Identifying Metabolic Indicators of Phytoplankton and Comparing the Temporal Changes in Algal Community Composition in Two Lake Ontario Areas of Concern

by

David McNabney

A thesis submitted to the
School of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the degree of

Master of Science in Applied Bioscience

Faculty of Science

University of Ontario Institute of Technology (Ontario Tech University)

Oshawa, Ontario, Canada

December 2021

THESIS EXAMINATION INFORMATION
Submitted by: **David McNabney**

Master of Science in Applied Bioscience

Identifying Metabolic Indicators of Phytoplankton and Comparing the Temporal Changes in Algal Community Composition in Two Lake Ontario Areas of Concern

An oral defense of this thesis took place on December 2, 2021, in front of the following examining committee:

Examining Committee:

Chair of Examining Committee	Dr. Janice Strap
Research Supervisor	Dr. Denina Simmons
Examining Committee Member	Dr. Andrea Kirkwood
Thesis Examiner	Dr. Robert Bailey, Ontario Tech University

The above committee determined that the thesis is acceptable in form and content and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate during an oral examination. A signed copy of the Certificate of Approval is available from the School of Graduate and Postdoctoral Studies.

ABSTRACT

Freshwater algal blooms transition into harmful algal blooms (HABs) when cyanobacteria produce toxins that impact ecosystem health. The point at which algal blooms transition into HABs is not fully understood, but I believe that the transition can be predicted by combining molecular and community-level information. This study consisted of ten weeks of sampling in 2020 from August to October in Hamilton Harbour and the Bay of Quinte, two Lake Ontario areas of concern (AOCs). Phytoplankton genus composition was obtained via microscopy and water quality was assessed. Metabolomics were performed using liquid chromatography tandem mass spectroscopy (LC-MS/MS) to identify changes in metabolites over time. The results from this study identified taurine, serine, and glycine in large amounts in both locations and that phytoplankton community composition is primarily affected by nitrates, total phosphorus, and turbidity. These findings will improve our understanding of these AOCs as well as the behaviour of phytoplankton overall.

Keywords: phytoplankton; algal blooms; Lake Ontario; Area of Concern; metabolomics.

AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Ontario Institute of Technology (Ontario Tech University) to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Ontario Institute of Technology (Ontario Tech University) to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

_____ DAVID McNABNEY

STATEMENT OF CONTRIBUTIONS

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 have used standard referencing practices to acknowledge ideas, research techniques, or other materials that belong to others. Furthermore, I hereby certify that I am the sole source of the creative works and/or inventive knowledge described in this thesis.

ACKNOWLEDGEMENTS

I want to start by acknowledging that the work presented in this thesis was completed while the world was dealing with the adjustments of the novel COVID-19 pandemic. My research proposal needed to be heavily altered and my objectives needed to shift from focusing on fish and harmful algal blooms to the phytoplankton within the blooms as a result of resource and collaboration limitations. It was difficult to get into laboratories on campus during lockdowns and there were a lot of unknowns during the formation of this thesis that impeded my ability to plan ahead and be prepared. While I understand that I have not been as negatively impacted as some others during the pandemic, I still believe that it is important to recognize what was happening during the completion of this thesis.

With respect to getting the samples for my thesis, I would like to give a huge thank you to my peers who woke up very early all throughout the summer of 2020 to drive one to two hours with me to sample. I know the weather did not always cooperate for us, but I hope you all enjoyed the experience and learned a thing or two about field sampling. Tina Flaherty, Shreya Jain, Chase Tudor, Simon Pollard, and Emily Hassal, thank you for your help—I really could not have collected the data I needed for my thesis without you!

I also want to thank my family for always being there for me during the last two and a half years. I would not have been able to finish this degree without your support...even if you had no idea what I was researching or why I was looking at lake water; I know you knew what I was doing!

Next, I would like to thank my partner, Josh McMeekin, for his constant support and encouragement during my studies. Thank you for always being there to listen to my problems, offer solutions, and celebrate my successes with me.

Finally, I would like to extend my deep gratitude to my supervisor, Dr. Denina Simmons, for giving me the opportunity to be a part of her lab. I have greatly developed my technical and critical thinking skills as a scientist under her guidance. It is from her wealth of knowledge, enthusiasm for her field of research, and ability to problem solve just about anything that I have gained a huge amount of respect for what it takes to do research. I would not have been able to collect my samples, analyze the data, and write this thesis without her guidance and mentorship. I am truly thankful to have been a part of Nina's first batch of graduate students to proudly hold a Master's of Science degree from her supervision!

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

CHAPTER 3

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 3 (continued)

CHAPTER 3 (continued)

LIST OF APPENDICES

APPENDIX A

APPENDIX D

APPENDIX D (continued)

LIST OF ABBREVIATIONS

HAB	Harmful Algal Bloom
LC-MS/MS	Liquid-Chromatography tandem Mass Spectrometry
WQ	Water Quality
TP	Total Phosphorus
BF	Bayfront Harbour
CB	Carroll's Bay
YC	Yacht Club
GL	Glenora
HB	Hay Bay
PH	Picton Harbour
NP	Napanee
WC	Wilton Creek
PAH	Polyaromatic Hydrocarbon
PAE	Phthalic Acid Ethers
BETX	Benzene, Toluene, Ethylbenzene, and Xylene
MC	Microcystin
ATX-a	Anatoxin-a
VFDF	Very Fast Death Factor
DO	Dissolved Oxygen
HH	Hamilton Harbour
BQ	Bay of Quinte
AOC	Area of Concern
PCB	Polychlorinated Biphenyl
IJC	International Joint Commission
RAP	Remedial Action Plan
BUI	Beneficial Use Impairment
MC-LR	Microcystin Leucine-Arginine
MC-RR	Microcystin Arginine-Arginine
MC-YR	Microcystin Tyrosine-Arginine
NI	Neare Island
GC	Golf Course
COND	Conductivity
TDS	Total Dissolved Solids
TURB	Turbidity
HPLC	High Performance Liquid Chromatography
PFTE	Polytetrafluoroethylene
Q-TOF	Quadrupole-Time of Flight
PPE	Polyphenylene ether
PCDL	Personal Compound Database and Library

LIST OF ABBREVIATIONS (continued)

CSV	Comma Separated Value
PCA	Principal Components Analysis
RDA	Redundancy Analysis
AABA	Alpha-Aminobutyric Acid
GABA	Gamma-Aminobutyric Acid
ISS	Internal Suspended Solids
GBT	Glycine Betaine

CHAPTER 1: INTRODUCTION

1.1 Freshwater Phytoplankton

Phytoplankton are very important components of any freshwater ecosystem. These microscopic organisms are great sources of primary production and are a part of the base of the trophic pyramid in the aquatic food web. Primary production is the incorporation of organic carbon, such as carbon dioxide, into the food web of an ecosystem in the form of complex organic compounds like glucose through the photosynthesis of plants and plant-like organisms (Van Damme *et al.*, 2010). Due to their abundance and ability to efficiently photosynthesize, freshwater phytoplankton greatly contribute to the energy available in the aquatic food web. Freshwater phytoplankton serve as an important source of energy for higher trophic level organisms that prey on them, such as zooplankton and other macroinvertebrates (Yokoyama *et al.*, 2005). For example, other small aquatic zooplankton graze on phytoplankton, transferring the energy produced by the photosynthetic phytoplankton up a trophic level. This energy transfer is perpetuated when larger zooplankton, insects, and fish prey upon the phytoplankton-consuming zooplankton, followed by other fish, aquatic birds, and mammals preying on those organisms. Without the primary production of freshwater phytoplankton, an aquatic ecosystem could collapse.

Phytoplankton are autotrophs, specifically photoautotrophs, meaning that they produce their own energy through photosynthesis. However, some phytoplankton can be mixotrophs, such as some members of the dinoflagellate genus *Ceratium sp.* that can produce their own energy through photosynthesis or by consuming other plankton (Hansen & Nielsen, 1997; Smalley *et al.*, 2003). Moreover, there are also some phytoplankton that

are obligate heterotrophs, meaning that they do not photosynthesize and instead must obtain organic matter from their environment. This type of energy acquisition has been observed in the colourless algae *Prototheca zopfii*, which are unable to photosynthesize and instead metabolise a variety of fatty acids like propionic acid, iso-butyric acid, and oleic acid, in addition to a number of carbohydrates, such as glucose, fructose, and mannose for energy (Barker, 1935). The different energy acquisition methods of phytoplankton allow them to fill a variety of ecological niches in many different environments. Phytoplankton are versatile and diverse organisms, which is reflected by the high diversity of several phyla of phytoplankton.

The many different types of phytoplankton that can be found in freshwater environments include: green algae, chrysophytes, cryptophytes, diatoms, dinoflagellates, euglenids, and cyanobacteria.

Green algae, or Chlorophytes, are a diverse group of algae very common in aquatic environments. Chlorophytes can be both unicellular and multicellular and are most commonly found in freshwater environments, but can also be found in marine environments and on land (Lee, 2018). Chlorophytes are important primary producers due to their efficient means of photosynthesis. Although, green algae are aptly named due to the green chlorophyll pigments they possess, not all appear to be green. For example, *Chlamydomonas nivalis* is a type of green algae that resides in snow and contains a high amount of a red carotenoid that gives it a red colour (Williams *et al.*, 2003). Some green algae, such as *Chlorella vulgaris* have been used to develop biosensors that can detect

pesticides (Gossett *et al.*, 2019). Overall, green algae are very diverse group of algae important for making up a large proportion of phytoplankton biomass in freshwater environments. They are strong contributors of primary production, diversity, and ability to adapt to many different types of waters.

Chrysophytes are a common group of phytoplankton from the Chrysophyta phylum that are often found in freshwater lakes. These unicellular phytoplankton are usually motile through the use of a flagellum, and have a characteristic golden-brown colouration due to pigments necessary for photosynthesis at specific wavelengths of light. One marine chrysophyte, *Prymnesium parvum*, is known for causing fish kills (Guo *et al.*, 1996). While typically found in marine or estuarine waters, *P. parvum* can also be found in lakes with high mineral content (Rodgers *et al.*, 2010). The number of different types of toxins that cause such harm to fish, known as prymnesins, is not currently known (Manning & La Claire, 2010). Moreover, the effects of prymnesins in humans is not fully understood. Nonetheless, chrysophytes are integral members in the aquatic food web that contribute to energy transfer and biodiversity.

Cryptophytes are another common group of freshwater phytoplankton from the phylum Cryptophyta. In addition to freshwater, these phytoplankton can also be found in marine and brackish waters, as well. Cryptophytes are also motile phytoplankton thanks to two unequal flagella and a unique pocket-like structure. This group of phytoplankton are known for their unique extrusomes, also known as ejectosomes, which can be found in the pocket-like structures of this group of algae. These ejectosomes are organelles that eject their

contents when a cryptophyte is stressed from either mechanical, chemical, or light stress. This ejection of contents propels the cell away from the source of the stress, serving as a quick departure. The pollution-tolerant species *Chilomonas paramecium*, for example, expresses this behaviour upon exposure to copper (Abraham-Peskir, 1998). Although cryptophytes are not known to be toxic, important feeding relationships are thought to exist between cryptophytes and toxin-producing dinoflagellates, such as those from the *Dinophysis sp.* genus (Nishitani *et al.*, 2005). Otherwise, cryptophytes contribute to biodiversity and serve as important prey for higher trophic levels.

Diatoms are a unique type of algae found in both fresh and marine waters, as well as in soil and on damp surfaces. Members of the Bacillariophyta phylum, diatoms are believed to be the most species rich group of algae in the world with possibly over 260 genera and more than 100,000 species (Mann, 1999; Round, 1991). Diatoms are surrounded by a cell wall, like other algae, however diatom cell walls are made of hydrated silicon dioxide, also known as silica. This type of cell wall, called a frustule, protects the organism and allows for diatoms to grow in very unique and geometric shapes and sizes (Friedrichs *et al.*, 2013; Round *et al.*, 1990). Specifically, diatoms grow in two varieties, pennate and centric. Pennate diatoms are longer in shape, sometimes described as oblong, and propel themselves by secreting mucous. Centric diatoms, on the other hand, are symmetrical around a central point and are usually circular in shape. Centric diatoms are often not motile and will free-float near the surface of waters. Diatoms can be found in large numbers, contributing to bloom formations in arctic marine regions among other locations (Dore *et al.*, 2008). Diatoms can be used to infer nutrient concentrations and therefore assess

environmental conditions. Common water quality (WQ) metrics, such as total phosphorus (TP) and turbidity, have been shown to correlate with diatoms in the presence of agricultural runoff (Pan *et al.*, 1996). Additionally, since diatoms have silica-based frustules they can be preserved fairly well in sediments. Researchers have used these characteristics to examine sediments in Lake Ontario over time and suggest that pollution abatement efforts have returned nutrient and heavy metal concentrations to nearly pre-human impact levels (Pillsbury *et al.*, 2021). Thus, diatoms are important members of any aquatic ecosystem and can be very useful for environmental health monitoring efforts.

Dinoflagellates from the phylum Dinophyta are another type of phytoplankton that can be found in both freshwater and marine environments. These phytoplankton are unique because they come in many different unique shapes and sizes, partially because of their protective outer armour, known as a thecate. Plates made of cellulose, called theca, serve as the protection in thecate dinoflagellates (Lau *et al.*, 2007). Dinoflagellates that do not have theca surrounding them are known as athecate, however some athecate dinoflagellates have been reclassified to thecate upon the discovery of thin, difficult to observe theca (Carty & Parrow, 2015). Dinoflagellates are very important members of any aquatic ecosystem because they can grow to be quite large, serving as important biomass for other trophic levels, and they are also significant primary producers. However, not all dinoflagellates are autotrophic as different species can be mixotrophs or heterotrophs. One heterotrophic dinoflagellate in particular, *Noctiluca scintillans,* is a marine alga that are capable of bioluminescence and emit a cyan-coloured light at night when they are jostled by waves in warmer waters and consume other phytoplankton (Buskey, *et al.*, 1992).

Marine dinoflagellates are also notorious for causing toxic blooms, and some species, such as *Cochlodinium sp.* are responsible for the red tides in places like Korea (Park *et al.*, 2013). While many marine species of dinoflagellates are known to produce toxins, in freshwater environments there is only one species of dinoflagellates known to be able to produce toxins. A toxin known as glenodinine was isolated in the 1960s from Lake Sagami, Japan by Hashimoto *et al.* (1968) from the dinoflagellate *Peridinium polonicum*. This toxin is not known to be harmful to humans, but has been observed to harm and kill fish. Given their limited ability to produce toxins and their roles as significant sources of biomass and primary production, freshwater dinoflagellates are integral components of any freshwater aquatic ecosystem that pose a low-risk to humans.

Euglenids from the Euglenophyta phylum are a type of phytoplankton commonly used in laboratory studies. These phytoplankton have flagella to move around, and some even have a rudimentary eye spot. There are some euglenids that are capable of producing toxins, however these toxins and their impact on the environment are not well-studied. Euglenids can be found in both marine and freshwater ecosystems and are useful as bioindicators as their presence in an aquatic system can indicate the presence of a metal or nutrient. For example, *Euglena viridis* and *Euglena sanguinea* are two common bioindicators of degraded WQ, such as degradation caused by high levels of nitrogen in freshwater lakes (Amengual-Morroet *et al.*, 2011; Sen *et al.*, 2013). Moreover, assessing the DNA of certain *Euglena sp.* can serve as early warning signs of exposure to organic pollutants. Li *et al.* (2014) describe how breaks in DNA identified from a comet assay could be used as a bioindicator to exposure from polyaromatic hydrocarbons (PAHs), phthalic acid esters

(PAEs), and benzene, toluene, ethylbenzene, and xylene (BETXs) in *E. gracilis*. These characteristics of euglenids make them highly suitable for biomonitoring studies, and are favoured by researchers around the globe. Euglenids in their natural habitats, however, contribute greatly to primary production and the transfer of energy up the trophic pyramid.

Cyanobacteria are different from all other algae because they are prokaryotes and not eukaryotes. Also known as blue-green algae due to their colouration, cyanobacteria are very significant primary producers and are very efficient at producing oxygen through their cellular processes. Cyanobacteria are also thought to have been the driving force for oxygenating the earth some 3-billion years ago and are thought to also be the precursor to the modern-day chloroplast organelle found in many other photosynthetic organisms (Schirrmeister *et al.*, 2016; Xiong *et al.*, 2016; Huisman *et al.*, 2018). Cyanobacteria are great at growing rapidly because they can assimilate nitrogen from a variety of different nitrogen-containing compounds, like ammonium, nitrate, nitrite, urea, and atmospheric nitrogen (N_2) (Flores & Herrero, 2005). This ability makes cyanobacteria very versatile and capable of existing in both marine and freshwater environments, as well as some soils. However, cyanobacteria can grow so rapidly and efficiently that they can quickly degrade aesthetics of natural landscapes, impair drinking quality of water, and give off unpleasant odours (Wang *et al.*, 2019). Moreover, cyanobacteria can produce certain kinds of compounds, known as cyanotoxins, that are toxic to a wide variety of organisms. Cyanobacteria produce cyanotoxins for a variety of different reasons, including: allelopathy, defence mechanisms against grazing, and assisting with general physiological functions (Holland & Kinnear, 2013). Several genera of cyanobacteria contain species that

produce cyanotoxins, namely: *Dolichospermum sp.* (formerly *Anabaena sp.*), *Aphanizomenon sp.*, *Microcystis sp.*, and *Nodularia sp.* Cyanotoxins come in a variety of different chemical structures, including: cyclic peptides, alkaloids, amino acids, and lipopolysaccharides. A colonial cyanobacteria, *Microcystis aeruginosa*, is responsible for producing microcystins (MCs), a class of cyclic peptide cyanotoxins. These hepatotoxic molecules consist of a number of different cyclic peptides and their toxicity depends on which amino acids are present in their structure. For instance, MC Leucine-Arginine (MC-LR) is the most toxic form of microcystin, however other almost-as-toxic forms of microcystin can be derived by replacing the leucine for another arginine (MC-RR) or tyrosine (MC-YR) (Pichardo *et al.*, 2007). These hepatotoxins inhibit protein phosphatases in a number of organisms, resulting in the severe impairment of a variety of biological systems or even death (Xiong *et al.*, 2016). Another toxic molecule produced by *Dolichospermum sp*., Anatoxin-a (ATX-a), is a bicyclic amine alkaloid that is acutely neurotoxic upon exposure. ATX-a is also known as Very Fast Death Factor (VFDF) since its exposure to animals can lead to rapid impairment or death (Quiblier *et al.*, 2016; Wood *et al.*, 2020). The structures of each of these cyanotoxins can be found in Figure 1.1. The aforementioned cyanotoxins have a range of effects, from generalized irritants, to hepatotoxic, endotoxic, and neurotoxic effects that can impair or kill aquatic life, as well as humans and other mammals. Currently, it is not clear why there are cyanotoxin-producing (toxigenic) and non-cyanotoxin-producing (non-toxigenic) species of cyanobacteria and why non-toxigenic species might not take advantage of the benefits of cyanotoxin production.

Like all other life on earth, freshwater phytoplankton require carbon, nitrogen, and phosphorus to grow. Carbon can be found in excess in aquatic ecosystems, both in aqueous forms in the water and in gaseous forms in the atmosphere, so nitrogen and phosphorus are often the nutrients limiting the growth and proliferation of freshwater phytoplankton. In fact, when nitrogen or phosphorus are in excess, this can lead to the exponential growth of freshwater phytoplankton, creating an algal bloom. While algal blooms are common and relatively harmless, these can become destructive if cyanobacteria grow exponentially and begin to produce cyanotoxins under the right conditions, leading to what is known as a harmful algal bloom (HAB). HABs consist of cyanobacteria, other algae, bacteria and other microorganisms that grow in large, buoyant surface aggregations and are considered harmful because of the cyanotoxins produced by cyanobacteria.

Freshwater HABs typically form when three conditions are met: when the air and waters are warm, such as in late summer and early fall; when the water is still and undisturbed from waves and wind; and when nutrients are in excess, such as nitrogen and phosphorus from surrounding land-use runoff and other point sources (Paerl & Scott, 2010; Ho & Michalak, 2015). When these conditions are met, cyanobacteria and other phytoplankton use up the excess nutrients to grow exponentially, resulting in unsightly masses of slimy, blue-green mats of algae and other aquatic life. However, after a period of time the excess nutrients will eventually be depleted and exponential growth will come to a plateau before declining. During the plateau, the microorganisms in the HAB begin to die and senesce, sinking to the bottom of the water column in the process. During senescence, benthic bacteria decompose the formerly prolific phytoplankton, using up the available dissolved

oxygen (DO) in the water in the process. The depletion of DO in the water leads to the formation of zones of hypoxia, where the levels of DO can become too low to sustain aerobic respiration and metabolic activities as a result (Dyble *et al.*, 2006). The zones of hypoxia can move throughout a body of water with currents, choking out fish and other aquatic life of available oxygen necessary to sustain life (Zhou *et al.*, 2015).

Not only are the health of fish and other aquatic life affected by HABs, but so are the health of terrestrial animals. HABs have been responsible for the deaths of numerous dogs and livestock upon the ingestion of water originating from a HAB (Puschner *et al.*, 2008; Dreher *et al.*, 2019) HABs can also cause dermatitis in humans, as well as a host of respiratory ailments (Vranješ & Jovanović, 2011; Codd *et al.*, 2020). Microcystins have been detected in the nasal cavities of residents nearby frequent HABs in Florida by Schaefer *et al.* (2020), however the long-term effects of these exposures are unknown at this time. Furthermore, the presence of HABs in aquatic ecosystems can lead to the closure of commercial fisheries, tourism, and recreational spaces, which often results in millions of dollars in economic loss (Itakura & Imani, 2014). Each of the aforementioned environmental conditions necessary for HAB development can be met in some of the coastal regions of the Great Lakes in Ontario, Canada. For example, HABs have been forming annually since the 1970s in Hamilton Harbour (HH) and the Bay of Quinte (BQ), two sheltered coastal Areas of Concern (AOCs) that have high amounts of urban and agricultural land-use, respectively (Kim *et al.*, 2016; Yauk & Quinn, 1996).

1.2 Hamilton Harbour Area of Concern

HH is found on the western end of Lake Ontario. The harbour is surrounded by a highly urbanized landscape with two major cities immediately along the shore of the harbour: the City of Hamilton and the City of Burlington (Figure 1.2). At the east end of HH is sandbar, a narrow stretch of land known as the Beach Strip, that separates the 2,150-ha embayment from Lake Ontario, effectively sheltering it from the larger waves of the lake (Hall *et al.*, 2006). HH drains a 270,908-ha watershed (Figure 1.3), with three major streams feeding into it, including Grindstone, Red Hill, and Spencer (Figure 1.2). To the west of HH is a 240-ha shallow marsh known as Cootes Paradise where the mouth of Spencer Creek meets the harbour. Cootes Paradise is an important part of the local landscape as it is recognized nationally as an important fish nursery. The shore to the south of HH is heavily industrialized and home to two steel mills (Sofowote *et al.*, 2008). This industrialized area has been a major source of pollution to the harbour, historically. Remediation efforts have been made but the effects of the pollution of heavy metals and toxic compounds can still be seen and all of these compounds are still present in varying amounts. There is a remediation project currently in progress at a site known as Randle Reef, which is home to the largest contaminated site on the Canadian side of the Great Lakes. The sediment is so heavily contaminated that the most cost-effective and environmentally-friendly solution to the contamination is to encase the entire site in cement (Milani & Grapentine, 2016).

Since the nineteenth century, HH has been subjected to a great deal of anthropogenic pollution and cultural eutrophication, and it is because of these stressors that the health of the Harbour has been so poor. For example, waste products from industries including heavy

metals like steel, iron, and copper, as well as harmful chemicals including polychlorinated biphenyls (PCBs) and PAHs have been released into the embayment untreated since industrialization began in the mid-nineteenth century (HHRAP, 1992; Sofowote *et al.*, 2008). These harmful waste products have specifically impaired the health of fish and wildlife in the Harbour, even causing some species to disappear altogether (Bowlby & Hoyle, 2016). Non-industrial forms of pollution in HH include: agricultural nutrient runoff, road salt runoff, untreated sewage discharge from secondary wastewater treatment plants, and stormwater overflows in events of high rainfall (Hall *et al.*, 2006). Large inputs of excessive nutrients into the embayment have resulted in HABs that occur almost every year in the late summer. HABs occur in HH because conditions several conditions are met that are conducive to the growth of cyanobacteria, such as urban nutrient runoff, the Beach Strip that shelters the embayment from the wind and waves of Lake Ontario, and warm summer temperatures (Perri *et al.*, 2015). For example, the summertime hypoxia events in HH have occurred nearly every year for the past several decades with varying degrees of severity. The oxygen-poor bottom waters caused by the zones of hypoxia during cyanobacteria senescence in the Harbour are detrimental to some fish species that reside in the colder, deeper waters, and it has been reported that some cold-water fish species, such as Cisco, can no longer reside in these depths due to the lack of oxygen (Bowlby & Hoyle, 2016). The loss of Cisco among other fish species has negative impacts on the Harbour, such as food web instability and loss of biodiversity.

In 1985, due to a combination of the previously mentioned issues, HH was identified by the International Joint Commission (IJC) as one of the most degraded bodies of water in

the Great Lakes (IJC, 1987). Two years later in 1987, HH was designated as one of forty-two AOCs in the Great Lakes (IJC, 1987). A Remedial Action Plan (RAP) was created in 1992 to improve the environmental health of this area with a series of beneficial use impairments (BUIs) to gauge the health of the area. As of 2018, ten of the fourteen BUIs are either impaired or require further assessment and are detailed in Table 1.1 (Government of Canada, 2021; HHRAP, 2018). All BUIs need to not impaired in order to de-list the harbour as an AOC.

1.3 Bay of Quinte Area of Concern

BQ is a 64-km long, narrow bay on the northern shore of Lake Ontario, approximately 150-km east of Toronto. BQ has a characteristic "Z" shape to its landscape and is 25,400-ha in area. Although the Bay is mainly surrounded by natural and agricultural landscapes, several towns and cities can be found along the shore of the Bay, including: Picton to the far southeast, Deseronto and Napanee to the northeast, Belleville to the north, and Trenton to the west. Several smaller hamlets and villages exist around BQ, including one hamlet just north east of Picton, Glenora, which is where a ferry crosses daily (Figure 1.4). The embayment serves as a drainage area for a 722,557-ha watershed where several rivers including the Trent, the Moira, the Salmon, and the Napanee River can be found (Figure 1.5).

BQ has historically received several sources of excess nutrients and waste products. For example, the landscape surrounding BQ has rich soils that are suitable for farming and over time thousands of hectares of land have been converted to agriculture use, taking advantage

of the fertile land (Wile, 1975). However, as a result of the large amounts of farming BQ has been negatively impacted by excessive nutrient runoff, which has contributed to the formation of HABs in the area. Seasonal crops farmed in large-scale around BQ include: corn, wheat, sunflowers, and soybeans. In addition to those crops, there are also a vast number of vineyards that grow various types of grapes and other fruits for the production of artisanal wines and other beverages. A large proportion of the farmers in BQ have been fertilizing the aforementioned crops with fertilizers containing nitrogen and phosphorus to enhance their yield (BQRAP, 2018). As a result, phosphorus has become a large problem in BQ, contributing to the development of HABs annually (BQRAP, 2018). In fact, at one point in time the phosphorus loads in BQ were measured to be 215-kg of phosphorus per day, but this has since been reduced to 15-kg per day (Government of Canada, 2017). While the use of phosphorus and nitrogen in farming is not a problem, the issue is that excess fertilizer is unable to remain in the soil or be assimilated by the crops is washed away during rainfall events and are deposited into the waters of BQ.

In addition to excessive fertilizer application, nutrients have historically been deposited into the Bay by conventional tillage farming in at least 27,000-ha of agricultural land in BQ. Conventional tillage farming is a technique where farmers completely remove harvested crop stalks and remaining biomass, allowing exposed soil to be eroded, blown away by winds, and deposited into nearby bodies of water (Grantz *et al.*, 1998). However, this form of farming has been phased-out in recent years in favour of conservation tillage, where the stalks and roots of harvested crops are kept in the soil until the next farming

season (BQRAP, 2018). Conservation tillage farming practices have been beneficial as they greatly reduce the potential for nutrient-rich soil to erode and drift into the Bay.

Moreover, phosphorus and nitrogen have also been deposited into the Bay through wastewater treatment plants. Historically, sewage was not fully treated to remove excess pharmaceuticals, solids, and nutrients among other waste products in primary wastewater treatment plants, meaning a lot of these waste products made their way into the bodies of water nearby the facilities (Nicholls & Hurley, 1989) These urban sources of phosphorus in BQ have been mitigated in recent years thanks to the upgrade of wastewater treatment facilities to at least secondary wastewater treatment plants, where more of the sewage input is treated to remove excess nutrients and waste before release into the Bay.

BQ has also had a significant presence of industrial factories and facilities in the past, which have deposited several different types of chemical compounds and heavy metals that are still present in BQ to this day. For example, Bakelite Thermosets Limited had a factory in Belleville, Ontario where a number of chemicals were made, including formaldehyde and Bakelite plastic. Bakelite is one of the first plastics made from synthetic components that is revered for its beauty and many uses in decorative art pieces and jewelry. However, the process by which Bakelite is made creates a lot of waste. Upon the realization of necessary environmental controls for pollution in the 1970s, Bakelite Thermosets Limited disposed of their hazardous waste chemicals on the site of the factory that eventually leached into the surrounding soils and the waters of BQ. Despite being demolished in 2005, the hazardous waste remains on the former site of the factory as the clean-up of the waste is far too expensive.

Furthermore, the counties surrounding the BQ have developed greatly over recent decades, leading to the depletion of shoreline wildlife habitat. Moreover, contaminated waste from several industries in the past have contaminated fish during the twentieth century (Government of Canada, 2017). It is for these reasons that BQ was designated as an AOC in 1987 alongside HH (IJC, 1987). As of 2018, eleven of the fourteen BUIs are designated as impaired or impaired but undergoing a re-designation process and are detailed in Table 1.1 (BQRAP, 2018).

1.4 Omics and Phytoplankton

Proteomics and metabolomics are the large-scale studies of proteomes and metabolomes. Proteomes are the entire set of proteins in a biological system. Similarly, metabolomes are the entire set of metabolites produced in a biological system. The words "proteome" and "metabolome" are both combinations of the words "protein" and "metabolite" with the word "genome." The term "genome," the entire set of genes in a biological system, was introduced in 1920 by Hans Winkler (Cristescu, 2019), and the term "proteome" was first coined by Wasinger *et al.* (1995). A few years later in 1998, Oliver *et al.* first used the term "metabolome." The late 1990s were the beginning of several new branches in biology, informally known as "omics." In addition to proteomics and metabolomics, which aim to characterize and quantify the function and dynamics of biological systems, omics also includes genomics and transcriptomics, which are the studies of the genes of a biological system and the RNA transcripts in which they produce.

Proteomics and metabolomics can be either targeted or non-targeted. Targeted approaches look for certain proteins or metabolites only, where the way in which the compounds behave is already known (such as mass and retention times). Non-targeted approaches, conversely, aim to see any proteins or metabolites present in the sample. Targeted approaches have higher selectivity and sensitivity than non-targeted approaches because targeted methods limit the amount of data being measured. Non-targeted approaches therefore have lower selectivity and sensitivity than targeted approached, but are far more useful for looking at samples with a wider lens to identify unknown compounds, like harmful chemicals in an environmental sample, metabolites of an unknown metabolic profile, or a biomarker for disease in a blood plasma sample.

Proteomics is a really useful approach to evaluate the function of biological systems. Proteins represent the final product of gene expression and indicate a cell's priorities and actions to stimuli. Proteomics can be used to explain which nutrients are being utilized by certain organisms. For example, Diaz *et al.* (2018) used proteomics to discern which form of dissolved organic phosphorus was preferred by a genus of marine diatom based on the abundances of different phosphorus-degrading enzymes. This information helps to resolve how nutritionally-available phosphorus is used in a marine ecosystem. Furthermore, metaproteomics, or the proteomic approach to evaluate all of the proteins in a microbial community, is becoming increasingly popular because this technique provides insights into the activities of all microorganisms in a community. For instance, Georges *et al.* (2014) used metaproteomics to identify proteins that may potentially serve as indicators of long-

term environmental changes in organic nutrient scavenging linked to seasonal changes from winter to spring,

Metabolomics is a great technique that can identify the dynamics and physiological behaviour within a biological system. Metabolomics focuses on metabolites, the end products of all cellular processes representing: products of metabolic reactions, outputs from enzymatic reactions, and precursors to other chemical reactions. Changes in environmental conditions can be reflected in metabolomes of phytoplankton. Kuhlisch *et al.* (2020) evaluated the metabolome of an arctic marine phytoplankton and identified that changes to nutrient availability and light can be seen in their metabolome. The changes in the metabolites produced may have a direct effect on the quality of the phytoplankton as food, thereby affecting zooplankton grazing behaviours, contributing to the understanding of arctic marine food webs. Investigating the metabolites produced by phytoplankton also gives great insight into what the algae are doing with the resources available to them. For example, two common forms of cyanotoxins (MC-LR and MC-RR) require arginine to be constructed, therefore I hypothesize that arginine is an important amino acid in the production of cyanotoxins and could be an important metabolite to look for during a metabolomics approach. Dai *et al.* (2009) identified the importance of arginine, leucine, and alanine as amino acids that were readily biosynthesized by *M. aeruginosa* to produce MCs. Other potential metabolites that would give insight into cyanobacteria dynamics include acetates and other carbon-containing molecules that comprise cyanotoxins (Tillett *et al.*, 2000; Moore *et al.*, 1991). Thus, the presence of the aforementioned amino acids and

carbon-containing molecules could be an indication that they are being used by the cyanobacteria to synthesize cyanotoxins.

Proteomics and metabolomics are very appealing techniques to apply to HAB communities to elucidate function and behaviour because of the ability of these techniques to acquire vast amounts of information. For example, it is not clear why there are toxigenic and non-toxigenic species of cyanobacteria and why non-toxigenic species might not take advantage of the benefits of cyanotoxin production. I believe that the use of non-targeted -omics approaches such as proteomics and metabolomics may enable a better understanding of the metabolic switch that allows cyanobacteria to produce toxins.

1.5 Research Objectives

The goal of this thesis was to characterize the phytoplankton community composition in waters sampled from two Lake Ontario AOCs and assess how the metabolites produced by these communities changed over time to relate this information to the formation of HABs. Understanding the community composition of the phytoplankton in the two Lake Ontario AOCs of this study will serve as baseline information for future research in these areas and this information will better inform future management decisions regarding HABs. Additionally, associated metabolites for phytoplankton and how their composition changes over time could contribute to the formation of a metabolite-based predictor of HABs. In order to achieve these long-term goals, the following research objectives were formulated:

1. Characterize and assess temporal changes in both the WQ conditions and phytoplankton community composition in HH and BQ over one summer (August to October);
2. Identify the metabolites of phytoplankton related to phytoplankton bloom formations and assess how phytoplankton metabolites change over time.
3. Identify relationships between phytoplankton community composition, WQ conditions, and phytoplankton metabolites.

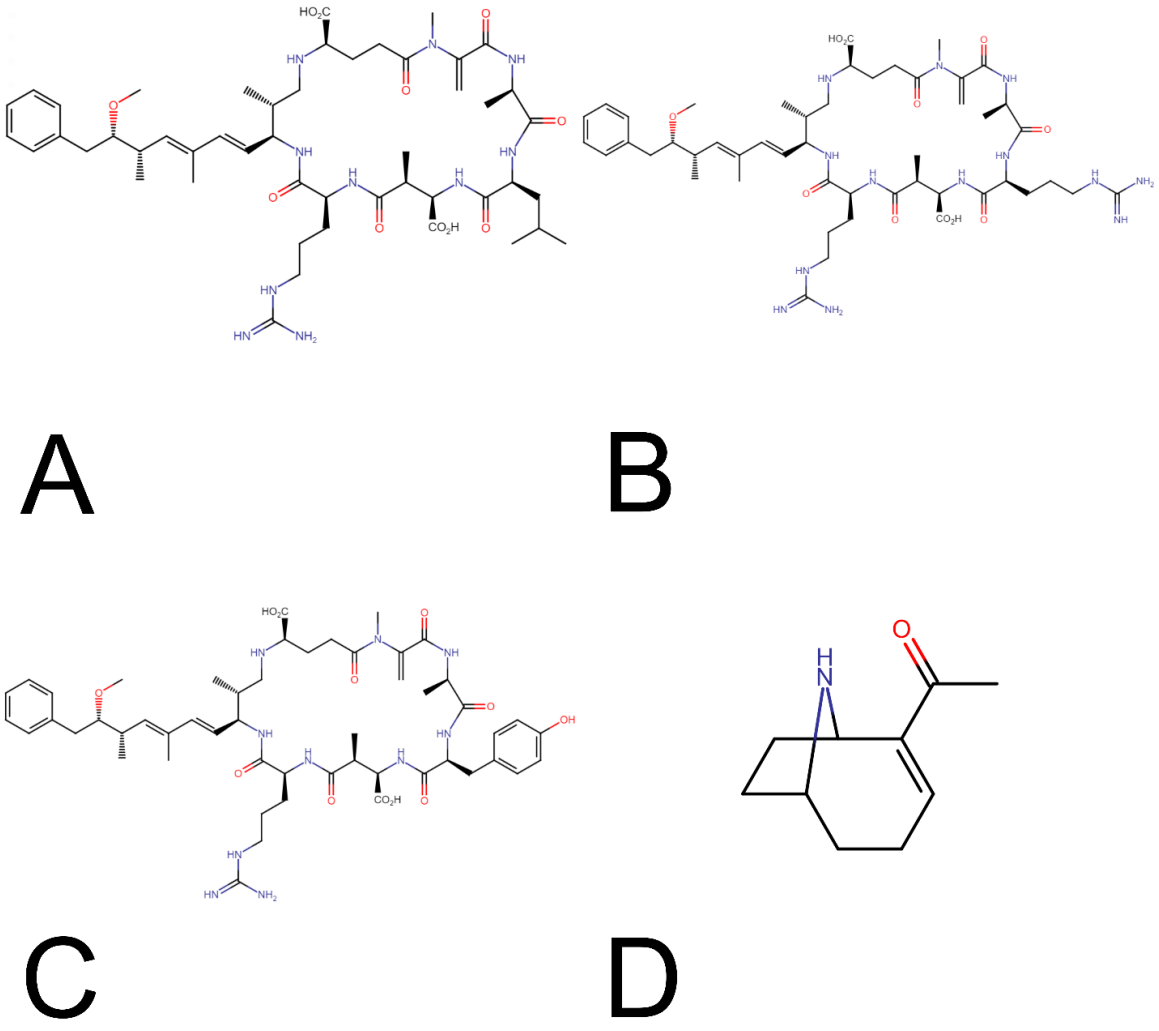

Figure 1.1. Molecular structures of four common and toxic cyanotoxins, including: MC-LR (A), MC-RR (B), MC-YR (C), and ATX-a (D).

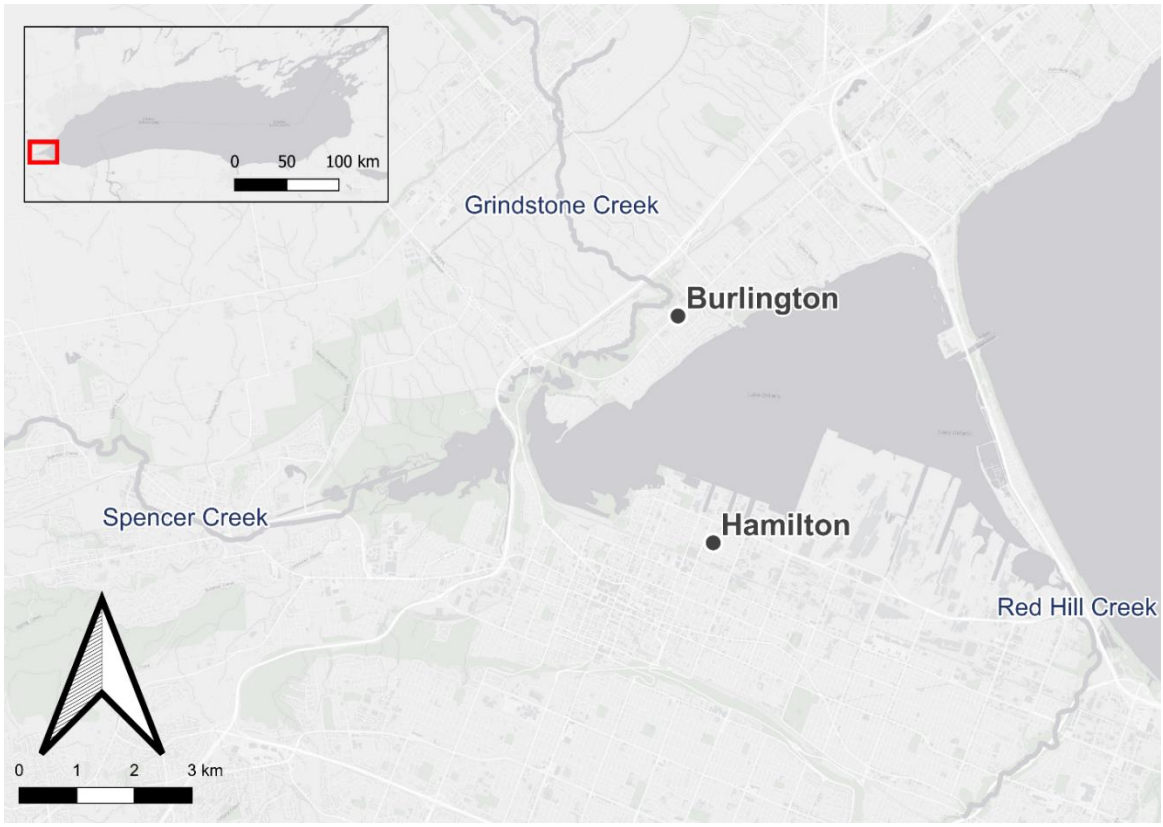

Figure 1.2. Location overview of HH. Major cities surrounding the harbour include Burlington, ON to the north and Hamilton, ON to the south. Three major rivers feed into the harbour, including: Spencer Creek, Grindstone Creek, and Red Hill Creek.

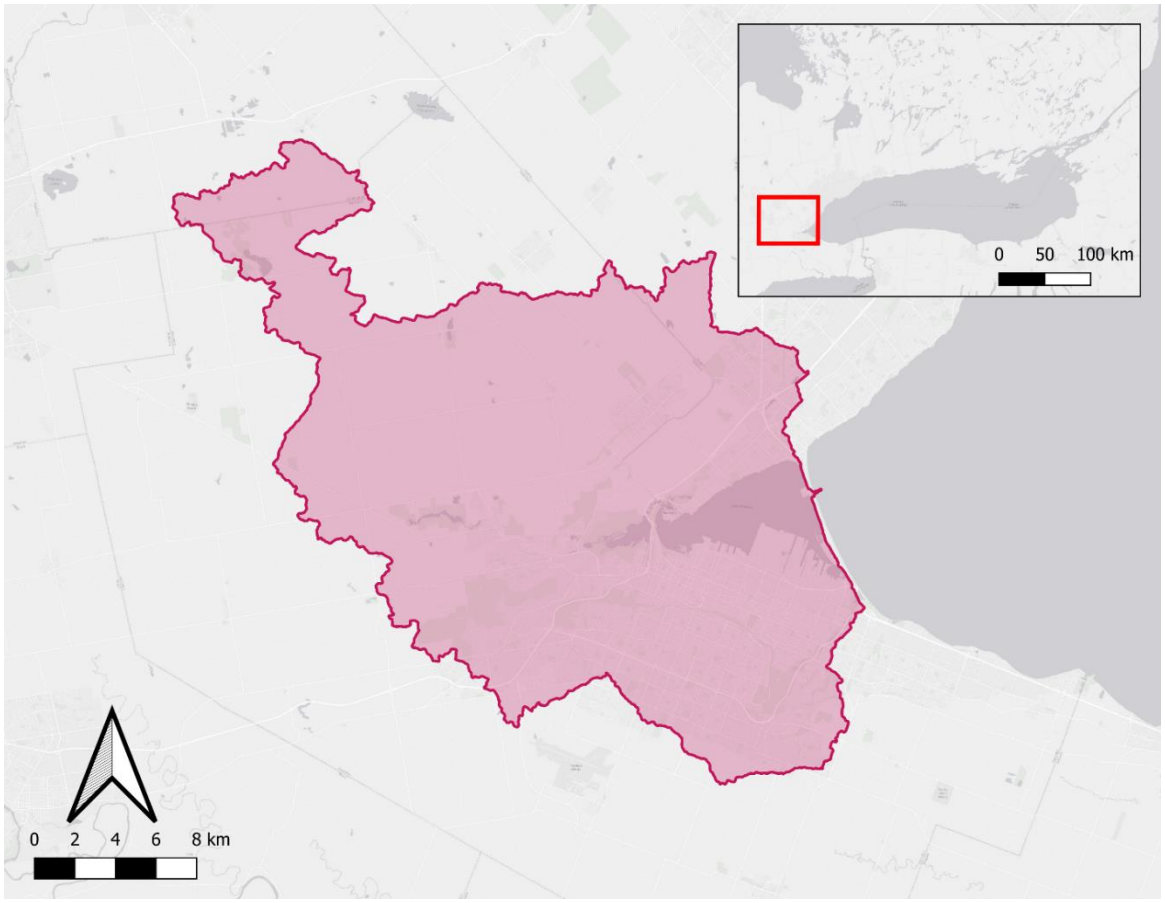

Figure 1.3. HH watershed overview map.

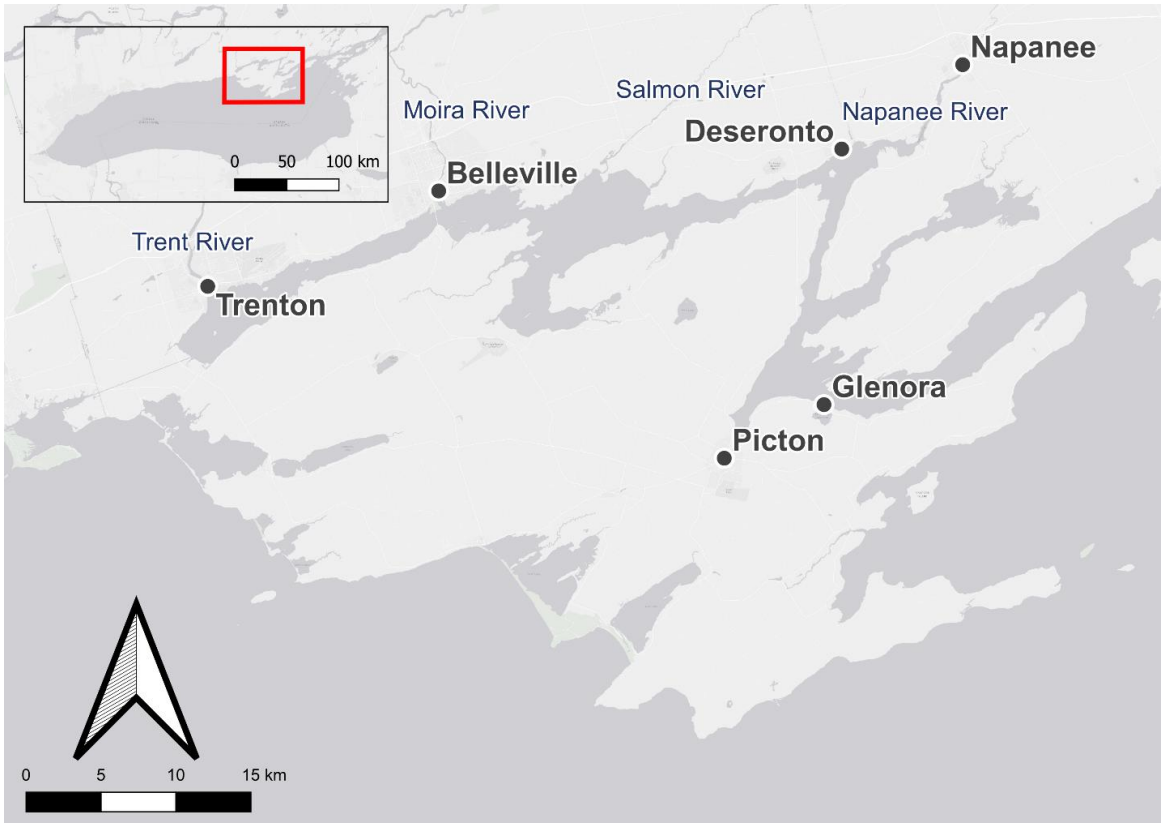

Figure 1.4. Location overview of BQ. Major communities surrounding the Bay include Trenton, ON, Belleville, ON, Deseronto, ON, Napanee, ON, Picton, ON, and Glenora, ON. Four major rivers feed into the bay, including: Trent, Moira, Salmon, and Napanee River.

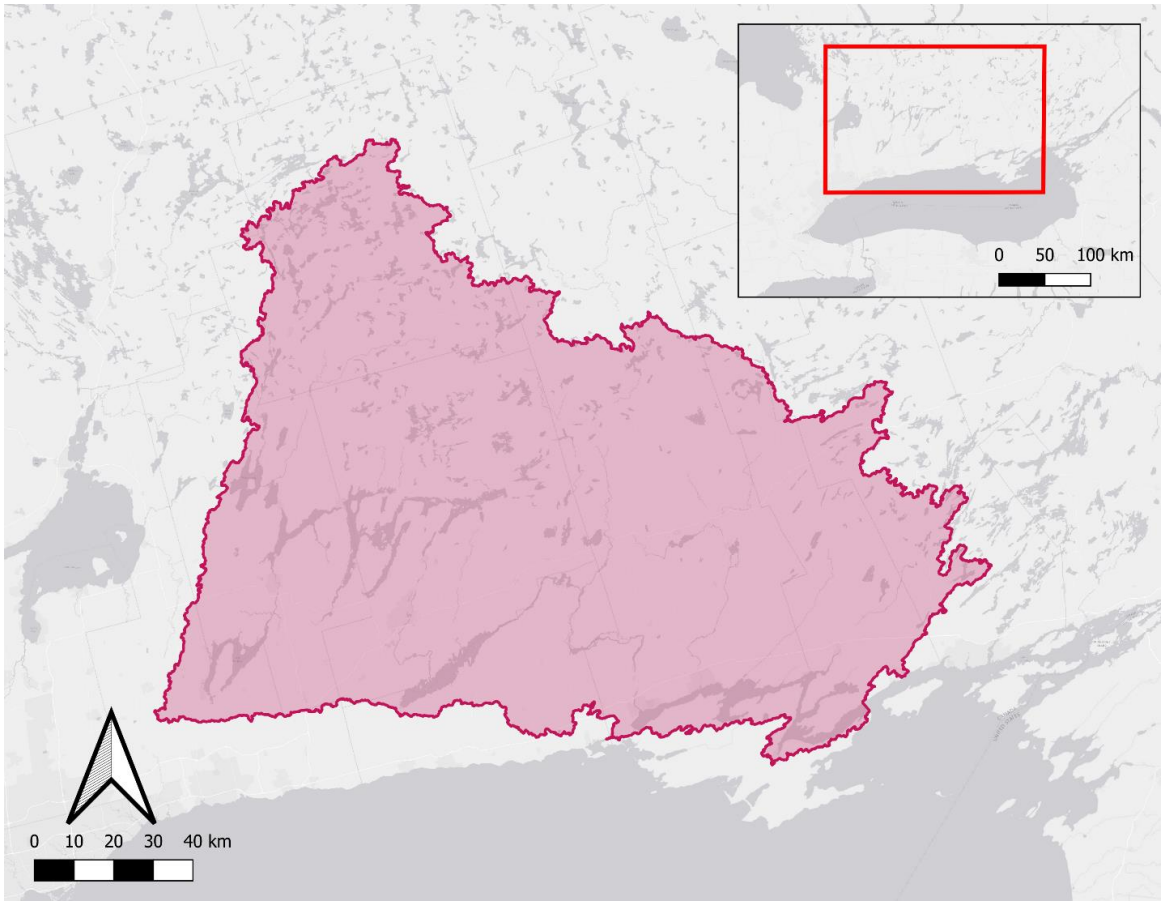

Figure 1.5. BQ watershed overview map.

Table 1.1. BUI status for HH and BQ as of 2018. Both locations share one impaired BUI: Eutrophication and Undesirable Algae (Government of Canada, 2017; HHRAP, 2018; & BQRAP, 2018).

Beneficial Use Impairment	HH Status	BQ Status
1. Restriction on Fish and Wildlife Consumption	↓	•
2. Tainting of Fish and Wildlife Flavour	↑	↑
3. Degradation of Fish and Wildlife Populations	↓	→
4. Fish Tumours or Other Deformities	←	→
5. Bird or Animal Deformities	←	↑
6. Degradation of Benthos	↓	→
7. Restrictions on Dredging Activities	—	→
8. Eutrophication or Undesirable Algae	↓	↓
9. Restrictions on Drinking Water Consumption	↑	↓
10. Beach Closings	↓	→
11. Degradation of Aesthetics	↓	→
12. Added Costs to Agriculture and Industry	↑	↑
13. Degradation of Phytoplankton and Zooplankton Populations	←	↓
14. Loss of Fish and Wildlife Habitat	↓	→

↑ Not Impaired
↓ Impaired
→ Impaired; Re-designation in Process
← Requires Further Assessment
• Impaired in Some Areas
— Deferred

CHAPTER 2: METHODOLOGY

2.1 Study Areas

As previously mentioned, the two locations in which all eight study sites of this study can be found are HH and BQ, chosen for their contrasting landscapes and land-use. HH is primarily surrounded by an urbanized landscape with numerous populated cities and industrial complexes immediately along its shore. Conversely, BQ is primarily surrounded by an agricultural landscape with fewer urbanized towns immediately along its shore.

2.1.1 Hamilton Harbour

The three sampling sites in HH include: Bayfront Harbour (BF), Carroll's Bay (CB), and Yacht Club (YC). These three sites were located on the western side of HH to hone in on the influences from the nearby City of Hamilton. The industrialized southern shore of HH was avoided due to heavy ship traffic and other safety concerns. Initially, only CB was sampled alongside two additional sampling locations, one near the Burlington Golf & Country Club (GC) and another by Neare Island next to the Canadian Centre for Inland Waters building (NI), however these sampling sites were omitted after week four because phytoplankton seemed to be absent from these locations. The data from GC and NI are entirely excluded from this thesis. Instead, BF and YC were chosen in addition to CB for sampling and is why there are no data available for analyses before week five for BF and YC. Algal blooms were anecdotally observed in BF and YC during the first five weeks of sampling, thus these two sites were chosen for sampling. CB was thought to be a suitable sampling location as preliminary investigations into satellite imagery revealed the site around CB turbid and I thought it would be conducive to the growth of phytoplankton and

presence of excess nutrients. A figure detailing the sampling sites in HH can be seen in Figure 2.1.

2.1.2 Bay of Quinte

As previously mentioned, BQ is separated into three main sections: the upper, middle, and lower bay (Figure 2.2). Of the three sections, the middle bay was chosen for the sampling site locations. The middle bay was ideal for sampling as this unique portion of BQ encompasses a wide variety of landscapes, including several urbanized towns including Deseronto and Picton, as well as an abundance of agricultural land (Figure 2.3). Additionally, there are several unique hydrological features in the middle bay, including a deeper central open area and a long bay jutting off of the middle bay known as Hay Bay.

The five study sites in BQ include: Picton Harbour (PH), Glenora (GL), Hay Bay (HB), Wilton Creek (WC), and Napanee (NP). Initially, only PH, HB, and WC were sampled, however GL and NP were added starting on week two of sampling. All sites were chosen as they are similar to the sampling locations from previous studies, namely those of Millard and Sager (1994) and Munawar *et al.* (2018), allowing for potential comparisons to historical information. PH was selected as a site for sampling because it was near Picton, a major town in Prince Edward County, and any influences from this more urban area could have an effect on the types of algae found. GL was located near a frequently-ran ferry crossing in markedly deeper water than other locations in the middle bay. NP is near the town of Deseronto, ON in Hastings County and is also near the outflow of the Napanee River. HB was chosen due to a significant presence of historical algal blooms, in addition

to being located in the unique bay jutting off of the middle bay. WC was also chosen for sampling to assess any spatial relationships with HB as WC is near the end of the Hay Bay jutting off of the middle bay. Moreover, WC is also near the outflow of a stream of the same name which is surrounded by vast amounts of agricultural land, and whose influences were thought to have an effect on the phytoplankton found in this location. Figure 2.3 details the locations of the sampling sites in BQ

2.2 Water Quality Data and Sample Acquisition

2.2.1 Water Quality Data Acquisition

I sampled all but 2 sites weekly for ten weeks in 2020, starting in the first week of August and ending in the first week of October. I sampled HH weekly on Mondays and BQ on Thursdays; if weather was not forecasted to be conducive to sampling, sampling was pushed back a day. A detailed look at each date when sampling occurred can be found in Appendix A. I conducted all sampling on a Boston Whaler Dauntless 170 research vessel.

Due to equipment allocation issues from COVID-19 lab closures, I was unable to measure all WQ parameters at all BQ sites on week one and only PH on week two. I was able to collect phytoplankton microscopy samples during all ten weeks at all sites, except for week one in NP and GL in BQ. Similarly, I was not able to collect neither WQ nor phytoplankton microscopy samples at BF and YC in HH prior to week six for reasons previously described.

I measured the following WQ parameters on-site during each sampling event: site depth (m), Secchi depth (m), water temperature (TEMP, °C), and DO (mg/L). Ideally, I would have used a multi-parameter probe to collect further on-site data, such as pH, conductivity (COND), total dissolved solids (TDS), and turbidity (TURB), however access to such a probe was not possible due to laboratory closures caused by the COVID-19 pandemic lockdown restrictions. Therefore, I collected additional water samples in 50 mL centrifuge tubes and stored them in coolers containing ice for transport back to the lab for further WQ parameter measurements on the same day as sampling. The additional WQ parameters I measured at the lab upon returning to campus included: COND ($\mu S\ cm^{-1}$), pH, TDS (ppm), nitrates (NO_3, ppm), TP (ppb), and absorbance @ 750 nm (proxy for TURB). I froze the 50 mL centrifuge tubes for NO_3 and TP at -20 °C until nutrient analyses could be conducted.

I recorded both site depth and TEMP using the research vessel's on-board depth transducer and thermometer, respectively. To measure DO, a HANNA Oxy-Check Dissolved Oxygen Meter was used on-site. Immediately upon returning to the laboratory, I removed the 50 mL centrifuge tubes from the coolers and recorded pH, COND, and TDS using a Mettler Toledo S40 SevenMulti pH meter. TDS was found to be collinear to COND and was therefore removed from all analyses included in this paper.

I measured TURB from the 50 mL centrifuge tubes after immediately returning to the lab where absorbance was recorded as a proxy for TURB since COVID-19 lockdown restrictions made access to a turbidimeter difficult. To measure TURB, I transferred water

samples to 1 cm plastic cuvettes and measured absorbance at a wavelength of 750 nm on a Varian Cary Eclipse Fluorescence Spectrophotometer. I will be referring to the absorbance readings as TURB for the remainder of this paper.

I measured TP using the Ontario Ministry of the Environment (1983) modified Ascorbic Acid method, which was originally developed by Murphy and Riley (1962). I also measured NO_3-N using a LaMotte Nitrate-Nitrogen test kit and obtained NO_3 by multiplying these data by 4.4 as per the instructions of the test kit.

2.2.2 Phytoplankton Sample Acquisition

I scooped water at the surface of the water using a bailer at each sampling site into 500 mL plastic jars for phytoplankton microscopy. After, I added Lugol's iodine to each jar until the water inside resembled weakly brewed tea (approximately 5% of the total volume) as per the instructions by Guillard and Sieracki (2005), then the jars were kept in iced coolers for transport back to campus. Upon returning to the campus, I wrapped the jars in tinfoil to prevent light from deteriorating the iodine and put them in a refrigerator at 4 °C.

2.2.3 Omics Sample Acquisition

I collected the water samples for omics analyses in 10 L plastic carboys. To do this, I scooped water from surface level and funneled into the carboys, then kept the carboys on ice in coolers for transport back to campus. Upon returning to campus, I placed the carboys in a refrigerator at 4 °C I could process them. To process the water, I removed the 10 L carboys from storage and transferred the water to twenty 50 mL centrifuge tubes (1 L total

at a time) and centrifuged them at 1715 g at 4 °C for 15 minutes. Once a 15-minute cycle was complete, I discarded the supernatant leaving a small pellet of algal in each tube. I then re-filled the tubes with fresh sample water and repeated the process until sizable algal pellets could be seen in all twenty 50 mL centrifuge tubes. Once enough pellets had formed in each of the tubes, I transferred the pellets to several 2 mL microcentrifuge tubes and spun them at 1710 g at 4 °C for 15 minutes. I poured off the remaining supernatant from the microcentrifuge tubes prior to storing the samples at -80 °C until -omics analyses could occur. The amount of water I centrifuged for each sample varied from 1 L to 10 L total depending on the density of phytoplankton during collection. A detailed table of how much water was centrifuged for each site and week can be found in Appendix A.

2.3 Sample Preparation and Data Acquisition

2.3.1 Phytoplankton Microscopy

I used the Utermöhl method (Utermöhl, 1931) to settle the Lugol's-preserved phytoplankton samples for a minimum of 24-hours. I identified the phytoplankton to genus level by following dichotomous keys from Prescott (1962), Sheath and Wehr (2003), and Baker *et al.* (2012) using an EVOS xl-core inverted microscope at 400x magnification. During phytoplankton identification, I counted a minimum of 300 phytoplankton cells per sample. To get enough cells counted, I typically counted phytoplankton in one transect of the slide, but I occasionally counted phytoplankton in a quarter of a transect for phytoplankton-dense samples, and up to two transects in phytoplankton-rare samples. Additionally, I used the algal cell counts used to calculate total cells present, cells per millilitre, and cells per litre. Furthermore, I measured algal cell lengths and widths using a

standard ruler (mm) on the microscope's LCD screen, which I then converted to micrometers to calculate both phytoplankton biomass (mg/L) and phytoplankton biovolume (μm^3).

2.3.2 Phytoplankton Omics Sample Extraction

For metabolomics, I removed the frozen cell pellet samples from the -80 °C freezer and kept them on ice for the duration of the liquid-liquid phase extraction. I thawed the pellet samples on ice, and used a clean stainless-steel micro spatula to weigh 50 mg of each pellet immediately upon complete thaw. I quenched metabolic activity using 100 µL of ice-cold methanol and then I sonicated the samples five times for 1 second each. After sonication, I then lysed the samples in a ball mill (Qiagen Tissue Lyzer) with one 5 mm stainless-steel ball for 2 minutes at 60 Hz. I then removed the stainless-steel balls and transferred the cellular slurries to 2 mL low retention tubes. Next, I added 50 µL of ice-cold chloroform to the samples, followed by 20 seconds of vortexing; then an addition of 50 µL of ice-cold MilliQ water and 20 seconds of vortexing; and one more 50 µL addition of ice-cold chloroform followed by 20 seconds of vortexing. I spun the samples at 5,000 rpm at 4 °C for 30 minutes to separate the solutions into three phases: an upper aqueous phase layer containing metabolites, a middle disk of proteins and cellular debris, and a lower organic phase layer containing lipids.

2.3.3 Phytoplankton Metabolites

From the liquid-liquid phase extraction, I separated a 50 µL aliquot from the aqueous phase to use for metabolomics. I added 25 µL of a mixed stable isotope labelled standard

(Canonical Amino Acid Mix, Cambridge Isotope Laboratories) to each 50 μL aliquot, then vortexed the samples for 3 seconds to mix the solutions. I then evaporated the solutions in a SpeedVac™ SPD 1030/2030 centrifugal evaporator for 30-minute intervals until the samples were at near-dryness. I re-suspended the samples in 10 μL of 90% acetonitrile 10% MilliQ H_2O and removed 2 μL from each re-suspended sample to pooled together to make a composite sample containing all of the aqueous phases. I took the remaining 8 μL of each sample and transferred them to 250 μL spring-bottom polyphenylene ether (PPE) spring inserts (Canadian Life Sciences, 6 × 29 mm) and placed the inserts in 2 mL pre-labelled screw thread high-performance liquid chromatography (HPLC) vials (Chromatographic Specialties, 12 × 32 mm), capped with blue 9 mm polytetrafluoroethylene (PFTE)/silicone/PFTE screw caps.

5 μL of the metabolomics samples were separated on a HILIC zwitterion column (Agilent InfinityLab Poroshell 120 HILIC-Z, 2.1 × 150 mm, 2.7 μm, PEEK lined) using an Agilent Infinity 1260 HPLC coupled to an Agilent 6545 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) mass spectrometer for detection with 10mM ammonium formate with 0.1% v/v formic acid as mobile phase A, and 90% acetonitrile in 10mM ammonium formate with 0.1% v/v formic acid as mobile phase B. Detailed methods information is available in Appendix B.

2.3.4 Phytoplankton Proteins

I dried the protein disks that were separated from the liquid-liquid phase extraction samples in a SpeedVac™ SPD 1030/2030 centrifugal evaporator for 30-minute intervals. Then, I

resuspended the disks in 50 µL of 100 mM AB buffer and vortexed them to mix. After this, I added 2.65 µL of 100 mM of tris(2-carboxyethyl)phosphine in 100 mM AB buffer, vortexed to mix, and incubated the sampled at room temperature for 45 minutes to reduce the proteins. I then alkylated the samples with an addition of 2.8 µL of 200 mM iodoacetamide in 100 mM AB buffer, followed by a gentle vortex, and an incubation in the dark at room temperature for 45 minutes. At the end of the second incubation, I added 50 µL of chemical digestion solution (20% formic acid v/v) to each sample and vortexed them for 5 seconds. I placed lid locks were on each tube and incubated the tubes at 115 °C for 30 minutes in a heating block (VWR 96 place heating block). I then dried the samples in a SpeedVac™ SPD 1030/2030 centrifugal evaporator for 40 min and stored the dried samples at 4 °C overnight. After storage, I resuspended the samples in 20 µL of 95% H_2O 5% acetonitrile and 0.1% formic acid, making sure to vortex the samples until dried pellets were completely dissolved. I then centrifuged the samples for 10 min at 14,000 g at 4 °C, then added 20 µL of the sample supernatant to 250 µL spring-bottom polyphenylene ether (PPE) spring inserts (Canadian Life Sciences, 6 × 29 mm). Next, I took the inserts and placed them into 2 mL pre-labelled screw thread HPLC vials (Chromatographic Specialties, 12 × 32 mm) capped with blue 9 mm PFTE/silicone/PFTE screw caps. The samples were stored at 4 °C until instrumental analysis.

2 µL of the peptide solution from each sample was injected and then separated by reverse phase liquid chromatography using a Zorbax, 300SB-C18, 1.0 × 50 mm 3.5 µm column (Agilent Technologies Canada Inc., Mississauga, ON) using an Agilent 1260 Infinity Binary LC. The Agilent 6545 Accurate-Mass Q-TOF was used as the detector in tandem

to the Agilent 1200 series liquid chromatography system (see detailed instrumental methods are available in Appendix C). Each analytical run included a solvent blank and a BSA digest standard (Agilent Technologies Canada Inc., Mississauga, ON) injection every 10 samples in order to monitor baseline, carry-over, drift, and sensitivity during the runtime. The protein samples were injected once per each individual sample.

2.3.5 Phytoplankton Metabolite Identification and Quantification

I ran the spectral data in MassHunter Profinder (version 10.0) using metabolite information from a database of the masses and retention times of thirty-three metabolites and amino acids of interest and twenty standard amino acids using MassHunter Personal Compound Database and Library (PCDL) Manager Software (version B.08.00). Chromatographic peaks were automatically integrated by Profinder for most sites and locations, otherwise peaks were manually integrated when they were not recognized by the program automatically. Once the data from all sites and locations were inspected, I exported the chromatographic peak table data and chromatographic peak area tables as comma separated values (CSV) files for statistical analyses. All endogenous metabolites were quantified using the twenty-seven isotope-labelled internal standards, where I divided the endogenous metabolite by its internal standard counterpart and multiplied by the internal standard concentration of 0.416 mM. Several metabolites did not have internal standard counterparts, thus surrogate internal standards that were both similar in structure and in physical properties were used for quantification instead. A table detailing which internal standards were used to quantify each of the metabolites of interest can be found in Appendix C.

2.3.6 Phytoplankton Protein Identification and Quantification

I analyzed the spectral files for each sample using Spectrum Mill Software (Version B.04.01.141). Peptides were searched against a concatenated freshwater algae database containing proteins from seven algal classes: Bacillariophyceae, Charophyceae, Chlorophyceae, Cryptophyceae, Cyanobacteria, and Dinophyceae (downloaded from Uniprot, June 2021).

Proteins were then manually validated and accepted when at least one peptide had a peptide score (quality of the raw match between the observed spectrum and the theoretical spectrum) greater than 5 and a %SPI (percent of the spectral intensity that are accounted for by theoretical fragments) of greater than 60% (these settings are recommended by the manufacturer for validating results obtained with an Agilent Q-TOF mass spectrometer). After peptides were sequenced and identified by Spectrum Mill at the MS/MS level, quantification at the MS1 level was performed using the DDA workflow in Skyline 20.2 (MacCoss Lab Software) with a score of 0.9, retention time window of 5 mins, and 5 missed cleavages with transition settings for TOF (Pino *et al.*, 2020). Data were sorted and manually consolidated using Excel.

2.4 Data Analysis

To compare the differences in both the aquatic environments and the surrounding landscapes of HH and BQ, I used a Principal Components Analysis (PCA) to examine the covariation of the following WQ parameters: COND, pH, NO_3, TP, TEMP, TURB, DEPTH, and Secchi depth. I center-standardized all WQ data in RStudio (version 4.0.2., R

Core Team, 2020) prior to creating the PCA. The PCA allowed me to identify how the WQ parameters distinguished the aquatic environments in each of the sites HH and BQ. To investigate the surrounding landscapes of HH and BQ, I analyzed the watersheds of each location to categorize the land-uses in QGIS (version 13.10, QGIS Development Team, 2020). I mapped out the watersheds of each location using tertiary watershed shapefiles that I obtained from the Canadian Open Government website (Government of Canada, 2016). I obtained water shapefiles from Natural Resources Canada (NRCAN, 2021), and regional and municipal roads data from the Quick OSM plugin in QGIS. I calculated the areas of each of the land-uses directly in QGIS and normalized the data to percentages in Microsoft Excel. To easily compare the land-use area percentages, I created column plots in RStudio.

To classify the phytoplankton community composition profiles between HH and BQ, I created stacked bar plots for all eight sites that showed the relative abundance of the biomass (mg/L) for each of the seven phytoplankton groups of interest over the ten weeks of sampling in RStudio. To discern the relationships between the phytoplankton groups of interest and the eight WQ parameters, I used a Redundancy Analysis (RDA) where the biomasses of the phytoplankton groups were the species variables and the WQ parameters were the environmental variables. I conducted the RDA in RStudio (version 4.0.2, R Core Team, 2020), and performed a Hellinger transformation on the species data to give lower weights to rarer phytoplankton groups. To get a better idea about which types of phytoplankton within the eight groups of interest were being affected by the WQ, I created a Kendall correlation matrix using the phytoplankton biomass and WQ data. Kendall

correlation matrices are ideal for data that are non-parametric, such as the phytoplankton biomass and WQ data I acquired, and are useful in identifying significant trends in data. To get more detail about each of the phytoplankton groups of interest, I divided the groups into the orders of phytoplankton identified and then I created the Kendall correlation matrix in RStudio. Then I used a multiple linear regression to model the predictive relationship between the WQ parameters and phytoplankton biomass. I included the weekly mean biomass of each phytoplankton group as the dependent variable and the weekly WQ parameters as the independent variables using SigmaPlot (version 14.5, Systat Software, 2020). Variance inflation factors were examined to determine which variables were colinear—and I subsequently removed COND due to being colinear with NO_3.

I used the Simpson's Diversity index to investigate the phytoplankton biodiversity of each of the sites because it considers both species richness and evenness, two important metrics when quantifying biological diversity, and it yields an easy-to-understand index between 0 and 1, where 0 is no diversity and 1 is infinite diversity. I calculated Simpson's Diversity indices for each of the sites for all of the weeks that data were available and created line plots to visualize the change in indices over time using RStudio. To see if any WQ parameters could be used as a predictor for a higher Simpson's Diversity index, I conducted a multiple linear with WQ parameters as the independent variables (SigmaPlot version 14.5, Systat Software, 2020).

To see how the types of metabolites found in both HH and BQ differed, I normalized the metabolite concentration data to percentages, and created stacked bar plots to assess their

abundance using RStudio. To discern how the phytoplankton groups of interest were associated with the metabolites, I used an RDA because it allows for the easy interpretation of a large amount of multivariate environmental data with respect to a set of species, such as the WQ parameters and the seven phytoplankton groups of interest. I pooled the data from HH and BQ together and performed a Hellinger transformation on the phytoplankton biomass (mg/L) data and metabolite concentration (mM) data in RStudio to give low weights to rarer phytoplankton and metabolites, respectively, prior to making the RDA. From the R vegan package (Oksanen *et al.*, 2020), I used the "vif.cca()" function to check if any of the metabolite variables were redundant. From this, I removed the following redundant metabolites: glutamine, isoleucine, hydroxyproline, ornithine, phenylalanine, serine, tyrosine, valine, and 1-methylhistidine because they would not contribute to the model in any meaningful way. To examine which of the relationships identified in the RDA were significant, I used another Kendall correlation matrix as a suitable way of examining trends in non-parametric data, such as the phytoplankton biomass and metabolite concentration data. Similarly, I also created another Kendall correlation matrix to identify significant non-parametric relationships between the WQ of interest and the metabolites identified in both locations.

I pooled the metabolite data together and conducted pathway analyses in the online statistical program, Metaboanalyst 5.0 (Pang *et al.*, 2021). Metaboanalyst 5.0 was ideal to use because it can access multiple metabolite pathway databases, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, as well as to create meaningful visualizations. Prior to assessing metabolite pathways, I normalized the metabolite

concentration data by median and log-transformed them using the built-in tools in Metaboanalyst 5.0. I selected the Global Test pathway enrichment method that uses a generalized linear model to estimate a Q-statistic for each of the metabolites in the set, allowing an outcome pathway to be predicted based on the metabolite concentration profiles provided. I searched the metabolite data against the *Chlorella variabilis* KEGG pathway library (obtained in October 2019) as it was the closest reference species taxonomically to the phytoplankton genera from which I obtained my data.

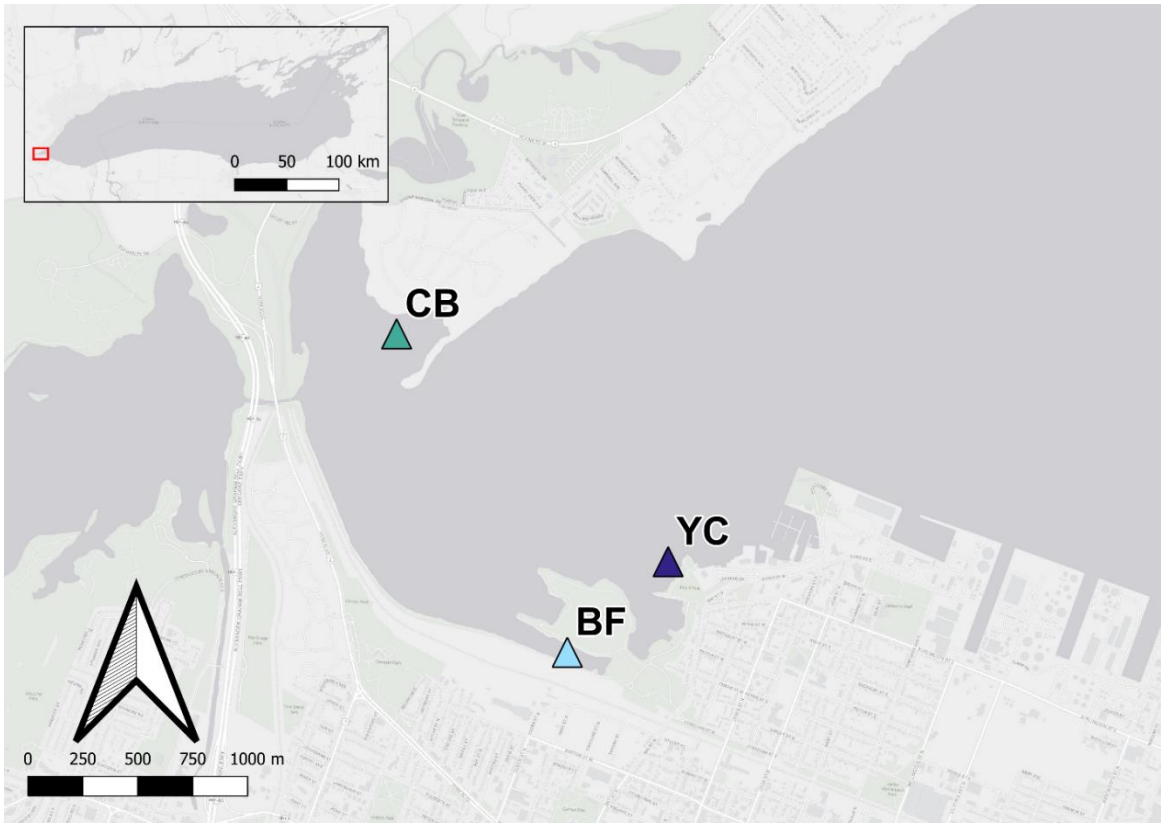

Figure 2.1. Site overview map for HH. Sites include Bayfront Harbour (BF), Carroll's Bay (CB), and Yacht Club (YC).

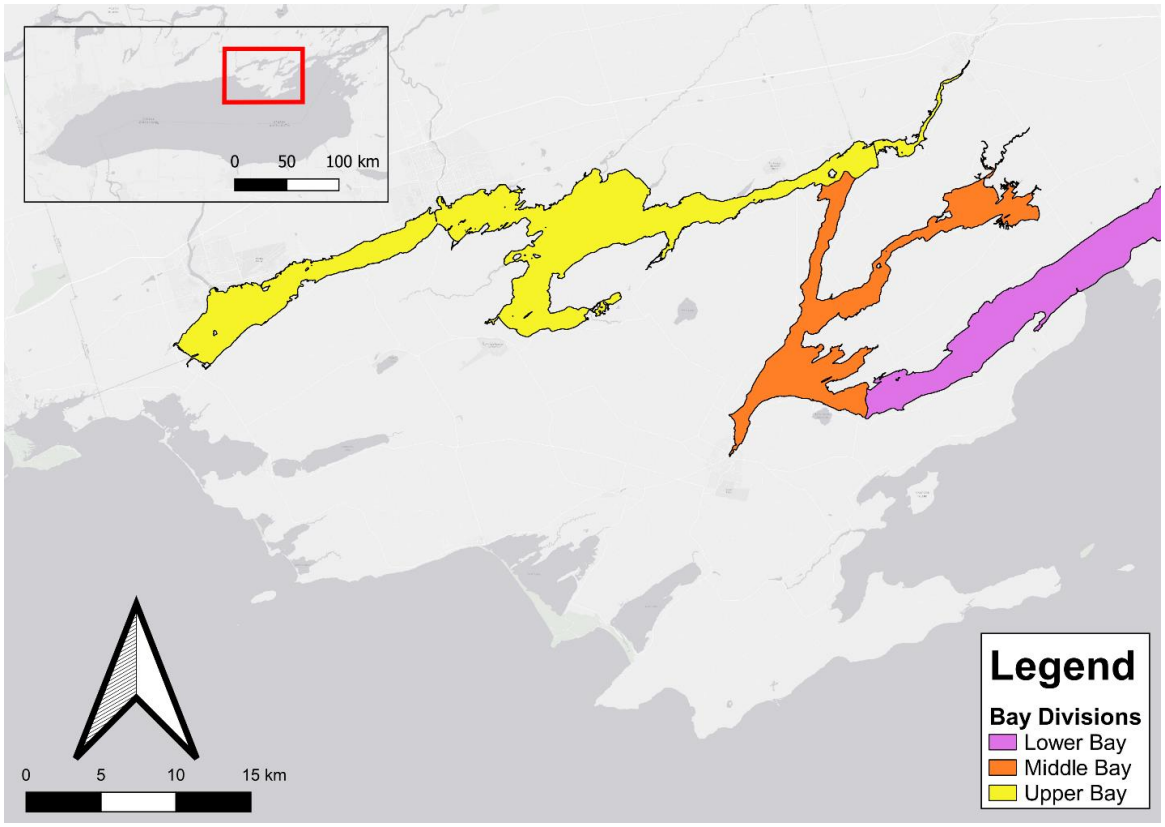

Figure 2.2. Map overview of BQ. The bay is separated into three sections, the Upper, Middle, and Lower Bays.

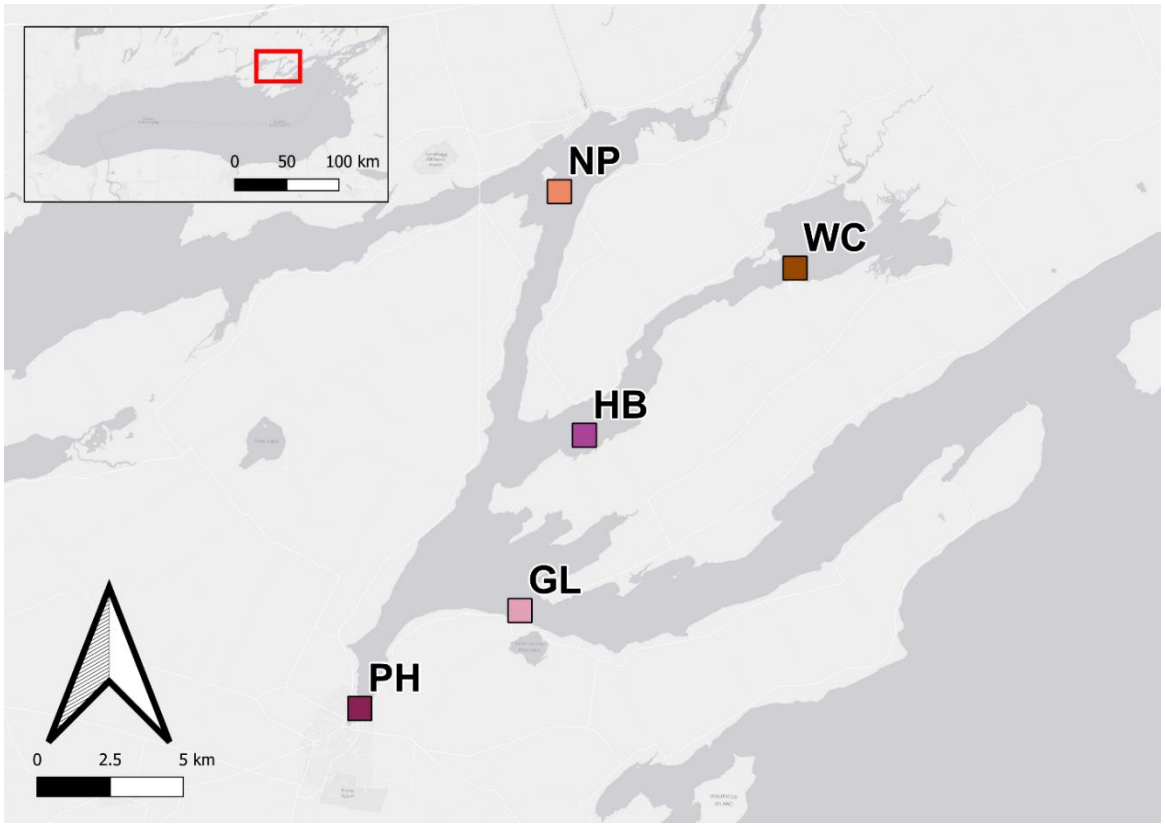

Figure 2.3. Site overview map for BQ. Sites include Picton Harbour (PH), Glenora (GL), Hay Bay (HB), Wilton Creek (WC), and Napanee (NP).

CHAPTER 3: RESULTS

3.1 Characterizing Water Quality in Hamilton Harbour and Bay of Quinte

There is a clear distinction in the land-use profiles of HH and BQ based on the land-use bar plots (Figure 3.1) and land cover summary statistics (Table 3.1). The majority of the land-use in HH is made up of agriculture whereas the land-use in BQ is primarily natural land-use. HH is far more developed than BQ, with 29% of the total area in the watershed draining into HH being developed compared to 1.5% in BQ. Despite the watershed of BQ being significantly larger than the watershed of HH, both have similar proportions of their land-use being water (9.3% and 5%, respectively). Full land-use overview maps can be found in Figure D2 and D3 in Appendix D.

The WQ parameter summary statistics for HH (Table 3.2) and BQ (Table 3.3.) also show that the two locations have distinct profiles. Namely, HH has much higher levels of COND, NO_3, and TDS, whereas BQ has greater light penetration (Secchi depth) and deeper site depths (Depth) on average, with the exception of YC in HH which has the deepest average site depth overall. Additionally, the highest average TP levels are seen in both HH and BQ, particularly in BF, NP, and PH. Trend plots showing the changes in WQ parameters over time can be seen in Figures C4-C11.

The PCA performed on the WQ parameters of the eight study sites explained 50.40% of the variation observed (Figure 3.2). The PCA also shows a clear separation in the WQ profiles of the two locations. The sites in HH are more strongly associated with higher levels of NO_3, COND, and TDS. Conversely, the WQ profiles of the sites in BQ are weakly

associated with NO_3, COND, and TDS, but have higher secchi depth readings than the sites in HH. The three sites of HH overlap, indicating that they have similar WQ profiles. Similarly, the five sites in BQ also overlap, with HB, GL, and WC in particular having similarly-size ellipsoids and clustering even more, demonstrating that they are more similar to each other than they are to PH and NP. There is also a TP gradient present for three of the eight sites: BF, PH, and NP. The three aforementioned sites have more stretched-out ellipsoids in the direction of the TURB and TP eigenvectors, indicating a stronger association with these two WQ parameters. Also, some of the weeks sampled in BF, PH, and NP seem to have lower levels of DO and pH.

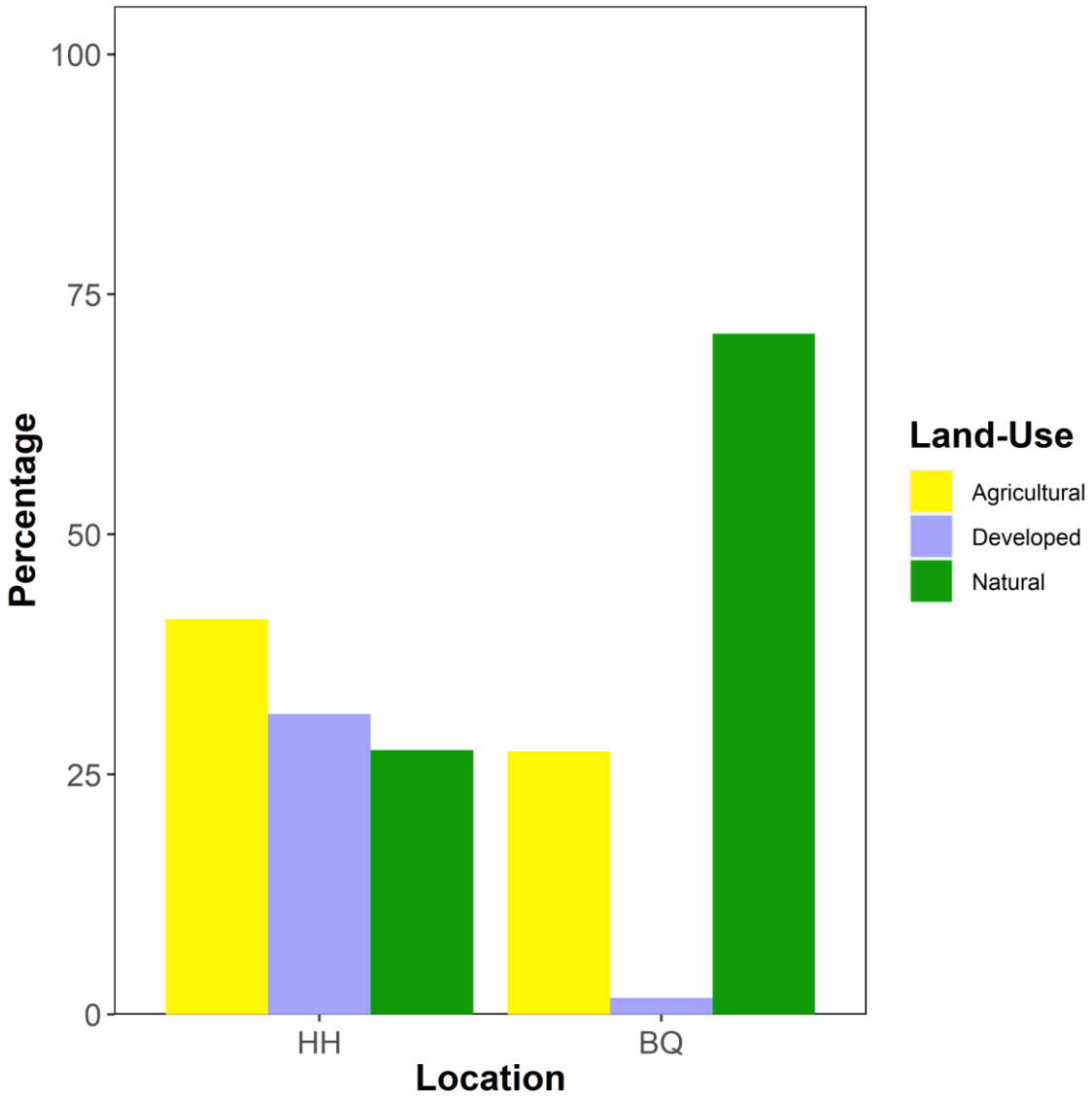

Figure 3.1. Land-use profiles for each of the two study locations. Land-use is shown as proportion of the total watershed area.

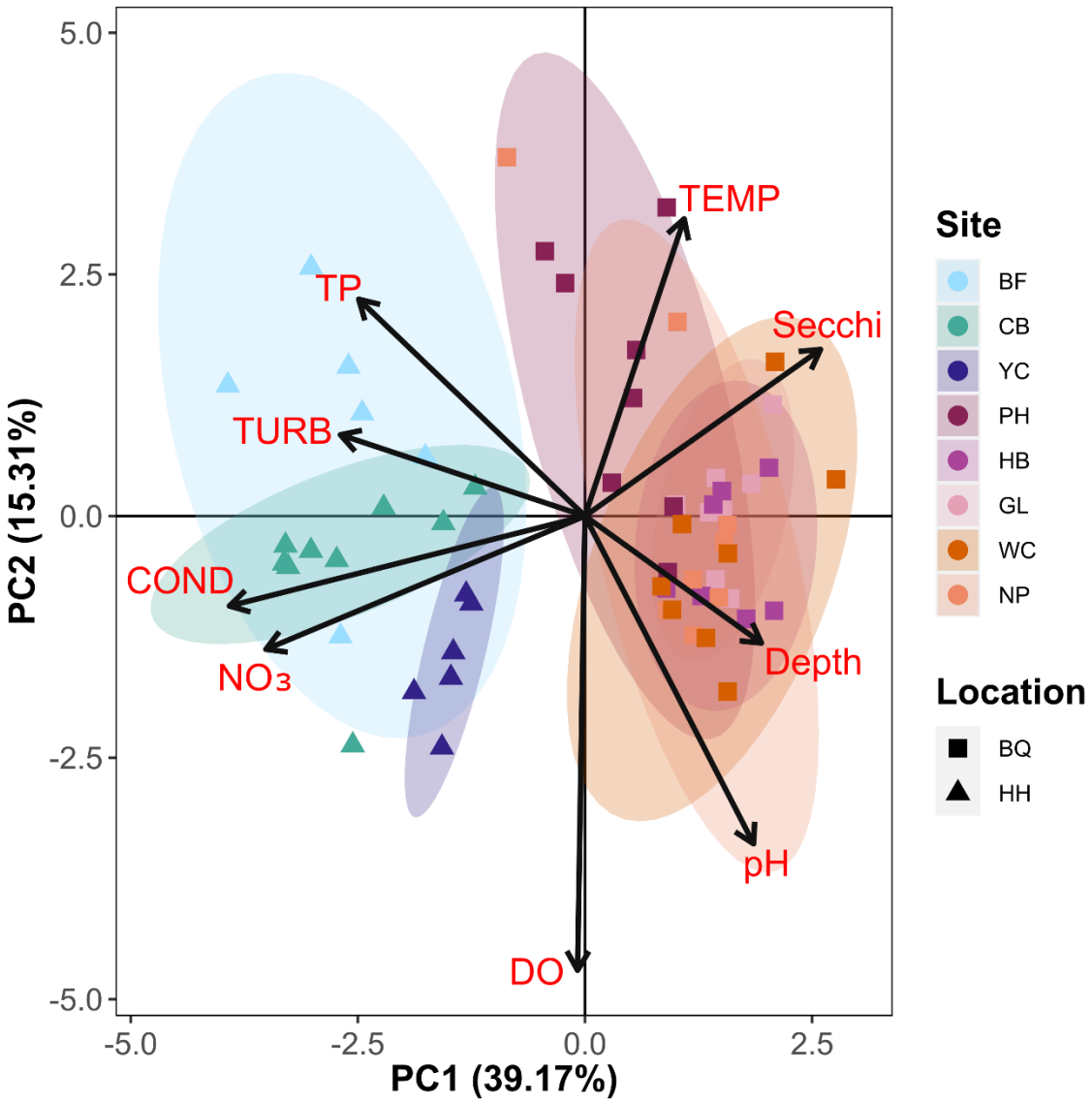

Figure 3.2. PCA displaying WQ parameters as eigenvectors and the study sites as points. Triangles represent the sites of HH and squares represent the sites of BQ. Each point represents one week from August to October in 2020

Table 3.1. Watershed land cover profiles for the two Lake Ontario AOCs, in addition to area and percent land-use. The total watershed area of HH is 52,847.8 ha and the total area of the BQ watershed is 1,864,474.0 ha.

Location	Land Cover	Area (ha)	Percent Land-Use (%)
HH	Agricultural	20,662.4	39.1%
	Developed	15,710.6	29.7%
	Natural	13,816.8	26.1%
	Water	2,658.1	5.0%
BQ	Agricultural	463,323.3	24.9%
	Developed	28,748.8	1.5%
	Natural	1,199,086.9	64.3%
	Water	173,315.	9.3%

Table 3.2. Mean and standard deviation of WQ parameters measured in HH. Sample sizes are as follows: BF = 6, CB = 9, YC = 6.

WQ Parameter	BF	CB	YC
Secchi (m)	1.08 ± 0.24	0.69 ± 0.17	1.24 ± 0.28
Depth (m)	2.28 ± 0.50	1.53 ± 0.44	7.93 ± 0.82
TEMP (°C)	18.3 ± 2.51	20.06 ± 3.20	18.72 ± 1.87
DO (ppm)	7.27 ± 1.68	8.18 ± 1.27	8.62 ± 0.71
pH	7.6 ± 0.56	7.84 ± 0.49	7.99 ± 0.43
TURB	0.062 ± 0.08	0.041 ± 0.07	0.0083 ± 0.0077
NO_3 (mg/L)	11.73 ± 3.59	12.22 ± 2.93	13.93 ± 3.31
COND (µs/cm)	462.5 ± 39.38	475.11 ± 67.25	451 ± 45.93
TP (µg/L)	335.81 ± 213.74	47.55 ± 14.16	72.76 ± 58.29

Table 3.3. Mean and standard deviation of WQ parameters measured in BQ. The sample size for each site is 8.

WQ Parameter	GL	HB	NP	PH	WC
Secchi (m)	1.51 ± 0.27	1.24 ± 0.35	1.01 ± 0.11	1.16 ± 0.15	1.31 ± 0.64
Depth (m)	6.26 ± 1.13	6.04 ± 1.24	4.15 ± 1.06	2.15 ± 0.48	2.87 ± 0.45
TEMP (°C)	20.44 ± 2.93	20.6 ± 3.25	20.53 ± 3.62	20.08 ± 2.99	19.9 ± 3.58
DO (ppm)	7.61 ± 0.61	8.16 ± 0.82	8.34 ± 0.75	6.88 ± 2.02	8.48 ± 1.13
pH	7.94 ± 0.44	8.22 ± 0.46	8.27 ± 0.48	7.37 ± 0.49	8.43 ± 0.37
TURB	0.0047 ± 0.002	0.0066 ± 0.0014	0.041 ± 0.09	0.02 ± 0.034	0.0054 ± 0.0049
NO_3 (mg/L)	0.69 ± 0.34	0.44 ± 0	0.69 ± 0.34	0.99 ± 0.59	0.69 ± 0.34
COND (µs/cm)	257.13 ± 14	239.75 ± 25.5	230.13 ± 17.37	269.63 ± 30.5	229.38 ± 15.17
TP (µg/L)	28.74 ± 7.44	41.99 ± 13.91	165.59 ± 219.87	105.66 ± 167.99	26.24 ± 18.95

3.2 Phytoplankton Communities in Hamilton Harbour and Bay of Quinte

Phytoplankton relative abundance based on average biomass (mg/L) for seven classes of phytoplankton in HH and BQ are shown in Figure 3.3 and Figure 3.4, respectively. In HH, dinoflagellates dominate the majority of the weeks sampled, particularly in CB and YC, with diatoms dominating the phytoplankton composition on Sept 15 in YC and Oct 05 in CB and YC. BF has a much more varied composition of phytoplankton classes compared to CB and YC, with a larger abundance of greens, diatoms, and cyanobacteria, with the latter dominating Oct 05. In contrast, diatoms appear to dominate the majority of the weeks sampled and dinoflagellates are not seen with large abundances as often in BQ. Namely, dinoflagellates are seen earlier in the 10 weeks of sampling in GL and PH, with the phytoplankton group dominating Aug 26 in HB. Cyanobacteria dominate more weeks in BQ, particularly Aug 26 in PH, Aug 13 in HB, Aug 20 in GL and NP, and Sept 10 in HB.

The summary statistics of the proportion of phytoplankton biomass (mg/L) in HH and BQ for all weeks of phytoplankton microscopy are shown in Table 3.4 and Table 3.5, respectively. The phytoplankton that dominate BF in terms of biomass are split between cyanobacteria, dinoflagellates, diatoms, and greens. CB and YC are both mainly dominated by dinoflagellates over time. Diatoms take up the largest amount of biomass over time in all five BQ sites. Throughout all eight sites, chrysophytes, cryptophytes, and euglenids contributed very little proportionally to phytoplankton biomass overall.

The RDA conducted on the biomass of phytoplankton (mg/L) and several WQ parameters can be found in Figure 3.5. The RDA shows that dinoflagellates are strongly associated

with COND and NO_3, while negatively associated with Secchi depth. The opposite is true for diatoms, greens, and cyanobacteria as the RDA shows that these three phytoplankton classes are negatively associated with NO_3 and COND yet positively related to Secchi depth. Cyanobacteria also appear to be positively associated with several other WQ parameters, including TP, TURB, TEMP, site depth, and pH. Cryptophytes, euglenids, and greens all appear to have a negative association with the aforementioned WQ parameters. Chrysophytes appear to not have any strongly positive or negative associations with any of the WQ parameters of this study. The "permutest()" function from the vegan R package (Oksanen *et al.*, 2020) identified Secchi depth, site depth, TURB, and NO_3 to be significant.

The Kendall correlation matrix in Figure 3.6 that shows the relationships between the WQ parameters of this study and the biomass (mg/L) of the families of the seven major classes of phytoplankton. Numerous positive relationships are present among many phytoplankton families and both COND and NO_3. Several significant negative relationships exist between pH and several families of phytoplankton among the greens and diatoms. According to the Kendall correlation matrix, few significant relationships exist between phytoplankton families and Secchi depth, site depth, TURB, and DO.

The MLR performed on the biomass of phytoplankton (mg/L) and the WQ parameters (Table 3.7) indicates several significant relationships. No significant relationships between WQ and euglenids were detected, so that class of phytoplankton were omitted from the table. COND was removed because it was found to be collinear with NO_3. NO_3 was found to have a significant relationship with all other classes of phytoplankton, except for

cyanobacteria. No significant relationships were present between phytoplankton biomass and Secchi depth or TP. Dinoflagellates had the most significant relationships with WQ parameters than all other classes of phytoplankton. Specifically, the biomass of dinoflagellates was found to be significant related with TEMP, DO, pH, TURB, and NO_3.

There are marked differences in the diversity of phytoplankton between HH and BQ, as shown by the Simpson's Diversity Indices based off of genus diversity (Table 3.6) and the corresponding line graphs (Figure 3.7 and Figure 3.8). HH has many more sites above 0.85 on the Simpson's Diversity Index scale than the sites in BQ. Almost half of the weeks among all of the BQ sites are below 0.70, with all weeks in WC below 0.70. One week in HH is below 0.70: CB on week 1. YC has the highest diversity values among all sites and weeks, with the highest Simpson's Diversity Index being 0.92. A MLR was conducted to relate WQ parameters with the Simpson's Diversity Indices. NO_3 was removed due to being collinear with COND (Table 3.8). The only WQ parameter that was shown to have a significant relationship with Simpson's Diversity Index was COND.

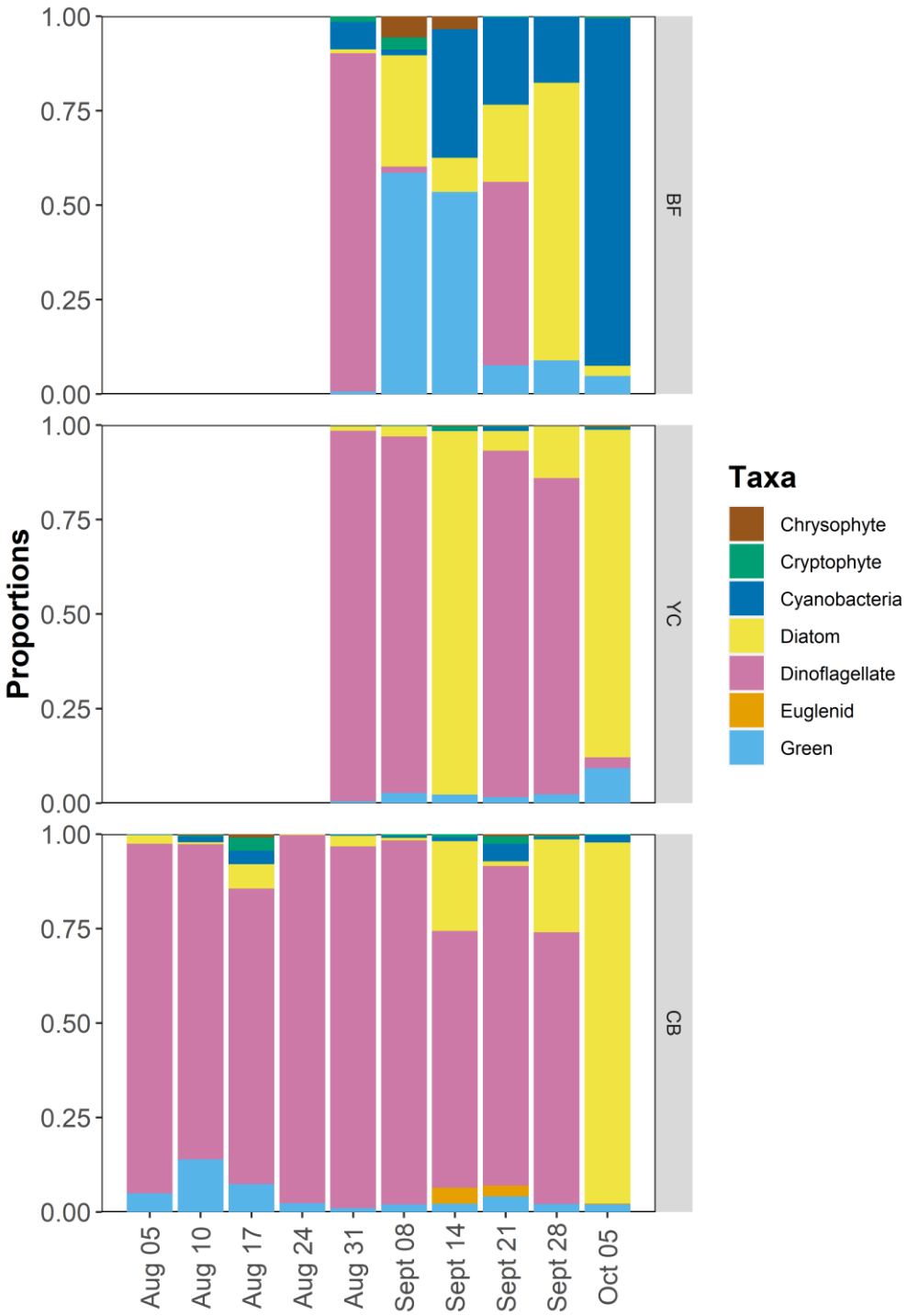

Figure 3.3. Relative abundance plot of seven major phytoplankton classes in HH calculated using weekly biomass (mg/L) totals.

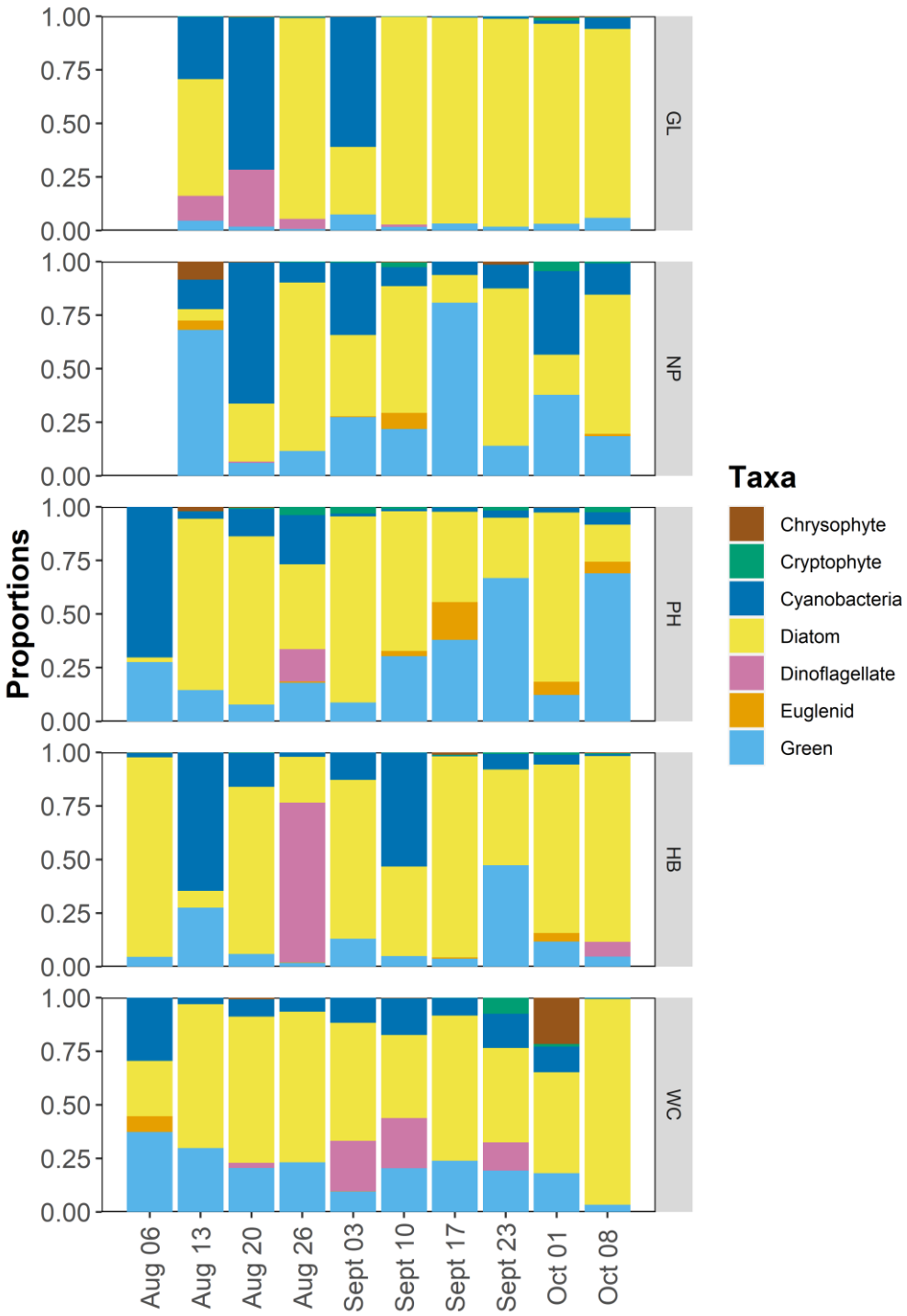

Figure 3.4. Relative abundance plot of seven major phytoplankton classes in BQ calculated using weekly biomass (mg/L) totals.

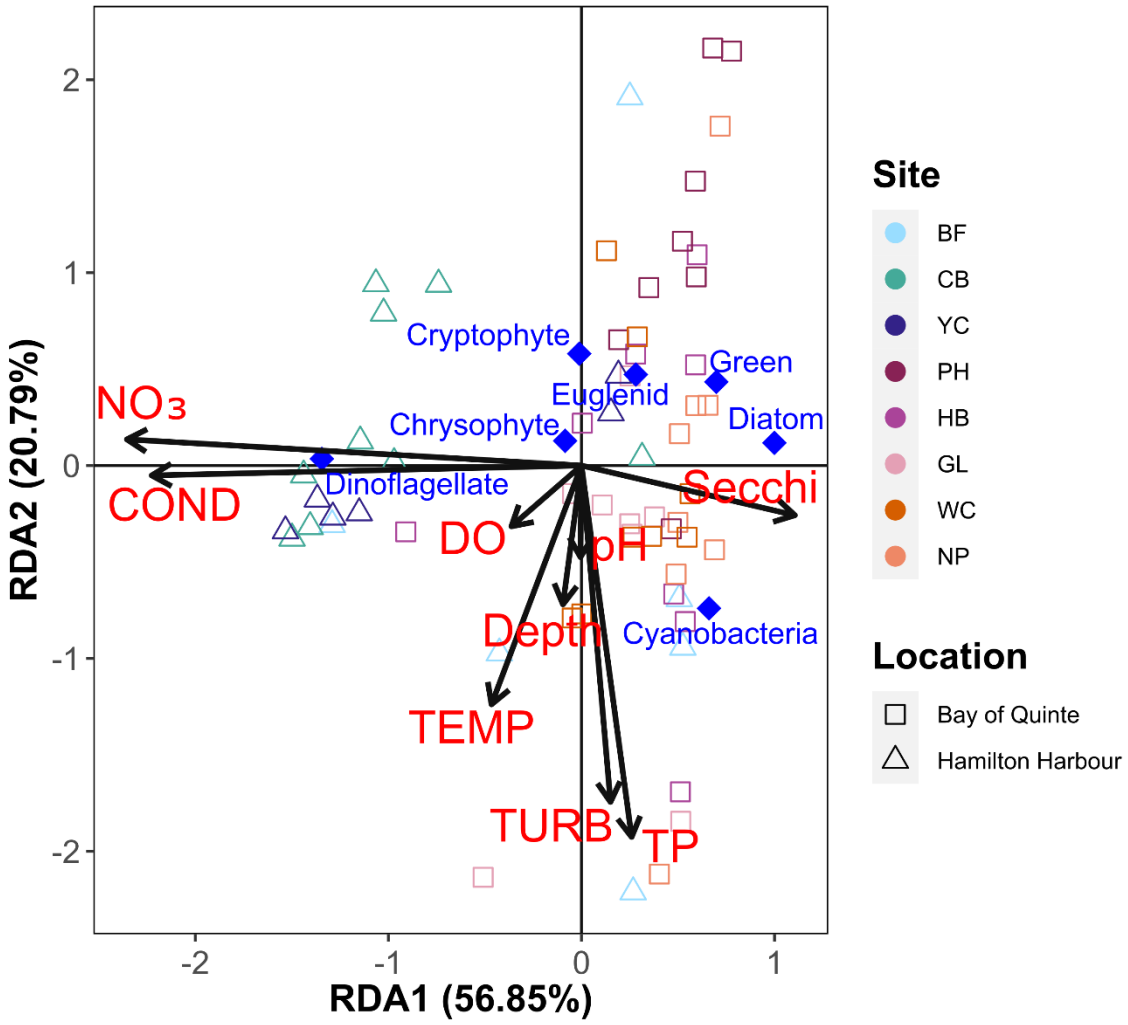

Figure 3.5. RDA of seven major phytoplankton classes in relation to WQ parameters. Sites are represented by the coloured diamonds. The "permutest()" function from the vegan R package (Oksanen *et al.*, 2020) indicates that Secchi depth, site depth, and TURB are significant to a 0.05 level and NO_3 is significant to a 0.001 level.

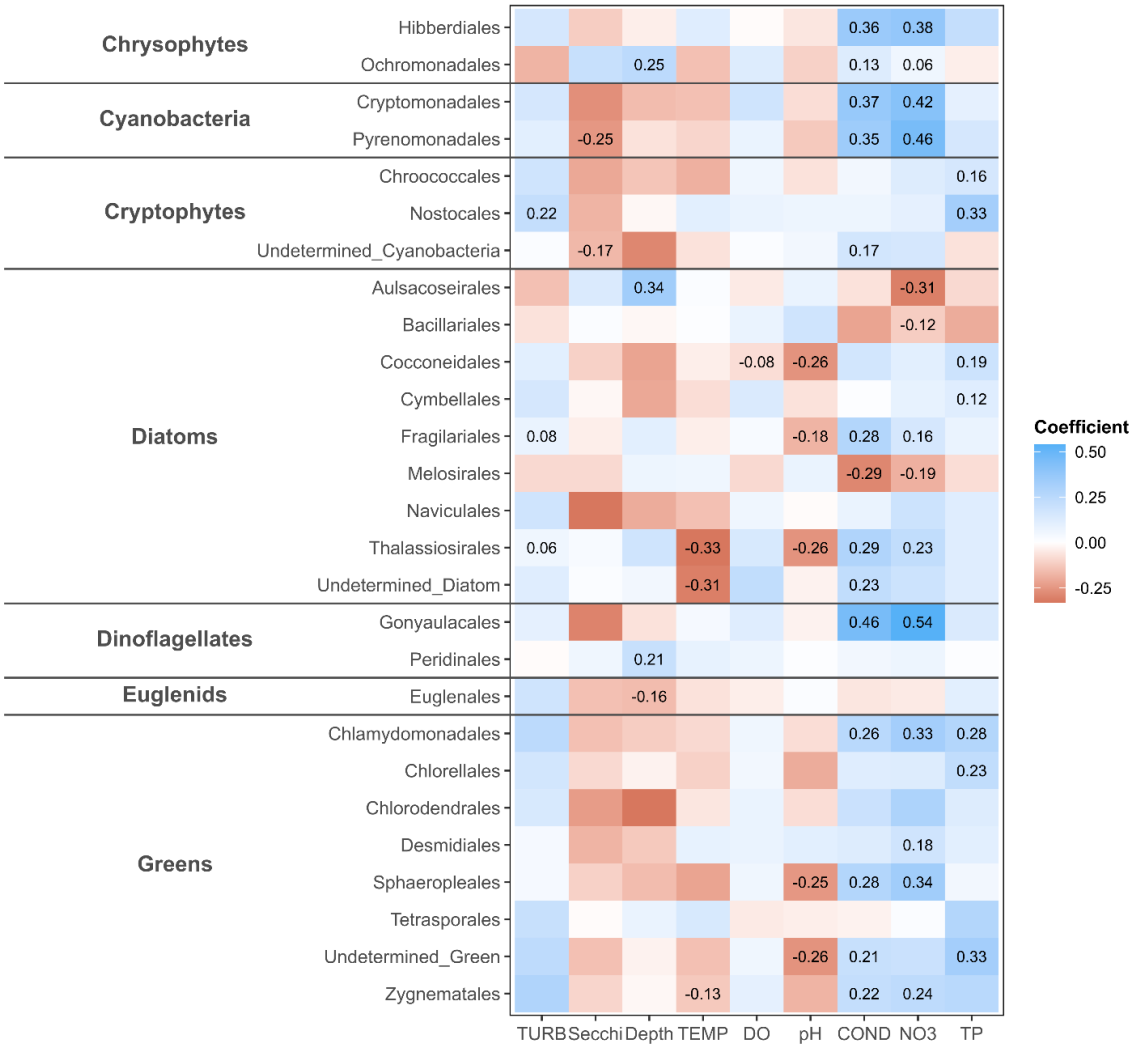

Figure 3.6. A Kendall correlation matrix displaying the relationships between WQ parameters and phytoplankton families from seven major groups of phytoplankton. Only coefficients with a significant negative or positive relationship (p-value < 0.05) are displayed in the coloured boxes.

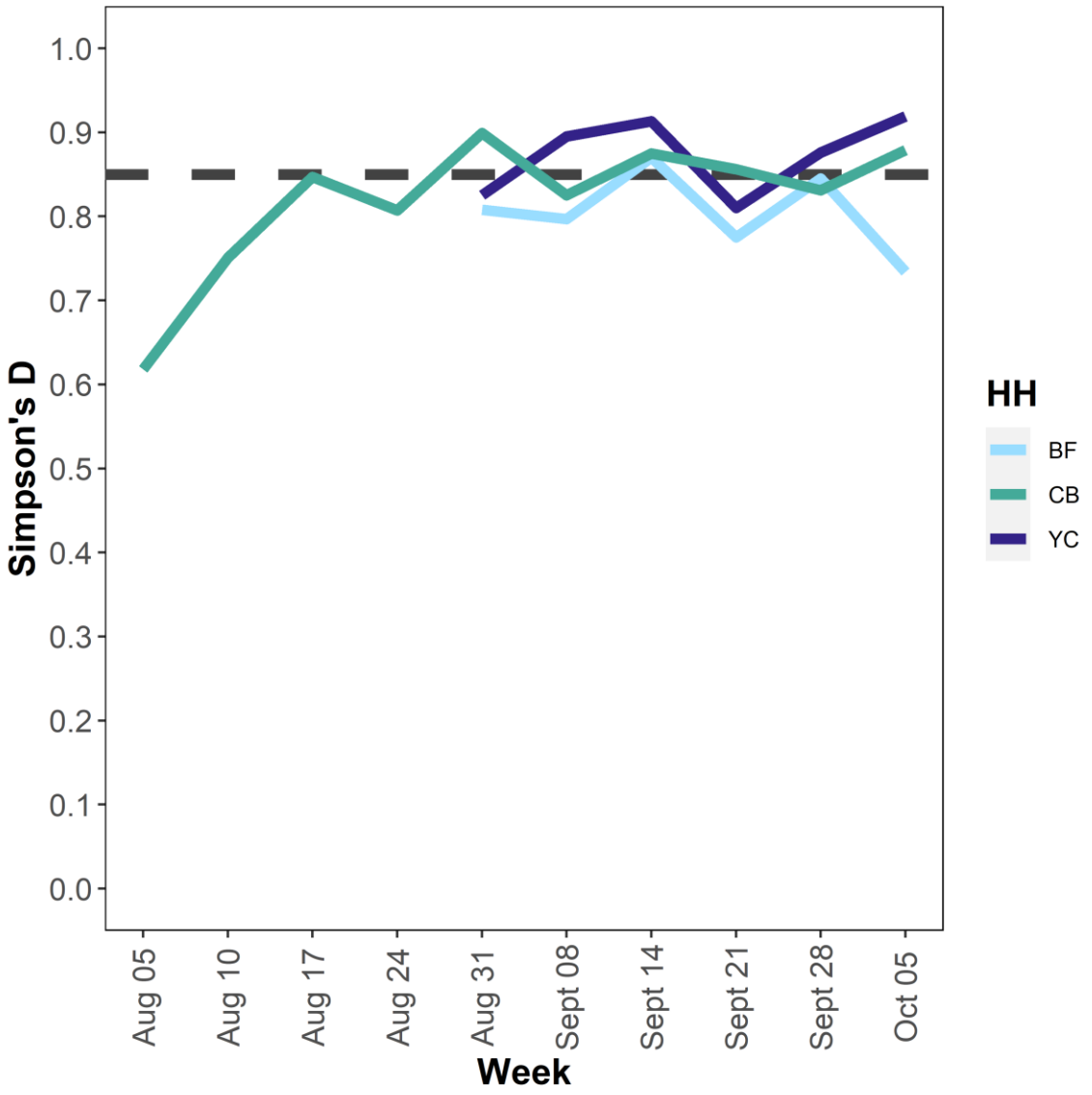

Figure 3.7. Weekly Simpson's Diversity indices for the three sites in HH plotted over the ten weeks from August to October, 2020. The dashed line at 0.85 indicates an arbitrary cut-off of markedly higher diversity (>0.85) and markedly lower diversity (<0.85). No data is available for BF and YC from August 05 to August 24, 2020.

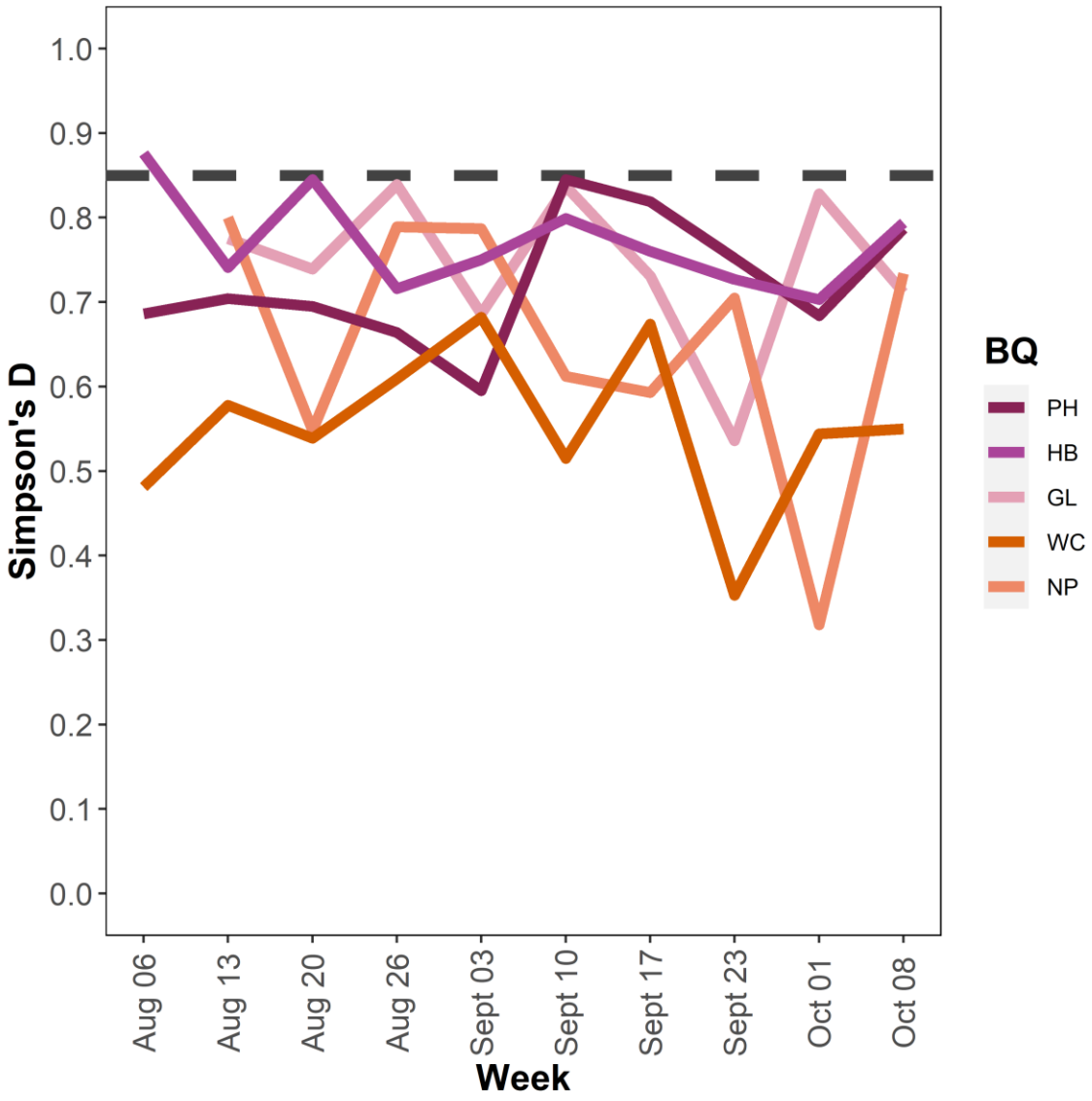

Figure 3.8. Weekly Simpson's Diversity indices for the five sites in BQ plotted over the ten weeks from August to October, 2020. The dashed line at 0.85 indicates an arbitrary cut-off of markedly higher diversity (>0.85) and markedly lower diversity (<0.85). No data is available for GL and NP on August 06, 2020.

Table 3.4. Summary of the phytoplankton biomasses expressed as proportions over the sampling period in HH. Sample sizes are as follows: BF = 6, CB = 10, YC = 6.

Phytoplankton Group	**BF**			**CB**			**YC**		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Chrysophyte	1.5	0.0	5.6	0.2	0.0	0.9	0.2	0.0	0.4
Cryptophyte	0.9	0.0	3.2	0.8	0.0	3.5	0.3	0.0	1.1
Cyanobacteria	29.2	1.5	92.0	1.4	0.0	4.6	0.4	0.0	1.1
Diatom	22.7	1.0	73.5	15.8	0.2	95.5	34.3	1.2	96.2
Dinoflagellate	23.3	0.0	89.5	76.8	0.0	97.4	61.8	0.0	98.1
Euglenid	0.0	0.0	0.0	0.7	0.0	4.2	0.0	0.0	0.0
Green	22.4	0.7	58.6	4.2	1.0	13.9	3.1	0.5	9.2

Table 3.5. Summary of the phytoplankton biomasses expressed as proportions over the sampling period in BQ. Sample sizes are as follows: GL = 9, HB = 10, NP = 9, PH = 10, WC = 10.

Phytoplankton Group	GL			HB			NP			PH			WC		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Chrysophyte	0.2	0.0	0.7	0.2	0.0	1.1	1.2	0.0	8.5	0.2	0.0	2.0	2.3	0.0	21.7
Cryptophyte	0.3	0.0	1.4	0.4	0.0	1.1	0.9	0.0	4.5	1.3	0.0	3.9	0.9	0.0	7.5
Cyanobacteria	18.9	0.2	70.9	16.3	0.4	64.5	22.5	6.2	65.9	12.5	0.9	70.3	11.3	0.6	29.5
Diatom	72.3	0.0	97.0	62.0	7.9	93.8	42.0	5.3	78.7	51.9	2.1	86.8	58.0	25.8	96.1
Dinoflagellate	4.9	0.0	26.6	8.1	0.0	74.5	0.1	0.0	0.6	1.5	0.0	15.1	6.3	0.0	23.6
Euglenid	0.0	0.0	0.0	0.5	0.0	4.0	1.5	0.0	7.5	3.2	0.0	17.6	0.7	0.0	7.3
Green	3.4	0.7	7.5	12.5	1.8	47.3	31.9	6.1	80.8	29.3	7.7	69.0	20.5	3.3	37.3

Table 3.6. Multiple linear regression for the biomass of six phytoplankton classes and WQ parameters. Euglenids showed no significant relationships with any WQ parameters and were removed. All phytoplankton groups shown had significant models. All WQ parameters with significant relationships to biomass are indicated in bold, non-significant relationships are indicated by "N.S."

Phytoplankton Group	**Secchi Depth**	**Depth**	**TEMP**	**DO**	**pH**	**TURB**	**TP**	**NO_3**	**Model R^2**	**Adj R^2**
Chrysophyte	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	**0.034**	1	<0.001
Cryptophyte	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	**<0.001**	0.519	0.456
Cyanobacteria	N.S.	N.S.	N.S.	N.S.	N.S.	**<0.001**	**<0.001**	N.S.	0.675	0.625
Diatom	N.S.	**0.011**	N.S.	N.S.	N.S.	N.S.	N.S.	**0.046**	0.239	0.138
Dinoflagellate	N.S.	N.S.	**<0.001**	**<0.001**	**<0.001**	**<0.001**	N.S.	**<0.001**	0.414	0.336
Green	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	**0.03**	0.137	0.023

Table 3.7. Simpson's Diversity Index summary table for the eight sites of this study. All indices below 0.70 are highlighted in red and all indices above 0.85 are shown in blue.

Location	Site	Week									
		1	2	3	4	5	6	7	8	9	10
HH	BF	N/A	N/A	N/A	N/A	0.81	0.80	**0.87**	0.77	0.84	0.73
	CB	**0.62**	0.75	0.85	0.81	**0.90**	0.83	**0.87**	**0.86**	0.83	**0.88**
	YC	N/A	N/A	N/A	N/A	0.82	**0.90**	**0.91**	0.81	**0.88**	**0.92**
BQ	GL	N/A	0.77	0.74	0.84	**0.69**	0.84	0.73	**0.54**	0.83	0.71
	HB	**0.88**	0.74	0.85	0.72	0.75	0.80	0.76	0.73	0.70	0.79
	NP	N/A	0.80	**0.55**	0.79	0.79	**0.61**	**0.59**	0.71	**0.32**	0.73
	PH	**0.69**	0.70	**0.70**	**0.66**	**0.60**	0.85	0.82	0.75	**0.68**	0.79
	WC	**0.48**	**0.58**	**0.54**	**0.61**	**0.68**	**0.51**	**0.67**	**0.35**	**0.54**	**0.55**

Table 3.8. Multiple linear regression for Simpson's Diversity Index and WQ parameters. COND was removed due to being collinear with NO_3. All significant relationships are indicated in bold.

WQ Parameter	Estimate ± SE	p-val	Model R^2	Adj R^2
Depth	$1.32 * 10^{-2} \pm 6.78 * 10^{-3}$	0.057	0.408	0.317
TEMP	$1.68 * 10^{-3} \pm 5.77 * 10^{-3}$	0.772		
Secchi depth	$-3.82 * 10^{-2} \pm 4.29 * 10^{-2}$	0.376		
TURB	$-1.79 * 10^{-1} \pm 3.82 * 10^{-1}$	0.640		
pH	$-4.80 * 10^{-2} \pm 3.42 * 10^{-2}$	0.167		
DO	$7.35 * 10^{-3} \pm 1.56 * 10^{-2}$	0.640		
TP	$5.12 * 10^{-6} \pm 1.24 * 10^{-4}$	0.967		
NO_3	**$1.16 * 10^{-2} \pm 2.53 * 10^{-3}$**	**<0.001**		

3.3 Phytoplankton Metabolites

Metabolite relative abundance plots for 33 amino acids and metabolites can be seen in Figure 3.9 and Figure 3.10 for HH and BQ, respectively. Glycine is the dominant metabolite in HH in all three sites. Abundance of glycine over all weeks looks to be relatively similar between BF, YC, and CB. β-Alanine is present in all three sites in HH, with the greatest abundances in CB and YC on August 31st. Serine also has a larger abundance in HH; however, serine is seen mostly in YC and CB. All other metabolites are in smaller abundances throughout all three sites. BQ has a more diverse and mixed range of abundances among all of the metabolites. Metabolites such as β-alanine, ornithine, and taurine are seen in greater abundances in BQ than they are in HH. Glycine is seen in BQ in larger abundances, especially in GL and PH, but not as much throughout the weeks in NP. α-Aminobutyric acid (AABA) is seen in fairly great abundances in all five sites in BQ, with the most seen in PH, GL, and HB. The abundance of AABA in HB appears to increase over time, with the abundance peaking on October 8th. Similarly large abundances of AABA on October 8th can also be seen in PH and GL. Supplementary figures, including a PCA and dendrogram, indicating the differences in metabolite profiles between HH and BQ are in Appendix C.

Figure 3.11 shows an RDA for the metabolites and the biomass (mg/L) of seven major classes of phytoplankton. 9 metabolites were removed due to collinearity: glutamine, isoleucine, hydroxy-proline, ornithine, phenylalanine, serine, tyrosine, valine, and 1-methyl-histidine. The RDA shows that numerous metabolites are positively associated with the biomass of cyanobacteria, including: leucine, arginine, lysine, proline, and histidine.

On the other hand, it appears that hydroxy-proline is negatively associated with cyanobacteria but positively associated with chrysophytes. Anserine is strongly positively associated with the biomass of dinoflagellates but negatively associated with greens and diatoms. Conversely, citrulline, sarcosine, aspartate, γ-aminobutyric acid (GABA) and AABA are negatively associated with dinoflagellates but positively associated with greens and diatoms. The biomass of euglenids appear to have weak relationships with the majority of the metabolites.

The Kendall Correlation matrix in Figure 3.12 indicates many significant relationships between metabolites and WQ parameters. Fourteen metabolites have significant relationships with TP, twelve are positive, one (hydroxyproline) is negative, and one (isoleucine) is neither positive nor negative. Twelve metabolites are significantly related to COND, with only three being positive. Similarly, eleven of the twelve metabolites significantly related to COND are significantly related to NO_3, with d-hydroxylysine not having a significant relationship with NO_3. Eleven metabolites are positively related to TURB. Seven metabolites are significantly positively related to Secchi depth, with only one being negative. Six metabolites are significantly related to temperature, with three being positive and three being negative. Only asparagine has a significant negative relationship with DO and only citrulline has a significantly positive relationship with pH. No significant relationships exist between any metabolites and site depth.

The Kendall Correlation matrix in Figure 3.13 shows non-parametric relationships between seven phytoplankton groups and thirty-three metabolites. The biomass of cyanobacteria is

significantly related to a number of metabolites, including: d-hydroxylysine, histidine, lysine, methionine, and valine. The biomass of dinoflagellates is significantly related to four metabolites, including beta-alanine, methionine, phenylalanine, and tryptophan. There are several significant relationships between the biomass of cryptophytes and metabolites, including a strongly negative relationship with citrulline and sarcosine, and a few other relationships. There are other significantly negative relationships with green algae and asparagine, citrulline, AABA, and sarcosine. Few significant relationships exist between metabolites and the chrysophyte and euglenid phytoplankton groups. No significant relationships were found for the biomass of diatoms and any metabolites.

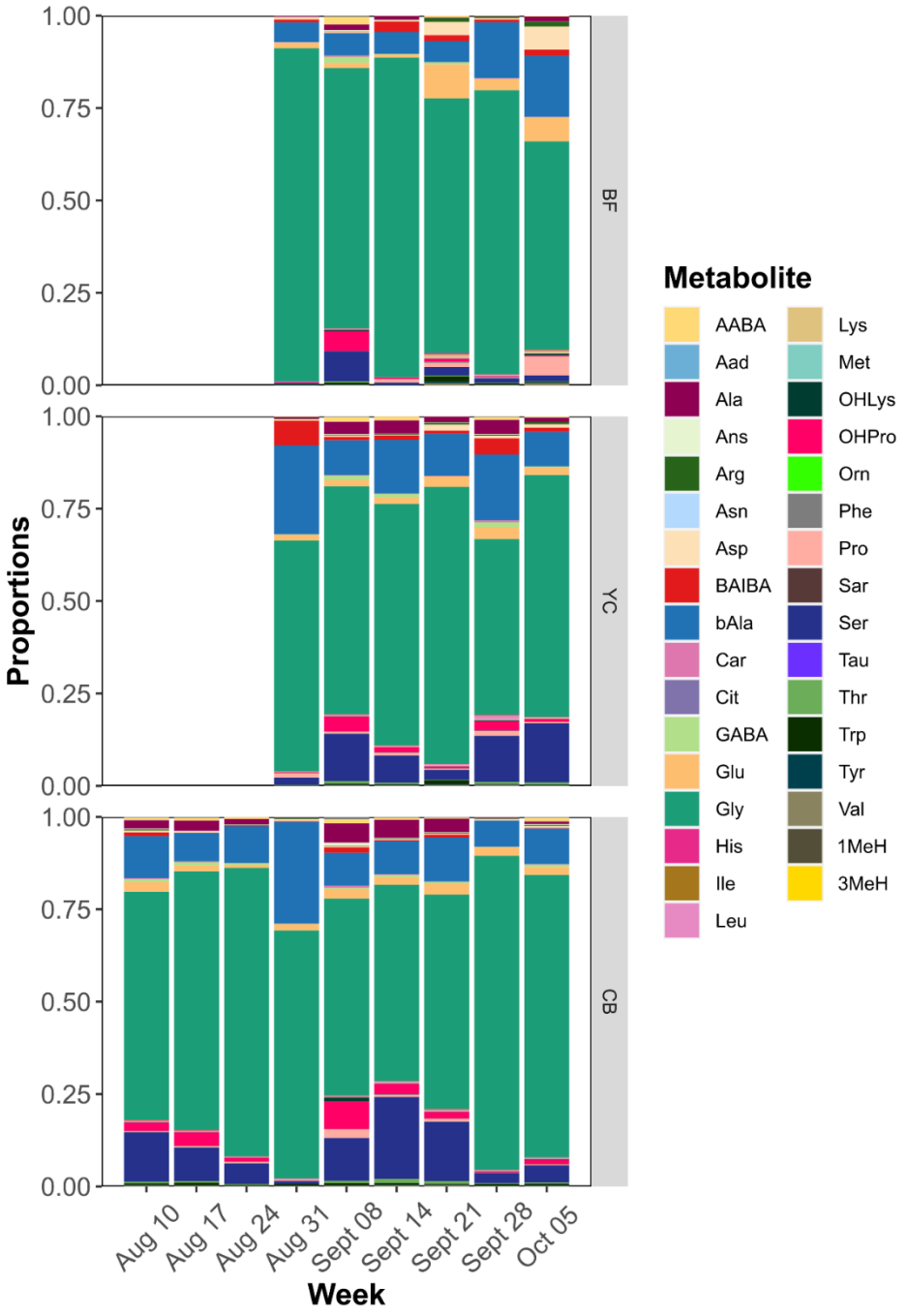

Figure 3.9. Relative abundance plot of 33 metabolites for HH calculated using weekly totals for HH.

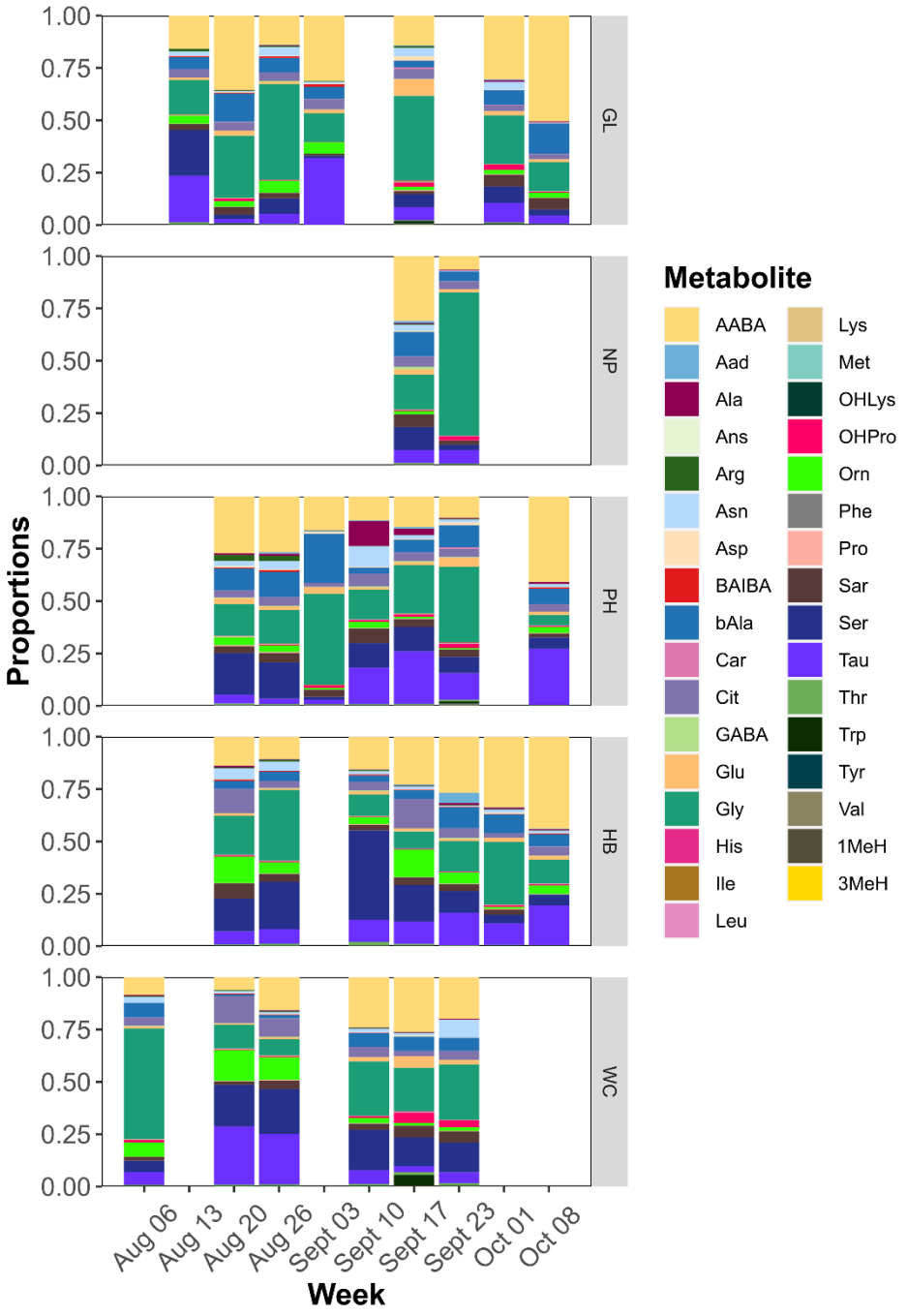

Figure 3.10. Relative abundance plot of 33 metabolites calculated using weekly totals for BQ.

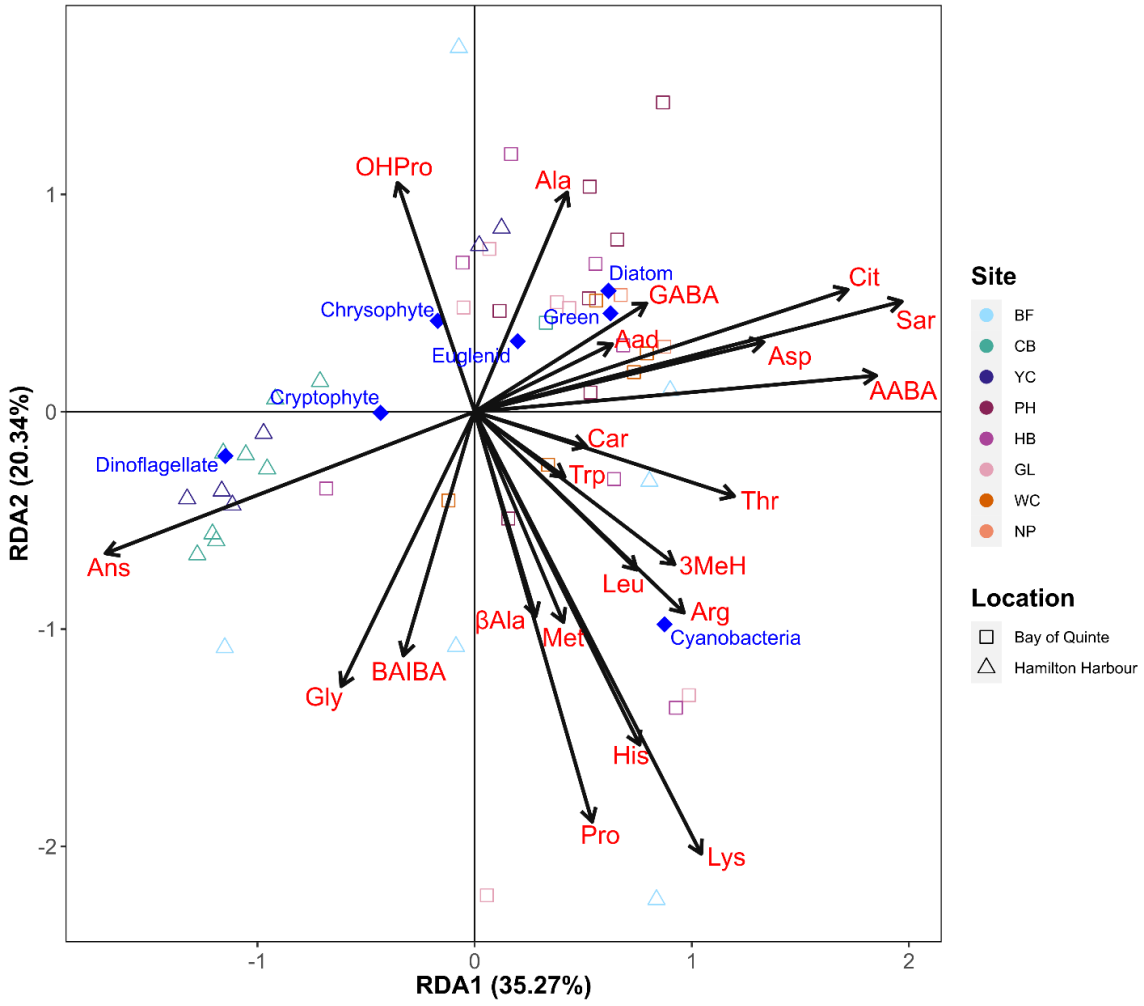

Figure 3.11. RDA of seven major phytoplankton classes in relation to 22 metabolites. The eigenvectors represent the metabolites and the coloured squares and triangles correspond to the sites and weeks sampled.

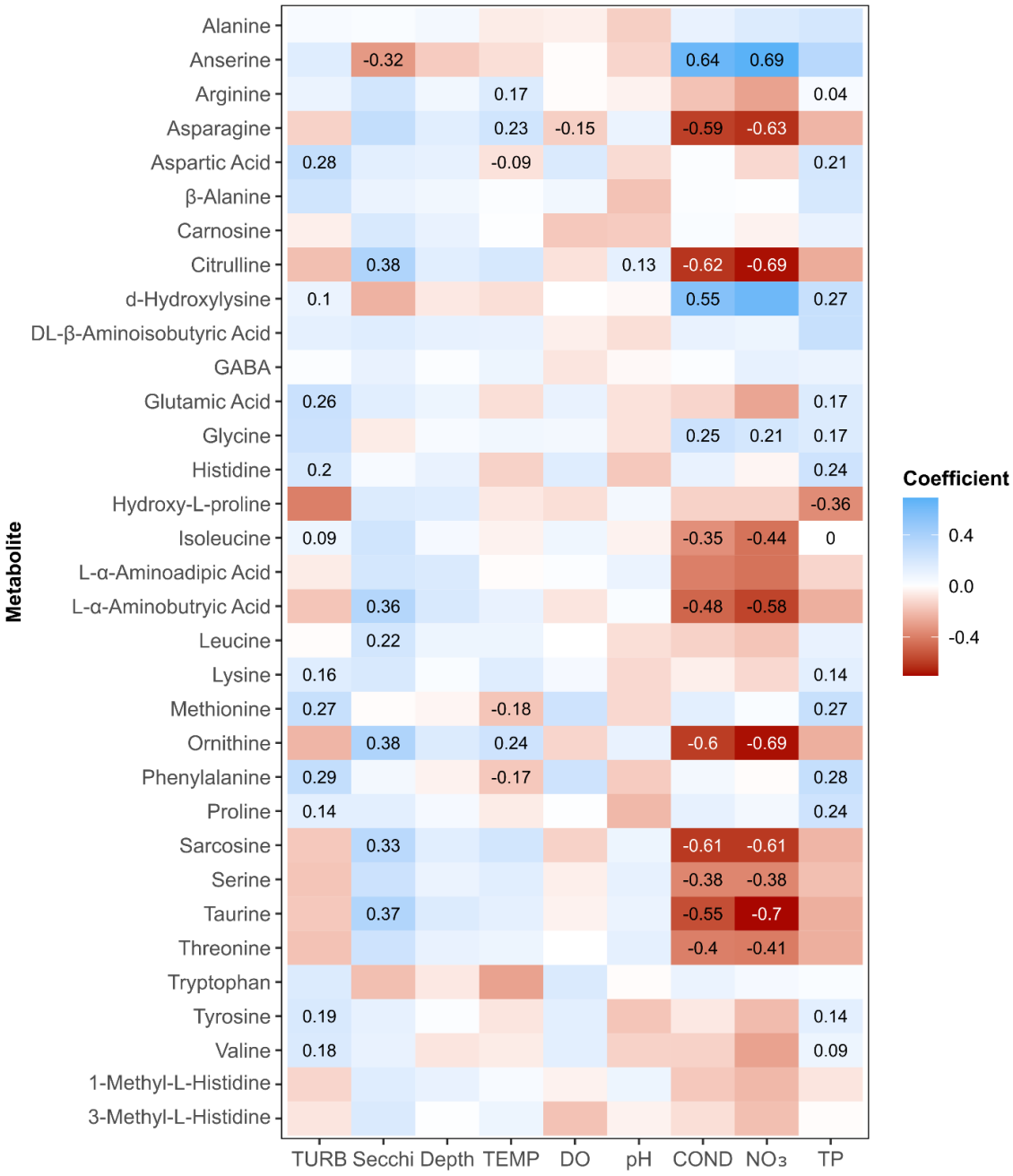

Figure 3.12. A Kendall correlation matrix displaying the relationships between WQ parameters and 33 metabolites. Only coefficients with a significant negative or positive relationship (p-value < 0.05) are displayed in the coloured boxes.

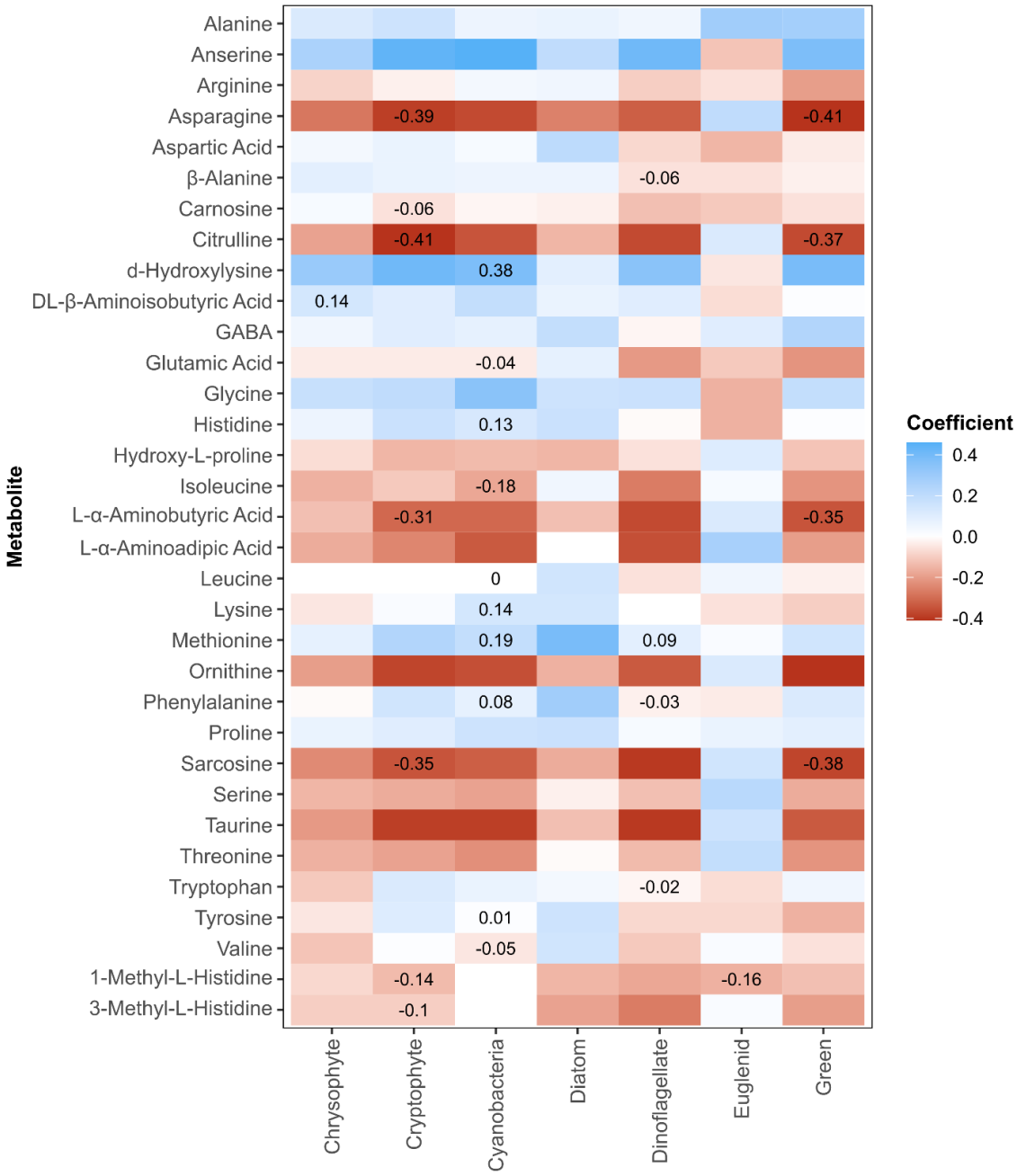

Figure 3.13. A Kendall correlation matrix displaying the relationships between 33 metabolites and phytoplankton families from seven major classes of phytoplankton. Only coefficients with a significant negative or positive relationship (p-value < 0.05) are displayed in the coloured boxes.

3.5 Phytoplankton Metabolic Pathways

In total, 33 pathways were found to be the most relevant to the metabolites found in HH and BQ. Table 3.10 summarizes each pathway, the metabolites found within, an adjusted p-value, and an impact rating on the percentage of matched metabolites to the pathway. The top four pathways with respect to impact and adjusted p-value are shown in Table 3.14 (Arginine Biosynthesis), Table 3.15 (Arginine and Proline Metabolism), Table 3.6 (Alanine, Aspartate, and Glutamate Metabolism), and Table 3.7 (β-Alanine Metabolism). The aforementioned figures indicate which metabolites from our study were apart of the pathway, and to what significance indicated by colour. One other additional pathway that had a high adjusted p-value, but a low impact score, is Glutathione Metabolism.

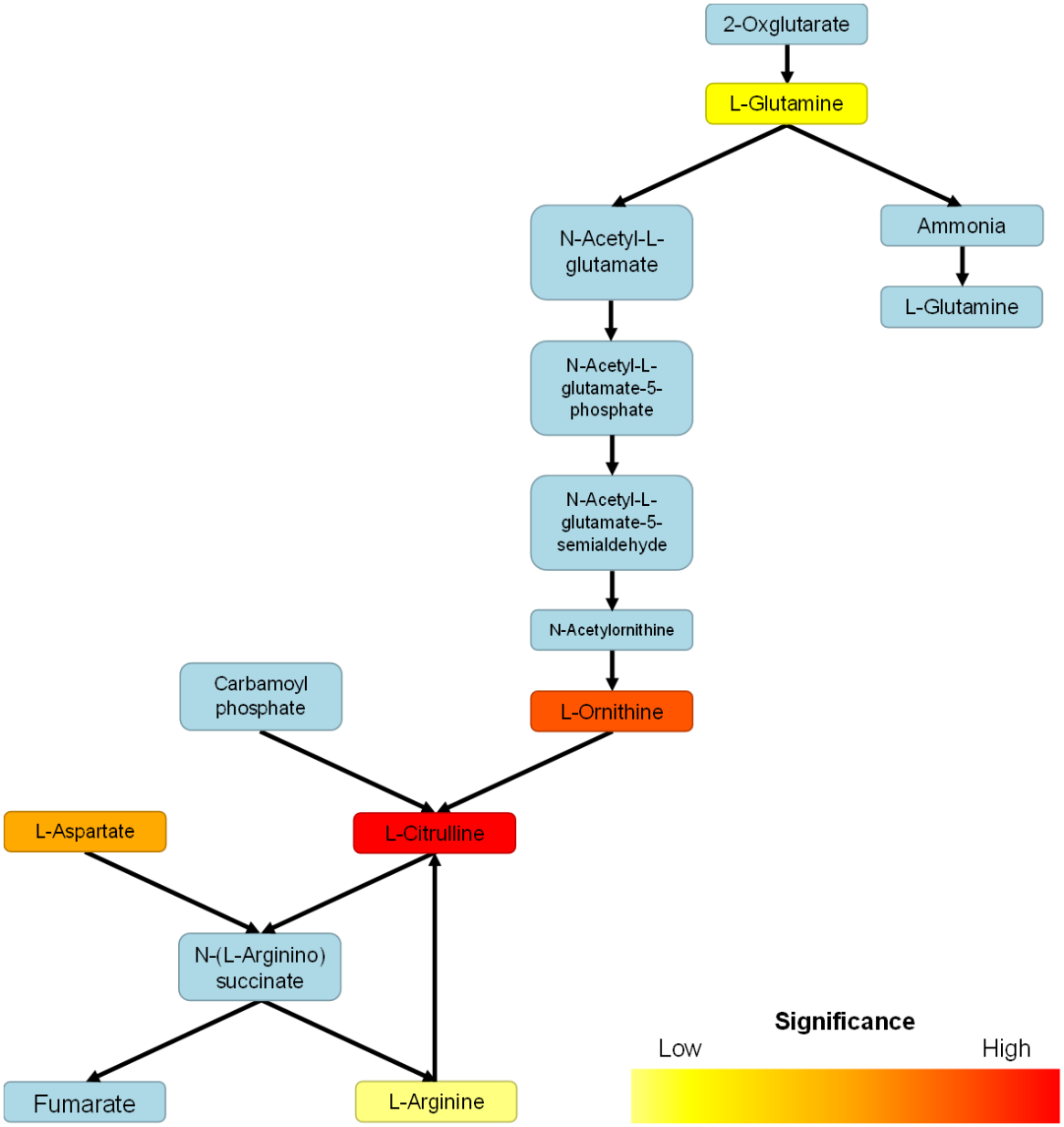

Figure 3.14. Pathway overview of Arginine Biosynthesis. Metabolites in light blue boxes were not apart of the metabolite dataset used in the pathway analysis, but are included as background for enrichment.

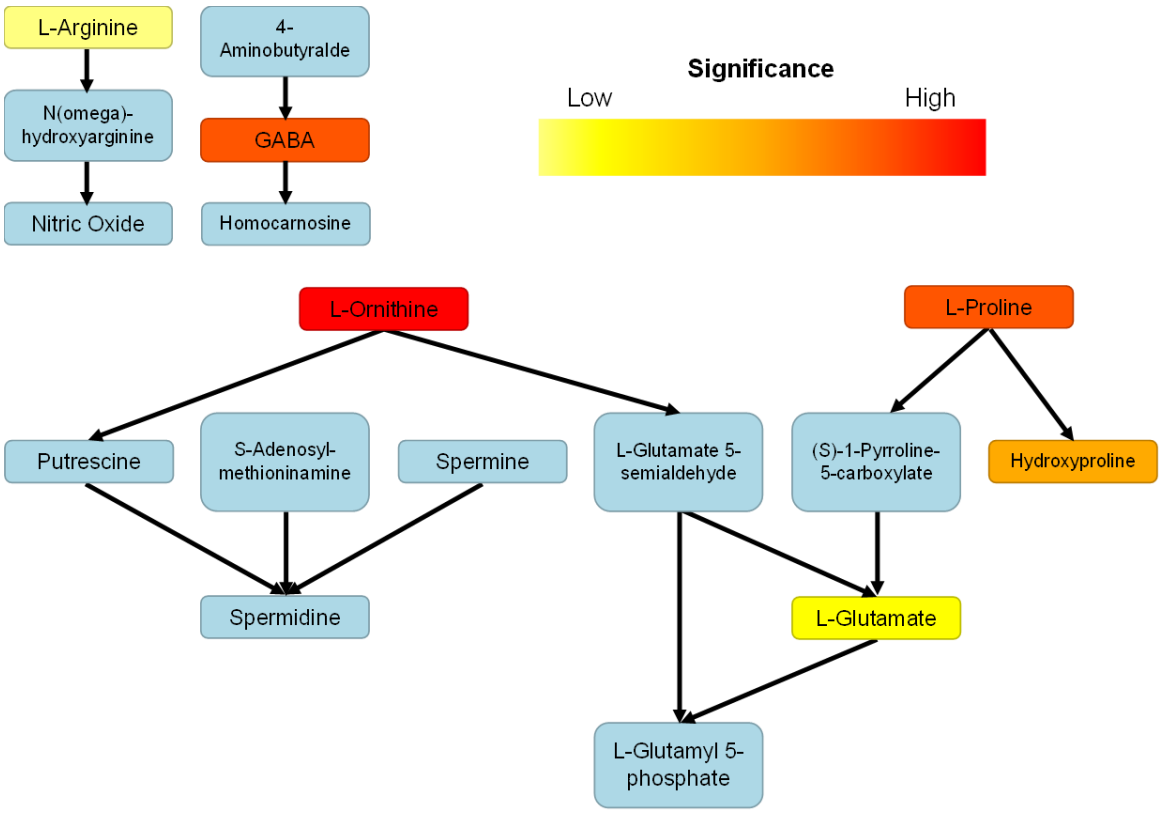

Figure 3.15. Pathway overview of Arginine and Proline Metabolism. Metabolites in light blue boxes were not apart of the metabolite dataset used in the pathway analysis, but are included as background for enrichment. Other coloured boxes represent different levels of significance of the metabolites in the dataset used, from yellow to red.

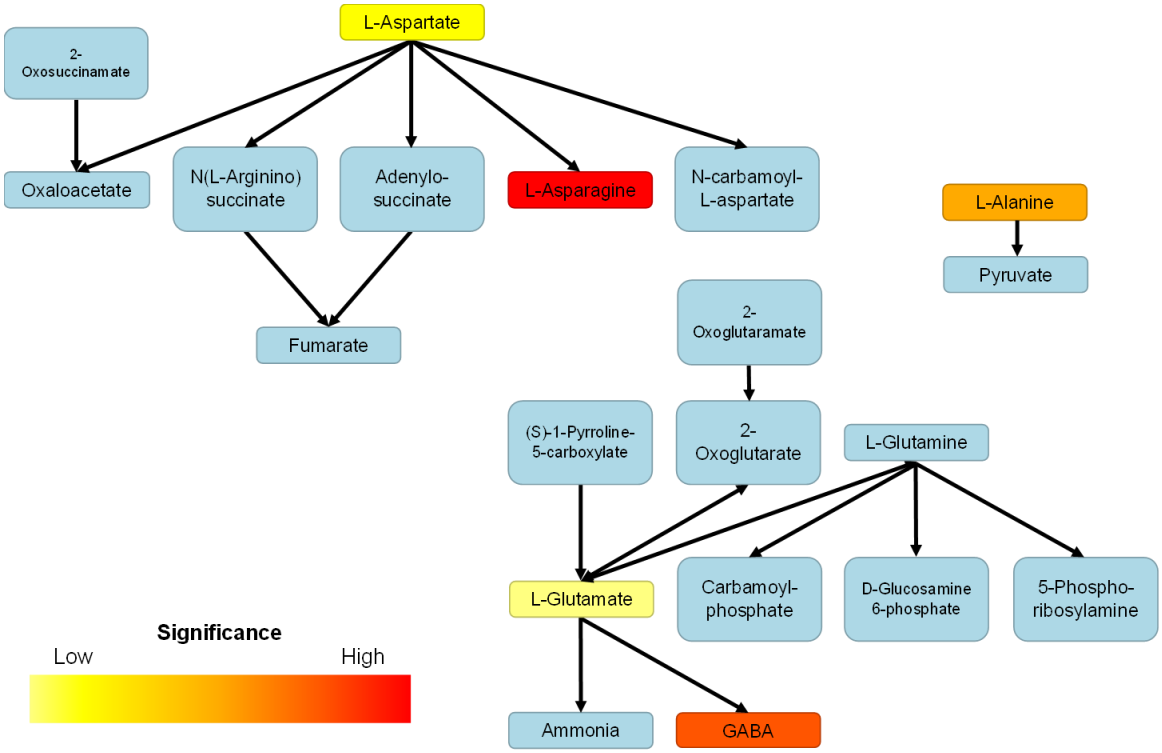

Figure 3.16. Pathway overview of Alanine, Aspartate, and Glutamate Metabolism. Metabolites in light blue boxes were not apart of the metabolite dataset used in the pathway analysis, but are included as background for enrichment. Other coloured boxes represent different levels of significance of the metabolites in the dataset used, from yellow to red.

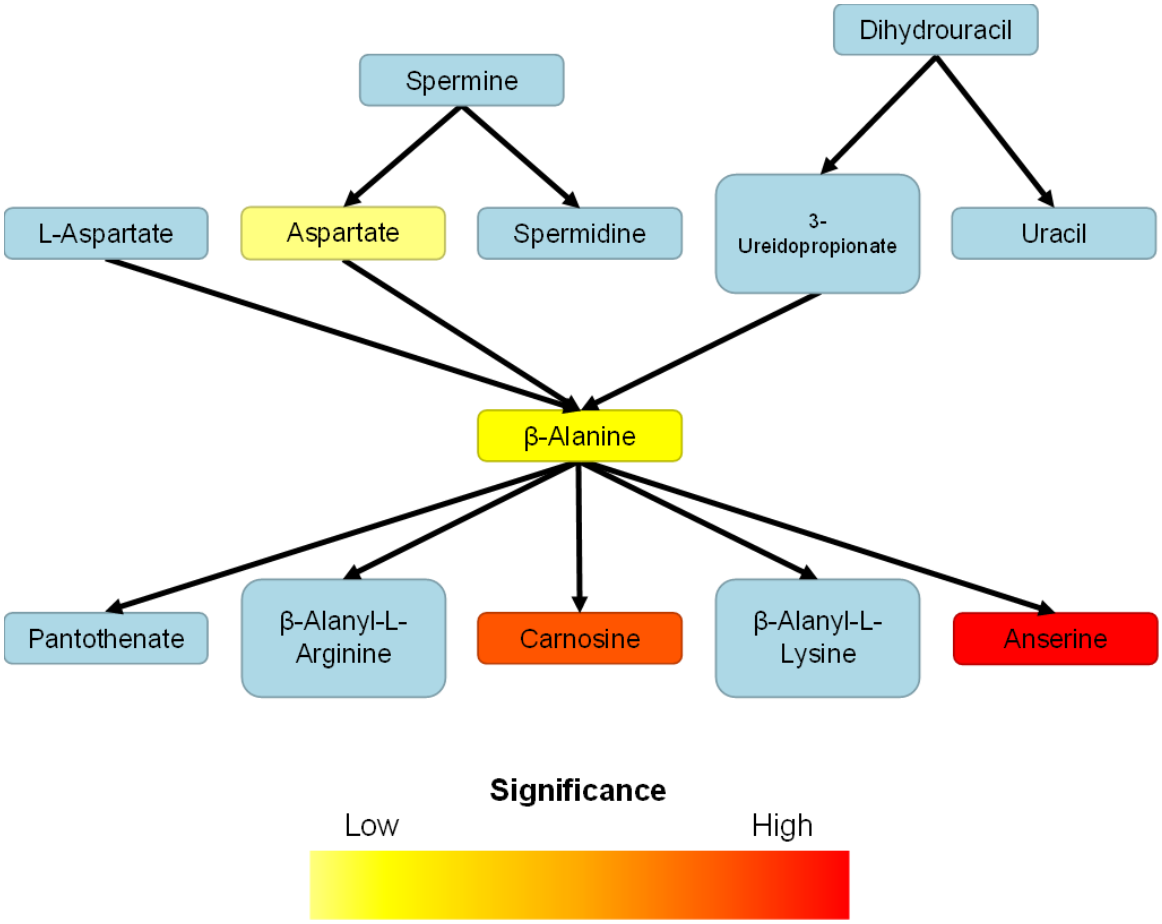

Figure 3.17. Pathway overview of β-Alanine Metabolism. Metabolites in light blue boxes were not apart of the metabolite dataset used in the pathway analysis, but are included as background for enrichment. Other coloured boxes represent different levels of significance of the metabolites in the dataset used, from yellow to red.

Table 3.9. Pathway analysis summary table. Match status represents how many metabolites matched to the number of metabolites in the pathway. $-\log_{10}(p)$ is the transformed p-value from the initial enrichment analysis. Impact is the pathway impact value calculated from the pathway topology analysis. Pathways in bold had the highest Impact and $-\log_{10}(p)$ values.

Pathway	Match Status	$-\log_{10}(p)$	Impact	Metabolites Involved in Pathway
Cysteine and methionine metabolism	3/43	20.80	0.19	Met, Asp, AABA
Glutathione metabolism	3/27	19.00	0.12	Gly, Glu, Orn
Arginine and proline metabolism	**6/31**	**18.90**	**0.45**	**Arg, GABA, Pro, Glu, Orn, OHPro**
beta-Alanine metabolism	**4/20**	**18.40**	**0.49**	**βAla, Asp, Car, Ans**
Arginine biosynthesis	**5/17**	**18.40**	**0.37**	**Glu, Arg, Cit, Asp, Orn**
Taurine and hypotaurine metabolism	1/7	17.10	0.2	Tau
Histidine metabolism	4/18	16.90	0.18	Car, His, 1MeH, Ans
Aminoacyl-tRNA biosynthesis	17/48	15.50	0	Asp, His, Phe, Arg, Gly, Asp, Met, Val, Ala, Lys, Ile, Leu, Thr, Trp, Tyr, Pro, Glu
Alanine, aspartate and glutamate metabolism	**5/20**	**15.10**	**0.53**	**Asp, Ala, Glu, Asp, GABA**
Phenylalanine metabolism	1/7	7.35	0.4	Phe
Tropane, piperidine and pyridine alkaloid biosynthesis	1/4	7.35	0	Phe
Cyanoamino acid metabolism	1/15	6.68	0.02	Gly
Thiamine metabolism	1/22	6.68	0.02	Gly
Glyoxylate and dicarboxylate metabolism	2/31	5.85	0.23	Gly, Glu
Glycine, serine and threonine metabolism	4/28	4.94	0.18	Asp, Gly, Thr, Try
Lysine biosynthesis	3/11	4.81	0.02	Asp, Aad, Lys
Lysine degradation	1/11	4.66	0.02	Aad
Butanoate metabolism	2/10	3.77	0.02	Glu, GABA
Valine, leucine and isoleucine biosynthesis	4/21	3.70	0	Thr, Leu, Ile, Val
Phenylalanine, tyrosine and tryptophan biosynthesis	3/22	3.57	0.02	Tyr, Trp, Phe
Isoquinoline alkaloid biosynthesis	1/5	3.12	0.6	Tyr
Tyrosine metabolism	1/18	3.12	0.39	Tyr
Betalain biosynthesis	1/4	3.12	0.02	Tyr
Valine, leucine and isoleucine degradation	3/37	3.12	0	Val, Ile, Leu
Carbon xation in photosynthetic organisms	2/21	2.89	0	Asp, Ala
Pantothenate and CoA biosynthesis	2/22	2.62	0.12	βAla, Val
Tryptophan metabolism	1/28	2.19	0.19	Trp
Selenocompound metabolism	1/16	2.17	0	Ala
Monobactam biosynthesis	1/8	2.01	0	Asp
Nicotinate and nicotinamide metabolism	1/12	2.01	0	Asp
Pyrimidine metabolism	1/38	1.86	0	βAla
Porphyrin and chlorophyll metabolism	1/54	0.96	0	Glu
Nitrogen metabolism	1/11	0.96	0	Glu

3.6 Phytoplankton Proteins

BF week 10 (October 05, 2020) had the greatest number of proteins discovered after proteomic analysis, with a total of 93. Most of the proteins were related to general cellular functions and maintaining physiology. Table 3.9 shows two proteins of the 93 total that were not related to either of the aforementioned cellular maintenance categories. Both are related to tRNAs, specifically for the amino acid, arginine.

Table 3.10. Interesting proteins discovered from BF on week 10 related to phytoplankton blooms. Two proteins, both related to arginine, are shown with a brief description.

Protein	Description
Arg_tRNA_synt_N	Arginyl tRNA synthetase N terminal domain
tRNA-synt_1d	tRNA synthetases class I (R)

CHAPTER 4: DISCUSSION

4.1 Water Quality

The watershed land-cover profiles of HH and BQ (Figure 3.1 and Table 3.1) demonstrate that the two embayments are surrounded by two very different landscapes influencing their WQ as HH has a much higher developed land-cover (29%) than BQ (1.5%). The difference in land-cover is reflected in the WQ profiles of each location (Table 3.2). The greater the developed land-cover, the more runoff carrying contaminants like sulfate and sodium from road deicers (Tong & Chen, 2002). This is reflected in Table 3.2 where levels of COND are almost twice as high in HH than in BQ. The higher levels of COND in HH also contribute to the distinction in the PCA between the three sites in HH and the five sites in BQ in Figure 3.2. HH likely has higher levels of COND due to the ability of road salts and road deicers to conduct electricity, so higher inputs of those compounds in a more urban environment would increase levels of COND in the HH embayment. Additionally, HH has a much more significant and ongoing history of heavy metal deposition than BQ. Metals such as copper, lead, and zinc have remained at relatively constant levels in HH since 1990 (Milani *et al.*, 2017). It is highly likely that the presence of these metals is contributing to the greater ability of the water in HH to conduct electricity, thus the higher levels of COND.

Additionally, a greater proportion of the watershed draining into HH is agricultural land-use (39.1%) compared to BQ (29.7%) (Figure 3.1 and Table 3.1). Tong and Chen (2002) also note that a greater agricultural land-cover, the more nutrients and sediments will be present in runoff acting as contaminants. The WQ in HH corroborates this information as all three sites in HH had higher levels of TP and NO_3 than the five sites in BQ. Interestingly,

despite the watershed draining into BQ having a considerable proportion of its area being agricultural land-use, the levels of TP were still lower in BQ than in HH, and the levels of NO_3 in HH were greatly lower than the levels of NO_3 in HH (Table 3.2 and Table 3.3). This could be due to the fact that a smaller proportion of the watershed draining into BQ has proportionally lower inputs of nutrient runoff. Another explanation may be that nutrients making their way into the Bay are diluted because the embayment in BQ is much larger and deeper overall than the embayment in HH.

Average TP levels in all eight sites in both HH and BQ exceed the Ontario provincial WQ objective guideline of 20 µg/L during ice-free periods to avoid nuisance concentrations of algae in lakes (Government of Ontario, 2021). In three sites in particular, PH, NP, and BF, the average TP concentrations exceeded 100 µg/L, with the average TP concentration in BF exceeding 300 µg/L (Table 3.2 and Table 3.3). These levels of TP are similar to the findings of Watson *et al.* (2009), however average levels of TP in NP and PH exceeded 80 µg/L. The findings from Watson *et al.* (2009) are also ideal for comparison because the samples they observed to be greater than 80 µg/L were sampled between 0.1 to 1 m in depth, similar to the methods of our study. These levels of TP are all far beyond the recommended levels in lakes to avoid excess algae. A TP model by Carle *et al.* (2005) shows a significant positive correlation between TP and household density, which is an indication of urbanization. Moreover, urban lawn care fertilizer applications could also contribute to the higher levels of TP seen in sites near towns. Bierman *et al.* (2010) demonstrated that phosphorus runoff from lawn care fertilizer applications in suburban communities greatly contributes to the degradation of WQ.

The PCA (Figure 3.2) indicates a positive correlation between DO and pH, with particularly low levels of both WQ parameters seen in BF, NP, and PH during some of the weeks earlier in the summer. Wang *et al.* (2016) determined that pH and DO are strongly positively correlated together and that lower levels of these two WQ parameters, in addition to increased levels of TP, are indicative of an algal bloom. Considering the phytoplankton information in Figure 3.3, it would appear that the increased abundance of cyanobacteria and the lower levels of DO and pH would suggest that a cyanobacteria algal bloom may have been present during those weeks. However, better identification into the specific species of cyanobacteria present during those weeks would be necessary to confirm whether or not a HAB could have been present.

WQ parameters in Munawar *et al.* (2017) focused more on nutrients compared to our study, including parameters such as: TP, Nitrate + Nitrite, and silica. No data are available for pH, DO, Secchi depth, or TURB, so comparisons on these parameters from previously documented data cannot be made. TP levels are slightly higher in YC and CB, and much higher in BF compared to the historical data from Munawar *et al.* (2017). More recent data from Munawar and Fitzpatrick (2018) indicates that the mean levels of TP in Table 3.2 are comparable to the levels of TP from 2016, with the exception of BF where TP was far greater. Additionally, warmer average temperatures were measured in this study compared to the average temperatures presented by Munawar *et al.* (2017) and Munawar and Fitzpatrick (2018). However, these discrepancies could be due to BF, CB, and YC being closer to shore with lower depths compared to the sampling site from Munawar *et al.* (2017) and Munawar and Fitzpatrick (2018), which was in open, deeper water in the center of HH.

WQ parameters are different in nearshore locations when compared to offshore locations of the same body of water. This is demonstrated in the higher variability of parameters, and higher levels of mean values of WQ in nearshore locations, with the exception of nitrates and nitrites (Yurista *et al.*, 2016). Therefore, making the comparisons between the sampling locations is difficult to interpret since the nearshore sites of this study are inherently different from the monitoring site presented by Munawar *et al.* (2017) and Munawar and Fitzpatrick (2018).

Figure 3.2, Table 3.2, and Table 3.3 also indicate greater readings of Secchi depth in the five sites of BQ and lower readings in HH. Higher Secchi depth readings indicate less turbid water; more photosynthetically active radiation can penetrate deeper into the water column. Things like inorganic suspended solids (ISS) and other dissolved organic compounds can impact water clarity, and thus the ability of light to penetrate deeper into a water column. Secchi depth is often utilized as a proxy for TURB, as both metrics measure the clarity of water. TURB and Secchi depth can be used to indicate eutrophication (Carlson, 1977) as another factor that would reduce water clarity are algal cells that can obstruct light penetration. The presence of more algae would indicate that things like excess nutrients are present, and thus the water could be eutrophic. However, it is recommended by Zou *et al.* (2020) to base eutrophication estimations on chlorophyll-a metrics as Secchi depth is strongly influenced by inorganic suspended solids that may not be indicative of algae levels. Therefore, due to the higher Secchi depth readings overall in BQ, that location could be considered less eutrophic than HH. However, as per the recommendation by Zou *et al.* (2020), Secchi depth should not be a metric for indicating eutrophication and future

studies in both BQ and HH should include metrics such as chlorophyll-a to account for ISS and to make the distinction with more confidence.

4.2 Phytoplankton

Dinoflagellates had the highest amount of biomass overall in HH (Table 3.3). Munawar *et al.* (2017) identified the dinoflagellate *Ceratium furcoides* to be the dominating species of phytoplankton during two of the seven algal bloom events documented in 2002, 2004, and 2006. These dinoflagellates made up a significant proportion of phytoplankton biomass in numerous weeks during our study, and is similar to the findings of Munawar *et al.* (2017). More recent data from 2016 shows a similar composition of phytoplankton, with high relative composition of dinoflagellates in August and September and cyanobacteria dominating in October (Munawar & Fitzpatrick, 2018). Specifically, several species of dinoflagellates from the genera *Ceratium sp*. and *Gymnodinium sp*. were found to make up greater than 5% of the total phytoplankton biomass in that study. These observations are somewhat corroborated in the data collected in our study, as dinoflagellates dominated a number of weeks in HH, with *Ceratium sp*. being the dominant genus during a number of those weeks (Figure 3.3). However, the data collected in Munawar *et al.* (2017) and Munawar and Fitzpatrick (2018) were sampled in a site centrally located in the deep, open water of HH, which is not necessarily comparable to our sites nearer to the shallower shore. Different phytoplankton communities can be found in shallower, nearshore waters due to difference in water depth and increased nutrient availability compared to offshore sites (Kovalenko *et al.*, 2019). Therefore, I must take caution when comparing between the aforementioned studies that sampled from different locations. However, the data collected

by Munawar and Fitzpatrick (2018) were also sampled from the middle of HH and are not necessarily comparable to the data from BF, CB, and YC.

On average, diatoms dominated all sites in BQ (Table 3.4). The domination of this group of phytoplankton was also observed by Munawar and Fitzpatrick (2018) and Munawar *et al.* (2018). However, Munawar and Fitzpatrick (2018) presented a much greater proportion of cyanobacteria in August and September of 2015 in BQ, which was not observed in our study. This discrepancy may be attributed to the differences in sampling locations, as Munawar and Fitzpatrick (2018) only used data sampled near the Moira River by Belleville, ON, located in the upper bay in BQ. Improved spatial resolution is presented by Munawar *et al.* (2018) where phytoplankton ecology is examined in seventeen sites across the Bay in 2010, including all five sample locations used in our study. In September 2010, diatoms dominated the HB, PH, GL, and NP, and were the second most abundant after cyanobacteria in WC (Munawar *et al.* 2018). This is substantiated for HB and GL, but only somewhat for PH and NP (Figure 3.4). Additionally, there is not a great abundance of cyanobacteria in WC during the month of September in our study, which slightly contrasts Munawar *et al.* (2018). However, the differences in phytoplankton biomass abundances appear to be small, and could be due to a number of factors, including: lower water levels, changes in nutrient input, and frequency of heavy rainfall events. Diatoms, greens, and cyanobacteria were positively related to Secchi depth (Figure 3.11), which would explain why these three types of phytoplankton were observed in greater proportions in BQ. However, the Kendall correlation matrix (Figure 3.9) did not indicate any significant relationships between Secchi depth and both cyanobacteria and diatoms. Moreover, Table

3.4 did not indicate any significant relationships between phytoplankton biomass and Secchi depth for all groups of algae. There may be relationships present, but none are statistically significant and may be due to random chance.

The Simpson's Diversity Index values for the phytoplankton genera identified in this study are much higher in HH than they are in BQ (Table 3.5). This may be due to a number of things, particularly the allochthonous inputs of carbon and other nutrients from wastewater, or urban runoff that further contribute to nutrient levels and inputs of metals influencing COND (Tong & Chen, 2002; Bowen & Currie, 2017). Since HH is rather eutrophic, it has a highly productive microbial and phytoplankton community (Bowen & Currie, 2017; Munawar *et al.*, 2017). Algal blooms can support a highly diverse composition of phytoplankton to maintain high levels of primary productivity, which may explain why higher Simpson's Diversity Index values are seen in HH than in BQ, which has fewer inputs from human activity relative to its catchment area. Additionally, freshwater algal diversity globally may be determined by nutrient limitations (Interlandi & Kilham, 2001). As lower levels of TP and NO_3 are characteristic of the WQ in BQ (Table 3.3), this may be why a lower diversity of phytoplankton is reflected in Table 3.5. The multiple linear regression (Table 3.6) assumes relationships when indicating which environmental parameters are significant. Given NO_3 was shown to be significant for five of the seven phytoplankton groups, it may be an indication that NO_3—and COND as it was shown to be collinear with NO_3—are potentially drivers of phytoplankton abundance. Positive relationship between dinoflagellates and COND and NO_3 (Figure 3.9) was also observed by da Silva *et al.* (2012), where the highest densities of the dinoflagellate, *C. furcatus*, were positively

related to conductivity and NO_3 concentrations. Furthermore, the presence of metals in HH, specifically metals that serve as micronutrients, may also be contributing to an increase in biodiversity. Micronutrients required by phytoplankton include: cobalt, cadmium, zinc, iron, manganese, selenium, molybdenum, and nickel (Baeyens *et al.*, 2018). Additionally, vanadium is a micronutrient that has also been found to be essential in green algae (Arnon & Wessel, 1953; Nalewajko *et al.*, 1995). Figure D14 in Appendix D details six of the aforementioned micronutrients analyzed by a collaborator at the University of Toronto that were detected in HH and BQ from the water samples I collected. On average, HH has higher concentrations of cadmium, copper, iron, and nickel, whereas BQ has higher concentrations of zinc and vanadium (Figure D14). Despite the concentrations of these metals being below Ontario WQ safety guidelines, it could be possible that the higher concentrations of metals in HH are enough to be more bioavailable to phytoplankton and thus more conducive to their growth and proliferation.

4.3 Metabolites

Compared to the concentrations in BQ, the relatively high amounts of glycine in HH could be indicative of the response of phytoplankton to environmental stressors. For example, glycine could have been actively used to biosynthesize glutathione. Glutathione is an important antioxidant in phytoplankton that helps prevent damage caused by reactive oxygen radicals, peroxides, and heavy metals (Pompella *et al.*, 1995). Glutathione is also a key precursor in the formation of phytochelatins, important oligomers that are responsible for binding to metals to detoxify them in plant and algal cells (Ahner *et al.*, 2002). Because much greater levels of glycine can be found in HH, an area with a history of heavy metal

contamination (HHRAP, 1992; Sofowote *et al.*, 2008), it is possible that the phytoplankton are actively using to make glutathione in response to oxidative stress caused by heavy metals in the waters of HH. One other explanation for the greater amounts of glycine in HH could be that glycine is an important precursor for glycine betaine (GBT). GBT is an osmolyte that has an important role in acclimating phytoplankton to saline environments (Reed *et al.*, 1984; Dickson & Kirst, 1986). There has been extensive research on the role of GBT in marine and estuarine phytoplankton, but its role in freshwater phytoplankton is not as well researched. Much greater levels of glycine in HH may be present because there are much greater levels of COND compared to BQ (Figure 3.2 & Table 3.2). COND is related to the ability of water to conduct electricity, which is directly related to the salinity of the water. Since the levels of COND are much greater in HH, phytoplankton may be producing glycine to synthesize GBT in order to osmotically acclimate to their environment. One other explanation for the greater amounts of glycine in HH could be the amino acid's relationship with photosynthesis. Intracellular glycine has been shown to increase during photorespiration in the absence of light (Leboulanger *et al.*, 1998). This would mean that glycine levels could be higher in the dark, such as in water with lower light penetration or higher turbidity. This may help to explain the higher presence of glycine during the day observed in our study, since HH had lower Secchi depth readings and higher levels of TURB compared to BQ. There are also instances of higher levels of glycine in BQ, such as NP on Sept 23, WC on Aug 06, and GL on Aug 26 (Figure 3.10). However, these may be one-off instances of higher levels of glycine produced in response to temporary low water clarity or an unknown temporary stressor since these levels are not maintained temporally and are inconsistent with other weekly metabolite profiles.

α-Aminobutyric acid (AABA) is an analogue of the neurotransmitter γ-Aminobutyric acid (GABA), and this metabolite was found to be in higher levels in BQ than in HH. AABA made up a larger proportion of all metabolites in HB and GL, with proportions increasing over the ten weeks in HB. There does not appear to be any literature available for the role of AABA in phytoplankton, especially in freshwater environments. Based on our results, I believe that AABA is positively related to the biomass of cyanobacteria and diatoms, but negatively related to dinoflagellates (Figure 3.11). The environments in which phytoplankton are found may also affect their ability or need to produce AABA. For example, AABA is negatively related to both COND and NO_3, yet positively related to Secchi depth (Figure 3.12). AABA may be a metabolite that is produced by phytoplankton in response to greater photosynthetically active radiation penetration in waters that have a high clarity, noted by a greater Secchi depth reading, opposite of what was observed in our study by glycine. Moreover, waters that are high in COND and NO_3 may not be conducive to the phytoplankton that produce AABA. Dinoflagellates that were found to have greater biomass in HH, which subsequently has higher levels of COND and NO_3, and this group of algae may not need to produce AABA for their cellular processes. In contrast, diatoms were found in greater biomass in the less turbid waters of BQ, so maybe diatoms need to produce AABA in response to the greater light availability. This hypothesis is backed up by larger amounts of serine found in the phytoplankton in BQ. Serine is involved in the production of mycosporine methylamine-serine, a photoprotective compound in phytoplankton (Pallela *et al.*, 2010). Mycosporine amino acids are secondary metabolites often described as microbial sunscreens that are produced by aquatic organisms found in waters with high amounts of sunlight. Another metabolite that was found in larger amounts

in the phytoplankton of BQ is taurine, which is also a component in another photo-protective metabolite, mycosporine-taurine. Mycosporine-taurine is capable of dissipating harmful ultraviolet radiation as heat energy, and this compound—in addition to mycosporine-serine and mycosporine-glycine—are found in microalgae (Sen and Mallick, 2021). As mycosporines are fairly novel molecules, more research is needed to investigate their roles and behaviours in phytoplankton.

Anserine was one of the metabolites that had a strong, positive relationship with dinoflagellates (Figure 3.6). Anserine has been shown to have antioxidant properties in a red alga (Tamura *et al.*, 1998), and although dinoflagellates are not red alga, perhaps anserine is functioning in a similar manner to combat oxidative stress. Anserine is a di-peptide that is derived from carnosine, which was not detected in high concentrations in algae from either HH or BQ (Figure 3.9 and Figure 3.10). The absence of carnosine may be indicative of its consumption in the biosynthesis of anserine. A negative relationship can be seen in Figure 3.11, apparent by the arrows of anserine and carnosine in opposite directions, contributing to this hypothesis.

The biomass of cyanobacteria was shown to be positively related to the majority of the metabolites incorporated in the RDA (Figure 3.11). Among the metabolites that were positively associated with the biomass of cyanobacteria are three amino acids: leucine, arginine, and tyrosine (tyrosine was removed due to being collinear with arginine and leucine). These three amino acids are imperative in the three most toxic forms of microcystins, MC-LR, MC-RR, and MC-YR. Thus, seeing this positive relationship

between the aforementioned amino acids and cyanobacteria could be an indication that cyanobacteria are producing, or about to produce MCs. Hydroxyproline was significantly negatively related to TP (Figure 3.12), as the metabolite was only found in HH (Figure D15), which overall had much higher levels of TP compared to BQ. There appears to be no literature available to explain the negative relationship between cyanobacteria, TP, and hydroxy-proline. However, hydroxyproline has been demonstrated to be a significant component in the glycoproteins that comprise the cell walls of green algae (Pivokonsky *et al.*, 2014). Therefore, it is possible that a negative relationship would exist between cyanobacteria and hydroxyproline because cyanobacteria do not have cell walls with glycoproteins and therefore would have no need hydroxyproline.

The top four pathways identified by the 33 metabolites of this study are indicative of a few key processes in phytoplankton. Although there were not a large number of hits matching to the metabolites making up the entire pathway, the four pathways identified in Table 3.9 and detailed in Figures 3.14 through Figure 3.17 had the greatest significance and impact. The impact is a relation between the number of metabolites in a given pathway versus the metabolites that match within. There are inherent limitations when providing a subset of metabolites to a pathway analysis program, as many of the precursors, products, and in-between compounds are not necessarily a part of the limited input list. Arginine biosynthesis (Figure 3.14) is an interesting pathway to be both impactful and significant among all other pathways. Given arginine is positively related to cyanobacteria (Figure 3.11), it can be hypothesized that this group of algae are making arginine for the production of MCs. However, since there is no significant relationship between cyanobacteria biomass

and arginine is not apparent in Figure 3.13, more research needs to be conducted with a greater number of input metabolites to determine whether or not the arginine biosynthesis pathway is in fact significant. There are several other reasons beyond the scope of this study that could explain why the arginine synthesis pathway was found to be significant but a relationship between cyanobacteria and arginine was not. For example, arginine is a very important precursor to the nitrogen reservoir, cyanophycin, in cyanobacteria, so the arginine being biosynthesized could be going immediately to the production of that compound and not MCs (Zhang & Yang, 2019). The second pathway found to be significant, Arginine and Proline Metabolism, is related to Arginine Biosynthesis (Figure 3.15). The Arginine and Proline Metabolism pathway indicates that arginine is not only being produced but consumed. Arginine and proline metabolism pathways have been shown to be involved in low temperature stress responses in algae (He *et al.*, 2018). Low water temperatures were observed later on in October during the sampling season, so arginine consumption may relate to the change in season in addition to MC and cyanophycin production. The other two pathways related to alanine, aspartate, glutamate, and β-alanine metabolism (Figure 3.16 and Figure 3.17), signify that many amino acids are being used up by the phytoplankton observed in this study. All of the metabolites from these pathways are common in a number of cellular maintenance and physiological behaviours in algae.

4.4 Proteins

The proteomics conducted in this research project were not very successful. No proteins were detected in the majority of sites during most weeks of sampling, several sites had

anywhere from 1-50 proteins for a few weeks, but not enough to gain any meaning from them or perform statistics on the results. That is with the exception of BF week 10, where 96 proteins were identified after proteomic analysis. There were many proteins related to generic cellular processes and signalling, however two proteins that stood out among the rest were tRNA for arginine (Table 3.6). This is fairly substantial, as this indicates that arginine was being used in the cellular processes of the phytoplankton during that week. This is corroborated by the metabolic pathways related to arginine biosynthesis and metabolism being significant (Figure 3.14 and Figure 3.15).

There are several factors that could have contributed to the proteomics of this research project being unsuccessful. First, the goal of performing proteomics on wild algae collected from real environments is a goal that is novel to our knowledge. There are a variety of studies that perform proteomics on algae of a known strain and origin, grown under controlled conditions in a laboratory. Some studies conduct proteomics on phytoplankton communities grown in mesocosms, such as Russo *et al.* (2016), but these studies still have control over the samples of algae they collect. Secondly, it is likely that there are better ways of collecting wild phytoplankton samples that are both effective, consistent, and do not require a lot of time to collect. McCain *et al.* (2021) were able to collect samples for 'omics analyses from the arctic, however they used three separate filters and it took 3 hours to acquire these samples. In addition to the previously mentioned study not being published during the planning of our study, and the fact that sampling each site could take up to 3 hours, this method of sampling would not have been suitable for sampling at the eight sites I selected. I considered and attempted filtering for this study, albeit with different filters

and with lower volumes of water, but it was immediately apparent that it would take a very long time. Moreover, the logistics of acquiring filter papers during the COVID-19 pandemic were difficult and early attempts caused algae to stick to the filter that were unsuitable for proteomic analyses. Furthermore, there is room to improve the laboratory sample preparation of phytoplankton sampled in the field. For example, it is possible that cell lysis methods were inadequate to release the proteins from the algal cells. It is also possible that the Bligh Dyer liquid-liquid phase extraction method could have been optimized to separate the proteins more effectively for phytoplankton cells, but I was sample- and time-limited. During the initial liquid-liquid phase extraction, the protein disc was rather large, at least when compared to a disc from plasma, so it is possible that other contaminants like sediment particulate matter, cellular debris, and other soluble organic matter could have contaminated the protein disc and rendered it unsuitable for the digestion protocol, or even that there was much less protein present than estimated by mass and gross spectrophotometric measurement. Other sample preparation methods could be used in future studies with algae. For example, ultrapore filters could be used to concentrate the proteins, or trypsin could be used to digest the proteins instead of formic acid. Further method development is required to be able to consistently and effectively perform proteomics on wild phytoplankton samples acquired in the field.

CHAPTER 5: CONCLUSION

The WQ conditions in two Lake Ontario AOCs were successfully characterized. HH is an embayment heavily influenced by human activity and urban land-use. There are far greater levels of COND, NO_3, and TP compared to BQ. These WQ parameters also drive the phytoplankton communities to thrive in the waters of HH and supports a greater phytoplankton biodiversity than BQ. Although phytoplankton community composition differed from previous studies conducted over the last 18 years, this is most likely attributed to the differences in sample locations and sampling methods. However, the data I collected in this study provide a different perspective into the phytoplankton communities in such a diverse and complex ecosystem. The metabolite profiles of HH are novel for this area and reflect a community of phytoplankton constantly responding to oxidative stresses and other environmental stressors. Given the high COND and presence of heavy metals, among other historical contaminant inputs in HH, it makes sense to see glycine present in high levels as it is a precursor to several oxidative stress compounds.

BQ has a much larger watershed and catchment area compared to HH, with less influence from urbanization. This is reflected in the WQ of BQ, which is more representative of a landscape not as impacted by road salt runoff, deicers, metals, lawn care fertilizers, and other human influences. In particular, BQ has much greater depths to which a Secchi disc can be read, indicating light can penetrate deeper into the water there than in HH. However, there are some exceptions, such as in NP and PH, which are two sample locations close to towns with greater populations. The phytoplankton community composition profiles are similar to previously published papers, likely due to the similarities in sample locations.

However, TP levels do not appear to have improved greatly from recently published data in BQ. The metabolite profile in BQ is a bit more complex to interpret, but the metabolites are indicative of algal responses to light stress including serine and taurine, two metabolic precursors to ultraviolet radiation absorbing compounds. However, much more research needs to be conducted to identify the relationship between the metabolites of this study and the physiology and stress response of phytoplankton, particularly in diatoms since this group of phytoplankton was seen in the greatest amount in BQ.

Metabolite pathway analyses identified several interesting pathways related to some key amino acids in toxin production, but also general cell function. More method development is necessary in order to fine-tune the sample acquisition and processing to have better success with proteomics conducted on phytoplankton sampled in the field. Doing so could provide great insight into the functionality of phytoplankton and how they behave during a bloom.

Overall, phytoplankton and the metabolites in which they produce were successfully identified. The metabolite profiles, WQ, and phytoplankton community composition were compared and all are indicative of two very different AOCs in Lake Ontario. More research needs to be done to provide more detailed information about a number of metabolites and what they mean for phytoplankton. For example, a non-targeted metabolomic search of the phytoplankton metabolite data might be able to recognize secondary metabolites and precursor metabolites that would be better connected to arginine biosynthesis and metabolism. Also, this study could be repeated on a finer temporal scale to omit effects

from time as a variable. The results of this study represent the preliminary information regarding a number of relationships between WQ and phytoplankton metabolites within the context of two freshwater locations in Lake Ontario, such as with glycine, AABA, and photo-protective mycosporine amino acids. These findings will contribute greatly to the future actions of conservation authorities and municipalities to better manage and deal with nuisance phytoplankton and other algal blooms in HH and BQ in the future.

REFERENCES

Abraham-Peskir, J. A. (1998). Structural changes in fully hydrated *Chilomonas paramecium* exposed to copper. *European Journal of Protistology*, 34(1), 51-57. doi:10.1016/S0932-4739(98)80039-4

Ahner, B. A., Wei, L. P., Oleson, J. R., & Ogura, N. (2002). Glutathione and other low molecular weight thiols in marine phytoplankton under metal stress. *Marine Ecology Progress Series*, 232, 93-103. doi:10.3354/meps232093

Amengual-Morro, C., Niell, G. M., & Martínez-Taberner, A. (2011). Phytoplankton as bioindicator for waste stabilization ponds. *Journal of Environmental Management*, 95, 1-6. doi:10.1016/j.jenvman.2011.07.008

Arnon, D. I., & Wessel, G. (1953). Vanadium as an Essential Element for Green Plants. *Nature*, 172(4388), 1039-1040. doi:10.1038/1721039a0

Baeyens, W., Gao, Y., Davison, W., Galceran, J., Leermakers, M., Puy, J., . . . Beguery, L. (2018). *In situ* measurements of micronutrient dynamics in open seawater show that complex dissociation rates may limit diatom growth. *Scientific Reports*, 8, 16125. doi:10.1038/s41598-018-34465-w

Baker, A. L., & al., e. (2012). Pycokey -- an image based key to Algae (PS Protists), Cyanobacteria, and other aquatic objects. *University of New Hampshire Center for Freshwater Biology*. Retrieved from http://cfb.unh.edu/phycokey/phycokey.htm

Barker, H. A. (1935). The metabolism of the colorless alga, *Prototheca zopfii krüger*. *Journal of Cellular and Comparative Physiology*, 7(1), 73-93. doi:10.1002/jcp.1030070105

Bierman, P. M., Horgan, B. P., Rosen, C. J., & Hollman, A. B. (2010). Phosphorus runoff from turfgrass as affected by phosphorus fertilization and clipping management. *Journal of Environmental Quality,* 39(1), 282-292. doi:10.2134/jeq2008.0505

Bowen, K. L., & Currie, W. J. (2017). Elevated zooplankton production in a eutrophic Lake Ontario embayment: Hamilton Harbour 2002-2014. *Aquatic Ecosystem Health & Management,* 20(3), 230-241. doi:10.1080/14634988.2017.1294425

Bowlby, J. N., & Hoyle, J. A. (2017). Developing restoration targets for nearshore fish populations in two Areas of Concern in Lake Ontario. *Aquatic Ecosystem Health and Management,* 20(2), 242-251. doi:10.1080/14634988.2017.1295760

BQRAP (Bay of Quinte Remedial Action Plan). (2018). *Bay of Quinte Remedial Action Plan (BQRAP).* 1-2.

Buskey, E. J., Strom, S., & Coulter, C. (1992). Bioluminescence of heterotrophic dinoflagellates from Texas coastal waters. *Journal of Experimental Marine Biology and Ecology,* 159(1), 37-49. doi:10.1016/0022-0981(92)90256-A

Carle, M. V., Halpin, P. N., & Stow, C. A. (2005). Patterns of watershed urbanization and impacts on water quality. *Journal of the American Water Resources Association,* 41(3), 693-708. doi:10.1111/j.1752-1688.2005.tb03764.x

Carlson, R. E. (1977). A trophic state index for lakes. *Limnology and Oceanography,* 22, 361-369. doi:10.1016/j.scitotenv.2018.07.208

Carty, S., & Parrow, M. W. (2015). Dinoflagellates. In J. D. Wehr, R. G. Sheath, & J. P. Kociolek (Eds.), *Freshwater Algae of North America* (Second ed., pp. 773-807). Academic Press.

Codd, G. A., Testai, R., Funari, R., & Svirčev, Z. (2020). Cyanobacteria, Cyanotoxins, and Human Health. In A. E. Hiskia, T. M. Triantis, M. G. Antoniou, T. Kaloudis, & D. D. Dionysiou, *Water Treatment for Purification from Cyanobacteria and Cyanotoxins* (pp. 37-68). doi:10.1002/9781118928677.ch2

Cristescu, M. E. (2019). The concept of genome after one century of usage. *Genome*, 62(10), iii-v. doi:10.1139/gen-2019-0129

da Silva, L. C., Leone, I. C., dos Santos-Wisniewski, M. J., Peret, A. C., & Rocha, O. (2012). Invasion of the dinoflagellate *Ceratium furcoides (Levander) Langhans* 1925 at tropical reservoir and its relation to environmental variables. *Biota Neotropica*, 12(2), 93-100. doi:10.1590/S1676-06032012000200010

Dai, R., Liu, H., Qu, J., Zhao, X., & Hou, Y. (2009). Effects of amino acids on microcystin production of the *Microcystis aeruginosa*. *Journal of Hazardous Materials*, 161(2-3), 730-736. doi:10.1016/j.jhazmat.2008.04.015

Diaz, J. M., Holland, A., Sanders, J. G., Bulski, K., Mollett, D., Chou, C., . . . Duahmel, S. (2018). Dissolved Organic Phosphorus Utilization by Phytoplankton Reveals Preferential Degradation of Polyphosphates Over Phosphomonoesters. *Frontiers in Marine Science*, 5, 380. doi:10.3389/fmars.2018.00380

Dickson, D. M., & Kirst, G. O. (1986). The role of β-dimethylsulphoniopropionate, glycine betaine and homarine in the osmoacclimation of *Platymonas subcordiformis*. *Planta*, 167(4), 536-543. doi:10.1007/BF00391230

Dore, J. E., Letelier, R. M., Church, M. J., Lukas, R., & Karl, D. M. (2008). Summer phytoplankton blooms in the oligotrophic North Pacific Subtropical Gyre: Historical perspective and recent observations. *Progress in Oceanography,* 76, 2-38. doi:10.1016/j.pocean.2007.10.002

Dreher, T. W., Colalrt, L. P., Mueller, R. S., Halsey, K. H., Bildfell, R. J., Schreder, P., ... Ferry, R. (2019). *Anabaena/Dolichospermum* as the source of lethal microcystin levels responsible for a large cattle toxicosis event. *Toxicon X*, 1, 100003. doi:10.1016/j.toxcx.2018.100003

Dyble, J., Tester, P. A., & Litaker, R. W. (2006). Effects of light intensity on cylindrospermopsin production in the cyanobacterial HAB species *Cylindrospermopsis raciborskii*. *African Journal of Marine Science*, 28(2), 309-312. doi:10.2989/18142320609504168

Flores, E., & Herrero, A. (2005). Nitrogen assimilation and nitrogen control in cyanobacteria. *Biochemical Society Transactions*, 33(1), 164-167. doi:10.1042/BST0330164

Friedrichs, L., Hörnig, M., Schulze, L., Bertram, A., Jansen, S., & Hamm, C. (2013). Size and biomechanic properties of diatom frustules influence food uptake by copepods. *Marine Ecology Process Series*, 481, 41-51. doi:10.3354/meps10227

Georges, A. A., El-Swiad, H., Craig, S. E., Li, W. K., & Walsh, D. A. (2014). Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *The ISME Journal*, 8, 1301-1313. doi:10.1038/ismej.2013.234

Gosset, A., Oestreicher, V., Perullini, M., Bilmes, S. A., Jobbágy, M., Dulhoste, S., Bayard, R., & Durrieu, C. (2019). Optimization of sensors based on encapsulated algae for pesticide detection in water. *Analytical Methods*. 48, 6193-6203. doi:10.1039/c9ay02145k

Government of Canada. (2017, February 20). Bay of Quinte: Area of Concern. Retrieved from Government of Canada: https://www.canada.ca/en/environment-climate-change/services/great-lakes-protection/areas-concern/bay-of-quinte.html

Government of Canada. (2021, January 04). Hamilton Harbour: Area of Concern. Retrieved from Government of Canada: https://www.canada.ca/en/environment-climate-change/services/great-lakes-protection/areas-concern/hamilton-harbour.html

Government of Ontario. (2016, January 21). Ontario Land Cover Compilation v.2.0. Retrieved from Ontario GeoHub: https://geohub.lio.gov.on.ca/documents/7aa998fdf100434da27a41f1c637382c/about

Government of Ontario. (2021). Lakeshore Capacity Assessment Handbook: Protecting Water Quality in Inland Lakes on Ontario's Precambrian Shield Appendix A: Rationale for a Revised Phosphorus Criterion for Precambian Shield Lakes in Ontario. Retrieved from https://www.ontario.ca/page/lakeshore-capacity-assessment-handbook-protecting-water-quality-inland-lakes-ontarios-precambrian

Grantz, D. A., Vaughn, D. L., Farber, R. J., Kim, B., Ashbaugh, L., VanCuren, T., & Campbell, R. (1998). Wind barriers suppress fugitive dust and soil-derived airborne particles in arid regions. *Journal of Environmental Quality*, 27(4), 946-952. doi:10.2134/jeq1998.00472425002700040031x

Guillard, R. L., & Sieracki, M. S. (2005). Counting Cells in Cultures with the Light Microscope. In R. A. Anderson, *Algal Culturing Techniques* (pp. 239-252). New York: Elseviewr Inc.

Guo, M., Harrison, P. J., & Taylor, F. J. (1996). Fish kills related to *Prymnesium parvum N. Carter* (Haptophyta) in the People's Republic of China. *Journal of Applied Phycology*, 8(2), 111-117. doi:10.1007/BF02186313

Hall, J. D., O'Connor, K., & Ranieri, J. (2006). Progress Toward Delisting a Great Lakes Area of Concern: The Role of Integrated Research and Monitoring in the Hamilton Harbour Remedial Action Plan. *Environmental Monitoring and Assessment*, 113, 227-243. doi: 10.1007/s10661-005-9082-8

Hansen, P. J., & Nielsen, T. G. (1997). Mixotrophic feeding of *Fragilidium subglobosum* (Dinophyceae) on three species of *Ceratium*: effects of prey concentration, prey species and light intensity. *Marine Ecology Progress Series*, 147, 187-196. doi:10.3354/meps147187

Hashimoto, Y., Okaichi, T., Fang, L. D., & Noguchi, T. (1968). Glenodinine, an ixhthyotoxic substance produces by a dinoflagellate, *Peridinium polonicum*. *Bulletin of the Japanese Society of Scientific Fisheries*, 34(6), 528.

He, Y., Hu, C., Wang, Y., Cui, D., Sun, X., Li, Y., & Xu, N. (2018). The metabolic survival strategy of marine macroalga *Ulva prolifera* under temperature stress. *Journal of Applied Phycology*, 30, 3611-3621. doi:10.1007/s10811-018-1493-3

HHRAP (Hamilton Harbour Remedial Action Plan). (1992). Remedial Action Plan for Hamilton Harbour Stage I Report, second ed.

HHRAP (Hamilton Harbour Remedial Action Plan). (2018). Contaminant Loadings and Concentrations to Hamilton Harbour: 2008-2016 Update.

Ho, J. C., & Michalchuk, M. A. (2015). Challenges in tracking harmful algal blooms: A synthesis of evidence from Lake Erie. *Journal of Great Lakes Research*, 41(2), 317-325. doi:10.1016/j.jglr.2015.01.001

Holland, A., & Kinnear, S. (2013). Interpreting the Possible Ecological Role(s) of Cyanotoxins: Compounds for Competitive Advantage and/or Physiological Aide? *Marine Drugs*, 11(7), 2239-2258. doi:10.3390/md11072239

Hudon, C., De Sève, M., & Cattaneo, A. (2014). Increasing occurrence of the benthic filamentous cyanobacterium *Lyngbya wollei*: a symptom of freshwater ecosystem degradation. *Freshwater Science*, 33(2), 606-618. doi:10.1086/675932

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

Interlandi, S. J., & Kilham, S. S. (2001). Limiting resources and the regulations of diversity in phytoplankton communities. *Ecology*, 82(5), 1270-1282. doi:10.1890/0012-9658(2001)082[1270:LRATRO]2.0.CO;2

International Joint Commission United States and Canada (IJC). (1987). Revised Great Lakes Water Quality Agreement of 1978: Agreement, with Annexes and Terms of Reference, Between the United States and Canada, Signed at Ottawa November 22, 1978 and Phosphorus Load Reduction Supplement signed October 16, 1983: as Amended by. Windsor, ON, Canada: Protocol Signed November 18, 1987.

Itakura, S., & Imani, I. (2014). Economic impacts of harmful algal blooms on fisheries and aquaculture in western Japan - an overview of interannual variability and interspecies comparison. *PICES Scientific Report*, 47(17), 17-26.

Kim, D., Kaluskar, S., Mugalingham, S., & Arhonditsis. (2016). Evaluating the relationships between watershed physiology, land use patterns, and phosphorus loading in the bay of Quinte basin, Ontario, Canada. *Journal of Great Lakes Research*, 42, 972-984. doi:10.1016/j.jglr.2016.07.008

Kovalenko, K. E., Reavie, E. D., Bramburger, A. J., Cotter, A., & Sierszen, M. E. (2019). Nearshore-offshore trends in Lake Superior by phytoplankton. *Journal of Great Lakes Research*, 45(6), 1197-1204. doi:10.1016/j.jglr.2019.09.016

Kuhlisch, C., Althammer, J., Sazhin, A. F., Jakobsen, H. H., Nejstgaard, J. C., & Pohnert, G. (2020). Metabolomics-derived marker metabolites to characterize *Phaeocystis pouchetii* physiology in natural plankton communities. *Scientific Reports*, 10, 20444. doi:10.1038/s41598-020-77169-w

Lau, R. K., Kwok, A. C., Chan, W. K., Zhang, T. Y., & Wong, J. T. (2007). Mechanism characterization of cellulosic thecal plates in dinoflagellates by nanoidentification. *Journal of Nanoscience and Technology*, 7, 452-457.

Leboulanger, C., Martin-Jézéquel, V., Descolas-Gros, C., & Sciandra, A. (1998). Photorespiration in continuous culture of *Dunaliella tertiolecta* (Chlorophyta): Relationships between serine, glycine, and extracellular glycolate. *Journal of Phycology*, 34(4), 651-654. doi:10.1046/j.1529-8817.1998.340651.x

Lee, R. (2018). Phycology (5th ed.). Cambridge: Cambridge University Press. doi:10.1017/9781316407219

Li, M., Gao, X., Wu, B., Qian, X., Giesy, J. P., & Cui, Y. (2014). Microalga Euglena as a bioindicator for testing genotoxic potentials of organic pollutants in Taihu Lake, China. Ecotoxicology, 23, 633-640. doi:10.1007/s10646-014-1214-x

Mann, D. G. (1999). The species concept in diatoms. *Phycologia*, 38(6), 437-495. doi:10.2216/i0031-8884-38-6-437.1

Manning, S. R., & La Calire, J. W. (2010). Prymnesins: Toxic Metabolites of the Golden Alga, *Prymnesium parvum Carter* (Haptophyta). *Marine Drugs*, 8(3), 678-704. doi:10.3390/md8030678

Milani, D., & Grapentine, L. C. (2016). Prioritization of sites for sediment remedial action at Randle Reef, Hamilton Harbour. *Aquatic Ecosystem Health & Management*, 19(2), 150-160. doi:10.1080/14634988.2015.1106224

Milani, D., Grapentine, L., Burniston, D. A., Graham, M., & Marvin, C. (2017). Trends in sediment quality in Hamilton Harbour, Lake Ontario. *Aquatic Ecosystem Health & Management*, 20(3), 295-307. doi:10.1080/14634988.2017.1302780

Millard, E. S., & Sager, P. E. (1994). Comparison of Phosphorus, Light Climate, and Photosynthesis between Two Culturally Eutrophied Bays: Green Bay, Lake

Michigan, and the Bay of Quinte, Lake Ontario. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(11), 2579-2590. doi:10.1139/f94-258

Moore, R. E., Chen, J. L., Moore, B. S., Patterson, G. M., & Carmichael, W. W. (1991). Biosynthesis of microcystin-LR. Origin of the carbons in the adda and masp units. *Journal of the American Chemical Society*, 113(13), 5083-5084. doi:10.1021/ja00013a066

Munawar, M., & Fitzpatrick, M. A. (2018). Eutrophication in three Canadian Areas of Concern: Phytoplankton and major nutrient interactions. *Aquatic Ecosystem and Management*, 21(4), 421-437. doi:10.1080/14634988.2018.1530895

Munawar, M., Fitzpatrick, M., Niblock, H., Kling, H., Lorimer, J., & Rozon, R. (2018). Phytoplankton ecology in the Bay of Quinte, Lake Ontario: Spatial distribution, dynamics and heterogeneity. *Aquatic Ecosystem Health & Management*, 21(2), 213-226. doi:10.1080/14634988.2018.1474058

Munawar, M., Fitzpatrick, M., Niblock, H., Kling, H., Rozon, R., & Lorimer, J. (2017). Phytoplankton ecology of a culturally eutrophic embayment: Hamilton Harbour, Lake Ontario. *Aquatic Ecosystem Heath and Management*, 20(3), 201-213. doi:10.1080/14634988.2017.1307678

Murphy, J., & Riley, J. P. (1962). A modified single solution method for the determination of phosphate in natural waters. *Analytica Chemica Acta*, 27, 31-36. doi:10.1016/S0003-2670(00)88444-5

Nalewajko, C., Lee, K., & Jack, T. R. (1995). Effects of vanadium on freshwater phytoplankton photosynthesis. *Water, Air, and Soil Pollution*, 81, 93-105. doi:https://doi.org/10.1007/BF00477258

Nicholls, K. H., & Hurley, D. A. (1989). Recent Changes in the Phytoplankton of the Bay of Quinte, Lake Ontario: the Relative Importance of Fish, Nutrients, and Other Factors. *Canadian Journal of Fisheries and Aquatic Science*, 46(5), 770-779. doi:10.1139/f89-095

Nishitani, G., Yamaguchi, M., Ishikawa, A., Yanagiya, S., Mitsuya, T., & Imai, I. (2005). Relationships between occurrences of toxic *Dinophysis species* (Dinophyceae) and small phytoplanktons in Japanese coastal waters. Harmful Algae, 4(4), 755-762. doi:10.1016/j.hal.2004.11.003

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. M., Szoecs, E., & Wagner, H. (2020). vegan: Community Ecology Package. R package version 2.5-7. https://CRAN.R-project.org/package=vegan

Oliver, S. G., Winson, M. K., Kell, D. B., & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, 16(9), 373-378. doi:10.1016/S0167-7799(98)01214-1

Ontario Ministry of Environment. (1983). Handbook of Analytical Methods for Environmental Samples. Toronto, Ontario, Canada: Laboratory Services and Applied Research Branch, Ontario Ministry of Environment.

Ontario Ministry of Natural Resources and Forestry. (2021, June 25). Ontario Watershed Boundaries (OWB) [Data set]. Retrieved from Ontario GeoHub: https://geohub.lio.gov.on.ca/maps/mnrf::ontario-watershed-boundaries-owb/explore?location=39.548770%2C-81.397584%2C5.00

Paerl, H. W., & Scott, J. T. (2010). Throwing Fuel on the Fire Synergistic Effects of Excessive Nitrogen Inputs and Global Warming on Harmful Algal Blooms. *Environmental Science & Technology*, 44(20), 7756-7758. doi:10.1021/es102665e

Pallela, R., Na-Young, Y., & Kim, S. (2010). Anti-photoaging and photoprotective compounds derived from marine organisms. *Marine Drugs*, 8(4), 1189-1202. doi:10.3390/md8041189

Pan, R., Stevenson, R. J., Hill, B. H., Herlihy, A. T., & Collins, G. B. (1996). Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. *Journal of the North American Benthological Society*, 15(4), 481-495. doi:10.2307/1467800

Pang, Z., Chong, J., Zhou, G., de Lima Morais, D. A., Chang, L., Barrette, M., . . . Xia, J. (2021). Metaboanalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Research*, 49(W1), W388-W396. doi:10.1093/nar/gkab382

Park, J., Jeong, H. J., Yoo, Y. D., & Yoon, E. Y. (2013). Mixotrophic dinoflagellate red tides in Korean waters: Distribution and ecophysiology. *Harmful Algae*, 30(1), S28-S40. doi:10.1016/j.hal.2013.10.004

Perri, K. A., Sullivan, J. M., & Boyer, G. L. (2015). Harmful algal blooms in Sodus Bay, Lake Ontario: A comparison of nutrients, marina presence, and cyanobacterial toxins. *Journal of Great Lakes Research*, 41(2), 326-337. doi:10.1016/j.jglr.2015.03.022

Pichardo, S., Jos, A., Zurita, J. L., Salguero, M., Cameán, A. M., & Repetto, G. (2007). Acute and subacute toxic effects produced by microcystin-YR on the fish cell lines RTG-2 and PLHC-1. *Toxicity in Vitro*, 21(8), 1460-1467. doi:10.1016/j.tiv.2007.06.012

Pillsbury, R. W., Reavie, E. D., & Estepp, L. R. (2021). Diatom and geochemical paleolimnology reveals a history of multiple stressors and recovery on Lake Ontario. *Journal of Great Lakes Research*, 47(5), 1316-1326. doi:10.1016/j.jglr.2021.07.006

Pino, L. K., Searle, B. C., Bollinger, J. G., MacLean, B., & MacCoss, M. J. (2020). The Skyline ecosystem: informatics for quantitative mass spectrometry proteomics. *Mass Spectrometry Reviews*, 39, 229-244. doi:10.1002/mas.21540

Pivokonsky, M., Safarikova, J., Baresova, M., Pivokonska, L., & Kopecka, I. (2014). A comparison of the character of algal extracellular versus cellular organic matter produced by cyanobacterium, diatom and alga. *Water Research*, 51(15), 37-46. doi:10.1016/j.watres.2013.12.022

Pompella, A., Visvikis, A., Paolicchi, A., De Tata, V., & Casini, A. F. (2003). The changing faces of glutathione, a cellular protagonist. *Biochemical Pharmacology*, 66(8), 1499-1503. doi:10.1016/S0006-2952(03)00504-5

Prescott, G. W. (1962). Algae of the western Great Lakes area. Dubuque, Iowa, W. C. Brown Co.

Puschner, B., Hoff, B., & Tor, E. R. (2008). Diagnosis of anatoxin-a poisoning in dogs from North America. *Journal of Veterinary Diagnostic Investigation*, 20(1), 89-92. doi:10.1177/104063870802000119

Quiblier, C., Wood, S., Echenique-Subiabre, I., Heath, M., Villeneuve, A., & Humbert, J. -F. (2013). A review of current knowledge on toxic benthic freshwater cyanobacteria Ecology, toxin production and risk management. *Water Research*, 47(15), 5464-5479. doi:10.1016/j.watres.2013.06.042

Reed, R. H. (1984). Transient breakdown in the selective permeability of the plasma membrane of *Chlorella emersonii* in response to hyperosmotic shock: implications for cell water relations and osmotic adjustment. *Journal of Membrane Biology*, 82(1), 83-88. doi:10.1007/BF01870734

Rodgers, J. H., Johnson, B. M., & Bishop, W. M. (2010). Comparison of Three Algaecides for Controlling the Density of *Prymnesium parvum*. *Journal of the American Water Resources Association*, 46(1), 153-160. doi:10.1111/j.1752-1688.2009.00399.x

Round, F. E. (1991). Diatoms in river water-monitoring studies. *Journal of Applied Phycology*, 3, 129-145. doi:10.1007/BF00003695

Round, F. E., Crawford, R. M., & Mann, D. G. (1990). The Diatoms. Biology and Morphology of the Genera. *Cambridge University Press*, 346(6285), 747. doi:10.1038/346619a0

Russo, D. A., Beckerman, A. P., & Pandhal, J. (2016). A metaproteomics analysis of the response of a freshwater microbial community under nutrient enrichment. *Frontiers in Microbiology*, 7, 1172. doi:10.3389/fmicb.2016.01172

Schaefer, A. M., Yrastorza, L., Stockley, N., Harvey, K., Harris, N., Grady, R., . . . Reif, J. S. (2020). Exposure to microcystin among coastal residents during a cyanobacterial bloom in Florida. *Harmful Algae*, 92, 101769. doi:10.1016/j.hal.2020.101769

Sen, B., Alp, T. A., Sonmez, F., Turan Kocer, M. A., & Canpolar, O. (2013). Relationship of Algae to Water Pollution and Waste Water Treatment. In W. Elshorbagy, Water Treatment (pp. 335-354). doi:10.5772/51927

Sen, S., & Mallick, N. (2021). Mycosporine-like amino acids: Algal metabolites shaping the safety and sustainability profiles of commercial sunscreens. *Algal Research*, 58, 102425.doi:10.1016/j.algal.2021.102425

Sheath, R. G., & Wehr, J. D. (Eds.). (2003). Freshwater algae of North America: ecology and classification. Academic Press.

Smalley, G. W., Coats, D. W., & Stoecker, D. K. (2003). Feeding in the mixotrophic dinoflagellate *Ceratium furca* is influenced by intracellular nutrient concentrations. *Marine Ecology Progress Series*, 262, 137-151. doi:10.3354/meps262137

Sofowote, U. M., McCarry, B. E., & Marvin, C. H. (2008). Source Apportionment of PAH in Hamilton Harbour Suspended Sediments: Comparison of Two Factor Analysis Methods. *Environmental Science and Technology*, 42(16), 6007-6014. doi:10.1021/es800219z

Tamura, Y., Takenaka, S., Sugiyama, S., & Nakayama, R. (1998). Occurrence of anserine as an antioxidative dipeptide in a red alga, *Porphyra yezoenis*. *Bioscience, Biotechnology & Biochemistry*, 62(3), 561-563. doi:10.1271/bbb.62.561

Tillett, D., Dittmann, E., Erhard, M., von Döhren, H., Börner, T., & Neilan, B. (2000). Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide–polyketide synthetase system. *Chemistry & Biology*, 7(10), 753-764. doi:10.1016/S1074-5521(00)00021-1

Tong, S. T., & Chen, W. (2002). Modeling the relationship between land use and surface water quality. *Journal of Environmental Management*, 66, 377-393. doi:10.1006/jema.2002.0593

Utermöhl, H. (1931). Neue Wege in der quantitativen Erfassung des Plankton. (Mit besonderer Berücksichtigung des Ultraplanktons.). *Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen*, 5(2), 567-597. doi:10.1080/03680770.1931.11898492

Van Damme, S., Struyf, E., Maris, T., Cox, T., & Meire, P. (2010). Characteristic aspects of the tidal freshwater zone that affect aquatic primary production. In Water quality and the estuarine environment: Spatio temporal patterns and opportunities for restoration with emphasis on nitrogen removal (pp. 15-44).

Vranješ, N., & Jovanović, M. (2011). Cyanotoxins: a dermatological problem. *Archive of Oncology*, 19(3-4), 64-66. doi:10.2298/AOO1104064V

Wang, J., Jiang, X., Zheng, B., Chen, C., Kang, X., Zhang, C., . . . Wang, W. W. (2016). Effect of algal bloom on phosphorus exchange at the sediment– water interface in Meiliang Bay of Taihu Lake, China. *Environmental Earth Science*, 75(57), 1-9. doi:10.1007/s12665-015-4810-z

Wang, Z., Song, G., Li, Y., Yu, G., Hou, X., Gan, Z., & Li, R. (2019). The diversity, origin, and evolutionary analysis of geosmin synthase gene in cyanobacteria. *Science of the Total Environment*, 689(1), 789-796. doi:10.1016/j.scitotenv.2019.06.468

Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R.,. . . Humphrey-Smith, I. (1995). Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis*, 16, 1090-1094. doi:10.1002/elps.11501601185

Watson, S., Broisko, J., & Lalor, J. (2009). Bay of Quinte Harmful Algal Blooms: Phase 1 - 2009. Bay of Quinte Remedial Action Plan: Monitoring Report #20, 27-50.

Wile, I. (1975). Lake restoration through mechanical harvesting of aquatic vegetation. *SIL Proceedings*, 19(1), 1922-2010. doi:10.1080/03680770.1974.11896109

Williams, W. E., Gorton, H. L., Vogelmann, T. C. (2003). Surface gas-exchange processes of snow algae. *Proceedings of the National Academy of Science of the United States of America*. 2, 562-566. doi:10.1073/pnas.0235560100

Wood, S. A., Kelly, L. T., Bouma-Gregson, L., Humbert, J. F., Laughinghouse, I. V., Lazorchak, J., . . . Davis, T. W. (2020). Toxic benthic freshwater cyanobacterial proliferations: Challenges and solutions for enhancing knowledge and improving monitoring and mitigation. *Freshwater Biology*, 80(10), 88-95. doi:10.1111/fwb.13532

Xiong, W., Chen, Z., & Ge, F. (2016). Proteomic analysis of post translational modifications in cyanobacteria. *Journal of Proteomics*, 134(16), 57-64. doi:10.1016/j.jprot.2015.07.037

Yauk, C. L., & Quinn, J. S. (1996). Multilocus DNA fingerprinting reveals high rate of heritable genetic mutation in herring gulls nesting in an industrialized urban site. *Proceedings of the National Academy of Sciences*, 93(22), 12137-12141. doi:10.1073/pnas.93.22.12137

Yokoyama, H., Tamaki, A., Koyama, K., Ishihi, Y., Shimoda, K., & Harada, K. (2005). Isotopic evidence for phytoplankton as a major food source for macrobenthos on an intertidal sandflat in Ariake Sound, Japan. *Marine Ecology Progress Series*, 304, 101-116. doi:10.3354/meps304101

Yurista, P. M., Kelly, J. R., & Scharold, J. V. (2016). Great Lake nearshore-offshore: distinct water quality regions. *Journal of Great Lakes Research*, 42(2), 375-385. doi:10.1016/j.jglr.2015.12.002

Zhang, H., & Yang, C. (2019). Arginine and nitrogen mobilization in cyanobacteria. Molecular Microbiology, 111(4), 863-867. doi:doi:10.1111/mmi.14204

Zhou, Y., Michalak, A. M., Beletsky, D., Rao, Y. R., & Richards, R. P. (2015). Record-Breaking Lake Erie Hypoxia during 2012 Drought. *Environmental Science and Technology*, 49(2), 800-807. doi:10.1021/es503981n

Zou, W., Zhu, G., Cai, Y., Vilmi, A., Xu, H., Zhu, M., . . . Qin, B. (2020). Relationships between nutrient, chlorophyll a and Secchi depth in lakes of the Chinese Eastern Plains ecoregion: Implications for eutrophication management. *Journal of Environmental Management*, 260, 109923. doi:10.1016/j.jenvman.2019.109923

APPENDIX A

Figure A1. Detailed sampling dates for all eight locations in 2020. N/A represents neither phytoplankton nor WQ samples were taken.

Location	Site	Week									
		1	2	3	4	5	6	7	8	9	10
HH	CB	Aug 05*	Aug 10	Aug 17	Aug 24	Aug 31	Sep 08	Sep 14	Sep 21	Sep 28	Oct 05
	BF	N/A	N/A	N/A	N/A	Aug 31	Sep 08	Sep 14	Sep 21	Sep 28	Oct 05
	YC	N/A	N/A	N/A	N/A	Aug 31	Sep 08	Sep 14	Sep 21	Sep 28	Oct 05
BQ	GL	N/A	Aug 13*	Aug 20	Aug 26	Sep 03	Sep 10	Sep 17	Sep 23	Oct 01	Oct 08
	HB	Aug 06*	Aug 13*	Aug 20	Aug 26	Sep 03	Sep 10	Sep 17	Sep 23	Oct 01	Oct 08
	NP	N/A	Aug 13*	Aug 20	Aug 26	Sep 03	Sep 10	Sep 17	Sep 23	Oct 01	Oct 08
	PH	Aug 06*	Aug 13*	Aug 20	Aug 26	Sep 03	Sep 10	Sep 17	Sep 23	Oct 01	Oct 08
	WC	Aug 06*	Aug 13*	Aug 20	Aug 26	Sep 03	Sep 10	Sep 17	Sep 23	Oct 01	Oct 08

* Insufficient WQ samples collected to be included in data analyses

Figure A2. Total number of litres per site and week centrifuged to obtain algal pellets.

N/A = no water was able to be sampled that week.

Location	Site	Week									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
HH	**CB**	9	9	10	9	9	8	9	9	10	10
	BF	N/A	N/A	N/A	N/A	1	5	4	3	1	1
	YC	N/A	N/A	N/A	N/A	4	3	2	2	2	2
BQ	**GL**	N/A	5	5	5	5	8	10	10	10	10
	HB	N/A	3	5	5	6	1	10	10	10	10
	NP	N/A	N/A	3	4	7	8	10	10	10	10
	PH	N/A	1	10	10	10	10	10	10	10	10
	WC	10	N/A	9	8	9	8	10	10	10	10

APPENDIX B

Acquisition Method Info:

Method Name:

metaboHILICneg_msonly_low_pH.m

Device List:

- Multisampler
- Binary Pump
- Column Comp.
- Q-TOF

TOF/Q-TOF Mass Spectrometer

Component Name:	MS Q-TOF
Component Model:	G6545A
Ion Source:	Dual AJS ESI
Stop Time (min):	No Limit/As Pump
Can wait for temp.:	Enable
Fast Polarity:	FALSE
MS Abs. threshold:	1000
MS Rel. threshold(%):	0.01
MS/MS Abs. threshold:	5
MS/MS Rel. threshold(%):	0.01

Time Segments

Time Segment #	Start Time (min)	Diverter Valve State	Storage Mode	Ion Mode
1	0	MS	Both	Dual AJS ESI

Time Segment 1:

Acquisition Mode MS1:

Min Range (m/z)	60
Max Range (m/z)	1600
Scan Rate (spectra/sec)	1

Instrument Parameters:

Parameter	**Value**
Gas Temp (°C)	225
Gas Flow (l/min)	10
Nebulizer (psig)	35

SheathGasTemp	350
SheathGasFlow	12

Scan Segments:

Scan Seg #	Ion Polarity	Collision Energy
1	Negative	0

Scan Segment 1

Scan Source Parameters:

Parameter	Value
VCap	3500
Nozzle Voltage (V)	0
Fragmentor	125
Skimmer1	45
OctopoleRFPeak	750

ReferenceMasses

Ref Mass Enabled	Disabled
Ref Nebulizer (psig)	

Chromatograms

Chrom Type	Label	Offset	Y-Range
TIC	TIC	15	10000000

Name: Multisampler **Module: G7167A**

Sampling Speed	
Draw Speed	100.0 µL/min
Eject Speed	400.0 µL/min
Wait Time After Drawing	1.2 s
Injection	
Needle Wash Mode	Standard Wash
Injection Volume	5.00 µL
Standard Needle Wash	
Needle Wash Mode	Flush Port
Duration	3 s
High Throughput	
Injection Valve to Bypass for Delay Volume Reduction	No
Sample Flush-Out Factor	5
Overlapped Injection	
Overlap Injection Enabled	No
Needle Height Position	
Draw Position Offset	-1.0 mm
Use Vial/Well Bottom Sensing	No
Stop Time	

Stoptime Mode	No Limit
Post Time	
Posttime Mode	Off

Name: Binary Pump G7112B	**Module:**
Flow	0.250 mL/min
Use Solvent Types	Yes
Low Pressure Limit	0.00 bar
High Pressure Limit	500.00 bar
Maximum Flow Gradient	100.000
mL/min²	
Stroke A	
Automatic Stroke Calculation A	Yes
Stroke B	
Automatic Stroke Calculation B	Yes
Stop Time	
Stoptime Mode	Time set
Stoptime	25.00 min
Post Time	
Posttime Mode	Time set
Posttime	1.00 min

Solvent Composition

Channel	Solvent 1	Name 1	Solvent 2	Name 2	Selected	Used	Percent (%)
A	H20		H20		Ch. 2	Yes	2.00%
B	premixed ACN(90%) - H20(10%)		premixed ACN(90%) - H20(10%)		Ch. 2	Yes	98.00%

Timetable

Time (min)	A (%)	B (%)	Flow (mL/min)	Pressure (bar)
0.00 min	2.00%	98.00%	0.250 mL/min	500.00 bar
3.00 min	2.00%	98.00%	0.250 mL/min	500.00 bar
11.00 min	30.00%	70.00%	0.250 mL/min	500.00 bar
12.00 min	40.00%	60.00%	0.250 mL/min	500.00 bar
16.00 min	95.00%	5.00%	0.250 mL/min	500.00 bar
18.00 min	95.00%	5.00%	0.250 mL/min	500.00 bar
19.00 min	2.00%	98.00%	0.250 mL/min	500.00 bar
25.00 min	2.00%	98.00%	0.250 mL/min	500.00 bar

Name: Column Comp.	**Module: G7116A**
Left Temperature Control	
Temperature Control Mode	Temperature Set
Temperature	50.0 °C
Enable Analysis Left Temperature	
Enable Analysis Left Temperature On	Yes
Enable Analysis Left Temperature Value	0.8 °C
Left Temp. Equilibration Time	0.0 min
Right Temperature Control	
Right temperature Control Mode	Temperature Set
Right temperature	50.0 °C
Enable Analysis Right Temperature	
Enable Analysis Right Temperature On	Yes
Enable Analysis Right Temperature Value	0.8 °C
Right Temp. Equilibration Time	0.0 min
Enforce column for run	
Enforce column for run enabled	No
Stop Time	
Stoptime Mode	As pump/injector
Post Time	
Posttime Mode	Off
Timetable	
Valve Position	Position 1 (Port 1->6)
Position Switch After Run	Do not switch

APPENDIX C

Acquisition Method Info:

Method Name:

Agilent_training_peptides_slope4.m

Device List:

- Multisampler
- Binary Pump
- Column Comp.
- Q-TOF

TOF/Q-TOF Mass Spectrometer

Component Name:	MS Q-TOF
Component Model:	G6545A
Ion Source:	Dual AJS ESI
Stop Time (min):	No Limit/As Pump
Can wait for temp.:	Enable
Fast Polarity:	FALSE
MS Abs. threshold:	500
MS Rel. threshold(%):	0.01
MS/MS Abs. threshold:	5
MS/MS Rel. threshold(%):	0.01

Time Segments

Time Segment #	Start Time (min)	Diverter Valve State	Storage Mode	Ion Mode
1	0	MS	Both	Dual AJS ESI

Time Segment 1

Acquisition Mode AutoMS2

MS Min Range (m/z)	200
MS Max Range (m/z)	3000
MS Scan Rate (spectra/sec)	3.00
MS/MS Min Range (m/z)	50
MS/MS Max Range (m/z)	3000
MS/MS Scan Rate (spectra/sec)	2.00
Isolation Width MS/MS	Medium (~4 amu)
Decision Engine	Native

Ramped Collision Energy

Charge	4
Slope	2
Offset	All

Auto MS/MS Preferred/Exclude Table

Mass	921.9686
Delta Mass (ppm)	100
Charge	1
Type	Exclude
Retention Time (min)	0
Delta Ret. Time (min)	
Isolation Width	Narrow (~1.3 amu)
Collision Energy	

Precursor Selection

Max Precursors Per Cycle	10
Threshold (Abs)	500
Threshold (Rel) (%)	0.010
Precursor abundance-based scan speed	Yes
Target (counts/spectrum)	25000.000
Use MS/MS accumulation time limit	Yes
Use dynamic precursor rejection	No
Purity Stringency (%)	100.000
Purity Cut off (%)	30.000
Isotope Model	Peptides
Active exclusion enabled	Yes
Active exclusion excluded after (spectra)	2
Active exclusion released after (min)	0.20
Sort precursors	By abundance only

Static Exclusion Ranges

StartMZ	25
EndMZ	300

Charge State Preference

Selected Charge Preference: 2, 3, >3

Instrument Parameters

Parameter	**Value**
Gas Temp (°C)	325
Gas Flow (l/min)	8
Nebulizer (psig)	35
SheathGasTemp	350
SheathGasFlow	11

Scan Segments

Scan Segment #	1
Ion Polarity	Positive

Scan Segment 1

Scan Source Parameters

Parameter	Value
VCap	4500
Nozzle Voltage (V)	1000
Fragmentor	180
Skimmer1	65
OctopoleRFPeak	750

ReferenceMasses

Ref Mass Enabled	Disabled
Ref Nebulizer (psig)	

Chromatograms

Chrom Type	Label	Offset	Y-Range
TIC	TIC	15	10000000
TIC	TIC	15	10000000

Name: Multisampler **Module: G7167A**

Sampling Speed	
Draw Speed	100.0 µL/min
Eject Speed	400.0 µL/min
Wait Time After Drawing	1.2 s
Injection	
Needle Wash Mode	Standard Wash
Injection Volume	2.00 µL
Standard Needle Wash	
Needle Wash Mode	Flush Port
Duration	10 s
High Throughput	
Injection Valve to Bypass for Delay Volume Reduction	No
Sample Flush-Out Factor	5
Overlapped Injection	
Overlap Injection Enabled	No
Needle Height Position	
Draw Position Offset	-1.0 mm
Use Vial/Well Bottom Sensing	No
Stop Time	
Stoptime Mode	No Limit

Post Time

Posttime Mode	Off

Name: Binary Pump **Module: G7112B**

Flow	0.000 mL/min
Use Solvent Types	Yes
Low Pressure Limit	0.00 bar
High Pressure Limit	400.00 bar
Maximum Flow Gradient	100.000 mL/min²

Stroke A

Automatic Stroke Calculation A	Yes

Stroke B

Automatic Stroke Calculation B	Yes

Stop Time

Stoptime Mode	Time set
Stoptime	50.00 min

Post Time

Posttime Mode	Off

Solvent Composition

Channel	Solvent 1	Name 1	Solvent 2	Name 2	Selected	Used	Percent (%)
A	H_2O		H_2O		Ch. 1	Yes	98.00%
B	premixed ACN(95%) - H_2O(5%)		ACN		Ch. 1	Yes	2.00%

Timetable

Time (min)	A (%)	B (%)	Flow (mL/min)	Pressure (bar)
0.00 min	98.00%	2.00%	0.100 mL/min	400.00 bar
2.00 min	98.00%	2.00%	0.100 mL/min	400.00 bar
27.00 min	60.00%	2.00%	0.100 mL/min	400.00 bar
32.00 min	40.00%	60.0%	0.100 mL/min	400.00 bar
32.01 min	15.00%	85.0%	0.100 mL/min	400.00 bar
37.00 min	15.00%	85.0%	0.100 mL/min	400.00 bar
37.01 min	98.00%	2.00%	0.100 mL/min	400.00 bar

Name: Column Comp. G7116A	**Module:**
Left Temperature Control	
Temperature Control Mode	Temperature Set
Temperature	40.0 °C
Enable Analysis Left Temperature	
Enable Analysis Left Temperature On	Yes
Enable Analysis Left Temperature Value	1.0 °C
Left Temp. Equilibration Time	0.0 min
Right Temperature Control	
Right temperature Control Mode	Temperature Set
Right temperature	40.0 °C
Enable Analysis Right Temperature	
Enable Analysis Right Temperature On	Yes
Enable Analysis Right Temperature Value	0.8 °C
Right Temp. Equilibration Time	0.0 min
Enforce column for run	
Enforce column for run enabled	No
Stop Time	
Stoptime Mode	As pump/injector
Post Time	
Posttime Mode	Off
Timetable	
Valve Position	Position 2 (Port 1->2)
Position Switch After Run	Do not switch

APPENDIX D

Figure D1. List of surrogate isotope-labelled standards used to quantify metabolites without a counterpart standard.

Metabolite	Surrogate Isotope-Labelled Standard
1-Methyl-L-histidine	Histidine-13C6 15N3
3-Methyl-L-histidine	Histidine-13C6 15N3
d-Hydroxylysine	Lysine-13C6 15N2
DL-b-Aminoisobutyric Acid	Serine-13C3 15N
g-Amino-n-butyric Acid	Serine-13C3 15N
Hydroxy-L-proline	Proline-13C5 15N
L-a-Aminoadipic Acid	Serine-13C3 15N
L-a-Amino-n-butyric Acid	Alanine-13C3 15N
L-Anserine	Histidine-13C6 15N3
L-Carnosine	Histidine-13C6 15N3
Taurine	Cystine 13C6 15N2
b-Alanine	Alanine-13C3 15N
L-Ornithine	Glutamine 13C5 15N2
L-Sarcosine	Alanine-13C3 15N
L-Citrulline	Glutamine 13C5 15N2
L-Carnosine	Tyrosine-13C9 15N

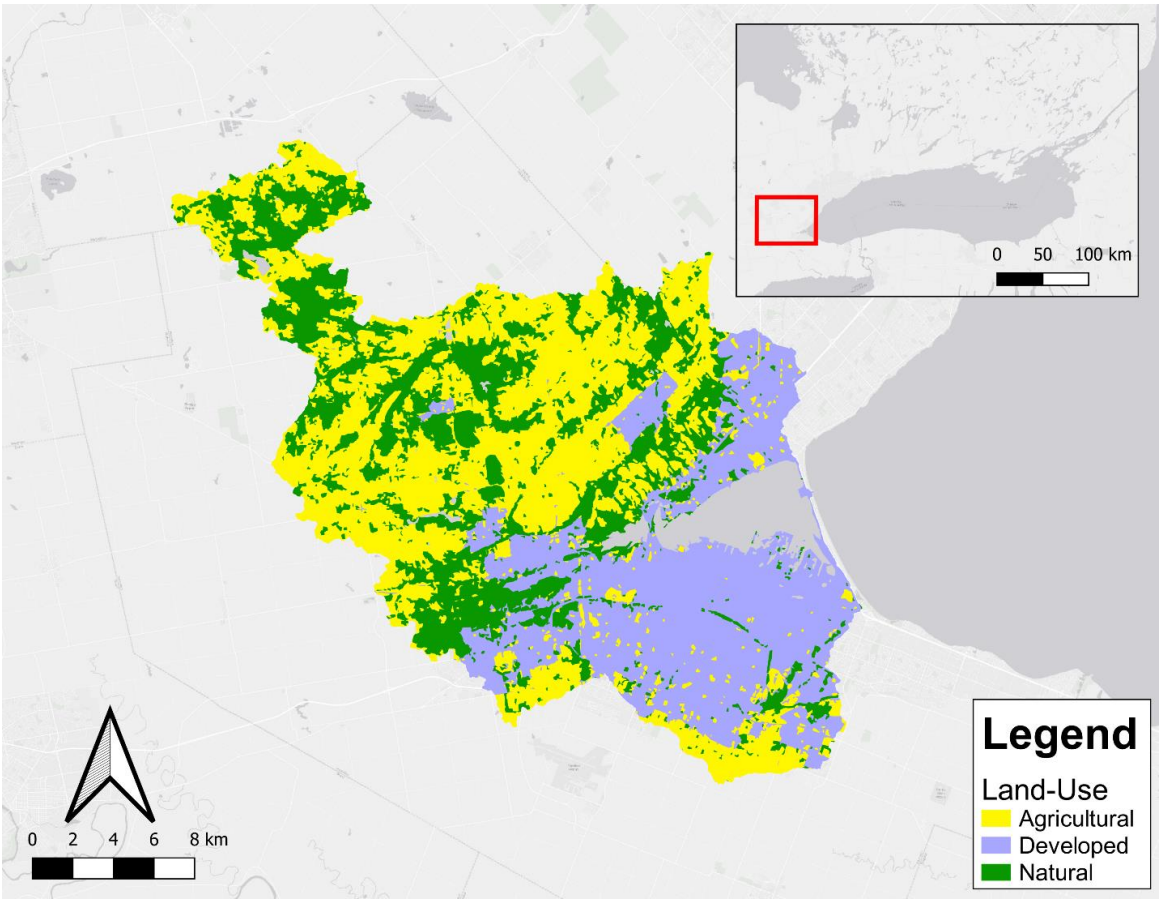

Figure D2. Land-use map of the watershed draining into HH. The total area of the watershed is 52,848 ha, comprised of: 20,662.4 ha (39.1%) of agricultural land; 15,710.6 ha (29.7%) of developed land; 13,816.8 ha (26.1%) of natural land; and 2,658.1 ha (5%) of water.

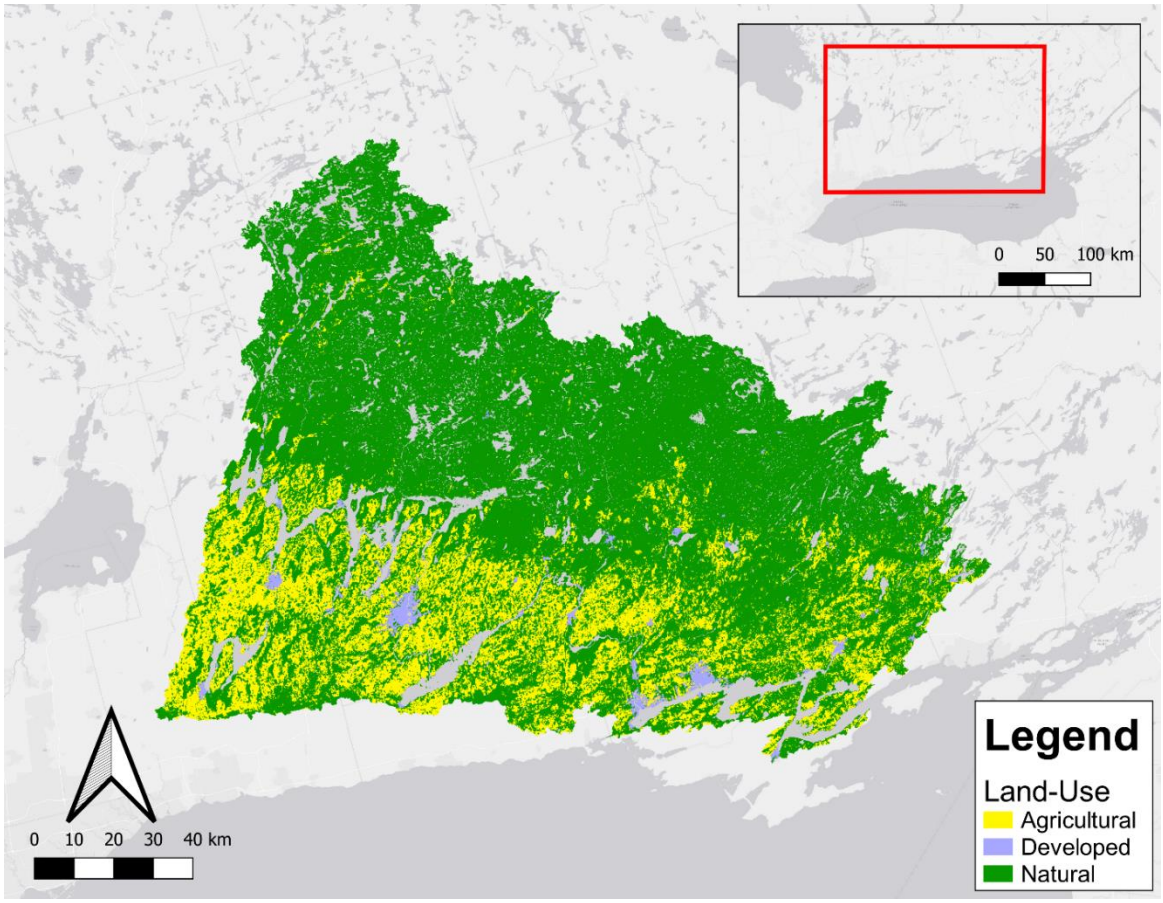

Figure D3. Land-use map of the watershed draining into BW. The total area of the watershed is 1,864,474 ha, comprised of: 1,199,086.9 ha (64.3%) of natural land; 463,323.3 ha (24.9%) of agricultural land; 173,315 ha 9.3%) of water; and 28,748.8 ha (1.5%) of developed land.

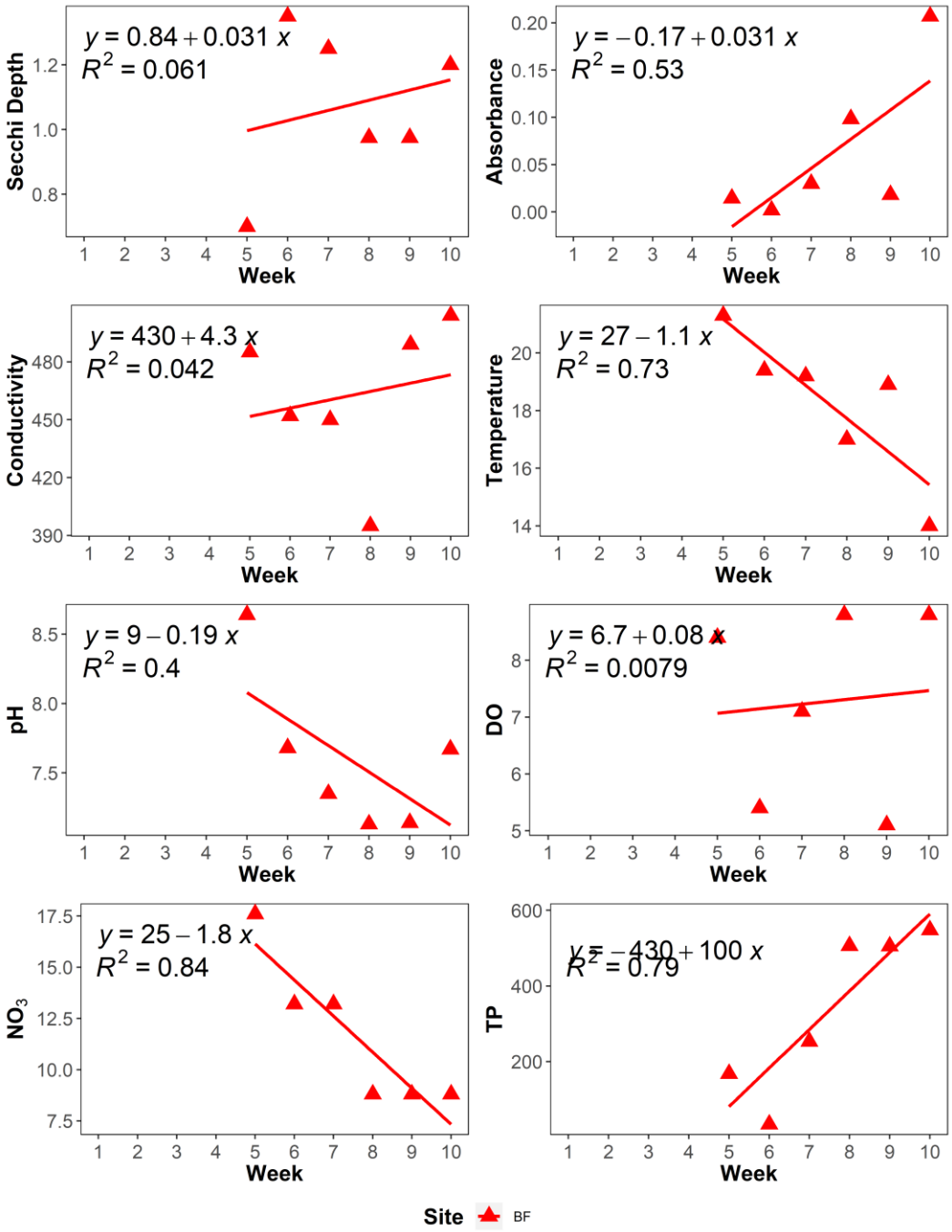

Figure D4. WQ trendline plots for BF in HH. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

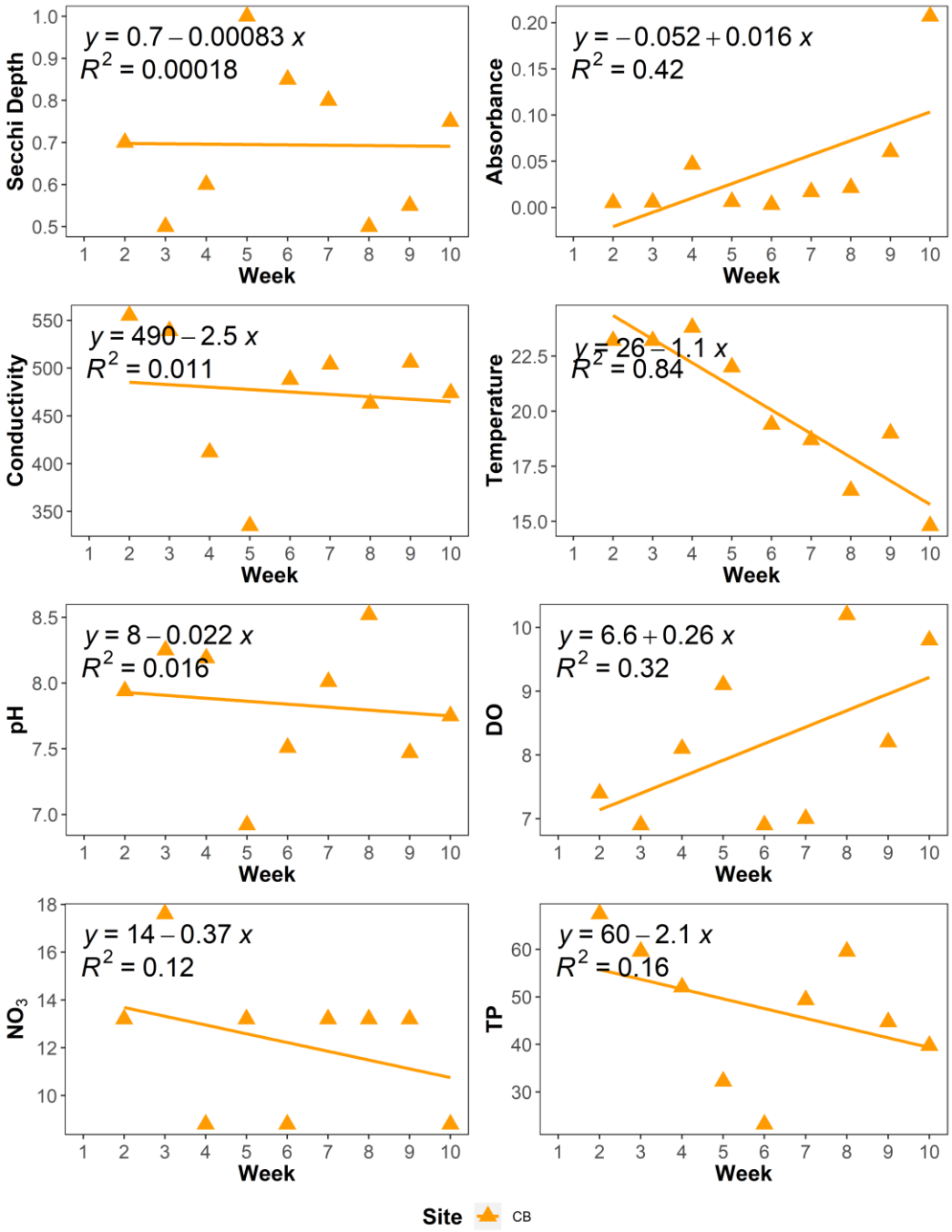

Figure D5. WQ trendline plots for CB in HH. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

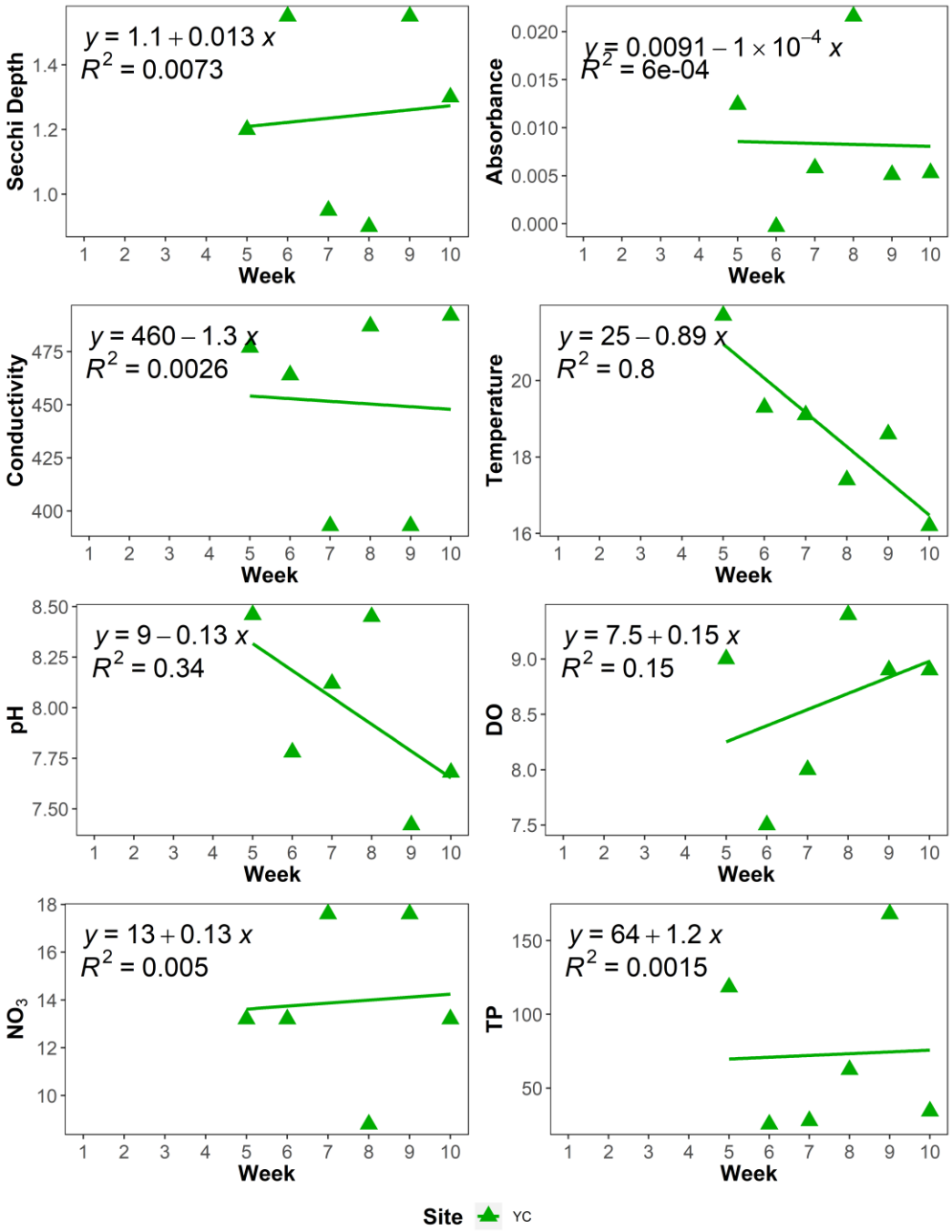

Figure D6. WQ trendline plots for YC in HH. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

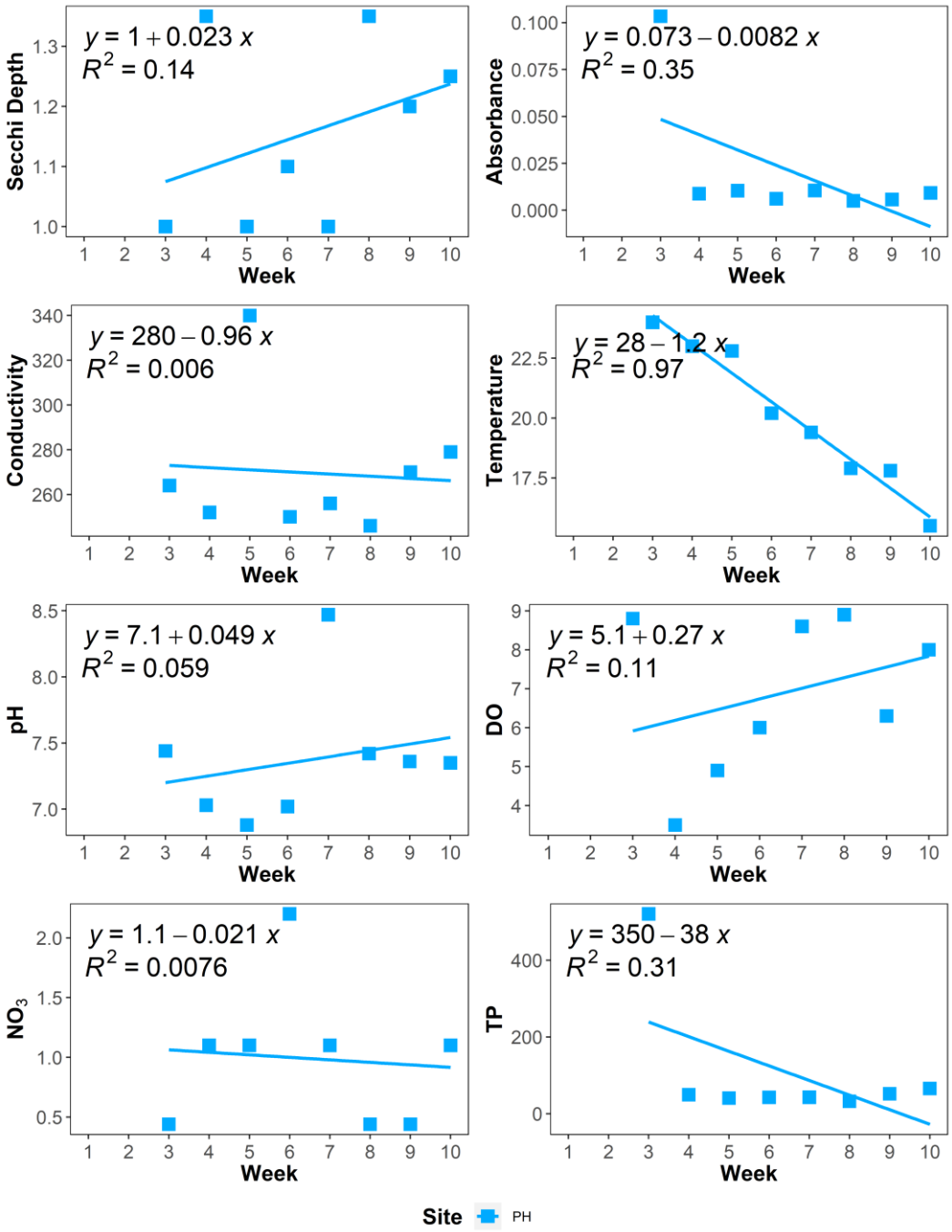

Figure D7. WQ trendline plots for PH in BQ. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

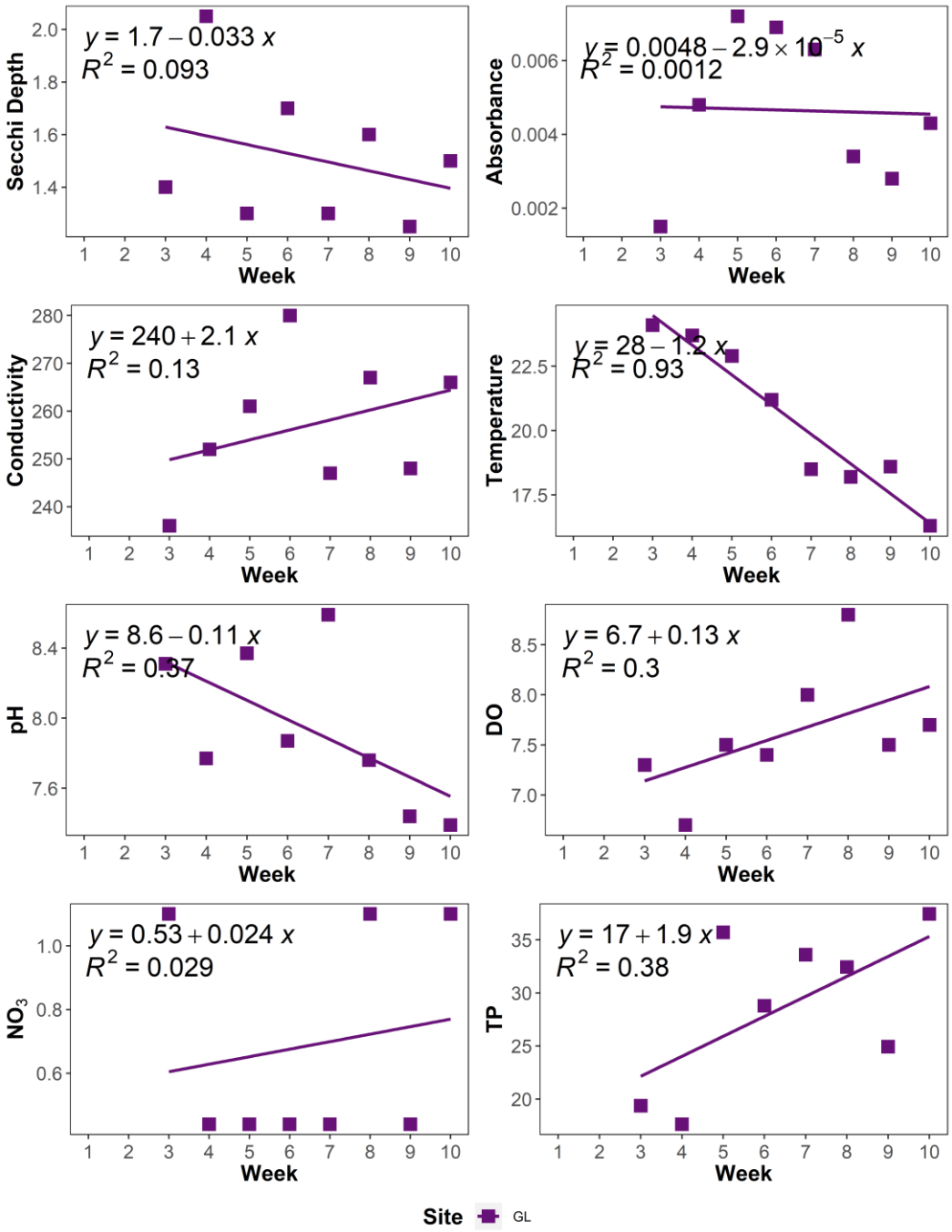

Figure D8. WQ trendline plots for GL in BQ Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

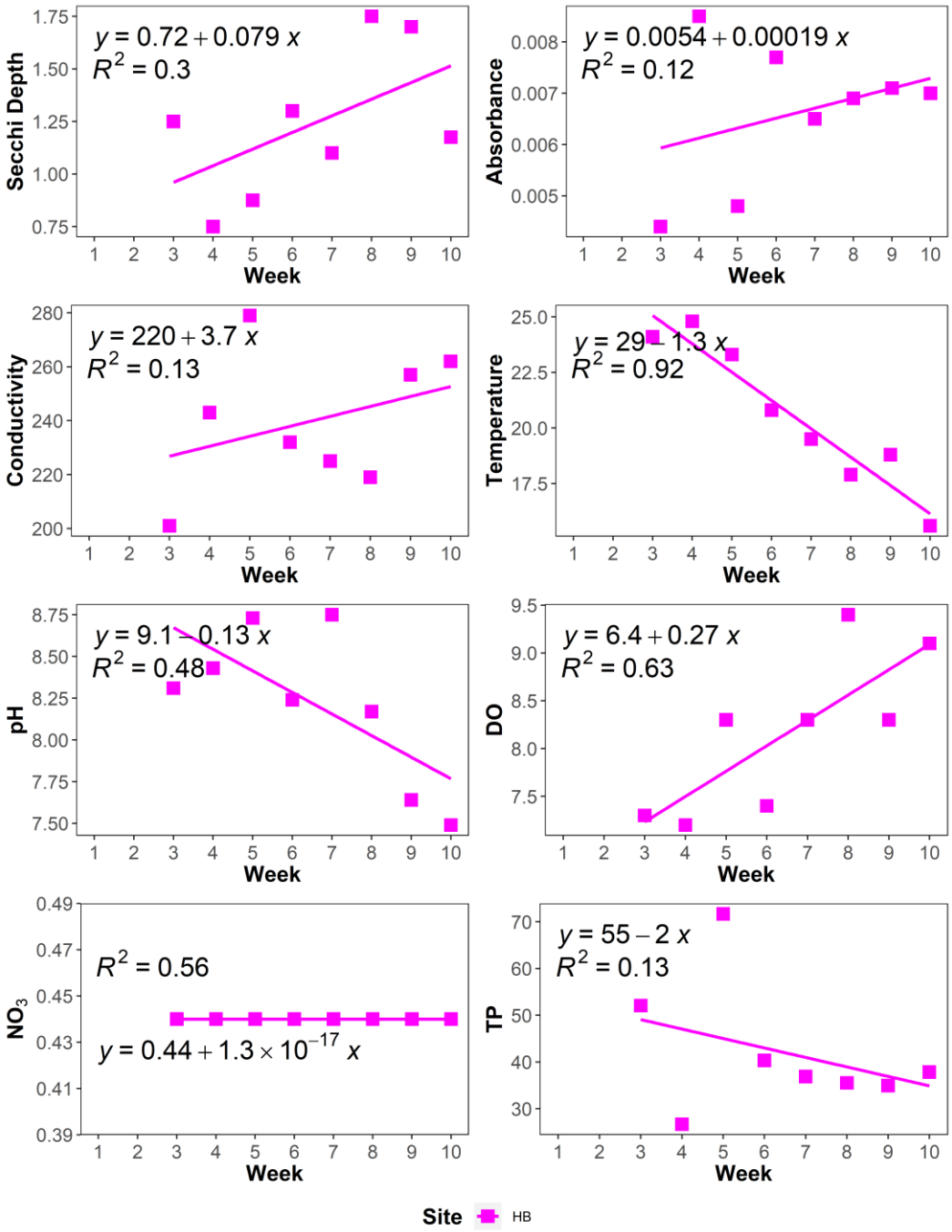

Figure D9. WQ trendline plots for HB in BQ. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

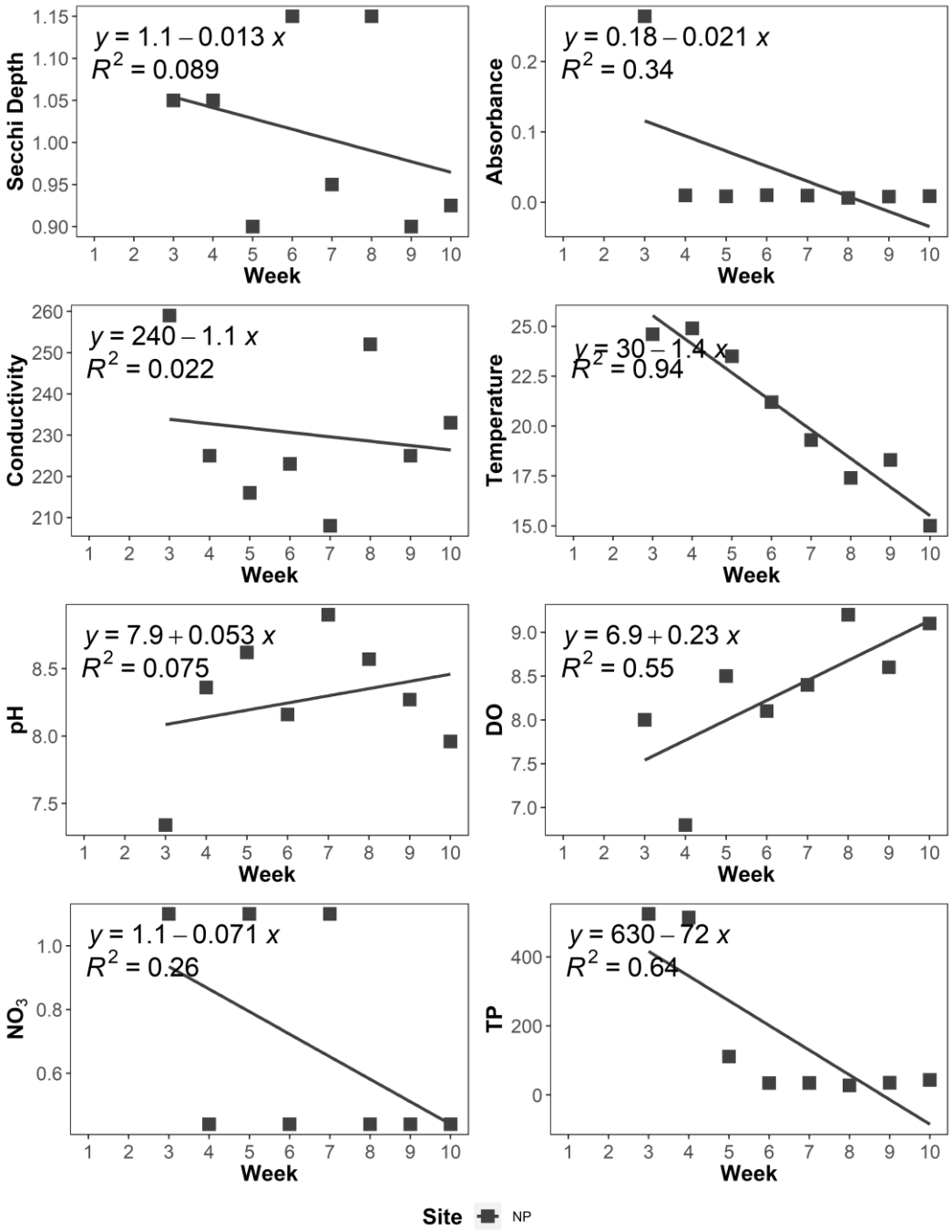

Figure D10. WQ trendline plots for NP in BQ. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

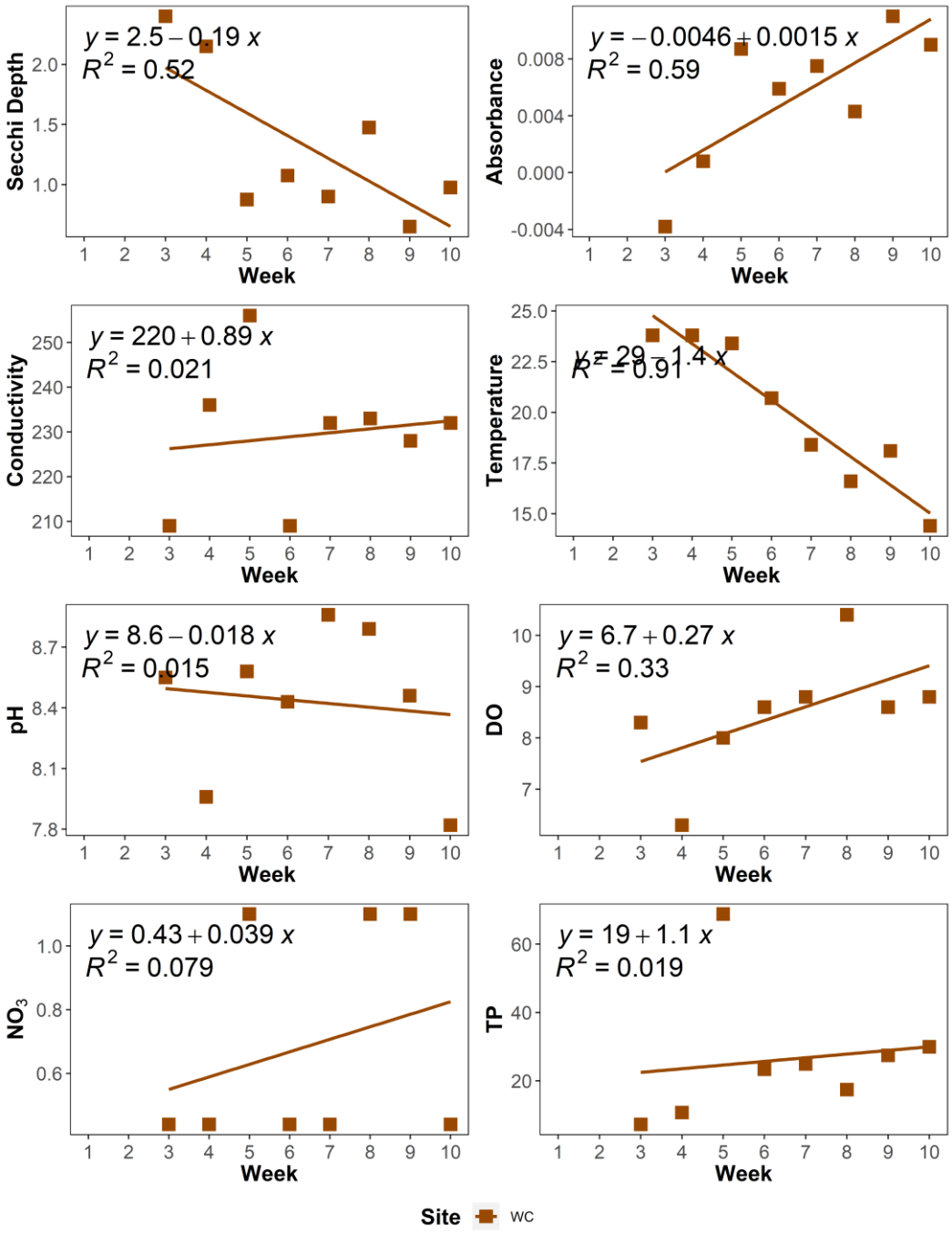

Figure D11. WQ trendline plots for WQ in BQ. Units: Secchi depth (m), Absorbance (TURB), Conductivity (µs/cm), Temperature (°C), DO (ppm), NO_3 (mg/L), and TP (µg/L).

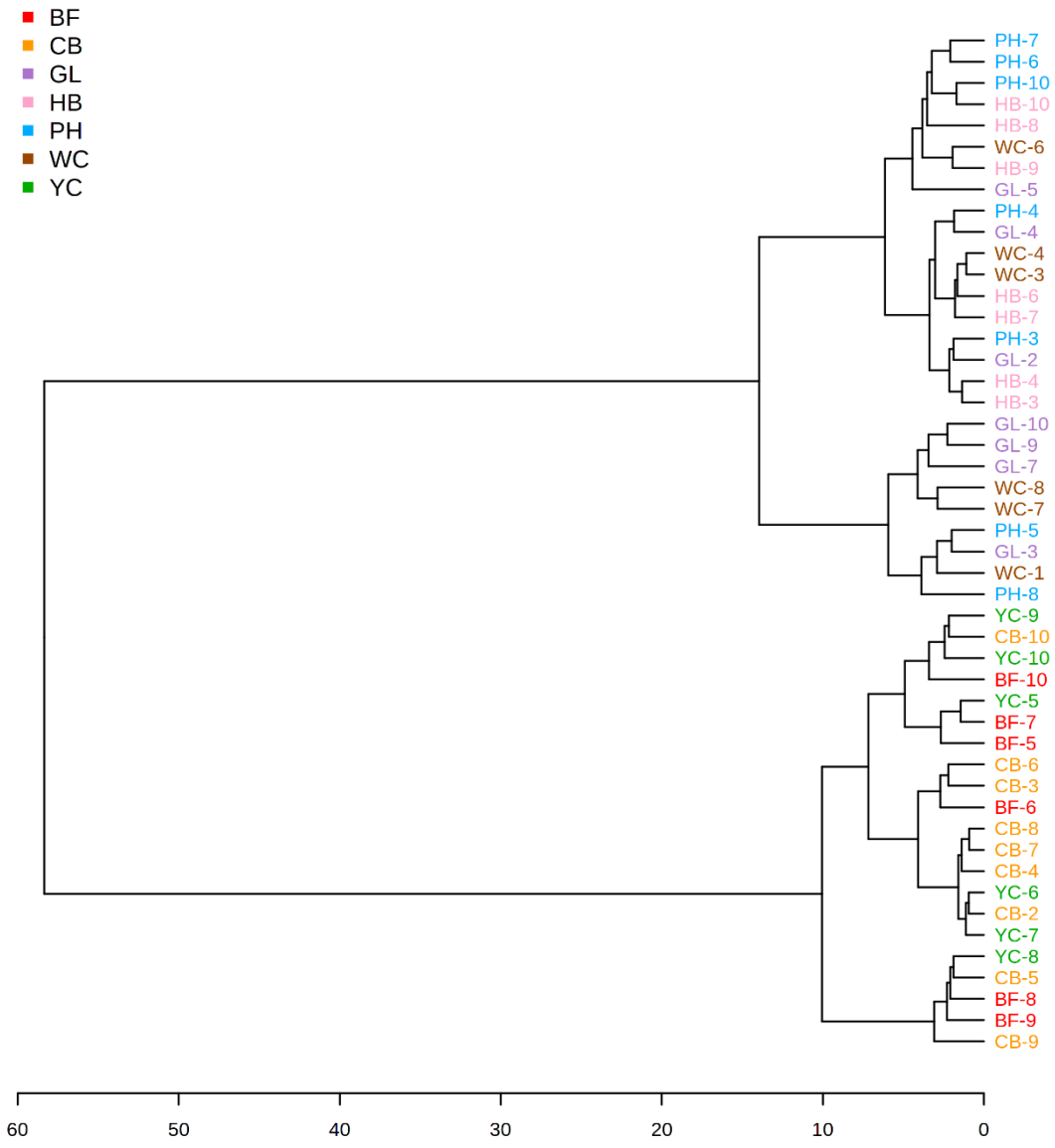

Figure D12. Dendrogram showing the separation in metabolite profiles between HH and BQ sites.

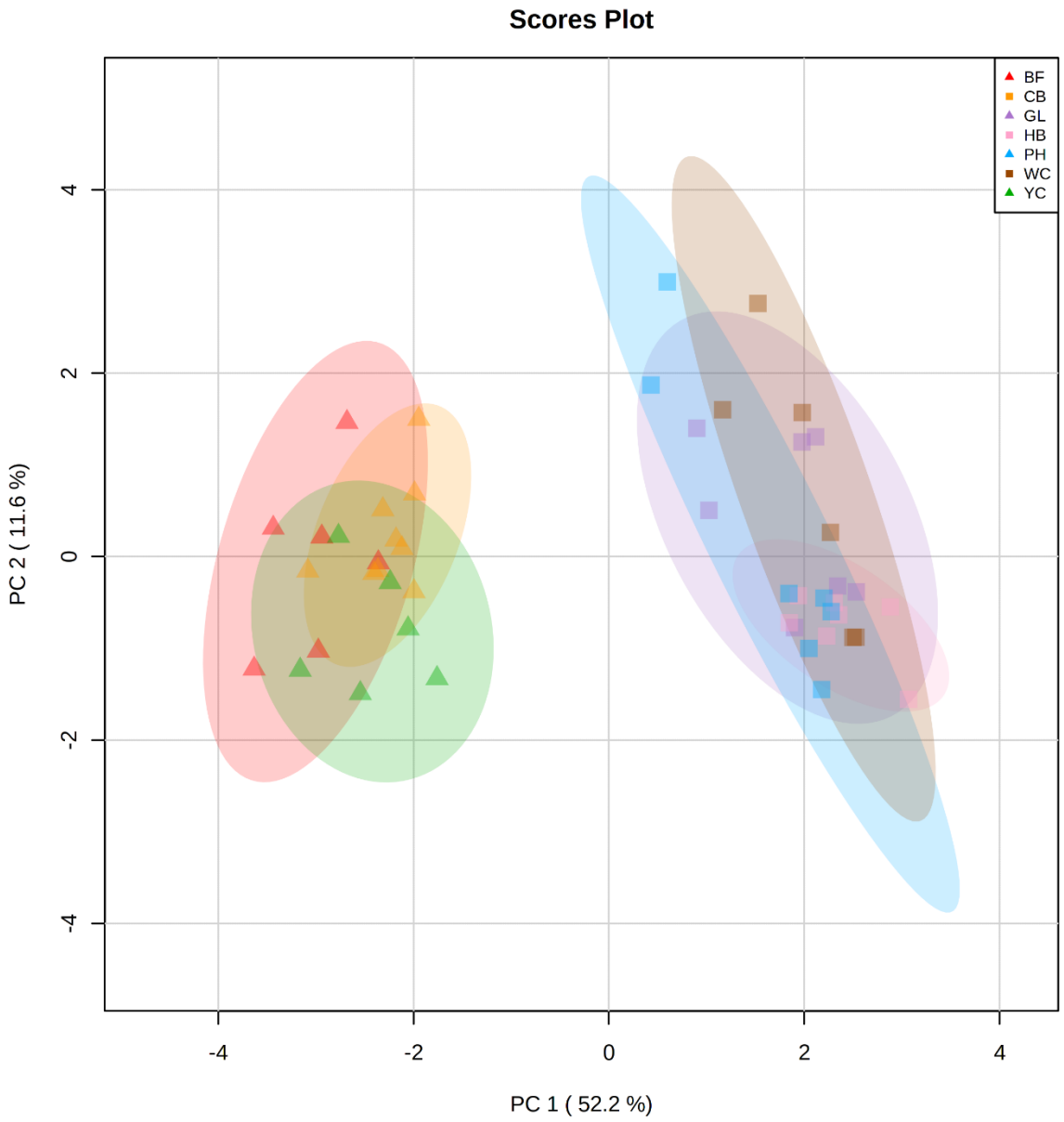

Figure D13. PCA showing the separation between in metabolite profiles between HH and BQ.

Figure D14. Average concentration (ppb) and standard deviation for six metals measured in HH and BQ. Metal samples were taken on week 2, week 6, and week 10.

Metal	HH	BQ
Cd	0.023 ± 0.009	0.018 ± 0.006
Cu^{2+}	0.721 ± 0.21	0.104 ± 0.079
Fe^{2+}	125.98 ± 60	73.84 ± 33
Ni^{2+}	1.337 ± 0.025	0.104 ± 0.015
V	0.711 ± 0.13	1.144 ± 0.941
Zn^{2+}	6.915 ± 0.682	14.527 ± 3.408

Figure D15. Summary table of all 33 metabolites used in this study. Codes, R-group structure, surrogate for quantification (if applicable), and the location in which each metabolite was found are detailed. 21 of the 22 proteinogenic amino acids are bolded (cystine was not detected).

Metabolite	Shorthand	R Group Classification	Surrogate for Quantification (if applicable)	Location in which Found
1-Methyl-L-histidine	1MeH	Hydrophobic	Histidine-13C6 15N3	Both
3-Methyl-L-histidine	3MeH	Hydrophobic	Histidine-13C6 15N3	Both
α-Aminoadipic Acid	Aad	Polar, uncharged	Serine-13C3 15N	Both
α-Aminobutryic Acid	AABA	Nonpolar, aliphatic	Alanine-13C3 15N	Both
β-Alanine	bAla	Nonpolar, aliphatic	Alanine-13C3 15N	Both
γ-Aminobutyric Acid	GABA	Polar, uncharged	Serine-13C3 15N	Both
d-Hydroxylysine	OHLys	Positively charged	Lysine-13C6 15N2	HH only
DL-β-Aminoisobutyric Acid	BAIBA	Polar, uncharged	Serine-13C3 15N	Both
Alanine	Ala	Nonpolar, aliphatic	-	Both
Anserine	Ans	Positively charged	Histidine-13C6 15N3	Both
Arginine	Arg	Positively charged	-	Both
Asparagine	Asp	Polar, uncharged	-	Both
Aspartic Acid	Asp	Negatively charged	-	Both
Carnosine	Car	Positively charged	Histidine-13C6 15N3	BQ only
Citrulline	Cit	Polar, uncharged	Glutamine 13C5 15N2	BQ only
Glutamic Acid	Glu	Negatively charged	-	Both
Glycine	Gly	Nonpolar, aliphatic	-	Both
Histidine	His	Positively charged	-	Both
Hydroxy-L-proline	OHPro	Polar, uncharged	Proline-13C5 15N	Both
Isoleucine	Ile	Nonpolar, aliphatic	-	Both
Leucine	Leu	Nonpolar, aliphatic	-	Both
Lysine	Lys	Positively charged	-	Both
Methionine	Met	Nonpolar, aliphatic	-	Both
Ornithine	Orn	Polar, uncharged	Glutamine 13C5 15N2	BQ only
Phenylalanine	Phe	Aromatic	-	Both
Proline	Pro	Polar, uncharged	-	Both
Sarcosine	Sar	Nonpolar, aliphatic	Alanine-13C3 15N	BQ only
Serine	Ser	Polar, uncharged	-	Both
Taurine	Tau	Polar, uncharged	Cystine 13C6 15N2	BQ only
Threonine	Thr	Polar, uncharged	-	Both
Tryptophan	Trp	Aromatic	-	Both
Tyrosine	Tyr	Aromatic	-	Both
Valine	Val	Nonpolar, aliphatic	-	Both