Title*: Expression of* Microcystis *biosynthetic gene clusters in natural populations suggests temporally dynamic synthesis of novel and known secondary metabolites in western Lake Erie*

Running Title: *Novel & known* Microcystis *biosynthetic genes*

Authors: Colleen E. Yancey[1], Fengan Yu[2], Ashootosh Tripathi[2,3,4], David H. Sherman[2,4], Gregory J. Dick[1,5,*]

(1) Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, 48109, USA

(2) Life Sciences Institute, University of Michigan, Ann Arbor, 48109, USA

(3) Natural Products Discovery Core, University of Michigan, Ann Arbor, 48109, USA

(4) Department of Medicinal Chemistry, University of Michigan, Ann Arbor, 48109, USA

(5) Cooperative Institute for Great Lakes Research (CIGLR), University of Michigan, 4840 South State Road, Ann Arbor, MI 48108 USA

***Corresponding Author:** Dr. Gregory J. Dick, Address: 1100 North University Ave, Rm. 2014, Ann Arbor, MI, USA, 48109 | phone: 1-734-763-3228, fax:1-734-763-4690 | gdick@umich.edu

Abstract

Microcystis spp. produces diverse secondary metabolites within freshwater cyanoHABs around the world. In addition to the biosynthetic gene clusters (BGCs) encoding known compounds, *Microcystis* genomes harbor numerous BGCs of unknown function, indicating a poorly understood chemical repertoire. While recent studies show that *Microcystis* produces several metabolites in the lab and field, little work has focused on analyzing the abundance and expression of its broader suite of BGCs during cyanoHAB events. Here, we use metagenomic and metatranscriptomic approaches to track the relative abundance of *Microcystis* BGCs and their transcripts throughout the 2014 western Lake Erie cyanoHAB. Results indicate the presence of several transcriptionally active BGCs that are predicted to synthesize both known and novel secondary metabolites. The abundance and expression of these BGCs shifted throughout the bloom, with transcript abundance levels correlating with temperature, nitrate, and phosphorus concentrations, and the abundance of co-occurring predatory and competitive eukaryotic microorganisms, suggesting the importance of both abiotic and biotic controls in regulating expression. This work highlights the need for understanding the chemical ecology and potential risks to human and environmental health posed by secondary metabolites that are produced but often unmonitored. It also indicates the prospects for identifying pharmaceutical-like molecules from cyanoHAB-derived BGCs.

Importance

Microcystis spp. dominate cyanobacterial harmful algal blooms (cyanoHABs) worldwide and pose significant threats to water quality through the production of secondary metabolites, many of which are toxic. While the toxicity and biochemistry of microcystins and several other compounds have been studied, the broader suite of secondary metabolites produced by

Microcystis remains poorly understood, leaving gaps in our understanding of their impacts on human and ecosystem health. We use community DNA and RNA sequences to track the diversity of genes encoding synthesis of secondary metabolites in natural *Microcystis* populations and assess patterns of transcription in western Lake Erie cyanoHABs. Our results reveal the presence of both known gene clusters that encode toxic secondary metabolites as well as novel ones that may encode cryptic compounds. This research highlights the need for targeted studies of the secondary metabolite diversity in western Lake Erie, a vital freshwater source to the United States and Canada.

Introduction

CyanoHABs, which are dense proliferations of cyanobacteria that can discolor water and produce toxins, occur annually and globally, and are expected to increase in severity with climate change (1–3). These blooms, which can persist through the early summer into late fall in temperate environments, produce a range of secondary metabolites that can be deleterious to ecosystem function and human health (4–6). CyanoHABs result largely from anthropogenic eutrophication via nutrient runoff (2, 7, 8) and are expected to become more toxic with increasing nutrient loading and continued, intensifying, climate change (2, 9, 10).

Microcystis is a non-N-fixing, potentially toxic cyanobacterium that often dominates freshwater cyanoHABs on every continent except Antarctica (5). Blooms made up of largely *Microcystis* are expected to expand and increase in severity over the next several years (2, 6). This expansion is a concern because *Microcystis* produces diverse bioactive secondary metabolites with an array of ecological and physiological functionalities (11–13). Many *Microcystis* secondary metabolites possess toxic or inhibitory properties toward diverse cell types (13–15) and/or antibiotic, antifungal, or cytotoxic properties that may be useful in

pharmaceutical discovery (16–18). Secondary metabolites likely provide *Microcystis* with ecological advantages such as grazer defense, allelopathy, quorum sensing, and/or protection against reactive oxygen species (ROS) and may influence cyanoHAB community composition (17–20). Despite their significance to ecology, human, and ecosystem health, the physiological and environmental controls on biosynthesis of *Microcystis* secondary metabolites remain poorly understood.

Microcystins, the most extensively studied family of secondary metabolites produced by *Microcystis* (21–23), are structurally related hepatotoxins that are responsible for several drinking crises (4, 24, 25) and livestock poisoning events (26) around the world. However, *Microcystis* genomes also contain additional BGCs (18, 20, 27) encoding secondary metabolites such as aeruginosins, anabaenopeptins, cyanobactins, cyanopeptolins, microginins, and microviridins (11, 13), which can be toxic and may occur in the environment and drinking water treatment plants at frequencies and concentrations that equal or exceed those of microcystins (28–30). In addition, many BGCs identified from *Microcystis* genomes do not have associated products, and thus have been designated "orphan" clusters (12, 31), highlighting the need to identify and determine the structural characteristics and biological properties of these unknown compounds. While *Microcystis* genomes harbor a diverse and variable suite of BGCs (19, 20) and some metabolites have been shown to be expressed and produced in culture (18), little is known about the occurrence or expression of these *Microcystis* BGCs in natural environments. Understanding their distribution and expression in relation to the biotic and abiotic environment may shed light on the ecological role of these molecules and inform risks to public health.

In this study we used metagenomics and metatranscriptomics to assess the abundance, diversity, and expression of BGCs in *Microcystis* populations in a time series at three stations

during the 2014 western Lake Erie bloom, which show a succession of *Microcystis* strains alongside changing nutrient availability, microcystin concentrations, and microbial communities (32, 33). Western Lake Erie is subject to annual cyanoHABs that have intensified in recent decades (6, 34), and the 2014 bloom caused a drinking water crisis in which Toledo residents lost access to potable water due to dangerously high levels of microcystins (4). While microcystins are heavily monitored and studied in these waters, monitoring of "other" secondary metabolites has been limited to saxitoxin, anatoxin, and cylindrospermopsin (35). Motivated by the known diversity of *Microcystis* BGCs and the influence of nutrient stoichiometry on biosynthesis of secondary metabolites (36, 37), we hypothesized that the relative abundance and transcription of BGCs would vary across the bloom season as nutrient availability changes.

Results

The 2014 Western Lake Erie cyanoHAB

The 2014 western Lake Erie cyanoHAB was notable due to high levels of microcystins at station WE12, the Toledo Drinking water crib, which led to a drinking water crisis event in Toledo, OH, USA. This study used samples and data collected and described previously (32, 33, 38, 39). Briefly, samples were collected from three core stations as part of greater sampling efforts performed by the NOAA Great Lakes Environmental Research Laboratory (GLERL) to monitor cyanoHAB development in the western basin of Lake Erie. WE2 and WE12 are close to the Ohio coast and Maumee River inlet and are considered nearshore stations. WE4 is located more centrally in the basin and is considered offshore (Fig 1A).

The 2014 bloom consisted of high levels of nitrate, particulate microcystin, and cyanobacteria biomass in August at the nearshore stations WE2 and WE12. A secondary peak in

cyanobacteria biomass, where microcystin concentrations were low, was observed in late September at WE12 and WE2. The trends at the nearshore stations suggest that high nitrate concentrations may support initial bloom development, proliferation, and/or production of microcystin, but that biomass and microcystin production decreases as nitrate is being depleted. At the offshore station, WE4, microcystin concentrations were considerably lower (Fig 1B). Soluble reactive phosphorus (SRP) and nitrate concentrations (NO_3^-) were also variable across stations and sampling times, with marked decreases in nitrate as the bloom persisted at stations WE2 and WE12 from 4 August, onwards. Notably, at all three sampling stations, the lowest measurement of nitrate occurred one sampling point before the lowest measurement of particulate microcystin (Fig 1B).

Microcystis MAG Generation

Ten *Microcystis* metagenome-assembled genomes (MAGs) were generated from fifteen metagenomic samples from the 2014 western Lake Erie cyanoHAB (Table 1). Samples that failed to generate a *Microcystis* MAG were likely due to low abundance of the organism as seen in early phases of the bloom (early July), or poor assembly of *Microcystis* contiguous sequences due to high strain heterogeneity (40). Most MAGs had high completion (above 90%) and variable redundancy given the high strain heterogeneity observed (41), except for the MAG generated from August 4th, at Station 12, at 40% completion (Table 1). This MAG was kept for further analysis as this sample was taken directly during the time of the Toledo Drinking Water Crisis. Five fragmented BGCs (<5kb) identified in this MAG were removed from further analysis due to incompletion, while one complete BGC (T1PKS) was identified and used for subsequent analysis. Fragmented BGCs were determined to be incomplete due to the fragmentation or contiguous sequence breaks present within gene open reading frames.

BGC Diversity and Abundance

Nineteen distinct BGCs, representing multiple classes of putative secondary metabolite synthesis, were identified (Table 2, Fig. 2). Of the nineteen BGCs, eleven clusters are thought to putatively encode known synthesis products, while eight are considered cryptic (Table 2). Within the cryptic clusters, PKS-M 1 and 2 appear to be fragmented or incomplete microginin encoding BGCs, four have been identified in previous work as "orphan clusters" (31), and two are novel, and to our knowledge, have not been previously discussed in the literature (Table 2).

The relative abundance of these BGCs, which was estimated by quantifying metagenomic reads mapped to assembled BGCs and normalized to reads mapped to *Microcystis* 16S rRNA genes, showed differences across stations and sampling dates (Fig 2). The relative abundance of the complete *mcy* operon, which was previously analyzed (33) is included for reference as well. Relative abundances of identified BGCs, which putatively produce aeruginosins, anabaenopeptin, cyclophane-like 1, cyanopeptolin 3, and microviridin B, and a cyclophane-like molecule (PKSmod-NRSlike-T3PKS), increased at station WE12 in late September through October, during later phases of the bloom. BGCs predicted to putatively encode aeruginosins, cyanobactins (58-75% similarity to piricyclamides encoding clusters), microviridins, and cyanopeptolins each yielded multiple distinct clusters with conserved gene content, but gene rearrangements, insertions, and duplications were observed within these clusters (Fig. S1). Gene order and orientation of BGCs encoding aeruginosin and microviridin were more conserved (Fig S1A & S1C), while rearrangements, insertions, and deletions were more common in the cyanobactin and cyanopeptolin BGCs (Fig S1B & S1D). Deeper annotations for clusters thought to encode anabaenopeptin, microviridin, aeruginosin, and a piricyclamide-like cyanobactin as well as chemical structures of potential congeners produced are shown in Fig. S2. Some clusters,

such as the PKSmod-NRPSlike-T3PKS 1, which may encode a cyclophane-like metabolite, were rare at most sampling times and stations, except on 20 October when it was one of the most abundant clusters. The cyanobactin clusters, which encode ribosomally synthesized cyanobacterial derived macrocyclic metabolites (42), shifted in relative abundance during the bloom with 2 and 3 being more abundant early, and 1 being the most abundant during late phases (Fig. 2).

Two gene clusters, PKS-M 1 and 2, are believed to be part of the microginin synthesis pathway, which was recently confirmed to be present in *Microcystis* (43); however, both clusters contain a fragmented hybrid NRPS-PKS encoding gene and lack the additional two NRPS encoding genes required for microginin synthesis (Fig 3A). When compared to the microginin pathway in *Microcystis aeruginosa* LEGE 91341 (43), PKS-M 1 is shown to contain additional synthesis genes thought to encode a fatty acyl-AMP ligase and acyl carrier protein, as well as a hypothetical protein. PKS-M 2, however, lacks these putatively identified genes, and contains a putative endonuclease (Fig. 3A). Both core biosynthesis genes present in both PKS-M clusters contain domains for ketosynthase, acyltransferase, PP-binding, and aminotransferase activity, but lack the condensation, AMP binding, and peptidyl-carrier protein domains observed in the complete clusters (Fig. 3A). To determine if these incomplete clusters were a result of natural pathway truncations or poor metagenomic assembly, mapping of western Lake Erie (WLE) metagenomic reads against the reference microginin BGC from LEGE 91341 was completed. Mapping revealed read coverage across the entire length of the BGC, suggesting the presence of this complete BGC in natural populations, although there are some notable variations in sequence similarity and regions of non-uniform coverage (Fig 3B). These gene clusters were most abundant at nearshore stations in later phases of the bloom (Fig. 2).

Three cryptic polyketide synthase (PKS) or hybrid nonribosomal peptide synthetase (NRPS)-PKS cluster classes were also identified and varied greatly in abundance (Fig. 2,4). These systems putatively encode unknown metabolites and have either very low percent similarities (≤40%; Table S1) or no similarity at all to previously described clusters in the MiBIG database, suggesting high potential for biosynthesis of unique structures. The cryptic type III polyketide synthase (T3PKS) identified reached peak relative abundance at the nearshore stations on 4 August (Fig. 2) and contains a putative naringenin-chalone synthase (Fig 4A). The cluster encoding a predicted iterative PKS was greatest in relative abundance on 20 October at nearshore stations and contains a PKS encoding a putative enediyne biosynthesis system (Fig. 2, 4A). MIC 1 was previously described in *Microcystis aeruginosa* NIES-843 (31), and putatively encodes two modular T1PKSs, an NRPS like enzyme, an additional PKS, and a halogenase (Fig. 4C). Deeper annotations for genes identified in the MIC 1 cluster are described in Table S2.

Protein Phylogenetics for Biosynthesis Genes

Phylogenetic trees of various predicted proteins from PKS-containing clusters were analyzed for potential insights into the corresponding biosynthetic pathways. The enediyne biosynthesis protein sequence (PKSE) from the PKS iterative cluster, which was initially identified in *Microcystis aeruginosa* strain PCC 7806 (31), was investigated more deeply as enediyne synthesis is a critical step in the formation of anticancer compounds such as calicheamicin (44). While there were no closely related homologous sequences from known biosynthesis pathways, the enediyne protein sequence had over 97% identity and 100% alignment to a putatively identified PKSE in *Microcystis aeruginosa* strain NIES-843 (45), with a bootstrap of 1 (Fig 5). Other putative PKSE enzymes identified from a variety of cyanobacteria (45), formed a subclade with the PKSE identified in W. Lake Erie MAGs (Fig 5). These protein sequences form a larger

clade with known PKSE enzyme sequences that generate a variety of enediyne-containing metabolites with pharmaceutical relevance (bootstrap=0.81, Fig. 5).

Protein sequences from the T3PKS gene cluster (Fig. 3B), also provided clues into potential functions of biosynthesis enzymes. An exporter protein sequence shared high sequence similarity to proteins involved in siderophore-like export from other *Microcystis* strains (~95-97% identity, bootstrap = 1, Fig S3A). Distantly related clades contained protein sequences with annotations for export of siderophores and/or cyclic peptides from a variety of cyanobacteria including benthic *Leptolyngbya* sp. (bootstrap = 0.894, ~63% identity, Fig S3A). The core synthesis enzyme within this cluster, annotated by antiSMASH as naringenin-chalcone synthase, was most closely related to other *Microcystis* naringenin-chalcone synthase and T3PKS enzymes (~98% identity, bootstrap = 0.877, Fig. S3B) as well as naringenin-chalcone synthases from *Chamaesiphon* sp. from the cyanobacteria Order *Synechococcales* (bootstrap = 0.688, 57-62% identity, Fig. S3B), none of which are linked to a known synthesized compound. Its protein sequence was also similar to an experimentally confirmed α-pyrone synthesis polyketide synthase (PKS18) (Fig. S3B) from the actinobacterium *Nocardia seriolae* (50.6% identity, 97% alignment length).

Transcriptional activity of BGCs

Quantification of relative abundances of metatranscriptomic sequences, which were available for seven of the 15 samples analyzed, revealed that BGC clusters were differentially expressed across stations and sampling times (Fig. 6). During early phases of the bloom, there was little expression of BGCs as observed on 21 July at WE2 and 29 July at WE4, likely due to low biomass of *Microcystis*. At WE12 during the peak bloom phase (4 August), clusters T3PKS, PKS-M 1, and *mcn* 2 increased in relative transcript abundance compared to other clusters (Fig.

6). While the relative abundance of transcripts for *mcy* genes encoding microcystin (33) was highest at WE12 during August, other BGCs had similar, and in some cases greater, relative transcript abundances than the microcystin encoding gene cluster during other phases of the bloom (Fig. 6). The relative abundance of BGC transcripts shifted and greatly increased, especially in comparison to the *mcy* operon, at all three stations during middle and late phases of the bloom, with high relative abundance of transcripts for T3PKS and MIC1 at WE12 on 25 August, WE2 on 6 October, and 8 September at WE4. On 23 September at WE12, the PKSmod-NRPS-T3PKS 1 gene cluster had the highest relative transcript abundance compared to other gene clusters. There was also a marked increase in relative transcript abundance of *mcn* 2 during later phases of the bloom at both WE2 and WE4. Greater transcriptional abundance at WE2 was also observed for clusters PKSmod-NRPSlike-T3PKS 1 and *aer* 2 on 6 October as well. (Fig. 6).

Correlation Analysis between transcript abundance and both abiotic and biotic variables

To explore potential controls on the expression of BGCs, we performed correlational analysis of BGC transcript abundance with a variety of available abiotic conditions and relative abundances of known *Microcystis* predators and competitors in metagenomic data. Relative abundances of BGC transcripts were significantly correlated with both abiotic conditions (Table 3) and relative abundance of predatory and competitive organisms (Table 4). The transcript abundance of all BGC classes examined was significantly and negatively correlated with nitrate concentration while the transcript abundances of genes for cyanobactins (Pearson's R=-0.531, p=0.029), PKSmod-NRPSlike-T3PKS (Pearson's R=0.598, p=0.040), and *mcn* (cyanopeptolin) (Pearson's R=-0.412, p=0.089) BGCs were negatively correlated to temperature. Transcript abundances for *aer*, *mcn*, *mdn*, and PKSmod-NRPSlike-T3PKS clusters had positive and

252 significant or near significant correlations with both soluble reactive phosphorus (SRP) and total
253 phosphorus (TP) concentrations (Table 3).

254 The relative metagenomic abundance of ciliates and diatoms exhibited some of the
255 strongest positive correlations with relative transcript abundance of BGCs, including those
256 predicted for *aer*, *mdn*, and PKSmod-NRPSlike-T3PKS clusters (Pearson's R=0.653, p=0.041 to
257 Pearson's R=0.517, p=0.003, Table 4). Cyanobactin BGC expression was significantly and
258 positively correlated with the relative abundance of the diatom *Skeletonema* (Pearson's R=
259 0.617, p=0.033). Cyanopeptolin encoding BGC expression was also significantly and positively
260 correlated with the abundance of *Skeletonema* (Pearson's R= 0.741, p=0.006) as well as the
261 betaproteobacterium, *Paraburkholderi*a (Pearson's R= 0.488, p= 0.025). All other biotic
262 correlations are summarized in Table S4. Taken together, these results suggest that both abiotic
263 and biotic variables may be important in stimulating biosynthesis in natural cyanoHAB
264 *Microcystis* populations.

265 **Discussion**

266 *Microcystis* blooms are renowned for their production of microcystins, toxins that
267 threaten drinking water supplies (46–48). However, *Microcystis* genomes contain many other
268 BGCs that encode known and unknown products, and their genetic diversity, biosynthesis, and
269 ecological functions in the environment are poorly understood. In this study, we tracked spatial
270 and temporal shifts in BGC abundance and transcriptional activity through seasonal changes in
271 the biotic and abiotic environment of the 2014 western Lake Erie bloom, providing insights into
272 the biosynthetic potential and expression of diverse *Microcystis* populations in their natural
273 habitat. Our results demonstrate that *Microcystis* contains a diverse and highly dynamic suite of
274 BGCs that putatively encode both known cyanotoxins and cryptic compounds and are

transcriptionally active in natural blooms. The transcriptional activity of these putative BGCs is related to both abiotic conditions (nutrients, temperature) as well as the composition of the biological community. While previous studies have identified BGCs in *Microcystis* isolates (20, 27, 31) or described metabolomic profiles in both field and culture studies (12, 30), our work focused on the relative abundance, diversity, and drivers of transcriptional activity of BGCs within natural populations as they occur *in situ.*

Microcystis MAGs displayed extensive genetic diversity at the sub-species level, which is increasingly recognized to be prevalent in natural blooms (32, 33, 38), and encodes important biosynthetic and physiological diversity (4, 20, 33). Although these MAGs contained some redundancy and short contiguous sequences due to high strain heterogeneity present in natural populations (19, 33), they were sufficient to track spatial and temporal shifts in *Microcystis* gene content. Our approach to estimating relative abundance of BGCs also had limitations. It was based on normalization of BGC abundance to *Microcystis* 16S rRNA abundance. In some instances, the estimated relative abundance of some BGCs was greater than 1, implying multiple copies of a BGC per cell. We discussed potential explanations for this observation previously (33), including uncertainty of 16S rRNA gene and BGC gene copies per *Microcystis* genome (49, 50) in natural populations, presence of multiple copies of BGCs per cell due to replication of genes near the origin of replication during rapid growth (51–53), or non-specific mapping of 16S rRNA and/or BGC genes. In addition, BGCs may be present on plasmids (54), highlighting the potential for multiple copies of these genes per genome. Our previous analysis showed that non-specific mapping of metagenomic sequence reads to *Microcystis* 16S rRNA and BGC genes was not a substantial issue under the stringent thresholds that we used here (33), but the current data is not sufficient to assess other potential issues. Regardless of these caveats, the results of this

298 study do not depend on absolute calculations but rather relative differences, highlighting
299 dynamics in time and space.

300 The diversity of BGCs we observed, and their spatiotemporal shift in abundance, reflects the
301 presence and shifting abundance of diverse *Microcystis* strains with varying BGC content in their
302 genomes (19, 20, 55). Whereas we identified and deeply analyzed nineteen BGCs, *Microcystis*
303 typically has only 5-7 NRPS, PKS, or hybrid NRPS/PKS BGCs per genome (20, 56), and BGC
304 content varies across clades (20). Western Lake Erie *Microcystis* populations contain multiple
305 strains in varying abundance (33, 38, 57), often with microcystin producing strains dominating in
306 early and peak bloom phases and non-microcystin producing strains dominating later phases of
307 blooms (32, 33). Our results show that the "non-toxic" phases of the bloom, so-called because of
308 low concentrations of microcystins, can be enriched in other BGCs (Fig. 2, Table 2) that are
309 transcriptionally active, including those that putatively encode anabaenopeptin, aeruginosin,
310 aeruginoguanidine, cyanopeptolin, and cyanobactins (Fig. 6). Even more striking is the pervasive
311 abundance and transcriptional activity of cryptic, putative, BGCs containing PKS modules with
312 unknown products (Figs. 2,4,6). BGCs with the highest transcript abundance included those
313 encoding unknown compounds such as T3PKS, PKSmod-NRPSlike-T3PKS 1 (thought to
314 putatively synthesize a cyclophane-like product), and MIC 1 (Fig. 6), indicating that they may be
315 functionally and ecologically important. The lack of functional knowledge of these abundant and
316 transcriptionally active PKS containing clusters highlights the need to investigate further their
317 potential for generating unique metabolites.

318 The apparent "replacement" of *mcy* genes by other BGCs that we observed from early to late
319 blooms in the field is consistent with studies of cultured *Microcystis* strains showing that non-
320 microcystin producing strains often contain other BGCs in their place (20). Our finding of a

strong temporal shift in *Microcystis* BGC content and expression suggests environmental drivers of BGC content and expression. Strains containing other BGCs instead of those encoding nitrogen-rich microcystins may be more fit during nitrogen deplete phases of the bloom that occur later in the season (Fig. 1B). For example, cyclophane-like molecules, which may be encoded by cryptic PKSmod-NRPSlike-T3PKS clusters, are nitrogen poor, in many cases lacking N atoms altogether (Table 3, (58)). Strains with BGCs that putatively encode such molecules may have a competitive advantage over microcystin producers with higher N demands (36, 37, 59). However, inconsistent with this conclusion is the increase in abundance of BGCs putatively encoding other N-rich molecules, such as aeruginosin (Table 3), while nitrate concentrations decrease. These results suggest that the different secondary metabolites may have different fitness tradeoffs and/or environmental/physiological controls on gene expression. The high potential for biosynthesis of other secondary metabolites when microcystins are at low concentrations suggests functionally redundancy or niche differentiation among secondary metabolites.

BGCs that were more highly expressed in later bloom phases included *apn* (anabaenopeptin), *aer* (aeruginosin), *mcn* (cyanopeptolin), and cyanobactin classes, which have inhibitory and toxic effects towards eukaryotes (13, 60–62). The T1PKS-M BGCs likely represent poorly assembled microginin synthesis pathways, which, when synthesized, have inhibitory properties against angiotensin-converting enzymes (63, 64). Other transcriptionally active BGCs encode unknown compounds with unknown toxicology. Given the continued emergence and discovery of novel cyanotoxins and their underpinning BGCs (65, 66), and the potential for toxicological synergism between multiple compounds (67, 68), these results underscore the need to assess potential threats to human and ecosystem health.

In addition to potential risks of unmonitored, toxic, cyanobacteria secondary metabolites, these molecules may be an untapped source for drug discovery (69). For example, the T3PKS BGC was one of the most abundant and transcriptionally active PKSs during early phases of the bloom (Fig. 2,6) and contained a putative naringenin-chalcone synthase encoding gene (Fig. 4), which may have the potential to generate antibiotics or other compounds with wide ranges of biological activity and pharmaceutical relevance (70, 71). The cluster bearing an iterative PKS (Fig 4A, 5) is predicted to encode biosynthesis of an enediyne-containing molecule, which are excellent candidate antibiotics and anticancer compounds with rare structural characteristics (72, 73). Previous work has shown that enediyne gene clusters usually contain a set of five conserved biosynthetic genes included the enediyne PKS, a thioesterase, and three other genes that encode proteins of unknown function (73). The iterative PKS cluster contains a putative enediyne PKS, thioesterase, and two of the three genes of unknown function that are part of the enediyne biosynthesis cassette (Fig 4A). This, taken together with the presence of a putative enediyne cluster identified in *Microcystis* strain NIES-843 (45), which shares 97% identity with our cluster, supports the conclusion that *Microcystis* are capable of synthesizing uncharacterized enediyne metabolites. Finally, the presence of a halogenase encoding gene within the MIC1 cluster (Fig. 3D) is noteworthy due to the broad pharmaceutical application of halogenated compounds, which have largely been studied in marine organisms (74).

Correlation of transcript abundance with abiotic and biotic factors provided insights into environmental conditions that may influence expression of BGCs in natural environments. Negative correlations between relative abundance of transcripts and nitrate and temperature suggest a functional need for synthesized compounds in conditions of lower nitrate concentrations and temperature (Table 3). Positive correlations of BGC transcript abundance

with abundances of known eukaryotic predators (Table 4) are consistent with the hypothesis that BGC products may deter grazers (ciliates) that commonly feed on cyanobacteria in aquatic systems (75, 76). Of the cyanobacteria commonly found in western Lake Erie cyanoHABs, *Microcystis* are the most resistant to grazing by both daphnid and protozoan microorganisms (77) consistent with findings that *Microcystis* is resistant to grazing due to production of secondary metabolites (78–80). Likewise, significant correlations between expression of BGCs and abundance of photosynthetic competitors (diatoms) is consistent with a role of secondary metabolites in allelopathic interactions (81, 82). Overall, these results are also consistent with recent findings that community composition and interspecies interactions influence expression of BGCs and secondary metabolites affect interspecies interaction networks (83). *Microcystis* secondary metabolites are likely multi-functional and may serve as nitrogen storage compounds (84, 85), aid in grazer defense (14, 80, 86), protection from reactive oxygen species (87, 88), and/or cell communication/phycosphere recruitment (89, 90). While correlations uncovered here cannot determine true drivers of BGC abundance or expression, they provide hypotheses that can be tested with targeted experiments on the ecological and cellular role of the diversity of secondary metabolites encoded in *Microcystis* genomes.

This study shows that natural populations of *Microcystis* within blooms contain and express a diverse and dynamic suite of genes encoding secondary metabolites. Several of these BGCs have no known products and are expressed most highly during the late phases of the bloom, which is dominated by "nontoxigenic" *Microcystis* strains that are incapable of producing microcystins. These results suggest the potential for diverse metabolite production beyond microcystins in western Lake Erie, with implications for understanding *Microcystis* physiology and diversity, community ecology, and water quality. Seasonal shifts in the abundance and

expression of BGCs along with correlations with both abiotic and biotic variables suggest that the BGCs underpin adaptations to changing environmental conditions and may encode important niche differentiation among *Microcystis* strains. Several clusters identified may encode toxic compounds while others putatively encode cryptic compounds that may be toxic, highlighting the need to identify their products, determine their bioactivities, and assess potential threats to human and ecosystem health. Others show the potential to produce metabolites with interesting chemical features that suggest avenues for exploring biotechnological and medicinal applications. Finally, the abundant environmental expression of BGCs encoding secondary metabolites that are of unknown function, and their correlation with biotic and abiotic conditions, suggest that the BGCs may play important roles in supporting the dominance of *Microcystis* over a wide range of conditions, underscoring the need to understand the functioning of these putative secondary metabolites in the ecophysiology and community interactions of *Microcystis* blooms.

Materials and Methods

Study Site and Sample Collection

Samples were collected weekly from various NOAA Great Lakes Environmental Research Laboratory (GLERL) sampling stations throughout the western Basin of Lake Erie from mid-June through late October 2014 (91). The long-term NOAA GLERL stations selected for sampling were WE2, WE4, and WE12. WE2 is close to the inlet for the Maumee River inlet (41° 45.743'N, 83° 19.874' W), WE4 is considered an offshore site closer to the center of the basin (41° 49.595'N, 83° 11.698'W), and WE12 is proximal to the Toledo drinking water inlet (41° 42.535'N, 83° 14.989'W) (Fig. 1A).

Twenty liters of depth integrated samples were collected from each station for biological and chemical analysis. Integrated depths were defined as the surface of the water to 1m above the lake floor. While samples were collected, measurements of pH and water temperature were performed as well. To capture *Microcystis* aggregates, 2L of integrated depth collected water were filtered through a 100 µm polycarbonate mesh filter. The biomass on the filter was then collected and filtered through a 0.22 µm filter. The filter-bound biomass was preserved in 1 mL of RNALater™ (Invitrogen™, Ambion™) and placed on ice. Upon arrival to the lab, samples were stored in the -80°C freezer until DNA and RNA extractions could be completed.

Sample Processing and Sequencing

Processing and sequencing of these samples was described previously (33). Briefly, DNA was extracted with the Qiagen DNeasy Blood and Tissue Extraction Kit (Qiagen, Hilden, Germany) and quantified with the Quant-iT™ PicoGreen™ dsDNA Assay Kit (Eugene, OR, USA). The Qiagen RNeasy Kit (Qiagen, Hilden, Germany) was used to complete RNA extraction. Shotgun DNA and RNA sequencing was conducted at the University of Michigan DNA Sequencing Core on the Illumina® HiSeq™ platform (2000 PE 100, Illumina, Inc., San Diego, CA, USA).

Bioinformatic Analysis

Shotgun reads were assembled *de novo* to recover *Microcystis* Metagenome Assembled Genomes (MAGs). Metagenomic raw read sequences obtained from the University of Michigan Omics Core were analyzed and processed through the following metagenomic workflow. Raw reads for metagenomic analysis can be found at NCBI BioProject Accession: PRJNA464361.

The IMG-JGI supported software, BBTools was used to complete trimming, adaptor removal, and quality check of raw reads. Quality checked (QC) reads were normalized to 100x coverage using bbnorm from the BBTools package (92). 100x coverage was chosen based on the assembly quality of the following housekeeping genes: a global nitrogen regulator *ntcA*, the beta subunit phycocyanin gene *cpcB*, the photosystem I P700 chlorophyll apoprotein *psaB*, the photosystem II P680 reaction center D1 protein *psbA2*, as well as the assembly quality of the *mcy* operon. Assembly quality was assessed based on contiguous sequence length, and De Bruijn Graph assembly visualized in Bandage (93).

De novo assemblies were generated for each sample using MEGAHIT (94). To obtain differential coverage for binning software, mapping of QC reads was completed using Bowtie2 (95). A multiple binning approach was used in which bins generated from Concoct, Metabat, Tetra ESOM, and Vizbin were assessed and chosen with DASTool (96–100). To improve accuracy of binning software, differential coverage of all 15 samples was used by mapping reads from each sample to each assembly using Bowtie2. The minimum cut off for contiguous sequence length was 1 kb for Concoct, 1.5 kb for Metabat, and 3.5 kb for Tetra ESOM and Vizbin. The cut off for Metabat was selected based on minimum cut off length the program required and that for ESOM and Vizbin was selected based on clarity of maps. *Microcystis* MAGs were assessed and manually refined using Anvi'o v. 5, "Margaret" (101). Completion and strain heterogeneity of each MAG was assessed using CheckM with default parameters (102) as well as metrics generated through Anvi'o outputs. MAGs were mined for BGCs with antiSMASH 6.0 using a full Hidden Markov Modelled search, orthologous gene prediction, gene border prediction, PFAM analysis, and COG identification (42). For further BGC annotation,

identified BGC sequences from all *Microcystis* MAGs found by antiSMASH were run through PRISM (103) software as well.

For downstream analysis, BGC sequences identified by antiSMASH from all *Microcystis* MAGs were pooled together and deduplicated using the dedupe tool from the BBTools package (92). Dedupe was set to deduplicate sequences that were at least 95% identical and had no more than 10 substitutions or deletions. This deduplicated set of BGC clusters was used for subsequent analysis with special attention to specific clusters discussed in greater detail. BGC sequences longer than 10,000 bp were saved for future gene and expression quantification calculations. In addition to these BGCs we also closely examined three identified cyanobactin BGC clusters that were ~7 kb in length. To compare the similarities between clusters of the same class or identification, the clinker software (104) was used to generate gene maps to visually inspect for similarities and differences in gene cluster structure. In addition to using annotations provided by antiSMASH, we aligned our BGCs to a previously described gene clusters found in *Microcystis* (31, 43, 105) to assess the presence of BGCs known to be present in *Microcystis* genomes but are not deposited in the MiBIG database. BGC identified from w. Lake Erie that shared ≥95% identity to previously described clusters they were annotated as such.

Two of the clusters identified from w. Lake Erie MAGs, PKS-M 1 &2, were shown to have 57% similarity to the recently characterized microginin biosynthesis pathway from *Microcystis aeruginosa* LEGE 91341 (43), but were incomplete as they lacked two core NRPS encoding genes, and the hybrid PKS-NRPS encoding gene was shown to be fragmented. To assess if these patterns reflect truncated microginin pathways observed in natural populations, or are a result of poor assembly, metagenomic reads were mapped to the microginin biosynthesis pathway found in LEGE 91341 to assess the completion and presence of this cluster. Read mapping and

coverage plots were then used to visualize and assess the completeness of this pathway and microginin producing potential.

BGC Cluster Quantification

To estimate the relative abundance of BGC clusters the following steps were completed. QC metagenomic reads were aligned to all identified BGC clusters found in any cyanobacterial MAG (including *Microcystis*, *Anabaena, Pseudanabaena*, and *Cyanobium)* generated from this sample set using Basic Local Alignment Search Tool (BLAST) v. 2.8.1 (106). Using all identified BGC clusters expanded our database of biosynthetic genes to allow for competitive mapping. Reads were kept for quantification if they were 95% identical to the query sequence, and at least 80% in alignment length. If reads identically mapped to more than one sequence, i.e., they had identical bit scores, the read count was divided among the total number of ambiguously mapped regions using the blast_hit_counts.sh script found at https://github.com/Geo-omics/scripts/tree/master/scripts. Reads mapped to a BGC were then normalized by BGC length.

To compare across samples, reads were also normalized by the 16S rRNA gene V4 variable region. This is based on the common assumption that all *Microcystis* cells contain one 16S rRNA gene, but do not contain all BGCs; therefore, we used this ratio to estimate the relative abundance of BGCs in *Microcystis* populations. All V4 regions for the entire Phylum of Cyanobacteria were used in the mapping databases to ensure competitive, sensitive, and accurate mapping. These sequences were pulled from the SILVA v.1381.1 database (107) accessed in February 2021. This database can be found at: https://github.com/ceyancey/mcyGenotypes-databases. QC metagenomic reads were aligned to this database using BLAST and kept for downstream analysis if they met the same parameters listed above. These alignment parameters, databases, and mapping tool quality were tested in full to ensure specificity and sensitivity.

Reads were then quantified in the same fashion described above and normalized by 16S rRNA V4 region length. The equation below summarizes how relative BGC cluster abundance was calculated:

$$Coverage\ ratio\ (Microcystis\ BGC{:}\ 16s){:}$$

$$= \frac{Reads\ mapped\ to\ BGC\ \div\ length\ of\ BGC\ (bp)}{Reads\ mapped\ to\ V4\ region\ \div length\ of\ V4\ region\ (bp)} = \frac{BGC\ coverage}{16s\ V4\ coverage}$$

To assess the completeness of cryptic PKS-containing clusters across sampling points, which have not been extensively characterized previously, we additionally calculated the coverage and percentage of nucleotides with reads mapped within a cluster. If at least 80% of the nucleotides within a BGC were covered, it was considered complete for that sampling point. If the percentage of nucleotides within gene clusters were less than 80%, they were considered absent or incomplete, and relative abundance was not calculated. A summary of coverage and completeness for gene clusters MIC 1, iterative PKS, and T3PKS is observed in Table S5.

Biosynthesis Gene Protein Phylogenies

To explore the novelty and relatedness of coarsely annotated PKS clusters from our *Microcystis* MAGs to previously identified biosynthesis pathways, protein phylogenies were generated. For the iterative PKS gene cluster, we assessed phylogenetic relationships for the PKSE gene, believed to be critical in enediyne biosynthesis. To generate a phylogenetic tree, we used PKSE amino acid sequences from biosynthesis pathways known to encode enediyne-containing metabolites in addition to sequences that have been putatively identified as PKSEs in various cyanobacterial taxa (45, 73). The AufC amino acid sequence from aurafuron biosynthesis was used as an outgroup. For the T3PKS cluster, the core T3PKS and a transport related gene were selected for analysis. To find related protein sequences, these sequences were searched

against the nonredundant protein database on NCBI (accessed July 2022) using BLASTP (108). Protein sequences were used to build phylogenetic trees if they had at least a 70% alignment length and at least 30% identity to the target sequences. Protein sequences were aligned using MUSCLE (109) and FastTree (110) was used to construct phylogenetic trees. Bootstrap values equal to or greater than 0.5 were reported. PaperBLAST (https://papers.genomics.lbl.gov/cgi-bin/litSearch.cgi) was used to query published manuscripts to investigate whether these protein sequences had been investigated previously.

BGC Cluster Transcript Abundance

The relative transcript abundance of BGCs *in situ* was also determined. Metatranscriptomic raw reads can be accessed at NCBI under BioProject Accession: PRJNA370007. Alignments were completed using BLAST with identical parameters mentioned previously. However, to normalize gene expression, reads were mapped to reference genome *Microcystis aeruginosa* PCC 7806SL (Accession: CP020771.1, GI: 1181755937) instead of the V4 region. Competitive genome mapping was achieved by creating a database that contains reference genomes for other common cyanobacteria found in western Lake Erie cyanoHABs: *Anabaena sp.* 90 (GCA_000312705.1), *Cyanobium sp.* NIES-981 (GCA_900088535.1), and *Pseudanabaena sp.* PCC 7367 (GCA_000317065.1). The following equation summarizes calculations completed to relatively quantify BGC expression:

$$Relative\ Transcript\ Abundance = \frac{reads\ mapped\ to\ BGC \div length\ of\ BGC\ (bp)}{reads\ mapped\ to\ PCC\ 7806SL \div length\ of\ PCC\ 7806\ SL\ (bp)}$$

BGC Expression Correlation Analysis

To assess the relationships between expression and 1) abiotic environmental conditions and 2) abundance of known competitors and predators of *Microcystis*, simple linear regressions

were generated. Correlations were generated for all abiotic variables, and for all organisms identified in each metagenome that had at least 200 genes assigned to their taxonomic identification. Organism abundance and analysis was completed at the genus level. Pearson's correlation coefficient, R, as well as p-values were determined to assess the strength of correlation and the significance of individual variables, respectively. Abiotic environmental condition data was provided by CIGLR and can be found in Table S6. The relative abundance of organisms was determined by aligning QC reads to the UniRef100 database using DIAMOND (111). Relative abundance of each organism was then determine by calculating RPKM of all genes assigned to the taxa using the AnnotateContigs.pl script found at https://github.com/TealFurnholm/Strain-Level_Metagenome_Analysis/wiki/Step-2.-Gene-Annotation-and-Community-Analysis which summarizes the complete community analysis workflow and outputs.

Data Visualization

All figures were generated using R and R Studio v. 3.5.1 (112). Packages used for data visualization included ggplot2 (113), ggthemes (114), and ggpubr (115). To generate the map in Figure 1, QGIS was used using the Quick Map Services Plugin (QGIS Development Team, 2020). The base map used is an open-source resource, the Open Street Map (OSM, https://wiki.osmfoundation.org/wiki/Main_Page). Clinker (104) software was used to generate BGC comparison figures observed in Figure S1. Metagenomic and metatranscriptomic raw reads can be found on NCBI under the BioProject Accessions PRJNA464361 and PRJNA370007, respectively.

Conflict of Interest

The authors state no conflict of interest.

Acknowledgements

We thank Robert Hein, Teal Furnholm, and Derek Smith for the bioinformatic support. We also thank the field crew at CIGLR/GLERL including Paul Den Uyl, Dack Stuart, and Kent Baker for allowing us to sample with them and assisting in field measurements.

Funding

This work was supported by NIH and NSF awards to the Great Lakes Center for Fresh Waters and Human Health (NIH: 1P01ES028939-01, NSF: OCE-1840715), the Cooperative Institute of Great Lakes Research (NA17OAR4320152), and the Hans W. Vahlteich Professorship. Support was also provided by UM Natural Products Biosciences Initiative, NOAA OAR 'Omics and NOAA OAR Ocean Technology Development Initiative.

Data availability

All data pertinent to this manuscript is accessible for review within the manuscript. Metagenomic and metatranscriptomic raw reads can be found on NCBI under the BioProject Accessions PRJNA464361 and PRJNA370007, respectively.

References

1. Ho JC, Michalak AM, Pahlevan N. 2019. Widespread global increase in intense lake phytoplankton blooms since the 1980s. Nature 574:667–670.

2. Huisman J, Codd GA, Paerl HW, Ibelings BW, H Verspagen JM, Visser PM. 2018. Cyanobacterial blooms. Nat Rev Microbiol 16:471-483.

3. Griffith AW, Gobler CJ. 2020. Harmful algal blooms: A climate change co-stressor in marine and freshwater ecosystems. Harmful Algae 91:101590.

4. Steffen MM, Davis TW, McKay RML, Bullerjahn GS, Krausfeldt LE, Stough JMA, Neitzey ML, Gilbert NE, Boyer GL, Johengen TH, Gossiaux DC, Burtner AM, Palladino D, Rowe MD, Dick GJ, Meyer KA, Levy S, Boone BE, Stumpf RP, Wynne TT, Zimba P v., Gutierrez D, Wilhelm SW. 2017. Ecophysiological

Examination of the Lake Erie *Microcystis* Bloom in 2014: Linkages between Biology and the Water Supply Shutdown of Toledo, OH. Environ Sci Technol 51:6745–6755.

5. Harke MJ, Steffen MM, Otten TG, Wilhelm SW, Wood SA, Paerl HW. 2016. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. Harmful Algae 54:4–20.

6. Watson SB, Miller C, Arhonditsis G, Boyer GL, Carmichael W, Charlton MN, Confesor R, Depew DC, Höök TO, Ludsin SA, Matisoff G, McElmurry SP, Murray MW, Peter Richards R, Rao YR, Steffen MM, Wilhelm SW. 2016. The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia. Harmful Algae 56:44–66.

7. Heisler J, Glibert PM, Burkholder JM, Anderson DM, Cochlan W, Dennison WC, Dortch Q, Gobler CJ, Heil CA, Humphries E, Lewitus A, Magnien R, Marshall HG, Sellner K, Stockwell DA, Stoecker DK, Suddleson M. 2008. Eutrophication and harmful algal blooms: A scientific consensus. Harmful Algae 8:3–13.

8. Paerl HW, Otten TG, Kudela R. 2018. Mitigating the Expansion of Harmful Algal Blooms Across the Freshwater-to-Marine Continuum. Environ Sci Technol 52:5519–5529.

9. Krausfeldt LE, Farmer AT, Castro Gonzalez HF, Zepernick BN, Campagna SR, Wilhelm SW. 2019. Urea Is Both a Carbon and Nitrogen Source for *Microcystis aeruginosa*: Tracking 13C Incorporation at Bloom pH Conditions. Front Microbiol 10:1064.

10. Chapra SC, Boehlert B, Fant C, Victor J. Bierman Jr, Henderson J, Mills D, Mas DML, Rennels L, Jantarasami L, Martinich J, Strzepek KM, Paerl HW. 2017. Climate Change Impacts on Harmful Algal Blooms in U.S. Freshwaters: A Screening-Level Assessment. Environ Sci Technol 51:8933–8943.

11. Kehr J-C, Gatte Picchi D, Dittmann E. 2011. Natural product biosyntheses in cyanobacteria: A treasure trove of unique enzymes. Beilstein J. Org. Chem 7:1622–35.

12. le Manach S, Duval C, Marie A, Djediat C, Catherine A, Edery M, Bernard C, Marie B. 2019. Global Metabolomic Characterizations of Microcystis spp. Highlights Clonal Diversity in Natural Bloom-Forming Populations and Expands Metabolite Structural Diversity. Front Microbiol 10:791.

13. Welker M, von Döhren H. 2006. Cyanobacterial peptides - Nature's own combinatorial biosynthesis. FEMS Microbiol Rev 4:530-563.

14. Wiegand C, Pflugmacher S. 2005. Ecotoxicological effects of selected cyanobacterial secondary metabolites a short review. Toxicol Appl Pharmacol 203:201–218.

15. Pearson LA, Dittmann E, Mazmouz R, Ongley SE, D'Agostino PM, Neilan BA. 2016. The genetics, biosynthesis and regulation of toxic specialized metabolites of cyanobacteria. Harmful Algae 54:98–111.

16. Newman DJ, Cragg GM. 2020. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. J Nat Prod 83:770–803.

17. Ehrenreich IM, Waterbury JB, Webb EA. 2005. Distribution and diversity of natural product genes in marine and freshwater cyanobacterial cultures and genomes. Appl Environ Microbiol 71:7401–13.

18. Dittmann E, Gugger M, Sivonen K, Fewer DP. 2015. Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. Trends Microbiol 23:642–652.

19. Dick GJ, Duhaime MB, Evans JT, Errera RM, Godwin CM, Kharbush JJ, Nitschky HS, Powers MA, Vanderploeg HA, Schmidt KC, Smith DJ, Yancey CE, Zwiers CC, Denef VJ. 2021. The genetic and ecophysiological diversity of *Microcystis*. Environ Microbiol 1462-2920.15615.

20. Pérez-Carrascal OM, Terrat Y, Giani A, Fortin N, Greer CW, Tromas N, Shapiro BJ. 2019. Coherence of *Microcystis* species revealed through population genomics. ISME 13:2887–2900.

21. Tillett D, Dittmann E, Erhard M, Von Döhren H, Börner T, Neilan BA. 2000. Structural organization of microcystin biosynthesis in Microcystis aeruginosa PCC7806: An integrated peptide-polyketide synthetase system. Chem Biol 7:753–764.

22. Carmichael WW. 1992. Cyanobacteria secondary metabolites-the cyanotoxins. J. Appl Bacteriol 72:445–459.

23. Hughes EO, Gorham PR, Zehnder A. 1958. Toxicity of a unialgal culture of *Microcystis aeruginosa*. Can J Microbiol 4:225–236.

24. Qin B, Zhu G, Gao G, Zhang Y, Li W, Paerl HW, Carmichael WW. 2010. A Drinking Water Crisis in Lake Taihu, China: Linkage to Climatic Variability and Lake Management. Environ Manage 45:105–112.

25. Kotut K, Ballot A, Krienitz L. 2006. Toxic cyanobacteria and their toxins in standing waters of Kenya: implications for water resource use. J Water Health 4:233–245.

26. Stewart I, Seawright AA, Shaw GR. 2008. Cyanobacterial poisoning in livestock, wild mammals and birds--an overview. Adv Exp Med Biol 613-637.

27. Pearson LA, Crosbie ND, Neilan BA. 2020. Distribution and conservation of known secondary metabolite biosynthesis gene clusters in the genomes of geographically diverse *Microcystis aeruginosa* strains. Mar Freshw Res 71:701.

28. Kust A, Řeháková K, Vrba J, Maicher V, Mareš J, Hrouzek P, Chiriac MC, Benedová Z, Tesařová B, Saurav K. 2020. Insight into Unprecedented Diversity of Cyanopeptides in Eutrophic Ponds Using an MS/MS Networking Approach. Toxins 12:561.

29. Beversdorf LJ, Weirich CA, Bartlett SL, Miller TR. 2017. Variable Cyanobacterial Toxin and Metabolite Profiles across Six Eutrophic Lakes of Differing Physiochemical Characteristics. Toxins 9:62.

30. Janssen EML. 2019. Cyanobacterial peptides beyond microcystins – A review on co-occurrence, toxicity, and challenges for risk assessment. Water Res 151:488-499.

31. Humbert J-F, Barbe V, Latifi A, Gugger M, Calteau A, Coursin T, Lajus A, Castelli V, Oztas S, Samson G, Longin C, Medigue C, de Marsac NT. 2013. A Tribute to Disorder in the Genome of the Bloom-Forming Freshwater Cyanobacterium *Microcystis aeruginosa*. PLoS One 8:e70747.

32. Berry MA, Davis TW, Cory RM, Duhaime MB, Johengen TH, Kling GW, Marino JA, den Uyl PA, Gossiaux D, Dick GJ, Denef VJ. 2017. Cyanobacterial harmful algal blooms are a biological disturbance to Western Lake Erie bacterial communities. Environ Microbiol 19:1149–1162.

33. Yancey CE, Smith DJ, Uyl PA den, Mohamed OG, Yu F, Ruberg SA, Chaffin JD, Goodwin KD, Tripathi A, Sherman DH, Dick GJ. 2022. 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. Appl Environ Microbiol 88:e02464-21.

34. Steffen MM, Belisle BS, Watson SB, Boyer GL, Wilhelm SW. 2014. Status, causes and controls of cyanobacterial blooms in Lake Erie. J Great Lakes Res 40:215–225.

35. Lake Erie Programs | Ohio Environmental Protection Agency. https://epa.ohio.gov/divisions-and-offices/surface-water/reports-data/lake-erie-programs. Retrieved 5 June 2022.

36. van de Waal DB, Verspagen JMH, Lürling M, van Donk E, Visser PM, Huisman J. 2009. The ecological stoichiometry of toxins produced by harmful cyanobacteria: An experimental test of the carbon-nutrient balance hypothesis. Ecol Lett 12:1326–1335.

37. Wagner ND, Osburn FS, Wang J, Taylor RB, Boedecker AR, Chambliss CK, Brooks BW, Scott JT. 2019. Biological stoichiometry regulates toxin production in *Microcystis aeruginosa* (UTEX 2385). Toxins (Basel) 11:601.

38. Smith DJ, Tan JY, Powers MA, Lin XN, Davis TW, Dick GJ. 2021. Individual Microcystis colonies harbour distinct bacterial communities that differ by *Microcystis* oligotype and with time. Environ Microbiol 23:3020–3036.

39. Smith DJ, Berry MA, Cory RM, Johengen TH, Kling GW, Davis TW, Dick GJ. 2022. Heterotrophic Bacteria Dominate Catalase Expression during *Microcystis* Blooms. Appl Environ Microbiol 88.

40. Dick GJ. 2018. Genomic Approaches in Earth and Environmental Sciences, 1st ed. John Wiley & Sons, Ltd.

41. Chen LX, Anantharaman K, Shaiber A, Murat Eren A, Banfield JF. 2020. Accurate and complete genomes from metagenomes. Genome Res 30:315–333.

42. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35.

43. Eusébio N, Castelo-Branco R, Sousa D, Preto M, D'agostino P, Gulder TAM, Leão PN. 2022. Discovery and heterologous expression of microginins from *Microcystis aeruginosa* LEGE 91341. ACS Synth Biol 10:3493-3503.

44. Nicolaou KC, Dai W -M. 1991. Chemistry and Biology of the Enediyne Anticancer Antibiotics. Angew. Chem., Int. Ed 30:1387–1416.

45. Rudolf JD, Yan X, Shen B. 2016. Genome neighborhood network reveals insights into enediyne biosynthesis and facilitates prediction and prioritization for discovery. J Ind Microbiol Biotechnol 43:261–276.

46. Dawson RM. 1998. The toxicology of microcystins. Toxicon. 36:953-962.

47. Campos A, Vasconcelos V. 2010. Molecular Mechanisms of Microcystin Toxicity in Animal Cells. International Journal of Molecular Sciences 11:268–287.

48. Bouaïcha N, Miles CO, Beach DG, Labidi Z, Djabri A, Benayache NY, Nguyen-Quang T. 2019. Structural diversity, characterization and toxicology of microcystins. Toxins (Basel) 11:714.

49. Vaitomaa J, Rantala A, Halinen K, Rouhiainen L, Tallberg P, Mokelke L, Sivonen K. 2003. Quantitative real-time PCR for determination of microcystin synthetase e copy numbers for microcystis and anabaena in lakes. Appl Environ Microbiol 69:7289–7297.

50. Davis TW, Berry DL, Boyer GL, Gobler CJ. 2009. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. Harmful Algae 8:715–725.

51. Roth JR, Benson N, Galitski T, Haack K, Lawrence JG, Miesel L. Rearrangements of the Bacterial Chromosome: Formation and Applications.

52. Rocha EPC. 2004. The replication-related organization of bacterial genomes. Microbiol 150:1609–1627.

53. Couturier E, Rocha EPC. 2006. Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. Mol Microbiol 59:1506–1518.

54. Popin RV, Alvarenga DO, Castelo-Branco R, Fewer DP, Sivonen K. 2021. Mining of Cyanobacterial Genomes Indicates Natural Product Biosynthetic Gene Clusters Located in Conjugative Plasmids. Front Microbiol 12:3353.

55. Cai H, McLimans CJ, Beyer JE, Krumholz LR, Hambright KD. 2023. *Microcystis* pangenome reveals cryptic diversity within and across morphospecies. Sci Adv 9:eadd3783.

56. Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau De Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport KW, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M, Kerfeld CA. 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci 110:1053–1058.

57. Rinta-Kanto JM, Konopko EA, DeBruyn JM, Bourbonniere RA, Boyer GL, Wilhelm SW. 2009. Lake Erie *Microcystis*: Relationship between microcystin production, dynamics of genotypes and environmental parameters in a large lake. Harmful Algae 8:665–673.

58. Martins TP, Rouger C, Glasser NR, Freitas S, de Fraissinette NB, Balskus EP, Tasdemir D, Leão PN. 2019. Chemistry, bioactivity and biosynthesis of cyanobacterial alkylresorcinols. Nat Prod Rep 36:1437–1461.

59. van de Waal D, Smith V, Declerck S, Stam E, Esler J. 2014. Stoichiometric regulation of phytoplankton toxins. Ecol Lett 17:736–742.

60. Lenz KA, Miller TR, Ma H. 2019. Anabaenopeptins and cyanopeptolins induce systemic toxicity effects in a model organism the nematode *Caenorhabditis elegans*. Chemosphere 214:60–69.

61. Murakami M, Ishida K, Okino T, Okita Y, Matsuda H, Yamaguchi K. 1995. Aeruginosins 98-A and B, trypsin inhibitors from the blue-green alga *Microcystis aeruginosa* (NIES-98). Tetrahedron Lett 36:2785–2788.

62. Leikoski N, Fewer DP, Jokela J, Alakoski P, Wahlsten M, Sivonen K. 2012. Analysis of an Inactive Cyanobactin Biosynthetic Gene Cluster Leads to Discovery of New Natural Products from Strains of the Genus *Microcystis*. PLoS One 7:e43002.

63. Okino T, Matsuda H, Murakami M, Yamaguchi K. 1993. Microginin, an angiotensin-converting enzyme inhibitor from the blue-green alga *Microcystis aeruginosa*. Tetrahedron Lett 34:501–504.

64. Rounge TB, Rohrlack T, Nederbragt AJ, Kristensen T, Jakobsen KS. 2009. A genome-wide analysis of nonribosomal peptide synthetase gene clusters and their peptides in a *Planktothrix rubescens* strain. BMC Genomics 10:396.

65. Lima ST, Fallon TR, Cordoza JL, Chekan JR, Delbaje E, Hopiavuori AR, Alvarenga DO, Wood SM, Luhavaya H, Baumgartner JT, Dörr FA, Etchegaray A, Pinto E, McKinnie SMK, Fiore MF, Moore BS. 2022. Biosynthesis of Guanitoxin Enables Global Environmental Detection in Freshwater Cyanobacteria. J Am Chem Soc 144:9372–9379.

66. Breinlinger S, Phillips TJ, Haram BN, Mareš J, Martínez Yerena JA, Hrouzek P, Sobotka R, Henderson WM, Schmieder P, Williams SM, Lauderdale JD, Wilde HD, Gerrin W, Kust A, Washington JW, Wagner C, Geier B, Liebeke M, Enke H, Niedermeyer THJ, Wilde SB. 2021. Hunting the eagle killer: A cyanobacterial neurotoxin causes vacuolar myelinopathy. Science 371:6536.

67. Fernandes K, Gomes A, Calado L, Yasui G, Assis D, Henry T, Fonseca A, Pinto E. 2019. Toxicity of Cyanopeptides from Two Microcystis Strains on Larval Development of *Astyanax altiparanae*. Toxins (Basel) 11:220.

68. Pawlik-Skowrońska B, Bownik A. 2022. Synergistic toxicity of some cyanobacterial oligopeptides to physiological activities of *Daphnia magna* (Crustacea). Toxicon 206:74–84.

69. Aesoy R, Herfindal L. 2022. Cyanobacterial anticancer compounds in clinical use: Lessons from the dolastatins and cryptophycins. The Pharmacological Potential of Cyanobacteria 55–79.

70. Bhat ZS, Rather MA, Maqbool M, Lah HU, Yousuf SK, Ahmad Z. 2017. α-pyrones: Small molecules with versatile structural diversity reflected in multiple pharmacological activities-an update. Biomedicine & Pharmacotherapy 91:265–277.

71. Austin B, M, Noel J, P. 2003. The chalcone synthase superfamily of type III polyketide synthases. Nat Prod Rep 20:79–110.

72. Lanen SG van, Shen B. 2008. Biosynthesis of Enediyne Antitumor Antibiotics. Curr Top Med Chem 8:448.

73. Shen B, Hindra, Yan X, Huang T, Ge H, Yang D, Teng Q, Rudolf JD, Lohman JR. 2015. Enediynes: Exploration of microbial genomics to discover new anticancer drug leads. Bioorg Med Chem Lett 25:9–15.

74. Gribble GW. 2015. Biological Activity of Recently Discovered Halogenated Marine Natural Products. Marine Drugs 13:4044–4136.

75. Calbet A, Landry MR. 2004. Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. Limnol Oceanogr 49:51–57.

76. Gobler CJ, Davis TW, Deonarine SN, Saxton MA, Lavrentyev PJ, Jochem FJ, Wilhelm SW. 2008. Grazing and virus-induced mortality of microbial populations before and during the onset of annual hypoxia in Lake Erie. Aquatic Microbial Ecology 51:117–128.

77. Ladds M, Jankowiak J, Gobler CJ. 2021. Novel high throughput sequencing - fluorometric approach demonstrates Microcystis blooms across western Lake Erie are promoted by grazing resistance and nutrient enhanced growth. Harmful Algae 110:102126.

78. Davis TW, Gobler CJ. 2011. Grazing by mesozooplankton and microzooplankton on toxic and non-toxic strains of *Microcystis* in the Transquaking River, a tributary of Chesapeake Bay. J Plankton Res 33:415–430.

79. Yang Z, Kong F, Shi X, Cao H. 2006. Morphological Response of Microcystis aeruginosa to Grazing by Different Sorts of Zooplankton. Hydrobiologia 563:225–230.

80. Sadler T, von Elert E. 2014. Physiological interaction of Daphnia and Microcystis with regard to cyanobacterial secondary metabolites. Aquat Toxicol 156:96–105.

81. Chia MA, Jankowiak JG, Kramer BJ, Goleski JA, Huang IS, Zimba P v., do Carmo Bittencourt-Oliveira M, Gobler CJ. 2018. Succession and toxicity of *Microcystis* and *Anabaena* (Dolichospermum) blooms are controlled by nutrient-dependent allelopathic interactions. Harmful Algae 74:67–77.

82. Pflugmacher S. 2002. Possible allelopathic effects of cyanotoxins, with reference to microcystin-LR, in aquatic ecosystems. Environ Toxicol 17:407–413.

83. Chevrette MG, Thomas CS, Hurley A, Rosario-Melendez N, Sankaran K, Tu Y, Hall A, Magesh S, Handelsman J. 2022. Microbiome composition modulates secondary metabolism in a multispecies bacterial community. Proc Natl Acad Sci 119:e2212930119.

84. Harke MJ, Gobler CJ. 2013. Global Transcriptional Responses of the Toxic Cyanobacterium, *Microcystis aeruginosa*, to Nitrogen Stress, Phosphorus Stress, and Growth on Organic Matter. PLoS One 8:e69834.

85. Wang J, Wagner ND, Fulton JM, Scott JT. 2021. Diazotrophs modulate phycobiliproteins and nitrogen stoichiometry differently than other cyanobacteria in response to light and nitrogen availability. Limnol Oceanogr 66:2333–2345.

86. Penn K, Wang J, Fernando SC, Thompson JR. 2014. Secondary metabolite gene expression and interplay of bacterial functions in a tropical freshwater cyanobacterial bloom. ISME J 8:1866–1878.

87. Zilliges Y, Kehr JC, Meissner S, Ishida K, Mikkat S, Hagemann M, Kaplan A, Börner T, Dittmann E. 2011. The Cyanobacterial Hepatotoxin Microcystin Binds to Proteins and Increases the Fitness of *Microcystis* under Oxidative Stress Conditions. PLoS One 6:e17615.

88. Alexova R, Fujii M, Birch D, Cheng J, Waite TD, Ferrari BC, Neilan BA. 2011. Iron uptake and toxin synthesis in the bloom-forming *Microcystis aeruginosa* under iron limitation. Environ Microbiol 13:1064–1077.

89. Gan N, Xiao Y, Zhu L, Wu Z, Liu J, Hu C, Song L. 2012. The role of microcystins in maintaining colonies of bloom-forming Microcystis spp. Environ Microbiol 14:730–742.

90. Baran R, Brodie EL, Mayberry-Lewis J, Hummel E, da Rocha UN, Chakraborty R, Bowen BP, Karaoz U, Cadillo-Quiroz H, Garcia-Pichel F, Northen TR. 2015. Exometabolite niche partitioning among sympatric soil bacteria. Nat Commun 2015 6:1 6:1–9.

91. Cooperative Institute for Great Lakes Research U of MNGLERL. Physical, chemical, and biological water quality monitoring data to support detection of Harmful Algal Blooms (HABs) in western Lake Erie, collected by the Great Lakes Environmental Research Laboratory and the Cooperative Institute for Great Lakes Research since 2012. https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:GLERL-CIGLR-HAB-LakeErie-water-qual. Retrieved 6 April 2022.

92. Bushnell B. BBTools User Guide - DOE Joint Genome Institute. https://sourceforge.net/projects/bbmap/. Retrieved 8 April 2021.

93. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. Bioinformatics 31:3350–3352.

94. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676.

95. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

96. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Loman NJ, Andersson AF, Quince C. 2013. CONCOCT: Clustering cONtigs on COverage and ComposiTion.

97. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165.

98. Ultsch A, Mörchen F. 2009. ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM Analysis of high dimensional and big data View project Genetic Foundations of Leukemia View project ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM.

99. Laczny CC, Sternal T, Plugaru V, Gawron P, Atashpendar A, Margossian H, Coronado S, der Maaten L, Vlassis N, Wilmes P. 2015. VizBin - an application for reference-independent visualization and human-augmented binning of metagenomic data. Microbiome 3:1.

100. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843.

101. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. PeerJ 3:e1319.

102. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–55.

103. Skinnider MA, Johnston CW, Gunabalasingam M, Merwin NJ, Kieliszek AM, MacLellan RJ, Li H, Ranieri MRM, Webster ALH, Cao MPT, Pfeifle A, Spencer N, To QH, Wallace DP, Dejong CA, Magarvey NA. 2020. Comprehensive prediction of secondary metabolite structure and biological activity from microbial genome sequences. Nat Commun 11:1–9.

104. Gilchrist CLM, Chooi Y-H. 2021. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37:2473-2475.

105. Pancrace C, Ishida K, Briand E, Pichi DG, Weiz AR, Guljamow A, Scalvenzi T, Sassoon N, Hertweck C, Dittmann E, Gugger M. 2019. Unique Biosynthetic Pathway in Bloom-Forming Cyanobacterial Genus *Microcystis* Jointly Assembles Cytotoxic Aeruginoguanidines and Microguanidines. ACS Chem Biol 14:67–75.

106. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

107. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.

108. Madden T. 2013. The BLAST Sequence Analysis Tool. The NCBI Handbook, 2nd ed.

109. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.

110. Price MN, Dehal PS, Arkin AP. 2009. FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. Mol Biol Evol 26:1641–1650.

111. Buchfink B, Reuter K, Drost HG. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 2021 18:4 18:366–368.

112. Allaire JJ. 2015. RStudio: Integrated Development Environment for R.

113. Wickham H. 2011. ggplot2. Wiley Interdiscip Rev Comput Stat 3:180–185.

114. Arnold JB. 2021. ggthemes: Extra Themes, Scales and Geoms for "ggplot2." https://cran.r-project.org/web/packages/ggthemes/index.html. Retrieved 14 January 2021.

115. Kassambara A. 2020. ggpubr: "ggplot2" Based Publication Ready Plots. https://cran.r-project.org/web/packages/ggpubr/index.html. Retrieved 14 January 2021.

116. QGIS Development Team. 2020. QGIS Geo Graphic Information System. Open Source Geospatial 742 Foundation Project. https://www.qgis.org/en/site/. Retrieved 9 March 2021.

117. Ishida K, Okita Y, Matsuda H, Okino T, Murakami M. 1999. Aeruginosins, protease inhibitors from the cyanobacterium *Microcystis aeruginosa*. Tetrahedron 55:10971–10988.

118. Ishida K, Welker M, Christiansen G, Cadel-Six S, Bouchier C, Dittmann E, Hertweck C, de Marsac NT. 2009. Plasticity and evolution of aeruginosin biosynthesis in cyanobacteria. Appl Environ Microbiol 75:2017–2026.

119. Cadel-Six S, Dauga C, Castets AM, Rippka R, Bouchier C, Tandeau De Marsac N, Welker M. 2008. Halogenase genes in nonribosomal peptide synthetase gene clusters of *Microcystis* (cyanobacteria): sporadic distribution and evolution. Mol Biol Evol 25:2031–2041.

120. Rounge TB, Rohrlack T, Tooming-Klunderud A, Kristensen T, Jakobsen KS. 2007. Comparison of cyanopeptolin genes in *Planktothrix*, *Microcystis*, and *Anabaena* strains: Evidence for independent evolution within each genus. Appl Environ Microbiol 73:7322–7330.

121. Nishizawa T, Ueda A, Nakano T, Nishizawa A, Miura T, Asayama M, Fujii K, Harada KI, Shirai M. 2011. Characterization of the locus of genes encoding enzymes producing heptadepsipeptide micropeptin in the unicellular cyanobacterium *Microcystis*. J Biochem 149:475–485.

122. Ziemert N, Ishida K, Liaimer A, Hertweck C, Dittmann E. 2008. Ribosomal Synthesis of Tricyclic Depsipeptides in Bloom-Forming Cyanobacteria. Angewandte Chemie 120:7870–7873.

Tables and Figure Captions

Table 1: Summary and Statistics of *Microcystis* MAGs

Station	Date	Completion (%)	Redundancy (%)	N50	Size (Mbp)	Number of contigs	GC Content (%)
WE12	4-Aug	40.29	2.16	2907	0.816	335	39.67
	25-Aug	89.21	35.25	2353	5.99	2767	43.17
	23-Sep	100	4.32	8601	5.31	903	42.79
	29-Sep	93.53	10.23	2548	4.27	1888	43.4
	20-Oct	96.4	12.95	4155	4.92	1461	42.83
WE2	21-Jul	97.84	4.32	5000	4.27	1102	42.9
	6-Oct	97.12	10.07	5061	5.39	1404	43.28
	20-Oct	99.28	5.76	5472	4.69	1108	42.8
WE4	29-Jul	92.81	20.86	1786	3.93	2234	43.53
	8-Sep	94.96	9.67	2668	4.29	1797	43.49

Table 2: Summary of Identified BGCs and putative synthesis products from WLE MAGs.

BGC name	Putative synthesis product	Described in
aer 1	aeruginosin	Ishida et al., 1999; Ishida et al., 2009; Cadel-Six et al., 2008, Humbert et al., 2013 (31, 117–119)
aer 2	aeruginosin	
agd	aeruginoguanidine	Pancrace et al., 2019 (105)
apn	anabaenopeptin	Welker and Dohren 2006; Humbert et al., 2013 (13, 31)
cyanobactin 1	piricyclamide	

cyanobactin 2	piricyclamide	Leikoski et al., 2012 (62)
cyanobactin 3	piricyclamide	
mcn 1	cyanopeptolin	Tooming-Klunderud et al., 2007; Nishizawa et al., 2011, Humbert et al., 2013 (31, 120, 121)
mcn 2	cyanopeptolin	
mdn 1	microviridin	Ziemert et al., 2008; Humbert et al., 2013 (31, 122)
mdn 2	microviridin	
MIC1 (T1PKS-NRPS-hgIEKS)	unknown	Humbert et. al., 2013 (31)
Iterative PKS	unknown, possible enediyne	Humbert et. al., 2013 (31)
PKS-M 1	microginin	Humbert et al., 2013; Eusébio et al., 2022 (31, 43)
PKS-M 2	microginin	
PKSmod-NRPSlike-T3PKS 1	cyclophane-like	Humbert et. al., 2013 (31)
PKSmod-NRPSlike-T3PKS 2	cyclophane-like	
T1PKS 1	unknown	Novel
T3PKS	unknown	Novel

884

885 **Table 3:** Abiotic and BGC expression Correlations.

	T1PKSmod-NRPSlike T3PKS (cyclophane-like)		**cyanobactin**		***mcn* (cyanopeptolin)**		***mdn* (microvirdin)**		***aer* (aeruginosin)**	
C:N ratio	~0	C-rich	3-6	N-rich	4.7	N-rich	5.25	N-rich	~6	N-rich
	R-squared	p-value	R-squared	p-value	R-squared	p-value	R-squared	p-value	R-squared	p-value
NO_3	-0.551	0.041	-0.418	0.06	-0.434	0.049	-0.749	0.002	-0.48	0.028
Temperature	-0.598	0.04	-0.531	0.029	-0.412	0.089	-0.423	0.17	-0.375	0.1249
SRP	0.769	0.0013	0.251	0.27	0.216	0.35	0.358	0.21	0.364	0.1
TP	0.601	0.023	0.114	0.62	0.227	0.32	0.537	0.048	0.449	0.04

886

887 **Table 4:** Pearson's correlations for BGC expression and selected organism abundance

		***aer* (aeruginosin)**		**cyanobactin**		***mcn* (cyanopeptolin)**		***mdn* (microvirdin)**		**T1PKSmod-NRPSlike T3PKS (cyclophane-like)**	
	Genus	R	p-value	R	p-value	R	p-value	R	p-value	R	p-value
Ciliates	*ICHTHYOPHTHIRIUS*	0.553	0.03	0.245	0.378	0.336	0.221	0.653	0.041	0.79	0.00

			3							9	6
	PARAMECIUM	0.562	0.029	0.256	0.379	0.340	0.215	0.657	0.039	0.810	0.005
	PSEUDOCOHNILEMBUS	0.549	0.034	0.230	0.410	0.326	0.236	0.641	0.046	0.803	0.005
	TETRAHYMENA	0.563	0.029	0.238	0.393	0.335	0.222	0.659	0.038	0.817	0.004
Diatoms	*CRASPEDOSTAUROS*	0.567	0.054	0.298	0.347	0.436	0.156	0.414	0.308	0.853	0.007
	SKELETONEMA	0.630	0.028	0.617	0.033	0.741	0.006	0.832	0.010	0.031	0.941
Beta-proteobacteria	*PARABURKHOLDERIA*	0.296	0.285	0.081	0.773	0.117	0.678	0.380	0.279	0.640	0.046
	RHODOFERAX	0.409	0.066	0.419	0.059	0.488	0.025	0.669	0.009	0.171	0.559
Proteobacteria	*OLIGOFLEXUS*	0.505	0.094	0.057	0.860	0.191	0.553	0.530	0.177	0.828	0.011
Cytophagia	*CHRYSEOTALEA*	0.405	0.096	0.149	0.555	0.276	0.030	0.360	0.250	0.448	0.144

888

889 **Figure 1:** Overview of the 2014 W. Lake Erie cyanoHAB. A) Map of western Lake Erie and the
890 sampling sites used for this study. WE2 and WE12 are close to the Ohio coast and Maumee
891 River and are considered nearshore stations, while WE4 is located more centrally in the basin
892 and considered offshore. B) Various bloom metrics for the 2014 bloom including cyanobacterial
893 biomass as measured by both chlorophyll α (blue) and phycocyanin (green) concentrations, as
894 well as measured concentrations of particulate microcystins, soluble reactive phosphorus, and
895 nitrate. Figure 1A was generated via QGIS using the Open Street Map (OSM)
896 (https://wiki.osmfoundation.org/wiki/Main_Page).

897 **Figure 2:** Relative abundance of biosynthetic gene clusters (BGCs), including the complete mcy
898 operon described previously (33). Relative abundance is presented as estimated relative
899 proportion of the *Microcystis* population containing the BGC (see methods). Predicted secondary
900 metabolites are indicated on the right vertical axis) an in the legend. On 4 August at WE2, the
901 *mcy* operon and T3PKS gene clusters had a relative proportion greater than 1, indicating more
902 copies of BGC genes than 16S rRNA genes.

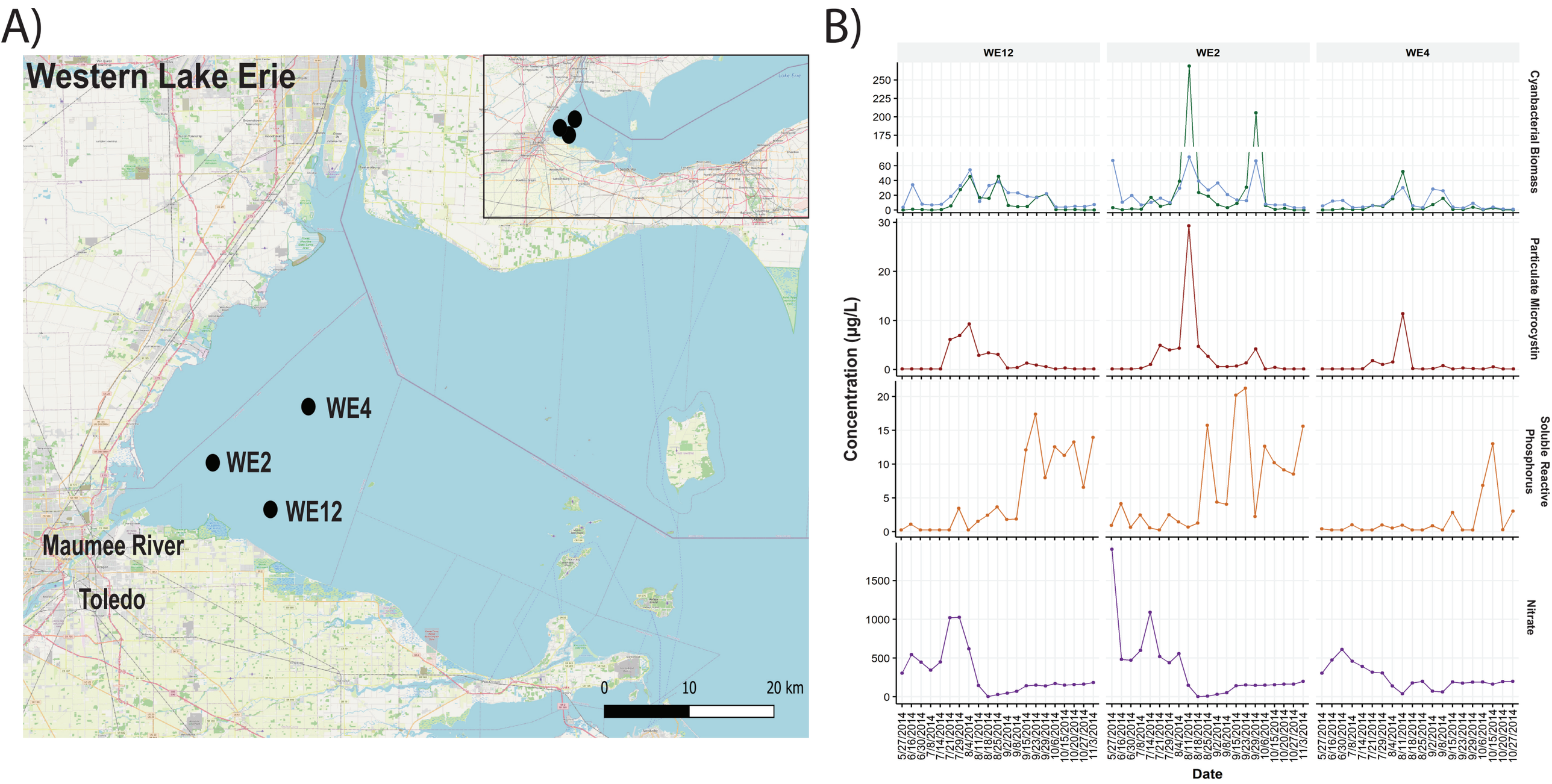
Figure 3: Comparison of PKS-M BGCs identified in WLE MAGs to the known microginins encoding BGC from Microcystis aeruginosa LEGE 91341. A) Both PKS-M gene clusters are incomplete/fragmented compared to the complete cluster where PKS-M contains conserved additional synthesis and hypothetical protein encoding genes while PKS-M 2 lacks these genes and contains a putative nuclease. Both PKS-M clusters contain a fragmented hybrid NRPS-PKS encoding gene found in the microginin encoding pathway. B) Read mapping from a sample collected on 4-August at WE2 against the complete microginin biosynthesis pathway in LEGE 91341. Coverage is observed throughout with notable variation in sequence identity and evenness.

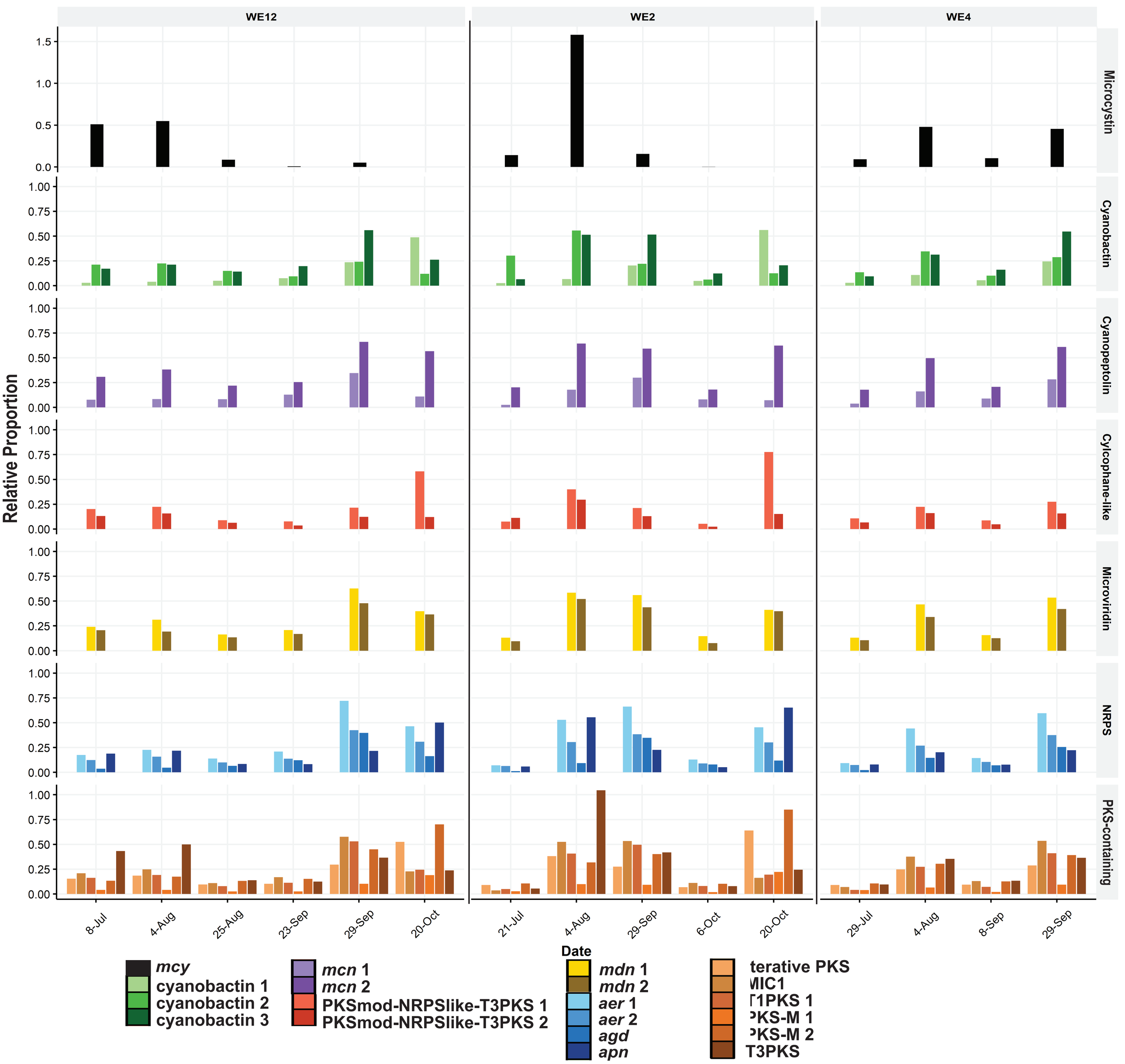
Figure 4: Gene schematics for select cryptically annotated PKS-containing gene clusters. A) The iterative PKS gene cluster contains transport genes, a putative enediyne biosynthesis gene, a thioesterase, 2/3 genes of unknown function that are part of the enediyne biosynthesis cassette, and a putative nucleases B) The T3PKS cluster contains several transport genes and a cytochrome B5 binding gene. C) The MIC cluster contains several core biosynthesis and transport genes, and a tryptophan halogenase.

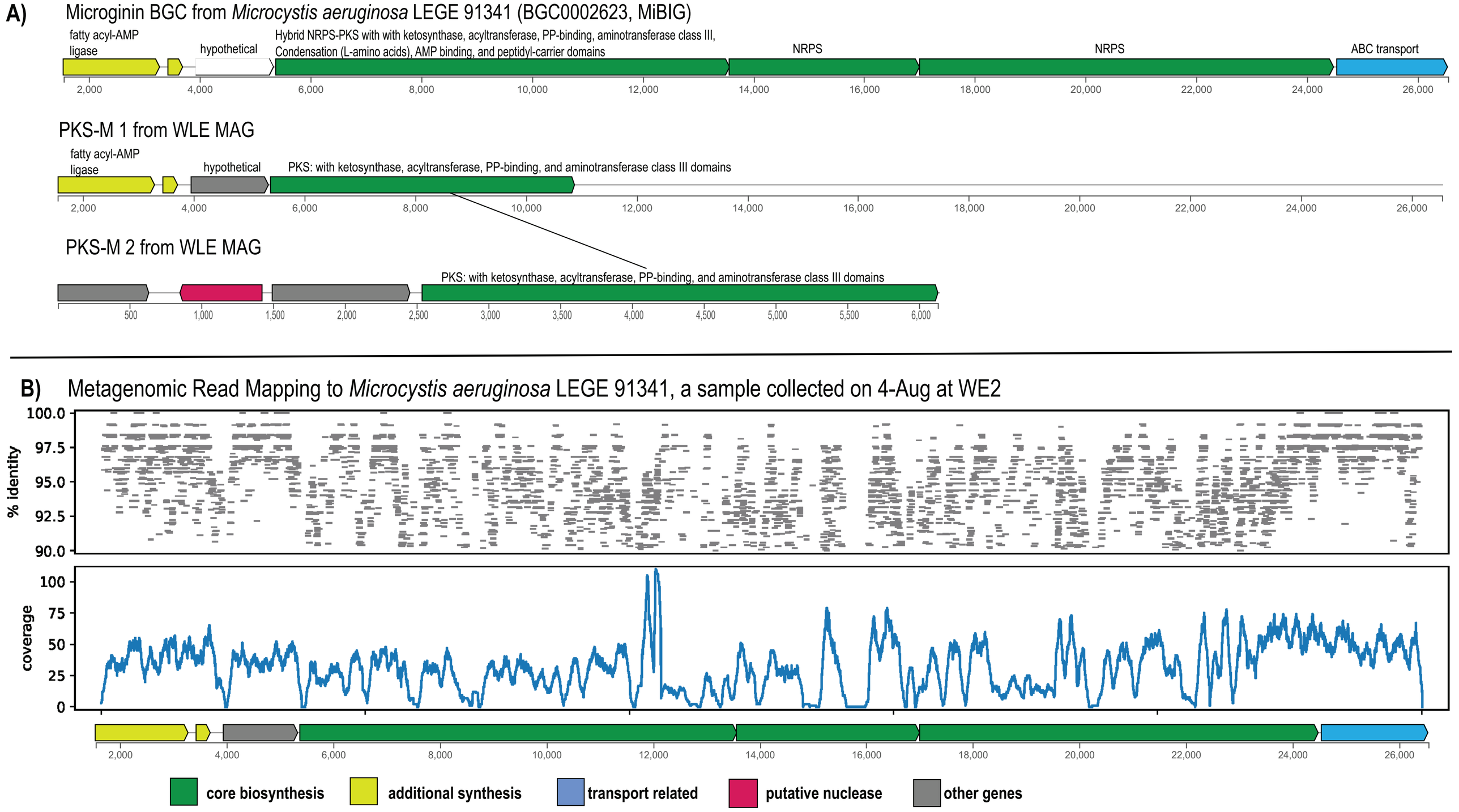
Figure 5: Phylogenetic tree consisting of protein sequences from the putatively identified PKSE from the WLE MAGs, putatively identified PKSEs from previous studies (45, 73), and PKSEs from known biosynthesis pathways that synthesize enediyne containing molecules. The PKSE from the WLE MAGs is colored red, the putative PKSEs identified in other cyanobacteria are colored in green, and a black circle indicates a known synthesized product associated with the PKSE enzyme.

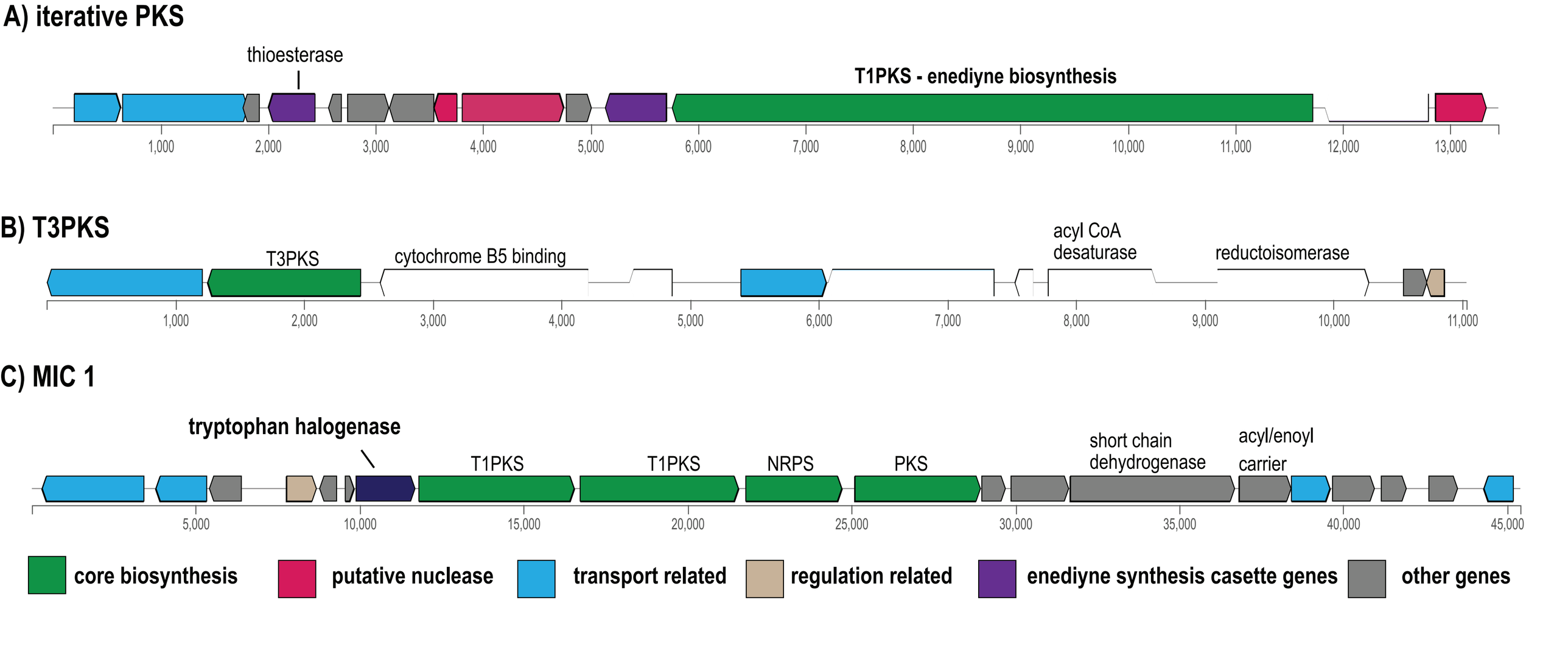
Figure 6: Relative abundance of transcripts from BGCs throughout the 2014 Bloom separated by sampling station. Relative transcript abundance was calculated by determining the number of

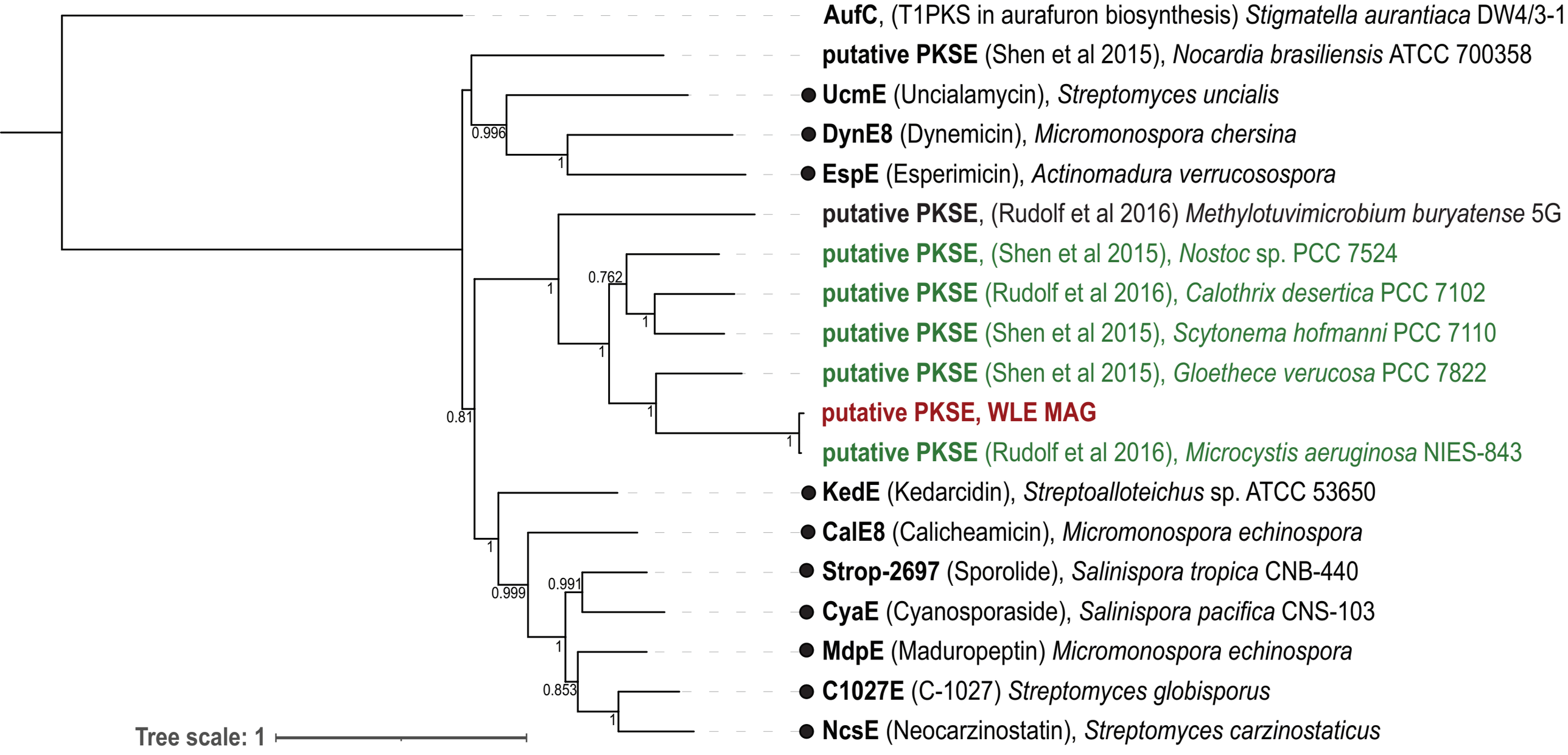
926 reads mapped to a BGC via specific cut off parameters, normalized to the number of reads
927 mapped to an entire reference *Microcystis* genome to determine "expression effort" (see
928 Experimental Procedure).

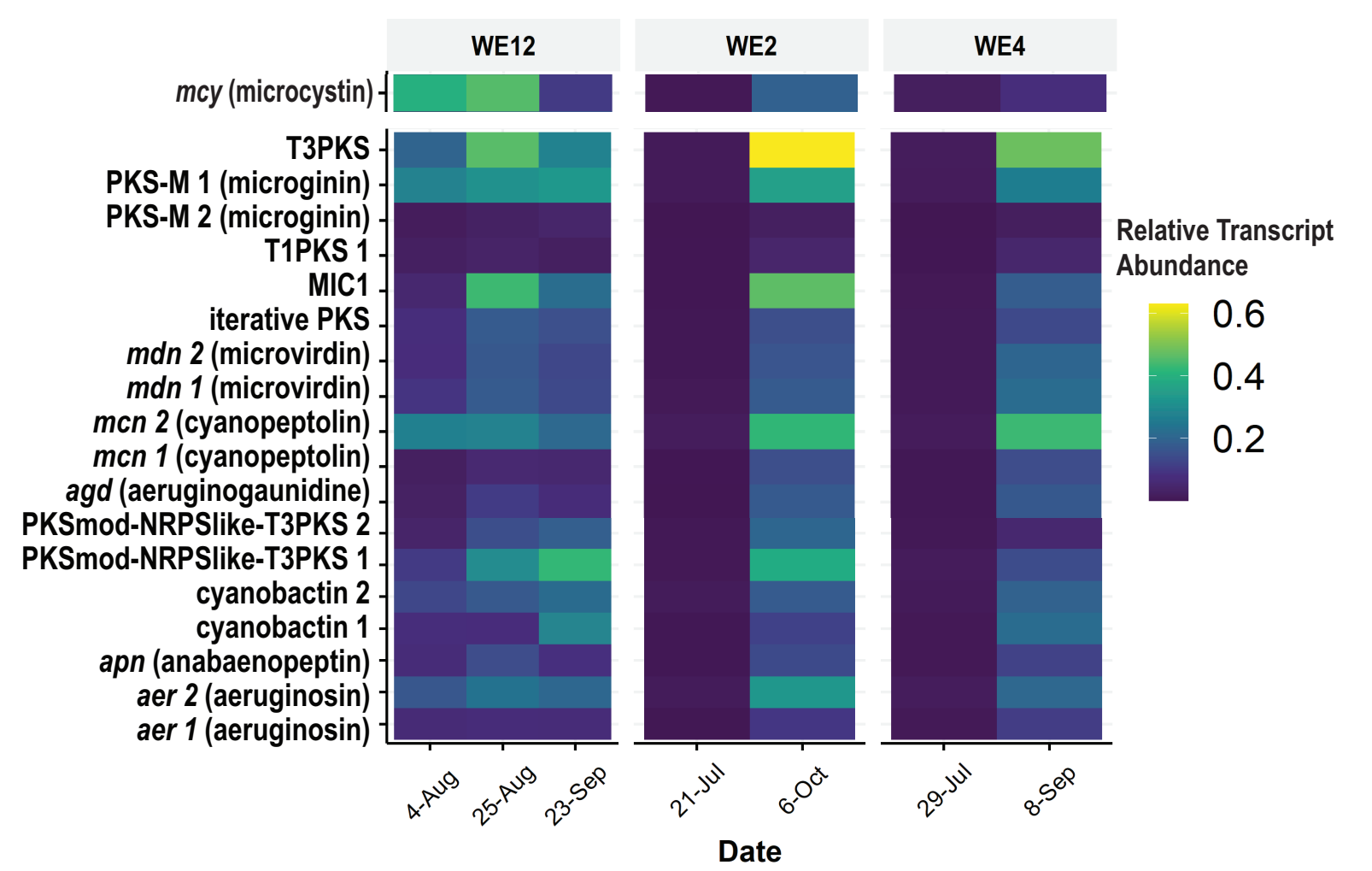

A)
Western Lake Erie
WE4
WE2
WE12
Maumee River
Toledo
0
10
20 km
B)
WE12
WE2
WE4
Cyanbacterial Biomass
Particulate Microcystin
Soluble Reactive Phosphorus
Nitrate
Concentration (µg/L)
Date


WE12
WE2
WE4
Microcystin
Cyanobactin
Cyanopeptolin
Cylcophane-like
Microviridin
NRPS
PKS-containing
Relative Proportion
0.0
0.5
1.0
1.5
0.00
0.25
0.50
0.75
1.00
8-Jul
4-Aug
25-Aug
23-Sep
29-Sep
20-Oct
21-Jul
4-Aug
29-Sep
6-Oct
20-Oct
29-Jul
4-Aug
8-Sep
29-Sep
Date
mcy
cyanobactin 1
cyanobactin 2
cyanobactin 3
mcn 1
mcn 2
PKSmod-NRPSlike-T3PKS 1
PKSmod-NRPSlike-T3PKS 2
mdn 1
mdn 2
aer 1
aer 2
agd
apn
terative PKS
MIC1
Γ1PKS 1
PKS-M 1
PKS-M 2
T3PKS

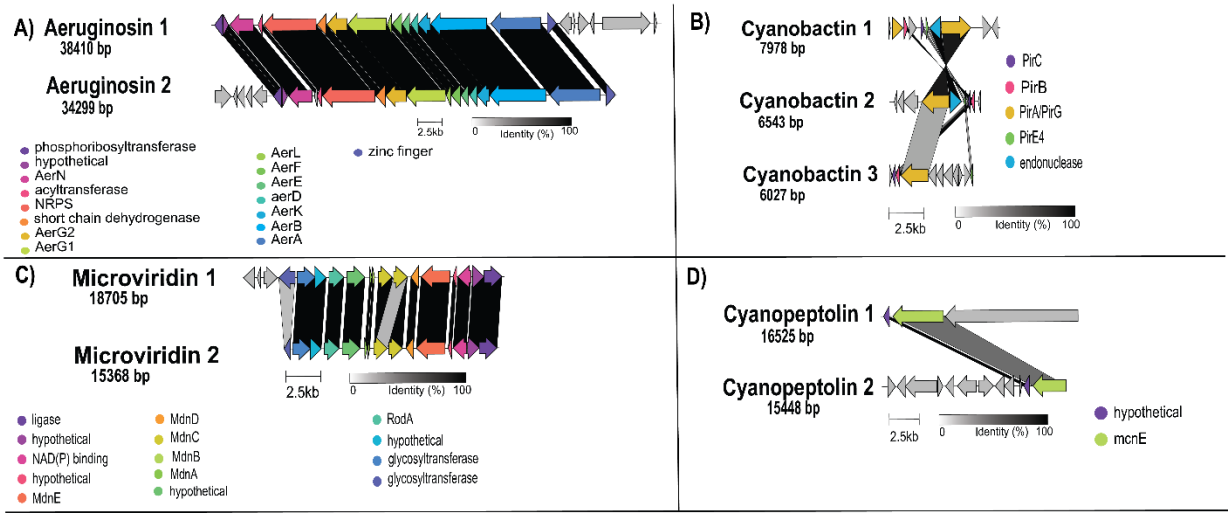

A)
Microginin BGC from Microcystis aeruginosa LEGE 91341 (BGC0002623, MiBIG)
fatty acyl-AMP ligase
hypothetical
Hybrid NRPS-PKS with with ketosynthase, acyltransferase, PP-binding, aminotransferase class III, Condensation (L-amino acids), AMP binding, and peptidyl-carrier domains
NRPS
NRPS
ABC transport
2,000 4,000 6,000 8,000 10,000 12,000 14,000 16,000 18,000 20,000 22,000 24,000 26,000
PKS-M 1 from WLE MAG
fatty acyl-AMP ligase
hypothetical
PKS: with ketosynthase, acyltransferase, PP-binding, and aminotransferase class III domains
2,000 4,000 6,000 8,000 10,000 12,000 14,000 16,000 18,000 20,000 22,000 24,000 26,000
PKS-M 2 from WLE MAG
PKS: with ketosynthase, acyltransferase, PP-binding, and aminotransferase class III domains
500 1,000 1,500 2,000 2,500 3,000 3,500 4,000 4,500 5,000 5,500 6,000
B) Metagenomic Read Mapping to Microcystis aeruginosa LEGE 91341, a sample collected on 4-Aug at WE2
% identity
100.0 97.5 95.0 92.5 90.0
coverage
100 75 50 25 0
2,000 4,000 6,000 8,000 10,000 12,000 14,000 16,000 18,000 20,000 22,000 24,000 26,000
core biosynthesis
additional synthesis
transport related
putative nuclease
other genes


A) iterative PKS
thioesterase
T1PKS - enediyne biosynthesis
1,000
2,000
3,000
4,000
5,000
6,000
7,000
8,000
9,000
10,000
11,000
12,000
13,000
B) T3PKS
T3PKS
cytochrome B5 binding
acyl CoA desaturase
reductoisomerase
1,000
2,000
3,000
4,000
5,000
6,000
7,000
8,000
9,000
10,000
11,000
C) MIC 1
tryptophan halogenase
T1PKS
T1PKS
NRPS
PKS
short chain dehydrogenase
acyl/enoyl carrier
5,000
10,000
15,000
20,000
25,000
30,000
35,000
40,000
45,000
core biosynthesis
putative nuclease
transport related
regulation related
enediyne synthesis casette genes
other genes

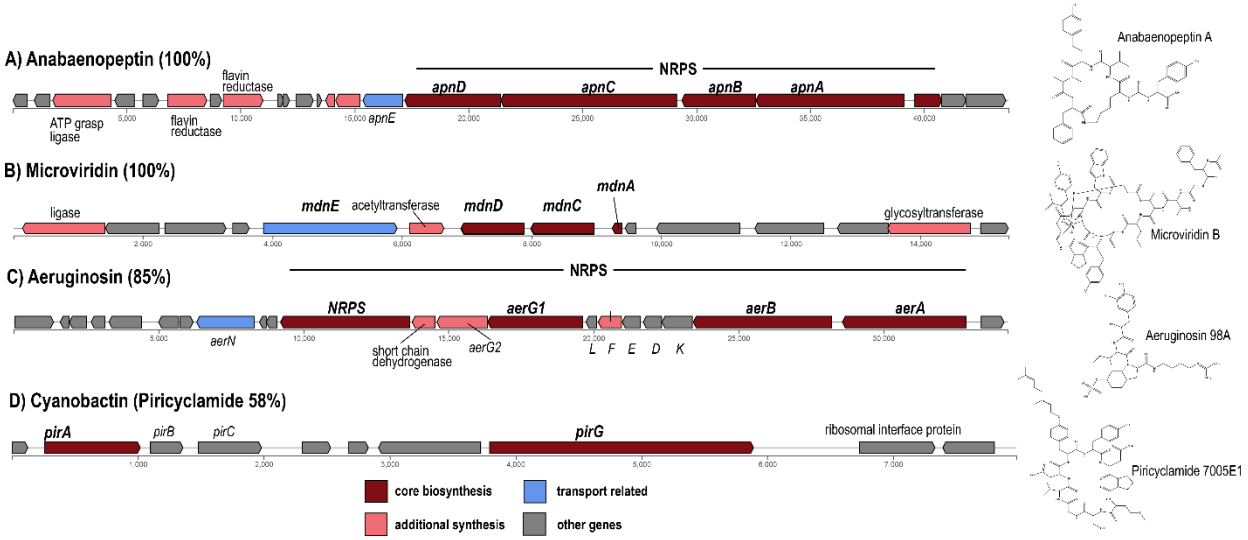

AufC, (T1PKS in aurafuron biosynthesis) Stigmatella aurantiaca DW4/3-1
putative PKSE (Shen et al 2015), Nocardia brasiliensis ATCC 700358
UcmE (Uncialamycin), Streptomyces uncialis
DynE8 (Dynemicin), Micromonospora chersina
EspE (Esperimicin), Actinomadura verrucosospora
putative PKSE, (Rudolf et al 2016) Methylotuvimicrobium buryatense 5G
putative PKSE, (Shen et al 2015), Nostoc sp. PCC 7524
putative PKSE (Rudolf et al 2016), Calothrix desertica PCC 7102
putative PKSE (Shen et al 2015), Scytonema hofmanni PCC 7110
putative PKSE (Shen et al 2015), Gloethece verucosa PCC 7822
putative PKSE, WLE MAG
putative PKSE (Rudolf et al 2016), Microcystis aeruginosa NIES-843
KedE (Kedarcidin), Streptoalloteichus sp. ATCC 53650
CalE8 (Calicheamicin), Micromonospora echinospora
Strop-2697 (Sporolide), Salinispora tropica CNB-440
CyaE (Cyanosporaside), Salinispora pacifica CNS-103
MdpE (Maduropeptin) Micromonospora echinospora
C1027E (C-1027) Streptomyces globisporus
NcsE (Neocarzinostatin), Streptomyces carzinostaticus
0.996
1
0.762
1
1
1
0.81
1
1
0.999
0.991
1
0.853
1
Tree scale: 1


WE12
WE2
WE4
mcy (microcystin)
T3PKS
PKS-M 1 (microginin)
PKS-M 2 (microginin)
T1PKS 1
MIC1
iterative PKS
mdn 2 (microvirdin)
mdn 1 (microvirdin)
mcn 2 (cyanopeptolin)
mcn 1 (cyanopeptolin)
agd (aeruginogaunidine)
PKSmod-NRPSlike-T3PKS 2
PKSmod-NRPSlike-T3PKS 1
cyanobactin 2
cyanobactin 1
apn (anabaenopeptin)
aer 2 (aeruginosin)
aer 1 (aeruginosin)
4-Aug
25-Aug
23-Sep
21-Jul
6-Oct
29-Jul
8-Sep
Date
Relative Transcript Abundance
0.6
0.4
0.2

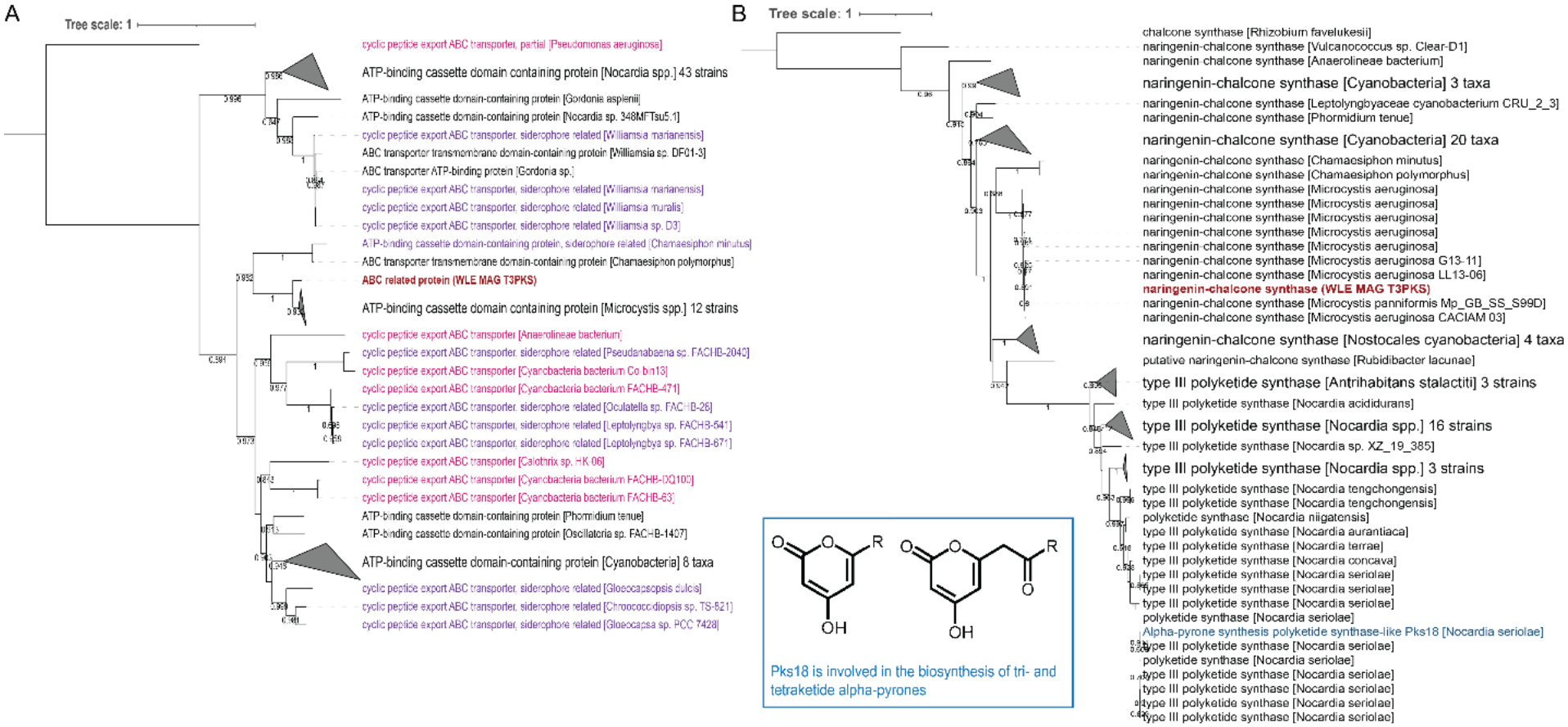
SUPPLEMENTALS



Supplemental Figure 1: Comparison of gene order, orientation, and sequence similarity between distinct BGCs from different *Microcystis* MAGs. Clinker was used to visually compare similarities and differences of the A) aeruginosin, B) cyanobactin, C) microviridin B, and D) micropeptin BGC classes. Homologous genes are the same color and percent similarity between genes is shown via a bar connecting homologous genes.



Supplemental Figure 2: Gene Schematics for identified BGC clusters including A) anabaenopeptin, B) microviridin B, C) aeruginosin, and D) a cyanobactin with 58% similarity to piricyclamide. Gene schematics were rendered through AntiSMASH v.6. Examples of chemical congeners that may be synthesized by these BGCs are shown. Percent similarity represents the percentage of genes within the closest known compound on the MiBIG database that have a significant BLAST hit


A
Tree scale: 1
cyclic peptide export ABC transporter, partial [Pseudomonas aeruginosa]
ATP-binding cassette domain containing protein [Nocardia spp.] 43 strains
ATP-binding cassette domain-containing protein [Gordonia asplenii]
ATP-binding cassette domain-containing protein [Nocardia sp. 348MFTsu5.1]
cyclic peptide export ABC transporter, siderophore related [Williamsia marianensis]
ABC transporter transmembrane domain-containing protein [Williamsia sp. DF01-3]
ABC transporter ATP-binding protein [Gordonia sp.]
cyclic peptide export ABC transporter, siderophore related [Williamsia marianensis]
cyclic peptide export ABC transporter, siderophore related [Williamsia muralis]
cyclic peptide export ABC transporter, siderophore related [Williamsia sp. D3]
ATP-binding cassette domain-containing protein, siderophore related [Chamaesiphon minutus]
ABC transporter transmembrane domain-containing protein [Chamaesiphon polymorphus]
ABC related protein (WLE MAG T3PKS)
ATP-binding cassette domain containing protein [Microcystis spp.] 12 strains
cyclic peptide export ABC transporter [Anaerolineae bacterium]
cyclic peptide export ABC transporter, siderophore related [Pseudanabaena sp. FACHB-2040]
cyclic peptide export ABC transporter [Cyanobacteria bacterium Co bin13]
cyclic peptide export ABC transporter [Cyanobacteria bacterium FACHB-471]
cyclic peptide export ABC transporter, siderophore related [Oculatella sp. FACHB-26]
cyclic peptide export ABC transporter, siderophore related [Leptolyngbya sp. FACHB-541]
cyclic peptide export ABC transporter, siderophore related [Leptolyngbya sp. FACHB-671]
cyclic peptide export ABC transporter [Calothrix sp. HK-06]
cyclic peptide export ABC transporter [Cyanobacteria bacterium FACHB-DQ100]
cyclic peptide export ABC transporter [Cyanobacteria bacterium FACHB-63]
ATP-binding cassette domain-containing protein [Phormidium tenue]
ATP-binding cassette domain-containing protein [Oscillatoria sp. FACHB-1407]
ATP-binding cassette domain-containing protein [Cyanobacteria] 8 taxa
cyclic peptide export ABC transporter, siderophore related [Gloeocapsopsis dulcis]
cyclic peptide export ABC transporter, siderophore related [Chroococcidiopsis sp. TS-821]
cyclic peptide export ABC transporter, siderophore related [Gloeocapsa sp. PCC 7428]
B
Tree scale: 1
chalcone synthase [Rhizobium favelukesii]
naringenin-chalcone synthase [Vulcanococcus sp. Clear-D1]
naringenin-chalcone synthase [Anaerolineae bacterium]
naringenin-chalcone synthase [Cyanobacteria] 3 taxa
naringenin-chalcone synthase [Leptolyngbyaceae cyanobacterium CRU_2_3]
naringenin-chalcone synthase [Phormidium tenue]
naringenin-chalcone synthase [Cyanobacteria] 20 taxa
naringenin-chalcone synthase [Chamaesiphon minutus]
naringenin-chalcone synthase [Chamaesiphon polymorphus]
naringenin-chalcone synthase [Microcystis aeruginosa]
naringenin-chalcone synthase [Microcystis aeruginosa]
naringenin-chalcone synthase [Microcystis aeruginosa]
naringenin-chalcone synthase [Microcystis aeruginosa]
naringenin-chalcone synthase [Microcystis aeruginosa]
naringenin-chalcone synthase [Microcystis aeruginosa G13-11]
naringenin-chalcone synthase [Microcystis aeruginosa LL13-06]
naringenin-chalcone synthase (WLE MAG T3PKS)
naringenin-chalcone synthase [Microcystis panniformis Mp_GB_SS_S99D]
naringenin-chalcone synthase [Microcystis aeruginosa CACIAM 03]
naringenin-chalcone synthase [Nostocales cyanobacteria] 4 taxa
putative naringenin-chalcone synthase [Rubidibacter lacunae]
type III polyketide synthase [Antrihabitans stalactiti] 3 strains
type III polyketide synthase [Nocardia acididurans]
type III polyketide synthase [Nocardia spp.] 16 strains
type III polyketide synthase [Nocardia sp. XZ_19_385]
type III polyketide synthase [Nocardia spp.] 3 strains
type III polyketide synthase [Nocardia tengchongensis]
type III polyketide synthase [Nocardia tengchongensis]
polyketide synthase [Nocardia niigatensis]
type III polyketide synthase [Nocardia aurantiaca]
type III polyketide synthase [Nocardia terrae]
type III polyketide synthase [Nocardia concava]
type III polyketide synthase [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
polyketide synthase [Nocardia seriolae]
Alpha-pyrone synthesis polyketide synthase-like Pks18 [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
polyketide synthase [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
type III polyketide synthase [Nocardia seriolae]
O
O
R
OH
O
O
R
O
OH
Pks18 is involved in the biosynthesis of tri- and tetraketide alpha-pyrones

Supplemental Figure 3: Phylogenies for A) a transport protein and B) core biosynthesis naringenin-chalcone synthase from the T3PKS gene cluster. The transport protein has high homology to proteins involved in siderophore transport while the naringenin chalcone synthase is similar to cyanobacterial naringenin chalcone synthases. Bootstrap values are included at nodes where values are greater than 0.5. Proteins that share the same functional annotation of interest are colored. Dark red and bolded protein sequences are from the W. Lake Erie cyanoHAB as identified through antiSMASH.

Table S1: PKS Cluster Similarities to known compounds as identified via antiSMASH.

Cluster ID	Known Compound Cluster Hit	Similarity (%)
T1Iterative PKS	Calicheamicin	4
T3PKS	NA	NA
MIC 1	Phenalamide	40

Table S2: Annotations for genes identified in the MIC BGC. NRPS/PKS domains are provided when applicable as well as proposed functions. DCL = Condensation linking L-amino acids to a peptide ending with a D-amino acid. LCL = Condensation linking L-amino acid to a peptide ending with an L-amino acid as described via antiSMASH output.

Gene	Length (bp)	Function	Domains	Proposed Function
1	3099	Transport-related	---	Permease, efflux RND
2	1545	Transport-related	---	Periplasmic adapter, efflux RND
3	975	Other/Hypothetical	---	Unknown/Hypothetical
4	903	Regulatory	---	Transcriptional Regulator
5	519	Other/Hypothetical	---	Unknown/Hypothetical
6	261	Additional Biosynthesis	Acyl Carrier	Acyl Carrier Protein
7	1797	Additional Biosynthesis	---	Halogenase
8	4734	Core - T1PKS	Hybrid Ketosynthase, Acyltransferase, Ketoreductase, PP-binding	T1PKS
9	4824	Core - T1PKS	Hybrid Ketosynthase, Acyltransferase, PP-binding, Condensation DCL, PP-binding	T1PKS
10	2925	Core - NRPS	AMP-binding, Condensation LCL	NRPS
11	3822	Core - PKS	Hybrid Ketosynthase, Acyltransferase, PP-binding	PKS
12	720	Other/Hypothetical	---	Unknown/Hypothetical
13	1758	Additional Biosynthesis	Ketoreductase	Short chain Dehydrogenase/Reductase

14	5004	Additional Biosynthesis	Acyltransferase	Short chain Dehydrogenase/Reductase
15	1575	Additional Biosynthesis	---	Enoyl reductase
16	1173	Transport-related	---	MFS transporter
17	1278	Other	---	FAD-dependent oxidoreductase
18	765	Other/Hypothetical	---	Unknown/Hypothetical
19	876	Other/Hypothetical	---	Unknown/Hypothetical
20	894	Transport-related	---	ABC transporter

Table S3: Correlations of BGC relative transcript abundance and available abiotic variables

Variable	R-squared	p-value
NO3	-0.418	0.59
NH4	0.188	0.41
Temperature	-0.514	0.029
pH	-0.129	0.58
SRP	0.251	0.27
TP	0.114	0.62
PC	-0.0536	0.82
chlA	-0.0603	0.8

Aeruginosins (*aer*)		
Variable	R-squared	p-value
NO3	-0.48	0.028
NH4	0.0988	0.67
Temperature	-0.375	0.1249
pH	0.0413	0.859
SRP	0.364	0.1
TP	0.449	0.04
PC	0.252	0.27
chlA	0.162	0.48

PKSmod-NRPSlike-T3PKS

Variable	R-squared	p-value
NO3	-0.551	0.041
NH4	0.0802	0.79
Temperature	-0.598	0.04
pH	-0.238	0.41
SRP	0.769	0.0013
TP	0.601	0.023
PC	0.129	0.66
chlA	-0.0753	0.8

Cyanopeptolin *(mcn)*

Variable	R-squared	p-value
NO3	-0.434	0.049
NH4	0.178	0.44
Temperature	-0.412	0.089
pH	-0.00607	0.98
SRP	0.216	0.35
TP	0.227	0.32
PC	0.0991	0.669
chlA	0.0839	0.72

Microviridin (*mdn)*

Variable	R-squared	p-value
NO3	-0.749	0.002
NH4	0.00647	0.98
Temperature	-0.423	0.17
pH	0.144	0.62
SRP	0.358	0.21
TP	0.537	0.048
PC	0.312	0.28
chlA	0.218	0.45

Table S4: Correlations of BGC relative transcript abundance and relative abundance of select organisms

BGC Class	Organism Genus	R-squared	p-value
Cyanobactin	*UNCLASSIFIED_GEMMATIMONADACEAE_GENUS*	0.532002	0.023055
Cyanobactin	*HERBASPIRILLUM*	-0.64425	0.023743
Cyanobactin	*WOLBACHIA*	-0.52319	0.025879
Cyanobactin	*UNCLASSIFIED_CYTOPHAGALES_GENUS*	0.444163	0.04368

Cyanobactin	*RHODOFERAX*	0.418508	0.059008
Cyanobactin	*GEMMATIMONAS*	-0.41654	0.060332
Cyanobactin	*ROSEOMONAS*	-0.4043	0.069096
Cyanobactin	*UNCLASSIFIED_VERRUCOMICROBIACEAE_GENUS*	-0.45766	0.086274
Cyanobactin	*UNCLASSIFIED_ARMATIMONADETES_GENUS*	-0.378	0.09112
Cyanobactin	*ACINETOBACTER*	-0.50547	0.093656
Cyanobactin	*NOVIHERBASPIRILLUM*	-0.48984	0.105981
Cyanobactin	*UNCLASSIFIED_GEMMATIMONADETES_GENUS*	0.352644	0.116898
Cyanobactin	*UNCLASSIFIED_SYNECHOCOCCALES_GENUS*	-0.41704	0.121975
Cyanobactin	*BREVUNDIMONAS*	-0.34195	0.129211
Cyanobactin	*PIRELLULA*	-0.35852	0.144021
Cyanobactin	*BDELLOVIBRIO*	-0.35723	0.145583
Cyanobactin	*UNCLASSIFIED_PHYCISPHAERAE_GENUS*	-0.35563	0.147523
Cyanobactin	*SKELETONEMA*	0.617171	0.03252
Cyanobactin	*APOSTICHOPUS*	-0.51817	0.08438
Cyanobactin	*PHYTOPHTHORA*	0.372823	0.12757
Cyanobactin	*EURYTEMORA*	0.32275	0.153585
aer (aeruginosin)	*UNCLASSIFIED_BATHYARCHAEOTA_GENUS*	-0.54321	0.010932
aer (aeruginosin)	*UNCLASSIFIED_EURYARCHAEOTA_GENUS*	-0.5217	0.081919
aer (aeruginosin)	*WOLBACHIA*	-0.68302	0.001783
aer (aeruginosin)	*UNCLASSIFIED_VERRUCOMICROBIACEAE_GENUS*	-0.71213	0.002897
aer (aeruginosin)	*GEMMATIMONAS*	-0.56891	0.007115
aer (aeruginosin)	*ACINETOBACTER*	-0.72704	0.007383
aer (aeruginosin)	*PAUCIBACTER*	-0.71693	0.008691
aer (aeruginosin)	*UNCLASSIFIED_ARMATIMONADETES_GENUS*	-0.54466	0.010681
aer (aeruginosin)	*UNCLASSIFIED_SYNECHOCOCCALES_GENUS*	-0.62251	0.013194
aer (aeruginosin)	*UNCLASSIFIED_CHLOROFLEXI_GENUS*	-0.50549	0.019408
aer (aeruginosin)	*BRYOBACTER*	-0.58735	0.021321
aer (aeruginosin)	*UNCLASSIFIED_VERRUCOMICROBIALES_GENUS*	-0.49482	0.02258
aer (aeruginosin)	*UNCLASSIFIED_VERRUCOMICROBIA_GENUS*	-0.49006	0.024122
aer (aeruginosin)	*ROSEOMONAS*	-0.48363	0.026336
aer (aeruginosin)	*POLYNUCLEOBACTER*	-0.56913	0.02681
aer (aeruginosin)	*METHYLOBACTERIUM*	-0.47444	0.029776
aer (aeruginosin)	*UNCLASSIFIED_ACTINOMYCETIA_GENUS*	-0.47312	0.0303
aer (aeruginosin)	*RUNELLA*	-0.62123	0.031077
aer (aeruginosin)	*UNCLASSIFIED_GEMMATIMONADACEAE_GENUS*	0.497769	0.035545
aer (aeruginosin)	*UNCLASSIFIED_RHODOBACTERACEAE_GENUS*	-0.49136	0.03837
aer (aeruginosin)	*SYNECHOCOCCUS*	-0.53619	0.03936
aer (aeruginosin)	*LIMNOHABITANS*	-0.43737	0.047397
aer (aeruginosin)	*UNCLASSIFIED_OPITUTAE_GENUS*	-0.51859	0.047632
aer (aeruginosin)	*GEMMOBACTER*	-0.43359	0.049571
aer (aeruginosin)	*BREVUNDIMONAS*	-0.43254	0.050189

aer (aeruginosin)	*UNCLASSIFIED_FLAVOBACTERIACEAE_GENUS*	-0.50918	0.052554
aer (aeruginosin)	*HERBASPIRILLUM*	-0.5664	0.054859
aer (aeruginosin)	*PLANKTOPHILA*	-0.55568	0.060676
aer (aeruginosin)	*UNCLASSIFIED_PELAGIBACTERALES_GENUS*	-0.54773	0.065257
aer (aeruginosin)	*RHODOFERAX*	0.408822	0.065753
aer (aeruginosin)	*BELNAPIA*	-0.48663	0.065843
aer (aeruginosin)	*UNCLASSIFIED_BRYOBACTERALES_GENUS*	-0.46307	0.082158
aer (aeruginosin)	*TABRIZICOLA*	-0.37622	0.092781
aer (aeruginosin)	*OLIGOFLEXUS*	0.504783	0.094177
aer (aeruginosin)	*CHRYSEOTALEA*	0.404708	0.095728
aer (aeruginosin)	*UNCLASSIFIED_PHYCISPHAERAE_GENUS*	-0.39564	0.104123
aer (aeruginosin)	*SICCIRUBRICOCCUS*	-0.43209	0.10774
aer (aeruginosin)	*AQUINCOLA*	-0.36081	0.108089
aer (aeruginosin)	*UNCLASSIFIED_CYTOPHAGALES_GENUS*	0.360119	0.108811
aer (aeruginosin)	*ROSEOCOCCUS*	-0.42753	0.111925
aer (aeruginosin)	*BOSEA*	0.381182	0.118589
aer (aeruginosin)	*ALGORIPHAGUS*	-0.47269	0.120699
aer (aeruginosin)	*PARACRAUROCOCCUS*	-0.41354	0.125466
aer (aeruginosin)	*DERXIA*	-0.40095	0.138566
aer (aeruginosin)	*NOVIHERBASPIRILLUM*	-0.44759	0.144545
aer (aeruginosin)	*CHITINOPHAGA*	-0.32747	0.147314
aer (aeruginosin)	*PIRELLULA*	-0.35561	0.147549
aer (aeruginosin)	*APOSTICHOPUS*	-0.70705	0.010128
aer (aeruginosin)	*SKELETONEMA*	0.630377	0.027992
aer (aeruginosin)	*TETRAHYMENA*	0.562994	0.028878
aer (aeruginosin)	*PARAMECIUM*	0.561877	0.029267
aer (aeruginosin)	*TIGRIOPUS*	-0.46945	0.031789
aer (aeruginosin)	*PARAMURICEA*	-0.46756	0.032577
aer (aeruginosin)	*ICHTHYOPHTHIRIUS*	0.552709	0.032614
aer (aeruginosin)	*PSEUDOCOHNILEMBUS*	0.549228	0.033957
aer (aeruginosin)	*NYMPHAEA*	0.544111	0.036006
aer (aeruginosin)	*LEPEOPHTHEIRUS*	-0.46573	0.051427
aer (aeruginosin)	*CRASPEDOSTAUROS*	0.567259	0.054411
aer (aeruginosin)	*ATTHEYA*	-0.38152	0.087909
aer (aeruginosin)	*STYLONYCHIA*	-0.50671	0.092719
aer (aeruginosin)	*CHAETOCEROS*	-0.50033	0.097601
aer (aeruginosin)	*NOCCAEA*	0.440112	0.100646
aer (aeruginosin)	*DITYLUM*	-0.36555	0.103194
aer (aeruginosin)	*UNCLASSIFIED_SIPHOVIRIDAE_GENUS*	-0.58451	0.022115
aer (aeruginosin)	*CYANOPHAGE*	-0.44199	0.044845
aer (aeruginosin)	*UNCLASSIFIED_CAUDOVIRALES_GENUS*	-0.42915	0.052221
aer (aeruginosin)	*UNCLASSIFIED_PODOVIRIDAE_GENUS*	-0.39876	0.101175

mcn (cyanopeptolin)	*UNCLASSIFIED_BATHYARCHAEOTA_GENUS*	-0.38504	0.084774
mcn (cyanopeptolin)	*UNCLASSIFIED_GEMMATIMONADACEAE_GENUS*	0.625287	0.00552
mcn (cyanopeptolin)	*WOLBACHIA*	-0.62255	0.005792
mcn (cyanopeptolin)	*HERBASPIRILLUM*	-0.69002	0.013013
mcn (cyanopeptolin)	*GEMMATIMONAS*	-0.49778	0.021662
mcn (cyanopeptolin)	*ACINETOBACTER*	-0.64436	0.023713
mcn (cyanopeptolin)	*RHODOFERAX*	0.488081	0.024786
mcn (cyanopeptolin)	*PAUCIBACTER*	-0.61405	0.033661
mcn (cyanopeptolin)	*UNCLASSIFIED_VERRUCOMICROBIACEAE_GENUS*	-0.54704	0.03482
mcn (cyanopeptolin)	*UNCLASSIFIED_ARMATIMONADETES_GENUS*	-0.45981	0.035975
mcn (cyanopeptolin)	*ROSEOMONAS*	-0.44897	0.04119
mcn (cyanopeptolin)	*POLYNUCLEOBACTER*	-0.51391	0.050033
mcn (cyanopeptolin)	*UNCLASSIFIED_CYTOPHAGALES_GENUS*	0.427287	0.053361
mcn (cyanopeptolin)	*PLANKTOPHILA*	-0.54341	0.067844
mcn (cyanopeptolin)	*UNCLASSIFIED_CHLOROFLEXI_GENUS*	-0.40392	0.06938
mcn (cyanopeptolin)	*BREVUNDIMONAS*	-0.4015	0.071227
mcn (cyanopeptolin)	*UNCLASSIFIED_PHYCISPHAERAE_GENUS*	-0.43177	0.073576
mcn (cyanopeptolin)	*UNCLASSIFIED_PELAGIBACTERALES_GENUS*	-0.52941	0.076714
mcn (cyanopeptolin)	*NOVIHERBASPIRILLUM*	-0.52646	0.078678
mcn (cyanopeptolin)	*BELNAPIA*	-0.45468	0.088609
mcn (cyanopeptolin)	*UNCLASSIFIED_SYNECHOCOCCALES_GENUS*	-0.45449	0.088754
mcn (cyanopeptolin)	*ROSEOCOCCUS*	-0.44649	0.095243
mcn (cyanopeptolin)	*UNCLASSIFIED_ACTINOMYCETIA_GENUS*	-0.36592	0.102817
mcn (cyanopeptolin)	*SICCIRUBRICOCCUS*	-0.43354	0.106437
mcn (cyanopeptolin)	*UNCLASSIFIED_OPITUTAE_GENUS*	-0.42291	0.116282
mcn (cyanopeptolin)	*PARACRAUROCOCCUS*	-0.41768	0.121342
mcn (cyanopeptolin)	*METHYLOBACTERIUM*	-0.33637	0.135985
mcn (cyanopeptolin)	*SKELETONEMA*	0.741023	0.005825
mcn (cyanopeptolin)	*APOSTICHOPUS*	-0.68998	0.01302
mcn (cyanopeptolin)	*TIGRIOPUS*	-0.37636	0.092651
mcn (cyanopeptolin)	*PARAMURICEA*	-0.36919	0.09955
mcn (cyanopeptolin)	*PSEUDO*	0.346343	0.124046
mcn (cyanopeptolin)	*CYANOPHAGE*	-0.34948	0.120445
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_BATHYARCHAEOTA_GENUS*	-0.56172	0.036579
PKSmod-NRPSlike-T3PKS	*WOLBACHIA*	-0.76316	0.003881
PKSmod-NRPSlike-T3PKS	*OLIGOFLEXUS*	0.827902	0.011155
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_SYNECHOCOCCALES_GENUS*	-0.73805	0.014805
PKSmod-NRPSlike-T3PKS	*PAUCIBACTER*	-0.80483	0.015971
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_BDELLOVIBRIONALES_GENUS*	0.710863	0.021193

PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_VERRUCOMICROBIALES_GENUS*	-0.6066	0.021449
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_VERRUCOMICROBIA_GENUS*	-0.60468	0.021977
PKSmod-NRPSlike-T3PKS	*METHYLOBACTERIUM*	-0.59006	0.026336
PKSmod-NRPSlike-T3PKS	*BOSEA*	0.628276	0.02868
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_XANTHOMONADALES_GENUS*	0.754947	0.030359
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_VERRUCOMICROBIACEAE_GENUS*	-0.6773	0.031427
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_ARMATIMONADETES_GENUS*	-0.56951	0.033514
PKSmod-NRPSlike-T3PKS	*GEMMATIMONAS*	-0.55262	0.040423
PKSmod-NRPSlike-T3PKS	*PARABURKHOLDERIA*	0.639625	0.046424
PKSmod-NRPSlike-T3PKS	*OHTAEKWANGIA*	0.629832	0.050985
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_HYPHOMICROBIALES_GENUS*	-0.53039	0.05104
PKSmod-NRPSlike-T3PKS	*ROSEOMONAS*	-0.50999	0.062449
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_PLANCTOMYCETACEAE_GENUS*	0.50495	0.065528
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_OPITUTAE_GENUS*	-0.57355	0.083004
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_CHLOROFLEXI_GENUS*	-0.46759	0.091806
PKSmod-NRPSlike-T3PKS	*SYNECHOCOCCUS*	-0.55893	0.093031
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_ACTINOMYCETIA_GENUS*	-0.46576	0.093257
PKSmod-NRPSlike-T3PKS	*ACINETOBACTER*	-0.62657	0.096445
PKSmod-NRPSlike-T3PKS	*RHODOPIRELLULA*	0.624078	0.098181
PKSmod-NRPSlike-T3PKS	*POLYNUCLEOBACTER*	-0.54784	0.10113
PKSmod-NRPSlike-T3PKS	*VAMPIROVIBRIO*	0.61026	0.108112
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_FLAVOBACTERIACEAE_GENUS*	-0.53047	0.114691
PKSmod-NRPSlike-T3PKS	*CYANOBIUM*	-0.51312	0.129322
PKSmod-NRPSlike-T3PKS	*LIMNOHABITANS*	-0.42267	0.132157
PKSmod-NRPSlike-T3PKS	*CHRYSEOTALEA*	0.447747	0.14439
PKSmod-NRPSlike-T3PKS	*TETRAHYMENA*	0.817487	0.00387
PKSmod-NRPSlike-T3PKS	*PARAMECIUM*	0.809702	0.004528
PKSmod-NRPSlike-T3PKS	*PSEUDOCOHNILEMBUS*	0.803325	0.005124

PKSmod-NRPSlike-T3PKS	*ICHTHYOPHTHIRIUS*	0.799409	0.005517
PKSmod-NRPSlike-T3PKS	*NYMPHAEA*	0.787085	0.006891
PKSmod-NRPSlike-T3PKS	*CRASPEDOSTAUROS*	0.853161	0.007069
PKSmod-NRPSlike-T3PKS	*CHAETOCEROS*	-0.78724	0.020399
PKSmod-NRPSlike-T3PKS	*ATTHEYA*	-0.58158	0.029142
PKSmod-NRPSlike-T3PKS	*DITYLUM*	-0.57856	0.030194
PKSmod-NRPSlike-T3PKS	*NOCCAEA*	0.673847	0.032641
PKSmod-NRPSlike-T3PKS	*PHAEODACTYLUM*	-0.55405	0.0398
PKSmod-NRPSlike-T3PKS	*ODONTELLA*	-0.5466	0.065931
PKSmod-NRPSlike-T3PKS	*GRAMMATOPHORA*	-0.45155	0.140597
PKSmod-NRPSlike-T3PKS	*PARAMURICEA*	-0.41046	0.144905
PKSmod-NRPSlike-T3PKS	*FISTULIFERA*	-0.4076	0.148004
PKSmod-NRPSlike-T3PKS	*TIGRIOPUS*	-0.40737	0.148253
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_SIPHOVIRIDAE_GENUS*	-0.5798	0.078941
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_PODOVIRIDAE_GENUS*	-0.525	0.079661
PKSmod-NRPSlike-T3PKS	*UNCLASSIFIED_CAUDOVIRALES_GENUS*	-0.43758	0.117642
mdn (microviridin)	*UNCLASSIFIED_BATHYARCHAEOTA_GENUS*	-0.54366	0.044485
mdn (microviridin)	*UNCLASSIFIED_EURYARCHAEOTA_GENUS*	-0.58147	0.130566
mdn (microviridin)	*WOLBACHIA*	-0.80994	0.001406
mdn (microviridin)	*UNCLASSIFIED_VERRUCOMICROBIACEAE_GENUS*	-0.82399	0.003375
mdn (microviridin)	*PAUCIBACTER*	-0.87028	0.00494
mdn (microviridin)	*GEMMATIMONAS*	-0.68815	0.006514
mdn (microviridin)	*ACINETOBACTER*	-0.84757	0.007873
mdn (microviridin)	*RHODOFERAX*	0.66903	0.008884
mdn (microviridin)	*UNCLASSIFIED_ARMATIMONADETES_GENUS*	-0.64982	0.011887
mdn (microviridin)	*UNCLASSIFIED_CHLOROFLEXI_GENUS*	-0.63376	0.014948
mdn (microviridin)	*POLYNUCLEOBACTER*	-0.71234	0.020805
mdn (microviridin)	*UNCLASSIFIED_SYNECHOCOCCALES_GENUS*	-0.70906	0.021674
mdn (microviridin)	*ROSEOMONAS*	-0.59088	0.026076
mdn (microviridin)	*HERBASPIRILLUM*	-0.75803	0.029302
mdn (microviridin)	*UNCLASSIFIED_ACTINOMYCETIA_GENUS*	-0.56311	0.036019
mdn (microviridin)	*UNCLASSIFIED_GEMMATIMONADACEAE_GENUS*	0.601612	0.038501
mdn (microviridin)	*LIMNOHABITANS*	-0.54686	0.043
mdn (microviridin)	*BRYOBACTER*	-0.64031	0.046116

mdn (microviridin)	*BREVUNDIMONAS*	-0.51833	0.057584
mdn (microviridin)	*PLANKTOPHILA*	-0.67286	0.067456
mdn (microviridin)	*UNCLASSIFIED_OPITUTAE_GENUS*	-0.59786	0.067924
mdn (microviridin)	*SYNECHOCOCCUS*	-0.59707	0.06838
mdn (microviridin)	*METHYLOBACTERIUM*	-0.48917	0.075866
mdn (microviridin)	*UNCLASSIFIED_VERRUCOMICROBIALES_GENUS*	-0.48596	0.078103
mdn (microviridin)	*UNCLASSIFIED_PELAGIBACTERALES_GENUS*	-0.64992	0.081073
mdn (microviridin)	*UNCLASSIFIED_VERRUCOMICROBIA_GENUS*	-0.47108	0.089076
mdn (microviridin)	*UNCLASSIFIED_FLAVOBACTERIACEAE_GENUS*	-0.55257	0.097624
mdn (microviridin)	*RUNELLA*	-0.6133	0.105882
mdn (microviridin)	*UNCLASSIFIED_CYTOPHAGALES_GENUS*	0.440241	0.11517
mdn (microviridin)	*BELNAPIA*	-0.52926	0.115681
mdn (microviridin)	*NOVIHERBASPIRILLUM*	-0.58682	0.126215
mdn (microviridin)	*VAMPIROVIBRIO*	-0.57153	0.13887
mdn (microviridin)	*UNCLASSIFIED_RHODOBACTERACEAE_GENUS*	-0.4489	0.143238
mdn (microviridin)	*APOSTICHOPUS*	-0.97738	2.84E-05
mdn (microviridin)	*SKELETONEMA*	0.831539	0.010493
mdn (microviridin)	*PARAMECIUM*	0.659467	0.038036
mdn (microviridin)	*TETRAHYMENA*	0.656704	0.039138
mdn (microviridin)	*TIGRIOPUS*	-0.55241	0.040515
mdn (microviridin)	*ICHTHYOPHTHIRIUS*	0.653081	0.040614
mdn (microviridin)	*NYMPHAEA*	0.651423	0.041302
mdn (microviridin)	*PSEUDOCOHNILEMBUS*	0.641096	0.045763
mdn (microviridin)	*NOCCAEA*	0.588531	0.073478
mdn (microviridin)	*PARAMURICEA*	-0.49062	0.07487
mdn (microviridin)	*LEPEOPHTHEIRUS*	-0.49739	0.099902
mdn (microviridin)	*EURYTEMORA*	0.442545	0.113059
mdn (microviridin)	*CYANOPHAGE*	-0.53873	0.046846
mdn (microviridin)	*UNCLASSIFIED_CAUDOVIRALES_GENUS*	-0.48634	0.077838
mdn (microviridin)	*UNCLASSIFIED_SIPHOVIRIDAE_GENUS*	-0.54786	0.101118
mdn (microviridin)	*UNCLASSIFIED_PODOVIRIDAE_GENUS*	-0.45023	0.141909

Table S5: Coverage and Completeness metrics for cryptic PKS containing cluster T3PKS, MIC 1 and Iterative PKS.

Station	Date	BGC Name	Bases with Coverage	Average Coverage	Status
WE12	7/8/2014	T3PKS	83.798	6	Complete
	7/8/2014	MIC 1	81.36	5.96	Complete
	7/8/2014	Iterative PKS	28.455	1.647	Absent
	8/4/2014	T3PKS	100	2431	Complete
	8/4/2014	MIC 1	100	28.85	Complete
	8/4/2014	Iterative PKS	100	923.7	Complete
	8/25/2014	T3PKS	100	110.9	Complete
	8/25/2014	MIC 1	100	239.2	Complete
	8/25/2014	Iterative PKS	100	83.51	Complete
	9/23/2014	T3PKS	87.26	64.67	Complete
	9/23/2014	MIC 1	100	339.4	Complete
	9/23/2014	Iterative PKS	57.053	62.49	Absent
	9/29/2014	T3PKS	100	188.5	Complete
	9/29/2014	MIC 1	100	827.3	Complete
	9/29/2014	Iterative PKS	100	170.6	Complete
	10/20/2014	T3PKS	79.603	21.62	Complete
	10/20/2014	MIC 1	98.53	28.11	Complete
	10/20/2014	Iterative PKS	100	57.804	Complete
WE2	7/21/2014	T3PKS	47.75	9.126	Absent
	7/21/2014	MIC 1	100	11.089	Complete
	7/21/2014	Iterative PKS	91.82	16.79	Complete
	8/4/2014	T3PKS	100	1217	Complete
	8/4/2014	MIC 1	100	1275	Complete
	8/4/2014	Iterative PKS	100	303.4	Complete
	9/29/2014	T3PKS	100	521.4	Complete
	9/29/2014	MIC 1	100	1599	Complete
	9/29/2014	Iterative PKS	100	312.9	Complete
	10/6/2014	T3PKS	85.303	167.1	Complete
	10/6/2014	MIC 1	100	49.09	Complete

	10/6/2014	Iterative PKS	100	214.4	Complete
	10/20/2014	T3PKS	93.98	81.19	Complete
	10/20/2014	MIC 1	100	53.84	Complete
	10/20/2014	Iterative PKS	100	230.2	Complete
WE4	7/29/2014	T3PKS	98.85	31.6	Complete
	7/29/2014	MIC 1	100	52.59	Complete
	7/29/2014	Iterative PKS	98.75	33.4	Complete
	8/4/2014	T3PKS	100	29.42	Complete
	8/4/2014	MIC 1	100	70.94	Complete
	8/4/2014	Iterative PKS	100	21.82	Complete
	9/8/2014	T3PKS	97.49	28.78	Complete
	9/8/2014	MIC 1	100	71.88	Complete
	9/8/2014	Iterative PKS	94.43	23.79	Complete
	9/29/2014	T3PKS	99.6	24.26	Complete
	9/29/2014	MIC 1	100	89.52	Complete
	9/29/2014	Iterative PKS	100	19.56	Complete

Table S6: Paired environmental variables for the 2014 cyanoHAB.

Station	Date	NO3 (mg/L)	NH4 (ug/L)	Temperature (°C)	Latitude	Longitidue	Distance to Maumee River (km)	Shore	pH	Particulate_Microcystins (ug/L)	Phycocyanin (ug/L)	CHLa (ug/L)	SRP (ug/L)	TP (ug/L)	DOC (uM)
WE12	8-Jul-14	0.34	16.16	23.5	NA	NA	17.44	Near	8.33	NA	0.12	6.98	0.2	34	214
WE12	4-Aug-14	0.62	3.25	24.6	41 42.157	83 15.781	16.35	Near	9.29	9.28	45.43	54.46	0.1	44.4	384
WE12	25-Aug-14	0.03	2.58	25	41 42.298	83 15.377	16.88	Near	9.29	3.1	45.72	38.28	3.68	98.1	415
WE12	23-Sep-14	0.15	3.14	17.9	41 42.196	83 15.499	16.72	Near	8.33	0.9	17.55	16.97	17.39	60.3	495
WE12	29-Sep-14	0.14	0.44	19.5	41 42.185	83 15.546	16.7	Near	8.74	0.6	21.87	22.25	7.98	69.2	398
WE12	20-Oct-14	0.16	5.18	13	41 42.269	83 15.663	16.54	Near	7.98	NA	0.54	5	13.27	45.2	419
WE2	21-Jul-14	0.52	2.93	23.5	41 45.873	83 19.849	12.23	Near	8.38	4.94	5.13	16.26	0.1	32.1	311
WE2	4-Aug-14	0.56	3.03	23.1	41 45.912	83 19.835	12.3	Near	9.2	4.33	39.01	29.44	1.45	32	462
WE2	29-Sep-14	0.15	0.1	19.4	41 45.884	83 20.107	11.95	Near	8.84	4.2	205.5	66.63	21.22	65.3	601
WE2	6-Oct-14	0.15	11.18	15.3	41 45.795	83 19.942	12.08	Near	8.04	NA	6.24	8.06	12.65	53	502
WE2	20-Oct-14	0.16	6.57	12.5	41 45.827	83 19.855	12.2	Near	8.04	NA	2.12	7.23	9.16	44.3	287
WE4	29-Jul-14	0.31	9.39	22.6	41 49.637	83 11.713	25.53	Off	8.59	1.01	4.77	6.12	1.01	16.5	328
WE4	4-Aug-14	0.14	1.64	24.6	41 49.714	83 11.654	25.64	Off	9.02	1.54	15.38	18.11	0.55	22.1	332
WE4	8-Sep-14	0.07	5.28	NA	41 49.585	83 11.629	25.55	Off	8.86	0.8	15.81	26.24	0.27	30.6	328
WE4	29-Sep-14	0.19	0.33	19.9	41 49.608	83 11.684	25.44	Off	8.58	0.2	3.65	9.21	0.1	17.7	464