Noncontiguous finished genome sequence and description of Gabonia massiliensis gen. nov., sp. nov.

G. Mourembou1,2, J. Rathored1, A. Ndjoayi-Mbiguino1, J. B. Lekana-Douki1,2, F. Fenollar1, C. Robert1, P.-E. Fournier1, D. Raoult1,6 and J. C. Lagier1

1) Aix Marseille Université, URMITE, UM63, CNRS 7278, INSERM 1095, Marseille, France, 2) Ecole Doctorale Régionale d’Afrique Centrale, 3) Unité de Parasitologie Médicale (UPARAM) CIRMF, Franceville, 4) Département de Microbiologie, Laboratoire National de référence IST/sida, Faculté de Médecine, 5) Département de Parasitologie Mycologie et de Médecine Tropicale, Université des Sciences de la Santé B.P., Libreville, Gabon and 6) Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

Culturomics coupled with taxonogenomics is currently used to isolate and characterize new bacteria. Here we describe the features and complete genome sequence of Gabonia massiliensis strain GM3, an anaerobic Gram negative, non-spore-forming and catalase-positive bacillus isolated from a stool specimen of a healthy Gabonese male youth. Belonging to a new genus called Gabonia, it exhibits a genome of 4 261 752 bp including 37.9% GC content and 3,288 predicted genes.

Keywords: Culturomics, Gabonia massiliensis gen. nov. et sp. nov., genome, taxonogenomics, gut microbiota

Original Submission: 28 September 2015; Revised Submission: 3 November 2015; Accepted: 4 November 2015
Article published online: 24 November 2015

Introduction

Culturomics was recently introduced in our laboratory (URMITE, Marseille, France) as an alternative method to expand the human gut repertoire [1], thanks to the multiplication of culture conditions with a rapid identification method by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) [1]. Consequently, new genera and species of bacteria are found thanks to this technique [2–6].

The characterization of new bacteria was previously based on a combination of phylogenetic and genotypic characteristics, including 16S rRNA sequence similarity, GC percentage and DNA-DNA hybridization [7–10]. Unfortunately, significant limits such as their threshold values, which were not applicable to all species or genera, did not promote their utilization [11]. The advent of high-throughput sequencing techniques made many bacterial genome sequences available in public databases [12]. Our laboratory recently developed a taxonogenomic approach based on a systematic comparison of phenotypic (especially the MALDI-TOF spectrum) characteristics and genomic characteristics with the phylogenetically closest bacteria found in the databases [2–6]. Here we present an analysis of the characteristics that allowed us to describe Gabonia massiliensis strain GM3, a bacterium isolated from a stool specimen from a healthy Gabonese male youth and classified into the Porphyromonadaceae family, created in 2011 by Krieg [13].

We describe the classification, biochemical features and complete genomic sequencing and annotation of G. massiliensis gen. nov., sp. nov. strain GM3 (= CSUR P1909 = DSM 100571).
Material and Methods

Ethics and sample collection
The stool sample was collected in Lébamba (Gabon) in January 2015 after approval from the National Ethics Committee of Gabon (registration 0023/2013/SG/CNE) and IFR48 (Marseille, France) (registration 09-022) and after receipt of the patient’s signed consent. The specimen was taken from a healthy 16-year-old Gabonese male youth (body mass index, 19.03 kg/m²). He was a member of the Nzebi tribe, which is one of the biggest tribes in Gabon. The stool sample was stored at −80°C until it was sent to URMITE (Marseille, France).

Strain isolation
In April 2015, the stool sample was diluted in phosphate-buffered saline (Life Technologies, Carlsbad, CA, USA). Obtained inoculum was preincubated anaerobically at 37°C in an anaerobic blood culture bottle (bioMérieux, Marcy l’Étoile, France) enriched with blood and rumen [14]. After the preincubation stage, 100 μL was incubated for 24 to 48 hours on 5% sheep’s blood–enriched Columbia agar (bioMérieux) at 37°C under anaerobic conditions using an anaerobic atmosphere generator system (GENbag anaer; bioMérieux). The bacteria colonies obtained were isolated on 5% sheep’s blood–enriched Columbia agar solid medium (bioMérieux) and identified by MALDI-TOF and 16S rRNA sequencing.

MALDI-TOF and 16S rRNA identifications
MALDI-TOF identification consisted of picking one isolated bacterial colony with a pipette tip from a culture agar plate and spreading it as a thin film on an MTP 384 MALDI-TOF target plate (Bruker Daltonics, Leipzig, Germany) [2–6]. Twelve depositions from 12 distinct isolated colonies were performed for strain GM3. Then 2 μL of matrix solution (saturated solution of α-cyano-4-hydroxycinnamic acid in 50% acetonitrile and 2.5% trifluoroacetic acid) was overlaid on each smear and allowed to dry for 5 minutes. Measurements and MALDI-TOF analysis were conducted as previously described [2–6]. Identification depended on the score obtained; a score of ≥2 with a validly published species enabled identification at the species level; a score of ≥1.7 but <2 enabled identification at the genus level; and a score of <1.7 did not enable any identification [2–6]. If the colonies corresponded to a new bacterium, confirmation of the identification had to occur with a 16S rRNA PCR coupled with sequencing. That was performed using GeneAmp PCR System 2720 thermal cyclers (Applied Biosystems, Bedford, MA, USA) and ABI Prism 3130xl Genetic Analyzer capillary sequencer (Applied Biosystems), respectively [15]. If the confirmation was positive, the spectrum of the new bacterium was entered into the Bruker database. The 16S rRNA nucleotide sequence was corrected using Chromas Pro 1.34 software (Technelysium, Tewantin, Australia), and the BLASTn searches were performed by the National Center for Biotechnology Information tool (http://blast.ncbi.nlm.nih.gov/gate1.inist.fr/Blast.cgi). For phylogeny, sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using the neighbour-joining method within MEGA software. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 1000 times to generate a majority consensus tree. The 16S rRNA for G. massiliensis strain GM3 was deposited in GenBank under accession number LN849789.

Growth conditions
Strain GM3 was cultivated on 5% sheep’s blood–enriched Colombia agar (bioMérieux) in anaerobic conditions at different temperatures (28, 37, 45 and 55°C) to assess its range of growth temperatures. The GENbag anaer and GENbag microaer systems (bioMérieux) were respectively used to assess the ability of the bacterium to grow anaerobically or in microaerophilic conditions at 37°C, using the same medium culture. For halophilia and pH testing, four NaCl concentrations (0, 5, 15 and 45%) and three different pHs (5, 7 and 8.5) were tested.

Biochemical, sporulation and motility assays
API Gallery systems were performed with API ZYM, API 20 A and API 50CH (bioMérieux) for biochemical assays. Oxidase (Becton Dickinson, Le Pont de Claix, France) and catalase assays (bioMérieux) were done separately. Sporulation was tested by performing thermal shock on bacterial colonies (diluted in phosphate-buffered saline) at 80°C for 10 minutes. The motility of strain GM3 was tested by observing its fresh colony between blades and slats using a DM1000 photonic microscope (Leica Microsystems, Nanterre, France) with a 40× objective lens.

Antibiotic susceptibility
The antibiotic susceptibility of G. massiliensis was tested using SirScan Discs of amoxicillin, doxycycline, rifampicin, nitrofurantoin, vancomycin, amoxicillin/clavulanic acid, clindamycin, gentamicin, imipenem, erythromycin, metronidazole, trimethoprim/sulfamethoxazole, amikacin, ciprofloxacin and tobramycin antibiotics (i2a, Montpellier, France).

Microscopy
G. massiliensis strain GM3 was observed after negative colouration using a Tecnai G20 (FEI, Limeil-Brevannes, France) transmission electron microscope at an operating voltage of 60 kV. Gram coloration was performed using a Color Gram 2 Kit.
This coloration was observed by using the DM1000 photonic microscope (Leica Microsystems) with a 100× oil-immersion objective lens.

Genome sequencing and assembly
Using the mate-pair strategy, the genomic DNA (gDNA) of G. massiliensis GM3 was sequenced on the MiSeq sequencer (Illumina, San Diego, CA, USA) [2–6]. The gDNA was bar-coded 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 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 of 4.08 kb. No size selection was performed, and only 464 ng of tagmented fragments were circularized [2–6]. The circularized DNA was mechanically sheared to small fragments with an optimal size of 569 bp in microtubes on the Covaris S2 device (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. An automated cluster generation and sequencing run were performed in a single 39-hour run on 2 × 251 bp. A total of 10.1 Gb of information was obtained from a 1189K/mm² cluster density with a cluster passing quality control filters of 99.1% (22 579 000 clusters).

The reads obtained were trimmed and then assembled using the CLC genomics WB4 software [2–6].

Genome annotation and comparison
Using Prodigal (http://prodigal.ornl.gov/) with default parameters, we predicted open reading frames (ORFs). However, if simultaneously fragmented and tagged with a mate-pair junction adapter. The pattern of 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 of 4.08 kb. No size selection was performed, and only 464 ng of tagmented fragments were circularized [2–6]. The circularized DNA was mechanically sheared to small fragments with an optimal size of 569 bp in microtubes on the Covaris S2 device (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. An automated cluster generation and sequencing run were performed in a single 39-hour run on 2 × 251 bp. A total of 10.1 Gb of information was obtained from a 1189K/mm² cluster density with a cluster passing quality control filters of 99.1% (22 579 000 clusters). The reads obtained were trimmed and then assembled using the CLC genomics WB4 software [2–6].

Genome annotation and comparison
Using Prodigal (http://prodigal.ornl.gov/) with default parameters, we predicted open reading frames (ORFs). However, if
they spanned a sequencing gap region, the predicted ORFs were excluded. The predicted bacterial protein sequences were searched against the GenBank [16] and Clusters of Orthologous Groups (COGs) databases using BLASTP. Both tRNAs and rRNAs were predicted using the tRNAScan-SE [17] and RNAmmer [18] tools, respectively. Signal peptides and the number of transmembrane helices were predicted using SignalP [19] and TMHMM [20], respectively. Mobile genetic elements were predicted using PHAST [21] and RAST [22]. ORFans were identified when their BLASTP E value was lower than 1e-03 for an alignment length greater than 80 amino acids. If alignment lengths were smaller than 80 amino acids, we used an E value of 1e-05. We used such parameter thresholds in previous works to define ORFans. Artemis [23] and a DNAPlotter [24] were used for data management and visualization of genomic features, respectively. Finally, we used the Mauve alignment tool (version 2.3.1) for multiple genomic sequence alignments [25].

The mean level of nucleotide sequence similarity at the genome level between G. massiliensis and other bacteria was estimated using the average genomic identity of orthologous gene sequences (AGIOS) homemade software [6]. Overall, this software is combined with other software: Proteinortho [26] (to detect orthologous proteins between genomes compared two by two, and then retrieves the corresponding genes) and the Needleman-Wunsch global alignment algorithm (to determine the mean percentage of nucleotide sequence identity among orthologous ORFs).

Annotation and comparison analyses were performed using the multiagent software system DAGOBHAH [27] that includes Figenix [28].

G. massiliensis strain GM3 was compared with Bacteroides salantrionis strain DSM 18170, Alistipes shahii strain WAL 8301, Alistipes finegoldii strain DSM 17242, Barnesiella viscericola strain DSM 18177, Ornithobacterium rhinotracheale strain ORT-UMN 88, Bacteroides fragilis strain YCH46, Bacteroides dorei strain HS1 and Helicobacter pylori strain 26695.

Results

MALDI-TOF and phylogenic analysis
The spectrum resulting from the 12 clean G. massiliensis spots did not identify bacteria because there was no spectrum match with those of Bruker database. Using 16S rRNA sequence, the phylogenic analysis revealed that G. massiliensis gen. nov., sp.
strain GM3, exhibited 91% identification with Barnesiella viscericola (AB267809.1) (Fig. 1). The similarity of 16S rRNA in these two bacteria led to classification of this new bacterium into the Porphyromonadaceae family, created in 2011 by Krieg (Table 1) [13]. However, this 16S rRNA nucleotide similarity to Barnesiella viscericola was lower than the threshold of 95% recommended by Stackebrandt and Ebers [8] to delineate a new genus. Consequently, G. massiliensis gen. nov., sp. nov., strain GM3, was classified as new genus called Gabonia. The spectrum resulting from G. massiliensis was thus added into the Bruker database (Fig. 2), and a gel view was performed in order to see the spectra differences between other close bacteria (Fig. 3). Then the 16S rRNA sequence of G. massiliensis was deposited in GenBank under accession number LN849789.

FIG. 3. Gel view comparing Gabonia massiliensis strain GM3 with other members of Porphyromonadaceae 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 running spectrum number originating from subsequent spectra loading. Peak intensity is expressed by greyscale scheme code. Color bar and right y-axis indicate relation between color peak is displayed with and peak intensity in arbitrary units. Displayed species are indicated at left.

FIG. 4. Gram staining of Gabonia massiliensis strain GM3.
TABLE 2. Differential characteristics of *G. massiliensis* strain GM3 (data from this study) with *Barnesiella viscericola* strain C46T [31], *Barnesiella intestinihominis* strain YIT 11860 [32], *Coprobacter fastidiosus* strain NSB1 [33] and *Porphyromonas gingivicanis* strain JCM 15907 [34]

| Property                     | G. massiliensis | B. viscericola | B. intestinihominis | C. fastidiosus | P. gingivicanis |
|------------------------------|-----------------|----------------|--------------------|---------------|----------------|
| Cell diameter (μm)           | 0.5             | 0.8–1.6        | 0.4–1.0            | 0.2–0.3       | NA             |
| Oxygen requirement           | −               | −              | −                  | −             | −              |
| Gram stain                   | −               | −              | −                  | −             | −              |
| Mobility                     | −               | −              | −                  | −             | −              |
| Endospore formation          | −               | −              | NA                | −             | −              |
| Indole                       | −               | −              | −                  | −             | −              |
| Production of:               |                 |                |                    |               | NA             |
| Alkaline phosphatase         | +               | +              | +                  | +             | NA             |
| Catalase                     | +               | −              | −                  | +             | NA             |
| Oxidase                      | −               | NA             | −                  | NA            | NA             |
| Urease                       | −               | −              | −                  | −             | −              |
| β-Galactosidase              | +               | +              | +                  | +             | NA             |
| N-acetyl-glucosaminidase     | +               | +              | +                  | +             | NA             |
| Acid from:                   |                 |                |                    |               | NA             |
| D-Arabinose                  | −               | −              | −                  | −             | NA             |
| Ribose                       | −               | −              | −                  | NA            | NA             |
| Mannose                      | +               | +              | +                  | +             | NA             |
| Mannitol                     | −               | −              | −                  | −             | −              |
| Sucrose                      | +               | +              | −                  | −             | −              |
| D-Glucose                    | −               | −              | −                  | +             | +              |
| D-Fructose                   | −               | −              | −                  | −             | −              |
| D-Maltose                    | +               | +              | +                  | +             | NA             |
| G+C content (%)              | 37.9            | 51.7           | 43.9               | 38.3          | 42.7           |
| Habitat                      | Human gut       | Chicken gut    | Human gut          | Human gut     | Beagles        |

NA, not applicable.

**FIG. 6.** Graphical circular map of genome. From outside to center: contigs (red/grey), Clusters of Orthologous Groups (COGs) database category of genes on forward strand (three circles), genes on forward (blue circle) and reverse strands (red circle), COGs category on reverse strand (three circles), GC content.

© 2015 New Microbes and New Infections published by Elsevier Ltd on behalf of European Society of Clinical Microbiology and Infectious Diseases, NMNI, 9, 35–44

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

*G. massiliensis* gen. nov., sp. nov., strain GM3 (= CSUR P1909 = DSM 100571), is a Gram-negative bacillus (Fig. 4). Strain GM3 is a non motile, non-spore-forming, oxidase-negative but catalase-positive bacterium. Growth was observed from 28 to 45°C, with optimal growth at 37°C, and colonies were obtained after 48 hours of culture. *G. massiliensis* grew in salinity conditions of 0 and at a pH of 7. It exhibited translucent colonies with a diameter of 1.5 mm on 5% sheep’s blood-enriched Colombia agar. Strain GM3 is a preferentially anaerobic bacillus but was able to grow in microaerophilic atmosphere and unable to grow in aerobic conditions. Individual cells had a mean diameter of 0.5 μm and a length of 1.25 μm when observed with electron microscopy (Fig. 5).

Using API 20A, we observed positive reactions with D-glucose, D-lactose, D-maltose, D-cellbiose, D-mannose, D-raffinose, D-ribose and D-trehalose but not with D-mannitol, D-sucrose, salicin, D-xyllose, L-arabinose, glycerol, D-melezitose and D-sorbitol. Negative reactions were recorded for indole formation, urease and gelatin using the same gallery. Using API 50CH, we concluded that *G. massiliensis* is able to ferment inositol, methyl-α-glucopyranoside, methyl-α-glucopyranoside, N-acetylglucosamine, esculin ferric citrate, D-cellbiose, D-maltose, D-lactose, D-sucrose, inulin, starch, D-arabitol and potassium-5-ketogluconate. Results from the API ZYM Gallery showed enzymatic activities of alkaline phosphatase, esterase (C4), esterase lipase (C8), acid phosphatase, naphtol-AS-BI phosphohydrolase, α- and β-galactosidase and N-acetyl-β-gluconamidase. Antibiotic susceptibility testing revealed that *G. massiliensis* strain GM3 is sensitive to nitrofurantoin, doxycycline, rifampicin, amoxicillin/clavulanic acid, and imipenem. However, it remained resistant against amoxicillin, vancomycin, clindamycin, gentamicin, erythromycin/metronidazole, tobramycin, trimethoprim/sulfamethoxazole, amikacin and ciprofloxacin. The main phenotypic and biochemical characteristics of *G. massiliensis* compared to the closest bacteria are summarized in Table 2.

**Genome properties**

The genome was 4,261,752 bp long with 37.9% GC content. It was composed of 19 scaffolds (composed of 66 contigs) (Fig. 6). Of the 3,288 predicted genes, 3,227 were protein-coding genes and 61 were RNA genes (two genes were 5S rRNA, two genes were 16S rRNA, two genes were 23S rRNA and 55 genes were tRNA genes). A total of 1,807 genes (56.00%) were assigned as putative function (by COGs or by NR blast), and 113 genes (3.50%) were identified as ORFans (Table 3). The remaining genes were annotated as encoding hypothetical proteins (1,229 genes, 38.08%). Nucleotide content and gene count levels of the genome are summarized in Table 3, while the distribution of genes into COGs functional categories is presented in Table 4. The genome sequence has been deposited in GenBank under accession number CYPV000000000.

**Genome comparison**

The draft genome of *Gabonia massiliensis* strain GM3 was larger in size (4.26 Mb) than *Bacteroides salantronis* strain DSM 18170, *Alistipes shahii* strain WAL 8301, *Alistipes finegoldii* strain DSM 17242, *Barnesiella viscericola* strain DSM 18177 and *Ornithobacterium rhinotracheale* strain ORT-UMN 88 (4.24, 3.76, 3.73, 3.59, and 3.54 Mb, respectively).

---

**Table 3. Nucleotide content and gene count levels of chromosome**

| Attribute               | Genome (total) | % of total |
|-------------------------|----------------|------------|
| Size (bp)               | 4,261,752      | 100        |
| G+C content (bp)        | 1,611,220      | 37.9       |
| Coding sequence size (bp)| 3,718,423     | 87.3       |
| Extrachromosomal elements | 0             | 0          |
| Total genes             | 3,288          | 100        |
| RNA genes               | 61             | 1.9        |
| Protein-coding genes    | 3,227          | 98.1       |
| Genes with function prediction | 1,807       | 55.9       |
| Genes assigned to COGs  | 1,380          | 42.8       |
| Protein with peptide signal | 1,024       | 31.7       |
| Genes with transmembrane helices | 627       | 19.4       |
| No of antibiotic resistant genes | 0      | 0          |
| Genes associated with PKS or NRPS | 7         | 0.21       |
| No. of genes with Pfam-A domains | 2,907   | 88.41      |
| No. of CRISPRs           | 1              | 0.03       |

**Table 4. Number of genes associated with 25 general COGs functional categories**

| Code | Value | % of total | Description |
|------|-------|------------|-------------|
| J    | 138   | 4.3        | Translation |
| A    | 0     | 0.0        | RNA processing and modification |
| K    | 81    | 2.5        | Transcription |
| L    | 87    | 2.7        | Replication, recombination and repair |
| B    | 0     | 0.0        | Chromatin structure and dynamics |
| D    | 13    | 0.4        | Cell cycle control, mitosis and meiosis |
| V    | 39    | 1.2        | Defense mechanisms |
| T    | 26    | 0.8        | Signal transduction mechanisms |
| M    | 119   | 3.7        | Cell wall/membrane biogenesis |
| N    | 0     | 0.0        | Cell motility |
| Z    | 0     | 0.0        | Cytoskeleton |
| W    | 0     | 0.0        | Extracellular structures |
| U    | 16    | 0.50       | Intracellular trafficking and secretion |
| O    | 52    | 1.6        | Posttranslational modification, protein turnover, chaperones |
| C    | 110   | 3.4        | Energy production and conversion |
| G    | 149   | 4.6        | Carbohydrate transport and metabolism |
| E    | 130   | 4.0        | Amino acid transport and metabolism |
| F    | 57    | 1.8        | Nucleotide transport and metabolism |
| H    | 76    | 2.4        | Coenzyme transport and metabolism |
| I    | 50    | 1.5        | Lipid transport and metabolism |
| P    | 73    | 2.3        | Inorganic ion transport and metabolism |
| Q    | 19    | 0.6        | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 206   | 6.4        | General function prediction only |
| S    | 78    | 2.4        | Function unknown |
|      | 572   | 17.0       | Not in COGs |

**Note:** Total is based on either size of genome (bp) or total number of protein coding genes in annotated genome. COGs, Clusters of Orthologous Groups database.
3.08 and 2.4 Mb, respectively) but smaller than Bacteroides fragilis strain YCH46 and Bacteroides dorei strain HSI (5.28 and 5.24 Mb). The G+C content (37.9%) of Gabonia massiliensis strain GM3 is larger than that of Ornithobacterium rhinotracheale strain ORT-UMN 88 (37.2%) but smaller than those of Alistipes shahii strain WAL 8301, Alistipes finegoldii strain DSM 17242, Barnesiella viscerciola strain DSM 18177, Bacteroides fragilis strain YCH46 and Bacteroides dorei strain HSI (57.2, 56.6, 51.7, 46.6, 43.3 and 41.6, respectively). The G+C content (3228) of Gabonia massiliensis strain GM3 is larger compared to those of Ornithobacterium rhinotracheale strain ORT-UMN 88, Barnesiella viscerciola strain DSM 18177, Alistipes shahii strain WAL 8301 and Alistipes finegoldii strain DSM 17242 (2,289, 2,557, 3,090 and 3,226) but smaller than those of Bacteroides fragilis strain YCH46, Bacteroides dorei strain HSI, Bacteroides salanitronis strain DSM 18170 and Bacteroides fragilis (4,670, 4,024 and 3,543). Similarly, the number of protein-coding genes (3,227) of Gabonia massiliensis strain GM3 were higher than those of Ornithobacterium rhinotracheale strain ORT-UMN 88, Barnesiella viscerciola strain DSM 18177, Alistipes shahii strain WAL 8301 and Alistipes finegoldii strain DSM 17242 (2,210, 2,464, 2,996 and 3,108) but lower than those of Bacteroides fragilis strain YCH46, Bacteroides dorei strain HSI and Bacteroides salanitronis strain DSM 18170 (4,578, 4,024 and 3,543). However, the distribution of genes into COGs categories was similar in all genomes compared (Fig. 7). In addition, Gabonia massiliensis strain GM3 shared 3,231, 3,132, 4,269, 4,373, 3,686, 2,499, 1,579 and 2,312 orthologous genes with Alistipes finegoldii strain DSM 17242, Alistipes shahii strain WAL 8301, Barnesiella viscerciola strain DSM 18177, Helicobacter pylori strain 26695 and Ornithobacterium rhinotracheale strain ORT-UMN 88 (Table 5). Although Gabonia massiliensis is a draft genome, making the comparison complex, we propose a summary of the comparison of genome content of Gabonia massiliensis and other bacteria in Table 6.

Among species with standing in nomenclature, AGIOS values ranged from 83.25 between Alistipes shahii strain WAL 8301 and Alistipes finegoldii strain DSM 17242 to 56.41 between Gabonia massiliensis strain GM3 and Alistipes finegoldii strain DSM 17242 (Table 5).
Alistipes finegoldii 100% ± 0.00 30.5% ± 2.99 39.1% ± 2.54 35% ± 2.54 30.1% ± 2.54 25.6% ± 2.54 24.7% ± 2.52
Alistipes shahii 100% ± 0.00 28.8% ± 2.54 45.4% ± 2.54 31.4% ± 2.54 26.7% ± 2.54 25.1% ± 2.52 15.3% ± 2.52
Bacteroides dorei 100% ± 0.00 26.9% ± 2.59 22.3% ± 2.58 23% ± 2.55 23.2% ± 2.54 27.1% ± 2.53 15.6% ± 2.52
Bacteroides fragilis 100% ± 0.00 28.4% ± 2.56 34.8% ± 2.56 24.1% ± 2.54 24.1% ± 2.54 41% ± 2.54 15.4% ± 2.52
Bacteroides salanitronis 100% ± 0.00 30.6% ± 2.58 22.9% ± 2.54 20.9% ± 2.54 38% ± 2.55 31.2% ± 2.52 15.3% ± 2.52
Barnesiella viscericola 100% ± 0.00 38% ± 2.55 23.9% ± 2.52 23.9% ± 2.52 23.9% ± 2.52 17.9% ± 2.52
Gabonia massiliensis 100% ± 0.00 20.9% ± 2.54 38% ± 2.55 23.9% ± 2.52 23.9% ± 2.52 17.9% ± 2.52
Ornithobacterium rhinotracheale 100% ± 0.00 22.9% ± 2.54 38% ± 2.55 23.9% ± 2.52 23.9% ± 2.52 17.9% ± 2.52
Helicobacter pylori 100% ± 0.00 17.9% ± 2.52

To evaluate genomic similarity among the studied strains, we determined two parameters, AGIOS (Table 5) [6], which was designed to be independent from DNA-DNA hybridization (DDH), and digital DDH, which exhibits a high correlation with DDH [29,30] (Table 7).

Conclusion

On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the creation of Gabonia massiliensis gen. nov., sp. nov. that contains the strain GM3. This bacterium was isolated from a stool sample from a healthy Gabonese male youth in Marseille, France.

Taxonomic and nomenclatural proposals

Description of Gabonia gen. nov.

Gabonia (Ga.bo.nia. N.L. gen. n. Gabonia, the Latin name of Gabon, the sub-Saharan African country where the stool specimen was collected).

This is a Gram-negative, non-spore-forming and non motile bacillus. G. massiliensis is preferentially anaerobic. It is oxidase and indole negative but catalase positive. It showed positive activity for alkaline phosphatase, esterase (C4), esterase lipase (C8), acid phosphatase, naphthol-AS-BI-phosphohydrolase, α- and β-galactosidase and N-acetyl-β-glucosaminidase. Its habitat is the human gut. Its type species is Gabonia massiliensis strain GM3.

Description of Gabonia massiliensis gen. nov., sp. nov.

Gabonia massiliensis (ma.si.li.en sis. L. masc. adj. massiliensis, of Massilia, the Latin name of Marseille, where G. massiliensis was isolated).

G. massiliensis exhibited translucent colonies with a 1.5 mm diameter. Individual cells exhibited a diameter of 0.5 μm and a length of 1.25 μm. G. massiliensis is preferentially anaerobic and grows at an optimal temperature of 37°C. It is a Gram-negative, non-spore-forming and nonmotile bacillus. It is oxidase negative but catalase positive. G. massiliensis showed positive reactions for alkaline phosphatase, esterase (C4), esterase lipase (C8), acid phosphatase, naphthol-AS-BI-phosphohydrolase, α- and β-galactosidase and N-acetyl-β-glucosaminidase. It was sensitive for nitrofurantoin, doxycycline, rifampicin, amoxicillin/clavulanic acid and imipenem. G. massiliensis type strain GM3 (= CSUR P1909 = DSM 100571) was isolated from a stool sample of a healthy Gabonese male youth. This strain exhibited a GC content of 37.9%. Its 16S rRNA sequence was deposited in GenBank under accession number LN849789, and the whole
genome shotgun sequence has been deposited in GenBank under accession number CYPV000000000.

Acknowledgements

The authors thank the Xegen Company (www.xegen.fr) for automating the genomic annotation process and C. Andrieu for administrative assistance. Funded by the Méditerranée Infection Foundation.

Conflict of Interest

None declared.

References

[1] Lagier JC, Armoougom F, Million M, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 2012;18:1185–93.
[2] Lagier JC, El Karkouri K, Nguyen TT, Armoougom F, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Aneurococcus senegalensis sp. nov. Stand Genomic Sci 2012;6:116–25.
[3] Lagier JC, Armoougom F, Mishra AK, Nguyen TT, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Alistipes timonensis sp. nov. Stand Genomic Sci 2012;6:315–24.
[4] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Clusteridium sene galense sp. nov. Stand Genomic Sci 2012;6:386–95.
[5] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Pectinophilus timonensis sp. nov. Stand Genomic Sci 2012;7:11–11.
[6] Mishra AK, Lagier JC, Robert C, Raoult D, Fournier PE. Non-contiguous-finished genome sequence and description of Pectinophilus timonensis sp. nov. Stand Genomic Sci 2013;7:370–81.
[7] Tindall BJ, Rosselló-Mora R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. Int J Syst Evol Microbiol 2010;60:249–66.
[8] Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 2006;33:152–5.
[9] Wayne LG, Brenner DJ, Colwell PR, et al. Report of the ad hoc committee on reconciliation of approaches to bacterial systematic. Int J Syst Bacteriol 1987;37:463–4.
[10] Rossello-Mora R. DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In: Stackebrandt E, editor. Molecular identification, systematics, and population structure of prokaryotes. Berlin: Springer; 2006. p. 23–50.
[11] Welker M, Moore ER. Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. Syst Appl Microbiol 2011;34:2–11.
[12] Ramasamy D, Mishra AK, Lagier JC, et al. A polyphasic strategy incorporating genomic data for the taxonomic description of new bacterial species. Int J Syst Evol Microbiol 2014;64:384–91.
[13] Krieg NR. “Family IV. Porphyromonadaceae fam. nov.” In: Krieg NR, Staley JT, Brown DR, et al, editors. Bergey’s manual of systematic bacteriology. 2nd ed. New York: Springer-Verlag; 2011. 4:61.
[14] Lagier JC, Hugon P, Khelifa 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.
[15] Nikamga VD, Huyyn HT, Abboudharam G, Ruym D, Drancourt M. Diversity of human-associated Methanobrevibacter smithii isolates revealed by multiplexar sequence typing. Curr Microbiol 2015;70:810–5.
[16] Benson DA, Karchs-Mizrahi I, Clark K, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucl Acids Res 2012;40:D48–53.
[17] Lowe TM, Eddy SR. rRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 1997;25:955–64.
[18] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rogers T, Ussery DW. Rfam: consistent and rapid annotation of ribosomal RNA genes. Nucl Acids Res 2007;35:3100–8.
[19] Bentdtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[20] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001;305:567–80.
[21] Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast and accurate method for identifying bacterial species. Nucleic Acids Res 2009;37:19–20.
[22] Darling AC, Mau B, Blatner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 2004;14:1394–403.
[23] Lechner M, Findeis S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co)orthologs in large-scale analysis. BMC Bioinformatics 2011;12:124.
[24] Gouret P, Pagnanini J, Dainat J et al. Integration of evolutionary biology concepts for functional annotation and automation of complex research in evolution: the multi-agent software system DAGOBAH. New York: Springer; p. 71–87.
[25] Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, Danchin EG. FigenHX: intelligent automation of genomic annotation: expertise integration in a new software platform. BMC Bioinformatics 2005;6:198.
[26] Auch AF, Van Jan M, Klenk HP, Goker M. Digital DNA-DNA hybridization for microbial species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013;14:60.
[27] Auker A, Fons L, Tassetti M, et al. Genome sequence-based species delimitation with confidence intervals. BMC Bioinformatics 2013;14:60.
[28] Salamot M, Lan PT, Benno Y. Barnesiella viscercalis gen. nov., sp. nov., a novel member of the family Porphyromonadaceae isolated from chicken caecum. Int J Syst Evol Microbiol 2007;57:342–6.
[29] Morotomi M, Nagai F, Sakon H, Tanaka R, Diolateri saccinophilus sp. nov. and Barnesiella intestinohemomas sp. nov., isolated from human faeces. Int J Syst Evol Microbiol 2008;58:2716–20.
[30] Shkurov AN, Khokhlova EV, Chaplin AV, et al. Coprobacter fastidiosus gen. nov., sp. nov., a novel member of the family Porphyromonadaceae isolated from infant faeces. Int J Syst Evol Microbiol 2013;63:4181–8.
[31] Hirassawa M, Takada K. Porphyromonas gingivalis sp. nov. and Porphyromonas creticians sp. nov., isolated from beagles. Int J Syst Bacteriol 1994;44:637–40.