Massilioclostridium coli gen. nov., sp. nov., a new member of the Clostridiaceae family isolated from the left colon of a 27-year-old woman

C. Lo, M. Mailhe, D. Ricaboni, V. Vitton, A. Benezech, C. Michelle, N. Armstrong, F. Bittar, Pierre-Edouard Fournier, Didier Raoult, et al.

Massilioclostridium coli gen. nov., sp. nov., a new member of the *Clostridiaceae* family isolated from the left colon of a 27-year-old woman

C. I. Lo[1], M. Mailhe[2], D. Ricaboni[2,4], V. Vitton[3], A. Benezech[3], C. Michelle[2], N. Armstrong[2], F. Bittar[2], P.-E. Fournier[2], D. Raoult[2] and J.-C. Lagier[2]

1) *Campus commun IRD-UCAD de Hann et Hôpital Principal de Dakar, Senegal,* 2) *Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095,* 3) *Service de Gastroentérologie, Hopital Nord, Assistance Publique-Hopitaux de Marseille, Marseille, France and* 4) *Département des sciences cliniques et biomédicales, Luigi Sacco, Division des Maladies Infectieuses III, Université de Milan, Milan, Italy*

Abstract

Massilioclostridium coli strain Marseille-P2976^T (= CSUR P2976 = DSM 103344) is a new bacterial genus isolated from the left colon of a patient who underwent colonoscopy for colorectal cancer screening. *Massilioclostridium coli* is a Gram-negative bacillus, strict anaerobic, nonsporogenous and nonmotile organism. We describe here the strain Marseille-P2976^T and provide its complete annotated genome sequence according to taxonogenomics concepts. Its genome is 2 985 330 bp long and contains 2562 predicted genes and 75 RNA genes.

Keywords: Colon, culturomics, *Massilioclostridium coli*, microbiota, taxonogenomics
Original Submission: 21 November 2016; **Revised Submission:** 2 January 2017; **Accepted:** 9 January 2017
Article published online: 16 January 2017

Corresponding author: J.-C. Lagier, Aix-Marseille Université, URMITE, UM63, CNRS7278, IRD198, Inserm 1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385, Marseille Cedex 05, France
E-mail: jclagier@yahoo.fr
C.I. Lo and M. Mailhe contributed equally to this article, and both should be considered first author.

Introduction

The genus *Clostridium* (Prazmowski, 1880) is an obligate anaerobic rod-shaped bacilli capable of producing endospores [1]. Currently it contains 211 species with validly published names [2]. The members of this genus are mostly present in the environment or associated with the commensal digestive flora of humans and animals, but several are major human pathogens, such as *C. difficile, C. botulinum, C. perfringens* and *C. tetani* (http://www.hemltd.ru/export/sites/HemLtd/publications/sections/Normativ/foreign/Infections/medicine/NHS041/article.pdf) [3,4].

We refer to a bacterial classification currently based on a polyphasic approach with phenotypic and genotypic characteristics such as DNA-DNA hybridization, G + C content and 16S rRNA sequence similarity [5,6]. Indeed, the greatest limitation of this classification is the high cost of the DNA-DNA hybridization technique and its low reproducibility [5,7]. In our laboratory, we elaborated a new concept of bacterial description [8–12], with the recent development of genome sequencing technology [13]. This concept, which regroups aspects of the bacteria, is termed taxonogenomics [14]. It is a combination of its proteomic description and its matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) profile [15] associated with a phenotypic description and the sequencing, annotation and comparison of the complete genome of the new bacterial species [16].

Here we describe a new *Massilioclostridium coli* sp. nov., strain Marseille-P2976^T (= CSUR P2976 = DSM 103344) according the concept of taxonogenomics.

Methods

Organism information

Strain Marseille-P2976^T was isolated from 27-year-old woman. A liquid sample was taken from the left colon of the patient as

http://dx.doi.org/10.1016/j.nmni.2017.01.004

she underwent a colonoscopy for colorectal cancer screening. The sample was immediately incubated at 37°C in a blood culture bottle under anaerobic conditions. It was enriched with blood and filtered rumen. After 3 days of preincubation, the strain Marseille-P2976^T was isolated on 5% sheep's blood–enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) at 37°C in anaerobic atmosphere. Written informed consent was obtained from the patient, and the study was approved by the ethics committee of the Institut Hospitalo-Universitaire Méditerranée Infection, Marseille, France, under agreement 2016-010.

Strain identification by MALDI-TOF MS and 16S rRNA sequencing

In the context of the rebirth of culture in microbiology, our sample was cultured using the 18 culture conditions of culturomics [17,18]. Colonies were obtained by spreading samples on petri dishes. They were then purified by subculture and identified by MALDI-TOF MS [19,20]. Some colonies were put down in duplicate on a MTP 96 MALDI-TOF MS target plate (Bruker Daltonics, Leipzig, Germany), which was analyzed with a Microflex spectrometer (Bruker). The 12 spectra obtained were matched against the references of the 7567 bacteria contained in the database by standard pattern matching (with default parameter settings), with MALDI BioTyper database 2.0 software (Bruker). An identification score of >1.9 with a validated species allows identification at the species level, and a score of <1.7 does not enable identification. When identification by MALDI-TOF MS failed, the 16S rRNA was sequenced [21]. Stackebrandt and Ebers [22] suggest similarity levels of 98.7% of the 16S rRNA sequence as a threshold to define a new species without performing DNA-DNA hybridization.

Growth conditions

In order to test the growth conditions of this strain, different temperatures (25, 28, 37, 45 and 56°C) and atmospheres (anaerobic, microaerophilic and aerobic) were analyzed. GENbag anaer and GENbag miroaer systems (bioMérieux) were used respectively to test anaerobic and microaerophilic growth. Only the anaerobic growth was observed with and without 5% of CO_2.

Morphologic, biochemical and antibiotic susceptibility testing

Different phenotypic characteristics such as Gram staining, motility, catalase, oxidase and sporulation were studied as previously described [18]. To carry out a biochemical description, we used, according to the manufacturer's instructions, API 20A (bioMérieux) to identify anaerobes, API ZYM (bioMérieux) to detect enzymatic activities and API 50CH (bioMérieux) to evaluate the capacity to ferment different carbohydrates. Cellular fatty acid methyl ester (FAME) analysis was performed by gas chromatography/mass spectrometry (GC/MS). Two samples were prepared with approximately 20 mg of bacterial biomass per tube collected from several culture plates. FAMEs were prepared as described by Sasser et al. [23]. GC/MS analyses were carried out as previously described [24]. Briefly, FAMEs were separated using an Elite 5-MS column and monitored by MS (Clarus 500-SQ 8 S; PerkinElmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA) and the FAME mass spectral database (Wiley, Chichester, UK).

The tests of antibiotic susceptibility were realized using the disk diffusion method according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) 2015 recommendations (http://www.eucast.org/). The resistance and susceptibility of strain Marseille-P2976^T were estimated with 15 antibiotic treatments: vancomycin 30 µg, rifampicin 30 µg, doxycycline 30 IU, erythromycin 15 IU, amoxicillin 25 µg, gentamicin 15 µg, ceftriaxone 30 µg, amoxicillin 20 µg + clavulanic acid 10 µg, penicillin G 10 µg, gentamicin 15 µg, trimethoprim 1.25 µg + sulfamethoxazole 23.75 µg, oxacillin 5 µg, imipenem 10 µg, tobramycin 10 g, metronidazole 4 µg, fosfomycin 200 µg and daptomycin in stripe 0.016–256 µg (bioMérieux). The bacterial suspension (0.5 McFarland) was made in 2 mL NaCl 0.85% medium. Then this suspension was seeded onto petri dishes with Mueller-Hinton + 5% sheep's blood (Becton Dickinson (BD), San Diego, CA, USA). These 50 different antibiotic dishes (SirScan) were deposited on petri dishes. Electron photomicrography of strain Marseille-P2976^T was performed by doing a negative staining as previously described [25].

Growth conditions and genomic DNA preparation

Massilioclostridium coli strain Marseille-P2976T (= CSUR P2976 = DSM 103344) was grown on 5% sheep's blood–enriched Columbia agar (bioMérieux) at 37°C in anaerobic atmosphere. Bacteria grown on four petri dishes were recovered and suspended in 4 × 100 µL of Tris-EDTA (TE) buffer. Then 200 µL of this suspension was diluted in 1 mL TE buffer for lysis treatment. Thirty minutes' incubation with 2.5 µg/µL lysozyme at 37°C followed by an overnight incubation with 20 µg/µL proteinase K at 37°C was necessary for a complete lysis. We purified extracted DNA using three successive phenol–chloroform extractions and ethanol precipitations at −20°C overnight. Then the DNA was resuspended in 160 µL TE buffer after centrifugation.

Genome sequencing and assembly

Genomic DNA of *Massilioclostridium coli* was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The genomic DNA (gDNA) was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). gDNA was quantified by a Qubit assay with the high sensitivity kit (Life Technologies, Carlsbad, CA, USA) to 60.7 ng/μL. The mate pair library was prepared with 1.5 μg of genomic DNA using the Nextera mate pair Illumina guide. The genomic DNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments ranged in size from 1.5 to 11 kb with an optimal size at 7.68 kb. No size selection was performed, and 600 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with optima on a bimodal curve at 990 and 1536 bp on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured at 29.27 nmol/L.

The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and a sequencing run were performed in a single 39-hour run at a 2 × 151 bp read length.

The total information of 7.9 Gb was obtained from a 863K/mm^2 cluster density with a cluster passing quality control filters of 94% (15 627 000 passing filter paired reads). Within this run, the index representation for *Massilioclostridium coli* was determined to be 10.02%. The 1 565 833 paired reads were trimmed, then assembled into six scaffolds.

Genome annotation and comparison

Open reading frame (ORF) prediction was carried out using Prodigal [26] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region (contains N). The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COGs) using BLASTP (*E* value 1e-03, coverage 70%, identity percent 30%). If no hit was found, it searched against the NR database using BLASTP with an *E* value of 1e-03 coverage 70% and identity percent of 30%. If sequence lengths were smaller than 80 amino acids, we used an *E* value of 1e-05. The tRNAScanSE tool [27] was used to find tRNA genes, whereas rRNAs were found by using RNAmmer [28]. Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [29]. ORFans were identified if all the performed BLASTP procedures did not give positive results (*E* value smaller than 1e-03 for ORFs with sequence size superior to 80 aa or *E* value smaller than 1e-05 for ORFs with sequence length smaller than 80 aa). Such threshold parameters have already been used in previous works to define ORFans.

Genomes were automatically retrieved from the 16S RNA tree using Xegen software (PhyloPattern [30]). For each selected genome, complete genome sequence, proteome and ORFeome genome sequence were retrieved from the National Center for Biotechnology Information FTP site. All proteomes were analyzed with proteinOrtho [31]. Then for each couple of genomes, a similarity score was computed. This score is the mean value of nucleotide similarity between all couples of orthologues between the two genomes studied (average genomic identity of orthologous gene sequences, AGIOS) [17].

An annotation of the entire proteome was performed to define the distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins (using the same method as for the genome annotation). To evaluate the genomic similarity among the compared strains, we determined two parameters: digital DNA-DNA hybridization, which exhibits a high correlation with DNA-DNA hybridization (DDH) [32,33], and AGIOS [17], which was designed to be independent from DDH.

Results

Strain identification and phylogenetic analyses

Strain Marseille-P2976^T, the details of which are provided in Table 1, was first isolated in April 2016 by culturing a sample in anaerobic atmosphere on 5% sheep's blood–enriched Colombia agar (bioMérieux) at 37°C after 24 hours' incubation.

A significant score was not obtained when we used MALDI-TOF MS for identification for strain Marseille-P2976^T. This

TABLE 1. Classification and general features of *Massilioclostridium coli* strain Marseille-P2976^T

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Firmicutes* Class: *Clostridia* Order: *Clostridiales* Family: *Clostridiaceae* Genus: *Massilioclostridium* Species: *Massilioclostridium coli* Type strain: Marseille-P2976^T
Gram stain	Negative
Cell shape	Rod
Motility	Nonmotile
Sporulation	Non–spore forming
Temperature range	Mesophilic
Optimum temperature	37°C

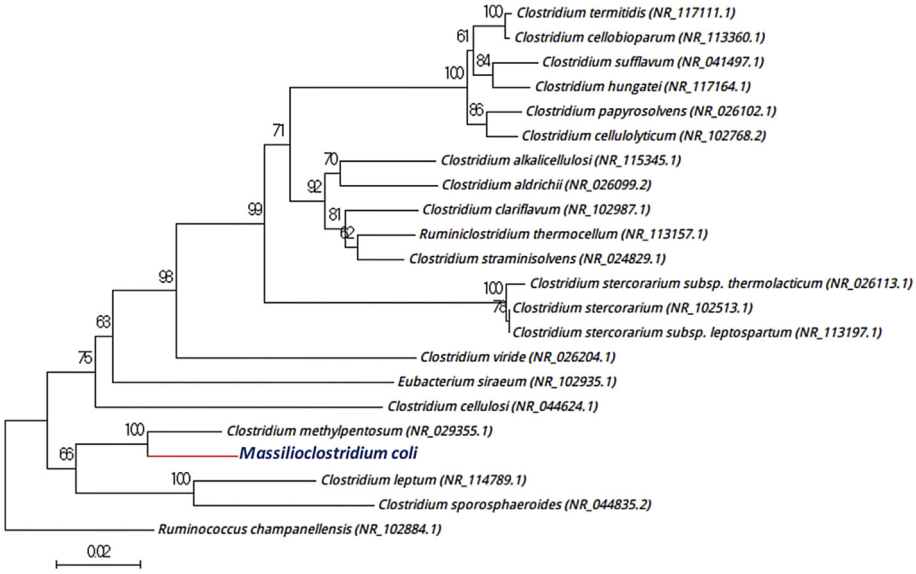

FIG. 1. Phylogenetic trees highlighting position of *Massilioclostridium coli* strain Marseille-P2976^T (= CSUR P2976 = DSM 103344) relative to other strains within the genus *Clostridium*. Sequences of 16S rRNA gene were aligned by CLUSTALW. Scale bar represents 2% nucleotide sequence divergence.

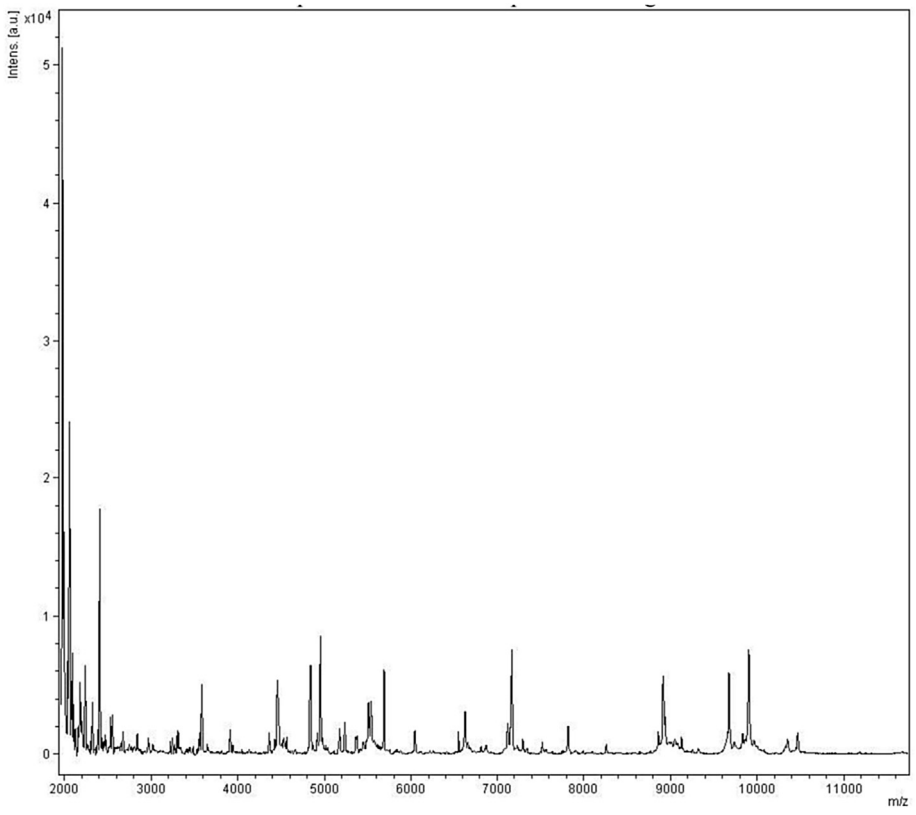

FIG. 2. Reference mass spectrum from *Massilioclostridium coli* strain Marseille-P2976^T. Spectra from 12 individual colonies were compared and reference spectrum generated.

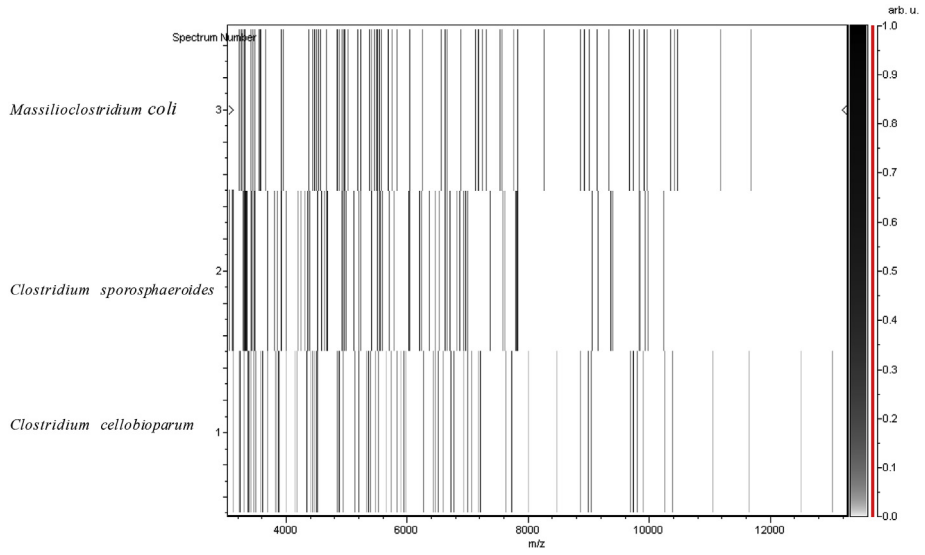

FIG. 3. Gel view comparing *Massilioclostridium coli* strain Marseille-P2976^T (= CSUR P2976 = DSM 103344) to other species within *Clostridiaceae* family. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. *X*-axis records m/z value. Left *y*-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Colour bar and right *y*-axis indicate relation between colour peak displayed and peak intensity in arbitrary units.

proves that the spectrum of this isolate did not match any spectra in our MALDI-TOF database. The nucleotide sequences of the 16S rRNA genes of strain Marseille-P2976^T (GenBank accession no. LT598551) showed a 95.1% similarity level with *Clostridium methylpentosum* [34], the phylogenetically closest species with a validly published name (Fig. 1), therefore defining it as a new bacterial genus. The spectra of strain Marseille-P2976^T (Fig. 2) were added to our MALDI-TOF database. The reference spectrum for Marseille-P2976^T was then compared to the spectra of some *Clostridium* species available in our database, and the differences were illustrated in a gel view photograph (Fig. 3).

Phenotypic description

The strain's growth was observed from 37 to 45°C on 5% sheep's blood–enriched Columbia agar (bioMérieux). Optimal growth was achieved at 37°C in anaerobic condition after 24 hours' incubation. Cells were not motile, and we did not

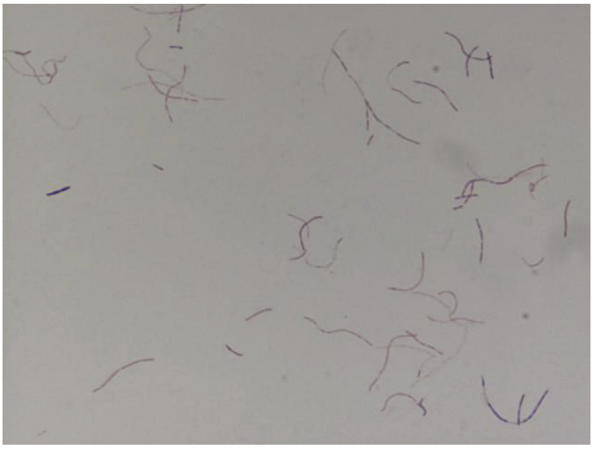

FIG. 4. Gram staining of *Massilioclostridium coli* strain Marseille-P2976^T.

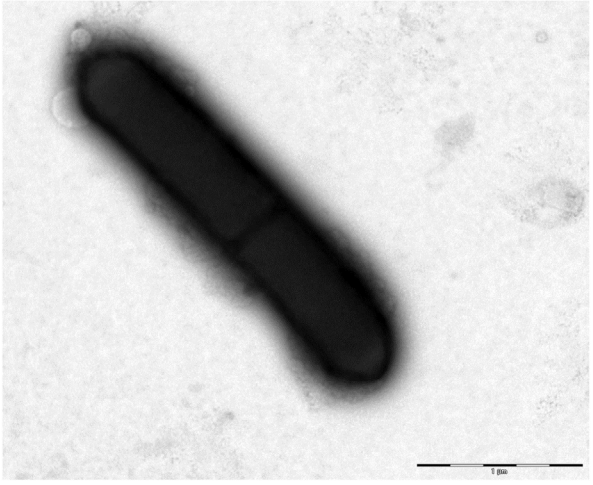

FIG. 5. Transmission electron microscopy of *Massilioclostridium coli* strain Marseille-P2976^T using Morgani 268D (Philips, Amsterdam, The Netherlands) at operating voltage of 60 kV. Scale bar = 1 µm.

TABLE 2. Cellular fatty acid composition (%)

Fatty acid	Name	Mean relative %[a]
16:0 iso	14-methyl-Pentadecanoic acid	33.2 ± 1.3
17:0 anteiso	14-methyl-Hexadecanoic acid	20.4 ± 0.1
16:0	Hexadecanoic acid	15.3 ± 0.6
15:0 anteiso	12-methyl-Tetradecanoic acid	13.4 ± 0.4
15:0 iso	13-methyl-Tetradecanoic acid	5.0 ± 0.1
17:0 iso	15-methyl-Hexadecanoic acid	2.0 ± 0.2
18:1n9	9-Octadecenoic acid	1.1 ± 0.1
18:2n6	9,12-Octadecadienoic acid	TR
14:0	Tetradecanoic acid	TR
17:0	Heptadecanoic acid	TR
15:0	Pentadecanoic acid	TR

TR, trace amounts <1%.
[a]Mean peak area percentage.

TABLE 4. Nucleotide content and gene count levels of genome

Attribute	Genome (Total)	
	Value	% of total[a]
Size (bp)	2 985 330	100.0
G + C content (bp)	1 193 310	39.9
Coding region (bp)	2 541 428	85.1
Total of genes	2562	100.0
RNA genes	75	2.9
Protein-coding genes	2487	100.0
Protein with function prediction	1599	64.2
Protein assigned to COGs	1369	55.0
Genes with peptid signals	299	12.0
Genes with transmembrane helices	644	25.8

COGs, Clusters of Orthologous Groups database.
[a]Total is based on either size of genome in base pairs or total number of protein-coding genes in annotated genome.

observe any spore formation after thermal shock. They were Gram-negative bacilli (Fig. 4). On 5% sheep's blood–enriched Columbia agar, colonies were circular, thin and translucent with an entire margin of 1.0 mm in diameter after 24 hours at 37°C. Under electron microscopy, the cells had a mean diameter of 2.5 µm and a length of 4 µm (Fig. 5). Oxidase and catalase tests were both negative.

The major fatty acid was 14-methyl-pentadecanoic acid (33%). All described fatty acids were saturated structures except two (18:1n9 and 18:2n6). Almost half of the listed fatty acids are branched structures (iso and anteiso) (Table 2).

Strain Marseille-P2976^T was able to ferment D-tagatose and potassium 5-ketogluconate; ribose, maltose, glucose, D-mannose, D-ribose, D-galactose, D-glucose, D-fructose, D-maltose, D-saccharose, arbutin and ferric citrate esculin were not hydrolyzed. Indole and urea were negative. Acid phosphatase, alkaline phosphatase, esterase (C4) and naphthol-AS-BI-phosphohydrolase activities were positive. Leucine arylamidase, pyrazinamidase, cystine arylamidase, trypsin, acid phosphatase, β-glucosidase and esterase lipase (C8) activities were negative.

Strain Marseille-P2976^T is susceptible to vancomycin, amoxicillin + clavulanic acid, daptomycin, metronidazole, imipenem, tobramycin and ceftriaxone but resistant to gentamicin, colistin, trimethoprim + sulfamethoxazole, oxacillin, fosfomycin, doxycycline, erythromycin and amoxicillin.

The biochemical and phenotypic features of strain Marseille-P2976^T were compared to those of other close representative strains in the *Clostridiaceae* family (Table 3).

TABLE 3. Differential characteristics of *Massilioclostridium coli* strain Marseille-P2976^T, *Clostridium thermocellum* strain JW20^T, *Clostridium difficile* strain 630, *Clostridium dakarense* strain FF1^T, *Clostridium beijerinckii* strain NCIMB 8052^T [4,35–37]

Property	*M. coli*	*C. thermocellum*	*C. difficile*	*C. dakarense*	*C. beijerinckii*
Cell diameter (µm)	2.5	NA	2.5	1.2	1.5
Oxygen requirement	−	−	−	−	−
Gram stain	−	+	−	+	+
Salt requirement	−	−	−	−	−
Motility	−	−	−	+	+
Endospore formation	−	+	+	+	+
Production of:					
Acid phosphatase	+	−	NA	+	+
Catalase	−	−	−	−	−
Oxidase	−	−	NA	−	−
Indole	−	−	NA	−	NA
Urease	−	−	NA	−	−
β-Galactosidase	−	−	NA	−	+
N-acetyl-glucosamine	−	−	NA	+	NA
Ribose	−	NA	NA	−	NA
Pyrazinamidase	−	NA	NA	NA	NA
Pyrrolidinyl arylamidase	−	NA	NA	NA	NA
Mannose	−	−	NA	−	+
Mannitol	−	−	+	−	+
Sucrose	−	−	+	−	+
D-Glucose	−	−	+	+	+
D-Fructose	−	−	+	−	+
D-Maltose	−	−	+	+	+
D-Lactose	−	−	+	−	+
Habitat	Human	Louisiana cotton bale	Poplar wood	Human gut	Human gut

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

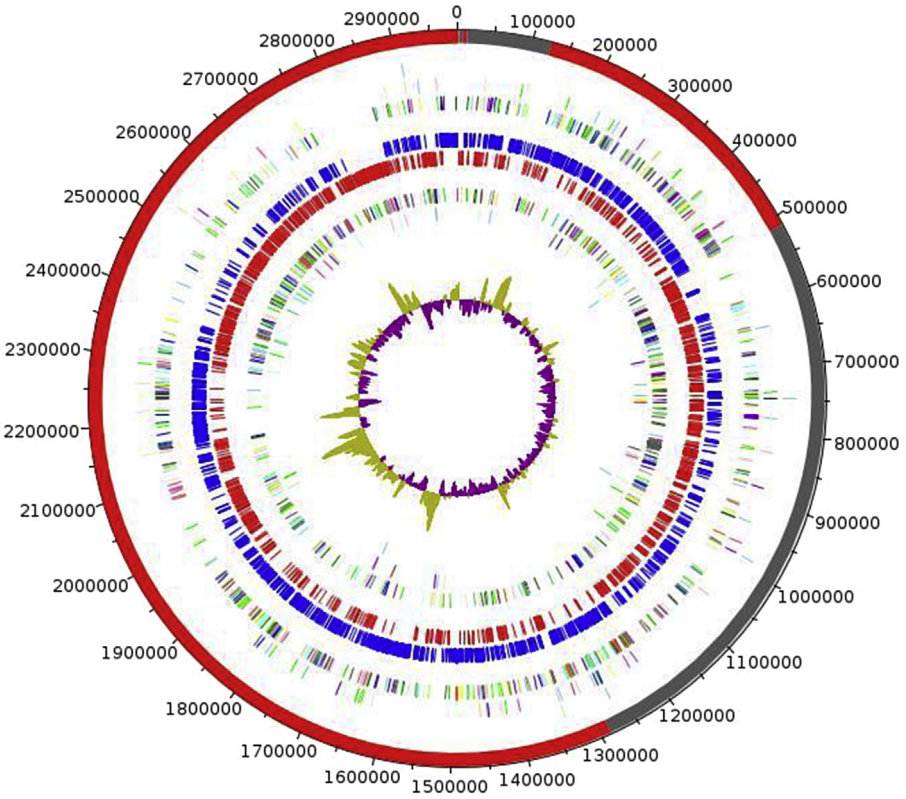

FIG. 6. Graphical circular map of chromosome. From outside to center: genes on forward strain coloured by COGs categories (only gene assigned to COGs), RNA genes (tRNAs green, rRNAs red), GC content and GC skew. COGs, Clusters of Orthologous Groups database.

TABLE 5. Number of genes associated with 25 general COGs functional categories

Code	Value	% of total[a]	Description
J	182	7.31	Translation
A	0	0	RNA processing and modification
K	134	5.38	Transcription
L	78	3.13	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	33	1.32	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	56	2.25	Defense mechanisms
T	67	2.69	Signal transduction mechanisms
M	85	3.41	Cell wall/membrane biogenesis
N	6	0.24	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	25	1.00	Intracellular trafficking and secretion
O	62	2.49	Posttranslational modification, protein turnover, chaperones
X	23	0.92	Mobilome: prophages, transposons
C	84	3.37	Energy production and conversion
G	83	3.33	Carbohydrate transport and metabolism
E	132	5.30	Amino acid transport and metabolism
F	55	2.21	Nucleotide transport and metabolism
H	67	2.69	Coenzyme transport and metabolism
I	42	1.68	Lipid transport and metabolism
P	63	2.53	Inorganic ion transport and metabolism
Q	23	0.92	Secondary metabolites biosynthesis, transport and catabolism
R	123	4.94	General function prediction only
S	86	3.45	Function unknown
—	1118	44.9	Not in COGs

COGs, Clusters of Orthologous Groups database.
[a]Total is based on total number of protein-coding genes in annotated genome.

Genome properties

The genome is 2 985 330 bp long with 39.97% GC content (Table 4, Fig. 6). It is composed of seven scaffolds (composed of seven contigs). Of the 2562 predicted genes, 2487 were protein-coding genes and 75 were RNAs (five genes are 5S rRNA, five genes are 16S rRNA, five genes are 23S rRNA and 60 genes are tRNA genes). A total of 1599 genes (64.29%) were assigned as putative function (by COGs or by NR BLAST). Among them, 224 genes were identified as ORFans (9.01%). The remaining genes were annotated as hypothetical proteins (578 genes, 23.24%).

TABLE 6. Genome comparison of closely related species to *Massilioclostridium coli* strain Marseille-P2976^T

Species (strain)	Size (Mb)	G + C (%)	Total genes
Flavonifractor plautii strain (CCUG 28093)	3.82	61.0	4278
Pseudoflavonifractor capillosus (ATCC 29799)	4.24	59.1	4829
Oscillibacter valericigenes (NBRC_101213)	4.47	53.1	4723
Ruminococcus albus (DSM 20455)	4.33	43.6	3991
Oscillibacter ruminantium (GH1)	3.10	54.9	2929
Intestinimonas butyriciproducens (SRB-521−5-I)	3.37	59.0	3381
Massilioclostridium coli (Marseille-P2976)	2.98	39.9	2487
Ethanoligenens harbinense (YUAN-3)	3.01	55.5	2701

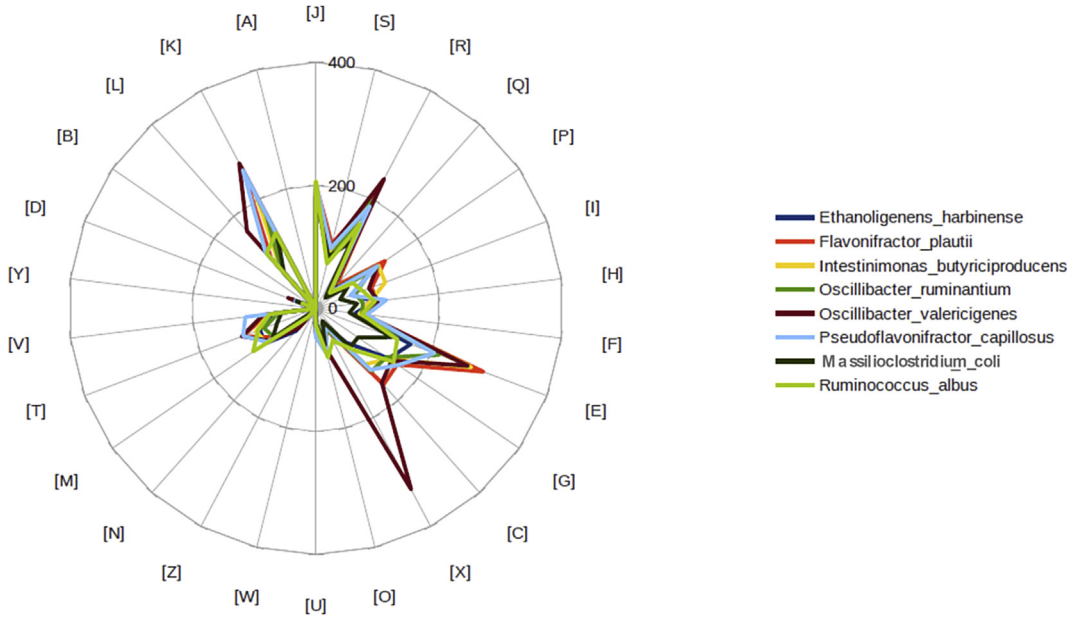

FIG. 7. Distribution of functional classes of predicted genes according to COGs proteins. COGs, Clusters of Orthologous Groups database.

The National Center for Biotechnology Information project ID is PRJEB15310, and the genome is deposited under accession number FMIZ01000000. The distribution of genes into COGs functional categories is presented in Table 5.

Genome comparison

Massilioclostridium coli genomic characteristics were compared to other close species (Table 6). The draft genome sequence of *Massilioclostridium coli* is smaller than that of *Flavonifractor plautii, Pseudoflavonifractor capillosus, Oscillibacter valericigenes, Ruminococcus albus, Oscillibacter ruminantium, Intestinimonas butyriciproducens* and *Ethanoligenens harbinense* (2.98, 3.82, 4.24, 4.47, 4.33, 3.10, 3.37 and 3.01 MB respectively).

The G + C content of *Massilioclostridium coli* is smaller than that of *Flavonifractor plautii, Pseudoflavonifractor capillosus, Oscillibacter valericigenes, Ruminococcus albus, Oscillibacter ruminantium, Intestinimonas butyriciproducens* and *Ethanoligenens harbinense* (39.97, 61.07, 59.11, 53.19, 43.62, 54.97, 59.06 and 55.56% respectively).

The gene content of *Massilioclostridium coli* is smaller than that of *Flavonifractor plautii, Pseudoflavonifractor capillosus, Oscillibacter valericigenes, Ruminococcus albus, Oscillibacter ruminantium, Intestinimonas butyriciproducens* and *Ethanoligenens harbinense* (2487, 4278, 4829, 4723, 3991, 2929, 3381 and 2701 genes respectively).

Fig. 7 shows the comparison of gene distribution into COG categories of Marseille-P2976^T with the other finished genomes cited above. *Massilioclostridium coli* strain Marseille-P2976^T shared 731, 829, 791, 754, 804, 727 and 750 orthologous genes with *E. harbinense, F. plautii, I. butyriciproducens, O. valericigenes, P. capillosus, R. albus* and *O. ruminantium* respectively (Table 7). Among species with standing in nomenclature, AGIOS values ranged from 49.96% between *F. plautii* and Marseille-P2976^T to 65.90% between *O. valericigenes* and *O. ruminantium*. When Marseille-P2976^T was compared to the other species, AGIOS values ranged from 49.96% with *F. plautii* to 61.76% with *R. albus*.

TABLE 7. Numbers of orthologous protein shared between genomes (upper right)[a]

	Ethanoligenens harbinense YUAN 3	*Flavonifractor plautii* CCUG 28093	*Intestinimonas butyriciproducens* SRB 521	*Oscillibacter valericigenes* NBRC 101213	*Pseudoflavonifractor capillosus* ATCC 29799	*Ruminococcus albus* ATCC 27210	*Oscillibacter ruminantium* GH1	*Massilioclostridium coli* Marseille-P2976
E. harbinense	**2701**	778	741	759	751	700	736	731
F. plautii	56.08	**4278**	1278	1108	1273	745	1077	829
I. butyriciproducens	56.06	63.35	**3381**	1033	1234	733	1012	791
O. valericigenes	56.75	59.35	57.74	**4723**	988	699	1228	754
P. capillosus	57.62	64.32	63.74	59.87	**4829**	765	979	804
R. albus	54.56	52.44	53.61	54.53	54.01	**3991**	690	727
O. ruminantium	56.60	58.14	57.85	65.90	58.53	59.75	**2929**	750
M. coli	53.40	49.96	51.11	52.55	51.23	61.76	58.23	**2487**

[a]Average percentage similarity of nucleotides corresponding to orthologous protein shared between genomes (lower left) and numbers of proteins per genome (bold).

Conclusion

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of *Massilioclostridium coli* which contains the type strain Marseille-P2976^T. This bacterial strain was isolated from the left colon of a woman seeking care at our hospital.

Description of *Massilioclostridium* gen. nov.

Massilioclostridium (mas.il'io, L. gen. masc., from *massilio,* 'of Massilia,' the Latin name of Marseille, France) is a Gram-negative bacilli. It is indole, catalase, oxidase and urea negative. Cells of this genus have a mean diameter and length of 2.5 µm and 4 µm respectively. It ferments only D-tagatose and potassium 5-ketogluconate. Ribose, D-galactose, maltose, D-glucose, D-fructose, glucose, D-mannose, D-ribose, D-maltose and D-saccharose were not used. Positive reactions were noted from acid phosphatase, alkaline phosphatase, esterase (C4) and naphthol-AS-BI-phosphohydrolase, while negative reactions were observed from leucine arylamidase, pyrazinamidase, cystine arylamidase, trypsin, acid phosphatase, β-glucosidase and esterase lipase (C8). Habitat and type species are respectively humans and *Massilioclostridium coli.*

Description of *Massilioclostridium coli* gen. nov., sp. nov.

Massilioclostridium coli (coli, M. L. adj. *coli,* pertaining to the colon, where strain Marseille-P2976^T was isolated). The strain grows at temperatures ranging between 37 and 45°C in anaerobic conditions (at an optimum temperature of 37°C). Salinity range growth was tested between 10 and 20% (no growth was observed), while pH growth is about 5 to 8 (with an optimum of 7). The potential pathogenicity of the type strain Marseille-P2976^T (= CSUR P2976 = DSM 103344) is unknown, but it was isolated from the left colon of a woman seeking care at our hospital. This strain exhibited a G + C content of 39.97%. The genome and 16S rRNA sequences of *M. coli* were deposited in GenBank under accession numbers LT598551 and FMIZ01000000, respectively.

Acknowledgements

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process. This study was funded by the Fondation Méditerranée Infection.

Conflict of interest

None declared.

References

[1] Prazmowski A. Untersuchung über die Entwickelungsgeschichte und Fermentwirking einiger Bakterien-Arten. Unpublished doctoral dissertation. Germany: University of Leipzig; 1880.
[2] Parte AC. LPSN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res 2014;42:D613–6.
[3] Wells CL, Wilkins TD. Clostridia: spore forming anaerobic bacilli. In: Baron S, editor. Baron's medical microbiology. 4th ed. Galveston, TX: University of Texas Medical Branch; 1996.
[4] Rodriguez C, Taminiau B, Van Broeck J, Delmée M, Daube G. *Clostridium difficile* infection and intestinal microbiota interactions. Microb Pathog 2015;89:201–9.
[5] Rosselló-Móra R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Stackebrandt E, editor. Molecular identification, systematics, and population structure of prokaryotes. Berlin: Springer; 2006. p. 23–50.
[6] Wayne LG, Brenner DJ, Colwell PR, Grimont PAD, Kandler O, Krichevsky MI, et al. Report of the Ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol 1987;37:463–4.
[7] Viale AM, Arakaki AK, Soncini FC, Ferreyra RG. Evolutionnary relationships among eubacterial groups as inferred from GroEL (Chaperonin) sequences comparison. Int J Syst Bacteriol 1944;44:527–33.
[8] Seck E, Rathored J, Khelaifia S, Croce O, Robert C, Couderc C, et al. *Virgibacillus senegalensis* sp. nov., a new moderately halophilic bacterium isolated from human gut. New Microbes New Infect 2015;8:116–26.
[9] Tidjani Alou M, Nguyen TT, Armstrong N, Rathored J, Khelaifia S, Raoult D, et al. *Numidum massiliense* gen. nov., sp. nov., a new member of the *Bacillaceae* family isolated from the human gut. New Microbes New Infect 2016;12:76–85.
[10] Lo CI, Padhmanabhan R, Mediannikov O, Caputo A, Michelle C, Faye N, et al. High-quality genome sequence and description of *Bacillus ndiopicus* strain FF3T sp. nov. New Microbes New Infect 2015;8:154–63.
[11] Lo CI, Sankar SA, Ehounoud CB, Mediannikov O, Labas N, Caputo A, et al. High-quality genome sequencing and description of *Dermabacter indicis* sp.nov. New Microbes New Infect 2016;11:59–67.
[12] Lagier JC, Armougom F, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. Stand Genomic Sci 2012;6:315–24.
[13] Reddy TB, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, et al. The Genome OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta) genome project classification. Nucleic Acids Res 2015;43(Database issue):D1099–106.
[14] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 2015;36:73–8.
[15] Fall B, Lo CI, Samb-Ba B, Perrot N, Diawara S, Gueye MW, et al. The ongoing revolution of MALDI-TOF mass spectrometry for microbiology reaches tropical Africa. Am J Trop Med Hyg 2015;92:641–7.
[16] 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(Pt 2):384–91.

[17] Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 2015;28:237–64.
[18] Lagier JC, Khelaifia S, Tidjani Alou M, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 2016;1:16203.
[19] Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, et al. Identification of rare pathogenic bacteria in a clinical microbiology laboratory: impact of matrix-assisted laser desorption ionization–time of flight mass spectrometry. J Clin Microbiol 2013;51:2182–94.
[20] Lo CI, Fall B, Sambe-Ba B, Diawara S, Gueye MW, Mediannikov O, et al. MALDI-TOF mass spectrometry: a powerful tool for clinical microbiology at Hôpital Principal de Dakar, Senegal (West Africa). PLoS One 2015;10:e0145889.
[21] Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 2000;38:3623–30.
[22] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[23] Sasser M, Kunitsky C, Jackoway G, Ezzell JW, Teska JD, Harper B, et al. Identification of *Bacillus anthracis* from culture using gas chromatographic analysis of fatty acid methyl esters. J AOAC Int 2005;88:178–81.
[24] Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N, et al. Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes New Infect 2016;10:66–76.
[25] Cresci M, Ibrahima Lo C, Khelaifia S, Mouelhi D, Delerce J, Di Pinto F, et al. *Corynebacterium phoceense* sp. nov., strain MC1^{T} a new bacterial species isolated from human urine. New Microbes New Infect 2016;14:73–82.
[26] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 2010;11:119.
[27] Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 1997;25:955–64.
[28] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[29] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–36.
[30] Gouret P, Thompson JD, Pontarotti P. PhyloPattern: regular expressions to identify complex patterns in phylogenetic trees. BMC Bioinform 2009;10:298.
[31] Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinform 2011;12:124.
[32] Auch AF, von Jan M, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2010;2:117–37.
[33] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 2013;14:60.
[34] Himelbloom BH, Canale-Parola E. *Clostridium methylpentosum* sp. nov.: a ring-shaped intestinal bacterium that ferments only methylpentoses and pentoses. Arch Microbiol 1989;151:287–93.
[35] Freier D, Mothershed CP, Wiegel J. Characterization of *Clostridium thermocellum* JW20. Appl Environ Microbiol 1988;54:204–11.
[36] Lo CI, Mishra AK, Padhmanabhan R, Samb-Ba B, Gassama-Sow A, Robert C, et al. Non-contiguous finished genome sequence and description of *Clostridium dakarense* sp. nov. Stand Genomic Sci 2013;9:14–27.
[37] Keis S, Shaheen R, David T. Jones Emended descriptions of *Clostridium acetobutylicum* and *Clostridium beijerinckii*, and descriptions of *Clostridium saccharoperbutylacetonicum* sp. nov. and *Clostridium saccharobutylicum* sp. nov. Int J Syst Evol Microbiol 2001;51:2095–103.