**Rubeoparvulum massiliense gen. nov., sp. nov., a new bacterial genus isolated from the human gut of a Senegalese infant with severe acute malnutrition**

M. Tidjani Alou¹, J. Rathored¹, J.-C. Lagier¹,²,³, S. Khelaïfa¹, C. Michelle¹, C. Sokhna¹, A. Diallo³, A. B. Diallo⁴, P.-E. Fournier¹, D. Raoult¹⁵ and S. Edouard¹

¹) Aix-Marseille Université, URMITE, UM63, CNRS7278, INSERM 1095, Faculté de médecine, 2) Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes IRD 198, CNRS 7278, Aix-Marseille Université, Marseille, France, 3) Campus Commun UCAD-IRD of Hann, Dakar, Senegal, 4) Laboratoire de microbiologie, département de biologie, Université Abdou Moumouni de Niamey, Niamey, Niger and 5) Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

**Abstract**

*R. massiliense* strain mt6T was isolated from the gut microbiota of a severely malnourished boy from Senegal and consisted of facultative anaerobic, spore-forming, nonmotile and Gram-negative rods. *R. massiliense* showed a 92% similarity with the 16S rRNA of *Bacillus mannanlyticus*. The genome of strain mt6T is 2,843,796 bp long with a 43.75% G+C content. It contains 2,735 protein-coding genes and 76 RNA genes, among which are nine rRNA genes.

© 2016 Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases.

**Keywords:** Culturomics, genome, gut microbiota, *Rubeoparvulum massiliense*, taxonogenomics

**Original Submission:** 18 October 2016; **Accepted:** 9 November 2016

**Article published online:** 17 November 2016

**Corresponding author:** S. Edouard, Aix-Marseille Université, URMITE, UM63, CNRS7278, INSERM 1095, Faculté de médecine, 27 Boulevard Jean Moulin, 13385 Marseille cedex 05, France
E-mail: sophie.edouard@univ-amu.fr

**Introduction**

The human microbiome is defined as the sum of all microbes colonizing the human body [1]. The gut microbiota is one of the largest microbial ecosystems of the human body, consisting of $10^{14}$ microbial cells with a microbiome 150 times larger than the human genome [2]. The gastrointestinal microbiota colonization starts before birth with the maternal microbiota, and its early composition is influenced by the mode of birth. Its composition matures rapidly for the first year and reaches adult form by 3 years [2,3]. A disruption of its equilibrium has been proven to be implicated in a growing number of pathologies such as inflammatory bowel disease, irritable bowel syndrome, obesity [3,4] and severe acute malnutrition [5–7].

A new cultural approach, microbial culturomics, based on the multiplication of culture conditions with a variation of temperature, media and atmosphere, was developed in our laboratory in order to explore as exhaustively as possible a microbial ecosystem [8,9]. Using this new approach, we isolated a new member of the *Bacillaceae* family. At this time, 52 validated genera are part of the *Bacillaceae* family, which was created in 1895 by Fisher; *Bacillus* is its type genus, described by Cohn in 1872 [10]. Most species of this family are found in the environment (soil, water and plants) and are opportunistic pathogens in humans, except *Bacillus anthracis*, which is well known as being highly pathogenic. The *Bacillaceae* family includes Gram-positive, rod-shaped, mostly aerobic and facultative anaerobic genera [11]. By adding the description of the assembled and annotated genome of the species and the proteomic description of the strain with the matrix-assisted desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) profile to the classical description principles (phylogenetic relationships based on the 16S rRNA sequence,
phenotypic and genotypic characteristics), a new concept of description called taxonogenomics was developed in our laboratory [12].

Here we describe the genus *Rubeoparvulum*, the type species of which is *Rubeoparvulum massilense* strain mt6 (= CSUR P1473 = DSM 100479) from a stool sample collected in a 2-month-old infant living in Senegal and presenting with kwashiorkor, a type of severe acute malnutrition.

**Materials and Methods**

**Ethics and sample collection**

The strain mt6 was isolated from a stool taken from a severely malnourished 2-month-old boy with a height-for-age score of −5.87 who had nutritional edema. Collection was performed in Senegal in April 2014. This sampling was undertaken as part of an exploratory study of the human gut microbiota in African children with malnutrition. The study was approved by the local IFR 48 ethics committee under agreement 09-022. The boy’s parents provided informed consent. The sample was stored at −80°C after collection.

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

In order to explore as exhaustively as possible the bacterial diversity of the faecal sample, the culturomics concept was used to culture this sample using 18 culture conditions [8]. The obtained 16S rRNA sequence was deposited in GenBank (http://blast.ncbi.nlm.nih.gov/blast.cgi) to determine the percentage of sequence similarity with the closest bacteria. A new species or genus was defined by a similarity level of the 16S rRNA sequence under 98.65% or 95% respectively [15].

**Growth conditions**

The ideal growth conditions of strain mt6 were determined by testing different culture conditions. Five growth temperatures (25, 30, 37, 45 and 56°C) were tested under anaerobic and microaerophilic atmospheres using GENbag anaer and GENbag microer systems respectively (bioMérieux, Marcy l’Étoile, France). Aerobic growth was tested with and without 5% CO2. Growth was also tested at various pHs (6, 6.5, 7, 7.5, 8 and 8.5) using a pH-adjusted Colombia agar (bioMérieux). Salt tolerance was also tested with 0.5, 1, 5, 7.5 and 10% (w/v) NaCl.

**Morphologic, biochemical and antibiotic susceptibility tests**

Phenotypic characteristics (Gram staining, sporulation, motility) were determined as previously described [8]. The catalase (bioMérieux) and oxidase (Becton Dickinson, Le Pont de Claix, France) activities were also tested. Cell morphology was observed after negative staining of bacteria using a Tecnai G20 transmission electron microscope (FEI Company, Limel-Brevannes, France). The biochemical features of strain mt6 were investigated with API 50CH, API ZYM and API 20A strips (bioMérieux) according to the manufacturer’s instructions. Cellular fatty acid methyl ester (FAME) analysis was performed by gas chromatography mass spectrometry (GC/MS). Strain mt6 was grown on 5% sheep’s blood–enriched Colombia agar (bioMérieux) for the fatty acid analysis, which was carried out by GC/MS. Approximately 67 mg of bacterial biomass was each collected from several culture plates. Cellular FAMEs were prepared as described by Sasser (http://www.midi-inc.com/pdf/MIS_Technote_101.pdf).

Briefly, GC/MS analyses were realized by a Clarus 500 gas chromatograph equipped with a SQ8S MS detector (Perkin Elmer, Courtaboeuf, France). A total of 2 μL of FAME extracts were volatized at 250°C (split 20 mL/min) in a Focus liner with wool and separated on an Elite-5MS column (30 m, 0.25 mm i.d., 0.25 mm film thickness) using a linear temperature gradient (70 to 290°C at 6°C/min), allowing the detection of three to C24 fatty acid methyl esters. Helium flowing at 1.2 mL/min was used as carrier gas. The MS inlet line was set at 250°C and EI source at 200°C. Full scan monitoring was performed from 45 to 500 m/z. All data were collected and processed using Turbomass 6.1 (Perkin Elmer). FAMEs were identified by a spectral database search using MS Search 2.0 operated with the Standard Reference Database 1A (National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA) and the FAMEs mass spectral database (Wiley, Chichester, UK). Retention time correlations with estimated nonpolar retention indexes from the NIST database were obtained using a 37-component FAME mix (Supelco; Sigma-Aldrich, Saint-Quentin Fallavier, France); FAME identifications were confirmed using this index. Antibiotic susceptibility testing was performed using a disk...
diffusion method according to European Committee on Antimicrobial Susceptibility Testing (EUCAST) 2015 recommendations [17]. Inhibition diameters were measured using the Scan1200 scanner (Interscience, Saint-Nom-La Bretêche, France).

**Genomic DNA (gDNA) preparation**

For gDNA preparation, *R. massiliense* strain mt6<sup>T</sup> was cultured on 5% sheep’s blood–enriched Columbia agar (bioMérieux) at 37°C aerobically. Bacteria grown on three petri dishes were resuspended in 4 × 100 μL of Tris-EDTA (TE) buffer. Then 200 μL of this suspension was diluted in 1 mL TE buffer for lysis treatment, which included a 30-minute incubation with 2.5 μg/μL lysozyme at 37°C, followed by an overnight incubation with 20 μg/μL protease K at 37°C. 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.

**Genome sequencing and assembly**

Using the mate-pair strategy, the gDNA of *R. massiliense* strain mt6<sup>T</sup> was sequenced on the MiSeq sequencer (Illumina, San Diego, CA, USA). The gDNA was barcoded in order to be mixed with 11 other projects with the Nextera Mate-Pair sample prep kit (Illumina). The mate-pair library was prepared with 1 μg of gDNA using the Nextera Mate-Pair Illumina guide, and the gDNA sample was simultaneously fragmented and tagged with a mate-pair junction adapter. The pattern of the fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 labchip. The DNA fragments ranged in size from 1 to 10 kb, with an optimal size at 4.08 kb. No size selection was performed, and only 464 ng of tagedmented fragments were circularized. The circularized DNA was mechanically sheared to small fragments with an optimal size at 569 bp on the Covaris S2 device in microtubes (Covaris, Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies), and the final library concentration was measured at 24.4 nmol/L. The libraries were normalized at 2 nM and pooled. After a denaturation step and dilution at 15 pM, the pool of libraries was loaded onto the reagent cartridge and then onto the instrument along with the flow cell. Automated cluster generation and sequencing run were performed in a single 39-hour run at 2 × 251 bp. Total information of 10.1 Gb was obtained from a 1189K/mm<sup>2</sup> cluster density with a cluster passing quality control filters of 99.1% (22 579 000 clusters). The reads obtained were trimmed; assembly was performed by CLC genomicsWB4 software.

**Genome annotation and comparison**

Open reading frames (ORFs) were predicted using Prodigal [18] with default parameters, but the predicted ORFs were excluded if they spanned a sequencing gap region. The predicted bacterial protein sequences were searched against the GenBank [19] and the Clusters of Orthologous Groups (COGs) databases using BLASTP (E value 1e-03, coverage 0.7 and identity percentage 30%). If no hit was found, it was searched against the NR database using BLASTP with an E value of 1e-03, a coverage of 0.7 and an identity percentage of 30%, and if the sequence length was smaller than 80 aa, we used an E value of 1e-05. The tRNAscanSE tool [20] was used to find tRNA genes, while ribosomal RNAs were found using RNAmmer [21]. Lipoprotein signal peptides and the number of transmembrane helices were predicted using Phobius [22]. Mobile genetic elements were predicted using PHAST [23] and RAST [24]. ORFans were identified if all the BLASTP performed did not give positive results (E value smaller than 1e-03 for ORFs with sequence size larger 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 studies to define ORFans. Artemis [25] and DNA Plotter [26] were used for data management and the visualization of genomic features respectively. The Mauve alignment tool (version 2.3.1) was used for multiple genomic sequence alignment [27].

Comparator species for genomic comparison were identified in the 16S RNA tree using Phylopattern software [28]. The genome of strain mt6<sup>T</sup> was compared to those of *Alkaliphilus metalliredigens* strain QYMF, *Clostridium acetici*um strain DSM 1496, *Alkaliphilus transvaalensis* strain SAGM1 and *Alkaliphilus oremlandii* strain OhILAs.

For each selected genome, the complete genome sequence, proteome genome sequence and Orfeome genome sequence were retrieved from the FTP of NCBI. An annotation of the entire proteome was performed to define the distribution of 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 processes were performed in the multiagent software system DAGOBH [29], which includes Figenix [30] libraries that provide pipeline analysis. To evaluate the genomic similarity between studied genomes, we determined two parameters, digital DNA-DNA hybridization (DH), which exhibits a high correlation with DDH [31,32] and average genomic identity of orthologous gene sequences (AGIOS) [33], which was designed to be independent from DDH [33]. The AGIOS score is the mean value of nucleotide similarity between all couples of orthologous proteins between the two studied genomes [33].

© 2016 Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, NMNI, 15, 49–60.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).
Results

Strain identification and phylogenetic analyses

The mt6 strain was isolated after a 30-day preincubation at 37°C in an anaerobic blood culture bottle supplemented with 5 mL of rumen fluid filter-sterilized through a 0.2 μm pore filter (Thermo Fisher Scientific, Villebon sur Yvette, France). Strain mt6T was subcultured on 5% blood-enriched Colombia agar (bioMérieux) in an anaerobic atmosphere at 37°C. The bacterium could not be identified by MALDI-TOF MS (score under 1.7), but sequencing of the 16S rRNA revealed that strain mt6T’s nucleotide sequence had a 92% similarity level with Bacillus mannanlyticus, the phylogenetically closest species with a validly published name (Fig. 1). According to Kim et al. [16], a new genus can be defined by a similarity level threshold lower than 95%. Consequently, strain mt6T was classified as a new genus called Rubeoparvulum, its type species being Rubeoparvulum massiliense (Table 1). The 16S rRNA sequence of strain mt6 was deposited in GenBank under accession number LN828926.

Phenotypic description

The growth of the mt6 stains occurred between 25 and 56°C on 5% sheep’s blood–enriched Colombia agar. Optimal growth was achieved at 37°C after 48 hours of incubation in both
anaerobic and microaerophilic conditions. The cell growth was weaker in aerobic conditions. Strain mt6 was able to grow at pH values ranging from 6 to 8.5 and 0.5 to 5% NaCl concentrations. Cells were spore forming and motile, and they formed translucent colonies with a mean diameter of 0.5 mm on blood-enriched Colombia agar. Microscopic observations showed Gram-stain-negative, rod-shaped cells (Fig. 2), and electron microscopy showed rods with a mean diameter of 1 μm and a mean length of 6.8 μm (Fig. 3). Our MALDI-TOF MS database was incremented with the reference spectrum obtained for strain mt6T (Fig. 4). Comparisons of the aforementioned spectrum to that of other known species of the Bacillaceae family are represented in the gel view (Fig. 5). Reference spectra are available in our online database (http://www.mediterranee-infection.com/article.php?laref=256&titre=urms-database).

Strain mt6T was negative for catalase activity and positive for oxidase activity. Using an API ZYM strip, positive reactions were recorded for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, α-chymotrypsin, acid phosphatase and naphthol-AS-BI-phosphohydrolase. Nitrate reduction was observed; urease, β-glucosidase and protease activities were positive using an API 20 NE strip. All other reactions were negative on both strips. An API 50CH was used to test the carbohydrates metabolism. The following carbohydrates were metabolized by strain mt6T: glycerol, d-lactose, d-fucose, d-mannose, d-cellobiose, d-maltose, salicin, d-arabitol, N-acetyl-glucosamine and potassium-5-ketogluconate. Amygdalin, arbutin, d-fructose, inulin, d-sucrose, d-raffinose, erythritol, d-arabinoose, d-ribose, d-xyllose, l-xyllose, d-adenitol, methyl-β-D-xyllopyranoside, d-glucose, d-galactose, l-sorbose, l-rhamnose, dulcitol, inositol, d-mannitol, d-sorbitol, methyl-α-D-mannopyranoside, methyl-α-D-glucopyranoside, esculin ferric citrate, d-melibiose, d-trehalose, d-melezitose, starch, glycogen, xylitol, gentiobiose, d-turanose, d-lyxose, d-tagatose, l-fucose, l-arabitol, potassium gluconate and potassium 2-ketogluconate showed negative reactions. Differences between the biochemical features of close members of the family Bacillaceae are listed in Table 2.

The major cellular fatty acids found for strain mt6 were 12-methyl-tetradecanoic acid (25%), 10-methyldecanoic acid (15%) and hexadecanoic acid (12%). This strain was composed of many branched structures (iso/anteiso). A specific 3-OH fatty acid was also described (<1%, Table 3).

Among tested antibiotics, cells were susceptible to amoxicillin, gentamicin, ceftriaxone, ciprofloxacin, penicillin,
FIG. 4. Reference mass spectrum from *Rubeoparvulum massiliense* strain mt6T (= CSUR P1473 = DSM 100479). Spectra from 12 individual colonies were compared and reference spectrum generated.

FIG. 5. Gel view comparing *Rubeoparvulum massiliense* strain mt6T (= CSUR P1473 = DSM 100479) to other species within *Bacillaceae* family. Gel view displays raw spectra of loaded spectrum files arranged in pseudo-gel-like look. x-axis records m/z value. Left y-axis displays running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Colour bar and right y-axis indicate relation between colour peak is displayed, with peak intensity in arbitrary units. Displayed species are indicated at left.

© 2016 Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, NMNI, 15, 49–60

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0).
TABLE 2. Differential characteristics of Rubeoparvulum massiliense strain mt6<sup>T</sup> CSUR P1473 = DSM 100479, Bacillus mannanlyticus strain AM-001<sup>T</sup> DSM 16130<sup>T</sup>, Tepidibacillus fermentans strain STGH<sup>T</sup> DSM 23802<sup>T</sup>, Pullulanibacillus uraniitolerans strain UG-2<sup>T</sup> DSM 19429<sup>T</sup>, Alkalibacillus haloalkaliphilus DSM 5271<sup>T</sup>, Tenuibacillus halotolerans strain YIM 94025<sup>T</sup> KCTC 33046<sup>T</sup>, Thalassobacillus devorans strain G-19.<sup>T</sup> DSM 16966<sup>T</sup>, Salinibacillus aidingensis strain 25-7<sup>T</sup> JCM 12389<sup>T</sup>, Salinibacillus kushneri strain 8-2<sup>T</sup> JCM 12390<sup>T</sup>, Ornithinibacillus bavariensis strain WSBC 24001<sup>T</sup> DSM 15681<sup>T</sup> [34–41]

| Property                  | Rubeoparvulum massiliense | Bacillus mannanlyticus | Tepidibacillus fermentans | Pullulanibacillus uraniitolerans | Alkalibacillus haloalkaliphilus | Tenuibacillus halotolerans | Thalassobacillus devorans | Salinibacillus kushneri | Salinibacillus aidingensis | Ornithinibacillus bavariensis |
|---------------------------|---------------------------|------------------------|---------------------------|----------------------------------|-------------------------------|---------------------------|---------------------------|-------------------------|-----------------------------|-----------------------------|
| Cell diameter (μm)        | 1.0                       | 0.6–0.8                | 0.3                       | 1.0                              | 0.3–0.5                       | 0.2–0.3                   | 1.0–1.2                   | 0.4–0.6                 | 0.3–0.5                     | 0.4                         |
| Oxygen requirement        | ++                        | +/−                    | +/−                       | +/−                               | +                             | +                         | +                         | +                       | +                          | +                           |
| Gram stain                | −/−                       | +/−                    | +/−                       | +/−                               | −/−                           | NA                        | NA                        | +/−                     | +/−                         | NA                          |
| Salt requirement          | −/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | +/−                         | −/−                         |
| Indole                    | −/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | −/−                     | −/−                         | −/−                         |
| Production of:            |                           |                        |                           |                                   |                               |                           |                           |                         |                             |                             |
| Catalase                  | −/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Oxidase                   | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| NitrOxidase reductase     | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Urease                    | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Acid from:                |                           |                        |                           |                                   |                               |                           |                           |                         |                             |                             |
| L-Arabinose               | −/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Ribose                    | −/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Mannose                   | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Maltose                   | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Sucrose                   | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| d-Glucose                 | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| d-Fructose                | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| d-Maltose                 | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| d-Lactose                 | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| d-Hexose                  | +/−                       | +/−                    | +/−                       | +/−                               | −/−                           | −/−                       | −/−                       | +/−                     | −/−                         | −/−                         |
| Habitats                  | Human stool               | Industry               | Gas storage               | Mill tailing effluent             | Salt lake                     | Salt lake                 | Hypersaline environments | Neutral saline lake       | Neutral saline lake         | Pasteurized milk            |

*, positive result; −, negative result; NA, data not available.
Table 3. Cellular fatty acid composition (%) of Rubeoparvulum massiliense strain mt6T

| Fatty acid | Name | Mean relative %a |
|------------|------|------------------|
| 15:0 anteiso | 12-Methyl-tetradecanoic acid | 25.2 ± 0.3 |
| 13:0 anteiso | 10-Methyl-dodecanoic acid | 15.2 ± 0.2 |
| 16:0 | Hexadecanoic acid | 12.0 ± 0.6 |
| 18:1n9 | 9-Octadecenoic acid | 9.3 ± 0.4 |
| 13:0 iso | 11-Methyl-tetradecanoic acid | 7.1 ± 0.1 |
| 18:0 | Octadecanoic acid | 6.3 ± 0.1 |
| 15:0 iso | 13-Methyl-tetradecanoic acid | 5.8 ± 0.1 |
| 14:0 iso | 12-Methyl-tridecanoic acid | 5.6 ± 0.3 |
| 18:2n6 | 9,12-Octadecadienoic acid | 5.3 ± 0.2 |
| 5:0 iso | 3-Methyl-butanoyl acid | 1.7 ± 0.1 |
| 14:0 | Tetradecanoic acid | 1.6 ± 0.1 |
| 11:0 anteiso | 8-Methyl-decanoic acid | 1.1 ± 0.1 |
| 18:1n6 | 12-Octadecenoic acid | TR |
| 15:0 | Pentadecanoic acid | TR |
| 12:0 iso | 10-Methyl-undecanoic acid | TR |
| 17:0 | Heptadecanoic acid | TR |
| 18:1n7 | 11-Octadecenoic acid | TR |
| 13:0 | Tridecanoic acid | TR |
| 15:0 3-OH | 2-Hydroxy-12-methyl-tetradecanoic acid | TR |
| anteiso | acid | TR |
| 11:0 iso | 9-Methyl-decanoic acid | TR |
| 16:0 iso | 14-Methyl-pentadecanoic acid | TR |
| 10:0 | Decanoic acid | TR |
| 16:1n7 | 9-Hexadecenoic acid | TR |
| 20:4n6 | 5,8,11,14-Eicosatetraenoic acid | TR |
| 17:0 anteiso | 14-Methyl-hexadecanoic acid | TR |
| 12:0 | Dodecanoic acid | TR |
| 17:0 iso | 15-Methyl-hexadecanoic acid | TR |

TR, trace amounts <1%. aMean peak area percentage.

Imipenem, tobramycin and oxacillin but were resistant to metronidazole, trimethoprim/sulfamethoxazole, rifampicin, doxycycline, vancomycin, nitrofurantoin and erythromycin.

**Genome properties**

With an estimated size of 2,843,796 bp, the R. massiliense genome had a G+C content of 43.75% (Table 4, Fig. 6). It was composed of six scaffolds composed of six contigs. Out of 2811 predicted genes, 2735 were protein-coding genes, and 76 were RNAs (seven 5S rRNA, one 16S rRNA gene, one 23S rRNA gene, 67 tRNA genes). A putative function was assigned to 1873 genes (66.63%) by COGs or NR blast. A total of 233 genes (8.28%) were identified as ORFans. The remaining 402 genes (14.70%) were annotated as hypothetical proteins. Table 4 shows the statistics of the genome, while Table 5 presents the distribution of genes into COGs functional categories.

**Genome comparison**

The genome of strain mt6T was compared to those of closely related species (Table 6) by comparing their main genomic characteristics (size, G+C content, protein-coding genes, total number of genes). The genome size of strain mt6T (2.84 Mb) is smaller than B. agri (5.51 Mb), B. borstelensis (5.16 Mb), B. mannanilyticus (4.53 Mb), B. thermoruber (4.43 Mb) and C. thermarum (2.9Mb). Strain mt6T had a higher G+C content (43.75%) than B. mannanilyticus (39.6%) but lower than B. thermoruber (58.4%), B. agri (54.2%), B. borstelensis (52%) and C. thermarum (47.6%). Strain mt6T has the smallest number of protein-coding genes as well as the smallest number of total genes than all of the other compared genomes, as summarized in Table 6.

 Among species with standing in nomenclature, AGIOS values ranged from 75.55 between B. borstelensis and B. thermoruber to 59.20 between B. thermoruber and B. mannanilyticus. The comparison of the AGIOS value of strain mt6T with the other species gave AGIOS values ranging from 60.30 with B. thermoruber to 63.12 with B. mannanilyticus (Tables 7 and 8). In addition, strain mt6T shared 1296, 1316, 1039, 1079 and 1605 orthologous genes with B. borstelensis, C. thermarum, B. mannanilyticus and B. agri respectively. Finally, we observed that in each COGs categories, all compared genomes have nearly the same number of genes (Fig. 7).

**Conclusion**

The proteomic analysis of strain mt6T with its MALDI-TOF MS spectrum, the 92% similarity level of the 16S rRNA nucleotide sequence to Bacillus mannanilyticus and the analysis of its complete assembled and annotated genome allowed us to propose the creation of a new genus called Rubeoparvulum gen. nov. Rubeoparvulum massiliense sp. nov. and strain mt6T are the type species and type strain respectively of Rubeoparvulum gen. nov.
**Description of Rubeoparvulum gen. nov.**

*Rubeoparvulum* (ru.be.o.o, adj. ‘red’; par.vu.lum, n. ‘infant’) strain mt6\textsuperscript{T} was isolated from the stool of a patient with kwashiorkor. The term ‘red infant’ refers to the hair discoloration observed in kwashiorkor patients.

Cells are rod-shaped, Gram-stain-negative bacteria. Optimal growth in anaerobic and microaerophilic conditions is at 37°C. The organism is catalase negative and oxidase positive; nitrate reduction, urease, β-glucosidase and alkaline phosphatase were positive. The type species is *Rubeoparvulum massiliense* strain mt6\textsuperscript{T}.

**Description of Rubeoparvulum massiliense strain mt6\textsuperscript{T} gen. nov., sp. nov.**

*Rubeoparvulum massiliense* (mas.si.li.en’se, L. adj. massiliense, of Massilia, the old Greek and Roman name of Marseille, France, where the strain was isolated).

Cells are spore-forming, motile and facultative anaerobe, Gram-stain-negative, rod-shaped bacilli with a mean diameter of 1 μm and a mean length of 6.8 μm. Colonies were small (mean diameter of 0.5 mm) and translucent on 5% sheep’s blood–

**TABLE 5. Number of genes associated with 25 general COGs functional categories**

| Code | Value | % of total | Description                                      |
|------|-------|------------|--------------------------------------------------|
| J    | 150   | 5.48       | Translation                                      |
| A    | 0     | 0          | RNA processing and modification                  |
| K    | 141   | 5.15       | Transcription                                    |
| L    | 139   | 5.08       | Replication, recombination and repair            |
| B    | 1     | 0.03       | Chromatin structure and dynamics                 |
| D    | 28    | 1.02       | Cell cycle control, mitosis and meiosis          |
| Y    | 0     | 0          | Nuclear structure                                |
| V    | 50    | 1.82       | Defense mechanisms                               |
| T    | 90    | 3.29       | Signal transduction mechanisms                   |
| M    | 82    | 2.99       | Cell wall/membrane biogenesis                    |
| N    | 49    | 1.79       | Cell motility                                    |
| Z    | 0     | 0          | Cytoskeleton                                     |
| W    | 0     | 0          | Extracellular structures                         |
| U    | 39    | 1.42       | Intracellular trafficking and secretion          |
| O    | 82    | 2.99       | Posttranslational modification, protein          |
|      |       |            | turnover, chaperones                             |
| C    | 129   | 4.71       | Energy production and conversion                 |
| G    | 81    | 2.96       | Carbohydrate transport and metabolism            |
| E    | 227   | 8.29       | Amino acid transport and metabolism              |
| F    | 62    | 2.36       | Nucleotide transport and metabolism              |
| H    | 76    | 2.77       | Coenzyme transport and metabolism                |
| I    | 66    | 2.41       | Lipid transport and metabolism                   |
| P    | 155   | 5.66       | Inorganic ion transport and metabolism           |
| Q    | 30    | 1.09       | Secondary metabolites biosynthesis, transport    |
| R    | 273   | 9.98       | General function prediction only                 |
| S    | 189   | 6.91       | Function unknown                                 |
| —    | 862   | 31.51      | Not in COGs                                     |

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

© 2016 Published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases. NMNI, 15, 49–60.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
enriched Colombia agar. The organism is catalase negative and oxidase positive. Positive reactions were recorded for alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, α-chymotrypsin, acid phosphatase and naphtol-AS-BI-phosphohydrolase. Urease, β-glucosidase, protease activities and nitrate reduction were also positive. Glycerol, D-lactose, D-fucose, D-mannose, D-cellobiose, D-maltose, salicin, N-acetylglucosamine, potassium-5-ketogluconate and D-arabitol were metabolized. Cells were susceptible to amoxicillin, gentamicin, ceftriaxone, ciprofloxacin, penicillin, imipenem, tobramycin and oxacillin but were resistant to metronidazole, trimethoprim/sulfamethoxazole, rifampicin, doxycycline, vancomycin, nitrofurantoin and erythromycin.

The G+C content of the genome is 43.75%. The 16S rRNA gene sequence and whole-genome shotgun sequence of *R. massiliense* strain mt6T are deposited in European Molecular Biology Laboratory/European Bioinformatics Institute under accession numbers LN828926 and CVPE00000000 respectively. The type strain mt6T (= CSUR P1473 = DSM 100479) was isolated from the faecal matter of a 2-month-old boy from Senegal with kwashiorkor.

### TABLE 6. Genome comparison of closely related species to *Rubeoparvulum massiliense* strain mt6T

| Organism                                      | INSDC          | Size (Mb) | G+C (%) | Protein-coding genes | Total genes |
|-----------------------------------------------|----------------|-----------|---------|----------------------|-------------|
| *Rubeoparvulum massiliense* strain mt6T       | CVPE0000000000 | 2.84      | 43.75   | 2735                 | 2811        |
| *Bacillus mannanilyticus* strain AM-001       | BAMO0000000000 | 4.23      | 39.6    | 3846                 | 4454        |
| *Brevibacillus agri* strain DSM 6348          | JATL0000000000 | 5.51      | 54.2    | 5047                 | 5297        |
| *Brevibacillus borstelensis* strain DSM 6347T | APBN0000000000 | 5.16      | 52.0    | 4817                 | 5039        |
| *Brevibacillus thermonubere* strain DSM 7064  | ATNE0000000000 | 4.43      | 58.4    | 4072                 | 4269        |
| *Caldalkalibacillus thermarum* strain HA6     | AFCE0000000000 | 2.9       | 47.6    | 2741                 | 2969        |

INSDC, International Nucleotide Sequence Database Collaboration.

### TABLE 7. Numbers of orthologous protein shared between genomes (upper right)*

|                   | *B. mannanilyticus* strain AM-001 | *B. agri* strain DSM 6348 | *B. borstelensis* strain DSM 6347T | *B. thermoruber* strain DSM 7064 | *C. thermarum* strain HA6 | *R. massiliense* strain mt6T |
|-------------------|-----------------------------------|---------------------------|-----------------------------------|----------------------------------|---------------------------|----------------------------|
| *B. mannanilyticus* strain AM-001 | 4842                             | 1605                      | 1606                              | 1533                             | 1178                      | 1079                      |
| *B. agri* strain DSM 6348 | 60.62                            | 5273                      | 2713                              | 2625                             | 1368                      | 1286                      |
| *B. borstelensis* strain DSM 6347T | 61.19                            | 73.07                     | 5019                              | 2645                             | 1376                      | 1316                      |
| *B. thermoruber* strain DSM 7064 | 59.20                            | 75.04                     | 75.55                             | 4253                             | 135                       | 1269                      |
| *C. thermarum* strain HA6 | 66.10                            | 62.97                     | 63.15                              | 63.13                             | 2986                      | 1039                      |
| *R. massiliense* strain mt6T | 63.12                            | 61.17                     | 61.57                             | 60.30                             | 62.87                      | 2733                      |

*Average percentage similarity of nucleotides corresponding to orthologous protein shared between genomes (lower left) and numbers of proteins per genome (bold).*

### TABLE 8. Pairwise comparison of *Rubeoparvulum massiliense* strain mt6T with other species using GGDC, formula 2 (DDH estimates based on identities/HSP length),* upper right

|                   | *B. mannanilyticus* | *B. agri* | *B. borstelensis* | *B. thermoruber* | *C. thermarum* | *R. massiliense* |
|-------------------|---------------------|-----------|-------------------|------------------|---------------|-----------------|
| *B. mannanilyticus* | 100% ± 00           | 32.9% ± 2.52 | 32.2% ± 2.52      | 31.6% ± 2.52     | 26.2% ± 2.52  | 26.3% ± 2.52    |
| *B. agri*         | 100% ± 00           | 18.8% ± 2.70 | 18.5% ± 2.70      | 19.4% ± 2.73     | 30% ± 2.52    | 35% ± 2.52      |
| *B. borstelensis* | 100% ± 00           | 20.2% ± 2.80 | 20.3% ± 2.80      | 20.5% ± 2.80     | 29.9% ± 2.52  | 30.7% ± 2.52    |
| *B. thermoruber*  | 100% ± 00           | 20.2% ± 2.80 | 20.3% ± 2.80      | 20.5% ± 2.80     | 29.9% ± 2.52  | 30.7% ± 2.52    |
| *C. thermarum*    | 100% ± 00           | 35.1% ± 2.52 | 35.1% ± 2.52      | 35.1% ± 2.52     | 35.1% ± 2.52  | 35.1% ± 2.52    |
| *R. massiliense*  | 100% ± 00           | 23.2% ± 2.52 | 23.2% ± 2.52      | 23.2% ± 2.52     | 23.2% ± 2.52  | 23.2% ± 2.52    |

*Confidence intervals indicate inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical test data sets (which are always limited in size). These results are in accordance with 16S rRNA (Fig. 1) and phylogenomic analyses as well as GGDC results.*
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 thank K. Griffiths for English-language review and C. Andrieu for administrative assistance.

Conflict of Interest

None declared.

References

[1] Fricke WF. The more the merrier! Reduced fecal microbiota diversity in preterm infants treated with antibiotics. J Pediatr 2014;165:8–10.
[2] Simpson HL, Campbell BJ. Review article: dietary fibre—microbiota interactions. Aliment Pharmacol Ther 2015;42:158–79.
[3] Gritz EC, Bhandari V. The human neonatal gut microbiome: a brief review. Front Pediatr 2015:3:60.
[4] Salazar N, Arboleya S, Valdés L, Stanton C, Ross P, Ruiz L, et al. The human intestinal microbiome at extreme ages of life. Dietary intervention as a way to counteract alterations. Front Genet 2014;5:406.
[5] Million M, Diallo A, Raoult D. Gut microbiota and malnutrition. Microb Pathog 2016.
[6] Million M, Tidjani Alou M, Khelaifi S, Bachar D, Lagier JC, Dione N, et al. Increased gut redox and depletion of anaerobic and methanogenic prokaryotes in severe acute malnutrition. Sci Rep 2016;6:26051.
[7] Tidjani Alou M, Lagier JC, Raoult D. Diet influence on the gut microbiota and dysbiosis related to nutritional disorders. Hum Microbiome J 2016:1:3–11.
[8] Lagier JC, Hugon P, Khelaifi S, Fournier PE, La Scola B, Raoult D. The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. Clin Microbiol Rev 2015;28:237–64.
[9] Lagier JC, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[10] Parte AC. LPSN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res 2014;42:D613–6.
[11] De Vos P, Garrity GM, Jones D, Krieg NR, Ludwig W, Rainey FA, et al., editors. Systematic bacteriology. New York: Springer; 2009.
[12] Fournier PE, Lagier JC, Dubourg G, Raoult D. From culturomics to taxonomogenomics: a need to change the taxonomy of prokaryotes in clinical microbiology. Anaerobe 2015;36:73–8.
[13] Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Infect Dis 2009;49:543–51.
[14] Seng P, Abat C, Rolain JM, Colson P, Lagier JC, Gouriet F, et al. 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;51:2182–94.
[15] Drancourt M, Bollet C, Carlino A, Martelin R, Gayral JP, Raoult D. 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 2000;38:3623–30.
[16] Kim M, Oh HS, Park SC, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 2014;64:346–51.
[17] Matuschek E, Brown DFJ, Kahlmeter G. Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect 2014;20:O255–66.

FIG. 7. Distribution of functional classes of predicted genes according COGs of protein. COGs, Clusters of Orthologous Groups database.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).