



# Draft Whole-Genome Sequences of the Polar Cyanobacterium *Leptolyngbya* sp. Strain Cla-17 and Its Associated Flavobacterium

 Bérangère Péquin,<sup>a</sup> Julien Tremblay,<sup>b</sup> Christine Maynard,<sup>b</sup> Jessica Wasserscheid,<sup>b</sup> Charles W. Greer<sup>a,b</sup>

<sup>a</sup>McGill University, Department of Natural Resource Sciences, Ste-Anne-de-Bellevue, Quebec, Canada

<sup>b</sup>Energy, Mining, and Environment Research Centre, National Research Council Canada, Montreal, Quebec, Canada

**ABSTRACT** Draft whole-genome sequences of a coculture are presented. One component was a polar cyanobacterium, *Leptolyngbya* sp. strain Cla-17. The second was a heterotrophic bacterium, *Flavobacterium saccharophilum*, found in the phycosphere of the cyanobacterium.

Cyanobacteria and some heterotrophic bacteria interact closely in the phycosphere microenvironment (1), which is conducive to various molecular exchanges between species (2). Because of these relationships, many cyanobacteria cannot grow in axenic cultures (3). Here, we report the draft whole-genome sequences of an Arctic cyanobacterium and its associated flavobacterium.

The cyanobacterium Cla-17 (strain PCCC\_Cla17 from the Polar Cyanobacteria Culture Collection [PCCC]) was cultured from Char Lake snow (74°42'30"N, 94°53'0"W) in 2008 by Harding and coworkers (4). Briefly, snow was melted and filtered through 0.2- $\mu$ m polycarbonate filters, and then the filters were incubated in liquid BG-11 medium (5). The culture was first exposed to natural light from Rolute Bay at 16°C; once back in the university laboratory, it was cultivated at 10°C under continuous light at 50  $\mu$ mol photons  $m^{-2} s^{-1}$ . The culture was kindly provided by Professor W. Vincent (Centre d'Études Nordiques [CEN], Université Laval, Québec, Canada). For sequencing, the culture was grown on liquid BG-11 medium at 14°C with a 12-h/12-h light/dark cycle at 5 to 28  $\mu E m^{-2} s^{-1}$  irradiation for 5 weeks. As reported here, this cyanobacterium grows in a mixed culture with a heterotrophic flavobacterium.

DNA was extracted using the DNeasy UltraClean microbial kit (Qiagen). A short-read library was prepared using the QIAseq FX DNA library kit (Qiagen) and sequenced on a MiSeq system (Illumina) using v2 chemistry (2  $\times$  250 bp). A 20-kb SMRTbell library was prepared and sequenced with one single-molecule real-time (SMRT) Cell on a Sequel system (Pacific Biosciences [PacBio]) using v3.0 chemistry at the Génome Québec Innovation Center (McGill University, Montréal, Canada). Genomic analyses were performed with the default settings for all software unless otherwise noted.

The *de novo* assembly was carried out using the Hierarchical Genome Assembly Process (HGAP4) (6) in SMRT Link v6.0.0. The coverage cutoff value was set to 30 $\times$  and the estimated genome size to 5 million bp. Raw subreads (1,210,976 subreads) with a read quality (RQ) value of <0.7 (pbcoretools v1.5.0) were omitted, and the remaining subreads (151,243 subreads) were used as input for the FALCON assembler (falcon-kit v1.2.2) (7). Polishing of the assembly was performed with Arrow v2.2.2 in SMRT Link v6.0.0 to give a total of 24 contigs, representing 12,039,402 bp. The polished assembly was scaffolded using SSPACE-LongRead (8) with default parameters and gave 10 scaffolds, totaling 12,173,746 bp. The presence/absence of circularity of the genomes and the overlapping ends were assessed and, if necessary, removed in the postprocessing steps of the SSPACE-LongRead. The genomes were not rotated to a certain base.

To further correct for artifacts, MiSeq sequencing data generated with the same starting DNA were aligned against scaffolds (BWA v0.7.17) (9). Ten consensus scaffolds were generated

**Editor** J. Cameron Thrash, University of Southern California

© Crown copyright 2022. 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 Bérangère Péquin, [berangere.pequin@mcgill.ca](mailto:berangere.pequin@mcgill.ca).

The authors declare no conflict of interest.

**Received** 24 January 2022

**Accepted** 3 June 2022

**Published** 27 June 2022

**TABLE 1** Genome statistics for each strain

Parameter	Data for:	
	<i>Leptolyngbya</i> sp.	<i>Flavobacterium saccharophilum</i>
No. of reads		
Illumina		2,640,000
PacBio		151,471
PacBio raw read $N_{50}$ (bp)		10,231
Total genome size (bp)	6,647,508	5,526,236
Size of putative chromosomal contigs (bp)	Contig 1: 2,498,688; contig 2: 2,920,439	Circular contig: 5,526,236
Size of other types of contigs (bp)	Contig 3: 94,057	None
Total no. and size (bp) of plasmids	6, ranging from 49,058 to 480,286	None
Chromosome $N_{50}$ (bp)	2,498,684	5,526,236
Coverage (×)		
Illumina	108	108
PacBio	400	400
GC content (%)	Contig 1: 46.81; contig 2: 47.34	35.75
No. of genes	6,226	4,570
No. of rRNAs	3	5
No. of tRNAs	204	56
CheckM completeness (%)	96.62	99.65
CheckM contamination (%)	0.12	0.71
Accession no.		
Assembly	<a href="https://ncbi.nlm.nih.gov/assembly/GCA_016807185.1">GCA_016807185.1</a>	<a href="https://ncbi.nlm.nih.gov/assembly/GCA_013112255.2">GCA_013112255.2</a>
BioSample	<a href="https://ncbi.nlm.nih.gov/biosample/SAMN17309407">SAMN17309407</a>	<a href="https://ncbi.nlm.nih.gov/biosample/SAMN14367226">SAMN14367226</a>

with bcftools v1.9 (10), more precisely with (i) bcftools mpileup, (ii) bcftools call, and (iii) bcftools norm. Plasmids were identified with PlasFlow v1.1 (11) and were not circularized. Final corrected scaffolds were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (12). The genomes were estimated to be highly complete by CheckM v1.0.4 (13). Finally, the taxonomic assignment was performed with CAT v5.2.3 (14). The results showed two very different bacterial species that are phylogenetically distant (*Flavobacterium* versus *Leptolyngbya*). The genome statistics for each strain are indicated in Table 1.

It is known that *Flavobacterium* species are regularly found with cyanobacteria. They can inhibit or enhance the growth of cyanobacteria and degrade compounds synthesized by cyanobacteria (15). However, the ecological implications of these close interactions remain largely unknown.

**Data availability.** This whole-genome shotgun project has been deposited in GenBank under the accession number [JAEVYN000000000](https://ncbi.nlm.nih.gov/assembly/JAEVYN000000000). Raw reads are available under the BioProject accession number [PRJNA612312](https://ncbi.nlm.nih.gov/bioproject/PRJNA612312). The SRA accession numbers for the raw PacBio Sequel and MiSeq data are [SRR11301546](https://ncbi.nlm.nih.gov/sra/SRR11301546) and [SRR11301545](https://ncbi.nlm.nih.gov/sra/SRR11301545), respectively.

## ACKNOWLEDGMENTS

The study was funded by GENICE, a Genome Canada project.

We acknowledge Compute Canada for access to the University of Waterloo High-Performance Computing (HPC) infrastructure (Graham system).

## REFERENCES

- Seymour JR, Amin SA, Raina J-B, Stocker R. 2017. Zooming in on the phycosphere: the ecological interface for phytoplankton-bacteria relationships. *Nat Microbiol* 2:17065. <https://doi.org/10.1038/nmicrobiol.2017.65>.
- Cai H, Jiang H, Krumholz LR, Yang Z. 2014. Bacterial community composition of size-fractionated aggregates within the phycosphere of cyanobacterial blooms in a eutrophic freshwater lake. *PLoS One* 9:e102879. <https://doi.org/10.1371/journal.pone.0102879>.
- Cole JJ. 1982. Interactions between bacteria and algae in aquatic ecosystems. *Annu Rev Ecol Syst* 13:291–314. <https://doi.org/10.1146/annurev.es.13.110182.001451>.
- Harding T, Jungblut AD, Lovejoy C, Vincent WF. 2011. Microbes in high Arctic snow and implications for the cold biosphere. *Appl Environ Microbiol* 77:3234–3243. <https://doi.org/10.1128/AEM.02611-10>.
- Hughes EO, Gorham PR, Zehnder A. 1958. Toxicity of a unialgal culture of *Microcystis aeruginosa*. *Can J Microbiol* 4:225–236. <https://doi.org/10.1139/m58-024>.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, long-read sequencing with single-molecule real-time (SMRT) technology. *Nat Methods* 10:560–561. <https://doi.org/10.1038/nmeth.2330>.

- finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
7. Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, Cramer GR, Delledonne M, Luo C, Ecker JR, Cantu D, Rank DR, Schatz MC. 2016. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods* 13:1050–1054. <https://doi.org/10.1038/nmeth.4035>.
  8. Boetzer M, Pirovano W. 2014. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* 15:211. <https://doi.org/10.1186/1471-2105-15-211>.
  9. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
  10. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
  11. 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>.
  12. Tatusova T, DiCuccio M, Badretdin A, Chetverin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
  13. 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>.
  14. von Meijenfeldt FAB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE. 2019. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol* 20:217. <https://doi.org/10.1186/s13059-019-1817-x>.
  15. Berg KA, Lyra C, Sivonen K, Paulin L, Suomalainen S, Tuomi P, Rapala J. 2009. High diversity of cultivable heterotrophic bacteria in association with cyanobacterial water blooms. *ISME J* 3:314–325. <https://doi.org/10.1038/ismej.2008.110>.