PLOS | ONE
TENTH ANNIVERSARY

# Genome sequences of lower Great Lakes *Microcystis* sp. reveal strain-specific genes that are present and expressed in western Lake Erie blooms

Kevin Anthony Meyer, Timothy W. Davis, Susan B. Watson, Vincent J. Denef, Michelle A. Berry, Gregory J. Dick

## Abstract

Blooms of the potentially toxic cyanobacterium *Microcystis* are increasing worldwide. In the Laurentian Great Lakes they pose major socioeconomic, ecological, and human health threats, particularly in western Lake Erie. However, the interpretation of "omics" data is constrained by the highly variable genome of *Microcystis* and the small number of reference genome sequences from strains isolated from the Great Lakes. To address this, we sequenced two *Microcystis* isolates from Lake Erie (*Microcystis aeruginosa* LE3 and *M. wesenbergii* LE013-01) and one from upstream Lake St. Clair (*M. cf aeruginosa* LSC13-02), and compared these data to the genomes of seventeen *Microcystis* spp. from across the globe as well as one metagenome and seven metatranscriptomes from a 2014 Lake Erie *Microcystis* bloom. For the publically available strains analyzed, the core genome is ~1900 genes, representing ~11% of total genes in the pan-genome and ~45% of each strain's genome. The flexible genome content was related to *Microcystis* subclades defined by phylogenetic analysis of both housekeeping genes and total core genes. To our knowledge this is the first evidence that the flexible genome is linked to the core genome of the *Microcystis* species complex. The majority of strain-specific genes were present and expressed in bloom communities in Lake Erie. Roughly 8% of these genes from the lower Great Lakes are involved in genome plasticity (rapid gain, loss, or rearrangement of genes) and resistance to foreign genetic elements (such as CRISPR-Cas systems). Intriguingly, strain-specific genes from *Microcystis* cultured from around the world were also present and expressed in the Lake Erie blooms, suggesting that the *Microcystis* pangenome is truly global. The presence and expression of flexible genes, including strain-specific genes, suggests that strain-level genomic diversity may be important in maintaining *Microcystis* abundance during bloom events.

## Introduction

Genetic and physiological differentiation can result in the emergence of closely related populations that are ecologically distinct [1, 2]. This differentiation is encoded in part by "flexible" genes that are frequently gained and lost [3, 4]. In instances where this divergence yields differences in niche occupation the populations represent different ecotypes [5]. Ecotypes have been identified and tracked in populations over seasonal cycles [4] and multiple years [6], and across environmental [7–9] and geographic gradients [10–12]. Some of the best-documented studies of bacterial ecotypes were done on the abundant and widespread cyanobacteria *Prochlorococcus* and *Synechococcus*, for which ecotypes have been defined that have distinct adaptations to light, nutrient availability, and phage resistance [7, 8, 13].

Whereas cyanobacteria such as *Prochlorococcus* and *Synechococcus* are globally important primary producers [14, 15], some cyanobacteria have an innate ability to form large, harmful, and sometimes toxic blooms that impact ecosystem dynamics and threaten freshwater ecosystems, regional economies, recreational activities, and drinking water supplies [16–19]. Among these harmful cyanobacteria, *Microcystis* is one of the most widely reported, described from large blooms in lakes, ponds, reservoirs, and rivers of every continent except Antarctica [18, 20, 21]. However, factors contributing to the worldwide success of *Microcystis* are largely uncharacterized and there remains considerable variance in its response to management efforts [22]. Based on average nucleotide identity, 16S rRNA sequence homology, and DNA-DNA hybridization, all known strains of *Microcystis* belong to the same species complex [21]. However, unlike *Prochlorococcus* and *Synechococcus*, no patterns have been observed in the biogeographic distribution of *Microcystis* genotypes. Genetic variation between strains from the same continent can be higher than between continents [23] and some strains have a large geographic distribution [23, 24].

The dynamic genotypic and phenotypic nature of M*icrocystis* is evident in the high abundance of insertion sequences, transposable elements, restriction enzymes, and genes likely acquired through lateral transfer [25, 26]. This includes the phylogenetic distribution of *mcyA-J* genes encoding biosynthesis of microcystins, the most commonly detected cyanotoxins globally [21, 27, 28]. Phylogenetic analysis of both individual genes and multiple loci within that gene cluster have found toxicity to be polyphyletic with clusters of toxic strains, non-toxic strains, and a combination of both [29–31]. Thus, toxicity appears to be a genetic element that is stable only in the short-term [30]. Indeed, *Microcystis* has lost the ability to produce microcystins in multiple instances over evolutionary time, indicating that toxicity was not essential, or potentially too costly, for survival in particular ecosystems or under certain environmental conditions [32].

Despite the high genetic variability between strains and apparent lack of biogeography, subclades of strains based on phylogenetic relationships of core genes are well preserved, suggesting some degree of phylogenetic cohesion [30, 31]. Further, individual genes display biogeographical patterns [33], and hydrologically linked systems show genetic connectivity in toxic strains of *Microcystis* through the *mcyA* gene [34]. These conflicting patterns likely result from a balance between genome plasticity (rapid gain, loss, or rearrangement of genes), which generates potentially useful variation [35], and genome stability (restriction of gene transfer and maintenance of core genes), which preserves useful variants and resists harmful elements [36].

One hypothesis developed by previous studies is that the conserved subclades of *Microcystis* represent ecotypes (physiologically different strains) [26, 37] or cryptic ecotypes (a phylogenetic cluster that is new or low frequency but ecologically distinct) [30, 38]. In this case, the high genetic diversity observed in co-occurring *Microcystis* strains may represent distinct ecotypes that provide the genetic variation needed for ecological divergence through selection processes [30]. An alternative hypothesis is that stable niches are a prerequisite for ecotype differentiation, and that potential niches available to *Microcystis* are insufficiently stable. As a result, genome plasticity represents an efficient strategy for dynamic environments [39] by rapidly generating variants that may be adapted to new environmental conditions [31, 40]. The genome plasticity of *Microcystis* may therefore have been selected for through both gene content and transcription [33, 40], creating phenotypic variation within a localized population or bloom adapted for specific environmental conditions [38].

The Laurentian Great Lakes have a history of large-scale cyanobacterial blooms dominated by *Microcystis*, with the largest blooms typically occurring in western Lake Erie [22, 41–44]. Currently, only one genome of cultured *Microcystis* from the Laurentian Great Lakes (Lake Michigan) has been described [31]. This represents a critical gap for genomic and metagenomic studies that rely on mapping sequence reads to available genomes from reference strains [21, 45–47]. Our goal was to close this knowledge gap by sequencing the genomes of three lower Great Lakes (LGL) *Microcystis* strains, using comparative genomics to place them in the framework of other *Microcystis* genomes. Further, analysis of the content and expression of their strain-specific genes at several stages of the 2014 Lake Erie bloom reveals how genetic variation manifests in natural communities across changing environmental conditions, providing insights into whether differences in gene content and expression of LGL *Microcystis* strains could represent ecotypes and how these differences contribute to bloom formation, proliferation, and toxicity.

## Methods

### *Microcystis* strain culture and DNA extraction

Whole genome analysis of 20 *Microcystis* strains was conducted (Table 1). Of these genomes, three are newly sequenced isolates from the LGL (*Microcystis cf aeruginosa* LSC13-02, *Microcystis aeruginosa* LE3, and *Microcystis wesenbergii* LE013-01) while the other 17 have been previously compared and were obtained from the Joint Genome Institute Integrated Microbial Genomes & Microbiomes (JGI IMG/M) databases [21, 31, 36, 48–54]. Isolates LE3 and LE013-01 were isolated from western Lake Erie [43], and LSC13-02 was isolated from Lake St. Clair, which is both physically and genetically connected to Lake Erie via the Detroit River [34].

| Strain | Culture Reference | Isolation date | Genome Size (Mbp) | mcy genes? | Strain-specific genes | % core genes | Strain-specific % present | Strain-specific % expressed |
|---|---|---|---|---|---|---|---|---|
| Microcystis aeruginosa | PCC 9807 | 1973 | 5.1 | Yes | 594 | 36.9 | 98.0 | 95.3 |
| Microcystis aeruginosa | PCC 9443 | 1954 | 5.1 | Yes | 683 | 36.7 | 97.2 | 81.1 |
| Microcystis aeruginosa | PCC 9432 | Sep 1954 | 5.0 | No | 316 | 38.2 | 98.7 | 78.5 |
| Microcystis aeruginosa | PCC 7005 | 1946 | 4.9 | No | 561 | 37.0 | 97.2 | 82.0 |
| Microcystis aeruginosa | PCC 9806 | 1973 | 5.0 | Yes | 448 | 37.4 | 98.0 | 78.6 |
| Microcystis aeruginosa | PCC 7941 | Sep 1954 | 4.9 | Yes | 226 | 40.7 | 96.9 | 81.9 |
| Microcystis aeruginosa | LE3 | 1996 | 4.6 | Yes | 189 | 43.0 | 99.5 | 88.9 |
| Microcystis aeruginosa | TAIHU98 | 1997 | 4.9 | No | 331 | 39.6 | 99.1 | 89.7 |
| Microcystis aeruginosa | SPC 777 | Feb 2000 | 5.5 | Yes | 471 | 38.6 | 92.6 | 96.8 |
| Microcystis aeruginosa | PCC 7806 | 1972 | 5.3 | Yes | 392 | 37.1 | 99.5 | 90.3 |
| Microcystis aeruginosa | DIANCHI905 | 1998 | 4.9 | Yes | 283 | 39.7 | 97.2 | 91.6 |
| Microcystis aeruginosa | PCC 9806 | Aug 1975 | 4.2 | No | 398 | 45.5 | 98.2 | 84.7 |
| Microcystis sp. | T1-4 | Unknown | 4.7 | No | 525 | 40.9 | 99.2 | 90.4 |
| Microcystis aeruginosa | PCC 9701 | 1996 | 4.7 | No | 442 | 40.3 | 98.1 | 88.2 |
| Microcystis wesenbergii | LE13-01 | Aug 2013 | 4.8 | No | 222 | 49.6 | 100 | 83.6 |
| Microcystis aeruginosa | NIES-44 | Aug 1974 | 4.6 | No | 292 | 43.3 | 94.9 | 87.3 |
| Microcystis of aeruginosa | LSC13-02 | Aug 2013 | 4.3 | Yes | 458 | 44.9 | 99.8 | 89.1 |
| Microcystis aeruginosa | PCC 9717 | 1996 | 5.2 | Yes | 642 | 36.0 | 97.5 | 96.1 |
| Microcystis aeruginosa | PCC 9809 | Aug 1982 | 4.9 | Yes | 526 | 37.4 | 98.1 | 85.6 |
| Microcystis aeruginosa | NIES-843 | Aug 1997 | 5.8 | Yes | 526 | 33.7 | 99.4 | 89.9 |

Microcystis aeruginosa strain LE3 and Microcystis wesenbergii strain LE13-01 were isolated from Lake Erie in the summer of 1996 and August, 2013, respectively. Microcystis of aeruginosa strain LSC13-02 was isolated from Lake St. Clair in August, 2013. Strain-specific % present and expressed columns represent the proportion of strain-specific genes found in the metagenome and metatranscriptomes from a 2014 Lake Erie Microcystis bloom. Horizontal bars indicate subclades as assigned by Humbert et al. (2013) [31] and as presented in Fig 1.

https://doi.org/10.1371/journal.pone.0183859.t001

**Table 1. Comparative genomic features of 17 publicly available *Microcystis* genomes (Downloaded from JGI IMG/M) [52] and three strains isolated from the lower Great Lakes (in bold).**
https://doi.org/10.1371/journal.pone.0183859.t001

Individual isolates (15mL) were grown in BG-11 medium [55] in an incubation room (20°C, 38 µE m$^{-2}$ s$^{-2}$, 12h Light:Dark cycle). Cells were spun down at ~9,400 × g for 10 minutes in a benchtop centrifuge, the supernatant was decanted and the pelleted cellular material was frozen at -80°C until subsequent DNA extraction, purification, and sequencing. Frozen cell pellets were thawed at room temperature and then extracted using a Qiagen DNeasy® Blood and Tissue Kit (Qiagen, Hilden, Germany). Briefly, samples were incubated with 100µL Qiagen ATL tissue lysis buffer, 300µL Qiagen AL lysis buffer, and 30µL proteinase K at 56°C for 1 hour with agitation, followed by mixing with a vortexer at maximum speed for 10 minutes. Lysates were homogenized with a QiaShredder™ spin-column before purification according to the DNeasy® protocol.

### Lake Erie sampling and RNA extractions

Field samples were collected in conjunction with the joint NOAA Great Lakes Environmental Research Laboratory / University of Michigan Cooperative Institute for Great Lakes Research weekly sampling program for western Lake Erie. In 2014, six sites were sampled bi-monthly in June then weekly from July through October. From this sampling effort we focused on representative samples to capture each stage of the *Microcystis* bloom; early bloom (late July-early August), mid-bloom (late August), and late/post-peak bloom (September-October). Bloom stages were determined by phycocyanin fluorescence and relative abundance [56]. Samples were chosen from three of the regularly sampled stations (WE2, near to the mouth of Maumee River, 41° 45.743' N, 83°19.874' W; WE4, offshore towards the center of the western basin, 41°49.595' N, 83°11.698' W; and WE12, adjacent to the water intake crib for the city of Toledo, 41°42.535' N, 83°14.989' W) [56]. Metagenomic data were generated from samples collected from station WE12 in early blooms stages, and metatranscriptomic data were generated from early (1 sample each from all three stations, WE2 21.July.2014, WE4 29.July.2014, and WE12 4.August.2014), middle (1 sample from station WE12 only, 25.August.2014), and late (1 sample each from all three stations, WE2 6.October.2014, WE4 8.September.2014, WE12 23.September.2014) bloom stage samples. All samples were collected shortly after arriving on-station using a peristaltic pump to obtain 2L of water integrated from 0.1m below the surface to 1m above the bottom [56]. This was then filtered onto 100µm polycarbonate 47mm filter in a Swinnex™ filter holder using a sterile 60mL syringe. A pore size of 100 µm was used to maximize the amount of *Microcystis* colonies retained on the filter while excluding smaller particles. Previous work has shown that in Lake Erie the > 100 µm *Microcystis* community comprised over 90% of all *Microcystis* cells in the water column [57, 58]. Filters were then immersed in 1 mL RNALater™ (Invitrogen™ Ambion™) and placed on ice during transport before being stored at -80°C. RNA was extracted from samples using RNeasy Mini Kit (QIAGEN) according to manufacturer's instructions.

### Assembly and binning

Shotgun sequencing of DNA and RNA was performed on the Illumina® HiSeq™ platform (2000 PE 100, Illumina, Inc., San Diego, CA, USA) at the University of Michigan DNA Sequencing Core. Sequence reads were put through a quality control pipeline that consisted of two runs of FASTQC version 0.10.0 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), dereplication (100% identity over 100% of length), adapter removal using Scythe [59], and read trimming using Sickle [60].

The genome of each *Microcystis* isolate was assembled *de novo*. We used the iterative de Bruijn graph approach for uneven sequencing depths (IDBA-UD) [61] for all assemblies with the following parameters: minimum kmer size 52, maximum kmer size 92, step size 8, minimum contig 500. Read coverage was generated by mapping paired end reads to the assembled contigs using the Burrows-Wheeler Aligner (BWA version 0.7.9a-r786) [62] with default parameters. Paired forward and reverse alignments were generated in SAM format and read counts extracted using SAMtools 1.0 [63]. Because the *Microcystis* cultures were not axenic, assembled contigs were binned into putative taxonomic groups with emergent self-organizing maps (ESOM) of tetranucleotide frequencies (Robust ZT transformation) using Databionics ESOM Tools [64] (http://databionic-esom.sourceforge.net). Only contigs longer than 4kb were considered for ESOM binning, and longer contigs were chopped into 10kb windows for this analysis. The other ESOM parameters were as follows: training with a K-Batch algorithm (k = 0.15%) for 40 training epochs, a standard best match search method with a local best match search radius of 8, a Gaussian weight initialization, Euclidean data space function, a starting training radius of 204 with linear cooling to 1, and a starting learning rate of 0.5 with linear cooling to 0.1. Bin taxonomy was identified using a combination of: (i) BLASTN of contigs [65] against the Silva SSU Database version 119 [66, 67]; (ii) ESOM binning with *Microcystis aeruginosa* DIANCHI905 and PCC 9808 as reference genomes; (iii) phylogenetic analysis using the full marker set in the PhyloSift pipeline [68].

The three newly sequenced *Microcystis* genomes were individually submitted to the Integrated Microbial Genomes database (US DOE JGI/IMG) for gene calling [52–54]. All newly sequenced genome files are available from the Integrated Microbial Genomes Database (US DOE JGI/IMG: Taxon IDs 2606217223, 2606217222, and 2606217221) and NCBI (Genome accession numbers MTBS00000000, MTBU00000000, and MTBT00000000; Sequence Read Archive accession numbers SRR5144737, SRR5145065, and SRR5145066; BioProjects PRJNA340013, PRJNA340089, and PRJNA340134; BioSamples SAMN05629279, SAMN05645801, and SAMN05645814). Gene annotations are available from JGI/IMG (analysis project IDs GA0066243, GA0066240, and GA0066226).

### Phylogenetic analysis

Nucleic acid sequences for six housekeeping genes (Cell division protein FtsZ, glutamine synthetase, Glutamyl-tRNA Synthetase, glucose-6-phosphate isomerase, DNA repair protein recA, and triose phosphate isomerase) [31], as well as the ribosomal protein S3, global nitrogen regulator (*ntcA*), and phycocyanin subunit B gene (*cpcB*) were obtained from the NCBI Nucleotide database or IMG for all 20 *Microcystis* genomes. Sequences were aligned using the MUSCLE tool [69] in MEGA6 (Version 6.06; Build 6140226) [70] and concatenated using ARB v5.5 [71]. Phylogenetic inferences were made on concatenated alignments with Randomized axelerated maximum likelihood (RAxML) using RAxML-VI-HPC v8.1.15 with a total of 1000 iterations [72]. A second, maximum parsimony tree was created through GET_HOMOLOGUES with the PARS program of the PHYLIP suite [73] using presence/absence data of the flexible genome compiled into a pan-genome matrix from OMCL clusters of only the flexible genes [74]. Phylogenetic trees were formatted using FigTree v1.4.2 [http://tree.bio.ed.ac.uk] to identify previously established *Microcystis* sp. subclades [31] and joined using Adobe® Illustrator® CS6.

### Core and pan-genome analysis

In order to standardize the gene calling and annotation for the purpose of comparative genomics, we used the Prokaryotic dynamic programming gene-finding algorithm (Prodigal) to predict genes in all 20 *Microcystis* genomes [75]. Prodigal was first run in training mode to establish *Microcystis aeruginosa* NIES-843 (the only complete *Microcystis* genome used in this study) [25] as a reference genome followed by a run in normal mode with all 20 genomes. Annotations were then re-assigned by a BLASTN of annotated genome sequences obtained from IMG or NCBI to the gene sequences called by Prodigal, keeping only the BLAST hits with a minimum bitscore of 100 and percent identify match of 95%.

Comparative genomics analysis was done using the GET_HOMOLOGUES software package [74]. Orthologous gene families were identified using the OrthoMCL clustering algorithm (OMCL) with a sequence cluster reporting value of t = 0 and no Pfam-domain composition requirements [74, 76, 77]. This approach uses the exponential decay models of Tettelin [78] and Willenbrock [79] to calculate the core-genome size and the exponential model of Tettelin to estimate the pan-genome size (the sum of both genes shared amongst all strains and genes strain-specific to individual strains) [78]. Strain-specific genes, genes found in only a single *Microcystis* isolate, were identified using the *parse_pangenome_matrix.pl* script with the pan-genome set of each individual strain being compared to those of all other genomes [76].

### Metagenomic and metatranscriptomic analysis

Raw reads from the 2014 bloom metagenomic and metatranscriptomic data were mapped to the strain-specific genes identified in GET_HOMOLOGUES, called genes from IMG, and the housekeeping, ribosomal protein S3, *ntcA*, and *cpcB* genes used in the phylogenetic analysis using the BWA mapper with default parameters [62]. Strain-specific gene reads were then normalized by the average read coverage of the genes used in the phylogenetic analysis. Annotation data for strain-specific genes were obtained from IMG and the proportion of genes in each COG was calculated using the normalized read coverage.

## Results

### Features and phylogeny of three new draft genomes from lower Great Lakes *Microcystis* strains

Basic features of the three new draft genomes of *Microcystis* strains from the lower Great Lakes (LGL) are shown in Table 1. Approximately 70% of the coding DNA sequences in the LGL genomes had functional annotations. Comparative genomics showed that every *Microcystis* strain had strain-specific genes (not found in any other *Microcystis* strain used in this study). The number of strain-specific genes for the LGL strains of *Microcystis* was 189, 222, and 458, in LE3, LE013-01, and LSC13-02, respectively. The most abundant functional categories for completely strain-specific genes were tranposases (2.9%), transferases (5.9%), and endonucleases (2.6%).

Two methods were used to evaluate the genetic relationships of sequenced *Microcystis* strains: maximum likelihood analysis of nine core genes and maximum parsimony based on the content of flexible genes within each genome (Fig 1). There was good overall congruence between the trees produced by these two methods. Only the placement of two pairs of strains differed, *Microcystis aeruginosa* SPC 777 and TAIHU98, and PCC 9701 and PCC 9806, and both of these differences were within subclades defined previously [31].
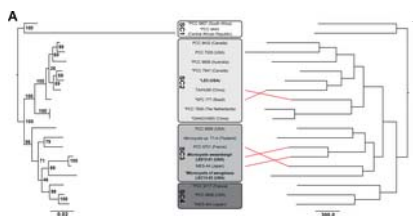
**Fig 1.**
Cladograms of twenty *Microcystis* strains based on A- phylogenetic analysis using RAxML inference on concatenated alignments of six housekeeping genes (*ftsZ*, *glnA*, *gltX*, *pgi*, *recA*, and *tpi*), ribosomal protein S3, global nitrogen regulator (*ntcA*), and phycocyanin subunit B (*cpcB*) genes and B- parsimony search of presence/absence of flexible genes within each genome. SC = Subclade as assigned by Humbert et al. (2013) [31], red lines indicate where phylogenetic assignments differ between trees A and B, *strains which are toxic based on the presence of the *mcy* gene operon.
https://doi.org/10.1371/journal.pone.0183859.g001

Based on 20 genomes of *Microcystis* the core-genome was estimated to be 2008 and 1924 genes with residual standard errors of 174 and 134 for the Tettelin and Willenbrock models, respectively. The pan-genome for these 20 *Microcystis* strains was estimated to be 8620 genes with a residual standard error of 196. Within any given *Microcystis* strain, the genome consisted of 34–49% core genes and 51–66% flexible genes.

### Occurrence of foreign genetic defense systems in lower Great Lakes isolates of *Microcystis*

Multiple strain-specific genes associated with genomic plasticity were found in LGL *Microcystis* isolates, several of which defend against foreign genetic elements (Table 2). Strain LE3 had four strain-specific genes identified as *csc* and *cas* components of the CRISPR-Cas system (Fig 2), which defends against invading genetic elements [80]. The CRISPR-Cas system found in strain LE3 has an architecture and array repeat similar to that of the *Microcystis* Subtype I-D CRISPR-Cas 2 and DR1 [36]. The LE3 CRISPR-Cas 2 differed from those in 11 of the 19 other *Microcystis* genomes (NIES-843, PCC 9443, PCC 7806, PCC 9808, PCC 7941, PCC 9717, PCC 9809, PCC 9807, DIANCHI905, PCC 9432, and PCC 9701) in that *cas*3 was not adjacent to the transcriptional regulator but was instead between *csc*1 and *uma*2 family endonucleases adjacent to the *cas*6 gene. There were also second copies of *csc*1, *csc*2, *csc*3, and *cas*3 near the beginning of the scaffold. Subtype III-B CRISPR-Cas 5 and Subtype III-B CRISPR-Cas 6 (with DR3 and DR4 repeat arrays) were also found, but with none of the associated genes being classified as strain-specific. Both CRISPR-Cas 5 and CRISPR-Cas 6 were most similar to that found in *Microcystis* TAIHU98. In addition there was a polyphosphate kinase (*ppk*) gene located upstream of two CRISPR arrays (DR2 repeat type) with 3 hypothetical proteins and an mRNA interferase (RelE/StbE) between them, which could be the remnants of a Subtype III-B CRISPR-Cas 4 system. Strain LE3 also had 81 restriction enzymes and 6 DNA restriction-modification genes, of which 3 restriction endonucleases were strain-specific.
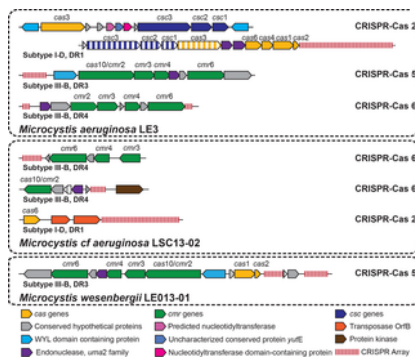


**Fig 2. Organization and repeat structure of CRISPR-Cas arrays identified in lower Great Lakes *Microcystis* strains.**
Genes colored with vertical stripes indicate genes strain-specific to that particular strain of *Microcystis*. CRISPR-Cas categories, subtypes, and repeat types were based on architecture described by Yang et al. (2015) [36].
https://doi.org/10.1371/journal.pone.0183859.g002



| Strain | CRISPR-Cas systems | Restriction-modification | Strain-specific restriction-modification | Transposases | Strain-specific transposases |
|---|---|---|---|---|---|
| LE3 | 3 | 87 | 3 | 48 | 1 |
| LE013-01 | 1 | 86 | 5 | 47 | 5 |
| LSC13-02 | 2 | 95 | 20 | 49 | 13 |

https://doi.org/10.1371/journal.pone.0183859.t002

**Table 2. Genes associated with genomic plasticity and resistance to foreign genetic elements in lower Great Lakes *Microcystis* strains.**
https://doi.org/10.1371/journal.pone.0183859.t002

Strain LSC13-02 contained two versions of Subtype III-B CRISPR-Cas 6, both with the DR4 repeat [36]. One had an organization most similar to that of PCC 9807 and the second had an organization similar to that of *Microcystis* PCC 7941, PCC 7806, and DIANCHI905. The second CRISPR-Cas 6 array differed from previously described arrays with a protein kinase oriented away from the CRISPR array (as opposed to a polyphosphate kinase running towards the CRISPR array) and a strain-specific hypothetical protein between the CRISPR array and *cas10/cmr*2. There is also a large CRISPR array that is downstream of *Cas6* with two transposases between the *Cas6* protein and the CRISPR array that could be a partial Subtype I-D CRISPR-Cas 2 given that it has a DR1 repeat [36]. Strain LSC13-02 also had 91 restriction enzymes and 4 restriction-modification proteins. 20 of these were strain-specific to strain LSC13-02.

Strain LE013-01 possessed Subtype III-B CRISPR-Cas 5 with the same organization and repeat type (DR3) as four other strains of *Microcystis* (PCC 9701, PCC 7941, DIANCHI905, and PCC 9432) with a strain-specific hypothetical protein between that CRISPR array and a second array. LE013-01 also had 83 restriction enzymes and 3 restriction-modification proteins, 5 of which were strain-

specific.

**Occurrence and expression of strain-specific genes in western Lake Erie**

One metagenomic and five metatranscriptomic data sets from western Lake Erie were mapped against strain-specific gene sequences for 20 *Microcystis* strains to determine the presence and expression of those genes in the environment under different bloom conditions. These bloom conditions do not necessarily represent natural cell growth cycles but are rather operationally defined by trends in biomass and pigment through the season on a basin-wide scale. Over 92% of strain-specific genes from each strain studied here were present in Lake Erie metagenomic data (Table 1). Of these genes found in the LGL strain metagenomes, 168 of 189 LE3, 186 of 222 LE013-01, and 408 of 458 LSC13-02 genes were also present in the metatranscriptomic data of at least one of the three bloom stages. There was no clear enrichment of strain-specific genes from the LGL strains (LE3, LE013-01, and LSC13-02) relative to other *Microcystis* strains from around the world in the Lake Erie metagenome or metatranscriptome.

To evaluate the relative abundance of these strain-specific genes and their transcripts in Lake Erie *Microcystis* cells, we next compared their coverage in the metagenomic and metatranscriptomic data to that of phylogenetic marker genes. Because these marker genes are present in the *Microcystis* genomes in a single copy, for the metagenome this ratio provides an estimate of the fraction of cells containing the strain-specific genes. For the metatranscriptome the ratio provides a qualitative estimate of expression of the strain-specific genes relative to housekeeping genes. We focused on samples in which *Microcystis* was abundant (based on phycocyanin and 16S rRNA gene data [56]), including the metagenomic data from the early bloom and metatranscriptomic data from the early, mid, and late blooms. The frequency of strain-specific genes varied widely in the 2014 *Microcystis* bloom. Approximately 16% of strain-specific genes were present at > 50% of the relative abundance of phylogenetic marker genes, suggesting they are present in the majority of *Microcystis* cells in Lake Erie. The remaining 84% of strain-specific genes are present at < 50% of phylogenetic marker gene relative abundance, indicating that they are more rare in the Lake Erie *Microcystis* populations (Fig 3).
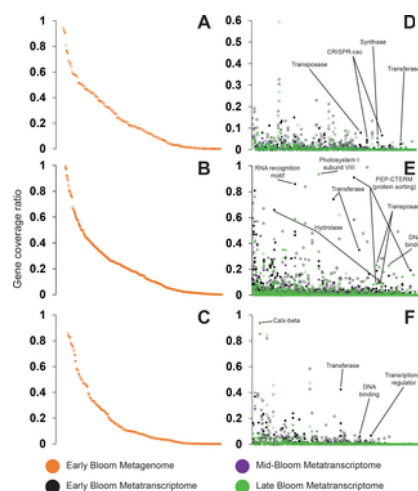


**Fig 3. Rank abundance of the coverage of strain-specific genes from LGL *Microcystis* isolates in metagenomic and metatranscriptomic data from multiple bloom stages of the 2014 Lake Erie *Microcystis* bloom.**
Coverage is presented as the ratio of strain-specific genes compared to average coverage for marker genes used in the phylogenetic analysis. (A, B, C) Metagenomic data, (D, E, F) Metatranscriptomic data with genes ranked by metagenomic abundance. (A, D) strain LE3; (B, E) strain LSC13-02; (C, F) strain LE013-01. Closed circles have an annotated function, open circles are hypothetical proteins.
https://doi.org/10.1371/journal.pone.0183859.g003

Expression patterns of strain-specific genes varied across western Lake Erie and throughout the bloom event. Most strain-specific genes were expressed at levels < 50% of the average of marker genes or not at all. At station WE12 (Toledo water intake crib), the number of strain specific genes expressed at any level decreased from 72% in the early bloom to 58% in the late bloom. In contrast, stations WE2 (Maumee River outflow) and WE4 (offshore, towards central Lake Erie) showed an opposite pattern. At station WE2 31% of strain-specific genes were expressed in the early bloom and 58% in the late bloom and at Station WE4 50% and 67% of strain-specific genes were expressed in the early and late blooms, respectively.

We next analyzed the functional profile of the strain-specific genes in the Lake Erie metagenome and metatranscriptomes. Strain-specific genes in the metagenomic data were dominated by functional categories related to (i) "general function", (ii) replication, recombination and repair; (iii) mobilome; prophages, transposons, and (iv) energy production and conversion (Fig 4). The relative abundance of functional categories was variable across bloom stages in the metatranscriptomic data (Fig 4). Of the strain-specific genes expressed in the metatranscriptomic data with a functional annotation, many were associated with genomic plasticity including transposases, endonucleases, transcriptases, integrases, and CRISPR-Cas proteins. Transposases included two different IS4 family transposases and a DDE domain transpose at the start of a large operon for pigment synthesis in strain LE3, DDE domain transposases in strain LE013-01, and an IS605 OrfB family transposase in strain LSC13-02. Transcripts mapped to strain-specific restriction endonucleases in strains LE3 and LSC13-02 and DNA binding helix-turn-helix endonucleases in all three LGL

strains. Transcripts related to foreign genetic elements included Subtype I-D CRISPR-Cas 2 proteins in strain LE3, a reverse transcriptase and a phase integrase in strain LSC13-02, and a plasmid stabilization ParE toxin of the ParDE toxin-antitoxin system in strain LE013-01.
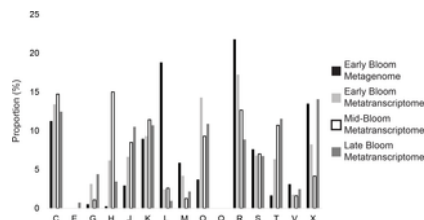


**Fig 4.**
Relative abundance of strain-specific genes in (A) metagenomes and (B) metatranscriptomes from multiple bloom stages of the 2014 Lake Erie *Microcystis* bloom, annotated by COG category. Relative abundance was determined by normalizing the metagenomic and metatranscriptomic strain-specific gene read coverage by the read coverage of phylogenetic marker genes. COG categories are: C- Energy production and conversion, E- Amino acid transport and metabolism, G- Carbohydrate transport and metabolism, H- Coenzyme transport and metabolism, J- Translation, ribosomal structure, and biogenesis, K- Transcription, L- Replication, recombination, and repair, M- Cell wall/membrane/envelope biogenesis, O- Posttranslational modification, protein turnover, chaperones, Q- Secondary metabolites biosynthesis, transport, and catabolism, R- General function prediction only, S- Function unknown, T- Signal transduction mechanisms, V- Defense mechanisms, X- Mobilome: prophages, transposons.
https://doi.org/10.1371/journal.pone.0183859.g004

The next most abundant type of annotated strain-specific genes found in the Lake Erie metatranscriptomic data were associated with transcription or metabolism. Methyltransferases were identified in all three LGL strains, matching *ecoRI* adenine-specific methyltransferase in strain LSC13-02, *fkbM* domain methyltransferases in both LE3 and LE013-01, and DNA methyltransferase (*dcm*) in LE3. Transcripts also matched to transcriptional regulators in the *xre* and *snf*2 families, a colonic acid biosynthesis glycosyltransferase, dolichyl-phosphate-mannose protein, and ribosomal S18 acetylase in strain LE013-01. There were also transcripts that mapped to strain-specific genes in strain LE3 identified as 2Fe-2S ferredoxin, endopeptidase, and protein phosphatase 2C, associated with electron transfer, peptide bond cleavage, and $Mg^{2+}/Mn^{2+}$ -dependent enzymes involved in stress signaling, respectively.

Several strain-specific genes associated with cell walls and transporters were also found in metatranscriptomic data. Two glycosyltransferases involved in cell wall biosynthesis, identified in strain LE013-01, were found in early bloom stages. Late bloom stages contained a strain-specific N-acetylmuramoyl-L-alanine amidase and a penicillin-insensitive murein endopeptidase, identified in LE3, which cleave links in cell wall peptides. There were also tetratricopeptide repeats (TPR) associated with *xisH* and *xisJ* excision proteins, identified in strain LE3, which are associated with lysogeny in cyanobacteria (given that *Microcystis* is a non-heterocystous cyanobacterium). Transporter related proteins included a cation transporter/ATPase identified in strain LE013-01, a wholly strain-specific scaffold in strain LSC13-02 which contained two genes of the Calx-beta domain of sodium-calcium exchangers, and HEAT repeat domain proteins found in strain LSC13-02, which form cytoplasmic intracellular transporters, that were found in all bloom stage metatranscriptomes. Throughout the bloom there was expression of CARDB proteins, which are cell adhesion related domain proteins, identified from the strain LSC13-02 genome.

## Discussion

We compared genomes of *Microcystis* strains isolated from the lower Great Lakes (LGL) to those from strains that were isolated across the globe and to genes and transcripts from communities of mixed strains in blooms of Lake Erie. This improves the molecular understanding of *Microcystis* by demonstrating the diversity and characteristics of the flexible genome of *Microcystis* and the presence and expression of a majority of the flexible genome during a naturally occurring bloom. We identified genes specific to individual strains, which may confer adaptive or functional advantages that play a role in intraspecific competition [33], adaptation in a changing environment [31, 40, 81], and the composition and succession of blooms [82].

Our comparative analysis of 20 *Microcystis* genomes was consistent with results from an earlier study that identified core and pan-genomes of twelve *Microcystis* genomes [31]. Our findings confirm that the core genome of *Microcystis* is both known and described [30, 31] and indicate that the worldwide genomic diversity of *Microcystis* likely exceeds our capability to sequence and identify previously unknown flexible genes [42, 83].

Based on concatenated alignment of housekeeping genes, our phylogenetic results identified four subclades that are congruent with previous studies that used concatenated alignments of multiple different genes including seven housekeeping genes or all 1,989 core genes [30, 31]. In addition, we found that phylogenies based on the content of the flexible genome are congruent with those from housekeeping and core genes (Fig 1). To our knowledge this is the first evidence that the flexible genome is linked to the core genome of *Microcystis*. Taken together, the absence of a consistent geographic distribution of subclades, the open pan-genome with high potential for gene acquisition [26], genome re-arrangement [32, 42, 84] and loss [33, 85, 86], and the consistent phylogenetic relationships of *Microcystis* core and flexible genes suggest genetic cohesiveness of sub-clades that is independent of geography [21].

To explore the environmental relevance of the strain-specific genes identified here, we assessed their presence in metagenomic and metatranscriptomic data at various stages of a 2014 western Lake Erie *Microcystis* bloom. This included samples of the early bloom that were characterized by elevated toxicity and high environmental nitrate concentration and late stages of the bloom that

were characterized by lower toxicity and nitrate concentrations [53, 87]. Intriguingly, the Lake Erie blooms contained strain-specific genes from not only the Great Lakes strains but also from strains isolated from continents across the world. This suggests that the pangenome is truly global, and highlights the value of using the entire pangenome for analyzing and interpreting metagenomic and metatranscriptomic datasets. Because the flexible genome is linked to phylogeny it may be more appropriate to focus on the subclade structure rather than geographic origin when considering and employing reference strains and genomes. The abundance and substantial expression of some of the flexible genes in the environment (Figs 3 and 4) suggests that these genes may be ecologically important, but a key question that remains open is to what extent the flexible genome is linked to phenotype.

Despite strain LE3 having a larger genome than LE013-01 and LSC13-02 it had the fewest strain-specific genes of the three newly sequenced strains. Strain LE3 has been in culture since 1996 while both LE013-01 and LSC13-02 were isolated in 2013. Both physiology and community composition (for non-axenic cultures) can be altered by time in culture [88] and culturing conditions [89, 90] but to our knowledge the impact of time spent in culture on the genomic features or gene composition of cyanobacterial isolates is unknown.

Previous work has shown that the *Microcystis* genome has a large proportion of repeat sequences, transposases, restriction enzymes [26] and insertion sequences [91] compared to other cyanobacteria. This indicates a capacity for *Microcystis* to adapt to different environments through internal genome re-arrangement as well as horizontal gene transfer [39, 40]. These genomic changes can lead to adaptive diversification but may also cause loss of physiological function [40, 83, 91]. A balance is therefore required for *Microcystis* to be able to regulate DNA incorporation and transposition so that essential pathways (core genes) are preserved while allowing for adaptation to environment-specific stressors such as nutrient gradients or phages [40, 47, 87, 92]. Such a combination would ensure genetic stability through the presence and inheritance of genes that are resistant to lateral transfers (thereby remaining for a longer evolutionary period) [93] while allowing for adaptation and survival in an unstable environment via the plastic portions of the genome [40].

Of the strain-specific genes identified in metatranscriptomic data, those found in strain LSC13-02 were most represented followed by strain LE013-01 and then strain LE3, though it should be noted that the metatranscriptomic data represent multiple *Microcystis* genotypes [24]. While the replication of our analyses was not sufficient to assess its statistical significance, the strain-specific genes in the metatranscriptomic data indicated a shift from coenzyme transport & metabolism, and replication, recombination, & repair early in the bloom to translation, ribosomal structure & biogenesis, signal transduction defense mechanisms, and prophages & transposons in the late bloom. The repeatability and physiological significance of this shift (if any) remain to be investigated, but it could reflect shifts in *Microcystis* population structure within blooms [13], for example between strains being better adapted to withstand either top-down [47] or bottom-up [40, 87] stressors.

The three LGL *Microcystis* strains had different genetic capacities for defense against foreign DNA by using restriction-modification and CRISPR-Cas systems (Table 2, Fig 2), similar to previous findings [36, 80, 92]. Strain LE3 has the most diversity in CRISPR-Cas Subtypes, the most DNA restriction-modification enzymes and the fewest strain-specific genes while strain LSC13-02 had the greatest number of restriction enzymes, a strain-specific restriction-modification protein, less CRISPR-Cas diversity, and the most strain-specific genes. CRISPR-Cas systems are a biochemical mechanism for regulation of genetic exchange in a host, but can also serve a role in determining population structure by changing host-phage dynamics within a bloom [36, 94]. Evidence of viral infection was found in the genomes of strains LSC13-02 and LE013-01, which contained DNA-binding prophage proteins and phage integrases, and phage tail sheath proteins, respectively. Viruses can be a significant top-down selective pressure that can drive microdiversity within an environment even at the earliest stages of divergence, yielding a greater number of subpopulations [13, 94]. Resistance to viral infection through a diversity of CRISPR-Cas systems can sustain subpopulation diversity and overall population stability [95]. This is relevant to our study as *Microcystis* spp. often dominate the bloom biomass in western Lake Erie from mid-July through October [41]. Furthermore, previous studies have found that the dominance of different *Microcystis* strains changes both spatially and temporally during the course of the bloom [82, 96–100].

In conclusion, our results demonstrate that strain-specific genes observed in the newly sequenced genomes of cultured strains of LGL *Microcystis* strains are also present and expressed in high abundance in bloom communities from these same bodies of water. The high degree of genomic plasticity in *Microcystis* has been attributed to a strategy for adapting to variable environments, such as those found in many freshwater ecosystems [31]. Our work contributes to the available data of *Microcystis* strains from the LGL and our capacity to interpret the ecological relevance of metagenomic and metatranscriptomic data [101]. Based on our data on the abundance and expression of the flexible genome in natural communities and its phylogenetic link to the core genome we hypothesize that differences in the flexible portion of the *Microcystis* genome encodes ecologically relevant variation between strains. The nature of these phenotypes, and whether they qualify as ecotypes, remains to be determined. However, previous studies suggest that key traits that vary between strains are toxicity, competition for light, and capability for uptake of nitrogen and carbon, and accordingly that shifts in the strain composition of *Microcystis* blooms are associated with changes in the availability of light, nitrogen, and/or carbon [33, 47, 87]. Blooms are typically associated with many microbial antagonists that are potential top-down controls on *Microcystis* [102]. Our results are consistent with a variable genomic capacity to resist foreign genetic elements and represents an adaptive strategy for *Microcystis* in freshwater ecosystems. Such differences between strains would lead to population re-structuring when exposed to either top-down or bottom-up controls and may be a driver of bloom composition throughout bloom events.

## Supporting information

**S1 Fig. Estimates of the core and pan-genome of *Microcystis* based on 3 newly sequenced genomes and 17 publicly available genomes (obtained from IMG).**

A- Core-genome based on exponential decay models of Tetteline et al. (Blue) and Willenbrock et al. (Red) fitted to ten random resamplings of OMCL core-genome clusters. B- Estimate of the pan-genome based on the model of Tettelin et al. fitted to ten random resamplings of OMCL gene clusters.
https://doi.org/10.1371/journal.pone.0183859.s001
(DOCX)

**S1 Text. Comparative genomics pipeline and example commands and scripts.**

For additional scripts check https://github.com/Geo-omics, or contact the corresponding author.
https://doi.org/10.1371/journal.pone.0183859.s002
(DOCX)

## Acknowledgments

## References

1. Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF. Resource partitioning and sympatric differentiation among closely related bacterioplankton. Science. 2008; 320(5879):1081–5. pmid:18497299
   View Article  •  PubMed/NCBI  •  Google Scholar

2. Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC, et al. Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. Proc Natl Acad Sci U S A [Internet]. 2010; 107(6):2383–90. Available from: <Go to ISI>://000274408100008\nhttp://www.pnas.org/content/107/6/2383.full.pdf. pmid:20133593
   View Article  •  PubMed/NCBI  •  Google Scholar

3. Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S, et al. Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. PLoS Genet. 2007; 3(12):2515–28.
   View Article  •  PubMed/NCBI  •  Google Scholar

4. Kashtan N, Roggensack SE, Rodrigue S, Thompson JW, Biller SJ, Coe A, et al. Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. Science [Internet]. 2014; 344(6182):416–20. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24763590. pmid:24763590
   View Article  •  PubMed/NCBI  •  Google Scholar

5. Cohan F. Bacterial species and speciation. Syst Biol. 2001; 50(4):513–24. pmid:12116650
   View Article  •  PubMed/NCBI  •  Google Scholar

6. Malmstrom RR, Coe A, Kettler GC, Martiny AC, Frias-Lopez J, Zinser ER, et al. Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific oceans. ISME J [Internet]. Nature Publishing Group; 2010; 4(10):1252–64. Available from: http://dx.doi.org/10.1038/ismej.2010.60. pmid:20463762
   View Article  •  PubMed/NCBI  •  Google Scholar

7. Moore LR, Rocap G, Chisholm SW. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. Nature. 1998; 393(6684):464–7. pmid:9624000
   View Article  •  PubMed/NCBI  •  Google Scholar

8. Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, Chisholm SW. Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. Science. 2006; 311(5768):1737–40. pmid:16556835
   View Article  •  PubMed/NCBI  •  Google Scholar

9. Sohm JA, Ahlgren NA, Thomson ZJ, Williams C, Moffett JW, Saito MA, et al. Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. ISME J [Internet]. Nature Publishing Group; 2015; 1–13. Available from: http://www.nature.com/doifinder/10.1038/ismej.2015.115.
   View Article  •  PubMed/NCBI  •  Google Scholar

10. Crosbie ND, Pöckl M, Weisse T, Po M. Dispersal and Phylogenetic Diversity of Nonmarine Picocyanobacteria, Inferred from 16S rRNA Gene and *cpcBA*-Intergenic Spacer Sequence Analyses. Appl Environ Microbiol. 2003; 69(9):5716–21. pmid:12957969
    View Article  •  PubMed/NCBI  •  Google Scholar

11. Martiny AC, Tai APK, Veneziano D, Primeau F, Chisholm SW. Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. Environ Microbiol. 2009; 11(4):823–32. pmid:19021692
    View Article  •  PubMed/NCBI  •  Google Scholar

12. Mella-Flores D, Mazard S, Humily F, Partensky F, Mahé F, Bariat L, et al. Is the distribution of *Prochlorococcus* and *Synechococcus* ecotypes in the Mediterranean Sea affected by global warming? Biogeosciences. 2011; 8(9):2785–804.
    View Article  •  PubMed/NCBI  •  Google Scholar

13. Avrani S, Wurtzel O, Sharon I, Sorek R, Lindell D. Genomic island variability facilitates *Prochlorococcus*-virus coexistence. Nature [Internet]. Nature Publishing Group; 2011; 474(7353):604–8. Available from: http://dx.doi.org/10.1038/nature10172. pmid:21720364
   View Article  •  PubMed/NCBI  •  Google Scholar

14. Campbell L, Nolla HA, Vaulot D. The importance of *Prochlorococcus* to community structure in the central North Pacific Ocean.0 Limnol. Oceanogr. 1994; 39: 954–961.
   View Article  •  PubMed/NCBI  •  Google Scholar

15. Shalapyonok R, Olson R, Shalapyonok L. Arabian Sea phytoplankton during Southwest and Northeast monsoons 1995: composition, size structure and biomass from individual cell properties measured by flow cytometry. Deep-Sea Research II 2001; 48(6–7): 1231–1261.
   View Article  •  PubMed/NCBI  •  Google Scholar

16. Paerl HW, Huisman J Blooms like it hot. Science 2008:320:57–58. pmid:18388279
   View Article  •  PubMed/NCBI  •  Google Scholar

17. Dodds WK, Bouska WW, Eitzmann JL, Pilger TJ, Pitts KL, Riley AJ, et al. Eutrophication of U.S. freshwaters: Analysis of potential economic damages. Environ Sci Technol 2009; 43:12–19. pmid:19209578
   View Article  •  PubMed/NCBI  •  Google Scholar

18. O'Neil JM, Davis TW, Burford MA, Gobler CJ. The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. Harmful Algae [Internet]. Elsevier B.V.; 2012; 14:313–34. Available from: http://dx.doi.org/10.1016/j.hal.2011.10.027.
   View Article  •  PubMed/NCBI  •  Google Scholar

19. Simm S, Keller M, Selymesi M, Schleiff E. The composition of the global and feature specific cyanobacterial core-genomes. Front. Microbiol. 2015; 6.
   View Article  •  PubMed/NCBI  •  Google Scholar

20. Fristachi A, Sinclair JL. Occurrence of cyanobacterial harmful algal blooms workgroup report. In: Hudnell KH, editor. Cyanobacterial harmful algal blooms: State of the science and research needs. New York, Springer; 2008. P. 45–103.

21. Harke MJ, Steffen MM, Gobler CJ, Otten TG, Wilhelm SW, Wood SA, et al. A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. Harmful Algae [Internet]. Elsevier B.V.; 2016; 54:4–20. Available from: http://dx.doi.org/10.1016/j.hal.2015.12.007. pmid:28073480
   View Article  •  PubMed/NCBI  •  Google Scholar

22. Bullerjahn GS, McKay RM, Davis TW, Baker DB, Boyer GL, D'Anglada L V., et al. Global solutions to regional problems: Collecting global expertise to address the problem of harmful cyanobacterial blooms. A Lake Erie case study. Harmful Algae [Internet]. Elsevier B.V.; 2016; 54:223–38. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1568988315301657. pmid:28073479
   View Article  •  PubMed/NCBI  •  Google Scholar

23. van Gremberghe I, Leliaert F, Mergeay J, Vanormelingen P, van der Gucht K, Debeer AE, et al. Lack of phylogeographic structure in the freshwater cyanobacterium *Microcystis aeruginosa* suggests global dispersal. PLoS One. 2011; 6(5).
   View Article  •  PubMed/NCBI  •  Google Scholar

24. Janse I, Kardinaal WEA, Meima M, Fastner J, Visser PM, Zwart G. Toxic and nontoxic *Microcystis* colonies in natural populations can be differentiated on the basis of rRNA gene internal transcribed spacer diversity. Appl Environ Microbiol. 2004; 70(7):3979–87. pmid:15240273
   View Article  •  PubMed/NCBI  •  Google Scholar

25. Kaneko T, Nakajima N, Okamoto S, Suzuki I, Tanabe Y, Tamaoki M, et al. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. DNA Res. 2007; 14(6):247–56. pmid:18192279
   View Article  •  PubMed/NCBI  •  Google Scholar

26. Frangeul L, Quillardet P, Castets A-M, Humbert J-F, Matthijs HCP, Cortez D, et al. Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium. BMC Genomics. 2008; 9:274. pmid:18534010
   View Article  •  PubMed/NCBI  •  Google Scholar

27. Zurawell RW, Chen H, Burke JM, Prepas EE. Hepatotoxic cyanobacteria: a review of the biological importance of microcystins in freshwater environments. J Toxicol Environ Health B Crit Rev. 2005; 8(1):1–37. pmid:15762553
   View Article  •  PubMed/NCBI  •  Google Scholar

28. Tillett D, Parker DL, Neilan BA. Detection of toxigenicity by a probe for the Microcystin Synthetase A Gene (*mcyA*) of the cyanobacterial genus *Microcystis*: Comparison of toxicities with 16S rRNA and phycocyanin operon (Phycocyanin intergenic spacer) Phylogenies. Appl Environ Microbiol. 2001; 67(6):2810–8. pmid:11375198
   View Article  •  PubMed/NCBI  •  Google Scholar

29. Otsuka S, Suda S, Li R, Watanabe M, Oyaizu H, Matsumoto S, et al. Phylogenetic relationships between toxic and non-toxic strains of the genus *Microcystis* based on 16S to 23S internal transcribed spacer sequence. FEMS Microbiol Lett. 1999; 172(1):15–21. pmid:10079523
   View Article   •   PubMed/NCBI   •   Google Scholar

30. Tanabe Y, Kasai F, Watanabe MM. Multilocus sequence typing (MLST) reveals high genetic diversity and clonal population structure of the toxic cyanobacterium *Microcystis aeruginosa*. Microbiology. 2007; 153:3695–703. pmid:17975077
   View Article   •   PubMed/NCBI   •   Google Scholar

31. Humbert JF, Barbe V, Latifi A, Gugger M, Calteau A, Coursin T, et al. A tribute to disorder in the genome of the bloom-forming freshwater cyanobacterium *Microcystis aeruginosa*. PLoS One. 2013; 8(8).
   View Article   •   PubMed/NCBI   •   Google Scholar

32. Rantala A, Fewer DP, Hisbergues M, Rouhiainen L, Vaitomaa J, Börner T, et al. Phylogenetic evidence for the early evolution of microcystin synthesis. Proc Natl Acad Sci U S A. 2004; 101(2):568–73. pmid:14701903
   View Article   •   PubMed/NCBI   •   Google Scholar

33. Sandrini G, Matthijs HCP, Verspagen JMH, Muyzer G, Huisman J. Genetic diversity of inorganic carbon uptake systems causes variation in $CO_2$ response of the cyanobacterium *Microcystis*. ISME J [Internet]. Nature Publishing Group; 2014; 8(3):589–600. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24132080. pmid:24132080
   View Article   •   PubMed/NCBI   •   Google Scholar

34. Davis TW, Watson SB, Rozmarynowycz MJ, Ciborowski JJH, McKay RM, Bullerjahn GS. Phylogenies of microcystin-producing cyanobacteria in the lower Laurentian Great Lakes suggest extensive genetic connectivity. PLoS One. 2014; 9(9):e106093. pmid:25207941
   View Article   •   PubMed/NCBI   •   Google Scholar

35. Lefébure T, Stanhope MJ. Evolution of the core and pan-genome of Streptococcus: positive selection, recombination, and genome composition. Genome Biol [Internet]. 2007; 8(5):R71. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17475002. pmid:17475002
   View Article   •   PubMed/NCBI   •   Google Scholar

36. Yang C, Lin F, Li Q, Li T, Zhao J. Comparative genomics reveals diversified CRISPR-Cas systems of globally distributed *Microcystis aeruginosa*, a freshwater bloom-forming cyanobacterium. Front Microbiol [Internet]. 2015; 6(May). Available from: http://www.frontiersin.org/Evolutionary_and_Genomic_Microbiology/10.3389/fmicb.2015.00394/abstract.
   View Article   •   PubMed/NCBI   •   Google Scholar

37. Briand E, Escoffier N, Straub C, Sabart M, Quiblier C, Humbert J-F. Spatiotemporal changes in the genetic diversity of a bloom-forming *Microcystis aeruginosa* (cyanobacteria) population. ISME J. 2009; 3(4):419–29. pmid:19092863
   View Article   •   PubMed/NCBI   •   Google Scholar

38. Tanabe Y, Watanabe MM. Local expansion of a panmictic lineage of water bloom-forming cyanobacterium *Microcystis aeruginosa*. PLoS One. 2011; 6(2).
   View Article   •   PubMed/NCBI   •   Google Scholar

39. Polz MF, Alm EJ, Hanage WP. Horizontal gene transfer and the evolution of bacterial and archaeal population structure. Trends Genet [Internet]. 2013; 29(3):170–5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3760709&tool=pmcentrez&rendertype=abstract\nhttp://www.sciencedirect.com/science/article/pii/S0168952512002107. pmid:23332119
   View Article   •   PubMed/NCBI   •   Google Scholar

40. Steffen MM, Dearth SP, Dill BD, Li Z, Larsen KM, Campagna SR, et al. Nutrients drive transcriptional changes that maintain metabolic homeostasis but alter genome architecture in *Microcystis*. ISME J [Internet]. Nature Publishing Group; 2014; 1–13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24858783.
   View Article   •   PubMed/NCBI   •   Google Scholar

41. Stumpf RP, Wynne TT, Baker DB, Fahnenstiel GL. Interannual variability of cyanobacterial blooms in Lake Erie. PLoS One. 2012; 7(8).
   View Article   •   PubMed/NCBI   •   Google Scholar

42. Steffen MM, Belisle BS, Watson SB, Boyer GL, Wilhelm SW. Status, causes and controls of cyanobacterial blooms in Lake Erie. J Great Lakes Res [Internet]. Elsevier B.V.; 2014; 40:215–25. Available from: http://dx.doi.org/10.1016/j.jglr.2013.12.012.
   View Article   •   PubMed/NCBI   •   Google Scholar

43. Brittain SM, Wang J, Babcock-Jackson L, Carmichael WW, Rinehart KL, Culver DA. Isolation and characterization of microcystins, cyclic heptapeptide hepatotoxins from a Lake Erie Strain of *Microcystis aeruginosa*. J Great Lakes Res. 2000; 26(3):241–9.
   View Article   •   PubMed/NCBI   •   Google Scholar

44. Watson SB, Miller C, Arhonditsis G, Boyer GL, Carmichael W, Charlton MN, et al. The re-eutrophication of Lake Erie: Harmful algal blooms and hypoxia. Harmful Algae [Internet]. 2016; 56(May):44–66. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1568988315301141.
   View Article   •   PubMed/NCBI   •   Google Scholar

45.  Harke MJ, Berry DL, Ammerman JW, Gobler CJ. Molecular response of the bloom-forming cyanobacterium, *Microcystis aeruginosa*, to phosphorus limitation. Microb Ecol. 2012; 63(1):188–98. pmid:21720829
     View Article   •   PubMed/NCBI   •   Google Scholar

46.  Harke MJ, Gobler CJ. Global Transcriptional Responses of the Toxic Cyanobacterium, *Microcystis aeruginosa*, to nitrogen stress, phosphorus stress, and growth on organic matter. PLoS One. 2013; 8(7).
     View Article   •   PubMed/NCBI   •   Google Scholar

47.  Steffen MM, Belisle BS, Watson SB, Boyer GL, Bourbonniere RA, Wilhelm SW. Metatranscriptomic evidence for co-occurring top-down and bottom-up controls on toxic cyanobacterial communities. Appl Environ Microbiol. 2015; (February).
     View Article   •   PubMed/NCBI   •   Google Scholar

48.  Stanier RY, Kunisawa R, Mandel M, Cohen-Bazire G. Purification and properties of unicellular blue-green algae (order *Chroococcales*). Bacteriol Rev. 1971; 35(2):171–205. pmid:4998365
     View Article   •   PubMed/NCBI   •   Google Scholar

49.  Sant'Anna CL, de Carvalho LR, Fiore MF, Silva-Stenico ME, Lorenzi AS, Rios FR, et al. Highly toxic *Microcystis aeruginosa* strain, isolated from São Paulo-Brazil, produce hepatotoxins and paralytic shellfish poison neurotoxins. Neurotox Res [Internet]. 2011; 19(3):389–402. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20376712. pmid:20376712
     View Article   •   PubMed/NCBI   •   Google Scholar

50.  Fiore MF, Alvarenga DO, Varani AM, Hoff-Risseti C, Crespim E, Ramos RTJ, et al. Draft genome sequence of the Brazilian toxic bloom-forming cyanobacterium *Microcystis aeruginosa* Strain SPC777. Genome Announc. 2013; 1(4).
     View Article   •   PubMed/NCBI   •   Google Scholar

51.  Okano K, Miyata N, Ozaki Y. Whole Genome Sequence of the Non-Microcystin-Producing *Microcystis aeruginosa* Strain NIES-44. Genome Announc. 2015; 3(2):e00135–15. pmid:25792056
     View Article   •   PubMed/NCBI   •   Google Scholar

52.  Chen I-MA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, et al. IMG/M: integrated genome and metagenome comparative data analysis system. Nucleic Acids Res [Internet]. 2017; 45:D507–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27738135. pmid:27738135
     View Article   •   PubMed/NCBI   •   Google Scholar

53.  Markowitz VM, Chen I-MA, Chu K, Szeto E, Palaniappan K, Pillay M, et al. IMG/M 4 version of the integrated metagenome comparative analysis system. Nucleic Acids Res. 2014; 42(October 2013):568–73.
     View Article   •   PubMed/NCBI   •   Google Scholar

54.  Markowitz VM, Mavromatis K, Ivanova NN, Chen I-MA, Chu K, Kyrpides NC. IMG ER: A system for microbial genome annotation expert review and curation. Bioinformatics. 2009; 25(17):2271–8. pmid:19561336
     View Article   •   PubMed/NCBI   •   Google Scholar

55.  Allen MM. Simple conditions for growth of unicellular blue-green algae on plates. J. Phycol. 1968; 4:1–3.
     View Article   •   PubMed/NCBI   •   Google Scholar

56.  Berry MA, Cory RM, Davis TW, Duhaime MB, Johengen TH, Kling GW, et al. Cyanobacterial harmful algal blooms are a biological disturbance to Western Lake Erie bacterial communities. Environ. Microbiol. 2017; 19(3):1149–1162. pmid:28026093
     View Article   •   PubMed/NCBI   •   Google Scholar

57.  Chaffin JD, Bridgeman TB, Heckathorn SA, 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.
     View Article   •   PubMed/NCBI   •   Google Scholar

58.  Bridgeman TB, Chaffin JD, Filbrun JE. A novel method for tracking western Lake Erie *Microcystis* blooms, 2002–2011. J Great Lakes Res. 2013; 39(1):83–9.
     View Article   •   PubMed/NCBI   •   Google Scholar

59.  Buffalo, V. Scythe–A Bayesian adapter trimmer [software]. 2014 [cited 2016 August]. Available from: https://github.com/vsbuffalo/scythe.https://github.com/vsbuffalo/scythe

60.  Joshi NA, Fass JN. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files. Version 1.33 [software]. Available from: https://github.com/najoshi/sickle.

61.  Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: A *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics. 2012; 28(11):1420–8. pmid:22495754
     View Article   •   PubMed/NCBI   •   Google Scholar

62. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25(14):1754–60. pmid:19451168
View Article  •  PubMed/NCBI  •  Google Scholar

63. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25(16):2078–9. pmid:19505943
View Article  •  PubMed/NCBI  •  Google Scholar

64. Dick GJ, Andersson AF, Baker BJ, Simmons SL, Thomas BC, Yelton AP, et al. Community-wide analysis of microbial genome sequence signatures. Genome Biol. 2009; 10(8):R85. pmid:19698104
View Article  •  PubMed/NCBI  •  Google Scholar

65. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. J Mol Biol. 1990; 215(3):403–10. pmid:2231712
View Article  •  PubMed/NCBI  •  Google Scholar

66. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 2013; 41(D1):590–6.
View Article  •  PubMed/NCBI  •  Google Scholar

67. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. Nucl. Acids. Res. 2014; 42:D643–48. pmid:24293649
View Article  •  PubMed/NCBI  •  Google Scholar

68. Darling AE, Jospin G, Lowe E, Matsen FA, Bik HM, Eisen JA. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ [software]. 2014; 2:e243. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3897386&tool=pmcentrez&rendertype=abstract. pmid:24482762
View Article  •  PubMed/NCBI  •  Google Scholar

69. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32(5):1792–1797. pmid:15034147
View Article  •  PubMed/NCBI  •  Google Scholar

70. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; 30(12):2725–9. pmid:24132122
View Article  •  PubMed/NCBI  •  Google Scholar

71. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar , et al. ARB: a software environment for sequence data. Nucleic Acids Res [Internet]. 2004; 32(4):1363–71. Available from: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkh293. pmid:14985472
View Article  •  PubMed/NCBI  •  Google Scholar

72. Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006; 22(21):2688–90. pmid:16928733
View Article  •  PubMed/NCBI  •  Google Scholar

73. Felsenstein J. PHYLIP (Phylogeny Inference Package). Version 3.6 [software] 2005. Available from: http://evolution.genetics.washington.edu/phylip.html.

74. Contreras-Moreira B, Vinuesa P. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl Environ Microbiol. 2013; 79(24):7696–701. pmid:24096415
View Article  •  PubMed/NCBI  •  Google Scholar

75. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010; 11(119).
View Article  •  PubMed/NCBI  •  Google Scholar

76. Vinuesa P, Contreras-Moreira B. Robust identification of orthologues and paralogues for microbial pan-genomics using GET_HOMOLOGUES: A case study of pIncA/C Plasmids. In: Mengoni A, Galardini M, Fondi M, editors. Bacterial Pangenomics: Methods and Protocols. New York: Humana Press; 2015. P. 203–32.

77. Fischer S, Brunk BP, Chen F, Gao X, Harb OS, Iodice JB, et al. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. Curr Protoc Bioinforma. 2011; 6.12(SUPPL.35):1–19.
View Article  •  PubMed/NCBI  •  Google Scholar

78. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". Proc Natl Acad Sci U S A. 2005; 102(39):13950–5. pmid:16172379
View Article  •  PubMed/NCBI  •  Google Scholar

79. Willenbrock H, Hallin PF, Wassenaar TM, Ussery DW. Characterization of probiotic *Escherichia coli* isolates with a novel pan-genome microarray. Genome Biol. 2007; 8(12):R267. pmid:18088402

View Article    •    PubMed/NCBI    •    Google Scholar

80. Scholz I, Lange SJ, Hein S, Hess WR, Backofen R. CRISPR-Cas Systems in the Cyanobacterium *Synechocystis* sp. PCC6803 Exhibit Distinct Processing Pathways Involving at Least Two Cas6 and a Cmr2 Protein. PLoS One. 2013; 8(2).
   View Article    •    PubMed/NCBI    •    Google Scholar

81. Horst GP, Sarnelle O, White JD, Hamilton SK, Kaul RB, Bressie JD. Nitrogen availability increases the toxin quota of a harmful cyanobacterium, *Microcystis aeruginosa*. Water Res [Internet]. Elsevier Ltd; 2014; 54:188–98. Available from: http://dx.doi.org/10.1016/j.watres.2014.01.063. pmid:24568788
   View Article    •    PubMed/NCBI    •    Google Scholar

82. Welker M, Šejnohová L, Némethová D, von Döhren H, Jarkovsky J, Maršálek B. Seasonal shifts in chemotype composition of *Microcystis* sp. communities in the pelagial and the sediment of a shallow reservoir. Limnol Oceanogr. 2007; 52(2):609–19.
   View Article    •    PubMed/NCBI    •    Google Scholar

83. Mlouka A, Comte K, Castets AM, 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):2355–65. pmid:15060038
   View Article    •    PubMed/NCBI    •    Google Scholar

84. Kurmayer R, Blom JF, Deng L, Pernthaler J. Integrating phylogeny, geographic niche partitioning and secondary metabolite synthesis in bloom-forming *Planktothrix*. ISME J [Internet]. Nature Publishing Group; 2014; 9(4):909–21. Available from: http://www.nature.com/doifinder/10.1038/ismej.2014.189.
   View Article    •    PubMed/NCBI    •    Google Scholar

85. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, et al. Genome streamlining in a cosmopolitan oceanic bacterium. Science. 2005; 309(5738):1242–5. pmid:16109880
   View Article    •    PubMed/NCBI    •    Google Scholar

86. Visser PM, Verspagen JMH, Sandrini G, Stal LJ, Matthijs HCP, Davis TW, et al. How rising $CO_2$ and global warming may stimulate harmful cyanobacterial blooms. Harmful Algae [Internet]. Elsevier B.V.; 2016; 54:145–59. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1568988315301694. pmid:28073473
   View Article    •    PubMed/NCBI    •    Google Scholar

87. Gobler CJ, Burkholder JM, Davis TW, Harke MJ, Stow CA, Van de Waal DB. The dual role of nitrogen supply in controlling the growth and toxicity of cyanobacterial blooms. Harmful Algae [Internet]. Elsevier B.V.; 2016; 54:87–97. Available from: http://dx.doi.org/10.1016/j.hal.2016.01.010 pmid:28073483
   View Article    •    PubMed/NCBI    •    Google Scholar

88. Ferguson RL, Buckley EN, Palumbo AV. Response of marine bacterioplankton to differential filtration and confinement. Appl Environ Microbiol [Internet]. 1984 Jan; 47(1):49–55. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=239610&tool=pmcentrez&rendertype=abstract. pmid:6696422
   View Article    •    PubMed/NCBI    •    Google Scholar

89. Massana R, Pedrós-Alió C, Casamayor EO, Gasol JM. Changes in marine bacterioplankton phylogenetic composition during incubations designed to measure biogeochemically significant parameters. Limnol Oceanogr. 2001; 46(5):1181–8.
   View Article    •    PubMed/NCBI    •    Google Scholar

90. Meyer KA. Microbial interactions and ecology within blooms of the toxic dinoflagellate Karenia brevis on the West Florida Shelf [dissertation]. College Park (MD): The University of Maryland; 2013.

91. Lin S, Haas S, Zemojtel T, Xiao P, Vingron M, Li R. Genome-wide comparison of cyanobacterial transposable elements, potential genetic diversity indicators. Gene [Internet]. Elsevier B.V.; 2011; 473(2):139–49. Available from: http://dx.doi.org/10.1016/j.gene.2010.11.011. pmid:21156198
   View Article    •    PubMed/NCBI    •    Google Scholar

92. Kuno S, Yoshida T, Kaneko T, Sako Y. Intricate interactions between the bloom-forming cyanobacterium *Microcystis aeruginosa* and foreign genetic elements, revealed by diversified clustered regularly interspaced short palindromic repeat (CRISPR) signatures. Appl Environ Microbiol. 2012; 78(15):5353–60. pmid:22636003
   View Article    •    PubMed/NCBI    •    Google Scholar

93. Daubin V, Gouy M, Perrière G. A phylogenomic approach to bacterial phylogeny: Evidence of a core of genes sharing a common history. Genome Res. 2002; 12(7):1080–90. pmid:12097345
   View Article    •    PubMed/NCBI    •    Google Scholar

94. Kuno S, Sako Y, Yoshida T. Diversification of CRISPR within coexisting genotypes in a natural population of the bloom-forming cyanobacterium *Microcystis aeruginosa*. Microbiology. 2014; 160(2014):903–16.
   View Article    •    PubMed/NCBI    •    Google Scholar

95. Steffen MM, Davis TW, McKay RM, Bullerjahn GS, Krausfeldt LE, Stough JMA, et al. Ecophysiological examination of the Lake Erie Microcystis bloom in 2014: linkages between biology and the water supply shutdown of Toledo, Ohio. Environ Sci Technol. 2017.
View Article  •  PubMed/NCBI  •  Google Scholar

96. Childs LM, England WE, Young MJ, Weitz JS, Whitaker RJ. CRISPR-Induced Distributed Immunity in Microbial Populations. PLoS One [Internet]. 2014; 9(7):e101710. Available from: http://dx.plos.org/10.1371/journal.pone.0101710. pmid:25000306
View Article  •  PubMed/NCBI  •  Google Scholar

97. Welker M, Von Döhren H, Täuscher H, Steinberg CEW, Erhard M. Toxic *Microcystis* in shallow lake Müggelsee (Germany)–dynamics, distribution, diversity. Arch für Hydrobiol. 2003; 157(2):227–48.
View Article  •  PubMed/NCBI  •  Google Scholar

98. Davis TW, Berry DL, Boyer GL, Gobler CJ. The effects of temperature and nutrients on the growth and dynamics of toxic and non-toxic strains of *Microcystis* during cyanobacteria blooms. Harmful Algae. 2009; 8(5):715–25.
View Article  •  PubMed/NCBI  •  Google Scholar

99. Davis TW, Harke MJ, Marcoval MA, Goleski J, Orano-Dawson C, Berry DL, et al. Effects of nitrogenous compounds and phosphorus on the growth of toxic and non-toxic strains of *Microcystis* during cyanobacterial blooms. Aquat Microb Ecol. 2010; 61(2):149–62.
View Article  •  PubMed/NCBI  •  Google Scholar

100. Otten TG, Xu H, Qin B, Zhu G, Paerl HW. Spatiotemporal patterns and ecophysiology of toxigenic *Microcystis* blooms in Lake Taihu, China: Implications for water quality management. Environ Sci Technol. 2012; 46(6):3480–8. pmid:22324444
View Article  •  PubMed/NCBI  •  Google Scholar

101. Scanlan DJ, Ostrowski M, Mazard S, Dufresne A, Garczarek L, Hess WR, et al. Ecological genomics of marine picocyanobacteria. Microbiol Mol Biol Rev. 2009; 73(2):249–99. pmid:19487728
View Article  •  PubMed/NCBI  •  Google Scholar

102. van Wichelen J, Vanormelingen P, Codd GA, Vyverman W. The common bloom-forming cyanobacterium *Microcystis* is prone to a wide array of microbial antagonists. Harmful Algae [Internet]. Elsevier B.V.; 2016; 55:97–111. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1568988315301682. pmid:28073551
View Article  •  PubMed/NCBI  •  Google Scholar