

Complete genome sequence of *Eubacterium brachy* GTC20104 isolated from a human pleural effusion

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

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT The genus *Eubacterium* comprises gram-positive anaerobic bacteria. It has long been recognized as a component of human oral and gut microbiota. Here, we report the complete genome sequence of a novel strain, *Eubacterium brachy* GTC 20104, isolated from a human pleural effusion, consisting of a 1,593,027-bp circular chromosome.

KEYWORDS *Eubacterium*, anaerobe, genome

Eubacterium brachy is an anaerobic coccobacillus frequently isolated from human oral lesions (1). Its pathogenic potential has been suggested by its high prevalence of 35–42% in subgingival periodontitis samples (1). *E. brachy* has been reported in other body sites, including empyema fluid (2). Strain GTC20104 was isolated from a human pleural effusion sample in 2020. Isolation and maintenance were performed using Brucella HK agar plates (Kyokuto Pharmaceutical Industrial, Tokyo, Japan) supplemented with 5% sheep blood and incubated at 37°C for 48 h under anaerobic conditions (82% N₂, 10% CO₂, and 8% H₂). This study was conducted following the principles of the Declaration of Helsinki. Identification using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Bruker, Bremen, Germany) was unsuccessful.

The strain was subcultured under the same conditions as the initial isolation, and a single colony was cultured in Brucella broth for DNA extraction. Genomic DNA was extracted from the resulting cell pellet using Quick Taq HS DyeMix (TOYOBO Co., Ltd., Osaka, Japan). The entire isolate genome was sequenced using both long- (Oxford Nanopore Technologies, Tokyo, Japan) and short-read sequencing (Illumina Inc., San Diego, CA, USA) (3–5). For long-read sequencing, a library was prepared using a ligation sequencing kit (SQK-NBD114-24; Oxford Nanopore Technologies) without DNA shearing. The sequencing run generated 4,951,692 raw reads, with an N50 read length of 4,877 bp (Table 1). Small DNA fragments were removed (Short Read Eliminator XS; Circulomics, Baltimore, MD, USA), and sequencing was performed using a PromethION 2i instrument with a FLO-PRO114M flow cell. Long-read data were base-called in super-accuracy mode using the default base-calling pipeline in MinKNOW version 23.11.8 (Oxford Nanopore Technologies). The raw long reads were trimmed and quality-filtered using NanoFilt v.2.7.1 (6) with the parameters “-q 10 -headcrop 50 -l 1000.” An Illumina DNA Prep (M) Tagmentation Kit (Illumina) was used for library construction for short-read sequencing, followed by 2 × 151 bp paired-end sequencing on a NovaSeq 6000 platform (Illumina). Raw short reads were processed using fastp v.0.20.1 (3) with “-q 30 -n 20 -t 1 -T 1” parameters. Short-read quality was assessed using fastp, and mean long-read quality was assessed using NanoPlot v.1.32.1 (6). High-quality short reads (>87% bases >Q30 on average) and long reads (mean read quality of 16.7) were assembled using Unicycler v.0.4.8 with the default settings (7). Assembly circularization and rotation were performed automatically using the Unicycler to start with dnaA on the forward strand. The assembled contig graph was visualized and assessed for integrity using BlobTools v.1.0.

Editor Zhenjiang Zech Xu, Nanchang University, Nanchang, Jiangxi, China

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

The authors declare no conflict of interest.

Received 6 July 2025

Accepted 17 August 2025

Published 12 September 2025

Copyright © 2025 Yonetamari et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

TABLE 1 Information regarding the complete genome sequence of *Eubacterium brachy* GTC20104, isolated from a human pleural effusion

Parameter	Strain name <i>Eubacterium brachy</i> GTC20104
Illumina sequencing ^a	
No. of reads	4,951,692
Size (kb)	740,355
Avg coverage (x)	465
DRA accession no.	DRX640103
ONT sequencing ^a	
No. of reads	395,990
Size (kb)	1,336,347
Average read length (bp)	3,374.7
Average coverage (x)	839
N50	4,877
DRA accession no.	DRX640102
Assembly	
Estimated genome completeness (%) ^b	100.0
Estimated genome contamination (%) ^b	0.0
Genome structure	One chromosome
DDBJ/GenBank accession no.	AP040261
Genome size (bp)	1,593,027
GC content (%)	38.0
(chromosome/ plasmid name)	
No. of coding sequences ^c	1,399
Number of rRNAs ^c	6
Number of tRNAs ^c	37
Number of CRISPRs ^c	2

^aDRA, DNA DataBank of Japan (DDBJ) sequence read archive.^bDetermined using CheckM.^cDFAST, DDBJ fast annotation and submission tool.

Genome annotation was performed using DFAST ([DDBJ Fast Annotation and Submission Tool](#)) with default settings to predict coding sequences, rRNAs, tRNAs, and clustered regularly interspaced short palindromic repeats (CRISPR) elements.

The taxonomic identification of strain GTC20104 was performed using the [Type Strain Genome Server TYGS](#) (8). Genome- and rRNA-based comparisons supported *E. brachy* ATCC 33089 as the most closely related strain. The digital DNA–DNA hybridization value (88.6%) exceeded the species delineation threshold, supporting the classification of strain GTC20104 as *E. brachy*. This whole-genome sequence adds valuable insight to the limited genomic data available for the genus *Eubacterium*.

ACKNOWLEDGMENTS

We thank Kyoko Hatazaki, Akiko Katano, and Ayako Nagasawa for technical support. The authors declare no competing financial interests. This study was approved by the institutional ethics committee.

AUTHOR AFFILIATIONS

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

²Division of Clinical Laboratory, Gifu University Hospital, Gifu, Japan

³Division of Anaerobe Research, Life Science Research Center, Gifu University, Gifu City, Gifu, Japan

⁴Institute for Glyco-core Research iGCORE, Gifu University, Gifu City, Gifu, Japan

⁵Gifu University Center for Conservation of Microbial Genetic Resource, 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, Project administration, Resources, Supervision, Writing – review and editing | Yoshinori Muto, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – review and editing | Kaori Tanaka, Conceptualization, Project administration, Supervision, Writing – review and editing.

DATA AVAILABILITY

The complete genome sequence of *Eubacterium* sp. GTC20104, isolated from human pleural effusion, was deposited in the DNA Data Bank of Japan (DDBJ; <https://www.ddbj.nig.ac.jp/index.html>) under the accession number [AP040261](#). Raw sequence data for GTC20104 have been deposited in the DDBJ Sequence Read Archive under accession numbers [DRX640102](#) and [DRX640103](#).

REFERENCES

1. Kim JY, Kang B, Oh S, Gil Y, Choi IG, Chang IS. 2023. Genome-based reclassification of strain KIST612, previously classified as *Eubacterium limosum*, into a new strain of *Eubacterium callanderi* J Microbiol Biotechnol 33:1084–1090. <https://doi.org/10.4014/jmb.2304.04011>
2. Yamakawa H, Hagiwara E, Hayashi M, Katano T, Isomoto K, Otoshi R, Shintani R, Ikeda S, Tanaka K, Ogura T. 2017. A case of relapsed lung abscess caused by *Eubacterium brachy* infection following an initial diagnosis of pulmonary actinomycosis. Respir Med Case Rep 22:171–174. <https://doi.org/10.1016/j.rmcr.2017.08.013>
3. 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>
4. Vu H, Muto Y, Hayashi M, Noguchi H, Tanaka K, Yamamoto Y. 2022. Complete genome sequences of three *Phocaeicola vulgatus* strains isolated from a healthy Japanese individual. Microbiol Resour Announc 11:e0112421. <https://doi.org/10.1128/mra.01124-21>
5. 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>
6. 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>
7. 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>
8. 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>