

# Complete genome sequence of *Cellulomonas* sp. GTC 20042 isolated from human bile

Jun Yonetamari,<sup>1,2</sup> Masahiro Hayashi,<sup>3,4,5</sup> Yoshinori Muto,<sup>4</sup> Tetsuya Iida,<sup>6</sup> Kaori Tanaka<sup>1,3,4,5</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** The genus *Cellulomonas* comprises aerobic, Gram-positive, rod-shaped bacteria commonly present in the soil and aquatic environments. Herein, we report the isolation of *Cellulomonas* sp. GTC20042 from human bile and the complete genome sequence of this strain, consisting of a 4,145,753-bp circular chromosome.

**KEYWORDS** *Cellulomonas*, genome analysis

*Cellulomonas*, proposed by Bergey et al. (1), includes aerobic, Gram-positive rods, mostly isolated from soil (2–4), with rare human sources (5). In 2021, *Cellulomonas* sp. GTC20042 was isolated from human bile using sheep blood agar plates (Kyokuto Pharmaceutical Industrial Co., Ltd., Tokyo, Japan), incubated at 37°C for 48 h under aerobic conditions. The strain was maintained under the same conditions.

The sample was cultured aerobically on Vitalmedia sheep blood agar plates (Kyokuto) and incubated at 37°C for 48 h. Genomic DNA was extracted from the bacterial pellet using Quick Taq HS DyeMix (TOYOBO Co., Ltd., Osaka, Japan). The cells were suspended in 100 µL of distilled water, heated at 100°C for 10 min, and centrifuged at 15,000 rpm for 5 min. The genomic DNA-containing supernatant was used for all genome analyses.

The complete genome of the isolate was sequenced using a combination of long- (Oxford Nanopore Technologies [ONT], Tokyo, Japan) and short-read sequencing (Illumina, San Diego, CA, USA) (6–8). For long-read sequencing, genomic DNA (1,000 ng) was treated with the Short Read Eliminator XS (Circulomics, Japan Genetic) to remove fragments below 10 kb. After end-repair and a second SRE XS treatment, the library was constructed using the Native Barcode Sequencing Kit (SQK-NBD114-24; ONT) without mechanical DNA shearing. The final library was purified using AMPure XP beads. Sequencing was performed using a PromethION 2i instrument with a FLO-PRO114M flow cell, and data were base-called in super accuracy mode using Dorado v7.2.14 implemented in MinKNOW v23.11.8 (all by Oxford Nanopore Technologies; <https://nanopore-tech.com>). The run generated 2,761,482 raw reads with an N50 read length of 11,797 bp (Table 1). The raw long reads were trimmed and quality-filtered using NanoFilt v2.7.1 (7) with the parameters “-q 10 -headcrop 50 -l 1000.” For short-read sequencing, the Illumina DNA Prep (M) Tagmentation Kit (Illumina) was used for library construction, followed by 2 × 151-bp paired-end sequencing on the NovaSeq 6000 platform (Illumina). Raw short reads were processed using fastp v0.20.1 (6) with the parameters “-q 30 -n 20 -t 1 -T 1.” The short-read and mean long-read sequence qualities were assessed using fastp and NanoPlot v1.32.1, respectively (7). High-quality short (>92% bases >Q30 on average) and long (mean read quality: 14.3) reads were assembled using Unicycler v0.4.8 with default settings (8). Circular contigs generated by Unicycler were automatically trimmed for overlapping ends, and the chromosomes were rotated to start at the *dnaA* gene when detected, as part of the default pipeline (8). The assembled contig graph was visualized and assessed for integrity using BlobTools v1.0 (9). Genome completeness and contamination were evaluated using CheckM v1.2.2 (10), and functional annotation

**Editor** David Rasko, University of Maryland Baltimore, Baltimore, Maryland, USA

Address correspondence to Masahiro Hayashi, hayashi.masahiro.w8@f.gifu-u.ac.jp.

The authors declare no conflict of interest.

See the funding table on p. 3.

**Received** 1 October 2025

**Accepted** 7 December 2025

**Published** 12 January 2026

Copyright © 2026 Yonetamari et al. 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/).

**TABLE 1** Complete genome sequence information of *Cellulomonas* sp. GTC20042 isolated from human bile

|                                                 | Strain name                      |
|-------------------------------------------------|----------------------------------|
| Parameter                                       | <i>Cellulomonas</i> sp. GTC20042 |
| Illumina sequencing <sup>a</sup>                |                                  |
| No. of reads                                    | 2,761,482                        |
| Size (bp)                                       | 408,271,901                      |
| Avg coverage (x)                                | 91                               |
| DRA accession no.                               | <a href="#">DRR720122</a>        |
| ONT sequencing <sup>a</sup>                     |                                  |
| No. of reads                                    | 104,041                          |
| Size (bp)                                       | 694,823,651                      |
| Average read length (bp)                        | 10,144                           |
| Average coverage (x)                            | 119                              |
| N50                                             | 11,797                           |
| DRA accession no.                               | <a href="#">DRR720121</a>        |
| Assembly                                        |                                  |
| Estimated genome completeness (%) <sup>b</sup>  | 93.66                            |
| Estimated genome contamination (%) <sup>b</sup> | 3.93                             |
| Genome structure                                | 1 chromosome                     |
| DDBJ/GenBank accession no.                      | <a href="#">AP043678</a>         |
| Genome size (bp)                                | 4,145,753                        |
| GC content (%)                                  |                                  |
| (chromosome/plasmid name)                       | 74.0                             |
| No. of coding sequences <sup>c</sup>            | 3,826                            |
| Number of rRNAs <sup>c</sup>                    | 9                                |
| Number of tRNAs <sup>c</sup>                    | 48                               |
| Number of CRISPRs <sup>c</sup>                  | 2                                |

<sup>a</sup>DRA, DNA DataBank of Japan (DDBJ) Sequence Read Archive.<sup>b</sup>Determined using CheckM.<sup>c</sup>DFAST, DDBJ Fast Annotation and Submission Tool.

was performed using DFAST (11). Default parameters were used for all software unless otherwise specified.

The taxonomic identification of *Cellulomonas* sp. GTC20042 was performed using GTDB-Tk v2.4.0 (12), classifying the strain within the genus *Cellulomonas*.

BLASTn analysis showed no identical species. The genome-derived 16S rRNA gene was most similar (98.4%) to *Cellulomonas pakistanensis* JCM 18755 (accession number: [AB618146](#)), suggesting this is a novel *Cellulomonas* species.

## ACKNOWLEDGMENTS

We thank Kyoko Hatazaki, Akiko Katano, and Ayako Nagasawa for their technical support. We declare no competing financial interests.

This study was approved by the ethics committee of our institute. This study was supported by the Grant for Joint Research Project of the Research Institute for Microbial Diseases, The University of Osaka (JRPRIMD25B3).

## AUTHOR AFFILIATIONS

<sup>1</sup>United Graduate School of Drug Discovery and Medical Information Sciences, Gifu University, Gifu City, Gifu, Japan

<sup>2</sup>Division of Clinical Laboratory, Gifu University Hospital, Gifu, Japan

<sup>3</sup>Division of Anaerobe Research, Life Science Research Center, Gifu University, Gifu City, Gifu, Japan

<sup>4</sup>Institute for Glyco-core Research iGCORE, Gifu University, Gifu City, Gifu, Japan

<sup>5</sup>Gifu University Center for Conservation of Microbial Genetic Resource, Gifu, Japan

<sup>6</sup>Pathogenic Microbes Repository Unit Research Institute For Microbial Diseases, Osaka University, Suita, Osaka, Japan

## AUTHOR ORCIDS

Masahiro Hayashi  <http://orcid.org/0000-0001-6317-9077>

Tetsuya Iida  <http://orcid.org/0000-0001-5134-1794>

## FUNDING

| Funder                                                                                                     | Grant(s)    | Author(s)        |
|------------------------------------------------------------------------------------------------------------|-------------|------------------|
| Grant for Joint Research Project of the Research Institute for Microbial Diseases, The University of Osaka | JRPRIMD25B3 | Masahiro Hayashi |

## AUTHOR CONTRIBUTIONS

Jun Yonetamari, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft | Masahiro Hayashi, Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review and editing | Yoshinori Muto, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – review and editing | Tetsuya Iida, Conceptualization, Funding acquisition, Project administration, Supervision, Validation, Writing – review and editing | Kaori Tanaka, Project administration, Supervision, Writing – review and editing

## DATA AVAILABILITY

The complete genome sequence of *Cellulomonas* sp. GTC20042, isolated from human bile, has been deposited in the DNA DataBank of Japan (DDBJ; <https://www.ddbj.nig.ac.jp/index.html>) under the following accession number: [AP043678](https://www.ddbj.nig.ac.jp/index.html). Raw sequence data for GTC20042 have been deposited in the DDBJ Sequence Read Archive under the accession numbers [DRR720121](https://www.ddbj.nig.ac.jp/index.html) and [DRR720122](https://www.ddbj.nig.ac.jp/index.html).

This study was performed in accordance with the Declaration of Helsinki.

## REFERENCES

1. Society of American B. 1923. Bergey's manual of determinative bacteriology: a key for the identification of organisms of the class schizomycetes. Williams & Wilkins.
2. Shi Z, Luo G, Wang G. 2012. *Cellulomonas carbonis* sp. nov., isolated from coal mine soil. *Int J Syst Evol Microbiol* 62:2004–2010. <https://doi.org/10.1099/ijs.0.034934-0>
3. Lee HJ, Kim SY, Whang KS. 2020. *Cellulomonas citrea* sp. nov., isolated from paddy soil. *Int J Syst Evol Microbiol* 70:5304–5311. <https://doi.org/10.1099/ijs.0.004409>
4. An DS, Im WT, Yang HC, Kang MS, Kim KK, Jin L, Kim MK, Lee ST. 2005. *Cellulomonas terrae* sp. nov., a cellulolytic and xylanolytic bacterium isolated from soil. *Int J Syst Evol Microbiol* 55:1705–1709. <https://doi.org/10.1099/ijs.0.03696-0>
5. Lai P-C, Chen Y-S, Lee SS-J. 2009. Infective endocarditis and osteomyelitis caused by *Cellulomonas*: a case report and review of the literature. *Diagn Microbiol Infect Dis* 65:184–187. <https://doi.org/10.1016/j.diagmicrobio.2009.06.001>
6. Hayashi M, Yonetamari J, Muto Y, Tanaka K. 2024. Complete genome sequence of *Peptostreptococcus porci* isolated from porcine endocarditis in Japan. *Microbiol Resour Announc* 13:e0020124. <https://doi.org/10.1128/mra.00201-24>
7. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
8. 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>
9. Laetsch DR, Blaxter ML. 2017. BlobTools: interrogation of genome assemblies. *F1000Res* 6:1287. <https://doi.org/10.12688/f1000research.12232.1>
10. 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>
11. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>
12. Parks DH, Chuvochina M, Chaumeil PA, Rinke C, Mussig AJ, Hugenholtz P. 2020. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* 38:1079–1086. <https://doi.org/10.1038/s41587-020-0501-8>