Beyond Microcystins: Uncovering the Drivers and Dangers of Co-Occurring Cyanobacteria and Cyanopeptides in Eutrophic Great Lakes Basins

by

Lauren N. Hart

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2025

Doctoral Committee:

Professor Gregory J. Dick, Co-Chair
Professor David H. Sherman, Co-Chair
Professor Marcy J. Balunas
Professor Matthew R. Chapman
Dr. Reagan M. Errera, NOAA Great Lakes Environmental Research Laboratory
Assistant Professor Jenan J. Kharbush

Lauren N. Hart

lnhart@umich.edu

ORCID iD: 0000-0001-8553-2839

© Lauren N. Hart 2025

Dedication

To my parents and brothers, who nurtured my curiosity for the natural world from a young age, who stood behind me through every challenge with unwavering love, belief, and humor, and who taught me the quiet power of grit.

And to all Women in STEM—those who broke barriers, the brilliant collaborators, and the fierce thinkers still to come. This is for those who built a path, walked beside me on it, and inspired me to leave it wider for others.

Acknowledgements

The completion of this dissertation would not have been possible without the guidance of various mentors, colleagues, and funding agencies. First, I would like to thank my mentors Greg Dick and David Sherman. You played instrumental roles in my scientific training and development as a collaborator. Through patience, thoughtful feedback, curiosity-inducing conversations, and willingness to accommodate my learning and personal needs, I became a better scientist and mentor. Within the laboratories of Greg and David were highly collaborative and friendly research groups who were invaluable to my success. While there are far too many colleagues and friends to list, I would like to personally thank Colleen Yancey, Katherine Lev, Brittany Zepernick, and Kate Brown for your collaboration, mentorship, and friendship. Additionally, I would like to thank the undergraduates in Greg's lab for maintaining the western Lake Erie culture collection, and Natalie Keating and Ryann Jibson for providing me an opportunity to practice mentorship in a fun and creative atmosphere. I would like to thank the members of my dissertation committee who brought fresh perspectives, insightful feedback and questions, and encouragement every time we met. Thank you to Rose Cory, Jenan Kharbush, and Reagan Errera for your invaluable perspectives on the broader impacts of my research as it relates to nutrient cycling and water chemistry, and for serving as role models of determined and collaborative women in STEM.

I would like to thank the staff at University of Michigan that assisted me in navigating the PhD journey. Thank you to collaborators at the Life Sciences Institute, including Pam Schultz who coordinated shipments of samples from Kenya and Wendy Feng who provided patient and

informative mass spectrometry training and help. I thank the Program in Chemical Biology administrative team, especially Traci Swan, Cameron Yaxley, and Alison Narayan for helping me navigate this journey and accommodating my learning needs.

I would also like to thank the mentors, friends, and family who shaped my path and made this journey so meaningful. First, thank you Tony Gonzalez, Brian Sedio, and Colin Morrison for sparking my interest in chemical ecology and helping me build the confidence and skills to take this step in my career. Your creative approaches to science and mentorship are something I carry with me to this day. Thank you to my friends Caitlin, Elena, Noelle, Meaghan, Colleen, and Helena for a never-ending supply of support and laughs, even from more than a thousand miles away. Thank you to my parents, brothers Brian and Eric, and my partner Will for your endless support, I couldn't have done this without you.

This work would not have been possible without extensive assistance with field, lab, and bioinformatic support that began long before I started my research, particularly NOAA Great Lakes Environmental Research Laboratory (NOAA GLERL), the Cooperative Institute for Great Lakes Research (CIGLR), the Great Lakes Center for Fresh Waters and Human Health, and Kenya Marine and Fisheries Research Institute (KMFRI). From these networks, I would like to thank David Kennedy, Steven Haller, Dragan Isailović, Andrew Fribley, Hans Paerl, Justin Chaffin, Steve Willhelm, Eric Owino, James Achiya, and Casey Godwin for their close collaboration on these projects. Thank you to Captains Kent Baker and Frederick Okello for informative, efficient, and safe research cruises on western Lake Erie and Winam Gulf. Thank you to Paul Den Uyl, Andrew Camilleri, and James Achiya for collaborating in the field together and assisting in field measurements. Thank you to the many folks of NOAA GLERL/CIGLR HAB monitoring team for providing robust physicochemical measurements from all cruises. Thank you to Karly Clippinger,

Nathan LaFramboise, NOAA GLERL, the Kling Lab, and past members of Rose Cory's Lab for their contributions to curating the fluorometric dissolved organic matter dataset utilized in Chapter 2. Thank you to collaborators at Kisii University and Technical University of Kenya for field and data analysis support. Thank you to Sara Rivera, Nicole Rappuhn, McKenzie Powers, Katherine Polik, Helena Nitschky, Natalie Keating, and undergraduates in Greg Dick's lab for assistance with sample collection and maintenance of the western Lake Erie culture collection. Thank you to Aaron Robida and the University of Michigan Center for Chemical Genomics for assistance with human cell line maintenance. Thank you to Robert Hein for extensive and timely bioinformatic support. Thank you to the network of collaborators that supported my academic journey and led to the accomplishments of this dissertation. I would especially like to thank George Bullerjahn and Mike McKay, who played pivotal roles in my journey, not only as mentors, but as a spark for my curiosity and a reliable source of encouragement and good humor.

This dissertation was supported by funds received through: NIH and NSF awards to the Great Lakes Center for Fresh Waters and Human Health at Bowling Green State University (BGSU) and University of Michigan (UM) (NIH Awards: 1P01ES0289-01 (BGSU), 5P01ES028939-02 (UM), NSF Awards: 1840715 (BGSU), OCE-1840715 (UM)), NSF IRES (1953468), NIH F31 (1F31ES036421-01), Great Lakes Restoration Initiative through U.S. E.P.A. and the NOAA Omics program (NOAA 'Omics OAR NO_0012, NO_0045, NO_0086) awarded to Cooperative Institute for Great Lakes Research (CIGLR) through the NOAA Cooperative Agreement with the University of Michigan (NA22OAR4320150 and NA17OAR4320152), the USGS Environmental Health Program Toxic Substances Hydrology Program, and University of Michigan Rackham Graduate Student Research Grants.

Table of Contents

Dedication ii

Acknowledgements iii

List of Tables xii

List of Figures xiv

Abstract xxv

Chapter 1 Introduction 1

1.1 Introduction and Overview 1

1.2 CyanoHABs in Eutrophic Great Lakes 4

1.2.1 CyanoHABs in Western Lake Erie, United States: Past and Present 4

1.2.2 CyanoHABs in Winam Gulf, Lake Victoria, Kenya: Past and Present 6

1.2.3 Implications for Environmental, Human, and Economic Prosperity 8

1.3 *Microcystis* and *Dolichospermum* Bloom Formation: Ecophysiology and Drivers of Dominance 12

1.3.1 Ecophysiology of Microcystis 12

1.3.2 Ecophysiology of Dolichospermum 16

1.3.3 Dueling Bloom Dynamics of Microcystis and Dolichospermum in Shared Waters 20

1.4 The Diverse Biosynthetic Repertoire of *Microcystis* 21

1.4.1 Characterized Secondary Metabolites 22

1.4.2 Cryptic Microcystis Secondary Metabolism 23

1.5 The Understudied Biosynthetic Potential of *Dolichospermum* 25

1.6 Environmental Drivers of *Microcystis* and *Dolichospermum* Secondary Metabolites 27

1.7 Combinatorial Toxic Effects of Cyanobacterial Secondary Metabolites on Human and Environmental Health 30

1.7.1 Combinatorial Toxic Effects of Cyanotoxins on Aquatic Organisms 30

1.7.2 Combinatorial Toxic Effects of Cyanotoxins on Mammalian Cell Models 34

1.8 Dissertation Outline and Research Questions 36

1.9 References 38

1.10 Figures 61

Chapter 2 Diverse Cyanopeptides Follow Distinct Temporal Succession Patterns in Lake Erie Harmful Algal Blooms 68

2.1 Abstract 68

2.2 Introduction 69

2.3 Materials and Methods 72

2.3.1 Sample Collection 72

2.3.2 DNA Extraction and Sequencing 72

2.3.3 Metabolite Extraction and Data Collection 73

2.3.4 Metagenomic Data Processing and Analysis 73

2.3.5 BGC Detection from Metagenomes 74

2.3.6 Biosynthesis Module Analysis 75

2.3.7 Hierarchical General Additive Models (HGAMs) 75

2.3.8 Microcystis Isolate Analysis 76

2.3.9 Metabolomics Analysis 77

2.3.10 Figure Rendering 78

2.4 Results 78

2.4.1 Succession of Microcystis to Dolichospermum is Coincident with Changes in Temperature and Nutrient Availability 78

2.4.2 Cryptic BGCs are the Most Abundant BGCs Encoded by Microcystis and Dolichospermum 80

2.4.3 BGC Co-Variance Forms Distinct Biosynthesis Modules with Varying Temporal Patterns 81

2.4.4 Biosynthesis Modules Reflect Mosaics of Microcystis Isolates and their BGC Content 81

2.4.5 Uncoupled Temporal Patterns of Cyanopeptides and Associated BGCs 82

2.4.6 Unique Cyanopeptide Maxima are Associated with Shifts in Nitrogen, DOM, and Grazer Abundance 83

2.5 Discussion 84

2.6 Data Availability 88

2.7 References 89

2.8 Figures 97

2.9 Supplementary Information 104

2.9.1 SI Tables 104

2.9.2 SI Figures 134

2.9.3 SI Notes 159

2.9.4 Supplemental Notes References 163

Chapter 3 Metagenomics Reveals Spatial Variation in Cyanobacterial Composition, Function, and Biosynthetic Potential in the Winam Gulf, Lake Victoria, Kenya 165

3.1 Abstract 165

3.2 Introduction 166

3.3 Results 170

3.3.1 Water Physicochemistry Indicates Nitrogen Limitations Throughout the Gulf 170

3.3.2 River and Gulf Sites Vary in Microbial Community Makeup 170

3.3.3 Dolichospermum Dominates Large CyanoHAB Events Detected in Distinct Gulf Regions 171

3.3.4 Relationships Between Dominant Cyanobacterial Genera and Environmental Conditions 172

3.3.5 Phylogenomic Analysis Reveals Low Intra-Genus Genetic Diversity and Novel Clades of Bloom-Forming Cyanobacteria from the Gulf 173

3.3.6 Nitrogen Fixation Genes Harbored Primarily by Dolichospermum and Identified at Sites with Large Bloom Events 174

3.3.7 Cyanotoxin- and Cyanopeptide-Encoding BGCs Detected in Regions with and without Visible CyanoHABs 174

3.3.8 Genus-Specific Biosynthetic Potential Revealed Through Over 300 Cryptic BGCs 175

3.3.9 Biosynthetic Potential Converges with Site-Specific Cyanobacterial Composition 176

3.4 Discussion 177

3.4.1 Nitrogen Limitation May Provide a Competitive Advantage for Diazotrophs During CyanoHABs 177

3.4.2 Non-diazotrophic Cyanobacteria were Identified at Lower Abundance in Gulf Regions with Distinct Physicochemical Conditions 179

3.4.3 Low Intra-Genus Diversity Among Bloom-Forming Cyanobacteria in the Gulf Compared to Temperate Systems 180

3.4.4 BGCs Encoding the Synthesis of Cyanotoxins, Cyanopeptides, and Unknown Metabolites Found Across the Gulf, Even in Areas Without Visible CyanoHABs 181

3.5 Conclusion 183

3.6 Materials and Methods 183

3.6.1 Data Collection 183

3.6.2 Bioinformatic Analysis 186

3.7 Data Availability and Inclusion and Ethics Statement 191

3.8 References 192

3.9 Figures and Tables 200

3.10 Supplementary Information 214

Chapter 4 Cyanopeptide Mixtures Induce Variable Synergistic and Antagonistic Effects Across Diverse Human Cell Lines 254

4.1 Abstract 254

4.2 Introduction 255
4.3 Materials & Methods 258
4.3.1 Cell Line Rationale 258
4.3.2 Hep-3B Cells 258
4.3.3 A549 Cells 259
4.3.4 HK-2 Cells 259
4.3.5 Microcystis Culturing 259
4.3.6 Microcystis Culture Biomass Collection and Metabolite Extraction 260
4.3.7 Targeted Metabolomics Data Collection and Analysis 261
4.3.8 Cell Line Exposures to Individual Pure Cyanopeptide Standards 261
4.3.9 Cell Line Exposures to Combinations of Pure Cyanopeptide Standards 262
4.3.10 Quantifying Cyanopeptide Combination Activity 263
4.3.11 Cell Line Exposures to Culture Extracts 264
4.3.12 Statistical Analysis 265
4.3.13 Figure Rendering 266
4.4 Results 266
4.4.1 Individual Cyanopeptides Exhibit Dose-Dependent Responses Across Different Cell Lines 266
4.4.2 Cyanopeptide Combinations Exhibit Dose-Dependent Synergistic and Antagonistic Responses Across Different Cell Lines 267
4.4.3 Microcystis Culture Extracts Display Dose-Dependent Effects Not Explained by Cyanopeptide Concentrations Alone 268
4.5 Discussion 270
4.6 Conclusion 276
4.7 References 277
4.8 Figures 286

4.9 Supplementary Information 294

4.9.1 SI Tables 294

4.9.2 SI Figures 330

Chapter 5 Conclusions 333

5.1 Summary and Overview 333

5.2 Looking Beyond Microcystins in Freshwater CyanoHABs: Conclusions, Opportunities, and Future Directions 335

5.2.1 Implications for Water Quality Monitoring in a Changing Climate 335

5.2.2 Insights into the Chemical Ecology and Functions of the Diverse Biosynthetic Arsenals of Bloom-Forming Cyanobacteria 338

5.2.3 Improved Understanding of Global CyanoHABs through Collaboration 339

5.2.4 Toxicity beyond Microcystins: Assessing, Monitoring, and Managing the Combinatorial Effects of Co-Occurring Cyanobacterial Metabolites 341

5.3 Continuing Advancements for Future Applications 342

5.4 Concluding Remarks 344

5.5 References 345

Appendix 351

List of Tables

Table S2.9.1: Metadata for all samples used in this study, including sampling and physicochemical parameter information. 104

Table S2.9.2: HPLC Solvent Gradient. 124

Table S2.9.3: Biosynthetic gene cluster (BGC) database information 125

Table S2.9.4: *Microcystis* genomes used for phylogenomic analysis 127

Table 2.9.5: MZmine Parameters used in the feature-finding processing wizard. 132

Table S2.9.6: Hierarchical general additive models (HGAM) equation and results for the relative abundance (%) of *Microcystis* and *Dolichospermum* relative to the whole microbial community. 132

Table S2.9.7: R^2 and *P*-values for NMDS vectors in Figure 2C. 133

Table S2.9.8: Hierarchical general additive models (HGAM) equation and results for the FPKM of BGCs within modules 4, 5, 6 and 7. 133

Table S2.9.9: Residual standard errors of loess fits of metabolite, dissolved organic matter (DOM), gene, nitrogen, and grazer abundance patterns. 134

Table 3.9.1: Summary of sampled stations in this study. 201

Table 3.9.2: Summary and statistics of high-quality, dereplicated cyanobacterial MAGs. 205

Table S3.10.1: Table of spearman correlation coefficients between highly abundant cyanobacterial genera and associated metadata variables. 214

Table S3.10.2: Reference genomes used for phylogenomic analysis. 220

Table S3.10.3: Manual annotation notes for dereplicated set of BGCs identified from high quality cyanobacterial MAGs from the sampled sites in Winam Gulf. 224

Table S4.9.1: Single cyanotoxin exposure viability data. 294

Table S4.9.2: Combination cyanotoxin exposure viability data. 294

Table S4.9.3: Combination cyanotoxin exposure Loewe, Bliss, HAS, and ZIP synergy scores 294

Table S4.9.4: *Microcystis* isolate extract exposure viability data. 296

Table S4.9.5: Concentration of cyanotoxins detected via targeted LC-MS in each crude extract in the 100 μM dose exposed to each cell line in viability experiments. 320

List of Figures

Figure 1.10.1: Photos of cyanoHABs in global, eutrophic Great Lakes. **(A)** Aerial image of western Lake Erie in August of 2019. Photo taken by Zachary Haslick (NOAA). **(B)** Scum and foam from cyanoHAB in western Lake Erie in July 2021. Photo by Colleen E. Yancey. **(C)** CyanoHAB scum and water hyacinth along shores of Winam Gulf, Lake Victoria, Kenya. A person is seen washing dishes near the shore. Photo taken by Julia Obuya. 61

Figure 1.10.2: Representative structures of cyanobacterial metabolites. **(A)** Hepatotoxins: Microcystin LR, Anabaenopeptin A, Aeruginosin KB 676, Microviridin (PubChem ID: 156580592), Microginin 713, Cyanopeptolin A, Aerucyclamide A. Colored structural moieties represent variable parts that can be substituted with alternative building blocks, leading to congener diversity within each metabolite class. **(B)** Neurotoxins: Saxitoxin, Anatoxin-A, β-N-methylamino-L-alanine (BMAA), and guanitoxin (previously anatoxin-(S)). 62

Figure 1.10.3: Map of western Lake Erie, Lake Erie, U.S. with the Detroit and Maumee Rivers noted. Map pulled from Stadia Maps (https://stadiamaps.com/, accessed January 2025). 63

Figure 1.10.4: Cyanotoxins monitored by the Environmental Protection Agency (EPA) and their allowed levels. Drinking water and recreation levels come from EPA, World Health Organization (WHO), and Oregon Health Authority (USA). 64

Figure 1.10.5: Map of Winam Gulf, Lake Victoria, Kenya with the Kibuon and Nyando Rivers noted. Map pulled from Stadia Maps (https://stadiamaps.com/, accessed January 2025). 65

Figure 1.10.6: Microscopy photos taken of **(A)** Western Lake Erie culture collection *Microcystis* isolates LE18-22.4 and LE19-197.1 (24) and **(B)** samples taken from a plankton net tow sample from western Lake Erie containing *Microcystis* cells and other organisms. Photos taken by Helena Nitschky and Nicole Rappuhn. *Microcystis* cells generally range in size from between 1.5 to 7 μm (283). 66

Figure 1.10.7: Microscopy photos taken of **(A)** ADA cultures maintained at Kellogg Biological Station at Michigan State University and **(B)** samples taken from a plankton net tow sample from western Lake Erie containing ADA cells and other organisms. Photos taken by Helena Nitschky and Carol Rosenbaum. *Dolichospermum* cells generally range in size from between 3 to 10 μm (284). 67

Figure 2.8.1: Map of western Lake Erie showing NOAA long-term monitoring stations evaluated in this study and nearby rivers and cities. Near-shore sites WE2 and WE12 border the Ohio coastline, WE8 lies off the Michigan shoreline, and WE4 represents an offshore station. 97

Figure 2.8.2: (a) Temporal patterns of particulate microcystin (μg/L) and the relative abundance of *Microcystis* and *Dolichospermum.* All individual measurements are shown as points colored by year. Particulate microcystin concentrations are shown as a black smoothed line fitted using locally weighted regression (loess). *Microcystis* and *Dolichospermum* abundance are also black smoothed lines fit using generalized additive models (GAMs) with a smooth on DOY with a cubic regression function. **(b)** Samples were categorized based on their relative abundance differences between *Microcystis* and *Dolichospermum.* Pie charts show the distribution of samples in each stage for each year. Sample size (n) in stage 1, 2, and 3, are 85, 61, and 24, respectively. **(c)** Nonmetric MultiDimensional Scaling (NMDS) ordination of samples based on cyanobacterial community (genus level), with points colored by stage. Ellipses represent 95% confidence intervals around samples grouped by stages 1, 2, and 3. Arrows indicate the direction and strength of environmental variables, with longer arrows representing strong correlations with NMDS axes, and bolded variables indicating key drivers of ordination structure *(P*<0.05). Abbreviations: Soluble reactive phosphorus (SRP), Particulate Organic Nitrogen (PON), Particulate Organic Carbon (POC), Colored dissolved organic matter (CDOM). **(d)** Random forest abiotic environmental variable importance based on mean decrease in accuracy for predicting stage, where bars represent the relative importance of each abiotic variable in classifying samples into stages..... 98

Figure 2.8.3: Relative abundance of BGCs from *Microcystis* and the ADA clade. Tiles are colored by log-transformed average fragments per kilobase million (FPKM) for each BGC across sampling dates within each year, with warmer colors indicating higher relative abundance. Black squares denote values of zero. BGC pathways are indicated by the color of labels on the horizontal axis..... 99

Figure 2.8.4: (a) Biosynthesis modules determined by hierarchical clustering (average linkage) of temporal distributions of relative abundance of BGCs based on Bray-Curtis distance. The dark gray subtree contains *Microcystis* BGCs that occur consistently across samples, whereas the light gray subtree contains ADA BGCs and sporadically occurring *Microcystis* BGCs (zero observations in two or more years of the dataset). BGC pathways are color-coded according to the legend. An asterisk (*) denotes significantly supported modules based on permutation-derived clustering quality. Scale bars for trees are shown, where the length is equal to a branch length of 0.1. **(b)** Partial effects of DOY from module-specific HGAMs with the log-transformed average FPKM of BGCs from modules 4, 5, 6, and 7 as the response. Shaded ribbons indicate standard errors around the fitted temporal trends..... 100

Figure 2.8.5: (a) Percent of each biosynthesis module present in isolates from the WLECC. Isolates are hierarchically clustered using average-linkage. Tiles are filled in by the percent of each biosynthesis module found in each isolate, with brighter tiles indicating more of the module present in that isolate. **(b)** Relative abundance of six WLECC isolates across field samples. Points represent individual observations, relative to the whole microbial community, and smoothed lines were fitted using a loess function..... 101

Figure 2.8.6: (a) Representative chemical structures of putatively annotated metabolites across the dataset: Microcystin LR, Anabaenopeptin A, Aeruginosin 98A, and Aerucyclamide A. **(b)** Temporal distribution of the normalized abundance of metabolites (Microcystin = sum of MC-LR, MC-YR, MC-HilR, MC-RR, and MC-HtyR; Aeruginosin 98A; Anabaenopeptin = sum of

AP-A and AP-B; and Aerucyclamide A), plotted with the average FPKM of associated BGCs. Data were Z-score scaled with all individual observations less than or equal to one shown (6 observations excluded), where black dots correspond to BGC abundance and colored dots correspond to metabolite abundance. Smoothed lines for BGC and metabolite abundance, represented as black dashed and filled colored lines, respectively, were fitted using a loess function. 102

Figure 2.8.7: (a) Temporal patterns of the normalized abundance of metabolites (Microcystin = sum of MC-LR, MC-YR, MC-HilR, MC-RR, and MC-HtyR; Aeruginosin 98A; Anabaenopeptin = sum of AP-A and AP-B; and Aerucyclamide A) plotted with nitrogen forms/sources, DOM components, and grazer abundance. Metabolite data were Z-score scaled and plotted together, whereas all other variables were plotted in their original scales. Scales for each variable are as follows: Ammonia (µg L^{-1}), Nitrate + Nitrate (mg L^{-1}), *Daphnia*, *Dreissena*, and Amoeba (% Relative Abundance). Temporal patterns were visualized using a loess function. Dashed lines indicate predicted peak dates for each metabolite and are colored accordingly. Abbreviations: Particulate organic carbon: particulate organic nitrogen (POC/PON). **(b)** Predicted maxima and minima for environmental variables and metabolites were estimated using 1,000 bootstrapped GAMs with a cyclic smoother on DOY. Median peak dates and 95% confidence intervals are shown as points with horizontal bars. Maxima are indicated by solid lines with circles, whereas minima, predicted only for nitrogen- and DOM-related variables, are shown as dashed lines with stars. Metabolite maxima are shown as solid-colored lines in the top panel separated from the environmental variable predictions..... 103

Figure S2.9.1: Comparison of nitrate + nitrite (mg/L), extracted phycocyanin (µg/L), temperature (°C), particulate microcystin (µg/L), particulate organic carbon: particulate organic nitrogen ratio (POC:PON), and soluble reactive phosphorus (µg P/L) from all stations between the years of 2016 and 2022. Individual observations are plotted as filled in circles, with observations below the detection limit plotted as unfilled triangles. Values under the detection limit were only identified for particulate microcystins and nitrate + nitrite and were imputed as half the limit of detection. Smoothed lines were fit using a loess function..... 135

Figure S2.9.2: (a-jj) Schematics of biosynthetic gene clusters used in the study, generated via antiSMASH v7 (1). 141

Figure S2.9.3: Mirror matches generated by GNPS of library matches detected and used in this study. The black spectrum on top represents the sample spectrum whereas the bottom green spectrum represents the database spectrum for the library match. All metabolite predictions shown here meet at least level 3 criteria as outlined in Schymasnki et al. 2014 (2). **(a)** Spectra putatively annotated as microcystins (MC) annotated as the congeners LR (cosine = 0.836), YR (cosine = 0.822), HilR (cosine = 0.619), RR (cosine = 0.756), HtyR (cosine = 0.736) are shown, **(b)** Aeruginosin 98A (cosine = 0.626), **(c)** Anabaenopeptins A (cosine = 0.855) and B (cosine = 0.912), and **(d)** Aerucyclamide A are shown (cosine = 0.2509). 146

Figure S2.9.4: Fixed and partial effects of day of year (DOY), year, and model diagnostics for hierarchical generalized additive models (HGAMs) fit to log-transformed relative abundance of **(a)** *Microcystis* and **(b)** *Dolichospermum***.** The first panels display the fixed effect smooths for DOY based on the HGAM equation described in SI Table 2.9.6. Diagnostic plots (second

panels) indicate that residuals are homoscedastic and approximately normally distributed (top right panel) and fitted values closely follow observed values (bottom right panel), supporting model adequacy for both taxa. The third panels show partial effect smooths of DOY stratified by year, with year-specific smooths illustrating interannual variation in seasonal patterns of each organism. Tick marks along the x-axis indicate the distribution of observed values. **(c)** A pairwise correlation plot of log-transformed *Microcystis* and *Dolichospermum* relative abundance reveals a modest but significant negative association (Pearson's r = –0.26, $p = 6x10^{-4}$), suggesting potential ecological interaction or shared environmental controls between the two taxa. Because these two response variables are derived from the same biological system, it is likely they are correlated in some manner. However, models were fit independently, and potential interdependence was not explicitly accounted for in the analysis 147

Figure S2.9.5: Seasonal pattern of relative abundance (%) of **(a)** *Microcystis* and **(b)** *Dolichospermum* across all stations and years. Points are colored by stage, where Stage 1 indicates *Microcystis* > *Dolichospermum* by 3% or more, Stage 2 indicates approximately equal populations of *Microcystis* and *Dolichospermum* (±3%), and Stage 3 indicates *Dolichospermum* > *Microcystis* by 3% or more. These stages were determined based on the distribution of their relative abundance relative to one another. Seasonal patterns faceted by station (WE12, WE2, WE4, and WE8) and year for **(c)** *Microcystis* and **(d)** *Dolichospermum* relative abundance (%) are also shown, with points colored by stage..... 148

Figure S2.9.6: NMDS stress plot assessing the goodness-of-fit of a non-metric multidimensional scaling (NMDS) ordination based on physicochemical parameters. The plot compares observed dissimilarities (x-axis) to distances between points in the ordination space (y-axis). Blue points represent pairwise distances; the red line shows a monotonic fit between observed and ordination distances. The non-metric fit ($R^2 = 0.995$) and linear fit ($R^2 = 0.982$) indicate an excellent correspondence between the original distance matrix and the reduced-dimensional NMDS space..... 149

Figure 2.9.7: Boxplots showing the distribution of 18 physicochemical parameters across stages. Each point represents an individual sample, grouped by stage classification. Horizontal bars denote statistically significant pairwise differences between stages for each parameter, based on Kruskal–Wallis tests followed by Dunn's post hoc comparisons. Significance levels are indicated as: **** $P<0.0001$; *** $P<0.001$; ** $P<0.01$; * $P<0.5$; n.s. = not significant). ... 150

Figure S2.9.8: Patterns in surface water temperature from NOAA GLERL WE2 Nutrient Buoy from 2015 through 2023. Data was collected from Great Lakes Observing System (GLOS) archive (3). **(a)** Surface water temperatures (°C) in September and October colored by year. **(b)** Frequency of 24-hour temperature changes in September classified as warming (red) or cooling (blue) days. **(c)** Density distributions of 24-hour temperature changes (magnitude), with a 30-minute observation tolerance, for each year. Vertical dashed lines indicate a change of zero (no net warming or cooling)..... 151

Figure S2.9.9: Temporal patterns of average FPKM of BGCs from *Microcystis*. Colored lines represent individual BGCs, smoothed using generalized additive models (GAMs) with a smooth on day of year with a cubic regression function. 152

Figure S2.9.10: Residual diagnostic plots are based on models using log-transformed biosynthesis module abundance (modules 4, 5, 6, and 7) as the response variable within the equation at the bottom of the figure. Residuals appear randomly scattered around zero (homoscedastic) across predictors (top right panel), and fitted values closely follow observed values (bottom right panel), supporting model adequacy for the modules. 153

Figure 2.9.11: Log FPKM of BGCs from modules 8 and 9 belonging to the ADA clade. Individual observations are shown as points colored by year, and smoothed lines are fitted using a loess function. Shaded ribbons indicate standard errors around the fitted temporal patterns. 154

Figure 2.9.12: (a) Co-occurrence matrix of BGCs across publicly available genomes of *Microcystis* deposited in NCBI. Co-occurrence of BGCs across genomes was calculated and hierarchically clustered using average-linkage. The cluster each BGC belongs to is indicated as the colored bar on the right. Tiles are brighter in the heatmap indicating more co-occurrences between BGCs. **(b)** Presence/absence matrix of BGCs in WLECC isolates, organized by which cluster the BGC is in (colored bar). Green tile indicates the presence of the BGC in the isolate, and a white tile indicates the absence of the BGC in the isolate. 155

Figure S2.9.13: Phylogenomic analysis of publicly available *Microcystis* genomes with *Synechocystis* sp. PCC 6803 used as the outgroup. Genomes were used to build a maximum likelihood phylogenomic tree based on the prepackaged single-copy gene-set for cyanobacteria (251 target genes) in GToTree v1.7.05 (4). Lime green boxes indicate the detection of the anabaenopeptin BGC within a genome, and blue stars indicate the detection of the microcystin BGC within a genome. Clades shaded as blue indicate phylogenomically related genomes mostly containing the microcystin BGC, clades shaded as green indicate phylogenomically related genomes mostly containing the anabaenopeptin BGC, clades in purple indicate phylogenomically related genomes that have either the anabaenopeptin or microcystin BGCs, and clades in white contain phylogenomically related without either of the BGCs. 156

Figure S2.9.14: (a) Log-transformed normalized relative abundance (AUC) of metabolites putatively annotated as Aerucyclamide A, Aeruginosin 98A, Anabaenopeptin A and B, and Microcystin LR, YR, HilR, RR, and HtyR. All individual observations are shown, and smoothed lines for each metabolite were fitted using a loess function. **(b)** Representative metabolites from each metabolite class above. 157

Figure S2.9.15: The top row displays partial effects of day of year on log-transformed metabolite abundance from GAMs using a cyclic cubic spline (±95% confidence interval), capturing seasonal patterns for each metabolite (Microcystins, Anabaenopeptins, Aeruginosins, and Aerucyclamides from left to right). The bottom five rows of panels display the relationship between each metabolites' GAM residuals and an abiotic variable. Residuals represent variability unexplained by seasonal trends, and linear models (±95% confidence intervals) are shown with R^2 and *P*-values from the fitted linear regression. Each row represents an abiotic variable (Nitrate + Nitrite, Ammonia, Terrestrial/Humic fraction of DOM, Microbial: Terrestrial Ratio of DOM, and Aromatic Amino Acid Portion of DOM from top to bottom)... 158

Figure 3.9.1: True color images of Winam Gulf, Lake Victoria, Kenya. The images were captured by Sentinel-2 on June 30th, 2022 (top panel) and May 26th, 2023 (bottom panel) and obtained as Level-2A (Surface Reflectance - atmospherically corrected by ESA using Sen2Cor) from Google Earth Engine. The images were collated together using QGIS LTR version 3.28.14. Sites sampled in 2022 are shown in the top panel, while sites sampled in 2023 are shown in the bottom panel. The location descriptor of a site is indicated by a symbol (star, circle, square, or triangle) with the site number indicated next to it on the map. Corresponding information for each site is reported in Table 1. 200

Figure 3.9.2: Turbidity (Nephelometric Turbidity Unit (NTU)), Dissolved Oxygen (mg/L), and Total Nitrogen: Total Phosphorus (TN:TP) (molar) ratio measurements. Individual data points for each variable are plotted on top of each box plot. The red line in the TN:TP panel indicates the threshold for nitrogen-limitation, with all values below it indicating nitrogen-limiting scenarios according to Guildford and Hecky (2000) (34). TN:TP was only measured for sites in 2023. 202

Figure 3.9.3: Microbial and archaeal composition makeup in Winam Gulf, LV. Relative abundance (percentage) of bacterial and archaeal phyla (gray panel) and cyanobacterial genera (green panel) from sampled sites in 2022 and 2023 is shown. The phyla and genera representing less than 1 percent of the population were grouped in the "Other" category. Cyanobacteria with no taxonomic assignment for genus were labeled as "unclassified cyanobacteria". Unmapped reads were excluded from this analysis and mapped reads were normalized to 100 percent for phyla and genera. Station location indicators are shown using symbols after the name of the site number and name on the y-axis. 203

Figure 3.9.4: Cyanobacterial composition of Winam Gulf in relation to physicochemical parameters. **(A)** Non-metric multidimensional scaling (NMDS) analysis of the relative abundance of all cyanobacterial genera from each site, in relation to all other microbial community members (stress = 0.074). Labeled arrows show the direction and strength of the relationship between the ordination axes and associated environmental parameters and the relative abundance of the top 6 most abundant cyanobacterial genera across the sites. Shapes of ordination points indicate station location. Samples from 2022 are colored in amber and samples from 2023 in teal. **(B)** Spearman correlation coefficients between the 6 most abundant cyanobacterial genera across the sample sites and environmental parameters. Rows and columns are hierarchically clustered using correlation. Gray squares indicate self-correlation. 204

Figure 3.9.5: Phylogenomic analysis of dominant bloom-forming cyanobacteria from sampled sites with taxonomic annotations from GTDB (85). **(A)** *Microcystis* phylogenomic tree built with 5 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup. Colored ranges indicate clades as outlined in Pérez-Carrascal et al. 2019 (36). **(B)** *Dolichospermum* phylogenomic tree built with 5 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup The presence of a red circle is indicative of the presence of the saxitoxin gene cluster within the genome (38). **(C)** *Planktothrix* phylogenomic tree built with 3 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup. Benthic or planktic designations for reference genomes are indicated as colored ranges (54). **(D)** *Sphaerospermopsis* phylogenomic tree built with 4

representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup. 209

Figure 3.9.6: Relative abundance of *nifH, nifD,* and *nifK* genes across sites. Reads were aligned to *nifH, nifD,* and *nifK* genes from NFixDB and normalized using the total fragments per kilobase per million mapped reads (FPKM) metric (Equation 3.6.1) (96). The y-axis shows site names with their location descriptor symbol. The x-axis shows the FPKM of each *nif* gene in each sample. Bars are colored based on the genus the *nif* gene was extracted from. 210

Figure 3.9.7: Known cyanopeptide BGC relative abundance across sites. Station numbers, names and their location indicator symbols are shown on the y-axis, with BGCs encoding the biosynthesis of cyanopeptides from MIBiG (accessed December 2023) shown on the x-axis (41). The FPKM (total fragments per kilobase per million mapped reads) of each whole BGC is shown as bubble size (Equation 3.6.2). Bubbles are colored based on the cyanobacterial genus the reference BGC originates from. 211

Figure 3.9.8: BGC richness and diversity in dominant cyanobacterial genera from the Gulf. The top panel shows the average number of BGCs per MAG from 14 different cyanobacterial genera. The bottom panel shows the number of BGCs (bubble size) from each biosynthetic class (y-axis) identified from each cyanobacterial genus (x-axis) by antiSMASH v7.0 on the x-axis (102) 212

Figure 3.9.9: Total BGC relative abundance across the Winam Gulf. Station numbers, names, and their location indicator symbols are shown on the y-axis, with BGC groups shown on the x-axis. BGCs are grouped into broader biosynthetic or product classifications, labeled as the top facet strips (Table S3). The FPKM of each BGC was calculated (Equation 3.6.2) and the size of a point corresponds to the total FPKM of all BGCs of a particular BGC type indicated on the x-axis. Bubbles are colored to differentiate BGC groups. Samples on the y-axis are hierarchically clustered using the complete-linkage method. 213

Figure S3.10.1: Read mapping (via blastn) from Site 27. Bridge Island to four genes from the cylindrospermopsin BGC from *Cylindrospermopsis raciborskii AWT205* (MIBiG BGC: BGC0000978). Percent identity of reads mapped to these genes is shown in the top quadrants, and percent coverage by mapped reads to these genes is shown in the bottom quadrant..... 215

Figure S3.10.2: Gene cluster families (GCFs) for identified BGCs via antiSMASH and BiG-SCAPE. BGCs identified from high quality cyanobacterial MAGs via antiSMASH v7 were manually annotated (SI Table 2) and analyzed via BiG-SCAPE. Nodes are colored based on the cyanobacterial genus of the MAG they originate from (based on GTDB-Tk taxonomic identification). Nodes are labeled based on manual annotation of BGCs. Square nodes with a red outline are from the MiBIG database, linking BGCs to known biosynthesis products. Edges between nodes indicate similarity of BGCs calculated through a distance matrix that considers number of shared PFAM domains, pairs of adjacent PFAM domains, and sequence similarities between protein sequences. Gray circles around nodes indicate the putative synthesis product of those nodes based on antiSMASH annotation and manual inspection of core and additional biosynthesis genes. 216

Figure S3.10.3: Genome profile of MAG samp_4317_maxbin_26, annotated as Microcystis sp. (NCBI: SAMN41711332). The contig with the mcy operon is indicated in the Anvi'o MAG profile. Gene similarity image from clinker shows the reference mcy operon from Microcystis aeruginosa FACHB-1326 (GenBank: OQ291092.1) and its close similarity to the mcy operon identified in samp_4317_maxbin_26 on contig 177 below. 217

Figure S3.10.4: Biosynthetic potential composition in relation to dominant cyanobacterial genera relative abundances in the Winam Gulf. NMDS analysis of the relative abundance of all identified BGCs from each sample site. Arrows and labels represent the direction and strength of the relationship between the ordination axes and the relative abundance of the top 6 most abundant cyanobacterial genera across the samples. Station location is indicated by the shape of ordination points, and samples from 2022 are represented in amber, while samples from 2023 are in teal. Samples taken from rivers (n=4) were excluded from this ordination because no BGCs were identified in them. 218

Figure S3.10.5: Microbial and archaeal composition makeup in Winam Gulf, LV. (**A**) Relative abundance (percentage) of bacterial and archaeal phyla from 2022 samples with replicates is shown. The phyla, annotated with GTDB-Tk release 214, representing less than 1 percent of the population were grouped in the "Other" category (87). (**B**) Relative abundance (percentage) of cyanobacterial genera from 2022 samples with replicates is shown. Cyanobacteria with no genus taxonomic assignment from GTDB-Tk were labeled as "unclassified cyanobacteria". Unmapped reads were excluded from this analysis and mapped reads were normalized to 100 percent. Metagenomic samples from the same sampling site are grouped together in a facet. .. 219

Figure 4.8.1: Chemical structures of microcystins (LY, LA, LF, LR, Asp^3-LR, RR, YR, and HilR) and anabaenopeptins A and B. Microcystins are grouped based on amino acid substitutions at the variable X and Z positions: leucine (Leu) residues at position X are highlighted in orange, and arginine (Arg) residues at position Z are highlighted in blue. Structural differences between anabaenopeptins A and B are shown in magenta. 286

Figure 4.8.2: Workflow for preparing cyanotoxin standards and culture-derived metabolite extracts and assessing their effects on cell viability in human cell lines using the CellTiter-Glo assay. Pure standards (MC-LR, MC-LA, MC-RR, AP-A, AP-B) and extracts from seven WLECC *Microcystis* isolates were prepared in biological triplicate. After metabolite extraction and cyanotoxin quantification, human cell lines (A549, HK-2, and Hep-3B) were plated and incubated for 24 hours, followed by toxin treatment and a 48-hour incubation. Cell viability was measured via luminescence. 287

Figure 4.8.3: Linear dose-dependent cell viability response by human **(a)** lung (A549), **(b)** kidney (HK2), and **(c)** liver (Hep-3B) cells to pure MC and AP congeners. Viability is shown relative to DMSO (negative control) and compared to doxorubicin (positive control, data shown in black). 288

Figure 4.8.4: Dose-dependent cell viability response by human **(a)** lung (A549), **(b)** kidney (HK2), and **(c)** liver (Hep-3B) cells to combinations of MC-LA and AP-B. Viability is shown relative to DMSO (negative control), with lighter tiles indicating more cell viability and darker tiles indicating lower cell viability. 2D contour plots show Loewe synergy scores for MC-LA

and AP-B combinations in **(d)** A549, **(e)** HK2, and **(f)** Hep-3B cells. Scores <1 indicate synergistic effects (green), ~1 indicate additive effects (white), and >1 indicate antagonistic effects (red). Actual Loewe scores are displayed within each plot. Mean Loewe scores for each cell line are: A549: 0.45, HK2: -0.48, and Hep-3B: -9.33. 289

Figure 4.8.5: Concentrations of select congeners of microcystins and anabaenopeptins in crude extracts from western Lake Erie culture collection *Microcystis* cultures used in human cell line exposures. Biological duplicates or triplicates are shown for each isolate. Bars represent the mean concentration in parts per billion (ppb) for each cyanotoxin, with standard error indicated with error bars. 290

Figure 4.8.6: Dose-dependent cell viability response by human lung (A549), kidney (HK2), and liver (Hep-3B) cells to *Microcystis* isolate extracts from WLECC. Extracts were dosed between 12.5 and 100 µM, based on the crude dry weight of extracted metabolites. Each strain was grown in biological duplicate or triplicate, and cell viability was assessed in technical quadruples across two plates. Viability is shown relative to DMSO (negative vehicle control). A red square indicates the detection of microcystins in the extract, and a blue circle indicates the detection of anabaenopeptins in the extract. A unfilled blue circle indicates the presence of the anabaenopeptin biosynthetic gene cluster (BGCs) but no anabaenopeptin A or B detection. Significant effects were determined by Tukey's multiple comparisons post-hoc tests against the vehicle control (Significance: $P < 0.01$ (*), $P < 0.001$ (**), $P < 0.0001$ (***). 291

Figure 4.8.7: Pearson correlation coefficients (R) between cell viability effects exhibited by WLECC *Microcystis* isolates at four doses between 12.5 and 100 µM and cyanotoxin concentrations (MC and AP congeners) in each extract. Each tile shows the correlation coefficient (R), colored by strength and direction (blue (negative) to red (positive)). "Sum APs" indicate the sum of all AP congeners measured in the extract, and "Sum MCs" indicate the sum of all MC congeners measured in the extract. 292

Figure 4.8.8: Comparative toxicity and transportability of microcystin and anabaenopeptin congeners. Percent cell viability at 100 µM for three microcystin and two anabaenopeptin congeners in human lung (A549), liver (HK2), and kidney (Hep-3B) cells from the current study (orange arrow). OATP1B3 selectivity, a key transporter for MC-LR, from Niedermeyer et al. (2014) is indicated with a green arrow for eight microcystin congeners. Intraperitoneal LD_{50} values (µg/kg) in murine models for six microcystin congeners are shown as summarized in Chorus and Welker (2021) (red arrow). Congeners used in the current study are bolded. Congeners on the left tend to be less toxic across metrics, while those on the right tend to be more toxic. Protein phosphatase 1 and 2 inhibition across these congeners do not vary significantly (Fischer et al. (2010) and Blom and Jüttner (2005)). 293

Figure S4.9.1: Mean viability and ±95% confidence interval of (a) lung (A549), (b) kidney (HK2), and (c) liver (Hep-3B) cell lines following exposure to four doses (100, 50, 25, 12.5 µM) crude extracts of *Microcystis* isolates from the WLECC relative to the DMSO vehicle control (shown in black). The dotted line represents the mean viability of DMSO vehicle controls. Isolate effects below the dotted line indicate significant cytotoxicity relative to control. 331

Figure S4.9.2: Principal Coordinates Analysis (PCoA) of viability responses to WLECC *Microcystis* isolate extracts in human cell lines lung (A549), liver (Hep-3B), and kidney (HK2), based on Euclidean distances of mean viability. Each point represents one isolate–cell line combination, colored by isolate identity. Shapes and sizes of points vary across panels where in **(a)** shapes indicate the cell line tested, **(b)** point size is scaled to total microcystin content (ppb) in the extract, and **(c)** point size is scaled to total anabaenopeptin content (ppb) in the extract. PCoA axes represent the proportion of total variance explained (PCoA 1: 66.8%, PCoA 2: 28.3%). 332

Appendix Figure A.6.1: Map of western Lake Erie showing NOAA long-term monitoring stations evaluated in this study. Near-shore sites WE2 and WE12 border the Ohio coastline, WE9 is located between the Maumee and Detroit River mouths, and WE4 represents an offshore station..... 358

Appendix Figure A.6.2: Quantification of cyanotoxins via liquid chromatography-mass spectrometry (LC-MS), including anabaenopeptins (ABN-A and ABN-B) and microcystin (MC) congeners (Asp^3-LR, HilR, LA, LR, LY, RR, YR). Each vertical panel represents a different cyanotoxin, each with a unique y-axis scale. Biological replicates (samples collected from the same station and date) are shown along the x-axis as replicates and are grouped by collection date and station. All extracts were collected from the surface using a 100 µM plankton net tow, while samples with a * (sample 11, collected from WE4 on October 19th, 2023) were collected from a depth integrated sample and were not filtered using any mesh. ... 359

Appendix Figure A.6.3: (a) Cell viability response by HK-2 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 12.5 and 100 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control. **(b)** Cell viability response by HK-2 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 1.25 and 10 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control. 360

Appendix Figure A.6.4: (a) Cell viability response by A549 (Lung) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 12.5 and 100 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control. **(b)** Cell viability response by HK2 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 1.25 and 10 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control.. 361

Appendix Figure A.6.5: (a) Cell viability response by Hep-G2 (Liver) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 3.9 and 500 µM (shown on the x-axis), relative to the crude weight of dried extract after metabolite extraction. Cell viability response, shown as the percent viability relative to the DMSO control, is shown on the y-axis.

Biological replicates share the same color but differ in shape. **(b)** Cell viability response by Hek-293 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 3.9 and 500 μM (shown on the x-axis), relative to the crude weight of dried extract after metabolite extraction. Cell viability response, shown as the percent viability relative to the DMSO control, is shown on the y-axis. Biological replicates share the same color but differ in shape. 362

Abstract

Toxic cyanobacterial harmful algal blooms (cyanoHABs) are proliferating in freshwater ecosystems due to excess nutrient inputs, causing eutrophication and threatening human, wildlife, and ecosystem health. These blooms are dominated by cyanobacteria with diverse repertoires of biosynthetic gene clusters (BGCs) that encode for the synthesis of secondary metabolites, bioactive compounds that confer functional advantages beyond growth or reproduction such as defense or competition. Among these are cyanopeptides, a broad class of peptidic metabolites with diverse bioactivities, including disruption of enzymatic and cellular functions. Microcystins, a hepatotoxic subset of cyanopeptides, are classified as cyanotoxins due to their potent toxicity to humans and wildlife. While monitoring often focuses on cyanotoxins, other cyanopeptides, such as anabaenopeptins, may also contribute to bloom toxicity. The cosmopolitan bloom-forming genera *Microcystis* and *Dolichospermum* frequently co-occur in freshwater bodies and dominate different bloom phases, each encoding numerous BGCs for both known and uncharacterized metabolites. Intraspecific variability in secondary metabolite production shapes bloom toxicity and drives ecological interactions across trophic levels during cyanoHABs. However, the temporal dynamics, environmental drivers, and health risks of individual and co-occurring cyanopeptides beyond microcystins remain poorly defined. Emerging evidence suggests these compounds may act synergistically, amplifying their toxic effects towards humans, wildlife, and aquatic organisms.

To address these gaps, this thesis integrates multi-omics (metagenomics and metabolomics) and toxicological approaches to investigate the dynamics, drivers, and understudied risks of natural cyanoHAB communities and their bioactive metabolites from two geographically distant but similarly eutrophic Great Lakes basins, western Lake Erie (WLE) (USA) and the Winam Gulf of Lake Victoria (Kenya).

Chapter 2 investigated the temporal patterns of *Microcystis* and *Dolichospermum* dominance, BGC abundance, and cyanopeptide production in WLE. Lower temperatures and nitrogen availability were most closely associated with late season *Dolichospermum* dominance. Beyond taxonomic shifts, co-varying modules of BGCs and their associated cyanopeptides exhibited distinct temporal patterns, revealing biosynthetic successions across the bloom season that extended beyond microcystins. The cyanopeptide phases were linked to changes in the environment, including eDNA detection of grazers, dissolved organic matter composition, and nitrogen availability, suggesting both top-down and bottom-up controls on cyanoHAB ecology and toxicity.

Chapter 3 characterized the cyanobacterial community and its BGC repertoire in the Winam Gulf of Lake Victoria using metagenomic sequencing. Extensive spatial variation in cyanobacterial composition was observed, with *Dolichospermum* as the dominant genus during most large bloom events, while *Microcystis* dominated non-visible blooms in turbid regions at the time of sampling. Microcystin production was attributed specifically to *Microcystis*, resolving a long-standing ambiguity in the region. Additionally, over 300 novel cyanobacterial BGCs with unknown products were identified, revealing a vast and previously unexplored chemical landscape.

Chapter 4 evaluated the toxicological effects of two commonly co-occurring cyanopeptide classes, anabaenopeptins and microcystins, on human lung, kidney, and liver cell lines. Varying synergistic interactions between the cyanopeptides were observed across all three cell lines, highlighting the importance of investigating mechanisms of toxicity for co-occurring metabolites. Toxicity assays using *Microcystis* culture extracts demonstrated that microcystin and anabaenopeptin concentrations alone poorly explained the variation in cytotoxicity across cell lines. This suggests that additional compounds or interactive effects between metabolites may play a more substantial role in overall toxicity.

Overall, these findings underscore the need to expand research, monitoring, and management efforts beyond microcystins to better capture the complexity of cyanoHAB toxicity and its risks to public and ecological health.

Chapter 1 Introduction

1.1 Introduction and Overview

Freshwater ecosystems are essential to global environmental and human well-being, yet they are increasingly threatened by human activity. They play a vital role in regulating Earth's climate, supporting biodiversity, and providing essential societal resources (1). However, freshwater ecosystems face increasing threats from eutrophication, driven by excessive nitrogen and phosphorus inputs from industrial, urban, and agricultural sources (2). Rapid urbanization, invasive species, and pollution are further accelerating declines in water quality and biodiversity within these systems (3). Worldwide, these anthropogenic pressures are leading to the widespread proliferation of cyanobacterial harmful algal blooms (cyanoHABs) (**Figure 1.1**) (4,5). Nuisance cyanobacteria are responsible for these blooms, relying on photosynthesis and nutrient enrichment to proliferate and produce toxins that pose serious risks to human health, aquatic ecosystems, and wildlife (6,7). CyanoHABs and their toxins are detrimental to the prosperity of ecosystems and societies, costing the United States approximately 2.2 billion dollars annually (8). The issues they pose are multi-faceted, resulting in degraded drinking water and hypoxic ecosystems, thus reducing biodiversity, health, tourism, and recreation in effected areas (8). Alarmingly, cyanoHAB proliferation and toxicity are expected to intensify with rising temperatures, carbon emissions, and nutrient inputs. Temperate freshwater systems are projected to experience more than twice as many cyanoHAB-affected days by 2050 (5,9). Together, these trends emphasize the need to understand the dynamics and drivers of bloom-forming cyanobacteria to prevent, mitigate, and manage their global consequences.

CyanoHABs consist of diverse bloom-forming cyanobacteria with distinct ecophysiological adaptations (10–12). *Microcystis*, the most widespread bloom-forming genus, is found on nearly every continent and frequently produces the hepatotoxin microcystin (MC) (13). Unlike *Microcystis,* which is unicellular, colony forming, and non-diazotrophic, other globally dominant genera are filamentous and can often fix atmospheric nitrogen (diazotrophic), including *Dolichospermum* (previously *Anabaena), Anabaena, Aphanizomenon*, *Planktothrix, Raphidopsis, Nodularia,* and *Phormidium* (14). Diazotrophic and non-diazotrophic genera often bloom in succession of one another (12). These genera-specific temporal patterns are mediated by strain-specific morphological and physiological adaptations which control buoyancy, provide resistance to grazing, determine nutrient uptake efficiency, and regulate growth and temperature response (12,15–19). Additionally, each bloom-forming genus, species, and strain poses distinct threats to human and environmental health through unique toxin production abilities (20–22). However, resolving phylogenetic relationships at the species and strain levels in common bloom-forming genera such as *Microcystis* and *Dolichospermum* remains challenging due to the limited availability of complete genomes, highly plastic and open pangenomes, extensive repetitive regions, phenotypic diversity, and low genomic synteny (10,13,23). While genus-specific ecophysiological traits likely drive the dominance of particular cyanobacteria under specific environmental conditions, our understanding of how these traits respond to, and shape, bloom dynamics remains limited, constraining our ability to predict and manage blooms effectively.

Beyond genus-level variation, cyanobacterial strains exhibit substantial heterogeneity in key functional traits including carbon concentration, nutrient metabolism, and host-microbiome interactions, with secondary metabolism standing out as the most diverse trait of them all (20–22,24–27). Cyanobacterial secondary metabolism, the act of producing molecules for purposes

other than growth and reproduction, is highly diverse, documented to synthesize 2,425 unique molecules (as of September 2025), many of which having negative effects on phytoplankton, microbes, plants, invertebrates, and vertebrates, including humans (7). The ability to synthesize secondary metabolites is encoded in the genome through biosynthetic gene clusters (BGCs) which are made up of tightly linked sets of genes that encode the enzymes and protein necessary to synthesize a secondary metabolite (21,22,28). Freshwater cyanoHAB research on secondary metabolism largely focuses on the synthesis of MCs, accounting for over 90% of current literature but only 13% of known cyanobacterial metabolites (7,29). However, bloom-forming genera produce a vast array of other structurally complex hepatotoxins, neurotoxins, taste and odor compounds, and molecules which remain to be characterized and linked to their corresponding BGC (**Figure 1.2)**. Secondary metabolites produced within cyanoHABs are temporally variant and are hypothesized to serve diverse ecological functions including nutrient storage, grazer defense, allelopathy, photosynthetic support, phycosphere recruitment, and oxidative stress protection (30–34). Few studies investigate the temporal distributions and putative drivers of cyanobacterial secondary metabolism beyond MCs, leaving a large gap in understanding the functional roles, harmful effects, and potential mitigation strategies of these toxic molecules.

One reason for this gap is that robust phenological modeling of cyanobacterial dominance and secondary metabolism requires extensive multi-omics data, including data analyzing multiple types of macromolecules such as DNA, RNA, proteins, metabolites, to reveal meaningful and conserved temporal patterns (35–37). Producing these long-term, multi-omics (metagenomics, metatranscriptomics, proteomics, metabolomics) datasets is costly and requires close-knit collaboration. Eutrophic Great Lakes basins such as western Lake Erie (U.S.) and the

Winam Gulf of Lake Victoria (Kenya) are ideal systems for generating such valuable datasets, as they are well-monitored and experience frequent, severe cyanoHABs with significant local impacts (38–41). These systems provide an opportunity to comprehensively study the spatiotemporal dynamics of cyanobacterial secondary metabolism across a vast latitudinal gradient.

1.2 CyanoHABs in Eutrophic Great Lakes

1.2.1 CyanoHABs in Western Lake Erie, United States: Past and Present

CyanoHABs have been detected and cause detrimental local effects in all Laurentian Great Lakes and their surrounding tributaries (42). Of these water bodies, Lake Erie, which is divided into eastern, central, and western basins and provides drinking water for ~11 million people, experiences some of the most prolific cyanoHAB events (43). These events are concentrated in Sandusky Bay and western Lake Erie (WLE) **(Figure 1.3)**, the shallowest, warmest, and most biologically productive area in the Great Lakes (44,45).

CyanoHABs in WLE have been documented for more than a century, but the phytoplankton composition has been constantly shifting in response to changes in nutrient accumulation driven by lakeside population, land use, and lake management changes (43,44). In the mid-1900's, rapid increases in nutrient concentrations, both nitrogen and phosphorus, were reported and hypothesized to originate mainly from fertilizers (46). As WLE became increasingly eutrophic in the 1950's, the dominant phytoplankton switched from diatoms to filamentous, nitrogen-fixing (diazotrophic) cyanobacteria from the genera *Anabaena, Dolichospermum,* and *Aphanizomenon*, now considered the ADA clade, accompanied by a relatively small population of *Microcystis* (43,47). Policymakers and researchers recognized the declining trophic state of WLE in the early 1970's and passed the binational Great Lakes Water

Quality Agreement (GLWQA) in 1972 to reduce phosphorus delivery to the western and central basins of Lake Erie by 40% (48). The GLWQA reduced point-source phosphorus pollution in WLE, improving water clarity and lowering algal biomass through the 1980s, an process that was also aided by the introduction of invasive Dreissenid mussels (49–51).

Although total phosphorus inputs from point sources have substantially declined in western Lake Erie (WLE), increasing nutrient inputs from non-point sources, particularly since the mid-1990s, have reshaped cyanoHAB phytoplankton composition. Specifically, rising levels of bioavailable phosphorus and chemically reduced nitrogen forms (e.g., urea, ammonia, and dissolved organic nitrogen [DON]) from non-point source, agricultural runoff have fueled the proliferation of non-diazotrophic bloom-forming genera like *Microcystis* during summer months (42,52–55). *Microcystis* rapidly respond to externally loaded, chemically reduced nitrogen, likely due to its energetic favorability and *Microcystis'* superior ability to scavenge these resources compared to its eukaryotic competitors (56–58). Moreover, external loading and internal recycling of bioavailable nitrogen increases the production of the nitrogen-rich hepatotoxin, MC (59). Researchers collectively hypothesize that the rising levels of DON and bioavailable phosphorus are driving the increased intensity and toxicity of *Microcystis*-dominated cyanoHABs in WLE and surrounding freshwater ecosystems (60). In the last five years, the diazotrophic genus *Dolichospermum* has been blooming in mass in the colder fall months (Chapter 2, (61)). These shifts over the past two decades likely reflect the combined influence of legacy phosphorus accumulation in the Lake Erie watershed, climate variability, evolving agricultural practices (e.g., no-till farming), and altered nutrient dynamics driven by invasive species (62–65). Together with rising inputs of bioavailable nitrogen and phosphorus, these factors have contributed to increasingly variable and intense phytoplankton blooms in WLE.

This evolving nutrient landscape underscores the need for a dual-nutrient management approach that addresses both nitrogen and phosphorus inputs to mitigate cyanoHAB impacts.

Starting in the mid-1990's, monitoring of WLE consisted mostly of opportunistic sampling of a limited set of physicochemical parameters by National Oceanic and Atmospheric Administration Great Lakes Environmental Research Laboratory (NOAA GLERL) and United States Geological Survey (USGS), and cyanotoxin (MC, cylindrospermopsins (CYL), anatoxin-a (ATX), and saxitoxin (SXT)) monitoring by the Environmental Protection Agency (EPA) (**Figure 1.4**) (41). More intense monitoring, collaboration with universities, and the use of molecular tools to monitor the genomes and metabolites of bloom-forming cyanobacteria in WLE began in 2014 due to the Toledo drinking water crisis, in which MCs made it into the Toledo's processed water and resulted in a three day "do not drink" advisory (66). This event cost the city an estimated 65 million dollars, damaged residents' trust with their water source, and underscored the importance of monitoring and managing WLE, which is now declared an impaired water source under the Clean Water Act (67). Today, monitoring of WLE consists of satellite monitoring, early warning forecasting, measurement of over 50 physicochemical and biological water quality measurements, buoy deployment, cyanotoxin monitoring (MC, CYL, ATX, SXT), autonomous vehicle sampling, sample collection for DNA, RNA, and metabolomics analysis, and *in situ* experimental studies (41,68). With substantial investments in monitoring infrastructure and long-term data collection, WLE serves as an ideal system for studying bloom dynamics and drivers, supporting early warning modeling and informed management planning.

1.2.2 CyanoHABs in Winam Gulf, Lake Victoria, Kenya: Past and Present

The African Great Lakes are among the most ecologically and socioeconomically vital freshwater systems on the planet yet remain critically understudied. Together, Lakes Victoria,

Tanganyika, Malawi, Turkana, Albert, Kivu, and Edward hold nearly a quarter of Earth's unfrozen freshwater, supporting millions of people through fishing, agriculture, and drinking water (69). However, these lakes are increasingly threatened by cyanoHABs, which jeopardize biodiversity, ecosystem productivity, and the resources communities depend on for daily life (70). For example, Lake Victoria, the world's second-largest lake by surface area, supports over half of Kenya's population and is experiencing widespread eutrophication due to rapid population growth and land-use changes, resulting in widespread cyanoHABs concentrated in its shallow, nutrient-rich gulfs (71).

Some of the most prolific and dangerous cyanoHAB events in Lake Victoria take place in the Winam Gulf (Nyanza Gulf), a shallow inlet in the southwestern corner of Kenya (**Figure 1.5**) (39,72). The gulf provides water for drinking, bathing, daily tasks, subsistence fishing, and aquaculture to approximately 4.5 million people from Kenya's third largest city, Kisumu, and smaller nearby fishing villages such as Homa and Kendu Bays (73,74). Substantial population growth in these lakeside communities and increasing sewage and urban runoff, mainly via the Nyando, Sondu-Miriu, Awach-Tende, Awach-Kibuon, Kibos, and Kisat rivers, have led to year-around cyanoHABs in the gulf. Despite the abundant sources of both nitrogen and phosphorus to the gulf, total nitrogen: total phosphorus (TN:TP) ratios indicate nitrogen limitation relative to phosphorus (75). These conditions should select for diazotroph dominance in the gulf, although there are plenty of non-diazotrophic cyanobacteria (i.e., *Microcystis*) present at high densities, suggesting non-diazotrophs may be employing alternative adaptation strategies to proliferate in the nitrogen limited gulf (76,77).

Monitoring in this region is mostly conducted by the Kenya Marine and Fisheries Research Institute (KMFRI) and local universities (39,78,79). Quarterly cruises assess key water

quality indicators, classify phytoplankton morphology through microscopy, and survey fish populations. Prior to 2022, studies on cyanoHABs in this region identified *Microcystis* as the dominant bloom-forming cyanobacterium via microscopy (39). A collaborative grant between 2022 and 2024 funded through the National Science Foundation International Research Experience for Students (NSF-IRES) initiated the first 16s ribosomal RNA and metagenomic sequencing projects studying the molecular composition of Winam Gulf cyanoHABs (link). The first study used 16s rRNA sequencing and quantitative polymerase chain reaction (qPCR) approaches to study the microbial community and toxin biosynthesis genes along a spatial gradient in the gulf in 2022 (76). The results identified *Dolichospermum* as the main bloom-former across cyanoHAB events in the gulf, inconsistent with previous reports of *Microcystis* dominance. qPCR analysis revealed alarmingly high levels of toxin biosynthesis genes for MC and CYL, *mcyE* and *cyrA*, near Homa Bay and the Nyando river mouth, respectively. Prior to this study, no genomic investigations had characterized the taxonomic composition or biosynthetic potential of cyanoHABs in the Winam Gulf, Lake Victoria's most eutrophic and bloom-prone region. This gap highlights the need for increased molecular-based monitoring in the Winam Gulf to better understand the dynamics and threats of cyanoHABs locally, while also addressing the global data scarcity for this ecologically and economically critical freshwater system.

1.2.3 Implications for Environmental, Human, and Economic Prosperity

Human-induced eutrophication is deteriorating water quality worldwide primarily due to the rapid expansion of toxic cyanoHABs, which have severe consequences for environmental health and stability (80). The environmental consequences of cyanoHABs are far-reaching, affecting water quality, biodiversity, and ecosystem stability. Bloom-forming cyanobacteria

outcompete beneficial phytoplankton, altering the energy flux of the ecosystem and reducing resources available to higher trophic levels (81). The die off and subsequent decomposition of cyanoHABs results in hypoxic dead zones. Hypoxia can cause major fish kills and also selects for aquatic organisms adapted to hypoxia, leading to a loss of biodiversity and ecosystem sustainability (81). Toxins produced by bloom-forming cyanobacteria greatly affect aquatic organisms, livestock, domestic animals, and wildlife. Cyanotoxins accumulate in fish and bivalves, causing death and biomagnifying toxins throughout the food web (82). Animal sickness and death due to cyanotoxin exposure have been reported in over 50 countries, but this number is known to be significantly underreported (83,84). Cyanotoxin exposures have caused massive die-offs of wading wildlife such as birds and non-wading animals in African countries such as zebra, impalas, giraffes, and elephants (85). Livestock are especially vulnerable to cyanotoxin exposure, often developing skin irritation, organ damage, liver injury, and even death due to drinking from and swimming in visibly contaminated, foul-smelling, cyanoHAB-laden ponds (86). Domestic animals such as cats and dogs are particularly affected by cyanobacterial neurotoxins, ATX and SXT, and die every year due to cyanoHABs around the world (85,86). CyanoHABs negatively affect every part of an ecosystem, reducing its overall quality and safety.

Beyond environmental and megafauna disruptions, cyanoHABs pose serious risks to human health through a variety of exposure routes. 80% of exposures are caused by contaminated drinking water, with over 2,000 human poisonings annually (87,88). The effects of human cyanotoxin exposure greatly depends on the route of exposure and type of toxin (80). The main routes of exposure are dermal, oral, inhalation, parenteral, and exposure medias such as food, water, dietary supplements, aerosols, and dust (80). Acute exposure to cyanotoxins causes symptoms such as conjunctivitis, vomiting, blistered mouth, rashes, abdominal pain, fever, skin

irritation, hepatic failure, and in extreme cases, death (89,90). One instance of acute cyanotoxin poisoning took place at a hemodialysis center in Brazil where the facilities' reverse osmosis center failed and resulted in the delivery of MCs and CYLs directly to 126 patients. This led to acute liver failure in 100 patients, with 60 of the patients succumbing to the exposure (90).

Long-term, chronic exposure to cyanotoxins leads to liver and kidney damage and cancer (4). Cyanotoxin exposures have even more detrimental effects towards populations with pre-existing conditions such as non-alcoholic fatty liver disease, liver failure, dermal ailments such as eczema and psoriasis, various cancers, immunomodulating diseases, and asthma (91–93). Inhalation of cyanotoxins leads to approximately 15% of asthma cases in the coastal regions of the U.S., and contributes to the symptoms of bronchitis, allergies, rhinitis, and dermatitis (93,94). The Asthma and Allergy Foundation of America has declared the Ohio-Lake Erie region one of the most challenging places to live with asthma in the U.S., and researchers believe the aerosolization of cyanoHABs and their toxins may be exacerbating asthmatic conditions in these lakeside communities (95). The complex, chronic, and irreversible effects of cyanotoxin exposure warrant greater public awareness and education to help prevent these consequences.

The consequences of cyanoHABs on the environment, human health, and society are costly to socioeconomic sectors such as fisheries, tourism, public health, real estate, and recreational centers and disrupt social and cultural practices (80). As of 2009, they were estimated to cost the U.S. an estimated $2.2 billion annually, with nearly half of the losses stemming from declines in waterfront property values (96). CyanoHABs also increase water treatment costs, with effects tricking down into consumer water bills. Water supply has been shut down for over 48 hours in major cities such as Toledo, Ohio, and Wuxi, China due to cyanotoxins and taste and odor compounds making their way into treated water supply (97,98).

Emergency response measures to these crises are expensive, costing Toledo over $65 million, and an additional $100 annually in water bills for every family in Toledo thereafter (97). Cyanotoxin poisoning from contaminated water increases medical costs and contributes to lost workdays, harming local and national economies. Additionally, cyanoHABs' toxicity and unsightly appearance reduces tourism, business revenue, fishing, and boating activities (74,80). These economic costs affect the global economy, but they especially affect vulnerable communities relying directly on lakes for their livelihood, emphasizing the importance of managing these environmental events.

The Winam Gulf region is severely threatened by the public health risks posed by cyanoHABs, as nearly half of the population relies on untreated lake water for drinking and daily use, with limited access to consistent water treatment or filtration. (38). Studies have found that over 60% of samples taken from the gulf exceed maximum MC levels for children and immunocompromised individuals recommended by the US EPA (74). This is of particular concern due to the high prevalence of immunomodulating diseases such as HIV and malaria in this area (99,100). Fishing and fish consumption are deeply rooted in the region's culture, making cyanotoxin ingestion, primarily MCs, a significant concern, with reported MC levels in local fish reaching levels up to 89 μg/kg. Alarmingly, data shows that small-scale fishing communities remain unaware of cyanoHABs in the region, and their water use remains unchanged despite awareness, as it is often their only accessible water source (38). Beyond human health effects, cyanoHABs affect the livelihood of these vulnerable communities, resulting in fish die-offs, livestock poisoning, and decreases in tourism, all of which are major contributors to the maintenance and growth of the local economy. The Winam Gulf is just one example of how

cyanoHABs obstruct the safe use of locally and culturally important water bodies, leading to a decline in the environmental, human, and economic health of an emerging region.

1.3 *Microcystis* and *Dolichospermum* Bloom Formation: Ecophysiology and Drivers of Dominance

Microcystis and *Dolichospermum* are cosmopolitan bloom-forming cyanobacteria, frequently co-occurring in nutrient-rich freshwater systems prone to cyanoHABs. Alongside other genera such as *Planktothrix*, *Aphanizomenon*, *Anabaena*, *Raphidiopsis*, *Nodularia*, and *Oscillatoria*, they contribute significantly to bloom dynamics and toxicity in freshwater lakes (101). Their frequent dominance in systems like the Laurentian and African Great Lakes (11,13,40,76,102), coupled with their genetic diversity, morphological variation, and toxin-producing potential, make them ideal models for understanding the ecological and functional complexity of cyanoHABs (10,23).

1.3.1 Ecophysiology of Microcystis

Microcystis spp. employ diverse functional traits to fill changing ecological niches in cyanoHABs worldwide. *Microcystis* is a buoyant, unicellular, coccoid, colony forming cyanobacterium **(Figure 1.6)** that, often in temperate systems, overwinters in the benthos and rises to the epilimnion in the summer to form blooms (103–105). They thrive across diverse physicochemical conditions and have been documented in 108 countries across every continent except Antarctica, with 79 of these countries also reporting the presence of MC (13,106). The study of *Microcystis* spp. often focuses on the presence or absence of the *mcy* operon, which encodes the synthesis of MC, due to its potential human and ecosystem health risks and roles in fitness (10,107,108). Beyond the *mcy* operon, *Microcystis* spp. greatly differ in their morphology, physiology, and ecology due to genome mutations and rearrangements and flexible

gene variation, which is equal to ~50% of the genome, and phenotypic plasticity (10,109). The highly plastic genetic makeup, characterized by high amounts of repeat sequences, low synteny, and phenotypic diversity of *Microcystis* spp. allows it to dominate in a wide-range of ecosystems (109,110). Specifically, *Microcystis* spp. leverage this diversity in its response to light and temperature, buoyancy regulation, nutrient cycling, secondary metabolite synthesis, and colony formation to enhance nutrient uptake, decrease the negative effects of unfavorable physicochemical conditions, mitigate inter- and intra-species competition, and deter predation (7,10,17,107,111–115).

Microcystis generally like it hot and sunny. Blooms dominated by this genera often occur at temperatures above 20°C, but their temperature optimum is strain-specific and ranges between 25°C and 30°C (10). Because of their preference for warmer temperatures, *Microcystis* bloom occurrence and intensity are expected to increase in tandem with global temperatures (5). Light intensity also greatly impacts the growth rate of *Microcystis*, with an optimum around 50 µm photons m^{-2} s^{-1}. The effects of light on *Microcystis* spp. fitness are difficult to disentangle, as they depend on light quality (e.g., wavelength), photoacclimation, colony size, and water depth, with individual cells experiencing light differently due to microenvironmental variation. Water temperature and light intensity dually impact strain-specific physiological responses and fitness in cyanoHABs.

Microcystis uses buoyant gas vesicles and high pH tolerance to outcompete other microorganisms during stressful conditions such as nutrient and light limitation (108,116). Buoyant behavior in *Microcystis* is driven by the balance between carbon fixation and nutrient acquisition, allowing cells to sink for nutrient uptake from the benthos and rise to the surface to form scums and photosynthesize (13,117). With the ability to rise to the surface and form dense

scums, *Microcystis* increases light attenuation and limitation, ultimately limiting resources for competing photosynthetic organisms (118). *Microcystis* spp. are well-adapted to both low and high CO_2 concentrations due to strain-specific carbon concentrating mechanisms and encoding of both low- and high-affinity bicarbonate uptake systems (114,119). As blooms progress, photosynthesis drives pH to levels above 9, creating unfavorable conditions for competitors like diatoms whose siliceous frustules dissolve at such levels (114,120). *Microcystis* spp. CCM diversity and buoyant properties enable strains to thrive in conditions with fluctuating pH and climate-driven acidification (114,121).

As light requirements and carbon fixation strategies are strain-specific, so are nutrient requirements, even displaying intraspecific niche divergence within ecosystems of different trophic statuses (10,122). *Microcystis* excels at scavenging phosphorus in low concentrations, particularly in shallow, turbid waters like WLE and the Winam Gulf, where it can access phosphorus-rich benthos (116,123). It can also endure phosphorus deprivation by storing intracellular polyphosphate bodies and accumulating phosphorus on its surface (116,124). While research and policy focus on the role of phosphorus in driving and sustaining *Microcystis* blooms, *Microcystis* is non-diazotrophic and depends entirely on external and regenerated nitrogen sources (65). Nitrogen is critical not only for global bloom expansion but also for bloom toxicity, as most *Microcystis*-produced hepatotoxins are nitrogen-rich (7,125,126). The form and concentration of nitrogen influence *Microcystis* spp. growth and behavior in a strain-specific manner. Field studies suggest that *mcy*-encoding strains, which require more nitrogen than non-*mcy* strains, preferentially use inorganic nitrogen sources like nitrate and ammonia, whereas non-*mcy* strains favor organic forms such as urea and glutamic acid (127). This preference may explain the shift from *mcy* to non-*mcy* strains observed in many temperate lakes following peak

bloom conditions, when external nitrate is depleted and the community relies on organic nitrogen, favoring non-*mcy* strains (108,111,127).

Microcystis is well-known for its ability to form dense blooms in freshwaters and for its dynamic production of MC. The factors that control MC production are complex as the *mcy* operon is phylogenetically scattered and is not directly related to the presence of specific functional genes (10). High light conditions, up to a certain maximum, favor *mcy*-encoding strains and MC production, contributing to the toxigenic to non-toxigenic bloom succession as light limitation increases later in the season (128,129). This supports the proposed role of MC in protecting *Microcystis* and key proteins like rubisco from oxidative stress, as MCs have been shown to bind such proteins during early and peak bloom periods characterized by high light and reactive oxygen species (ROS) (81,130). Although extracellular transport of MCs remains unconfirmed, extracellular MCs increase colony size and exopolysaccharide substance (EPS) production. These effects enhance *Microcystis* fitness by improving nutrient and light acquisition via microbial interactions while also deterring grazers who typically avoid feasting on larger colonies (131). Although MC production can deter grazers in experimental settings, it likely did not evolve for this purpose, as MC biosynthesis predates metazoans (132–134). Instead, physical defense mechanisms, such as increased colony size and EPS composition may be stronger predictors of grazer deterrence (135–138). Moreover, *Microcystis* spp. produce many other bioactive secondary metabolites that have detrimental behavioral and lethal effects on common phytoplankton grazers such as daphnia (139,140). Compared to other algae, *Microcystis* have lower rates of mortality due to zooplankton grazing, however, the factors mediating this trait likely go far beyond the ability to produce MC including colony formation, EPS makeup, and the diverse secondary metabolite repertoire of *Microcystis* (115,141).

Microcystis spp. harbor a strain-specific phycosphere of associated microbes that both benefit from and support the cyanobacteria within a colony (17,142). The phycosphere enhances *Microcystis* spp. fitness by exchanging vitamins, cofactors, regenerated nutrients, and oxidative stress protection in return for organic matter and sulfur (143–145). As a result, photosynthetic performance and growth rates are higher in xenic versus axenic cultures. Colony formation further enhances fitness by reducing resource demands, particularly for limited nutrients like iron and organic nitrogen (146,147). Within colonies, light and oxidative stress are mitigated with photoprotectants produced by *Microcystis* and ROS decomposition by heterotrophic bacteria (148,149). In addition to chemical exchanges, *Microcystis* spp. EPS mucilage is strain-specific, attracting specific bacterial populations who adhere to and use the mucilage as a carbon source while also maintaining colony integrity (17,150). The phycosphere of *Microcystis* spp. remains understudied, especially in terms of how it changes temporally and in response to strain successions.

Given the strain-specific functional traits that shape *Microcystis* spp. ecophysiology across diverse latitudes, access to locally relevant isolates and genomes is crucial. The Western Lake Erie Culture Collection (WLECC) has begun addressing this need in the Laurentian Great Lakes region by curating and sequencing 21 *Microcystis* cultures from WLE (24). This collection serves as a vital resource for Great Lakes researchers studying strain-specific fitness the region, while also providing a model for developing other regional culture collections.

1.3.2 Ecophysiology of Dolichospermum

As globally distributed bloom-forming cyanobacteria, *Microcystis* and *Dolichospermum* share many functional traits and responses to changing physicochemical conditions, but they also employ unique adaptations to outcompete one another and other phytoplankton (12). Notably,

they share genetic features, including buoyancy-regulating genes and diverse repertoires of biosynthetic gene clusters (BGCs) and mobile genetic elements (11,22). However, *Dolichospermum* is distinct for its ability to form secondary nitrogen-fixing cells, enabling it to thrive under nitrogen-limited conditions that typically lead to decline of *Microcystis* blooms (151). Moreover, this genus can produce hepatotoxins and potent neurotoxins such as ATX, guanitoxin (formerly anatoxin-a(S)), and SXT **(Figure 1.2B**) (22,152). Because of these fitness adaptations, *Dolichospermum* is one of the most ubiquitous bloom-forming cyanobacteria in freshwater and brackish ecosystems worldwide, tolerating salinities up to 15 g/L NaCl (153–155). Its range and bloom duration are expanding, driven by rising temperatures and increasingly common dry periods followed by intense rainfall, rapidly elevating water levels and resuspending labile nutrients in the littoral zone (156).

Studying *Dolichospermum*, a member of the order Nostocales, is complicated by its taxonomic ambiguity and overlap with related genera (157). Originally classified as *Anabaena* Bory ex Bornet et Flahault 1888, planktonic strains were reclassified as *Dolichospermum* by Wacklin et al. (2009) (158,159). However, this distinction can sometimes be problematic, as there is no clear phylogenetic separation between benthic and planktonic strains, and some Nostocales even oscillate between these two lifestyles (160,161). With the advent of genome sequencing, phylogenomic analyses have revealed a unified Nostocales clade encompassing *Anabaena, Dolichospermum*, and *Aphanizomenon*—collectively termed the ADA clade. This group shares a large pangenome, including core genes for photosynthesis, cell differentiation, and the assimilation of nitrogen, phosphorus, and sulfur (157). As of 2021, 40 ADA genomes are publicly available, and phylogenomics has identified eight sub-clades within this group (23).

Unlike *Microcystis*, the ability to produce neurotoxins and hepatotoxins is phylogenetically ordered in the ADA clade, suggesting a strong evolutionary basis for toxin diversity (22,157).

For the purposes of this thesis, *Dolichospermum*, including literature referring to planktonic *Anabaena*, will be the focus. *Dolichospermum* is characterized by uniseriate trichomes capable of forming gas vesicles and differentiating into heterocysts and akinetes (**Figure 1.7**) (117,162). Heterocysts are specialized cells found every 10–20 vegetative cells that fix atmospheric nitrogen under nitrogen-deprived conditions (151,163,164). Nitrogen limitation activates the *ntcA* gene, which induces *hetR* expression, ultimately leading to heterocyst formation within 18–24 hours (18). These cells also contain the *nif* operon, encoding nitrogenase subunits essential for nitrogen fixation (165,166). Fixed nitrogen from heterocysts and fixed carbon from vegetative cells are exchanged along the filament via the continuous periplasm (18). This diazotrophic capability can supply up to 25 to 50% of the total nitrogen in lakes, ameliorating nitrogen-deplete scenarios (167). Heterocyst formation often precedes bloom formation during periods of low dissolved inorganic nitrogen (DIN), particularly nitrate, and *Dolichospermum* proliferation is more strongly correlated with increases in phosphorus relative to DIN (168,169). Although less studied than in *Microcystis* spp., nitrogen form also influences *Dolichospermum* spp. dynamics. Culture-based experiments on *Dolichospermum* FACHB-1250 showed enhanced biomass accumulation on ammonia and urea, but inhibition of growth in the presence of nitrate (170). Both nitrogen and phosphorus, and their different forms, regulate *Dolichospermum* and *Microcystis* blooms, highlighting the need for dual-nutrient management strategies to mitigate toxic cyanobacterial proliferation.

The *Dolichospermum* life cycle is also distinct from that of *Microcystis* as it forms akinetes, which are dormant, spore-like cells that develop under light or phosphorus limitation to

survive harsh conditions such as overwintering, dry seasons, and flooding (151). Akinetes settle in the sediment and are recruited in late spring and mid- to late autumn into the hypolimnion, which is nutrient replete due to upwelling and mixing (171–173). Beyond nutrients, light and temperature also play a role in akinete formation and recruitment (174). Wind-driven turbulence can lift akinetes into shallow zones where they are exposed to light, warmer temperatures, and labile nutrients, fueling growth into vegetative cells (175). *Dolichospermum* is better adapted to cooler temperatures than *Microcystis*, with optimal growth and heterocyst formation around 20°C (176). However, *Dolichospermum* blooms also like it cold, being the most documented taxa to bloom in waters below 15°C (177). Moreover, *Dolichospermum* spp. have even been observed to bloom at water temperatures between 2-4°C, although over 25% of the cells were akinetes, indicating bloom senescence (178). Together, akinete and heterocyst formation give *Dolichospermum* a competitive advantage during early and late bloom phases characterized by low temperatures and nutrient concentrations.

Another competitive advantage for *Dolichospermum* is its extracellular organic matter (EOM), particularly its EPS, which is rich in nitrogenous compounds including proteins containing -NH functional groups (179). Unlike non-diazotrophs, *Dolichospermum* increases EPS production under nitrogen limitation, likely due to its ability to fix nitrogen (180). Like *Microcystis* spp., *Dolichospermum* spp. hosts a distinct phycosphere within their EPS that shapes nutrient cycling by associated bacteria (181). A comparative study found that *Dolichospermum* EOM composition is simple, made up of mainly small molecules and α-glucosidase genes for carbohydrate breakdown. Meanwhile, *Microcystis* EOM was richer in organic acids and genes for nitrate-to-ammonium conversion. These differences suggest that *Microcystis* promotes microbial nitrogen recycling within its EOM, while *Dolichospermum* relies on internal processes

for nutrients like nitrogen fixation and phosphorus hydrolysis, but may also benefit from carbohydrate breakdown by its phycosphere (182). These distinct phycosphere dynamics may explain in part how both species coexist and fill different niches within cyanoHABs.

Dolichospermum uniquely produces both hepatotoxins and neurotoxins while supplying its own nitrogen through nitrogen fixation. Its blooms are expanding globally, driven in part by imbalanced nutrient reductions during different bloom phases, shifting phytoplankton communities in its favor (167,183). Despite its growing impact, the genetics and ecophysiology of *Dolichospermum* remain far less studied than *Microcystis*, leaving significant gaps in our understanding of its ecology (184).

1.3.3 Dueling Bloom Dynamics of Microcystis and Dolichospermum in Shared Waters

Microcystis and *Dolichospermum* are two of the most widespread bloom-forming cyanobacteria globally and frequently co-occur or succeed one another in shared waters (11,13). In temperate lakes, seasonal phytoplankton cycles typically begin with spring diatom blooms that sequester phosphorus, followed by nitrogen- and phosphorus-rich runoff fueling diazotrophic blooms, and finally non-diazotrophic blooms in warmer summer waters (185). Temperature and nutrients are major drivers of the transition from diazotroph- to non-diazotroph-dominated blooms (186). In systems where *Microcystis* and *Dolichospermum* co-occur, such as Winam Gulf, studies suggest they may suppress each other's growth via nutrient-dependent allelochemical interactions—e.g., ATX and MC production (33,76). In these tropical and subtropical systems, wet and dry seasons strongly influence nutrient fluxes and understanding how their co-occurrence shapes bloom toxicity is especially critical (39).

Nutrient management policies often focus on phosphorus reduction, such as enhanced wastewater treatment or phosphorus detergent bans. Experimental lake studies show that

phosphorus-only reductions may shift phytoplankton dynamics rather than eliminate eutrophication-driven cyanoHABs, potentially exacerbating bloom severity or toxicity (187,188). For example, WLE has been experiencing intense, *Microcystis*-dominated cyanoHABs in the summer for decades (42). However, *Dolichospermum* becomes abundant later in the season, bringing with it the ability to produce the potent neurotoxin, SXT (62,189). The production of SXT in Lake Erie has been documented in the central basin and Maumee River, making its production in WLE a novel and alarming occurrence (190–192). Managing nitrogen and phosphorus loads to the basin, as well as understanding the role of regenerated nutrients on bloom dynamics, is vital in mitigating the consequences of cyanoHABs.

Despite regional variation, the shifting dynamics between *Microcystis* and *Dolichospermum* in shared waters underscore the urgent need for long-term monitoring of taxa and toxin production in major water bodies. Frequent and consistent monitoring is especially important, as these dynamics are likely evolving in response to rapidly shifting anthropogenic stressors (65).

1.4 The Diverse Biosynthetic Repertoire of *Microcystis*

Bacteria, fungi, and plants are prolific producers of secondary metabolites, a trait largely attributed to the abundance of BGCs encoded in their genomes (193). BGCs are groups of physically co-localized genes that collectively encode the enzymatic machinery for the synthesis of specific secondary metabolites and their congeners (28). Recent advances in genomics, bioinformatics, synthetic biology, and analytical chemistry, and the creation of gold standard BGC detection tools like antiSMASH now allow for the rapid discovery and analysis of BGCs across thousands of genomes (194,195). However, the rate of BGC discovery far exceeds our ability to experimentally characterize their metabolic products, resulting in a large number of uncharacterized BGCs (196). Confirmed BGCs, typically validated through molecular and

synthetic biology approaches such as gene knockouts or heterologous expression in non-native host organisms, are curated in the MIBiG database, which contains over 2,500 entries as of 2025. (193).

Microcystis biosynthetic potential is well documented in the MIBiG database, characterized to produce a wide array of secondary metabolites from over seven chemical classes, many of which co-occur in cyanoHABs worldwide. However, more than 90% of research has focused on MCs, likely because they are potent hepatotoxins that frequently reach dangerous, micromolar concentrations in blooms (7). The following sections explore both characterized and cryptic secondary metabolism in *Microcystis*, focusing on biosynthetic gene clusters (BGCs) with known and unknown synthesis products, respectively.

1.4.1 Characterized Secondary Metabolites

Microcystis secondary metabolism relies on three main biosynthetic pathways: nonribosomal peptide synthetases (NRPS), ribosomally synthesized and post-translationally modified peptides (RiPPs), and polyketide synthases (PKS) (21). Each pathway encodes a variety of core mega-synthases, accessory proteins, and tailoring enzymes, producing a chemically diverse range of metabolites (28). Each BGC can synthesize a diverse panel of congeners with some structural similarities and differences, arising from factors like substrate availability, enzyme promiscuity, and environmental conditions (197,198). NRPS and RiPP BGCs encode the production of molecules rich in amino acids and peptidic bonds, termed cyanopeptides. Peptides comprise over 60% of the known biologically active compounds produced by cyanobacteria, ranging in size between 400 and 1900 Daltons (199,200). In *Microcystis*, NRPS pathways produce compounds like anabaenopeptins and cyanopeptolins (**Figure 1.2A**), while hybrid NRPS-PKS pathways synthesize MCs, aeruginosins, and

microginins (**Figure 1.2A**) (21,198,201–203). RiPP pathways produce cyclic peptides such as microviridins, aerucyclamides, and piricyclamides (**Figure 1.2A**) (140,204,205). These molecules can have both proteinogenic and non-proteinogenic amino acids, and other non-amino acid moieties, incorporated into their structures with high variability across almost all positions resulting in a diverse congener suite for each metabolite class (206,207).

While hybrid NRPS-PKS pathways are well studied in *Microcystis* due to their role in MC synthesis, PKS-only pathways remain poorly characterized. In fact, there are no *Microcystis* BGCs annotated solely as PKS pathways in the MIBiG database (193). Some evidence suggests *Microcystis* may produce cyclophane-like compounds (e.g., merocyclophanes) using a hybrid Type I and Type III PKS system paired with a NRPS domain, though synthesis has not been confirmed (208,209). These cryptic PKS pathways are discussed in more detail in the following section on cryptic biosynthesis.

As with other functional traits, secondary metabolism in *Microcystis* is highly strain-specific. Over 50 unique BGCs were identified from a metabologenomics analysis of 21 *Microcystis* isolates from the WLE culture collection, with each isolate encoding approximately 5-15 BGCs each. While some phylogenetic patterns of secondary metabolism emerged, no two strains shared the exact same BGC profile, each differing by the presence or absence of at least one BGC (21). This diversity likely contributes to strain-specific niche partitioning within eutrophic freshwater systems.

1.4.2 Cryptic Microcystis Secondary Metabolism

Microcystis genomes encode a wealth of biosynthetic gene clusters (BGCs), many of which remain cryptic, or "orphan", due to the lack of experimentally confirmed products (109). Despite encoding up to 50 unique BGCs across only 21 isolates, only seven of these have been

definitively linked to their corresponding metabolite classes to date in *Microcystis* (**Figure 1.2A**) (21,193). Linking BGCs to their products is non-trivial, typically requiring heterologous expression in a non-native host followed by structural validation using techniques such as NMR spectroscopy (210). Orphan BGCs in *Microcystis* have been recognized since at least 2013 with few successful linking studies thereafter, highlighting the persistent gap between BGC discovery and functional characterization.

Characterized *Microcystis* BGCs typically produce cyanopeptides, while cryptic clusters are predicted to encode a more diverse set of products from PKS, terpene, and RiPP bacteriocin and cyanobactin pathways (211). These pathways are predicted to produce molecules that are less nitrogen-rich than cyanopeptides, potentially playing different or parallel ecological functions. One notable example of a cryptic BGC encoded by *Microcystis* is a hybrid Type I/III PKS BGC that closely resembles the [7,7] paracyclophane-producing BGCs in *Nostoc* (212). This BGC, termed PmNT3 by Yancey and Hart et al. (2024), has been studied across three research articles but remains unlinked to a confirmed compound (21,208,213). Interestingly, in *Microcystis* PCC 7806, PmNT3 has an inverse transcriptional relationship with the *mcy* operon, suggesting a possible regulatory or reciprocal role with MC production (213). This underscores how characterizing cryptic BGCs can reveal insights into bloom toxicity and the evolutionary pressures shaping cyanobacterial secondary metabolism.

Cyanobacterial metabolites serve as a dual source of environmental toxins and bioactive compounds with pharmaceutical potential, and linking BGCs to their products provides access to these molecules for further functional and therapeutic investigation. Notable cyanobacterial therapeutic leads include cryptophycins, curacin A, and dolostatins, which exhibit antiproliferative, antimitotic, and anticancer properties, respectively (214–217). Fortunately,

emerging tools like NPLinker and synthetic biology platforms such as CRAGE (Chassis-Independent Recombinase-Assisted Genome Engineering) are accelerating efforts to connect BGCs to their chemical products (218,219). Despite these advancements, the cryptic world of biosynthesis encoded by *Microcystis* remains mostly uncharted, warranting *in silico* and *in vivo* characterization of these BGCs.

1.5 The Understudied Biosynthetic Potential of *Dolichospermum*

The secondary metabolism of *Dolichospermum* is significantly less characterized than that of *Microcystis*, despite its capacity to produce a broader array of cyanotoxins than any other cyanobacterial genus (22). Alongside the broader ADA clade, *Dolichospermum* is well known for synthesizing MCs and potent neurotoxins including saxitoxin (SXT), anatoxin-a (ATX), guanitoxin, and β-N-methylamino-L-alanine (BMAA) (**Figure 1.2B**) (152,220–223). The research bias toward *Microcystis* likely stems from its status as the primary microcystin (MC) producer in most freshwater systems and its dominance among bloom-forming genera, leading to a greater number of publicly available genomes (>150 for *Microcystis* vs. <40 for *Dolichospermum*) and a more extensively studied role in bloom dynamics (224). However, the ability of *Dolichospermum* to produce a diverse array of hepatotoxins, neurotoxins, and other chemically complex metabolites underscores the urgent need to investigate its biosynthetic potential.

The increasing availability of ADA clade genomes has enabled deeper exploration of their secondary metabolite biosynthetic potential. Österholm et al. (2020) analyzed BGC content across 31 publicly available ADA clade genomes and revealed that biosynthetic potential clustered phylogenetically by sub-clade. (22). Among the genomes annotated as *Dolichospermum*, one sub-clade included strains encoding BGCs for MC, anabaenopeptin, and

hassallidin, while a separate sub-clade encoded BGCs for anacyclamide and SXT. BGCs encoding the production of ATX, microviridins, pseudospumigins, aeruginosins, and anabaenolysins were primarily associated with *Aphanizomenon* and benthic strains of *Anabaena.* These findings contrast with phylogenetic studies of *Microcystis*, where the ability to produce cyanotoxins and toxic cyanopeptides is less phylogenetically ordered. This discrepancy may reflect genus-specific evolutionary dynamics underlying biosynthetic potential and could offer valuable insight into the ecological and functional roles of these metabolites.

In lakes where *Dolichospermum* is abundant at specific phases of the bloom, such as WLE, its region-specific biosynthetic potential remains largely understudied. A study analyzing metagenomic data from the 2014 WLE cyanoHAB season identified BGCs from ADA metagenome-assembled genomes (MAGs) putatively encoding the synthesis of aeruginosins, pseudospumigins, cyanopeptolins, and geosmin. These BGCs were increasingly transcriptionally active later in the bloom season (September/October). However, no MAGs in the study contained neurotoxin-encoding BGCs (62). More recently, a study in WLE analyzing data from 2016–2022 detected the sporadic presence of SXT BGCs in MAGs annotated as *Dolichospermum sp000312705* in the Genome Taxonomy Database (GTDB), suggesting the emergence of neurotoxin-producing strains in this critical drinking water source (189). The biosynthetic potential of *Dolichospermum* in the Winam Gulf is even less studied than in WLE, remaining largely unexplored despite its potential public health and ecological relevance. The first targeted investigation of *Dolichospermum* biosynthetic potential in this region is presented in Chapter 3 of this thesis (225).

1.6 Environmental Drivers of *Microcystis* and *Dolichospermum* Secondary Metabolites

The expansion, intensification, and increasing toxicity of cyanoHABs worldwide is closely linked to anthropogenic eutrophication and climate change related events like increasing temperatures and CO_2 (125). These abiotic drivers, along with biotic factors such as grazing and allelopathy, shape cyanobacterial patterns of synthesizing secondary metabolites. While the exact contributions of these factors remain poorly constrained, substantial research has explored how they affect toxin biosynthesis, particularly MCs, and to a lesser extent, ATX, SXT, and anabaenopeptins. However, comparing these studies remains challenging due to variability in experimental design, analytical methods (e.g., inclusion or exclusion of protein-bound toxins), and limited transcriptomic research beyond MCs, owing to the recent elucidation and complexity of other BGCs (226,227).

Experimental studies show that altering growth conditions such as temperature, light intensity, and CO_2 levels trigger strain- and toxin-specific responses in cyanobacteria. While MC production is often positively correlated with growth rate, toxins like SXT, ATX, and anabaenopeptins show an inverse relationship (228,229). SXT and ATX are typically produced under lower light and temperature, while higher temperatures enhance anabaenopeptin production but suppress growth of the producing strain (228–230). Congener-specific patterns also emerge, with CO_2 enrichment favoring the production of homoanatoxin, a methylated variant of anatoxin-a, over anatoxin-a itself (226). For anabaenopeptins, increased light intensity selects for the production of anabaenopeptolide 90B over anabaenopeptins A and C, all three of which have very similar carbon-to-nitrogen (C:N) ratios (229). This may indicate that strains producing anabaenopeptins maintain the production of molecules with similar properties and functions to improve survival under varying conditions.

For MCs, toxin production and growth rate appear to be linked, with fast growing *Microcystis* cells containing higher levels of intracellular MCs (231). This could be explained by the experimental evidence linking active photosynthesis with MC production (232,233). However, this relationship may be nutrient-dependent, as experimental studies have shown a much stronger linkage between CO_2 enrichment and MC production when cultures were grown with dissolved organic phosphorus rather than dissolved inorganic phosphorus (233,234). High light and UV-B also upregulate *mcy* expression and select for toxic strains of *Microcystis* (233,235). It has been hypothesized that MCs provide intracellular protection under high light conditions that generate harmful levels of reactive oxygen species. Under such stress, MC has been shown to bind specifically to proteins involved in photosynthesis, such as Rubisco (130). Taken together, these data suggest that climate-driven abiotic stressors shape the toxin composition of cyanoHABs, and that these toxins may, in turn, enhance the fitness of specific cyanobacterial strains under varying environmental conditions.

Nitrogen and phosphorus availability are key components mediating cyanobacterial secondary metabolism and growth. Many cyanotoxins are nitrogen-rich, with SXT having a C:N ratio below 1.5, and some cyanopeptides like microviridins containing over 16 nitrogen atoms. While phosphorus limitation generally increases toxin production across multiple metabolite classes (e.g., MCs, anabaenopeptins, SXT, ATX), the effects of nitrogen are more complex and congener-specific (229,236,237). For example, SXT and ATX are often produced at higher levels under nitrogen-limited conditions (230,238). This pattern may reflect a growth versus defense tradeoff, where bloom-forming cyanobacteria shift toward increased toxin production under nutrient stress to deter grazing, store nutrients, or reduce competition. Although SXT and ATX are nitrogen-rich, their relatively simple biosynthetic pathways may still offer an efficient

defense investment when nitrogen is scarce (239). These responses may also be influenced by co-varying physicochemical factors such as temperature or conductivity (61). For MCs, the response to changing nitrogen conditions is highly strain- and context-dependent. Although the presence of three *ntcA* binding sites in the *mcyA*/*D* promoter region suggests that nitrogen may regulate MC production, transcriptomic studies show inconsistent patterns under nitrogen depletion (227,231,240). These discrepancies hint at more complex, nutrient-dependent interactions involving nitrogen form, carbon metabolism, or micronutrients such as iron. For example, urea can act as both a nitrogen and carbon source for MC production, especially under high pH conditions (241). Field studies indicate that reduced nitrogen forms (e.g., ammonium, urea) enhance *Microcystis* growth and toxicity (53).

Nitrogen availability also influences MC congener composition (242,243). Under low nitrogen conditions, more toxic congeners such as MC-LA, which contain leucine and alanine, are favored over less toxic, arginine-rich congeners like MC-RR in large eutrophic basins like WLE (197). Modeling of *Microcystis* populations with and without the *mcy* operon suggests that phosphorus reduction alone may inadvertently promote toxic bloom conditions when nitrogen concentrations remain unregulated and unmanaged (244,245). These findings underscore the importance of nutrient dynamics, specifically concentration, chemical form, and N:P ratio, in shaping cyanotoxin composition and overall bloom toxicity. Together, they highlight the urgent need for dual nutrient management strategies in eutrophic lakes dominated by toxin-producing cyanobacteria.

In addition to complex abiotic drivers, cyanobacterial toxin production is also shaped by biotic interactions, particularly those involving grazers, which cyanotoxins may help deter. *Microcystis'* ability to produce MCs and form large colonies protects them from common grazers

such as mussels and daphnia (115). Daphnia are common grazers of phytoplankton and are highly sensitive to MCs, though their survival depends on the specific congener ingested (246). Interestingly, *Microcystis* has been shown to increase MC production upon direct contact with daphnia in the water column, suggesting that cyanotoxin synthesis may function as an inducible defense mechanism (247). Protozoan grazers such as amoebae selectively consume certain *Microcystis* strains, though the extent to which this selectivity is influenced by toxin-producing capability remains poorly understood (248). These findings support the hypothesis that cyanotoxins provide a defensive advantage in freshwater ecosystems. However, grazing pressure and toxin production are both likely co-regulated by changing environmental conditions, adding to the complexity of disentangling the biotic and abiotic drivers of cyanotoxin production. Although considerable work has examined temporal shifts in bloom taxonomic composition and the strain-specific BGC profiles of bloom-forming cyanobacteria, far less is known about how BGC composition and the synthesis of metabolites—especially those beyond the traditionally studied cyanotoxins—change over time or responds to environmental and ecological pressures (12,21,22).

1.7 Combinatorial Toxic Effects of Cyanobacterial Secondary Metabolites on Human and Environmental Health

1.7.1 Combinatorial Toxic Effects of Cyanotoxins on Aquatic Organisms

CyanoHABs are made up of complex mixtures of metabolites that have adverse, toxic effects towards organisms within the water column that often cannot be explained by the presence of an individual cyanotoxin (249,250). However, most studies either focus on the effects of individual cyanotoxins or uncharacterized cyanobacterial extracts against mammalian cell models and aquatic organisms. Since cyanotoxins and cyanopeptides almost always co-occur

within cyanoHABs, testing the toxicity of individual metabolites is not representative of their *in-situ* effects. Naturally occurring mixtures of cyanobacterial metabolites should be tested for their adverse effects, but without accurate characterization of the molecules present in the extract, the toxic effects of cyanobacterial metabolite mixtures remain elusive. Moreover, studies often fraction whole extracts, which may remove the toxic effects of co-occurring and potentially synergistic molecules when they are separated. The combinatorial toxic effects of co-occurring cyanotoxins and complex cyanobacterial extracts on aquatic organisms such as plants, zooplankton, microcrustaceans, *Daphnia*, and fish have been extensively studied with a range of results indicating some additive and synergistic toxic effects.

Duckweeds (*Lemna minor* and *Spirodela polyrhiza*) play an important role in freshwater ecosystems, and are even hypothesized as a potential bioremediation strategy for cyanoHABs due to their ability to compete with cyanobacteria for nutrients (251). However, the combination of MC-LR with aeruginosins, anabaenopeptins, cylindrospermopsins, and microginins exhibit synergistic deleterious effects towards the growth and oxygen evolution ability of duckweed (252). Increased research on the effects of cyanotoxins on vital freshwater flora are needed to understand this trophic interaction and to provide necessary data to inform flora-based bioremediation of cyanoHABs.

Zooplankton feed on primary consumers like toxigenic cyanobacteria and are greatly affected by ingesting cyanotoxins. For example, ternary mixtures of MCs, anabaenopeptins, and microginins synergistically inhibited the swimming speed, degraded F-actin, and triggered cytotoxicity in the rotifer *Brachionus calyciflorus* (253). *B. calyciflorus* can thrive well under eutrophic conditions that drive cyanoHABs, leading to their high population densities particularly in the absence of large zooplankton such as daphnia (254,255). The toxic effects of

cyanotoxins on rotifers affects multiple trophic levels within an ecosystem by reducing the fitness of competing zooplankton and decreasing primary producer consumption.

The viability, physiology, and behavior of large zooplankton such as daphnia, which are common water fleas and known cyanobacteria grazers, are affected by extracts of *Microcystis* from culture- and field-based studies. In these studies, intracellular cyanobacterial extracts with lower MC content were more toxic than high MC containing extracts, suggesting that other components of the extracts contributed to toxicity individually or through an interactive mechanism (256). Anabaenopeptin-B and MC-LR exhibit synergistic toxic effects towards daphnia (257,258). Additionally, combinations of aeruginosin, anabaenopeptins, and microginins have synergistic and additive toxic effects on daphnia viability and behavior (259). Beyond intracellular metabolites, exuded metabolites from *Microcystis* can also inhibit the viability of *Daphnia*. Exudates from two MC-producing *Microcystis* isolates were collected at multiple growth phases and exposed to daphnia. Exposure to exudates from the stationary growth phase significantly impaired daphnia behavior, viability, and stress response compared to those from the exponential phase (249). Exudates from the low-MC producing strain had more toxic effects than the higher-MC producing strain, suggesting that *Microcystis* strains with lower hepatotoxin output may produce more inhibitory extracellular metabolites to deter predators. These findings underscore the importance of studying the toxic effects of both intra- and extra-cellular cyanobacterial metabolites, and the strain-specific and bloom phase-dependent impacts on daphnia.

Shrimp play a vital role in freshwater ecosystems because they decompose organic matter, promote nutrient cycling, and channel molecules and resources to higher trophic levels (260–262). Because shrimp connect many different trophic levels, the toxic effects and toxin

accumulation within these organisms are important to understand. Analyses show the accumulation of anabaenopeptins, spumigins, and aeruginosins in blue mussels and shrimp collected from a cyanoHAB event in the Baltic Sea (263). Moreover, three independent studies show synergistic impairments of shrimp, mainly *Thamnocephalus platyurus,* caused by the mixture of cyanopeptolins, aeruginosins, and spumigins (139,263,264). Because shrimp can channel molecules through multiple trophic systems via bioaccumulation and transformation, it is important to understand their relationship with cyanotoxins and cyanotoxin mixtures.

Freshwater ecosystems host a diverse array of fish that accumulate and are effected by cyanotoxins, including species consumed by humans and other predators such as rainbow trout, pike, bass, and carp, as well as species like zebrafish, which serve as effective models to assess toxicological hazards posed by biomolecules (265–267). Analysis of cyanotoxins within commonly fished species such as bass, trout, and pike along the Pike River near Lake Champlain (Canada) revealed the presence of MCs, cylindrospermopsins, and anabaenopeptins in 30% of fish muscle samples, representing the first report of anabaenopeptins in wildlife (268). The maximum MCs detected in a fish sample was 1.15-fold higher than the adult daily intake recommended by the World Health Organization (WHO) (268,269). Moreover, an analysis of the ternary mixture of MCs, anabaenopeptins, and cylindrospermopsins on a rainbow trout gill cell model (RTgill-W1) revealed synergistic toxic effects associated with apoptosis, reduced glutathione levels, and oxidative stress (270). Additional cyanotoxins such as cyanopeptolins and microginins were revealed to have significantly more toxic effects on zebrafish larvae (*Danio rerio*) than MCs (271). These data suggest that there may be interactive effects of cyanobacterial metabolites on heavily fished and eaten species of fish in freshwater bodies.

Although there are many studies investigating the combinatorial toxic effects of cyanotoxins on aquatic organisms, significant gaps remain in understanding the precise contributions of individual cyanotoxins and their interactions to overall toxicity. These uncertainties arise due to challenges such as extract fractionation (disrupting potential metabolite interactions), limited technology for identifying uncharacterized compounds within extracts, and the high technical and financial costs associated with mechanistic and pathway analyses of these interactions at the cellular level. Further research should be conducted not only to elucidate the effects of these cyanotoxin mixtures on aquatic organisms, but also how those effects may impede ecosystem function and homeostasis, and how cyanotoxins may thus travel up through the trophic system towards human, wildlife, and domesticated animal consumption.

1.7.2 Combinatorial Toxic Effects of Cyanotoxins on Mammalian Cell Models

In contrast to the extensive research on the effects of cyanobacterial metabolite mixtures on aquatic organisms, studies using mammalian cell models remain limited and challenging, largely due to the lack of consensus on an appropriate in vitro model system. This challenge stems from the fact that MCs—and likely other cyanopeptides, though their mode of entry remains unconfirmed—require organic anion transporting polypeptides (OATPs) for uptake. However, OATPs are often absent or expressed at low levels in commonly used mammalian cell models, limiting their utility in toxicological studies focused on cyanobacterial metabolites (272,273). For the study of cyanobacterial metabolite mixtures with known hepatotoxins, the human hepatocellular carcinoma HepG2 cell line (American Type Culture Collection (ATCC) HB-8065), cervical cancer cell line HeLa (ATCC CCL-2), primary human hepatocytes, and engineered TransportoCells^{TM} OATP1A2 and OATP1B1 (Corning, NY, USA) cell lines are often used (274). Studies investigating the toxic effects of cyanobacterial isolate extracts towards

these cell lines have found strong linear relationships between the toxic effects in HepG2, HeLa, and primary hepatocyte cell lines, confirming these cell lines as good, comparable models for studying complex cyanobacterial metabolite mixtures (274). In a study where these cell lines were exposed to cyanobacterial extracts collected across a spatiotemporal gradient of cyanoHABs in the Indian River Lagoon (Florida, USA), HepG2 cells were far less sensitive to extracts than the engineered TransportoCells likely due to less OATp activity in the HepG2 cells (275). Another study that exposed extracts from MC- and retinoic acid-producing cyanobacterial isolates to HepG2 cells found that MC and retinoic acid content was a poor predictor of an extracts effect on cell viability and genotoxicity (276). These data indicate that the whole cyanobacterial extracts likely contribute more to toxicity than the parts that make up the extract, underscoring the need to investigate these combinatorial effects more in depth on standard cell lines.

Cyanobacteria produce potent neurotoxins that frequently co-occur with hepatotoxins such as MCs and CYLs during cyanoHABs, and their combined presence is hypothesized to pose heightened risks to human health through additive or synergistic effects (277). A study tested the individual and combined effects of pure cyanotoxins MC-LR, ATX, and CYL against three murine neural and immune cell models (278). The results found that mixtures of these molecules at low doses, threefold less concentrated than individual compounds, were three to fifteen times more potent at inducing apoptosis and inflammation in the cell models (278). Another study found that the mixture of the neurotoxin BMAA and two of its isomers, aminoethyl glycine (AEG) and 2,4-diaminobutyric acid (2,4DAB), significantly decreased cell viability by 19% and increased apoptosis induction by 110% as compared to the molecules individually (279). The neurotoxic, neuroinflammatory, and neurodegenerative effects of cyanotoxin mixtures remains

understudied, but further *in vivo* research is needed to improve safety limits and guidelines regarding neurotoxin limit levels in surface waters. Additionally, cyanotoxins are known to be aerosolized, and this could provide further complications for neuroinflammation and degeneration for lakeside populations (280–282).

As previously mentioned, there are limited studies elucidating the combinatorial toxic effects of cyanotoxins on mammalian cell models, underpinning the vulnerability these naturally occurring mixtures pose towards human health. Research is still needed to understand how co-occurring cyanotoxins may interact within humans, and how their route of exposure mediates or amplifies these mechanisms. Moreover, studies regarding the mode of entry and mechanism of action on human metabolic pathways by cyanotoxin mixtures are scarce. Future research with these foci would be pivotal in determining accurate protective measures for human exposure to cyanotoxin mixtures.

1.8 Dissertation Outline and Research Questions

This dissertation investigates the patterns and co-varying environmental conditions of *Microcystis* and *Dolichospermum* abundance*,* biosynthetic potential, and cyanopeptide synthesis in two geographically distant Great Lakes basins. Further, it identifies the overlooked threats posed by co-occurring and unmonitored cyanobacterial metabolites beyond MCs in these freshwater bodies. The integration of multi-omics and toxicology approaches serve as an example for future studies on cyanobacterial secondary metabolism, early warning bloom toxicity models, and provides testable hypotheses about the ecology and dangers of co-occurring cyanotoxins.

In Chapter two, the temporal distributions and drivers of succession between *Microcystis* and *Dolichospermum* dominance and biosynthetic potential in WLE are qualitatively and

quantitatively characterized. This study reveals the co-variance of temperature, nitrogen concentration and form, and grazer populations with temporal patterns and successions of *Microcystis* and *Dolichospermum*, biosynthetic potential, and cyanopeptide synthesis. Co-varying BGCs from these two genera cluster into temporally variant modules of biosynthesis representative of functionally distinct subpopulations that may be predictive of niche-controlled bloom dynamics. Further temporal analysis of cyanopeptides synthesized by BGCs within the modules indicated the presence of three phases of cyanoHABs in WLE that significantly co-vary with nitrogen form and grazer abundance. This study leverages long-term, multi-omic data to identify the taxonomic and biosynthetic successions that take place in WLE cyanoHABs and the potential drivers of such seasonality. These findings will enhance early warning models of bloom toxicity and help shift the field beyond a MC-centric view of cyanoHABs by highlighting the broader chemical ecology and dynamic succession of diverse cyanobacterial metabolites.

In Chapter three, the cyanobacterial community, functional genes, and biosynthetic potential over two years along a spatial gradient in the Winam Gulf, Lake Victoria, Kenya were characterized using metagenomic sequencing. *Dolichospermum* was the main bloom-forming cyanobacterium found in large cyanoHAB events located in distinct areas in the gulf, while *Microcystis* was found at lower concentrations mostly in turbid regions like river mouths. The production of MC in the gulf was attributed to *Microcystis*, clarifying the toxin-producing cyanobacteria present during sampling periods for the first time in this region. Beyond MCs, over 300 unique, cryptic BGCs were identified from cyanobacteria, highlighting the underexplored cyanobacterial secondary metabolism in the gulf. This study provides a valuable metagenomic dataset to this region and clarifies long-standing questions for the gulf such as what organism was responsible for producing MC during the sampling periods. This research provides

insights into potential management strategies to prevent both *Microcystis* and *Dolichospermum* blooms that are spatially stratified in this region and context to the many microscopic characterization studies of cyanoHABs in the gulf.

In Chapter four, multi-omic and toxicological approaches were used to investigate the individual and combinatorial effects of co-occurring cyanobacterial metabolites on human lung, liver, and kidney cell lines. A gradient of toxic effects was observed among three MC congeners and anabaenopeptins, with MC-LA exhibiting the highest toxicity, followed by MC-LA and anabaenopeptins A and B with similar effects, and MC-RR showing the lowest toxicity across all cell types. The combination of MC-LA and anabaenopeptin B demonstrated dose-dependent synergistic effects in the cell lines. These findings suggest the need for further toxicological evaluation of anabaenopeptins, which are not currently regulated or routinely monitored like MCs. Beyond these known cyanopeptides, the toxicity of *Microcystis* culture extracts could not be explained by MC and anabaenopeptin concentrations, suggesting that other metabolites and mixture effects contribute substantially to the observed variability in toxicity. This study underscores the complex and potentially hazardous combinatorial effects of both characterized and uncharacterized cyanobacterial secondary metabolites, emphasizing the importance of broadening the scope of toxin discovery, monitoring and regulation.

In Chapter 5, the conclusions from chapters two, three, and four are synthesized and avenues for continued research on these topics are discussed. New research questions and approaches are proposed to address remaining gaps and testable hypotheses proposed in this dissertation.

1.9 References

1. Carpenter SR, Fisher SG, Grimm NB, Kitchell JF. Global change and freshwater ecosystems. Annual Review of Ecology and Systematics. 1992;23:119–39.

2. Smith VH, Joye SB, Howarth RW. Eutrophication of freshwater and marine ecosystems. Limnology & Oceanography. 2006 Jan;51(1.2):351–5.

3. Zipperer WC, Northrop R, Andreu M. Urban development and environmental degradation. Oxford Research Encyclopedia of Environmental Science [Internet]. 2020;2020. Available from: https://research.fs.usda.gov/treesearch/60941

4. Huisman J, Codd GA, Paerl HW, Ibelings BW, Verspagen JMH, Visser PM. Cyanobacterial blooms. Nat Rev Microbiol. 2018 Aug;16(8):471–83.

5. Paerl HW, Huisman J. Climate. Blooms like it hot. Science. 2008 Apr 4;320(5872):57–8.

6. Heisler J, Glibert P, Burkholder J, Anderson D, Cochlan W, Dennison W, et al. Eutrophication and harmful algal blooms: A scientific consensus. Harmful Algae. 2008 Dec;8(1):3–13.

7. Janssen EML. Cyanobacterial peptides beyond microcystins – A review on co-occurrence, toxicity, and challenges for risk assessment. Water Research. 2019 Mar 15;151:488–99.

8. US National Office for Harmful Algal Blooms. Proceedings of the workshop on the socio-economic effects of harmful algal blooms in the United States [Internet]. Woods Hole, MA, USA; 2021 Mar. Available from: https://hab.whoi.edu/wp-content/uploads/2021/04/HAB-Socioeconomics-Workshop-Proceedings_14.pdf

9. Chapra SC, Boehlert B, Fant C, Bierman VJ Jr, Henderson J, Mills D, et al. Climate change impacts on harmful algal blooms in U.S. freshwaters: A screening-level assessment. Environ Sci Technol. 20170725th ed. 2017 Aug 15;51(16):8933–43.

10. Dick GJ, Duhaime MB, Evans JT, Errera RM, Godwin CM, Kharbush JJ, et al. The genetic and ecophysiological diversity of *Microcystis*. Environ Microbiol. 20210614th ed. 2021 Dec;23(12):7278–313.

11. Li X, Dreher TW, Li R. An overview of diversity, occurrence, genetics and toxin production of bloom-forming *Dolichospermum* (*Anabaena*) species. Harmful Algae. 2016 Apr 1;54:54–68.

12. Paerl HW, Otten TG. Duelling 'CyanoHABs': Unravelling the environmental drivers controlling dominance and succession among diazotrophic and non-N2-fixing harmful cyanobacteria. Environmental Microbiology. 2016;18(2):316–24.

13. Harke MJ, Steffen MM, Gobler CJ, Otten TG, Wilhelm SW, Wood SA, et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. Harmful Algae. 2016 Apr;54:4–20.

14. Benayache NY, Nguyen-Quang T, Hushchyna K, McLellan K, Afri-Mehennaoui FZ, Bouaïcha N. An overview of cyanobacteria harmful algal bloom (cyanohab) issues in freshwater ecosystems. In: Gökçe D, editor. Limnology - Some New Aspects of Inland Water Ecology [Internet]. IntechOpen; 2019. Available from: https://www.intechopen.com/books/limnology-some-new-aspects-of-inland-water-ecology/an-overview-of-cyanobacteria-harmful-algal-bloom-cyanohab-issues-in-freshwater-ecosystems

15. Stark GF, Martin RM, Smith LE, Wei B, Hellweger FL, Bullerjahn GS, et al. Microcystin aids in cold temperature acclimation: differences between a toxic *Microcystis* wildtype and non-toxic mutant. Harmful Algae. 2023 Oct 22;102531.

16. Le VV, Srivastava A, Ko SR, Ahn CY, Oh HM. *Microcystis* colony formation: Extracellular polymeric substance, associated microorganisms, and its application. Bioresource Technology. 2022 Sep 1;360:127610.

17. Smith DJ, Tan JY, Powers MA, Lin XN, Davis TW, Dick GJ. Individual *Microcystis* colonies harbour distinct bacterial communities that differ by *Microcystis* oligotype and with time. Environmental Microbiology. 2021;23(6):3020–36.

18. Flores E, Herrero A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. Nat Rev Microbiol. 2010 Jan;8(1):39–50.

19. Reynolds CS, Oliver ,Rod L., and Walsby AE. Cyanobacterial dominance: The role of buoyancy regulation in dynamic lake environments. New Zealand Journal of Marine and Freshwater Research. 1987 Sep 1;21(3):379–90.

20. McKindles KM, McKay RM, Bullerjahn GS. Genomic comparison of *Planktothrix agardhii* isolates from a Lake Erie embayment. PLOS ONE. 2022 Aug 23;17(8):e0273454.

21. Yancey CE, Hart L, Hefferan S, Mohamed OG, Newmister SA, Tripathi A, et al. Metabologenomics reveals strain-level genetic and chemical diversity of *Microcystis* secondary metabolism. Van Der Hooft JJJ, editor. mSystems. 2024 Jul 23;9(7):e00334-24.

22. Österholm J, Popin RV, Fewer DP, Sivonen K. Phylogenomic analysis of secondary metabolism in the toxic cyanobacterial genera *Anabaena, Dolichospermum*, and *Aphanizomenon*. Toxins. 2020 Apr 11;12(4):248.

23. Dreher TW, Davis EW, Mueller RS, Otten TG. Comparative genomics of the ADA clade within the *Nostocales*. Harmful Algae. 2021 Apr 1;104:102037.

24. Yancey CE, Kiledal EA, Chaganti SR, Denef VJ, Errera RM, Evans JT, et al. The Western Lake Erie culture collection: A promising resource for evaluating the physiological and genetic diversity of *Microcystis* and its associated microbiome. Harmful Algae. 2023 Jul;126:102440.

25. Bolius S, Wiedner C, Weithoff G. High local trait variability in a globally invasive cyanobacterium. Freshwater Biology. 2017;62(11):1879–90.

26. Mangan NM, Brenner MP. Systems analysis of the CO2 concentrating mechanism in cyanobacteria. eLife. 3:e02043.

27. Baker D, Godwin CM, Khanam M, Burtner AM, Dick GJ, Denef VJ. Variation in resource competition traits among *Microcystis* strains is affected by their microbiomes. mLife. 2023 Dec 18;2(4):401–15.

28. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, et al. Minimum Information about a Biosynthetic Gene cluster. Nature Chemical Biology. 2015 Sep 1;11(9):625–31.

29. Jones MR, Pinto E, Torres MA, Dörr F, Mazur-Marzec H, Szubert K, et al. CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. Water Res. 20210308th ed. 2021 May 15;196:117017.

30. Krausfeldt LE, Farmer AT, Castro HF, Boyer GL, Campagna SR, Wilhelm SW. Nitrogen flux into metabolites and microcystins changes in response to different nitrogen sources in *Microcystis aeruginosa* NIES-843. Environ Microbiol. 2020 Jun;22(6):2419–31.

31. Barchewitz T, Guljamow A, Meissner S, Timm S, Henneberg M, Baumann O, et al. Non-canonical localization of RubisCO under high-light conditions in the toxic cyanobacterium *Microcystis aeruginosa* PCC7806. Environ Microbiol. 2019 Dec;21(12):4836–51.

32. Große R, Heuser M, Teikari JE, Ramakrishnan DK, Abdelfattah A, Dittmann E. Microcystin shapes the *Microcystis* phycosphere through community filtering and by influencing cross-feeding interactions. ISME Commun. 2024 Dec 24;5(1):170.

33. Chia MA, Jankowiak JG, Kramer BJ, Goleski JA, Huang IS, Zimba PV, et al. Succession and toxicity of *Microcystis* and *Anabaena (Dolichospermum)* blooms are controlled by nutrient-dependent allelopathic interactions. Harmful Algae. 2018 Apr 1;74:67–77.

34. Vilar MCP, da Costa Pena Rodrigues TF, da Silva Ferrão-Filho A, de Oliveira E Azevedo SMF. Grazer-Induced chemical defense in a microcystin-*producing Microcystis aeruginosa* (Cyanobacteria) exposed to *Daphnia gessneri* infochemicals. J Chem Ecol. 2021 Nov;47(10–11):847–58.

35. Maguire TJ, Isabwe A, Stow CA, Godwin CM. Defining algal bloom phenology in Lake Erie. Harmful Algae. 2024 Nov 1;139:102731.

36. Dugan HA, Feiner ZS, Winder M, Sosik HM, Stanley EH. Advancing phenology in limnology and oceanography. Limnol Oceanogr Letters. 2024 Oct;9(5):506–11.

37. Polansky L, Robbins MM. Generalized additive mixed models for disentangling long-term trends, local anomalies, and seasonality in fruit tree phenology. Ecology and Evolution. 2013;3(9):3141–51.

38. Obuya JA, Onyango HO, Olokotum M, Zepernick B, Natwora K, Otieno D, et al. Socioeconomic consequences of cyanobacterial harmful algal blooms in small-scale fishing communities of Winam Gulf, Lake Victoria. Journal of Great Lakes Research. 2023 Sep 30;102236.

39. Sitoki L, Kurmayer R, Rott E. Spatial variation of phytoplankton composition, biovolume, and resulting microcystin concentrations in the Nyanza Gulf (Lake Victoria, Kenya). Hydrobiologia. 2012;691(1):109–22.

40. Berry MA, Davis TW, Cory RM, Duhaime MB, Johengen TH, Kling GW, et al. Cyanobacterial harmful algal blooms are a biological disturbance to Western Lake Erie bacterial communities. Environ Microbiol. 20170216th ed. 2017 Mar;19(3):1149–62.

41. Boegehold AG, Burtner AM, Camilleri AC, Carter G, DenUyl P, Fanslow D, et al. Routine monitoring of western Lake Erie to track water quality changes associated with

cyanobacterial harmful algal blooms. Earth System Science Data. 2023 Aug 25;15(8):3853–68.

42. McKindles K, Frenken T, McKay RML, Bullerjahn GS. Binational Efforts Addressing Cyanobacterial Harmful Algal Blooms in the Great Lakes. In: Contaminants of the Great Lakes [Internet]. Springer, Cham; 2020. p. 109–33. Available from: https://link.springer.com/chapter/10.1007/698_2020_513

43. Munawar M, Munawar IF. A lakewide study of phytoplankton biomass and its species composition in lake erie, April–December 1970. J Fish Res Bd Can. 1976 Mar;33(3):581–600.

44. O'Donnell DR, Briland R, Budnik RR, Ludsin SA, Hood JM. Trends in Lake Erie phytoplankton biomass and community structure during a 20-year period of rapid environmental change. Journal of Great Lakes Research. 2023 Jun;49(3):672–84.

45. Munawar M, Munawar IF, Dermott R, Niblock H, Carou S. Is Lake Erie a resilient ecosystem? Aquatic Ecosystem Health & Management. 2002 Mar 1;5(1):79–93.

46. Ludsin SA, Kershner MW, Blocksom KA, Knight RL, Stein RA. Life after death in Lake Erie: Nutrient controls drive fish species richness, rehabilitation. ecological applications. 2001;11(3):731–46.

47. Brittain SM; W Jim; Babcock Jackson, Lisa; Carmichael, Wayne W; Rinehart, Kenneth L; Culver, David A. Isolation and characterization of microcystins, cyclic heptapeptide hepatotoxins from a Lake Erie strain of *Microcystis aeruginosa*. 2000;26:241–9.

48. LeFeuvre DrAR. The Great Lakes Water Quality Agreement: An exercise in co-operation and commitment. Canadian Water Resources Journal. 1991 Jan;16(3):261–5.

49. Makarewicz JC. Phytoplankton as indicators of environmental health. SIL Proceedings, 1922-2010. 1993 Sep 1;25(1):363–5.

50. Makarewicz JC, Bertram P. Evidence for the Restoration of the Lake Erie Ecosystem: Water quality, oxygen levels, and pelagic function appear to be improving. BioScience. 1991 Apr 1;41(4):216–23.

51. Nicholls KH, Hopkins GJ, Standke SJ. Reduced chlorophyll to phosphorus ratios in nearshore Great Lakes waters coincide with the establishment of dreissenid mussels. Can J Fish Aquat Sci. 1999 Jan;56(1):153–61.

52. Baker DB, Confesor R, Ewing DE, Johnson LT, Kramer JW, Merryfield BJ. Phosphorus loading to Lake Erie from the Maumee, Sandusky and Cuyahoga rivers: The importance of bioavailability. Journal of Great Lakes Research. 2014 Sep 1;40(3):502–17.

53. Newell SE, Davis TW, Johengen TH, Gossiaux D, Burtner A, Palladino D, et al. Reduced forms of nitrogen are a driver of non-nitrogen-fixing harmful cyanobacterial blooms and toxicity in Lake Erie. Harmful Algae. 20190102nd ed. 2019 Jan;81:86–93.

54. Millar N, Doll J, Roberston P. Management of nitrogen fertilizer to reduce nitrous oxide emissions from field crops [Internet]. Michigan State University; 2014 Nov. (Climate Change and Agriculture Fact Sheet Series). Available from:

https://www.canr.msu.edu/resources/management_of_nitrogen_fertilizer_to_reduce_nitrous_oxide_emissions_from_fi

55. Kane DD, Conroy JD, Peter Richards R, Baker DB, Culver DA. Re-eutrophication of Lake Erie: Correlations between tributary nutrient loads and phytoplankton biomass. Journal of Great Lakes Research. 2014 Sep 1;40(3):496–501.

56. Monchamp ME, Pick FR, Beisner BE, Maranger R. Nitrogen forms influence microcystin concentration and composition via changes in cyanobacterial community structure. PLOS ONE. 2014 Jan 10;9(1):e85573.

57. Chaffin JD, Bridgeman TB, Bade DL. Nitrogen constrains the growth of late summer cyanobacterial blooms in Lake Erie. AiM. 2013;03(06):16–26.

58. Chaffin JD, Davis TW, Smith DJ, Baer MM, Dick GJ. Interactions between nitrogen form, loading rate, and light intensity on *Microcystis* and *Planktothrix* growth and microcystin production. Harmful Algae. 2018 Mar 1;73:84–97.

59. Davis TW, Bullerjahn GS, Tuttle T, McKay RM, Watson SB. Effects of increasing nitrogen and phosphorus concentrations on phytoplankton community growth and toxicity during planktothrix blooms in Sandusky Bay, Lake Erie. Environ Sci Technol. 2015 Jun 16;49(12):7197–207.

60. Belisle BS, Steffen MM, Pound HL, Watson SB, DeBruyn JM, Bourbonniere RA, et al. Urea in Lake Erie: Organic nutrient sources as potentially important drivers of phytoplankton biomass. Journal of Great Lakes Research. 2016 Jun 1;42(3):599–607.

61. Den Uyl PA, Kiledal EA, Errera RM, Chaganti SR, Godwin CM, Raymond HA, et al. Genomic identification and characterization of saxitoxin producing cyanobacteria in western Lake Erie harmful algal blooms. Environ Sci Technol. 2025 Apr 10;

62. Yancey CE, Mathiesen O, Dick GJ. Transcriptionally active nitrogen fixation and biosynthesis of diverse secondary metabolites by Dolichospermum and *Aphanizomenon*-like cyanobacteria in western Lake Erie *Microcystis* blooms. Harmful Algae. 2023 May 1;124:102408.

63. Michalak AM, Anderson EJ, Beletsky D, Boland S, Bosch NS, Bridgeman TB, et al. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proceedings of the National Academy of Sciences. 2013 Apr 16;110(16):6448–52.

64. Ho JC, Michalak AM, Pahlevan N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature. 2019 Oct;574(7780):667–70.

65. Paerl HW, Huisman J. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. Environmental Microbiology Reports. 2009;1(1):27–37.

66. Steffen MM, Davis TW, McKay RML, Bullerjahn GS, Krausfeldt LE, Stough JMA, et al. Ecophysiological Examination of the Lake Erie *Microcystis* Bloom in 2014: Linkages between biology and the water supply shutdown of Toledo, OH. Environmental Science & Technology. 2017 Jun 20;51(12):6745–55.

67. Davis TW, Stumpf R, Bullerjahn GS, McKay RML, Chaffin JD, Bridgeman TB, et al. Science meets policy: A framework for determining impairment designation criteria for large waterbodies affected by cyanobacterial harmful algal blooms. Harmful Algae. 2019 Jan 1;81:59–64.

68. Zepernick BN, Boegehold AG, Kiledal EA, Chase EE, Hart LN, Houghton KA, et al. Diel metatranscriptomes capture cyanobacteria-dominated Lake Erie community response to episodic events. Microbiol Resour Announc. 13(11):e00659-24.

69. Lawrence TJ, Achieng AO, Chavula G, Haninga Haambiya L, Iteba J, Kayanda R, et al. Future success and ways forward for scientific approaches on the African Great Lakes. Journal of Great Lakes Research. 2023 Dec 1;49(6):102242.

70. Zhao X, Liu Y, Guo YM, Xu C, Chen L, Codd GA, et al. Meta-analysis reveals cyanotoxins risk across African inland waters. Journal of Hazardous Materials. 2023 Jun 5;451:131160.

71. Njagi DM, Routh J, Odhiambo M, Luo C, Basapuram LG, Olago D, et al. A century of human-induced environmental changes and the combined roles of nutrients and land use in Lake Victoria catchment on eutrophication. Science of The Total Environment. 2022 Aug 20;835:155425.

72. Frank TH, Cornelissen IJM, Vijverberg J, Nagelkerke LAJ. Spatial and seasonal variation in the phytoplankton community of Lake Victoria's Mwanza Gulf, compared to northern parts of the lake. Journal of Great Lakes Research. 2023 Apr 1;49(2):453–62.

73. Calamari D, Akech MO, Ochumba PBO. Pollution of Winam Gulf, Lake Victoria, Kenya: A case study for preliminary risk assessment. Lakes & Reservoirs: Science, Policy and Management for Sustainable Use. 1995;1(2):89–106.

74. Roegner A, Sitoki L, Weirich C, Corman J, Owage D, Umami M, et al. Harmful algal blooms threaten the health of peri-urban fisher communities: a case study in Kisumu Bay, Lake Victoria, Kenya. Expo Health. 2020 Dec;12(4):835–48.

75. Gikuma-Njuru P, Hecky RE. Nutrient concentrations in Nyanza Gulf, Lake Victoria, Kenya: light limits algal demand and abundance. Hydrobiologia. 2005 Feb 1;534(1):131–40.

76. Brown KM, Barker KB, Wagner RS, Ward CS, Sitoki L, Njiru J, et al. Bacterial community and cyanotoxin gene distribution of the Winam Gulf, Lake Victoria, Kenya. Environmental Microbiology Reports. 2024 Jun;16(3):e13297.

77. Suba VO, Khan MS, Miruka J, Mwalati L, Njoroge B, Oluoch E, et al. Seasonal phytoplankton ecosystem dynamics in response to environmental variables in Winam Gulf, Lake Victoria, Kenya. Aquat Sci. 2025 Jan;87(1):4.

78. Kundu R, Aura CM, Nyamweya C, Agembe S, Sitoki L, Lung'ayia HBO, et al. Changes in pollution indicators in Lake Victoria, Kenya and their implications for lake and catchment management. Lakes & Reservoirs: Science, Policy and Management for Sustainable Use. 2017;22(3):199–214.

79. Plisnier PD, Kayanda R, MacIntyre S, Obiero K, Okello W, Vodacek A, et al. Need for harmonized long-term multi-lake monitoring of African Great Lakes. Journal of Great Lakes Research. 2023 Dec 1;49(6):101988.

80. Igwaran A, Kayode AJ, Moloantoa KM, Khetsha ZP, Unuofin JO. Cyanobacteria Harmful Algae Blooms: Causes, impacts, and risk management. Water Air Soil Pollut. 2024 Jan 1;235(1):1–26.

81. Paerl HW, Otten TG. Harmful Cyanobacterial Blooms: Causes, consequences, and controls. Microb Ecol. 2013 May 1;65(4):995–1010.

82. Durán-Vinet B, Araya-Castro K, Chao T, Wood S, Gallardo V, Godoy K, et al. Potential applications of CRISPR/Cas for next-generation biomonitoring of harmful algae blooms: A review. Harmful Algae. 2021 Mar 1;103:102027.

83. Trevino-Garrison I, DeMent J, Ahmed F, Haines-Lieber P, Langer T, Ménager H, et al. Human illnesses and animal deaths associated with freshwater harmful algal blooms—Kansas. Toxins. 2015 Jan 30;7(2):353–66.

84. Lim MH, Tay HSM, Devotta DA, Mowe MAD, Mitrovic SM. Risk management of cyanotoxins in Singapore. Journal of Water Resource and Protection. 2020 Jun 4;12(6):512–25.

85. Wang H, Xu C, Liu Y, Jeppesen E, Svenning JC, Wu J, et al. From unusual suspect to serial killer: Cyanotoxins boosted by climate change may jeopardize megafauna. Innovation. 2021 May 28;2(2).

86. Backer LC, Landsberg JH, Miller M, Keel K, Taylor TK. Canine Cyanotoxin Poisonings in the United States (1920s–2012): Review of suspected and confirmed cases from three data sources. Toxins. 2013 Sep;5(9):1597–628.

87. Pal M, Yesankar PJ, Dwivedi A, Qureshi A. Biotic control of harmful algal blooms (HABs): A brief review. Journal of Environmental Management. 2020 Aug 15;268:110687.

88. Díez-Quijada Jiménez L, Guzmán-Guillén R, Cătunescu GM, Campos A, Vasconcelos V, Jos Á, et al. A new method for the simultaneous determination of cyanotoxins (Microcystins and Cylindrospermopsin) in mussels using SPE-UPLC-MS/MS. Environmental Research. 2020 Jun 1;185:109284.

89. Schaefer AM, Yrastorza L, Stockley N, Harvey K, Harris N, Grady R, et al. Exposure to microcystin among coastal residents during a cyanobacteria bloom in Florida. Harmful Algae. 2020 Feb 1;92:101769.

90. Pouria S, de Andrade A, Barbosa J, Cavalcanti R, Barreto V, Ward C, et al. Fatal microcystin intoxication in hemodialysis unit in Caruaru, Brazil. The Lancet. 1998 Jul 4;352(9121):21–6.

91. Hernandez BY, Zhu X, Nagata M, Loo L, Chan O, Wong LL. Cyanotoxin exposure and hepatocellular carcinoma. Toxicology. 2023 Mar 15;487:153470.

92. Dhillon J, Parker ER. Climate change, harmful algal blooms, and cutaneous disease. JAAD Reviews. 2025 Mar 4.

93. Walsh JJ, Lenes JM, Weisberg RH, Zheng L, Hu C, Fanning KA, et al. More surprises in the global greenhouse: Human health impacts from recent toxic marine aerosol formations, due to centennial alterations of world-wide coastal food webs. Marine Pollution Bulletin. 2017 Mar 15;116(1):9–40.

94. Hofbauer WK. Toxic or otherwise harmful algae and the built environment. Toxins. 2021 Jul;13(7):465.

95. Asthma and Allergy Foundation of America. Asthma Capitals: The most challenging places to live with asthma. 2018 p. 40. Available from: https://www.aafa.org/wp-content/uploads/2022/10/aafa-2018-asthma-capitals-report.pdf

96. Dodds WK, Bouska WW, Eitzmann JL, Pilger TJ, Pitts KL, Riley AJ, et al. Eutrophication of U.S. freshwaters: analysis of potential economic damages. Environ Sci Technol. 2009 Jan 1;43(1):12–9.

97. Alliance for the Great Lakes. Western Lake Erie Basin Drinking Water Systems: Harmful algal bloom cost of intervention [Internet]. Chicago, IL; 2022 May p. 5. Available from: https://greatlakes.org/wp-content/uploads/2022/05/FINAL-COI-Report-051622.pdf

98. Zhang X jian, Chen C, Ding J qing, Hou A, Li Y, Niu Z bin, et al. The 2007 water crisis in Wuxi, China: Analysis of the origin. Journal of Hazardous Materials. 2010 Oct 15;182(1):130–5.

99. Jegede FE, Oyeyi TI, Abdulrahman SA, Mbah HAkwen, Badru T, Agbakwuru C, et al. Effect of HIV and malaria parasites co-infection on immune-hematological profiles among patients attending anti-retroviral treatment (ART) clinic in Infectious Disease Hospital Kano, Nigeria. PLoS One. 2017 Mar 27;12(3):e0174233.

100. Oyieko J, Copeland NK, Otieno S, Kifude C, Ocholla S, Hutter J, et al. Longitudinal and cross-sectional analyses of asymptomatic HIV-1/malaria co-infection in Kisumu County, Kenya. Am J Trop Med Hyg. 2023 Jan 11;108(1):85–92.

101. Paerl HW, Fulton RS, Moisander PH, Dyble J. Harmful Freshwater Algal Blooms, With an Emphasis on Cyanobacteria. Scientific World Journal. 2001 Apr 4;1:76–113.

102. Sheik CS, Natwora KE, Alexson EE, Callaghan JD, Sailer A, Schreiner KM, et al. *Dolichospermum* blooms in Lake Superior: DNA-based approach provides insight to the past, present and future of blooms. Journal of Great Lakes Research. 2022 Oct 1;48(5):1191–205.

103. Reynolds CS, Rogers DA. Seasonal variations in the vertical distribution and buoyancy of *Microcystis aeruginosa* Kütz. emend. Elenkin in Rostherne Mere, England. Hydrobiologia. 1976 Jan 1;48(1):17–23.

104. Ibelings BW, Mur LR, Walsby AE. Diurnal changes in buoyancy and vertical distribution in populations of *Microcystis* in two shallow lakes. Journal of Plankton Research. 1991 Jan 1;13(2):419–36.

105. Komárek J, Komárková J. Review of the European *Microcystis* morphospecies (Cyanoprokaryotes) from nature. Czech Phycology. 2002;1–24.

106. Zurawell RW, Chen ,Huirong, Burke ,Janice M., and Prepas EE. Hepatotoxic Cyanobacteria: A review of the biological importance of microcystins in freshwater environments. Journal of Toxicology and Environmental Health, Part B. 2005 Dec 14;8(1):1–37.

107. Vézie C, Rapala J, Vaitomaa J, Seitsonen J, Sivonen K. Effect of Nitrogen and Phosphorus on Growth of Toxic and Nontoxic *Microcystis* strains and on intracellular microcystin concentrations. Microb Ecol. 2002 May;43(4):443–54.

108. Davis TW, Berry DL, Boyer GL, Gobler CJ. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. Harmful Algae. 2009 Jun 1;8(5):715–25.

109. Humbert JF, Barbe V, Latifi A, Gugger M, Calteau A, Coursin T, et al. A tribute to disorder in the genome of the bloom-forming freshwater cyanobacterium *Microcystis aeruginosa*. PLOS ONE. 2013 Aug 12;8(8):e70747.

110. Pérez-Carrascal OM, Terrat Y, Giani A, Fortin N, Greer CW, Tromas N, et al. Coherence of *Microcystis* species revealed through population genomics. Isme j. 2019 Dec;13(12):2887–900.

111. Gobler CJ, Burkholder JM, Davis TW, Harke MJ, Johengen T, Stow CA, et al. The dual role of nitrogen supply in controlling the growth and toxicity of cyanobacterial blooms. Harmful Algae. 2016 Apr 1;54:87–97.

112. Welker M, Maršálek B, Šejnohová L, von Döhren H. Detection and identification of oligopeptides in *Microcystis* (cyanobacteria) colonies: Toward an understanding of metabolic diversity. Peptides. 2006 Sep 1;27(9):2090–103.

113. Xiao Y, Gan N, Liu J, Zheng L, Song L. Heterogeneity of buoyancy in response to light between two buoyant types of cyanobacterium *Microcystis*. Hydrobiologia. 2012 Jan 1;679(1):297–311.

114. Sandrini G, Matthijs HCP, Verspagen JMH, Muyzer G, Huisman J. Genetic diversity of inorganic carbon uptake systems causes variation in CO2 response of the cyanobacterium *Microcystis*. The ISME Journal. 2014 Mar 1;8(3):589–600.

115. Vanderploeg HA, Liebig JR, Carmichael WW, Agy MA, Johengen TH, Fahnenstiel GL, et al. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. Can J Fish Aquat Sci. 2001 Jun;58(6):1208–21.

116. Saxton MA, Arnold RJ, Bourbonniere RA, McKay RM, Wilhelm SW. Plasticity of total and intracellular phosphorus quotas in *Microcystis aeruginosa* cultures and Lake Erie algal assemblages. Front Microbiol. 2012 Jan 18;3.

117. Walsby AE, Hayes PK, Boje R, Stal LJ. The selective advantage of buoyancy provided by gas vesicles for planktonic cyanobacteria in the Baltic Sea. New Phytologist. 1997;136(3):407–17.

118. Wu H, Wu X, Yang T, Wang C, Tian C, Xiao B, et al. Feedback regulation of surface scum formation and persistence by self-shading of *Microcystis* colonies: Numerical simulations and laboratory experiments. Water Research. 2021 Apr 15;194:116908.

119. Sandrini G, Cunsolo S, Schuurmans JM, Matthijs HCP, Huisman J. Changes in gene expression, cell physiology and toxicity of the harmful cyanobacterium *Microcystis aeruginosa* at elevated CO2. Front Microbiol. 2015 May 5;6.

120. Wilhelm SW, Bullerjahn GS, McKay RML. The complicated and confusing ecology of *Microcystis* blooms. Moran MA, editor. mBio. 2020 Jun 30;11(3):e00529-20.

121. Schindler DW. Effects of acid rain on freshwater ecosystems. Science. 1988 Jan 8;239(4836):149–57.

122. Jackrel SL, White JD, Evans JT, Buffin K, Hayden K, Sarnelle O, et al. Genome evolution and host-microbiome shifts correspond with intraspecific niche divergence within harmful algal bloom-forming *Microcystis aeruginosa*. Molecular Ecology. 2019 Sep;28(17):3994–4011.

123. Kromkamp J, van den Heuvel ,Andien, and Mur LR. Phosphorus uptake and photosynthesis by phosphate-limited cultures of the cyanobacterium *Microcystis aeruginosa*. British Phycological Journal. 1989 Jan 1;24(4):347–55.

124. Carr NG, Whitton BA. The biology of cyanobacteria. University of California Press; 1982. 708 p.

125. Paerl H. Mitigating harmful cyanobacterial blooms in a human- and climatically-impacted world. Life. 2014 Dec 15;4(4):988–1012.

126. Paerl HW, Scott JT, McCarthy MJ, Newell SE, Gardner WS, Havens KE, et al. It Takes Two to Tango: When and where dual nutrient (n & p) reductions are needed to protect lakes and downstream ecosystems. Environ Sci Technol. 2016 Oct 18;50(20):10805–13.

127. Davis TW, Harke MJ, Marcoval MA, Goleski J, Orano-Dawson C, Berry DL, et al. Effects of nitrogenous compounds and phosphorus on the growth of toxic and non-toxic strains of *Microcystis* during cyanobacterial blooms. Aquatic Microbial Ecology. 2010;61(2):149–62.

128. Kardinaal WEA, Tonk L, Janse I, Hol S, Slot P, Huisman J, et al. Competition for light between toxic and nontoxic strains of the harmful cyanobacterium *Microcystis*. Applied and Environmental Microbiology. 2007 May;73(9):2939–46.

129. Wiedner C, Visser PM, Fastner J, Metcalf JS, Codd GA, Mur LR. Effects of light on the microcystin content of *Microcystis* Strain PCC 7806. Applied and Environmental Microbiology. 2003 Mar;69(3):1475–81.

130. Zilliges Y, Kehr JC, Meissner S, Ishida K, Mikkat S, Hagemann M, et al. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of *Microcystis* under oxidative stress conditions. PLoS One. 2011 Mar 18;6(3):e17615.

131. Gan N, Xiao Y, Zhu L, Wu Z, Liu J, Hu C, et al. The role of microcystins in maintaining colonies of bloom-forming *Microcystis* spp. Environmental Microbiology. 2012 Mar;14(3):730–42.

132. Ger KA, Faassen EJ, Pennino MG, Lürling M. Effect of the toxin (microcystin) content of *Microcystis* on copepod grazing. Harmful Algae. 2016 Feb 1;52:34–45.

133. Rantala A, Rizzi E, Castiglioni B, De Bellis G, Sivonen K. Identification of hepatotoxin-producing cyanobacteria by DNA-chip. Environmental Microbiology. 2008;10(3):653–64.

134. Wilson AE, Sarnelle O, Tillmanns AR. Effects of cyanobacterial toxicity and morphology on the population growth of freshwater zooplankton: Meta-analyses of laboratory experiments. Limnology and Oceanography. 2006;51(4):1915–24.

135. Yang Z, Kong F, Shi X, Cao H. Morphological response of *Microcystis aeruginosa* to grazing by different sorts of zooplankton. Hydrobiologia. 2006 Jun 1;563(1):225–30.

136. Xiao M, Li M, Reynolds CS. Colony formation in the cyanobacterium *Microcystis*. Biological Reviews. 2018;93(3):1399–420.

137. White JD, Kaul RB, Knoll LB, Wilson AE, Sarnelle O. Large variation in vulnerability to grazing within a population of the colonial phytoplankter, *Microcystis aeruginosa*. Limnology and Oceanography. 2011;56(5):1714–24.

138. Xiao M, Li M, Duan P, Qu Z, Wu H. Insights into the relationship between colony formation and extracellular polymeric substances (EPS) composition of the cyanobacterium *Microcystis* spp. Harmful Algae. 2019 Mar 1;83:34–41.

139. Torres M de A, Dax A, Grand I, vom Berg C, Pinto E, Janssen EML. Lethal and behavioral effects of semi-purified microcystins, Micropeptin and apolar compounds from cyanobacteria on freshwater microcrustacean *Thamnocephalus platyurus*. Aquatic Toxicology. 2024 Aug 1;273:106983.

140. Rohrlack T, Christoffersen K, Kaebernick M, Neilan BA. Cyanobacterial protease inhibitor Microviridin J causes a lethal molting disruption in *Daphnia pulicaria*. Applied and Environmental Microbiology. 2004 Aug;70(8):5047–50.

141. Gobler CJ, Davis TW, Coyne KJ, Boyer GL. Interactive influences of nutrient loading, zooplankton grazing, and microcystin synthetase gene expression on cyanobacterial bloom dynamics in a eutrophic New York lake. Harmful Algae. 2007 Jan 1;6(1):119–33.

142. Cai H, Jiang H, Krumholz LR, Yang Z. Bacterial community composition of size-fractioned aggregates within the phycosphere of cyanobacterial blooms in a eutrophic freshwater lake. PLoS One. 20140821st ed. 2014;9(8):e102879.

143. Paerl HW, Millie DF. Physiological ecology of toxic aquatic cyanobacteria. Phycologia. 1996 Nov;35(sup6):160–7.

144. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature. 2005 Nov;438(7064):90–3.

145. Cook KV, Li C, Cai H, Krumholz LR, Hambright KD, Paerl HW, et al. The global *Microcystis* interactome. Limnology & Oceanography. 2020 Jan;65(S1).

146. Li W, Baliu-Rodriguez D, Premathilaka SH, Thenuwara SI, Kimbrel JA, Samo TJ, et al. Microbiome processing of organic nitrogen input supports growth and cyanotoxin

production of *Microcystis aeruginosa* cultures. The ISME Journal. 2024 Jan 1;18(1):wrae082.

147. Wang Y, Wu M, Yu J, Zhang J, Zhang R, Zhang L, et al. Differences in growth, pigment composition and photosynthetic rates of two phenotypes *Microcystis aeruginosa* strains under high and low iron conditions. Biochemical Systematics and Ecology. 2014 Aug 1;55:112–7.

148. Smith DJ, Berry MA, Cory RM, Johengen TH, Kling GW, Davis TW, et al. Heterotrophic bacteria dominate catalase expression during *Microcystis* blooms. Applied and Environmental Microbiology. 2022 Jul 5;88(14):e02544-21.

149. Sommaruga R, Chen Y, Liu Z. Multiple strategies of bloom-forming *Microcystis* to minimize damage by solar ultraviolet radiation in surface waters. Microb Ecol. 2009 May;57(4):667–74.

150. Li P, Cai Y, Shi L, Geng L, Xing P, Yu Y, et al. microbial degradation and preliminary chemical characterization of *Microcystis* exopolysaccharides from a cyanobacterial water bloom of Lake Taihu. International Review of Hydrobiology. 2009 Dec;94(6):645–55.

151. Stewart WDP. Nitrogen fixation by photosynthetic microorganisms. Annual Review of Microbiology. 1973 Oct 1;27(Volume 27,):283–316.

152. Lima ST, Fallon TR, Cordoza JL, Chekan JR, Delbaje E, Hopiavuori AR, et al. Biosynthesis of guanitoxin enables global environmental detection in freshwater cyanobacteria. J Am Chem Soc. 2022 Jun 1;144(21):9372–9.

153. Slim K, Fadel A, Atoui A, Lemaire BJ, Vinçon-Leite B, Tassin B. Global warming as a driving factor for cyanobacterial blooms in Lake Karaoun, Lebanon. Desalination and Water Treatment. 2014 Feb 1;52(10):2094–101.

154. Huber V, Wagner C, Gerten D, Adrian R. To bloom or not to bloom: contrasting responses of cyanobacteria to recent heat waves explained by critical thresholds of abiotic drivers. Oecologia. 2012 May 1;169(1):245–56.

155. Moisander PH, E. McClinton I, Paerl HW. Salinity effects on growth, photosynthetic parameters, and nitrogenase activity in estuarine planktonic cyanobacteria. Microbial Ecology. 2002;43(4):432–42.

156. Callieri C, Bertoni R, Contesini M, Bertoni F. Lake level fluctuations boost toxic cyanobacterial "Oligotrophic Blooms." PLOS ONE. 2014 Oct 8;9(10):e109526.

157. Driscoll CB, Meyer KA, Šulčius S, Brown NM, Dick GJ, Cao H, et al. A closely-related clade of globally distributed bloom-forming cyanobacteria within the *Nostocales*. Harmful Algae. 2018 Jul 1;77:93–107.

158. Wacklin P, Hoffmann L, Komárek J. Nomenclatural validation of the genetically revised cyanobacterial genus *Dolichospermum* (RALFS ex BORNET et FLAHAULT) comb. nova. Fottea. 2009 Jan 1;9(1):59–64.

159. Komárek J. Modern taxonomic revision of planktic nostocacean cyanobacteria: a short review of genera. Hydrobiologia. 2010 Feb 1;639(1):231–43.

160. Rajaniemi P, Hrouzek P, Kaštovská K, Willame R, Rantala A, Hoffmann L, et al. Phylogenetic and morphological evaluation of the genera *Anabaena, Aphanizomenon, Trichormus* and *Nostoc* (Nostocales, Cyanobacteria). International Journal of Systematic and Evolutionary Microbiology. 2005;55(1):11–26.

161. Halinen K, Fewer DP, Sihvonen LM, Lyra C, Eronen E, Sivonen K. Genetic diversity in strains of the genus *Anabaena* isolated from planktonic and benthic habitats of the Gulf of Finland (Baltic Sea). FEMS Microbiology Ecology. 2008;64(2):199–208.

162. Komárek J, Anagnostidis K. Modern approach to the classification system of Cyanophytes 4 - Nostocales. Algological Studies/Archiv für Hydrobiologie, Supplement Volumes. 1989 Jan 1;247–345.

163. Haselkorn R. Heterocysts. Annual review of plant biology. 1978 Jun 1;29(Volume 29, 1978):319–44.

164. Meeks JC, Elhai J. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiology and Molecular Biology Reviews. 2002 Mar;66(1):94–121.

165. Kumar K, Mella-Herrera RA, Golden JW. Cyanobacterial heterocysts. Cold Spring Harb Perspect Biol. 2010 Apr 1;2(4):a000315.

166. Mulligan ME, Haselkorn R. Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120. The *nifB-fdxN-nifS-nifU* operon. J Biol Chem. 1989 Nov 15;264(32):19200–7.

167. C M. Algal size. The physiological ecology of phytoplankton. 1980;433–63.

168. Wood SA, Rueckert A, Hamilton DP, Cary SC, Dietrich DR. Switching toxin production on and off: intermittent microcystin synthesis in a *Microcystis* bloom. Environmental Microbiology Reports. 2011;3(1):118–24.

169. Rapala J, Sivonen K, Lyra C, Niemelä SI. Variation of microcystins, cyanobacterial hepatotoxins, in *Anabaena* spp. as a function of growth stimuli. Appl Environ Microbiol. 1997 Jun;63(6):2206–12.

170. Liu Z, Xie F, Wang L, Yu L, Shi X, Yang Z, et al. Response of heterocyst differentiation of *Dolichospermum* to different forms of nitrogen deficiency. Harmful Algae. 2024 Nov 1;139:102742.

171. Head R m., Jones R i., Bailey-Watts A e. Vertical movements by planktonic cyanobacteria and the translocation of phosphorus: implications for lake restoration. Aquatic Conservation: Marine and Freshwater Ecosystems. 1999;9(1):111–20.

172. Cook CM, Vardaka E, Lanaras T. Toxic Cyanobacteria in Greek Freshwaters, 1987—2000: Occurrence, toxicity, and impacts in the Mediterranean region. Acta hydrochimica et hydrobiologica. 2004;32(2):107–24.

173. Kravchuk ES, Ivanova EA, Gladyshev MI. Seasonal dynamics of akinetes of *Anabaena flos-aquae* in bottom sediments and water column of small Siberian reservoir. Aquat Ecol. 2006 Sep 1;40(3):325–36.

174. Tsujimura S, Okubo T. Development of *Anabaena* blooms in a small reservoir with dense sediment akinete population, with special reference to temperature and irradiance. Journal of Plankton Research. 2003 Sep 1;25(9):1059–67.

175. Rengefors K, Gustafsson S, Ståhl-Delbanco A. Factors regulating the recruitment of cyanobacterial and eukaryotic phytoplankton from littoral and profundal sediments. Aquatic Microbial Ecology. 2004 Sep 14;36(3):213–26.

176. Deng D, Meng H, Ma Y, Guo Y, Wang Z, He H, et al. The cumulative impact of temperature and nitrogen availability on the potential nitrogen fixation and extracellular polymeric substances secretion by *Dolichospermum*. Harmful Algae. 2024 May 3;102633.

177. Reinl KL, Harris TD, North RL, Almela P, Berger SA, Bizic M, et al. Blooms also like it cold. Limnology and Oceanography Letters. 2023;8(4):546–64.

178. Tan X, Duan Z, Li N, Shen H, Zhen S, Xiao M. Photosynthetic activity of bloom-forming cyanobacteria in winter and their relationship with light intensity. Fresenius Environmental Bulletin. 2015;

179. Pereira SB, Mota R, Vieira CP, Vieira J, Tamagnini P. Phylum-wide analysis of genes/proteins related to the last steps of assembly and export of extracellular polymeric substances (EPS) in cyanobacteria. Sci Rep. 2015 Oct 6;5(1):14835.

180. Compaoré J, Stal LJ. Effect of temperature on the sensitivity of nitrogenase to oxygen in two heterocystous cyanobacteria. Journal of Phycology. 2010;46(6):1172–9.

181. Fuster M, Ruiz T, Lamarque A, Coulon M, Legrand B, Sabart M, et al. Cyanosphere dynamic during *Dolichospermum* Bloom: potential roles in cyanobacterial proliferation. Microb Ecol. 2024 Dec 1;87(1):1–10.

182. Yan Z, Liu Z, Jia Z, Song C, Cao X, Zhou Y. Metabolites of extracellular organic matter from *Microcystis* and *Dolichospermum* drive distinct modes of carbon, nitrogen, and phosphorus recycling. Science of The Total Environment. 2023 Mar 20;865:161124.

183. Capelli C, Ballot A, Cerasino L, Papini A, Salmaso N. Biogeography of bloom-forming microcystin producing and non-toxigenic populations of *Dolichospermum lemmermannii* (Cyanobacteria). Harmful Algae. 2017 Jul 1;67:1–12.

184. Liu Y, Xu Y, Xiao P, Pan Q, Yu G, Li R. Genetic analysis on *Dolichospermum* (Cyanobacteria; sensu *Anabaena*) populations based on the culture-independent clone libraries revealed the dominant genotypes existing in Lake Taihu, China. Harmful Algae. 2014 Jan 1;31:76–81.

185. Beversdorf LJ, Miller TR, McMahon KD. The role of nitrogen fixation in cyanobacterial bloom toxicity in a temperate, eutrophic lake. Gilbert JA, editor. PLoS ONE. 2013 Feb 6;8(2):e56103.

186. Liu Y, Zhu G, Fan Y, Guo C. Successional conditions of *Dolichospermum* and *Microcystis* in Taihu Lake, China. J Ocean Limnol. 2024 Nov 1;42(6):1777–88.

187. Paerl H. Nutrient and other environmental controls of harmful cyanobacterial blooms along the freshwater-marine continuum. Adv Exp Med Biol. 2008;619:217–37.

188. Schindler DW, Hecky RE, Findlay DL, Stainton MP, Parker BR, Paterson MJ, et al. Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences. 2008 Aug 12;105(32):11254–8.

189. Den Uyl P. Genetic characterization of saxitoxin-producing cyanobacteria associated with western Lake Erie harmful algal blooms [Internet]. University of Michigan; 2024. Available from: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/193158/MS_Thesis_DenUyl_DeepBlue_FinalSignatures.pdf?sequence=1&isAllowed=y

190. Chaffin JD, Mishra S, Kane DD, Bade DL, Stanislawczyk K, Slodysko KN, et al. Cyanobacterial blooms in the central basin of Lake Erie: Potentials for cyanotoxins and environmental drivers. Journal of Great Lakes Research. 2019 Apr 1;45(2):277–89.

191. Nauman C, Stanislawczyk K, Reitz LA, Chaffin JD. The spatiotemporal distribution of potential saxitoxin-producing cyanobacteria in western Lake Erie. Journal of Great Lakes Research. 2024 Jun 1;50(3):102342.

192. Laiveling A, Nauman C, Stanislawczyk K, Bair HB, Kane DD, Chaffin JD. Potamoplankton of the Maumee River during 2018 and 2019: The relationship between cyanobacterial toxins and environmental factors. Journal of Great Lakes Research. 2022 Dec 1;48(6):1587–98.

193. Zdouc MM, Blin K, Louwen NLL, Navarro J, Loureiro C, Bader CD, et al. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration. Nucleic Acids Research. 2025 Jan 6;53(D1):D678–90.

194. Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, Mavrommatis K, et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. Cell. 2014 Jul 17;158(2):412–21.

195. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research. 2023 Jul 5;51(W1):W46–50.

196. Hoskisson PA, Seipke RF. Cryptic or Silent? The known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. mBio. 2020 Oct 20;11(5):10.1128/mbio.02642-20.

197. Chaffin JD, Westrick JA, Reitz LA, Bridgeman TB. Microcystin congeners in Lake Erie follow the seasonal pattern of nitrogen availability. Harmful Algae. 2023 Aug 1;127:102466.

198. Tillett D, Dittmann E, Erhard M, von Döhren H, Börner T, Neilan BA. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. Chem Biol. 2000 Oct;7(10):753–64.

199. Welker M, Von Döhren H. Cyanobacterial peptides — Nature's own combinatorial biosynthesis. FEMS Microbiology Reviews. 2006 Jul 1;30(4):530–63.

200. E. Chlipala G, Mo S, Orjala J. Chemodiversity in freshwater and terrestrial cyanobacteria - A source for drug discovery. Current Drug Targets. 2011 Oct 1;12(11):1654–73.

201. Tooming-Klunderud A, Rohrlack T, Shalchian-Tabrizi K, Kristensen T, Jakobsen KS. Structural analysis of a non-ribosomal halogenated cyclic peptide and its putative operon from *Microcystis*: implications for evolution of cyanopeptolins. Microbiology (Reading). 2007 May;153(Pt 5):1382–93.

202. Ishida K, Christiansen G, Yoshida WY, Kurmayer R, Welker M, Valls N, et al. Biosynthesis and structure of aeruginoside 126A and 126B, cyanobacterial peptide glycosides bearing a 2-carboxy-6-hydroxyoctahydroindole moiety. Chem Biol. 2007 May;14(5):565–76.

203. Eusébio N, Castelo-Branco R, Sousa D, Preto M, D'Agostino P, Gulder TAM, et al. Discovery and heterologous expression of microginins from *Microcystis aeruginosa* LEGE 91341. ACS Synth Biol. 2022 Oct 21;11(10):3493–503.

204. Leikoski N, Fewer DP, Jokela J, Alakoski P, Wahlsten M, Sivonen K. Analysis of an inactive cyanobactin biosynthetic gene cluster leads to discovery of new natural products from strains of the genus *Microcystis*. PLoS One. 2012;7(8):e43002.

205. Portmann C, Blom JF, Gademann K, Jüttner F. Aerucyclamides A and B: Isolation and synthesis of toxic ribosomal heterocyclic peptides from the cyanobacterium *Microcystis aeruginosa* PCC 7806. J Nat Prod. 2008 Jul 1;71(7):1193–6.

206. Wenski SL, Thiengmag S, Helfrich EJN. Complex peptide natural products: Biosynthetic principles, challenges and opportunities for pathway engineering. Synthetic and Systems Biotechnology. 2022 Mar 1;7(1):631–47.

207. Walsh CT, O'Brien RV, Khosla C. Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. Angew Chem Int Ed Engl. 2013 Jul 8;52(28):7098–124.

208. Larsen J. Type III Polyketide biosynthesis in cyanobacteria [Internet]. School of Environmental and Life Sciences, University of Newcastle, Australia; 2023 [cited 2025 Mar 3]. Available from: https://nova.newcastle.edu.au/vital/access/services/Download/uon:50347/ATTACHMENT01

209. Larsen JS, Pearson LA, Neilan BA. Genome mining and evolutionary analysis reveal diverse type iii polyketide synthase pathways in cyanobacteria. Genome Biol Evol. 2021 Mar 19;13(4):evab056.

210. Ke J, Yoshikuni Y. Multi-chassis engineering for heterologous production of microbial natural products. Current Opinion in Biotechnology. 2020 Apr 1;62:88–97.

211. Yancey C. The Biosynthetic repertoire of *Microcystis* spp. in Western Lake Erie harmful algal blooms: Insights into biosynthesis of known and novel secondary metabolites in natural populations using a `multi-omic' approach [Internet] [Thesis]. University of Michigan; 2023 [cited 2025 Apr 9]. Available from: http://deepblue.lib.umich.edu/handle/2027.42/178036

212. May DS, Chen WL, Lantvit DD, Zhang X, Krunic A, Burdette JE, et al. Merocyclophanes C and D from the cultured freshwater cyanobacterium *Nostoc* sp. (UIC 10110). J Nat Prod. 2017 Apr 28;80(4):1073–80.

213. Makower AK, Schuurmans JM, Groth D, Zilliges Y, Matthijs HCP, Dittmann E. Transcriptomics-aided dissection of the intracellular and extracellular roles of microcystin in *Microcystis aeruginosa* PCC 7806. Appl Environ Microbiol. 2015 Jan 15;81(2):544–54.

214. Chang Z, Sitachitta N, Rossi JV, Roberts MA, Flatt PM, Jia J, et al. Biosynthetic pathway and gene cluster analysis of curacin a, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. J Nat Prod. 2004 Aug 1;67(8):1356–67.

215. Kallifidas D, Dhakal D, Chen M, Chen QY, Kokkaliari S, Colon Rosa NA, et al. Biosynthesis of Dolastatin 10 in marine cyanobacteria, a prototype for multiple approved cancer drugs. Org Lett. 2024 Feb 23;26(7):1321–5.

216. Weiss C, Figueras E, Borbely AN, Sewald N. Cryptophycins: Cytotoxic cyclodepsipeptides with potential for tumor targeting. Journal of Peptide Science. 2017;23(7–8):514–31.

217. Qamar H, Hussain K, Soni A, Khan A, Hussain T, Chénais B. Cyanobacteria as natural therapeutics and pharmaceutical potential: Role in antitumor activity and as nanovectors. Molecules. 2021 Jan;26(1):247.

218. Hjörleifsson Eldjárn G, Ramsay A, Van Der Hooft JJJ, Duncan KR, Soldatou S, Rousu J, et al. Ranking microbial metabolomic and genomic links in the NPLinker framework using complementary scoring functions. Nagarajan N, editor. PLoS Comput Biol. 2021 May 4;17(5):e1008920.

219. Wang G, Zhao Z, Ke J, Engel Y, Shi YM, Robinson D, et al. CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria. Nat Microbiol. 2019 Dec;4(12):2498–510.

220. Al-Tebrineh J, Mihali TK, Pomati F, Neilan BA. Detection of saxitoxin-producing cyanobacteria and *Anabaena circinalis* in environmental water blooms by quantitative PCR. Appl Environ Microbiol. 2010 Dec;76(23):7836–42.

221. Sanchez JA, Otero P, Alfonso A, Ramos V, Vasconcelos V, Aráoz R, et al. Detection of anatoxin-a and three analogs in *Anabaena* spp. Cultures: New fluorescence polarization assay and toxin profile by LC-MS/MS. Toxins. 2014 Feb;6(2):402–15.

222. Halinen K, Jokela J, Fewer DP, Wahlsten M, Sivonen K. Direct evidence for production of microcystins by *Anabaena* strains from the Baltic Sea. Applied and Environmental Microbiology. 2007 Oct 15;73(20):6543–50.

223. Peters SJ, Rodgers KJ, Mitrovic SM, Bishop DP. The changes in cyanobacterial concentration of β-Methylamino-L-Alanine during a bloom event. Molecules. 2022 Oct 30;27(21):7382.

224. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-d26.

225. Hart LN, Zepernick BN, Natwora KE, Brown KM, Obuya JA, Lomeo D, et al. Metagenomics reveals spatial variation in cyanobacterial composition, function, and biosynthetic potential in the Winam Gulf, Lake Victoria, Kenya. Applied and Environmental Microbiology. 2025 Jan 8;0(0):01507–24.

226. Boopathi T, Ki JS. Impact of environmental factors on the regulation of cyanotoxin production. Toxins. 2014 Jul;6(7):1951–78.

227. Neilan BA, Pearson LA, Muenchhoff J, Moffitt MC, Dittmann E. Environmental conditions that influence toxin biosynthesis in cyanobacteria. Environ Microbiol. 20120320th ed. 2013 May;15(5):1239–53.

228. Peary, J.A., Gorham PR. Influence of light and temperature on growth and toxin production by *Anabaena flos-aquae* NRC-44-1. Journal of Phycology. 1966;S2.

229. Tonk L, Welker M, Huisman J, Visser PM. Production of cyanopeptolins, anabaenopeptins, and microcystins by the harmful cyanobacteria *Anabaena* 90 and *Microcystis* PCC 7806. Harmful Algae. 2009 Jan 1;8(2):219–24.

230. Rapala J, Sivonen K, Luukkainen R, Niemelä SI. Anatoxin-a concentration in *Anabaena* and *Aphanizomenon* under different environmental conditions and comparison of growth by toxic and non-toxic *Anabaena*-strains — a laboratory study. J Appl Phycol. 1993 Dec 1;5(6):581–91.

231. Long BM, Jones GJ, Orr PT. Cellular microcystin content in N-limited *Microcystis aeruginosa* can be predicted from growth rate. Appl Environ Microbiol. 2001 Jan;67(1):278–83.

232. Sevilla E, Martin-Luna B, Bes MT, Fillat MF, Peleato ML. An active photosynthetic electron transfer chain required for *mcyD* transcription and microcystin synthesis in *Microcystis aeruginosa* PCC7806. Ecotoxicology. 2012 Apr 1;21(3):811–9.

233. Ding Y, Song L, Sedmak B. UVB radiation as a potential selective factor favoring microcystin producing bloom forming cyanobacteria. PLOS ONE. 2013 Sep 13;8(9):e73919.

234. Zhang J, Xu W, Zhang Q, Deng L, Ai F, Yin Y, et al. Effects of elevated atmospheric carbon dioxide to *Microcystis aeruginosa* under different forms of phosphorus sources. Chemosphere. 2025 Apr 1;374:144210.

235. Sevilla E, Martin-Luna B, Vela L, Teresa Bes M, Luisa Peleato M, Fillat MF. Microcystin-LR synthesis as response to nitrogen: transcriptional analysis of the *mcyD* gene in *Microcystis aeruginosa* PCC7806. Ecotoxicology. 2010 Oct 1;19(7):1167–73.

236. Oh HM, Lee SJ, Jang MH, Yoon BD. Microcystin production by *Microcystis aeruginosa* in a phosphorus-limited chemostat. Applied and Environmental Microbiology. 2000 Jan;66(1):176–9.

237. Kramer BJ, Jankowiak JG, Nanjappa D, Harke MJ, Gobler CJ. Nitrogen and phosphorus significantly alter growth, nitrogen fixation, anatoxin-a content, and the transcriptome of the bloom-forming cyanobacterium, *Dolichospermum*. Front Microbiol. 2022 Sep 7;13.

238. Yunes JS, De La Rocha S, Giroldo D, Silveira SB da, Comin R, Bicho M da S, et al. Release of carbohydrates and proteins by a subtropical strain of *Raphidiopsis Brookii* (Cyanobacteria) able to produce saxitoxin at three nitrate concentrations. J Phycol. 2009 Jun;45(3):585–91.

239. Cusick KD, Sayler GS. An overview on the marine neurotoxin, saxitoxin: Genetics, molecular targets, methods of detection and ecological functions. Mar Drugs. 2013 Mar 27;11(4):991–1018.

240. Ginn HP, Pearson LA, Neilan BA. NtcA from *Microcystis aeruginosa* PCC 7806 is autoregulatory and binds to the microcystin promoter. Appl Environ Microbiol. 2010 Jul;76(13):4362–8.

241. Krausfeldt LE, Farmer AT, Castro Gonzalez HF, Zepernick BN, Campagna SR, Wilhelm SW. Urea is both a carbon and nitrogen source for *Microcystis aeruginosa*: Tracking 13C incorporation at bloom pH conditions. Front Microbiol. 2019 May 17;10.

242. Van de Waal DB, Verspagen JMH, Lürling M, Van Donk E, Visser PM, Huisman J. The ecological stoichiometry of toxins produced by harmful cyanobacteria: an experimental test of the carbon-nutrient balance hypothesis. Ecology Letters. 2009;12(12):1326–35.

243. Wagner ND, Osburn FS, Wang J, Taylor RB, Boedecker AR, Chambliss CK, et al. Biological stoichiometry regulates toxin production in *Microcystis aeruginosa* (UTEX 2385). Toxins (Basel). 2019 Oct 16;11(10):601.

244. Schampera C, Hellweger FL. Nitrogen availability controls response of microcystin concentration to phosphorus reduction: Evidence from model application to multiple lakes. Harmful Algae. 2024 Nov;139:102711.

245. Hellweger FL, Martin RM, Eigemann F, Smith DJ, Dick GJ, Wilhelm SW. Models predict planned phosphorus load reduction will make Lake Erie more toxic. Science. 20220526th ed. 2022 May 27;376(6596):1001–5.

246. Rohrlack T, Dittmann E, Börner T, Christoffersen K. Effects of cell-bound microcystins on survival and feeding of *Daphnia* spp. Appl Environ Microbiol. 2001 Aug;67(8):3523–9.

247. Jang MH, Jung JM, Takamura N. Changes in microcystin production in cyanobacteria exposed to zooplankton at different population densities and infochemical concentrations. Limnology and Oceanography. 2007;52(4):1454–66.

248. Wichelen JV, Gremberghe IV, Vanormelingen P, Debeer AE, Leporcq B, Menzel D, et al. Strong effects of amoebae grazing on the biomass and genetic structure of a *Microcystis* bloom (Cyanobacteria).

249. Xu J, Chang X, MacIsaac HJ, Zhou Y, Li J, Wang T, et al. Is a lower-toxicity strain of *Microcystis aeruginosa* really less toxic? Aquatic Toxicology. 2023 Oct 1;263:106705.

250. Gutiérrez-Praena D, Guzmán-Guillén R, Pichardo S, Moreno FJ, Vasconcelos V, Jos Á, et al. Cytotoxic and morphological effects of microcystin-LR, cylindrospermopsin, and their combinations on the human hepatic cell line HepG2. Environmental Toxicology. 2019;34(3):240–51.

251. Rzodkiewicz LD, Turcotte MM. Two duckweed species exhibit variable tolerance to microcystin-LR exposure across genotypic lineages. Harmful Algae. 2024 Jan 1;131:102548.

252. Pawlik-Skowrońska B, Bownik A, Pogorzelec M, Kulczycka J, Sumińska A. First report on adverse effects of cyanobacterial anabaenopeptins, aeruginosins, microginin and their mixtures with microcystin and cylindrospermopsin on aquatic plant physiology: An experimental approach. Toxicon. 2023 Dec 1;236:107333.

253. Bownik A, Adamczuk M, Pawlik-Skowrońska B. Behavioral disturbances induced by cyanobacterial oligopeptides microginin-FR1, anabaenopeptin-A and microcystin-LR are associated with neuromotoric and cytotoxic changes in *Brachionus calyciflorus*. Journal of Hazardous Materials. 2022 Sep 15;438:129472.

254. Gulati RD. Zooplankton structure in the Loosdrecht lakes in relation to trophic status and recent restoration measures. In: Trophic Relationships in Inland Waters. Springer, Dordrecht; 1990. p. 173–88.

255. Gilbert JJ. Suppression of rotifer populations by *Daphnia*: A review of the evidence, the mechanisms, and the effects on zooplankton community structure1. Limnology & Oceanography. 1988 Nov;33(6):1286–303.

256. Okumura DT, Sotero-Santos RB, Takenaka RA, Rocha O. Evaluation of cyanobacteria toxicity in tropical reservoirs using crude extracts bioassay with cladocerans. Ecotoxicology. 2007 Mar;16(2):263–70.

257. Pawlik-Skowrońska B, Bownik A. Synergistic toxicity of some cyanobacterial oligopeptides to physiological activities of *Daphnia magna* (Crustacea). Toxicon. 20211220th ed. 2022 Jan 30;206:74–84.

258. Pawlik-Skowrońska B, Bownik A. Cyanobacterial anabaenopeptin-B, microcystins and their mixture cause toxic effects on the behavior of the freshwater crustacean *Daphnia magna* (Cladocera). Toxicon. 2021 Jul 30;198:1–11.

259. Bownik A, Adamczuk M, Pawlik-Skowrońska B, Mieczan T. Cyanobacterial metabolites: Aeruginosin 98A, microginin-FR1, anabaenopeptin-A, cylindrospermopsin and their mixtures affect behavioral and physiological responses of *Daphnia magna.* Environmental Toxicology and Pharmacology. 2023 Jun 1;100:104161.

260. Crowl TA, McDowell WH, Covich AP, Johnson SL. Freshwater shrimp effects on detrital processing and nutrients in a tropical headwater stream. Ecology. 2001;82(3):775–83.

261. Snyder MN, Small GE, Pringle CM. Diet-switching by omnivorous freshwater shrimp diminishes differences in nutrient recycling rates and body stoichiometry across a food quality gradient. Freshwater Biology. 2015;60(3):526–36.

262. Galanti LN, Amé MV, Wunderlin DA. Accumulation and detoxification dynamic of cyanotoxins in the freshwater shrimp *Palaemonetes argentinus*. Harmful Algae. 2013 Jul 1;27:88–97.

263. Mazur-Marzec H, Sutryk K, Hebel A, Hohlfeld N, Pietrasik A, Błaszczyk A. *Nodularia spumigena* peptides—accumulation and effect on aquatic invertebrates. Toxins (Basel). 2015 Oct 30;7(11):4404–20.

264. Médice RV, Arruda RS, Yoon J, Borges RM, Noyma NP, Lürling M, et al. Unlocking biological activity and metabolomics insights: Primary screening of cyanobacterial biomass from a tropical reservoir. Environ Toxicol Chem. 2024 Oct;43(10):2222–31.

265. Leach JH, Nepszy SJ. The fish community in Lake Erie. J Fish Res Bd Can. 1976 Mar;33(3):622–38.

266. Bambino K, Chu J. Chapter Nine - Zebrafish in toxicology and environmental health. In: Sadler KC, editor. Current Topics in Developmental Biology [Internet]. Academic Press; 2017. p. 331–67. (Zebrafish at the Interface of Development and Disease Research; vol. 124). Available from: https://www.sciencedirect.com/science/article/pii/S0070215316301867

267. Shahmohamadloo RS, Bhavsar SP, Ortiz Almirall X, Marklevitz SAC, Rudman SM, Sibley PK. Lake Erie fish safe to eat yet afflicted by algal hepatotoxins. Science of The Total Environment. 2023 Feb 25;861:160474.

268. Skafi M, Vo Duy S, Munoz G, Dinh QT, Simon DF, Juneau P, et al. Occurrence of microcystins, anabaenopeptins and other cyanotoxins in fish from a freshwater wildlife reserve impacted by harmful cyanobacterial blooms. Toxicon. 2021 Apr 30;194:44–52.

269. World Health Organization. Chemical safety of drinking-water Assessing priorities for risk management [Internet]. World Health Organization; 2007 [cited 2025 Mar 21]. 160 p. Available from: https://www.who.int/publications/i/item/924154676X

270. Bownik A, Pawlik-Skowrońska B. Responses of RTgill-W1 cells to cyanobacterial metabolites microcystin-LR, anabaenopeptin-A, cylindrospermopsin, their binary and ternary mixtures. Toxicon. 2024 Oct;249:108059.

271. Torres MDA, Jones MR, Vom Berg C, Pinto E, Janssen EML. Lethal and sublethal effects towards zebrafish larvae of microcystins and other cyanopeptides produced by cyanobacteria. Aquatic Toxicology. 2023 Oct;263:106689.

272. Monks NR, Moscow JA. Addressing the design of in vitro models for the investigation of microcystins: the essential role of transporter-mediated uptake. Cell Biol Int. 2010 Sep;34(9):907–9.

273. Monks NR, Liu S, Xu Y, Yu H, Bendelow AS, Moscow JA. Potent cytotoxicity of the phosphatase inhibitor microcystin LR and microcystin analogues in OATP1B1- and OATP1B3-expressing HeLa cells. Molecular Cancer Therapeutics. 2007 Feb 16;6(2):587–98.

274. Hrouzek P, Kapuścik A, Vacek J, Voráčová K, Paichlová J, Kosina P, et al. Cytotoxicity evaluation of large cyanobacterial strain set using selected human and murine in vitro cell models. Ecotoxicology and Environmental Safety. 2016 Feb 1;124:177–85.

275. Guzmán EA, Peterson TA, Winder PL, Francis KT, McFarland M, Roberts JC, et al. An assessment of potential threats to human health from algae blooms in the Indian River

Lagoon (USA) 2018–2021: Unique Patterns of cytotoxicity associated with toxins. Toxins. 2023 Nov;15(11):664.

276. Bittner M, Štern A, Smutná M, Hilscherová K, Žegura B. Cytotoxic and genotoxic effects of cyanobacterial and algal extracts—Microcystin and retinoic acid content. Toxins (Basel). 2021 Feb 2;13(2):107.

277. Laureano-Rosario AE, McFarland M, Bradshaw DJ, Metz J, Brewton RA, Pitts T, et al. Dynamics of microcystins and saxitoxin in the Indian River Lagoon, Florida. Harmful Algae. 2021 Mar 1;103:102012.

278. Takser L, Benachour N, Husk B, Cabana H, Gris D. Cyanotoxins at low doses induce apoptosis and inflammatory effects in murine brain cells: Potential implications for neurodegenerative diseases. Toxicol Rep. 2016 Jan 15;3:180–9.

279. Martin RM, Stallrich J, Bereman MS. Mixture designs to investigate adverse effects upon co-exposure to environmental cyanotoxins. Toxicology. 20190425th ed. 2019 Jun 1;421:74–83.

280. Shi JH, Birbeck JA, Olson NE, Parham RL, Holen AL, Ledsky IR, et al. Detection of aerosolized anabaenopeptins from cyanobacterial harmful algal blooms in atmospheric particulate matter. ACS Earth Space Chem. 2025 Feb 25;

281. May NW, Olson NE, Panas M, Axson JL, Tirella PS, Kirpes RM, et al. Aerosol emissions from great lakes harmful algal blooms. Environ Sci Technol. 2018 Jan 16;52(2):397–405.

282. Lim CC, Yoon J, Reynolds K, Gerald LB, Ault AP, Heo S, et al. Harmful algal bloom aerosols and human health. eBioMedicine. 2023 Jul 1;93.

283. Reynolds CS. The ecology of phytoplankton. Cambridge University Press; 2006. 437 p.

284. Guan Y, Zhang M, Yang Z, Shi X, Zhao X. Intra-annual variation and correlations of functional traits in *Microcystis* and *Dolichospermum* in Lake Chaohu. Ecological Indicators. 2020 Apr 1;111:106052.

1.10 Figures



Figure 1.10.1: Photos of cyanoHABs in global, eutrophic Great Lakes. **(A)** Aerial image of western Lake Erie in August of 2019. Photo taken by Zachary Haslick (NOAA). **(B)** Scum and foam from cyanoHAB in western Lake Erie in July 2021. Photo by Colleen E. Yancey. **(C)** CyanoHAB scum and water hyacinth along shores of Winam Gulf, Lake Victoria, Kenya. A person is seen washing dishes near the shore. Photo taken by Julia Obuya.

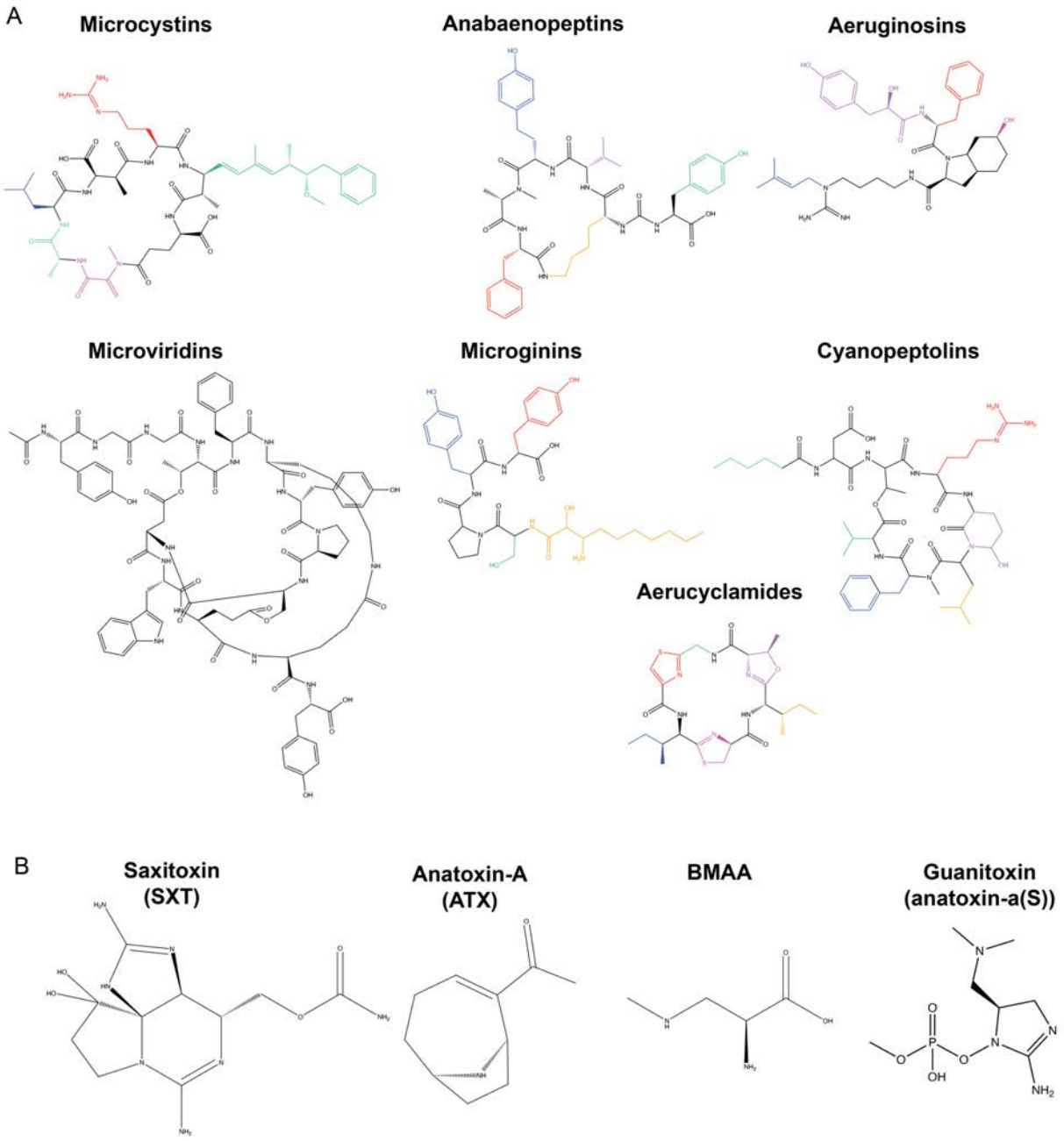


Figure 1.10.2: Representative structures of cyanobacterial metabolites. **(A)** Hepatotoxins: Microcystin LR, Anabaenopeptin A, Aeruginosin KB 676, Microviridin (PubChem ID: 156580592), Microginin 713, Cyanopeptolin A, Aerucyclamide A. Colored structural moieties represent variable parts that can be substituted with alternative building blocks, leading to congener diversity within each metabolite class. **(B)** Neurotoxins: Saxitoxin, Anatoxin-A, β-N-methylamino-L-alanine (BMAA), and guanitoxin (previously anatoxin-(S)).

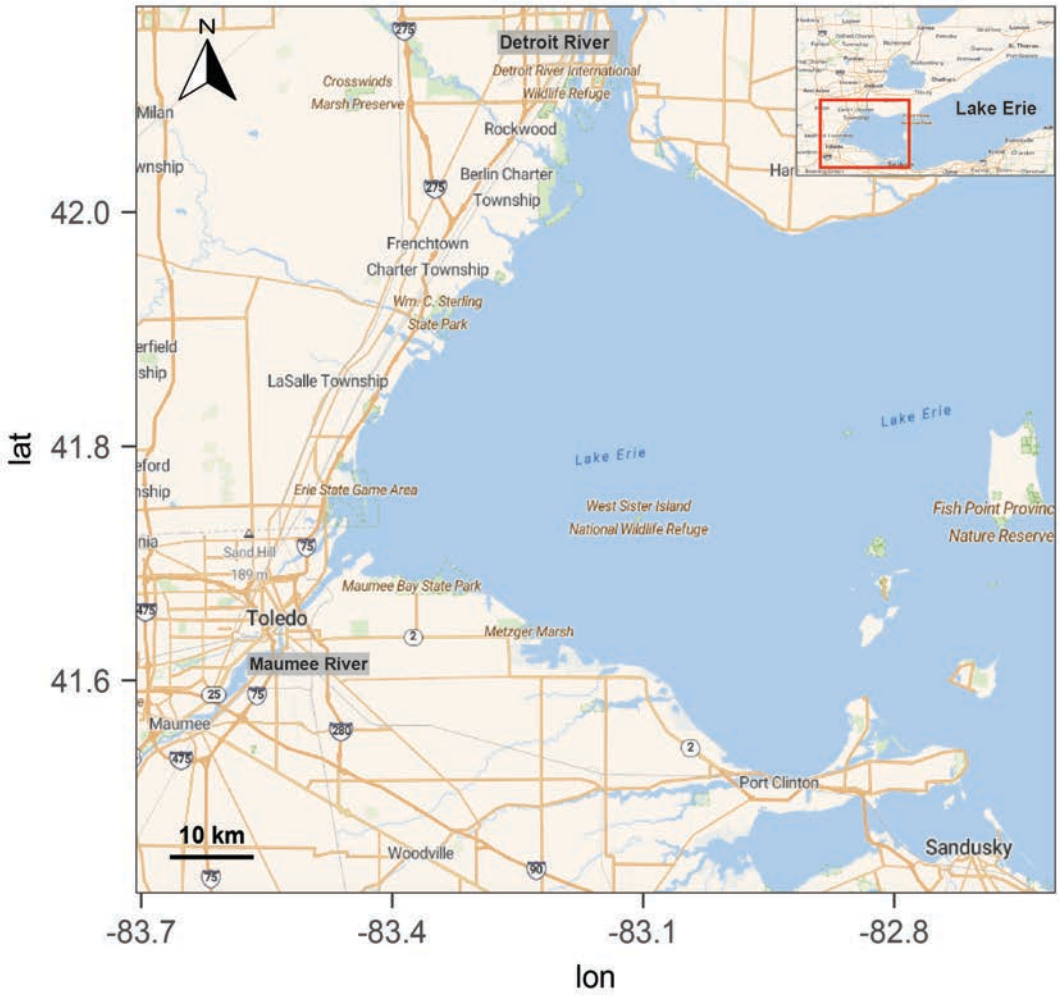


Figure 1.10.3: Map of western Lake Erie, Lake Erie, U.S. with the Detroit and Maumee Rivers noted. Map pulled from Stadia Maps (https://stadiamaps.com/, accessed January 2025).

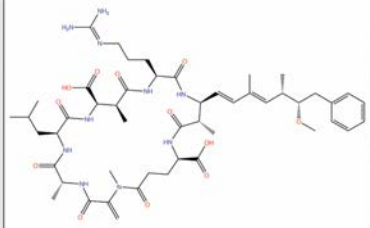
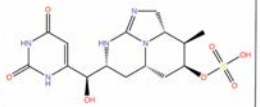
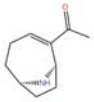
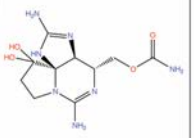
Cyanotoxin	Microcystin	Cylindrospermopsin	Anatoxin-A	Saxitoxin
U.S. Drinking Water Health Advisory (Adults)	3.0 µg/L EPA	1.6 µg/L EPA	3.0 µg/L Oregon State	1.0 µg/L Oregon State
Recreational Limits	8.0 µg/L EPA/WHO	15.0 µg/L EPA/WHO	15.0 µg/L Oregon State	8.0 µg/L Oregon State
Chemical Structure				

Figure 1.10.4: Cyanotoxins monitored by the Environmental Protection Agency (EPA) and their allowed levels. Drinking water and recreation levels come from EPA, World Health Organization (WHO), and Oregon Health Authority (USA).



Figure 1.10.5: Map of Winam Gulf, Lake Victoria, Kenya with the Kibuon and Nyando Rivers noted. Map pulled from Stadia Maps (https://stadiamaps.com/, accessed January 2025).

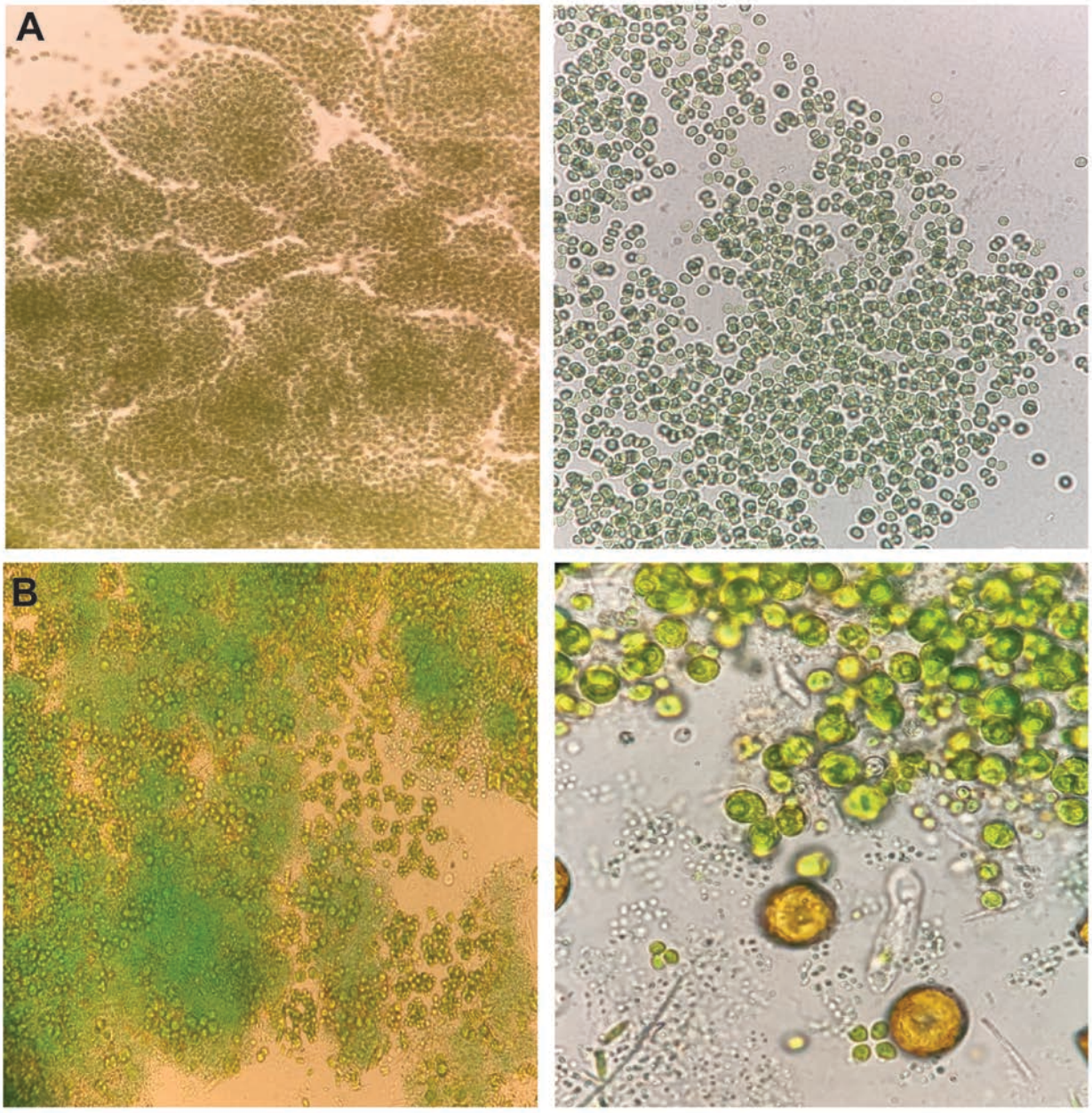


Figure 1.10.6: Microscopy photos taken of **(A)** Western Lake Erie culture collection *Microcystis* isolates LE18-22.4 and LE19-197.1 (24) and **(B)** samples taken from a plankton net tow sample from western Lake Erie containing *Microcystis* cells and other organisms. Photos taken by Helena Nitschky and Nicole Rappuhn. *Microcystis* cells generally range in size from between 1.5 to 7 µm (283).

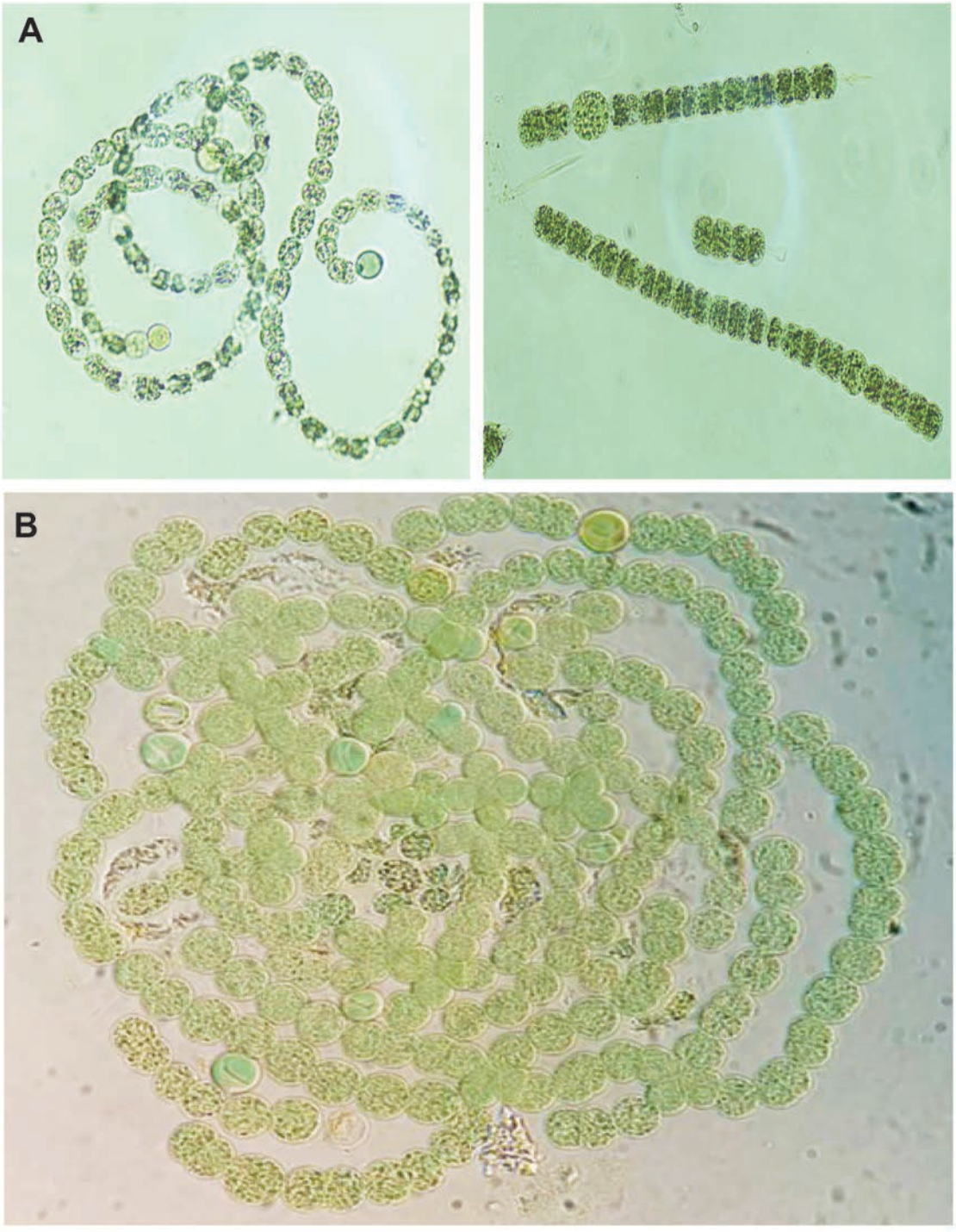


Figure 1.10.7: Microscopy photos taken of **(A)** ADA cultures maintained at Kellogg Biological Station at Michigan State University and **(B)** samples taken from a plankton net tow sample from western Lake Erie containing ADA cells and other organisms. Photos taken by Helena Nitschky and Carol Rosenbaum. *Dolichospermum* cells generally range in size from between 3 to 10 µm (284).

Chapter 2 Diverse Cyanopeptides Follow Distinct Temporal Succession Patterns in Lake Erie Harmful Algal Blooms

2.1 Abstract

Toxic cyanobacterial harmful algal blooms (cyanoHABs) threaten freshwater resources globally and are regarded as intensifying with increasing eutrophication. Bloom toxicity is strongly influenced by intraspecific variation in the biosynthetic repertoires of toxic cyanobacteria, yet few studies examine the vast diversity of cyanobacterial cyanopeptides beyond hepatotoxic microcystins. To understand the dynamics and drivers of intraspecific variation in cyanoHABs, we analyzed temporal patterns of cyanobacteria, cyanopeptides, and their biosynthetic gene clusters (BGCs) in western Lake Erie using a seven-year time series (2016-2022) of metagenomic and metabolomic data. Our findings demonstrate that shifts from *Microcystis* to *Dolichospermum* occur later in the bloom season, coinciding with lower temperatures. Modules of co-varying BGCs (biosynthesis modules) from these genera were identified with hierarchical clustering, with uncharacterized BGCs among the most abundant. Biosynthesis modules rich in nonribosomal peptide synthetases (NRPS) peaked in early August, coinciding with elevated levels of bioavailable nitrogen, warmer temperatures, and high *Microcystis* abundance. In contrast, modules rich in polyketide synthases (PKS) and ribosomally synthesized and post-translationally modified peptides (RiPPs) peaked following the *Microcystis* maximum in mid-August. Metabolomic analyses confirmed that cyanopeptides followed shared seasonal patterns with their associated biosynthesis modules, forming three cyanopeptide phases characterized by (1) microcystins, (2)

anabaenopeptins and aeruginosins, and (3) aerucyclamides. These cyanopeptide phases co-varied with bottom-up and top-down pressures, with later phases coinciding with increased microbially processed organic nitrogen and reduced eDNA detection of grazers. This study reveals consistent patterns of cyanopeptide succession and co-occurrence beyond microcystins, suggesting seasonal tradeoffs between biosynthetic resource demands and ecological controls.

2.2 Introduction

Cyanobacterial harmful algal blooms (cyanoHABs) degrade ecosystems and drinking water supplies globally through the production of toxins that disrupt food webs and harm humans and wildlife (1,2). In temperate and sub-tropical lakes and reservoirs, two genera of bloom-forming cyanobacteria, *Microcystis* and *Dolichospermum* (previously *Anabaena*), often dominate and co-occur within cyanoHABs (3–5). These blooms often follow a seasonal succession where *Dolichospermum*, a diazotrophic (N_2-fixing) genus, proliferates in the spring during periods of high nutrient loading, whereas *Microcystis,* a non-diazotrophic genus, dominates thereafter in warmer, less nutrient rich summer conditions (3,5). *Microcystis* can scavenge scarce nutrients throughout shallow and low-turbulence water columns, contributing to its ability to outcompete other phytoplankton during nutrient-depleted conditions (6–8). Conversely, *Dolichospermum* is better adapted to cooler temperatures and can persist under nitrogen (N) limitation, though it requires high phosphorus (P) availability to outcompete non-diazotrophic competitors (9,10). The ecological success of these two genera is also driven by a suite of adaptive traits beyond nutrient acquisition and temperature adaptation, including the production of toxic, N-rich metabolites that can deter grazing by organisms such as mussels or *Daphnia* and promote selective feeding on other phytoplankton, enhancing their competitive advantage (11).

Cyanobacteria rely on diverse biosynthetic pathways, encoded by BGCs, to synthesize bioactive secondary metabolites, mainly nonribosomal peptide synthetases (NRPS), polyketide synthases (PKS), and ribosomally synthesized and post-translationally modified peptides (RiPPs) (12). Within *Microcystis,* these BGCs encode the synthesis of a wide array of molecules, many of which remain uncharacterized (cryptic), but with most focus on the enzyme-inhibiting cyanopeptides from the classes of anabaenopeptins (NRPS), cyanopeptolins (NRPS), aeruginosins (NRPS-PKS), microcystins (NRPS-PKS), microginins (NRPS-PKS), cyanobactins (aerucyclamides, piricyclamides, etc.) (RiPPs), and microviridins (RiPPs) (13). Microcystins receive substantial attention, accounting for approximately 90% of the relevant literature due to their potency and high concentrations within blooms (13). However, other metabolites can reach even higher concentrations during blooms and comprise over 80% of characterized cyanobacterial secondary metabolites, including compounds equally or more toxic than some MC congeners, representing an underappreciated component of cyanoHAB toxicity (14–16). Members of the *Anabaena, Dolichospermum*, and *Aphanizomenon* (ADA) clade can produce an even greater diversity of toxins, including the neurotoxins anatoxin-a, guanitoxin, and saxitoxin (PKS-like) in addition to hepatotoxic microcystins (17). These metabolites can mediate ecosystems via lethal and behavioral effects on mammals and aquatic organisms such as zooplankton, microcrustaceans, and mussels (18–20).

Biosynthetic potential is strain-specific in bloom-forming cyanobacteria, yet efforts to understand intraspecific variability, particularly within the scope of *Microcystis,* have largely relied on culture- and qPCR-based approaches targeting genes such as the *mcy* operon responsible for MC synthesis (17,21–23). For example, “toxigenic,” or *mcy*-encoding strains of *Microcystis,* are often observed early in the season in temperate blooms, coinciding with periods

of high temperature and elevated inorganic N concentrations (24,25). Numerous studies across diverse habitats have documented this pattern, however less is known about the ecological strategies and toxigenic potential of strains present in eutrophic environments where N is predominantly present in recycled, organic forms (26).

Although some studies have examined cyanopeptide profiles of *Microcystis* isolates and single cells (27,28), these approaches have not resolved in situ cyanopeptide temporal dynamics and their environmental drivers from an integrated genomic and metabolomic perspective. Recent advances in metagenomic and metabolomic technology, along with expanding BGC and metabolite databases, now make it possible to study such dynamics with greater resolution (29–31). Despite these advancements, temporal patterns of cyanobacterial secondary metabolism remain poorly understood due to limited long-term datasets, underscoring the need for sustained sampling efforts coupled with physicochemical measurements and robust phenological modeling approaches to uncover the environmental drivers of cyanopeptide synthesis beyond microcystins.

The patterns and drivers of succession between *Microcystis* and *Dolichospermum* are well documented, however, the corresponding biosynthetic patterns—temporal patterns of BGCs and their products—remain understudied, leaving their dynamics in natural settings largely unexplored. In this study, we analyzed the temporal patterns of *Microcystis* and *Dolichospermum* dominance, biosynthetic potential, and cyanopeptide abundance using hierarchical general additive models (HGAMs) from seven years (2016-2022) of metagenomics and metabolomics data from the western basin of Lake Erie (WLE). WLE serves as a natural laboratory to model temperate eutrophic basins experiencing cyanoHABs because it is well-monitored for its prolific seasonal blooms, and it has spatiotemporal gradients of nutrient availability from dual-river (Maumee and Detroit Rivers) inputs. The co-occurrence of *Microcystis* and *Dolichospermum* in

this basin enabled us to identify environmental conditions favoring their dominance. We detected genus-specific seasonal and interannual patterns in BGCs and identified distinct biosynthesis modules—clusters of BGCs with shared temporal patterns. We validated the temporal trends of these biosynthesis modules using non-targeted metabolomics, which confirmed seasonally variable production of cyanopeptides of interest. Top-down and bottom-up ecological controls co-varied with bloom phases characterized by distinct cyanopeptides, implicating resource demand and grazing as putative drivers of temporal successions in biosynthetic potential and cyanopeptide abundance, expanding our understanding of intraspecific chemical diversity beyond microcystin production in cyanoHABs.

2.3 Materials and Methods

2.3.1 Sample Collection

Samples were collected from 2016 through 2022 at four NOAA Great Lakes Environmental Research Laboratory (GLERL) sampling stations (**Figure 2.8.1**) in WLE. Physicochemical parameters used in this study are described in Boegehold et al. (2023) (**SI Figure S2.9.1**) (32). Characterization of the five major fluorescence components of dissolved organic matter (FDOM) (C1-C5) was conducted as previously described **(SI Table S2.9.1**) **(**33). Lake water was collected from one meter below the surface, filtered through a 3 μm Isopore polycarbonate membrane filter (MilliporeSigma, Burlington, MA) for both DNA and metabolite samples, and frozen at -80°C.

2.3.2 DNA Extraction and Sequencing

DNA was extracted from filters using a DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) per the manufacturer's instructions. The quality and quantity of DNA was assessed using a

NanoDrop™ Life Spectrophotometer (Thermo Scientific, Waltham, MA) and stored at -20°C. DNA libraries were prepared and sequenced at the University of Michigan Advanced Genomics Core using the NovaSeq platform (Illumina, San Diego, CA) with a 300 cycle S4 flowcell to generate ~50 million 150 base pair paired-end reads per sample.

2.3.3 Metabolite Extraction and Data Collection

Metabolite filters were submerged in 1.5 mL of water, vortexed for 10 seconds, and subjected to three freeze-thaw cycles (−80°C for 20 min, 26°C water bath for 8 min). QuikLyse solution A (100 µL; Gold Standard Diagnostics, Horsham, PA) was added, followed by a 2-minute shake and 8-minute rest at room temperature. This step was repeated using 10 µL of QuikLyse solution B. Extracts were transferred to polypropylene microcentrifuge tubes, centrifuged at 5.0×g for 10 min at 25°C, and filtered through a 0.7 µm glass fiber filter into LC/MS vials. Samples were stored at −20°C until liquid chromatography/high resolution mass spectrometry (LC/HRMS).

Analytes were separated using the solvent gradient in **SI Table S2.9.2** on a Thermo Scientific Vanquish Bioinert HPLC with a Waters Atlantis T3 column (3 µm, 3.0 × 150 mm). Mass spectrometry was performed on a Thermo Q Exactive Orbitrap with HESI-II source in positive ion mode. All samples were run in duplicate. Full HPLC and LC/HRMS protocols are described in **SI Note 1**. Data are available in Lopez et al. 2025 (34).

2.3.4 Metagenomic Data Processing and Analysis

Metagenomic reads were processed, assembled and binned using the Great Lakes Atlas of Multi-Omics Research (GLAMR) pipeline as described previously (35) (https://github.com/Geo-

omics/GLAMR_omics_pipelines; **SI Note 2**). GLAMR quality-controlled (QC) reads were mapped to a dereplicated metagenome-assembled genome (MAG) set (98% average nucleotide identity (ANI) via dRep v3.2.0) using CoverM v0.6.1 (MiniMap2 v.2.24) to calculate MAG relative abundance (36–38). Samples were categorized into three bloom stages based on *Microcystis* and *Dolichospermum* relative abundance differences, where Stage 1 indicates *Microcystis* dominance (>3%), Stage 2 indicated relative equal abundance of *Microcystis* and *Dolichospermum* (±3%), and Stage 3 indicates *Dolichospermum* abundance (>3%).

Non-metric multidimensional scaling (NMDS) was performed on a Bray-Curtis distance matrix derived from the cyanobacterial composition of samples at the genus level (Stress = 0.071), with samples colored by stage. Environmental variables were fit as vectors and those with $P < 0.05$ were bolded. A random forest model (*randomForest* R package (39)) using 1,000 trees and six-variable splits assessed variable importance in classifying bloom stage.

To assess grazer abundance, QC reads were taxonomically classified using the easy taxonomy workflow (mode 3) in MMseqs2 v2.24 against the UniRef100 database (v2022_04) (17, 29–31). Relative abundance of *Daphnia* and *Dreissena* (mussels) was calculated via total sum scaling (sensitivity 4). Due to their benthic habitat, *Dreissena* detection via eDNA likely reflects gametes or veligers present in the water column.

2.3.5 BGC Detection from Metagenomes

A non-redundant, BGC database was developed to analyze temporal BGC patterns (**SI Table S2.9.3**), with files and code available on GitHub, **SI Note 3,** and **SI Figure S2.9.2.** QC reads were mapped to the database using MiniMap2 v2.24 (> 80% coverage and identity kept) (37). Only core BGC genes were considered due to their essential nature for biosynthesis (40). BGCs were

considered present if 90% of the core genes were covered. Average fragments per kilobase million (FPKM) values were calculated for each BGC in each sample using Equation 2.3.1 (41).

Equation 2.3.1: Average Fragments per kilobase million (FPKM) for each biosynthetic gene cluster.

$$Average\ FPKM\ = \frac{\sum_{i=1}^{no.of\ core\ genes\ in\ BGC}\left(\frac{\text{reads mapped to gene}_i}{\left(\frac{\text{gene length}_i}{1{,}000}\ \text{x total reads per sample}\right)}\text{x } 1{,}000{,}000\right)}{no.of\ core\ genes\ in\ BGC}$$

2.3.6 Biosynthesis Module Analysis

To prevent skewing of distance calculations and clustering by high-frequency features, BGCs were grouped based on taxonomic origin and occurrence frequency. ADA BGCs and infrequent *Microcystis* BGCs (absent in two or more years) were grouped, whereas consistently detected *Microcystis* BGCs were grouped. BGCs were hierarchically clustered (average linkage, Bray–Curtis distance), with the optimal number of clusters (k = 10) determined by silhouette scores (> 0.5). Cluster stability was assessed via 10,000 bootstraps, and clusters were defined using *cutree* (R *stats* package) (42). Trees were visualized in the Interactive Tree of Life (iTOL) (43). Permutation tests (n=10,000) based on silhouette scores assessed cluster quality, with $P < 0.05$ indicating cohesive clusters.

2.3.7 Hierarchical General Additive Models (HGAMs)

Independent models for *Microcystis*, *Dolichospermum*, and the summed relative abundance (FPKM) of BGCs from individual biosynthesis modules were fit to a smoothed spline of day of year (DOY) in a generalized additive model (GAM) using 170 observations. Because the data are

structured by site and year, each model was fit as a hierarchical GAM (HGAM). Response variables were natural log-transformed using ln(x + 1) to ensure model adequacy, resulting in residuals that were approximately normally distributed and centered near zero. Akaike Information Criterion (AIC) values were assessed during model selection to confirm that each term improved fit. The HGAM for *Microcystis* and *Dolichospermum* included a smooth on DOY with a cyclic cubic regression basis function, site-specific DOY smooths, a random smooth of year within site, and separate random smooths for site and year. These models assumed a Tweedie distribution, suitable for continuous data with many zeros such as relative abundance. The HGAM for biosynthesis modules included a smooth on DOY with a cyclic cubic regression basis function, a tensor interaction allowing DOY patterns to vary by site, a random smooth of year within site, and a separate random smooth for site. These models assumed a Gaussian distribution, appropriate for continuous data that were approximately symmetric and homoscedastic after transformation. Models were fit using the *mgcv* R package (44).

2.3.8 Microcystis Isolate Analysis

BGCs from the non-redundant database were identified in *Microcystis* genomes available in the NCBI database (**SI Table S2.9.4**) using blastn (45). BGC hits were considered if 50% of core genes and 50% of all genes within a BGC were present (≥70% identity and coverage). Genomes were used to build a maximum likelihood phylogenomic tree based on the prepackaged single-copy gene-set for cyanobacteria (251 target genes) using FastTree2 v2.1.11 via GToTree v.1.7.05 (43,46,47). The relative abundance of WLECC *Microcystis* isolates (NCBI BioProject: PRJNA903891) was calculated as described in Yancey et al. 2023A and **SI Note 2** (48).

2.3.9 Metabolomics Analysis

Raw data files (.d) were converted to .mzML format using MSConvert in Proteowizard v3 (49). A suite of *in silico* tools were used to identify and annotate MS/MS features (13,428 total), an ion signal of a specific mass-to-charge ratio (*m/z)* detected at a specific retention time (RT) (31). First, MZmine v4.3.0, using standard Orbitrap parameters (**SI Table 5**), was utilized for feature extraction across samples (50). Feature intensities across samples correspond to peak areas in extracted ion-chromatograms (EICs). The feature-based molecular networking (FBMN) workflow with standard parameters in GNPS2 was run (https://gnps2.org/status?task=8c4014d3f0c547b0b4e9fdcaf6aed61c) (31). Library matches, level three as outlined in Schymasnski et al. (2014), to cyanopeptides were kept and checked for characteristic fragment ions (**SI Figure S2.9.3**) (51,52). The normalized abundance of the cyanopeptides across samples were calculated after blank removal (20% threshold, removed 4,077 features) and imputation in the Functional Metabolomics Lab's FBMN-STATS guide (53) using Equation 2.3.2 where AUC indicates peak area and volume filtered indicates the volume of water filtered for the corresponding sample.

Equation 2.3.2: Normalized metabolite abundance.

$$\text{Normalized Metabolite Abundance} = \frac{\text{AUC} * 0.00161\ \text{uL}}{\text{Volume filtered (uL)}} * \text{Chlorophyll A } (\frac{\text{ug}}{\text{L}})$$

The normalized abundance of metabolite classes (congeners summed by metabolite class) and their associated BGC were Z-scored to compare relative abundance. The maxima and minima of metabolite abundance, physicochemical parameters, and grazer eDNA detection were estimated by fitting GAMs using DOY as a smooth term with a cyclic cubic regression basis, followed by

10,000 bootstrap iterations to generate 95% confidence intervals. Combinations of the five FDOM components are proxy estimates of fractions of the DOM pool: terrestrial humic DOM (Component 3 (C3)), microbially-derived DOM calculated as the ratio of microbially-derived to terrestrial humic DOM ((C1+C4+C5)/C3), and portion of aromatic amino acids within the total FDOM pool ((C4+C5)/C1+C2+C3+C4+C5). To assess how nitrogen and DOM variables relate to metabolite abundance independent of seasonal trends, we fit GAMs to log-transformed metabolite abundances using DOY as a smooth term with a cyclic cubic regression basis. Residuals from each GAM, representing variation not explained by seasonal progression, were then extracted and plotted against individual abiotic variables in pairwise scatterplots. Linear regressions were performed on each comparison, and the resulting R^2 and *P*-values are reported.

2.3.10 Figure Rendering

Figures were generated in R on R Studio using the packages *ggplot2* and *gratia* (54–56). BGC gene schematics were generated in antiSMASH v7 (29).

2.4 Results

2.4.1 Succession of Microcystis to Dolichospermum is Coincident with Changes in Temperature and Nutrient Availability

Microcystis and *Dolichospermum* showed distinct temporal patterns of relative abundance, with *Microcystis* peaking concurrently with peak particulate microcystin in late July through early August, whereas *Dolichospermum* abundance increased as microcystin levels declined in late August, reaching its peak in October (**Figure 2.8.2A**). Interannual variation accounted for 27.8% and 36.3% of the explained deviance in the HGAMs for *Microcystis* and *Dolichospermum*,

respectively (**SI Figure S2.9.4, SI Table S2.9.6**). Site-specific variation only contributed between 1 and 8 percent of explained deviance (**SI Table S2.9.6**).

Samples were grouped into three bloom stages based on *Microcystis* and *Dolichospermum* abundance to assess interannual variation in taxonomic succession patterns. Between 2016 and 2019, most samples were categorized as stages 1 and 2, indicating *Microcystis* dominance and approximately equal relative abundance of *Microcystis* and *Dolichospermum*, respectively (**Figure 2.8.2B**). Most stage 3 samples, indicating *Dolichospermum* dominance, were observed in September and October between 2020 and 2022. Transitions between stages 1, 2, and 3 often followed a sequential pattern, but not always (**SI Figure S2.9.5**). NMDS revealed associations between physicochemical measurements and bloom composition (**Figure 2.8.2C, SI Figure S2.9.6, SI Table S2.9.7**). The first NMDS axis appeared to separate samples based on indicators of bloom biomass, with higher values of phycocyanin, particulate microcystin, and particulate organic carbon and nitrogen (POC, PON) aligning mostly with stage 1 samples. The second axis was more strongly associated with seasonal patterns in environmental conditions and discriminated between samples dominated by *Microcystis* (stage 1) or *Dolichospermum* (stage 3), including variables such as temperature, nitrate + nitrite, soluble reactive phosphorus, and total phosphorus. Despite significant associations with physicochemical parameters ($P < 0.01$, **SI Table S2.9.7**), samples did not distinctly cluster by stage, likely due to biological and environmental variability across years. Random forest classification further supported these patterns, identifying temperature, particulate microcystin, and phycocyanin as the variables most strongly associated with bloom stage classification (**Figure 2.8.2D,** 58.3% accuracy, 95% Confidence Interval: 49–67%). Measures of organic matter (POC, PON, colored dissolved organic matter (CDOM)) and dissolved and particulate nutrients also contributed to model accuracy. Both nutrient

concentrations and biomass-associated metrics were among the most informative features and displayed statistically significant differences between stages (**SI Figure S2.9.7**). Analysis of surface water temperature data from WE2 between 2015 and 2023 revealed an increasing frequency of cooling days in September in recent years, suggesting that more frequent or abrupt cooling events in the fall may contribute to interannual variation in succession dynamics (**SI Figure S2.9.8**) (57).

2.4.2 Cryptic BGCs are the Most Abundant BGCs Encoded by Microcystis and Dolichospermum

Microcystis and *Dolichospermum* BGCs exhibited varied temporal patterns (**Figure 2.8.3**). ADA BGCs putatively encoding anacyclamide, cryptic NRPS (NRPS-1 and NRPS-2), and PKS (T1PKS) pathways were the most abundant among ADA BGCs. The BGCs were present across all years, however they increased in abundance in 2019, reaching maximum abundance in October of 2022. Conversely, the cryptic RiPPs (Lanthipeptide and LAP), saxitoxin, and NRPS-4 ADA BGCs occurred infrequently at low abundance in July of 2018 through 2021.

The most frequently observed and highly abundant *Microcystis* BGCs were PmNT3, LAP, T1PKS-Terpene, lanthipeptide-3, anabaenopeptin, and cyanopeptolin (**Figure 2.8.3, SI Figure S2.9.9**) (23,58). The abundance of these BGCs peaked between late July and early August. The BGCs putatively encoding synthesis of aeruginosins and microcystins, and the cryptic BGCs cyanobactin-1, cyanobactin-2, T3PKS-1, and NRPS-T1PKS were observed at lower abundance and followed similar temporal patterns. The cryptic BGCs RiPP-like-2, lanthipeptide-1, and lanthipeptide-2 were found at low abundance sporadically only between the years of 2019 and

2022. The other 13 BGCs, mainly from RiPP and PKS pathways, were found at lower abundance, often becoming most abundant in late August through early September.

2.4.3 BGC Co-Variance Forms Distinct Biosynthesis Modules with Varying Temporal Patterns

We used hierarchical clustering to identify sets of BGCs that share distinct temporal patterns, which we refer to as biosynthesis modules (**Figure 2.8.4A**). Modules often contained BGCs from similar pathways, such as module four (75% NRPS-PKS), module six (80% NRPS and PKS), and module seven (66% RiPP-Cyanobactin and PKS). Conversely, module five contained BGCs from a mixture of RiPP, NRPS-PKS, and PKS pathways. *Microcystis* biosynthesis modules followed distinct temporal patterns of relative abundance (**Figure 2.8.4B, SI Figure S2.9.10, SI Table S2.9.8**). Modules four and six exhibited peak abundance from late July through early August, whereas module seven showed a later peak spanning August to mid-September. Although temporal differences were modest and confidence intervals overlapped, module four, encoding the *mcy* operon, declined sharply following its peak, whereas module six showed a secondary, lower-intensity peak in mid-September. Modules eight and nine, composed of ADA BGCs, shared a similar composition, each encoding a trio of BGCs from NRPS, RiPP, and either PKS or another pathway. Module eight was observed at low abundance in June and July, whereas module nine was observed at higher abundance in September through October (**SI Figure S2.9.11**).

2.4.4 Biosynthesis Modules Reflect Mosaics of Microcystis Isolates and their BGC Content

To test the hypothesis that field-derived biosynthesis modules reflect the BGC content of locally isolated *Microcystis* cultures, we calculated the percent of each module represented in genomes of *Microcystis* cultures isolated from WLE (**Figure 2.8.5A**) (48). Some cultures encoded 100% of one or more modules, but no culture encoded only one module. Module one, containing only the aerucyclamide BGC, was not represented in any of the locally relevant cultures. All cultures contained some of modules five and six, with cultures often encoding more of module six and less of module four. Analysis of all publicly available *Microcystis* genomes revealed that BGCs from the same modules (microcystin & aeruginosin (four), T1PKS-Terpene & PmNT3 (six), cyanopeptolin & anabaenopeptin (six), etc.) co-occur most often (**SI Figure S2.9.12**). Conversely, the microcystin and anabaenopeptin BGCs did not co-occur in the publicly available genomes analyzed. Phylogenomic analysis revealed an evolutionary split between *Microcystis* genomes encoding either of these two BGCs (**SI Figure S2.9.13**).

Some cultures were found to encode nearly complete modules, such as LE18-22.4 (module four), LE19-10.1 (modules five and six), and LE19-12.2 and LE19-197.1 (modules six and seven). LE17-10 and LE17-20 are good representatives of *Microcystis* cultures with low biosynthetic potential in WLE, relative to other cultures available in the culture collection. The relative abundance of LE18-22.4, LE19-197.1, LE19-10.1, and LE17-10 peaked between mid-July through early-August, concurrent with peaks in microcystin and abundance of modules four, five, and six (**Figure 2.8.5B)**. LE17-20 and LE19-12.2 peaked later in the season between August and September, when microcystin levels declined and the relative abundance of module seven was highest.

2.4.5 Uncoupled Temporal Patterns of Cyanopeptides and Associated BGCs

Metabolomics data showed that the temporal patterns of four cyanopeptides varied, with microcystin peaking in late July, aeruginosin and anabaenopeptin peaking in mid-August, and aerucyclamide peaking from late-August through early September (**Figure 2.8.6, SI Table S2.9.9**, **SI Figure S2.9.14)**. Relationships between metabolites and BGCs also varied, with the abundance of microcystin and its BGC reaching a maximum at the same time, whereas the aeruginosin and anabaenopeptin metabolites peaked approximately fifteen days after their BGCs peaked. Lastly, the aerucyclamide BGC exhibited two distinct maxima, first between July and early August and again in late August through September; however, the aerucyclamide metabolite was only observed during the second, smaller BGC peak.

2.4.6 Unique Cyanopeptide Maxima are Associated with Shifts in Nitrogen, DOM, and Grazer Abundance

The temporal variability in cyanopeptide abundance suggests that environmental factors, including resource availability and grazing pressure, may influence their dynamics. To evaluate potential bottom-up and top-down controls, we examined the temporal patterns, maxima, and minima of cyanopeptides, nitrogen, dissolved organic matter (DOM), and grazer abundance (**Figure 2.8.7, SI Table S2.9.9**). Combined nitrate and nitrite concentrations, terrestrial humic fraction of DOM, ammonia, and *Daphnia* and *Dreissena* eDNA detection all peaked in early to mid-July, followed by temperature and microcystin peaks in late July. POC:PON ratio, microbially derived and aromatic amino acid fractions of DOM, and amoeba eDNA detection peaked in mid- to late-August, immediately following the observed peaks in anabaenopeptin, aeruginosin, and aerucyclamide and minima of combined nitrate and nitrite, terrestrial humic DOM, and ammonia. Ammonia concentrations rose again between September and October, although not to the peak

levels observed in July (**Figure 2.8.7A**). Because DOY accounted for up to half or more of the seasonal variation in cyanopeptide abundance, we evaluated the effects of nitrogen and DOM on the residual variation not explained by DOY **(SI Figure S2.9.15)**. Nitrogen and DOM variables showed only weak to moderate correlations with the residuals (all $R^2 < 0.40$), suggesting limited additional predictive power for these variables beyond what was already captured by DOY.

2.5 Discussion

A growing body of research has demonstrated that the production of cyanopeptides beyond microcystins by cyanobacteria is shaped by complex ecological interactions and environmental conditions, yet the temporal dynamics of these processes remain poorly understood. The BGC compositions that encode the synthesis of these metabolites vary within genera like *Microcystis* and *Dolichospermum*, which frequently co-occur and dominate freshwater cyanoHABs. Our analysis of a seven-year, metagenomic and metabolomic time series revealed coincident changes in environmental conditions and grazer abundances with three distinct phases characterized by diverse cyanopeptides in WLE.

Consistent with previous studies documenting *Dolichospermum* adaptation to low temperatures and nutrient availability (3,5), our results suggest that declining temperatures and nitrogen concentrations were most associated with the late season shift to *Dolichospermum* dominance. In contrast, the warm, N-replete yet P-depleted summer conditions favored *Microcystis* dominance, consistent with its exceptional ability to scavenge P at low concentrations and, under N sufficiency, to produce grazer-deterring secondary metabolites such as microcystins (3,7,59). Previous studies in WLE have not documented these patterns beyond a few years. In our study of seven years of data, we first observed the successional shift to *Dolichospermum* in 2019, with this pattern continuing through 2022. This interannual variability is not unprecedented as

ADA dominance was observed in WLE in 2011 (60) but not 2014 (24,61). However, WLE was dominated by ADA in the 1970's prior to total P load reductions catalyzed by the Great Lakes Water Quality Agreement (62). In addition to changing nitrogen availability, our data suggests the shift to *Dolichospermum* in 2019-2022 could be due to colder September temperatures (**SI Figure S2.9.15)** that limited *Microcystis* proliferation, promoted biomass decomposition, and released bioavailable P that facilitated *Dolichospermum* growth (63). Increased climate variability such as variable jet stream patterns causing abrupt fall cooling may further promote these successions, in addition to other seasonally-linked changes not measured in this study, such as light (64).

This study demonstrates clear BGC succession in cyanoHABs over multiple years, providing evidence for seasonally distinct biosynthesis modules—co-varying clusters of BGCs with coordinated temporal patterns. Biosynthesis modules, which partially reflect BGC composition in co-occurring strains, influence bloom toxicity beyond microcystins alone. We found that modules dominated by NRPS and NRPS–PKS BGCs, which rely on N-rich enzymatic machinery and non-proteinogenic amino acids, were most abundant under N-replete conditions, coinciding with peak microcystin production (65,66). In contrast, modules enriched in RiPPs, such as those predicted to produce cyanobactins using canonical amino acids and ribosomal machinery, and PKS BGCs, responsible for producing carbon-rich metabolites, were more abundant during periods with lower N availability (65). These findings expand upon previous work linking nutrient availability to microcystin production and congener differentiation (67–69) by showing that stoichiometric controls may extend more broadly across diverse classes of biosynthetic pathways and cyanopeptides.

To further explore the environmental controls on cyanopeptides, we integrated metagenomic and metabolomic data, identifying three temporally variant bloom phases defined by distinct

cyanopeptides. These phases mirrored patterns in biosynthesis module composition and corresponded to changes in concentration of inorganic N and DOM makeup. The early summer microcystin peak followed maxima in nitrate, ammonia, temperature, and terrestrial humic DOM, likely reflecting high nutrient loads delivered via spring runoff. This timing aligned with elevated levels of nitrogen-intensive NRPS and NRPS-PKS biosynthesis modules, supporting the hypothesis that microcystin production is promoted by N availability in concert with other cues, particularly temperature, which peaked just beforehand (59,67). Two additional cyanopeptide phases occurred later in the season under different conditions. The late August rise in anabaenopeptins and aeruginosins, and the subsequent peak in aerucyclamides, a less N-intensive RiPP, coincided with increases in reduced and organic nitrogen, including ammonia and aromatic amino acids. The co-variance of microbially processed DOM and inorganic N depletion with non-microcystin cyanopeptides suggests that strains lacking *mcy* gain a competitive advantage late in the bloom, a pattern observed in eutrophic lakes such as Erie (26,70), Taihu (71), and Champlain (72). Based on these data, we hypothesize that shifts in nitrogen and DOM pools may alter cyanopeptide profiles through mechanisms such as co-regulation by shared environmental signals (73), stimulation of biosynthesis via increased carbon flux through aromatic amino acid pathways (e.g., shikimate) (74), or competitive and protective responses to accumulating DOM components (75). Together, these findings demonstrate phase-specific environmental associations with cyanopeptide production, with early season cyanopeptides tied to external nutrient loading and warming, and late season cyanopeptides possibly linked to internal recycling and organic nutrient availability.

In addition to these metabolic triggers, ecological interactions, such as those within the phycosphere, may also influence cyanopeptide abundance under nutrient-depleted periods.

Previous studies have shown that microcystins can interfere with inorganic carbon metabolism and reduce microbiome richness (76). However, the late season production of anabaenopeptins, aeruginosins, and aerucyclamides suggests different interactions. We hypothesize that these cyanopeptide phases may also be driven by N remineralization and recycling mediated by the phycosphere, implying a potential symbiotic relationship between non-*mcy* encoding *Microcystis* strains and their associated microbiome (77,78).

Beyond interactions with the phycosphere, the shared temporal patterns of biosynthesis modules and their associated cyanopeptides suggest broader, multi-trophic ecological roles that may confer strain-specific competitive advantages. Certain BGC pairs frequently co-localize in *Microcystis* genomes, such as those encoding microcystin with aeruginosin, or anabaenopeptin with cyanopeptolins, whereas others, like microcystin and anabaenopeptin, rarely occur together, suggesting selective pressures favoring specific metabolite combinations. Although the ecological functions of most cyanopeptides remain unclear, they are predicted to influence bloom dynamics via grazing deterrence, allelopathy, quorum sensing, and nutrient regulation (79–84). We identified that microcystin abundance peaked during periods of high *Daphnia* and *Dreissena* abundance, also aligning with known periods of high grazing activity by these organisms, and further supporting microcystins' role as a grazer deterrent (85,86). As microcystin declined, anabaenopeptin and aeruginosin levels increased, resulting in the co-occurrence of these three cyanopeptides. Prior studies have demonstrated synergistic and additive toxicity of these compounds in *Daphnia* (16), supporting the ecological relevance of this pattern. This second cyanopeptide phase also coincided with increased amoeba abundance, against which anabaenopeptins have demonstrated inhibitory effects, suggesting a potential defensive role of anabaenopeptins and aeruginosins (87). Although microcystins are the most extensively studied, anabaenopeptins may serve similar ecological

functions or support strain-specific niche partitioning, which may explain their rare genomic co-occurrence despite overlapping temporal patterns with distinct grazer populations. These findings highlight how variation in biosynthesis may influence strain fitness and emphasizes the need to investigate the ecological roles and combinatorial effects of co-occurring metabolites within complex bloom communities.

This study reveals interannually conserved temporal patterns of cyanobacterial taxa, biosynthetic potential, and cyanopeptide abundance in freshwater cyanoHABs, suggesting intraspecific variability in secondary metabolism as a marker of distinct populations in a bloom. These successions co-varied with ecological controls, providing hypotheses about environmental drivers of strain-specific fitness. The identification of three bloom phases, each characterized by different biosynthesis modules, *Microcystis* strains, and cyanopeptides, revealed new insights about cyanoHAB chemical ecology beyond the dynamics microcystin alone. These results demonstrate that the absence of microcystins does not necessarily indicate reduced risk, as other bioactive cyanopeptides may occur at toxicologically relevant levels. Co-occurrence of cyanopeptides across phases warrants investigation into their individual and combinatorial toxic effects to better define cyanoHAB risk profiles. Integrating multi-omics analyses, strain-level fitness studies, and evolutionary frameworks will be critical for understanding cyanoHAB functional complexity and anticipating how anthropogenic pressures may intensify their ecological and public health impacts.

2.6 Data Availability

All physicochemical parameters measured in this study are freely available from NOAA National Centers for Environmental Information (NCEI) at https://doi.org/10.25921/11da-3x54 (Cooperative Institute for Great Lakes Research, University of Michigan, and NOAA Great Lakes

Environmental Research Laboratory, 2019, Accessed November 2024). Raw reads, assemblies, and MAGs used in this study were deposited on NCBI under BioProject PRJNA1190886, BioSamples SAMN45060126-SAMN45060294. All metabolomics datafiles used in this study can be found in Lopez et al. (2025) (34) and were deposited on the MassIVE database under doi: https://doi.org/doi:10.25345/C58W38F3P. All scripts for analysis are available on GitHub at: https://github.com/Geo-omics/2016-2022-Western-Lake-Erie-Time-Series-Metagenomics-and-Metabolomics.

2.7 References

1. Harke MJ, Steffen MM, Gobler CJ, Otten TG, Wilhelm SW, Wood SA, et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. Harmful Algae. 2016 Apr;54:4–20.

2. Chorus I, Welker M, editors. Toxic Cyanobacteria in water: A guide to their public health consequences, monitoring and management. Taylor & Francis; 2021.

3. Paerl HW, Otten TG. Duelling 'CyanoHABs': unravelling the environmental drivers controlling dominance and succession among diazotrophic and non-N2-fixing harmful cyanobacteria. Environmental Microbiology. 2016;18(2):316–24.

4. Hart LN, Zepernick BN, Natwora KE, Brown KM, Obuya JA, Lomeo D, et al. Metagenomics reveals spatial variation in cyanobacterial composition, function, and biosynthetic potential in the Winam Gulf, Lake Victoria, Kenya. Applied and Environmental Microbiology. 2025 Jan 8;0(0):01507–24.

5. Liu Y, Zhu G, Fan Y, Guo C. Successional conditions of *Dolichospermum* and *Microcystis* in Taihu Lake, China. J Ocean Limnol. 2024 Nov 1;42(6):1777–88.

6. Wan L, Chen X, Deng Q, Yang L, Li X, Zhang J, et al. Phosphorus strategy in bloom-forming cyanobacteria (*Dolichospermum* and *Microcystis*) and its role in their succession. Harmful Algae. 2019 Apr 1;84:46–55.

7. Harke MJ, Berry DL, Ammerman JW, Gobler CJ. Molecular response of the bloom-forming cyanobacterium, *Microcystis aeruginosa*, to phosphorus limitation. Microb Ecol. 20110701st ed. 2012 Jan;63(1):188–98.

8. Harke MJ, Gobler CJ. Global transcriptional responses of the toxic cyanobacterium, *Microcystis aeruginosa*, to nitrogen stress, phosphorus stress, and growth on organic matter. PLoS One. 20130723rd ed. 2013;8(7):e69834.

9. Paerl HW, Hall NS, Calandrino ES. Controlling harmful cyanobacterial blooms in a world experiencing anthropogenic and climatic-induced change. Science of The Total Environment. 2011 Apr 15;409(10):1739–45.

10. Deng D, Meng H, Ma Y, Guo Y, Wang Z, He H, et al. The cumulative impact of temperature and nitrogen availability on the potential nitrogen fixation and extracellular polymeric substances secretion by *Dolichospermum*. Harmful Algae. 2024 May 3;102633.

11. Ger KA, Urrutia-Cordero P, Frost PC, Hansson LA, Sarnelle O, Wilson AE, et al. The interaction between cyanobacteria and zooplankton in a more eutrophic world. Harmful Algae. 2016 Apr;54:128–44.

12. He Y, Chen Y, Tao H, Zhou X, Liu J, Liu Y, et al. Secondary metabolites from cyanobacteria: source, chemistry, bioactivities, biosynthesis and total synthesis. Phytochem Rev. 2024 Apr 22;

13. Janssen EML. Cyanobacterial peptides beyond microcystins – A review on co-occurrence, toxicity, and challenges for risk assessment. Water Research. 2019 Mar 15;151:488–99.

14. Zastepa A, Westrick JA, Miller TR, Liang A, Szlag DC, Chaffin JD. The Lake Erie Harmful Algal Blooms Grab: High-resolution mapping of toxic and bioactive metabolites (cyanotoxins/cyanopeptides) in cyanobacterial harmful algal blooms within the western basin. Aquatic Ecosystem Health & Management. 2024 Jan;27(1).

15. Portmann C, Blom JF, Gademann K, Jüttner F. Aerucyclamides A and B: Isolation and synthesis of toxic ribosomal heterocyclic peptides from the Cyanobacterium *Microcystis aeruginosa* PCC 7806. J Nat Prod. 2008 Jul 1;71(7):1193–6.

16. Pawlik-Skowrońska B, Bownik A. Synergistic toxicity of some cyanobacterial oligopeptides to physiological activities of *Daphnia magna* (Crustacea). Toxicon. 20211220th ed. 2022 Jan 30;206:74–84.

17. Österholm J, Popin RV, Fewer DP, Sivonen K. Phylogenomic analysis of secondary metabolism in the toxic cyanobacterial genera *Anabaena, Dolichospermum*, and *Aphanizomenon*. Toxins. 2020 Apr 11;12(4):248.

18. Juhel G, Ramsay RM, Davenport J, O'Halloran J&, Culloty SC. Effect of the microcystin-producing cyanobacterium, *Microcystis aeruginosa*, on immune functions of the zebra mussel *Dreissena polymorpha*. shre. 2015 Aug;34(2):433–42.

19. Pawlik-Skowrońska B, Toporowska M, Mazur-Marzec H. Effects of secondary metabolites produced by different cyanobacterial populations on the freshwater zooplankters *Brachionus calyciflorus* and *Daphnia pulex*. Environ Sci Pollut Res Int. 2019 Apr;26(12):11793–804.

20. Torres M de A, Dax A, Grand I, vom Berg C, Pinto E, Janssen EML. Lethal and behavioral effects of semi-purified microcystins, Micropeptin and apolar compounds from cyanobacteria on freshwater microcrustacean *Thamnocephalus platyurus*. Aquatic Toxicology. 2024 Aug 1;273:106983.

21. Al-Tebrineh J, Mihali TK, Pomati F, Neilan BA. Detection of saxitoxin-producing cyanobacteria and *Anabaena circinalis* in environmental water blooms by quantitative PCR. Appl Environ Microbiol. 2010 Dec;76(23):7836–42.

22. Otten TG, Xu H, Qin B, Zhu G, Paerl HW. Spatiotemporal patterns and ecophysiology of toxigenic *Microcystis* blooms in Lake Taihu, China: Implications for water quality management. Environ Sci Technol. 2012 Mar 20;46(6):3480–8.

23. Yancey CE, Hart L, Hefferan S, Mohamed OG, Newmister SA, Tripathi A, et al. Metabologenomics reveals strain-level genetic and chemical diversity of *Microcystis* secondary metabolism. mSystems. 2024 Jun 25;0(0):e00334-24.

24. Berry MA, Davis TW, Cory RM, Duhaime MB, Johengen TH, Kling GW, et al. Cyanobacterial harmful algal blooms are a biological disturbance to Western Lake Erie bacterial communities. Environ Microbiol. 20170216th ed. 2017 Mar;19(3):1149–62.

25. Tanvir RU, Hu Z, Zhang Y, Lu J. Cyanobacterial community succession and associated cyanotoxin production in hypereutrophic and eutrophic freshwaters. Environmental Pollution. 2021 Dec 1;290:118056.

26. Kharbush JJ, Robinson RS, Carter SJ. Patterns in sources and forms of nitrogen in a large eutrophic lake during a cyanobacterial harmful algal bloom. Limnology & Oceanography. 2023 Apr;68(4):803–15.

27. Lezcano MÁ, Agha R, Cirés S, Quesada A. Spatial-temporal survey of *Microcystis* oligopeptide chemotypes in reservoirs with dissimilar waterbody features and their relation to genetic variation. Harmful Algae. 2019 Jan 1;81:77–85.

28. Agha R, Àngeles Lezcano M, del Mar Labrador M, Cirés Samuel, Quesada A. Seasonal dynamics and sedimentation patterns of *Microcystis* oligopeptide-based chemotypes reveal subpopulations with different ecological traits. Limnology and Oceanography. 2014;59(3):861–71.

29. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualization. Nucleic Acids Research. 2023 Jul 5;51(W1):W46–50.

30. Terlouw BR, Blin K, Navarro-Muñoz JC, Avalon NE, Chevrette MG, Egbert S, et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. Nucleic Acids Research. 2023 Jan 6;51(D1):D603–10.

31. Nothias LF, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A, et al. Feature-based molecular networking in the GNPS analysis environment. Nat Methods. 2020 Sep;17(9):905–8.

32. Boegehold AG, Burtner AM, Camilleri AC, Carter G, DenUyl P, Fanslow D, et al. Routine monitoring of western Lake Erie to track water quality changes associated with cyanobacterial harmful algal blooms. Earth System Science Data. 2023 Aug 25;15(8):3853–68.

33. Cory RM, Davis TW, Dick GJ, Johengen T, Denef VJ, Berry MA, et al. Seasonal dynamics in dissolved organic matter, hydrogen peroxide, and cyanobacterial blooms in Lake Erie. Front Mar Sci. 2016 Apr 26;3.

34. Lopez AE, Laughrey ZR, Loftin KA. Qualitative Analysis of Cyanotoxins in Lake Erie samples using liquid chromatography/high resolution mass spectrometry (LC/HRMS) from July 2016 to October 2022: U.S. Geological Survey data release. 2025.

35. Den Uyl PA, Kiledal EA, Errera RM, Chaganti SR, Godwin CM, Raymond HA, et al. Genomic identification and characterization of saxitoxin producing cyanobacteria in western Lake Erie harmful algal blooms. Environ Sci Technol. 2025 Apr 10;

36. Woodcroft BJ. CoverM: Read coverage calculator for metagenomics [Internet]. [cited 2022 Jun 6]. Available from: https://github.com/wwood/CoverM

37. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018 Sep 15;34(18):3094–100.

38. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. The ISME Journal. 2017 Dec 1;11(12):2864–8.

39. Liaw A, Wiener M. Classification and Regression by "randomForest." R News. 2002;2(3):18–22.

40. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, et al. Minimum Information about a Biosynthetic Gene cluster. Nature Chemical Biology. 2015 Sep 1;11(9):625–31.

41. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010 May;28(5):511–5.

42. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available from: http://www.R-project.org/

43. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research. 2021 Jul 2;49(W1):W293–6.

44. Wood S. R package: “mgcv”: Generalized additive (mixed) models, some of their extensions and other generalized ridge regression with multiple smoothing parameter estimation [Internet]. 2023. Available from: https://doi.org/10.1201/9781315370279

45. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009 Dec 15;10(1):421.

46. Lee MD. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics. 2019 Oct 15;35(20):4162–4.

47. Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE. 2010;5(3):e9490.

48. Yancey CE, Kiledal EA, Chaganti SR, Denef VJ, Errera RM, Evans JT, et al. The Western Lake Erie culture collection: A promising resource for evaluating the physiological and genetic diversity of *Microcystis* and its associated microbiome. Harmful Algae. 2023 Jul;126:102440.

49. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol. 2012 Oct;30(10):918–20.

50. Schmid R, Heuckeroth S, Korf A, Smirnov A, Myers O, Dyrlund TS, et al. Integrative analysis of multimodal mass spectrometry data in MZmine 3. Nat Biotechnol. 2023 Apr;41(4):447–9.

51. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environ Sci Technol. 2014 Feb 18;48(4):2097–8.

52. McDonald K, DesRochers N, Renaud JB, Sumarah MW, McMullin DR. Metabolomics reveals strain-specific cyanopeptide profiles and their production dynamics in *Microcystis aeruginosa* and *M. flos-aquae*. Toxins. 2023 Mar 31;15(4):254.

53. Pakkir Shah AK, Walter A, Ottosson F, Russo F, Navarro-Diaz M, Boldt J, et al. Statistical analysis of feature-based molecular networking results from non-targeted metabolomics data. Nat Protoc. 2025 Jan;20(1):92–162.

54. Allaire J. RStudio: Integrated Development Environment for R. Boston, U.S.A: R Studio Inc.; 2015.

55. Kassambara A. R Package: ggpubr:“ggplot2” based publication ready plots. 2018.

56. Simspon G. R Package: “gratia”. ggplot-based graphics and other useful functions for GAMs fitted using mgcv. 2018.

57. NOAA. NOAA WE2 Nutrient Buoy Data from 2015-2023 [Internet]. Great Lakes Environmental Research Laboratory; [cited 2025 Jul 1]. Available from: https://seagull-erddap.glos.org/erddap/tabledap/GLERLWE2_archive.html

58. Larsen J. Type III polyketide biosynthesis in cyanobacteria [Internet]. School of Environmental and Life Sciences, University of Newcastle, Australia; 2023 [cited 2025 Mar 3]. Available from: https://nova.newcastle.edu.au/vital/access/services/Download/uon:50347/ATTACHMENT01

59. Harke MJ, Davis TW, Watson SB, Gobler CJ. Nutrient-controlled niche differentiation of western Lake Erie cyanobacterial populations revealed via metatranscriptomic surveys. Environmental Science & Technology. 2016 Jan 19;50(2):604–15.

60. Michalak AM, Anderson EJ, Beletsky D, Boland S, Bosch NS, Bridgeman TB, et al. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proceedings of the National Academy of Sciences. 2013 Apr 16;110(16):6448–52.

61. Yancey CE, Mathiesen O, Dick GJ. Transcriptionally active nitrogen fixation and biosynthesis of diverse secondary metabolites by *Dolichospermum* and *Aphanizomenon*-like cyanobacteria in western Lake Erie *Microcystis* blooms. Harmful Algae. 2023 May 1;124:102408.

62. Munawar M, Munawar IF. A lakewide study of phytoplankton biomass and its species composition in Lake Erie, April–December 1970. J Fish Res Bd Can. 1976 Mar;33(3):581–600.

63. Wang Y, Chen F. Decomposition and phosphorus release from four different size fractions of *Microcystis* spp. taken from Lake Taihu, China. Journal of Environmental Sciences. 2008 Jan 1;20(7):891–6.

64. Hanna E, Francis J, Wang M, Overland JE, Cohen J, Luo D, et al. Influence of high-latitude blocking and the northern stratospheric polar vortex on cold-air outbreaks under Arctic amplification of global warming. Environ Res: Climate. 2024 Dec;3(4):042004.

65. Wenski SL, Thiengmag S, Helfrich EJN. Complex peptide natural products: Biosynthetic principles, challenges and opportunities for pathway engineering. Synthetic and Systems Biotechnology. 2022 Mar 1;7(1):631–47.

66. Walsh CT, O’Brien RV, Khosla C. Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. Angew Chem Int Ed Engl. 2013 Jul 8;52(28):7098–124.

67. Chaffin JD, Westrick JA, Reitz LA, Bridgeman TB. Microcystin congeners in Lake Erie follow the seasonal pattern of nitrogen availability. Harmful Algae. 2023 Aug 1;127:102466.

68. Van de Waal DB, Smith VH, Declerck SAJ, Stam ECM, Elser JJ. Stoichiometric regulation of phytoplankton toxins. Ecology Letters. 2014;17(6):736–42.

69. Wagner ND, Osburn FS, Wang J, Taylor RB, Boedecker AR, Chambliss CK, et al. Biological Stoichiometry Regulates Toxin Production in *Microcystis aeruginosa* (UTEX 2385). Toxins. 2019;11(10).

70. Hoffman DK, McCarthy MJ, Boedecker AR, Myers JA, Newell SE. The role of internal nitrogen loading in supporting non-N-fixing harmful cyanobacterial blooms in the water column of a large eutrophic lake. Limnology and Oceanography. 2022;67(9):2028–41.

71. Hampel JJ, McCarthy MJ, Gardner WS, Zhang L, Xu H, Zhu G, et al. Nitrification and ammonium dynamics in Taihu Lake, China: Seasonal competition for ammonium between nitrifiers and cyanobacteria. Biogeosciences. 2018 Feb 6;15(3):733–48.

72. McCarthy MJ, Gardner WS, Lehmann MF, Bird DF. Implications of water column ammonium uptake and regeneration for the nitrogen budget in temperate, eutrophic Missisquoi Bay, Lake Champlain (Canada/USA). Hydrobiologia. 2013 Nov;718(1):173–88.

73. Kosakowska A, Nedzi M, Pempkowiak J. Responses of the toxic cyanobacterium *Microcystis aeruginosa* to iron and humic substances. Plant Physiol Biochem. 2007 May;45(5):365–70.

74. Shende VV, Bauman KD, Moore BS. The shikimate pathway: gateway to metabolic diversity. Nat Prod Rep. 2024 Apr 24;41(4):604–48.

75. Chen J, Zhou Y, Zhang Y. New insights into microbial degradation of cyanobacterial organic matter using a fractionation procedure. International Journal of Environmental Research and Public Health. 2022 Jan;19(12):6981.

76. Große R, Heuser M, Teikari JE, Ramakrishnan DK, Abdelfattah A, Dittmann E. Microcystin shapes the *Microcystis* phycosphere through community filtering and by influencing cross-feeding interactions. ISME Commun. 2024 Dec 24;5(1):170.

77. Wang K, Mou X, Cao H, Struewing I, Allen J, Lu J. Co-occurring microorganisms regulate the succession of cyanobacterial harmful algal blooms. Environmental Pollution. 2021 Nov 1;288:117682.

78. Li W, Baliu-Rodriguez D, Premathilaka SH, Thenuwara SI, Kimbrel JA, Samo TJ, et al. Microbiome processing of organic nitrogen input supports growth and cyanotoxin

production of *Microcystis aeruginosa* cultures. The ISME Journal. 2024 Jan 1;18(1):wrae082.

79. Ger KA, Faassen EJ, Pennino MG, Lürling M. Effect of the toxin (microcystin) content of *Microcystis* on copepod grazing. Harmful Algae. 2016 Feb 1;52:34–45.

80. Rohrlack T, Christoffersen K, Kaebernick M, Neilan BA. Cyanobacterial protease inhibitor microviridin J causes a lethal molting disruption in *Daphnia pulicaria*. Applied and Environmental Microbiology. 2004 Aug;70(8):5047–50.

81. Pereira DA, Giani A. Cell density-dependent oligopeptide production in cyanobacterial strains. FEMS Microbiology Ecology. 2014;88(1):175–83.

82. Sukenik A, Eshkol R, Livne A, Hadas O, Rom M, Tchernov D, et al. Inhibition of growth and photosynthesis of the dinoflagellate *Peridinium gatunense* by *Microcystis* sp. (cyanobacteria): A novel allelopathic mechanism. Limnology and Oceanography. 2002;47(6):1656–63.

83. Zilliges Y, Kehr JC, Meissner S, Ishida K, Mikkat S, Hagemann M, et al. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of *Microcystis* under oxidative stress conditions. PLoS One. 2011 Mar 18;6(3):e17615.

84. Krausfeldt LE, Farmer AT, Castro HF, Boyer GL, Campagna SR, Wilhelm SW. Nitrogen flux into metabolites and microcystins changes in response to different nitrogen sources in *Microcystis aeruginosa* NIES-843. Environ Microbiol. 2020 Jun;22(6):2419–31.

85. Vanderploeg HA, Liebig JR, Carmichael WW, Agy MA, Johengen TH, Fahnenstiel GL, et al. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. Can J Fish Aquat Sci. 2001 Jun;58(6):1208–21.

86. Boegehold AG, Glyshaw P, Vanderploeg HA, Errera R. *Microcystis* strains in Lake Erie explain interactions between a selective filter feeder and the phytoplankton community. Hydrobiologia. 2025 Apr 2;1–16.

87. Urrutia-Cordero P, Agha R, Cirés S, Lezcano MÁ, Sánchez-Contreras M, Waara KO, et al. Effects of harmful cyanobacteria on the freshwater pathogenic free-living amoeba *Acanthamoeba castellanii*. Aquatic Toxicology. 2013 Apr;130–131:9–17.

2.8 Figures

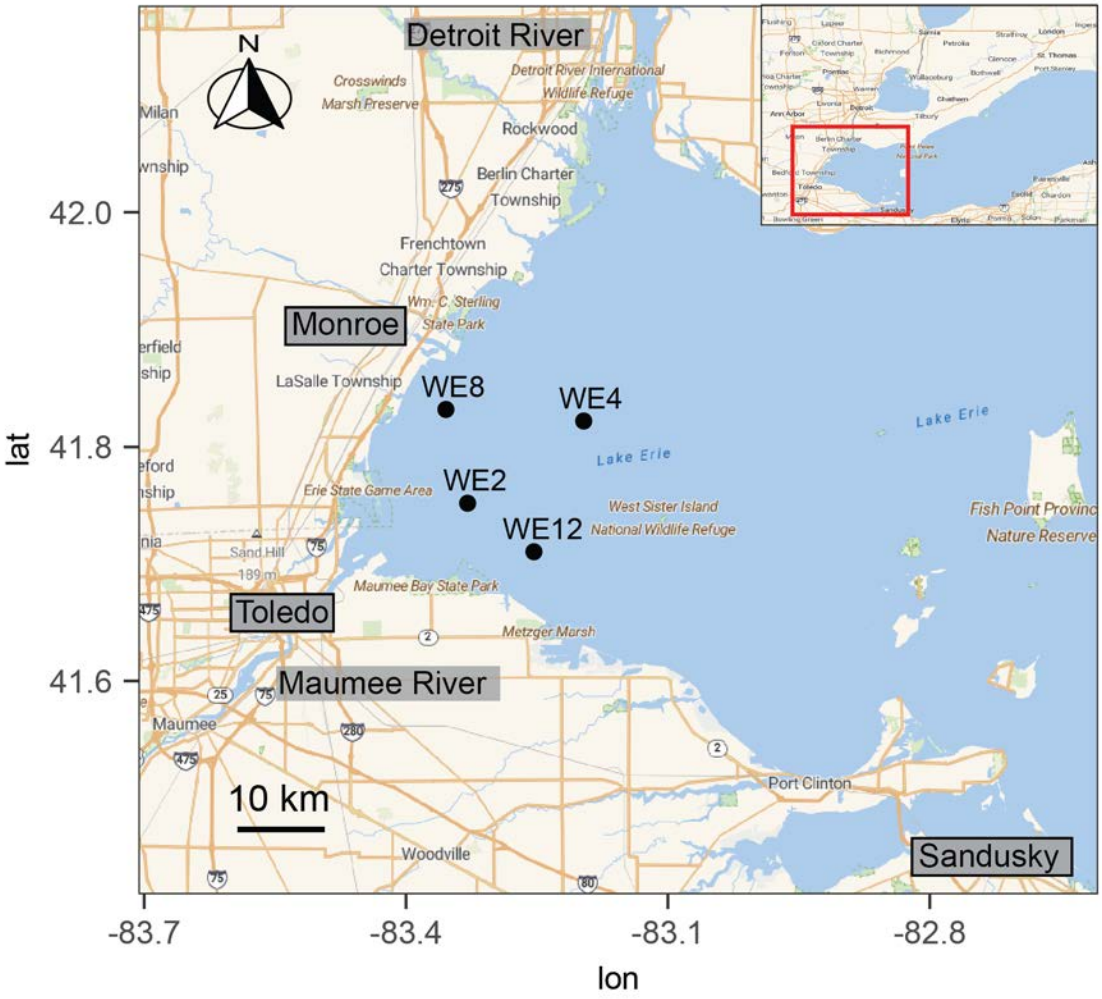


Figure 2.8.1: Map of western Lake Erie showing NOAA long-term monitoring stations evaluated in this study and nearby rivers and cities. Near-shore sites WE2 and WE12 border the Ohio coastline, WE8 lies off the Michigan shoreline, and WE4 represents an offshore station.

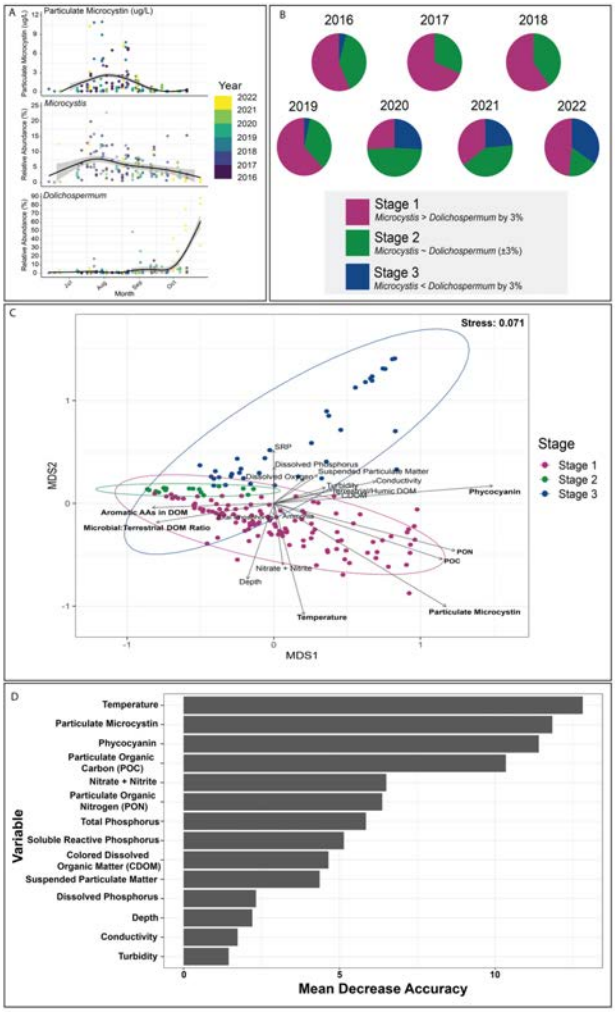


Figure 2.8.2: (a) Temporal patterns of particulate microcystin (µg/L) and the relative abundance of *Microcystis* and *Dolichospermum.* All individual measurements are shown as points colored by year. Particulate microcystin concentrations are shown as a black smoothed line fitted using locally weighted regression (loess). *Microcystis* and *Dolichospermum* abundance are also black smoothed lines fit using generalized additive models (GAMs) with a smooth on DOY with a cubic regression function. **(b)** Samples were categorized based on their relative abundance differences between *Microcystis* and *Dolichospermum.* Pie charts show the distribution of samples in each stage for each year. Sample size (n) in stage 1, 2, and 3, are 85, 61, and 24, respectively. **(c)** Nonmetric MultiDimensional Scaling (NMDS) ordination of samples based on cyanobacterial community (genus level), with points colored by stage. Ellipses represent 95% confidence intervals around samples grouped by stages 1, 2, and 3. Arrows indicate the direction and strength of environmental variables, with longer arrows representing strong correlations with NMDS axes, and bolded variables indicating key drivers of ordination structure *(P*<0.05). Abbreviations: Soluble reactive phosphorus (SRP), Particulate Organic Nitrogen (PON), Particulate Organic Carbon (POC), Colored dissolved organic matter (CDOM). **(d)** Random forest abiotic environmental variable importance based on mean decrease in accuracy for predicting stage, where bars represent the relative importance of each abiotic variable in classifying samples into stages.

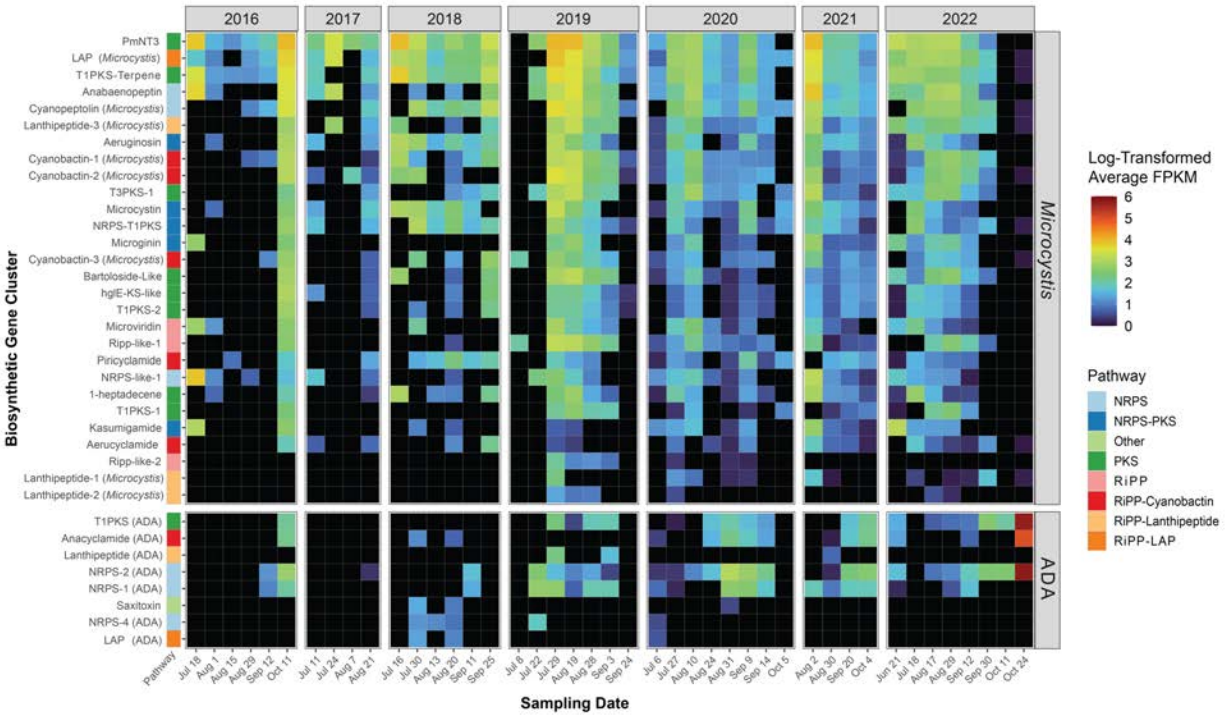


Figure 2.8.3: Relative abundance of BGCs from *Microcystis* and the ADA clade. Tiles are colored by log-transformed average fragments per kilobase million (FPKM) for each BGC across sampling dates within each year, with warmer colors indicating higher relative abundance. Black squares denote values of zero. BGC pathways are indicated by the color of labels on the horizontal axis.

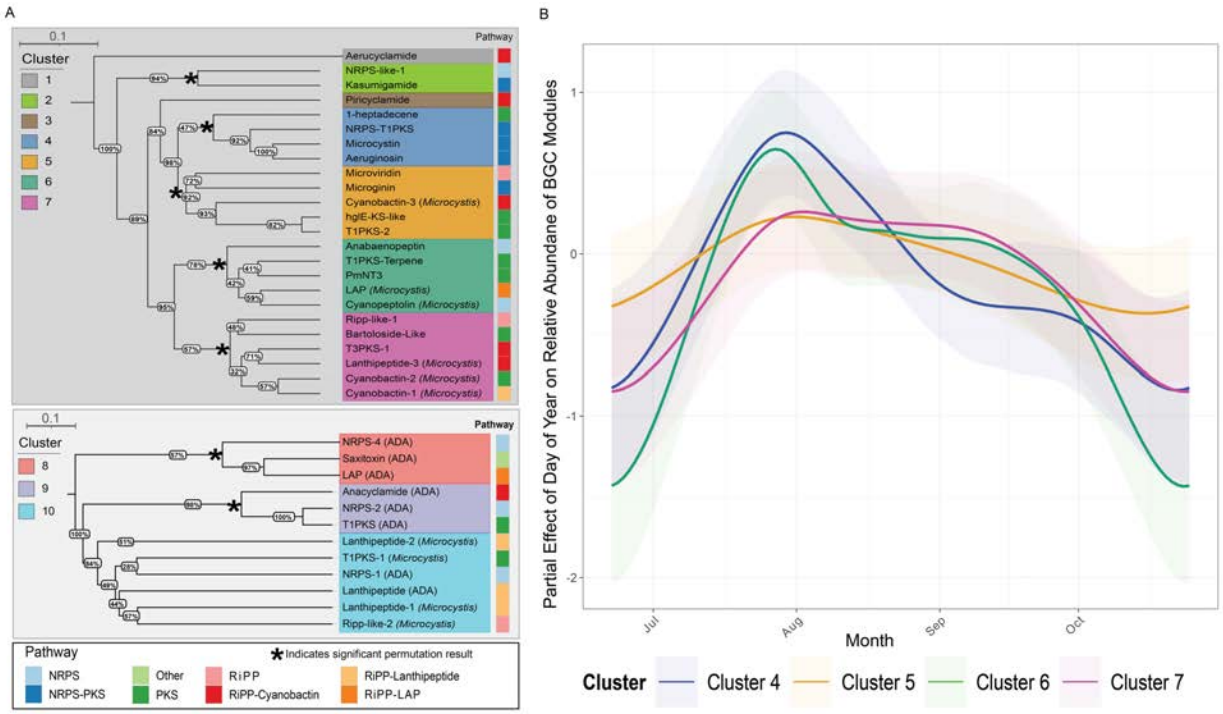


Figure 2.8.4: (a) Biosynthesis modules determined by hierarchical clustering (average linkage) of temporal distributions of relative abundance of BGCs based on Bray-Curtis distance. The dark gray subtree contains *Microcystis* BGCs that occur consistently across samples, whereas the light gray subtree contains ADA BGCs and sporadically occurring *Microcystis* BGCs (zero observations in two or more years of the dataset). BGC pathways are color-coded according to the legend. An asterisk (*) denotes significantly supported modules based on permutation-derived clustering quality. Scale bars for trees are shown, where the length is equal to a branch length of 0.1. **(b)** Partial effects of DOY from module-specific HGAMs with the log-transformed average FPKM of BGCs from modules 4, 5, 6, and 7 as the response. Shaded ribbons indicate standard errors around the fitted temporal trends.

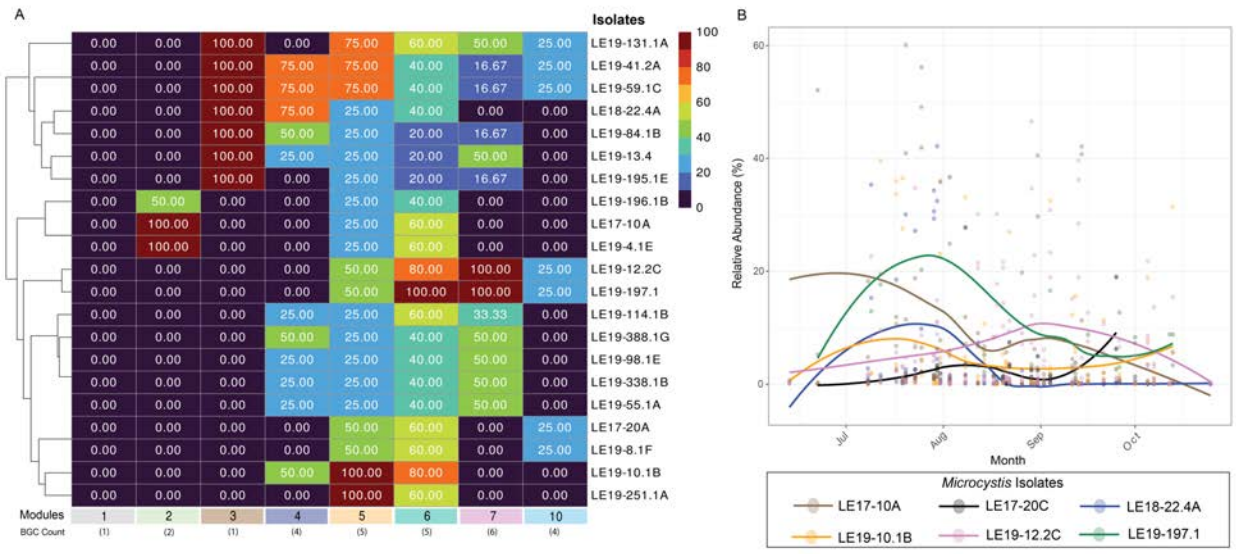


Figure 2.8.5: (a) Percent of each biosynthesis module present in isolates from the WLECC. Isolates are hierarchically clustered using average-linkage. Tiles are filled in by the percent of each biosynthesis module found in each isolate, with brighter tiles indicating more of the module present in that isolate. **(b)** Relative abundance of six WLECC isolates across field samples. Points represent individual observations, relative to the whole microbial community, and smoothed lines were fitted using a loess function.

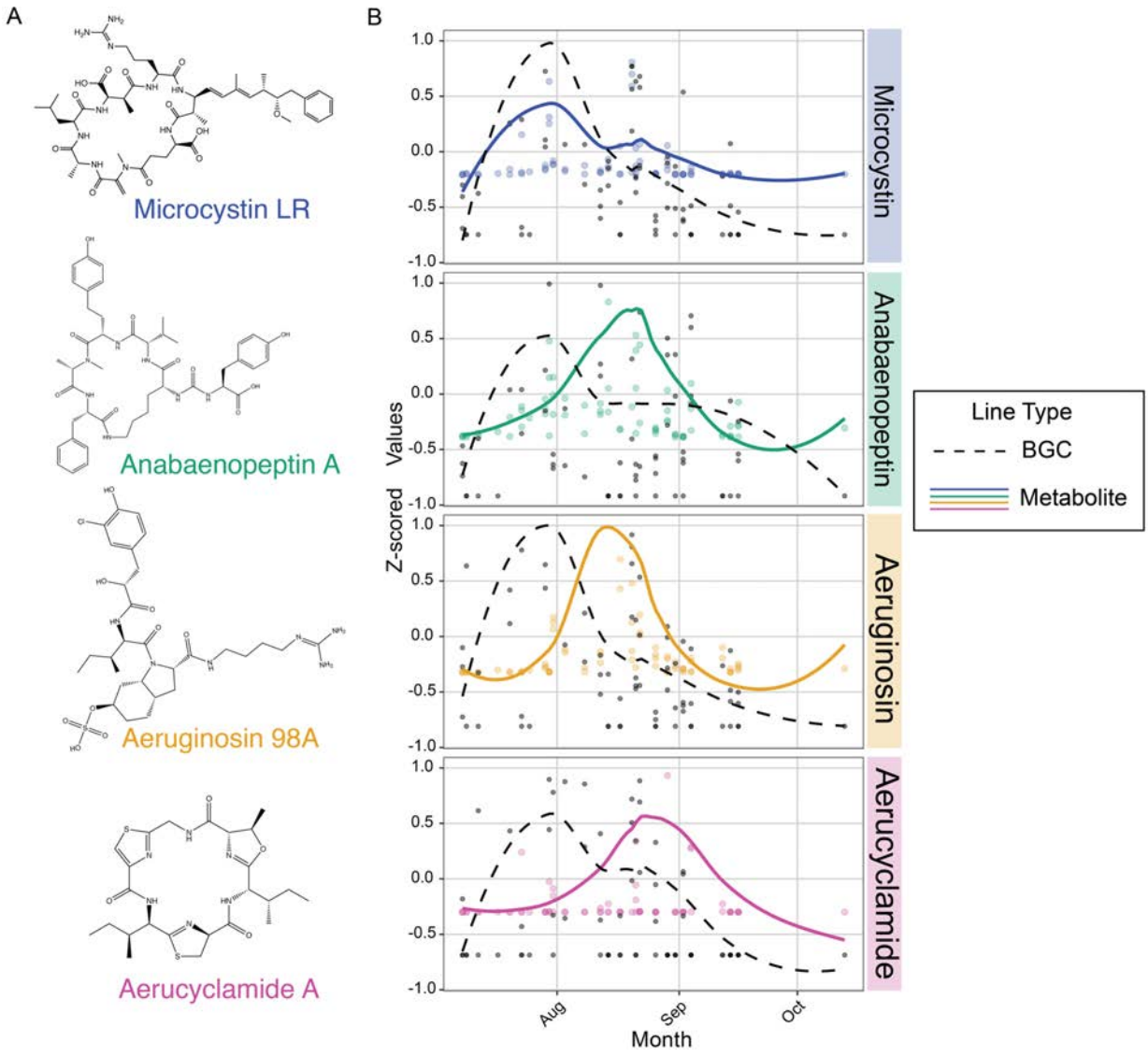


Figure 2.8.6: (a) Representative chemical structures of putatively annotated metabolites across the dataset: Microcystin LR, Anabaenopeptin A, Aeruginosin 98A, and Aerucyclamide A. **(b)** Temporal distribution of the normalized abundance of metabolites (Microcystin = sum of MC-LR, MC-YR, MC-HilR, MC-RR, and MC-HtyR; Aeruginosin 98A; Anabaenopeptin = sum of AP-A and AP-B; and Aerucyclamide A), plotted with the average FPKM of associated BGCs. Data were Z-score scaled with all individual observations less than or equal to one shown (6 observations excluded), where black dots correspond to BGC abundance and colored dots correspond to metabolite abundance. Smoothed lines for BGC and metabolite abundance, represented as black dashed and filled colored lines, respectively, were fitted using a loess function.

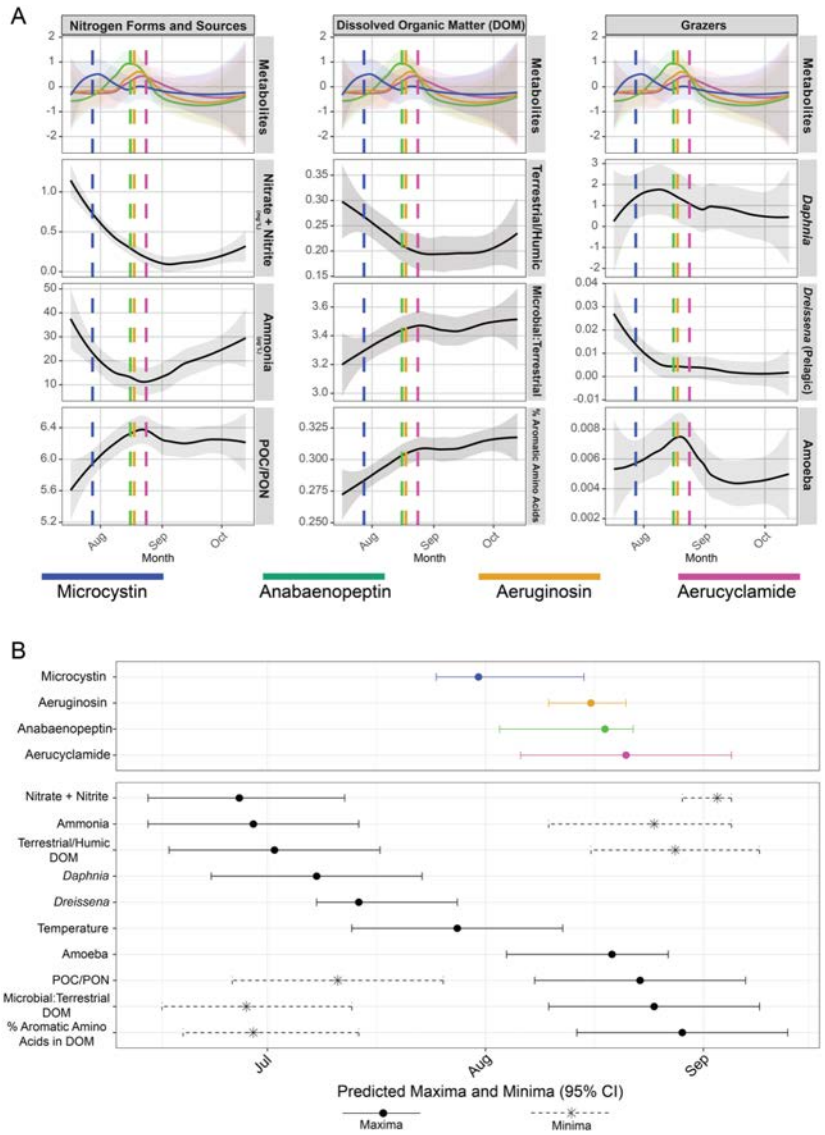


Figure 2.8.7: (a) Temporal patterns of the normalized abundance of metabolites (Microcystin = sum of MC-LR, MC-YR, MC-HilR, MC-RR, and MC-HtyR; Aeruginosin 98A; Anabaenopeptin = sum of AP-A and AP-B; and Aerucyclamide A) plotted with nitrogen forms/sources, DOM components, and grazer abundance. Metabolite data were Z-score scaled and plotted together, whereas all other variables were plotted in their original scales. Scales for each variable are as follows: Ammonia (µg L^{-1}), Nitrate + Nitrate (mg L^{-1}), *Daphnia*, *Dreissena*, and Amoeba (% Relative Abundance). Temporal patterns were visualized using a loess function. Dashed lines indicate predicted peak dates for each metabolite and are colored accordingly. Abbreviations: Particulate organic carbon: particulate organic nitrogen (POC/PON). **(b)** Predicted maxima and minima for environmental variables and metabolites were estimated using 1,000 bootstrapped GAMs with a cyclic smoother on DOY. Median peak dates and 95% confidence intervals are shown as points with horizontal bars. Maxima are indicated by solid lines with circles, whereas minima, predicted only for nitrogen- and DOM-related variables, are shown as dashed lines with stars. Metabolite maxima are shown as solid-colored lines in the top panel separated from the environmental variable predictions.

2.9 Supplementary Information

2.9.1 SI Tables

Table S2.9.1: Metadata for all samples used in this study, including sampling and physicochemical parameter information.

GLAMR_ SampleID	NCBI BioProject	NCBI Biosample	NCBI Accession	Chlorophyll A (ug/L)	MASSiVE DB ID	C1 (RU)	C2 (RU)	C3 (RU)	C4 (RU)	C5 (RU)	NOAA_Site	collection_date	Metabolomics volume filtered (mL)
samp_2042	PRJNA1190886	SAMN45060126	SRR31514654	4.42	MSV000097661	NA	NA	NA	NA	NA	WE4	2016-07-18T12:23:00Z	300
samp_2043	PRJNA1190886	SAMN45060127	SRR31514643	16.28	NA	NA	NA	NA	NA	NA	WE8	2016-07-18T13:12:00Z	100
samp_2044	PRJNA1190886	SAMN45060128	SRR31514728	7.93	NA	NA	NA	NA	NA	NA	WE12	2016-07-18T09:50:00Z	NA
samp_2045	PRJNA1190886	SAMN45060129	SRR31514716	22.69	MSV000097661	0.22351001	0.40080213	0.16113691	0.20244467	0.16762341	WE2	2016-08-01T08:25:00Z	75
samp_2047	PRJNA1190886	SAMN45060131	SRR31514666	15.3	NA	NA	NA	NA	NA	NA	WE8	2016-08-01T13:16:00Z	100
samp_2048	PRJNA1190886	SAMN45060132	SRR31514800	25.53	NA	0.21895659	0.41970956	0.17483684	0.18326945	0.18988966	WE12	2016-08-01T10:06:00Z	NA
samp_2049	PRJNA1190886	SAMN45060133	SRR31514798	19.64	MSV000097661	0.15655251	0.26818767	0.12207815	0.13894522	0.1372291	WE2	2016-08-15T08:35:00Z	80
samp_2051	PRJNA1190886	SAMN45060135	SRR31514794	46.33	NA	NA	NA	NA	NA	NA	WE8	2016-08-15T14:00:00Z	25
samp_2052	PRJNA1190886	SAMN45060136	SRR31514770	18.81	MSV000097661	0.13779923	0.23732496	0.10298506	0.12225352	0.09427645	WE12	2016-08-15T10:34:00Z	80
samp_2053	PRJNA1190886	SAMN45060137	SRR31514661	11.47	MSV000097661	0.15207383	0.27301959	0.11693199	0.17881542	0.14418313	WE2	2016-08-29T08:20:00Z	75
samp_2054	PRJNA1190886	SAMN45060138	SRR31514660	9.97	MSV000097661	0.17679072	0.29801903	0.12932853	0.20436449	0.16424687	WE4	2016-08-29T11:30:00Z	125
samp_2055	PRJNA1190886	SAMN45060139	SRR31514659	22.88	NA	NA	NA	NA	NA	NA	WE8	2016-08-29T12:20:00Z	NA
samp_2056	PRJNA1190886	SAMN45060140	SRR31514658	10.35	MSV000097661	NA	NA	NA	NA	NA	WE12	2016-08-29T10:05:00Z	80
samp_2057	PRJNA1190886	SAMN45060141	SRR31514657	22.16	MSV000097661	NA	NA	NA	NA	NA	WE2	2016-09-12T08:25:00Z	50
samp_2058	PRJNA1190886	SAMN45060142	SRR31514656	7.97	MSV000097661	NA	NA	NA	NA	NA	WE4	2016-09-12T12:29:00Z	150

samp_2059	PRJNA1190886	SAMN45060143	SRR31514655	10.85	NA	NA	NA	NA	NA	NA	WE8	2016-09-12T13:12:00Z	100
samp_2060	PRJNA1190886	SAMN45060144	SRR31514652	4.77	MSV000097661	0.15005652	0.25825964	0.11607598	0.13319336	0.10914125	WE12	2016-09-12T10:04:00Z	100
samp_2061	PRJNA1190886	SAMN45060145	SRR31514651	7.69	MSV000097661	0.22479311	0.38068094	0.17055193	0.21601059	0.10782203	WE2	2016-10-11T08:25:00Z	70
samp_2062	PRJNA1190886	SAMN45060146	SRR31514650	2.77	NA	NA	NA	NA	NA	NA	WE4	2016-10-11T12:50:00Z	NA
samp_2065	PRJNA1190886	SAMN45060148	SRR31514648	9.17	NA	0.55447537	0.96942573	0.40719774	0.36020969	0.18672574	WE2	2017-07-11T13:13:00Z	NA
samp_2067	PRJNA1190886	SAMN45060150	SRR31514646	2.28	NA	NA	NA	NA	NA	NA	WE8	2017-07-11T08:12:00Z	200
samp_2068	PRJNA1190886	SAMN45060151	SRR31514645	8.88	MSV000097661	0.19362998	0.33681004	0.13094566	0.16264632	0.1216456	WE12	2017-07-11T11:33:00Z	100
samp_2070	PRJNA1190886	SAMN45060153	SRR31514641	16.71	MSV000097661	0.26091395	0.45386735	0.17712924	0.19115258	0.18424208	WE4	2017-07-24T09:07:00Z	100
samp_2073	PRJNA1190886	SAMN45060156	SRR31514638	21.52	MSV000097661	0.44477503	0.80840778	0.3374859	0.31137765	0.21319479	WE2	2017-08-07T13:19:00Z	100
samp_2074	PRJNA1190886	SAMN45060157	SRR31514735	7.33	MSV000097661	0.2632535	0.46698889	0.18007316	0.21376804	0.11016626	WE4	2017-08-07T09:00:00Z	100
samp_2075	PRJNA1190886	SAMN45060158	SRR31514734	23.1	NA	NA	NA	NA	NA	NA	WE8	2017-08-07T08:15:00Z	NA
samp_2077	PRJNA1190886	SAMN45060160	SRR31514732	111.87	MSV000097661	0.35708941	0.65744296	0.27025225	0.26697909	0.21751072	WE2	2017-08-21T13:13:00Z	10
samp_2078	PRJNA1190886	SAMN45060161	SRR31514731	54.72	MSV000097661	0.44145892	0.77727813	0.323784	0.32954006	0.17761858	WE4	2017-08-21T08:59:00Z	10
samp_2079	PRJNA1190886	SAMN45060162	SRR31514730	31.45	NA	NA	NA	NA	NA	NA	WE8	2017-08-21T08:14:00Z	10
samp_2080	PRJNA1190886	SAMN45060163	SRR31514727	22.04	MSV000097661	0.11588003	0.2195184	0.08759668	0.10349796	0.13371699	WE12	2017-08-21T11:33:00Z	10
samp_2082	PRJNA1190886	SAMN45060165	SRR31514725	9.14	MSV000097661	0.23358073	0.44321124	0.17278718	0.24355021	0.09786177	WE4	2018-07-16T09:01:00Z	100
samp_2085	PRJNA1190886	SAMN45060168	SRR31514722	25.61	MSV000097661	0.25210919	0.4643288	0.18198427	0.22844178	0.10482629	WE2	2018-07-30T12:46:00Z	150
samp_2087	PRJNA1190886	SAMN45060170	SRR31514720	25.48	NA	NA	NA	NA	NA	NA	WE8	2018-07-30T08:08:00Z	100
samp_2088	PRJNA1190886	SAMN45060171	SRR31514719	13.13	MSV000097661	0.25701346	0.49431749	0.1851145	0.25546297	0.09920542	WE12	2018-07-30T11:04:00Z	100
samp_2089	PRJNA1190886	SAMN45060172	SRR31514717	19.43	MSV000097661	0.22509486	0.44170442	0.16312827	0.22321405	0.11276842	WE2	2018-08-	100

												13T12:27:00Z	
samp_2091	PRJNA190886	SAMN45060174	SRR31514714	18.85	NA	NA	NA	NA	NA	NA	WE8	2018-08-13T08:06:00Z	100
samp_2092	PRJNA190886	SAMN45060175	SRR31514713	19.56	MSV000097661	0.26743867	0.52737452	0.19641342	0.35013392	0.12755635	WE12	2018-08-13T10:51:00Z	100
samp_2093	PRJNA190886	SAMN45060176	SRR31514712	20.15	MSV000097661	0.21188314	0.40799649	0.15222571	0.22476646	0.11555636	WE2	2018-08-20T13:01:00Z	100
samp_2094	PRJNA190886	SAMN45060177	SRR31514685	12.32	MSV000097661	0.1470105	0.271167	0.10301277	0.14588934	0.09474307	WE4	2018-08-20T09:28:00Z	150
samp_2095	PRJNA190886	SAMN45060178	SRR31514684	33.88	NA	NA	NA	NA	NA	NA	WE8	2018-08-20T08:44:00Z	75
samp_2096	PRJNA190886	SAMN45060179	SRR31514683	15.08	MSV000097661	0.12409002	0.24767309	0.09505436	0.17082765	0.0904503	WE12	2018-08-20T11:21:00Z	150
samp_2097	PRJNA190886	SAMN45060180	SRR31514682	16.96	MSV000097661	0.30886586	0.54704255	0.22799342	0.2905749	0.16202631	WE2	2018-09-11T12:32:00Z	20
samp_2098	PRJNA190886	SAMN45060181	SRR31514680	14.3	MSV000097661	0.20817315	0.36546181	0.1564099	0.17768348	0.12482376	WE4	2018-09-11T09:02:00Z	25
samp_2101	PRJNA190886	SAMN45060184	SRR31514676	12.27	NA	0.19012862	0.342356	0.13950545	0.18676976	0.13231956	WE2	2018-09-25T08:20:00Z	NA
samp_2105	PRJNA190886	SAMN45060188	SRR31514672	28.35	MSV000097661	0.59532945	1.06892543	0.44366878	0.42126003	0.18676232	WE2	2019-07-08T13:04:00Z	100
samp_2107	PRJNA190886	SAMN45060190	SRR31514669	11.18	NA	NA	NA	NA	NA	NA	WE8	2019-07-08T08:26:00Z	150
samp_2109	PRJNA190886	SAMN45060192	SRR31514667	143.23	NA	0.58925174	1.05102803	0.42886006	0.43611264	0.14864482	WE2	2019-07-22T13:20:00Z	NA
samp_2110	PRJNA190886	SAMN45060193	SRR31514665	18.4	MSV000097661	0.22755268	0.43223895	0.16664064	0.19874041	0.10687989	WE4	2019-07-22T09:23:00Z	100
samp_2111	PRJNA190886	SAMN45060194	SRR31514664	49.79	NA	NA	NA	NA	NA	NA	WE8	2019-07-22T08:32:00Z	100
samp_2112	PRJNA190886	SAMN45060195	SRR31514663	13.98	MSV000097661	0.25248731	0.46347601	0.18088552	0.20675937	0.08753306	WE12	2019-07-22T11:26:00Z	75
samp_2113	PRJNA190886	SAMN45060196	SRR31514662	154.05	MSV000097661	0.88124269	1.6038317	0.65625548	0.62560035	0.25289479	WE2	2019-07-29T13:09:00Z	50
samp_2114	PRJNA190886	SAMN45060197	SRR31514711	52.36	MSV000097661	0.19306411	0.38785067	0.14923181	0.18388755	0.12715246	WE4	2019-07-29T09:10:00Z	100
samp_2115	PRJNA190886	SAMN45060198	SRR31514710	42.9	NA	NA	NA	NA	NA	NA	WE8	2019-07-29T08:23:00Z	100
samp_2116	PRJNA190886	SAMN45060199	SRR31514804	49.63	MSV000097661	0.36257989	0.68138225	0.26399731	0.28081493	0.15230117	WE12	2019-07-29T11:24:00Z	50

samp_2117	PRJNA1190886	SAMN45060200	SRR31514803	68.01	MSV000097661	0.27154376	0.4976523	0.20235239	0.23339903	0.16620707	WE2	2019-08-19T12:57:00Z	50
samp_2118	PRJNA1190886	SAMN45060201	SRR31514802	13.36	MSV000097661	0.0768894	0.16160214	0.05629276	0.07135446	0.10060078	WE4	2019-08-19T09:04:00Z	50
samp_2119	PRJNA1190886	SAMN45060202	SRR31514801	59.87	NA	NA	NA	NA	NA	NA	WE8	2019-08-19T08:15:00Z	50
samp_2120	PRJNA1190886	SAMN45060203	SRR31514799	72.35	MSV000097661	0.20179872	0.37442287	0.15274053	0.1661887	0.13200241	WE12	2019-08-19T10:55:00Z	50
samp_2131	PRJNA1190886	SAMN45060204	SRR31514797	52.5	MSV000097661	0.33411807	0.62155533	0.25612887	0.28576267	0.19205044	WE2	2019-08-28T08:59:00Z	50
samp_2132	PRJNA1190886	SAMN45060205	SRR31514796	39.63	MSV000097661	0.19849749	0.37184604	0.14608393	0.1898103	0.17705239	WE4	2019-08-28T11:56:00Z	50
samp_2133	PRJNA1190886	SAMN45060206	SRR31514795	54.02	NA	NA	NA	NA	NA	NA	WE8	2019-08-28T12:20:00Z	40
samp_2134	PRJNA1190886	SAMN45060207	SRR31514773	45.48	MSV000097661	0.33592558	0.61787679	0.25014657	0.28010777	0.19388346	WE12	2019-08-28T09:50:00Z	40
samp_2135	PRJNA1190886	SAMN45060208	SRR31514762	70.64	MSV000097661	0.43950072	0.79784875	0.33869023	0.34256688	0.25219745	WE2	2019-09-03T12:55:00Z	40
samp_2136	PRJNA1190886	SAMN45060209	SRR31514653	45.97	MSV000097661	0.24758528	0.45695634	0.18452425	0.2181774	0.19886744	WE4	2019-09-03T09:13:00Z	40
samp_2136	PRJNA1190886	SAMN45060209	SRR31514653	45.97	MSV000097661	0.24758528	0.45695634	0.18452425	0.2181774	0.19886744	WE4	2019-09-03T09:13:00Z	50
samp_2137	PRJNA1190886	SAMN45060210	SRR31514642	93.65	NA	NA	NA	NA	NA	NA	WE8	2019-09-03T08:13:00Z	40
samp_2138	PRJNA1190886	SAMN45060211	SRR31514729	46.75	MSV000097661	0.31278498	0.57142458	0.23620275	0.26655614	0.20326785	WE12	2019-09-03T11:09:00Z	50
samp_2139	PRJNA1190886	SAMN45060212	SRR31514718	3.83	MSV000097661	0.33640664	0.60492658	0.23508302	0.27333227	0.11490839	WE2	2020-07-06T08:32:00Z	250
samp_2141	PRJNA1190886	SAMN45060213	SRR31514681	10.83	MSV000097661	0.33953204	0.64248968	0.2436509	0.29905456	0.13485634	WE8	2020-07-06T08:00:00Z	250
samp_2142	PRJNA1190886	SAMN45060214	SRR31514670	1.86	MSV000097661	0.19495014	0.37361146	0.14316565	0.18389282	0.09252777	WE12	2020-07-06T09:07:00Z	250
samp_2143	PRJNA1190886	SAMN45060215	SRR31514805	28.47	NA	0.33607884	0.6380365	0.24279384	0.31283637	0.18340581	WE2	2020-07-28T15:45:00Z	50
samp_2144	PRJNA1190886	SAMN45060216	SRR31514759	16.77	NA	0.12782199	0.24855264	0.09088362	0.15120748	0.09752643	WE4	2020-07-28T12:45:00Z	200
samp_2145	PRJNA1190886	SAMN45060217	SRR31514748	14.3	NA	0.23986685	0.45825809	0.17818745	0.2400559	0.13264826	WE8	2020-07-28T12:07:00Z	100
samp_2147	PRJNA1190886	SAMN45060219	SRR31514700	20.48	MSV000097661	0.26344689	0.49909133	0.18666743	0.25479305	0.15802989	WE2	2020-08-	100

												10T11:40:00Z	
samp_2149	PRJNA190886	SAMN45060221	SRR31514776	15.67	MSV000097661	0.20627819	0.3935465	0.15144187	0.21523407	0.13207083	WE8	2020-08-10T08:03:00Z	100
samp_2150	PRJNA190886	SAMN45060222	SRR31514775	21.39	MSV000097661	0.25411204	0.47428203	0.18415987	0.25394673	0.15768387	WE12	2020-08-10T10:20:00Z	100
samp_2151	PRJNA190886	SAMN45060223	SRR31514774	12.92	MSV000097661	0.20048872	0.37048403	0.1473689	0.1839897	0.13131832	WE2	2020-08-24T11:30:00Z	100
samp_2152	PRJNA190886	SAMN45060224	SRR31514772	12.79	MSV000097661	0.05645926	0.12011817	0.04527454	0.04475656	0.06582042	WE4	2020-08-24T08:44:00Z	150
samp_2153	PRJNA190886	SAMN45060225	SRR31514771	43.53	MSV000097661	0.154992	0.29856527	0.11826856	0.14341296	0.11757006	WE8	2020-08-24T08:03:00Z	100
samp_2154	PRJNA190886	SAMN45060226	SRR31514769	11.47	MSV000097661	0.15399634	0.28867295	0.11474058	0.14137771	0.11372068	WE12	2020-08-24T10:15:00Z	100
samp_2155	PRJNA190886	SAMN45060227	SRR31514768	17.36	NA	0.138471	0.27482508	0.10570234	0.12091016	0.1209265	WE2	2020-09-01T15:10:00Z	100
samp_2156	PRJNA190886	SAMN45060228	SRR31514767	7.08	NA	0.04027485	0.07668245	0.02995493	0.01315781	0.06912681	WE4	2020-09-01T12:40:00Z	200
samp_2157	PRJNA190886	SAMN45060229	SRR31514766	18.85	NA	0.15832126	0.32375396	0.12484128	0.16938356	0.14605627	WE8	2020-09-01T12:03:00Z	100
samp_2158	PRJNA190886	SAMN45060230	SRR31514765	17.49	NA	0.16772248	0.32946469	0.12478838	0.17731158	0.13813026	WE12	2020-09-01T14:03:00Z	100
samp_2159	PRJNA190886	SAMN45060231	SRR31514764	17.68	NA	0.16535125	0.3291731	0.13159433	0.17048553	0.15863559	WE2	2020-09-15T16:37:00Z	100
samp_2160	PRJNA190886	SAMN45060232	SRR31514763	6.57	NA	NA	NA	NA	NA	NA	WE4	2020-09-15T13:15:00Z	200
samp_2161	PRJNA190886	SAMN45060233	SRR31514761	26.78	NA	0.17617579	0.34651394	0.13877454	0.18223431	0.14210564	WE8	2020-09-15T12:31:00Z	75
samp_2162	PRJNA190886	SAMN45060234	SRR31514760	25.42	NA	0.21849984	0.40975278	0.16175789	0.20395643	0.15021328	WE12	2020-09-15T15:03:00Z	100
samp_4334	PRJNA190886	SAMN45060235	SRR31514746	10.8	NA	0.32737561	0.57565094	0.22938364	0.26441137	0.25185109	WE2	2019-09-24T12:33:00	NA
samp_4336	PRJNA190886	SAMN45060237	SRR31514696	12.34	NA	NA	NA	NA	NA	NA	WE8	2019-09-24T8:17:00	NA
samp_4337	PRJNA190886	SAMN45060238	SRR31514686	13.46	NA	NA	NA	NA	NA	NA	WE12	2019-09-24T10:56:00	NA
samp_4338	PRJNA190886	SAMN45060239	SRR31514782	22.62	NA	0.10893656	0.21177654	0.08531873	0.08825699	0.1018985	WE2	2020-09-09T11:09:00	NA
samp_4339	PRJNA190886	SAMN45060240	SRR31514781	3.47	NA	0.04311027	0.08775543	0.03585794	0.04169641	0.05500984	WE4	2020-09-09T8:37:00	NA

samp_4340	PRJNA1190886	SAMN45060241	SRR31514780	19.71	NA	0.14246638	0.29476691	0.11608998	0.15611653	0.14146217	WE8	2020-09-09T8:02:00	NA
samp_4341	PRJNA1190886	SAMN45060242	SRR31514793	22.36	NA	0.0816416	0.17162154	0.06750442	0.10251356	0.08361454	WE12	2020-09-09T10:02:00	NA
samp_4342	PRJNA1190886	SAMN45060243	SRR31514789	8	NA	0.1429915	0.27774544	0.10758753	0.14532368	0.12371367	WE2	2020-10-05T11:34:00	NA
samp_4343	PRJNA1190886	SAMN45060244	SRR31514788	6.79	NA	0.1770522	0.33693688	0.13481054	0.17455376	0.15186627	WE8	2020-10-05T8:12:00	NA
samp_4344	PRJNA1190886	SAMN45060245	SRR31514787	9.65	NA	0.12934764	0.26138372	0.10065354	0.15484774	0.12602422	WE12	2020-10-05T10:18:00	NA
samp_4345	PRJNA1190886	SAMN45060246	SRR31514786	7.07	NA	0.09476252	0.15771029	0.07262143	0.09164191	0.0505145	WE2	2021-08-02T11:32:00	NA
samp_4346	PRJNA1190886	SAMN45060247	SRR31514785	6.05	NA	NA	NA	NA	NA	NA	WE4	2021-08-02T8:35:00	NA
samp_4347	PRJNA1190886	SAMN45060248	SRR31514784	9.28	NA	0.27536761	0.42842064	0.21620843	0.18324834	0.09908717	WE8	2021-08-02T7:56:00	NA
samp_4348	PRJNA1190886	SAMN45060249	SRR31514758	24.38	NA	0.39914302	0.58392553	0.31164203	0.20133831	0.08467972	WE12	2021-08-02T10:08:00	NA
samp_4349	PRJNA1190886	SAMN45060250	SRR31514757	26.26	NA	0.28435535	0.42994183	0.21022351	0.20988594	0.10759512	WE2	2021-08-30T11:03:00	NA
samp_4350	PRJNA1190886	SAMN45060251	SRR31514756	14.1	NA	NA	NA	NA	NA	NA	WE4	2021-08-30T8:30:00	NA
samp_4351	PRJNA1190886	SAMN45060252	SRR31514755	28.27	NA	0.30240088	0.45877169	0.22455266	0.21166589	0.12640404	WE8	2021-08-30T7:53:00	NA
samp_4352	PRJNA1190886	SAMN45060253	SRR31514754	21.12	NA	0.15577268	0.23596198	0.11394085	0.1219941	0.08272845	WE12	2021-08-30T9:55:00	NA
samp_4353	PRJNA1190886	SAMN45060254	SRR31514753	5.49	NA	0.36989747	0.55800854	0.27542275	0.28397992	0.15219524	WE2	2021-09-20T11:05:00	NA
samp_4354	PRJNA1190886	SAMN45060255	SRR31514752	7.73	NA	0.13772783	0.21499184	0.10240838	0.12592246	0.09718451	WE4	2021-09-20T8:31:00	NA
samp_4355	PRJNA1190886	SAMN45060256	SRR31514751	18.78	NA	0.27694269	0.41165643	0.20693008	0.21355258	0.15285326	WE8	2021-09-20T7:51:00	NA
samp_4356	PRJNA1190886	SAMN45060257	SRR31514750	1.05	NA	0.23923759	0.35992962	0.18120597	0.17763265	0.1230628	WE12	2021-09-20T10:00:00	NA
samp_4357	PRJNA1190886	SAMN45060258	SRR31514749	16.64	NA	0.30573078	0.45559256	0.22662188	0.23817533	0.13252341	WE2	2021-10-04T12:47:00	NA
samp_4358	PRJNA1190886	SAMN45060259	SRR31514747	13	NA	0.08836079	0.1412398	0.06717622	0.08052981	0.06479799	WE4	2021-10-04T9:36:00	NA
samp_4359	PRJNA1190886	SAMN45060260	SRR31514745	33	NA	0.38755054	0.57678887	0.28984997	0.28941448	0.16808289	WE8	2021-10-	NA

												04T9:18:00	
samp_4360	PRJNA1190886	SAMN45060261	SRR31514744	19.23	NA	0.15226611	0.22893164	0.11420109	0.1265861	0.10116213	WE12	2021-10-04T11:32:00	NA
samp_4361	PRJNA1190886	SAMN45060262	SRR31514743	3.37	NA	0.24945796	0.40772431	0.18740872	0.18882029	0.09742297	WE2	2022-06-21T12:39:00	NA
samp_4362	PRJNA1190886	SAMN45060263	SRR31514742	32.93	NA	0.45095294	0.71865715	0.333426	0.32820848	0.13827611	WE2	2022-07-18T12:24:00	NA
samp_4363	PRJNA1190886	SAMN45060264	SRR31514741	10.59	NA	0.0914507	0.16921562	0.07092479	0.10893242	0.06998942	WE4	2022-07-18T8:45:00	NA
samp_4364	PRJNA1190886	SAMN45060265	SRR31514740	143.46	NA	NA	NA	NA	NA	NA	WE8	2022-07-18T8:01:00	NA
samp_4365	PRJNA1190886	SAMN45060266	SRR31514739	12.04	NA	0.22903198	0.37614953	0.1763906	0.19276742	0.08201778	WE12	2022-07-18T10:34:00	NA
samp_4366	PRJNA1190886	SAMN45060267	SRR31514738	125.49	NA	0.48585922	0.75520786	0.3628483	0.38673334	0.21154754	WE2	2022-08-17T13:14:00	NA
samp_4367	PRJNA1190886	SAMN45060268	SRR31514736	37.25	NA	0.11838056	0.20250067	0.09014074	0.12751821	0.10587773	WE4	2022-08-17T9:00:00	NA
samp_4368	PRJNA1190886	SAMN45060269	SRR31514709	92.09	NA	NA	NA	NA	NA	NA	WE8	2022-08-17T8:12:00	NA
samp_4369	PRJNA1190886	SAMN45060270	SRR31514707	84.18	NA	0.39502273	0.5993388	0.29569067	0.29325962	0.21582206	WE12	2022-08-17T10:59:00	NA
samp_4370	PRJNA1190886	SAMN45060271	SRR31514706	86.72	NA	0.41354141	0.63691508	0.30392392	0.32163258	0.20663327	WE2	2022-08-29T12:35:00	NA
samp_4371	PRJNA1190886	SAMN45060272	SRR31514705	6.41	NA	0.0517289	0.09054019	0.03912936	0.04606955	0.05645184	WE4	2022-08-29T9:04:00	NA
samp_4372	PRJNA1190886	SAMN45060273	SRR31514704	79.35	NA	NA	NA	NA	NA	NA	WE8	2022-08-29T8:19:00	NA
samp_4373	PRJNA1190886	SAMN45060274	SRR31514703	47.5	NA	0.30049068	0.46547479	0.22352643	0.24077536	0.19173193	WE12	2022-08-29T9:55:00	NA
samp_4374	PRJNA1190886	SAMN45060275	SRR31514702	55.48	NA	0.30480117	0.46006347	0.22174221	0.23852501	0.17700519	WE2	2022-09-12T12:54:00	NA
samp_4375	PRJNA1190886	SAMN45060276	SRR31514701	24.51	NA	0.09752893	0.16039148	0.07420123	0.0829415	0.0750688	WE4	2022-09-12T9:02:00	NA
samp_4376	PRJNA1190886	SAMN45060277	SRR31514699	78.98	NA	NA	NA	NA	NA	NA	WE8	2022-09-12T8:11:00	NA
samp_4377	PRJNA1190886	SAMN45060278	SRR31514698	30.42	NA	0.21607286	0.3270305	0.15745755	0.16650265	0.14320468	WE12	2022-09-12T11:00:00	NA
samp_4378	PRJNA1190886	SAMN45060279	SRR31514697	12.52	NA	NA	NA	NA	NA	NA	WE2	2022-09-30T12:43:00	NA

samp_4379	PRJNA1190886	SAMN45060280	SRR31514695	35.88	NA	NA	NA	NA	NA	NA	WE8	2022-09-30T9:11:00	NA
samp_4380	PRJNA1190886	SAMN45060281	SRR31514694	8.31	NA	NA	NA	NA	NA	NA	WE12	2022-09-30T10:58:00	NA
samp_4381	PRJNA1190886	SAMN45060282	SRR31514693	54.02	NA	0.23998351	0.38614284	0.19749008	0.34943478	0.16008492	WE2	2022-10-11T14:23:00	NA
samp_4382	PRJNA1190886	SAMN45060283	SRR31514692	27.24	NA	NA	NA	NA	NA	NA	WE8	2022-10-11T9:22:00	NA
samp_4383	PRJNA1190886	SAMN45060284	SRR31514691	27.43	NA	0.17099821	0.26191026	0.13420047	0.13336341	0.10208512	WE12	2022-10-11T12:21:00	NA
samp_4384	PRJNA1190886	SAMN45060285	SRR31514690	20.74	NA	0.13830831	0.21707677	0.10982523	0.10940678	0.11351109	WE2	2022-10-24T13:36:00	NA
samp_4385	PRJNA1190886	SAMN45060286	SRR31514688	4.3	NA	0.05923287	0.14328373	0.06222033	0.24644909	0.07216346	WE4	2022-10-24T10:05:00	NA
samp_4386	PRJNA1190886	SAMN45060287	SRR31514687	12.07	NA	NA	NA	NA	NA	NA	WE8	2022-10-24T9:16:00	NA
samp_4387	PRJNA1190886	SAMN45060288	SRR31514783	26.59	NA	0.13255767	0.2056887	0.10402919	0.10980134	0.08984944	WE12	2022-10-24T12:00:00	NA
samp_4422	PRJNA1190886	SAMN45060291	SRR31514777	1.74	NA	0.03598948	0.08363668	0.02933112	0.05196671	0.06623324	WE4	2020-10-05T8:49:00	NA
samp_4425	PRJNA1190886	SAMN45060294	SRR31514790	3.67	NA	NA	NA	NA	NA	NA	WE4	2022-10-11T10:30:00	NA
samp_3427	NA	NA	NA	NA	NA	0.19859479	0.30930151	0.14465772	0.11706672	0.16054217	WE2	9/6/16	NA
samp_3428	NA	NA	NA	NA	NA	0.16152245	0.2415042	0.11490152	0.04962216	0.14039437	WE12	9/6/16	NA
samp_3429	NA	NA	NA	NA	NA	0.159625	0.23834571	0.11150558	0.03745349	0.13259034	WE15	9/6/16	NA
samp_3430	NA	NA	NA	NA	NA	0.10020087	0.15091614	0.06501205	0.02691908	0.11883577	WE4	9/6/16	NA
samp_3400	NA	NA	NA	NA	NA	0.34806655	0.59915837	0.24429819	0.31091643	0.14230859	WE12	7/11/16	NA
samp_3437	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	9/19/16	NA
samp_3411	NA	NA	NA	NA	NA	0.22351001	0.40080213	0.16113691	0.20244467	0.16762341	WE2	8/1/16	NA
samp_3434	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE15	9/12/16	NA
samp_3432	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	9/12/16	NA
samp_3407	NA	NA	NA	NA	NA	0.18138228	0.31600789	0.13048908	0.14959491	0.12054828	WE2	7/25/16	NA
samp_3433	NA	NA	NA	NA	NA	0.15005652	0.25825964	0.11607598	0.13319336	0.10914125	WE12	9/12/16	NA
samp_3435	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/12/16	NA
samp_509	NA	NA	NA	NA	NA	0.15954965	0.28324683	0.1158588	0.15365287	0.12326802	WE4	9/19/16	NA
samp_3439	NA	NA	NA	NA	NA	0.15954965	0.28324683	0.1158588	0.15365287	0.12326802	WE4	9/19/16	NA
samp_3420	NA	NA	NA	NA	NA	0.13779923	0.23732496	0.10298506	0.12225352	0.09427645	WE12	8/15/16	NA
samp_453	NA	NA	NA	NA	NA	0.12329477	0.20509475	0.09019331	0.096944	0.1105113	WE12	8/22/16	NA
samp_3424	NA	NA	NA	NA	NA	0.12329477	0.20509475	0.09019331	0.096944	0.1105113	WE12	8/22/16	NA

samp_3588	NA	NA	NA	NA	NA	0.123 29477	0.205 09475	0.090 19331	0.096 944	0.110 5113	WE12	8/22/16	NA
samp_3438	NA	NA	NA	NA	NA	0.131 97014	0.237 03304	0.103 42938	0.120 69711	0.113 31359	WE15	9/19/16	NA
samp_454	NA	NA	NA	NA	NA	0.134 98321	0.238 72711	0.103 9533	0.124 40356	0.129 25744	WE15	8/22/16	NA
samp_3425	NA	NA	NA	NA	NA	0.134 98321	0.238 72711	0.103 9533	0.124 40356	0.129 25744	WE15	8/22/16	NA
samp_3585	NA	NA	NA	NA	NA	0.134 98321	0.238 72711	0.103 9533	0.124 40356	0.129 25744	WE15	8/22/16	NA
samp_3442	NA	NA	NA	NA	NA	0.150 53632	0.260 51442	0.117 33798	0.142 43945	0.096 39059	WE12	10/3/16	NA
samp_3413	NA	NA	NA	NA	NA	0.084 97294	0.164 36658	0.067 63584	0.091 81695	0.089 94901	WE15	8/1/16	NA
samp_3408	NA	NA	NA	NA	NA	0.083 12922	0.150 97409	0.064 63342	0.083 49272	0.079 16719	WE12	7/25/16	NA
samp_3384	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE15	5/31/16	NA
samp_3402	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/11/16	NA
samp_3398	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/5/16	NA
samp_2046	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/1/16	250
samp_3414	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/1/16	NA
samp_451	NA	NA	NA	NA	NA	0.218 95659	0.419 70956	0.174 83684	0.183 26945	0.189 88966	WE12	8/1/16	NA
samp_3412	NA	NA	NA	NA	NA	0.218 95659	0.419 70956	0.174 83684	0.183 26945	0.189 88966	WE12	8/1/16	NA
samp_3440	NA	NA	NA	NA	NA	0.168 61518	0.292 10888	0.134 76308	0.164 1239	0.107 89069	WE2	10/3/16	NA
samp_3436	NA	NA	NA	NA	NA	0.162 56624	0.274 38105	0.129 82179	0.152 19185	0.119 93273	WE2	9/19/16	NA
samp_2050	NA	NA	NA	NA	NA	0.058 98854	0.117 74836	0.048 87148	0.069 90355	0.090 16306	WE4	8/15/16	80
samp_3422	NA	NA	NA	NA	NA	0.058 98854	0.117 74836	0.048 87148	0.069 90355	0.090 16306	WE4	8/15/16	NA
samp_3382	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	5/31/16	NA
samp_3421	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE15	8/15/16	NA
samp_3394	NA	NA	NA	NA	NA	0.083 68093	0.141 20833	0.057 84257	0.081 62554	0.101 07053	WE4	6/27/16	NA
samp_455	NA	NA	NA	NA	NA	0.087 32352	0.167 37658	0.069 4675	0.128 80763	0.112 31901	WE4	8/22/16	NA
samp_3426	NA	NA	NA	NA	NA	0.087 32352	0.167 37658	0.069 4675	0.128 80763	0.112 31901	WE4	8/22/16	NA
samp_3587	NA	NA	NA	NA	NA	0.087 32352	0.167 37658	0.069 4675	0.128 80763	0.112 31901	WE4	8/22/16	NA
samp_3397	NA	NA	NA	NA	NA	0.126 48681	0.228 54215	0.097 14788	0.125 13535	0.124 8641	WE15	7/5/16	NA
samp_3405	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE15	7/18/16	NA
samp_452	NA	NA	NA	NA	NA	0.173 40901	0.286 13217	0.131 84635	0.151 00728	0.150 98253	WE2	8/22/16	NA
samp_3423	NA	NA	NA	NA	NA	0.173 40901	0.286 13217	0.131 84635	0.151 00728	0.150 98253	WE2	8/22/16	NA
samp_3586	NA	NA	NA	NA	NA	0.173 40901	0.286 13217	0.131 84635	0.151 00728	0.150 98253	WE2	8/22/16	NA
samp_3391	NA	NA	NA	NA	NA	0.171 57786	0.294 08118	0.121 83903	0.139 01331	0.110 20655	WE12	6/27/16	NA
samp_3419	NA	NA	NA	NA	NA	0.156 55251	0.268 18767	0.122 07815	0.138 94522	0.137 2291	WE2	8/15/16	NA
samp_3418	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/9/16	NA
samp_3399	NA	NA	NA	NA	NA	0.201 30333	0.343 54104	0.138 13759	0.190 22224	0.124 25746	WE2	7/11/16	NA
samp_3392	NA	NA	NA	NA	NA	0.193 62468	0.335 33046	0.139 863	0.166 0779	0.126 04198	WE15	6/27/16	NA
samp_3406	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/18/16	NA
samp_3443	NA	NA	NA	NA	NA	0.124 22974	0.199 43903	0.097 69804	0.111 09073	0.110 03523	WE15	10/3/16	NA
samp_3410	NA	NA	NA	NA	NA	0.127 80001	0.210 46968	0.088 84625	0.121 95756	0.095 9252	WE4	7/25/16	NA
samp_3385	NA	NA	NA	NA	NA	0.061 26715	0.114 30109	0.044 93236	0.069 20825	0.129 33111	WE4	6/13/16	NA

samp_3389	NA	NA	NA	NA	NA	0.061 26715	0.114 30109	0.044 93236	0.069 20825	0.129 33111	WE4	6/13/16	NA
samp_3383	NA	NA	NA	NA	NA	0.276 29944	0.504 67808	0.202 30179	0.261 17513	0.135 54498	WE12	5/31/16	NA
samp_3444	NA	NA	NA	NA	NA	0.047 47189	0.068 70002	0.033 78331	0.043 18138	0.049 40848	WE4	10/3/16	NA
samp_3409	NA	NA	NA	NA	NA	0.134 19352	0.256 69306	0.104 94158	0.135 38304	0.123 77118	WE15	7/25/16	NA
samp_3388	NA	NA	NA	NA	NA	0.078 25639	0.142 78625	0.058 71374	0.084 1949	0.113 35647	WE15	6/13/16	NA
samp_3417	NA	NA	NA	NA	NA	0.123 0652	0.227 22925	0.093 67163	0.122 93694	0.114 36387	WE15	8/9/16	NA
samp_3387	NA	NA	NA	NA	NA	0.208 80728	0.355 27517	0.152 62263	0.165 36563	0.122 65459	WE12	6/13/16	NA
samp_3415	NA	NA	NA	NA	NA	0.226 28728	0.402 31166	0.165 88292	0.214 97232	0.153 13026	WE2	8/9/16	NA
samp_3452	NA	NA	NA	NA	NA	0.045 20378	0.093 37216	0.035 83631	0.077 70638	0.076 77847	WE4	10/17/16	NA
samp_3451	NA	NA	NA	NA	NA	0.115 53288	0.206 40379	0.092 38794	0.126 85017	0.100 6866	WE15	10/17/16	NA
samp_2041	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	7/18/16	125
samp_3403	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	7/18/16	NA
samp_3416	NA	NA	NA	NA	NA	0.134 8291	0.226 6153	0.103 61207	0.084 08674	0.130 08396	WE12	8/9/16	NA
samp_3450	NA	NA	NA	NA	NA	0.151 16962	0.263 65439	0.117 36342	0.157 4108	0.105 05064	WE12	10/17/16	NA
samp_3404	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	7/18/16	NA
samp_3445	NA	NA	NA	NA	NA	0.224 79311	0.380 68094	0.170 55193	0.216 01059	0.107 82203	WE2	10/11/16	NA
samp_3446	NA	NA	NA	NA	NA	0.217 70531	0.366 30904	0.164 6138	0.200 05032	0.140 52954	WE12	10/11/16	NA
samp_3386	NA	NA	NA	NA	NA	0.105 66527	0.191 04766	0.080 98043	0.088 9452	0.136 31758	WE2	6/13/16	NA
samp_3447	NA	NA	NA	NA	NA	0.118 21705	0.208 84869	0.087 19288	0.109 72716	0.120 43489	WE15	10/11/16	NA
samp_3390	NA	NA	NA	NA	NA	0.287 45631	0.472 33643	0.189 80859	0.195 16124	0.131 86097	WE2	6/27/16	NA
samp_3449	NA	NA	NA	NA	NA	0.206 78879	0.346 28875	0.155 78147	0.195 21839	0.123 27044	WE2	10/17/16	NA
samp_3396	NA	NA	NA	NA	NA	0.268 52421	0.440 2001	0.194 61521	0.214 35213	0.108 15756	WE12	7/5/16	NA
samp_3401	NA	NA	NA	NA	NA	0.121 60812	0.213 53261	0.086 18078	0.113 262	0.119 16334	WE15	7/11/16	NA
samp_3448	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	10/11/16	NA
samp_3395	NA	NA	NA	NA	NA	0.154 51662	0.266 75621	0.109 43383	0.131 90945	0.120 24301	WE2	7/5/16	NA
samp_3484	NA	NA	NA	NA	NA	0.444 77503	0.808 40778	0.337 4859	0.311 37765	0.213 19479	WE2	8/7/17	NA
samp_2076	NA	NA	NA	NA	NA	0.103 25624	0.221 86633	0.081 96223	0.158 51106	0.145 24503	WE12	8/7/17	100
samp_3485	NA	NA	NA	NA	NA	0.103 25624	0.221 86633	0.081 96223	0.158 51106	0.145 24503	WE12	8/7/17	NA
samp_3486	NA	NA	NA	NA	NA	0.252 37455	0.465 66044	0.191 10589	0.206 48865	0.126 34295	WE15	8/7/17	NA
samp_3487	NA	NA	NA	NA	NA	0.263 2535	0.466 98889	0.180 07316	0.213 76804	0.110 16626	WE4	8/7/17	NA
samp_3455	NA	NA	NA	NA	NA	0.417 56117	0.786 8368	0.335 80467	0.290 2949	0.177 22831	WE15	5/30/17	NA
samp_3489	NA	NA	NA	NA	NA	0.116 67806	0.226 63292	0.085 7682	0.131 58552	0.094 06005	WE12	8/14/17	NA
samp_3473	NA	NA	NA	NA	NA	1.008 1733	1.752 71008	0.835 99469	0.489 65211	0.193 16848	WE2	7/17/17	NA
samp_3482	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE15	7/31/17	NA
samp_3468	NA	NA	NA	NA	NA	0.114 36574	0.220 55277	0.081 13769	0.114 8257	0.098 05282	WE4	7/5/17	NA
samp_3488	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	8/14/17	NA
samp_3490	NA	NA	NA	NA	NA	0.171 09643	0.322 39505	0.126 59983	0.153 68896	0.126 61294	WE15	8/14/17	NA
samp_3491	NA	NA	NA	NA	NA	0.302 80473	0.631 99874	0.245 26065	0.443 10251	0.183 78522	WE4	8/14/17	NA

samp_3474	NA	NA	NA	NA	NA	0.731 61778	1.263 19811	0.591 97136	0.371 58757	0.195 124	WE12	7/17/17	NA
samp_3475	NA	NA	NA	NA	NA	0.093 02417	0.182 14133	0.068 83615	0.071 85885	0.137 23876	WE15	7/17/17	NA
samp_3476	NA	NA	NA	NA	NA	0.119 76685	0.215 79538	0.082 46335	0.118 88536	0.083 37279	WE4	7/17/17	NA
samp_3465	NA	NA	NA	NA	NA	0.788 78647	1.336 89478	0.614 03606	0.423 67613	0.213 79473	WE2	7/5/17	NA
samp_3467	NA	NA	NA	NA	NA	0.084 83195	0.154 14128	0.060 33496	0.068 40523	0.126 64844	WE15	7/5/17	NA
samp_3466	NA	NA	NA	NA	NA	0.216 97753	0.360 33124	0.150 73551	0.187 995	0.114 03716	WE12	7/5/17	NA
samp_3524	NA	NA	NA	NA	NA	0.078 46727	0.146 17964	0.060 33082	0.078 03574	0.137 58316	WE2	10/16/17	NA
samp_3525	NA	NA	NA	NA	NA	0.102 58724	0.181 99559	0.082 72462	0.067 51296	0.205 56319	WE12	10/16/17	NA
samp_3526	NA	NA	NA	NA	NA	0.042 04709	0.067 65403	0.027 99619	0.032 75102	0.081 93132	WE4	10/16/17	NA
samp_3520	NA	NA	NA	NA	NA	0.262 66121	0.461 22105	0.199 45585	0.237 76484	0.210 64527	WE2	10/10/17	NA
samp_3521	NA	NA	NA	NA	NA	0.167 0668	0.290 17172	0.124 99395	0.140 15085	0.176 03184	WE12	10/10/17	NA
samp_3522	NA	NA	NA	NA	NA	0.159 44866	0.288 16141	0.116 84482	0.163 14837	0.128 90577	WE15	10/10/17	NA
samp_3523	NA	NA	NA	NA	NA	0.058 7692	0.098 04259	0.041 16616	0.049 66398	0.099 47038	WE4	10/10/17	NA
samp_3512	NA	NA	NA	NA	NA	0.198 494	0.372 42236	0.156 38794	0.193 42064	0.209 61177	WE2	9/25/17	NA
samp_3513	NA	NA	NA	NA	NA	0.192 58078	0.339 62518	0.143 23658	0.169 40318	0.148 41593	WE12	9/25/17	NA
samp_3514	NA	NA	NA	NA	NA	0.176 04926	0.286 9962	0.114 93195	0.161 80899	0.147 15985	WE15	9/25/17	NA
samp_3515	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/25/17	NA
samp_3508	NA	NA	NA	NA	NA	0.209 06178	0.388 72051	0.169 78822	0.177 42014	0.157 27466	WE2	9/18/17	NA
samp_3509	NA	NA	NA	NA	NA	0.223 08103	0.387 87705	0.165 62707	0.199 12509	0.160 39223	WE12	9/18/17	NA
samp_3510	NA	NA	NA	NA	NA	0.198 47712	0.332 2099	0.135 36737	0.161 48873	0.146 925	WE15	9/18/17	NA
samp_3511	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/18/17	NA
samp_3504	NA	NA	NA	NA	NA	0.096 74328	0.166 27079	0.072 8333	0.070 70855	0.136 15437	WE2	9/11/17	NA
samp_3505	NA	NA	NA	NA	NA	0.380 175	0.675 13248	0.292 85345	0.274 69781	0.257 46247	WE12	9/11/17	NA
samp_3937	NA	NA	NA	NA	NA	0.380 175	0.675 13248	0.292 85345	0.274 69781	0.257 46247	WE12	9/11/17	NA
samp_3506	NA	NA	NA	NA	NA	0.214 39979	0.374 68989	0.161 22342	0.174 69275	0.171 68647	WE15	9/11/17	NA
samp_3507	NA	NA	NA	NA	NA	0.048 0218	0.080 21696	0.031 721	0.031 20818	0.132 54812	WE4	9/11/17	NA
samp_3500	NA	NA	NA	NA	NA	0.310 62099	0.541 10608	0.239 58333	0.221 68422	0.207 52781	WE2	9/5/17	NA
samp_3501	NA	NA	NA	NA	NA	0.240 06787	0.383 85677	0.170 67336	0.158 78175	0.183 631	WE12	9/5/17	NA
samp_3502	NA	NA	NA	NA	NA	0.186 44818	0.316 60661	0.139 56616	0.156 49007	0.166 96371	WE15	9/5/17	NA
samp_3503	NA	NA	NA	NA	NA	0.064 42869	0.096 93645	0.041 64789	0.033 96054	0.111 47674	WE4	9/5/17	NA
samp_3492	NA	NA	NA	NA	NA	0.357 08941	0.657 44296	0.270 25225	0.266 97909	0.217 51072	WE2	8/21/17	NA
samp_3493	NA	NA	NA	NA	NA	0.115 88003	0.219 5184	0.087 59668	0.103 49796	0.133 71699	WE12	8/21/17	NA
samp_3494	NA	NA	NA	NA	NA	0.191 02132	0.334 52889	0.136 67283	0.162 56034	0.136 83295	WE15	8/21/17	NA
samp_3495	NA	NA	NA	NA	NA	0.441 45892	0.777 27813	0.323 784	0.329 54006	0.177 61858	WE4	8/21/17	NA
samp_3453	NA	NA	NA	NA	NA	0.993 98591	1.774 64454	0.834 72179	0.500 50645	0.136 50945	WE2	5/30/17	NA
samp_3454	NA	NA	NA	NA	NA	0.693 56062	1.218 6018	0.582 37622	0.325 92986	0.164 965	WE12	5/30/17	NA
samp_3456	NA	NA	NA	NA	NA	0.071 56713	0.132 7038	0.053 488	0.062 48785	0.102 60155	WE4	5/30/17	NA
samp_3481	NA	NA	NA	NA	NA	0.164 70867	0.290 02577	0.115 57805	0.132 94711	0.129 66569	WE12	7/31/17	NA

samp_3483	NA	NA	NA	NA	NA	0.312 1884	0.539 45448	0.215 82926	0.230 42843	0.156 15585	WE4	7/31/17	NA
samp_3457	NA	NA	NA	NA	NA	0.598 2432	1.051 68653	0.483 10925	0.306 96677	0.142 47357	WE2	6/12/17	NA
samp_3458	NA	NA	NA	NA	NA	0.160 9896	0.287 38341	0.113 57762	0.108 19736	0.162 15999	WE12	6/12/17	NA
samp_3459	NA	NA	NA	NA	NA	0.659 83481	1.156 24509	0.526 17987	0.334 3024	0.251 6667	WE15	6/12/17	NA
samp_3460	NA	NA	NA	NA	NA	0.105 59902	0.180 62128	0.069 32895	0.072 76044	0.163 11838	WE4	6/12/17	NA
samp_3461	NA	NA	NA	NA	NA	0.150 24004	0.274 14607	0.109 2727	0.123 22292	0.133 64196	WE2	6/26/17	NA
samp_3462	NA	NA	NA	NA	NA	0.687 24135	1.176 88162	0.515 70139	0.404 02202	0.159 36311	WE12	6/26/17	NA
samp_3463	NA	NA	NA	NA	NA	0.080 52902	0.139 17748	0.054 48825	0.067 46263	0.126 84878	WE15	6/26/17	NA
samp_3464	NA	NA	NA	NA	NA	0.101 96054	0.177 57482	0.067 52228	0.090 79926	0.120 79347	WE4	6/26/17	NA
samp_3469	NA	NA	NA	NA	NA	0.554 47537	0.969 42573	0.407 19774	0.360 20969	0.186 72574	WE2	7/11/17	NA
samp_3470	NA	NA	NA	NA	NA	0.193 62998	0.336 81004	0.130 94566	0.162 64632	0.121 6456	WE12	7/11/17	NA
samp_3471	NA	NA	NA	NA	NA	0.101 60808	0.188 16196	0.069 37489	0.073 43083	0.189 61983	WE15	7/11/17	NA
samp_2066	NA	NA	NA	NA	NA	0.077 48348	0.131 6084	0.046 17646	0.050 65402	0.125 30779	WE4	7/11/17	200
samp_3472	NA	NA	NA	NA	NA	0.077 48348	0.131 6084	0.046 17646	0.050 65402	0.125 30779	WE4	7/11/17	NA
samp_2069	NA	NA	NA	NA	NA	0.589 72905	1.023 35638	0.464 74371	0.322 58617	0.208 31725	WE2	7/24/17	200
samp_3477	NA	NA	NA	NA	NA	0.589 72905	1.023 35638	0.464 74371	0.322 58617	0.208 31725	WE2	7/24/17	NA
samp_2072	NA	NA	NA	NA	NA	0.142 46276	0.265 89862	0.104 0299	0.227 20252	0.114 97783	WE12	7/24/17	200
samp_3478	NA	NA	NA	NA	NA	0.142 46276	0.265 89862	0.104 0299	0.227 20252	0.114 97783	WE12	7/24/17	NA
samp_3479	NA	NA	NA	NA	NA	0.260 91395	0.453 86735	0.177 12924	0.191 15258	0.184 24208	WE4	7/24/17	NA
samp_3480	NA	NA	NA	NA	NA	0.343 38137	0.587 93566	0.260 22837	0.233 34609	0.167 92329	WE2	7/31/17	NA
samp_2090	NA	NA	NA	NA	NA	0.132 69778	0.266 75424	0.094 66192	0.160 27049	0.085 04678	WE4	8/13/18	100
samp_3557	NA	NA	NA	NA	NA	0.132 69778	0.266 75424	0.094 66192	0.160 27049	0.085 04678	WE4	8/13/18	NA
samp_458	NA	NA	NA	NA	NA	0.211 88314	0.407 99649	0.152 22571	0.224 76646	0.115 55636	WE2	8/20/18	NA
samp_3558	NA	NA	NA	NA	NA	0.211 88314	0.407 99649	0.152 22571	0.224 76646	0.115 55636	WE2	8/20/18	NA
samp_3561	NA	NA	NA	NA	NA	0.158 64962	0.316 80368	0.123 76355	0.177 96834	0.108 54109	WE2	8/27/18	NA
samp_457	NA	NA	NA	NA	NA	0.225 09486	0.441 70442	0.163 12827	0.223 21405	0.112 76842	WE2	8/13/18	NA
samp_3555	NA	NA	NA	NA	NA	0.225 09486	0.441 70442	0.163 12827	0.223 21405	0.112 76842	WE2	8/13/18	NA
samp_3563	NA	NA	NA	NA	NA	0.146 21641	0.293 72716	0.108 19253	0.172 02859	0.107 5725	WE4	8/27/18	NA
samp_3553	NA	NA	NA	NA	NA	0.232 8003	0.476 94098	0.176 46305	0.355 29828	0.109 38072	WE12	8/6/18	NA
samp_3556	NA	NA	NA	NA	NA	0.267 43867	0.527 37452	0.196 41342	0.350 13392	0.127 55635	WE12	8/13/18	NA
samp_3527	NA	NA	NA	NA	NA	0.877 62521	1.679 45631	0.767 14571	0.502 98003	0.140 74679	WE2	5/21/18	NA
samp_3528	NA	NA	NA	NA	NA	0.353 15108	0.690 621	0.302 80077	0.337 60477	0.096 85271	WE12	5/21/18	NA
samp_3529	NA	NA	NA	NA	NA	0.420 00072	0.750 91759	0.341 07765	0.217 81277	0.293 28934	WE4	5/21/18	NA
samp_3530	NA	NA	NA	NA	NA	0.735 96583	1.346 42801	0.567 54383	0.480 74216	0.119 36027	WE2	6/12/18	NA
samp_4420	NA	NA	NA	NA	NA	0.735 96583	1.346 42801	0.567 54383	0.480 74216	0.119 36027	WE2	6/12/18	NA
samp_3531	NA	NA	NA	NA	NA	0.634 85334	1.136 78261	0.482 01751	0.417 43002	0.142 73763	WE12	6/12/18	NA
samp_3532	NA	NA	NA	NA	NA	0.097 66702	0.194 24415	0.075 09785	0.110 24801	0.074 75973	WE4	6/12/18	NA
samp_3533	NA	NA	NA	NA	NA	0.512 74784	0.930 37689	0.386 24307	0.376 01589	0.123 41789	WE2	6/25/18	NA

samp_3534	NA	NA	NA	NA	NA	0.936 83807	1.708 35743	0.714 86306	0.598 86409	0.195 60198	WE12	6/25/18	NA
samp_3535	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	6/25/18	NA
samp_3536	NA	NA	NA	NA	NA	0.280 36779	0.511 32983	0.216 47653	0.243 59206	0.094 33649	WE2	7/2/18	NA
samp_3538	NA	NA	NA	NA	NA	0.342 91073	0.642 6369	0.252 39447	0.287 47044	0.105 39331	WE12	7/2/18	NA
samp_3539	NA	NA	NA	NA	NA	0.193 61092	0.352 21595	0.141 36448	0.180 31378	0.070 4387	WE4	7/2/18	NA
samp_3540	NA	NA	NA	NA	NA	0.548 47837	0.991 66696	0.407 41293	0.435 79613	0.157 20578	WE2	7/9/18	NA
samp_3541	NA	NA	NA	NA	NA	0.279 17609	0.515 32528	0.206 74013	0.252 48393	0.085 94352	WE12	7/9/18	NA
samp_3542	NA	NA	NA	NA	NA	0.084 32799	0.163 60707	0.061 89783	0.108 5281	0.082 01628	WE4	7/9/18	NA
samp_2081	NA	NA	NA	NA	NA	0.235 80862	0.425 47356	0.168 96512	0.214 89752	0.084 9365	WE2	7/16/18	100
samp_3543	NA	NA	NA	NA	NA	0.235 80862	0.425 47356	0.168 96512	0.214 89752	0.084 9365	WE2	7/16/18	NA
samp_2084	NA	NA	NA	NA	NA	0.235 29874	0.459 50234	0.170 21667	0.310 09271	0.096 34426	WE12	7/16/18	125
samp_3544	NA	NA	NA	NA	NA	0.235 29874	0.459 50234	0.170 21667	0.310 09271	0.096 34426	WE12	7/16/18	NA
samp_3545	NA	NA	NA	NA	NA	0.233 58073	0.443 21124	0.172 78718	0.243 55021	0.097 86177	WE4	7/16/18	NA
samp_3546	NA	NA	NA	NA	NA	0.242 04848	0.446 21326	0.175 05871	0.217 50595	0.097 18762	WE2	7/23/18	NA
samp_3547	NA	NA	NA	NA	NA	0.218 47325	0.404 73449	0.157 35379	0.201 64436	0.089 54338	WE12	7/23/18	NA
samp_3548	NA	NA	NA	NA	NA	0.286 9245	0.516 81971	0.200 6874	0.239 53534	0.117 82121	WE4	7/23/18	NA
samp_3549	NA	NA	NA	NA	NA	0.252 10919	0.464 3288	0.181 98427	0.228 44178	0.104 82629	WE2	7/30/18	NA
samp_3550	NA	NA	NA	NA	NA	0.257 01346	0.494 31749	0.185 1145	0.255 46297	0.099 20542	WE12	7/30/18	NA
samp_2086	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/30/18	70
samp_3551	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/30/18	NA
samp_3552	NA	NA	NA	NA	NA	0.313 90545	0.587 07828	0.228 46087	0.341 52035	0.130 96684	WE2	8/6/18	NA
samp_3554	NA	NA	NA	NA	NA	0.104 56124	0.210 0995	0.076 98193	0.118 81686	0.080 82057	WE4	8/6/18	NA
samp_3559	NA	NA	NA	NA	NA	0.124 09002	0.247 67309	0.095 05436	0.170 82765	0.090 4503	WE12	8/20/18	NA
samp_3936	NA	NA	NA	NA	NA	0.124 09002	0.247 67309	0.095 05436	0.170 82765	0.090 4503	WE12	8/20/18	NA
samp_510	NA	NA	NA	NA	NA	0.147 0105	0.271 167	0.103 01277	0.145 88934	0.094 74307	WE4	8/20/18	NA
samp_3560	NA	NA	NA	NA	NA	0.147 0105	0.271 167	0.103 01277	0.145 88934	0.094 74307	WE4	8/20/18	NA
samp_3562	NA	NA	NA	NA	NA	0.107 47607	0.211 69631	0.084 92373	0.104 95642	0.091 20287	WE12	8/27/18	NA
samp_3564	NA	NA	NA	NA	NA	0.123 4591	0.250 62514	0.100 29544	0.138 16018	0.108 56518	WE2	9/4/18	NA
samp_3565	NA	NA	NA	NA	NA	0.105 49595	0.204 61754	0.082 97643	0.111 38778	0.095 00066	WE12	9/4/18	NA
samp_3566	NA	NA	NA	NA	NA	0.185 90458	0.392 29493	0.148 85494	0.320 87637	0.128 54817	WE4	9/4/18	NA
samp_3567	NA	NA	NA	NA	NA	0.308 86586	0.547 04255	0.227 99342	0.290 5749	0.162 02631	WE2	9/11/18	NA
samp_2100	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	9/11/18	20
samp_3568	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	9/11/18	NA
samp_3569	NA	NA	NA	NA	NA	0.208 17315	0.365 46181	0.156 4099	0.177 68348	0.124 82376	WE4	9/11/18	NA
samp_3570	NA	NA	NA	NA	NA	0.247 20533	0.441 35035	0.182 00495	0.247 39702	0.128 43252	WE2	9/17/18	NA
samp_3571	NA	NA	NA	NA	NA	0.238 87899	0.418 89999	0.178 15014	0.207 35683	0.134 39733	WE12	9/17/18	NA
samp_3572	NA	NA	NA	NA	NA	0.107 06906	0.197 23548	0.077 21691	0.123 89635	0.101 29429	WE4	9/17/18	NA
samp_3573	NA	NA	NA	NA	NA	0.190 12862	0.342 356	0.139 50545	0.186 76976	0.132 31956	WE2	9/25/18	NA

samp_2104	NA	NA	NA	NA	NA	0.284 0114	0.490 72925	0.209 48391	0.250 77527	0.140 4303	WE12	9/25/18	50
samp_3574	NA	NA	NA	NA	NA	0.284 0114	0.490 72925	0.209 48391	0.250 77527	0.140 4303	WE12	9/25/18	NA
samp_2102	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/25/18	150
samp_3575	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/25/18	NA
samp_3576	NA	NA	NA	NA	NA	0.194 55234	0.353 97104	0.146 73177	0.191 93913	0.125 14531	WE2	10/1/18	NA
samp_3577	NA	NA	NA	NA	NA	0.130 82526	0.235 53535	0.096 21163	0.124 22376	0.097 61895	WE12	10/1/18	NA
samp_3578	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	10/1/18	NA
samp_3581	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	5/21/19	NA
samp_3580	NA	NA	NA	NA	NA	0.942 26674	1.884 18807	0.897 27369	0.339 82118	0.214 51408	WE12	5/21/19	NA
samp_3579	NA	NA	NA	NA	NA	1.022 98099	1.925 02411	0.914 3174	0.372 76202	0.222 83867	WE2	5/21/19	NA
samp_3018	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	6/3/19	NA
samp_3021	NA	NA	NA	NA	NA	0.080 92593	0.169 46262	0.066 56455	0.103 28104	0.083 35374	WE4	6/17/19	NA
samp_461	NA	NA	NA	NA	NA	0.522 33181	0.840 25521	0.328 7056	0.323 70652	0.166 83198	WE4	7/1/19	NA
samp_3024	NA	NA	NA	NA	NA	0.522 33181	0.840 25521	0.328 7056	0.323 70652	0.166 83198	WE4	7/1/19	NA
samp_2106	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/8/19	250
samp_3027	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/8/19	NA
samp_3016	NA	NA	NA	NA	NA	0.386 65183	0.753 75018	0.320 51551	0.260 72229	0.121 90463	WE2	6/3/19	NA
samp_3019	NA	NA	NA	NA	NA	1.175 54538	2.105 32238	1.007 14276	0.502 3516	0.273 43219	WE2	6/17/19	NA
samp_4421	NA	NA	NA	NA	NA	1.175 54538	2.105 32238	1.007 14276	0.502 3516	0.273 43219	WE2	6/17/19	NA
samp_459	NA	NA	NA	NA	NA	0.675 16298	1.326 50342	0.568 47754	0.590 11136	0.234 82173	WE2	7/1/19	NA
samp_3022	NA	NA	NA	NA	NA	0.675 16298	1.326 50342	0.568 47754	0.590 11136	0.234 82173	WE2	7/1/19	NA
samp_3025	NA	NA	NA	NA	NA	0.595 32945	1.068 92543	0.443 66878	0.421 26003	0.186 76232	WE2	7/8/19	NA
samp_3017	NA	NA	NA	NA	NA	0.860 40833	1.588 17814	0.745 02567	0.410 1969	0.180 60394	WE12	6/3/19	NA
samp_3020	NA	NA	NA	NA	NA	0.918 01186	1.702 86361	0.757 21127	0.470 96665	0.213 6833	WE12	6/17/19	NA
samp_460	NA	NA	NA	NA	NA	0.517 44782	0.995 36696	0.413 52585	0.422 34556	0.145 99161	WE12	7/1/19	NA
samp_3023	NA	NA	NA	NA	NA	0.517 44782	0.995 36696	0.413 52585	0.422 34556	0.145 99161	WE12	7/1/19	NA
samp_2108	NA	NA	NA	NA	NA	0.801 43203	1.457 45327	0.619 68438	0.546 20719	0.246 8735	WE12	7/8/19	150
samp_3026	NA	NA	NA	NA	NA	0.801 43203	1.457 45327	0.619 68438	0.546 20719	0.246 8735	WE12	7/8/19	NA
samp_3030	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	7/15/19	NA
samp_3033	NA	NA	NA	NA	NA	0.227 55268	0.432 23895	0.166 64064	0.198 74041	0.106 87989	WE4	7/22/19	NA
samp_3028	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	7/15/19	NA
samp_3031	NA	NA	NA	NA	NA	0.589 25174	1.051 02803	0.428 86006	0.436 11264	0.148 64482	WE2	7/22/19	NA
samp_511	NA	NA	NA	NA	NA	0.544 31267	0.963 326	0.389 23891	0.378 66182	0.133 53834	WE12	7/15/19	NA
samp_3029	NA	NA	NA	NA	NA	0.544 31267	0.963 326	0.389 23891	0.378 66182	0.133 53834	WE12	7/15/19	NA
samp_3032	NA	NA	NA	NA	NA	0.252 48731	0.463 47601	0.180 88552	0.206 75937	0.087 53306	WE12	7/22/19	NA
samp_512	NA	NA	NA	NA	NA	0.193 06411	0.387 85067	0.149 23181	0.183 88755	0.127 15246	WE4	7/29/19	NA
samp_3036	NA	NA	NA	NA	NA	0.193 06411	0.387 85067	0.149 23181	0.183 88755	0.127 15246	WE4	7/29/19	NA
samp_3039	NA	NA	NA	NA	NA	0.071 71188	0.157 53409	0.056 52604	0.100 88595	0.086 61378	WE4	8/5/19	NA
samp_3285	NA	NA	NA	NA	NA	0.071 71188	0.157 53409	0.056 52604	0.100 88595	0.086 61378	WE4	8/5/19	NA

samp_466	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/12/19	NA
samp_3042	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/12/19	NA
samp_3288	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/12/19	NA
samp_3034	NA	NA	NA	NA	NA	0.881 24269	1.603 8317	0.656 25548	0.625 60035	0.252 89479	WE2	7/29/19	NA
samp_462	NA	NA	NA	NA	NA	0.420 88424	0.786 13968	0.313 78275	0.401 55168	0.215 73716	WE2	8/5/19	NA
samp_3037	NA	NA	NA	NA	NA	0.420 88424	0.786 13968	0.313 78275	0.401 55168	0.215 73716	WE2	8/5/19	NA
samp_3284	NA	NA	NA	NA	NA	0.420 88424	0.786 13968	0.313 78275	0.401 55168	0.215 73716	WE2	8/5/19	NA
samp_464	NA	NA	NA	NA	NA	0.292 44738	0.564 99041	0.219 73568	0.305 64246	0.260 14519	WE2	8/12/19	NA
samp_3040	NA	NA	NA	NA	NA	0.292 44738	0.564 99041	0.219 73568	0.305 64246	0.260 14519	WE2	8/12/19	NA
samp_3045	NA	NA	NA	NA	NA	0.076 8894	0.161 60214	0.056 29276	0.071 35446	0.100 60078	WE4	8/19/19	NA
samp_3043	NA	NA	NA	NA	NA	0.271 54376	0.496 76523	0.202 35239	0.233 39903	0.166 20707	WE2	8/19/19	NA
samp_508	NA	NA	NA	NA	NA	0.362 57989	0.681 38225	0.263 99731	0.280 81493	0.152 30117	WE12	7/29/19	NA
samp_3035	NA	NA	NA	NA	NA	0.362 57989	0.681 38225	0.263 99731	0.280 81493	0.152 30117	WE12	7/29/19	NA
samp_463	NA	NA	NA	NA	NA	0.394 3494	0.713 48145	0.285 98304	0.316 05431	0.160 48107	WE12	8/5/19	NA
samp_3038	NA	NA	NA	NA	NA	0.394 3494	0.713 48145	0.285 98304	0.316 05431	0.160 48107	WE12	8/5/19	NA
samp_3286	NA	NA	NA	NA	NA	0.394 3494	0.713 48145	0.285 98304	0.316 05431	0.160 48107	WE12	8/5/19	NA
samp_465	NA	NA	NA	NA	NA	0.434 44798	0.799 31803	0.313 04453	0.333 52359	0.228 43407	WE12	8/12/19	NA
samp_3041	NA	NA	NA	NA	NA	0.434 44798	0.799 31803	0.313 04453	0.333 52359	0.228 43407	WE12	8/12/19	NA
samp_3287	NA	NA	NA	NA	NA	0.434 44798	0.799 31803	0.313 04453	0.333 52359	0.228 43407	WE12	8/12/19	NA
samp_3044	NA	NA	NA	NA	NA	0.201 79872	0.374 42287	0.152 74053	0.166 1887	0.132 00241	WE12	8/19/19	NA
samp_3048	NA	NA	NA	NA	NA	0.198 49749	0.371 84604	0.146 08393	0.189 8103	0.177 05239	WE4	8/28/19	NA
samp_469	NA	NA	NA	NA	NA	0.247 58528	0.456 95634	0.184 52425	0.218 17774	0.198 86744	WE4	9/3/19	NA
samp_3051	NA	NA	NA	NA	NA	0.247 58528	0.456 95634	0.184 52425	0.218 17774	0.198 86744	WE4	9/3/19	NA
samp_3046	NA	NA	NA	NA	NA	0.334 11807	0.621 55533	0.256 12887	0.285 76267	0.192 05044	WE2	8/28/19	NA
samp_467	NA	NA	NA	NA	NA	0.439 50072	0.797 84875	0.338 69023	0.342 56688	0.252 19745	WE2	9/3/19	NA
samp_3049	NA	NA	NA	NA	NA	0.439 50072	0.797 84875	0.338 69023	0.342 56688	0.252 19745	WE2	9/3/19	NA
samp_3047	NA	NA	NA	NA	NA	0.335 92558	0.617 87679	0.250 14657	0.280 10777	0.193 88346	WE12	8/28/19	NA
samp_468	NA	NA	NA	NA	NA	0.312 78498	0.571 42458	0.236 20275	0.266 55614	0.203 26785	WE12	9/3/19	NA
samp_3050	NA	NA	NA	NA	NA	0.312 78498	0.571 42458	0.236 20275	0.266 55614	0.203 26785	WE12	9/3/19	NA
samp_3054	NA	NA	NA	NA	NA	0.069 37051	0.124 37782	0.045 5585	0.029 80735	0.069 15567	WE4	9/9/19	NA
samp_3057	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/16/19	NA
samp_3052	NA	NA	NA	NA	NA	0.519 15994	0.942 8356	0.371 06017	0.437 73328	0.308 55858	WE2	9/9/19	NA
samp_3055	NA	NA	NA	NA	NA	0.609 31202	1.059 78073	0.434 43157	0.467 30234	0.278 69792	WE2	9/16/19	NA
samp_3053	NA	NA	NA	NA	NA	0.496 04309	0.924 46555	0.383 63483	0.412 23334	0.177 12306	WE12	9/9/19	NA
samp_3056	NA	NA	NA	NA	NA	0.481 02592	0.855 12504	0.339 84993	0.405 73684	0.325 8476	WE12	9/16/19	NA
samp_3060	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/24/19	NA
samp_4335	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/24/19	NA
samp_3059	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	9/24/19	NA

samp_3058	NA	NA	NA	NA	NA	0.327 37561	0.575 65094	0.229 38364	0.264 41137	0.251 85109	WE2	9/24/19	NA
samp_3063	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	10/7/19	NA
samp_3061	NA	NA	NA	NA	NA	0.348 86687	0.628 39962	0.250 19451	0.275 23228	0.211 04697	WE2	10/7/19	NA
samp_3062	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	10/7/19	NA
samp_470	NA	NA	NA	NA	NA	0.336 40664	0.604 92658	0.235 08302	0.273 33227	0.114 90839	WE2	7/6/20	NA
samp_3064	NA	NA	NA	NA	NA	0.336 40664	0.604 92658	0.235 08302	0.273 33227	0.114 90839	WE2	7/6/20	NA
samp_3069	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE2	7/13/20	NA
samp_3065	NA	NA	NA	NA	NA	0.422 36785	0.768 04297	0.294 28084	0.360 0945	0.151 72637	WE6	7/6/20	NA
samp_471	NA	NA	NA	NA	NA	0.339 53204	0.642 48968	0.243 6509	0.299 05456	0.134 85634	WE8	7/6/20	NA
samp_3066	NA	NA	NA	NA	NA	0.339 53204	0.642 48968	0.243 6509	0.299 05456	0.134 85634	WE8	7/6/20	NA
samp_3289	NA	NA	NA	NA	NA	0.339 53204	0.642 48968	0.243 6509	0.299 05456	0.134 85634	WE8	7/6/20	NA
samp_3067	NA	NA	NA	NA	NA	0.801 55294	1.469 87713	0.572 55839	0.666 63114	0.205 89853	WE9	7/6/20	NA
samp_3070	NA	NA	NA	NA	NA	0.775 40308	1.354 91753	0.545 72282	0.627 62361	0.205 70177	WE9	7/13/20	NA
samp_472	NA	NA	NA	NA	NA	0.194 95014	0.373 61146	0.143 16565	0.183 89282	0.092 52777	WE12	7/6/20	NA
samp_3068	NA	NA	NA	NA	NA	0.194 95014	0.373 61146	0.143 16565	0.183 89282	0.092 52777	WE12	7/6/20	NA
samp_3071	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	7/13/20	NA
samp_474	NA	NA	NA	NA	NA	0.127 82199	0.248 55264	0.090 88362	0.151 20748	0.097 52643	WE4	7/28/20	NA
samp_3076	NA	NA	NA	NA	NA	0.127 82199	0.248 55264	0.090 88362	0.151 20748	0.097 52643	WE4	7/28/20	NA
samp_3072	NA	NA	NA	NA	NA	0.220 15785	0.413 61255	0.154 47904	0.214 48172	0.103 99599	WE2	7/20/20	NA
samp_473	NA	NA	NA	NA	NA	0.336 07884	0.638 0365	0.242 79384	0.312 83637	0.183 40581	WE2	7/28/20	NA
samp_3075	NA	NA	NA	NA	NA	0.336 07884	0.638 0365	0.242 79384	0.312 83637	0.183 40581	WE2	7/28/20	NA
samp_3074	NA	NA	NA	NA	NA	0.168 07014	0.310 8281	0.116 30406	0.178 36513	0.088 77625	WE12	7/20/20	NA
samp_476	NA	NA	NA	NA	NA	0.160 63796	0.299 7464	0.110 0415	0.169 15078	0.083 08382	WE12	7/28/20	NA
samp_2146	NA	NA	NA	NA	NA	0.160 63796	0.299 7464	0.110 0415	0.169 15078	0.083 08382	WE12	7/28/20	150
samp_3078	NA	NA	NA	NA	NA	0.160 63796	0.299 7464	0.110 0415	0.169 15078	0.083 08382	WE12	7/28/20	NA
samp_3073	NA	NA	NA	NA	NA	0.662 4728	1.206 78687	0.466 59411	0.565 5155	0.193 46005	WE9	7/20/20	NA
samp_475	NA	NA	NA	NA	NA	0.522 78532	0.950 09746	0.375 39845	0.460 1587	0.195 14696	WE9	7/28/20	NA
samp_3077	NA	NA	NA	NA	NA	0.522 78532	0.950 09746	0.375 39845	0.460 1587	0.195 14696	WE9	7/28/20	NA
samp_3082	NA	NA	NA	NA	NA	0.252 45713	0.487 36505	0.195 76388	0.242 27568	0.130 71682	WE2	8/4/20	NA
samp_477	NA	NA	NA	NA	NA	0.263 44689	0.499 09133	0.186 66743	0.254 79305	0.158 02989	WE2	8/10/20	NA
samp_3086	NA	NA	NA	NA	NA	0.263 44689	0.499 09133	0.186 66743	0.254 79305	0.158 02989	WE2	8/10/20	NA
samp_478	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/10/20	NA
samp_2148	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/10/20	200
samp_3087	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/10/20	NA
samp_3085	NA	NA	NA	NA	NA	0.341 91575	0.645 36006	0.252 17351	0.329 84859	0.160 4019	WE12	8/4/20	NA
samp_480	NA	NA	NA	NA	NA	0.254 11204	0.474 28203	0.184 15987	0.253 94673	0.157 68387	WE12	8/10/20	NA
samp_3089	NA	NA	NA	NA	NA	0.254 11204	0.474 28203	0.184 15987	0.253 94673	0.157 68387	WE12	8/10/20	NA
samp_3084	NA	NA	NA	NA	NA	0.520 10207	0.948 38057	0.384 51698	0.471 23976	0.208 84961	WE9	8/4/20	NA

samp_479	NA	NA	NA	NA	NA	0.411 95151	0.764 71934	0.298 04392	0.402 73765	0.201 92888	WE9	8/10/20	NA
samp_3088	NA	NA	NA	NA	NA	0.411 95151	0.764 71934	0.298 04392	0.402 73765	0.201 92888	WE9	8/10/20	NA
samp_3093	NA	NA	NA	NA	NA	0.218 93605	0.411 25771	0.160 38174	0.191 46725	0.140 47781	WE2	8/17/20	NA
samp_3096	NA	NA	NA	NA	NA	0.212 03917	0.393 63061	0.155 90105	0.181 85788	0.123 29767	WE12	8/17/20	NA
samp_3099	NA	NA	NA	NA	NA	0.153 99634	0.288 67295	0.114 74058	0.141 37771	0.113 72068	WE12	8/24/20	NA
samp_3095	NA	NA	NA	NA	NA	0.472 85767	0.888 46217	0.344 2171	0.448 70857	0.206 87629	WE9	8/17/20	NA
samp_3098	NA	NA	NA	NA	NA	0.487 71754	0.923 94817	0.357 03441	0.485 57305	0.217 60999	WE9	8/24/20	NA
samp_3100	NA	NA	NA	NA	NA	0.138 471	0.274 82508	0.105 70234	0.120 91016	0.120 9265	WE2	9/1/20	NA
samp_481	NA	NA	NA	NA	NA	0.108 93656	0.211 77654	0.085 31873	0.088 25699	0.101 8985	WE2	9/9/20	NA
samp_3104	NA	NA	NA	NA	NA	0.108 93656	0.211 77654	0.085 31873	0.088 25699	0.101 8985	WE2	9/9/20	NA
samp_3103	NA	NA	NA	NA	NA	0.167 72248	0.329 46469	0.124 78838	0.177 31158	0.138 13026	WE12	9/1/20	NA
samp_484	NA	NA	NA	NA	NA	0.081 6416	0.171 62154	0.067 50442	0.102 51356	0.083 61454	WE12	9/9/20	NA
samp_3107	NA	NA	NA	NA	NA	0.081 6416	0.171 62154	0.067 50442	0.102 51356	0.083 61454	WE12	9/9/20	NA
samp_3102	NA	NA	NA	NA	NA	0.477 26625	0.891 53896	0.351 7528	0.468 5121	0.234 88332	WE9	9/1/20	NA
samp_483	NA	NA	NA	NA	NA	0.412 11328	0.793 95766	0.309 05495	0.437 52891	0.220 35162	WE9	9/9/20	NA
samp_3106	NA	NA	NA	NA	NA	0.412 11328	0.793 95766	0.309 05495	0.437 52891	0.220 35162	WE9	9/9/20	NA
samp_3109	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/15/20	NA
samp_3113	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	9/21/20	NA
samp_3108	NA	NA	NA	NA	NA	0.165 35125	0.329 1731	0.131 59433	0.170 48553	0.158 63559	WE2	9/15/20	NA
samp_3112	NA	NA	NA	NA	NA	0.242 39787	0.468 04866	0.185 37256	0.274 31405	0.171 52223	WE2	9/21/20	NA
samp_3111	NA	NA	NA	NA	NA	0.218 49984	0.409 75278	0.161 75789	0.203 95643	0.150 21328	WE12	9/15/20	NA
samp_3115	NA	NA	NA	NA	NA	0.220 15121	0.412 5502	0.166 56597	0.223 88569	0.150 58274	WE12	9/21/20	NA
samp_3110	NA	NA	NA	NA	NA	0.588 90436	1.095 19613	0.416 49406	0.621 48143	0.235 71216	WE9	9/15/20	NA
samp_3114	NA	NA	NA	NA	NA	0.344 42078	0.655 09065	0.257 89801	0.361 29644	0.199 87376	WE9	9/21/20	NA
samp_485	NA	NA	NA	NA	NA	0.160 0693	0.310 55637	0.123 4436	0.159 11173	0.146 319	WE2	9/28/20	NA
samp_3119	NA	NA	NA	NA	NA	0.160 0693	0.310 55637	0.123 4436	0.159 11173	0.146 319	WE2	9/28/20	NA
samp_3291	NA	NA	NA	NA	NA	0.160 0693	0.310 55637	0.123 4436	0.159 11173	0.146 319	WE2	9/28/20	NA
samp_3123	NA	NA	NA	NA	NA	0.142 9915	0.277 74544	0.107 58753	0.145 32368	0.123 71367	WE2	10/5/20	NA
samp_487	NA	NA	NA	NA	NA	0.110 80028	0.224 48984	0.088 66638	0.122 92179	0.104 22513	WE12	9/28/20	NA
samp_3122	NA	NA	NA	NA	NA	0.110 80028	0.224 48984	0.088 66638	0.122 92179	0.104 22513	WE12	9/28/20	NA
samp_3121	NA	NA	NA	NA	NA	0.247 81315	0.496 10363	0.189 44884	0.261 0203	0.179 76329	WE9	9/28/20	NA
samp_3083	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/4/20	NA
samp_3094	NA	NA	NA	NA	NA	0.052 56728	0.106 43973	0.039 99629	0.036 34773	0.057 77923	WE4	8/17/20	NA
samp_3097	NA	NA	NA	NA	NA	0.056 45926	0.120 11817	0.045 27454	0.044 75656	0.065 82042	WE4	8/24/20	NA
samp_3101	NA	NA	NA	NA	NA	0.040 27485	0.076 68245	0.029 95493	0.013 15781	0.069 12681	WE4	9/1/20	NA
samp_482	NA	NA	NA	NA	NA	0.043 11027	0.087 75543	0.035 85794	0.041 69641	0.055 00984	WE4	9/9/20	NA
samp_3105	NA	NA	NA	NA	NA	0.043 11027	0.087 75543	0.035 85794	0.041 69641	0.055 00984	WE4	9/9/20	NA
samp_3290	NA	NA	NA	NA	NA	0.043 11027	0.087 75543	0.035 85794	0.041 69641	0.055 00984	WE4	9/9/20	NA

samp_486	NA	NA	NA	NA	NA	0.045 6643	0.101 22323	0.034 95676	0.064 25161	0.064 78561	WE4	9/28/20	NA
samp_3120	NA	NA	NA	NA	NA	0.045 6643	0.101 22323	0.034 95676	0.064 25161	0.064 78561	WE4	9/28/20	NA
samp_3124	NA	NA	NA	NA	NA	0.035 98948	0.083 63668	0.029 33112	0.051 96671	0.066 23324	WE4	10/5/20	NA
samp_3126	NA	NA	NA	NA	NA	0.129 34764	0.261 38372	0.100 65354	0.154 84774	0.126 02422	WE12	10/5/20	NA
samp_3125	NA	NA	NA	NA	NA	0.233 74057	0.463 96975	0.176 75947	0.249 14345	0.173 69291	WE9	10/5/20	NA
samp_3127	NA	NA	NA	NA	NA	0.120 13138	0.266 73186	0.102 62843	0.148 98292	0.070 28799	WE2	4/27/21	NA
samp_3130	NA	NA	NA	NA	NA	0.156 61894	0.329 57078	0.128 41262	0.168 70785	0.092 67993	WE12	4/27/21	NA
samp_3129	NA	NA	NA	NA	NA	0.820 77158	1.520 2737	0.618 48574	0.635 69768	0.211 85198	WE9	4/27/21	NA
samp_3128	NA	NA	NA	NA	NA	0.075 94794	0.155 48078	0.057 13806	0.097 97812	0.078 80697	WE4	4/27/21	NA
samp_3131	NA	NA	NA	NA	NA	0.402 40539	0.780 89826	0.335 04035	0.288 94406	0.109 58289	WE2	5/24/21	NA
samp_3134	NA	NA	NA	NA	NA	0.363 35566	0.713 43081	0.297 70269	0.282 36139	0.098 69606	WE12	5/24/21	NA
samp_3132	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	5/24/21	NA
samp_3133	NA	NA	NA	NA	NA	1.178 517	2.248 74229	1.004 16699	0.626 89636	0.258 97718	WE9	5/24/21	NA
samp_3135	NA	NA	NA	NA	NA	0.403 11247	0.752 25647	0.311 46719	0.299 88922	0.105 82275	WE2	6/7/21	NA
samp_3136	NA	NA	NA	NA	NA	0.100 90973	0.200 85275	0.079 64842	0.104 39276	0.082 06211	WE4	6/7/21	NA
samp_3137	NA	NA	NA	NA	NA	0.840 23145	1.527 05421	0.637 31436	0.565 16906	0.163 84334	WE9	6/7/21	NA
samp_3138	NA	NA	NA	NA	NA	0.193 70031	0.366 86897	0.155 09331	0.148 15862	0.078 37364	WE12	6/7/21	NA
samp_3210	NA	NA	NA	NA	NA	0.848 08319	1.530 38194	0.634 36837	0.530 073	0.221 29992	WE9	6/8/21	NA
samp_3211	NA	NA	NA	NA	NA	0.190 43502	0.355 64551	0.145 19213	0.150 28502	0.103 39485	WE12	6/8/21	NA
samp_3209	NA	NA	NA	NA	NA	0.087 46634	0.168 88874	0.063 78954	0.091 36162	0.095 06837	WE4	6/8/21	NA
samp_3208	NA	NA	NA	NA	NA	0.266 79396	0.501 92992	0.202 79934	0.196 61504	0.106 82109	WE2	6/8/21	NA
samp_3141	NA	NA	NA	NA	NA	0.705 96421	1.328 77867	0.505 21231	0.583 66769	0.160 77321	WE9	6/22/21	NA
samp_3143	NA	NA	NA	NA	NA	0.191 97058	0.391 42476	0.149 24201	0.246 84669	0.099 21913	WE2	7/6/21	NA
samp_3144	NA	NA	NA	NA	NA	0.179 97843	0.344 93461	0.131 91226	0.188 77765	0.077 97348	WE4	7/6/21	NA
samp_3145	NA	NA	NA	NA	NA	0.912 28321	1.637 33574	0.682 17786	0.600 62458	0.185 48872	WE9	7/6/21	NA
samp_3146	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE12	7/6/21	NA
samp_3212	NA	NA	NA	NA	NA	0.345 24041	0.526 79205	0.242 3079	0.151 59343	0.094 94845	WE2	7/14/21	NA
samp_3213	NA	NA	NA	NA	NA	0.169 02014	0.251 00385	0.112 60537	0.043 85764	0.095 91646	WE4	7/14/21	NA
samp_3215	NA	NA	NA	NA	NA	0.094 89314	0.138 80318	0.069 93665	0	0.088 87822	WE12	7/14/21	NA
samp_3150	NA	NA	NA	NA	NA	0.096 53642	0.140 43511	0.073 42316	0	0.041 40886	WE12	7/13/21	NA
samp_3154	NA	NA	NA	NA	NA	0.192 92917	0.299 93732	0.143 56166	0.139 07116	0.095 51779	WE12	7/19/21	NA
samp_3152	NA	NA	NA	NA	NA	0.057 12904	0.094 23003	0.040 38448	0.056 13587	0.052 28886	WE4	7/19/21	NA
samp_3142	NA	NA	NA	NA	NA	0.239 73145	0.429 67189	0.175 10042	0.135 80784	0.069 5178	WE12	6/22/21	NA
samp_3139	NA	NA	NA	NA	NA	0.145 3934	0.278 87201	0.110 75213	0.118 3546	0.059 74053	WE2	6/22/21	NA
samp_4423	NA	NA	NA	NA	NA	0.145 3934	0.278 87201	0.110 75213	0.118 3546	0.059 74053	WE2	6/22/21	NA
samp_3140	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	6/22/21	NA
samp_3148	NA	NA	NA	NA	NA	0.156 4986	0.228 88403	0.109 07514	0.034 68817	0.068 84539	WE4	7/13/21	NA
samp_3155	NA	NA	NA	NA	NA	0.329 91818	0.501 49051	0.257 28296	0.176 00141	0.146 62761	WE2	7/26/21	NA

samp_3156	NA	NA	NA	NA	NA	0.108 40452	0.181 60304	0.075 45254	0.121 49526	0.053 33806	WE4	7/26/21	NA
samp_3157	NA	NA	NA	NA	NA	1.590 85808	2.444 06485	1.398 99798	0.560 62202	0.206 43546	WE9	7/26/21	NA
samp_3158	NA	NA	NA	NA	NA	0.553 46468	0.850 25784	0.437 17799	0.347 21015	0.123 29686	WE12	7/26/21	NA
samp_3161	NA	NA	NA	NA	NA	1.284 02708	1.871 64937	1.086 82209	0.412 89203	0.174 75925	WE9	8/2/21	NA
samp_3162	NA	NA	NA	NA	NA	0.399 14302	0.583 92553	0.311 64203	0.201 33831	0.084 67972	WE12	8/2/21	NA
samp_3216	NA	NA	NA	NA	NA	0.115 95728	0.187 86117	0.082 15278	0.114 97002	0.114 01831	WE2	8/3/21	NA
samp_3293	NA	NA	NA	NA	NA	0.115 95728	0.187 86117	0.082 15278	0.114 97002	0.114 01831	WE2	8/3/21	NA
samp_3217	NA	NA	NA	NA	NA	0.053 77125	0.088 78301	0.039 64811	0.055 57253	0.047 30839	WE4	8/3/21	NA
samp_3218	NA	NA	NA	NA	NA	0.756 67747	1.098 3991	0.581 70453	0.352 99562	0.176 03179	WE16 -WSW	8/3/21	NA
samp_3219	NA	NA	NA	NA	NA	0.797 77084	1.151 04402	0.608 05231	0.368 65314	0.160 59518	WE12 -S	8/3/21	NA
samp_1418	NA	NA	NA	NA	NA	0.507 92004	0.756 95804	0.389 10554	0.290 48391	0.126 15639	WE2	8/9/21	NA
samp_3163	NA	NA	NA	NA	NA	0.507 92004	0.756 95804	0.389 10554	0.290 48391	0.126 15639	WE2	8/9/21	NA
samp_1416	NA	NA	NA	NA	NA	0.108 41747	0.174 78469	0.077 53578	0.110 44254	0.072 12022	WE4	8/9/21	NA
samp_3164	NA	NA	NA	NA	NA	0.108 41747	0.174 78469	0.077 53578	0.110 44254	0.072 12022	WE4	8/9/21	NA
samp_1414	NA	NA	NA	NA	NA	0.933 75113	1.383 48872	0.723 11132	0.456 85765	0.189 5359	WE6	8/9/21	NA
samp_1412	NA	NA	NA	NA	NA	0.348 10911	0.521 85949	0.265 50327	0.217 94725	0.112 47653	WE8	8/9/21	NA
samp_1410	NA	NA	NA	NA	NA	1.142 13886	1.649 56707	0.876 44388	0.515 35864	0.193 42206	WE9	8/9/21	NA
samp_3165	NA	NA	NA	NA	NA	1.142 13886	1.649 56707	0.876 44388	0.515 35864	0.193 42206	WE9	8/9/21	NA
samp_1420	NA	NA	NA	NA	NA	0.219 82382	0.332 9293	0.160 73073	0.162 65171	0.088 57667	WE16	8/9/21	NA
samp_488	NA	NA	NA	NA	NA	0.204 56021	0.330 67575	0.156 43068	0.165 22511	0.088 05989	WE2	8/16/21	NA
samp_1417	NA	NA	NA	NA	NA	0.204 56021	0.330 67575	0.156 43068	0.165 22511	0.088 05989	WE2	8/16/21	NA
samp_3167	NA	NA	NA	NA	NA	0.204 56021	0.330 67575	0.156 43068	0.165 22511	0.088 05989	WE2	8/16/21	NA
samp_1413	NA	NA	NA	NA	NA	0.717 94012	1.083 47132	0.546 39427	0.429 79523	0.163 42834	WE6	8/16/21	NA
samp_1411	NA	NA	NA	NA	NA	0.483 05594	0.732 84175	0.367 60602	0.319 18222	0.144 30252	WE8	8/16/21	NA
samp_1409	NA	NA	NA	NA	NA	1.096 40315	1.579 87618	0.823 93598	0.559 91328	0.198 11885	WE9	8/16/21	NA
samp_3169	NA	NA	NA	NA	NA	1.096 40315	1.579 87618	0.823 93598	0.559 91328	0.198 11885	WE9	8/16/21	NA
samp_490	NA	NA	NA	NA	NA	0.161 59066	0.253 17301	0.122 57109	0.125 53143	0.079 31005	WE12	8/16/21	NA
samp_1423	NA	NA	NA	NA	NA	0.161 59066	0.253 17301	0.122 57109	0.125 53143	0.079 31005	WE12	8/16/21	NA
samp_3170	NA	NA	NA	NA	NA	0.161 59066	0.253 17301	0.122 57109	0.125 53143	0.079 31005	WE12	8/16/21	NA
samp_3159	NA	NA	NA	NA	NA	0.094 76252	0.157 71029	0.072 62143	0.091 64191	0.050 5145	WE2	8/2/21	NA
samp_3160	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/2/21	NA
samp_1424	NA	NA	NA	NA	NA	0.136 3352	0.208 1936	0.101 50659	0.110 50191	0.060 57199	WE12	8/9/21	NA
samp_1498	NA	NA	NA	NA	NA	0.136 3352	0.208 1936	0.101 50659	0.110 50191	0.060 57199	WE12	8/9/21	NA
samp_3166	NA	NA	NA	NA	NA	0.136 3352	0.208 1936	0.101 50659	0.110 50191	0.060 57199	WE12	8/9/21	NA
samp_1422	NA	NA	NA	NA	NA	0.069 17946	0.109 57795	0.051 15211	0.068 11828	0.060 22833	WE13	8/9/21	NA
samp_489	NA	NA	NA	NA	NA	0.105 19439	0.174 92469	0.077 25376	0.103 58952	0.071 20775	WE4	8/16/21	NA
samp_1415	NA	NA	NA	NA	NA	0.105 19439	0.174 92469	0.077 25376	0.103 58952	0.071 20775	WE4	8/16/21	NA
samp_3168	NA	NA	NA	NA	NA	0.105 19439	0.174 92469	0.077 25376	0.103 58952	0.071 20775	WE4	8/16/21	NA

samp_3292	NA	NA	NA	NA	NA	0.105 19439	0.174 92469	0.077 25376	0.103 58952	0.071 20775	WE4	8/16/21	NA
samp_1421	NA	NA	NA	NA	NA	0.077 30914	0.123 74767	0.055 72555	0.076 74362	0.055 63001	WE13	8/16/21	NA
samp_1419	NA	NA	NA	NA	NA	0.113 48234	0.188 23678	0.089 59084	0.106 38912	0.066 97651	WE16	8/16/21	NA
samp_3171	NA	NA	NA	NA	NA	0.327 97771	0.510 9729	0.244 74129	0.236 92943	0.100 42577	WE2	8/23/21	NA
samp_3173	NA	NA	NA	NA	NA	1.075 54733	1.548 5239	0.789 68775	0.572 13063	0.235 50473	WE9	8/23/21	NA
samp_3174	NA	NA	NA	NA	NA	0.184 63252	0.282 57006	0.137 67294	0.132 09062	0.104 87934	WE12	8/23/21	NA
samp_3220	NA	NA	NA	NA	NA	0.279 00832	0.425 40524	0.203 45862	0.200 40626	0.146 62494	WE2	8/31/21	NA
samp_3222	NA	NA	NA	NA	NA	0.829 98748	1.202 18648	0.597 7176	0.502 52223	0.210 14933	WE9	8/31/21	NA
samp_3177	NA	NA	NA	NA	NA	0.795 68334	1.148 40097	0.574 55001	0.486 25941	0.184 54528	WE9	8/30/21	NA
samp_3179	NA	NA	NA	NA	NA	0.240 03262	0.370 07216	0.180 56813	0.191 17815	0.109 45175	WE2	9/7/21	NA
samp_3181	NA	NA	NA	NA	NA	0.816 09375	1.183 93742	0.597 09663	0.499 11315	0.217 56738	WE9	9/7/21	NA
samp_3182	NA	NA	NA	NA	NA	0.247 0596	0.365 73109	0.187 94112	0.176 39226	0.091 19114	WE12	9/7/21	NA
samp_3935	NA	NA	NA	NA	NA	0.247 0596	0.365 73109	0.187 94112	0.176 39226	0.091 19114	WE12	9/7/21	NA
samp_3183	NA	NA	NA	NA	NA	0.285 32743	0.415 99409	0.211 83289	0.216 39803	0.110 78372	WE2	9/13/21	NA
samp_516	NA	NA	NA	NA	NA	0.095 64616	0.152 18063	0.071 95936	0.098 21198	0.061 37568	WE4	9/13/21	NA
samp_3184	NA	NA	NA	NA	NA	0.095 64616	0.152 18063	0.071 95936	0.098 21198	0.061 37568	WE4	9/13/21	NA
samp_3185	NA	NA	NA	NA	NA	0.654 54577	0.959 27149	0.485 21237	0.457 56519	0.171 16021	WE9	9/13/21	NA
samp_3186	NA	NA	NA	NA	NA	0.252 96618	0.368 13426	0.189 24027	0.184 63368	0.139 80691	WE12	9/13/21	NA
samp_3187	NA	NA	NA	NA	NA	0.369 89747	0.558 00854	0.275 42275	0.283 97992	0.152 19524	WE2	9/20/21	NA
samp_3189	NA	NA	NA	NA	NA	0.702 15788	1.018 70783	0.513 5268	0.493 48206	0.205 19032	WE9	9/20/21	NA
samp_3190	NA	NA	NA	NA	NA	0.239 23759	0.359 92962	0.181 20597	0.177 63265	0.123 0628	WE12	9/20/21	NA
samp_3224	NA	NA	NA	NA	NA	0.405 75575	0.588 5441	0.298 58829	0.296 68329	0.150 27157	WE2	9/21/21	NA
samp_3225	NA	NA	NA	NA	NA	0.540 72905	0.785 37647	0.395 25705	0.375 40101	0.198 9182	WE9	9/21/21	NA
samp_3147	NA	NA	NA	NA	NA	0.399 34045	0.596 65012	0.287 16421	0.196 39433	0.099 58425	WE2	7/13/21	NA
samp_3149	NA	NA	NA	NA	NA	1.492 75036	2.349 24157	1.188 69093	0.671 11483	0.202 08939	WE9	7/13/21	NA
samp_3214	NA	NA	NA	NA	NA	1.589 49558	2.847 08654	1.304 478	0.938 55896	0.257 37598	WE9	7/14/21	NA
samp_3151	NA	NA	NA	NA	NA	0.156 89051	0.296 23058	0.130 8745	0.315 90162	0.071 61873	WE2	7/19/21	NA
samp_3153	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE9	7/19/21	NA
samp_3172	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/23/21	NA
samp_3221	NA	NA	NA	NA	NA	0.097 69728	0.154 24415	0.070 22934	0.084 62857	0.074 18537	WE4	8/31/21	NA
samp_3294	NA	NA	NA	NA	NA	0.097 69728	0.154 24415	0.070 22934	0.084 62857	0.074 18537	WE4	8/31/21	NA
samp_3223	NA	NA	NA	NA	NA	0.176 0373	0.267 02608	0.128 31416	0.132 08797	0.099 7292	WE12	8/31/21	NA
samp_3175	NA	NA	NA	NA	NA	0.284 35535	0.429 94183	0.210 22351	0.209 88594	0.107 59512	WE2	8/30/21	NA
samp_3176	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE4	8/30/21	NA
samp_3178	NA	NA	NA	NA	NA	0.155 77268	0.235 96198	0.113 94085	0.121 9941	0.082 72845	WE12	8/30/21	NA
samp_3180	NA	NA	NA	NA	NA	0.106 12137	0.173 01952	0.078 09255	0.103 35465	0.070 54408	WE4	9/7/21	NA
samp_3188	NA	NA	NA	NA	NA	0.137 72783	0.214 99184	0.102 40838	0.125 92246	0.097 18451	WE4	9/20/21	NA
samp_3226	NA	NA	NA	NA	NA	0.261 75711	0.383 23644	0.196 32988	0.192 8004	0.127 09554	WE12	9/21/21	NA

samp_3295	NA	NA	NA	NA	NA	0.261 75711	0.383 23644	0.196 32988	0.192 8004	0.127 09554	WE12	9/21/21	NA
samp_3191	NA	NA	NA	NA	NA	0.738 28189	1.045 43174	0.519 04037	0.525 67306	0.185 26549	WE2	9/27/21	NA
samp_3193	NA	NA	NA	NA	NA	1.213 43619	1.659 57102	1.035 24216	0.475 08341	0.246 01329	WE9	9/27/21	NA
samp_3192	NA	NA	NA	NA	NA	0.166 91888	0.237 02758	0.122 91036	0.120 10881	0.093 38059	WE4	9/27/21	NA
samp_3194	NA	NA	NA	NA	NA	0.196 53059	0.286 38601	0.148 8042	0.154 69067	0.102 83101	WE12	9/27/21	NA
samp_3197	NA	NA	NA	NA	NA	0.867 9709	1.237 78677	0.699 66551	0.440 45894	0.221 81129	WE9	10/4/21	NA
samp_3195	NA	NA	NA	NA	NA	0.305 73078	0.455 59256	0.226 62188	0.238 17533	0.132 52341	WE2	10/4/21	NA
samp_3196	NA	NA	NA	NA	NA	0.088 36079	0.141 2398	0.067 17622	0.080 52981	0.064 79799	WE4	10/4/21	NA
samp_3198	NA	NA	NA	NA	NA	0.152 26611	0.228 93164	0.114 20109	0.126 5861	0.101 16213	WE12	10/4/21	NA
samp_3199	NA	NA	NA	NA	NA	0.537 80189	0.785 41039	0.408 25532	0.318 65949	0.177 91027	WE2	10/12/21	NA
samp_3201	NA	NA	NA	NA	NA	1.410 57349	2.019 15385	1.196 96181	0.531 66295	0.250 31048	WE9	10/12/21	NA
samp_3202	NA	NA	NA	NA	NA	0.315 1123	0.460 14095	0.235 02013	0.200 51909	0.159 22228	WE12	10/12/21	NA
samp_3203	NA	NA	NA	NA	NA	0.479 58925	0.692 3179	0.362 57028	0.304 60802	0.166 98832	WE2	10/18/21	NA
samp_3205	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	WE9	10/18/21	NA
samp_3206	NA	NA	NA	NA	NA	0.472 87066	0.668 57853	0.361 7826	0.281 53107	0.163 5571	WE12	10/18/21	NA
samp_3200	NA	NA	NA	NA	NA	0.063 92646	0.098 97028	0.048 90245	0.029 64631	0.068 81178	WE4	10/12/21	NA
samp_3204	NA	NA	NA	NA	NA	0.070 86286	0.109 59088	0.052 46087	0.074 51895	0.042 74546	WE4	10/18/21	NA
samp_3207	NA	NA	NA	NA	NA	0.260 75601	0.383 87514	0.207 57475	0.161 26045	0.097 66742	WE2	10/27/21	NA

Table S2.9.2: HPLC Solvent Gradient.

Time (minutes)	% Mobile Phase B
0	0
2	0
2.1	20
10	100
14	100
14.1	0
17	0

A gradient was run over 17 minutes at a flow rate of 0.6 mL/min starting at 0% mobile phase B for 2 minutes, then increasing from 20% to 100% mobile phase B over 8 minutes, holding at 100% mobile phase B for 4 minutes, before decreasing back to 0% for 3 minutes.

Table S2.9.3: Biosynthetic gene cluster (BGC) database information

BGC Name in Fasta Amended to Gene	Name in Study	Cluster	Pathway	Organism	Detected Via
1-heptadecene_20_0	1-heptadecene	3	PKS	*Microcystis*	WLE CC
MicrocystisaeruginosaNIES-98_aeruginosin98-A,aeruginosin98-B,aeruginosin98-C2	Aeruginosin	3	NRPS-PKS	*Microcystis*	MIBiG
Anabaenopeptin_16_0	Anabaenopeptin	5	NRPS	*Microcystis*	WLE CC
Bartoloside-T1PKS_18_0	Bartoloside-Like	6	PKS	*Microcystis*	WLE CC
Cyanobactin_50_0	Cyanobactin-3 (Microcystis)	4	RiPP-Cyanobactin	*Microcystis*	WLE CC
Microcystissp.NIVA-CYA172/5_cyanopeptolin	Cyanopeptolin (Microcystis)	5	NRPS	*Microcystis*	MIBiG
MicrocystisaeruginosaNIES-87_kasumigamide	Kasumigamide	1	NRPS-PKS	*Microcystis*	MIBiG
RRE-containing_23_0	LAP (Microcystis)	5	RiPP-LAP	*Microcystis*	WLE CC
lanthipeptide_21_0	Lanthipeptide-3 (Microcystis)	6	RiPP-Lanthipeptide	*Microcystis*	WLE CC
MicrocystisaeruginosaPCC7806_microcystinLR	Microcystin	3	NRPS-PKS	*Microcystis*	MIBiG
MicrocystisaeruginosaLEGE91341_microginin	Microginin	4	NRPS-PKS	*Microcystis*	MIBiG
MicrocystisaeruginosaNIES-298_microviridinB	Microviridin	4	RiPP	*Microcystis*	MIBiG
NRPS-T1PKS_26_0	NRPS-T1PKS	3	NRPS-PKS	*Microcystis*	WLE CC
NRPS-like_34_0	NRPS-like-1	1	NRPS	*Microcystis*	WLE CC
T3PKS-Merocyclophane_5_0	PmNT3	5	PKS	*Microcystis*	WLE CC
T1PKS-Terpene_11_0	T1PKS-Terpene	5	PKS	*Microcystis*	WLE CC
Cyanobactin_90_0	Cyanobactin-1 (Microcystis)	6	RiPP-Cyanobactin	*Microcystis*	WLE CC
Cyanobactin_64_0	Cyanobactin-2 (Microcystis)	6	RiPP-Cyanobactin	*Microcystis*	WLE CC
MicrocystisaeruginosaPCC7005_piricyclamide7005E1,piricyclamide7005E2,piricyclamide7005E3,piricyclamide7005E4	Piricyclamide	2	RiPP-Cyanobactin	*Microcystis*	MIBiG
T1PKS_37_0	T1PKS-2	4	PKS	*Microcystis*	WLE CC
Novel-T3PKS_28_0	T3PPKS-1	6	PKS	*Microcystis*	WLE CC
hglE-KS_45_0	hglE-KS-like	4	PKS	*Microcystis*	WLE CC
RiPP-like_84_0	Ripp-like-1	6	RiPP	*Microcystis*	WLE CC
Anabaenasp.90_anacyclamideA10	Anacyclamide (ADA)	8	RiPP-Cyanobactin	ADA	MIBiG
samp_2073_477589_LAP	LAP (ADA)	7	RiPP-LAP	ADA	WLE MAGs (Denuyl et al. 2025,

					ES&T)
samp_471_14978_NRPS	NRPS-4 (ADA)	7	NRPS	ADA	WLE MAGs (Den uyl et al. 2025, *ES&T*)
DolichospermumcircinaleAWQC131C_decarbamoylgonyautoxin2,decarbamoylgonyautoxin3,decarbamoylsaxitoxin,gonyautoxin2,gonyautoxin3,gony	Saxitoxin	7	Other	ADA	MIBiG
MicrocystisaeruginosaPCC7806_microcyclamide	Aerucyclamide	9	RiPP-Cyanobactin	*Microcystis*	MIBiG
samp_4305_693378_NRPS	NRPS-1 (ADA)	9	NRPS	ADA	WLE MAGs (Den uyl et al. 2025, ES&T)
RiPP-like_80_0	Ripp-like-2	9	RiPP	*Microcystis*	WLE CC
lanthipeptide_87_0	Lanthipeptide-1 (Microcystis)	9	RiPP-Lanthipeptide	*Microcystis*	WLE CC
lanthipeptide_54_0	Lanthipeptide-2 (Microcystis)	9	RiPP-Lanthipeptide	*Microcystis*	WLE CC
samp_4305_474408_T1PKS	T1PKS (ADA)	8	PKS	ADA	WLE MAGs (Den uyl et al. 2025, ES&T)
T1PKS_51_0	T1PKS-1	9	PKS	*Microcystis*	WLE CC
samp_4380_82358_NRPS-like	NRPS-2 (ADA)	8	NRPS	ADA	WLE MAGs (Den uyl et al. 2025, ES&T)
samp_4305_554341_lanthipeptide-class-v	Lanthipeptide (ADA)	9	RiPP-Lanthipeptide	ADA	WLE MAGs (Den uyl et al. 2025, ES&T)

Acronyms: western Lake Erie culture collection (WLECC), Aphanizomenon, Dolichospermum, Anabaena (ADA), polyketide synthase (PKS), nonribosomal peptide synthetase (NRPS), Ribosomally synthesized and post-translationally modified peptides (RiPPs), WLE MAGs (Den uyl et al. 2025, ES&T) = western Lake Erie metagenome assembled genomes from Paul Den Uyl et al. study 2025 in ACS ES&T. Gene specific information found in seed5 --> BGCdb_seed5_key

Table S2.9.4: *Microcystis* genomes used for phylogenomic analysis

Genome Name
ATCC29413.fa
LE17-10A.fa
LE17-20C.fa
LE18-22.fa
LE19-10.fa
LE19-114.fa
LE19-12.fa
LE19-131.fa
LE19-13.fa
LE19-195.fa
LE19-196.fa
LE19-197.fa
LE19-251.fa
LE19-281.fa
LE19-338.fa
LE19-388.fa
LE19-41.fa
LE19-4.fa
LE19-55.fa
LE19-59.fa
LE19-84.fa
LE19-8.fa
LE19-98.fa
Microcystis_0824.fa
Microcystis_aeruginosa_11-30S32.fa
Microcystis_aeruginosa_BK11-02.fa
Microcystis_aeruginosa_BS11-05.fa
Microcystis_aeruginosa_BS13-02.fa
Microcystis_aeruginosa_BS13-10.fa
Microcystis_aeruginosa_CACIAM_03.fa
Microcystis_aeruginosa_CHAOHU_1326.fa
Microcystis_aeruginosa_DA14.fa
Microcystis_aeruginosa_DIANCHI905.fa
Microcystis_aeruginosa_EAWAG127a.fa

Microcystis_aeruginosa_F13-15.fa
Microcystis_aeruginosa_FACHB-524.fa
Microcystis_aeruginosa_FCY-26.fa
Microcystis_aeruginosa_FCY-27.fa
Microcystis_aeruginosa_FCY-28.fa
Microcystis_aeruginosa_FD4.fa
Microcystis_aeruginosa_G11-01.fa
Microcystis_aeruginosa_G11-04.fa
Microcystis_aeruginosa_G11-06.fa
Microcystis_aeruginosa_G11-09.fa
Microcystis_aeruginosa_G13-01.fa
Microcystis_aeruginosa_G13-03.fa
Microcystis_aeruginosa_G13-05.fa
Microcystis_aeruginosa_G13-07.fa
Microcystis_aeruginosa_G13-09.fa
Microcystis_aeruginosa_G13-10.fa
Microcystis_aeruginosa_G13-11.fa
Microcystis_aeruginosa_G13-12.fa
Microcystis_aeruginosa_K13-05.fa
Microcystis_aeruginosa_K13-06.fa
Microcystis_aeruginosa_K13-07.fa
Microcystis_aeruginosa_K13-10.fa
Microcystis_aeruginosa_KLA2.fa
Microcystis_aeruginosa_KW.fa
Microcystis_aeruginosa_L111-01.fa
Microcystis_aeruginosa_L211-07.fa
Microcystis_aeruginosa_L211-101.fa
Microcystis_aeruginosa_L211-11.fa
Microcystis_aeruginosa_L311-01.fa
Microcystis_aeruginosa_LE13-04.fa
Microcystis_aeruginosa_LE3.fa
Microcystis_aeruginosa_LG11-05.fa
Microcystis_aeruginosa_LG13-03.fa
Microcystis_aeruginosa_LG13-11.fa
Microcystis_aeruginosa_LG13-12.fa
Microcystis_aeruginosa_LG13-13.fa
Microcystis_aeruginosa_LL11-07.fa
Microcystis_aeruginosa_LL13-03.fa
Microcystis_aeruginosa_LL13-06.fa

Microcystis_aeruginosa_Ma_AC_P_19900807_S299.fa
Microcystis_aeruginosa_Ma_AC_P_19900807_S300.fa
Microcystis_aeruginosa_Ma_MB_F_20061100_S19D.fa
Microcystis_aeruginosa_Ma_MB_F_20061100_S19.fa
Microcystis_aeruginosa_Ma_MB_F_20061100_S20D.fa
Microcystis_aeruginosa_Ma_MB_F_20061100_S20.fa
Microcystis_aeruginosa_Ma_MB_S_20031200_S102D.fa
Microcystis_aeruginosa_Ma_MB_S_20031200_S102.fa
Microcystis_aeruginosa_Ma_OC_H_19870700_S124.fa
Microcystis_aeruginosa_Ma_OC_LR_19540900_S633.fa
Microcystis_aeruginosa_Ma_QC_B_20070730_S2.fa
Microcystis_aeruginosa_Ma_QC_C_20070703_M131.fa
Microcystis_aeruginosa_Ma_QC_C_20070823_S13D.fa
Microcystis_aeruginosa_Ma_QC_C_20070823_S13.fa
Microcystis_aeruginosa_Ma_QC_Ca_00000000_S207.fa
Microcystis_aeruginosa_Ma_QC_Ch_20071001_M135.fa
Microcystis_aeruginosa_Ma_QC_Ch_20071001_S25D.fa
Microcystis_aeruginosa_Ma_QC_Ch_20071001_S25.fa
Microcystis_aeruginosa_Ma_SC_T_19800800_S464.fa
Microcystis_aeruginosa_NaRes975.fa
Microcystis_aeruginosa_NIES-1211.fa
Microcystis_aeruginosa_NIES-2481.fa
Microcystis_aeruginosa_NIES-2519.fa
Microcystis_aeruginosa_NIES-2520.fa
Microcystis_aeruginosa_NIES-2521.fa
Microcystis_aeruginosa_NIES-2522.fa
Microcystis_aeruginosa_NIES-2549.fa
Microcystis_aeruginosa_NIES-298.fa
Microcystis_aeruginosa_NIES-3787.fa
Microcystis_aeruginosa_NIES-3804.fa
Microcystis_aeruginosa_NIES-3806.fa
Microcystis_aeruginosa_NIES-3807.fa
Microcystis_aeruginosa_NIES-4264.fa
Microcystis_aeruginosa_NIES-4285.fa
Microcystis_aeruginosa_NIES-4325.fa
Microcystis_aeruginosa_NIES-44.fa
Microcystis_aeruginosa_NIES-843.fa
Microcystis_aeruginosa_NIES-87.fa
Microcystis_aeruginosa_NIES-88.fa

Microcystis_aeruginosa_NIES-98.fa
Microcystis_aeruginosa_PCC_7005.fa
Microcystis_aeruginosa_PCC_7806SL.fa
Microcystis_aeruginosa_PCC_7941.fa
Microcystis_aeruginosa_PCC_9432.fa
Microcystis_aeruginosa_PCC_9443.fa
Microcystis_aeruginosa_PCC_9701.fa
Microcystis_aeruginosa_PCC_9717.fa
Microcystis_aeruginosa_PCC_9806.fa
Microcystis_aeruginosa_PCC_9807.fa
Microcystis_aeruginosa_PCC_9808.fa
Microcystis_aeruginosa_PCC_9809.fa
Microcystis_aeruginosa_S11-01.fa
Microcystis_aeruginosa_S11-05.fa
Microcystis_aeruginosa_Sj.fa
Microcystis_aeruginosa_SPC777.fa
Microcystis_aeruginosa_SX13-01.fa
Microcystis_aeruginosa_SX13-11.fa
Microcystis_aeruginosa_TA09.fa
Microcystis_aeruginosa_TAIHU98.fa
Microcystis_aeruginosa_W11-03.fa
Microcystis_aeruginosa_W11-06.fa
Microcystis_aeruginosa_W13-11.fa
Microcystis_aeruginosa_W13-13.fa
Microcystis_aeruginosa_W13-15.fa
Microcystis_aeruginosa_W13-16.fa
Microcystis_aeruginosa_W13-18.fa
Microcystis_flos-aquae_DF17.fa
Microcystis_flos-aquae_Ma_QC_C_20070823_S18D.fa
Microcystis_flos-aquae_Ma_QC_C_20070823_S18.fa
Microcystis_flos-aquae_Mf_QC_C_20070823_S10D.fa
Microcystis_flos-aquae_Mf_QC_C_20070823_S10.fa
Microcystis_flos-aquae_Mf_QC_C_20070823_S20D.fa
Microcystis_flos-aquae_Mf_QC_C_20070823_S20.fa
Microcystis_flos-aquae_Mf_QC_C_20070823_S20T.fa
Microcystis_flos-aquae_Mf_WU_F_19750830_S460.fa
Microcystis_flos-aquae_TF09.fa
Microcystis_LSC13-02.fa
Microcystis_MC19.fa

Microcystis_M_OC_Ca_00000000_C217Col.fa
Microcystis_M_OC_Ca_00000000_S217Cul.fa
Microcystis_M_QC_C_20170808_M2Col.fa
Microcystis_M_QC_C_20170808_M3Col.fa
Microcystis_M_QC_C_20170808_M9Col.fa
Microcystis_Msp_OC_L_20101000_S702.fa
Microcystis_novacekii_Mn_MB_F_20050700_S1D.fa
Microcystis_novacekii_Mn_MB_F_20050700_S1.fa
Microcystis_panniformis_FACHB-1757.fa
Microcystis_panniformis_Mp_GB_SS_20050300_S99D.fa
Microcystis_panniformis_Mp_GB_SS_20050300_S99.fa
Microcystis_panniformis_Mp_MB_F_20051200_S6D.fa
Microcystis_panniformis_Mp_MB_F_20051200_S6.fa
Microcystis_panniformis_Mp_MB_F_20051200_S9D.fa
Microcystis_panniformis_Mp_MB_F_20051200_S9.fa
Microcystis_panniformis_Mp_MB_F_20080800_S26D.fa
Microcystis_panniformis_Mp_MB_F_20080800_S26.fa
Microcystis_T1-4.fa
Microcystis_viridis_Mv_BB_P_19951000_S68D.fa
Microcystis_viridis_Mv_BB_P_19951000_S68.fa
Microcystis_viridis_Mv_BB_P_19951000_S69D.fa
Microcystis_viridis_Mv_BB_P_19951000_S69.fa
Microcystis_viridis_NIES-102.fa
Microcystis_wesenbergii_LE013-01.fa
Microcystis_wesenbergii_Mw_MB_S_20031200_S109D.fa
Microcystis_wesenbergii_Mw_MB_S_20031200_S109.fa
Microcystis_wesenbergii_Mw_QC_B_20070930_S4D.fa
Microcystis_wesenbergii_Mw_QC_B_20070930_S4.fa
Microcystis_wesenbergii_Mw_QC_S_20081001_S30D.fa
Microcystis_wesenbergii_Mw_QC_S_20081001_S30.fa
Microcystis_wesenbergii_TW10.fa
PCC_6803.fa

Table 2.9.5: MZmine Parameters used in the feature-finding processing wizard.

Parameter	Input
Smoothing	Checked
Stable ionization across samples	Checked
Crop retention time	0.50-18 minutes
Max peaks in chromatogram	15
Minimum consecutive scans	3
Approximate feature FWHM	0.10 minutes
RT Tolerance (intra sample)	0.08 minutes
RT Tolerance (sample to sample)	0.40 minutes
Ion mode	Positive, Factor of lowest signal
Noise threshold (MS1)	10
Noise threshold (MS2	5
Minimum feature height	5.00E+04
m/z tolerance (scan to scan)	0.0020 m/z
m/z tolerance (intra sample)	0.0015 m/z
m/z tolerance (sample to sample)	0.0015 m/z
Minumum samples per aligned feature	Max of 2 samples
Annotate lipids	unchecked
Files exported for molecular networking and SIRIUS	Checked

Table S2.9.6: Hierarchical general additive models (HGAM) equation and results for the relative abundance (%) of *Microcystis* and *Dolichospermum* relative to the whole microbial community.

HGAM Equati	log(Summed abundance per Genus) ~ s(yday, bs = "cc") + s(yday, site, bs = ("fs") + te(year, site, bs = "re") + s(yday, year, bs = "fs") + s(site, bs = "re") + s(year, bs = "re"), family = tw()								
Organism	s(yday)	s(yday, site)	te(year, site)	s(yday, year)	s(site)	s(year)	% Deviance Explained	AIC	% Additional Deviance Explained by Year or Site
Microcystis	5.82E-06	0.0474	0.4714	< 2e-16	0.4921	0.3125	57.70%	261.9082	Year = 27.8%, Site = 1.9%
ADA	5.98E-07	0.06813	2.97E-03	< 2e-16	0.2236	0.0000	76.10%	223.3221	Year = 36.3% , Site = 8.2%
									Calculated by reducing model to exclude terms containing either site or year to see how deviance explained changed
									Year = log(Summed BGC abundance per Genus) ~ s(yday, bs = "cc") + s(yday, site, bs = ("fs") + s(site, bs = "re") , family = tw()
									Site = log(Summed BGC abundance per Genus) ~ s(yday, bs = "cc") + s(yday, year, bs = "fs") + s(year, bs = "re"), family = tw()

Table S2.9.7: R^2 and *P*-values for NMDS vectors in Figure 2C.

Variable	NMDS1	NMDS2	R^2	*P* -value
Nitrate + Nitrite	0.10434	-0.99454	0.0465	0.144
Particulate Microcystin	0.76005	-0.64986	0.3046	0.001
Soluble Reactive Phosphorus (SRP)	-0.00402	0.99999	0.0342	0.195
Particulate Organic Carbon (POC)	0.90378	-0.42799	0.2065	0.001
Particulate Organic Nitrogen (PON)	0.93678	-0.34992	0.2218	0.001
Temperature	0.18667	-0.98242	0.1548	0.004
Depth	-0.23742	-0.97141	0.0731	0.051
Dissovled Oxygen	0.66417	0.74758	0.0128	0.578
Conductivity	0.95561	0.29465	0.0676	0.056
Turbidity	0.91657	0.39987	0.0189	0.47
Phycocyanin	0.99356	0.11327	0.2881	0.001
Total Phosphorus	0.0018	-1	0.0017	0.93
Dissolved Phosphorus	0.01121	0.99994	0.016	0.476
Ammonia	0.57752	-0.81638	0.0019	0.939
Colored Dissolved Organic Matter (CDOM)	0.97979	0.20002	0.0276	0.314
Susepended Particulate Matter	0.73293	0.68031	0.0212	0.426
Terrestrial/Humic DOM	0.94599	0.32421	0.0224	0.36
Microbial:Terrestrial DOM Ratio	-0.97419	-0.22575	0.0862	0.033
Aromatic AAs in DOM	-0.99874	-0.05014	0.0879	0.027

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Permutation: free
Number of permutations: 999

Table S2.9.8: Hierarchical general additive models (HGAM) equation and results for the FPKM of BGCs within modules 4, 5, 6 and 7.

HGAM Equation	log(Module abundance) ~ s(yday, bs = "cc") + ti(yday, site, bs = ("cc", "re") + s(year, site, bs = "re") + s(site, bs = "re"), fa						
Module	s(yday)	ti(yday, site)	s(year, site)	s(site)	% Deviance Explained	AIC	% Additional Deviance Explained by Year or Site
Module 4 (blue)	1.66E-05	0.554	0.138	0.464	26.30%	455.675	Year = 8.5, Site = 8.2%
Module 5 (orange)	0.0378	0.8034	0.0209	0.1445	28.20%	466.666	Year = 18.42, Site = 14.7%
Module 6 (green)	5.90E-07	0.8253	0.0706	0.5232	33.30%	411.156	Year = 11.2, Site = 7.1
Module 7 (magenta)	2.86E-03	0.55445	2.99E-05	0.43787	39.40%	468.768	Year = 32.79, Site = 15.4

Table S2.9.9: Residual standard errors of loess fits of metabolite, dissolved organic matter (DOM), gene, nitrogen, and grazer abundance patterns.

Variable	Variable Type	Figure	Line Color	RSE
Microcystin BGC	BGC	6B	Black-dashed	0.79540969
Aeruginosin BGC	BGC	6B	Black-dashed	0.82690029
Anabaenopeptin BGC	BGC	6B	Black-dashed	1.01694828
Aerucyclamide BGC	BGC	6B	Black-dashed	0.90128274
Microcystin	Normalized Abundance of Putatively Annotated Metabolite	7A	Blue	1.01268857
Aeruginosin	Normalized Abundance of Putatively Annotated Metabolite	7A	Yellow	0.8486629
Anabaenopeptin	Normalized Abundance of Putatively Annotated Metabolite	7A	Green	0.93646401
Aerucyclamide	Normalized Abundance of Putatively Annotated Metabolite	7A	Berry	1.01300712
Nitrate & Nitrite	Measured Physicochemical Parameter	7A	Black	0.67203844
Ammonia	Measured Physicochemical Parameter	7A	Black	13.7975771
POC:PON	Measured Physicochemical Parameter	7A	Black	1.17660177
Terrestrial/Humic DOM (Fmax3)	DOM Dervied from FDOM	7A	Black	0.22959764
Microbial:Terrestrial DOM (Fmax1+4+5)/Fmax3)	DOM Dervied from FDOM	7A	Black	0.69495006
% Aromatic Amino Acids in DOM (Fmax 4+5/Fmax3	DOM Dervied from FDOM	7A	Black	0.06240958
Daphnia	Calculated abundance of contigs annotated as Mussels	7A	Black	7.24011408
Mussels	Calculated abundance of contigs annotated as Daphnia	7A	Black	0.01470439
Amoeba	Calculated abundance of contigs annotated as Amoeba	7A	Black	0.00379085

2.9.2 SI Figures

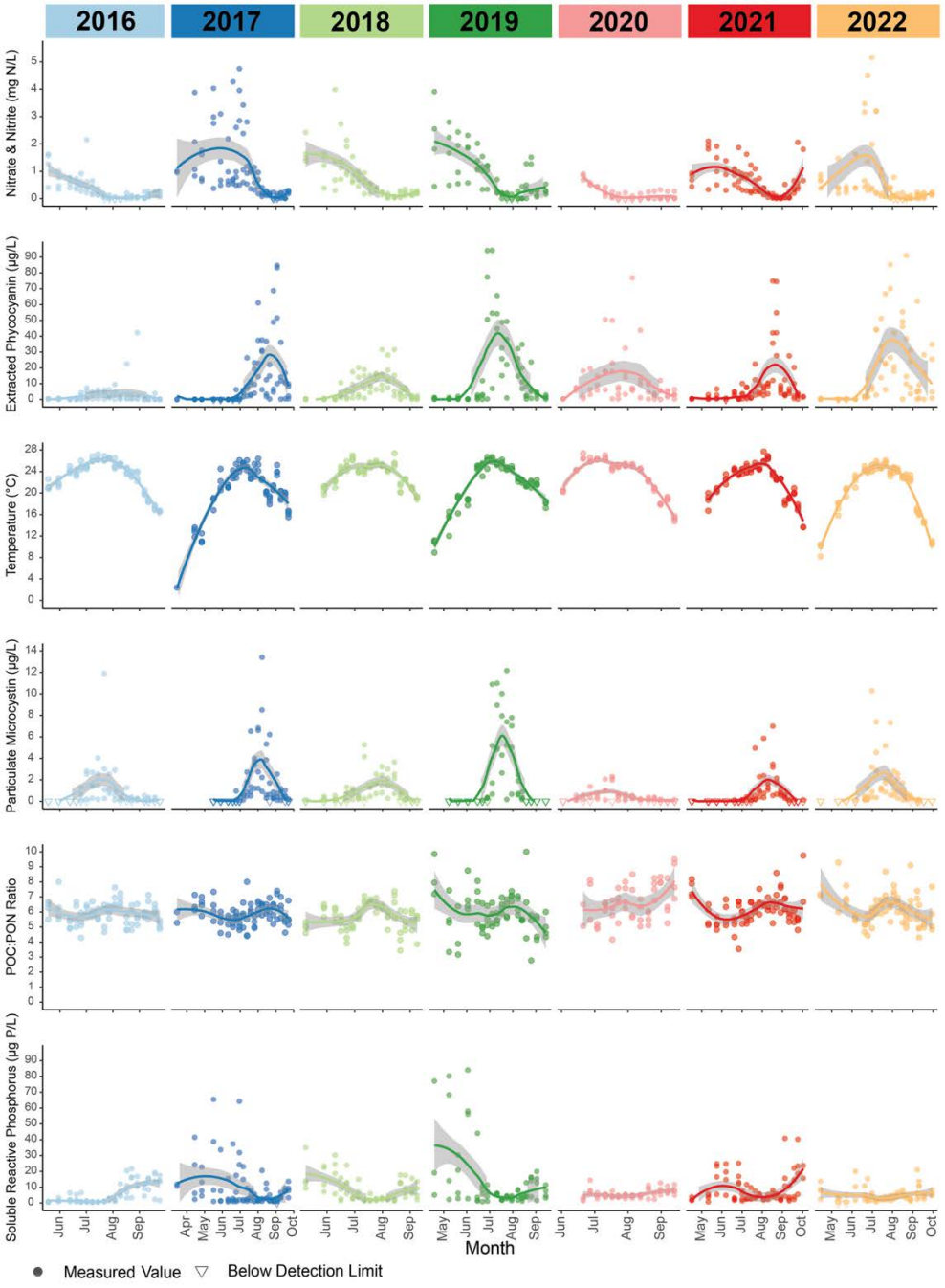
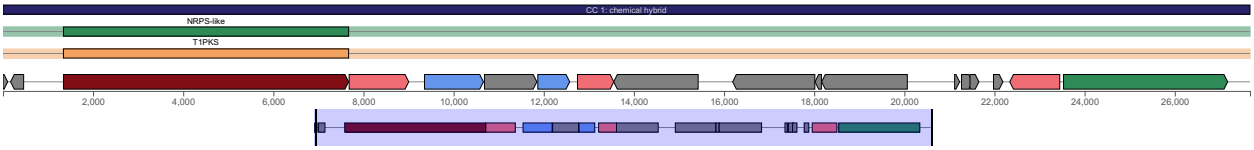
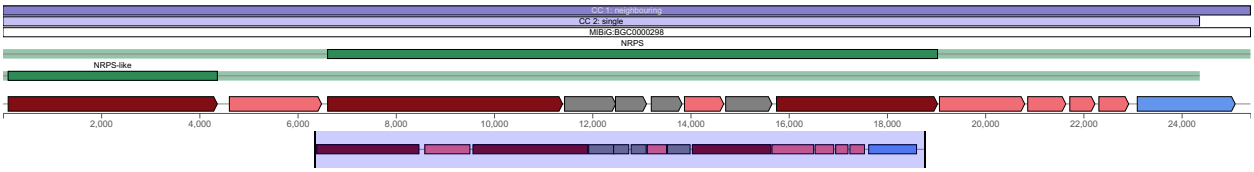


Figure S2.9.1: Comparison of nitrate + nitrite (mg/L), extracted phycocyanin (µg/L), temperature (°C), particulate microcystin (µg/L), particulate organic carbon: particulate organic nitrogen ratio (POC:PON), and soluble reactive phosphorus (µg P/L) from all stations between the years of 2016 and 2022. Individual observations are plotted as filled in circles, with observations below the detection limit plotted as unfilled triangles. Values under the detection limit were only identified for particulate microcystins and nitrate + nitrite and were imputed as half the limit of detection. Smoothed lines were fit using a loess function.

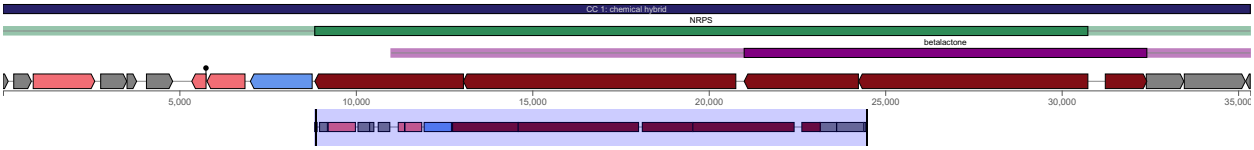
A. 1-heptadecene_20_0



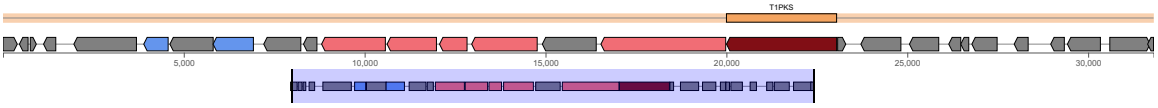
B. *Microcystis aeruginosa* NIES-98_aeruginosin98-A, aeruginosin98-B, aeruginosin 98-C2



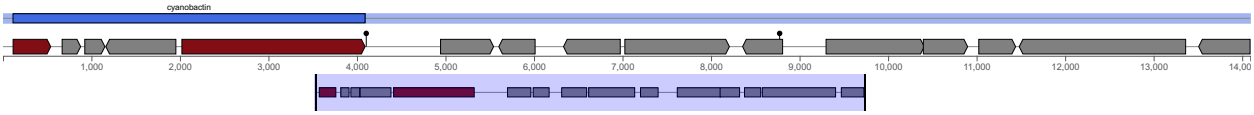
C. Anabaenopeptin_16_0



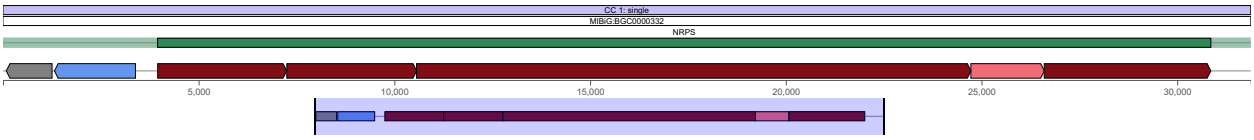
D. Bartoloside-T1PKS_18_0



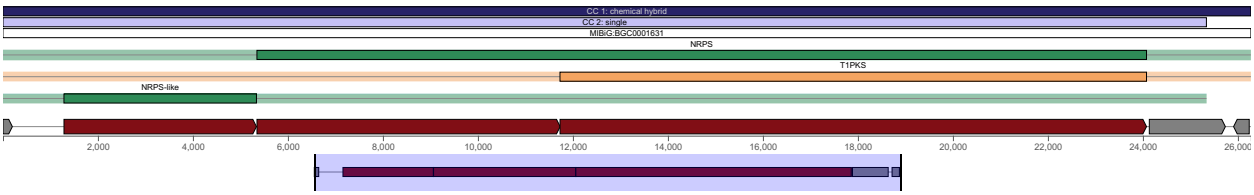
E. Cyanobactin_50_0



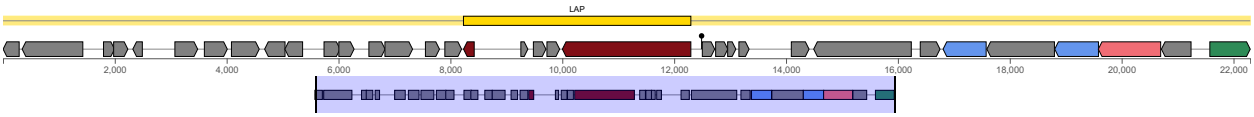
F. *Microcystis* sp. NIVA-CYA172/5_cyanopeptolin



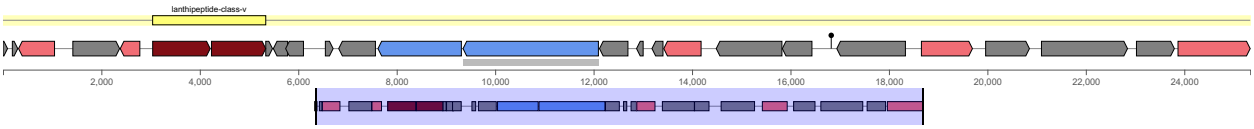
G. *Microcystis aeruginosa* NIES-87_kasumigamide



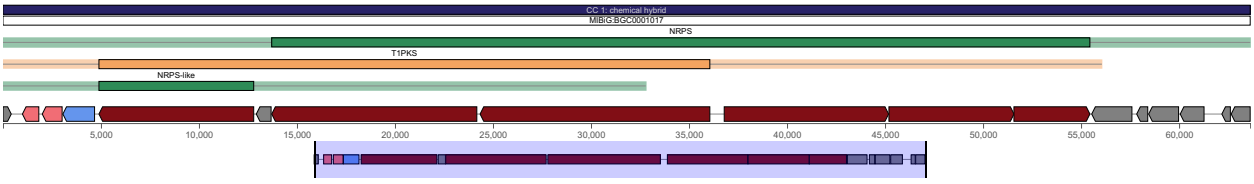
H. RRE-containing_23_0



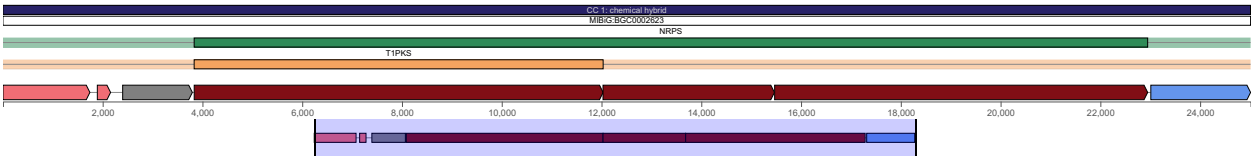
I. lanthipeptide_21_0



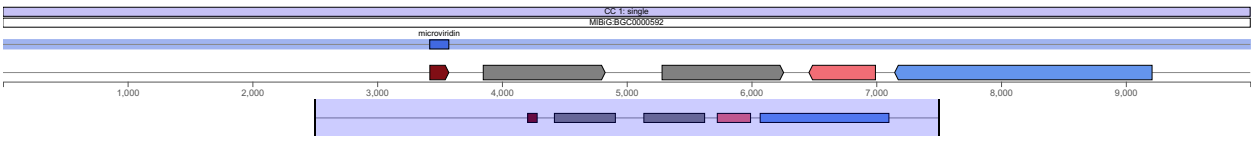
J. *Microcystis aeruginosa* PCC7806_microcystin LR



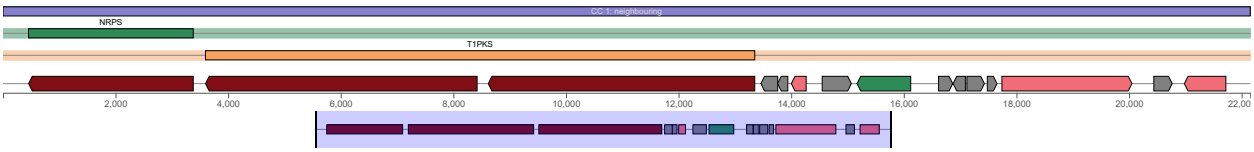
K. *Microcystis aeruginosa* LEGE91341_microginin



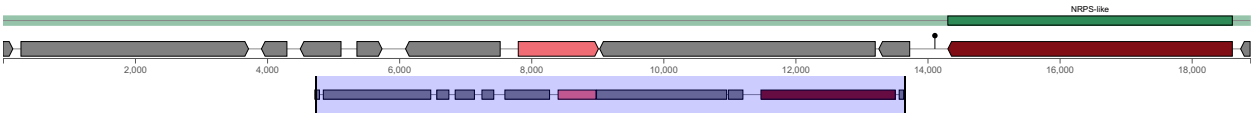
L. *Microcystis aeruginosa* NIES-298_microviridinB



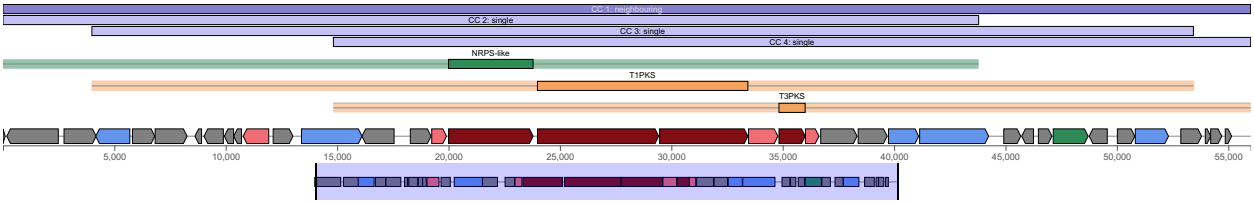
M. NRPS-T1PKS_26_0



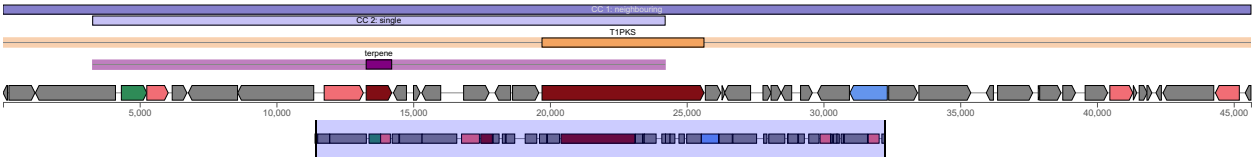
N. NRPS-like_34_0



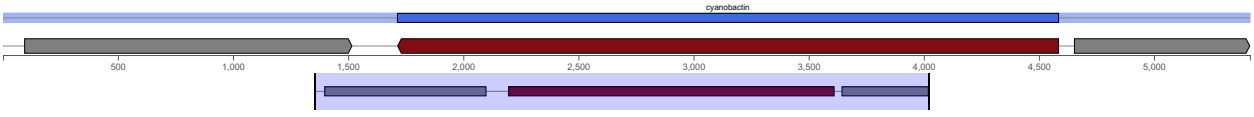
O. T3PKS-Merocyclophane_5_0



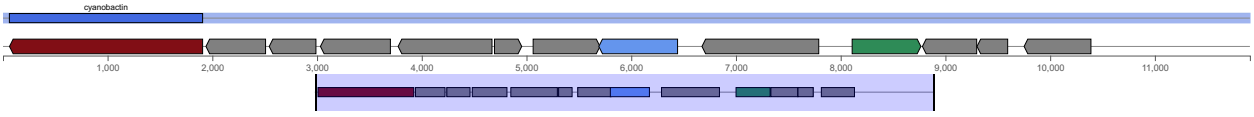
P. T1PKS-Terpene_11_0



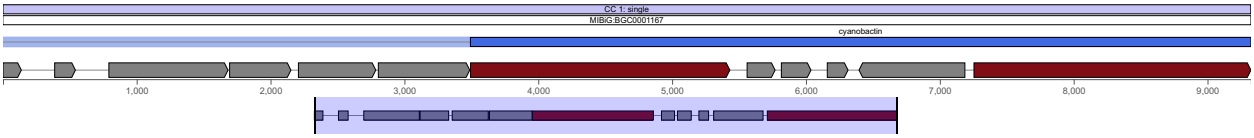
Q. Cyanobactin_90_0



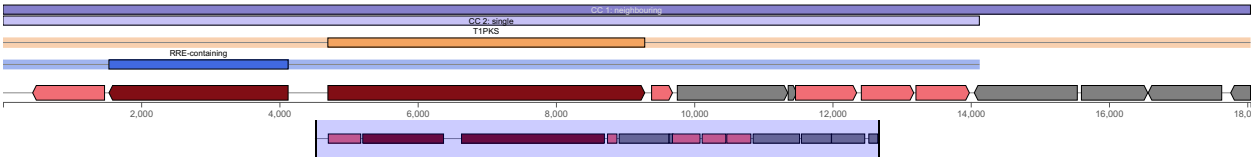
R. Cyanobactin_64_0 (Cyanobactin 2)



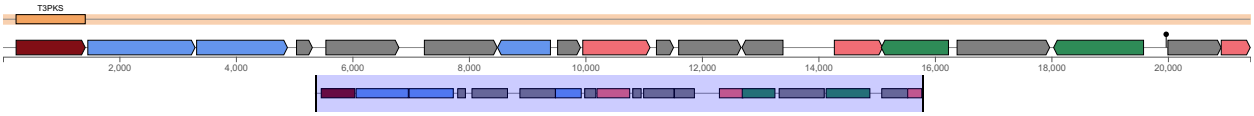
S. *Microcystis aeruginosa* PCC7005_piricyclamide7005E1, piricyclamide7005E2, piricyclamide7005E3, piricyclamide7005E4



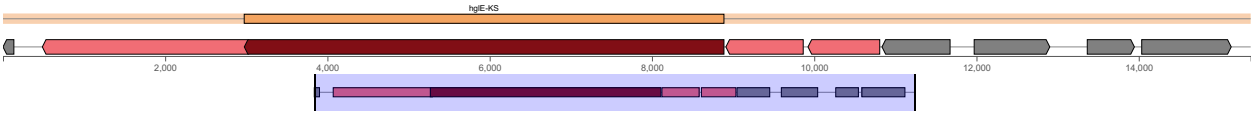
T. T1PKS_37_0



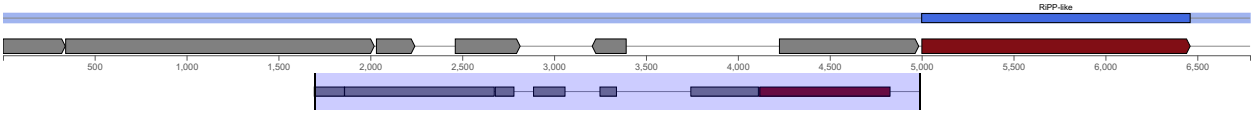
U. Novel-T3PKS_28_0



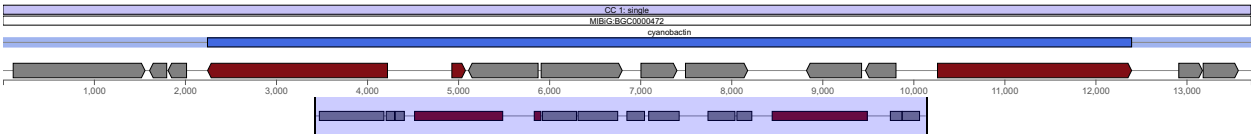
V. hglE-KS_45_0



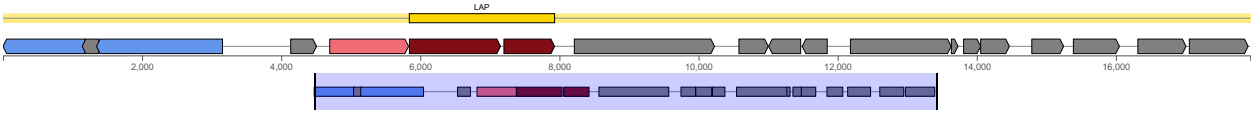
W. RiPP-like_84_0



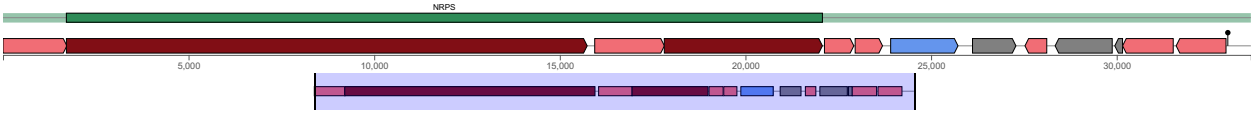
X. Anabaena sp.90_anacyclamide A10



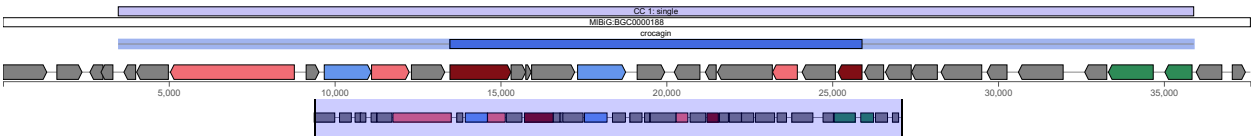
Y. samp_2073_477589_LAP



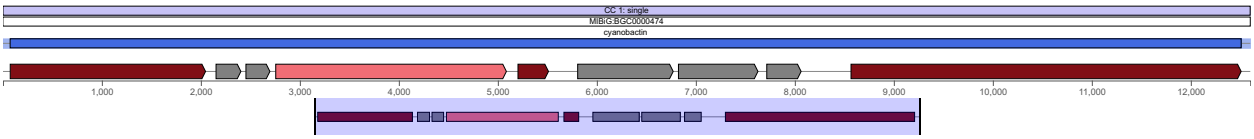
Z. samp_471_14978_NRPS

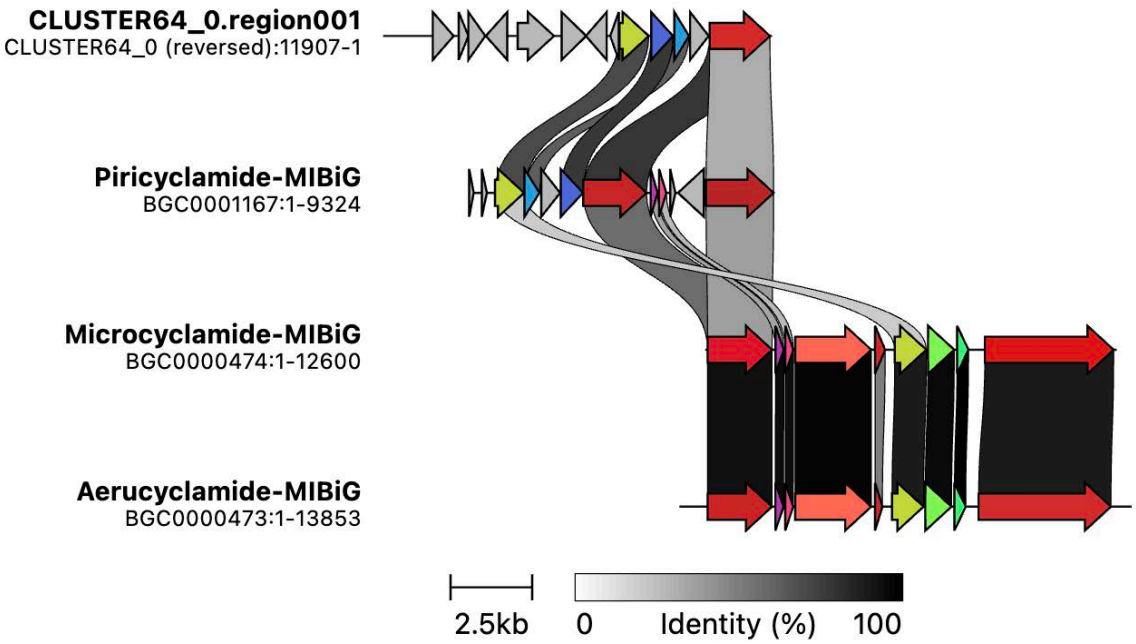


AA. *Dolichospermum circinale* AWQC131C_decarbamoylgonyautoxin 2, decarbamoylgonyautoxin 3, decarbamoylsaxitoxin, gonyautoxin 2, gonyautoxin 3

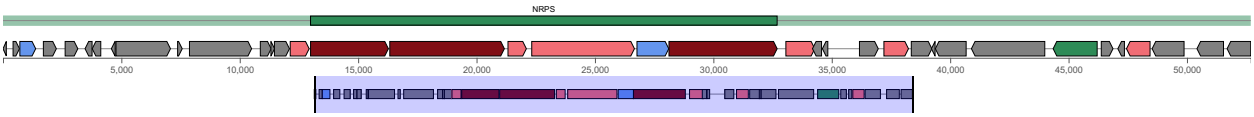


BB. ***Microcystis aeruginosa*** **PCC7806_aerucyclamide/microcyclamide**





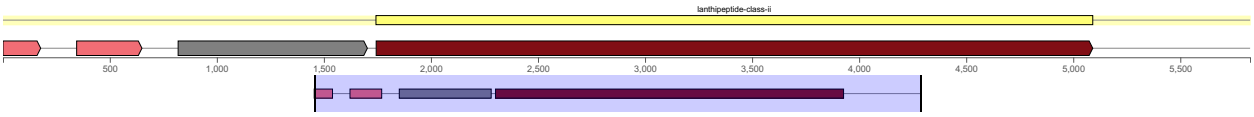
CC. **samp_4305_693378_NRPS**



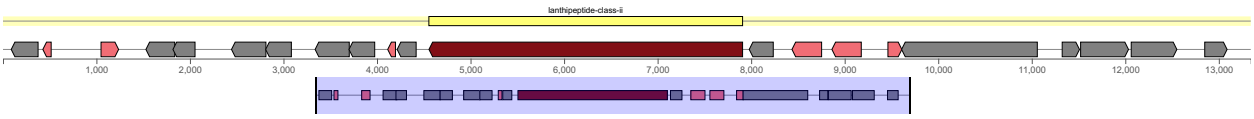
DD. **RiPP-like_80_0**



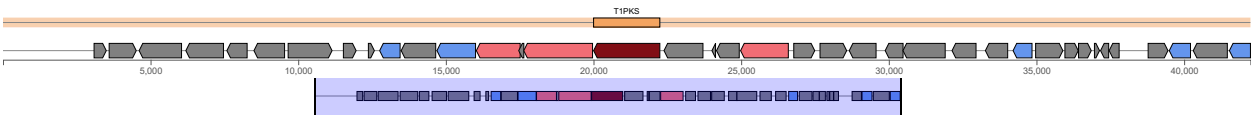
EE. **lanthipeptide_87_0**



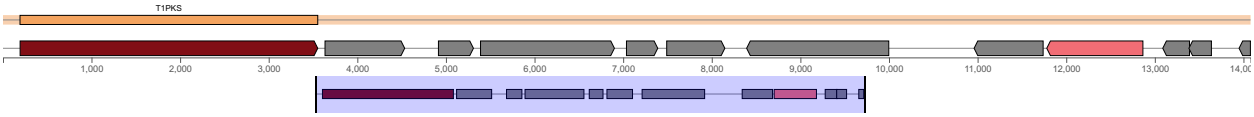
FF.lanthipeptide_54_0



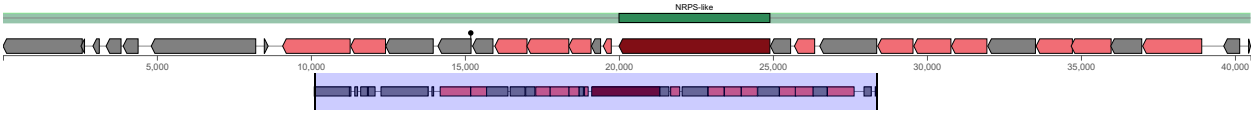
GG. samp_4305_474408_T1PKS



HH. T1PKS_51_0



II. samp_4380_82358_NRPS-like



JJ. samp_4305_554341_lanthipeptide-class-v

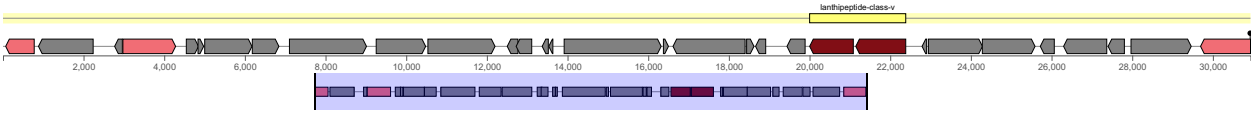
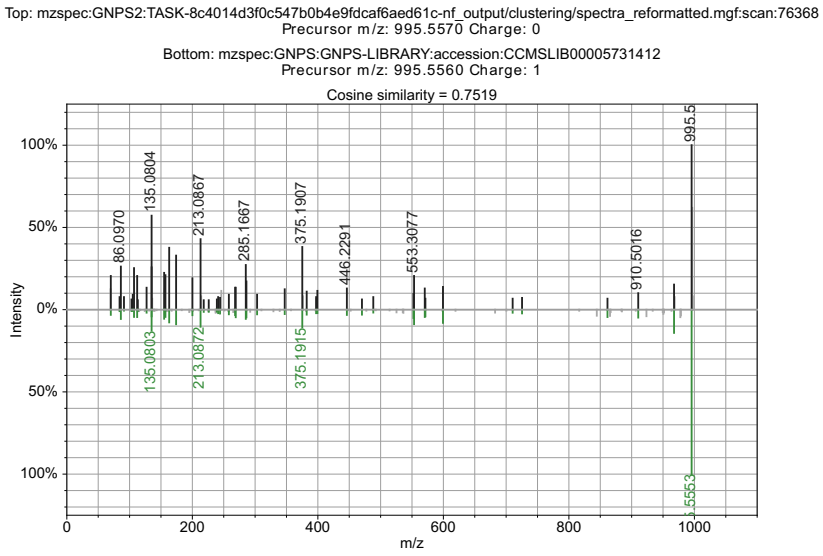


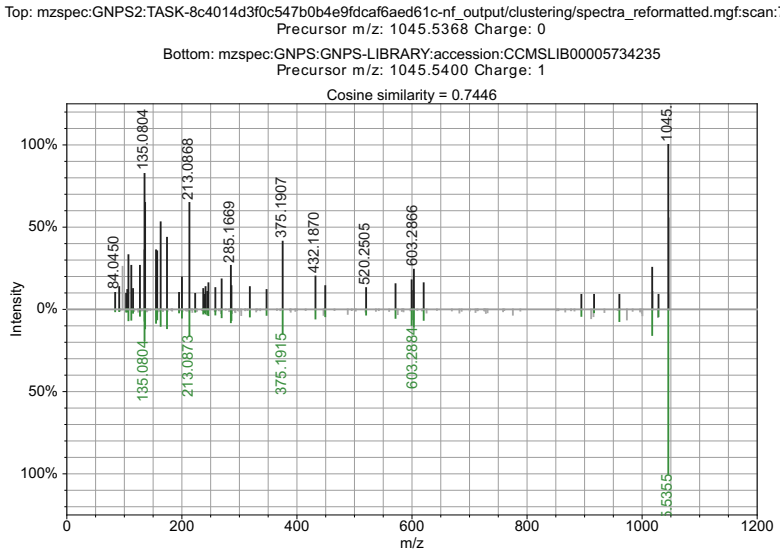
Figure S2.9.2: (a-jj) Schematics of biosynthetic gene clusters used in the study, generated via antiSMASH v7 (1).

A. Microcystins

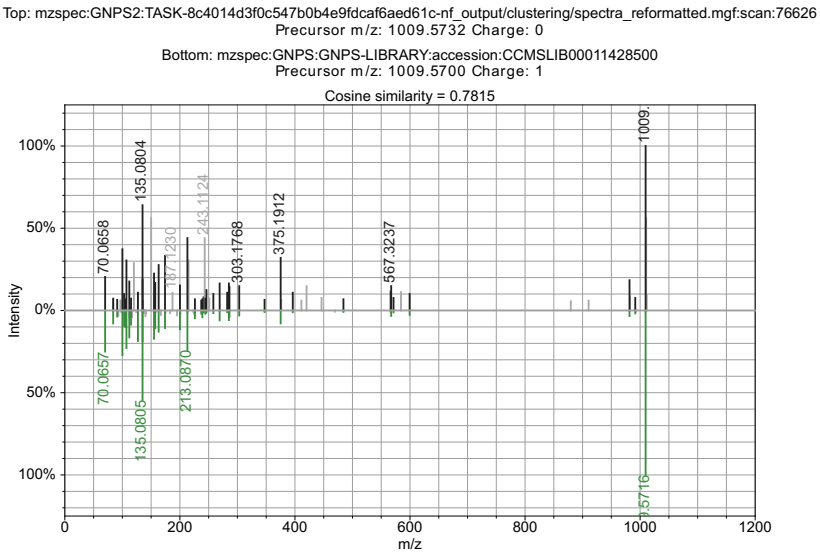
MC-LR



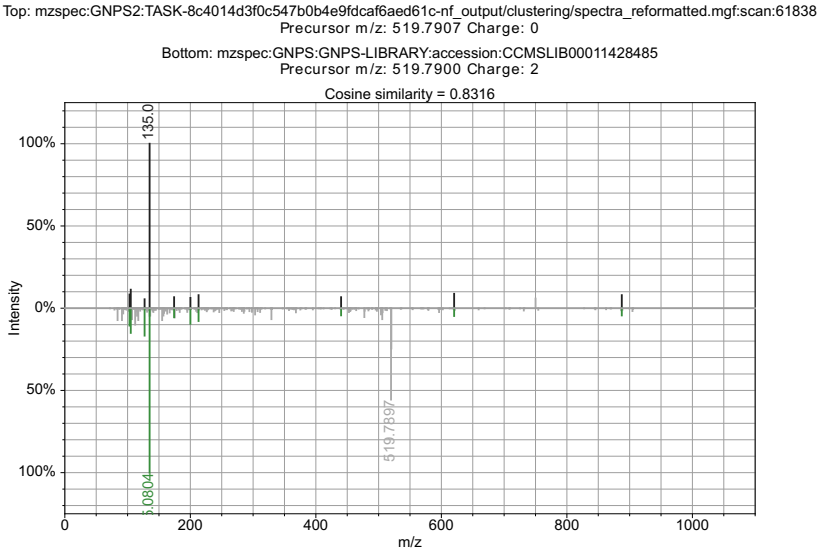
MC-YR



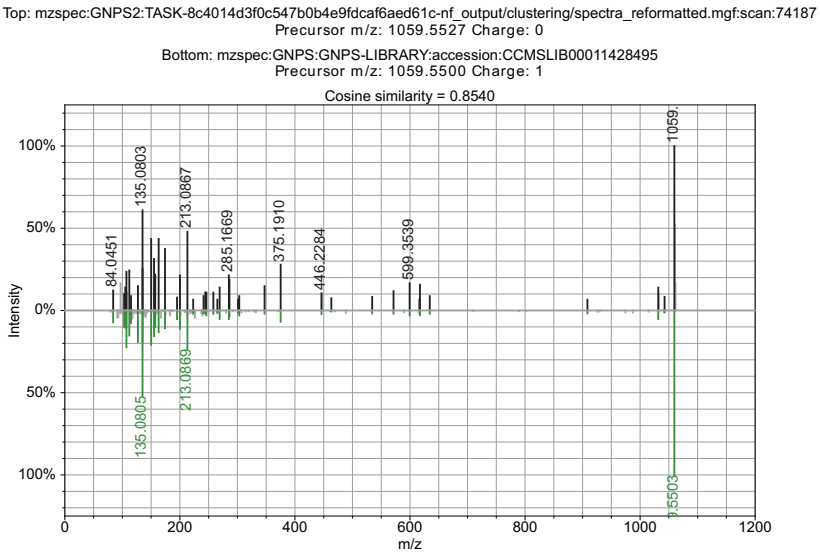
MC-HilR


Top: mzspec:GNPS2:TASK-8c4014d3f0c547b0b4e9fdcaf6aed61c-nf_output/clustering/spectra_reformatted.mgf:scan:76626
Precursor m/z: 1009.5732 Charge: 0
Bottom: mzspec:GNPS:GNPS-LIBRARY:accession:CCMSLIB00011428500
Precursor m/z: 1009.5700 Charge: 1
Cosine similarity = 0.7815
Intensity
m/z

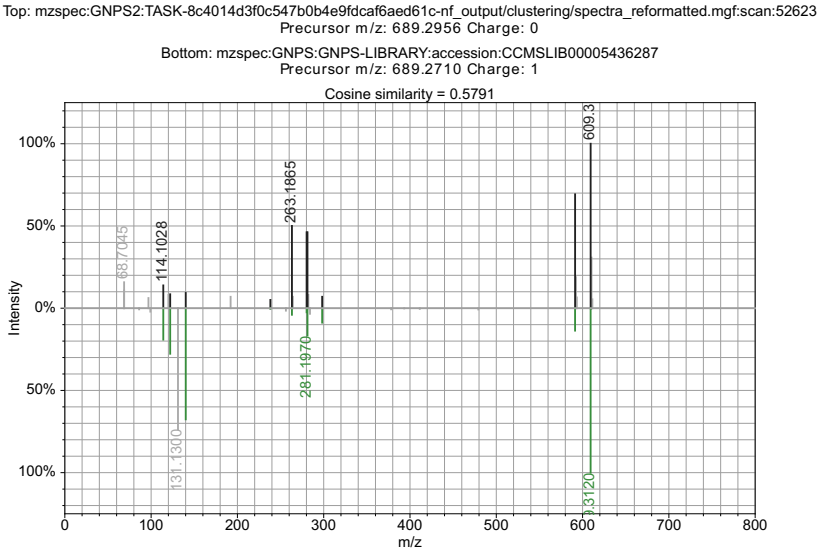
MC-RR


Top: mzspec:GNPS2:TASK-8c4014d3f0c547b0b4e9fdcaf6aed61c-nf_output/clustering/spectra_reformatted.mgf:scan:61838
Precursor m/z: 519.7907 Charge: 0
Bottom: mzspec:GNPS:GNPS-LIBRARY:accession:CCMSLIB00011428485
Precursor m/z: 519.7900 Charge: 2
Cosine similarity = 0.8316
Intensity
m/z

MC-HtyR

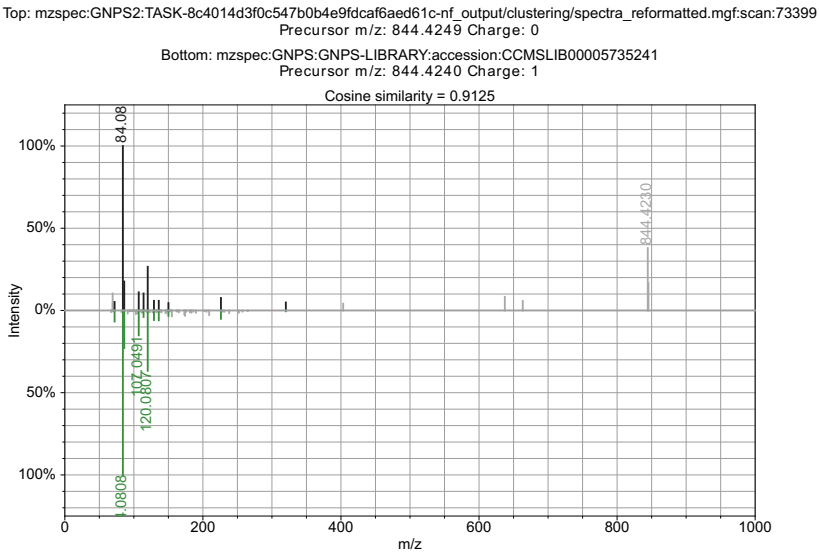


B. Aeruginosin 98A

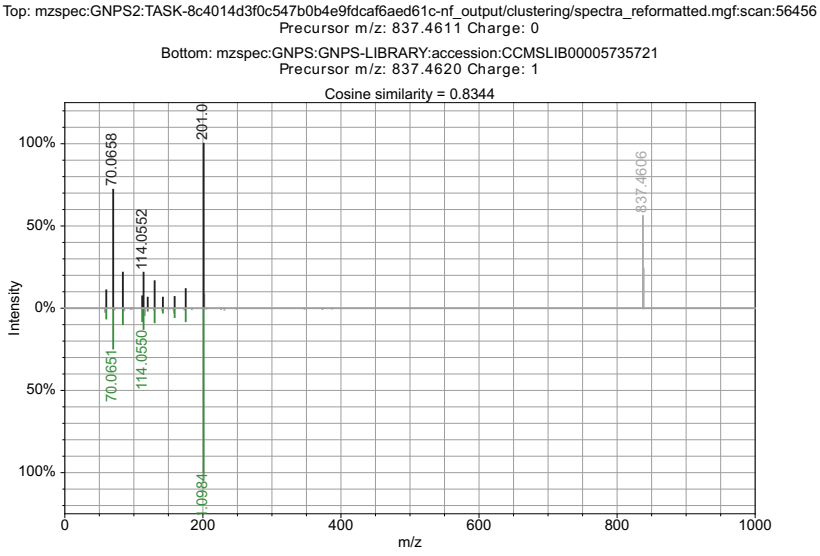


C. Anabaenopeptins

Anabaenopeptin A



Anabaenopeptin B



D. Aerucyclamide A

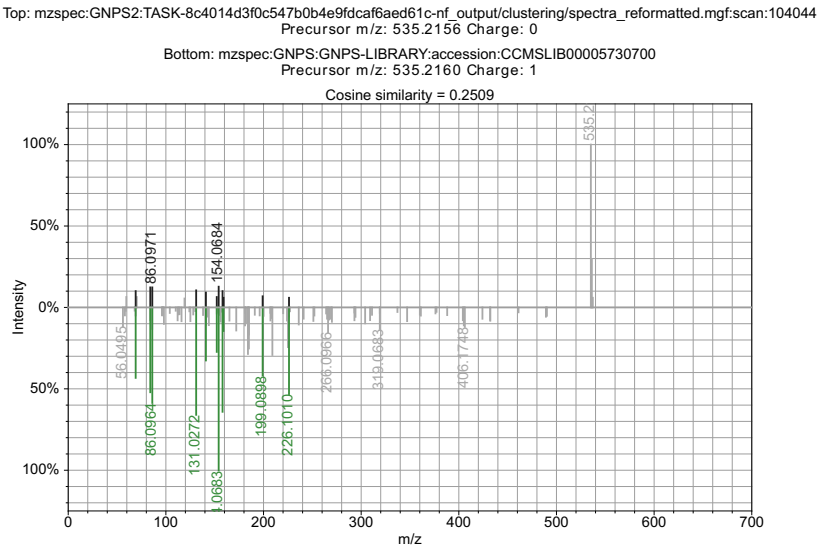


Figure S2.9.3: Mirror matches generated by GNPS of library matches detected and used in this study. The black spectrum on top represents the sample spectrum whereas the bottom green spectrum represents the database spectrum for the library match. All metabolite predictions shown here meet at least level 3 criteria as outlined in Schymasnki et al. 2014 (2). **(a)** Spectra putatively annotated as microcystins (MC) annotated as the congeners LR (cosine = 0.836), YR (cosine = 0.822), HilR (cosine = 0.619), RR (cosine = 0.756), HtyR (cosine = 0.736) are shown, **(b)** Aeruginosin 98A (cosine = 0.626), **(c)** Anabaenopeptins A (cosine = 0.855) and B (cosine = 0.912), and **(d)** Aerucyclamide A are shown (cosine = 0.2509).

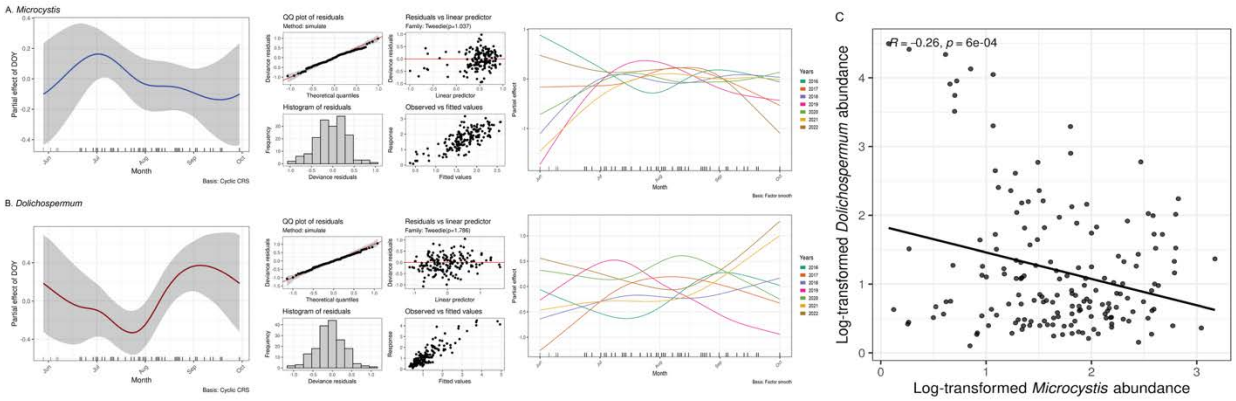


Figure S2.9.4: Fixed and partial effects of day of year (DOY), year, and model diagnostics for hierarchical generalized additive models (HGAMs) fit to log-transformed relative abundance of **(a)** *Microcystis* and **(b)** *Dolichospermum***.** The first panels display the fixed effect smooths for DOY based on the HGAM equation described in SI Table 2.9.6. Diagnostic plots (second panels) indicate that residuals are homoscedastic and approximately normally distributed (top right panel) and fitted values closely follow observed values (bottom right panel), supporting model adequacy for both taxa. The third panels show partial effect smooths of DOY stratified by year, with year-specific smooths illustrating interannual variation in seasonal patterns of each organism. Tick marks along the x-axis indicate the distribution of observed values. **(c)** A pairwise correlation plot of log-transformed *Microcystis* and *Dolichospermum* relative abundance reveals a modest but significant negative association (Pearson's r = –0.26, $p = 6x10^{-4}$), suggesting potential ecological interaction or shared environmental controls between the two taxa. Because these two response variables are derived from the same biological system, it is likely they are correlated in some manner. However, models were fit independently, and potential interdependence was not explicitly accounted for in the analysis

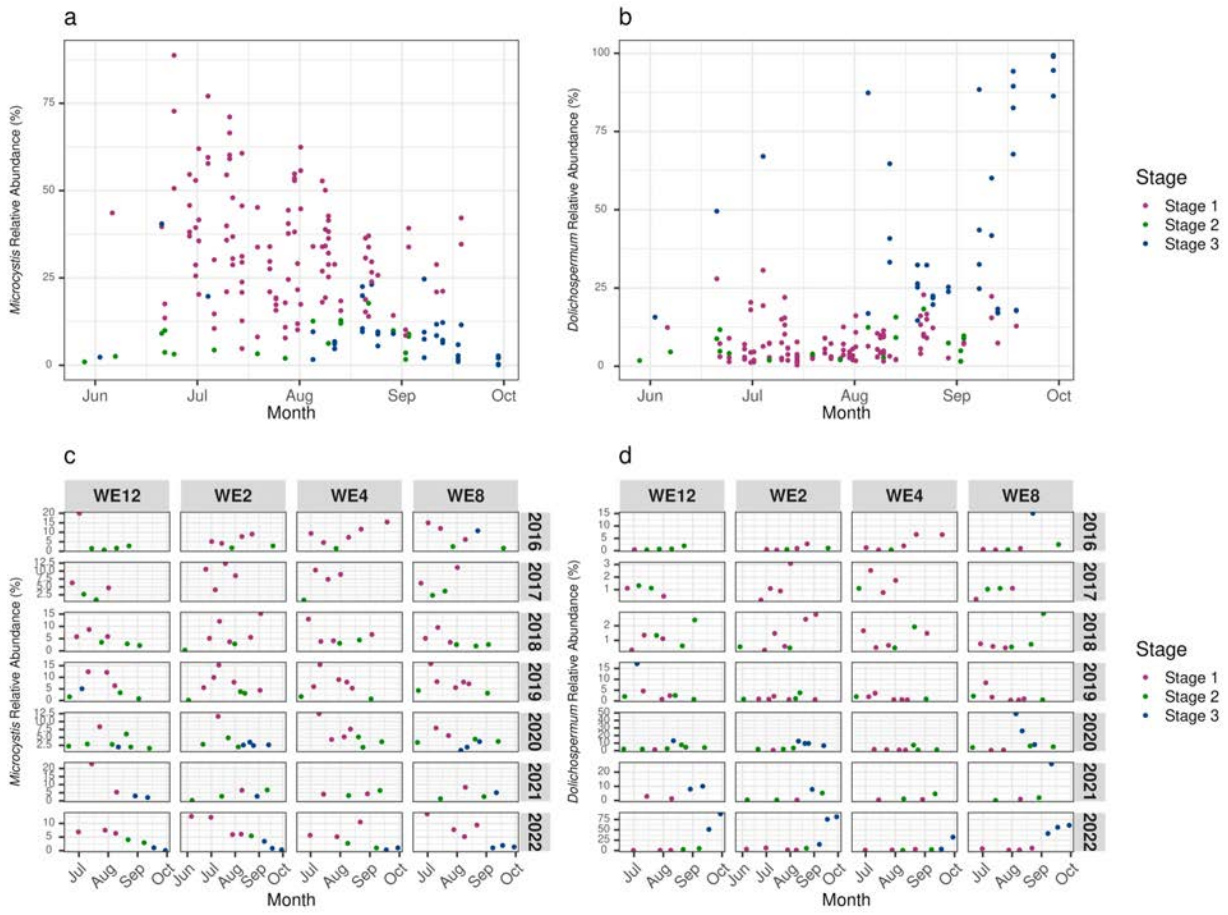


Figure S2.9.5: Seasonal pattern of relative abundance (%) of **(a)** *Microcystis* and **(b)** *Dolichospermum* across all stations and years. Points are colored by stage, where Stage 1 indicates *Microcystis* > *Dolichospermum* by 3% or more, Stage 2 indicates approximately equal populations of *Microcystis* and *Dolichospermum* (±3%), and Stage 3 indicates *Dolichospermum* > *Microcystis* by 3% or more. These stages were determined based on the distribution of their relative abundance relative to one another. Seasonal patterns faceted by station (WE12, WE2, WE4, and WE8) and year for **(c)** *Microcystis* and **(d)** *Dolichospermum* relative abundance (%) are also shown, with points colored by stage.

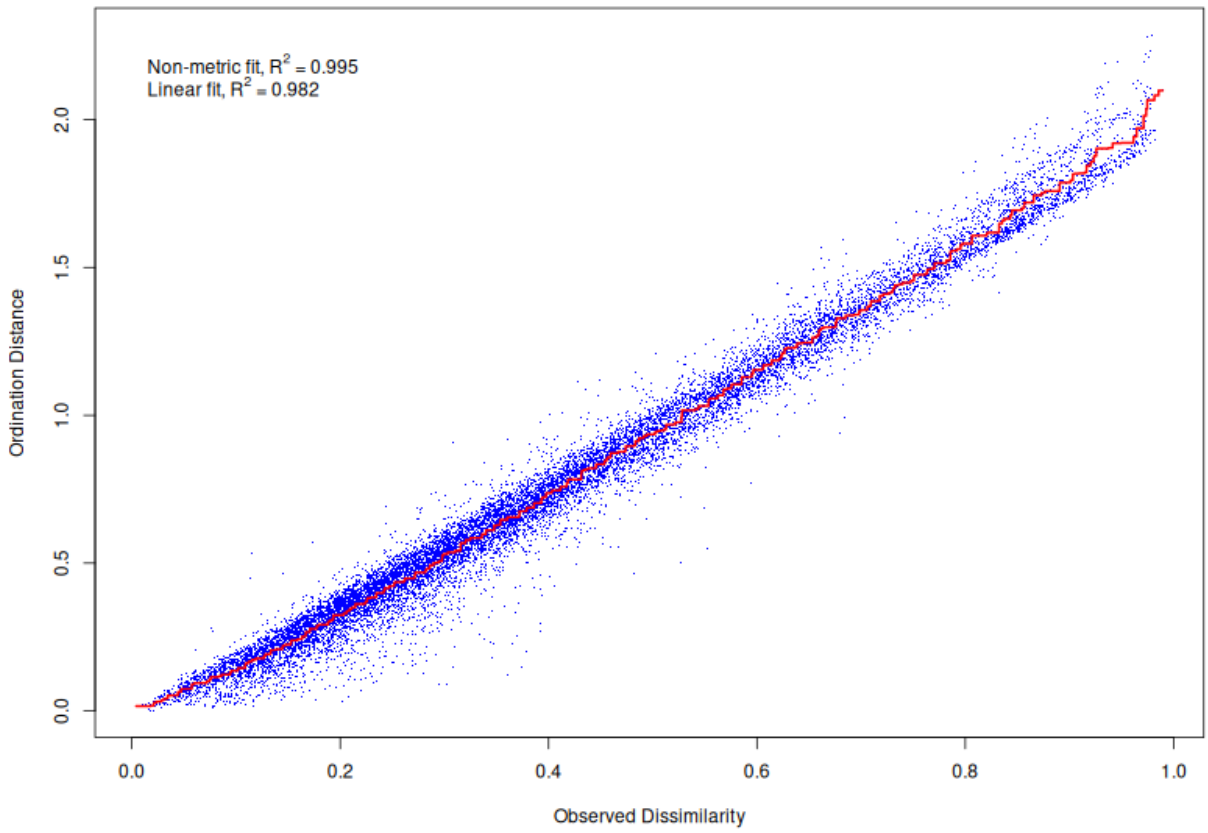


Figure S2.9.6: NMDS stress plot assessing the goodness-of-fit of a non-metric multidimensional scaling (NMDS) ordination based on physicochemical parameters. The plot compares observed dissimilarities (x-axis) to distances between points in the ordination space (y-axis). Blue points represent pairwise distances; the red line shows a monotonic fit between observed and ordination distances. The non-metric fit ($R^2 = 0.995$) and linear fit ($R^2 = 0.982$) indicate an excellent correspondence between the original distance matrix and the reduced-dimensional NMDS space.

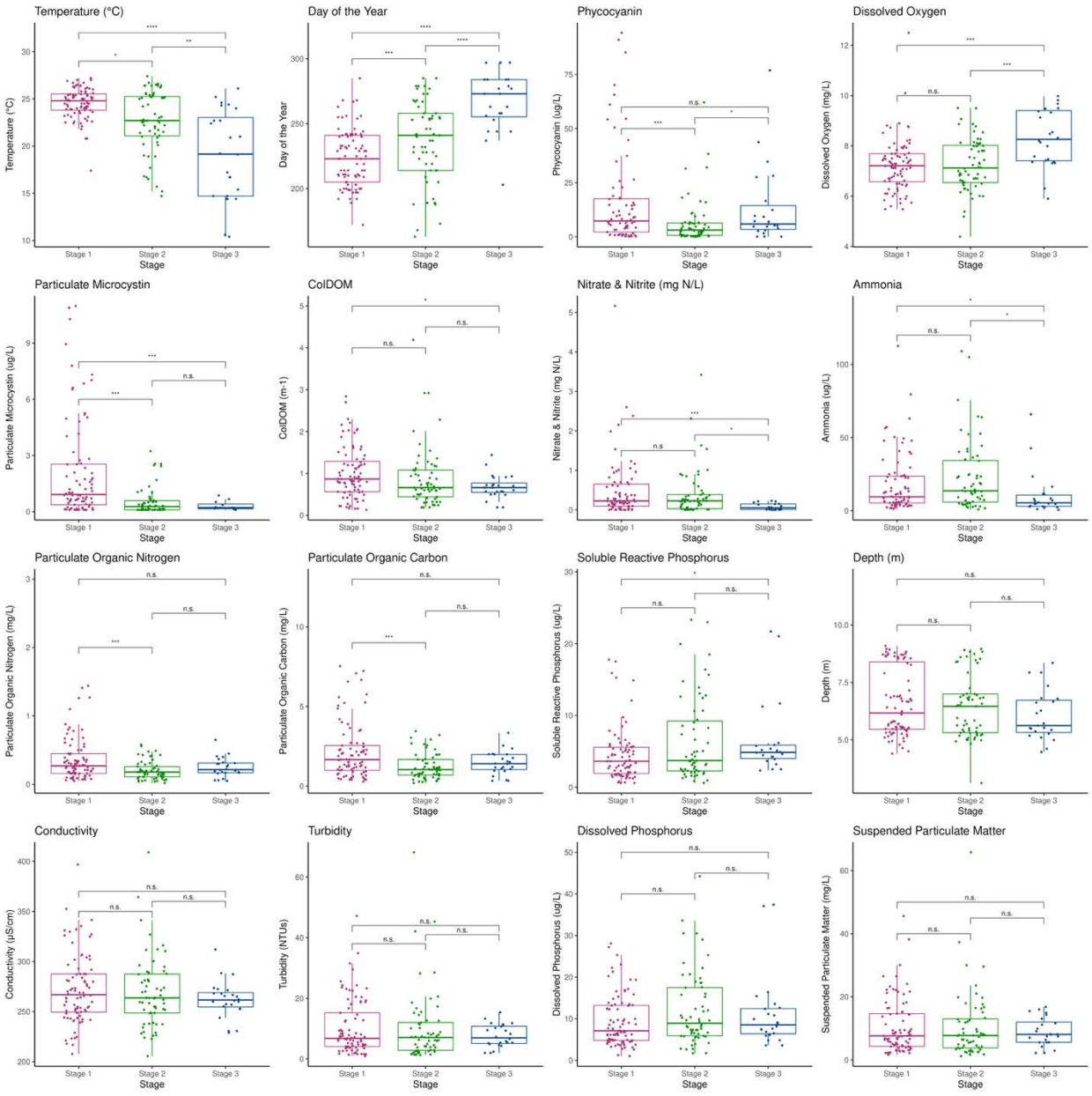


Figure 2.9.7: Boxplots showing the distribution of 18 physicochemical parameters across stages. Each point represents an individual sample, grouped by stage classification. Horizontal bars denote statistically significant pairwise differences between stages for each parameter, based on Kruskal–Wallis tests followed by Dunn's post hoc comparisons. Significance levels are indicated as: **** $P<0.0001$; *** $P<0.001$; ** $P<0.01$; * $P<0.5$; n.s. = not significant).

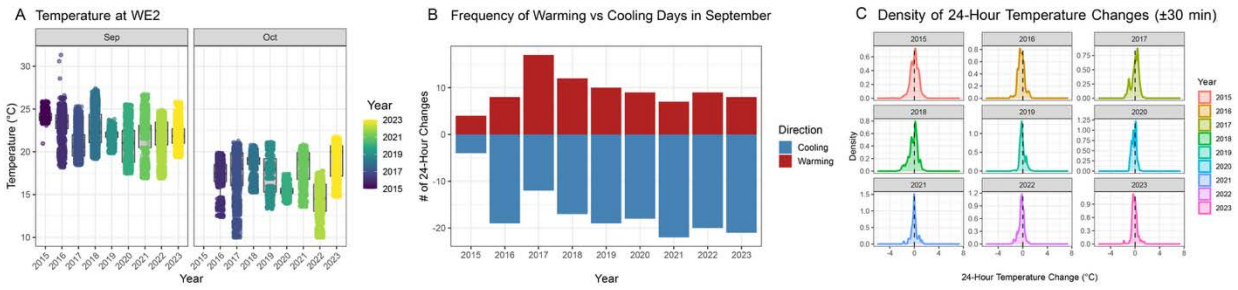


Figure S2.9.8: Patterns in surface water temperature from NOAA GLERL WE2 Nutrient Buoy from 2015 through 2023. Data was collected from Great Lakes Observing System (GLOS) archive (3). **(a)** Surface water temperatures (°C) in September and October colored by year. **(b)** Frequency of 24-hour temperature changes in September classified as warming (red) or cooling (blue) days. **(c)** Density distributions of 24-hour temperature changes (magnitude), with a 30-minute observation tolerance, for each year. Vertical dashed lines indicate a change of zero (no net warming or cooling).

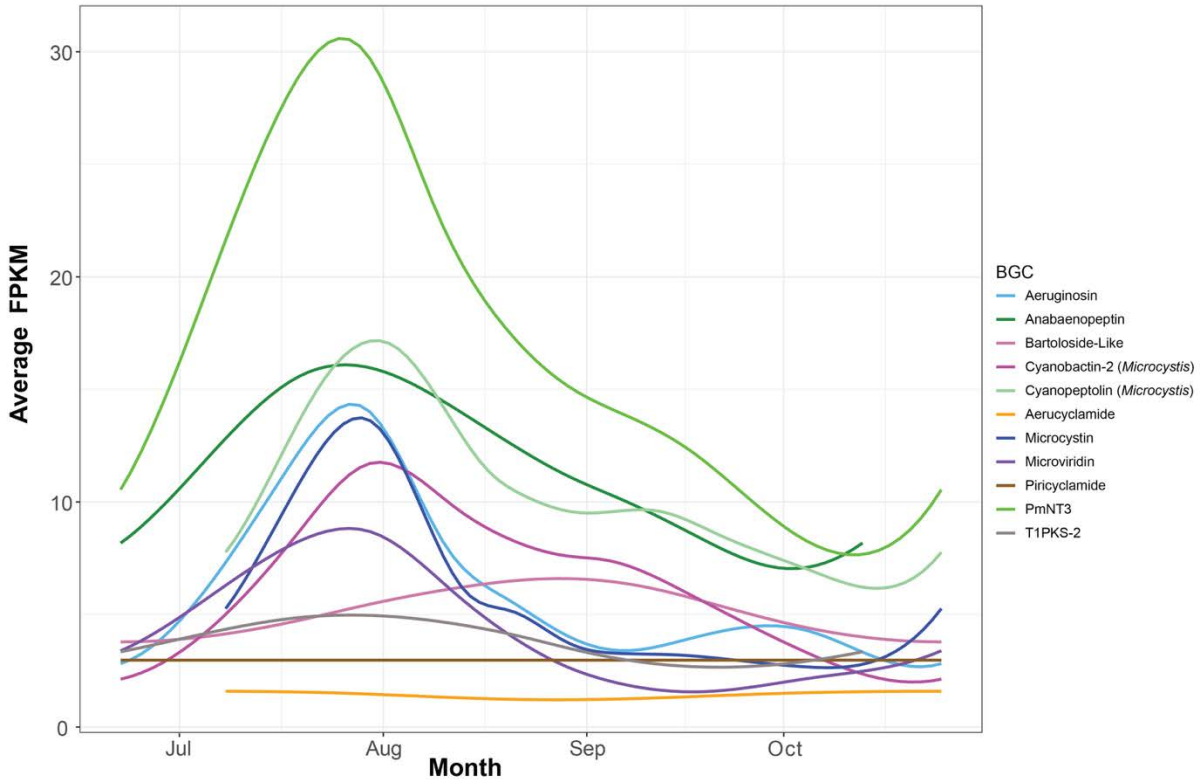


Figure S2.9.9: Temporal patterns of average FPKM of BGCs from *Microcystis*. Colored lines represent individual BGCs, smoothed using generalized additive models (GAMs) with a smooth on day of year with a cubic regression function.

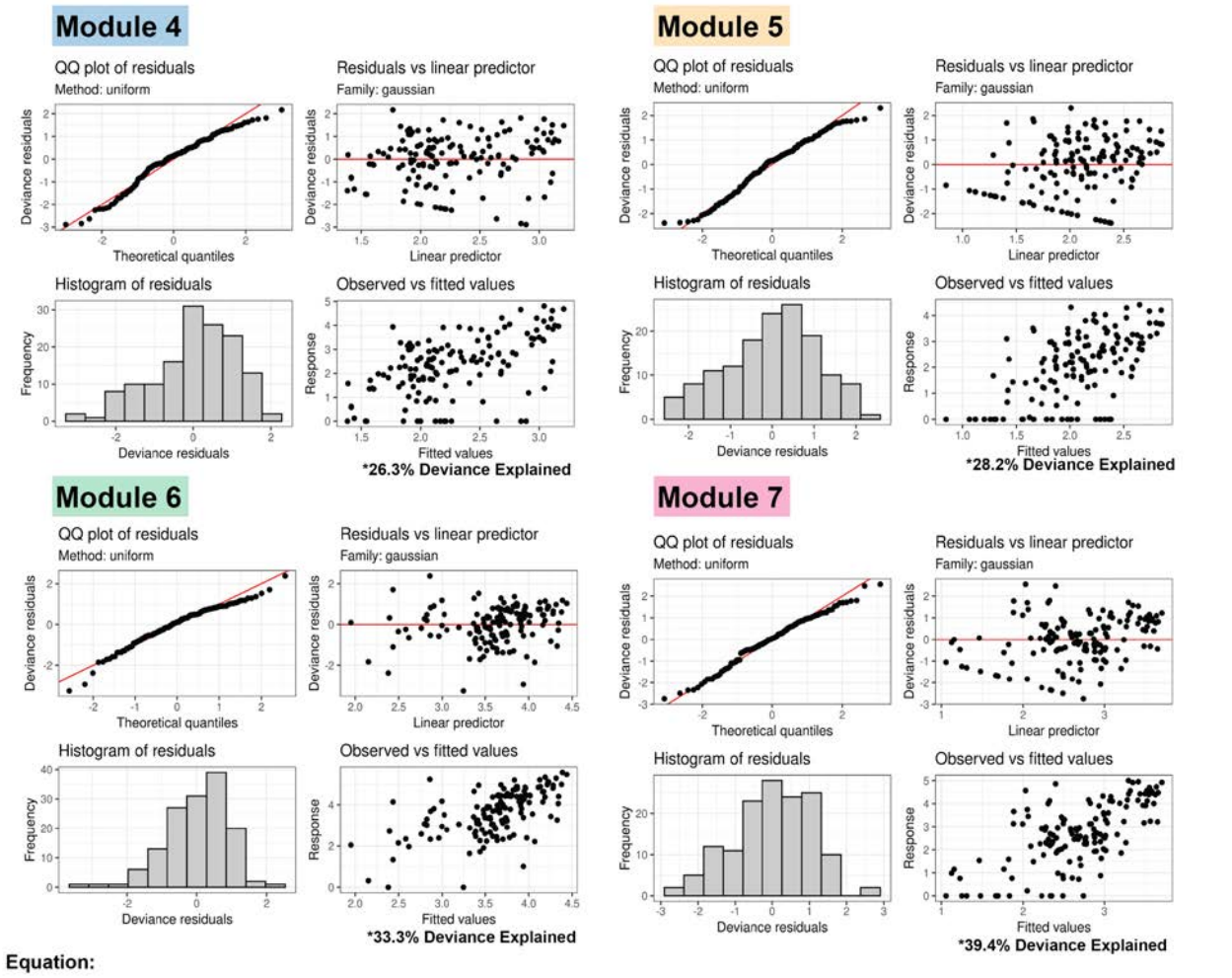


Figure S2.9.10: Residual diagnostic plots are based on models using log-transformed biosynthesis module abundance (modules 4, 5, 6, and 7) as the response variable within the equation at the bottom of the figure. Residuals appear randomly scattered around zero (homoscedastic) across predictors (top right panel), and fitted values closely follow observed values (bottom right panel), supporting model adequacy for the modules.

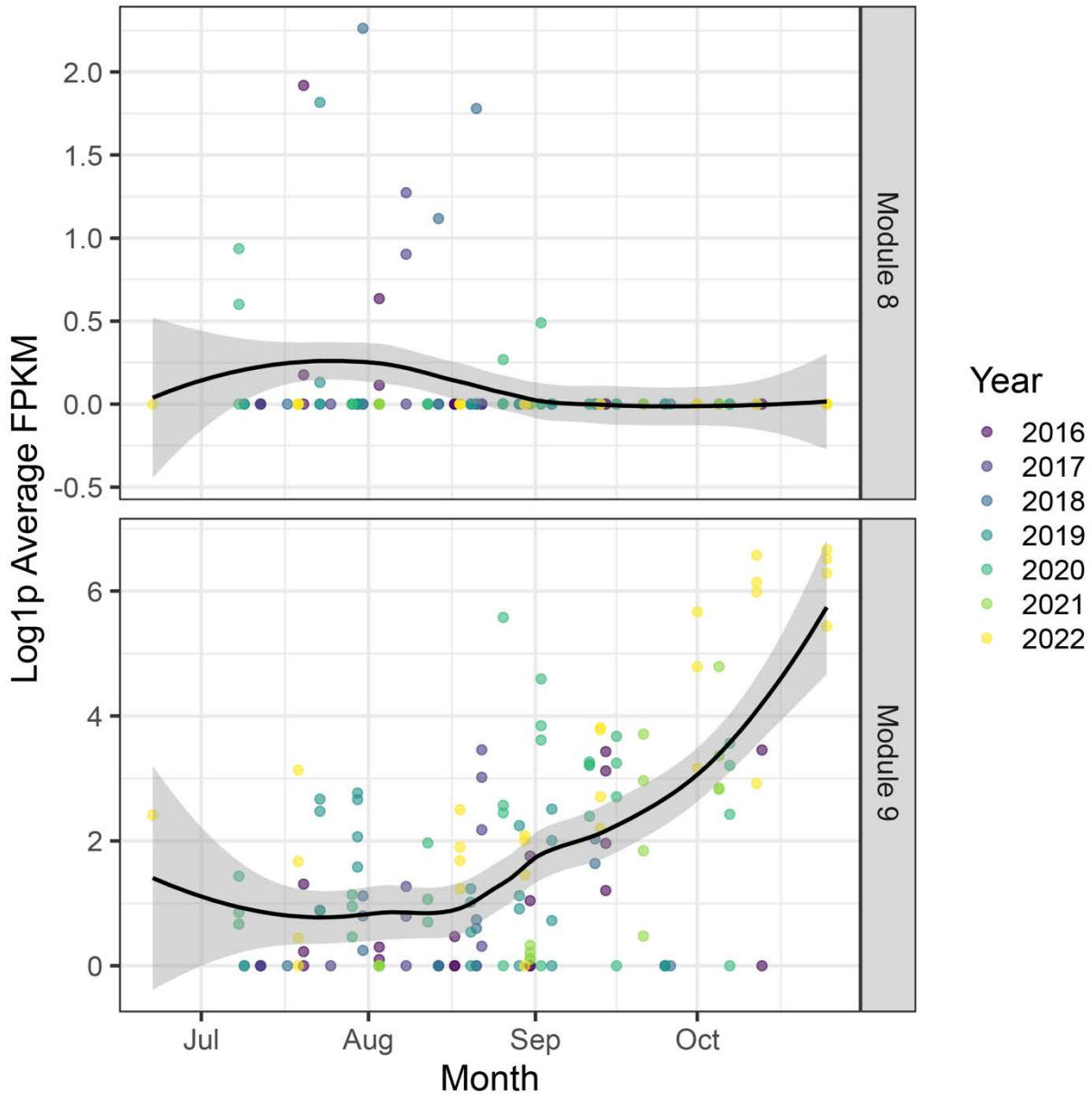


Figure 2.9.11: Log FPKM of BGCs from modules 8 and 9 belonging to the ADA clade. Individual observations are shown as points colored by year, and smoothed lines are fitted using a loess function. Shaded ribbons indicate standard errors around the fitted temporal patterns.

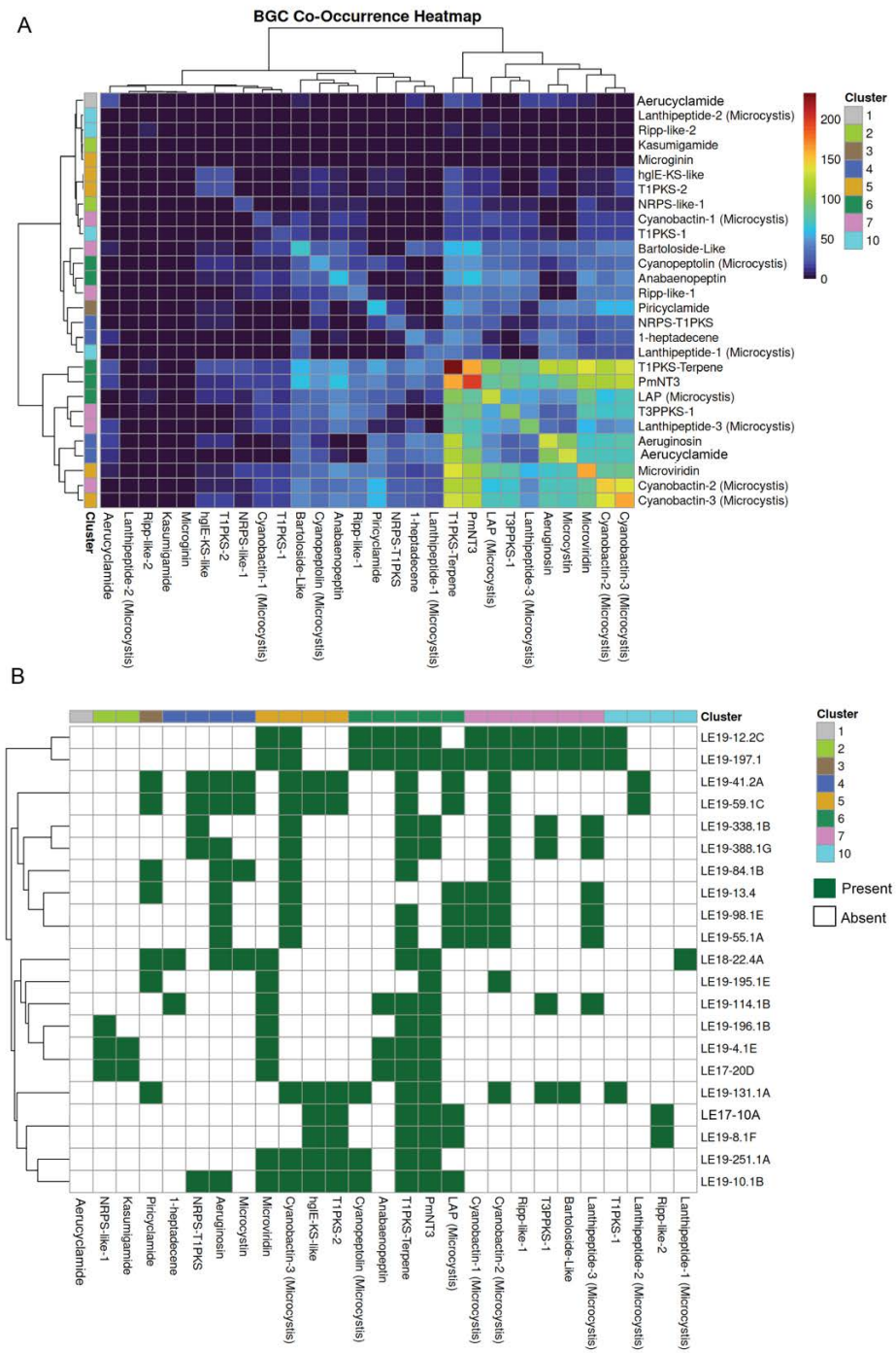


Figure 2.9.12: (a) Co-occurrence matrix of BGCs across publicly available genomes of *Microcystis* deposited in NCBI. Co-occurrence of BGCs across genomes was calculated and hierarchically clustered using average-linkage. The cluster each BGC belongs to is indicated as the colored bar on the right. Tiles are brighter in the heatmap indicating more co-occurrences between BGCs. **(b)** Presence/absence matrix of BGCs in WLECC isolates, organized by which cluster the BGC is in (colored bar). Green tile indicates the presence of the BGC in the isolate, and a white tile indicates the absence of the BGC in the isolate.

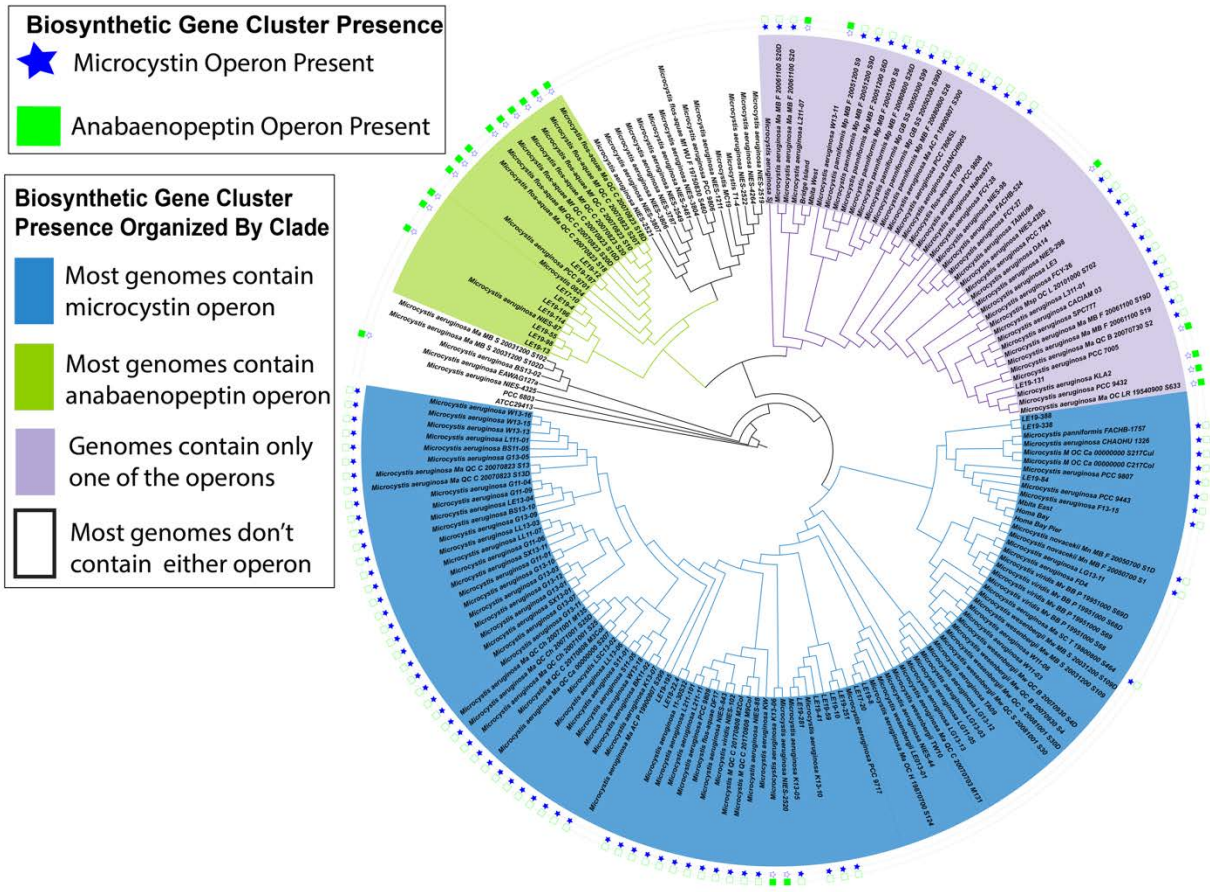


Figure S2.9.13: Phylogenomic analysis of publicly available *Microcystis* genomes with *Synechocystis* sp. PCC 6803 used as the outgroup. Genomes were used to build a maximum likelihood phylogenomic tree based on the prepackaged single-copy gene-set for cyanobacteria (251 target genes) in GToTree v1.7.05 (4). Lime green boxes indicate the detection of the anabaenopeptin BGC within a genome, and blue stars indicate the detection of the microcystin BGC within a genome. Clades shaded as blue indicate phylogenomically related genomes mostly containing the microcystin BGC, clades shaded as green indicate phylogenomically related genomes mostly containing the anabaenopeptin BGC, clades in purple indicate phylogenomically related genomes that have either the anabaenopeptin or microcystin BGCs, and clades in white contain phylogenomically related without either of the BGCs.

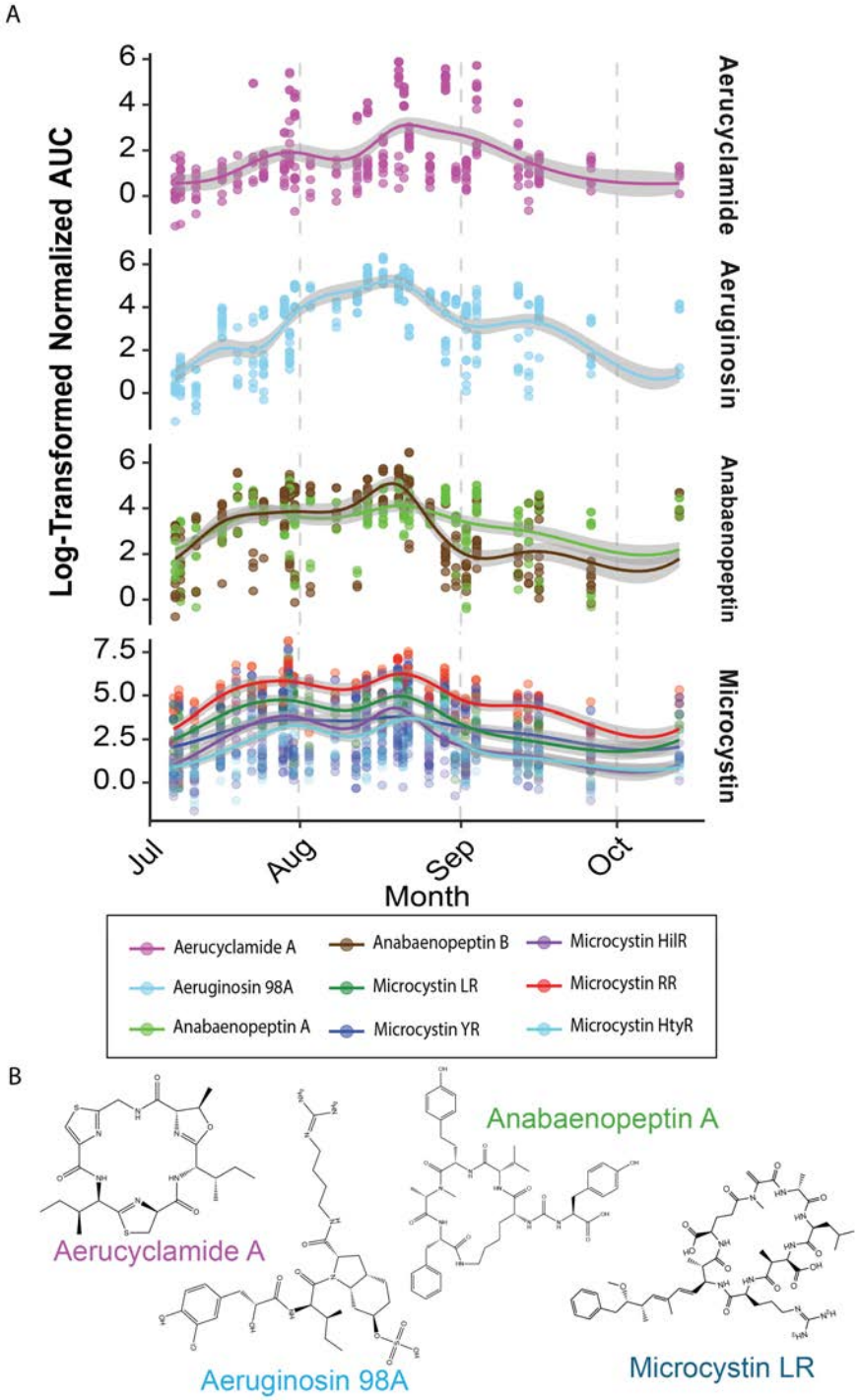


Figure S2.9.14: (a) Log-transformed normalized relative abundance (AUC) of metabolites putatively annotated as Aerucyclamide A, Aeruginosin 98A, Anabaenopeptin A and B, and Microcystin LR, YR, HilR, RR, and HtyR. All individual observations are shown, and smoothed lines for each metabolite were fitted using a loess function. **(b)** Representative metabolites from each metabolite class above.

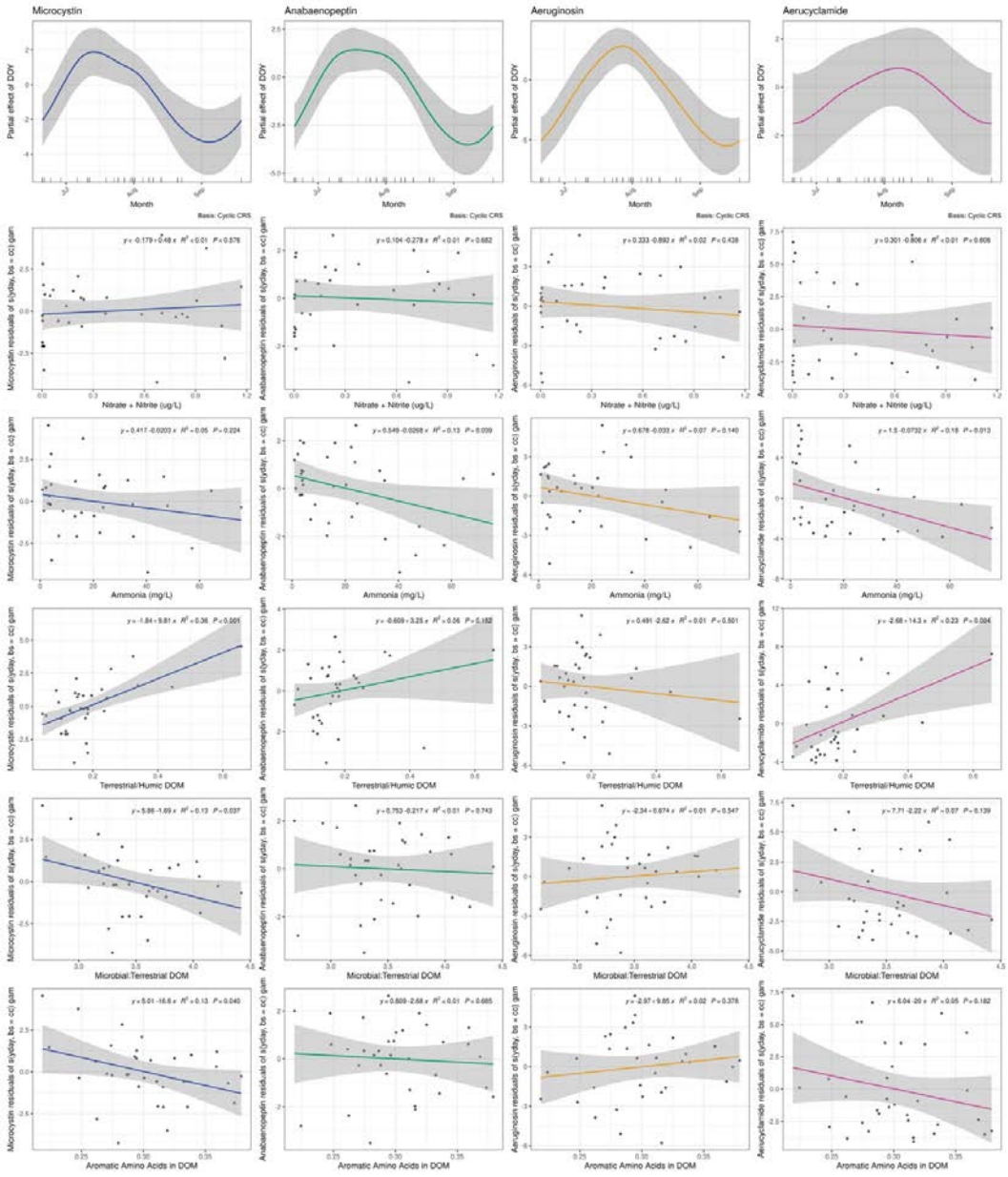


Figure S2.9.15: The top row displays partial effects of day of year on log-transformed metabolite abundance from GAMs using a cyclic cubic spline (±95% confidence interval), capturing seasonal patterns for each metabolite (Microcystins, Anabaenopeptins, Aeruginosins, and Aerucyclamides from left to right). The bottom five rows of panels display the relationship between each metabolites' GAM residuals and an abiotic variable. Residuals represent variability unexplained by seasonal trends, and linear models (±95% confidence intervals) are shown with R^2 and *P*-values from the fitted linear regression. Each row represents an abiotic variable (Nitrate + Nitrite, Ammonia, Terrestrial/Humic fraction of DOM, Microbial: Terrestrial Ratio of DOM, and Aromatic Amino Acid Portion of DOM from top to bottom).

2.9.3 SI Notes

SI Note 1: Metabolomics Data Collection

HPLC

Analytes were separated on a Thermo Scientific Vanquish Bioinert high-performance liquid chromatography (HPLC) system (ThermoFisher Scientific, San Jose, CA) with a Waters Atlantis T3 column (3um, 3.0x150mm) (Waters Corp., Milford, MA) at 30°C. A stacked injection method was used to add 2µL 5% EDTA, 2µL 5µg/L simetone (PubChem ID: 12654), and 50µL of sample. All reagents were LC/MS grade except for THF, which was >99.5% and stabilized with butylated hydroxytoluene (BHT). The mobile phases were: (A) water with 0.1% formic acid and 2mM ammonium formate and (B) 50:50 ACN:MeOH with 2mM ammonium formate, 0.1% THF and 0.1% formic acid. Samples were maintained at 4°C in the autosampler and the flow rate was 0.6 ml/min throughout the gradient in **SI Table 2**, with the first two minutes sent to waste.

Liquid Chromatography-High Resolution Mass Spectrometry (LC/HRMS)

Mass spectrometric analysis was carried out on a Thermo Fisher Scientific Q Exactive Orbitrap mass spectrometer equipped with a heated electrospray ionization (HESI-II) source in positive ion mode. The spray voltage, capillary temperature and auxiliary gas heater temperature were set to 3.5 kV, 275°C and 450°C, respectively. The sheath gas, auxiliary gas, sweep gas and S-lens RF level were set to 55, 15, 3 (arbitrary units) and 50.0 V, respectively. Nitrogen was used for spray stabilization, for collision-induced dissociation experiments in the higher energy collision dissociation (HCD) cell, and as the damping gas in the C-trap. The instrument was calibrated in positive mode every 7 days using the manufacturer's calibration solution. The Full MS/ddMS2

mode comprised a full MS scan followed by a data dependent MS2 scan with a fragmentation energy applied. The mass spectrometer acquired a full MS scan at a resolution of 70,000 (FWHM at 200 m/z). The automatic gain control (AGC) target was set at $3.0 \times 10e^6$ with a maximum injection time (IT) of 100 ms. The full MS scan ranges were set as 150–2000 m/z. For the dd-MS2 scans, the mass resolution was set at 35,000 FWHM, AGC target at $5 \times 10e^4$, maximum IT 50 ms, isolation window 5.0 m/z, intensity threshold $2 \times 10e^3$, exclude isotopes "on", and dynamic exclusion 2.0s. The fragmentation energy (NCE) was stepped with values of 15, 40, and 60. Instrument control and data processing were conducted with Xcalibur 4.2, Compound Discoverer 3.3, and Exactive Tune 2.11 software (Thermo Scientific).

The data generated from the quantitation runs was used to do a pre-emptive search of the CyanoMetDB using Compound Discoverer 3.3 (5). The compounds that were discovered from the database were added to the inclusion list. The inclusion list is a list of compounds that will be preferentially fragmented. A blank was also run. The m/z of analytes in the blank were placed in an exclusion list. These m/z will not be selected for fragmentation. If there was an analyte that was found in the exclusion list that may interfere with a desired analyte, either the compound was removed from the exclusion list or the time segment that it was part of the exclusion list was edited so that it would not interfere with the analytes of interest.

SI Note 2: Metagenomic processing information

Raw metagenomic reads were processed using fastp v0.23.2 (6). Clumpify was used to sort reads and improve compression efficiency (http://sourceforge.net/projects/bbmap/). BBmap was used to

remove human contaminant reads (GENCODE release 38). Metagenomic assemblies were produced using MEGAHIT v1.2.9 with the meta-sensitive preset (7). Multiple binning software were used to maximize metagenome assembled genome (MAG) recovery including CONCOCT v1.1.0, MetaDecoder v.1.0.13, VAMB v.3.08, MetaBAT2 v2.17, MaxBin 2.0 v2.2.7, and SemiBin v.1.03 (8–12). All MAGs were grouped and dereplicated using dRep v3.2.0 at 95 percent (13).

The relative abundance of WLECC *Microcystis* isolates (NCBI BioProject: PRJNA903891) was calculated as described in Yancey et al. 2023A (14). Briefly, sourmash gather v4.5.0 was used to compare k-mer sketches (k=31) of WLECC *Microcystis* MAGs to sampled metagenomes using the GTDB release 202 database and 159 *Microcystis* reference genomes (15). Relative abundance of isolates was calculated as described in Irber et al. 2022 (16).

SI Note 3: BGC Database Creation

To analyze the presence and abundance of BGCs across metagenomics samples in this study, a comprehensive, non-redundant, gene-based BGC database was created (Key to database in **SI Table 3**). All files and code to produce the final BGC database used in this study can be found on GitHub (https://github.com/Geo-omics/2016-2022-Western-Lake-Erie-Time-Series-Metagenomics-and-Metabolomics). First, all genes from BGCs belonging to the cyanobacterial genera *Planktothrix, Sphaerospermopsis, Dolichospermum, Anabena, Aphanizomenon, Microcystis, Cylindrospermopsis,* and *Cyanobium* were pulled into a fasta file from the MIBiG 3.0 database (17). All genes from BGCs identified from 21 isolates in the Western Lake Erie Culture Collection, described in Yancey and Hart et al. 2024 (18), were pulled into the same

fasta file, resulting in Seed 1 of the database. Seed 2 of the database was produced by removing any BGCs that putatively synthesize the same product and were derived from a genome from the same genera. This database was clustered using MMseqs2 at 95 percent minimum sequence id and 90 percent coverage to remove similar genes from the database (19). To ensure non-redundancy, different BGCs that shared genes 95 percent or more similar were marked to be removed from the database, with one of the BGCs being kept in the database. This non-redundant gene database resulted in seed 3. To ensure there was fair representation of geographically relevant BGCs from the *Anabaena, Aphanizomenon, Dolichospermum* (ADA) clade, BGCs were mined from nine highly curated ADA MAGs from WLE described in Den Uyl et al. 2024 (20). BGCs were detected using antiSMASH v7, run within the multismash tool (https://github.com/zreitz/multismash) (--minimal, --genefinding-tool prodigal, --cb-known, --cc-mibig, --tigrfam) (1,21). Identified BGCs smaller than 5,000 base pairs were excluded, and remaining BGCs were put into a fasta file and clustered at 95 percent identity and 90 percent coverage using MMseqs2 to reduce redundancy (19). Non-redundant ADA genes from the BGCs were merged with seed 3 into seed 4 of the database. Lastly, seed 4 was clustered using MMseqs2 at 95 percent identity and 90 percent coverage to ensure an entirely non-redundant database of genes within BGCs. Any BGCs putatively encoding the synthesis of terpenes were excluded in this study. This final database is seed 5 and was used for identifying BGCs across metagenomic samples in this study. Each BGC in seed 5 is visualized in **SI Figure 1**. A key to the genes within the seed 5 database are also provided in GitHub, where the type of gene (core, additional, transporter, regulatory, other) is annotated for each gene (as provided by antiSMASH v7) (1).

2.9.4 Supplemental Notes References

1. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research. 2023 Jul 5;51(W1):W46–50.

2. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. Environ Sci Technol. 2014 Feb 18;48(4):2097–8.

3. NOAA. NOAA WE2 Nutrient Buoy Data from 2015-2023 [Internet]. Great Lakes Environmental Research Laboratory; [cited 2025 Jul 1]. Available from: https://seagull-erddap.glos.org/erddap/tabledap/GLERLWE2_archive.html

4. Lee MD. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics. 2019 Oct 15;35(20):4162–4.

5. Jones MR, Pinto E, Torres MA, Dörr F, Mazur-Marzec H, Szubert K, et al. CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. Water Res. 20210308th ed. 2021 May 15;196:117017.

6. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884–90.

7. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 20150120th ed. 2015 May 15;31(10):1674–6.

8. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. Nat Methods. 20140914th ed. 2014 Nov;11(11):1144–6.

9. Liu CC, Dong SS, Chen JB, Wang C, Ning P, Guo Y, et al. MetaDecoder: a novel method for clustering metagenomic contigs. Microbiome. 2022 Mar 10;10(1):46.

10. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019 Jul 26;7:e7359.

11. Wu YW, Tang YH, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome. 2014 Dec;2(1):26.

12. Pan S, Zhu C, Zhao XM, Coelho LP. A deep siamese neural network improves metagenome-assembled genomes in microbiome datasets across different environments. Nat Commun. 2022 Apr 28;13(1):2326.

13. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. The ISME Journal. 2017 Dec 1;11(12):2864–8.

14. Yancey CE, Kiledal EA, Chaganti SR, Denef VJ, Errera RM, Evans JT, et al. The Western Lake Erie culture collection: A promising resource for evaluating the physiological and genetic diversity of *Microcystis* and its associated microbiome. Harmful Algae. 2023 Jul;126:102440.

15. Brown CT, Irber L. sourmash: a library for MinHash sketching of DNA. Journal of Open Source Software. 2016 Sep 14;1(5):27.

16. Irber L, Brooks PT, Reiter T, Pierce-Ward NT, Hera MR, Koslicki D, et al. Lightweight compositional analysis of metagenomes with FracMinHash and minimum metagenome covers. bioRxiv; 2022. p. 2022.01.11.475838.

17. Terlouw BR, Blin K, Navarro-Muñoz JC, Avalon NE, Chevrette MG, Egbert S, et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. Nucleic Acids Research. 2023 Jan 6;51(D1):D603–10.

18. Yancey CE, Hart L, Hefferan S, Mohamed OG, Newmister SA, Tripathi A, et al. Metabologenomics reveals strain-level genetic and chemical diversity of *Microcystis* secondary metabolism. mSystems. 2024 Jun 25;0(0):e00334-24.

19. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol. 2017 Nov;35(11):1026–8.

20. Den Uyl P. Genetic characterization of saxitoxin-producing cyanobacteria associated with western Lake Erie harmful algal blooms [Internet]. University of Michigan; 2024. Available from: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/193158/MS_Thesis_DenUyl_DeepBlue_FinalSignatures.pdf?sequence=1&isAllowed=y

21. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010 Mar 8;11:119.

Chapter 3 Metagenomics Reveals Spatial Variation in Cyanobacterial Composition, Function, and Biosynthetic Potential in the Winam Gulf, Lake Victoria, Kenya

3.1 Abstract

The Winam Gulf in the Kenyan region of Lake Victoria experiences prolific, year-round cyanobacterial harmful algal blooms (cyanoHABs) which pose threats to human, livestock and ecosystem health. To our knowledge, there is limited molecular research on the gulf's cyanoHABs, thus the strategies employed for survival and proliferation by toxigenic cyanobacteria in this region remain largely unexplored. Here, we used metagenomics to analyze the Winam Gulf's cyanobacterial composition, function, and biosynthetic potential. *Dolichospermum* was the dominant bloom-forming cyanobacterium, co-occurring with *Microcystis* at most sites. *Microcystis* and *Planktothrix* were more abundant in shallow and turbid sites. Metagenome assembled genomes (MAGs) of *Dolichospermum* harbored nitrogen fixation genes, suggesting diazotrophy as a potential mechanism supporting the proliferation of *Dolichospermum* in the nitrogen limited gulf. Over 300 biosynthetic gene clusters (BGCs) putatively encoding the synthesis of toxins and other secondary metabolites were identified across the gulf, even at sites where there were no visible cyanoHAB events. Almost all BGCs identified had no known synthesis product, indicating a diverse and novel biosynthetic repertoire capable of synthesizing harmful or potentially therapeutic metabolites. *Microcystis* MAGs

Materials from this chapter have been published in the following form:
Lauren N. Hart, Brittany N. Zepernick, Kaela E. Natwora, Katelyn M. Brown, Julia Akinyi Obuya, Davide Lomeo, Malcolm A. Barnard, Eric O. Okech, E. Anders Kiledal, Paul A. Den Uyl, 2022-2023 NSF IRES Research Consortium, Mark Olokotum, Steven W. Wilhelm, R. Michael McKay, Ken G. Drouillard, David H. Sherman, Lewis Sitoki, James Achiya, Albert Getabu, Kefa M. Otiso, George S. Bullerjahn, and Gregory J. Dick. (2025). Metagenomics Reveals Spatial Variation in Cyanobacterial Composition, Function, and Biosynthetic Potential in the Winam Gulf, Lake Victoria, Kenya. *American Society for Microbiology Applied and Environmental Microbiology (ASM AEM)*. https://journals.asm.org/doi/10.1128/aem.01507-24

contained *mcy* genes encoding the synthesis of hepatotoxic microcystins which are a concern for drinking water safety. These findings illustrate the spatial variation of bloom-forming cyanobacteria in the Winam Gulf and their available strategies to dominate different ecological niches. This study underscores the need for further use of genomic techniques to elucidate the dynamics and mitigate the potential harmful effects of cyanoHABs and their associated toxins on human, environmental, and economic health.

3.2 Introduction

Cyanobacterial harmful algal blooms (cyanoHABs) are a severe, global threat to freshwater systems, and are expected to intensify with climate change (1–3). The overgrowth of cyanobacteria and production of cyanotoxins threatens human and environmental health, freshwater supply, livestock, recreation, tourism, and fishing (4–6). Two of the most ubiquitous freshwater, bloom-forming cyanobacterial genera are *Microcystis,* a non-diazotroph notorious for its ability to produce toxic microcystins, and *Dolichospermum,* a filamentous diazotroph (7–10). These two genera can produce a variety of hepatotoxins, neurotoxins, and other secondary metabolites, including microcystins, saxitoxins, anatoxins, anabaenopeptins, aeruginosins, microginins, microviridins, and cyanopeptolins (10, 11).

Considerable attention has been directed towards understanding and monitoring *Microcystis*- and *Dolichospermum*-dominated cyanoHABs in North America and Europe (12–15). Yet, cyanoHAB characteristics, occurrences, and consequences in the global south remain less studied (16). Notably, Lake Victoria, the world's second largest lake by surface area, is experiencing rapid eutrophication and proliferation of toxic, *Microcystis*-dominated cyanoHABs (17, 18). CyanoHABs are prevalent throughout the Lake Victoria basin, which has a coastline

shared by Kenya, Uganda, and Tanzania. Moreover, most cyanoHABs are concentrated within embayments such as Winam Gulf (*i.e.,* Nyanza Gulf) (17, 18).

The Winam Gulf is a shallow inlet (average depth 7 m) near the northeastern corner of Lake Victoria that experiences year-round cyanoHABs (17). The gulf provides water for drinking, laundry, bathing, nearshore subsistence and commercial fishing, and aquaculture to Kisumu, Kenya's third largest city, and many nearby fishing villages such as Homa Bay and Kendu Bay (directly supporting ~4.5 million people) (19, 20). A 2015 study found that 50% of households sampled near Kisumu reported sole use of raw lake water for drinking and household use, with greater than 30 percent of water samples from the area exceeding World Health Organization (WHO) drinking water guidelines for microcystins (21). Sixty percent of samples exceeded microcystin thresholds for children and immunocompromised individuals suggested by the United States Environmental Protection Agency (USEPA), which is of particular concern due to the high prevalence of immunomodulating diseases in Kisumu like HIV and malaria (21–24). Thus, resource-dependent populations could face many routes and multifaceted effects of cyanotoxin exposure.

Since the 1980's, the Winam Gulf has suffered profound ecological changes due to invasive water hyacinth and other species, oxygen depletion, and increasing phytoplankton biomass (20). These changes are accompanied by substantial human population growth and increasing agricultural, industrial, urban, and wastewater contaminant runoff into the gulf (25). The main rivers flowing into the gulf are the Nyando and Sondu-Miriu Rivers which enter from the southeast corner, delivering loads of silt, nitrogen (N), and phosphorus (P) from agricultural, industrial, and wastewater sources. Smaller rivers such as the Awach-Tende, Awach-Kibuon, Kibos, and Kisat rivers deliver sewage and urban runoff from Kisumu (**Figure 3.9.1**). Sugar cane

processing, tea and cereal farming, and urban areas within the Winam Gulf catchment are the main sources of pollution, with Kisumu as a significant contributor (26, 27). Total nitrogen (TN) is higher in the gulf than the open lake, while total phosphorus (TP) is lower in the gulf than the open lake. Despite abundant sources of nitrogen delivery to the gulf, TN:TP ratios remain lower than 20 - indicative of nitrogen limitation relative to phosphorus, leading to potential selection for diazotrophic organisms (26). Despite this, non-diazotrophic cyanobacteria are commonly found in the gulf (i.e., *Microcystis*). This indicates that there is an abundance of available nutrients in the gulf, allowing non-diazotrophs to proliferate even under nitrogen-limited conditions by employing alternative strategies (28).

The climate of the Winam Gulf is characterized by two wet seasons, ranging from March to May and October to December. Two studies, from 2008-2009 and 2020, found maximum phytoplankton biovolumes coincide with wet (March, April, & May) and seasonal upwelling (July & August) seasons in the gulf (17, 29). Based on morphological classifications of phytoplankton, the 2008-2009 study found *Microcystis* dominates cyanoHABs in the gulf, with increased abundance of the diazotrophic cyanobacterial genus *Anabaena*, now renamed *Dolichospermum,* in September when nitrogen became more limited (30). Microcystin concentrations often exceeded WHO guidelines and increased in tandem with *Microcystis*, suggesting *Microcystis* to be the dominant microcystin producer in the gulf (17). The 2020 study identified *Microcystis* and *Dolichospermum* as the two main bloom-forming cyanobacteria at high abundance from May through July (29). Another study in 2015 and 2016 found similar phytoplankton patterns and identified *Microcystis* and *Dolichospermum* in community waters and near the drinking water intake in Kisumu (21).

Most molecular studies of the Winam Gulf have focused on fish pathogens such as *Escherichia coli, Vibrio cholerae,* and *Salmonella* spp. (31–33). A study by Brown et al. (2024) in 2022 used 16S rRNA gene sequencing to characterize the microbial community in the gulf, identifying the diazotrophic genus *Dolichospermum* as the dominant cyanobacterial organism (28). This study also found high copy numbers across the gulf of *cyrA* and *mcyE*, conserved genes in the biosynthetic gene clusters (BGCs) for cylindrospermopsin and microcystin, respectively (28). Despite these investigations, the producer(s) of these cyanotoxins in the gulf remains unknown. Additionally, the functional and biosynthetic capabilities of bloom-forming cyanobacteria that contribute to their fitness across different regions of the gulf have not been studied.

The aim of this study was to genetically characterize the functional and biosynthetic potential of bloom-forming cyanobacteria in the Winam Gulf using whole genome shotgun metagenomic sequencing. We conducted a two-year sampling expedition (2022-2023) from coastal, offshore, and river sites within and outside of the gulf. We found cyanoHAB events were not limited to coastal zones, or even within the gulf, and were most prolific near Homa Bay in 2022 and outside of the gulf near Bondo and the Yala River mouth in 2023. These cyanoHAB events were dominated by *Dolichospermum,* with *Microcystis* co-occurring but at lower abundance. Highly diverse and abundant BGCs were encoded by many genera of cyanobacteria, with the most biosynthetic potential encoded at sites with the highest abundance of bloom-forming cyanobacteria. Most BGCs could not be linked to a known biosynthesis product, highlighting the novel chemical repertoire encoded by cyanobacteria in the gulf. Lastly, our data suggests *Microcystis* is the main and sole producer of microcystin in the gulf during the sampling period.

3.3 Results

3.3.1 Water Physicochemistry Indicates Nitrogen Limitations Throughout the Gulf

Thirty-one sites throughout the gulf and nearby rivers were sampled for metagenomic analysis in May and June of 2022 and 2023 (**Figure 3.9.1, Table 3.9.1**). The sites sampled span spatial, temporal, and physical and chemical (physicochemical) gradients. Physicochemical analyses were performed at each site to infer water quality (**Figure 3.9.2**), excluding river sites. Turbidity varied across sites ranging between 10 and 48.1 nephelometric turbidity units (NTUs), with a median of 23.7 NTUs. Dissolved oxygen measured between 4.26 and 11.6 mg/L across sites, with sites within the Rusinga Channel (26, 27, and 31), on the higher end of the range measured at 9.9, 10.98, and 11.6 mg/L, respectively. Total nitrogen (TN) and total phosphorus (TP) were measured in 2023 and ratios (TN:TP) ranged between 5.1 and 25 with a median of 9.06. All sites except site 19 fell below a TN:TP ratio of 20, suggestive of nitrogen limitation (34) (**Figure 3.9.2**). Site 19 on the southern coast of the Gulf had a TN:TP ratio of 25, falling in the ambiguous nitrogen or phosphorus limited range.

3.3.2 River and Gulf Sites Vary in Microbial Community Makeup

Overall, the microbial community composition of the gulf closely resembled other eutrophic lakes with the phyla *Actinobacteria*, *Cyanobacteria*, *Proteobacteria*, and *Planctomycetota* being highly abundant (35) (**Figure 3.9.3**). The communities in 2022 and 2023 were similar in their compositions and dominant microbial phyla. The community composition at river sites (sites 12, 13, 14, and 16) diverged from that of gulf sites primarily because of lower *Cyanobacteria* ($P = 0.1$) and significantly higher *Bacteroidota* ($P < 0.001$) abundance observed in the samples. Cyanobacteria represented less than 1 percent of the total microbial community

in each river site. Cyanobacteria were present at all sites, with relative abundance ranging between approximately 1 percent at site 31 and 78 percent at site 28.

3.3.3 Dolichospermum Dominates Large CyanoHAB Events Detected in Distinct Gulf Regions

Dolichospermum, *Cyanobium, Vulcanococcus,* and *Microcystis* were the major cyanobacterial genera found in the gulf and rivers (**Figure 3.9.3)**. Minor cyanobacterial genera present across the sites were *Snowella, Synechococcus, Elainella, Pseudanabaena, Cuspidothrix, Nodosilinea*, and *Caenarcaniphilales*. *Sphaerospermopsis* and *Planktothrix* had high relative abundance at site 27 and 10, respectively, making up more than 36% and 65% of the cyanobacterial community at these two sites, respectively. S*phaerospermopsis* and *Planktothrix* were found at four other sites at an abundance below 5% and 20% of the total cyanobacterial community, respectively. Cyanobacteria unclassified at the genus level were a substantial portion of the community (1-16%) at most sites.

River sites had similar cyanobacterial compositions despite their lower cyanobacterial abundance. *Cyanobium* and *Vulcanococcus* represented greater than 50 percent of the cyanobacterial population in river sites, with *Dolichospermum* and *Microcystis* representing less than 15 percent individually.

Across the Gulf, there were two major scenarios, with cyanobacterial community composition dominated by either bloom-forming cyanobacterial genera (*Dolichospermum* and *Microcystis*; sites 1, 2, 3, 4, 7, 18-24), or picocyanobacterial genera (*Cyanobium* and *Vulcanococcus*; sites 5, 6, 8, 9, 25, 26, 30, and 31). In 2022, there was a large cyanoHAB event made up of *Dolichospermum* and *Microcystis* near Homa Bay, together making up approximately 98% of the cyanobacterial community. In 2023, *Dolichospermum* and *Microcystis* dominated sites along the southern and southwestern coast of the gulf, beginning near the Awach

River mouth through Homa Bay and towards Mbita, together making up between 77% and 95% of the cyanobacterial community at these sites. At the sites where *Dolichospermum* and *Microcystis* co-occurred, *Dolichospermum* was more abundant than *Microcystis* by approximately 25 to 60 percent. In a few instances, a single genus of cyanobacteria dominated the cyanobacterial assemblage: *Planktothrix* at site 10 in Kisumu harbor; *Microcystis* at site 15 near the Sondu River mouth; *Sphaerospermopsis* at site 27 off Rusinga Island in the open waters of Lake Victoria; and *Dolichospermum* at sites 28 and 29 outside the gulf near Bondo and the Yala River mouth.

3.3.4 Relationships Between Dominant Cyanobacterial Genera and Environmental Conditions

Cyanobacterial composition was assessed in relation to selected environmental parameters through Non-metric MultiDimensional Scaling (NMDS) (**Figure 3.9.4A)**. River sites and sites dominated by picocyanobacteria *Cyanobium* and *Vulcanococcus* clustered together and showed a positive relationship with dissolved oxygen (DO). Sites dominated by *Dolichospermum* and *Sphaerospermopsis* clustered near each other and had a positive relationship with dissolved phosphorus, pH, and depth. Sites with mixed communities of bloom-forming cyanobacteria including *Microcystis* and *Planktothrix* formed a large cluster and had a positive relationship with ammonia, nitrate, and turbidity.

The relative abundance of dominant cyanobacterial genera (*Dolichospermum, Cyanobium, Vulcanococcus, Microcystis, Planktothrix, and Sphaerospermopsis*) were assessed for their relationship with environmental parameters using Spearman's correlation analysis **(Figure 3.9.4B, Table S3.10.1).** *Dolichospermum, Microcystis,* and *Planktothrix* shared similar relationships with environmental parameters. These genera had strong positive correlations (greater than or equal to 0.5) with ammonia, temperature, chlorophyll, and turbidity ($P < 0.001$).

This group was strongly negatively correlated (less than -0.5) with depth and DO ($P < 0.001$). *Sphaerospermopsis, Cyanobium,* and *Vulcanococcus* clustered together based on their relationships with environmental parameters. *Sphaerospermopsis* was strongly positively correlated with DO, depth, and dissolved phosphorus, while strongly negatively correlated with ammonia, dissolved N:P, nitrate, and turbidity ($P < 0.001$). *Sphaerospermopsis'* negative correlation with dissolved N:P levels suggest it to be potentially more fit in nitrogen limiting waters, where this negative correlation most often occurs at dissolved N:P ratios less than 1 (indicative of nitrogen deficiency). *Cyanobium* and *Vulcanococcus* were strongly positively correlated with depth and strongly negatively correlated with ammonia, temperature, and cyanobacterial detection ("cyanos") via the use of an AlgaeTorch (bbe moldaenke GmnH, Schwentinental, Germany) ($P < 0.001$).

3.3.5 Phylogenomic Analysis Reveals Low Intra-Genus Genetic Diversity and Novel Clades of Bloom-Forming Cyanobacteria from the Gulf

Assembly and binning of metagenomic sequence reads produced 203 high quality cyanobacterial metagenome-assembled genomes (MAGs) (≥ 90% completion, ≤ 10% contamination). Dereplication at 99.5% average nucleotide identity produced a total of 17 representative MAGs of the genera *Microcystis* (5 MAGs)*, Dolichospermum* (5 MAGs)*, Planktothrix* (3 MAGs)*,* and *Sphaerospermopsis* (4 MAGs) (**Table 3.9.2**). Phylogenomic analysis showed that *Microcystis* MAGs from the gulf formed two distinct clades (**Figure 3.9.5A**). Three MAGs from sites 1, 2, and 5 clustered with the Mae4 clade (36), including strain LE19-84.1 isolated from Western Lake Erie (37). Two MAGs from sites 26 and 27 fell within the Mae5 clade. *Dolichospermum* MAGs from sites 1, 25, 26, 27, and 28 formed a single, tight sister clade to a clade of *Dolichospermum circinale* and two MAGs from Lake Erie, including

one MAG that has the saxitoxin BGC (**Figure 3.9.5B**) (38). *Planktothrix* MAGs from sites 10, 15, and 17 formed a single clade most closely related to *Planktothrix* strains with benthic properties (**Figure 3.9.5C**). *Sphaerospermopsis* MAGs from sites 25, 26, 27, and 30 formed a single, tight clade most closely related to *Sphaerospermopsis* sp. LEGE 00249, an isolate from a reservoir in Portugal (**Figure 3.9.5D**) (39).

3.3.6 Nitrogen Fixation Genes Harbored Primarily by Dolichospermum and Identified at Sites with Large Bloom Events

Due to nitrogen limitation in the gulf and its potential selection for diazotrophic organisms, we assessed the microbial community's ability to fix nitrogen by identifying *nif*HDK genes (40). *Dolichospermum* and *Anabaena,* two members of the *Anabaena/Dolichospermum/Aphanizomenon* (ADA) clade, had the most *nif*HDK gene matches, and very few hits were to bacteria outside of the Cyanobacteria phylum (**Figure 3.9.6**) (30). *nif*HDK genes belonging to ADA clade members were found at sites where there were large cyanoHAB events in 2022 and 2023. The only instance where the three ADA clade members did not harbor most of the *nif* genes was at site 27 near Bridge Island where *Spaherospermopsis,* a closely related genera, dominated the community (**Figure 3.9.3**) (10). At this site, *Sphaerospermopsis* and other *Nostacales* cyanobacteria were responsible for the *nif*HDK genes detected.

3.3.7 Cyanotoxin- and Cyanopeptide-Encoding BGCs Detected in Regions with and without Visible CyanoHABs

Mapping metagenomic sequence reads to biosynthetic gene clusters (BGCs) from the MiBIG database (41) revealed high spatiotemporal variation in genetic potential for biosynthesis of cyanotoxins and cyanopeptides (**Figure 3.9.7**). No cyanobacterial BGCs were found in the

four river sites, or at sites 6, 8, 9, 11, 19, 20, and 31, which had limited abundance of cyanobacteria. BGCs encoding metabolites from the classes aeruginosins, anacyclamide, microcystins, micropeptins, and microviridins were most common, detected in over half of the sites. The microcystin and *Anabaena* anacyclamide 1 BGCs were the most abundant BGCs across the dataset. Although genes encoding cylindrospermopsin (28, 42) did not assemble using our metagenomic approach, read mapping revealed the presence of four highly conserved cylindrospermopsin genes (*cyr*AEBC) at a low abundance at site 27 near Bridge Island (**Figure S3.10.1**). Deeper sequencing could fully elucidate this BGC and its taxonomic identity in the gulf.

3.3.8 Genus-Specific Biosynthetic Potential Revealed Through Over 300 Cryptic BGCs

BGCs were mined from cyanobacterial MAGs to assess total biosynthetic potential of the Winam Gulf cyanobacterial assemblage. 1,000 BGCs from cyanobacterial MAGs were identified, 300 of which were unique, from 11 different biosynthetic classes and 14 different cyanobacterial genera (**Figure 3.9.8)**. Notably, all but three of these BGCs did not match any known BGCs in the MiBIG database, indicating their cryptic nature (**Figure S3.10.2**). BGC richness differed between the 14 cyanobacterial genera. *Synechococcus, Vulcanococcus, Cyanobium,* and the unclassified cyanobacterial genera groups all encoded an average of one BGC per MAG, with *Cyanobium* encoding the largest range of BGCs per MAG, ranging between 1 and 8. *Pseudanabaena, Raphidiopsis, Nodosilinea, Snowella,* and *Elainella* encoded an average of 2 to 4 BGCs per MAG. *Planktothrix, Cuspidothrix, Dolichospermum, Microcystis,* and *Sphaerospermopsis* encoded the most BGCs, ranging between 6 to 10 BGCs per MAG on average.

Terpenes, the largest and most structurally diverse class of natural products including compounds such as geosmin, were encoded by all 14 cyanobacterial genera analyzed (43). The next most common biosynthetic class was non-ribosomal peptide synthetases (NRPS), a pathway known for its synthesis of cyanopeptides like anabaenopeptins and aeruginosins (44, 45). Cyanobactin BGCs encoded by *Microcystis* and *Dolichospermum* formed gene cluster families (GCFs) with known anacyclamide BGCs from the MiBIG database (**Figure S3.10.2**). BGCs putatively encoding microviridin from the RiPP biosynthetic class were encoded by *Planktothrix, Microcystis,* and *Sphaerospermopsis* (**Figure S3.10.2**). *Microcystis* was the only cyanobacterial genus with MAGs encoding microcystins (**Figures S3.10.2, S3.10.3**). BGCs that could not be grouped into these larger biosynthetic classes were categorized as "Other" and were putatively annotated as ribosomally-synthesized and post-translationally modified peptide (RiPP) recognition element (RRE)-containing, ectoine, lassopeptide, spliceotide, arylprolene, crocagin, and mycosporine-like.

3.3.9 Biosynthetic Potential Converges with Site-Specific Cyanobacterial Composition

The relative abundance of cyanobacterial BGCs varied across the gulf (**Figure 3.9.9**). BGC composition at each site converged with cyanobacterial composition, particularly with the dominating cyanobacterial genera at each site (**Figure S3.10.4**). Specifically, sites with high abundance of bloom-forming cyanobacteria (1-4, 23, and 27-29) had increased abundances of most BGC classes and clustered together based on the site's biosynthetic potential. BGCs putatively encoding pathways related to Anacyclamide-like, cyanobactin, heterocyst glycolipids, Lanthipeptide class V, NRPS, NRPS-like, RRE-containing, Type 1 polyketide synthase (T1PKS), RiPP-like, spliceotide, terpene biosynthetic classes and pathways were the most common and abundant BGC types across sites.

3.4 Discussion

The Winam Gulf of Lake Victoria experiences prolific, year-round cyanoHABs, previously shown to be dominated by *Microcystis* with periodic occurrences of *Dolichospermum* based on morphological classifications (17, 46). However, due to limited molecular research on these blooms, little is known about the bloom-forming cyanobacterial function and biosynthetic potential supporting year-around growth in this region. In this study, we employed metagenomics to establish a molecular snapshot of the microbial communities found along spatial and physicochemical gradients in the Winam Gulf between two consecutive years. Subsequently, we investigated the functional and biosynthetic potential encoded by bloom-forming cyanobacteria in the gulf to determine potential strategies used by these organisms to grow and produce toxic metabolites.

3.4.1 Nitrogen Limitation May Provide a Competitive Advantage for Diazotrophs During CyanoHABs

Two, large cyanoHAB events dominated mainly by *Dolichospermum* (56 to 91% of the cyanobacterial community) and *Microcystis* (14 to 39% of the cyanobacterial community) were detected near Homa Bay in 2022 and outside the gulf near Bondo and the Yala River mouth in 2023 (**Figure 3.9.1)**. Toxic cyanoHAB events, due to microcystins, in Homa Bay have presented exposure risks to this fishing community of ~60,000 residents (21, 28, 47). CyanoHABs have also been previously reported in the Yala River mouth region, which has excessive turbidity, declining water quality, and is a large source of anthropogenic pollution to Lake Victoria (48, 49). While this region is not as regularly monitored for cyanoHABs and cyanotoxins due to its location outside the gulf, our results provide further evidence that it should be considered in future monitoring schemes.

Organisms from the ADA clade and *Sphaerospermopsis* harbored essential genes for nitrogen fixation, such as the nitrogenase operon (*nif*HDK) and heterocyst glycolipid synthesis gene clusters, where cyanoHABs were detected in 2022 and 2023 (50). To our knowledge, this is the first report of the genus *Sphaerospermopsis*, a potentially toxigenic and diazotrophic cyanobacterium, in the Winam Gulf. These results suggest that diazotrophic cyanobacterial genera might be employing nitrogen fixation as a competitive advantage over non-diazotrophic competitors that are reliant on fixed forms of nitrogen. Low TN:TP ratios at cyanoHAB sites in 2023 (34) and 2005 (26) also indicate nitrogen limitation, with less severe nitrogen limitation closer to and outside the Rusinga channel (34). Although TN:TP was not measured in 2022, dissolved nitrate: soluble reactive phosphorus (DN:SRP) ratios were all below 1, suggestive of nitrogen deficiency (17). These results concur with Brown et al. 2024 which identified *Dolichospermum* as a dominant bloom-forming cyanobacteria in the gulf in 2022, but differ from previous reports that identified *Microcystis* as the dominating bloom-forming cyanobacteria in most seasons since 2008 in this region (28, 17, 51). Investigations of the phytoplankton community of the gulf in 2020 also identified *Microcystis* at slightly higher abundance than *Dolichospermum* in May and June (29). Data collection on an annual basis is vital to elucidate how changing landscapes and populations in this region are influencing cyanoHAB community structures with different toxin threats that require variant monitoring and management approaches.

According to additional studies on subtropical water bodies experiencing cyanoHABs made up of both *Microcystis* and diazotrophic cyanobacteria, high *nif*HDK (nitrogenase) gene expression was observed after large *Microcystis* blooms, with diazotrophs potentially acting as "replenishers" of bioavailable nitrogen (52). Gene expression data is needed to identify if

diazotrophic cyanobacteria in the Winam Gulf serve this role in the ecosystem and could help explain the gulf's seasonal bloom succession patterns.

3.4.2 Non-diazotrophic Cyanobacteria were Identified at Lower Abundance in Gulf Regions with Distinct Physicochemical Conditions

Microcystis and *Planktothrix*, non-diazotrophs, were at lower relative abundance at sites with large cyanoHAB events (between 1% and 39% of the cyanobacterial community). However, *Microcystis* was observed at higher relative abundance at sites in 2023 with higher turbidity on average (between 36% and 59% of the cyanobacterial community). Although these sites were nitrogen limited, the high turbidity at these sites may give *Microcystis* an advantage in comparison to deeper and less turbid regions of the gulf. This is due to *Microcystis'* ability to form gas vesicles and thus reach the surface to circumvent light limitation, plausibly caused by the shallow depth and well-mixed water column at these sites, therefore generating shading for their non-buoyant competitors (53). *Planktothrix,* which was found to have close genomic similarity to previously characterized benthic *Planktothrix* isolates*,* was the dominant cyanobacterial genera in Kisumu Harbor, the shallowest site in the study (5.4 meters), at the time of sampling (up to 67% of the cyanobacterial community). No *Planktothrix nif* genes were identified, indicating that *Planktothrix* was likely using strategies other than diazotrophy to overcome nitrogen limitation and proliferate in this region of the gulf (54, 55).

At sites where cyanobacteria were detected at lower abundance compared to other microbial phyla, *Cyanobium* and *Vulcanococcus* (picocyanobacteria) were the most abundant cyanobacterial genera, each making up almost half of the cyanobacterial community at these sites. There is little documentation of the picocyanobacterial community in the Winam Gulf likely due to their small cell size and thus difficult morphological classification (56). *Cyanobium*

is cosmopolitan, found across the globe, and is well known for its ability to scavenge nitrogen and use light efficiently, potentially leading to its success in this ecosystem, while *Vulcanococcus* is not as well documented (57, 58). While picocyanobacteria typically do not have many BGCs, MAGs recovered in this study do encode some BGCs from terpene, PKS, and NRPS biosynthetic pathways, with *Cyanobium* encoding a far more diverse set of BGCs than *Vulcanococcus.* This finding underscores the need to include picocyanobacteria in future investigations of cyanoHABs in the Winam Gulf and freshwater systems broadly as it relates to bloom dynamics and human health (59).

3.4.3 Low Intra-Genus Diversity Among Bloom-Forming Cyanobacteria in the Gulf Compared to Temperate Systems

The main bloom-forming cyanobacteria (*Dolichospermum, Microcystis, Sphaerospermopsis, and Planktothrix)* showed low intra-genus diversity across the sites, recovering only 3-5 MAGs from each genus. For *Microcystis* and *Planktothrix,* this result contrasts with findings from Lake Erie, where isolates from these genera share much lower average nucleotide identities (37, 60). Conversely, the results for *Microcystis* align with findings from subtropical lakes like Lake Okeechobee (Florida, USA) and Lake Taihu (China). Low intra-genus diversity was identified by Krausfeldt et al. (2024) (61) in Lake Okeechobee, where only three *Microcystis* MAGs, two of which were annotated as *Microcystis* panniformis (a commonly found species in Lake Taihu), were recovered from a large dataset of metagenomes (62). Three out of five of the non-redundant *Microcystis* MAGs in our study were annotated as *M. panniformis,* while the other two MAGs were annotated as *M. aeruginosa* and *Microcystis* sp. (**Figure 3.9.5**). *M. panniformis* and *M.aeruginosa* are known to share high levels of genomic similarity, but differ in their colony morphology, with *M. panniformis* often forming larger

colonies, potentially giving them an advantage in shading out their competitors (63). Together, these data show low intra-genus diversity and the common occurrence of *M. panniformis* across subtropical and tropical lakes, indicating a potential ecological pattern across lakes with shared climates (distinct dry and rainy seasons, warm temperatures, high humidity, and reduced seasonality). This data could not be compared directly to other studies on African Great Lakes due to limited data availability, highlighting these high-quality MAGs as a foundational resource for metagenomic research on cyanoHABs in this region.

3.4.4 BGCs Encoding the Synthesis of Cyanotoxins, Cyanopeptides, and Unknown Metabolites Found Across the Gulf, Even in Areas Without Visible CyanoHABs

The diverse makeup of cyanobacteria within the Winam Gulf pose threats not only due to their noxious smells and effects on fishing, tourism, recreation, and wildlife, but also due to their ability to produce toxins that are harmful to humans and wildlife (21, 47). We identified the *mcy* BGC responsible for the synthesis of microcystin at over half of the sites and definitively ascribed its source for the first time, solely to *Microcystis*, resolving speculation that *Microcystis* or *Dolichospermum* could harbor these genes (28, 64). The expression of the *mcy* BGC has been shown to increase in response to temperature, presence of reactive oxygen species (ROS), light, nitrogen forms and availabilities, and growth rate, though factors which constrain the biosynthesis of microcystins remain convoluted (65, 66). The attribution of microcystin synthesis to *Microcystis* within this tropical, warm, eutrophic basin is important in accurate cyanotoxin human exposure management and prevention, especially in areas where raw water is used for drinking and domestic tasks. A previous study in the Winam Gulf identified high levels of the amidinotransferase gene (*cyrA)* in the cylindrospermopsin-encoding BGC near the Nyando River mouth in June 2022 using quantitative polymerase chain reaction (qPCR) (28, 42). In this study,

these genes were not detected at high levels and our methods were not able to assemble the cylindrospermopsin BGC or link it to an organism.

The bloom-forming cyanobacteria present in the gulf encoded the ability to produce other nitrogen-rich, hepatotoxic, cyanopeptides. These BGCs were identified at over half of the sites in the study, including those that did not have a large, visible cyanoHAB event. Many of the sites in this study are heavily used for fishing, aquaculture, bathing, washing clothes, and for potable water, underscoring the importance of communicating the dangers of non-visible cyanoHABs to local residents (21, 47). With changing conditions across rainy and dry seasons in the gulf, the composition and expression of these BGCs may change, producing diverse cocktails of harmful molecules beyond microcystins. Such seasonal expression changes were observed in a study in a subtropical river (Florida, USA) where the cyanopeptolin BGC, also identified in our study, was 80 times more expressed than the *mcy* operon in the winter season, while this relationship inverted in the spring season (52). BGC expression studies in tropical regions are needed to begin making hypotheses about metabolites' complex functional roles and environmental triggers. Moreover, the human health effects of these molecular combinations remain unknown and BGC expression studies across both tropical and temperate systems indicate the expression of multiple BGCs at any given time (11, 67).

In addition to cyanopeptides and cyanotoxins, cyanobacteria in the gulf encoded over 300 unique BGCs, most of which remain cryptic. These findings illustrate the rich ability of Winam Gulf cyanobacteria to produce uncharacterized metabolites with unknown but potentially harmful or therapeutic properties, such as dolostatins which have been used as potent payloads in anticancer therapies. The number and diversity of BGCs per MAG in the Winam Gulf was consistent with results from cyanobacteria in Western Lake Erie (37). The diverse biosynthetic

repertoire of these bloom-forming cyanobacteria is both an opportunity for discovery and has implications for protecting human health.

3.5 Conclusion

This study characterizes the cyanobacterial assemblage, function, and biosynthetic potential across different regions of the Winam gulf. The metagenomic data revealed different survival strategies available for use by major, bloom-forming cyanobacteria to harness distinct environments and conditions around the gulf, including extensive use of BGCs. Omics techniques (metagenomics, metatranscriptomics, metabolomics) should be incorporated into current monitoring schemes in the gulf and other African Great Lakes impacted by toxic cyanobacteria to increase research capacity on cyanoHAB ecology and dynamics, inform water management plans, and protect human and ecosystem health. A multi-faceted monitoring plan that incorporates these techniques will look different in various regions depending on the level of cyanoHAB exposure potential and available resources, personnel, and expertise. Moreover, filling this data gap is vital in understanding the biogeographical differences in cyanoHABs, specifically how cyanoHAB events in the global south differ from those in more temperate regions of the world. For example, our results suggest low genomic diversity of dominant cyanobacterial taxa in the gulf, differing from what is observed in temperate regions like the Laurentian Great Lakes but converging with findings from subtropical lakes like Lake Okeechobee and Lake Taihu. Insights from a diverse array of geographic regions will shed valuable insight into how cyanoHABs may respond to a rapidly warming climate across latitudinal gradients.

3.6 Materials and Methods

3.6.1 Data Collection

Study Site and Sample Collection

This study was conducted on June 24-25th, 2022 and May 30th-June 3rd, 2023 in the Winam Gulf. Thirty-one water samples were collected from a depth of 1 meter using a 2.5 L Van Dorn sampler. Microbial biomass was collected by filtering water (between 60-300 mL) through a 0.22 µm PES Sterivex cartridge filter (MilliporeSigma, Burlington, MA) and preserved in 1.5 mL of Zymo DNA/RNA shield (Zymo, Orange, CA) and placed in a -20°C freezer. Water samples for dissolved nutrient analyses were filtered through a 0.2 µm filter while unfiltered water was collected for total nutrient analyses and frozen at -20° C. Water was filtered through GF/F filters and stored in a sterile tube for chlorophyll analysis. Nutrient, DNA, and chlorophyll samples were shipped from Kisumu, Kenya at room temperature back to *University of Tennessee Knoxville* in 2022 and *University of Michigan* in 2023 for DNA extraction, and *Baylor University* for nutrient and chlorophyll analyses in 2023. These shipments were accompanied by a letter of support from Kenya Marine and Fisheries Research Institute (KMFRI) and/or a USDA permit from D.H.S.. Physicochemical parameters (i.e., water temperature, dissolved oxygen, conductivity, total dissolved solids, turbidity, and pH) were collected using a multiparameter sonde (YSI Inc., Yellow Springs, OH) on site. Total cyanobacterial abundance was measured using an AlgaeTorch (bbe moldaenke GmnH, Schwentinental, Germany).

DNA Extraction and Sequencing

All sequencing data for this manuscript has been summarized previously in a Microbiology Resource Announcement (MRA) by Zepernick & Hart et al. (*In press,* ASM MRA, 2024) and is briefly described below (68). For samples from 2022, DNA was extracted from the Sterivex

cartridges after flushing all DNA/RNA shield out of the filter cartridge using phosphate-buffered saline (PBS), prepared according to Cold Spring Harbor's protocol (69). Duplicate samples from each site in 2022 were collected, extracted, and sequenced. DNA extractions were performed at the *University of Tennessee Knoxville* using standard phenol-chloroform methods and subsequent ethanol precipitation as reported previously (70). DNA quality was assessed *via* the Nanodrop ND-100 spectrophotometer (Thermo Fisher Scientific, Waltham, MA) to ensure 260/280 ratios were sufficient (260/280 > 1.80). Following, DNA concentration was quantified using the Qubit double-stranded DNA HS assay kit (Invitrogen, Carlsbad, CA). Extracted DNA was submitted to the *University of Minnesota Genomics Core* for genomic library construction and sequencing. Sequencing was completed on an Illumina Novaseq system equipped with a S4 flowcell with 150 cycles (150-base pair paired end reads) (Illumina, San Diego, CA).

For samples from 2023, DNA was extracted from the Sterivex cartridges after flushing all DNA/RNA shield out of the filter cartridge using phosphate-buffered saline (PBS), prepared according to Cold Spring Harbor's protocol (69). DNA extractions were performed at the *University of Michigan*. Filter cartridges were opened using a pipe cutter and the filter membrane was carefully removed using sterile forceps. DNA from the filters was extracted using the DNeasy Powerwater Sterivex Kit (Qiagen, Germantown, MD) according to the manufacturer's protocol. Quantity was assessed using the high sensitivity DNA qubit kit (Invitrogen, Carlsbad, CA). Extracted DNA was submitted to the *University of Michigan Advanced Genomics Core* for genomic library construction and sequencing. Sequencing was completed on an Illumina NovaX system equipped with a 10B flowcell with 300 cycles (150-base pair paired end reads) (Illumina, San Diego, CA).

Nutrient Analyses

Water samples filtered through Whatman GF/F filters pre-combusted at 550° C for 4 hours were collected in new 50 mL Falcon Tubes and frozen at -20° C until analysis. Filtered dissolved nutrient samples were analyzed for nitrate-nitrogen (NO_3-N), ammonium-nitrogen (NH_4-N) and soluble reactive phosphate (PO_4-P) using Cd reduction, phenate, and molybdate methods, respectively (71). Total dissolved nitrogen (TDN) and total dissolved phosphorus (TDP) were analyzed from filtered samples using the Cd reduction and molybdate methods (72). Unfiltered samples were analyzed for total nitrogen (TN) and phosphorus (TP) using the Cd reduction and molybdate methods (72). All soluble and total nutrients were analyzed using a Lachat QuickChem 8500 Series II autoanalyzer (Lachat Instruments, Hach Company) at the *Center for Reservoir and Aquatic Systems Research (CRASR) located at Baylor University, Waco, TX*. Chlorophyll a in extracts was measured using the non-acidification method of Welschmeyer on a Turner Designs Trilogy fluorometer calibrated with pure Chlorophyll a standards (Turner Designs, Sunnyvale, CA, USA) (73).

3.6.2 Bioinformatic Analysis

Metagenomic Data Processing

Metagenomes were processed with the Great Lakes Atlas for Multi-Omics Research (GLAMR) pipeline (GLAMR Portal, GLAMR Github). fastp v0.23.2 was used to deduplicate (--dup_calc_accuracy 6), remove adapters (--detect_adapter_for_pe), trim (--cut_front --cut_tail --cut_window_size=4 --cut_mean_quality 20), and filter (--length_required 50 --n_base_limit 5 --

low_complexity_filter --complexity_threshold 7) raw sequence reads (74). BBmap was used to remove human contaminant reads mapping to the GENCODE release 38 human genome (75). Each sample was assembled using Megahit v1.2.9 with the meta-sensitive parameter (76). Assembled contigs longer than 2000 base pairs were used for binning. Concoct v1.1.0, MaxBin2 v2.2.7, Metabat2 v2.17, MetaDecoder v1.0.13, SemiBin v1.0.3, and VAMB v3.0.8 were used to bin contigs into metagenome assembled genomes (MAGs) (77–82). MAGs were clustered and dereplicated at 98 percent identity using dRep v3.2.0 (83, 84). MAG quality was assessed with CheckM v.1.2.2 and MAG taxonomy was assigned using GTDBtk and the GTDB release 214 database (85, 86).

Community Composition analysis

MAGs from all 31 samples were grouped and dereplicated using dRep v3.2.0 at 98 percent (84). MAG quality statistics were calculated with checkM v1.2.2 and bakta v1.8.1 (86, 87)**.** Quality-controlled (QC) reads were mapped to the dereplicated set of MAGs using CoverM v0.6.1 (using the default mapper: MiniMap2 v2.24) to get the relative abundance of each MAG in each sample (Figure 3) (88, 89). Observations from duplicate samples from 2022 were averaged in all applicable metagenomic methods in the study due to the replicates similarity (**Figure S3.10.5**). A Bray-Curtis dissimilarity matrix was calculated from the relative abundance of all cyanobacterial genera in each sample and assessed using Non-metric Multidimensional Scaling (NMDS) using the vegan v2.6 package in the R statistical software (90). Cyanobacterial makeup was assessed in relation to environmental parameters using the envfit function in the vegan v2.6 package in the R statistical software (Figure 4A) (90). The correlation of the six most abundant genera with each

other and associated environmental parameters was assessed using Spearman's correlation coefficient in the R statistical software.

Phylogenomic analysis

High-quality MAGs (checkM: ≥ 90% completeness and ≤ 10% contamination) annotated by GTDBtk as *Microcystis, Dolichospermum, Planktothrix,* and *Sphaerospermopsis* were dereplicated at 99.5 percent identity with dRep v3.2.0 to produce representative bins for each of these genera (84–86). A tree was built using these MAGs and reference genomes from NCBI (**SI Table S3.10.2**) using GToTree v1.7.05 (91, 92). A maximum likelihood tree for each taxa using 251 marker genes was generated. Reference genomes for the *Microcystis* (**Figure 3.9.5A**), ADA (**Figure 3.9.5B**), and *Planktothrix* (**Figure 3.9.5C**) trees were sourced from prior studies (36, 37, 93, 94, 60). Reference genomes for the *Spaherospermopsis* tree (**Figure 3.9.5D**) were chosen from the NCBI genome database if they were characterized as either a reference genome or a strain (91). An outgroup represented by *Cyanobium gracile PCC 6307* (NCBI: PRJNA158695) was used in each of the four trees. Phylogenomic trees were visualized in ITOL (95).

Nitrogen fixation genomic potential detection

QC reads from each sample were mapped to the *nif*H, *nif*D, *nif*K, *Chl*N, and *Chl*B databases in NFixDB (accessed March 2024, Database) using MiniMap2 v2.24 (89, 96). Hits to *Chl*N and *Chl*B were excluded, but were used in the read mapping step to ensure a competitive database was used (97). Reads that mapped to the reference with 80 percent coverage and 90 percent identity were kept. Hits to *nif*H, *nif*D, and *nif*K were kept if 80 percent of the gene was covered

with 90 percent identity (Figure 6). *nif*HDK were used to ensure the identification of true nitrogen fixers, and not pseudo-*nif*H containing microbes (40). Total fragments per kilobase per million mapped reads (FPKM) was calculated for each gene in each sample using Equation 3.6.1 (98).

Equation 3.6.1: Total Fragments per Kilobase per Million Mapped Reads (FPKM) of *nif* genes

$$\text{Total Fragments per Kilobase per Million Mapped Reads (FPKM)} = \frac{total\ \#\ of\ mapped\ reads}{(length\ of\ nif\ gene/1000) * (total\ \#of\ reads\ per\ sample)} * 1{,}000{,}000$$

Cyanopeptide and cyanotoxin BGC detection.

QC reads from each sample were mapped to all BGCs in the MIBiG database from the cyanobacterial genera *Microcystis, Dolichospermum, Planktothrix, Sphaerospermopsis, Anabaena, Aphanizomenon, Cylindrospermopsis,* and *Cyanobium* using MiniMap2 v2.24 (41, 89). Reads that mapped to the reference with 80 percent coverage and 90 percent identity were kept. Hits to BGCs were kept if 70 percent of the BGC was covered by reads in that sample. Total Fragments per Kilobase per Million Mapped Reads (FPKM) was calculated for each BGC in each sample using Equation 3.6.2 (98).

Equation 3.6.2: Total Fragments per Kilobase per Million Mapped Reads (FPKM) of biosynthetic gene clusters

$$\text{Total Fragments per Kilobase per Million Mapped Reads (FPKM)} = \frac{total\ \#\ of\ mapped\ reads}{(length\ of\ BGC/1000) * (total\ \#\ of\ reads\ per\ sample)} * 1{,}000{,}000$$

Mcy and cyr operon identification and comparison to references

All assemblies in the dataset were searched for the presence of genes from the *mcy* (MiBIG BGC: BGC0001016) and *cyr* (MiBIG BGC: BGC0000978) operons via BLASTn (41, 99). If genes from either of these operons were present in the assembly of a sample, the sample's MAGs were analyzed for all genes from the operons using BLASTn. Hits to genes of interest were kept if they had 80 percent identity and coverage in the queried MAG. MAGs with multiple genes from the reference operon present were analyzed in Anvi'o to identify if the whole reference operon was present and if it was correctly assembled and binned into the MAG (100). The *cyr* operon could not be recovered from this dataset, but low levels of the core biosynthetic genes of the *cyr* operon (MiBIG BGC: BGC0000978) were detected via BLASTn at Site 27. Bridge Island (**Figure S3.10.1)** (99)**.** The *mcy* operon was detected in multiple MAGs, all identified as *Microcystis* by GTDB-Tk (85). One MAG from Site 1. Homa Bay (NCBI Accession: SAMN41711332) was chosen for MAG analysis in Anvi'o to confirm the taxonomic identity of the contig with the *mcy* operon (**Figure S3.10.3**). Additionally, the *mcy* operon detected in the MAG was compared to the *mcy* operon from *Microcystis aeruginosa* FACHB 1326 (GenBank: OQ291092.1) using the program "clinker" (**Figure S3.10.3)** (101).

Biosynthetic gene cluster mining and abundance analysis

BGCs were mined from medium- and high-quality cyanobacterial MAGs (≥70% completion, ≤30% contamination) using antiSMASH v7.0 (102). All BGCs greater than 5,000 base pairs long were extracted and dereplicated at 99 percent using dRep v3.2.0 (**SI Table S3.10.3)** (84). Gene clusters were grouped into gene cluster families (GCFs) using BiG-SCAPE and visualized in cytoscape v3.9.1 (**SI Figure S3.10.2**) (103, 104). QC reads from each sample were mapped to

the dereplicated set of BGCs using MiniMap2 v2.24 (89). Reads that mapped to the reference with 80 percent coverage and 90 percent identity were kept. Hits to BGCs were kept if 70 percent of the BGC was covered by reads in that sample (**Figure 3.9.9**). FPKM for each BGC was calculated using 3.6.2. Total FPKM for each BGC type (x-axis) was calculated by summing the FPKM of each BGC belonging to that BGC type per sample.

Statistical Analysis and Figure Rendering

All statistical analysis and plots were created using R and R studio v4.3.3 (105). The R packages vegan, tidyverse, and RColorBrewer were used for figure creation (90, 106, 107).

3.7 Data Availability and Inclusion and Ethics Statement

Raw sequences are available on the NCBI SRA under BioProject PRJNA1119566 and BioSamples SAMN41659670- SAMN41659717. All MAGs have been deposited in the NCBI database under BioProject PRJNA1119566 and BioSamples SAMN41711301-SAMN41712291. Physicochemical data is available on the Biological and Chemical Oceanography Data Management Office (BCO-DMO) site (108). All code used in the analysis and visualization of the data presented in this study can be found on GitHub.

All collaborators of this study fulfilled the criteria for authorship required by American Society for Microbiology Applied and Environmental Microbiology (ASM AEM) and have been included as authors. Their participation was critical for the resource acquisition, design, implementation, and dissemination of this study and its results. The roles and responsibilities of all authors were agreed upon prior to the beginning of the project. This work includes locally relevant findings and interpretations, as determined in collaboration with local researchers. Local

and regional research is heavily cited throughout the manuscript, illustrating how this study fills local literature gaps. All data produced in this study is accessible to all collaborators of the project and the public.

3.8 References

1. Huisman J, Codd GA, Paerl HW, Ibelings BW, Verspagen JMH, Visser PM. Cyanobacterial blooms. Nat Rev Microbiol. 2018 Aug;16(8):471–83.
2. Griffith AW, Gobler CJ. Harmful algal blooms: A climate change co-stressor in marine and freshwater ecosystems. Harmful Algae. 2020 Feb 11;91(91):101590.
3. Ho JC, Michalak AM, Pahlevan N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature. 2019 Oct;574(7780):667–70.
4. Paerl H. Mitigating harmful cyanobacterial blooms in a human- and climatically-impacted world. Life. 2014 Dec 15;4(4):988–1012.
5. Chapra SC, Boehlert B, Fant C, Bierman VJ Jr, Henderson J, Mills D, et al. Climate change impacts on harmful algal blooms in U.S. freshwaters: A screening-level assessment. Environ Sci Technol. 20170725th ed. 2017 Aug 15;51(16):8933–43.
6. Igwaran A, Kayode AJ, Moloantoa KM, Khetsha ZP, Unuofin JO. Cyanobacteria harmful algae blooms: causes, impacts, and risk management. Water Air Soil Pollut. 2024 Jan 9;235(1):71.
7. Le Manach S, Duval C, Marie A, Djediat C, Catherine A, Edery M, et al. Global metabolomic characterizations of *Microcystis* spp. highlights clonal diversity in natural bloom-forming populations and expands metabolite structural diversity. Front Microbiol. 20190416th ed. 2019;10:791.
8. Harke MJ, Steffen MM, Gobler CJ, Otten TG, Wilhelm SW, Wood SA, et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. Harmful Algae. 2016 Apr;54:4–20.
9. Capelli C, Ballot A, Cerasino L, Papini A, Salmaso N. Biogeography of bloom-forming microcystin producing and non-toxigenic populations of *Dolichospermum lemmermannii* (Cyanobacteria). Harmful Algae. 2017 Jul 1;67:1–12.
10. Österholm J, Popin RV, Fewer DP, Sivonen K. Phylogenomic analysis of secondary metabolism in the toxic cyanobacterial genera *Anabaena, Dolichospermum*, and *Aphanizomenon*. Toxins. 2020 Apr 11;12(4):248.
11. Yancey CE, Yu F, Tripathi A, Sherman DH, Dick GJ. Expression of *Microcystis* biosynthetic gene clusters in natural populations suggests temporally dynamic synthesis of novel and known secondary metabolites in western Lake Erie. Glass JB, editor. Appl Environ Microbiol. 2023 Apr 18;e02092-22.
12. Paerl HW, Paul VJ. Climate change: Links to global expansion of harmful cyanobacteria. Water Research. 2012 Apr 1;46(5):1349–63.

13. Kaloudis T, Hiskia A, Triantis TM. Cyanotoxins in bloom: ever-increasing occurrence and global distribution of freshwater cyanotoxins from planktic and benthic cyanobacteria. Toxins. 2022 Apr;14(4):264.

14. Burford MA, Carey CC, Hamilton DP, Huisman J, Paerl HW, Wood SA, et al. Perspective: Advancing the research agenda for improving understanding of cyanobacteria in a future of global change. Harmful Algae. 2020 Jan 1;91:101601.

15. Svirčev Z, Lalić D, Bojadžija Savić G, Tokodi N, Drobac Backović D, Chen L, et al. Global geographical and historical overview of cyanotoxin distribution and cyanobacterial poisonings. Arch Toxicol. 20190726th ed. 2019 Sep;93(9):2429–81.

16. Plisnier PD, Kayanda R, MacIntyre S, Obiero K, Okello W, Vodacek A, et al. Need for harmonized long-term multi-lake monitoring of African Great Lakes. Journal of Great Lakes Research. 2023 Dec 1;49(6):101988.

17. Sitoki L, Kurmayer R, Rott E. Spatial variation of phytoplankton composition, biovolume, and resulting microcystin concentrations in the Nyanza Gulf (Lake Victoria, Kenya). Hydrobiologia. 2012;691(1):109–22.

18. Frank TH, Cornelissen IJM, Vijverberg J, Nagelkerke LAJ. Spatial and seasonal variation in the phytoplankton community of Lake Victoria's Mwanza Gulf, compared to northern parts of the lake. Journal of Great Lakes Research. 2023 Apr 1;49(2):453–62.

19. Njuguna H, Romero JR, Khisa P. The effect of turbid inflows into Winam Gulf, Lake Victoria: a 3D modeling study with ELCOM-CAEDYM. In: Proceedings of the 11th World Lakes Conference - Volume 2. 2006. p. 90–4.

20. Calamari D, Akech MO, Ochumba PBO. Pollution of Winam Gulf, Lake Victoria, Kenya: A case study for preliminary risk assessment. Lakes & Reservoirs: Science, Policy and Management for Sustainable Use. 1995;1(2):89–106.

21. Roegner A, Sitoki L, Weirich C, Corman J, Owage D, Umami M, et al. Harmful algal blooms threaten the health of peri-urban fisher communities: a case study in Kisumu Bay, Lake Victoria, Kenya. Expo Health. 2020 Dec;12(4):835–48.

22. Codd GA, Morrison LF, Metcalf JS. Cyanobacterial toxins: risk management for health protection. Toxicol Appl Pharmacol. 2005 Mar 15;203(3):264–72.

23. Jegede FE, Oyeyi TI, Abdulrahman SA, Mbah HAkwen, Badru T, Agbakwuru C, et al. Effect of HIV and malaria parasites co-infection on immune-hematological profiles among patients attending anti-retroviral treatment (ART) clinic in Infectious Disease Hospital Kano, Nigeria. PLoS One. 2017 Mar 27;12(3):e0174233.

24. Oyieko J, Copeland NK, Otieno S, Kifude C, Ocholla S, Hutter J, et al. Longitudinal and cross-sectional analyses of asymptomatic HIV-1/malaria co-infection in Kisumu County, Kenya. Am J Trop Med Hyg. 2023 Jan 11;108(1):85–92.

25. Marriott AL, Osano OF, Coffey TJ, Humphrey OS, Ongore CO, Watts MJ, et al. Considerations for environmental biogeochemistry and food security for aquaculture around Lake Victoria, Kenya. Environ Geochem Health. 2023 Aug 1;45(8):6137–62.

26. Gikuma-Njuru P, Hecky RE. Nutrient concentrations in Nyanza Gulf, Lake Victoria, Kenya: light limits algal demand and abundance. Hydrobiologia. 2005 Feb 1;534(1):131–40.

27. Misigo AWS, Suzuki S. Spatial-temporal sediment hydrodynamics and nutrient loads in Nyanza Gulf, characterizing variation in water quality. World Journal of Engineering and Technology. 2018 May 22;6(2):98–115.

28. Brown KM, Barker KB, Wagner RS, Ward CS, Sitoki L, Njiru J, et al. Bacterial community and cyanotoxin gene distribution of the Winam Gulf, Lake Victoria, Kenya. Environmental Microbiology Reports. 2024;16(3):e13297.

29. Suba VO, Khan MS, Miruka J, Mwalati L, Njoroge B, Oluoch E, et al. Seasonal phytoplankton ecosystem dynamics in response to environmental variables in Winam Gulf, Lake Victoria, Kenya. Aquat Sci. 2025 Jan;87(1):4.

30. Driscoll CB, Meyer KA, Šulčius S, Brown NM, Dick GJ, Cao H, et al. A closely-related clade of globally distributed bloom-forming cyanobacteria within the *Nostocales*. Harmful Algae. 2018 Jul 1;77:93–107.

31. Hounmanou YMG, Baniga Z, García V, Dalsgaard A. *Salmonella Salamae* and *S. Waycross* isolated from Nile perch in Lake Victoria show limited human pathogenic potential. Sci Rep. 2022 Mar 10;12:4229.

32. Baniga Z, Hounmanou YMG, Kudirkiene E, Kusiluka LJM, Mdegela RH, Dalsgaard A. Genome-based analysis of extended-spectrum β-lactamase-producing *Escherichia coli* in the aquatic environment and Nile perch (Lates niloticus) of Lake Victoria, Tanzania. Frontiers in Microbiology. 2020.

33. Hounmanou YMG, Leekitcharoenphon P, Hendriksen RS, Dougnon TV, Mdegela RH, Olsen JE, et al. Surveillance and genomics of toxigenic *Vibrio cholerae* O1 from fish, phytoplankton, and water in Lake Victoria, Tanzania. Frontiers in Microbiology. 2019.

34. Guildford SJ, Hecky RE. Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: Is there a common relationship? Limnology and Oceanography. 2000;45(6):1213–23.

35. Yanez-Montalvo A, Aguila B, Gómez-Acata ES, Guerrero-Jacinto M, Oseguera LA, Falcón LI, et al. Shifts in water column microbial composition associated to lakes with different trophic conditions: "Lagunas de Montebello" National Park, Chiapas, México. PeerJ. 2022 Sep 16;10:e13999.

36. Pérez-Carrascal OM, Terrat Y, Giani A, Fortin N, Greer CW, Tromas N, et al. Coherence of Microcystis species revealed through population genomics. Isme j. 20190730th ed. 2019 Dec;13(12):2887–900.

37. Yancey CE, Kiledal EA, Chaganti SR, Denef VJ, Errera RM, Evans JT, et al. The Western Lake Erie culture collection: A promising resource for evaluating the physiological and genetic diversity of *Microcystis* and its associated microbiome. Harmful Algae. 2023 Jul;126:102440.

38. Den Uyl PA, Kiledal EA, Errera RM, Chaganti SR, Godwin CM, Raymond HA, et al. Genomic Identification and Characterization of Saxitoxin Producing Cyanobacteria in Western Lake Erie Harmful Algal Blooms. Environ Sci Technol. 2025 Apr 10.

39. Martins J, Leão PN, Ramos V, Vasconcelos V. N-terminal protease gene phylogeny reveals the potential for novel cyanobactin diversity in cyanobacteria. Marine Drugs. 2013 Dec;11(12):4902–16.

40. Mise K, Masuda Y, Senoo K, Itoh H. Undervalued pseudo-*nifH* sequences in public databases distort metagenomic insights into biological nitrogen fixers. mSphere. 2021 Nov 17;6(6):e00785-21.

41. Terlouw BR, Blin K, Navarro-Muñoz JC, Avalon NE, Chevrette MG, Egbert S, et al. MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. Nucleic Acids Research. 2023 Jan 6;51(D1):D603–10.

42. Mihali TK, Kellmann R, Muenchhoff J, Barrow KD, Neilan BA. Characterization of the gene cluster responsible for cylindrospermopsin biosynthesis. Appl Environ Microbiol. 2008 Feb;74(3):716–22.

43. Helfrich EJN, Lin GM, Voigt CA, Clardy J. Bacterial terpene biosynthesis: challenges and opportunities for pathway engineering. Beilstein J Org Chem. 2019 Nov 29;15:2889–906.

44. Rouhiainen L, Jokela J, Fewer DP, Urmann M, Sivonen K. Two alternative starter modules for the non-ribosomal biosynthesis of specific anabaenopeptin variants in *Anabaena* (Cyanobacteria). Chem Biol. 2010 Mar 26;17(3):265–73.

45. Ishida K, Christiansen G, Yoshida WY, Kurmayer R, Welker M, Valls N, et al. Biosynthesis and structure of aeruginoside 126A and 126B, cyanobacterial peptide glycosides bearing a 2-carboxy-6-hydroxyoctahydroindole moiety. Chem Biol. 2007 May;14(5):565–76.

46. Lung'Ayia HBO, M'Harzi A, Tackx M, Gichuki J, Symoens JJ. Phytoplankton community structure and environment in the Kenyan waters of Lake Victoria.

47. Obuya JA, Onyango HO, Olokotum M, Zepernick B, Natwora K, Otieno D, et al. Socioeconomic consequences of cyanobacterial harmful algal blooms in small-scale fishing communities of Winam Gulf, Lake Victoria. Journal of Great Lakes Research. 2023 Sep 30;102236.

48. Aura CM, Odoli C, Nyamweya CS, Njiru JM, Musa S, Miruka JB, et al. Application of phytoplankton community structure for ranking the major riverine catchments influencing the pollution status of a lake basin. Lakes & Reservoirs: Science, Policy and Management for Sustainable Use. 2020;25(1):3–17.

49. Wanjala WSN, Olutende OM, Rose W, Joab O, Omuterema OS, Steve O. Rivers of resilience: assessing hydrological baseline data for nature-based conservation in the Yala River watershed. Open Access Library Journal. 2024 Feb 4;11(2):1–22.

50. Bauersachs T, Compaoré J, Hopmans EC, Stal LJ, Schouten S, Sinninghe Damsté JS. Distribution of heterocyst glycolipids in cyanobacteria. Phytochemistry. 2009 Dec 1;70(17):2034–9.

51. Miruka JB, Getabu A, Sitoki L, James O, Mwamburi J, George O, et al. Water quality, phytoplankton composition and microcystin concentrations in Kisumu Bay (Kenya) of Lake Victoria after a prolonged water hyacinth infestation period. Lakes & Reservoirs: Science, Policy and Management for Sustainable Use. 2021;26(4):e12380.

52. Hancock TL, Dahedl EK, Kratz MA, Urakawa H. The synchronicity of bloom-forming cyanobacteria transcription patterns and hydrogen peroxide dynamics. Environmental Pollution. 2024 May 1;348:123812.

53. Paerl H. Nuisance phytoplankton blooms in coastal, estuarine, and inland waters. Limnol Oceanogr. 1988;33:823–47.

54. Pancrace C, Barny MA, Ueoka R, Calteau A, Scalvenzi T, Pédron J, et al. Insights into the Planktothrix genus: genomic and metabolic comparison of benthic and planktic strains. Sci Rep. 2017 Jan 24;7(1):41181.

55. Hampel JJ, McCarthy MJ, Neudeck M, Bullerjahn GS, McKay RML, Newell SE. Ammonium recycling supports toxic Planktothrix blooms in Sandusky Bay, Lake Erie: Evidence from stable isotope and metatranscriptome data. Harmful Algae. 2019 Jan;81:42–52.

56. Meriluoto J, Spoof L, Codd GA, editors. Handbook of cyanobacterial monitoring and cyanotoxin analysis. 1st ed. Wiley; 2016.

57. Cabello-Yeves PJ, Picazo A, Camacho A, Callieri C, Rosselli R, Roda-Garcia JJ, et al. Ecological and genomic features of two widespread freshwater picocyanobacteria. Environmental Microbiology. 2018;20(10):3757–71.

58. Burson A, Stomp M, Mekkes L, Huisman J. Stable coexistence of equivalent nutrient competitors through niche differentiation in the light spectrum. Ecology. 2019;100(12):e02873.

59. Zepernick BN, McKay RML, Martin RM, Bullerjahn GS, Paerl HW, Wilhelm SW. A tale of two blooms: do ecological paradigms for algal bloom success and succession require revisiting? Journal of Great Lakes Research. 2024 Jun 1;50(3):102336.

60. McKindles KM, McKay RM, Bullerjahn GS. Genomic comparison of *Planktothrix agardhii* isolates from a Lake Erie embayment. PLOS ONE. 2022 Aug 23;17(8):e0273454.

61. Krausfeldt LE, Shmakova E, Lee HW, Mazzei V, Loftin KA, Smith RP, et al. Microbial diversity, genomics, and phage–host interactions of cyanobacterial harmful algal blooms. mSystems. 2024 Jun 10;9(7):e00709-23.

62. Huo D, Gan N, Geng R, Cao Q, Song L, Yu G, et al. Cyanobacterial blooms in China: diversity, distribution, and cyanotoxins. Harmful Algae. 2021 Nov 1;109:102106.

63. Zhang JY, Guan R, Zhang HJ, Li H, Xiao P, Yu GL, et al. Complete genome sequence and genomic characterization of *Microcystis panniformis* FACHB 1757 by third-generation sequencing. Standards in Genomic Sciences. 2016 Jan 28;11(1):11.

64. Brown KM, Ward CS, Bullerjahn GS. Metagenome-assembled genome sequences of two cyanobacterial cultures from Homa Bay County, Kenya. Thrash JC, editor. Microbiol Resour Announc. 2024 Feb 20;e01205-23.

65. Dziallas C, Grossart HP. Increasing oxygen radicals and water temperature select for toxic Microcystis sp. PLoS One. 2011 Sep 28;6(9):e25569.

66. Wei N, Hu C, Dittmann E, Song L, Gan N. The biological functions of microcystins. Water Research. 2024 Jul 18;122119.

67. Harke MJ, Davis TW, Watson SB, Gobler CJ. Nutrient-controlled niche differentiation of western Lake Erie cyanobacterial populations revealed via metatranscriptomic surveys. Environmental Science & Technology. 2016 Jan 19;50(2):604–15.

68. Zepernick BN, Hart LN, Natwora KE, Brown KM, Obuya JA, Olokotum M, et al. Metagenomic sequencing of cyanobacterial dominated Lake Victoria - an African Great Lake. ASM MRA. 2024.

69. Phosphate-buffered saline (PBS). Cold Spring Harb Protoc. 2006 Jun 1;2006(1):pdb.rec8247.

70. Zepernick BN, Truchon AR, Gann ER, Wilhelm SW. Draft genome sequence of the freshwater diatom *Fragilaria crotonensis* SAG 28.96. Stajich JE, editor. Microbiol Resour Announc. 2022 Sep 15;11(9):e00289-22.

71. Taylor JM, Andersen IM, Hoke AK, Kelly PT, Scott JT. In-situ N2:Ar ratios describe the balance between nitrogen fixation and denitrification in shallow eutrophic experimental lakes. Biogeochemistry. 2023 Dec;166(3):283–301.

72. Bridgewater L, Association APH, Association AWW, Federation WE. Standard methods for the examination of water and wastewater. American Public Health Association; 2012. 724 p.

73. Welschmeyer NA. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. Limnology & Oceanography. 1994 Dec;39(8):1985–92.

74. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018 Sep 1;34(17):i884–90.

75. Bushnell, B. BBMap [Internet]. Joint Genome Institute; 2015. Available from: https://sourceforge.net/projects/bbmap/

76. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 20150120th ed. 2015 May 15;31(10):1674–6.

77. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. Nat Methods. 20140914th ed. 2014 Nov;11(11):1144–6.

78. Wu YW, Tang YH, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome. 2014 Dec;2(1):26.

79. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019 Jul 26;7:e7359.

80. Liu CC, Dong SS, Chen JB, Wang C, Ning P, Guo Y, et al. MetaDecoder: a novel method for clustering metagenomic contigs. Microbiome. 2022 Mar 10;10(1):46.

81. Pan S, Zhu C, Zhao XM, Coelho LP. A deep siamese neural network improves metagenome-assembled genomes in microbiome datasets across different environments. Nat Commun. 2022 Apr 28;13(1):2326.

82. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, et al. Improved metagenome binning and assembly using deep variational autoencoders. Nat Biotechnol. 2021 May;39(5):555–60.

83. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol. 2018 Jul;3(7):836–43.

84. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. The ISME Journal. 2017 Dec 1;11(12):2864–8.

85. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil PA, Hugenholtz P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Research. 2022 Jan 7;50(D1):D785–94.

86. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 20150514th ed. 2015 Jul;25(7):1043–55.

87. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom. 2021 Nov;7(11):000685.

88. Woodcroft BJ. CoverM: Read coverage calculator for metagenomics [Internet]. [cited 2022 Jun 6]. Available from: https://github.com/wwood/CoverM

89. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018 Sep 15;34(18):3094–100.

90. Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science [Internet]. 2003 Apr 9 [cited 2024 Apr 15]; Available from: https://onlinelibrary-wiley-com.proxy.lib.umich.edu/doi/10.1111/j.1654-1103.2003.tb02228.x

91. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. Nucleic Acids Res. 2022 Jan 7;50(D1):D20-d26.

92. Lee MD. GToTree: a user-friendly workflow for phylogenomics. Bioinformatics. 2019 Oct 15;35(20):4162–4.

93. Yancey CE, Mathiesen O, Dick GJ. Transcriptionally active nitrogen fixation and biosynthesis of diverse secondary metabolites by *Dolichospermum* and *Aphanizomenon*-like cyanobacteria in western Lake Erie Microcystis blooms. Harmful Algae. 2023 May 1;124:102408.

94. Dreher TW, Davis EW, Mueller RS, Otten TG. Comparative genomics of the ADA clade within the *Nostocales*. Harmful Algae. 2021 Apr 1;104:102037.

95. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research. 2021 Jul 2;49(W1):W293–6.

96. Bellanger M, Figueroa JL, Tiemann L, Friesen ML, White RA. NFixDB (Nitrogen Fixation DataBase) - a comprehensive integrated database for robust ‘omics analysis of diazotrophs. NAR Genomics and Bioinformatics. 2024. Jun;6(2):lqae06.

97. Dos Santos PC, Fang Z, Mason SW, Setubal JC, Dixon R. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. BMC Genomics. 2012 Dec;13(1):1–12.

98. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010 May;28(5):511–5.

99. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009 Dec 15;10(1):421.

100. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, et al. Community-led, integrated, reproducible multi-omics with anvi’o. Nat Microbiol. 2020 Dec 21;6(1):3–6.

101. Gilchrist CLM, Chooi YH. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics. 2021 Aug 25;37(16):2473–5.

102. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research. 2023 Jul 5;51(W1):W46–50.

103. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, et al. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol. 20191125th ed. 2020 Jan;16(1):60–8.

104. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003 Nov;13(11):2498–504.

105. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available from: http://www.R-project.org/

106. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. JOSS. 2019 Nov 21;4(43):1686.

107. Neuwirth E. RColorBrewer: ColorBrewer Palettes [Internet]. 2014. Available from: https://CRAN.R-project.org/package=RColorBrewer

108. Bullerjahn GS, Zepernick BN, Hart LN. Metagenomic Time Series of Winam Gulf, Lake Victoria from 2022-2023 (ASI Lake Victoria project). Biological and Chemical Oceanography Data Management Office (BCO-DMO); 2024.

3.9 Figures and Tables

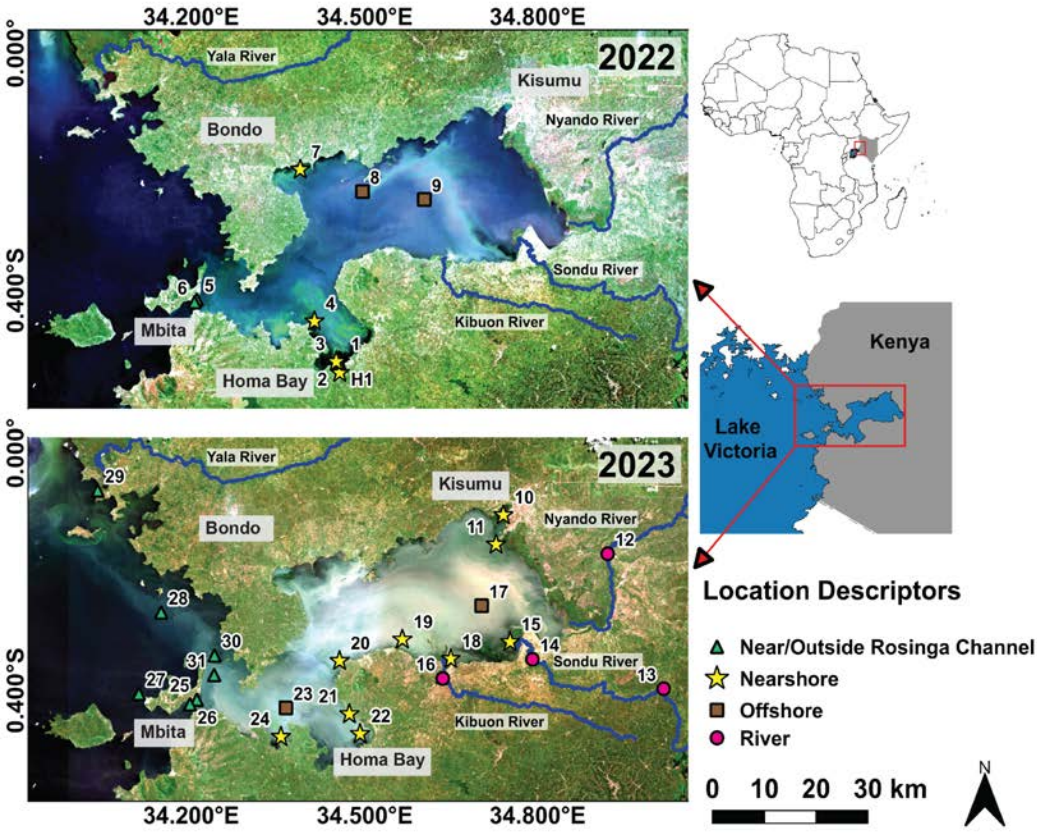


Figure 3.9.1: True color images of Winam Gulf, Lake Victoria, Kenya. The images were captured by Sentinel-2 on June 30th, 2022 (top panel) and May 26th, 2023 (bottom panel) and obtained as Level-2A (Surface Reflectance - atmospherically corrected by ESA using Sen2Cor) from Google Earth Engine. The images were collated together using QGIS LTR version 3.28.14. Sites sampled in 2022 are shown in the top panel, while sites sampled in 2023 are shown in the bottom panel. The location descriptor of a site is indicated by a symbol (star, circle, square, or triangle) with the site number indicated next to it on the map. Corresponding information for each site is reported in Table 1.

Table 3.9.1: Summary of sampled stations in this study.

Station Number	Site Name	Sample Date	Location Descriptor
1	Homa Bay	June 24th, 2022	Nearshore
2	Homa Bay Pier	June 24th, 2022	Nearshore
3	Homa Bay Water Intake	June 24th, 2022	Nearshore
4	Soklo	June 24th, 2022	Nearshore
5	Mbita East*	June 24th, 2022	Near/Outside Rusinga Channel
6	Mbita West*	June 24th, 2022	Near/Outside Rusinga Channel
7	Asembo Bay	June 25th, 2022	Nearshore
8	Ndere Island	June 25th, 2022	Offshore
9	Mid Gulf	June 25th, 2022	Offshore
10	Kisumu Harbor	May 29th, 2023	Nearshore
11	Dunga	May 29th, 2023	Nearshore
12	Nyando River	May 29th, 2023	River
13	Sondu River Bridge	May 29th, 2023	River
14	Sondu River Shore	May 29th, 2023	River
15	Sondu Miriu	May 29th, 2023	Nearshore
16	Kibuon River	May 30th, 2023	River
17	Southern Mid-Gulf	May 29th, 2023	Offshore
18	Awach River Mouth	May 30th, 2023	Nearshore
19	Bala Rawi	May 30th, 2023	Nearshore
20	Ingra	May 30th, 2023	Nearshore
21	Kowuor	May 30th, 2023	Nearshore
22	Oluch	May 31st, 2023	Nearshore
23	Sikli	May 31st, 2023	Offshore
24	Mirunda	May 31st, 2023	Nearshore
25	Mbita East*	June 1st, 2023	Near/Outside Rusinga Channel

26	Mbita West*	June 1st, 2023	Near/Outside Rusinga Channel
27	Bridge Island	June 1st, 2023	Near/Outside Rusinga Channel
28	Bondo	June 2nd, 2023	Near/Outside Rusinga Channel
29	Yala River Mouth	June 2nd, 2023	Near/Outside Rusinga Channel
30	Rusinga Channel	June 3rd, 2023	Near/Outside Rusinga Channel
31	South Rusinga	June 3rd, 2023	Near/Outside Rusinga Channel

*Indicates sampling at site in both years

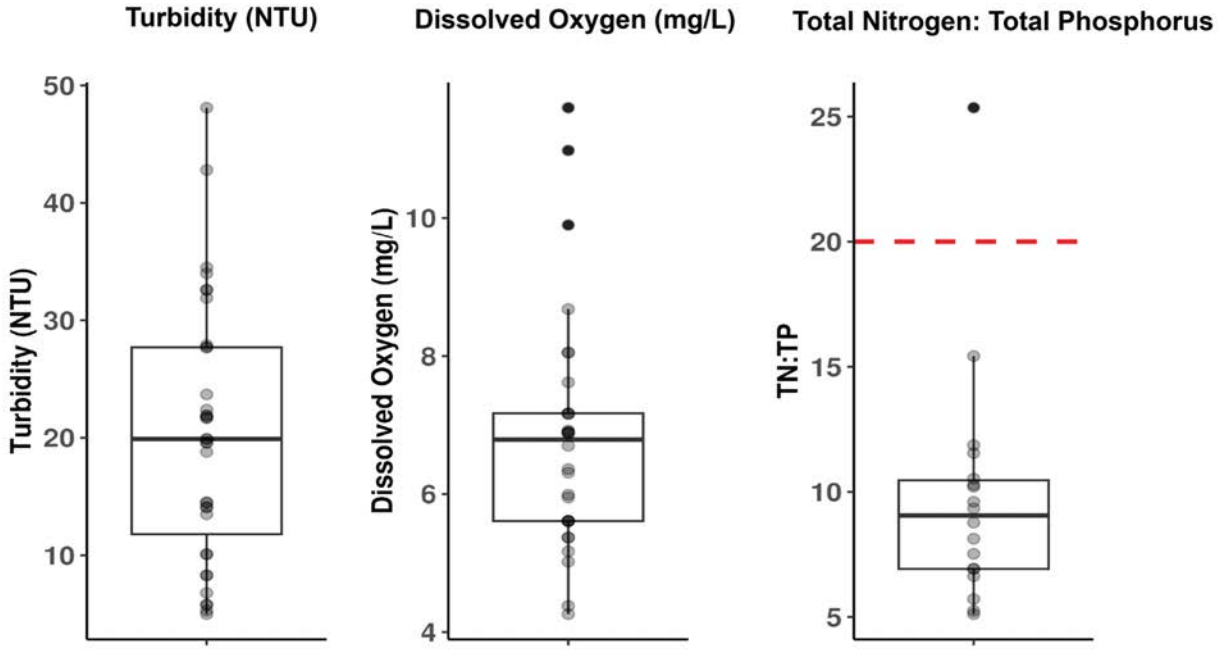


Figure 3.9.2: Turbidity (Nephelometric Turbidity Unit (NTU)), Dissolved Oxygen (mg/L), and Total Nitrogen: Total Phosphorus (TN:TP) (molar) ratio measurements. Individual data points for each variable are plotted on top of each box plot. The red line in the TN:TP panel indicates the threshold for nitrogen-limitation, with all values below it indicating nitrogen-limiting scenarios according to Guildford and Hecky (2000) (34). TN:TP was only measured for sites in 2023.

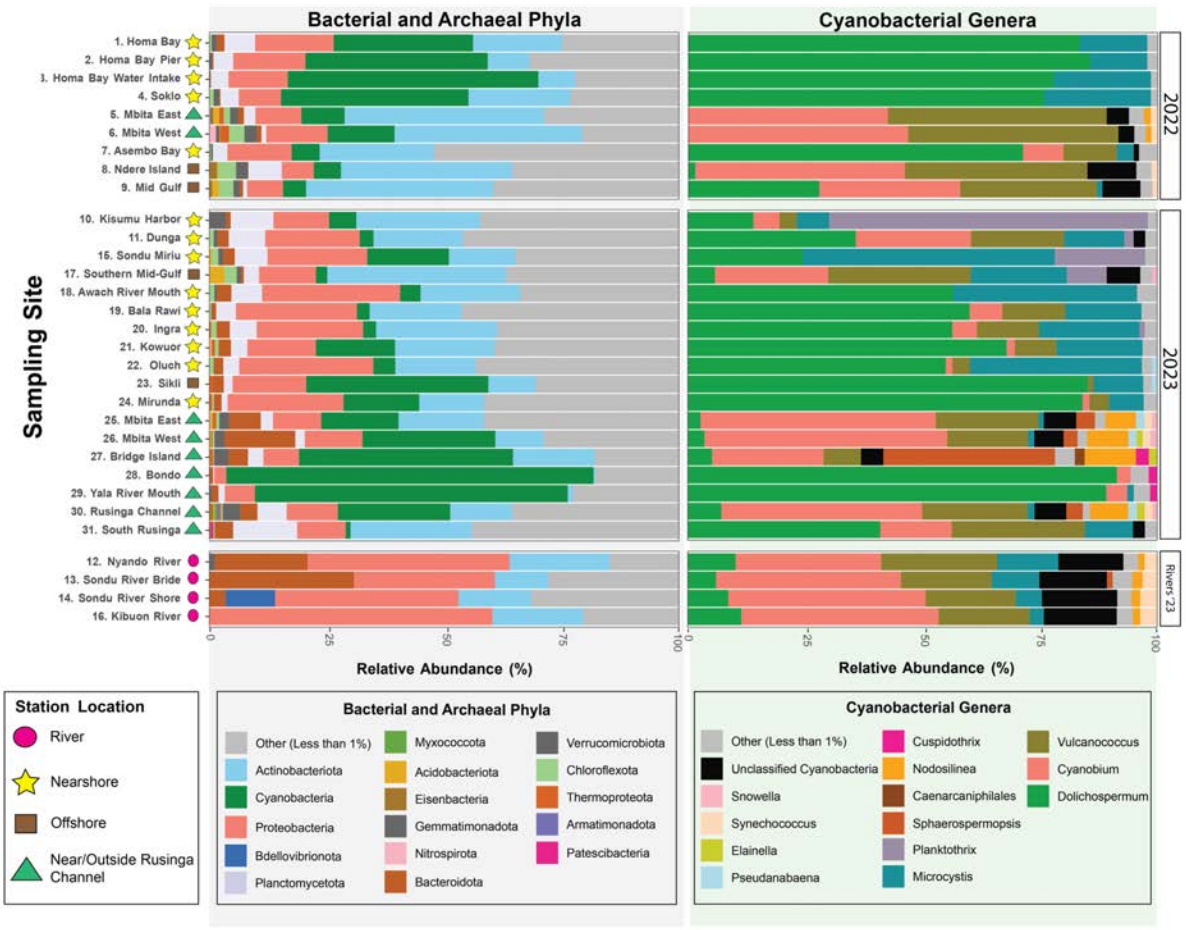


Figure 3.9.3: Microbial and archaeal composition makeup in Winam Gulf, LV. Relative abundance (percentage) of bacterial and archaeal phyla (gray panel) and cyanobacterial genera (green panel) from sampled sites in 2022 and 2023 is shown. The phyla and genera representing less than 1 percent of the population were grouped in the "Other" category. Cyanobacteria with no taxonomic assignment for genus were labeled as "unclassified cyanobacteria". Unmapped reads were excluded from this analysis and mapped reads were normalized to 100 percent for phyla and genera. Station location indicators are shown using symbols after the name of the site number and name on the y-axis.

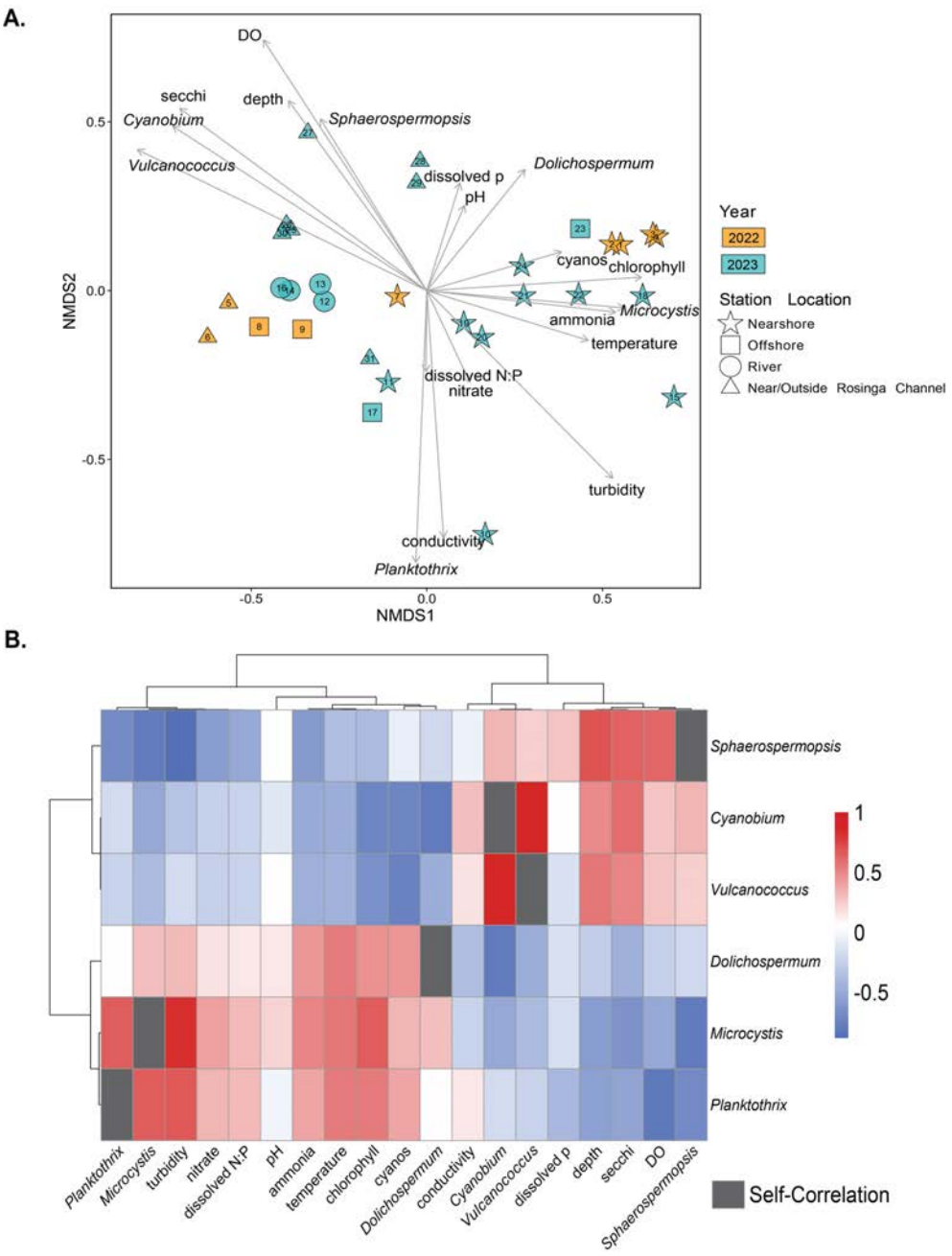


Figure 3.9.4: Cyanobacterial composition of Winam Gulf in relation to physicochemical parameters. **(A)** Non-metric multidimensional scaling (NMDS) analysis of the relative abundance of all cyanobacterial genera from each site, in relation to all other microbial community members (stress = 0.074). Labeled arrows show the direction and strength of the relationship between the ordination axes and associated environmental parameters and the relative abundance of the top 6 most abundant cyanobacterial genera across the sites. Shapes of ordination points indicate station location. Samples from 2022 are colored in amber and samples from 2023 in teal. **(B)** Spearman correlation coefficients between the 6 most abundant cyanobacterial genera across the sample sites and environmental parameters. Rows and columns are hierarchically clustered using correlation. Gray squares indicate self-correlation.

Table 3.9.2: Summary and statistics of high-quality, dereplicated cyanobacterial MAGs.

MAG	Site	Completeness (%)	Contamination (%)	Strain Heterogeneity (%)	Species *(Based on GTDBtk Annotations)*	Size (Mbp)	No. of Contigs	GC Content (%)	N50	NCBI Accession
samp_4315_VAMB_280	5. Mbita East	92.91	5.74	60	*Microcystis panniformis* A	4.230761	2568	43.5	2116	SAMN41711324
samp_4322_concoct_168	2. Homa Bay Pier	98.79	1.24	55.56	*Microcystis panniformis A*	5.518758	724	42.6	11220	SAMN41711381
samp_4325_VAMB_332	H1. Homa Bay Coast	95.24	6.77	37.21	*Microcystis* sp.	5.43631	4522	42.8	1747	SAMN41711408
samp_4444_concoct_1899	26. Mbita West	90.3	7.15	69.05	*Microcystis panniformis* A	3.847615	1713	43.2	2584	SAMN41711474
samp_4445_concoct_1450	27. Bridge Island	90.56	6.79	78.05	*Microcystis aeruginosa* H	3.729636	1804	43.1	2269	SAMN41711483
samp_4326_maxbin_35	H1. Homa Bay Coast	97.22	0.22	0	*Dolichospermum* sp.	4.077973	370	37.5	16022	SAMN41711415
samp_4443_metadecoder_676	25. Mbita East	93.74	1.01	33.33	*Dolichospermum* sp.	3.767753	443	37.2	11273	SAMN417

										11470
samp_4444_maxbin_761	26. Mbita West	91.61	0.81	33.33	*Dolichospermum* sp.	3.834118	345	37.1	15823	SAMN41711477
samp_4445_metadecoder_820	27. Bridge Island	95.11	0.44	50	*Dolichospermum* sp.	3.814968	232	37.1	24166	SAMN41711487
samp_4446_metabat2_471	28. Bondo	90.07	9.23	58.33	*Dolichospermum* sp.	4.539608	629	36.9	9935	SAMN41711489
samp_4432_maxbin_231	10. Kisumu Harbor	97.82	0.95	80	*Planktothrix* sp.	5.017227	441	39.3	19782	SAMN41711420
samp_4434_metabat2_31	15. Sondu Miriu	98.69	1.24	100	*Planktothrix* sp.	5.003513	421	39.2	18186	SAMN41711432
samp_4435_metabat2_87	17. Southern Mid-Gulf	98.47	1.69	63.64	*Planktothrix* sp.	5.119342	423	39.2	18944	SAMN41711440
samp_4443_maxbin_260	25. Mbita East	91.67	0.44	33.33	*Sphaerospermopsis kisseleviana* A	4.743136	543	37.6	11832	SAMN41711466
samp_4444_maxbin_786	26. Mbita West	92.89	1.12	83.33	*Sphaerospermopsis kisseleviana* A	4.915663	624	37.6	10851	SAMN

										417 114 78
samp_4445_VAMB_547	27. Bridge Island	89.96	7.53	77.55	*Sphaerospermopsis kisseleviana* A	6.110041	1452	37.5	6049	SAMN 417 114 82
samp_4448_concoct_1871	30. Rusinga Channel	93.7	1.78	77.78	*Sphaerospermopsis kisseleviana* A	5.020957	595	37.7	10901	SAMN 417 114 93

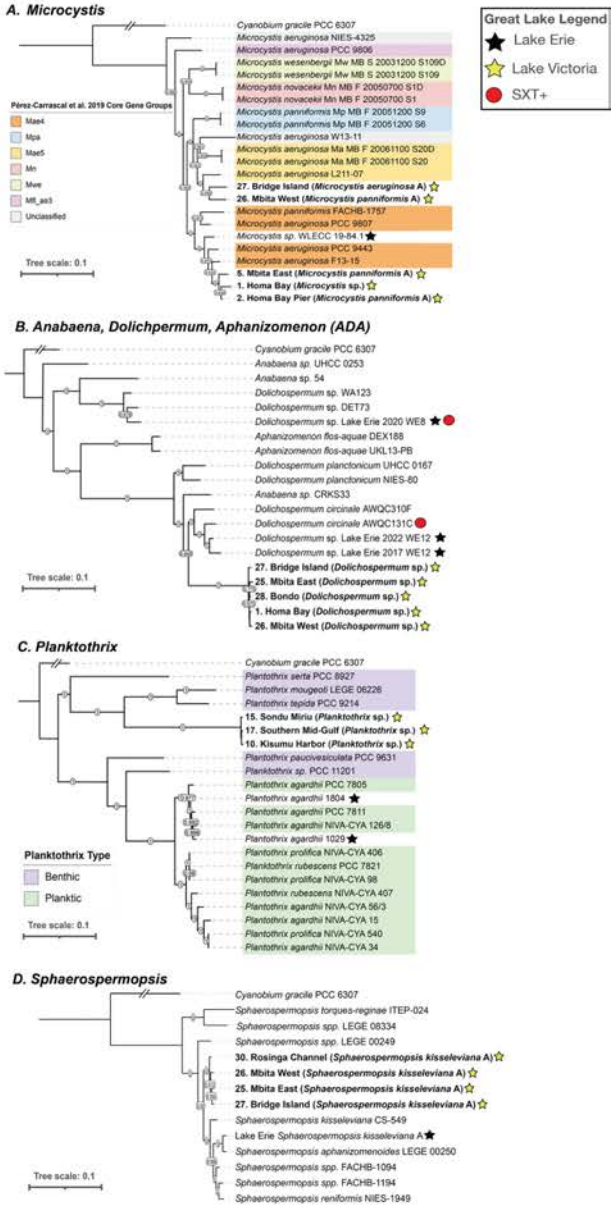


Figure 3.9.5: Phylogenomic analysis of dominant bloom-forming cyanobacteria from sampled sites with taxonomic annotations from GTDB (85). **(A)** *Microcystis* phylogenomic tree built with 5 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup. Colored ranges indicate clades as outlined in Pérez-Carrascal et al. 2019 (36). **(B)** *Dolichospermum* phylogenomic tree built with 5 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup The presence of a red circle is indicative of the presence of the saxitoxin gene cluster within the genome (38). **(C)** *Planktothrix* phylogenomic tree built with 3 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup. Benthic or planktic designations for reference genomes are indicated as colored ranges (54). **(D)** *Sphaerospermopsis* phylogenomic tree built with 4 representative MAGs, reference genomes, and *Cyanobium gracile* PCC 6307 as the outgroup.

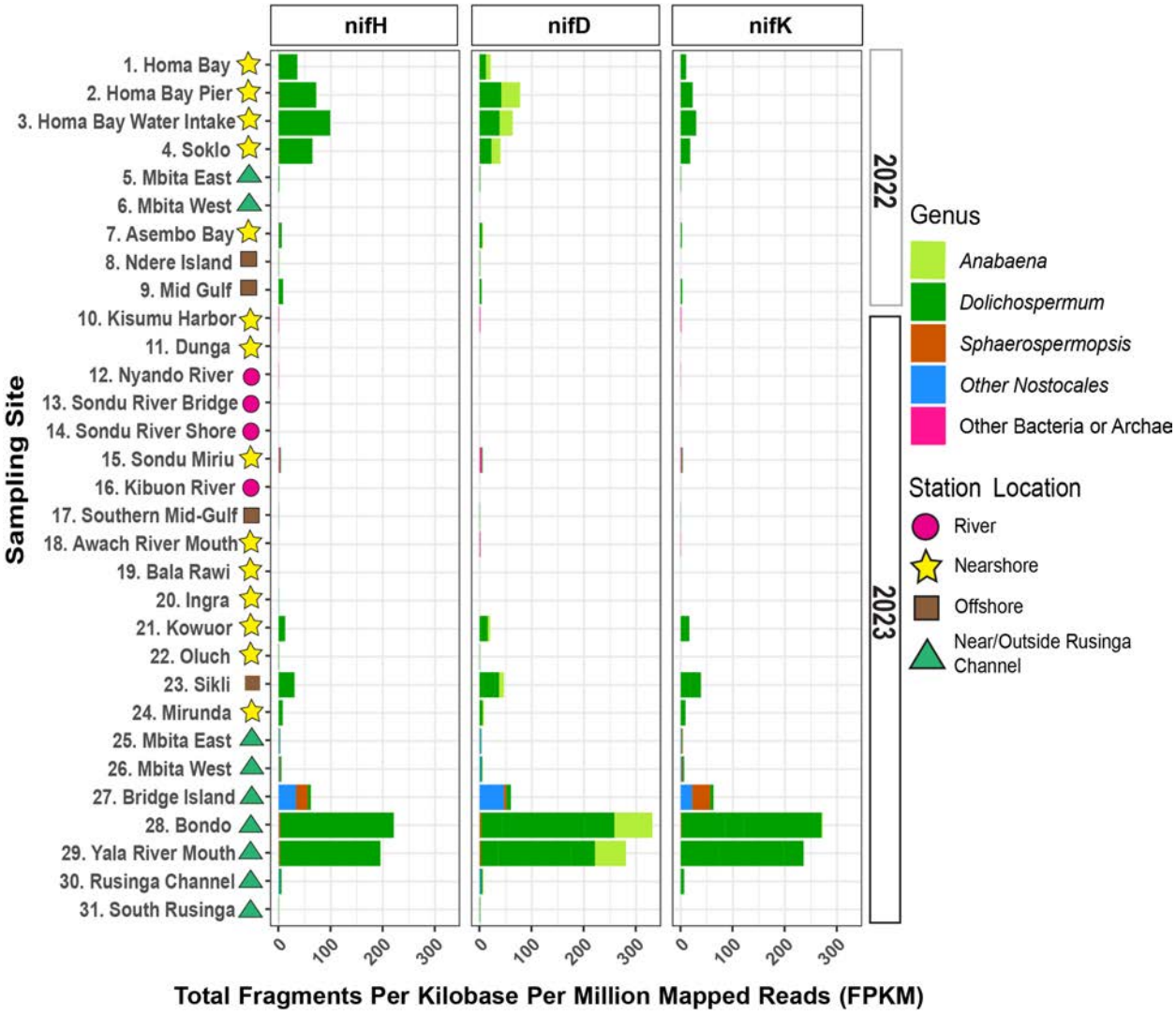


Figure 3.9.6: Relative abundance of *nifH, nifD,* and *nifK* genes across sites. Reads were aligned to *nifH, nifD,* and *nifK* genes from NFixDB and normalized using the total fragments per kilobase per million mapped reads (FPKM) metric (Equation 3.6.1) (96). The y-axis shows site names with their location descriptor symbol. The x-axis shows the FPKM of each *nif* gene in each sample. Bars are colored based on the genus the *nif* gene was extracted from.

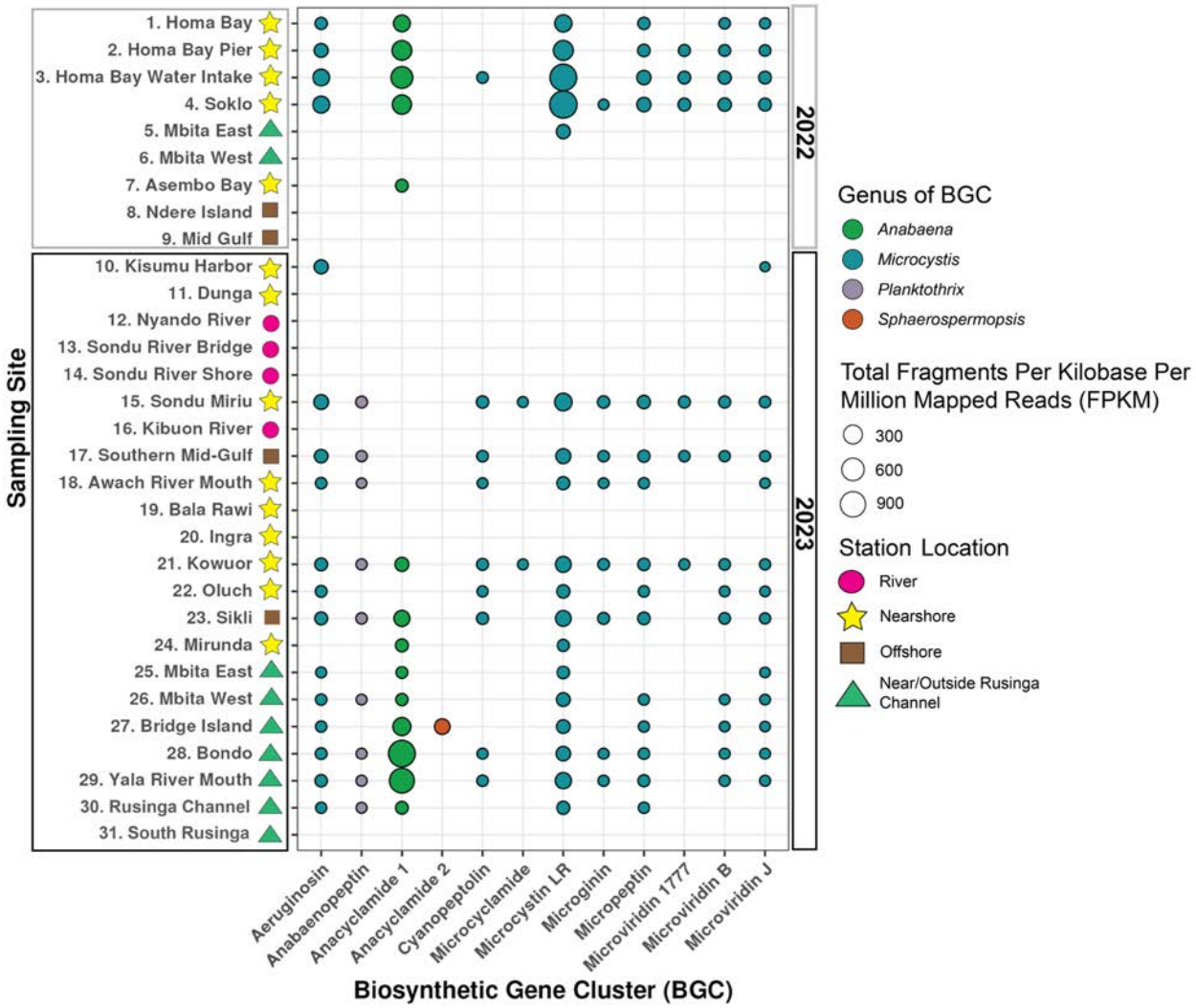


Figure 3.9.7: Known cyanopeptide BGC relative abundance across sites. Station numbers, names and their location indicator symbols are shown on the y-axis, with BGCs encoding the biosynthesis of cyanopeptides from MIBiG (accessed December 2023) shown on the x-axis (41). The FPKM (total fragments per kilobase per million mapped reads) of each whole BGC is shown as bubble size (Equation 3.6.2). Bubbles are colored based on the cyanobacterial genus the reference BGC originates from.

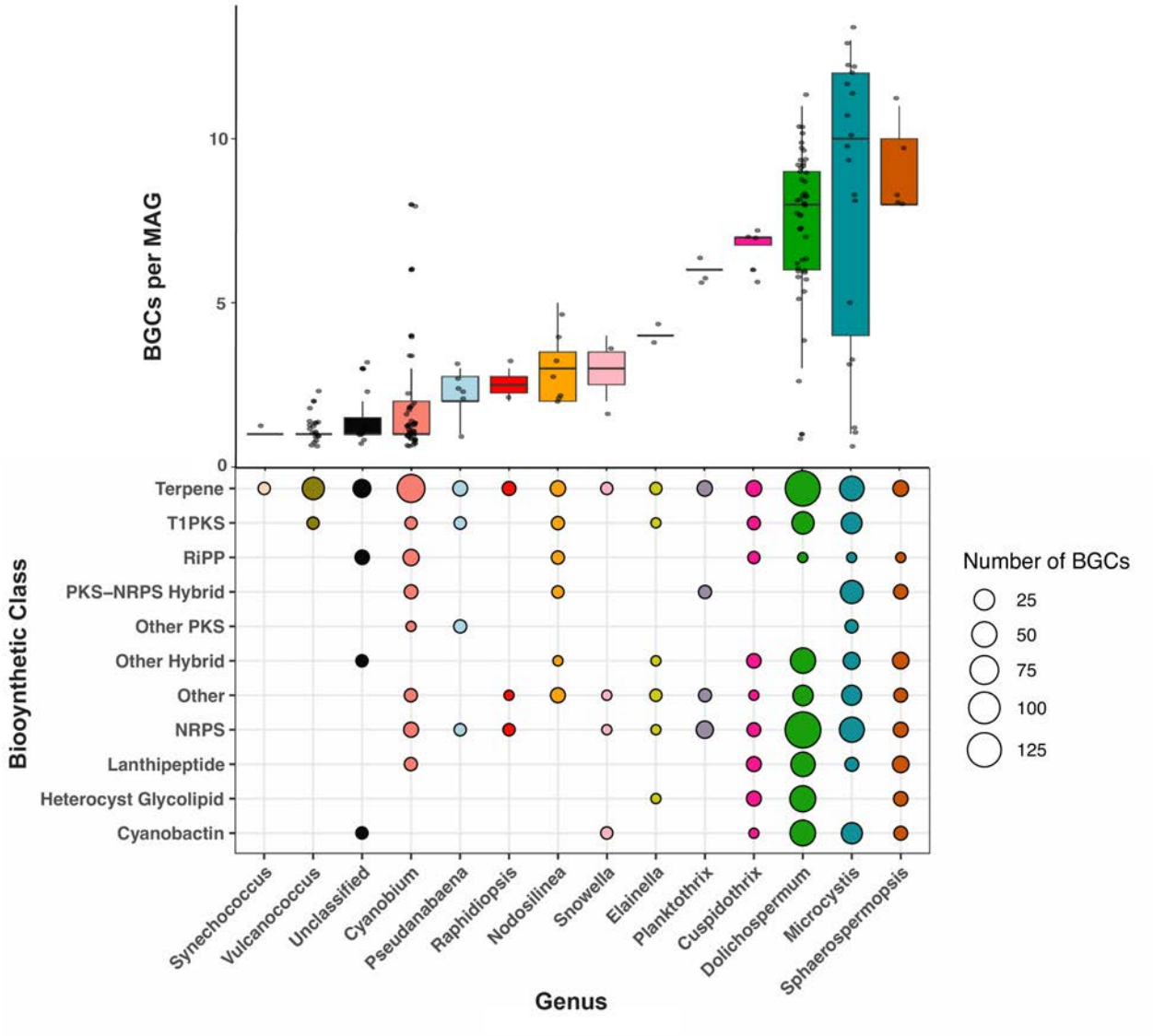


Figure 3.9.8: BGC richness and diversity in dominant cyanobacterial genera from the Gulf. The top panel shows the average number of BGCs per MAG from 14 different cyanobacterial genera. The bottom panel shows the number of BGCs (bubble size) from each biosynthetic class (y-axis) identified from each cyanobacterial genus (x-axis) by antiSMASH v7.0 on the x-axis (102)

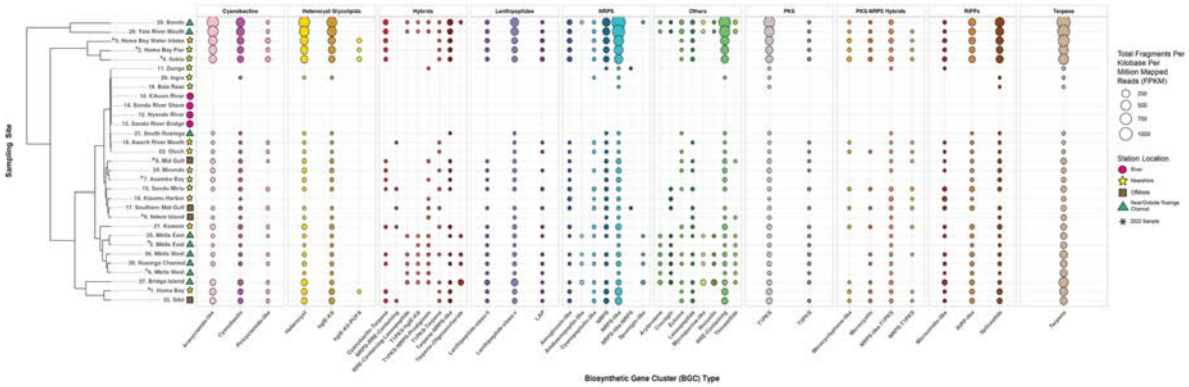


Figure 3.9.9: Total BGC relative abundance across the Winam Gulf. Station numbers, names, and their location indicator symbols are shown on the y-axis, with BGC groups shown on the x-axis. BGCs are grouped into broader biosynthetic or product classifications, labeled as the top facet strips (Table S3). The FPKM of each BGC was calculated (Equation 3.6.2) and the size of a point corresponds to the total FPKM of all BGCs of a particular BGC type indicated on the x-axis. Bubbles are colored to differentiate BGC groups. Samples on the y-axis are hierarchically clustered using the complete-linkage method.

3.10 Supplementary Information

Table S3.10.1: Table of spearman correlation coefficients between highly abundant cyanobacterial genera and associated metadata variables.

Variable	Dolichospermum	Microcystis	Planktothrix	Sphaerospermopsis	Cyanobium	Vulcanococcus
Depth	-0.2615	-0.5692	-0.5560	0.6835	0.4901	0.5604
pH	0.1367	0.2161	-0.0198	0.0595	-0.0970	0.0662
Secchi	-0.4636	-0.6336	-0.5298	0.6314	0.6027	0.5033
Turbidity	0.3036	0.8119	0.6711	-0.8823	-0.3190	-0.1716
Ammonia	0.4489	0.5089	0.3867	-0.5867	-0.4889	-0.4534
Nitrate	0.1622	0.4089	0.3245	-0.5689	-0.2445	-0.2467
Temperature	0.5325	0.5611	0.5457	-0.3564	-0.4708	-0.4378
DO	-0.2527	-0.5165	-0.7978	0.6220	0.2835	0.2835
Conductivity	-0.3630	-0.2398	0.1452	-0.0506	0.2904	0.1716
Dissolved N:P	0.1489	0.3156	0.3200	-0.5156	-0.2311	-0.2422
Cyanos (sonde)	0.4462	0.3231	0.3890	-0.0681	-0.7143	-0.7363
Chlorophyll	0.4681	0.6571	0.5473	-0.3890	-0.7143	-0.6527
Dissolved p	-0.1473	-0.1648	-0.3934	0.2747	0.0637	-0.1341

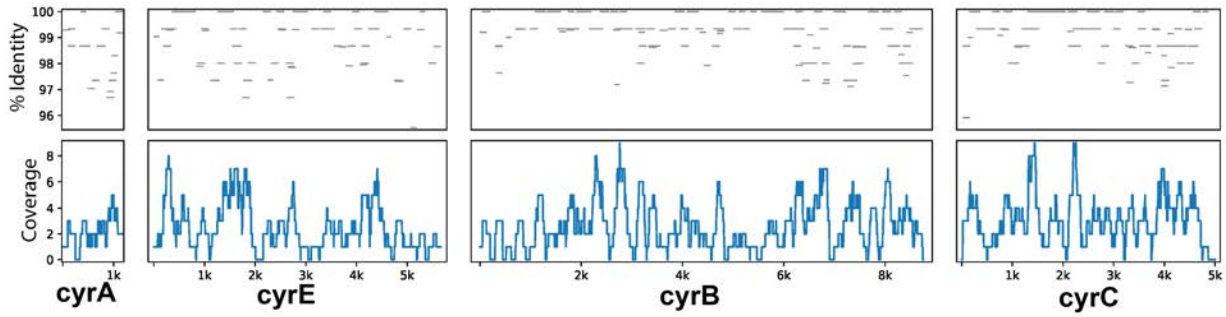


Figure S3.10.1: Read mapping (via blastn) from Site 27. Bridge Island to four genes from the cylindrospermopsin BGC from *Cylindrospermopsis raciborskii AWT205* (MIBiG BGC: BGC0000978). Percent identity of reads mapped to these genes is shown in the top quadrants, and percent coverage by mapped reads to these genes is shown in the bottom quadrant.

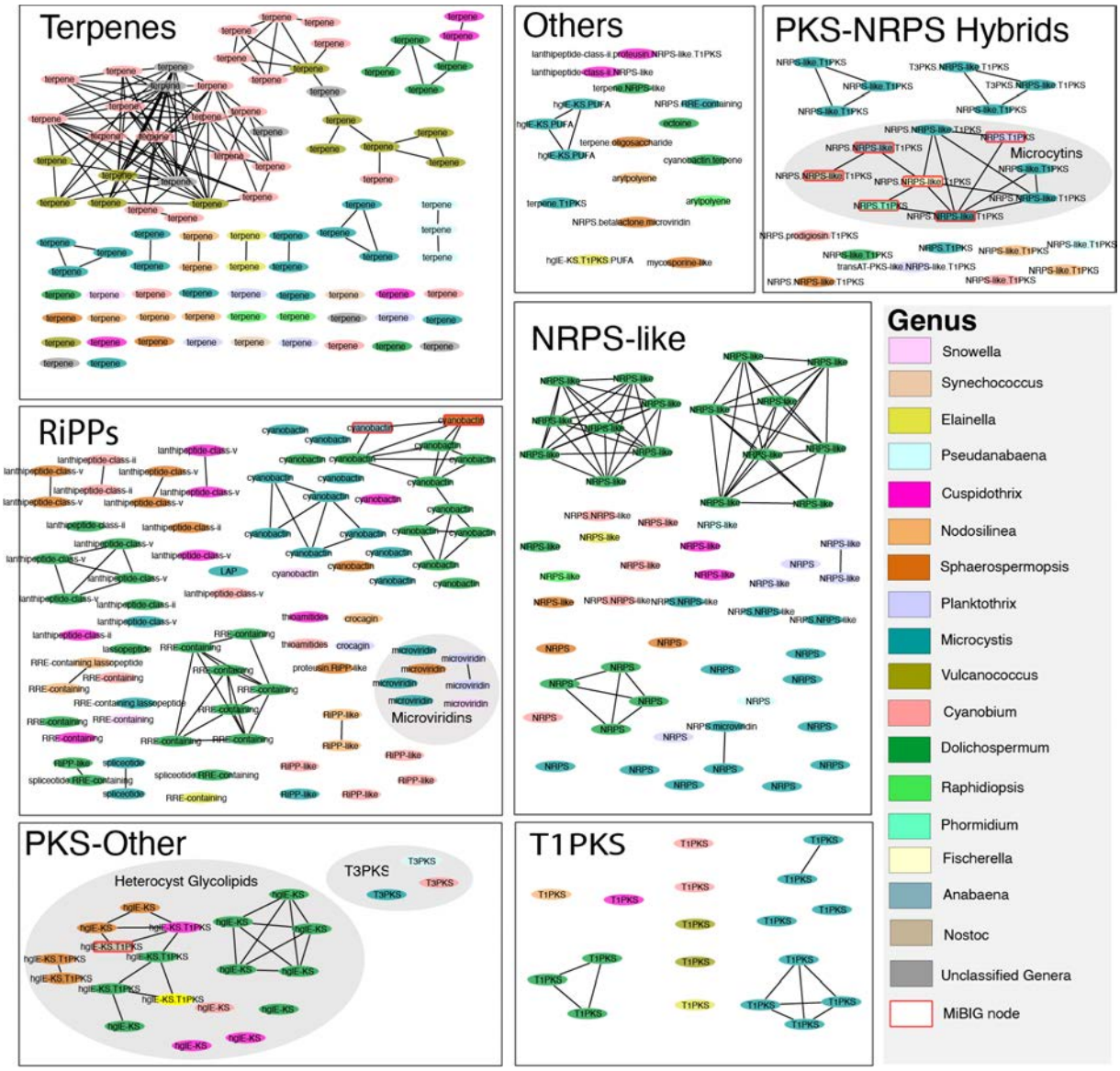


Figure S3.10.2: Gene cluster families (GCFs) for identified BGCs via antiSMASH and BiG-SCAPE. BGCs identified from high quality cyanobacterial MAGs via antiSMASH v7 were manually annotated (SI Table 2) and analyzed via BiG-SCAPE. Nodes are colored based on the cyanobacterial genus of the MAG they originate from (based on GTDB-Tk taxonomic identification). Nodes are labeled based on manual annotation of BGCs. Square nodes with a red outline are from the MiBIG database, linking BGCs to known biosynthesis products. Edges between nodes indicate similarity of BGCs calculated through a distance matrix that considers number of shared PFAM domains, pairs of adjacent PFAM domains, and sequence similarities between protein sequences. Gray circles around nodes indicate the putative synthesis product of those nodes based on antiSMASH annotation and manual inspection of core and additional biosynthesis genes.

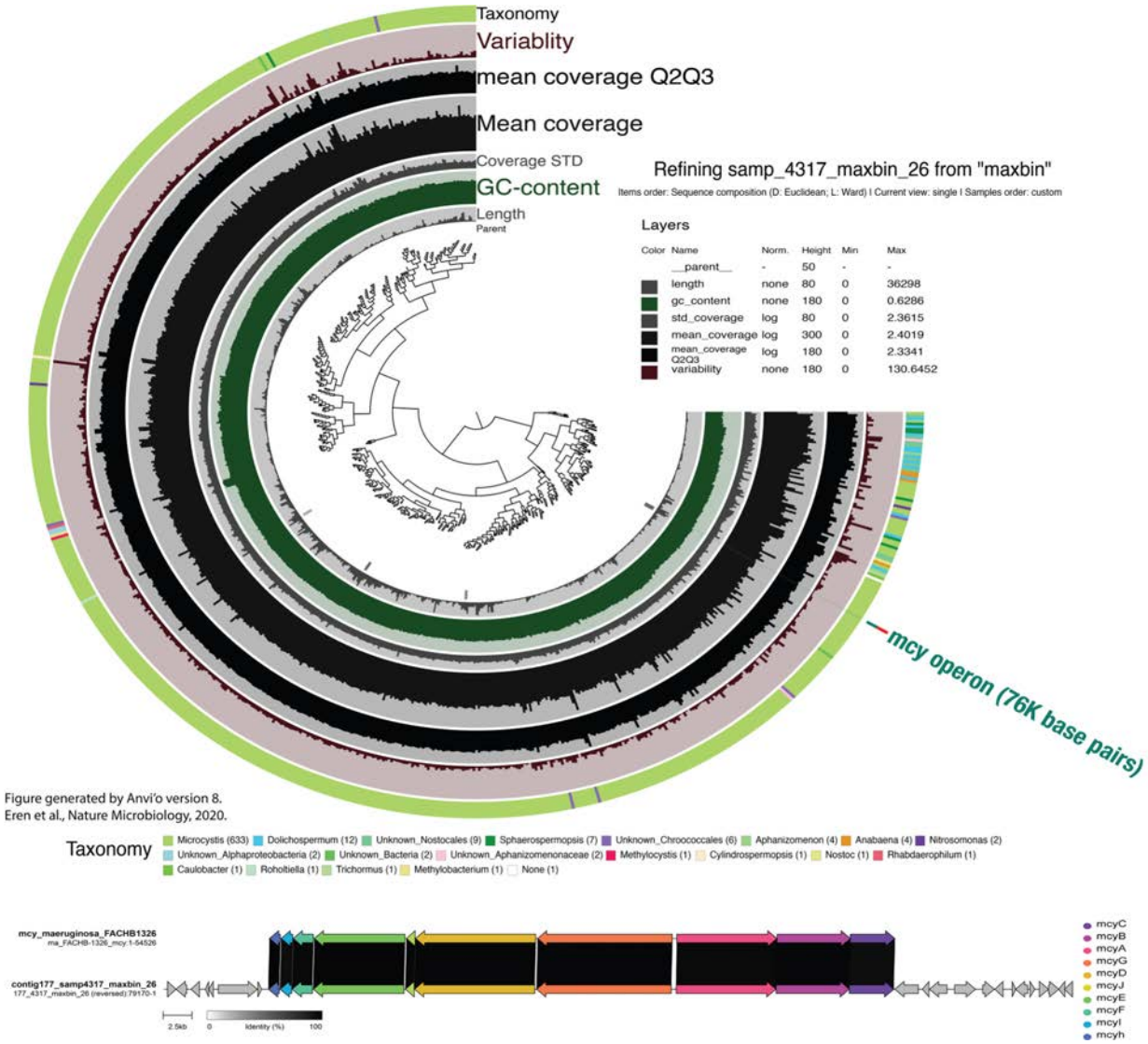


Figure S3.10.3: Genome profile of MAG samp_4317_maxbin_26, annotated as Microcystis sp. (NCBI: SAMN41711332). The contig with the mcy operon is indicated in the Anvi'o MAG profile. Gene similarity image from clinker shows the reference mcy operon from Microcystis aeruginosa FACHB-1326 (GenBank: OQ291092.1) and its close similarity to the mcy operon identified in samp_4317_maxbin_26 on contig 177 below.

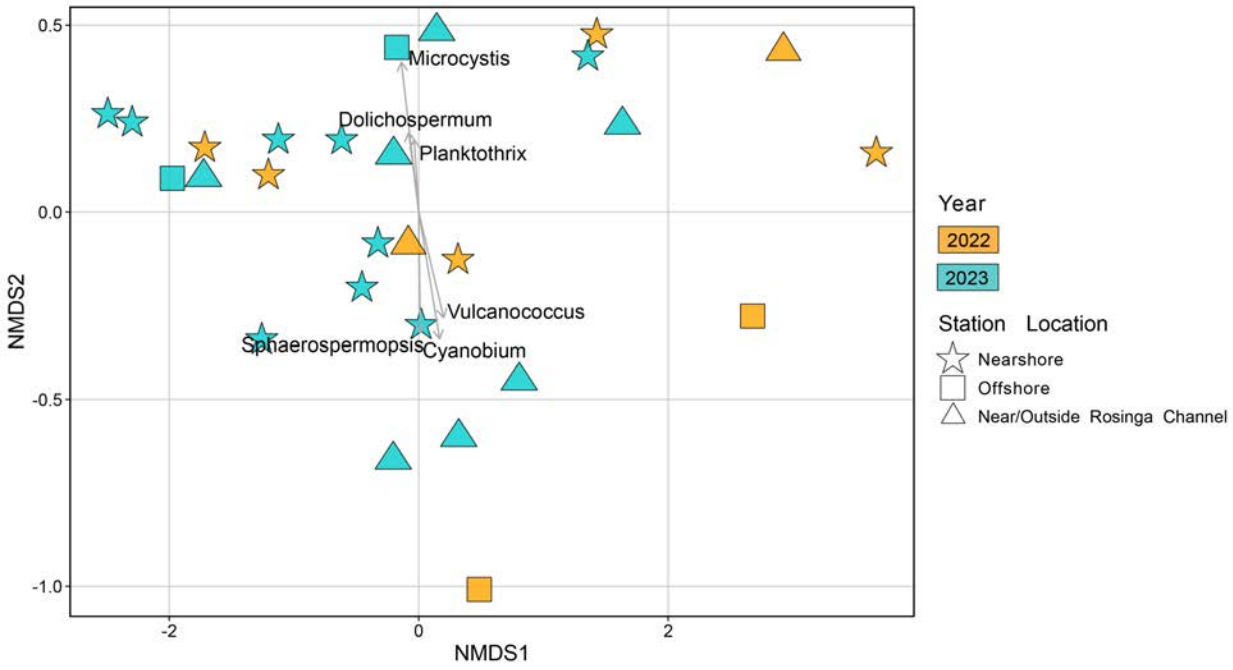


Figure S3.10.4: Biosynthetic potential composition in relation to dominant cyanobacterial genera relative abundances in the Winam Gulf. NMDS analysis of the relative abundance of all identified BGCs from each sample site. Arrows and labels represent the direction and strength of the relationship between the ordination axes and the relative abundance of the top 6 most abundant cyanobacterial genera across the samples. Station location is indicated by the shape of ordination points, and samples from 2022 are represented in amber, while samples from 2023 are in teal. Samples taken from rivers (n=4) were excluded from this ordination because no BGCs were identified in them.

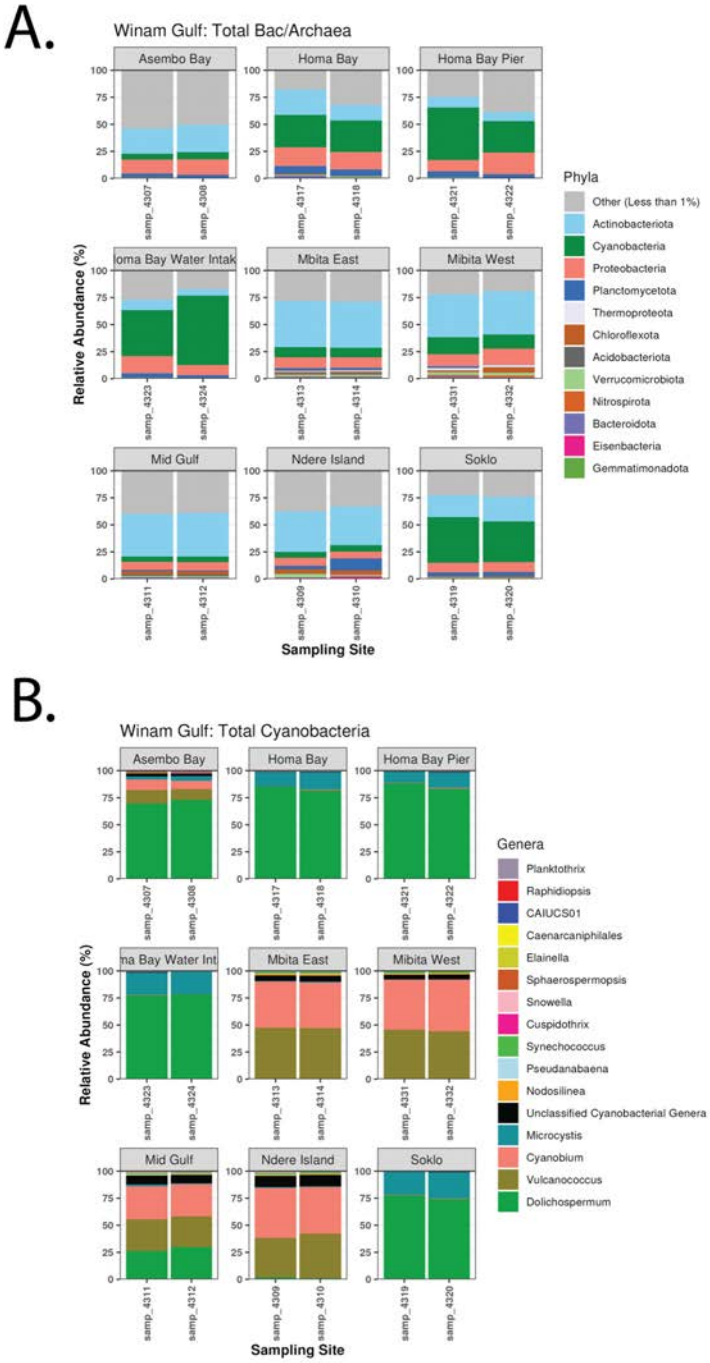


Figure S3.10.5: Microbial and archaeal composition makeup in Winam Gulf, LV. (**A**) Relative abundance (percentage) of bacterial and archaeal phyla from 2022 samples with replicates is shown. The phyla, annotated with GTDB-Tk release 214, representing less than 1 percent of the population were grouped in the "Other" category (87). (**B**) Relative abundance (percentage) of cyanobacterial genera from 2022 samples with replicates is shown. Cyanobacteria with no genus taxonomic assignment from GTDB-Tk were labeled as "unclassified cyanobacteria". Unmapped reads were excluded from this analysis and mapped reads were normalized to 100 percent. Metagenomic samples from the same sampling site are grouped together in a facet.

Table S3.10.2: Reference genomes used for phylogenomic analysis.

Tree	GenBank or NCBI	Name On Tree
All Trees	GCA_000316515.1	*Cyanobium gracile* PCC 6307
Microcystis	GCA_008974145.1	*Microcystis aeruginosa* NIES-4325
Microcystis	GCA_002095975.1	*Microcystis aeruginosa* PCC 7806SL
Microcystis	GCA_007096905.1	*Microcystis wesenbergii* Mw MB S 20031200 S109D
Microcystis	GCA_007096585.1	*Microcystis wesenbergii* Mw MB S 20031200 S109
Microcystis	GCA_007096485.1	*Microcystis novacekii* Mn MB F 20050700 S1D
Microcystis	GCA_007096045.1	*Microcystis novacekii* Mn MB F 20050700 S1
Microcystis	GCA_007096865.1	*Microcystis panniformis* Mp MB F 20051200 S9
Microcystis	GCA_007096905.1	*Microcystis panniformis* Mp MB F 20051200 S6
Microcystis	GCA_009995705.1	*Microcystis aeruginosa* W13-11
Microcystis	GCA_007095855.1	*Microcystis aeruginosa* Ma MB F 20061100 S20D
Microcystis	GCA_007095885.1	*Microcystis aeruginosa* Ma MB F 20061100 S20
Microcystis	GCA_009996085.1	*Microcystis aeruginosa* L211-07
Microcystis	SAMN41711483	27. Bridge Island (*Microcystis aeruginosa* A)
Microcystis	SAMN41711474	26. Mbita West (*Microcystis panniformis* A)
Microcystis	GCA_001264245.1	*Microcystis panniformis* FACHB-1757

Microcystis	GCA_000312225.1	*Microcystis aeruginosa* PCC 9807
Microcystis	GCA_000312185.1	*Microcystis aeruginosa* PCC 9443
Microcystis	GCA_027531875.1	*Microcystis* sp. WLECC 19-84.1
Microcystis	GCA_009996455.1	*Microcystis aeruginosa* F13-15
Microcystis	SAMN41711324	5. Mbita East (*Microcystis panniformis* A)
Microcystis	SAMN41711408	1. Homa Bay (*Microcystis* sp.)
Microcystis	SAMN41711381	2. Homa Bay Pier (*Microcystis panniformis* A)
ADA	GCA_009712085.1	*Anabaena* sp. UHCC 0253
ADA	GCA_017355725.1	*Anabaena* sp. 54
ADA	GCA_018447775.1	*Dolichospermum* sp. WA123
ADA	GCA_017355625.1	*Dolichospermum* sp. DET73
ADA	Den Uyl et al. 2025	*Dolichospermum* sp. Lake Erie 2020 WE8
ADA	GCA_017346855.1	*Aphanizomenon flos-aquae* DEX188
ADA	GCA_001593825.2	*Aphanizomenon flos-aquae* UKL13-PB
ADA	GCA_009712075.2	*Dolichospermum planctonicum* UHCC 0167
ADA	GCF_005402965.1	*Dolichospermum planctonicum* NIES-80
ADA	GCA_001672075.1	*Anabaena* sp. CRKS33
ADA	GCA_000426905.1	*Dolichospermum circinale* AWQC131C
ADA	GCA_000426925.1	*Dolichospermum circinale* AWQC310F
ADA	Den Uyl et al. 2025	*Dolichospermum* sp. Lake Erie 2022 WE12
ADA	Den Uyl et al. 2025	*Dolichospermum* sp. Lake Erie 2017 WE12
ADA	SAMN41711487	27. Bridge Island (*Dolichospermum* sp.)
ADA	SAMN41711470	25. Mbita East (*Dolichospermum* sp.)

ADA	SAMN41711489	28. Bondo (*Dolichospermum* sp.)
ADA	SAMN41711415	1. Homa Bay (*Dolichospermum* sp.)
ADA	SAMN41711477	26. Mbita West (*Dolichospermum* sp.)
Planktothrix	GCA_900010725.2	*Planktothrix serta* PCC 8927
Planktothrix	GCA_015207485.1	*Planktothrix mougeoti* LEGE 06226
Planktothrix	GCA_900009145.1	*Planktothrix tepida* PCC 9214
Planktothrix	SAMN41711432	15. Sondu Miriu (*Planktothrix* spp.)
Planktothrix	SAMN41711440	17. Southern Mid-Gulf (*Planktothrix* spp.)
Planktothrix	SAMN41711420	10. Kisumu Harbor (*Planktothrix* spp.)
Planktothrix	GCA_900009265.2	*Planktothrix paucivesiculata* PCC 9631
Planktothrix	GCA_963898545.1	*Planktothrix* sp. PCC 11201
Planktothrix	GCA_904830915.1	*Planktothrix agardhii* PCC 7805
Planktothrix	GCA_021554855.1	*Planktothrix agardhii* 1804
Planktothrix	GCA_904830885.1	*Planktothrix agardhii* PCC 7811
Planktothrix	GCA_904830765.1	*Planktothrix agardhii* NIVA-CYA 126/8
Planktothrix	GCA_021554935.1	*Planktothrix agardhii* 1029
Planktothrix	GCA_000464765.1	*Planktothrix prolifica* NIVA-CYA 406
Planktothrix	GCA_904830895.1	*Planktothrix rubescens* PCC 7821
Planktothrix	GCA_000464845.1	*Planktothrix prolifica* NIVA-CYA 98
Planktothrix	GCA_000464785.1	*Planktothrix rubescens* NIVA-CYA 407
Planktothrix	GCA_000464825.1	*Planktothrix agardhii* NIVA-CYA 56/3
Planktothrix	GCA_000464665.1	*Planktothrix agardhii* NIVA-CYA 15
Planktothrix	GCA_000464805.1	*Planktothrix prolifica* NIVA-CYA 540

Planktothrix	GCA_000464725.1	*Planktothrix agardhii* NIVA-CYA 34
Spaerospermopsis	GCA_019598945.1	*Sphaerospermopsis torques-reginae* ITEP-024
Spaerospermopsis	GCF_014299995.1	*Sphaerospermopsis* spp. LEGE 00249
Spaerospermopsis	GCF_015207205.1	*Sphaerospermopsis* spp. LEGE 08334
Spaerospermopsis	SAMN41711478	26. Mbita West (*Sphaerospermopsis kisseleviana* A)
Spaerospermopsis	SAMN41711493	30. Rosinga Channel (*Sphaerospermopsis kisseleviana* A)
Spaerospermopsis	SAMN41711482	27. Bridge Island (*Sphaerospermopsis kisseleviana* A)
Spaerospermopsis	SAMN41711466	25. Mbita East (*Sphaerospermopsis kisseleviana* A)
Spaerospermopsis	GCF_028329225.1	*Sphaerospermopsis kisseleviana* CS-549
Spaerospermopsis	Den Uyl et al. 2025	Lake Erie *Sphaerospermopsis kisseleviana* A
Spaerospermopsis	GCF_015207835.1	*Sphaerospermopsis aphanizomenoides* LEGE 00250
Spaerospermopsis	GCF_014696195.1	*Sphaerospermopsis* spp. FACHB-1094
Spaerospermopsis	GCF_014696235.1	*Sphaerospermopsis* spp. FACHB-1194
Spaerospermopsis	GCF_005402885.1	*Sphaerospermopsis reniformis* NIES-1949

Table S3.10.3: Manual annotation notes for dereplicated set of BGCs identified from high quality cyanobacterial MAGs from the sampled sites in Winam Gulf.

AntiSMASH Region ID	Region ID	Pathway	Length	Manual Annotated Descriptor	Start	End
286	samp_4313_maxbin_367_contig_143_T1PKS	T1PKS	9,649	1-heptadecene-like 1	2738	8005
32	samp_4317_concoct_68_contig_312_NRPS	NRPS	52,073	Aeruginosin-like 1	18821	43020
147	samp_4435_maxbin_422_contig_111_NRPS-like	NRPS-like	9,281	Aeruginosin-like 2	3300	9211
59	samp_4319_VAMB_955_contig_421_NRPS	NRPS,NRPS-like	39,125	Aeruginosin-like 3	1783	26466
282	samp_4322_metabat2_24_contig_109_NRPS-like	NRPS-like,NRPS	24,156	Aeruginosin-like 4	2	11716
231	samp_4445_concoct_1415_contig_72_NRPS	NRPS,betalactone,microviridin	39,273	Anabaenopeptin-like 1	2063	31577

50	samp_4321_maxbin_137_contig_1_cyanobactin	cyanobactin	7,410	Anacyclamide-like 1	41	6983
110	samp_4311_concoct_412_contig_314_cyanobactin	cyanobactin	16,985	Anacyclamide-like 4	1002	16941
125	samp_4447_metabat2_48_contig_296_cyanobactin	cyanobactin	11,671	Anacyclamide-like 5	2	5627
130	samp_4307_semibin_116_contig_228_cyanobactin	cyanobactin	14,564	Anacyclamide-like 6	6036	14222
239	samp_4322_concoct_184_contig_61_cyanobactin	cyanobactin	12,726	Anacyclamide-like 7	6036	11975
307	samp_4439_concoct_171_contig_184_cyanobactin	cyanobactin	19,842	Anacyclamide-like 8	1002	19843
81	samp_4445_concoct_1447_contig_530_arylpolyene	arylpolyene	7,501	Arylprolene 1	31	6029
139	samp_4446_VAMB_96_contig_914_arylpolyene	arylpolyene	6,570	Arylprolene 2	2230	6095

107	samp_4434_semibin_67_contig_284_crocagin	crocagin	16,312	Crocagin 1	1124	1593
277	samp_4444_metabat2_185_contig_143_crocagin	crocagin	11,426	Crocagin 2	54	10914
14.1	samp_4435_concoct_1904_contig_427_cyanobactin	cyanobactin	6,030	Cyanobactin 1	2376	6029
16.1	samp_4435_concoct_1904_contig_665_cyanobactin	cyanobactin	9,016	Cyanobactin 2	2	1582
76	samp_4434_VAMB_405_contig_1902_cyanobactin	cyanobactin	11,968	Cyanobactin 3	2	2388
85	samp_4446_metadecoder_636_contig_262_cyanobactin	cyanobactin	13,091	Cyanobactin 4	238	3520
92	samp_4319_maxbin_413_contig_252_cyanobactin	cyanobactin	12,310	Cyanobactin 5	197	6688
89	samp_4446_VAMB_97_contig_192_cyanobactin	cyanobactin,terpene	15,102	Cyanobactin-Terpene 1	5383	15009
180	samp_4434_metabat2_58_contig_120_NRPS-like	NRPS	9,595	Cyanopeptolin-like 1	3	9595
210	samp_4447_concoct_1018_contig_252_NRPS	NRPS	11,577	Cyanopeptolin-like 2	2	11424

184	samp_4434_metabat2_58_contig_530_NRPS-like	NRPS-like,NRPS	7,931	Cyanopeptolin-like 3	3	7932
88	samp_4446_VAMB_97_contig_68_ectoine	ectoine	8,340	Ectoine 1	3	7589
3	samp_4325_concoct_153_contig_525_hglE-KS	hglE-KS	41,480	Heterocyst Glycolipids 1	12604	40311
145	samp_4315_VAMB_280_contig_117_PUFA	PUFA,hglE-KS	23,844	Heterocyst Glycolipids 10	6950	23845
175	samp_4324_maxbin_53_contig_213_T1PKS	T1PKS,hglE-KS	27,185	Heterocyst Glycolipids 11	16683	26592
31	samp_4317_concoct_68_contig_274_hglE-KS	hglE-KS	35,844	Heterocyst Glycolipids 2	10938	23244

41	samp_4447_VAMB_472_contig_107_hglE-KS	hglE-KS	43,153	Heterocyst Glycolipids 3	16996	30197
64	samp_4441_maxbin_669_contig_29_hglE-KS	hglE-KS	30,794	Heterocyst Glycolipids 4	5888	18194
83	samp_4446_metadecoder_636_contig_129_hglE-KS	hglE-KS	32,272	Heterocyst Glycolipids 5	3	26463
86	samp_4312_concoct_454_contig_57_hglE-KS	hglE-KS	35,194	Heterocyst Glycolipids 6	8159	34594
119	samp_4448_VAMB_839_contig_2366_hglE-KS	hglE-KS	13,897	Heterocyst Glycolipids 7	2789	11716
255	samp_4443_semibin_319_contig_514_hglE-KS	hglE-KS	26,767	Heterocyst	3658	20242

				Glycolipids 8		
98	samp_4324_maxbin_57_contig_260_PUFA	PUFA,hglE-KS	25,891	Heterocyst Glycolipids 9	7964	25219
169	samp_4443_semibin_480_contig_250_hglE-KS	hglE-KS	13,416	hglE-KS 1	3	12147
173	samp_4324_maxbin_53_contig_124_hglE-KS	hglE-KS	35,259	hglE-KS 2	1770	33031
262	samp_4449_metabat2_97_contig_255_hglE-KS	hglE-KS	20,928	hglE-KS 3	3431	8572
58	samp_4319_VAMB_955_contig_274_PUFA	hglE-KS,PUFA	21,390	hglE-KS-PUFA 1	5712	20718
4	samp_4319_concoct_522_contig_1241_lanthipeptide-class-ii	lanthipeptide-class-ii	5,637	Lanthipeptide-class-ii 1	95	4232
191	samp_4321_concoct_295_contig_693_lanthipeptide-class-ii	lanthipeptide-class-ii	20,973	Lanthipeptide-class-ii 2	7627	10974
215	samp_4445_metabat2_158_contig_277_lanthipeptide-class-ii	lanthipeptide-class-ii	15,548	Lanthipeptide-class-ii 3	9658	13080

264	samp_4320_concoct_456_contig_168_lanthipeptide-class-ii	lanthipeptide-class-ii	5,002	Lanthipeptide-class-ii 4	515	3868
265	samp_4443_metabat2_135_contig_62_lanthipeptide-class-ii	lanthipeptide-class-ii	10,677	Lanthipeptide-class-ii 5	5738	9160
294	samp_4443_VAMB_834_contig_1768_lanthipeptide-class-ii	lanthipeptide-class-ii	5,429	Lanthipeptide-class-ii 6	265	3561
297	samp_4445_semibin_421_contig_114_lanthipeptide-class-ii	lanthipeptide-class-ii	13,458	Lanthipeptide-class-ii 7	10002	13184
5	samp_4447_maxbin_94_contig_103_lanthipeptide-class-v	lanthipeptide-class-v	7,430	Lanthipeptide-class-v 1	1400	7429
251	samp_4443_semibin_319_contig_215_lanthipeptide-class-v	lanthipeptide-class-v	11,116	Lanthipeptide-class-v 10	7965	10351
254	samp_4443_semibin_319_contig_425_lanthipeptide-class-v	lanthipeptide-class-v	30,784	Lanthipeptide-class-v 11	1345	15775

267	samp_4443_metabat2_135_contig_457_lanthipeptide-class-v	lanthipeptide-class-v	15,031	Lanthipeptide-class-v 12	2167	3953
295	samp_4445_semibin_421_contig_72_lanthipeptide-class-v	lanthipeptide-class-v	23,974	Lanthipeptide-class-v 13	19857	22339
11	samp_4435_concoct_1904_contig_96_lanthipeptide-class-v	lanthipeptide-class-v	6,277	Lanthipeptide-class-v 2	3718	6015
33	samp_4317_concoct_68_contig_348_lanthipeptide-class-v	lanthipeptide-class-v	35,821	Lanthipeptide-class-v 3	13423	14584
40	samp_4447_VAMB_472_contig_97_lanthipeptide-class-v	lanthipeptide-class-v	12,847	Lanthipeptide-class-v 4	9022	11499
75	samp_4434_VAMB_405_contig_1469_lanthipeptide-class-v	lanthipeptide-class-v	29,320	Lanthipeptide-class-v 5	6922	9321
113	samp_4448_VAMB_839_contig_384_lanthipeptide-class-v	lanthipeptide-class-v	28,727	Lanthipeptide-class-v 6	6298	8728

115	samp_4448_VAMB_839_contig_614_lanthipeptide-class-v	lanthipeptide-class-v	15,379	Lanthipeptide-class-v 7	767	3386
202	samp_4444_semibin_569_contig_43_lanthipeptide-class-v	lanthipeptide-class-v	29,238	Lanthipeptide-class-v 8	12887	15286
246	samp_4326_metabat2_21_contig_28_lanthipeptide-class-v	lanthipeptide-class-v	35,285	Lanthipeptide-class-v 9	12887	15286
270	samp_4317_maxbin_26_contig_153_LAP	LAP	10,015	LAP 1	231	2528
57	samp_4319_VAMB_955_contig_191_lassopeptide	lassopeptide	17,103	Lassopeptide 1	10002	14367
193	samp_4441_concoct_710_contig_258_lassopeptide	lassopeptide	5,836	Lassopeptide 2	2082	4679
133	samp_4447_VAMB_322_contig_258_lassopeptide	lassopeptide,RRE-containing	6,530	Lassopeptide 3	732	5305
189	samp_4321_concoct_295_contig_230_T1PKS	T1PKS,NRPS-like,T3PKS	39,886	Merocyclophane-like 1	6828	24205

62	samp_4319_VAMB_955_contig_842_T1PKS	T1PKS,T3PKS,NRPS-like	36,987	Merocyclophane-like 2	1027 4	27651
55	samp_4319_VAMB_955_contig_178_NRPS	NRPS,T1PKS,NRPS-like	76,737	Microcystin 1	2284	61400
271	samp_4317_maxbin_26_contig_177_NRPS	NRPS,T1PKS,NRPS-like	79,169	Microcystin 2	1096 8	70981
15	samp_4435_concoct_1904_contig_477_microviridin	microviridin	6,000	Microviridin-like 1	995	5360
21	samp_4435_concoct_1904_contig_1372_microviridin	microviridin	6,540	Microviridin-like 2	197	2243
28	samp_4444_metabat2_75_contig_746_microviridin	microviridin	5,407	Microviridin-like 3	3	5094
101	samp_4434_semibin_67_contig_107_microviridin	microviridin	10,330	Microviridin-like 4	856	4129
150	samp_4435_maxbin_422_contig_185_microviridin	microviridin	10,750	Microviridin-like 5	611	2905
269	samp_4317_maxbin_26_contig_19_microviridin	microviridin	4,999	Microviridin-like 6	856	3347
292	samp_4443_VAMB_834_contig_202_microviridin	microviridin	7,634	Microviridin-like 7	1	7635

308	samp_4434_semibin_67_contig_9_microviridin	microviridin	10,365	Microviridin-like 8	7841	10142
118	samp_4448_VAMB_839_contig_2328_mycosporine-like	mycosporine-like	20,029	Mycosporine-like 1	190	12468
253	samp_4443_semibin_319_contig_369_T1PKS	T1PKS,NRPS,NRPS-like	43,243	Nocuolin 1	1	43069
13.1	samp_4435_concoct_1904_contig_197_NRPS	NRPS	6,471	NRPS 1	2	6472
303	samp_4321_metadecoder_473_contig_307_NRPS	NRPS	42,552	NRPS 10	525	24094
24.1	samp_4435_concoct_1904_contig_2230_NRPS	NRPS	7,761	NRPS 2	1	7761
151	samp_4435_maxbin_422_contig_336_NRPS	NRPS	9,638	NRPS 3	2	9483
156	samp_4435_maxbin_440_contig_738_NRPS	NRPS	15,411	NRPS 4	1	15412
157	samp_4445_VAMB_917_contig_2786_NRPS	NRPS	7,328	NRPS 5	3	7329
158	samp_4445_VAMB_917_contig_7332_NRPS	NRPS	19,778	NRPS 6	3737	17544
174	samp_4324_maxbin_53_contig_182_NRPS	NRPS	42,852	NRPS 7	2093	35976

181	samp_4434_metabat2_58_contig_186_NRPS-like	NRPS	8,426	NRPS 8	2	8425
243	samp_4322_concoct_184_contig_432_NRPS	NRPS	53,470	NRPS 9	17553	41122
2	samp_4325_concoct_153_contig_28_NRPS-like	NRPS-like	23,373	NRPS-like 1	636	16795
82	samp_4446_metadecoder_636_contig_120_NRPS-like	NRPS-like	11,598	NRPS-like 10	1495	11599
87	samp_4312_concoct_454_contig_249_NRPS-like	NRPS-like	38,008	NRPS-like 11	2254	35887
106	samp_4434_semibin_67_contig_203_NRPS-like	NRPS-like	13,035	NRPS-like 12	2430	8722
111	samp_4311_concoct_412_contig_530_NRPS-like	NRPS-like	37,709	NRPS-like 13	550	31687
123	samp_4443_VAMB_823_contig_1189_NRPS-like	NRPS-like	20,262	NRPS-like 14	8304	20262
129	samp_4448_metadecoder_579_contig_113_NRPS-like	NRPS-like	37,929	NRPS-like 15	129	29411
149	samp_4435_maxbin_422_contig_182_NRPS-like	NRPS-like	13,486	NRPS-like 16	97	6389

159	samp_4323_semibin_61_contig_74_NRPS-like	NRPS-like	30,352	NRPS-like 17	12576	27698
160	samp_4323_semibin_61_contig_81_NRPS-like	NRPS-like	44,251	NRPS-like 18	10503	43704
165	samp_4320_metadecoder_702_contig_55_NRPS-like	NRPS-like	30,034	NRPS-like 19	11356	23292
6	samp_4447_maxbin_94_contig_134_NRPS-like	NRPS-like	9,572	NRPS-like 2	6808	9462
177	samp_4324_maxbin_53_contig_388_NRPS-like	NRPS-like	23,563	NRPS-like 20	826	17036
205	samp_4444_semibin_569_contig_375_NRPS-like	NRPS-like	12,862	NRPS-like 21	40	11746
238	samp_4322_concoct_184_contig_1_NRPS-like	NRPS-like	42,007	NRPS-like 22	8615	41528
242	samp_4322_concoct_184_contig_336_NRPS-like	NRPS-like	29,014	NRPS-like 23	2446	27080
290	samp_4448_concoct_1792_contig_649_NRPS-like	NRPS-like	16,566	NRPS-like 24	2178	16105
47.1	samp_4445_maxbin_221_contig_353_NRPS-like	NRPS-like,NRPS	11,957	NRPS-like 25	2688	10543

275	samp_4320_semibin_145_contig_392_NRPS	NRPS,microviridin	30,358	NRPS-like 26	6864	25371
8	samp_4435_semibin_217_contig_357_NRPS-like	NRPS-like	14,645	NRPS-like 3	2	2585
30	samp_4317_concoct_68_contig_195_NRPS-like	NRPS-like	42,981	NRPS-like 4	1914	39348
45	samp_4447_VAMB_472_contig_513_NRPS-like	NRPS-like	34,068	NRPS-like 5	127	31982
65	samp_4441_maxbin_669_contig_134_NRPS-like	NRPS-like	44,251	NRPS-like 6	3019	41270
73	samp_4434_VAMB_405_contig_1043_NRPS-like	NRPS-like	42,738	NRPS-like 7	11313	42739
77	samp_4434_VAMB_405_contig_2570_NRPS-like	NRPS-like	41,091	NRPS-like 8	8475	40470
80	samp_4445_concoct_1447_contig_373_NRPS-like	NRPS-like	9,478	NRPS-like 9	4432	9479
7	samp_4435_semibin_217_contig_325_NRPS-like	NRPS,NRPS-like	10,812	NRPS-like-NRPS Hybrid 1	3	10813

1	samp_4444_semibin_567_contig_259_T1PKS	T1PKS,NRPS-like	27,405	NRPS-like-T1PKS 1	1579 8	22205
219	samp_4445_metabat2_158_contig_347_T1PKS	NRPS-like,T1PKS	48,621	NRPS-like-T1PKS 10	8036	37332
273	samp_4317_maxbin_26_contig_563_NRPS-like	NRPS-like,T1PKS	18,054	NRPS-like-T1PKS 11	3656	15831
90	samp_4446_VAMB_97_contig_859_NRPS-like	T1PKS,NRPS-like	10,111	NRPS-like-T1PKS 2	268	10111
187	samp_4321_concoct_295_contig_143_NRPS-like	T1PKS,NRPS-like	17,701	NRPS-like-T1PKS 3	4291	16462
211	samp_4447_concoct_1018_contig_313_T1PKS	T1PKS,NRPS-like	20,620	NRPS-like-T1PKS 4	2	14444
296	samp_4445_semibin_421_contig_112_T1PKS	T1PKS,NRPS-like,lanthipeptide-class-ii,proteusin	31,052	NRPS-like-T1PKS 5	1839 9	31053

105	samp_4434_semibin_67_contig_183_T1PKS	T1PKS,transAT-PKS-like,NRPS-like	32,853	NRPS-like-T1PKS 6	3455	32623
19	samp_4435_concoct_1904_contig_1012_NRPS-like	NRPS-like,T1PKS	17,586	NRPS-like-T1PKS 7	5829	17585
69	samp_4445_semibin_342_contig_369_T1PKS	NRPS-like,T1PKS	16,133	NRPS-like-T1PKS 8	1042	11439
194	samp_4320_metadecoder_628_contig_90_NRPS-like	NRPS-like,T1PKS	23,217	NRPS-like-T1PKS 9	472	20998
182	samp_4434_metabat2_58_contig_297_NRPS	NRPS,RRE-containing	19,972	NRPS-RRE-Containing 1	2074	19647
18	samp_4435_concoct_1904_contig_899_T1PKS	NRPS,T1PKS	19,856	NRPS-T1PKS 1	2908	19532
63	samp_4319_VAMB_955_contig_876_cyanobactin	cyanobactin	13,696	Piricyclamide-like 1	1490	13697

97	samp_4324_maxbin_57_contig_142_cyanobactin	cyanobactin	7,630	Piricyclamide-like 2	10 3	359 4
99	samp_4324_maxbin_57_contig_424_cyanobactin	cyanobactin	14,175	Piricyclamide-like 3	24 5	549 1
135	samp_4447_VAMB_322_contig_935_cyanobactin	cyanobactin	10,753	Piricyclamide-like 4	34 39	107 16
190	samp_4321_concoct_295_contig_676_cyanobactin	cyanobactin	11,922	Piricyclamide-like 5	20 71	111 24
196	samp_4320_metadecoder_628_contig_277_cyanobactin	cyanobactin	10,866	Piricyclamide-like 6	3	832 5
17	samp_4435_concoct_1904_contig_753_RiPP-like	RiPP-like	5,343	RiPP-like 1	2	487 6
304	samp_4445_semibin_366_contig_44_RiPP-like	RiPP-like	6,633	RiPP-like 10	10 19	617 1
136	samp_4445_VAMB_519_contig_64_RiPP-like	RiPP-like	10,150	RiPP-like 2	29 25	603 3
153	samp_4435_maxbin_440_contig_82_RiPP-like	RiPP-like	7,495	RiPP-like 3	64 4	749 6

172	samp_4443_maxbin_277_contig_222_RiPP-like	RiPP-like	12,659	RiPP-like 4	4499	12659
221	samp_4445_metabat2_158_contig_426_RiPP-like	RiPP-like	5,844	RiPP-like 5	458	4603
229	samp_4445_metabat2_46_contig_296_RiPP-like	RiPP-like	6,450	RiPP-like 6	2537	4664
240	samp_4322_concoct_184_contig_235_RiPP-like	RiPP-like	6,580	RiPP-like 7	261	4684
259	samp_4444_metadecoder_351_contig_165_RiPP-like	RiPP-like	9,640	RiPP-like 8	3493	9155
278	samp_4444_metabat2_185_contig_184_RiPP-like	RiPP-like	8,597	RiPP-like 9	1058	6345
233	samp_4445_concoct_1415_contig_928_proteusin	proteusin,RiPP-like	13,963	RiPP-like-Proteusin	433	11882
27	samp_4444_metabat2_75_contig_553_RRE-containing	RRE-containing	8,650	RRE-Containing 1	396	8156
281	samp_4444_metabat2_185_contig_588_RRE-containing	RRE-containing	7,935	RRE-Containing 10	584	7804

300	samp_4445_semibin_421_contig_290_RRE-containing	RRE-containing	13,918	RRE-Containing 11	6548	13809
306	samp_4439_concoct_171_contig_155_RRE-containing	RRE-containing	20,449	RRE-Containing 12	3489	12101
29	samp_4317_concoct_68_contig_97_RRE-containing	RRE-containing	15,272	RRE-Containing 2	3564	15012
49	samp_4323_metabat2_108_contig_203_RRE-containing	RRE-containing	6,879	RRE-Containing 3	1324	6279
66	samp_4441_maxbin_669_contig_191_RRE-containing	RRE-containing	14,104	RRE-Containing 4	709	12149
84	samp_4446_metadecoder_636_contig_230_RRE-containing	RRE-containing	12,601	RRE-Containing 5	4060	12149
91	samp_4319_maxbin_413_contig_136_RRE-containing	RRE-containing	9,906	RRE-Containing 6	2891	9599

126	samp_4447_metabat2_48_contig_396_RRE-containing	RRE-containing	9,920	RRE-Containing 7	1138	7798
204	samp_4444_semibin_569_contig_329_RRE-containing	RRE-containing	11,369	RRE-Containing 8	679	11006
235	samp_4445_semibin_379_contig_460_RRE-containing	RRE-containing	15,237	RRE-Containing 9	4921	11350
171	samp_4443_maxbin_277_contig_110_RRE-containing	RRE-containing,lassopeptide	10,487	RRE-Containing-Lassopeptide 1	1027	10123
188	samp_4321_concoct_295_contig_156_spliceotide	spliceotide	16,462	Spliceotide 1	370	15938
141	samp_4308_VAMB_457_contig_1077_spliceotide	spliceotide,RRE-containing	13,000	Spliceotide 3	2	12451
247	samp_4326_metabat2_21_contig_315_spliceotide	spliceotide,RRE-containing	6,320	Spliceotide 4	742	6321
285	samp_4322_metabat2_24_contig_394_spliceotide	spliceotide	13,571	Splideotide 2	32	13170

232	samp_4445_concoct_1415_contig_907_NRPS-like	NRPS-like	7,289	Spumigin-like 1	1500	7289
61	samp_4319_VAMB_955_contig_507_T1PKS	T1PKS	24,287	T1PKS 1	4557	22260
284	samp_4322_metabat2_24_contig_185_T1PKS	T1PKS	25,243	T1PKS 10	3217	20824
120	samp_4435_VAMB_805_contig_117_T1PKS	T1PKS	13,791	T1PKS 2	698	12892
128	samp_4448_metadecoder_579_contig_38_T1PKS	T1PKS	32,899	T1PKS 3	3707	13687
140	samp_4308_VAMB_457_contig_88_T1PKS	T1PKS	30,254	T1PKS 4	20002	29978
213	samp_4447_concoct_1018_contig_1216_T1PKS	T1PKS	15,490	T1PKS 5	409	14577
263	samp_4449_metabat2_97_contig_277_T1PKS	T1PKS	8,484	T1PKS 6	1525	8385
268	samp_4443_metabat2_135_contig_520_T1PKS	T1PKS	26,563	T1PKS 7	5625	23185
274	samp_4317_maxbin_26_contig_612_T1PKS	T1PKS	21,481	T1PKS 8	1854	21481
279	samp_4444_metabat2_185_contig_246_T1PKS	T1PKS	14,674	T1PKS 9	1327	14673

234	samp_4445_semibin_379_contig_282_T1PKS	PUFA,T1PKS,hglE-KS	20,440	T1PKS-hglE-KS 1	2	14732
208	samp_4443_metadecoder_774_contig_17_T1PKS	T1PKS,prodigiosin,NRPS	59,771	T1PKS-NRPS-Prodigiosin 1	41 39	58628
183	samp_4434_metabat2_58_contig_508_T1PKS	T1PKS,terpene	17,306	T1PKS-Terpene 1	76	15578
179	samp_4434_metabat2_58_contig_50_T3PKS	T3PKS	11,732	T3PKS 1	3	11731
200	samp_4445_VAMB_546_contig_53_T3PKS	T3PKS	6,818	T3PKS 2	1	6817
257	samp_4443_semibin_466_contig_286_T3PKS	T3PKS	9,489	T3PKS 3	3	8405
9	samp_4444_semibin_482_contig_165_terpene	terpene	10,855	Terpene 1	2	10,857
43	samp_4447_VAMB_472_contig_346_terpene	terpene	11,178	Terpene 10	1	11,179
46	samp_4448_VAMB_889_contig_147_terpene	terpene	12,591	Terpene 11	1	12,592
48	samp_4445_maxbin_221_contig_370_terpene	terpene	10,028	Terpene 12	1	10,029

51	samp_4321_maxbin_137_contig_210_terpene	terpene	20,932	Terpene 13	2	20,934
52	samp_4321_maxbin_137_contig_261_terpene	terpene	20,830	Terpene 14	2	20,832
53	samp_4332_VAMB_1153_contig_44_terpene	terpene	11,172	Terpene 15	2	11,174
54	samp_4332_VAMB_1153_contig_99_terpene	terpene	17,780	Terpene 16	1	17,781
56	samp_4319_VAMB_955_contig_178_terpene	terpene	17,463	Terpene 17	2	17,465
60	samp_4319_VAMB_955_contig_444_terpene	terpene	11,734	Terpene 18	1	11,735
68	samp_4445_semibin_342_contig_168_terpene	terpene	20,970	Terpene 19	1	20,971
10	samp_4444_semibin_482_contig_216_terpene	terpene	14,292	Terpene 2	2	14,294
70	samp_4443_metabat2_120_contig_75_terpene	terpene	11,682	Terpene 20	1	11,683
71	samp_4443_metabat2_120_contig_110_terpene	terpene	14,552	Terpene 21	1	14,553
72	samp_4443_semibin_400_contig_169_terpene	terpene	9,348	Terpene 22	1	9,349

78	samp_4445_concoct_1447_contig_111_terpene	terpene	6,821	Terpene 23	1	6,822
79	samp_4445_concoct_1447_contig_277_terpene	terpene	7,390	Terpene 24	1	7,391
93	samp_4445_metabat2_36_contig_20_terpene	terpene	20,923	Terpene 25	2	20,925
94	samp_4445_metabat2_36_contig_61_terpene	terpene	21,286	Terpene 26	2	21,288
95	samp_4445_VAMB_1115_contig_244_terpene	terpene	16,729	Terpene 27	1	16,730
96	samp_4324_maxbin_57_contig_128_terpene	terpene	11,706	Terpene 28	1	11,707
100	samp_4324_maxbin_57_contig_447_terpene	terpene	9,786	Terpene 29	1	9,787
12	samp_4435_concoct_1904_contig_173_terpene	terpene	7,382	Terpene 3	1	7,383
103	samp_4434_semibin_67_contig_168_terpene	terpene	15,546	Terpene 30	2	15,548
104	samp_4434_semibin_67_contig_175_terpene	terpene	15,682	Terpene 31	1	15,683
108	samp_4445_metabat2_59_contig_300_terpene	terpene	8,368	Terpene 32	1	8,369

109	samp_4324_maxbin_57_contig_776_terpene	terpene	13,276	Terpene 33	2	13,278
117	samp_4448_VAMB_839_contig_1804_terpene	terpene	15,472	Terpene 34	2	15,474
121	samp_4435_VAMB_805_contig_126_terpene	terpene	10,169	Terpene 35	1	10,170
127	samp_4444_semibin_481_contig_179_terpene	terpene	6,019	Terpene 36	1	6,020
132	samp_4447_VAMB_322_contig_31_terpene	terpene	8,457	Terpene 37	1	8,458
134	samp_4447_VAMB_322_contig_532_terpene	terpene	9,555	Terpene 38	1	9,556
138	samp_4446_VAMB_96_contig_306_terpene	terpene	10,930	Terpene 39	1	10,931
23	samp_4435_concoct_1904_contig_2065_terpene	terpene	10,868	Terpene 4	1	10,869
142	samp_4443_VAMB_1051_contig_58_terpene	terpene	15,297	Terpene 40	2	15,299
143	samp_4331_metabat2_120_contig_60_terpene	terpene	14,576	Terpene 41	2	14,578
144	samp_4331_metabat2_120_contig_210_terpene	terpene	21,259	Terpene 42	2	21,261

148	samp_4435_maxbin_422_contig_157_terpene	terpene	11,041	Terpene 43	1	11,042
152	samp_4435_maxbin_422_contig_364_terpene	terpene	11,115	Terpene 44	1	11,116
154	samp_4435_maxbin_440_contig_357_terpene	terpene	7,278	Terpene 45	1	7,279
155	samp_4435_maxbin_440_contig_634_terpene	terpene	8,376	Terpene 46	1	8,377
163	samp_4446_VAMB_171_contig_132_terpene	terpene	10,302	Terpene 47	1	10,303
166	samp_4444_metabat2_78_contig_58_terpene	terpene	6,249	Terpene 48	1	6,250
167	samp_4444_metabat2_78_contig_113_terpene	terpene	13,323	Terpene 49	2	13,325
26	samp_4444_metabat2_75_contig_469_terpene	terpene	6,225	Terpene 5	1	6,226
168	samp_4443_semibin_480_contig_163_terpene	terpene	8,920	Terpene 50	1	8,921
170	samp_4443_maxbin_277_contig_38_terpene	terpene	12,666	Terpene 51	1	12,667
176	samp_4324_maxbin_53_contig_286_terpene	terpene	13,693	Terpene 52	1	13,694

186	samp_4321_concoct_295_contig_35_terpene	terpene	7,718	Terpene 53	1	7,719
192	samp_4444_metabat2_114_contig_228_terpene	terpene	20,935	Terpene 54	2	20,937
195	samp_4320_metadecoder_628_contig_198_terpene	terpene	12,243	Terpene 55	1	12,244
197	samp_4443_semibin_379_contig_6_terpene	terpene	14,240	Terpene 56	1	14,241
201	samp_4448_metabat2_83_contig_176_terpene	terpene	10,645	Terpene 57	1	10,646
203	samp_4444_semibin_569_contig_138_terpene	terpene	21,084	Terpene 58	2	21,086
207	samp_4443_concoct_2141_contig_81_terpene	terpene	14,258	Terpene 59	1	14,259
34	samp_4317_concoct_68_contig_406_terpene	terpene	19,770	Terpene 6	2	19,772
212	samp_4447_concoct_1018_contig_976_terpene	terpene	7,197	Terpene 60	1	7,198
218	samp_4445_metabat2_158_contig_342_terpene	terpene	15,839	Terpene 61	1	15,840
220	samp_4445_metabat2_158_contig_419_terpene	terpene	14,948	Terpene 62	1	14,949

223	samp_4445_concoct_1532_contig_132_terpene	terpene	14,195	Terpene 63	1	14,196
225	samp_4445_maxbin_504_contig_64_terpene	terpene	10,733	Terpene 64	1	10,734
226	samp_4444_semibin_500_contig_4_terpene	terpene	11,034	Terpene 65	1	11,035
228	samp_4445_metabat2_46_contig_72_terpene	terpene	12,144	Terpene 66	2	12,146
230	samp_4445_metabat2_46_contig_482_terpene	terpene	8,898	Terpene 67	1	8,899
236	samp_4445_semibin_379_contig_677_terpene	terpene	8,790	Terpene 68	1	8,791
241	samp_4322_concoct_184_contig_278_terpene	terpene	16,515	Terpene 69	1	16,516
35	samp_4444_metabat2_171_contig_181_terpene	terpene	10,534	Terpene 7	1	10,535
248	samp_4447_metabat2_80_contig_110_terpene	terpene	7,517	Terpene 70	1	7,518
249	samp_4443_semibin_319_contig_86_terpene	terpene	12,545	Terpene 71	1	12,546
256	samp_4443_semibin_466_contig_149_terpene	terpene	8,115	Terpene 72	1	8,116

258	samp_4444_metadecoder_351_contig_152_terpene	terpene	16,755	Terpene 73	1	16,756
260	samp_4313_concoct_735_contig_94_terpene	terpene	16,518	Terpene 74	1	16,519
261	samp_4313_concoct_735_contig_188_terpene	terpene	10,786	Terpene 75	2	10,788
276	samp_4318_metabat2_38_contig_498_terpene	terpene	14,796	Terpene 76	1	14,797
280	samp_4444_metabat2_185_contig_418_terpene	terpene	10,535	Terpene 77	1	10,536
283	samp_4322_metabat2_24_contig_181_terpene	terpene	12,648	Terpene 78	1	12,649
289	samp_4444_metabat2_168_contig_614_terpene	terpene	12,027	Terpene 79	2	12,029
36	samp_4444_maxbin_831_contig_107_terpene	terpene	20,950	Terpene 8	2	20,952
291	samp_4448_concoct_1792_contig_692_terpene	terpene	12,948	Terpene 80	1	12,949
299	samp_4445_semibin_421_contig_279_terpene	terpene	13,507	Terpene 81	1	13,508
301	samp_4445_semibin_421_contig_292_terpene	terpene	20,830	Terpene 82	2	20,832

305	samp_4445_semibin_366_contig_76_terpene	terpene	20,836	Terpene 83	2	20,838
39	samp_4447_VAMB_472_contig_32_terpene	terpene	21,166	Terpene 9	1	21,167
164	samp_4320_metadecoder_702_contig_50_terpene	terpene,NRPS-like	63,705	Terpene-NRPS-like 1	2	63,707
112	samp_4448_VAMB_839_contig_142_terpene	terpene,oligosaccharide	29,354	Terpene-Oligosaccharide 1	2	29,356
206	samp_4447_VAMB_291_contig_134_thioamitides	thioamitides	13,466	Thioamitide 1	1	13,467
217	samp_4445_metabat2_158_contig_330_thioamitides	thioamitides	15,513	Thioamitide 2	1	15,514

Chapter 4 Cyanopeptide Mixtures Induce Variable Synergistic and Antagonistic Effects Across Diverse Human Cell Lines

4.1 Abstract

Cyanobacterial harmful algal blooms (cyanoHABs) threaten human, animal, and ecosystem health and safety through production of toxic secondary metabolites. *Microcystis,* a cosmopolitan bloom-forming cyanobacterial genus, is well-known for producing hepatotoxic microcystins (MCs), but it can produce many other bioactive cyanopeptides, such as anabaenopeptins (APs), that occur at higher levels in blooms. The toxicological and ecological impacts of such co-occurring cyanopeptides in the natural environment remain understudied. Here we evaluated the effects of pure MCs and APs individually and in combination, as well as extracts of *Microcystis* cultures producing diverse suites of cyanopeptides, including strains with and without MCs and APs, on human lung (A549), kidney (HK2), and liver (Hep-3B) cell viability. Individual MC and AP congeners exhibited a gradient of toxic effects across cell lines; MC-LA caused the most toxic effects, MC-LR had comparable effects to AP-A and AP-B, and MC-RR caused the least toxicity. Combined exposure to MC-LA and AP-B produced dose-dependent synergistic effects across all three cell lines. *Microcystis* culture extracts significantly reduced cell viability in dose- and *Microcystis* strain-dependent patterns that could not be explained by cyanopeptide content alone, suggesting a role of other metabolites and their interactions within the mixtures. These findings demonstrate that mixtures of environmentally relevant cyanopeptides can have greater toxic threats than individual compounds and underscore

the importance of considering metabolites beyond MCs and their potential interactions in public health, future risk assessments, and management strategies for cyanoHAB-impacted waters.

4.2 Introduction

Cyanobacterial harmful algal blooms (cyanoHABs) are a growing, global threat to freshwater ecosystems and public health due to their production of toxic secondary metabolites (1–4). Among these, federal agencies in the United States currently monitor for four cyanotoxins: microcystins (MCs), cylindrospermopsin, anatoxin-a, and saxitoxin (5). These compounds can accumulate to micromolar concentrations in freshwater systems used for drinking, recreation, and agriculture, posing acute and chronic health risks to humans and animals depending on their structure, exposure route and duration, mode of entry, and mechanism of action (6–9). However, these four cyanotoxins and their related congeners represent only a small fraction, approximately 10%, of the more than 3,000 cyanobacterial metabolites identified to date (10,11). Many of these unmonitored cyanobacterial metabolites exhibit diverse and potent bioactivities, yet remain poorly characterized, particularly in the context of human health risk (12,13). Beyond the known and predicted risks of characterized cyanobacterial metabolites, advances in metabologenomics suggest that a wealth of uncharacterized cyanobacterial secondary metabolites remain undiscovered, representing an even greater, largely unassessed toxicological unknown (14,15).

MCs are the most well-studied class of cyanobacterial metabolites. They are cyclic heptapeptides with over 300 characterized congeners produced by various bloom-forming cyanobacterial genera, including *Microcystis, Dolichospermum, Planktothrix, Anabaena*, and *Oscillatoria* (10,11,16,17). Their structure consists of seven L- and D-amino acids, including the

distinctive Adda moiety [(2S,3S,8S,9S)-3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid], which is produced by a polyketide synthase and is critical for hepatotoxicity (18,19). MC congener diversity predominantly arises from variations at the second (X) and fourth (Z) amino acid positions (**Figure 4.8.1)**. However, modifications of all amino acids within the backbone have been reported. Despite structural diversity, all MCs potently inhibit serine/threonine protein phosphatases PP1 and PP2A in the low nanomolar range in vitro, leading to disruption of cytoskeletal integrity and cell signaling pathways (20–23). Their toxic potential in vivo, however, varies widely by congener (e.g., 50–600 μg/kg in mice) (24,25). The discrepancy between *in vitro* PP inhibition and *in vivo* toxicity could be explained by differences in cellular uptake for distinct congeners via Organic Anion Transporting Polypeptides (OATPs), primarily OATP1B1 and OATP1B3 (26–29). These transporters mediate the active entry of amphiphilic MCs into hepatocytes and other tissues, and OATP inhibitors ameliorate the hepatoxic effects of MCs (30–33). *In vitro* toxicology studies of MCs must therefore account for cell line-specific OATP expression. In addition to PP inhibition, MC exposure has been linked to damage in liver, kidney, lung, and skin tissues, induction of inflammatory responses, and oxidative stress (34–36).

Microcystis is one of the most common and abundant bloom-forming cyanobacteria globally, and is studied extensively for its ability to produce MCs (1,37). However, diverse *Microcystis* strains also produce a slurry of other bioactive cyanopeptides, including anabaenopeptins (APs), cyanopeptolins, microviridins, microginins, aeruginosins, and cyclamides (12,14,38). These metabolites target diverse enzymes such as serine proteases, peptidases, and phosphatases and may influence bloom ecology through allelopathy and grazing deterrence, yet only ~10% of studies on *Microcystis* secondary metabolism focus on these

compounds (13). More recent monitoring efforts suggest that several non-MC cyanopeptides, including APs, are often present at concentrations exceeding those of MCs during bloom events, raising concerns about their individual and combinatorial roles in human exposure and toxicity (6,39,40).

Anabaenopeptins (APs) are cyclic hexapeptides synthesized via nonribosomal peptide synthetases and are structurally characterized by an invariant lysine and a ureido linkage with over 120 AP congeners characterized to date (10,11). AP concentrations in environmental samples can reach levels greater than 1000 μg^{-1} L in some freshwater systems, with AP-A, AP-B, and AP-F among the most commonly detected congeners (6,41). While APs inhibit PP1 and PP2A in the mid to low micromolar range, they potently inhibit carboxypeptidases at low nanomolar concentrations (42,43). Despite their demonstrated bioactivity, APs remain unregulated, are not routinely monitored, and are still broadly categorized as non-toxic. One of the few in vivo studies on APs to date demonstrated that AP-A, B, and F reduced growth, hatching, reproduction, and lifespan in nematodes (*Caenorhabditis elegans*) to a similar extent as MC-RR (44). However, their mode of entry, mechanism of action, subcellular localization, and clearance in human models remain uncharacterized.

While MCs and APs are frequently detected together in environmental samples, their combined toxicity has not been extensively studied in human-relevant systems. Evidence from *Daphnia* and trout suggests that MC–AP combinations can exert synergistic toxic effects, raising concerns that co-exposure in humans and wildlife may also result in enhanced toxicity (45,46). Fish such as bass, trout, and pike have been shown to accumulate both MCs and APs, implying potential dietary exposure risks (47). Further, the presence and concentration of MCs in *Microcystis* extracts often poorly predict toxicity, suggesting that additional, unidentified

metabolites or interactions among metabolites may drive biological effects (48). These combinatorial effects may stem from multiple mechanisms, including additive or synergistic biochemical activity, competitive or facilitated uptake, or modulation of metabolic and signaling pathways (49). Yet, few studies have systematically explored these dynamics in human cell models.

In this study, we investigated the cytotoxic effects of environmentally relevant MC and AP congeners individually, in combination, and within complex *Microcystis* metabolite mixtures on human lung (A549), kidney (HK2), and liver (Hep-3B) cell lines. By comparing dose- and congener-dependent effects across cell types, and assessing the effects of *Microcystis* metabolite mixtures, we aimed to uncover how combinatorial exposure shapes toxicity in human-relevant systems.

4.3 Materials & Methods

4.3.1 Cell Line Rationale

In this study, we selected to use A549 (lung), HK2 (kidney), and Hep-3B (liver) human cell lines based on the known impacts of cyanotoxins and bioactive cyanopeptides on these organs. Each cell line expresses at least one OATP transporter (36,50,51), making them appropriate models for evaluating the toxicity of microcystins (MCs) (27,52).

4.3.2 Hep-3B Cells

Human hepatocellular carcinoma (liver epithelial) Hep-3B cells were acquired from ATCC (Virginia, USA, catalog no. HB-8064) (53). Cells were cultured under standard conditions and stored in Gibco Recovery Cell Culture Freezing Medium at 2 x 10^6 cells per aliquot. Cells were

cultured in T175 flasks in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies media, Carlsbad, CA, USA) supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin. Cultures were maintained at 37°C in a humid atmosphere with 5% CO_2.

4.3.3 A549 Cells

Human adenocarcinoma (lung tissue) A549 cells were acquired from ATCC (Virginia, USA, catalog no. CCL-185) (54). Cells were cultured under standard conditions and stored in Gibco Recovery Cell Culture Freezing Medium at 2 x 10^6 cells per aliquot. Cells were cultured in T175 flasks in the recommended F-12K medium (ATCC, Virginia, USA) supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin. Cultures were maintained at 37°C in a humid atmosphere with 5% CO_2.

4.3.4 HK-2 Cells

Human immortalized kidney HK-2 cells were acquired from American Type Cell Collection (ATCC, Virginia, USA, catalog no. CRL-2190). This cell line is a proximal tubular cell (PCT) line derived from a normal, male kidney (55). Cells were cultured under standard conditions and stored in Gibco Recovery Cell Culture Freezing Medium at 2 x 10^6 cells per aliquot. Cells were cultured in T175 flasks in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies media, Carlsbad, CA, USA) supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin. Cultures were maintained at 37°C in a humid atmosphere with 5% CO_2.

4.3.5 Microcystis Culturing

Seven *Microcystis* cultures from the western Lake Erie culture collection (WLECC) were selected for analysis in this study. Cultures were selected based on diversity of phylogeny (56) and biosynthetic potential (14) and their ability to grow in 500 mL volumes. The cultures selected were LE18-22.4, LE19-55.1, LE19-131.1, LE19-41.2, LE17-20, LE19-197.1, and LE19-84.1. Cultures were grown following the WLECC liquid transfer protocol (https://sites.lsa.umich.edu/wleculturecollection/). Cultures were grown without shaking at 25°C at 50 µmol photons m^{-2} s^{-1} light intensity on a 12:12 hour diel cycle in BG11-2N media (recipe available from the WLECC: https://docs.google.com/document/d/13Ggb_asGllLyJO4Zc3AJoQsQmr7ojAHgpdvVmgr_i4E/edit?tab=t.0#heading=h.cs9fserzkw1i). All cultures were started in a 10 mL volume at a 1:10 dilution in glass test tubes. Cultures were scaled up every two weeks at a 1:10 dilution, resulting in a final volume of 500 mL contained within two-liter glass bottles with caps lightly secured. Each culture was grown in triplicate to 500 mL if possible. One replicate from the strains LE19-41.2 and LE19-84.1 bleached during growth and resulted in these cultures only being grown in duplicate for the study. Uninoculated media controls were maintained in triplicate within the same incubator. All vessels used for culturing in this study were acid washed in 10% HCl overnight and extensively rinsed with Milli-Q water before the experiment.

4.3.6 Microcystis Culture Biomass Collection and Metabolite Extraction

Biomass was harvested after all strains were maintained at 500 mL for one week. Cultures were centrifuged in clean HPDE plastic bottles at 10,000 revolutions per minute (rpm) for 15 minutes. Supernatant was decanted and cell pellets were collected using a sterile metal scoopula. Cell pellets were put into glass scintillation vials and frozen at -20°C until extraction. Metabolites

were extracted from the cellular biomass using a freshly made 50:50 high-performance liquid chromatography (HPLC)-grade methanol/chloroform mixture. The solvent mixture was added to the thawed scintillation vials to completely cover the biomass. Vials were then sonicated in a water bath for 30 minutes. Methanol rinsed stir bars were then placed in each vial and samples were warmed at 40°C and slowly stirred for four hours with caps lightly screwed on to allow for gas exchange. For each sample, two clean, pre-weighed vials, one for cell viability assays and one for targeted metabolomics analysis, were prepared. Extraction mixtures were filtered to remove residual solids and then evenly divided between the two vials. After filtration, vials were dried under a gentle stream of N_2 gas at 30°C. Dried samples were re-weighed to calculate their crude extract weight and stored at -20°C until further analysis.

4.3.7 Targeted Metabolomics Data Collection and Analysis

Targeted metabolomics vials for each sample were delivered to the University of Toledo on ice and were stored in -20°C prior to analysis. Microcystin and anabaenopeptin congener composition was determined using liquid chromatography-mass spectrometry (LC-MS) via methods previously described in Baliu-Rodriquez et al. (2022) (17) and Flores and Caixach (2020) (39), respectively. Microcystin and anabaenopeptin congener identification was confirmed via standards for isoforms MC-LR, MC-RR, MC-YR, MC-LA, MC-LF, [Asp^3] MC-LR, MC-HilR, MC-LY, AP-A, and AP-B (**Figure 4.8.1**). Characteristic ions for AP-A and AP-B were 201.09 *m/z* $[CO+Arg-COOH]^+$ and 84.08 *m/z* $[Lys+H-CO-NH_3]^+$.

4.3.8 Cell Line Exposures to Individual Pure Cyanopeptide Standards

Pure cyanotoxin standards in powder form were purchased from Cayman Chemical (Ann Arbor, MI, USA) (MC-LR) and Enzo Life Sciences (Farmingdale, NY, USA) (MC-LA, MC-RR, AP-A, AP-B). Doxorubicin was purchased from AK Scientific (Union City, CA, USA) in powder form and used as a positive control in all assays. Powder stocks were stored at 4°C, and liquid stocks were made in dimethyl sulfoxide (DMSO) and stored at -20°C for the duration of experimental use. The viability and toxicity of cyanotoxin standards individually were assessed in Hep-3B, A549, and HK-2 cells using the CellTiter-Glo assay (Promega, Madison, WI, USA). Cells were seeded in 100 µL of medium at a density of 10,000 cells per well in opaque, 96-well tissue culture-treated plates after reaching desired confluency and viability >90% in flasks, and allowed to adhere for 24 hours. The A549 cell line was seeded at 4,000 cells per well due to the faster doubling time. After adherence, the culture media was replaced with fresh medium and cells were treated with four two-fold serial dilutions of single cyanotoxins and doxorubicin with a highest concentration of 100 µM. The plates were incubated at 37°C for an additional 48 hours. After incubation, 100 µL of CellTiter-Glo reagent was added to each well and shaken for five minutes at room temperature. Luminescence from each well was measured using a spectrophotometer. Results were normalized to a DMSO-vehicle control, and growth curves were visualized using GraphPad Prism 10 (San Diego, CA, USA). Each cyanotoxin was tested in duplicate on each plate, and each experiment was repeated on different days to ensure technical and biological replication. All data is shown in **SI Table S4.9.1**. Methods used for this process are visualized in **Figure 4.8.2.**

4.3.9 Cell Line Exposures to Combinations of Pure Cyanopeptide Standards

As above, cells were seeded in 100 µL of medium at a density of 10,000 or 4,000 cells per well in opaque, 96-well tissue culture-treated plates and allowed to adhere for 24 hours. The activity of MC-LA and AP-B in combination was evaluated using a checkerboard assay. Compounds were each four-fold serially diluted in 50 µL of culture medium within four separate flat-bottom microtiter plates starting at 100 µM. AP-B was transferred to wells containing MC-LA, so each row contained a serial dilution of MC-LA and each column contained a serial dilution of AP-B. Additionally, each plate included wells containing only single compounds, serially diluted four-fold. The plates were incubated at 37°C for 48 hours. After incubation, 100 µL of CellTiter-Glo reagent was added to each well and incubated for five minutes at room temperature. Luminescence from each well was measured using a spectrophotometer. Results were normalized to a DMSO-vehicle control, and effects were visualized using GraphPad Prism 10 (San Diego, CA, USA). The combination was tested in duplicate on each plate, and the experiment was repeated on different days to ensure technical and biological replication. All data is shown in **SI Table S4.9.2**.

4.3.10 Quantifying Cyanopeptide Combination Activity

SynergyFinder+ (57) was used to quantify the activity of MC-LA and AP-B in combination relative to their individual effects. The tool provides four scoring metrics to calculate the synergy score of an experiment based on specific assumptions: Bliss Independence (Bliss) (58), Loewe additivity (Loewe) (59), Zero Interaction Potency (ZIP) (60), and Highest Single Agent (HSA) (61), **SI Table S4.9.3**). The Loewe additivity model was used because it defines the expected effect of a combination as if a molecule were combined with itself and considers the dose-response curves of individual molecules. The expected effect (y_{loewe}) must satisfy Equation 4.3.1,

where x_1 and x_2 are doses for each molecule and X^1_{loewe} and X^2_{loewe} are the doses of molecule 1 and 2 alone that produce y_{loewe}.

Equation 4.3.1: Expected effect (y_{loewe})

$$\left(\frac{x_1}{X^1_{loewe}}\right) + \left(\frac{x_2}{X^2_{loewe}}\right) = 1$$

Using 4-parameter log-logistic (4PL) curves to describe dose-response curves, the parametric form of Equation 4.3.1 is derived in Equation 4.3.2, where E_{min} and E_{max} are the minimal and maximal effects of the molecules, $m_{1,2}$ are the doses of the molecules required to produce the midpoint effect of $E_{min} + E_{max}$ (EC_{50}), and $\lambda_{1,2}$ are the shape parameters indicating the slope of the dose-response curves.

Equation 4.3.2: Parametric form of expected effect equation

$$\left(\frac{x_1}{m_1\left(\frac{y_{loewe} - E^1_{min}}{E^1_{max} - y_{loewe}}\right)^{1/\lambda_1}}\right) + \left(\frac{x_2}{m_1\left(\frac{y_{loewe} - E^2_{min}}{E^2_{max} - y_{loewe}}\right)^{1/\lambda_2}}\right) = 1$$

A Loewe score greater than one is indicative of antagonistic effects, a Loewe score less than one is indicative of synergistic effects, and a Loewe score approximately equal to 1 is indicative of additive effects.

4.3.11 Cell Line Exposures to Culture Extracts

Extracts were reconstituted in 50/50 DMSO/water at a concentration of 100 mg/mL, based on the crude extract weight of the sample and stored at -20°C for the duration of experimental use. The viability and toxicity of all extracts were assessed in Hep-3B, A549, and HK-2 cells using the

CellTiter-Glo assay (Promega, Madison, WI, USA). As above, cells were seeded in 100 µL of medium at a density of 10,000 or 4,000 cells per well in opaque, 96-well tissue culture-treated plates and allowed to adhere for 24 hours. After adherence, culture media was replaced with 100 µL of fresh medium and cells were treated with four serial dilutions between 10 and 100 µM of each extract, doxorubicin (positive control), and the highest stock concentration of DMSO/water to account for a negative vehicle control. The plates were incubated at 37°C for 48 hours. After incubation, 100 µL of CellTiter-Glo reagent was added to each well and incubated for five minutes at room temperature. Luminescence from each well was measured using a spectrophotometer. Results were normalized to a DMSO-vehicle control, and growth curves were visualized using GraphPad Prism 10 (San Diego, CA, USA). Each extract was tested in duplicate on each plate, and each experiment was repeated on different days to ensure technical and biological replication. All data is shown in **SI Table S4.9.4.** Methods used for this process are visualized in **Figure 4.8.2.**

4.3.12 Statistical Analysis

Statistical analysis was performed using GraphPad Prism 10 (San Diego, CA, USA) and in R on R Studio using the *stats* package(62). For comparisons involving less than two groups, a one-way analysis of variance (ANOVA) was performed and followed up with Tukey's Honest Significant Difference (HSD) multiple comparisons post-hoc tests when applicable. All growth curve data is presented as the mean and standard error of the mean. A P-value of <0.05 was considered statistically significant in this study. To assess the relationship between cyanopeptide concentration in *Microcystis* culture extracts and cell viability, Pearson correlation coefficients

(R) were calculated between cell viability at four doses (12.5, 25, 50, and 100 μM) and parts per billion (ppb) of each cyanopeptide congener. The amount of each cyanopeptide congener present in the 100 μM dose in the cell viability assays was calculated using SI Equation 4.9.1 in **SI Table S4.9.5**.

4.3.13 Figure Rendering

All figures were rendered in GraphPad Prism10 software (San Diego, CA, USA) and in R on R Studio using the package *ggplot2* (62,63).

4.4 Results

4.4.1 Individual Cyanopeptides Exhibit Dose-Dependent Responses Across Different Cell Lines

A549 (lung), HK2 (kidney), and Hep-3B (liver) cells exposed to 0–100 μM of MC-RR, MC-LR, MC-LA, AP-A, and AP-B for 48 hours showed decreases in cell viability in a cell line-, congener-, and dose-dependent manner (**Figure 4.8.3**) relative to a vehicle control (0.5% DMSO) None of the cyanopeptides reduced cell viability to the same extent as the positive control (doxorubicin) at any dose tested. A549 cell viability decreased most strongly, followed by less pronounced effects in HK2 and Hep-3B cells. Across all three cell lines, MC-LA consistently caused the greatest decrease in viable cells. MC-LR, AP-A, and AP-B exhibited similar levels of cell viability reduction, while MC-RR induced the smallest decrease in cell viability out of all congeners tested across all cell lines. Significant decreases in cell viability by MC-LA were observed at 100 μM in all cell lines, at 50 μM in both A549 and HK2 cells, and at 12.5 and 25 μM in A549 cells ($P<0.05$, Tukey's HSD). Additionally, all cyanopeptides except

MC-RR significantly reduced viability at 100 µM in A549 cells (P<0.05, Tukey's HSD). The IC_{50} for each cyanopeptide against each cell line could not be calculated as they often did not cause cell viability decreases beyond 50 percent.

4.4.2 Cyanopeptide Combinations Exhibit Dose-Dependent Synergistic and Antagonistic Responses Across Different Cell Lines

The effects of MC-LA, the most toxic MC congener, and AP-B on cell viability were assessed in combination via checkerboard assays in A549, HK2, and Hep-3B cells (**Figure 4.8.4a–c**). Loewe synergy scores for each combination were calculated and visualized (**Figure 4.8.4d–f, SI Table S4.9.3**), with scores <1 indicating synergism (enhanced activity), >1 indicating antagonism (decreased activity), and ≈1 indicating additivity. The combined exposure to MC-LA and AP-B resulted in dose-dependent patterns of cell viability reduction that varied across all three cell lines.

In A549 cells, the greatest decrease in viable cells occurred following individual exposure to 100 µM MC-LA. Similar effects were observed for combinations of high doses, between 50 and 100 µM, of MC-LA and AP-B. Mild effects on cell viability were observed at the lowest dose of MC-LA. Dose-dependent cell viability was observed for AP-B, although maximum cell viability decreased with high doses of MC-LA regardless of the presence of AP-B. The mean Loewe synergy score for all combinations in A549 cells was 0.45, indicating neither strong synergism nor antagonism. Loewe synergy scores increased with higher concentrations of AP-B in the presence of 100 µM MC-LA, indicating less synergism and suggesting that AP-B may interfere with the toxic effects of MC-LA at high concentrations in this model.

In HK2 cells, cell viability decreased most following individual exposure to solely 100 μM MC-LA and by MC-LA between 12.5 and 100 μM when combined with 100 μM AP-B. However, combinations of 100 μM MC-LA and intermediate concentrations of AP-B (12.5-50 μM) did not reduce cell viability. The mean Loewe synergy score for all combinations in HK2 cells was -0.48, not indicative of strong synergistic or antagonistic mechanisms. However, lower scores (synergistic) were observed at doses of 50 μM or lower of MC-LA in combination with doses between 25 and 100 μM AP-B, indicating strong synergism at these doses.

In Hep-3B cells, the synergistic effect of MC-LA and AP-B was the most consistent of all the cell lines, with viability reduced under combined exposure to 100 μM MC-LA and AP-B and a gradual dose-dependent increase in viable cell reduction across increasing concentrations of both cyanopeptides. The mean Loewe synergy score for all combinations indicated synergism (-9.33), suggesting that combined exposure to both cyanopeptides increased cell viability reduction beyond their individual effects. Synergism was most pronounced at high combined doses (50-100 μM) and decreased in a dose-dependent manner with concentration of both compounds.

4.4.3 Microcystis Culture Extracts Display Dose-Dependent Effects Not Explained by Cyanopeptide Concentrations Alone

Extracts from seven *Microcystis* cultures were assessed for their cyanopeptide content (**Figure 4.8.5**) and dose-dependent effects on cell viability across three cell lines (A549, HK2, Hep-3B) (**Figure 4.8.6)**. No culture extract contained both MCs and APs, and none included all measured congeners. The rarest MC congeners were MC-LA, detected only in LE19-84.1, and MC-LF, present in LE19-55.1 and LE18-22.4. APs were only detected in extracts from LE19-

197.1, containing both AP-A and AP-B, while LE17-20 and LE19-131.1 had no detectable cyanopeptides. Some biological replicates from the same culture varied, including LE18-22.4 (A vs B and C), LE19-55.1 (C vs A and B), and LE19-84.1 (A vs B).

Culture extracts caused dose-dependent reductions in cell viability across all three cell lines (**Figure 4.8.6, SI Figure S4.9.1)**. At 100 µM crude extract, significant decreases in viability relative to vehicle controls were observed for LE19-131.1B, LE19-197.1B, and LE19-55.1B in all three cell lines (P<0.05, Tukey's HSD). Actual concentrations of MCs and APs present in these extracts at 100 µM doses ranged between 0.1 nM and 0.2 µM (**SI Table S4.9.5)**, orders of magnitude lower than the concentrations used in the present study's individual cyanopeptide experiments (**Figure 4.8.3)**. However, of these extracts, LE19-131.1B had no detectable cyanopeptides within the extracts but caused greater reductions in viable cells than pure MC-LA at 100 µM, suggesting the potential role of unidentified bioactive compounds or interactions within complex extract mixtures. Moreover, all LE17-20 replicates, lacking any detectable MCs or APs, induced significant viability reductions across several doses in all cell lines (P<0.05, Tukey's HSD). The most significant viability reductions effects were observed in Hep-3B cells with twelve significant effects by different cultures, replicates, and doses (P<0.05, Tukey's HSD), compared to four significant effects in both A549 and HK2 cells. These results indicate variability in extract toxicity based on culture, replicate, dose, and cell line (**SI Figure S4.9.2**).

To evaluate whether cyanopeptide content accounted for the variability in toxicity, Pearson correlation coefficients were calculated between cyanopeptide concentrations (ppb) and cell viability across four doses for the extracts. Weak correlations (R= -0.1 to 0.3) were observed between individual congeners, or total MC or AP levels, and viability outcomes. MC content and viability correlations increased at higher doses (50-100 µM), while AP content and viability

correlations decreased. Consistently low R values suggest that other factors, such as additional metabolites in the extracts and their interactions may better explained observed viability outcomes.

4.5 Discussion

Across the globe, cyanobacteria are producing cyanotoxins, bioactive cyanopeptides, and a vast array of uncharacterized secondary metabolites in freshwater systems relied on by millions of people, animals, and ecosystems (1,4,15). These chemical cocktails pose an emerging threat in increasingly eutrophic waters. While regulatory frameworks have focused on a limited subset of well-characterized cyanotoxins, the potential toxicity of unmonitored metabolites and synergistic toxicity from interactions among co-occurring metabolites remain poorly understood. To address this gap, we evaluated the cytotoxic effects of two of the most commonly detected and abundant cyanopeptide classes in freshwater ecosystems, MCs and ABNs (6,40), both individually, in combination, and within complex *Microcystis* metabolite mixtures using human lung, kidney, and liver cell lines.

Consistent with previous studies on in vitro PP1/PP2A inhibition and in vivo toxicity (24,64–66), MC-LA was more toxic than MC-LR, which was in turn more toxic than MC-RR. Although all three congeners exhibit similar low-nanomolar IC_{50} values against PP1/PP2A (64,65), their effects on human cell viability vary substantially, from 35–90% reduction at 100 μM in the present study (**Figure 4.8.8**). This extensive variability in congener-specific effects on human cell viability is consistent with murine intraperitoneal LD_{50} values, ranging from ~50 to 600 μg/kg (66), and could reflect differences in transporter-mediated uptake (27,28,67). Studies using cells transfected with OATP1B1 and OATP1B3 show enhanced transporter selectivity for

MC congeners containing leucine and acidic, aromatic residues in positions X and Z (28). This aligns with observations that MC-LW and MC-LF exhibit highly potent IC50 values of ~0.4–0.6 nM in primary human hepatocytes, whereas MC-RR is significantly less potent, with IC50 values ranging from 167 to 900 nM, representing a ~1000-fold decrease in cytotoxicity (67). MC congeners containing arginine in either the X or Z position (e.g., MC-RR, MC-LR) are less potent, possibly due to their dicationic nature at physiological pH, which may reduce OATP-mediated uptake and thus, membrane transport (28). These findings suggest that more toxic MC congeners such as MC-LA, MC-LW, and MC-LF warrant increased monitoring and toxicological evaluation. Risk assessments that focus solely on MC-LR may underestimate risk, particularly when more potent congeners are present. In western Lake Erie, a eutrophic basin of the Laurentian Great Lakes plagued by seasonal cyanoHABs, MC-RR and MC-LR are typically the dominant congeners, but MC-LA concentrations rise later in the season and are the fourth most common MC congener, warranting increased attention during the end of the bloom when attention and awareness of bloom conditions may decrease due to lower overall MC concentrations and less visible green scums (6,68).

The MC congeners in this study exhibited variable toxicity across cell lines, with A549 and HK2 cells generally more affected than Hep-3B cells. Unfortunately, these results cannot be directly compared to previous studies because to our knowledge, no studies have compared these specific congeners in A549, HK2, or Hep-3B cells at high concentrations (>50 μM). Differences between cell lines may stem from a combination of factors: variable expression of OATP transporters, differences in baseline metabolism, detoxification capacity, membrane composition and fluidity, susceptibility to stress, and intracellular accumulation (50,51,69,70). Among these, OATP expression is likely a major determinant. A549 and HK2 cells were more sensitive to

MCs than Hep-3B cells, which is consistent with reports that Hep-3B cells express OATP1B1 and OATP1B3 at lower levels (50,51,69,70). Given that MCs are widely recognized for their hepatotoxicity, it is critical that the field adopts standardized hepatic cell models for evaluating congener-specific toxicity mechanisms. For example, Hep-G2 cells are commonly used in assessing the toxicity of MCs, however, they exhibit low OATP expression and may not accurately reflect MC uptake or toxicity (71–73). To alleviate these issues, some studies recommend culturing hepatic cells in spheroid or aggregate formats to upregulate OATP expression and improve physiological relevance (74,75). Based on our findings, Hep-3B cells may serve as a more responsive hepatic model for MC toxicity assessment, although their sensitivity could likely be enhanced through similar 3D culture methods. Moreover, testing the effects of these cyanopeptides across a wider range of concentrations could further determine their potency in the selected cell lines.

Beyond hepatotoxicity, MCs also have nephrotoxic effects, including renal corpuscle enlargement, tubular damage, autophagy, and apoptosis (76,77). While Hek-293 cells transfected with OATP1B1/1B3 are often used to study MC-induced nephrotoxicity, we selected HK2 cells for their proximal tubule phenotype and endogenous OATP expression (28,50,67). Our findings demonstrate that HK2 cells respond differentially to various MC congeners, supporting their utility as a physiologically relevant model for nephrotoxicity assessment of MCs. MCs have also been shown to exert cytotoxic and pro-inflammatory effects in lung tissue, especially in A549 cells (36,70,78). However, studies using A549 cells have primarily focused on the toxic effects of MC-LR at lower concentrations (< 10 μM) (70). The inclusion of additional MC congeners at high micromolar concentrations broadens the assessment of MC toxicity in pulmonary models and supports the use of A549 cells as a relevant in vitro system for evaluating MC-induced

effects in lung cell models. Based on this study and supporting literature, using OATP transporter-expressing human cell lines is vital for understanding MC toxicity mechanisms (27–29,67). However, this focus introduces bias when studying other cyanopeptides like anabaenopeptins, whose uptake mechanisms remain unknown.

APs are the second most widely studied cyanopeptide class and are increasingly recognized for their high abundance in cyanoHABs, including those in the Laurentian Great Lakes and Mediterranean (6,13,40,41). APs often co-occur with MCs and exhibit broad inhibitory effects of PP1/2A, serine proteases, and carboxypeptidases (13,79). In this study, AP cytotoxicity was greater than MC-RR and comparable to MC-LR. Despite this, APs are not currently regulated by the World Health Organization (WHO) or other agencies as MCs are. Given their widespread occurrence in water bodies used for drinking, recreation, agriculture, and daily life, our findings support the need to further study APs effects on human cell lines, whole organs, and organismal models such as mice to establish health-based guideline values for APs. Establishing such guidelines would not only enhance protection of humans and wildlife from cyanotoxin exposure but also drive further investigation into the mechanisms of action and cellular entry of APs, which remain poorly understood. Moreover, APs have been detected in water samples, aerosols, and fish tissues alongside MCs, highlighting the importance of considering their combined effects (39,47,80).

Given the frequent co-occurrence of MCs and APs in cyanoHABs, we evaluated the combined effects of MC-LA and AP-B, the most potent congeners from each cyanopeptide class tested, across three human cell lines. In A549 cells, minimal synergy was observed, and no combination surpassed the cytotoxicity of MC-LA alone at 100 μM. This may reflect the high expression of multiple OATPs in A549 cells, which facilitates efficient MC uptake and potent

individual effects (69,70). MC-LA and AP-B may compete for similar transporters, as AP-B shares peptide-like structure, amphipathicity, and anionic charge at physiological pH with known OATP substrates (81). However, its lower intrinsic cytotoxicity, particularly regarding its ability to inhibit phosphatases, may dampen combined effects (65,82). Alternatively, MC-LA may dominate uptake and intracellular impact, leaving little opportunity for ABN-B to enhance toxicity unless dosed very highly. In HK2 cells, synergistic interactions were observed when MC-LA was combined with high AP-B concentrations. Unlike A549 and Hep-3B cells, HK2 cells do not express OATP1B1 or OATP1B3, the main OATPs studied for MC transport, but they do express the kidney-specific OATP4C1 (36,50,69). Tissue-specific transporter profiles likely shape these differential responses to MC and AP combinations (83).

Beyond differences in transporter profiles, kidney cells are highly sensitive to oxidative stress, which may be amplified by this co-exposure (84). Further research is needed to understand the effects of diverse cyanopeptides on toxicity, inflammation, and transporter regulation in kidney tissues. In Hep-3B liver cells, strong synergy was observed with the combination of 100 μM MC-LA and AP-B. This may stem from liver-specific enrichment of MC/AP targets like protein phosphatases and modest expression of OATP1B1/1B3, enabling co-accumulation (85). AP-B may also potentiate MC-LA toxicity by modulating stress pathways such as ROS or inflammatory signaling, although studies on AP bioactivity in liver cells are lacking. Taken together, these results raise the possibility that synergy between cyanopeptides depends not only on transporter availability, but also on congener-specific potency, intracellular signaling dynamics, and potential off-target effects. Future studies should clarify the uptake mechanisms of APs, their targets in different tissues, and the broader cellular responses induced by co-exposure.

In addition to MCs and APs, *Microcystis* can produce a chemically diverse repertoire of cyanopeptides and other secondary metabolites with known or predicted bioactivity (14,15,86). Known bioactive cyanopeptides include serine protease and peptidase inhibitors such as aeruginosins, micropeptins, microginins, and microviridins, and cyclamides (12). Beyond these characterized molecules, *Microcystis* can synthesize many uncharacterized metabolites that may exert individual or combinatorial effects on human health (14,87). In this study, we assessed the cytotoxic effects of metabolite extracts from *Microcystis* cultures containing different profiles of known and unknown cyanopeptides and metabolites. The concentrations of MCs and APs in these extracts were poorly correlated with observed cytotoxicity at any given dose. For instance, LE18-22.4 and LE19-41.2 contained the highest concentrations of MCs but produced minimal effects on cell viability across all cell lines. In contrast, LE19-55.1, which produces both MCs and the potent class of serine protease inhibitors, micropeptins (14), showed clear cytotoxicity in all three cell lines. The most toxic extract, LE19-131.1, exhibited cytotoxicity exceeding that of MC-LA alone at 100 μM, despite lacking detectable MCs or APs. Genomic analysis indicates LE19-131.1 harbors the biosynthetic potential to produce APs and microginins, both of which can inhibit essential peptidases (14,42,88). While no APs were detected in this extract in this study, only two AP congeners were targeted in our analysis, and other congeners may have been present. Similarly, LE19-197.1 displayed strong cytotoxicity and was confirmed to contain APs. Of the most toxic extracts, all except LE19-55.1 encode the cryptic biosynthetic gene cluster PmNT3, predicted to produce a cyclophane-like compound such as merocyclophane, a metabolite with known cytotoxic potential (14,89). Additionally, *Microcystis* produces non-cyanopeptide bioactive metabolites such as pheophorbide, a chlorophyll breakdown product with pro-apoptotic activity in the low micromolar range (90), which was previously detected in all

cultures tested in this study (14). *Microcystis* also synthesizes a diverse array of fatty acids and lipid derivatives (14,91), which may contribute to cytotoxicity either directly or by enhancing the effects of co-occurring toxins, potentially through membrane disruption or metabolite–metabolite interactions (92). Consistent with other studies showing that *Microcystis* extracts lacking MCs can sometimes be more toxic than those with high MC concentrations (48,93–96), our findings support the hypothesis that other metabolites, either alone or in interaction, may be primary drivers of toxicity in some strains. Together, these results emphasize the need for comprehensive metabolomic and toxicological profiling of *Microcystis* strains beyond MC-focused risk assessments. Uncharacterized or combinatorially acting metabolites may represent significant but overlooked contributors to the health risks posed by cyanoHABs.

4.6 Conclusion

This study demonstrates the cytotoxic effects of MCs and APs individually, in combination, and within complex *Microcystis* metabolite mixtures on human lung, kidney, and liver cell models. Although APs exhibited cytotoxicity comparable to some MC congeners and showed synergistic interactions with MCs across all cell lines, MC and AP concentrations within complex *Microcystis* mixtures were poorly correlated with overall cytotoxicity, suggesting the involvement of additional metabolites and interactive effects. These findings highlight APs as toxicologically relevant cyanopeptides that warrant further investigation and evaluation for regulation and routine monitoring, particularly given their high abundance in freshwater systems impacted by cyanoHABs. More broadly, this work contributes to a growing body of evidence that risk assessments focused solely on MCs may underestimate the full scope of toxicity in cyanoHABs (6,12,13,40). Reassessing our current toxicity frameworks to include a broader

range of cyanopeptides and metabolite interactions is essential for accurately evaluating health risks and protecting both human and environmental health in the face of increasingly frequent and chemically diverse blooms.

4.7 References

1. Harke MJ, Steffen MM, Gobler CJ, Otten TG, Wilhelm SW, Wood SA, et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. Harmful Algae. 2016 Apr;54:4–20.

2. Gobler CJ, Di Cecco GJ, Doherty OM, Kramer BJ. Decadal warming has intensified *Microcystis*-dominated cyanobacterial blooms in Lake Erie. Limnol Oceanogr Letters. 2024 Jun 4;lol2.10406.

3. Paerl HW, Huisman J. Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. Environmental Microbiology Reports. 2009;1(1):27–37.

4. Plaas HE, Paerl HW. Toxic Cyanobacteria: A Growing Threat to Water and Air Quality. Environ Sci Technol. 20201218th ed. 2021 Jan 5;55(1):44–64.

5. Boegehold AG, Burtner AM, Camilleri AC, Carter G, DenUyl P, Fanslow D, et al. Routine monitoring of western Lake Erie to track water quality changes associated with cyanobacterial harmful algal blooms. Earth System Science Data. 2023 Aug 25;15(8):3853–68.

6. Zastepa A, Westrick JA, Miller TR, Liang A, Szlag DC, Chaffin JD. The Lake Erie Harmful Algal Blooms Grab: High-resolution mapping of toxic and bioactive metabolites (cyanotoxins/cyanopeptides) in cyanobacterial harmful algal blooms within the western basin. Aquatic Ecosystem Health & Management. 2024 Jan;27(1).

7. Cusick KD, Sayler GS. An Overview on the Marine Neurotoxin, Saxitoxin: Genetics, Molecular Targets, Methods of Detection and Ecological Functions. Mar Drugs. 2013 Mar 27;11(4):991–1018.

8. Mchau GJ, Machunda R, Kimanya M, Makule E, Gong YY, Mpolya E, et al. First Report of the Co-occurrence of Cylindrospermopsin, Nodularin and Microcystins in the Freshwaters of Lake Victoria, Tanzania. Expo Health. 2021 Jun 1;13(2):185–94.

9. Rapala J, Sivonen K, Luukkainen R, Niemelä SI. Anatoxin-a concentration in *Anabaena* and *Aphanizomenon* under different environmental conditions and comparison of growth by toxic and non-toxic *Anabaena*-strains — a laboratory study. J Appl Phycol. 1993 Dec 1;5(6):581–91.

10. Jones MR, Pinto E, Torres MA, Dörr F, Mazur-Marzec H, Szubert K, et al. CyanoMetDB, a comprehensive public database of secondary metabolites from cyanobacteria. Water Res. 20210308th ed. 2021 May 15;196:117017.

11. Janssen EML, Jones MR, Pinto E, Dörr F, Torres MA, Rios Jacinavicius F, et al. S75 | CyanoMetDB | Comprehensive database of secondary metabolites from cyanobacteria [Internet]. Zenodo; 2024 [cited 2025 Jun 12]. Available from: https://zenodo.org/doi/10.5281/zenodo.13854577

12. Janssen EML. Cyanobacterial peptides beyond microcystins – A review on co-occurrence, toxicity, and challenges for risk assessment. Water Research. 2019 Mar 15;151:488–99.

13. Stringer BB, Szlag Silva RG, Kodanko JJ, Westrick JA. Structure, Toxicity, Prevalence, and Degradation of Six Understudied Freshwater Cyanopeptides. Toxins. 2025 May;17(5):233.

14. Yancey CE, Hart L, Hefferan S, Mohamed OG, Newmister SA, Tripathi A, et al. Metabologenomics reveals strain-level genetic and chemical diversity of *Microcystis* secondary metabolism. Van Der Hooft JJJ, editor. mSystems. 2024 Jul 23;9(7):e00334-24.

15. Kust A, Řeháková K, Vrba J, Maicher V, Mareš J, Hrouzek P, et al. Insight into Unprecedented Diversity of Cyanopeptides in Eutrophic Ponds Using an MS/MS Networking Approach. Toxins (Basel) [Internet]. 20200831st ed. 2020 Aug 31;12(9). Available from: https://mdpi-res.com/d_attachment/toxins/toxins-12-00561/article_deploy/toxins-12-00561.pdf?version=1598862592

16. Hisbergues M, Christiansen G, Rouhiainen L, Sivonen K, B□rner T. PCR-based identification of microcystin-producing genotypes of different cyanobacterial genera. Archives of Microbiology. 2003 Dec 1;180(6):402–10.

17. Baliu-Rodriguez D, Peraino NJ, Premathilaka SH, Birbeck JA, Baliu-Rodriguez T, Westrick JA, et al. Identification of Novel Microcystins Using High-Resolution MS and MSn with Python Code. Environ Sci Technol. 2022 Feb 1;56(3):1652–63.

18. Tillett D, Dittmann E, Erhard M, von Döhren H, Börner T, Neilan BA. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. Chem Biol. 2000 Oct;7(10):753–64.

19. Rudolph-Böhner S, Mierke DF, Moroder L. Molecular structure of the cyanobacterial tumor-promoting microcystins. FEBS Letters. 1994;349(3):319–23.

20. Goldberg J, Huang H bin, Kwon Y guen, Greengard P, Nairn AC, Kuriyan J. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. Nature. 1995 Aug;376(6543):745–53.

21. Gulledge BM, Aggen JB, Eng H, Sweimeh K, Chamberlin AR. Microcystin analogues comprised only of adda and a single additional amino acid retain moderate activity as

PP1/PP2A inhibitors. Bioorganic & Medicinal Chemistry Letters. 2003 Sep 1;13(17):2907–11.

22. Toivola DM, Eriksson JE. Toxins Affecting Cell Signalling and Alteration of Cytoskeletal Structure. Toxicology in Vitro. 1999 Aug 1;13(4):521–30.

23. Campos A, Vasconcelos V. Molecular Mechanisms of Microcystin Toxicity in Animal Cells. International Journal of Molecular Sciences. 2010 Jan;11(1):268–87.

24. Chernoff N, Hill D, Lang J, Schmid J, Farthing A, Huang H. Dose–Response Study of Microcystin Congeners MCLA, MCLR, MCLY, MCRR, and MCYR Administered Orally to Mice. Toxins (Basel). 2021 Jan 24;13(2):86.

25. Chernoff N, Hill D, Lang J, Schmid J, Le T, Farthing A, et al. The Comparative Toxicity of 10 Microcystin Congeners Administered Orally to Mice: Clinical Effects and Organ Toxicity. Toxins. 2020 Jun;12(6):403.

26. Blom JF, Jüttner F. High crustacean toxicity of microcystin congeners does not correlate with high protein phosphatase inhibitory activity. Toxicon. 2005 Sep 15;46(4):465–70.

27. Fischer WJ, Altheimer S, Cattori V, Meier PJ, Dietrich DR, Hagenbuch B. Organic anion transporting polypeptides expressed in liver and brain mediate uptake of microcystin. Toxicology and Applied Pharmacology. 2005 Mar 15;203(3):257–63.

28. Niedermeyer THJ, Daily A, Swiatecka-Hagenbruch M, Moscow JA. Selectivity and Potency of Microcystin Congeners against OATP1B1 and OATP1B3 Expressing Cancer Cells. PLoS One. 2014 Mar 10;9(3):e91476.

29. Sallandt LL, Wolf CA, Schuster S, Enke H, Enke D, Wolber G, et al. Derivatization of Microcystins Can Increase Target Inhibition while Reducing Cellular Uptake. J Nat Prod. 2024 Oct 20;acs.jnatprod.4c00688.

30. Eriksson JE, Grönberg L, Nygård S, Slotte JP, Meriluoto JAO. Hepatocellular uptake of 3H-dihydromicrocystin-LR, a cyclic peptide toxin. Biochimica et Biophysica Acta (BBA) - Biomembranes. 1990 Jun 11;1025(1):60–6.

31. Herfindal L, Myhren L, Kleppe R, Krakstad C, Selheim F, Jokela J, et al. Nostocyclopeptide-M1: A Potent, Nontoxic Inhibitor of the Hepatocyte Drug Transporters OATP1B3 and OATP1B1. Mol Pharmaceutics. 2011 Apr 4;8(2):360–7.

32. König J, Seithel A, Gradhand U, Fromm MF. Pharmacogenomics of human OATP transporters. Naunyn Schmied Arch Pharmacol. 2006 Mar 1;372(6):432–43.

33. Hagenbuch B, Gui C. Xenobiotic transporters of the human organic anion transporting polypeptides (OATP) family. Xenobiotica. 2008 Aug;38(7–8):778–801.

34. Lad A, Breidenbach JD, Su RC, Murray J, Kuang R, Mascarenhas A, et al. As We Drink and Breathe: Adverse Health Effects of Microcystins and Other Harmful Algal Bloom Toxins in the Liver, Gut, Lungs and Beyond. Life (Basel). 2022 Mar 14;12(3):418.

35. Botha N, Gehringer MM, Downing TG, Venter M van de, Shephard EG. The role of microcystin-LR in the induction of apoptosis and oxidative stress in CaCo2 cells. Toxicon. 2004 Jan 1;43(1):85–92.

36. Breidenbach JD, French BW, Gordon TT, Kleinhenz AL, Khalaf FK, Willey JC, et al. Microcystin-LR aerosol induces inflammatory responses in healthy human primary airway epithelium. Environment International. 2022 Nov 1;169:107531.

37. Brittain SM; W Jim; Babcock Jackson, Lisa; Carmichael, Wayne W; Rinehart, Kenneth L; Culver, David A. Isolation and characterization of microcystins, cyclic heptapeptide hepatotoxins from a Lake Erie strain of microcystis aeruginosa. 2000;26:241–9.

38. McDonald K, DesRochers N, Renaud JB, Sumarah MW, McMullin DR. Metabolomics Reveals Strain-Specific Cyanopeptide Profiles and Their Production Dynamics in *Microcystis aeruginosa* and *M. flos-aquae*. Toxins. 2023 Mar 31;15(4):254.

39. Flores C, Caixach J. High Levels of Anabaenopeptins Detected in a Cyanobacteria Bloom from N.E. Spanish Sau-Susqueda-El Pasteral Reservoirs System by LC–HRMS. Toxins. 2020 Aug 22;12(9):541.

40. Earnshaw CD, McMullin DR. Cyanopeptolins and Anabaenopeptins Are the Dominant Cyanopeptides from Planktothrix Strains Collected in Canadian Lakes. Toxins. 2024 Feb 17;16(2):110.

41. Gkelis S, Lanaras T, Sivonen K. Cyanobacterial Toxic and Bioactive Peptides in Freshwater Bodies of Greece: Concentrations, Occurrence Patterns, and Implications for Human Health. Mar Drugs. 2015 Oct 12;13(10):6319–35.

42. Itou Y, Suzuki S, Ishida K, Murakami M. Anabaenopeptins G and H, potent carboxypeptidase A inhibitors from the cyanobacterium Oscillatoria agardhii (NIES-595). Bioorg Med Chem Lett. 1999 May 3;9(9):1243–6.

43. Sano T, Usui T, Ueda K, Osada H, Kaya K. Isolation of New Protein Phosphatase Inhibitors from Two Cyanobacteria Species, Planktothrix spp. J Nat Prod. 2001 Aug 1;64(8):1052–5.

44. Lenz KA, Miller TR, Ma H. Anabaenopeptins and cyanopeptolins induce systemic toxicity effects in a model organism the nematode *Caenorhabditis elegans*. Chemosphere. 2019 Jan 1;214:60–9.

45. Pawlik-Skowrońska B, Bownik A. Synergistic toxicity of some cyanobacterial oligopeptides to physiological activities of *Daphnia magna* (Crustacea). Toxicon. 2022 Jan 30;206:74–84.

46. Bownik A, Pawlik-Skowrońska B. Responses of RTgill-W1 cells to cyanobacterial metabolites microcystin-LR, anabaenopeptin-A, cylindrospermopsin, their binary and ternary mixtures. Toxicon. 2024 Oct;249:108059.

47. Skafi M, Vo Duy S, Munoz G, Dinh QT, Simon DF, Juneau P, et al. Occurrence of microcystins, anabaenopeptins and other cyanotoxins in fish from a freshwater wildlife reserve impacted by harmful cyanobacterial blooms. Toxicon. 2021 Apr 30;194:44–52.

48. Bittner M, Štern A, Smutná M, Hilscherová K, Žegura B. Cytotoxic and Genotoxic Effects of Cyanobacterial and Algal Extracts—Microcystin and Retinoic Acid Content. Toxins (Basel). 2021 Feb 2;13(2):107.

49. Jia J, Zhu F, Ma X, Cao ZW, Li YX, Chen YZ. Mechanisms of drug combinations: interaction and network perspectives. Nat Rev Drug Discov. 2009 Feb;8(2):111–28.

50. Jenkinson SE, Chung GW, van Loon E, Bakar NS, Dalzell AM, Brown CDA. The limitations of renal epithelial cell line HK-2 as a model of drug transporter expression and function in the proximal tubule. Pflugers Arch. 2012 Dec;464(6):601–11.

51. Tsujimoto M, Hatozaki D, Shima D, Yokota H, Furukubo T, Izumi S, et al. Influence of Serum in Hemodialysis Patients on the Expression of Intestinal and Hepatic Transporters for the Excretion of Pravastatin. Therapeutic Apheresis and Dialysis. 2012;16(6):580–7.

52. Komatsu M, Furukawa T, Ikeda R, Takumi S, Nong Q, Aoyama K, et al. Involvement of Mitogen-Activated Protein Kinase Signaling Pathways in Microcystin-LR–Induced Apoptosis after its Selective Uptake Mediated by OATP1B1 and OATP1B3. Toxicological Sciences. 2007 Jun 1;97(2):407–16.

53. Knowles BB, Howe CC, Aden DP. Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. Science. 1980 Jul 25;209(4455):497–9.

54. Lieber M, Todaro G, Smith B, Szakal A, Nelson-Rees W. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. Int J Cancer. 1976 Jan 15;17(1):62–70.

55. Ryan MJ, Johnson G, Kirk J, Fuerstenberg SM, Zager RA, Torok-Storb B. HK-2: an immortalized proximal tubule epithelial cell line from normal adult human kidney. Kidney Int. 1994 Jan;45(1):48–57.

56. Yancey CE, Kiledal EA, Chaganti SR, Denef VJ, Errera RM, Evans JT, et al. The Western Lake Erie culture collection: A promising resource for evaluating the physiological and genetic diversity of *Microcystis* and its associated microbiome. Harmful Algae. 2023 Jul;126:102440.

57. Zheng S, Wang W, Aldahdooh J, Malyutina A, Shadbahr T, Tanoli Z, et al. SynergyFinder Plus: Toward Better Interpretation and Annotation of Drug Combination Screening Datasets. Genomics Proteomics Bioinformatics. 2022 Jun;20(3):587–96.

58. Bliss CI. The Toxicity of Poisons Applied Jointly. Annals of Applied Biology. 1939 Aug;26(3):585–615.

59. Loewe S. The problem of synergism and antagonism of combined drugs. Arzneimittelforschung. 1953 Jun;3(6):285–90.

60. Yadav B, Wennerberg K, Aittokallio T, Tang J. Searching for Drug Synergy in Complex Dose–Response Landscapes Using an Interaction Potency Model. Computational and Structural Biotechnology Journal. 2015;13:504–13.

61. Lehár J, Zimmermann GR, Krueger AS, Molnar RA, Ledell JT, Heilbut AM, et al. Chemical combination effects predict connectivity in biological systems. Molecular Systems Biology. 2007 Jan;3(1):80.

62. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2013. Available from: http://www.R-project.org/

63. Kassambara A. R Package: ggpubr:"ggplot2" based publication ready plots. 2018.

64. Aggen JB, Humphrey JM, Gauss CM, Huang HB, Nairn AC, Chamberlin AR. The design, synthesis, and biological evaluation of analogues of the serine-threonine protein phosphatase 1 and 2A selective inhibitor microcystin LA: rational modifications imparting PP1 selectivity. Bioorg Med Chem. 1999 Mar;7(3):543–64.

65. Chen YM, Lee TH, Lee SJ, Huang HB, Huang R, Chou HN. Comparison of protein phosphatase inhibition activities and mouse toxicities of microcystins. Toxicon. 2006 Jun 1;47(7):742–6.

66. Chorus I, Welker M, editors. Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management. Taylor & Francis; 2021.

67. Fischer A, Hoeger SJ, Stemmer K, Feurstein DJ, Knobeloch D, Nussler A, et al. The role of organic anion transporting polypeptides (OATPs/SLCOs) in the toxicity of different microcystin congeners in vitro: A comparison of primary human hepatocytes and OATP-transfected HEK293 cells. Toxicology and Applied Pharmacology. 2010 May 15;245(1):9–20.

68. Chaffin JD, Westrick JA, Reitz LA, Bridgeman TB. Microcystin congeners in Lake Erie follow the seasonal pattern of nitrogen availability. Harmful Algae. 2023 Aug 1;127:102466.

69. DepMap B. DepMap 24Q4 Public [Internet]. Figshare+; 2024 [cited 2025 Jun 16]. p. 30825074613 Bytes. Available from: https://plus.figshare.com/articles/dataset/DepMap_24Q4_Public/27993248/1

70. Wang H, Xu K, Wang B, Liu J, Wang X, Xing M, et al. Microcystin-LR induces a wide variety of biochemical changes in the A549 human non-small cell lung cancer cell line:

Roles for protein phosphatase 2A and its substrates. Environmental Toxicology. 2017;32(3):1065–78.

71. Ikehara T, Nakashima J, Nakashima S, Yasumoto T. Different responses of primary normal human hepatocytes and human hepatoma cells toward cyanobacterial hepatotoxin microcystin-LR. Toxicon. 2015 Oct 1;105:4–9.

72. Ulvestad M, Nordell P, Asplund A, Rehnström M, Jacobsson S, Holmgren G, et al. Drug metabolizing enzyme and transporter protein profiles of hepatocytes derived from human embryonic and induced pluripotent stem cells. Biochemical Pharmacology. 2013 Sep 1;86(5):691–702.

73. Gan N, Sun X, Song L. Activation of Nrf2 by Microcystin-LR Provides Advantages for Liver Cancer Cell Growth. Chem Res Toxicol. 2010 Sep 20;23(9):1477–84.

74. Roegner AF, Puschner B. Aggregate culture: A more accurate predictor of microcystin toxicity for risk assessment. Toxicon. 2014 Jun 1;83:1–14.

75. Chowdhury RR, Rose S, Ezan F, Sovadinová I, Babica P, Langouët S. Hepatotoxicity of cyanotoxin microcystin-LR in human: Insights into mechanisms of action in the 3D culture model Hepoid-HepaRG. Environmental Pollution. 2024 Feb 1;342:123047.

76. Yi X, Xu S, Huang F, Wen C, Zheng S, Feng H, et al. Effects of Chronic Exposure to Microcystin-LR on Kidney in Mice. International Journal of Environmental Research and Public Health. 2019 Jan;16(24):5030.

77. Alverca E, Andrade M, Dias E, Sam Bento F, Batoréu MCC, Jordan P, et al. Morphological and ultrastructural effects of microcystin-LR from *Microcystis aeruginosa* extract on a kidney cell line. Toxicon. 2009 Sep 1;54(3):283–94.

78. Teneva I, Klaczkowska D, Batsalova T, Kostova Z, Dzhambazov B. Influence of captopril on the cellular uptake and toxic potential of microcystin-LR in non-hepatic adhesive cell lines. Toxicon. 2016 Mar;111:50–7.

79. Halland N, Brönstrup M, Czech J, Czechtizky W, Evers A, Follmann M, et al. Novel Small Molecule Inhibitors of Activated Thrombin Activatable Fibrinolysis Inhibitor (TAFIa) from Natural Product Anabaenopeptin. J Med Chem. 2015 Jun 11;58(11):4839–44.

80. Shi JH, Birbeck JA, Olson NE, Parham RL, Holen AL, Ledsky IR, et al. Detection of Aerosolized Anabaenopeptins from Cyanobacterial Harmful Algal Blooms in Atmospheric Particulate Matter. ACS Earth Space Chem. 2025 Feb 25;

81. Ciută AD, Nosol K, Kowal J, Mukherjee S, Ramírez AS, Stieger B, et al. Structure of human drug transporters OATP1B1 and OATP1B3. Nat Commun. 2023 Sep 18;14(1):5774.

82. Schreuder H, Liesum A, Lönze P, Stump H, Hoffmann H, Schiell M, et al. Isolation, Co-Crystallization and Structure-Based Characterization of Anabaenopeptins as Highly

Potent Inhibitors of Activated Thrombin Activatable Fibrinolysis Inhibitor (TAFIa). Sci Rep. 2016 Sep 8;6:32958.

83. Feurstein D, Kleinteich J, Heussner AH, Stemmer K, Dietrich DR. Investigation of Microcystin Congener–Dependent Uptake into Primary Murine Neurons. Environ Health Perspect. 2010 Oct;118(10):1370–5.

84. Ozbek E. Induction of Oxidative Stress in Kidney. International Journal of Nephrology. 2012;2012(1):465897.

85. Mestareehi A, Abu-Farsakh N. Impact of Protein Phosphatase Expressions on the Prognosis of Hepatocellular Carcinoma Patients. ACS Omega. 2024 Mar 5;9(9):10299–331.

86. Le Manach S, Duval C, Marie A, Djediat C, Catherine A, Edery M, et al. Global metabolomic characterizations of *Microcystis* spp. highlights clonal diversity in natural bloom-forming populations and expands metabolite structural diversity. Front Microbiol. 20190416th ed. 2019;10:791.

87. Yancey C. The Biosynthetic Repertoire of *Microcystis* spp. in Western Lake Erie Harmful Algal Blooms: Insights into Biosynthesis of Known and Novel Secondary Metabolites in Natural Populations using a `Multi-Omic' Approach [Internet] [Thesis]. University of Michigan; 2023 [cited 2025 Apr 9]. Available from: http://deepblue.lib.umich.edu/handle/2027.42/178036

88. Lifshits M, Zafrir-Ilan E, Raveh A, Carmeli S. Protease inhibitors from three fishpond water blooms of *Microcystis* spp. Tetrahedron. 2011 Jun 3;67(22):4017–24.

89. May DS, Chen WL, Lantvit DD, Zhang X, Krunic A, Burdette JE, et al. Merocyclophanes C and D from the Cultured Freshwater Cyanobacterium *Nostoc* sp. (UIC 10110). J Nat Prod. 2017 Apr 28;80(4):1073–80.

90. Ma S, Weng M, Yang T, Ge L, Yang K. Triterpenes and Pheophorbides from Camellia ptilosperma and Their Cytotoxicity, Photocytotoxicity, and Photodynamic Antibacterial Activity. Molecules. 2023 Oct 12;28(20):7058.

91. Hayakawa K, Tsujimura S, Napolitano GE, Nakano S, Kumagai M, Nakajima T, et al. Fatty acid composition as an indicator of physiological condition of the cyanobacterium Microcystis aeruginosa. Limnology. 2002 Apr 1;3(1):29–35.

92. Maulucci G, Cohen ,O., Daniel ,B., Sansone ,A., Petropoulou ,P. I., Filou ,S., et al. Fatty acid-related modulations of membrane fluidity in cells: detection and implications. Free Radical Research. 2016 Nov 1;50(sup1):S40–50.

93. Freitas EC, Pinheiro C, Rocha O, Loureiro S. Can mixtures of cyanotoxins represent a risk to the zooplankton? The case study of *Daphnia magna* Straus exposed to hepatotoxic and neurotoxic cyanobacterial extracts. Harmful Algae. 2014 Jan 1;31:143–52.

94. Lopes VR, Schmidtke M, Helena Fernandes M, Martins R, Vasconcelos V. Cytotoxicity in L929 fibroblasts and inhibition of herpes simplex virus type 1 Kupka by estuarine cyanobacteria extracts. Toxicology in Vitro. 2011 Jun 1;25(4):944–50. 95. Okumura DT, Sotero-Santos RB, Takenaka RA, Rocha O. Evaluation of cyanobacteria toxicity in tropical reservoirs using crude extracts bioassay with cladocerans. Ecotoxicology. 2007 Mar;16(2):263–70.

95. Okumura DT, Sotero-Santos RB, Takenaka RA, Rocha O. Evaluation of cyanobacteria toxicity in tropical reservoirs using crude extracts bioassay with cladocerans. Ecotoxicology. 2007 Mar;16(2):263–70.

96. Štěpánková T, Ambrožová L, Bláha L, Giesy JP, Hilscherová K. *In vitro* modulation of intracellular receptor signaling and cytotoxicity induced by extracts of cyanobacteria, complex water blooms and their fractions. Aquatic Toxicology. 2011 Oct 1;105(3):497–507.

4.8 Figures

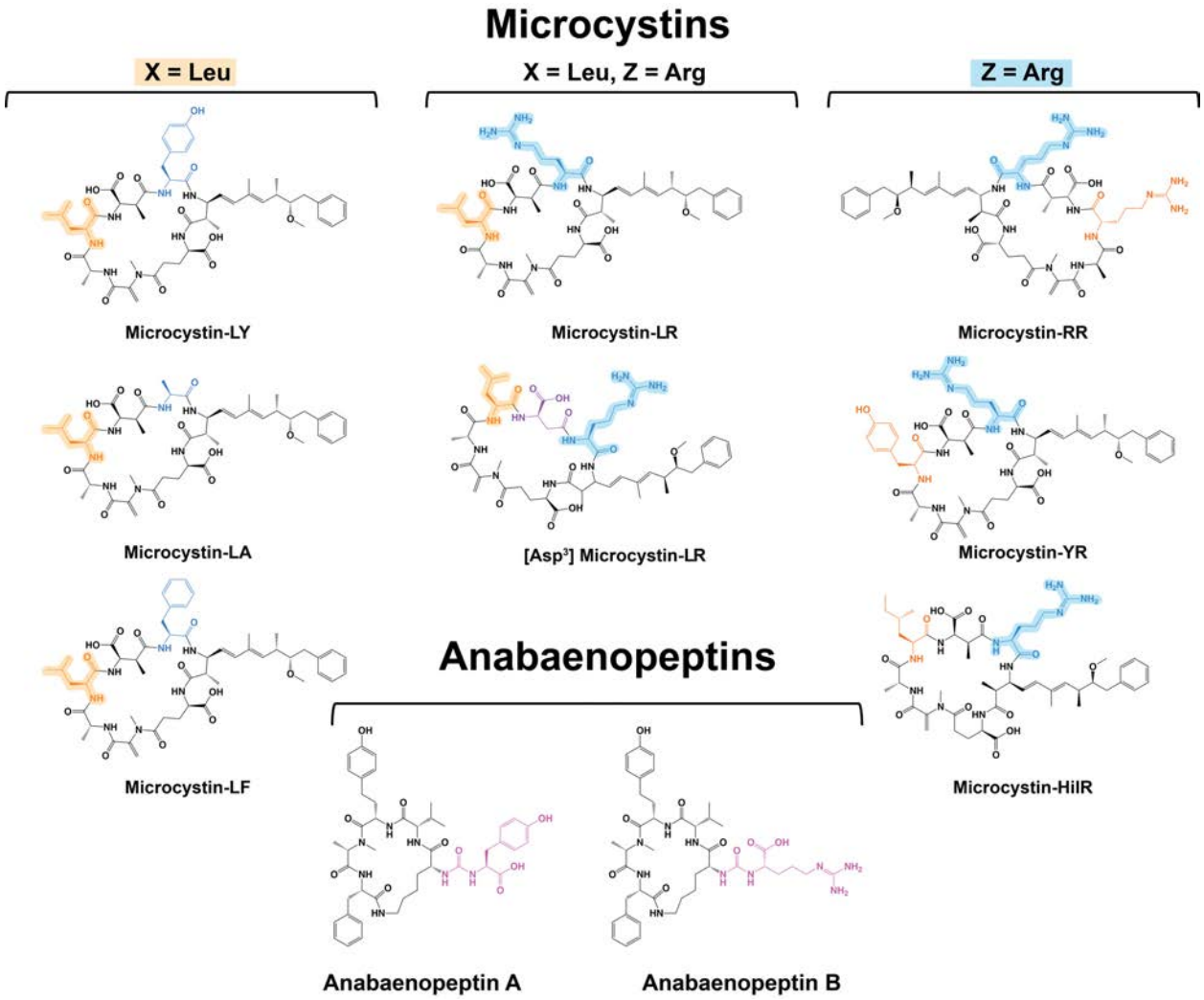


Figure 4.8.1: Chemical structures of microcystins (LY, LA, LF, LR, Asp^3-LR, RR, YR, and HilR) and anabaenopeptins A and B. Microcystins are grouped based on amino acid substitutions at the variable X and Z positions: leucine (Leu) residues at position X are highlighted in orange, and arginine (Arg) residues at position Z are highlighted in blue. Structural differences between anabaenopeptins A and B are shown in magenta.

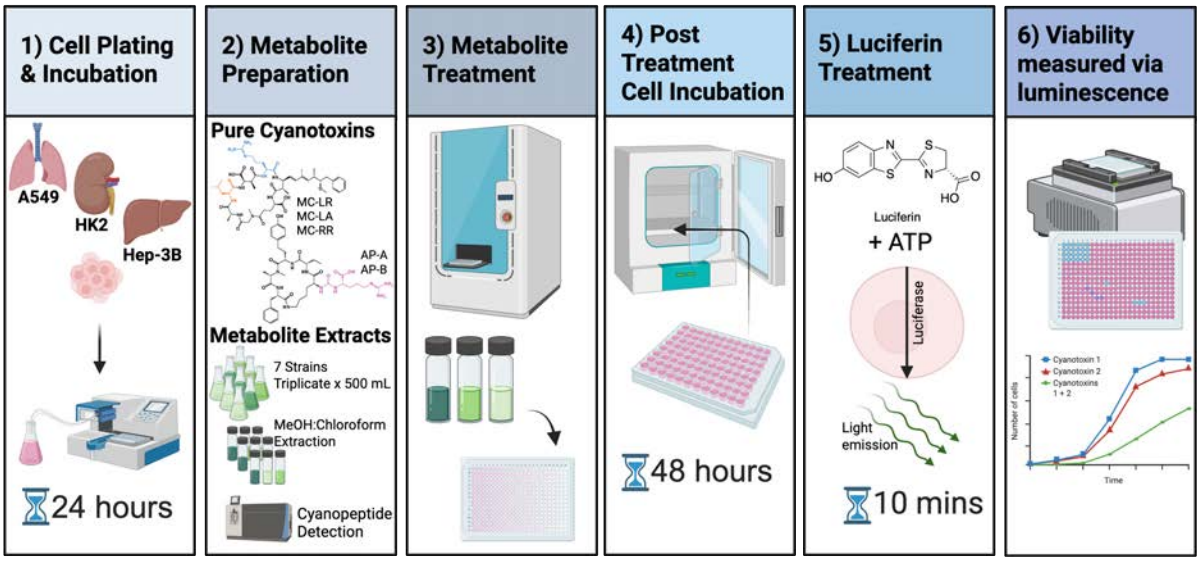


Figure 4.8.2: Workflow for preparing cyanotoxin standards and culture-derived metabolite extracts and assessing their effects on cell viability in human cell lines using the CellTiter-Glo assay. Pure standards (MC-LR, MC-LA, MC-RR, AP-A, AP-B) and extracts from seven WLECC *Microcystis* isolates were prepared in biological triplicate. After metabolite extraction and cyanotoxin quantification, human cell lines (A549, HK-2, and Hep-3B) were plated and incubated for 24 hours, followed by toxin treatment and a 48-hour incubation. Cell viability was measured via luminescence.

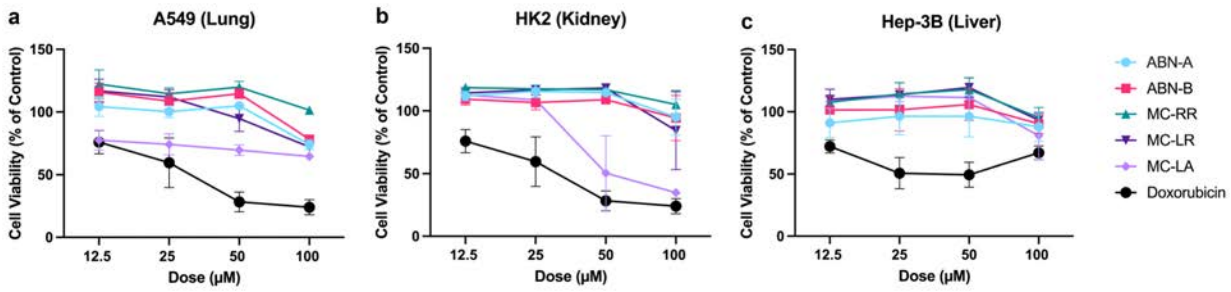


Figure 4.8.3: Linear dose-dependent cell viability response by human **(a)** lung (A549), **(b)** kidney (HK2), and **(c)** liver (Hep-3B) cells to pure MC and AP congeners. Viability is shown relative to DMSO (negative control) and compared to doxorubicin (positive control, data shown in black).

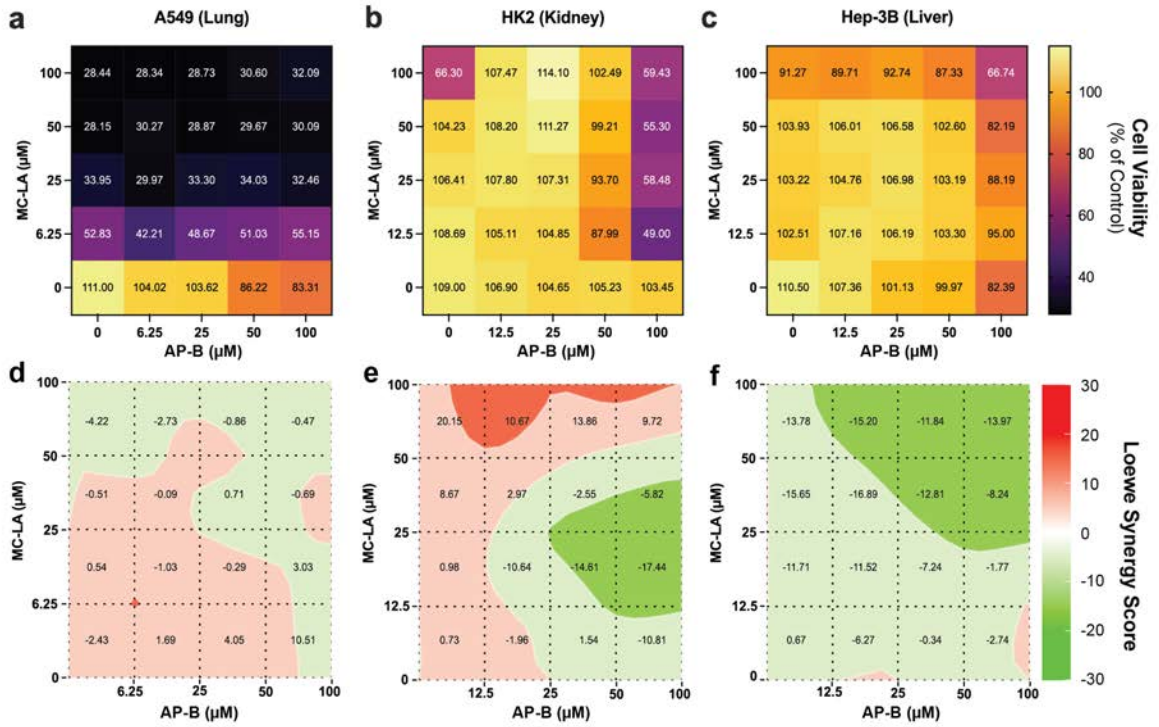


Figure 4.8.4: Dose-dependent cell viability response by human **(a)** lung (A549), **(b)** kidney (HK2), and **(c)** liver (Hep-3B) cells to combinations of MC-LA and AP-B. Viability is shown relative to DMSO (negative control), with lighter tiles indicating more cell viability and darker tiles indicating lower cell viability. 2D contour plots show Loewe synergy scores for MC-LA and AP-B combinations in **(d)** A549, **(e)** HK2, and **(f)** Hep-3B cells. Scores <1 indicate synergistic effects (green), ~1 indicate additive effects (white), and >1 indicate antagonistic effects (red). Actual Loewe scores are displayed within each plot. Mean Loewe scores for each cell line are: A549: 0.45, HK2: -0.48, and Hep-3B: -9.33.

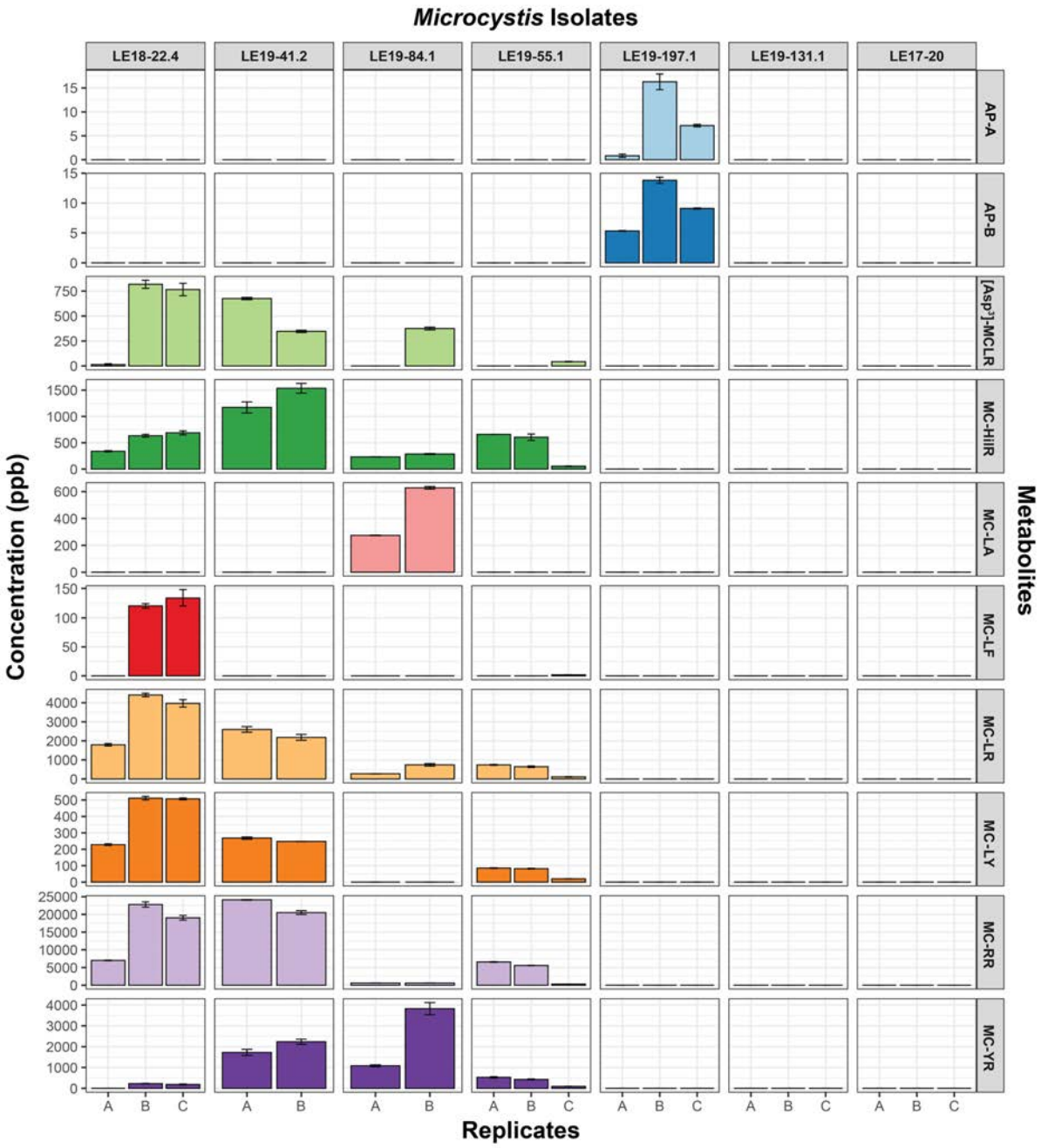


Figure 4.8.5: Concentrations of select congeners of microcystins and anabaenopeptins in crude extracts from western Lake Erie culture collection *Microcystis* cultures used in human cell line exposures. Biological duplicates or triplicates are shown for each isolate. Bars represent the mean concentration in parts per billion (ppb) for each cyanotoxin, with standard error indicated with error bars.

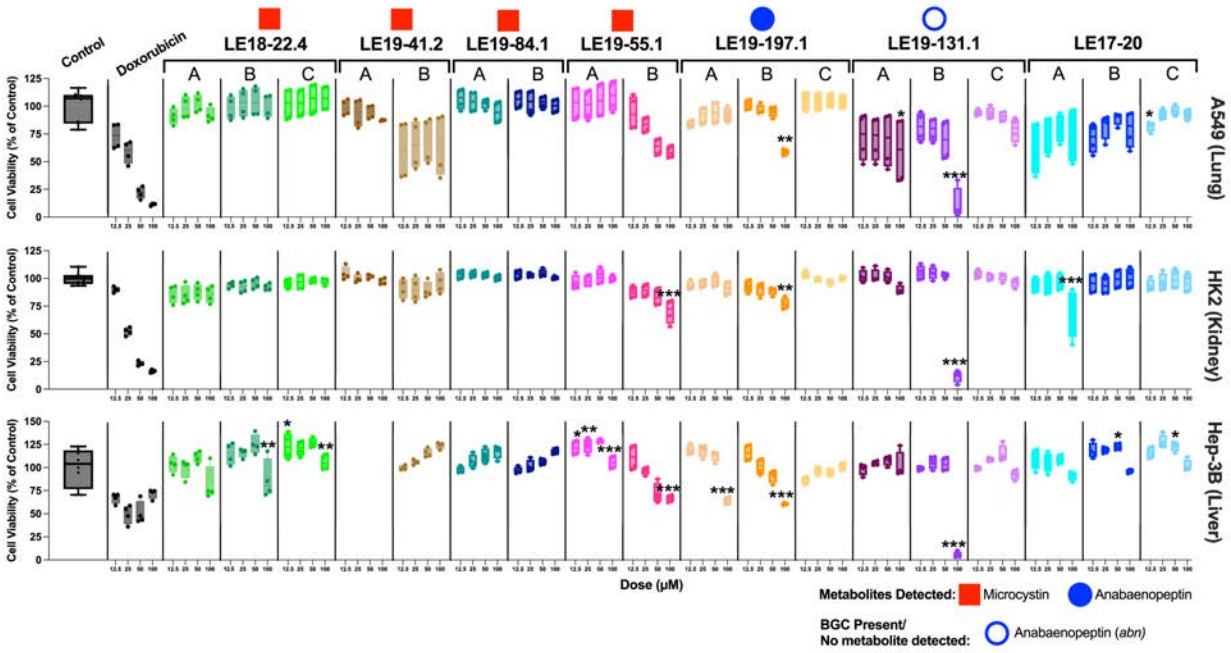


Figure 4.8.6: Dose-dependent cell viability response by human lung (A549), kidney (HK2), and liver (Hep-3B) cells to *Microcystis* isolate extracts from WLECC. Extracts were dosed between 12.5 and 100 μM, based on the crude dry weight of extracted metabolites. Each strain was grown in biological duplicate or triplicate, and cell viability was assessed in technical quadruples across two plates. Viability is shown relative to DMSO (negative vehicle control). A red square indicates the detection of microcystins in the extract, and a blue circle indicates the detection of anabaenopeptins in the extract. A unfilled blue circle indicates the presence of the anabaenopeptin biosynthetic gene cluster (BGCs) but no anabaenopeptin A or B detection. Significant effects were determined by Tukey's multiple comparisons post-hoc tests against the vehicle control (Significance: $P < 0.01$ (*), $P < 0.001$ (**), $P < 0.0001$ (***).

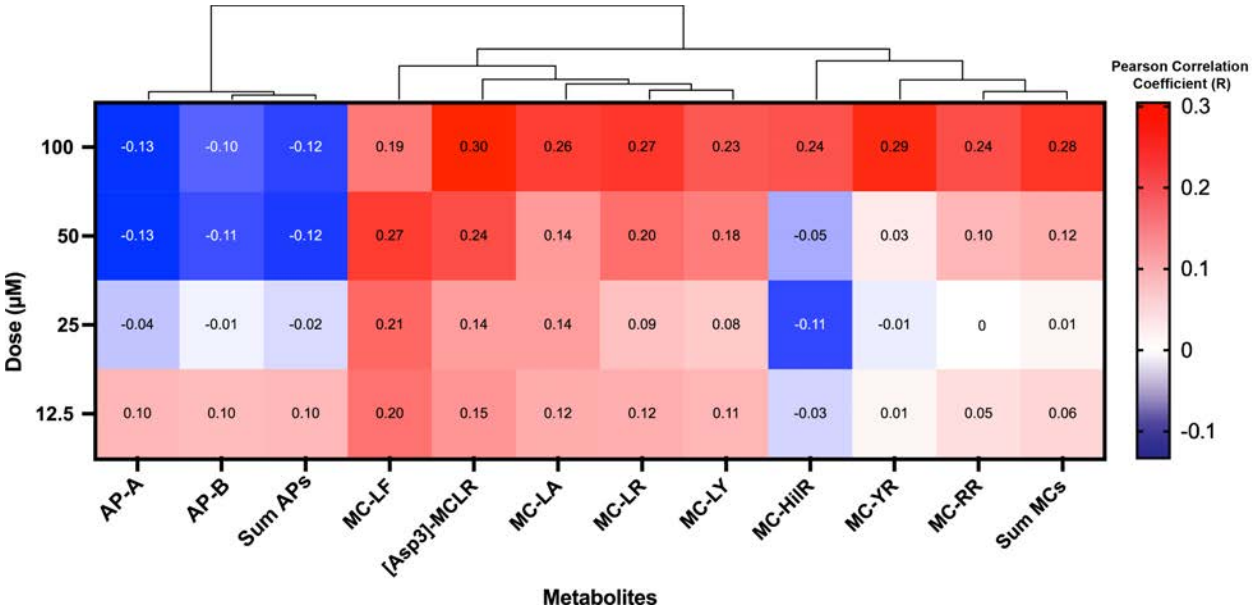


Figure 4.8.7: Pearson correlation coefficients (R) between cell viability effects exhibited by WLECC *Microcystis* isolates at four doses between 12.5 and 100 µM and cyanotoxin concentrations (MC and AP congeners) in each extract. Each tile shows the correlation coefficient (R), colored by strength and direction (blue (negative) to red (positive)). "Sum APs" indicate the sum of all AP congeners measured in the extract, and "Sum MCs" indicate the sum of all MC congeners measured in the extract.

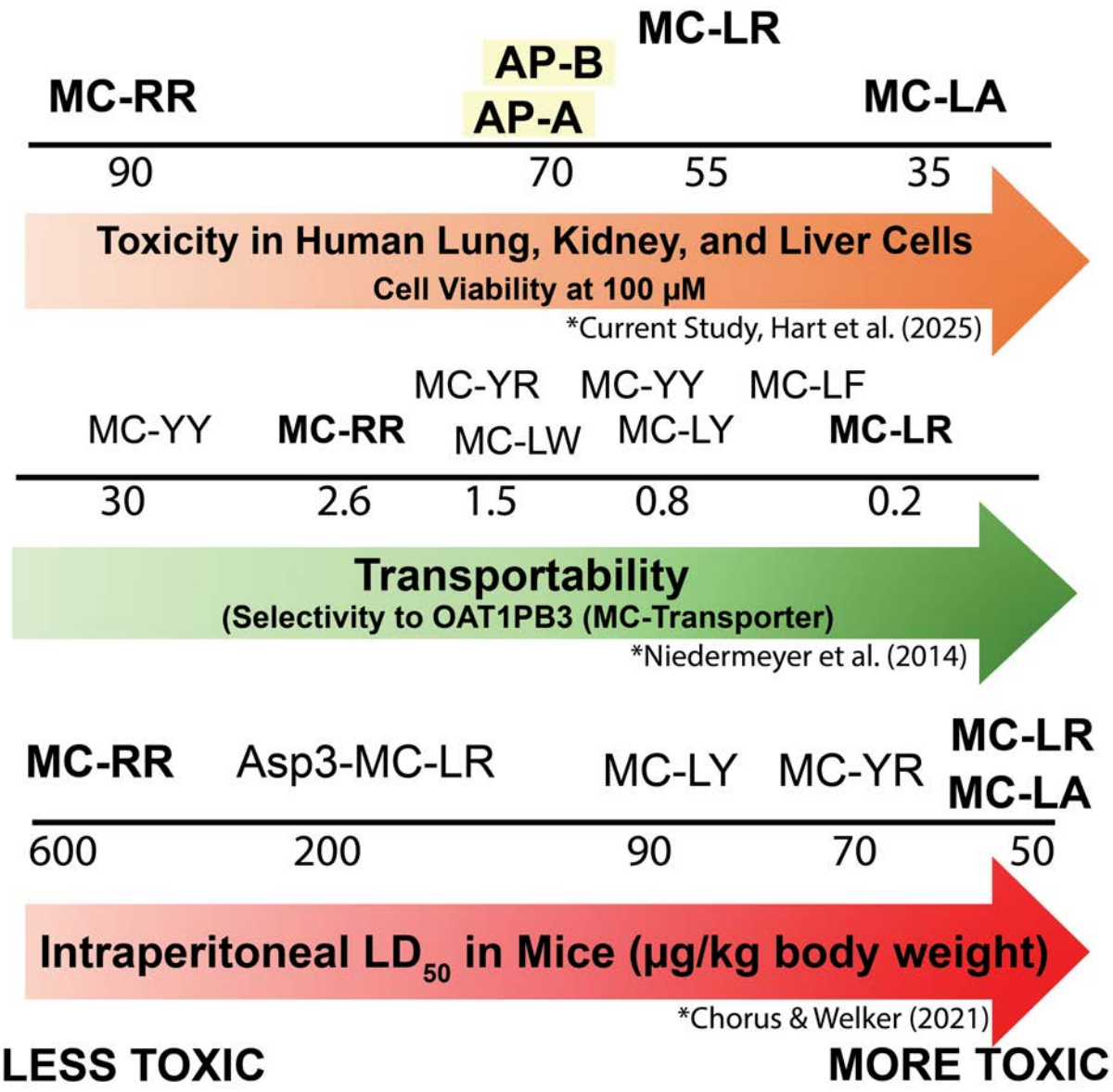


Figure 4.8.8: Comparative toxicity and transportability of microcystin and anabaenopeptin congeners. Percent cell viability at 100 µM for three microcystin and two anabaenopeptin congeners in human lung (A549), liver (HK2), and kidney (Hep-3B) cells from the current study (orange arrow). OATP1B3 selectivity, a key transporter for MC-LR, from Niedermeyer et al. (2014) is indicated with a green arrow for eight microcystin congeners. Intraperitoneal LD_{50} values (µg/kg) in murine models for six microcystin congeners are shown as summarized in Chorus and Welker (2021) (red arrow). Congeners used in the current study are bolded. Congeners on the left tend to be less toxic across metrics, while those on the right tend to be more toxic. Protein phosphatase 1 and 2 inhibition across these congeners do not vary significantly (Fischer et al. (2010) and Blom and Jüttner (2005)).

4.9 Supplementary Information

4.9.1 SI Tables

Table S4.9.1: Single cyanotoxin exposure viability data.

Hep3B																								
Dose	ABN-A				ABN-B				MC-RR				MC-LR				MC-LA				Doxorubicin			
100	94.8991	100.6	77.63806	78.48286	100.0256	91.06593	81.5142	88.7005	96.25415	105.2903	94.71253	85.22457	97.19833	95.63908	84.358	95.62013	100.0256	91.06593	75.68429	56.3517	59.17326	70.92256	71.14921	67.15919
50	104.0037	115.9318	83.16112	81.90772	111.8007	119.8969	90.82131	100.7731	119.6529	129.6738	115.5995	106.3837	124.2747	126.3653	108.9507	118.4044	111.8007	119.8969	108.4609	105.5707	47.66876	55.13896	36.12699	58.96524
25	106.2141	111.2572	81.83865	85.83997	118.1106	114.2213	86.68059	87.15323	121.5209	122.0308	110.2146	103.434	117.3484	124.8923	102.1249	110.2992	118.1106	114.2213	108.059	109.9622	43.38883	47.98089	69.12427	42.05404
12.5	97.48319	103.4387	77.35171	86.19748	108.3512	112.4005	88.36099	96.61565	109.1087	113.4362	104.7142	102.981	112.9049	119.3107	100.1154	107.3051	108.3512	112.4005	104.8368	107.9012	64.29916	74.15345	75.2196	74.95753
HK2																								
Dose	ABN-A				ABN-B				MC-RR				MC-LR				MC-LA				Doxorubicin			
100	79.19804	87.24676	109.6781	104.1664	98.98348	67.83928	103.2945	107.2818	88.54637	106.9997	112.3903	112.3107	43.79558	76.49427	109.3519	107.7826	31.77949	34.91996	34.42428	38.1613	18.92827	22.86942	32.71507	21.90304
50	112.016	111.9565	118.1193	116.2165	107.7493	105.6424	111.5822	110.7286	114.7376	114.8918	119.6843	118.4681	114.2295	119.5633	117.1793	121.8475	25.5038	75.41494	23.94168	77.13105	30.18226	32.74689	16.6802	33.70719
25	116.7798	109.9765	122.1048	113.6862	100.1027	109.6881	103.929	112.8896	115.7718	117.2371	119.1621	118.2429	118.8682	116.8623	116.1298	115.2881	104.7134	110.6922	106.3638	113.4161	67.21748	68.69332	29.98034	72.16463
12.5	112.0904	107.4007	116.8076	111.2245	103.5015	111.6523	107.807	114.0698	116.9291	119.2808	118.2423	119.9468	112.5215	114.1344	113.3926	117.262	106.8472	115.2176	107.916	120.8242	77.13615	82.56828	62.49356	81.69837
A549																								
Dose	ABN-A				ABN-B				MC-RR				MC-LR				MC-LA				Doxorubicin			
100	72.87601	78.87195	69.65161	74.90068	77.09696	78.52737	81.0311	76.78625	102.3295	101.921	100.8185	100.6216	73.14267	74.28958	72.7363	68.7561	67.15112	66.01378	63.03524	62.21739	18.92827	22.86942	32.71507	21.90304
50	103.4461	105.5641	102.8764	107.7193	118.6495	117.1079	111.0182	111.6937	123.3148	124.345	115.7554	115.7473	102.5219	103.7532	89.51189	82.90777	72.98901	73.48207	66.59413	65.30455	30.18226	32.74689	16.6802	33.70719
25	93.88906	100.0812	101.4097	105.2235	114.372	113.7168	101.2083	104.9747	116.4293	120.2139	110.9519	110.649	114.9704	118.3069	109.1265	105.2376	82.5204	80.11033	67.7981	66.28537	67.21748	68.69332	29.98034	72.16463
12.5	98.26342	106.6099	97.66809	114.5442	118.676	124.8606	109.0205	110.7892	131.8697	131.8748	109.6295	115.248	126.6264	123.0576	110.6577	107.1401	84.61535	84.20692	70.80704	70.54478	77.13615	82.56828	62.49356	81.69837

Table S4.9.2: Combination cyanotoxin exposure viability data.

A549	100	50	25	6.25	0	*ABN-B Down, MC-LA Across		
100	32.09093	30.08602	32.46468	55.15381	83.31326			
50	30.6014	29.6728	34.03422	51.03451	86.21544			
25	28.7256	28.87371	33.30115	48.67361	103.6216			
6.25	28.33911	30.26903	29.97384	42.21327	104.0169			
0	28.43981	28.15026	33.94773	52.82926	111			
HK2	100	50	25	12.5	0	*ABN-B Down, MC-LA Across		
100	59.43073	102.4863	114.099	107.4706	66.30062			
50	55.29857	99.21303	111.2715	108.1989	104.2317			
25	58.48274	93.69752	107.3067	107.7965	106.411			
12.5	49.0032	87.99492	104.8512	105.107	108.6911			
0	103.4486	105.2258	104.6498	106.9003	109			
Hep3B	100	50	25	12.5	0	*ABN-B Down, MC-LA Across		
100	66.7423	87.32582	92.74121	89.71163	91.26718			
50	82.19336	102.5987	106.5796	106.0064	103.9263			
25	88.19065	103.1921	106.9827	104.7632	103.2235			
12.5	94.99565	103.2985	106.1882	107.1571	102.5141			
0	82.39301	99.97144	101.132	107.3555	110.5			

Table S4.9.3: Combination cyanotoxin exposure Loewe, Bliss, HAS, and ZIP synergy scores.

A549					
conc1_ABNB	**conc2_MCLA**	**ZIP_synergy**	**HSA_synergy**	**Loewe_synergy**	**Bliss_synergy**
100	100	-8.5495367	-3.651127	-4.2242833	-8.3968048
50	100	-6.359725	-2.1615959	-2.7347521	-6.0818984
25	100	0.20402213	-0.2857894	-0.8589457	0.74419931
6.25	100	1.04626216	0.10069481	-0.4724615	1.24309426
0	100	0	0	0	0
100	50	-5.9282808	-1.9357637	-0.5053398	-6.6331251

50	50	-4.8656362	-1.522539	-0.092115	-5.4029283
25	50	0.82441851	-0.7234567	0.70696723	0.29604557
6.25	50	2.20729502	-2.1187692	-0.6883453	-0.9880007
0	50	0	0	0	0
100	25	-5.3443627	1.48305489	0.54406731	-4.1817159
50	25	-5.0248374	-0.0864889	-1.0254765	-4.7660343
25	25	1.41098833	0.64657767	-0.2924099	1.87604377
6.25	25	4.66779464	3.97389352	3.03490594	5.33754087
0	25	0	0	0	0
100	6.25	-9.8959655	-2.3245459	-2.433714	-11.140029
50	6.25	-5.6564567	1.7947492	1.6855811	-5.4875322
25	6.25	6.32529603	4.15564556	4.04647746	6.06893356
6.25	6.25	11.9154569	10.6159906	10.5068225	12.7380907
0	6.25	0	0	0	0
100	0	0	0	0	0
50	0	0	0	0	0
25	0	0	0	0	0
6.25	0	0	0	0	0
0	0	0	0	0	0
			Mean	0.44981115	
HK2					
conc1_ABNB	conc2_MCLA	ZIP_synergy	HSA_synergy	Loewe_synergy	Bliss_synergy
100	100	9.53764021	6.86988527	9.72387129	9.1563159
50	100	13.0339542	11.0020476	13.8560336	14.4667811
25	100	12.0568876	7.81787506	10.6718611	10.9007511
12.5	100	17.7460108	17.2974153	20.1514013	21.8723814
0	100	0	0	0	0
100	50	3.79123494	0.9622991	-5.8227263	5.33988771
50	50	8.56756041	5.01862218	-2.5494786	10.4655535
25	50	13.7773384	10.5341341	2.96603337	15.3807448
12.5	50	18.2055059	16.2367399	8.66863913	23.4290756
0	50	0	0	0	0
100	25	8.84834067	-10.650397	-17.435422	-4.01835
50	25	8.54279995	-6.0456708	-14.607909	0.70031253
25	25	9.84344705	-2.6568756	-10.643165	4.05218391
12.5	25	11.8647566	1.5597115	0.98195824	8.90242654
0	25	0	0	0	0
100	12.5	9.35488868	-4.0219808	-10.807006	4.96888394
50	12.5	9.18882052	-2.9731107	-11.535349	6.17221413
25	12.5	10.5628678	-3.1466107	-1.9632494	5.94865762
12.5	12.5	11.1627463	1.7933105	0.72617984	11.0841723

0	12.5	0	0	0	0
100	0	0	0	0	0
50	0	0	0	0	0
25	0	0	0	0	0
12.5	0	0	0	0	0
0	0	0	0	0	0
			Mean	-0.4761455	
Hep3B					
conc1_ABNB	conc2_MCLA	ZIP_synergy	HSA_synergy	Loewe_synergy	Bliss_synergy
100	100	9.14449481	15.6507144	-13.782953	8.45548041
50	100	3.37187578	0.19965257	-15.199534	3.43462713
25	100	2.1971999	-5.7976393	-11.840505	-3.1416838
12.5	100	-0.7956627	-12.602642	-13.967348	-10.531174
0	100	0	0	0	0
100	50	1.56204024	3.94135698	-15.653838	3.91528671
50	50	-1.044326	-2.6272441	-16.892214	1.29790744
25	50	0.89493031	-3.2206585	-12.807468	0.00194109
12.5	50	1.71479903	-3.3270726	-8.2357215	-0.8136591
0	50	0	0	0	0
100	25	-1.3278956	-1.4740314	-11.712905	-0.4409131
50	25	4.45932765	-5.4476094	-11.516794	-1.4768921
25	25	7.19719497	-5.8507766	-7.2418013	-2.5907669
12.5	25	8.06452318	-5.0562277	-1.769092	-2.5136368
0	25	0	0	0	0
100	6.25	8.24241231	1.55554673	0.67299351	8.26867737
50	6.25	6.63722372	-2.0801721	-6.2654985	5.56409449
25	6.25	8.64429932	-1.539637	-0.3440504	6.05293885
12.5	6.25	9.44118115	-4.6429476	-2.7379722	2.89744937
0	6.25	0	0	0	0
100	0	0	0	0	0
50	0	0	0	0	0
25	0	0	0	0	0
12.5	0	0	0	0	0
0	0	0	0	0	0
			Mean	-9.3309188	

Table S4.9.4: *Microcystis* isolate extract exposure viability data.

A549						
dose	**culture_name**	**culture**	**viability**	**bio_rep**	**plate**	**cell_line**

100	LE18-22.4	C1A	94.9251145	A	1	A549
50	LE18-22.4	C1A	110.709781	A	1	A549
25	LE18-22.4	C1A	109.088417	A	1	A549
12.5	LE18-22.4	C1A	96.3857236	A	1	A549
100	LE18-22.4	C1A	105.937581	A	1	A549
50	LE18-22.4	C1A	116.802012	A	1	A549
25	LE18-22.4	C1A	114.988093	A	1	A549
12.5	LE18-22.4	C1A	103.766844	A	1	A549
100	LE18-22.4	C1B	113.56696	B	1	A549
50	LE18-22.4	C1B	123.170574	B	1	A549
25	LE18-22.4	C1B	116.220619	B	1	A549
12.5	LE18-22.4	C1B	109.00793	B	1	A549
100	LE18-22.4	C1B	114.413499	B	1	A549
50	LE18-22.4	C1B	118.70848	B	1	A549
25	LE18-22.4	C1B	120.837547	B	1	A549
12.5	LE18-22.4	C1B	114.821417	B	1	A549
100	LE18-22.4	C1C	123.067718	C	1	A549
50	LE18-22.4	C1C	124.771104	C	1	A549
25	LE18-22.4	C1C	117.611926	C	1	A549
12.5	LE18-22.4	C1C	119.209166	C	1	A549
100	LE18-22.4	C1C	121.883659	C	1	A549
50	LE18-22.4	C1C	123.58573	C	1	A549
25	LE18-22.4	C1C	120.252427	C	1	A549
12.5	LE18-22.4	C1C	119.891661	C	1	A549
100	LE19-55.1	C2A	124.2952	A	1	A549
50	LE19-55.1	C2A	123.137458	A	1	A549
25	LE19-55.1	C2A	119.387685	A	1	A549
12.5	LE19-55.1	C2A	121.522235	A	1	A549
100	LE19-55.1	C2A	128.123706	A	1	A549
50	LE19-55.1	C2A	125.800767	A	1	A549
25	LE19-55.1	C2A	116.157458	A	1	A549
12.5	LE19-55.1	C2A	120.721751	A	1	A549
100	LE19-55.1	C2B	67.4202669	B	1	A549
50	LE19-55.1	C2B	73.8818171	B	1	A549
25	LE19-55.1	C2B	91.7299774	B	1	A549
12.5	LE19-55.1	C2B	115.483516	B	1	A549
100	LE19-55.1	C2B	63.5726805	B	1	A549
50	LE19-55.1	C2B	72.1915898	B	1	A549
25	LE19-55.1	C2B	94.0923919	B	1	A549
12.5	LE19-55.1	C2B	109.022624	B	1	A549
100	LE18-22.4	C1A	89.7857291	A	2	A549

50	LE18-22.4	C1A	98.6581603	A	2	A549
25	LE18-22.4	C1A	94.624302	A	2	A549
12.5	LE18-22.4	C1A	85.9089222	A	2	A549
100	LE18-22.4	C1A	92.5998678	A	2	A549
50	LE18-22.4	C1A	101.128454	A	2	A549
25	LE18-22.4	C1A	94.1776089	A	2	A549
12.5	LE18-22.4	C1A	91.5915422	A	2	A549
100	LE18-22.4	C1B	98.0399615	B	2	A549
50	LE18-22.4	C1B	96.8583234	B	2	A549
25	LE18-22.4	C1B	93.5346711	B	2	A549
12.5	LE18-22.4	C1B	91.2930056	B	2	A549
100	LE18-22.4	C1B	93.1816529	B	2	A549
50	LE18-22.4	C1B	97.0110663	B	2	A549
25	LE18-22.4	C1B	99.3536339	B	2	A549
12.5	LE18-22.4	C1B	95.4315184	B	2	A549
100	LE18-22.4	C1C	101.369173	C	2	A549
50	LE18-22.4	C1C	96.7364349	C	2	A549
25	LE18-22.4	C1C	97.6043871	C	2	A549
12.5	LE18-22.4	C1C	91.9401129	C	2	A549
100	LE18-22.4	C1C	102.039352	C	2	A549
50	LE18-22.4	C1C	100.564042	C	2	A549
25	LE18-22.4	C1C	94.4889321	C	2	A549
12.5	LE18-22.4	C1C	95.7704993	C	2	A549
100	LE19-55.1	C2A	98.4627494	A	2	A549
50	LE19-55.1	C2A	95.8162249	A	2	A549
25	LE19-55.1	C2A	96.0035747	A	2	A549
12.5	LE19-55.1	C2A	95.2479212	A	2	A549
100	LE19-55.1	C2A	105.176071	A	2	A549
50	LE19-55.1	C2A	95.1698124	A	2	A549
25	LE19-55.1	C2A	90.9787638	A	2	A549
12.5	LE19-55.1	C2A	91.4407451	A	2	A549
100	LE19-55.1	C2B	55.8772201	B	2	A549
50	LE19-55.1	C2B	61.2036253	B	2	A549
25	LE19-55.1	C2B	79.3373636	B	2	A549
12.5	LE19-55.1	C2B	83.8722575	B	2	A549
100	LE19-55.1	C2B	56.6731788	B	2	A549
50	LE19-55.1	C2B	57.9911597	B	2	A549
25	LE19-55.1	C2B	82.669077	B	2	A549
12.5	LE19-55.1	C2B	84.3443845	B	2	A549
100	LE19-131.1	C3A	34.1751665	A	1	A549
50	LE19-131.1	C3A	44.8774451	A	1	A549

25	LE19-131.1	C3A	49.7199987	A	1	A549
12.5	LE19-131.1	C3A	49.6539847	A	1	A549
100	LE19-131.1	C3A	37.2644268	A	1	A549
50	LE19-131.1	C3A	55.2114178	A	1	A549
25	LE19-131.1	C3A	63.1587547	A	1	A549
12.5	LE19-131.1	C3A	63.9487988	A	1	A549
100	LE19-131.1	C3B	34.6983766	B	1	A549
50	LE19-131.1	C3B	58.6644088	B	1	A549
25	LE19-131.1	C3B	69.8325439	B	1	A549
12.5	LE19-131.1	C3B	70.3183352	B	1	A549
100	LE19-131.1	C3B	1.54838836	B	1	A549
50	LE19-131.1	C3B	55.0400101	B	1	A549
25	LE19-131.1	C3B	75.6847525	B	1	A549
12.5	LE19-131.1	C3B	76.247179	B	1	A549
100	LE19-131.1	C3C	67.6065314	C	1	A549
50	LE19-131.1	C3C	91.2302383	C	1	A549
25	LE19-131.1	C3C	98.4398195	C	1	A549
12.5	LE19-131.1	C3C	95.2054591	C	1	A549
100	LE19-131.1	C3C	78.5264602	C	1	A549
50	LE19-131.1	C3C	99.2980019	C	1	A549
25	LE19-131.1	C3C	105.196616	C	1	A549
12.5	LE19-131.1	C3C	102.173207	C	1	A549
100	LE19-41.2	C4B	90.3491799	B	1	A549
50	LE19-41.2	C4B	104.828801	B	1	A549
25	LE19-41.2	C4B	108.818401	B	1	A549
12.5	LE19-41.2	C4B	110.992615	B	1	A549
100	LE19-41.2	C4B	91.1704336	B	1	A549
50	LE19-41.2	C4B	104.048724	B	1	A549
25	LE19-41.2	C4B	110.280023	B	1	A549
12.5	LE19-41.2	C4B	108.023618	B	1	A549
100	LE19-131.1	C3A	90.7461963	A	2	A549
50	LE19-131.1	C3A	95.5688446	A	2	A549
25	LE19-131.1	C3A	90.8653736	A	2	A549
12.5	LE19-131.1	C3A	92.2341204	A	2	A549
100	LE19-131.1	C3A	90.1095386	A	2	A549
50	LE19-131.1	C3A	94.1538008	A	2	A549
25	LE19-131.1	C3A	92.3329868	A	2	A549
12.5	LE19-131.1	C3A	95.6432184	A	2	A549
100	LE19-131.1	C3B	6.39943273	B	2	A549
50	LE19-131.1	C3B	87.5031699	B	2	A549
25	LE19-131.1	C3B	90.8904635	B	2	A549

12.5	LE19-131.1	C3B	99.2552763	B	2	A549
100	LE19-131.1	C3B	6.42168513	B	2	A549
50	LE19-131.1	C3B	90.5323342	B	2	A549
25	LE19-131.1	C3B	92.2073876	B	2	A549
12.5	LE19-131.1	C3B	94.1294576	B	2	A549
100	LE19-131.1	C3C	93.5374541	C	2	A549
50	LE19-131.1	C3C	98.2134457	C	2	A549
25	LE19-131.1	C3C	93.8355467	C	2	A549
12.5	LE19-131.1	C3C	97.4888238	C	2	A549
100	LE19-131.1	C3C	89.1774765	C	2	A549
50	LE19-131.1	C3C	89.9022477	C	2	A549
25	LE19-131.1	C3C	92.7238226	C	2	A549
12.5	LE19-131.1	C3C	97.0427303	C	2	A549
100	LE19-41.2	C4B	91.2653194	B	2	A549
50	LE19-41.2	C4B	94.3872271	B	2	A549
25	LE19-41.2	C4B	88.9159734	B	2	A549
12.5	LE19-41.2	C4B	97.8444142	B	2	A549
100	LE19-41.2	C4B	91.92184	B	2	A549
50	LE19-41.2	C4B	92.5107073	B	2	A549
25	LE19-41.2	C4B	84.010588	B	2	A549
12.5	LE19-41.2	C4B	95.800628	B	2	A549
100	LE19-41.2	C4C	36.77357	C	1	A549
50	LE19-41.2	C4C	50.7863539	C	1	A549
25	LE19-41.2	C4C	45.6312142	C	1	A549
12.5	LE19-41.2	C4C	37.9928686	C	1	A549
100	LE19-41.2	C4C	49.0270148	C	1	A549
50	LE19-41.2	C4C	56.0351225	C	1	A549
25	LE19-41.2	C4C	51.2845638	C	1	A549
12.5	LE19-41.2	C4C	39.0742568	C	1	A549
100	LE17-20	C5A	50.1095468	A	1	A549
50	LE17-20	C5A	68.3021296	A	1	A549
25	LE17-20	C5A	57.5179871	A	1	A549
12.5	LE17-20	C5A	37.9649271	A	1	A549
100	LE17-20	C5A	57.4338355	A	1	A549
50	LE17-20	C5A	76.6702243	A	1	A549
25	LE17-20	C5A	65.0790439	A	1	A549
12.5	LE17-20	C5A	48.7016833	A	1	A549
100	LE17-20	C5B	62.1059618	B	1	A549
50	LE17-20	C5B	83.9435319	B	1	A549
25	LE17-20	C5B	68.0210798	B	1	A549
12.5	LE17-20	C5B	57.9320453	B	1	A549

100	LE17-20	C5B	75.0837635	B	1	A549
50	LE17-20	C5B	95.089281	B	1	A549
25	LE17-20	C5B	90.6709166	B	1	A549
12.5	LE17-20	C5B	67.4728693	B	1	A549
100	LE17-20	C5C	90.6632367	C	1	A549
50	LE17-20	C5C	100.231512	C	1	A549
25	LE17-20	C5C	97.7195477	C	1	A549
12.5	LE17-20	C5C	78.6377363	C	1	A549
100	LE17-20	C5C	101.463066	C	1	A549
50	LE17-20	C5C	106.122283	C	1	A549
25	LE17-20	C5C	102.634161	C	1	A549
12.5	LE17-20	C5C	88.8525897	C	1	A549
100	LE19-197.1	C6A	97.8720009	A	1	A549
50	LE19-197.1	C6A	107.684888	A	1	A549
25	LE19-197.1	C6A	102.056702	A	1	A549
12.5	LE19-197.1	C6A	89.2878901	A	1	A549
100	LE19-197.1	C6A	104.664419	A	1	A549
50	LE19-197.1	C6A	107.768386	A	1	A549
25	LE19-197.1	C6A	103.175672	A	1	A549
12.5	LE19-197.1	C6A	91.1519692	A	1	A549
100	LE19-41.2	C4C	93.2650488	C	2	A549
50	LE19-41.2	C4C	88.7056957	C	2	A549
25	LE19-41.2	C4C	84.8255636	C	2	A549
12.5	LE19-41.2	C4C	87.693286	C	2	A549
100	LE19-41.2	C4C	94.7065266	C	2	A549
50	LE19-41.2	C4C	92.374206	C	2	A549
25	LE19-41.2	C4C	91.909295	C	2	A549
12.5	LE19-41.2	C4C	86.9337174	C	2	A549
100	LE17-20	C5A	97.3569521	A	2	A549
50	LE17-20	C5A	92.920958	A	2	A549
25	LE17-20	C5A	92.7620549	A	2	A549
12.5	LE17-20	C5A	86.9685148	A	2	A549
100	LE17-20	C5A	101.206767	A	2	A549
50	LE17-20	C5A	97.8668159	A	2	A549
25	LE17-20	C5A	91.9915841	A	2	A549
12.5	LE17-20	C5A	88.0031769	A	2	A549
100	LE17-20	C5B	96.6088831	B	2	A549
50	LE17-20	C5B	96.0967791	B	2	A549
25	LE17-20	C5B	88.5868171	B	2	A549
12.5	LE17-20	C5B	83.6452901	B	2	A549
100	LE17-20	C5B	97.9901749	B	2	A549

50	LE17-20	C5B	96.8644124	B	2	A549
25	LE17-20	C5B	93.9896707	B	2	A549
12.5	LE17-20	C5B	87.7921524	B	2	A549
100	LE17-20	C5C	95.4447389	C	2	A549
50	LE17-20	C5C	95.0185083	C	2	A549
25	LE17-20	C5C	91.9966618	C	2	A549
12.5	LE17-20	C5C	86.1169497	C	2	A549
100	LE17-20	C5C	99.7829121	C	2	A549
50	LE17-20	C5C	99.217492	C	2	A549
25	LE17-20	C5C	93.0867309	C	2	A549
12.5	LE17-20	C5C	87.9637498	C	2	A549
100	LE19-197.1	C6A	93.2968593	A	2	A549
50	LE19-197.1	C6A	89.5432224	A	2	A549
25	LE19-197.1	C6A	90.9167482	A	2	A549
12.5	LE19-197.1	C6A	85.4584876	A	2	A549
100	LE19-197.1	C6A	89.192411	A	2	A549
50	LE19-197.1	C6A	88.1516258	A	2	A549
25	LE19-197.1	C6A	87.1457873	A	2	A549
12.5	LE19-197.1	C6A	84.050015	A	2	A549
100	LE19-197.1	C6B	57.834926	B	1	A549
50	LE19-197.1	C6B	92.6815884	B	1	A549
25	LE19-197.1	C6B	98.9848155	B	1	A549
12.5	LE19-197.1	C6B	102.843026	B	1	A549
100	LE19-197.1	C6B	60.7852871	B	1	A549
50	LE19-197.1	C6B	103.54097	B	1	A549
25	LE19-197.1	C6B	107.808327	B	1	A549
12.5	LE19-197.1	C6B	110.799945	B	1	A549
100	LE19-197.1	C6C	101.536697	C	1	A549
50	LE19-197.1	C6C	114.465557	C	1	A549
25	LE19-197.1	C6C	115.405809	C	1	A549
12.5	LE19-197.1	C6C	116.480497	C	1	A549
100	LE19-197.1	C6C	101.478736	C	1	A549
50	LE19-197.1	C6C	117.000735	C	1	A549
25	LE19-197.1	C6C	117.372649	C	1	A549
12.5	LE19-197.1	C6C	117.6379	C	1	A549
100	LE19-84.1	C7C	100.044812	C	1	A549
50	LE19-84.1	C7C	112.321212	C	1	A549
25	LE19-84.1	C7C	117.350511	C	1	A549
12.5	LE19-84.1	C7C	117.089688	C	1	A549
100	LE19-84.1	C7C	98.061268	C	1	A549
50	LE19-84.1	C7C	113.786129	C	1	A549

25	LE19-84.1	C7C	118.866546	C	1	A549
12.5	LE19-84.1	C7C	118.616389	C	1	A549
100	LE19-84.1	C7B	91.1287244	B	1	A549
50	LE19-84.1	C7B	108.697058	B	1	A549
25	LE19-84.1	C7B	117.239621	B	1	A549
12.5	LE19-84.1	C7B	120.394455	B	1	A549
100	LE19-84.1	C7B	88.6800857	B	1	A549
50	LE19-84.1	C7B	110.965173	B	1	A549
25	LE19-84.1	C7B	117.223924	B	1	A549
12.5	LE19-84.1	C7B	118.272248	B	1	A549
100	Doxorubicin	Dox	12.6183281	NA	1	A549
50	Doxorubicin	Dox	16.1525225	NA	1	A549
25	Doxorubicin	Dox	48.1641415	NA	1	A549
12.5	Doxorubicin	Dox	65.0299028	NA	1	A549
100	Doxorubicin	Dox	10.0624218	NA	1	A549
50	Doxorubicin	Dox	20.8201709	NA	1	A549
25	Doxorubicin	Dox	57.34065	NA	1	A549
12.5	Doxorubicin	Dox	66.9725928	NA	1	A549
100	LE19-197.1	C6B	61.3624387	B	2	A549
50	LE19-197.1	C6B	99.097944	B	2	A549
25	LE19-197.1	C6B	99.044817	B	2	A549
12.5	LE19-197.1	C6B	108.090214	B	2	A549
100	LE19-197.1	C6B	64.2615996	B	2	A549
50	LE19-197.1	C6B	103.175488	B	2	A549
25	LE19-197.1	C6B	97.9560015	B	2	A549
12.5	LE19-197.1	C6B	100.815077	B	2	A549
100	LE19-197.1	C6C	112.707276	C	2	A549
50	LE19-197.1	C6C	102.913881	C	2	A549
25	LE19-197.1	C6C	101.12214	C	2	A549
12.5	LE19-197.1	C6C	98.9381795	C	2	A549
100	LE19-197.1	C6C	115.381078	C	2	A549
50	LE19-197.1	C6C	103.358652	C	2	A549
25	LE19-197.1	C6C	99.8867935	C	2	A549
12.5	LE19-197.1	C6C	105.03033	C	2	A549
100	LE19-84.1	C7C	109.658132	C	2	A549
50	LE19-84.1	C7C	101.359006	C	2	A549
25	LE19-84.1	C7C	95.9095575	C	2	A549
12.5	LE19-84.1	C7C	102.201749	C	2	A549
100	LE19-84.1	C7C	111.002034	C	2	A549
50	LE19-84.1	C7C	105.318597	C	2	A549
25	LE19-84.1	C7C	100.401186	C	2	A549

12.5	LE19-84.1	C7C	105.19393	C	2	A549
100	LE19-84.1	C7B	105.741119	B	2	A549
50	LE19-84.1	C7B	106.401658	B	2	A549
25	LE19-84.1	C7B	100.263669	B	2	A549
12.5	LE19-84.1	C7B	106.522489	B	2	A549
100	LE19-84.1	C7B	108.590797	B	2	A549
50	LE19-84.1	C7B	100.090671	B	2	A549
25	LE19-84.1	C7B	99.2818746	B	2	A549
12.5	LE19-84.1	C7B	99.8866017	B	2	A549
100	Doxorubicin	Dox	12.474869	NA	2	A549
50	Doxorubicin	Dox	25.2136467	NA	2	A549
25	Doxorubicin	Dox	68.3445141	NA	2	A549
12.5	Doxorubicin	Dox	86.6547211	NA	2	A549
100	Doxorubicin	Dox	12.2732933	NA	2	A549
50	Doxorubicin	Dox	29.0422422	NA	2	A549
25	Doxorubicin	Dox	70.609987	NA	2	A549
12.5	Doxorubicin	Dox	87.2803537	NA	2	A549
NA	Control (Vehicle)	NA	83.7156277	NA	NA	A549
NA	Control (Vehicle)	NA	109.628602	NA	NA	A549
NA	Control (Vehicle)	NA	110.719125	NA	NA	A549
NA	Control (Vehicle)	NA	116.554671	NA	NA	A549
NA	Control (Vehicle)	NA	78.9237121	NA	NA	A549
NA	Control (Vehicle)	NA	107.817323	NA	NA	A549
NA	Control (Vehicle)	NA	106.425099	NA	NA	A549
NA	Control (Vehicle)	NA	86.2158408	NA	NA	A549

HK2						
dose	**culture_name**	**culture**	**viability**	**bio_rep**	**plate**	**cell_line**
100	LE18-22.4	C1A	79.9943871	A	1	HK2
50	LE18-22.4	C1A	82.3997143	A	1	HK2
25	LE18-22.4	C1A	83.4631451	A	1	HK2
12.5	LE18-22.4	C1A	79.2998983	A	1	HK2
100	LE18-22.4	C1A	97.2728329	A	1	HK2
50	LE18-22.4	C1A	100.332102	A	1	HK2
25	LE18-22.4	C1A	101.19548	A	1	HK2
12.5	LE18-22.4	C1A	97.0307253	A	1	HK2

100	LE18-22.4	C1B	97.381572	B	1	HK2
50	LE18-22.4	C1B	102.655331	B	1	HK2
25	LE18-22.4	C1B	100.715453	B	1	HK2
12.5	LE18-22.4	C1B	99.5839979	B	1	HK2
100	LE18-22.4	C1B	100.233249	B	1	HK2
50	LE18-22.4	C1B	105.241275	B	1	HK2
25	LE18-22.4	C1B	102.171115	B	1	HK2
12.5	LE18-22.4	C1B	101.216088	B	1	HK2
100	LE18-22.4	C1C	97.1560516	C	1	HK2
50	LE18-22.4	C1C	102.240816	C	1	HK2
25	LE18-22.4	C1C	107.217678	C	1	HK2
12.5	LE18-22.4	C1C	102.958091	C	1	HK2
100	LE18-22.4	C1C	104.236319	C	1	HK2
50	LE18-22.4	C1C	105.208771	C	1	HK2
25	LE18-22.4	C1C	107.736241	C	1	HK2
12.5	LE18-22.4	C1C	104.163435	C	1	HK2
100	LE19-55.1	C2A	103.86922	A	1	HK2
50	LE19-55.1	C2A	115.034992	A	1	HK2
25	LE19-55.1	C2A	107.675086	A	1	HK2
12.5	LE19-55.1	C2A	109.868967	A	1	HK2
100	LE19-55.1	C2A	107.906805	A	1	HK2
50	LE19-55.1	C2A	108.455694	A	1	HK2
25	LE19-55.1	C2A	104.9546	A	1	HK2
12.5	LE19-55.1	C2A	100.221353	A	1	HK2
100	LE19-55.1	C2B	83.5745649	B	1	HK2
50	LE19-55.1	C2B	95.4435377	B	1	HK2
25	LE19-55.1	C2B	98.9727807	B	1	HK2
12.5	LE19-55.1	C2B	97.9565986	B	1	HK2
100	LE19-55.1	C2B	81.6797581	B	1	HK2
50	LE19-55.1	C2B	91.9385893	B	1	HK2
25	LE19-55.1	C2B	96.3456192	B	1	HK2
12.5	LE19-55.1	C2B	93.8828228	B	1	HK2
100	LE18-22.4	C1A	85.418261	A	2	HK2
50	LE18-22.4	C1A	88.9308889	A	2	HK2
25	LE18-22.4	C1A	80.6519016	A	2	HK2
12.5	LE18-22.4	C1A	86.785751	A	2	HK2
100	LE18-22.4	C1A	93.896251	A	2	HK2
50	LE18-22.4	C1A	94.6773326	A	2	HK2
25	LE18-22.4	C1A	88.1808595	A	2	HK2
12.5	LE18-22.4	C1A	95.6045685	A	2	HK2
100	LE18-22.4	C1B	93.3529873	B	2	HK2

50	LE18-22.4	C1B	94.1645239	B	2	HK2
25	LE18-22.4	C1B	90.7172847	B	2	HK2
12.5	LE18-22.4	C1B	94.426228	B	2	HK2
100	LE18-22.4	C1B	93.5354188	B	2	HK2
50	LE18-22.4	C1B	97.0026628	B	2	HK2
25	LE18-22.4	C1B	92.5520502	B	2	HK2
12.5	LE18-22.4	C1B	97.2600375	B	2	HK2
100	LE18-22.4	C1C	103.395081	C	2	HK2
50	LE18-22.4	C1C	99.9521714	C	2	HK2
25	LE18-22.4	C1C	94.6279179	C	2	HK2
12.5	LE18-22.4	C1C	98.3529845	C	2	HK2
100	LE18-22.4	C1C	103.199363	C	2	HK2
50	LE18-22.4	C1C	100.243286	C	2	HK2
25	LE18-22.4	C1C	94.7197308	C	2	HK2
12.5	LE18-22.4	C1C	94.0960001	C	2	HK2
100	LE19-55.1	C2A	101.216652	A	2	HK2
50	LE19-55.1	C2A	101.325932	A	2	HK2
25	LE19-55.1	C2A	97.3185589	A	2	HK2
12.5	LE19-55.1	C2A	94.7164464	A	2	HK2
100	LE19-55.1	C2A	103.89953	A	2	HK2
50	LE19-55.1	C2A	99.2369264	A	2	HK2
25	LE19-55.1	C2A	97.4593387	A	2	HK2
12.5	LE19-55.1	C2A	95.2989733	A	2	HK2
100	LE19-55.1	C2B	59.0254059	B	2	HK2
50	LE19-55.1	C2B	80.1095336	B	2	HK2
25	LE19-55.1	C2B	86.6800542	B	2	HK2
12.5	LE19-55.1	C2B	85.9324134	B	2	HK2
100	LE19-55.1	C2B	69.2897916	B	2	HK2
50	LE19-55.1	C2B	79.2190229	B	2	HK2
25	LE19-55.1	C2B	89.7613853	B	2	HK2
12.5	LE19-55.1	C2B	86.8723388	B	2	HK2
100	LE19-131.1	C3A	92.8160867	A	1	HK2
50	LE19-131.1	C3A	105.639859	A	1	HK2
25	LE19-131.1	C3A	103.888734	A	1	HK2
12.5	LE19-131.1	C3A	105.855242	A	1	HK2
100	LE19-131.1	C3A	99.9578618	A	1	HK2
50	LE19-131.1	C3A	112.906038	A	1	HK2
25	LE19-131.1	C3A	116.06028	A	1	HK2
12.5	LE19-131.1	C3A	114.520248	A	1	HK2
100	LE19-131.1	C3B	13.0754399	B	1	HK2
50	LE19-131.1	C3B	108.516874	B	1	HK2

25	LE19-131.1	C3B	116.44797	B	1	HK2
12.5	LE19-131.1	C3B	118.188549	B	1	HK2
100	LE19-131.1	C3B	13.4675976	B	1	HK2
50	LE19-131.1	C3B	108.499715	B	1	HK2
25	LE19-131.1	C3B	109.808457	B	1	HK2
12.5	LE19-131.1	C3B	111.349562	B	1	HK2
100	LE19-131.1	C3C	104.403687	C	1	HK2
50	LE19-131.1	C3C	99.9555382	C	1	HK2
25	LE19-131.1	C3C	108.614467	C	1	HK2
12.5	LE19-131.1	C3C	113.31321	C	1	HK2
100	LE19-131.1	C3C	99.964654	C	1	HK2
50	LE19-131.1	C3C	108.567815	C	1	HK2
25	LE19-131.1	C3C	108.849869	C	1	HK2
12.5	LE19-131.1	C3C	109.161057	C	1	HK2
100	LE19-41.2	C4B	104.926861	B	1	HK2
50	LE19-41.2	C4B	107.08498	B	1	HK2
25	LE19-41.2	C4B	109.480824	B	1	HK2
12.5	LE19-41.2	C4B	118.008557	B	1	HK2
100	LE19-41.2	C4B	103.139988	B	1	HK2
50	LE19-41.2	C4B	103.484064	B	1	HK2
25	LE19-41.2	C4B	100.585064	B	1	HK2
12.5	LE19-41.2	C4B	107.987443	B	1	HK2
100	LE19-131.1	C3A	91.121325	A	2	HK2
50	LE19-131.1	C3A	100.228693	A	2	HK2
25	LE19-131.1	C3A	102.603008	A	2	HK2
12.5	LE19-131.1	C3A	100.400618	A	2	HK2
100	LE19-131.1	C3A	90.4180348	A	2	HK2
50	LE19-131.1	C3A	104.331438	A	2	HK2
25	LE19-131.1	C3A	103.882393	A	2	HK2
12.5	LE19-131.1	C3A	104.45572	A	2	HK2
100	LE19-131.1	C3B	4.15155417	B	2	HK2
50	LE19-131.1	C3B	103.62334	B	2	HK2
25	LE19-131.1	C3B	108.993085	B	2	HK2
12.5	LE19-131.1	C3B	104.27243	B	2	HK2
100	LE19-131.1	C3B	17.366621	B	2	HK2
50	LE19-131.1	C3B	109.660534	B	2	HK2
25	LE19-131.1	C3B	101.906566	B	2	HK2
12.5	LE19-131.1	C3B	103.02233	B	2	HK2
100	LE19-131.1	C3C	101.638042	C	2	HK2
50	LE19-131.1	C3C	107.548914	C	2	HK2
25	LE19-131.1	C3C	102.451627	C	2	HK2

12.5	LE19-131.1	C3C	104.365095	C	2	HK2
100	LE19-131.1	C3C	94.3229202	C	2	HK2
50	LE19-131.1	C3C	107.845704	C	2	HK2
25	LE19-131.1	C3C	106.287888	C	2	HK2
12.5	LE19-131.1	C3C	107.861148	C	2	HK2
100	LE19-41.2	C4B	99.3080197	B	2	HK2
50	LE19-41.2	C4B	107.902235	B	2	HK2
25	LE19-41.2	C4B	102.791397	B	2	HK2
12.5	LE19-41.2	C4B	106.549272	B	2	HK2
100	LE19-41.2	C4B	99.0757751	B	2	HK2
50	LE19-41.2	C4B	106.366274	B	2	HK2
25	LE19-41.2	C4B	103.697355	B	2	HK2
12.5	LE19-41.2	C4B	103.582835	B	2	HK2
100	LE19-41.2	C4C	102.876702	C	1	HK2
50	LE19-41.2	C4C	91.3393286	C	1	HK2
25	LE19-41.2	C4C	99.4598894	C	1	HK2
12.5	LE19-41.2	C4C	100.200771	C	1	HK2
100	LE19-41.2	C4C	112.571614	C	1	HK2
50	LE19-41.2	C4C	104.291437	C	1	HK2
25	LE19-41.2	C4C	107.215461	C	1	HK2
12.5	LE19-41.2	C4C	104.395286	C	1	HK2
100	LE17-20	C5A	69.9695917	A	1	HK2
50	LE17-20	C5A	105.215171	A	1	HK2
25	LE17-20	C5A	104.128068	A	1	HK2
12.5	LE17-20	C5A	101.601564	A	1	HK2
100	LE17-20	C5A	42.0198626	A	1	HK2
50	LE17-20	C5A	109.651343	A	1	HK2
25	LE17-20	C5A	109.877272	A	1	HK2
12.5	LE17-20	C5A	105.917444	A	1	HK2
100	LE17-20	C5B	115.025549	B	1	HK2
50	LE17-20	C5B	112.480635	B	1	HK2
25	LE17-20	C5B	107.702173	B	1	HK2
12.5	LE17-20	C5B	107.452472	B	1	HK2
100	LE17-20	C5B	109.19752	B	1	HK2
50	LE17-20	C5B	110.523778	B	1	HK2
25	LE17-20	C5B	103.758253	B	1	HK2
12.5	LE17-20	C5B	108.119176	B	1	HK2
100	LE17-20	C5C	110.258169	C	1	HK2
50	LE17-20	C5C	113.86981	C	1	HK2
25	LE17-20	C5C	111.918673	C	1	HK2
12.5	LE17-20	C5C	105.762833	C	1	HK2

100	LE17-20	C5C	104.933117	C	1	HK2
50	LE17-20	C5C	109.065073	C	1	HK2
25	LE17-20	C5C	100.621527	C	1	HK2
12.5	LE17-20	C5C	92.900095	C	1	HK2
100	LE19-197.1	C6A	101.53829	A	1	HK2
50	LE19-197.1	C6A	109.177859	A	1	HK2
25	LE19-197.1	C6A	103.462794	A	1	HK2
12.5	LE19-197.1	C6A	96.59485	A	1	HK2
100	LE19-197.1	C6A	98.6509077	A	1	HK2
50	LE19-197.1	C6A	105.222678	A	1	HK2
25	LE19-197.1	C6A	102.328325	A	1	HK2
12.5	LE19-197.1	C6A	103.230073	A	1	HK2
100	LE19-41.2	C4C	90.0387797	C	2	HK2
50	LE19-41.2	C4C	87.0481571	C	2	HK2
25	LE19-41.2	C4C	82.0885549	C	2	HK2
12.5	LE19-41.2	C4C	83.7728381	C	2	HK2
100	LE19-41.2	C4C	94.4425392	C	2	HK2
50	LE19-41.2	C4C	89.8856498	C	2	HK2
25	LE19-41.2	C4C	86.9023121	C	2	HK2
12.5	LE19-41.2	C4C	91.7000424	C	2	HK2
100	LE17-20	C5A	84.1108604	A	2	HK2
50	LE17-20	C5A	93.1079342	A	2	HK2
25	LE17-20	C5A	89.5817212	A	2	HK2
12.5	LE17-20	C5A	91.1195766	A	2	HK2
100	LE17-20	C5A	93.9975737	A	2	HK2
50	LE17-20	C5A	95.6318821	A	2	HK2
25	LE17-20	C5A	91.1409944	A	2	HK2
12.5	LE17-20	C5A	91.9132004	A	2	HK2
100	LE17-20	C5B	96.6025305	B	2	HK2
50	LE17-20	C5B	96.583881	B	2	HK2
25	LE17-20	C5B	90.4468833	B	2	HK2
12.5	LE17-20	C5B	90.7152613	B	2	HK2
100	LE17-20	C5B	95.0476282	B	2	HK2
50	LE17-20	C5B	94.9733216	B	2	HK2
25	LE17-20	C5B	90.1272191	B	2	HK2
12.5	LE17-20	C5B	91.8884315	B	2	HK2
100	LE17-20	C5C	93.3790805	C	2	HK2
50	LE17-20	C5C	95.1231004	C	2	HK2
25	LE17-20	C5C	94.969242	C	2	HK2
12.5	LE17-20	C5C	97.6259219	C	2	HK2
100	LE17-20	C5C	91.3665369	C	2	HK2

50	LE17-20	C5C	95.853345	C	2	HK2
25	LE17-20	C5C	93.4455193	C	2	HK2
12.5	LE17-20	C5C	94.6026627	C	2	HK2
100	LE19-197.1	C6A	88.1363847	A	2	HK2
50	LE19-197.1	C6A	98.3252783	A	2	HK2
25	LE19-197.1	C6A	98.6858839	A	2	HK2
12.5	LE19-197.1	C6A	96.7158845	A	2	HK2
100	LE19-197.1	C6A	91.7240828	A	2	HK2
50	LE19-197.1	C6A	96.5050577	A	2	HK2
25	LE19-197.1	C6A	95.5550986	A	2	HK2
12.5	LE19-197.1	C6A	93.9360887	A	2	HK2
100	LE19-197.1	C6B	77.6947107	B	1	HK2
50	LE19-197.1	C6B	87.1164243	B	1	HK2
25	LE19-197.1	C6B	85.6513836	B	1	HK2
12.5	LE19-197.1	C6B	89.8039834	B	1	HK2
100	LE19-197.1	C6B	87.5376681	B	1	HK2
50	LE19-197.1	C6B	94.4707543	B	1	HK2
25	LE19-197.1	C6B	96.0996561	B	1	HK2
12.5	LE19-197.1	C6B	99.2641361	B	1	HK2
100	LE19-197.1	C6C	104.3256	C	1	HK2
50	LE19-197.1	C6C	97.0917387	C	1	HK2
25	LE19-197.1	C6C	100.905719	C	1	HK2
12.5	LE19-197.1	C6C	108.116596	C	1	HK2
100	LE19-197.1	C6C	107.074384	C	1	HK2
50	LE19-197.1	C6C	101.265341	C	1	HK2
25	LE19-197.1	C6C	106.122326	C	1	HK2
12.5	LE19-197.1	C6C	104.65035	C	1	HK2
100	LE19-84.1	C7C	103.914263	C	1	HK2
50	LE19-84.1	C7C	114.158651	C	1	HK2
25	LE19-84.1	C7C	109.485737	C	1	HK2
12.5	LE19-84.1	C7C	112.15586	C	1	HK2
100	LE19-84.1	C7C	106.879415	C	1	HK2
50	LE19-84.1	C7C	108.642061	C	1	HK2
25	LE19-84.1	C7C	109.554492	C	1	HK2
12.5	LE19-84.1	C7C	111.320109	C	1	HK2
100	LE19-84.1	C7B	107.833059	B	1	HK2
50	LE19-84.1	C7B	108.561418	B	1	HK2
25	LE19-84.1	C7B	111.813476	B	1	HK2
12.5	LE19-84.1	C7B	111.967232	B	1	HK2
100	LE19-84.1	C7B	105.332741	B	1	HK2
50	LE19-84.1	C7B	110.683687	B	1	HK2

25	LE19-84.1	C7B	108.242612	B	1	HK2
12.5	LE19-84.1	C7B	110.737977	B	1	HK2
100	Doxorubicin	Dox	18.7304879	X	1	HK2
50	Doxorubicin	Dox	23.8591211	X	1	HK2
25	Doxorubicin	Dox	53.4624936	X	1	HK2
12.5	Doxorubicin	Dox	93.912002	X	1	HK2
100	Doxorubicin	Dox	17.5224337	X	1	HK2
50	Doxorubicin	Dox	27.0529257	X	1	HK2
25	Doxorubicin	Dox	55.7704182	X	1	HK2
12.5	Doxorubicin	Dox	90.5959455	X	1	HK2
100	LE19-197.1	C6B	75.4489265	B	2	HK2
50	LE19-197.1	C6B	93.4637483	B	2	HK2
25	LE19-197.1	C6B	95.5316128	B	2	HK2
12.5	LE19-197.1	C6B	94.9409509	B	2	HK2
100	LE19-197.1	C6B	81.3034925	B	2	HK2
50	LE19-197.1	C6B	95.2762803	B	2	HK2
25	LE19-197.1	C6B	99.6273433	B	2	HK2
12.5	LE19-197.1	C6B	102.868495	B	2	HK2
100	LE19-197.1	C6C	101.676396	C	2	HK2
50	LE19-197.1	C6C	103.411853	C	2	HK2
25	LE19-197.1	C6C	101.815203	C	2	HK2
12.5	LE19-197.1	C6C	105.131421	C	2	HK2
100	LE19-197.1	C6C	103.954663	C	2	HK2
50	LE19-197.1	C6C	101.130846	C	2	HK2
25	LE19-197.1	C6C	105.198815	C	2	HK2
12.5	LE19-197.1	C6C	112.071789	C	2	HK2
100	LE19-84.1	C7C	103.217705	C	2	HK2
50	LE19-84.1	C7C	106.742499	C	2	HK2
25	LE19-84.1	C7C	105.821074	C	2	HK2
12.5	LE19-84.1	C7C	103.2367	C	2	HK2
100	LE19-84.1	C7C	104.268258	C	2	HK2
50	LE19-84.1	C7C	106.111656	C	2	HK2
25	LE19-84.1	C7C	105.515698	C	2	HK2
12.5	LE19-84.1	C7C	105.648113	C	2	HK2
100	LE19-84.1	C7B	103.298798	B	2	HK2
50	LE19-84.1	C7B	104.800475	B	2	HK2
25	LE19-84.1	C7B	106.25229	B	2	HK2
12.5	LE19-84.1	C7B	106.579035	B	2	HK2
100	LE19-84.1	C7B	101.885155	B	2	HK2
50	LE19-84.1	C7B	106.389636	B	2	HK2
25	LE19-84.1	C7B	105.248676	B	2	HK2

12.5	LE19-84.1	C7B	103.466828	B	2	HK2
100	Doxorubicin	Dox	15.4227769	X	2	HK2
50	Doxorubicin	Dox	24.004177	X	2	HK2
25	Doxorubicin	Dox	58.5034694	X	2	HK2
12.5	Doxorubicin	Dox	96.8904639	X	2	HK2
100	Doxorubicin	Dox	15.9683264	X	2	HK2
50	Doxorubicin	Dox	21.8201527	X	2	HK2
25	Doxorubicin	Dox	49.705822	X	2	HK2
12.5	Doxorubicin	Dox	93.6706809	X	2	HK2
NA	Control (Vehicle)	NA	101.5877	NA	NA	HK2
NA	Control (Vehicle)	NA	110.5116	NA	NA	HK2
NA	Control (Vehicle)	NA	101.7818	NA	NA	HK2
NA	Control (Vehicle)	NA	103.5619	NA	NA	HK2
NA	Control (Vehicle)	NA	98.1335	NA	NA	HK2
NA	Control (Vehicle)	NA	93.30079	NA	NA	HK2
NA	Control (Vehicle)	NA	96.77899	NA	NA	HK2
NA	Control (Vehicle)	NA	94.34373	NA	NA	HK2

Hep3B						
dose	**culture_name**	**culture**	**viability**	**bio_rep**	**plate**	**cell_line**
100	LE18-22.4	C1A	110.277961	A	1	Hep3B
50	LE18-22.4	C1A	109.243843	A	1	Hep3B
25	LE18-22.4	C1A	103.252633	A	1	Hep3B
12.5	LE18-22.4	C1A	92.581541	A	1	Hep3B
100	LE18-22.4	C1A	74.8621038	A	1	Hep3B
50	LE18-22.4	C1A	115.171005	A	1	Hep3B
25	LE18-22.4	C1A	106.297275	A	1	Hep3B
12.5	LE18-22.4	C1A	103.563869	A	1	Hep3B
100	LE18-22.4	C1B	85.8881318	B	1	Hep3B
50	LE18-22.4	C1B	114.587562	B	1	Hep3B
25	LE18-22.4	C1B	109.33558	B	1	Hep3B
12.5	LE18-22.4	C1B	102.032041	B	1	Hep3B
100	LE18-22.4	C1B	118.133585	B	1	Hep3B
50	LE18-22.4	C1B	125.189602	B	1	Hep3B
25	LE18-22.4	C1B	118.726034	B	1	Hep3B

12.5	LE18-22.4	C1B	109.411637	B	1	Hep3B
100	LE18-22.4	C1C	115.057585	C	1	Hep3B
50	LE18-22.4	C1C	122.29941	C	1	Hep3B
25	LE18-22.4	C1C	118.102228	C	1	Hep3B
12.5	LE18-22.4	C1C	114.910807	C	1	Hep3B
100	LE18-22.4	C1C	110.712957	C	1	Hep3B
50	LE18-22.4	C1C	127.71485	C	1	Hep3B
25	LE18-22.4	C1C	111.488879	C	1	Hep3B
12.5	LE18-22.4	C1C	109.295216	C	1	Hep3B
100	LE19-55.1	C2A	110.599204	A	1	Hep3B
50	LE19-55.1	C2A	127.2088	A	1	Hep3B
25	LE19-55.1	C2A	113.964089	A	1	Hep3B
12.5	LE19-55.1	C2A	116.517025	A	1	Hep3B
100	LE19-55.1	C2A	113.06741	A	1	Hep3B
50	LE19-55.1	C2A	128.771318	A	1	Hep3B
25	LE19-55.1	C2A	122.508235	A	1	Hep3B
12.5	LE19-55.1	C2A	114.129215	A	1	Hep3B
100	LE19-55.1	C2B	70.5785221	B	1	Hep3B
50	LE19-55.1	C2B	67.1932882	B	1	Hep3B
25	LE19-55.1	C2B	96.1742655	B	1	Hep3B
12.5	LE19-55.1	C2B	102.265551	B	1	Hep3B
100	LE19-55.1	C2B	65.2167894	B	1	Hep3B
50	LE19-55.1	C2B	62.4233377	B	1	Hep3B
25	LE19-55.1	C2B	91.2471958	B	1	Hep3B
12.5	LE19-55.1	C2B	97.97363	B	1	Hep3B
100	LE18-22.4	C1A	69.5814482	A	2	Hep3B
50	LE18-22.4	C1A	104.174093	A	2	Hep3B
25	LE18-22.4	C1A	84.6502702	A	2	Hep3B
12.5	LE18-22.4	C1A	105.41722	A	2	Hep3B
100	LE18-22.4	C1A	73.8311141	A	2	Hep3B
50	LE18-22.4	C1A	118.161278	A	2	Hep3B
25	LE18-22.4	C1A	104.144142	A	2	Hep3B
12.5	LE18-22.4	C1A	114.75364	A	2	Hep3B
100	LE18-22.4	C1B	70.8773755	B	2	Hep3B
50	LE18-22.4	C1B	123.13502	B	2	Hep3B
25	LE18-22.4	C1B	117.80526	B	2	Hep3B
12.5	LE18-22.4	C1B	123.426195	B	2	Hep3B
100	LE18-22.4	C1B	75.2227616	B	2	Hep3B
50	LE18-22.4	C1B	140.011409	B	2	Hep3B

25	LE18-22.4	C1B	120.180047	B	2	Hep3B
12.5	LE18-22.4	C1B	126.852359	B	2	Hep3B
100	LE18-22.4	C1C	99.4705286	C	2	Hep3B
50	LE18-22.4	C1C	126.298417	C	2	Hep3B
25	LE18-22.4	C1C	129.198737	C	2	Hep3B
12.5	LE18-22.4	C1C	130.978824	C	2	Hep3B
100	LE18-22.4	C1C	98.6213982	C	2	Hep3B
50	LE18-22.4	C1C	133.672574	C	2	Hep3B
25	LE18-22.4	C1C	125.988407	C	2	Hep3B
12.5	LE18-22.4	C1C	139.204581	C	2	Hep3B
100	LE19-55.1	C2A	99.0743707	A	2	Hep3B
50	LE19-55.1	C2A	131.686844	A	2	Hep3B
25	LE19-55.1	C2A	127.78177	A	2	Hep3B
12.5	LE19-55.1	C2A	130.004639	A	2	Hep3B
100	LE19-55.1	C2A	101.294152	A	2	Hep3B
50	LE19-55.1	C2A	127.922263	A	2	Hep3B
25	LE19-55.1	C2A	132.729885	A	2	Hep3B
12.5	LE19-55.1	C2A	126.657522	A	2	Hep3B
100	LE19-55.1	C2B	68.6776647	B	2	Hep3B
50	LE19-55.1	C2B	71.8722476	B	2	Hep3B
25	LE19-55.1	C2B	96.788203	B	2	Hep3B
12.5	LE19-55.1	C2B	121.67853	B	2	Hep3B
100	LE19-55.1	C2B	62.4252862	B	2	Hep3B
50	LE19-55.1	C2B	87.9523067	B	2	Hep3B
25	LE19-55.1	C2B	100.821418	B	2	Hep3B
12.5	LE19-55.1	C2B	124.454106	B	2	Hep3B
100	LE19-131.1	C3A	94.7681992	A	1	Hep3B
50	LE19-131.1	C3A	100.667673	A	1	Hep3B
25	LE19-131.1	C3A	102.745916	A	1	Hep3B
12.5	LE19-131.1	C3A	99.6212128	A	1	Hep3B
100	LE19-131.1	C3A	97.2944483	A	1	Hep3B
50	LE19-131.1	C3A	107.580582	A	1	Hep3B
25	LE19-131.1	C3A	106.640536	A	1	Hep3B
12.5	LE19-131.1	C3A	101.407567	A	1	Hep3B
100	LE19-131.1	C3B	10.262449	B	1	Hep3B
50	LE19-131.1	C3B	96.9521887	B	1	Hep3B
25	LE19-131.1	C3B	97.0666089	B	1	Hep3B
12.5	LE19-131.1	C3B	101.01994	B	1	Hep3B
100	LE19-131.1	C3B	3.37922925	B	1	Hep3B

50	LE19-131.1	C3B	108.889575	B	1	Hep3B
25	LE19-131.1	C3B	112.179403	B	1	Hep3B
12.5	LE19-131.1	C3B	101.391889	B	1	Hep3B
100	LE19-131.1	C3C	85.8784578	C	1	Hep3B
50	LE19-131.1	C3C	108.323812	C	1	Hep3B
25	LE19-131.1	C3C	106.849694	C	1	Hep3B
12.5	LE19-131.1	C3C	97.169687	C	1	Hep3B
100	LE19-131.1	C3C	91.7305623	C	1	Hep3B
50	LE19-131.1	C3C	111.002844	C	1	Hep3B
25	LE19-131.1	C3C	107.543554	C	1	Hep3B
12.5	LE19-131.1	C3C	101.873588	C	1	Hep3B
100	LE19-41.2	C4B	124.986782	B	1	Hep3B
50	LE19-41.2	C4B	111.937219	B	1	Hep3B
25	LE19-41.2	C4B	102.955074	B	1	Hep3B
12.5	LE19-41.2	C4B	98.499362	B	1	Hep3B
100	LE19-41.2	C4B	119.430568	B	1	Hep3B
50	LE19-41.2	C4B	117.545806	B	1	Hep3B
25	LE19-41.2	C4B	106.115137	B	1	Hep3B
12.5	LE19-41.2	C4B	103.333361	B	1	Hep3B
100	LE19-131.1	C3A	124.33955	A	2	Hep3B
50	LE19-131.1	C3A	110.661141	A	2	Hep3B
25	LE19-131.1	C3A	102.985311	A	2	Hep3B
12.5	LE19-131.1	C3A	103.419448	A	2	Hep3B
100	LE19-131.1	C3A	93.9718691	A	2	Hep3B
50	LE19-131.1	C3A	112.722829	A	2	Hep3B
25	LE19-131.1	C3A	107.787375	A	2	Hep3B
12.5	LE19-131.1	C3A	89.6715641	A	2	Hep3B
100	LE19-131.1	C3B	3.23564989	B	2	Hep3B
50	LE19-131.1	C3B	98.0702354	B	2	Hep3B
25	LE19-131.1	C3B	105.034957	B	2	Hep3B
12.5	LE19-131.1	C3B	96.5288322	B	2	Hep3B
100	LE19-131.1	C3B	4.46086783	B	2	Hep3B
50	LE19-131.1	C3B	111.518299	B	2	Hep3B
25	LE19-131.1	C3B	111.949657	B	2	Hep3B
12.5	LE19-131.1	C3B	100.231349	B	2	Hep3B
100	LE19-131.1	C3C	98.2965673	C	2	Hep3B
50	LE19-131.1	C3C	112.8133	C	2	Hep3B
25	LE19-131.1	C3C	109.6357	C	2	Hep3B
12.5	LE19-131.1	C3C	100.542594	C	2	Hep3B

100	LE19-131.1	C3C	95.0297311	C	2	Hep3B
50	LE19-131.1	C3C	128.725386	C	2	Hep3B
25	LE19-131.1	C3C	111.576966	C	2	Hep3B
12.5	LE19-131.1	C3C	100.030955	C	2	Hep3B
100	LE19-41.2	C4B	126.02793	B	2	Hep3B
50	LE19-41.2	C4B	122.47486	B	2	Hep3B
25	LE19-41.2	C4B	109.374477	B	2	Hep3B
12.5	LE19-41.2	C4B	98.133843	B	2	Hep3B
100	LE19-41.2	C4B	127.231225	B	2	Hep3B
50	LE19-41.2	C4B	113.29221	B	2	Hep3B
25	LE19-41.2	C4B	106.526339	B	2	Hep3B
12.5	LE19-41.2	C4B	99.931838	B	2	Hep3B
100	LE17-20	C5A	84.6324303	A	1	Hep3B
50	LE17-20	C5A	106.015736	A	1	Hep3B
25	LE17-20	C5A	105.277264	A	1	Hep3B
12.5	LE17-20	C5A	95.0311023	A	1	Hep3B
100	LE17-20	C5A	92.2181274	A	1	Hep3B
50	LE17-20	C5A	114.5492	A	1	Hep3B
25	LE17-20	C5A	117.78314	A	1	Hep3B
12.5	LE17-20	C5A	108.251965	A	1	Hep3B
100	LE17-20	C5B	97.8955086	B	1	Hep3B
50	LE17-20	C5B	119.861507	B	1	Hep3B
25	LE17-20	C5B	119.801454	B	1	Hep3B
12.5	LE17-20	C5B	110.288117	B	1	Hep3B
100	LE17-20	C5B	93.8345029	B	1	Hep3B
50	LE17-20	C5B	126.18966	B	1	Hep3B
25	LE17-20	C5B	120.8063	B	1	Hep3B
12.5	LE17-20	C5B	116.15844	B	1	Hep3B
100	LE17-20	C5C	100.133521	C	1	Hep3B
50	LE17-20	C5C	116.346923	C	1	Hep3B
25	LE17-20	C5C	127.039022	C	1	Hep3B
12.5	LE17-20	C5C	109.991122	C	1	Hep3B
100	LE17-20	C5C	111.820061	C	1	Hep3B
50	LE17-20	C5C	123.98643	C	1	Hep3B
25	LE17-20	C5C	123.151635	C	1	Hep3B
12.5	LE17-20	C5C	112.783881	C	1	Hep3B
100	LE19-197.1	C6A	60.0573996	A	1	Hep3B
50	LE19-197.1	C6A	104.67941	A	1	Hep3B
25	LE19-197.1	C6A	109.410809	A	1	Hep3B

12.5	LE19-197.1	C6A	114.440391	A	1	Hep3B
100	LE19-197.1	C6A	65.232655	A	1	Hep3B
50	LE19-197.1	C6A	116.761051	A	1	Hep3B
25	LE19-197.1	C6A	123.849676	A	1	Hep3B
12.5	LE19-197.1	C6A	117.219772	A	1	Hep3B
100	LE17-20	C5A	93.1374928	A	2	Hep3B
50	LE17-20	C5A	103.887143	A	2	Hep3B
25	LE17-20	C5A	92.5239956	A	2	Hep3B
12.5	LE17-20	C5A	115.994071	A	2	Hep3B
100	LE17-20	C5A	95.76422	A	2	Hep3B
50	LE17-20	C5A	110.202307	A	2	Hep3B
25	LE17-20	C5A	108.918598	A	2	Hep3B
12.5	LE17-20	C5A	120.30832	A	2	Hep3B
100	LE17-20	C5B	98.0759345	B	2	Hep3B
50	LE17-20	C5B	119.531483	B	2	Hep3B
25	LE17-20	C5B	115.700448	B	2	Hep3B
12.5	LE17-20	C5B	117.876694	B	2	Hep3B
100	LE17-20	C5B	96.443831	B	2	Hep3B
50	LE17-20	C5B	126.617975	B	2	Hep3B
25	LE17-20	C5B	122.726984	B	2	Hep3B
12.5	LE17-20	C5B	130.448687	B	2	Hep3B
100	LE17-20	C5C	100.024995	C	2	Hep3B
50	LE17-20	C5C	128.064377	C	2	Hep3B
25	LE17-20	C5C	123.158668	C	2	Hep3B
12.5	LE17-20	C5C	119.843903	C	2	Hep3B
100	LE17-20	C5C	97.5888313	C	2	Hep3B
50	LE17-20	C5C	125.27593	C	2	Hep3B
25	LE17-20	C5C	138.96732	C	2	Hep3B
12.5	LE17-20	C5C	122.981068	C	2	Hep3B
100	LE19-197.1	C6A	68.5496407	A	2	Hep3B
50	LE19-197.1	C6A	114.418359	A	2	Hep3B
25	LE19-197.1	C6A	119.597597	A	2	Hep3B
12.5	LE19-197.1	C6A	118.905347	A	2	Hep3B
100	LE19-197.1	C6A	61.9752065	A	2	Hep3B
50	LE19-197.1	C6A	115.782118	A	2	Hep3B
25	LE19-197.1	C6A	123.954302	A	2	Hep3B
12.5	LE19-197.1	C6A	126.487692	A	2	Hep3B
100	LE19-197.1	C6B	62.4351393	B	1	Hep3B
50	LE19-197.1	C6B	91.1526337	B	1	Hep3B

25	LE19-197.1	C6B	107.768866	B	1	Hep3B
12.5	LE19-197.1	C6B	107.749542	B	1	Hep3B
100	LE19-197.1	C6B	60.9067622	B	1	Hep3B
50	LE19-197.1	C6B	81.5899462	B	1	Hep3B
25	LE19-197.1	C6B	92.1898847	B	1	Hep3B
12.5	LE19-197.1	C6B	109.538644	B	1	Hep3B
100	LE19-197.1	C6C	97.1912244	C	1	Hep3B
50	LE19-197.1	C6C	91.8970523	C	1	Hep3B
25	LE19-197.1	C6C	91.1787954	C	1	Hep3B
12.5	LE19-197.1	C6C	81.531677	C	1	Hep3B
100	LE19-197.1	C6C	104.973431	C	1	Hep3B
50	LE19-197.1	C6C	99.1488908	C	1	Hep3B
25	LE19-197.1	C6C	99.2125112	C	1	Hep3B
12.5	LE19-197.1	C6C	89.0067816	C	1	Hep3B
100	LE19-84.1	C7C	117.921083	C	1	Hep3B
50	LE19-84.1	C7C	102.995252	C	1	Hep3B
25	LE19-84.1	C7C	96.4045903	C	1	Hep3B
12.5	LE19-84.1	C7C	95.7829532	C	1	Hep3B
100	LE19-84.1	C7C	114.692494	C	1	Hep3B
50	LE19-84.1	C7C	107.817622	C	1	Hep3B
25	LE19-84.1	C7C	101.209717	C	1	Hep3B
12.5	LE19-84.1	C7C	94.155578	C	1	Hep3B
100	LE19-84.1	C7B	107.161202	B	1	Hep3B
50	LE19-84.1	C7B	104.856298	B	1	Hep3B
25	LE19-84.1	C7B	104.92973	B	1	Hep3B
12.5	LE19-84.1	C7B	94.7742422	B	1	Hep3B
100	LE19-84.1	C7B	118.970523	B	1	Hep3B
50	LE19-84.1	C7B	109.892421	B	1	Hep3B
25	LE19-84.1	C7B	99.5389376	B	1	Hep3B
12.5	LE19-84.1	C7B	96.6602613	B	1	Hep3B
100	Doxorubicin	Dox	71.149207	X	1	Hep3B
50	Doxorubicin	Dox	36.1269931	X	1	Hep3B
25	Doxorubicin	Dox	69.1242667	X	1	Hep3B
12.5	Doxorubicin	Dox	75.2196005	X	1	Hep3B
100	Doxorubicin	Dox	67.1591914	X	1	Hep3B
50	Doxorubicin	Dox	58.9652376	X	1	Hep3B
25	Doxorubicin	Dox	42.0540354	X	1	Hep3B
12.5	Doxorubicin	Dox	74.9575347	X	1	Hep3B
100	LE19-197.1	C6B	59.7798394	B	2	Hep3B

50	LE19-197.1	C6B	96.7581957	B	2	Hep3B
25	LE19-197.1	C6B	109.695111	B	2	Hep3B
12.5	LE19-197.1	C6B	123.07635	B	2	Hep3B
100	LE19-197.1	C6B	59.4249639	B	2	Hep3B
50	LE19-197.1	C6B	94.5913483	B	2	Hep3B
25	LE19-197.1	C6B	96.8456992	B	2	Hep3B
12.5	LE19-197.1	C6B	125.197177	B	2	Hep3B
100	LE19-197.1	C6C	101.214396	C	2	Hep3B
50	LE19-197.1	C6C	90.6358633	C	2	Hep3B
25	LE19-197.1	C6C	98.1741329	C	2	Hep3B
12.5	LE19-197.1	C6C	85.7975646	C	2	Hep3B
100	LE19-197.1	C6C	106.173579	C	2	Hep3B
50	LE19-197.1	C6C	99.5651154	C	2	Hep3B
25	LE19-197.1	C6C	101.774094	C	2	Hep3B
12.5	LE19-197.1	C6C	90.9709695	C	2	Hep3B
100	LE19-84.1	C7C	115.998284	C	2	Hep3B
50	LE19-84.1	C7C	106.097094	C	2	Hep3B
25	LE19-84.1	C7C	99.548587	C	2	Hep3B
12.5	LE19-84.1	C7C	96.0095541	C	2	Hep3B
100	LE19-84.1	C7C	121.625087	C	2	Hep3B
50	LE19-84.1	C7C	110.191937	C	2	Hep3B
25	LE19-84.1	C7C	111.344391	C	2	Hep3B
12.5	LE19-84.1	C7C	102.326663	C	2	Hep3B
100	LE19-84.1	C7B	120.055532	B	2	Hep3B
50	LE19-84.1	C7B	123.088341	B	2	Hep3B
25	LE19-84.1	C7B	107.606369	B	2	Hep3B
12.5	LE19-84.1	C7B	95.3001271	B	2	Hep3B
100	LE19-84.1	C7B	123.028709	B	2	Hep3B
50	LE19-84.1	C7B	114.844533	B	2	Hep3B
25	LE19-84.1	C7B	116.968277	B	2	Hep3B
12.5	LE19-84.1	C7B	102.483521	B	2	Hep3B
100	Doxorubicin	Dox	59.1732577	X	2	Hep3B
50	Doxorubicin	Dox	47.6687606	X	2	Hep3B
25	Doxorubicin	Dox	43.3888262	X	2	Hep3B
12.5	Doxorubicin	Dox	64.2991642	X	2	Hep3B
100	Doxorubicin	Dox	70.9225609	X	2	Hep3B
50	Doxorubicin	Dox	55.138964	X	2	Hep3B
25	Doxorubicin	Dox	47.9808884	X	2	Hep3B
12.5	Doxorubicin	Dox	74.1534455	X	2	Hep3B

NA	Control (Vehicle)	NA	119.7498	NA	NA	Hep3B
NA	Control (Vehicle)	NA	122.8228	NA	NA	Hep3B
NA	Control (Vehicle)	NA	115.7631	NA	NA	Hep3B
NA	Control (Vehicle)	NA	70.48745	NA	NA	Hep3B
NA	Control (Vehicle)	NA	70.48745	NA	NA	Hep3B
NA	Control (Vehicle)	NA	108.0646	NA	NA	Hep3B
NA	Control (Vehicle)	NA	94.83625	NA	NA	Hep3B
NA	Control (Vehicle)	NA	99.84572	NA	NA	Hep3B

Table S4.9.5: Concentration of cyanotoxins detected via targeted LC-MS in each crude extract in the 100 μM dose exposed to each cell line in viability experiments.

Sample	Culture	Replicate	Metabolite	ppb	sd	MW (g/mol)	Final stock added for 100 uM (uL)	Final assay volume (uL)	Stock (mg/mL)	Toxin per mg crude (ug/mg)	Toxin conc in crude (ug/mL)	uM of toxin in stock	Dose (uM)	Final uM added
C1A	LE18-22.4	A	MCLR	1793.60124	65.6840507	995.2	1	100	100	0.01793601	1.79360124	1.80225205	100	0.01802252
C1B	LE18-22.4	B	MCLR	4406.8309	89.5388771	995.2	1	100	100	0.04406831	4.4068309	4.42808571	100	0.04428086
C1C	LE18-22.4	C	MCLR	3967.30283	197.069356	995.2	1	100	100	0.03967303	3.96730283	3.98643773	100	0.03986438
C2A	LE19-55.1	A	MCLR	738.445796	22.3324303	995.2	1	100	100	0.00738446	0.7384458	0.74200743	100	0.00742007
C2B	LE19-55.1	B	MCLR	642.814993	39.3716051	995.2	1	100	100	0.00642815	0.64281499	0.64591539	100	0.00645915
C2C	LE19-55.1	C	MCLR	111.644324	15.1565257	995.2	1	100	100	0.00111644	0.11164432	0.1121828	100	0.00112183
C3A	LE19-131.1	A	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C3B	LE19-131.1	B	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C3C	LE19-131.1	C	MCLR	0	0	995.2	1	100	100	0	0	0	100	0

C4B	LE19-41.2	B	MCLR	2599.44458	147.314708	995.2	1	100	100	0.02599445	2.59944458	2.61198209	100	0.02611982
C4C	LE19-41.2	C	MCLR	2183.61067	154.940874	995.2	1	100	100	0.02183611	2.18361067	2.19414255	100	0.02194143
C5A	LE17-20	A	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C5B	LE17-20	B	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C5C	LE17-20	C	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C6A	LE19-197.1	A	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C6B	LE19-197.1	B	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C6C	LE19-197.1	C	MCLR	0	0	995.2	1	100	100	0	0	0	100	0
C7B	LE19-84.1	B	MCLR	272.936695	4.771 91568	995.2	1	100	100	0.00272937	0.2729367	0.27425311	100	0.00274253
C7C	LE19-84.1	C	MCLR	739.881607	64.4604999	995.2	1	100	100	0.00739882	0.73988161	0.74345017	100	0.007 4345
C1A	LE18-22.4	A	MCRR	6994.90227	89.5386476	1038.2	1	100	100	0.06994902	6.99490227	6.73752868	101	0.06737529
C1B	LE18-22.4	B	MCRR	22776.7598	731.963202	1038.2	1	100	100	0.2277676	22.7767598	21.9387014	102	0.21938701
C1C	LE18-22.4	C	MCRR	19043.4746	660.582706	1038.2	1	100	100	0.19043475	19.04 34746	18.34 27804	103	0.183 4278
C2A	LE19-55.1	A	MCRR	6564.71981	82.6941843	1038.2	1	100	100	0.0656472	6.564 71981	6.323 17454	104	0.063 23175
C2B	LE19-55.1	B	MCRR	5573.30051	52.6052263	1038.2	1	100	100	0.05573301	5.573 30051	5.368 23398	105	0.053 68234
C2C	LE19-55.1	C	MCRR	299.520517	16.13 92132	1038.2	1	100	100	0.00299521	0.299 52052	0.288 49982	106	0.002 885
C3A	LE19-131.1	A	MCRR	9.001 38337	0.052 60896	1038.2	1	100	100	9.001 38E-05	0.009 00138	0.008 67018	107	8.670 18E-05
C3B	LE19-131.1	B	MCRR	0	0	1038.2	1	100	100	0	0	0	108	0
C3C	LE19-131.1	C	MCRR	0	0	1038.2	1	100	100	0	0	0	109	0

C4B	LE19-41.2	B	MCRR	24113.8748	92.4160925	1038.2	1	100	100	0.24113875	24.1138748	23.226618	110	0.23226618
C4C	LE19-41.2	C	MCRR	20505.8562	519.595862	1038.2	1	100	100	0.20505856	20.5058562	19.7513545	111	0.19751355
C5A	LE17-20	A	MCRR	0	0	1038.2	1	100	100	0	0	0	112	0
C5B	LE17-20	B	MCRR	0	0	1038.2	1	100	100	0	0	0	113	0
C5C	LE17-20	C	MCRR	0	0	1038.2	1	100	100	0	0	0	114	0
C6A	LE19-197.1	A	MCRR	0	0	1038.2	1	100	100	0	0	0	115	0
C6B	LE19-197.1	B	MCRR	0	0	1038.2	1	100	100	0	0	0	116	0
C6C	LE19-197.1	C	MCRR	0	0	1038.2	1	100	100	0	0	0	117	0
C7B	LE19-84.1	B	MCRR	616.036517	2.53298661	1038.2	1	100	100	0.00616037	0.61603652	0.59336979	118	0.0059337
C7C	LE19-84.1	C	MCRR	616.281813	1.88576478	1038.2	1	100	100	0.00616282	0.61628181	0.59360606	119	0.00593606
C1A	LE18-22.4	A	MCYR	0	0	1045.2	1	100	100	0	0	0	120	0
C1B	LE18-22.4	B	MCYR	225.600556	11.851421	1045.2	1	100	100	0.00225601	0.22560056	0.21584439	121	0.00215844
C1C	LE18-22.4	C	MCYR	183.663205	23.1771482	1045.2	1	100	100	0.00183663	0.18366321	0.17572063	122	0.00175721
C2A	LE19-55.1	A	MCYR	526.375385	33.7967618	1045.2	1	100	100	0.00526375	0.52637539	0.50361212	123	0.00503612
C2B	LE19-55.1	B	MCYR	426.772731	22.4155649	1045.2	1	100	100	0.00426773	0.42677273	0.40831681	124	0.00408317
C2C	LE19-55.1	C	MCYR	90.5210642	9.07511679	1045.2	1	100	100	0.00090521	0.09052106	0.08660645	125	0.00086607
C3A	LE19-131.1	A	MCYR	0	548.586172	1045.2	1	100	100	0	0	0	126	0
C3B	LE19-131.1	B	MCYR	0	0	1045.2	1	100	100	0	0	0	127	0
C3C	LE19-131.1	C	MCYR	0	0	1045.2	1	100	100	0	0	0	128	0

C4B	LE19-41.2	B	MCYR	1721.5937	147.608951	1045.2	1	100	100	0.01721594	1.7215937	1.64714284	129	0.01647143
C4C	LE19-41.2	C	MCYR	2235.90233	124.411402	1045.2	1	100	100	0.02235902	2.23590233	2.13921004	130	0.0213921
C5A	LE17-20	A	MCYR	0	0	1045.2	1	100	100	0	0	0	131	0
C5B	LE17-20	B	MCYR	0	0	1045.2	1	100	100	0	0	0	132	0
C5C	LE17-20	C	MCYR	0	0	1045.2	1	100	100	0	0	0	133	0
C6A	LE19-197.1	A	MCYR	0	0	1045.2	1	100	100	0	0	0	134	0
C6B	LE19-197.1	B	MCYR	0	0	1045.2	1	100	100	0	0	0	135	0
C6C	LE19-197.1	C	MCYR	0	0	1045.2	1	100	100	0	0	0	136	0
C7B	LE19-84.1	B	MCYR	1081.67109	43.7326338	1045.2	1	100	100	0.01081671	1.08167109	1.03489388	137	0.01034894
C7C	LE19-84.1	C	MCYR	3825.30804	290.4802	1045.2	1	100	100	0.03825308	3.82530804	3.6598814	138	0.03659881
C1A	LE18-22.4	A	MCLA	0	0	910.1	1	100	100	0	0	0	139	0
C1B	LE18-22.4	B	MCLA	0	0	910.1	1	100	100	0	0	0	140	0
C1C	LE18-22.4	C	MCLA	0	0	910.1	1	100	100	0	0	0	141	0
C2A	LE19-55.1	A	MCLA	0	0	910.1	1	100	100	0	0	0	142	0
C2B	LE19-55.1	B	MCLA	0	0	910.1	1	100	100	0	0	0	143	0
C2C	LE19-55.1	C	MCLA	0	0	910.1	1	100	100	0	0	0	144	0
C3A	LE19-131.1	A	MCLA	0	0	910.1	1	100	100	0	0	0	145	0
C3B	LE19-131.1	B	MCLA	0	0	910.1	1	100	100	0	0	0	146	0
C3C	LE19-131.1	C	MCLA	0	0	910.1	1	100	100	0	0	0	147	0

C4B	LE19-41.2	B	MCLA	0	0	910.1	1	100	100	0	0	0	148	0
C4C	LE19-41.2	C	MCLA	0	0	910.1	1	100	100	0	0	0	149	0
C5A	LE17-20	A	MCLA	0	0	910.1	1	100	100	0	0	0	150	0
C5B	LE17-20	B	MCLA	0	0	910.1	1	100	100	0	0	0	151	0
C5C	LE17-20	C	MCLA	0	0	910.1	1	100	100	0	0	0	152	0
C6A	LE19-197.1	A	MCLA	0	0	910.1	1	100	100	0	0	0	153	0
C6B	LE19-197.1	B	MCLA	0	0	910.1	1	100	100	0	0	0	154	0
C6C	LE19-197.1	C	MCLA	0	0	910.1	1	100	100	0	0	0	155	0
C7B	LE19-84.1	B	MCLA	274.147512	1.9360354	910.1	1	100	100	0.00274148	0.27414751	0.3012279	156	0.00301228
C7C	LE19-84.1	C	MCLA	628.135512	9.7652708	910.1	1	100	100	0.00628136	0.62813551	0.6901829 6	157	0.00690183
C1A	LE18-22.4	A	MCLF	0	0	986.2	1	100	100	0	0	0	158	0
C1B	LE18-22.4	B	MCLF	120.793415	3.92614094	986.2	1	100	100	0.00120793	0.12079342	0.1224836 9	159	0.00122484
C1C	LE18-22.4	C	MCLF	134.329132	13.6497032	986.2	1	100	100	0.00134329	0.13432913	0.1362088 1	160	0.00136209
C2A	LE19-55.1	A	MCLF	0	0	986.2	1	100	100	0	0	0	161	0
C2B	LE19-55.1	B	MCLF	0	0	986.2	1	100	100	0	0	0	162	0
C2C	LE19-55.1	C	MCLF	1.81116345	0.06484307	986.2	1	100	100	1.81116E-05	0.00181116	0.0018365 1	163	1.83651E-05
C3A	LE19-131.1	A	MCLF	0	0	986.2	1	100	100	0	0	0	164	0
C3B	LE19-131.1	B	MCLF	0	0	986.2	1	100	100	0	0	0	165	0
C3C	LE19-131.1	C	MCLF	0	0	986.2	1	100	100	0	0	0	166	0

C4B	LE19-41.2	B	MCLF	0	0	986.2	1	100	100	0	0	0	167	0
C4C	LE19-41.2	C	MCLF	0	0	986.2	1	100	100	0	0	0	168	0
C5A	LE17-20	A	MCLF	0	0	986.2	1	100	100	0	0	0	169	0
C5B	LE17-20	B	MCLF	0	0	986.2	1	100	100	0	0	0	170	0
C5C	LE17-20	C	MCLF	0	0	986.2	1	100	100	0	0	0	171	0
C6A	LE19-197.1	A	MCLF	0	0	986.2	1	100	100	0	0	0	172	0
C6B	LE19-197.1	B	MCLF	0	0	986.2	1	100	100	0	0	0	173	0
C6C	LE19-197.1	C	MCLF	0	0	986.2	1	100	100	0	0	0	174	0
C7B	LE19-84.1	B	MCLF	0	0	986.2	1	100	100	0	0	0	175	0
C7C	LE19-84.1	C	MCLF	0	0	986.2	1	100	100	0	0	0	176	0
C1A	LE18-22.4	A	Asp3 MCLR	14.2282312	8.32401468	981.15	1	100	100	0.00014228	0.01422823	0.01450159	177	0.00014502
C1B	LE18-22.4	B	Asp3 MCLR	818.213283	40.2404669	981.15	1	100	100	0.00818213	0.81821328	0.83393292	178	0.00833933
C1C	LE18-22.4	C	Asp3 MCLR	765.512574	61.929256	981.15	1	100	100	0.00765513	0.76551257	0.78021972	179	0.0078022
C2A	LE19-55.1	A	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	180	0
C2B	LE19-55.1	B	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	181	0
C2C	LE19-55.1	C	Asp3 MCLR	43.3669781	2.5391738	981.15	1	100	100	0.00043367	0.04336698	0.04420015	182	0.000442
C3A	LE19-131.1	A	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	183	0
C3B	LE19-131.1	B	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	184	0
C3C	LE19-131.1	C	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	185	0

C4 B	LE19-41.2	B	Asp3 MCLR	675.502562	11.720664	981.15	1	100	100	0.00675503	0.67550256	0.68848042	186	0.006 8848
C4 C	LE19-41.2	C	Asp3 MCLR	346.526315	11.2472885	981.15	1	100	100	0.00346526	0.34652632	0.35318383	187	0.003 53184
C5 A	LE17-20	A	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	188	0
C5 B	LE17-20	B	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	189	0
C5 C	LE17-20	C	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	190	0
C6 A	LE19-197.1	A	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	191	0
C6 B	LE19-197.1	B	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	192	0
C6 C	LE19-197.1	C	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	193	0
C7 B	LE19-84.1	B	Asp3 MCLR	0	0	981.15	1	100	100	0	0	0	194	0
C7 C	LE19-84.1	C	Asp3 MCLR	374.633005	13.7894701	981.15	1	100	100	0.00374633	0.37463301	0.38183051	195	0.003 81831
C1 A	LE18-22.4	A	MCHilR	338.507937	12.4282363	1009.2	1	100	100	0.00338508	0.33850794	0.33542205	196	0.003 35422
C1 B	LE18-22.4	B	MCHilR	633.613422	25.6536637	1009.2	1	100	100	0.00633613	0.63361342	0.62783732	197	0.006 27837
C1 C	LE18-22.4	C	MCHilR	688.303746	37.6970167	1009.2	1	100	100	0.00688304	0.68830375	0.68202908	198	0.006 82029
C2 A	LE19-55.1	A	MCHilR	658.00103	1.72934843	1009.2	1	100	100	0.00658001	0.65800103	0.65200261	199	0.006 52003
C2 B	LE19-55.1	B	MCHilR	604.865668	60.8824365	1009.2	1	100	100	0.00604866	0.60486567	0.59935163	200	0.005 99352
C2 C	LE19-55.1	C	MCHilR	53.8211296	4.44291329	1009.2	1	100	100	0.00053821	0.05382113	0.05333049	201	0.000 53331
C3 A	LE19-131.1	A	MCHilR	0	0	1009.2	1	100	100	0	0	0	202	0
C3 B	LE19-131.1	B	MCHilR	0	0	1009.2	1	100	100	0	0	0	203	0
C3 C	LE19-131.1	C	MCHilR	0	0	1009.2	1	100	100	0	0	0	204	0

C4B	LE19-41.2	B	MCHilR	1171.06563	105.187795	1009.2	1	100	100	0.01171066	1.17106563	1.16039004	205	0.0116039
C4C	LE19-41.2	C	MCHilR	1535.42413	92.346639	1009.2	1	100	100	0.01535424	1.53542413	1.52142701	206	0.01521427
C5A	LE17-20	A	MCHilR	0	0	1009.2	1	100	100	0	0	0	207	0
C5B	LE17-20	B	MCHilR	0	0	1009.2	1	100	100	0	0	0	208	0
C5C	LE17-20	C	MCHilR	0	0	1009.2	1	100	100	0	0	0	209	0
C6A	LE19-197.1	A	MCHilR	0	0	1009.2	1	100	100	0	0	0	210	0
C6B	LE19-197.1	B	MCHilR	0	0	1009.2	1	100	100	0	0	0	211	0
C6C	LE19-197.1	C	MCHilR	0	0	1009.2	1	100	100	0	0	0	212	0
C7B	LE19-84.1	B	MCHilR	231.183997	2.24546573	1009.2	1	100	100	0.00231184	0.231184	0.22907649	213	0.00229077
C7C	LE19-84.1	C	MCHilR	286.169096	9.44882869	1009.2	1	100	100	0.00286169	0.2861691	0.28356034	214	0.0028356
C1A	LE18-22.4	A	MCLY	228.207198	4.95223631	1002.2	1	100	100	0.00228207	0.2282072	0.22770624	215	0.00227706
C1B	LE18-22.4	B	MCLY	511.276214	10.8707233	1002.2	1	100	100	0.00511276	0.51127621	0.51015388	216	0.00510154
C1C	LE18-22.4	C	MCLY	507.020751	5.67767204	1002.2	1	100	100	0.00507021	0.50702075	0.50590775	217	0.00505908
C2A	LE19-55.1	A	MCLY	85.3378672	2.34186725	1002.2	1	100	100	0.00085338	0.08533787	0.08515054	218	0.00085151
C2B	LE19-55.1	B	MCLY	81.7868745	2.36162592	1002.2	1	100	100	0.00081787	0.08178687	0.08160734	219	0.00081607
C2C	LE19-55.1	C	MCLY	20.0676932	0.39985302	1002.2	1	100	100	0.00020068	0.02006769	0.02002364	220	0.00020024
C3A	LE19-131.1	A	MCLY	0	0	1002.2	1	100	100	0	0	0	221	0
C3B	LE19-131.1	B	MCLY	0	0	1002.2	1	100	100	0	0	0	222	0
C3C	LE19-131.1	C	MCLY	0	0	1002.2	1	100	100	0	0	0	223	0

C4B	LE19-41.2	B	MCLY	267.979271	7.15805018	1002.2	1	100	100	0.00267979	0.26797927	0.26739101	224	0.00267391
C4C	LE19-41.2	C	MCLY	247.5012	0.72645965	1002.2	1	100	100	0.00247501	0.2475012	0.24695789	225	0.00246958
C5A	LE17-20	A	MCLY	0	0	1002.2	1	100	100	0	0	0	226	0
C5B	LE17-20	B	MCLY	0	0	1002.2	1	100	100	0	0	0	227	0
C5C	LE17-20	C	MCLY	0	0	1002.2	1	100	100	0	0	0	228	0
C6A	LE19-197.1	A	MCLY	0	0	1002.2	1	100	100	0	0	0	229	0
C6B	LE19-197.1	B	MCLY	0	0	1002.2	1	100	100	0	0	0	230	0
C6C	LE19-197.1	C	MCLY	0	0	1002.2	1	100	100	0	0	0	231	0
C7B	LE19-84.1	B	MCLY	0	0	1002.2	1	100	100	0	0	0	232	0
C7C	LE19-84.1	C	MCLY	0	0	1002.2	1	100	100	0	0	0	233	0
C1A	LE18-22.4	A	ABNA	0	0	844	1	100	100	0	0	0	234	0
C1B	LE18-22.4	B	ABNA	0	0	844	1	100	100	0	0	0	235	0
C1C	LE18-22.4	C	ABNA	0	0	844	1	100	100	0	0	0	236	0
C2A	LE19-55.1	A	ABNA	0	0	844	1	100	100	0	0	0	237	0
C2B	LE19-55.1	B	ABNA	0	0	844	1	100	100	0	0	0	238	0
C2C	LE19-55.1	C	ABNA	0	0	844	1	100	100	0	0	0	239	0
C3A	LE19-131.1	A	ABNA	0	0	844	1	100	100	0	0	0	240	0
C3B	LE19-131.1	B	ABNA	0	0	844	1	100	100	0	0	0	241	0
C3C	LE19-131.1	C	ABNA	0	0	844	1	100	100	0	0	0	242	0

C4B	LE19-41.2	B	ABNA	0	0	844	1	100	100	0	0	0	243	0
C4C	LE19-41.2	C	ABNA	0	0	844	1	100	100	0	0	0	244	0
C5A	LE17-20	A	ABNA	0	0	844	1	100	100	0	0	0	245	0
C5B	LE17-20	B	ABNA	0	0	844	1	100	100	0	0	0	246	0
C5C	LE17-20	C	ABNA	0	0	844	1	100	100	0	0	0	247	0
C6A	LE19-197.1	A	ABNA	0.83234972	0.34691938	844	1	100	100	8.3235E-06	0.00083235	0.0009862	248	9.86196E-06
C6B	LE19-197.1	B	ABNA	16.2971254	1.62311045	844	1	100	100	0.00016297	0.01629713	0.01930939	249	0.00019309
C6C	LE19-197.1	C	ABNA	7.13777948	0.260429	844	1	100	100	7.13778E-05	0.00713778	0.00845709	250	8.45708E-05
C7B	LE19-84.1	B	ABNA	0	0	844	1	100	100	0	0	0	251	0
C7C	LE19-84.1	C	ABNA	0	0	844	1	100	100	0	0	0	252	0
C1A	LE18-22.4	A	ABNB	0	0	837	1	100	100	0	0	0	253	0
C1B	LE18-22.4	B	ABNB	0	0	837	1	100	100	0	0	0	254	0
C1C	LE18-22.4	C	ABNB	0	0	837	1	100	100	0	0	0	255	0
C2A	LE19-55.1	A	ABNB	0	0	837	1	100	100	0	0	0	256	0
C2B	LE19-55.1	B	ABNB	0	0	837	1	100	100	0	0	0	257	0
C2C	LE19-55.1	C	ABNB	0	0	837	1	100	100	0	0	0	258	0
C3A	LE19-131.1	A	ABNB	0	0	837	1	100	100	0	0	0	259	0
C3B	LE19-131.1	B	ABNB	0	0	837	1	100	100	0	0	0	260	0
C3C	LE19-131.1	C	ABNB	0	0	837	1	100	100	0	0	0	261	0

C4 B	LE19-41.2	B	ABNB	0	0	837	1	100	100	0	0	0	262	0
C4 C	LE19-41.2	C	ABNB	0	0	837	1	100	100	0	0	0	263	0
C5 A	LE17-20	A	ABNB	0	0	837	1	100	100	0	0	0	264	0
C5 B	LE17-20	B	ABNB	0	0	837	1	100	100	0	0	0	265	0
C5 C	LE17-20	C	ABNB	0	0	837	1	100	100	0	0	0	266	0
C6 A	LE19-197.1	A	ABNB	5.33872621	0.06317831	837	1	100	100	5.33873E-05	0.00533873	0.00637841	267	6.37841E-05
C6 B	LE19-197.1	B	ABNB	13.8225447	0.51809544	837	1	100	100	0.00013823	0.01382255	0.01651439	268	0.00016514
C6 C	LE19-197.1	C	ABNB	9.09958833	0.09290935	837	1	100	100	9.09959E-05	0.00909959	0.01087167	269	0.00010872
C7 B	LE19-84.1	B	ABNB	0	0	837	1	100	100	0	0	0	270	0
C7 C	LE19-84.1	C	ABNB	0	0	837	1	100	100	0	0	0	271	0

Equation 4.9.1: Calculation of toxin added to cells from *Microcystis* culture extracts.

$$ug\ CP\ per\ mg\ crude = \frac{CP\ (\frac{ug}{L})}{100{,}000\frac{mg}{L}\ stock}$$

$$\frac{ug}{mL}\ CP\ in\ crude = \left(\frac{ug}{mg}\ CP\ in\ stock\right) * 100\frac{mg}{mL}\ stock$$

$$uM\ CP\ in\ stock = \frac{\frac{ug}{mL}\ toxin}{MW\ of\ CP\ (\frac{g}{mol})} * 1000$$

$$CP\ added\ to\ 100uM\ dose = uM\ toxin\ in\ stock * \frac{1uL\ added}{100uL\ total\ assay\ volume}$$

*CP = cyanopeptide

4.9.2 SI Figures

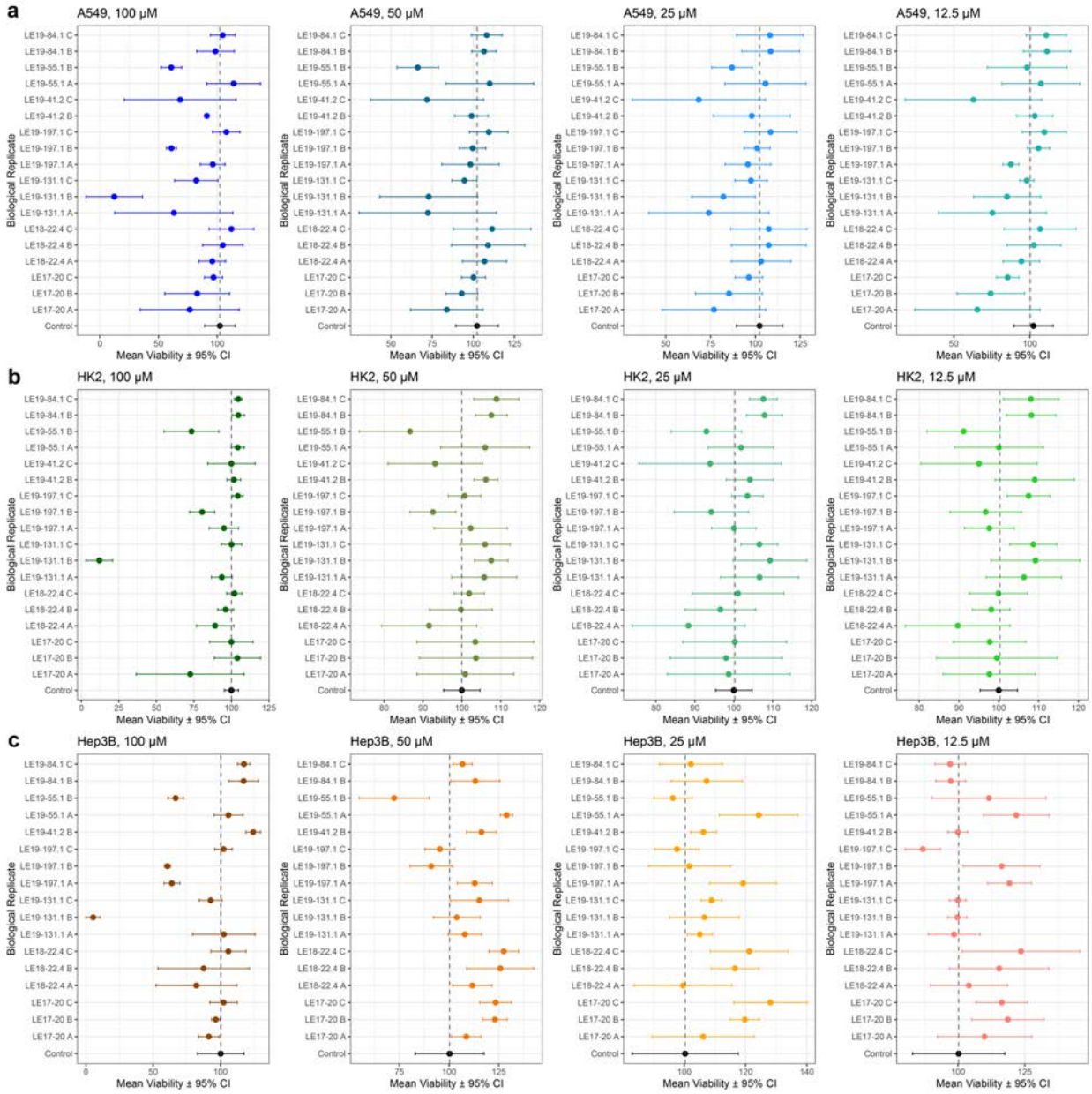


Figure S4.9.1: Mean viability and ±95% confidence interval of (a) lung (A549), (b) kidney (HK2), and (c) liver (Hep-3B) cell lines following exposure to four doses (100, 50, 25, 12.5 µM) crude extracts of *Microcystis* isolates from the WLECC relative to the DMSO vehicle control (shown in black). The dotted line represents the mean viability of DMSO vehicle controls. Isolate effects below the dotted line indicate significant cytotoxicity relative to control.

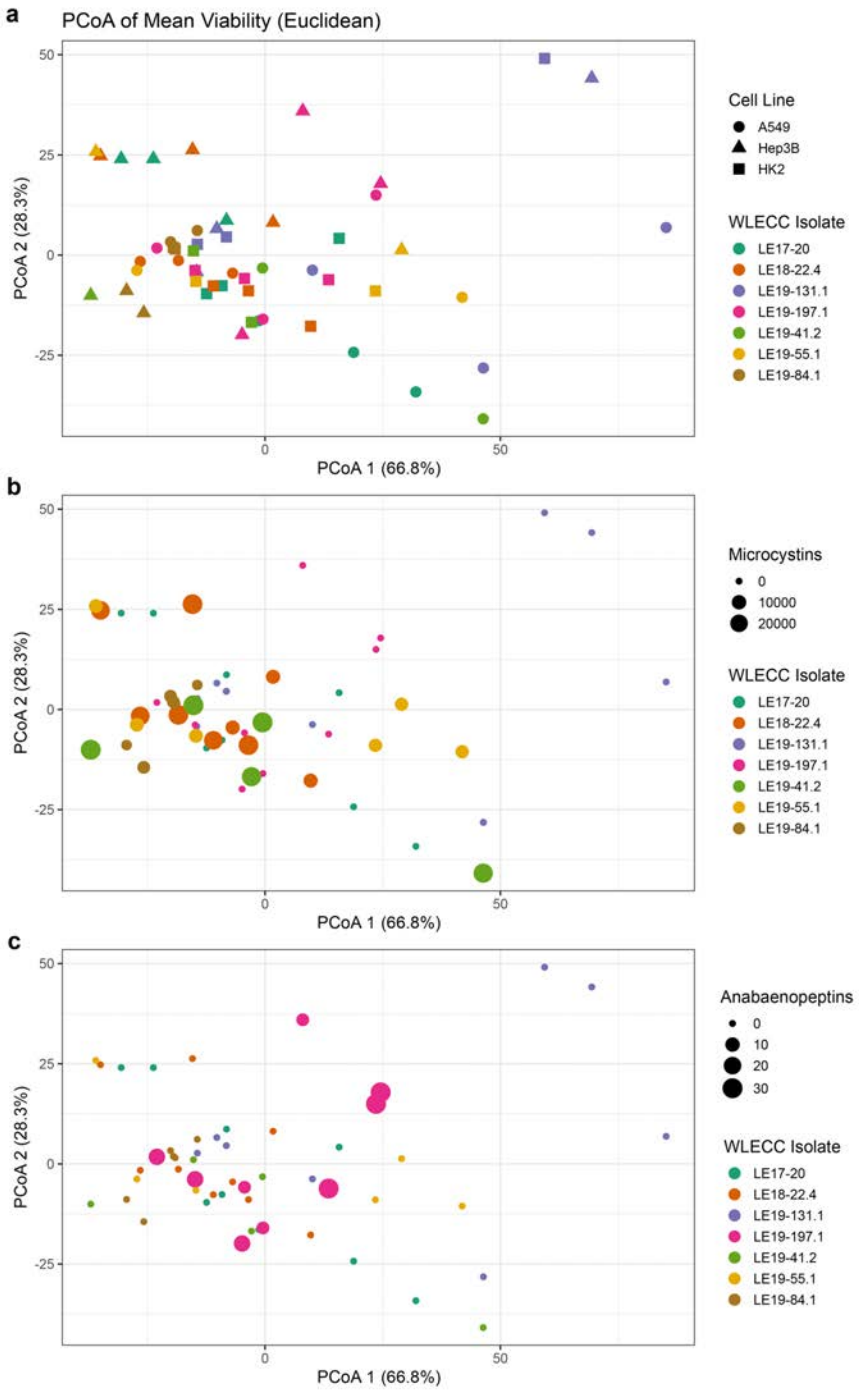


Figure S4.9.2: Principal Coordinates Analysis (PCoA) of viability responses to WLECC *Microcystis* isolate extracts in human cell lines lung (A549), liver (Hep-3B), and kidney (HK2), based on Euclidean distances of mean viability. Each point represents one isolate–cell line combination, colored by isolate identity. Shapes and sizes of points vary across panels where in **(a)** shapes indicate the cell line tested, **(b)** point size is scaled to total microcystin content (ppb) in the extract, and **(c)** point size is scaled to total anabaenopeptin content (ppb) in the extract. PCoA axes represent the proportion of total variance explained (PCoA 1: 66.8%, PCoA 2: 28.3%).

Chapter 5 Conclusions

5.1 Summary and Overview

This dissertation advances our understanding of the environmental and ecological drivers shaping the spatiotemporal dynamics of bloom-forming cyanobacteria and their biosynthesis in large, eutrophic Great Lakes basins. It also provides evidence that the complex chemical ecology of cyanoHABs poses unexplored risks to human and environmental health through the co-occurrence of bioactive secondary metabolites in freshwater bodies relied on for drinking, subsistence, and recreation.

In Chapter 2, integrated metagenomic and metabolomic analyses revealed temporal distributions, successions, and simultaneous occurrences of multiple cyanopeptides in western Lake Erie (WLE) cyanoHABs. These phases co-varied with shifts in nitrogen form and availability, dissolved organic matter makeup, and eDNA detection of grazers, providing testable hypotheses for the mechanisms driving biosynthetic successions within cyanoHABs. In Chapter 3, metagenomic analyses revealed spatial differences in cyanobacterial dominance and biosynthetic potential across the Winam Gulf, Lake Victoria (Kenya). *Dolichospermum* abundance was identified at sites with visible, large cyanoHAB events, while *Microcystis* was identified to encode the *mcy* operon, enabling microcystin production, at high abundance in turbid regions near river mouths. The *Microcystis*-dominated blooms could go unnoticed by lakeside communities using the water for daily tasks like laundry or bathing, due to low visibility and the absence of green scums, underscoring the hidden health risks these events pose. Moreover, over 300 biosynthetic gene clusters (BGCs) without known synthesis products were

identified in the gulf, revealing a widely unexplored chemical space encoded by these organisms. Together, Chapters 2 and 3 demonstrate the power of spatiotemporally resolved multi-omic datasets to uncover the distribution, dynamics, and co-varying environmental pressures of cyanobacterial metabolites. These data reveal the importance of evaluating not only well-established cyanotoxins like microcystins, but also lesser studied yet potentially toxic cyanobacterial secondary metabolites in future monitoring and management efforts.

Chapter 4 examines the toxicity of co-occurring cyanopeptides, specifically microcystins and anabaenopeptins, as well as complex *Microcystis* metabolite mixtures in human liver, kidney, and lung cells. Anabaenopeptins exhibited toxic effects comparable to microcystin-LR across all three cell types, underscoring the need for further investigation and recognition of anabaenopeptins as cyanobacterial metabolites of concern. The data also revealed dose-dependent synergistic and antagonistic responses by microcystin and anabaenopeptin mixtures, with toxicity varying by metabolite ratios and cell type. Beyond these two cyanopeptides, which did not fully account for the observed variation in toxicity of *Microcystis* metabolite extracts, other cyanobacterial metabolites within the extract and their interactions likely contributed to the observed toxic effects on human cells and warrant closer investigation. These results underscore the need for expanded research on the mechanisms of entry and action for lesser studied and co-occurring cyanobacterial metabolites. Continued evaluation of these metabolites and their mixtures will clarify whether high-abundance, lesser-studied cyanopeptides pose risks that merit inclusion in routine monitoring and regulatory oversight.

This dissertation challenges the field to move beyond the conventional toxic/non-toxic paradigm centered on the presence of the *mcy* operon and its ability to encode the synthesis of microcystins. Many cyanopeptides such as the highly abundant anabaenopeptins, aeruginosins,

and aerucyclamides likely contribute to the strain-specific adaptive success of *Microcystis* and *Dolichospermum* under certain conditions and play critical ecological roles, such as allelopathy, grazing deterrence, oxidative stress protection, phycosphere recruitment, and nutrient storage and cycling (1–7). Chapters 2 and 3 generate testable hypotheses for these functions in two vital freshwater bodies. Future work should combine field studies and controlled culture experiments to further investigate how these molecules respond to environmental change.

Notably, non-microcystin cyanopeptides often occur at higher levels than microcystins and may exhibit greater toxicity toward aquatic organisms (8,9). This dissertation provides evidence that they mark functionally distinct cyanobacterial populations and contribute to bloom toxicity, strain heterogeneity, and multi-trophic impacts on ecosystems. Despite this, their potential human and environmental health effects remain largely unexamined due to a lack of standardized toxicological methods for assessing the effects of metabolite combinations and complex mixtures. Chapter 4 provides evidence and argues for developing such methods and shifting toward testing combinations of metabolites across relevant biological targets beyond the liver.

5.2 Looking Beyond Microcystins in Freshwater CyanoHABs: Conclusions, Opportunities, and Future Directions

5.2.1 Implications for Water Quality Monitoring in a Changing Climate

As of 2025, the Ohio Environmental Protection Agency (EPA) monitors four cyanotoxins in western Lake Erie (WLE): microcystins, cylindrospermopsins, saxitoxins, and anatoxins (**Figure 1.4**). While these are critical due to their potent toxicity and prevalence in waters used for drinking, recreation, and fishing, they represent only a fraction of the diverse and potentially harmful cyanobacterial metabolites produced throughout a cyanoHAB season. Chapters 2, 3, and

4 demonstrate that the dynamics and consequences of this chemical diversity remains understudied despite the frequent occurrence and co-occurrence of potentially toxic metabolites in cyanoHABs, increasing risks to human and aquatic health.

Due to the vast chemical diversity of cyanobacteria, identifying which metabolites to prioritize for monitoring and investigation remains challenging. A two-pronged approach to identify the most common and abundant metabolites with toxic or therapeutic activities could narrow down the list of metabolites for further investigation. First, nontargeted liquid chromatography-mass spectrometry (LC-MS) applied to culture and field-derived samples can reveal the most common and abundant individual and co-occurring metabolites (10–12). In parallel, samples could be investigated for their activity via bioactivity-guided fractionation to isolate metabolites with activities of interest (13–15). Used together, these approaches offer a promising framework for characterizing underexplored freshwater chemical space and identifying individual metabolites or mixtures of concern. Once novel metabolites are characterized, their role in culture- and field-based studies could be explored to expand our understanding of cyanoHAB chemical ecology beyond the paradigm of microcystins.

Not only is the cyanobacterial metabolome relatively unexplored, but it is also responding to environmental variability such as warming temperatures, increased nutrient loading, and more intense and frequent weather events (16–18). These changes can shift bloom composition which in turn influence bloom toxicity and metabolite profiles. For example, *Dolichospermum* abundance has increased in both WLE and Winam Gulf in recent years (Chapters 2 & 3), plausibly increasing the potential for other toxins to be produced in WLE such as saxitoxin (19). Rising temperatures and elevated nitrogen:phosphorus ratios have been linked to increased detection of the saxitoxin gene *sxtA* (19), while microcystin concentrations rise with

nitrogen and light availability (20–22). Some models even predict that reducing phosphorus alone, without reducing nitrogen, could yield more toxic blooms (23). The succession from *Microcystis* to *Dolichospermum* in these basins may also shift bloom chemistry, as nutrient-dependent co-culture studies suggest these genera can engage in allelopathic interactions that elevate microcystin, saxitoxin, and anatoxin production (4). However, most of this work focuses on well-characterized cyanotoxins, leaving the responses of other cyanopeptides, like anabaenopeptins, aeruginosins, and aerucyclamides, to environmental stressors largely unknown. Chapter 2 shows these metabolites increase after the peak of nitrate and microcystin abundance, suggesting dependence on regenerated nitrogen sources and late-season nutrient pulses, among many other co-variant environmental conditions later in the season. Understanding how climate and nutrient dynamics influence these metabolites is critical for anticipating future bloom toxicity and tracking functionally distinct populations of bloom-forming cyanobacteria.

Beyond cyanobacteria, cyanoHABs also contain a complex microbial community with broad implications for water quality (24–26). They are hotspots for antimicrobial resistance (AMR), fueled by antibiotic runoff from agriculture and wastewater reuse (27). AMR is a growing global threat, contributing to nearly 5 million deaths in 2019 alone (28). Lakes and rivers act as AMR reservoirs and transmission routes, and cyanoHABs may accelerate resistance via increased microbial biomass and mobile genetic elements. The co-occurrence of *Microcystis* and AMR-carrying bacteria, such as Proteobacteria, may create a positive feedback loop enhancing bloom persistence, toxicity, and resistance potential (29,30). Future monitoring and management must consider the broader microbiome of cyanoHABs, not just cyanobacteria, and the ways these communities interact under changing environmental conditions. As climate

change continues to reshape freshwater ecosystems, proactive metabolite monitoring and collaborative, capacity-building science will be vital for protecting global water security.

5.2.2 Insights into the Chemical Ecology and Functions of the Diverse Biosynthetic Arsenals of Bloom-Forming Cyanobacteria

The ecological functions of microcystins and other cyanobacterial secondary metabolites remain poorly understood. Hypothesized roles include allelopathy, grazing deterrence, oxidative stress mitigation, nutrient storage and metabolism, photosynthetic support, and microbiome recruitment and maintenance (31). Studies investigating the plausibility of each of these hypotheses have been conducted for microcystins (1,32–34), however, even the ecological role of this well-studied compound remains unknown. This suggests that focusing solely on microcystins may potentially obscure broader ecological patterns. Other commonly produced cyanopeptides, such as anabaenopeptins, aeruginosins, micropeptins, and cyclamides, may perform redundant, complementary, or competitive functions that could provide greater insights into the ecological role of microcystins and other cyanopeptides.

The ecological roles of individual cyanobacterial metabolites may be less informative for understanding strain-specific niche partitioning, which ultimately shapes bloom toxicity and chemical composition, than the collective effects of distinct suites of secondary metabolites produced by each strain. In Chapter 2, seven years of data revealed consistent temporal co-variance among certain BGCs and their associated metabolites, suggesting that specific strains and thus BGCs and metabolites co-occur more frequently than others. Studying these co-varying metabolite suites may clarify how entire packages of biosynthetic potential and metabolite synthesis confers ecological advantages that exceed those of single compounds. Integrating temporal patterns of BGC abundance, expression, and metabolite production with environmental

parameters offers a framework for identifying drivers of metabolite synthesis, forming hypotheses that can be experimentally tested under controlled conditions.

For example, Chapter 2 showed that microcystins peaked during periods of elevated nitrate and eDNA detection of grazers (mussels and *Daphnia*), while anabaenopeptins, aeruginosins, and aerucyclamides peaked later in the season during periods of high reduced organic nitrogen and eDNA detection of amoebas. These patterns suggest distinct ecological roles for strains that can produce different cyanopeptides. We hypothesize that microcystin-producing strains may be more resistant to grazing by *Daphnia* and mussels, but primarily under conditions of high inorganic nitrogen. In contrast, strains producing anabaenopeptins, aeruginosins, and aerucyclamides may be less defended against grazing but potentially more capable of fostering symbiotic interactions with microbiome members to enhance access to reduced organic nitrogen. Experimental validation of these hypotheses would help confirm field observations and further elucidate the complex chemical ecology underlying cyanoHAB dynamics

5.2.3 Improved Understanding of Global CyanoHABs through Collaboration

As of June 2025, only six metagenome-based studies have been conducted on the microbial communities of Lake Victoria, Kenya, half of which stem from the work presented in Chapter 3, compared to over 25 on Lake Erie, USA. Notably, Lake Erie has five times more NCBI BioProjects than Lake Victoria, underscoring a stark disparity in publicly available genomic data. Our collaborative project between 2022 and 2023 generated the first known metagenomic sequencing data from Winam Gulf, Lake Victoria, an area with year-round intense cyanoHABs effecting dense human and wildlife populations. This work was supported by the NSF International Research Experiences for Students (IRES) program (Award #1953468), which

supported North American and African students to jointly investigate biological, chemical, limnological, and socioeconomic aspects of cyanoHABs and water quality in the region. Close collaboration with local experts, students, and faculty ensured sampling captured the spatial and temporal complexity of cyanoHABs and addressed regionally relevant questions. Metagenomic analyses provided key insights into the bloom ecology of the gulf, including identification of *Microcystis* as the primary microcystin producer, detection of other potentially toxigenic cyanobacteria like *Dolichospermum* and *Sphaerospermopsis*, and the influence of spatially variable nitrogen: phosphorus ratios on bloom dominance. Importantly, toxigenic *Microcystis* was detected in highly turbid waters frequently used by lakeside communities for daily tasks, identifying serious but often unrecognized exposure risks due to a lack of green, visible scums. All resulting data were made publicly available and directly shared with Kenyan collaborators to support continued research and management.

This work was presented at the February 2024 African Center for Aquatic Research and Education (ACARE) meeting in Kisumu, Kenya, where a North American–African Great Lakes Metagenomics Working Group was established. The group, comprising over 30 researchers (65% faculty/staff, 35% students), aims to foster collaborative, genomics-driven research across the African and North American Great Lakes through shared datasets, tool tutorials, and joint deliverables such as datasets, reports and peer-reviewed manuscripts. Members represent Lakes Victoria, Tanganyika, Albert, Kivu, Edward, Erie, and Superior. Through this initiative and the NSF-IRES program, students and faculty have exchanged tools and tutorials (e.g., KBase), and conceptualized collaborative grants and manuscripts, building capacity and helping to close the genomic data gap in the Global South. These models of collaboration can be extended to

underrepresented freshwater systems in Africa, South America, Pacific Islands, and Southeast Asia, helping advance more equitable and informed global lake management.

5.2.4 Toxicity beyond Microcystins: Assessing, Monitoring, and Managing the Combinatorial Effects of Co-Occurring Cyanobacterial Metabolites

Cyanobacteria produce a complex mixture of metabolites during cyanoHABs, many of which remain uncharacterized, unmonitored, and unmanaged, threatening human and environmental health. While individual cyanotoxins such as microcystins, cylindrospermopsins, and anatoxins are well-studied, growing evidence shows that mixtures of these compounds can act synergistically in human and murine models (35–37). Chapter 4 of this thesis demonstrated potentially synergistic effects between microcystins and anabaenopeptins in human lung, liver, and kidney cell lines, underscoring the need to study both the individual toxic mechanisms of anabaenopeptins and its' interactions with other co-occurring metabolites, an increasingly common scenario in natural blooms (8,38,39). Experiments using crude extracts from *Microcystis* cultures have repeatedly shown that microcystin concentrations alone do not account for observed toxicity (40–42), a finding echoed in Chapter 4. This thesis contributes to a growing body of evidence that cyanoHAB toxicity cannot be fully understood through a microcystin-centric lens. Interactions among known and unknown metabolites within the cyanobacterial metabolome likely contribute to the observed toxic effects.

To address this complexity, coordinated long-term monitoring and toxicity testing efforts are needed. Integrating acute toxicity testing with multi-omics sampling (e.g., metagenomics, metatranscriptomics, metabolomics) on a standardized time course will support the development of interannual toxicity datasets across vital freshwater bodies. Such efforts require standardization of field extract testing, including defined cell lines, passage numbers, dosing

protocols, and specifying cytotoxicity and pro-inflammatory assays. Establishing this standardized framework will enhance our ability to identify metabolites of concern, support informed regulatory decisions, and ultimately better protect lakeside communities. Collaborative, standardized, and interdisciplinary approaches are essential to fully capture and mitigate the evolving risks of cyanoHAB toxicity.

5.3 Continuing Advancements for Future Applications

Advancing our understanding of cyanobacterial biosynthetic potential and its spatiotemporal dynamics requires two major priorities: obtaining more high-quality genomes and elucidating the synthesis products of cryptic biosynthetic gene clusters (BGCs). As of June 2025, there are only nine *Microcystis* and three *Dolichospermum* genomes in the in Joint Genome Institute Integrated Microbial Genomes (JGI IMG/MER) database labeled as "Finished," reflecting the highest sequencing quality (43). These genomes were largely generated using long-read sequencing platforms such as PacBio. Continued improvements in long-read sequencing technologies such as PacBio and Oxford Nanopore will be essential for producing high-quality and complete genomes that capture large BGCs, often exceeding 50 kilobases. There are currently 146 characterized cyanobacterial BGCs in the MIBiG database, compared to over 500 for Ascomycota and more than 1,000 for Actinomycetota (44). Metabologenomics studies and data within this thesis (Chapters 2 and 3) confirm the presence and high abundance of many cryptic BGCs in both cyanobacterial cultures and environmental cyanoHAB samples (25,45–47), underscoring the need for further elucidation of the synthesis products of these BGCs. As bioinformatics tools such as antiSMASH, GECCO, DeepBGC, and SanntiS (48–50,50–52) and synthetic biology approaches for BGC characterization continue to evolve, the discovery and annotation of new BGCs from cyanobacterial will accelerate.

Beyond biosynthetic potential, advances in non-targeted metabolomics workflows and databases, especially for environmental samples, are enabling more confident compound identification through secondary fragmentation (48,53). High-resolution liquid chromatography-mass spectrometry (HRLC-MS), paired with tandem mass spectrometry (MS/MS), provides detailed snapshots of metabolite profiles at specific points in time (54). One of the most powerful workflows for analyzing non-targeted metabolomics data is feature-based molecular networking via the Global Natural Products Social Molecular Networking (GNPS) platform (55). This is supported by the MassIVE database, which contains over 1,800 public datasets and 1.2 billion MS/MS spectra (56). Researchers can use this platform to annotate compounds in their own datasets, compare unidentified features to others globally, and track molecular signatures across studies. These tools have rapidly advanced the field and enabled much of the work in this thesis such as tracking the temporal patterns cyanopeptides in WLE in Chapter 2 while also opening new opportunities for novel molecule discovery which is a vast opportunity as suggested by the cryptic biosynthetic potential revealed in Winam Gulf in Chapter 3. Continued use of non-targeted metabolomics and the curation of new scripts and tools such as MassQL (57) to automatically detect and annotate cyanopeptides beyond microcystins, especially those with characteristic diagnostic ions (12), will expand the available data on lesser-studied compounds and enable more researchers to explore them in their own samples. In parallel, further elucidation of cyanopeptide congeners and other cyanobacterial secondary metabolites and deposition of annotated spectra into public databases will accelerate efforts to uncover the vast and largely untapped chemical space of cyanobacteria.

Ultimately, long-term, spatially and seasonally resolved time series of multi-omics data—paired with detailed physicochemical metadata—can enable predictive modeling of taxa, gene

expression, and metabolite production across bloom events. These models require extensive datasets and consistent sampling across multiple years to generate accurate, meaningful prediction. This represents a critical frontier for the field, where we have the ability not only to describe bloom dynamics, but to predict what is happening now, what triggered it, and what may happen next. Current cyanoHAB forecasting models primarily rely on satellite imagery, hydrodynamic data, microcystin levels, and chlorophyll concentrations to predict bloom movement and severity through hindcasts and various computer models (58–60). By integrating biological insights from multi-omics datasets and agent-based models, we will be able to model the in-situ behavior of cyanobacterial communities in real time, transforming bloom forecasting into a tool for proactive public health protection. Harnessing advances in computing, machine learning, and systems ecology, we can move toward more holistic and predictive management of cyanoHABs.

Together, improvements in sequencing technology, bioinformatic pipelines, metabolomics infrastructure, and data sharing platforms are driving the next generation of multi-omics integration. These advances will enable more holistic, accessible studies of cyanobacterial chemical ecology in aquatic ecosystems.

5.4 Concluding Remarks

This dissertation evaluated the biosynthetic potential, chemical diversity, and ecological consequences of cyanobacterial secondary metabolism in two geographically distinct eutrophic Great Lakes basins, western Lake Erie and the Winam Gulf. The findings make clear that bloom-forming cyanobacteria proliferating in freshwater cyanoHABs encode the ability to produce a wide array of bioactive metabolites beyond microcystins, most of which remain uncharacterized. This understudied chemical diversity carries implications for both human and ecosystem health

that remain largely unexamined. To fully grasp cyanoHAB toxicity and its drivers, future work must move beyond the narrow framework of *mcy*-encoding versus non-*mcy* encoding organisms and subsequent microcystin detection. A more expansive view, one that considers the full suite of secondary metabolites produced by different strains and genera, may reveal the ecological roles of these compounds and uncover new strategies to mitigate bloom proliferation and toxicity. For the protection of our most vital natural resource, I encourage researchers to use this work as a stepping stone toward a more holistic understanding of cyanoHAB chemical complexity, its spatiotemporal dynamics, and the evolving threats it poses to both human and environmental well-being.

5.5 References

1. Große R, Heuser M, Teikari JE, Ramakrishnan DK, Abdelfattah A, Dittmann E. Microcystin shapes the *Microcystis* phycosphere through community filtering and by influencing cross-feeding interactions. ISME Commun. 2024 Dec 24;5(1):170.

2. Barchewitz T, Guljamow A, Meissner S, Timm S, Henneberg M, Baumann O, et al. Non-canonical localization of RubisCO under high-light conditions in the toxic cyanobacterium *Microcystis aeruginosa* PCC7806. Environ Microbiol. 2019 Dec;21(12):4836–51.

3. Harke MJ, Davis TW, Watson SB, Gobler CJ. Nutrient-controlled niche differentiation of western Lake Erie cyanobacterial populations revealed via metatranscriptomic surveys. Environmental Science & Technology. 2016 Jan 19;50(2):604–15.

4. Chia MA, Jankowiak JG, Kramer BJ, Goleski JA, Huang IS, Zimba PV, et al. Succession and toxicity of *Microcystis* and *Anabaena* (*Dolichospermum*) blooms are controlled by nutrient-dependent allelopathic interactions. Harmful Algae. 2018 Apr 1;74:67–77.

5. Wang L, Zi J, Xu R, Hilt S, Hou X, Chang X. Allelopathic effects of *Microcystis aeruginosa* on green algae and a diatom: Evidence from exudates addition and co-culturing. Harmful Algae. 2017 Jan 1;61:56–62.

6. Freitas EC, Pinheiro C, Rocha O, Loureiro S. Can mixtures of cyanotoxins represent a risk to the zooplankton? The case study of *Daphnia magna Straus* exposed to hepatotoxic and neurotoxic cyanobacterial extracts. Harmful Algae. 2014 Jan;31:143–52.

7. Bojadzija Savic G, Colinet H, Bormans M, Edwards C, Lawton LA, Briand E, et al. Cell free *Microcystis aeruginosa* spent medium affects *Daphnia magna* survival and stress response. Toxicon. 2021 May;195:37–47.

8. Zastepa A, Westrick JA, Miller TR, Liang A, Szlag DC, Chaffin JD. The Lake Erie Harmful Algal Blooms Grab: High-resolution mapping of toxic and bioactive metabolites (cyanotoxins/cyanopeptides) in cyanobacterial harmful algal blooms within the western basin. Aquatic Ecosystem Health & Management. 2024 Jan;27(1).

9. Lenz KA, Miller TR, Ma H. Anabaenopeptins and cyanopeptolins induce systemic toxicity effects in a model organism the nematode *Caenorhabditis elegans*. Chemosphere. 2019 Jan 1;214:60–9.

10. Kust A, Řeháková K, Vrba J, Maicher V, Mareš J, Hrouzek P, et al. Insight into unprecedented diversity of cyanopeptides in eutrophic ponds using an MS/MS networking approach. Toxins (Basel). 20200831st ed. 2020 Aug 31;12(9).

11. Le Manach S, Duval C, Marie A, Djediat C, Catherine A, Edery M, et al. Global metabolomic characterizations of *Microcystis* spp. highlights clonal diversity in natural bloom-forming populations and expands metabolite structural diversity. Front Microbiol. 20190416th ed. 2019;10:791.

12. McDonald K, DesRochers N, Renaud JB, Sumarah MW, McMullin DR. Metabolomics reveals strain-specific cyanopeptide profiles and their production dynamics in *Microcystis aeruginosa* and *M. flos-aquae*. Toxins. 2023 Mar 31;15(4):254.

13. Weiss MB, Médice RV, Jacinavicius FR, Pinto E, Crnkovic CM. Metabolomics applied to cyanobacterial toxins and natural products. In: Pacheco Fill T, editor. Microbial Natural Products Chemistry [Internet]. Cham: Springer International Publishing; 2023 [cited 2025 Jun 18]. p. 21–49. (Advances in Experimental Medicine and Biology; vol. 1439).

14. Falch B, König G, Wright A, Sticher O, Angerhofer C, Pezzuto J, et al. Biological activities of cyanobacteria: Evaluation of extracts and pure compounds. Planta Med. 1995 Aug;61(04):321–8.

15. Médice RV, Arruda RS, Yoon J, Borges RM, Noyma NP, Lurling M, et al. Unlocking biological activity and metabolomics insights: primary screening of cyanobacterial biomass from a tropical reservoir. In Review; 2024 Mar [cited 2024 Mar 12].

16. Paerl H. Mitigating harmful cyanobacterial blooms in a human- and climatically-impacted world. Life. 2014 Dec 15;4(4):988–1012.

17. Chapra SC, Boehlert B, Fant C, Bierman VJ Jr, Henderson J, Mills D, et al. Climate change impacts on harmful algal blooms in U.S. freshwaters: A screening-level assessment. Environ Sci Technol. 20170725th ed. 2017 Aug 15;51(16):8933–43.

18. Liu X, Zhang J, Wu Y, Yu Y, Sun J, Mao D, et al. Intensified effect of nitrogen forms on dominant phytoplankton species succession by climate change. Water Research. 2024 Aug 3;122214.

19. Den Uyl PA, Kiledal EA, Errera RM, Chaganti SR, Godwin CM, Raymond HA, et al. Genomic identification and characterization of saxitoxin producing cyanobacteria in western Lake Erie harmful algal blooms. Environ Sci Technol. 2025 Apr 10;

20. Chaffin JD, Davis TW, Smith DJ, Baer MM, Dick GJ. Interactions between nitrogen form, loading rate, and light intensity on *Microcystis* and *Planktothrix* growth and microcystin production. Harmful Algae. 2018 Mar 1;73:84–97.

21. Chaffin JD, Westrick JA, Reitz LA, Bridgeman TB. Microcystin congeners in Lake Erie follow the seasonal pattern of nitrogen availability. Harmful Algae. 2023 Aug 1;127:102466.

22. Davis TW, Harke MJ, Marcoval MA, Goleski J, Orano-Dawson C, Berry DL, et al. Effects of nitrogenous compounds and phosphorus on the growth of toxic and non-toxic strains of *Microcystis* during cyanobacterial blooms. Aquatic Microbial Ecology. 2010;61(2):149–62.

23. Hellweger FL, Martin RM, Eigemann F, Smith DJ, Dick GJ, Wilhelm SW. Models predict planned phosphorus load reduction will make Lake Erie more toxic. Science. 20220526th ed. 2022 May 27;376(6596):1001–5.

24. Delmont TO, Pierella Karlusich JJ, Veseli I, Fuessel J, Eren AM, Foster RA, et al. Heterotrophic bacterial diazotrophs are more abundant than their cyanobacterial counterparts in metagenomes covering most of the sunlit ocean. The ISME Journal. 2022 Apr 1;16(4):927–36.

25. Yancey CE, Kiledal EA, Chaganti SR, Denef VJ, Errera RM, Evans JT, et al. The Western Lake Erie culture collection: A promising resource for evaluating the physiological and genetic diversity of *Microcystis* and its associated microbiome. Harmful Algae. 2023 Jul;126:102440.

26. Jankowiak JG, Gobler CJ. The Composition and function of microbiomes within *Microcystis* colonies are significantly different than native bacterial assemblages in two North American lakes. Front Microbiol. 20200528th ed. 2020;11:1016.

27. Amarasiri M, Sano D, Suzuki S. Understanding human health risks caused by antibiotic resistant bacteria (ARB) and antibiotic resistance genes (ARG) in water environments: Current knowledge and questions to be answered. Critical Reviews in Environmental Science and Technology. 2020 Oct 1;50(19):2016–59.

28. Naghavi M, Vollset SE, Ikuta KS, Swetschinski LR, Gray AP, Wool EE, et al. Global burden of bacterial antimicrobial resistance 1990–2021: a systematic analysis with forecasts to 2050. The Lancet. 2024 Sep;404(10459):1199–226.

29. Kim MJ, Kang D, Lee G, Kim K, Kim J, Shin JH, et al. Interplays between cyanobacterial blooms and antibiotic resistance genes. Environ Int. 2023 Oct 18;181:108268.

30. Ji W, Ma J, Zheng Z, Al-Herrawy AZ, Xie B, Wu D. Algae blooms with resistance in fresh water: Potential interplay between *Microcystis* and antibiotic resistance genes. Science of The Total Environment. 2024 May 25;173528.

31. Wei N, Hu C, Dittmann E, Song L, Gan N. The biological functions of microcystins. Water Research. 2024 Jul 18;122119.

32. Zilliges Y, Kehr JC, Meissner S, Ishida K, Mikkat S, Hagemann M, et al. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of *Microcystis* under oxidative stress conditions. PLoS One. 2011 Mar 18;6(3):e17615.

33. Vanderploeg HA, Liebig JR, Carmichael WW, Agy MA, Johengen TH, Fahnenstiel GL, et al. Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. Can J Fish Aquat Sci. 2001 Jun;58(6):1208–21.

34. Smith DJ, Berry MA, Cory RM, Johengen TH, Kling GW, Davis TW, et al. Heterotrophic bacteria dominate catalase expression during *Microcystis* blooms. Applied and Environmental Microbiology. 2022 Jul 5;88(14):e02544-21.

35. Takser L, Benachour N, Husk B, Cabana H, Gris D. Cyanotoxins at low doses induce apoptosis and inflammatory effects in murine brain cells: Potential implications for neurodegenerative diseases. Toxicol Rep. 20160115th ed. 2016;3:180–9.

36. Martin RM, Stallrich J, Bereman MS. Mixture designs to investigate adverse effects upon co-exposure to environmental cyanotoxins. Toxicology. 20190425th ed. 2019 Jun 1;421:74–83.

37. Diez-Quijada L, Puerto M, Gutiérrez-Praena D, Turkina MV, Campos A, Vasconcelos V, et al. In vitro toxicity evaluation of cyanotoxins cylindrospermopsin and microcystin-lr on human kidney HEK293 cells. Toxins. 2022 Jul;14(7):429.

38. Earnshaw CD, McMullin DR. Cyanopeptolins and anabaenopeptins are the dominant cyanopeptides from *Planktothrix* strains collected in Canadian lakes. Toxins. 2024 Feb 17;16(2):110.

39. Gkelis S, Lanaras T, Sivonen K. The presence of microcystins and other cyanobacterial bioactive peptides in aquatic fauna collected from Greek freshwaters. Aquatic Toxicology. 2006 Jun 10;78(1):32–41.

40. Bittner M, Štern A, Smutná M, Hilscherová K, Žegura B. Cytotoxic and genotoxic effects of cyanobacterial and algal extracts—Microcystin and retinoic acid content. Toxins (Basel). 2021 Feb 2;13(2):107.

41. Guzmán EA, Peterson TA, Winder PL, Francis KT, McFarland M, Roberts JC, et al. An assessment of potential threats to human health from algae blooms in the Indian River Lagoon (USA) 2018–2021: Unique Patterns of Cytotoxicity Associated with Toxins. Toxins. 2023 Nov;15(11):664.

42. Hrouzek P, Kapuścik A, Vacek J, Voráčová K, Paichlová J, Kosina P, et al. Cytotoxicity evaluation of large cyanobacterial strain set using selected human and murine *in vitro* cell models. Ecotoxicology and Environmental Safety. 2016 Feb 1;124:177–85.

43. Chen IMA, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, et al. The IMG/M data management and analysis system v.7: content updates and new features. Nucleic Acids Research. 2023 Jan 6;51(D1):D723–32.

44. Zdouc MM, Blin K, Louwen NLL, Navarro J, Loureiro C, Bader CD, et al. MIBiG 4.0: advancing biosynthetic gene cluster curation through global collaboration. Nucleic Acids Research. 2025 Jan 6;53(D1):D678–90.

45. Yancey CE, Hart L, Hefferan S, Mohamed OG, Newmister SA, Tripathi A, et al. Metabologenomics reveals strain-level genetic and chemical diversity of *Microcystis* secondary metabolism. Van Der Hooft JJJ, editor. mSystems. 2024 Jul 23;9(7):e00334-24.

46. Yancey C E, Smith Derek J, Den Uyl Paul A, Mohamed Osama G, Yu F, Ruberg Steven A, et al. Metagenomic and metatranscriptomic insights into population diversity of microcystis blooms: Spatial and temporal dynamics of *mcy* genotypes, including a partial operon that can be abundant and expressed. Applied and Environmental Microbiology. 2022;88(9):e02464-21.

47. Larsen JS, Pearson LA, Neilan BA. Genome mining and evolutionary analysis reveal diverse type III polyketide synthase pathways in Cyanobacteria. Genome Biol Evol. 2021 Mar 19;13(4):evab056.

48. Dinglasan JLN, Otani H, Doering DT, Udwary D, Mouncey NJ. Microbial secondary metabolites: Advancements to accelerate discovery towards application. Nat Rev Microbiol. 2025 Jun;23(6):338–54.

49. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Research. 2023 Jul 5;51(W1):W46–50.

50. Carroll LM, Larralde M, Fleck JS, Ponnudurai R, Milanese A, Cappio E, et al. Accurate de novo identification of biosynthetic gene clusters with GECCO. bioRxiv. 2021;2021.05.03.442509.

51. Sanchez S, Rogers JD, Rogers AB, Nassar M, McEntyre J, Welch M, et al. Expansion of novel biosynthetic gene clusters from diverse environments using SanntiS. 2023 [cited 2025 Jun 18].

52. Hannigan GD, Prihoda D, Palicka A, Soukup J, Klempir O, Rampula L, et al. A deep learning genome-mining strategy for biosynthetic gene cluster prediction. Nucleic Acids Res. 2019 Oct 10;47(18):e110.

53. Ye D, Li X, Shen J, Xia X. Microbial metabolomics: From novel technologies to diversified applications. TrAC Trends in Analytical Chemistry. 2022 Mar;148:116540.

54. Kind T, Tsugawa H, Cajka T, Ma Y, Lai Z, Mehta SS, et al. Identification of small molecules using accurate mass MS/MS search. Mass Spectrometry Reviews. 2018 Jul;37(4):513–32.

55. Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016 Aug 9;34(8):828–37.

56. Leao TF, Clark CM, Bauermeister A, Elijah EO, Gentry E, Husband M, et al. Quick-start for untargeted metabolomics analysis in GNPS. Nat Metab. 2021 Jul;3(7):880–2.

57. Jarmusch AK, Aron AT, Petras D, Phelan VV, Bittremieux W, Acharya DD, et al. A universal language for finding mass spectrometry data patterns. bioRxiv; 2022 [cited 2025 Jun 19]. p. 2022.08.06.503000.

58. Liu Q, Rowe MD, Anderson EJ, Stow CA, Stumpf RP, Johengen TH. Probabilistic forecast of microcystin toxin using satellite remote sensing, *in situ* observations and numerical modeling. Environmental Modelling & Software. 2020 Jun 1;128:104705.

59. Anderson EJ, Rowe M, Xu J, Zhang A, Lang G, Kelley JGW, et al. Using the Next-Generation Great Lakes Operational Forecasting System (GLOFS) to Predict Harmful Algal Bloom (HAB) Transport with the HAB Tracker. In 2017 [cited 2025 Jun 19]. p. 9B.1. Available from: https://ui.adsabs.harvard.edu/abs/2017AMS....9715763A

60. Liu Q, Rowe MD, Stumpf RP, Errera R, Godwin C, Chaffin JD, et al. Ten-Year Hindcast Assessment of an Improved Probabilistic Forecast System for Cyanotoxin (Microcystins) Risk Level in Lake Erie. Water Resources Research. 2025;61(4):e2024WR038952.

Appendix

Appendix A: Effects of Field-Derived CyanoHAB Extracts on Human Cell Viability

A.1 Introduction

This study was conducted in parallel with the cell viability assays described in Chapter 4, using the same methodology. However, in this case, field-derived cyanobacterial harmful algal bloom (cyanoHAB) metabolite mixtures collected from four stations in western Lake Erie (WLE) between August and October 2023 (**Figure A.6.1**) were tested on four human cell lines (A549, HK-2, Hep-G2, and HEK-293) to assess their effects on viability. Preliminary results revealed a dose-dependent response: minimal effects were observed at concentrations up to 100 μM, while some extracts caused viability reductions exceeding 50% at the highest concentration (500 μM) in two of the cell lines.

A.2 Results

Cyanotoxin Concentrations vary Spatiotemporally in WLE

The concentration of anabaenopeptin (ABN) and microcystin (MC) congeners were quantified in each field extract (**Figure A.6.2**). ABN-B and all MC congeners were detected at their highest levels in replicates from samples 1, 2, and 3. These samples were collected in August and September of 2023 from stations WE4, WE12, and WE9. ABN-A was only detected in samples 7 and 8, collected in October 2023 from stations WE2 and WE9. Sample 7 was the only sample to have replicates containing all measured cyanotoxins in the study, although often detected at very low concentrations.

Field-Derived Metabolite Mixtures Elicit Dose-Dependent Responses Across Four Human Cell Lines

The effects of field-derived cyanoHAB metabolite mixtures were evaluated across four human cell lines at varying concentrations (**Figure A.6.3, Figure A.6.4, Figure A.6.5**). At lower extract doses (≤100 μM), cell viability was largely unaffected in all cell lines. Among them, A549 cells showed the greatest sensitivity at these lower concentrations. At higher doses (>100 μM), Hep-G2 and HEK-293 cells exhibited pronounced decreases in viability (**Figure A6.5**). This trend was consistent across all extracts, though some mixtures caused stronger cytotoxic effects than others. Notably, replicates from sample 6 induced some of the most substantial reductions in cell viability—even at low doses—despite not having the highest cyanotoxin concentrations compared to other samples.

A.3 Discussion

Within cyanoHABs, hundreds of cyanobacterial derived metabolites can be present at any given time, many of which exhibit toxic effects towards humans, wildlife, domestic animals, and aquatic organisms (1–3). However, the combinatorial and interactive effects of such mixtures remain understudied, especially in western Lake Erie (WLE), a critical drinking water source for Toledo and its' surrounding municipalities. Most existing literature focuses on the effects of microcystins, largely overlooking the remaining 80% of characterized cyanobacterial metabolites, as well as numerous uncharacterized compounds (1). In this study, we tested the effects of field-derived cyanoHAB metabolite mixtures collected across a bloom season in WLE on the cell viability of human lung, liver, and kidney cells. Although some samples yielded less material and could not be tested at the highest concentration across all cell lines, we were still

able to observe distinct effects of extracts from different stations and bloom periods, as well as differential responses among cell lines exposed to the same extract.

We observed that some samples with relatively low cyanotoxin concentrations induced more pronounced toxic effects than samples containing higher concentrations of cyanotoxins. Several factors may explain this unexpected outcome: (1) the presence of other bioactive molecules in the low-cyanotoxin extracts that exert stronger cytotoxic effects on human cell lines, or (2) potential interactions within the high-cyanotoxin extracts that interfere with or suppress the activity of known cyanotoxins (MCs and ABNs), thereby reducing their overall impact on cell viability. These results suggest that complex drug–drug or drug–target interactions may be occurring within these crude mixtures, contributing to extract-specific cytotoxic profiles across different cell lines. To better understand these dynamics, future studies should focus on collecting larger volumes of material across a broader range of stations and bloom periods in WLE to ensure sufficient sample availability for IC_{50} determination across all cell lines. Additionally, fractionation of the crude extracts, followed by individual testing of the fractions and non-targeted metabolomic profiling, will be essential to identify the compounds most responsible for driving toxicity in each complex extract.

In addition to extract-specific effects on cell viability, the results suggest that some human cell lines were more strongly affected by the toxic effects of the extracts than others. This variability may be attributed to differences in expression of organic anion transporting polypeptides (OATPs), which are required for the cellular uptake of microcystins to exert their protein phosphatase inhibition activity (4). Notably, we observed more pronounced cytotoxic effects at low extract doses in the A549 lung cell line, which is known to express OATP1B3, the primary transporter for microcystin-LR (5). In contrast, the HK-2 kidney cell line, which showed

weaker responses to the extracts, expresses only OATP4C1, a transporter not yet confirmed to mediate uptake of microcystins or other cyanopeptides (6). HEK-293 cells do not express OATPs, while Hep-G2 cells are known to express OATP1B3 but at relatively low levels (7,8). Interestingly, greater toxic effects in HEK-293 cells compared to Hep-G2 were observed, suggesting that other bioactive molecules within the extracts, either transported by different carrier proteins or capable of passive diffusion, may be responsible for these effects. At higher extract doses, additional modes of entry and toxicity may be at play, including synergistic or antagonistic interactions between different compounds within the mixtures. Future studies should focus on elucidating the uptake mechanisms of other abundant cyanopeptides, such as anabaenopeptins, both in the presence and absence of microcystins, to better understand potential interactive toxic mechanisms in human cell models.

A.4 Materials and Methods

Cell Culturing and Targeted Metabolomics Analysis

Mammalian cell culturing for the cell lines A549 (lung) and HK-2 (kidney) and targeted metabolomics analysis of extracts were conducted as outlined in the methods section of Chapter 4 (Section 4.6) of this thesis.

Hep-G2 and Hek-293 Cell Culturing

Human hepatocellular carcinoma (liver epithelial) Hep-G2 (Virginia, USA, catalog no. HB-8065) and human immortalized embryonic kidney Hek-293 (Virginia, USA, catalog no. CRL-1573) cells were acquired from ATCC. Cells were cultured under standard conditions and stored in Gibco Recovery Cell Culture Freezing Medium at 2 x 10^6 cells per aliquot. Cells were

cultured in T175 flasks in Dulbecco's Modified Eagle Medium (DMEM, Life Technologies media, Carlsbad, CA, USA) supplemented with 10% Fetal Bovine Serum (FBS) and 1% penicillin/streptomycin. Cultures were maintained at 37°C in a humid atmosphere with 5% CO_2.

Field Sample Collection

Surface water samples (<1 meter depth) were collected using a 100 μm mesh plankton net tow. In some cases, depth-integrated samples were collected using a peristaltic pump from the surface to 1 meter above the lake bottom. All samples were collected between August and October 2023 from four National Oceanic and Atmospheric Administration Great Lakes Environmental Research Laboratory (NOAA GLERL) regulator monitoring stations: WE2, WE12, WE4, and WE9. At each station, four biological replicates were collected, resulting in four 1 L samples per site. Samples were stored in high-density polyethylene (HDPE) bottles at room temperature and returned to the lab, where they were left on the benchtop for four hours to allow biomass separation. Once biomass had settled or floated, it was pipetted into 50 mL Falcon tubes, centrifuged at 10,000 rpm for 15 minutes, and the supernatant was discarded. Pelleted biomass was stored at –20 °C until metabolite extraction.

Metabolite Extraction

Cell pellets were thawed and put into glass scintillation vials. Metabolites were extracted from the cellular biomass using a freshly made 50:50 high-performance liquid chromatography (HPLC)-grade methanol/chloroform mixture. The solvent mixture was added to the thawed scintillation vials to cover the biomass completely. Vials were then sonicated in a water bath for 30 minutes. Methanol rinsed stir bars were placed in each vial and samples were warmed at 40°C

and slowly stirred for four hours with caps lightly screwed on to allow for gas exchange. For each sample, two clean, pre-weighed vials, one for toxicity assays and one for targeted metabolomics analysis, were prepared. Extraction mixtures were filtered to remove residual solids and then evenly divided between the two vials. After filtration, vials were dried under a gentle stream of N_2 gas at 30°C. Dried samples were weighed again to calculate their crude extract weight and stored at -20°C until further analysis.

Cell Line Exposures to Field Extracts

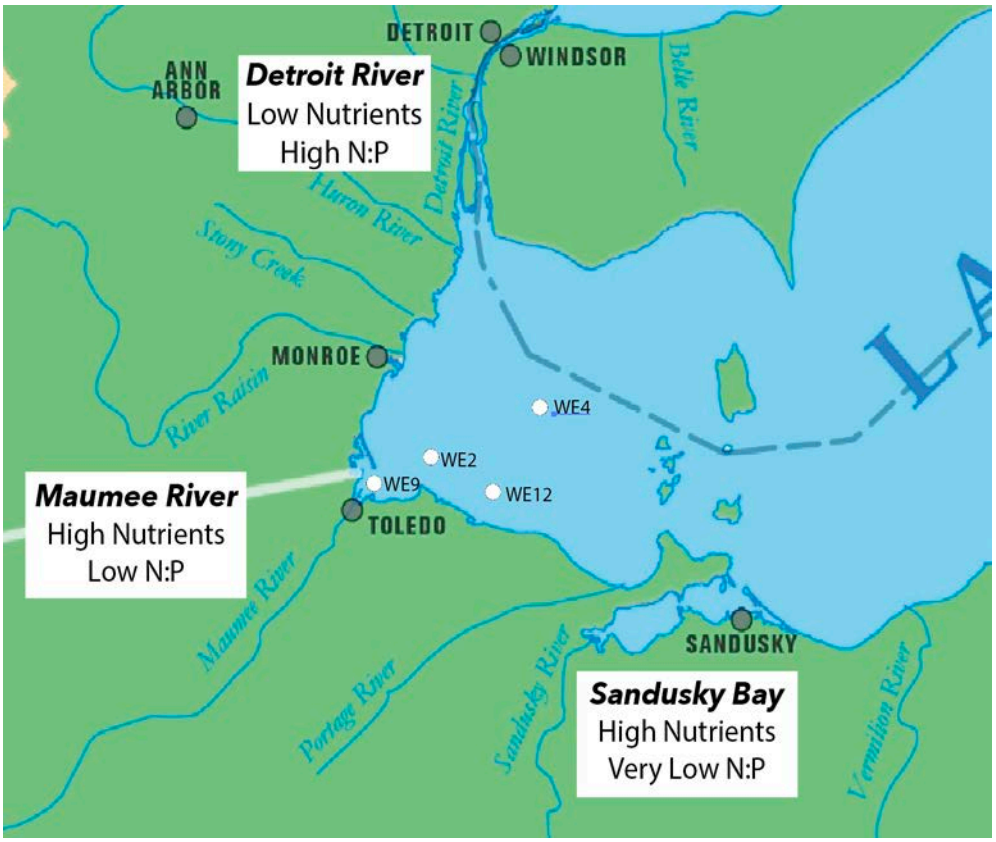
Extracts were reconstituted in 50/50 DMSO/water at a concentration of 100 mg/mL, based on the crude extract weight of the sample and stored at -20°C for the duration of experimental use. The viability and toxicity of all extracts were assessed in Hep-3B, A549, Hep-G2, and Hek-293 cells using the CellTiter-Glo assay (Promega, Madison, WI, USA). Cells were seeded in 20 µL of medium at a density of 10,000 cells per well in opaque, 96-well tissue culture-treated plates after reaching desired confluency and allowed to adhere for 24 hours. After adherence, cells were treated with serial dilutions between 1.25 and 500 µM of each extract, doxorubicin (positive control), and the highest stock concentration of DMSO/water to account for a negative vehicle control. Note that all extracts against each cell line were not tested at all doses or against all cell lines due to material limitation. Following treatment, 20 µL of fresh medium was added to each well and plates were incubated for an additional 48 hours. After incubation, 20 µL of CellTiter-Glo reagent was added to each well and incubated for five minutes at room temperature. Luminescence from each well was measured using a spectrophotometer. Results were normalized to a DMSO-vehicle control, and growth curves were visualized using GraphPad

Prism 10 (San Diego, CA, USA). Each extract was tested in duplicate on each plate, and each plate was repeated on different days to ensure technical and biological replication.

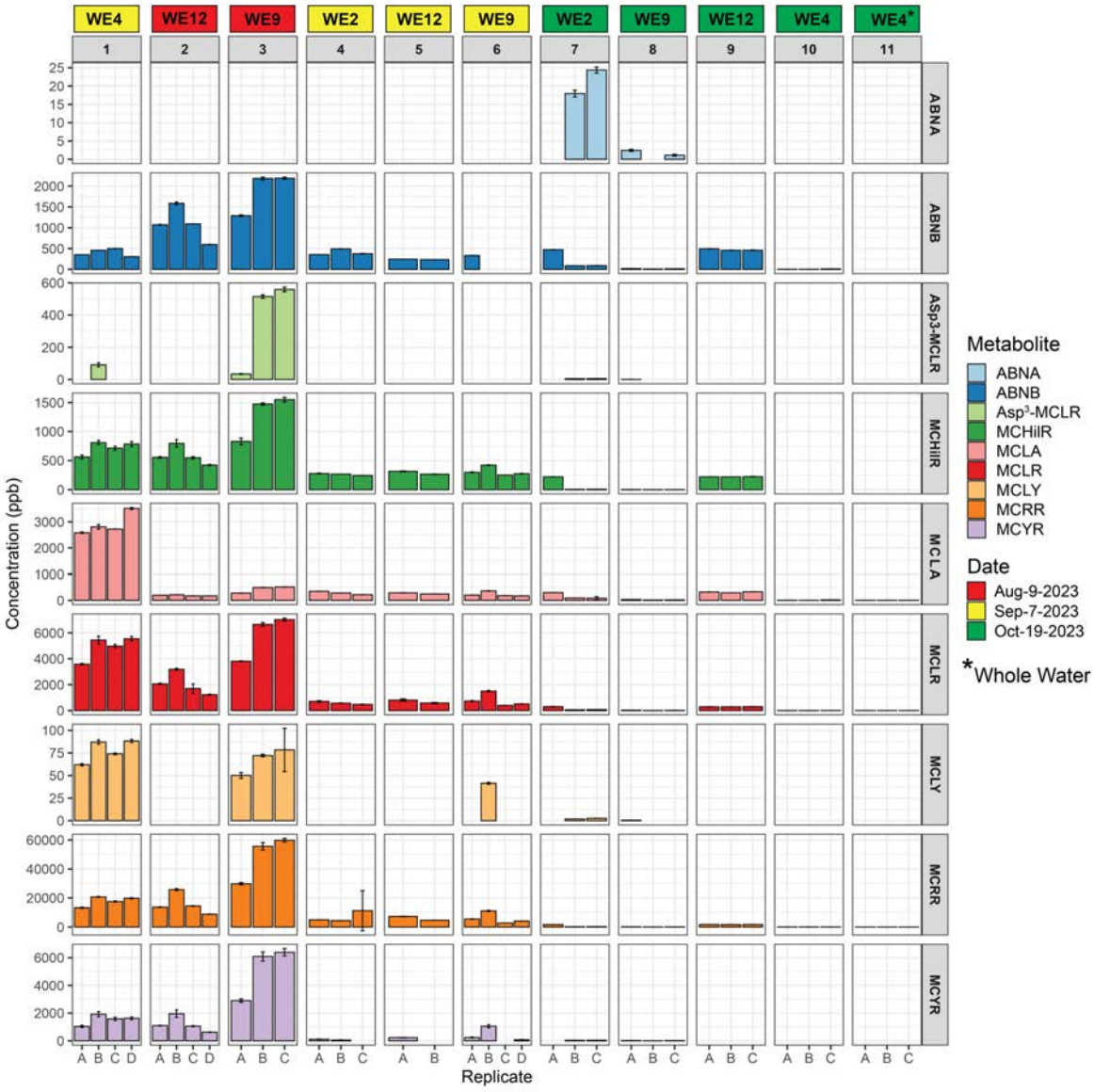
A.5 References

1. Janssen EML. Cyanobacterial peptides beyond microcystins – A review on co-occurrence, toxicity, and challenges for risk assessment. Water Research. 2019 Mar 15;151:488–99.
2. Yancey CE, Hart L, Hefferan S, Mohamed OG, Newmister SA, Tripathi A, et al. Metabologenomics reveals strain-level genetic and chemical diversity of Microcystis secondary metabolism. Van Der Hooft JJJ, editor. mSystems. 2024 Jul 23;9(7):e00334-24.
3. Kust A, Řeháková K, Vrba J, Maicher V, Mareš J, Hrouzek P, et al. Insight into Unprecedented Diversity of Cyanopeptides in Eutrophic Ponds Using an MS/MS Networking Approach. Toxins (Basel). 20200831st ed. 2020 Aug 31;12(9).
4. Breidenbach JD, French BW, Gordon TT, Kleinhenz AL, Khalaf FK, Willey JC, et al. Microcystin-LR aerosol induces inflammatory responses in healthy human primary airway epithelium. Environment International. 2022 Nov 1;169:107531.
5. Wang H, Xu K, Wang B, Liu J, Wang X, Xing M, et al. Microcystin-LR induces a wide variety of biochemical changes in the A549 human non-small cell lung cancer cell line: Roles for protein phosphatase 2A and its substrates. Environmental Toxicology. 2017;32(3):1065–78.
6. Jenkinson SE, Chung GW, van Loon E, Bakar NS, Dalzell AM, Brown CDA. The limitations of renal epithelial cell line HK-2 as a model of drug transporter expression and function in the proximal tubule. Pflugers Arch. 2012 Dec;464(6):601–11.
7. Fischer A, Hoeger SJ, Stemmer K, Feurstein DJ, Knobeloch D, Nussler A, et al. The role of organic anion transporting polypeptides (OATPs/SLCOs) in the toxicity of different microcystin congeners in vitro: A comparison of primary human hepatocytes and OATP-transfected HEK293 cells. Toxicology and Applied Pharmacology. 2010 May 15;245(1):9–20.
8. Tsujimoto M, Hatozaki D, Shima D, Yokota H, Furukubo T, Izumi S, et al. Influence of Serum in Hemodialysis Patients on the Expression of Intestinal and Hepatic Transporters for the Excretion of Pravastatin. Therapeutic Apheresis and Dialysis. 2012;16(6):580–7.

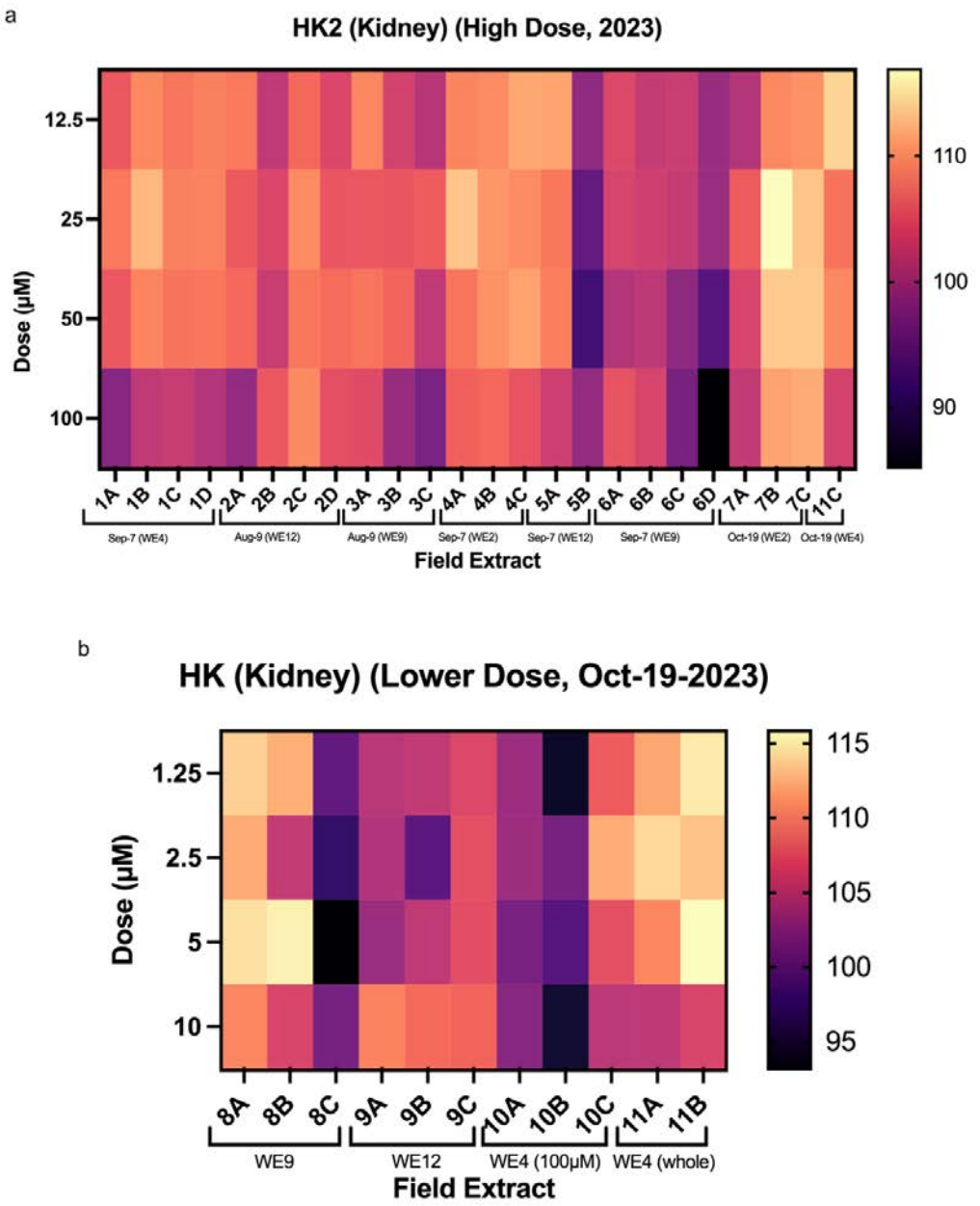
A.6 Figures



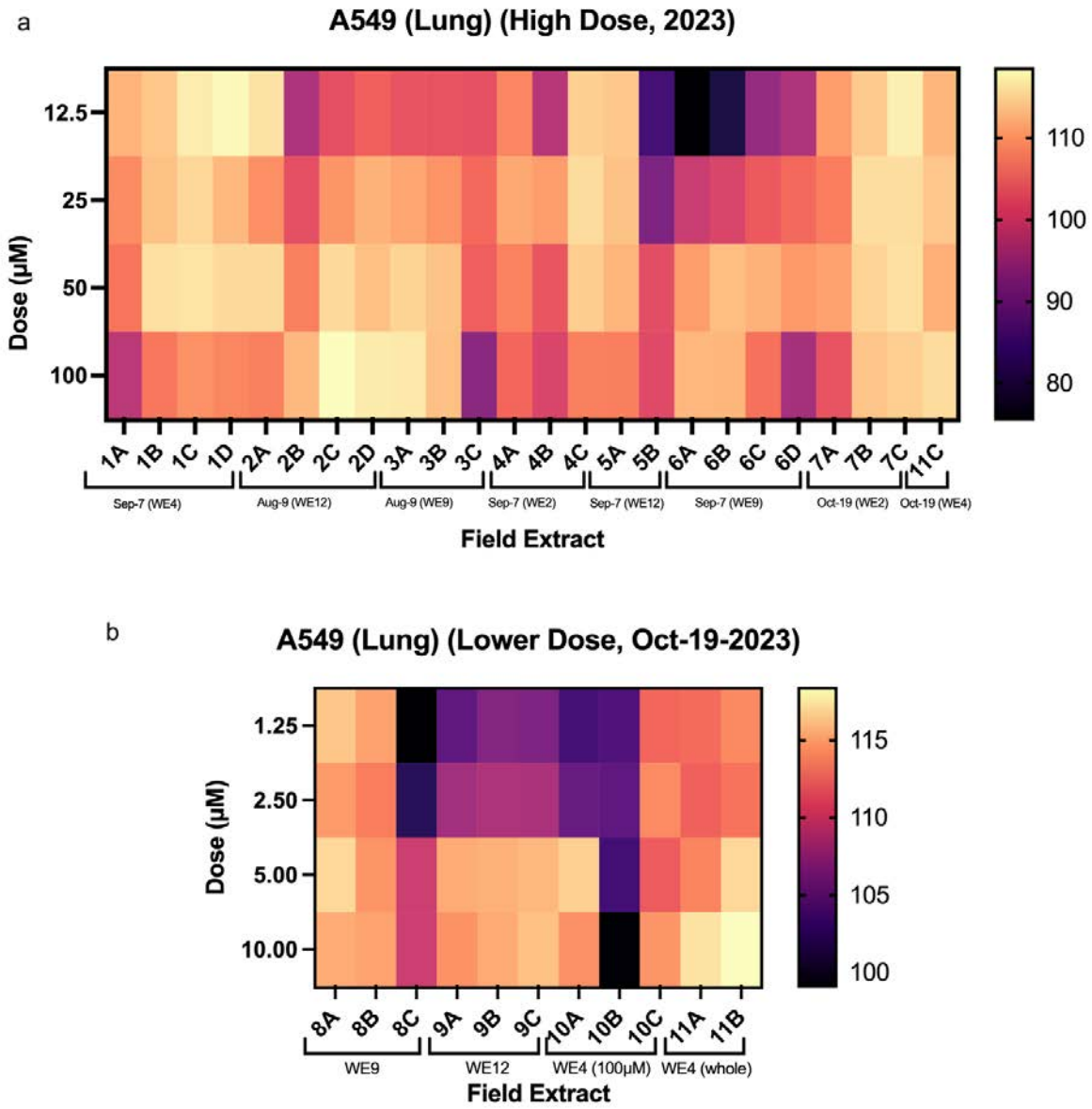
Appendix Figure A.6.1: Map of western Lake Erie showing NOAA long-term monitoring stations evaluated in this study. Near-shore sites WE2 and WE12 border the Ohio coastline, WE9 is located between the Maumee and Detroit River mouths, and WE4 represents an offshore station.



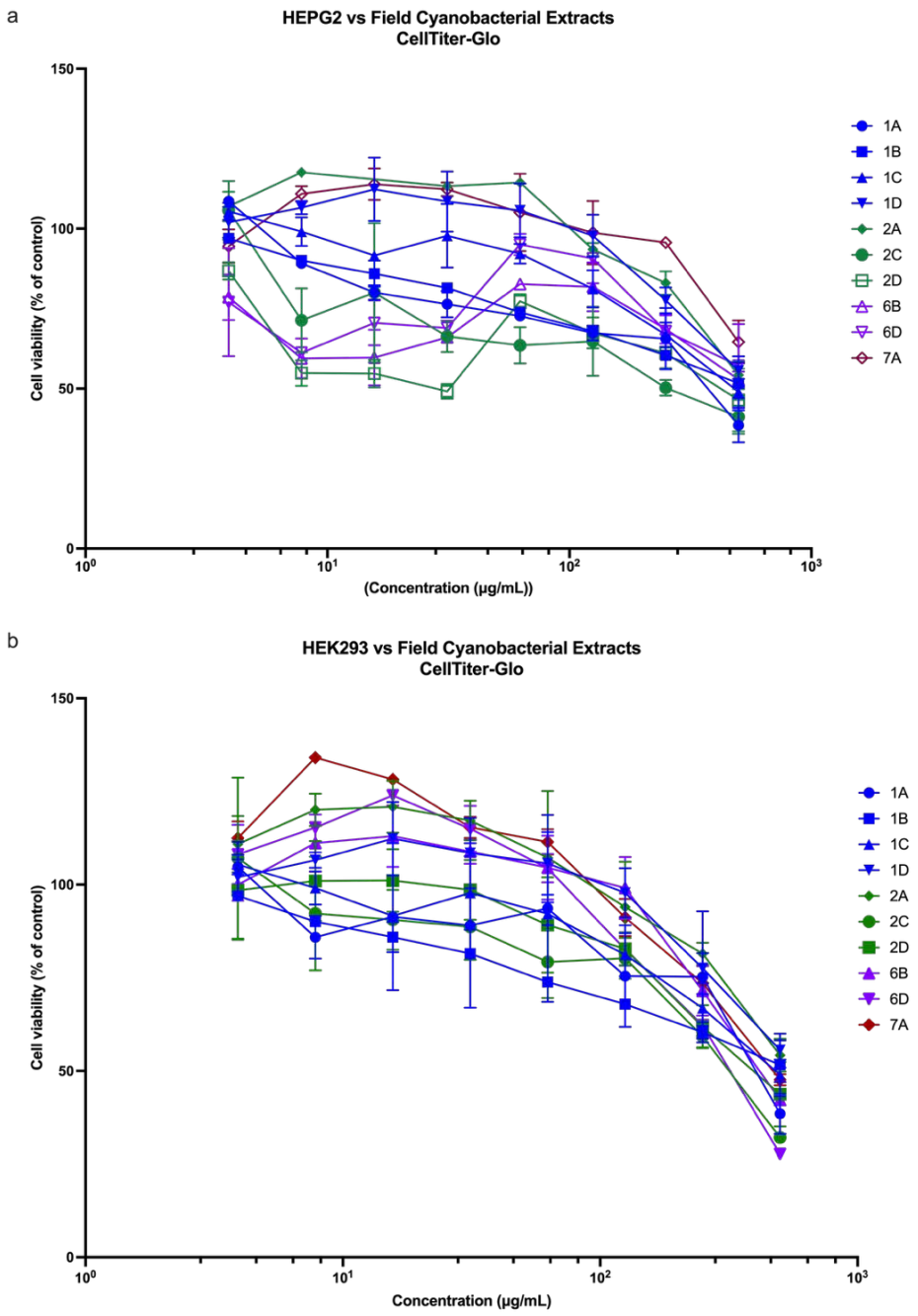
Appendix Figure A.6.2: Quantification of cyanotoxins via liquid chromatography-mass spectrometry (LC-MS), including anabaenopeptins (ABN-A and ABN-B) and microcystin (MC) congeners (Asp^3-LR, HilR, LA, LR, LY, RR, YR). Each vertical panel represents a different cyanotoxin, each with a unique y-axis scale. Biological replicates (samples collected from the same station and date) are shown along the x-axis as replicates and are grouped by collection date and station. All extracts were collected from the surface using a 100 µM plankton net tow, while samples with a * (sample 11, collected from WE4 on October 19th, 2023) were collected from a depth integrated sample and were not filtered using any mesh.



Appendix Figure A.6.3: (a) Cell viability response by HK-2 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 12.5 and 100 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control. **(b)** Cell viability response by HK-2 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 1.25 and 10 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control.



Appendix Figure A.6.4: (a) Cell viability response by A549 (Lung) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 12.5 and 100 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control. **(b)** Cell viability response by HK2 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 1.25 and 10 µM, relative to the crude weight of dried extract after metabolite extraction. Hotter tiles indicate more viability relative to the DMSO control, and darker tiles indicate less viability relative to the DMSO control.



Appendix Figure A.6.5: (a) Cell viability response by Hep-G2 (Liver) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 3.9 and 500 µM (shown on the x-axis), relative to the crude weight of dried extract after metabolite extraction. Cell viability response, shown as the percent viability relative to the DMSO control, is shown on the y-axis. Biological replicates share the same color but differ in shape. **(b)** Cell viability response by Hek-293 (Kidney) cells to cyanoHAB field extracts collected in 2023. Extracts were dosed between 3.9 and 500 µM (shown on the x-axis), relative to the crude weight of dried extract after metabolite extraction. Cell viability response, shown as the percent viability relative to the DMSO control, is shown on the y-axis. Biological replicates share the same color but differ in shape.