Non contiguous-finished genome sequence and description of *Alistipes obesi* sp. nov

Perrine Hugon[1]*, Dhamodharan Ramasamy[1]*, Jean-Christophe Lagier[1], Romain Rivet[1], Carine Couderc[1], Didier Raoult[1] and Pierre-Edouard Fournier[1]¶

[1]Aix-Marseille Université, Faculté de médecine, Marseille, France

***Corresponding author**: Pierre-Edouard Fournier (pierre-edouard.fournier@univmed.fr)

Keywords: *Alistipes obesi*, genome, culturomics, taxono-genomics

Alistipes obesi sp. nov. strain ph8^T is the type strain of *A. obesi*, a new species within the genus *Alistipes*. This strain, whose genome is described here, was isolated from the fecal flora of a 26-year-old woman suffering from morbid obesity. *A. obesi* is an obligately anaerobic rod. Here we describe the features of this organism, together with the complete genome sequence and annotation. The 3,162,233 bp long genome (1 chromosome but no plasmid) contains 2,623 protein-coding and 49 RNA genes, including three rRNA genes.

Introduction

Alistipes obesi strain ph8^T (CSUR= P186, DSMZ= 25724) is the type strain of *A. obesi* sp. nov. This bacterium is a Gram-negative, anaerobic, indole-negative bacillus and was isolated from the stool of a French patient suffering from morbid obesity as part of a culturomics study aiming at cultivating individually all species within human feces [1].

We recently proposed that genomic and proteomic data, which do not suffer from the lack of reproducibility and inter-laboratory comparability that the "gold standard" DNA-DNA hybridization and G+C content determination does [2], may be included in the official description of new bacterial species [3-14].

The genus *Alistipes* (Rautio *et al*. 2003) [15] is currently comprised of five species, including *A. finegoldii* (Rautio *et al.* 2003) [15], *A. indistinctus* (Nagai *et al.* 2010) [16], *A. onderdonkii* (Song *et al.* 2006) [17], *A. putredinis* (Rautio *et al.* 2003) [15], and *A. shahii* (Song *et al.* 2006) [17]. In addition, we recently described two new species, *A. senegalensis* (Mishra *et al.* 2012) [6] and *A. timonensis* (Lagier *et al.* 2012) [7] that were isolated from the digestive microbiota from an asymptomatic Senegalese patient [1]. Members of the genus *Alistipes* are strictly anaerobic Gram-negative rods that are closely related to the *Bacteroides fragilis* group, with which they share the characteristic of bile-resistance and indole-positivity. Most *Alistipes* species have been isolated from human specimens, including the normal intestinal flora [17] and in cases of bacteremia, appendicitis, perirectal and brain abscess [18-20]. A 16S rRNA phylogenetic analysis revealed that *A. obesi* was closely related to *A. shahii*, *A. senegalensis* and *A. timonensis*. To the best of our knowledge, *A. obesi* sp. nov. is the first *Alistipes* species isolated from the digestive flora of an obese patient.

Here we present a summary classification and a set of features for *A. obesi* sp. nov. strain ph8^T together with the description of the complete genome sequencing and annotation. These characteristics support the circumscription of the species *A. obesi*.

Classification and features

A stool sample was collected from an obese, 26-year-old woman living in Marseille, France, who suffered from morbid obesity: BMI=48.2 (118.8 kg, 1.57 meter). At the time of stool sample collection she was not a drug user and was not on a diet. The patient gave an informed and signed consent, and the agreement of the local ethics committee of the IFR48 (Marseille, France) was obtained under agreement 11-017. The fecal specimen was preserved at -80°C after collection. Strain ph8 (Table 1) was isolated in 2011 by anaerobic cultivation at 37°C on 5% sheep blood-enriched Columbia agar (BioMerieux, Marcy l'Etoile, France), after 11 days of preincubation of the stool sample with addition of rumen fluid in an anaerobic blood culture bottle.

This strain exhibited a 93.5% 16S rRNA sequence similarity with *A. shahii* (Song *et al.* 2006) [17], the phylogenetically closest validated *Alistipes* species (Figure 1), and 94.26 and 93.38% with *A. senegalensis* (Mishra *et al.* 2012) [6] and *A. timonensis* (Lagier *et al.* 2012) [7], respectively. Among validly published *Alistipes* species [31], the percentage of 16S rRNA sequence similarity ranges from 90.5% between *A. indistinctus* (Nagai *et al.* 2010) [16] and *A. shahii* (Song *et al*. 2006) [17], to 96.8% between *A. finegoldii* (Rautio *et al.* 2003) [8] and *A. onderdonkii* (Song *et al.* 2006) [17]. As a consequence, and despite the fact that strain ph8 exhibited a 16SrRNA sequence similarity with the nearest validly published species lower than the 95.0% cutoff usually regarded as a threshold for the creation of new genus [32], we considered it as a new species within the genus *Alistipes*.

Table 1. Classification and general features of *Alistipes obesi* strain ph8^T according to the MIGS recommendations [21]

MIGS ID	Property	Term	Evidence code[a]
		Domain *Bacteria*	TAS [22]
		Phylum *Bacteroidetes*	TAS [23,24]
		Class *Bacteroidia*	TAS [23,25]
	Current classification	Order *Bacteroidales*	TAS [23,26]
		Family *Rikenellaceae*	TAS [23,27]
		Genus *Alistipes*	TAS [28,29]
		Species *Alistipes obesi*	IDA
		Type strain ph8^T	IDA
	Gram stain	Negative	IDA
	Cell shape	Rod	IDA
	Motility	Motile	IDA
	Sporulation	Non sporulating	IDA
	Temperature range	Mesophile	IDA
	Optimum temperature	37°C	IDA
MIGS-6.3	Salinity	Unknown	IDA
MIGS-22	Oxygen requirement	Anaerobic	IDA
	Carbon source	Unknown	NAS
	Energy source	Unknown	NAS
MIGS-6	Habitat	Human gut	IDA
MIGS-15	Biotic relationship	Free living	IDA
	Pathogenicity	Unknown	NAS
	Biosafety level	2	
MIGS-14	Isolation	Human feces	
MIGS-4	Geographic location	France	IDA
MIGS-5	Sample collection time	January 2011	IDA
MIGS-4.1	Latitude	43.296482	IDA
MIGS-4.1	Longitude	5.36978	IDA
MIGS-4.3	Depth	Surface	IDA
MIGS-4.4	Altitude	0 m above sea level	IDA

Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [30]. If the code is IDA, then the property was directly observed for a live isolate by one of the authors or an expert mentioned in the acknowledgements.

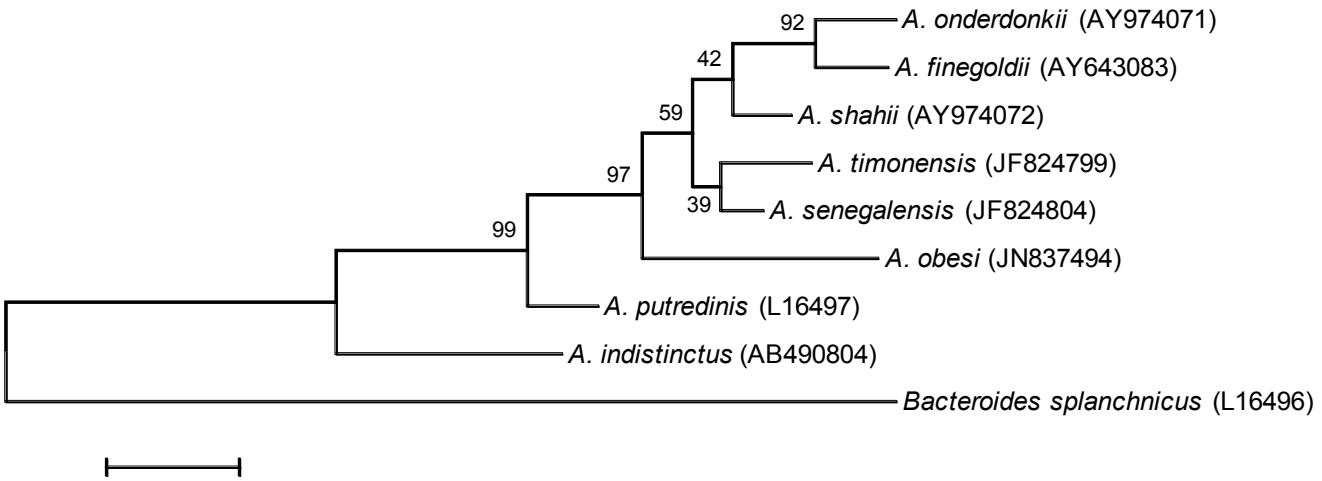

Figure 1. Phylogenetic tree highlighting the position of *Alistipes obesi* strain ph8^T relative to other type strains within the genus *Alistipes*. Genbank accession numbers are indicated in parentheses. Sequences were aligned using ClustalW, and phylogenetic inferences obtained using the maximum-likelihood method within the MEGA software. Numbers at the nodes are bootstrap values obtained by repeating the analysis 500 times to generate a majority consensus tree. *Bacteroides splanchnicus* was used as an outgroup. The scale bar represents a 2% nucleotide sequence divergence.

Different growth temperatures (25, 30, 37, 45°C) were tested; no growth occurred at 25°C or 30°C, growth occurred between 37°C and 45°C, and optimal growth was observed at 37°C. Colonies were 0.5 mm in diameter on 5% blood-enriched Columbia agar, are translucent and light grey. Growth of the strain was tested under anaerobic and microaerophilic conditions using GENbag anaer and GENbag microaer systems, respectively (BioMerieux), and in the presence of air, with or without 5% CO_2. Optimal growth was achieved anaerobically. No growth was observed under aerobic and microaerophilic conditions. Gram staining showed Gram-negative rods (Figure 2). A motility test was positive. Cells grown on agar and diameter ranged from 0.44 µm to 0.76 µm, with a mean diameter of 0.61 µm by electron microscopy (Figure 3). Comparison between seven *Alistipes* strains is presented in Table 2.

Strain ph8^T exhibited catalase activity but was not oxidase positive. Using the API RAPID ID 32A (BioMerieux), a positive reaction was obtained for α-galactosidase, β-galactosidase, N-acetyl-β-glucosaminidase, alkaline phosphatase, leucyl glycine arylamidase, and alanine arylamidase. All other tested reactions were negative, notably nitrate reduction, indole formation, urease, arginine dihydrolase, α- and β-glucosidase, 6-phospho-β-galactosidase, arginine arylamidase, proline arylamidase, phenylalanine arylamidase, leucine arylamidase, pyroglutamic acid arylamidase, tyrosine arylamidase, glycine arylamidase, histidine arylamidase, glutamyl glutamic acid arylamidase, serine arylamidase, and mannose and raffinose fermentation. Using the Api Zym system (BioMerieux), esterase, esterase lipase, acid phosphatase, Naphtol-AS-BI phosphohydrolase and α-galactosidase activities were positive. *A. obesi* is susceptible to imipenem, ciprofloxacin, metronidazole, nitrofurantoin and rifampicin, but resistant to penicillin G, amoxicillin, amoxicillin-clavulanic acid, erythromycin, vancomycin, gentamicin 15 and gentamycin 500, doxycycline, ceftriaxone and trimethoprim/sulfamethoxazole. By comparison with *A. senegalensis*, *A. obesi* differed in motility, α-galactosidase, β-galactosidase, indole production, β-glucuronidase, arginine arylamidase, glycine arylamidase, proline arylamidase and mannose fermentation [6]. By comparison with *A. timonensis*, *A. obesi* differed in motility, indole production, β-glucuronidase and N-acetyl-β-glucosaminidase [7]. By comparison with *A.putredinis*, *A.obesi* differed in motility, α-galactosidase, β-galactosidase N-acetyl-β-glucosaminidase and indole production [15]. By comparison with *A.finegoldii*, *A.obesi* differed in catalase, α-glucosidase and indole production [15]. Finally, A. obesidiffered in indole production, catalase, esterase, esterase lipase and alpha-glucosidase with *A. shahii* [17], and alpha-glucosidase, esterase, esterase lipase and acid phosphatase with *A. indistinctus* [16].

Table 2. Differential characteristics of *Alistipes* strains.†

Properties	*A.obesi*	*A.timonensis*	*A.senegalensis*	*A.putredinis*	*A.finegoldii*	*A.shahii*	*A.indistinctus*
Cell diameter (μm)	0.61	0.62	0.56	0.40	0.20	0.15	0.60
Oxygen requirement	anaerobic	anaerobic	anaerobic facultative	anaerobic	anaerobic	anaerobic	anaerobic
Gram stain	-	-	-	-	-	-	-
Salt requirement	na	na	na	na	na	na	na
Motility	+	-	-	-	na	na	na
Endospore formation	na	na	na	-	-	na	-
Production of							
Alkaline phosphatase	+	na	na	na	+	+	+
Acid phosphatase	+	na	na	na	+	+	w
Catalase	+	+	+	var	-	-	+
Oxidase	-	-	-	na	na	na	-
Nitrate reductase	-	na	na	-	-	-	-
Urease	-	na	na	na	na	na	-
αgalactosidase	+	+	w	-	+	+	+
βgalactosidase	+	+	w	-	+	+	+
βglucuronidase	-	+	w	na	na	-	+/-
αglucosidase	-	na	na	na	+	+	+
Esterase	+	na	na	na	+	w	w
Esterase lipase	+	na	na	na	+	w	w
Indole	-	w	w	+	+	+	-
N-acetyl-β glucosaminidase	+	w	na	-	+	+	+
Arginine arylamidase	-	na	w	na	na	na	-
glutamic acid decarboxylase	na	+	na	+	na	na	-
Leucyl glycine arylamidase	+	+	+	na	na	+	+
Alanine arylamidase	+	+	+	na	na	+	+
Proline arylamidase	-	na	+	na	na	na	-
Glycine arylamidase	-	na	w	na	na	na	-
Utilization of							
D-mannose	-	-	+	na	na	+	
Habitat	human gut	human gut	human gut	human	Human gut	human gut	human gut

var: variable
w: weak
na: data not available
+/-: depending on tests used

†*A. obesi* sp. nov strain ph8^T, *Alistipes timonensis* strain JC136 ^{T}, *Alistipes senegalensis* strain JC50 ^{T}, *Alistipes putredinis* strain ATCC 29800^T, *Alistipes finegoldii* strain AHN 2437^T, *Alistipes shahii* strain ATCC BAA-1179^T and *Alistipes indistinctus* strain YIT 12060^T.

Matrix-assisted laser-desorption/ionization time-of-flight (MALDI-TOF) MS protein analysis was carried out as previously described [33]. Briefly, a pipette tip was used to pick one isolated bacterial colony from a culture agar plate, and to spread it as a thin film on a MTP 384 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany). Four distinct deposits were done for strain ph8 from four isolated colonies. Each smear was overlaid with 2µL of matrix solution (saturated solution of alpha-cyano-4-hydroxycinnamic acid) in 50% acetonitrile, 2.5% tri-fluoracetic-acid, and allowed to dry for five minutes. Measurements were performed with a Microflex spectrometer (Bruker). Spectra were recorded in the positive linear mode for the mass range of 2,000 to 20,000 Da (parameter settings: ion source 1 (IS1), 20 kV; IS2, 18.5 kV; lens, 7 kV). A spectrum was obtained after 675 shots at a variable laser power. The time of acquisition was between 30 seconds and 1 minute per spot. The four ph8 spectra were imported into the MALDI BioTyper software (version 2.0, Bruker) and analyzed by standard pattern matching (with default parameter settings) against the main spectra of 3,769 bacteria including the spectra from *A. finegoldii*, *A. onderdonkii*, *A. shahii*, ***A. senegalensis*** and *A. timonensis* used as reference data, in the BioTyper database. The method of identification included the m/z from 3,000 to 15,000 Da. For every spectrum, 100 peaks at most were taken into account and compared with spectra in the database. A score enabled either an identification, or non-identification, from the tested species: a score > 2 with a validly published species enabled a presumed identification at the species level, a score > 1.7 but < 2 enabled a presumed identification at the genus level; and a score < 1.7 did not enable an identification. For strain ph8, the obtained score was 1.1, suggesting that this isolate was not a member of a known species. We incremented our database with the spectrum from strain ph8 (Figure 4). Finally, the gel view allows us to highlight the spectra differences with other of *Alistipes* genera members (Figure 5).

Genome sequencing information

The organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA similarity to other members of the genus *Alistipes*, and is part of a study of the human digestive flora aiming at isolating all bacterial species contained within human feces. It was the seventh genome of an *Alistipes* species and the first genome of *Alistipes obesi* sp. nov. The EMBL accession number is CAHA00000000 and consists of 59 contigs. Table 3 shows the project information and its association with MIGS version 2.0 compliance [21].

Figure 2. Gram staining of *A. obesi* strain ph8^T

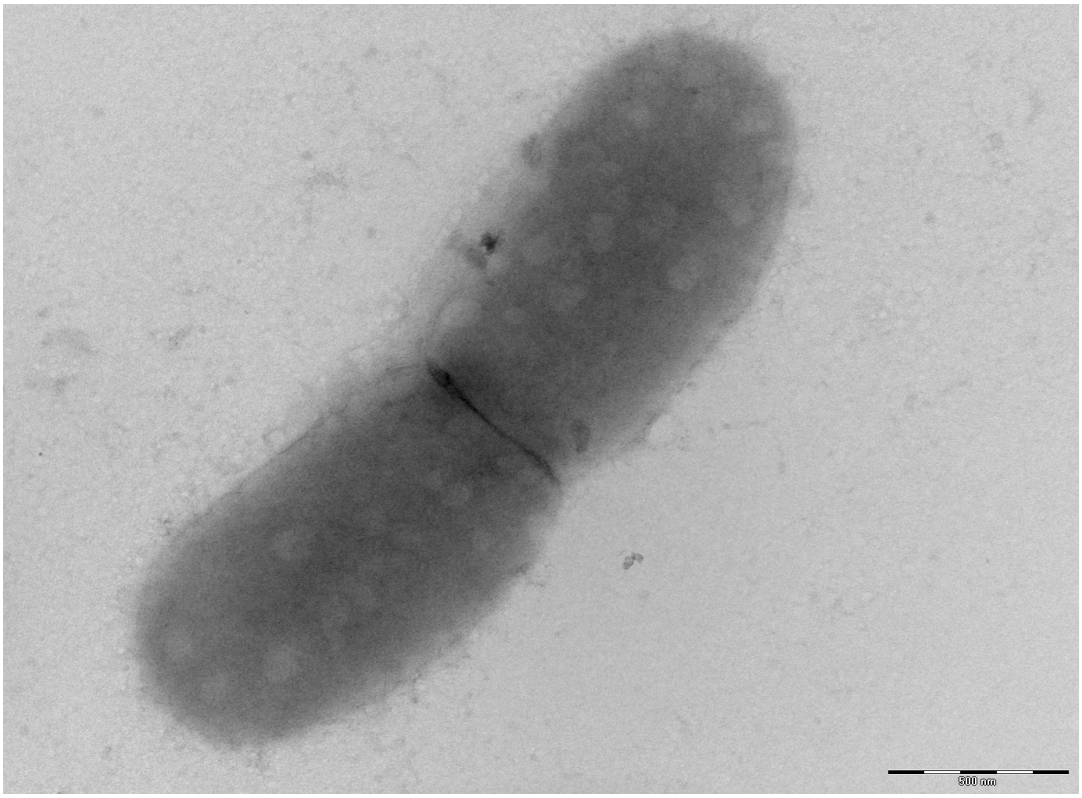

Figure 3. Transmission electron microscopy of *A. obesi* strain ph8^T, using a Morgani 268D (Philips) at an operating voltage of 60kV. The scale bar represents 500 nm.

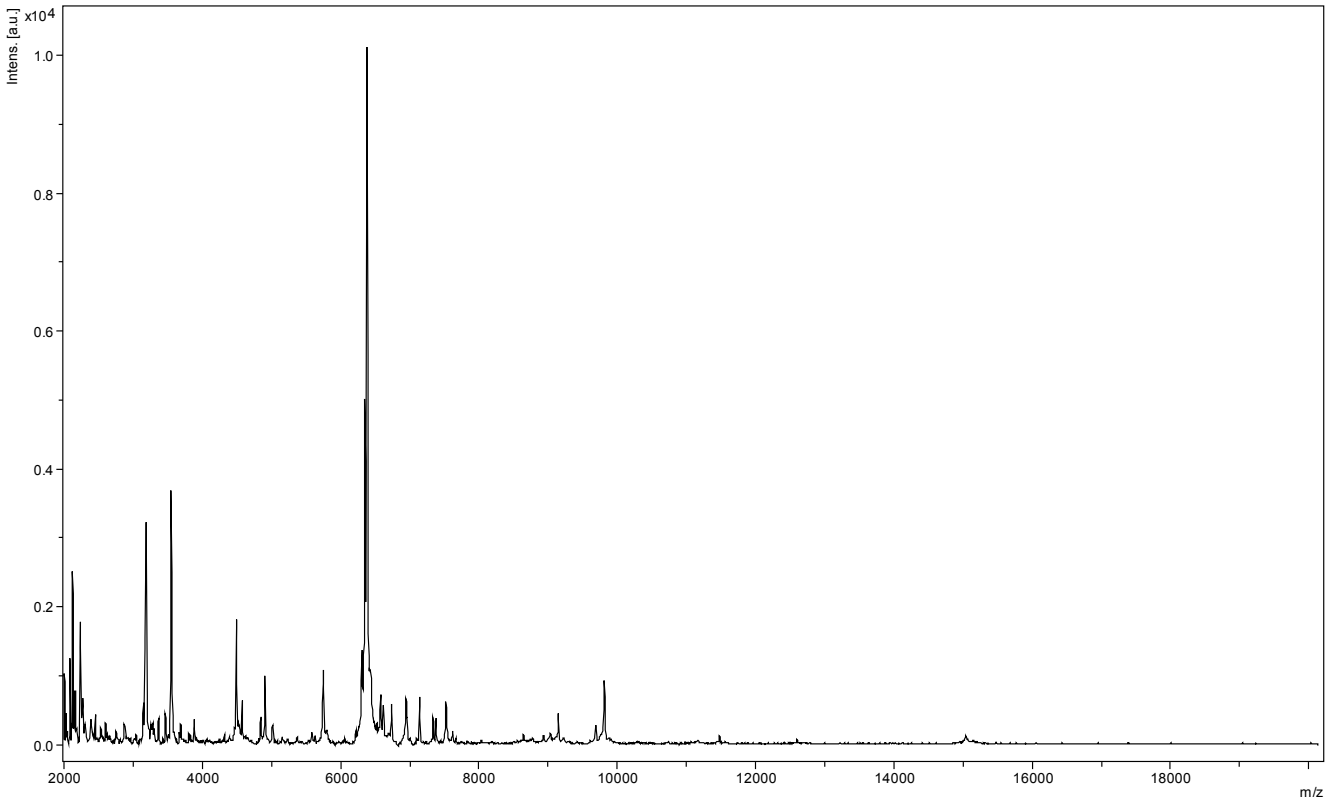

Figure 4. Reference mass spectrum from *A. obesi* strain ph8^T. Spectra from 4 individual colonies were compared and a reference spectrum was generated.

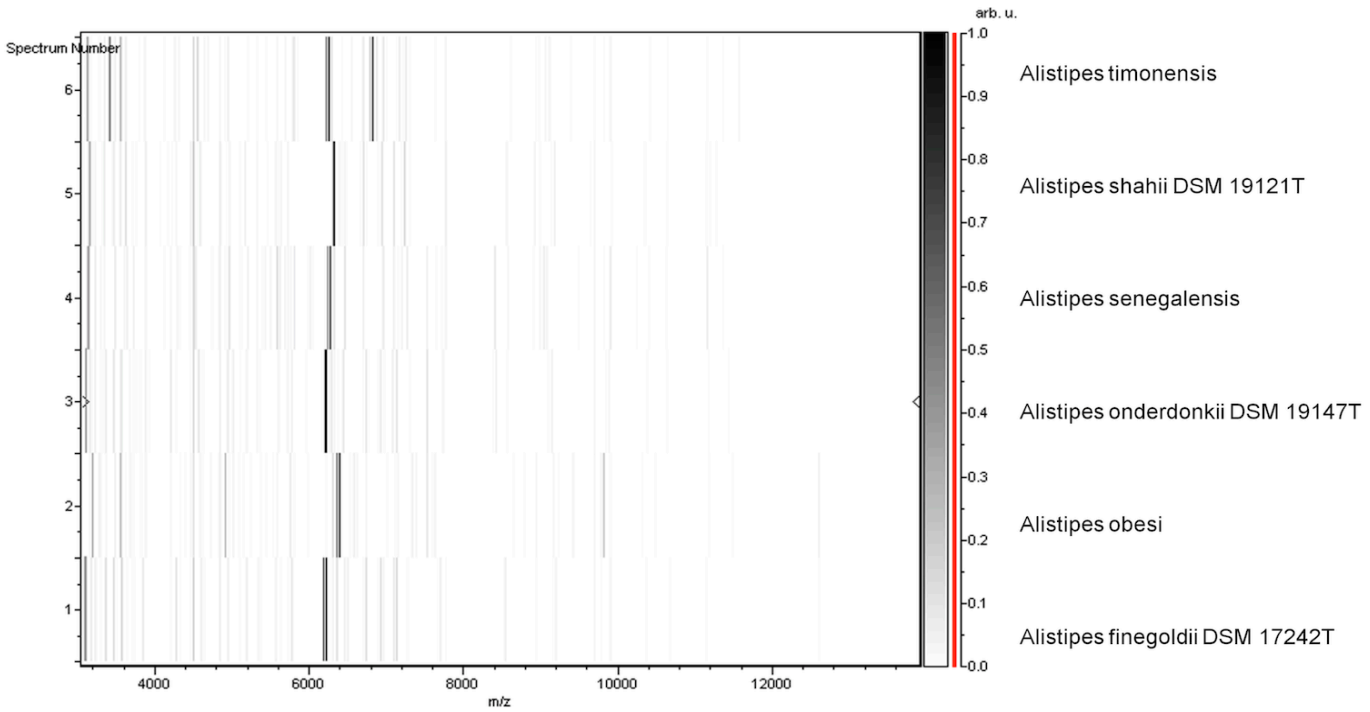

Figure 5. Gel view comparing *Alistipes obesi* ph8^T spectra with other members into *Alistipes* genera (*Alistipes timonensis*, *Alistipes senegalensis*, *Alistipes shahii*, *Alistipes onderdonkii* and *Alistipes finegoldii*). The Gel View displays the raw spectra of all loaded spectrum files arranged 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 a peak is displayed with and the peak intensity in arbitrary units.

Table 3. Project information

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality draft
MIGS-28	Libraries used	One paired end 3-kb library and one Shotgun library
MIGS-29	Sequencing platforms	454 GS FLX Titanium
MIGS-31.2	Fold coverage	18×
MIGS-30	Assemblers	Newbler version 2.5.3
MIGS-32	Gene calling method	Prodigal
	EMBL ID	CAHA00000000
	EMBL Date of Release	May 30, 2012
	Project relevance	Study of the human gut microbiome

Growth conditions and DNA isolation

A. obesi sp. nov. strain ph8^T (CSUR= P186, DSM=25724), was grown anaerobically on 5% sheep blood-enriched Columbia agar (BioMerieux) at 37°C. Four petri dishes were spread and resuspended in 3×100µl of G2 buffer (EZ1 DNA Tissue kit, Qiagen, Hilden, Germany). A first mechanical lysis was performed by glass powder on the Fastprep-24 device (MP Biomedicals, Santa Ana, CA, USA) using 2×20 seconds cycles. DNA was then treated with 2.5 µg/µL lysozyme for 30 minutes at 37°C and extracted using the BioRobot EZ 1 Advanced XL (Qiagen). The DNA concentration was measured at 120.7 ng/µL using the Genios fluorometer (Tecan, Lyon, France).

Genome sequencing and assembly

Five µg of DNA was mechanically fragmented on a Hydroshear device (Digilab, Holliston, MA,USA) with an enrichment size at 3-4kb. The DNA fragmentation was visualized using the 2100 BioAnalyzer (Agilent, Massy, France) on a DNA labchip 7500 with an optimal size of 3.213 kb. A 3-kb paired-end library was constructed using the GS Rapid library Prep kit (Roche) according to the 454 GS FLX Titanium paired-end protocol. Circularization and nebulization were performed and generated a pattern with an optimal size of 454 bp. After PCR amplification through 15 cycles followed by double size selection, the single stranded paired-end library was then quantified using the Genios fluorometer (Tecan) at 390 pg/µL. The library concentration equivalence was calculated as 1.58E+08 molecules/µL. The library was stored at -20°C until further use.

The paired-end library was clonally-amplified with 1cpb in 4 SV-emPCR reactions with the GS Titanium SV emPCR Kit (Lib-L) v2 (Roche). The yield of the emPCR was 18.6%, in the 5 to 20% range recommended by the Roche procedure. Approximately 790,000 beads were loaded on one quarter region of a GS Titanium PicoTiterPlate PTP Kit 70×75 and sequenced with the GS FLX Titanium Sequencing Kit XLR70 (Roche). The run was performed overnight and then analyzed on the cluster through the gsRunBrowser and Newbler assembler (Roche). A total of 192,803 passed filter wells were obtained and generated 56.6 Mb with a length average of 293.5 bp. The passed filter sequences were assembled using Newbler with 90% identity and 40 bp as overlap. The final assembly identified 3 scaffolds and 59 contigs (>1,500bp), for a genome size of 3.16 Mb.

Genome annotation

Open Reading Frames (ORFs) were predicted using Prodigal [34] with default parameters but the predicted ORFs were excluded if they spanned a sequencing gap. The predicted bacterial protein sequences were searched against the GenBank database and the Clusters of Orthologous Groups (COG) databases using BLASTP. The tRNAScanSE tool [35] was used to find tRNA genes, whereas ribosomal RNAs were found by using RNAmmer [36] and BLASTn against GenBank. Signal peptides and transmembrane helices were predicted using SignalP [37] and TMHMM [38], respectively. ORFans were identified if their BLASTP *E*-value was lower than 1e-03 for alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, we used an *E*-value of 1e-05. Such parameter thresholds have already been used in previous works to define ORFans. To estimate the mean level of nucleotide sequence similarity at the genome level between *Alistipes obesi* strain ph8T and other members of the *Alistipes* genera, we compared genomes two by two and determined the mean percentage of nucleotide sequence identity among orthologous ORFs using BLASTn. Orthologous genes were detected using the Proteinortho software [39]. We compared *A. obesi* strain ph8^T with *A. finegoldii* strain AHN 2437 (GenBank accession number CP003274), *A. indistinctus* strain YIT 12060 (ADLD00000000), *A. putredinis* strain DSM 17216 (ABFK00000000), *A. senegalensis* strain JC50^T (CAHI00000000), *A. shahii* strain WAL 8301 (FP929032) and *A. timonensis* strain JC136^T (CAEG00000000).

Genome properties

The genome is 3,162,233 bp long (1 chromosome, but no plasmid) with a 58.6% G+C content (Table 4, Figure 6). Of the 2,672 predicted genes, 2,623 were protein-coding genes and 49 were RNAs. A total of 1,409 genes (52.75%) were assigned a putative function. One hundred twenty-seven genes were identified as ORFans (4.8%). The remaining genes were annotated as hypothetical proteins. The distribution of genes into COGs functional categories is presented in Table 5. The properties and the statistics of the genome are summarized in Tables 4 and 5.

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

Attribute	Value	% of total[a]
Genome size (bp)	3,162,233	
DNA coding region (bp)	2,799,840	88.53
G+C content (bp)	1,853,068	58.6
Total genes	2,672	100
RNA genes	49	1.83
Protein-coding genes	2,623	98.16
Genes with function prediction	1,409	52.75
Genes assigned to COGs	1,559	58.36
Genes with peptide signals	429	16.06
Genes with transmembrane helices	512	19.16

[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.

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

Code	Value	%age[a]	Description
J	134	5.02	Translation
A	1	0.04	RNA processing and modification
K	67	2.51	Transcription
L	114	4.27	Replication, recombination and repair
B	0	0	Chromatin structure and dynamics
D	19	0.71	Cell cycle control, mitosis and meiosis
Y	0	0	Nuclear structure
V	42	1.57	Defense mechanisms
T	25	0.94	Signal transduction mechanisms
M	129	4.83	Cell wall/membrane biogenesis
N	5	0.19	Cell motility
Z	0	0	Cytoskeleton
W	0	0	Extracellular structures
U	19	0.71	Intracellular trafficking and secretion
O	41	1.54	Posttranslational modification, protein turnover, chaperones
C	103	3.86	Energy production and conversion
G	108	4.04	Carbohydrate transport and metabolism
E	90	3.37	Amino acid transport and metabolism
F	45	1.68	Nucleotide transport and metabolism
H	55	2.06	Coenzyme transport and metabolism
I	38	1.42	Lipid transport and metabolism
P	80	3.0	Inorganic ion transport and metabolism
Q	6	0.22	Secondary metabolites biosynthesis, transport and catabolism
R	202	7.56	General function prediction only
S	86	3.22	Function unknown
-	150	5.62	Not in COGs

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

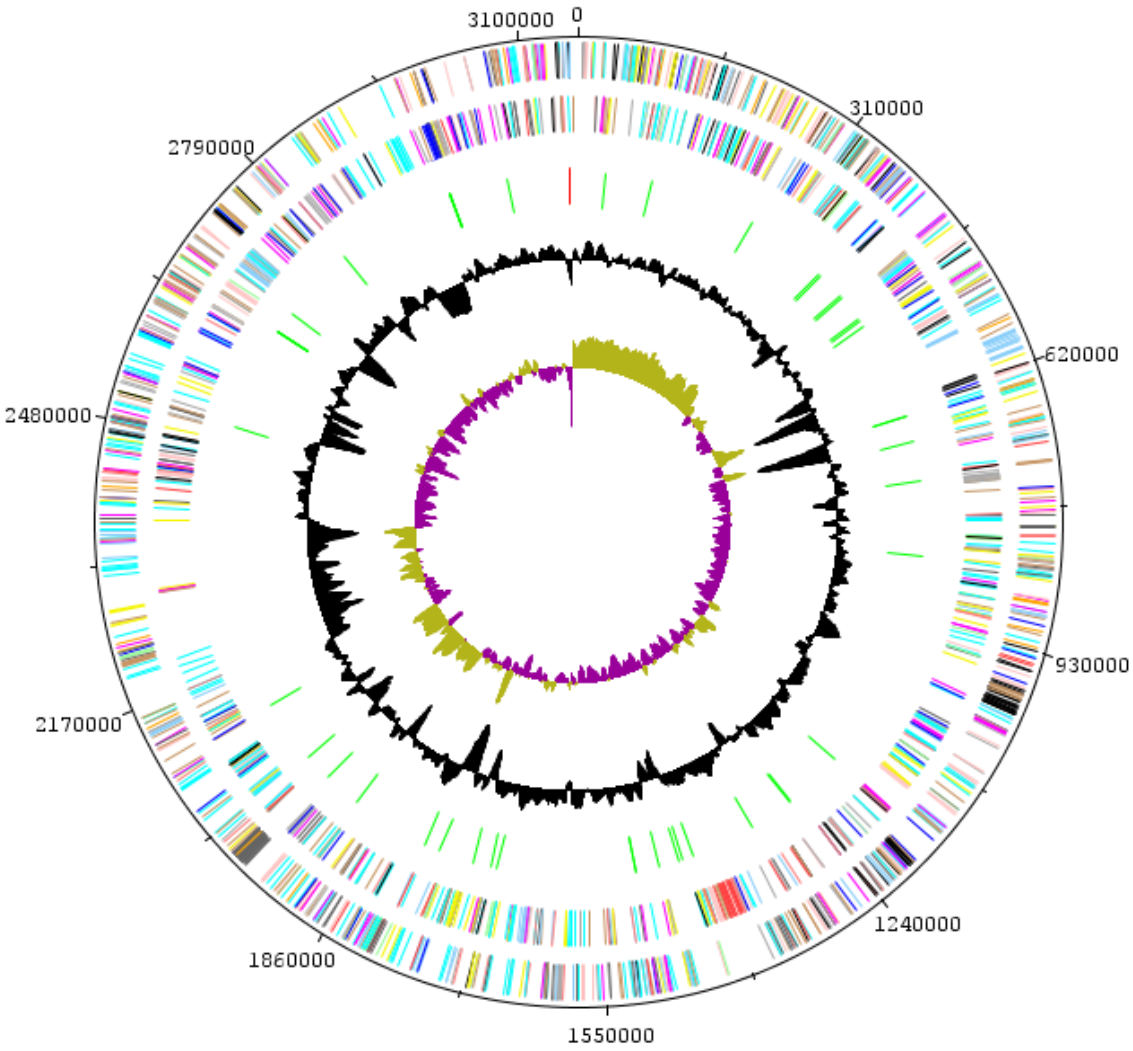

Figure 6. Graphical circular map of the chromosome. From outside to the center: Genes on forward strand (colored by COG categories), genes on reverse strand (colored by COG categories), RNA genes (tRNAs green, rRNAs red), GC content, GC skew.

Comparison with the genomes of other *Alistipes* species

Currently, the genomes from six validly published *Alistipes* species: *A. finegoldii* strain DSM 17242 (GenBank accession number CP003274), *A. indistinctus* strain YIT 12060 (ADLD00000000), *A. putredinis* strain DSM 17216 (ABFK00000000) and *A. shahii* strain WAL 8301 (FP929032), *A. senegalensis* strain JC50^T (CAHI00000000) and *A. timonensis* strain JC136^T (CAEG00000000) are available. The draft genome of *A. obesi* is larger than *A. indistinctus* and *A. putredinis* (3.16 *vs* 2.85 and 2.55Mb, respectively), but of a smaller size than *A. finegoldii*, *A. shahii*, *A. senegalensis* and *A. timonensis* (3.73, 3.76, 4.01 and 3.49Mb, respectively). With the exception of *A. timonensis,* which exhibits a value of 58.8%, *A. obesi* has a higher G+C content than *A. finegoldii*, *A. indistinctus*, *A. putredinis A. shahii* and *A. senegalensis* (58.6 *vs* 56.6, 54.8, 53.3, 57.6 and 58.4%, respectively. *A. obesi* has more predicted genes than *A. indistinctus* and *A. putredinis* (2,619 *vs* 2,342 and 2,335 respectively), but fewer than *A. finegoldii*, *A. shahii*, *A. senegalensis* or *A. timonensis* (3,231, 3,132, 3,161, and 2,709, respectively). In addition, A. *obesi* shared 1,417,1,462, 1,429, 1,431, 1,135 and 1,218 orthologous genes with *A. timonensis*, *A. senegalensis*, *A. shahii*, *A. finegoldii*, *A. indistinctus* and *A. putredinis* respectively. The average nucleotide sequence identity ranged from 69.70 to 90.98 % among *Alistipes* species, and from 70.00 to 76.81% between A. *obesi* and other species, thus confirming its new species status (Table 6).

Table 6. Number of orthologous genes (upper right), average nucleotide identity levels (lower left) between pairs of genomes and numbers of proteins per genome (bold) [39].

	A. obesi	***A. timonensis***	***A. senegalensis***	***A. shahii***	***A. finegoldii***	***A. indistinctus***	***A. putredinis***
A. obesi	**2,619**	1,417	1,462	1,429	1,431	1,135	1,218
A. timonensis	76.19	**2,709**	1,764	1,650	1,585	1,210	1,238
A. senegalensis	76.40	90.98	**3,161**	1,739	1,660	1,218	1,277
A. shahii	76.81	80.03	86.33	**3,132**	1,674	1,155	1,270
A. finegoldii	76.23	81.14	82.04	82.90	**3,231**	1,202	1,303
A. indistinctus	70.00	70.05	70.02	70.00	69.91	**2,342**	1,038
A. putredinis	74.49	75.21	75.32	75.50	76.23	69.70	**2,335**

Conclusion

On the basis of phenotypic, phylogenetic and genomic analyses (taxono-genomics), we formally propose the creation of *A. obesi* sp. nov., which contains strain ph8^T. This bacterium has been cultivated from an obese patient in Marseille, France.

Description of *Alistipes obesi* sp. nov.

Alistipes obesi (o.be.si. L. masc. gen. adj. *obesi* of an obese, the disease presented by the patient from whom the type strain ph8^T was isolated). Colonies are 0.5 mm in diameter and are translucent and light grey on blood-enriched Columbia agar. Cells are rod-shaped with a mean diameter of 0.61 µm. Optimal growth is achieved anaerobically. No growth is observed in aerobic or microaerophilic conditions. Growth occurs between 37°C-45°C, with optimal growth observed at 37°C.

Cells stain Gram negative and are motile. Catalase, α-galactosidase, β-galactosidase, N-acetyl-β-glucosaminidase, acid phosphatase, alkaline phosphatase, leucyl glycine arylamidase, alanine arylamidase, esterase, esterase lipase and Naphtol-AS-BI phosphohydrolase activities are present. Oxidase, nitrate reduction, indole formation, urease, arginine dihydrolase, α- and β-glucosidase, 6-phospho-β-galactosidase, arginine arylamidase, proline arylamidase, phenylalanine arylamidase, leucine arylamidase, pyroglutamic acid arylamidase, tyrosine arylamidase, glycine arylamidase, histidine arylamidase, glutamyl glutamic acid arylamidase and serine arylamidase, mannose and raffinose fermentation activities are absent. Cells are susceptible to imipenem, ciprofloxacin, metronidazole, nitrofurantoin and rifampicin, but resistant to penicillin G, amoxicillin, amoxicillin-clavulanic acid, erythromycin, vancomycin, gentamicin, doxycycline, ceftriaxone and trimethoprim/sulfamethoxazole. The G+C content of the genome is 58.6%. The 16S rRNA and genome sequences are deposited in Genbank under accession numbers JN837494 and CAHA00000000, respectively. The type strain ph8^T (CSUR= P186 = DSM 25724) was isolated from the fecal flora of an obese patient in France.

Acknowledgements

The authors thank the XEGEN company (www.xegen.fr) for automating the genomic annotation process. This work was funded by the Mediterranee-Infection foundation.

References

1. Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, Bittar F, Fournous G, Gimenez G, Maraninchi M, *et al*. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012; **18**:1185-1193. PubMed

2. Rossello-Mora R. DNA-DNA Reassociation Methods Applied to Microbial Taxonomy and Their Critical Evaluation. In: Stackebrandt E (ed), Molecular Identification, Systematics, and population Structure of Prokaryotes. Springer, Berlin, 2006, p. 23-50.

3. Kokcha S, Mishra AK, Lagier JC, Million M, Leroy Q, Raoult D, Fournier PE. Non contiguous-finished genome sequence and description of *Bacillus timonensis* sp. nov. *Stand Genomic Sci* 2012; **6**:346-355. PubMed http://dx.doi.org/10.4056/sigs.2776064

4. Lagier JC, El Karkouri K, Nguyen TT, Armougom F, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. *Stand Genomic Sci* 2012; **6**:116-125. PubMed http://dx.doi.org/10.4056/sigs.2415480

5. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Clostridium senegalense* sp. nov. *Stand Genomic Sci* 2012; **6**:386-395. PubMed

6. Mishra AK, Gimenez G, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes senegalensis* sp. nov. *Stand Genomic Sci* 2012; **6**:304-314. http://dx.doi.org/10.4056/sigs.2625821

7. Lagier JC, Armougom F, Mishra AK, Ngyuen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Alistipes timonensis* sp. nov. *Stand Genomic Sci* 2012; **6**:315-324. PubMed http://dx.doi.org/10.4056/sigs.2685917

8. Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Peptoniphilus timonensis* sp. nov. *Stand Genomic Sci* 2012; **7**:1-11. http://dx.doi.org/10.4056/sigs.2956294

9. Mishra AK, Lagier JC, Rivet R, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Paenibacillus senegalensis* sp. nov. *Stand Genomic Sci* 2012; **7**:70-81. http://dx.doi.org/10.4056/sigs.3054650

10. Lagier JC, Gimenez G, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Herbaspirillum massiliense* sp. nov. *Stand Genomic Sci* 2012; **7**:200-209. PubMed http://dx.doi.org/10.4056/sigs.3086474

11. Roux V, El Karkouri K, Lagier JC, Robert C, Raoult D. Non-contiguous finished genome sequence and description of *Kurthia massiliensis* sp. nov. *Stand Genomic Sci* 2012; **7**:221-232. PubMed http://dx.doi.org/10.4056/sigs.3206554

12. Kokcha S, Ramasamy D, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Brevibacterium senegalense* sp. nov. *Stand Genomic Sci* 2012; **7**:233-245. PubMed http://dx.doi.org/10.4056/sigs.3256677

13. Ramasamy D, Kokcha S, Lagier JC, N'Guyen TT, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Aeromicrobium massilense* sp. nov. *Stand Genomic Sci* 2012; **7**:246-257. PubMed http://dx.doi.org/10.4056/sigs.3306717

14. Lagier JC, Ramasamy D, Rivet R, Raoult D, Fournier PE. Non-contiguous finished genome sequence and description of *Cellulomonas massiliensis* sp. nov. *Stand Genomic Sci* 2012; **7**:258-270. PubMed http://dx.doi.org/10.4056/sigs.3316719

15. Rautio M, Eerola E, Vaisanen-Tunkelrott ML, Molitoris D, Lawson P, Collins MD, Jousimies-Somer S. Reclassification of *Bacteroides putredinis* (Weinberg et al., 1937) in a new genus *Alistipes* gen. nov., as *Alistipes putredinis* comb. nov., and description of *Alistipes finegoldii* sp. nov., from human sources. *Syst Appl Microbiol* 2003; **26**:182-188. PubMed http://dx.doi.org/10.1078/072320203322346029

16. Nagai F, Morotomi M, Watanabe Y, Sakon H, Tanaka R. *Alistipes indinstinctus* sp. nov. and *Odoribacter laneus* sp. nov., common members of the human intestinal microbiota isolated from faeces. *Int J Syst Evol Microbiol* 2010; **60**:1296-1302. PubMed http://dx.doi.org/10.1099/ijs.0.014571-0

17. Song Y, Kononen E, Rautio M, Liu C, Bryk A, Eerola E, Finegold SM. *Alistipes onderdonkii* sp. nov. and *Alistipes shahii* sp. nov., of human origin. *Int J Syst Evol Microbiol* 2006; **56**:1985-1990. PubMed http://dx.doi.org/10.1099/ijs.0.64318-0

18. Fenner L, Roux V, Ananian P, Raoult D. *Alistipes finegoldii* in blood cultures from colon cancer patients. *Emerg Infect Dis* 2007; **13**:1260-1262. PubMed http://dx.doi.org/10.3201/eid1308.060662

19. Brook I. Clinical review: bacteremia caused by anaerobic bacteria in children. *Crit Care* 2002; **6**:205-211. PubMed http://dx.doi.org/10.1186/cc1490

20. Rautio M, Saxen H, Siitonen A, Nikku R, Jousimies-Somer H. Bacteriology of histopathologically defined appendicitis in children. *Pediatr Infect Dis J* 2000; **19**:1078-1083. PubMed http://dx.doi.org/10.1097/00006454-200011000-00010

21. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, *et al*. The minimum information about a genome sequence (MIGS) specification.

Nat Biotechnol 2008; **26**:541-547. PubMed http://dx.doi.org/10.1038/nbt1360

22. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea, Bacteria*, and *Eukarya*. *Proc Natl Acad Sci USA* 1990; **87**:4576-4579. PubMed http://dx.doi.org/10.1073/pnas.87.12.4576

23. Validation List No. 143. *Int J Syst Evol Microbiol* 2012; **62**:1-4.

24. Krieg NR, Ludwig W, Euzéby J, Whitman WB. Phylum XIV. *Bacteroidetes* phyl. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 4, Springer, New York, 2011, p. 25.

25. Krieg NR. Class I. *Bacteroidia* class. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 4, Springer, New York, 2011, p. 25.

26. Krieg NR. Order I. *Bacteroidales* ord. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 4, Springer, New York, 2011, p. 25.

27. Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB. Family III. *Rikenellaceae* fam. nov. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, Ludwig W, Whitman WB (eds), Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 4, Springer, New York, 2011, p. 54.

28. Validation List no. 94. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. *Int J Syst Evol Microbiol* 2003; **53**:1701-1702. PubMed http://dx.doi.org/10.1099/ijs.0.03001-0

29. Rautio M, Eerola E, Väisänen-Tunkelrott ML, Molitoris D, Lawson P, Collins MD, Jousimies-Somer H. Reclassification of *Bacteroides putredinis* (Weinberg et al., 1937) in a new genus *Alistipes* gen. nov., as *Alistipes putredinis* comb. nov., and description of *Alistipes finegoldii* sp. nov., from human sources. *Syst Appl Microbiol* 2003; **26**:182-188. PubMed http://dx.doi.org/10.1078/072320203322346029

30. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25-29. PubMed http://dx.doi.org/10.1038/75556

31. List of Prokaryotic names with Standing in Nomenclature. *http://www.bacterio.cict.fr*

32. Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006; **33**:152-155.

33. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, Raoult D. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis* 2009; **49**:543-551. PubMed http://dx.doi.org/10.1086/600885

34. Prodigal. http://prodigal.ornl.gov/

35. Lowe TM, Eddy SR. t-RNAscan-SE: a program for imroved detection of transfer RNA gene in genomic sequence. *Nucleic Acids Res* 1997; **25**:955-964. PubMed

36. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007; **35**:3100-3108. PubMed http://dx.doi.org/10.1093/nar/gkm160

37. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**:783-795. PubMed http://dx.doi.org/10.1016/j.jmb.2004.05.028

38. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001; **305**:567-580. PubMed http://dx.doi.org/10.1006/jmbi.2000.4315

39. Lechner M, Findeib S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* 2011; **12**:124. PubMed http://dx.doi.org/10.1186/1471-2105-12-124