cyanobacteria genus composition (( n_1 ))</td>
<td>No</td>
</tr>
<tr>
<td>% cyanobacteria genus composition during bloom (( n_2 ))</td>
<td>No</td>
</tr>
<tr>
<td>Biomass of cyanobacteria genus composition during bloom (( n_2 ))</td>
<td>Yes Microcystis*** Anabaena***</td>
</tr>
</tbody>
</table>

Table 2.4: PERMANOVA results showing significance of CCA data results for phytoplankton percent community composition. 2016 and 2017 were analyzed together. * p<0.05 n=1,110

<table>
<thead>
<tr>
<th>Spatial-Temporal Variation</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>0.0001 *</td>
</tr>
<tr>
<td>Site</td>
<td>0.004 *</td>
</tr>
<tr>
<td>Depth</td>
<td>0.3018</td>
</tr>
</tbody>
</table>
**Figure 2.2:** Seasonal changes in phytoplankton biomass at Colchester Inshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m$^3$). Yellow box indicates average historical seasonal biomass in western Lake Erie.
Figure 2.3: Seasonal changes in phytoplankton biomass at Colchester Offshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m$^3$). Yellow box indicates average seasonal biomass in western Lake Erie.
Figure 2.4: Seasonal changes in phytoplankton biomass at Pelee Inshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication ($3g/m^3$). Yellow box indicates average seasonal biomass in western Lake Erie.
Figure 2.5: Seasonal changes in phytoplankton biomass at Pelee Offshore from June to Oct 2016. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie.
Figure 2.6: Seasonal changes in phytoplankton biomass at Colchester Inshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie.
Figure 2.7: Seasonal changes in phytoplankton biomass at Colchester Offshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie.
Figure 2.8: Seasonal changes in phytoplankton biomass at Pelee Inshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m$^3$). Yellow box indicates average seasonal biomass in western Lake Erie.
Figure 2.9: Seasonal changes in phytoplankton biomass at Pelee Offshore from June to Oct 2017. Depths are separated by black line. Red line indicates when biomass indicated eutrophication (3g/m³). Yellow box indicates average seasonal biomass in western Lake Erie.
a) Colchester Inshore

b) Colchester Offshore

c) Pelee Inshore

d) Pelee Offshore

Figure 2.10: Pie charts showing yearly changes in % cyanobacteria from 2016 and 2017 at a) Colchester Inshore; b) Colchester Offshore; c) Pelee Inshore; and d) Pelee Offshore. Chryo=Chrysophytes, Crypto=Cryptophytes, Chloro=chlorophytes, and Cyanos=cyanophytes.
Figure 2.11: Pie charts showing biomass changes in cyanobacteria community composition from 2016 to 2017 at a) Colchester Inshore; b) Colchester Offshore; c) Pelee Inshore; and d) Pelee Offshore.
**Figure 2.12:** One-way ANOVA Kruskal Wallis showing significant changes in biomass for cyanobacteria genus types during bloom conditions in 2016 and 2017. ***p<0.0001

**Figure 2.13:** One-way ANOVA Kruskal Wallis showing significant changes in biomass for whole community composition to group during bloom conditions in 2016 and 2017.
Figure 2.14: One-way ANOVA Kruskal Wallis showing significant changes in % community composition for whole community to phytoplankton group during bloom conditions in 2016 and 2017.

Figure 2.15: Box and whisker plot showing the distribution of wind speeds 60 hours before the end of the sampling dates in 2016. Average speeds were 21-Jun= 11.3km/h, 19-Jul=14.9km/h, 26-Aug= 14.6km/h, 12-Sep= 13.5km/h, and 4-Oct= 7.1km/h.
Figure 2.16: Box and whisker plot showing the distribution of wind speeds 60 hours before the end of the sampling dates in 2017. Average speeds were 8-Jun= 15.6km/h, 21-Jun=16.2km/h, 6-Jul= 8.7km/h, 26-Jul=12.3km/h, 8-Aug=9.4km/h, 30-Aug= 6.9km/h, 15-Sep=7.6km/h, 27-Sep=8.5km/h.

Figure 2.17: CCA plots for monthly distribution of phytoplankton communities in 2016 and 2017.
**Figure 2.18:** CCA plot for site location and distribution of phytoplankton communities in 2016 and 2017.

**Figure 2.19:** CCA plot for depth distribution of phytoplankton communities in 2016 and 2017.
CHAPTER THREE: GENOMIC TECHNIQUES AS A TOOL TO TRACK AND UNDERSTAND PHYTOPLANKTON DYNAMICS

3.1 Introduction

Microbial communities, especially phytoplankton communities, respond to small changes in their environments (Needham et al. 2018; Cram et al. 2014; Gilbert al 2012; Fuhrman et al 2006). These responses can occur on multiple temporal and spatial scales and can include a range of adaptations from the physiological level to ecosystem adjustments (Berdjeb et al. 2018; Fuhrman et al 2015). The multiple response scales of phytoplankton communities represent how adaptable and resilient the communities can be (Lyria Berdjeb et al. 2018), showing how important they are to ecological and biogeochemical processes (Berdjeb et al. 2018; Worden et al. 2015; Caron et al. 2009; Sherr et al. 2007). These communities and their growth are highly influenced by light and nutrient availability (Purina et al. 2018; Smayda and Reynolds 2001). Because of this, phytoplankton community composition is used as a bio-indicator to determine the water quality of an aquatic ecosystem (Sun et al. 2018; Borics et al. 2014; Lugoli et al. 2012; Conforti et al. 1995).

Globally, phytoplankton communities have been changing in response to many factors (Hallegraeff 1993). These factors include nutrient availability (Anderson et al. 2002), habitat alteration (Hallegraeff 1993), and increased temperatures or climate change (Hallegraeff 1993). Trends have been observed where phytoplankton communities shift to more problematic species, such as toxic species (Heisler et al. 2008; Glibert et al. 2005; Anderson et al. 2002; Hallegraeff 1993; Smayda 1990; Anderson 1989). Such shifts can be seen in the coastal waters of Australia (McLeod et al. 2012) and the Gulf of Mexico with dinoflagellate dominated “red tides” (Walsh et al. 2006); and Taihu (Paerl et al. 2011), Lake Victoria (Paerl and Huisman 2008), Lake Mead
(Beaver et al 2018), and Lake Erie (Paerl and Huisman 2008) with Microcystis dominated cyanobacteria HABs.

With increasing bloom occurrence across the globe (Paerl and Huisman 2008), quick and more economical genetic analyses of phytoplankton communities are becoming more common for early detection and hazard assessment (Visco et al 2015; Bohmann et al 2014; Baird and Hajibabaei 2012; Taberlet et al 2012). For prokaryotic and eukaryotic analyses of phytoplankton communities, taxonomic studies are done using next generation sequencing (NGS) (Visco et al 2015). This analysis uses targeted rRNA sequences (Needham et al 2018) which allow for a more robust analysis of the community as rRNA sequences are present through the three life domains (Needham et al 2018; Needham and Fuhrman 2016; Parada et al 2016). It also uses longer sequences (Hajibabaei et al 2011) which can provide identification for nearly 90% of species (Hajibabaei et al 2011; Meusnier et al 2008; Hajibabaei et al 2007; Hajibabaei et al 2006) even with degraded DNA in environmental samples (Meusnier et al 2008).

In Norway, NGS has been used to characterize the microbial community during phytoplankton bloom conditions. In a study conducted by Parulekar et al (2017), the microbial community was analyzed in Lake Akersvannet, a eutrophic lake in Norway, from samples collected from July to August in 2013. The 16S rRNA V1-V3 and V3-V4 regions were targeted and community composition determined. It was found that about 31% and 42% of the bacterial community consisted of cyanobacteria species. More than 97% of the cyanobacteria species present were Aphanizomenon (95.8%) and Microcystis (2.1%). Lake Akersvannet is very similar to the Western Basin of Lake Erie as it is shallow (at 6m deep and max depth at 13m) and is dominated by cyanobacteria HABs during the summer months. In Lake Erie few studies have focused on the phytoplankton community composition within spatial and vertical scales using
next generation sequencing. One study in Lake Erie by Chaudhary et al (2009) looked at bacterial diversity along the coastline of the Erie Marsh preserve in Michigan. This study found that about 47% of the bacterial community was made up of cyanobacteria species. These species present were primarily *Limnothrix*, *Synechocystis*, and *Planktothrix*-like species. Many genetic studies performed in western Lake Erie focused on diversity of gene expression such as those by Harke et al (2016) who concluded that the prokaryotic community was dominated by *Microcystis* spp. It was observed that more genes associated with toxin production in *Microcystis* were expressed the further the samples were collected from the Maumee River. For other toxin producing cyanobacteria, such as *Planktothrix* and *Anabaena*, genes were expressed more in nutrient rich environments especially in the nearshore regions of western Lake Erie. While this study and studies similar to this provide insight into the dynamics of the prokaryotic component of the algal assemblage, it leaves out a large component of the phytoplankton community including eukaryotic taxa. With this part of community composition not being presented, potential shifts in composition and relative abundance associated with HABs are not quantified. This study aims to fill that knowledge gap as well as look at potential factors regulating the composition and relative abundance in the algal assemblage on multiple spatial and temporal scales.

The objective of this chapter is to determine how physical and chemical processes influence phytoplankton community composition and HAB formation. Furthermore, we aim to determine if metabarcoding through next generation sequencing can be used as an early warning tool for HABs and to determine what factors influence HAB growth by answering the following questions:
a) Do we see presence of cyanobacterial DNA, especially *Microcystis* and *Anabaena*, early in the sample season?

b) Are there differences in community composition based on site location, depth, and sample date?

c) What are the most influential growth factors for HAB/cyanobacteria formation?

d) Do genomic approaches regarding community composition and HAB formation provide appropriate data to determine the viability of meta-barcoding as an early response tracking tool?

### 3.2 Methods

#### 3.2.1 Sampling Protocol

Sampling occurred at four sites located in the western basin (*Colchester Inshore: N41°58.801’, W82°56.183’; Pelee Inshore: N41°50.290’, W82°40.268’, Colchester Offshore: N41°51.396’, W82°59.137’; Pelee Offshore: N41°50.290’, W82°39.837’*) at four different depth (1, 3, 7, and 10 m) (Fig 2.1). Water samples were collected monthly from June to October in 2016, and bi-monthly from June to end of September in 2017. Water samples were collected and stored in glass amber bottles until lab analysis of chlorophyll-α and nutrients (SRP, TP, NO₂ and NO₃) as described in section 2.2. Water samples collected for DNA analysis were stored in 1L glass amber bottles, kept cold in a 4 °C refrigerator and filtered within 24 hours of collection.

#### 3.2.2 Water Column Profiles

The RBR *maestro* logger (Brancker) was used to collect environmental parameters such as temperature throughout the water column from surface to *Z*ₘₐₓ. LI-250A light meter attached to a LI-193 spherical quantum sensor was used to collect irradiance from surface to *Z*ₘₐₓ. This
data was then used to calculate light attenuation ($\varepsilon_{\text{par}}$) and euphotic depth ($Z_{\text{eu}}$) using formulas stated in section 2.2 equations a) through c).

### 3.2.3 Nutrients
SRP, TP, and NO$_2$ and NO$_3^-$ concentrations were determined using EPA methods 365.4 1974, 365.1 Revision 2.0 1993, and 353.2 Revision 2.0 1993 as stated in section 2.2

### 3.2.4 Chlorophyll-α Extraction
Uncorrected chlorophyll-α concentrations were analyzed and determined using the Chlorophyll extractive method (EPA method 445.0, Revision 1.2 1997) as per section 2.2

### 3.2.5 DNA Extraction, PCRs and Next Generation Sequencing
Phytoplankton DNA was extracted using sucrose lysis buffer method (Shakraki et al 2018) on the TECAN Freedom EVO DNA extraction robot. In total, DNA was extracted from 219 samples in 2016 and 336 samples in 2017. The V5 and V6 regions (Table 3.1) of the bacterial 16S variable SSU rRNA gene were targeted and PCR amplified. For first PCR, 25 uL reactions were amplified in 2.5 μL of 10x Taq Buffer, 3.5 μL of 20mM MgSO$_4$, 0.5 μL of 10 mM dNTP, 0.5 μL of 10 mM forward primer V5F, 0.5 μL of 10 mM reverse primer V6R, 0.1 μL of 1KU Taq Polymerase, 1 μL of DNA, and 16.4 μL of DI water. For the first thermal cycle, PCR was amplified at 94 °C for 2 mins, followed by 28 cycles of 94 °C at 30 s, 55 °C at 30 s, 72 °C at 60 s, and finally ending at 72 °C for 8 mins. The first PCR products then underwent purification using Agencourt AMPure XP beads. The second PCR was then conducted using 2.5 μL of Taq Buffer, 3.5 μL of MgSO$_4$, 0.5 μL of dNTP, 0.5 μL of 10mM UniB, 0.1 μL of Taq Polymerase, 7.4 μL of DI water, 10 μL of purified PCR product, and 0.5 μL of 10mM UniA Barcodes. For the second PCR thermal cycle 25 uL reactions underwent one cycle at 95 °C for 2
mins, then eight cycles at 95 °C for 30 s, 60 °C at 30 s, 72 °C at 1 min, and finally one cycle at 72 °C for 5 mins. PCR product was normalized and pooled together using band intensity data obtained using GEL DOC XR+ under Trans UV conditions. The product was then purified using Epoch Life Science Nucleic Acid Purification kit. The purified product was prepared for next generation sequencing (NGS) and run on Agilent 2100 Bioanalyzer with Agilent High Sensitivity DNA Assay kit. After bio-analysis was completed, samples were diluted to a concentration of around 60 pmol/μL. Pooled samples were sequenced on 318 v2 chips using Ion Torrent

For eukaryotic phytoplankton, the V9 region (Table 3.1) of eukaryotic 18S variable SSU rRNA gene was targeted and PCR amplified. 18S rRNA gene PCR products underwent the same process as 16S rRNA gene products. However, different primers were used for PCR amplification during the first PCR process. Forward primer, V9-1391F and reverse primer, V9-1774-1797R were used (Table 3.1).

3.2.6 16S rRNA gene and 18S rRNA gene Taxonomy and Bioinformatics

After NGS runs were completed, raw data were de-multiplexed, quality filtered, and trimmed of the adaptor, barcode and primer sequences using Quantitative Insights into Microbial Ecology (QIIME) 1.8 for both 16S rRNA gene and 18S rRNA gene targets. (Caporaso et al 2010) Operational taxonomic units (OTUs) were organized based on a similarity sequence of 97% or greater among reads and assigned their taxonomic name using SILVA database with 80% or greater confidence level (Edgar 2010). Focus was placed on bacterial and eukaryotic DNA associated with chloroplasts, such that Cyanobacteria, Bacillariophyta, Chlorophyta, Cryptophyta, Chrysophyta were all targeted in the taxonomic classification.
Abundance and rarity of the OTUs were organized based on presence in the community such that abundant was defined as any OTU where it was more than 2000 reads for each OTU and rare was less than 2000 reads for each OTU. (Logares et al 2014). This number was chosen due to rarefaction curves that showed that the majority of OTUs had over 2000 reads and plateaued therefore sample analyses would not be affected.

### 3.2.7 Qualitative and Statistical Analyses

After rRNA reads and OTUs counts were obtained, the data were organized as follows. To determine phytoplankton community composition, only prokaryotic (16S rRNA) and eukaryotic (18S rRNA) OTUs with presence of chlorophyll were analyzed for rRNA reads and percent frequency in the whole community. This was found to be *Cyanobacteria, Bacillariophyta, Chlorophyta, Cryptophyta, Euglenophyta, and Chrysophyta*. The reads for each group were then calculated and organized in a stacked percent graph to determine their abundance in the phytoplankton community. This was performed for both 2016 and 2017. For cyanobacterial growth, both *Microcystis* and *Anabaena* were analyzed. *Microcystis* and *Anabaena* were chosen because they made up large frequencies in the bacterial community for both 2016 and 2017 and are cyanobacteria species associated with HABs. These qualitative line graphs were made to show changes in frequencies in both species across site, depth, and sample dates in 2016 and 2017.

Non-metric multidimensional scaling (nMDS) using the Bray-Curtis index on two dimensions was used to show the phytoplankton microbial community composition and how they grouped based on the spatial and temporal parameters. This analysis uses lack of fit or ‘stress’ between variables to determine dissimilarities (Paliy and Shankar 2016). It then repositioned the variables to minimize stress (Paliy and Shankar 2016; Dugard et al 2014) and
show data in simple clusters. In this study, nMDS was used to look at clusters and their relationship with spatial and temporal variables. Stress values were found through Shepard plot analysis. Canonical correlation analysis (CCAs) was conducted for both 16S and 18S rRNA data using percent community composition (frequency in community) to determine which environmental factors were the most influential for community structure. CCAs were used due to their ability to show influential environmental parameters over large datasets that do not follow linear patterns but rather unimodal patterns (Palmer 1993). For 16S rRNA gene data, all cyanobacterial species found in samples were analyzed using environmental parameters; nutrients (TP, NO$_2^-$, NO$_3^-$ and SRP), Z$_{eu}$, temperature of water, chl-$\alpha$ concentrations, air temperature, wave height, and wind speed. For 18S rRNA gene, all species classified as phytoplankton were analyzed using same environmental parameter as 16S rRNA gene for analysis. Four different factors (depth, site, month, and season) were analyzed for CCAs in 16S and 18S to determine if environmental factors were more influential for temporal and/or spatial variations. After CCAs were completed, one-way permutational multivariate analysis of variance (PERMANOVAs) were performed to determine significance of clusters and changes in community structure. Bray-Curtis was used for clustering of similarity index and sequential Bonferroni p-value corrected was used to determine significance.

3.3 Results
3.3.1 rRNA gene reads and OTU counts

After next generation sequencing of both 16S and 18S rRNA genes for both 2016 and 2017, total reads were determined. In 2016, 16S rRNA genes had 8,760,250 reads and 18S rRNA had 7,160,256 reads. In 2017, 16S rRNA genes had 9,044,476 reads and 18S rRNA genes had 8,430,371. After data processing and omission of OTUs under 2000 reads, final reads were
3.3.2 Community Composition and Change in Harmful Cyanobacteria Frequency

Community composition was determined after rRNA reads and OTU counts were obtained. Annual averaged microbial phytoplankton community composition was determined in 2016 and 2017. 16S rRNA (prokaryotic) and 18S rRNA (eukaryotic) phytoplankton composition were analyzed differently and are therefore shown separately 16S rRNA and 18S rRNA both show changes in abundance from 2016 to 2017. For 16S rRNA in 2016 cyanobacteria had 61,200 out of 2,445,487 reads and for 2017 it had 111,002 out of 4,234,562 reads. The graphs were organized into the 4 most abundant cyanobacteria classes for both sample years. The cyanobacteria classes that had increased reads from 2016 to 2017 were Oscillatoriophycideae (Microcystis is species in this class) which saw an increase of 12.8% to 17.3% in the community and Synechococcolophycideae which increased from 12.1% to 15.5% (Fig 3.1). The two classes that saw decreases were Nostocophycideae (Anabaena is a species in this class) which saw a decrease from 5.2% to 2.6% and “Other” which saw a decrease of 69.9% to 64.5% (Fig 3.1). “Other” was a classification due to QIIME software which does not have the ability to identify
those OTUs yet. For 18S rRNA in 2016 phytoplankton had 1,273,026 out of 4,395,641 reads and 2017 had 268,446 out of 3,026,120 reads. 18S rRNA classes that increased were *Chlorophyceae* which increased from 15.2% to 27.0% and *Chrysophyceae* which increased from 9.3% to 29.3% (Fig 3.2). Other classes decreased in 2017. Such decreases were found with *Bacillariophyceae* (diatoms) which decreased from 42.9% to 31.2 %, *Cryptophyceae* which decreased 16.3% to 0.6%, and Dinoflagellates which decreased 16.4% to 12.0% (Fig 3.2).

Tracking *Microcystis* and *Anabaena* frequencies within the whole bacterial community were conducted for both sample years. In 2016, frequency in the community for *Microcystis* reached its peak of 1.2% on July 19th at Pelee Offshore at a depth of 10m. At Colchester Inshore in 2016, frequencies were below 0.4% for all depths and samples dates (Fig 3.3). The largest frequency was 0.35% at 7 m for June 21st. For Colchester Offshore (Fig 3.3), the largest frequency was 0.62% at 1m on June 21st. For Pelee Inshore (Fig 3.3), the largest peak was 0.57% at 3m on July 19th. Pelee Offshore (Fig 3.3) saw the largest frequency of *Microcystis* in 2016 it was 1.2% at 10 m on July 19th. In 2017, frequencies of *Microcystis* were much larger with peaks reaching nearly 5% of the whole cyanobacterial community. For Colchester Inshore (Fig 3.4) the *Microcystis* bloom was observed on September 27th with the largest frequency of 4.58% at 1m. Colchester Offshore (Fig 3.4) also had large frequencies on September 27th with the largest at 2.91% at 7 m. At Pelee Inshore (Fig 3.4) had its peak at 2.96% on September 27th. For Pelee Offshore (Fig 3.4), peak frequency was 1.66% at 7m on September 27th.

For *Anabaena* in 2016, its peak was at 0.89% on July 19th at Pelee Offshore (Fig 3.3) at a depth of 3m. Colchester Inshore (Fig 3.3) showed extremely low presence of *Anabaena*, with its largest frequency at 0.02% of the community at a depth of 3m on June 21st. Colchester Offshore (Fig 3.3) had its peak frequency of 0.83% at 10m on July 19th. At Pelee Inshore (Fig
3.3), *Anabaena* frequencies were on average larger than at the Colchester sites. Peak frequency was about 0.45% on August 26th at 1m. Pelee Offshore (Fig 3.3) had the largest *Anabaena* frequency of all four sites which was 0.89% on July 19th at 3m. In 2017, the largest *Anabaena* frequency observed was 0.51% at 1m at Pelee Offshore (Fig 3.4) on July 26th. Colchester Inshore (Fig 3.4) and Offshore (Fig 3.4) showed very similar frequency patterns to 2016 where frequencies were extremely low. The largest frequency observed at Colchester Inshore for was 0.37% at 1 m September 27th. For Colchester Offshore, the largest frequency observed was 0.33% at 7m September 27th. Pelee Inshore (Fig 3.4) had peak frequencies at 0.43% at 1m on July 26th. For Pelee Offshore (Fig 3.4) peak frequency was 0.51% on July 26th at 1m.

*Microcystis* growth was different in 2016 compared to 2017. In 2016, a decline over the sample season was observed. This was especially evident at the Pelee sites. In 2017, growth was apparent and slow compared to 2016. While frequencies were higher overall in 2017, large (peak) abundances weren’t observed until September 27th due to the Colchester Harbour bloom during this sample date. The only sites that did not follow this pattern were Pelee Inshore and Offshore at surface depths. For *Anabaena*, frequencies were higher than *Microcystis* in 2016 at Colchester and Pelee Offshore but lower at the inshore sites. In 2017, *Anabaena* frequencies were higher than *Microcystis* during July sample dates at all sites and they were much lower at these sites during the Colchester Harbour bloom. In both years, Pelee sites consistently had larger frequencies of *Anabaena* compared to Colchester sites throughout the sample season.

### 3.3.3 Spatial-Temporal Factors and Clustering

Chemical and physical parameters were examined to determine their relative importance in determining the above spatial and temporal patterns observed in the algal community. nMDS results are displayed for both spatial (depth and location) and temporal (sample date and season)
to determine the relative importance of drivers determining composition and relative abundance of microbial communities. 16S 2016 and 18S 2017 Shepard stress scores were all under 0.15 which indicated good data. 16S 2017 and 18S 2016 had stress scores between 0.2 and 0.3 which indicated it’s acceptable but unlikely to represent the data well (Table 3.3).

For 16S in 2016 temporal nMDS plots (Fig 3.5) showed strong clusters indicating strong temporal changes in the microbial community. This was not the same for 16S in 2017, however, where there was more overlap (Fig 3.6) indicating that community composition was more similar based on sample date. For 16S spatial variation, both vertically and horizontally, the communities were all very similar for both years. For 18S 2016 depth, site, and season had considerable overlap whereas month (Fig 3.7) formed distinct clusters. All plots for 18S 2017 had weak clusters which indicated that the spatial and temporal (Fig 3.8) factors all had communities that were very similar (Spatial graphs in Appendix A).

3.3.4 Microbial Clustering
16S Clusters for 2016 and 2017

For 16S OTU counts, 1220 cyanobacterial OTUs were identified in 2016 and 3995 OTUs in 2017. Four CCA plots were used to determine how community structure responds to environmental factors. The four CCA plots clustered depth (1m, 3m, 7m, 10m), site (Colchester Inshore, Colchester Offshore, Pelee Inshore, Pelee Offshore), month (June, July, August, September, October), and season (spring, summer, autumn). PERMANOVAs determined if the cluster patterns from the CCA plots were significant.

In 2016, clusters were not easily defined and there was overlap among the depths (Fig 3.9). Community structure based on depth was more with associated chl-α concentrations, euphotic depth, and water temperature but these relationships were not significant (Table 3.2).
CCA plots for month showed stronger clusters especially for June which was more associated with TP loading more than other environmental parameters (Appendix A). This relationship was significant (p<0.0001) (Table 3.2). CCA clusters for season (Fig 3.11) were also significant (p<0.0001) (Table 3.2). Spring was highly associated with TP loadings, where summer was associated with chl-α, euphotic depth (Z_{eu}) and water temperature. CCA clusters for site (Fig 3.10) were found to be significant (p<0.0001) (Table 3.2) and community structure was more related at the Colchester sites compared to the Pelee sites and favoured chl-α, water temperature, Z_{eu}, and wave height. In 2016, 16S communities were influenced by multiple factors. TP was highly associated with June/Spring, and depth and site where highly influenced by chl-a concentrations, euphotic depth (Z_{eu}), and water temperature.

In 2017, community structure and distribution were found to be significant (p<0.0188) (Table 3.2) with depth (Fig 3.12) and the community favoured air and water temperature, wind speed, and wave height. Community structure was found to be more significant (p<0.0001) (Table 3.2) with month especially for June which was strongly associated with SRP loadings (Appendix A), whereas September was associated with wind speed, wave height, and water temperature. For season CCA plots (Fig 3.14), spring was highly associated with SRP and NO₂-NO₃⁻ loadings and autumn was associated with chl-α and Z_{eu}, and these clusters were found to be significant (p<0.0001) (Table 3.2). Site location (Fig 3.13) clusters had a large overlap and do not have a significant relationship (Table 3.2). In 2017, 16S communities had different environmental influence than 2016. June/Spring was highly influenced by with SRP and NO₂-NO₃⁻ loadings, and depth and autumnal composition was highly influenced by water temperature, wind speed, and wave height.
18S Clusters for 2016 and 2017

For 18S OTU counts, 127 eukaryotic phytoplankton OTUs were identified in 2016 and 96 OTUs in 2017. The same methods as 16S rRNA were used for 18S rRNA CCA plot structures and PERMANOVA analyses to determine clustering and significance of environmental factors. In 2016, all depths (Fig 3.15) clustered in the same pattern and were found to have poor associations with all environmental factors, therefore considered insignificant (Table 3.2). Monthly clustering (Appendix A) was found to be significant (p<0.0001) (Table 3.2) and June was very associated with TP loadings. Other months that had environmental associations were both September and October which clustered towards water temperature. Seasonal clustering (Fig 3.17) was found to be similar to month. Spring was strongly associated with TP loadings. Seasonal clusters were found to be significant (p<0.0001) (Table 3.2). Site location (Fig 3.16) was found to be significant (p<0.0001) (Table 3.2). Both Colchester and Pelee Inshore clusters were found to have associations with wave height and chl-α, whereas Pelee Inshore and Offshore clusters were not strongly associated with any environmental parameters. In 2016, 18S communities had no influences with depth, however monthly/seasonal communities were highly influenced by TP in June, and nearshore sites were influenced by wave height and chl-α.

In 2017, depth (Fig 3.18) was found to be significant (p<0.0005) (Table 3.2) and were associated with Z_{eu}, water and air temperature, wave height, and wind speed. Monthly distribution (Appendix A) was found to be significant (p<0.0001) (Table 3.2). June was highly associated with SRP and NO2-NO3-, and September was associated with Z_{eu} and water temperature. Seasonal distribution (Fig 3.20) was also found to be significant (p<0.0001) (Table 3.2). Spring was highly associated with SRP and NO2-NO3-. Site location (Fig 3.19) especially Pelee sites were more associated with water temperature and Z_{eu} compared to the Colchester
sites. These site location clusters were found to be significant (p<0.0001) (Table 3.2). In 2017, 18S communities were influenced by similar environmental factors to 16S communities in 2017. Depth was highly influenced by euphotic depth (Zeu), water and air temperature, wave height and speed. Monthly/seasonal communities were highly influenced by SRP and NO₂-NO₃⁻ loadings especially in June. Pelee sites were more influenced by water temperature and euphotic depth than Colchester sites.

3.4 Discussion
3.4.1 Harmful Algal Growth

Similar to microscopic observations in Chapter Two, the genomic approach confirmed the abundance of cyanobacteria was much higher in 2017. Cyanobacterial abundancies were nearly 20 times higher during the bloom period in 2017, especially true for the Colchester Inshore site (Fig 3.2 and Fig 3.3). Interestingly, 2016 and 2017 showed growth patterns that are not usually indicative of cyanobacterial or Microcystis growth in Lake Erie. Microcystis is generally most abundant during summer months and visible at the surface during late summer periods when waters are calmer (Reynolds 1973; Reynolds 1971). In 2016, this pattern was not represented. At all sites, abundance was generally below 1% and peak abundance periods were observed in July. Anabaena also exhibited a similar pattern to this. This could be due to these two cyanobacterial genera sharing similar spatial-temporal niches (Tromas et al 2018). In 2017, a different abundance pattern was observed. Both Microcystis and Anabaena frequencies were quite low. However, Microcystis experienced a large spike in abundance on September 27th during the Colchester Harbour Bloom. Anabaena abundance, however, stayed low during this time period. This may have been because of Microcystis’ ability to dominate during bloom conditions and during nitrogen limiting conditions especially from dissolved nitrogen (Tromas et
al 2018, Dai et al 2009; Moisander et al 2009) which was observed during this sample period. Surprisingly, Colchester Offshore and the Pelee sites had larger abundances of *Microcystis* lower in the water column compared to the surface. This could have been because of *Microcystis*’ ability to sink in the water column during autumn months (Verspagen et al 2004; Visser et al 1995) and that these sites did not experience the surface bloom to the same extent as Colchester Inshore.

### 3.4.2 Spatial and Temporal Variation

nMDS was used to help better understand the similarities/dissimilarities between the spatial (site location and depth) and temporal (sample date and season) variation in the phytoplankton microbial communities. Temporal variation for 16S (prokaryotes) in 2016 had very distinct clusters. This indicated that community composition was very different for each sample period. This could be because certain cyanobacterial communities are more prevalent during different times of the year. For example, bloom forming cyanobacteria like *Microcystis*, are more likely to be present in the community in the late summer and autumnal periods than the spring (Reynolds 1971). The spatial factors for 16S 2016 had less distinct clusters especially with depth. Lake Erie is considered a well-mixed system (Makarewicz et al 1999) and in 2016 did not experience a cyanobacterial bloom which could have contributed to a highly similar community composition throughout the water column. Site, while having clusters that overlapped, had smaller distinct communities especially with the nearshore environments like Colchester and Pelee Inshore. Nearshore environments generally have different microbial communities compared to offshore environments (Leon et al 2011) due to environmental factors such as the difference in nutrient availability as described by the nearshore shunt (Heckey et al 2004) and water temperature (Burns et al 2005, McCormick and Fahnenstiel 1999).
18S 2016 phytoplankton (eukaryotic) community composition was very similar for both spatial and temporal factors except month which showed distinct community structures. Interestingly seasonal variation in community composition had more overlap, however spring time had a relatively distinct community composition. Peaks in diatom biomass are generally seen in during the spring/June period (Lashaway and Carrick 2010; Solovieva et al 2005; Montagnes and Franklin 2001) where this relationship was clearly seen. In 2017 the 18S community composition was very similar for all spatial-temporal factors with very few distinct clusters. Like 2016, spring/June had a smaller, distinct cluster compared to the rest of the sample period.

3.4.3 Environmental Factors influencing Microbial Community Composition

CCAs were performed to help identify the most influential environmental factors for phytoplankton community composition and structure. Monthly and seasonal variation in prokaryotic (16S) and eukaryotic (18S) communities in 2016 were well defined and controlled by TP during the June/Spring period. However, in 2017, these same temporal factors were influenced more by NO$_2$-NO$_3$- and SRP. In more recent years, research has looked at how bioavailable nutrients, such as NO$_2$-NO$_3$- and SRP, control phytoplankton communities and influence cyanobacteria HABs by increasing bloom size (Conroy et al 2005; Mitrovic et al 2001) and toxicity (Beversdorf et al 2013; Downing et al 2005).

After the influence of the nutrients in the spring/June period, 16S and 18S communities in both sample years tended to be influenced by water temperature, wind speed, wave height, and $Z_{eu}$. These physical factors are known to affect community composition in Lake Erie (Millie et al 2009) causing different phytoplankton groups to become more prominent during different parts of the growing period (Millie et al 2009; Lean et al 1983). These physical factors are also
becoming increasingly more influential factors for bloom forming cyanobacteria (Millie et al 2009). *Microcystis* likes warmer water temperature (Paerl and Huisman 2008) and deeper euphotic depths- light penetration (Paerl 1996). This data suggests that the community also was influenced by wind events because they allow for horizontal and vertical transport of HABs (Wu and Kong 2009). These graphs indicate that phytoplankton communities are favouring these conditions and suggesting that phytoplankton communities could become more cyanobacterial dominant in the future.

3.5 Conclusion

In this chapter, NGS helped determine phytoplankton community composition, factors regulating composition, and growth patterns of two dominant cyanobacteria genera in a no bloom and bloom year. Observations were similar to that observed using a microscopic approach, but unlike traditional techniques, NGS allowed for larger and quicker analyses of phytoplankton communities. NGS therefore is an ideal tool to use for tracking changes in composition and HAB growth. Temporal variability tended to dominate over spatial heterogeneity (horizontal and vertical patchiness) using the genomic approach. Nutrients influenced composition in algal assemblages during the spring/June period, whereas more physical factors influenced growth and distribution patterns later in the year, where temperature, euphotic depth, and wind events regulated these community structures. However, more research and sample season should be completed before such assumption can be made about community composition and shifts within the community.

Growth trends were also observed in both sample years to determine if genomic techniques could be used as a potential tracking tool for early response to HAB growth. 2016 had a no/low bloom year and data indicated a strange growing pattern with small peaks observed in
early summer which is not generally observed for cyanobacteria growth. However, 2017, showed a more common growth pattern where larger cyanobacteria peaks were observed in late summer and early autumn. These two years not enough to determine if genomic techniques are a good early warning tool especially since there was no observable pattern between the two years. Research and cyanobacteria tracking needs to be on going to determine if this tool could be appropriate for bloom tracking and see if there are observable patterns and shifts in community composition.
3.6 References:


Table 3.1: Primers and primer sequences used for DNA PCR amplification for 16S rRNA and 18S rRNA in 2016 and 2017

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>V5F</td>
<td>acctgcctgccgATTAGATACCCNGGTAG</td>
</tr>
<tr>
<td>V6R</td>
<td>acgccacgcacgCGACAGCCCATGACANCCT</td>
</tr>
<tr>
<td>UniA (Barcodes)</td>
<td>CCATCTCATCCCTGCGTGTCTGGACTCAGXXXXXXXXXGATaccgtgcctgccg</td>
</tr>
<tr>
<td>UniB</td>
<td>CCTCTCTATGCGCAGTGGTGATacgcacgacg</td>
</tr>
<tr>
<td>V9-1391F</td>
<td>acctgcctgccgGTACACACCGCCTGATcctgccg</td>
</tr>
<tr>
<td>V9-1774-179R</td>
<td>acgccacgcacgCGCTGATCCTTCTGACACCTTCACCTAC</td>
</tr>
</tbody>
</table>

Table 3.2: PERMANOVA results showing significance of CCA data results for 16S and 18S rRNA gene data for 2016 and 2017. * p<0.05

<table>
<thead>
<tr>
<th></th>
<th>depth</th>
<th>site</th>
<th>month</th>
<th>season</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>16S</td>
<td>0.8533</td>
<td>0.0001*</td>
<td>0.0001*</td>
</tr>
<tr>
<td></td>
<td>18S</td>
<td>0.2175</td>
<td>0.0001*</td>
<td>0.0001*</td>
</tr>
<tr>
<td>2017</td>
<td>16S</td>
<td>0.0188*</td>
<td>0.2715</td>
<td>0.0001*</td>
</tr>
<tr>
<td></td>
<td>18S</td>
<td>0.0005*</td>
<td>0.0001*</td>
<td>0.0001*</td>
</tr>
</tbody>
</table>
Table 3.3: Shepard plot values of stress for nMDS analyses using Bray-Curtis similarity index.

<table>
<thead>
<tr>
<th></th>
<th>2016</th>
<th>16S</th>
<th>18S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth</td>
<td></td>
<td>0.0961</td>
<td>0.2229</td>
</tr>
<tr>
<td>Month</td>
<td></td>
<td>0.09907</td>
<td>0.2215</td>
</tr>
<tr>
<td>Seasonal</td>
<td></td>
<td>0.09907</td>
<td>0.2242</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td>0.09907</td>
<td>0.2201</td>
</tr>
<tr>
<td></td>
<td>2017</td>
<td>16S</td>
<td>18S</td>
</tr>
<tr>
<td>Depth</td>
<td></td>
<td>0.2844</td>
<td>0.1281</td>
</tr>
<tr>
<td>Month</td>
<td></td>
<td>0.2844</td>
<td>0.1281</td>
</tr>
<tr>
<td>Seasonal</td>
<td></td>
<td>0.2844</td>
<td>0.1281</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td>0.2844</td>
<td>0.1281</td>
</tr>
</tbody>
</table>
Figure 3.1: Phytoplankton community composition determined by 16S rRNA reads in 2016 and 2017. Note: Other indicates cyanobacteria species that could not be identified to genus/species from their OTUs.

Figure 3.2: Phytoplankton community composition determined by 18S rRNA reads in 2016 and 2017.
Figure 3.3: Relative abundance of *Microcystis* and *Anabaena* through the sample season at all sample locations in 2016. Depths shown are 1m, 3m, 7m, and 10m. *Microcystis* is orange, and *Anabaena* is grey.
Figure 3.4: Relative abundance of *Microcystis* and *Anabaena* through the sample season at all sample locations in 2017. Depths shown are 1m, 3m, 7m, and 10m. *Microcystis* is orange, and *Anabaena* is grey.
Figure 3.5: nMDS plot of 16S rRNA 2016 Month.

Figure 3.6: nMDS plot of 16S rRNA 2017 Month.
Figure 3.7: nMDS plots of 18S rRNA 2016 Month.

Figure 3.8: nMDS plots of 18S rRNA 2017 Month.
Figure 3.9: 16S rRNA CCA for depth in 2016.

Figure 3.10: 16S rRNA CCA for site in 2016.
Figure 3.11: 16S rRNA CCA for season in 2016.

Figure 3.12: 16S rRNA CCA for depth in 2017.
Figure 3.13: 16S rRNA CCA for site in 2017.

Figure 3.14: 16S rRNA CCA for season in 2017.
**Figure 3.15:** 18S rRNA CCA for depth in 2016.

**Figure 3.16:** 18S rRNA CCA for site in 2016.
Figure 3.17: 18S rRNA CCA for season in 2016.

Figure 3.18: 18S rRNA CCA for depth in 2017.
Figure 3.19: 18S rRNA CCA for site in 2017.

Figure 3.20: 18S rRNA CCA for season in 2017.
CHAPTER FOUR: GENERAL DISCUSSION

4.1 Discussion

This dissertation examined and determined the phytoplankton community composition and relative abundance in the western basin of Lake Erie during 2016 and 2017 using traditional microscopy and genomic techniques. It identified how a complex interaction of physical and chemical factors regulate the phytoplankton assemblage focusing with specific reference to harmful algal blooms (HABs). HABs have become increasingly more problematic in Lake Erie (Watson et al. 2016) with drinking water concerns such as the Toledo Water Crisis in 2014 (Steffen et al. 2016) and beach fouling (Watson et al. 2016; Higgins et al. 2008; Carmichael et al. 1997). While biomass stayed relatively the same (6 g/m$^3$) since the mid-1990s (Conroy et al. 2005a), there has been an increase in HABs abundance over this period. This is the first study in Lake Erie that will quantify spatial and temporal changes for the entire phytoplankton community since the early 1990s. Furthermore, this is the first study in Lake Erie that used two different techniques, microscopic and genomic, to determine phytoplankton community composition and factors regulating composition.

In Chapter Two, phytoplankton community composition and relative abundance was determined through traditional microscopic techniques. Analyses of annual biomass showed that it ranged between 4-6 g/m$^3$ in 2016 and 2017 which is consistent with Conroy et al. (2005a) who stated that biomass has had a similar biomass range since the mid-1990s. Therefore, there is little evidence that annual biomass has increased since the mid-1990s. As most management models are based on predicting total algal biomass, the use of these models to manage cyanobacterial blooms should be viewed with caution. This research further concluded that different environmental factors, physical and chemical as well as propagule sources (lake, sediment, tributary) can influence community composition and relative abundance from year to year. 2016
had a small bloom that was diatom dominant, whereas 2017 had a large bloom that was cyanobacteria (*Microcystis*) dominant, although both years achieved a similar total average biomass of algae. It was suggested that these two years had different seeding sources with storm resuspension playing a key role in 2016 and tributary inputs dominating in 2017. Specifically, bioavailable nutrients such as NO$_2$-NO$_3^-$ and SRP were determined to be very influential in HAB growth during the 2017 sample year. Recent research has found that these nutrients are more influential in HAB formation (Conroy et al 2005b; Mitrovic et al 2001) and toxin production (Beversdorf et al 2013; Downing et al 2005) than previously suggested. With the assumed change in chemical influence from TP to more bioavailable nutrients and the legislated 40% reduction in P loading in Lake Erie, nutrient loading may be a more difficult predictor of phytoplankton biomass and HAB distribution (Bertani et al 2016). This chapter also found that wind, water temperature, and euphotic depth were influential. HABs are associated with warmer temperatures (Paerl and Huisman 2008) and a deeper euphotic depth (Paerl 1996) however HAB species such as *Microcystis* prefer calm conditions ((Reynolds 1973; Reynolds 1971). Persistent wind can transport HABs horizontally (Wu and Kong 2009), potentially increasing spatial coverage of the HABs and developing areas of very high abundances as a result of cells piling up in a specific area. Interestingly, chlorophyll-$\alpha$ was not a good predictor of overall biomass. Hillis (2017 *unpublished*) and Holland (1993) stated that it has become an increasingly less accurate proxy for phytoplankton biomass especially in Lake Erie.

In Chapter Three, phytoplankton community composition and relative abundance was determined using genomic techniques. Like Chapter Two, this chapter concluded phytoplankton community composition was regulated by physical and chemical drivers. This is the first study in Lake Erie that examined both 16S (prokaryotic) and 18S (eukaryotic) rRNA genes as a means to
determine the composition of the algal assemblage. It was concluded that community composition had low spatial variation, opposite to what was observed microscopically. Using the genomic approach suggested that the algal community composition had high temporal variability. Environmental factors, such as nutrient concentrations and loadings, have more of an influence on community composition and distribution during the spring/June whereas physical factors such as temperature, euphotic depth, and wind are more influential during the late summer months.

These two data chapters have helped better understand the phytoplankton community and identified key environmental factors that regulate phytoplankton composition and relative abundance in Lake Erie. Lake Erie has a very diverse phytoplankton community with highly eutrophic and highly oligotrophic species commonly observed in the same water sample. For both sampling seasons, Diatoms, Cyanobacteria, Cryptophytes, Chrysophytes, Chlorophytes, and Dinoflagellates were commonly present throughout the basin. Overall diatoms contributed the largest portion of biomass and composition in 2016 and 2017, which is revealed using both traditional and genomic analyses. Cyanobacteria had the second largest portion of biomass and composition overall, however, during the 2017 Colchester Harbour bloom it was over 90% of the biomass and composition. While both 2016 and 2017, had similar phytoplankton composition, structure of that community and relative abundances were different. Such variation in abundances between the years could make prediction of community composition difficult in the future especially with changing nutrient management plans and the unpredictability of climate change. Nutrients helped regulate community composition in spring (June) whereas physical factors were more influential for community composition later in the summer and fall (September/October). We also found that while there was slight variation in depth distribution of
community composition, we did see a more variation in community composition based on sample location with a slight east to west gradient. This suggests that the western basin is not as well-mixed as previously thought.

4.2: Final Conclusion and Future Work

Phytoplankton compositional analyses are important for continued understanding of community dynamics in Lake Erie especially in regard to HABs and shifts towards HABs. With recent policies, such as the 40% reduction of TP loading, trying to mitigate their effect on the Lake Erie through nutrient control (Annex 4 Objectives and Targets Task Team 2015) might be ambitious. This thesis suggests that other physical and chemical factors play important roles in regulating HAB growth. SRP and NO$_2$-NO$_3^-$ were shown to be highly influential in community regulation. Other factors such as wind, euphotic depth, and temperature were also important as physical drivers as they can determine propagule source and in lake production potential at critical times of the year.

Community composition was very diverse and there was presence of oligotrophic and eutrophic phytoplankton throughout the season, even during the same sample period. Oligotrophic *Dinobryon* was found as late as July 26$^{th}$ in 2017 and eutrophic cyanobacteria *Microcystis* was found as early as April 28$^{th}$ in 2017. While cyanobacteria regularly dominate the community composition during bloom periods that was proven to not always be the case. Diatoms were also present during HABs, even in low quantities during 2017, however throughout the sample season they consistently made up the majority of the biomass. Biomass was relatively within common ranges of 4-6 g/m$^3$, however large spikes in biomass especially during the Colchester Harbour Bloom were infrequent but expected because of biomass being distributed by the wind.
Both the microscopic and genomic approaches revealed that there was a compositional shift of diatom dominant communities in 2016 to a diatom-cyanobacteria dominant assemblage in 2017. They concluded that similar chemical and physical factors were regulating community composition and relative abundance. The genomic approach, however, demonstrated temporal changes in composition were more significant than spatial changes, although this was not as strongly observed using microscopic techniques. Therefore, it can be suggested that genomic analyses of Lake Erie phytoplankton communities might be a viable option for an early response and tracking tool for HAB growth and formation.

Future work in the western basin of Lake Erie should include more research conducted throughout the basin. Based on high inter-annual differences observed in this study, long term data sets must be collected as they will provide more insight into compositional shifts in phytoplankton communities and the effects of stressors acting in different time scale such as climate change. As well, further testing of genomic techniques such as NGS, meta-transcriptomics, and even quantitative real time polymerase chain reaction (qrt-PCR) may provide more insight into the dynamics of genera associated with HABs. These techniques will be very beneficial not only to researchers but managers that need to provide knowledge to policy makers about HABs and their growth patterns.

I think that this research is not only important to scientists, but also to policy makers and the general public. Ecosystems and human health are at risk with the increased presence of HABs in the western basin of Lake Erie. More knowledge and understanding about phytoplankton and factors regulating their composition is critical to prevent further problems with HABs. Therefore, I recommend governments, policy makers, scientists, and the public work
together to learn more and do more to prevent HABs from being a harbinger of the death of Lake Erie.
4.3: References
Annex 4 Objectives and Targets Task Team. Phosphorus loading targets for Lake Erie. 2015. 70.


APPENDIX A

Initial and Final Reads and OTU counts for 16S rRNA and 18S rRNA in 2016 and 2017.

<table>
<thead>
<tr>
<th></th>
<th>Initial Reads</th>
<th>Final Reads</th>
<th>Initial OTUS</th>
<th>Final OTUS (over 2000 Hits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S 2016</td>
<td>8760250</td>
<td>2445487</td>
<td>64470</td>
<td>402</td>
</tr>
<tr>
<td>16S 2017</td>
<td>9044476</td>
<td>4234562</td>
<td>189055</td>
<td>366</td>
</tr>
<tr>
<td>18S 2016</td>
<td>7160256</td>
<td>4320013</td>
<td>4734</td>
<td>54</td>
</tr>
<tr>
<td>18S 2017</td>
<td>8430371</td>
<td>3026120</td>
<td>2954</td>
<td>36</td>
</tr>
</tbody>
</table>

CCA plots Month

a) 16S rRNA CCA for month in 2016.
b) 16S rRNA CCA for month in 2017.

c) 18S rRNA CCA for month in 2016.
d) 18S rRNA CCA for month in 2017.

**nMDS plots**

a) 16S rRNA 2016 Depth.
b) 16S rRNA 2016 Season.

c) 16S rRNA 2016 Site.
d) 16S rRNA 2017 Depth.

e) 16S rRNA 2017 Season.
f) 16S rRNA 2017 Site.

g) 18S rRNA 2016 Depth.
h) 18S rRNA 2016 Season

i) 18S rRNA 2016 Site.
j) 18S rRNA 2017 Depth.

k) 18S rRNA 2017 Season.
1) 18S rRNA 2017 Site.
VITA AUCTORIS

NAME: Jessica M. Owen

PLACE OF BIRTH: Leamington, ON

YEAR OF BIRTH: 1993

EDUCATION: General Amherst High School, Amherstburg, ON, 2011

University of Windsor, B.Sc., Windsor, ON, 2015

University of Windsor, M.Sc., Windsor, ON, 2019