

## Genome sequence and description of *Bacteroides timonensis* sp. nov.

Dhamodharan Ramasamy<sup>1</sup>, Jean-Christophe Lagier<sup>1</sup>, Morgane Rossi-Tamisier<sup>1</sup>, Anne Pfeiderer<sup>1</sup>, Caroline Michelle<sup>1</sup>, Carine Couderc<sup>1</sup>, Didier Raoult<sup>1,2</sup> and Pierre-Edouard Fournier<sup>1\*</sup>

<sup>1</sup> Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, Aix-Marseille Université, Marseille, France

<sup>2</sup> King Fahd Medical Research Center, King Abdul Aziz University, Jeddah, Saudi Arabia

\*Correspondence: Pierre-Edouard Fournier (pierre-edouard.fournier@univ-amu.fr)

Keywords: *Bacteroides timonensis*, genome, culturomics, taxono-genomics

*Bacteroides timonensis* strain AP1<sup>T</sup> (= CSUR P194 = DSM 26083) is the type strain of *B. timonensis* sp. nov. This strain, whose genome is described here, was isolated from the fecal flora of a 21-year-old French Caucasoid female who suffered from severe anorexia nervosa. *Bacteroides timonensis* is a Gram-negative, obligate anaerobic bacillus. Here we describe the features of this organism, together with the complete genome sequence and annotation. The 7,130,768 bp long genome (1 chromosome, no plasmid) exhibits a G+C content of 43.3% and contains 5,786 protein-coding and 59 RNA genes, including 2 rRNA genes.

### Introduction

*Bacteroides timonensis* strain AP1<sup>T</sup> (= CSUR P194 = DSM 26083) is the type strain of *B. timonensis* sp. nov. This bacterium was isolated from the stool sample of a 21-year-old French Caucasoid female in an effort of cultivating individually all bacterial species within human feces [1]. It is a Gram-negative, anaerobic, indole-positive rod-shaped bacillus.

The conventional genetic parameters used in the delineation of bacterial species include 16S rRNA sequence identity and phylogeny [2,3], genomic G + C content diversity and DNA–DNA hybridization (DDH) [4,5]. These tools have limitations, notably because their cutoff values vary across species or genera [6]. With the introduction of high-throughput sequencing techniques [7], a wealth of genomic data was made available for many bacterial species. We recently proposed to include genomic data in a polyphasic approach to describe new bacterial taxa (taxono-genomics) [8]. This strategy combines phenotypic characteristics, notably the MALDI-TOF MS spectrum, and genomic analysis [8–37].

Here, we present a summary classification and a set of features for *B. timonensis* sp. nov. strain AP1<sup>T</sup> (= CSUR P194 = DSM 26083) together with the description of the complete genome sequencing and annotation. These characteristics

support the circumscription of the type species, *B. timonensis*.

The genus *Bacteroides* (Castellani and Chalmers 1919) was created in 1919 [38]. Currently, it is one of the largest genera among the human gut microbiota [39], and consists of 91 species and 5 subspecies with validly published names [40]. *Bacteroides* species are Gram-negative, non-spore-forming, non-motile and anaerobic rods that are generally isolated from the gastrointestinal tract of mammals [41]. They have symbiotic relationships with humans and play many beneficial roles on normal intestinal physiology and function. Several *Bacteroides* species are identified as opportunistic pathogens when isolated from anaerobic infections [42].

### Classification and features

A stool sample was collected from 21-year-old French Caucasoid female who suffered from severe restrictive anorexia nervosa from the age of 12 years. At the time of sample collection, she had been hospitalized for recent aggravation of her medical condition (BMI: 10.4 kg/m<sup>2</sup>). The patient's written consent and the agreement of the local ethics committee of the IFR48 (Marseille, France) were obtained under agreement number 09-022. The feces sample of this patient



was stored at -80°C immediately after collection. Strain AP1<sup>T</sup> (Table 1) was isolated in November 2011 after 1 month of incubation in Columbia agar (BioMerieux, Marcy l'Etoile, France). Several other new bacterial species were isolated from this stool specimen using various culture conditions.

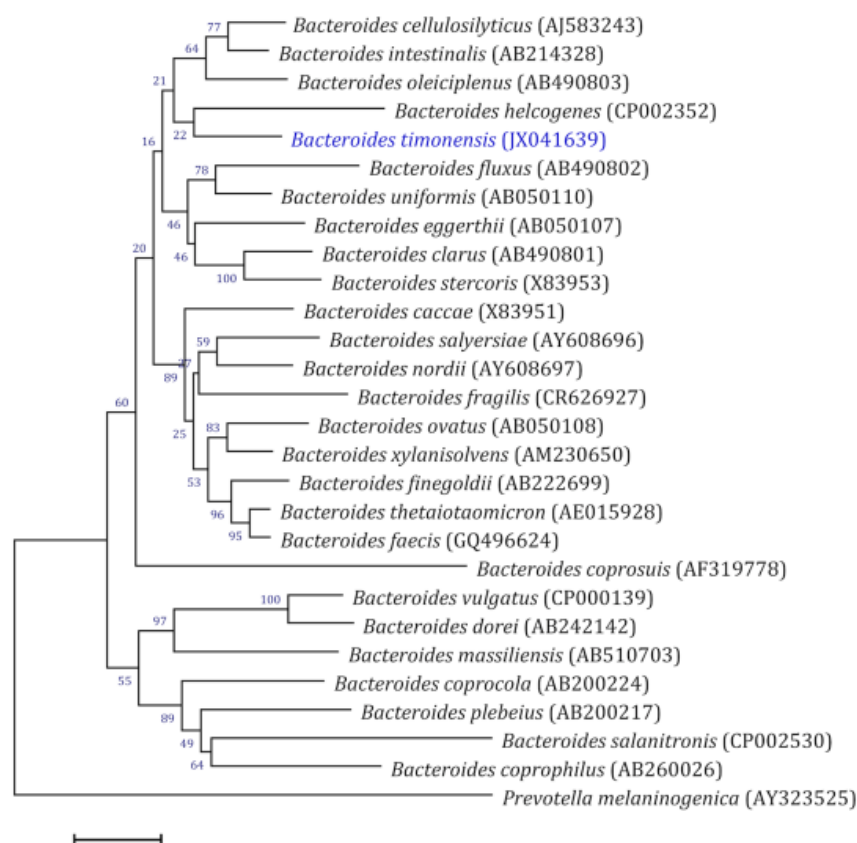
When compared to sequences available in GenBank, the 16S rRNA gene sequence of *B. timonensis* strain AP1<sup>T</sup> (GenBank accession number JX041639) exhibited an identity of

97.00% with *Bacteroides cellulosilyticus* (Figure 1). This value was the highest similarity observed, but was lower than the 97.8% 16S rRNA gene sequence threshold recommended by Stackebrandt and Ebers (2006) to delineate a new species without carrying out DNA-DNA hybridization [3], and was in the 74.8 to 98.7% range of 16S rRNA identity values observed among 41 *Bacteroides* species with validly published names [56].

**Table 1.** Classification and general features of *Bacteroides timonensis* strain AP1<sup>T</sup> according to the MIGS recommendations [43]

| MIGS ID  | Property               | Term                                  | Evidence code <sup>a</sup> |
|----------|------------------------|---------------------------------------|----------------------------|
|          | Current classification | Domain <i>Bacteria</i>                | TAS [44]                   |
|          |                        | Phylum <i>Bacteroidetes</i>           | TAS [45,46]                |
|          |                        | Class <i>Bacteroidia</i>              | TAS [45,47]                |
|          |                        | Order <i>Bacteroidales</i>            | TAS [45,48]                |
|          |                        | Family <i>Bacteroidaceae</i>          | TAS [49,50]                |
|          |                        | Genus <i>Bacteroides</i>              | IDA [49,51-54]             |
|          |                        | Species <i>Bacteroides timonensis</i> | IDA                        |
|          |                        | Type strain AP1 <sup>T</sup>          | IDA                        |
|          |                        | Gram stain                            | Negative                   |
|          |                        | Cell shape                            | Rod                        |
|          | Motility               | Non motile                            | IDA                        |
|          |                        | Sporulation                           | Non sporulating            |
|          |                        | Temperature range                     | Mesophile                  |
|          |                        | Optimum temperature                   | 37°C                       |
|          |                        | Salinity                              | Unknown                    |
| MIGS-6.3 |                        |                                       | IDA                        |
| MIGS-22  | Oxygen requirement     | Anaerobic                             | IDA                        |
|          |                        | Carbon source                         | Unknown                    |
|          |                        | Energy source                         | Unknown                    |
| MIGS-6   | Habitat                | Human gut                             | IDA                        |
| MIGS-15  | Biotic relationship    | Free living                           | IDA                        |
|          |                        | Pathogenicity                         | Unknown                    |
|          |                        | Biosafety level                       | 2                          |
| MIGS-14  | Isolation              | Human feces                           |                            |
| MIGS-4   | Geographic location    | France                                | IDA                        |
|          |                        | Sample collection                     |                            |
| MIGS-5   | time                   | November 2011                         | IDA                        |
| MIGS-4.1 | Latitude               | 43.296482                             | IDA                        |
| MIGS-4.1 | Longitude              | 5.36978                               | IDA                        |
| MIGS-4.3 | Depth                  | surface                               | IDA                        |
| MIGS-4.4 | Altitude               | 0 m above sea level                   | IDA                        |

Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [55]. If the evidence is IDA, then the property was directly observed for a live isolate by one of the authors or an expert mentioned in the acknowledgements.

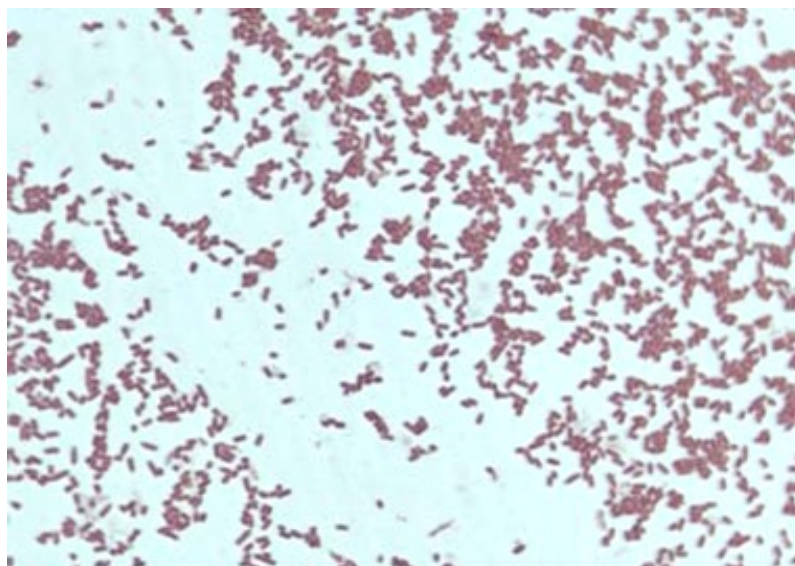


**Figure 1.** Phylogenetic tree highlighting the position of *Bacteroides timonensis* strain AP1<sup>T</sup> relative to other type strains within the *Bacteroides* genus. GenBank accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using the maximum-likelihood method within the MEGA software. Numbers at the nodes are percentages of bootstrap values obtained from 500 replicates. *Prevotella melaninogenica* was used as outgroup. The scale bar represents a 2% nucleotide sequence divergence.

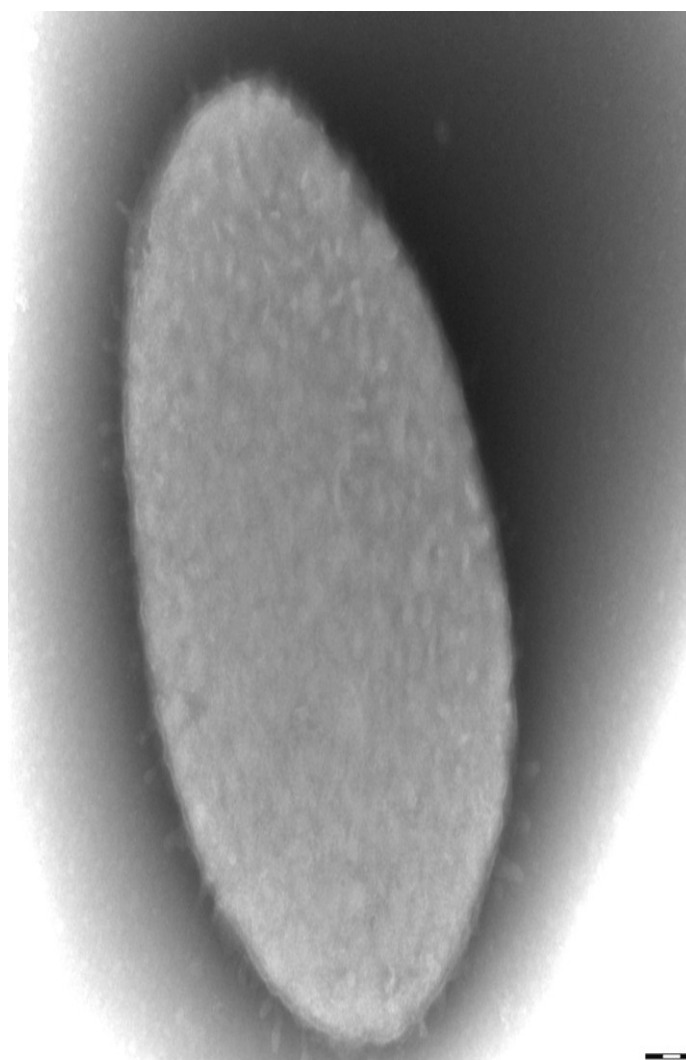
Four different growth temperatures (25, 30, 37, 45°C) were tested; growth occurred between 25 and 37°C, but optimal growth was observed at 37°C, 24 hours after inoculation. No growth occurred at 45°C. Colonies were translucent and approximately 0.3 mm in diameter on 5% sheep blood-enriched Columbia agar (BioMerieux). Growth of the strain was tested in the same agar under anaerobic and microaerophilic conditions using GENbag anaer and GENbag microaer systems, respectively (BioMerieux), and under aerobic conditions, with or without 5% CO<sub>2</sub>. Growth was observed under anaerobic and microaerophilic conditions, and only weakly with 5% CO<sub>2</sub>. No growth occurred under aerobic condition without CO<sub>2</sub>. Gram staining showed short Gram-negative rods unable to form spores (Figure 1). A motility test was negative. Cells grown on agar are translucent and exhibit a mean diameter of 0.88 µm in electron microscopy (Figure 2, Figure 3).

Strain AP1<sup>T</sup> exhibited catalase but no oxidase activity (Table 2). Using an API Rapid ID 32A strip (BioMerieux), positive reactions were obtained

for arginine dihydrolase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, α-arabinosidase, N-acetyl-β-glucosaminidase, glutamic acid decarboxylase, α-fucosidase, nitrate reduction, indole production, alkaline phosphatase, proline arylamidase, leucyl glycine arylamidase, alanine arylamidase, glutamyl glutamic acid arylamidase, and fermentation of mannose and raffinose. Weak activities were observed for glycine arylamidase and serine arylamidase. Negative reactions were obtained for urease, β-galactosidase-6-phosphatase, β-glucuronidase, arginine arylamidase, phenylalanine arylamidase, leucine arylamidase, pyroglutamic acid arylamidase, tyrosine arylamidase and histidine arylamidase. Using an API 50CH strip (Biomerieux), strain AP1<sup>T</sup> was asaccharolytic. *B. timonensis* is susceptible to amoxicillin-clavulanate, ceftriaxon, imipenem, trimethoprim-sulfamethoxazole, metronidazole and doxycycline but resistant to amoxicillin, vancomycin and gentamicin. By comparison with other *Bacteroides* species, *B. timonensis* dif-



**Figure 2.** Gram staining of *B. timonensis* strain AP1<sup>T</sup>



**Figure 3.** Transmission electron microscopy of *B. timonensis* strain AP1<sup>T</sup>, made using a Morgani 268D (Philips) at an operating voltage of 60kV. The scale bar represents 200 nm

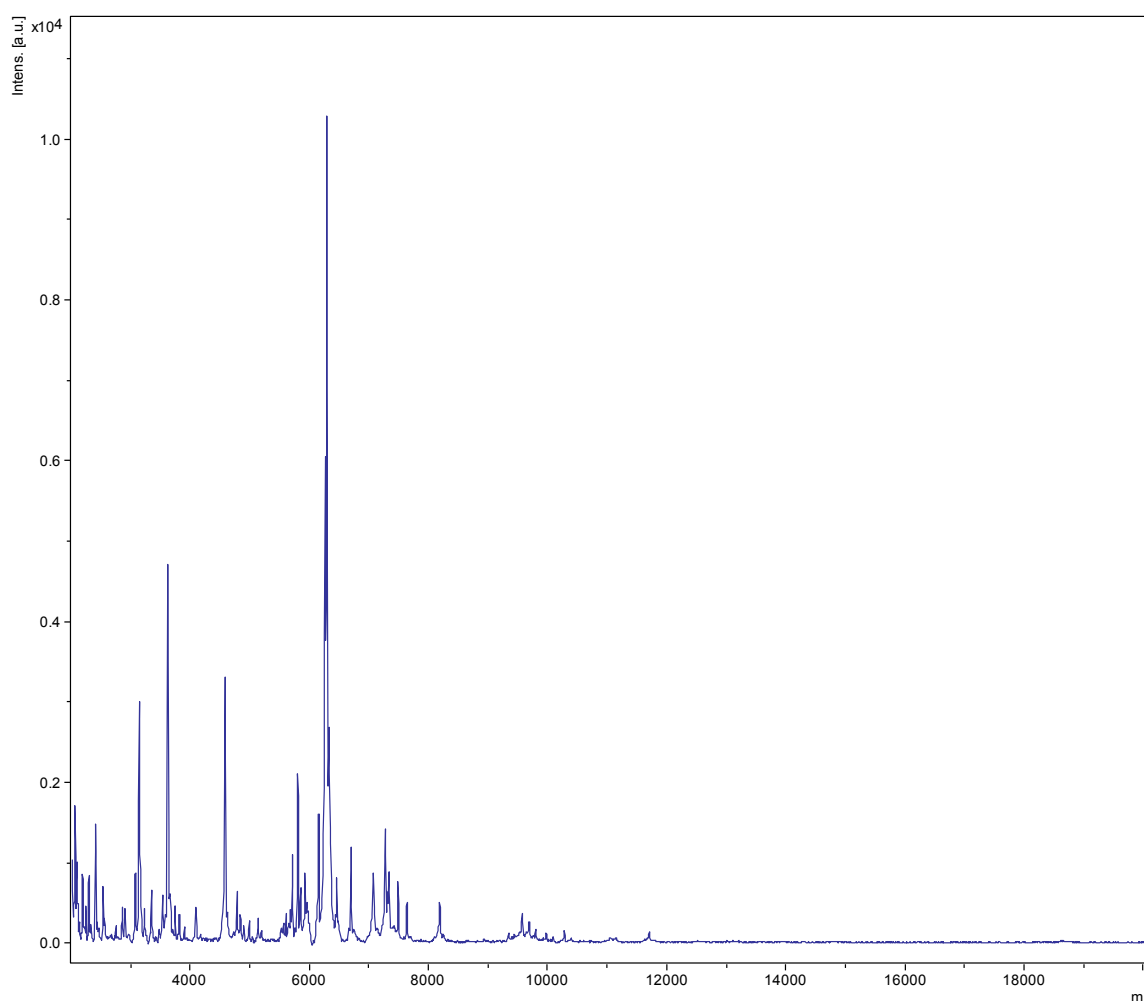
**Table 2.** Differential characteristics of *Bacteroides* species [57-67]<sup>†</sup>.

| Properties           | <i>B. timonensis</i> | <i>B. cellulosycticus</i> | <i>B. intestinalis</i> | <i>B. fragilis</i> | <i>B. vulgatus</i> | <i>B. thetaiotaomicron</i> | <i>B. salanitronis</i> | <i>B. helcogenes</i> | <i>B. finegoldii</i> | <i>B. uniformis</i> |
|----------------------|----------------------|---------------------------|------------------------|--------------------|--------------------|----------------------------|------------------------|----------------------|----------------------|---------------------|
| Cell diameter (µm)   | 0.88                 | 2-5                       | 1-3                    | 1.3                | 0.5-0.8            | 0.7-2                      | 2-3                    | 1-2                  | 1-2                  | 0.5-2               |
| Oxygen requirement   | anaerobic            | anaerobic                 | anaerobic              | anaerobic          | anaerobic          | anaerobic                  | anaerobic              | anaerobic            | anaerobic            | anaerobic           |
| Gram stain           | -                    | -                         | -                      | -                  | +                  | -                          | -                      | -                    | -                    | -                   |
| Salt requirement     | +                    | +                         | +                      | +                  | na                 | na                         | na                     | na                   | +                    | +                   |
| Motility             | -                    | -                         | -                      | -                  | -                  | -                          | -                      | -                    | -                    | -                   |
| Endospore formation  | -                    | -                         | -                      | +                  | na                 | -                          | -                      | -                    | -                    | -                   |
| Indole               | +                    | +                         | +                      | -                  | -                  | +                          | -                      | -                    | -                    | +                   |
| <b>Production of</b> |                      |                           |                        |                    |                    |                            |                        |                      |                      |                     |
| Alkaline phosphatase | +                    | +                         | +                      | Na                 | +                  | +                          | +                      | na                   | +                    | na                  |
| Catalase             | +                    | -                         | +                      | +                  | na                 | +                          | -                      | -                    | na                   | -                   |
| Oxidase              | -                    | +                         | na                     | +                  | na                 | na                         | na                     | na                   | na                   | na                  |
| Nitrate reductase    | +                    | na                        | -                      | na                 | -                  | +                          | -                      | -                    | -                    | -                   |
| Urease               | -                    | na                        | -                      | na                 | na                 | -                          | na                     | -                    | -                    | na                  |
| α-galactosidase      | +                    | -                         | +                      | na                 | -                  | +                          | +                      | +                    | +                    | +                   |
| N-acetyl-glucosamine | +                    | +                         | +                      | na                 | na                 | +                          | na                     | na                   | +                    | +                   |
| <b>Acid from</b>     |                      |                           |                        |                    |                    |                            |                        |                      |                      |                     |
| L-Arabinose          | -                    | w                         | +                      | -                  | -                  | +                          | +                      | -                    | -                    | +                   |
| Ribose               | -                    | +                         | na                     | -                  | +                  | na                         | na                     | -                    | na                   | na                  |
| Mannose              | -                    | +                         | +                      | +                  | +                  | +                          | +                      | +                    | +                    | +                   |
| Mannitol             | -                    | -                         | -                      | -                  | -                  | -                          | -                      | -                    | -                    | -                   |
| Sucrose              | -                    | +                         | +                      | +                  | +                  | +                          | +                      | +                    | +                    | +                   |
| D-glucose            | -                    | +                         | +                      | +                  | +                  | +                          | +                      | +                    | +                    | +                   |
| D-fructose           | -                    | +                         | +                      | +                  | +                  | +                          | +                      | +                    | na                   | na                  |
| D-maltose            | -                    | +                         | +                      | +                  | +                  | +                          | +                      | +                    | +                    | +                   |
| D-lactose            | -                    | w                         | +                      | +                  | +                  | +                          | +                      | +                    | +                    | +                   |
| Habitat              | human gut            | human gut                 | human gut              | human gut          | human gut          | human gut                  | human gut              | pig gut              | human gut            | human gut           |

<sup>†</sup>*Bacteroides timonensis* strain AP1<sup>T</sup>, *B. cellulosycticus* strain DSM 14838, *B. intestinalis* strain DSM 17393, *B. fragilis* strain YCH46, *B. vulgatus* strain ATCC 8482, *B. thetaiotaomicron* strain VPI-5482, *B. salanitronis* strain DSM 18170, *B. helcogenes* strain P 36-108, *B. finegoldii* strain DSM 17565 and *B. uniformis* strain ATCC 8492. na = data not available; w = weak, v = variable reaction

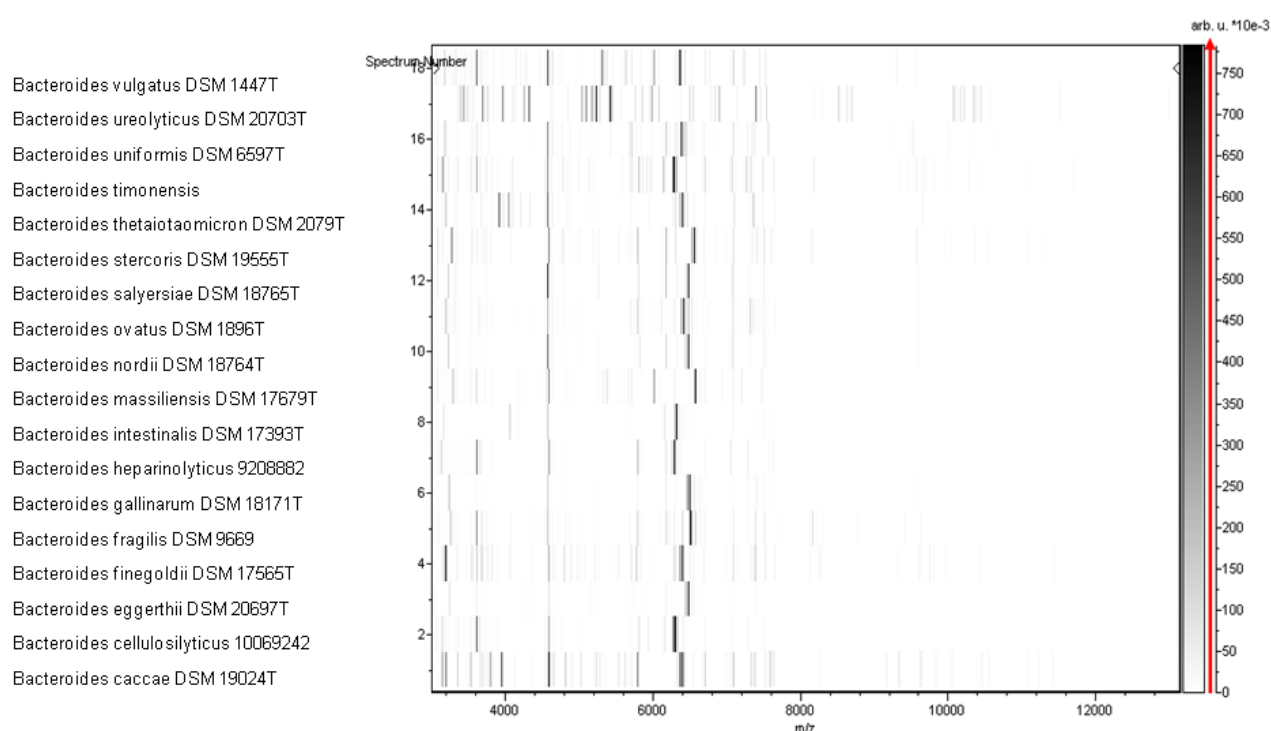
Matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) MS protein analysis was carried out as previously described [68]. Briefly, a pipette tip was used to pick one isolated bacterial colony from a culture agar plate and spread it as a thin film on a MTP 384 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany). Twelve distinct deposits from twelve isolated colonies were performed for strain AP1<sup>T</sup>. Each smear was overlaid with 2 µL of matrix solution (saturated solution of alpha-cyano-4-hydroxycinnamic acid) in 50% acetonitrile, 2.5% tri-fluoroacetic acid, and allowed to dry for 5 minutes. Measurements were performed with a Microflex spectrometer (Bruker). Spectra were recorded in the positive linear mode for the mass range of 2,000 to 20,000 Da (parameter settings: ion source 1 (ISI), 20kV; IS2, 18.5 kV; lens, 7 kV). A spectrum was obtained after 675 shots with variable laser power. The time of acquisition was between 30 seconds and 1 minute per spot. The twelve AP1<sup>T</sup> spectra were imported

into the MALDI BioTyper software (version 2.0, Bruker) and analyzed by standard pattern matching (with default parameter settings) against the main spectra of 3,769 bacteria, including 129 spectra from 98 *Bacteroides* species. The method of identification included the m/z from 3,000 to 15,000 Da. For every spectrum, a maximum of 100 peaks were compared with spectra in database. The resulting score enabled the identification of tested species, or not: a score  $\geq 2$  with a validly published species enabled identification at the species level, a score  $\geq 1.7$  but  $< 2$  enabled identification at the genus level, and a score  $< 1.7$  did not enable any identification. No significant MALDI-TOF score was obtained for strain AP1<sup>T</sup> against the Bruker database, suggesting that our isolate was not a member of a known species. We added the spectrum from strain AP1<sup>T</sup> to our database (Figure 4). Finally, the gel view showed the spectral differences with other members of the genus *Bacteroides* (Figure 5).



**Figure 4.** Reference mass spectrum from *B. timonensis* strain AP1<sup>T</sup>. Spectra from 12 individual colonies were compared and a reference spectrum was generated.





**Figure 5.** Gel view comparing *B. timonensis* strain AP1<sup>T</sup> to other *Bacteroides* species. The gel view displays the raw spectra of loaded spectrum files as a pseudo-electrophoretic gel. The x-axis records the m/z value. The left y-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a grey scale scheme code. The grey scale bar on the right y-axis indicate the relation between the shade of grey a peak is displayed with and the peak intensity in arbitrary units. Displayed species are detailed in the left column.

## Genome sequencing information

### Genome project history

The organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA gene sequence similarity to members of the genus *Bacteroides*, and is part of a study of the human digestive flora aiming at isolating all bacterial species within human feces [1]. It was

the ninety-ninth genome of a *Bacteroides* species and the first genome of *B. timonensis* sp. nov. The GenBank accession number is CBVI000000000 and consists of 211 contigs. Table 3 shows the project information and its association with MGS version 2.0 compliance [43].

**Table 3.** Project information

| MIGS ID   | Property            | Term                           |
|-----------|---------------------|--------------------------------|
| MIGS-31   | Finishing quality   | High-quality draft             |
| MIGS-28   | Libraries used      | 454 GS paired-end 3-kb library |
| MIGS-29   | Sequencing platform | 454 GS FLX Titanium            |
| MIGS-31.2 | Fold coverage       | 35.76                          |
| MIGS-30   | Assemblers          | gsAssembler                    |
| MIGS-32   | Gene calling method | PRODIGAL                       |

### Growth conditions and DNA isolation

*B. timonensis* sp. nov., strain AP1<sup>T</sup> (= CSUR P194 = DSM 26083) was grown on 5% sheep blood-enriched Columbia agar (BioMerieux) at 37°C in anaerobic atmosphere. Bacteria grown on four

Petri dishes were harvested and resuspended in 4x100μL of TE buffer. Then, 200μL of this suspension was diluted in 1ml TE buffer for lysis treatment that included a 30- minute incubation with 2.5 μg/μL lysozyme at 37°C, followed by an

overnight incubation with 20 µg/µL proteinase K at 37°C. Extracted DNA was then purified using 3 successive phenol-chloroform extractions and ethanol precipitation at -20°C overnight. After centrifugation, the DNA was resuspended in 160 µL TE buffer. The yield and concentration was measured by the Quant-it Picogreen kit (Invitrogen) on the Genios-Tecan fluorometer at 88.6 ng/µL.

### Genome sequencing and assembly

Five µg of DNA was mechanically fragmented on Covaris device (KBioScience-LGC Genomics, Teddington, UK) using miniTUBE-blue. The DNA fragmentation was visualized through an Agilent 2100 BioAnalyzer on a DNA labchip 7500 with an average size of 2.950kb. A 3 kb paired-end library was constructed according to the 454 GS FLX Titanium paired-end protocol (Roche). Circularization and nebulization were performed and generated a pattern with a mean size of 513 bp. After PCR amplification through 17 cycles followed by double size selection, the single stranded paired-end library was quantified with the Quant-it Ribogreen kit (Invitrogen) on the Genios Tecan fluorometer at 243 pg/µL. The library concentration equivalence was calculated as  $8.69 \times 10^8$  molecules/µL. The library was stored at -20°C until further use.

The paired-end library was clonally amplified with 0.5cpb and 1cbp in 8 SV-emPCR reactions with the GS Titanium SV emPCR Kit (Lib-L) v2 (Roche). The yields of the emPCR reactions were 4.65 and 7.29% respectively, within the recommended range of 5 to 20% from the Roche procedure. Approximately 790,000 beads were loaded on a 1/4 region of a GS Titanium PicoTiterPlate PTP Kit 70×75 and sequenced with the GS Titanium Sequencing Kit XLR70 (Roche). The run was performed overnight and then analyzed on the cluster through the gsRunBrowser and Newbler assembler (Roche). A total of 802,249 passed filter wells were obtained and generated 255Mb with a length average of 314 bp. These sequences were assembled using Newbler (Roche) with 90% identity and 40bp as overlap. The final assembly identified 63 scaffolds and 211 large contigs (>1,500bp) generating a genome size of 7.13 Mb which corresponds to a coverage of 35.76× genome equivalent.

### Genome annotation

Open Reading Frames (ORFs) were predicted using Prodigal [69] with default parameters. However, the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched

against the GenBank [70] and Clusters of Orthologous Groups (COG) databases using BLASTP. The tRNAs and rRNAs were predicted using the tRNAScan-SE [71] and RNAmmer [72] tools, respectively. Signal peptides and numbers of transmembrane helices were predicted using SignalP [73] and TMHMM [74], respectively. Mobile genetic elements were predicted using PHAST [75] and RAST [76]. ORFans were identified if their BLASTP *E*-value was lower than  $1e-03$  for alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, we used an *E*-value of  $1e-05$ . Such parameter thresholds have already been used in previous works to define ORFans. Artemis [77] and DNA Plotter [78] were used for data management and visualization of genomic features, respectively. The Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [79].

To estimate the mean level of nucleotide sequence similarity at the genome level between *B. timonensis* and 9 other members of the genus *Bacteroides* (Table 6), we used the Average Genomic Identity Of gene Sequences (AGIOS) in-house software [8]. Briefly, this software uses the Proteinortho software [80] for the pairwise detection of orthologous proteins between genomes, then retrieves the corresponding genes and determines the mean percentage of nucleotide sequence identity among orthologous ORFs using the Needleman-Wunsch global alignment algorithm. *B. timonensis* strain AP1<sup>T</sup> was compared to *B. intestinalis* strain DSM 17393 (GenBank accession number NZ\_ABJL000000000), *B. cellulosilyticus* strain DSM 14838 (NZ\_ACCH000000000), *B. fragilis* strain YCH46 (NC\_006347), *B. vulgatus* strain ATCC 8482 (NC\_009614), *B. thetaiotaomicron* strain VPI-5482 (NC\_004663), *B. salanitronis* strain DSM 18170 (NC\_015164), *B. helcogenes* strain P36-108 (NC\_014933), *B. finegoldii* strain DSM 17565 (NZ\_ABXI000000000) and *B. uniformis* strain ATCC 8492 (AAYH000000000).

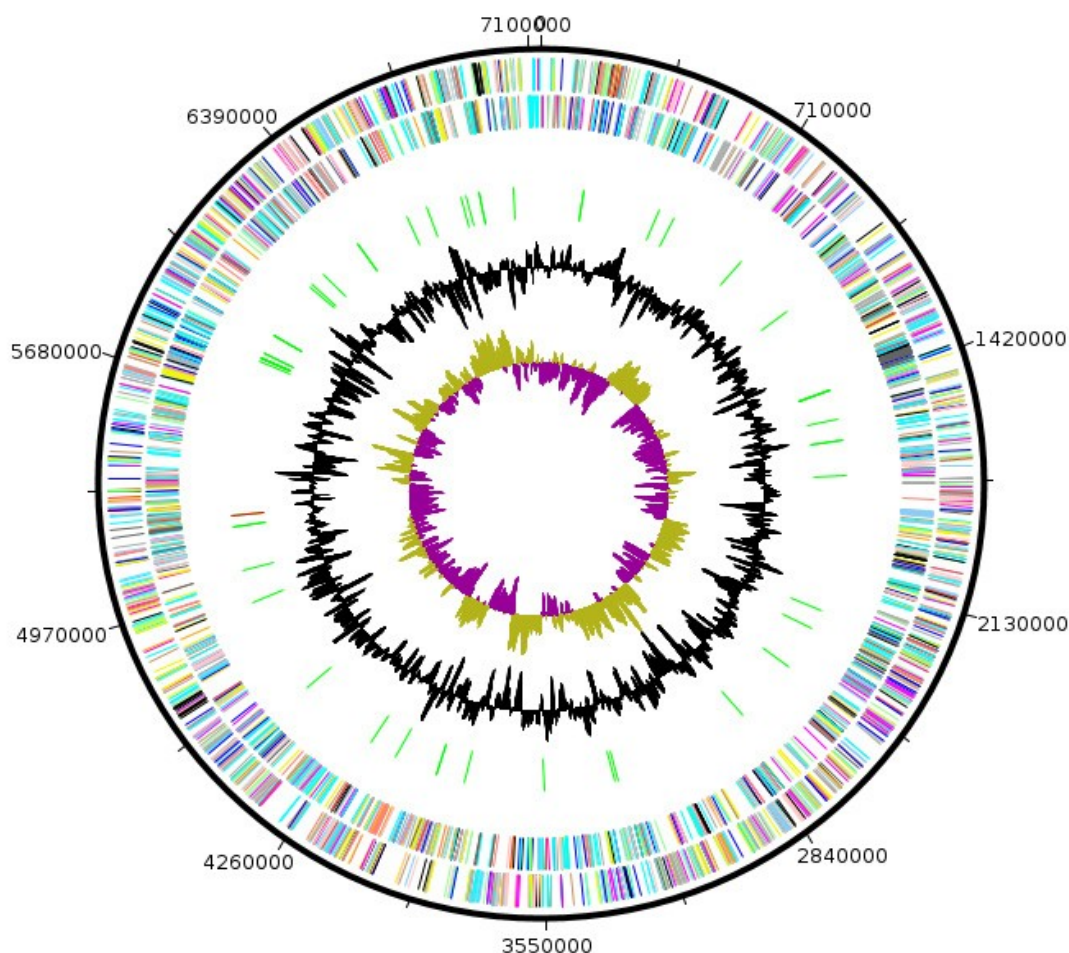
### Genome properties

The genome is 7,130,768 bp long (1 chromosome, but no plasmid) with a 43.3% G+C content (Figure 6 and Table 4). Of the 5,845 predicted genes, 5,786 were protein-coding genes and 59 were RNAs, including 1 complete rRNA operon. A total of 3,111 genes (53.22%) were assigned a putative function and 3,283 genes were identified as ORFans (56.16%). Strain AP1<sup>T</sup> possesses a variety of mobile genetic elements. These include 6 prophages of 13.70, 14.60, 10.51, 8.18, 9.91 and 12.79 Kb, respectively) and 91 trans-



posable elements belonging to 18 transposon families that include the putative mobilization protein BF0133, the putative conjugative transposon mobilization protein BF0132, the hypothetical protein clustered with conjugative transposons BF0131, TraA-CTn, TraB-CTn, TraD-CTn, TraE-CTn, TraF-CTn, TraG-CTn, TraH-

CTn, TraI-CTn, TraJ-CTn, TraK-CTn, TraL-CTn, TraM-CTn, TraN-CTn, TraO-CTn and TraQ-CTn. The properties and statistics of the genome are summarized in Tables 4 and 5. The distribution of genes into COGs functional categories is presented in Table 5.



**Figure 6.** Graphical circular map of the chromosome. From the outside in: open reading frames oriented in the forward (colored by COG categories) direction, open reading frames oriented in the reverse (colored by COG categories) direction, RNA operon (red), and tRNAs (green), GC content plot, and GC skew (purple: negative values, olive: positive values).

**Table 4.** Nucleotide content and gene count levels of the genome

| Attribute                        | Value     | % of total <sup>a</sup> |
|----------------------------------|-----------|-------------------------|
| Genome size (bp)                 | 7,130,768 |                         |
| DNA coding region (bp)           | 6,434,142 | 90.23                   |
| DNA G+C content (bp)             | 3,087,622 | 43.30                   |
| Number of replicons              | 1         |                         |
| Extra chromosomal element        | 0         |                         |
| Total genes                      | 5,845     | 100                     |
| RNA genes                        | 59        | 1.01                    |
| Protein-coding genes             | 5,786     | 98.99                   |
| Genes with function prediction   | 3,111     | 53.22                   |
| Genes assigned to COGs           | 2,820     | 48.24                   |
| Genes with peptide signals       | 435       | 7.44                    |
| Genes with transmembrane helices | 456       | 7.80                    |

<sup>a</sup> The total is based on either the size of the genome in base pairs or the total number of protein-coding genes in the annotated genome

**Table 5.** Number of genes associated with the 25 general COG functional categories

| Code | Value | %age <sup>a</sup> | Description                                                  |
|------|-------|-------------------|--------------------------------------------------------------|
| J    | 156   | 2.66              | Translation                                                  |
| A    | 0     | 0                 | RNA processing and modification                              |
| K    | 234   | 4.00              | Transcription                                                |
| L    | 200   | 3.42              | Replication, recombination and repair                        |
| B    | 0     | 0                 | Chromatin structure and dynamics                             |
| D    | 27    | 0.46              | Cell cycle control, mitosis and meiosis                      |
| Y    | 0     | 0                 | Nuclear structure                                            |
| V    | 107   | 1.83              | Defense mechanisms                                           |
| T    | 240   | 4.22              | Signal transduction mechanisms                               |
| M    | 361   | 6.17              | Cell wall/membrane biogenesis                                |
| N    | 5     | 0.08              | Cell motility                                                |
| Z    | 0     | 0                 | Cytoskeleton                                                 |
| W    | 0     | 0                 | Extracellular structures                                     |
| U    | 65    | 1.11              | Intracellular trafficking and secretion                      |
| O    | 89    | 1.52              | Posttranslational modification, protein turnover, chaperones |
| C    | 168   | 2.87              | Energy production and conversion                             |
| G    | 369   | 6.31              | Carbohydrate transport and metabolism                        |
| E    | 212   | 3.62              | Amino acid transport and metabolism                          |
| F    | 73    | 1.25              | Nucleotide transport and metabolism                          |
| H    | 130   | 2.22              | Coenzyme transport and metabolism                            |
| I    | 87    | 1.48              | Lipid transport and metabolism                               |
| P    | 202   | 3.42              | Inorganic ion transport and metabolism                       |
| Q    | 47    | 0.80              | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 518   | 8.86              | General function prediction only                             |
| S    | 197   | 3.37              | Function unknown                                             |
| -    | 2966  | 51.26             | Not in COGs                                                  |

<sup>a</sup> The total is based on the total number of protein-coding genes in the annotated genome.

## Genome comparison with other

### *Bacteroides* genomes

Here, we compare the genome of *B. timonensis* with those of *B. intestinalis*, DSM 17393, *B. cellulosilyticus* DSM 14838, *B. fragilis* YCH46, *B. vulgatus* ATCC 8482, *B. thetaiotaomicron* VPI-5482, *B. salanitronis* DSM 18170, *B. helcogenes* P 36-108, *B. finegoldii* DSM 17565 and *B. uniformis* ATCC 8492. The draft genome of *B. timonensis* (7.13Mb) is larger than all other studied genomes (Table 6A). It also exhibits a higher G+C content than all other genomes except *B. salanitronis*, *B. helcogenes* and *B. uniformis* (43.3, 46.4, 44.7 and 46.4%, respectively). *B. timonensis* has a higher gene content (5,786) than any other compared genome. The distribution of genes into COG categories was similar in all 10 compared genomes except in the N cate-

gory (cell motility) for which *B. fragilis*, *B. vulgatus*, *B. salanitronis*, *B. helcogenes* and *B. uniformis* were underrepresented (Figure 7). In addition, *B. timonensis* shared 2,956, 3,081, 2,159, 2,099, 2,379, 1,721, 2,001, 2,039 and 2,268 orthologous genes with *B. intestinalis*, *B. cellulosilyticus*, *B. fragilis*, *B. vulgatus*, *B. thetaiotaomicron*, *B. salanitronis*, *B. helcogenes*, *B. finegoldii* and *B. uniformis*, respectively. Among compared genomes except *B. timonensis*, AGIOS values ranged from 70.16 between *B. salanitronis* and *B. cellulosilyticus* to 88.16% between *B. intestinalis* and *B. cellulosilyticus*. When *B. timonensis* was compared to other species, AGIOS values ranged from 70.29 with *B. salanitronis* to 93.61% with *B. cellulosilyticus* (Table 6B).

**Table 6A.** Genomic comparison of *B. timonensis* with 9 other *Bacteroides* species<sup>†</sup>.

| Species                    | Strain    | Genome accession number | Genome size (Mb) | G+C content |
|----------------------------|-----------|-------------------------|------------------|-------------|
| <i>B. timonensis</i>       | AP1       | CBVI010000000           | 7.13             | 43.3        |
| <i>B. intestinalis</i>     | DSM 17393 | NZ_ABJL00000000         | 6.05             | 42.8        |
| <i>B. cellulosilyticus</i> | DSM 14838 | NZ_ACCH00000000         | 6.87             | 42.7        |
| <i>B. fragilis</i>         | YCH46     | NC_006347               | 5.28             | 43.2        |
| <i>B. vulgatus</i>         | ATCC 8482 | NC_009614               | 5.16             | 42.2        |
| <i>B. thetaiotaomicron</i> | VPI-5482  | NC_004663               | 6.26             | 42.8        |
| <i>B. salanitronis</i>     | DSM 18170 | NC_015164               | 4.24             | 46.4        |
| <i>B. helcogenes</i>       | P 36-108  | NC_014933               | 4.0              | 44.7        |
| <i>B. finegoldii</i>       | DSM 17565 | NZ_ABXI00000000         | 4.89             | 42.9        |
| <i>B. uniformis</i>        | ATCC 8492 | AAYH000000000           | 4.72             | 46.4        |

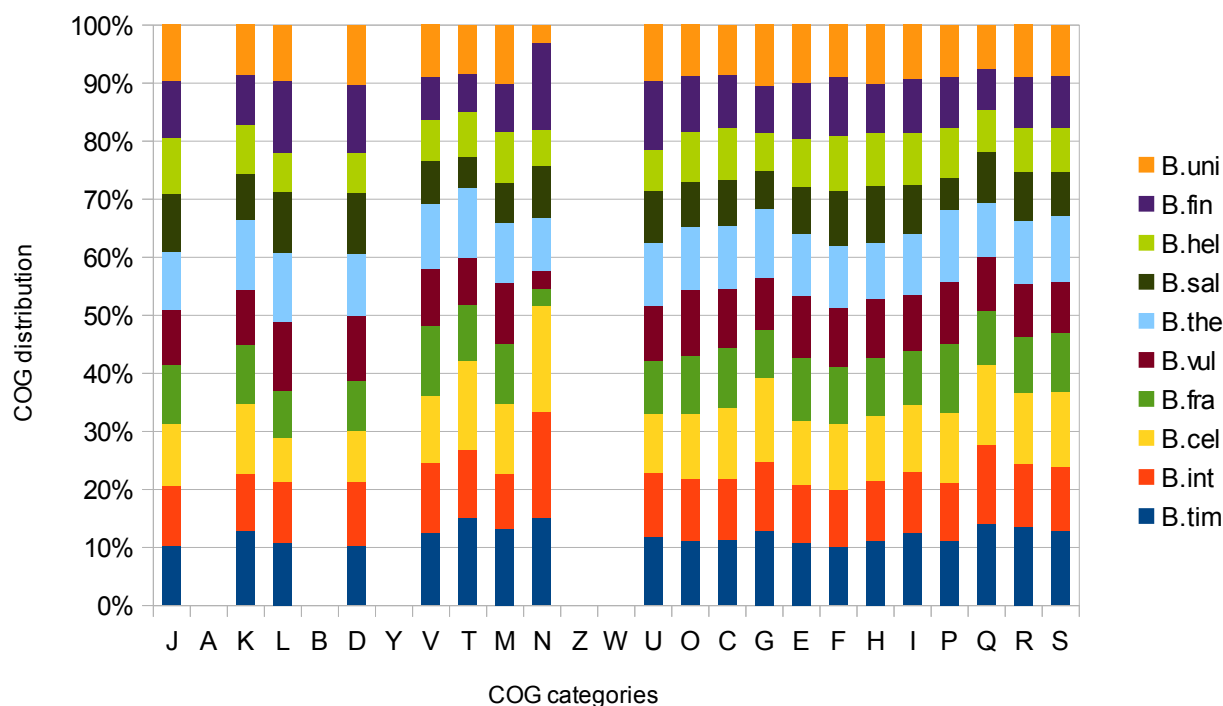
<sup>†</sup>Species, Strain, GenBank accession number, genome size and G+C content of all compared genomes.

**Table 6B.** Genomic comparison of *B. timonensis* with 9 other *Bacteroides* species<sup>†</sup>.

|                       | <i>B. tim</i> | <i>B. int</i> | <i>B. cel</i> | <i>B. fra</i> | <i>B. vul</i> | <i>B. the</i> | <i>B. sal</i> | <i>B. hel</i> | <i>B. fin</i> | <i>B. uni</i> |
|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| <b><i>B. tim</i></b>  | <b>5,786</b>  | 2,956         | 3,081         | 2,159         | 2,099         | 2,379         | 1,721         | 2,001         | 2,039         | 2,268         |
| <b><i>B. int</i></b>  | 87.73         | <b>4,911</b>  | 2,967         | 2,085         | 2,036         | 2,361         | 1,667         | 1,963         | 2,066         | 2,278         |
| <b><i>B. cell</i></b> | 93.61         | 88.16         | <b>5,719</b>  | 2,130         | 2,078         | 2,380         | 1,655         | 1,990         | 2,017         | 2,231         |
| <b><i>B. fra</i></b>  | 73.76         | 74.43         | 73.92         | <b>4,184</b>  | 1,927         | 2,174         | 1,517         | 1,893         | 1,880         | 1,995         |
| <b><i>B. vul</i></b>  | 71.91         | 71.74         | 71.48         | 71.87         | <b>4,066</b>  | 2,100         | 1,638         | 1,743         | 1,859         | 1,898         |
| <b><i>B. the</i></b>  | 73.99         | 74.65         | 73.87         | 75.42         | 72.21         | <b>4,778</b>  | 1,601         | 1,891         | 2,191         | 2,039         |
| <b><i>B. sal</i></b>  | 70.29         | 70.65         | 70.16         | 70.35         | 72.18         | 70.50         | <b>3,553</b>  | 1,466         | 1,580         | 1,584         |
| <b><i>B. hel</i></b>  | 76.40         | 76.51         | 76.41         | 74.15         | 71.62         | 73.64         | 70.68         | <b>3,244</b>  | 1,703         | 1,930         |
| <b><i>B. fin</i></b>  | 74.28         | 75.01         | 74.45         | 75.72         | 72.22         | 81.24         | 70.77         | 73.99         | <b>4,485</b>  | 1,920         |
| <b><i>B. uni</i></b>  | 77.08         | 76.83         | 76.80         | 74.25         | 72.45         | 74.36         | 71.32         | 79.37         | 74.77         | <b>4,663</b>  |

<sup>†</sup> numbers of orthologous proteins shared between genomes (above diagonal), AGIOS values (below diagonal) and numbers of proteins per genome (bold numbers).

*B. tim* = *B. timonensis*, *B. int* = *B. intestinalis*, *B. cel* = *B. cellulosilyticus*, *B. fra* = *B. fragilis*, *B. vul* = *B. vulgatus*, *B. the* = *B. thetaiotaomicron*, *B. sal* = *B. salanitronis*, *B. hel* = *B. helcogenes*, *B. uni* = *B. uniformis*, *B. fin* = *B. finegoldii*.



**Figure 7.** Distribution of predicted genes of *B. timonensis* and 9 other *Bacteroides* species into COG categories. B. uni = *B. uniformis*, B. fin = *B. finegoldii*, B. hel = *B. helcogenes*, B. sal = *B. salanitronis*, B. the = *B. thetaiotaomicron*, B. vul = *B. vulgatus*, B. fra = *B. fragilis*, B. cel = *B. cellulosilyticus*, B. int = *B. intestinalis*, B. tim = *B. timonensis*.

## Conclusion

On the basis of phenotypic, phylogenetic and genomic analyses (taxono-genomics), we formally propose the creation of *Bacteroides timonensis* sp. nov. that contains strain AP1<sup>T</sup>. This strain was isolated from the fecal flora of a 21-year-old woman who suffered from severe anorexia nervosa.

## Description of *B. timonensis* sp. nov.

*Bacteroides timonensis* (tim.o.nen'sis. L. masc. adj. timonensis, of Timone, the name of the hospital where strain AP1<sup>T</sup> was first cultivated).

Colonies are translucent and 0.3 mm in diameter on blood-enriched Columbia agar. Cells are rod-shaped with a mean diameter of 0.88  $\mu\text{m}$ . Optimal growth is achieved anaerobically, although the strain is able to grow under microaerophilic conditions, and weakly with 5% CO<sub>2</sub>. Growth occurs between 25°C and 37°C, with optimal growth at 37°C. Cells stain Gram-negative and are not motile. Positive reactions for catalase, arginine dihydrolase,  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\alpha$ -glucosidase,  $\beta$ -glucosidase,  $\alpha$ -arabinosidase, N-acetyl- $\beta$ -glucosaminidase, glutamic acid decarboxylase,  $\alpha$ -fucosidase, nitrate reduction, indole production, alkaline phosphatase, proline arylamidase, leucyl glycine

arylamidase, alanine arylamidase, glutamyl glutamic acid arylamidase, and fermentation of mannose and raffinose.

Weak activities are observed for glycine arylamidase and serine arylamidase. Negative reactions are obtained for urease,  $\beta$ -galactosidase-6-phosphatase,  $\beta$ -glucuronidase, arginine arylamidase, phenylalanine arylamidase, leucine arylamidase, pyroglutamic acid arylamidase, tyrosine arylamidase and histidine arylamidase. Using an API 50CH strip (Biomerieux), strain AP1<sup>T</sup> is asaccharolytic. Cells are susceptible to susceptible to amoxicillin-clavulanate, ceftriaxone, imipenem, trimethoprim-sulfamethoxazole, metronidazole and doxycycline but resistant to amoxicillin, vancomycin and gentamicin.

The 16S rRNA and genome sequences are deposited in GenBank under accession numbers JX041639 and CBVI000000000, respectively. The G+C content of the genome is 43.3%. The habitat of the organism is the digestive tract. The type strain AP1<sup>T</sup> (= CSUR P194 = DSMZ 26083) was isolated from the fecal flora of a French Caucasoid female who suffered from a severe restrictive form of anorexia nervosa. This strain has been found in Marseille, France.

## Acknowledgements

The authors thank the Xegen Company ([www.xegen.fr](http://www.xegen.fr)) for automating the genomic an-

notation process. This study was funded by the Mediterranean-Infection Foundation.

## References

1. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, *et al.* Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012; **18**:1185-1193. [PubMed](#)
2. Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 2010; **60**:249-266. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.016949-0>
3. Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006; **33**:152-155.
4. Wayne LG, Brenner DJ, Colwell PR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, *et al.* Report of the ad hoc committee on reconciliation of approaches to bacterial systematic. *Int J Syst Bacteriol* 1987; **37**:463-464. <http://dx.doi.org/10.1099/00207713-37-4-463>
5. Rossello-Mora R. DNA-DNA Reassociation Methods Applied to Microbial Taxonomy and Their Critical Evaluation. In: Stackebrandt E (ed), *Molecular Identification, Systematics, and population Structure of Prokaryotes*. Springer, Berlin, 2006; p. 23-50.
6. Welker M, Moore ER. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Syst Appl Microbiol* 2011; **34**:2-11. [PubMed](#) <http://dx.doi.org/10.1016/j.syapm.2010.11.013>
7. Kokcha S, Mishra AK, Lagier JC, Million M, Leroy Q, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Bacillus timonensis* sp. nov. *Stand Genomic Sci* 2012; **6**:346-355. [PubMed](#) <http://dx.doi.org/10.4056/sigs.2776064>
8. Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi-Tamisier M, Sentausa E, Raoult D, Fournier PE. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. *Int J Syst Evol Microbiol* 2013; **64**:384-391. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.057091-0>
9. Lagier JC, El Karkouri K, Nguyen TT, Armougom F, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. *Stand Genomic Sci* 2012; **6**:116-125. [PubMed](#) <http://dx.doi.org/10.4056/sigs.2415480>
10. Mishra AK, Gimenez G, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes senegalensis* sp. nov. *Stand Genomic Sci* 2012; **6**:304-314. <http://dx.doi.org/10.4056/sigs.2625821>
11. 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**:315-324. [PubMed](#) <http://dx.doi.org/10.4056/sigs.2685971>
12. 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**:386-395. [PubMed](#)
13. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. *Stand Genomic Sci* 2012; **7**:1-11. [PubMed](#) <http://dx.doi.org/10.4056/sigs.2956294>
14. Mishra AK, Lagier JC, Rivet R, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Paenibacillus senegalensis* sp. nov. *Stand Genomic Sci* 2012; **7**:70-81. [PubMed](#) <http://dx.doi.org/10.4056/sigs.3056450>
15. Lagier JC, Gimenez G, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. nov. *Stand Genomic Sci* 2012; **7**:200-209. [PubMed](#)
16. Kokcha S, Ramasamy D, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. *Stand Genomic Sci* 2012; **7**:233-245. [PubMed](#) <http://dx.doi.org/10.4056/sigs.3256677>
17. Ramasamy D, Kokcha S, Lagier JC, N'Guyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Aeromicrobium massiliense* sp. nov. *Stand Genomic Sci* 2012; **7**:246-257. [PubMed](#) <http://dx.doi.org/10.4056/sigs.3306717>
18. Lagier JC, Ramasamy D, Rivet R, Raoult D, Fournier PE. Non-contiguous finished genome



- sequence and description of *Cellulomonas massiliensis* sp. nov. *Stand Genomic Sci* 2012; 7:258-270. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3316719>
19. Lagier JC, Karkouri K, Rivet R, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov. *Stand Genomic Sci* 2013; 7:343-356. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3246665>
20. Mishra AK, Hugon P, Nguyen TT, Robert C, Couderc C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Peptoniphilus obesi* sp. nov. *Stand Genomic Sci* 2013; 7:357-369. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.32766871>
21. 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; 7:370-381. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3366764>
22. Lagier JC, Karkouri K, Mishra AK, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Enterobacter massiliensis* sp. nov. *Stand Genomic Sci* 2013; 7:399-412. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3396830>
23. Hugon P, Ramasamy D, Rivet R, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Alistipes obesi* sp. nov. *Stand Genomic Sci* 2013; 7:427-439. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3336746>
24. Hugon P, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Brevibacillus massiliensis* sp. nov. *Stand Genomic Sci* 2013; 8:1-14. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3466975>
25. Mishra AK, Hugon P, Nguyen TT, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Enorma massiliensis* gen. nov., sp. nov., a new member of the Family *Coriobacteriaceae*. *Stand Genomic Sci* 2013; 8:290-305. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3426906>
26. Ramasamy D, Lagier JC, Gorlas A, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus massiliosenegalensis* sp. Nov. *Stand Genomic Sci* 2013; 8:264-278. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3496989>
27. Ramasamy D, Lagier JC, Nguyen TT, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Dielma fastidiosa* gen. nov., sp. nov., a new member of the Family *Erysipelotrichaceae*. *Stand Genomic Sci* 2013; 8:336-351. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3567059>
28. Mishra AK, Pflieger A, Lagier JC, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus massiliorexius* sp. nov. *Stand Genomic Sci* 2013; 8:465-479. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4087826>
29. Hugon P, Ramasamy D, Robert C, Couderc C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Kallipyga massiliensis* gen. nov., sp. nov., a new member of the family *Clostridiales* *Incertae Sedis* XI. *Stand Genomic Sci* 2013; 8:500-515. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4047997>
30. Padmanabhan R, Lagier JC, Dangui NPM, Michelle C, Couderc C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Megasphaera massiliensis*. *Stand Genomic Sci* 2013; 8:525-538. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4077819>
31. Mishra AK, Edouard S, Dangui NPM, Lagier JC, Caputo A, Blanc-Tailleur C, Ravaut I, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Nosocomiicoccus massiliensis* sp. nov. *Stand Genomic Sci* 2013; 9:205-219. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4378121>
32. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Genome sequence and description of *Timonella senegalensis* gen. nov., sp. nov., a new member of the suborder *Micrococcineae*. *Stand Genomic Sci* 2013; 8:318-335. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3476977>
33. Keita MB, Diene SM, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus massiliogorillae* sp. nov. *Stand Genomic Sci* 2013; 9:93-105. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4388124>
34. Mediannikov O, El Karkouri K, Robert C, Fournier PE, Raoult D. Non contiguous-finished genome sequence and description of *Bartonella florentiae* sp. nov. *Stand Genomic Sci* 2013; 9:185-196. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4358060>
35. Lo CI, Mishra AK, Padmanabhan R, Samb Ba B, Gassama Sow A, Robert C, Couderc C, Faye N, Raoult D, Fournier PE, Fenollar F. Non contiguous-finished genome sequence and description of *Clostridium dakarensis* sp. nov. *Stand Genomic Sci* 2013; 9:14-27. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.4097825>
36. Mishra AK, Hugon P, Robert C, Raoult D, Fournier PE. Non contiguous-finished genome se-



- quence and description of *Peptoniphilus grossensis* sp. nov. *Stand Genomic Sci* 2012; **7**:320-330. [PubMed](#)
37. Mediannikov O, El Karkouri K, Diatta G, Robert C, Fournier PE, Raoult D. Non contiguous-finished genome sequence and description of *Bartonella senegalensis* sp. nov. *Stand Genomic Sci* 2013; **8**:279-289. [PubMed](#)  
<http://dx.doi.org/10.4056/sigs.3807472>
  38. Garrity GM, Holt JG. Taxonomic Outline of the *Archaea* and *Bacteria*. In: Garrity GM, Boone DR, Castenholz RW (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 1, Springer, New York, 2001, p. 155-166.
  39. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA. Diversity of the human intestinal microbial flora. *Science* 2005; **308**:1635-1638. [PubMed](#)  
<http://dx.doi.org/10.1126/science.1110591>
  40. List of Prokaryotic names with standing nomenclature (LPSN). <http://www.bacterio.cict.fr>.
  41. Smith CJ, Rocha ER, Paster BJ. 2005. The medically important *Bacteroides* spp. in health and disease. In *The Prokaryotes*, an evolving electronic resource for the microbiological community, Release 3.19 (18.3.2005) (<http://141-150-157-117:8080/prokPUB/index.htm>). Edited by M. Dworkin. New York: Springer.
  42. Finegold SM, George WL. 1989. *Anaerobic Infections in Humans*. San Diego: Academic Press.
  43. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, et al. The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 2008; **26**:541-547. [PubMed](#) <http://dx.doi.org/10.1038/nbt1360>
  44. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. [PubMed](#) <http://dx.doi.org/10.1073/pnas.87.12.4576>
  45. Validation List No. 143. *Int J Syst Evol Microbiol* 2012; **62**:1-4. <http://dx.doi.org/10.1099/ijs.0.039487-0>
  46. Krieg NR, Ludwig W, Euzéby J, Whitman WB. Phylum XIV. *Bacteroidetes* phyl. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 4, Springer, New York, 2011, p. 25.
  47. Krieg NR. Class I. *Bacteroidia* class. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 4, Springer, New York, 2011, p. 25.
  48. Krieg NR. Order I. *Bacteroidales* ord. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), *Bergey's Manual of Systematic Bacteriology*, Second Edition, Volume 4, Springer, New York, 2011, p. 25.
  49. Skerman VBD, McGowan V, Sneath PHA. Approved Lists of Bacterial Names. *Int J Syst Bacteriol* 1980; **30**:225-420. <http://dx.doi.org/10.1099/00207713-30-1-225>
  50. Pribram E. Klassifikation der Schizomyceten. *Klassifikation der Schizomyceten (Bakterien)*, Franz Deuticke, Leipzig, 1933, p. 1-143.
  51. Castellani A, Chalmers AJ. Genus *Bacteroides* Castellani and Chalmers, 1918. *Manual of Tropical Medicine*, Third Edition, Williams, Wood and Co., New York, 1919, p. 959-960.
  52. Holdeman LV, Moore WEC. Genus I. *Bacteroides* Castellani and Chalmers 1919, 959. In: Buchanan RE, Gibbons NE (eds), *Bergey's Manual of Determinative Bacteriology*, Eighth Edition, The Williams and Wilkins Co., Baltimore, 1974, p. 385-404.
  53. Cato EP, Kelley RW, Moore WEC, Holdeman LV. *Bacteroides zoogloeoformans*, Weinberg, Nativelle, and Prévot 1937) corrig. comb. nov.: emended description. *Int J Syst Bacteriol* 1982; **32**:271-274. <http://dx.doi.org/10.1099/00207713-32-3-271>
  54. Shah HN, Collins MD. Proposal to restrict the genus *Bacteroides* (Castellani and Chalmers) to *Bacteroides fragilis* and closely related species. *Int J Syst Bacteriol* 1989; **39**:85-87. <http://dx.doi.org/10.1099/00207713-39-1-85>
  55. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29. [PubMed](#) <http://dx.doi.org/10.1038/75556>
  56. 16S Yourself database. (<http://www.mediterranee-infection.com/article.php?larub=152&titre=16s-yourself>).
  57. Eggerth AH, Gagnon BH. The *Bacteroides* of Human Feces. *J Bacteriol* 1933; **25**:389-413. [PubMed](#)
  58. Shah HN. 1992. The genus *Bacteroides* and related taxa. In *The Prokaryotes*, 2nd edn, pp. 3593-3607. Edited by Balows A, Truper HG,

- Dworkin M, Harder M & Schleifer KH. New York: Springer.
59. Bakir MA, Kitahara M, Sakamoto M, Matsumoto M, Benno Y. *Bacteroides intestinalis* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 2006; **56**:151-154. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.63914-0>
60. Robert C, Chassard C, Lawson PA, Bernalier-Donadille A. *Bacteroides cellulosilyticus* sp. nov., a cellulolytic bacterium from the human gut microbial community. *Int J Syst Evol Microbiol* 2007; **57**:1516-1520. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.64998-0>
61. Johnson JL. Taxonomy of the *Bacteroides* I. Deoxyribonucleic acid homologies among *Bacteroides fragilis* and other saccharolytic *Bacteroides* species. *Int J Syst Evol Microbiol* 1978; **28**:245-256.
62. Cato EP, Johnson JL. Reinstatement of species rank for *Bacteroides fragilis*, *B. ovatus*, *B. distasonis*, *B. thetaiotaomicron*, and *B. vulgatus*: Designation of Neotype Strains for *Bacteroides fragilis* (Veillon and Zuber) Castellani and Chalmers and *Bacteroides thetaiotaomicron* (Distaso) Castellani and Chalmers. *Int J Syst Bacteriol* 1976; **26**:230-237. <http://dx.doi.org/10.1099/00207713-26-2-230>
63. Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV, Gordon JI. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. *Science* 2003; **299**:2074-2076. [PubMed](#) <http://dx.doi.org/10.1126/science.1080029>
64. Lan PT, Sakamoto M, Sakata S, Benno Y. *Bacteroides barnesiae* sp. nov., *Bacteroides salanitronis* sp. nov. and *Bacteroides gallinarum* sp. nov., isolated from chicken caecum. *Int J Syst Evol Microbiol* 2006; **56**:2853-2859. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.64517-0>
65. Benno Y, Watabe J, Mitsuoka T. *Bacteroides pyogenes* sp. nov., *Bacteroides suis* sp. nov., and *Bacteroides helcogenes* sp. nov., New Species from Abscesses and Feces of Pigs. *Syst Appl Microbio* 1983; **14**:396-407.
66. Pati A, Gronow S, Zeytun A, Lapidus A, Nolan M, Hammon N, Deshpande S, Cheng JF, Tapia R, Han C, et al. Complete genome sequence of *Bacteroides helcogenes* type strain (P 36-108). *Stand Genomic Sci* 2011; **4**:45-53. [PubMed](#) <http://dx.doi.org/10.4056/sigs.1513795>
67. Bakir MA, Kitahara M, Sakamoto M, Matsumoto M, Benno Y. *Bacteroides finegoldii* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 2006; **56**:931-935. [PubMed](#) <http://dx.doi.org/10.1099/ijs.0.64084-0>
68. Seng P, Drancourt M, Gouriet F, La SB, Fournier PE, Rolain JM, Raoult D. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis* 2009; **49**:543-551. [PubMed](#) <http://dx.doi.org/10.1086/600885>
69. Prodigal. <http://prodigal.ornl.gov/>
70. Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* 2012; **40**:D48-D53. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkr1202>
71. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 1997; **25**:955-964. [PubMed](#) <http://dx.doi.org/10.1093/nar/25.5.0955>
72. 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**:3100-3108. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkm160>
73. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**:783-795. [PubMed](#) <http://dx.doi.org/10.1016/j.jmb.2004.05.028>
74. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001; **305**:567-580. [PubMed](#) <http://dx.doi.org/10.1006/jmbi.2000.4315>
75. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. *Nucleic Acids Res* 2011; **39**:W347-W352. [PubMed](#) <http://dx.doi.org/10.1093/nar/gkr485>
76. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 2008; **9**:75. [PubMed](#) <http://dx.doi.org/10.1186/1471-2164-9-75>
77. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. Artemis: sequence visualization and annotation. *Bioinformatics* 2000; **16**:944-945. [PubMed](#) <http://dx.doi.org/10.1093/bioinformatics/16.10.944>
78. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 2009; **25**:119-120. [PubMed](#) <http://dx.doi.org/10.1093/bioinformatics/btn578>

- 
79. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004; **14**:1394-1403. [PubMed](#) <http://dx.doi.org/10.1101/gr.2289704>
80. Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* 2011; **12**:124. [PubMed](#) <http://dx.doi.org/10.1186/1471-2105-12-124>