Butyricimonas phoceensis sp. nov., a new anaerobic species isolated from the human gut microbiota of a French morbidly obese patient

A. H. Togo[1], A. Diop[1], G. Dubourg[1], T. T. Nguyen[1], C. Andrieu[1], A. Caputo[1], C. Couderc[1], P.-E. Fournier[1], M. Maraninchi[2,3], R. Valero[2,3], D. Raoult[1,4] and M. Million[1]

1) *Aix Marseille Université, URMITE, Institut Hospitalier Universitaire Méditerranée-Infection, UM63, CNRS7278, IRD198, INSERM1095,* 2) *Aix Marseille Université, NORT "Nutrition, Obesity and Risk of Thrombosis", INSERM1062, INRA1260,* 3) *APHM, CHU Hôpital de la Conception, Service Nutrition, Maladies Métaboliques et Endocrinologie, F-13385 Marseille, France and* 4) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Butyricimonas phoceensis strain AT9 (= CSUR 2478 = DSM 100838) was isolated from a stool sample from a morbidly obese French patient living in Marseille using the culturomics approach. The genome of this Gram-negative-staining, anaerobic and non–spore forming rod bacillus is 4 736 949 bp long and contains 3947 protein-coding genes. Genomic analysis identified 173 genes as ORFans (4.5%) and 1650 orthologous proteins (42%) not shared with the closest phylogenetic species, *Butyricimonas virosa*. Its major fatty acid was the branched acid iso-C15:0 (62.3%).

Keywords: Butyrate, *Butyricimonas phoceensis*sp. nov., culturomics, genome, obesity, taxonogenomics
Original Submission: 18 May 2016; **Revised Submission:** 11 July 2016; **Accepted:** 25 July 2016
Article published online: 9 August 2016

Corresponding author: M. Million, Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095, Marseille, France
E-mail: matthieumillion@gmail.com

Introduction

Butyricimonas phoceensis strain AT9 (= CSUR P2478 = DSM 100838 was isolated from the faeces of a 57-year-old French woman living in Marseille with class III morbid obesity (body mass index (BMI) 55.8 kg/m^2). This isolate is part of an exploratory study of the gut flora from obese patients before and after bariatric surgery. Bariatric surgery is the most effective treatment for morbid obesity for sustainable weight loss and leads to an enrichment of the gut flora [1]. The goal of our study was to compare microbial diversity of the gut flora in obese patients before and after bariatric surgery by culturomics. The aim of culturomics is to exhaustively explore the microbial ecosystem of gut flora by using different culture conditions followed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) identification [2].

The conventional approaches used in the delineation of bacterial species are 16S rRNA sequence identity and phylogeny [3], genomic (G+C content) diversity and DNA-DNA hybridization (DDH) [4,5]. However, these approaches present some difficulties, mainly as a result of their cutoff values, which change according to species or genera [6]. The accession of new technology tools, such as high-throughput sequencing, has made available nucleotide sequence libraries for many bacterial species [7]. We recently suggested incorporating genomic data in a polyphasic taxonogenomics approach to describe new bacteria. This procedure considers phenotypic characteristics, genomic analysis and the MALDI-TOF MS spectrum comparison [8,9].

Here we propose a classification and a set of characteristics for *Butyricimonas phoceensis* strain AT9, together with the description of complete genome sequencing, annotation and comparison as new species belonging to the genus *Butyricimonas*. The genus *Butyricimonas* was established in 2009 by

http://dx.doi.org/10.1016/j.nmni.2016.07.010

Sakamoto and encompasses four described species (*B. faecihominis, B. synergistica, B. paravirosa* and *B. virosa*). They were isolated from rat or human faeces and belong to the family *Porphyromonadaceae* [10,11]. The family *Porphyromonadaceae* contains 11 genera: *Porphyromonas* (type genus), *Barnesiella, Butyricimonas, Dysgonomonas, Macellibacteroides, Odoribacter, Paludibacter, Parabacteroides, Petrimonas, Proteiniphilum* and *Tannerella* [12]. *Butyricimonas virosa* bacteraemia has been described in patients with colon cancer [13,14] and in patients with posttraumatic chronic bone and joint infection [14].

Materials and Methods

Sample collection

A stool sample was collected from a 57-year-old obese French woman (BMI 55.8 kg/m^2; 150 kg, 1.64 m tall) in June 2012. Written informed consent was obtained from the patient at the nutrition, metabolic disease and endocrinology service at La Timone Hospital (Marseille, France). The study and assent procedure were approved by the local ethics committee (IFR 48, no. 09-022, 2010). The stool sample was stored at −80°C after collection.

Isolation and identification of strain

Strain isolation was performed in May 2015. Stool extract was preincubated in blood culture bottles enriched with lamb rumen juice and sheep's blood in anaerobic atmosphere as described elsewhere [2]. The culture was followed closely for 30 days. At different time points (days 1, 3, 7, 10, 15, 21 and 30), a seeding of the preincubated product was performed on sheep's blood–enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) during 48 hours of incubation in an anaerobic atmosphere at 37°C. Colonies that emerged were cultivated in the same isolated conditions.

The colonies were then identified by MALDI-TOF MS as previously described [15]. Briefly, one isolated bacterial colony was picked up with a pipette tip from a culture agar plate and spread as a thin smear on a MTP 384 MALDI-TOF MS target plate (Bruker Daltonics, Leipzig, Germany). Measurement and identification were performed as previously described [16]. When a bacterium was unidentifiable, 16S rRNA gene amplification and sequencing were performed.

The 16S rRNA PCR coupled with sequencing were performed using GeneAmp PCR System 2720 thermal cyclers (Applied Biosystems, Bedford, MA, USA) and ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems) respectively [17]. Chromas Pro 1.34 software (Technelysium, Tewantin, Australia) was used to correct sequences, and BLASTn searches were performed at the National Center for Biotechnology Information (NCBI) website (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi).

Phylogenetic analysis

A custom Python script was used to automatically retrieve all species from the same family of the new species and download 16S sequences from NCBI by parsing NCBI results and NCBI taxonomy page. The scripts also remove species that are not found on the List of Prokaryotic Names With Standing in Nomenclature (LPSN) website (http://www.bacterio.net/). The script retains the most appropriate 16S sequence (the longest sequence with the smallest number of degenerate nucleotides) whilst also retaining one sequence from another genus as an outside group. It then aligns and trims the extremities of the sequences. Sequences were aligned using Muscle v3.8.31 with default parameters, and phylogenetic inferences were obtained using neighbour-joining method with 500 bootstrap replicates within MEGA6 software.

Phenotypic and biochemical characterization

Growth conditions. Different growth temperatures (28, 37, 45 and 55 °C) were tested on sheep's blood–enriched Columbia agar (bioMérieux). Growth of this strain was tested under anaerobic conditions using the GENbag anaer system (bioMérieux), microaerophilic conditions using the GENbag microaer system (bioMérieux) and under aerobic conditions with or without 5% CO_2. The tolerance to salt of this strain over a range salt concentrations (0–100 g/L) on Schaedler agar with 5% sheep's blood (bioMérieux) under anaerobic atmosphere was performed.

Microscopy. A heat shock at 80°C for 20 minutes was performed for the sporulation test. A fresh colony was observed between blades and slats using a photonic microscope Leica DM 1000 (Leica Microsystems, Nanterre, France) at 40× to assess the motility of the bacteria. Gram staining was performed and observed using a photonic microscope Leica DM 2500 with a 100× oil-immersion objective lens. Transmission electron microscopy using a Tecnai G20 device (FEI Company, Limeil-Brevannes, France) at an operating voltage of 60 kV was performed to observe strain AT9 after negative colouration.

Biochemical assays. Biochemical assays were performed using API Gallery systems (API ZYM, API 20A and API 50CH) according to the manufacturer's instructions (bioMérieux). Detection of catalase (bioMérieux) and oxidase (Becton Dickinson, Le Pont de Claix, France) was also performed according to the manufacturer's instructions.

Antibiotic susceptibility. The antibiotic susceptibility of the strain was tested using a disk diffusion method [18] for 21 antibiotics

including the following: amoxicillin 25 µg/mL, amoxicillin–clavulanic acid 30 µg/mL, ceftriaxone 30 µg, ciprofloxacin 5 µg, clindamycin (DA15), colistin (CT50), Dalacin 15 µg/mL, doripenem 10 µg/mL, doxycycline 30 IU, erythromycin 15 IU, fosfomycin 10 µg, gentamicin 500 µg, gentamicin 15 µg, imipenem 10 µg/mL, metronidazole 4 µg/mL, oxacillin 5 µg, penicillin G 10 IU, rifampicin 30 µg, sulfamethoxazole 23.75 µg, trimethoprim 1.25 µg, teicoplanin (TEC30) and vancomycin 30 µg (i2a, Montpellier, France). The 1200 scan was used for the interpretation of results (Interscience, Saint-Nom-La-Bretèche, France).

Fatty acid analysis. Fresh colonies from a plate of Columbia agar with 5% sheep's blood were collected after 48 hours' incubation at 37°C for fatty acid analysis. Cellular fatty acid analysis was performed by gas chromatography/mass spectrometry (GC/MS). Two samples were prepared with approximately 100 mg of bacterial biomass each collected from a culture plate. Cellular fatty acid methyl esters were prepared as described by Sasser [19]. GC/MS analyses were carried out as previously described [20]. Briefly, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by a Clarus 500 gas chromatograph equipped with a SQ8S MS detector (PerkinElmer, Courtaboeuf, France). Fatty acid methyl esters were identified by using the spectral database search using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology, Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK).

Genome sequencing and assembly

Genomic DNA (gDNA) of strain AT9 was sequenced using MiSeq Technology (Illumina, San Diego, CA, USA) with the mate-pair strategy. The gDNA was barcoded so it could be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The gDNA was quantified by a Qubit assay with the high-sensitivity kit (Thermo Fisher Scientific Life Sciences, Waltham, MA, USA) to 325 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 lab chip. The DNA fragments ranged in size from 1.5 to 11 kb with an optimal size at 4.8 kb. No size selection was performed, and 600 ng of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal at 966 bp on the Covaris S2 device 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 24.3 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 sequencing run were performed in a single 39-hour run at a 2 × 151 bp read length. Total information of 8.9 Gb was obtained from a 1009K/mm^2 cluster density, with a cluster passing quality control filters of 91.5% (17 486 000 passing filter-paired reads). Within this run, the index representation for strain AT9 was determined to be 8.38%. The 1 465 998 paired reads were trimmed then assembled in six scaffolds using Spades software [21].

Genome annotation and comparison

Open reading frames (ORFs) were predicted using Prodigal [22] with default parameters. Nevertheless, the predicted ORFs were excluded if they spanned a sequencing gap region (contains N). The predicted bacterial protein sequences were searched against the GenBank and Clusters of Orthologous Groups (COGs) databases using BLASTP (*E* value 1e-03 coverage). If no hit was found, it searched against the nr (nonredundant) database using BLASTP with an *E* value of 1e-03, coverage 70% and identity 30%. If the sequence length was smaller than 80 amino acids, we used an *E* value of 1e-05. The tRNAs and rRNAs were predicted using the tRNA Scan-SE and RNAmmer tools respectively [23,24]. Phobius was used to foresee the signal peptides and number of transmembrane helices respectively [25]. Mobile genetic elements were foretold using PHAST and RAST [26,27]. ORFans were identified if none of the BLASTP runs provided positive results (*E* value was lower than 1e-03 for an alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, we used an *E* value of 1e-05). Artemis and DNA Plotter were used for data management and visualization of genomic features respectively [28,29]. Genomes were automatically retrieved from the 16S rRNA tree using Xegen software (PhyloPattern) [30]. For each selected genome, complete genome sequence, proteome genome sequence and Orfeome genome sequence were retrieved from the NCBI FTP site. All proteomes were analysed 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 couple of orthologous genes between the two genomes studied (average genomic identity of orthologous gene sequences (AGIOS)) [7]. For the genomic comparison of strain AT9, we used *Butyricimonas virosa* (type) strain JCM15149T (Genbank project number: JAEW00000000), *Odoribacter laneus* strain YIT12061 (ADMC00000000), *Bacteroides plebeius* strain DSM17135

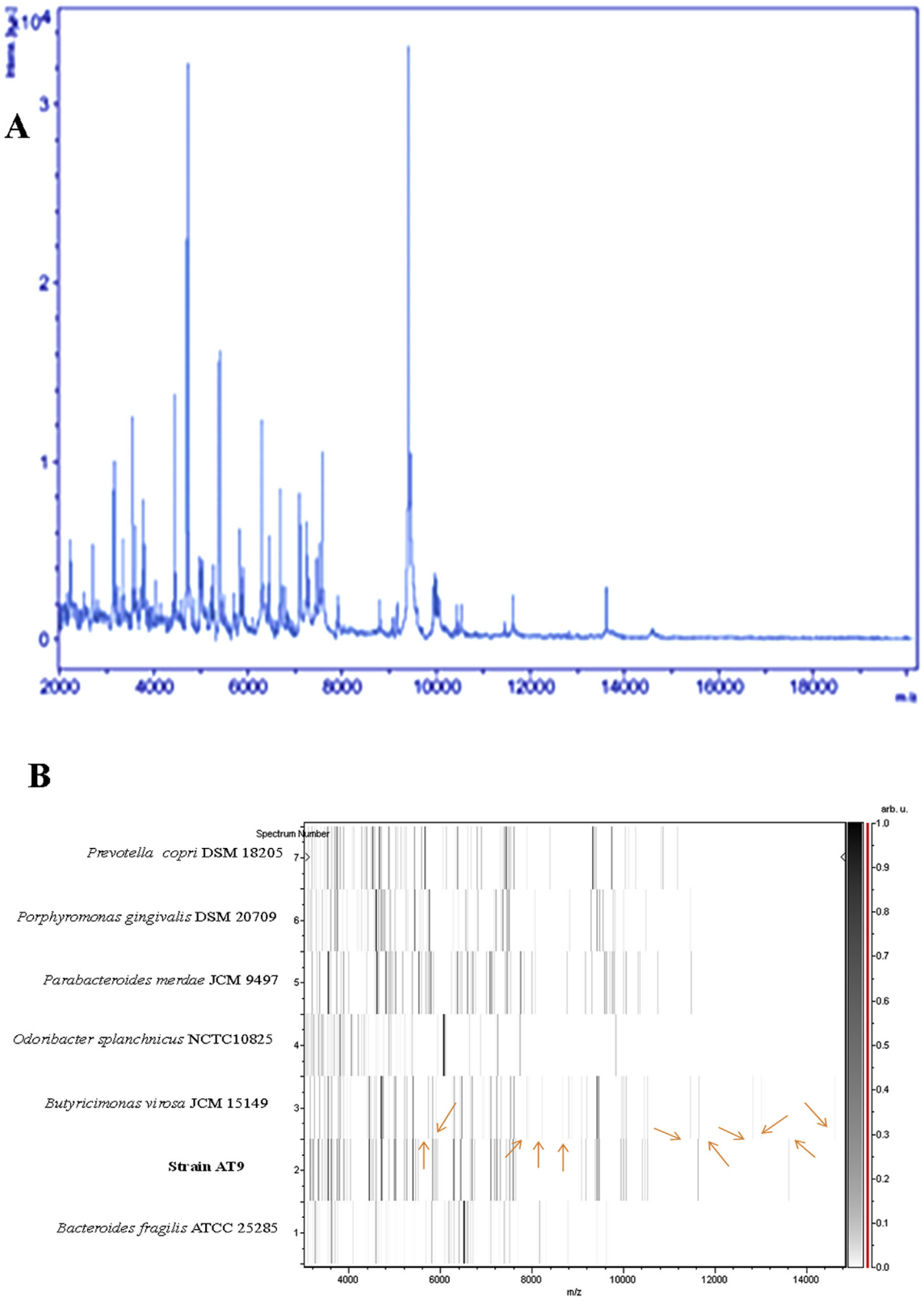

FIG. 1. MALDI-TOF MS analysis of *Butyricimonas phoceensis* strain AT9. (a) Reference mass spectrum from strain AT9. (b) Gel view comparing strain AT9 to other close species. Gel view displays raw spectra of loaded spectrum files arranged in pseudo–gel-like look. The *x*-axis records m/z value. Left *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 is displayed with and peak intensity in arbitrary units. Displayed species are indicated at left. Arrows indicated discordant peaks between strain AT9 and its closest phylogenetic neighbour, *Butyricimonas virosa*.

(ABQC00000000), *Paraprevotella clara* strain YIT11840 (AFFY00000000), *Parabacteroides merdae* ATCC43184 (AAXE00000000), *Porphyromonas catoniae* ATCC 51270 (JDFF00000000) and *Odoribacter splanchnicus* strain DSM20712 (CP002544). 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). The genome of strain AT9 was locally aligned pairwise using the BLAT algorithm [28,29] against each of the selected genomes previously cited, and DDH values were estimated from a generalized model [32]. Annotation and comparison processes were performed in the multiagent software system DAGOBAH [33], which includes Figenix [34] libraries that provide pipeline analysis.

Results

Phylogenetic analysis

The spectrum generated from clean strain AT9 spots did not match with those identified from the Bruker database even when two strains of *Butyricimonas virosa*, including the type strain (JCM15149T), were included in the database (Fig. 1a). The phylogenetic analysis, performed using 16S rRNA gene sequences, showed that our strain AT9 exhibited 98.3, 97.8, 97.5 and 94.2% similarity with *Butyricimonas virosa* JCM 15149T, *Butyricimonas faecihominis* JCM 18676T, *Butyricimonas paravirosa* JCM 18677T and *Butyricimonas synergistica* JCM 15148T respectively [9,10] (Table 1). However, this percentage remains lower than the 98.7% 16S rRNA gene sequence threshold recommended by Kim *et al.* [35] to delineate a new species. The neighbour-joining phylogenetic tree (Fig. 2), based on 16S rRNA gene sequences, shows the relationships between strain AT9 and some related taxa. The 16S rRNA sequence of strain AT9 was deposited in European Molecular Biology Laboratory–European Bioinformatics Institute (EMBL-EBI) under accession number LN881597. A gel view was performed in order to see the spectra differences of strain AT9 with other related bacteria. Eleven discordant peaks were found when we compared strain AT9 and the *B. virosa* JCM15149T profile (Fig. 1b).

Phenotypic and biochemical characterization

The growth of strain AT9 occurred between 28 to 37°C, but optimal growth was observed at 37°C after 48 hours' incubation in anaerobic atmosphere. It is an anaerobic bacillus, but it

TABLE 1. Percentage 16S rRNA gene similarity within *Butyricimonas* genus

	B. faecihominis JCM 18676T	*B. paravirosa* JCM 18677T	*B. synergistica* JCM 15148T	*B. virosa* JCM 15149T	*B. phoceensis* strain AT9
B. faecihominis JCM 18676T	100	97.30	94.07	96.84	97.77
B. paravirosa JCM 18677T		100	94.75	96.84	97.51
B. synergistica JCM 15148T			100	94.22	94.20
B. virosa JCM 15149T				100	98.38
B. phoceensis strain AT9					100

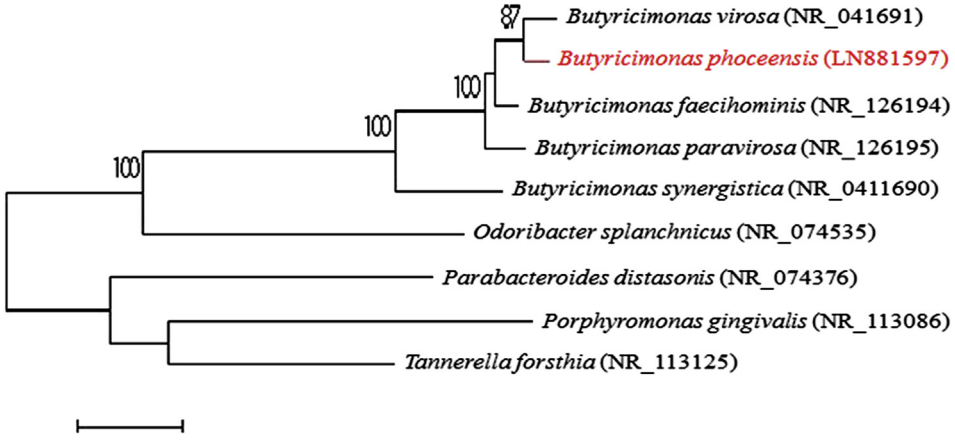

FIG. 2. Phylogenetic tree based on 16S rRNA highlighting position of *Butyricimonas phoceensis* strain AT9 relative to other close species. Corresponding GenBank accession numbers for 16S rRNA genes are indicated at right of strains in tree. Sequences were aligned using Muscle v3.8.31 with default parameters, and phylogenetic inferences were obtained using neighbour-joining method with 500 bootstrap replicates within MEGA6 software. Scale bar represents 2% nucleotide sequence divergence.

can also grow in microaerophilic atmospheres at 28°C. The colonies were ~1 to 2 mm in diameter and opalescent on 5% sheep's blood–enriched Columbia agar. Growth of this isolate was observed using 5 g of salt on Schaedler agar with 5% sheep's blood but not with 10 g/L of salt. This bacterium is not able to form spores. It is a Gram-negative stain (Fig. 3a); it is a motile rod-shaped bacterium that is catalase positive and oxidase negative. Cell diameter ranges 0.5 to 1.5 µm, with a mean diameter of 1 µm by electron microscopy (Fig. 3b). Table 2 summarizes the classification and main features of strain AT9.

Using the API ZYM strip, we observed that strain AT9 possesses alkaline phosphatase, esterase (C4), esterase lipase (C8), naphthol-AS-BI-phosphohydrolase, phosphatase acid and *N*-acetyl-β-glucosaminidase activities; there were no activities for the other enzymes tested. Using API 20A strip, positive reactions were obtained for indole, D-glucose, D-lactose, glycerol and D-mannose. Using the API 50 CH strip, positive reactions were observed only with esculin ferric citrate and potassium 2-ketogluconate. The differences of characteristics compared to other representatives of the genus *Butyricimonas* are detailed in Table 3.

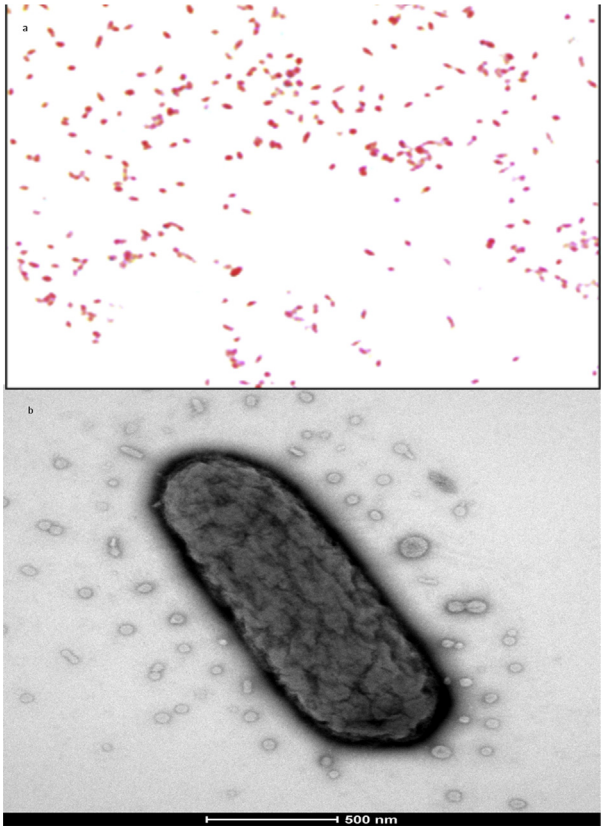

FIG. 3. Phenotypic features of *Butyricimonas phoceensis* strain AT9. (a) Gram stain. (b) Transmission electron microscopy using Tecnai G20 (FEI Company) at operating voltage of 60 kV. Scale bar = 500 nm.

TABLE 2. **Classification and general features of *Butyricimonas phoceensis* strain AT9**

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Bacteroidetes* Class: *Bacteroidia* Order: *Bacteroidales* Family: *Porphyromonadaceae* Genus: *Butyricimonas* Species: *B. phoceensis* Type strain: AT9
Gram stain	Negative
Cell shape	Rod
Motility	Motile
Sporulation	Non–spore forming
Temperature range	Mesophile
Optimum temperature	37°C
Oxygen requirement	Anaerobic
Carbon source	Unknown
Energy source	Unknown
Habitat	Human gut
Biotic relationship	Free living
Pathogenicity	Unknown
Isolation	Human faeces

Of the 21 antibiotics tested, strain AT9 was susceptible to gentamicin 500 µg, vancomycin, doxycycline, trimethoprim–sulfamethoxazole, rifampicin, amoxicillin 25 µg/mL, metronidazole 4 µg/mL, amoxicillin–clavulanic acid 30 µg/mL, imipenem 10 µg/mL, penicillin G, teicoplanin and doripenem 10 µg/mL and was resistant to erythromycin, oxacillin, gentamicin 15 µg, colistin, ceftriaxone, ciprofloxacin, clindamycin, dalacin 15 µg/mL and fosfomycin. Analysis of the total cellular fatty acid composition demonstrated that the major fatty acid detected was the branched iso-C15:0 acid (62.3%). Hydroxy and cyclo fatty acids were also detected (Table 4).

Genome properties

The draft genome of strain AT9 (Fig. 4) (accession no. FBYB00000000) is 4 736 949 bp long with 42.51% G+C content (Table 5). It is composed of six scaffolds comprising seven contigs. Of the 4007 predicted genes, 3947 were protein-coding genes and 60 were RNAs (four genes 5S rRNA, one 16S rRNA, one 23S rRNA and 54 tRNA). A total of 2386 genes (60.45%) were assigned as putative functions (by COGs or by NR BLAST), 178 genes (4.51%) were identified as ORFans and ten genes were associated with polyketide synthase or non-ribosomal peptide synthetase [36]. Using ARG-ANNOT [37], three genes associated with resistance were found, including *TetQ, TetX* (which confers resistance to tetracycline) and *ErmF* (which confers resistance to erythromycin). This could represent the *in silico/in vitro* discordance for antibiotic resistance prediction, as strain AT9 was resistant to erythromycin but susceptible to doxycycline. The remaining 1316 genes (33.34%) were annotated as hypothetical proteins. Genome statistics are provided in Table 5. Table 6 lists the distribution of genes into COGs functional categories of strain AT9.

TABLE 3. Differential characteristics of strain *Butyricimonas phoceensis* strain AT9 with *Butyricimonas* species

Property	Strain AT9	*B. virosa*	*B. faecihominis*	*B. paravirosa*	*B. synergistica*
Cell diameter width/length (μm)	0.5/1.75	0.6–0.8/2.5–5	0.7–1/3–5	0.8–1/2–12.4	0.5–1/3–6
Oxygen requirement	–	–	–	–	–
Gram stain	–	–	–	–	–
Motility	+	–	–	–	–
Spore formation	–	–	–	–	–
Production of:					
Catalase	+	+	+	+	–
Oxidase	–	–	–	–	–
Urease	–	–	–	–	–
Indole	+	+	+	+	+
β-Galactosidase	+	+	+	+	+
N-acetyl-glucosaminidase	+	+	+	+	+
Utilization of:					
L-Arabinose	–	–	+	–	–
D-Mannose	+	–	+	+	+
D-Mannitol	–	–	–	–	–
D-Glucose	+	+	+	+	+
D-Maltose	–	–	+	–	+
Isolation source	Human faeces	Rat faeces	Human faeces	Rat faeces	Human faeces
DNA G+C content (mol%)	42.5	46.5	45.2	44.9	46.4

TABLE 4. Cellular fatty acid profiles of strain *Butyricimonas phoceensis* strain AT9 compared to other closely related *Butyricimonas* species

Fatty acid	Strain AT9	*B. faecihominis* JCM 18676T	*B. paravirosa* JCM 18677T	*B. synergistica* JCM 15148T	*B. virosa* JCM 15149T
C4:0	TR	NA	NA	NA	NA
C12: 0	NA	TR	TR	NA	NA
C14:0	TR	TR	1.8	NA	1.3
C15: 0	TR	TR	NA	NA	NA
C16:0	3.7	2.8	3.2	2.4	2.1
C18:0	TR	TR	TR	1.0	TR
iso-C5: 0	2.9	NA	NA	NA	NA
iso-C11: 0	NA	TR	TR	NA	NA
iso-C13: 0	NA	1.0	1.0	NA	TR
iso-C15: 0	62.3	64.6	57.6	61.8	68.6
anteiso-C15: 0	1.2	1.8	1.7	2.0	1.5
iso-C17: 0	NA	1.0	TR	NA	TR
C14:0 3-OH	TR	NA	NA	NA	NA
C16: 0 3-OH	4.8	1.7	6.3	1.6	5.2
C17: 0 3-OH	9.0	NA	NA	NA	NA
iso-C15: 0 3-OH	NA	TR	1.8	1.6	1.7
iso-C17: 0 3-OH	NA	5.3	10.6	14.9	10.4
C18: 2n6	2.9	NA	NA	NA	NA
C18:1n5	2.1	NA	NA	NA	NA
C16:1n7	TR	NA	NA	NA	NA
iso-C17: 0	TR	1.0	TR	NA	TR
iso-C15:1n5	TR	NA	NA	NA	NA
C18: 1ω9c	NA	8.3	9.5	12.6	6.0
C18: 2ω6, 9c	NA	1.4	1.5	2.3	1.2
C9, 10-methylene C16:0	7.0	NA	NA	NA	NA

Number are percentages. NA, not available; TR, trace amounts <1%.

Genome comparison

The draft genome (4.74 Mb) sequence of strain AT9 is smaller than those of *Butyricimonas synergistica* (4.77 Mb), but larger than those of *Butyricimonas virosa, Porphyromonas catoniae, Bacteroides plebeius, Paraprevotella clara, Odoribacter laneus, Parabacteroides merdae* and *Odoribacter splanchnicus* (4.72, 2.04, 3.27, 3.65, 4.43, 3.77 and 4.39 MB respectively).

The G+C content of strain AT9 (42.5%) is smaller than those of *Butyricimonas virosa, Odoribacter splanchnicus, Bacteroides plebeius, Parabacteroides merdae, Paraprevotella clara, Butyricimonas synergistica* and *Porphyromonas catoniae* (46.5, 43.4, 44.3, 44.8, 45.3, 48.1, 46.4 and 51.0% respectively) but larger than those of *Odoribacter laneus* (40.55). Fig. 5 shows that the distribution of genes into COGs categories was similar in all genomes compared. In addition, strain AT9 shared 2297, 1535, 742, 1720, 999, 1173, 2108 and 960 orthologous genes with *B. virosa*, *O. laneus*, *P. catoniae*, *O. splanchnicus*, *B. plebeius*, *P. merdae*, *B. synergistica* and *P. clara* respectively (Table 6). Accordingly, strain AT9 has 1650 (42%) of 3947 orthologous proteins not shared with its closest phylogenetic neighbour, *B. virosa*. The AGIOS values ranged from 53.3 to 76.2% among the compared closest species except strain AT9. When strain AT9 was compared to other close species, the AGIOS values ranged from 53.5% with *P. catoniae* to 97.7% with *B. virosa* (Table 7).

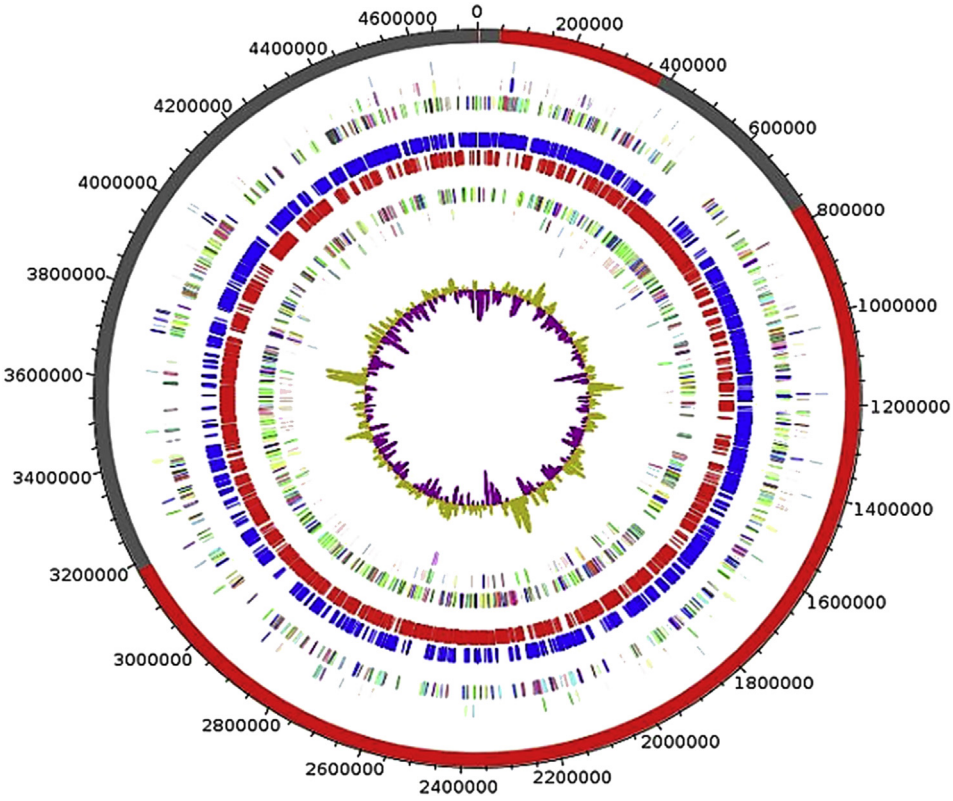

FIG. 4. Graphical circular map of genome of *Butyricimonas phoceensis* strain AT9. From outside to centre: contigs (red/grey), COGs category of genes on forward strand (three circles), genes on forward strand (blue circle), genes on reverse strand (red circle), COGs category on reverse strand (three circles), GC content.

The DDH value was 80.2% ± 2.7 with *B. virosa*, 17.7% ± 2.2 with *O. laneus*, 21.4% ± 2.3 with *B. plebeius*, 20.2% ± 2.3 with *P. clara*, 19.1% ± 2.2 with *P. merdae*, 18.3% ± 2.2 with *P. catoniae* and 17.3% ± 2.2 with *O. splanchnicus* (Table 8).

TABLE 5. Nucleotide content and gene count levels of the genome of *Butyricimonas phoceensis* strain AT9

	Genome (total)	
Attribute	**Value**	**% of total**
Size (bp)	4 736 949	100
G+C content (bp)	2 013 756	42.51
Coding region (bp)	4 330 163	91.40
Total genes	4007	100
RNA genes	60	1.50
Protein-coding genes	3947	98.50
Genes with function prediction	2386	60.45
Genes assigned to COGs	1880	47.63
Genes with peptide signals	1185	30.02
Gene associated to PKS or NRPS	10	0.25
Genes associated to ORFan	178	4.51
Genes associated to mobilome	1109	28.10
Genes associated to toxin/antitoxin	70	1.8
Genes associated to resistance genes	3	0.076
Genes with paralogues (*E* value 1e-10)	1449	36.71
Genes with paralogues (*E* value 1e-25)	1098	27.82
Gene associated to hypothetical protein	1316	33.34
Genes larger than 5000 nucleotides	5	0

COGs, Clusters of Orthologous Groups database; PKS, polyketide synthase; NRPS, nonribosomal peptide synthase.

Discussion

Strain AT9 is part of an exploratory culturomics study of the gut flora from obese patients before and after bariatric surgery. The aim of culturomics is to exhaustively explore the microbial ecosystem of gut flora by using different culture conditions followed by MALDI-TOF MS identification [2]. The phylogenetic analysis, performed using 16S rRNA sequences, showed that strain AT9 exhibited 98.3% similarity with *Butyricimonas virosa*. However, this percentage remains lower than the 98.7% 16S rRNA gene sequence threshold recommended to delineate a new species [3,38].

The genus *Butyricimonas* was established in 2009 by Sakamoto and includes four described species [9–11]. All the species of the genus *Butyricimonas* are anaerobic. These bacteria are isolated in human or rat faeces. To evaluate the genomic similarity with other closest species, we determined two parameters: DDH [39] and AGIOS [7]. Although the values of DDH (80.2%) and AGIOS (97.7%) were very high between strain AT9 and *Butyricimonas virosa* (type strain JCM15149T), we found several discrepancies justifying the description of a new species, including motility, D-mannose utilization (absent in

TABLE 6. Number of genes associated with the 25 general COGs functional categories of *Butyricimonas phoceensis* strain AT9

Code	Value	% value	Description
J	193	4.89	Translation
A	0	0	RNA processing and modification
K	192	4.87	Transcription
L	111	2.81	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	23	0.58	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	85	2.15	Defence mechanisms
T	174	4.41	Signal transduction mechanisms
M	200	5.06	Cell wall/membrane biogenesis
N	20	0.51	Cell motility
Z	4	0.10	Cytoskeleton
W	3	0.07	Extracellular structures
U	28	0.71	Intracellular trafficking and secretion
O	91	2.30	Posttranslational modification, protein turnover, chaperones
X	32	0.81	Mobilome: prophages, transposons
C	122	3.09	Energy production and conversion
G	92	2.33	Carbohydrate transport and metabolism
E	120	3.04	Amino acid transport and metabolism
F	60	1.52	Nucleotide transport and metabolism
H	99	2.51	Coenzyme transport and metabolism
I	69	1.75	Lipid transport and metabolism
P	199	5.04	Inorganic ion transport and metabolism
Q	26	0.66	Secondary metabolites biosynthesis, transport and catabolism
R	150	3.80	General function prediction only
S	67	1.69	Function unknown
—	2067	52.36	Not in COGs

COGs, Clusters of Orthologous Groups database.

B. virosa but present in *B. paravirosa, B. synergistica* and *B. faecihominis*), MALDI-TOF MS spectrum (11 different peaks), different GC% (42.5 vs. 46.5% for *B. virosa*), high proportion of orthologous proteins not shared between the two species (1650/3947 (42%)) and different COGs repartition ((D) cell cycle control 110 vs. 124, (P) transport of inorganic ions 32 vs. 20 for strain AT9 and *B. virosa* respectively).

Conclusion

On the basis of phenotypic, chemotaxonomic, phylogenetic and genomic information, a novel species belonging to the genus *Butyricimonas* is proposed with the name *Butyricimonas phoceensis* sp. nov. The type strain is AT9. This bacterium was isolated from the faeces of a 57-year-old obese French woman living in Marseille after bariatric surgery. The isolation of this new species demonstrates that microbial culturomics extends the repertoire of human gut anaerobes, which are of critical importance to decipher the links among gut microbiota, health and disease, including obesity.

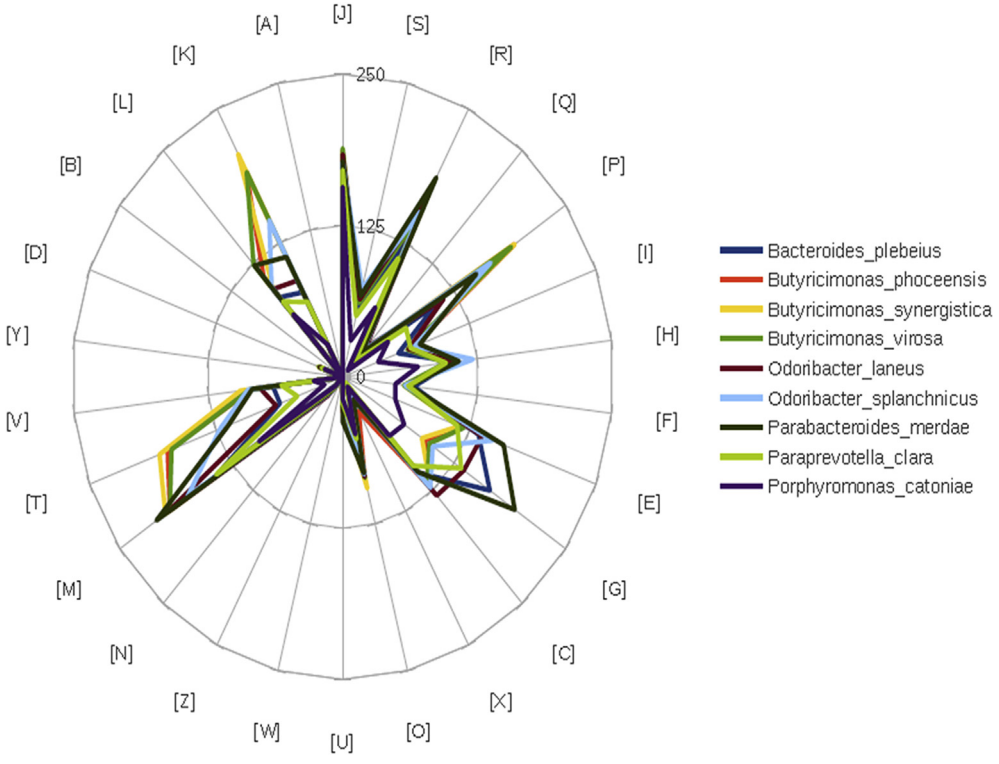

FIG. 5. Distribution of functional classes of predicted genes according to clusters of orthologous groups of proteins from *Butyricimonas phoceensis* strain AT9.

TABLE 7. Numbers of orthologous proteins shared between genomes (upper right), average percentage similarity of nucleotides corresponding to orthologous protein shared between genomes (lower left) and numbers of proteins per genome (bold)

	Odoribacter laneus	*Butyricimonas phoceensis* strain AT9	*Porphyromonas catoniae*	*Odoribacter splanchnicus*	*Bacteroides plebeius*	*Parabacteroides merdae*	*Butyricimonas virosa*	*Butyricimonas synergistica*	*Paraprevotella clara*
O. laneus	**3103**	1535	745	1472	1005	1187	1519	1480	964
B. phoceensis strain AT9	57.52	**3947**	742	1720	999	1173	2297	2108	960
P. catoniae	53.37	53.50	**1597**	737	726	826	729	725	746
O. splanchnicus	59.11	68.17	53.65	**3497**	977	1149	1702	1604	960
B. plebeius	55.52	62.17	55.11	62.84	**2643**	1175	986	963	1059
P. merdae	55.38	63.08	55.30	63.65	66.34	**4384**	1154	1130	1123
B. virosa	57.47	97.79	53.41	68.18	62.13	62.87	**3934**	2086	950
B. synergistica	57.22	76.18	53.75	68.24	62.15	62.93	76.24	**3874**	926
P. clara	54.31	61.92	54.84	62.34	68.02	65.40	61.87	62.09	**2847**

TABLE 8. Pairwise comparison of *Butyricimonas phoceensis* strain AT9 with other species using GGDC, formula 2 (DDH estimates based on identities/HSP length)[a]

	Strain AT9	*Odoribacter laneus*	*Bacteroides plebeius*	*Butyricimonas virosa*	*Paraprevotella clara*	*Parabacteroides merdae*	*Porphyromonas catoniae*	*Odoribacter splanchnicus*
Strain AT9	100% ± 0	17.7% ± 2.2	21.4% ± 2.3	80.2% ± 2.7	20.2% ± 2.3	19.1% ± 2.3	18.3% ± 2.3	17.3% ± 2.2
O. laneus		100% ± 0	19% ± 2.3	18.2% ± 2.3	20.5% ± 2.3	18.9% ± 2.3	19.6% ± 2.3	18.2% ± 2.3
B. plebeius			100% ± 0	19.9% ± 2.3	20.3% ± 2.3	21.5% ± 2.3	17.6% ± 2.2	18.4% ± 2.3
B. virosa				100% ± 0	20.3% ± 2.3	19.4% ± 2.3	19.0% ± 2.3	17.4% ± 2.2
P. clara					100% ± 0	18.9% ± 2.3	17.8% ± 2.2	17.7% ± 2;2
P. merdae						100% ± 0	17.6% ± 2.2	21.5% ± 2.3
P. catoniae							100% ± 0	18% ± 2.2
O. splanchnicus								100% ± 0

DDH, DNA-DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; HSP, high-scoring segment pairs.
[a]Confidence intervals indicate inherent uncertainly in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size). These results are in accordance with 16S rRNA and phylogenomic analyses as well as GGDC results.

Taxonomic and nomenclatural proposals

Description of strain AT9 sp. nov.. Butyricimonas phoceensis (*pho.ce.en.sis*, N.L. gen. n. *phoceensis*, based on the acronym of the Phocean city where the type strain was isolated). Cells are Gram-negative-staining, non–spore forming, motile, rod-shaped bacilli, with a size of 0.5 to 1.5 µm in diameter. Colonies are opalescent with a diameter of 1 to 2 mm on 5% sheep's blood–enriched Columbia agar. The strain is oxidase negative and catalase positive. It has an optimum growth temperature of 37°C and is anaerobic, but it is able to grow in microaerophilic condition at 28°C. Using API Gallery systems, positive reactions were observed for alkaline phosphatase, esterase (C4), esterase lipase (C8), naphthol-AS-BI-phosphohydrolase, phosphatase acid, *N*-acetyl-β-glucosaminidase, indole, D-glucose, D-lactose, glycerol and D-mannose, esculin ferric citrate and potassium 2-ketogluconate. Cells are susceptible to gentamicin 500 µg, vancomycin, doxycycline, trimethoprim–sulfamethoxazole, rifampicin, penicillin G and teicoplanin. The major fatty acid detected was iso-C15:0. The length of the genome is 4 736 949 bp with 42.51% G+C content. The 16S rRNA gene sequence and whole-genome shotgun sequence of *B. phoceensis* strain AT9 were deposited in EMBL-EBI under accession numbers LN881597 and FBYB00000000, respectively. The type strain AT9 (= CSUR P2478 = DSM 100838) was isolated from the stool sample of a French obese woman. The habitat of this microorganism is the human digestive gut.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process and K. Griffiths for English-language review. This study was funded by the Fondation Méditerranée Infection.

Conflict of Interest

None declared.

References

[1] Zhang H, DiBaise JK, Zuccolo A, Kudrna D, Braidotti M, Yu Y, et al. Human gut microbiota in obesity and after gastric bypass. Proc Natl Acad Sci U S A 2009;106:2365–70.

[2] 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.

[3] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[4] Garrity GM, Trüper HG, Whitman WB, Grimont PAD, Nesme X, Frederiksen W, et al. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. Int J Syst Evol Microbiol 2002;52:1043–7.
[5] Rosselló-Mora R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Molecular identification, systematics, and population structure of prokaryotes. New York: Springer; 2006. p. 23–50.
[6] Welker M, Moore ERB. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[7] 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.
[8] Kokcha S, Mishra AK, Lagier JC, Million M, Leroy Q, Raoult D, et al. Non-contiguous-finished genome sequence and description of *Bacillus timonensis* sp. nov. Stand Genomic Sci 2012;6:346–55.
[9] Mishra AK, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of *Peptoniphilus senegalensis* sp. nov. Stand Genomic Sci 2013;7:370–81.
[10] Sakamoto M, Takagaki A, Matsumoto K, Kato Y, Goto K, Benno Y. *Butyricimonas synergistica* gen. nov., sp. nov. and *Butyricimonas virosa* sp. nov., butyric acid–producing bacteria in the family *'Porphyromonadaceae'* isolated from rat faeces. Int J Syst Evol Microbiol 2009;59(Pt 7):1748–53.
[11] Sakamoto M, Tanaka Y, Benno Y, Ohkuma M. *Butyricimonas faecihominis* sp. nov. and *Butyricimonas paravirosa* sp. nov., isolated from human faeces, and emended description of the genus *Butyricimonas*. Int J Syst Evol Microbiol 2014;64(Pt 9):2992–7.
[12] Sakamoto M. The Family *Porphyromonadaceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes—other major lineages of bacteria and the Archaea. Berlin: Springer; 2014. p. 811–24.
[13] Ulger Toprak N, Bozan T, Birkan Y, Isbir S, Soyletir G. *Butyricimonas virosa*: the first clinical case of bacteraemia. New Microbes New Infect 2015;4:7–8.
[14] Ferry T, Laurent F, Ragois P, Chidiac C, Lyon BJI Study Group. Post-traumatic chronic bone and joint infection caused by *Butyricimonas* spp., and treated with high doses of ertapenem administered subcutaneously in a 30-year-old obese man. BMJ Case Rep 2015;2015:212359.
[15] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543–51.
[16] Hugon P, Ramasamy D, Lagier JC, Rivet R, Couderc C, Raoult D, et al. Non-contiguous-finished genome sequence and description of *Alistipes obesi* sp. nov. Stand Genomic Sci 2013;7:427–39.
[17] Nkamga VD, Huynh HTT, Aboudharam G, Ruimy R, Drancourt M. Diversity of human-associated *Methanobrevibacter smithii* isolates revealed by multispacer sequence typing. Curr Microbiol 2015;70:810–5.
[18] Le Page S, van Belkum A, Fulchiron C, Huguet R, Raoult D, Rolain JM. Evaluation of the PREVI® Isola automated seeder system compared to reference manual inoculation for antibiotic susceptibility testing by the disk diffusion method. Eur J Clin Microbiol Infect Dis 2015;34:1859–69.
[19] Sasser M. Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). Technical note 101. Newark, DE: MIDI; 2006.
[20] 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.
[21] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol J Comput Mol Cell Biol 2012;19:455–77.
[22] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010;11:119.
[23] 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.
[24] Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[25] Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–36.
[26] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. Nucleic Acids Res 2011;39(Web Server issue):W347–52.
[27] Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 2014;42(Database issue):D206–14.
[28] Kent WJ. BLAT—the BLAST-like alignment tool. Genome Res 2002;12:656–64.
[29] 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;2:117–34.
[30] Gouret P, Thompson JD, Pontarotti P. PhyloPattern: regular expressions to identify complex patterns in phylogenetic trees. BMC Bioinformatics 2009;10:298.
[31] Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[32] Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013;14:60.
[33] Gouret P, Paganini J, Dainat J, Louati D, Darbo E, et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBAH. In: Springer-Verlag, editor. Evolutionary biology—concepts, biodiversity, macroevolution and genome evolution. Amsterdam: Springer-Verlag; 2011. p. 71–87.
[34] Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, Danchin EG. FIGENIX: intelligent automation of genomic annotation: expertise integration in a new software platform. BMC Bioinformatics 2005;6:198.
[35] Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 2014;64(Pt 2):346–51.
[36] Conway KR, Boddy CN. ClusterMine360: a database of microbial PKS/NRPS biosynthesis. Nucleic Acids Res 2013;41(Database issue):D402–7.
[37] 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.
[38] Konstantinidis KT, Stackebrandt E. Defining taxonomic ranks. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E, editors. The Prokaryotes. New York: Springer; 2006.
[39] Auch AF, Klenk HP, Göker M. Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci 2010;2:142–8.