

Draft Genome Sequence of *Microcystis aeruginosa* NIES-98, a Non-Microcystin-Producing Cyanobacterium from Lake Kasumigaura, Japan

Haruyo Yamaguchi,^a Shigekatsu Suzuki,^a Tomoharu Sano,^b Yuuhiko Tanabe,^c Nobuyoshi Nakajima,^a Masanobu Kawachi^a

Center for Environmental Biology and Ecosystem Studies, National Institute for Environmental Studies, Ibaraki, Japan^a; Center for Environmental Measurement and Analysis, National Institute for Environmental Studies, Ibaraki, Japan^b; Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan^c

***Microcystis aeruginosa* is a well-known bloom-forming cyanobacterium. We newly sequenced the whole genome of *M. aeruginosa* NIES-98, which is a non-microcystin-producing strain isolated from Lake Kasumigaura, Japan. The genome contains approximately 5.0 Mbp, with an average G+C content of 42.41% and 5,140 predicted protein-coding genes.**

Received 13 September 2016 Accepted 21 September 2016 Published 10 November 2016

Citation Yamaguchi H, Suzuki S, Sano T, Tanabe Y, Nakajima N, Kawachi M. 2016. Draft genome sequence of *Microcystis aeruginosa* NIES-98, a non-microcystin-producing cyanobacterium from Lake Kasumigaura, Japan. *Genome Announc* 4(6):e01187-16. doi:10.1128/genomeA.01187-16.

Copyright © 2016 Yamaguchi 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 Haruyo Yamaguchi, yamaguchi.haruyo@nies.go.jp.

Microcystis aeruginosa is a well-known bloom-forming cyanobacterium in freshwater lakes worldwide (1). The species is genetically divided into at least eight clades (groups A to G and X) based on a multilocus phylogenetic analysis (2). Group A, X, and some group B strains produce hepatotoxic cyanotoxins called microcystins. *M. aeruginosa* NIES-98 is a non-microcystin-producing strain isolated from Lake Kasumigaura, Japan (2). The strain is available via the Microbial Culture Collection at the National Institute for Environmental Studies in Japan (<http://mcc.nies.go.jp/>). According to the previous phylogenetic analysis by Tanabe and Watanabe (2), NIES-98 is a group B strain. Thus far, whole-genome sequencing of only three strains of *M. aeruginosa* from Lake Kasumigaura have been reported: NIES-843 (group A) (3), NIES-44 (group E) (4), and NIES-2549 (group G) (5). Additional genetic information on *M. aeruginosa* is required to determine the spatial and seasonal dynamics in Lake Kasumigaura and the evolutionary history of this species.

Our analysis of volatile compounds using gas chromatography-mass spectrometry (GC-MS) showed that NIES-98 produces 3,5-dimethylanisole, which is known as a kind of fungal volatile compound (6); however, the other major volatile compounds from this strain remain unidentified.

Genomic DNA was extracted from 50 ml of the axenic NIES-98 culture using a DNeasy plant minikit (Qiagen) and fragmented to approximately 550-bp segments using a Covaris M220 ultrasonicator (Covaris). The genomic library was constructed using a TruSeq Nano DNA library prep kit for NeoPrep (Illumina) and sequenced by the MiSeq platform (Illumina) using the 600-cycle MiSeq reagent kit version 3. The number of resultant paired-end reads was 5,681,573. Low-quality reads/bases were filtered using Trimmomatic version 0.36 (7), and *de novo* assembly was performed using SPAdes 3.9.0, with k-mer of 21, 33, 55, 77, 99, and 127 (8). The resulting draft genome comprises 497 scaffolds (>200 bp in size) of 4,983,728 bp, with an average genome coverage of approximately 322.1×. The maximum scaffold length was 199,403 bp, and the mean size of the scaffolds was 10,027 bp. The

draft genome of NIES-98 was annotated using the RAST server (9). The genome comprises 5,140 predicted protein-coding sequences (CDSs) including 2,236 hypothetical proteins, and 50 RNA genes. The number of CDSs of NIES-98 is larger than those of NIES-44 (4,790 CDSs) and NIES-2549 (4,282 CDSs) but smaller than that of NIES-843 (6,312 CDSs). The G+C content of the genome is 42.41%. Functional annotation based on COG categories using COGNIZER (10) showed that NIES-98 has a smaller number of genes for transposase and inactivated derivatives, including those in category L, than the three aforementioned strains. A microcystin biosynthesis gene cluster was not detected in the genome of NIES-98, as shown in a previous study (11). antiSMASH (12) predicted 11 secondary metabolite gene clusters in the genome, including aeruginosin, puwainaphycins, and micropeptin biosynthetic gene clusters. This genomic sequence provides basic information for better understanding of the ecology and evolution of *M. aeruginosa*.

Accession number(s). This whole-genome shotgun project has been deposited in GenBank under the accession no. [MDZH00000000](https://www.ncbi.nlm.nih.gov/nuclink/MDZH00000000). The version described in this paper is the first version.

REFERENCES

1. Carmichael WW. 1996. Toxic *Microcystis* and the environment, p 1–11. In Watanabe MF, Harada K, Carmichael WW, Fujiki H (ed), *Toxic Microcystis*. CRC Press, Boca Raton, FL.
2. Tanabe Y, Watanabe MM. 2011. Local expansion of a panmictic lineage of water bloom-forming cyanobacterium *Microcystis aeruginosa*. *PLoS One* 6:e17085. <http://dx.doi.org/10.1371/journal.pone.0017085>.
3. Kaneko T, Nakajima N, Okamoto S, Suzuki I, Tanabe Y, Tamaoki M, Nakamura Y, Kasai F, Watanabe A, Kawashima K, Kishida Y, Ono A, Shimizu Y, Takahashi C, Minami C, Fujishiro T, Kohara M, Katoh M, Nakazaki N, Nakayama S, Yamada M, Tabata S, Watanabe MM. 2007. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Res* 14:247–256. <http://dx.doi.org/10.1093/dnares/dsm026>.
4. Okano K, Miyata N, Ozaki Y. 2015. Genome sequence of *Microcystis aeruginosa* strain NIES-44. *Genome Announc* 3(2):e00135-15. <http://dx.doi.org/10.1128/genomeA.00135-15>.

5. Yamaguchi H, Suzuki S, Tanabe Y, Osana Y, Shimura Y, Ishida K, Kawachi M. 2015. Complete genome sequence of *Microcystis aeruginosa* NIES-2549, a bloom-forming cyanobacterium from lake Kasumigaura, Japan. *Genome Announc* 3(3):e00551-15. <http://dx.doi.org/10.1128/genomeA.00551-15>.
6. Müller A, Faubert P, Hagen M, zu Castell W, Polle A, Schnitzler JP, Rosenkranz M. 2013. Volatile profiles of fungi—chemotyping of species and ecological functions. *Fungal Genet Biol* 54:25–33. <http://dx.doi.org/10.1016/j.fgb.2013.02.005>.
7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120 <http://dx.doi.org/10.1093/bioinformatics/btu170>.
8. 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>.
9. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano 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>.
10. Bose T, Haque MM, Reddy C, Mande SS. 2015. COGNIZER: a framework for functional annotation of metagenomic datasets. *PLoS One* 10:e0142102. <http://dx.doi.org/10.1371/journal.pone.0142102>.
11. Tanabe Y, Sano T, Kasai F, Watanabe MM. 2009. Recombination, cryptic clades and neutral molecular divergence of the microcystin synthetase (*mcy*) genes of toxic cyanobacterium *Microcystis aeruginosa*. *BMC Evol Biol* 9:115. <http://dx.doi.org/10.1186/1471-2148-9-115>.
12. 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>.