

# Whole-Genome Sequence of *Microcystis aeruginosa* TAIHU98, a Nontoxic Bloom-Forming Strain Isolated from Taihu Lake, China

Chen Yang,<sup>a</sup> Wei Zhang,<sup>b</sup> Minglei Ren,<sup>a</sup> Lirong Song,<sup>a</sup> Tao Li,<sup>a</sup> Jindong Zhao<sup>a,b</sup>

Key Laboratory of Algal Biology, Institute of Hydrobiology, Chinese Academy of Science, Wuhan, China<sup>a</sup>; College of Life Science, Peking University, Beijing, China<sup>b</sup>

***Microcystis aeruginosa* is a dominant bloom-forming cyanobacterium in many freshwater lakes. This report describes the first whole-genome sequence of the nontoxic strain of *M. aeruginosa* TAIHU98, which was isolated from Taihu Lake in eastern China.**

Received 25 April 2013 Accepted 1 May 2013 Published 13 June 2013

**Citation** Yang C, Zhang W, Ren M, Song L, Li T, Zhao J. 2013. Whole-genome sequence of *Microcystis aeruginosa* TAIHU98, a nontoxic bloom-forming strain isolated from Taihu Lake, China. *Genome Announc.* 1(3):e00333-13. doi:10.1128/genomeA.00333-13.

**Copyright** © 2013 Yang 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/3.0/).

Address correspondence to Tao Li, [litao@ihb.ac.cn](mailto:litao@ihb.ac.cn), or Jindong Zhao, [jzhao@ihb.ac.cn](mailto:jzhao@ihb.ac.cn).

*Microcystis aeruginosa* is one of the most ecologically harmful and dominant bloom-forming cyanobacteria in freshwater lakes (1). The production of microcystin, a cyclic heptapeptide made by *M. aeruginosa* and some other toxic cyanobacteria, is now a major concern for water safety (2). Also, the accumulation of *M. aeruginosa* mass on the surface water has deteriorative effects on freshwater ecosystems, such as blocking light for other photosynthetic organisms and causing hypoxia in the water. In many freshwater lakes, such as Taihu Lake in eastern China, *Microcystis* water bloom becomes dominant in the summer and lasts until early winter (3, 4). Despite the concentration of microcystin being highest in the summer in Taihu Lake (5), toxic and nontoxic cells of *M. aeruginosa* coexist during the entire period of the water bloom (6). To understand the mechanism of water-bloom formation and its ecological effects, we isolated both toxic and nontoxic *M. aeruginosa* strains from Taihu Lake water bloom. Here, we report the whole-genome sequence of the nontoxic *M. aeruginosa* strain TAIHU98.

Whole-genome sequencing of TAIHU98 was performed with a combination of Genome Sequencer FLX (Roche) (400-bp single-end library and 3-kb paired-end library, 1,070,319 reads) and Genome Analyzer IIx (Illumina) (3-kb mate-pair library, 6,867,274 reads). All 454 reads were assembled into 395 contigs by Newbler v2.0.01.14. Paired-end and Solexa mate-pair reads were used to order these contigs into 50 supercontigs within 6 scaffolds. Gaps were then filled by sequencing PCR products with the help of Phred-Phrap-Consed v23.0. The prediction of protein-encoding sequences (CDSs) was generated by Glimmer 3.0 (7) and GeneMark 2.5 (8). tRNAs and rRNAs were detected using tRNAscan-SE 1.23 (9) and RNAmmer 1.2 (10), respectively. Functional annotation of the genome was performed by searching against NCBI nonredundant protein, InterPro, and COG databases (11).

The genome of TAIHU98 comprises 4 supercontigs with a total length of 4,849,611 bp and an average G + C content of 42.45%. It contains 5,356 protein-coding genes and 48 genes coding for RNA (two sets of rRNA genes and 42 tRNA genes). According to annotation results, 2,660 putative genes show similarity to the genes

with known functions, and the remaining 2,696 genes were determined as encoding hypothetical proteins or were assigned putative functions. A total of 349 copies of insertion sequence (IS) transposase genes are assigned to 19 families, confirming that the genome was as highly plastic as *M. aeruginosa* strains NIES843 (12) and PCC7806 (13).

Comparative analysis revealed that the nucleic acid base sequence of TAIHU98 bears a similarity to those of *M. aeruginosa* PCC7806 (76.5%) and *M. aeruginosa* NIES843 (64.58%). The three genomes only share 2,511 CDSs involved in cell structure components and primary metabolism processes, while TAIHU98 has 1,559 strain-specific genes, mainly with unknown functions. This large variation in the genomes indicates that each strain has gained a considerable number of genes during evolution.

The TAIHU98 genome is missing all microcystin synthetase (*mcy*) genes (14) and cyanopeptolin synthetase (*mcp*) genes (15), while the whole aeruginosin synthetase (*aer*) gene cluster (16) is present.

**Nucleotide sequence accession numbers.** This Whole-Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [ANKQ000000000](https://www.ncbi.nlm.nih.gov/nuccore/ANKQ000000000). The version described in this paper is the first version, accession no. [ANKQ010000000](https://www.ncbi.nlm.nih.gov/nuccore/ANKQ010000000).

## ACKNOWLEDGMENTS

This work was supported by the National High-Technology Research and Development Program of China (863 Program, 2009AA02Z30) and the National Basic Research Program of China (973 Program, 2008CB418001-1).

We thank the Freshwater Algae Collection of the Institution of Hydrobiology, Chinese Academy of Sciences (FACHB collection), for providing the strain *M. aeruginosa* TAIHU98 (FACHB-1752).

## REFERENCES

1. Wilson AE, Sarnelle O, Neilan BA, Salmon TP, Gehring MM, Hay ME. 2005. Genetic variation of the bloom-forming cyanobacterium *Microcystis aeruginosa* within and among lakes: implications for harmful algal blooms. *Appl. Environ. Microbiol.* 71:6126–6133.
2. Gilroy DJ, Kauffman KW, Hall RA, Huang X, Chu FS. 2000. Assessing

- potential health risks from microcystin toxins in blue-green algae dietary supplements. *Environ. Health Perspect.* 108:435–439.
3. Chen YW, Qin BQ, Teubner K, Dokulil MT. 2003. Long-term dynamics of phytoplankton assemblages: *Microcystis*-domination in Lake Taihu, a large shallow lake in China. *J. Plankton Res.* 25:445–453.
  4. Tan X, Kong F, Zeng Q, Cao H, Qian S, Zhang M. 2009. Seasonal variation of *Microcystis* in lake Taihu and its relationships with environmental factors. *J. Environ. Sci. (China)* 21:892–899.
  5. Xu Q, Chen W, Gao G. 2008. Seasonal variations in microcystin concentrations in Lake Taihu, China. *Environ. Monit. Assess.* 145:75–79.
  6. Li D, Kong F, Shi X, Ye L, Yu Y, Yang Z. 2012. Quantification of microcystin-producing and non-microcystin producing *Microcystis* populations during the 2009 and 2010 blooms in Lake Taihu using quantitative real-time PCR. *J. Environ. Sci. (China)* 24:284–290.
  7. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679.
  8. Lukashin AV, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26:1107–1115.
  9. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964.
  10. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
  11. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28:33–36.
  12. 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.
  13. Frangeul L, Quillardet P, Castets AM, Humbert JF, Matthijs HCP, Cortez D, Tolonen A, Zhang CC, Gribaldo S, Kehr JC, Zilliges Y, Ziemert N, Becker S, Talla E, Latifi A, Billault A, Lepelletier A, Dittmann E, Bouchier C, de Marsac NT. 2008. Highly plastic genome of *Microcystis aeruginosa* PCC 7806, a ubiquitous toxic freshwater cyanobacterium. *BMC Genomics* 9:274.
  14. Tillett D, Dittmann E, Erhard M, von Döhren H, Börner T, Neilan BA. 2000. Structural organization of microcystin biosynthesis in *Microcystis aeruginosa* PCC7806: an integrated peptide-polyketide synthetase system. *Chem. Biol.* 7:753–764.
  15. Nishizawa T, Ueda A, Nakano T, Nishizawa A, Miura T, Asayama M, Fujii K, Harada K, Shirai M. 2011. Characterization of the locus of genes encoding enzymes producing heptadepsipeptide micropeptin in the unicellular cyanobacterium *Microcystis*. *J. Biochem.* 149:475–485.
  16. Ishida K, Welker M, Christiansen G, Cadel-Six S, Bouchier C, Dittmann E, Hertweck C, de Marsac NT. 2009. Plasticity and evolution of aeruginosin biosynthesis in cyanobacteria. *Appl. Environ. Microbiol.* 75: 2017–2026.