

Complete genome sequence of two *Hominenteromicrobium mulieris* strains CIP 112194 and CIP 112527, isolated from a human fecal sample

Gérald Touak,¹ Damien Rei,² Dominique Clermont,¹ Christiane Bouchier¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT The human gut microbiota is dominated by obligately anaerobic bacteria. A method based on species-targeted sorting using flow cytometry under anaerobic conditions has been performed to analyze the gut microbiome. This report describes the complete genome sequences of two *Hominenteromicrobium mulieris* strains CIP 112194 and CIP 112527 that were isolated from the feces of a healthy adult from France. This volunteer is a young adult having a vegetarian diet.

KEYWORDS gut microbiome, anaerobes, genomes

To study microbial communities, Afrizal et al. (1) developed an application of single-cell dispensing, which allows label-free sorting of microscopic particles. Based on this approach, 82 human gut bacterial strains have been characterized, including novel species such as *Hominenteromicrobium mulieris* (1). We report here the complete genome sequences of two strains of *H. mulieris* CIP 112194 and CIP 112527 isolated from human feces belonging to the ancillary cohort “COSIMMGEN J” initiated for the study and approved by the COMITE DE PROTECTION DES PERSONNES Ile de France 1 - DOSSIER: 2018-fév. -14819. These strains were isolated from a healthy volunteer’s fecal sample, targeted by two polyclonal antibodies against two *Faecalibacterium prausnitzii* phylogroups and sorted using flow cytometry under anaerobic conditions (2). They were grown in pre-reduced tryptone, peptone, glucose yeast extract medium (supplemented with 0.1% l-cysteine-HCl and vitamins) (3) 48 h at 37°C and under strict anaerobic conditions (5% H₂/5% CO₂/90% N₂). Further analysis using 16S rRNA gene sequencing assigned the isolates to *Hominenteromicrobium* genus, which belongs to the *Oscillospira*-ceae family along with *F. prausnitzii*.

Genomic DNA used for Illumina and Nanopore sequencing was extracted from an 8 mL overnight culture of the strains CIP 112194 and CIP 112527. Genomic-tip 100G kit (Qiagen) and Nanobind CBB kit (Pacific Biosciences) were used for CIP 112194 and CIP 112527, respectively. Illumina sequencing was performed by the Mutualized Platform for Microbiology (Institut Pasteur, Paris, France) following their standard workflow for library preparation (Nextera tagmentation kit, Illumina) and an Illumina NextSeq 550 device using a 2 × 150 bp protocol. Paired-end reads were trimmed using fqCleanerER v23.12 workflow with a Phred quality score of 25 and read length ≥ 100 bases (4). The same gDNA aliquots were also used for long-read sequencing using the ligation sequencing gDNA kit (SQK-NBD114.24, Oxford Nanopore Technologies). gDNAs were not sheared or size selected to perform libraries. Libraries were sequenced on a MinION Mk1C device for 48 h (Oxford Nanopore Technologies, UK) with R10.4 flow cell (FLO-MIN114, Oxford Nanopore Technologies). The Dorado v0.3.0 tool was used to perform base calling. A command-line tool Hybracter hybrid v0.7.3 has been used for *de novo* assembly method (5). Long reads are filtered and subsampled using Filtlong v0.2.1

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

Address correspondence to Christiane Bouchier, christianebouchier@pasteur.fr.

The authors declare no conflict of interest.

Received 28 November 2024

Accepted 17 March 2025

Published 7 April 2025

Copyright © 2025 Touak 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 Summary of DNA sequence assembly data

Feature	Data per strain	
	CIP 112194	CIP 112527
Sequencing statistics		
Number of filtered ONT reads	41,843	11,389
Filtered ONT read N50	8,351	15,403
Genome statistics		
Genome size (bp)	3,146,805	3,184,959
GC content (%)	52.4	52.41
No. of tRNAs	57	57
No. of rRNAs	6	6
Average coverage (x) ^a	108	54
Average coverage (x) ^b	284	461
No. of CDSs ^c	3,030	3,088
Accession no.	CP173189	CP173188
Sequence read archive accession		
Filtered ONT reads	SRR31246135	SRR31246133
Illumina paired-end reads	SRR31246134	SRR31246132

^aCoverage after subsampling of length and quality-filtered ONT reads using Filtlong (see text for details).

^bCoverage after polishing with Illumina paired-end reads (see text for details).

^cCDSs, coding sequences.

with parameter “--min_quality 10 min_length 10000” for CIP 112527 and “--min_quality 10 min_length 5000” for CIP 112194. This retained 11,389 reads (175,875,634 bp with an N50 of 15,403 bp) and 41,843 reads (345,761,802 bp with an N50 of 8,351 bp) for CIP 112527 and CIP 112194, respectively. Genome assemblies were generated using Flye v2.9 yielding a single contig, marked as circular by Flye, for each strain. The contigs were finally polished using Illumina-trimmed paired-end reads with pypolca v0.3.1 for CIP 112194 and polypolish v0.6.0 for CIP 112527 (6). The contigs were reoriented to begin by the dnaA chromosomal replication initiator gene using dnaapler v0.7.0 (7).

Both genome assemblies consist of a single circular chromosome with a length of ~3.1 Mbp and a G+C content of 52.4% (Table 1). More information is shown in Table 1. A pairwise genome comparison between the draft genome of *H. mulieris* type strain CLA-AAH250 (GCF_020687165) and genome sequences of CIP 112194 and CIP 112527 reported an average nucleotide identity based on MUMmer of 98.88% (with 85.37% aligned nucleotides) and 98.98% (with 84.03% aligned nucleotides), respectively, (8) indicating that the two strains belong to *Hominenteromicrobium mulieris* species.

The annotation of the genomes was performed using the NCBI Prokaryotic Genome Annotation Pipeline (version 6.7, March 2024).

ACKNOWLEDGMENTS

The authors acknowledge Samuel Bellais of BIOASTER (Paris) for the targeted cell sorting under anaerobic conditions and Rémy Artus of ICAREb-biobank (Institut Pasteur, Paris) for his help in obtaining fecal samples.

AUTHOR AFFILIATIONS

¹CIP, CRBIP, DT, Microbiology Department, Institut Pasteur, Paris, France

²NBX Biosciences, SAS, Centre ADVISOR, Montpellier, France

AUTHOR ORCIDs

Christiane Bouchier  <http://orcid.org/0000-0003-3549-1543>

AUTHOR CONTRIBUTIONS

Gérald Touak, Methodology, Validation, Visualization | Damien Rei, Conceptualization, Funding acquisition, Visualization | Dominique Clermont, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing | Christiane Bouchier, Formal analysis, Methodology, Validation, Writing – original draft

DATA AVAILABILITY

The collection of volunteer's samples has been reviewed by the COMITE DE PROTECTION DES PERSONNES Ile de France 1 - N°IRB / IORG #: IORG0009918 - N° ID-RCB: 0-A01353-36. The genome sequences were deposited at GenBank under the accession numbers [CP173188](#) and [CP173189](#).

The Illumina paired-end reads can be found at SRA accession numbers [SRR31246134](#) and [SRR31246132](#). The ONT-filtered reads from the MinION MK1C run can be found at SRA accession numbers [SRR31246135](#) and [SRR31246133](#).

REFERENCES

1. Afrizal A, Hitch TCA, Viehof A, Treichel N, Riedel T, Abt B, Buhl EM, Kohlheyer D, Overmann J, Clavel T. 2022. Anaerobic single-cell dispensing facilitates the cultivation of human gut bacteria. *Environ Microbiol* 24:3861–3881. <https://doi.org/10.1111/1462-2920.15935>
2. Bellais S, Nehlich M, Ania M, Duquenoy A, Mazier W, van den Engh G, Baijer J, Treichel NS, Clavel T, Belotserkovsky I, Thomas V. 2022. Species-targeted sorting and cultivation of commensal bacteria from the gut microbiome using flow cytometry under anaerobic conditions. *Microbiome* 10:24. <https://doi.org/10.1186/s40168-021-01206-7>
3. Le Gratiet T, Poezevara T, Rouxel S, Houard E, Mazuet C, Chemaly M, Maréchal CL. 2020. Development of an innovative and quick method for the isolation of *Clostridium botulinum* strains involved in avian botulism outbreaks. *Toxins (Basel)* 12:42. <https://doi.org/10.3390/toxins12010042>
4. Alexis CRISCUOLO. 2021. fqCleanER. Available from: <https://gitlab.pasteur.fr/GIPhy/fqCleanER>
5. Bouras G, Houtak G, Wick RR, Mallawaarachchi V, Roach MJ, Papudeshi B, Judd LM, Sheppard AE, Edwards RA, Vreugde S. 2024. Hybracter: enabling scalable, automated, complete and accurate bacterial genome assemblies. *Microb Genom* 10:001244. <https://doi.org/10.1099/mgen.0.001244>
6. Bouras G, Judd LM, Edwards RA, Vreugde S, Stinear TP, Wick RR. 2024. How low can you go? Short-read polishing of Oxford Nanopore bacterial genome assemblies. *Microb Genom* 10:001254. <https://doi.org/10.1099/mgen.0.001254>
7. Bouras G, Grigson SR, Papudeshi B, Mallawaarachchi V, Roach MJ. 2024. Dnaapler: a tool to reorient circular microbial genomes. *JOSS* 9:5968. <https://doi.org/10.21105/joss.05968>
8. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32:929–931. <https://doi.org/10.1093/bioinformatics/btv681>