Anaerobe 43 (2017) 47–55

Contents lists available at ScienceDirect

Anaerobe

journal homepage: www.elsevier.com/locate/anaerobe

Anaerobes in the microbiome

Blautia massiliensis sp. nov., isolated from a fresh human fecal sample and emended description of the genus *Blautia*

Guillaume A. Durand [a, b], Thao Pham [a], Sokhna Ndongo [a], Sory Ibrahima Traore [a], Grégory Dubourg [a, b], Jean-Christophe Lagier [a, b], Caroline Michelle [a], Nicholas Armstrong [a], Pierre-Edouard Fournier [a, b], Didier Raoult [a, b, c], Matthieu Million [a, b, *]

[a] URMITE UM63, CNRS7278, IRD198, INSERM1095, Faculté de Médecine, Aix Marseille Université, 27 Boulevard Jean Moulin, 13385, Marseille Cedex 5, France
[b] Pôle des Maladies Infectieuses, Hôpital La Timone, Assistance Publique-Hôpitaux de Marseille, Marseille, France
[c] Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

Article history:
Received 15 October 2016
Received in revised form
1 December 2016
Accepted 2 December 2016
Available online 5 December 2016

Handling Editor: Elisabeth Nagy

Keywords:
Blautia massiliensis sp. nov.
Taxono-genomics
Culturomics
Anaerobe
Gut microbiota

ABSTRACT

The strain GD9^T is the type strain of the newly proposed species *Blautia massiliensis* sp. nov., belonging to the family *Lachnospiraceae*. It was isolated from a fresh stool sample collected from a healthy human using the culturomics strategy. Cells are Gram-negative rods, oxygen intolerant, non-motile and non-spore forming. The 16S rRNA gene sequencing showed that strain GD9^T was closely related to *Blautia luti*, with a 97.8% sequence similarity. Major fatty acids were C14:0 (19.8%) and C16:0 (53.2%). Strain GD9^T exhibits a genome of 3,717,339 bp that contains 3,346 protein-coding genes and 81 RNAs genes including 63 tRNAs. The features of this organism are described here, with its complete genome sequence and annotation. Compared with other *Blautia* species which are Gram positive, the strain was Gram negative justifying an emended description of the genus *Blautia*.

1. Introduction

The description of the human gut microbiota as detailed in the Human Microbiome Project [1] is an important concern for microbiologists, because of its importance to humans (53% of human-linked species were cultivated from the gut [2]) and its potential link to diseases, such as chronic inflammatory bowel disease [3–6]. The use of molecular tools with pyrosequencing is a widespread approach, but suffers from several biases including an extraction bias or the detection of the most abundant species, thus neglecting minority but important species (the 'depth' bias). The combination of cultivation methods in tandem with pyrosequencing has been demonstrated to be successful in the isolation of novel taxa [7,8]. The species described here was isolated from fresh human stool of a healthy volunteer using the anaerobic conditions of the culturomics approach. The species belongs to *Blautia*, a genus including 13 species, proposed by Liu et *al.* [9]. Million et al. [10] had previously shown that malnutrition was linked to a decrease of the relative abundance of the *Blautia* genus in the human gut. Here, we describe the type strain GD9^T of the species for which we propose the name *Blautia massiliensis* (=CSUR P2132, = DSM 101187).

2. Materials and methods

2.1. Ethics and sample collection

The stool specimen was collected at La Timone hospital in Marseille (France) in April 2015 from a 28-year-old healthy French male (BMI 23.2 kg/m^2), without current treatment. Informed and

* Corresponding author. URMITE UM63, CNRS7278, IRD198, INSERM1095, Aix Marseille Université, Fondation Méditerranée Infection, 19-21 boulevard Jean Moulin, 13385, Marseille Cedex 5, France.
E-mail address: matthieumillion@gmail.com (M. Million).

http://dx.doi.org/10.1016/j.anaerobe.2016.12.001
1075-9964/

Abbreviations

URMITE	Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes
CSUR	Collection de Souches de l'Unité des Rickettsies
DSM	Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany
ATCC	American Type Culture Collection
MALDI-TOF MS	Matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry
FAME	Fatty Acid Methyl Ester
GC/MS	Gas Chromatography/Mass Spectrometry
TE buffer	Tris-EDTA buffer
AGIOS	Average Genomic Identity Of the gene Sequence
GGDC	Genome-to-Genome Distance Calculator
dDDH	digital DNA:DNA hybridization

signed consent, approved by the Institut Fédératif de Recherche 48 (Faculty of Medicine, Marseille, France), was obtained under agreement number 09-022.

2.2. Isolation, identification and growth conditions of the strain

The stool specimen was directly inoculated on Columbia agar supplemented with sheep blood (Cos) after dilutions, and then incubated at 37 °C in anaerobic conditions. Subculturing was performed on days one, two, five and ten on Cos at 37 °C. Identification was performed using MALDI-TOF MS, comparing the spectrum with our database (which includes the Bruker database and our own collection), as previously described [11]. When the identification failed (score < 1.7), the 16S rRNA gene was amplified and sequenced, as previously described [12]. The 16S rRNA sequence was compared to the nucleotide database using the BLAST similarities web-service. In case of a sequence similarity value lower than 98.65%, the species was suspected to be novel, albeit without performing DNA-DNA hybridization, as previously suggested [13,14]. Characterization of growth conditions was tested as previously described [11]. Sporulation and different culture conditions were tested in order to determine the best culture conditions [15].

2.3. Morphological and biochemical characterization

Morphological characterization was first performed by microscopic observation of Gram staining and motility of the fresh sample. Negative staining was performed after bacterial fixation in glutaraldehyde 2.5%. This solution was deposited on carbon formvar film incubated for one second on ammonium molybdate 1%, dried on blotting paper and finally observed using TECNAI G20 transmission electron microscope (FEI Company, Limeil-Brevannes, France) at an operating voltage of 200 keV. Biochemical features, such as oxidase, catalase, API 20A, API ZYM and 50CH galleries (Biomerieux, Marcy l'Etoile, France) were investigated, according to the manufacturer's instructions. Cellular FAME analysis was performed by GC/MS. Two samples were prepared with approximately 30 mg of bacterial biomass per tube harvested from several culture plates, grown under anaerobic conditions on Cos agar for 48 h. FAME were prepared as described previously [16]. GC/MS analyses were carried out as described before [17]. Briefly, fatty acid methyl esters were separated using an Elite 5-MS column and monitored by mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). Spectral database search was performed using MS Search 2.0 operated with the Standard Reference Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK). The profile obtained was compared to *Blautia faecis* M25^T, *Blautia stercoris* GAM6-1^T, *Blautia coccoides* DSM 935^T, *Blautia hansenii* DSM 20583^T, *Blautia schinkii* DSM 10518^T, *Blautia obeum* DSM 25238^T, and *Blautia glucerasea* DSM 22028^T [18].

2.4. Antibiotic susceptibility

The susceptibility to classical antibiotics was tested with the diffusion method according to CASFM/Eucast 2015 recommendations for fastidious anaerobes [19]. A suspension of 1 McFarland of species was grown on Wilkins-Chalgren agar (Sigma Aldrich, Steinheim, Germany) supplemented with 5% sheep blood. Incubations were performed under anaerobic conditions at 37 °C and reading was done after 48 h using Sirscan system© (i2a, Montpellier, France). Inhibition diameters were controlled by manual measurement using a ruler.

2.5. Genome sequencing, annotation and comparison

The strain GD9^T was cultivated on Cos at 37 °C under anaerobic atmosphere and then resuspended in 400 μL of TE buffer. Then, 200 μL of this suspension was diluted in 1 ml TE buffer for lysis treatment that included a 30 min incubation with 2.5 μg/μL lysozyme at 37 °C, followed by an overnight incubation with 20 μg/μL proteinase K at 37 °C [20]. Extracted DNA was then purified using three successive phenol-chloroform extractions and ethanol precipitations at −20 °C overnight. After centrifugation, the DNA was resuspended in 160 μL TE buffer. Quantification and sequencing of the whole genome was done on the MiSeq Technology (IlluminaInc, San Diego, CA, USA) with the mate pair strategy as previously described [21]. The open reading frames (ORFs) were predicted using Prodigal [22] with default parameters but the predicted ORFs were excluded if they were spanning a sequencing gap region (contain N). The predicted bacterial protein sequences were searched against the Clusters of Orthologous Groups (COG) using BLASTP (E-value 1e-03, coverage 0.7 and identity percent 30%). If no hit was found, a search was performed against the NR database using BLASTP with an E-value of $1e^{-03}$ coverage 0.7 and an identity percent of 30%. If the sequence lengths were smaller than 80 amino acids, we used an E-value of $1e^{-05}$. The tRNAScanSE tool [23] was used to find tRNA genes, whereas ribosomal RNAs were found using RNAmmer [24]. Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [25]. ORFans were identified if all the performed BLASTP did not give positive results (*E*-value smaller than $1e^{-03}$ for ORFs with sequence size greater than 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 [26,27]. For each selected genome, the complete genome sequence, proteome genome sequence and ORFeome genome sequence were retrieved from the FTP of NCBI. All proteomes were analyzed with proteinOrtho [28]. Then, for each genomes pair, a similarity score was computed. This score is the mean value of nucleotide similarity between all pairs of orthologues between the two genomes studied (AGIOS) [29]. An annotation of the entire proteome was performed to define the distribution of the functional classes of predicted genes according to the clusters of orthologous groups of proteins (using the same method as for the genome annotation). Annotation and comparison of genome size, G + C content, and gene content with other close species were performed in the Multi-Agent software system DAGOBAH [30], that include Figenix libraries. To

assess the affiliation of our novel strain to the type strains of known species with available genome, the Genome-to-Genome Distance Calculator web service was used to calculate digital DNA:DNA hybridization estimates (dDDH) with confidence intervals under recommended settings (Formula 2, BLAST+) [31,32].

3. Results and discussion

3.1. Classification and features

The first isolation of type strain GD9^T occurred after direct inoculation of fresh stool on Cos agar, without enrichment into blood bottles. The MALDI-TOF spectrum neither matched against our database nor Brucker's one (Supplementary Fig. 1). The 16S rRNA gene is 1,493bp long (accession number: AA00076 from 16S IHU bank, LN890282 from EBI Sequence Database), with BLASTN search against reference sequences indicating *Blautia luti* DSM 14534^T (NR_041960) as the most closely cultured species at 97.8% (Fig. 1) [33]. The pairwise comparison of our strain with all type strains of the genus *Blautia* is represented in Table 1. Close species on the basis of 16S rRNA tree and their presence into our MALDI-TOF spectrum database were compared at the protein level with *B. massiliensis* GD9^T and represented in a gel view (Fig. 2).

Optimal growth was at 37 °C after 48 h under anaerobic conditions. Colonies appeared to be smooth, white, non-hemolytic, non-motile, non spore-forming and 1 mm in size. Optical microscopic observation showed bacilli with Gram-negative staining (Supplementary Fig. 2). Electronic microscopy showed small rods of about 1 μm (Supplementary Fig. 3). Classification of the strain and main characteristics are presented in Table 2.

We observed no production reaction for catalase and oxidase. Using API 20A, positive reactions were found with D-Glucose, D-Mannitol, D-Lactose, D-Saccharose, D-Maltose, Salicine, D-Xylose, L-Arabinose, Esculine, Glycerol, D-Cellobiose, D-Mannose, D-Melezitose, D-Raffinose, D-Sorbitol, L-Rhamnose and D-Trehalose. Using API ZYM strips, positive reactions were observed with α-Galactosidase, β-Galactosidase, α-Glucosidase and β-Glucosidase. Using API 50CH strips, positive reactions were found with Erythritol, D-Arabinose, D-Ribose, L-Xylose, L-Rhamnose, Dulcitol, *N*-Acetylglucosamine, Amygdaline, Arbutine, Inuline, Amidon, Glycogene, Gentiobiose, D-Lyxose, D-Tagatose, D-Fucose, L-Fucose, Potassium Gluconate and Potassium 5-Cetogluconate. A comparison of phenotypic and biochemical characteristics was made with other representatives of the family *Lachnospiraceae* (Table 3). The major fatty acids found for this strain were C16:0 (53%) and C14:0 (20%). Saturated fatty acids were the most abundant and represented 83% of the fatty acids found (Table 4). The cellular fatty acids of *Blautia massiliensis* GD9^T were compared with the profiles of 7 other species of the genus *Blautia* retrieved from the literature: *B. faecis* M25^T, *B. stercoris* GAM6-1^T, *B. coccoides* DSM 935^T, *B. hansenii* DSM 20583^T, *B. schinkii* DSM 10518^T, *B. obeum* ATCC 29174^T, and *B. glucerasea* DSM 22028^T (profiles described by Park et al., 2013

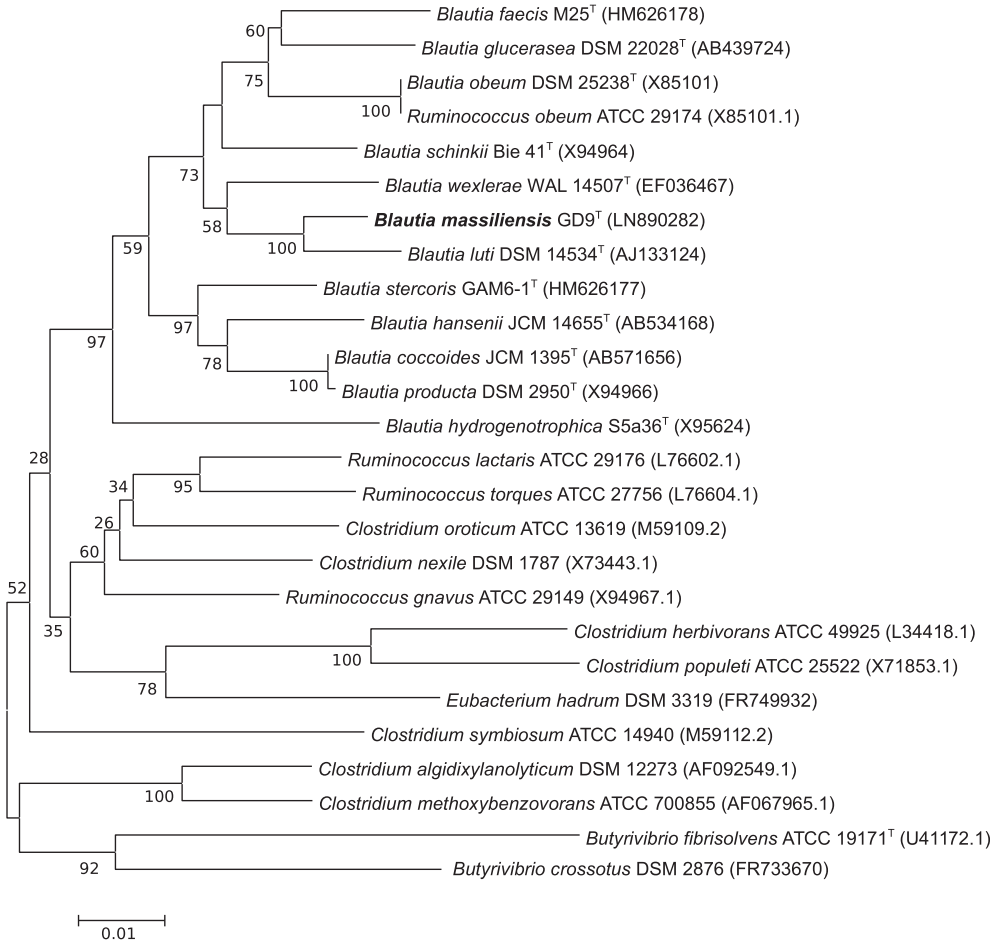

Fig. 1. Phylogenetic tree highlighting the position of *B. massiliensis* strain GD9^T relative to other phylogenetically close type strains. Genbank accession numbers of the 16S rRNA gene reference sequences are indicated in parenthesis. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained with kimura two parameter model using neighbor-joining method with 1000 bootstrap replicates, within MEGA6 software. The scale bar represents a 1% nucleotide sequence divergence.

Table 1
Pairwise comparison between *Blautia massiliensis* and type strains within the genus *Blautia*.

Type strains (sequence accession number)[a]	BLAST similarity compared with *B. massiliensis* (LN890282)
B. coccoides (AB571656)	1409/1498 (94.1%)
B. faecis (HM626178)	1216/1264 (96.2%)
B. glucerasea (AB439724)	1392/1470 (94.7%)
B. hansenii (AB534168)	1395/1493 (93.4%)
B. hydrogenitrophica (X95624)	1357/1464 (92.7%)
B. luti (AJ133124)	1307/1336 (97.8%)
B. obeum (X85101)	1398/1460 (95.7%)
B. producta (X94966)	1385/1478 (93.7%)
B. schinkii (X94964)	1390/1475 (94.2%)
B. stercoris (HM626177)	1256/1317 (95.4%)
B. wexlerae (EF036467)	1391/1440 (96.6%)

[a] The sequences were those proposed as reference by LPSN.

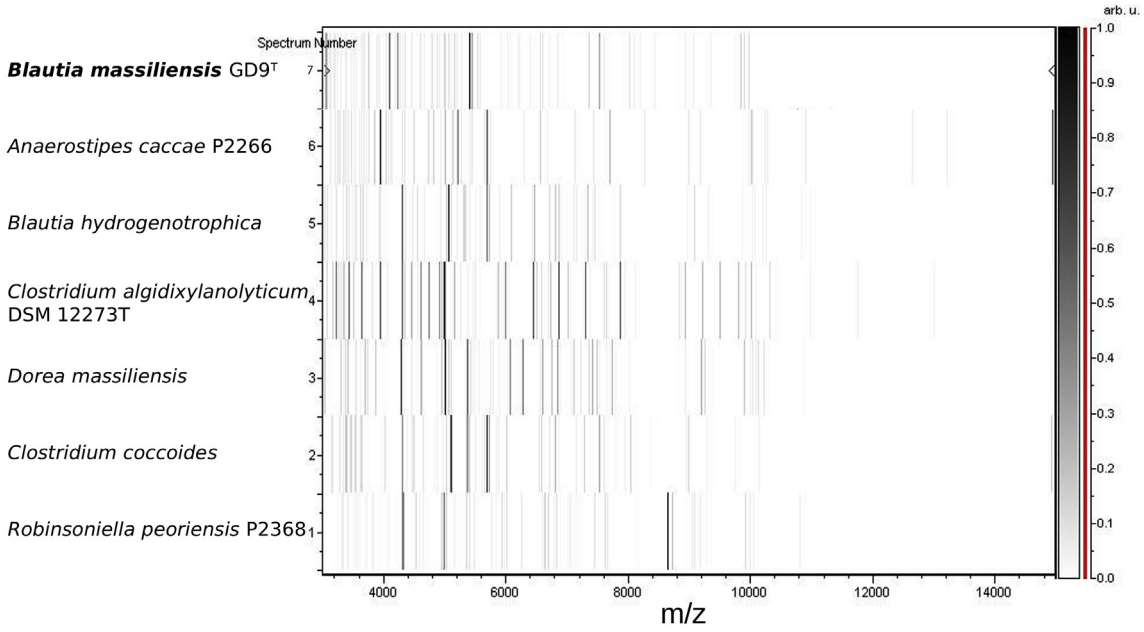

Fig. 2. Gel view comparing *B. massiliensis* strain GD9^T to other phylogenetically close species. The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel like look. The x-axis records the m/z value. The left y-axis displays the running spectrum number originating from subsequent spectra loading. The peak intensity is expressed by a gray scale scheme code. The color bar and the right y-axis indicate the relation between the color of the peak and its intensity, in arbitrary units. Displayed species are indicated on the left. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Classification and general features of *Blautia massiliensis* strain GD9^T.

Property	Term
Current classification	Domain: *Bacteria* Phylum: *Firmicutes* Class: *Clostridia* Order: *Clostridiales* Family: ***Lachnospiraceae*** Genus: *Blautia* Species: *Blautia massiliensis* Type strain: GD9^T
Gram stain	Negative
Cell shape	Rod
Motility	Non-motile
Sporulation	Non-sporulating
Temperature range	Mesophilic
Optimum temperature	37 °C

[18]). Regarding others species among the *Blautia* genus, fatty acid profiles are comparable with C16:0 and C14:0 as the most prevalent fatty acids, except for *B. schinkii*, *B. glucerasea* and *B. coccoides* showing higher relative abundance of C12:0 and C16:1n7. Moreover, Liu et al. also described C16:0 and C14:0 as the most abundant fatty acids for *Blautia* species [9].

Antibiotic susceptibility was observed for the class of beta-lactam compounds that included amoxicillin, amoxicillin with clavulanic acid, ticarcillin with clavulanic acid, tazocillin, ceftriaxone, cefepime and imipenem. For aminoglycoside antibiotics, resistance was observed for gentamicin, tobramycin and amikacin. The GD9^T strain was also susceptible to tigecyclin, rifampicin, nitrofurantoin, but resistant to sulfamethoxazole, aztreonam and ciprofloxacin.

3.2. Genomic characterization and comparison

The genome is 3,717,339 bp long with a 43.99% GC content (Table 5). It is composed of five scaffolds (composed of nine contigs). Of the 3,427 predicted genes, 3,346 were protein-coding genes, and 81 were RNAs (eight genes are 5S rRNA, six genes are 16S rRNA, four genes are 23S rRNA, 63 genes are TRNA genes). A total of 2,570 genes (76.81%) were assigned as putative function (by

Table 3
Comparison of *B. massiliensis* strain GD9^T with close species *Blautia luti* DSM 14534^T, *Blautia coccoides* DSM 935^T, *Blautia hensenii* DSM 20583^T, *Blautia schinkii* DSM 10518^T, *Blautia obeum* DSM 25238^T, na: non-available data.

Properties	*B. massiliensis*	*B. luti*	*B. coccoides*	*B. hensenii*	*B. schinkii*	*B. obeum*
Cell diameter (μm)	0.5	0.7–0.9	1–1.5	1–1.5	1–1.5	1–1.5
Oxygen requirement	anaerobic	anaerobic	anaerobic	anaerobic	anaerobic	anaerobic
Gram stain	–	+	+	+	+	+
Motility	–	–	–	–	–	–
Endospore formation	–	–	–	–	–	–
Indole	–	–	–	–	–	–
Production of						
Catalase	–	–	–	–	–	–
Oxidase	–	NA	–	–	–	–
Urease	–	–	NA	NA	NA	NA
Gelatinase	–	–	–	–	–	–
α-Galactosidase	+	+	+	–	+	+
β-Galactosidase	+	+	+	+	–	–
α -Glucosidase	+	+	+	–	+	+
β -Glucosidase	+	+	+	–	+	–
Acid from						
L-Arabinose	+	+	+	–	+	+
Mannose	+	+	+	–	+	+
Mannitol	+	–	+	–	NA	NA
Cellobiose	+	+	+	–	+	+
Raffinose	+	+	+	+	+	+
Saccharose	+	+	+	–	+	+
D-maltose	+	+	+	+	+	+
D-lactose	+	+	+	+	NA	NA
Habitat	human gut	human gut	human gut	human gut	human gut	human gut

Table 4
Cellular fatty acid profiles of *Blautia massiliensis* strain GD9^T compared with other *Blautia* type strains of closely related species.

Fatty acid	NAME	1	2	3	4	5	6	7	8
C5:0 anteiso	2-methyl-Butanoic acid	ND	ND	ND	ND	ND	ND	ND	ND
C12:0	Dodecanoic acid	ND	ND	2.3	3.4	2.6	26.1	2.1	50.2
C14:0	Tetradecanoic acid	19.8	39.2	47.3	35.9	19.2	31.6	29.3	15.7
C15:0	Pentadecanoic acid	4.3	1.3	0.7	1.1	0.6	0.6	0.9	ND
C15:0 iso	13-methyl-tetradecanoic	ND	ND	ND	ND	ND	ND	ND	ND
C16:0	Hexadecanoic acid	53.5	44.3	37.4	34.1	41.4	27.0	32.3	17.6
C16:1n5	11-Hexadecenoic acid	ND	0.6	0.6	0.6	ND	0.5	ND	ND
C16:1n7	9-Hexadecenoic acid	2.6	2.4	3.0	15.4	3.2	3.7	6.8	ND
C16:1n9	7-Hexadecenoic acid	ND	ND	0.4	ND	1.5	1.0	0.4	ND
C17:0	Heptadecanoic acid	1.3	ND	ND	ND	ND	ND	ND	ND
C17:1n7	10-Heptadecenoic acid	0.7	ND	ND	ND	ND	ND	ND	ND
C18:0	Octadecanoic acid	4.7	2.3	3.5	1.8	10.4	2.1	5.4	1.4
C18:1n7	11-Octadecenoic acid	3.3	ND	ND	ND	ND	ND	ND	ND
C18:1n9	9-Octadecenoic acid	6.1	8.8	18.3	7.7	20.8	7.0	22.9	13.9
C18:2n6	9,12-Octadecadienoic acid	1.7	0.7	0.8	ND	0.3	0.5	ND	1.3
C20:4n6	5,8,11,14-Eicosatetraenoic acid	2.1	ND	ND	ND	ND	ND	ND	ND

Strains: 1, *B. massiliensis* GD9^T; 2, *B. faecis* M25^T; 3, *B. stercoris* GAM6-1^T; 4, *B. coccoides* DSM 935^T; 5, *B. hansenii* DSM 20583^T; 6, *B. schinkii* DSM 10518^T; 7, *B. obeum* DSM 25238^T; 8, *B. glucerasea* DSM 22028^T. Strains 2 to 8 were described by Park, 2013 [18]. Values represent the percentage of total identified fatty acid methyl esters only (aldehydes, dimethyl acetals and unidentified "summed features" described previously were not included). ND, Not detected.

Table 5
Nucleotide content and gene count levels of the genome.

Attribute	Genome (total)	
	Value	% of total[a]
Size (bp)	3,717,339	100
G + C content (bp)	1,634,069	43.98
Coding region (bp)	3,269,367	87.94
Total genes	3427	100
RNA genes	81	2.36
Protein-coding genes	3346	97.63
Proteins with function prediction	2570	76.8
Genes assigned to COGs	1978	59.11
Genes with peptide signals	361	10.78
Genes with transmembrane helices	742	22.17
Genes with Pfam domains	3091	90

[a] 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.

cogs or by NR blast). 101 genes were identified as ORFans (3.02%). The remaining genes were annotated as hypothetical proteins (581 genes, 17.36%, Supplementary Fig. 4). The genome sequence has been deposited in EMBL-EBI under accession number PRJEB11857.

The draft genome sequence of *B. massiliensis* is smaller than that of *Blautia schinkii* and *Blautia obeum* (3.72, 6.68 and 3.76 MB respectively), but larger than that of *Blautia wexlerae* DSM 19850^T, *Blautia hydrogenotrophica*, and *Blautia hansenii* (3.58, 3.35 and 3.05 MB respectively). The G + C content of *B. massiliensis* (44.0%) is smaller than that of *B. schinkii* and *B. hydrogenotrophica* (44, 46 and 45 respectively), but larger than that of *B. wexlerae*, *B. hansenii* and *B. obeum* (42, 39 and 42% respectively). The gene content of *B. massiliensis* is smaller than that of *B. schinkii* (3346 and 5851 respectively), but larger than that of *B. wexlerae* (3,297), *B. hydrogenotrophica* (3,087), *B. hansenii* (3,171) and *B. obeum* (3,155).

Table 6
Pairwise comparison of *B. massiliensis* GD9^T with other species using the GGDC web-service, formula 2 (upper right)[a].

	B. massiliensis GD9^T	*B. obeum* DSM 25238^T	*B. schinkii* DSM 10518^T	*B. hansenii* DSM 20583^T	*B. hydrogenotrophica* DSM 10507^T	*B. wexlerae* DSM 19850^T
B. massiliensis GD9^T	100% ± 00	31.8 ± 2.5	20.8% ± 2.4	36.9% ± 2.6	21.5% ± 2.4	31.1% ± 2.5
B. obeum DSM 25238^T		100% ± 00	20.5% ± 2.4	27.1% ± 2.5	20.8%± ± 2.4	27.9% ± 2.5
B. schinkii DSM 10518^T			100% ± 00	23.9% ± 2.5	20.4% ± 2.4	23.2% ± 2.5
B. hansenii DSM 20583^T				100% ± 00	22.4% ± 2.5	38.0% ± 2.5
B. hydrogenotrophica DSM 10507^T					100% ± 00	19.9% ± 2.4
B. wexlerae DSM 19850^T						100% ± 00

[a] dDDH values are DDH estimates based on identities/HSP length. The confidence intervals indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test datasets (which are always limited in size).

Table 7
Numbers of orthologous protein shared between genomes (upper right), AGIOS values (lower left) and numbers of proteins per genome (bold numbers).

	Blautia obeum DSM 25238^T	*Blautia schinkii* DSM 10518^T	*Blautia hansenii* DSM 20583^T	*Blautia hydrogenotrophica* DSM 10507^T	*Blautia wexlerae* DSM 19850^T	*Blautia massiliensis* GD9^T
Blautia obeum DSM 25238^T	**3155**	1389	1086	1111	1393	1450
Blautia schinkii DSM 10518^T	65.32	**5851**	1313	1371	1556	1586
Blautia hansenii DSM 20583^T	63.63	66.87	**3171**	1128	1224	1225
Blautia hydrogenotrophica DSM 10507^T	63.25	68.34	67.33	**3087**	1227	1232
Blautia wexlerae DSM 19850^T	67.75	72.33	68.37	68.41	**3297**	1600
Blautia massiliensis GD9^T	68.08	71.80	68.44	68.18	74.75	**3346**

DNA-DNA hybridization (DDH) is currently considered as the "gold standard" criterion for species delineation of prokaryotes. However, this tool suffers from limitations, notably the 70% cutoff value that is not applicable to all prokaryotic genera [34], and the need of special facilities that are available in a limited number of laboratories. Moreover, determining DDH is a labor-intensive method that lacks reproducibility and cannot be used to establish a comparative reference database incrementally [35,36]. Therefore, in order to evaluate the genomic similarity among studied strains, we used two parameters: digital DDH (dDDH) that exhibits a high correlation with DDH [32,37] and AGIOS [29] that was designed to be independent from DDH. When considering only *Blautia* species with available genome and standing in nomenclature, dDDH values ranged from 19.9 ± 2.4% between *B. hydrogenotrophica* and *B. wexlerae* to 38.0 ± 2.5% between *B. hansenii* and *B. wexlerae*. When comparing strain GD9^T, dDDH values ranged from 20.8% ± 2.4 with *B. schinkii* strain DSM10518^T to 36.9% ± 2.6 with *B. hansenii* strain DSM 20583^T with a probability of error of 1.19%

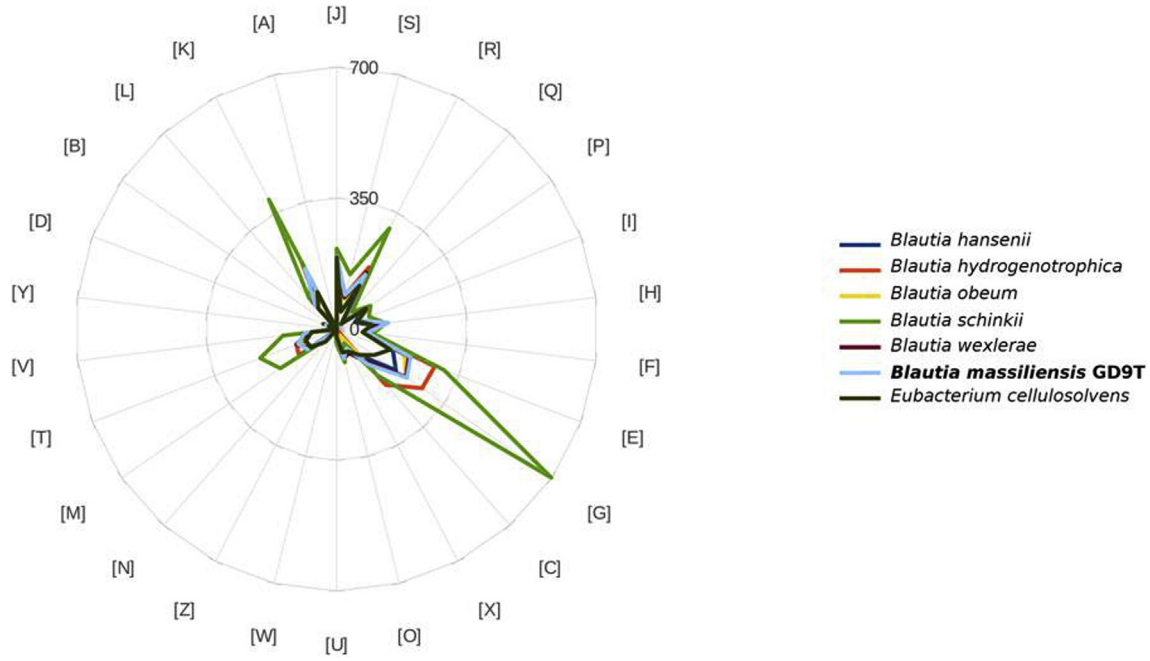

Fig. 3. Distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins. The functional classes of predicted genes was assessed using the NCBI COGs database updated in 2014 [43] and the Blastp [44] tool (Evalue 1e-03, coverage 0,7 and identity percent 30%). The tRNA genes were predicted with the tRNAscan-SE tool [23] and the RNA genes were predicted with the Rnammer [24] tool. The figure was made using the DNAPlotter [45] tool.

Table 8
Number of genes associated with the 25 general COG functional categories.

Code	Value	% of total[a]	Description
[J]	195	5.83	Translation
[A]	0	0	RNA processing and modification
[K]	182	5.44	Transcription
[L]	89	2.66	Replication, recombination and repair
[B]	0	0	Chromatin structure and dynamics
[D]	32	0.96	Cell cycle control, mitosis and meiosis
[Y]	0	0	Nuclear structure
[V]	86	2.57	Defense mechanisms
[T]	111	3.32	Signal transduction mechanisms
[M]	95	2.84	Cellwall/membrane biogenesis
[N]	13	0.39	Cellmotility
[Z]	0	0	Cytoskeleton
[W]	4	0.12	Extracellular structures
[U]	25	0.75	Intracellular trafficking and secretion
[O]	81	2.42	Post translational modification, protein turnover, chaperones
[X]	58	1.73	Mobilome: prophages, transposons
[C]	125	3.74	Energy production and conversion
[G]	229	6.84	Carbohydrate transport and metabolism
[E]	212	6.34	Amino acid transport and metabolism
[F]	82	2.45	Nucleotide transport and metabolism
[H]	139	4.15	Coenzyme transport and metabolism
[I]	65	1.94	Lipid transport and metabolism
[P]	94	2.81	Inorganic ion transport and metabolism
[Q]	29	0.87	Secondary metabolites biosynthesis, transport and catabolism
[R]	165	4.93	General function prediction only
[S]	94	2.81	Function unknown
–	1368	40.89	Not in COGs

[a] The total is based on the total number of protein coding genes in the annotated genome.

according to GGDC (Table 6). Among compared *Blautia* species, AGIOS values ranged from 63.25 between *B. hydrogenotrophica* and *B. obeum* to 72.33% between *B. schinkii* and *B. wexlerae* among compared species except *B. massiliensis* and from 68.08 with *B. obeum* to 74.75% with *B. wexlerae* (Table 7). Fig. 3 demonstrates that strain GD9^T exhibits a similar distribution of genes into COG categories when compared to other *Blautia* species. *B. massiliensis* also shared 1450; 1586; 1225; 1232 and 1600 orthologous proteins with *B. obeum*, *B. schinkii* DSM 10518, *B. hansenii* DSM 20583, *B. hydrogenotrophica* DSM 10507^T and *B. wexlerae* DSM 19850 respectively (Table 8).

3.3. Importance of *B. massiliensis* in human gut and possible role in human health

Blautia massiliensis is likely to be an important species for human health because it was isolated and detected frequently and abundantly in human gut culturomes and metagenomes. *B. massiliensis* was isolated from the fecal samples from three other healthy individuals by three other researchers (TP, SD, ST) in our culturomics team. These strains were identified by MALDI-TOF MS and confirmed by sequencing of the 16S rRNA (unpublished data). *B. massiliensis* corresponded to 4.25% of all reads (200,239/4,716,269 reads) in an ongoing study on gut microbiota and bariatric surgery in our lab, ranking 6th in relative abundance. In addition, we investigated the presence of 16S rRNA from *B. massiliensis* in the high-throughput DNA and RNA sequence read archive (SRA) using an open resource online [38]. We found sequences with a similarity greater than 99% with *B. massiliensis* in several gut metagenomes (human, mouse and primate), skin and vaginal metagenomes; as well as from environmental samples (wastewater, bioreactor, coral). Sequences corresponding to *B. massiliensis* were found in 11% (9752/88,579) of all metagenomes and 41% (6839/16,667) of human gut metagenomes in this database. Among the 6839 positive samples from human gut, the mean relative abundance was 0.0088 but sequences corresponding to strain GD9 accounted for as much as 20% of all reads in 4 human gut metagenomes (SRR2143746, 27%; SRR2658700, 27%; SRR515348, 24%; SRR578392, 22%).

The specific role of *Blautia* species in human health remains to be determined. Jenq et al. found a lower mortality due to a graft-versus-host disease after allogenic blood/marrow transplantation among patients with high abundance of *Blautia* sequences obtained by pyrosequencing [39]. Chen et al. have shown that colorectal cancer was associated with lower *Blautia* in their digestive tract [40]. Touyama et al. calculated that *B. wexlerae* and *B. luti* were found at the concentration of 5.10^9 bacteria per gram of stool [41], suggesting that *Blautia* are important members of the healthy human mature anaerobic gut microbiota.

4. Conclusion

Anaerobic culturomics conditions applied to fresh human stool have permitted the culture of a new species belonging to *Blautia*. This taxogenomic study confirmed the new species *Blautia massiliensis* sp. nov., which appeared close to *B. hensenii* strain DSM 20583 on genome-based analysis. These species are both anaerobes inhabitants of the human gut, indole and catalase negative, sharing the same repartition of cellular fatty acids C14:0 and C16:0. However, *B. massiliensis* has lower C18:0 than *B. hensenii*, but higher G + C content. *B. massiliensis* is the only species with Gram-negative staining among the *Blautia* genus. The closest species on the 16S rRNA sequence identity was *B. luti*, for which no genome was available. However, beyond the 16S rRNA divergence (2.2%), *B. massiliensis* GD9^T differed from *B. luti* by several phenotypic characteristics including a lower diameter (0.5 versus 0.7–0.9 μm), a negative Gram-staining (verified several times with positive controls), an absence of bacterial chains (chains up to 10 cells for *B. luti* [42]), and the use of mannitol, melezitose, rhamnose and salicin (not found for *B. luti* [42]). This led us to propose *B. massiliensis* as a new species.

4.1. Taxonomic and nomenclatural proposals

4.1.1. Emended description of the genus Blautia

Blautia (Blauti.a. N.L. fem. n. Blautia in honour of Michael Blaut, a German microbiologist, in recognition of his many contributions to human gastrointestinal microbiology). Gram-positive or Gram-negative staining, non-motile. Coccoid or ovalshaped, pointed ends are often observed. Spores are not normally observed, but may be produced by some strains. Chemo-organotrophic and obligately anaerobic having a fermentative type of catabolism. Some species use H2/CO2 as major energy sources. The major end products of glucose metabolism are acetate, ethanol, hydrogen, lactate and succinate. The G + C content of the DNA is 37–47 mol%. Isolated from animal and human faeces. The type species of the genus is *Blautia coccoides* (Kaneuchi, Benno & Mitsuoka, 1976).

4.1.2. Description of Blautia massiliensis, sp. nov

Blautia massiliensis (ma.si.li.en'sis. L. fem. adj. *Massiliensis*, from the Latin *Massilia,* the city where the bacteria was isolated, Marseille) presented white smooth colonies of 1 mm diameter. The bacteria appear to be 1 μm long and with Gram-negative staining. Their metabolism is oxygen intolerant, and growth is optimal at 37 °C. *B. massiliensis* is non spore-forming and non-motile. Oxidase and catalase are negative. Positive reactions observed using API 20A are D-Glucose, D-Mannitol, D-Lactose, D-Sucrose, D-Maltose, Salicine, D-Xylose, L-Arabinose, Esculine, Glycerol, D-Cellobiose, D-Mannose, D-Melezitose, D-Raffinose, D-Sorbitol, L-Rhamnose, and D-Trehalose. Using API 50CH strips, positive reactions are found with Erythitol, D-Arabinose, D-Ribose, L-Xylose, L-Rhamnose, Dulcitol, *N*-Acetylglucosamine, Amygdaline, Arbutine, Inuline, Amidon, Glycogene, Gentiobiose, D-Lyxose, D-Tagatose, D-Fucose, L-Fucose, Potassium Gluconate and Potassium 5-Cetogluconate. Antibiotic resistance for aminoglycosides, sulfamethoxazole, aztreonam, and ciprofloxacin are observed. The major fatty acids are C16:0 and C14:0. This strain exhibits a G + C content of 44.0%. Its 16S rRNA sequence is deposited in EMBL-EBI under the accession number LN890282, and the whole genome shotgun sequence is deposited in EMBL-EBI under accession number PRJEB11857. The type species is *Blautia massiliensis* strain GD9^T (= CSUR P2132, = DSM 101187) isolated from the stool of a healthy person from Marseille, France.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process. This study was funded by the “Fondation Méditerranée Infection”. We also thank Karolina Griffiths and Magdalen Lardière for English reviewing, Nicholas Armstrong and Magali Richez who performed the fatty acids analyses.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.anaerobe.2016.12.001.

References

[1] P.J. Turnbaugh, R.E. Ley, M. Hamady, C.M. Fraser-Liggett, R. Knight, J.I. Gordon, The human microbiome project, Nature 449 (2007) 804–810.
[2] P. Hugon, J.-C. Dufour, P. Colson, P.-E. Fournier, K. Sallah, D. Raoult, A comprehensive repertoire of prokaryotic species identified in human beings, Lancet Infect. Dis. 15 (2015) 1211–1219.
[3] R.H. Siggers, J. Siggers, M. Boye, T. Thymann, L. Mølbak, T. Leser, B.B. Jensen, P.T. Sangild, Early administration of probiotics alters bacterial colonization and limits diet-induced gut dysfunction and severity of necrotizing enterocolitis in preterm pigs, J. Nutr. 138 (2008) 1437–1444.
[4] G. De Hertogh, J. Aerssens, R. de Hoogt, P. Peeters, P. Verhasselt, P. Van Eyken, N. Ectors, S. Vermeire, P. Rutgeerts, B. Coulie, K. Geboes, Validation of 16S rDNA sequencing in microdissected bowel biopsies from Crohn's disease patients to assess bacterial flora diversity, J. Pathol. 209 (2006) 532–539.
[5] C. Manichanh, L. Rigottier-Gois, E. Bonnaud, K. Gloux, E. Pelletier, L. Frangeul, R. Nalin, C. Jarrin, P. Chardon, P. Marteau, J. Roca, J. Dore, Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach, Gut 55 (2006) 205–211.
[6] P.D. Scanlan, F. Shanahan, C. O'Mahony, J.R. Marchesi, Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's disease, J. Clin. Microbiol. 44 (2006) 3980–3988.
[7] J.-C. Lagier, P. Hugon, S. Khelaifia, P.-E. Fournier, B. La Scola, D. Raoult, The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota, Clin. Microbiol. Rev. 28 (2015) 237–264.
[8] J.-C. Lagier, S. Khelaifia, M.T. Alou, S. Ndongo, N. Dione, P. Hugon, A. Caputo, F. Cadoret, S.I. Traore, E.H. Seck, G. Dubourg, G. Durand, G. Mourembou, E. Guilhot, A. Togo, S. Bellali, D. Bachar, N. Cassir, F. Bittar, J. Delerce, M. Mailhe, D. Ricaboni, M. Bilen, N.P.M. Dangui Nieko, N.M. Dia Badiane, C. Valles, D. Mouelhi, K. Diop, M. Million, D. Musso, J. Abrahão, E.I. Azhar, F. Bibi, M. Yasir, A. Diallo, C. Sokhna, F. Djossou, V. Vitton, C. Robert, J.M. Rolain, B. La Scola, P.-E. Fournier, A. Levasseur, D. Raoult, Culture of previously uncultured members of the human gut microbiota by culturomics, Nat. Microbiol. 1 (2016) 16203, http://dx.doi.org/10.1038/nmicrobiol.2016.203.
[9] C. Liu, S.M. Finegold, Y. Song, P.A. Lawson, Reclassification of *Clostridium coccoides, Ruminococcus hansenii, Ruminococcus hydrogenotrophicus, Ruminococcus luti, Ruminococcus productus* and *Ruminococcus schinkii* as *Blautia coccoides* gen. nov., comb. nov., *Blautia hansenii* comb. nov., *Blautia hydrogenotrophica* comb. nov., *Blautia luti* comb. nov., *Blautia producta* comb. nov., *Blautia schinkii* comb. nov. and description of *Blautia wexlerae* sp. nov., isolated from human faeces, Int. J. Syst. Evol. Microbiol. 58 (2008) 1896–1902.
[10] M. Million, A. Diallo, D. Raoult, Gut microbiota and malnutrition, Microb. Pathog. (2016 Feb 4), http://dx.doi.org/10.1016/j.micpath.2016.02.003 pii: S0882-4010(15)30212-6. [Epub ahead of print].
[11] J.-C. Lagier, K. Elkarkouri, R. Rivet, C. Couderc, D. Raoult, P.-E. Fournier, Non contiguous-finished genome sequence and description of *Senegalemassilia anaerobia* gen. nov., sp. nov, Stand. Genomic. Sci. 7 (2013) 343–356.
[12] G. Mourembou, J. Rathored, J.B. Lekana-Douki, A. Ndjoyi-Mbiguino, F. Fenollar, C. Michelle, P.-E. Fournier, D. Raoult, J.-C. Lagier, Noncontiguous finished genome sequence and description of *Kallipyga gabonensis* sp. nov, New Microbes. New Infect. 9 (2016) 15–23.
[13] E. Stackebrandt, J. Ebers, Taxonomic parameters revisited: tarnished gold standards, Microbiol. Today 33 (2006) 152–155.
[14] M. Kim, H.-S. Oh, S.-C. Park, J. Chun, Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes, Int. J. Syst. Evol. Microbiol. 64 (2014) 346–351.
[15] R. Aghnatios, C. Cayrou, M. Garibal, C. Robert, S. Azza, D. Raoult, M. Drancourt, Draft genome of *Gemmata massiliana* sp. nov, a water-borne Planctomycetes species exhibiting two variants, Stand. Genomic. Sci. 10 (2015).
[16] Myron Sasser, Bacterial Identification by Gas Chromatographic Analysis of Fatty Acids Methyl Esters (GC-FAME), MIDI, 2006.
[17] N. Dione, et al., Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes. New Infect 10 (2016) 66–76.
[18] S.-K. Park, M.-S. Kim, J.-W. Bae, *Blautia faecis* sp. nov., isolated from human faeces, Int. J. Syst. Evol. Microbiol. 63 (2013) 599–603.
[19] European Committee on Antimicrobial Susceptibility Testing, Comité de l'Antibiogramme de la Société Française de Microbiologie, CASFM-EUCAST recommendations (n.d.), http://www.sfm-microbiologie.org/UserFiles/files/casfm/CASFM_EUCAST_V1_2015.pdf, 2015. Last access 2016/11/28.
[20] B. Sengüven, E. Baris, T. Oygur, M. Berktas, Comparison of methods for the extraction of DNA from formalin-fixed, paraffin-embedded archival tissues, Int. J. Med. Sci. 11 (2014) 494–499.
[21] J.-C. Lagier, F. Bibi, D. Ramasamy, E.I. Azhar, C. Robert, M. Yasir, A.A. Jiman-Fatani, K.Z. Alshali, P.-E. Fournier, D. Raoult, Non contiguous-finished genome sequence and description of *Clostridium jeddahense* sp. nov, Stand. Genomic. Sci. 9 (2014) 1003–1019.
[22] Prodigual, (n.d.). http://prodigal.ornl.gov/. Last access 2016/11/28.
[23] T.M. Lowe, S.R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, Nucleic Acids Res. 25 (1997) 955–964.
[24] K. Lagesen, P. Hallin, E.A. Rødland, H.-H. Staerfeldt, T. Rognes, D.W. Ussery, RNAmmer: consistent and rapid annotation of ribosomal RNA genes, Nucleic Acids Res. 35 (2007) 3100–3108.
[25] J.D. Bendtsen, H. Nielsen, G. von Heijne, S. Brunak, Improved prediction of signal peptides: SignalP 3.0, J. Mol. Biol. 340 (2004) 783–795.
[26] J.-C. Lagier, F. Armougom, M. Million, P. Hugon, I. Pagnier, C. Robert, F. Bittar, G. Fournous, G. Gimenez, M. Maraninchi, J.-F. Trape, E.V. Koonin, B. La Scola, D. Raoult, Microbial culturomics: paradigm shift in the human gut microbiome study, Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. 18 (2012) 1185–1193.
[27] A.H. Togo, S. Khelaifia, J.-C. Lagier, A. Caputo, C. Robert, P.-E. Fournier,

M. Maraninchi, R. Valero, D. Raoult, M. Million, Noncontiguous finished genome sequence and description of *Paenibacillus ihumii* sp. nov. strain AT5, New Microbes. New Infect. 10 (2016) 142–150.
[28] M. Lechner, S. Findeiss, L. Steiner, M. Marz, P.F. Stadler, S.J. Prohaska, Proteinortho: detection of (co-)orthologs in large-scale analysis, BMC Bioinforma. 12 (2011) 124.
[29] D. Ramasamy, A.K. Mishra, J.-C. Lagier, R. Padhmanabhan, M. Rossi, E. Sentausa, D. Raoult, P.-E. Fournier, A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species, Int. J. Syst. Evol. Microbiol. 64 (2014) 384–391.
[30] P. Gouret, J. Paganini, J. Dainat, D. Louati, E. Darbo, P. Pontarotti, P. Pontarotti, Integration of Evolutionary Biology Concepts for Functional Annotation and Automation of Complex Research in Evolution: the Multi-agent Software System DAGOBAH, in: Evol. Biol. – Concepts Biodivers, Macroevolution Genome Evol., Springer, Berlin Heidelberg, 2011, pp. 71–87.
[31] A.F. Auch, H.-P. Klenk, M. Göker, Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs, Stand. Genomic. Sci. 2 (2010) 142–148.
[32] J.P. Meier-Kolthoff, A.F. Auch, H.-P. Klenk, M. Göker, Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC Bioinforma. 14 (2013) 60.
[33] P. Yarza, M. Richter, J. Peplies, J. Euzeby, R. Amann, K.-H. Schleifer, W. Ludwig, F.O. Glöckner, R. Rosselló-Móra, The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains, Syst. Appl. Microbiol. 31 (2008) 241–250.
[34] P.-E. Fournier, D. Raoult, Current knowledge on phylogeny and taxonomy of *Rickettsia* spp. Ann. N. Y. Acad. Sci. 1166 (2009) 1–11.
[35] H.-J. Busse, B.J. Tindall, W. Ludwig, R. Rosselló-Móra, P. Kämpfer, Notes on the characterization of prokaryote strains for taxonomic purposes, Int. J. Syst. Evol. Microbiol. 60 (2010) 249–266.
[36] D. Ramasamy, A.K. Mishra, J.-C. Lagier, R. Padhmanabhan, M. Rossi, E. Sentausa, D. Raoult, P.-E. Fournier, A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species, Int. J. Syst. Evol. Microbiol. 64 (2014) 384–391.
[37] A.F. Auch, M. von Jan, H.-P. Klenk, M. Göker, Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison, Stand. Genomic. Sci. 2 (2010) 117–134.
[38] I. Lagkouvardos, D. Joseph, M. Kapfhammer, S. Giritli, M. Horn, D. Haller, T. Clavel, IMNGS: a comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies, Sci. Rep. 6 (2016) 33721.
[39] R.R. Jenq, Y. Taur, S.M. Devlin, D.M. Ponce, J.D. Goldberg, K.F. Ahr, E.R. Littmann, L. Ling, A.C. Gobourne, L.C. Miller, M.D. Docampo, J.U. Peled, N. Arpaia, J.R. Cross, T.K. Peets, M.A. Lumish, Y. Shono, J.A. Dudakov, H. Poeck, A.M. Hanash, J.N. Barker, M.-A. Perales, S.A. Giralt, E.G. Pamer, M.R.M. van den Brink, Intestinal *Blautia* is associated with reduced death from graft-versus-host disease, biol. Blood marrow transplant, J. Am. Soc. Blood Marrow Transpl. 21 (2015) 1373–1383.
[40] W. Chen, F. Liu, Z. Ling, X. Tong, C. Xiang, Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer, PloS One 7 (2012) e39743.
[41] M. Touyama, J.S. Jin, R. Kibe, H. Hayashi, Y. Benno, Quantification of *Blautia wexlerae* and *Blautia luti* in human faeces by real-time PCR using specific primers, Benef. Microbes. 6 (2015) 583–590.
[42] R. Simmering, D. Taras, A. Schwiertz, G, Le Blay, B. Gruhl, P.A. Lawson, M.D. Collins, M. Blaut, *Ruminococcus luti* sp. nov., isolated from a human faecal sample, Syst. Appl. Microbiol. 25 (2002) 189–193.
[43] M.Y. Galperin, K.S. Makarova, Y.I. Wolf, E.V. Koonin, Expanded microbial genome coverage and improved protein family annotation in the COG database, Nucleic Acids Res. 43 (2015) D261–D269.
[44] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T.L. Madden, BLAST+: architecture and applications, BMC Bioinforma. 10 (2009) 421.
[45] T. Carver, N. Thomson, A. Bleasby, M. Berriman, J. Parkhill, DNAPlotter: circular and linear interactive genome visualization, Bioinforma. Oxf, Engl 25 (2009) 119–120.