microbiology research

MDPI

Article

Negativibacillus massiliensis gen. nov., sp. nov., a New Bacterial Genus Isolated from a Human Left Colon Sample

Camille Valles [1,2], Morgane Mailhe [2], Davide Ricaboni [1], Nicholas Armstrong [1], Stéphane Alibar [2], Véronique Vitton [3], Jean-Christophe Lagier [1,2], Didier Raoult [1,2] and Maryam Tidjani Alou [1,2,*]

1 Faculty of Medicine, Aix Marseille University, IRD, AP-HM, MEPHI, 13005 Marseille, France; vallescamille@gmail.com (C.V.); davide.ricaboni@unimi.it (D.R.); nicholas.armstrong@univ-amu.fr (N.A.); jclagier@yahoo.fr (J.-C.L.); didier.raoult@gmail.com (D.R.)
2 IHU-Méditerranée Infection, 13005 Marseille, France; morgane.mailhe@gmail.com (M.M.); stephane.alib@gmail.com (S.A.)
3 Service de Gastroentérologie, Hôpital Nord, Assistance Publique-Hôpitaux de Marseille, 13015 Marseille, France; veronique.vitton@ap-hm.fr
* Correspondence: tidjani_maryam@hotmail.com

Abstract: A new genus, a member of the *Ruminococcaceae* family, was isolated from the left colon of a healthy woman. Strain Marseille P3213 was a non-motile, spore-forming, Gram-stain negative, rod-shaped bacterium. This strictly anaerobic species reached optimal growth after an incubation of 72 h at 37 °C. The 16S rRNA gene sequence of this strain shared a 93.52% similarity level with *Harryflintia acetispora* strain V20-281a, its closest phylogenetic neighbor with standing in the nomenclature. Its genome had a size of 2.87 Mb, with a 45.81% G + C content. We hereby propose the creation of *Negativibacillus massiliensis* strain P3213^T as the 43rd genus within the *Ruminococcaceae* family.

Keywords: *Negativibacillus massiliensis*; new genus; culturomics; taxono-genomics; genome

Citation: Valles, C.; Mailhe, M.; Ricaboni, D.; Armstrong, N.; Alibar, S.; Vitton, V.; Lagier, J.-C.; Raoult, D.; Tidjani Alou, M. *Negativibacillus massiliensis* gen. nov., sp. nov., a New Bacterial Genus Isolated from a Human Left Colon Sample. *Microbiol. Res.* **2021**, *12*, 29–42. https://doi.org/10.3390/microbiolres12010004

Academic Editor: Elisa Bona

Received: 25 December 2020
Accepted: 4 February 2021
Published: 9 February 2021

1. Introduction

The bacterial diversity of the gut microbiota has a tremendous impact on physiological functions and disease susceptibility. Therefore, studies regarding its role and diversity are of great importance for human health [1,2], and they have increased exponentially over the two last decades [3]. These studies were mostly conducted using culture independent approaches [1]. However, the last decade has witnessed a turning point in the study of gut microbiota diversity with a rebirth of culture methods [2,4], with high-throughput culture methods such as culturomics [2,5]. The culturomics method was coined for the exploration of gut microbiota diversity, and is based on the multiplication of culture conditions with a variety of physic-chemical parameters, such as the culture medium, atmosphere, temperature and pH [5]. This technique is complementary to metagenomics studies [2,5], and has led to an increased diversity of the cultured human intestinal microbiota, and therefore to the increase of the repertoire of known bacterial species to man [6].

In 2016, as a part of a culturomics study focused on the modifications of human gut microbiota diversity along the different anatomical sites of the gastrointestinal tract [7], a new member of the *Ruminococcaceae* family was isolated. The family *Ruminococcaceae*—formerly known as *Clostridium* cluster III [8]—was first coined in 2010, and presently consists of 41 genera which are all strictly anaerobic [9]. These bacterial species are morphologically diverse, including cocci and bacilli, as well as intermediate forms [8]. This family, of which the members share the ability to break down cellulose [10], includes numerous commensals of the human gut [11].

In this study, we present a complete description of this new member of the *Ruminococcaceae* family according the concept of taxonogenomics. This innovative concept uses a combination of phenotypic, proteomic and genomic characteristics [12,13] to classify and describe new bacterial species.

2. Material and Methods

2.1. Sample Collection

The left colon sample was collected from a seventy-six-year-old woman who had a body mass index of 26 kg/m^2. This woman was hospitalized for a colonoscopy and fibroscopy in order to assess the condition of her esophagus and observe colonic polyps. The samples were immediately cultured without prior storage. Written consent was obtained from the patient for this study, which was validated by the ethics committee of the Institut Fédératif de Recherche IFR48 under number 09-022.

2.2. Bacterial Strain and Identification

The bacterial diversity of this sample was studied using the standardized culturomics technique consisting of 18 culture conditions [14].

For each of the 18 culture conditions, the samples were incubated in a liquid medium for enrichment. Subsequently, at different timepoints (incubation day 1, 3, 7, 10, 15, 21, 30), this enriched culture was diluted and seeded on 5% sheep-blood–enriched Columbia agar (COS, bioMerieux Craponne France). The colonies obtained after 48 to 72 h of incubation were then subcultured on COS plates. Each colony was identified using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) [15]. Each colony was tested in duplicate; the obtained spectra were matched with the database consisting of the protein profiles of numerous species. The bacteria were considered to be identified if the score obtained was higher than 1.9. Between 1.7 and 1.9, the colony was considered identified at the genus level. If the score was lower than 1.7, the 16S rRNA gene was then sequenced using fD1-rP2 primers, as previously described [16], using a 3130-XLsequencer (Applied Biosciences, Saint Aubin, France). The obtained spectra were then added to the in-lab database publicly available at https://www.mediterranee-infection.com/acces-ressources/base-de-donnees/urms-data-base (accessed on 25 December 2020). The obtained sequences were matched with the NCBI database using BLASTn. As defined by Stackebrandt and Ebers [17], the thresholds to define a new species and a new genus were 98.65% and 95%, respectively.

2.3. Phenotypic Characteristics

2.3.1. Optimal Growth Conditions

In order to determine the optimal growth conditions of strain P3213, different culture parameters were evaluated: incubation temperatures and atmospheres, as well as pH and salt tolerance. The optimal growth temperature was established by testing the growth at 25 °C, 28 °C, 37 °C, 45 °C and 56 °C in three atmospheres—namely aerobic, anaerobic and microaerophilic—using anaeroGEN (Beckton Dickinson, Arcueil, France) and campyGen (Beckton Dickinson, Arcueil, France) generators, respectively. Different pHs were tested—5, 5.5, 6, 6.5, 7, 7.5, and 8—using a solid medium containing, per liter: 42 g Columbia agar (ThermoFisher Scientific, Dardilly, France), 10 g NaCl, 5 g $MgCl_2$ (MP Biomedicals, Illkirch, France), 5 g $MgSO_4$, 2 g KCl, 1 g $CaCl_2$ (VWR, Briaire, France) and 2 g glucose (MP Biomedicals, Illkirch, France). The pH was adjusted by adding a solution of NaOH or HCl.

The salt tolerance was evaluated using a medium containing, per liter: 42 g Columbia agar, 10 g $MgCl_2$, 10 g $MgSO_4$, 4 g KCl, 1 g $CaCl_2$, 0.5 g $NaHCO_3$, 5 g yeast extract and 2 g glucose. Different quantities of NaCl were added to reach concentrations of 7.5%, 10%, 15% and 20%.

2.3.2. Morphologic and Biochemical Characteristics, and Antibiotic Susceptibility

Phenotypic characteristics like Gram staining (bioMerieux, Craponne, France), oxidase (Beckton Dickinson, Arcueil, France), catalase (bioMerieux, Craponne, France), motility and sporulation were determined according to the manufacturer's instructions. Motility was determined by observing the fresh bacterial colonies using an optical microscope, magnification ×100. In order to test the spore-forming ability of strain P3213, a thermic

shock was used. In fact, a suspension of 10^8 cfu/mL was heated in a dry bath for 20 min at 80 °C. In total, 50 µL of this bacterial suspension was inoculated on COS plates.

Morphologic observations were also carried out by performing negative staining. Detection form var-coated grids were deposited on a 40 µL bacterial suspension drop and incubated at 37 °C for 30 min. This was then followed by a 10-s incubation on ammonium molybdate 1%. The grids were dried on blotting paper and observed with a Tecnai G20 transmission electron microscope (FEI Company, Limeil Brevannes, France).

In order to determine the metabolic features of strain P3213, three API strips were used. The 50CH API strip allowed the evaluation of the capacity of the studied strain to metabolize carbohydrates and their derivatives, such as heterosids, polyalcohols, uronic acids. The enzymatic capacity was evaluated using an API zym strip, while an API 20A strip was used to complete the metabolic profile of the strain. All of the strips were used according to the manufacturer's instructions (bioMerieux, Craponne, France). The antibiotic susceptibility was determined using the disk diffusion method according to the European Committee recommendations on antimicrobial susceptibility testing 2015 [18].

2.3.3. Fatty Acid Methyl Ester (FAME) Analysis by GC/MS

The cellular fatty acid methyl ester (FAME) analysis was performed by GC/MS in duplicate, with approximately 40 mg bacterial biomass per tube harvested from several culture plates. The fatty acid methyl esters were prepared as described by Sasser and colleagues [19]. The GC/MS analyses were carried out as described before [20]. Briefly, the 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). The obtained spectra were matched with the Standard Reference Database 1A (NIST, Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK) using MS Search 2.0.

2.4. Genomic Characteristics

2.4.1. Genome Sequencing

For the genomic DNA (gDNA) extraction, strain P3213 was cultured on three COS plates at 37 °C anaerobically. The plates were harvested and resuspended in 400 of Tris-EDTA (TE) buffer. The chemical lysis was carried out by adding 1 mL TE buffer; subsequently, a 30 min incubation at 37 °C with 2.5 µg/µL lysozyme (Sigma Aldrich, Saint-Quentin Fallavier, France) was performed, followed by an overnight incubation at 37 °C with 20 µg/µL proteinase K (Euromedex, Souffelweyersheim, France). The gDNA was purified using three consecutive phenol-chloroform extractions with an ethanol precipitation performed overnight at −20 °C. After centrifugation, the DNA was resuspended in 160 µL elution buffer. The gDNA was quantified using a Qubit assay with a high sensitivity kit (Life technologies, Carlsbad, CA, USA) to a concentration of 88.3 ng/µL, and then sequenced using the MiSeq Technology (Illumina Inc, San Diego, CA, USA) with the mate pair strategy. After barcoding the gDNA using the Nextera Mate Pair sample prep kit (Illumina, Evry, France) in order to mix it with gDNA from 11 other projects, 1.5 µg was used to prepare the mate pair library using the Nextera mate pair Illumina guide. The gDNA was simultaneously fragmented and labeled using a mate pair junction adapter. The fragmentation pattern was checked using an Agilent 2100 BioAnalyzer (Agilent Technologies Inc, Santa Clara, CA, USA) with a DNA 7500 labchip. DNA fragments ranging in size from 1.5 kb up to 11 kb were obtained, with an average size at 6.639 kb. The labeled fragments were circularized before mechanical shearing into small fragments, with an optimal size of 843 bp, on the Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). Using a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc., Santa Clara, CA, USA), the library profile was visualized and then normalized at 2 nM, and pooled. After the DNA denaturation and dilution at 15 pM, the pool of libraries was loaded onto the instrument, along with the flow cell for the automated cluster generation and sequencing run.

2.4.2. Genome Assemblage, Annotation and Comparison

The 1,423,574 paired reads were trimmed and assembled into eight scaffolds using the software SPAdes 3.9.0. The contigs with a size below 800 bp, as well as those with a coverage of less than 25% of the average coverage, were trimmed. The Open Reading Frames (ORFs) prediction was achieved using Prodigal [21], with the default parameters. The protein coding genes were predicted by matching against the NR database using BLASTP with an E-value of 10^{-3} (10^{-5} for a sequence shorter than 80 amino acids, coverage 0.7, and an identity percentage of 30%). The predicted protein coding genes were then matched against the Clusters of Orthologous Groups (COG) using BLASTP (E-value 10^{-3}, coverage 0.7 and identity percent 30%) in order to infer the functional abilities of the described organism. The transfer RNA genes (tRNA) were predicted using the tRNAScanSE tool [22], whereas the ribosomal RNAs (rRNA) were predicted using RNAmmer [23]. The lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [24]. The genes were considered to be ORFans if no hits were obtained using BLASTP (an E-value smaller than 10^{-3} for ORFs with a sequence size higher than 80 aa, or E-value smaller than 1×10^{5} for ORFs with sequence size lower than 80 aa).

For the genome comparison, the genomes were selected from the 16S rRNA phylogenetic tree using Phylopattern, an XEGEN software [25]. The retrieved sequences from the FTP of NCBI included the complete genome sequence, proteome genome sequence (all gene sequences encoding proteins in a genome) and Orfeome genome sequence (all gene sequences encoding orphan proteins in a genome). The proteomes were analyzed using proteinOrtho [26]. Moreover, the distribution into functional classes of the predicted genes according to the clusters of orthologous groups of proteins was performed as described above. The genomic similarity between the compared genomes was evaluated using two parameters: digital DNA–DNA Hybridization (dDDH), a parameter highly correlated with DDH [27,28], and OrthoANI [29]. OrthoANI, a similarity score consisting of the mean value of nucleotide similarity between two compared genomes determined using the OAT software, was computed for each couple of genomes. The dDDH was determined using Type Strain Genome Server TYGS (https://tygs.dsmz.de/, accessed on 25 December 2020) [30] and interpreted with the d4 formula, as recommended. The Multi-Agent software system DAGOBAH [31], including Figenix [32] libraries, was used to perform the annotation and comparison.

3. Results

3.1. Strain Identification

Strain P3213 was isolated after 14 days of preincubation in an anaerobic blood culture bottle supplemented with 5 mL sheep blood and 5 mL sterile rumen fluid, followed by a 72 h incubation in an anaerobic atmosphere at 37 °C on COS plates.

A score under 1.7 was obtained for strain P3213 after the MALDI-TOF MS analysis. Therefore, the 16S rRNA gene was sequenced. The sequence—which is available under accession number LT598596 Marseille-P3213—exhibited a 93.52% identity with *Harriflyntia acetispora* strain V20-281a (GenBank accession no. KU999999), the phylogenetically-closest species with a validly-published name, as shown in the phylogenetic tree in Figure 1. The reference spectra for *Negativibacillus massiliensis* (Figure 2) were incremented in the MALDI-TOF database.

The GenBank accession numbers for the 16S rRNA gene are indicated in parenthesis. The sequences were aligned, and the phylogenetic inferences were obtained using the maximum-likelihood method within the MEGA7 software. The numbers at the nodes are the percentages of the bootstrap values obtained by repeating the analysis 1000 times, in order to generate a majority consensus tree. *Catabacter hongkongensis* was used as an outgroup. The scale bar represents a 1% nucleotide sequence divergence.

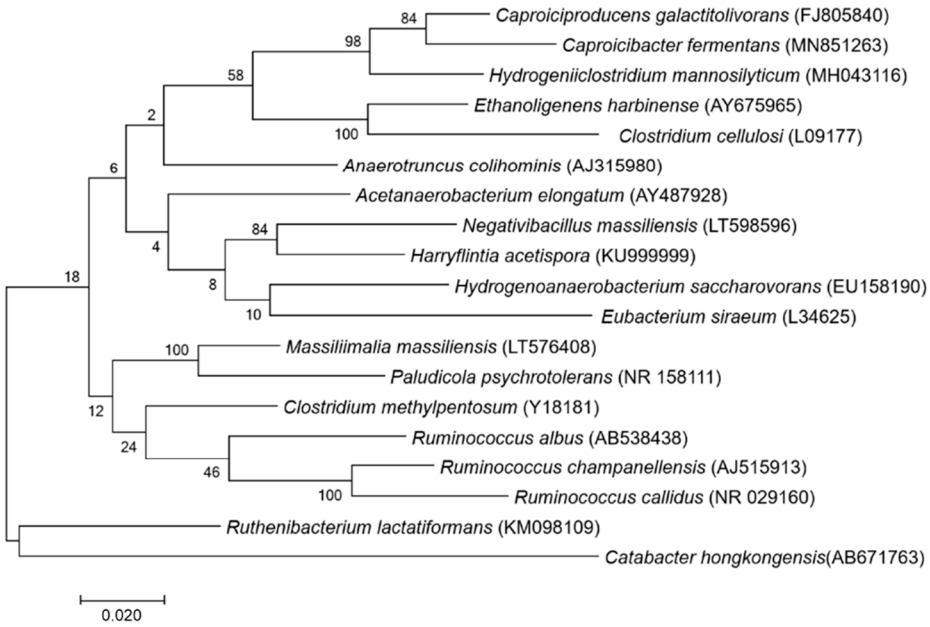

Figure 1. Phylogenetic tree highlighting the position of *Negativibacillus massiliensis* strain marseille-P3213 relative to other closely-related strains.

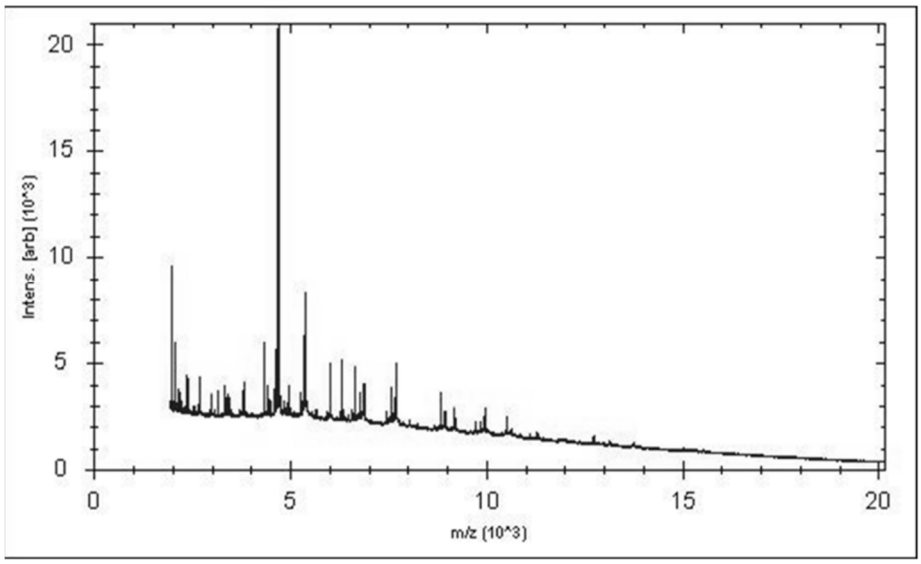

Figure 2. Reference MALDI-TOF spectrum of *Negativibacillus massiliensis* strain marseille-P3213 (=CSUR P3213 = DSM 103594). The spectra from 12 individual colonies were compared, and the reference spectrum was generated.

3.2. Phenotypic Description

3.2.1. Morphologic and Biochemical Characteristics

Strain P3213 is a Gram-stain negative, non-motile, spore-forming bacillus (Figure 3). The negative staining also showed a bacillus with a diameter ranging from 0.5 to 0.8 µm, and a length ranging from 3.0 to 4.5 µm (Figure 4). The catalase and oxidase tests were negative. The major cellular fatty acids were hexadecanoic acid (50%) and tetradecanoic acid (21%), both of which are saturated fatty acids (Table 1).

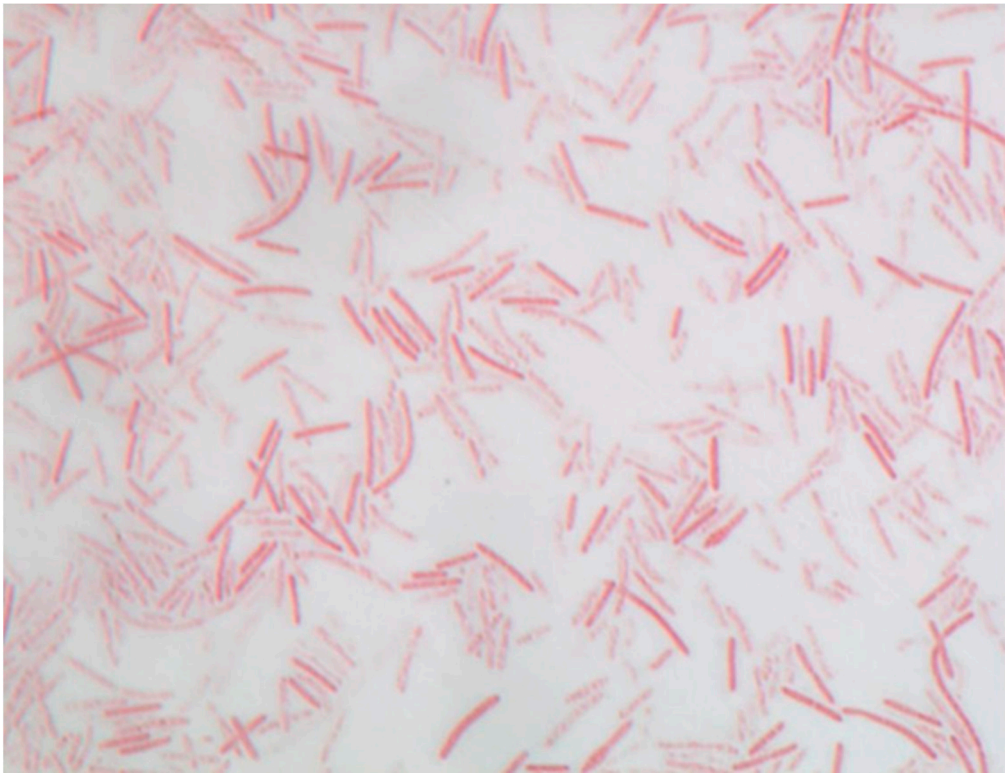

Figure 3. Gram staining of *Negativibacillus massiliensis* strain P3213 (=CSUR P3213, =DSM 103594) at a 100× magnification.

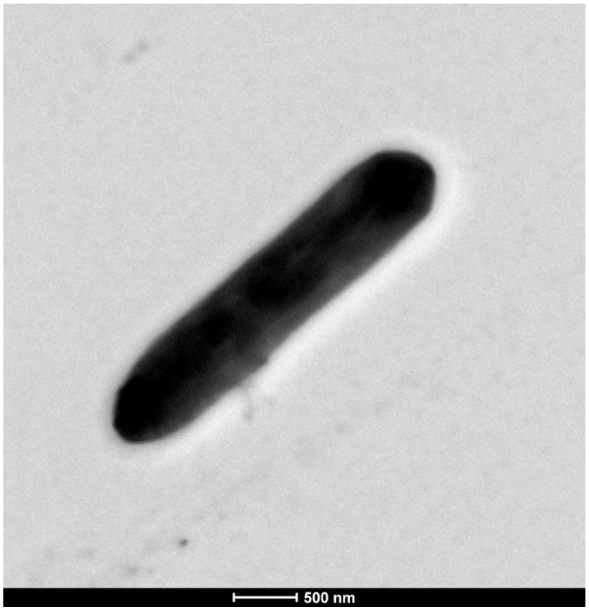

Figure 4. Transmission electron microscopy of *Negativibacillus massiliensis* strain P3213 (=CSUR P3213, = DSM 103594) using a Tecnai G20 transmission electron microscope (FEI Company) at an operating voltage of 60 kV. Scale bar = 500 nm.

Table 1. Cellular fatty acid composition (%) of *Negativibacillus massiliensis* P3213.

Fatty Acids	Name	Mean Relative % [(a)]
16:0	Hexadecanoic acid	49.8 ± 0.5
14:0	Tetradecanoic acid	21.0 ± 0.8
18:1n9	9-Octadecenoic acid	9.0 ± 0.3
18:2n6	9,12-Octadecadienoic acid	8.4 ± 0.1
18:0	Octadecanoic acid	7.4 ± 0.1
12:0	Dodecanoic acid	1.6 ± 0.2
18:1n7	11-Octadecenoic acid	1.5 ± 0.1
15:0	Pentadecanoic acid	TR
17:0	Heptadecanoic acid	TR
16:1n7	9-Hexadecenoic acid	TR
15:0 iso	13-methyl-tetradecanoic acid	TR

[a] mean peak area percentage. TR = trace amounts < 1%.

Strain P3213 was able to metabolize D-arabinose, D-ribose, D-glucose, dulcitol, D-mannitol, arbutine, D-lactose, D-trehalose, xylitol, gentiobiose, D-lyxose, D-tagatose, D-arabitol, L-arabitol and potassium 5-ketogluconate. Strain P3213 exhibited the following activities: alkaline phospatase, leucine arylamidase, alpha-chymotrypsin, acid phosphatase, naphthol phosphohydrolase, alpha-galactosidase, and beta-galactosidase, but not C4 esterase, C8 esterase lipase, C14 lipase, valine arylamidase, cystine arylamidase, trypsin, beta-glucuronidase, alpha-glucuronidase, beta-glucosidase, N-acetyl-beta-glucosaminidase, alpha-mannosidase, or alpha-fucosidase. No indole formation or urease activity was observed for strain P3213.

3.2.2. Optimal Growth Conditions

Growth was observed after 48 h at 37 °C and 45 °C, but not at 25, 28 or 56 °C. This strain is strictly anaerobic, as no growth was observed aerobically or microaerobically. The optimal growth was obtained at 37 °C after 48 h. Strain P3213 was able to grow at all of the tested pH, ranging from 5 to 8, with an optimum pH level at 7.5. Conversely, no growth was observed at the NaCl concentrations tested. However, strain P3213 was able to grow with a NaCl concentration at 10 g/L contained in the modified Columbia agar medium used to assess its pH tolerance.

3.2.3. Antibiotic Susceptivity

Strain P3213 was susceptible to cefotaxime 50 µg, but resistant to vancomycin 30 µg, cefalexin 15 µg, amoxicillin 30 µg, amoxicillin and clavulanic acid combination 20–10 µg, penicillin G 1UI, daptomycin 15 µg, metronidazole 4 µg, trimethoprim sulfamethoxazole 1.25–23.75 µg, oxacillin 5 µg, imipenem 10 µg, ceftriaxone 10 µg, rifampicin 30 µg, doxycycline 30 µg, erythromycin 15 µg, tobramycin 10 µg and fosfomycin 50 µg.

3.3. Genomic Characteristics

3.3.1. Genome Properties

The genome of strain P3213 is 2,876,881 bp long, with a 45.41% GC content. It consists of 8 scaffolds (consisting of eight contigs). Of the 2779 predicted genes, 2716 were protein-coding genes, and 63 were RNAs (five copies of 5S rRNA, one copy of 16S rRNA, one copy of 23S rRNA, and 56 tRNA genes). A total of 2453 genes (90.32%) were assigned as a putative function (by cogs or by NR blast). The comparison with the Cluster of Orthologous Groups (COGs) database allowed us to assign a function to 1827 predicted proteins (67.27%), with transcription [K] (172 proteins, 8.56%) and translation [I] (159 proteins, 7.91%) being the most represented functions. Moreover, 9.66% of the proteins, i.e., 194, were assigned to an unknown function [S]. In total, 21 genes were identified as ORFans (0.77%). The remaining 148 genes (5.45%) were annotated as hypothetical proteins (Figure 5, Table 2).

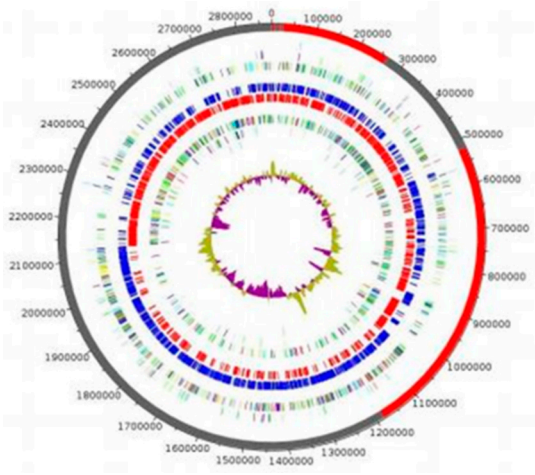

Figure 5. Graphical circular map of the genome. From the outer to the inner circles: Contigs (red/grey), COG category on the forward strand (three circles: forward strand (blue circle), and reverse strand (red circle), GC content).

Table 2. Nucleotide content and gene count levels of the genome.

Attribute	Genome (Total)	
	Value	**% of Total**
Size	2,876,881	100
G + C content (%)	1,306,499	45.41
Coding region (bp)	2,512,042	87.31
Total genes	2779	100
RNA genes	63	2.27
Protein-coding genes	2716	97.73
Genes with function prediction	2453	90.32
Genes assigned to COGs	1827	65.74
Genes with peptide signals	345	12.70
Genes with transmembrane helices	631	23.23
CRISPR repeats		
ORFans genes	21	0.77
Genes associated with PKS or NRPS	4	0.15
N° of antibiotic resistance genes	2	0.074

3.3.2. Genomic Comparison

Strain P3213 had a smaller draft genome sequence than those of Harryflintia acetispora, Ruminococcus callidus, Hydrogenoanaerobacterium saccharovorans, Clostridium methylpentosum, Anaerotruncus colihominis, and Ruminococcus albus (2.930, 3.1, 3.180, 3.48, 3.720, and 3.85 MB, respectively), but larger than those of Massiliimalia massiliensis, Caproiciproducens galactitovorans, and Ruminococcus champanellensis (2.84, 2.58 and 2.570 MB respectively) (Table 3). The G + C content of strain P3213 is smaller than those of Harryflintia acetispora, Anaerotruncus colohominis, Ruminococcus champanellensis, Clostridium methylpentosum, Ruminococcus callidus, Caproiciproducens galactitovorans and Massiliimalia massiliensis, (60%, 54.2%, 53.3%, 50.7%, 49%, 45.4% 48.1%, and 47.2%, respectively), but larger than those of Hydrogenoanaerobacterium saccharovorans and Ruminococcus albus (42.7% and 45.2% respectively). The repartition between the COG categories was similar for all of the compared genomes, except for the fewer carbohydrate transport and metabolism-related proteins for strain P3213 (Figure 6, Table 4). The degree of genomic similarity of strain P3213 to its closely related species with an available genome (Table 3, Table S1) was estimated using the OrthoANI software [33]. The highest OrthoANI value was observed between Ruminococcus champanellensis and Ruminococcus callidus (69.64%), while the lowest value was 63.91%, observed between Anaerotruncus colihominis

and Ruminococcus albus. The genome of strain P3213 exhibited the highest OrthoANI value, 66.41%, with Massiliimalia massiliensis (Figure 7). The degree of genomic similarity of strain P3213 with closely-related species was also calculated using dDDH. The values among the compared genomes (Table 5) ranged from 17.5% between Acetanaerobacterium elongatum and Caproiciproducens galactitolivorans, to 35.9% between Anaerotruncus colihominis and Ruminococcus champanellensis. For strain P3213, the highest value observed was 35.6%, with Harryflintia acetispora, while the lowest was 24%, with Acetanaerobacterium elongatum. Moreover, the TYGS server clearly highlights strain P3213 as a distinct and novel genus within the Ruminococcaceae family (Figure S1).

Table 3. Genomes used for the genomic comparison, including the closest species to *Negativibacillus massiliensis* CSUR P3213 within the available genomes.

Name of Organisms	RefSeq	Size (Mb)	G + C (%)	Protein-Coding Genes	Total Genes (ORFs)
Anaerotruncus colihominis	NZ_DS544194	3.72	54.2	3525	3704
Ruminococcus champanellensis	NC_021039	2.57	53.3	2276	2383
Harryflintia acetispora	NZ_SLUK01000001	2.93	60.0	2704	2786
Negativibacillus massiliensis	NZ_FTRU01000008	2.88	45.4	268	2781
Hydrogenoanaerobcaterium saccharovorans	NZ_RKRD01000001	3.18	42.7	2906	2994
Caproiciproducens galactitovorans	NZ_SRMQ01000001	2.58	48.1	2413	2535
Ruminococcus albus	NZ_FOAT01000022	3.85	45.2	3334	3484
Clostridium methylpentosum	PRJNA30029	3.48	50.7	3907	3964
Massiliimalia massiliensis	NZ_FUHT00000000.1	2.84	47.2	2681	2769
Ruminococcus callidus	PRJNA18185	3.1	49.0	2719	2866

Table 4. Number of genes associated with the 25 general COG functional categories.

		Negativibacillus massiliensis	
Code	**Description**	**Value**	**% of Total**
[A]	Rna processing and modification	0	0.00
[B]	Chromatin structure and dynamics	0	0.00
[C]	Energy production and conversion	91	4.53
[D]	Cell cycle control. mitosis and meiosis	26	1.29
[E]	Amino acid transport and metabolism	156	7.77
[F]	Nucleotide transport and metabolism	58	2.89
[G]	Carbohydrate transport and metabolism	89	4.43
[H]	Coenzyme transport and metabolism	59	2.94
[I]	Lipid transport and metabolism	33	1.64
[J]	Translation	159	7.91
[K]	Transcription	172	8.56
[L]	Replication. recombination and repair	126	6.27
[M]	Cell wall/membrane biogenesis	94	4.68
[N]	Cell motility	2	0.10
[O]	Posttanslational modification. protein turnover.chaperones	51	2.54
[P]	Inorganic ion transport and metabolism	107	5.33
[Q]	Secondary metabolites biosynthesis. transport and catabolism	10	0.50
[R]	General function prediction only	232	11.55
[T]	Signal transduction mechanisms	49	2.44
[U]	Intracellular trafficking and secretion	24	1.19
[V]	Defense mechanisms	95	4.73
[W]	Extracellular structures	0	0.00
[Y]	Nuclear structure	0	0.00
[Z]	Cytoskeleton	0	0.00
[S]	Function unknown	194	9.66
_	Not in COGs	182	9.06

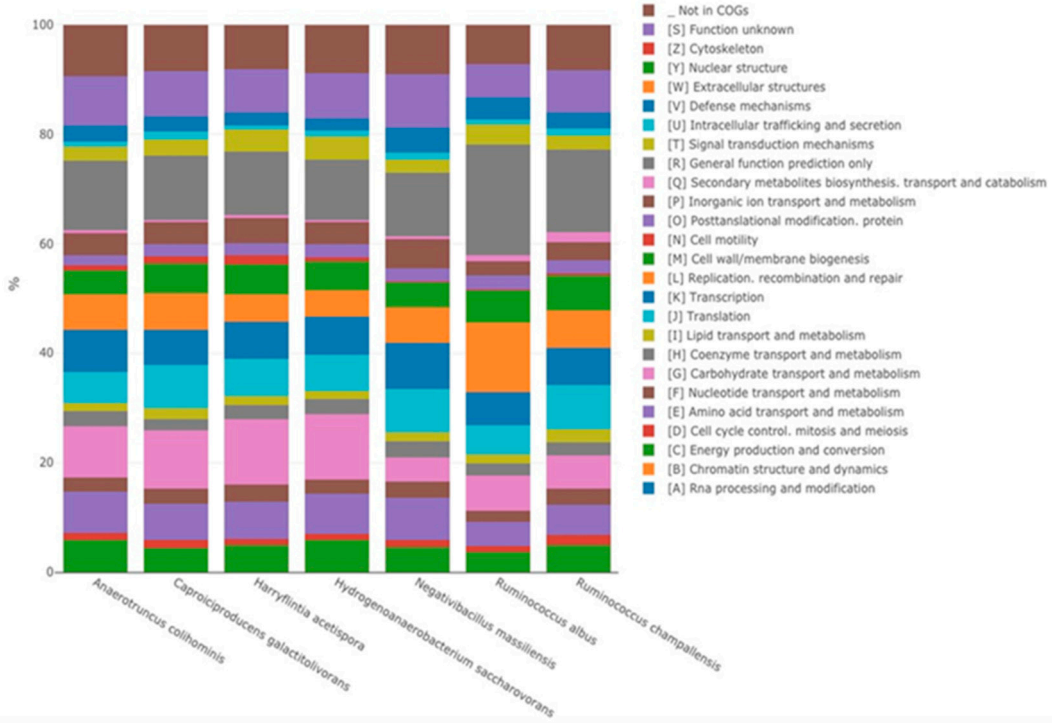

Figure 6. Distribution of functional classes of the predicted genes according to the clusters of orthologous groups of proteins.

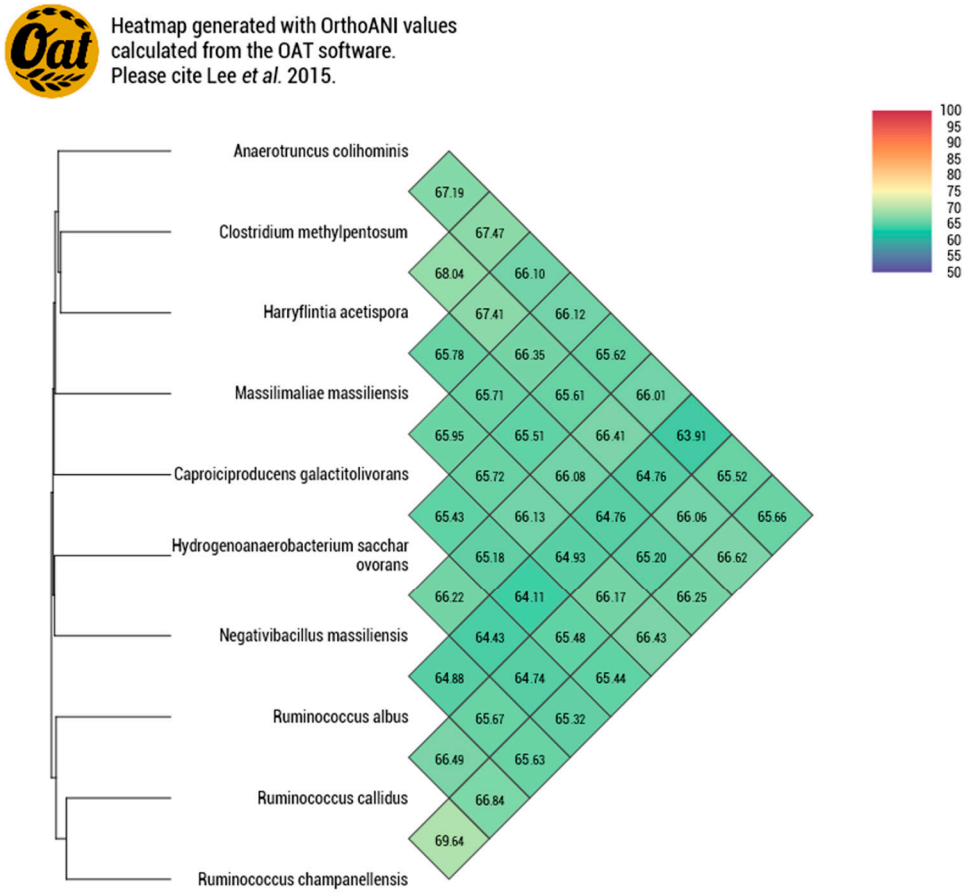

Figure 7. Heatmap generated with the OrthoANI values, calculated using the OAT software, between *Negativibacillus massiliensis* and other closely-related species with standing in the nomenclature.

Table 5. dDDH values obtained by comparing all of the studied genomes.

	Acetanaerobacterium elongatum	*Anaerotruncus colihominis*	*Caproiciproducens galactitolivorans*	*Clostridium methylpentosum*	*Ethanoligenens harbinense*	*Harryflintia acetispora*	*Hydrogenoanaerobacterium saccharovorans*	*Massilimaliae massiliensis*	*Negativibacillus massiliensis*	*Ruminococcus albus*	*Ruminococcus callidus*	*Ruminococcus champanellensis*
Acetanaerobacterium elongatum		20.60%	23%	17.50%	20%	18.70%	19.40%	37.50%	24%	26.60%	29.40%	33.80%
Anaerotruncus colihominis			35%	21.10%	19.60%	22.60%	27.20%	33.90%	32.30%	28.80%	26.80%	35.90%
Caproiciproducens galactitolivorans				25.10%	23.60%	30.20%	27.80%	32.90%	28.60%	26%	31.30%	27.70%
Clostridium methylpentosum					18.20%	20.30%	21.10%	22.90%	27.80%	25.80%	24%	28.10%
Ethanoligenens harbinense						18.80%	25.50%	31.40%	27.90%	27.90%	29.70%	30.40%
Harryflintia acetispora							24.20%	24.50%	35.60%	29.30%	23.40%	20.2%
Hydrogenoanaerobacterium saccharovorans								23.70%	26.20%	27.50%	21.90%	35.50%
Massilimaliae massiliensis									25.20%	34.90%	20.20%	29.70%
Negativibacillus massiliensis										24.30%	25.50%	24.40%
Ruminococcus albus											22.70%	24.50%
Ruminococcus callidus												20.90%
Ruminococcus champanellensis												

4. Conclusions

Strain Marseille-P3213^{T} exhibits a 16S rRNA sequence divergence under 5% with its phylogenetically closest genus with standing in the nomenclature. In addition, the highest OrthoANI and dDDH values observed were well under 95% and 70%, respectively, with the closest species with a validly-published name. We consequently suggest the creation of a new genus *Negativibacillus* gen. nov., type species *Negativibacillus massiliensis* sp. nov., type strain 3213 within the *Ruminococcaceae* family.

Description of *Negativibacillus* gen. nov.: Ne.ga.ti.vi.ba.cil'lus, L. adj. *negativus,* 'negative'; L. masc. n. *bacillus*, 'a small staff'; N.L. masc. n. named for the Gram-stain negative cell wall structure of the hereby-described bacilli.

The cells are strictly anaerobic, Gram-stain negative, spore-forming and non-motile bacilli.

Description of *Negativibacillus massiliensis* sp. Nov: mas.si.li.en'sis, L. masc. adj. massiliensis from Massilia, the Roman name of Marseille.

The cells have a diameter ranging from 0.5 to 0.8 μm, and a length ranging from 3.0 to 4.5 μm. The catalase and oxidase tests were negative. The major cellular fatty acids were hexadecanoic acid and tetradecanoic acid

Strain P3213 was able to metabolize D-arabinose, D-ribose, D-glucose, dulcitol, D-mannitol, arbutine, D-lactose, D-trehalose, xylitol, gentiobiose, D-lyxose, D-tagatose, D-arabitol, L-arabitol, and potassium 5-ketogluconate. Strain P3213 exhibited the following activities: alkaline phospatase, leucine arylamidase, alpha-chymotrypsin, acid phosphatase, naphthol phosphohydrolase, alpha-galactosidase, and beta-galactosidase, but not C4 esterase, C8 esterase lipase, C14 lipase, valine arylamidase, cystine arylamidase, trypsin, beta-glucuronidase, alpha-glucuronidase, beta-glucosidase, N-acetyl-beta-glucosaminidase, alpha-mannosidase, or alpha-fucosidase. No indole formation or urease activity was observed for strain P3213. Strain P3213 was susceptible to cefotaxime, but resistant to vancomycin, cefalexin, amoxicillin and clavulanic acid, penicillin G, daptomycin, metronidazole, trimethoprim sulfamethoxazole, oxacillin, imipenem, ceftriaxone, rifampicin, doxycycline, erythromycin, tobramycin, fosfomycin, and amoxicillin.

The genome of strain P3213 is 2,876,881 bp long, with a 45.41% GC content, and is accessible under FTRU00000000 in the GenBank collection. The 16S rRNA sequence is also accessible in the GenBank collection under accession number NR_147378. The type

strain Marseille-P3213 (=CSUR P3213 = DSM 103594) was isolated from the left colon of a French woman.

4.1. Nucleotide Sequence Accession Number

The 16S rRNA gene and genome sequences were deposited in Genbank under accession numbers NR_147378 and FTRU00000000, respectively.

4.2. Deposit in Culture Collections

Strain Marseille-P5551^T was deposited in the Collection de Souches de l'Unité des Rickettsies under the number CSUR P3213, and in the DSMZ collection under the number DSM 103594.

Supplementary Materials: The following are available online at https://www.mdpi.com/2036-7481/12/1/4/s1, Figure S1: Phylogenetic tree based on the genomic sequences of closely related species with available genomes., Table S1: List of type species of validly published genera within the *Ruminococcaceae* family and their 16S similarity with *Negativibacillus massiliensis*.

Author Contributions: Formal analysis: C.V., N.A. Funding acquisition: D.R. (Didier Raoult) Investigation: C.V., M.M., D.R. (Davide Ricaboni) and N.A. Methodology: D.R. (Didier Raoult) Resources: S.A., V.V. Supervision: V.V., J.-C.L., D.R. (Didier Raoult) and M.T.A. Validation: J.-C.L., D.R. (Didier Raoult) and M.T.A. Visualization: C.V. Writing—original draft: C.V., N.A. and M.T.A. Writing—Review and Editing: C.V., J.-C.L., D.R. (Didier Raoult) and M.T.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection and the National Research Agency under the program "Investissements d'avenir", reference ANR-10-IAHU-03.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of Institut Fédératif de Recherche IFR48 under number 09-022.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The 16S rRNA gene and genome sequences were deposited in Genbank under accession numbers NR_147378 and FTRU00000000, respectively. Strain Marseille-P5551^T was deposited in the Collection de Souches de l'Unité des Rickettsies under the number CSUR P3213, and in the DSMZ collection under the number DSM 103594.

Acknowledgments: We thank Aurelia Caputo and the genomics platform for sequencing and submitting the genome and 16S rRNA gene sequences. This work was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection and the National Research Agency under the program 'Investissements d'avenir', reference ANR-10-IAHU-03.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rajilić-Stojanović, M.; de Vos, W.M. The First 1000 Cultured Species of the Human Gastrointestinal Microbiota. *FEMS Microbiol. Rev.* **2014**, *38*, 996–1047. [CrossRef]
2. Lagier, J.-C.; Hugon, P.; Khelaifia, S.; Fournier, P.-E.; 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*, 237–264. [CrossRef] [PubMed]
3. Lagier, J.-C.; Million, M.; Hugon, P.; Armougom, F.; Raoult, D. Human Gut Microbiota: Repertoire and Variations. *Front. Cell. Infect. Microbiol.* **2012**, *2*. [CrossRef]
4. Tidjani Alou, M.; Naud, S.; Khelaifia, S.; Bonnet, M.; Lagier, J.-C.; Raoult, D. State of the Art in the Culture of the Human Microbiota: New Interests and Strategies. *Clin. Microbiol. Rev.* **2020**, *34*. [CrossRef] [PubMed]
5. Lagier, J.-C.; 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. [CrossRef] [PubMed]
6. Bilen, M.; Dufour, J.-C.; Lagier, J.-C.; Cadoret, F.; Daoud, Z.; Dubourg, G.; Raoult, D. The Contribution of Culturomics to the Repertoire of Isolated Human Bacterial and Archaeal Species. *Microbiome* **2018**, *6*, 94. [CrossRef] [PubMed]

7. Mailhe, M.; Ricaboni, D.; Vitton, V.; Gonzalez, J.-M.; Bachar, D.; Dubourg, G.; Cadoret, F.; Robert, C.; Delerce, J.; Levasseur, A.; et al. Repertoire of the Gut Microbiota from Stomach to Colon Using Culturomics and Next-Generation Sequencing. *BMC Microbiol.* **2018**, *18*, 157. [CrossRef]
8. Vos, P.; Garrity, G.; Jones, D.; Krieg, N.R.; Ludwig, W.; Rainey, F.A.; Schleifer, K.-H.; Whitman, W. Bergey's Manual of Systematic Bacteriology: Volume 3: The Firmicutes. In *Bergey's Manual of Systematic Bacteriology*, 2nd ed.; Springer: New York, NY, USA, 2009; ISBN 978-0-387-95041-9.
9. Parte, A.C.; Sardà Carbasse, J.; Meier-Kolthoff, J.P.; Reimer, L.C.; Göker, M. List of Prokaryotic Names with Standing in Nomenclature (LPSN) Moves to the DSMZ. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 5607–5612. [CrossRef]
10. Robert, C.; Bernalier-Donadille, A. The Cellulolytic Microflora of the Human Colon: Evidence of Microcrystalline Cellulose-Degrading Bacteria in Methane-Excreting Subjects. *FEMS Microbiol. Ecol.* **2003**, *46*, 81–89. [CrossRef]
11. Okeke, F.; Roland, B.C.; Mullin, G.E. The Role of the Gut Microbiome in the Pathogenesis and Treatment of Obesity. *Glob. Adv. Health Med.* **2014**, *3*, 44–57. [CrossRef]
12. Kim, M.; Oh, H.-S.; Park, S.-C.; 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*, 346–351. [CrossRef]
13. Fournier, P.-E.; Lagier, J.-C.; Dubourg, G.; Raoult, D. From Culturomics to Taxonomogenomics: A Need to Change the Taxonomy of Prokaryotes in Clinical Microbiology. *Anaerobe* **2015**, *36*, 73–78. [CrossRef]
14. Lagier, J.-C.; Khelaifia, S.; Alou, M.T.; Ndongo, S.; Dione, N.; Hugon, P.; Caputo, A.; Cadoret, F.; Traore, S.I.; Seck, E.H.; et al. Culture of Previously Uncultured Members of the Human Gut Microbiota by Culturomics. *Nat. Microbiol.* **2016**, *1*, 16203. [CrossRef] [PubMed]
15. Seng, P.; Abat, C.; Rolain, J.M.; Colson, P.; Lagier, J.-C.; Gouriet, F.; Fournier, P.E.; Drancourt, M.; La Scola, B.; Raoult, D. Identification of Rare Pathogenic Bacteria in a Clinical Microbiology Laboratory: Impact of Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *J. Clin. Microbiol.* **2013**, *51*, 2182–2194. [CrossRef] [PubMed]
16. Drancourt, M.; Bollet, C.; Carlioz, A.; Martelin, R.; Gayral, J.P.; Raoult, D. 16S Ribosomal DNA Sequence Analysis of a Large Collection of Environmental and Clinical Unidentifiable Bacterial Isolates. *J. Clin. Microbiol.* **2000**, *38*, 3623–3630. [CrossRef] [PubMed]
17. Stackebrandt, E.; Ebers, J. Taxonomic Parameters Revisited: Tarnished Gold Standards. *Microbiol. Today* **2006**, *33*, 152.
18. Brown, D.F.J.; Brown, L. Evaluation of the E Test, a Novel Method of Quantifying Antimicrobial Activity. *J. Antimicrob. Chemother.* **1991**, *27*, 185–190. [CrossRef]
19. Sasser, M. *Bacterial Identification by Gas Chromatographic Analysis of Fatty Acid Methyl Esters (GC-FAME)*; MIDI, Inc.: Newark, DE, USA, 1990; p. 6.
20. Dione, N.; Sankar, S.A.; Lagier, J.-C.; Khelaifia, S.; Michele, C.; Armstrong, N.; Richez, M.; Abrahão, J.; Raoult, D.; Fournier, P.-E. Genome Sequence and Description of *Anaerosalibacter massiliensis* sp. Nov. *New Microbes New Infect.* **2016**, *10*, 66–76. [CrossRef]
21. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. *BMC Bioinform.* **2010**, *11*, 119. [CrossRef]
22. Lowe, T.M.; Eddy, S.R. TRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964. [CrossRef]
23. Lagesen, K.; Hallin, P.; Rødland, E.A.; Stærfeldt, H.-H.; Rognes, T.; Ussery, D.W. RNAmmer: Consistent and Rapid Annotation of Ribosomal RNA Genes. *Nucleic Acids Res.* **2007**, *35*, 3100–3108. [CrossRef]
24. Käll, L.; Krogh, A.; Sonnhammer, E.L.L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *J. Mol. Biol.* **2004**, *338*, 1027–1036. [CrossRef] [PubMed]
25. Gouret, P.; Thompson, J.D.; Pontarotti, P. PhyloPattern: Regular Expressions to Identify Complex Patterns in Phylogenetic Trees. *BMC Bioinform.* **2009**, *10*, 298. [CrossRef]
26. Lechner, M.; Findeiß, S.; Steiner, L.; Marz, M.; Stadler, P.F.; Prohaska, S.J. Proteinortho: Detection of (Co-)Orthologs in Large-Scale Analysis. *BMC Bioinform.* **2011**, *12*, 124. [CrossRef]
27. Auch, A.F.; von Jan, M.; Klenk, H.-P.; Göker, M. Digital DNA-DNA Hybridization for Microbial Species Delineation by Means of Genome-to-Genome Sequence Comparison. *Stand Genom. Sci.* **2010**, *2*, 117–134. [CrossRef]
28. Ramasamy, D.; Mishra, A.K.; Lagier, J.-C.; Padhmanabhan, R.; Rossi, M.; Sentausa, E.; Raoult, D.; Fournier, P.-E. A Polyphasic Strategy Incorporating Genomic Data for the Taxonomic Description of Novel Bacterial Species. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 384–391. [CrossRef] [PubMed]
29. Meier-Kolthoff, J.P.; Auch, A.F.; Klenk, H.-P.; Göker, M. Genome Sequence-Based Species Delimitation with Confidence Intervals and Improved Distance Functions. *BMC Bioinform.* **2013**, *14*, 60. [CrossRef]
30. Meier-Kolthoff, J.P.; Göker, M. TYGS Is an Automated High-Throughput Platform for State-of-the-Art Genome-Based Taxonomy. *Nat. Commun.* **2019**, *10*. [CrossRef]
31. Integration of Evolutionary Biology Concepts for Functional Annotation and Automation of Complex Research in Evolution: The Multi-Agent Software System DAGOBAH | SpringerLink. Available online: https://link.springer.com/chapter/10.1007/978-3-642-20763-1_5 (accessed on 9 December 2020).

32. Gouret, P.; Vitiello, V.; Balandraud, N.; Gilles, A.; Pontarotti, P.; Danchin, E.G. FIGENIX: Intelligent automation of genomic annotation: Expertise integration in a new software platform. *BMC Bioinform.* **2005**, *6*, 198. [CrossRef] [PubMed]
33. Lee, I.; Ouk Kim, Y.; Park, S.-C.; Chun, J. OrthoANI: An Improved Algorithm and Software for Calculating Average Nucleotide Identity. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 1100–1103. [CrossRef] [PubMed]