Genome sequence and description of *Clostridium niameyense* sp. nov., isolated from a human with marasmus in Nigeria

S. Chabou, M. Tidjani Alou, S. Khelaifia, J.-C. Lagier, N. Labas, T. Cimmino, S. Diene, D. Raoult and J.-M. Rolain

Aix-Marseille Université, IRD, APHM, MEPHI, IHU Méditerranée Infection, Marseille, France

Abstract

Clostridium niameyense sp. nov. strain MT5 is the type strain of *C. niameyense* sp. nov., a new species within the genus *Clostridia. C. niameyense* is a Gram-positive, anaerobic bacillus. The strain MT5 (= CSUR P1468 = DSMZ 100441), whose genome is described here, was isolated from a faecal sample collected from a patient with anorexia and marasmus living in Nigeria. The genome is 2 542 841 bp long with 27.44% G + C content and consists of six scaffolds.

Keywords: *Clostridium niameyense*, culturomics, genome, marasmus and anorexia, taxonogenomics

Original Submission: 27 December 2017; **Revised Submission:** 22 February 2018; **Accepted:** 1 March 2018

Article published online: 14 March 2018

Corresponding author. J.-M. Rolain, IRD, APHM, MEPHI, IHU Méditerranée Infection, Faculté de médecine, 19-21 Boulevard Jean Moulin, 13385, Marseille cedex 5, France.

E-mail: jean-marc.rolain@univ-amu.fr

Introduction

Clostridium niameyense strain MT5 (= CSUR P1468 = DSMZ 100441) is the type strain of *C. niameyense* sp. nov. This bacterium is a Gram-positive, anaerobic, motile bacterium that was isolated from the stool of an anorexic patient with marasmus living in Nigeria as part of a culturomics study aiming at cultivating all species individually within human faeces [1,2].

The usual parameters used to delineate a bacterial species include 16S rDNA sequence identity and phylogeny [3], genomic G + C content diversity and DNA-DNA hybridization [3]. Recently a new approach has been proposed, taxonogenomics, which includes genomic data in a polyphasic taxonomy to describe new bacterial species [4]. This strategy is based on a combination of genomic and phenotypic characteristics, including matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) as well as spectrum and genomic analyses [5].

Clostridium is a genus of Gram-positive bacteria which includes several significant human pathogens and consists of obligate anaerobic bacteria [6]. These bacteria are the cause of some human diseases such as botulism (*C. botulinum*), tetanus (*C. tetani*), pseudomembranous colitis (*C. difficile*) and food poisoning–gas gangrene (*C. perfringens*) [7]. The pathogenic clostridial species have evolved several mechanisms to survive inside and outside a number of hosts, as evidenced by various diseases often linked to their protein toxins and spores [8].

Here we present a summary classification and set of features for *C. niameyense* sp. nov. strain MT5 together with the description of the complete genome sequence and annotation. These characteristics support the circumscription of the species *C. niameyense*.

Phenotypic and biochemical characterization

Clostridium niameyense sp. nov. strain MT5 was isolated from a faecal sample collected from an anorexic patient with marasmus living in Nigeria. The faecal specimen was preserved at −80 °C after collection and sent to Marseille. The MT5 strain was cultivated on 5% sheep's blood–enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) after 5 days' preincubation in a blood culture bottle with rumen fluid.

https://doi.org/10.1016/j.nmni.2018.03.001

For the growth of *C. niameyense,* we tested different temperatures (25, 30, 37 °C). However, optimal growth occurred at 37 °C, 24 hours after inoculation. Colonies were opaque with a smooth aspect, irregular and 1 mm in diameter on blood-enriched Columbia agar. Growth was observed only under anaerobic conditions, and no growth occurred under aerobic or microaerophilic conditions. Gram staining revealed a Gram-positive nonsporulating bacillus; the motility test was positive (Fig. 1).

The obtained colonies were identified using a Microflex spectrometer (Bruker Daltonics, Bremen, Germany). MALDI-TOF MS identification, measurement and analysis were performed as previously described [1]. No significant MALDI-TOF MS score was obtained for strain MT5 against the Bruker database, suggesting that our isolate was not a member of a known species. We added the spectrum from strain MT5 to our database (Fig. 2).

Biochemical features of the strains were determined using API50-CHL and the API ZYM (bioMérieux). Oxydase (Becton Dickinson, Pont-de-Claix, France) and catalase assays (bioMérieux) were performed separately. API ZYM presented positive reactions for α-glucosidase (maltase) and β-glucosidase (cellulase). The differential phenotypic characteristics with other *Clostridium* species are summarized in Table 1.

Antibiotic susceptibility of our isolate was assessed using the disc diffusion method on Muller-Hinton agar plates supplemented with 5% blood (bioMérieux) using the standard disc diffusion procedure as recommended by the European Committee on Antimicrobial Susceptibility Testing (EUCAST; http://www.eucast.org/). *C. niameyense* was likely resistant to penicillin (14 mm), amoxicillin (13 mm), trimethoprim/sulfamethoxazole (6 mm), colistin (6 mm) and metronidazole (6 mm) but susceptible to vancomycin (22 mm), ertapenem (20 mm), fosfomycin (27 mm), teicoplanin (19 mm), oxacillin (22 mm), cefotaxime (20 mm), rifampicin (25 mm), gentamycin (27 mm), doxycycline (25 mm), pristinamycin (22 mm), erythromycin (26 mm), lincomycin (27 mm), imipenem (22 mm), tazocillin (24 mm), ofloxacin (25 mm), meropenem (20 mm) and ciprofloxacin (19 mm).

16S rRNA gene sequencing and phylogenetic analyses

The organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA similarity to members of the genus *Clostridium* [9] and is part of a study of the human digestive flora aiming to isolate all bacterial species in human faeces [1]. A phylogenetic tree was constructed using the maximum-likelihood method in the MEGA6 software package (https://www.megasoftware.net/). Phylogenetic analysis of the complete sequence of the 16S rRNA gene of *C. niameyense* strain MT5 (= CSUR P1468 = DSMZ 100441) exhibited a 98.3% nucleotide sequence similarity with *Clostridium botulinum* A (L37585). This value was lower than the 98.7% 16S rRNA gene sequence threshold recommended by Stackebrandt and Ebers [10] to delineate a new species without carrying out DNA-DNA hybridization (Fig. 3).

Genome properties

The genome is 2 542 841 bp long with 27.44% G+C content (Table 2). It is composed of six scaffolds (composed of seven contigs). Of the 2489 predicted genes, 2410 were protein-coding genes and 79 were RNAs (five genes are 5S rRNA, one gene is 16S rRNA, two genes are 23S rRNA and 71 genes are tRNA genes). A total of 1790 genes (74.27%) were assigned

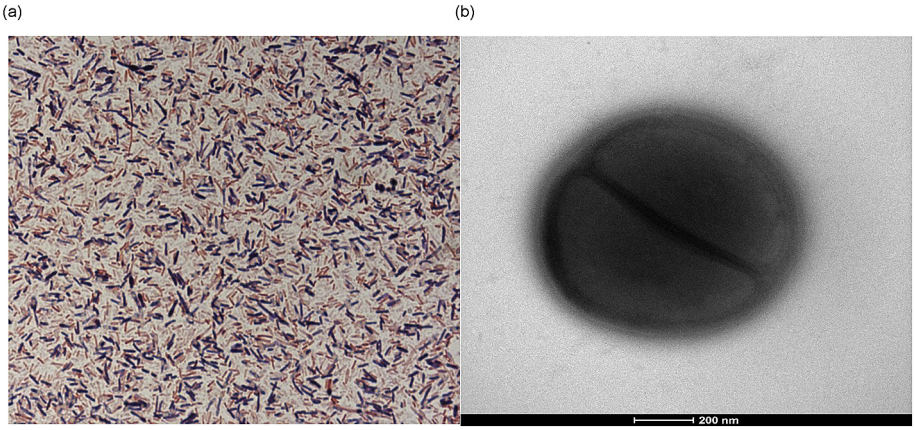

FIG. 1. (a) Gram staining of *Clostridium niameyense* strain MT sp. (b) Transmission electron microscopy of *Clostridium niameyense* sp. nov. strain MT5 using TechnaiG^2 Cryo device (FEI Company, Limeil-Brévannes, France) at operating voltage of 200 keV. Scale bar = 200 nm.

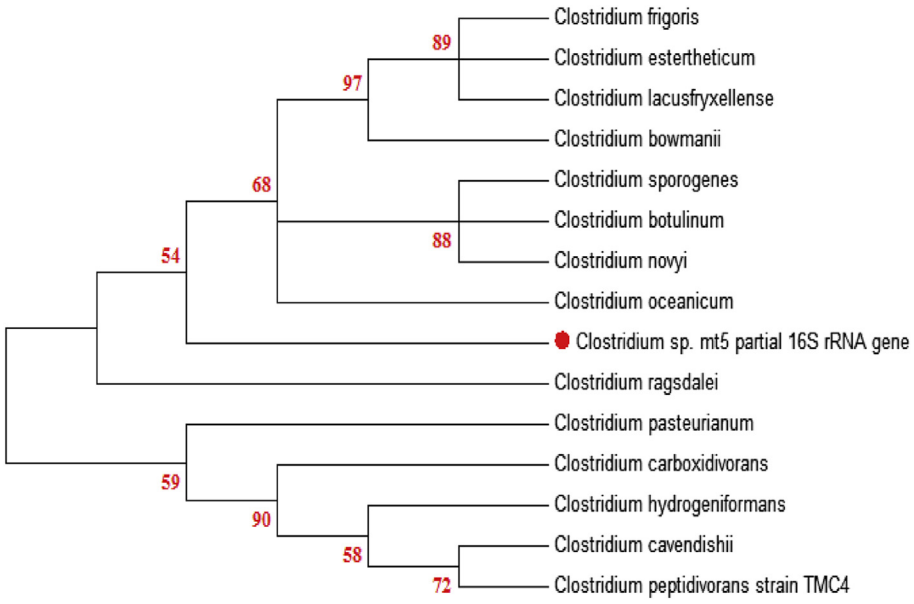

FIG. 2. Phylogenetic tree highlighting position of *Clostridium niameyense* sp. nov. strain MT5 relative to other type strains within *Clostridium* genus. GenBank accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALW and phylogenetic inferences obtained using maximum-likelihood method within MEGA software.

a putative function by Clusters of Orthologous Groups database (COGs) or by NR BLAST. Fifty-nine genes were identified as ORFans (2.45%). The remaining genes were annotated as hypothetical proteins (401 genes, >16.64%). The properties of the genome are summarized in Table 2. A genomic comparison of *C. niameyense* with seven other *Clostridium* species is provided in Table 3. The distribution of genes into COGs functional categories is presented in Table 4 (Fig. 4). The genome was assembled and annotated by Xegen (http://www.xegen.fr/).

Genome annotation

The ARG-ANNOT database for acquired antibiotic resistance genes (ARGs) was used for a BLAST search using the Bio-Edit interface. The assembled sequences were searched against the ARG database under moderately stringent conditions (*E* value of 10^{-5}) for the *in silico* ARG prediction [11]. The results obtained by ARG-ANNOT demonstrated that there are no known antibiotic-resistance genes in the genome of *C. niameyense* strain MT5.

TABLE 1. Differential characteristics of *Clostridium niameyense*, *C. beijerinckii* strain NCIMB 8052, *C. disporicum* NCIB12424, *C. carboxidivorans* strain P7, *C. dakarense* strain FF1, *C. difficile* strain B1 and *C. saudii* JCCT

Property	*C. niameyense*	*C. beijerinckii*	*C. disporicum*	*C. carboxidivorans*	*C. dakarense*	*C. difficile*	*C. saudii*
Cell diameter (μm)	1	1.7	1.5	1.5	1.2	3.0	1.0
Oxygen requirement	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic	Strictly anaerobic
Gram strain	positive	Variable	Positive	Positive	Positive	Positive	Positive
Motility	Motile	Motile	NA	Motile	Motile	Motile	Motile
Endospore training	NA	+	NA	+	+	+	+
Indole	−	NA	−	−	+	NA	−
Production of:							
Alkaline phosphatase	−	NA	NA	NA	+	NA	−
Catalase	−	−	−	−	−	NA	−
Oxidase	−	NA	NA	−	−	NA	−
Nitrate reduction	−	−	NA	−	−	−	−
Urease	−	−	NA	−	−	NA	−
β-Galactosidase	+	NA	NA	NA	−	NA	−
N-Acetyl-glucosamine	−	NA	NA	NA		NA	−
Acid from:							
L-Arabinose	−	+	NA	+	−	−	−
Ribose	−	−	+	+	−	−	−
Mannose	+	+	+	+	−	+	−
Mannitol	+	+	+	+	−	+	−
Sucrose	−	+	+	+	−	+	−
D-Glucose	+	+	+	+	+	NA	−
D-Fructose	+	+	+	+	−	+	−
D-Maltose	−	+	+	+	+	−	−
D-Lactose	−	+	+	+	−	−	−
G+C satisfied (%)	27.44	28	29	31	27.98	28	28
Habitat	Human gut	Human gut	Rat gut	Environment	Human gut	Human gut	Human gut

+, positive result; −, negative result; NA, data not available.

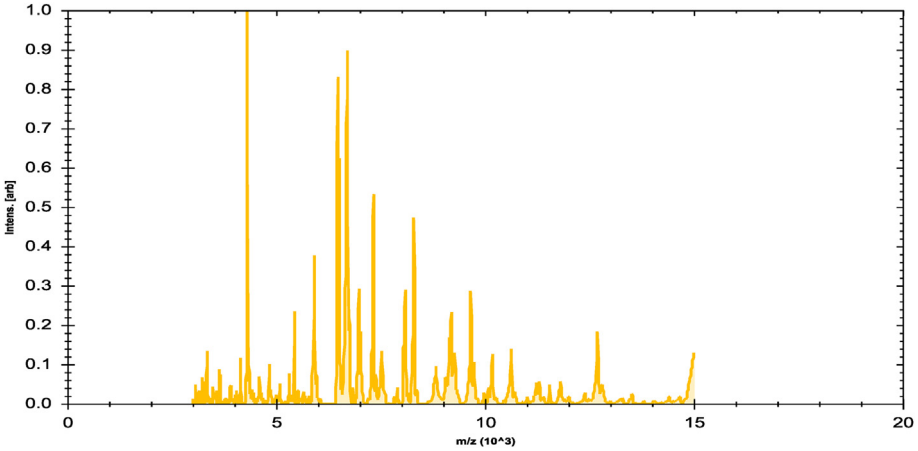

FIG. 3. Reference mass spectrum from *C. Clostridium niameyense* sp. nov. strain MT5. Spectra from 12 individual colonies were compared and reference spectrum generated.

TABLE 2. Nucleotide content and gene count levels of genome

Attribute	Value	% of total
Genome size (bp)	2 542 841	100
No. of G+C (bp)	697 772	27.44
Genes	2489	100
Protein genes	2410	96.82
RNA genes	79	3.17
tRNA genes	71	2.85
RNA (5S, 16S, 23S) genes	8	0.32
Coding sequence size	2 232 886	87.81
Coding sequence gene protein size	2 219 514	87.28
Coding sequence tRNA gene size	5558	0.21
Coding sequence (5S, 16S, 23S) genes size	7814	0.307
Protein-coding gene	2410	100

These sequences were submitted to Rapid Annotation Using Subsystem Technology (RAST) [12]. An exhaustive search of the bacteriocin database available in our laboratories (Bacteriocins of the URMITE, BUR; http://drissifatima.wix.com/bacteriocins) was performed by collecting all currently available sequences from the databases and from the National Center for Biotechnology Information. Protein sequences from this database allowed putative bacteriocins from human gut microbiota to be identified using BLASTp methodology [13].

TABLE 3. Genomic comparison of *C. niameyense* with seven other *Clostridium* species

Species	Strain	Genome accession number	Genome size (Mb)	G+C content
C. niameyense	MT5	CVPI00000000	2.54	27.4
C. perfringens	ATCC 13124	ATCC 13124	3.26	28.4
C. difficile	B1	NC_017179	4.46	28.4
C. dakarense	FF1	CBTZ010000000	3.73	27.9
C. leptum	DSM753	ABCB02000000	3.27	50.2
C. botulinum	ATCC 3502	NC_009495	3.90	28.2
C. beijerinckii	NCIMB 8052	NC_009617	6.0	29.0
C. senegalense	DSM 25507	CAEV01000001	3.89	26.8

The analysis of the resistome did not identify antibiotic resistance genes, and the absence of bacteriocin was confirmed.

Analysis of the presence of polyketide synthases and nonribosomal peptide synthetases was performed by discriminating the genes with large sizes using a database realized in our laboratory; predicted proteins were compared against the nonredundant GenBank database using BLASTp and finally examined using antiSMASH [14].

Analysis revealed five clusters of clusters of differentiation (CDs) (11 642 bp) predicted to code for enzymes involved in the nonribosomal synthesis of peptides (Fig. 5), which in turn

TABLE 4. Number of genes associated with 25 general COGs functional categories

Code	Value	% of total	Description
J	152	6.3070545	Translation
A	0	0	RNA processing and modification
K	138	5.726141	Transcription
L	99	4.107884	Replication, recombination and repair
B	1	0.041493777	Chromatin structure and dynamics
D	27	1.120332	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	53	2.19917	Defense mechanisms
T	97	4.0248966	Signal transduction mechanisms
M	81	3.360996	Cell wall/membrane biogenesis
N	60	2.4896266	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	37	1.5352697	Intracellular trafficking and secretion
O	56	2.3236516	Posttranslational modification, protein turnover, chaperones
C	95	3.9419088	Energy production and conversion
G	83	3.4439836	Carbohydrate transport and metabolism
E	138	5.726141	Amino acid transport and metabolism
F	60	2.4896266	Nucleotide transport and metabolism
H	75	3.1120331	Coenzyme transport and metabolism
I	44	1.8257263	Lipid transport and metabolism
P	87	3.6099586	Inorganic ion transport and metabolism
Q	24	0.9958507	Secondary metabolites biosynthesis, transport and catabolism
R	208	8.630706	General function prediction only
S	160	6.6390047	Function unknown
—	812	33.692947	Not in COGs

COGs, Clusters of Orthologous Groups database.

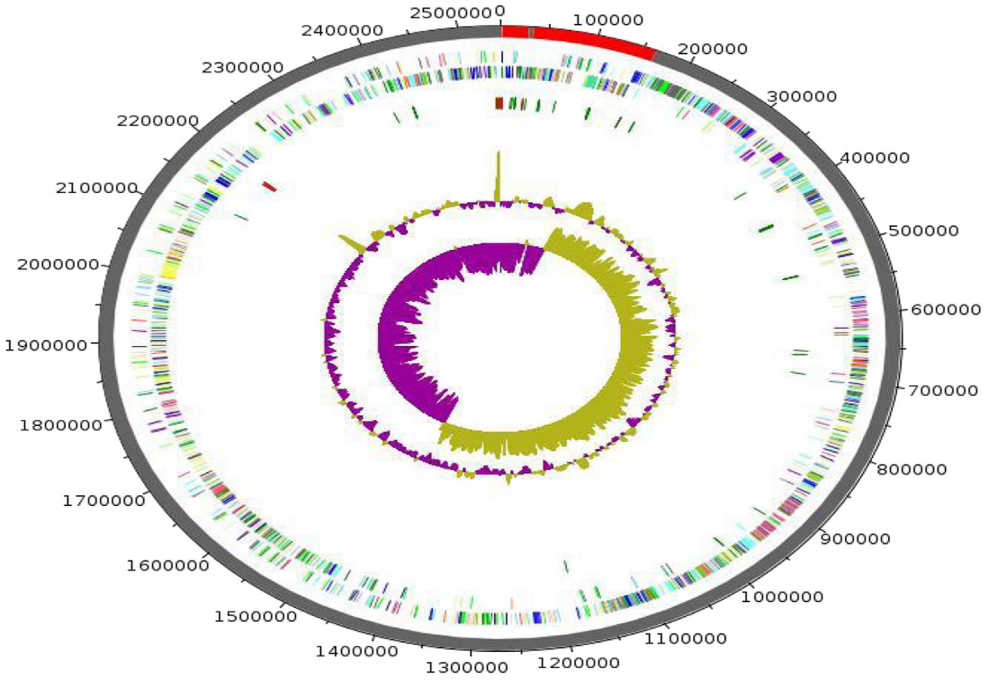

FIG. 4. Graphical circular map of chromosome. From outside to centre: genes on forward strand coloured by COGs categories (only genes assigned to COGs), genes on reverse strand coloured by COGs categories (only gene assigned to COGs), RNA genes (tRNAs green, rRNAs red), G+C content and G+C skew. COGs, Clusters of Orthologous Groups database.

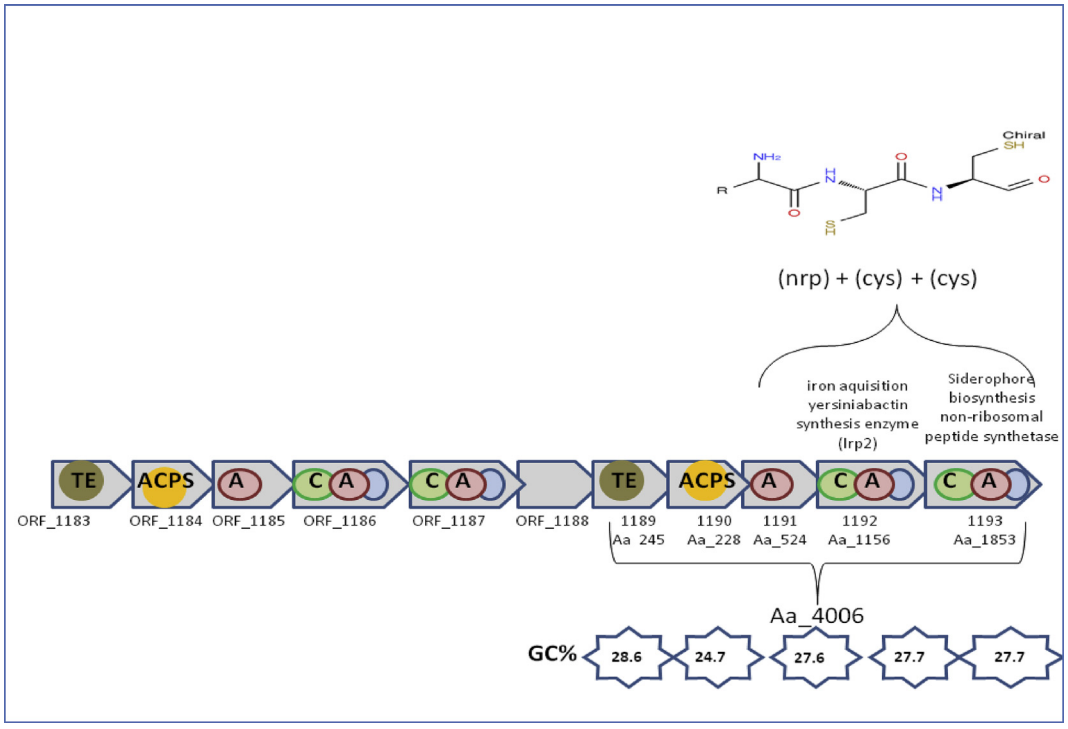

FIG. 5. CDS clusters predicted to encode siderophore biosynthesis nonribosomal peptide synthetase. A, C, ACPS and TE are domains of non-ribosomal peptide synthetases. A, AMP binding, adenylation domain; C, heterocyclization; TE, thioesterase domain.

were predicted to catalyze the synthesis of a siderophore (Fig. 5), which could be used under iron-deficient growth conditions.

This mechanism is already known in *Clostridium* spp., particularly in in *C. kluyveri* DSM555. The alignment of non-ribosomal synthesis in *C. kluyveri* DSM555 (WP_012101903.1) and *C. niameyense* MT5 (WP_050607297.1) showed a 31% similarity with 99% coverage.

Description of *Clostridium niameyense* sp. nov. strain MT5

Clostridium niameyense sp. nov. strain MT5 is the type strain of *C. niameyense* sp. nov., a new species within the genus *Clostridia*. The strain MT5 (= CSUR P1468 = DSMZ 100441), whose genome is described here, was isolated from a faecal sample collected from an anorexic patient with marasmus living in Nigeria. Optimal growth occurred at 37 °C, 24 hours after inoculation. The colonies were 0.1 to 0.3 mm in diameter on blood-enriched agar. *C. niameyense* is a Gram-positive, obligate anaerobic bacterium with a mean diameter of 1.3 µm.

C. niameyense has positive reactions for glycerol, erythritol, D-glucose, D-fructose, D-mannitol, D-saccharose, D-trehalose D-maltose, D-melezitose, amidone D-turanose and potassium 5-cétogluconate. Also, API ZYM presented positive reactions for α-glucosidase (maltase) and 17-β-glucosidase (cellulase).

C. niameyense is resistant to penicillin (14 mm), amoxicillin (13 mm), trimethoprim/sulfamethoxazole (6 mm), colistin (6 mm) and metronidazole (6 mm) but susceptible for another antibiotic.

The G+C content of the genome is 27.44%. The 16S rDNA and genome sequences are deposited in GenBank under accession numbers LN827532 and CVPI00000000, respectively. The type strain is MT5 (= CSUR P1468 = DSMZ 100441).

Acknowledgements

We thank L. Hadjadj (Aix-Marseille Univ., IRD, APHM, MEPHI, IHU Méditerranée Infection, Marseille, France, IHU Méditerranée Infection, Marseille, France.) for technical assistance. This project was funded by the French Centre National de la Recherche Scientifique (CNRS): UM63, Infectiopole Sud Fundation and IHU Méditerranée Infection.

Conflict of interest

None declared.

References

[1] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.

[2] Pfleiderer A, Lagier JC, Armougom F, Robert C, Vialettes B, Raoult D. Culturomics identified 11 new bacterial species from a single anorexia nervosa stool sample. Eur J Clin Microbiol Infect Dis 2013;32:1471–81.

[3] Tindall BJ, Rosselló-Móra R, Busse H-J, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2010;60:249–66.

[4] Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.

[5] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Clostridium senegalense* sp. nov. Stand Genomic Sci 2012;6:401–10.

[6] Angelakis E, Bibi F, Ramasamy D, Azhar EI, Jiman-Fatani AA, Aboushoushah SM, et al. Non-contiguous finished genome sequence and description of *Clostridium saudii* sp. nov. Stand Genomic Sci 2014;9:8.

[7] Wells CL, Wilkins TD. Clostridia: sporeforming anaerobic bacilli. In: Baron S, editor. Medical microbiology. 4th ed. Galveston, TX: University of Texas Medical Branch at Galveston; 1996.

[8] Krautz-Peterson G, Zhang Y, Chen K, Oyler GA, Feng H, Shoemaker CB. Retargeting *Clostridium difficile* toxin B to neuronal cells as a potential vehicle for cytosolic delivery of therapeutic biomolecules to treat botulism. J Toxicol 2012;2012:760142.

[9] Kalia VC, Mukherjee T, Bhushan A, Joshi J, Shankar P, Huma N. Analysis of the unexplored features of rrs (16S rDNA) of the genus *Clostridium*. BMC Genomics 2011; 2011;12:18.

[10] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.

[11] Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, et al. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58:212–20.

[12] Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics 2008;9:75.

[13] Drissi F, Buffet S, Raoult D, Merhej V. Common occurrence of antibacterial agents in human intestinal microbiota. Front Microbiol 2015;6:441.

[14] Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 2015;43:W237–43.