9-1-2020

Complete genome sequence of microcystis aeruginosa FD4, isolated from a subtropical river in Southwest Florida

Hidetoshi Urakawa
Florida Gulf Coast University

Taylor L. Hancock
Florida Gulf Coast University

Jacob H. Steele
Florida Gulf Coast University

Elizabeth K. Dahedl
Florida Gulf Coast University

Haruka E. Urakawa
Florida Gulf Coast University

See next page for additional authors

NSUWorks Citation

Urakawa, Hidetoshi; Taylor L. Hancock; Jacob H. Steele; Elizabeth K. Dahedl; Haruka E. Urakawa; Luka K. Ndungu; Lauren E. Krausfeldt; Barry H. Rosen; and Jose V. Lopez. 2020. "Complete genome sequence of microcystis aeruginosa FD4, isolated from a subtropical river in Southwest Florida." *Microbiology Resource Announcements* 9, (38). doi:10.1128/MRA.00813-20.

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Microcystis aeruginosa* FD4, Isolated from a Subtropical River in Southwest Florida

Hidetoshi Urakawa,[a] Taylor L. Hancock,[a] Jacob H. Steele,[a] Elizabeth K. Dahedl,[a] Haruka E. Urakawa,[b] Luka K. Ndungu,[a] Lauren E. Krausfeldt,[c] Barry H. Rosen,[a] Jose V. Lopez[c]

[a]Department of Ecology and Environmental Studies, The Water School, Florida Gulf Coast University, Fort Myers, Florida, USA
[b]Department of Marine and Earth Sciences, The Water School, Florida Gulf Coast University, Fort Myers, Florida, USA
[c]Department of Biological Sciences, Nova Southeastern University, Dania Beach, Florida, USA

ABSTRACT We report the first complete genome of *Microcystis aeruginosa* from North America. A harmful bloom that occurred in the Caloosahatchee River in 2018 led to a state of emergency declaration in Florida. Although strain FD4 was isolated from this toxic bloom, the genome did not have a microcystin biosynthetic gene cluster.

Lake Okeechobee is the largest lake in the southeastern United States (1,900 km^2) and serves as a hub for water flow from the north to the Everglades in southern Florida. When flood control releases are necessary, water is directed to the Atlantic and Gulf coasts through two waterways, the St. Lucie River and the Caloosahatchee River, respectively (1). Since the 1980s, Lake Okeechobee and its waterways have suffered from chronic eutrophication problems and harmful cyanobacterial blooms (2). A recent *Microcystis aeruginosa* bloom that occurred in the Caloosahatchee River in 2018 led to a state of emergency declaration in Florida (3).

M. aeruginosa FD4 was isolated from surface water collected from the Caloosahatchee River at the Fort Denaud Bridge (26.7444N, 81.5103W) on 27 June 2018 during the bloom (3). The water quality parameters determined were as follows: water temperature, 32.4°C; pH 7.7; dissolved oxygen, 3.0 mg/liter; total iron, 1.32 mg/liter; chlorophyll *a*, 72.5 μg/liter; microcystin, 450.5 μg/liter; and hydrogen peroxide, 9.2 μg/liter (3). Surface scum was originally incubated with ultrapure water with pyruvic acid (8.8 μg/liter) at room temperature and then transferred to pyruvic acid-amended 10% BG-11 with germanium dioxide (10 mg/liter) and cycloheximide (100 mg/liter) to inhibit the growth of diatoms and other eukaryotes, respectively. Once we microscopically confirmed that the culture was unialgal, strain FD4 was maintained in BG-11 medium at 25°C under fluorescent light with a 12:12-h light/dark cycle. Genomic DNA was extracted using the Quick-DNA miniprep plus kit (Zymo Research). The DNA was made into SMRTbell libraries using the Express Template prep kit 2.0 (Pacific Biosciences). The sample was multiplexed with other samples into a single library and size selected using BluePippin (Sage Sciences) according to the manufacturer's recommendations using the 0.75% DF Marker S1 high-pass 6-kb to 10-kb v3 run protocol and S1 marker (BPstart value of 8,000). The size-selected SMRTbell library was annealed and bound according to the SMRT Link setup and sequenced with the PacBio Sequel II system (Pacific Biosciences). Raw PacBio reads were converted to FASTA format with samtools fasta and then assembled with Flye version 2.6 (4). Default parameters were used for all software. The assembled genome was annotated with Prokka version 1.11 (5). The genome was annotated with SEED Viewer (6) and the NCBI Prokaryotic Genome Annotation Pipeline (GeneMark S-2+ version 4.10) (7). The final assembly of the genome comprised 5.45 Mbp at 125-fold coverage (N_{50} and N_{90}, 10,352 and 6,453 bp,

Citation Urakawa H, Hancock TL, Steele JH, Dahedl EK, Urakawa HE, Ndungu LK, Krausfeldt LE, Rosen BH, Lopez JV. 2020. Complete genome sequence of *Microcystis aeruginosa* FD4, isolated from a subtropical river in southwest Florida. Microbiol Resour Announc 9:e00813-20. https://doi.org/10.1128/MRA.00813-20.

Editor Frank J. Stewart, Georgia Institute of Technology

Address correspondence to Hidetoshi Urakawa, hurakawa@fgcu.edu.

Received 14 July 2020
Accepted 26 August 2020
Published 17 September 2020

TABLE 1 Genome statistics of *Microcystis aeruginosa* FD4[a]

Attribute[b]	Value
Total genome size (bp)	5,493,112
Chromosome size (bp)	5,449,501
Plasmid size (bp)	43,611
G+C content (%)	42.59
Total no. of genes	5,455
Total no. of CDSs	5,403
No. of coding genes	4,679
No. of CDSs with protein	4,679
No. of RNA genes	52
No. of rRNA sets (5S, 16S, 23S)	2
No. of tRNAs	42
No. of ncRNAs	4
Total no. of pseudogenes	724
No. of CRISPR arrays	3

[a] Annotation is based on the total genome and the NCBI Prokaryotic Genome Annotation Pipeline.
[b] CDS, coding sequences; ncRNAs, noncoding RNAs; CRISPR, clustered regularly interspaced short palindromic repeat.

respectively) and consisted of two completely closed contigs, one chromosome, and one plasmid (Table 1).

Although strain FD4 was isolated from the toxic algal bloom (3), the genome did not have a microcystin biosynthetic gene cluster. It was confirmed by the annotations of the predicted open reading frames and homology searches against the genome. The presence of nine secondary metabolite gene clusters, including piricyclamide (8), micropeptin, and aeruginosin, were identified using antiSMASH version 5.1.2 (9). These numbers were quite small in comparison with those of *M. aeruginosa* NIES-2481 (GenBank accession number CP012375.1), in which 28 secondary metabolite gene clusters were found (10). Haft et al. (11) found that all bacteria with a short C-terminal homology domain that includes a highly conserved motif proline-glutamate-proline triad (PEP-CTERM) have both an outer membrane and exopolysaccharide production genes. Notably, 62 clusters of PEP-CTERM sorting domain-containing protein were found along with genes of exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase, polysaccharide pyruvyl transferase (CsaB), and WecB/TagA/CpsF family glycosyltransferase in the genome, suggesting the possible association of these genes for *Microcystis* colony formation (11). Consistent with the colony-forming ability of strain FD4, we found a coding gene of gas vesicle protein GvpC and *psb* and *apc* photoregulation clusters, which confer an ecological advantage to *M. aeruginosa* FD4 to compete with other phytoplankton through surface scum formation (12). Kardinaal and colleagues (13) reported that nontoxic strains of *Microcystis* were better competitors for light than toxic strains. Further genome annotation and genome comparisons with other strains of *M. aeruginosa* will provide additional insights into the ecological adaptation of this cyanobacterium.

Data availability. The genome sequence information has been deposited under BioProject number PRJNA595771 (GenBank accession numbers CP046973.1 [chromosome] and CP046974.1 [plasmid]). The PacBio reads have been deposited in the SRA under the accession number SRR12188899.

ACKNOWLEDGMENTS

We thank Rick D. Bartleson at Sanibel-Captiva Conservation Foundation for field assistance.

This research was supported by the NSF Division of Environmental Biology grant DEB-1664052 and the Aquatic Nuisance Species Research Program, U.S. Army Corps of Engineers (cooperative agreement number W912HZ-19-2-0014). A part of this work was also supported via funding provided by FGCU via The Water School, the Blair Foundation, the Seidler Fund, and the FGCU Center for Environmental and Sustainability Education.

REFERENCES

1. Khare Y, Naja GM, Stainback GA, Martinez CJ, Paudel R, Van Lent T. 2019. A phased assessment of restoration alternatives to achieve phosphorus water quality targets for Lake Okeechobee, Florida, USA. Water 11:327. https://doi.org/10.3390/w11020327.
2. Kramer BJ, Davis TW, Meyer KA, Rosen BH, Goleski JA, Dick GJ, Oh G, Gobler CJ. 2018. Nitrogen limitation, toxin synthesis potential, and toxicity of cyanobacterial populations in Lake Okeechobee and the St. Lucie River Estuary, Florida, during the 2016 state of emergency event. PLoS One 13:e0196278. https://doi.org/10.1371/journal.pone.0196278.
3. Ndungu LK, Steele JH, Hancock TL, Bartleson RD, Milbrandt EC, Parsons ML, Urakawa H. 2019. Hydrogen peroxide measurements in subtropical aquatic systems and their implications for cyanobacterial blooms. Ecol Eng 138:444–453. https://doi.org/10.1016/j.ecoleng.2019.07.011.
4. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
5. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
6. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang H-Y, Cohoon M, de Crécy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res 33:5691–5702. https://doi.org/10.1093/nar/gki866.
7. 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.
8. Leikoski N, Fewer DP, Jokela J, Alakoski P, Wahlsten M, Sivonen K. 2012. Analysis of an inactive cyanobactin biosynthetic gene cluster leads to discovery of new natural products from strains of the genus *Microcystis*. PLoS One 7:e43002. https://doi.org/10.1371/journal.pone.0043002.
9. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0—updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 47:W81–W87. https://doi.org/10.1093/nar/gkz310.
10. Yamaguchi H, Suzuki S, Osana Y, Kawachi M. 2018. Complete genome sequence of *Microcystis aeruginosa* NIES-2481 and common genomic features of group G *M. aeruginosa*. J Genomics 6:30–33. https://doi.org/10.7150/jgen.24935.
11. Haft DH, Paulsen IT, Ward N, Selengut JD. 2006. Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. BMC Biol : 4:29. https://doi.org/10.1186/1741-7007-4-29.
12. Zhang J-Y, Guan R, Zhang H-J, Li H, Xiao P, Yu G-L, Du L, Cao D-M, Zhu B-C, Li R-H, Lu Z-H. 2016. Complete genome sequence and genomic characterization of *Microcystis panniformis* FACHB 1757 by third-generation sequencing. Stand Genomic Sci 11:11.
13. Kardinaal WEA, Tonk L, Janse I, Hol S, Slot P, Huisman J, Visser PM. 2007. Competition for light between toxic and nontoxic strains of the harmful cyanobacterium *Microcystis*. Appl Environ Microbiol 73:2939–2946. https://doi.org/10.1128/AEM.02892-06.