

# Closed circular genome sequence of a *Microcystis aeruginosa* PCC7806 $\Delta$ *mcyB* (UTK) non-toxic mutant

Gwendolyn F. Stark,<sup>1</sup> Alexander R. Truchon,<sup>1</sup> Elke Dittmann,<sup>2</sup> Steven W. Wilhelm<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** Here we report the complete, closed genome of the non-toxic *Microcystis aeruginosa* PCC7806  $\Delta$ *mcyB* mutant strain. This genome is 5,103,923 bp long, with a GC content of 42.07%. Compared to the published wild-type genome (*Microcystis aeruginosa* PCC7806SL), there is evidence of accumulated mutations beyond the inserted chloramphenicol resistance marker.

**KEYWORDS** cyanobacteria, harmful algal blooms, cyanotoxins, microcystin

*Microcystis aeruginosa* PCC7806  $\Delta$ *mcyB* is a non-toxic strain available at the Pasteur Culture Collection (1). This mutant is widely used in comparative studies with the toxic wild-type strain PCC7806 (2–15). Except for a chloramphenicol cassette inserted in the *mcyB* gene, the mutant and wild-type genomes were thought to be identical, although this had not been confirmed. Whole-genome sequencing revealed that our strain, which has undergone live passage and cryopreservation (–150°C) for the last 7 years after an aliquot was transferred from Potsdam (Germany) to Knoxville (USA), has accumulated mutations relative to the published 7806SL genome—we have added the “UTK” annotation to signify our specific isolate. Of note, our mutant strain has an additional putative transposase gene (*tpnA*), not seen in the 7806SL wild type, inserted in an SLC13-family permease gene. There is also a region of the genome that assembled in a manner suggesting a large inversion (~2.5 Mbp) has occurred. This paper serves as a cautionary tale for labs to routinely re-sequence strains.

Axenic cultures of *Microcystis aeruginosa* PCC7806  $\Delta$ *mcyB* (UTK) were maintained in 50 mL of modified CT medium (B-glycerophosphate was replaced with 1 mL of Na<sub>2</sub>HPO<sub>4</sub> (23.19 mg/mL) and grown under a 14-h/10-h day/night cycle (~35  $\mu$ mol photons/m<sup>2</sup>/s). Late-log phase cells were pelleted (15,000  $\times$  g, 10 min) and resuspended in 400  $\mu$ L TE buffer. RNase (0.6  $\mu$ L of 10 mg/mL) and 120  $\mu$ L of lysozyme (20 mg/mL) were added, and cells were incubated at 37°C for 30 min. Then, 6  $\mu$ L of proteinase K (20 mg/mL) in 3 mM CaCl<sub>2</sub> and 200 mM Tris buffer and 39.5  $\mu$ L of 20% SDS were added before incubation at 55°C for 2 h. DNA was extracted via a phenol-chloroform method (16). Half the extracted DNA was used for long-read sequencing (DNA was not sheared), and the other half was used for Illumina short-read sequencing. Long reads were sequenced in-house using a MinION R9.4.1 flow cell (Oxford Nanopore Technologies). Library prep for high-molecular weight DNA was prepared with a ligation sequencing kit, SQK-LSK110, per manufacturer protocols, which enriched for reads >3 kb, but no formal size selection was performed (Oxford Nanopore Technologies). Long-read sequencing resulted in 320,921 raw reads (N<sub>50</sub> = 17.55 kb). For Illumina sequencing, DNA was sent to SeqCenter (Pittsburgh, PA). Sample libraries were prepared using an Illumina DNA prep kit and IDT’s 10-bp unique dual indices and were processed on an Illumina NextSeq 2000, generating 8,278,500 151-bp paired-end reads. Demultiplexing, quality control, and adapter trimming were performed with bcl-convert using default parameters (v.3.9.3) (Illumina) (Table 1).

**Editor** Elinne Becket, California State University San Marcos, San Marcos, California, USA

Address correspondence to Steven W. Wilhelm, wilhelm@utk.edu.

The authors declare no conflict of interest.

See the funding table on p. 2.

**Received** 1 August 2023

**Accepted** 14 September 2023

**Published** 19 October 2023

Copyright © 2023 Stark et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

**TABLE 1** Genomic comparison of PCC7806 *ΔmcyB*-UTK to the published annotated genome 7806SL<sup>a</sup>

|                         | <i>Microcystis aeruginosa</i> PCC7806<br><i>ΔmcyB</i> -UTK | <i>Microcystis aeruginosa</i><br>PCC7806SL |
|-------------------------|------------------------------------------------------------|--------------------------------------------|
| Genome size (total bp)  | 5,103,923                                                  | 5,139,339                                  |
| GC content (%)          | 42.07                                                      | 42.09                                      |
| <i>tnpA</i> transposons | 16                                                         | 15                                         |

<sup>a</sup> We note the version of Prokaryote Genome Annotation Pipeline (PGAP) we used to annotate PCC7806 *ΔmcyB*-UTK failed to annotate one copy of the 243-bp *tnpA* gene that was annotated in the PCC7806SL published genome: we confirmed its presence manually.

For genome assembly, default parameters were used for all software. Bases were called using Guppy (v.6.01) (17), with the dna\_r9.4.1\_450bps\_fast.cfg config file, and adapter sequences were trimmed using Porechop (v.0.2.4) (18), then filtered using NanoFilt (v.2.8.0) at a quality of 9 and a length of 500 (19). Illumina reads were trimmed using CLC Genomics Workbench (v. 21.0.4). Genome assembly using long and short reads was performed *de novo* using Unicycler (v.0.4.9b), using a predicted length of 5 Mbp (20). A circular contig was assembled, circularized, trimmed, and rotated by Unicycler. The final genome assembly was annotated using National Center for Biotechnology Information PGAP (annotation software revision: 2022-12-13.build6469), and genomic GC content was determined with QUAST (v.4.4)(21).

## ACKNOWLEDGMENTS

We thank Robbie Martin and Liz Denison for their assistance.

This work was supported by grants from the National Institute of Environmental Health Sciences (grant number 1P01ES02328939-01) and the National Science Foundation (OCE-1840715). We also acknowledge the support of the Simons Foundation (735077).

## AUTHOR AFFILIATIONS

<sup>1</sup>Department of Microbiology, University of Tennessee Knoxville, Knoxville, Tennessee, USA

<sup>2</sup>Department of Microbiology, Institute for Biochemistry and Biology, University of Potsdam, Potsdam, Germany

## AUTHOR ORCIDs

Steven W. Wilhelm  <http://orcid.org/0000-0001-6283-8077>

## FUNDING

| Funder                                                                  | Grant(s)          | Author(s)         |
|-------------------------------------------------------------------------|-------------------|-------------------|
| HHS   NIH   National Institute of Environmental Health Sciences (NIEHS) | 1P01ES02328939-01 | Steven W. Wilhelm |
| National Science Foundation (NSF)                                       | OCE-1840715       | Steven W. Wilhelm |
| Simons Foundation (SF)                                                  | 735077            | Steven W. Wilhelm |

## AUTHOR CONTRIBUTIONS

Gwendolyn F. Stark, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Alexander R. Truchon, Investigation, Methodology, Writing – review and editing | Elke Dittmann, Resources, Writing – review and editing | Steven W. Wilhelm, Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Writing – review and editing

## DATA AVAILABILITY

All sequencing data are available on National Center for Biotechnology Information under the BioProject ID [PRJNA995633](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA995633). Raw reads can be found at the Sequence Read Archive (SRA) under accession numbers [SRX21194812](https://www.ncbi.nlm.nih.gov/sra/SRX21194812) (Nanopore) and [SRX21194811](https://www.ncbi.nlm.nih.gov/sra/SRX21194811) (Illumina). The annotated genome assembly can be found under [GCA\\_030553035.1](https://www.ncbi.nlm.nih.gov/genome/GCA_030553035.1).

## REFERENCES

- Dittmann E, Neilan BA, Erhard M, von Döhren H, Börner T. 1997. Insertional mutagenesis of a peptide synthetase gene that is responsible for hepatotoxin production in the cyanobacterium *Microcystis aeruginosa* PCC 7806. *Mol Microbiol* 26:779–787. <https://doi.org/10.1046/j.1365-2958.1997.6131982.x>
- Barchewitz T, Guljamow A, Meissner S, Timm S, Henneberg M, Baumann O, Hagemann M, Dittmann E. 2019. Non-canonical localization of RubisCO under high-light conditions in the toxic cyanobacterium *Microcystis aeruginosa* PCC7806. *Environ Microbiol* 21:4836–4851. <https://doi.org/10.1111/1462-2920.15301>
- Zuo J, Hu L, Shen W, Zeng J, Li L, Song L, Gan N. 2021. The involvement of  $\alpha$ -proteobacteria *Phenylobacterium* in maintaining the dominance of toxic *Microcystis* blooms in lake Taihu, China. *Environ Microbiol* 23:1066–1078. <https://doi.org/10.1111/1462-2920.15301>
- Mallia V, Verhaegen S, Styriahave B, Eriksen GS, Johannsen ML, Ropstad E, Uhlig S. 2020. Microcystins and *Microcystis aeruginosa* PCC7806 extracts modulate steroidogenesis differentially in the human H295R adrenal model. *PLoS One* 15:e0244000. <https://doi.org/10.1371/journal.pone.0244000>
- Tillett D, Dittmann E, Erhard M, von Döhren H, Börner T, Neilan BA. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide–polyketide synthetase system. *Chem Biol* 7:753–764. [https://doi.org/10.1016/s1074-5521\(00\)00021-1](https://doi.org/10.1016/s1074-5521(00)00021-1)
- Hesse K, Dittmann E, Börner T. 2001. Consequences of impaired microcystin production for light-dependent growth and pigmentation of *Microcystis aeruginosa* PCC 7806. *FEMS Microbiol. Ecol* 37:39–43. <https://doi.org/10.1111/j.1574-6941.2001.tb00851.x>
- Jähnichen S, Ihle T, Petzoldt T, Benndorf J. 2007. Impact of inorganic carbon availability on microcystin production by *Microcystis aeruginosa* PCC 7806. *Appl Environ Microbiol* 73:6994–7002. <https://doi.org/10.1128/AEM.01253-07>
- Briand E, Bormans M, Quiblier C, Salençon M-J, Humbert J-F. 2012. Evidence of the cost of the production of microcystins by *Microcystis aeruginosa* under differing light and nitrate environmental conditions. *PLoS One* 7:e29981. <https://doi.org/10.1371/journal.pone.0029981>
- Zilliges Y, Kehr J-C, Mikkat S, Bouchier C, de Marsac NT, Börner T, Dittmann E. 2008. An extracellular glycoprotein is implicated in cell-cell contacts in the toxic cyanobacterium *Microcystis aeruginosa* PCC 7806. *J Bacteriol* 190:2871–2879. <https://doi.org/10.1128/JB.01867-07>
- Zilliges Y, Kehr J-C, Meissner S, Ishida K, Mikkat S, Hagemann M, Kaplan A, Börner T, Dittmann E. 2011. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of *Microcystis* under oxidative stress conditions. *PLoS One* 6:e17615. <https://doi.org/10.1371/journal.pone.0017615>
- Van de Waal DB, Verspagen JM, Finke JF, Vournazou V, Immers AK, Kardinaal WEA, Tonk L, Becker S, Van Donk E, Visser PM, Huisman J. 2011. Reversal in competitive dominance of a toxic versus non-toxic Cyanobacterium in response to rising CO<sub>2</sub>. *ISME J* 5:1438–1450. <https://doi.org/10.1038/ismej.2011.28>
- Meissner S, Steinhäuser D, Dittmann E. 2015. Metabolomic analysis indicates a pivotal role of the hepatotoxin microcystin in high light adaptation of *Microcystis*. *Environ Microbiol* 17:1497–1509. <https://doi.org/10.1111/1462-2920.12565>
- Urrutia-Cordero P, Agha R, Cirés S, Lezcano MÁ, Sánchez-Contreras M, Waara KO, Utkilen H, Quesada A. 2013. Effects of harmful cyanobacteria on the freshwater pathogenic free-living amoeba *Acanthamoeba castellanii*. *Aquat Toxicol* 130–131:9–17. <https://doi.org/10.1016/j.aquatox.2012.12.019>
- Dittmann E, Erhard M, Kaebbernick M, Scheler C, Neilan BA, von Döhren H, Börner T. 2001. Altered expression of two light-dependent genes in a Microcystin-lacking mutant of *Microcystis aeruginosa* PCC7806. *Microbiology (Reading)* 147:3113–3119. <https://doi.org/10.1099/00221287-147-11-3113>
- Alexova R, Dang TC, Fujii M, Raftery MJ, Waite TD, Ferrari BC, Neilan BA. 2016. Specific global responses to N and Fe nutrition in toxic and non-toxic *Microcystis aeruginosa*. *Environ Microbiol* 18:401–413. <https://doi.org/10.1111/1462-2920.12958>
- Martin RM, Wilhelm SW. 2018. Phenol/chloroform extraction of DNA from Cyanobacteria V.2. <https://doi.org/10.17504/protocols.io.ptndnme>
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for oxford nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex minion sequencing. *Microb Genom* 3:e000132. <https://doi.org/10.1099/mgen.0.000132>
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. Nanopack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>