AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Genome Sequence of *Cupriavidus campinensis* Strain G5, a Member of a Bacterial Consortium Capable of Polyethylene Degradation

Beate Schneider,[a] Friedhelm Pfeiffer,[b] Mike Dyall-Smith,[b,c] Hans-Jörg Kunte[a]

[a]Bundesanstalt für Materialforschung und -prüfung, Division of Biodeterioration and Reference Organisms, Berlin, Germany
[b]Computational Biology Group, Max Planck Institute of Biochemistry, Martinsried, Germany
[c]Veterinary Biosciences, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, Australia

ABSTRACT Nine different bacterial isolates were recovered from landfills. Each isolate was obtained in pure culture. As a consortium, the bacteria degrade polyethylene. The complete genome sequence of strain G5 was determined by PacBio sequencing. Using the TYGS for taxonomic classification, strain G5 was assigned to the species *Cupriavidus campinensis*.

As part of a bacterial consortium, *Cupriavidus campinensis* G5 degrades polyethylene and was isolated from a former plastic landfill (Niemegk-Neuendorf, Germany [52°04′59″N, 12°41′59″E]) using a modified iChip procedure (1). A bacterial suspension from the landfill was diluted in soft agar, loaded into an iChip, and incubated for 4 weeks (2, 3). The emerging microcolony of strain G5 was streaked on agar, and cells from a single colony were inoculated into casein-soy broth (casein peptone, 15 g L^{-1}; soymeal peptone, 5 g L^{-1}; yeast extract, 0.5 g L^{-1}; NaCl, 5 g L^{-1}; glucose, 10 mM [pH 7.2]). After growth for 24 h at 30°C, cells were centrifuged, washed twice with phosphate-buffered saline, and kept at −80°C.

The following procedures were performed at Genewiz/Azenta (Leipzig, Germany), including DNA extraction using the Genomic-tip 100/G kit (Qiagen, Hilden, Germany), DNA shearing to ~10 kb using the g-TUBE device without size selection, quality (NanoDrop and pulsed-field gel electrophoresis analyses) and quantity (Qubit 2.0) assessments, and sequencing library preparation using the SMRTbell library preparation kit v2.0 (Pacific Biosciences [PacBio], Menlo Park, CA, USA). The library was sequenced on a PacBio Sequel sequencer (4) using one single-molecule real-time (SMRT) cell. The sequencing yielded 68,653 reads, from which 389,796 subreads were obtained (average length, 5,724 bp; read N_{50}, 7,235 bp; total size, 2.231 Mbp). An assembly was performed within the SMRT Link Suite v5.0.0.6792 at Genewiz/Azenta, including checks for quality and coverage using default parameters. Subread coverage was reduced to 200-fold by random down-sampling, and the genome was automatically assembled using Canu v1.7 (5) with a coverage cutoff set to 40-fold and default parameters otherwise. The assembly was evaluated for overlapping contigs, which eventually were merged, and for potential assembly artifacts, which were dropped. Contigs were polished with raw subreads using Arrow v2.3.2 (https://github.com/PacificBiosciences/pbbioconda) with default parameters. Initial assembly generated 13 contigs, with a total length of 6,468,862 bp.

A supervised assembly using Geneious v10.2 (6) for mapping of PacBio reads to contigs, with subsequent integration and editing of contigs, was performed. Long duplications, such as two highly similar 27-kb prophages, were resolved by mapping adjacent unique regions to closely related genomes and by analyzing long PacBio reads; this resulted in two replicons. Chromosome 1 (GenBank accession number CP097330), which was obtained with 331-fold average coverage, is complete, circular, and 3,780,571 bp long, with a G+C content of 66.2%. Chromosome 2 (GenBank accession number CP097331), which was obtained with 339-fold average coverage, is complete, circular, and 2,524,655 bp long, with a G+C content of 66.7%. Start bases were set according to related chromosomes that are sufficiently similar

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Address correspondence to Hans-Jörg Kunte, hans-joerg.kunte@bam.de.

The authors declare no conflict of interest.

Received 31 May 2022
Accepted 6 September 2022
Published 20 September 2022

and have been scrutinized for replication origin consensus sequences (7, 8). For GenBank accession number CP097330, the start base is consistent with that of GenBank accession number CP000352, upstream of *dnaA* (7). For GenBank accession number CP097331, the start base is consistent with that of GenBank accession number CU633750, upstream of *repA* (8). No additional replicons were detected.

The TYGS (9) assigned strain G5 to the species *Cupriavidus campinensis*, with a digital DNA-DNA hybridization (dDDH) value (d4) of 96.7% (95% confidence interval, 95.5 to 97.6%) (10). The small G+C content difference (0.20%) between the genome and the type strain, *Cupriavidus campinensis* LMG19282, supports the assignment. The genome was submitted to GenBank and annotated using PGAP v6.1 (11).

Data availability. The annotated genome has been deposited in GenBank under the BioProject accession number PRJNA837128 and the BioSample accession number SAMN28191058. The nucleotide sequence accession numbers are CP097330 (major chromosome) and CP097331 (minor chromosome). The Sequence Read Archive (SRA) accession number is SRR19184626.

ACKNOWLEDGMENTS

The isolation and sequencing were part of the BMBF-funded project ENSURE (Federal Ministry of Education and Research grant 02WPL1449C).

REFERENCES

1. Berdy B, Spoering AL, Ling LL, Epstein SS. 2017. *In situ* cultivation of previously uncultivable microorganisms using the ichip. Nat Protoc 12:2232–2242. https://doi.org/10.1038/nprot.2017.074.
2. Schneider B, Pfeiffer F, Dyall-Smith M, Kunte H-J. 2022. Genome sequence of *Pseudomonas veronii* strain G2, a member of a bacterial consortium capable of polyethylene degradation. Microbiol Resour Announc 11:e00365-22. https://doi.org/10.1128/mra.00365-22.
3. Schneider B, Pfeiffer F, Dyall-Smith M, Kunte H-J. 2022. Genome sequence of *Micromonospora aurantiaca* strain G9, a member of a bacterial consortium capable of polyethylene degradation. Microbiol Resour Announc 11:e01148-21. https://doi.org/10.1128/mra.01148-21.
4. Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Genomics Proteomics Bioinformatics 13:278–289. https://doi.org/10.1016/j.gpb.2015.08.002.
5. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
6. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
7. Janssen PJ, Van Houdt R, Moors H, Monsieurs P, Morin N, Michaux A, Benotmane MA, Leys N, Vallaeys T, Lapidus A, Monchy S, Medigue C, Taghavi S, McCorkle S, Dunn J, van der Lelie D, Mergeay M. 2010. The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. PLoS One 5:e10433. https://doi.org/10.1371/journal.pone.0010433.
8. Amadou C, Pascal G, Mangenot S, Glew M, Bontemps C, Capela D, Carrere S, Cruveiller S, Dossat C, Lajus A, Marchetti M, Poinsot V, Rouy Z, Servin B, Saad M, Schenowitz C, Barbe V, Batut J, Medigue C, Masson-Boivin C. 2008. Genome sequence of the β-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. Genome Res 18:1472–1483. https://doi.org/10.1101/gr.076448.108.
9. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10:2182. https://doi.org/10.1038/s41467-019-10210-3.
10. Meier-Kolthoff JP, Auch AF, Klenk HP, Goker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.
11. 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.