Environmental Science & Technology

pubs.acs.org/est

Article

Synthesis of a Truncated Microcystin Tetrapeptide Molecule from a Partial Mcy Gene Cluster in Microcystis Cultures and Blooms

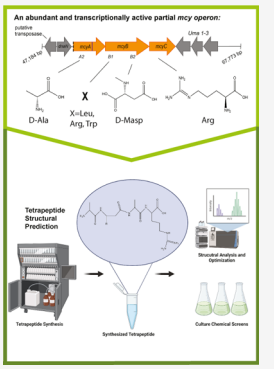
Colleen E. Yancey, Lauren Hart, Apurva Chandrakant Lad, Johnna A. Birbeck, Siliang Song, Osama G. Mohamed, Andrew M. Fribley, Steven T. Haller, Ashootosh Tripathi, David J. Kennedy, Judy A. Westrick, David H. Sherman, and Gregory J. Dick*

Cite This: *Environ. Sci. Technol.* 2024, 58, 19936−19947

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: *Microcystis* spp. threaten freshwater ecosystems through the proliferation of cyanobacterial harmful algal blooms (cyanoHABs) and production of the hepatotoxin, microcystin. While microcystin and its biosynthesis pathway, encoded by the *mcy* genes, have been well studied for over 50 years, a recent study found that *Microcystis* populations in western Lake Erie contain a transcriptionally active partial *mcy* operon, in which the A2 domain of *mcyA* and *mcyB-C* are present but the *mcyD-J* genes are absent. Here, we investigate the potential biosynthetic products and the evolutionary history of this partial operon. Our results reveal two candidate tetrapeptide constructs, with an X variable position, to be produced by strains with the partial operon. The partial operon appears necessary and sufficient for tetrapeptide biosynthesis and likely evolved from a single ancestor hundreds to tens of thousands of years ago. Bioactivity screens using Hep3B cells indicate a mild elevation of some markers of hepatotoxicity and inflammation, suggesting the need to further assess the effects of these novel secondary metabolites on freshwater ecosystems and public health. The need to assess these effects is even more pressing given the detection of tetrapeptides in both culture and western Lake Erie, which is a vital source of fresh water. Results from this study emphasize previous findings in which novel bacterial secondary metabolites may be derived from the molecular evolution of existing biosynthetic machinery under different environmental forcings.



KEYWORDS: *harmful algal blooms, Lake Erie, secondary metabolites, tetrapeptide, multiomics*

INTRODUCTION

Microcystin (MC), or the "fast-death-factor" was first identified in 1959[1] and is a potent hepatotoxin produced by cyanobacteria around the world within cyanobacteria harmful algal blooms (cyanoHABs).[2−4] This cyanotoxin is responsible for the illness and death of humans,[5,6] as well as livestock including sheep, cows, and birds.[7,8] Dangerously high levels of MCs have also led to drinking water crises such as those in Toledo (USA) in 2014[9] and Wuxi (China) in 2007.[10] While the toxicity of this molecule is caused via phosphatase inhibition,[11−13] its functional role in natural communities remains elusive, with hypothesized functions including iron acquisition, grazer defense, allelopathy, oxidative stress protection, and others.[14]

MCs have been well studied over the last 50 years and account for about 90% of secondary metabolite research on cyanobacteria.[15] Following structural elucidation in the mid-1980s,[16] over 270 congeners have been identified.[2] To date, all characterized congeners are cyclic heptapeptides that derive from complete *mcy* operons in which *mcy* genes (A-J) are present. Understanding congener diversity is important as slight structural changes can greatly impact potency and toxicity.[2,3] For example, a substitution from arginine in MC-LR to alanine in MC-LA results in significant increases in toxicity and lethality in mice.[12]

The cyclic heptapeptide structure of MCs contains an unusual, nonproteinogenic amino acid, 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyl- 4,6-decadienoic acid, (Adda),[1,16,17] which is often the target for detection assays.[18] The X and Y positions are highly variable and can be substituted for various amino acids, including leucine, arginine, tryptophan, and aspartic acid[2]. MCs are biosynthesized by a nonribosomal peptide synthetase (NRPS), polyketide synthase (PKS) multienzyme complex that contains hybrid NRPS-PKS, NRPS, tailoring, and ABC transporter modules.[19−21] Within *Microcystis*, this enzyme complex is encoded by the *mcy* gene cluster, which consists of 10 genes and is about 55 kilobases in length.[19] Other cyanobacterial taxa are also known to produce MCs, and their biosynthetic machinery is largely the same as that in *Microcystis*, although there is some variation observed in gene cluster structure.[22] The genetic substructure of the *mcy* gene cluster is responsible, in part, for variation in the congener

Received: January 5, 2024
Revised: September 19, 2024
Accepted: September 19, 2024
Published: October 29, 2024

production. For example, rearrangements and point mutations in the highly dynamic *mcyABC* region can lead to the production of MC-LR, MC-RR, or congeners with demethylated Adda domains.[21,23] Complete loss of MC synthesis can also occur via transposable elements that cause insertions and deletions within the gene cluster.[24,25]

Beyond the genetic architecture of the *mcy* gene cluster, several other factors influence differential isoform production. The concentration and form of nitrogen can influence the production of congeners with N-rich amino acids.[26] Similarly, amino acid availability within the environment can also impact structure, particularly which amino acids are incorporated in the variable X and Y positions.[27] The C:N ratio of environmental nutrients affects congener production as well.[28] Finally, the relaxed substrate binding specificity of the multienzyme biosynthetic complex allows for further variations of the MC structure.[23] While several factors influence the MC congener produced, the genes encoding its molecular assembly are valuable indicators of biosynthetic potential.

Recently, we detected a partial *mcy* genotype in which only the *mcyA2* domain, *mcyB,* and *mcyC* genes are present.[29] This partial genotype is transcriptionally active, persistent across multiple years, and occasionally the most abundant *mcy* genotype in western Lake Erie cyanoHABs, suggesting that it is ecologically successful.[29] To our knowledge, only one axenic *Microcystis* culture is known to contain this genotype, PCC 9717 (NCBI Genbank: GCA_00312165.1), but the partial operon has also been detected in two xenic *Microcystis* strains isolated from western Lake Erie.[30] Here, based on the known biosynthetic pathway of MC, we predicted that the partial *mcy* genotype produces tetrapeptides (TPs), and tested this hypothesis by screening cultured *Microcystis* strains with varying *mcy* genotypes. We present evidence to support this hypothesis along with an assessment of the evolutionary history of the partial haplotype and environmental occurrence and bioactivity of the tetrapeptide in terms of the potential for hepatoxicity.

MATERIALS AND METHODS

Partial Operon Annotation. The structure of the partial operon has previously been described and detected in the 2014 western Lake Erie cyanoHAB event.[29] To determine whether the culture isolates used within this study maintained the same genetic architecture observed in the field, several approaches were taken. The genome assembly of PCC 9717 (NCBI Genbank: GCA_00312165.1) was annotated with antiSMASH v. 6[31] to identify and annotate the partial operon and surrounding genes. Predicted amino acid sequences from genes in the partial operon were aligned to the nonredundant protein database on NCBI using BLASTp[32] to annotate the proposed function and identity. Strains from the Western Lake Erie Culture Collection (WLECC),[30] LE19−10.1 and LE19−251.1, were also annotated with antiSMASH v. 6, but due to poor assembly, the contiguous sequences containing partial *mcy* genes were fragmented and incomplete. To overcome limitations in assembly, paired-end read mapping of WLECC strains onto the PCC 9717 partial operon and flanking genes was also completed to determine sequence structure and orientation using bbmap[33] and visualized in Tablet v. 1.21.02.08.[34] Assembly statistics for PCC 9717, LE19−10.1, and LE19−251.1 are listed in Table S1. Additional alignments using BLASTn were completed to compare assembled genes found in WLECC strains and PCC 9717 and identify insertions or deletions. For both mapping approaches, a 95% identity and 80% alignment length cut-off was used to count mapped reads. The singular top hit for each read was also mapped to prevent multimapping. In addition to mapping to PCC 9717, reads from LE19−10.1 and LE19−251.1 were mapped to the complete *mcy* operon from PCC 7806 (Genbank Accession: AF183408.1) to confirm the content and structure of this gene cluster (Figure S1).

Structure Prediction and Chemical Synthesis. The product of the partial *mcy* operon was predicted based on the previously described biosynthesis pathways encoded in the complete *mcy* operon.[19,21,23] We hypothesized a tetrapeptide-construct (TP) to be the biosynthesis product of partial *mcy* operons based on the presence of *mcyA2*, *mcyB*, and *mcyC*, all of which encode nonribosomal peptide synthetases (NRPSs), which are modular enzymes that incorporate amino acids into secondary metabolite structures.[35] The predicted product therefore was: D-ala-X-D-MAsp-Arg, where X = leucine (L), arginine (R), or tryptophan (Y) based on common MC congeners found in the 2014 western Lake Erie cyanoHAB.[9] TPs were chemically synthesized at the University of Michigan Proteomics and Peptide Synthesis Core. A summary of the three tetrapeptides including purity, molecular weight, and yield is summarized in Table S2.

Standard Optimization. Prepared TPs were analyzed using an online concentration HPLC triple quadrupole mass spectrometry method to detect and quantify in both field and culture samples. TP standards were prepared as stated below and were analyzed on a Thermo TSQ Altis with an EQuan Max Plus system at the Lumigen Instrument Center at Wayne State University, Detroit, Michigan, USA. The method used was a modified method from Birbeck et al. 2019.[36] Briefly, the online concentration column was Hypersil GOLD aQ 20 mm × 2.1 mm, 12 μm, and the analytical column was an Accucore aQ 50 mm × 2.1 mm, 2.6 μm. The loading mobile phase for the online concentration was 0.1% formic acid in Milli-Q water and flowed at 1.5 mL/min during the injection of a 1 mL sample onto the column. Sample elution from the online concentration column onto the analytical column was completed by using a gradient with 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The gradient flow was 0.5 mL/min, started at 10% B, and was held for 1 min. The gradient was increased from 10 to 50% B from 1 to 6.5 min, then increased to 98% B from 6.51 to 7.5 min, and brought down back to 10% B to equilibrate for the next injection. Mass spectrometer settings were the same as Birbeck et al., 2019.[36]

Detection of the Tetrapeptide in Field Samples and Culture Isolates using HPLC/MS. Field data from the 2018 HAB Grab[37] was screened for the detection of the TP constructs. This sample set was chosen based on the availability of samples for analysis. During the 2018 HAB Grab sampling period MC-RR and MC-LR were the most abundant MC congeners detected and were especially concentrated around the shoreline and shallow regions of the western basin with peak concentrations ranging from 8,000 to 12,000 ppt (Westrick and Birbeck, unpublished). For detection from the field, standard curves were generated for each TP with a range of 0.5−500 parts per trillion (ppt) for the LR and YR tetrapeptides and 50−500 ppt for the RR tetrapeptide. Both the particulate and dissolved fractions were analyzed for each sample, as described above.

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

Select culture isolates of *Microcystis* from WLECC, PCC, or NIES were also screened for the detection of the TP constructs (Table 1). A variety of *mcy* genotypes were screened including

Table 1. Detection of the RR-tetrapeptide in *Microcystis* Culture Isolates from the Western Lake Erie Culture Collection (LE), Pasteur Culture Collection (PCC), or the National Intitute of Environmental Studies (NIES); ND Indicates that the RR-tetrapeptide was Not Detected in that Culture

culture	RR tetrapeptide concentration (ppt)	*mcy* genotype	culture status
blank control	ND		
BG-112N control	ND		
Milli-Q control	ND		
LE19−251.1	946.98	partial	xenic
LE19−10.1	4458.23	partial	xenic
PCC 9717	2135.77	partial	axenic
PCC 7806 Δ*mcyB*	ND	Δ*mcyB*	axenic
PCC 9806	ND	absent	axenic
PCC 7005	ND	absent	axenic
LE19−196.1	ND	absent	xenic
PCC7806	ND	complete	axenic
NIES-843	ND	complete	axenic
LE18−22.4	ND	complete	xenic
LE19−195.1	ND	complete	xenic
limit of detection (ppt)	10		

the complete (all *mcy* genes present), the absent (no *mcy* genes present), the partial (*mcyA2, mcyB-C* present), and an *mcyB* knockout mutant (PCC 7806, Δ*mcyB*) to assess the necessity and sufficiency of the partial operon in tetrapeptide biosynthesis. Both axenic and xenic *Microcystis* isolates were screened for TPs to determine if the presence of associated bacteria, known to coexist with *Microcystis* in natural cyanoHABs, had an impact on biosynthesis. PCC and NIES strains were obtained from their respective culture collections, while WLECC strains were isolated from western Lake Erie and cultivated by the Geomicrobiology Lab at the University of Michigan.[30] A total of 11 *Microcystis* isolates were grown in BG11−2N[38,39] at ∼23 °C (room temperature) and 40 μmol photons meter^{-2} second^{-1} on a 12:12 h light:dark cycle. Cultures were transferred every 2 weeks to fresh media until sufficient biomass was accumulated (around 25 mL of dense culture). Cyanobacterial biomass was then collected onto GF/F filters (Whatman, Maidstone, United Kingdom) and stored at −80 °C until mass spectrometry analysis could be completed as described in the section above. For detection in culture isolates, the limit of detection was 5 ppt for the YR construct and 10 ppt for the LR and RR constructs.

Phylogenetic and Evolutionary Analysis of the Partial Haplotype. Several phylogenetic approaches were used to assess the evolutionary history of the partial haplotype. Nucleotide sequences for individual genes were aligned using MUSCLE[40] as part of the MEGA11 workflow using default parameters. One thousand iterations were generated for each tree. Phylogenetic trees were reconstructed using MrBayes 3.2.7a, which implements Bayesian inference using Markov Chain Monte Carlo modeling approaches to reconstruct phylogenetic trees.[41] Sequences from the conserved *mcyA* region as well as the *mcyB* region were used for the analysis. *mcyC* was omitted due to poor assembly quality from the WLECC strains. For both *mcyA* and *mcyB*, consensus trees and time of divergence trees were reconstructed. In reconstructions of the time of divergence trees, the substitution rate in MrBayes was calibrated based on the previously estimated mutation rate for the intergenic spacer sequence (ITS) region.[43] Based on consensus trees, in which the partial haplotypes were observed to be monophyletic, we prespecified the partial haplotypes to form a monophyletic group in the time of divergence trees, which assumes that a single ancestor underwent a deletion event, in which the partial haplotype was the result. This was verified by confirming sequences upstream and downstream of the partial operon in PCC 9717, LE19−251.1, and LE19−10.1 were conserved, as these are the only publicly available strains with this genotype, as evidenced by read mapping of WLECC strain reads onto the PCC 9717 assembly of the partial operon (Table S3). In addition to single gene analyses, the same phylogenomic approaches were used on concatenated *mcyA* and *mcyB* sequences (from isolates in which these genes were assembled completely), since adjacent genes are expected to have the same evolutionary history. For all trees, *mcy* sequences from *Planktothrix agardhii* NIVA-CYA 126/8 (Genbank Accession: AJ441056.1) and *Anabaena* sp. 90 (Genbank Accession: AY212249.1) were used as outgroups. A complete list of Genbank Accession identifiers for sequences used in phylogenomic analysis can be found in Table S4.

Bioassay Screening for Antibacterial Activity and Cytotoxicity. Bioassays were conducted to test for antibacterial activity against methicillin-resistant *Staphylococcus aureus* (MRSA) and for cytotoxicity against SW48 and HCT15 (both colorectal cancer) cell lines. In preparation for bioassay screening, stock solutions of each TP were suspended in DMSO for a final concentration of 10 mM and stored at −20 °C until use. HCT15 (CCl-225) and SW-48 (CCL-231) cells were purchased from the American Type Culture Collection (ATCC). All cell lines were *Mycoplasma* free and independently authenticated by short tandem repeat profiling, performed by ATCC. Cells were grown and cultured according to ATCC recommendations. HCT15 cells were cultured in RPMI1640 (30−2001) supplemented with 10% fetal bovine serum (FBS) (30−2020). SW48 cells were cultured in Leibovitz's L-15 medium (30−2008) containing 10% FBS. SW-48 cells were grown and treated in an incubator set under atmospheric conditions (no supplemental CO_2 addition). For cell-based assays, cells were expanded and frozen into single-use aliquots. For each assay, cells were thawed at 37 °C for 1 min and then immediately resuspended into 10 mL of complete growth medium. Cells were then spun down at 300 × *g* for 5 min and then resuspended into cell-specific growth medium and plated at 2,500 cells per well into Greiner 781080 white cell culture 384-well plates with total volume per well at 40 μL. Natural product extracts or fractions were dissolved in DMSO at 15 mg/mL and delivered into the assay plates using Echo 655 acoustic liquid handler instrumentation (Beckman Coulter). Extract and fraction testing concentrations were at 0.25%. For primary screening assays extract testing was performed $n = 1$ at 0.25% final extract testing concentration (where original fraction is defined at 100%). Validation assay and fraction studies were performed in triplicate at similar testing concentrations. Negative controls medium only plus matching 0.25% DMSO were included in columns 1 and 2. The positive control for these studies was a 10 μM treatment with staurosporine in columns 23 and 24 of each assay plate. On each plate, samples were interrogated in wells A03 to P22.

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

The high-throughput data software MScreen[42] was utilized for primary hit, validation selection, and for analysis of concentration response curve results (1). Following compound addition, cells were cultured for 48 h at either 5% CO_2 at 37 °C for HCT15 cells or atmospheric air at 37 °C for SW-48 cells. Cell viability was measured using a CellTiter-Glo luminescent kit (catalog no. G7571) from Promega as directed using a PHERAstar instrument from BMG Labtech.

Hep3B Cells. Human hepatocellular carcinoma Hep3B (liver epithelial) cells were acquired from ATCC (catalog no. HB-8064, ATCC, VA, USA). The cells were grown in the recommended EMEM media (cat. no. 30−2003, ATCC, VA, USA) supplemented with 10% FBS (cat. no. FBS-BBT, Rocky Mountain Biologicals, Montana, USA,) and 1% penicillin-streptomycin solution (cat. no. PSL01−100 ML, Caisson Laboratories, UT, USA,). The cells were grown and maintained in T-75 culture flasks incubated at 37 °C with 5% CO_2. For the purposes of experiments, the cells were grown and treated in 6-well plates.

Hep3B Cell Line Exposures to MCs and Tetrapeptides. After seeding the cells in 6-well plates and reaching the desired confluency, the FBS-containing growth medium was replaced with 1 mL/well of serum-free media followed by incubation at 37 °C with 5% CO_2 for 3 h. For exposure to the tetrapeptides and their appropriate controls, lyophilized MC-LR (cat. no. 10007188, Cayman Chemicals, MI, USA) was dissolved in 1 mL of Ultrapure Distilled water (cat. no. 10977−015, Invitrogen, MA, USA) to obtain a stock concentration of 1 mg/mL. For the experiment, the cells were treated with a final concentration of 1 or 10 μM MC-LR. Similarly, lyophilized MC-RR (cat. no. 10007868, Cayman Chemicals, MI, USA) was dissolved in 1 mL of Ultrapure Distilled water to obtain a stock concentration of 100 μg/mL, and the cells were treated with a final concentration of 1 and 10 μM MC-RR exposure. For both tetrapeptide exposures, the above-mentioned 10 mM tetrapeptide stocks were diluted to 1000 μM working stocks using Ultrapure Distilled water, and the cells were treated with a final concentration of 1 or 10 μM. An irrelevant control peptide (~20 amino acids) was used as a negative control at the same concentration as the tetrapeptides and full-length MCs and cells were collected at 6 and 24 h after exposure. After the exposures, the cells were incubated at 37 °C with 5% CO_2 for appropriate time before proceeding to further experimental analysis.

RNA Extraction and RealTime−PCR analysis. RNA was extracted from the cells using the RNeasy Plus Mini Kit (cat. no. 74136, Qiagen, MD, USA) as per the manufacturer's protocol. RNA quality and purity were confirmed using nanodrop to measure 260/280 ratios and 260/230 ratios, which fell between the acceptable range indicating RNA purity. Approximately 500 ng of the extracted RNA was taken further to synthesize cDNA using QIAGEN's RT^2 First Strand Kit (cat. no. 330401, Qiagen, MD, USA). Automated liquid handling workflow system QIAgility (for qPCR sample and reagent loading) was used as described previously.[44,45] The reaction setup using the automated liquid handling system QIAgility is described in Table S5. All samples were run in duplicate. qPCR was performed using a Qiagen Rotor-Gene Q thermo-cycler. The cycle threshold values obtained in the process were used to calculate the fold change in the gene expression. 18S rRNA (cat. no. 4319413E, Thermo Fisher Scientific) was used as a housekeeping gene for normalization. The following Taqman primers obtained from Thermo Fisher Scientific were used to assess the hepatotoxicity in Hep3B cells: plasminogen activator inhibitor-1 (PAI-1 aka SerpinE1) (Hs00167155_m1), Alkaline Phosphatase (ALPL) (Hs01029144_m1), and transforming growth factor-beta 1 (Tgf-β1) (Hs00998133_m1). The program settings for the qPCR runs are described in Table 2.

Table 2. qPCR Instrument Cycle Settings for Inflammation and Hepatoxicity Studies

		temperature (°C)	time (seconds)	
initial denaturation		95	300 (5 min)	
cycling	denature	95	30	40×
	annealing	60	90	
	extension	72	30	

Statistical Analysis. Statistical analysis was performed using GraphPad PRISM 7 software (San Diego, CA, USA). For comparisons involving >2 groups, a one-way analysis of variance (ANOVA) was performed along with Tukey's multiple comparisons posthoc tests where applicable. All the data are presented as mean ± standard error of mean (S.E.M.) and a p-value of <0.05 was considered statistically significant.

RESULTS

Partial *mcy* Gene Cluster. The partial *mcy* gene cluster, which was previously described in western Lake Erie cyanoHABs,[29] differs from the complete *mcy* operon (Figure 1A) as it lacks *mcyD-J* and the A1 domain of *mcyA* (Figure 1B). The intact partial operon, which encodes NRPS enzymes, was hypothesized to incorporate four amino acids into a tetrapeptide molecule based on the known biosynthesis pathway in the complete *mcy* operon.[2,19] This tetrapeptide is predicted to contain a D-alanine (D-Ala), via the *mcyA2* domain, an X variable amino acid (Leu, Arg, Trp) via the *mcyB1* domain, *N*-methyl-D-aspartic acid (D-MAsp) via the *mcyB2* domain, and an arginine via *mcyC* (Arg) (Figure 1B).

The partial *mcy* operon has been identified in culture isolates PCC 9717, LE19−251.1, and LE19−10.1, the latter two being isolated from western Lake Erie.[29,30,46] Currently, the partial *mcy* operon has only been detected in these three publicly available culture isolates. The order and structure of genes in the partial operon are conserved with the complete operon as evidenced by the complete assembly in PCC 9717 and paired-end read mapping for LE19−251.1 and LE19−10.1 onto the complete *mcy* operon from PCC 7806 and flanking genes (Figure 1B, Figure S1, Table S1). The *mcy* gene sequences in the partial operons of these strains share 99−100% sequence identity at the nucleotide level. An 11 base pair insertion at position 1793 of the *mcyA* gene of PCC 9717 is not present in the western Lake Erie strains (Figure 1B, Table S6). Complete assembly of the partial operon and neighboring genes within PCC 9717 reveals the presence of several flanking genes that are also adjacent to the complete *mcy* operon;[19] we identified *uma1−3* genes upstream of *mcyC* and *dnaN*, which encodes a DNA polymerase III beta subunit, downstream of the truncated *mcyA*. We also identified a putative transposase adjacent to *dnaN* (Figure 1, Figure S1, and Table S7).

Detection of Tetrapeptides in Western Lake Erie. The three synthesized TP constructs were quantified in western Lake Erie samples (see Materials and Methods, Tables S2 and S8, Figure 2A). The LR-tetrapeptide was detected in samples

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

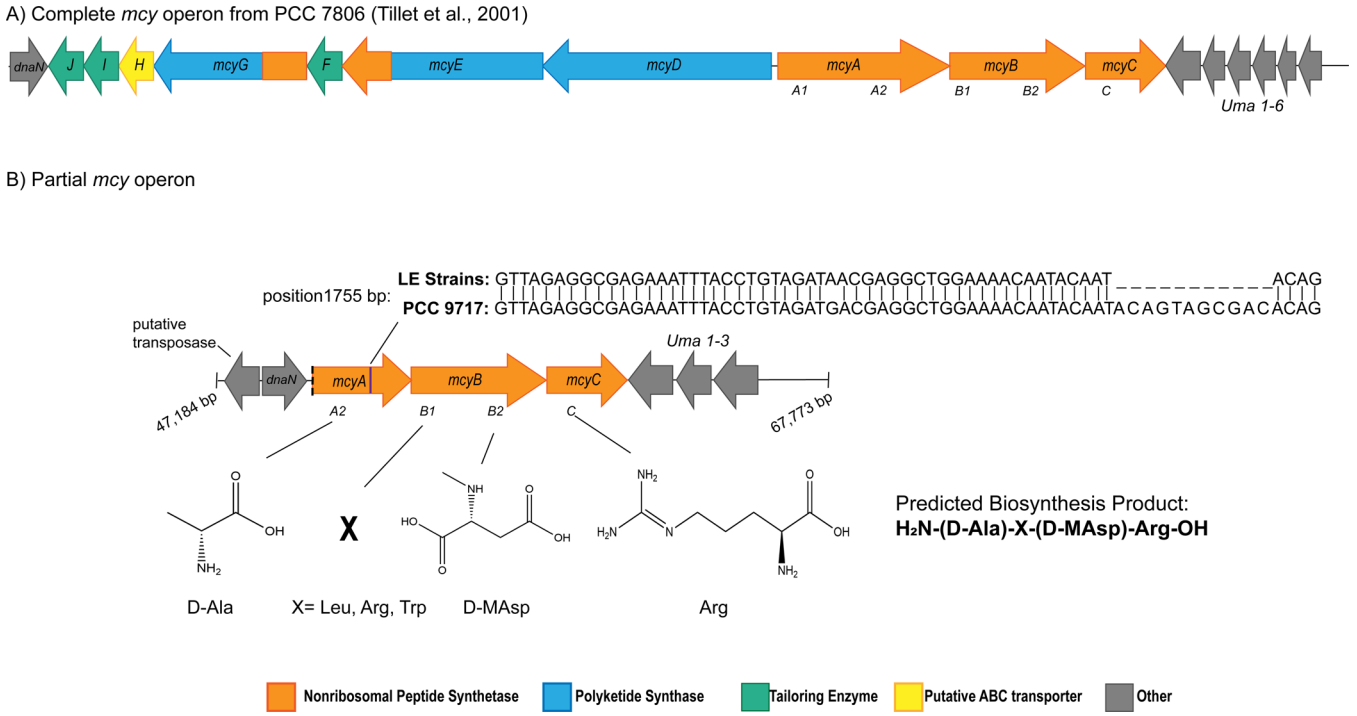


Figure 1. Gene schematics for (A) the complete *mcy* operon and neighboring genes in *Microcystis aeruginosa* PCC 7806 (Genbank Accession: AF183408.1), which is known to encode the biosynthesis of microcystin, and (B) the partial *mcy* operon. Based on the presence of conserved domains in gene *mcyA-C*, a tetrapeptide construct is the predicted biosynthesis product of this partial operon consisting of D-ala, an *X* variable position that may contain leu, arg, or trp, D-MAsp, and arg. The partial operon is detected in PCC 9717 (NCBI Genbank: GCA 00312165.1), LE19−251.1, and LE19−10.1. These sequences are highly conserved, with an 11 bp insertion observed in *mcyA* in PCC 9717 at position 1755 bp. The partial operon is found on a contiguous sequence that is over 70,000 base pairs in length in PCC 9717. The two genes upstream of the partial operon include *dnaN*, which encodes a DNA polymerase III beta subunit and a putative transposase. The three genes downstream of the partial operon putatively encode *uma1−3*..

collected from western Lake Erie during the 2018 HABs Grab, a binational sampling effort aimed at resolving high spatial resolution during a single day.[37] Of the 123 samples collected during sampling, 29 had detectable amounts of the LR-tetrapeptide ranging from about 12−19 ppt (Figure 2B,C, limit of detection = 0.5 ppt). The LR-tetrapeptide was detected in samples largely along the southern coast of Ohio with little detection in the center of the western basin (Figure 2C). Detection only occurred in the particulate fraction containing microbial biomass, not in the dissolved filtrate fraction, providing evidence that the tetrapeptide is intracellular and suggesting it is a biosynthesis rather than a degradation product. The partial genotype was detected in 5 of 26 available metagenomes from the HABs Grab but at low levels of relative abundance[29,37] (Table S8), which may explain low concentrations of the detected LR-tetrapeptide compared to concentrations of microcystin L,R (8,000−12,000 ppt, Zastepa et al., 2024, accepted at AEHM).

Detection of Tetrapeptides in Culture Isolates. Culture isolates were screened for TP constructs to determine if the partial operon can be attributed to their biosynthesis. Of all the isolates screened, the RR-tetrapeptide was only detected in strains containing the partial operon genotype, regardless of culture status (axenic vs xenic). We did not detect the LR- or YR-tetrapeptide from any of the culture isolates (Table 1). The concentration of the RR-tetrapeptide was high compared to field samples, ranging from 946.98 to 4458.23 ppt (Table 1, limit of detection = 10 ppt). Based on the presence of the RR-tetrapeptide in cultures only containing the partial operon, these results suggest that this genotype is necessary and sufficient for tetrapeptide biosynthesis. This result provides evidence that the tetrapeptide is a biosynthetic product and not a degradation product of complete microcystin molecules.

Phylogenomic and Evolutionary Analysis of the Partial Haplotype. To investigate the evolutionary history of the partial *mcy* gene cluster, various phylogenomic approaches were used to compare its relatedness to complete gene cluster sequences and estimate the time of divergence. Due to the lower quality of assemblies from WLECC strains,[30] concatenated *mcyA* and *mcyB* (Figure 3) sequences were analyzed as well as individual *mcyA* and *mcyB* genes (Figure S2), while *mcyC* was omitted. Consensus trees for concatenated genes reveal that the partial haplotype sequences are found within monophyletic groups (posterior probability = 99), suggesting a shared deletion event in the common ancestor of the three strains (Figure 3A). Observed differences in clustering based on *mcyA* and *mcyB* sequences may be explained by the dynamic and mosaic nature of this region of the *mcy* operon,[23] especially in *mcyB* where known recombination and replacements of entire domains have been observed.[21]

Time trees reveal that estimated times of divergence between the complete and partial haplotypes range from hundreds to tens of thousands of years based on which sequence was analyzed. These trees were built on the assumption that the deletion happened once within a common ancestor, which is corroborated by conserved flanking regions on either end of the partial *mcy* operons in all three strains with this genotype (Figure 1, Tables S3 and S4). For the concatenated sequences, the estimated time of divergence

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

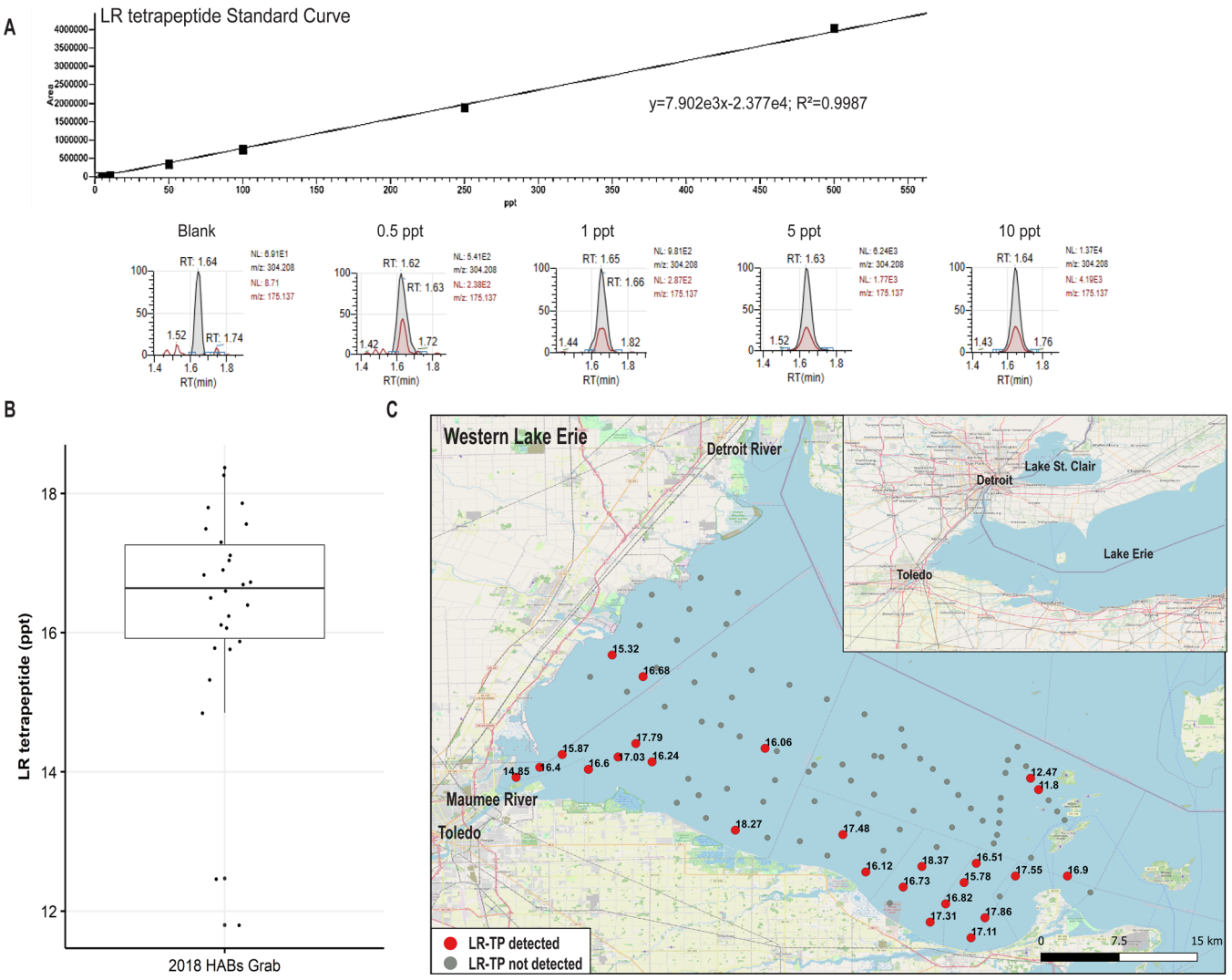


Figure 2. Detection of the LR-tetrapeptide during the 2018 HAB Grab using HPLC/MS. (A) Standard Curve for the LR tetrapeptide for quantification in HAB Grab 2018 samples. Various concentrations of the LR tetrapeptide (blank, 0.5 ppt, 1 ppt, 5 ppt, 10 ppt,) demonstrate accurate quantification of the LR tetrapeptide at the lowest reported values (~12 ppt). The retention time for the LR tetrapeptide was 1.62−1.65 min. (B) Quantification of the LR tetrapeptide from 29 HAB Grab Samples, ranging from around 12 to 19 ppt. (C) Map of western Lake Erie with sample points where the LR TP was detected. Samples in which the LR-TP were detected are colored red. Concentrations at each time point are denoted in ppt. Samples where the tetrapeptide was not detected are colored gray. The Open Street Map [OSM] was used as the basis for the rendered map (https://wiki.osmfoundation.org/wiki/Main_Page).

between the complete and partial haplotype is 2,300−13,900 years ago (Figure 3B). For the individual *mcyA* sequence, the estimated time of divergence was calculated to be about 3,400 to 31,700 years ago (Figure S2A). For the *mcyB* sequence, the estimated time of divergence was determined to be 100−14,300 years ago (Figure S2B). The variation in divergence times inferred from *mcyA* and *mcyB* may be explained by different rates of substitution or recombination between the two genes.[21,23] These estimations are in agreement with previous calculations of times of divergence based on upper and lower limits of observed bacterial mutation rates.[29]

Bioassay Screens to Assess Activity. Bioassays to test for antibacterial activity against methicillin-resistant *Staphylococcus aureus* (MRSA), as measured by growth inhibition, and for cytotoxicity against SW48 and HCT15 cell lines were both negative (data not shown).

Tetrapeptide MC-RR Shows a Mild Hepatotoxic Response in Hep3B Cells. In order to screen for potential hepatotoxic effects of the LR and RR tetrapeptides, Hep3B human hepatocellular carcinoma cells were exposed to the TPs (as well as their full-length MC-LR and MC-RR counterparts) at two doses (1 and 10 μM) for either 6 or 24 h and compared with both vehicle (water) and irrelevant peptide controls (Figure S3). Quantitative PCR analysis was performed for markers of hepatotoxicity including PAI-1/SerpinE1, ALPL, and Tgfβ1 (Figure 4A−C). After 6 h of exposure, TP-LR (1 μM) exhibited significant increases in both ALPL (Figure 4A) and SerpinE1 (Figure 4B), while TP-RR (10 μM) exhibited significant increases in ALPL (Figure 4A) relative to both vehicle and control peptide. After 24 h of exposure, TP-RR (1 μM) exhibited significant increases in both SerpinE1 (Figure 4B) and Tgfβ1 (Figure 4C), relative to both vehicle and control peptide.

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

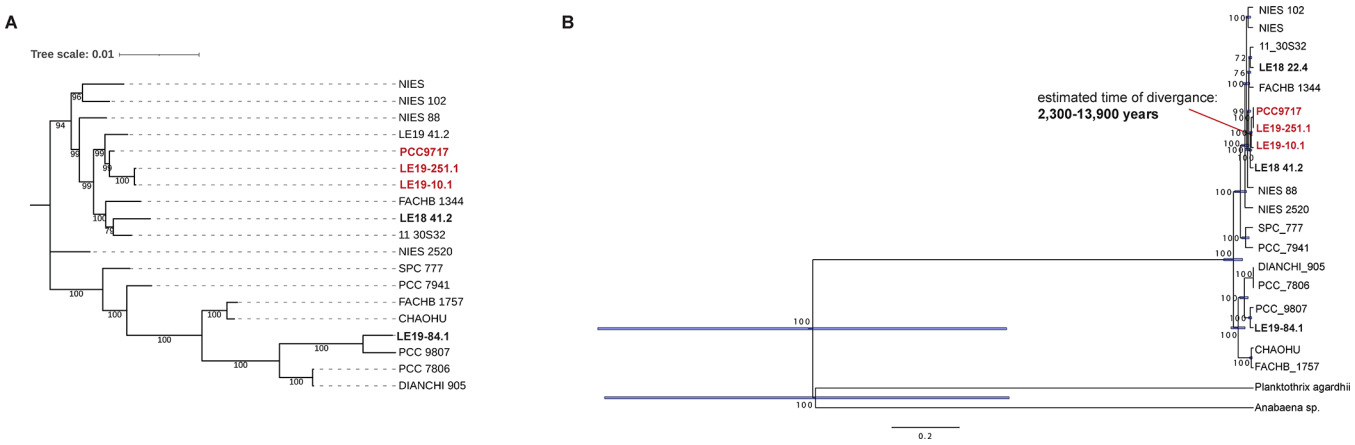


Figure 3. Phylogenomic analyses of conserved sequences in the complete and partial haplotypes using concatenated *mcyA* and *mcyB* sequences from both the WLECC and publicly available *Microcystis* isolate genomes. Isolates with the partial gene cluster are bolded and red, while isolates from the WLECC with the complete operon are bolded and black. (A) The tree on the left is a consensus tree constructed by MRBAYES. Posterior probabilities are shown at each node. (B) The tree on the right represents time of divergence generated by MRBAYES. Based on the time tree and rate dating the partial haplotype diverged thousands of years ago from the complete haplotype. The red line points to the node representing the last common ancestor shared between strains with the complete or partial haplotype. To generate both trees, 1000 iterations were completed, and *Planktothrix agardhii* NIVA-CYA 126/8 (Genbank Accession: AJ441056.1) and *Anabaena* sp. 90 (Genbank Accession: AY212249.1) were used as outgroups. Note: outgroups have been hidden on consensus trees.

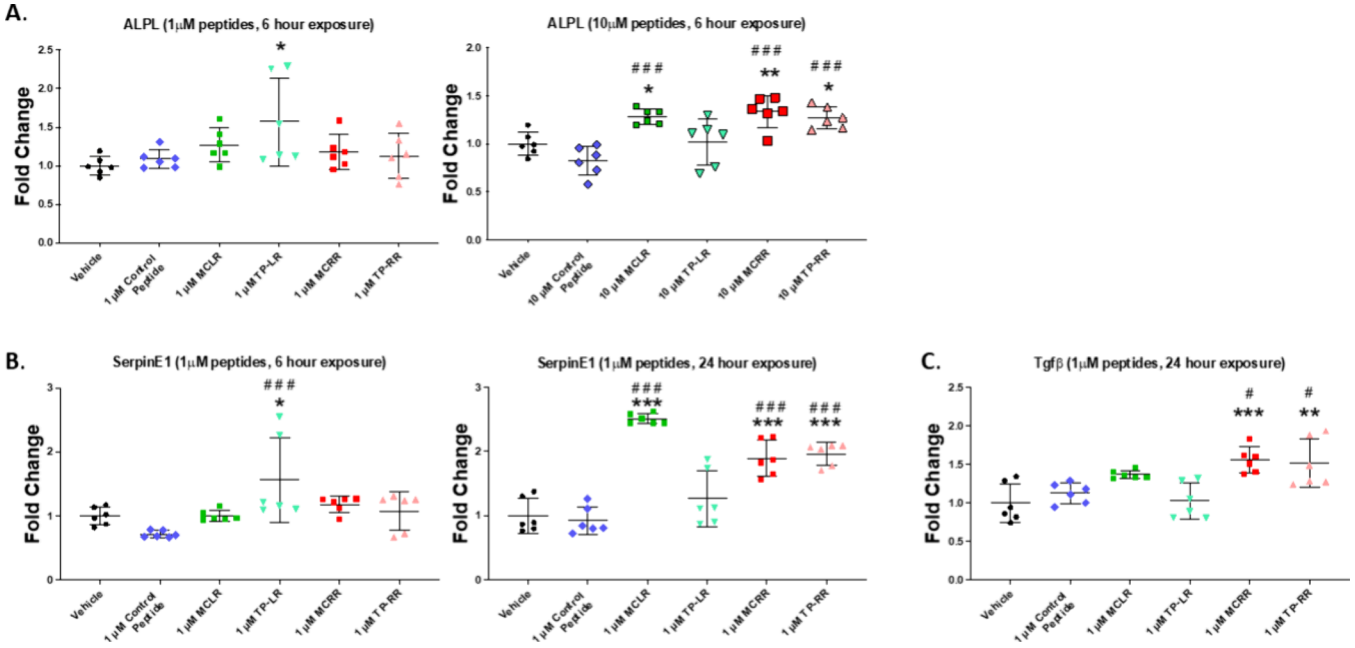


Figure 4. Quantitative PCR analysis of markers of hepatotoxicity in Hep3B cells after exposure for either 6 or 24 h to the indicated concentrations of microcystin TPs or full length microcystins vs vehicle (water) or irrelevant peptide controls. (A) Alkaline phosphatase (ALPL) expression after 6 h exposure to 1 μM (left) and 10 μM (right) of indicated peptides. (B) Plasminogen activator inhibitor-1 (PAI-1 aka SerpinE1) expression after 6 h (left) and 24 h (right) exposure to 1 μM of indicated peptides. (C) Transforming growth factor-beta 1 (Tgf-β1) expression after 24 h exposure to 1 μM of indicated peptides; $^{*}p \leq 0.05$, ** $p \leq 0.01$, $^{***}p \leq 0.001$vs vehicle; $^{\#}p \leq 0.05$, $^{\#\#}$ $p \leq 0.01$, $^{\#\#\#}p \leq 0.001$ vs control peptide.

DISCUSSION

Our findings provide evidence that a tetrapeptide construct is biosynthesized from transcriptionally active partial *mcy* operons that have been observed in both field and culture samples.[29,30] These results challenge the notion that modified *mcy* gene clusters, either through the insertion of transposable elements or deletion of entire genes, are nonfunctional,[24,25,47] transient artifacts of genome streamlining.[48−50] While the RR tetrapeptide was exclusively detected in culture, the LR tetrapeptide was detected only in the field, and the YR tetrapeptide was not detected at all; these results highlight the possibility for truncated *mcy* products to be biosynthesized under different conditions when varying substrates are available. The functional role and direct biomolecular target(s) of the tetrapeptide have yet to be established, and available field data on its environmental distribution is limited. However, these findings suggest the potential for a novel secondary metabolite to influence observed community dynamics within cyanoHABs where strains with the partial operon are present.

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

Interestingly, we detect only the LR tetrapeptide in the field and only the RR tetrapeptide in culture. This may be due to different substrates available in the natural freshwater environment versus nutrient-replete culture media. During the 2018 HAB Grab in which the LR tetrapeptide was detected (Figure 2), total nitrogen measurements at the time of sampling were comparatively lower to other bloom years.[9,29,37] Cultures in which the RR tetrapeptide was detected were grown in nutrient-replete BG-112N, which contains 2 mM of $NaNO_3$.[38,39] Differences in available nitrogen may explain the incorporation of N-rich arginine into tetrapeptide constructs grown in culture compared to N-poor leucine in tetrapeptides detected in the field.[26−28,51]

The presence of additional variants of the tetrapeptide cannot be ruled out. LR, RR, and YR tetrapeptides were chosen for chemical synthesis, as these were the most abundant, complete MC congeners produced in western Lake Erie during the 2014 bloom.[9] However, in more recent blooms in western Lake Erie, the most abundant MC congeners have been LR, RR, and LA, while YR is in much lower abundance.[52] These recent trends, and the known diversity of MC congeners,[2] highlight the possibility of other tetrapeptide constructs to be biosynthesized. Future studies should consider the possibility of tetrapeptide constructs with LA in the X and Z variable positions as it has become more abundant in western Lake Erie blooms,[37,52] and complete MC-LA has been shown to persist longer in surface waters and has similar toxicity as MC-LR.[53]

Another need for future investigations is the source of D-MeAsp, which is part of the tetrapeptide construct (Figure 1B). Previous work determined that an Asp racemase, encoded by *mcyF*, is necessary to switch the stereochemistry of the MeAsp from L to D conformation.[54,55] *mcyF* encodes the only known aspartate racemase in *Microcystis* genomes,[54] and it is lacking from strains with the partial genotype (Figure 1B). Other racemases encoded elsewhere in *Microcystis* genomes, outside of the *mcy* operon, may provide D-glutamate for complete MC biosynthesis.[55] However, searches for *mcyF* homologues in genomes that contain the partial genotype do not return any hits (data not shown), suggesting these strains lack aspartate racemases altogether, although novel isomerases cannot be ruled out. Alternatively, D-amino acids for tetrapeptide biosynthesis may be obtained exogenously from the environment, potentially through exometabolite exchange, as commonly observed in microbial communities.[56−58] However, it is unknown if *Microcystis* contains specific transporters for this function. If such transporters exist, this mechanism is feasible given that partial operon genotypes co-occur with *Microcystis* strains that have complete *mcy* operon, including the *mcyF* racemase in Western Lake Erie Blooms.[29]

Phylogenomic analysis, based on gene sequences of *mcyA* and *mcyB*, and conservation of flanking genes in strains with the partial operon provide evidence that this deletion occurred once in a common ancestor, hundreds to thousands of years ago, and has persisted through time (Figure 3, Table S3 and Figure S2). This is further supported by the fact that the partial operon was observed in western Lake Erie cyanoHABs from 2014 and 2018[29,30] and in PCC 9717, which was isolated from a water dam in Rochereau, La Vendée, France in 1996 (Institut Pasteur, Paris, France). It is well-established that the evolution and modification of biosynthetic gene clusters can give rise to new biosynthesis pathways and secondary metabolites. For example, the nodularin biosynthesis pathway is believed to be derived from the *mcy* operon via deletions of the second domain in *mcyA* and the first domain in *mcyB* and fusion of remaining modules into a singular gene, *ndaA*.[22,59] Horizontal gene transfer, mutation, and recombination of biosynthetic gene clusters,[60,61] along with evolutionary pressures due to limited pools of substrates in natural communities, may explain the modification and diversification of produced secondary metabolites that share similar evolutionary routes.[22,60]

While the deletion event in the partial operon was substantial, eliminating over half of the genes needed for complete MC biosynthesis, its persistence through time, transcriptional activity in the field,[29] and detection of proposed biosynthesis products provide evidence that the partial operon is functional and benefits fitness. Although the functional role of complete MCs is complex and likely multifacted;[14] the partial operon may encode a truncated microcystin-like molecule that serves a critical function in changing environments. One explanation for this functional intermediate is the trade-off between producing N-rich MCs over none at all under depleted conditions. Environmental stress, including nitrogen limitation, has previously been shown to select for gene loss and genome streamlining in bacteria in marine systems, especially genes related to nitrogen metabolism.[62,63] The partial operon has been shown to dominate blooms in the late season in western Lake Erie, when bioavailable N is considerably lower,[29,37] suggesting a functional role under environmental stress and nitrogen limitation.

We evaluated the biological activity of the tetrapeptides using Hep3B cells, a cell line derived from human hepatocellular carcinoma, and a well-established screening model for drug toxicity. Our preliminary studies showed that a final concentration of 10 μM could elicit a significant hepatotoxic response from the full-length MCs without affecting the viability of the cells, and thus, we assessed potential hepatotoxicity after short-term (6 h) and long-term (24 h) exposure of both low (1 μM) and high (10 μM) dose of the tetrapeptides. Previous work has demonstrated that microcystins effectively target and inhibit protein phosphatases 1 and 2A, although structural differences in congeners have varied effects on protein phosphatases.[64]

Hepatotoxicity was assessed by gene expression of several common markers identified in previous studies including SerpinE1, ALPL, and Tgf-β1.[44,64] Results of this analysis demonstrated that both tetrapeptide-LR and tetrapeptide-RR showed mild but significant upregulation of these markers, albeit with slightly different doses and time points. It should be noted that the tetrapeptides did not appear to have a standard dose- and time-dependent relationship with the various markers of cytotoxicity. Interestingly, we and others have observed a similar trend with full-length MC congeners wherein increases in the toxin concentration did not translate to an increase in the genetic expression of hepatotoxicity markers or the associated enzymatic activity of common hepatic injury enzymes such as alkaline phosphatase, but in some cases, there was actually a suppression of this phenotype.[65,66] Based on these studies and the fact that the putative target of these toxins are ubiquitous protein phosphatases which control a wide spectrum of cellular processes, further study is warranted to understand the exact mechanism whereby MC tetrapeptides may exert their cytotoxic effects. The fact that these effects were noted in the absence of the ubiquitous Adda domain in full-length microcystin congeners suggests the need to characterize the

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

impact of the tetrapeptides more thoroughly and other potential tetrapeptide congeners on both aquatic ecosystems as well as human and animal health, as they may have the potential to alter protein phosphatase activity.

Results from this study highlight the dynamic nature of biosynthetic gene clusters encoded in *Microcystis* genomes and that pathways that have been well studied for over 30 years are still rich sources for discovery, particularly through the use of multiomic approaches. To the best of our knowledge, these findings provide the first evidence of a tetrapeptide biosynthesis product from the partial *mcyA-C* gene cluster. Future work is needed to further assess bioactivity, as preliminary results suggest the potential for some tetrapeptide congeners to exhibit mild levels of hepatoxicity. Further screens will help determine threats to ecosystem function and access to clean drinking water. Additional work should also screen for the presence of the partial operon in *Microcystis*-dominated cyanoHABs around the world to assess its presence and persistence on a global scale.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.4c00039.

(Table S1) Assembly metrics for culture isolates with partial *mcy* operon; (Table S2) summary of three synthesized tetrapeptides; (Table S3) paired-end read mapping and coverage of WLECC strains onto partial operon from PCC 9717; (Table S4) Genbank Accession identifiers for sequences used in phylogenomic analysis; (Table S5) reaction setup using automated liquid handling workflow system QIAgility for qPCR; (Table S6) BLASTn results for partial *mcyA* genes; (Table S7) BLASTn summaries from querying the NR database with the *mcy* operon and flanking genes from PCC 7806; (Table S8) quantification and of the LR tetrapeptide and relative abundance of the partial genotype; (Figure S1) read mapping plots of WLECC isolates reads onto the *mcy* operon and flanking genes from PCC 7806; (Figure S2) phylogenetic analysis for individual *mcyA* and *mcyB* genes; and (Figure S3) complete quantitative qPCR results on HEP3B cells (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Gregory J. Dick − *Department of Earth and Environmental Science, University of Michigan, Ann Arbor, Michigan 48109, United States; Cooperative Institute for Great Lakes Research, University of Michigan, Ann Arbor, Michigan 48108, United States;* orcid.org/0000-0001-7666-6288; Email: gdick@umich.edu

Authors

Colleen E. Yancey − *Department of Earth and Environmental Science, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0002-6078-9483

Lauren Hart − *Program in Chemical Biology and Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0001-8553-2839

Apurva Chandrakant Lad − *Department of Medicine, University of Toledo, Toledo, Ohio 43614, United States;* orcid.org/0000-0001-6073-7432

Johnna A. Birbeck − *Lumigen Instrument Center, Wayne State University, Detroit, Michigan 48202, United States*

Siliang Song − *Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0002-4658-5062

Osama G. Mohamed − *Life Sciences Institute and Natural Products Discovery Core, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0003-3216-6948

Andrew M. Fribley − *Department of Pediatrics, Wayne State University, Detroit, Michigan 48202, United States;* orcid.org/0000-0002-2046-1600

Steven T. Haller − *Department of Medicine, University of Toledo, Toledo, Ohio 43614, United States;* orcid.org/0000-0002-6919-6342

Ashootosh Tripathi − *Life Sciences Institute and Natural Products Discovery Core, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0003-2297-8088

David J. Kennedy − *Department of Medicine, University of Toledo, Toledo, Ohio 43614, United States;* orcid.org/0000-0001-5265-0142

Judy A. Westrick − *Lumigen Instrument Center and Department of Chemistry, Wayne State University, Detroit, Michigan 48202, United States;* orcid.org/0000-0002-1712-5737

David H. Sherman − *Life Sciences Institute, University of Michigan, Ann Arbor, Michigan 48109, United States;* orcid.org/0000-0001-8334-3647

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.4c00039

Funding

This work was supported by NIH and NSF awards to the Great Lakes Center for Fresh Waters and Human Health (NIH: 1P01ES028939−01, and 2P01ES028939−06; NSF: OCE-1840715 and OCE-2418066), the NOAA Great Lakes Omics program distributed through the UM Cooperative Institute for Great Lakes Research (NA17OAR4320152 and NA22OAR4320150., and the Hans W. Vahlteich Professorship. This material is based upon work supported by the U.S. Geological Survey under Grant/Cooperative Agreement No G23AC00555−00. Support was also provided by UM Natural Products Biosciences Initiative, the NOAA OAR Ocean Technology Development Initiative, and NIEHS 1R01ES034017. In addition, his research was funded by a Harmful Algal Bloom Research Initiative grant from the Ohio Department of Higher Education, the David and Helen Boone Foundation Research Fund, and the University of Toledo Women, and the Philanthropy Genetic Analysis Instrumentation Center. Support was also provided by the Center for Urban Response to Environmental Stressors (CURES, NIEHS P30 ES036084, and 1R01ES034017)

Notes

The authors declare the following competing financial interest(s): C.E.Y. was a graduate student at the time this research was completed, but is currently employed at New England Biolabs, Ipswich, MA, USA.

■ ACKNOWLEDGMENTS

We thank Robert Hein for the bioinformatic support. We thank Henriette Remmer for assistance in the chemical

synthesis of the tetrapeptides at the Proteomics and Peptide Synthesis Core at the University of Michigan. We thank McKenzie Powers, Katherine Polik, and Helena Nitschky for assistance in culturing and maintenance of the WLECC.

REFERENCES

(1) Bishop, C. T.; Anet, E. F. L. J.; Gorham, P. R. Isolation and Identification of the Fast-Death Factor in Microcystis Aeruginosa NRC-1. *Can. J. Biochem. Physiol.* **1959**, *37* (3), 453−471.
(2) Bouaïcha, N.; Miles, C. O.; Beach, D. G.; Labidi, Z.; Djabri, A.; Benayache, N. Y.; Nguyen-Quang, T. Structural Diversity, Characterization and Toxicology of Microcystins. *Toxins* **2019**, *11* (12), 714.
(3) Díez-Quijada, L.; Prieto, A. I.; Guzmán-Guillén, R.; Jos, A.; Cameán, A. M. Occurrence and Toxicity of Microcystin Congeners Other than MC-LR and MC-RR: A Review. *Food Chem. Toxicol.* **2019**, *125*, 106−132.
(4) Harke, M. J.; Steffen, M. M.; Gobler, C. J.; Otten, T. G.; Wilhelm, S. W.; Wood, S. A.; Paerl, H. W. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, Microcystis spp. *Harmful Algae* **2016**, *54*, 4−20.
(5) Jochimsen, E. M.; Carmichael, W. W.; An, J.; Cardo, D. M.; Cookson, S. T.; Holmes, C. E. M.; Antunes, M. B.; de Melo Filho, D. A.; Lyra, T. M.; Barreto, V. S. T.; Azevedo, S. M. F. O.; Jarvis, W. R. Liver Failure and Death after Exposure to Microcystins at a Hemodialysis Center in Brazil. *New England Journal of Medicine* **1998**, *338* (13), 873−878.
(6) Pouria, S.; De Andrade, A.; Barbosa, J.; Cavalcanti, R. L.; Barreto, V. T. S.; Ward, C. J.; Preiser, W.; Poon, G. K.; Neild, G. H.; Codd, G. A. Fatal Microcystin Intoxication in Haemodialysis Unit in Caruaru. *Brazil. Lancet* **1998**, *352* (9121), 21−26.
(7) Carbis, C. R.; Simons, J. A.; Mitchell, G. F.; Anderson, J. W.; McCauley, I. A Biochemical Profile for Predicting the Chronic Exposure of Sheep to Microscystis Aeruginosa, an Hepatotoxic Species of Blue-Green Alga. *Research in Veterinary Science* **1994**, *57* (3), 310−316.
(8) Stewart, I.; Seawright, A. A.; Shaw, G. R. Cyanobacterial Poisoning in Livestock, Wild Mammals and Birds–an Overview. *Advances in experimental medicine and biology* **2008**, *619*, 613−637.
(9) Steffen, M. M.; Davis, T. W.; McKay, R. M. L.; Bullerjahn, G. S.; Krausfeldt, L. E.; Stough, J. M. A.; Neitzey, M. L.; Gilbert, N. E.; Boyer, G. L.; Johengen, T. H.; Gossiaux, D. C.; Burtner, A. M.; Palladino, D.; Rowe, M. D.; Dick, G. J.; Meyer, K. A.; Levy, S.; Boone, B. E.; Stumpf, R. P.; Wynne, T. T.; Zimba, P. V.; Gutierrez, D.; Wilhelm, S. W. Ecophysiological Examination of the Lake Erie *Microcystis* Bloom in 2014: Linkages between Biology and the Water Supply Shutdown of Toledo. *OH. Environmental Science & Technology* **2017**, *51* (12), 6745−6755.
(10) Qin, B.; Zhu, G.; Gao, G.; Zhang, Y.; Li, W.; Paerl, H. W.; Carmichael, W. W. A Drinking Water Crisis in Lake Taihu, China: Linkage to Climatic Variability and Lake Management. *Environmental Management* **2010**, *45* (1), 105−112.
(11) Dawson, R. M. The Toxicology of Microcystins. *Toxicon* **1998**, *36* (7), 953−962.
(12) Chernoff, N.; Hill, D.; Lang, J.; Schmid, J.; Le, T.; Farthing, A.; Huang, H. The Comparative Toxicity of 10 Microcystin Congeners Administered Orally to Mice: Clinical Effects and Organ Toxicity. *Toxins* **2020**, *12* (6), 403.
(13) Honkanen, R. E.; Zwiller, J.; Moore, R. E.; Daily, S. L.; Khatra, B. S.; Dukelow, M.; Boynton, A. L. Characterization of Microcystin-LR, a Potent Inhibitor of Type 1 and Type 2A Protein Phosphatases. *J. Biol. Chem.* **1990**, *265* (32), 19401−19404.
(14) Omidi, A.; Esterhuizen-Londt, M.; Pflugmacher, S. Still Challenging: The Ecological Function of the Cyanobacterial Toxin Microcystin − What We Know so Far. *Toxin Reviews* **2018**, *37* (2), 87−105.
(15) Janssen, E. M. L. Cyanobacterial Peptides beyond Microcystins − A Review on Co-Occurrence, Toxicity, and Challenges for Risk Assessment. *Water Res.* **2019**, *151*, 488−499.
(16) Botes, D. P.; Tuinman, A. A.; Wessels, P. L.; Viljoen, C. C.; Kruger, H.; Williams, D. H.; Santikarn, S.; Smith, R. J.; Hammond, S. J. The Structure of Cyanoginosin-LA, a Cyclic Heptapeptide Toxin from the Cyanobacterium Microcystis Aeruginosa. *Journal of the Chemical Society, Perkin Transactions 1* **1984**, *0*, 2311−2318.
(17) Rinehart, K. L.; Namikoshi, M.; Choi, B. W. Structure and Biosynthesis of Toxins from Blue-Green Algae (Cyanobacteria). *J. Appl. Phycol* **1994**, *6* (2), 159−176.
(18) Zeck, A.; Weller, M. G.; Bursill, D.; Niessner, R. Generic Microcystin Immunoassay Based on Monoclonal Antibodies against Adda. *Analyst* **2001**, *126* (11), 2002−2007.
(19) Tillett, D.; Dittmann, E.; Erhard, M.; Von Döhren, H.; Börner, T.; Neilan, B. A. Structural Organization of Microcystin Biosynthesis in Microcystis Aeruginosa PCC7806: An Integrated Peptide-Polyketide Synthetase System. *Chemistry and Biology* **2000**, *7* (10), 753−764.
(20) Welker, M.; Von Döhren, H. Cyanobacterial Peptides - Nature's Own Combinatorial Biosynthesis. *FEMS Microbiology Reviews* **2006**, *30* (4), 530−563.
(21) Mikalsen, B.; Boison, G.; Skulberg, O. M.; Fastner, J.; Davies, W.; Gabrielsen, T. M.; Rudi, K.; Jakobsen, K. S. Natural Variation in the Microcystin Synthetase Operon mcyABC and Impact on Microcystin Production in Microcystis Strains. *Journal of bacteriology* **2003**, *185* (9), 2774−2785.
(22) Dittmann, E.; Fewer, D. P.; Neilan, B. A. Cyanobacterial Toxins: Biosynthetic Routes and Evolutionary Roots. *FEMS Microbiology Reviews* **2013**, *37* (1), 23−43.
(23) Tooming-Klunderud, A.; Mikalsen, B.; Kristensen, T.; Jakobsen, K. S. The Mosaic Structure of the mcyABC Operon in Microcystis. *Microbiology* **2008**, *154* (7), 1886−1899.
(24) Pearson, L. A.; Hisbergues, M.; Börner, T.; Dittmann, E.; Neilan, B. A. Inactivation of an ABC Transporter Gene, mcyH, Results in Loss of Microcystin Production in the Cyanobacterium Microcystis Aeruginosa PCC 7806. *Applied and environmental microbiology* **2004**, *70* (11), 6370−6378.
(25) Christiansen, G.; Kurmayer, R.; Liu, Q.; Börner, T. Transposons Inactivate Biosynthesis of the Nonribosomal Peptide Microcystin in Naturally Occurring Planktothrix Spp. *Applied and environmental microbiology* **2006**, *72* (1), 117−123.
(26) Puddick, J.; Prinsep, M.; Wood, S.; Kaufononga, S.; Cary, S.; Hamilton, D. High Levels of Structural Diversity Observed in Microcystins from Microcystis CAWBG11 and Characterization of Six New Microcystin Congeners. *Marine Drugs* **2014**, *12* (11), 5372−5395.
(27) Tonk, L.; Van De Waal, D. B.; Slot, P.; Huisman, J.; Matthijs, H. C. P.; Visser, P. M. Amino Acid Availability Determines the Ratio of Microcystin Variants in the Cyanobacterium Planktothrix Agardhii. *FEMS Microbiology Ecology* **2008**, *65* (3), 383−390.
(28) Liu, J.; Van Oosterhout, E.; Faassen, E. J.; Lürling, M.; Helmsing, N. R.; Van de Waal, D. B. Elevated pCO2 Causes a Shift towards More Toxic Microcystin Variants in Nitrogen-Limited Microcystis Aeruginosa. *FEMS Microbiol. Ecol.* **2016**, *92* (2), No. fiv159.
(29) Yancey, C. E.; Smith, D. J.; Uyl, P. A. D.; Mohamed, O. G.; Yu, F.; Ruberg, S. A.; Chaffin, J. D.; Goodwin, K. D.; Tripathi, A.; Sherman, D. H.; Dick, G. J. 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.* **2022**, *88*, No. e0246421.
(30) Yancey, C. E.; Kiledal, E. A.; Chaganti, S. R.; Denef, V. J.; Errera, R. M.; Evans, J. T.; Hart, L. N.; Isailovic, D.; James, W. S.; Kharbush, J. J.; Kimbrel, J. A.; Li, W.; Mayali, X.; Nitschky, H.; Polik, C. A.; Powers, M. A.; Premathilaka, S. H.; Rappuhn, N. A.; Reitz, L. A.; Rivera, S. R.; Dick, G. J. The Western Lake Erie Culture Collection: A Promising Resource for Evaluating the Physiological and Genetic Diversity of Microcystis and Its Associated Microbiome. *Harmful Algae* **2023**, *126*, No. 102440.

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947

(31) Blin, K.; Shaw, S.; Kloosterman, A. M.; Charlop-Powers, Z.; Van Wezel, G. P.; Medema, M. H.; Weber, T. antiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities. *Nucleic Acids Res.* **2021**, *49* (W1), W29−W35.
(32) Madden, T. The BLAST Sequence Analysis Tool. In *The NCBI handbook*; 2013.
(33) Bushnell, B. *BBTools User Guide - DOE Joint Genome Institute*; https://sourceforge.net/projects/bbmap/ (accessed 2021−04−09).
(34) Milne, I.; Stephen, G.; Bayer, M.; Cock, P. J. A.; Pritchard, L.; Cardle, L.; Shawand, P. D.; Marshall, D. Using Tablet for Visual Exploration of Second-Generation Sequencing Data. *Briefings Bioinf.* **2013**, *14* (2), 193−202.
(35) Miller, B. R.; Gulick, A. M. Structural Biology of Non-Ribosomal Peptide Synthetases. *Methods Mol. Biol.* **2016**, *1401*, 3−29.
(36) Birbeck, J. A.; Westrick, J. A.; O'neill, G. M.; Spies, B.; Szlag, D. C. Comparative Analysis of Microcystin Prevalence in Michigan Lakes by Online Concentration LC/MS/MS and ELISA. *Toxins* **2019**, *11* (1), 13.
(37) Chaffin, J. D.; Bratton, J. F.; Verhamme, E. M.; Bair, H. B.; Beecher, A. A.; Binding, C. E.; Birbeck, J. A.; Bridgeman, T. B.; Chang, X.; Crossman, J.; Currie, W. J. S.; Davis, T. W.; Dick, G. J.; Drouillard, K. G.; Errera, R. M.; Frenken, T.; MacIsaac, H. J.; McClure, A.; McKay, R. M.; Reitz, L. A.; Domingo, J. W. S.; Stanislawczyk, K.; Stumpf, R. P.; Swan, Z. D.; Snyder, B. K.; Westrick, J. A.; Xue, P.; Yancey, C. E.; Zastepa, A.; Zhou, X. The Lake Erie HABs Grab: A Binational Collaboration to Characterize the Western Basin Cyanobacterial Harmful Algal Blooms at an Unprecedented High-Resolution Spatial Scale. *Harmful Algae* **2021**, *108*, No. 102080.
(38) 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, N. T. A Tribute to Disorder in the Genome of the Bloom-Forming Freshwater Cyanobacterium Microcystis Aeruginosa. *PLoS One* **2013**, *8* (8), No. e70747.
(39) Allen, M. M. Simple Conditions for Growth of Unicellular Blue-Green Algae on Plates1, 2. *Journal of Phycology* **1968**, *4* (1), 1−4.
(40) Edgar, R. C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32* (5), 1792−1797.
(41) Huelsenbeck, J. P.; Ronquist, F. MRBAYES: Bayesian Inference of Phylogenetic Trees. *Bioinformatics* **2001**, *17* (8), 754−755.
(42) Jacob, R. T.; Larsen, M. J.; Larsen, S. D.; Kirchhoff, P. D.; Sherman, D. H.; Neubig, R. R. MScreen: an integrated compound management and high-throughput screening data storage and analysis system. *J. Biomol Screen* **2012**, *17*, 1080−1087.
(43) Segawa, T.; Takeuchi, N.; Fujita, K.; Aizen, V. B.; Willerslev, E.; Yonezawa, T. Demographic Analysis of Cyanobacteria Based on the Mutation Rates Estimated from an Ancient Ice Core. *Heredity* **2018**, *120* (6), 562−573.
(44) Lad, A.; Su, R. C.; Breidenbach, J. D.; Stemmer, P. M.; Carruthers, N. J.; Sanchez, N. K.; Khalaf, F. K.; Zhang, S.; Kleinhenz, A. L.; Dube, P.; Mohammed, C. J.; Westrick, J. A.; Crawford, E. L.; Palagama, D.; Baliu-Rodriguez, D.; Isailovic, D.; Levison, B.; Modyanov, N.; Gohara, A. F.; Malhotra, D.; Haller, S. T.; Kennedy, D. J. Chronic Low Dose Oral Exposure to Microcystin-LR Exacerbates Hepatic Injury in a Murine Model of Non-Alcoholic Fatty Liver Disease. *Toxins* **2019**, *11* (9), 486.
(45) Lad, A.; Hunyadi, J.; Connolly, J.; Breidenbach, J. D.; Khalaf, F. K.; Dube, P.; Zhang, S.; Kleinhenz, A. L.; Baliu-Rodriguez, D.; Isailovic, D.; Hinds, T. D.; Gatto-Weis, C.; Stanoszek, L. M.; Blomquist, T. M.; Malhotra, D.; Haller, S. T.; Kennedy, D. J. Antioxidant Therapy Significantly Attenuates Hepatotoxicity Following Low Dose Exposure to Microcystin-LR in a Murine Model of Diet-Induced Non-Alcoholic Fatty Liver Disease. *Antioxidants* **2022**, *11* (8), 1625.
(46) Pearson, L. A.; Dittmann, E.; Mazmouz, R.; Ongley, S. E.; D'Agostino, P. M.; Neilan, B. A. The Genetics, Biosynthesis and Regulation of Toxic Specialized Metabolites of Cyanobacteria. *Harmful Algae* **2016**, *54*, 98−111.
(47) Christiansen, G.; Molitor, C.; Philmus, B.; Kurmayer, R. Nontoxic Strains of Cyanobacteria Are the Result of Major Gene Deletion Events Induced by a Transposable Element. *Mol. Biol. Evol.* **2008**, *25* (8), 1695−1704.
(48) Giovannoni, S. J.; Tripp, H. J.; Givan, S.; Podar, M.; Vergin, K. L.; Baptista, D.; Bibbs, L.; Eads, J.; Richardson, T. H.; Noordewier, M.; Rappé, M. S.; Short, J. M.; Carrington, J. C.; Mathur, E. J. Genome Streamlining in a Cosmopolitan Oceanic Bacterium. *Science* **2005**, *309* (5738), 1242−1245.
(49) Lynch, M. Streamlining and Simplification of Microbial Genome Architecture. *Annu. Rev. Microbiol.* **2006**, *60* (1), 327−349.
(50) Lee, M.-C.; Marx, C. J. Repeated, Selection-Driven Genome Reduction of Accessory Genes in Experimental Populations. *PLOS Genetics* **2012**, *8* (5), No. e1002651.
(51) Puddick, J.; Prinsep, M. R.; Wood, S. A.; Cary, S. C.; Hamilton, D. P. Modulation of Microcystin Congener Abundance Following Nitrogen Depletion of a Microcystis Batch Culture. *Aquatic Ecology* **2016**, *50* (2), 235−246.
(52) Palagama, D. S. W.; Baliu-Rodriguez, D.; Snyder, B. K.; Thornburg, J. A.; Bridgeman, T. B.; Isailovic, D. Identification and Quantification of Microcystins in Western Lake Erie during 2016 and 2017 Harmful Algal Blooms. *Journal of Great Lakes Research* **2020**, *46* (2), 289−301.
(53) Zastepa, A.; Pick, F. R.; Blais, J. M. Fate and Persistence of Particulate and Dissolved Microcystin-LA from Microcystis Blooms. *Human and Ecological Risk Assessment: An International Journal* **2014**, *20* (6), 1670−1686.
(54) Cao, D.-D.; Zhang, C.-P.; Zhou, K.; Jiang, Y.-L.; Tan, X.-F.; Xie, J.; Ren, Y.-M.; Chen, Y.; Zhou, C.-Z.; Hou, W.-T. Structural Insights into the Catalysis and Substrate Specificity of Cyanobacterial Aspartate Racemase McyF. *Biochem. Biophys. Res. Commun.* **2019**, *514* (4), 1108−1114.
(55) Sielaff, H.; Dittmann, E.; Tandeau De Marsac, N.; Bouchier, C.; Von Döhren, H.; Borner, T.; Schwecke, T. The mcyF Gene of the Microcystin Biosynthetic Gene Cluster from Microcystis Aeruginosa Encodes an Aspartate Racemase. *Biochem. J.* **2003**, *373* (3), 909−916.
(56) Baran, R.; Brodie, E. L.; Mayberry-Lewis, J.; Hummel, E.; Da Rocha, U. N.; Chakraborty, R.; Bowen, B. P.; Karaoz, U.; Cadillo-Quiroz, H.; Garcia-Pichel, F.; Northen, T. R. Exometabolite Niche Partitioning among Sympatric Soil Bacteria. *Nat. Commun.* **2015**, *6* (1), 1−9.
(57) Phelan, V. V.; Liu, W.-T.; Pogliano, K.; Dorrestein, P. C. Microbial Metabolic Exchange—the Chemotype-to-Phenotype Link. *Nat. Chem. Biol.* **2012**, *8* (1), 26−35.
(58) Zelezniak, A.; Andrejev, S.; Ponomarova, O.; Mende, D. R.; Bork, P.; Patil, K. R. Metabolic Dependencies Drive Species Co-Occurrence in Diverse Microbial Communities. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (20), 6449−6454.
(59) Moffitt, M. C.; Neilan, B. A. Characterization of the Nodularin Synthetase Gene Cluster and Proposed Theory of the Evolution of Cyanobacterial Hepatotoxins. *Appl. Environ. Microbiol.* **2014**, *70* (11), 6353−6362.
(60) Fewer, D. P.; Metsä-Ketelä, M. A Pharmaceutical Model for the Molecular Evolution of Microbial Natural Products. *FEBS Journal* **2020**, *287* (7), 1429−1449.
(61) Gu, L.; Wang, B.; Kulkarni, A.; Geders, T. W.; Grindberg, R. V.; Gerwick, L.; Håkansson, K.; Wipf, P.; Smith, J. L.; Gerwick, W. H.; Sherman, D. H. Metamorphic Enzyme Assembly in Polyketide Diversification. *Nature* **2009**, *459* (7247), 731−735.
(62) Simonsen, A. K. Environmental Stress Leads to Genome Streamlining in a Widely Distributed Species of Soil Bacteria. *ISME J.* **2022**, *16* (2), 423−434.
(63) Grzymski, J. J.; Dussaq, A. M. The Significance of Nitrogen Cost Minimization in Proteomes of Marine Microorganisms. *ISME J.* **2012**, *6* (1), 71−80.
(64) Chen, Y.-M.; Lee, T.-H.; Lee, S.-J.; Huang, H.-B.; Huang, R.; Chou, H.-N. Comparison of Protein Phosphatase Inhibition Activities and Mouse Toxicities of Microcystins. *Toxicon* **2006**, *47* (7), 742−746.

(65) Su, R. C.; Lad, A.; Breidenbach, J. D.; Kleinhenz, A. L.; Modyanov, N.; Malhotra, D.; Haller, S. T.; Kennedy, D. J. Assessment of Diagnostic Biomarkers of Liver Injury in the Setting of Microcystin-LR (MC-LR) Hepatotoxicity. *Chemosphere* **2020**, *257*, No. 127111.
(66) Solter, P.; Liu, Z.; Guzman, R. Decreased Hepatic ALT Synthesis Is an Outcome of Subchronic Microcystin-LR Toxicity. *Toxicol. Appl. Pharmacol.* **2000**, *164* (2), 216−220.

https://doi.org/10.1021/acs.est.4c00039
Environ. Sci. Technol. 2024, 58, 19936−19947