AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequence of *Microcystis aeruginosa* CACIAM 03, a Cyanobacterium Isolated from an Amazonian Freshwater Environment

Wendel Oliveira Castro,[a] Alex Ranieri Jerônimo Lima,[a] Pablo Henrique Gonçalves Moraes,[a] Andrei Santos Siqueira,[a] Délia Cristina Figueira Aguiar,[a] Anna Rafaella Ferreira Baraúna,[a] Luisa Carício Martins,[b] Hellen Thais Fuzii,[b] Clayton Pereira Silva de Lima,[c] João Lídio Silva Gonçalves Vianez-Júnior,[c] Márcio Roberto Teixeira Nunes,[c] Leonardo Teixeira Dall'Agnol,[d] Evonnildo Costa Gonçalves[a]

Laboratório de Tecnologia Biomolecular, Instituto de Ciências Biológicas (ICB), Universidade Federal do Pará (UFPA), Belém, Pará, Brazil[a]; Laboratório de Patologia Clínica e Doenças Tropicais, Núcleo de Medicina Tropical, Universidade Federal do Pará, Belém, Pará, Brazil[b]; Centro de Inovações Tecnológicas (CIT), Instituto Evandro Chagas (IEC), Ananindeua, Pará, Brazil[c]; Universidade Federal do Maranhão (UFMA), Campus de Bacabal, Maranhão, Brazil[d]

Given its toxigenic potential, *Microcystis aeruginosa* is an important bloom-forming cyanobacterium. Here, we present a draft genome and annotation of the strain CACIAM 03, which was isolated from an Amazonian freshwater environment.

Received 23 September 2016 **Accepted** 27 September 2016 **Published** 17 November 2016

Citation Castro WO, Lima ARJ, Moraes PHG, Siqueira AS, Aguiar DCF, Baraúna ARF, Martins LC, Fuzii HT, de Lima CPS, Vianez-Júnior JLSG, Nunes MRT, Dall'Agnol LT, Gonçalves EC. 2016. Draft genome sequence of *Microcystis aeruginosa* CACIAM 03, a cyanobacterium isolated from an Amazonian freshwater environment. Genome Announc 4(6): e01299-16. doi:10.1128/genomeA.01299-16.

Address correspondence to Evonnildo Costa Gonçalves, ecostag@ufpa.br.

Microcystis aeruginosa is a bloom-forming freshwater cyanobacterium that may have toxic activity and thus has economic and ecological importance worldwide (1). This species is well known for synthesizing powerful monocyclic heptapeptides known as microcystins, which are a group of hepatotoxins that cause several problems to drinking water supplies, aquatic organisms, human health, and the environment (2).

Genomic data available for cyanobacteria from the Amazonian environment are scarce, with only a few genomes sequenced to date. To improve the genomic data of different cyanobacterial strains, we recovered the draft genome of *Microcystis aeruginosa* CACIAM 03 from the total DNA obtained from a nonaxenic culture. This strain was isolated from a water sample from the Tucuruí hydroelectric power station reservoir (3°50′04.9″S, 49°42′32.2″W) in Pará, Brazil.

After DNA extraction of the cyanobacterial culture, two sequencing runs were performed on the GS FLX 454 (Roche Life Sciences) platform using nonpaired libraries, and one sequencing run was carried out on the Illumina MiSeq platform using a paired-end library with a 150-bp read length. All the raw reads were quality-filtered with a minimum Phred score of 20. A coassembly of all reads was performed by three assembly software programs: Newbler version 2.9 (which was parameterized with minimum overlap of 40 bp, minimum overlap identity of 90%, heterozygote mode, and extended low-depth overlap options on), CLC Genomics Workbench (http://www.clcbio.com) (default parameters), and SPAdes version 3.9 (3) (with parameter flag for metagenome and *k*-mer sizes of 21, 33, 55, 77, 99, and 127). These assemblers produced 1,484, 14,250, and 5,742 scaffolds, respectively.

MaxBin version 2.2.1 (4) was used to bin each set of assembled scaffolds. To classify taxonomically the obtained bins, we performed a BLASTp analysis for each bin in the sequences containing hidden Markov models for essential genes identified by MaxBin version 2.2.1 against the NCBI nonredundant database. The results were visualized on MEGAN 5 (5). The bins identified as *M. aeruginosa* were integrated and subsequently processed using the hybrid assembly program CISA version 1.3.1 (6), producing 249 scaffolds (>1.852 bp length) with an N_{50} value of 33,157, a total of 5.0 Mb, and a GC content of 42.9%.

The scaffolds were annotated by the NCBI Prokaryotic Genome Annotation Pipeline (7). This process identified 4,197 coding sequences, 44 tRNA genes (including 1 tmRNA), 8 rRNA genes, 614 pseudogenes, and 4 noncoding RNA genes.

Preliminary analysis with antiSMASH version 3.0 (8) revealed gene clusters involved in the nonribosomal biosynthesis of microcystin, aeruginosin, and terpene, as well as gene clusters involved in ribosome synthesis of peptides such as yersiniabactin (bacteriocin), micropeptin (microviridin), microcyclamide (cyanobactin), and microcyclamide (bacteriocin), which are of great biotechnological value.

This report can improve the genomic data about *M. aeruginosa* by including the first genome of this species obtained from an Amazonian environment.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number MCIH00000000. The version described in this paper is the first version, MCIH01000000.

ACKNOWLEDGMENTS

We are grateful to Centrais Elétricas do Norte do Brasil S/A–Eletronorte for logistical support in collecting samples.

FUNDING INFORMATION

This work, including the efforts of Evonnildo Costa Gonçalves, was funded by MCTI | Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (554321/2010-6 and 140218/2016-5). This work, including the efforts of Evonnildo Costa Gonçalves, was funded by Fundação Amazônia Paraense de Amparo à Pesquisa (FAPESPA) (ICAAF 099/2014).

Financial support was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; process numbers 554321/2010-6 and 140218/2016-5) and Fundação Amazônia de Amparo a Estudos e Pesquisa (FAPESPA; ICAAF 099/2014).

REFERENCES

1. **Straub C, Quillardet P, Vergalli J, de Marsac NT, Humbert J.** 2011. A day in the life of *Microcystis aeruginosa* strain PCC 7806 as revealed by a transcriptomic analysis. PLoS One **6:**e16208. http://dx.doi.org/10.1371/journal.pone.0016208.
2. **Boopathi T, Ki J-S.** 2014. Impact of environmental factors on the regulation of cyanotoxin production. Toxins (Basel) **6:**1951–1978. http://dx.doi.org/10.3390/toxins6071951.
3. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19:**455–477. http://dx.doi.org/10.1089/cmb.2012.0021.
4. **Wu YW, Simmons BA, Singer SW.** 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics **32:**605–607. http://dx.doi.org/10.1093/bioinformatics/btv638.
5. **Huson DH, Auch AF, Qi J, Schuster SC.** 2007. MEGAN analysis of metagenomic data. Genome Res **17:**377–386. http://dx.doi.org/10.1101/gr.5969107.
6. **Lin SH, Liao YC.** 2013. CISA: contig integrator for sequence assembly of bacterial genomes. PLoS One **8:**e60843. http://dx.doi.org/10.1371/journal.pone.0060843.
7. **Tatusova T, Ciufo S, Federhen S, Fedorov B, McVeigh R, O'Neill K, Tolstoy I, Zaslavsky L.** 2015. Update on RefSeq microbial genomes resources. Nucleic Acids Res **43:**D599–D605. http://dx.doi.org/10.1093/nar/gku1062.
8. **Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH.** 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res **43:**W237–W243. http://dx.doi.org/10.1093/nar/gkv437.