AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Draft genome sequences of three *Planktothrix rubescens* strains cultivated from a eutrophic lake in Norwalk, Ohio (USA)

Ryan S. Wagner,[1,2] **George S. Bullerjahn**[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Draft genomes were generated for three filamentous toxin-producing cyanobacterial strains cultivated from aquatic sources in Ohio sequenced by NovaSeq S4. Here, we report the classification and genome statistics of *Planktothrix rubescens* PR221, PR222, and PR223.

KEYWORDS cyanobacteria, cyanotoxin, freshwater, harmful algal blooms

Planktothrix spp. are filamentous cyanobacteria commonly found in eutrophic waters, and many species produce toxins (1). Two of the most common taxa are *Planktothrix agardhii* and *Planktothrix rubescens*, which occupy different niches in freshwaters. *P. rubescens* is found in the metalimnion while *P. agardhii* the epilimnion (2, 3). Both taxa contain gas vesicles allowing them to migrate through the water column to their optimal light intensity (4). Because *P. rubescens* has been shown to be toxic (1), understanding under what conditions *P. rubescens* grows will be valuable for protecting water supplies (5).

Skinn Lake is a former rock quarry located in Norwalk, Ohio (41.22 N, −82.63 W), with a depth of 12 m. The lake sees annual winter *Planktothrix rubescens* blooms, which become visible during mixing events. Grab samples were taken from the shoreline using a bottle on a pole, placed on ice, and brought back for culturing of dominant taxa. Preliminary identification used microscopy to identify species. After identification and cultivation from three separate grab samples, preparation of xenic, unialgal cultures was done from each sample using single filament picking with a glass pipette in liquid JM media (6) and incubated at 12°C at 10 µmol quanta m^{-2} s^{-1} with a light dark cycle of 14–10 h. After reculturing of selected single filaments, DNA was extracted by filtering biomass through a 0.22-µm Sterivex filter cartridge. Filters were frozen at −80°C before DNA extraction. Sterivex cartridges were then cut open, and the filters were removed for extraction with the DNeasy PowerWater Kit (Qiagen, Germantown, MD) according to the manufacturer's instructions. DNA libraries were prepared and sequenced by the University of Minnesota Genomics Center (Minneapolis, MN). Libraries were created using the Illumina NexteraXT Kit and sequenced on a NovaSeq S4 to generate 150-bp paired-end reads yielding 307,960,113, 285,974,114, and 278,521,261 total sequences for PR221, PR222, and PR223 respectively (Table 1).

Reads were uploaded to the DOE Systems Biology Knowledgebase (KBase) v2.7.11 (7) for further analysis, and all parameters were run using default settings. Reads were then trimmed to remove low-quality base calls, and adapters were clipped using Trimmomatic v0.36 (8). After trimming, the quality of reads was assessed using FastQC v0.11.9 (9). Samples were assembled with a minimum of 2,000 bp using metaSPAdes v3.15.3 (10). MaxBin2 v2.2.4 (11), MetaBAT2 v1.7 (12), and CONCOCT v1.1 (13) were used to bin assemblies into contigs using default settings. Contigs were optimized using the DAS-tool v1.1.2 (14) and quality assessed using CheckM v1.0.18 (15). Next, bins were filtered for completeness (90%) and contamination (≤5%) with CheckM. Extracted bins

Editor Julia A. Maresca, SUNY College of Environmental Science and Forestry, Syracuse, New York, USA

Address correspondence to George S. Bullerjahn, bullerj@bgsu.edu.

The authors declare no conflict of interest.

Received 12 July 2024
Accepted 16 August 2024
Published 9 September 2024

TABLE 1 Genome statistics on Skinn Lake *P. rubescens*

Characteristic	*Planktothrix rubescens*		
	PR221	PR222	PR223
Genome length	5,534,738	5,533,187	5,567,603
No. of contigs	134	125	140
GC content (%)	39.42	39.38	39.43
Completeness (%)	99.34	99.56	99.78
Contamination (%)	2.29	2.29	2.29
N_{50} (bp)	66,822	66,509	68,205
FastANI reference	GCF_900009275.1	GCF_900009275.1	GCF_900009275.1
FastANI organism name	*Planktothrix rubescens* NIVA-CYA 18	*Planktothrix rubescens* NIVA-CYA 18	*Planktothrix rubescens* NIVA-CYA 18
FastANI ANI (%)	98.23	98.27	98.29
Assembly depth of coverage	392.846	414.208	387.251
NCBI PGAP			
CDS	4,971	4,988	5,041
tRNAs	38	38	38
rRNAs	2	3	3
ncRNAs	4	4	4
Totals	5,015	5,033	5,086

were run through QUAST v4.4 (16) for final quality assessment. Taxonomy was assigned using average nucleotide identity against the publicly available reference genome, *Planktothrix rubescens* NIVA-CYA 18 with FASTANI v0.1.3 (17) and through GTDB-Tk v2.3.2 (18). Genomes were submitted to NCBI for the Prokaryotic Genome Annotation Pipeline (PGAP) v6.7 (19).

ACKNOWLEDGMENTS

We would also like to thank the Skinn family, without whom this research would not be possible.

This work was supported by grants from the NIEHS (P01ES028939) and NSF (OCE-1840715) through the Great Lakes Center for Freshwaters and Human Health at Bowling Green State University.

AUTHOR AFFILIATIONS

[1]Department of Biological Sciences, Bowling Green State University, Bowling Green, Ohio, USA
[2]Great Lakes Center for Fresh Waters and Human Health, Bowling Green State University, Bowling Green, Ohio, USA

AUTHOR ORCIDs

George S. Bullerjahn http://orcid.org/0000-0001-5319-7896

DATA AVAILABILITY

The metagenome-assembled genomes have been deposited in GenBank under the accession numbers JAYWIL000000000, JAYWIM000000000, and JAYWIN000000000 for PR221, PR222, and PR223, respectively. The raw sequence files are available as sequence read archives under SRR27742818, SRR27742817, and SRR27742816. The BioProject can be found under the reference PRJNA1050598.

REFERENCES

1. Kurmayer R, Gumpenberger M. 2006. Diversity of microcystin genotypes among populations of the filamentous cyanobacteria *Planktothrix rubescens* and *Planktothrix agardhii*. Mol Ecol 15:3849–3861. https://doi.org/10.1111/j.1365-294X.2006.03044.x

2. Oberhaus L, Briand JF, Leboulanger C, Jacquet S, Humbert JF. 2007. Comparative effects of the quality and quantity of light and temperature on the growth of *Planktothrix agardhii* and *P. rubescens* 1. J Phycol 43:1191–1199. https://doi.org/10.1111/j.1529-8817.2007.00414.x
3. Lenard T, Poniewozik M. 2022. *Planktothrix agardhii* versus *Planktothrix rubescens*: separation of ecological niches and consequences of cyanobacterial dominance in freshwater. Int J Environ Res Public Health 19:14897. https://doi.org/10.3390/ijerph192214897
4. Walsby AE, Avery A, Schanz F. 1998. The critical pressures of gas vesicles in *Planktorhrix rubescens* in relation tothe depth of winter mixing in Lake Zürich, Switzerland. J Plankton Res 20:1357–1375. https://doi.org/10.1093/plankt/20.7.1357
5. Erratt KJ, Creed IF, Freeman EC, Trick CG, Westrick J, Birbeck JA, Watson LC, Zastepa A. 2022. Deep cyanobacteria layers: an overlooked aspect of managing risks of cyanobacteria. Environ Sci Technol 56:17902–17912. https://doi.org/10.1021/acs.est.2c06928
6. Jaworski GHM, Talling JF, Heaney SI. 1981. The influence of carbon dioxide-depletion on growth and sinking rate of two planktonic diatoms in culture. Brit Phycol J 16:395–410. https://doi.org/10.1080/-00071618100650461
7. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, 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
8. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
9. Andrews S. Babraham bioinformatics - FastQC a quality control tool for high throughput sequence data. Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc. Retrieved 22 Jan 2024.
10. 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
11. 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
12. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165. https://doi.org/10.7717/peerj.1165
13. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146. https://doi.org/10.1038/nmeth.3103
14. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843. https://doi.org/10.1038/s41564-018-0171-1
15. 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
16. 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
17. 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
18. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
19. 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