



# Metagenome-Assembled Genome Sequences of *Raphidiopsis raciborskii* and *Planktothrix agardhii* from a Cyanobacterial Bloom in Kissena Lake, New York, USA

Robbie M. Martin,<sup>a</sup> Michael Kausch,<sup>b</sup> Kimarie Yap,<sup>b</sup> John D. Wehr,<sup>b</sup> Gregory L. Boyer,<sup>c</sup>  Steven W. Wilhelm<sup>a</sup>

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

<sup>b</sup>Department of Biological Sciences, Fordham University, Bronx, New York City, New York, USA

<sup>c</sup>Department of Chemistry, State University of New York College of Environmental Science and Forestry, Syracuse, New York, USA

**ABSTRACT** *Raphidiopsis raciborskii* and *Planktothrix agardhii* are filamentous, potentially toxin-producing cyanobacteria that form nuisance blooms in fresh waters. Here, we report high-quality metagenome-assembled genome sequences of *R. raciborskii* and *P. agardhii* collected from a bloom in Kissena Lake, New York.

**R***aphidiopsis raciborskii* (formerly *Cylindrospermopsis raciborskii*) (1) is a filamentous, nitrogen-fixing cyanobacterium that forms nuisance blooms in fresh waters and is considered an invasive species in temperate ecosystems (2). Its geographic range has expanded in recent decades due to rising lake temperatures and its tolerance of diverse light and nutrient conditions (2). *Planktothrix agardhii* is a filamentous cyanobacterium found in eutrophic, polymictic lakes of temperate zones. Some strains produce microcystin (3). Both are species of concern in North American lakes (4, 5), yet genome sequences from North American strains are underrepresented in publicly available nucleotide archives. Here, we present metagenome-assembled genome sequences of *R. raciborskii* and *P. agardhii* from Kissena Lake, NY.

Kissena Lake is a 3.4-ha natural lake located in Queens, NY (40.748°N, 73.806°W). Kissena Lake is subject to dense cyanobacterial populations through most of the summer months. Samples were collected on 26 September 2019 during a cyanobacterial bloom microscopically determined as dominated by *R. raciborskii*. From a shoreline position, multiple casts of a phytoplankton net were used to concentrate 1.5 liters of surface seston. The seston was transported to the lab on ice. Two samples (KL1 and KL2) were collected by filtering 300 ml of seston each onto 0.4- $\mu$ m-pore-size polycarbonate filters. The filters were immediately stored at  $-20^{\circ}\text{C}$ .

Total DNA was extracted using a ZymoBIOMICS DNA/RNA miniprep kit according to the manufacturer's instructions. DNA libraries were prepared and sequenced by the Microbial Genome Sequencing Center, LLC (Pittsburgh, PA). Libraries were prepared using the Illumina Nextera kit and sequenced on the Illumina NextSeq 550 genome sequencer, producing 150-bp paired-end reads. Sequencing produced a total of 14,107,668 (KL1) and 13,236,752 (KL2) reads.

The reads were trimmed with CLC Genomics Workbench v10.1.1 using default settings. metaSPAdes, MaxBin2, CheckM, and FastANI were implemented in KBase using default parameters unless noted (6–10). *De novo* assembly was conducted with metaSPAdes v3.13.0 using a minimum contig length of 2,000 bp. MaxBin2 v2.2.4 was used to recover metagenome-assembled genomes (MAGs). MAG quality was evaluated with CheckM v1.0.18. The sequencing depth of each MAG was estimated by mapping source reads to their respective MAGs. The genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.13 (11). Taxonomy was assigned by average nucleotide identity (ANI) comparison calculated by FastANI v0.1.2.

**Citation** Martin RM, Kausch M, Yap K, Wehr JD, Boyer GL, Wilhelm SW. 2021. Metagenome-assembled genome sequences of *Raphidiopsis raciborskii* and *Planktothrix agardhii* from a cyanobacterial bloom in Kissena Lake, New York, USA. Microbiol Resour Announc 10:e01380-20. <https://doi.org/10.1128/MRA.01380-20>.

**Editor** Frank J. Stewart, Georgia Institute of Technology

**Copyright** © 2021 Martin 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/).

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

**Received** 1 December 2020

**Accepted** 18 December 2020

**Published** 14 January 2021

From KL1, we recovered a MAG of *Raphidiopsis raciborskii* comprised of 100 contigs with a combined length of 3,657,391 bp and 40.1% GC content. The sequencing depth was ~260-fold. CheckM indicated that the genome was 99.7% complete with 0.3% contamination. The ANI was 99% with *R. raciborskii* UTEX LB2897 (USA), 96% with strains from Brazil (ITEP-A1, CYLP, CYRF), and 94% with CS-505 (Australia). PGAP predicted 3,208 coding DNA sequences (CDS), 3,115 proteins, 38 tRNAs, and 4 noncoding RNAs (ncRNAs). From KL2, we recovered a MAG of *Planktothrix agardhii* comprised of 334 contigs with a combined length of 4,765,279 bp and 39.6% GC content. The sequencing depth was ~17-fold. CheckM indicated the genome to be 99.6% complete with 0.65% contamination. The ANI was 97% with *P. agardhii* NIVA-CYA 15 (Norway). PGAP predicted 4,402 CDS, 4,285 proteins, 37 tRNAs, and 4 rRNAs.

**Data availability.** The sequences are deposited at DDBJ/ENA/GenBank under the accession no. [JADQCS000000000](https://doi.org/10.1093/bioinformatics/btv638) and [JADQCT000000000](https://doi.org/10.1093/bioinformatics/btv638). The reads are deposited at the NCBI Sequence Read Archive under the accession no. [PRJNA675899](https://doi.org/10.1093/bioinformatics/btv638).

## ACKNOWLEDGMENTS

This work was supported by grants from NIEHS (1P01ES028939-01) and NSF (OCE-1840715) to S.W.W. and G.L.B. through the Great Lakes Center for Fresh Waters and Human Health at Bowling Green State University.

## REFERENCES

1. Aguilera A, Gómez EB, Kaštovský J, Echenique RO, Salerno GL. 2018. The polyphasic analysis of two native *Raphidiopsis* isolates supports the unification of the genera *Raphidiopsis* and *Cylindrospermopsis* (Nostocales, Cyanobacteria). *Phycologia* 57:130–146. <https://doi.org/10.2216/17-2.1>.
2. Antunes JT, Leao PN, Vasconcelos VM. 2015. *Cylindrospermopsis raciborskii*: review of the distribution, phylogeography, and ecophysiology of a global invasive species. *Front Microbiol* 6:473. <https://doi.org/10.3389/fmicb.2015.00473>.
3. Kurmayer R, Deng L, Entfellner E. 2016. Role of toxic and bioactive secondary metabolites in colonization and bloom formation by filamentous cyanobacteria *Planktothrix*. *Harmful Algae* 54:69–86. <https://doi.org/10.1016/j.hal.2016.01.004>.
4. McKay RM, Frenken T, Diep N, Cody WR, Crevecoeur S, Dove A, Drouillard KG, Ortiz X, Wintermute J, Zastepa A. 2020. Bloom announcement: an early autumn cyanobacterial bloom co-dominated by *Aphanizomenon flos-aquae* and *Planktothrix agardhii* in an agriculturally-influenced Great Lakes tributary (Thames River, Ontario, Canada). *Data Brief* 30:105585. <https://doi.org/10.1016/j.dib.2020.105585>.
5. Chapman AD, Schelske CL. 1997. Recent appearance of *Cylindrospermopsis* (Cyanobacteria) in five hypereutrophic Florida lakes. *J Phycol* 33:191–195. <https://doi.org/10.1111/j.0022-3646.1997.00191.x>.
6. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>.
7. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607. <https://doi.org/10.1093/bioinformatics/btv638>.
8. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
9. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
10. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, et al. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>.
11. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.