



# Whole-Genome Sequencing and Annotation of *Clostridium tyrobutyricum* Strain Cirm BIA 2237, Isolated from Silage

Edouard Munier,<sup>a,b</sup> Hélène Licandro-Séraut,<sup>b</sup> Christine Achilleos,<sup>a</sup> Rémy Cachon,<sup>b</sup> Eric Beuvier<sup>a</sup>

<sup>a</sup>URTAL, INRA, Poligny, France

<sup>b</sup>UMR Procédés Alimentaires et Microbiologiques, Université Bourgogne Franche-Comté, AgroSup Dijon, PAM UMR A 02.102, Dijon, France

**ABSTRACT** *Clostridium tyrobutyricum* is the main bacterial species leading to the late blowing defect, a major cause of spoilage in semihard and hard cheeses. This study reports the complete genome sequencing, assembly, and annotation of *C. tyrobutyricum* strain Cirm BIA 2237, formerly called CNRZ 608, isolated from silage.

The strict anaerobic endospore-forming bacterium *Clostridium tyrobutyricum* is known to be the leading cause of the late blowing defect in semihard and hard cheeses (1, 2), which results in huge losses of products and profits in the cheese industry. Spores of *C. tyrobutyricum* contaminate the milk during milking and survive throughout the pasteurization process. Butyric fermentation due to *C. tyrobutyricum* produces butyric acid, hydrogen, and carbon dioxide, resulting in characteristic defects in cheeses, such as late blowing and altered flavor. This metabolic pathway remains poorly understood for *C. tyrobutyricum*. To date, the genome sequence of the type strain ATCC 25755 has been sequenced and annotated (3), but its role in the late blowing defect has not been studied. Some projects of *C. tyrobutyricum* draft genome sequencing from the late blowing defect in cheese have been initiated (4, 5).

We report the whole-genome sequencing, assembly, and annotation of *C. tyrobutyricum* strain Cirm BIA 2237 (formerly CNRZ 608 [6]), belonging to the French National Institute for Agricultural Research collection. This strain was isolated from silage and is known to cause the late blowing defect in semihard and hard cheeses. It was also used as a reference strain in many studies (7–14).

*C. tyrobutyricum* was grown in reinforced clostridial medium (RCM) (Biokar, France) at 37°C in an anaerobic chamber. The genomic DNA isolation and sequencing were performed by Genoscreen, France. Extraction of the genomic DNA was performed using the Qiagen Gentra Puregene kit, and the DNA was quantified by fluorescence using the Qubit double-stranded DNA broad-range (dsDNA BR) kit. The sequencing libraries were prepared with the Illumina Nextera XT sample prep kit and sequenced using an Illumina HiSeq 2 × 100-bp sequencing platform. A total of 10,595,027 raw reads were obtained. Raw data quality analysis was performed with FastQC v0.11.5 (15), and elimination of sequences containing undetermined or small-sized (<60 bp) nucleotides or a Phred score under 30 was performed with Prinseq v0.20.4 (16). The *de novo* assembly, which consists of 93 scaffolds, was performed using SPAdes v3.10.1 (k-mer sizes of 21, 33, 55, and 77) (17). The assembly was verified by calculating the reads mapped back to contigs (RMBC) index using Bowtie v2.1.0 with stringent mapping parameters (18) and by evaluating the depth of scaffold sequencing using SAMtools v1.5 (19). Scaffolds with a value of less than 200 bp were removed. Annotation was performed using Prokka v1.11 (20) with the hidden Markov model (HMM) (21) and UniProt (22) databases by predicting open reading frames (ORFs) and HMMER v3.1b2 (23) with the eggNOG v4.5 (24) database for a supplementary annotation from ORFs.

The strain Cirm BIA 2237 genome is composed of a chromosome of 3.16 Mb with a

**Citation** Munier E, Licandro-Séraut H, Achilleos C, Cachon R, Beuvier E. 2019. Whole-genome sequencing and annotation of *Clostridium tyrobutyricum* strain Cirm BIA 2237, isolated from silage. *Microbiol Resour Announc* 8:e00492-19. <https://doi.org/10.1128/MRA.00492-19>.

**Editor** Vincent Bruno, University of Maryland School of Medicine

**Copyright** © 2019 Munier 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/).

Address correspondence to Edouard Munier, [edouard.munier@inra.fr](mailto:edouard.munier@inra.fr).

**Received** 2 May 2019

**Accepted** 9 July 2019

**Published** 25 July 2019

GC content of 30.75% and does not contain plasmids. It comprises 3,182 coding sequences.

The genome analysis showed that this strain is composed of genes encoding the required enzymes involved in butyric fermentation from lactate and acetate. The genome annotation will ease studies of the mechanisms causing the defect of late blowing during cheese ripening.

**Data availability.** The complete genome sequence of *C. tyrobutyricum* strain Cirm BIA 2237 has been deposited in DDBJ/EMBL/GenBank (25) under the accession number [CP038158](https://doi.org/10.1093/nar/gkx1195), and the raw reads have been deposited in the NCBI Sequence Read Archive under the accession number [PRJNA524264](https://doi.org/10.1093/nar/gkx1195).

## ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## REFERENCES

- Klijn N, Nieuwenhof FF, Hoolwerf JD, Van Der Waals CB, Weerkamp AH. 1995. Identification of *Clostridium tyrobutyricum* as the causative agent of late blowing in cheese by species-specific PCR amplification. *Appl Environ Microbiol* 61:2919–2924.
- Brändle J, Domig KJ, Kneifel W. 2016. Relevance and analysis of butyric acid producing clostridia in milk and cheese. *Food Control* 67:96–113. <https://doi.org/10.1016/j.foodcont.2016.02.038>.
- Jiang L, Zhu L, Xu X, Li Y, Li S, Huang H. 2013. Genome sequence of *Clostridium tyrobutyricum* ATCC 25755, a butyric acid-overproducing strain. *Genome Announc* 1:e00308-13. <https://doi.org/10.1128/genomeA.00308-13>.
- Bassi D, Fontana C, Gazzola S, Pietta E, Puglisi E, Cappa F, Cocconcetti PS. 2013. Draft genome sequence of *Clostridium tyrobutyricum* strain UC7086, isolated from Grana Padano cheese with late-blowing defect. *Genome Announc* 1:e00614-13. <https://doi.org/10.1128/genomeA.00614-13>.
- Storari M, Wüthrich D, Bruggmann R, Berthoud H, Arias-Roth E. 2015. Draft genome sequences of *Clostridium tyrobutyricum* strains FAM22552 and FAM22553, isolated from Swiss semihard red-smear cheese. *Genome Announc* 3:e00078-15. <https://doi.org/10.1128/genomeA.00078-15>.
- Roux C, Bergere JL. 1977. Taxonomic characters of "*Clostridium tyrobutyricum*." *Ann Microbiol (Paris)* 128A:267–276. (In French.)
- Bourgeois CM, Le Parc O, Abgrall B, Cleret J-J. 1984. Membrane filtration of milk for counting spores of *Clostridium tyrobutyricum*. *J Dairy Sci* 67:2493–2499. [https://doi.org/10.3168/jds.S0022-0302\(84\)81605-7](https://doi.org/10.3168/jds.S0022-0302(84)81605-7).
- Bergère J-L, Hayes H, Commissaire J. 1986. Major protein components in the cell envelope of *Clostridium tyrobutyricum*. *Ann Inst Pasteur Microbiol* 137:271–282. [https://doi.org/10.1016/S0769-2609\(86\)80117-X](https://doi.org/10.1016/S0769-2609(86)80117-X).
- Bergère J-L, Vassal L, Pitel M, Cathelin M. 1978. Bactéries lactiques susceptibles d'inhiber la croissance de *Clostridium tyrobutyricum* en culture et dans le fromage. *Lait* 58:215–233. <https://doi.org/10.1051/lait:1978575-57613>.
- Mathot AG, Beliard E, Thuault D. 2003. *Streptococcus thermophilus* 580 produces a bacteriocin potentially suitable for inhibition of *Clostridium tyrobutyricum* in hard cheese. *J Dairy Sci* 86:3068–3074. [https://doi.org/10.3168/jds.S0022-0302\(03\)73906-X](https://doi.org/10.3168/jds.S0022-0302(03)73906-X).
- Thuault D, Beliard E, Le Guern J, Bourgeois C-M. 1991. Inhibition of *Clostridium tyrobutyricum* by bacteriocin-like substances produced by lactic acid bacteria. *J Dairy Sci* 74:1145–1150. [https://doi.org/10.3168/jds.S0022-0302\(91\)78266-0](https://doi.org/10.3168/jds.S0022-0302(91)78266-0).
- Abgrall B, Bourgeois C, Bourva F. 1985. Dénombrement des spores de *Clostridium tyrobutyricum* par filtration sur membrane et culture sur milieu gélosé. *Lait* 65:45–53. <https://doi.org/10.1051/lait:1985647-6484>.
- Le Bourhis A-G, Doré J, Carlier J-P, Chamba J-F, Popoff M-R, Tholozan J-L. 2007. Contribution of *C. beijerinckii* and *C. sporogenes* in association with *C. tyrobutyricum* to the butyric fermentation in Emmental type cheese. *Int J Food Microbiol* 113:154–163. <https://doi.org/10.1016/j.ijfoodmicro.2006.06.027>.
- Bergère J-L. 1985. Mise en évidence de deux groupes sérologiques principaux chez *Clostridium tyrobutyricum*. *Ann Inst Pasteur Microbiol* 136:75–91. [https://doi.org/10.1016/S0769-2609\(85\)80008-9](https://doi.org/10.1016/S0769-2609(85)80008-9).
- Brown J, Pirrung M, McCue LA. 2017. FQC Dashboard: integrates FastQC results into a Web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* 33:3137–3139. <https://doi.org/10.1093/bioinformatics/btx373>.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Bateman A, Haft DH. 2002. HMM-based databases in InterPro. *Brief Bioinform* 3:236–245. <https://doi.org/10.1093/bib/3.3.236>.
- The UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515. <https://doi.org/10.1093/nar/gky1049>.
- Prakash A, Jeffryes M, Bateman A, Finn RD. 2017. The HMMER Web server for protein sequence similarity search. *Curr Protoc Bioinformatics* 60:3.15.1–3.15.23. <https://doi.org/10.1002/cpbi.40>.
- Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C, Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol Biol Evol* 34:2115–2122. <https://doi.org/10.1093/molbev/msx148>.
- Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2012. GenBank. *Nucleic Acids Res* 41:D36–D42. <https://doi.org/10.1093/nar/gks1195>.