Article

Ecophysiological examination of the Lake Erie *Microcystis* bloom in 2014: linkages between biology and the water supply shutdown of Toledo, Ohio

Morgan Michelle Steffen, Timothy W. Davis, Robert Michael McKay, George S. Bullerjahn, Lauren E. Krausfeldt, Joshua M.A. Stough, Michelle L. Neitzey, Naomi E. Gilbert, Gregory L. Boyer, Thomas H. Johengen, Duane C. Gossiaux, Ashley M. Burtner, Danna Palladino, Mark Rowe, Gregory J. Dick, Kevin Meyer, Shawn Levy, Braden Boone, Richard Stumpf, Timothy Wynne, Paul V. Zimba, Danielle B. Gutierrez, and Steven W. Wilhelm

Environ. Sci. Technol., **Just Accepted Manuscript** • Publication Date (Web): 23 May 2017

Ecophysiological examination of the Lake Erie *Microcystis* bloom in 2014: linkages between biology and the water supply shutdown of Toledo, OH

Morgan M. Steffen [a*], Timothy W. Davis [b], R. Michael L. McKay [d], George S. Bullerjahn [d], Lauren E. Krausfeldt [c], Joshua M.A. Stough [c], Michelle L. Neitzey [a], Naomi E. Gilbert [a], Gregory L. Boyer [e], Thomas H. Johengen [f], Duane C. Gossiaux [b], Ashley M. Burtner [f], Danna Palladino [f], Mark D. Rowe [f], Gregory J. Dick [g], Kevin A. Meyer [g], Shawn Levy [h], Braden Boone [h], Richard P. Stumpf [i], Timothy T. Wynne [i], Paul V. Zimba [j], Danielle Gutierrez [j], and Steven W. Wilhelm [c]

a. Department of Biology, James Madison University, Harrisonburg, VA 22807;

b. NOAA–GLERL, 4840 South State Rd., Ann Arbor, MI, 48108, USA

c. Department of Microbiology, University of Tennessee, Knoxville, TN 37996;

d. Department of Biological Sciences, Bowling Green State University, Bowling Green, OH 43403

e. Department of Chemistry, State University of New York, Environmental Science and Forestry, Syracuse, NY 13210

f. Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, MI, 48108

g. Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, 48109

h. Genomic Service Laboratory, Hudson Alpha Institute for Biotechnology, Huntsville, AL 35806

i. NOAA National Ocean Service, National Centers for Coastal Ocean Sciences, Silver Spring, MD 20910

j. Department of Life Sciences, Texas A&M Corpus Christi, Corpus Christi, TX 78412

*author for correspondence: steffemm@jmu.edu

Abstract

Annual cyanobacterial blooms dominated by *Microcystis* have occurred in western Lake Erie (USA/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 re-occurring Lake Erie blooms and thus may re-occur 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[2]. 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 reductions, with some notable successes[4, 5]. This approach alone is not sufficient in all systems. Evidence now suggests that additional factors, including increases in temperature, the availability and chemistry of other nutrients such as nitrogen, and top down controls such as viral lysis have a compounding effect on the success of bloom-forming organisms[6-9]. To this end, restoration of ecosystem function may require a more comprehensive management strategy that incorporates these and additional factors[10].

There is a long history documenting the impact of cyanobacterial blooms on human uses of water resources. One of the first well-documented cases occurred on Palm Island, Australia in 1979 when a bloom of what was likely *Cylindrospermopsis raciborskii* caused an outbreak of human hepatoenteritis[11]. Other examples are specific to the consequences of toxic blooms for potable water resources. For more than two decades, annual blooms dominated by *Microcystis* spp. have plagued Lake Erie, the shallowest and most productive of the Laurentian Great Lakes[12]. In 2014, crisis arose in the city of Toledo, OH, USA, when a cyanobacterial bloom in western Lake Erie

overwhelmed the city's water treatment system, resulting in microcystins persisting in finished water at concentrations above the World Health Organization's guideline level for microcystin-LR for safe drinking water (1 ug L^{-1}). This caused the city of Toledo to issue a 'do not drink' advisory which spanned a weekend in early August, leaving >400,000 residents without access to potable water (Figure S1). Local businesses and industries were forced to purge water systems and other costly precautions in the following weeks, resulting in at least $65 million in related economic losses[13]. Conventional water treatment processes can effectively remove cyanobacterial toxins, but public water systems affected by cHABs must invest in enhanced monitoring and adjustments or upgrades to existing treatment processes that can incur multi-million dollar costs. To ensure a continuous supply of safe drinking water, a better understanding of the drivers of such events is needed for management considerations. Here we report a comprehensive survey of geochemical, ecophysiological, and hydrodynamic conditions of the western basin of Lake Erie at the time of the drinking water crisis in Toledo. Within the context of this study we linked traditional water quality parameters, algal pigment concentrations, modeling and metatranscriptomics to query the entire microbial community, with a focus on *Microcystis*, to ascertain the environmental cues (nutrient stress, etc.) to which populations were responding at the time of the 2014 Toledo bloom event.

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 (EI and WI, respectively; Figure 1). At each site, integrated 0.5 m to 1.5 m water was collected using a 1m 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 eight hours of collection. The RNA samples were processed, as described below, immediately on-board 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 one 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[14] as modified by Davis and Simmons[15] on a QuAAtro AutoAnalyzer (Seal Analytical Inc., Mequon, WI) according to methods detailed by manufacturer and is in compliance with EPA Methods 365.4, 350.1, and 353.1. NH_4 was determined by the Bethelot reaction in which

ammonium ions react with salicylate and free chlorine to form a blue-green colored complex. NO_3+NO_2 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[16]. SiO_2 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 concentrating 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 10AU fluorometer (Turner Designs) [17].

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 (Abraxis LLC; Warminster, PA) as per the manufacturer's instructions. The samples were then centrifuged for 5 min at 2,000 x *g* to pellet cellular debris. The concentrations of microcystins (reported as microcystin-LR equivalents) were measured using a microcystin enzyme-linked immunosorbent assay (Abraxis LLC) following the methodologies of Fischer et al. [18]. 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 analyzed. 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.[19] Briefly, concentrations of microcystins were determined using an LC-MS screening method against 14 common congeners (RR, dRR, mRR, YR, LR, mLR, 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, homoanatoxin-a, cylindrospermopsin and deoxycylindrospermopsin presence was screened for using LC-MS and if present, confirmed by LC-MS/MS[19]. 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 ºC 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 minutes longer than recommended and all wash buffers were allowed to sit for one

minute 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)[20]. Any additional DNase treatments needed were performed using the Turbo DNase kit (Thermo Fisher Scientific, Waltham, MA). RNA was stored at -80 ºC 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 Bioanalysis (Agilent Technologies, Santa Clara, CA). Ribosomal 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[21], *Anabaena cylindrica* PCC 7122 (NC_01977.1- chromosome), and *Planktothrix agardhii* NVA CYA 126/8 (CM002803.1- chromosome only). While some members of the genus *Anabaena* were recently re-classified as *Dolichospermum*[22, 23], 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 v.8.0 (Qiagen, Redwood City, CA) using default quality settings, with all failed reads discarded prior to downstream

analysis (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)[20]. 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[24]. For identity annotation, the M5nr database was used for Best Hit Annotation and the default identity increased to 65%[25]. Functional annotation was performed using the SEED database, again increasing default identity to 65%[26, 27]. Paired end reads were mapped back to contigs to assess how well assemblies represented the sequence libraries (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[28]. 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 SRP094616, 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.[29], which considers 3D advection[30] and random-walk vertical mixing of buoyant particles [31]. The LPD was forced by 3D 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 [9] 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.[9] showed that while the 2012 bloom was spatially smaller than either 2013 or 2014, the basin averaged phycocyanin concentrations peaked higher than in 2013 or 2014. Furthermore, basin-averaged 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, WI, EI) from the western basin of Lake Erie in response to the Toledo do-not-drink advisory that had been announced two 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 from samples collected at each station. Recruitment of transcripts (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, Table S2), a pattern also observed from DNA samples collected during this time period[32]. 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),[33] 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 WI, and EI 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). The genes in cluster II were 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 [34].

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]. *urtA*, the gene encoding the substrate (urea) binding component of the urea transporter, was highly transcribed across the western basin (Figure 3)[42]. While detected at each station, transcripts of this gene were most abundant at stations WE02, WE08, and WI (Figure 3). Notably, along with WE06, stations WE02 and WE08 were also among the stations where P stress appears to be the most severe, as the expression of two indicator genes (MAE_13810 and *pstS*) was highest. Indeed, the expression of these particular genes may have driven the overall gene clustering (Figure 2B, Supplemental Figure 1). 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 (Figure 2B, 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 database[26] 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, 4C). 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 on-going 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[20]. 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 two orders of magnitude. And for samples collected at similar near-shore locations just three weeks later, the same low-level of virus activity was observed[48]. Indeed, in looking at historical samples from this region and calendar period in previous years, the low active-infection relationship was observed in 2012[38], 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[49] that are capable of rapidly changing the spatial distribution of cHABs. We used a hydrodynamic model[50] 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[50] were well-mixed through the water column on 29 July and through 1 August (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, WI, EI) or Off Shore (OS: WE02, WE04, WE08) groups and the transcriptional fingerprints of the *Microcystis* population were compared between the two groups of stations (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 (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*) (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 (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[20] 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 (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 (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, while *gvpC, gvpF, 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 [7]. 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.

Acknowledgements

We thank Taylor Tuttle for assistance with sampling. GJD and KAM were supported by the Erb Family Foundation. GSB was supported by the Ohio Water Resources 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-SMDUNSOL14-0001). This work was also supported by funding from the National Science Foundation (IOS1451528, DEB1240870). CSCOR HAB Event Response Program publication #18, NOAA GLERL publication number 1856.

Supporting Information

The supporting information includes supplementary data and is designated with an S in the text of this publication.

Works Cited

1. 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 commonly toxic cyanobacterium, *Microcystis*. *Harmful Algae* **2016,** *54*, 4-20.
2. Charlton, M.; Milne, J., Review of thirty years of change in Lake Erie water quality. *NWRI contribution* **2004**, (04-167).
3. Schindler, D., Evolution of phosphorus limitation in lakes. *Science* **1977,** *195*, (4275), 260-262.
4. Scavia, D.; DePinto, J. V.; Bertani, I., A multi-model approach to evaluating target phosphorus loads for Lake Erie. *J Great Lakes Res* **2016,** *42,* (6), 1139-1150.
5. Scavia, D.; David Allan, J.; Arend, K. K.; Bartell, S.; Beletsky, D.; Bosch, N. S.; Brandt, S. B.; Briland, R. D.; Daloğlu, I.; DePinto, J. V.; Dolan, D. M.; Evans, M. A.; Farmer, T. M.; Goto, D.; Han, H.; Höök, T. O.; Knight, R.; Ludsin, S. A.; Mason, D.; Michalak, A. M.; Peter Richards, R.; Roberts, J. J.; Rucinski, D. K.; Rutherford, E.; Schwab, D. J.; Sesterhenn, T. M.; Zhang, H.; Zhou, Y., Assessing and addressing the re-eutrophication of Lake Erie: Central basin hypoxia. *J Great Lakes Res* **2014,** *40*, (2), 226-246.
6. Chaffin, J. D.; Bridgeman, T. B.; Bade, D. L., Nitrogen constrains the growth of late summer cyanobacterial blooms in Lake Erie. *Adv Microbiol* **2013,** *3*, 16.
7. Steffen, M. M.; Belisle, B. S.; Watson, S. B.; Boyer, G. L.; Wilhelm, S. W., Status, causes and controls of cyanobacterial blooms in Lake Erie. *J Great Lakes Res* **2014,** *40*, (2), 215-225.
8. Bullerjahn, G. S.; McKay, R. M.; Davis, T. W.; Baker, D. B.; Boyer, G. L.; D'Anglada, L. V.; Doucette, G. J.; Ho, J. C.; Irwin, E. G.; Kling, C. L.; Kudela, R. M.; Kurmayer, R.; Michalak, A. M.; Ortiz, J. D.; Otten, T. G.; Paerl, H. W.; Qin, B.; Sohngen, B. L.; Stumpf, R. P.; Visser, P. M.; Wilhelm, S. W., Global solutions to regional problems: Collecting global expertise to address the problem of harmful cyanobacterial blooms. A Lake Erie case study. *Harmful Algae* **2016,** *54*, 223-238.
9. Gobler, C. J.; Burkholder, J. M.; Davis, T. W.; Harke, M. J.; Johengen, T.; Stow, C. A.; Van de Waal, D. B., The dual role of nitrogen supply in controlling the growth and toxicity of cyanobacterial blooms. *Harmful Algae* **2016,** *54*, 87-97.
10. Paerl, H. W.; SCott, J. T.; McCarthy, M. J.; Newell, S. E.; Gardner, W. S.; Havens, K. E.; Hoffman, D. K.; Wilhelm, S. W.; Wurtsbaugh, W. A., It takes two to tango: When and where dual nutrient (N & P) reductions are needed to protect lakes and downstream ecosystems. *Environ Sci Technol* **2016,** *50*, (20), 10805-10813
11. Byth, S., Palm Island mystery disease. *The Medical Journal of Australia* **1980,** *2*, (1), 40, 42.
12. Michalak, A. M.; Anderson, E. J.; Beletsky, D.; Boland, S.; Bosch, N. S.; Bridgeman, T. B.; Chaffin, J. D.; Cho, K.; Confesor, R.; Daloğlu, I.; DePinto, J. V.; Evans, M. A.; Fahnenstiel, G. L.; He, L.; Ho, J. C.; Jenkins, L.; Johengen, T. H.; Kuo, K. C.; LaPorte, E.; Liu, X.; McWilliams, M. R.; Moore, M. R.; Posselt, D. J.; Richards, R. P.; Scavia, D.; Steiner, A. L.; Verhamme, E.; Wright, D. M.; Zagorski, M. A., Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc Natl Acad Sci USA* **2013,** *110*, (16), 6448-6452.
13. Bingham, M.; Sinha, S. K.; Lupi, F. *Economic benefits of reducing harmful algal blooms in Lake Erie*; Environmental Consulting and Technology, Inc.: 2015; p 66.
14. Clesceri, L. S.; Greenberg, A. G.; Eaton, A. D.; Eds., *Standard Methods for Examination of Water and Wastewater*. 20th ed.; American Public Health Association: Washington, D.C., 1998.
15. Davis, C. O.; Simmons, M. S. *Water chemistry and phytoplankton field and laboratory procedures. Special Report No. 70*; Michigan Univ., Ann Arbor (USA). Great Lakes Research Div.: 1979.

16. Menzel, D. W.; Corwin, N., The measurement of total phosphorus in seawater based on the liberation of organically bound fractions by persulfate oxidation. *Limnol. Oceanogr.* **1965,** *10*, (2), 280-282.
17. Speziale, B. J.; Schreiner, S. P.; Giammatteo, P. A.; Schindler, J. E., Comparison of N,N-dimethylformamide, dimethyl sulfoxide, and acetone for extraction of phytoplankton chlorophyll. *Can J Fish Aquat Sci* **1984,** *41*, (10), 1519-1522.
18. Fischer, W. J.; Garthwaite, I.; Miles, C. O.; Ross, K. M.; Aggen, J. B.; Chamberlin, A. R.; Towers, N. R.; Dietrich, D. R., Congener-independent immunoassay for microcystins and nodularins. *Environ Sci Technol* **2001,** *35*, (24), 4849-4856.
19. Boyer, G. L., The occurrence of cyanobacterial toxins in New York lakes: Lessons from the MERHAB-Lower Great Lakes program. *Lake Reservoir Manage* **2007,** *23*, (2), 153-160.
20. Steffen, M. M.; Belisle, B. S.; Watson, S. B.; Boyer, G. L.; Bourbonniere, R. A.; Wilhelm, S. W., Metatranscriptomic evidence for co-occurring top-down and bottom-up controls on toxic cyanobacterial communities. *Appl Environ Microbiol* **2015,** *81*, (9), 3268-3276.
21. Kaneko, T.; Nakajima, N.; Okamoto, S.; Suzuki, I.; Tanabe, Y.; Tamaoki, M.; Nakamura, Y.; Kasai, F.; Watanabe, A.; Kawashima, K.; Kishida, Y.; Ono, A.; Shimizu, Y.; Takahashi, C.; Minami, C.; Fujishiro, T.; Kohara, M.; Katoh, M.; Nakazaki, N.; Nakayama, S.; Yamada, M.; Tabata, S.; Watanabe, M. M., Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Res* **2007,** *14*, (6), 247-256.
22. Wacklin, P.; Hoffmann, L.; Komárek, J., Nomenclatural validation of the genetically revised cyanobacterial genus *Dolichospermum* (Ralfs ex Bornet et Flahault) comb. nova. *Fottea* **2009,** *9*, (1), 59-64.
23. Komárek, J., Modern taxonomic revision of planktic nostocacean cyanobacteria: A short review of genera. *Hydrobiologia* **2010,** *639*, (1), 231-243.
24. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A.; Wilkening, J.; Edwards, R., The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinf* **2008,** *9*, (1), 1-8.
25. Wilke, A.; Harrison, T.; Wilkening, J.; Field, D.; Glass, E. M.; Kyrpides, N.; Mavrommatis, K.; Meyer, F., The M5nr: A novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinf* **2012,** *13*, (1), 1-5.
26. Overbeek, R.; Begley, T.; Butler, R. M.; Choudhuri, J. V.; Chuang, H.-Y.; Cohoon, M.; de Crécy-Lagard, V.; Diaz, N.; Disz, T.; Edwards, R.; Fonstein, M.; Frank, E. D.; Gerdes, S.; Glass, E. M.; Goesmann, A.; Hanson, A.; Iwata-Reuyl, D.; Jensen, R.; Jamshidi, N.; Krause, L.; Kubal, M.; Larsen, N.; Linke, B.; McHardy, A. C.; Meyer, F.; Neuweger, H.; Olsen, G.; Olson, R.; Osterman, A.; Portnoy, V.; Pusch, G. D.; Rodionov, D. A.; Rückert, C.; Steiner, J.; Stevens, R.; Thiele, I.; Vassieva, O.; Ye, Y.; Zagnitko, O.; Vonstein, V., The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **2005,** *33*, (17), 5691-5702.
27. Mitra, S.; Rupek, P.; Richter, D. C.; Urich, T.; Gilbert, J. A.; Meyer, F.; Wilke, A.; Huson, D. H., Functional analysis of metagenomes and metatranscriptomes using SEED and KEGG. *BMC Bioinf* **2011,** *12*, (1), 1-8.
28. Baggerly, K. A.; Deng, L.; Morris, J. S.; Aldaz, C. M., Differential expression in SAGE: Accounting for normal between-library variation. *Bioinformatics* **2003,** *19*, (12), 1477-1483.
29. Rowe, M. D.; Anderson, E. J.; Wynne, T. T.; Stumpf, R. P.; Fanslow, D. L.; Kijanka, K.; Vanderploeg, H. A.; Strickler, J. R.; Davis, T. W., Vertical distribution of buoyant *Microcystis* blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. *J Geophys Res: Oceans* **2016,** *121,* (7), 5296-5314.
30. Huret, M.; Runge, J. A.; Chen, C.; Cowles, G.; Xu, Q.; Pringle, J. M., Dispersal modeling of fish early life stages: sensitivity with application to Atlantic cod in the western Gulf of Maine. *Mar Ecol: Prog Ser* **2007,** *347*, 261-274.

31. Gräwe, U., Implementation of high-order particle-tracking schemes in a water column model. *Ocean Model* **2011,** *36*, (1), 80-89.
32. Berry, M. A.; Davis, T. W.; Cory, R. M.; Duhaime, M. B.; Johengen, T. H.; Kling, G. W.; Marino, J. A.; Den Uyl, P. A.; Gossiaux, D.; Dick, G. J.; Denef, V. J., Cyanobacterial harmful algal blooms are a biological disturbance to western Lake Erie bacterial communities. *Environ Microbiol* **2017,** *19*, (3), 1149-1162.
33. Harke, M. J.; Davis, T. W.; Watson, S. B.; Gobler, C. J., Nutrient-controlled niche differentiation of Western Lake Erie cyanobacterial populations revealed via metatranscriptomic surveys. *Environ Sci Technol* **2016,** *50*, (2), 604-615.
34. Harke, M. J.; Berry, D. L.; Ammerman, J. W.; Gobler, C. J., Molecular response of the bloom-forming cyanobacterium, *Microcystis aeruginosa*, to phosphorus limitation. *Microbl Ecol* **2011,** *63*, (1), 188-198.
35. Glibert, P. M.; Harrison, J.; Heil, C.; Seitzinger, S., Escalating worldwide use of urea – A global change contributing to coastal eutrophication. *Biogeochemistry* **2006,** *77*, (3), 441-463.
36. Solomon, C. M.; Collier, J. L.; Berg, G. M.; Glibert, P. M., Role of urea in microbial metabolism in aquatic systems: A biochemical and molecular review. *Aquat Microb Ecol* **2010,** *59*, (1), 67-88.
37. Bogard, M. J.; Donald, D. B.; Finlay, K.; Leavitt, P. R., Distribution and regulation of urea in lakes of central North America. *Freshwater Biol* **2012,** *57*, (6), 1277-1292.
38. Belisle, B. S.; Steffen
, M. M.; Pound, H. L.; Watson, S. B.; DeBruyn, J. M.; Bourbonniere, R. A.; Boyer, G. L.; Wilhelm, S. W., Urea in Lake Erie: Organic nutrient sources as potentially important drivers of phytoplankton biomass. *J Great Lakes Res* **2016,** *42,* (3), 599-607.
39. Finlay, K.; Patoine, A.; Donald, D. B.; Bogard, M. J.; Leavitt, P. R., Experimental evidence that pollution with urea can degrade water quality in phosphorus-rich lakes of the Northern Great Plains. *Limnol Oceanogr* **2010,** *55*, (3), 1213-1230.
40. Steffen, M. M.; Dearth, S. P.; Dill, B. D.; Li, Z.; Larsen, K. M.; Campagna, S. R.; Wilhelm, S. W., Nutrients drive transcriptional changes that maintain metabolic homeostasis but alter genome architecture in *Microcystis*. *ISME J* **2014,** *8*, (10), 2080-2092.
41. Quintero, M. J.; Muro-Pastor, A. M.; Herrero, A.; Flores, E., Arginine catabolism in the cyanobacterium *Synechocystis* sp. strain PCC 6803 involves the urea cycle and arginase pathway. *J Bacteriol* **2000,** *182*, (4), 1008-1015.
42. Valladares, A.; Montesinos, M. L.; Herrero, A.; Flores, E., An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol Microbiol* **2002,** *43*, (3), 703-715.
43. Chaffin, J. D.; Bridgeman, T. B., Organic and inorganic nitrogen utilization by nitrogen-stressed cyanobacteria during bloom conditions. *J Appl Phycol* **2013,** *26*, (1), 299-309.
44. Davis, T. W.; Bullerjahn, G. S.; Tuttle, T.; McKay, R. M.; Watson, S. B., Effects of increasing nitrogen and phosphorus concentrations on phytoplankton community growth and toxicity during *Planktothrix* blooms in Sandusky Bay, Lake Erie. *Environ Sci Technol* **2015,** *49*, (12), 7197-7207.
45. Mou, X.; Jacob, J.; Lu, X.; Robbins, S.; Sun, S.; Ortiz, J. D., Diversity and distribution of free-living and particle-associated bacterioplankton in Sandusky Bay and adjacent waters of Lake Erie Western Basin. *J Great Lakes Res* **2013,** *39*, (2), 352-357.
46. Steffen, M. M.; Li, Z.; Effler, T. C.; Hauser, L. J.; Boyer, G. L.; Wilhelm, S. W., Comparative metagenomics of toxic freshwater cyanobacteria bloom communities on two continents. *PLoS ONE* **2012,** *7*, (8), e44002.
47. Burke, C.; Steinberg, P.; Rusch, D.; Kjelleberg, S.; Thomas, T., Bacterial community assembly based on functional genes rather than species. *Proc Natl Acad Sci USA* **2011,** *108*, (34), 14288-14293.
48. Davenport, E. J. Diel regulation of metabolic functions of a western Lake Erie *Microcystis* bloom informed by metatranscriptomic analysis. Bowling Green State University, 2016.

49. Beletsky, D.; Hawley, N.; Rao, Y. R., Modeling summer circulation and thermal structure of Lake Erie. *J Geophys Res: Oceans* **2013,** *118*, (11), 6238-6252.
50. Mlouka, A.; Comte, K.; Castets, A.-M.; Bouchier, C.; Tandeau de Marsac, N., The gas vesicle gene cluster from *Microcystis aeruginosa* and DNA rearrangements that lead to loss of cell buoyancy. *J Bacteriol* **2004,** *186*, (8), 2355-2365.
51. Xu, B.-Y.; Dai, Y.-N.; Zhou, K.; Liu, Y.-T.; Sun, Q.; Ren, Y.-M.; Chen, Y.; Zhou, C.-Z., Structure of the gas vesicle protein GvpF from the cyanobacterium *Microcystis aeruginosa*. *Acta Crystallogr, Sect D: Biol Crystallogr* **2014,** *70*, (11), 3013-3022.
52. Wynne, T.; Stumpf, R., Spatial and temporal patterns in the seasonal distribution of toxic cyanobacteria in western Lake Erie from 2002–2014. *Toxins* **2015,** *7*, (5), 1649.
53. Chaffin, J. D.; Bridgeman, T. B.; Heckathorn, S. A.; Mishra, S., Assessment of *Microcystis* growth rate potential and nutrient status across a trophic gradient in western Lake Erie. *J Great Lakes Res* **2011,** *37*, (1), 92-100.
54. Watson, S. B.; Miller, C.; Arhonditsis, G.; Boyer, G. L.; Carmichael, W.; Charlton, M. N.; Confesor, R.; Depew, D. C.; Höök, T. O.; Ludsin, S. A.; Matisoff, G.; McElmurry, S. P.; Murray, M. W.; Peter Richards, R.; Rao, Y. R.; Steffen, M. M.; Wilhelm, S. W., The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia. *Harmful Algae* **2016,** *56*, 44-66.
55. (NOAA), N. O. A. A. *Experimental Lake Erie Harmful Algal Bloom Bulletin*; 1 November, 2016, 2016.
56. Nakao, M.; Okamoto, S.; Kohara, M.; Fujishiro, T.; Fujisawa, T.; Sato, S.; Tabata, S.; Kaneko, T.; Nakamura, Y., CyanoBase: The cyanobacteria genome database update 2010. *Nucleic Acids Res* **2009**.

Tables

Table 1. Environmental conditions at the time of sample collection.

Station	*Latitude/ Longitude*	*Nitrate (mg/L)*	*Ammonia (mg/L)*	*PON (mg/L)*	*TN (mg/L)*	*TP (µg/L)*	*Chl a (µg/L)*	*PMCS-ELISA (µg/L)*	*PMCS-LCMS (µg/L)*
WE02	41 45.912 / 83 19.835	*0.56*	0.0030	0.32	0.88	31.98	29.44	4.33	0.91
WE04	41 49.714 / 83 11.654	*0.14*	0.0016	0.24	0.38	22.09	18.11	1.54	0.52
WE06	41 42.679 / 83 22.631	*0.29*	0.0018	0.86	1.15	62.19	71.62	10.14	3.45
WE08	41 50.254 / 83 21.823	*0.09*	0.0013	0.41	0.50	40.07	46.53	4.31	1.10
WE12	41 42.157 / 83 15.781	*0.62*	0.0033	0.68	1.30	44.37	54.46	9.28	2.91
WE13	41 44.539 / 83 08.286	*0.22*	0.0020	0.14	0.35	14.98	6.74	1.17	0.40
WI	41 42.913 / 83 19.722	*0.58*	0.0026	0.56	1.15	60.24	51.20	8.19	1.10
EI	41 40.561 / 83 14.329	*0.54*	0.0027	0.74	1.29	53.01	69.44	10.58	1.24

Figure Legends

Figure 1. Percentage of filtered sequence reads mapped to the genomes of *M. aeruginosa* NIES 843, *A. cylindrical* PCC 7122, and *P. agardhii* NVA CYA 128/6 (chromosome only). Percentages represent mean number of reads mapped between duplicate samples. See Table S2 for full details.

Figure 2. Transcriptional profiles for *M. aeruginosa* NIES 843. Values are normalized mean expression values (TCPM) for each station for duplicate samples. A) Log transformed [Log(X+1)] TCPM for all genes in the NIES 843 genome. B) TCPM for genes involved in nitrogen and phosphorus transport and metabolism. Annotated transcripts from the *Microcystis* genome without gene names include MAE_18310 and MAE_38290 (phosphate-binding periplasmic proteins), MAE_17690 and MAE_40010 (ammonium/methylammonium permease), MAE_16640 and MAE_50240 (alkaline phosphatase), MAE_ 09250 and MAE_09280 (phosphate transport system ATP-binding protein), MAE_12590 and MAE_40020 (Ammonium transporter or transport protein) MAE_09260, 09270, 18280, 18290, 18300 (phosphate transport system permease protein).

Figure 3. Schematic for expression of urea metabolism by *Microcystis*. The metabolic pathways were based on those proposed for *Synechocystis* WH6803[41] and the genome of *M. aeruginosa* NIES 843[21]. Colored dots denote station. Heat maps correspond to normalized mean expression value (TCPM) for genes involved in urea metabolism for

M. aeruginosa NIES 843. Urea is transported into the cell by the transport protein encoded by *urtABCDE* and hydrolyzed by the urease enzyme complex encoded by *ureABCDEFG*. The activity of arginase (ARG1 MAE_47100; ARG2, MAE_47180) breaks arginine down into urea and ornithine as part of the urea cycle, and this urea by-product can subsequently be hydrolyzed by urease. *argF* (MAE_54100) encodes ornithine carbamoyl transferase. The required input, carbamoyl phosphate, is synthesized by the carbamoyl phosphate synthase, encoded by *carA* (MAE_28430) and *carB* (MAE_50420). *argG* (MAE_02090) encodes argininosuccinate synthase, and *argH* (MAE_19870) encodes argininosuccinate lyase. Annotations are available from the *Microcystis* CyanoBase page[56].

Figure 4. Metatranscriptome-derived microbial community identity and function. A) Bacterial phylum abundance in assembled contigs. Values represent mean % abundance in Bacteria domain at the phylum level as annotated by M5nr database [25]. All phyla with <1% abundance were binned into the “Other” category. B) Functional profile of Proteobacteria at each station. C) Functional profile of Cyanobacteria at each station. Values are mean % abundance in annotated hits (65% Identity cutoff) to the SEED database using the workbench function in MG-RAST.

Figure 5. Transcriptional expression of *Microcystis* phage Ma-LMM01 tail sheath gene (*gp091*) normalized by *M. aeruginosa* expression of RNA polymerase beta subunit (*rpoB*) in Lake Erie metatranscriptomes isolated from 2012-2014.

Figure 6. Hydrodynamic model simulation showing transport of microcystin-contaminated water from Maumee Bay east along the southern shoreline toward the Toledo water intake in the days preceding the incident. Surface microcystin concentration observed on A) July 21, 2014 (spatially interpolated), and simulated microcystin concentration resulting from hydrodynamic transport of the July 21 concentrations on B) 24 July, 2014; C) 29 July, 2014; D) 4 August, 2014 (date of sample collection for genetic analysis). Observed concentrations are indicated as symbols on the same color scale as the simulated values; 29 July and 4 August microcystin observations were independent of the model and can be used for skill assessment.

TOC Figure.

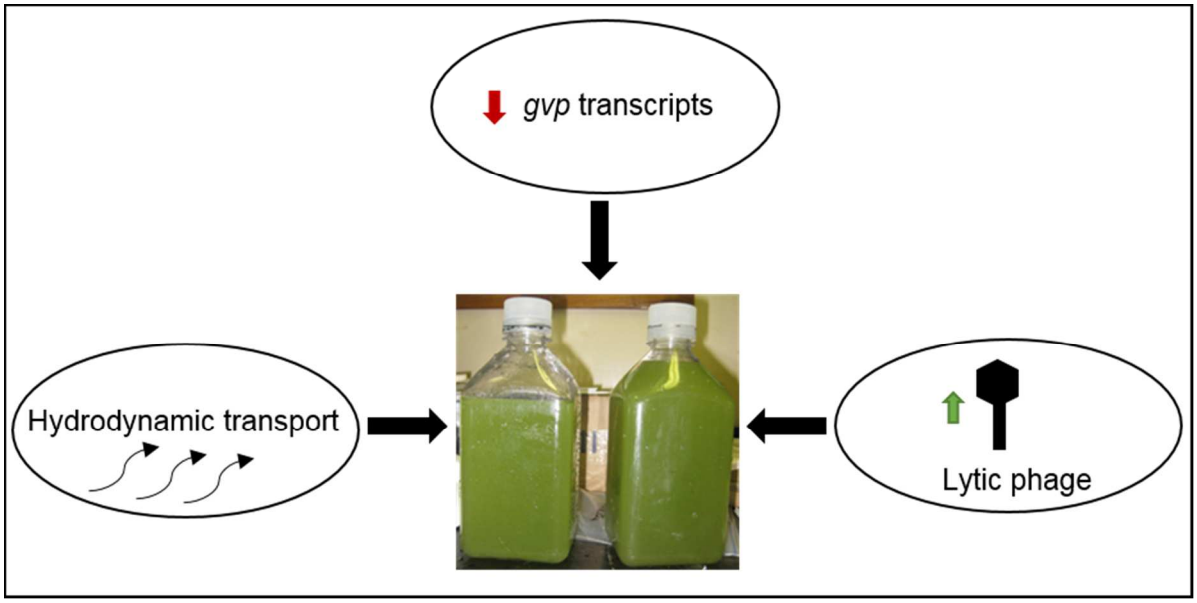


Figure 1.

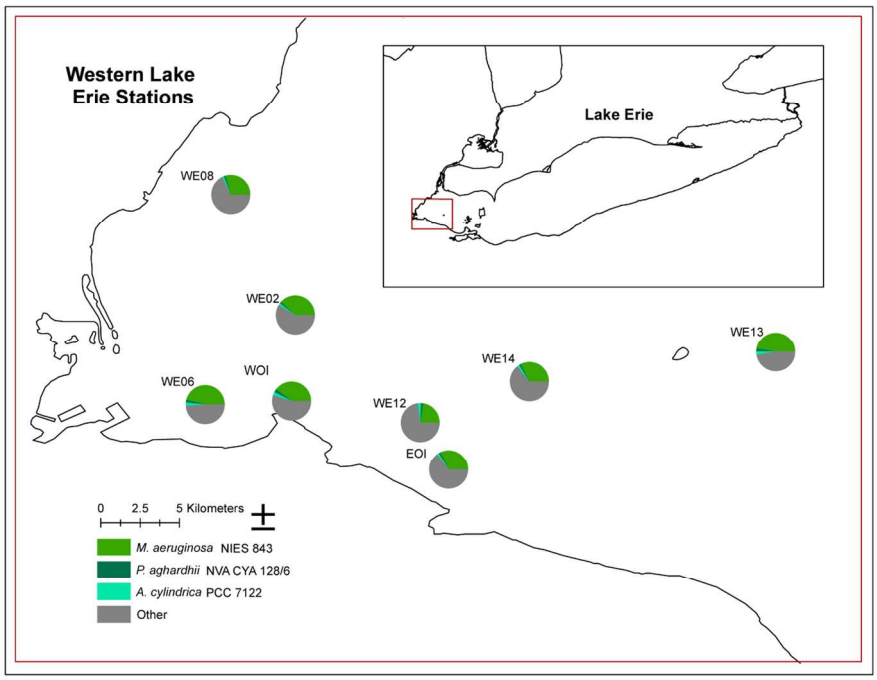

Western Lake Erie Stations
Lake Erie
WE08
WE02
WE06
WOI
WE12
EOI
WE14
WE13
0 2.5 5 Kilometers
M. aeruginosa NIES 843
P. aghardhii NVA CYA 128/6
A. cylindrica PCC 7122
Other

Figure 2.

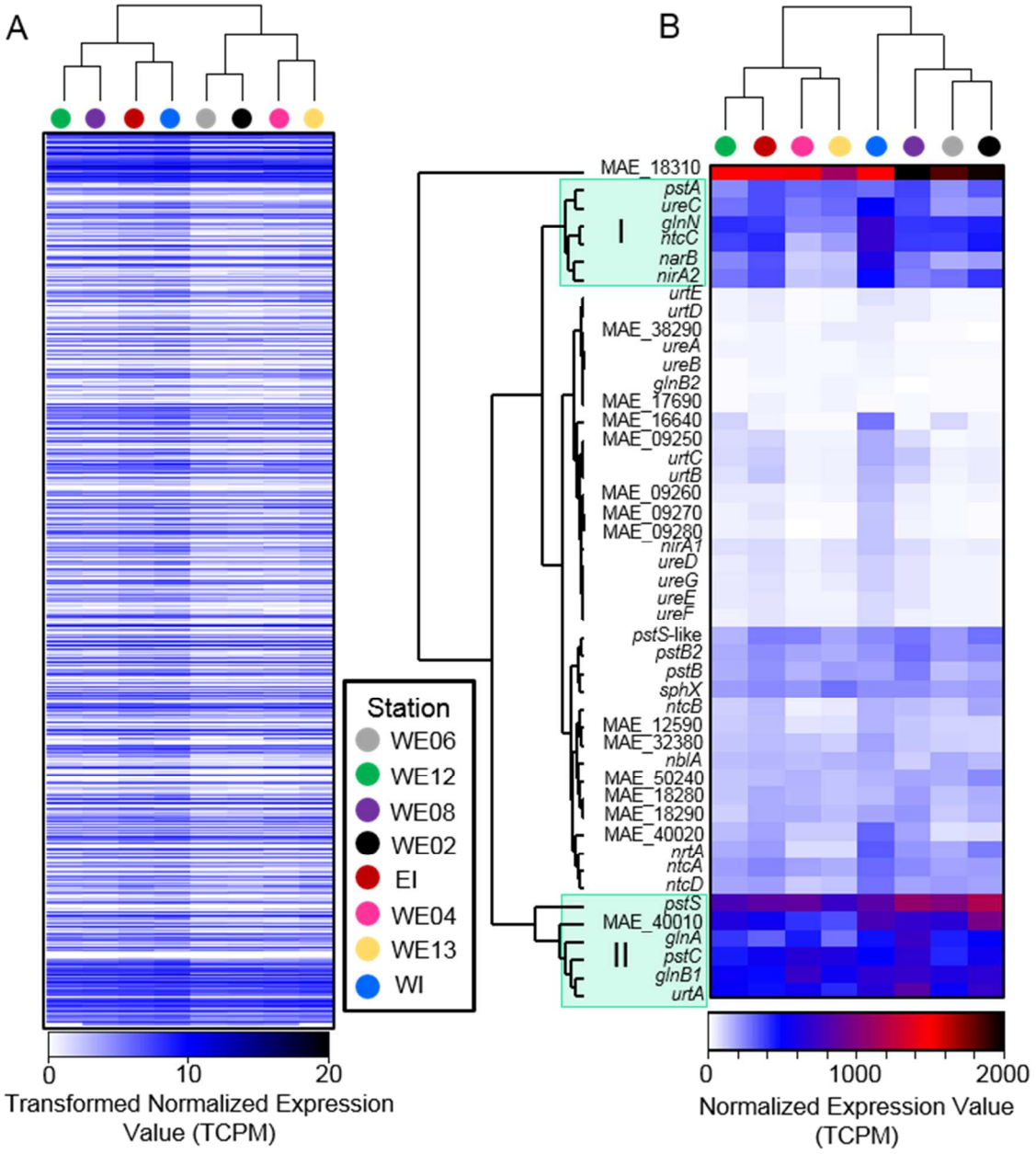

A
B
Station
WE06
WE12
WE08
WE02
EI
WE04
WE13
WI
0
10
20
Transformed Normalized Expression Value (TCPM)
0
1000
2000
Normalized Expression Value (TCPM)
I
II
MAE_18310
pstA
ureC
glnN
ntcC
narB
nirA2
urtE
urtD
MAE_38290
ureA
ureB
glnB2
MAE_17690
MAE_16640
MAE_09250
urtC
urtB
MAE_09260
MAE_09270
MAE_09280
nirA1
ureD
ureG
ureE
ureF
pstS-like
pstB2
pstB
sphX
ntcB
MAE_12590
MAE_32380
nblA
MAE_50240
MAE_18280
MAE_18290
MAE_40020
nrtA
ntcA
ntcD
pstS
MAE_40010
glnA
pstC
glnB1
urtA

Figure 3.

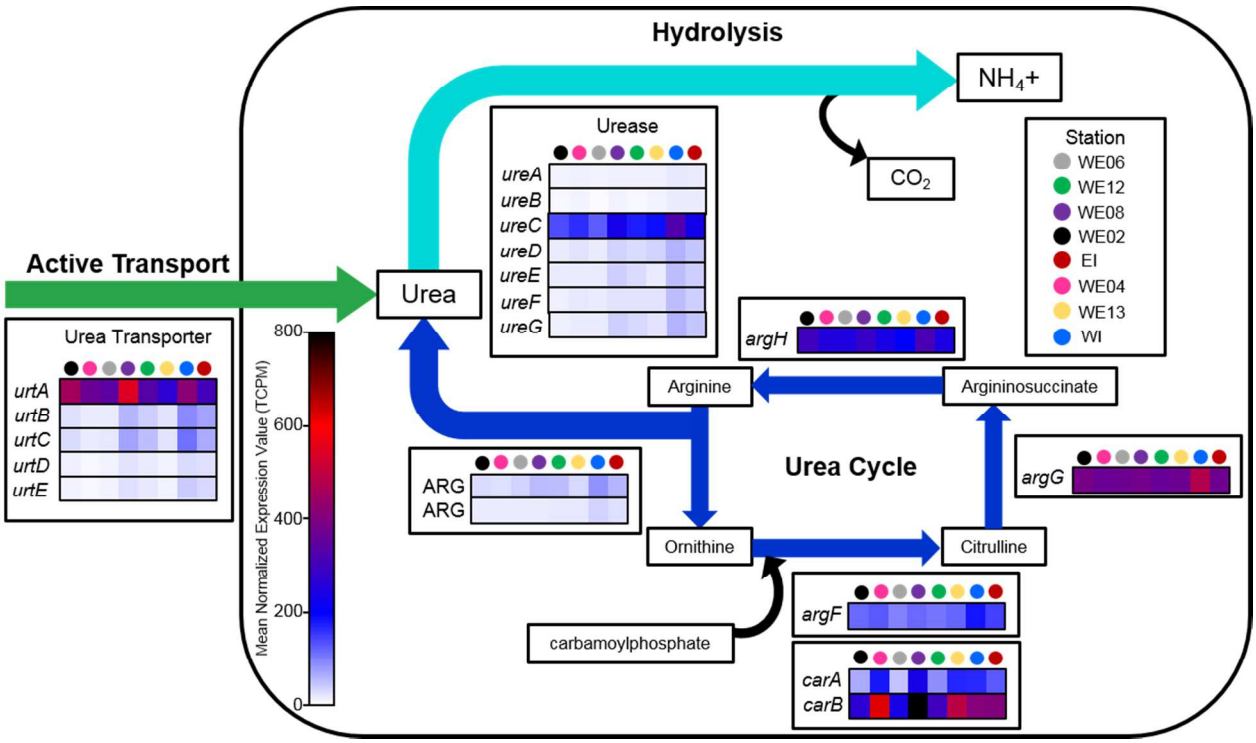

Hydrolysis
NH4+
CO2
Urease
ureA
ureB
ureC
ureD
ureE
ureF
ureG
Station
WE06
WE12
WE08
WE02
EI
WE04
WE13
WI
Active Transport
Urea
Urea Transporter
urtA
urtB
urtC
urtD
urtE
Mean Normalized Expression Value (TCPM)
800
600
400
200
0
argH
Arginine
Argininosuccinate
ARG
ARG
Urea Cycle
argG
Ornithine
Citrulline
argF
carbamoylphosphate
carA
carB

Figure 4

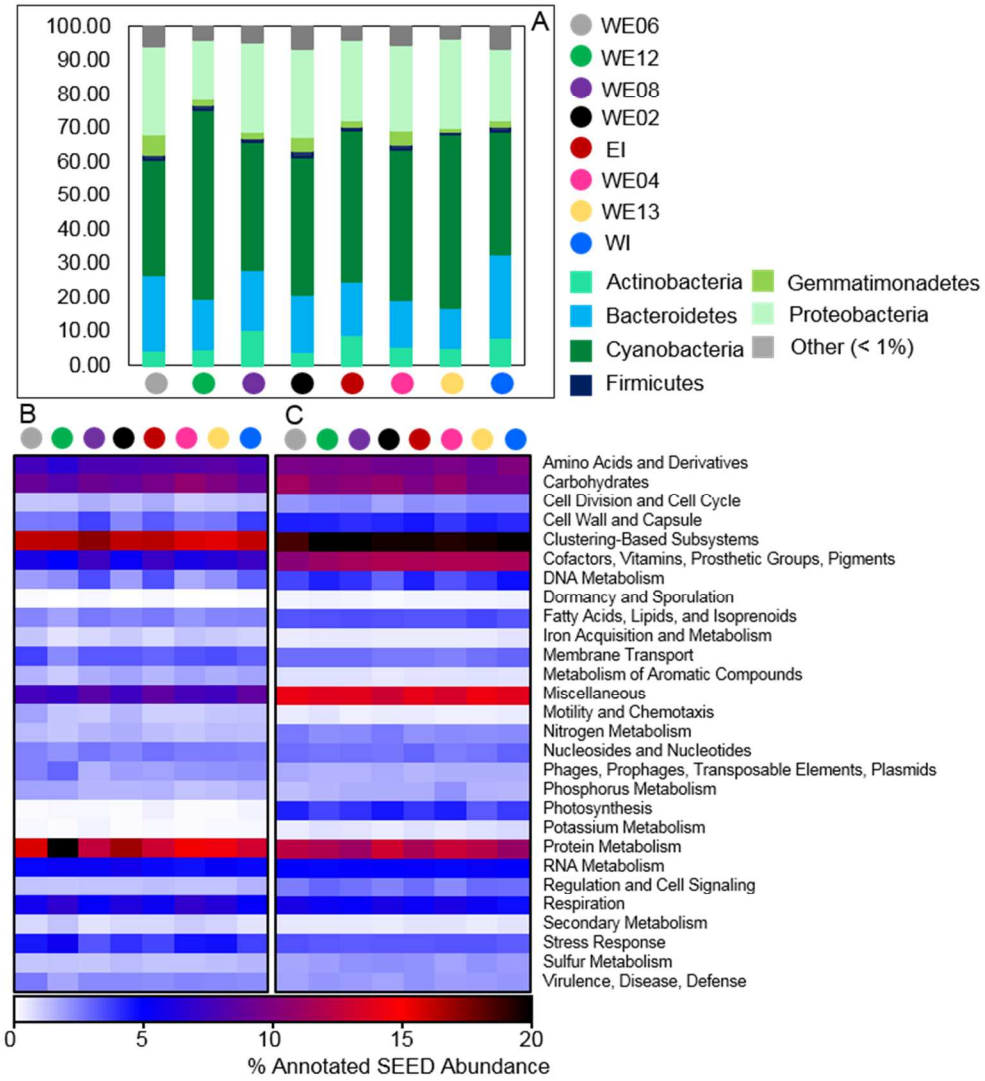

A
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
WE06
WE12
WE08
WE02
EI
WE04
WE13
WI
Actinobacteria
Bacteroidetes
Cyanobacteria
Firmicutes
Gemmatimonadetes
Proteobacteria
Other (< 1%)
B
C
Amino Acids and Derivatives
Carbohydrates
Cell Division and Cell Cycle
Cell Wall and Capsule
Clustering-Based Subsystems
Cofactors, Vitamins, Prosthetic Groups, Pigments
DNA Metabolism
Dormancy and Sporulation
Fatty Acids, Lipids, and Isoprenoids
Iron Acquisition and Metabolism
Membrane Transport
Metabolism of Aromatic Compounds
Miscellaneous
Motility and Chemotaxis
Nitrogen Metabolism
Nucleosides and Nucleotides
Phages, Prophages, Transposable Elements, Plasmids
Phosphorus Metabolism
Photosynthesis
Potassium Metabolism
Protein Metabolism
RNA Metabolism
Regulation and Cell Signaling
Respiration
Secondary Metabolism
Stress Response
Sulfur Metabolism
Virulence, Disease, Defense
0
5
10
15
20
% Annotated SEED Abundance

Figure 5

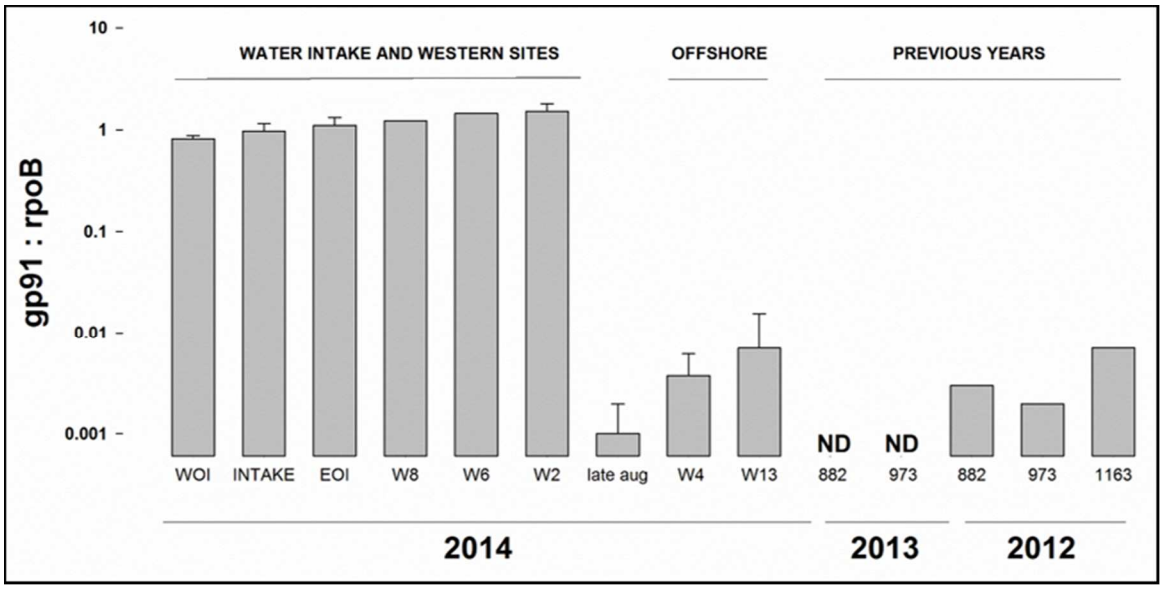
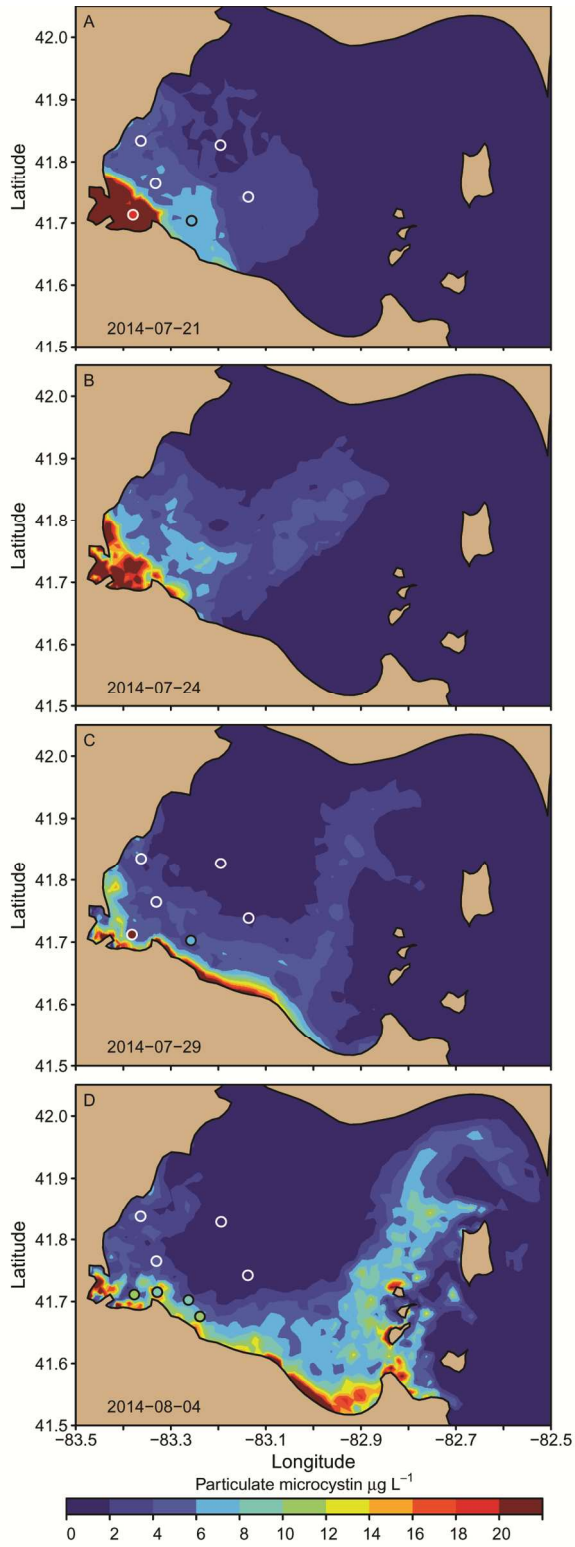

WATER INTAKE AND WESTERN SITES
OFFSHORE
PREVIOUS YEARS
gp91 : rpoB
10
1
0.1
0.01
0.001
WOI
INTAKE
EOI
W8
W6
W2
late aug
W4
W13
ND
882
ND
973
882
973
1163
2014
2013
2012

Figure 6.


A
B
C
D
2014−07−21
2014−07−24
2014−07−29
2014−08−04
Latitude
42.0
41.9
41.8
41.7
41.6
41.5
−83.5
−83.3
−83.1
−82.9
−82.7
−82.5
Longitude
Particulate microcystin µg L−1
0
2
4
6
8
10
12
14
16
18
20