AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Phormidium yuhuli* AB48, Isolated from an Industrial Photobioreactor Environment

Yilin Qiu,[a] Avery J. C. Noonan,[a,b] Kalen Dofher,[a] Moritz Koch,[c] Brandon Kieft,[c] Xuan Lin,[e] Ryan M. Ziels,[b,e] Steven J. Hallam[a,b,c,d,f,g]

[a]Genome Science and Technology Program, University of British Columbia, Vancouver, British Columbia, Canada
[b]ECOSCOPE Training Program, University of British Columbia, Vancouver, British Columbia, Canada
[c]Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada
[d]Graduate Program in Bioinformatics, University of British Columbia, Vancouver, British Columbia, Canada
[e]Department of Civil Engineering, University of British Columbia, Vancouver, British Columbia, Canada
[f]Bradshaw Research Institute for Minerals and Mining, University of British Columbia, Vancouver, British Columbia, Canada
[g]Life Sciences Institute, University of British Columbia, Vancouver, British Columbia, Canada

ABSTRACT We report the genome of *Phormidium yuhuli* AB48, which includes a circular chromosome and a circular plasmid (4,747,469 bp and 51,599 bp, respectively). This is currently the only closed reference genome of an isolate of the *Phormidium* genus, based on the Genome Taxonomy Database (GTDB), providing a potential model system for sustainable biotechnology innovation.

Cyanobacteria have the potential to be harnessed as climate-positive microbial cell factories. Therefore, closed reference genomes for cyanobacterial isolates and mobile genetic elements are needed to provide biological parts for phylogenetic reconstruction, metabolic modeling, and pathway engineering. Members of the *Phormidium* genus, which are capable of forming dense biofilms, have recently emerged as candidate cyanobacterial chassis to realize this potential (1–3).

Here, we report the complete genome sequence of *Phormidium yuhuli* AB48, which was first described by Koch and colleagues (4). *P. yuhuli* AB48 was isolated from an industrial photobioreactor environment operated under elevated temperature (35°C to 45°C), salinity (10 g/L), and alkalinity (pH 9 to pH 11). Grown as a biofilm under these conditions, AB48 was found to co-occur with several other microorganisms (5) prior to isolation using a gradient culture method (4). To prepare genomic DNA for sequencing, the isolate was cultured under continuous light in 50 mL of Zarrouk medium (6) in an Erlenmeyer flask for 96 h prior to biomass collection. Genomic DNA extracted using a cetyltrimethylammonium bromide (CTAB)-chloroform extraction protocol (7) was sequenced on both Illumina HiSeq and Oxford Nanopore Technologies (ONT) MinION platforms. Default parameters were used for software related to sequencing data processing and analysis, unless otherwise noted.

For short-read sequencing, the NEBNext Ultra DNA library preparation kit (New England Biolabs) was used. DNA libraries were sequenced on the Illumina HiSeq platform, and 283,172,660 paired-end 150-bp reads (42,475.9 Mbp) were subsampled to 5,659,080 reads (848.9 Mbp) by BBMap (v.38.18) prior to assembly, for an average coverage of 170×. Trimming and quality filtering were performed using BBDuk (v.38.93), which removed 2,818 reads (4.3 Mbp). For long-read sequencing, size selection was performed using the Circulomics short-read eliminator XS kit (Pacific Biosciences [PacBio]) to enrich for fragments longer than 10 kbp. The sequencing library was prepared using the NEBNext companion module for ONT ligation sequencing (New England Biolabs), ONT ligation sequencing kit (SQK-LSK109), and Flongle sequencing expansion kit (EXP-FSE001) and then was sequenced with a Flongle flow cell (FLO-

Editor J. Cameron Thrash, University of Southern California

Address correspondence to Steven J. Hallam, shallam@mail.ubc.ca.

The authors declare a conflict of interest.

Received 3 August 2022
Accepted 5 December 2022
Published 10 January 2023

FLG001) on the MinION Mk1b sequencer. Base calling was performed using Guppy (v.5.0.16) and resulted in 130,887 ONT reads (834.3 Mbp), with an N_{50} value of 9,790 bp. A total of 64,617 reads (410.8 Mbp) passed quality filtering (Q values of >10 [N_{50}, 9,600 bp]), for an average coverage of 175.97×. Adapters were trimmed using Porechop (v.0.3.2) (8). Resulting short- and long-read data were hybrid assembled using Unicycler (v.0.4.8) (9). Two circular contigs, of 4,747,469 bp (GC content, 51.68%) and 51,599 bp (GC content, 48.61%), were assembled, corresponding to the isolated cyanobacterial genome and its associated plasmid, respectively, based on contig circularization prediction with Unicycler and classification using Plasflow (v.1.1) (9, 10).

Phylogenetic classification using GTDB-Tk (v.0.3.2) (11) indicated that the isolate is a new member of the *Phormidium_A* genus. The isolate genome was compared to all 12 isolate genomes and metagenome-assembled genomes (MAGs) within this Genome Taxonomy Database (GTDB)-defined genus on the basis of average nucleotide identity (ANI), using FastANI (v.1.32) (12). The genome of *Sodalinema gerasimenkoae* sp. strain IPPAS B-353 (13) (identity, 87.3%) was the most similar. Koch et al. designated this isolate a new species in the genus *Phormidium*, with the name *Phormidium yuhuli* AB48 (4). Open reading frame (ORF) prediction and genome annotation for *Phormidium yuhuli* AB48 were performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (v.6.2) (14). The complete genome contains 4,187 genes, 4,133 coding sequences (CDSs), 6 rRNA genes (two sets of 5S, 16S, and 23S rRNA genes, consistent with other members of the genus [1]), 3 ncRNAs, and 45 tRNA genes.

Data availability. Data from this project are publicly accessible through the NCBI under the BioProject accession number PRJNA834472 and the BioSample accession number SAMN28044976. Raw sequencing data are available from the Sequence Read Archive (SRA) under the SRA accession numbers SRX15625011 and SRX15625009. This BioProject also includes BioSample accession number SAMN28044958; this represents a distinct sample and contains a PacBio data set and metagenome assembly from the photobioreactor enrichment from which *P. yuhuli* AB48 was isolated (5).

ACKNOWLEDGMENTS

We thank Soheyl Mottahedeh, Manisha Shastri, Craig Fourie, and Kevin Wilson at AlgaBloom for providing photobioreactor biomass, Sean Formby for invaluable input on genome assembly best practices, and Tanja Woyke at the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) and Tom Pfeifer in the Biofactorial Automation Core Facility at the University of British Columbia for technical advice and support. We also thank Jill Campbell and the Musqueam Indian Band for assisting with the naming of *Phormidium yuhuli* AB48.

This work was performed under the auspices of the Natural Sciences and Engineering Research Council (NSERC) of Canada, Genome British Columbia, the Canada Foundation for Innovation (CFI), the G. Unger Vetlesen and Ambrose Monell Foundations, and Facilities Integrating Collaborations for User Science (FICUS) JGI-Environmental Molecular Science Laboratory (EMSL) project 50967, which was supported by the Office of Science of the U.S. DOE (contract DE-AC02-05CH11231) with essential automation support through the Biofactorial Automation Core Facility in the Life Sciences Institute at the University of British Columbia. A.J.C.N. was supported by the NSERC CREATE Ecosystem Services, Commercialization Platforms, and Entrepreneurship (ECOSCOPE) training program at the University of British Columbia and the Mitacs Globalink program. Y.Q., A.J.C.N., and K.D. were also supported by the NSERC CREATE Genome Science and Technology (GSAT) training program at the University of British Columbia.

S.J.H. is a cofounder of Koonkie, Inc., a bioinformatics consulting company that designs and provides scalable algorithmic and data analytic solutions in the cloud.

REFERENCES

1. Ataeian M, Vadlamani A, Haines M, Mosier D, Dong X, Kleiner M, Strous M, Hawley AK. 2021. Proteome and strain analysis of cyanobacterium *Candidatus* "Phormidium alkaliphilum" reveals traits for success in biotechnology. iScience 24:103405. https://doi.org/10.1016/j.isci.2021.103405.
2. Nies F, Mielke M, Pochert J, Lamparter T. 2020. Natural transformation of the filamentous cyanobacterium *Phormidium lacuna*. PLoS One 15:e0234440. https://doi.org/10.1371/journal.pone.0234440.
3. Nies F, Wörner S, Wunsch N, Armant O, Sharma V, Hesselschwerdt A, Falk

F, Weber N, Weiß J, Trautmann A, Posten C, Prakash T, Lamparter T. 2017. Characterization of *Phormidium lacuna* strains from the North Sea and the Mediterranean Sea for biotechnological applications. Process Biochem 59:194–206. https://doi.org/10.1016/j.procbio.2017.05.015.
4. Koch M, Noonan AJC, Qiu Y, Dofher K, Kieft B, Mottahedeh S, Shastri M, Hallam SJ. 2022. The survivor strain: isolation and characterization of *Phormidium yuhuli* AB48, a filamentous phototactic cyanobacterium with biotechnological potential. Front Bioeng Biotechnol 10:932695. https://doi.org/10.3389/fbioe.2022.932695.
5. Noonan AJC, Qiu Y, Kieft B, Formby S, Liu T, Dofher K, Koch M, Hallam SJ. 2022. Metagenome-assembled genomes for "*Candidatus* Phormidium sp. strain AB48" and co-occurring microorganisms from an industrial photobioreactor environment. Microbiol Resour Announc 11:e00447-22. https://doi.org/10.1128/mra.00447-22.
6. Delrue F, Alaux E, Moudjaoui L, Gaignard C, Fleury G, Perilhou A, Richaud P, Petitjean M, Sassi J-F. 2017. Optimization of *Arthrospira platensis* (spirulina) growth: from laboratory scale to pilot scale. Fermentation 3:59. https://doi.org/10.3390/fermentation3040059.
7. Morin N, Vallaeys T, Hendrickx L, Natalie L, Wilmotte A. 2010. An efficient DNA isolation protocol for filamentous cyanobacteria of the genus *Arthrospira*. J Microbiol Methods 80:148–154. https://doi.org/10.1016/j.mimet.2009.11.012.
8. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 3:e000132. https://doi.org/10.1099/mgen.0.000132.
9. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
10. Krawczyk PS, Lipinski L, Dziembowski A. 2018. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. Nucleic Acids Res 46:e35. https://doi.org/10.1093/nar/gkx1321.
11. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
12. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/s41467-018-07641-9.
13. Kupriyanova EV, Cho SM, Park Y-I, Pronina NA, Los DA. 2016. The complete genome of a cyanobacterium from a soda lake reveals the presence of the components of CO_2-concentrating mechanism. Photosynth Res 130:151–165. https://doi.org/10.1007/s11120-016-0235-0.
14. 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.