Ecophysiological Examination of the Lake Erie Microcystis Bloom in 2014: Linkages between Biology and the Water Supply Shutdown of Toledo, OH


†Department of Biology, James Madison University, Harrisonburg, Virginia 22807, United States
‡NOAA−GLERL, 4840 South State Rd., Ann Arbor, Michigan 48108, United States
§Department of Microbiology, University of Tennessee, Knoxville, Tennessee 37996, United States
∥Department of Biological Sciences, Bowling Green State University, Bowling Green, Ohio 43403, United States
⊥Department of Chemistry, State University of New York, Environmental Science and Forestry, Syracuse, New York 13210, United States
#Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, Michigan 48108, United States
∇Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, Michigan 48109, United States
○Genomic Service Laboratory, Hudson Alpha Institute for Biotechnology, Huntsville, Alabama 35806, United States
§NOAA National Ocean Service, National Centers for Coastal Ocean Sciences, Silver Spring, Maryland 20910, United States
∞Department of Life Sciences, Texas A&M Corpus Christi, Corpus Christi, Texas 78412, United States

Supporting Information

ABSTRACT: Annual cyanobacterial blooms dominated by Microcystis have occurred in western Lake Erie (U.S./Canada) during summer months since 1995. The production of toxins by bloom-forming cyanobacteria can lead to drinking water crises, such as the one experienced by the city of Toledo in August of 2014, when the city was rendered without drinking water for >2 days. It is important to understand the conditions and environmental cues that were driving this specific bloom to provide a scientific framework for management of future bloom events. To this end, samples were collected and metatranscriptomes generated coincident with the collection of environmental metrics for eight sites located in the western basin of Lake Erie, including a station proximal to the water intake for the city of Toledo. These data were used to generate a basin-wide ecophysiological fingerprint of Lake Erie Microcystis populations in August 2014 for comparison to previous bloom communities. Our observations and analyses indicate that, at the time of sample collection, Microcystis populations were under dual nitrogen (N) and phosphorus (P) stress, as genes involved in scavenging of these nutrients were being actively transcribed. Targeted analysis of urea transport and hydrolysis suggests a potentially important role for exogenous urea as a nitrogen source during the 2014 event. Finally, simulation data suggest a wind event caused microcystin-rich water from Maumee Bay to be transported east along the southern shoreline past the Toledo water intake. Coupled with a significant cyanophage infection, these results reveal that a combination of biological and environmental factors led to the disruption of the Toledo water supply. This scenario was not atypical of reoccurring Lake Erie blooms and thus may reoccur in the future.

INTRODUCTION

The threat posed by cyanobacterial harmful algal blooms (cHABs) to freshwater ecosystems is well documented.1 Accumulation of nuisance biomass, hypoxic zones, reduction in water clarity, and the production of cyanobacterial toxins (microcystins, anatoxins) are all consequences of freshwater cHABs.1 Decades of research have shown that nutrient loading is likely the primary driver of bloom development. Phosphorus, in particular, is often suggested to be the principal limiting nutrient for primary production in many fresh waters.3 Conventional management strategies have thus focused on phosphorus load...
MATERIALS AND METHODS

Water Collection. On 4 August, 2014, water samples were collected from NOAA Great Lakes Environmental Research Laboratory’s (GLERL) eight weekly water quality monitoring sites throughout the western basin of Lake Erie, including at the Toledo Water intake (WE12, Figure 1) as well as two additional stations 1–2 miles east and west of the intake (EOI and WOI, respectively; Figure 1). At each site, integrated 0.5–1.5 m water was collected using a 1 m long Niskin bottle and served as the surface sample. This depth range was chosen because previous data has shown that 0.5 m below the surface is below any surface scum formation and it allowed for the collection of additional depth discrete samples throughout the water column if warranted. The depth has been consistent for all sites and years. Once all of the sites had been sampled, the water was kept cool and transported to NOAA-GLERL for processing of total and dissolved nutrients, chlorophyll (chl) a and particulate (intracellular) microcystins within 8 h of collection. The RNA samples were processed, as described below, immediately onboard the vessel following the completion of the water collection at each station.

Nutrients. For total phosphorus (TP) samples, duplicate 50 mL aliquots of whole lake water were collected into acid-washed glass culture tubes and stored at 4 °C until analysis within 1 week. For dissolved nutrients, duplicate whole water samples were collected in a triple rinsed (ultrapure water) 20 mL syringe and filtered through a 0.22 μm nylon filter, after a 3 mL rinse of the...
filter with whole lake water, into a 15 mL collection tube and stored at −20 °C until analysis. Nutrient concentrations were determined using standard automated colorimetric procedures as modified by Davis and Simmons on a QuAAtro AutoAnalyzer (Seal Analytical Inc., Mequon, WI) according to methods detailed by manufacturer and in compliance with EPA Methods 365.4, 350.1, and 353.1. NH₄ was determined by the Berthelot reaction in which ammonium ions react with salicylate and free chlorine to form a blue-green colored complex. NO₃+NO₂ was determined by the cadmium reduction method. SRP was determined by the molybdate/ascorbic acid method. TP and TDP used the same analysis following a persulfate digestion adapted from Menzel and Corwin. SiO₂ was determined by the reduction of a silico molybdate in an ascorbic acid solution to moldbenum blue.

Chlorophyll a. Chl a biomass was measured by concentrat- ing lake water on a 47 mm diameter GF/F filter (Whatman, GE Healthcare Bio-Sciences, Pittsburgh, PA). Samples were extracted with N,N-dimethylformamide under low light levels and analyzed with a 10 AU fluorometer (Turner Designs).

Particulate Microcystins: ELISA Assay. Particulate microcystins (MCs) were measured by filtering whole lake water onto a 25 mm, 3 μm polycarbonate membrane and kept at −20 °C until analysis. Particulate MCs were extracted from samples using a combination of physical and chemical lysis techniques. All samples were resuspended in 1 mL molecular grade water (pH 7; Sigma-Aldrich, St. Louis, MO) and subjected to three freeze/thaw cycles before the addition of the QuikLyse reagents (Abraaxis LLC; Warminister, PA) as per the manufacturer’s instructions. The samples were then centrifuged for 5 min at 2000g to pellet cellular debris. The concentrations of microcystins (reported as microcystin-LR equivalents) were measured using a microcystin enzyme-linked immunosorbent assay (Abraaxis LLC) following the methodologies of Fischer et al. This assay is largely congener-independent as it detects the ADDA moiety, which is found in almost all MCs. These analyses yielded a detection limit of 0.04 μg/L.

LC-MS and HPLC-PDA Cyanotoxin Analysis. Duplicate samples were filtered onto 47 mm GF/C filters (nominal pore size 1.2 μm) for additional toxin analysis via LC-MS and high-performance liquid chromatography with photodiode array detection (HPLC-PDA). Samples were immediately frozen at −20 °C until analysis. The filters were extracted in 50% methanol using ultrasound and clarified by centrifugation. Concentrations and congener ratios of microcystins were analyzed using methods detailed by Boyer et al. Briefly, concentrations of microcystins were determined using an LC-MS screening method against 14 common congeners (RR, dRR, mRR, YR, LR, mL, dLR, AR, FR, LA, LW, LF, WR, and NOD-R). Microcystins were also analyzed by high-performance liquid chromatography with photodiode array detection (HPLC-PDA) to detect other congeners for which we did not have standards. HPLC-PDA should detect any congener containing the ADDA group in high enough concentrations. Anatoxin-a, homoaatox- in-a, cylindrospermopsin, and deoxyxylindospermopsin presence was screened using a method for LC-MS and if present, confirmed by LC-MS/MS. The presence of BMAA (free) was screened for using LC-MS. Detection limits for each method were calculated from the instrument detection limits that day incorporating sample volumes provided with the sample. The smaller the sample volume provided, the higher our overall detection limit per liter starting water.

RNA Extraction and Sequencing. Seston was collected on Sterivex cartridge filters (0.22 μm; EMD Millipore, Billerica, MA) and stored at −80 °C prior to extraction. RNA was extracted using the MoBio DNA isolation kit for Sterivex modified for RNA (MO BIO Laboratories, Inc., Carlsbad, CA). To optimize the protocol, Sterivex were vortexed for 5 min longer than recommended and all wash buffers were allowed to sit for 1 min before being pulled through the binding column using a vacuum manifold. DNase treatment was performed as recommended in the protocol using the MoBio On-Spin Column DNase kit. This protocol was optimized by allowing the DNase solution to sit for 15 min longer than recommended. RNA was checked for DNA contamination using universal 16S primers (27F and 1522R). Any additional DNase treatments needed were performed using the Turbo DNase kit (Thermo Fisher Scientific, Waltham, MA). RNA was stored at −80 °C until sent to HudsonAlpha Institute for Biotechnology (Huntsville, AL) for sequencing. Total RNA concentrations and quality were assessed fluorometrically via RiboGreen (Life Technologies, Carlsbad, CA) followed by integrity measurement via Bioanlysis (Agilent Technologies, Santa Clara, CA). Ribosonal RNA reduction was done using the Illumina Ribo-Zero Epidemiology rRNA removal kit (San Diego, CA) followed by first- and second-strand cDNA synthesis (New England Biolabs, Ipswitch, MA) and library preparation (Kapa Biosystems, Wilmington, MA). Sequencing was done on the Illumina HiSeq platform for 100-bp paired-end sequencing by the HudsonAlpha Genomic Services Laboratory.

Transcriptome Analysis. Targeted analysis of cyanobacterial populations was performed using the genomes of the model organisms M. aeruginosa NIES 843, Anabaena cylindrica PCC 7122 (NC_019771.1- chromosome), and Planktothrix agardhii NVA CYA 126/8 (CM002803.1- chromosome). While some members of the genus Anabaena were recently reclassified as Dolichospermum,20,21 we will use Anabaena for consistency with the model genome and with previous analyses in this system.20 Fastq files were imported into CLC Genomics Workbench v8.0 (Qiagen, Redwood City, CA) using default quality settings, with all failed reads discarded prior to downstream analysis (SI Table S1). RNA-Seq Analysis within the Transcriptomics module was used for mapping and calculation of expression values. Paired-end reads from two separate lanes per sample were pooled for this analysis. Duplicate sequence libraries were generated for all sites excluding WE06 and WE08, which only had single libraries due to loss of biological samples. Analysis was performed as previously described.20 Only those reads that mapped non-redundantly to a single cyanobacterial genome were considered for expression analysis to exclude potential false signals from highly conserved genes (Table S2). Expression values were calculated from the number of reads mapped to each gene within the model genomes, and then normalized per 1 000 000 reads to generate the expression value of Total Counts per Million (TCPM).

For community analyses, reads were assembled into contigs using the CLC Genomics Workbench de novo assembly function. A minimum contig length of 200 bp was used, with all contigs below this threshold disregarded. Contigs were uploaded into the MG-RAST pipeline for analysis. For identity annotation, the M5nr database was used for Best Hit Annotation and the default identity increased to 65%. Functional annotation was performed using the SEED database, again increasing default identity to 65%. Paired end reads were
mapped back to contigs to assess how well assemblies represented the sequence libraries (SI Table S1).

All comparative analyses were performed in the Primer 7.0.10 (Primer-e, Quest Res Ltd. Auckland, NZ) statistical package. Clustering was generated through Bray–Curtis resemblance analysis and subsequent clustering using complete linkage. Statistical analysis of Southern Shore (SS) vs Off Shore (OS) populations was performed in CLC Genomics Workbench using Baggerly’s test. For all statistical tests, a p-value of p < 0.05 was used to indicate significance. Raw sequences are available from the NCBI sequence read archive under SRR094616, and contigs are available from MG-RAST under Project ID 17333.

Toxin Simulation Analysis. To simulate the effect of hydrodynamic transport on the distribution of microcystins in western Lake Erie between 21 July and 4 August, 2014, we used the Lagrangian particle dispersion (LPD) model described by Rowe et al. which considers 3D advection and random-walk vertical mixing of buoyant particles. The LPD was forced by 3D simulated Lagrangian particle positions, under the assumption that subsequent microcystin concentrations were calculated from within 12 km of stations by nearest neighbor interpolation; concentrations measured in western Lake Erie on 21 July 2014 concentration was initialized in proportion to Community Ocean Model (FVCOM). Lagrangian particle currents and vertical turbulent diffusivity from Finite Volume Community Ocean Model (FVCOM). Lagrangian particle concentration was initialized in proportion to Microcystis concentrations measured in western Lake Erie on 21 July 2014 within 12 km of stations by nearest neighbor interpolation; subsequent microcystin concentrations were calculated from simulated Lagrangian particle positions, under the assumption that each Lagrangian particle represented a fixed mass of microcystin. The LPD model simulation considered transport only; biochemical production and loss of microcystins were not simulated.

RESULTS AND DISCUSSION

Bloom Conditions. The 2014 Lake Erie bloom received unprecedented public attention due to the detection of microcystins in the finished water supply of the city of Toledo, OH in August. Gobler et al. describes the seasonal trends in algal pigments, particulate microcystins, nitrate and SRP concentrations from data collected at NOAA GLERL’s core monitoring stations (WE2, WE4, WE6, WE8) from 2012–2014. Overall, Gobler et al. showed that while the 2012 bloom was spatially smaller than either 2013 or 2014, the basin averaged phyocyanin concentrations peaked higher than in 2013 or 2014. Furthermore, basin-averaged particulate MC concentrations were higher in 2013 and 2014 than in 2012, with 2014 peaking at nearly twice the concentrations of 2013 and an order of magnitude higher than 2012. In light of this recent synthesis, we will only briefly describe seasonal trends in relation to our focused analysis of the period surrounding the Toledo water crisis. To understand whether conditions at the water intake were aligned to yield a bloom of particularly high toxicity, we processed samples of opportunity collected on August 4, 2014 (stations denoted as WE02, WE04, WE06, WE08, WE12—the site of the Toledo water intake crib, WE13, WOI, EOI) from the western basin of Lake Erie in response to the Toledo do-not-drink advisory that had been announced 2 days prior (Table 1, Figure 1). Overall, the average Chl-a concentration (46.1 μg/L) for 29 July, the week prior to the 4 August sampling event in 2014, was higher than the basin averages for 2012 and 2013 (14.8 and 22.4 μg/L, respectively). The increase in basin average was not due to a basin-wide increase in Chl a biomass but due to an increase in pigments at the Maumee Bay station (WE6) in 2014 (126.1 μg/L) compared to 35.6 μg/L and 15.2 μg/L for 2012 and 2013, respectively. On 4 August, Chl a biomass, in conjunction with the modeling results described below, clearly show that Maumee Bay was flushed as Chl a concentration at WE6 decreased to 71.6 μg/L while Chl a biomass increased at all other stations. The basin-averaged particulate microcystins, as measured by ELISA, showed similar trends as the Chl a data described above with microcystins at WE6 decreasing from 37.1 μg/L to 10.1 μg/L from 29 July to 4 August, and providing more evidence of a Maumee Bay flushing event. The most abundant microcystin congeners found in the 4 August samples were microcystin-LR (65–85%), microcystin-RR (15–30%) and microcystin-YR (10–15%). These data are similar to seasonal trends that showed microcystin-LR was detected on every date sampled from July through August 2014. Microcystins RR and YR were also detected frequently throughout the 2014 sampling period but were more sporadic during July as the bloom was developing (data not shown). Similar to the data collected on 4 August, microcystin-LR was always the most prevalent followed by microcystin-RR then microcystin-YR. Measurable concentrations of anatoxin-a (0.06 μg/L) occurred at WE2 on 29 July. Cylindrospermopsin (CYN) was detected three times during 2014 (8, 14, and 21 July) but these could not be quantified due to lack of a CYN standard. Furthermore, during the 4 August sampling, no other cyanobacterial toxins, other than microcystins, were detected (<0.01 μg/L) at any of our sampling sites.

Cyanobacterial Physiological Ecology. To assess the ecophysiological status of bloom communities, shotgun metatranscriptomes were generated from total mRNA extracted

<table>
<thead>
<tr>
<th>station</th>
<th>latitude/longitude</th>
<th>nitrate (mg/L)</th>
<th>ammonia (mg/L)</th>
<th>PON (mg/L)</th>
<th>TN (mg/L)</th>
<th>TP (μg/L)</th>
<th>Chl a (μg/L)</th>
<th>PMCS-ELISA (μg/L)</th>
<th>PMCS-LCMS (μg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WE02</td>
<td>41 45.912/83 19.835</td>
<td>0.56</td>
<td>0.0030</td>
<td>0.32</td>
<td>0.88</td>
<td>31.98</td>
<td>29.44</td>
<td>4.33</td>
<td>0.91</td>
</tr>
<tr>
<td>WE04</td>
<td>41 49.714/83 11.654</td>
<td>0.14</td>
<td>0.0016</td>
<td>0.24</td>
<td>0.38</td>
<td>22.09</td>
<td>18.11</td>
<td>1.54</td>
<td>0.52</td>
</tr>
<tr>
<td>WE06</td>
<td>41 42.679/83 22.631</td>
<td>0.29</td>
<td>0.0018</td>
<td>0.86</td>
<td>1.15</td>
<td>62.19</td>
<td>71.62</td>
<td>10.14</td>
<td>3.45</td>
</tr>
<tr>
<td>WE08</td>
<td>41 50.254/83 21.823</td>
<td>0.09</td>
<td>0.0013</td>
<td>0.41</td>
<td>0.50</td>
<td>40.07</td>
<td>46.53</td>
<td>4.31</td>
<td>1.10</td>
</tr>
<tr>
<td>WE12</td>
<td>41 42.157/83 15.781</td>
<td>0.62</td>
<td>0.0033</td>
<td>0.68</td>
<td>1.30</td>
<td>44.37</td>
<td>54.46</td>
<td>9.28</td>
<td>2.91</td>
</tr>
<tr>
<td>WE13</td>
<td>41 44.539/83 08.286</td>
<td>0.22</td>
<td>0.0020</td>
<td>0.14</td>
<td>0.35</td>
<td>14.98</td>
<td>6.74</td>
<td>1.17</td>
<td>0.40</td>
</tr>
<tr>
<td>WOI</td>
<td>41 42.913/83 19.722</td>
<td>0.58</td>
<td>0.0026</td>
<td>0.56</td>
<td>1.15</td>
<td>60.24</td>
<td>51.20</td>
<td>8.19</td>
<td>1.10</td>
</tr>
<tr>
<td>EOI</td>
<td>41 40.561/83 14.329</td>
<td>0.54</td>
<td>0.0027</td>
<td>0.74</td>
<td>1.29</td>
<td>53.01</td>
<td>69.44</td>
<td>10.58</td>
<td>1.24</td>
</tr>
</tbody>
</table>
from samples collected at each station. Recruitment of transcripts (SI Table S2) to model cyanobacterial genomes of *M. aeruginosa* NIES 843, *P. agardhii* NVA CYA 128/6, and *A. cylindrica* PCC 7122 revealed a clear dominance of *Microcystis* over other cyanobacteria across the western basin (Figure 1, SI Table S2), a pattern also observed from DNA samples collected during this time period. The percentage of reads mapped to *Microcystis* was greatest at station WE13 (47.4%) and WE06 (46.2%) (Figure 1). Notably, the fewest reads mapped to *Microcystis* at the site of the Toledo water intake (station WE12, 24.4%) and station WE08 (29.8%) (Figure 1). Both *Planktothrix* and *Anabaena* appear to have made only a minor contribution to total community expression, as they comprise less than 3.5% of total mRNA at each station indicating *Microcystis* was the dominant potentially toxic-cyanobacterium at all sites. This result is similar to Harke et al. (2016), who also showed this pattern in the open waters of western Lake Erie.

Across all stations in this study, only minor deviations in gene expression for the chosen *Microcystis* model were noted (Figure 2A). Cluster analysis of genome-wide expression revealed two distinct groups: Stations WE12, WE08, and WOI, and EOI clustered together, as did Stations WE02, WE04, WE06, and WE13 (Figure 2A). A much finer examination was necessary to resolve the difference in transcriptional response to environmental conditions between sites. A series of 47 genes involved in nitrogen (N) and phosphorus (P) metabolism were selected for analysis to examine active nutrient metabolism by *Microcystis* (Figure 2B). Based on these expression profiles, it appears that *Microcystis* populations were experiencing both N- and P-stress at the time of sampling, as cells were actively transcribing genes indicative of nutrient stress, including those involved in the transport of phosphate, ammonium, nitrate/nitrite, and urea (Figure 2B). The most highly expressed gene across all sites is involved in phosphorus acquisition: MAE_18310 encodes the substrate binding component of a phosphate transporter. Outside of this single gene, two small clusters of genes showed higher relative transcription when compared to the others: these were involved in both N and P acquisition and metabolism (Figure 2B). Cluster I includes genes involved in the transport and metabolism of both nitrate and nitrite, as well as urea metabolism (Figure 2B). Cluster II was even more highly transcribed, and encode proteins involved in the high-affinity transport of phosphate (*pstS*, *pstC*), ammonium (MAE_40010), and urea (MAE_06220) (Figure 2B). Increased transcription of *pstS* has previously been demonstrated under P-stress in culture and has been shown to be up-regulated in western Lake Erie during time of low-P concentrations, as well.
Urea has recently become of interest to the study of chABs due to its nearly ubiquitous usage as an N-rich fertilizer and its presence in aquatic systems.35−38 Initial research efforts have indicated that urea-rich waters may preferentially select for organisms such as Microcystis, even when P is abundant.39 Based on the expression patterns of genes in this proposed cyanobacterial metabolic network, it appears urea is a key nutrient in terms of its ability to shape cell physiology in the natural environment, something that has previously been suggested in culture (Figure 3).21,40,41

The arginases that produce urea (ARG1, ARG2) during the conversion of arginine to ornithine in the cellular urea cycle appear to be effectively inactive, with only weak transcription by Microcystis of both genes encoding putative arginases at all stations (Figure 3). This suggests import of extracellular urea from the environment was a much larger contributor to the cellular pool of urea than internal biological production, supporting the accumulating evidence for a role of urea as a driver for the success of Microcystis in systems such as Lake Erie.6,38,43 Taken together, these data provide strong evidence that Microcystis cells were actively scavenging both N and P during this time (Figures 2B and 4), an observation in line with other recent surveys of bloom populations across the lake.38

Community Function within a Microcystis Bloom. The composition of the microbial community within Lake Erie is variable, particularly with regards to the dominant cyanobacteria across the western basin. For example, despite the widespread distribution of Microcystis, the filamentous organism Planktothrix dominates toxic populations in Sandusky Bay.44,45 Understanding the interactions between Microcystis, other toxic bloom forming cyanobacteria, and the environment is critical in the development of future mitigation strategies. Moreover, mounting evidence has demonstrated that toxin producing cyanobacteria interact with other members of the co-occurring microbial community,45,46 and that while there is phylogenetic variation across locations, biological functions are often conserved.46,47 Hierarchical classification of assembled contigs using the SEED database26 support this observation for the August 2014 Lake Erie populations. The second most abundant bacterial phylum after the Cyanobacteria was Proteobacteria (Figure 4A). Compared to cyanobacterial populations, the Proteobacteria produced fewer transcripts assigned to five functional categories (Figure 4B, C). However, the Proteobacteria had comparatively increased function in the protein metabolism, respiration, and stress response categories. The functional similarities across stations suggest that environmental conditions and the functional response of the bloom community to these conditions were not unique to station WE12 in August of 2014, but rather common to multiple sites across Lake Erie. This
result is not surprising as transport of bloom biomass throughout western Lake Erie is an ongoing process therefore similarities between sampling sites is expected.

In addition to interactions with other bacteria, transcriptional evidence suggests viral activity was significant at the time of sample collection. Previously we have demonstrated that the signatures of dsDNA phage infecting *Microcystis* are present in Lake Erie metatranscriptomes, implying active, ongoing infections of the community. To determine whether viral effects were shaping the community, we examined the occurrence of the transcript from *gp91* (which encodes the *Microcystis* phage tail sheath protein) relative to a conserved
marker for active *Microcystis* cell density (*rpoB*) (Figure 5). Surprisingly, the occurrence of virus transcripts relative to the host marker occurred at ∼1:1 ratio across the near-shore during the *Microcystis* bloom in August 2014. From samples collected concurrently at offshore sites (W4, W13), this relationship decreased by 2 orders of magnitude. And for samples collected at similar near-shore locations just 3 weeks later, the same low-level of virus activity was observed. Indeed, in looking at historical samples from this region and calendar period in previous years, the low active-infection relationship was observed in 2012 whereas in samples collected in 2013, we did not detect any signatures for this virus (Figure 5). The presence of these virus signatures raises the intriguing hypothesis that lysis of cells may facilitate movement of microcystins from the particulate to dissolved phase, elevating the opportunity for dissolved microcystins to enter water treatment facilities such as Toledo.

**Influence of Hydrodynamic Transport on the Spatial Distribution of *Microcystis***. Currents in the southern half of the western basin of Lake Erie are generally weak during summer, but wind events can cause alongshore currents that are capable of rapidly changing the spatial distribution of CChABs. We used a hydrodynamic model to visualize the influence of currents and vertical mixing on the spatial and vertical distribution of buoyant *Microcystis* colonies (and associated particulate microcystins) preceding the August 2014 incident at the Toledo water intake. The spatial distribution of particulate microcystins was initialized in the model by interpolation of values measured at six stations on 21 July; a particulate microcystin concentration of 20 μg/L was observed at station WE06 in Maumee Bay, and 6 μg/L at station WE12 near the Toledo water intake, with lower values to the north (Figure 6A). The model simulation predicted the spatial distribution of particulate microcystins on subsequent days, as it was modified by advection and vertical mixing of buoyant *Microcystis* colonies. Three days later on 24 July, weak transport resulted in little movement (Figure 6B). However, on 28 July, strong wind from the northwest (9.6 m/s at Toledo Harbor Light) produced currents in the model that flushed microcystin-rich water from Maumee Bay eastward along the southern shoreline (Figure 6C). Furthermore, the northerly winds constrained the biomass along the south shore as it moved eastward. Simulated *Microcystis* colonies were well-mixed through the water column on 29 July and through 1 August (SI Figure S5). After 1 August, lighter winds (2.5–4.1 m/s) allowed buoyant *Microcystis* colonies to accumulate within 1–2 m of the surface in the model, resulting in elevated surface concentrations (Figure 6D vs 6C). Observed microcystin concentrations on 4 August had decreased to 10 μg/L at WE6 (Maumee Bay) which was a 50% decrease from the previous week providing further evidence that Maumee Bay was flushed as previously discussed. Concentrations of microcystins were 8–11 μg/L at three stations along the southern shoreline with lower values to the north, consistent with the transport pattern indicated by the model (Figure 6D).

**Geographic Partitioning of Sites.** To determine whether transcriptional signals in *Microcystis* populations were consistent with the estimated spatial pattern of microcystins concentration produced at the end of the hydrodynamic simulation (Figure 6D), stations were divided into Southern Shore (SS: WE06, WE12, WOI, EOI) or Off Shore (OS: WE02, WE04, WE08) groups and the transcriptional fingerprints of the *Microcystis* population were compared between the two groups of stations (SI Figure S4A). A total of 73 genes were differentially expressed (*p* < 0.05) between the two groups. Of these 73 genes, 57 (78%)
had significantly more transcripts detected at SS stations compared to OS stations, and 16 (22%) were significantly overrepresented at OS stations (SI Figure S4A). While a majority (56%) of these genes are annotated as “hypothetical”, there are several genes of known function represented. These include four genes involved in construction of gas vesicle proteins (gvpAI, gvpJ, gvpK, gvpN) (SI Figure S4B).

The genes responsible for encoding the gas vesicle proteins in *Microcystis* have been identified, although only a subset are fully characterized.50,51 The primary structural genes are *gvpA*, which encodes the primary component of the vesicle wall, and *gvpC*, which strengthens the protein wall encoded by *gvpA*.50 In this study, gas vesicle genes gvpAI, gvpJ, gvpK, and gvpN were all significantly upregulated at OS stations, indicating increased transcription of gas vesicle genes in these populations compared to SS populations (SI Figure S4B).

**Comparing 2014 to Previous Years.** The transcript profiles from the 2014 WE12 *Microcystis* population were compared to similar transcript profiles of bloom populations collected from the Environment and Climate Change Canada station 973 (41°47′30″ N, 83°19′58″ W), located in the western basin in 2012 and 2013. As a “low bloom” year,52 2012 was included to serve as a contrasting population to the 2014 sample.12,20 Large differences in *Microcystis* gene expression exist between years (SI Figure S3). However, there are several factors that may, at least in part, account for this. The available 2012 and 2013 samples were collected using a 20 μm mesh net, enriching for large colonies and filaments of cyanobacteria, unlike the 2014 samples, which were filtered onto 0.2 μm filters. Other differences may lie in dates of sample collection (July vs August) and onset of bloom development. These differences highlight the need to standardize sample collection for molecular analyses of bloom communities, something that has been increasingly recognized but has yet to be accomplished. A previous study determined that biomass captured in plankton nets of various mesh sizes (112 μm, 53 μm, and 30 μm) captures at least 93% of *Microcystis* biomass, and samples collected in this manner would be representative of the Lake Erie *Microcystis* community.53 As the 2012 and 2013 samples were collected using a 20 μm plankton net, we could likely identify true expression differences that are not artifacts of sample collection strategy. To further control for these differences in our analysis, we only included those *Microcystis*-specific genes that had conserved significant over- or under-represented transcripts in 2012 and 2013 when compared to WE12 2014. Compared to *Microcystis* populations sampled in 2012 and 2013, 121 genes were significantly over-represented in the 2014 samples (SI Figure S3; Table S3). These genes included 10 involved in P acquisition, implicating a stronger potential P-stress at the time of sampling in 2014 relative to the previous two summers. Interestingly, one of the *Microcystis* gvpA genes was also significantly upregulated, whereas gvpC, gvpE, gvpG, gvpJ, gvpK, and gvpN were among the 266 genes which were downregulated in 2014 relative to 2012 and 2013, suggesting gas vesicle construction was not as active in 2014. These temporal and geographic differences in transcript levels of genes regulating gas vesicle production provide further insight into how cells and/or toxin may have been introduced into the water intake of the water treatment facility.

**Lessons Learned from the Toledo Water Crisis.** Once declared a “dead lake”, the ecological status of Lake Erie improved dramatically after the implementation of phosphorus reduction strategies in the late twentieth century. However, recent re-eutrophication has received national attention, especially when microcystin concentrations in Toledo’s drinking water exceeded the World Health Organization’s provisional drinking water guideline.54 A major unanswered question remains whether the 2014 Toledo Event was a “common” bloom scenario that has the potential for a repeat event or was a singular event unique to that site. In hindsight, 2014 was a fairly typical bloom according to NOAA’s cyanobacterial index.55 Overall our observations point to three new hypotheses derived from the data generated from these samples. These hypotheses provide a framework for future empirical testing and may in fact reveal features of this and other blooms that may exacerbate introduction of *Microcystis* cells or their toxins into water supplies:

1. Based on hydrodynamic transport modeling of microcystins prior to the 2014 event, it appears that the source of the high toxicity water that entered the water intake originated from Maumee Bay and conditions were sufficient to not only flush the Bay, but to introduce toxic cells deeper into the water column
2. Populations in the western basin of Lake Erie had down-regulated a majority of their gas-vesicle production genes: given the assumed linkage between this process and cell buoyancy, *Microcystis* populations would have been less resistant to the deep mixing events described
3. A broad scale infection of the *Microcystis* community by a lytic cyanophage may have contributed to the redistribution of toxins from the particulate to dissolved phase in the system. Coupled with the mixing events that were occurring, it is likely that this event further enhanced the introduction of toxin to the water supply intake (albeit in a dissolved relative to particulate state).

Transcriptomes are often considered proxies for what the cells are “trying to do” and is the measure of function most immediately tied to environmental conditions of current ‘omics approaches. Our data suggest that the microbial community structure and functional potential at station WE12 were similar to those populations dispersed across the western basin of Lake Erie during the Toledo 2014 event. Combined with simulation and wind data, our analysis implies that while the introduction of this bloom into the Toledo water intake was site specific, the conditions which led to its occurrence were not particularly unique, other than the evidence for viral lysis. Given that lysis is likely a regular process occurring in a bloom, this suggests a strong chance that this event may recur in the future if significant changes in the ecosystem dynamics of western Lake Erie do not happen.

**ASSOCIATED CONTENT**

Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.est.7b00856.

Additional information as noted in the text (PDF)

**AUTHOR INFORMATION**

Corresponding Author
E-mail: steffemm@jmu.edu.

ORCID®
Morgan M. Steffen: 0000-0001-9755-9879
Danielle Gutierrez: 0000-0001-6355-2134

Notes
The authors declare no competing financial interest.
ACKNOWLEDGMENTS

We thank Taylor Tuttle for assistance with sampling. MDR thanks Eric Anderson of NOAA GLERL for providing hydrodynamic data from the Lake Erie Operational Forecasting System (https://tidesandcurrents.noaa.gov/ofis/ofeis/ofeis.html). GJD and KAM were supported by a grant from the Erb Family Foundation made through the University of Michigan Marine Science Center, USGS 104b Program and NOAA’s Ohio Sea Grant College Program, R/ER-104 (jointly with RMLM). The work conducted by the U.S. DOE Joint Genome Institute (RMLM), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. DOE under Contract No. DE-AC02-05CH11231. TWD was supported by Great Lakes Restoration Initiative through the U.S. Environmental Protection Agency and National Oceanic and Atmospheric Administration. THJ, MDR, AMB, and DP were supported by an award to Cooperative Institute for Limnology and Ecosystems Research (CILER) through the NOAA Cooperative Agreement with the University of Michigan (NA12OAR4320071). This is CILER contribution number 1108. RPS and TTW were partially supported by NASA Public Health and Water Quality (NNH08ZDA001N) and the NASA Ocean Biology and Biochemistry Programs (proposal 14-SMDUNSL14-0001). This work was also supported by funding from the National Science Foundation (IOS1415128, DEB1240870). CSCOR HAB Event Response Program publication #18, NOAA GLERL publication # 1856.

REFERENCES


(48) Davenport, E. J. Del Regulation of Metabolic Functions of a Western Lake Erie Microcystis Bloom Informed by Metatranscriptomic Analysis; Bowling Green State University, 2016.


