**Clostridium massiliamazoniense** sp. nov., New Bacterial Species Isolated from Stool Sample of a Volunteer Brazilian

Niokhor Dione1,2 · Cheikh Ibrahima Lo2,3 · Didier Raoult1,2 · Florence Fenollar2,3 · Pierre-Edouard Fournier2,3

Received: 27 January 2020 / Accepted: 18 June 2020 / Published online: 1 July 2020
© The Author(s) 2020

**Abstract**

The study of the gut microbiota by the “culturomics concept” permitted us to isolate, from human stool sample, an unknown anaerobic bacterium within the genus *Clostridium* for which we propose the name *Clostridium massiliamazoniense* sp. nov. It was isolated from the fecal flora of a healthy 49-year-old Brazilian male. Here, we describe the characteristics of this organism and its complete genome sequencing and annotation. *Clostridium massiliamazoniense* sp. nov., ND2T (= CSURP1360 = DSMZ 27309) is a Gram-positive, obligate anaerobic member of *Firmicutes* with a 3,732,600 bp-long genome and a G+C content of 27.6%.

**Introduction**

Characterization of bacteria by mass spectrometry and high-throughput genome sequencing has allowed improving their proteomic and genetic characterization [1]. Therefore, it was proposed to include these data into the taxono-genomic description of new bacterial taxa [2, 3]. This strategy combines the analysis of the complete genome sequence, MALDI-TOF spectrum, and main phenotypic features of any bacterium of interest [4].

The genus *Clostridium* (Prazmowski, 1880) was created in 1880. It is classified within the phylum *Firmicutes*. It contains anaerobic rod-shaped bacilli that are able to produce endospores. Members of the *Clostridium* genus are essentially environmental bacteria or belong to the gut flora of mammals. However, several species are major human pathogens, including *C. botulinum*, *C. difficile*, *C. tetani*, and *C. perfringens* [5, 6]. In addition, some species, such as *C. butyricum* and *C. pasteurianum*, are able to fix nitrogen and have agricultural and industrial applications [7, 8]. Herein, we describe a new bacterial species, *Clostridium massiliamazoniense* sp. nov., ND2T using the taxono-genomics approach.

**Material and Methods**

**Organism Informations and Collection**

A stool sample was collected from a 49-year-old volunteer and healthy male living in Brazil. The patient gave an informed and signed consent, and the agreement of the ethics committee of the Institut Fédératif de Recherche 48 (Aix-Marseille University, Marseille, France) was obtained under number 09-022. At the time of sampling, the patient was not under antibiotics. The stool sample was stored at −80 °C until further investigation.

**Strain Isolation and Identification by MALDI-TOF MS and 16S rRNA Sequencing**

Strain ND2 was isolated from stool in November 2013 by culture on 5% sheep blood-enriched Columbia agar (bioMérieux, Marcy l’Étoile, France) at 37 °C in anaerobic atmosphere after an initial thermal shock (15 min at 65 °C for in a dry bath) followed by 3 days of stool sample pre-incubation in an anaerobic blood culture bottle enriched with 5% sheep blood (Thermo Fisher Scientific, Waltham, USA) and 5% bovine rumen fluid (obtained from a local abattoir) filtered on a 0.2 µm filter.
Strain ND2 was screened by MALDI-TOF MS using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany), as previously described [9]. Eighteen distinct deposits were made on MTP 384 MALDI-TOF target plate for strain ND2, from eighteen isolated colonies. Spectra were recorded in the positive linear mode for the mass range of 2000 to 20,000 Da (parameter settings: ion source 1 (IS1), 20 kV; IS2, 18.5 kV; lens, 7 kV). A spectrum was obtained after 675 shots at a variable laser power. The eighteen spectra were imported into the MALDI BioTyper software (version 3.0, Bruker Daltonics) and analyzed by standard pattern matching (with default parameter settings) against the main spectra of 5625 bacteria including 216 spectra from Clostridium species with validly published names that are part of the reference data contained in the BioTyper database. An isolate is considered correctly identified by MALDI-TOF MS at the species level if it has a score of > 1.9 or at the genus level if it only has a score > 1.7. The spectra obtained for strain ND2 were added to our database.

The spectrum of strain ND2 did not match those of any species in the Bruker database. Therefore, we sequenced its 16S rRNA gene as previously described [10].

**Phenotypic Characteristics**

**Culture Conditions**

Different growth temperatures (25, 30, 37, 45 °C) were tested. The growth of strain ND2 was also tested on 5% sheep blood-enriched Columbia agar (bioMérieux) in aerobic atmosphere, with or without 5% CO2, and under anaerobic and microaerophilic conditions using GENbag anaer and GENbag microaer systems (bioMérieux), respectively.

**Biochemical, Sporulation, and Motility Assays**

The biochemical characteristics of strain ND2 were studied using the API Rapid 32A, API ZYM and API 50 CH strips (bioMérieux). The ability to form spores was tested following a thermal shock and the motility assay was performed by direct examination of a fresh colony using a DM 1000 optical microscope (Leica, Nanterre, France) at a × 400 magnification.

**Antibiotic Susceptibility**

The susceptibility of strain ND2 to antimicrobial agents was tested using the disk diffusion method [11]. We tested the following antibiotics: amoxicillin, amoxicillin/clavulanic acid, imipenem clindamycin, doxycycline, rifampicin, tetrnodazole, trimethoprim/sulfamethoxazole, and vancomycin.

**Transmission Electron Microscopy**

A 3.5 µL drop of bacterial suspension was applied for 30 s to the top of a formvar carbon 400 mesh nickel grid (FCF400-Ni, EMS) that was previously glow discharged. After drying using filter paper, bacteria were immediately stained with 1% ammonium molybdate for 1 sec. Electron micrographs were acquired on a Tecnai G20 transmission electron microscope (FEI Company, Limeil-Brevannes, France) operated at 200 keV.

**FAME Analysis by GC/MS**

Three samples were prepared with approximately 50 mg of bacterial biomass harvested from several culture plates. Cellular fatty acid methyl esters were prepared as described by Sasser [12]. GC/MS analyses were carried out as described elsewhere [13].

**Genomic DNA Preparation**

Strain ND2 was cultivated on 5% sheep blood-enriched Columbia agar (bioMérieux, Marcy l’Etoile, France) at 37 °C in anaerobic condition. After 24 h of incubation, colonies grown on three culture plates were collected and suspended in 4 x 100 µL of TE buffer. For cell lysis, 200 µL of this suspension and 2.5 µg/µL lysozyme were added in 1 mL TE buffer and incubated during 30 min at 37 °C. Then, 20 µg/µL proteinase K was added for an overnight incubation at 37 °C. Extracted genomic DNA (gDNA) was then purified as previously described by Lo et al. [17].

**Genome Sequencing and Assembly**

A MiSeq sequencer (Illumina, San Diego, CA, USA) was used for sequencing the genomic DNA (gDNA) of strain ND2. The gDNA was barcoded and pooled 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 fragmentation pattern of gDNA was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) with a DNA 7500 labchip. The sequencing run was performed as reported in previous studies [14–18]. Then, the obtained reads were trimmed and assembled using the CLC genomics Workbench (CLCbio, Seoul, Korea).

**Genome Annotation and Analysis**

The online website Prodigal was used to detect Open Reading Frames (ORFs) (https://prodigalornl.gov/). The
predicted bacterial protein sequences were identified by comparison with the GenBank [19] and Clusters of Orthologous Groups (COG) databases using BLASTP. The RNAs, signal peptides and transmembrane helices were detected using tRNAscan-SE [20], RNaMMer [21], SignalP [22] and TMHMM [23], respectively. The presence of mobile genetic structures was revealed using PHAST [24] and RAST [25]. In order to visualize and manage genomic data, Artemis [26] and DNA Plotter [27] were used successively. The Mauve alignment tool (version 2.3.1) [28] enabled performing multiple genomic sequence alignment between strain ND2 and closely related species. Finally, the average nucleotide identity between genomes was evaluated using the GGDC software [29].

Results

Phylogenetic Analysis

The MALDI-TOF MS score obtained from strain ND2T colonies were < 1.7 and did not enable any identification at the species level. This strain exhibited a 95.4% nucleotide sequence similarity with C. perfringens ATCC 13124T, the phylogenetically closest bacterial species with a validly published name (Fig. 1). This similarity value is lower than the 98.7% threshold used to delineate a new species without carrying out DNA-DNA hybridization [30, 31].

Phenotypic Description

Among the incubation temperatures tested (25, 30, 37, 45 °C) growth was observed at 37 and 45 °C, with an optimal growth at 37 °C, 24 h after inoculation. At 45 °C, growth was observed but was slower. Growth was only observed in anaerobic atmosphere on 5% sheep blood-enriched Columbia agar (bioMérieux, Marcy l’Etoile, France). Colonies were dark gray with irregular edges and had a diameter of 1.0 mm.

Strain ND2T possessed acid phosphatase, naphthol-AS-BI-phosphohydrolase, and galactosidase. It was able to ferment D-glucose and D-maltose. The phenotypic and biochemical characteristics of strain ND2T were compared with those of other related species (Table 1). It was susceptible to amoxicillin, imipenem, doxycycline, rifampicin, metronidazole, and vancomycin. The major fatty acid was octadecanoic acid (14.4% relative abundance). The other abundant fatty acids were mainly saturated species while minor ones were unsaturated (Supplement Table S2).

![Phylogenetic tree](image-url)

**Fig. 1** Phylogenetic tree obtained from the comparison of 16S rRNA sequences highlighting the position of *C. massiliamazoniense* ND2T relative to other type strains within the genus *Clostridium*. GenBank accession numbers are indicated in parentheses. Sequences were aligned using CLUSTALW, and phylogenetic inferences were obtained using the maximum-likelihood method within the MEGA software. Numbers at the nodes are percentages of bootstrap values greater than 70% obtained by repeating 1000 times the analysis to generate a majority consensus tree. *Bacillus subtilis* was used as outgroup. The scale bar represents a 2% nucleotide sequence divergence.
**Table 1** Differential characteristics of *C. massiliamazoniense* ND2<sup>T</sup> (1), *C. dakarense* strain FF1 (2), *C. perfringens* DSM 599<sup>T</sup> (3), *C. liquoris* DSM 100320<sup>T</sup> (4), and *C. frigidicarnis* DSM 12271<sup>T</sup>

| Properties               | 1     | 2     | 3     | 4     |
|--------------------------|-------|-------|-------|-------|
| Cell diameter (µm)       | 0.9   | 1.3   | 0.5–1.0| 1.3–1.6|
| Oxygen requirement       | −     | −     | −     | −     |
| Gram stain               | +     | +     | +     | +     |
| Motility                 | −     | −     | −     | +     |
| Endospore formation      | +     | +     | +     | +     |
| Production of            |       |       |       |       |
| Acid phosphatase         | +     | +     | Na    | Na    |
| Catalase                 | −     | −     | −     | −     |
| Oxidase                  | −     | Na    | Na    | −     |
| Nitrate reductase        | −     | +     | Na    | −     |
| Urease                   | −     | Na    | Na    | −     |
| β-galactosidase          | +     | +     | Na    | Na    |
| N-acetyl-glucosamine     | −     | Na    | Na    | −     |
| Acid from                |       |       |       |       |
| L-Arabinose              | −     | −     | Na    | −     |
| D-Ribose                 | −     | +     | Na    | Na    |
| D-Mannose                | −     | +     | w     | +     |
| D-Mannitol               | −     | −     | +     | −     |
| Succrose                 | −     | +     | Na    | Na    |
| D-glucose                | +     | +     | +     | +     |
| D-fructose               | −     | +     | Na    | +     |
| D-maltose                | +     | +     | +     | +     |
| D-lactose                | −     | +     | +     | −     |
| G+C content (mol%)       | 27.6  | 27    | 34.4  | 27.3  |
| Habitat                  | Human gut | Colonic flora | Fermentation pit | Bovine feces |

*Na* data not available, *w* weak

**Genome Properties**

The genome was 3,732,600-bp long with a G+C content of 27.6% (Fig. 2). It was composed of 12 contigs. Among the 3518 predicted genes, 3403 were protein-coding genes, and 115 were RNAs (10 genes are 5S rRNA, 11 genes are 16S rRNA, 7 genes are 23S rRNA, 83 genes are tRNA genes and 4 genes are ncRNAs). A total of 2305 genes (65.53%) were assigned a putative function (by COGs or by BLAST against NR), and 397 (11.2%) genes were identified as ORFans. The remaining genes were annotated as hypothetical proteins. The statistics of the genome are detailed in Table 2 while the distribution of genes into COG functional categories is presented in Table 3.

**Genome Comparison**

The genome size of *C. massiliamazoniense* ND2<sup>T</sup> (3.7 Mb) was larger than those of *Clostridium perfringens* ATCC 13124, *Clostridium liquoris* DSM 100320, *Clostridium fallax* DSM 2631, and *Clostridium cadaveris* LH052 (3.2, 2.8, 2.7, and 3.5 Mb, respectively) but smaller than those of *Clostridium frigidicarnis* DSM 12271 (4.3 Mb).

Strain ND2 had dDDH values of 28.6, 31.8, 28.8, 28.4, 32.7, and 32.5 with *C. cadaveris* LH052, *C. fallax* DSM 2631, *C. frigidicarnis* DSM 12271, *C. liquoris* DSM 100320, *C. perfringens* ATCC 13124, and *C. tarantellae* DSM 3997. These values far below the recommended 70% threshold value suggest that the strain ND2 is new bacterial species (Table 4).

The degree of genomic similarity of the strain ND2 compared to its related species was calculated with Orthologous ANI Tool (OAT) software [32]. *C. massiliamazoniense* shared with *C. liquoris* the lowest OrthoANI value (68.56%) and with *C. perfringens* the highest value (72.72%). Analysis of all studied genomes showed that 74.34% was the highest value shared between *C. perfringens* and *C. tarantellae* (Fig. 3).

**Conclusion**

We formally propose the creation of *Clostridium massiliamazoniense* sp. nov., that contains strain ND2<sup>T</sup> as type strain, in agreement with the results obtained by phenotypic,
The strain was isolated from the fecal flora of a healthy 49-year-old Brazilian man.

**Taxonomic and Nomenclatural Proposals**

**Description of Clostridium massiliamazoniense sp. nov.**

Clostridium massiliamazoniense (ma.si.li.a.ma.zo.ni.e.n.se. N. L. gen. neut. n. massiliamazoniense, a combination of Massilia, the Latin name of Marseille where strain ND2T was first isolated and characterized, and Amazonia the origin of the patient from whom C. massiliamazoniense was first collected).

Colonies are dark gray with irregular edges and a diameter of 0.6 to 1 mm. Growth is observed at 37 and 45 °C in anaerobic atmosphere only, but is optimal at 37 °C. Cells are Gram-positive rods and have a length ranging from 1.90 to 3.0 μm and a width of 0.8 to 1.0 μm. They are not motile and form subterminal spores.
Catalase and oxidase negative. Indole is produced and nitrate is not reduced. Positive reactions were obtained with esterase (C4), valine arylamidase, α-chymotrypsin, acid phosphatase, naphthol-AS-BI-phosphohydrolase, β-galactosidase, alkaline phosphatase, d-glucose, and d-maltose. The type strain ND2T (= CSUR P1360 = DSM 27309) was isolated from the fecal flora of a healthy 49-year-old Brazilian male living in the Amazonian part of Brazil. The major fatty acids were octadecanoic acid (14.4%) and hexadecanoic acid (12.9%). The genome is 3,732,600 bp long, with a G+C content of 27.6%. The 16S rRNA and genome sequences were deposited in GenBank under accession numbers HG315672 and CYSO00000000, respectively.

---

| Code | Value | % of total | Description |
|------|-------|------------|-------------|
| J    | 156   | 4.58       | Translation |
| A    | 0     | 0.0        | RNA processing and modification |
| K    | 223   | 6.55       | Transcription |
| L    | 158   | 4.64       | Replication, recombination and repair |
| B    | 2     | 0.06       | Chromatin structure and dynamics |
| D    | 30    | 0.88       | Cell cycle control, mitosis and meiosis |
| Y    | 0     | 0.0        | Nuclear structure |
| V    | 66    | 1.94       | Defense mechanisms |
| T    | 125   | 3.67       | Signal transduction mechanisms |
| M    | 148   | 4.35       | Cell wall/membrane biogenesis |
| N    | 55    | 1.62       | Cell motility |
| Z    | 0     | 0.0        | Cytoskeleton |
| W    | 0     | 0.0        | Extracellular structures |
| U    | 41    | 1.20       | Intracellular trafficking and secretion |
| O    | 86    | 2.53       | Post-translational modification, protein turnover, chaperones |
| C    | 154   | 4.53       | Energy production and conversion |
| G    | 170   | 5.0        | Carbohydrate transport and metabolism |
| E    | 229   | 6.73       | Amino acid transport and metabolism |
| F    | 68    | 2.0        | Nucleotide transport and metabolism |
| H    | 103   | 3.03       | Coenzyme transport and metabolism |
| I    | 59    | 1.73       | Lipid transport and metabolism |
| P    | 155   | 4.55       | Inorganic ion transport and metabolism |
| Q    | 28    | 0.82       | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 325   | 9.55       | General function prediction only |
| S    | 228   | 6.70       | Function unknown |
|     | 230   | 6.53       | Not in COGs |

**Table 3** Number of genes associated with the 25 general COG functional categories

| Code | Value | % of total | Description |
|------|-------|------------|-------------|
| J    | 156   | 4.58       | Translation |
| A    | 0     | 0.0        | RNA processing and modification |
| K    | 223   | 6.55       | Transcription |
| L    | 158   | 4.64       | Replication, recombination and repair |
| B    | 2     | 0.06       | Chromatin structure and dynamics |
| D    | 30    | 0.88       | Cell cycle control, mitosis and meiosis |
| Y    | 0     | 0.0        | Nuclear structure |
| V    | 66    | 1.94       | Defense mechanisms |
| T    | 125   | 3.67       | Signal transduction mechanisms |
| M    | 148   | 4.35       | Cell wall/membrane biogenesis |
| N    | 55    | 1.62       | Cell motility |
| Z    | 0     | 0.0        | Cytoskeleton |
| W    | 0     | 0.0        | Extracellular structures |
| U    | 41    | 1.20       | Intracellular trafficking and secretion |
| O    | 86    | 2.53       | Post-translational modification, protein turnover, chaperones |
| C    | 154   | 4.53       | Energy production and conversion |
| G    | 170   | 5.0        | Carbohydrate transport and metabolism |
| E    | 229   | 6.73       | Amino acid transport and metabolism |
| F    | 68    | 2.0        | Nucleotide transport and metabolism |
| H    | 103   | 3.03       | Coenzyme transport and metabolism |
| I    | 59    | 1.73       | Lipid transport and metabolism |
| P    | 155   | 4.55       | Inorganic ion transport and metabolism |
| Q    | 28    | 0.82       | Secondary metabolites biosynthesis, transport and catabolism |
| R    | 325   | 9.55       | General function prediction only |
| S    | 228   | 6.70       | Function unknown |
|     | 230   | 6.53       | Not in COGs |

**Table 4** Pairwise comparison of *C. massiliamazoniense ND2T* with *C. cadaveris* LH052 (*Cca*), *C. fallax* DSM 2631 (*Cfa*), *C. frigidicarnis* DSM 12271 (*Cfr*), *C. liquoris* DSM 100320 (*Cli*), *C. perfingens* ATCC 13124 (*Cpe*), and *C. tarantellae* DSM 3997 (*Cta*) using GGDC, formula 2 (DDH estimates based on identities/HSP length)

| Cca   | Cfa  | Cfr  | Cli  | Cma  | Cpe  | Cta  |
|-------|------|------|------|------|------|------|
| 100%  | 29.2 | 30.8 | 28.6 | 28.6 | 29.3 | 28.9 |
| 100%  | 30.2 | 29.2 | 31.8 | 31.8 | 31.8 | 32.2 |
| 100%  | 28.8 | 28.8 | 29.5 | 29.5 | 29.8 |
| 100%  | 28.4 | 28.4 | 28.7 | 28.7 | 28.9 |
| 100%  | 32.7 | 32.7 | 32.5 |
| 100%  | 33.8 |
| 100%  |      |      |      |      |      |      |
Acknowledgements The authors thank Catherine Robert for sequencing the genome and Aurelia Caputo for submitting the genomic sequence to GenBank and also thank Claudia Andrieu for administrative assistance.

Author Contributions ND and CIL performed the technical characterization on strain ND2 and drafted the manuscript. DR and FF conceived the study and aided to draft the manuscript. PEF conceived the study, participated in its design and coordination, and helped to draft the manuscript. All authors read and approved the final manuscript.

Funding This study was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection, the National Research Agency under the program « Investissements d’avenir», reference ANR-10-IAHU-03, the Région Provence Alpes Côte d’Azur, and European funding FEDER PRIMI.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Seng P, Rolain J-M, Fournier PE, La Scola B, Drancourt M, Raoult D (2010) MALDI-TOF-mass spectrometry applications in clinical microbiology. Future Microbiol 5:1733–1754
2. Ramasamy D, Mishra AK, Lagier J-C, Padmanabhan R, Rossi M, Sentausa E et al (2014) A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. Int J Syst Evol Microbiol 64:384–391
3. Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, Robert C et al (2012) Microbial culturomics: paradigm shift in the human gut microbiome study. Clin Microbiol Infect 18:1185–1193
4. Pfleiderer A, Lagier J-C, Armougom F, Robert C, Vialettes B, Raoult D (2013) Culturomics identified 11 new bacterial species from a single anorexia nervosa stool sample. Eur J Clin Microbiol Infect 32:1471–1481
5. Evaluations and Standards Laboratory (2008) Identification of Clostridium species, p 14
6. Wells CL, Wilkins TD (1996) Clostridia: Spore forming Anaerobic Bacilli. In: Baron S, et al. (eds) Baron’s medical microbiology (4th edn). University of Texas Medical Branch, Galveston
7. Keis S, Shaheen R, Jones DT (2001) Emended descriptions of Clostridium acetobutylicum and Clostridium beijerinckii, and descriptions of Clostridium saccharoperbutylicum sp. nov., and Clostridium saccharobutylicum sp. nov. Int J Syst Evol Microbiol 51(6):2095–2103
8. Carnahan JE, Castle JE (1963) Nitrogen fixation. Annu Rev Plant Physiol 14(1):125–136
9. Lo CI, Fall B, Sambe-Ba B, Diawara S, Gueye MW, Medianikov O et al (2015) MALDI-TOF mass spectrometry: a powerful tool for clinical microbiology at Hôpital Principal de Dakar, Senegal (West Africa). PLoS ONE 10(12):e0145889. https://doi.org/10.1371/journal.pone.0145889
10. Drancourt M, Bollet C, Carlioz A, Martelin R, Gayral JP, Raoult D (2000) 16S ribosomal DNA sequence analysis of a large collection of environmental and clinical unidentifiable bacterial isolates. J Clin Microbiol 38:3623–3630
11. Matuschek E, Brown DFJ, Kahlmeter G (2014) Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. Clin Microbiol Infect 20(4):O25566

12. Sasser M (2006) Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). MIDI, Technical Note, New York, NY

13. Dione N, Sankar SA, Lagier J-C, Khelafia S, Michele C, Armstrong N et al (2016) Genome sequence and description of Anaerosalibacter massiliamazoniense sp. nov. New Microbes New Infect 10:66–76

14. Kokcha S, Mishra AK, Lagier J-C, Million M, Leroy Q, Raoult D (2012) Non contiguous-finished genome sequence and description of Bacillus timonensis sp. nov. Stand Genomic Sci 6(3):346–355

15. Lagier J-C, El Karkouri K, Nguyen T-T, Armougom F, Raoult D, Fournier P-E (2012) Non-contiguous finished genome sequence and description of Anaerococcus senegalensis sp. nov. Stand Genomic Sci 6(1):116–125

16. Mishra AK, Gimenez G, Lagier J-C, Robert C, Raoult D, Fournier P-E (2012) Genome sequence and description of Alistipes senegalensis sp. nov. Stand Genomic Sci 6(3):1–16

17. Lo CI, Sankar SA, Fall B, Sambe-Ba B, Diawara S, Gueye MW et al (2016) High-quality draft genome sequence and description of Haemophilus massiliensis sp. nov. Stand Genomic Sci 11:31

18. Mishra AK, Lagier J-C, Robert C, Raoult D, Fournier P-E (2012) Non-contiguous finished genome sequence and description of Clostridium senegalense sp. nov. Stand Genomic Sci 6(3):386–395

19. Benson DA, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2015) GenBank. Nucleic Acids Res 43(Database issue):D30-5

20. Lowe TM, Eddy SR (1997) RNA scan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964

21. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rogne T, Ussery DW (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35(9):3100–3108

22. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340(4):783–795

23. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305(3):567–580

24. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS (2011) PHAST: a fast phage search tool. Nucleic Acids Res 39.W347–W352

25. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T et al (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206–D214

26. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA et al (2000) Artemis: sequence visualization and annotation. Bioinforma Oxf Engl 16(10):944–945

27. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization. Bioinforma Oxf Engl 25(1):119–120

28. Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14(7):1394–1403

29. Auch AF, von Jan M, Klenk H-P, Göker M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2(1):117–127

30. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinf 14:60

31. Stackebrandt E, Ebers J (2006) Taxonomic parameters revisited: tarnished gold standards. Microbiol Today 33:152–155

32. Lee I, Ouk Kim Y, Park SC, Chun J (2016) OrthoANI: An improved algorithm and software for calculating average nucleotide identity. Int J Syst Evol Microbiol 66:1100–1103

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.