ORIGINAL ARTICLE

Noncontiguous finished genome sequences and description of *Bacteroides mediterraneensis* sp. nov., *Bacteroides ihuae* sp. nov., *Bacteroides togonis* sp. nov., *Bacteroides ndongoniae* sp. nov., *Bacteroides ilei* sp. nov. and *Bacteroides congonensis* sp. nov. identified by culturomics

C. Andrieu[1], M. Mailhe[1], D. Ricaboni[1,4], M. D. M. Fonkou[1], M. Bilen[1], F. Cadoret[1], E. Tomei[2], N. Armstrong[1], V. Vitton[3], A. Benezech[3], B. Davoust[1], A. Levasseur[1], J.-C. Lagier[1], P.-E. Fournier[2] and D. Raoult[1,5]

1) *Aix-Marseille Université, IRD, APHM, MEPHI, IHU-Méditerranée Infection,* 2) *Aix-Marseille Université, IRD, AP-HM, SSA, VITROME, IHU-Méditerranée Infection,* 3) *Service de Gastroentérologie, Hopital Nord, Assistance Publique-Hopitaux de Marseille, Marseille, France,* 4) *Département des Sciences Cliniques et Biomédicales, Luigi Sacco, Division des Maladies Infectieuses III, Université de Milan, Milan, Italy and* 5) *Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia*

Abstract

Culturomics is a concept developing different culture conditions in order to enlarge our knowledge of the human microbiota through the discovery of previously uncultured bacteria. This enabled us to isolate six new species of the *Bacteroides* genus: *Bacteroides mediterraneensis* strain Marseille-P2644, *Bacteroides ihuae* strain Marseille-P2824, *Bacteroides togonis* strain Marseille-P3166, *Bacteroides ndongoniae* strain Marseille-P3108, *Bacteroides ilei* strain Marseille-P3208 and *Bacteroides congonensis* strain Marseille-P3132. Those bacteria are Gram-negative anaerobic bacilli. We describe here their phenotypic features, together with phylogenetic analysis, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry spectrum, fatty acid composition, and genome sequencing and annotation.

Keywords: *Bacteroides* species, culturomics, emerging bacteria, human microbiota, taxonogenomics
Original Submission: 20 April 2018; **Revised Submission:** 8 June 2018; **Accepted:** 13 June 2018
Article published online: 21 June 2018

Corresponding author: D. Raoult, Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), CNRS 7278, IRD 198, INSERM 1095, UM63, Institut Hospitalo-Universitaire Méditerranée-Infection, Faculté de médecine, 27 Boulevard Jean Moulin, 13385, Marseille Cedex 5, France.
E-mail: didier.raoult@gmail.com

The first two authors contributed equally to this article and both should be considered first author.

Introduction

The *Bacteroides* genus was discovered in 1919 and defined as the type genus of the *Bacteroidaceae* family [1]. The first type species, *Bacteroides fragilis*, was discovered in 1898 and approved in 1980 [2]. This genus currently comprises 52 validated species (http://www.bacterio.net/) (Fig. 1) including *Bacteroides fragilis* and *Bacteroides thetaiotaomicron,* the oldest and most studied species. Bacteria of this genus are non-sporulating, anaerobic, Gram negative and rod shaped [3]. *Bacteroides* is one of the major lineages present in the human colon and is involved in 'good' processes (production of energy source, activation of immune response) as well as 'bad' processes leading to diseases (abscess, reservoir of antibiotic resistance determinants, bacteraemia) [3]. More recent studies have attempted to elucidate the implication of those bacteria in colorectal cancer [4,5].

Consequently, enlarging our knowledge of bacteria that colonize the human gut, but more generally of the human microbiota, is foundational to better understand the multiple functions in which they are involved and enable the treatment of diseases. This is the reason why we used the culturomics concept, consisting of the use of multiple growth conditions, in

https://doi.org/10.1016/j.nmni.2018.06.006

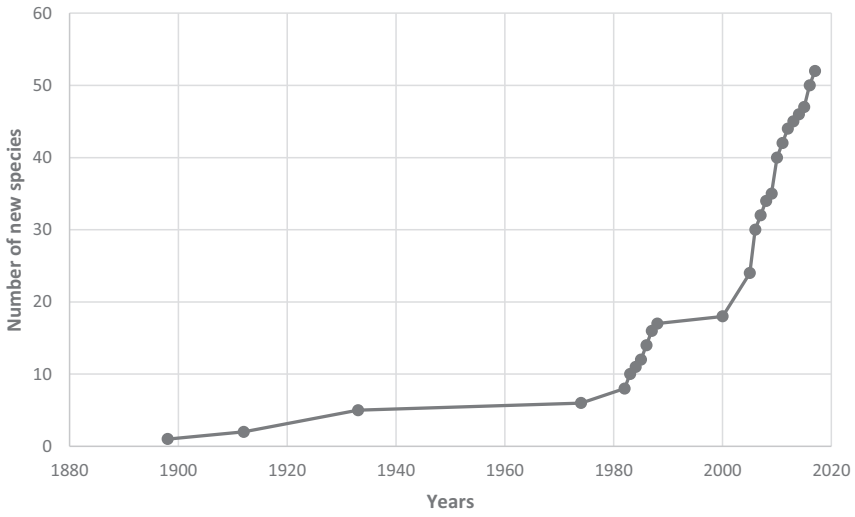

FIG. 1. Identification of *Bacteroides* new species since first one, *Bacteroides fragilis*, was described by Veillon and Zuber in 1898 [2].

order to identify new bacteria previously uncultivable with classic methods [6–9]. This enabled us to isolate six new species of the *Bacteroides* genus. We used a taxonogenomics approach, including matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) spectrum, phylogenetic analysis, main phenotypic description and genome sequencing [10,11], to describe the following bacteria: *Bacteroides mediterraneensis* strain Marseille-P2644 (= CSUR P2644 = DSM 103033), *Bacteroides ihuae* strain Marseille-P2824 (= CSUR P2824 = CCUG 70550), *Bacteroides togonis* strain Marseille-P3166 (= CSUR P3166 = DSM 103637), *Bacteroides ndongoniae* strain Marseille-P3108 (= CSUR P3108 = DSM 103636), *Bacteroides ilei* strain Marseille-P3208 (= CSUR P3208 = CCUG 69964) and *Bacteroides congonensis* strain Marseille-P3132 (= CSUR P3132 = CCUG 70144).

Materials and methods

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

According to the culturomics approach, we tested 18 conditions on samples to isolate these strains, as previously described [7,8]. The samples' origins and conditions of isolation are summarized in Table 1.

Purified colonies were identified by MALDI-TOF MS using a Microflex LT spectrometer and an MSP 96 MALDI-TOF target plate (Bruker Daltonics, Bremen, Germany), as previously described [12,13]. The obtained spectra were imported into MALDI Biotyper 3.0 software (Bruker Daltonics) and analysed by standard pattern matching (with default parameter settings) against the main spectra of the 7537 bacteria included in the databases (Bruker and constantly updated Microbes Evolution Phylogeny and Infections (MEPHI) databases). The resulting score enabled the identification (or not) of tested species: a score of $\geq$2 with a validly published species enabled identification at the species level, a score of $\geq$1.7 but <2 enabled identification at the genus level and a score of <1.7 did not enable any identification. No significant scores were obtained for the studied strains, suggesting that our isolates were not members of known species.

Consequently, sequencing of the 16S rRNA gene was performed in order to identify these strains. DNA was previously extracted by EZ1 DNA Tissue Kit using BioRobot EZ1 Advanced XL (Qiagen, Courtaboeuf, France). The amplification and purification of the 16S rRNA gene was performed as previously described [14,15] by use of the universal primers pair fD1 and rP2 (Eurogentec, Angers, France). Sequencing was then done using the Big Dye Terminator v1.1 Cycle Sequencing Kit and ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems; Thermo Fisher Scientific, Waltham, MA, USA), as previously described [14,15]. The 16S rRNA nucleotide sequences were assembled and corrected using CodonCode Aligner software (http://www.codoncode.com), and BLASTn searches were performed against the National Center for Biotechnology Information (NCBI) GenBank database (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi) to determine the percentage of similarity with the closest bacteria. A similarity threshold of <98.65% enabled us to define a new species, whereas a threshold of <95% enabled us to define a new genus without performing DNA-DNA hybridization [16].

Phylogenetic analysis

A custom Python script was used to automatically retrieve all species from the same family as that of the new species, and we downloaded 16S sequences from NCBI by parsing NCBI eUtils

TABLE 1. Sample information of six *Bacteroides* species

Characteristic	*B. ilei*	*B. mediterraneensis*	*B. ndongoniae*	*B. togonis*	*B. congonensis*	*B. ihuae*
Strain	Marseille-P3208	Marseille-P2644	Marseille-P3108	Marseille-P3166	Marseille-P3132	Marseille-P2824
Sample origin	Human ileum	Human ileum	Human right colon	Human right colon	Human stool	Human sputum
Patient information	76-year-old woman with oesophagitis (Marseille, France), no antibiotics	58-year-old woman realizing colorectal cancer screening (Marseille, France), no antibiotics	76-year-old woman with oesophagitis (Marseille, France), no antibiotics	76-year-old woman with oesophagitis (Marseille, France), no antibiotics	35-year-old healthy Pygmy woman (Congo), no treatment	Healthy 27-year-old woman (Marseille, France), no treatment
Authorization/ consent	Committee of 'IHU Méditerranée Infection,' No. 2016-010, signed consent	Committee of 'IHU Méditerranée Infection,' No. 2016-010, signed consent	Committee of 'IHU Méditerranée Infection,' No. 2016-010, signed consent	Committee of 'IHU Méditerranée Infection,' No. 2016-010, signed consent	Committee of 'IHU Méditerranée Infection,' No. 2016-011, signed consent	Committee of 'IHU Méditerranée Infection,' No. 2016-011, signed consent
Storage	No storage, fresh sample	No storage, fresh sample	No storage, fresh sample	No storage, fresh sample	+4°C	No storage, fresh sample
Isolation conditions	7 days on COS, 37°C, anaerobic	1 day on COS, 37°C, anaerobic	7 days on COS, 37°C, anaerobic	3 days on COS, 37°C, anaerobic	Blood culture (+5% sheep's blood + 5% rumen) + 10 days on COS, 37°C, anaerobic	Blood culture (+5% rumen) + 10 days on COS, 30°C, anaerobic

COS, Columbia medium supplemented with 5% sheep's blood.

results and the NCBI taxonomy page. It only kept sequences from type strains. The script then divided 16S sequences into two groups: one containing the sequences of strains from the same genus (group A) and one containing the others (group B). Finally, it only kept the 49 closest strains from group A and one from group B (outgroup).

Phenotypic, biochemical and antibiotic susceptibility tests

Ideal growth conditions of the studied strains were determined by testing five growth temperatures (25, 30, 37, 42, 56°C) in an aerobic atmosphere with or without 5% CO_2 and under anaerobic and microaerophilic conditions using the GENbag anaer and GENbag microaer systems, respectively (bioMérieux, Marcy l'Étoile, France). Different pH values (5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5) and NaCl concentrations (5, 10, 50, 75, 100, 150, 200 g/L) were also tested. Phenotypic characteristics such as Gram staining, motility, sporulation, and catalase and oxidase activities were tested as previously described [7].

The biochemical analysis was carried out using API 50CH, API 20A, API ZYM strips (bioMérieux) in an anaerobic atmosphere, according to the manufacturer's instructions. Cellular fatty acid methyl ester (FAME) analysis was performed by gas chromatography/mass spectrometry (GC/MS). Two samples were prepared with approximately 20 to 70 mg (according to bacteria) of bacterial biomass per tube collected from several culture plates. FAME were prepared as described by Sasser [17]. GC/MS analyses were carried out as previously described [18]. Briefly, FAME were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500–SQ 8 S; Perkin Elmer, Courtaboeuf, France). A spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology, Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK). Antibiotic susceptibility was tested using the disc diffusion method [19] and according to European Committee on Antimicrobial Susceptibility Testing 2018 recommendations.

Microscopy

Negative staining was performed in order to observe the cells' morphology. The cells were fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer for at least 1 hour at 4°C. A drop of cell suspension was deposited for approximately 5 minutes on glow-discharged Formvar carbon film with 400 mesh nickel grids (FCF400-Ni; Electron Microscopy Sciences (EMS), Hatfield, PA, USA). The grids were dried on blotting paper, and the cells were negatively stained for 10 seconds with 1% ammonium molybdate solution in filtered water at room temperature. Electron micrographs were acquired with a Tecnai G20 Cryo (FEI Company, Limeil-Brévannes, France) transmission electron microscope operated at 200 keV or with a Morgagni 268D (Philips, Amsterdam, The Netherlands) transmission electron microscope operated at 80 keV.

DNA extraction and genome sequencing

Genomic DNA (gDNA) of strains Marseille-P3132, Marseille-P3166 and Marseille-P3208 were first extracted by a mechanical treatment using acid-washed glass beads (G4649–500g; MilliporeSigma, St. Louis, MO, USA) and a FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5) for 3 × 30 seconds. Then for all the studied *Bacteroides* strains a 2-hour lysozyme incubation at 37°C was done and gDNA was extracted using the EZ1 biorobot (Qiagen) with the EZ1 DNA tissues kit. The elution volume was 50 µL. gDNA was quantified by a Qubit assay with the High Sensitivity kit (Life Technologies, Carlsbad, CA, USA) (Supplementary Table S1).

gDNA was sequenced on the MiSeq Technology (Illumina, San Diego, CA, USA) with the mate pair strategy. The gDNA

was barcoded in order to be mixed with 11 other projects with the Nextera Mate Pair sample prep kit (Illumina). The mate pair library was prepared with 1.5 µg of gDNA using the Nextera mate pair Illumina guide. The gDNA sample was simultaneously fragmented and tagged with a mate pair junction adapter. The pattern of fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments ranged in size from 1.5 to 11 kb (with an optimal size at 7.710, 6.047, 7.937, 7.840, 10.380 and 6.752 kb for strain Marseille-P2644, Marseille-P2824, Marseille-P3132, Marseille-P3108, Marseille-P3166 and Marseille-P3208, respectively). No size selection was performed, and 600 ng (610.4 ng for strain Marseille-P3132 and 369.6 ng for strain Marseille-P3108) of tagmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments (with an optimal size of 1209, 997, 1140 and 1167 bp for strain Marseille-P3132, Marseille-P3108, Marseille-P3166 and Marseille-P3208, with optima on a bimodal curve at 975 and 1514 bp for strain Marseille-P2824 and with optima on a trimodal curve at 675, 1252 and 2049 bp for strain Marseille-P2644) on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final concentration library was measured (Supplementary Table S1).

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 in a 2 × 151 bp (2 × 251 bp for strains Marseille-P3108, Marseille-P3166 and Marseille-P3208). The paired reads were finally trimmed and assembled. Complementary information is available in Supplementary Table S1.

Genome assembly, annotation and comparison

The genomes' assembly was performed with a pipeline that enabled us to create an assembly with different software packages (Velvet [20], Spades [21] and Soap Denovo [22]), on trimmed (MiSeq and Trimmomatic [23] software) or untrimmed data (only MiSeq software). For each of the six assemblies performed, GapCloser [22] was used to reduce gaps. Then contamination with Phage Phix was identified (BLASTn against Phage Phix174 DNA sequence) and eliminated. Finally, scaffolds under 800 bp in size were removed and scaffolds with a depth value lower than 25% of the mean depth were removed (identified as possible contaminants). The best assembly was selected by using different criteria (number of scaffolds, N50, number of N). For all studied strains, Spades gave the best assembly (with a depth coverage of 153, 44, 647, 99 and 94, respectively, for strains Marseille-P2824, Marseille-P3132, Marseille-P3108, Marseille-P3166 and Marseille-P3208), except for strain Marseille-P2644, which obtained the best assembly with Velvet (with a depth coverage of 174).

Open reading frames (ORFs) were predicted using Prodigal [24] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region (contained N). The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COGs) database using BLASTP (*E* value of 1e-03, coverage 0.7 and identity percentage 30%). If no hit was found, we searched against the NR database using BLASTP (*E* value of 1e-03, coverage 0.7 and identity percent of 30%). If the sequence length was smaller than 80 aa, we used an *E* value of 1e-05. The tRNAScanSE [25] tool was used to find transfer RNA genes, whereas ribosomal RNA genes were found by using RNAmmer [26]. Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [27]. ORFans were identified if all the BLASTP searches we performed failed to provide positive results (*E* value smaller than 1e-03 for ORFs with sequence size superior to 80 aa or *E* value smaller than 1e-05 for ORFs with sequence length smaller than 80 aa). Such parameter thresholds have already been used in previous works to define ORFans. Pfam-conserved domains (Pfam-A and Pfam-B domains) were searched on each protein with an HMMscan of the HMMER3 suite [28]. Number of genes associated to polyketide synthase (PKS) or nonribosomal peptide synthase (NRPS) and antibiotic resistance genes were found using the BLAST program against the ClusterMine360 specific database [29] and the Homemade database [30], respectively. The annotation process was performed in the multi-agent software system DAGOBAH [31], which includes Figenix [32] libraries, which provided the pipeline analysis.

Species that must be compared were automatically retrieved from the 16S RNA tree using PhyloPattern. For each selected species, the complete genome sequence, proteome sequence and ORFeome sequence were retrieved from the NCBI's FTP site. If one specific strain did not have a complete and available genome, a complete genome of the same species was used. If ORFeome and proteome were not predicted, Prodigal was used with default parameters to predict them. All proteomes were analysed with proteinOrtho [33]. Then, for each couple of genomes, a similarity score was computed. This score is the mean value of nucleotide similarity between all couples of orthologous genes between the two genomes studied (average genomic identity of orthologous gene sequences (AGIOS) tool). An annotation of all proteomes was also realized to define the distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins (with same method as that used for genome annotation). The comparison

process was performed in DAGOBAH [31], which includes Figenix [32] libraries, which provided pipeline analysis, and by using PhyloPattern [34] for tree manipulation.

Results

Strain identification and phylogenetic analysis

The studied strains could not be identified using MALDI-TOF MS, and therefore their 16S rRNA gene was sequenced. Strain Marseille-P3132 (accession no. LT598566) revealed 96.74% sequence similarity with the 16S rRNA of *Bacteroides thetaiotaomicron* strain VPI-5482, the closest species with a validly published name. We consequently propose that our strain is a representative strain of a new species within the *Bacteroides* genus for which we suggest the name *'Bacteroides congonensis'* strain Marseille-P3132. Strain Marseille-P2824 (accession no. LT576386) revealed 95.62% sequence similarity with the 16S rRNA of Bacteroides graminisolvens strain JCM15093, the closest species with a validly published name. We consequently propose that our strain is a representative strain of a new species within the genus *Bacteroides* for which we suggest the name *'Bacteroides ihuae'* strain Marseille-P2824. Strain Marseille-P3208 (accession no. LT623890) revealed 94.29% sequence similarity with the 16S rRNA of *Bacteroides coprophilus* strain CB42, the closest species with a validly published name. Although this value is under the threshold of 95%, we proposed this strain as a new species and not a new genus according to the low percentage of similarity already existing between two validly published *Bacteroides* species (74.8–98.7%) [35]. We consequently propose that our strain is a representative strain of a new species within the genus *Bacteroides* for which we suggest the name *'Bacteroides ilei'* strain Marseille-P3208. Strain Marseille-P3108 (accession no. LT615364) revealed 93.91% sequence similarity with the 16S rRNA of *Bacteroides eggerthii* strain JCM12986, the closest species with a validly published name. We consequently propose that our strain is a representative strain of a new species within the genus *Bacteroides* for which we suggest the name *'Bacteroides ndongoniae'* strain Marseille-P3108. Strain Marseille-P3166 (accession no. LT631521) revealed 93.14% sequence similarity with the 16S rRNA of *Bacteroides eggerthii* strain JCM12986, the closest species with a validly published name. We consequently propose that our strain is a representative strain of a new species within the genus *Bacteroides* for which we suggest the name *'Bacteroides togonis'* strain Marseille-P3166. Strain Marseille-P2644 (accession no. LT558804) revealed a 94.61% sequence similarity with the 16S rRNA of *Bacteroides coprocola* strain M16, the closest species with a validly published name. We consequently propose that our strain is a representative strain of a new species within the genus *Bacteroides* for which we suggest the name *'Bacteroides mediterraneensis'* strain Marseille-P2644.

A phylogenetic tree showing the position of the studied strains in the *Bacteroides* genus is shown in Fig. 2. We observed that *B. ilei* and *B. mediterraneensis,* which are closely related, exhibited a 95.31% similarity, which confirms that they belong to two distinct species. The same observation was made for *B. ndongoniae* and *B. togonis,* which exhibited a 93.05% 16S rRNA similarity. All spectra (Fig. 3(A)) were added to MEPHI database (http://www.mediterranee-infection.com/article.php?larub=280&titre=urms-database). A gel view was made in order to compare proteomic profiles between the studied strains and their closest species (Fig. 3(B)). The analysis revealed a common general profile between *Bacteroides* strains, including the studied strains, whereas the outsider species *Parabacteroides goldsteinii* profile shows several unique differences.

Phenotypic features

The main phenotypic results of each studied strain compared to published data of close species are presented in Table 2 [36–38]. The results show that the morphologic description, notably Gram staining, cell shape, spore formation and atmosphere, supports the fact that these six new species are members of the *Bacteroides* genus. These observations were confirmed by electronic microscopy (Fig. 4), which revealed a similar morphology. Regarding the biochemical characteristics, nearly all compared strains were negative for oxidase and urease activities and mannitol metabolism, whereas they were positive for β-galactosidase and alkaline phosphatase activities (Table 2). A more detailed analysis of biochemical features of the studied strains is available in Supplementary Tables S2, S3 and S4, corresponding to API strip results. Finally, we observed that the main fatty acid is C15:0 anteiso for nearly all compared strains (Table 2).

The composition in cellular fatty acids of the studied strains is presented in detail in Table 3; results of antibiotic susceptibility testing are presented in Table 4. We observed that the major cellular fatty acid of the studied *Bacteroides* strains is 12-methyl-tetradecanoic acid except for the Marseille-P3166 and Marseille-P3108 strains (Table 3). Table 4 shows that the strains often have the same antibiotic susceptibility profile, except strain Marseille-P3132. Nevertheless, all strains are susceptible to imipenem, amoxicillin, amoxicillin/clavulanic acid, metronidazole and rifampicin and resistant to gentamicin.

Genome description and comparison

The genomes of the studied strains are represented in a map in Supplementary Fig. 1. The detail of the genomes' content and statistics are summarized in Table 5, and the distribution of

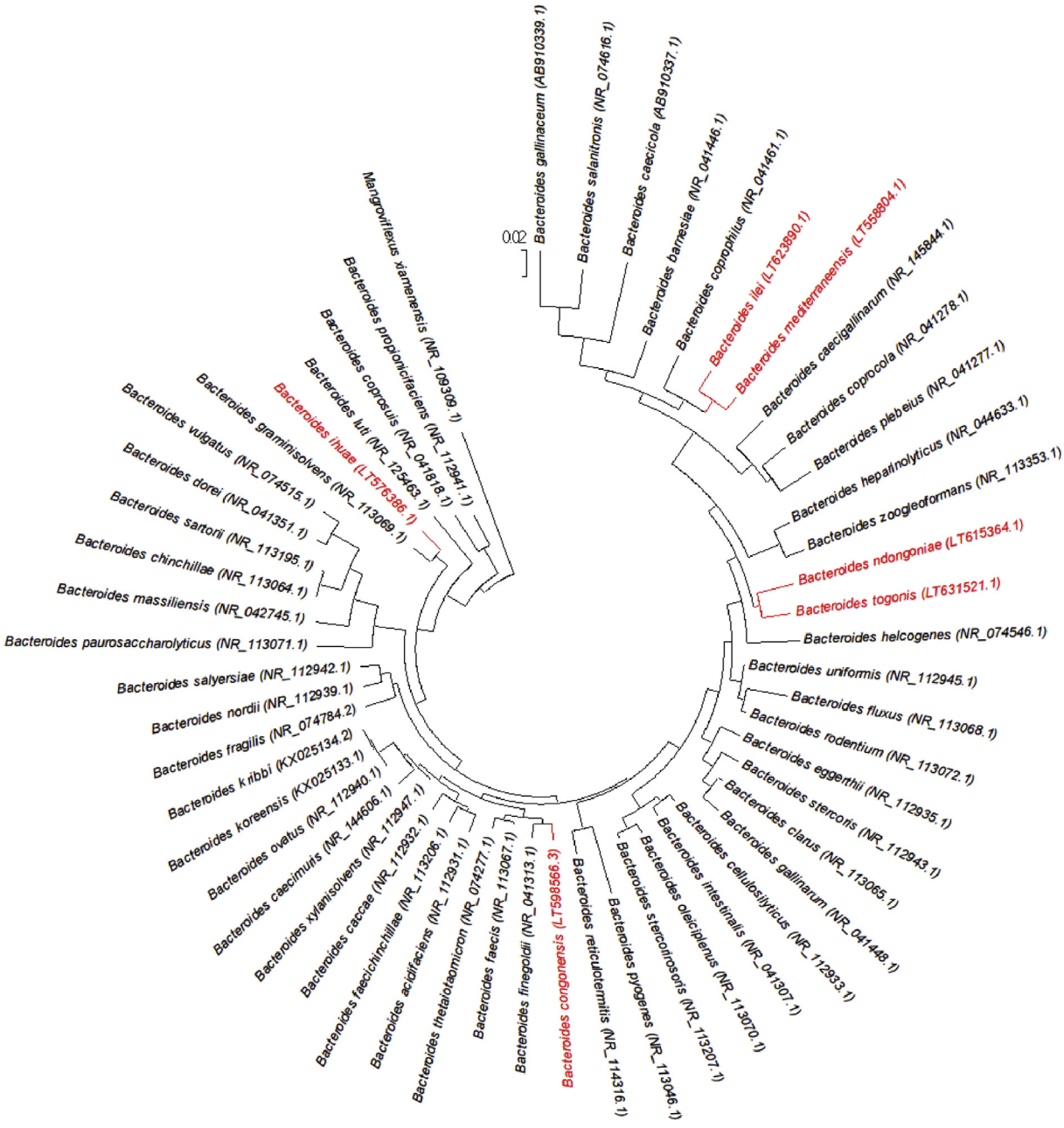

FIG. 2. Phylogenetic tree highlighting position of studied strains relative to other *Bacteroides* species. Respective GenBank accession numbers for 16S rRNA genes are indicated in parentheses. Sequences were aligned by Muscle v3.8.31 with default parameters and phylogenetic inferences were obtained using neighbour-joining method with 500 bootstrap replicates within MEGA6 software. Only bootstrap values > 95% are shown. *Mangroviflexus xiamenensis* was used as outgroup. Scale bar represents a 0.02% nucleotide sequence divergence.

predicted genes in COGs categories is shown in Table 6. We observed that the gene distribution in the 25 general COGs functional categories of the six new *Bacteroides* species is similar. The genomic characteristics of the studied strains were compared to the available genomes of closely related species in Table 7. This enabled us to observe that for all studied strains, the genome's size, percentage of G+C content and number of total genes are in the same range as those of other *Bacteroides* species.

Furthermore, digital DNA-DNA hybridization values (Table 8) between compared species except for the studied strains ranged from 17.90% to 34.50%. When the studied strains were compared to other *Bacteroides* species, the values were approximately in the same range: 19.70–29.80% for strain

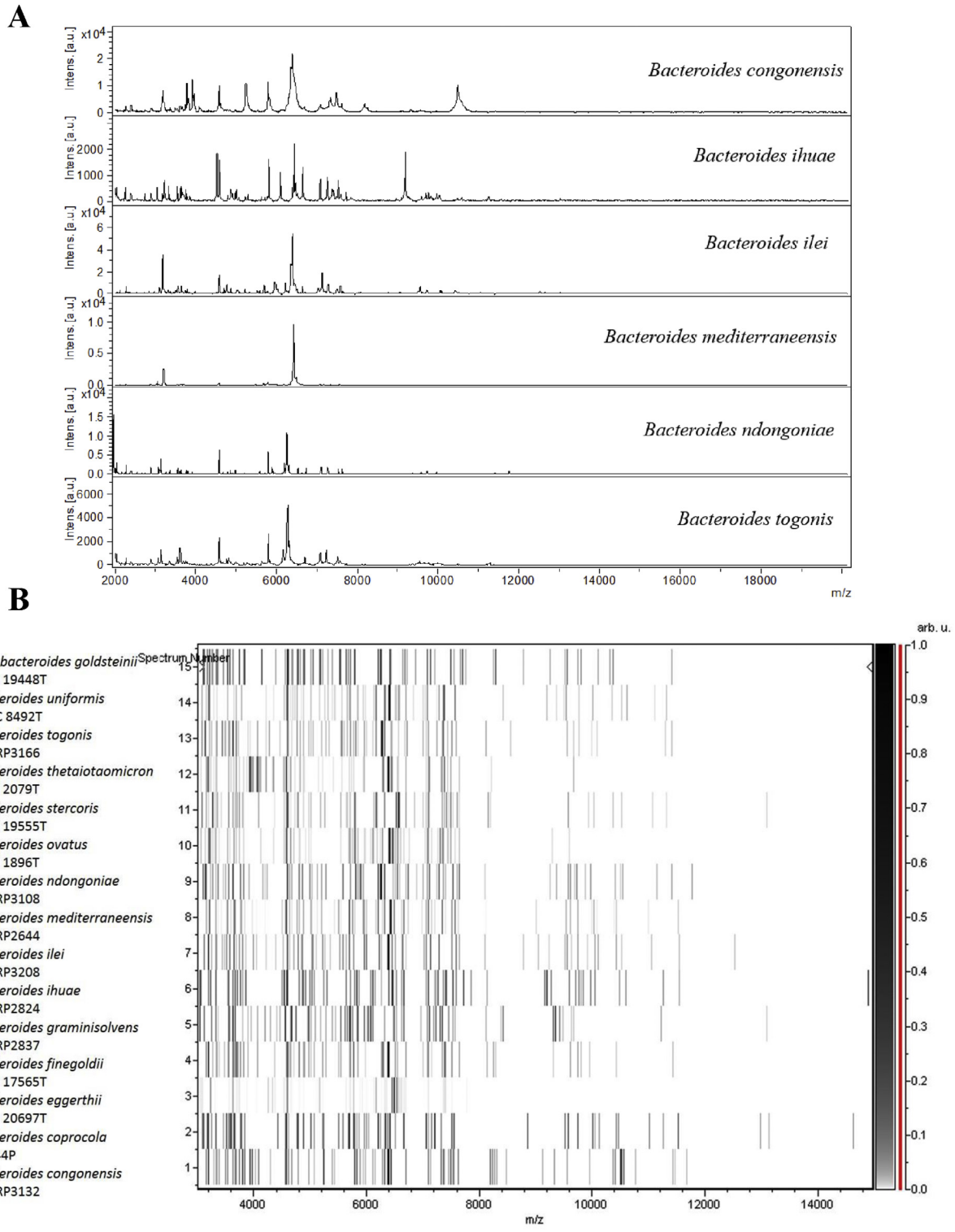

FIG. 3. (A) Reference mass spectra. Spectra from 12 individual colonies were compared and each reference spectrum generated. (B) Gel view comparing studied strains to other species within *Bacteroides* genus. Gel view displays raw spectra of loaded spectrum files arranged in a pseudo-gel–like look. *X*-axis records *m/z* value. Left *y*-axis displays running spectrum number originating from subsequent spectra loading. peak intensity is expressed by greyscale scheme code. Colour bar and right *y*-axis indicate relation between colour and peak intensity in arbitrary units. Displayed species are indicated at left.

TABLE 2. Comparison of phenotypic characteristics between nine *Bacteroides* species [36–38]

Property	*B. ilei*	*B. mediterraneensis*	*B. coprocola*	*B. ndongoniae*	*B. togonis*	*B. finegoldii*	*B. congonensis*	*B. ihuae*	*B. graminisolvens*
Strain	Marseille-P3208	Marseille-P2644	M16^T	Marseille-P3108	Marseille-P3166	199^T	Marseille-P3132	Marseille-P2824	XDT-1^T
Optimal temperature	37°C	37°C	37°C	37°C	37°C	37°C	37°C	30°C	30–35°C
Atmosphere	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic
pH range	5–7.5	5–8,	NA	7–7.5	7–7.5	NA	6–8.5	6–8.5	6.1–8.2
Colony aspect	Circular, translucent	Circular, grey, translucent, shiny	Disc shaped and greyish-white	Circular, grey, translucent, small	Small, circular and white	Circular, translucent-whitish, raised, convex	Smooth	Mucous	Thin, smooth-surfaced, with pearl-like, iridescent lustre
Cell shape	Fusiform	Rod	Rod	Rod	Rod	Rod	Rod	Rod	Rod
Cell length (µm)	3.9–5	0.9–2.3	1–4,	2.2–3.1	1.4–3.2	1.5–4.5	2–2.1	1.2–1.8	1.2–4.5
Cell width (µm)	0.5–0.6	0.6–0.7	0.8	0.6–1.1	0.6–0.7	0.8	2–2.1	0.5–0.6	0.4–0.6
Gram staining	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative	Negative
Salt tolerance (g.L^{-1})	0	0	NA	0	0	NA	50	5	0–40
Motility	–	+	–	–	+	–	+	+	–
Endospore formation	–	–	–	–	–	–	–	–	–
Major cellular fatty acid	15:0 anteiso	15:0 anteiso	15:0 anteiso	16:0.	15:0 iso	15:0 anteiso	15:0 anteiso	15:0 anteiso	15:0 anteiso
Production of:									
Alkaline phosphatase	+	+	+	+	+	+	+	+	NA
Catalase	–	+	–	+	–	NA	+	–	–
Oxidase	–	–	NA	–	–	NA	–	–	–
Urease	–	–	–	–	–	–	–	–	–
β-Galactosidase	+	+	+	+	+	+	+	+	NA
N-Acetyl-glucosamine	–	+	NA	–	+	NA	+	+	NA
Acid from:									
L-Arabinose	–	+	–	–	–	+	+	+	+
Ribose	–	–	NA	–	–	NA	+	+	w
Mannose	–	–	+	–	–	+	+	+	+
Mannitol	–	–	–	–	–	–	–	–	–
D-Saccharose	–	+	+	–	–	+	+	+	+
D-Glucose	–	+	+	–	–	+	+	+	+
D-Fructose	–	–	NA	–	–	NA	+	+	+
D-Maltose	–	+	+	–	–	+	+	+	+
D-Lactose	–	–	+	–	–	+	+	+	+
Habitat	Human ileum	Human ileum	Human stool	Human right colon	Human right colon	Human faeces	Human stool	Human sputum	Rice-straw residue

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

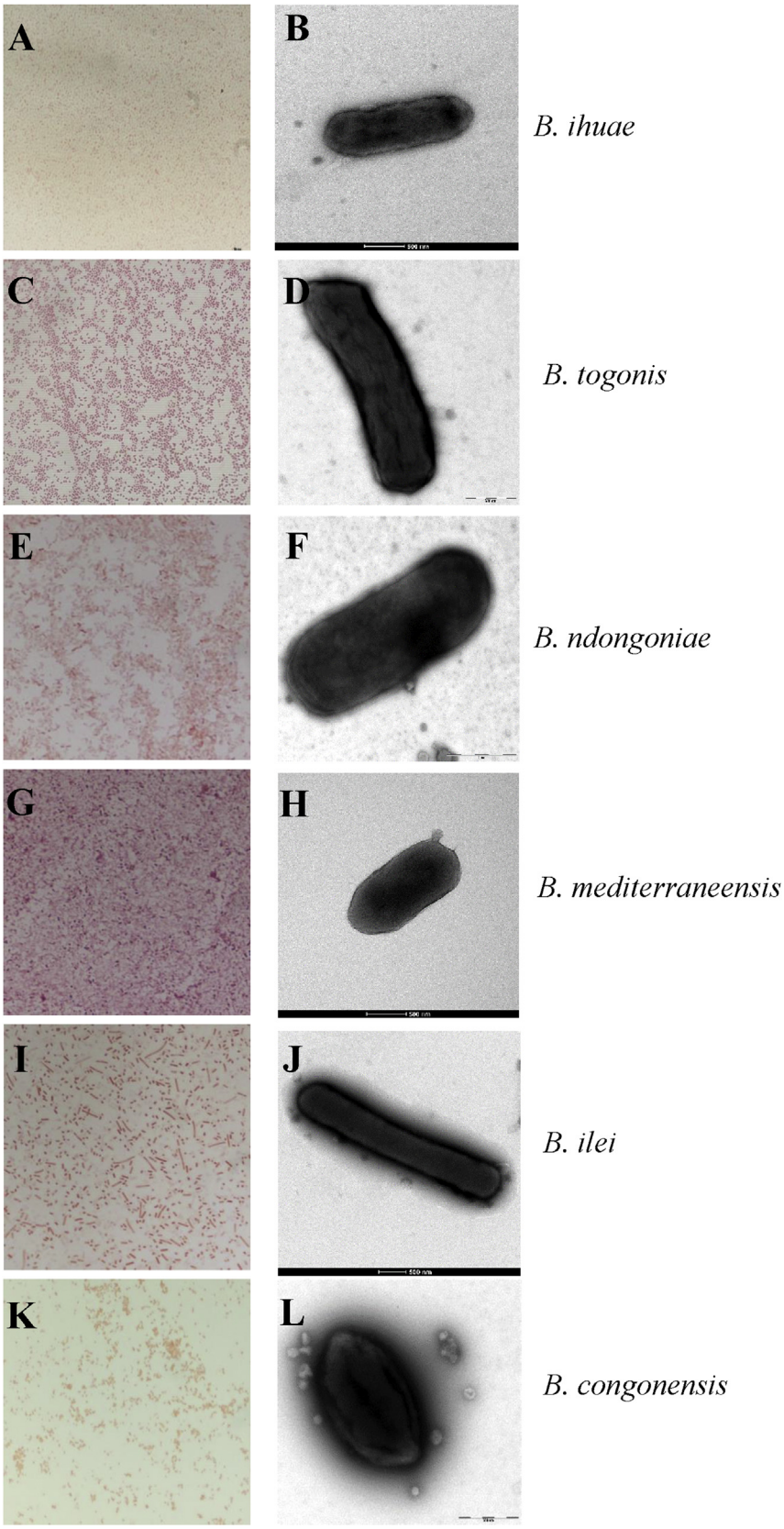

FIG. 4. Gram staining and electron micrographs of *Bacteroides* new species: *B. ihuae* (A, B), *B. togonis* (C, D), *B. ndongoniae* (E, F), *B. mediterraneensis* (G, H), *B. ilei* (I, J) and *B. congonensis* (K, L), respectively.

TABLE 3. Cellular fatty acid composition (in %[a]) of five *Bacteroides* species

Fatty acid		*B. ilei*	*B. mediterraneensis*	*B. ndongoniae*	*B. togonis*	*B. congonensis*	*B. ihuae*
Strain	IUPAC name	Marseille-P3208	Marseille-P2644	Marseille-P3108	Marseille-P3166	Marseille-P3132	Marseille-P2824
5:0 iso	3-methyl-Butanoic acid	No	No	No	No	1.4 ± 0.3	TR
5:0 anteiso	2-methyl-Butanoic acid	No	1.4 ± 0.2	No	TR	No	No
13:0 iso	11-methyl-Dodecanoic acid	TR	No	No	No	TR	TR
13:0 anteiso	10-methyl-Dodecanoic acid	TR	No	No	No	TR	TR
14:0.	Tetradecanoic acid	3.9 ± 0.2	1.2 ± 0.1	8.6 ± 0.2	4.7 ± 0.5	1.7 ± 0.1	1.4 ± 0.4
14:0 3-OH	3-hydroxy-Tetradecanoic acid	TR	No	No	No	No	No
14:0 iso	12-methyl-Tridecanoic acid	1.0 ± 0.1	No	No	No	1.8 ± 0.1	No
15:0.	Pentadecanoic acid	1.5 ± 0.2	1.4 ± 0.2	2.1 ± 0.2	2.4 ± 0.1	1.2 ± 0.1	TR
15:0 iso	13-methyl-Tetradecanoic acid	23.2 ± 0.7	13.9 ± 0.8	12.9 ± 0.5	26.4 ± 0.3	13.8 ± 0.2	4.0 ± 0.3
15:0 anteiso	12-methyl-Tetradecanoic acid	46.6 ± 3.2	39.2 ± 0.2	17.6 ± 0.7	15.1 ± 0.4	59.5 ± 1.5	46.7 ± 3.3
15:0 3-OH	3-hydroxy-Pentadecanoic acid	TR	TR	No	TR	TR	TR
15:0 3-OH iso	3-hydroxy-13-methyl-Tetradecanoic acid	1.2 ± 0.2	1.7 ± 0.1	No	TR	TR	TR
15:0 3-OH anteiso	3-hydroxy-12-methyl-Tetradecanoic acid	TR	TR	No	No	No	No
16:0.	Hexadecanoic acid	8.3 ± 0.6	4.4 ± 0.1	26.3 ± 1.3	15.8 ± 0.5	4.3 ± 0.2	4.0 ± 0.1
16:0 3-OH	3-hydroxy-Hexadecanoic acid	1.1 ± 0.3	4.4 ± 0.1	4.2 ± 0.7	5.2 ± 0.3	3.9 ± 0.3	2.5 ± 0.1
16:0 3-OH iso	3-hydroxy-14-methyl-Pentadecanoic acid	No	TR	No	No	TR	TR
16:0 9,10-methylene	2-hexyl-Cyclopropaneoctanoic acid	No	No	No	No	TR	TR
17:0	Heptadecanoic acid	No	TR	No	TR	No	No
17:0 iso	15-methyl-Hexadecanoic acid	No	TR	No	1.2 ± 0.1	No	No
17:0 anteiso	14-methyl-Hexadecanoic acid	No	TR	No	TR	No	TR
17:1n7 iso	15-Methylhexadec-9-enoic acid	No	No	No	No	No	TR
17:0 3-OH iso	3-hydroxy-15-methyl-Hexadecanoic acid	TR	16.3 ± 0.5	TR	6.6 ± 0.5	7.3 ± 0.4	17.8 ± 1.0
17:0 3-OH anteiso	3-hydroxy-14-methyl-Hexadecanoic acid	TR	5.2 ± 0.8	TR	TR	1.6 ± 0.1	12.8 ± 2.1
17:1n7 anteiso	14-Methylhexadec-9-enoic acid	No	No	No	No	No	TR
18:0.	Octadecanoic acid	2.2 ± 0.4	2.0 ± 0.1	10.0 ± 0.4	6.8 ± 0.1	TR	1.4 ± 0.9
18:1n7	11-Octadecenoic acid	No	No	No	No	No	TR
18:1n9	9-Octadecenoic acid	4.9 ± 0.6	3.2 ± 0.2	11.3 ± 0.7	8.7 ± 0.5	1.0 ± 0.2	2.1 ± 0.2
18:2n6	9,12-Octadecadienoic acid	3.5 ± 0.5	2.5 ± 0.1	5.8 ± 0.5	4.6 ± 0.1	TR	2.7 ± 0.3
20:4n6	5,8,11,14-Eicosatetraenoic acid	No	TR	No	No	No	No

[a]Mean peak area percentage ±standard deviation; TR, trace amounts <1%; no, not present.

TABLE 4. Antibiotic resistance test results of six *Bacteroides* species

Antibiotic	Disc load (µg/mL)	*B. ilei* Marseille-P3208	*B. mediterraneensis* Marseille-P2644	*B. ndongoniae* Marseille-P3108	*B. togonis* Marseille-P3166	*B. congonensis* Marseille-P3132	*B. ihuae* Marseille-P2824
Amoxicillin	25	S	S	S	S	S	S
Clindamycin	15	S	S	R	R	R	S
Imipenem	10	S	S	S	S	S	S
Gentamicin	15	R	R	R	R	R	R
Amoxicillin/clavulanic acid	30	S	S	S	S	S	S
Penicillin	10U	S	S	S	S	NA	S
Metronidazole	4	S	S	S	S	S	S
Trimethoprim/sulfamethoxazole	25	S	R	R	R	S	S
Oxacillin	5	S	R	R	R	R	R
Tobramycin	10	S	S	S	S	R	S
Ceftriaxone	30	S	S	S	S	R	R
Rifampicin	30	S	S	S	S	S	S
Doxycycline	30	S	S	S	S	NA	S
Erythromycin	15	S	R	R	R	R	S

NA, not available; R, resistant; S, susceptible.

Marseille-P2644, 18.50–20.40% for strain Marseille-P2824, 18.60–25.60% for strain Marseille-P3132, 19.40–50.00% for strain Marseille-P3108, 17.90–25.70% for strain Marseille-P3166 and 19.40–50.00% for strain Marseille-P3208. This confirms their new *Bacteroides* species status. These results are supported by values lower than 70%.

Conclusion

According to their phylogenetic, phenotypic and biochemical features and genomic comparisons, we propose the creation of six new species as members of the *Bacteroides* genus: *Bacteroides mediterraneensis* strain Marseille-P2644, *Bacteroides ihuae* strain Marseille-P2824, *Bacteroides togonis* strain Marseille-P3166, *Bacteroides ndongoniae* strain Marseille-P3108, *Bacteroides ilei* strain Marseille-P3208 and *Bacteroides congonensis* strain Marseille-P3132.

Description of *Bacteroides ilei* sp. nov.

Bacteroides ilei (i'le.i, L. gen. n., *ilei,* 'of the ileum,' the part of the digestive tract from which the bacterium was collected).

Cells are Gram-negative bacilli and have a fusiform shape with a length of 4 to 5 µm and a width of 0.5 to 0.6 µm. The

TABLE 5. Nucleotide content and gene count levels of genomes of six *Bacteroides* species

Attribute	*B. ilei*		*B. mediterraneensis*		*B. ndongoniae*		*B. togonis*		*B. congonensis*		*B. ihuae*	
	14; 18		18; 24		14; 14		23; 30		11; 11		6; 6	
Scaffolds/contigs	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]
Size (bp)	3 974 619	100	4 075 657	100	4 949 270	100	3 977 096	100	6 373 337	100	4 063 701	100
G + C content (mol%)	1 789 237	45.02	1 933 958	47.47	2 359 608	47.67	1 918 144	48.25	2 738 514	42.96	1 613 861	39.71
Coding region (bp)	3 615 554	90.96	3 687 552	90.47	4 396 977	88.84	3 590 610	90.28	5 824 153	91.38	3 669 831	90.30
Total genes	3517	100	3515	100	4285	100	3487	100	5067	100	3396	100
Protein-coding genes	3425	97.38	3419	97.26	4205	98.13	3419	98.04	4988	98.44	3332	98.11
RNA genes	92	2.61	96(6 5S, 3 16S, 1 23S rRNA, 86 tRNA)	2.73	80 (5 5S, 5 16S, 5 23S rRNA, 65 tRNA)	1.86	68 (4 5S, 1 16S, 1 23S rRNA, 62 tRNA)	1.95	79 (4 5S, 3 16S, 1 23S rRNA, 71 tRNA)	1.55	64 (4 5S, 1 16S, 1 23S rRNA, 58 tRNA)	1.88
Proteins with function prediction	2551	74.48	2814	82.30	3234	76.90	283	82.77	3815	76.48	2366	71.00
Proteins assigned to COGs	1636	47.76	1672	48.90	1796	42.71	1638	47.90	2211	44.32	1804	54.14
Proteins with peptide signals	842	24.58	907	26.52	1124	26.73	877	25.65	1701	34.10	978	29.35
No. of protein associated to ORFan	178	5.19	143	4.18	178	4.23	133	3.89	126	2.52	201	6.03
Genes with transmembrane helices	660	19.27	680	19.88	805	19.14	667	19.50	902	18.08	641	19.23
Genes associated with PKS or NRPS	10	0.29	8	0.23	9	0.21	9	0.26	9	0.18	6	0.18
No. of antibiotic resistance genes	0	0	2	0.05	4	0.09	3	0.08	1	0.02	0	0
No. of genes associated with Pfam-A domains	3018	85	3052	86	3654	85	2998	85	4494	88	3011	88

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

type strain did not exhibit catalase or oxidase activities. *Bacteroides ilei* is non–spore forming and nonmotile. Colonies are circular, small and translucent with a diameter of 0.8 to 1 mm. Optimum growth occurred in an anaerobic atmosphere at 37° C and pH 7 on Columbia agar enriched with 5% sheep's blood after 7 days of growth. Growth was observed at pH values between 5 to 8 and without NaCl.

Results of the API ZYM strip show that the strain possesses an activity for the following enzymes: alkaline phosphatase, esterase (C4), α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase and α-fucosidase. The other enzymes of the strip have no activities (Supplementary Table S2). Results of the API 50CH strip show that the strain is able to metabolize the following substrates: amygdalin, esculin ferric citrate, glycogen and potassium 5-ketogluconate. The others are negative (Supplementary Table S3). Finally, the results of the API 20A strip show a positive reaction only for β-glucosidase (Supplementary Table S4).

Bacteroides ilei is resistant to gentamicin and susceptible to other tested antibiotics (Table 4).

The major fatty acid is 12-methyl-tetradecanoic acid (47%).

The genome of strain Marseille-P3208 is 3 974 619 bp long with a 45.02 mol% G+C content. In the European Molecular Biology Laboratory–European Bioinformatics Institute (EMBL-EBI) database, the 16S rRNA gene and genome sequences are available under accession numbers LT623890 and FQSC00000000, respectively. The strain Marseille-P3208^T is the type strain of the species *Bacteroides ilei* (= CSUR P3208 = CCUG 69964) and was isolated from the ileum of a 76-year-old woman with oesophagitis in Marseille, France. This bacteria was also found in the left colon of the same patient.

Description of *Bacteroides mediterraneensis* sp. nov.

Bacteroides mediterraneensis (me.di.ter.ra.ne.en'sis, L. masc. adj., *mediterraneensis,* 'of Mediterranea,' the sea bordering Marseille, where the strain was isolated).

Cells are Gram-negative bacilli and are rod shaped with a length of 0.9 to 2.3 μm and a width of 0.6 to 0.7 μm. The type strain exhibited a catalase activity but no oxidase activity. *Bacteroides mediterraneensis* is non–spore forming but motile. Colonies are circular, translucent, grey and shiny, with a diameter of 0.6 mm. Optimum growth occurred in an anaerobic atmosphere at 37°C and pH 7 on Columbia agar enriched with 5% sheep's blood after 1 day of growth. Growth was observed at pH values between 5 to 8 and without NaCl.

Results of the API ZYM strip show that the strain possesses an activity for the following enzymes: alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-

TABLE 6. Number of genes associated with 25 general COGs functional categories for six *Bacteroides* species

Code	Description	*B. ilei*		*B. mediterraneensis*		*B. ndongoniae*		*B. togonis*		*B. congonensis*		*B. ihuae*	
		Value	% of total[a]	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]	Value	% of total[a]
J	Translation	180	5.25	178	5.20	182	4.32	179	5.23	194	3.88	190	5.70
A	RNA processing and modification	0	0	0	0	0	0	0	0	0	0	0	0
K	Transcription	76	2.21	93	2.72	103	2.44	107	3.12	162	3.24	115	3.45
L	Replication, recombination and repair	108	3.15	112	3.27	135	3.21	112	3.27	131	2.62	92	2.76
B	Chromatin structure and dynamics	0	0	0	0	0	0	0	0	0	0	0	0
D	Cell cycle control, mitosis and meiosis	26	0.75	23	0.67	29	0.68	27	0.78	25	0.50	25	0.75
Y	Nuclear structure	0	0	0	0	0	0	0	0	0	0	0	0
V	Defense mechanisms	70	2.04	83	2.42	101	2.40	86	2.51	88	1.76	83	2.49
T	Signal transduction mechanisms	62	1.81	71	2.07	82	1.95	80	2.33	118	2.36	76	2.28
M	Cell wall/membrane biogenesis	142	4.14	168	4.91	178	4.23	161	4.70	227	4.55	170	5.10
N	Cell motility	11	0.32	11	0.32	15	0.35	11	0.32	20	0.40	14	0.42
Z	Cytoskeleton	0	0	0	0	0	0	0	0	0	0	0	0
W	Extracellular structures	0	0	0	0	0	0	0	0	0	0	0	0
U	Intracellular trafficking and secretion	25	0.72	33	0.96	37	0.87	29	0.84	34	0.68	23	0.69
O	Post-translational modification, protein turnover, chaperones	75	2.18	80	2.33	76	1.80	70	2.04	81	1.62	77	2.31
X	Mobilome: prophages, transposons	47	1.37	14	0.40	35	0.83	21	0.61	51	1.02	16	0.48
C	Energy production and conversion	112	3.27	112	3.27	105	2.49	99	2.89	125	2.50	114	3.42
G	Carbohydrate transport and metabolism	162	4.72	156	4.56	176	4.18	154	4.50	257	5.15	190	5.70
E	Amino acid transport and metabolism	134	3.91	132	3.86	128	3.04	118	3.45	164	3.28	149	4.47
F	Nucleotide transport and metabolism	70	2.04	64	1.87	63	1.49	64	1.87	69	1.38	71	2.13
H	Coenzyme transport and metabolism	109	3.18	121	3.53	103	2.44	97	2.83	114	2.28	116	3.48
I	Lipid transport and metabolism	57	1.66	63	1.84	60	1.42	55	1.60	86	1.72	70	2.10
P	Inorganic ion transport and metabolism	102	2.97	111	3.24	110	2.61	106	3.10	225	4.51	148	4.44
Q	Secondary metabolites biosynthesis, transport and catabolism	20	0.58	23	0.67	19	0.45	16	0.46	32	0.64	28	0.84
R	General function prediction only	158	4.61	144	4.21	169	4.01	153	4.47	192	3.84	162	4.86
S	Unknown function	60	1.75	56	1.63	68	1.61	70	2.04	90	1.80	72	2.16
—	Not in COGs	1789	52.23	1747	51.09	2409	57.28	1781	52.09	2777	55.67	1528	45.85

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

TABLE 7. Genome comparison of closely related *Bacteroides* species.

Organism	Strain	INSDC	Size (Mb)	G+C (mol%)	Total genes
Bacteroides ilei	Marseille-P3208^T	FQSC00000000	3.97	45.02	3425
Bacteroides mediterraneensis	Marseille-P2644^T	FQRZ00000000	4.07	47.47	3419
Bacteroides coprophilus	DSM 18228	ACBW00000000	3.87	45.71	3939
Bacteroides coprocola	DSM 17136	ABIY00000000	4.30	41.86	4291
Bacteroides ndongoniae	Marseille-P3108^T	FNVV00000000	4.94	47.67	4205
Bacteroides togonis	Marseille-P3166^T	FQXX00000000	3.97	48.25	3419
Bacteroides congonensis	Marseille-P3132^T	FQXY00000000	6.37	42.96	4988
Bacteroides ihuae	Marseille-P2824^T	FNVX00000000	4.06	39.71	3332
Bacteroides graminisolvens	JCM15093	BAJS00000000	3.68	41.5	3403
Bacteroides eggerthii	JCM12986	ABVO00000000	4.19	44.6	3488
Bacteroides thetaiotaomicron	VPI5482	PUEO00000000	6.29	42.86	4825

INSDC, International Nucleotide Sequence Database Collaboration

phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase and α-fucosidase. The other enzymes of the strip have no activity (Supplementary Table S2). Results of the API 50CH strip show that the strain is able to metabolize the following substrates: D-arabinose, L-arabinose, D-xylose, D-glucose, methyl-αD-glucopyranoside, *N*-acetylglucosamine, esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-melibiose, D-saccharose, inulin, D-raffinose, amidon and potassium 5-ketogluconate. The others are negative (Supplementary Table S3). Finally, the results of the API 20A strip show positive reactions for the following: glucose, lactose, saccharose, maltose, salicin, xylose, arabinose, cellobiose, mannose, raffinose and β-glucosidase (Supplementary Table S4).

Bacteroides mediterraneensis was resistant to gentamicin, trimethoprim/sulfamethoxazole, oxacillin and erythromycin and susceptible to other tested antibiotics (Table 4). The major fatty acid is 12-methyl-tetradecanoic acid (39%).

The genome of strain Marseille-P2644 is 4 075 657 bp long with a 47.47 mol% G+C content. In the EMBL-EBI database, the 16S rRNA gene and genome sequences are available under accession numbers LT558804 and FQRZ00000000, respectively. The strain Marseille-P2644^T is the type strain of the species *Bacteroides mediterraneensis* (= CSUR P2644 = DSM 103033) and was isolated from the ileum of a 58-year-old woman undergoing colorectal cancer screening in Marseille, France. This bacterium was also found in the stomach and the right and left colon of the same patient.

Description of *Bacteroides ndongoniae* sp. nov.

Bacteroides ndongoniae (ndon.go'ni.ae, N.L. fem. gen. n., *ndongoniae,* 'of Ndongo,' the family name of Sokhna Ndongo, a Senegalese microbiologist involved in culturomics and in the description of new bacterial species).

TABLE 8. Pairwise comparison of *Bacteroides* species using GGDC, formula 2 (DDH estimates based on identities/HSP length),[a] upper right

	1	2	3	4	5	6	7	8	9	10	11
1	**100%**	25.60% ± 2.3	29.80% ± 2.4	24.60% ± 2.4	20.20% ± 2.3	25.00% ± 2.4	24.60% ± 2.4	23.40% ± 2.3	22.20% ± 2.3	19.70% ± 2.3	22.80% ± 2.3
2		**100%**	50.00% ± 2.6	22.80% ± 2.4	19.40% ± 2.3	25.20% ± 2.1	24.10% ± 2.4	24.90% ± 2.4	21.60% ± 2.3	19.70% ± 2.3	29.10% ± 2.4
3			**100%**	21.70% ± 2.3	19.50% ± 2.2	25.70% ± 2.4	26.00% ± 2.4	31.80% ± 2.5	21.20% ± 2.3	19.40% ± 2.2	24.20% ± 2.3
4				**100%**	19.10% ± 2.3	21.80% ± 2.3	24.00% ± 2.3	22.00% ± 2.3	25.60% ± 2.4	18.60% ± 2.2	20.70% ± 2.3
5					**100%**	19.00% ± 2.2	20.40% ± 2.3	18.60% ± 2.2	19.80% ± 2.3	18.50% ± 2.2	19.60% ± 2.3
6						**100%**	25.70% ± 2.4	24.60% ± 2.4	20.80% ± 2.3	17.90% ± 2.2	22.80% ± 2.4
7							**100%**	25.00% ± 2.4	23.30% ± 2.3	20.20% ± 2.3	34.50% ± 2.5
8								**100%**	21.60% ± 2.3	18.80% ± 2.2	22.70% ± 2.3
9									**100%**	18.50% ± 2.3	22.30% ± 2.3
10										**100%**	18.60% ± 2.3
11											**100%**

Bold indicates comparison between strain and itself.
DDH, DNA-DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; HSP, high-scoring segment pairs.
1, *B. mediterraneensis*; 2, *B. ilei*; 3, *B. ndongoniae*; 4, *B. congonensis*; 5, *B. ihuae*; 6, *B. togonis*; 7, *B. coprocola*; 8, *B. coprophilus*; 9, *B. thetaiotaomicron*; 10, *B. graminisolvens*; 11, *B. eggerthii*.
[a]Confidence intervals indicate inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size).

Cells are Gram-negative bacilli and are rod shaped with a length of 2.2 to 3.1 µm and a width of 0.6 to 1.1 µm. The type strain exhibits a catalase activity but no oxidase activity. *Bacteroides ndongoniae* is non–spore forming and also nonmotile. Colonies are circular, translucent, small and grey, with a diameter of 1 to 1.2 mm. Optimum growth occurred in an anaerobic atmosphere at 37°C and pH 7 on Columbia agar enriched with 5% sheep's blood after 7 days of growth. Growth was observed at pH values between 7 to 7.5 and without NaCl.

Results of the API ZYM strip show that the strain possesses an activity for the following enzymes: alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase and α-fucosidase. The other enzymes of the strip have no activity (Supplementary Table S2). Results of the API 50CH strip show that the strain is able to metabolize the following substrates: esculin ferric citrate, D-turanose, potassium 5-ketogluconate. The others are negative (Supplementary Table S3). Finally, the results of the API 20A strip show positive reactions for the following: glucose, lactose, saccharose, maltose, xylose, arabinose, cellobiose, mannose, raffinose, rhamnose and β-glucosidase (Supplementary Table S4).

Bacteroides ndongoniae was resistant to clindamycin, gentamicin, trimethoprim/sulfamethoxazole, oxacillin and erythromycin and susceptible to other tested antibiotics (Table 4). The major fatty acid is hexadecanoic acid (26%).

The genome of strain Marseille-P3108 is 4 949 270 bp long with a 47.67 mol% G+C content. In the EMBL-EBI database, the 16S rRNA gene and genome sequences are available under accession numbers LT615364 and FNVV00000000, respectively. The strain Marseille-P3108^T is the type strain of the new species *Bacteroides ndongoniae* (= CSUR P3108 = DSM 103636) and was isolated from the right colon of a 76-year-old woman with oesophagitis in Marseille, France. This bacterium was also found in the left colon of the same patient.

Description of *Bacteroides togonis* sp. nov.

Bacteroides togonis (to.go'nis, N. L. masc. gen. n., *togonis,* named in honor of Amadou Togo, a Malian microbiologist involved in culturomics and in the description of new bacterial species).

Cells are Gram-negative bacilli and are rod shaped with a length of 1.4 to 3.2 µm and a width of 0.6 to 0.7 µm. The type strain did not exhibit activities for catalase and oxidase. *Bacteroides togonis* is non–spore forming but motile. Colonies are small, circular and white, with a diameter of 0.6 to 0.8 mm. Optimum growth occurred in an anaerobic atmosphere at 37°C and pH 7 on Columbia agar enriched with 5% sheep's blood after 3 days of growth. Growth was observed at pH values between 7 to 7.5 and without NaCl.

Results of the API ZYM strip show that the strain possesses activity for the following enzymes: alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, valine arylamidase, cystine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase, *N*-acetyl-β-glucosaminidase and α-fucosidase. The other enzymes of the strip had no activity (Supplementary Table S2). Results of the API 50CH strip show that the strain is able to metabolize the following substrates: D-arabinose, *N*-acetylglucosamine, amygdalin, esculin ferric citrate, D-melibiose, D-melezitose, xylitol, gentiobiose, D-tagatose, potassium gluconate, potassium 5-ketogluconate. The others were negative (Supplementary Table S3). Finally, the results of the API 20A strip show positive reactions for the following: glucose, lactose, saccharose, maltose, xylose, arabinose, mannose, raffinose, rhamnose and β-glucosidase (Supplementary Table S4).

Bacteroides togonis was resistant to amoxicillin, clindamycin, gentamicin, trimethoprim/sulfamethoxazole, oxacillin and erythromycin and susceptible to other tested antibiotics (Table 4). The major fatty acid is 13-methyl-tetradecanoic acid (26%).

The genome of strain Marseille-P3166 is 3 977 096 bp long with a 48.26 mol% G+C content. In the EMBL-EBI database, the 16S rRNA gene and genome sequences are available under accession numbers LT631521 and FQXX00000000, respectively. The strain Marseille-P3166^T is the type strain of the species *Bacteroides togonis* (= CSUR P3166 = DSM 103637) and was isolated from the right colon of a 76-year-old woman with oesophagitis in Marseille, France.

Description of *Bacteroides congonensis* sp. nov.

Bacteroides congonensis (con.go'ne.n'sis, L. masc. adj., *congonensis,* 'from Congo,' the country where the sample was collected).

Cells are Gram-negative bacilli and are rod shaped with a length of 1 to 1.2 µm and a width of 0.6 to 0.7 µm. The type strain exhibits catalase activity but no oxidase activity. *Bacteroides congonensis* is non−spore forming and motile. Colonies are smooth, with a diameter of 0.8 to 1 mm. Optimum growth occurred in an anaerobic atmosphere at 37°C and pH 7 on Columbia agar enriched with 5% sheep's blood after 10 days of growth. Growth was observed at pH values between 6 to 8.5 and at a concentration of 50 g/L of NaCl.

Results of the API ZYM strip show that the strain possesses an activity for the following enzymes: alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, *N*-acetyl-β-glucosaminidase and α-fucosidase. The other enzymes of the strip had no activity (Supplementary Table S2). Results of the API 50CH strip showed that the strain is able to metabolize most of the substrates except: erythritol, L-xylose, D-adonitol, methyl-βD-xylopyranoside, L-sorbose, dulcitol, inositol, D-mannitol, methyl-αD-mannopyranoside, methyl-αD-glucopyranoside, arbutin, salicin, D-trehalose, inulin, D-melezitose, D-turanose, D-lyxose, D-tagatose, D-fucose, D-arabitol, L-arabitol, potassium gluconate and potassium 2-ketogluconate (Supplementary Table S3). Finally, the results of the API 20A strip show positive reactions for all the studied reactions except for the formation of indole, urease, hydrolysis of gelatin and esculin (Supplementary Table S4).

Bacteroides congonensis was susceptible to rifampicin, amoxicillin, imipenem, amoxicillin/clavulanic acid, metronidazole and trimethoprim/sulfamethoxazole and resistant to other tested antibiotics (Table 4). The major fatty acid is 12-methyl-tetradecanoic acid (60%).

The genome of strain Marseille-P3132 is 6 373 337 bp long with a 42.96 mol% G+C content. In the EMBL-EBI database, the 16S rRNA gene and genome sequences are available under accession numbers LT598566 and FQXY00000000, respectively. The strain Marseille-P3132^T is the type strain of the species *Bacteroides congonensis* (= CSUR P3132 = CCUG 70144) and was isolated from the stool of a 35-year-old healthy pygmy woman in Congo.

Description of *Bacteroides ihuae* sp. nov.

Bacteroides ihuae (i.hu'ae, N. L. gen. n., ihuae, 'of IHU' (Institut Hospitalo-Universitaire), where the type strain was cultivated).

Cells are Gram-negative bacilli and are rod shaped with a length of 1 to 1.6 µm and a width of 0.5 to 0.6 µm. The type strain does not exhibit activities for catalase or oxidase. *B. ihuae* is non−spore forming and motile. Colonies are mucous, with a diameter of 1 to 3 mm. Optimum growth occurred in an anaerobic atmosphere at 30°C and pH 7 on Columbia agar enriched with 5% sheep's blood after 10 days of growth. Growth was observed at pH values between 6 to 8.5 and at a concentration of 5g/l of NaCl.

Results of the API ZYM strip show that the strain possesses an activity for the following enzymes: alkaline phosphatase, esterase (C4), esterase lipase (C8), acid phosphatase, naphthol-AS-BI-phosphohydrolase, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase and *N*-acetyl-β-glucosaminidase. The other enzymes of the strip had no activity (Supplementary Table S2). Results of the API 50CH strip show that the strain is able to metabolize most of the substrates except: erythritol, D-arabinose, L-xylose, D-adonitol, methyl-βD-xylopyranoside, L-sorbose, dulcitol, inositol, methyl-αD-mannopyranoside, D-melezitose, D-lyxose, D-fucose, L-fucose, D-arabitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate and potassium 5-ketogluconate (Supplementary Table S3). Finally, the results of the API 20A strip show positive reactions for all studied reactions except for indole formation, urease, mannitol, hydrolysis of gelatin, glycerol, melezitose, sorbitol and rhamnose (Supplementary Table S4).

B. ihuae was resistant to gentamicin, oxacillin, fosfomycin and ceftriaxone and susceptible to others tested antibiotics (Table 4). The major fatty acid is 12-methyl-tetradecanoic acid (47%).

The genome of strain Marseille-P2824 is 4 063 701 bp long with a 39.71 mol% G+C content. In the EMBL-EBI database, the 16S rRNA gene and genome sequences are available under accession numbers LT576386 and FNVX00000000, respectively. The strain Marseille-P2824^T is the type strain of the species *Bacteroides ihuae* (= CSUR P2824 = CCUG 70550) was isolated from the sputum of a healthy 27-year-old woman in Marseille, France.

Acknowledgements

The authors thank the Xegen Company (http://www.xegen.fr/) for automating the genomic annotation process. We thank M. Lardière for English-language editorial work, A. Caputo for submitting the genomic sequences to GenBank and A. Oren for providing etymology corrections.

This study was funded by the Fondation Méditerranée Infection. This work was also supported by the French government under the 'Investissements d'avenir' program managed by the Agence Nationale de la Recherche (Méditerranée Infection 10-IAHU-03).

Conflict of interest

None declared.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.nmni.2018.06.006.

References

[1] Castellani A, Chalmers AJ. Manual of tropical medicine. 3rd ed. New York: Williams Wood; 1919.
[2] Veillon A, Zuber A. Recherches sur quelques microbes strictement anaérobies et leur rôle en pathologie. In: Masson G, editor. Archives de médecine expérimentale et d'anatomie pathologique, vol. 10; 1898. p. 517–45. Paris.
[3] Wexler HM. Bacteroides: the good, the bad, and the nitty-gritty. Clin Microbiol Rev 2007;20:593–621.
[4] Akin H, Tözün N. Diet, microbiota, and colorectal cancer. J Clin Gastroenterol 2014;48:S67–9.
[5] Gagnière J, Raisch J, Veziant J, Barnich N, Bonnet R, Buc E, et al. Gut microbiota imbalance and colorectal cancer. World J Gastroenterol 2016;22:501–18.
[6] 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.
[7] Lagier JC, Hugon P, Khelaifia S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 2015;28:237–64.
[8] Lagier JC, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 2016;1:16203.
[9] Lagier JC, Edouard S, Pagnier I, Mediannikov O, Drancourt M, Raoult D. Current and past strategies for bacterial culture in clinical microbiology. Clin Microbiol Rev 2015;28:208–36.
[10] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 2015;36:73–8.
[11] Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[12] 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.
[13] Dubourg G, Cimmino T, Senkar SA, Lagier JC, Robert C, Flaudrops C, et al. Noncontiguous finished genome sequence and description of *Paenibacillus antibioticophila* sp. nov. GD11(T), the type strain of *Paenibacillus antibioticophila*. New Microbe New Infect 2015;8:137–47.
[14] Morel AS, Dubourg G, Prudent E, Edouard S, Gouriet F, Casalta JP, et al. Complementarity between targeted real-time specific PCR and conventional broad-range 16S rDNA PCR in the syndrome-driven diagnosis of infectious diseases. Eur J Clin Microbiol Infect Dis 2015;34:561–70.
[15] Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 2000;38:3623–30.
[16] 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;64:346–51.
[17] Sasser M. Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). Newark, NY: Microbial ID; 2006.
[18] Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N, et al. Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbe New Infect 2016;10:66–76.
[19] Matuschek E, Brown DF, Kahlmeter G. Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect 2014;20:O255–66.
[20] Zerbino DR, Birney E. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 2008;18:821–9.
[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 2012;19:455–77.
[22] Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. GigaScience 2012;1:18.
[23] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–20.
[24] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 2010;11:119.
[25] 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.
[26] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[27] Käll L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 2004;338:1027–36.
[28] Eddy SR. Accelerated profile HMM searches. PLoS Comput Biol 2011;7: e1002195.
[29] Conway KR, Boddy CN. ClusterMine360: a database of microbial PKS/NRPS biosynthesis. Nucleic Acids Res 2013;41:D402–7.
[30] Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 2014;58:212–20.

[31] Gouret P, Paganini J, Dainat J, Louati D, Darbo E, Pontarotti P, et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBAH. In: Pontarotti P, editor. Evolutionary biology—concepts, biodiversity, macroevolution and genome evolution. Berlin: Springer; 2011. p. 71–87.

[32] 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 Bioinform 2005;6:198.

[33] Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinform 2011;12:124.

[34] Gouret P, Thompson JD, Pontarotti P. PhyloPattern: regular expressions to identify complex patterns in phylogenetic trees. BMC Bioinform 2009;10:298.

[35] Rossi-Tamisier M, Benamar S, Raoult D, Fournier PE. Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. Int J Syst Evol Microbiol 2015;65:1929–34.

[36] Nishiyama T, Ueki A, Kaku N, Watanabe K, Ueki K. *Bacteroides graminisolvens* sp. nov., a xylanolytic anaerobe isolated from a methanogenic reactor treating cattle waste. Int J Syst Evol Microbiol 2009;59:1901–7.

[37] Bakir MA, Kitahara M, Sakamoto M, Matsumoto M, Benno Y. *Bacteroides finegoldii* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol 2006;56:931–5.

[38] Kitahara M, Sakamoto M, Ike M, Sakata S, Benno Y. *Bacteroides plebeius* sp. nov. and *Bacteroides coprocola* sp. nov., isolated from human faeces. Int J Syst Evol Microbiol 2005;55:2143–7.