

Complete genome sequence of *Varibaculum* sp. GTC20041 isolated from a clinical specimen in Japan

Masahiro Hayashi,^{1,2,3} Jun Yonetamari,^{2,4,5} Yoshinori Muto,² Kaori Tanaka^{1,2,3,5,6}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report the complete genome sequence of *Varibaculum* strain GTC20041 isolated from a clinical specimen in Japan. The genome comprised a circular chromosome with a length of 2,266,073 bp.

KEYWORDS genomes, *Varibaculum*, anaerobes

Varibaculum is a gram-positive, facultatively anaerobic, nonspore-forming, and nonmotile bacterial genus from the family *Actinomycetaceae*. *Varibaculum* species are commonly found in various human samples (1, 2). In particular, strain GTC20041 was identified in a human clinical specimen in Aichi, Japan, in 2021. More specifically, it was isolated from the ascites of an adult patient via anaerobic culture from an ascites sample submitted for clinical testing. The strain was maintained on Brucella HK agar supplemented with 5% laked sheep blood at 37°C for 48 h under anaerobic conditions (82% N₂, 10% CO₂, and 8% H₂). The strain was repeatedly subcultured three times under the same conditions for purification. The present study was conducted in accordance with the provisions of the Declaration of Helsinki. Identification via matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) (MALDI Biotyper, Bruker, Germany 12438 MSP) in our laboratory confirmed the classification of strain GTC20041 solely as *Varibaculum* at the genus level.

To obtain the complete genome sequence, the genome of GTC20041 was sequenced using PromethION 2i (Oxford Nanopore Technologies, Oxford, UK) for long-read sequencing and the NovaSeq 6000 system (Illumina Inc., USA) for short-read sequencing as previously described (3–5). Genomic DNA was extracted from the pellet using NucleoBond HMW DNA (MACHEREY-NAGEL, Japan) according to the manufacturer's instruction. For long-read sequencing, a library was constructed using a Native Barcode Sequencing Kit (SQK-NBD114-24; Oxford Nanopore Technologies) without shearing, followed by the removal of small fragments using a Short Read Eliminator XS (Circulomics, Baltimore, MD, USA). Sequencing was performed using a FLO-PRO114M flow cell. The MinkNOW v.23.11.8 software was used for base calling and raw reads were trimmed and quality filtered using NanoFilt v.2.7.1 (6) with the parameters "-l 1000 -q 10 -headcrop 50." For short-read sequencing, the DNA Prep (M) Tagmentation Kit (Illumina Inc.) was used for library construction. Subsequently, 2 × 151 bp paired-end sequencing was performed using the NovaSeq 6000 platform. Raw sequencing reads were processed using fastp v.0.20.1 (7) with the parameters "-q 30 n 20 t 1 T 1." The quality of short reads was also assessed using fastp v.0.20.1 (7). The mean read quality of long reads was scored using NanoPlot 1.32.1 (7). Short (>91.8% of bases > Q30 averaged) and long reads (mean read quality of 15.9) were assembled using Unicycler v.0.4.8 (8) with default settings. In particular, assembly circularization and rotation were performed automatically. The assembly was rotated to start with the *dnaA* gene on the forward strand. Annotation of the assembled genomes was performed using the DDBJ Fast Annotation and Submission

Editor Catherine Putonti, Loyola University Chicago, Chicago, Illinois, USA

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

The authors declare no conflict of interest.

Received 20 June 2025

Accepted 15 August 2025

Published 12 September 2025

Copyright © 2025 Hayashi 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 Information on the complete genome sequence of a *Varibaculum* sp. strain isolated from a human clinical specimen in Japan

	Strain name
Parameter	<i>Varibaculum</i> sp. GTC20041
NovaSeq6000 sequencing ^a	
No. of reads	6,577,930
Size (kb)	983,073
Avg coverage (×)	434
DRA accession no.	DRR683174
ONT sequencing ^a	
No. of reads	213,594
Size (kb)	800,004
Avg read length (bp)	7,249
Avg coverage (×)	353
N50	7,318
DRA accession no.	DRR683173
Assembly	
Assembly N50 (bp) ^c	2,266,073
Estimated genome completeness (%) ^d	99.05
Estimated genome contamination (%) ^d	0.71
Genome structure	1 chromosome
DDBJ/GenBank accession no.	AP041144
Genome size (bp)	2,266,073
GC content (%)	
(chromosome/plasmid name)	52.6
No. of coding sequences ^b	1,937
Number of rRNAs ^b	6
Number of tRNAs ^b	49
Number of CRISPRs ^b	2

^aDRA, DDBJ Sequence Read Archive.^bDFAST, DDBJ Fast Annotation and Submission Tool.^cDetermined with seqkit v2.3.0.^dDetermined with CheckM v1.2.2.

Tool (DFAST) (<https://dfast.ddbj.nig.ac.jp/>) (9). Average nucleotide identity (ANI) analysis was conducted using PyANI v.0.2.12 (10).

Table 1 summarizes the genomic information. Quality assessment and genome statistics were computed using CheckM (v.1.2.2) (11) and seqKit (v.2.9.0) (12), respectively. CheckM indicated that the genome was 99.05% complete with 0.71% contamination. DFAST predicted 1,937 coding sequences, six ribosomal RNAs, 49 transfer RNAs, and 2 CRISPR sequences.

BLASTN analysis (13) of the 16S rRNA gene sequence identified *Varibaculum cambriense* CCUG 44998^T (Acc.no. [AJ428402](#)) as the closest match, with 98.6% identity. However, species identification via the genome sequence using GTDB-tK (v.2.4.0) was unsuccessful (14). ANI analysis of the complete genome sequence revealed that GTC20041 is most closely related to *V. cambriense* ([ASM42006v1](#)), with an ANI value of 88.1%. Based on these results, we could not assign the strain to a species at this time.

ACKNOWLEDGMENTS

We thank Kyoko Hatazaki, Akiko Katano, and Ayako Nagasawa for their technical support. This study was exempted from institutional ethics committee approval.

AUTHOR AFFILIATIONS

¹Institute for Glyco-core Research iGCORE, Gifu University, Gifu City, Gifu, Japan

²Division of Anaerobe Research, Life Science Research Center, Gifu University, Gifu City, Gifu, Japan

³Gifu University Center for Conservation of Microbial Genetic Resource, Gifu, Japan

⁴Division of Clinical Laboratory, Gifu University Hospital, Gifu, Japan

⁵United Graduate School of Drug Discovery and Medical Information Sciences, Gifu University, Gifu City, Gifu, Japan

⁶Center for One Medicine Innovative Translational Research (COMIT), Gifu University, Gifu, Japan

AUTHOR ORCIDS

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

AUTHOR CONTRIBUTIONS

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

DATA AVAILABILITY

The genome sequence of *Varibaculum* sp. (GTC20041) has been deposited into DDBJ (<https://www.ddbj.nig.ac.jp/index.html>) under the accession number [AP041144](#). Raw sequence data for GTC20041 have been deposited into the DDBJ/Sequence Read Archive under the accession numbers [DRR683173](#) and [DRR683174](#).

REFERENCES

- Gibson GR, Macfarlane GT, Cummings JH. 1988. Occurrence of sulphate-reducing bacteria in human faeces and the relationship of dissimilatory sulphate reduction to methanogenesis in the large gut. *J Appl Bacteriol* 65:103–111. <https://doi.org/10.1111/j.1365-2672.1988.tb01498.x>
- Singh SB, Carroll-Portillo A, Lin HC. 2023. *Desulfovibrio* in the gut: the enemy within? *Microorganisms* 11:1772. <https://doi.org/10.3390/microorganisms11071772>
- Sydenham TV, Overballe-Petersen S, Hasman H, Wexler H, Kemp M, Justesen US. 2019. Complete hybrid genome assembly of clinical multidrug-resistant *Bacteroides fragilis* isolates enables comprehensive identification of antimicrobial-resistance genes and plasmids. *Microb Genom* 5:e000312. <https://doi.org/10.1099/mgen.0.000312>
- Hayashi M, Niwa A, Yonetamari J, Muto Y, Yokoyama S, Nakamura M, Yokobori Y, Ogawa M, Ichioka R, Kikuchi R, Okura H, Ogura S, Tetsuka N, Baba H, Tanaka K. 2024. Complete genome sequence of *Pigmentibacter ruber* isolated from a human patient in Japan. *Microbiol Resour Announc* 13:e0013324. <https://doi.org/10.1128/mra.00133-24>
- 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>
- 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>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- 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>
- 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>
- Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8:12–24. <https://doi.org/10.1039/C5AY02550H>
- 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>
- Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11:e0163962. <https://doi.org/10.1371/journal.pone.0163962>
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
- 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/bty2848>