AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Draft genome sequence of the anatoxin-a producing cyanobacterium *Tychonema bourrellyi* B0820 isolated from the epilimnion of the deep Alpine Lake Garda

Nico Salmaso,[1,2] Adriano Boscaini,[1] Leonardo Cerasino,[1] Massimo Pindo,[1] Federica Pinto,[3] Nicola Segata,[3] Claudio Donati[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report the draft genome sequence of strain B0820 of the cyanobacterium *Tychonema bourrellyi* isolated from the epilimnion of Lake Garda and assembled from a metagenome of a non-axenic culture. The strain analyzed was shown to produce anatoxin-a, a potent neurotoxin that can cause fatal intoxication in exposed organisms.

KEYWORDS cyanobacteria, *Tychonema*, draft genome, full shotgun, anatoxin-a, lake, non-axenic culture

Following the discovery (1) of anatoxin-a (ATX)-producing strains of *Tychonema bourrellyi* (2, 3) in the epilimnion of the southern Alpine lakes, other ATX-producing *Tychonema* species in freshwater mats were described in Germany, where they caused dog deaths (4, 5).

Tychonema strain B0820 was collected from the epilimnion of Lake Garda (45.69 N, 10.72 E) in August 2020. Under the microscope, a single filament was isolated, washed repeatedly with sterilized Z8 medium, and cultured in the same medium at 20°C with 25 µmol m^{-2} s^{-1} illumination (6). The 16S rRNA (OR088434) and ATX (OR124837) sequences [Sanger sequencing (6)] showed 100% identity with previously reported *T. bourrellyi* genes. Using standard LC-MS methods (7), strain B0820 was shown to produce anatoxin-a.

For metagenomic analysis, to reduce the bacteria in the non-axenic culture, filaments were diluted in 1 L of sterilized Z8 medium, filtered through sterilized 80 µm mesh, resuspended in 100 mL Z8 medium, and filtered through 5 µm cellulose nitrate membranes (Whatman). DNA was extracted using PowerWater Kit (Qiagen). The library was prepared using Kapa HyperPlus Kit (Roche) and processed by 150-bp paired-end sequencing on an Illumina NovaSeq 6000, which generated 98,364,812 paired-end reads. Unless otherwise stated, the software was run with default parameters. Reads were checked with FastQC 0.11.9 (8). Removal of residual adapters and trimming (trimq = 20 maq = 20 maxns = 0 minlen = 35, ftm = 5, tpe, tbo) were performed using bbduk 38.95 (https://jgi.doe.gov). After correction with metaSPAdes 3.15.3 (9) (--only-error-correction), the paired reads were assembled with megahit 1.2.9 (10) (--presets meta-sensitive), yielding 6,411 contigs larger than 1,000 bp, with a total size of 160 Mbp and N50 of 201,376 bp. The draft genomes were binned using CONCOCT 1.1.0 (11), MetaBAT 2 2.12.1 (12), and MaxBin 2 2.2.7 (13), and results combined using DAS Tool 1.1.6 (14). The *T. bourrellyi* bin was checked and confirmed with anvi'o 7.1 (15). The genome was 5.37 Mbp assembled into 223 contigs, with N50 39,298 bp, GC content 44.6%, and coverage 336×. Using the phylum Cyanobacteria and the order Oscillatoriales as marker lineages, CheckM 1.2.2 (16) estimated a completeness of 98.9% and 99.3%, respectively, with almost no evidence of contamination (0.07% and 0.08%) and no strain heterogeneity. The taxonomy of the draft genome was confirmed by GTDB-Tk 2.1.1 (17) and a

Editor Elinne Becket, California State University, San Marcos, California, USA

Address correspondence to Nico Salmaso, nico.salmaso@fmach.it.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 7 September 2023
Accepted 18 September 2023
Published 19 October 2023

phylogenomic analysis using PhyloPhlAn 3.0.3 (18), which indicated coincidence (ANI = 99.6) with a strain previously isolated from Lake Garda (19).

The genome annotation was completed by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP 6.5) (20), which identified 5,125 coding sequences and 52 tRNAs. After blastn (nr/nt) analysis, the 16S rRNA and *rbcX* genes from the draft genome were 99.9%–100% and 100% identical, respectively, to previously reported *T. bourrellyi* genes. A polyketide synthase protein gene cluster associated with ATX biosynthesis was identified by blastn (99%-100% query coverage) with 93.1% and 93.4% identity to the anatoxin-a synthetase gene clusters of *Anabaena* sp. 37 (JF803645) and *Cuspidothrix issatschenkoi* (LT984882), respectively; the same sequence showed 99.8%–100% identity to *anaF* sequences previously determined in *T. bourrellyi* isolates (query coverage 3%–7%).

ACKNOWLEDGMENTS

The authors acknowledge the financial support of NBFC to Fondazione Edmund Mach (C. Donati, N. Salmaso), funded by the Italian Ministry of University and Research, PNRR, Missione 4 Componente 2, "Dalla ricerca all'impresa", Investimento 1.4, Project CN00000033, and of HORIZON-MSCA-2021-SE-01, AlgaeNet4AV, Project ID 101086437, to Fondazione Edmund Mach (L. Cerasino, N. Salmaso).

AUTHOR AFFILIATIONS

[1]Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige, Italy
[2]NBFC, National Biodiversity Future Center, Palermo, Italy
[3]Department of Cellular, Computational and Integrative Biology, University of Trento, Trento, Italy

AUTHOR ORCIDs

Nico Salmaso http://orcid.org/0000-0001-6057-3642
Adriano Boscaini http://orcid.org/0000-0001-8696-2056
Leonardo Cerasino http://orcid.org/0000-0002-0839-9832
Massimo Pindo http://orcid.org/0000-0002-8516-9010
Federica Pinto http://orcid.org/0000-0002-4539-4811
Nicola Segata http://orcid.org/0000-0002-1583-5794
Claudio Donati http://orcid.org/0000-0001-8688-1651

FUNDING

Funder	Grant(s)	Author(s)
Italian Ministry of University and Research - NBFC	PNRR Missione 4 Componente 2 Dalla ricerca all'impresa Investimento 1.4 Project CN00000033	Nico Salmaso Claudio Donati
EC | Horizon Europe | Excellent Science | HORIZON EUROPE Marie Sklodowska-Curie Actions (MSCA)	HORIZON MSCA 2021 SE 01 AlgaeNet4AV Project ID 101086437	Nico Salmaso Leonardo Cerasino

AUTHOR CONTRIBUTIONS

Nico Salmaso, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review and editing | Adriano Boscaini, Conceptualization, Data curation, Investigation, Methodology, Resources, Validation, Writing – review and editing | Leonardo Cerasino, Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Writing – review and editing | Massimo Pindo, Conceptualization, Data curation, Investigation, Methodology, Validation, Writing – review and editing | Federica Pinto, Conceptualization, Investigation, Methodology, Writing – review and

editing | Nicola Segata, Conceptualization, Investigation, Methodology, Writing – review and editing | Claudio Donati, Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Software, Validation, Writing – review and editing

DATA AVAILABILITY

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAVCBQ000000000. The version described in this paper is version JAVCBQ000000000.1. The sequences are publicly available under BioProject accession number PRJNA979152, BioSample accession number SAMN35566361, and SRA accession number SRR25705839. Sanger sequences of 16S rRNA and anaF genes of strain B0820 were deposited to GenBank with accession numbers OR088434 and OR124837, respectively.

REFERENCES

1. Shams S, Capelli C, Cerasino L, Ballot A, Dietrich DR, Sivonen K, Salmaso N. 2015. Anatoxin-a producing *Tychonema* (Cyanobacteria) in European waterbodies. Water Res 69:68–79. https://doi.org/10.1016/j.watres.2014.11.006
2. Komárek J, Anagnostidis K. 2005. Cyanoprokaryota. part 2: Oscillatoriales. Springer Spektrum.
3. Kurmayer R, Sivonen K, Wilmotte A, Salmaso N. 2017. Molecular tools for the detection and quantification of toxigenic cyanobacteria. Wiley, Chichester. https://doi.org/10.1002/9781119332169
4. Bauer F, Fastner J, Bartha-Dima B, Breuer W, Falkenau A, Mayer C, Raeder U. 2020. Mass occurrence of Anatoxin-a- and Dihydroanatoxin-a-producing *Tychonema* sp.. Toxins 12:10.3390/toxins12110726. https://doi.org/10.3390/toxins12110726
5. Fastner J, Beulker C, Geiser B, Hoffmann A, Kröger R, Teske K, Hoppe J, Mundhenk L, Neurath H, Sagebiel D, Chorus I. 2018. Fatal neurotoxicosis in dogs associated with tychoplanktic, anatoxin-a producing *Tychonema* sp. In mesotrophic Lake Tegel, Berlin. Toxins 10:10.3390/toxins10020060. https://doi.org/10.3390/toxins10020060
6. Salmaso N, Cerasino L, Boscaini A, Capelli C. 2016. Planktic *Tychonema* (Cyanobacteria) in the large lakes south of the alps: phylogenetic assessment and toxigenic potential. FEMS Microbiol Ecol 92:fiw155. https://doi.org/10.1093/femsec/fiw155
7. Cerasino L, Capelli C, Salmaso N. 2017. A comparative study of the metabolic profiles of common nuisance cyanobacteria in Southern perialpine lakes. Adv Ocean Limnol 8:22–32. https://doi.org/10.4081/aiol.2017.6381
8. Andrews S. FastQC: a quality control tool for high throughput sequence data. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
9. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. Genome Res 27:824–834. https://doi.org/10.1101/gr.213959.116
10. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033
11. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146. https://doi.org/10.1038/nmeth.3103
12. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. Metabat 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359
13. Wu Y-W, Simmons BA, Singer SW. 2016. Maxbin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638
14. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843. https://doi.org/10.1038/s41564-018-0171-1
15. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o. Nat Microbiol 6:3–6. https://doi.org/10.1038/s41564-020-00834-3
16. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. Checkm: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114
17. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2022. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics 38:5315–5316. https://doi.org/10.1093/bioinformatics/btac672
18. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, Zhu Q, Bolzan M, Cumbo F, May U, Sanders JG, Zolfo M, Kopylova E, Pasolli E, Knight R, Mirarab S, Huttenhower C, Segata N. 2020. Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat Commun 11:2500. https://doi.org/10.1038/s41467-020-16366-7
19. Pinto F, Tett A, Armanini F, Asnicar F, Boscaini A, Pasolli E, Zolfo M, Donati C, Salmaso N, Segata N. 2017. Draft genome sequence of the planktic cyanobacterium *Tychonema bourrellyi*. Genome Announcements 5:e01294–17. https://doi.org/10.1128/genomeA.01294-17
20. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. Refseq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. https://doi.org/10.1093/nar/gkaa1105