Characterization of Microbial Communities of the Utikuma Region Study Area (Alberta, Canada)

by

Arnold Kwok

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, 2025

© Arnold Kwok 2025

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 Contribution

The environmental data was provided by, and the samples were collected in collaboration with, the Devito Lab from the University of Alberta (Edmonton, Alberta) and members active during the Summer of 2018: Dr. Kevin Devito, Dr. Emily Pugh, Megan Schmidt, Garret Malishewski, and Jessica Ritz. The pond samples were extracted for DNA in the lab and sent out for sequencing.

The DNA sequencing of the 16S rRNA gene sequence was conducted by Dr. Trevor Charles and Dr. Jiujun Cheng of Metagenom Bio Inc. (Waterloo, Ontario).

Abstract

Understanding the drivers of bacterial communities and the dynamics of boreal ecosystems is critical for safeguarding water quality and ecosystem resilience in the face of environmental change. This study investigates the interplay between cyanobacterial dynamics and ecohydrologic factors in the Utikuma Region Study Area (URSA) of northern Alberta. Cyanobacteria play a pivotal role in oxygenic photosynthesis and nutrient cycling, yet their unchecked proliferation can lead to harmful algal blooms with significant ecological and economic consequences.

Field sampling was conducted in 34 URSA ponds using a two-phase filtration process designed to capture both larger microbial and picocyanobacterial cells. DNA was extracted and analyzed through 16S rRNA gene sequencing, with taxonomic assignments derived from the curated CyanoSeq 1.3 database and complementary methods. Advanced bioinformatics pipelines provided high-resolution community profiling, while statistical analyses—including redundancy analysis and PERMANOVA—linked microbial diversity patterns to water chemistry parameters (e.g., pH, nutrients, trace metals) and ecohydrologic variables such as hydrologic relief areas, hydrologic units, and wildfire disturbances.

Key findings indicate that environmental factors explain up to 23.9% of the variance in microbial composition based on clr-transformed RDA. Notably, landscape-scale drivers and wildfire history significantly influenced community structure, with their interactions accounting for up to 30.1% of the variance (PERMANOVA). Dominant phyla observed included Pseudomonadota, Bacteroidota, Actinomycetota, Cyanobacteriota, and Verrucomicrobiota, while analysis of 165 cyanobacterial Amplicon Sequence Variants highlighted taxa such as Prochlorococcaceae and Aphanizomenonaceae.

These results, while based on a specific sampling period, provide valuable foundational insights into the relationships between ecohydrologic drivers and microbial community structure for a region previously uncharacterized in this regard. This work sets the stage for further research to refine taxonomic methods and inform targeted water quality management strategies in boreal ecosystems.

Acknowledgements

I want to express my sincere gratitude to my supervisor Dr. Kirsten Müller and my committee members Dr. Monica Emelko and Dr. Jonathan Witt. Thank you for the opportunities you have provided me, repeatedly through my strife, and for all the support. The journey to write this work has been difficult and I only have Dr. Kirsten Müller to thank for the continuous support.

I would like to give thanks to the members of the *for*Water NSERC network and collaborators for making the project possible and the Devito lab of the University of Alberta and to Dr. Trevor Charles and Dr. Jiujun Cheng of Metagenom Bio Inc. for their work to make the project possible, all of which was required for this work.

To the members of the Muller Lab, I need to give thanks to Dr. Vanessa Poletto Borges and MASc Elyse Batista for my chapter from my BIOL499 until my first break from my research. Without them the journey would have felt difficult, and I will always have fond memories of working in the lab and waxed philosophical on the importance of our work, the work, and life in general.

A special thanks to my friends and family who varied in watching my descent into madness; Brian Huang, Peter Zhu, Amy Zhang, Wilbert Lee, and Edens for their support at different points of my thesis and being there to check my sanity through the nth iteration of writing, data pipelines, visualizations, and reorienting me between the states and throes of the work.

Special thanks to Dr. Dragana Miskovic. My academic journey and perspective on life, work, and the pursuit of meaning would definitely be different if it was not for her guidance when I was lost in my undergraduate studies. Never will I forget your initial welcome that led me down this road, the pains of Embryology, nor will I forget the twilight marking and midday tea sessions.

Finally, I would like to thank my mother, Lucille Wong, for her support and continue willingness to support my endeavours no matter what they are.

"You have three choices in this life. Be good, get good, or give up."

- *Hugh Laurie (as Dr. Gregory House)*

Table of Contents

Author's Declaration ii
Statement of Contribution iii
Abstract iv
Acknowledgements v
List of Figures viii
List of Tables ix
Chapter 1. Overview 1
1.1 The Importance of Cyanobacteria 1
1.2 Boreal Plains Ecohydrology and Utikuma Region Study Area 2
1.2.1 Wildfire as a Dominant Disturbance in Boreal Ecosystems 3
1.3 Microbial Communities in Aquatic Ecosystems and the Role of Cyanobacteria 4
1.4 Molecular Analysis of Cyanobacterial Community Composition 6
1.5 Linking Ecohydrologic Factors to Microbial Community Structure 9
1.6 Research Questions and Objectives 10
Chapter 2. Materials and Methods 11
2.1 Sample Collection: URSA 11
2.2 DNA Extraction 13
2.3 Water Chemistry & Hydrologic Data 14
2.4 Amplicon Sequencing 18
2.5 Sequence Analyses 18
2.6 Microbial Diversity 19
2.7 Cyanobacterial Classification Qualifiers 21
Chapter 3. Results 23
3.1 Water Chemistry and Ecohydrology Factors 23
3.1.1 Water Chemistry Redundancy Analysis 23
3.1.2 PERMANOVA 25
3.2 Bacterial Community Composition 27
3.3 Cyanobacterial Community Composition 29
3.4 Cyanobacterial Phylogeny 34
3.5 Alpha Diversity 42
3.6 Beta Diversity 44
Chapter 4. Discussion 47
4.1 Microbial Communities of the Utikuma Region Study Area 48
4.2 Diversity and Influence of Ecohydrologic Factors 51
4.3 Implications for Ecosystem Function and Management 53
4.4 Taxonomic Challenges of Cyanobacteria 55
4.4.1 Picocyanobacteria and Prochlorococcaceae 58
Chapter 5. Conclusions 61
5.1 Summary of Key Findings 61
5.2 Main Implications 62
5.3 Significance of the Research 63
5.4 Future Research Directions 64
5.5 Concluding Remarks 66
References 68

Appendix A. Variance Inflation Factor for Water Chemistry Data of the URSA Ponds 79
Appendix B. Cyanoseq Classifier and Blast 16S Ribosomal Database Comparisons..... 80
Appendix C. Redundant Amplicon Sequence Variants Removed from Phylogenetic Analysis... 92
Appendix D. Centre Log-Ratio Redundancy Analysis Summary 102
Appendix E. Raw Reads Redundancy Analysis Summary..... 103

List of Figures

Figure 1 Two Examples of Pond Sampling. 12

Figure 2 GF/C Whatman Filter and vacuum filtration setup. 13

Figure 3 Images of different catchments of the URSA. 15

Figure 4 Map of the Utikuma Region Study Area. 17

Figure 5 RDA of Microbial Composition in URSA Ponds. Relationship between environmental variables and microbial communities in URSA ponds, highlighting 2013 wildfire burn regions. 24

Figure 6 Relative abundance of phyla from URSA 2018. 28

Figure 7 Cyanobacterial Genus-level Reads using CyanoSeq v1.3 classifier. 33

Figure 8 Cyanobacterial Phylogenetic Tree of ASVs from Utikuma Region Study Area 2018. .. 41

Figure 9 Alpha diversity of bacterial reads from the Utikuma Region Study Area. 43

Figure 10 Beta diversity of bacterial communities using Bray-Curtis dissimilarity. a) non-metric and b) PCoA. The categories were chosen using the PERMANOVA to delineate on hydrologic relief areas and regions affected by the 2013 wildfire 46

List of Tables

Table 2.1 GPS Locations of the Sample Ponds of the Utikuma Region Study Area. The landforms are characterized as HRAs and HUs as described by Devito et al. (2016). Two depths were not recorded for the 2018 sampling season. 16

Table 3.1 PERMANOVA Results for Environmental Variables and Interactions. The '+' denotes additive effects and '*' possible interactions between predictors. Df: Degrees of freedom; SumOfSqs: Sum of Squares; R2: Proportion of variance explained; F: F-statistic; P: Probability value from permutation test. * indicates significance at $P < 0.05$. 27

Table 3.2 Relative Abundances of Phyla across the Ponds of the Utikuma Region Study Area .. 27

Table 3.3 Summary of Cyanobacterial Classification. Unassigned has no classification at the family level. 31

Table 3.4 Redundant ASVs Removed from Phylogenetic Analysis. The (*) denotes sets containing ASVs. 40

Table A.1 VIF for Water Chemistry Data of the URSA Ponds 79

Table B.1 Cyanoseq Classifier and Blast 16S Ribosomal Database Comparisons 80

Table C.1 CyanoSeq v1.3 Sequences removed due to redundancy 92

Table C.2 CyanoSeq v1.3 Sequences removed due to redundancy. 101

Table D.1 ANOVA Permutation Test Results for CLR-RDA Model 102

Table D.2 CLR-RDA Model Summary - Variance Explained 102

Table D.3 Eigenvalues for CLR-RDA Axes 102

Table E.1 ANOVA Permutation Test Results for Raw Read RDA Model 103

Table E.2 Raw Read RDA Model Summary - Variance Explained 103

Table E.3 Eigenvalues for Raw Read RDA Axes 103

Chapter 1. Overview

1.1 The Importance of Cyanobacteria

Cyanobacteria, commonly known as blue-green algae, are photosynthetic prokaryotes recognized as primary drivers of life on Earth, shaping the planet's atmosphere through oxygenic photosynthesis over 3.5 billion years ago (Lyons et al., 2014; Sánchez-Baracaldo et al., 2022). These organisms inhabit a diverse range of environments, spanning oceans and freshwater, to terrestrial biocrusts and symbiotic associations with fungi and plants (Usher et al., 2007). In freshwater lakes, ponds, and wetlands, Cyanobacteria are key primary producers, contributing significantly to carbon fixation, nutrient cycling (Falkowski et al., 2008; Paerl & Otten, 2013; Ji et al., 2019), and ultimately supporting the food web (Falkowski et al., 2008). However crucial they are, Cyanobacteria can overproliferate, resulting in harmful algal blooms (HABs), posing a threat to water quality (Winter et al., 2011; Nwankwegu et al., 2019; Paerl et al., 2011; Yang et al., 2008; Zamyadi et al., 2012) and human and animal health (Chorus & Welker, 2021; Wood, 2016). These blooms produce potent toxins (Kinnear, 2010; Du et al., 2019; Huisman et al., 2018), with the genetics, biosynthesis, and environmental regulation of these toxins being crucial areas of study for predicting and managing toxic events (Pearson et al., 2016). Economic impacts from HABs continue to increase due to eutrophication of aquatic environments and can be costly to treat (Chorus & Welker, 2021; Ho & Michalak, 2020). Understanding the factors that govern cyanobacterial community composition is thus crucial for managing water resources, anticipating harmful algal blooms (Paerl & Paul, 2012; Paerl et al., 2011; Wells et al., 2020; Igwaran et al., 2024; Song et al., 2023), and protecting ecosystem health (Paerl & Otten, 2013; Nwankwegu et al., 2019). The need for standardized procedures for cyanotoxin analysis is an area of critical need

to effectively monitor and manage HABs (Igwaran et al., 2024; Du et al., 2019; Zamyadi et al., 2012).

1.2 Boreal Plains Ecohydrology and Utikuma Region Study Area

The Utikuma Region Study Area (URSA), a Boreal Plains landscape with extensive peatlands and numerous shallow ponds, experiences a dynamic sub-humid climate characterized by distinct wet and dry periods that significantly influence regional hydrology and ecosystem processes (Devito et al., 2012; Leader et al., 2024a, 2024b; Pugh et al., 2021). Within URSA, peatland-dominated areas notably exhibit ecohydrologic self-regulation, where features like moss surface resistance to evaporation during droughts help conserve water and influence local water movement (Kettridge & Waddington, 2014; Waddington et al., 2015). This inherent landscape resilience interacts with broader climatic shifts to shape water availability and storage across the region. Understanding these foundational ecohydrologic characteristics is crucial for contextualizing aquatic ecosystem dynamics.

In URSA's low-relief, glacially derived terrain, where surficial geology often overrides topography as the primary hydrologic driver (Devito et al., 2005), the landscape is conceptualized into Hydrologic Response Areas (HRAs) and finer-scale Hydrologic Units (HUs) (Devito et al., 2012; Petrone et al., 2016). HRAs, defined by surficial geology, are associated with differing permeability, water storage, surface water interactions, and consequently, distinct broad-scale biogeochemical conditions (Devito et al., 2012). These conditions influence factors such as nutrient budgets (Plach et al., 2016) and pond Dissolved Organic Carbon (DOC) characteristics (Pugh et al., 2021). While groundwater interactions vary by HRA, with coarse outwash areas often showing significant interaction with regional groundwater systems (Smerdon et al., 2005;

Hokanson et al., 2021), detailed mechanistic discussions of specific groundwater sourcing are beyond the scope of this high-level overview.

HUs, further defined by soil properties and dominant vegetation like wetlands or forests (Devito et al., 2012), contribute to a varied mosaic of environments. This landscape heterogeneity, driven by HU characteristics, results in localized variations in conditions such as organic matter inputs and hydrologic connectivity, which can even oppose local topography (Devito et al., 2012; Hokanson et al., 2020; Petrone et al., 2016). Such environmental diversity provides a range of niches and can foster distinct microbial microcosms, particularly at HU interfaces, potentially contributing to overall microbial diversity across URSA (Devito et al., 2012). The dynamic interplay of these landscape units (HRAs and HUs) with URSA's variable climate (Leader et al., 2024a, 2024b) ultimately shapes the diverse aquatic habitats under investigation in this study.

1.2.1 Wildfire as a Dominant Disturbance in Boreal Ecosystems

Wildfire is a paramount natural disturbance in boreal ecosystems like URSA, posing recognized threats to water resources (Paul et al., 2022; Robinne et al., 2016, 2019). The ecological impact and recovery trajectories following fire are significantly shaped by fire return intervals, where alterations to these historical cycles can induce lasting ecosystem shifts; the 2011 Utikuma Complex wildfire serves as a pertinent regional example of such an event. Fires alter landscapes by reducing vegetation and impacting soil properties (Certini, 2005; Emelko et al., 2024), leading to increased surface runoff and erosion (Silins et al., 2009; Alessio et al., 2021; Emelko et al., 2024). These impacts on sediment production can be substantial, persisting for years and affecting water quality even at large basin scales (Silins et al., 2009; Emmerton et al., 2020).

Increased runoff post-wildfire often transports elevated concentrations of sediments and associated contaminants such as metals (Abraham et al., 2017; Paul et al., 2022) and phosphorus

(Emelko et al., 2016; Gustine et al., 2022; Silins et al., 2014; Paul et al., 2022). Concurrently, other contaminants, such as Polycyclic Aromatic Hydrocarbons (PAHs), may be released (Crouch et al., 2006; Mansilha et al., 2014). Concentrations of nutrients such as nitrogen (Bladon et al., 2008; Gustine et al., 2022) are also elevated and organic carbon export is altered, with Dissolved Organic Carbon (DOC) concentrations often increasing and its chemical nature shifting towards more aromatic, treatment-resistant forms (Emelko et al., 2011; Emelko et al., 2024). These changes collectively redefine the biogeochemical fluxes within affected catchments.

Notably, altered sediment-phosphorus dynamics can create a long-term legacy, as bioavailable phosphorus desorbs from P-enriched fine sediments mobilized post-fire (Emelko et al., 2016; Stone et al., 2011). This sustained nutrient enrichment can, in turn, promote algal growth (Silins et al., 2014) and trigger broader ecosystem responses, including shifts in aquatic productivity (Martens et al., 2019; Silins et al., 2014), with impacts potentially magnified by convergence with other anthropogenic disturbances (Watt et al., 2021). Such widespread water quality alterations are particularly concerning due to the potential for increased proliferation of toxin-producing cyanobacteria, which affect human and animal health (Cheung et al., 2013; Zamyadi et al., 2012). Consequently, these multifaceted impacts underscore that wildfires can severely challenge drinking water treatment plant capacities, often leading to increased operational costs and potential service disruptions (Emelko et al., 2011).

1.3 Microbial Communities in Aquatic Ecosystems and the Role of Cyanobacteria

Microbial communities are essential components of aquatic ecosystems, driving key processes such as nutrient cycling (Ji et al., 2019; Yang et al., 2008), organic matter decomposition, and primary production (Falkowski et al., 2008). These communities are highly diverse,

encompassing bacteria, archaea, fungi, algae, protists, and viruses, each playing unique roles in ecosystem functioning (Malki et al., 2020; Vettorazzo et al., 2024; Acinas et al., 2009). Their metabolic capabilities allow them to break down complex organic molecules, including recalcitrant compounds not easily degraded through conventional decomposition pathways. Microbial communities are essential for new biomass creation and trophic transfer of nutrients and energy through grazing food webs (Falkowski et al., 2008). Temperature, pH, light, and nutrient availability are all important in structuring microbial assemblages for aquatic systems (Yang et al., 2008; Paerl & Otten 2013; Ho & Michalak, 2020; Zilius et al., 2014). Functional genes have also been recognized to play an important role in the processes of microbial communities (Levy-Booth et al., 2014). Understanding the intricacies within aquatic ecosystems is essential to managing water bodies (Igwaran et al., 2024; Wells et al., 2020; Winter et al., 2011; Ho & Michalak, 2020; Nwankwegu et al., 2019).

Cyanobacteria, photosynthetic bacteria (Demoulin et al., 2019), are important primary producers in freshwater ecosystems, contributing to carbon fixation and oxygen production (Falkowski et al., 2008; Paerl & Otten, 2013). They also interact closely with other microbial groups, forming complex relationships through competition for resources, mutualistic relationships including specialized cells that can fix nitrogen (Herrero et al., 2016; Foster et al., 2011), and viral lysis that can influence nutrient cycling and food web dynamics (Śliwińska-Wilczewska et al., 2018; Zhang et al., 2022; Zwart et al., 2005). These interactions between Cyanobacteria and other species are critical in maintaining a healthy food web (Mazard et al., 2016). While this study focuses on molecular characterization, traditional methods for identifying Cyanobacteria often include microscopy (Callieri, 2008; Putland & Rivkin, 1999; Stockner, 1988), sometimes with plankton nets. However, standard plankton nets (e.g., with mesh sizes of 20 μm

or larger) are generally unsuitable for capturing the smallest Cyanobacteria (Callieri, 2008; Stockner, 1988). These picocyanobacteria, typically 0.2 to 2 μm in size (Callieri, 2008; Stockner, 1988; Zhang et al., 2022), would largely pass through such nets, leading to their significant underrepresentation or complete omission if net-based sampling were the primary collection method. Microscopy can detect them, especially with techniques like epifluorescence (Putland & Rivkin, 1999; Stockner, 1988), but quantification and differentiation from other picoplankton can be challenging (Zhang et al., 2022). Given their ancient lineage (Callieri et al., 2022; Demoulin et al., 2019), global ubiquity (Callieri, 2008; Stockner, 1988), and ability to thrive even in oligotrophic conditions (Stockner, 1988; Zhang et al., 2022), picocyanobacteria are expected to be a persistent and foundational component of the microbial communities in the URSA ponds. Picocyanobacteria can account for a substantial part of cyanobacterial biomass and are a key component to monitor in both spatial and temporal contexts (Batista & Giani, 2019; Nwosu et al., 2021; Cameron et al., 2022; Cameron et al., 2024). Picocyanobacteria that thrive at different depths and thermal conditions are useful components to analyze within lakes (Nwosu et al., 2021; Cameron et al., 2022; Cameron et al., 2024; Sim et al., 2023). In cases of low levels of phosphorus, picocyanobacteria can continue to proliferate in aquatic ecosystems, demonstrating the importance of the bacteria even under climate stressors (Śliwińska-Wilczewska et al., 2018).

1.4 Molecular Analysis of Cyanobacterial Community Composition

Taxonomic classification of cyanobacteria, also known as Cyanobacteriota under the Genome Taxonomy Database (GTDB), is challenging due to their diverse morphology, epigenetic traits, and capability for horizontal gene transfer results in difficulty for taxonomic classification (Lefler et al., 2023; Strunecký et al., 2023; Komárek, 2016; Acinas et al., 2009; Altschul et al., 1997; Moreira et al., 2013). Classifying Cyanobacteria has traditionally required a polyphasic

taxonomic approach (Komárek, 2016; Komárek et al., 2020; Sciuto & Moro, 2016), which considers multiple lines of evidence, including morphology, physiology, ecology, and molecular data (Komárek, 2016; Acinas et al., 2009; Altschul et al., 1997; Yutin & Galperin, 2013). This integrated approach is necessary because relying on a single characteristic can be insufficient for accurate species delineation, given the variability and plasticity of cyanobacterial traits (Komárek, 2016; Acinas et al., 2009).

The 16S ribosomal RNA (rRNA) gene is widely employed for microbial community analysis (Caporaso et al., 2011) due to its ubiquitous presence in bacteria (Janda & Abbott, 2007; Johnson et al., 2019; Caporaso et al., 2011) and archaea, and its structure, which includes both conserved regions suitable for universal primer design and hypervariable regions that offer phylogenetic information (Johnson et al., 2019; Caporaso et al., 2011). This approach allows for high-throughput sequencing providing a culture-independent and cost-effective means to profile complex microbial assemblages from numerous environmental samples (Straub et al., 2020). However, the 16S rRNA gene method also has inherent limitations (Janda & Abbott, 2007). These include potential biases introduced during PCR amplification (such as primer mismatches and chimera formation) (Caporaso et al., 2011; Johnson et al., 2019; Janda & Abbott, 2007), variations in gene copy number among different taxa which can affect perceived relative abundances (Johnson et al., 2019; Janda & Abbott, 2007), and issues related to the completeness and accuracy of reference databases used for taxonomic assignment. Furthermore, while broadly effective for higher taxonomic ranks, its resolution can be limited for distinguishing closely related species or strains (Janda & Abbott, 2007; Johnson et al., 2019; Caporaso et al., 2011). While 16S rRNA gene sequencing offers an advantage over traditional microscopic identification (MacKeigan et al., 2022), its utility for cyanobacterial classification is limited by the challenges of taxonomic resolution and the

occurrence of horizontal gene transfer (Strunecký et al., 2023; Werner et al., 2024; Acinas et al., 2009). Furthermore, interpretation of microbial diversity in samples can also lead to some errors (Pacheco et al., 2016; Jung, 2023). Increasing the OTU clustering reveals the cryptic cyanobacterial genera (Li et al., 2020). Furthermore, reliance on morphology alone or single genetic markers can be insufficient for accurate classification, as demonstrated in cases like the bloom-forming genera *Microcystis* and *Radiocystis*, where polyphasic characterization including multiple genetic loci and secondary structure analysis was necessary for correct taxonomic placement (Lv et al., 2022). High-throughput sequencing, although with limitations, can provide a comprehensive perspective to assess the cyanobacterial populations (Zhang et al., 2022).

To improve taxonomic assignments, this study utilizes the CyanoSeq database (version 1.3; Jung 2023; Lefler et al., 2023). This database is curated for cyanobacterial 16S rRNA gene sequences and utilizes the GTDB taxonomy of purely molecular evidence to resolve taxonomic issues (Parks et al., 2022). The use of the GTDB taxonomy has led to the term Cyanobacteriota to describe this phylum (Parks et al., 2018; Strunecký et al., 2023). In comparison to the NCBI database, the GTDB taxonomy aims to mitigate the NCBI issues with resolution and solve issues related to polyphasic taxonomy within the Cyanobacteria (Lefler et al., 2023). CyanoSeq can improve the resolution in cyanobacterial communities, making it an essential for identifying novel bacteria (Li et al., 2020). In addition, CyanoSeq aims to improve taxonomic resolution and accuracy by incorporating curated cyanobacterial sequences aiming to improve accuracy of classification (Jung, 2023; Lefler et al., 2023). Other tools like Cydrasil (Roush, 2021) and the SILVA database (Glöckner et al., 2017; Yilmaz et al., 2014; Quast et al., 2013) can be used in this context. When doing metabarcoding studies, it is useful to include a primer to account for bacterial contaminants in the process (Altschul et al., 1997; Rastogi et al., 2010).

1.5 Linking Ecohydrologic Factors to Microbial Community Structure

Hydrologic connectivity is recognized as a driver of microbial community assembly for freshwater systems (Tsuji et al., 2020; Malki et al., 2020; Vettorazzo et al., 2024). The forces that direct the interplay of bacterioplankton community composition are affected by the residence time and local water chemistry (e.g., pH, nutrient concentrations) that cause sequential shifts (Niño-García et al., 2016; Zheng et al., 2022). In the context of the URSA, these features are highlighted by the unique mosaic of forestlands and peatlands of the boreal landscape (Devito et al., 2012; Hokanson et al., 2020). The hydrogeologic conditions of complex water connectivity and glacial deposits influence the local nutrient dynamics, with nitrogen limitation being a common feature of lakes in the URSA, and consequently microbial dispersal (Tsuji et al., 2020; Malki et al., 2020; Vettorazzo et al., 2024). The hydrogeomorphic setting and wet/dry periods significantly influence DOC characteristics in these lakes (Pugh et al., 2021), and the presence of peatlands within the Boreal Plains directly influences nutrient dynamics through production of dissolved organic matter and nitrogen availability (Herndon et al., 2023). High levels of nutrients can cause a cascade of harmful effects in the food web and have many unintended consequences that are still being investigated (Sukenik, 2012). The capturing of environmental factors into hydrologic relief areas and hydrologic units reflects these characteristics and thereby influence microbial community structure. Thus, integrating hydrologic frameworks with nutrient dynamics offers an improved understanding of how microbial community structures could progress leading to the potential leverage of geohydrology for use in effective water management strategies (Nwankwegu et al., 2019), building upon established frameworks for understanding Boreal Plain hydrology (Devito et al., 2005).

1.6 Research Questions and Objectives

The Utikuma Region Study Area (URSA) presents a complex mosaic of ecohydrologic features whose influence on aquatic microbial communities remains largely uncharacterized. This study addresses this knowledge gap by undertaking one of the first comprehensive molecular investigations of aquatic microbial communities (including cyanobacteria) within this specific boreal landscape. Given the challenges in accurately classifying cyanobacteria, the intricate interplay between microbial groups, and the need to understand their environmental drivers, this research aims to utilize 16S rRNA gene sequencing and curated databases to:

1. Characterize the bacterial community composition in 34 ponds of the URSA using 16S rRNA gene sequencing and the CyanoSeq 1.3 database to identify the presence of problematic taxa.
2. Examine the influence of environmental spatial variables (HRAs and HUs) on the observed bacterial communities.

This thesis aims to provide an initial assessment of cyanobacterial communities in the URSA and their connections to ecohydrologic factors, contributing to a better understanding of the structure and function of these understudied boreal ecosystems, and informing future water resource management strategies.

Chapter 2. Materials and Methods

2.1 Sample Collection: URSA

The samples were collected in early 2018, from July 4th to 9th, from the ponds of the Utikuma Region Study Area (URSA). A total of 34 ponds were sampled from various relief areas; these ponds are shallow, assumed to be polymictic, and non-stratified during the sampling period (see Table 2.1 for pond characteristics). For this study, all 34 sampled water bodies in the URSA are referred to as ponds. This classification is based observed characteristics of shallow depth and assumed polymictic nature that distinguishes them from larger, deeper, and potentially stratified water bodies that might be termed lakes within the broader region. Each pond was point-sampled using the following procedure:

1. Accessing Ponds: Ponds were entered using the most available method (Figure 1).
2. Bailer Rinsing: Bailers were rinsed three times at each pond.
3. Water Collection: Water was collected at an approximate depth of 1 meter. The containers are opaque with no headspace to avoid fractionation. Samples were kept in iced coolers.
4. Sample Storage: The containers were stored in a refrigerator within 12 hours, and the samples were filtered within 48 hours of collection.

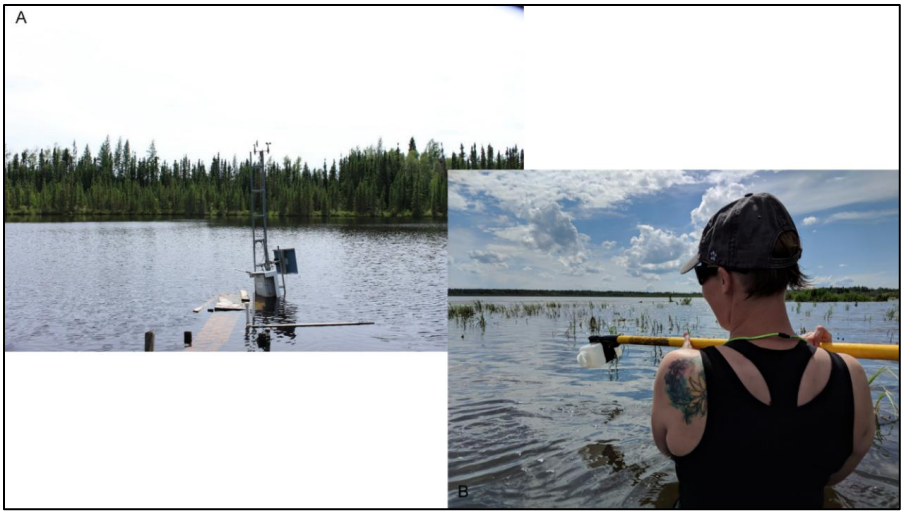


Figure 1 Two Examples of Pond Sampling. a) A permanent sampling platform with instrumentation at one pond. b) Wading into a shallower pond (approx. 3m from shore) to collect a sample using a bailer, typical for sites without platforms.

The filtering process occurred in two phases to target larger particulates and bacterial cells, and the subsequent filtration aimed to capture picocyanobacterial that are reported at 0.2 µm in size:

1. The first phase involved filtering 500 mL of the collected samples using 1.2 µm membranes 47-mm glass fiber; Whatman GF/C. The used filters were stored in 50-mm polystyrene petri dishes (Figure 2).
2. Then, 250 mL of filtrate was passed through a 0.2-µm Sterivex filter column (EMD Millipore Corporation, MA, USA). Both the glass filters and the Sterivex filters were stored in Whirl-Paks and kept at -20°C until DNA extraction.

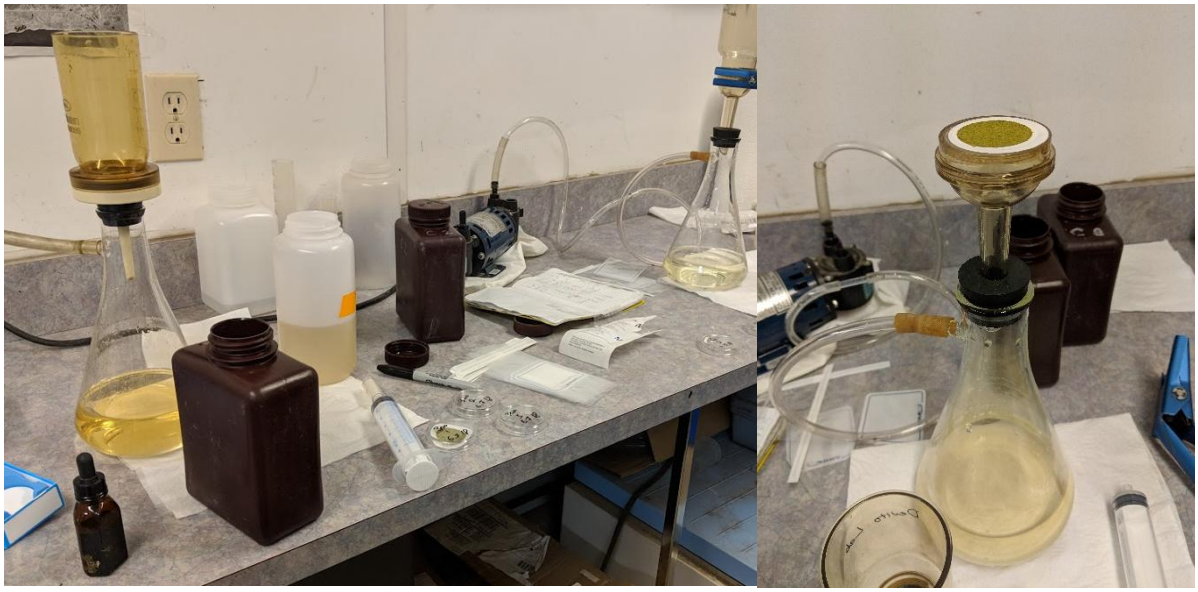


Figure 2 GF/C Whatman Filter and vacuum filtration setup. Utikuma Region Study Area remote laboratory.

2.2 DNA Extraction

DNA was extracted from the glass-membrane filters and the Sterivex filters. These two filter types were combined to capture both larger microbial cells and picocyanobacterial cells, thus increasing the overall breadth of the DNA pool. The Qiagen DNeasy PowerSoil Kit™ was used for extraction with three modifications.

1. The filters were excised and crushed using a mortar and pestle to create a mixture that substituted the expected soil sample for the kit.
2. An additional 5 minutes were added to the bead-beating phase.
3. The samples sat in the elution solution for an additional 5 minutes before the final centrifugation.

These modifications to the protocol were implemented to maximize the recovery of DNA from the potentially resilient microbial cells captured on the filters.

2.3 Water Chemistry & Hydrologic Data

The Devito Lab at the University of Alberta collected samples for water chemistry analyses during the sampling period. Field measurements included temperature and depth. Additional analyses were prepared by filtering samples to 0.45 μm within 24 hours of collection and storing them at 4°C. Water chemistry analyses for the 2018 season were conducted by the Limnology Laboratory and Earth and Atmospheric Sciences High Volume Laboratory at the University of Alberta. For water chemistry variables with concentrations below the analytical detection limit (LOD), values were substituted with LOD/2 for quantitative analyses. This approach is chosen because it retains all data points for statistical analyses, ensures positive non-zero values which are necessary for certain data transformations while still acknowledging the analyte's presence at a low, unquantifiable concentration. Missing pH measurements for ponds 42 and 111 were imputed using the mean pH across all other samples prior to statistical analysis.



Figure 3 Images of different catchments of the URSA. Top. Pond 40; Bottom Left, Pond 2; Bottom Right, Pond 302.

Table 2.1 GPS Locations of the Sample Ponds of the Utikuma Region Study Area. The landforms are characterized as HRAs and HUs as described by Devito et al. (2016). Two depths were not recorded for the 2018 sampling season.

ID	Longitude	Latitude	Sample Date	Hydrologic Relief Area	Hydrologic Units	Depth (m)
1	115.617	56.108	8-07-18	Coarse Outwash	Wetland	2.5
2	115.603	56.096	8-07-18	Coarse Outwash	Wetland	4.1
5	115.567	56.116	7-07-18	Coarse Outwash	Wetland	3.3
7	115.582	56.087	8-07-18	Coarse Outwash	Forestland	1.8
11	115.558	56.084	8-07-18	Coarse Outwash	Wetland	1.5
15	115.569	56.086	8-07-18	Coarse Outwash	Forestland	0.5
16	115.511	56.108	7-07-18	Coarse Outwash	Wetland	2.0
17	115.548	56.093	8-07-18	Coarse Outwash	Forestland	5.5
19	115.538	56.083	8-07-18	Fine Hummocky	Forestland	1.95
27	115.519	56.068	8-07-18	Fine Hummocky	Forestland	1.0
39	115.483	56.062	6-07-18	FH-CO	Wetland	2.5
40	115.477	56.074	6-07-18	FH-CO	Forestland	1.0
42	115.477	56.06	6-07-18	Fine Hummocky	Mixed	2.3
43	115.47	56.077	6-07-18	Fine Hummocky	Forestland	0.8
48	115.468	56.072	6-07-18	Fine Hummocky	Forestland	0.9
59	115.377	56.072	8-07-18	Fine Hummocky	Mixed	1.3
111	115.423	56.027	8-07-18	Fine Hummocky	Mixed	2.0
112	115.43	56.012	8-07-18	Fine Hummocky	Forestland	1.6
118	115.382	56.019	7-07-18	Fine Hummocky	Wetland	1.6
121	115.342	56.01	6-07-18	Fine Plain	Wetland	1.5
122	115.348	56.005	7-07-18	Fine Plain	Wetland	1.6
160	115.334	55.963	6-07-18	Fine Plain	Wetland	-
162	115.301	55.975	6-07-18	Fine Plain	Mixed	2.0
168	115.204	55.985	6-07-18	Fine Plain	Wetland	1.6
169	115.191	55.968	6-07-18	Fine Plain	Wetland	1.7
171	115.193	55.983	6-07-18	Fine Plain	Wetland	2.8
201	115.709	56.121	7-07-18	Fine Plain	Mixed	1.3
205	115.154	55.964	9-07-18	Fine Plain	Wetland	3.0
206	115.703	56.13	7-07-18	Fine Plain	Forestland	1.5
208	115.707	56.115	7-07-18	Fine Plain	Wetland	1.2
300	115.122	55.871	6-07-18	Fine Plain	Wetland	1.5
301	115.118	55.819	6-07-18	Fine Plain	Wetland	-
302	115.090	55.830	6-07-18	Fine Plain	Wetland	3.0
303	115.132	55.829	6-07-18	Fine Plain	Wetland	2.8

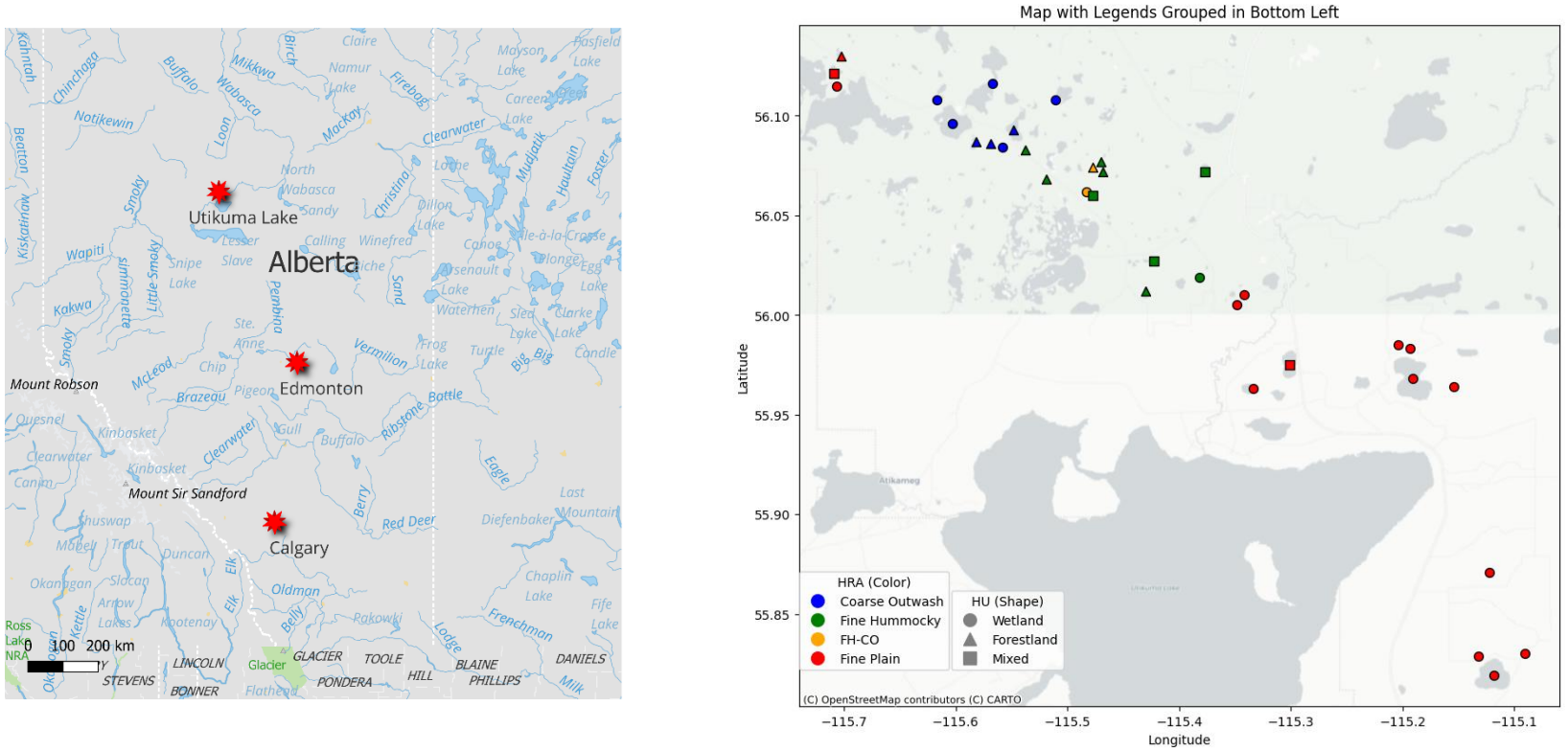


Figure 4 Map of the Utikuma Region Study Area. Colours correspond to landform classifications described in 2018 from Table 2.1. Hydrologic Relief Areas are designated by colour. FH-CO refers to a mixed region of Fine Hummocky sediment and Coarse Outwash. Hydrologic Units are described by shape**.**

2.4 Amplicon Sequencing

For each pond, DNA extractions from the 47-mm glass fiber and Sterivex filters were combined at a 1:1 ratio for a final volume of 20 µL per site. Aljoudi (2021) observed differences between communities from GF/F filter samples and Sterivex samples. To account for both communities, the DNA extractions were mixed for each site. Samples were required to have a final concentration of 1-10 ng/µL, confirmed using a NanoDrop 2000 spectrophotometer (Thermo Scientific, DE, USA) before submission. DNA samples were stored in microcentrifuge tubes at -20°C until sent for next-generation DNA sequencing.

The 34 DNA extracts from the URSA region were sent to Metagenom Bio Inc. at the University of Waterloo, Ontario, Canada, for amplification of the V4 region of the 16S rRNA gene. The initial PCR reaction targeted the 16S rRNA V4 regions as outlined by Walters et al. (2015): 515FB (5'-GTGYCAGCMGCCGCGGTAA) and 806RB (5'-GGACTACNVGGGTWTCTAAT-3'), using protocols from previous Metagenom Bio Inc. publications. An Illumina MiSeq and MiSeq Reagent Kit v2 (Illumina, Inc., San Diego, CA, USA) produced 2 x 250 reads. The V4 region was selected as it is a commonly targeted hypervariable region providing a good compromise between phylogenetic information content for bacteria and archaea and amplicon length suitable for widely used short-read Illumina sequencing platforms.

2.5 Sequence Analyses

All sequences were demultiplexed in a 2 x 250 bp read format and input into a QIIME2 pipeline (version 2024.2; Bolyen et al., 2019). The tool DADA2 was used to trim 15 bp from the 5' ends to remove primers (Callahan et al., 2016). Forward and reverse reads were truncated to 250 bp after inspecting the Phred scores produced by the QIIME2 demux summary to filter out poor sequence quality. DADA2 was used to merge paired reads and remove chimeric sequences and

singletons (Callahan et al., 2016). The removal of singletons is a potential source of bias for diversity analyses (Schmidt et al., 2022). A feature table containing amplicon sequence variants (ASVs) was produced and taxonomically classified using the q2-feature-classifier (Bokulich et al., 2018) naïve Bayesian classifier based on the CyanoSeq v1.3 (Lefler et al., 2023). The resulting ASV data and frequency data were exported for further analysis in R Statistical Software (v4.4.2; R Core Team 2024).

It is important to note that the CyanoSeq v1.3 database utilizes the Genome Taxonomy Database (GTDB) framework (Parks et al., 2022). Within this framework, phylum nomenclature may differ from traditional taxonomies. Specifically, the phyla traditionally known as Proteobacteria, Actinobacteria, Bacteroidetes, and Verrucomicrobia are referred to as Pseudomonadota, Actinomycetota, Bacteroidota, and Verrucomicrobiota, respectively. These GTDB phylum names are used consistently throughout this thesis following the classification method employed.

2.6 Microbial Diversity

The generated files contained GTDB taxonomy. The *phyloseq* (v1.42.0) (McMurdie & Holmes, 2013) objects were created using QIIME2 outputs. The data manipulations were done using *dplyr* (v1.1.4) (Wickham et al., 2023) and *Tidyverse* (v2.0.0) (Wickham et al., 2019). Visualizations and plotting of all figures in the R software used *ggplot2* (v3.5.1) (Wickham, 2016) and *ggh4x* (v0.3.0) (Brand, 2024).

The output ASVs were additionally filtered to remove sequences identified as chloroplast and mitochondrial to focus on the bacterial ASVs. The microbial diversity of the samples was analyzed in three main categories: α-diversity (within a sample), β-diversity (among samples), and other analyses. The initial phylum abundance was calculated using relative abundances by creating

proportions within each sample and then all samples were averaged again. Project-wide taxa that were below 1% at the phylum-level were aggregated into a single group of *other*.

$$Relative\ Abundance_{phylum} = {\#\ of\ Reads_{phylum}}/{Total\ \#\ of\ Reads}$$

The microbial diversity of the samples was analyzed in three main categories: α-diversity (within a sample), β-diversity (among samples), and other analyses. Alpha diversity quantifies the number and relative abundance of different taxa within a single sample. Beta diversity measures how the microbial composition differs between various samples.

A preliminary analysis utilized a rarefaction curve to visualize the normalized library sizes and a depth of 10,000 reads per sample was chosen for inclusion of all samples to represent microbial diversity for comparison. The rarefied library was created using sampling without replacement with 100 repetitions to normalize data for detecting differential abundances of ASVs between samples. A repeated subsampling approach from the R package *mirlyn* (Cameron et al., 2021). The alpha diversity was measured using the Shannon Diversity Index in the R package phyloseq (McMurdie & Holmes, 2013).

Beta diversity analyses were performed on the ASVs using Bray-Curtis dissimilarity (Bray & Curtis, 1957). Given the 21 water chemistry variables provided (Table A.1), a variance inflation factor (VIF) analysis was conducted to reduce multicollinearity and focus on the most informative parameters for the redundancy analysis (RDA). Because microbiome data are compositional, a center log-ratio (clr) transformation was applied to mitigate biases associated with analyzing such data directly. For each pond, the center log-ratio is applied using the formula

$$\mathrm{clr}(x)_i = \ln(x_i + 1) - \frac{1}{n}\sum_{j=1}^{n} \ln\left(x_j + 1\right)$$

The raw ASVs and the clr transformed reads were then analyzed using RDA with the vegan package (v2.6) (Oksanen et al., 2013).

Subsequently, permutational multivariate analysis of variance (PERMANOVA) was chosen to evaluate the influence of categorical variables (HRAs, HUs, and burn regions) on these community structures. The 'burn regions' classification, indicating whether ponds were affected by the 2011 Utikuma Complex wildfire, was based on information provided by the Devito Lab (University of Alberta) from their field observations and records for the URSA sites. PERMANOVA is suitable because it analyzes multivariate data based on a selected dissimilarity measure, is robust for ecological data not meeting parametric assumptions, and tests for differences between predefined groups. Bray-Curtis dissimilarity is a common choice that effectively uses species abundance data to compare community compositions and appropriately handles shared absences of taxa. Categorical variables (HRAs, HUs, and burn regions) were evaluated in all possible combinations using PERMANOVA with 999 permutations to inform subsequent visualizations. The samples showed significant effect from burns and HRAs at a value of $P < 0.001$ with an R^2 value of 0.30. Both nonmetric multidimensional scaling (NMDS) and principal coordinate analysis (PCoA) were conducted on the overall sample and a subset sample of just the cyanobacterial ASVs visualized on the subset groupings of burns and HRAs.

2.7 Cyanobacterial Classification Qualifiers

The resulting Cyanobacteriota ASVs were extracted from the QIIME2 sequences for classification analysis by BLAST against the NCBI 16S_ribosomal database, taking matches with the highest percentage identity and compared against the classifier.

The CyanoSeq sequences are a database of 4291 sequences with an assortment of ranges and lengths of ribosomal rRNA sequences. As a result, a manual curation of this database was taken and trimmed to match sequences at a length of 252 at the V4 region. The database and Cyanobacteriota sequences were aligned using the MUSCLE super5 algorithm (Edgar, 2022), then manually trimmed to the V4 region. The MUSCLE algorithm was chosen for its balance of speed and accuracy in generating multiple sequence alignments (Edgar, 2004; Edgar, 2022). However, it is important to acknowledge that all alignment algorithms, including MUSCLE, employ heuristics and rely on parameters such as gap opening and extension penalties, which critically influence how insertions and deletions (indels) are handled. Aligning hypervariable regions, such as those within the 16S rRNA V4 gene, can be particularly challenging due to frequent indels. If default parameters were used for MUSCLE super5, these are generalized and may not be optimal for this specific dataset, potentially impacting alignment accuracy in variable regions (Edgar, 2022). Such alignment inaccuracies or biases, especially in the placement of gaps, can significantly affect downstream phylogenetic inferences (Edgar, 2022), potentially influencing tree topology and branch support. The Maximum-Likelihood method with the General Time Reversible model and 1000 bootstrap replicates (Nei & Kumar, 2000). There were 40 maximum-likelihood trees were produced and tested in RAxML. The tree topology with the highest likelihood score, selected from 40 independent search replicates, was used for analysis, with branch support assessed using 1000 bootstrap replicates. Actinomycetota and Firmicutes were selected as outgroups for rooting the cyanobacterial phylogenetic tree because they are established as relatively distant, yet related, bacterial phyla to Cyanobacteriota in broad scale phylogenomic studies, providing a robust anchor for inferring evolutionary relationships within the ingroup.

Chapter 3. Results

3.1 Water Chemistry and Ecohydrology Factors

3.1.1 Water Chemistry Redundancy Analysis

A determination of the significance of relationships in environmental samples can be done through a redundancy analysis (RDA). To inform the environmental variables to test, a variance inflation factor (VIF) was conducted with a total of 20 variables to reduce effects of collinearity. The selection criteria for RDA were categories with scores of less than 10 (i.e. pH, Fe mg/L, NO2/NO3 µg/L, Chlorophyll *a* µg/L, and Al µg/L; Table A.1). The redundancy analysis conducted on the raw frequencies indicates that the environmental variables explain 20.6% of the total variance in the ASV composition leaving 79.4% unexplained. An ANOVA permutation test was conducted on the RDA model with 6 degrees of freedom with an F value of 1.1648 and a p-value of 0.279 suggesting the overall model is not statistically significant (Figure 5a).

A centre log-ratio (clr) transformed redundancy analysis was also conducted to normalize frequencies across the samples to mitigate compositional effects. The clr transformed RDA indicates the environmental factors accounted for 23.9% of the total variance with a remaining 76.1% unexplained. Notably, the permutation test with 6 degrees of freedom has an F value of 1.4146 with a statistically significant p-value of 0.011 (Table D.1). This affirms that the environmental variables had a significant effect on the microbial community composition under the clr transformation (Figure 5b).

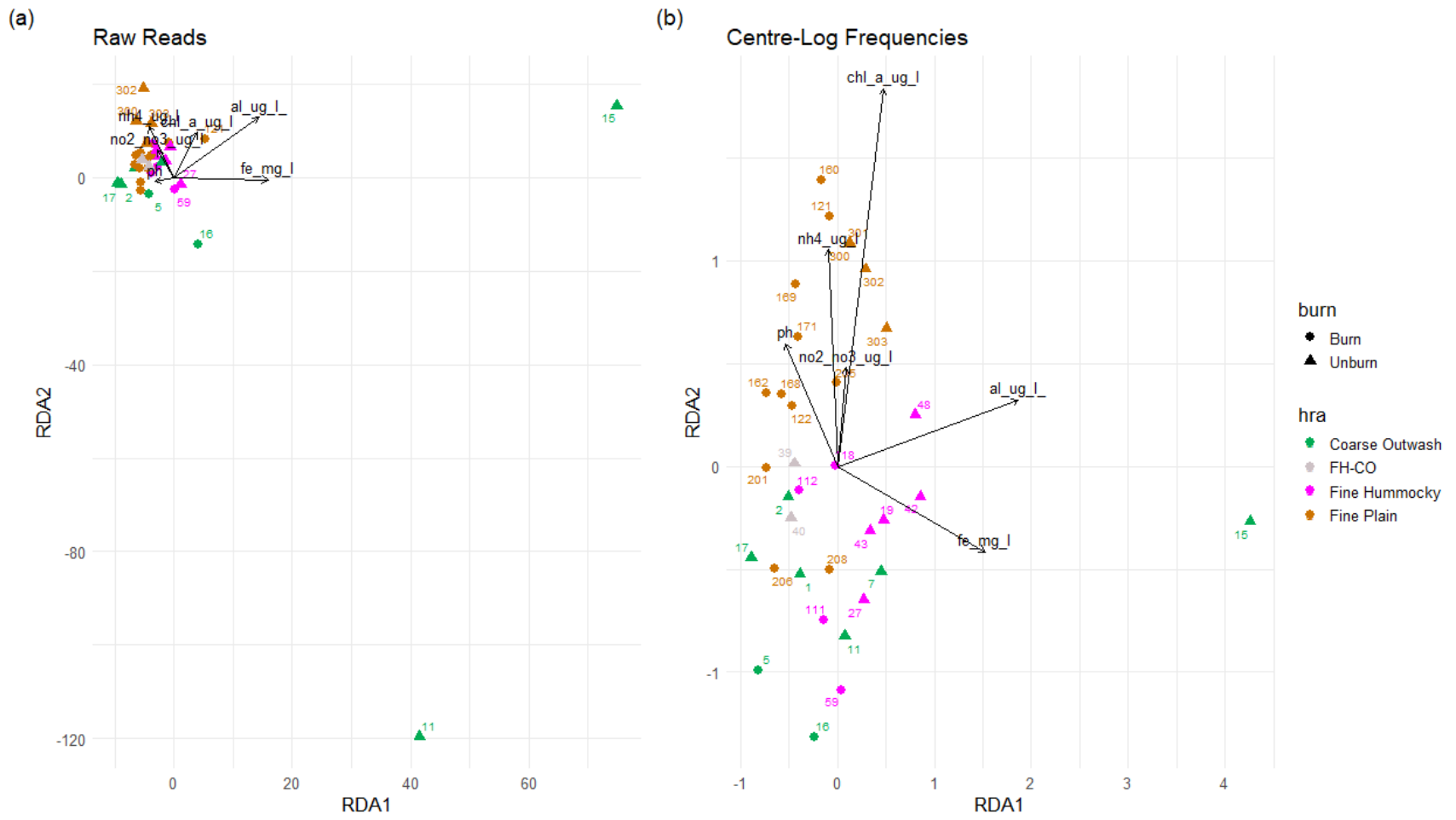


Figure 5 RDA of Microbial Composition in URSA Ponds. Relationship between environmental variables and microbial communities in URSA ponds, highlighting 2013 wildfire burn regions. a) Raw data RDA: 20.6% variance explained ($R^2 = 0.206$, $p = 0.279$). b) clr-transformed RDA: 23.9% variance explained ($R^2 = 0.239$, $p = 0.011$). Points: pond samples; Vectors: environmental variables (pH, Fe, NO2/NO3, Chlorophyll a, Al); Shapes: burn regions.

3.1.2 PERMANOVA

To identify the ecohydrologic effects on microbial community structure within the Utikuma Region Study Area (URSA), a PERMANOVA was executed. This analytical approach quantified the independent and interactive effects of hydrologic relief areas (HRAs), hydrologic units (HUs), and historical wildfire events on observed microbial community variation. Given the URSA's heterogeneous landscape, characterized by a mosaic of forests, peatlands, and variable hydrologic connectivity, the identification of these determinants is essential for interpreting cyanobacterial community dynamics and informing water resource management. It was hypothesized that landscape-scale factors, represented by HRAs, and disturbance-driven factors, such as wildfires, would exert influence on microbial community composition.

Results from the PERMANOVA, as detailed in Table 3.1, demonstrated significant influences of HRAs and wildfire history on microbial community composition. HRAs, which encapsulate broad landscape-scale hydrologic and geological characteristics, accounted for 16.0% of the variance ($R^2 = 0.159996$, $F = 1.904699$, $P = 0.001$), indicating a strong influence of overarching geological and hydrologic settings on microbial assemblages. Similarly, areas impacted by the 2013 wildfire, a significant ecological disturbance, accounted for 5.3% of the variance ($R^2 = 0.053168$, $F = 1.796917$, $P = 0.008$), demonstrating the impact of fire on microbial communities. In contrast, HUs, which represent finer-scale hydrologic variations, did not significantly influence community structure when considered independently ($R^2 = 0.072201$, $F = 1.206207$, $P = 0.139$). This suggests that while finer-scale hydrologic variations may contribute, their influence is less pronounced than landscape and disturbance factors.

The combined effects of these ecohydrologic factors exhibited enhanced explanatory power. Notably, the interaction between HRAs and wildfire history explained 22.1% of the variance (R^2

= 0.22077, P = 0.001), indicating that wildfire impact is modulated by the broader hydrologic and geological context. The additive effects of HUs and HRAs accounted for 21.5% of the variance (R^2 = 0.214836, P = 0.001), and the combination of HUs and wildfire accounted for 12.1% (R^2 = 0.12128, P = 0.022), further emphasizing the interconnectedness of these variables. This implies that microbial communities are shaped by a combination of landscape-scale hydrology, fine-scale hydrology, and disturbance regimes.

Further analysis of interaction terms revealed complex relationships. The interaction between HRAs and wildfire history was particularly robust, explaining 30.1% of the variance (R^2 = 0.301376, P = 0.001), highlighting that fire effects are contingent on specific hydrologic relief. The three-way interaction of HUs, HRAs, and wildfire explained 46.9% of the variance (R^2 = 0.468639, P = 0.001), revealing a complex relationship. However, the interaction between HUs and wildfire history was not significant (R^2 = 0.172175, P = 0.098), and due to sample size, the significant HU x HRA interaction will not be used in further analysis (R^2=0.34568, P=0.022). These findings underscore the importance of considering interaction effects when analyzing ecohydrologic influences on microbial communities.

These results underscore the significant role of both landscape-scale (HRAs) and disturbance-driven (wildfire) ecohydrologic factors in shaping microbial communities within the URSA. Given the significant interaction between HRAs and wildfire, subsequent analyses will focus on these factors to explore their influence on cyanobacterial community composition, providing insights into how hydrologic connectivity and disturbance regimes interact to structure microbial communities. This approach will contribute to a more detailed understanding of the URSA's microbial ecology, informing water resource management and conservation strategies in this complex boreal ecosystem.

Table 3.1 PERMANOVA Results for Environmental Variables and Interactions. The '+' denotes additive effects and '*' possible interactions between predictors. Df: Degrees of freedom; SumOfSqs: Sum of Squares; R2: Proportion of variance explained; F: F-statistic; P: Probability value from permutation test. * indicates significance at $P < 0.05$.

Model	Df	SumOfSqs	R2	F	P
hu	2	0.75	0.07	1.21	0.139
hra	3	1.65	0.16	1.90	0.001*
burn	1	0.55	0.05	1.80	0.008*
hu + hra	5	2.22	0.21	1.53	0.001*
hu + burn	3	1.25	0.12	1.38	0.022
hra + burn	4	2.28	0.22	2.05	0.001*
hu * hra	10	3.57	0.35	1.22	0.022
hu * burn	5	1.78	0.17	1.16	0.098
hra * burn	6	3.11	0.30	1.94	0.001*
hu * hra * burn	13	4.84	0.47	1.36	0.001*

3.2 Bacterial Community Composition

The bacterial community composition was assigned using the CyanoSeq 1.3 database. A total of 34 ponds were sequenced with 3910 ASVs. Of these 3881 bacterial and 29 archaeal ASVs with total reads of 972,074 across all the samples. After removal of non-target sequences of 123 chloroplast and 80 mitochondrial ASVs, a total of 3708 ASVs and 951075 reads remained. A total of 50 phyla were observed. The proportional analysis of the phyla was proportioned per pond and across the URSA transect (Table 3.2). The analysis focused on phyla with greater than 1% average proportional reads. There are predominantly five major phyla that comprised of over 98% of the reads.

Table 3.2 Relative Abundances of Phyla across the Ponds of the Utikuma Region Study Area

Phylum	Relative Abundance (%)
Pseudomonadota	49.7
Bacteroidota	26.7
Actinomycetota	10.2
Cyanobacteriota	5.6
Verrucomicrobiota	5.5
Other (45)	2.0

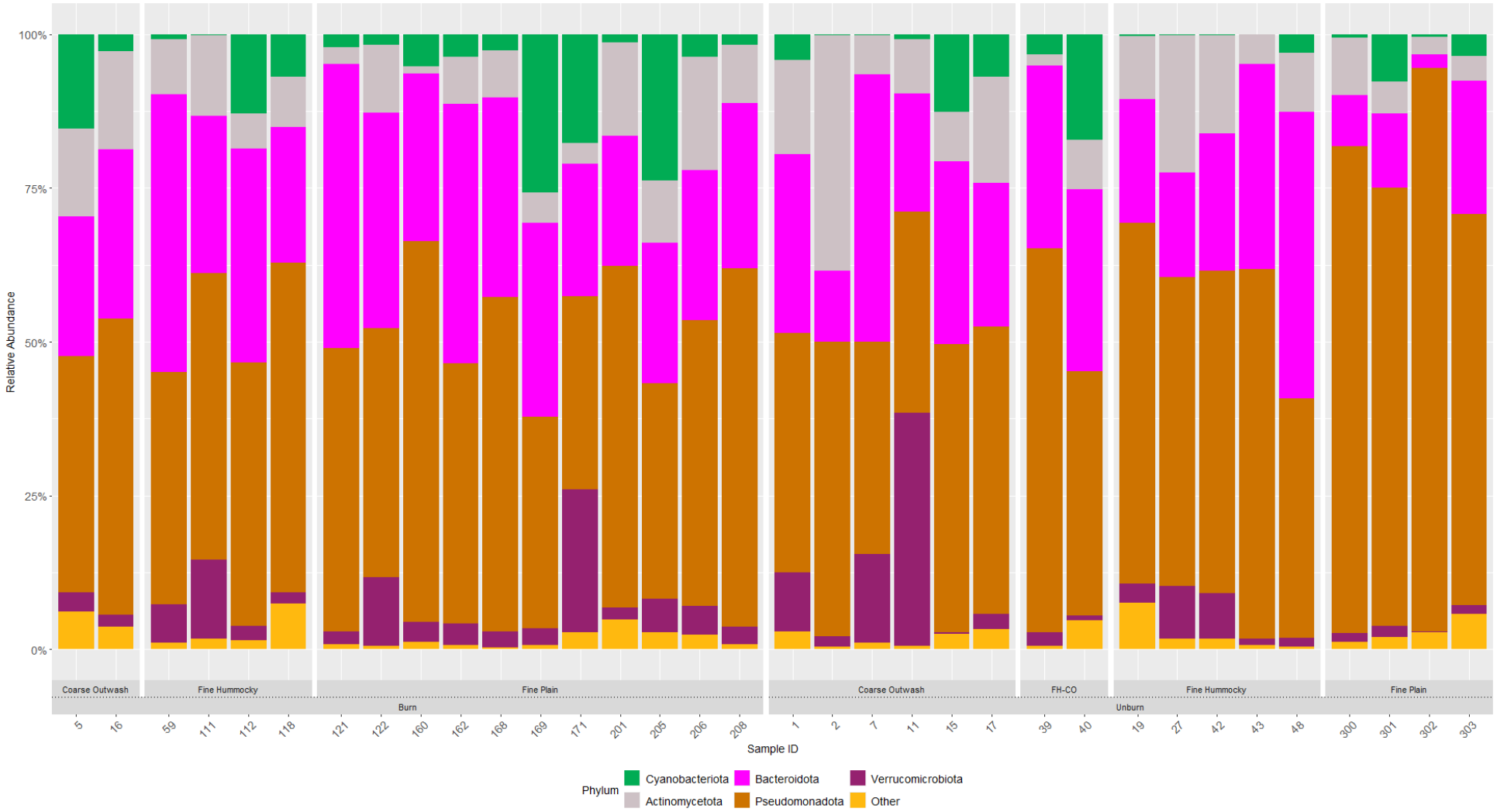


Figure 6 Relative abundance of phyla from URSA 2018. Samples are grouped by their hydrologic relief areas (HRAs) and areas affected by the 2013 wildfires.

3.3 Cyanobacterial Community Composition

A total of 165 unique cyanobacterial Amplicon Sequence Variants (ASVs) were analyzed from the samples, yielding a frequency total of 34,766 reads. Cyanobacteriota represented a mean relative proportion of 3.6% of all reads across the Utikuma Region Study Area (URSA) ponds after filtering for chloroplast and mitochondrial sequences. This proportion indicates a significant presence of cyanobacteria, considering the total read count, and highlights their potential ecological roles in the aquatic ecosystems through processes like nitrogen fixation and primary production.

The cyanobacterial community was characterized based on ASV absolute frequencies, with a summary of taxonomic classifications presented in Table 3.3. Overall, ASVs associated with the Prochlorococcaceae were the most abundant, exhibiting a mean relative abundance of 22.1% and accounting for 35.3% of the total reads (12,261 reads) across 21 samples, represented by 23 unique ASVs. Additionally, ASVs from the Aphanizomenonaceae were also prominent, constituting a mean relative abundance of 12.9% and 16.9% of the reads (5,883 reads), with detections in 22 samples and 13 unique ASVs. The dominance of these two families, known for their diverse ecological roles and potential for bloom formation, underscores their importance within the studied environments.

At the genus level, specific taxa contributed significantly to the overall cyanobacterial abundance. *Vulcanococcus* (Prochlorococcaceae) was highly abundant, contributing 9.4% of the overall abundance (4,669 reads) and present in 14 samples with 6 unique ASVs. *Microcystis* (Microcystaceae) and *Aphanizomenon* (Aphanizomenonaceae) were also highly abundant, exhibiting mean relative abundances of 8.9% (1,879 reads for *Microcystis* and 1,437 reads for *Aphanizomenon*) in 18 and 16 samples, respectively. Other genera, including *Cyanobium* (6.4%

abundance, 2,377 reads, 16 samples, 7 ASVs), *Nodosilinea* (3.5% abundance, 86 reads, 5 samples, 1 ASV), *Limnothrix* (3.1% abundance, 195 reads, 10 samples, 6 ASVs), and *Geminocystis* (3.0% abundance, 10 reads, 2 samples, 2 ASVs), were detected at lower frequencies, indicating a diverse community with varying levels of abundance. *Lacustricoccus* (Prochlorococcaceae) accounted for 2.3% of the mean relative abundance (1,971 reads), was found in 7 samples, and included 6 unique ASVs. Finally, taxa that could not be assigned beyond the family level, noted as Unassigned*, made up 2.2% of the mean relative abundance (425 reads), representing 19 unique ASVs across 17 samples and contributing 1.2% to the total reads.

The distribution of cyanobacterial abundance across samples is further illustrated in Figure 7, which complements the detailed numerical data presented in the supplementary table. As shown in Figure 7, *Vulcanococcus* and unassigned Prochlorococcaceae ASVs consistently exhibit high abundance in several wetland samples. For example, in sample 5, *Vulcanococcus* reached an abundance of 1157 reads, as seen in both the table and the prominent peak in Figure 7. This pattern highlights the dominance of these taxa in specific environmental conditions within the URSA ponds. Furthermore, sample 5 is shown to contain the highest relative abundance of cyanobacterial reads in Figure 6.

Table 3.3 Summary of Cyanobacterial Classification. Unassigned has no classification at the family level.

Taxonomic Classification	Mean Relative Abundance	Total Reads	Sample Count	Unique ASVs	Family	Proportional Reads %
Prochlorococcaceae	**22.1**	**12261**	**21**	**23**	**Prochlorococcaceae**	**35.3**
Aphanizomenonaceae	**12.9**	**5883**	**22**	**13**	**Aphanizomenonaceae**	**16.9**
Vulcanococcus	**9.4**	**4669**	**14**	**6**	**Prochlorococcaceae**	**13.4**
Microcystis	**8.9**	**1879**	**18**	**4**	**Microcystaceae**	**5.4**
Aphanizomenon	**8.9**	**1437**	**16**	**6**	**Aphanizomenonaceae**	**4.1**
Cyanobium	**6.4**	**2377**	**16**	**7**	**Prochlorococcaceae**	**6.8**
Nodosilinea	**3.5**	**86**	**5**	**1**	**Nodosilineaceae**	**0.2**
Limnothrix	**3.1**	**195**	**10**	**6**	**Pseudanabaenaceae**	**0.6**
Geminocystis	**3.0**	**10**	**2**	**2**	**Geminocystaceae**	**0.0**
Lacustricoccus	**2.3**	**1971**	**7**	**6**	**Prochlorococcaceae**	**5.7**
Unassigned	**2.2**	**425**	**17**	**19**	**Unassigned**	**1.2**

Additionally, all 165 unique ASVs were identified and classified against two databases: the CyanoSeq v1.3 and the NCBI BLAST reference (Table B.1). CyanoSeq assignments yielded 44 distinct genus-level names, including well-known cyanobacterial taxa such as *Aphanizomenon*, *Microcystis*, *Nodosilinea*, *Planktothrix*, and *Wollea*, alongside more recently described or less frequently encountered genera like *Johanseniella* and *Romeriopsis*. In contrast, NCBI BLAST results showed 69 genus or species-level designations, with some ASVs matching well-established genera (e.g., *Nostoc*, *Oscillatoria*, *Microcoleus*, *Cylindrospermum*) and others aligning with more specialized or newly described taxa (e.g., *Konicacronema*, *Vampirovibrio*, *Myxacorys*). While both databases recognized many of the same broad cyanobacterial groups, there were discrepancies in nomenclature and specificity. For instance, certain ASVs classified under Prochlorotrichaceae_X or Microcystaceae_X by GTDB were placed into more specific genera or species by NCBI BLAST (e.g., *Microcoleus anatoxicus* or *Microcystis aeruginosa*). Conversely, several ASVs had higher-level classifications in NCBI (e.g., *Pseudanabaena* or *Apatinema*) but were split into finer genera within the GTDB framework (e.g., *Vulcanococcus*, *Limnococcus*). Such differences likely stem from variations in database curation, naming conventions, and updates to taxonomic assignments. In addition, some taxa recognized at the genus level by GTDB were categorized as “unassigned” or placed within provisional lineages by NCBI BLAST, reflecting the dynamic nature of cyanobacterial taxonomy. Several newly described genera appeared only in one database or under different synonyms in the other, underscoring the challenge of maintaining consistent taxonomic references. Overall, the combined results highlight a rich cyanobacterial diversity in the dataset, with both GTDB and NCBI classifications providing complementary perspectives on lineage identity and evolutionary relationships.



Figure 7 Cyanobacterial Genus-level Reads using CyanoSeq v1.3 classifier.

3.4 Cyanobacterial Phylogeny

A maximum likelihood phylogenetic tree was constructed to explore the evolutionary relationships among the 165 cyanobacterial Amplicon Sequence Variants (ASVs) identified in this study, along with reference sequences from the CyanoSeq v1.3 database. The tree was rooted using the Least Common Ancestor (LCA) of Actinomycetota (NR_116348.1) to Firmicutes (NR_118736.1) containing 34 sequences from the CyanoSeq set to contextualize the rest of the cyanobacterial lineage. The final alignment length used for tree construction was 252 base pairs, corresponding to the trimmed V4 region of the 16S rRNA gene. It is important to note that 33 sets of genera could not be resolved at the V4 region based on the CyanoSeq reference sequences after the trimming. These unresolved genera, listed in Table C.1 and Table C.2, represent instances where multiple genera share identical V4 sequences, indicating a limitation in the discriminatory power of this region for certain cyanobacterial taxa at as low as the genus level. This lack of resolution must be considered when interpreting the phylogenetic relationships depicted in Figure 8, as some closely related genera may be grouped together due to the limitations of the V4 region.

The multiple sequence alignment started with 4280 sequences, comprising the 165 ASVs and the CyanoSeq v1.3 reference sequences. After manual trimming and alignment to a 252 bp length of the V4 region, a total of 579 sets of identical sequences were identified, and 2082 redundant sequences were removed, as outlined in the Tables 3.4, C.1 and C.2. This process ensured that only unique sequences were used for phylogenetic tree construction, resulting in a final dataset of 2198 sequences. Specifically, a total of 40 ASVs were removed due to redundancy, with their identical sequences detailed in Table 3.4. This table highlights the diversity of cyanobacterial taxa present in the initial dataset, including genera such *as Lacustricoccus, Microcystis, Vulcanococcus, Arthrospira, Dolichospermum, Pseudanabaena, Woronichinia,*

Heteroscytonema, Radiocystis, Gloeotrichia, Johanseniella, Limnococcus, Limnoraphis, RGUW01, Cuspidothrix, Trichormus, Nostoc, Coelosphaerium, Snowella, Limnothrix, Regnicoccus, Cyanodorina, Aphanizomenon, and *Cyanobium*. Additionally, a noteworthy feature of the tree is a highly supported clade (94% bootstrap support) that groups non-cyanobacterial reference sequences, specifically Bacteroidota (NR_109522) and Tenericutes (NR_041845), amidst the Cyanobacteriota ASVs from URSA (Figure 8). Resolving the true phylogenetic placement and nature of these sequences would necessitate further investigation, potentially using more comprehensive genomic markers or isolation and sequencing of the organisms in question.

The resulting phylogenetic tree (Figure 8) illustrates a diverse array of cyanobacterial orders, as indicated by the color-coded outer ring. Bootstrap support values, represented by triangles at the nodes where size denotes abundance with a cutoff at 75%. This provides an assessment of the robustness of the branching patterns. The tree was annotated using the CyanoSeq v1.3 taxonomy, utilizing order names such as Synechococcales, Nostocales, Oscillatoriales, Pseudanabaenales, Chroococcales, and Gloeobacterales, among others. Chloroplast sequences were retained to ensure all chloroplast ASVs were properly filtered out.

Beyond the detailed analyses of Synechococcales, Chroococcales, Nostocales, Pseudanabaenales, Leptolyngbyales, and *Oscillatoria*les, several additional CYAN ASVs were observed across the phylogenetic tree, distributed among various other orders. These ASVs, while fewer in number, contribute to the overall diversity of the cyanobacterial community. CYAN148 was found within Acaryochloridales, specifically in clade 1 (node INT1939), where it clustered with reference sequence MW172408.1, classified as *Petrachloros* (Petrachlorosaceae). This grouping received strong bootstrap support (82.0), indicating a robust phylogenetic association within this order. CYAN123 was associated with Candidatus-Gastranaerophilales, specifically

within clade 2 (node INT471). This clade, containing CYAN123 and two reference sequences, was supported by a bootstrap value of 79.0, suggesting a well-supported lineage. CYAN116, also within Candidatus-Gastranaerophilales, was located in clade 3 (node INT538). However, this clade exhibited lower bootstrap support (45.0), suggesting a less robust phylogenetic association. CYAN134 was placed within Prochlorothrichales, specifically in clade 1 (node INT1840). This ASV clustered with reference sequence AB863129.1, classified as Prochlorotrichaceae_X (Prochlorotrichaceae), and this grouping was supported by a bootstrap value of 77.0, indicating a moderately strong phylogenetic relationship. These examples illustrate how the limited V4 region can yield both robust and weak phylogenetic signals, thereby complicating the task of classification based on this marker.

Synechococcales represented the most abundant order, with 32 ASVs and 129 reference sequences identified. Phylogenetic analysis revealed that Synechococcales is not monophyletic, with 37 distinct clusters observed. Clade composition varied considerably, reflecting a complex evolutionary history. For example, Clade 1 (node AB610893.1) contained only the reference sequence AB610893.1, classified as *Vulcanococcus* (Prochlorococcaceae), with a bootstrap support of 1.0, indicating a highly specific, but potentially poorly resolved, lineage. Clade 15 (node INT12) contained ASVs CYAN048, CYAN071, and CYAN124, along with reference sequences AF330249.1 (*Vulcanococcus*), AM710351.1 (*Cyanobium*), and D50615.1 (*Vulcanococcus*), but had a low bootstrap support of 24.0, suggesting a grouping with weak statistical confidence. In contrast, Clade 21 (node INT2095) contained nine ASVs (CYAN007, CYAN009, CYAN021, CYAN083, CYAN093, CYAN114, CYAN117, CYAN143, CYAN163) and six reference sequences, with a bootstrap support of 1.0, again indicating a potentially unreliable grouping. These observations underscore that the current naming system, when applied to the short V4

region, often results in the appearance of non-monophyly and ambiguous evolutionary relationships, suggesting that a revision of nomenclature may be necessary.

Chroococcales represented 13 ASVs and 174 reference sequences. Phylogenetic analysis revealed that Chroococcales is not monophyletic, with 23 distinct clusters observed. This fragmentation underscores the complex phylogenetic relationships within this group. Clade 9 (node INT161) contained ASVs CYAN125 and CYAN165, along with eight reference sequences, and had a low bootstrap support of 34.0, suggesting a grouping with limited statistical robustness. Clade 13 (node INT1731) contained five ASVs (CYAN014, CYAN065, CYAN076, CYAN119, CYAN154) and 25 reference sequences, but had a low bootstrap support of 2.0, indicating a highly uncertain phylogenetic relationship. In contrast, Clade 18 (node INT200) contained only reference sequences, MF072345.1 and MF072346.1, both classified as *Cryptococcum* (Chroococcaceae), and had a strong bootstrap support of 100.0, suggesting a well-defined and highly supported lineage. The high number of distinct clusters and varying clade compositions, often with low bootstrap support, highlights the complex and potentially poorly resolved phylogenetic relationships within Chroococcales or more likely the challenges in resolving cyanobacterial phylogeny reflect both the limitations of the 16S V4 region.

Nostocales accounted for the second-highest ASV count, with 30 ASVs and 588 reference sequences. Despite its abundance, the phylogenetic tree showed that Nostocales was divided into only two distinct clusters, indicating a non-monophyletic structure. Clade 1 (node INT1475) contained all 30 ASVs (CYAN004, CYAN024, etc.) and a large number of reference sequences, suggesting a dominant and relatively homogeneous lineage. Clade 2 (node INT353) contained only two reference sequences, KC875344.1 and KC875345.1, both classified as *Chakia* (Scytonemataceae) and had a high bootstrap support of 99.0, indicating a distinct but less

represented lineage. The separation of Nostocales into these two distinct clusters, despite its high sequence diversity, demonstrates its non-monophyletic nature underscoring that even within a limited marker, some robust phylogenetic signals can be discerned if the sequences are sufficiently divergent.

Pseudanabaenales was represented by 5 ASVs and 33 reference sequences. Phylogenetic analysis showed that Pseudanabaenales is not monophyletic, with 4 distinct clusters observed. Clade 2 (node INT432) contained four ASVs (CYAN070, CYAN129, CYAN130, CYAN161) and 26 reference sequences, with a high bootstrap support of 97.0, suggesting a well-supported and robust lineage. Clade 3 (node INT433) contained ASV CYAN082 and reference sequence KF417657.1, classified as Pseudanabaenaceae_X (Pseudanabaenaceae), with a bootstrap support of 100.0, indicating a highly specific and well-defined clade. However, the overall complexity and fragmentation seen within Pseudanabaenales highlight that short sequence lengths may force artificial divisions that do not accurately reflect true evolutionary relationships, prompting a re-evaluation of current taxonomic labels.

Leptolyngbyales were well represented with 11 ASVs and 199 reference sequences, forming 41 distinct clusters. This high number of clusters indicates a highly diverse and potentially rapidly evolving group. For example, clade 8 (node INT100) grouped ASV CYAN106 with several references but showed low bootstrap support of 3.0, suggesting a poorly resolved grouping. Clade 9 (node INT114) included two ASVs with moderate support of 81.0, indicating a more robust lineage. Multiple polyphyletic groups involving Leptolyngbyales, notably with Gloeomargaritales, Oculatellales, and Nodosilineales. These results suggest that many of the apparent divisions within Leptolyngbyales are artifacts of the short V4 region, underlining the need for additional markers or longer sequences to refine the taxonomic framework in this lineage.

Oscillatoriales were represented by 5 ASVs and 210 reference sequences, resolving into 11 distinct clusters. Several clades showed strong support, such as clade 1 (node INT143) with a bootstrap of 96.0 and clade 3 (node INT1599) with a bootstrap of 99.0, indicating well-defined lineages. However, other clades, such as clade 10 (node INT792), received extremely low bootstrap support of 10.0, suggesting poorly resolved relationships. Polyphyly was noted with Chroococcales (node INT144; bootstrap 49.0) and with Coleofasciculales (node INT1619; bootstrap 7.0), reflecting a highly reticulated evolutionary history. These observations collectively point to a critical need for a revised nomenclatural approach that can accommodate the diversity apparent in cyanobacterial communities while accounting for the limited resolving power of the V4 region.

These findings stress that while the 16S V4 region is a valuable tool for preliminary surveys, its limitations must be acknowledged when drawing conclusions about evolutionary relationships and taxonomic classifications in cyanobacteria. The current limitations of the tools and nomenclature present numerous polyphyletic groups across many orders. Examples include Gloeomargaritales/Leptolyngbyales (INT2053), Leptolyngbyales/Oculatellales (INT122), and Candidatus-Gastranaerophilales/Chloroplast (INT1828). Low bootstrap support was common in polyphyletic nodes. This observation is compounded by the clade containing the Tenericutes and Firmicutes amidst the Cyanobacteriota. Together, the lack of resolution of the 16S rRNA gene and classification tools shows a need for further taxonomic revision and careful phylogenetic interpretation.

Table 3.4 Redundant ASVs Removed from Phylogenetic Analysis. The (*) denotes sets containing ASVs.

Classification	Sets	# of Sequences/Set	URSA CYAN ID
g__*Lacustricoccus*	**5**	**1*,2*,1*,1*,2***	**CYAN022, CYAN038, CYAN072, CYAN016, CYAN077**
g__*Microcystis*	**2**	**3*,2***	**CYAN028, CYAN013**
g__*Vulcanococcus*	**6**	**2,3,2*,1*,1*,1***	**CYAN020, CYAN003, CYAN023, CYAN012**
g__*Arthrospira*	**2**	**8,3***	**CYAN144**
g__*Dolichospermum*	**19**	**10,17,2,2,2,2,2,2,2,3,3,5, 7,1*,1*,1*,1*,1*,14***	**CYAN025, CYAN045, CYAN005, CYAN019, CYAN047, CYAN079**
g__*Pseudanabaena*	**7**	**2,2,2,2,3*,1*,1***	**CYAN051, CYAN139, CYAN075**
g__*Woronichinia*	**1**	**2***	**CYAN054**
g__*Heteroscytonema*	**2**	**15,2***	**CYAN155**
g__*Radiocystis*	**1**	**2***	**CYAN063**
g__*Gloeotrichia* **g__*Johanseniella***	**1**	**5***	**CYAN052**
g__*Limnococcus*	**1**	**1***	**CYAN105**
g__*Limnoraphis*	**1**	**18***	**CYAN078**
g__*RGUW01*	**1**	**1***	**CYAN035**
g__*Cuspidothrix*	**5**	**2,3,33,7,8***	**CYAN055**
g__*Trichormus*	**2**	**2,2***	**CYAN080**
g__*Nostoc*	**4**	**2,2,1*,2***	**CYAN097, CYAN032**
g__*Coelosphaerium* **g__*Snowella***	**1**	**3***	**CYAN090**
g__*Limnothrix* **g__*Pseudanabaena***	**1**	**3***	**CYAN094**
g__*Regnicoccus*	**4**	**2,5,2*,3***	**CYAN102, CYAN099**
g__*Cyanodorina*	**1**	**4***	**CYAN067**
g__*Aphanizomenon*	**3**	**2,3,7***	**CYAN011**
g__*Snowella*	**2**	**2,1***	**CYAN101**
g__*Cyanobium*	**3**	**2,2,6***	**CYAN043**

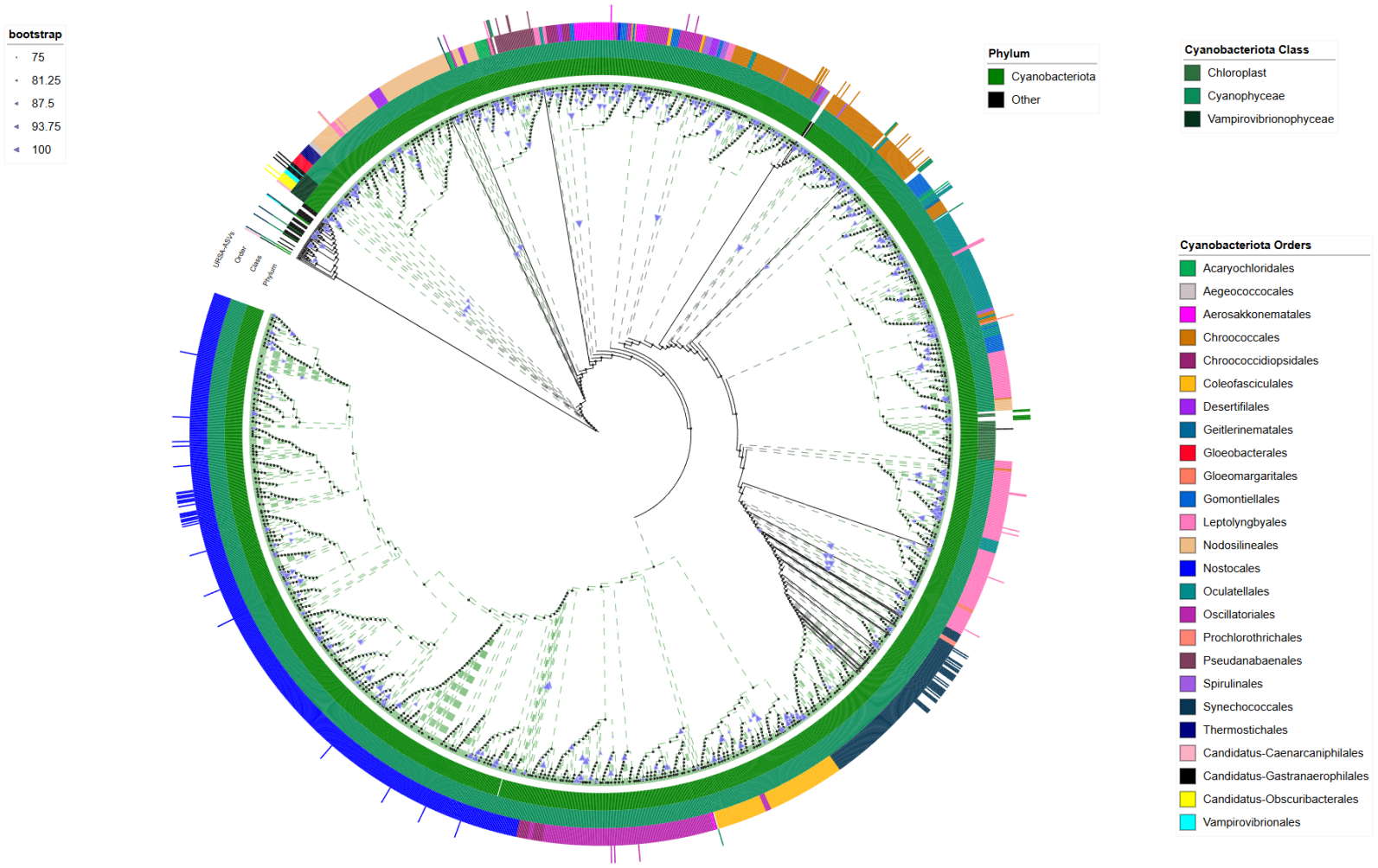


Figure 8 Cyanobacterial Phylogenetic Tree of ASVs from Utikuma Region Study Area 2018. Reference sequences were taken from CyanoSeq v1.3 (Lefler et al., 2023), combined with 165 URSA Cyanobacteriota ASVs aligned using MUSCLE super5 (Edgar, 2022). GTR+Gamma Maximum likelihood tree was constructed using RAxML with support of 1000 bootstraps (Stamatakis, 2014). A total of 2198 sequences were used. Internal nodes are represented by ◆. Bootstrap values are represented by size with a cutoff of 75% support. The rings represent taxonomy (inner to outer: Phyla; Class; Order; URSA CyanoSeq Classification).

3.5 Alpha Diversity

Alpha diversity, as measured by the Shannon Index, was analyzed for bacterial communities across the Utikuma Region Study Area (URSA) ponds. The R package *mirlyn* was utilized for rarefaction and normalization of library size, employing a rarefied library size of 10,000 reads. This threshold ensured that no samples were excluded from the analysis, providing a comprehensive overview of bacterial diversity across all sampled ponds (Figure 9).

The Shannon Index, which accounts for both the richness and evenness of species within a community, revealed a consistent pattern of bacterial diversity across the various pond types. Specifically, bacterial community diversity fell within a range of approximately 2.5 to 5.5, indicating a moderate to high level of diversity within these aquatic ecosystems. Notably, the 'Fine Hummocky' and 'Fine Plain' ponds tended to exhibit slightly higher Shannon Index values compared to the 'Coarse Outwash' ponds, although the differences were not substantial. This suggests that while all pond types support diverse bacterial communities, subtle variations in environmental characteristics may influence the fine-scale structure of these communities.

In summary, the analysis of bacterial alpha diversity in the URSA ponds demonstrates a consistent pattern of moderate to high diversity across different pond types. While subtle variations were observed, the overall findings suggest that these aquatic ecosystems support diverse and resilient bacterial communities. Further investigations, such as beta diversity analyses and functional profiling, may provide deeper insights into the factors shaping bacterial community structure and function in these environments.

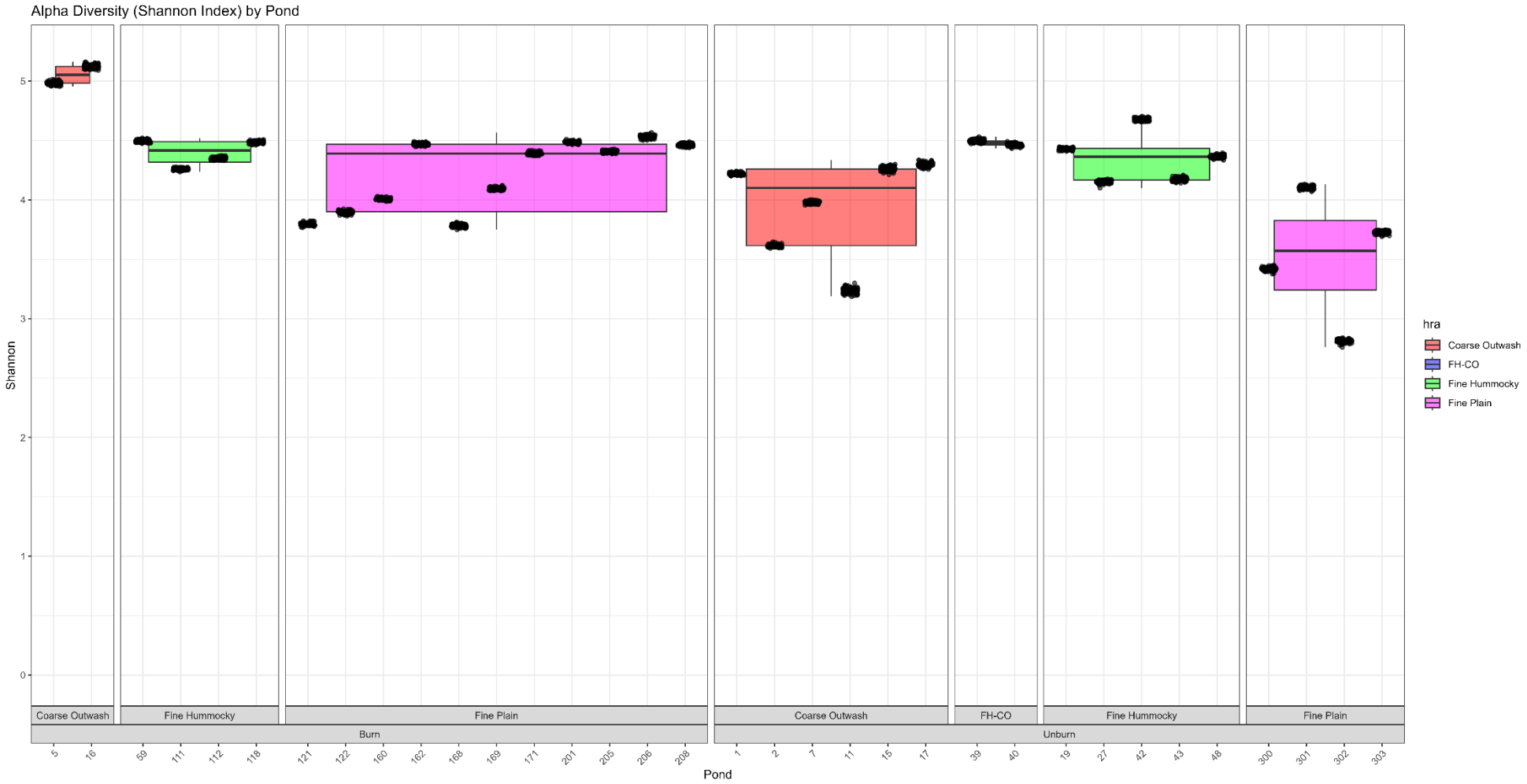


Figure 9 Alpha diversity of bacterial reads from the Utikuma Region Study Area. Shannon Index by pond with rarefied library size of 10,000 with R package *mirlyn*. No samples were excluded after rarefying the library size.

3.6 Beta Diversity

Beta diversity analysis, focusing on the overall bacterial communities, was conducted to explore the compositional differences among the sampled ponds within the Utikuma Region Study Area (URSA). To visualize these differences, both Non-metric Multidimensional Scaling (NMDS) and Principal Coordinates Analysis (PCoA) were employed using Bray-Curtis dissimilarity matrices (Figure 10). Prior to ordination, PERMANOVA was used to assess the statistical significance of categorical variables, specifically hydrologic relief areas (HRA) and areas affected by the 2013 wildfire. These variables were subsequently used to delineate groups within the ordination plots. Convex hulls were utilized to represent the clustering of samples within each category suitable given the sample size.

The NMDS and PCoA ordinations revealed a consistent pattern: the 300-series sampling sites formed a distinct cluster, separate from the remaining URSA ponds, suggesting a unique bacterial community composition within these sites. However, beyond this observation, the ordinations did not clearly delineate patterns based on the informed categories of HRA and burn interactions. The complexity of the NMDS plot, where the 'Fine Plains Burned' convex hull encompassed the 160s/171 series ponds, highlights the intricate nature of bacterial community variation. A systematic examination of each pond within the NMDS ordination revealed a general trend reflecting geospatial proximity, with the notable exception of pond 15. This pond, which exhibited an irregular position relative to its geographically close neighbors (ponds 7 and 11), was recorded to have a shallow depth of 0.3 m and the highest aluminum concentration (68.3 µg/L) among the sampled ponds. These unique environmental characteristics likely contributed to the distinct bacterial community composition observed in pond 15.

The observed geospatial trend in the NMDS ordination, coupled with the utilization of ASV frequencies suggest that spatial factors, such as environmental gradients or dispersal mechanisms, play a role in shaping bacterial diversity within the URSA ponds. This highlights the importance of considering the spatial context when interpreting the patterns seen in bacterial communities. Furthermore, the distinct clustering of the 300-series ponds in both ordinations, despite their varied Shannon diversity (Figure 9), suggests that while these ponds harbor diverse bacterial communities and share a unique compositional profile. This observation underscores the importance of considering both alpha and beta diversity metrics to fully understand bacterial community structure.

In summary, the beta diversity analysis revealed a clear differentiation of the 300-series ponds and highlighted the potential influence of spatial factors and localized environmental conditions on bacterial community composition. Future studies incorporating environmental data and functional analyses may provide a more nuanced understanding of the factors driving bacterial community variation within the URSA ponds

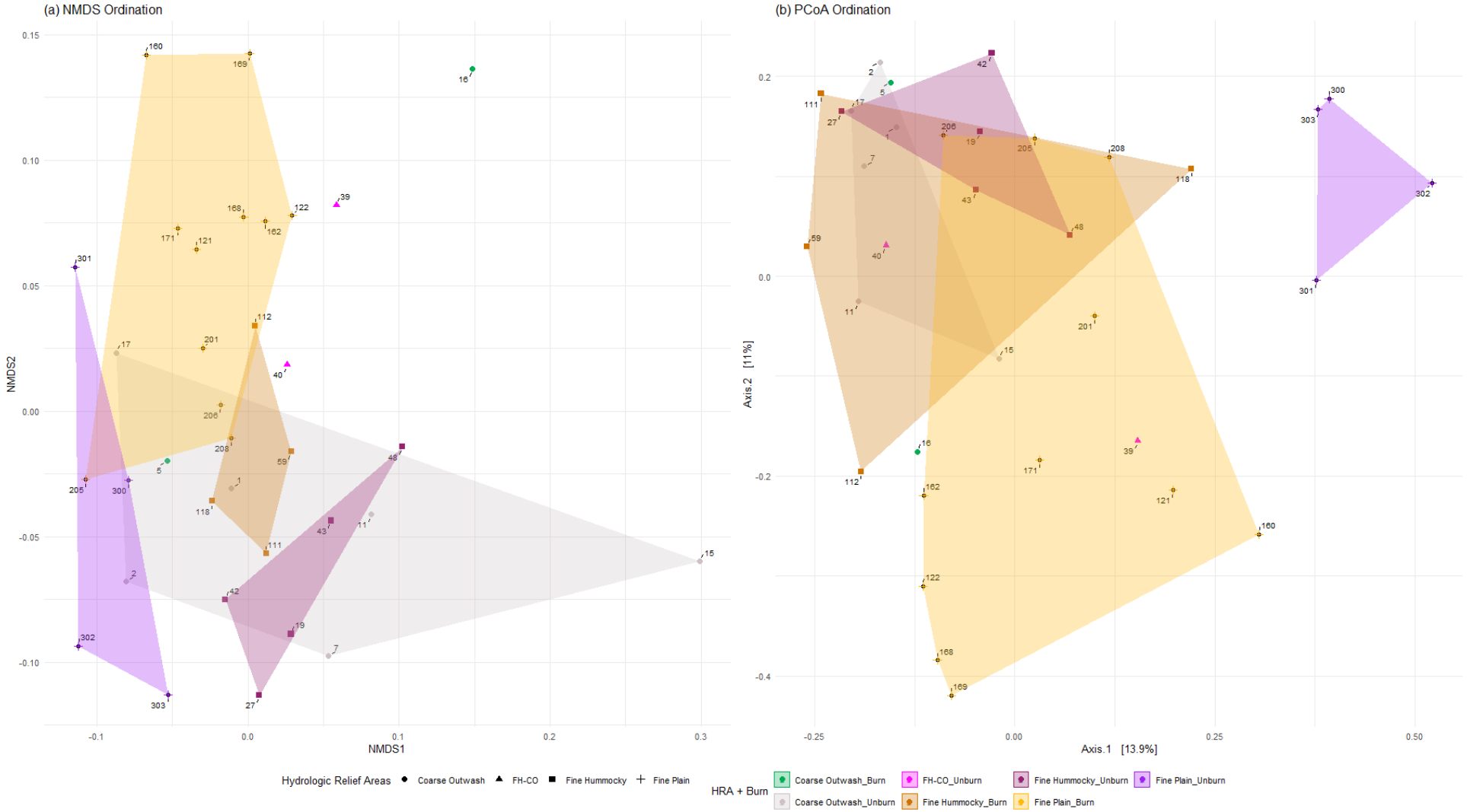


Figure 10 Beta diversity of bacterial communities using Bray-Curtis dissimilarity. a) non-metric and b) PCoA. The categories were chosen using the PERMANOVA to delineate on hydrologic relief areas and regions affected by the 2013 wildfire

Chapter 4. Discussion

This study provides novel insights into the microbial ecology of the Utikuma Region Study Area (URSA) boreal pond landscape, a system characterized by its intricate, low-relief ecohydrology and recognized sensitivity to environmental change (Devito et al., 2012; Turetsky et al., 2011). The URSA's complex ecohydrologic dynamics, shaped by natural factors like its unique geological history and disturbances such as wildfires, make it a critical area for understanding how landscape-scale processes influence water quality, nutrient cycling, and aquatic biodiversity (Devito et al., 2005; Hokanson et al., 2020). A key contribution of this research is its comprehensive molecular characterization of aquatic microbial communities within this specific, complex landscape, an area previously largely unexplored from a microbial perspective. By employing quantitative 16S rRNA gene sequencing and robust multivariate statistical analyses, this study has successfully begun to unravel the relationships between microbial community structure and key ecohydrologic delineations, namely Hydrologic Relief Areas (HRAs) and Hydrologic Units (HUs), as well as the influence of historical wildfire. This investigation offers one of the first such characterizations for the aquatic ecosystems of this particular region, providing a foundational understanding with practical implications for water resource management and the conservation of these sensitive boreal environments. The following sections will discuss the specific findings regarding microbial community composition, the influence of identified environmental drivers, the implications of these findings for ecosystem function, the challenges encountered in cyanobacterial taxonomy, and directions for future research.

4.1 Microbial Communities of the Utikuma Region Study Area

The microbial community composition observed in the URSA is consistent with findings reported in diverse freshwater and boreal ecosystems (Falkowski, Fenchel, & Delong, 2008; Paerl & Otten, 2013; Turetsky et al., 2011). Pseudomonadota, Actinomycetota, Bacteroidota, and Verrucomicrobiota dominate the URSA ponds, a pattern repeatedly observed in studies across temperate and boreal lakes (Newton et al., 2011; Niño-García et al., 2016; Lee et al., 2016). These phyla are known to play pivotal roles in carbon and nutrient cycling, and their prevalence is expected given the unique hydrological, chemical, and disturbance regimes (Turetsky et al., 2011). In addition, the influence of high dissolved organic carbon (DOC) from surrounding peatlands, which shows high variability based on landscape and climate (Pugh et al., 2021), and the impacts of wildfires (e.g., altering dissolved organic matter composition; Olefeldt et al., 2013) act as environmental filters that further select for these dominant groups (Grosbois et al., 2024; Hokanson et al., 2020; Petrone et al., 2016). Together, these multiple lines of evidence strongly support that the microbial assemblages found in URSA are an anticipated outcome of boreal freshwater processes (Falkowski et al., 2008; Paerl & Otten, 2013).

Pseudomonadota, accounting for nearly 50% of the URSA community, have been documented as a dominant phylum in a wide array of freshwater environments (Newton et al., 2011; Lee et al., 2016). Studies in recreational lakes (Lee et al., 2016) and natural freshwater networks (Niño-García et al., 2016) consistently report high relative abundances of Pseudomonadota, which reflects their metabolic versatility and ability to thrive in both oligotrophic and disturbed conditions (Falkowski et al., 2008). Their dominance in URSA is further corroborated by research indicating that fluctuations in nutrient inputs and environmental disturbances, such as wildfires, favor rapid metabolic responses and niche differentiation within

Pseudomonadota (Paerl & Otten, 2013; Petrone et al., 2016). The robust presence of Pseudomonadota is therefore not surprising given similar observations in both temperate and boreal lakes (Newton et al., 2011; Lee et al., 2016). This consistency emphasizes that environmental filtering in freshwater systems is a well-documented phenomenon, reinforcing the URSA findings (Falkowski et al., 2008; Niño-García et al., 2016).

Actinomycetota constitute approximately 10% of the microbial community, a proportion that aligns with prior studies in oligotrophic freshwater ecosystems (Newton et al., 2011; Turetsky et al., 2011). Actinomycetota are renowned for their efficiency in decomposing complex organic compounds under low nutrient conditions and are often linked with high DOC environments found in boreal regions (Grosbois et al., 2024; Li et al., 2012; Newton et al., 2011; Turetsky et al., 2011). Subsequent studies have shown that Actinomycetota are well adapted to nutrient-poor waters and consistently contribute to the recycling of organic matter (Falkowski et al., 2008; Paerl & Otten, 2013). The selective pressure from high allochthonous DOC inputs from peatlands in the URSA is expected to favor such oligotrophic specialists (Turetsky et al., 2011; Devito et al., 2012), making their presence a predictable feature. Thus, the findings are supported by a robust literature base that documents similar patterns across boreal and temperate freshwater systems (Newton et al., 2011; Turetsky et al., 2011).

Bacteroidota and Verrucomicrobia, comprising roughly 27% and 5.5% of the URSA community respectively, further validate established ecological patterns in freshwater habitats (Newton et al., 2011). Bacteroidota are known for their ability to break down high-molecular-weight organic matter and are commonly linked with environments characterized by high DOC, as evidenced in both eutrophic and boreal lakes (Newton et al., 2011; Paerl & Otten, 2013). Similarly, Verrucomicrobia have been frequently observed in freshwater systems and are

associated with various biogeochemical processes, although their ecological roles are still being clarified (Niño-García et al., 2016; He et al., 2017). The relative abundances observed in URSA are therefore in line with these previous studies, suggesting that these phyla are integral to the degradation and recycling of organic matter in boreal systems (Newton et al., 2011; Falkowski et al., 2008). Such findings further underscore that the environmental drivers present in the URSA, including high organic matter inputs and distinct hydrologic connectivity, consistently foster similar microbial community structures (Turetsky et al., 2011; Devito et al., 2012).

Cyanobacteria, while representing a smaller fraction of the overall microbial community (approximately 5.6%), are critical to primary production and nutrient cycling in freshwater environments (Paerl & Otten, 2013; Sánchez-Baracaldo et al., 2022). Although cyanobacterial blooms are frequently reported in nutrient-rich and eutrophic systems (Chen et al., 2024), boreal ecosystems typically exhibit lower cyanobacterial dominance due to their nutrient-limited conditions and high DOC levels (Paerl & Otten, 2013; Sánchez-Baracaldo et al., 2022). The URSA findings are thus consistent with this paradigm, as the modest abundance of Cyanobacteria supports the expected ecological dynamics in nutrient-poor, high DOC boreal ponds (Grosbois et al., 2024; Paerl & Otten, 2013). Moreover, the use of refined taxonomic tools like CyanoSeq has enabled detailed identification of cyanobacterial families, such as Prochlorococcaceae and Aphanizomenonaceae, confirming that even low-abundance groups are functionally significant (Sánchez-Baracaldo et al., 2022; Newton et al., 2011). This alignment with previous literature reinforces that the cyanobacterial community structure in URSA is an anticipated outcome of the unique biogeochemical and hydrologic conditions in boreal lakes (Paerl & Otten, 2013; Turetsky et al., 2011).

In summary, the microbial community structure observed in the URSA is strongly supported by a wealth of literature that documents similar patterns across diverse freshwater and boreal ecosystems (Falkowski et al., 2008; Newton et al., 2011; Paerl & Otten, 2013; Turetsky et al., 2011; Lee et al., 2016). The dominance of Pseudomonadota, coupled with the expected abundances of Actinomycetota, Bacteroidota, and Verrucomicrobiota, is well documented in studies spanning recreational lakes, temperate systems, and boreal networks (Newton et al., 2011; Niño-García et al., 2016; Newton et al., 2011). Additionally, the modest presence of Cyanobacteria in URSA aligns with observations in nutrient-limited, high DOC environments (Grosbois et al., 2024; Sánchez-Baracaldo et al., 2022). This evidence supports that the ecohydrologic and biogeochemical characteristics inherent to boreal systems naturally select for these microbial groups. Future research should incorporate metagenomic, metatranscriptomic, and isotopic techniques to further illuminate the functional dynamics and seasonal shifts of these communities, thereby enhancing the understanding of boreal lake ecosystems and informing effective resource management strategies (Niño-García et al., 2016; Paerl & Otten, 2013).

4.2 Diversity and Influence of Ecohydrologic Factors

The URSA ponds harbored diverse microbial communities, with Shannon Index values generally ranging from approximately 2.5 to 5.5 (Figure 9). Analysis of beta diversity revealed distinct clustering of the 300-series ponds and a general geospatial trend among other sites, though clear delineations based solely on HRA and burn status interactions were not consistently observed beyond this finding (Figure 10). Microbial community composition in these ponds showed significant associations with overall water chemistry conditions, which collectively explained 23.9% of microbial community variance based on clr-transformed RDA (Table D.1; Figure 5b). Landscape-scale ecohydrologic classifications were also strongly linked to microbial community

structure. HRAs, defined by surficial geology (Devito et al., 2005, 2012), were significantly associated with microbial community variance, explaining approximately 16% of this variation (Table 3.1). HRAs are associated with broad zones of differing biogeochemical conditions and processes (Devito et al., 2012; Plach et al., 2016) that shape pond environments, influencing nutrient budgets (Plach et al., 2016) and climate-responsive water level dynamics (Leader et al., 2024a, 2024b).

Finer-scale HUs, characterized by soil and vegetation (Devito et al., 2012), also appeared to contribute to microbial community differences, particularly through interactions; the three-way interaction among HUs, HRAs, and wildfire history was associated with approximately 47% of microbial community variance (Table 3.1). HUs contribute to an environmental mosaic (Devito et al., 2012; Petrone et al., 2016), associated with varied local niches and observed microbial heterogeneity across the URSA landscape. Interactions between Wetland HUs (often peatlands exhibiting ecohydrologic self-regulation; Kettridge & Waddington, 2014; Waddington et al., 2015) and Forestland HUs can create unique hydrologic connections and gradients, sometimes opposing local topography (Hokanson et al., 2020; Petrone et al., 2016), likely fostering diverse conditions for microbial life. These varied environmental settings created by HRAs and HUs provide a foundational context for understanding microbial distributions.

A particularly strong finding was the persistent association between historical wildfire and microbial community composition. The 2011 Utikuma Complex wildfire, especially its interaction with HRA classification, was significantly linked to microbial community structure approximately seven years post-event, with this interaction explaining approximately 30.1% of the variance (Table 3.1). Wildfires cause substantial and lasting alterations to watershed hydrology and biogeochemistry by delivering significant inputs of sediment, ash, nutrients (from fire-altered

soils; Certini, 2005), and altered DOC to aquatic systems, thereby inducing sustained shifts in the biogeochemical environment of ponds (Emelko et al., 2016; Bladon et al., 2014; Emelko et al., 2016; Paul et al., 2022; Silins et al., 2009; Stone et al., 2011). The varying association between wildfire history and microbial communities across different HRAs suggests the broader landscape context is linked to how fire-induced biogeochemical changes are expressed or persist.

This landscape-mediated response to disturbance, alongside baseline environmental settings, contributes to the complex structuring of aquatic microbial life. HRAs, with their inherent variations in characteristics like hydrologic connectivity and water storage (Devito et al., 2012; Leader et al., 2024a; Smerdon et al., 2005), are associated with a differential expression of wildfire's long-term effects on pond environments. For instance, peatlands within certain HRAs can undergo significant long-term vegetation and hydrologic shifts depending on fire severity and return intervals, which in turn alter inputs to associated ponds for decades. The detection of such a long-term microbial legacy underscores the lasting impact of wildfire on these boreal aquatic ecosystems, an impact associated with landscape-scale features. Overall, microbial communities in URSA ponds appear shaped by prevailing water chemistry, the environmental mosaic linked to HRAs and HUs (Devito et al., 2005, 2012), and the persistent biogeochemical impacts of major disturbances like wildfire, the expression of which is also associated with landscape setting.

4.3 Implications for Ecosystem Function and Management

The results indicate that ecohydrologic factors, such as hydrologic relief areas (HRAs) and wildfire history, significantly influence microbial community composition and beta diversity in the URSA ponds, thereby indirectly affecting carbon cycling processes (Walsh et al., 2025; Bladon et al., 2008). Microbial assemblages in URSA ponds with a history of wildfire disturbance were observed to differ compositionally from those in areas without such history. This association

between wildfire history and distinct microbial community characteristics is consistent with broader ecological literature. That literature demonstrates that landscape disturbances, including wildfires (Robinne et al., 2019) and forest management (Foster et al., 2024), can causally alter nutrient dynamics and modulate carbon storage in aquatic ecosystems (Watt et al., 2021)—processes which are, in turn, understood to shape microbial community structure. Alterations in phosphorus bioavailability, which are known to affect carbon fixation rates, may influence the long-term capacity for carbon sequestration in freshwater systems (Watt et al., 2021). The significant interaction between HRAs and wildfire could explain a large portion of the variance in microbial structure, suggesting that disturbances may reduce the capacity for stable carbon storage by favoring taxa less involved in carbon stabilization, potentially shifting them from sinks to sources (Paerl & Otten, 2013; Turetsky et al., 2011; Foster et al., 2024). Northern peatlands are recognized as critical carbon reservoirs that are vulnerable to warming and disturbance, thereby underscoring the broader implications for carbon storage in boreal aquatic systems (Yu, 2012). These findings highlight the potential for climate warming and disturbance-induced shifts in microbial communities to create feedback loops that weaken carbon sequestration and accelerate ecosystem-level carbon losses (Turetsky et al., 2011).

The redundancy analysis and PERMANOVA results demonstrated that key water chemistry parameters—such as pH, Fe, NO_2/NO_3, Chlorophyll *a*, and Al—explain a significant proportion of the variance in microbial community composition following appropriate normalization of the data (Paerl et al., 2011; Ji et al., 2019). Altered nutrient bioavailability due to landscape disturbances has been shown to have cascading effects on water quality and ecosystem function, particularly in aquatic systems (Orlova et al., 2020). Sediment dynamics, including the erosion and transport of fine-grained material, play a critical role in mediating these nutrient effects, as shifts

in hydrologic connectivity can influence microbial community structure (Krishnappan et al., 2020). Long-term shifts in nutrient patterns, driven by anthropogenic pressures and climate change, have been associated with increases in harmful algal blooms (Song et al., 2023), which further impact ecological processes and nutrient cycling in aquatic environments (Zhang et al., 2022). Nutrient stress itself can influence toxin production in key genera like *Microcystis* (Pimentel & Giani, 2014). Wildfire-induced changes in sediment properties can lead to prolonged phosphorus release (Bladon et al., 2014), potentially altering downstream nutrient cycling and increasing risks to oligotrophic systems (Watt et al., 2021). An integrated approach that links ecohydrologic frameworks with microbial diversity metrics provides valuable insights into the interactions among nutrient cycling, sediment dynamics, and overall ecosystem resilience, highlighting areas for targeted management and future research (Bladon et al., 2014; Krishnappan et al., 2020; Watt et al., 2021).

4.4 Taxonomic Challenges of Cyanobacteria

The phylogenetic analysis of cyanobacterial ASVs in the URSA ponds revealed several instances of polyphyletic groupings that have significant implications for understanding community diversity and function. Within the Synechococcales, the results identified 37 distinct clusters. One striking case is Clade 15 (node INT12), which grouped ASVs such as CYAN048, CYAN071, and CYAN124 with reference sequences for both *Vulcanococcus* and *Cyanobium*. This clade exhibited very low bootstrap support (24.0), indicating a weak phylogenetic resolution and highlighting the inability of the 16S rRNA V4 region to confidently delineate closely related taxa (Di Cesare et al., 2018; Komárek, 2023; Komárek et al., 2020). Similarly, in the Chroococcales, Clade 9 (node INT161) contained ASVs CYAN125 and CYAN165 along with several reference sequences, yet this grouping also had low bootstrap support (34.0), suggesting limited statistical

robustness (Müller et al., 2017). In contrast, Nostocales presented a more complex scenario where a dominant cluster exhibited strong support, but a secondary cluster—comprising reference sequences such as KC875344.1 and KC875345.1—displayed high bootstrap support (99.0) that indicated distinct evolutionary signals. These polyphyletic patterns, as also emphasized by Strunecký et al. (2023) Komárek (2016; 2023), Komárek et al. (2020), and Moreira et al. (2013), directly affect the capacity to accurately assess the diversity and ecological roles of Cyanobacteria in the URSA ponds. Furthermore, the observed phylogenetic ambiguities and instances of polyphyly may not solely stem from the inherent limitations of the V4 marker but could also be influenced by uncertainties in the multiple sequence alignment. The precise placement of gaps (indels) in these variable regions, dictated by alignment algorithm parameters, can significantly impact inferred homologies and thus the resulting tree topology.

In addition to the polyphyletic groupings, the analysis revealed a noteworthy proportion of "Unassigned" taxa that could not be confidently classified beyond the family level. The presence of these unassigned ASVs suggests the potential existence of novel cyanobacterial lineages or an underrepresentation of certain taxa in current reference databases (Jung, 2023; Li et al., 2020). Strunecký et al. (2023) discuss how the limitations of the V4 region and the dynamic nature of cyanobacterial evolution contribute to such unresolved classifications. This lack of taxonomic resolution may obscure important ecological functions, such as roles in nitrogen fixation or primary production, or the potential for toxin production under certain conditions (Pearson et al., 2016; Pimentel & Giani, 2014) and could result in an underestimation of true diversity. These challenges are compounded by the general limitations inherent to the 16S rRNA gene marker itself, particularly when using short-read amplicons like the V4 region. While cost-effective for broad surveys, this region may not always possess sufficient phylogenetic signal to resolve closely

related taxa or delineate species boundaries accurately, contributing to the observed ambiguities and the difficulties in achieving a stable cyanobacterial taxonomy based solely on this marker. Moreover, discrepancies in nomenclature between databases—where, for example, the curated CyanoSeq 1.3 database and NCBI BLAST produce different genus-level designations—further complicate interpretations (Di Cesare et al., 2018; Roush et al., 2021). Hence, addressing the unassigned taxa is critical for refining understanding of cyanobacterial community dynamics and their ecological impacts in the URSA.

A significant challenge in the work stems from discrepancies between taxonomic databases. The CyanoSeq database, which utilizes a hybrid nomenclature combining names from both NCBI and the Genome Taxonomy Database (GTDB), yields different taxonomic assignments compared to NCBI BLAST alone (Jung, 2023; Komárek, 2016). One example is classification with CyanoSeq results in 44 distinct genus-level names, whereas NCBI BLAST produced 69 designations. These differences are attributable to variations in curation standards, naming conventions, and database updates. Discrepancies are compounded by the inherent limitations of short-read sequencing regions, which often fail to resolve fine-scale differences among cyanobacterial taxa (Strunecký et al., 2023). The mixed nomenclature and inconsistent taxonomic assignments not only hinder cross-study comparisons but also affect downstream ecological interpretations. This situation underscores the urgent need for standardized taxonomic frameworks and integrated multi-marker approaches that can harmonize results across different databases (Di Cesare et al., 2018; Parks et al., 2022).

Complexity compounds when attempting to compare the results with other studies that utilize different analytical workflows and taxonomic classifiers. Variability in molecular techniques, primer sets, and bioinformatics pipelines often leads to discrepancies in taxonomic

resolution and community profiles (Komárek, 2016; MacKeigan et al., 2022). Some studies may use full-length 16S rRNA gene sequencing combined with multi-marker approaches, while others rely solely on the V4 region; this disparity makes direct comparisons challenging. The issues raised by Strunecký et al. (2023) further illustrate that even within similar methodologies, rapid evolutionary dynamics and horizontal gene transfer can produce conflicting taxonomic outcomes. To overcome these challenges, future research should integrate environmental data with metagenomic and metatranscriptomic approaches, offering a more holistic understanding of cyanobacterial community dynamics. Additionally, the development of a unified, continuously updated taxonomic database is essential to standardize nomenclature and improve comparability across studies. In conclusion, addressing these taxonomic challenges is critical not only for advancing understanding of cyanobacterial evolution and ecology but also for informing effective water resource management strategies in the URSA and similar ecosystems (Li et al., 2020; Parks et al., 2022; Strunecký et al., 2023).

4.4.1 Picocyanobacteria and Prochlorococcaceae

Picocyanobacteria, often represented within the Prochlorococcaceae family, are fundamental to aquatic ecosystems due to their substantial role in primary production and nutrient cycling, even under oligotrophic conditions (Callieri et al., 2022; Scanlan et al., 2009). In the URSA ponds, the cyanobacterial community was notably dominated by ASVs associated with Prochlorococcaceae, which contributed 22.1% of the relative abundance and included genera such as *Vulcanococcus* and *Cyanobium*. For example, in sample 5, *Vulcanococcus* reached an abundance of 1,157 reads, underscoring its ecological prominence within these systems. This high abundance suggests that picocyanobacteria play a critical role in carbon fixation and nutrient dynamics, acting as essential components of the microbial loop in boreal freshwater ecosystems. However, as highlighted in the

literature, the classification of these small Cyanobacteria remains challenging due to the limited resolution of the 16S rRNA V4 region and the potential confounding effects of horizontal gene transfer (Callieri, Cabello-Yeves, & Bertoni, 2022; Strunecký et al., 2023; Komárek et al., 2020). These taxonomic challenges complicate the ability to distinguish closely related picocyanobacterial taxa and have important implications for understanding their specific ecological roles and adaptive strategies in fluctuating environments.

The taxonomic ambiguity within Prochlorococcaceae is further compounded by discrepancies between different reference databases. This study observed that the curated CyanoSeq 1.3 database and NCBI BLAST sometimes provided conflicting assignments, with some ASVs being classified under *Vulcanococcus* in one system and as *Cyanobium* in the other. This inconsistency mirrors findings from previous research that emphasize the limitations of using a single genetic marker, like the V4 region, which can result in unresolved or mis-assigned taxa (Callieri, Cabello-Yeves, & Bertoni, 2022; Lefler et al., 2023). As picocyanobacteria are key indicators of ecosystem health due to their sensitivity to nutrient availability and light conditions, these classification challenges hinder accurate assessment of both diversity and function. Moreover, the presence of "Unassigned" taxa within the Prochlorococcaceae suggests that there may be novel or under-characterized lineages that current databases do not yet capture. This underscores the need for a polyphasic approach that combines molecular, morphological, and ecological data to refine cyanobacterial taxonomy (Strunecký et al., 2023; Lv et al., 2022).

Addressing these challenges is crucial for improving ecological assessments and management strategies. Integrating multi-marker and whole-genome sequencing approaches can help overcome the resolution limits inherent in the 16S rRNA V4 region and provide more robust phylogenetic placements for picocyanobacteria (Callieri, Cabello-Yeves, & Bertoni, 2022; Scanlan

et al., 2009). These integrative methods have allowed for cross-validation of taxonomic assignments, thereby reducing ambiguities that arise from database discrepancies. Furthermore, incorporating environmental data with metagenomic approaches can elucidate the adaptive traits that enable picocyanobacteria to thrive under nutrient-limited conditions, shedding light on their ecological functions and contributions to biogeochemical cycles. Ultimately, refining the taxonomic framework for Prochlorococcaceae enhances understanding of picocyanobacterial diversity in the URSA ponds and informs future research directions and water resource management practices in similar aquatic ecosystems (Callieri, Cabello-Yeves, & Bertoni, 2022; Scanlan et al., 2009; Strunecký et al., 2023).

The discussion has highlighted the dominant microbial groups within the URSA ponds, the significant influence of ecohydrologic factors and wildfire history on community structure, and the persistent taxonomic challenges, particularly for cyanobacteria. These insights collectively underscore the complexity of these boreal aquatic ecosystems. The research presented provides a basis for continued exploration into the boreal ecosystem and elucidation of the important effects of geospatial and hydrologic features on the formation of microbial communities by analyzing bacterial community compositions, abundance, and the impacts of the watershed. These findings mirror patterns observed in other temperate and boreal ecosystems. The identification of Cyanobacteria in these regions helps build an understanding of what to expect in remote areas away from anthropogenic effects, which could help inform the benefits of environmental features. Ultimately, this work provides critical insights that can inform targeted water quality management and conservation strategies, while also setting the stage for future research aimed at unraveling the functional dynamics underlying these microbial communities, as will be further elaborated in the concluding chapter.

Chapter 5. Conclusions

This study embarked on a molecular characterization of bacterial and cyanobacterial communities within the ponds of the Utikuma Region Study Area (URSA), aiming to understand their composition, diversity, and the environmental factors that shape them. The research provides foundational insights into these sensitive boreal aquatic ecosystems. This chapter summarizes the principal findings, draws overarching conclusions, discusses methodological considerations, highlights the significance of the work, and suggests avenues for future research.

5.1 Summary of Key Findings

The investigation into the drivers of microbial community composition and diversity in the URSA ponds revealed several interconnected findings. Key water chemistry parameters were identified as significant correlates, collectively accounting for approximately 24% of the observed bacterial community variation using a clr-transformed RDA model. Beyond local water chemistry, the broader landscape context played a crucial role. PERMANOVA demonstrated that HRAs representing distinct broad-scale geological and topographical zones were a highly significant factor, independently explaining approximately 16% of community variance. Notably, historical disturbance from the 2011 Utikuma Complex wildfire also emerged as a significant individual factor (explaining ~5% of variance) and, more importantly, its interaction with HRAs was found to explain a substantial portion (over 30%) of the community variance. A three-way interaction among HRAs, finer-scale HUs, and this wildfire history was associated with nearly half (approximately 47%) of the variation in microbial community composition. These strong interactive effects suggest that microbial communities in URSA ponds are associated with a complex interplay of underlying geology, surficial hydrologic units, and the lasting legacy of historical disturbances like wildfire. In terms of microbial diversity, within-pond alpha diversity

was generally moderate to high. Beta diversity analyses indicated distinct clustering patterns, such as the separation of the '300-series' ponds, and a general geospatial trend; however, clear delineations based solely on HRA and wildfire history interactions were not consistently observed across all sites.

5.2 Main Implications

Based on the findings of this research, several main conclusions and implications can be drawn regarding the microbial ecology of the URSA ponds. Firstly, the overall bacterial communities identified are characteristic of those expected within boreal aquatic ecosystems, adapted to prevalent conditions such as high dissolved organic carbon (DOC) concentrations and specific nutrient regimes. Secondly, cyanobacteria, while a modest fraction of the total community, include dominant picocyanobacteria (e.g., Prochlorococcaceae) typical of nutrient-poor boreal waters, and the widespread detection of potentially toxigenic genera (e.g., *Microcystis*, *Aphanizomenon*) provides a crucial baseline for water quality management.

A third major conclusion is that microbial community structure in URSA ponds is significantly shaped by a hierarchy of factors, from local water chemistry to broader landscape-scale ecohydrologic variables. The pronounced influence of Hydrologic Relief Areas (HRAs), and especially their strong interaction with the legacy of the 2011 wildfire—an impact detectable approximately seven years post-event—underscores not only the critical importance of URSA's unique ecohydrologic context for predicting microbial community patterns but also the potential for major disturbances to impart long-lasting ecological signatures on these aquatic systems. These strong correlations suggest that effective water resource management in such boreal regions must incorporate this landscape-level perspective and an understanding of disturbance history. Finally, this study reiterated the known taxonomic challenges associated with using the 16S rRNA gene

V4 region for detailed microbial classification, particularly for complex groups like cyanobacteria, highlighting the need for careful interpretation and integrated taxonomic approaches in future studies.

5.3 Significance of the Research

This research holds particular significance as it provides one of the first detailed molecular characterizations of the bacterial and cyanobacterial communities specifically within the complex pond ecosystems of URSA, a landscape previously largely uncharacterized from a microbial perspective. This work establishes an important ecological baseline for a region of the Boreal Plains known for its intricate hydrology and sensitivity to environmental change. A key novel contribution is the successful quantitative demonstration of associations between microbial community structure and unique ecohydrologic classifications (HRAs and HUs) and, critically, the detection of a persistent ecological legacy from a wildfire event that occurred approximately seven years prior to sampling. Identifying such a long-term impact from a single-timepoint, spatially distributed survey is a noteworthy finding, as studies demonstrating such lasting microbial distinctions in comparable aquatic systems, particularly from snapshot data, are uncommon in the literature on wildfire effects. The study's ability to discern these associations within a landscape characterized by complex hydrogeology, using robust statistical approaches, underscores its significance. This understanding of how local environmental variables and broader landscape-scale ecohydrologic factors—and their interaction with major disturbances—are associated with microbial community structure is crucial for developing more nuanced models of boreal ecosystem function and for informing future monitoring programs and targeted water quality management strategies within URSA and similar boreal landscapes.

5.4 Future Research Directions

While this study provides valuable foundational insights, future research should aim to build upon this work to achieve a deeper and more dynamic understanding of URSA's microbial communities and their environmental drivers. The current findings highlight the significant influence of water chemistry variables and ecohydrologic factors on microbial community composition, yet several avenues exist to refine and expand upon these results.

To better integrate water chemistry and ecohydrologic drivers, future investigations would benefit from increasing temporal resolution through continuous monitoring and high-frequency sampling. This approach would provide a more dynamic understanding of how variables such as pH, NH_4, NO_2/NO_3, Al, Fe, and Chlorophyll a fluctuate over time and affect microbial dynamics (Zamyadi et al., 2012; Peura et al., 2015). Furthermore, studies may benefit from integrating multi-omics approaches, including metagenomics, metatranscriptomics, and metabolomics, to capture both the functional potential and active processes within these communities (Niño-García et al., 2016), such as pathways for toxin biosynthesis or adaptation (Pearson et al., 2016; Moreira et al., 2013). Enhanced spatial resolution in ecohydrologic classifications, for instance through more detailed delineations of hydrologic relief areas and hydrologic units, would also improve the ability to understand associations between landscape features and microbial responses to environmental disturbances like wildfires (Ho & Michalak, 2020). Research could also focus on developing predictive models that couple environmental parameters with microbial community shifts to anticipate harmful algal bloom formation and nutrient cycling changes in boreal ecosystems. The incorporation of advanced statistical and computational tools to analyze interactions among chemical drivers and ecohydrologic factors holds promise for identifying critical thresholds and feedback mechanisms in these systems (Zheng et al., 2022). Importantly, the patterns and

significant environmental correlates—including specific HRAs, wildfire history, and key water chemistry variables—identified in this exploratory study can now serve to inform more targeted, hypothesis-driven sampling strategies for future research aimed at deeper mechanistic understanding of the specific biogeochemical processes within different landscape units that drive these microbial patterns.

Regarding the standardization and enhancement of cyanobacterial classification, it is crucial to adopt more integrative and multi-marker approaches to overcome the challenges highlighted by polyphyletic observations and the prevalence of unassigned taxa. While this study demonstrated the benefits of incorporating the curated CyanoSeq 1.3 database, expanding analyses to include additional genetic markers (e.g., ITS regions, functional genes) or full-length 16S rRNA gene sequencing could improve taxonomic resolution and reduce ambiguities (Li et al., 2020; Parks et al., 2022). The successful application of such polyphasic methods, for example, using ITS or protein-coding gene data alongside 16S rRNA as demonstrated by Lv et al. (2022) in resolving Microcystis/Radiocystis taxonomy or as applied broadly by Komárek et al. (2020) for Synechococcus-like taxa, is essential for robust classification and accurate ecological assessment in future studies. Integrating environmental data with metagenomic and metatranscriptomic approaches may also provide a more holistic understanding of cyanobacterial community dynamics and functional roles, ultimately bridging the gap between molecular data and ecosystem processes (MacKeigan et al., 2022; Roush et al., 2021). Furthermore, the development of a unified, regularly updated taxonomic database is essential to harmonize nomenclature and facilitate more reliable comparisons across studies that use different workflows and classifiers (Komárek, 2016; Komárek et al., 2023; Jung, 2023). Standardization in molecular workflows and classification

methodologies will not only enhance reproducibility but also enable researchers to draw more consistent and accurate ecological conclusions.

Finally, addressing methodological considerations and striving for a broader ecosystem understanding will be key. This includes continually addressing the limitations inherent in current methodological approaches, such as the challenges associated with compositional data transformations (e.g., clr transformations) and taxonomic resolution using 16S rRNA gene sequencing. Extending studies to include a broader range of ecosystems and long-term monitoring would support the generalization of these findings and inform more effective water quality management and conservation strategies (Martiny et al., 2006). Field or mesocosm experiments could also be designed to directly test the impact of specific environmental changes on community structure and function, moving beyond correlative findings.

5.5 Concluding Remarks

Bacterial communities of the Utikuma Region Study Area ponds are diverse, structured by a complex interplay of local water chemistry, landscape-scale ecohydrologic features (HRAs, HUs), and historical wildfire disturbances. This research provides critical first insights by quantitatively demonstrating microbial structure to unique landscape classifications and demonstrating a persistent, multi-year wildfire legacy in this complex boreal setting. It establishes an important ecological baseline for a region susceptible to environmental change. While highlighting potentially toxigenic Cyanobacteria and challenges in their taxonomic resolution with standard markers, the findings advocate for considering the broader ecohydrologic context, including disturbance history, to understand and manage these vital boreal aquatic ecosystems. Adopting more integrative, higher-resolution approaches in the future, guided by these initial

patterns, will facilitate a more complete understanding of the functional roles, resilience, and future trajectory of these important microbial communities.

References

Abraham, J., Dowling, K., & Florentine, S. (2017). Risk of post-fire metal mobilization into surface water resources: A review. *Science of the Total Environment*, *599–600*, 1740–1755. https://doi.org/10.1016/j.scitotenv.2017.05.096

Acinas, S. G., Haverkamp, T. H. A., Huisman, J., & Stal, L. J. (2009). Phenotypic and genetic diversification of *Pseudanabaena* spp. (cyanobacteria). *The ISME Journal*, *3*(1), 31-46. https://doi.org/10.1038/ismej.2008.78

Alessio, P., Dunne, T., & Morell, K. (2021). Post-wildfire generation of debris-flow slurry by rill erosion on colluvial hillslopes. *Journal of Geophysical Research: Earth Surface*, *126*(e2021JF006108). https://doi.org/10.1029/2021JF006108

Aljoudi, G. (2021). *Bacterial communities in the Glenmore Reservoir with an emphasis on the Cyanobacteria* (Master's thesis, University of Waterloo).

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389-3402. https://doi.org/10.1093/nar/25.17.3389

Batista, A. M. M., & Giani, A. (2019). Spatiotemporal variability of cyanobacterial community in a Brazilian oligomesotrophic reservoir: The picocyanobacterial dominance. *Ecohydrology & Hydrobiology*, *19*(4), 566-576. https://doi.org/10.1016/j.ecohyd.2019.03.002

Bladon, K. D., Emelko, M. B., Silins, U., & Stone, M. (2014). Wildfire and the future of water supply. *Environmental Science & Technology*, *48*(16), 8936–8943. https://doi.org/10.1021/es500130g

Bladon, K. D., Silins, U., Wagner, M. J., Stone, M., Emelko, M. B., Mendoza, C. A., Devito, K. J., & Boon, S. (2008). Wildfire impacts on nitrogen concentration and production from headwater streams in southern Alberta's Rocky Mountains. *Canadian Journal of Forest Research*, *38*(9), 2359-2371. https://doi.org/10.1139/x08-071

Bokulich, N. A., Kaehler, B. D., Rideout, J. R., Dillon, M., Bolyen, E., Knight, R., Huttley, G. A., & Caporaso, J. G. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, *6*(90). https://doi.org/10.1186/s40168-018-0470-z

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., ... 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

Brand, T. (2024). *ggh4x* (Version 0.2.8) [R package]. Comprehensive R Archive Network. https://teunbrand.github.io/ggh4x/

Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, *27*(4), 326-349.

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*, 581–583. https://doi.org/10.1038/nmeth.3869

Callieri, C. (2008). Picophytoplankton in freshwater ecosystems: The importance of small-sized phototrophs. *Freshwater Reviews, 1*(1), 1–28. https://doi.org/10.1608/FRJ-1.1.1

Callieri, C., Cabello-Yeves, P. J., & Bertoni, F. (2022). The "dark side" of picocyanobacteria: Life as we do not know it (yet). *Microorganisms*, *10*(3), 546. https://doi.org/10.3390/microorganisms10030546

Cameron, E. S., Krishna, A., Emelko, M. B., & Müller, K. M. (2024). Sporadic diurnal fluctuations of cyanobacterial populations in oligotrophic temperate systems can prevent accurate characterization of change and risk in aquatic systems. *Water Research*, *252*, 121199. https://doi.org/10.1016/j.watres.2024.121199

Cameron, E. S., Müller, K. M., Stone, M., Buttle, J., Leach, J., Webster, K., & Emelko, M. B. (2022). Early seasonal increases and persistence in relative abundance of potentially toxic Cyanobacteria: Concerning impacts of extended ice-free periods in northern temperate lakes. *BioRxiv*. https://doi.org/10.1101/2022.12.20.521158

Cameron, E. S., Schmidt, P. J., Tremblay, B. J.-M., Emelko, M. B., & Müller, K. M. (2021). Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Scientific Reports*, *11*, 22302. https://doi.org/10.1038/s41598-021-01735-5

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(Suppl 1), 4516–4522. https://doi.org/10.1073/pnas.1000080107

Certini, G. (2005). Effects of fire on properties of forest soils: a review. Oecologia, 143(1), 1–10. https://doi.org/10.1007/s00442-004-1788-8

Chen, L., Zhang, F., Dong, Z., Zhang, W., Sun, T., & Chen, L. (2024). Response and acclimation of Cyanobacteria to acidification: A comprehensive review. *Science of The Total Environment*, *945*, 173978. https://doi.org/10.1016/j.scitotenv.2024.173978

Cheung, M. Y., Liang, S., & Lee, J. (2013). Toxin-producing Cyanobacteria in freshwater: A review of the problems, impact on drinking water safety, and efforts for protecting public health. *Journal of Microbiology*, *51*(1), 1–10. https://doi.org/10.1007/s12275-013-2549-3

Chorus, I., & Welker, M. (Eds.). (2021). *Toxic Cyanobacteria in water*. CRC Press. https://doi.org/10.1201/9781003081449

Crouch, R. L., Timmenga, H. J., Barber, T. R., & Fuchsman, P. C. (2006). Post-fire surface water quality: Comparison of fire retardant versus wildfire-related effects. *Chemosphere*, *62*(6), 874–889. https://doi.org/10.1016/j.chemosphere.2005.05.031

Demoulin, C. F., Lara, Y. J., Cornet, L., François, C., Baurain, D., Wilmotte, A., & Javaux, E. J. (2019). Cyanobacteria evolution: Insight from the fossil record. *Free Radical Biology and Medicine*, *140*, 206-223. https://doi.org/10.1016/j.freeradbiomed.2019.05.007

Devito, K. J., Mendoza, C., Petrone, R. M., Kettridge, N., & Waddington, J. M. (2016). Utikuma Region Study Area (URSA) – Part 1: Hydrogeological and ecohydrologic studies (HEAD). *The Forestry Chronicle*, *92*(1), 7-17. https://doi.org/10.5558/tfc2016-017

Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., & Smerdon, B. (2005). A framework for broad-scale classification of hydrologic response units on the Boreal Plain: Is topography the last thing to consider? *Hydrologic Processes*, *19*(8), 1705–1714. https://doi.org/10.1002/hyp.5881

Devito, K., Mendoza, C., & Qualizza, C. (2012). *Conceptualizing water movement in the Boreal Plains. Implications for watershed reconstruction*. Canadian Oil Sands Network for Research and Development, Environmental and Reclamation Research Group.

Di Cesare, A., Cabello-Yeves, P. J., Chrismas, N. A. M., Sánchez-Baracaldo, P., Salcher, M. M., & Callieri, C. (2018). Genome analysis of the freshwater planktonic *Vulcanococcus limneticus* sp. nov. reveals horizontal transfer of the nitrogenase operon and alternative pathways of nitrogen utilization. *BMC Genomics*, *19*, 259. https://doi.org/10.1186/s12864-018-4648-3

Du, X., Liu, H., Yuan, L., Wang, Y., Ma, Y., Wang, R., Chen, X., Losiewicz, M., Guo, H., & Zhang, H. (2019). The diversity of cyanobacterial toxins: Structural characterization, distribution and identification—A systematic review. *Toxins*, *11*(9), 530. https://doi.org/10.3390/toxins11090530

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

Edgar, R. C. (2022). Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nature Communications*, *13*, 6968. https://doi.org/10.1038/s41467-022-34664-z

Emelko, M. B., Chow, A., Amiri, F., Chen, H., & Stone, M. (2024). Climate change implications for DBP occurrence in municipal drinking water. In *Handbook of Environmental Chemistry* (Vol. 135, pp. 357–376). https://doi.org/10.1007/698_2024_1168

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

Emmerton, C. A., Cooke, C. A., Hustins, S., Silins, U., Emelko, M. B., Lewis, T., Kruk, M. K., Taube, N., Zhu, D., Jackson, B., Stone, M., Kerr, J. G., & Orwin, J. F. (2020). Severe western Canadian wildfire affects water quality even at large basin scales. *Water Research*, *183*, 116071. https://doi.org/10.1016/j.watres.2020.116071

Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science*, *320*(5879), 1034-1039. https://doi.org/10.1126/science.1153213

Foster, D. E., Duinker, P. N., Jamieson, R. C., Keys, K., & Steenberg, J. W. N. (2024). Where does the carbon go? Long-term effects of forest management on the carbon budget of a temperate-forest water-supply watershed. *Journal of Environmental Management*, *352*, 120007. https://doi.org/10.1016/j.jenvman.2023.120007

Foster, R. A., Kuypers, M. M. M., Vagner, T., Paerl, R. W., Musat, N., & Zehr, J. P. (2011). Nitrogen fixation and transfer in open ocean diatom–cyanobacterial symbioses. *The ISME Journal*, *5*(9), 1484-1493. https://doi.org/10.1038/ismej.2011.26

Glöckner, F. O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., Bruns, G., Yarza, P., Peplies, J., Westram, R., & Ludwig, W. (2017). 25 years of serving the community with ribosomal RNA gene reference databases and tools. *Journal of Biotechnology*, *261*, 169-176. https://doi.org/10.1016/j.jbiotec.2017.06.1198

Grosbois, G., Mou, T. A., & Girona, M. M. (2024). Cyanobacteria in winter: Seasonal dynamics of harmful algal blooms and their driving factors in boreal lakes. *Heliyon*, *10*(24), e40687. https://doi.org/10.1016/j.heliyon.2024.e40687

Gustine, R. N., Hanan, E. J., Robichaud, P. R., & Elliot, W. J. (2022). From burned slopes to streams: how wildfire affects nitrogen cycling and retention in forests and fire-prone watersheds. Biogeochemistry, 157(1), 51-68.

He, S., Stevens, S. L. R., Chan, L.-K., Bertilsson, S., Glavina del Rio, T., Tringe, S. G., Malmstrom, R. R., & McMahon, K. D. (2017). Ecophysiology of freshwater Verrucomicrobia inferred from metagenome-assembled genomes. *mSphere*, *2*(5), e00277-17. https://doi.org/10.1128/mSphere.00277-17

Herndon, E., Richardson, J., Carrell, A. A., Pierce, E., & Weston, D. (2023). Sulfur speciation in *Sphagnum* peat moss modified by mutualistic interactions with cyanobacteria. *New Phytologist*, *241*(5), 1998-2008. https://doi.org/10.1111/nph.19476

Herrero, A., Stavans, J., & Flores, E. (2016). The multicellular nature of filamentous heterocyst-forming cyanobacteria. *FEMS Microbiology Reviews*, *40*(6), 831-854. https://doi.org/10.1093/femsre/fuw029

Ho, J. C., & Michalak, A. M. (2020). Exploring temperature and precipitation impacts on harmful algal blooms across continental U.S. lakes. *Limnology and Oceanography*, *65*(5), 992-1009. https://doi.org/10.1002/lno.11365

Hokanson, K. J., Peterson, E. S., Devito, K. J., & Mendoza, C. A. (2020). Forestland-peatland hydrologic connectivity in water-limited environments: Hydraulic gradients often oppose topography. *Environmental Research Letters*, *15*(3), 034021. https://doi.org/10.1088/1748-9326/ab699a

Hokanson, K. J., Rostron, B. J., Devito, K. J., Hopkinson, C., & Mendoza, C. A. (2021). Landscape controls of surface-water/groundwater interactions on shallow outwash lakes: How the long-term groundwater signal overrides interannual variability due to evaporative effects. *Hydrogeology Journal*, *30*(1), 251-264. https://doi.org/10.1007/s10040-021-02422-z

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

Igwaran, A., Kayode, A. J., Moloantoa, K. M., Khetsha, Z. P., & Unuofin, J. O. (2024). Cyanobacteria harmful algae blooms: Causes, impacts, and risk management. *Water, Air, & Soil Pollution*, *235*(1), 26. https://doi.org/10.1007/s11270-023-06782-y

Janda, J. M., & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *Journal of Clinical Microbiology*, *45*(9), 2761–2764. https://doi.org/10.1128/JCM.01228-07

Ji, B., Liang, J., Ma, Y., Zhu, L., & Liu, Y. (2019). Bacterial community and eutrophic index analysis of the East Lake. *Environmental Pollution*, *252*, 682-688. https://doi.org/10.1016/j.envpol.2019.05.138

Johnson, J. S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, *10*(1), 5029. https://doi.org/10.1038/s41467-019-13036-1

Jung, P. (2023). The right way for cyanobacteria: The metabarcoding database CyanoSeq. *Journal of Phycology*, *59*(3), 467-469. https://doi.org/10.1111/jpy.13338

Kettridge, N., & Waddington, J. M. (2014). Towards quantifying the negative feedback regulation of peatland evaporation to drought. *Hydrologic Processes, 28*(11), 3728–3740. https://doi.org/10.1002/hyp.9898

Kinnear, S. (2010). Cylindrospermopsin: A decade of progress on bioaccumulation research. *Marine Drugs*, *8*(3), 542-564. https://doi.org/10.3390/md8030542

Komárek, J. (2016). A polyphasic approach for the taxonomy of cyanobacteria: Principles and applications. *European Journal of Phycology*, *51*(3), 346-353. https://doi.org/10.1080/09670262.2016.1163738

Komárek, J. (2023). Taxonomic review of Cyanobacteria 2021/2022 according to polyphasic evaluation. *Fottea*, *23*(1), 141-148. https://doi.org/10.5507/fot.2022.017

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

Krishnappan, B. G., Stone, M., Granger, S., Upadhayay, H., Tang, Q., Zhang, Y., & Collins, A. (2020). Experimental investigation of erosion characteristics of fine-grained cohesive sediments. *Water*, *12*(5), 1511. https://doi.org/10.3390/w12051511

Leader, S., Kettridge, N., Hannah, D. M., Mendoza, C., & Devito, K. J. (2024b). Multi-year climate memory in shallow lake water levels. *Journal of Hydrology, 642*, 131764. https://doi.org/10.1016/j.jhydrol.2024.131764

Leader, S., Kettridge, N., Hannah, D., Mendoza, C., & Devito, K. J. (2024a). Diverse response of shallow lake water levels to decadal weather patterns in a heterogeneous glacial Boreal Plains landscape. *Hydrologic Processes, 38*(4), e15125. https://doi.org/10.1002/hyp.15125

Lee, C. S., Kim, M., Lee, C., Yu, Z., & Lee, J. (2016). The microbiota of recreational freshwaters and the implications for environmental and public health. *Frontiers in Microbiology*, *7*, 1826. https://doi.org/10.3389/fmicb.2016.01826

Lefler, F. W., Berthold, D. E., & Laughinghouse, H. D. (2023). CyanoSeq: A database of cyanobacterial 16S rRNA gene sequences with curated taxonomy. *Journal of Phycology*, *59*(3), 470-480. https://doi.org/10.1111/jpy.13335

Levy-Booth, D. J., Prescott, C. E., & Grayston, S. J. (2014). Microbial functional genes involved in nitrogen fixation, nitrification and denitrification in forest ecosystems. *Soil Biology and Biochemistry*, *75*, 11-25. https://doi.org/10.1016/j.soilbio.2014.03.021

Li, D., Sharp, J. O., Saikaly, P. E., Ali, S., Alidina, M., Alarawi, M. S., Keller, S., Hoppe-Jones, C., & Drewes, J. E. (2012). Dissolved organic carbon influences microbial community composition and diversity in managed aquifer recharge systems. *Applied and Environmental Microbiology*, *78*(19), 6819–6828. https://doi.org/10.1128/aem.01223-12

Li, X., Huo, S., & Xi, B. (2020). Updating the resolution for 16S rRNA OTUs clustering reveals the cryptic cyanobacterial genus and species. *Ecological Indicators*, *117*, 106695. https://doi.org/10.1016/j.ecolind.2020.106695

Lv, X., Cheng, Y., Zhang, S., Hu, Z., Xiao, P., Zhang, H., Geng, R., & Li, R. (2022). Polyphasic characterization and taxonomic evaluation of a bloom-forming strain morphologically resembling *Radiocystis fernandoi* (Chroococcales, Cyanobacteria) from Lake Erhai, China. *Diversity*, *14*(10), 816. https://doi.org/10.3390/d14100816

Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature*, *506*(7488), 307-315. https://doi.org/10.1038/nature13068

MacKeigan, P. W., Garner, R. E., Monchamp, M.-È., Walsh, D. A., Onana, V. E., Kraemer, S. A., Pick, F. R., Beisner, B. E., Agbeti, M. D., da Costa, N. B., Shapiro, B. J., & Gregory-Eaves,

I. (2022). Comparing microscopy and DNA metabarcoding techniques for identifying Cyanobacteria assemblages across hundreds of lakes. *Harmful Algae*, *113*, 102187. https://doi.org/10.1016/j.hal.2022.102187

Malki, K., Rosario, K., Sawaya, N. A., Székely, A. J., Tisza, M. J., & Breitbart, M. (2020). Prokaryotic and viral community composition of freshwater springs in Florida, USA. *mBio*, *11*(2), e00436-20. https://doi.org/10.1128/mbio.00436-20

Mansilha, C., Carvalho, A., Guimarães, P., & Espinha Marques, J. (2014). Water quality concerns due to forest fires: Polycyclic Aromatic Hydrocarbons (PAH) contamination of groundwater from mountain areas. *Journal of Toxicology and Environmental Health, Part A*, *77*(14–16), 806–815. https://doi.org/10.1080/15287394.2014.909301

Martens, A. M., Silins, U., Proctor, H. C., Williams, C. H. S., Wagner, M. J., Emelko, M. B., & Stone, M. (2019). Long-term impact of severe wildfire and post-wildfire salvage logging on macroinvertebrate assemblage structure in Alberta's Rocky Mountains. *International Journal of Wildland Fire*, *28*(10), 738–749. https://doi.org/10.1071/WF18177

Martiny, J. B. H., Bohannan, B. J. M., Brown, J. H., Colwell, R. K., Fuhrman, J. A., Green, J. L., Horner-Devine, M. C., Kane, M., Krumins, J. A., Kuske, C. R., Morin, P. J., Naeem, S., Øvreås, L., Reysenbach, A.-L., Smith, V. H., & Staley, J. T. (2006). Microbial biogeography: Putting microorganisms on the map. *Nature Reviews Microbiology*, *4*(2), 102-112. https://doi.org/10.1038/nrmicro1341

Mazard, S., Penesyan, A., Ostrowski, M., Paulsen, I., & Egan, S. (2016). Tiny microbes with a big impact: The role of Cyanobacteria and their metabolites in shaping our future. *Marine Drugs*, *14*(5), 97. https://doi.org/10.3390/md14050097

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

Moreira, C., Vasconcelos, V., & Antunes, A. (2013). Phylogeny and biogeography of Cyanobacteria and their produced toxins. *Marine Drugs*, *11*(11), 4350-4369. https://doi.org/10.3390/md11114350

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

Nei, M., & Kumar, S. (2000). *Molecular evolution and phylogenetics*. Oxford University Press. https://doi.org/10.1093/oso/9780195135848.001.0001

Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., & Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews*, *75*(1), 14–49. https://doi.org/10.1128/MMBR.00028-10

Niño-García, J. P., Ruiz-González, C., & del Giorgio, P. A. (2016). Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. *The ISME Journal*, *10*(7), 1755-1766. https://doi.org/10.1038/ismej.2015.226

Nwankwegu, A. S., Li, Y., Huang, Y., Wei, J., Norgbey, E., Sarpong, L., Lai, Q., & Wang, K. (2019). Harmful algal blooms under changing climate and constantly increasing anthropogenic actions: The review of management implications. *3 Biotech*, *9*(12), 449. https://doi.org/10.1007/s13205-019-1976-1

Nwosu, E. C., Roeser, P., Yang, S., Pinkerneil, S., Ganzert, L., Dittmann, E., Brauer, A., Wagner, D., & Liebner, S. (2021). Species-level spatio-temporal dynamics of Cyanobacteria in a

hard-water temperate lake in the southern Baltics. *Frontiers in Microbiology*, *12*, 761259. https://doi.org/10.3389/fmicb.2021.761259

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Oksanen, M. J. (2013). Package 'vegan'. *Community ecology package, version*, *2*(9), 1-295.

Olefeldt, D., Turetsky, M. R., & Blodau, C. (2013). Altered composition and microbial versus UV-mediated degradation of dissolved organic matter in boreal soils following wildfire. *Ecosystems*, *16*(8), 1396-1412. https://doi.org/10.1007/s10021-013-9691-y

Orlova, J., Olefeldt, D., Yasinski, J. H., & Anderson, A. E. (2020). Effects of prescribed burn on nutrient and dissolved organic matter characteristics in peatland shallow groundwater. *Fire*, *3*(3), 53. https://doi.org/10.3390/fire3030053

Pacheco, A. B. F., Guedes, I. A., & Azevedo, S. M. F. O. (2016). Is qPCR a reliable indicator of cyanotoxin risk in freshwater? *Toxins*, *8*(6), 172. https://doi.org/10.3390/toxins8060172

Paerl, H. W., & Otten, T. G. (2013). Harmful cyanobacterial blooms: Causes, consequences, and controls. *Microbial Ecology*, *65*(4), 995-1010. https://doi.org/10.1007/s00248-012-0159-y

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

Paerl, H. W., Hall, N. S., & Calandrino, E. S. (2011). Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. *Science of The Total Environment*, *409*(10), 1739-1745. https://doi.org/10.1016/j.scitotenv.2011.02.001

Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P.-A., & Hugenholtz, P. (2022). GTDB: An ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research*, *50*(D1), D785-D794. https://doi.org/10.1093/nar/gkab776

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*, *36*(10), 996-1004. https://doi.org/10.1038/nbt.4229

Paul, M. J., LeDuc, S. D., Lassiter, M. G., Moorhead, L. C., Noyes, P. D., & Leibowitz, S. G. (2022). Wildfire induces changes in receiving waters: A review with considerations for water quality management. *Water Resources Research*, *58*(e2021WR030699). https://doi.org/10.1029/2021WR030699

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

Petrone, R. M., Landine, C., Ahmed, M., Mendoza, C. A., & Devito, K. (2016). Utikuma Region Study Area (URSA) – Part 2: Hydrologic connectivity in a complex landscape (HEAD). *The Forestry Chronicle*, *92*(1), 18-25. https://doi.org/10.5558/tfc2016-018

Peura, S., Sinclair, L., Bertilsson, S., & Eiler, A. (2015). Metagenomic insights into strategies of aerobic and anaerobic carbon and nitrogen transformation in boreal lakes. *Scientific Reports*, *5*, 12102. https://doi.org/10.1038/srep12102

Pimentel, J. S. M., & Giani, A. (2014). Microcystin production and regulation under nutrient stress conditions in toxic *Microcystis* strains. *Applied and Environmental Microbiology*, *80*(18), 5836-5843. https://doi.org/10.1128/AEM.01009-14

Plach, J. M., Ferone, J.-M., Gibbons, Z., Smerdon, B. D., Mertens, A., Mendoza, C. A., Petrone, R. M., & Devito, K. J. (2016). Influence of glacial landform hydrology on phosphorus

budgets of shallow lakes on the Boreal Plain, Canada. *Journal of Hydrology*, *535*, 191–203. https://doi.org/10.1016/j.jhydrol.2016.01.041

Pugh, E. A., Olefeldt, D., Leader, S. N., Hokanson, K. J., & Devito, K. J. (2021). Characteristics of dissolved organic carbon in boreal lakes: High spatial and inter-annual variability controlled by landscape attributes and wet-dry periods. *Water Resources Research*, *57*(11), e2021WR030021. https://doi.org/10.1029/2021wr030021

Putland, J. N., & Rivkin, R. B. (1999). Influence of storage mode and duration on the microscopic enumeration of Synechococcus from a cold coastal ocean environment. *Aquatic Microbial Ecology, 17*, 191–199.

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*(D1), D590-D596. https://doi.org/10.1093/nar/gks1219

R Core Team. (2024). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rastogi, G., Tech, J. J., Coaker, G. L., & Leveau, J. H. J. (2010). A PCR-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments. *Journal of Microbiological Methods*, *83*(2), 127-132. https://doi.org/10.1016/j.mimet.2010.08

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), 384. https://doi.org/10.3390/f10050384

Robinne, F.-N., Miller, C., Parisien, M.-A., Emelko, M. B., Bladon, K. D., Silins, U., & Flannigan, M. (2016). A global index for mapping the exposure of water resources to wildfire. *Forests*, *7*(1), 22. https://doi.org/10.3390/f7010022

Roush, D., Giraldo-Silva, A., & Garcia-Pichel, F. (2021). Cydrasil 3, a curated 16S rRNA gene reference package and web app for cyanobacterial phylogenetic placement. *Scientific Data*, *8*(1), 230. https://doi.org/10.1038/s41597-021-01015-5

Sánchez-Baracaldo, P., Bianchini, G., Wilson, J. D., & Knoll, A. H. (2022). Cyanobacteria and biogeochemical cycles through Earth history. *Trends in Microbiology*, *30*(2), 143-157. https://doi.org/10.1016/j.tim.2021.05.008

Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., Post, A. F., Hagemann, M., Paulsen, I., & Partensky, F. (2009). Ecological genomics of marine picocyanobacteria. *Microbiology and Molecular Biology Reviews*, *73*(2), 249-299. https://doi.org/10.1128/MMBR.00035-08

Schmidt, P. J., Cameron, E. S., Müller, K. M., & Emelko, M. B. (2022). Ensuring that fundamentals of quantitative microbiology are reflected in microbial diversity analyses based on next-generation sequencing. *Frontiers in Microbiology*, *13*, 728146. https://doi.org/10.3389/fmicb.2022.728146

Sciuto, K., & Moro, I. (2016). Detection of the new cosmopolitan genus *Thermoleptolyngbya* (Cyanobacteria, Leptolyngbyaceae) using the 16S rRNA gene and 16S–23S ITS region. *Molecular Phylogenetics and Evolution*, *105*, 15-35. https://doi.org/10.1016/j.ympev.2016.08.010

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

Sim, Z. Y., Goh, K. C., He, Y., & Gin, K. Y. H. (2023). Present and future potential role of toxin-producing *Synechococcus* in the tropical region. *Science of The Total Environment*, *896*, 165230. https://doi.org/10.1016/j.scitotenv.2023.165230

Śliwińska-Wilczewska, S., Maculewicz, J., Barreiro Felpeto, A., & Latała, A. (2018). Allelopathic and bloom-forming picocyanobacteria in a changing world. *Toxins*, *10*(1), 48. https://doi.org/10.3390/toxins10010048

Smerdon, B. D., Devito, K. J., & Mendoza, C. A. (2005). Interaction of groundwater and shallow lakes on outwash sediments in the sub-humid Boreal Plains of Canada. *Journal of Hydrology*, *314*(1–4), 246–262. https://doi.org/10.1016/j.jhydrol.2005.04.001

Song, L., Jia, Y., Qin, B., Li, R., Carmichael, W. W., Gan, N., Xu, H., Shan, K., & Sukenik, A. (2023). Harmful cyanobacterial blooms: Biological traits, mechanisms, risks, and control strategies. *Annual Review of Environment and Resources*, *48*(1), 123-147. https://doi.org/10.1146/annurev-environ-112320-081653

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312-1313. https://doi.org/10.1093/bioinformatics/btu033

Stockner, J. G. (1988). Phototrophic picoplankton: An overview from marine and freshwater ecosystems. *Limnology and Oceanography,* *33*(4, part 2), 765–775. https://doi.org/10.4319/lo.1988.33.4part2.0765

Stone, M., Emelko, M. B., Droppo, I. G., & Silins, U. (2011). Biostabilization and erodibility of cohesive sediment deposits in wildfire-affected streams. Water Research, 45(2), 521–534. https://doi.org/10.1016/j.watres.2010.09.016

Straub, D., Blackwell, N., Langarica-Fuentes, A., Peltzer, A., Nahnsen, S., & Kleindienst, S. (2020). Interpretations of environmental microbial community studies are biased by the selected 16S rRNA (gene) amplicon sequencing pipeline. *Frontiers in Microbiology*, *11*, Article 550420. https://doi.org/10.3389/fmicb.2020.550420

Strunecký, O., Ivanova, A. P., & Mareš, J. (2023). An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. *Journal of Phycology*, *59*(1), 12-51. https://doi.org/10.1111/jpy.13304

Sukenik, A., Hadas, O., Kaplan, A., & Quesada, A. (2012). Invasion of Nostocales (Cyanobacteria) to subtropical and temperate freshwater lakes – Physiological, regional, and global driving forces. *Frontiers in Microbiology*, *3*, 86. https://doi.org/10.3389/fmicb.2012.00086

Tsuji, J. M., Tran, N., Schiff, S. L., Venkiteswaran, J. J., Molot, L. A., Tank, M., Hanada, S., & Neufeld, J. D. (2020). Anoxygenic photosynthesis and iron–sulfur metabolic potential of *Chlorobia* populations from seasonally anoxic Boreal Shield lakes. *The ISME Journal*, *14*(11), 2732-2747. https://doi.org/10.1038/s41396-020-0725-0

Turetsky, M. R., Donahue, W. F., & Benscoter, B. W. (2011). Experimental drying intensifies burning and carbon losses in a northern peatland. *Nature Communications*, *2*(1), 514. https://doi.org/10.1038/ncomms1523

Usher, K. M., Fromm, K. M., & Cockell, C. S. (2007). Algae on ice: Characterization of algal communities from cryoconite holes on glaciers in Svalbard. *FEMS Microbiology Ecology*, *62*(2), 110-121. https://doi.org/10.1111/j.1574-6941.2007.00368.x

Vettorazzo, S., Boscaini, A., Cerasino, L., & Salmaso, N. (2024). From small water bodies to lakes: Exploring the diversity of freshwater bacteria in an Alpine Biosphere Reserve. *Science of The Total Environment*, *954*, 176495. https://doi.org/10.1016/j.scitotenv.2024.176495

Waddington, J. M., Morris, P. J., Kettridge, N., Granath, G., Thompson, D. K., & Moore, P. A. (2015). Hydrological feedbacks in northern peatlands. *Ecohydrology*, *8*(1), 113–127. https://doi.org/10.1002/eco.1493

Walsh, C. R., Grey, F., Patterson, R. T., Ralchenko, M., Patterson, C. W., Reinhardt, E. G., Grey, D., Grey, H., & Thunder, D. (2025). Late 20th century hypereutrophication of northern Alberta's Utikuma Lake. *Environments*, *12*(2), 63. https://doi.org/10.3390/environments12020063

Walters, W., Hyde, E. R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J. A., Jansson, J. K., Caporaso, J. G., 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. *mSystems*, *1*(1), e00009-15. https://doi.org/10.1128/mSystems.00009-15

Watt, C., Emelko, M. B., Silins, U., Collins, A. L., & Stone, M. (2021). Anthropogenic and climate-exacerbated landscape disturbances converge to alter phosphorus bioavailability in an oligotrophic river. *Water*, *13*(22), 3151. https://doi.org/10.3390/w13223151

Wells, M. L., Karlson, B., Wulff, A., Kudela, R., Trick, C., Asnaghi, V., Berdalet, E., Cochlan, W., Davidson, K., De Rijcke, M., Dutkiewicz, S., Hallegraeff, G., Flynn, K. J., Legrand, C., Paerl, H., Silke, J., Suikkanen, S., Thompson, P., & Trainer, V. L. (2020). Future HAB science: Directions and challenges in a changing climate. *Harmful Algae*, *91*, 101632. https://doi.org/10.1016/j.hal.2019.101632

Werner, J. J., Jungblut, A. D., & Neilan, B. A. (2024). Towards resolving the global biogeography of Cyanobacteria using a DNA metabarcoding approach in biological soil crusts. *Environmental Microbiology Reports*, *16*(4), e13337. https://doi.org/10.1111/1758-2229.13337

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (Version 1.1.4) [R package]. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr

Winter, J. G., DeSellas, A. M., Fletcher, R., Heintsch, L., Morley, A., Nakamoto, L., & Utsumi, K. (2011). Algal blooms in Ontario, Canada: Increases in reports since 1994. *Lake and Reservoir Management*, *27*(2), 107-114. https://doi.org/10.1080/07438141.2011.557765

Wood, R. (2016). Acute animal and human poisonings from cyanotoxin exposure — A review of the literature. *Environment International*, *91*, 276-282. https://doi.org/10.1016/j.envint.2016.02.026

Yang, X. E., Wu, X., Hao, H. L., & He, Z. L. (2008). Mechanisms and assessment of water eutrophication. *Journal of Zhejiang University SCIENCE B*, *9*(3), 197-209. https://doi.org/10.1631/jzus.b0710626

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F. O. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, *42*(D1), D643-D648. https://doi.org/10.1093/nar/gkt1209

Yu, Z. C. (2012). Northern peatland carbon stocks and dynamics: A review. *Biogeosciences*, *9*(10), 4071-4085. https://doi.org/10.5194/bg-9-4071-2012

Yutin, N., & Galperin, M. Y. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environmental Microbiology*, *15*(10), 2631-2641. https://doi.org/10.1111/1462-2920.12173

Zamyadi, A., McQuaid, N., Prévost, M., & Dorner, S. (2012). Monitoring of potentially toxic Cyanobacteria using an online multi-probe in drinking water sources. *Journal of Environmental Monitoring*, *14*(2), 579-588. https://doi.org/10.1039/c1em10819k

Zhang, D., Te, S. H., He, Y., & Gin, K. Y.-H. (2022). Cyanophage dynamics in a tropical urban freshwater lake. *Ecological Indicators*, *142*, 109257. https://doi.org/10.1016/j.ecolind.2022.109257

Zhang, J., Li, F., Long, L., & Huang, S. (2022). Quantification of marine picocyanobacteria on water column particles and in sediments using real-time PCR reveals their role in carbon export. *mSphere, 7*(6), e00499-22. https://doi.org/10.1128/msphere.00499-22

Zheng, F., Zhang, T., Yin, S., Qin, G., Chen, J., Zhang, J., Zhao, D., Leng, X., An, S., & Xia, L. (2022). Comparison and interpretation of freshwater bacterial structure and interactions with organic to nutrient imbalances in restored wetlands. *Frontiers in Microbiology*, *13*, 946537. https://doi.org/10.3389/fmicb.2022.946537

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

Zwart, G., Kamst-van Agterveld, M. P., van der Werff-Staverman, I., Hagen, F., Hoogveld, H. L., & Gons, H. J. (2005). Molecular characterization of cyanobacterial diversity in a shallow eutrophic lake. *Environmental Microbiology*, *7*(3), 365-377. https://doi.org/10.1111/j.1462-2920.2004.00715.x

Appendix A. Variance Inflation Factor for Water Chemistry Data of the URSA Ponds

Table A.1 VIF for Water Chemistry Data of the URSA Ponds

Variable	VIF
pH	**2.6**
Iron (Fe) (mg/L)	**3.6**
Nitrate + Nitrite (NO2/NO3) (µg/L)	**4.1**
Ammonium (NH4) (µg/L)	**5.0**
Chlorophyll *a* (Chl-a) (µg/L)	**6.2**
Aluminum (Al) (µg/L)	**8.0**
Potassium (K) (mg/L)	**14.8**
Dissolved Organic Carbon (DOC) (mg/L)	**17.7**
Chloride (Cl) (mg/L)	**60.2**
Sodium (Na) (mg/L)	**61.5**
Total Dissolved Nitrogen (TDN) (µg/L)	**86.7**
Total Nitrogen (TN) (µg/L)	**128.0**
Magnesium (Mg) (mg/L)	**158.0**
Total Phosphorus (TP) (µg/L)	**169.0**
Sulfate (SO4) (mg/L)	**170.4**
Soluble Reactive Phosphorus (SRP) (µg/L)	**183.7**
Calcium (Ca) (mg/L)	**369.2**
Bicarbonate ($HCO3^-$) (mg/L)	**464.6**
Total Dissolved Phosphorus (TDP) (µg/L)	**466.1**
Alkalinity (as CaCO3) (mg/L)	**894.1**

Appendix B. Cyanoseq Classifier and Blast 16S Ribosomal Database Comparisons

Table B.1 Cyanoseq Classifier and Blast 16S Ribosomal Database Comparisons

URSA Identifier	Accession	Percent Identity	E-Value	NCBI	GTDB CyanoSeq
CYAN001	**NR_114406**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN001	**NR_102447**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN002	**NR_114406**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN002	**NR_102447**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN003	**NR_125481**	**98.419**	**4.35E-125**	***Synechococcus rubescens***	**f__Prochlorococcaceae**
CYAN004	**NR_177802**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN004	**NR_177801**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN004	**NR_177800**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN005	**NR_177802**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN005	**NR_177801**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN005	**NR_177800**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN006	**NR_114406**	**96.443**	**9.48E-117**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN006	**NR_102447**	**96.443**	**9.48E-117**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN007	**NR_114406**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN007	**NR_102447**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN008	**NR_114406**	**96.443**	**9.48E-117**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN008	**NR_102447**	**96.443**	**9.48E-117**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN009	**NR_114406**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN009	**NR_102447**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN010	**NR_114406**	**99.605**	**4.32E-130**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN010	**NR_102447**	**99.605**	**4.32E-130**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN011	**NR_177802**	**99.209**	**2.01E-128**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***

CYAN011	**NR_177801**	**99.209**	**2.01E-128**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN011	**NR_177800**	**99.209**	**2.01E-128**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN012	**NR_114406**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN012	**NR_102447**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN013	**NR_186878**	**99.605**	**4.32E-130**	***Microcystis fernandoi***	**g__*Microcystis***
CYAN013	**NR_074314**	**99.605**	**4.32E-130**	***Microcystis aeruginosa***	**g__*Microcystis***
CYAN014	**NR_186878**	**99.605**	**4.32E-130**	***Microcystis fernandoi***	**g__*Microcystis***
CYAN014	**NR_074314**	**99.605**	**4.32E-130**	***Microcystis aeruginosa***	**g__*Microcystis***
CYAN015	**NR_172611**	**95.635**	**7.38E-113**	***Konicacronema caatingensis***	**g__*Oscillatoria***
CYAN015	**NR_176558**	**95.635**	**7.38E-113**	***Funiculus tenuis***	**g__*Oscillatoria***
CYAN016	**NR_125481**	**99.605**	**4.32E-130**	***Synechococcus rubescens***	**g__*Lacustricoccus***
CYAN017	**NR_114406**	**96.838**	**2.04E-118**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN017	**NR_102447**	**96.838**	**2.04E-118**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN018	**NR_125481**	**98.419**	**4.35E-125**	***Synechococcus rubescens***	**g__*Lacustricoccus***
CYAN019	**NR_177802**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN019	**NR_177801**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN019	**NR_177800**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN020	**NR_114406**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN020	**NR_102447**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN021	**NR_114406**	**99.209**	**2.01E-128**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN021	**NR_102447**	**99.209**	**2.01E-128**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN022	**NR_125481**	**100**	**9.28E-132**	***Synechococcus rubescens***	**g__*Lacustricoccus***
CYAN023	**NR_114406**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN023	**NR_102447**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN024	**NR_177802**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN024	**NR_177801**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN024	**NR_177800**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN025	**NR_177802**	**97.233**	**4.38E-120**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN025	**NR_177801**	**97.233**	**4.38E-120**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN025	**NR_177800**	**97.233**	**4.38E-120**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**

CYAN026	**NR_114406**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN026	**NR_102447**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN027	**NR_114406**	**99.605**	**4.32E-130**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN027	**NR_102447**	**99.605**	**4.32E-130**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN028	**NR_186878**	**100**	**9.28E-132**	***Microcystis fernandoi***	**g__*Microcystis***
CYAN028	**NR_074314**	**100**	**9.28E-132**	***Microcystis aeruginosa***	**g__*Microcystis***
CYAN029	**NR_114406**	**96.443**	**9.48E-117**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN029	**NR_102447**	**96.443**	**9.48E-117**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN030	**NR_114406**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN030	**NR_102447**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN031	**NR_177802**	**99.209**	**2.01E-128**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN031	**NR_177801**	**99.209**	**2.01E-128**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN031	**NR_177800**	**99.209**	**2.01E-128**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN032	**NR_176571**	**99.216**	**5.58E-129**	***Nostoc favosum***	**g__*Nostoc***
CYAN033	**NR_177802**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN033	**NR_177801**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN033	**NR_177800**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN034	**NR_172612**	**89.328**	**3.56E-86**	***Trichocoleus caatingensis***	**p__Cyanobacteriota**
CYAN035	**NR_114406**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**g__*RGUW01***
CYAN035	**NR_102447**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**g__*RGUW01***
CYAN036	**NR_177802**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN036	**NR_177801**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN036	**NR_177800**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN037	**NR_177799**	**92.52**	**4.50E-100**	***Desmonostoc vinosum***	**g__*Wollea***
CYAN038	**NR_125481**	**99.605**	**4.32E-130**	***Synechococcus rubescens***	**g__*Lacustricoccus***
CYAN039	**NR_125481**	**98.024**	**2.02E-123**	***Synechococcus rubescens***	**g__*Vulcanococcus***
CYAN040	**NR_191017**	**95.635**	**7.38E-113**	***Konicacronema protuberans***	**g__*Oscillatoria***
CYAN040	**NR_172611**	**95.635**	**7.38E-113**	***Konicacronema caatingensis***	**g__*Oscillatoria***
CYAN040	**NR_176558**	**95.635**	**7.38E-113**	***Funiculus tenuis***	**g__*Oscillatoria***
CYAN041	**NR_177794**	**98.419**	**4.35E-125**	***Desmonostoc geniculatum***	**f__Nostocaceae**

CYAN041	**NR_176981**	**98.419**	**4.35E-125**	***Desmonostoc muscorum***	**f__Nostocaceae**
CYAN042	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN042	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**g__*Vulcanococcus***
CYAN043	**NR_114406**	**100**	**9.28E-132**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN043	**NR_102447**	**100**	**9.28E-132**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN044	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN044	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN045	**NR_177802**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN045	**NR_177801**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN045	**NR_177800**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN046	**NR_177802**	**95.652**	**2.05E-113**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN046	**NR_177801**	**95.652**	**2.05E-113**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN046	**NR_177800**	**95.652**	**2.05E-113**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN047	**NR_177802**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN047	**NR_177801**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN047	**NR_177800**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN048	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN048	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN049	**NR_114406**	**96.838**	**2.04E-118**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN049	**NR_102447**	**96.838**	**2.04E-118**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN050	**NR_177802**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN050	**NR_177801**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN050	**NR_177800**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN051	**NR_176478**	**98.419**	**4.35E-125**	***Arthronema gygaxiana***	**g__*Limnothrix***
CYAN052	**NR_177794**	**98.419**	**4.35E-125**	***Desmonostoc geniculatum***	**g__*Johanseniella***
CYAN052	**NR_176981**	**98.419**	**4.35E-125**	***Desmonostoc muscorum***	**g__*Johanseniella***
CYAN053	**NR_114406**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN053	**NR_102447**	**98.419**	**4.35E-125**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN054	**NR_177738**	**92.913**	**9.68E-102**	***Rippkaea orientalis***	**g__*Woronichinia***
CYAN055	**NR_177716**	**96.457**	**9.48E-117**	***Neowollea manoromense***	**g__*Cuspidothrix***

CYAN056	**NR_177036**	**98.419**	**4.35E-125**	***Nodosilinea radiophila***	**g__*Nodosilinea***
CYAN056	**NR_177034**	**98.419**	**4.35E-125**	***Nodosilinea ramsarensis***	**g__*Nodosilinea***
CYAN056	**NR_177033**	**98.419**	**4.35E-125**	***Nodosilinea ramsarensis***	**g__*Nodosilinea***
CYAN056	**NR_177031**	**98.419**	**4.35E-125**	***Nodosilinea ramsarensis***	**g__*Nodosilinea***
CYAN057	**NR_177802**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN057	**NR_177801**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN057	**NR_177800**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN058	**NR_191019**	**94.071**	**9.61E-107**	***Limnonema hangangris***	**g__Limnonema**
CYAN059	**NR_177802**	**96.047**	**4.41E-115**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN059	**NR_177801**	**96.047**	**4.41E-115**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN059	**NR_177800**	**96.047**	**4.41E-115**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN060	**NR_177802**	**97.233**	**4.38E-120**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN060	**NR_177801**	**97.233**	**4.38E-120**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN060	**NR_177800**	**97.233**	**4.38E-120**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN061	**NR_189725**	**99.605**	**4.32E-130**	***Nostoc montejanii***	**g__*Nostoc***
CYAN061	**NR_189724**	**99.605**	**4.32E-130**	***Nostoc montejanii***	**g__*Nostoc***
CYAN061	**NR_114430**	**99.605**	**4.32E-130**	***Nostoc punctiforme***	**g__*Nostoc***
CYAN061	**NR_074317**	**99.605**	**4.32E-130**	***Nostoc punctiforme***	**g__*Nostoc***
CYAN062	**NR_177802**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN062	**NR_177801**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN062	**NR_177800**	**96.443**	**9.48E-117**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN063	**NR_177668**	**89.804**	**2.12E-88**	***Chalicogloea cavernicola***	**g__*Radiocystis***
CYAN063	**NR_186890**	**89.804**	**2.12E-88**	***Chalicogloea cavernicola***	**g__*Radiocystis***
CYAN064	**NR_177802**	**94.862**	**4.44E-110**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN064	**NR_177801**	**94.862**	**4.44E-110**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN064	**NR_177800**	**94.862**	**4.44E-110**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN065	**NR_186878**	**99.209**	**2.01E-128**	***Microcystis fernandoi***	**g__*Microcystis***
CYAN065	**NR_074314**	**99.209**	**2.01E-128**	***Microcystis aeruginosa***	**g__*Microcystis***
CYAN066	**NR_177802**	**98.814**	**9.34E-127**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN066	**NR_177801**	**98.814**	**9.34E-127**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***

CYAN066	**NR_177800**	**98.814**	**9.34E-127**	***Dolichospermum brachiatum***	**g__*Aphanizomenon***
CYAN067	**NR_177824**	**92.49**	**4.50E-100**	***Gloeothece aurea***	**g__*Cyanodorina***
CYAN067	**NR_177822**	**92.49**	**4.50E-100**	***Gloeothece aurea***	**g__*Cyanodorina***
CYAN067	**NR_177821**	**92.49**	**4.50E-100**	***Gloeothece aurea***	**g__*Cyanodorina***
CYAN067	**NR_177820**	**92.49**	**4.50E-100**	***Gloeothece aurea***	**g__*Cyanodorina***
CYAN067	**NR_177818**	**92.49**	**4.50E-100**	***Gloeothece aurea***	**g__*Cyanodorina***
CYAN067	**NR_177817**	**92.49**	**4.50E-100**	***Gloeothece aurea***	**g__*Cyanodorina***
CYAN068	**NR_177802**	**98.024**	**2.02E-123**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN068	**NR_177801**	**98.024**	**2.02E-123**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN068	**NR_177800**	**98.024**	**2.02E-123**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN069	**NR_177802**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN069	**NR_177801**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN069	**NR_177800**	**98.419**	**4.35E-125**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN070	**NR_177730**	**96.047**	**4.41E-115**	***Pseudanabaena foetida***	**g__*Limnothrix***
CYAN071	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN071	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN072	**NR_125481**	**96.047**	**4.41E-115**	***Synechococcus rubescens***	**g__*Lacustricoccus***
CYAN073	**NR_114406**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN073	**NR_102447**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN074	**NR_172606**	**100**	**9.28E-132**	***Microcoleus anatoxicus***	**g__Tychonema**
CYAN075	**NR_177733**	**100**	**9.28E-132**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN075	**NR_177732**	**100**	**9.28E-132**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN075	**NR_177731**	**100**	**9.28E-132**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN075	**NR_177729**	**100**	**9.28E-132**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN075	**NR_177728**	**100**	**9.28E-132**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN076	**NR_197632**	**90.909**	**2.11E-93**	***Gloeothece guizhouensis***	**g__*Snowella***
CYAN077	**NR_125481**	**98.419**	**4.35E-125**	***Synechococcus rubescens***	**g__*Lacustricoccus***
CYAN078	**NR_118325**	**100**	**9.28E-132**	***Limnoraphis robusta***	**g__*Limnoraphis***
CYAN079	**NR_177802**	**100**	**9.28E-132**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN079	**NR_177801**	**100**	**9.28E-132**	***Dolichospermum brachiatum***	**g__*Dolichospermum***

CYAN079	**NR_177800**	**100**	**9.28E-132**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN080	**NR_191042**	**96.838**	**2.04E-118**	***Azorothrix ramosa***	**g__*Trichormus***
CYAN080	**NR_117192**	**96.838**	**2.04E-118**	***Hassallia antarctica***	**g__*Trichormus***
CYAN080	**NR_117191**	**96.838**	**2.04E-118**	***Hassallia andreassenii***	**g__*Trichormus***
CYAN081	**NR_112168**	**95.238**	**3.43E-111**	***Oscillatoria nigro-viridis***	**c__Cyanophyceae**
CYAN081	**NR_102469**	**95.238**	**3.43E-111**	***Oscillatoria nigro-viridis***	**c__Cyanophyceae**
CYAN081	**NR_176477**	**95.238**	**3.43E-111**	***Microcoleus acremannii***	**c__Cyanophyceae**
CYAN082	**NR_189176**	**92.49**	**4.50E-100**	***Tumidithrix elongata***	**g__Pseudanabaenaceae_X**
CYAN083	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN083	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN084	**NR_191023**	**98.814**	**9.34E-127**	***Ahomia soli***	**o__Nostocales**
CYAN084	**NR_191022**	**98.814**	**9.34E-127**	***Ahomia purpurea***	**o__Nostocales**
CYAN085	**NR_189725**	**94.862**	**4.44E-110**	***Nostoc montejanii***	**f__Aphanizomenonaceae**
CYAN085	**NR_189724**	**94.862**	**4.44E-110**	***Nostoc montejanii***	**f__Aphanizomenonaceae**
CYAN085	**NR_177802**	**94.862**	**4.44E-110**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN085	**NR_177801**	**94.862**	**4.44E-110**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN085	**NR_177800**	**94.862**	**4.44E-110**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN085	**NR_102457**	**94.862**	**4.44E-110**	***Anabaena cylindrica***	**f__Aphanizomenonaceae**
CYAN086	**NR_177901**	**96.838**	**2.04E-118**	***Planktothrix iranica***	**g__*Planktothrix***
CYAN086	**NR_177900**	**96.838**	**2.04E-118**	***Planktothrix iranica***	**g__*Planktothrix***
CYAN086	**NR_177899**	**96.838**	**2.04E-118**	***Planktothrix iranica***	**g__*Planktothrix***
CYAN086	**NR_177898**	**96.838**	**2.04E-118**	***Planktothrix iranica***	**g__*Planktothrix***
CYAN086	**NR_177723**	**96.838**	**2.04E-118**	***Planktothrix serta***	**g__*Planktothrix***
CYAN086	**NR_177722**	**96.838**	**2.04E-118**	***Planktothrix paucivesiculata***	**g__*Planktothrix***
CYAN086	**NR_112126**	**96.838**	**2.04E-118**	***Planktothrix agardhii***	**g__*Planktothrix***
CYAN086	**NR_112127**	**96.838**	**2.04E-118**	***Planktothrix rubescens***	**g__*Planktothrix***
CYAN086	**NR_112124**	**96.838**	**2.04E-118**	***Planktothrix rubescens***	**g__*Planktothrix***
CYAN087	**NR_177802**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN087	**NR_177801**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN087	**NR_177800**	**96.838**	**2.04E-118**	***Dolichospermum brachiatum***	**g__*Dolichospermum***

CYAN088	**NR_148664**	**92.885**	**9.68E-102**	***Foliisarcina bertiogensis***	**g__Microcystaceae_X**
CYAN089	**NR_189723**	**96.838**	**2.04E-118**	***Stigonema lichenoides***	**o__Nostocales**
CYAN090	**NR_197631**	**90.551**	**9.82E-92**	***Gloeothece guizhouensis***	**g__*Snowella***
CYAN091	**NR_191042**	**99.209**	**2.01E-128**	***Azorothrix ramosa***	**f__Tolypothrichaceae**
CYAN091	**NR_117192**	**99.209**	**2.01E-128**	***Hassallia antarctica***	**f__Tolypothrichaceae**
CYAN092	**NR_114406**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN092	**NR_102447**	**98.814**	**9.34E-127**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN093	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN093	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN094	**NR_176478**	**98.024**	**2.02E-123**	***Arthronema gygaxiana***	**g__*Limnothrix***
CYAN095	**NR_177848**	**91.304**	**4.54E-95**	***Jaaginema litorale***	**g__Microcystaceae_X**
CYAN095	**NR_177816**	**91.304**	**4.54E-95**	***Gloeothece aequatorialis***	**g__Microcystaceae_X**
CYAN095	**NR_177814**	**91.304**	**4.54E-95**	***Gloeothece aequatorialis***	**g__Microcystaceae_X**
CYAN095	**NR_177812**	**91.304**	**4.54E-95**	***Gloeothece aequatorialis***	**g__Microcystaceae_X**
CYAN095	**NR_177811**	**91.304**	**4.54E-95**	***Gloeothece aequatorialis***	**g__Microcystaceae_X**
CYAN095	**NR_177810**	**91.304**	**4.54E-95**	***Gloeothece aequatorialis***	**g__Microcystaceae_X**
CYAN095	**NR_177809**	**91.304**	**4.54E-95**	***Gloeothece aequatorialis***	**g__Microcystaceae_X**
CYAN095	**NR_177693**	**91.304**	**4.54E-95**	***Jaaginema litorale***	**g__Microcystaceae_X**
CYAN096	**NR_186901**	**93.254**	**7.48E-103**	***Trichothermofontia sichuanensis***	**c__Cyanophyceae**
CYAN097	**NR_176514**	**96.063**	**4.41E-115**	***Ahomia thailandicus***	**g__*Nostoc***
CYAN097	**NR_102457**	**96.063**	**4.41E-115**	***Anabaena cylindrica***	**g__*Nostoc***
CYAN098	**NR_177802**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN098	**NR_177801**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN098	**NR_177800**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**g__*Dolichospermum***
CYAN099	**NR_191040**	**96.443**	**9.48E-117**	***Ciimarium marinum***	**g__*Regnicoccus***
CYAN100	**NR_191038**	**91.373**	**4.54E-95**	***Komarkovaeasiopsis thermale***	**f__Leptolyngbyaceae**
CYAN101	**NR_172679**	**91.7**	**9.75E-97**	***Myxacorys californica***	**g__*Snowella***
CYAN102	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**g__*Regnicoccus***
CYAN102	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**g__*Regnicoccus***
CYAN103	**NR_191044**	**98.814**	**9.34E-127**	***Shahulinema minutum***	**g__*Shahulinema***

CYAN104	**NR_125682**	**96.838**	**5.66E-119**	***Cylindrospermum alatosporum***	**g__Nodulariaceae_XXX**
CYAN104	**NR_125686**	**96.838**	**2.04E-118**	***Cylindrospermum licheniforme***	**g__Nodulariaceae_XXX**
CYAN104	**NR_125683**	**96.838**	**2.04E-118**	***Cylindrospermum pellucidum***	**g__Nodulariaceae_XXX**
CYAN104	**NR_102462**	**96.838**	**2.04E-118**	***Cylindrospermum stagnale***	**g__Nodulariaceae_XXX**
CYAN104	**NR_114701**	**96.838**	**2.04E-118**	***Cylindrospermum stagnale***	**g__Nodulariaceae_XXX**
CYAN105	**NR_176486**	**100**	**9.28E-132**	***Limnococcus limneticus***	**g__*Limnococcus***
CYAN106	**NR_186862**	**98.419**	**4.35E-125**	***Apatinema mutabilis***	**g__*Apatinema***
CYAN107	**NR_125697**	**88.142**	**3.58E-81**	***Trichocoleus desertorum***	**p__Cyanobacteriota**
CYAN108	**NR_186901**	**93.254**	**7.48E-103**	***Trichothermofontia sichuanensis***	**c__Cyanophyceae**
CYAN109	**NR_114406**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN109	**NR_102447**	**98.024**	**2.02E-123**	***Cyanobium gracile***	**g__*Cyanobium***
CYAN110	**NR_177013**	**93.676**	**4.47E-105**	***Phormidesmis communis***	**f__Leptolyngbyaceae**
CYAN111	**NR_176486**	**99.605**	**4.32E-130**	***Limnococcus limneticus***	**g__*Limnococcus***
CYAN112	**NR_177802**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN112	**NR_177801**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN112	**NR_177800**	**97.628**	**9.41E-122**	***Dolichospermum brachiatum***	**f__Aphanizomenonaceae**
CYAN113	**NR_191044**	**99.209**	**2.01E-128**	***Shahulinema minutum***	**g__*Shahulinema***
CYAN114	**NR_114406**	**90.514**	**3.53E-91**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN114	**NR_102447**	**90.514**	**3.53E-91**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN115	**NR_177716**	**95.257**	**9.54E-112**	***Neowollea manoromense***	**g__*Neowollea***
CYAN116	**NR_104911**	**87.402**	**2.16E-78**	***Vampirovibrio chlorellavorus***	**g__Candidatus-Gastranaerophilales**
CYAN117	**NR_114406**	**97.233**	**4.38E-120**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN117	**NR_102447**	**97.233**	**4.38E-120**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN118	**NR_176486**	**99.605**	**4.32E-130**	***Limnococcus limneticus***	**g__*Limnococcus***
CYAN119	**NR_197632**	**93.281**	**2.08E-103**	***Gloeothece guizhouensis***	**g__*Woronichinia***
CYAN120	**NR_186901**	**92.46**	**1.62E-99**	***Trichothermofontia sichuanensis***	**c__Cyanophyceae**
CYAN121	**NR_184627**	**95.652**	**2.05E-113**	***Romeriopsis navalis***	**g__*Romeriopsis***
CYAN122	**NR_189176**	**91.7**	**9.75E-97**	***Tumidithrix elongata***	**c__Cyanophyceae**
CYAN123	**NR_104911**	**86.561**	**1.68E-74**	***Vampirovibrio chlorellavorus***	**g__Candidatus-Gastranaerophilales**

CYAN124	**NR_114406**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN124	**NR_102447**	**97.628**	**9.41E-122**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN125	**NR_197630**	**99.209**	**2.01E-128**	***Geminocystis urbisnovae***	**g__*Geminocystis***
CYAN126	**NR_186901**	**92.095**	**7.54E-98**	***Trichothermofontia sichuanensis***	**c__Cyanophyceae**
CYAN127	**NR_191044**	**91.016**	**2.12E-93**	***Shahulinema minutum***	**g__*Chamaethrix***
CYAN127	**NR_191039**	**91.016**	**2.12E-93**	***Gansulinema desertorum***	**g__*Chamaethrix***
CYAN128	**NR_172612**	**88.142**	**3.58E-81**	***Trichocoleus caatingensis***	**p__Cyanobacteriota**
CYAN129	**NR_176478**	**95.161**	**2.07E-108**	***Arthronema gygaxiana***	**g__*Limnothrix***
CYAN130	**NR_177733**	**98.814**	**9.34E-127**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN130	**NR_177732**	**98.814**	**9.34E-127**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN130	**NR_177731**	**98.814**	**9.34E-127**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN130	**NR_177729**	**98.814**	**9.34E-127**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN130	**NR_177728**	**98.814**	**9.34E-127**	***Pseudanabaena foetida***	**g__*Pseudanabaena***
CYAN131	**NR_104911**	**85.375**	**1.69E-69**	***Vampirovibrio chlorellavorus***	**g__Candidatus-Gastranaerophilales**
CYAN132	**NR_177852**	**96.063**	**1.23E-115**	***Kamptonema animale***	**g__Kamptonema**
CYAN133	**NR_172702**	**85.039**	**7.86E-68**	***Tildeniella torsiva***	**g__Candidatus-Obscuribacter**
CYAN134	**NR_172589**	**91.732**	**9.75E-97**	***Tildeniella alaskaensis***	**g__Prochlorotrichaceae_X**
CYAN135	**NR_177013**	**93.281**	**2.08E-103**	***Phormidesmis communis***	**f__Leptolyngbyaceae**
CYAN136	**NR_191039**	**92.941**	**9.68E-102**	***Gansulinema desertorum***	**g__*Chamaethrix***
CYAN137	**NR_176486**	**99.209**	**2.01E-128**	***Limnococcus limneticus***	**g__*Limnococcus***
CYAN138	**NR_189725**	**98.024**	**2.02E-123**	***Nostoc montejanii***	**g__*Johanseniella***
CYAN138	**NR_189724**	**98.024**	**2.02E-123**	***Nostoc montejanii***	**g__*Johanseniella***
CYAN138	**NR_177794**	**98.024**	**2.02E-123**	***Desmonostoc geniculatum***	**g__*Johanseniella***
CYAN138	**NR_172622**	**98.024**	**2.02E-123**	***Amazonocrinis nigriterrae***	**g__*Johanseniella***
CYAN138	**NR_114430**	**98.024**	**2.02E-123**	***Nostoc punctiforme***	**g__*Johanseniella***
CYAN138	**NR_074317**	**98.024**	**2.02E-123**	***Nostoc punctiforme***	**g__*Johanseniella***
CYAN138	**NR_176981**	**98.024**	**2.02E-123**	***Desmonostoc muscorum***	**g__*Johanseniella***
CYAN139	**NR_176478**	**98.024**	**2.02E-123**	***Arthronema gygaxiana***	**g__*Limnothrix***
CYAN140	**NR_176545**	**87.008**	**1.00E-76**	***Cryptochroococcus tibeticus***	**p__Cyanobacteriota**

CYAN140	**NR_176544**	**87.008**	**1.00E-76**	***Cryptochroococcus tibeticus***	**p__Cyanobacteriota**
CYAN140	**NR_176543**	**87.008**	**1.00E-76**	***Cryptochroococcus tibeticus***	**p__Cyanobacteriota**
CYAN140	**NR_172711**	**87.008**	**1.00E-76**	***Pseudochroococcus couteii***	**p__Cyanobacteriota**
CYAN141	**NR_172612**	**88.189**	**9.96E-82**	***Trichocoleus caatingensis***	**p__Cyanobacteriota**
CYAN142	**NR_114995**	**96.047**	**4.41E-115**	***Dulcicalothrix desertica***	**f__Rivulariaceae**
CYAN143	**NR_114406**	**92**	**1.28E-85**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN143	**NR_102447**	**92**	**1.28E-85**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN144	**NR_186887**	**99.605**	**4.32E-130**	***Microcoleusiopsis ganfuensis***	**g__*Arthrospira***
CYAN144	**NR_186886**	**99.605**	**4.32E-130**	***Microcoleusiopsis ganfuensis***	**g__*Arthrospira***
CYAN145	**NR_172612**	**92.46**	**1.62E-99**	***Trichocoleus caatingensis***	**c__Cyanophyceae**
CYAN146	**NR_177860**	**99.605**	**4.32E-130**	***Leptodesmis paradoxa***	**g__*Leptodesmis***
CYAN147	**NR_189686**	**91.304**	**4.54E-95**	***Scytolyngbya timoleontis***	**g__*Limnolyngbya***
CYAN147	**NR_189685**	**91.304**	**4.54E-95**	***Scytolyngbya timoleontis***	**g__*Limnolyngbya***
CYAN147	**NR_189684**	**91.304**	**4.54E-95**	***Scytolyngbya timoleontis***	**g__*Limnolyngbya***
CYAN147	**NR_172591**	**91.304**	**4.54E-95**	***Stenomitos hiloensis***	**g__*Limnolyngbya***
CYAN148	**NR_177874**	**90.551**	**3.53E-91**	***Monilinema alkalinum***	**g__*Petrachloros***
CYAN149	**NR_176486**	**96.443**	**9.48E-117**	***Limnococcus limneticus***	**f__Cyanothrichaceae**
CYAN150	**NR_177802**	**91.6**	**1.63E-94**	***Dolichospermum brachiatum***	**o__Nostocales**
CYAN150	**NR_177801**	**91.6**	**1.63E-94**	***Dolichospermum brachiatum***	**o__Nostocales**
CYAN150	**NR_177800**	**91.6**	**1.63E-94**	***Dolichospermum brachiatum***	**o__Nostocales**
CYAN151	**NR_191044**	**98.814**	**9.34E-127**	***Shahulinema minutum***	**g__*Shahulinema***
CYAN152	**NR_115288**	**93.281**	**2.08E-103**	***Crocosphaera watsonii***	**c__Cyanophyceae**
CYAN153	**NR_176580**	**86.275**	**2.17E-73**	***Sirenicapillaria stauglerae***	**p__Cyanobacteriota**
CYAN154	**NR_177763**	**91.304**	**4.54E-95**	***Annamia dubia***	**g__Microcystaceae_X**
CYAN154	**NR_177026**	**91.304**	**4.54E-95**	***Annamia dubia***	**g__Microcystaceae_X**
CYAN154	**NR_186889**	**91.304**	**4.54E-95**	***Annamia toxica***	**g__Microcystaceae_X**
CYAN155	**NR_102457**	**95.652**	**2.05E-113**	***Anabaena cylindrica***	**g__*Heteroscytonema***
CYAN156	**NR_104911**	**85.771**	**1.01E-71**	***Vampirovibrio chlorellavorus***	**g__Candidatus-Obscuribacter**
CYAN157	**NR_189176**	**94.862**	**4.44E-110**	***Tumidithrix elongata***	**g__*Tumidithrix***
CYAN158	**NR_176550**	**90.945**	**2.11E-93**	***Oculatella crustae-formantes***	**c__Cyanophyceae**

CYAN159	**NR_189176**	**91.2**	**2.11E-93**	***Tumidithrix elongata***	**c__Cyanophyceae**
CYAN160	**NR_172598**	**95.257**	**9.54E-112**	***Cymatolege isodiametrica***	**g__Heteroleibleinia**
CYAN160	**NR_172572**	**95.257**	**9.54E-112**	***Cymatolege spiroidea***	**g__Heteroleibleinia**
CYAN161	**NR_176478**	**90.196**	**1.64E-89**	***Arthronema gygaxiana***	**g__*Limnothrix***
CYAN162	**NR_191044**	**98.419**	**4.35E-125**	***Shahulinema minutum***	**g__*Shahulinema***
CYAN163	**NR_114406**	**94**	**9.96E-82**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN163	**NR_102447**	**94**	**9.96E-82**	***Cyanobium gracile***	**f__Prochlorococcaceae**
CYAN164	**NR_172612**	**88.933**	**1.66E-84**	***Trichocoleus caatingensis***	**f__Prochlorococcaceae**
CYAN165	**NR_197630**	**97.233**	**4.38E-120**	***Geminocystis urbisnovae***	**g__*Geminocystis***

Appendix C. Redundant Amplicon Sequence Variants Removed from Phylogenetic Analysis

Table C.1 CyanoSeq v1.3 Sequences removed due to redundancy

Classification	Sets	Sequences per set
g__AG-363-K07/g__AG-409-L18/g__AG-409-B05/g__AG-409-J16/g__CACEWU01/g__AG-409-P03/g__AG-463-P14/g__AG-409-O23/g__AG-409-M05/g__AG-409-J19/g__AG-409-B13	1	13
g__AG-409-A19/g__AG-363-P08/g__AG-363-J23	1	3
g__AG-409-N21/g__AG-363-P08	1	2
g__Acaryochloris	2	11,3
g__Aegeococcus	2	2,3
g__Aerofilum	1	4
g__Aerosakkonema	1	8
g__Aetokthonos	1	2
g__Affixifilum	2	3,9
g__Ahomia/g__Amazonocrinis	3	4,4,6
g__Ahomia/g__Aulosira	1	11
g__Albertania	2	3,7
g__Albertania/g__Egbenema	1	3
g__Aliinostoc	2	3,4
g__Aliterella	3	2,2,2
g__Alkalinema	1	8
g__Allocoleopsis	1	3
g__Altericista	3	2,2,2
g__Amazoninema	2	3,4
g__Amazonocrinis	2	2,2
g__Ammassolinea	1	6

g__Amphiheterocytum	1	3
g__Anabaena	1	37
g__Anabaenopsis	3	4,5,8
g__Anagnostidinema	1	2
g__Ancylothrix	2	4,4
g__Annamia	1	2
g__Apatinema	1	2
g__*Aphanizomenonaceae*_XXX	1	30
g__Aphanothece	1	5
g__Argonema	2	13,2
g__Atlanticothrix/g__Halotia	1	21
g__Aulosira	1	15
g__Aulosira/g__Parakomarekiella/g__Roholtiella	1	21
g__Azorthrix/g__Dactylothamnos	1	11
g__Baaleninema	2	4,6
g__Brasilonema	11	2,2,2,2,2,3,3,3,4,4,7
g__CACIRS01/g__AG-363-J23/g__Prochlorococcus_B	1	3
g__Caldora	2	22,3
g__Calenema	1	2
g__Candidatus-Atelocyanobacterium	1	2
g__Candidatus_Sivonenia	1	2
g__Capilliphycus	3	2,2,3
g__Cartusia	1	2
g__Cephalothrix	1	19
g__Chakia	1	3
g__Chamaesiphon	7	2,2,3,5,5,5,7
g__Chlorogloea	1	2
g__Chlorogloeopsis	1	8
g__Chondrocystis	1	2
g__Chroakolemma	1	7

g__Chroococcidiopsis	1	3
g__Chroococcus	1	5
g__Chroogloeocystis/g__Gloeocapsa	1	5
g__Chrysosporum	3	2,5,8
g__Ciimarium/g__Cyanobium_A	1	2
g__Coleofasciculus	6	11,16,2,3,4,4
g__Compactococcus	1	2
g__Compactonostoc/g__Komarekiella	1	3
g__Constrictifilum	1	2
g__Conticribra/g__Thalassiosira	1	2
g__Copelandiella	1	4
g__Crassifilum	1	5
g__Crinalium	1	3
g__Crocosphaera	2	2,3
g__Cronbergia	1	6
g__Crustifilum	1	2
g__Cryptochroococcus	1	11
g__Cryptococcum	1	4
g__Cyanoarbor	1	3
g__Cyanobium/g__Ciimarium	1	2
g__Cyanomargarita	2	3,4
g__Cyanospira	2	2,5
g__Cyanothece	2	2,8
g__Cyclotella	1	2
g__Cylindrospermum	6	2,2,2,3,5,8
g__Cymatolege	4	2,2,2,4
g__Dactylothamnos/g__Rexia	1	3
g__Dapis	5	2,2,2,8,8
g__Dendronalium	1	3
g__Dermocarpella	1	2

g__Desertifilum	1	8
g__Desikacharya	1	2
g__Desmonostoc	2	2,4
g__Drouetiella	1	2
g__Dulcicalothrix	1	2
g__Edaphophycus	1	2
g__Euhalothece	1	2
g__Euhalothece/g__Halothece	1	2
g__Euryhalinema	1	8
g__Fischerella/g__Hapalosiphon/g__Westiellopsis	1	37
g__Floridaenema	2	2,3
g__Foliisarcina	1	3
g__Fortiea	1	2
g__Fulbrightiella	3	2,2,6
g__Geitleria	1	3
g__Geitleribactron	1	2
g__Geitlerinema	2	2,2
g__Gibliniella	1	3
g__Gloeobacter	2	2,7
g__Gloeocapsopsis	1	2
g__Gloeothece	1	3
g__Godleya/g__Toxopsis	1	6
g__Goleter	2	2,9
g__Gracilinea	1	2
g__Haloleptolyngbya	1	2
g__Halospirulina	2	2,3
g__Halotia	3	2,3,3
g__Hassallia	1	4
g__Hillbrichtia	1	2
g__Hormoscilla	1	4

g__Hydrocoryne	8	2,23,3,3,3,5,6,7
g__Inacoccus	1	3
g__Iningainema	1	4
g__Inmanicoccus	1	3
g__Iphinoe	1	2
g__Jaaginema	5	2,2,3,3,4
g__Jacksonvillea	1	4
g__Kaiparowitsia	1	3
g__Kamptonema	1	4
g__Karukerafilum	1	2
g__Kastovskya	1	10
g__Katagnymene	1	2
g__Khargia	3	2,2,4
g__Klebsiella/g__Kluyvera	1	2
g__Klisinema	2	2,7
g__Koinonema	1	2
g__Komarekiella	1	9
g__Komarkovaea	1	2
g__Komarkovaeasiopsis	1	2
g__Komvophoron	3	2,2,2
g__Konicacronema	5	2,3,3,5,7
g__Kovacikia	2	2,2
g__Kryptousia	2	2,5
g__Lagosinema	1	2
g__Laspinema	1	3
g__Leptodesmis	2	2,2
g__Leptolyngbya	2	2,5
g__Leptolyngbya/g__Planktolyngbya	1	3
g__Leptothoe	6	2,2,2,2,4,8
g__Limnolyngbya	2	2,3

g__Limnospira	1	2
g__Lusitaniella	1	7
g__Macrochaete	4	2,2,3,5
g__Maricoleus	1	2
g__Marileptolyngbya	1	2
g__Mastigocladus	1	2
g__Merismopedia	1	3
g__Microcoleus	1	6
g__Microcrocis	1	8
g__Microseira	7	2,2,2,3,3,4,4
g__Minunostoc/g__Desikacharya	1	2
g__Mojavia	1	3
g__Monilinema	1	5
g__Moorena	4	38,4,6,7
g__Myxacorys	8	18,3,3,3,3,4,7,9
g__Myxosarcina	1	2
g__Neochroococcus	1	2
g__Neocylindrospermum	1	2
g__Neolyngbya	7	17,2,4,4,6,6,9
g__Neowestiellopsis	1	16
g__Nodosilinea	21	10,2,2,2,2,2,2,2,2,2,3,3,3,4,5,5,5,6,6,7,8
g__Nodularia	2	2,5
g__Nodulariaceae_XXX	1	4
g__Nunduva	2	3,7
g__Oculatella	9	2,2,2,2,2,3,4,4,9
g__Odorella	1	5
g__Okeania	6	18,2,4,5,5,6
g__Onodrimia	1	3
g__Ophiophycus	1	2
g__Oscillatoria	3	2,3,6

g__Oxynema	1	8
g__Pantanalinema	2	2,9
g__Parasynechococcus	2	2,4
g__Parasynechococcus/g__Prochlorococcus	1	11
g__Parathermosynechococcus	1	2
g__Pegethrix	4	2,2,2,2
g__Persinema	1	8
g__Petrachloros	1	2
g__Phaeodactylum	1	2
g__Phayaothrix	1	8
g__Phormidesmis	3	2,2,2
g__Phyllonema	2	3,4
g__Picosynechococcus	1	4
g__Pinocchia	2	4,4
g__Pinocchia/g__Trichothermofontia	1	5
g__Planktothricoides	3	13,2,3
g__Planktothrix	5	2,33,4,6,6
g__Plectolyngbya	1	6
g__Plectolyngbya/g__Radiculonema	2	2,2
g__Pleurocapsa	1	2
g__Porphyrosiphon	1	2
g__Potamolinea	3	2,24,3
g__Potamosiphon	2	2,3
g__Prochlorococcus	2	2,2
g__Prochlorococcus_B	5	2,2,3,3,4
g__Prochlorococcus_C/g__AG-409-D09	1	2
g__Prochloron	1	3
g__Prochlorothrix	2	2,3
g__Pseudanabaena/g__**Pseudanabaenaceae*****_*X***	1	2
g__**Pseudanabaenaceae*****_*X*/*g__Tumidithrix***	1	2

g__Pseudoaliinostoc	1	2
g__Pseudocalidococcus/g__Stenotopis	3	11,2,4
g__Pseudochroococcus	1	2
g__Pseudocyanosarcina	1	2
g__Pseudosynechococcus	1	3
g__Pycnacronema	5	2,2,3,4,6
g__Ramsaria	2	11,2
g__Raphidiopsis	6	17,2,2,2,2,58
g__Reptodigitus	1	3
g__Rhizonema	1	2
g__Roholtiella	1	2
g__Romeriopsis	2	2,3
g__Roseofilum	6	2,2,2,2,5,6
g__Scytolyngbya	1	4
g__Scytonematopsis	3	2,2,2
g__Shahulinema	1	2
g__Sherwoodiella	4	2,2,2,2
g__Siamcapillus	1	2
g__Sinocapsa	3	2,2,3
g__Sirenicapillaria	2	2,4
g__Sodaleptolyngbya	1	3
g__Sodalinema	3	15,20,3
g__Solentia	1	2
g__Sphaerospermopsis	3	13,44,5
g__Spirirestis	1	2
g__Spirulina/g__Thermospirulina	1	3
g__Stanieria	1	6
g__Stenomitos	2	2,3
g__Stigonema	3	3,4,5
g__Symploca	2	2,3

g__Symplocastrum	3	2,3,5
g__Synechococcus	1	2
g__Synechocystis	1	4
g__Synechospongium	1	2
g__Tenebriella	3	11,3,5
g__Thainema	3	2,2,2
g__Thalassoporum	1	2
g__Thermocoleostomius	1	3
g__Thermoleptolyngbya	3	11,2,3
g__Thermospirulina	1	2
g__Thermosynechococcus	1	2
g__Tigrinifilum	2	4,9
g__Tildeniella	1	8
g__Timaviella	4	2,2,3,3
g__Toxifilum	1	5
g__Toxopsis	1	3
g__Trichocoleus	8	10,2,2,2,2,4,7,9
g__Trichocoleusaceae_X	1	2
g__Trichodesmium	3	2,2,2
g__Trichothermofontia	1	2
g__Trichotorquatus	3	18,2,3
g__Tumidithrix	1	2
g__Tychonema	1	7
g__Umezakia	1	22
g__Venetifunis	1	2
g__Vermifilum	2	3,4
g__Wilmottia	1	2
g__Xenococcus	1	3

Table C.2 CyanoSeq v1.3 Sequences removed due to redundancy.

Classification	Sets	Sequences per set	URSA IDs
g__Lacustricoccus	**5**	**1*,2*,1*,1*,2***	**CYAN022,CYAN03, CYAN072, CYAN016 CYAN077**
g__Microcystis	**2**	**3*,2***	**CYAN028,CYAN013**
g__Vulcanococcus	**6**	**2,3,2*,1*,1*,1***	**CYAN020,CYAN003, CYAN012**
g__Arthrospira	**2**	**8,3***	**CYAN144**
g__Dolichospermum	**19**	**10,17,2,2,2,2,2,2,2,3,3,5,7,1*,1*,1*,1*,1*,14***	**CYAN025,CYAN045, CYAN005,CYAN019 CYAN047,CYAN079**
g__Pseudanabaena	**7**	**2,2,2,2,3*,1*,1***	**CYAN051,CYAN139, CYAN075**
g__Woronichinia	**1**	**2***	**CYAN054**
g__Heteroscytonema	**2**	**15,2***	**CYAN155**
g__Radiocystis	**1**	**2***	**CYAN063**
g__Gloeotrichia/g__Johanseniella	**1**	**5***	**CYAN052**
g__Limnococcus	**1**	**1***	**CYAN105**
g__Limnoraphis	**1**	**18***	**CYAN078**
g__RGUW01	**1**	**1***	**CYAN035**
g__Cuspidothrix	**5**	**2,3,33,7,8***	**CYAN055**
g__Trichormus	**2**	**2,2***	**CYAN080**
g__Nostoc	**4**	**2,2,1*,2***	**CYAN097 CYAN032**
g__Coelosphaerium g__Snowella	**1**	**3***	**CYAN090**
g__Limnothrix g__Pseudanabaena	**1**	**3***	**CYAN094**
g__Regnicoccus	**4**	**2,5,2*,3***	**CYAN102,CYAN099**
g__Cyanodorina	**1**	**4***	**CYAN067**
g__Aphanizomenon	**3**	**2,3,7***	**CYAN011**
g__Snowella	**2**	**2,1***	**CYAN101**
g__Cyanobium	**3**	**2,2,6***	**CYAN043**

*Note: The * represents sets that contain URSA IDs.*

Appendix D. Centre Log-Ratio Redundancy Analysis Summary

This section provides the summary statistics for the Centre Log-Ratio (CLR) transformed Redundancy Analysis (RDA) assessing the relationship between environmental variables and microbial community composition in the URSA ponds.

Table D.1 ANOVA Permutation Test Results for CLR-RDA Model

Term	Df	Variance	F	Pr(>F)
Model	6	550.12	1.4146	0.011
Residual	27	1749.95		

Note: Based on 999 permutations.

Table D.2 CLR-RDA Model Summary - Variance Explained

Component	Inertia	Proportion	Rank
Total	2300.0701	1.0000	
Constrained	550.1185	0.2392	6
Unconstrained	1749.9515	0.7608	27

Table D.3 Eigenvalues for CLR-RDA Axes

Axis	Constrained Eigenvalue	Axis	Unconstrained Eigenvalue
RDA1	181.32	**PC1**	205.00
RDA2	103.86	**PC2**	147.80
RDA3	86.88	**PC3**	129.74
RDA4	74.36	**PC4**	104.59
RDA5	63.91	**PC5**	104.03
RDA6	39.78	**PC6**	84.92
		PC7	75.56
		PC8	73.52

Appendix E. Raw Reads Redundancy Analysis Summary

This section provides the summary statistics for the raw reads assessing the relationship between environmental variables and non-normalized microbial community composition in the URSA ponds.

Table E.1 ANOVA Permutation Test Results for Raw Read RDA Model

Term	Df	Variance	F	Pr(>F)
Model	6	11503733	1.1648	0.279
Residual	27	44444076		

Note: Based on 999 permutations.

Table E.2 Raw Read RDA Model Summary - Variance Explained

Component	Inertia	Proportion	Rank
Total	5.595e+07	1.000e+00	
Constrained	1.150e+07	2.056e-01	6
Unconstrained	4.444e+07	7.944e-01	27

Table E.3 Eigenvalues for Raw Read RDA Axes

Axis	Constrained Eigenvalue	Axis	Unconstrained Eigenvalue
RDA1	6332784	**PC1**	29040222
RDA2	2605206	**PC2**	3730308
RDA3	1268338	**PC3**	2742539
RDA4	760407	**PC4**	1787649
RDA5	364270	**PC5**	1572741
RDA6	172727	**PC6**	817527
		PC7	579658
		PC8	565481