Complete genome of Flavobacterium pectinovorum str. ZE23VCel01 obtained through Nanopore Q20+ chemistry

Serra Moncadas, Lucas ; Schnellmann, Vanessa ; Hofer, Cyrill ; Rain-Franco, Angel ; Andrei, Adrian-Stefan

Abstract: This study reports the complete genome of Flavobacterium pectinovorum str. ZE23VCel01 isolated from a freshwater environment. By means of Nanopore Q20+ chemistry, the chromosome was assembled as a circular element with a length of 5.9 Mbp, a GC content of 33.58%, and a coverage of 122×.

DOI: https://doi.org/10.1128/mra.00715-23

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Computational Biology | Announcement

Complete genome of *Flavobacterium pectinovorum* str. ZE23VCel01 obtained through Nanopore Q20+ chemistry

Lucas Serra Moncadas,[1] Vanessa Schnellmann,[2] Cyrill Hofer,[1] Angel Rain-Franco,[3] Adrian-Stefan Andrei[2]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT This study reports the complete genome of *Flavobacterium pectinovorum* str. ZE23VCel01 isolated from a freshwater environment. By means of Nanopore Q20+ chemistry, the chromosome was assembled as a circular element with a length of 5.9 Mbp, a GC content of 33.58%, and a coverage of 122×.

KEYWORDS genomics, DNA sequencing, Nanopore, *Flavobacterium*

Members within the *Flavobacterium* genus primarily occupy freshwater environments and hold significant ecological relevance, as they are accountable for serious diseases in freshwater fishes. By means of Nanopore Q20+ technology, it is possible to gain a novel perspective on their biological features and ecological impact on aquatic environments (1).

The isolated strain from the surface waters of Lake Zurich (47°18′N, 8°34′E, Switzerland) was cultivated in agar plates (15 g/L) supplemented with artificial lake water (2) using cellobiose as a carbon source (100 µM final concentration). After 21 days of cultivation, colonies were purified through sequential streaking cycles and cryopreserved as described elsewhere (2). The isolated strain was reactivated in artificial lake water supplemented with glucose (100 µM) for 1 week and grown in commercial LB medium (Difco, 240210) for 2 days. From the resuscitated culture (3), DNA extraction was performed using the Quick-DNA HMW MagBead kit (Zymo Research). The obtained DNA was purified with Beckman Coulter AMPure XP magnetic beads and subsequently loaded on a FLO-MIN114 (R10.4.1) new chemistry flow cell and used for sequencing on a Nanopore minION Mk1B. The sequencing 1D library was constructed with the SQK-NBD114.24 Native Barcoding Kit 24 V14 (ONT, Oxford, UK) in conformity with the manufacturer's instructions and long DNA fragment selection (without any prior DNA fragmentation).

Sequencing resulted in 1,108,124 raw reads (mean read length 3.9 kbp; median read quality 14.3; N50 5.4 kbp) that were basecalled with Guppy v.6.4.6 (superaccurate basecalling, 400 bp) before quality filtration with chopper v.0.2.0 (4). A minimum threshold of Q17 was used, followed by the removal of the first 10 bases and reads < 1,000 bp (164,407 reads; mean read length 4.5 kbp; median read quality 17.9; N50 5.5 kbp). Assembly was conducted with Flye 2.9.2-b1786 (5) (-nano—corr option), leading to a circular chromosome (5,936,092 bp). The obtained circular genome was subsequently classified using GTDB-Tk v2.2.6 (6) combined with a comparison of its 16S rRNA gene (predicted with barrnap 0.9) against the SILVA database (v138.1) (7). Genome completeness (99.22%) and contamination (1.88%) were assessed with CheckM v1.1.3 (8). Prokka v1.12 (9) and NCBI's PGAP pipeline (10) were used to predict coding DNA sequences and tRNAs (4,970 genes; 30 pseudo-genes). BlastKOALA (11) was used to assign KO identifiers to orthologous genes. Metabolic reconstruction and general biological functions were conducted with the online KEGG mapping tools using the previously obtained KO numbers. PFAM domains were identified in the proteome using the script pfam_scan.pl

Editor John J. Dennehy, Queens College, Queens, New York, USA

Address correspondence to Adrian-Stefan Andrei, stefan.andrei@limnol.uzh.ch, or Lucas Serra Moncadas, lucas.serramoncadas@uzh.ch.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 4 August 2023
Accepted 3 November 2023
Published 6 December 2023

with the PFAM database release 32 (12). Default parameters were used except where otherwise specified.

Taxonomical classification suggests that the genome belongs to a new *Flavobacterium pectinovorum* strain by GTDB v214.1 (95.17% similarity to GCA_900142715, 98.13% coverage) and SILVA v138.1 (98.87% similarity) databases. The genome-inferred metabolic reconstructions pictured a microorganism with the metabolic capability to synthesize most of the proteinogenic amino acids apart from arginine.

ACKNOWLEDGMENTS

This work was accomplished within the "Microbial smackdown" module of the block course BIO 290: Aquatic microbial ecology (Limnological Station, University of Zurich).

This study was supported by the Ambizione grant PZ00P3193240 (Swiss National Science Foundation).

We would like to thank Bettina Sieber for help with the strain isolation.

AUTHOR AFFILIATIONS

[1]Microbial Evogenomics Lab, Limnological Station, Department of Plant and Microbial Biology, University of Zurich, Kilchberg, Switzerland
[2]Faculty of Sciences, University of Zurich, Zurich, Switzerland
[3]Limnological Station, Department of Plant and Microbial Biology, University of Zurich, Kilchberg, Switzerland

AUTHOR ORCIDs

Angel Rain-Franco http://orcid.org/0000-0002-0389-7154
Adrian-Stefan Andrei http://orcid.org/0000-0003-1425-7168

FUNDING

Funder	Grant(s)	Author(s)
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (SNF)	193240	Lucas Serra Moncadas
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (SNF)	193240	Adrian-Stefan Andrei
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung (SNF)	193240	Cyrill Hofer

AUTHOR CONTRIBUTIONS

Lucas Serra Moncadas, Supervision, Writing – original draft, Writing – review and editing | Vanessa Schnellmann, Formal analysis, Software, Writing – original draft | Cyrill Hofer, Supervision, Writing – review and editing | Angel Rain-Franco, Conceptualization, Investigation, Supervision | Adrian-Stefan Andrei, Conceptualization, Funding acquisition, Validation, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

Sequence data are available through the National Center for Biotechnology Information (NCBI) via the BioProject PRJNA991734 (Biosample SAMN36315996, Accession GCA_030505355.1).

REFERENCES

1. Loch TP, Faisal M. 2015. Emerging flavobacterial infections in fish: a review. J Adv Res 6:283–300. https://doi.org/10.1016/j.jare.2014.10.009
2. Le Moigne A, Randegger F, Gupta A, Petchey OL, Pernthaler J. 2023. Stochasticity causes high β-diversity and functional divergence of bacterial assemblages in closed systems. Ecology 104:e4005. https://doi.org/10.1002/ecy.4005

3. Rain-Franco A, de Moraes GP, Beier S. 2020. Cryopreservation and resuscitation of natural aquatic prokaryotic communities. Front Microbiol 11:597653. https://doi.org/10.3389/fmicb.2020.597653
4. De Coster W, Rademakers R. 2023. NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics 39:btad311. https://doi.org/10.1093/bioinformatics/btad311
5. Kolmogorov M, Bickhart DM, Behsaz B, Gurevich A, Rayko M, Shin SB, Kuhn K, Yuan J, Polevikov E, Smith TPL, Pevzner PA. 2020. metaFlye: scalable long-read metagenome assembly using repeat graphs. Nat Methods 17:1103–1110. https://doi.org/10.1038/s41592-020-00971-x
6. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH, Hancock J. 2020. GTDB-TK: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
7. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–6. https://doi.org/10.1093/nar/gks1219
8. 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
9. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153
10. Tatusova T, DiCuccio M, Badretdin A, Chetvernin 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
11. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol 428:726–731. https://doi.org/10.1016/j.jmb.2015.11.006
12. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. 2014. Pfam: the protein families database. Nucleic Acids Res 42:D222–30. https://doi.org/10.1093/nar/gkt1223