

Draft Genome Sequence of the Cyanobacterium *Aphanizomenon flos-aquae* Strain 2012/KM1/D3, Isolated from the Curonian Lagoon (Baltic Sea)

Sigitas Šulčius,^a Gediminas Alzbutas,^b Kotryna Kvederavičiūtė,^c Judita Koreivienė,^d Linas Zakrys,^b Arvydas Lubys,^b Ričardas Paškauskas^{a,d}

Marine Science and Technology Center, Klaipėda University, Klaipėda, Lithuania^a; Thermo Fisher Scientific Baltics, Vilnius, Lithuania^b; Institute of Biotechnology, Vilnius University, Vilnius, Lithuania^c; Institute of Botany, Nature Research Centre, Vilnius, Lithuania^d

We report here the *de novo* genome assembly of a cyanobacterium, *Aphanizomenon flos-aquae* strain 2012/KM1/D3, a harmful bloom-forming species in temperate aquatic ecosystems. The genome is 5.7 Mb with a G+C content of 38.2%, and it is enriched mostly with genes involved in amino acid and carbohydrate metabolism.

Received 24 November 2014 Accepted 9 December 2014 Published 15 January 2015

Citation Šulčius S, Alzbutas G, Kvederavičiūtė K, Koreivienė J, Zakrys L, Lubys A, Paškauskas R. 2015. Draft genome sequence of the cyanobacterium *Aphanizomenon flos-aquae* strain 2012/KM1/D3, isolated from the Curonian Lagoon (Baltic Sea). *Genome Announc* 3(1):e01392-14. doi:10.1128/genomeA.01392-14.

Copyright © 2015 Šulčius et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Sigitas Šulčius, sigitas.sulcius@gmail.com.

The filamentous and nitrogen-fixing cyanobacterium *Aphanizomenon flos-aquae* is globally distributed in temperate lakes and brackish water ecosystems, among other locations in the Curonian Lagoon of the Baltic Sea (1). Blooms of *A. flos-aquae* affect water quality and the recreational value of the ecosystem, and they endanger aquatic biota, animal health, and human health (2–4). There is presently only one draft genome sequence available for this species (5); yet, little is known about the genetic capabilities of this cyanobacterium.

The clonal *Aphanizomenon flos-aquae* strain 2012/KM1/D3 was isolated from the Curonian Lagoon during the bloom in 2012 and is maintained as a unicyanobacterial yet nonaxenic culture. DNA was extracted using a cetyltrimethylammonium bromide (CTAB) protocol (6), with some modifications. Cell lysis was facilitated by five freeze (in liquid nitrogen)/thaw (at 65°C) cycles following treatment with lysozyme (50 mg/ml), RNase (100 mg/ml), proteinase K (20 mg/ml), and SDS (10%). The draft genome sequence was performed by Thermo Fisher Scientific Baltics (<http://www.thermofisher.lt>) using the Personal Genome Machine (PGM) with the Ion PGM sequencing 400 kit and Thermo Scientific MuSeek library preparation kit (catalog no. 4480829). A total of 651 Mb of DNA sequence and 2,531,067 reads were generated. To filter out cyanobacterial reads and remove sequences belonging to associated bacteria (7), phylogenetic classification of all reads was done using the Kraken (8) and PhymmBL (9) programs. Genome assembly was carried out by SeqMan (DNASTar LaserGene version 11.1.0; Madison, WI, USA), SPAdes version 3.1.1 (10), and MIRA version 4.0.2 (11), and the final genome assembly was created by merging the results using CISA (12). Genome annotation was carried out using RAST (13) and the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). The genome was validated with QCAST (14).

The *A. flos-aquae* draft genome is 5,741,771 bp distributed in 325 contigs, with an average G+C content of 38.2% and an N_{50} value of 25,535. The largest contig is 120,451 bp. PGAP identified

5,478 protein-coding genes, of which 4,415 had a predicted function and 1,010 are pseudogenes. The genome contains 52 RNA genes, of which 38 are tRNA and 14 are rRNA genes. The calculation of the average nucleotide identity (ANI) based on the BLAST algorithm showed that the closest sequences match to *A. flos-aquae* strain NIES-81 (5) at 96.24% similarity. These results corresponded well to the tetranucleotide signature frequency correlation coefficient (0.997), which indicates the taxonomic relatedness of two strains (15). The genome contains fragments of cyanotoxin gene clusters similar to those observed for other *Aphanizomenon* strains (16) and possesses a high number of clustered regularly interspaced short palindromic repeats (17). The genome also contains genes involved in the metabolism of glycine betaine, which has been shown to be a very efficient osmolyte under conditions of saline stress, which occurs often in transitional aquatic ecosystems. This study will further facilitate our understanding of the genomic capabilities of *A. flos-aquae* to interact with the surrounding environment.

Nucleotide sequence and accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [JSDP00000000](https://www.ncbi.nlm.nih.gov/nuccore/JSDP00000000). The version described in this paper is version [JSDP01000000](https://www.ncbi.nlm.nih.gov/nuccore/JSDP01000000).

ACKNOWLEDGMENTS

We thank Ramūnas Stepanauskas (Single Cell Genomic Center, Bigelow Laboratory for Ocean Sciences) for help with tetramer analysis and useful comments on the manuscript.

This research was funded by a grant (no. MIP-036/2012) from the Research Council of Lithuania.

REFERENCES

1. Jaanus A, Andersson A, Olenina I, Toming K, Kaljurand K. 2011. Changes in phytoplankton communities along a north-south gradient in the Baltic Sea between 1990 and 2008. *Boreal Environ Res* 16:191–208.
2. Falconer IR. 1999. An overview of problems caused by toxic blue-green algae (cyanobacteria) in drinking and recreational water. *Environ Toxicol* 14:5–12.

3. Preussel K, Stüken A, Wiedner C, Chorus I, Fastner J. 2006. First report on cylindrospermopsin producing *Aphanizomenon flos-aquae* (cyanobacteria) isolated from two German lakes. *Toxicon* 47:156–162. <http://dx.doi.org/10.1016/j.toxicon.2005.10.013>.
4. Wu X, Jiang J, Wan Y, Giesy JP, Hu J. 2012. Cyanobacteria blooms produce teratogenic retinoic acids. *Proc Natl Acad Sci U. S. A.* 109: 9477–9482. <http://dx.doi.org/10.1073/pnas.1200062109>.
5. Cao H, Shimura Y, Masanobu K, Yin Y. 2014. Draft genome sequence of the toxic bloom-forming cyanobacterium *Aphanizomenon flos-aquae* NIES-81. *Genome Announc* 2(1):e00044-14. <http://dx.doi.org/10.1128/genomeA.00044-14>.
6. Wilson K. 2001. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* Chapter 2:Unit 2.4. <http://dx.doi.org/10.1002/0471142727.mb0204s56>.
7. Paul R, Jinkerson RE, Buss K, Steel J, Mohr R, Hess WR, Chen M, Fromme P. 2014. Draft genome sequence of the filamentous cyanobacterium *Leptolyngbya* sp. strain Heron Island J, exhibiting chromatic acclimation. *Genome Announc* 2(1):e01166-13. <http://dx.doi.org/10.1128/genomeA.01166-13>.
8. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <http://dx.doi.org/10.1186/gb-2014-15-3-r46>.
9. Brady A, Salzberg S. 2011. PhymmBL expanded: confidence scores, custom databases, parallelization and more. *Nat Methods* 8:367. <http://dx.doi.org/10.1038/nmeth0511-367>.
10. 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>.
11. Chevreaux B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, Suhai S. 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 14:1147–1159. <http://dx.doi.org/10.1101/gr.1917404>.
12. 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>.
13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <http://dx.doi.org/10.1186/1471-2164-9-75>.
14. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <http://dx.doi.org/10.1093/bioinformatics/btt086>.
15. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U. S. A.* 106: 19126–19131. <http://dx.doi.org/10.1073/pnas.0906412106>.
16. Stüken A, Jakobsen KS. 2010. The cylindrospermopsin gene cluster of *Aphanizomenon* sp. strain 10e6: organization and recombination. *Microbiology* 156:2438–2451. <http://dx.doi.org/10.1099/mic.0.036988-0>.
17. Sorek R, Kunin V, Hugenholtz P. 2008. CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* 6:181–186. <http://dx.doi.org/10.1038/nrmicro1793>.