



**HAL**  
open science

## Detailed description of *Senegalia massiliensis* strain SIT17T, a bacterium isolated from the human gut

S.I. Traore, I.I. Ngom, C.I. Lo, F. Di Pinto, C. Sokhna, P.-E. Fournier, D.  
Raoult, F. Fenollar

► **To cite this version:**

S.I. Traore, I.I. Ngom, C.I. Lo, F. Di Pinto, C. Sokhna, et al.. Detailed description of *Senegalia massiliensis* strain SIT17T, a bacterium isolated from the human gut. *New Microbes and New Infections*, 2020, 37, pp.100700. 10.1016/j.nmni.2020.100700 . hal-03159664

**HAL Id: hal-03159664**

<https://hal-amu.archives-ouvertes.fr/hal-03159664>

Submitted on 22 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International  
License

## Detailed description of *Senegalia massiliensis* strain Marseille-P2130<sup>T</sup>, a bacterium isolated from the human gut.

Sory Ibrahima TRAORE<sup>1,2</sup>, Issa Isaac NGOM<sup>1,2</sup>, Cheikh Ibrahima LO<sup>2,3</sup>, Fabrizio DI PINTO<sup>1,2</sup>, Cheikh SOKHNA<sup>2,4</sup>, Pierre-Edouard FOURNIER<sup>2,3</sup>, Didier RAOULT<sup>1,2</sup> and Florence FENOLLAR<sup>2,3\*</sup>

<sup>1</sup> Aix Marseille Univ, IRD, AP-HM, MEΦI, Marseille, France

<sup>2</sup> IHU-Méditerranée Infection, Marseille, France

<sup>3</sup> Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, Marseille, France

<sup>4</sup> Campus Commun UCAD-IRD of Hann, Dakar, Senegal

\* Corresponding author : Florence FENOLLAR, Institut Hospitalo-Universitaire Méditerranée-Infection, 19-21 Boulevard Jean Moulin, 13385, Marseille cedex 05, France ; tel +33 413 732 401 ; fax : +33 413 732 402 ; e-mail: [florence.fenollar@univ-amu.fr](mailto:florence.fenollar@univ-amu.fr).

**Running title:** *Senegalia massiliensis*

**Keywords:** *Senegalia massiliensis*; culturomics; taxonogenomics; human gut; Bacteria

1 **Abstract**

2 Strain Marseille-P2130<sup>T</sup> was isolated from the stool of healthy 13-month-old Senegalese  
3 boy. It is a Gram-positive, aero-anaerobic rod-shaped, non-spore-forming and mobile  
4 bacillus. It exhibited a 92.74% 16S rRNA gene sequence similarity with the *Brassicibacter*  
5 *thermophilus* strain Cel2f, the phylogenetically closest-related species. Its genome is about  
6 2.87 Mb long with 27.39 mol% G+C content. [Therefore, we provide more details of](#)  
7 [Senegalia massiliensis strain Marseille-P2130T \(= CSURP2130 =DSM 103071\) which its](#)  
8 [creation was announced before.](#)

## 9 Introduction

10 Recently, the culturomics concept developed in our laboratory allowed to change the  
11 paradigm of the human gut microbiota [1]. Indeed, by this method, more than 50% of the  
12 microorganisms present in the human gut microbiota are known. In order to improve culture  
13 and bacterial identification, culturomics is associated with a new process named taxono-  
14 genomics to provide exhaustive information and to better characterize bacterial species [4-  
15 5]. Combining phenotypic characteristics and genomic analysis and comparison, this  
16 polyphasic approach exceeds the limits of conventional methods long used for the  
17 description of new species [7-8].

18 Here, we present the classification and features of *Senegalia massiliensis* strain  
19 Marseille-P2130<sup>T</sup>, including a description of the complete genome sequencing and  
20 annotation.

### 21 Isolation and growth conditions

22 Strain Marseille-P2130<sup>T</sup> was first isolated in 2015 from the stool of a healthy 13-  
23 month-old Senegalese boy [11]. The sample was collected in Senegal and was then frozen at  
24 -80°C. Subsequently, it was transported in dry ice to Marseille, where the bacterial culture  
25 was started. The initial growth of bacterial cells was obtained on Columbia agar with 5% sheep's  
26 blood after 2 days of anaerobic incubation at 37°C. The identification of strain Marseille-  
27 P2130<sup>T</sup> using Matrix Assisted Laser Desorption Ionization -Time of Flight Mass Spectrometry  
28 (MALDI-TOF MS) was unsuccessful. The process was performed on a Microflex LT spectrometer  
29 (Bruker, Daltonics, Bremen, Germany) as previously described [15]. The spectra obtained were  
30 imported and analyzed using the Biotyper 3.0 software against the Bruker database permanently  
31 improved with the local MEPHI database (**Figure 1**).

### 32 Strain identification and phylogenetic analysis

33 In order to identify the strain Marseille-P2130<sup>T</sup>, the 16S rRNA gene was amplified  
34 using the fD1 and rP2 primer pair (Eurogentec, Angers, France) and sequenced using the

35 Big Dye® Terminator v1.1 Cycle Sequencing Kit and 3500xLGenetic Analyzer capillary  
36 sequencer (ThermoFisher, Saint-Aubin, France), as previously reported [16]. The 16S rRNA  
37 nucleotide sequences were assembled and corrected using CodonCode Aligner software  
38 (<http://www.codoncode.com>). The polymerase chain reaction-amplified genes coding for  
39 16S rRNA of *Senegalia massiliensis* yielded 92.74% similarity level with *Brassicibacter*  
40 *thermophilus* strain Cel2f (GenBank accession no: NR137216) [9], the phylogenetically  
41 closest species with standing in nomenclature (**Figure 2**). This value was lower than the  
42 95%, the recommended threshold for delineating a new bacterial genus based on 16S rRNA  
43 gene sequence without DNA-DNA hybridization [36]. Classification and general features  
44 are summarized in **Table 1**.

#### 45 **Phenotypic and biochemical characteristics**

46 Colonies of the strain Marseille-P2130<sup>T</sup> were grey and translucent with a size of 0.5-  
47 1 mm on Columbia agar with 5% sheep's blood. Growth was observed from 28 to 45°C,  
48 with an optimal growth at 37°C, and colonies were obtained after 48 hours of culture.  
49 Bacterial cells were Gram-positive, rod-shaped and motile, but non spore forming (**Figure**  
50 **3A**). Observed under electronic microscopy, the cells presented a mean diameter of 0.4 µm  
51 and a mean length of 3.2 µm (**Figure 3B**). Bacterium had a catalase positive but no oxidase  
52 negative activities. *Senegalia massiliensis* is able to grow in an environment with a pH  
53 ranging from 6 to 8.5, with an optimal value of 7. Strain Marseille-P2130<sup>T</sup> is an anaerobic  
54 bacterium that can grow in a microaerophilic atmosphere. On the other hand, no growth was  
55 observed under aerobic conditions. The biochemical and phenotypic features of strain  
56 Marseille-P2130<sup>T</sup> were compared to other close representative strains in the *Clostridiaceae*  
57 family (**Table 2**). Using API ZYM strips (bioMérieux), positive reactions were observed for  
58 esterase, esterase lipase, alkaline phosphatase, α-chymotrypsin, acid phosphatase, naphthol-  
59 AS-BI-phosphohydrolase and β-galactosidase. However, we noted that the enzymatic

60 activities for lipase, leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin,  $\alpha$ -  
61 galactosidase,  $\beta$ -glucuronidase,  $\beta$ -glucosidase, N-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase  
62 and  $\alpha$ -fucosidase, were negative. Using API 50 CH, positives reactions were observed for  
63 glycerol, D-ribose, L-xylose, D-galactose, D-glucose, D-fructose, D-mannose, L-rhamnose,  
64 inositol, D-mannitol, D-sorbitol, methyl  $\alpha$ -D-glucopyranoside, N-acetylglucosamine,  
65 amygdalin, arbutin, salicin, D-cellobiose, D-maltose, D-lactose, D-sucrose, D-trehalose, D-  
66 melezitose, D-raffinose, D-turanose, D-xylose, D-fucose, L-fucose, D-arabitol, potassium  
67 gluconate and starch. However, there was no metabolism for the following carbohydrates:  
68 erythritol, L-arabinose, D-adonitol, methyl  $\beta$ -D-xylopyranoside, methyl  $\alpha$ -D-  
69 mannopyranoside, D-arabinose, inulin and glycogen. Cellular fatty acid methyl esters  
70 (FAME) analysis of the strain Marseille-P2130<sup>T</sup> was carried out by operating gas  
71 chromatography/mass spectrometry (GC/MS) as previously described [22-23]. The result  
72 showed that Hexadecanoic acid (32.6%), 9-Octadecenoic acid (21.6%) and 13-methyl-  
73 tetradecanoic acid (11.9%), are the most abundant fatty. Other saturated and unsaturated  
74 fatty acids are also found (**Table 3**).

### 75 **Genomic properties and comparison**

76 The genome of strain Marseille-P2130 is 2,866,883 bp long with 27.39 mol% G+C  
77 content and contains 2,933 coding genes (**Figure 4**). It is composed of 22 contigs. By  
78 comparing it with others related species, its genome (2.87 Mbp) is smaller than *Alkaliphilus*  
79 *oremlandii* strain OhILAsm, *Proteiniborus ethanoligenes* strain DSM 21650,  
80 *Clostridiisalibacter paucivorans* strain DSM 22131, *Alkaliphilus transvaalensis* strain  
81 ATCC 700919, *Paramaledivibacter caminithermalis* strain DSM 15212, *Alkaliphilus*  
82 *peptidifermentans* strain DSM 18978 and *Alkaliphilus metalliredigens* strain L21-TH-D2  
83 (3.12, 3.16, 3.24, 4.02, 4.05, 4.45 and 4.93 Mbp, respectively), but larger than the genome of  
84 *Caldisalibacter kiritimatiensis* (2.79 Mbp). The G+C content of strain Marseille-P2130<sup>T</sup>

85 (27.39 mol%) is smaller than *A. oremlandii*, *P. ethanoligenes*, *C. paucivorans*, *A.*  
86 *transvaalensis*, *P. caminithermalis*, *A. peptidifermentans*, *A. metalliredigens* and *C.*  
87 *kiritimatiensis* (36.3, 32.6, 31.4, 34, 30.5, 34.1, 36.8 and 30.1 mol%, respectively). The gene  
88 content of strain Marseille-P2130<sup>T</sup> (2,933 genes) is larger than *A. oremlandii* (n=2,898), *C.*  
89 *kiritimatiensis* (n=2,557) and *P. ethanoligenes* (n=2,846), but smaller than *C. paucivorans*  
90 (n=3,014), *A. peptidifermentans* (n=4,072), *A. metalliredigens* (n=4,641), *A. transvaalensis*  
91 (n=3,640) and *P. caminithermalis* (n=3,543). Distribution of functional classes of predicted  
92 genes according to the clusters of orthologous groups (COGs) was reported in Table 4.  
93 Results from pairwise genome comparison obtained from analysis of the digital DNA-DNA  
94 hybridization (dDDH) using GGDC software are shown in Table 5. OrthoANI values among  
95 the closely related species ranged from 64.37%, between *A. metalliredigens* and *C.*  
96 *paucivorans*, to 70.05 % between *C. kiritimatiensis* and *S. massiliensis*. When *S.*  
97 *massiliensis* was compared to these closely species, values ranged from 65.93% with *A.*  
98 *metalliredigens* to 70.05 % with *C. kiritimatiensis* (**Figure 5**).

## 99 **Conclusion**

100 On the basis of phenotypic, phylogenetic and genomic analyses, we formally propose the  
101 creation of *Senegalia massiliensis* gen. nov. sp. nov., that contains the strain Marseille-  
102 P2130<sup>T</sup>. Thus, the combination of culturomics and taxono-genomics would contribute to a  
103 better knowledge of the associated human microorganisms and could help to better  
104 understand physiological functioning in health and disease.

## 105 **Description of *Senegalia* gen. nov.**

106 *Senegalia* (Se.ne.ga.lia. L. gen. n. Senegalia) is the Latin name of Senegal, where the stool  
107 specimen was collected. Cells are Gram-positive, non spore-forming, motile and aero-  
108 anaerobic bacilli. The type species is *Senegalia massiliensis* sp. nov.

## 109 **Description of *Senegalia massiliensis* sp. nov.**

110 *Senegalia massiliensis* gen. nov., sp. nov. (mas.si.li.en'sis. L. fem. adj., from *massiliensis*, of  
111 Massilia), the Latin name of Marseille where the strain was first isolated. It is classified as a  
112 member of the family *Clostridiaceae* within the phylum *Firmicutes*. The strain Marseille-  
113 P2130<sup>T</sup> designed the type strain of *Senegalia massiliensis* gen. nov., sp. nov., and was  
114 deposited in CSUR (CSURP2130) and DSMZ (DSM 103071) collections. It is Gram  
115 positive bacilli, motile, catalase positive, oxidase negative and non-spore forming. Strain  
116 Marseille-P2130<sup>T</sup> was first isolated from the stool of a healthy 13-month-old Senegalese  
117 boy. Its genome is 2,866,883 bp long with 27.39 mol% G+C content and possesses 2933  
118 coding genes. The genome and 16S rRNA sequences of the strain Marseille-P2130<sup>T</sup> are both  
119 deposited in GenBank under accession numbers UZAQ000000000 and LN881608,  
120 respectively.



121 **Acknowledgements**

122 The authors thank Catherine Robert for sequencing the genome and Aurelia Caputo for  
123 submitting the genomic sequence to GenBank.

124 **Conflict of interest**

125 The authors declare no conflict of interest.

126 **Funding sources**

127 This study was supported by the Institut Hospitalo-Universitaire (IHU) Méditerranée  
128 Infection, the National Research Agency under the program « Investissements d'avenir »,  
129 reference ANR-10-IAHU-03, the Région Provence-Alpes-Côte d'Azur and European  
130 funding FEDER PRIM1.

131 **Ethics and consent**

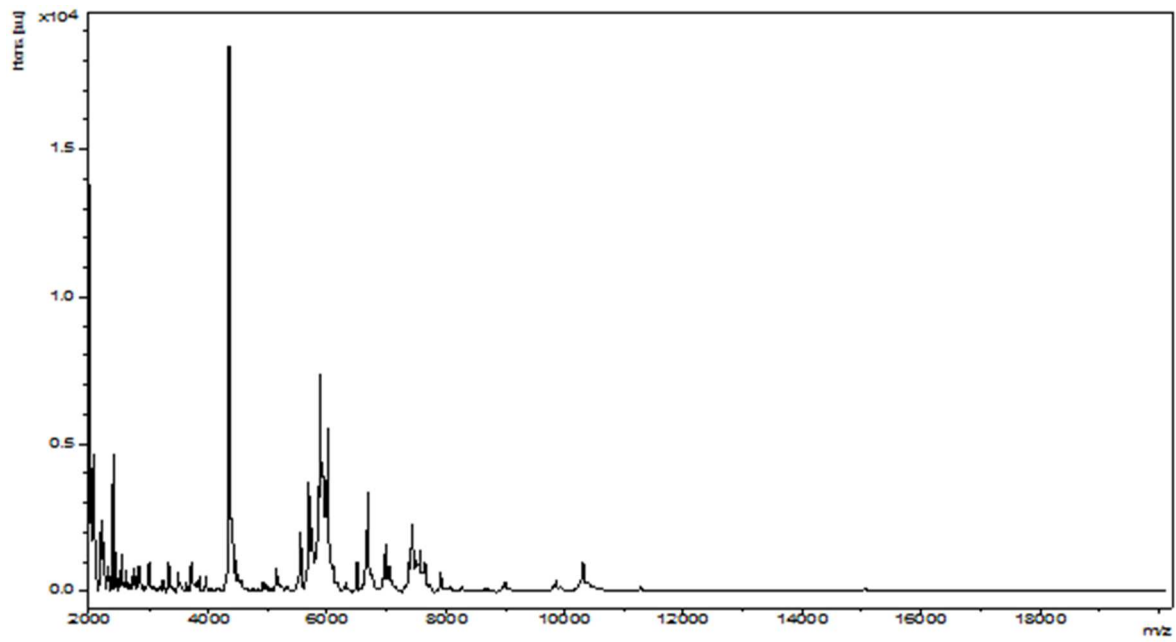
132 The child's parents provided a signed informed consent and the study was approved by the  
133 ethics committee of the Institut Fédératif de Recherche IFR48 under number 09-022.

134 **References**

- 135 1. Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial  
136 culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* 2012;  
137 18(12):1185-93. doi: 10.1111/1469-0691.12023.
- 138 2. Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously  
139 uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* 2016;  
140 1:16203. doi: 10.1038/nmicrobiol.2016.203.
- 141 3. Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, La Scola B, Raoult D. The rebirth of culture  
142 in microbiology through the example of culturomics to study human gut microbiota. *Clin*  
143 *Microbiol Rev* 2015; 28(1):237-64. doi: 10.1128/CMR.00014-14.
- 144 4. Seng P, Drancourt M, Gouriet F, La Scola B, Fournier P-E, Rolain J-M, et al. Ongoing  
145 Revolution in Bacteriology: Routine Identification of Bacteria by Matrix-Assisted Laser  
146 Desorption Ionization Time-of-Flight Mass Spectrometry. *Clin Infect Dis* 2009; 49 (4):543-  
147 51. doi: 10.1086/600885.
- 148 5. Fournier PE, Drancourt M. New Microbes New Infections promotes modern prokaryotic  
149 taxonomy: a new section "TaxonoGenomics: new genomes of microorganisms in humans".  
150 *New Microbes New Infect* 2015; 7:48-9. doi: 10.1016/j.nmni.2015.06.001.
- 151 6. Fournier P-E, Lagier J-C, Dubourg G, Raoult D. From culturomics to taxonomogenomics: A  
152 need to change the taxonomy of prokaryotes in clinical microbiology. *Anaerobe* 2015;  
153 36:73-8. doi: 10.1016/j.anaerobe.2015.10.011.
- 154 7. Ramasamy D, Mishra AK, Lagier JC, Padhmanabhan R, Rossi M, Sentausa E, et al. A  
155 polyphasic strategy incorporating genomic data for the taxonomic description of novel  
156 bacterial species. *Int J Syst Evol Microbiol* 2014; 64 (Pt 2):384-91. doi:  
157 10.1099/ijs.0.057091-0.
- 158 8. Vandamme P, Pot B, Gillis M, de Vos P, Kersters K, Swings J. Polyphasic taxonomy, a  
159 consensus approach to bacterial systematics. *Microbiol Rev.* 1996; 60 (2):407-38.
- 160 9. Wang B, Ji SQ, Tian XX, Qu LY, Li FL. *Brassicibacter thermophilus* sp. nov., a  
161 thermophilic bacterium isolated from coastal sediment. *Int J Syst Evol Microbiol* 2015;  
162 65(9):2870-4. doi: 10.1099/ijs.0.000348.
- 163 10. Rezgui R, Ben Ali Gam Z, Ben Hamed S, Fardeau ML, Cayol JL, Maaroufi A, et al.  
164 *Sporosalibacterium faouarense* gen. nov., sp. nov., a moderately halophilic bacterium  
165 isolated from oil-contaminated soil. *Int J Syst Evol Microbiol* 2011; 61(Pt 1):99-104. doi:  
166 10.1099/ijs.0.017715-0.
- 167 11. Traore SI, Cadoret F, Fournier P-E, Raoult D. "*Senegalia massiliensis*," a new bacterium  
168 isolated from the human gastrointestinal tract. *New Microbes New Infect* 2016; 12:88-9. doi:  
169 10.1016/j.nmni.2016.05.005.
- 170 12. Seng P, Rolain JM, Fournier P-E, La Scola B, Drancourt M, Raoult D. MALDI-TOF-mass  
171 spectrometry applications in clinical microbiology. *Future Microbiol* 2010; 5 (11):1733-54.  
172 doi: 10.2217/fmb.10.127.
- 173 13. Fall B, Lo CI, Samb-Ba B, Perrot N, Diawara S, Gueye MW et al. The ongoing revolution of  
174 MALDI-TOF mass spectrometry for microbiology reaches tropical Africa. *Am J Trop Med*  
175 *Hyg* 2015; 92(3):641-7. doi: 10.4269/ajtmh.14-0406.
- 176 14. Lo CI, Sankar SA, Fall B, Sambe-Ba B, Diawara S, Gueye MW. High-quality draft genome  
177 sequence and description of *Haemophilus massiliensis* sp. nov. *Stand Genomic Sci.* 2016;  
178 11:31. doi: 10.1186/s40793-016-0150-1.

- 179 15. Sankar SA, Lagier JC, Pontarotti P, Raoult D, Fournier PE. The human gut microbiome, a  
180 taxonomic conundrum. *Syst Appl Microbiol*. 2015; 38(4):276-86. doi:  
181 10.1016/j.syapm.2015.03.004.
- 182 16. Stackebrandt E, Ebers J. Taxonomic parameters revisited : tarnished gold standards.  
183 *Microbiol Today*. 2006; 6:152–5.
- 184 17. Diop K, Diop A, Levasseur A, Mediannikov O, Robert C, Armstrong N, Couderc C.  
185 Microbial Culturomics Broadens Human Vaginal Flora Diversity: Genome Sequence and  
186 Description of *Prevotella lascolaii* sp. nov., isolated from a Patient with Bacterial Vaginosis.  
187 *OMICS*. 2018; 22(3):210-222. doi: 10.1089/omi.2017.0151.
- 188 18. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.  
189 *Nucleic Acids Res* 2004; 32(5):1792-7.
- 190 19. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for  
191 large alignments. *PLoS One* 2010; 5(3):e9490. doi: 10.1371/journal.pone.0009490.
- 192 20. Citron DM, Ostovari MI, Karlsson A, Goldstein EJ. Evaluation of the E test for  
193 susceptibility testing of anaerobic bacteria. *J Clin Microbiol* 1991; 29(10):2197-203.
- 194 21. Matuschek E, Brown DF, Kahlmeter G. Development of the EUCAST disk diffusion  
195 antimicrobial susceptibility testing method and its implementation in routine microbiology  
196 laboratories. *Clin Microbiol Infect*. 2014; 20(4):O255-66. doi: 10.1111/1469-0691.12373.
- 197 22. Sasser M (2006). Bacterial identification by gas chromatographic analysis of fatty acids  
198 methyl esters (GC-FAME). [http://midi-inc.com/pdf/MIS\\_Technote\\_101.pdf](http://midi-inc.com/pdf/MIS_Technote_101.pdf).
- 199 23. Dione N, Sankar SA, Lagier JC, Khelaifia S, Michele C, Armstrong N, et al. Genome  
200 sequence and description of *Anaerosalibacter massiliensis* sp. nov. *New Microbes New*  
201 *Infect*. 2016; 10:66-76. doi: 10.1016/j.nmni.2016.01.002.
- 202 24. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic  
203 gene recognition and translation initiation site identification. *BMC Bioinformatics*  
204 2010;11:119. doi: 10.1186/1471-2105-11-119.
- 205 25. Lowe TM, Eddy SR. tRNAscan-SE: A Program for Improved Detection of Transfer RNA  
206 Genes in Genomic Sequence. *Nucleic Acids Res*. 1997; 25(5):955-64.
- 207 26. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. RNAmmer:  
208 consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;  
209 35(9):3100-8.
- 210 27. Dyrlov Bendtsen J, Nielsen H, von Heijne G, Brunak S. Improved Prediction of Signal  
211 Peptides: SignalP 3.0. *J Mol Biol*. 2004; 340(4):783-95.
- 212 28. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein  
213 topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;  
214 305(3):567-80.
- 215 29. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis:  
216 sequence visualization and annotation. *Bioinformatics* 2000; 16(10):944-5.
- 217 30. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear  
218 interactive genome visualization. *Bioinformatics*. 2009; 25(1):119-20. doi:  
219 10.1093/bioinformatics/btn578.
- 220 31. Gouret P, Paganini J, Dainat J, Louati D, Darbo E, Pontarotti P, et al. Integration of  
221 evolutionary biology concepts for functional annotation and automation of complex research  
222 in evolution: the multi-agent software system DAGOBAH. In: *Evolutionary biology-*  
223 *concepts biodiversity, macroevolution and genome evolution* (p. 71-87). Editions Springer.  
224 DOI: 10.1007/978-3-642-20763-1\_5.

- 225 32. Gouret P, Vitiello V, Balandraud N, Gilles A, Pontarotti P, Danchin EG. FIGENIX:  
226 Intelligent automation of genomic annotation: expertise integration in a new software  
227 platform. BMC Bioinformatics 2005; 6: 198.
- 228 33. Lee I, Ouk Kim Y, Park SC, Chun J. OrthoANI: An improved algorithm and software for  
229 calculating average nucleotide identity. Int J Syst Evol Microbiol 2016; 66(2):1100-1103.  
230 doi: 10.1099/ijsem.0.000760.
- 231 34. Auch AF, Jan M von, Klenk HP, Göker M. Digital DNA-DNA hybridization for microbial  
232 species delineation by means of genome-to-genome sequence comparison. Stand Genomic  
233 Sci. 2010; 2(1):117-34. doi: 10.4056/sigs.531120.
- 234 35. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species  
235 delimitation with confidence intervals and improved distance functions. BMC  
236 Bioinformatics 2013; 14:60. doi: 10.1186/1471-2105-14-60.
- 237 36. Meier-Kolthoff JP, Klenk HP, Göker M. Taxonomic use of DNA G+C content and DNA-  
238 DNA hybridization in the genomic age. Int J Syst Evol Microbiol 2014; 64(Pt 2):352-6. doi:  
239 10.1099/ijms.0.056994-0.
- 240 37. Skerman VBD, McGowan V and Sneath PHA. Approved Lists of Bacterial Names. Int J  
241 Syst Bacteriol. 1980; 30:225-420.
- 242 38. Liebgott PP, Joseph M, Fardeau ML, et al. *Clostridiisalibacter paucivorans* gen. nov., sp.  
243 nov., a novel moderately halophilic bacterium isolated from olive mill wastewater. Int J Syst  
244 Evol Microbiol. 2008;58(Pt 1):61–67. doi:10.1099/ijms.0.65182-0
- 245 39. Takai K, Moser DP, Onstott TC, Spoelstra N, Pfiffner SM, Dohnalkova A, Fredrickson JK.  
246 *Alkaliphilus transvaalensis* gen. nov., sp. nov., an Extremely Alkaliphilic Bacterium Isolated  
247 From a Deep South African Gold Mine. Int J Syst Evol Microbiol 2001; 51 (Pt 4), 1245-56.
- 248 40. Niu L, Song L, Dong X. *Proteiniborus ethanoligenes* gen. nov., sp. nov., an anaerobic  
249 protein-utilizing bacterium. Int J Syst Evol Microbiol. 2008; 58(Pt 1):12–16.  
250 doi:10.1099/ijms.0.65108-0.
- 251 41. Rezgui R, Ben Ali Gam Z, Ben Hamed S, et al. *Sporosalibacterium faouarensense* gen. nov.,  
252 sp. nov., a moderately halophilic bacterium isolated from oil-contaminated soil. Int J Syst  
253 Evol Microbiol. 2011; 61(Pt 1):99–104. doi:10.1099/ijms.0.017715-0.



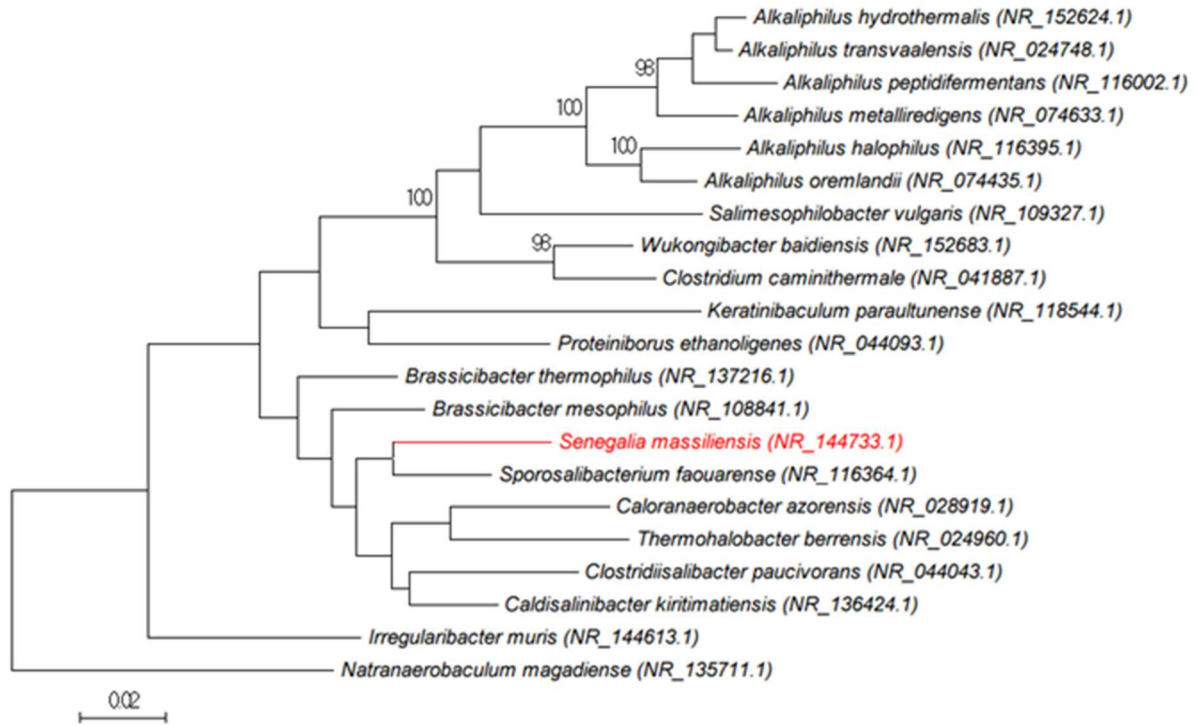
255

256

**Figure 1.** Reference mass spectrum from *Senegalia massiliensis* strain Marseille-

257

P2130<sup>T</sup>.



258

259 **Figure 2.** Phylogenetic tree highlighting the position of *Senegalia massiliensis* strain

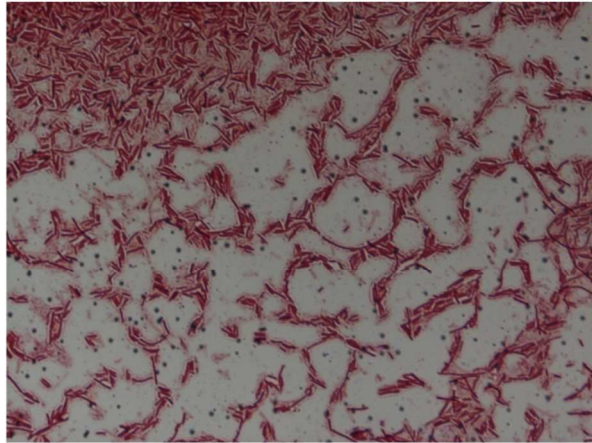
260 Marseille-P2130 relative to other close species. Sequences were aligned using CLUSTALW,

261 with default parameters, and phylogenetic inferences were obtained using the maximum

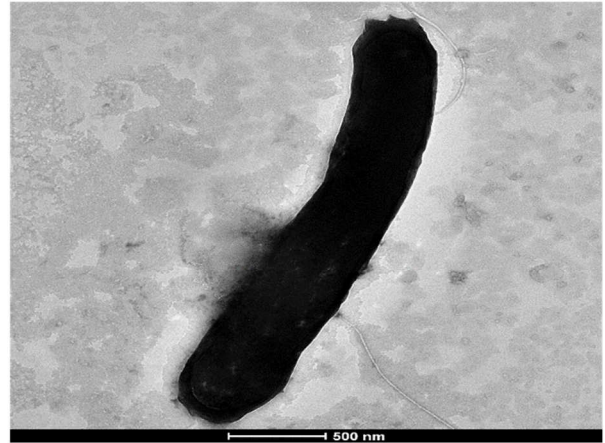
262 likelihood method within the MEGA7 software. Numbers at the nodes are percentages of

263 bootstrap values obtained by repeating the analysis 500 times to generate a majority

264 consensus tree. The scale bar indicates a 2% nucleotide sequence divergence.



**A**



**B**

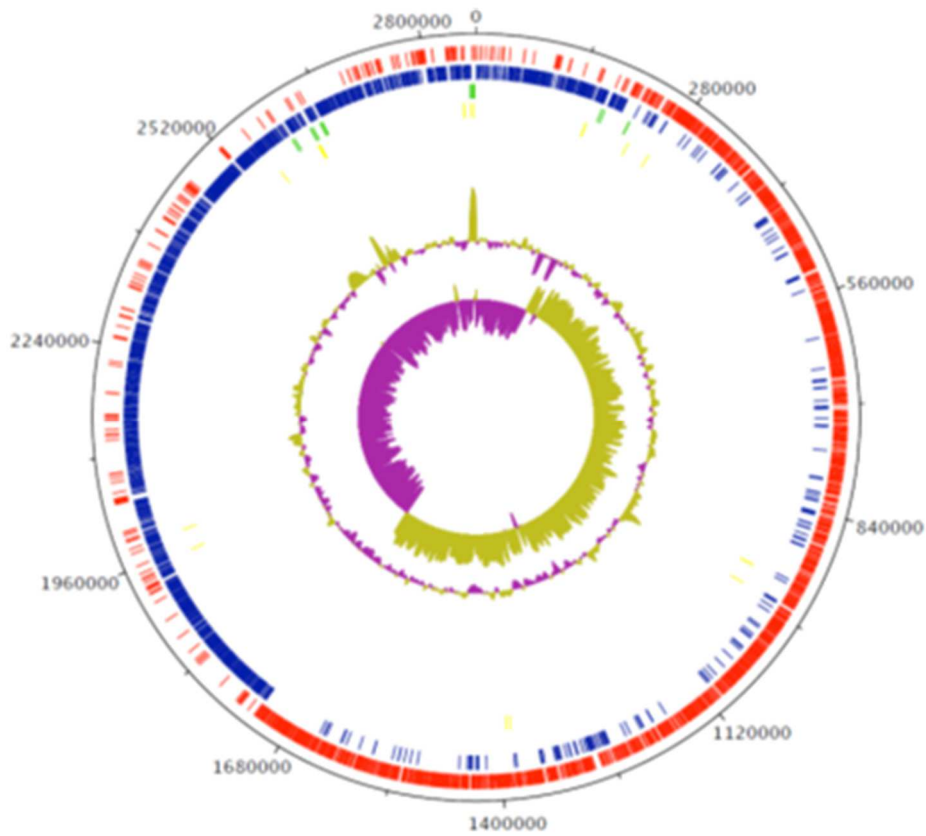
265

266 **Figure 3.** The morphology of bacterial cells of strain Marseille-P2130:

267 A. Gram staining of *Senegalia massiliensis* strain Marseille-P2130.

268 B. Transmission electron microscopy of *Senegalia massiliensis* strain Marseille-P2130 using

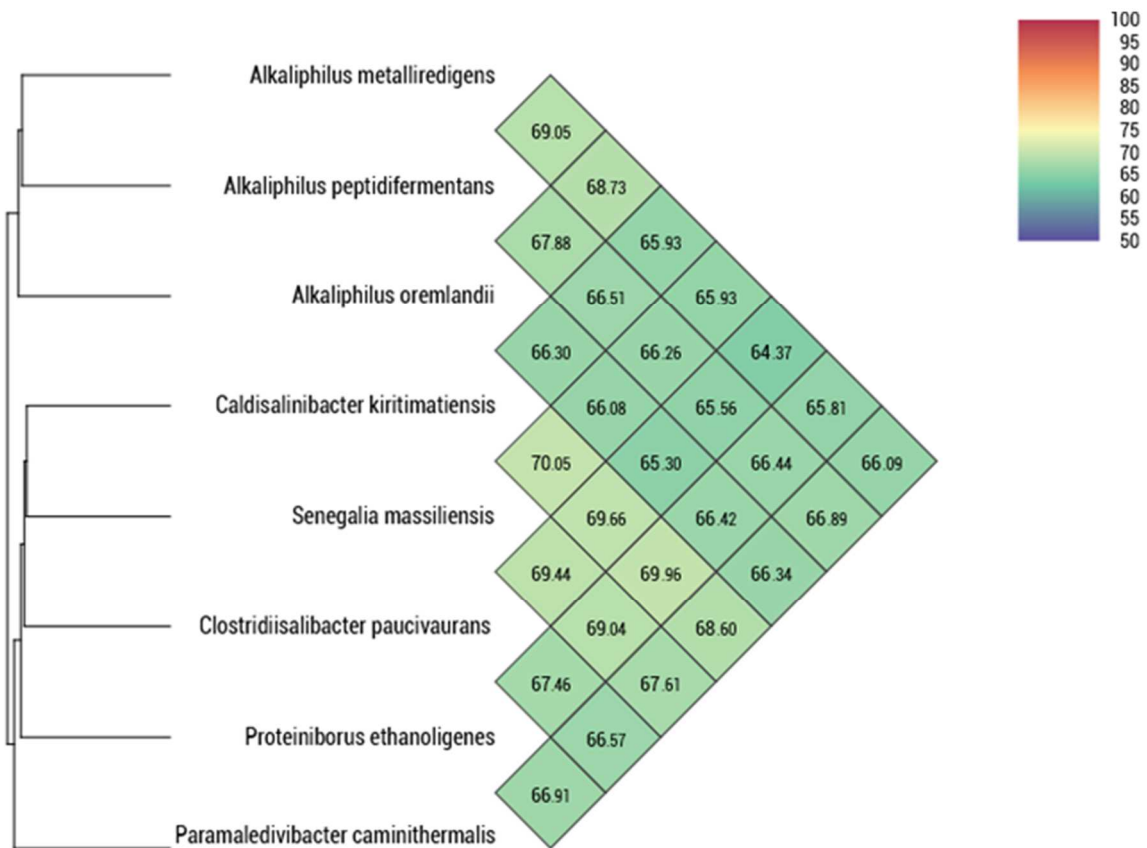
269 Tecnai G20 electron microscope (FEI Company). The scale bar is equal to 500 nm.



270

271 **Figure 4.** Graphical circular map of the chromosome. From outside to the center: Contigs  
 272 (red), COG category of genes on the forward strand (three circles), genes on forward strand  
 273 (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand  
 274 (three circles), G+C content.





275

276 **Figure 5.** Heatmap generated with OrthoANI values calculated using the OAT software

277 between *Senegalia massiliensis* and other closely related species with standing in

278 nomenclature.

279 **Table 1.** Classification and general features of *Senegalia massiliensis* strain Marseille-P2130.

<b>Property</b>	<b>Terms</b>
Current Classification	Domain : <i>Bacteria</i> Phylum : <i>Firmicutes</i> Class : <i>Clostridia</i> Order : <i>Clostridiales</i> Family : <i>Clostridiaceae</i> Genus : <i>Senegalia</i> Species : <i>Senegalia massiliensis</i> Type: strain Marseille-P2130
Gram stain	Positive
Cell shape	Rod
Motility	Motile
Sporulation	No sporulating
Temperature range	28-45°C
Optimum temperature	37°C
pH range (optimum)	7
Oxygen requirement	Anaerobic
Carbone source	Unknown
Habitat	Human gut
Biotic relationship	Free-living
Pathogenicity	Unknown

**Table 2.** Differential characteristics of *Senegalia massiliensis* strain Marseille-P2130 (data from this study) compared to other close bacteria

<b>Properties</b>	<i>Senegalia massiliensis</i>	<i>Clostridiisalibacter paucivorans</i>	<i>Alkaliphilus oremlandii</i>	<i>Alkaliphilus transvaalensis</i>	<i>Proteiniborus ethanoligenes</i>	<i>Sporosalibacterium faouarensense</i>
Cell diameter (µm)	0.3-0.5	0.5	0.5	0.4–0.7	0.5-0.6	0.5
Oxygen requirement	Aero-anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic	Anaerobic
Shape	bacilli	bacilli	bacilli	bacilli	bacilli	bacilli
Gram stain	+	+	+	+	+	+
Motility	+	-	+	+	-	+
Indole	-	-	NA	+	-	NA
<b>Production of</b>						
Alkaline phosphatase	+	NA	NA	NA	NA	NA
Catalase	+	NA	NA	NA	NA	NA
Oxidase	-	NA	NA	NA	NA	NA
Nitrate reductase	-	NA	-	+	+	NA
Urease	-	NA	NA	NA	NA	NA
β-galactosidase	+	NA	NA	NA	NA	NA
N-acetyl glucosamine	-	NA	NA	NA	NA	NA
<b>Acid from</b>						
L-arabinose	-	+	NA	NA	-	+
Ribose	+	-	NA	-	-	-
Mannose	+	-	NA	-	-	-
Mannitol	+	NA	NA	NA	-	+
D-glucose	-	NA	NA	-	-	+
D-fructose	+	-	+	-	-	+
D-maltose	+	NA	NA	-	-	-
D-lactose	+	NA	NA	-	-	-
G+C content (%)	27.4	33.0	36.1	36.4	38.0	37.7
Habitat	Human colon	Wastewater	Environment	Environment	Environment	Soil

282 **Table 3.** Cellular fatty acid profiles (%) of *Senegalia massiliensis* strain Marseille-P2130 compared  
 283 with other species. **1**, *Senegalia massiliensis* strain Marseille-P2130<sup>T</sup>; **2**, *Clostridiisalibacter*  
 284 *paucivorans* strain 37HS60<sup>T</sup> [38]; **3**, *Alkaliphilus transvaalensis* strain SAGM1<sup>T</sup> [39]; **4**,  
 285 *Proteiniborus ethanoligenes* strain GW<sup>T</sup> [40]; **5**, *Sporosalibacterium faouarense* strain SOL3f37<sup>T</sup> [41];  
 286 TR, trace amounts < 1%. -, not detected;

<b>Fatty acids</b>	<b>Names</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
12:00	Dodecanoic acid	1.2	-	-	-	-
13:00	Tridecanoic acid	TR	-	2.3	-	4.4
14:00	Tetradecanoic acid	9.2	14.3	1.7	15.58	21.6
15:0 anteiso	12-methyl-tetradecanoic acid	1.0	1.5	2.8	-	3.9
15:0 iso	13-methyl-tetradecanoic acid	11.9	6.6	51.6	4.30	41
16:00	Hexadecanoic acid	32.6	7.6	3.9	25.40	1.2
16:1n5	11-Hexadecanoic acid	TR	-	1.9	6.18	-
17:00	Heptadecanoic acid	TR	-	-	-	0.6
17:1n7	10-Heptadecenoic acid	TR	19.3	12.2	9.49	-
18:00	Octadecanoic acid	4.7	-	7.2	12.03	1.3
18:1n7	11-Octadecenoic acid	1.7	-	2.0	-	-
18:1n9	9-Octadecenoic acid	21.6	-	1.1	11.20	-

287 **Table 4.** Distribution of functional classes of predicted genes according to the clusters of orthologous  
 288 groups (COG) of proteins of *Senegalia massiliensis* strain Marseille-P2130.

289

<b>Code</b>	<b>Value</b>	<b>Description</b>
[J]	244	Translation, ribosomal structure and biogenesis
[A]	0	RNA processing and modification
[K]	225	Transcription
[L]	114	Replication, recombination and repair
[B]	1	Chromatin structure and dynamics
[D]	52	Cell cycle control, cell division, chromosome partitioning
[Y]	0	Nuclear structure
[V]	80	Defense mechanisms
[T]	181	Signal transduction mechanisms
[M]	148	Cell wall/membrane/envelope biogenesis
[N]	71	Cell motility
[Z]	0	Cytoskeleton
[W]	11	Extracellular structures
[U]	28	Intracellular trafficking, secretion, and vesicular transport
[O]	116	Posttranslational modification, protein turnover, chaperones
[X]	26	Mobilome: prophages, transposons
[C]	173	Energy production and conversion
[G]	144	Carbohydrate transport and metabolism
[E]	202	Amino acid transport and metabolism
[F]	92	Nucleotide transport and metabolism
[H]	124	Coenzyme transport and metabolism
[I]	89	Lipid transport and metabolism
[P]	145	Inorganic ion transport and metabolism
[Q]	34	Secondary metabolites biosynthesis, transport and catabolism
[R]	255	General function prediction only
[S]	203	Function unknown
—	563	Hypothetical protein

290

291 **Table 5.** Pairwise comparison of *Senegalia massiliensis* strain Marseille-P2130 with other species using GGDC formula 2 (DDH estimates based on identities  
 292 / HSP length)\*.

	<i>Senegalia massiliensis</i>	<i>Alkaliphilus metalliredigens</i>	<i>Alkaliphilus oremlandii</i>	<i>Alkaliphilus transvaalensis</i>	<i>Alkaliphilus peptidifermentans</i>	<i>Proteiniborus ethanoligenes</i>	<i>Clostridiisalibacter paucivorans</i>	<i>Paramaledivibacter caminithermalis</i>	<i>Caldsalinibacter kiritimatiensis</i>
<i>Caldsalinibacter kiritimatiensis</i>	19.7±2.4	37.2±5	34±4.9	17.5±4.4	18±4.5	18.9±4.6	16.9±4.4	23.6±4.7	100%
<i>Paramaledivibacter caminithermalis</i>	15.2±4.3	28.4±4.9	26.4±4.8	16±4.3	19.5±4.6	19±4.5	23.6±4.7	100%	
<i>Clostridiisalibacter paucivorans</i>	17.5±4.5	25±4.8	16.7±4.4	31±4.9	17.8±5.4	18.6±5.5	100%		
<i>Proteiniborus ethanoligenes</i>	17.3±4.1	19.9±4.6	19.8±4.6	18.7±4.6	16.2±4.3	100%			
<i>Alkaliphilus peptidifermentans</i>	17.2±4.4	25.5±5.2	24.8±4.8	19.9±4.7	100%				
<i>Alkaliphilus transvaalensis</i>	27.4±4.8	23.5±4.8	22.2±4.7	100%					
<i>Alkaliphilus oremlandii</i>	29.9±4.9	26.8±4.9	100%						
<i>Alkaliphilus metalliredigens</i>	33.9±4.9	100%							
<i>Senegalia massiliensis</i>	100%								

293 \*The confidence intervals indicate the inherent uncertainty in estimating DDH values from intergenomic distances based on models derived from empirical  
 294 test data sets (which are always limited in size) These results are in accordance with the 16S rRNA and phylogenomic analyses as well as the GGDC results