Identification and characterization of toxic cyanobacteria in two forested maritime watersheds in North America

by

Timothy James Shardlow

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2021

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The following collected my samples and provided environmental data: members of the Jamieson Lab from Dalhousie University, Wendy Krkošek, Beth Lowe and Alanna Fowler of Halifax Water, MSc candidate Alyssa Bourgeois of the Tank Lab at the University of Alberta, members of the Comox Valley Regional District (CVRD), MSc student Hannah McSorley of the Johnson Lab at the University of British Columbia and members of the Capital Regional District (CRD).

Dr. Trevor Charles, Dr. Michael Lynch and Dr. Jiujun Cheng of Metagenom Bio Inc. (Waterloo, Ontario) sequenced the V4 region of the 16S rRNA gene from DNA samples. They also sequenced the AMT region of the *mcy*E gene from DNA samples and *geo*A genes using novel *geo*A primers they developed.

Abstract

Healthy forested watersheds naturally provide high quality drinking water to communities. However, the integrity and quality of these water sources may be threatened by climate change-exacerbated disturbances such as wildfires and hurricanes, which can lead to increased delivery of nutrients to receiving waters and increase in water temperatures. This can result in the proliferation of cyanobacteria that may threaten water quality through the production of toxins as well as taste and odour (T&O) compounds. Hence, it is important to detect the presence of potentially harmful cyanobacteria prior to any associated water quality shifts. In this study, a synoptic field sampling campaign was conducted in 2019 and involved the collection of one water sample per month in various watersheds. To evaluate the composition and relative abundance of cyanobacteria, water samples were collected from the Pockwock Lake watershed (Nova Scotia, Canada) in June, August, September and October, the Comox Lake watershed (British Columbia, Canada) in May and September, and the Leech River and Sooke River watersheds (British Columbia, Canada) in July and August. Microbial DNA was extracted from water samples for 16S rRNA gene sequencing and assigned taxonomy in QIIME2 using a SILVA classifier and resulting cyanobacteria ASVs were analyzed using the R package *mirlyn*. Lakes within the same watershed typically contained similar communities, though monthly variations in diversity were observed in some lakes. Most cyanobacteria ASVs resolved to the genus-level were assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1); across all samples, the relative abundance of reads from these genera was 51% and 42%, respectively. Other genera represented in the samples included cyanobacteria strains known to form blooms and produce geosmin, a terpene with an earthy odour, and microcystin, a regulated hepatotoxin. Findings from this study provide insights into the presence of cyanobacteria in water resources replied upon as drinking water supplies and underscores some potential commonalities for their potential to deteriorate water quality and challenge drinking water treatability in diverse forested watersheds in Canada.

Acknowledgements

To my supervisors Dr. Kirsten Müller and Dr. Monica Emelko, thank you for the opportunities you have provided me and all of your support. I feel like I have come a long way since being a 4th year student desperately looking for a BIOL 499 research position and I have both of you to thank. And to my committee members Dr. Hugh Broders and Dr. Dail Laughinghouse, thank you for your contributions and feedback.

I would like to thank members of the *for*Water NSERC Network and our collaborators for all of the assistance they provided to make my project possible. Also, a special thank you to Dr. Trevor Charles, Dr. Michael Lynch and Dr. Jiujun Cheng of Metagenom Bio Inc. for their endless patience with my work requests and questions.

I would like to thank Ellen Cameron for answering my endless questions regarding microbial community analyses, use of various computational tools, how to use R and many more. Our 12-hour car rides to your field sampling site were a lot of fun and I'll miss them. And to all of my lab mates from the Müller lab, you were always supportive and helpful every step of the way through my studies, going back to when I was a BIOL 499 student, for which I am truly grateful.

To all my friends and family, thank you for your support.

"From the day of your birth, till you ride in the hearse, nothing's so bad that it couldn't be worse." – Bruce McCoskery

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMT	Aminotransferase
*ana*C	anatoxin-a synthetase C adenylating protein
*ana*F	anatoxin-a synthetase F polyketide synthase
ASV	Amplicon Sequence Variant
BLAST	Basic Local Alignment Search Tool
cnb	Cyclic Nucleotide Binding Protein
CRD	Capital Regional District
CVRD	Comox Valley Regional District
DADA2	Divisive Amplicon Denoising Algorithm 2
*geo*A	Gene encoding geosmin synthase
HQP	Highly Qualified Personnel
HRM	Halifax Regional Municipality
MC	Microcystin
*mcy*E	Gene encoding microcystin synthetase
MEGA	Molecular Evolutionary Genetics Analysis
MIB	2-Methylisoborneol
mirlyn	Multiple Iterations of Rarefaction for Library Normalization
MSA	Multiple Sequence Alignment
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
N	Nitrogen
NCBI	National Center for Biotechnology Information
ND	No Data
NRPS	Non-Ribosomal Peptide Synthetase
P	Phosphorus
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PKS	Polyketide Synthase
QIIME2	Quantitative Insights Into Microbial Ecology 2
rRNA	Ribosomal RNA
V4	Variable Region 4

Chapter 1: Introduction and Literature Review

1.1 Overview

Cyanobacteria are a concern for water quality as they can form blooms and scum on surface waters and are increasing in frequency and severity globally (Paerl, 2018). Climate change-exacerbated landscape disturbances and anthropogenic activities leading to nutrient inputs and increasing water temperatures are widely recognized drivers of cyanobacterial proliferation that can cause shifts in quality of water supplies that last for decades or longer (Emelko *et al*., 2016; Paerl, 2018). Cyanobacterial blooms can threaten the integrity and quality of drinking water sources as they can clog treatment process (Emelko *et al*., 2011) and potentially produce cyanotoxins, such as microcystin which causes liver damage in humans (Carmichael, 1992). Cyanobacteria are also able to produce the taste and odour compounds geosmin and 2-methylisoborneol (MIB) and, while not toxic, these compounds result in the public questioning the quality and safety of drinking water (Giglio *et al*., 2008; 2010). Across North America, communities rely on forested watersheds as high-quality supplies of drinking water. As the frequency and severity of cyanobacterial blooms is expected to increase (Chapra *et al.*, 2017), drinking water source quality may deteriorate (Lopes *et al*., 2018) and drinking water treatability may be challenged to the point of inability to meet demands, service disruptions, or even outages (Emelko *et al*., 2011). Hence, these forested watersheds are valuable resources that must be protected and maintained. This study examined the cyanobacteria composition and abundance in drinking water sources within forested watersheds in two maritime regions of Canada. The purpose of this study was to obtain a baseline perspective of the cyanobacterial communities present in maritime watersheds and determine if cyanobacteria capable of reducing drinking water quality were present.

1.2 Forested Watersheds

Forested watersheds can provide communities with what is often an under-appreciated resource: high quality drinking water (Ernst, 2004). Forested watersheds are critical to communities that rely on them for not only providing them with their drinking

water, but to naturally maintain high source quality (Ernst, 2004; Lopes *et al*., 2018). Healthy forests naturally provide high-quality source water by reducing runoff and filtering out nutrients such as nitrogen (N) and phosphorus (P) (Lopes *et al*., 2018). However, anthropogenic and climate change-exacerbated landscape disturbances threaten forest health and integrity, as well as the services they produce, including the provision of high-quality source water (Emelko *et al*., 2011; Robinne *et al*., 2019). These disturbances result in reduced forest cover, which in turn deteriorates source water quality and increases water treatment costs and operational challenges (Ernst, 2004; Postel and Thompson, 2005; Emelko *et al.,* 2011; Price *et al.,* 2017). For example, Ernst (2004) noted that watersheds with a 10% forest cover, compared to those with 60% forest cover, have a 211% increase in water treatment costs. By protecting forested watersheds, the cost to maintain safe drinking water quality can be reduced (Ernst, 2004). Given that healthy forests filter solids and associated contaminants, such as P, which is typically limiting in fresh water, deterioration of forests can lead to proliferation of cyanobacteria (Emelko *et al.,* 2011; 2016; Silins *et al.,* 2014). This in turn can challenge drinking water treatment or lead to service outages (Emelko *et al.,* 2011). It is for these reasons that forested watersheds provide a valuable resource of high-quality source water which emphasizes their importance.

1.2.1 Atlantic Maritime Ecozone Watershed

Within the Atlantic maritime ecozone is the Pockwock Lake watershed, a largely forested watershed (>90% forest cover) in Nova Scotia, Canada and contains Pockwock Lake, the primary drinking water source for the Halifax Regional Municipality (HRM) (Dunnington *et al*., 2018). Dunnington *et al*. (2018) determined that anthropogenic nutrient inputs and deforestation are primary disturbances to Pockwock Lake. As this water source is located off the Atlantic coast, climate change-exacerbated disturbances such as hurricanes and major storm events are common and are likely to become more frequent (Klotzbach *et al*., 2018). Additionally, Pockwock Lake has previously experienced water quality shifts from lake acidification through sulfate deposition, causing alterations to pH levels to shift closer to 6, which potentially support cyanobacteria growth and geosmin production

(Anderson *et al*., 2017). In Fall 2012, unpleasant tastes and odours were reported in Pockwock Lake; they were later attributed to geosmin produced by the cyanobacteria *Anabaena* (Anderson *et al*., 2017). It is thought that Island Lake, which feeds into Pockwock Lake, is where the presence of geosmin originates from (W. Krkošek, personal communication, June 5, 2019). However, analysis of cyanobacteria communities was not conducted in the study by Anderson *et al*. (2017) nor have they been reported in literature. Accordingly, an improved understanding of the cyanobacteria present in this drinking water source and the factors that drive proliferation is essential to ensuring future water security in the area.

1.2.2 Pacific Maritime Ecozone Watersheds

Disturbances common to the Pacific maritime ecozone include wildfires which are becoming more frequent due to climate change-exacerbated disturbances and periods of drought (Mitton and Ferrenberg, 2012; Talucci and Krawchuk, 2019). Additionally, pests such as the Mountain Pine Beetle, may increase wildfire susceptibility as they create dry conditions from high pine tree mortality within this region (Mitton and Ferrenberg, 2012; Talucci and Krawchuk, 2019). Within this ecozone is the Comox Lake watershed on Vancouver Island, British Columbia, Canada and includes Comox Lake which is utilized by the Comox Valley Regional District (CVRD) to supply drinking water to the communities of Courtenay and Comox (Chandran and Mazumder, 2015a). Previous studies from Comox Lake involved identifying temporal variations, diversity and presence of pathogenic *Escherichia coli* within this water body (Chandran and Mazumder, 2015a; 2015b). To my knowledge, there is no published literature that examined the cyanobacterial communities from Comox Lake.

Other watersheds in this ecozone are the Sooke River and Leech River watersheds. Within the Sooke River watershed is the Sooke Lake Reservoir which it utilized by the Greater Victoria Drinking Water Supply System to provide drinking water for the Greater Victoria Region. Adjacent to the Sooke Lake Reservoir is Deception Reservoir which is planned to one day be used as a supplementary drinking water source to the Greater Victoria

Region (H. McSorley, personal communication, January 2, 2020). Within the Leech River watershed is Jarvis Lake and Weeks Lake, which waters from these lakes flow downstream through the Leech River and via a tunnel transfer into Deception Reservoir (H. McSorley, personal communication, January 2, 2020). Jarvis Lake and Weeks Lake were previously recreational sites prior to being purchased by the CRD, and forest harvesting occurred northwest of Weeks Lake in 2018 and 2019 (H. McSorley, personal communication, January 2, 2020).

In the 1980s, following flushing of water through the tunnel that connects Leech River to Deception Reservoir, there were odour problems with the water (H. McSorley, personal communication, January 2, 2020). While no published literature is available regarding the source of these odours problems, and no literature currently exists to describe cyanobacterial communities in these water bodies, it is very possible these issues were associated with taste/odour producing cyanobacteria. Additionally, the drinking water supply in this region is unfiltered; the only treatment is disinfection. Thus, in the event of elevated turbidity and proliferation of cyanobacteria, disinfection alone would be insufficient, and the treatment system would likely be shutdown (Emelko *et al.*, 2011). Notably, cyanobacteria blooms can challenge treatment to the point of service disruptions even when more extensive treatment is available; thus, a better understanding of the factors that contribute to blooms is critical from risk management and drinking water perspectives (Emelko *et al.*, 2011; Skwaruk *et al.*, 2020).

1.3 Cyanobacteria

Cyanobacteria are photosynthetic microorganisms that are present in the fossil record dating back 3.5 billion years and these organisms played an important role in shifting oxygen levels on Earth from being anoxic to oxygenic through the process of photosynthesis (Schirrmeister *et al.*, 2015). Currently these organisms are globally ubiquitous and are observed as a frequent component of freshwater habitats (Walter *et al.*, 2017). In addition, cyanobacteria in these habitats have numerous adaptations to allow for their successful survival, such as specialized gas vesicles to allow for buoyancy to exploit resources in the

water column, such as warm water temperatures and light availability, or the ability to fix atmospheric nitrogen, enabling these organisms to inhabit low nutrient water sources (Paerl, 2018). This is particularly a concern as the frequency of cyanobacteria blooms are increasing in freshwater habitats, thus resulting in drinking water quality reductions and concerns associated with human health impacts (Chapra *et al.*, 2017).

1.3.1 Cyanobacterial Blooms

Cyanobacteria pose a global concern as they can form blooms in aquatic habitats, reducing water quality through alterations to food webs by creating hypoxic conditions (Zilius *et al*., 2014), potentially produce toxins (Chorus and Bartram, 1999) and taste and odour compounds (Giglio *et al*., 2008; 2010). Hypoxic conditions arise when cyanobacterial blooms die off and other bacteria degrade cyanobacterial cells, thereby decreasing dissolved oxygen levels and leading to death of fish and invertebrates (Zilius *et al*., 2014). Toxins that cyanobacteria produce can also reduce water quality as they can pose human health concerns if consumed, which is particularly concerning if present in drinking water sources (Chorus and Bartram, 1999; Otten *et al*., 2016; Müller *et al.*, 2017). There can also be significant water treatment costs to address quality issues associated with cyanobacterial blooms and the presence of toxins within drinking water sources; in some cases, these disruptions can lead to outages and plant shutdowns (Emelko *et al.*, 2011; Otten *et al*., 2016). Additionally, the presence of taste and odour compounds that cyanobacteria produce can result in high water treatment costs (Dunlap *et al*., 2015). While not harmful, taste and odour compounds are unpleasant to consume and give a negative public perception of the quality of their drinking water (Giglio *et al.*, 2008; 2010). Thus, an improved understanding of the potential for cyanobacterial bloom occurrence in drinking water supplies is critical for drinking water treatment infrastructure designs, assurance of operation response capacity to these events, and risk management (Emelko *et al*., 2011; Nunes *et al.*, 2018).

1.3.2 Factors that Influence Cyanobacterial Bloom Formation

1.3.2.1 Nutrient Input

Nutrient inputs into water sources provides a readily available resource for cyanobacteria to utilize. In water sources that are nutrient-rich, cyanobacterial abundance increases as nitrogen (N) and phosphorus (P) are bioavailable (Richardson *et al*., 2019). In freshwater sources, P is typically a limiting nutrient for cyanobacterial growth and primary productivity, thus, reductions in P inputs into water sources has traditionally been the focus to prevent the formation of cyanobacterial blooms (Schindler, 1977; Schindler *et al.,* 2008). Nitrogen can also be limited in some freshwater sources; however, certain genera of cyanobacteria can fix atmospheric N_2 using the nitrogenase enzyme within specialized cells called heterocytes to supply their own usable source of N rather than relying on external inputs (Findlay *et al.,* 1994). Recently, the combined removal and reduction of N and P has been identified as the most successful method for limiting cyanobacteria biomass in freshwater sources (Schindler *et al.,* 2008; Paerl *et al.,* 2016). However, climate change-exacerbated disturbances can influence anthropogenic loading of these nutrients into water sources, making them readily available and potentially causing the formation of cyanobacterial blooms (Schindler, 1977; Schindler *et al.,* 2008; Paerl *et al*., 2019). Therefore, it is important to understand how loading of N and P occurs and the factors that influence it.

The frequency of storm events such as hurricanes are increasing due to climate change, resulting in increased susceptibility of cyanobacterial blooms in water sources. This is due to storm events resuspending nutrients from sediment into the water column from high winds, which is more common in shallower lakes, and from precipitation-mediated inputs of nutrient-rich landscape runoff, influencing cyanobacterial bloom formation (Paerl *et al*., 2016; 2019). An example of this is from Lake Okeechobee, Florida, USA in 2005 when high rainfall caused nutrient-rich runoff to flow into this shallow lake as well as high winds resuspended nutrients from sediment, resulting in a lake-wide bloom (Phlips *et al.,* 2020).

Additionally, wildfires are becoming more frequent and severe due to climate change through alterations in weather patterns, periods of drought and dry conditions brought about

by pests (Mitton and Ferrenberg, 2012; Talucci and Krawchuk, 2019). Post-wildfire occurrence, burned watersheds can experience increased nutrient inputs through nutrient-rich runoff because of decreased precipitation interception and increase in rainfall that comes into contact with the ground surface (Silins *et al.*, 2009; 2014; Williams *et al*., 2019). The compounded effects of climate change-exacerbated disturbances are significant as they can substantially impact water quality in forested watersheds, leading to conditions that are more likely to promote cyanobacterial proliferation through nutrient inputs (Silins *et al*., 2009, 2014; Emelko *et al.*, 2011, 2016).

1.3.2.2 Increasing Water Temperature

Increased water temperatures have been influenced by climate change through global warming which poses an issue to water sources as cyanobacterial blooms may become more frequent and intense as cyanobacteria grow optimally at higher water temperatures (>25°C) (De Senerpont Domis *et al*., 2007; Chapra *et al.*, 2017). Enzymatic activity of nitrogenase is also increased at warmer temperatures, increasing nitrogen fixation rate which can benefit growth of N_2 fixing cyanobacteria (Brauer *et al.*, 2013). Additionally, warmer temperatures can extend the growing capabilities of cyanobacteria as vertical stratification of water sources increases with higher temperatures, allowing for longer growing periods, particularly for buoyant cyanobacteria as they can exploit warm surface waters for longer (Wagner and Adrian, 2009). Therefore, the ability of cyanobacteria to exploit warmer water temperatures is a concern as global temperatures continue to rise.

Climate change-exacerbated disturbances such as wildfires and human activities including forest harvesting can result in substantial canopy loss in forested watersheds. As forests inherently provide shade to local landscapes, providing cooler temperatures through canopy cover, this loss in coverage trees provide results in more direct sunlight that penetrate water sources (Ellison *et al*., 2017). With more direct sunlight, surface water temperatures can increase, again providing optimal growth temperatures for cyanobacteria (De Senerpont Domis *et al*., 2007). These impacts can also propagate for tens of kilometers downstream and last for decades or more, particularly after severe wildfire in some physiographic settings

(Silins *et al.*, 2014; Wagner *et al.*, 2014; Emelko *et al.*, 2016; Ellison *et al*., 2017). Losses in tree cover may especially compromise drinking water treatability for communities reliant on source water that originates in forested environments.

1.3.3 Cyanotoxins

The presence of cyanobacteria in drinking water supplies and recreational waters is cause for concern as some genera are capable of secreting toxic secondary metabolites called cyanotoxins (Du *et al.*, 2019). These toxins are grouped into the categories of cytotoxins, dermatotoxins, hepatotoxins and neurotoxins (Corbel *et al*., 2014). Structurally, these toxins exist as alkaloids, cyclic peptides, lipopeptides, lipoglycans or non-protein amino acids (Du *et al.*, 2019). Among the cyanotoxins is microcystin (MC), a cyclic peptide hepatotoxin that acts on hepatocytes of mammals and can cause liver damage and induce tumour cell production (Carmichael, 1992). Microcystin has also been the most common cyanotoxin observed in water sources, particularly MC-LR, one of the more toxic variants (Chorus and Bartram, 1999). The production of microcystin has been observed in several genera (Table 1.1) and MC and toxigenic cyanobacteria are spread globally in a wide distribution of habitats (Cotruvo, 2017). For this reason, the World Health Organization (WHO) has labelled MC as a public health concern (Cotruvo, 2017) and it is regulated in many distributed drinking water supplies globally, typically with an allowable limit of 1 µg/L (Cotruvo, 2017).

Table 1.1 Genera of cyanobacteria observed to produce microcystin.

Toxin	Genera
Microcystin	*Anabaena, Anabaenopsis, Annamia, Aphanocapsa, Arthrospira, Calothrix, Dolichospermum, Fischerella, Haphalosiphon, Leptolyngbya, Merismopedia, Microcystis, Nostoc, Oscillatoria, Phormidium, Planktothrix, Plectonema, Pseudanabaena, Radiocystis, Synechococcus*

Table adapted from Chorus and Welker (2021). Microcystin production in these cyanobacteria were verified in cultured strains by NMR, mass spectrometry, HPLC-PDA, ELISA, toxicity testing or molceular detection of *mcy* genes (Chorus and Welker, 2021).

1.3.4 Taste and Odour Compounds

In addition to certain genera of cyanobacteria being capable of decreasing drinking water quality through the production of toxins, some genera can produce the taste and odour compounds geosmin and methylisoborneol (MIB) (Giglio *et al.*, 2008; 2010). While non-toxic, these compounds contain earthy/muddy odours which are unpleasant to consume, thus reducing the quality of drinking water and public perception of the treatment of their water (Giglio *et al.*, 2008; 2010; John *et al*., 2018). Geosmin is a common taste and odour compound detected in water sources and is produced by several microorganisms including species of *Streptomyces* within the phylum Actinobacteria (Giglio *et al.*, 2008). The production of geosmin by cyanobacteria is more recently characterized and has been found in several genera (Table 1.2), primarily from the orders Nostocales, Oscillatoriales and Synechococcales (Izaguirre and Taylor, 2004; Wang *et al*., 2019; Churro *et al*., 2020).

Table 1.2 Genera of cyanobacteria observed to produce geosmin.

Taste/Odour Compound	Genera
Geosmin	*Anabaena*, *Aphanizomenon*, *Dolichospermum*, *Lyngbya*, *Leptolyngbya*, *Microcoleus*, *Nostoc*, *Oscillatoria*, *Phormidium*, *Planktothrix*, *Pseudanabaena*, *Synechococcus*, *Tychonema*

Data obtained from Izaguirre and Taylor (2004), Wang *et al.* (2019) and Churro *et al*. (2020).

1.4 Community Analysis of Bacteria

It is important to analyze the microbial communities present within environmental samples to identify bacteria that can cause water quality issues such as cyanobacteria. This can be achieved by molecular tools and methods to characterize microbial communities without the need for cultivation and observation (Hug *et al*., 2016). Rather, these molecular tools utilize high-throughput sequencing of molecular markers such as the 16S rRNA gene for phylogenetic analysis (Zhou *et al*., 2015). These methods are widely utilized as they can identify a range of taxa that are present within largely diverse environments (Zhou *et al*.,

2015). Analyzing microbial communities from environmental samples using molecular tools provides effective methods for identifying potential water quality reducing taxa, such as cyanobacteria within microbial community profiles (Hug *et al*., 2016).

1.4.1 16S rRNA Gene Sequencing

Community analysis of bacteria within environmental samples involves culture-independent methods including the use of molecular tools such as amplification and sequencing of the 16S rRNA gene (Figure 1.1) (Yarza *et al.*, 2014). The 16S rRNA gene is used as a phylogenetic marker to reveal bacterial identity as it is present in all bacteria and contains both highly conserved but also highly variable regions (Yang *et al.*, 2016). Using PCR primers that are specific to the V4 variable region of the 16S rRNA gene, these regions can be amplified (Walters *et al.*, 2015). Resulting amplified V4 regions of the 16S rRNA gene can then be sequenced using Next-Generation Sequencing (NGS) technology (e.g., Illumina MiSeq). Sequencing of the V4 region of the 16S rRNA gene yields sequences that can be used for downstream studies, such as using computational tools to assign taxonomy to the sequences, building microbial community profiles and generating phylogenetic trees to uncover evolutionary relationships (Zhou *et al*., 2015; Yang *et al.*, 2016). This makes the amplification and sequencing of the 16S rRNA gene essential for identifying taxa such as cyanobacteria from environmental samples.

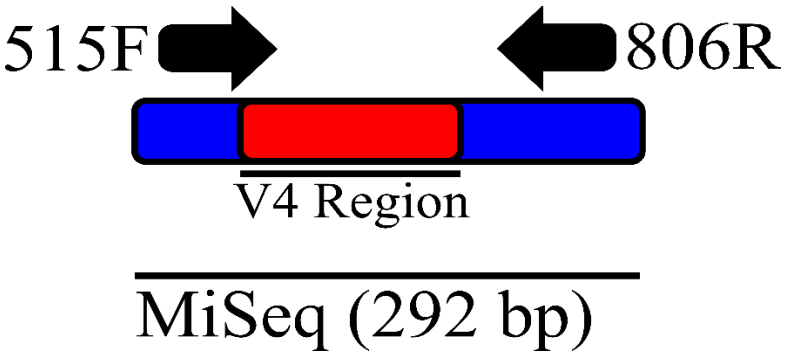


Figure 1.1 16S rRNA gene V4 region and primers used for sequencing. Imaged adapted from Shahi *et al.* (2017). Blue represents conserved regions and red is the variable (V) region 4. Arrows represent forward and reverse primers.

1.4.2 Toxin Marker Gene

Identifying cyanobacterial species capable of producing toxins such as microcystin is the one step in determining the potential presence of these toxins. It is also necessary to determine if the cyanobacteria contain the toxin genes. The synthesis of MC involves the *mcy* gene cluster which encodes for various enzymes (Figure 1.2), necessary for the formation of microcystin (Jun *et al*., 2018). To determine if cyanobacteria have the potential to synthesize microcystin, Jungblut and Neilan (2006) designed primers that were specific to an aminotransferase (AMT) region located between the non-ribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) domains of the *mcy*E gene for PCR amplification and isolation. The AMT region makes for an effective toxin marker as it is highly conserved within the *mcy*E gene, which is integral in the formation of microcystin (Jungblut and Neilan, 2006). If cyanobacteria can produce microcystin, they should contain the *mcy*E gene and therefore, the AMT region within this gene (Jungblut and Neilan, 2006). If the AMT region can be amplified and isolated, this can indicate that microcystin can potentially be produced, making this region an effective marker for microcystin (Jungblut and Neilan, 2006).

Figure 1.2 The *mcy* gene cluster in *Microcystis aeruginosa* PCC 7806. Image adapted from Pearson *et al*. (2016). Red indicates a polyketide synthase (PKS) and blue indicates a non-ribosomal peptide synthetase (NRPS) in which the aminotransferase (AMT) region falls in-between within the *mcy*E gene.

1.4.3 Taste and Odour Marker Gene

Despite a range of taxa that produce geosmin, there are no widely utilized universal primers to target all geosmin producers. This is owing to the number of cyanobacteria that can produce geosmin, and lack of current gene sequence data which makes identifying

conserved regions difficult (John *et al*., 2018). While difficult to create universal primers, studies have used primers to target the geosmin synthase gene, *geo*A, (Figure 1.3) to identify taste and odour producers within aquatic environments and drinking water sources. The production of geosmin involves *geo*A which was first identified within species of *Streptomyces*, including in the type species for geosmin production *Streptomyces coelicolor* A3 (Wang *et al*., 2019). Since this discovery, homologous genes sharing sequences similarities were identified in cyanobacteria (Giglio *et al.,* 2008). Although universal primers do not exist, the *geo*A gene has been widely utilized for the basis of producing custom primers to identify geosmin producing cyanobacteria (Giglio *et al.,* 2008; Suurnäkki *et al*., 2015; John *et al*., 2018).

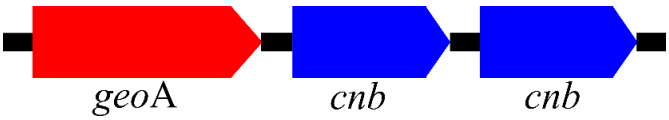


Figure 1.3 Geosmin synthetase gene cluster in *Anabeana ucrainica* CHAB 1432. Image adapted from Wang *et al.* (2015). The red arrow indicates the geosmin synthase (*geo*) gene and the blue arrows indicate cyclic nucleotide-binding protein (*cnb*) genes.

1.5 Objectives

The objectives of this study were to characterize the cyanobacterial communities present using 16S rRNA gene sequencing on water column samples collected from watersheds utilized as drinking water sources in the Atlantic and Pacific maritime ecozones. The resulting community profiles and phylogenetic characterization allows for identification of taxa potentially capable of producing microcystin, geosmin or forming blooms. This information is critical for understanding how future disturbances specific to each ecozone impacts watersheds and may affect the quality and treatability of drinking water. In addition, the characterization of cyanobacterial communities from the lakes in this study using any method (molecular, microscopy, etc.) has, to my knowledge, never been conducted. Hence, this study provides a novel and baseline understanding of the cyanobacterial communities

present within these source waters that could pose drinking water quality and treatment challenges if a proliferation event were to occur. This baseline data will allow us to observe how various disturbances, including climate impacts, specific to each ecozone impacts watersheds and how this may affect the quality and treatability of drinking water.

Chapter 2: Materials and Methods

2.1 Sample Collection

Collection of water samples from study lakes (Table 2.1) in forested watersheds in the Atlantic and Pacific maritime ecozones (Figure 2.1) was conducted during the summer and fall of 2019 by highly qualified personnel (HQP), who are part of the *for*Water NSERC Network. Sample site selection and sampling timeframes were determined by *for*Water NSERC Network for Forested Drinking Water Source Protection Technologies (i.e., "*for*Water Network") collaborators investigating hydrology and water quality at those watershed research observatories. One litre (1 L) water samples were collected from the surface and were vacuum filtered through 47 mm diameter, 1.2 µm pore size GF/C Whatman filters (Whatman plc, Buckinghamshire, United Kingdom) by HQP. According to those who filtered the water samples, there were some clogging of the filters primarily from organic and particulate matter. However, the full 1 L of water collected were still passed through the filters. Following vacuum filtration, the filters were then placed in petri dishes and kept cold and couriered to the Müller lab for cyanobacterial community analysis. Upon being received in coolers, samples were frozen at -20℃ until DNA could be extracted.

Table 2.1 Lakes in the Atlantic and Pacific maritime ecozones where water samples were collected.

Ecozone	Watershed	Sample Lakes	Max Depth (m)	Surface Area (hectares)	Water Source For	Population Water Supplies
Atlantic maritime	Pockwock Lake watershed	Pockwock Lake	47[A]	800.8[A]	Halifax Regional Municipality	~380,000[A]
		Island Lake	13.4	N/A	N/A	N/A
Pacific maritime	Comox Lake watershed	Comox Lake	109[B]	2100[C]	Courtenay and Comox	~38,000[C]
	Sooke River watershed	Deception Reservoir	6.5	59.5[D]	Greater Victoria[*]	~350,000[D*]
	Leech River watershed	Jarvis Lake	7	14.2[D]	N/A	N/A
		Weeks Lake	9.6	27.6[D]	N/A	N/A

Data obtained from (A) Tropea *et al.* (2007), (B) Epps and Phippen (2011), (C) Chandran and Mazumder (2015a) and (D) Barlak (2019). Remaining data obtained from HQP. It should be noted (*) that Deception Reservoir is not currrently a primary drinking water source but may be utilized as a future supplementary drinking water source.

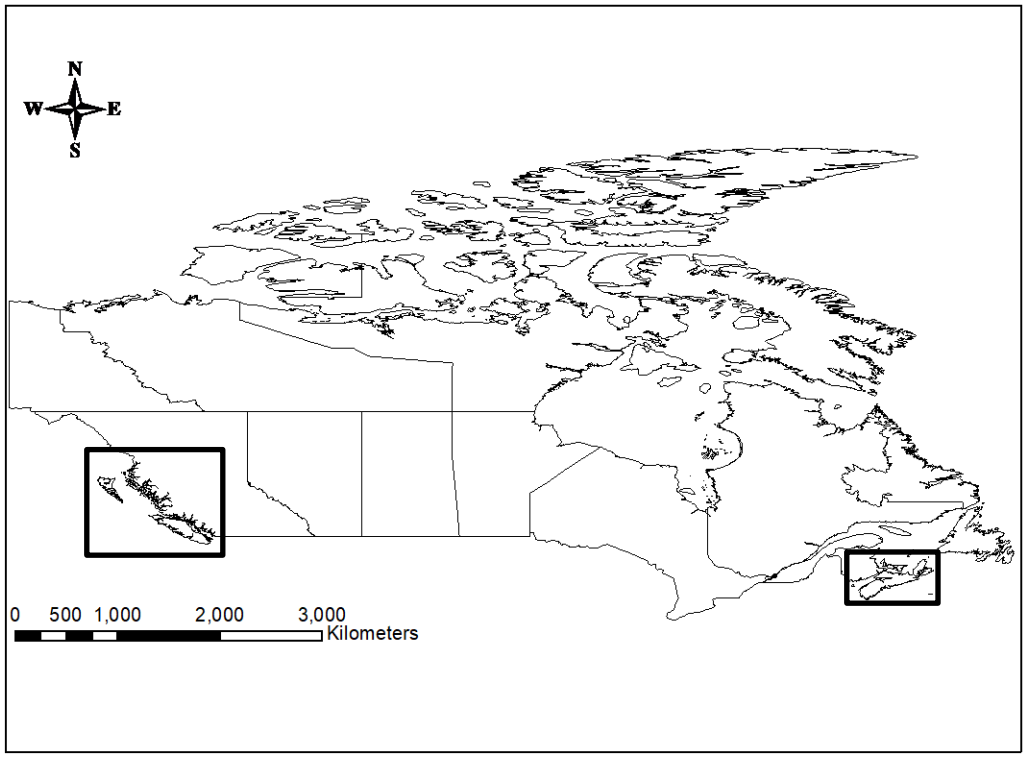


Figure 2.1 Map of Canada and the Atlantic and Pacific maritime ecozones. The Atlantic maritime ecozone is located on the east coast of Canada and the Pacific maritime ecozone is located on the west coast of Canada. Map was generated using ArcGIS (v. 10.5) (ESRI, 2011).

2.2 Sample Locations

2.2.1 Atlantic Maritime Ecozone

In the Atlantic maritime ecozone, water samples were collected from Pockwock Lake and Island Lake of the Pockwock Lake watershed, Nova Scotia, Canada (Figure 2.2). Samples were collected by members of the Jamieson Lab from Dalhousie University, Nova Scotia, Canada in collaboration with partners at Halifax Water (Table 2.2). Pockwock Lake was selected as a sample site as it is the primary drinking water source for the Halifax Regional Municipality (HRM) while Island Lake waters flow downstream into Pockwock Lake and is a potential origin point for cyanobacteria and associated taste and odour compounds (Figure 2.3).

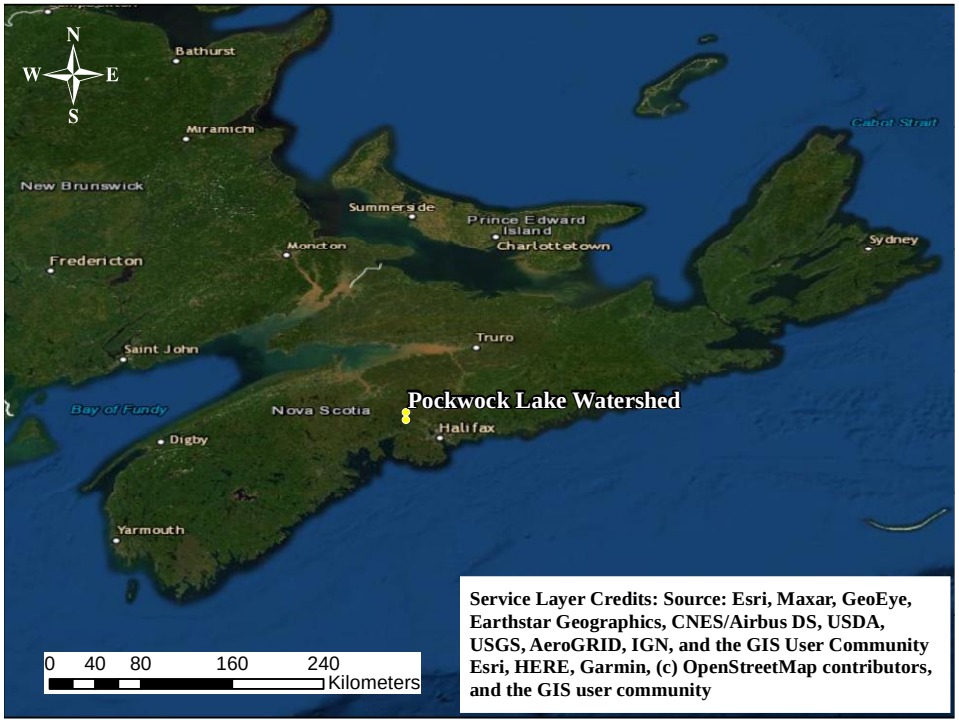


Figure 2.2 Location of the Pockwock Lake watershed in Nova Scotia, Canada within the Atlantic maritime ecozone. Map was generated using ArcGIS (v. 10.5) (ESRI, 2011).

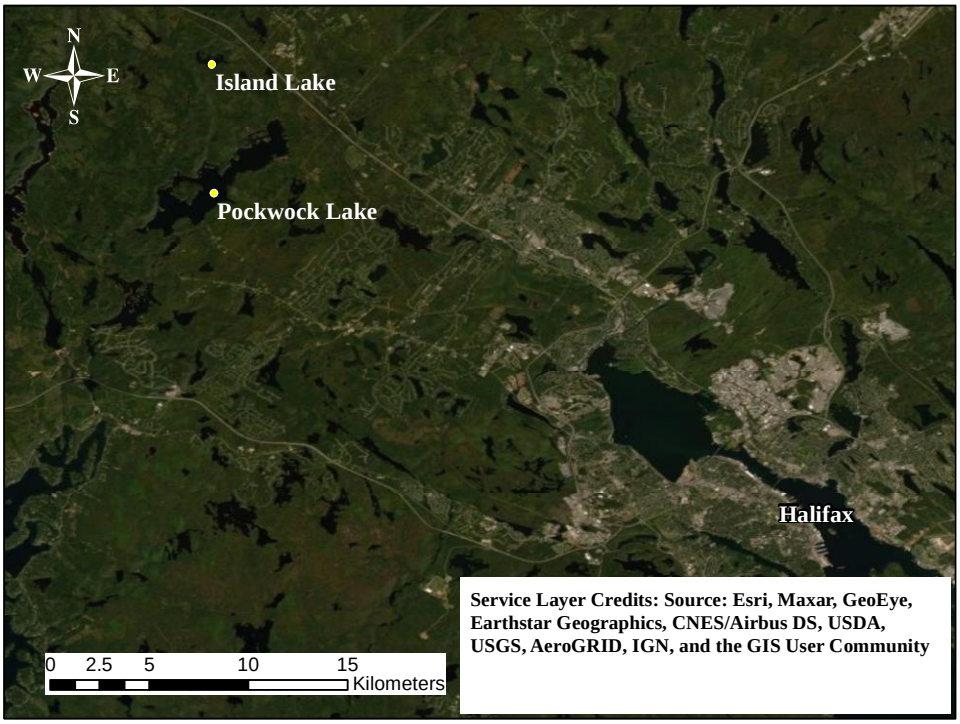


Figure 2.3 Pockwock Lake and Island Lake, Nova Scotia, Canada of the Pockwock Lake watershed located in the Atlantic maritime ecozone. The yellow marker indicates the location where water samples were collected. Map was generated using ArcGIS (v. 10.5) (ESRI, 2011).

Table 2.2 Location and collection date of water samples from the Pockwock Lake watershed for characterizing the cyanobacterial communities.

Sample Site	Collection Date (Month Day/Year)
Pockwock Lake	June 14/19 August 19/19 October 30/19
Island Lake	June 19/19 September 18/19 October 29/19

2.2.2 Pacific Maritime Ecozone

In the Pacific maritime ecozone (Figure 2.4), water samples were collected from Comox Lake (Figure 2.5) of the Comox Lake watershed located in Vancouver Island, British Columbia, Canada. Samples were collected by members of the Tank Lab at the University of Alberta, Canada in collaboration with partners at the Comox Valley Regional District (CVRD). Comox Lake was selected as a sample lake as it is the primary drinking water source for the municipalities of Courtenay and Comox. Water samples were collected at various sites within Comox Lake (Table 2.3). Boston Creek flows into Comox Lake and is located approximately at the mid-point. Cruikshank River and Upper Puntledge sample sites are the intake points which water flows into Comox Lake. Lake Outlet is the outflow point which water flows downstream from Comox Lake towards the water distribution plant.

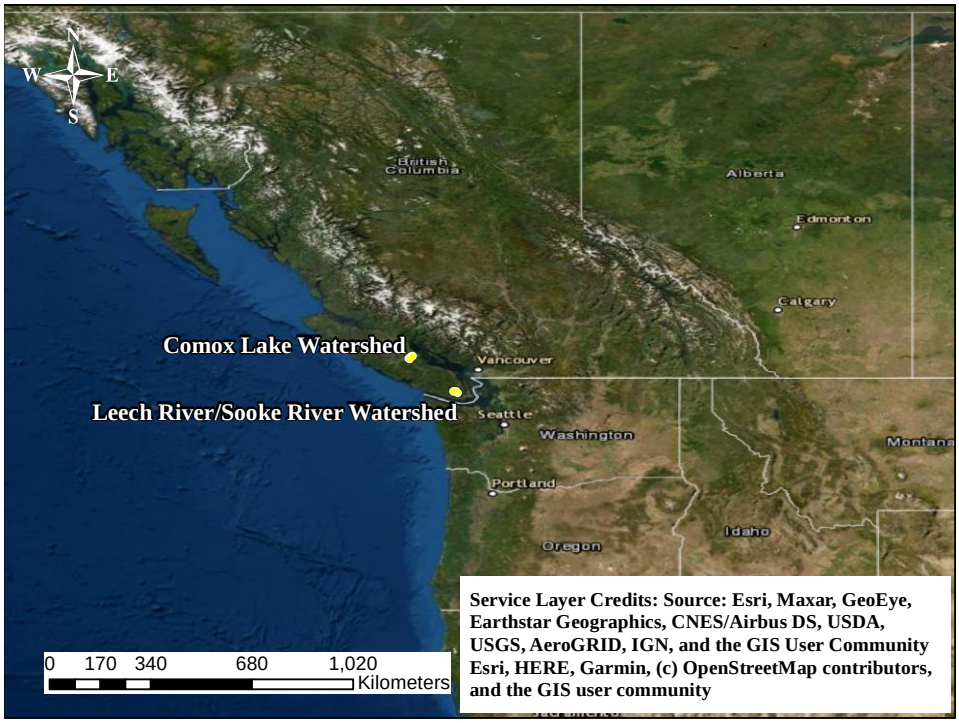


Figure 2.4 Location of Comox Lake and the Sooke River and Leech River watersheds within Vancouver Island, British Columbia, Canada of the Pacific maritime ecozone. Map was generated using ArcGIS (v. 10.5) (ESRI, 2011).

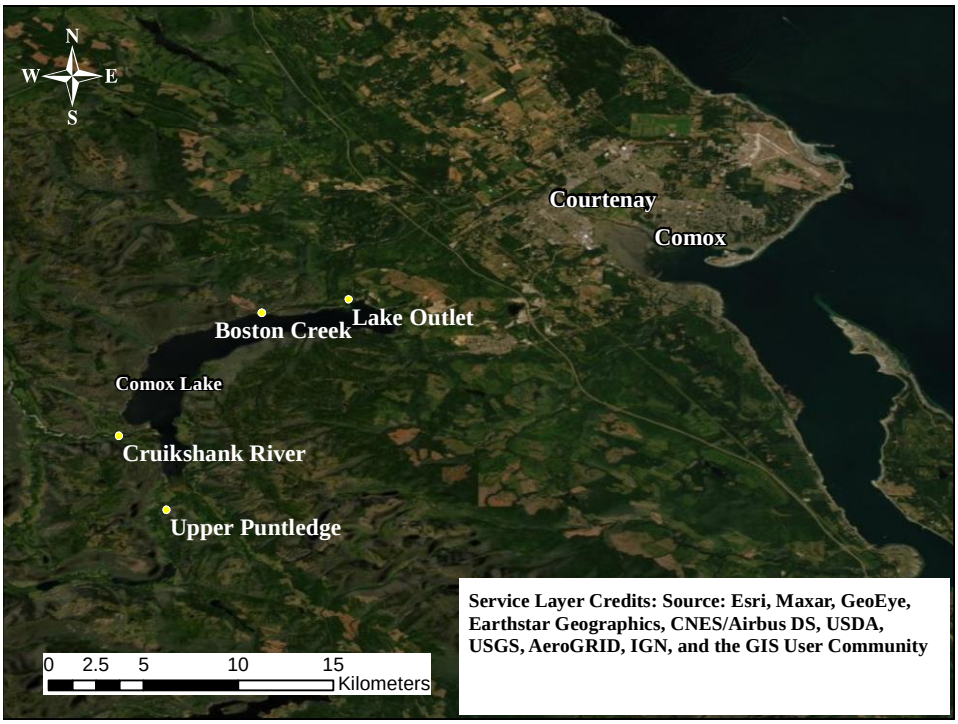


Figure 2.5 Comox Lake, British Columbia, Canada located in the Pacific maritime ecozone. The yellow markers indicate the locations where the water samples were collected. Map was generated using ArcGIS (v. 10.5) (ESRI, 2011).

Table 2.3 Location and collection date of water samples from the Comox Lake watershed for characterizing the cyanobacterial communities.

Sample Site	Collection Date (Month Day/Year)
Boston Creek	
Cruikshank River	May 29/19
Upper Puntledge	September 4/19
Lake Outlet	

From the Leech River and Sooke River watersheds (Figure 2.4) located in Vancouver Island, British Columbia, Canada, samples were collected from Jarvis Lake, Weeks Lake and Deception Reservoir (Figure 2.6) by members of the Johnson Lab at the University of British Columbia (UBC) in collaboration with the Capital Regional District (CRD) (Table 2.4). Jarvis Lake and Weeks Lake fall under the Leech River watershed and flow downstream into Deception Reservoir of the Sooke River watershed and acts as a holding basin. These lakes were selected as sample sites as Jarvis Lake and Weeks Lake flow into the Leech River which flows into Deception Reservoir which will potentially be used as a future secondary drinking water source for the Greater Victoria Region.

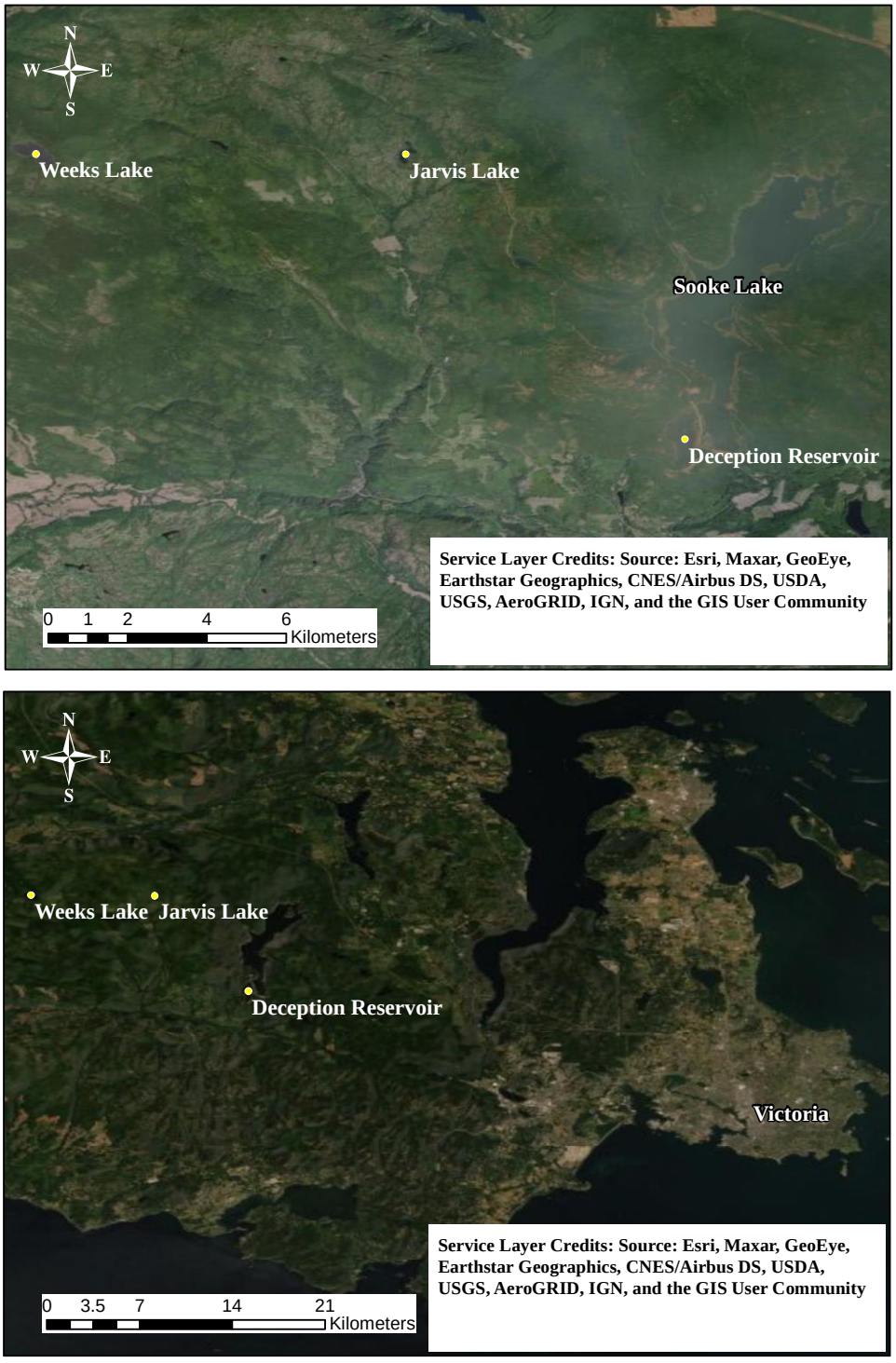


Figure 2.6 Location of Deception Reservoir, Jarvis Lake and Weeks Lake within the Sooke River watershed and Leech River watershed in southeastern Vancouver Island. Maps were generated using ArcGIS (v. 10.5) (ESRI, 2011).

Table 2.4 Location and collection date of water samples were collected from the Leech River and Sooke River watershed for characterizing the cyanobacterial communities.

Sample Site	Collection Date (Month Day/Year)
Jarvis Lake	July 25/19 August 29/19
Weeks Lake	July 26/19 August 29/19
Deception Reservoir	July 25/19 August 8/19

2.3 DNA Isolation and Amplicon Sequencing

DNA isolation from the filters (Figure 2.7) was performed using the Qiagen DNeasy PowerSoil Kit (QIAGEN Inc., Venlo, Netherlands) as per the kit protocols with the following deviation: extraction of microbial DNA was from GF/C Whatman filters instead of soil. To confirm successful extraction of DNA from filters, quantification and purity of extracted DNA was measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific, DE, USA). Isolated DNA extracts were then kept at -20°C until ready for sequencing.

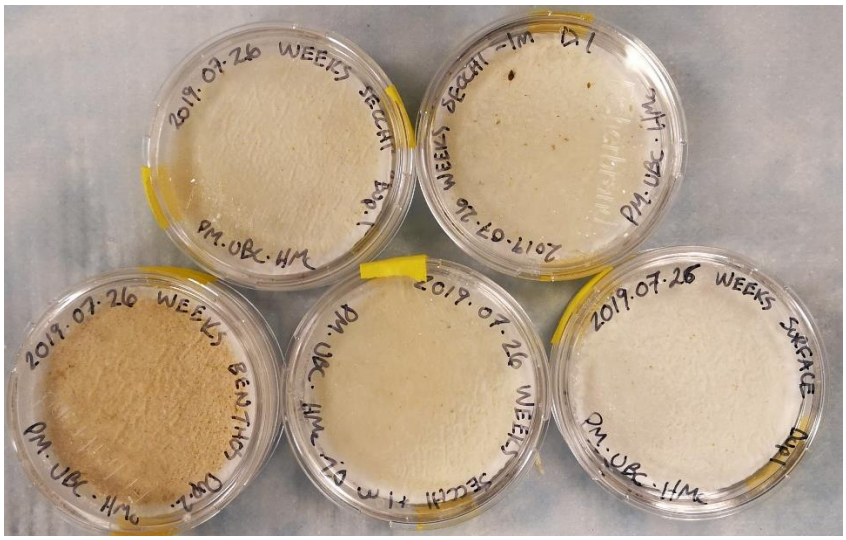


Figure 2.7 GF/C Whatman filters used for cyanobacteria sample collection.

Following bacterial DNA extraction, 20 µL of the DNA extracts were submitted to the commercial laboratory Metagenom Bio Inc. (Waterloo, Ontario, Canada) for 16S rRNA gene amplicon sequencing. Amplification and sequencing protocols were determined by Metagenom Bio Inc. (https://metagenom.com/) and follow those as outlined in previous publications. Primers used to capture the V4 region of the 16S rRNA gene were 515F 5'-GTGYCAGCMGCCGCGGTAA-3' (Parada *et al.*, 2016) and 806R 5'-GGACTACNVGGGTWTCTAAT-3' (Apprill *et al*., 2015). Resulting PCR products were sequenced using an Illumina MiSeq and the MiSeq Reagent Kit v2 (Illumina, Inc., Sand Diego, CA, USA) for 2 sets of 250 cycles.

2.4 Sequence Analysis

Resulting paired-end demultiplexed forward and reverse sequence reads were acquired from Metagenom Bio Inc. and analyzed using QIIME2 (v. 2018.8) (Bolyen *et al.,* 2019). Paired-end reads were imported into QIIME2 and the DADA2 pipeline (Callahan *et al.,* 2016) was used for trimming primers and truncating sequences to 250 base pairs to filter low quality reads. Sequence quality control using DADA2 dereplicated and denoised reads to remove any Illumina sequencing errors. Following quality control, paired-end reads were merged to generate an Amplicon Sequence Variant (ASV) table. These ASVs were assigned taxonomy using a Bayes classifier pre-trained on the SILVA databases (v. 132) (Quast *et al.,* 2013). Further manual filtering of ASVs included removing those that were assigned taxonomic classification to mitochondria and chloroplasts. Resulting ASVs were used in subsequent analyses.

2.5 Community Analyses

Files generated using QIIME2 were imported into R (v. 4.0.2) (R Core Team, 2020) and the package *qiime2R* (v. 0.99.23) (Bisanz, 2018) was used to create *phyloseq* (v. 1.32.0) (McMurdie and Holmes, 2013) objects from the QIIME2 files. Using the R package *mirlyn* (Multiple Iterations of Rarefaction for Library Normalization) (Cameron and Tremblay, 2020), relative abundance of ASVs within each filtered water sample were observed by

generating taxonomic barplots. These plots were generated by averaging the frequency of an ASV compared to the total number of ASVs present in each sample.

For diversity analyses, a rarified depth was determined using the R package *mirlyn* as a method of normalizing ASV libraries. Rarefaction curves were first generated to determine an appropriate normalized library size for analyzing both the whole bacterial communities (including cyanobacteria) and those just assigned to cyanobacteria. Based on rarefaction curves, for whole bacterial communities, samples were rarified to 5000 sequences and for cyanobacteria communities, 860 sequences. Rarefied depths were determined to retain as many samples as possible and omitting those with low sequence reads. Following library normalization, alpha diversity metrics (Shannon diversity index) were generated to observe diversity indices within samples. To determine variations in community composition among samples, beta diversity metrics (Bray-Curtis distances) were generated using Hellinger transformations and visualized using Principal Component Analysis (PCA) plots. For both alpha and beta diversity analyses, 100 iterations of rarefying library sizes were implemented to account for both potential loss in community diversity and for artificial variation that can be introduced through subsampling (Cameron *et al*., 2020).

2.6 Phylogenetic Analysis of Cyanobacteria

The 16S rRNA gene sequences assigned to cyanobacteria from sample lakes were compared to reference sequences from the National Center for Biotechnology Information (NCBI) nucleotide database at https://www.ncbi.nlm.nih.gov/ by using the Basic Local Alignment Search Tool (BLAST) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). This allowed for observing sequence similarities to representative and previously characterized cyanobacterial 16S rRNA gene sequences. For phylogenetic analysis, cyanobacterial reference sequences were obtained and selected from NCBI GenBank to be used in a phylogenetic tree based on being previously cited in literature, or by sharing the same taxonomic classification to ASVs based on assignment by the SILVA classifier down to the most resolved taxonomic level (genus or species).

A Multiple Sequence Alignment (MSA) was performed using the ClustalW algorithm in MEGA X (Kumar *et al*., 2018) to align reference sequences obtained from NCBI and 16S rRNA gene sequences assigned to cyanobacteria from samples. To improve taxonomic resolution, reference sequences were then trimmed to match the length of 16S rRNA gene sequences from sample lakes to only include the V4 region. Another MSA was performed to re-align sequences to improve alignment accuracy. Following the MSA, a phylogenetic tree was constructed using the Maximum Likelihood (ML) algorithm and a bootstrap value of 1000. The ML phylogenetic tree was constructed using the Kimura-2 parameter model (Kimura, 1980) with a discrete gamma distribution (G) and rate invariable site (I) model. The resulting phylogenetic tree inferred the evolutionary relatedness of cyanobacterial 16S rRNA gene sequences from water samples to previously characterized cyanobacteria based on sequence alignment similarity and clustering patterns. In a duplicate phylogenetic tree, cyanobacteria ASVs were coloured coded based on the watershed they were observed in to view biogeographic distributions.

As there were 26 ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1), which accounted for the majority (51%) of all cyanobacterial sequence reads, plus several variable branching patterns were observed in the phylogenetic tree, a subsequent tree of *Cyanobium* ASVs was constructed. Further reference sequences from this genus were included in this phylogenetic tree for taxonomic resolution and were obtained from Genuário *et al*. (2016). This phylogenetic tree was constructed using the same methods as described above with the purpose of further analyzing the taxonomic diversity within this genus.

The ASVs assigned to cyanobacteria by the SILVA classifier were searched against NCBI using BLAST for taxonomic comparisons at the genus and species level. The BLAST algorithm was set to exclude uncultured/environmental sequences for improved taxonomic resolution. The top match(es) to those in NCBI were based on query cover, E-value and percent (%) similarity. Taxonomic similarities between assignment by SILVA and BLAST were then identified to observe those that matched to at least the genus-level. Identifying the cyanobacteria ASVs that contained taxonomic matches/mismatches to those in NCBI allowed for potentially explaining variations observed within the phylogenetic tree.

2.7 Geosmin Primer Design

While no universal primers exist for the *geo*A gene, previous studies have used various primer sets to detect the *geo*A gene in cyanobacteria and *Streptomyces* (Giglio *et al*., 2008, Auffret *et al*., 2011, Suurnäkki *et al*., 2015; John *et al*., 2018). Using the gene regions these primer sets captured as templates, and additional reference sequences, new primers were designed by Metagenom Bio Inc. for the purpose of detecting *geo*A in both cyanobacteria and *Streptomyces* as these bacteria are well characterized geosmin producers.

For primer design, the *geo*A sequences from the model organisms *Nostoc punctiforme* PCC 73102 strain ATCC 29133 (CP001037.1) and *Streptomyces coelicolor* A3(2) (CP042324.1) (Giglio *et al*., 2008) were first obtained from NCBI. Then, the presence of *geo*A in these organisms was confirmed *in silico* using the primer set 250F/971R (Giglio *et al*., 2008) in Primer-BLAST (Ye *et al*., 2012). *Nostoc punctiforme* PCC 73102 strain ATCC 29133 (CP001037.1) and *Streptomyces coelicolor* A3(2) (CP042324.1) were then used as BLAST inputs to obtain further reference *geo*A sequences. There were 133 reference sequences including those from both cyanobacteria and *Streptomyces* that were obtained from NCBI and aligned by an MSA using the ClustalW algorithm in MEGA X (Supplementary Table 1). These aligned *geo*A sequences were sent to Metagenom Bio Inc. to be used as templates for novel *geo*A primer design.

Using the *geo*A reference sequences provided, Metagenom Bio Inc. visualized aligned sequences using MUSCLE (Edgar, 2004) and primers were constructed based on the resulting alignment using PrimerProspector (Walters *et al*., 2011) and openPrimeR (Kreer *et al*., 2020). The focus of primer design was to obtain amplicons that were appropriate length for Illumina MiSeq 250 x 2 and provided effective taxonomic resolution. From this, the primer pair geoA-297f (5'-RTCGAGTACATCGAGATGCG-3') and geoA-552r (5'-CGBGAGGTGAGGAYGTCGTT-3') were constructed by Metagenom Bio Inc.

2.8 Microcystin and Geosmin Gene Amplicon Sequencing

To validate the ability of primers (Table 2.5) to capture the *mcy*E and *geo*A genes, 20 µL of the DNA extract from the Weeks Lake sample collected in August was submitted to

the commercial laboratory Metagenom Bio Inc. This sample was used to validate the primers for isolation and amplification of the *mcy*E and *geo*A genes as it contained four ASVs assigned to the well characterized microcystin producing genus *Microcystis* with a frequency of 1,444 sequence reads as well as 12,959 sequence reads from the Actinobacteria.

Amplification and sequencing protocols for capturing the *mcy*E and *geo*A genes were determined by Metagenom Bio Inc. Primers used to capture these genes were the primer pairs HEPF 5'-TTTGGGGTTAACTTTTTTGGGCATAGTC-3'/HEPR 5'-AATTCTTGAGGCTGTAAATCGGGTTT-3' for the *mcy*E gene (Jungblut and Neilan, 2006) and geoA-297f 5'-RTCGAGTACATCGAGATGCG-3'/geoA-552r 5'-CGBGAGGTGAGGAYGTCGTT-3' for the *geo*A gene. Resulting PCR products were sequenced using an Illumina MiSeq and the MiSeq Reagent Kit v2 (Illumina, Inc., Sand Diego, CA, USA) for 2 sets of 250 cycles.

2.8.1 Microcystin and Geosmin Sequence Analysis

Once sequences of the AMT region of the *mcy*E gene were received from Metagenom Bio Inc., sequences were aligned against reference sequences from *Nodularia spumigena* strain NSOR10 (AY210783.2), *Anabaena* sp. 90 (AY212249.1) and *Microcystis aeruginosa* PCC 7806 (AF183408.1) which were used to create the HEPF/HEPR primers (Jungblut and Neilan, 2006), plus additional sequences under Genbank accession numbers AY817157-AY817171. Similarly, once *geo*A sequences were received from Metagenom Bio Inc., sequences were aligned against reference sequences that were used for primer design. An MSA of these sequences against their respective reference sequences were performed using the ClustalW algorithm in MEGA X.

The purpose of aligning these target genes to reference sequences was to observe regions of conservation and variability and to construct a subsequent phylogenetic tree. Sequences were also to be searched against the NCBI database using BLAST to observe sequence similarities against previously characterized *mcy*E and *geo*A genes. However, the sequences obtained were likely artefacts as they did not align with well characterized reference sequences for either gene, therefore they were not used for subsequent analyses. No

further samples were submitted for *mcy*E or *geo*A gene amplicon sequencing. This was partly due to time constraints, plus the need to further validate the ability of these primers to capture the target genes, particularly for the geoA-297f/geoA-552r primers as they are novel and were designed in this study.

Table 2.5 PCR primers utilized for capturing target marker genes from extracted microbial DNA from water samples.

Target Gene	Primers	Est. Product Size	Reference
16S rRNA (V4 region)	515F: 5'-GTGYCAGCMGCCGCGGTAA-3' 806R: 5'-GGACTACNVGGGTWTCTAAT-3'	292	Parada *et al.*, 2016 Apprill *et al*., 2015
*mcy*E (AMT region)	HEPF: 5'-TTTGGGGTTAACTTTTTTGGGCATAGTC-3' HEPR: 5'-AATTCTTGAGGCTGTAAATCGGGTTT-3'	472	Jungblut and Neilan, 2006
*geo*A	geoA-297f: 5'-RTCGAGTACATCGAGATGCG-3' geoA-552r: 5'-CGBGAGGTGAGGAYGTCGTT-3'	249	This study

Chapter 3: Results

3.1 Microbial Community Composition

Composition of microbial community profiles were analyzed from resulting taxonomic assignments of the V4 region of the 16S rRNA gene sequences. Across all watersheds and a total of 20 water samples, there were amplicon sequence variants (ASVs) from 44 bacterial and four archaeal phyla observed with two additional phyla labelled as ambiguous and unclassified. Prior to filtering out chloroplast and mitochondria hits, there were 4,829 unique ASVs observed across all samples with a total frequency of 347,982 reads. After filtering out those assigned to chloroplast and mitochondria, there were 4,301 unique ASVs with a frequency of 314,499 reads. The phyla that composed 95% of reads observed across all of the sample sites were *Proteobacteria*, *Actinobacteria*, *Cyanobacteria*, *Planctomycetes*, *Bacteriodetes* and *Verrucomicrobia*, in descending order of most reads (Table 3.1). While reads from these phyla were observed in every sample between watersheds and largely comprised the majority of the community composition within each sample, there were monthly variations observed in the abundance of reads, including those assigned to cyanobacteria.

Table 3.1 Relative abundance of the dominant bacteria phyla collectively observed among sample lakes.

Phylum	Relative Abundance (%)
Proteobacteria	42.4
Actinobacteria	17.8
Cyanobacteria	14.5
Planctomycetes	7.8
Bacteroidetes	6.6
Verrucomicrobia	6.1
Other	4.8

Other refers to all of the phyla that composed the remaining 5% of reads among samples.

3.2 Cyanobacteria Community Composition

From the dataset, 113 cyanobacterial ASVs with a frequency of 45,751 reads were observed. Of the 113 ASVs assigned to cyanobacteria, 79 were observed to be resolved to the genus or species-level while 34 were unresolved at the genus-level. From the 79 ASVs that were resolved to at least the genus-level, these fell under 23 known cyanobacteria genera based on taxonomic classification by SILVA. The cyanobacteria genera with the most reads collectively across all samples were assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1). For these genera, 26 ASVs were assigned to *Cyanobium* PCC-6307 (NR_102447.1) with 23,524 reads and for *Rhabdogloea smithii* SAG 47.91 (KM020002.1) there were nine ASVs with 19,018 reads. Together, these two genera comprised 93% of cyanobacteria reads observed across sample sites with *Cyanobium* PCC-6307 (NR_102447.1) contributing 51% of these reads and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) with 42%. The other 44 ASVs were assigned to cyanobacteria that comprised of the remaining 7% of reads observed.

3.2.1 Pockwock Lake Watershed Cyanobacteria Reads

Cyanobacteria reads in June samples from Island Lake and Pockwock Lake were in low abundance compared to other months and were grouped in phyla observed to be less than 1% abundant (Figure 3.1). While these samples contained a relatively low number of cyanobacteria reads in June, the number of cyanobacteria reads largely increased in the subsequent sampling month. Cyanobacteria reads comprised 12% of the microbial community in the September sample from Island Lake and 30% in the August sample from Pockwock Lake (Table 3.2). From the October samples, cyanobacteria reads comprised 5% of the microbial community in Island Lake and 11% in Pockwock Lake (Table 3.2). Although the number of cyanobacteria reads were generally higher from Pockwock Lake samples compared to Island Lake samples (excluding June), the relative abundance of cyanobacteria from these samples followed the same trend. It was observed from the Pockwock Lake watershed that the relative abundance of cyanobacteria largely increased from June to August/September and then decreased in October.

Table 3.2 Number of cyanobacteria reads, total number of reads and relative abundances of cyanobacteria from Pockwock Lake watershed samples across sampling months in 2019.

Watershed	Sample Site	Month	Cyanobacteria Reads	Total Reads	Relative Abundance (%)
Pockwock Lake Watershed	Island Lake	June	47	29058	<1
		September	2300	18619	12
		October	1096	23247	5
	Pockwock Lake	June	33	17983	<1
		August	6318	21211	30
		October	1901	17924	11

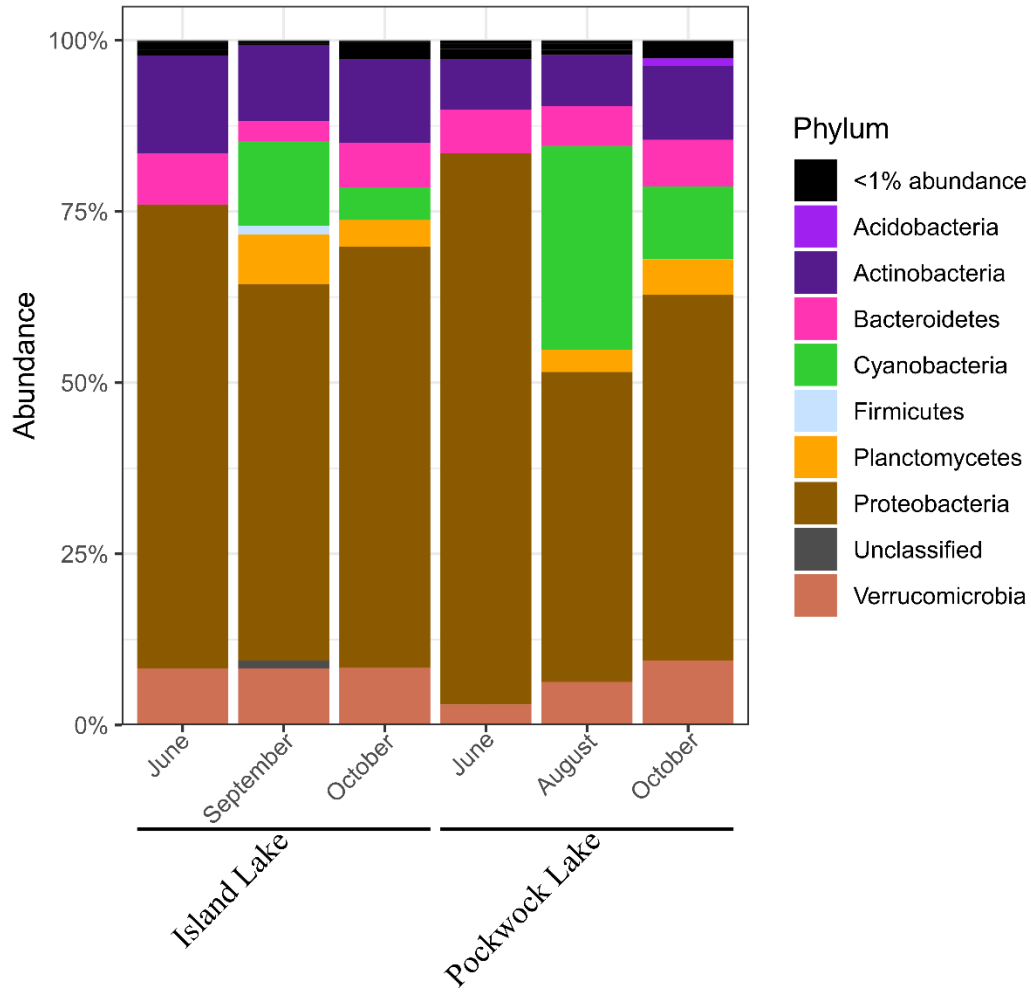


Figure 3.1 Relative abundance of microbial phyla from Pockwock Lake watershed samples across sampling months in 2019. Samples are grouped together based on sample site and ordered by sampling month. Cyanobacteria reads are present in June samples though in relatively low abundance and fall within phyla with <1% abundance.

3.2.2 Pockwock Lake Watershed Cyanobacteria Community

The cyanobacteria community composition in the Pockwock Lake watershed primarily comprised of reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) (Figure 3.2). These genera comprised the majority of cyanobacteria reads in each of the monthly Island Lake and Pockwock Lake samples. It should be noted that although *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads appear to dominate the June samples

from Island Lake and Pockwock Lake, these samples contained relatively low cyanobacteria reads and that cyanobacteria composed of <1% of the microbial community in these samples (Table 3.2). Overall, communities were largely dominated by reads of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) with relative abundance of *Cyanobium* PCC-6307 (NR_102447.1) reads reaching up to 70% in the October Island Lake sample and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) up to 66% in the August Pockwock Lake sample (Table 3.3). Apart from the September Island Lake sample, cyanobacteria genera with the highest proportion in samples was either clearly *Cyanobium* PCC-6307 (NR_102447.1) or *Rhabdogloea smithii* SAG 47.91 (KM020002.1) and that a co-dominance between these genera was not observed but rather shifted between them. Additionally, the Island Lake sample in June and the Pockwock Lake sample in October were the only ones from the Pockwock Lake watershed to have other genera present above a 1% relative abundance which belonged to ASVs assigned to *Aphanizomenon* NIES81 (AJ293131.1) (Figure 3.2).

Table 3.3 Relative abundance of sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) from the Pockwock Lake watershed samples across sampling months in 2019.

Watershed	**Sample Site**	**Month**	**Relative Abundance (%) of *Cyanobium* PCC-6307**	**Relative Abundance (%) of *Rhabdogloea smithii* SAG 47.91**	**Relative Abundance (%) of Other Genera**
Pockwock Lake Watershed	Island Lake	June	30	51	19
		September	47	52	<1
		October	70	29	1
	Pockwock Lake	June	67	12	21
		August	33	66	<1
		October	55	33	13

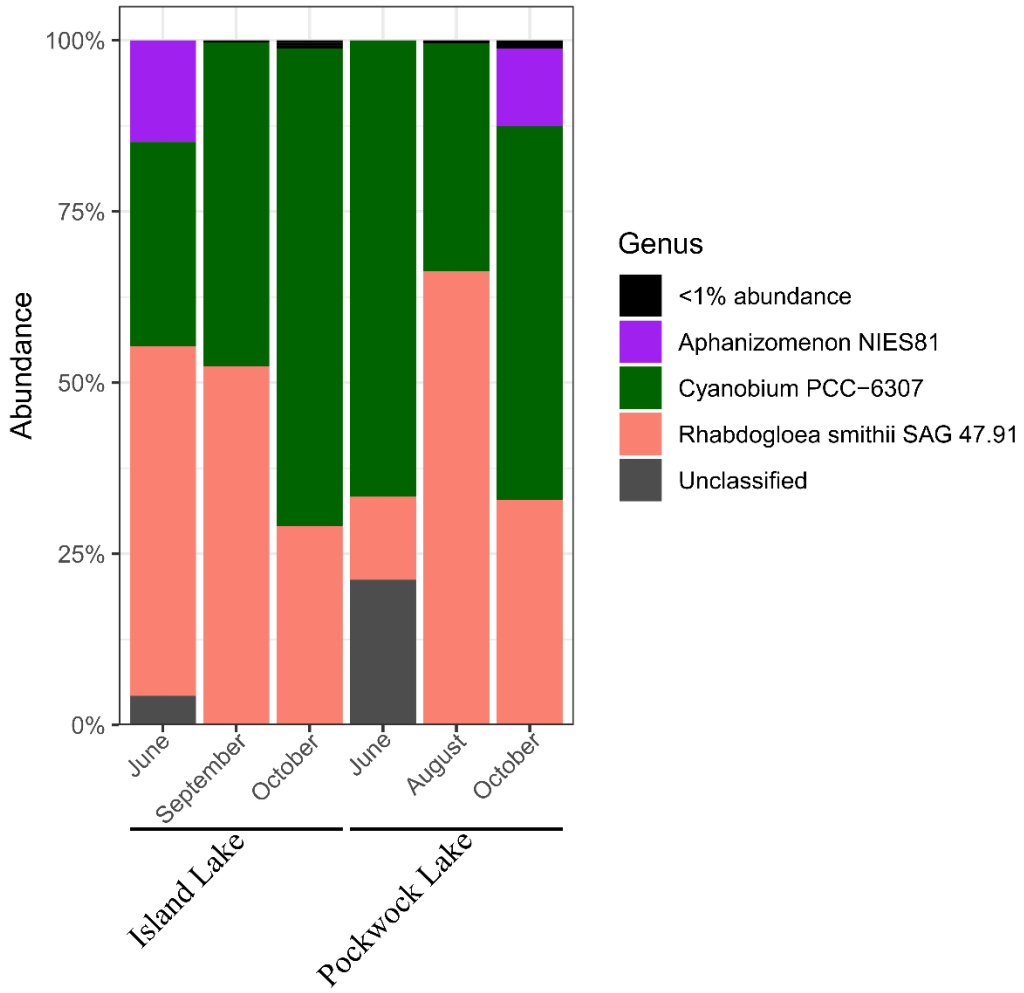


Figure 3.2 Relative abundance of cyanobacteria genera from Pockwock Lake watershed samples across sampling months in 2019. Samples are grouped together based on sample site and ordered by sampling month. Genera that are below 1% relative abundance within each sample are grouped together.

3.2.3 Comox Lake Watershed Cyanobacteria Reads

The Comox Lake watershed contained ASVs from a range of different phyla (Figure 3.3). For cyanobacteria, Boston Creek and Cruikshank River samples were observed to have low relative abundance of cyanobacterial reads while in Lake Outlet and Upper Puntledge samples they were notably higher (Figure 3.3). Cyanobacteria reads in the May and September samples from Boston Creek and Cruikshank River were relatively low and contributed a very small proportion of reads to the total microbial community (Table 3.4).

Comparatively, May and September samples from Lake Outlet and Upper Puntledge were observed to have the highest relative abundances of cyanobacteria reads in this watershed (Table 3.4). In May, cyanobacteria reads from Lake Outlet and Upper Puntledge comprised 42% and 20% of the total microbial communities from within these samples, respectively (Table 3.4). Similarly, in the September samples, cyanobacteria reads from Lake Outlet and Upper Puntledge comprised 12% and 8% of the total microbial communities, respectively (Table 3.4).

Table 3.4 Number of cyanobacteria reads, total number of reads and relative abundances of cyanobacteria from Comox Lake watershed samples across sampling months in 2019.

Watershed	Sample Site	Month	Cyanobacteria Reads	Total Reads	Relative Abundance (%)
Comox Lake Watershed	Boston Creek	May	82	6321	1
		September	179	6697	3
	Cruikshank River	May	83	5379	2
		September	97	6780	1
	Lake Outlet	May	2865	6803	42
		September	860	7193	12
	Upper Puntledge	May	2014	10023	20
		September	898	10924	8

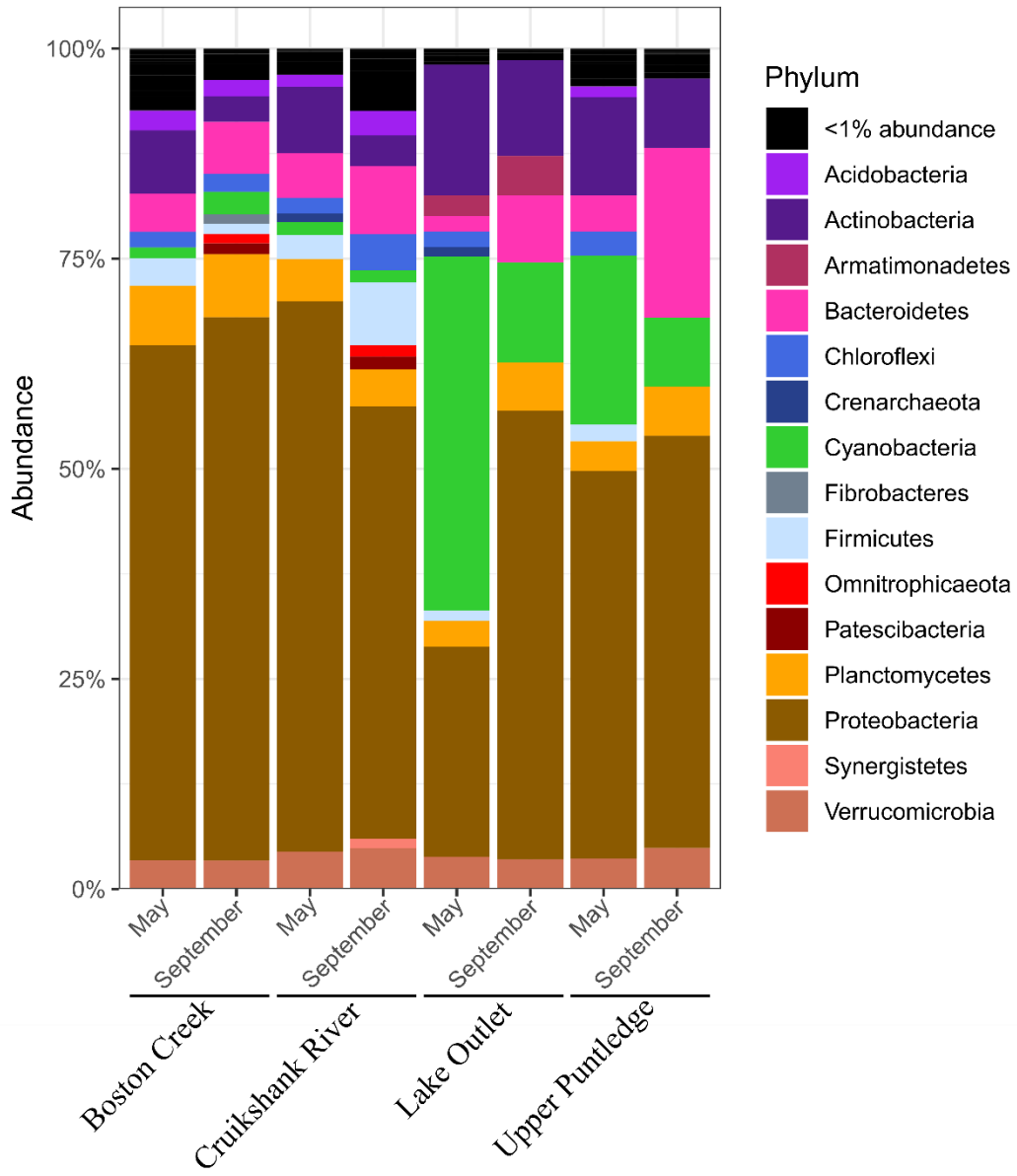


Figure 3.3 Relative abundance of microbial phyla from Comox Lake watershed samples across sampling months in 2019. Samples are grouped together based on sample site and ordered by sampling month. Phyla that are below 1% relative abundance within each sample are grouped together.

3.2.4 Comox Lake Watershed Cyanobacteria Community

Similar to the microbial community at the phylum-level, the Comox Lake watershed contained ASVs from a range of different cyanobacteria genera (Figure 3.4). Interestingly,

the Comox Lake watershed was the only watershed to not have cyanobacterial communities with an abundance of reads assigned to *Rhabdogloea smithii* SAG 47.91 (KM020002.1). Like the other watersheds, they contained a number of *Cyanobium* PCC-6307 (NR_102447.1) reads (Figure 3.4). The samples from Lake Outlet and Upper Puntledge in both May and September were dominated by a large proportion of *Cyanobium* PCC-6307 (NR_102447.1) reads with the relative abundance ranging from 94 - 99% of the cyanobacteria community (Table 3.5). The samples from Boston Creek in May and Cruikshank River in May and September, while containing some reads from *Cyanobium* PCC-6307, primarily contained reads from a more diverse range of cyanobacteria genera (Figure 3.4). When comparing all samples from the Comox Lake watershed, among samples with less reads assigned to *Cyanobium* PCC-6307 (NR_102447.1), the more diverse the cyanobacterial community (Figure 3.4).

Table 3.5 Relative abundance of sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) from the Comox Lake watershed samples across sampling months in 2019.

Watershed	**Sample Site**	**Month**	**Relative Abundance (%) of *Cyanobium* PCC-6307**	**Relative Abundance (%) of *Rhabdogloea smithii* SAG 47.91**	**Relative Abundance (%) of Other Genera**
Comox Lake Watershed	Boston Creek	May	0	0	100
		September	13	0	87
	Cruikshank River	May	39	0	61
		September	6	0	94
	Lake Outlet	May	99	<1	<1
		September	97	0	3
	Upper Puntledge	May	96	0	4
		September	94	0	6

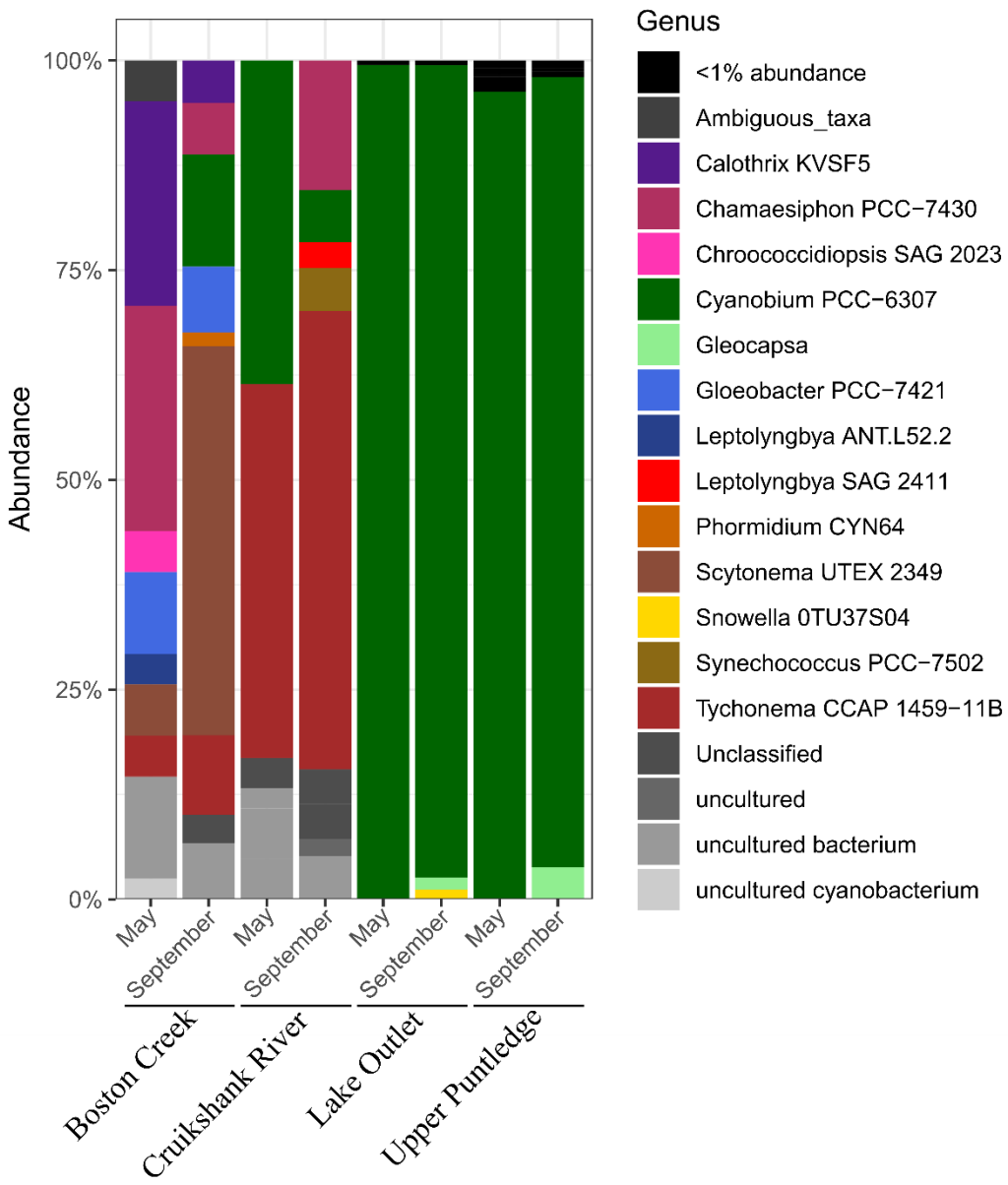


Figure 3.4 Relative abundance of cyanobacteria genera from Comox Lake watershed samples across sampling months in 2019. Samples are grouped together based on sample site and ordered by sampling month. Genera that are below 1% relative abundance within each sample are grouped together.

3.2.5 Leech River and Sooke River Watershed Cyanobacteria Reads

The relative abundance of reads assigned to cyanobacterial from the Leech River and Sooke River watersheds were similar among sample sites in both sampling months (Figure

3.5). In the July samples from Jarvis Lake, Weeks Lake and Deception Reservoir, the relative abundance of cyanobacteria reads was similar at 30%, 23% and 34%, respectively (Table 3.6). Some monthly variations in relative abundance from these sample sites were observed as the number of cyanobacteria reads decreased between July and August. However, in August, the relative abundance of cyanobacteria reads from Jarvis Lake, Weeks Lake, and Deception Reservoir again remained similar with relative abundances of cyanobacteria at 23%, 16% and 11%, respectively (Table 3.6).

Table 3.6 Number of cyanobacteria reads, total number of reads and relative abundances of cyanobacteria from Leech River/Sooke River watershed samples across sampling months in 2019.

Watershed	Sample Site	Month	Cyanobacteria Reads	Total Reads	Relative Abundance (%)
Leech River/ Sooke River Watersheds	Jarvis Lake	July	5762	19470	30
		August	4261	18533	23
	Weeks Lake	July	3588	15541	23
		August	5800	36349	16
	Deception Reservoir	July	5231	15329	34
		August	2336	21115	11

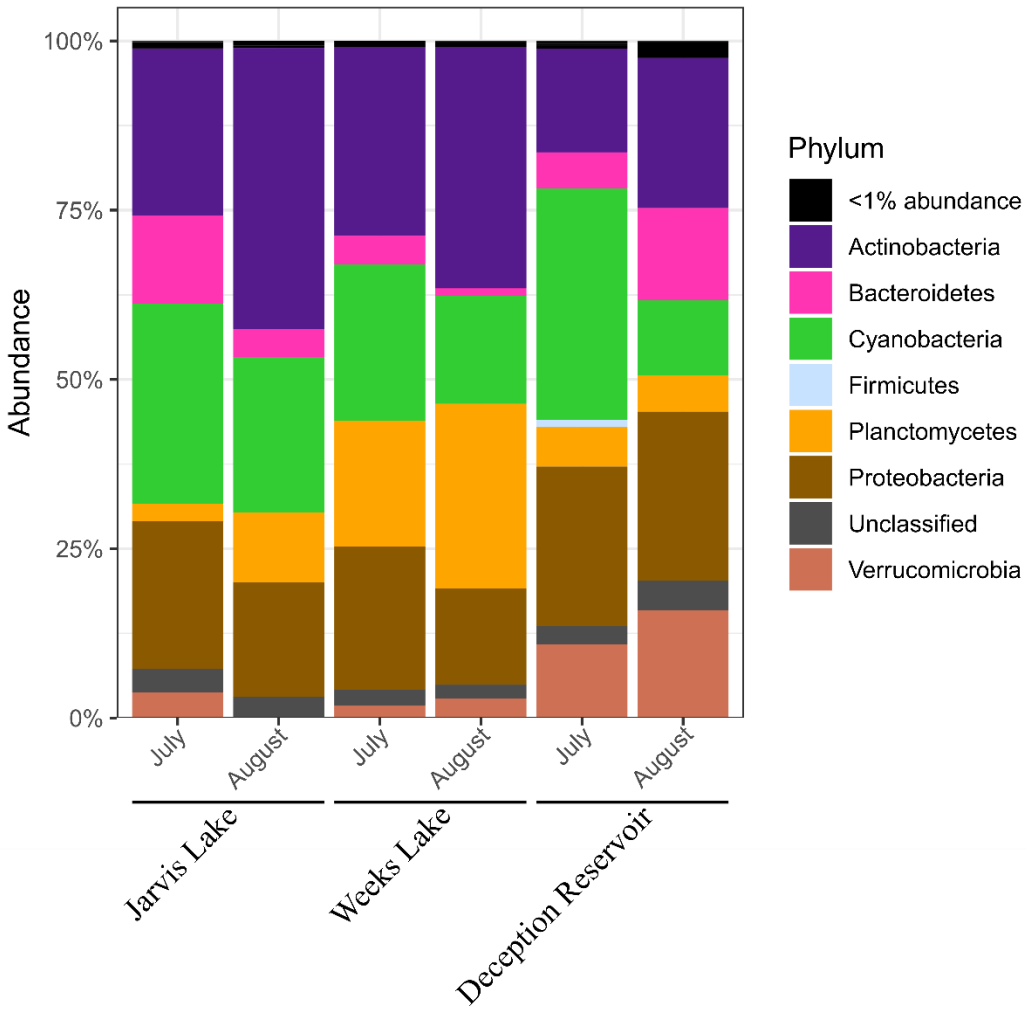


Figure 3.5 Relative abundance of microbial phyla from sample lakes from Leech River/Sooke River watershed samples across sampling months in 2019. Samples are grouped together based on sample site and ordered by sampling month. Phyla that are below 1% relative abundance within each sample are grouped together.

3.2.6 Leech River and Sooke River Watershed Cyanobacteria Community

The cyanobacteria community composition of the Leech River/Sooke River watersheds was predominantly comprised of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads (Figure 3.6). Among samples from these watersheds, the cyanobacteria genera with the highest relative abundance were either *Cyanobium* PCC-6307 (NR_102447.1) or *Rhabdogloea smithii* SAG 47.91, again indicating

a lack of co-dominance. In the Jarvis Lake samples, it was observed in both July and August that the cyanobacteria community was mainly comprised of *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads, with little monthly variation. The Weeks Lake samples, while also experiencing little monthly variation between July and August, was instead mainly comprised of *Cyanobium* PCC-6307 (NR_102447.1) reads. It was the Deception Reservoir samples that experienced a large shift in relative abundance, from primarily *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads in July to primarily *Cyanobium* PCC-6307 (NR_102447.1) reads in August. It was in these samples that relative abundance was the highest, with *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads up to 72% in July and with *Cyanobium* PCC-6307 (NR_102447.1) reads at 91% in August (Table 3.7). What was also observed were reads assigned to *Microcystis* PCC-7914 (no GenBank accession number) observed in Jarvis Lake and Weeks Lake samples in both July and August with the highest relative abundance in the August Weeks Lake sample (Figure 3.6).

Table 3.7 Relative abundance of sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) from the Leech River/Sooke River watershed samples across sampling months in 2019.

Watershed	**Sample Site**	**Month**	**Relative Abundance (%) of *Cyanobium* PCC-6307**	**Relative Abundance (%) of *Rhabdogloea smithii* SAG 47.91**	**Relative Abundance (%) of Other Genera**
Leech River/Sooke River Watershed	Jarvis Lake	July	28	70	2
		August	37	61	3
	Weeks Lake	July	53	35	11
		August	57	16	27
	Deception Reservoir	July	27	72	1
		August	91	3	6

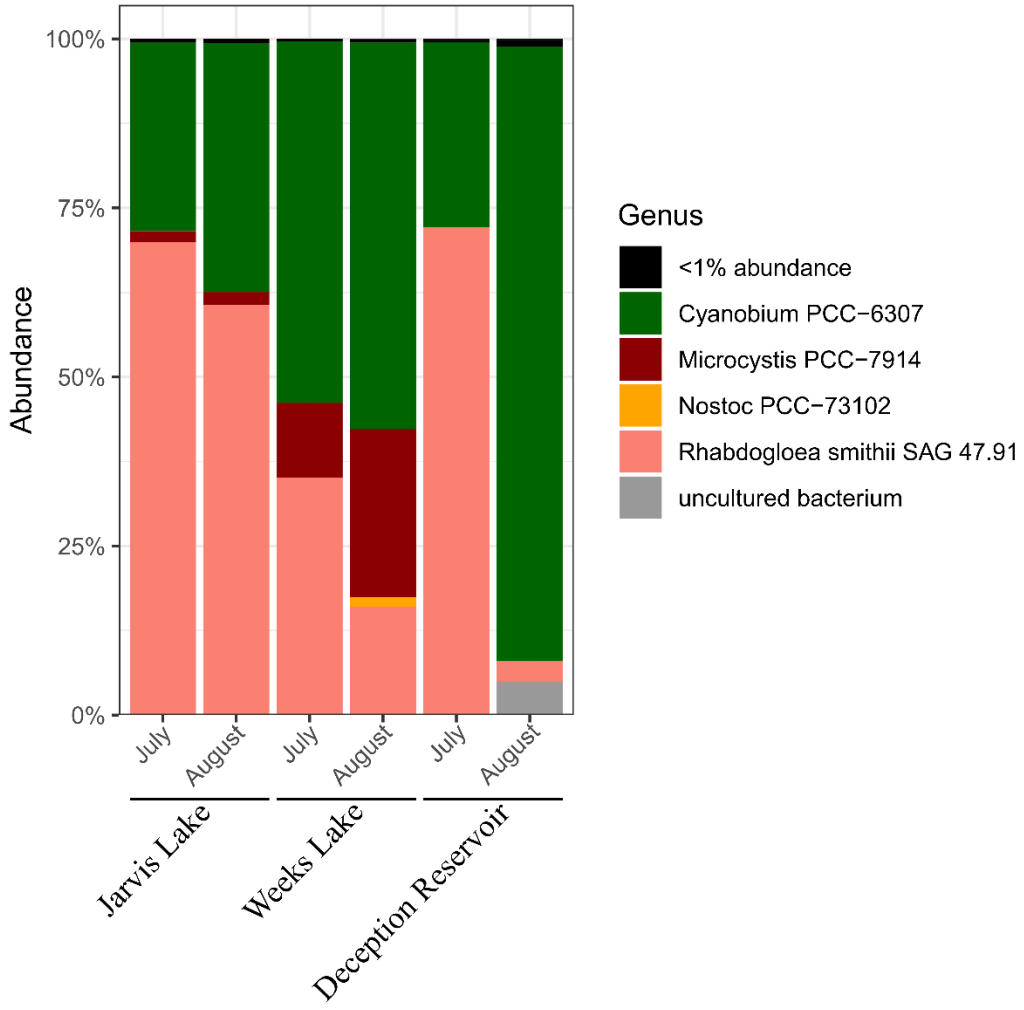


Figure 3.6 Relative abundance of cyanobacteria genera from Leech River/Sooke River watershed samples across sampling months in 2019. Samples are grouped together based on sample site and ordered by sampling month. Genera that are below 1% relative abundance within each sample are grouped together.

3.3 Comparisons of Cyanobacteria Among Watersheds

Among the watersheds, samples from the Leech River/Sooke River watersheds had the highest relative abundance of cyanobacterial reads (59%), followed by the Pockwock Lake watershed (26%) and then the Comox Lake watershed (15%) (Table 3.8). *Cyanobium* PCC-6307 (NR_102447.1) reads were present and consistently observed in every watershed.

In 19 of 20 samples, *Cyanobium* PCC-6307 (NR_102447.1) reads were present and comprised from as low as 6% of the cyanobacterial community up to 99%. *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads were present in 13 of 20 samples with the majority coming from the Pockwock Lake and Leech River/Sooke River watersheds. Only one sample from the Comox Lake watershed contained *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads but were less than 1% abundant among the cyanobacteria community. It was however the Comox Lake watershed that was observed to have a wide range of cyanobacteria ASVs compared to the other watersheds. In contrast, both the Pockwock Lake watershed and Leech River/Sooke River watersheds did not have a range of cyanobacteria ASVs and instead were rather saturated with both *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) ASVs. As previously stated, in samples that contained both *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) reads, the relative abundances of these genera were clearly dominated by one or the other, expressing a lack of co-dominance which may indicate potential competition among these organisms.

Table 3.8 Total number and the relative abundance of cyanobacteria reads per watershed.

Ecozone	Watershed	Cyanobacteria Reads	Relative Abundance (%)
Pacific	Leech River/Sooke River Watershed	26978	59
Atlantic	Pockwock Lake Watershed	11695	26
Pacific	Comox Lake Watershed	7078	15
	Total:	45751	100

3.4 Alpha Diversity of Communities

For alpha diversity analysis of the entire bacterial communities, no samples were excluded as all were above the rarefied library size (5000). Analysis of alpha diversity of cyanobacteria communities among sample sites excluded May and September samples from

Boston Creek and Cruikshank River, and June samples from Island Lake and Pockwock Lake as these samples contained a low number of cyanobacteria reads and therefore fell below the rarefied library size (860). Following rarefaction and normalization of library size using the R package *mirlyn*, among the watersheds, some monthly variations in whole bacterial community diversity indices were observed within samples (Figure 3.7) When comparing these indices to the alpha diversity of cyanobacterial communities, there were more monthly variations observed (Figure 3.8).

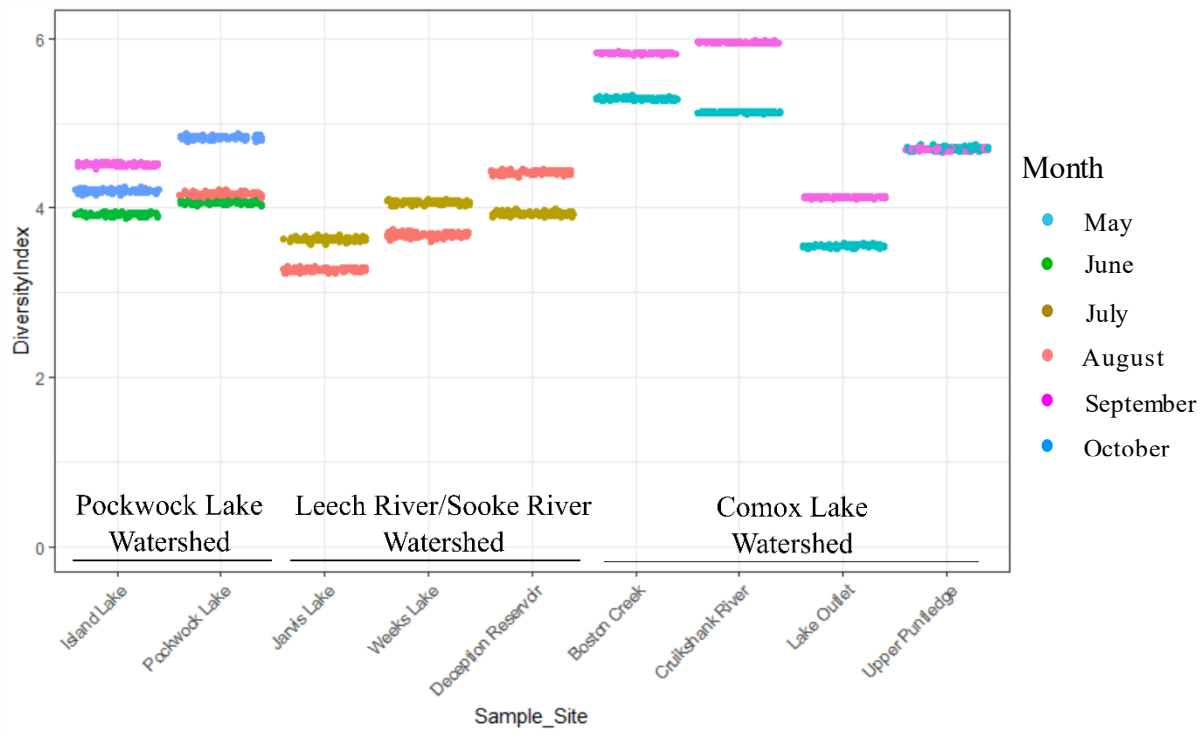


Figure 3.7 Alpha diversity of bacterial communities from sample sites. Diversity plot was constructed with a rarefied library size of 5000 sequences and replicated with 100 iterations generated using the Shannon diversity index in the R package *mirlyn*. No samples were excluded following library normalization.

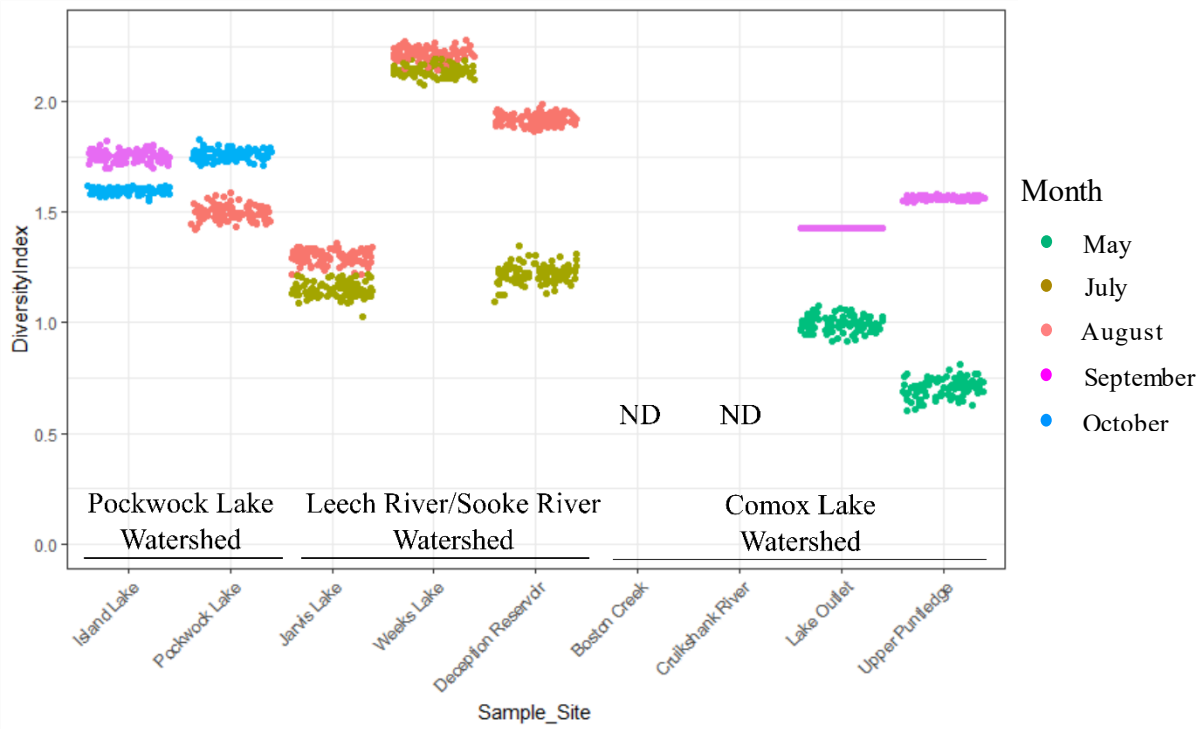


Figure 3.8 Alpha diversity of cyanobacterial communities from sample sites. Diversity plot was constructed with a rarefied library size of 860 sequences and replicated with 100 iterations generated using the Shannon diversity index in the R package *mirlyn*. Boston Creek and Cruikshank River are listed as ND (no data) as these samples were excluded following rarefaction as well as June samples from Island Lake and Pockwock Lake.

3.4.1 Pockwock Lake Watershed Community Diversity

Island Lake and Pockwock Lake samples had similar alpha diversity indices for the whole bacterial communities with little variation between the sampling months. Similarities in alpha diversity also were observed for the cyanobacterial communities, though excluding June samples due to low cyanobacteria sequences that therefore did not contribute much to the diversity of the entire bacterial communities this month. For the cyanobacterial communities, similarities in diversity indices could be contributed to these samples being saturated with reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogleoa smithii* SAG 47.91 (KM020002.1). In the September and October Island Lake samples, the total number of cyanobacteria ASVs was the same with *Cyanobium* PCC-6307

(NR_102447.1) and *Rhabdogleoa smithii* SAG 47.91 (KM020002.1) comprising most of these ASVs (Table 3.9). The August and October Pockwock Lake samples were similar in that the same total number of cyanobacteria ASVs were observed which were also primarily comprised of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogleoa smithii* SAG 47.91 (KM020002.1) ASVs.

Table 3.9 Number of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) ASVs compared to the total number of cyanobacteria ASVs from the Pockwock Lake watershed.

Watershed	Sample Site	Month	*Cyanobium* PCC-6307 ASVs	*Rhabdogloea smithii* SAG 47.91 ASVs	Total ASVs
Pockwock Lake Watershed	Island Lake	June	1	1	4
		September	6	3	10
		October	3	4	10
	Pockwock Lake	June	1	1	3
		August	4	6	12
		October	4	2	12

Total number of ASVs includes those unresolved to the genus-level.

3.4.2 Comox Lake Watershed Community Diversity

Alpha diversity indices of whole bacterial communities were highest in the Comox Lake watershed samples Boston Creek and Cruikshank River in May and September. However, for cyanobacterial communities, these samples were excluded due to low cyanobacteria sequences. The cyanobacterial composition therefore did not contribute much to the diversity of the entire bacterial communities within Boston Creek and Cruikshank River. Little monthly variations in bacterial communities from Lake Outlet and Upper Puntledge samples were observed but were more variable in cyanobacterial communities between sampling months. From these samples, the number of *Cyanobium* PCC-6307 (NR_102447.1) ASVs remained consistent but the increased diversity in the September

samples is likely due to the range of taxa other ASVs were assigned to including those that fell under 1% abundant (Table 3.10).

Table 3.10 Number of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) ASVs compared to the total number of cyanobacteria ASVs from the Comox Lake watershed.

Watershed	Sample Site	Month	*Cyanobium* PCC-6307 ASVs	*Rhabdogloea smithii* SAG 47.91 ASVs	Total ASVs
Comox Lake Watershed	Boston Creek	May	0	0	12
		September	2	0	12
	Cruikshank River	May	2	0	8
		September	1	0	9
	Lake Outlet	May	6	1	9
		September	6	0	9
	Upper Puntledge	May	6	0	20
		September	6	0	13

Total number of ASVs includes those unresolved to the genus-level.

3.4.3 Leech River/Sooke River Watershed Community Diversity

There were no samples from the Leech River/Sooke River watershed excluded from alpha diversity analyses as samples were consistently abundant with reads. Monthly variations and diversity indices of the bacterial communities were consistent from each sample site but were more variable for cyanobacteria. Jarvis Lake and Weeks Lake contained consistent cyanobacterial diversity indices between sampling months with Weeks Lake samples overall being more diverse. Cyanobacterial community diversity in Deception Reservoir, compared to the other samples, contained a much larger difference in diversity indices between sampling months. From Jarvis Lake, the number of ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1), *Rhabdogleoa smithii* SAG 47.91 (KM020002.1) and other genera was relatively consistent (Table 3.11). Comparatively, Weeks Lake and Deception Reservoir were similar with number of ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogleoa smithii* SAG 47.91 (KM020002.1) in both sampling

months. Therefore, variations in diversity could be attributed to the number of ASVs assigned to other taxa present in these samples.

Table 3.11 Number of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) ASVs compared to the total number of cyanobacteria ASVs from the Leech River/Sooke River watershed.

Watershed	Sample Site	Month	*Cyanobium* PCC-6307 ASVs	*Rhabdogloea smithii* SAG 47.91 ASVs	Total ASVs
Leech River/ Sooke River Watersheds	Jarvis Lake	July	3	3	9
		August	2	4	8
	Weeks Lake	July	8	3	16
		August	7	3	18
	Deception Reservoir	July	7	2	11
		August	8	1	15

Total number of ASVs includes those unresolved to the genus-level.

3.5 Beta Diversity of Communities

From beta diversity analysis, for both whole bacterial (Figure 3.9) and cyanobacterial community composition (Figure 3.10), samples generally grouped together based on the watershed they are present in as three groups were observed in the PCA plots. The only exception to this was Jarvis Lake in July for cyanobacterial diversity as this sample did not group well with any other samples, including those from the Leech River/Sooke River watersheds. While samples generally grouped together by watershed, from positions in the PCA plots, Comox Lake watershed and Leech River/Sooke River watershed samples were more similar based on bacterial community composition, and Pockwock Lake watershed and Leech River/Sooke River watershed samples were more similar based on cyanobacterial community composition. These observations indicate that, while samples from the same watershed were similar in composition, there are also some similarities in composition of communities between watersheds. For cyanobacteria, differences in composition between

watersheds can attributed to the various taxa ASVs were assigned to (Supplementary Tables 2 – 4).

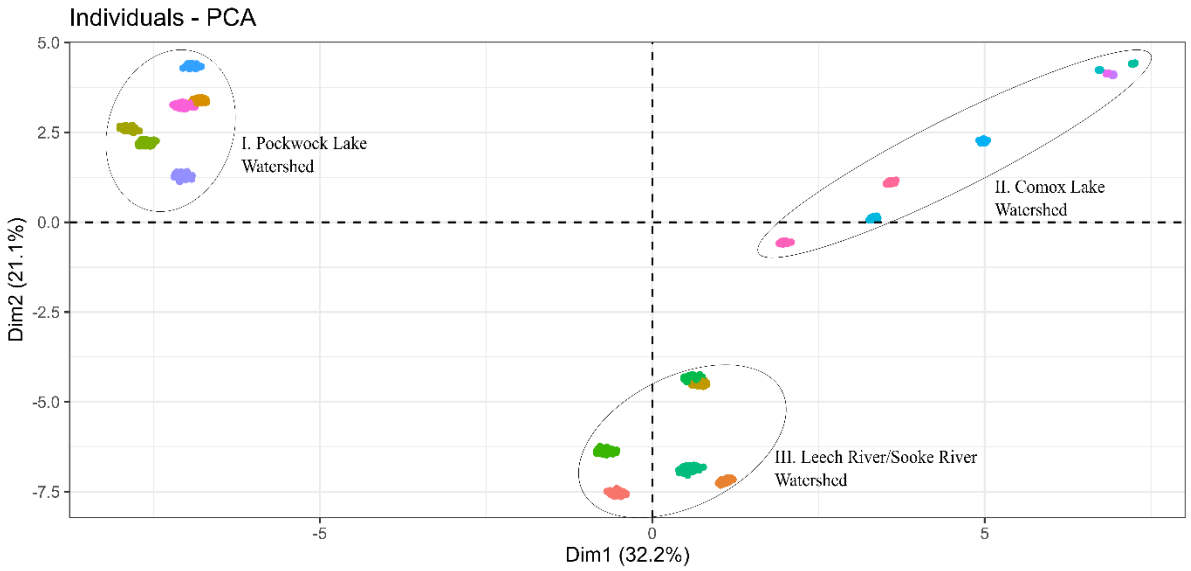


Figure 3.9 Beta diversity the bacterial communities of sample lakes within the watersheds from the Atlantic and Pacific maritime ecozones. PCA plot was constructed with a rarefied library size of 5000 sequences and replicated with 100 iterations generated using Bray-Curtis distances in the R package *mirlyn*. Three distinct groups were observed which represented each watershed (I: Pockwock Lake watershed, II: Comox Lake watershed, III: Leech River/Sooke River watershed).

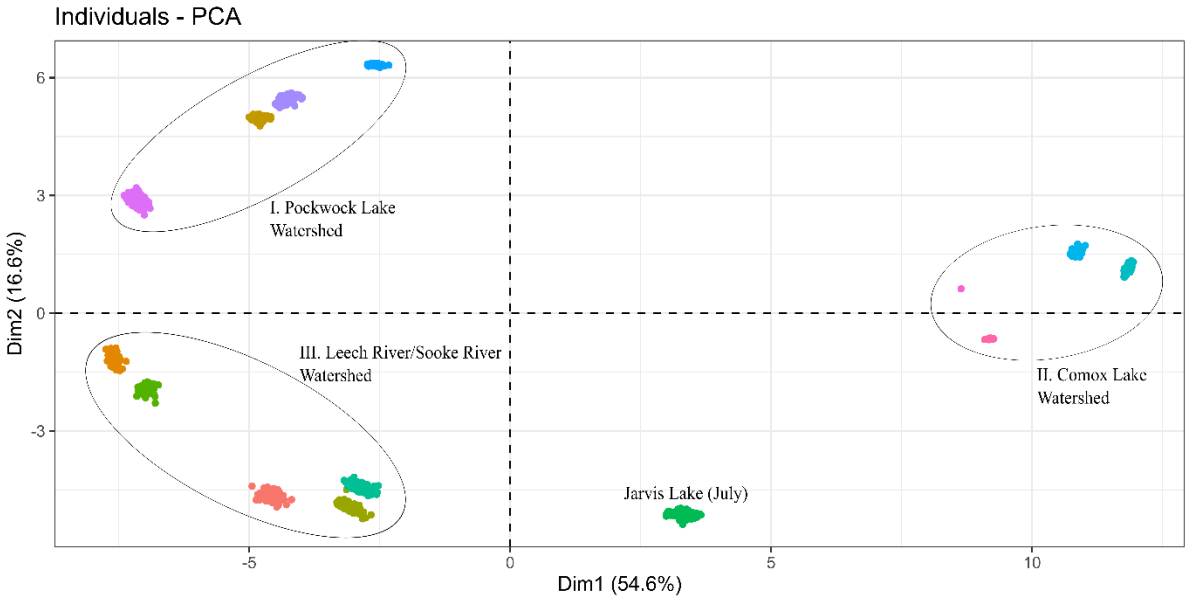


Figure 3.10 Beta diversity the cyanobacterial communities of sample lakes within the watersheds from the Atlantic and Pacific maritime ecozones. PCA plot was constructed with a rarefied library size of 860 sequences and replicated with 100 iterations generated using Bray-Curtis distances in the R package *mirlyn*. Three distinct groups were observed which represented each watershed (I: Pockwock Lake watershed, II: Comox Lake watershed, III: Leech River/Sooke River watershed) excluding Jarvis Lake in July.

3.6 Cyanobacteria Classification

There were 113 cyanobacterial ASVs observed with 79 ASVs containing taxonomic resolution at the species and genus-levels. Of these 79 ASVs, at the genus-level, there were 26 ASVs were assigned to *Cyanobium* PCC-6307 (NR_102447.1) and nine ASVs to *Rhabdogloea smithii* SAG 47.91 (KM020002.1), making these taxa the most diverse. Additionally, five ASVs were assigned to *Aphanizomenon* NIES81 (AJ293131.1) and *Microcystis* PCC-7914 (no GenBank accession number) and three ASVs assigned to *Calothrix* KVSF5 (EU022730.1), *Chamaesiphon* PCC-7430 (AY170472.1), *Gloeobacter* PCC-7421 (NR_074282.1), *Scytonema* UTEX 2349 (NZ_ALWD00000000.1) and *Tychonema* CCAP 1459-11B (AB045897.1), and two ASVs assigned to *Gleocapsa* (no GenBank accession number and incorrectly spelled in SILVA with the correct spelling being *Gloeocapsa*), *Leptolyngbya* ANT.L52.2 (AY493575.1), *Kamptonema* PCC-6407

(AM398782.1) and *Synechococcus* PCC-7502 (AF448080.1). There were 10 other genera that contained one ASV each.

Taxonomic classification of these 79 ASVs contained some inconsistences from SILVA between genus and species. Of the five ASVs assigned to *Microcystis* PCC-7914 at the genus-level, at the species-level, these ASVs were assigned to either *Cyanothece aeruginosa* SAG 87.79 (KM019992.1) or *Radiocystis* sp. JJ30-12 (AM710388.1). The one ASV assigned to *Calothrix* PCC-6303 (NC_019751.1) at the genus-level was assigned to *Macrochaete psychrophila* CCALA 32 (KT336439.2) at the species-level. One of the ASVs assigned to *Kamptonema* PCC-6407 (AM398782.1) at the genus-level was assigned to *Oscillatoriales cyanobacterium* USR001 (MBRE01000011.1) at the species-level. The last inconsistent classification was from one ASV assigned to *Cyanobium* PCC-6307 (NR_102447.1) at the genus-level and *Synechococcus* sp. LEGE 06306 (HM217052.1) at the species-level.

3.6.1 Cyanobacteria Phylogeny

A phylogenetic tree of the 79 cyanobacteria ASVs constructed using reference sequences from NCBI allowed for observing evolutionary relatedness of sequences obtained from this study to those that have been previously characterized. Reference sequences utilized for phylogenetic analysis were selected based on the SILVA classification of cyanobacteria ASVs. The SILVA classification of the 79 cyanobacteria ASVs with the associated identifier and the accession code for the reference sequences used in the phylogenetic tree are available in a supplementary table (Supplementary Table 5). The remaining 34 cyanobacteria ASVs that were excluded from phylogenetic analysis due to being unresolved to the genus-level are available in a supplementary table (Supplementary Table 6).

Observations of the phylogeny of ASVs and reference sequences included clustering of sequences within six different orders (Figure 3.11). The order Synechococcales was the most diverse and contained five separate clusters, two of which included ASV 14 – 39 which were assigned to *Cyanobium* PCC-6307 (NR_102447.1) and ASV 59 – 68 which were

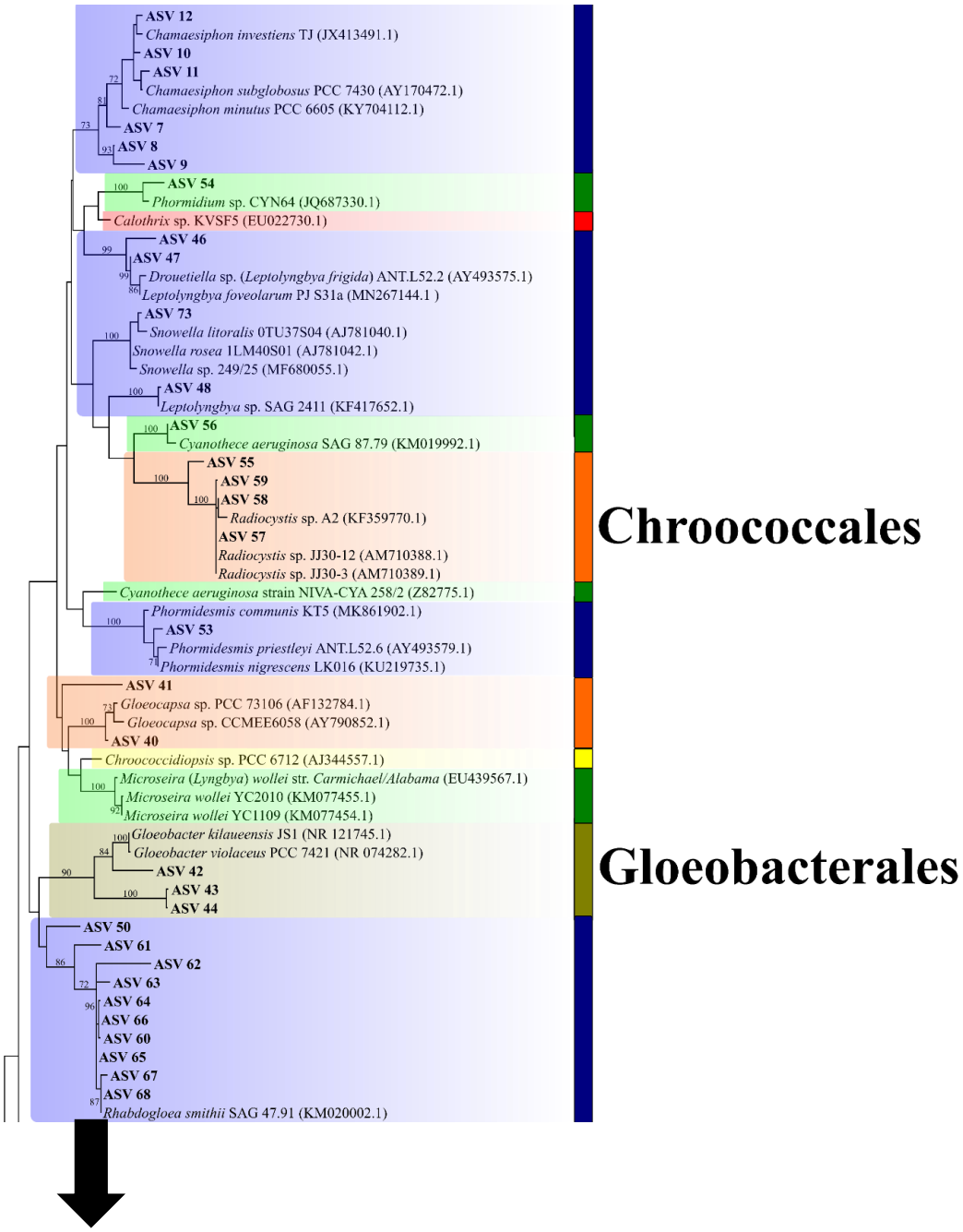
assigned to *Rhabdogloea smithii* SAG 47.91 (KM020002.1), both of which clustered with the respective reference sequences. Within the cluster of ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1), including ASV 74 – 75, were alignments with sequences from the genus *Synechococcus*, suggesting sequence similarity between these genera. Additional clusters under the order Synechococcales was observed with clustering of ASV 7 – 12 with the genus *Chamaesiphon*, ASV 46 – 48 with *Leptolyngbya*, ASV 53 with *Phormidesmis* and ASV 73 with *Snowella*.

The order Nostocales contained one large cluster with ASV 2 – 6 clustering with reference sequences from the genus *Aphanizomenon,* ASV 49 with *Calothrix* and *Macrochaete*, ASV 51 with *Nostoc,* and ASV 69 – 72 with *Scytonema*. Interestingly, the reference sequence *Calothrix* KVSF5 (EU022730.1) did not cluster within the order Nostocales but this sequence was still included in the phylogenetic tree as ASV 7 – 9 were assigned to these taxa by SILVA.

The order Oscillatoriales contained two clusters with ASV 45 and ASV 52 clustering with reference sequences from the genera *Kamptonema, Oscillatoria* and *Phormidium* and ASV 76 – 79 with *Phormidium* and *Tychonema.* Two ASVs aligned with reference sequences outside of these clusters which was ASV 54 with *Phormidium* sp. CYN64 (JQ687330.1) and ASV 56 with *Cyanothece aeruginosa* SAG 87.79 (KM019992.1). Reference sequences from the genus *Microseira* formed a separate cluster and did not align with ASV 50 despite classification by SILVA.

The order Chroococcidiopsidales contained two clusters with ASV 13 clustering with reference sequences from the genus *Chroococcidiopsis* and ASV 1 with *Aliterella* and *Gloeocapsa*. Two distinct clusters of the order Chroococcales was observed with ASV 55 – 59 clustering with reference sequences from the genus *Radiocystis* and ASV 40 – 41 with *Gloeocapsa*. The order Gloeobacterales contained one cluster with ASV 42 – 44 clustering with reference sequences from the genus *Gloeobacter*.

.


ASV 12
Chamaesiphon investiens TJ (JX413491.1)
ASV 10
ASV 11
Chamaesiphon subglobosus PCC 7430 (AY170472.1)
Chamaesiphon minutus PCC 6605 (KY704112.1)
ASV 7
ASV 8
ASV 9
ASV 54
Phormidium sp. CYN64 (JQ687330.1)
Calothrix sp. KVSF5 (EU022730.1)
ASV 46
ASV 47
Drouetiella sp. (Leptolyngbya frigida) ANT.L52.2 (AY493575.1)
Leptolyngbya foveolarum PJ S31a (MN267144.1)
ASV 73
Snowella litoralis 0TU37S04 (AJ781040.1)
Snowella rosea 1LM40S01 (AJ781042.1)
Snowella sp. 249/25 (MF680055.1)
ASV 48
Leptolyngbya sp. SAG 2411 (KF417652.1)
ASV 56
Cyanothece aeruginosa SAG 87.79 (KM019992.1)
ASV 55
ASV 59
ASV 58
Radiocystis sp. A2 (KF359770.1)
ASV 57
Radiocystis sp. JJ30-12 (AM710388.1)
Radiocystis sp. JJ30-3 (AM710389.1)
Cyanothece aeruginosa strain NIVA-CYA 258/2 (Z82775.1)
Phormidesmis communis KT5 (MK861902.1)
ASV 53
Phormidesmis priestleyi ANT.L52.6 (AY493579.1)
Phormidesmis nigrescens LK016 (KU219735.1)
ASV 41
Gloeocapsa sp. PCC 73106 (AF132784.1)
Gloeocapsa sp. CCMEE6058 (AY790852.1)
ASV 40
Chroococcidiopsis sp. PCC 6712 (AJ344557.1)
Microseira (Lyngbya) wollei str. Carmichael/Alabama (EU439567.1)
Microseira wollei YC2010 (KM077455.1)
Microseira wollei YC1109 (KM077454.1)
Gloeobacter kilaueensis JS1 (NR 121745.1)
Gloeobacter violaceus PCC 7421 (NR 074282.1)
ASV 42
ASV 43
ASV 44
ASV 50
ASV 61
ASV 62
ASV 63
ASV 64
ASV 66
ASV 60
ASV 65
ASV 67
ASV 68
Rhabdogloea smithii SAG 47.91 (KM020002.1)
Chroococcales
Gloeobacterales

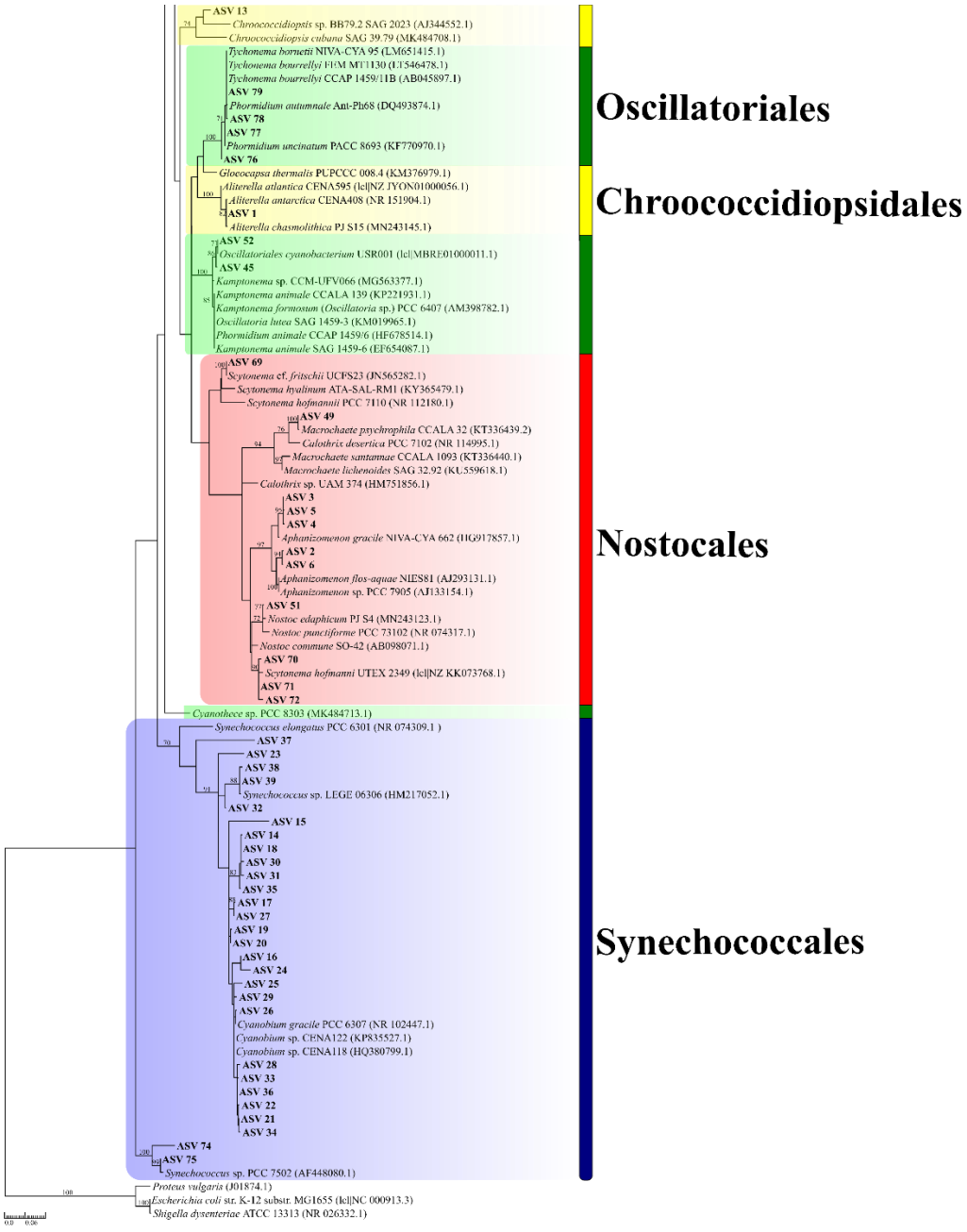


Figure 3.11 Phylogenetic tree of ASVs from sample lakes assigned to cyanobacteria. Reference sequences are included for taxonomic resolution. Cyanobacteria ASV are bolded. Phylogenetic tree was constructed in MEGA X using the Maximum Likelihood method and a bootstrap of 1000 with values of at least 70% provided on branches. There were 34 sequences from sample lakes that were omitted from this tree as they were unresolved to the genus-level and all of them grouped together separately from any reference sequences. Colour shading indicates the order that sequences clustered within.

3.6.2 *Cyanobium* Phylogenetic Diversity

While clusters of ASVs and reference sequences were observed, the ASVs that were assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1), the two genera with the most abundant and diverse ASVs, were observed to have variations in sequence similarities. Within the phylogenetic tree, clusters ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) contained various branching patterns from the reference sequences which may indicate these genera are more diverse than currently characterized (Figure 3.11). Further phylogenetic analysis of the 26 ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1), and two ASVs to *Synechococcus* PCC-7502 (AF448080.1) as these genera seem to share sequence similarity, with additional reference sequences obtained from Genuário *et al*. (2016), indicated these genera share sequence similarity and species diversity that has yet to be characterized (Figure 3.12).

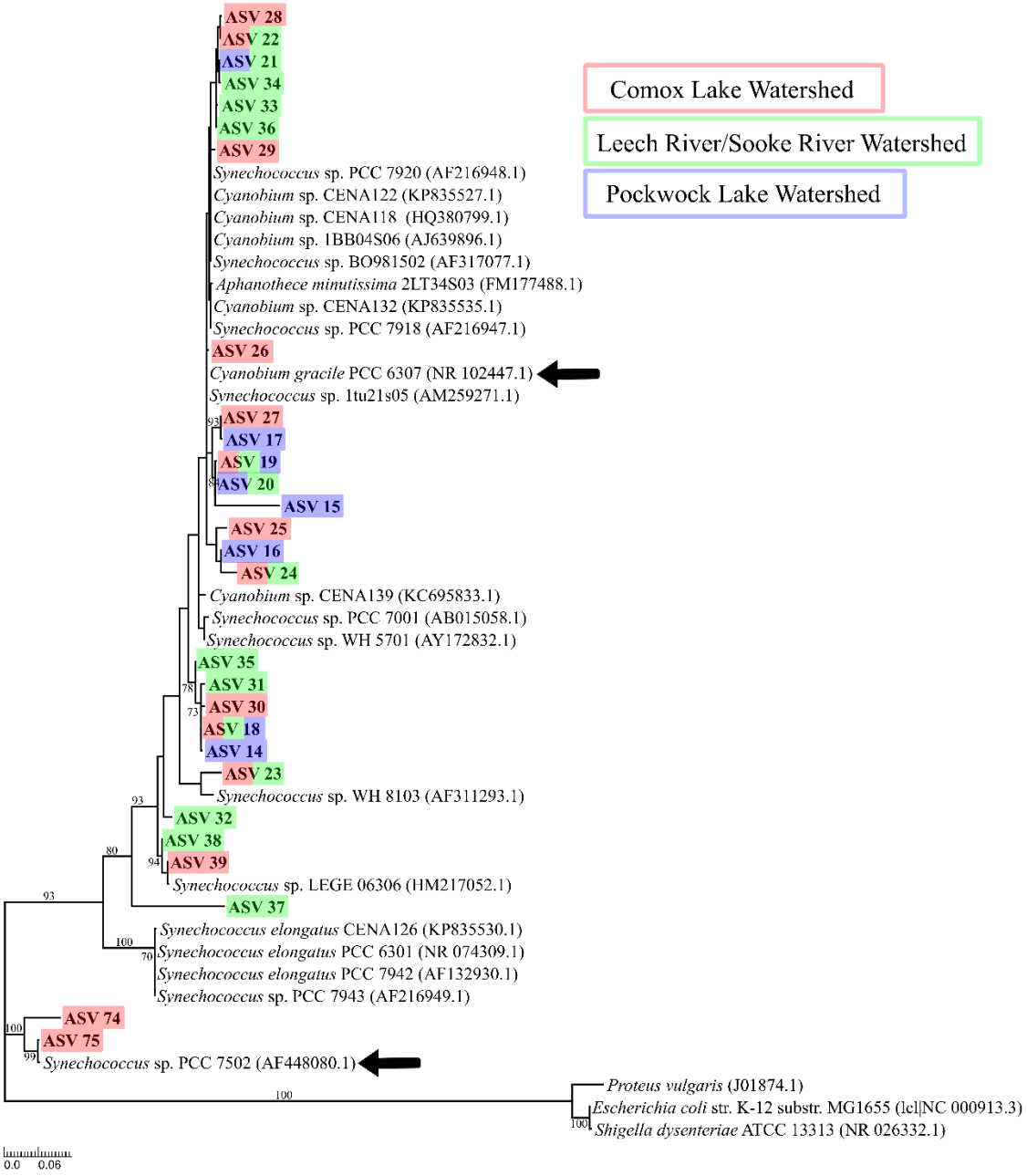
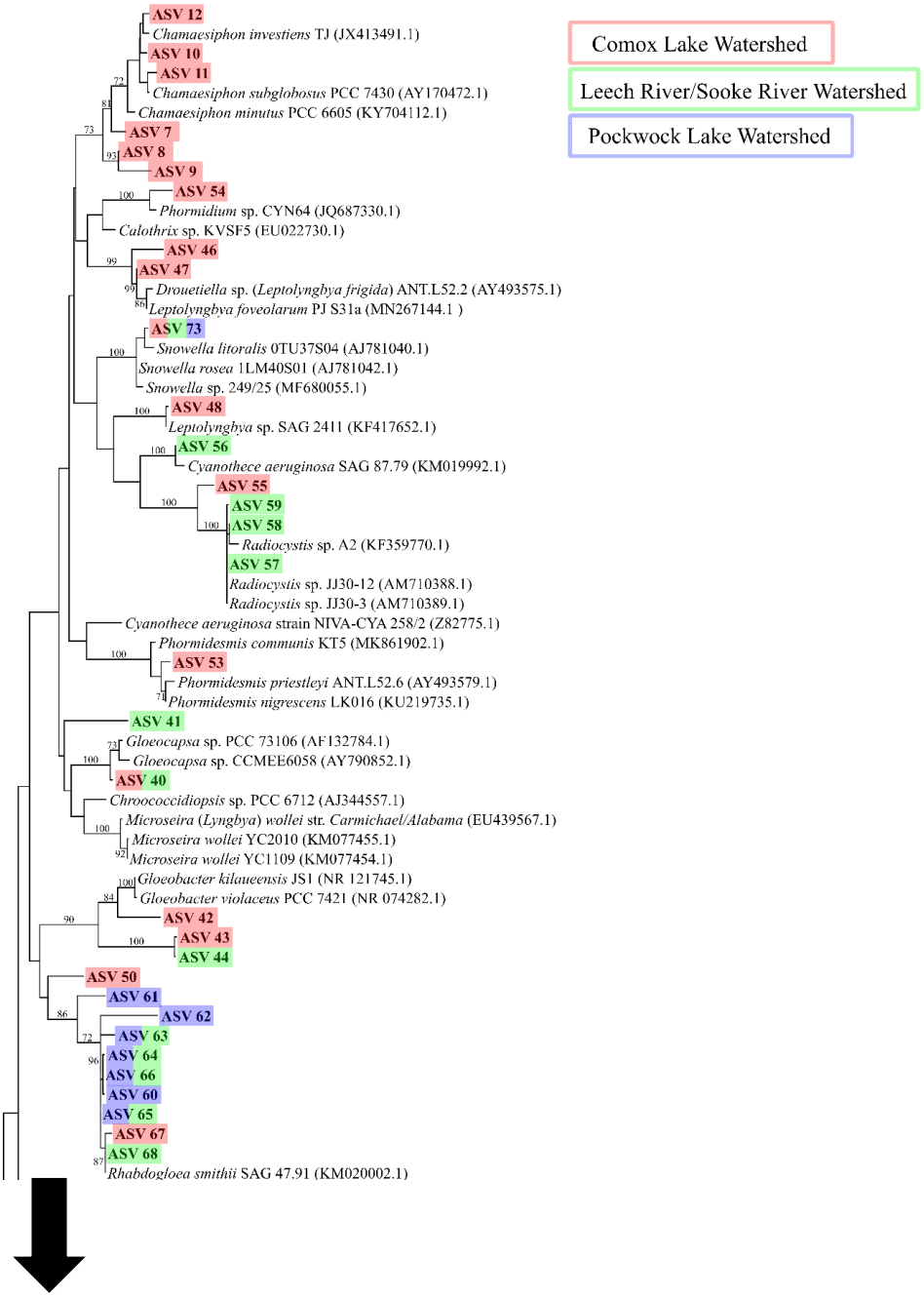


Figure 3.122 Phylogenetic tree of ASVs from sample lakes assigned to *Cyanobium* PCC-6307. Reference sequences are included for taxonomic resolution and were obtained from Genuário *et al*. (2016). The ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1) observed from sample lakes are bolded. Reference sequences with arrows are the taxa these ASVs were assigned to by SILVA. Phylogenetic tree was constructed in MEGA X using the Maximum Likelihood method and a bootstrap of 1000 with values of at least 70% provided on branches. Colour shading of ASVs indicates the watershed(s) observed in.

3.7 Biogeographic Distribution of Cyanobacteria ASVs

To visualize the geographical distribution of cyanobacteria ASVs, each of the ASVs observed among sample lakes were colour shaded within a phylogenetic tree based on watershed they were observed in (Figure 3.13). The Comox Lake watershed contained the most diverse genera of cyanobacteria based on the number of ASVs from this watershed and from the clustering patterns of ASVs to reference sequences. Most of the cyanobacteria ASVs from the Comox Lake watershed samples were unique to this watershed and were not observed in the other watersheds. The Leech River/Sooke River watershed also contained some cyanobacteria ASVs unique to only these watersheds but shared most ASVs with those also observed from samples in the Pockwock Lake watershed including those assigned to *Aphanizomenon* NIES81 (AJ293131.1), *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 41.97 (KM020002.1). It is interesting to note that the Leech River/Sooke River watersheds and Pockwock Lake watershed share similar cyanobacteria ASVs despite being in different ecozones, the Pacific maritime and Atlantic maritime, respectively.


Comox Lake Watershed
Leech River/Sooke River Watershed
Pockwock Lake Watershed
ASV 12
Chamaesiphon investiens TJ (JX413491.1)
ASV 10
ASV 11
Chamaesiphon subglobosus PCC 7430 (AY170472.1)
Chamaesiphon minutus PCC 6605 (KY704112.1)
ASV 7
ASV 8
ASV 9
ASV 54
Phormidium sp. CYN64 (JQ687330.1)
Calothrix sp. KVSF5 (EU022730.1)
ASV 46
ASV 47
Drouetiella sp. (Leptolyngbya frigida) ANT.L52.2 (AY493575.1)
Leptolyngbya foveolarum PJ S31a (MN267144.1)
ASV 73
Snowella litoralis 0TU37S04 (AJ781040.1)
Snowella rosea 1LM40S01 (AJ781042.1)
Snowella sp. 249/25 (MF680055.1)
ASV 48
Leptolyngbya sp. SAG 2411 (KF417652.1)
ASV 56
Cyanothece aeruginosa SAG 87.79 (KM019992.1)
ASV 55
ASV 59
ASV 58
Radiocystis sp. A2 (KF359770.1)
ASV 57
Radiocystis sp. JJ30-12 (AM710388.1)
Radiocystis sp. JJ30-3 (AM710389.1)
Cyanothece aeruginosa strain NIVA-CYA 258/2 (Z82775.1)
Phormidesmis communis KT5 (MK861902.1)
ASV 53
Phormidesmis priestleyi ANT.L52.6 (AY493579.1)
Phormidesmis nigrescens LK016 (KU219735.1)
ASV 41
Gloeocapsa sp. PCC 73106 (AF132784.1)
Gloeocapsa sp. CCMEE6058 (AY790852.1)
ASV 40
Chroococcidiopsis sp. PCC 6712 (AJ344557.1)
Microseira (Lyngbya) wollei str. Carmichael/Alabama (EU439567.1)
Microseira wollei YC2010 (KM077455.1)
Microseira wollei YC1109 (KM077454.1)
Gloeobacter kilaueensis JS1 (NR 121745.1)
Gloeobacter violaceus PCC 7421 (NR 074282.1)
ASV 42
ASV 43
ASV 44
ASV 50
ASV 61
ASV 62
ASV 63
ASV 64
ASV 66
ASV 60
ASV 65
ASV 67
ASV 68
Rhabdogloea smithii SAG 47.91 (KM020002.1)

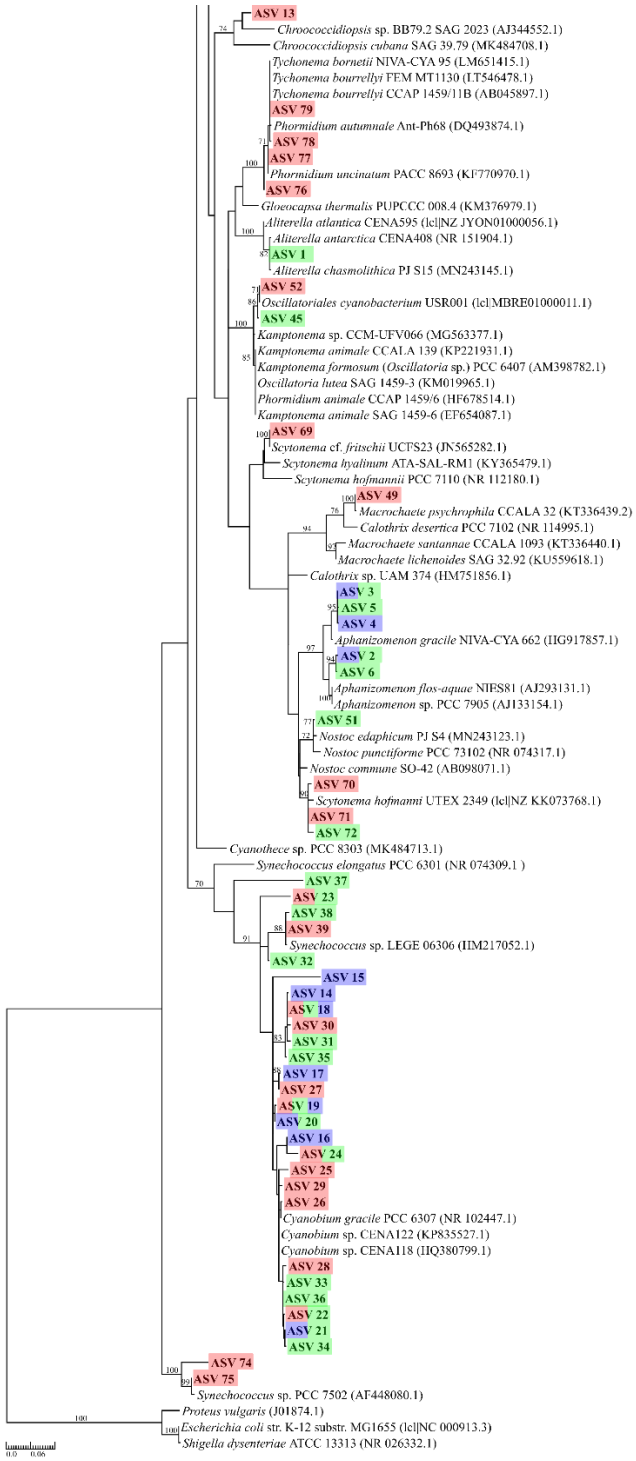


Figure 3.133 Phylogenetic tree of ASVs assigned to cyanobacteria and their associated watershed(s). This is the same phylogenetic tree provided in Figure 3.11 instead with colour shading of ASVs indicating the watershed(s) these ASVs were observed in.

3.8 Comparison of Taxonomic Classification by SILVA and BLAST

Observations from the phylogenetic tree of cyanobacteria ASVs indicated that, while clusters were observed with reference sequences, taxonomic classification by SILVA may not have correctly assigned taxonomy to all ASVs. To evaluate the accuracy of taxonomic assignment of cyanobacteria ASVs by SILVA, the 79 ASVs from the phylogenetic tree were searched against sequences in NCBI using BLAST. The top match(es) of each ASV from the BLAST output were obtained and were dependent on percent similarity, query cover and E-value. From the BLAST output, taxonomy was compared against SILVA taxonomic assignment which identified 34 ASVs that shared at least the same genus with the top match from BLAST, 11 ASVs which matched with multiple different genera but at least one with the same genus or species, and 34 ASVs that were mismatches (Table 3.12). The ASVs that shared the same taxonomic assignment between SILVA and BLAST to the genus or species level were shaded in green. The ASVs that shared taxonomic assignment at the genus or species level but contained multiple matches were shaded in grey. Those that did not match taxonomy between SILVA and BLAST were left unshaded (white).

Table 3.12 Comparison of taxonomic assignment to cyanobacteria ASVs from sample lakes by SILVA and BLAST.

	SILVA Classification		**BLAST Classification**	
ASV ID	**Genus**	**Species**	**Scientific Name**	**Similarity (%)**
ASV 1	*Aliterella* CENA595	unclassified	*Aliterella antarctica* *Aliterella chasmolithica* *Aliterella* sp.	99.6%
ASV 2	*Aphanizomenon* NIES81	unclassified	*Dolichospermum lemmermannii*	99.6%
ASV 3	*Aphanizomenon* NIES81	unclassified	*Anabaena* sp.	98.81%
ASV 4	*Aphanizomenon* NIES81	unclassified	*Anabaena* sp.	98.81%
ASV 5	*Aphanizomenon* NIES81	unclassified	*Anabaena* sp.	98.42%
ASV 6	*Aphanizomenon* NIES81	unclassified	*Dolichospermum lemmermannii*	99.6%
ASV 7	*Calothrix* KVSF5	unclassified	*Chamaesiphon* sp.	96.08%
ASV 8	*Calothrix* KVSF5	unclassified	*Chamaesiphon* cf. *incrustans* str. Ch. *fontanile*	94.86%
ASV 9	*Calothrix* KVSF5	unclassified	*Chamaesiphon* cf. *incrustans* str. Ch. *fontanile*	92.09%
ASV 10	*Chamaesiphon* PCC-7430	unclassified	*Placoma regulare*	98.42%
ASV 11	*Chamaesiphon* PCC-7430	unclassified	*Chamaesiphon* sp. *Chamaesiphon subglobosus* *Synechocystis pevalekii*	98.02%
ASV 12	*Chamaesiphon* PCC-7430	unclassified	*Chamaesiphon investiens*	98.81%
ASV 13	*Chroococcidiopsis* SAG 2023	unclassified	*Chroococcidiopsis* sp.	92.89%
ASV 14	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	98.42%

ASV 15	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	92.91%
ASV 16	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	99.21%
ASV 17	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	99.6%
ASV 18	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	98.81%
ASV 19	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	100%
ASV 20	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	99.6%
ASV 21	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	99.6%
ASV 22	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp. *Synechococcus* sp.	99.6%
ASV 23	*Cyanobium* PCC-6307	unclassified	*Candidatus Atelocyanobacterium thalassa* *Synechococcus rubescens* *Synechococcus* sp.	96.05%
ASV 24	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	100%
ASV 25	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	99.21%
ASV 26	*Cyanobium* PCC-6307	unclassified	*Aphanocapsa salina* *Aphanothece* sp. *Cyanobium gracile* *Cyanobium* sp. *Synechococcus elongatus* *Synechococcus* sp.	99.6%

ASV 27	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	100%
ASV 28	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	100%
ASV 29	*Cyanobium* PCC-6307	unclassified	*Cyanobium* cf. *plancticum* *Cyanobium* sp. *Synechococcus* sp.	99.21%
ASV 30	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	98.02%
ASV 31	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	99.6%
ASV 32	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	96.44%
ASV 33	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	99.6%
ASV 34	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	99.21%
ASV 35	*Cyanobium* PCC-6307	unclassified	*Cyanobium* sp.	98.02%
ASV 36	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	100%
ASV 37	*Cyanobium* PCC-6307	unclassified	*Synura uvella*	89.33%
ASV 38	*Cyanobium* PCC-6307	unclassified	*Synechococcus* sp.	98.81%
ASV 39	*Cyanobium* PCC-6307	*Synechococcus* sp. LEGE 06306	*Synechococcus* sp.	99.6%
ASV 40	*Gloeocapsa*	unclassified	*Limnococcus limneticus*	100%
	Gloeocapsa	unclassified	*Chroococcus minutus*	100%

ASV 41			*Chroococcus* sp.	
ASV 42	*Gloeobacter* PCC-7421	unclassified	*Aphanothece caldoriarum* var. *cavernarum* *Gloeobacter kilaueensis*	91.73%
ASV 43	*Gloeobacter* PCC-7421	unclassified	*Cyanosarcina* sp. *Synechococcus* sp.	88.19%
ASV 44	*Gloeobacter* PCC-7421	unclassified	*Cyanosarcina* sp. *Synechococcus* sp.	87.75%
ASV 45	*Kamptonema* PCC-6407	unclassified	*Kamptonema formosum*	98.81%
ASV 46	*Leptolyngbya* ANT.L52.2	unclassified	*Leptolyngbya* sp.	98.81%
ASV 47	*Leptolyngbya* ANT.L52.2	unclassified	*Leptolyngbya* sp.	98.81%
ASV 48	*Leptolyngbya* SAG 2411	unclassified	*Leptolyngbya* sp.	100%
ASV 49	*Calothrix* PCC-6303	*Macrochaete psychrophila* CCALA 32	*Macrochaete psychrophila*	99.6%
ASV 50	*Microseira Carmichael-Alabama*	unclassified	*Phormidium* cf. *nigrum* *Phormidium* sp.	94.07%
ASV 51	*Nostoc* PCC-73102	unclassified	*Aulosira terrestre*	100%
ASV 52	*Kamptonema* PCC-6407	*Oscillatoriales cyanobacterium* USR001	*Kamptonema formosum*	99.61%
ASV 53	*Phormidesmis* ANT.L52.6	unclassified	*Phormidesmis* sp.	99.21%
ASV 54	*Phormidium* CYN64	unclassified	*Timaviella obliquedivisa* *Timaviella* sp.	97.23%

ASV 55	*Microcystis* PCC-7914	*Radiocystis* sp. JJ30-12	*Radiocystis* sp.	94.14%
ASV 56	*Microcystis* PCC-7914	*Cyanothece aeruginosa* SAG 87.79	*Chlorogloea purpurea*	98.42%
ASV 57	*Microcystis* PCC-7914	*Radiocystis* sp. JJ30-12	*Radiocystis* sp.	100%
ASV 58	*Microcystis* PCC-7914	*Radiocystis* sp. JJ30-12	*Radiocystis* sp.	99.6%
ASV 59	*Microcystis* PCC-7914	*Radiocystis* sp. JJ30-12	*Radiocystis* sp.	99.6%
ASV 60	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii*	98.42%
ASV 61	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii*	91.7%
ASV 62	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* *Synechocystis fuscopigmentosa*	91.3%
ASV 63	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* *Synechocystis fuscopigmentosa*	97.63%
ASV 64	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii*	98.42%
ASV 65	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii*	99.21%
ASV 66	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii*	98.81%
ASV 67	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* *Synechocystis fuscopigmentosa*	98.81%
ASV 68	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii* SAG 47.91	*Rhabdogloea smithii*	100%
ASV 69	*Scytonema* PCC-7110	unclassified	*Scytonema* cf. *fritschii*	100%

ASV 70	*Scytonema* UTEX 2349	unclassified	*Coleodesmium* sp. *Dactylothamnos antarcticus* *Hassallia antarctica* *Tolypothrix* sp. *Tolypothrix tenuis*	99.21%
ASV 71	*Scytonema* UTEX 2349	unclassified	*Coleodesmium* sp. *Dactylothamnos antarcticus* *Hassallia antarctica* *Tolypothrix* sp. *Tolypothrix tenuis*	100%
ASV 72	*Scytonema* UTEX 2349	unclassified	*Hassallia andreassenii*	100%
ASV 73	*Snowella* 0TU37S04	unclassified	*Snowella litoralis* *Snowella rosea*	98.02%
ASV 74	*Synechococcus* PCC-7502	unclassified	*Synechococcus* sp.	93.7%
ASV 75	*Synechococcus* PCC-7502	unclassified	*Synechococcus* sp.	99.21%
ASV 76	*Tychonema* CCAP 1459-11B	unclassified	*Microcoleus vaginatus* *Phormidium autumnale* *Phormidium* sp.	99.6%
ASV 77	*Tychonema* CCAP 1459-11B	unclassified	*Microcoleus anatoxicus* *Microcoleus* sp. *Microcoleus vaginatus* *Oscillatoria limosa* *Phormidium autumnale* *Phormidium* cf. *subfuscum* *Phormidium* cf. *uncinatum* *Phormidium* sp.	100%

			Phormidium uncinatum *Tychonema bourrellyi*	
ASV 78	*Tychonema* CCAP 1459-11B	unclassified	*Microcoleus anatoxicus* *Microcoleus* sp. *Phormidium autumnale* *Phormidium* cf. *autumnale* *Phormidium* cf. *irriguum* *Phormidium* cf. *uncinatum* *Phormidium* sp. *Tychonema bornetii* *Tychonema bourrellyi* *Tychonema* sp. *Tychonema tenue* *Wilmottia murrayi*	99.6%
ASV 79	*Tychonema* CCAP 1459-11B	unclassified	*Microcoleus anatoxicus* *Microcoleus* sp. *Phormidium autumnale* *Phormidium* cf. *autumnale* *Phormidium* cf. *irriguum* *Phormidium* cf. *uncinatum* *Phormidium* sp. *Tychonema bornetii* *Tychonema bourrellyi* *Tychonema* sp. *Tychonema tenue* *Wilmottia murrayi*	100%

The ASVs with multiple top matches from BLAST contained the same percent similarity as well as query cover and E-value (not shown).

3.9 *mcy*E and *geo*A Marker Gene Detection

Sequences obtained from the August sample from Weeks Lake using the primers HEPF/HEPR were likely artefacts and did not represent the aminotransferase (AMT) region from the *mcy*E gene. The primers HEPF/HEPR are well defined primers for the detection of the microcystin toxin gene which may indicate that no cyanobacteria from this sample, despite containing many reads primarily assigned to *Cyanobium* PCC-6307 (NR_102447.1) and some to *Microcystis* PCC-7914 (no GenBank accession number), are toxin producers. Several *geo*A sequences were obtained from the August Weeks Lake sample, however, similar to the *mcy*E gene, the sequences obtained using the geoA-297f/geoA-552r primers may have yielded artefacts as they did not align with well characterized geosmin genes from cyanobacteria and *Streptomyces*, including those that were used to create these primers.

Chapter 4: Discussion

4.1 Cyanobacterial Communities from the Pockwock Lake Watershed

Sequence reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) primarily comprised of the cyanobacteria communities from the Pockwock Lake watershed. There were also sequences assigned to *Aphanizomenon* NIES81 (AJ293131.1) observed in the June sample from Island Lake, as well as the October sample from Pockwock Lake. Although there were reads assigned to *Aphanizomenon* NIES81 (AJ293131.1), phylogenetic tree alignments and BLAST results indicate that these sequences may have closer alignments to *Anabaena* and *Dolichospermum,* with sequence similarities to these genera ≥98%. The genera *Anabaena, Aphanizomenon* and *Dolichospermum* can threaten water quality as species can produce geosmin (Li *et al.*, 2016; Wang *et al*., 2019; Churro *et al*., 2020) and various toxins including anatoxin-a, cylindrospermopsins microcystin and saxitoxin (Lyra *et al.,* 2001; Al-Tebrineh *et al.,* 2010; Engström-Öst *et al.*, 2011; Cirés and Ballot, 2016; Huisman *et al*., 2018; Du *et al*., 2019). Furthermore, *Anabaena, Aphanizomenon* and *Dolichospermum* are capable of nitrogen (N_2) fixation, allowing these genera to potentially form blooms regardless of bioavailable external sources of N and P (Yema *et al.*, 2016) or light availability (Bradburn *et al.*, 2012).

The findings in this study are significant as geosmin has been previously detected in Pockwock Lake, thought to originate in Island Lake, and was associated with the cyanobacteria *Anabaena* (Anderson *et al*., 2017). Observations in this study may reflect this as there were reads assigned to *Aphanizomenon* NIES81 (AJ293131.1) with high sequence similarity to *Anabaena* and *Dolichospermum*. Therefore, this study indicates that Island Lake may still be a source of geosmin producers with the potential for them to flow into Pockwock Lake, the primary drinking water source for the Halifax Regional Municipality (HRM). The presence of these genera can also be cause for concern due to their bloom forming capabilities, influenced by the ability of N_2 fixation. As these taxa are potentially geosmin and toxin producers and bloom formers, their presence is significant as they can pose water quality issues and significant water treatment costs if a proliferation event were to occur,

thereby threatening the integrity of the HRM drinking water source (Emelko *et al.*, 2011; Dunlap *et al*., 2015).

4.2 Cyanobacterial Communities from the Comox Lake Watershed

Boston Creek in May primarily contained reads assigned to *Calothrix* KVSF5 (EU022730.1) and *Chamaesiphon* PCC-7430 (AY170472.1) compared to September when it was primarily reads assigned to *Scytonema* UTEX 2349 (NZ_ALWD00000000.1). *Calothrix* is a genetically diverse group of benthic cyanobacteria that has been observed in a range of aquatic habitats, including freshwater and marine (Sihvonen *et al*., 2007; Berrendero *et al*., 2011). *Chamaesiphon* has previously been characterized as a biofilm forming and epilithic cyanobacteria within lotic systems such as streams and rivers and can tolerate various levels of exposure to light, pH ranges and nutrient concentrations, including oligotrophic and mesotrophic conditions (Loza *et al.*, 2013; Kurmayer *et al*., 2018). *Scytonema* has been identified in the littoral zone of water sources with a range of trophic levels with some species capable of producing the neurotoxin saxitoxin (Smith *et al*., 2012).

From phylogenetic analysis of the ASVs from Boston Creek, sequences assigned to *Calothrix* KVSF5 (EU022730.1) did not cluster with other *Calothrix* reference sequences. There were three ASVs assigned to *Calothrix* KVSF5 (EU022730.1) and observed through using BLAST, all shared sequence similarity with *Chamaesiphon* with 92 – 96% similarity. Additionally, these ASVs clustered closer to those assigned to *Chamaesiphon* PCC-7430 (AY170472.1) in the phylogenetic tree. The ASVs that were assigned to *Chamaesiphon* PCC-7430 (AY170472.1) did cluster with *Chamaesiphon* reference sequences, though one ASV, observed through BLAST, contained sequence similarity (98%) with *Placoma regulare* (KF264594.1), a cyanobacterium that has been observed in small streams attached to rocks and bryophytes as well as in rivers with moderate nutrient concentrations (Broady and Ingerfeld, 1991; Carmona-Jiménez and Caro-Borrero, 2017).

The Boston Creek sample in September contained ASVs primarily assigned to *Scytonema* UTEX 2349 (NZ_ALWD00000000.1) and, while these ASVs clustered together in the phylogenetic tree, they did not cluster well with other *Scytonema* reference sequences.

Rather, the three ASVs assigned to *Scytonema* UTEX 2349 (NZ_ALWD00000000.1), using BLAST, were observed to align closer with a range of genera including *Coleodesmium, Dactylothamnos, Hassallia* and *Tolypothrix* with sequence similarities of 99 – 100%. The genera *Coleodesmium Dactylothamnos*, *Hassallia* and *Tolypothrix* are from the recently described Tolypothrichaceae family (Hauer *et al*., 2014) with *Coleodesmium* and *Tolypothrix* being observed to be benthic (Monteagudo and Moreno, 2016) while *Hassallia* and *Dactylothamnos* have been observed in the littoral zone from streams and lakes and attached to rocks (Komárek *et al*., 2015).

The Cruikshank River sample in both May and September contained cyanobacteria communities primarily composed of reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Tychonema* CCAP 1459-11B (AB045897.1) in May and mainly *Tychonema* CCAP 1459-11B (AB045897.1) in September. Presence of species of *Tychonema* have been observed to be planktonic and can be significant as some have been identified to contain *ana*C and *ana*F genes to produce the neurotoxin anatoxin-a (Salmaso *et al*., 2016). From the phylogenetic tree, the four ASVs assigned to this genus did cluster together with *Tychonema* reference sequences but BLAST results indicated a more complicated taxonomy. From BLAST, while these ASVs aligned with taxa from the genus *Tychonema,* alignments to a range of genera were also observed including to *Microcoleus, Phormidium, Oscillatoria* and *Wilmottia* with 99 – 100% sequence similarity. *Wilmottia* is a recently described genus previously classified under *Phormidium* and have been observed as being planktic, benthic or attached to sediment within rivers, streams and lakes (Comte *et al*., 2007; Hašler *et al*., 2012; Stoyanov *et al*., 2014; Heath *et al*., 2015; Salmaso *et al*., 2016). The genera *Microcoleus, Oscillatoria* and *Phormidium* have been observed as epipelic from growing on sediment (Hašler *et al.*, 2012) with some species of *Oscillatoria* also being planktonic (Izaguirre and Taylor, 2004). This is significant as species of *Microcoleus, Phormidium, Oscillatoria* and *Tychonema* have been observed to produce geosmin (Izaguirre and Taylor, 2004; Wang *et al*., 2019; Churro *et al*., 2020) and with some species of *Phormidium* and *Tychonema* beinh toxin producers (Teneva *et al.,* 2005; Shams *et al.,* 2015; Salmaso *et al*., 2016).

The samples in which cyanobacteria communities were more similar and consistent were from Lake Outlet and Upper Puntledge in May and September. It was from these sample sites that reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) were observed to compose of 94 – 99% of the cyanobacterial communities. What is interesting to note about this is that these sample sites are the primary intake and output points in Comox Lake. Additionally, these sites are located on opposite ends of Comox Lake and yet are both saturated with reads assigned to *Cyanobium* PCC-6307 (NR_102447.1). These findings provide insights into the diversity cyanobacterial communities that can be present in shallow rivers and creeks, but also the prevalence of the picocyanobacteria *Cyanobium* PCC-6307 (NR_102447.1).

While Lake Outlet and Upper Puntledge samples from the Comox Lake watershed were overly saturated with reads assigned to *Cyanobium* PCC-6307 (NR_102447.1), Boston Creek and Cruikshank River were much more diverse and complex. Various cyanobacteria genera were observed from these sites, most likely owing to these sites being shallow and dynamic, potentially having high stream flow and interactions with terrestrial environments. This was highlighted by ASVs containing sequence similarities to genera typically observed within shallower water sources, within the water/terrestrial interface and attached to rocks and sediment. While there were alignments of ASVs to *Scytonema* UTEX 2349 (NZ_ALWD00000000.1) and *Tychonema* CCAP 1459-11B (AB045897.1) which are potential geosmin and toxin producers (Smith *et al*., 2012; Salmaso *et al*., 2016; Churro *et al*., 2020), these sample sites contained a relatively low number of cyanobacteria reads, therefore any potential presence of these genera were in low abundance.

4.3 Cyanobacterial Communities from the Leech River/Sooke River Watershed

Cyanobacterial communities from the Leech River/Sooke River watershed were like that of the Pockwock Lake watershed as samples from Jarvis Lake, Weeks Lake and Deception Reservoir were saturated with reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1). From Weeks Lake in August, there were also some reads assigned to *Nostoc* PCC-73102 (NR_074317.1). Species

from this genus have been associated with the production of the taste/odour compound geosmin (Giglio *et al*., 2008; Wang *et al*., 2019; Churro *et al*., 2020). However, the BLAST result of the ASV assigned to *Nostoc* PCC-73102 (NR_074317.1) had sequence similarity (100%) to *Aulosira terrestre* FACHB-256 (JX872521.1), a freshwater cyanobacteria isolated from rivers and soils (Lukešová *et al*., 2009).

Within the Jarvis Lake and Weeks Lake samples were reads assigned to *Microcystis* PCC-7914 (no GenBank accession number). The genus *Microcystis* is known to be a common bloom former and producer of the hepatotoxin microcystin, which was first isolated in *Microcystis aeruginosa,* and can be fatal to mammals if consumed as it causes hemorrhaging of the liver (Carmichael, 1992). However, one ASV assigned to *Microcystis* PCC-7914 at the genus level was assigned to *Cyanothece aeruginosa* SAG 87.79 (KM019992.1) at the species level which has been observed to grow in acidic, cool and oligotrophic freshwater habitats (Mareš *et al*., 2019). The BLAST result of the ASV assigned to *Cyanothece aeruginosa* SAG 87.79 (KM019992.1) indicated higher sequence similarity (98.42%) to *Chlorogloea purpurea* SAG 13.99 (KM019990.1), which is a freshwater epilithic cyanobacterium (Saha *et al*., 2007; Marter *et al*., 2021).

The other ASVs assigned to *Microcystis* PCC-7914 were assigned to *Radiocystis* sp. JJ30-12 (AM710388.1) at the species level and based on phylogenetic analysis and BLAST results, these ASVs shared ≥99% sequence similarity with *Radiocystis*. These findings are significant as *Radiocystis* is a bloom forming, toxin producing cyanobacteria typically observed in tropic and subtropic regions in Brazil (Sant'Anna *et al.,* 2008; Paulino *et al*., 2017). Detecting ASVs with high sequence similarity (≥99) to *Radiocystis* may indicate climate change effects influencing their distribution and invasiveness into temperate regions in North America. Alternatively, *Radiocystis* and *Microcystis* are morphologically similar (Sant'Anna *et al.,* 2008) and observing these taxonomic discrepancies may highlight errors in classification of sequences that can exist within SILVA and BLAST.

Observing ASVs with high sequence similarity to *Radiocystis* sp. JJ30-12 (AM710388.1) in Jarvis Lake and Weeks Lake provides insight into the use of these lakes in

future water supply plans. Jarvis Lake and Weeks Lake drain into Leech River, which is planned to feed into Deception Reservoir for inter-basin transfers between Leech River and the Sooke Lake Reservoir, the primary drinking water source for the Greater Victoria Region. If potential toxin producers such as *Microcystis* and *Radiocystis* are present in these lakes, and if they proliferate, this may cause significant water quality issues, treatment costs, and potential challenges to the treatment system, which have the potential to lead to a shutdown (Emelko *et al.*, 2011; Dunlap *et al*., 2015).

4.4 Cyanobacteria Diversity Among Watersheds

Alpha diversity of cyanobacterial communities did indicate some monthly variations in community composition. However, as most samples were primarily comprised of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1), monthly variations that were observed are likely contributed to less abundant taxa or those that were labelled as uncultured. Variations in diversity could also be associated with the number of different ASVs that were assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1). This is most likely what occurred in the Pockwock Lake watershed and Leech River/Sooke River watershed samples as at least half of the cyanobacteria ASVs from these sites were assigned to these two taxa.

In the Comox Lake watershed, Lake Outlet and Upper Puntledge were dominated by sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1) and again monthly variations are likely contributed to less abundant taxa or the range of ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1). In comparison, Boston Creek and Cruikshank River were not primarily comprised of sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1). Rather, these sites were more diverse due to containing a range of benthic or epilithic taxa which were unique to these sites. These sites also contained the most diverse bacterial communities among all samples yet were excluded for diversity analysis of cyanobacterial communities due to low sequence reads. Cyanobacteria therefore did not have a strong influence on the diversity from these sites.

From beta diversity analysis of whole bacterial and cyanobacterial communities, samples generally grouped together by watershed, indicating community composition is unique to each watershed. However, based on clustering positions within the PCA plots, the Comox Lake watershed and Leech River/Sooke River watershed samples were more similar based on bacterial community composition while the Pockwock Lake watershed and Leech River/Sooke River watershed samples were more similar based on cyanobacterial community composition. These watersheds were similar in composition of cyanobacteria communities as both were primarily comprised of ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1). Comox Lake watershed samples varied from the other watersheds as they were primarily comprised of just *Cyanobium* PCC-6307 (NR_102447.1) plus several other, albeit less abundant, taxa.

4.5 Associations Between Environmental Factors and Water Quality Reducing Cyanobacteria

While available data on environmental factors from these watersheds were limited or was below the reporting detection limit (RDL), the genera that reads were assigned to do have various responses to nutrient concentrations and water temperatures. In Island Lake and Pockwock Lake, there were reads assigned to *Aphanizomenon* NIES81 (AJ293131.1) which contains species capable of growing at temperatures ranging from 10 – 25℃ and have been observed to be a bloom forming cyanobacteria in eutrophic waters (Wu *et al.*, 2010). From Boston Creek and Cruikshank River in the Comox Lake watershed, temperature and nutrient concentrations could potentially influence the presence of *Tychonema*, as this has been observed previously (Salmaso *et al*., 2016). *Tychonema* grows optimally at temperatures of 11 – 17℃, up to 25℃ and may utilize phosphorus for growth (Salmaso *et al*., 2016). Species of *Scytonema* may also benefit from warm environmental conditions as they have been observed to grow at temperatures from 4 – 40℃ and are also capable of nitrogen fixation (Giraldo-Silva *et al.*, 2020). From Jarvis Lake and Weeks Lake in the Leech River watershed, the presence of sequences assigned to *Radiocystis* sp. JJ30-12 (AM710388.1) which has been

observed to have optimal growth rates occurred at 25 – 30℃ and produces higher quantities of microcystin at 20℃ (Jacinavicius *et al.,* 2018).

For the sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1), the presence of this taxa could potentially be influenced by temperature. From a study on biomass and pigment production for *Cyanobium* sp. LEGE 06113, it was observed that optimal growth temperature was 20℃ (Pagels *et al*., 2020). While there could be influence of other factors on growth of *Cyanobium* PCC-6307 (NR_102447.1) such as nitrogen and phosphorus concentrations, this requires further studies. Temperature and nutrient concentrations may also influence growth of *Rhabdogloea smithii* SAG 47.91 (KM020002.1) though this research on optimal growth conditions for this genus is required as currently none exist.

4.6 Classification of Cyanobacteria

Of the 113 cyanobacteria ASVs identified, 79 were resolved to the genus-level using a SILVA classifier. As already briefly discussed, there were some inconsistencies in taxonomic assignments. These included ASVs assigned to *Microcystis* PCC-7914 at the genus-level assigned to either *Cyanothece aeruginosa* SAG 87.79 (KM019992.1) or *Radiocystis* sp. JJ30-12 (AM710388.1) at the species-level; *Calothrix* PCC-6303 (NC_019751.1) at the genus-level assigned to *Macrochaete psychrophila* CCALA 32 (KT336439.2) at the species-level; *Kamptonema* PCC-6407 (AM398782.1) at the genus-level assigned to *Oscillatoriales cyanobacterium* USR001 (MBRE01000011.1) at the species-level and *Cyanobium* PCC-6307 (NR_102447.1) at the genus-level and *Synechococcus* sp. LEGE 06306 (HM217052.1) at the species-level.

To improve taxonomic resolution, a phylogeny was constructed using reference sequences obtained based on SILVA classification of the ASVs. Additionally, these 79 cyanobacteria ASVs were search against sequences in NCBI by BLAST to further validate taxonomic assignments. Using a similar method as Li *et al.* (2019), taxonomy of ASVs were compared between SILVA and BLAST which identified 34 ASVs sharing at least the same genus with the top match from BLAST, 11 ASVs matching with multiple different genera but at least one with the same genus or species, and 34 ASVs that were mismatches. These

findings highlight some of the challenges in taxonomic classification of cyanobacteria. As some cyanobacteria share similar morphologies, it can be difficult to discern between genera or species simply through morphology (Hoffmann *et al.*, 2005; Komárek *et al*., 2014). This was highlighted in a study by Li *et al*. (2019) who observed that molecular (16S rRNA) data revealed the presence of certain genera that went undetected through morphological identification or were mis-identified due to indistinguishable phenotypic variations. Observing misidentified and inconsistent taxonomic assignments emphasizes the need to carefully construct phylogenies when characterizing communities and that a combination of morphological and molecular characterization can allow for improved classification of cyanobacteria (Hoffmann *et al.*, 2005; Komárek *et al*., 2014; Li *et al*., 2019).

4.6.1 Taxonomic State of *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1)

Although reads assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) saturated most samples, these genera do not have well resolved species taxonomic classification (Komárek *et al*., 2014; Komárek *et al*., 2020). The genus *Cyanobium* has species identification primarily based on morphology which can be problematic for these cyanobacteria as they are classified as picocyanobacteria due to their small cell sizes, making morphology-based classification difficult (Li *et al*., 2019; Komárek *et al*., 2020). This issue is highlighted in studies in which morphology-based identification is utilized compared to molecular characterization. For example, Li *et al*. (2019) observed no *Cyanobium* based on morphology while metabarcoding methods revealed that *Cyanobium* was present in most samples. The lack of morphological identification of *Cyanobium* was explained by these coccoid shaped cyanobacteria being misidentified as *Microcystis* (Li *et al*., 2019).

In this study, phylogenetic analysis of the ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1), with additional reference sequences, highlighted that taxonomic resolution of this genus is lacking. The phylogenetic tree contained variable branching patterns within this genus which suggests more species diversity than what currently exists in public

databases. Additionally, BLAST results did not improve taxonomic resolution of ASVs assigned to *Cyanobium* PCC-6307 (NR_102447.1). However, BLAST results did identify sequence similarities between *Cyanobium* and *Synechococcus*.

For the genus *Rhabdogloea*, it awaits modern molecular taxonomic classification and species identification that has yet to occur owing to being difficult to cultivate and, therefore, no well characterized reference strains being sequenced and available (Komárek *et al*., 2014). From phylogenetic analysis, the cluster of ASVs assigned to this genus contained many branching patterns, indicating species variation that likely exists but is not characterized. Again, BLAST results did not improve taxonomic resolution due to *Rhabdogloea smithii* SAG 47.91 (KM020002.1) being the only *Rhabdogloea* taxon in NCBI. This further emphasizes the need for improved taxonomic classification and species identification for the genus *Rhabdogloea* as other species likely exist but are not characterized.

4.7 Conclusions and Future Research

4.7.1 The Underestimated Prevalence of Picocyanobacteria in Watersheds

Sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1) were overly abundant in this study. However, the classification of these sequences in the context of a larger phylogenetic tree indicates that there is considerable resolution lacking in species delineation and more diversity than expected (Komárek *et al*., 2014; Komárek *et al*., 2020). As previously discussed, *Cyanobium* has only recently been characterized and that no reference strains of the genus *Rhabdogloea* exists (Komárek *et al*., 2014; Komárek *et al*., 2020). Due to the ubiquitous nature of these organisms, efforts should be made to determine the taxonomy and implications of these genera within water bodies in which they are observed.

4.7.2 Potential Presence of Geosmin and Microcystin Producers

There were sequences that were assigned to some well characterized geosmin and microcystin producers in this study. From the Pockwock Lake watershed, there were sequences assigned to *Aphanizomenon* NIES81 (AJ293131.1), a known geosmin and

microcystin producer which can form blooms (Wu *et al.,* 2010; Wang *et al*., 2019). Multiple potential toxin producers were observed from the Comox Lake watershed, including sequences assigned to *Scytonema,* which can produce saxitoxin (Smith *et al*., 2012), and *Tychonema* which can produce anatoxin-a (Salmaso *et al*., 2016) as well as geosmin (Churro *et al*., 2020), though contained low sequence reads. Finally, from the Leech River/Sooke River watershed, sequences were assigned to *Radiocystis* sp. JJ30-12 (AM710388.1) which has been identified as a microcystin producer (Paulino *et al*., 2017).

For the genus *Cyanobium,* currently only one species has been confirmed to contain microcystin, determined by Bláha and Maršálek (1999), and that is *Cyanobium rubescens* SAG 381 (Jakubowska and Szeląg-Wasielewska, 2015). However, it has been observed that some strains of *Cyanobium* may contain at least one of the NRPS or PKS genes, which are indicators of potential microcystin production (Genuário *et al*., 2016). These observations suggest that microcystin production by other species or strains of *Cyanobium* may be possible, including those that have yet to be characterized. To my knowledge, no literature exists on the production of taste and odour compounds in *Cyanobium*, or the production of toxins or taste and odour compounds in *Rhabdogloea* Again, this provides insights for future work on determining the potential ability of *Cyanobium* and *Rhabdogloea* to produce compounds that can reduce drinking water quality as findings in this study indicate they can be prevalent in freshwater sources.

An attempt to isolate and sequence the *mcy*E and *geo*A genes was not successful in this study and the ability of primers to capture these genes requires further validation. An alternative method to identify *mcy*E and *geo*A genes and water quality reducing cyanobacteria can involve a genomics approach. It has been demonstrated in a previous study that using shotgun metagenomics, cyanobacteria that produce taste and odour compounds and toxins can be identified (Otten *et al*., 2017). In this study, Otten *et al.* (2017) collected water samples from a drinking water reservoir, extracted DNA for shotgun sequencing and screened contigs for genes involved in the synthesis of taste and odour compounds and toxins. The contigs that contained these genes were then searched within NCBI using BLAST to identify sequence similarity to those that have been previously characterized

(Otten *et al*., 2017). This approach can provide an effective method to identify cyanobacteria that contain genes for compounds that reduce drinking water quality and potentially identify these genes in cyanobacteria they have not previously been observed in. A focus of future work should involve continuing to attempt to isolate and sequence the *mcy*E and *geo*A genes which can be accomplished using a genomics approach.

4.7.3 Biogeographic Distribution of Cyanobacteria

When considering biogeography of cyanobacteria, environmental factors such as nutrient availability and water temperature are main drivers that influence distribution within habitats at the local and regional scales, which is the thought that everything is everywhere, but the environment selects (Baas-Becking, 1934). Some cyanobacteria have been observed to have worldwide distributions with their proliferation primarily influenced by environmental factors rather than geographic region (Bonilla *et al.,* 2011), while others may have strong dispersal abilities and share high genetic similarity with strains from different geographic regions (van Gremberghe *et al*., 2011). In this present study, some cyanobacteria were observed within samples that corresponded to location within the water source, such as benthic and epilithic taxa in shallower depths compared to planktonic taxa in lakes. However, there were no clear biogeographic trends between sequences assigned to *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1). This could be attributed to these taxa potentially having the ability for dispersal or a global distribution that has gone undetected. Another explanation is that they are native to these water sources, or that they are invasive and are only now proliferating by outcompeting native cyanobacteria because of climate change-exacerbated disturbances or changes in environmental conditions (Mehnert *et al.,* 2011). To better understand biogeographic distributions of these cyanobacteria, improved taxonomic resolution and monitoring methods are needed to associate species to certain habitats and geographic regions.

4.7.4 Critical Need for Baseline Data to Understand Climate Change-Exacerbated Disturbance Impacts on Forested Watersheds and Cyanobacterial Blooms

In this study, cyanobacteria were detected in every water sample collected across multiple watersheds in two maritime ecozones in Canada. Although some cyanobacteria communities within samples were dominated by *Cyanobium* PCC-6307 (NR_102447.1) and *Rhabdogloea smithii* SAG 47.91 (KM020002.1), others were more diverse and complex. Some genera observed included cyanobacteria that have the potential to reduce water quality through the ability of producing the taste and odour compound geosmin, the toxin microcystin and being known bloom-formers, given ideal environmental conditions. To my knowledge, studies on the cyanobacterial communities in lakes within the three watersheds analyzed in this study have not been made available in published literature.

Investigations such as the one described herein provide information that is essential to understanding the habitat and ecology of cyanobacteria and the response of these microorganisms to changing environments that are impacted by anthropogenic and climate change-exacerbated disturbances. It also serves to provide a baseline study on the cyanobacteria present within these lakes which have not been the subject of such studies previously. This creates foundational knowledge of the communities from these watersheds and allows for monitoring how composition may shift due to changing environmental conditions as they can be impacted by climate change-exacerbated disturbances. The impact of this study is further emphasized within the scope of drinking water quality and security.

It is critical to understand the impacts climate change-exacerbated disturbances have on source waters as these events can result in the proliferation in cyanobacteria. This in turn provides information for researchers, policy makers and water treatment specialists to make informed decisions regarding drinking water treatment infrastructure, forested watershed protection and risk management for taste and odour events, presence of toxins and blooms (Emelko *et al*., 2011; Nunes *et al.,* 2018). Future work on cyanobacteria communities from these watersheds should involve analyzing seasonal and yearly trends of community composition, identifying environmental factors that may influence growth and identify the

genes for geosmin and microcystin production to determine if the cyanobacteria present pose a potential risk to water quality.

Bibliography

Al-Tebrineh, J., Mihali, T. K., Pomati, F., & Neilan, B. A. (2010). Detection of Saxitoxin-Producing Cyanobacteria and *Anabaena circinalis* in Environmental Water Blooms by Quantitative PCR. *Applied and Environmental Microbiology*, *76*(23), 7836–7842. https://doi.org/10.1128/AEM.00174-10

Anderson, L. E., Krkošek, W. H., Stoddart, A. K., Trueman, B. F., & Gagnon, G. A. (2017). Lake recovery through reduced sulfate deposition: A new paradigm for drinking water treatment. *Environmental Science and Technology*, *51*(3), 1414–1422. https://doi.org/10.1021/acs.est.6b04889

Apprill, A., Mcnally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, *75*(2), 129–137. https://doi.org/10.3354/ame01753

Arima, H., Horiguchi, N., Takaichi, S., Kofuji, R., Ishida, K. I., Wada, K., & Sakamoto, T. (2012). Molecular genetic and chemotaxonomic characterization of the terrestrial cyanobacterium *Nostoc commune* and its neighboring species. *FEMS Microbiology Ecology*, *79*(1), 34–45. https://doi.org/10.1111/j.1574-6941.2011.01195.x

Auffret, M., Pilote, A., Proulx, É., Proulx, D., Vandenberg, G., & Villemur, R. (2011). Establishment of a real-time PCR method for quantification of geosmin-producing *Streptomyces* spp. in recirculating aquaculture systems. *Water Research*, *45*(20), 6753–6762. https://doi.org/10.1016/j.watres.2011.10.020

Baas-Becking, L. G. M. B. (1934). Geobiologie of Inleiding Tot de Milieukunde. W.P. Van Stockum & Zoon, The Hague.

Barlak, R. (2019). Water quality assessment and proposed objectives for sooke watersheds, inlet, harbour and basin: Technical report. *Ministry of Environment and Climate Change Strategy*. Victoria, British Columbia, Canada.

Berrendero, E., Perona, E., & Mateo, P. (2011). Phenotypic variability and phylogenetic relationships of the genera *Tolypothrix* and *Calothrix* (Nostocales, Cyanobacteria) from running water. *International Journal of Systematic and Evolutionary Microbiology*, *61*(12), 3039–3051. https://doi.org/10.1099/ijs.0.027581-0

Berrendero Gómez, E., Johansen, J. R., Kaštovský, J., Bohunická, M., & Čapková, K. (2016). *Macrochaete* gen. nov. (Nostocales, Cyanobacteria), a taxon morphologically and molecularly distinct from Calothrix. *Journal of Phycology*, *52*(4), 638–655. https://doi.org/10.1111/jpy.12425

Bláha, L., & Maršálek, B. (1999). Microcystin production and toxicity of picocyanobacteria as a risk factor for drinking water treatment plants. *Algological Studies*, *92*, 95–108. https://doi.org/10.1127/algol_stud/92/1999/95

Bisanz, J.E. (2018). qiime2R: Importing QIIME2 artifacts and associated data into R sessions. https://github.com/jbisanz/qiime2R.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K. Bin, Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A. V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bonilla, S., Aubriot, L., Soares, M. C. S., González-Piana, M., Fabre, A., Huszar, V. L. M., Lürling, M., Antoniades, D., Padisák, J., & Kruk, C. (2012). What drives the distribution of the bloom-forming cyanobacteria *Planktothrix agardhii* and *Cylindrospermopsis raciborskii*? *FEMS Microbiology Ecology*, *79*(3), 594–607. https://doi.org/10.1111/j.1574-6941.2011.01242.x

Bradburn, M. J., Lewis Jr, W. M., & McCutchan Jr, J. H. (2012). Comparative adaptations of *Aphanizomenon* and *Anabaena* for nitrogen fixation under weak irradiance. *Freshwater Biology*, *57*(5), 1042–1049. https://doi.org/https://doi.org/10.1111/j.1365-2427.2012.02765.x

Brauer, V. S., Stomp, M., Rosso, C., Van Beusekom, S. A. M., Emmerich, B., Stal, L. J., & Huisman, J. (2013). Low temperature delays timing and enhances the cost of nitrogen fixation in the unicellular cyanobacterium *Cyanothece*. *ISME Journal*, *7*(11), 2105–2115. https://doi.org/10.1038/ismej.2013.103

Broady, P. A., & Ingerfeld, M. (1991). *Placoma regulare* sp. nov. (Entophysalidaceae, Cyanobacteria) from New Zealand streams. *Phycologia*, *30*(6), 547–555. https://doi.org/10.2216/i0031-8884-30-6-547.1

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High resolution sample inference from Illumina amplicon data.

Nature Methods, *13*(7), 4–5. https://doi.org/10.1038/nmeth.3869.DADA2

Cameron, E. S., Schmidt, P. J., Tremblay, B. J. M., Emelko, M. B., & Müller, K. M. (2020). To rarefy or not to rarefy: Enhancing microbial community analysis through next-generation sequencing. *BioRxiv*. https://doi.org/10.1101/2020.09.09.290049

Cameron, E.S., Tremblay, B.J.-M. 2020. mirlyn: Multiple Iterations of Rarefying for Library 515 Normalization. https://github.com/escamero/mirlyn/

Carmichael, W. W. (1992). Cyanobacteria secondary metabolites—the cyanotoxins. *Journal of Applied Bacteriology*, *72*(6), 445–459. https://doi.org/10.1111/j.1365-2672.1992.tb01858.x

Carmona-Jiménez, J., & Caro-Borrero, A. (2017). The last peri-urban rivers of the Mexico Basin: Establishment of potential reference conditions through the evaluation of ecological quality and biological indicators. *Revista Mexicana de Biodiversidad, 88*(2), 425–436. https://doi.org/10.1016/j.rmb.2017.03.019

Casero, M. C., Ballot, A., Agha, R., Quesada, A., & Cirés, S. (2014). Characterization of saxitoxin production and release and phylogeny of sxt genes in paralytic shellfish poisoning toxin-producing *Aphanizomenon gracile*. *Harmful Algae*, *37*, 28–37. https://doi.org/https://doi.org/10.1016/j.hal.2014.05.006

Chandran, A., & Mazumder, A. (2015a). Investigation on the temporal variation and source tracking of faecal bacteria in a forest dominated watershed (Comox Lake), British Columbia, Canada. *Journal of Applied Microbiology*, *119*(6), 1718–1728. https://doi.org/10.1111/jam.12969

Chandran, A., & Mazumder, A. (2015b). Pathogenic potential, genetic diversity, and population structure of *Escherichia coli* strains isolated from a forest-dominated watershed (Comox Lake) in British Columbia, Canada. *Applied and Environmental Microbiology*, *81*(5), 1788–1798. https://doi.org/10.1128/AEM.03738-14

Chapra, S. C., Boehlert, B., Fant, C., Bierman, V. J., Henderson, J., Mills, D., Mas, D. M. L., Rennels, L., Jantarasami, L., Martinich, J., Strzepek, K. M., & Paerl, H. W. (2017). Climate change impacts on harmful algal blooms in U.S. freshwaters: A screening-level assessment. *Environmental Science & Technology*, *51*, 8933–8943. ttps://doi.org/10.1021/acs.est.7b01498

Chorus, I. & Bartram, J. (1999). Toxic cyanobacteria in water: A guide to their public health consequences, monitoring and management. E & FN Spon, London, New York. CRC Press.

Chorus, I., & Welker, M. (2021). Toxic cyanobacteria in water: A guide to their public health consequences, monitoring and management (2nd ed.). CRC Press. https://doi.org/10.1201/9781003081449

Churro, C., Semedo-Aguiar, A. P., Silva, A. D., Pereira-Leal, J. B., & Leite, R. B. (2020). A novel cyanobacterial geosmin producer, revising GeoA distribution and dispersion patterns in Bacteria. *Scientific Reports*, *10*(1), 1–18. https://doi.org/10.1038/s41598-

020-64774-y

Cirés, S., & Ballot, A. (2016). A review of the phylogeny, ecology and toxin production of bloom-forming *Aphanizomenon* spp. and related species within the Nostocales (cyanobacteria). *Harmful Algae*, *54*, 21–43. https://doi.org/10.1016/j.hal.2015.09.007

Comte, K., Šabacká, M., Carré-Mlouka, A., Elster, J., & Komárek, J. (2007). Relationships between the Arctic and the Antarctic cyanobacteria; three *Phormidium*-like strains evaluated by a polyphasic approach. *FEMS Microbiology Ecology*, *59*(2), 366–376. https://doi.org/10.1111/j.1574-6941.2006.00257.x

Corbel, S., Mougin, C., & Bouaïcha, N. (2014). Cyanobacterial toxins: Modes of actions, fate in aquatic and soil ecosystems, phytotoxicity and bioaccumulation in agricultural crops. *Chemosphere*, *96*, 1–15. https://doi.org/10.1016/j.chemosphere.2013.07.056

Cotruvo, J. A. (2017). 2017 WHO guidelines for drinking water quality: First addendum to the fourth edition. *American Water Works Association*, *109*(7), 44-51.

De Senerpont Domis, L. N., Mooij, W. M., & Huisman, J. (2007). Climate-induced shifts in an experimental phytoplankton community: A mechanistic approach. *Hydrobiologia*, *584*(1), 403–413. https://doi.org/10.1007/s10750-007-0609-6

De Wever, A., Benzerara, K., Coutaud, M., Caumes, G., Poinsot, M., Skouri-Panet, F., Laurent, T., Duprat, E., & Gugger, M. (2019). Evidence of high Ca uptake by cyanobacteria forming intracellular $CaCO_3$ and impact on their growth. *Geobiology*, *17*(6), 676–690. https://doi.org/https://doi.org/10.1111/gbi.12358

Du, X., Liu, H., Yuan, L., Wang, Y., Ma, Y., Wang, R., Chen, X., Losiewicz, M. D., Guo, H., & Zhang, H. (2019). The diversity of cyanobacterial toxins on structural characterization, distribution and identification: A systematic review. *Toxins*, *11*(530), 1–34. https://doi.org/10.3390/toxins11090530

Dunlap, C. R., Sklenar, K. S., & Blake, L. J. (2015). A costly endeavor: Addressing algae problems in a water supply. *Journal - American Water Works Association*, *107*(5), E255–E262. https://doi.org/10.5942/jawwa.2015.107.0055

Dunnington, D. W., Spooner, I. S., Krkošek, W. H., Gagnon, G. A., Cornett, R. J., Kurek, J., White, C. E., Misiuk, B., & Tymstra, D. (2018). Anthropogenic activity in the Halifax region, Nova Scotia, Canada, as recorded by bulk geochemistry of lake sediments. *Lake and Reservoir Management*, *34*(4), 334–348. https://doi.org/10.1080/10402381.2018.1461715

Edgar, R. C. (2004). MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*(1), 113. https://doi.org/10.1186/1471-2105-5-113

Ellison, D., Morris, C. E., Locatelli, B., Sheil, D., Cohen, J., Murdiyarso, D., Gutierrez, V., Noordwijk, M. van, Creed, I. F., Pokorny, J., Gaveau, D., Spracklen, D. V., Tobella, A. B., Ilstedt, U., Teuling, A. J., Gebrehiwot, S. G., Sands, D. C., Muys, B., Verbist, B., Springgay E., Sugandi, Y., Sullivan C. A. (2017). Trees, forests and water: Cool insights

for a hot world. *Global Environmental Change*, *43*, 51–61. https://doi.org/10.1016/j.gloenvcha.2017.01.002

Emelko, M. B., Silins, U., Bladon, K. D., & Stone, M. (2011). Implications of land disturbance on drinking water treatability in a changing climate: Demonstrating the need for "source water supply and protection" strategies. *Water Research*, *45*(2), 461–472. https://doi.org/10.1016/j.watres.2010.08.051

Emelko, M. B., Stone, M., Silins, U., Allin, D., Collins, A. L., Williams, C. H. S., Martens, A. M., & Bladon, K. D. (2016). Sediment-phosphorus dynamics can shift aquatic ecology and cause downstream legacy effects after wildfire in large river systems. *Global Change Biology*, *22*(3), 1168–1184. https://doi.org/10.1111/gcb.13073

Engström-Öst, J., Repka, S., & Mikkonen, M. (2011). Interactions between plankton and cyanobacterium *Anabaena* with focus on salinity, growth and toxin production. *Harmful Algae*, *10*(5), 530–535. https://doi.org/10.1016/j.hal.2011.04.002

Epps, D., & Phippen, B. (2011). Water quality assessment and objectives for Comox Lake: Technical report. *Ministry of Environment and Climate Change Strategy*. Victoria, British Columbia, Canada.

Ernst, C. (2004). *Protecting the Source: Land Conservation and the Future of America's Drinking Water*. Trust for Public Land, Washington, D.C.

ESRI. (2011). ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Findlay, D. L., Hecky, R. E., Hendzel, L. L., Stainton, M. P., & Regehr, G. W. (1994). Relationship between N_2-fixation and heterocyst abundance and its relevance to the nitrogen budget of Lake 227. *Canadian Journal of Fisheries and Aquatic Sciences*, *51*(10), 2254–2266. https://doi.org/10.1139/f94-229

Fewer, D., Friedl, T., & Büdel, B. (2002). *Chroococcidiopsis* and heterocyst-differentiating cyanobacteria are each other's closest living relatives. *Molecular Phylogenetics and Evolution*, *23*(1), 82–90. https://doi.org/https://doi.org/10.1006/mpev.2001.1075

Genuário, D. B., Lorenzi, A. S., Agujaro, L. F., Isaac, R. de L., Azevedo, M. T. de P., Cantúsio Neto, R., & Fiore, M. F. (2016). Cyanobacterial community and microcystin production in a recreational reservoir with constant *Microcystis* blooms. *Hydrobiologia*, *779*(1), 105–125. https://doi.org/10.1007/s10750-016-2802-y

Giglio, S., Jiang, J., Saint, C. P., Cane, D. E., & Monis, P. T. (2008). Isolation and characterization of the gene associated with geosmin production in cyanobacteria. *Environmental Science and Technology*, *42*(21), 8027–8032. https://doi.org/10.1021/es801465w

Giglio, S., Chou, W. K. W., Ikeda, H., Cane, D. E., & Monis, P. T. (2010). Biosynthesis of 2-methylisoborneol in cyanobacteria. *Environmental Science and Technology*, *45*(3), 992–998. https://doi.org/10.1021/es102992p

Giraldo-Silva, A., Fernandes, V. M. C., Bethany, J., & Garcia-Pichel, F. (2020). Niche partitioning with temperature among heterocystous cyanobacteria (*Scytonema* spp., *Nostoc* spp., and *Tolypothrix* spp.) from biological soil crusts. *Microorganisms*, *8*(3). https://doi.org/10.3390/microorganisms8030396

Gugger, M., Lyra, C., Henriksen, P., Couté, A., Humbert, J. F., & Sivonen, K. (2002). Phylogenetic comparison of the cyanobacterial genera *Anabaena* and *Aphanizomenon*. *International Journal of Systematic and Evolutionary Microbiology*, *52*(5), 1867–1880. https://doi.org/10.1099/ijs.0.02270-0

Hašler, P., Dvořák, P., Johansen, J. R., Kitner, M., Ondřej, V., & Poulíčková, A. (2012). Morphological and molecular study of epipelic filamentous genera *Phormidium*, *Microcoleus* and *Geitlerinema* (Oscillatoriales, Cyanophyta/ cyanobacteria). *Fottea*, *12*(2), 341–356. https://doi.org/10.5507/fot.2012.024

Hauer, T., Bohunická, M., Johansen, J. R., Mareš, J., & Berrendero-Gomez, E. (2014). Reassessment of the cyanobacterial family Microchaetaceae and establishment of new families Tolypothrichaceae and Godleyaceae. *Journal of Phycology*, *50*(6), 1089–1100. https://doi.org/10.1111/jpy.12241

Heath, M. W., Wood, S. A., Brasell, K. A., Young, R. G., & Ryan, K. G. (2015). Development of habitat suitability criteria and in-stream habitat assessment for the benthic cyanobacteria *Phormidium*. *River Research and Applications*, *31*(1), 98–108. https://doi.org/https://doi.org/10.1002/rra.2722

Hoffmann, L., Komárek, J., & Kaštovský, J. (2005). System of Cyanoprokaryotes (Cyanobacteria) - State in 2004. *Algological Studies*, *117*, 95–115. https://doi.org/10.1127/1864-1318/2005/0117-0095

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hernsdorf, A. W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D. A., Finstad, K. M., Amundson, R., Thomas, B. C., & Banfield, J. F. (2016). A new view of the tree of life. *Nature Microbiology*, *1*(5), 1–6. https://doi.org/10.1038/nmicrobiol.2016.48

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

Izaguirre, G., & Taylor, W. D. (2004). A guide to geosmin- and MIB-producing cyanobacteria in the United States. *Water Science and Technology*, *49*(9), 19–24. https://doi.org/10.2166/wst.2004.0524

Jacinavicius, F. R., De Carvalho, L. R., Carneiro, R. L., & Sant'Anna, C. L. (2018). The influence of temperature on *Radiocystis fernandoi* strain (cyanobacteria) growth and microcystin production. *Revista Brasileira de Botanica*, *41*(3), 675–680. https://doi.org/10.1007/s40415-018-0490-8

Jakubowska, N., & Szeląg-Wasielewska, E. (2015). Toxic picoplanktonic cyanobacteria -

Review. *Marine Drugs*, *13*(3), 1497–1518. https://doi.org/10.3390/md13031497

Johansen, J. R., Mareš, J., Pietrasiak, N., Bohunická, M., Zima, J., Štenclová, L., & Hauer, T. (2017). Highly divergent 16S rRNA sequences in ribosomal operons of *Scytonema hyalinum* (Cyanobacteria). *PLoS ONE*, *12*(10), 1–16. https://doi.org/10.1371/journal.pone.0186393

John, N., Koehler, A. V., Ansell, B. R. E., Baker, L., Crosbie, N. D., & Jex, A. R. (2018). An improved method for PCR-based detection and routine monitoring of geosmin-producing cyanobacterial blooms. *Water Research*, *136*, 34–40. https://doi.org/10.1016/j.watres.2018.02.041

Jun, Z., Liting, C., Kun, S., & Lili, H. (2018). Assessment of different mcy genes for detecting the toxic to non-toxic *Microcystis* ratio in the field by multiplex qPCR. *Journal of Oceanology and Limnology*, *36*(4), 1132–1144.

Jung, P., Briegel-Williams, L., Schermer, M., & Büdel, B. (2019). Strong in combination: Polyphasic approach enhances arguments for cold-assigned cyanobacterial endemism. *MicrobiologyOpen*, *8*(5), 1–14. https://doi.org/10.1002/mbo3.729

Jung, P., Mikhailyuk, T., Emrich, D., Baumann, K., Dultz, S., & Büdel, B. (2020). Shifting boundaries: Ecological and geographical range extension based on three new species in the cyanobacterial genera *Cyanocohniella*, *Oculatella*, and, *Aliterella*. *Journal of Phycology*, *56*(5), 1216–1231. https://doi.org/10.1111/jpy.13025

Jungblut, A. D., & Neilan, B. A. (2006). Molecular identification and evolution of the cyclic peptide hepatotoxins, microcystin and nodularin, synthetase genes in three orders of cyanobacteria. *Archives of Microbiology*, *185*(2), 107–114. https://doi.org/10.1007/s00203-005-0073-5

Kellmann, R., Mihali, T. K., Young, J. J., Pickford, R., Pomati, F., & Neilan, B. A. (2008). Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Applied and Environmental Microbiology*, *74*(13), 4044–4053. https://doi.org/10.1128/AEM.00353-08

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, *16*(2), 111–120. https://doi.org/10.1007/BF01731581

Klotzbach, P. J., Schreck, C. J., Collins, J. M., Bell, M. M., Blake, E. S., & Roache, D. (2018). The extremely active 2017 North Atlantic hurricane season. *Monthly Weather Review*, *146*(10), 3425–3443. https://doi.org/10.1175/MWR-D-18-0078.1

Komárek, J., Genuário, D. B., Fiore, M. F., & Elster, J. (2015). Heterocytous cyanobacteria of the Ulu Peninsula, James Ross Island, Antarctica. *Polar Biology*, *38*(4), 475–492. https://doi.org/10.1007/s00300-014-1609-4

Komárek, J., Johansen, J. R., Šmarda, J., & Strunecký, O. (2020). Phylogeny and taxonomy of synechococcus–like cyanobacteria. *Fottea*, *20*(2), 171–191. https://doi.org/10.5507/fot.2020.006

Komárek, J., Kaštovský, J., Mareš, J., & Johansen, J. R. (2014). Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. *Preslia*, *86*, 295–335.

Kreer, C., Döring, M., Lehnen, N., Ercanoglu, M. S., Gieselmann, L., Luca, D., Jain, K., Schommers, P., Pfeifer, N., & Klein, F. (2020). openPrimeR for multiplex amplification of highly diverse templates. *Journal of Immunological Methods*, *480*, 112752. https://doi.org/10.1016/j.jim.2020.112752

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Kurmayer, R., Christiansen, G., Holzinger, A., & Rott, E. (2018). Single colony genetic analysis of epilithic stream algae of the genus *Chamaesiphon* spp. *Hydrobiologia*, *811*(1), 61–75. https://doi.org/10.1007/s10750-017-3295-z

Li, X., Dreher, T. W., & Li, R. (2016). An overview of diversity, occurrence, genetics and toxin production of bloom-forming *Dolichospermum* (*Anabaena*) species. *Harmful Algae*, *54*, 54–68. https://doi.org/10.1016/j.hal.2015.10.015

Li, X., Huo, S., Zhang, J., Ma, C., Xiao, Z., & Zhang, H. (2019). Metabarcoding reveals a more complex cyanobacterial community than morphological identification. *Ecological Indicators*, *107*, 105653. https://doi.org/10.1016/j.ecolind.2019.105653

Lopes, A. F., Macdonald, J. L., Quinteiro, P., Arroja, L., Carvalho-Santos, C., Cunha-e-Sá, M. A., & Dias, A. C. (2018). Surface vs. groundwater: The effect of forest cover on the costs of drinking water. *Water Resources and Economics*, *28*, 100123. https://doi.org/10.1016/j.wre.2018.06.002

Lopes, V. R., Ramos, V., Martins, A., Sousa, M., Welker, M., Antunes, A., & Vasconcelos, V. M. (2012). Phylogenetic, chemical and morphological diversity of cyanobacteria from Portuguese temperate estuaries. *Marine Environmental Research*, *73*, 7–16. https://doi.org/10.1016/j.marenvres.2011.10.005

Loza, V., Perona, E., & Mateo, P. (2013). Molecular fingerprinting of cyanobacteria from river biofilms as a water quality monitoring tool. *Applied and Environmental Microbiology*, *79*(5), 1459–1472. https://doi.org/10.1128/AEM.03351-12

Lukešová, A., Johansen, J. R., Martin, M. P., & Casamatta, D. A. (2009). *Aulosira bohemensis* sp. nov.: Further phylogenetic uncertainty at the base of the Nostocales (Cyanobacteria). *Phycologia*, *48*(2), 118–129. https://doi.org/10.2216/08-56.1

Lyra, C., Suomalainen, S., Gugger, M., Vezie, C., Sundman, P., Paulin, L., & Sivonen, K. (2001). Molecular characterization of planktic cyanobacteria of *Anabaena*, *Aphanizomenon*, *Microcystis* and *Planktothrix* genera. *International Journal of Systematic and Evolutionary Microbiology*, *51*(2), 513–526. https://doi.org/10.1099/00207713-51-2-513

Mareš, J., Johansen, J. R., Hauer, T., Zima, J., Ventura, S., Cuzman, O., Tiribilli, B., &

Kaštovský, J. (2019). Taxonomic resolution of the genus *Cyanothece* (Chroococcales, Cyanobacteria), with a treatment on *Gloeothece* and three new genera, *Crocosphaera*, *Rippkaea*, and *Zehria*. *Journal of Phycology*, *55*(3), 578–610. https://doi.org/10.1111/jpy.12853

Marquardt, J., & Palinska, K. A. (2007). Genotypic and phenotypic diversity of cyanobacteria assigned to the genus *Phormidium* (Oscillatoriales) from different habitats and geographical sites. *Archives of Microbiology*, *187*(5), 397–413. https://doi.org/10.1007/s00203-006-0204-7

Marter, P., Huang, S., Brinkmann, H., Pradella, S., Jarek, M., Rohde, M., Bunk, B., & Petersen, J. (2021). Filling the gaps in the cyanobacterial tree of life—metagenome analysis of *Stigonema ocellatum* DSM 106950, *Chlorogloea purpurea* SAG 13.99 and *Gomphosphaeria aponina* DSM 107014. *Genes*, *12*(389). https://doi.org/https://doi.org/10.3390/genes12030389

McGregor, G. B., & Sendall, B. C. (2015). Phylogeny and toxicology of *Lyngbya wollei* (Cyanobacteria, Oscillatoriales) from north-eastern Australia, with a description of *Microseira* gen. nov. *Journal of Phycology*, *51*(1), 109–119. https://doi.org/10.1111/jpy.12256

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

Mehnert, G., Leunert, F., Cirés, S., Jöhnk, K., Rücker, J., Nixdorf, B., & Wiedner, C. (2011). Competitiveness of invasive and native cyanobacteria from temperate freshwaters under various light and temperature conditions. *Journal of Plankton Research*, *32*(7), 1009–1021. https://doi.org/https://doi.org/10.1093/plankt/fbq033

Mitton, J. B., & Ferrenberg, S. M. (2012). Mountain pine beetle develops an unprecedented summer generation in response to climate warming. *The American Naturalist*, *179*(5), E163–E171. https://doi.org/10.1086/665007

Monteagudo, L., & Moreno, J. L. (2016). Benthic freshwater cyanobacteria as indicators of anthropogenic pressures. *Ecological Indicators*, *67*, 693–702. https://doi.org/10.1016/j.ecolind.2016.03.035

Müller, K. M., Chhun, A., Guildford, S. J., Yakobowski, S. J., & Jonlija, M. (2017). Molecular and ecological characterization of toxic cyanobacteria from the Bay of Quinte (Lake Ontario) and Maumee Bay (Lake Erie). *Journal of Great Lakes Research*, *43*(6), 1067–1083. https://doi.org/10.1016/j.jglr.2017.03.021

Nakamura, Y., Kaneko, T., Sato, S., Mimuro, M., Miyashita, H., Tsuchiya, T., Sasamoto, S., Watanabe, A., Kawashima, K., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Nakazaki, N., Shimpo, S., Takeuchi, C., Yamada, M., & Tabata, S. (2003). Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. *DNA Research*, *10*(4), 137–145. https://doi.org/10.1093/dnares/10.4.137

Norris, T. B., & Castenholz, R. W. (2006). Endolithic photosynthetic communities within ancient and recent travertine deposits in Yellowstone National Park. *FEMS Microbiology Ecology*, *57*(3), 470–483. https://doi.org/10.1111/j.1574-6941.2006.00134.x

Nunes, J. P., Doerr, S. H., Sheridan, G., Neris, J., Santín, C., Emelko, M. B., Silins, U., Robichaud, P. R., Elliot, W. J., & Keizer, J. (2018). Assessing water contamination risk from vegetation fires: Challenges, opportunities and a framework for progress. *Hydrological Processes*, *32*(5), 687–694. https://doi.org/10.1002/hyp.11434

Obuekwe, I. S., Vaz, M. G. M. V., Genuário, D. B., Castro, N. V., Almeida, A. V. M., Veloso, R. W., Pinto, G. N., Alvarenga, L. V., Mello, J. V., Nunes-Nesi, A., & Araújo, W. L. (2019). Arsenic-contaminated sediment from mining areas as source of morphological and phylogenetic distinct cyanobacterial lineages. *Algal Research*, *42*, 101589. https://doi.org/10.1016/j.algal.2019.101589

Otten, T. G., Graham, J. L., Harris, T. D., & Dreher, T. W. (2016). Elucidation of tasteand odor-producing bacteria and toxigenic cyanobacteria in a midwestern drinking water supply reservoir by shotgun metagenomic analysis. *Applied and Environmental Microbiology*, *82*(17), 5410–5420. https://doi.org/10.1128/AEM.01334-16

Paerl, H. W. (2018). Mitigating toxic planktonic cyanobacterial blooms in aquatic ecosystems facing increasing anthropogenic and climatic pressures. *Toxins*, 1–16. https://doi.org/10.3390/toxins10020076

Paerl, H. W., Hall, N. S., Hounshell, A. G., Luettich, R. A., Rossignol, K. L., Osburn, C. L., & Bales, J. (2019). Recent increase in catastrophic tropical cyclone flooding in coastal North Carolina, USA: Long-term observations suggest a regime shift. *Scientific Reports*. https://doi.org/10.1038/s41598-019-46928-9

Paerl, H. W., Scott, J. T., McCarthy, M. J., Newell, S. E., Gardner, W. S., Havens, K. E., Hoffman, D. K., Wilhelm, S. W., & Wurtsbaugh, W. A. (2016). It takes two to tango: When and where dual nutrient (N & P) reductions are needed to protect lakes and downstream ecosystems. *Environmental Science and Technology*, *50*(20), 10805–10813. https://doi.org/10.1021/acs.est.6b02575

Pagels, F., Salvaterra, D., Amaro, H. M., Lopes, G., Sousa-Pinto, I., Vasconcelos, V., & Guedes, A. C. (2020). Factorial optimization of upstream process for *Cyanobium* sp. pigments production. *Journal of Applied Phycology*, *32*(6), 3861–3872. https://doi.org/10.1007/s10811-020-02260-8

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/https://doi.org/10.1111/1462-2920.13023

Paulino, M. G., Rossi, P. A., Venturini, F. P., Tavares, D., Elisabete da Silva Souza, N., Sakuragui, M. M., Moraes, G., Terezan, A. P., Fernandes, J. B., Giani, A., & Fernandes,

M. N. (2017). Hepatotoxicity and metabolic effects of cellular extract of cyanobacterium *Radiocystis fernandoi* containing microcystins RR and YR on neotropical fish (*Hoplias malabaricus*). *Chemosphere*, *175*, 431–439. https://doi.org/10.1016/j.chemosphere.2017.02.089

Pearson, L. A., Dittmann, E., Mazmouz, R., Ongley, S. E., D'Agostino, P. M., & Neilan, B. A. (2016). The genetics, biosynthesis and regulation of toxic specialized metabolites of cyanobacteria. *Harmful Algae*, *54*, 98–111. https://doi.org/10.1016/j.hal.2015.11.002

Phlips, E. J., Badylak, S., Nelson, N. G., & Havens, K. E. (2020). Hurricanes, El Niño and harmful algal blooms in two sub-tropical Florida estuaries: Direct and indirect impacts. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-58771-4

Postel, S. L., & Thompson, B. H. (2005). Watershed protection: Capturing the benefits of nature's water supply services. *Natural Resources Forum*, *29*(2), 98–108. https://doi.org/10.1111/j.1477-8947.2005.00119.x

Price, J. I., Renzetti, S., Dupont, D., Adamowicz, W., & Emelko, M. B. (2017). Production costs, inefficiency, and source water quality: A stochastic cost frontier analysis of canadian water utilities. *Land Economics*. https://doi.org/10.3368/le.93.1.1

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

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Raabová, L., Kovacik, L., Elster, J., & Strunecký, O. (2019). Review of the genus *Phormidesmis* (Cyanobacteria) based on environmental, morphological, and molecular data with description of a new genus *Leptodesmis*. *Phytotaxa*, *395*(1), 1–16. https://doi.org/10.11646/phytotaxa.395.1.1

Rajaniemi-Wacklin, P., Rantala, A., Mugnai, M. A., Turicchia, S., Ventura, S., Komárková, J., Lepistö, L., & Sivonen, K. (2006). Correspondence between phylogeny and morphology of *Snowella* spp. and *Woronichinia naegeliana*, cyanobacteria commonly occurring in lakes. *Journal of Phycology*, *42*(1), 226–232. https://doi.org/https://doi.org/10.1111/j.1529-8817.2006.00179.x

Richardson, J., Feuchtmayr, H., Miller, C., Hunter, P. D., Maberly, S. C., & Carvalho, L. (2019). Response of cyanobacteria and phytoplankton abundance to warming, extreme rainfall events and nutrient enrichment. *Global Change Biology*, *25*, 3365–3380. https://doi.org/10.1111/gcb.14701

Rigonato, J., Alvarenga, D. O., Branco, L. H. Z., Varani, A. M., Brandini, F. P., & Fiore, M. F. (2016a). Draft genome sequence of a novel culturable marine chroococcalean cyanobacterium from the South Atlantic Ocean. *Genome Announcements*, *3*(2), 3–4.

https://doi.org/10.1128/genomeA.00384-15

Rigonato, J., Gama, W. A., Alvarenga, D. O., Branco, L. H. Z., Brandini, F. P., Genuário, D. B., & Fiore, M. F. (2016b). *Aliterella atlantica* gen. nov., sp. nov., and *Aliterella antarctica* sp. nov., novel members of coccoid Cyanobacteria. *International Journal of Systematic and Evolutionary Microbiology*, *66*(8), 2853–2861. https://doi.org/10.1099/ijsem.0.001066

Robinne, F. N., Bladon, K. D., Silins, U., Emelko, M. B., Flannigan, M. D., Parisien, M. A., Wang, X., Kienzle, S. W., & Dupont, D. P. (2019). A regional-scale index for assessing the exposure of drinking-water sources to wildfires. *Forests*, *10*(5), 1–21. https://doi.org/10.3390/f10050384

Rudi, K., Skulberg, O. M., Larsen, F., & Jakobsen, K. S. (1997). Strain characterization and classification of oxyphotobacteria in clone cultures on the basis of 16S rRNA sequences from the variable regions V6, V7, and V8. *Applied and Environmental Microbiology*, *63*(7), 2593–2599. https://doi.org/10.1128/aem.63.7.2593-2599.1997

Saha, S. K., Das, R., Bora, K. N., & Uma, L. (2007). Biodiversity of epilithic cyanobacteria from freshwater streams of Kakoijana reserve forest, Assam, India. *Indian Journal of Microbiology*, *47*(3), 219–232. https://doi.org/10.1007/s12088-007-0043-5

Salmaso, N., Cerasino, L., Boscaini, A., & Capelli, C. (2016). Planktic *Tychonema* (cyanobacteria) in the large lakes south of the alps: Phylogenetic assessment and toxigenic potential. *FEMS Microbiology Ecology*, *92*(10), 1–14. https://doi.org/10.1093/femsec/fiw155

Sant'Anna, C. L. A., Teresade P. Werner, V. R., Dogo, C. R., & Rios, F. R. de C. (2008). Review of toxic species of Cyanobacteria in Brazil. *Algological Studies*, *126*(April), 251–265. https://doi.org/10.1127/1864-1318/2008/0126-0251

Saw, J. H. W., Schatz, M., Brown, M. V., Kunkel, D. D., Foster, J. S., Shick, H., Christensen, S., Hou, S., Wan, X., & Donachie, S. P. (2013). Cultivation and complete genome sequencing of *Gloeobacter kilaueensis* sp. nov., from a lava cave in Kīlauea Caldera, Hawai'i. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0076376

Schindler, D. W. (1977). Evolution of phosphorus limitation in lakes: Natural mechanisms compensate for deficiencies of nitrogen and carbon in eutrophied lakes. *Science,* 195, 260–262.

Schindler, D. W., Hecky, R. E., Findlay, D. L., Stainton, M. P., Parker, B. R., Paterson, M. J., Beaty, K. G., Lyng, M., & Kasian, S. E. M. (2008). Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11254–11258. https://doi.org/10.1073/pnas.0805108105

Schirrmeister, B. E., Gugger, M., & Donoghue, P. C. J. (2015). Cyanobacteria and the Great Oxidation Event: Evidence from genes and fossils. *Palaeontology*, *58*(5), 769–785. https://doi.org/10.1111/pala.12178

Shahi, S. K., Freedman, S. N., & Mangalam, A. K. (2017). Gut microbiome in multiple sclerosis: The players involved and the roles they play. *Gut Microbes*, *8*(6), 607–615. https://doi.org/10.1080/19490976.2017.1349041

Shams, S., Capelli, C., Cerasino, L., Ballot, A., Dietrich, D. R., Sivonen, K., & Salmaso, N. (2015). Anatoxin-a producing *Tychonema* (Cyanobacteria) in European waterbodies. *Water Research*, *69*, 68–79. https://doi.org/10.1016/j.watres.2014.11.006

Shishido, T. K., Jokela, J., Fewer, D. P., Wahlsten, M., Fiore, M. F., & Sivonen, K. (2017). Simultaneous production of anabaenopeptins and namalides by the cyanobacterium *Nostoc* sp. CENA543. *ACS Chemical Biology*, *12*(11), 2746–2755. https://doi.org/10.1021/acschembio.7b00570

Sihvonen, L. M., Lyra, C., Fewer, D. P., Rajaniemi-Wacklin, P., Lehtimäki, J. M., Wahlsten, M., & Sivonen, K. (2007). Strains of the cyanobacterial genera *Calothrix* and *Rivularia* isolated from the Baltic Sea display cryptic diversity and are distantly related to *Gloeotrichia* and *Tolypothrix*. *FEMS Microbiology Ecology*, *61*(1), 74–84. https://doi.org/10.1111/j.1574-6941.2007.00321.x

Silins, U., Bladon, K. D., Kelly, E. N., Esch, E., Spence, J. R., Stone, M., Emelko, M. B., Boon, S., Wagner, M. J., Williams, C. H. S., & Tichkowsky, I. (2014). Five-year legacy of wildfire and salvage logging impacts on nutrient runoff and aquatic plant, invertebrate, and fish productivity. *Ecohydrology*, *7*(6), 1508–1523. https://doi.org/10.1002/eco.1474

Silins, U., Stone, M., Emelko, M. B., & Bladon, K. D. (2009). Sediment production following severe wildfire and post-fire salvage logging in the Rocky Mountain headwaters of the Oldman River Basin, Alberta. *Catena*, *79*(3), 189–197. https://doi.org/10.1016/j.catena.2009.04.001

Singh, Y., Gulati, A., Singh, D. P., & Khattar, J. I. S. (2018). Cyanobacterial community structure in hot water springs of Indian North-Western Himalayas: A morphological, molecular and ecological approach. *Algal Research*, *29*, 179–192. https://doi.org/10.1016/j.algal.2017.11.023

Skwaruk, J. S., Emelko, M. B., Silins, U., & Stone, M. (2020). Treatment of severely-deteriorated post-fire runoff: A comparison of conventional and high-rate clarification to demonstrate key drinking water treatment capabilities and challenges. *ChemRxiv*, 1–35. https://doi.org/10.26434/chemrxiv.13350785.v1

Smith, F. M. J., Wood, S. A., Wilks, T., Kelly, D., Broady, P. A., Williamson, W., & Gaw, S. (2012). Survey of *Scytonema* (cyanobacteria) and associated saxitoxins in the littoral zone of recreational lakes in Canterbury, New Zealand. *Phycologia*, *51*(5), 542–551. https://doi.org/10.2216/11-84.1

Stoyanov, P., Moten, D., Mladenov, R., Dzhambazov, B., & Teneva, I. (2014). Phylogenetic relationships of some filamentous cyanoprokaryotic species. *Evolutionary Bioinformatics*, *10*, 39–49. https://doi.org/10.4137/EBo.s13748

Strunecký, O., Komárek, J., & Šmarda, J. (2014). *Kamptonema* (Microcoleaceae, Cyanobacteria), a new genus derived from the polyphyletic *Phormidium* on the basis of combined molecular and cytomorphological markers. *Preslia*, *86*(2), 193–207.

Strunecký, O., Raabová, L., Bernardova, A., Ivanova, A. P., Semanova, A., Crossley, J., & Kaftan, D. (2020). Diversity of cyanobacteria at the Alaska North Slope with description of two new genera: *Gibliniella* and *Shackletoniella*. *FEMS Microbiology Ecology*, *96*(3). https://doi.org/10.1093/femsec/fiz189

Suda, S., Watanabe, M. M., Otsuka, S., Mahakahant, A., Yongmanitchai, W., Nopartnaraporn, N., Liu, Y., & Day, J. G. (2002). Taxonomic revision of water-bloom-forming species of oscillatorioid cyanobacteria. *International Journal of Systematic and Evolutionary Microbiology*, *52*(5), 1577–1595. https://doi.org/10.1016/j.arth.2005.03.039

Sugita, C., Ogata, K., Shikata, M., Jikuya, H., Takano, J., Furumichi, M., Kanehisa, M., Omata, T., Sugiura, M., & Sugita, M. (2007). Complete nucleotide sequence of the freshwater unicellular cyanobacterium *Synechococcus elongatus* PCC 6301 chromosome: Gene content and organization. *Photosynthesis Research*, *93*(1–3), 55–67. https://doi.org/10.1007/s11120-006-9122-4

Suurnäkki, S., Gomez-Saez, G. V., Rantala-Ylinen, A., Jokela, J., Fewer, D. P., & Sivonen, K. (2015). Identification of geosmin and 2-methylisoborneol in cyanobacteria and molecular detection methods for the producers of these compounds. *Water Research*, *68*(9), 56–66. https://doi.org/10.1016/j.watres.2014.09.037

Talucci, A. C., & Krawchuk, M. A. (2019). Dead forests burning: The influence of beetle outbreaks on fire severity and legacy structure in sub-boreal forests. *Ecosphere*, *10*(5), 1–17. https://doi.org/10.1002/ecs2.2744

Taton, A., Grubisic, S., Ertz, D., Hodgson, D. A., Piccardi, R., Biondi, N., Tredici, M. R., Mainini, M., Losi, D., Marinelli, F., & Wilmotte, A. (2006). Polyphasic study of antarctic cyanobacterial strains. *Journal of Phycology*, *42*(6), 1257–1270. https://doi.org/10.1111/j.1529-8817.2006.00278.x

Te, S. H., Tan, B. F., Thompson, J. R., & Gin, K. Y. H. (2016). Draft genome sequences of two benthic cyanobacteria, *Oscillatoriales* USR 001 and *Nostoc* sp. MBR 210, isolated from tropical freshwater lakes. *Genome Announcements*, *4*(5), 4–5. https://doi.org/10.1128/genomeA.01115-16

Teneva, I., Dzhambazov, B., Koleva, L., Mladenov, R., & Schirmer, K. (2005). Toxic potential of five freshwater Phormidium species (Cyanoprokaryota). *Toxicon*, *45*(6), 711–725. https://doi.org/https://doi.org/10.1016/j.toxicon.2005.01.018

Tomitani, A., Knoll, A. H., Cavanaugh, C. M., & Ohno, T. (2006). The evolutionary diversification of cyanobacteria: Molecular-phylogenetic and paleontological perspectives. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(14), 5442–5447. https://doi.org/10.1073/pnas.0600999103

Tropea, A. E., Ginn, B. K., Cumming, B. F., & Smol, J. P. (2007). Tracking long-term acidification trends in pockwock lake (Halifax, Nova Scotia), the water supply for a major eastern Canadian city. *Lake and Reservoir Management*, *23*(3), 279–286. https://doi.org/10.1080/07438140709354016

Turner, S. (1997). Molecular systematics of oxygenic photosynthetic bacteria. In D. Bhattacharya (Ed.), *Origins of Algae and their Plastids* (pp. 13–52). https://doi.org/10.1007/978-3-7091-6542-3_2

Turner, S., Pryer, K. M., Miao, V. P. W., & Palmer, J. D. (1999). Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *Journal of Eukaryotic Microbiology*, *46*(4), 327–338. https://doi.org/10.1111/j.1550-7408.1999.tb04612.x

van Gremberghe, I., Leliaert, F., Mergeay, J., Vanormelingen, P., van der Gucht, K., Debeer, A. E., Lacerot, G., de Meester, L., & Vyverman, W. (2011). Lack of phylogeographic structure in the freshwater cyanobacterium *Microcystis aeruginosa* suggests global dispersal. *PLoS ONE*, *6*(5), e19561. https://doi.org/10.1371/journal.pone.0019561

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

Wagner, M. J., Bladon, K. D., Silins, U., Williams, C. H. S., Martens, A. M., Boon, S., MacDonald, R. J., Stone, M., Emelko, M. B., & Anderson, A. (2014). Catchment-scale stream temperature response to land disturbance by wildfire governed by surface-subsurface energy exchange and atmospheric controls. *Journal of Hydrology*, *517*, 328–338. https://doi.org/10.1016/j.jhydrol.2014.05.006

Walter, J. M., Coutinho, F. H., Dutilh, B. E., Swings, J., Thompson, F. L., & Thompson, C. C. (2017). Ecogenomics and taxonomy of cyanobacteria phylum. *Frontiers in Microbiology*, *8*, 1–18. https://doi.org/10.3389/fmicb.2017.02132

Walters, W. A., Caporaso, J. G., Lauber, C. L., Berg-Lyons, D., Fierer, N., & Knight, R. (2011). PrimerProspector: De novo design and taxonomic analysis of barcoded polymerase chain reaction primers. *Bioinformatics*, *27*(8), 1159–1161. https://doi.org/10.1093/bioinformatics/btr087

Walters, W., Hyde, E. R., Berg-lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J. A., Jansson, J. K., Caporaso, G. J., Fuhrman, J. A., Apprill, A., & Knight, R. (2015). Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *American Society for Microbiology*, *1*(1), e0009-15. https://doi.org/10.1128/mSystems.00009-15

Wang, Z., Shao, J., Xu, Y., Yan, B., & Li, R. (2015). Genetic basis for geosmin production by the water bloom-forming cyanobacterium, *Anabaena ucrainica*. *Water (Switzerland)*, *7*(1), 175–187. https://doi.org/10.3390/w7010175

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*(7), 789–796. https://doi.org/10.1016/j.scitotenv.2019.06.468

Williams, C. H. S., Silins, U., Spencer, S. A., Wagner, M. J., Stone, M., & Emelko, M. B. (2019). Net precipitation in burned and unburned subalpine forest stands after wildfire in the northern Rocky Mountains. *International Journal of Wildland Fire*, *28*(10), 750–760. https://doi.org/10.1071/WF18181

Wu, W., Li, G., Li, D., & Liu, Y. (2010). Temperature may be the dominating factor on the alternant succession of *Aphanizomenon flos-aquae* and *Microcystis aeruginosa* in Dianchi Lake. *Fresenius Environmental Bulletin*, *19*(5), 846–853.

Yang, B., Wang, Y., & Qian, P.-Y. (2016). Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics*, *17*(135), 1–8. https://doi.org/10.1186/s12859-016-0992-y

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K. H., Whitman, W. B., Euzéby, J., Amann, R., & Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, *12*(9), 635–645. https://doi.org/10.1038/nrmicro3330

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134. https://doi.org/10.1186/1471-2105-13-134

Yema, L., Litchman, E., & de Tezanos Pinto, P. (2016). The role of heterocytes in the physiology and ecology of bloom-forming harmful cyanobacteria. *Harmful Algae*, *60*, 131–138. https://doi.org/https://doi.org/10.1016/j.hal.2016.11.007

Zhou, J., He, Z., Yang, Y., Deng, Y., Tringe, S., & Alvarez-Cohen, L. (2015). High-throughput metagenomic technologies for complex microbial community analysis: Open and closed formats. *MBio*, *6*(1), e02288-14. https://doi.org/10.1128/mBio.02288-14

Zilius, M., Bartoli, M., Bresciani, M., Katarzyte, M., Ruginis, T., Petkuviene, J., Lubiene, I., Giardino, C., Bukaveckas, P. A., de Wit, R., & Razinkovas-Baziukas, A. (2014). Feedback mechanisms between cyanobacterial blooms, transient hypoxia, and benthic phosphorus regeneration in shallow coastal environments. *Estuaries and Coasts*, *37*(3), 680–694. https://doi.org/10.1007/s12237-013-9717-x

Appendix A

Supplementary Tables

Supplementary Table 1 Reference sequences for *geo*A primer design**.**

Accession No.	Name
AB981724.1	Streptomyces_cyaneogriseus_subsp._noncyanogenus_scya_02397_gene_for_terpene_synthase_complete_cds
AL939126.1	Streptomyces_coelicolor_A3(2) complete_genome_segment_23/29
AP018178.1	Calothrix_sp._NIES-2100_DNA_nearly_complete_genome
AP018186.1	Nostoc_sp._NIES-2111_plasmid_plasmid2_DNA_complete_genome
AP018248.1	Tolypothrix_tenuis_PCC_7101_DNA_nearly_complete_genome
AP018307.1	Aulosira_laxa_NIES-50_DNA_nearly_complete_genome
AP018517.1	Streptomyces_rochei_7434AN4_DNA_complete_genome
AP019621.1	Streptomyces_avermitilis_MC3_DNA_complete_genome
BA000030.4	Streptomyces_avermitilis_MA-4680___NBRC_14893_DNA_complete_genome
CP001037.1	Nostoc_punctiforme_PCC_73102_strain_ATCC_29133_chromosome_complete_genome
CP003275.1	Streptomyces_hygroscopicus_subsp._jinggangensis_5008_complete_genome
CP003642.1	Cylindrospermum_stagnale_PCC_7417_complete_genome
CP003720.1	Streptomyces_hygroscopicus_subsp._jinggangensis_TL01_complete_genome
CP003943.1	Calothrix_sp._PCC_7507_chromosome_complete_genome
CP006259.1	Streptomyces_collinus_Tu_365_chromosome_complete_genome
CP009124.1	Streptomyces_lividans_TK24_complete_genome
CP009438.1	Streptomyces_glaucescens_strain_GLA.O_complete_genome
CP009754.1	Streptomyces_sp._CCM_MD2014_chromosome_complete_genome
CP010849.1	Streptomyces_cyaneogriseus_subsp._noncyanogenus_strain_NMWT_1_complete_genome
CP011497.1	Streptomyces_incarnatus_strain_NRRL_8089_sequence
CP011799.1	Streptomyces_sp._PBH53_genome
CP012382.1	Streptomyces_ambofaciens_ATCC_23877_complete_genome
CP012949.1	Streptomyces_ambofaciens_strain_DSM_40697_complete_genome

CP013142.1 Streptomyces_sp._4F_complete_genome
CP013219.1 Streptomyces_hygroscopicus_subsp._limoneus_strain_KCTC_1717_chromosome_I_complete_sequence
CP013743.1 Streptomyces_sp._CdTB01_complete_genome
CP015098.1 Streptomyces_sp._S10(2016) complete_genome
CP015588.1 Streptomyces_alfalfae_strain_ACCC40021_chromosome_complete_genome
CP015849.1 Streptomyces_sp._SAT1_complete_genome
CP015866.1 Streptomyces_parvulus_strain_2297_complete_genome
CP016438.1 Streptomyces_lincolnensis_strain_NRRL_2936_complete_genome
CP016795.1 Streptomyces_olivaceus_strain_KLBMP_5084_chromosome_complete_genome
CP017248.1 Streptomyces_fodineus_strain_TW1S1_chromosome_complete_genome
CP019724.1 Streptomyces_pactum_strain_ACT12_complete_genome
CP021080.1 Streptomyces_pluripotens_strain_MUSC_135_chromosome_complete_genome
CP021978.1 Streptomyces_hawaiiensis_strain_ATCC_12236_chromosome_complete_genome
CP022310.1 Streptomyces_asterosporus_strain_DSM_41452_chromosome_complete_genome
CP022433.1 Streptomyces_pluripotens_strain_MUSC_137_complete_genome
CP022744.1 Streptomyces_lincolnensis_strain_LC-G_chromosome_complete_genome
CP023407.1 Streptomyces_fungicidicus_strain_TXX3120_chromosome_complete_genome
CP023689.1 Streptomyces_chartreusis_strain_ATCC_14922_chromosome_complete_genome
CP023694.1 Streptomyces_coeruleorubidus_strain_ATCC_13740_chromosome_complete_genome
CP023695.1 Streptomyces_alboniger_strain_ATCC_12461_chromosome_complete_genome
CP023697.1 Streptomyces_prasinus_strain_ATCC_13879_chromosome_complete_genome
CP023703.1 Streptomyces_galilaeus_strain_ATCC_14969_chromosome_complete_genome
CP026121.1 Streptomyces_sp._Go-475_chromosome_complete_genome
CP026652.1 Streptomyces_dengpaensis_strain_XZHG99_chromosome_complete_genome
CP026681.1 Nostoc_sp._Peltigera_membranacea_cyanobiont_N6_chromosome_complete_genome
CP026730.1 Streptomyces_sp._CB09001_chromosome_complete_genome
CP027297.1 Streptomyces_sp._SGAir0924_chromosome_complete_genome
CP028369.1 Streptomyces_sp._P3_chromosome_complete_genome
CP028719.1 Streptomyces_sp._endophyte_N2_chromosome_complete_genome
CP028834.1 Streptomyces_sp._M2_chromosome_complete_genome

CP029043.1 Streptomyces_nigra_strain_452_chromosome_complete_genome
CP029078.1 Streptomyces_griseoviridis_strain_K61_chromosome_complete_genome
CP029601.1 Streptomyces_sp._WAC_01438_chromosome_complete_genome
CP029617.1 Streptomyces_sp._WAC_01529_chromosome_complete_genome
CP029624.1 Streptomyces_sp._ETH9427_chromosome
CP029788.1 Streptomyces_actuosus_strain_ATCC_25421_chromosome_complete_genome
CP030073.1 Streptomyces_sp._ZFG47_chromosome_complete_genome
CP030118.1 Brasilonema_sennae_CENA114_chromosome
CP030121.1 Brasilonema_octagenarum_UFV-E1_chromosome
CP031969.1 Streptomyces_sp._CC0208_chromosome_complete_genome
CP032229.1 Streptomyces_seoulensis_strain_KCTC_9819_chromosome_complete_genome
CP032266.1 Streptomyces_fradiae_strain_NKZ-259_chromosome_complete_genome
CP032427.1 Streptomyces_griseorubiginosus_strain_3E-1_chromosome_complete_genome
CP033073.1 Streptomyces_sp._Z022_chromosome_complete_genome
CP034353.1 Streptomyces_sp._KPB2_chromosome_complete_genome
CP034463.1 Streptomyces_aquilus_strain_GGCR-6_chromosome_complete_genome
CP034539.1 Streptomyces_sp._MK-45_chromosome_complete_genome
CP034687.1 Streptomyces_griseoviridis_strain_F1-27_chromosome_complete_genome
CP036534.1 Streptomyces_sp._VN1_chromosome_complete_genome
CP039123.1 Streptomyces_sp._SS52_chromosome_complete_genome
CP040941.1 Streptomyces_variabilis_strain_ARRS001_chromosome
CP041168.1 Streptomyces_griseorubiginosus_strain_BTU6_chromosome_complete_genome
CP041602.2 Streptomyces_sp._RLB3-6_chromosome
CP041604.2 Streptomyces_sp._S1A1-7_chromosome
CP041607.2 Streptomyces_sp._S1D4-14_chromosome_complete_genome
CP041609.2 Streptomyces_sp._S1D4-20_chromosome
CP041610.2 Streptomyces_sp._RLB3-17_chromosome_complete_genome
CP041611.1 Streptomyces_sp._S1A1-3_chromosome
CP041612.2 Streptomyces_sp._S1A1-8_chromosome_complete_genome
CP041613.2 Streptomyces_sp._S1D4-23_chromosome

CP041650.2	Streptomyces_sp._RLB1-8_chromosome_complete_genome
CP041651.1	Streptomyces_sp._RLB3-5_chromosome
CP041654.1	Streptomyces_sp._RLB1-9_chromosome_complete_genome
CP042278.1	Streptomyces_sp._WAC6273_substr. delta_orf15_pCRISPR-Cas9_chromosome_complete_genome
CP042324.1	Streptomyces_coelicolor_A3(2) strain_CFB_NBC_0001_chromosome_complete_genome
CP042594.1	Streptomyces_albogriseolus_strain_LBX-2_chromosome_complete_genome
CP043959.1	Streptomyces_tendae_strain_139_chromosome_complete_genome
CP045547.1	Streptomyces_sp._SYP-A7193_chromosome_complete_genome
CP045643.1	Streptomyces_sp._QMT-28_chromosome_complete_genome
CP045740.1	Streptomyces_sp._SUK_48_chromosome_complete_genome
CP046703.1	Nostoc_sp._ATCC_53789_chromosome_complete_genome
CP047020.1	Streptomyces_sp._T44_chromosome
CP047144.1	Streptomyces_sp._HF10_chromosome_complete_genome
CP049780.1	Streptomyces_sp._JB150_chromosome
CP050504.1	Streptomyces_sp._DSM_40868_chromosome_complete_genome
CP050692.1	Streptomyces_antibioticus_strain_DSM_41481_chromosome_complete_genome
CP050975.1	Streptomyces_sp._RPA4-2_chromosome_complete_genome
CP051010.1	Streptomyces_sp._S1D4-11_chromosome_complete_genome
CP053109.1	Streptomyces_sp._Z423-1_chromosome
CP053189.1	Streptomyces_sp._jing01_chromosome_complete_genome
FJ010202.1	Nostoc_punctiforme_PCC_73102_NJ2_protein_gene_partial_cds
FJ010203.1	Nostoc_punctiforme_PCC_73102_NPUNMOD_protein_gene_complete_cds
HE971709.1	Streptomyces_davawensis_strain_JCM_4913_complete_genome
HQ404996.1	Anabaena_ucrainica_CHAB1432_geosmin_synthesis_operon_complete_sequence
HQ404997.1	Anabaena_ucrainica_CHAB2155_geosmin_synthesis_operon_complete_sequence
JX962775.1	Oscillatoria_sp._PCC_6506_putative_geosmin_synthase_(geoL)_gene_complete_cds
JX966093.1	Streptomyces_fradiae_strain_HX_putative_germacradienol_synthase_(geoA)_gene_partial_cds
KF170339.1	Streptomyces_ansochromogenes_clone_terp3_metabolite_biosynthetic_gene_cluster_complete_sequence
KJ658370.1	Nostoc_sp._UK4_geosmin_synthase_(geoA)_gene_partial_cds
KJ658373.1	Oscillatoria_sp._327/2_geosmin_synthase_(geoA)_gene_partial_cds

LC331271.1	Coelosphaerium_sp._G2_geoA_gene_for_geosmin_synthase_complete_cds
LC331272.1	Coelosphaerium_sp._G3_geoA_gene_for_geosmin_synthase_complete_cds
LN997842.1	Streptomyces_reticuli_genome_assembly_TUE45_chromosome_
LT629768.1	Streptomyces_sp._2114.2_genome_assembly_chromosome
LT670819.1	Streptomyces_sp._3124.6_genome_assembly_chromosome
LT962942.1	Streptomyces_chartreusis_NRRL_3882_isolate_NRRL3882_genome_assembly_chromosome
LT963352.1	Streptomyces_chartreusis_NRRL_3882_isolate_NRRL3882_genome_assembly_chromosome
MK213943.1	Anabaena_minutissima_FACHB_250_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213944.1	Nostoc_commune_FACHB_261_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213945.1	Lyngbya_kuetzingii_FACHB_388_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213947.1	Dolichospermum_ucrainicum_CHAB1434_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213948.1	Aphanizomenon_sp._CHAB_1684_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213949.1	Calothrix_sp._CHAB_2384_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213951.1	Anabaena_circinalis_CHAB_3585_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213952.1	Scytonema_sp._CHAB_3651_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213953.1	Phormidium_sp._D6_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213954.1	Anabaena_planctonica_SDZ-1_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213955.1	Nodularia_sp._Su-A_putative_geosmin_synthase_(geo)_gene_complete_cds
MK213957.1	Cylindrospermum_sp._CHAB_2115_putative_geosmin_synthase_(geo)_gene_partial_cds
MN708236.1	Nostoc_sp._UIC10630_BGC2_biosynthetic_gene_cluster_complete_sequence

Supplementary Table 2 Number of cyanobacteria ASVs from all genera observed from the Pockwock Lake watershed.

	Island Lake			Pockwock Lake		
Genus	**June**	**September**	**October**	**June**	**August**	**October**
Aphanizomenon NIES81	1	0	0	0	0	3
Cyanobium PCC-6307	1	6	3	1	4	4
Rhabdogloea smithii SAG 47.91	1	3	4	1	6	2
Snowella 0TU37S04	0	1	1	0	1	1
unclassified	1	0	2	1	1	2

Supplementary Table 3 Number of cyanobacteria ASVs from all genera observed from the Comox Lake watershed.

	Boston Creek		Cruikshank River		Lake Outlet		Upper Puntledge	
Genus	**May**	**September**	**May**	**September**	**May**	**September**	**May**	**September**
Calothrix KVSF5	2	1	0	0	0	0	0	0
Calothrix PCC-6303	0	0	0	0	0	0	0	1
Chamaesiphon PCC-7430	1	1	0	1	0	0	0	0
Chroococcidiopsis SAG 2023	1	0	0	0	0	0	0	0
Cyanobium PCC-6307	0	2	2	1	6	6	6	6
Gloeobacter PCC-7421	1	1	0	0	1	0	2	0
Gloeocapsa	0	0	0	0	0	1	1	1
Kamptonema PCC-6407	0	0	0	0	0	0	1	0
Leptolyngbya ANT.L52.2	1	0	0	0	1	0	0	0
Leptolyngbya SAG 2411	0	0	0	1	0	0	0	0
Microcystis PCC-7914	0	0	0	0	0	0	1	0
Microseira Carmichael-Alabama	0	0	0	0	0	0	1	0
Phormidesmis ANT.L52.6	0	0	0	0	0	0	1	0
Phormidium CYN64	0	1	0	0	0	0	0	0
Rhabdogloea smithii SAG 47.91	0	0	0	0	1	0	0	0
Scytonema PCC-7110	0	0	0	0	0	1	0	0
Scytonema UTEX 2349	1	1	0	0	0	0	0	0
Snowella 0TU37S04	0	0	0	0	0	1	0	0
Synechococcus PCC-7502	0	0	0	1	0	0	1	0
Tychonema CCAP 1459-11B	1	1	1	1	0	0	1	1
unclassified	4	4	5	4	0	0	5	4

Supplementary Table 4 Number of cyanobacteria ASVs from all genera observed from the Leech River/Sooke River watershed.

	Jarvis Lake		Weeks Lake		Deception Reservoir	
Genus	**July**	**August**	**July**	**August**	**July**	**August**
Aliterella CENA595	0	0	0	1	0	0
Aphanizomenon NIES81	0	0	1	1	1	2
Cyanobium PCC-6307	3	2	8	7	7	8
Gloeobacter PCC-7421	0	0	0	1	0	0
Gloeocapsa	0	0	1	0	0	1
Kamptonema PCC-6407	1	0	0	0	0	0
Microcystis PCC-7914	1	1	3	4	1	0
Nostoc PCC-73102	0	0	0	1	0	0
Rhabdogloea smithii SAG 47.91	3	4	3	3	2	1
Scytonema UTEX 2349	0	0	0	0	0	1
Snowella 0TU37S04	1	1	0	0	0	0
unclassified	0	0	0	0	0	2

Supplementary Table 5 Taxonomic assignment and ID of cyanobacteria ASVs and reference sequences for phylogenetic analyses.

SILVA Taxonomy	ASV ID	Reference Sequence Accession No.	Reference
Aliterella CENA595	ASV 1	MN243145.1 NR_151904.1 NZ_JYON01000060.1	Jung *et al*., 2020 Rigonato *et al*., 2016a Rigonato *et al*., 2016b
Aphanizomenon NIES81	ASV 2-6	HG917857.1 AJ293131.1 AJ133154.1	Casero *et al*., 2014 Gugger *et al*., 2002 Lyra *et al*., 2001
Calothrix KVSF5	ASV 7-9	EU022730.1 HM751856.1 NR_114995.1	Unpublished Berrendero *et al.*, 2011 Sihvonen *et al*., 2007
Chamaesiphon PCC-7430	ASV 10-12	KY704112.1 JX413491.1 AY170472.1	Kurmayer *et al*., 2018 Loza *et al.*, 2013 Turner, 1997
Chroococcidiopsis SAG 2023	ASV 13	MK484708.1 AJ344552.1 AJ344557.1	De Wever *et al*., 2019 Fewer *et al.*, 2002
Cyanobium PCC-6307	ASV 14-38	NR_102447.1 HQ380799.1 KP835527.1	Unpublished Genuário *et al*. 2016
Cyanothece aeruginosa SAG 87.79[A]	ASV 56	KM019992.1 MK484713.1 Z82775.1	Unpublished De Wever *et al*., 2019 Rudi *et al*., 1997
Gloeocapsa	ASV 40-41	KM376979.1 AF132784.1 AY790852.1	Singh *et al*., 2018 Turner *et al*., 1999 Norris and Catenholz, 2006
Gloeobacter PCC-7421	ASV 42-44	NR_121745.1 NR_074282.1	Saw *et al*., 2013 Nakamura *et al*., 2003

Kamptonema PCC-6407	ASV 45	MG563377.1 KP221931.1 AM398782.1	Obuekwe *et al*., 2019 Strunecký *et al.*, 2014 Marquardt and Palinska, 2007
Leptolyngbya ANT.L52.2	ASV 46-47	MN267144.1 AY493575.1	Jung *et al*., 2019 Taton *et al*., 2006
Leptolyngbya SAG 2411	ASV 48	KF417652.1	Unpublished
Macrochaete psychrophila CCALA 32[B]	ASV 49	KT336439.2 KT336440.1 KU559618.1	Berrendero Gómez *et al*., 2016
Microseira Carmichael-Alabama	ASV 50	KM077455.1 KM077454.1 EU439567.1	McGregor and Sendall, 2015 Kellmann *et al*., 2008
Nostoc PCC-73102	ASV 51	NR_074317.1 MN243123.1 AB098071.1	Unpublished Jung *et al*., 2019 Arima *et al*., 2012
Oscillatoriales cyanobacterium USR001[C]	ASV 52	MBRE01000011.1 EF654087.1 HF678514.1 KM019965.1	Te *et al*., 2016
Phormidesmis ANT.L52.6	ASV 53	MK861902.1 KU219735.1 AY493579.1	Strunecký *et al*., 2020 Raabová *et al*., 2019 Taton *et al*., 2006
Phormidium CYN64	ASV 54	JQ687330.1 KF770970.1 DQ493874.1	Unpublished Stoyanov *et al*., 2014 Comte *et al*., 2007
Radiocystis sp. JJ30-12[A]	ASV 55-59	AM710388.1 KF359770.1 AM710389.1	Unpublished
Rhabdogloea smithii SAG 47.91	ASV 60-68	KM020002.1	Unpublished
Scytonema PCC-7110	ASV 69	KY365479.1	Johansen *et al*., 2017

		JN565282.1 NR_112180.1	Smith *et al*., 2012 Tomitani *et al*., 2006
Scytonema UTEX 2349	ASV 70-72	NZ_ALWD00000000.1	Unpublished
Snowella 0TU37S04	ASV 73	MF680055.1 AJ781040.1 AJ781042.1	Shishido *et al*., 2017 Rajaniemi-Wacklin *et al*., 2006
Synechococcus PCC-7502	ASV 74-75	AF448080.1 NR_074309.1	Unpublished Sugita *et al*., 2007
Synechococcus sp. LEGE 06306[D]	ASV 39	HM217052.1	Lopes *et al*., 2012
Tychonema CCAP 1459-11B	ASV 76-79	LT546478.1 LM651415.1 AB045897.1	Salmaso *et al*., 2016 Shams *et al*., 2015 Suda *et al*., 2002

[A]Taxonomic output by SILVA was *Microcystis* PCC-7914 at the genus-level but *Cyanothece aeruginosa* SAG 87.79 and *Radiocystis* sp. JJ30-12 at the species-level.

[B]Taxonomic output by SILVA was *Calothrix* PCC-6303 at the genus-level but *Macrochaete psychrophila* CCALA 32 at the species-level.

[C]Taxonomic output by SILVA was *Kamptonema* PCC-6407 at the genus-level but *Oscillatoriales cyanobacterium* USR001 at the species-level.

[D]Taxonomic output by SILVA was *Cyanobium* PCC-6307 at the genus-level but *Synechococcus* sp. LEGE 06306 at the species-level.

Supplementary Table 6 Taxonomic assignment of cyanobacteria ASVs unresolved to the genus-level excluded from phylogenetic analyses.

ASV ID	SILVA Classification
ASV 80-82	Unclassified Caenarcaniphilales
ASV 83	Unclassified Chroococcidiopsaceae
ASV 84	Unclassified Cyanobacteria
ASV 85-87	Unclassified Gastranaerophilales
ASV 88	Unclassified Microcystaceae
ASV 89-95	Unclassified Obscuribacterales
ASV 96-97	Unclassified Pseudanabaenaceae
ASV 98-99	Unclassified SepB-3
ASV 100-105	Unclassified Sericytochromatia
ASV 106-113	Unclassified Vampirovibrionales

The taxonomic ranks of these 34 cyanobacterial ASVs provided are the most resolved that were obtained from SILVA classification.