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

Version of Record: https://www.sciencedirect.com/science/article/pii/S2052297520300524
Manuscript_880a65660b134064ab9e634779c56a70

Detailed description of *Senegalia massiliensis* strain Marseille-P2130^{T}, a bacterium isolated from the human gut.

Sory Ibrahima TRAORE[1,2], Issa Isaac NGOM[1,2], Cheikh Ibrahima LO[2,3], Fabrizio DI PINTO[1,2], Cheikh SOKHNA[2,4], Pierre-Edouard FOURNIER[2,3], Didier RAOULT[1,2] and Florence FENOLLAR[2,3*]

[1] Aix Marseille Univ, IRD, AP-HM, MEΦI, Marseille, France

[2] IHU-Méditerranée Infection, Marseille, France

[3] Aix Marseille Univ, IRD, AP-HM, SSA, VITROME, Marseille, France

[4] 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.

Running title: *Senegalia massiliensis*

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

Abstract

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

Introduction

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

Here, we present the classification and features of *Senegalia massiliensis* strain Marseille-P2130^T, including a description of the complete genome sequencing and annotation.

Isolation and growth conditions

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

Strain identification and phylogenetic analysis

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

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

Phenotypic and biochemical characteristics

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

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

Genomic properties and comparison

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

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

Conclusion

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

Description of *Senegalia* gen. nov.

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

Description of *Senegalia massiliensis* sp. nov.

Senegalia massiliensis gen. nov., sp. nov. (mas.si.li.en'sis. L. fem. adj., from *massiliensis*, of Massilia), the Latin name of Marseille where the strain was first isolated. It is classified as a member of the family *Clostridiaceae* within the phylum *Firmicutes*. The strain Marseille-P2130^{T} designed the type strain of *Senegalia massiliensis* gen. nov., sp. nov., and was deposited in CSUR (CSURP2130) and DSMZ (DSM 103071) collections. It is Gram positive bacilli, motile, catalase positive, oxidase negative and non-spore forming. Strain Marseille-P2130^{T} was first isolated from the stool of a healthy 13-month-old Senegalese boy. Its genome is 2,866,883 bp long with 27.39 mol% G+C content and possesses 2933 coding genes. The genome and 16S rRNA sequences of the strain Marseille-P2130^{T} are both deposited in GenBank under accession numbers UZAQ00000000 and LN881608, respectively.

Acknowledgements

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

Conflict of interest

The authors declare no conflict of interest.

Funding sources

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.

Ethics and consent

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

References

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

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

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

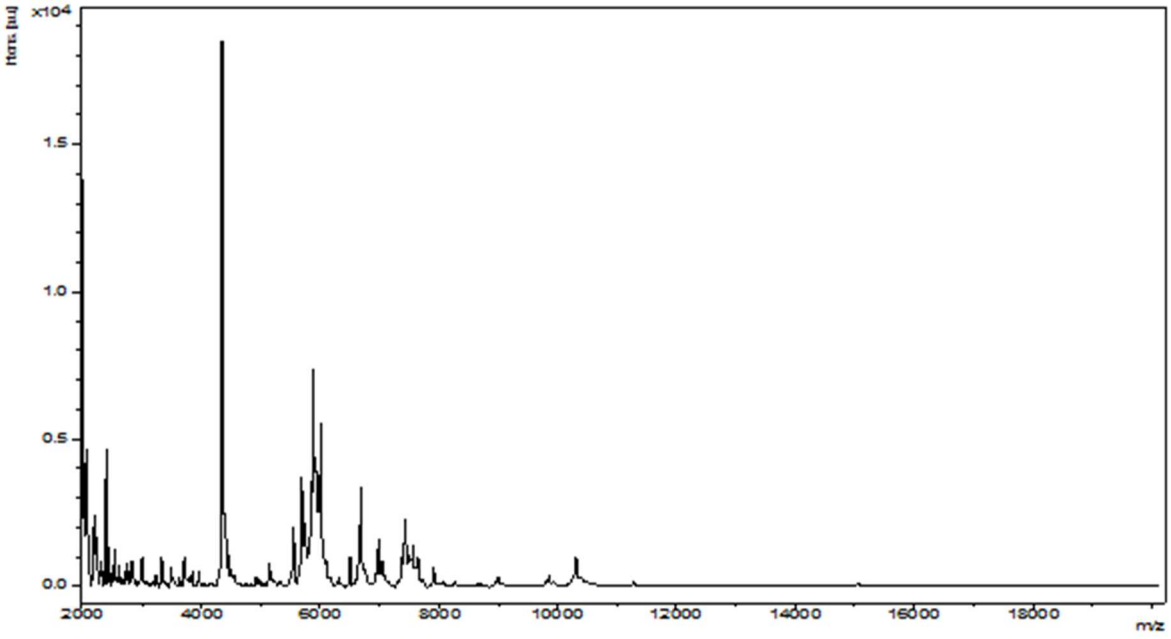


Figure 1. Reference mass spectrum from *Senegalia massiliensis* strain Marseille-P2130^{T}.

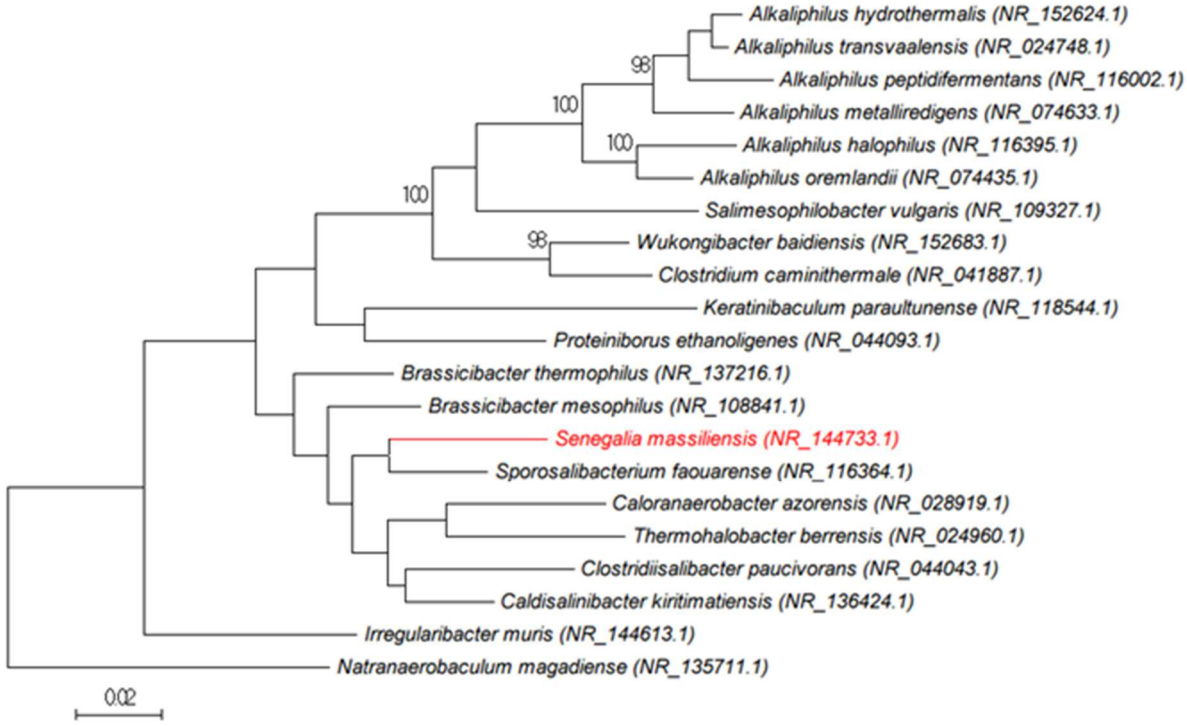
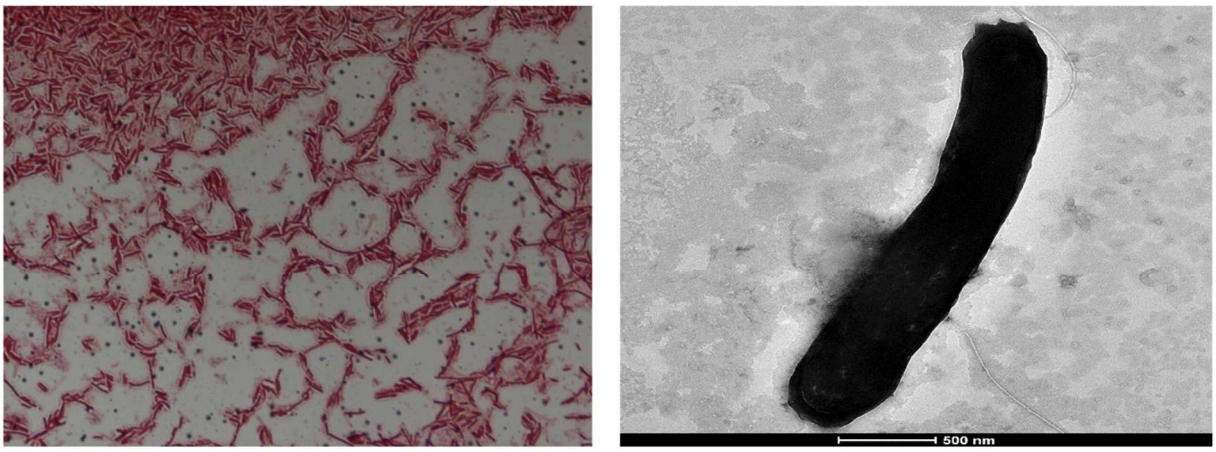


Figure 2. Phylogenetic tree highlighting the position of *Senegalia massiliensis* strain Marseille-P2130 relative to other close species. Sequences were aligned using CLUSTALW, with default parameters, and phylogenetic inferences were obtained using the maximum likelihood method within the MEGA7 software. Numbers at the nodes are percentages of bootstrap values obtained by repeating the analysis 500 times to generate a majority consensus tree. The scale bar indicates a 2% nucleotide sequence divergence.



A B

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

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

B. Transmission electron microscopy of *Senegalia massiliensis* strain Marseille-P2130 using Tecnai G20 electron microscope (FEI Company). The scale bar is equal to500 nm.

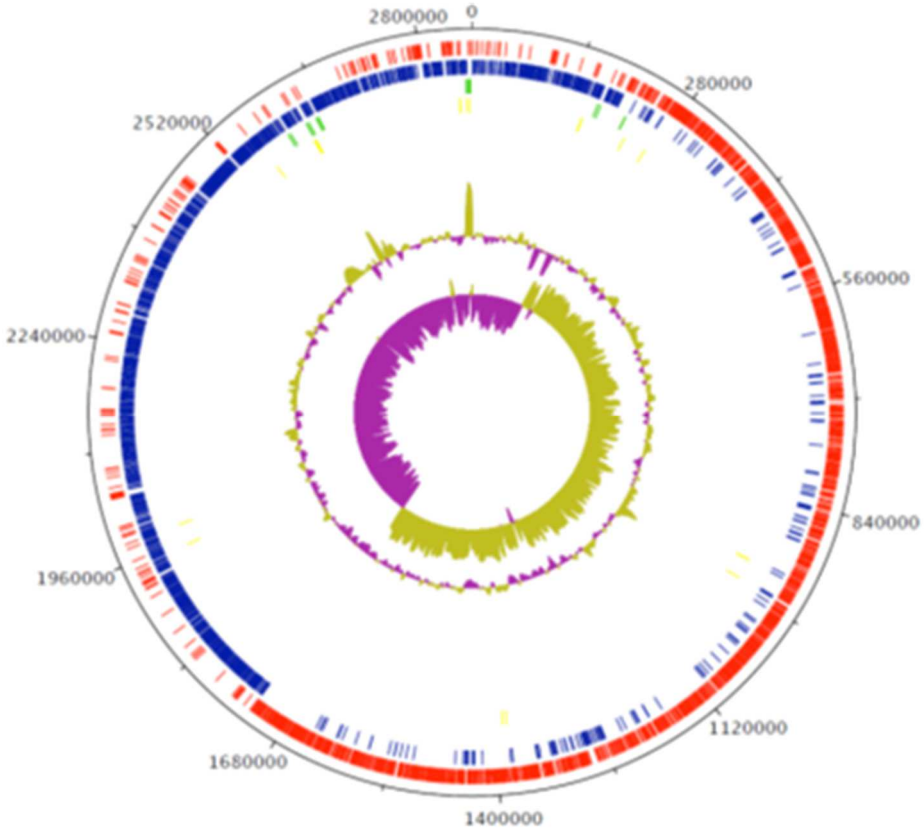


Figure 4. Graphical circular map of the chromosome. From outside to the center: Contigs (red), COG category of genes on the forward strand (three circles), genes on forward strand (blue circle), genes on the reverse strand (red circle), COG category on the reverse strand (three circles), G+C content.

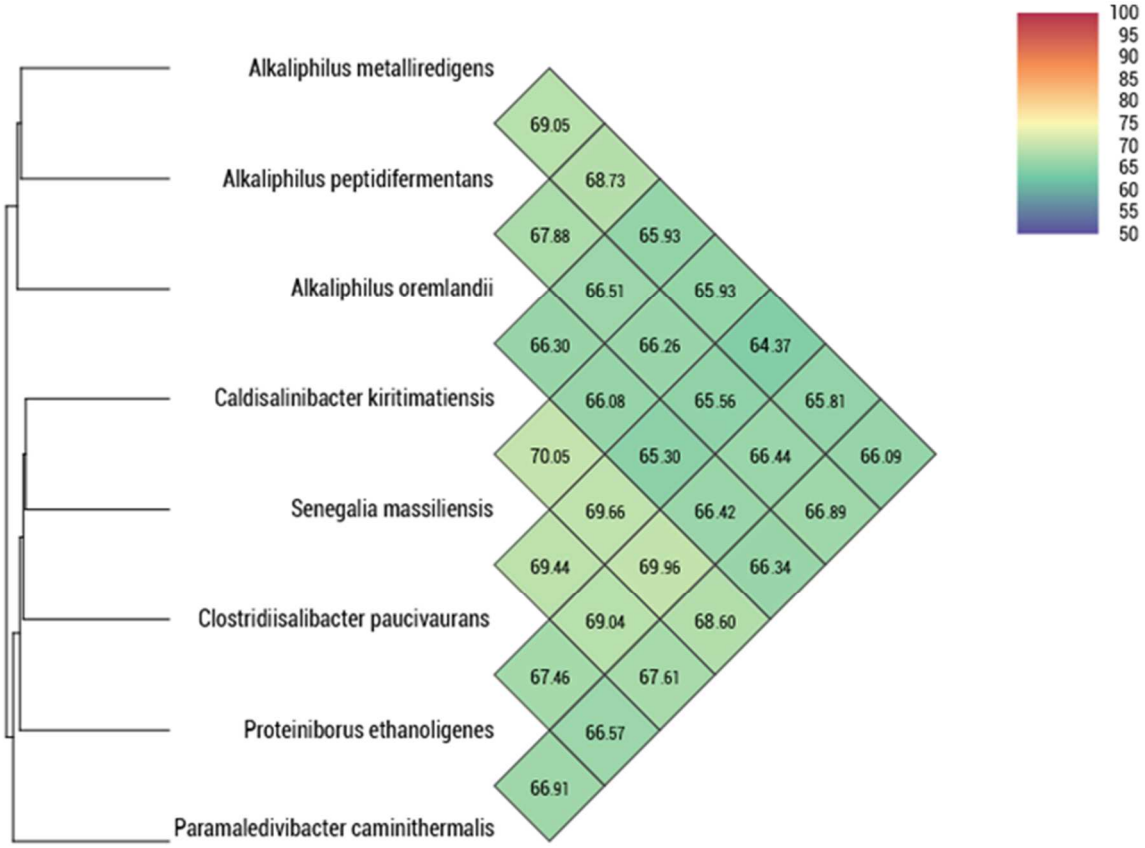


Figure 5. Heatmap generated with OrthoANI values calculated using the OAT software between *Senegalia massiliensis* and other closely related species with standing in nomenclature.

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

Property	**Terms**
Current Classification	Domain : *Bacteria* Phylum : *Firmicutes* Class : *Clostridia* Order : *Clostridiales* Family : *Clostridiaceae* Genus : *Senegalia* Species : *Senegalia massiliensis* 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

280

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

Properties	***Senegalia massiliensis***	***Clostridiisalibacter paucivorans***	***Alkaliphilus oremlandii***	***Alkaliphilus transvaalensis***	***Proteiniborus ethanoligenes***	***Sporosalibacterium faouarense***
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
Production of						
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
ß-galactosidase	+	NA	NA	NA	NA	NA
N-acetyl glucosamine	-	NA	NA	NA	NA	NA
Acid from						
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

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

Fatty acids	**Names**	**1**	**2**	**3**	**4**	**5**
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	-

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

Code	Value	Description
[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	Posttanslational 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

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

	Senegalia massiliensis	*Alkaliphilus metalliredigens*	*Alkaliphilus oremlandii*	*Alkaliphilus transvaalensis*	*Alkaliphilus peptidifermentans*	*Proteiniborus ethanoligenes*	*Clostridiisalibacter paucivorans*	*Paramaledivibacter caminithermalis*	*Caldisalinibacter kiritimatiensis*
Caldisalinibacter kiritimatiensis	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%
Paramaledivibacter caminithermalis	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%	
Clostridiisalibacter paucivorans	17.5±4.5	25±4.8	16.7±4.4	31±4.9	17.8±5.4	18.6±5.5	100%		
Proteiniborus ethanoligenes	17.3±4.1	19.9±4.6	19.8±4.6	18.7±4.6	16.2±4.3	100%			
Alkaliphilus peptidifermentans	17.2±4.4	25.5±5.2	24.8±4.8	19.9±4.7	100%				
Alkaliphilus transvaalensis	27.4±4.8	23.5±4.8	22.2±4.7	100%					
Alkaliphilus oremlandii	29.9±4.9	26.8±4.9	100%						
Alkaliphilus metalliredigens	33.9±4.9	100%							
Senegalia massiliensis	100%								

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