Antonie van Leeuwenhoek (2019) 112:935–945
https://doi.org/10.1007/s10482-019-01227-1

ORIGINAL PAPER

# *Dysgonomonas massiliensis* sp. nov., a new species isolated from the human gut and its taxonogenomic description

**Melhem Bilen · Maxime Descartes Mbogning Fonkou · Grégory Dubourg · Enora Tomei · Magali Richez · Jérémy Delerce · Anthony Levasseur · Ziad Daoud · Didier Raoult · Frédéric Cadoret**

Received: 1 October 2018 / Accepted: 5 January 2019 / Published online: 17 January 2019


**Abstract** Culturomics has allowed the isolation of a significant number of new bacterial species from the human gut microbiota and proved to be a valuable complement to culture-independent techniques. Using this culture-based approach, a new bacterial species has been isolated from a stool sample of a 39-year-old healthy Pygmy male and described using the taxonogenomic strategy. Cells of strain Marseille-P4356$^T$ are Gram-stain negative cocci. The strain grows optimally at 37 °C and is catalase positive but oxidase negative. Its 16S rRNA gene sequence exhibited 92.96% sequence similarity with *Dysgonomonas gadei* strain JCM 16698T (NR_113134.1), currently its phylogenetically closest species that has been validly named. The genome of strain Marseille-P4356$^T$ is 3,472,011 bp long with 37.3 mol% G+C content. Phenotypic, biochemical, proteomic, genomic and phylogenetic analyses, clearly demonstrate that strain Marseille-P4356$^T$ (= CCUG 71356$^T$ = CSUR P4356$^T$) represents a new species within the genus *Dysgonomonas*, for which we propose the name *Dysgonomonas massiliensis* sp. nov.

**Keywords** Culturomics · Taxono-genomics · Pygmy · *Dysgonomonas massiliensis* · Gut microbiota

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s10482-019-01227-1) contains supplementary material, which is available to authorized users.

M. Bilen · M. D. M. Fonkou · G. Dubourg · E. Tomei · M. Richez · J. Delerce · A. Levasseur · D. Raoult · F. Cadoret (✉)
IRD, APHM, MEPHI, IHU-Méditerranée Infection, Aix Marseille Université, Marseille, France
e-mail: frederic.n.cadoret@gmail.com

M. Bilen · Z. Daoud
Clinical Microbiology Department, Faculty of Medicine and Medical Sciences, University of Balamand, PO Box 33, Amioun, Lebanon

D. Raoult
Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

## Introduction

The human gut harbours a large number of bacteria and other organisms that have been extensively studied and correlated with health and diseases (Clemente et al. 2012). Besides the fundamental roles that the human gut plays (metabolic functions, energy harvesting, immunity, etc.), it is a good target for the development of therapeutic strategies, especially when specific bacterial profiles can be drawn in certain diseases (Million et al. 2016; Tidjani Alou et al. 2017). With the advancement of culture-independent techniques, our knowledge has been enhanced in terms of human gut microbiota composition, although several drawbacks have emerged such

as the accumulation of unidentified sequences, the inability to provide material for further investigations, failure to identify certain pathogenic species and depth bias (Greub 2012). Thus, culturomics was established with the purpose of culturing what was previously called 'un-cultivable' organisms, complementing metagenomics by providing identities to the operational taxonomic units (OTUs) and providing material for further investigations (Lagier et al. 2012a). This technique has succeeded in describing the human gut microbiota by isolating a significant number of new bacterial species and describing these via the 'taxonogenomic' approach (Fournier and Drancourt 2015; Lagier et al. 2016; Bilen et al. 2018). Herein, we describe using taxonogenomics, a new bacterial species, *Dysgonomonas massiliensis* sp. nov. represented by strain Marseille-P4356$^{T}$ (= CCUG 71356$^{T}$ = CSUR P4356$^{T}$), isolated from a stool sample of a healthy Pygmy male. The genus *Dysgonomonas* was first isolated from a human gall bladder and members are known to be Gram-stain negative, coccobacillus-shaped (Hofstad et al. 2000) and associated with gastroenteritis in immunocompromised patients (Murray et al. 2013).

## Materials and methods

Before starting this project, approval from the ethics committee under the number 09-022 from the Institut Fédératif de Recherche 48 was obtained along with a signed informed consent from the donor.

### Strain isolation and identification

Stool samples were taken from a 39-year-old healthy Pygmy male according to the Nagoya protocol. The shipment of samples was performed using a C-Top Ae-Ana medium (Culture Top, Marseille, France) and stored at − 80 °C at the URMITE laboratory (Marseille, France). Diluted in phosphate buffer saline, the stool sample was incubated in anaerobic culture bottles (BD BACTEC®, Plus Anaerobic/F Media, Le Pont de Claix, France) supplemented with 5% (v/v) sheep blood and 5% (v/v) sterile-filtered cow rumen at 37 °C. The growth of colonies was evaluated by subculturing samples on 5% sheep blood–enriched Columbia agar (bioMérieux, Marcy l'Etoile, France) and incubating them at 37 °C under anaerobic conditions. The identification of colonies was performed as previously described using MALDI-TOF MS (Matrix assisted laser desorption ionization time of flight mass spectrometry; Seng et al. 2010; Lagier et al. 2012b). When MALDI-TOF MS failed to identify the studied organism, 16S rRNA gene sequencing was performed for further analysis as previously described (Drancourt et al. 2004). Sequences were optimised and assembled with CodonCode Aligner tool (http://www.codoncode.com) and a Blast analysis was performed using the NCBI database and BLAST engine (http://blast.ncbi.nlm.nih.gov.gate1.inist.fr/Blast.cgi). A 16S rRNA sequence divergence of more than 5% was required to classify isolates into a new genus and a similarity of less than 98.65% was required to propose a new species (Kim et al. 2014). The 16S rRNA gene sequence of newly isolated bacterial species and the accompanying MALDI-TOF mass spectrum were deposited in EMBL-EBI and UMRS databases, respectively.

### Growth conditions, phenotypic, biochemical and antibiotic resistance analyses

Optimal growth conditions were evaluated by performing culture assays at different temperatures (25, 37, 45 and 55 °C), pH (6, 6.5, 7 and 8.5), oxygen environments (anaerobic (GENbag anaer, bioMérieux), aerobic and micro-aerophilic (GENbag Microaer, bioMérieux)) and NaCl concentrations (0, 5, 10, 50, 75 and 100 g $L^{-1}$).

In addition, API strips (20A, 50CH and ZYM; bioMerieux, France) were used according to the manufacturer's guidelines in order to determine the major enzymatic and biochemical activities of the unidentified organism. Spore-forming ability was evaluated by performing a culture assay using a bacterial suspension, previously exposed to a 20 min heat shock at 80 °C. A DM1000 photonic microscope (Leica Microsystems, Nanterre, France) was used as previously described for motility and Gram assessments and cell images were obtained as previously (Bilen et al. 2017).

Bile solubility test was performed by checking the turbidity after adding 0.5 mL of 2% sodium desoxycholate (Sigma Aldrich) to 0.5 mL of 1 McFarland bacterial suspension (in normal saline). The latter was incubated at 37 °C for 10 min before reading the results. A tube with 0.5 mL of 1 McFarland bacterial

suspension was taken as a control with normal saline added instead 2% sodium desoxycholate.

The antibiotic resistance profile of strain Marseille-P4356$^T$ was determined by performing different E-tests with: vancomycin, imipenem, rifampicin, amoxicillin, benzylpenecilin, minocycline, colistin, amikacin, metranidazol, teicoplanin, ertapenem, daptomycin, ceftriaxone, doxycycline, fosfomycin and kanamycin (bioMérieux, France).GC/MS and fatty acid analysis of strain Marseille-P4356$^T$ was carried out using 50 mg of bacterial biomass per tube as previously described (Dione et al. 2016). A Clarus 500 chromatography system was used connected to a SQ8s mass spectrometer (Perkin Elmer, Courtaboeuf, France) in order to analyse short chain fatty acids, as previously described (Zhao et al. 2006; Togo et al. 2017). Acetic, propionic, butyric, isobutyric, valeric, isovaleric, caproic, isocaproic, enanthic and isoenanthic acids were used as reference standards (Sigma Aldrich, Lyon, France).

## Genome sequencing and analyses

For genomic analyses, genomic DNA of strain Marseille-P4356$^T$ was first mechanically extracted using an acid-washed glass beads (G4649-500 g Sigma) treatment in a FastPrep BIO 101 instrument (Qbiogene, Strasbourg, France) at maximum speed (6.5) for 90 s. After 2 h of lysozyme incubation at 37 °C, an EZ1 biorobot (Qiagen) was used for DNA extraction with the EZ1 DNA tissues kit. The elution volume was 50 µl. Qubit assay was used for genomic DNA quantification with the high sensitivity kit (Life technologies, Carlsbad, CA, USA) to 71.6 ng/µl. Next, MiSeq Technology (Illumina Inc, San Diego, CA, USA) was used for genomic DNA sequencing by means of the Nextera mater pair strategy prep kit (Illumina). The studied organism, along with 11 other projects, was barcoded and mixed within the same run.1.5 µg of genomic DNA was used for library preparation using the Nextera mate pair Illumina guide. Tagmentation with a mate pair junction adapter was performed after fragmentation, the pattern of which was previously validated on an Agilent 2100 BioAnalyzer (Agilent Technologies Inc, Santa Clara, CA, USA) with a DNA 7500 labchip. Optimal DNA fragments were 8.933 kb and ranged from 1.5 kb up to 11 kb. 600 ng of the tagged fragmented DNA were circularised with no size selection and then mechanically sheared to small fragments with an optimal size at 721 bp on a Covaris device S2 in T6 tubes (Covaris, Woburn, MA, USA). The final library concentration was measured at 20.98 nmol/l and its profile was visualised on a High Sensitivity Bioanalyzer LabChip (Agilent Technologies Inc, Santa Clara, CA, USA). Normalization of the library at 2 nM was performed before pooling, followed by a denaturation and dilution step at 18 pM. Cluster generation and sequencing were set automatically for a single 39-h run in a 2 × 151-bp. Total information of 6 Gb was obtained from a 623 K/mm2 cluster density with a cluster passing quality control filters of 97.5% (11,904,000 passing filter paired reads). Within this run, the index representation for strain Marseille-P4356$^T$ was determined to be 7.32%. The 871,666 paired reads were trimmed and assembled as previously described (Lagier et al. 2012b). The Rast tool was used for genome annotation and analysis (Aziz et al. 2008; Overbeek et al. 2014; Brettin et al. 2015). Phages were detected using PHAST tool (Zhou et al. 2011), rRNA was detected with RNAmmer (Lagesen et al. 2007) and Artemis was used for genome circular representation (Carver et al. 2009).

For phylogenetic analyses, the NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/nucleotide/) was use to retrieve 16S rRNA gene sequences of the phylogenetically closely related species with standing in nomenclature of strain Marseille-P4356$^T$ after performing a Blast search at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch. Alignment was done using the ClustalW tool in MEGA (Kumar et al. 2016), which was adapted for phylogenetic inference generation within the maximum-likelihood method with 500 bootstraps (Tamura and Nei 1993; Tateno et al. 1994).

To further confirm the novelty of strain Marseille-P4356$^T$, OrthoANI and dDDH (Digital DNA:DNA hybridization) estimates were calculated between the genome of this strain and the genomes of its phylogenetically close species using OAT software (Lee et al. 2016) and Genome-to-Genome Distance Calculator 2.1 (http://ggdc.dsmz.de), respectively.

## Results

### Strain identification

MALDI-TOF MS failed to identify strain Marseille-P4356$^T$ due to the absence of a matching spectrum in the current database. Therefore 16S rRNA gene sequencing was performed and revealed that this strain exhibited 92.96% sequence similarity with *Dysgonomonas gadei* strain JCM 16698$^T$ (NR_113134.1) and 93.57% *Dysgonomonas mossii* strain JCM 16699$^T$ (NR_113135.1), currently the phylogenetically closest species with standing in nomenclature (Fig. 1; Supplementary Table 1). Even though the 16S rRNA gene sequence of strain Marseille-P4356$^T$ diverges by more than 5% with its phylogenetically close species, this strain falls between other *Dysgonomonas* species in the phylogenetic tree (Fig. 1). Accordingly, we classify strain Marseille-P4356$^T$ as a new bacterial species belonging to the genus *Dysgonomonas*. The 16S rRNA gene sequence of this strain has been deposited in GenBank with the accession number LT934441 and the mass spectrum generated by MALDI-TOF MS (Fig. 2) uploaded to the URMS database (http://www.mediterranee-infection.com/article.php?larub=280&titre=urms-database). The 16S rRNA gene similarities between strain Marseille-P4356$^T$ and closely related species are listed in Supplementary Table 1.

### Main characteristics of Strain Marseille-P4356$^{T'}$

Cells of strain Marseille-P4356$^T$ were observed to be Gram-stain negative, catalase positive but oxidase negative cocci. After 48 h of growth under anaerobic conditions on COS medium + 5% (v/v) sheep blood (bioMerieux, France) at 37 °C, the strain was observed to form smooth grey colonies of 0.03–1 mm diameter. Under electron microscope, cells were found to have a mean diameter of 0.6 µm (Table 1). Strain Marseille-P4356$^T$ was found to be able to grow at temperatures between 25 and 37 °C under aerobic and anaerobic conditions, optimally at 37 °C under anaerobic conditions to the strain can grow at pH between 6 and 8.5 and NaCl concentration below to 50 g $L^{-1}$. Strain Marseille-P4356$^T$ is non-motile, asporogenous and bile soluble.

In antibiotic susceptibility tests, strain Marseille-P4356$^T$ exhibited the following minimal inhibitory concentrations: vancomycin (> 32 µg $mL^{-1}$), rifampicin (2 µg $mL^{-1}$), benzylpenecilin (0.032 µg $mL^{-1}$), amoxicillin (0.023 µg $mL^{-1}$), imipenem (0.094 µg $mL^{-1}$), amikacin (> 256 µg $mL^{-1}$), minocycline (16 µg $mL^{-1}$), teicoplanin (3 µg $mL^{-1}$), colistin (> 256 µg $mL^{-1}$), daptomycin (> 256 µg $mL^{-1}$), metronidazole (1.5 µg $mL^{-1}$), ceftriaxone (0.38 µg $mL^{-1}$), ertapenem (0.016 µg $mL^{-1}$), fosfomycin (< 1024 µg $mL^{-1}$), doxycycline (24 µg $mL^{-1}$) and anamycin (4 µg $mL^{-1}$) respectively. These indicate possible resistance mechanisms towards metronidazole, doxycycline, amikacin, colistin, fosfomycin and vancomycin. The closely related species *D. gadei* exhibited minimal inhibitory concentrations as follows: metronidazole (1.5 µg $mL^{-1}$), doxycycline (0.19 µg $mL^{-1}$), imipenem (0.5 µg $mL^{-1}$), ceftriaxone (> 256 µg $mL^{-1}$), vancomycin (48 µg $mL^{-1}$), teicoplanin (> 256 µg $mL^{-1}$), amoxicillin (6 µg $mL^{-1}$), benzylpenicillin (32 µg $mL^{-1}$) and azithromicin (16 µg $mL^{-1}$), showing also a diverse resistance profile (Hofstad et al. 2000). To have an unidentified bacterial species with such a wide resistance profile should encourage the scientific-medical community to enhance their work on characterising the human microbiota and studying unknown commensal bacteria as they may have the capacity to become pathogenic (Isenberg 1988). Using API 50CH tests, positive reactions were observed with L-arabinose, D-xylose, D-galactose, D-glucose, D-fructose, D-mannose, D-mannitol, *N*-acetylglucosamine, amygdalin, arbutin, aesculin ferric citrate, salicin, D-maltose, D-lactose, starch, glycogen, gentiobiose, D-arabitol and potassium gluconate. For API ZYM tests, positive reactions were observed with esterase (C4), esterase lipase (C8), lipase (C14), valine arylamidase, α-chymotrypsin, naphthol-AS-BI-phosphohydrolase, α-galactosidase, α-glucosidase, β-glucosidase and *N*-acetyl-β-glucosaminidase. In API 20A tests, positive reactions were observed with glucose, mannitol, lactose, maltose, xylose, arabinose and mannose. A comparison of the general phenotypic and biochemical characteristics between this strain and close relatives is shown in Table 1 (Hofstad et al. 2000; Lawson et al. 2002, 2010; Kodama et al. 2012; Yang et al. 2014; Kita et al. 2015; Pramono et al. 2015). All *Dysgonomonas* species are facultatively anaerobic, Gram-stain negative, non-motile and produce acids from D-mannose, D-glucose, D-maltose and D-lactose.

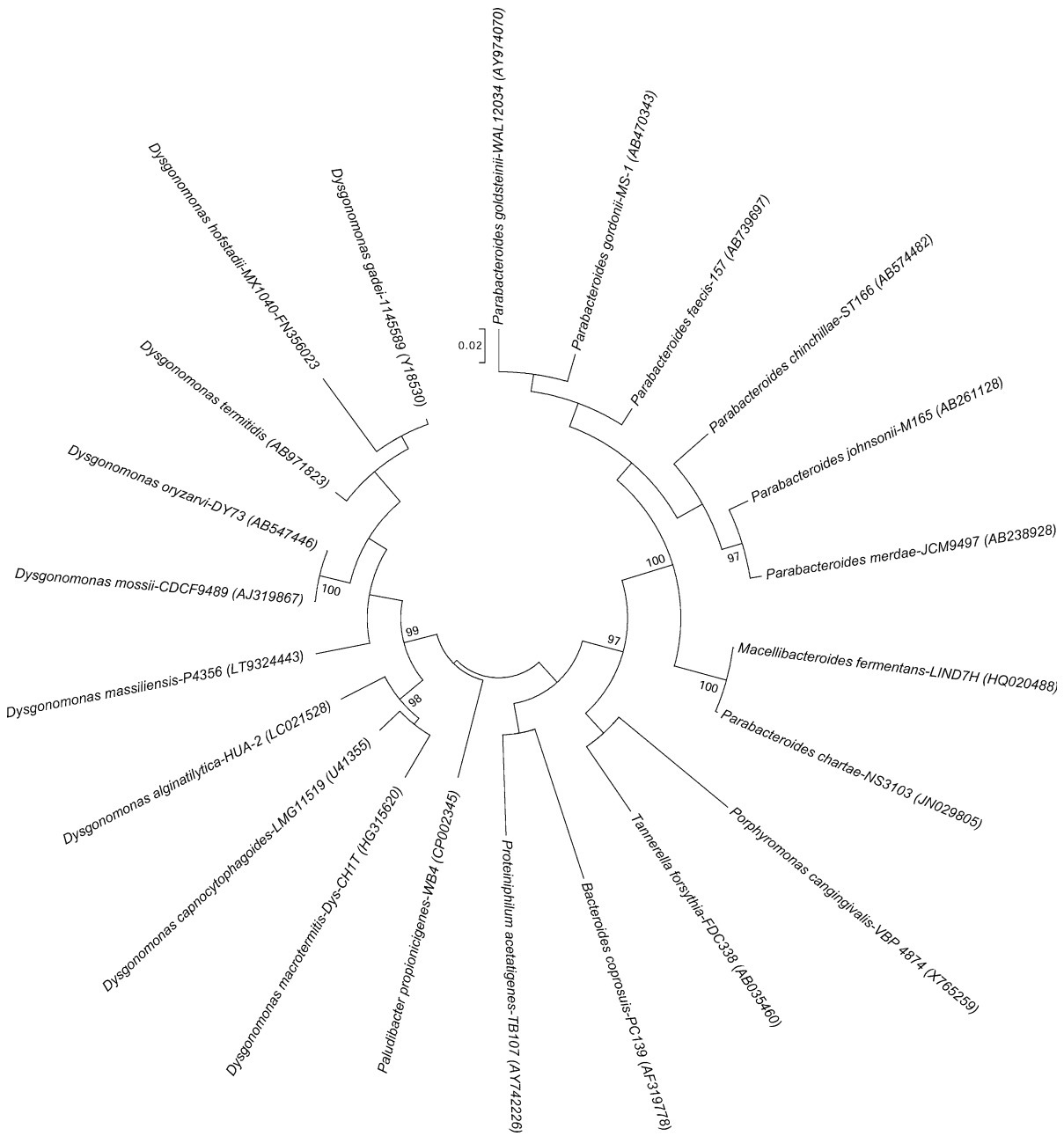


**Fig. 1** Phylogenetic tree representing the position of strain Marseille-P4356$^T$ relative to other closely related species. 16S rRNA sequences were recovered after a nucleotide Blast search of the Silva project "The All-Species Living Tree" database (LTPs121). Muscle was used for sequence alignment and Fast tree for approximately maximum likelihood sequence tree generation. Local values obtained with the Shimodaira–Hasegawa test are shown on the nodes

Only strain Marseille-P4356$^T$ is alkaline phosphatase-positive.

The most abundant fatty acid identified in strain Marseille-P4356$^T$ is anteiso-C15:0 (12-methyl-tetradecanoic acid). Several specific 3-hydroxy structures were also identified. Minor amounts of unsaturated and other saturated fatty acids were also detected. The closely related species, *D. gadei*

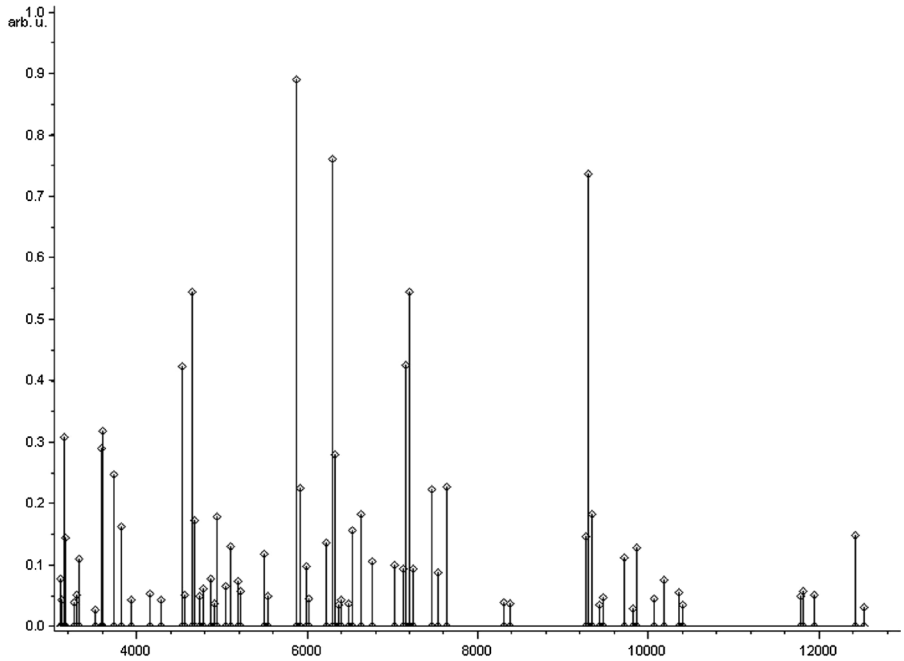


**Fig. 2** Reference MALDI-TOF mass spectrum representing *D. massiliensis* strain Marseille-P4356$^T$ obtained after comparing 12 spectra

contains iso-C14:0 (12-methyl-tridecanoic acid), anteiso-C15:0 (12-methyl-tetradecanoic acid), C16:0 (hexadecanoic acid) and iso-3OH C16:0 (3-hydroxy-14-methyl-Pentadecanoic acid) as the major fatty acids (Hofstad et al. 2000) which were also all detected in strain Marseille-P4356$^T$ (Table 2). These results are in accordance with the known fatty of acid profiles of these organisms, which are generally characterised by the presence of a large amounts of anteiso- and iso-methyl branched, straight- chain saturated and 3-hydroxy long-chain fatty acids (Wallace et al. 1989). Regarding short chain fatty acid fermentation products, after 72 h of culture in a hemoculture flask supplemented with blood, production of acetic (> 10 mM), propanoic (6.8 ± 0.3 mM), isobutanoic (0.5 ± 0.1 mM) and isopentanoic (1.1 ± 0.1 mM) acids was detected. Butanoic, pentanoic, hexanoic, isohexanoic and heptanoic acids were not detected.

## Genomic analyses and comparison

The draft genome of strain Marseille-P4356$^T$ is 3,472,011 bp long with 37.3 mol% G+C content. It is composed of 26 contigs (26 scaffolds). 2994 genes were detected with 2907 coding DNA sequences. 51 genes (RNA), 7 rRNA (5, 1, 1 for 5S, 16S, 23S rRNA, respectively) and 42 tRNAs were detected. No CRISPRs repeats were found. 1395 proteins were annotated as hypothetical proteins. Using the PHAST tool, 4 prophages regions were identified, of which one region is intact, 3 regions are incomplete (not 100% compatible with phage sequences). The intact region was located in the region 3097681–3103890 with 43.95 mol% G+C content and shares a high number of proteins (70%) with Enterobacteria phage phiX174 sensu lato (NC_001422.1). The importance of tracking phage sequences is to check for horizontal gene transfer scenarios that might have occurred and led to specific virulence traits for the studied microorganism (Arber 2014; Penadés et al. 2015). Based on the RAST annotation, 41 proteins were correlated with virulence factors and defense, of which 28 were related to antibiotic and toxic compound resistance as follows: Copper homeostasis (2), Cobalt-Zinc-cadmium-resistance (4), vancomycin resistance (1), tripartite multidrug resistance system (3), streptothricin resistance (1), tetracycline resistance-ribosome protection type

**Table 1** Differential characteristics of *Dysgonomonas massiliensis* (DM), *Dysgonomonas gadei* (DG), *Dysgonomonas capnocytophagoides* (DC), *Dysgonomonas macrotermitis* (DMM), *Dysgonomonas hofstadii* (DH), *Dysgonomonas oryzarvi* (DO), *Dysgonomonas alginatilytica* (DA), *Dysgonomonas mossii* (DMMM) and *Dysgonomonas termitidis* (DT). Data for reference strains are taken from the original work reporting its first isolation and description (Hofstad et al. 2000; Lawson et al. 2002, 2010; Kodama et al. 2012; Yang et al. 2014; Kita et al. 2015; Pramono et al. 2015)

| Properties | DM | DG | DC | DMM | DH |
|---|---|---|---|---|---|
| Oxygen requirement | Facultative anaerobic | Facultative anaerobic | Facultative anaerobic | Facultative anaerobic | Facultative anaerobic |
| Major fatty acid | Antesio C15:0, C15:0, C16:0 | Anteiso-C15 : 0, C16 : 0, iso-C14 : 0 | Iso-C14 : 0, anteiso-C15 : 0, iso-C16 : 0 3-OH | Anteiso-C15 : 0, iso-C17 : 0 3-OH, C16 : 0 | Anteiso-C15 : 0, iso-C14 : 0, iso-C16 : 0 3-OH |
| Production of | | | | | |
| Alkaline phosphatase | − | + | + | + | + |
| Catalase | + | + | − | + | − |
| β-galactosidase | _ | + | + | + | − |
| Acid production from | | | | | |
| *N*-acetyl-β-glucosamine | + | + | − | + | + |
| L-Arabinose | + | + | + | + | − |
| D-Mannitol | + | + | + | + | − |
| G+C content (mol%) | 57.3 | 39.6 | 37.7 | 40 | 45.3 |
| Habitat | Human gut | Human clinical sample | Human clinical sample | Termite hindgut | Human clinical sample |

| Properties | DO | DA | DMMM | DT |
|---|---|---|---|---|
| Oxygen requirement | Facultative anaerobic | Facultative anaerobic | Facultative anaerobic | Facultative anaerobic |
| Major fatty acid | Anteiso-C15 : 0, iso-C17 : 03-OH, C16 : 0 3-OH | Anteiso-C15: 0, C15:0, iso-C14:0 | Na | Anteiso-C15:0; and C18:1v9c |
| Production of | | | | |
| Alkaline phosphatase | + | + | + | + |
| Catalase | − | − | − | − |
| β-galactosidase | + | − | + | + |
| Acid production from | | | | |
| *N*-acetyl-β-glucosamine | + | + | + | + |
| L-Arabinose | + | + | + | + |
| D-Mannitol | − | + | + | − |
| G+C content (mol%) | 37.5 | 37.5 | 38.5 | 41.8 |
| Habitat | Microbial fuel cell | Microbial consortium | Human clinical source | Termite gut |

**Table 2** Fatty acids profile of strain Marseille-P4356

| Fatty acids | Name | Mean relative (%)[a] |
|---|---|---|
| 15:0 anteiso | 12-Methyl-tetradecanoic acid | 72.9 ± 0.8 |
| 15:0 | Pentadecanoic acid | 4.0 ± 0.3 |
| 16:0 | Hexadecanoic acid | 3.8 ± 0.1 |
| 15:0 iso | 13-Methyl-tetradecanoic acid | 2.9 ± 0.1 |
| 17:0 3-OH iso | 3-Hydroxy-15-methyl-hexadecanoic acid | 2.9 ± 0.4 |
| 14:0 anteiso | 11-Methyl-tridecanoic acid | 2.7 ± 0.2 |
| 17:0 3-OH anteiso | 3-Hydroxy-14-methyl-hexadecanoic acid | 2.4 ± 0.9 |
| 18:2ω6 | 9,12-Octadecadienoic acid | 1.4 ± 0.2 |
| 18:1ω9 | 9-Octadecenoic acid | 1.3 ± 0.1 |
| 5:0 anteiso | 2-Methyl-butanoic acid | 1.0 ± 0.1 |

[a]Mean peak area percentage

(4), fluoroquinolones resistance (4), beta-lactamase (1) and multidrug resistance pumps (8). No spore protection system was detected, which is in accordance with our phenotypic results. Moreover, no motility features were detected, which is also in accordance with our observations. A graphical representation of the draft genome of strain Marseille-P4356$^T$ is shown in Fig. 3. When compared to phylogenetically closely related species with standing in nomenclature, the draft genome sequence of strain Marseille-P4356$^T$ size is smaller than that of, *Dysgonomonas macrotermitis* strain DSM 27370,

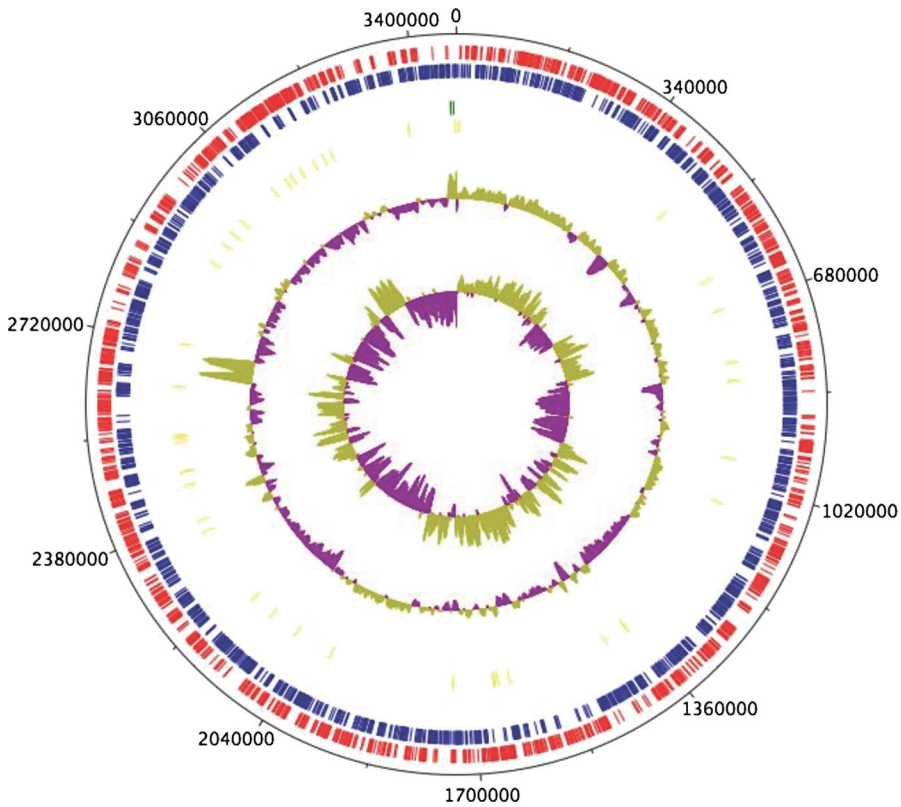


**Fig. 3** Circular representation of the *D. massiliensis* strain Marseille-P4356$^T$ draft genome. From the outer strand to the inner strand: Coding DNA sequences on the forward strand, coding DNA sequences on the reverse strand, rRNA, tRNA and GC plot and skew

*Dysgonomonas hofstadii* MX 1040$^T$, *Dysgonomonas capnocytophagoides* DSM 22835, *Dysgonomonas gadei* ATCC BAA-286$^T$ (*D. gadei*), *Dysgonomonas mossii* DSM 22836 and *Paludibacter propionicigenes* WB4 (4.7, 5.0, 4.4, 5.2, 3.9 and 3.7 Mb, respectively).

The G+C content of strain Marseille-P4356$^T$ is higher than that of *D. macrotermitis*, *D. hofstadii*, *D. capnocytophagoides*, *D. gadei*, *D. mossii* and *P. propionicigenes* (40, 45.3, 37.7, 39.6, 37.5 and 38.9%, respectively) (Supplementary Table 2).

The OrthoANI values (%) between strain Marseille-P4356$^T$, *P. propionicigenes*, *D. macrotermitis*, *D. hofstadii*, *D. capnocytophagoides*, *D. gadei* and *D. mossi* were 66.54, 69.06, 69.67, 69.38, 69.76 and 69.77, respectively (Supplementary Table 3). These values are below the cutoff threshold (95–96%), determined to discriminate bacterial species (Lee et al. 2016). Additionally, dDDH estimates between strain Marseille-P4356$^T$, *P. propionicigenes*, *D. macrotermitis*, *D. hofstadii*, *D. capnocytophagoides*, *D. gadei* and *D. mossi* were 18.4 [16.2–20.7], 21.3 [19–23.7%], 20 [17.8–22.4], 21.9 [19.6–24.3], 20 [17.8–22.4%] and 21.8 [19.5–24.2%], respectively (Supplementary Table 4). dDDH estimates of all species between each other were below 30% and below the considered threshold with intervals suggested to delimitate bacterial species (Meier-Kolthoff et al. 2013).

Phylogenetic, phenotypic and biochemical analyses demonstrated that strain Marseille-P4356$^T$ merits recognition as a novel member of the genus *Dysgonomonas* for which the name *Dsygonomonas massiliensis* is proposed. The type strain is *Dysgonomonas massiliensis* strain Marseille-P4356$^T$ (= CSUR P4356$^T$ = CCUG 71356$^T$), from the human gut. The Digital Protologue database (Rosselló-Móra et al. 2017) TaxoNumber for strain Marseille-P4356$^T$ is TA00816.

## Description of *Dysgonomonas massiliensis* sp. nov.

*Dysgonomonas massiliensis* (mas.si.li.en'sis. L. fem. adj., *massiliensis*, pertaining to Massilia, the ancient name of the city of Marseille, where this bacterium was characterised).

Cells are Gram-stain negative cocci, 0.6 µm in diameter, non-motile and asporogenous. Forms smooth grey colonies, 0.03–1 mm diameter. Catalase positive and oxidase negative. Grows optimally at 37 °C under anaerobic conditions and tolerates pH of 6–8.5 and NaCl concentration below 50 g $L^{-1}$. The major biochemical test reactions are given in Table 1. The most abundant fatty acid is anteiso-C15:0.

The type strain is Marseille-P4356$^T$ (= CSUR P4356$^T$ = CCUG 71356$^T$), which was isolated from the stool sample of a healthy 38-year-old pygmy male from the Democratic Republic of the Congo. The draft genome of strain Marseille-P4356$^T$ is 3.5 Mb long with 37.3 mol% of G+C content. The 16S rRNA gene and genome sequences of the type strain have been deposited in EMBL-EBI under accession numbers LT934443 and OEPV00000000, respectively.

**Acknowledgements** This study was supported by IHU Méditerranée Infection, Marseille, France and by the French Government under the «Investissements d'avenir» (Investments for the Future) program managed by the Agence Nationale de la Recherche (ANR, fr: National Agency for Research), (reference: Méditerranée Infection 10-IAHU- 03). This work was supported by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMI.

**Authors' contribution** MB: Isolated, described and wrote the manuscript; MDMF: helped in the taxonogenomics description, GD: critical analysis of the work and wrote the manuscript, ET: Genomic analysis, MR: helped in the taxonogenomics description, JD: genomic analysis, AL: helped in the genomic analyses, ZD: writing an critical analysis of the manuscript, DR: designed the project, helped in writing, reviewing and critical analysis; FC: study design, data analysis and writing the manuscript.

**Compliance with ethical standards**

**Conflict of interest** The authors declare no conflict of interest.

## References

Arber W (2014) Horizontal gene transfer among bacteria and its role in biological evolution. Life Open Access J 4:217–224. https://doi.org/10.3390/life4020217

Aziz RK, Bartels D, Best AA et al (2008) The RAST Server: rapid annotations using subsystems technology. BMC Genom 9:75. https://doi.org/10.1186/1471-2164-9-75

Bilen M, Cadoret F, Richez M et al (2017) *Libanicoccus massiliensis* gen. nov., sp. nov., a new bacterium isolated from human stool. New Microbes New Infect 21:63–71. https://doi.org/10.1016/j.nmni.2017.11.001

Bilen M, Dufour J-C, Lagier J-C et al (2018) The contribution of culturomics to the repertoire of isolated human bacterial

and archaeal species. Microbiome 6:94. https://doi.org/10.1186/s40168-018-0485-5

Brettin T, Davis JJ, Disz T et al (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365

Carver T, Thomson N, Bleasby A et al (2009) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25:119–120. https://doi.org/10.1093/bioinformatics/btn578

Clemente JC, Ursell LK, Parfrey LW, Knight R (2012) The impact of the gut microbiota on human health: an integrative view. Cell 148:1258–1270. https://doi.org/10.1016/j.cell.2012.01.035

Dione N, Sankar SA, Lagier J-C et al (2016) Genome sequence and description of *Anaerosalibacter massiliensis* sp. nov. New Microbes New Infect 10:66–76. https://doi.org/10.1016/j.nmni.2016.01.002

Drancourt M, Berger P, Raoult D (2004) Systematic 16S rRNA gene sequencing of atypical clinical isolates identified 27 new bacterial species associated with humans. J Clin Microbiol 42:2197–2202

Fournier P-E, Drancourt M (2015) New Microbes New Infections promotes modern prokaryotic taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans". New Microbes New Infect 7:48–49. https://doi.org/10.1016/j.nmni.2015.06.001

Greub G (2012) Culturomics: a new approach to study the human microbiome. Clin Microbiol Infect 18:1157–1159. https://doi.org/10.1111/1469-0691.12032

Hofstad T, Olsen I, Eribe ER et al (2000) *Dysgonomonas* gen. nov. to accommodate *Dysgonomonas gadei* sp. nov., an organism isolated from a human gall bladder, and *Dysgonomonas capnocytophagoides* (formerly CDC group DF-3). Int J Syst Evol Microbiol 50:2189–2195. https://doi.org/10.1099/00207713-50-6-2189

Isenberg HD (1988) Pathogenicity and virulence: another view. Clin Microbiol Rev 1:40–53

Kim M, Oh H-S, Park S-C, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. https://doi.org/10.1099/ijs.0.059774-0

Kita A, Miura T, Okamura Y et al (2015) *Dysgonomonas alginatilytica* sp. nov., an alginate-degrading bacterium isolated from a microbial consortium. Int J Syst Evol Microbiol 65:3570–3575. https://doi.org/10.1099/ijsem.0.000459

Kodama Y, Shimoyama T, Watanabe K (2012) *Dysgonomonas oryzarvi* sp. nov., isolated from a microbial fuel cell. Int J Syst Evol Microbiol 62:3055–3059. https://doi.org/10.1099/ijs.0.039040-0

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054

Lagesen K, Hallin P, Rødland EA et al (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160

Lagier J-C, Armougom F, Million M et al (2012a) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193. https://doi.org/10.1111/1469-0691.12023

Lagier J-C, El Karkouri K, Nguyen T-T et al (2012b) Non-contiguous finished genome sequence and description of *Anaerococcus senegalensis* sp. nov. Stand Genomic Sci 6:116–125. https://doi.org/10.4056/sigs.2415480

Lagier J-C, Khelaifia S, Alou MT et al (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol 1:16203. https://doi.org/10.1038/nmicrobiol.2016.203

Lawson PA, Falsen E, Inganäs E et al (2002) *Dysgonomonas mossii* sp. nov., from human sources. Syst Appl Microbiol 25:194–197. https://doi.org/10.1078/0723-2020-00107

Lawson PA, Carlson P, Wernersson S et al (2010) *Dysgonomonas hofstadii* sp. nov., isolated from a human clinical source. Anaerobe 16:161–164. https://doi.org/10.1016/j.anaerobe.2009.06.005

Lee I, Ouk Kim Y, Park S-C, Chun J (2016) OrthoANI: an improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–1103. https://doi.org/10.1099/ijsem.0.000760

Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinform 14:60. https://doi.org/10.1186/1471-2105-14-60

Million M, Tidjani Alou M, Khelaifia S et al (2016) Increased gut redox and Depletion of anaerobic and methanogenic prokaryotes in severe acute malnutrition. Sci Rep 6:26051. https://doi.org/10.1038/srep26051

Murray PR, Rosenthal KS, Pfaller MA (2013) Medical microbiology, with student consult online Access, 7. Medical Microbiology. Elsevier Health Sciences

Overbeek R, Olson R, Pusch GD et al (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226

Penadés JR, Chen J, Quiles-Puchalt N et al (2015) Bacteriophage-mediated spread of bacterial virulence genes. Curr Opin Microbiol 23:171–178. https://doi.org/10.1016/j.mib.2014.11.019

Pramono AK, Sakamoto M, Iino T et al (2015) *Dysgonomonas termitidis* sp. nov., isolated from the gut of the subterranean termite Reticulitermes speratus. Int J Syst Evol Microbiol 65:681–685. https://doi.org/10.1099/ijs.0.070391-0

Rosselló-Móra R, Trujillo ME, Sutcliffe IC (2017) Introducing a digital protologue: a timely move towards a database-driven systematics of archaea and bacteria. Antonie Van Leeuwenhoek 110:455–456. https://doi.org/10.1007/s10482-017-0841-7

Seng P, Rolain J-M, Fournier PE et al (2010) MALDI-TOF-mass spectrometry applications in clinical microbiology. Future Microbiol 5:1733–1754. https://doi.org/10.2217/fmb.10.127

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526. https://doi.org/10.1093/oxfordjournals.molbev.a040023

Tateno Y, Takezaki N, Nei M (1994) Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. Mol Biol Evol 11:261–277. https://doi.org/10.1093/oxfordjournals.molbev.a040108

Tidjani Alou M, Million M, Traore SI et al (2017) Gut bacteria missing in severe acute malnutrition, can we identify potential probiotics by culturomics? Front Microbiol 8:899. https://doi.org/10.3389/fmicb.2017.00899

Togo AH, Durand G, Khelaifia S et al (2017) *Fournierella massiliensis* gen. nov., sp. nov., a new human-associated member of the family *Ruminococcaceae*. Int J Syst Evol Microbiol 67:1393–1399. https://doi.org/10.1099/ijsem.0.001826

Wallace PL, Hollis DG, Weaver RE, Moss CW (1989) Characterization of CDC group DF-3 by cellular fatty acid analysis. J Clin Microbiol 27:735–737

Yang Y, Zhang N, Ji S et al (2014) *Dysgonomonas macrotermitis* sp. nov., isolated from the hindgut of a fungus-growing termite. Int J Syst Evol Microbiol 64:2956–2961. https://doi.org/10.1099/ijs.0.061739-0

Zhao G, Nyman M, Jönsson JA (2006) Rapid determination of short-chain fatty acids in colonic contents and faeces of humans and rats by acidified water-extraction and direct-injection gas chromatography. Biomed Chromatogr BMC 20:674–682. https://doi.org/10.1002/bmc.580

Zhou Y, Liang Y, Lynch KH et al (2011) PHAST: a fast phage search tool. Nucleic Acids Res 39:W347–W352. https://doi.org/10.1093/nar/gkr485