



**HAL**  
open science

# Genomic description and characterization of *Nigeribacterium massiliense* gen. nov., sp. nov., isolated from the human gut

Mudra Khare, Dhiraj Sinha, Rita Zgheib, Sory Ibrahima Traore,  
Jean-Christophe Lagier, Didier Raoult, Pierre-Edouard Fournier

## ► To cite this version:

Mudra Khare, Dhiraj Sinha, Rita Zgheib, Sory Ibrahima Traore, Jean-Christophe Lagier, et al.. Genomic description and characterization of *Nigeribacterium massiliense* gen. nov., sp. nov., isolated from the human gut. *Microbes and Infection*, 2021, 23 (8), pp.104842. 10.1016/j.micinf.2021.104842 . hal-03636904

HAL Id: hal-03636904

<https://amu.hal.science/hal-03636904>

Submitted on 5 Jan 2024

**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

1 **Genomic description and characterization of *Nigeribacterium massiliense***  
2 **gen. nov., sp. nov., isolated from the human gut**

3 Mudra Khare<sup>1,2</sup>, Dhiraj Sinha<sup>1,2</sup>, Rita Zgheib<sup>1,2</sup>, Sory Ibrahima Traore<sup>2,3</sup>, Jean-Christophe  
4 Lagier<sup>2,3</sup>, Didier Raoult<sup>2,3</sup>, Fournier Pierre Edouard<sup>2,3\*</sup>

5 <sup>1</sup> Aix Marseille University, Institut de Recherche pour le Développement (IRD), Service de  
6 Santé des Armées, Assistance Publique-Hôpitaux de Marseille, UMR Vecteurs Infections  
7 Tropicales et Méditerranéennes. (VITROME), Institut Hospitalo-Universitaire Méditerranée  
8 Infection, Marseille, France

9 <sup>2</sup> Institut Méditerranée-Infection, Marseille, France.

10 <sup>3</sup> Aix-Marseille Université, Institut de Recherche pour le Développement (IRD), UMR Microbes  
11 Evolution Phylogeny and Infections (MEPHI), Institut Hospitalo-Universitaire Méditerranée-  
12 Infection, Marseille, France.

13

14

15 \* **Corresponding author.** Pierre-Edouard FOURNIER, Institut Hospitalo-Universitaire  
16 Méditerranée-Infection, 19-21 Boulevard Jean Moulin, 13385, Marseille cedex 05, France; tel  
17 +33 413 732 401; fax: +33 413 732 402

18 e-mail: [pierre-edouard.fournier@univ-amu.fr](mailto:pierre-edouard.fournier@univ-amu.fr)

19

20

21 **Abstract**

22 Strain Marseille-P1302 was isolated from the stool of a 2-year-old Nigerian boy suffering from  
23 Kwashiorkor, a form of severe acute malnutrition. The strain grows in aerobic atmosphere and  
24 bacterial cells are Gram-positive cocci ranging in diameter from 0.8 to 1 µm. Strain Marseille-  
25 P1302 exhibits a 16S rRNA sequence similarity of 94.97% with *Brevilactibacter flavus* strain  
26 VG341<sup>T</sup>, but phylogenetically-closest species with standing in nomenclature is *Brevilactibacter*  
27 *sinopodophylli* strains KCTC 33808<sup>T</sup> with the sequence similarity of 93.41%. The draft genome  
28 of strain Marseille-P1302 is 2,934,258bp-long with a 70.38% G+C content, and contains 2,704  
29 protein-coding genes and 55 RNAs that includes 9 rRNA genes. On the basis of these data, we  
30 propose the creation of the new genus *Nigeribacterium* gen. nov., with strain Marseille-P1302<sup>T</sup>  
31 (= CSUR P1302 = DSM 29084) being the type strain of new species *Nigeribacterium*.  
32 *massiliense* gen. nov., sp. nov.

33 **Keywords** Culturomics, taxonomy, genomics, taxono-genomics, *Nigeribacterium massiliense*

34

## 35 **Introduction**

36 The human microbiome consists of a diversified array of microorganisms that include  
37 predominantly bacteria, viruses and fungi, but also archaea and protozoa. The main human  
38 habitats of microorganisms are the gastrointestinal system, oral cavity, genital organs,  
39 respiratory tract and skin [1]. The microbiota cell count in a human being is estimated to be  
40  $\sim 10^{13}$ - $10^{14}$ , with around 10:1 microbial cells to human cells [2]. Most of these microorganisms  
41 inhabit the colon [3], followed by the oral, skin and urogenital tracts [4]. The gastrointestinal  
42 microbiota is mainly composed of bacteria that belong to the phyla Firmicutes, Bacteroidetes  
43 and Actinobacteria [5]. Currently, for the study of the human gut microbiota, stool samples are  
44 the most representative [6]. In the past two decades, metagenomics has enabled exploring the  
45 composition of the gut microbiota very rapidly as well as extensively but this strategy has  
46 several limitations. The DNA extraction efficiency may vary among microorganisms; the  
47 results of these analyses are only snapshots of the real composition of the gut microbiota; this  
48 technique does not discriminate among living and dead microorganisms and does not allow  
49 appreciating their metabolic activity; and the interpretation of sequences allocated to unknown  
50 species is difficult and does not allow creating new species [7]. Because of these limitations  
51 and the need to obtain isolates from human-associated microorganisms, culturomics was  
52 developed. Culturomics is a strategy based on the diversification of culture conditions to mimic  
53 as much as possible each bacterial species' natural growth conditions [8]. To date, culturomics,  
54 combined to taxono-genomics [9], has enabled the characterization and creation of more than  
55 500 human-associated bacterial species [8]. Taxono-genomics was developed to incorporate  
56 into the taxonomic description of new bacterial taxa their peptidic profile using MALDI-TOF  
57 MS [10] and their genomic characteristics [6,11,12].

58 Using taxono-genomics, we herein report the description of strain Marseille-P1302 that was  
59 cultivated from the stool of a child suffering from Kwashiorkor. On the basis of its  
60 characteristics, we conclude that strain Marseille-P1302 belongs to a new species in a new  
61 genus for which we propose the name *Nigeribacterium massiliense* gen. nov., sp. nov.

## 62 **Materials and Methods**

### 63 **Sample collection**

64 The bacterial strain was isolated in 2014 from the stool of a 2-year-old child suffering from  
65 Kwashiorkor, a form of acute severe malnutrition in Niger. The patient was not taking any  
66 antibiotics at the time of sampling and the fecal sample was frozen at -80°C and was  
67 transported to Marseille, France, for diagnostic purposes. The child's parents provided an  
68 informed and signed consent. The study was validated by the ethics committee of the Institute  
69 Federatif de Recherche (Marseille, France) IFR48 under agreement number 09-022. All  
70 methods were performed in accordance with the relevant guidelines and regulations.

### 71 **Strain isolation and identification**

72 Culturomics was used to explore the microbial diversity of the child's gut microbiota [13]. One  
73 g of stool was pre-incubated at 37°C in anaerobic blood culture bottle (bioMérieux, Marcy  
74 l'Etoile, France) for 2 days. Then, the culture medium was spread on a 5% sheep blood-  
75 enriched Columbia agar plate (bioMérieux) and incubated at 37°C in aerobic condition.

76 Identification of the obtained colonies was attempted using MALDI-TOF-MS as previously  
77 described [10] using a Microflex spectrometer (Bruker Daltonics, Leipzig, Germany). The  
78 resultant spectrum was compared to the spectra contained in the Bruker database. If the score  
79 was  $\geq 2.0$ , the strain was considered as identified at the species level, but if the score was  $< 2$  the

80 strain was considered as unidentified. Strains considered as unidentified were further  
81 characterized by 16S rRNA gene amplification and sequencing using the Fd1 – rP2 universal  
82 primers [14]. If the 16S rRNA sequence similarity with the phylogenetically closest species  
83 with a validly published name was lower than 98.65% or 95%, the tested strain was considered  
84 as belonging to a new species and/or a new genus, respectively [15,16].

## 85 **Growth condition and Bacterial Morphology**

86 The optimal growth conditions of strain Marseille-P1302 were tested on 5% sheep blood-  
87 enriched Columbia agar (bioMérieux) at various temperatures (20, 25, 28, 37, 42 and 56°C),  
88 and also in various atmospheres (aerobic, anaerobic and microaerophilic). Gram-stain, motility,  
89 sporulation, catalase and oxidase tests were conducted according to the previously described  
90 protocols [17]. For transmission electron microscopy (TEM), a colony was collected from agar  
91 and immersed into a 2.5% glutaraldehyde fixative solution. Then, a drop of a suspension was  
92 directly deposited on a poly-L-lysine coated microscopic slide for 5 minutes and treated with  
93 1% phosphotungstic acid aqueous solution (pH 2.0) for 3 minutes to increase the TEM image  
94 contrast. The slide was gently washed in water, air dried and examined in a tabletop TEM  
95 (Tecnai G20). Acquisitions are shown in Figure 2.

## 96 **Isolation and sequencing of genomic DNA**

97 After the lysozyme incubation at 37°C for 30 minutes Genomic DNA (gDNA) of strain  
98 Marseille-P1302 was extracted using the EZ1 biorobot (Qiagen, Hilden, Germany) with the  
99 EZ1 DNA Tissue kit. The elution volume was 50µL. The genomic DNA was quantified by a  
100 Qubit assay with the high sensitivity kit (Life technologies, Carlsbad, CA, USA) at 105.7ng/µl.  
101 For genome sequencing, a mate-pair library was prepared with 1.5µg of gDNA using the  
102 Nextera mate-pair Illumina guide (Illumina, San Diego, CA, USA). The gDNA was

103 simultaneously fragmented and tagged with mate-pair junction adapters. The pattern of  
104 fragmentation was validated on an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa  
105 Clara, CA, USA) with a DNA 7500 labchip kit. The fragmented DNA ranged in size from 1.5  
106 kb up to 11 kb with an optimal size at 6.662 kb. No size selection was performed and 376.1 ng  
107 of tagmented fragments were circularized. The circularized DNA was mechanically sheared to  
108 small fragments with an optimal of 529 bp on the Covaris device S2 in T6 tubes (Covaris,  
109 Woburn, MA, USA). The library profile was visualized on a High Sensitivity Bioanalyzer  
110 LabChip kit (Agilent Technologies) and the final concentration library was measured at 4.162  
111 nmol/l. After a denaturation step and dilution at 15pM, the library was loaded onto the flow cell  
112 and then onto a MiSeq sequencer (Illumina). Automated cluster generation and sequencing run  
113 were performed in a single 39-hour run in a 2x251-bp format. Total information of 8.9 Gb was  
114 obtained from a 1023 K/mm<sup>2</sup> cluster density with a cluster passing quality control filters of  
115 89.6 %. The 1 178 689 reads were filtered according to the read qualities. The sequence  
116 assembly protocol was followed as previously described [18]

## 117 **Phenotypic and Biochemical features**

118 Biochemical properties of strain Marseille-P1302 were tested using API ZYM and API 50CH  
119 strips (bioMérieux) according to the manufacturer's instructions. Minimal inhibitory  
120 concentrations (MICs) of amoxicillin, vancomycin, ciprofloxacin, trimethoprim-  
121 sulfamethoxazole, rifampin erythromycin and colistin were determined according to EUCAST  
122 recommendations using E-test gradient strips (bioMerieux) [19,20].

123 Cellular fatty acid methyl ester (FAME) analysis was performed by Gas Chromatography/  
124 Mass Spectrometry (GC/MS). Two samples were prepared with approximately 80 mg of  
125 bacterial biomass per tube harvested from several culture plates. Fatty acid methyl esters were  
126 prepared as described by Sasser in 2006 [21]. GC/MS analyses were carried out as described

127 before [22]. Briefly, FAMEs were separated using an Elite 5-MS column and monitored by  
128 mass spectrometry (Clarus 500 - SQ 8 S, Perkin Elmer, Courtaboeuf, France). Spectral  
129 database search was performed using MS Search 2.0 operated with the Standard Reference  
130 Database 1A (NIST, Gaithersburg, USA) and the FAMEs mass spectral database (Wiley,  
131 Chichester, UK).

## 132 **Genome annotation and comparison**

133 The whole genome as well as 16S rRNA sequences (accession numbers CCYM000000000.1 and  
134 LK985392.1, respectively) of strain Marseille-P1302 were submitted to GenBank  
135 (www.ncbi.nlm.nih.gov). The 16S rRNA sequence from strain Marseille-P1302 was compared  
136 to those of other bacteria present in GenBank using BLASTn. Since for further analysis we  
137 have used whole genome sequences for the comparative analysis and these were not available  
138 for *Brevilactibacter sinopodophylli* strain TEYR-7<sup>T</sup> and *Propionicecella superfundia* strain BL-  
139 10<sup>T</sup>. 16S rRNA sequence extracted from *Brevilactibacter sinopodophylli* strain KCTC 33808<sup>T</sup>  
140 and *Propionicecella superfundia* strain DSM 22317<sup>T</sup>, respectively, were also used in the 16S  
141 phylogenetic analysis. Strain KCTC 33808 is also a type strain of *Brevilactibacter*  
142 *sinopodophylli*, while strain DSM 22317 is type strain of *Propionicecella superfundia*. A  
143 complete 16S rRNA-based phylogenetic tree was constructed by using the MEGA 7 software  
144 [23].

145 Genome annotation of strain Marseille-P1302 was obtained with the help of the PROKKA  
146 pipeline that includes several tools such as Aragorn, Barrnap and Prodigal to identify tRNAs,  
147 tmRNAs, rRNAs and coding sequences (CDS), respectively [24]. The predicted protein  
148 sequences were compared to the Clusters of Orthologous Groups (COGs) database using  
149 BLASTP (E value of 1e-03, coverage 0.7 and identity percent 30%). For the identification of  
150 antimicrobial resistance markers, plasmids and virulence genes, the contigs were screened



151 using CARD [24,25], Resfinder [26], ARG-ANNOT [27], PlasmidFinder [28] and VFDB [29],  
152 respectively, in the ABRicate pipeline (<https://github.com/tseemann/abricate>) [30] with a  
153 minimum identity of 60% and coverage of 20%. For the evaluation of sequence similarity at the  
154 genomic level between strain Marseille-P1302 and closely related species, we used the  
155 Genome-to-Genome Distance Calculator (GGDC) web server (<http://ggdc.dsmz.de>) to calculate  
156 digital DNA-DNA hybridization values values (dDDH) [31,32]. The recommended formula 2  
157 was considered for the interpretation of results. Moreover we also used the GGDC for the  
158 determination of the G+C content to compare the tested strains. The average nucleotide identity  
159 at the genome level was also estimated using the OAU tool that uses USEARCH algorithm for  
160 fast similarity search [33,34]. Morpheus software used for creating the heat map for ANI values  
161 (Morpheus, <https://software.broadinstitute.org/morpheus/>). The GView tool [35] was used for  
162 the visualisation of genomic features.

## 163 **Results and Discussion**

164 Strain identification and phylogenetic analysis

165 Strain Marseille-P1302 was isolated in 2014 and its growth was observed on 5% sheep blood-  
166 enriched Columbia agar at 37°C (bioMérieux) in aerobic atmosphere.

167 The MALDI-TOFF-MS spectrum of the novel isolate obtained a score <1.9, suggesting that the  
168 corresponding species was not represented in the reference database, and could belong to a new  
169 species. The spectrum was added to our database. The highest 16S rRNA sequence identity of  
170 strain Marseille-P1302 was 94.97% with *Brevilactibacter flavus* strain VG341<sup>T</sup>(MG799849.1)  
171 following to this 93.41% and 93.37% with *Brevilactibacter sinopodophylli* strains KCTC  
172 33808<sup>T</sup> and TEYR-7<sup>T</sup> (NR\_158077.1), 93% and 92.7% with *Propionicicella superfundia*  
173 strains DSM 22317<sup>T</sup> and BL-10<sup>T</sup> (NR-043609.1), 91.01% with *Microlunatus phosphovorus*

174 strain NM-1<sup>T</sup> (NR\_074642.2) and 93.03% with *Propioniclava tarda* strain WR061<sup>T</sup>  
175 (NR\_112669.1) (Figure 1).

176 As these values are lower than the 95% threshold for delineating a new genus as suggested by  
177 Stackebrandt and Ebers [15], strain Marseille-P1302 was considered as representative of a  
178 putatively new genus within the family *Propionibacteriaceae* in the phylum Actinobacteria.

179 Phenotypic and chemotaxonomic features

180 The growth of strain Marseille-P1302, was observed on 5% sheep blood–enriched Columbia  
181 agar at 37°C (bioMérieux). Optimal growth of strain Marseille P1302 was obtained in aerobic  
182 atmosphere, but weak growth was also observed in microaerophilic and anaerobic conditions.

183 The colonies of strain Marseille-P1302 were β-haemolytic. On blood-enriched agar, colonies  
184 were smooth and transparent, with a diameter of 0.2-0.4mm. Bacterial cells were non motile,  
185 non sporulating and coccoid with a diameter of 0.8-1.0µm (Figure 2). Bacterial Cells were  
186 Gram-positive (Figure 3) and exhibited catalase but no oxidase activity (Table1).

187 Using an API 50CH strip (bioMérieux), after 24 hours of incubation at 37°C, strain Marseille-  
188 P1302 was able to metabolize D-galactose, D-glucose, D-fructose, N-acetylglucosamine,  
189 esculine, D-maltose, D-lactose, D-saccharose, D-turanose and D-arabitol. In contrast, glycerol,  
190 erythritol, D-arabinose, L-arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, Methyl-βD-  
191 xylopyranoside, D-mannose, L-sorbose, L-rhamnose, Dulcitol, Inositol, D-mannitol, D-  
192 sorbitol, Methyl-αD-mannopyranoside, Methyl-αD-glucopyranoside, Amygdaline, Arbutine,  
193 Salicine, D-celiobiose, D-melibiose, D-trehalose, Inuline, D-melezitose, D-raffinose, Amidon,  
194 Glycogene, Xylitol, Gentiobiose, D-lyxose, D-tagatose, D-fucose, L-fucose, L-arabitol,  
195 Potassium gluconate, potassium 2-cetogluconate, potassium 5-cetogluconate were not  
196 metabolized (Table1).

197 An API ZYM strip (bioMérieux) showed the presence of leucine arylamidase, valine  
198 arylamidase, naphthol-AS-B1-phosphohydrolase,  $\beta$ -galactosidase,  $\alpha$ -glucosidase and  $\beta$ -  
199 glucosidase activities whereas negative reactions were obtained for alkaline phosphatase,  
200 esterase (C4), esterase lipase (C8), lipase (C14), cystine arylamidase, trypsin,  $\alpha$ -chymotrypsin,  
201 acid phosphatase,  $\alpha$ -galactosidase,  $\beta$ -glucuronidase, N-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -  
202 mannosidase and  $\alpha$ -fucosidase. By comparison with compared species, strain Marseille-P1302  
203 differed in a combination of alkaline phosphatase, esterase (C4), esterase lipase (C8) and N-  
204 acetyl- $\beta$ -glucosaminidase activities. The phenotypic comparison with other related strains is  
205 detailed in Table 1 [36-46].

206 Strain Marseille-P1302 was resistant to vancomycin, ciprofloxacin, erythromycin and  
207 trimethoprim-sulphamethoxazole but was susceptible to amoxicillin, rifampin and colistin.

208 FAME analysis revealed that the major cellular fatty acids found in strain Marseille-P1302 was  
209 12-methyl-tetradecanoic acid (89%), followed by 12-methyl-Tridecanoic acid (4%) and  
210 Hexadecanoic acid (2%). These fatty acids were principally saturated structures, half of them  
211 presenting branches. Minor amounts of unsaturated fatty acids were also detected (Table 2).

### 212 **Genomic characteristics and comparison**

213 The genome size of strain Marseille-P1302 is 2,934,258-bp long with a 70.38% G+C content.

214 A total of 2704 coding sequences, 9 rRNAs and 46 tRNAs were identified. The gene  
215 distribution in functional categories is presented in Table 3. Nine antibiotic resistance genes  
216 were detected, with identities ranging from 66.81% to 82.09%. According to CARD  
217 (comprehensive antibiotic resistance database) gene *Streptomyces\_cinnamoneus\_EF-*  
218 *Tu\_mutants\_conferring\_resistance\_to\_elfamycin* showing maximum coverage and identity  
219 percentage of 99% and 87.02% respectively, and it is responsible for the elfamycin resistance.

220 While the other database such as Resfinder and ARG-ANNOT showing *otr(A)\_1* and

221 (Tet)otrA respectively with very low coverage value of 15.06% and identity percentage of  
222 74.03% . Both the genes are responsible for the oxytetracycline and tetracycline resistance. The  
223 VFDB database showing the eight virulence genes among the only one gene *relA* has maximum  
224 coverage with identity of 68.03% and related to bacterial stringent response and involved in  
225 quorum sensing. Details of these resistance genes and virulence factors are presented in Table  
226 4. No plasmid was detected.

227 Comparison between the draft genomes of strain Marseille-P1302 and other phylogenetically  
228 closely related strains is presented in Table 5. The genome size (Mb) and number of protein-  
229 coding genes of strain Marseille-P1302 (2.93 and 2704 respectively) is smaller than those of all  
230 compared strains except *Cutibacterium acnes* strain ATCC 6919<sup>T</sup> (2.49 and 2308,  
231 respectively). Similarly, the lowest G+C content value was recorded for *Cutibacterium acnes*  
232 strain ATCC 6919<sup>T</sup> (60%) while strain Marseille-P1302 had a value of 70.38%, which is within  
233 the range of all other strains (67% for *Luteococcus japonicus* strain DSM 10546<sup>T</sup> to 77.32% for  
234 *Aeromicrobium massiliense* strain JC 14<sup>T</sup>). The distribution of genes into COG categories is  
235 shown in Figure 5. Strain Marseille-P1302 showing 2168 functional orthologous genes  
236 distributed in different categories of COG while the lowest number of orthologous genes  
237 (1785) found in strain *Cutibacterium acnes* strain ATCC 6919<sup>T</sup> and the highest number of  
238 orthologous genes (6313) were carried by *Amycolatopsis albispota* strain WP1<sup>T</sup>.

239 The dDDH values of strain Marseille-P1302 with other closely related strains ranged from  
240 17.9% to 22.8% (Table 6). The lowest value was recorded with *Microbacterium mangrove*  
241 strain MUSC115<sup>T</sup> and the highest value with *Cutibacterium acnes* strain ATCC6919<sup>T</sup> (Table  
242 6). OrthoANI values of strain Marseille-P1302 with other closely related strains varied from  
243 59.09% with *Microbacterium mangrovi* strain MUSC115<sup>T</sup> to 75.05% with *Brevilactibacter*

244 *sinopodophylli* strain KCTC 33808<sup>T</sup>. Values observed among other compared species were in  
245 the same range (Figure 6).

246 Based on these phenotypic and genomic characteristics, we consider strain Marseille-P1302 as  
247 belonging to a new genus and a new species, for which we propose the names *Nigeribacterium*  
248 gen. nov. and *Nigeribacterium massiliense* gen. nov., sp. nov., respectively. Strain Marseille-  
249 P1302<sup>T</sup> (= CSUR P1302) is the type strain of *Nigeribacterium massiliense* gen. nov., sp. nov.,  
250 the type species of the genus *Nigeribacterium* gen. nov.

#### 251 **Description of *Nigeribacterium* gen. nov.**

252 *Nigeribacterium* (ni.ge.ri.bac.te.ri'um. N.L. neutr. gen. n, combination of Niger, the country  
253 where the patient from whom the type strain was isolated, lived, and bacterium). Bacterial cells  
254 are aerobic, Gram-positive, non motile, non-spore-forming and cocci-shaped. Cells do not have  
255 any oxidase activity but are catalase-positive. *Nigeribacterium massiliense* gen. nov., sp. nov. is  
256 the type species of the genus *Nigeribacterium* gen. nov.

#### 257 **Description of *Nigeribacterium massiliense* gen. nov., sp. nov.**

258 *Nigeribacterium massiliense* (mas.il.i.en'se. L. Gen. neutr. n. *massiliense*, of Massilia, the Latin  
259 name of Marseille, the place where the strain was isolated). In addition to the description of the  
260 genus, cells have a diameter of 0.8µm to 1µm. Colonies grown on 5% sheep blood-enriched  
261 Columbia agar are small, smooth and transparent, varying in size from 0.2mm to 0.4mm in  
262 diameter(Figure 2). Growth is observed at 28 and 37°C, but optimal growth occurs at 37°C in  
263 aerobic atmosphere. Cells show weak growth in anaerobic and microaerophilic conditions.  
264 Using an API ZYM strip, leucine arylamidase, valine arylamidase, naphthol-AS-B1-  
265 phosphohydrolase, β-galactosidase, α-glucosidase and β-glucosidase activities are observed, but  
266 negative results are obtained for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase

267 (C14), cystine arylamidase, trypsin,  $\alpha$ -chymotrypsin, acid phosphatase,  $\alpha$ -galactosidase,  $\beta$ -  
268 glucuronidase, N-acetyl- $\beta$ -glucosaminidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase activities. Using  
269 an API 50CH strip, cells are able to metabolize D- galactose, D-glucose, D-fructose, N-  
270 acetylglucosamine, esculine, D-maltose, D-lactose, D-saccharose, D-turanose and D-arabitol  
271 (Table 1). The most abundant fatty acids are 12-methyl-tetradecanoic acid (89%) while some  
272 traces of 12-methyl-Tridecanoic acid (4%) and Hexadecanoic acid (2%) are also detected  
273 (Table 2). The genome is 2,934,258-bp long, its G+C content is 70.38% with genes distributed  
274 in different COG categories shown in Figure 4 and in Table 3.

275 The type strain, Marseille-P1302<sup>T</sup>, isolated from the fecal sample of a patient from Niger and  
276 suffering from Kwashiorkor, was deposited in the Collection de Souches de l'Unité des  
277 Rickettsies (CSUR, WDCM 875) and Deutsche Sammlung von Mikroorganismen und  
278 Zellkulturen (DSMZ) collections, under accession numbers CSUR P1302 and DSM 29084,  
279 respectively. The 16S rRNA and genome sequences are available in GenBank under accession  
280 numbers LK985392.1 and NZ\_CCYM01000003., respectively.

281

282 **Author contribution statement**

283 MK performed the genomic analysis and phenotypic tests, DS participated in the genomic  
284 analysis, RZ participated in the phenotypic analysis, SIT isolated the bacterium, DR, JCL and  
285 PEF designed the study and wrote the article. All authors reviewed the manuscript.

286 **Acknowledgments**

287 The study was supported by the Méditerranée Infection foundation, the National Research  
288 Agency under the program “Investissements d’avenir”, reference ANR-10-IAHU-03 and by  
289 Région Provence Alpes Côte d’Azur and European funding FEDER PRIMI.

290 **Additional information**

291 There are no competing interest.

292

293 **References**

- 294 1. Lloyd-Price J, Abu-Ali G, Huttenhower C (2016) The healthy human microbiome.  
295 Genome Medicine 8:51. <https://doi.org/10.1186/s13073-016-0307-y>
- 296 2. Sender R, Fuchs S, Milo R (2016) Are We Really Vastly Outnumbered? Revisiting the  
297 Ratio of Bacterial to Host Cells in Humans. Cell 164:337–340
- 298 3. Sender R, Fuchs S, Milo R (2016) Revised Estimates for the Number of Human and  
299 Bacteria Cells in the Body. PLoS Biology 14:.  
300 <https://doi.org/10.1371/journal.pbio.1002533>
- 301 4. Gupta VK, Paul S, Dutta C (2017) Geography, ethnicity or subsistence-specific  
302 variations in human microbiome composition and diversity. Frontiers in Microbiology 8
- 303 5. Tap J, Mondot S, Levenez F, et al (2009) Towards the human intestinal microbiota  
304 phylogenetic core. Environmental Microbiology 11:2574–2584.  
305 <https://doi.org/10.1111/j.1462-2920.2009.01982.x>
- 306 6. Bilen M, Mbogning Fonkou MD, Khelaifia S, et al (2019) Taxonogenomics description  
307 of Parabacteroides timonensis sp. nov. isolated from a human stool sample.  
308 MicrobiologyOpen 8: <https://doi.org/10.1002/mbo3.702>
- 309 7. Wang WL, Xu SY, Ren ZG, et al (2015) Application of metagenomics in the human gut  
310 microbiome. World Journal of Gastroenterology 21:803–814
- 311 8. Lagier JC, Dubourg G, Million M, et al (2018) Culturing the human microbiota and  
312 culturomics. Nature Reviews Microbiology 16:540–550
- 313 9. Ramasamy D, Mishra AK, Lagier JC, et al (2014) A polyphasic strategy incorporating  
314 genomic data for the taxonomic description of novel bacterial species. International  
315 Journal of Systematic and Evolutionary Microbiology 64:384–391
- 316 10. Seng P, Drancourt M, Gourié F, et al (2009) Ongoing revolution in bacteriology:



- 317 Routine identification of bacteria by matrix-assisted laser desorption ionization time-of-  
318 flight mass spectrometry. *Clinical Infectious Diseases* 49:543–551.  
319 <https://doi.org/10.1086/600885>
- 320 11. Fournier PE, Lagier JC, Dubourg G, Raoult D (2015) From culturomics to  
321 taxonomogenomics: A need to change the taxonomy of prokaryotes in clinical  
322 microbiology. *Anaerobe* 36:73–78. <https://doi.org/10.1016/j.anaerobe.2015.10.011>
- 323 12. Lagier JC, Hugon P, Khelaifia S, et al (2015) The rebirth of culture in microbiology  
324 through the example of culturomics to study human gut microbiota. *Clinical*  
325 *Microbiology Reviews* 28:237–264
- 326 13. Lagier J-CC, Armougom F, Million M, et al (2012) Microbial culturomics: Paradigm  
327 shift in the human gut microbiome study. *Clinical Microbiology and Infection* 18:1185–  
328 1193. <https://doi.org/10.1111/1469-0691.12023>
- 329 14. Drancourt M, Bollet C, Carlioz A, et al (2000) 16S ribosomal DNA sequence analysis of  
330 a large collection of environmental and clinical unidentifiable bacterial isolates. *Journal*  
331 *of Clinical Microbiology* 38:3623–3630. [https://doi.org/10.1128/jcm.38.10.3623-](https://doi.org/10.1128/jcm.38.10.3623-3630.2000)  
332 [3630.2000](https://doi.org/10.1128/jcm.38.10.3623-3630.2000)
- 333 15. Erko S, Ebers J (2006) Taxonomic parameters revisited: tarnished gold standards.  
334 *Microbiology Today* 33:152–155
- 335 16. Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between  
336 average nucleotide identity and 16S rRNA gene sequence similarity for species  
337 demarcation of prokaryotes. *International Journal of Systematic and Evolutionary*  
338 *Microbiology* 64:346–351. <https://doi.org/10.1099/ijs.0.059774-0>
- 339 17. Wormser GP, Stratton C (2008) *Manual of Clinical Microbiology*, 9th Edition Edited by  
340 Patrick R. Murray, Ellen Jo Baron, James H. Jorgensen, Marie Louise Landry, and  
341 Michael A. Pfaller Washington, DC: ASM Press, 2007 2488 pp., illustrated. \$209.95

- 342 (hardcover). *Clinical Infectious Diseases* 46:153–153. <https://doi.org/10.1086/524076>
- 343 18. Anani H, Abdallah RA, Khoder M, et al (2019) *Colibacter massiliensis* gen. nov. sp.  
344 nov., a novel Gram-stain-positive anaerobic diplococcal bacterium, isolated from the  
345 human left colon. *Scientific Reports* 9:. <https://doi.org/10.1038/s41598-019-53791-1>
- 346 19. Alcock BP, Raphenya AR, Lau TTY, et al (2020) CARD 2020: Antibiotic resistance  
347 surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids*  
348 *Research* 48:D517–D525. <https://doi.org/10.1093/nar/gkz935>
- 349 20. Macià MD, Rojo-Molinero E, Oliver A (2014) Antimicrobial susceptibility testing in  
350 biofilm-growing bacteria. *Clinical Microbiology and Infection* 20:981–990
- 351 21. Sasser M (2001) Identification of Bacteria by Gas Chromatography of Cellular Fatty  
352 Acids. *Technical Note* 101:1–6
- 353 22. Dione N, Sankar SA, Lagier JC, et al (2016) Genome sequence and description of  
354 *Anaerosalibacter massiliensis* sp. nov. *New Microbes and New Infections* 10:66–76.  
355 <https://doi.org/10.1016/j.nmni.2016.01.002>
- 356 23. Kumar S, Stecher G, Li M, et al (2018) MEGA X: Molecular evolutionary genetics  
357 analysis across computing platforms. *Molecular Biology and Evolution* 35:1547–1549.  
358 <https://doi.org/10.1093/molbev/msy096>
- 359 24. Seemann T (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*  
360 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
- 361 25. Jia B, Raphenya AR, Alcock B, et al (2017) CARD 2017: Expansion and model-centric  
362 curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*  
363 45:D566–D573. <https://doi.org/10.1093/nar/gkw1004>
- 364 26. Prasertsee T, Chuammitri P, Deudom M, et al (2019) Core genome sequence analysis to  
365 characterize *Salmonella enterica* serovar Rissen ST469 from a swine production chain.  
366 *International Journal of Food Microbiology* 304:68–74.

- 367 <https://doi.org/10.1016/j.ijfoodmicro.2019.05.022>
- 368 27. Brambila-Tapia AJL, Perez-Rueda E, Barrios H, et al (2017) Identification of beta-  
369 Lactamases and beta-Lactam-Related Proteins in Human Pathogenic Bacteria using a  
370 Computational Search Approach. *Current Microbiology* 74:915–920.  
371 <https://doi.org/10.1007/s00284-017-1265-3>
- 372 28. Carattoli A, Zankari E, García-Fernández A, et al (2014) In Silico detection and typing  
373 of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrobial  
374 Agents and Chemotherapy* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>
- 375 29. Chen L, Yang J, Yu J, et al (2005) VFDB: A reference database for bacterial virulence  
376 factors. *Nucleic Acids Research* 33:D325-8. <https://doi.org/10.1093/nar/gki008>
- 377 30. Arnott A, Wang Q, Bachmann N, et al (2018) Multidrug-resistant *Salmonella enterica*  
378 4,[5],12: I:- sequence type 34, New South Wales, Australia, 2016-2017. *Emerging  
379 Infectious Diseases* 24:751–753. <https://doi.org/10.3201/eid2404.171619>
- 380 31. Meier-Kolthoff JP, Auch AF, Klenk H-PP, Göker M (2013) Genome sequence-based  
381 species delimitation with confidence intervals and improved distance functions. *BMC  
382 Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>
- 383 32. Meier-Kolthoff JP, Göker M, Spröer C, Klenk HP (2013) When should a DDH  
384 experiment be mandatory in microbial taxonomy? *Archives of Microbiology* 195:413–  
385 418. <https://doi.org/10.1007/s00203-013-0888-4>
- 386 33. Yoon SH, Ha S min, Lim J, et al (2017) A large-scale evaluation of algorithms to  
387 calculate average nucleotide identity. *Antonie van Leeuwenhoek, International Journal  
388 of General and Molecular Microbiology* 110:1281–1286.  
389 <https://doi.org/10.1007/s10482-017-0844-4>
- 390 34. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST.  
391 *Bioinformatics* 26:2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>

- 392 35. Petkau A, Stuart-Edwards M, Stothard P, van Domselaar G (2010) Interactive microbial  
393 genome visualization with GView. *Bioinformatics* 26:3125–3126.  
394 <https://doi.org/10.1093/bioinformatics/btq588>
- 395 36. Zhang G, Wan L, Li J, Zhou Y (2016) *Amycolatopsis albispora* sp. Nov., isolated from  
396 deep-sea sediment. *International Journal of Systematic and Evolutionary Microbiology*  
397 66:3860–3864. <https://doi.org/10.1099/ijsem.0.001277>
- 398 37. Lee LH, Azman AS, Zainal N, et al (2014) *Microbacterium mangrovi* sp. nov., an  
399 amyolytic actinobacterium isolated from mangrove forest soil. *International journal of*  
400 *systematic and evolutionary microbiology* 64:3513–3519.  
401 <https://doi.org/10.1099/ijms.0.062414-0>
- 402 38. Dekio I, McDowell A, Sakamoto M, et al (2019) Proposal of new combination,  
403 *cutibacterium acnes* subsp. *Elongatum* comb. nov., and emended descriptions of the  
404 genus *cutibacterium*, *cutibacterium acnes* subsp. *acnes* and *cutibacterium acnes* subsp.  
405 *defendens*. *International Journal of Systematic and Evolutionary Microbiology* 69:1087–  
406 1092. <https://doi.org/10.1099/ijsem.0.003274>
- 407 39. Ramasamy D, Kokcha S, Lagier JC, et al (2012) Genome sequence and description of  
408 *Aeromicrobium massiliense* sp. nov. *Standards in Genomic Sciences* 7:246–257.  
409 <https://doi.org/10.4056/sigs.3306717>
- 410 40. Tamura T, Takeuchi M, Yokota A (1994) *Luteococcus japonicus* gen. nov., sp. nov., a  
411 new gram-positive coccus with LL-diaminopimelic acid in the cell wall. *International*  
412 *Journal of Systematic Bacteriology* 44:348–356. [https://doi.org/10.1099/00207713-44-2-](https://doi.org/10.1099/00207713-44-2-348)  
413 348
- 414 41. Udvary DW, Zeigler L, Asolkar RN, et al (2007) Genome sequencing reveals complex  
415 secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proceedings of*  
416 *the National Academy of Sciences of the United States of America* 104:10376–10381.

417 <https://doi.org/10.1073/pnas.0700962104>

418 42. Lee DW, Lee SD (2008) *Tessaracoccus flavescens* sp. nov., isolated from marine  
419 sediment. *International Journal of Systematic and Evolutionary Microbiology* 58:785–  
420 789. <https://doi.org/10.1099/ijs.0.64868-0>

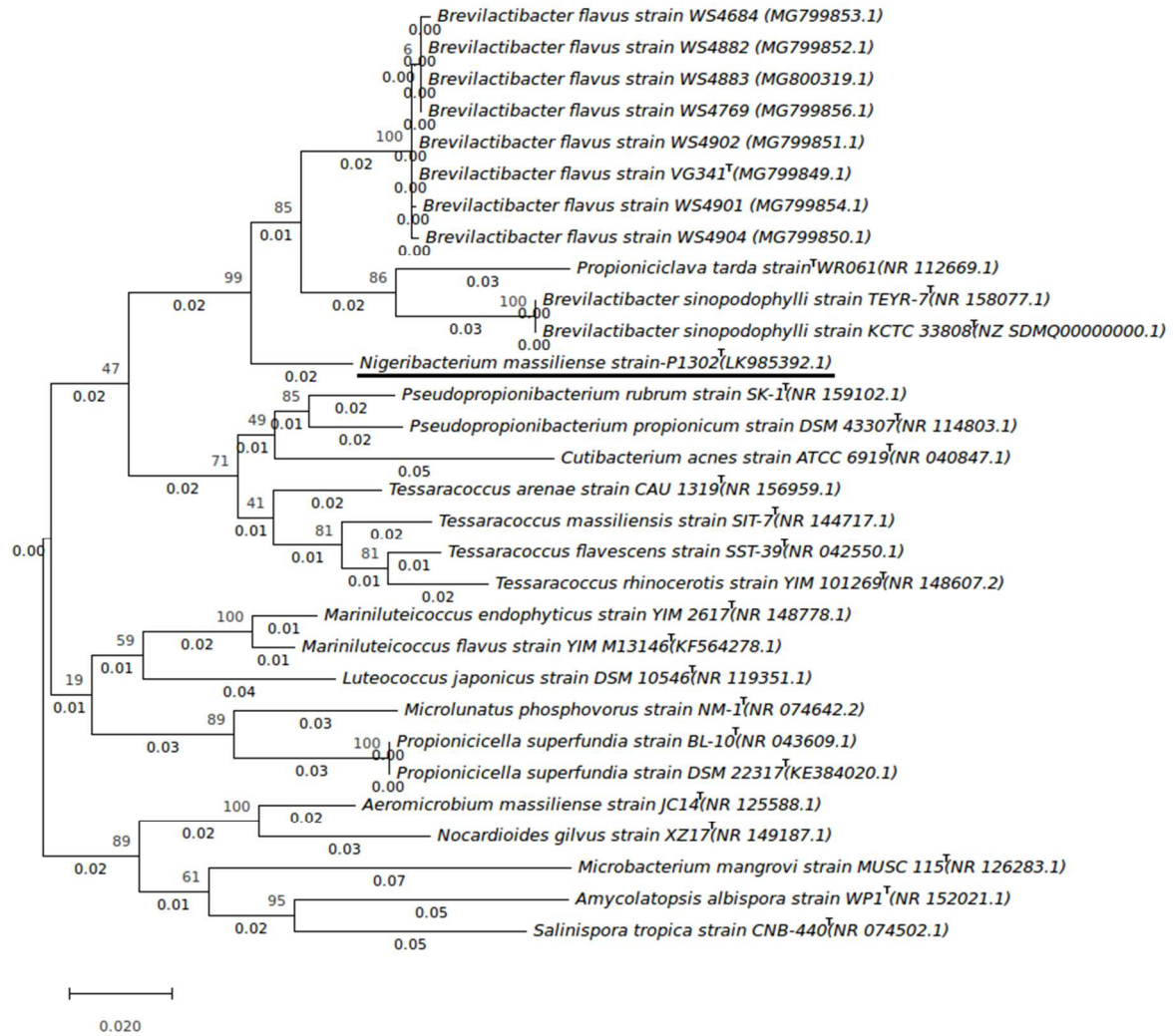
421 43. Zhang HX, Wang K, Xu ZX, et al (2016) *Nocardioides gilvus* sp. nov., isolated from  
422 Namtso Lake. *Antonie van Leeuwenhoek, International Journal of General and*  
423 *Molecular Microbiology* 109:1367–1374. <https://doi.org/10.1007/s10482-016-0735-0>

424 44. Nakamura K, Hiraishi A, Yoshimi Y, et al (1995) *Microlunatus phosphovorus* gen. nov.,  
425 sp. nov., a new gram-positive polyphosphate-accumulating bacterium isolated from  
426 activated sludge. *International Journal of Systematic Bacteriology* 45:17–22.  
427 <https://doi.org/10.1099/00207713-45-1-17>

428 45. Wenning M, Breitenwieser F, Huptas C, et al (2020) *Brevilactibacter flavus* gen. Nov.,  
429 sp. nov., a novel bacterium of the family propionibacteriaceae isolated from raw milk  
430 and dairy products and reclassification of *propioniciclava sinopodophylli* as  
431 *brevilactibacter sinopodophylli* comb. nov. *International Journal of Systematic and*  
432 *Evolutionary Microbiology* 70:2186–2193. <https://doi.org/10.1099/ijsem.0.003909>

433 46. Bae HS, Moe WM, Yan J, et al (2006) *Propionicicella superfundia* gen. nov., sp. nov., a  
434 chlorosolvent-tolerant propionate-forming, facultative anaerobic bacterium isolated from  
435 contaminated groundwater. *Systematic and Applied Microbiology* 29:404–413.  
436 <https://doi.org/10.1016/j.syapm.2005.11.004>

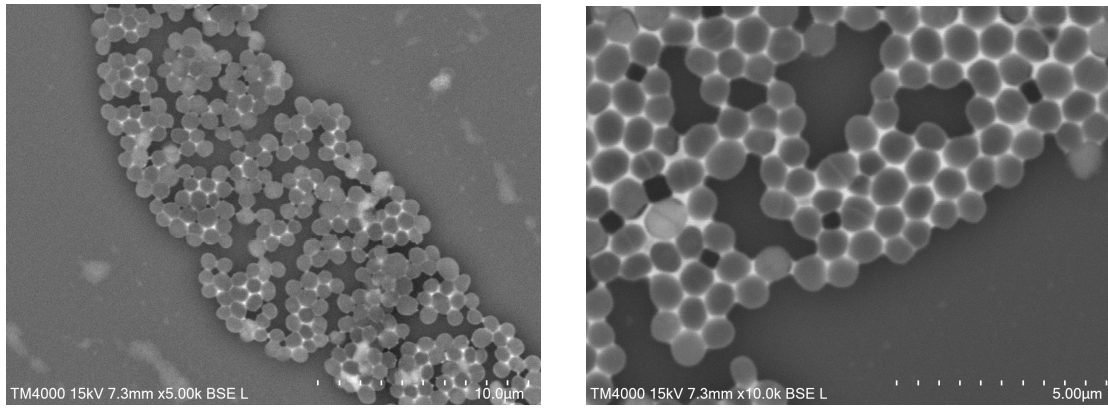
437



438 **Figure 1.** 16S rRNA-based phylogenetic tree highlighting the position of strain Marseille-  
 439 P1302 relative to other closely related strains within the class Actinobacteria. Sequences were  
 440 aligned by ClustalW with default parameters and Phylogenetic inferences were obtained using  
 441 the maximum-likelihood method and the MEGA version 7 software. Numbers at the nodes are  
 442 bootstrap values generated from 1000 replicates. The scale bar represents 2% of nucleotide  
 443 divergence. Superscript Ts represent type strains.

445

446



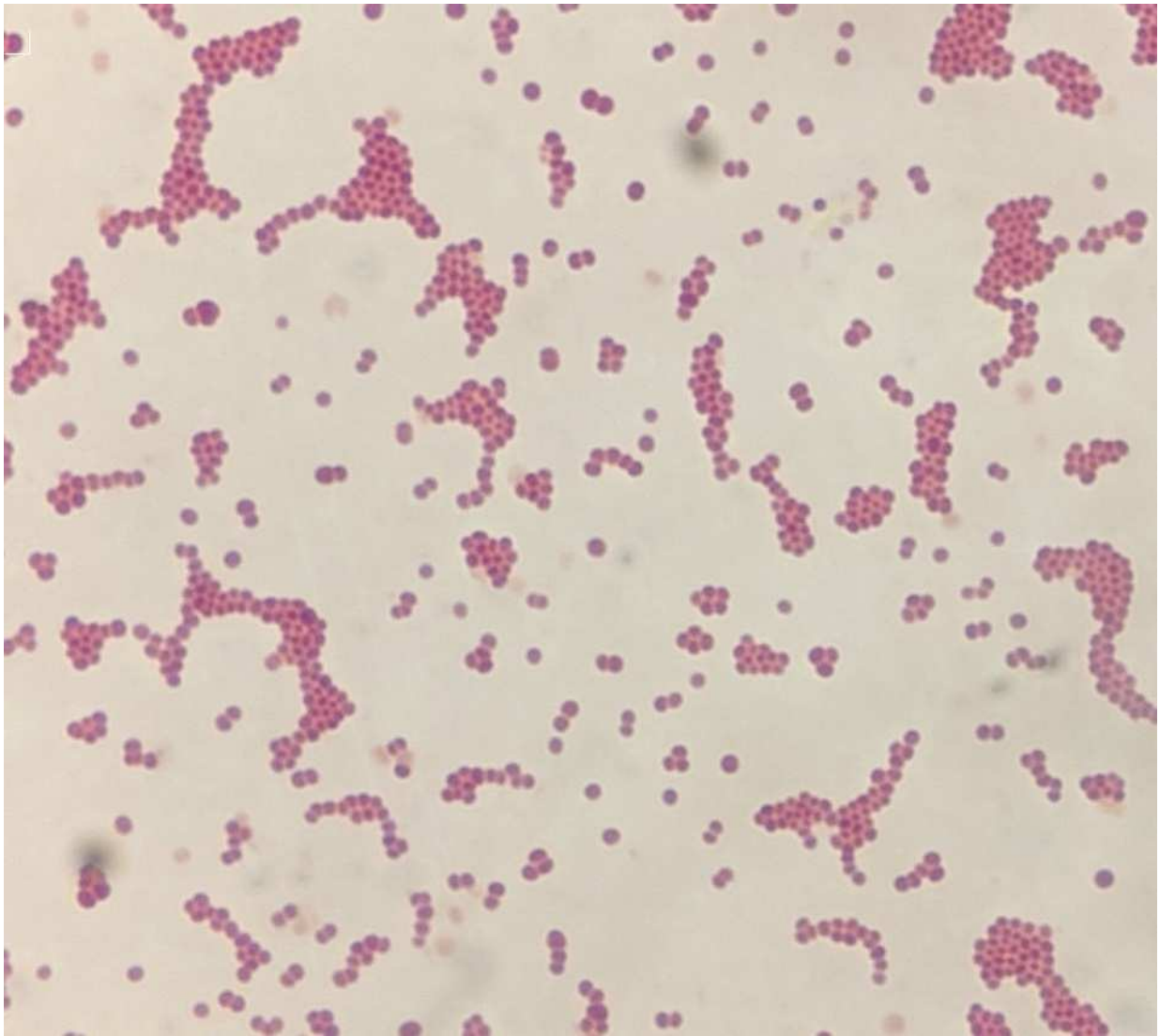
447

448

449 **Figure 2.** Transmission electron microscopy of strain Marseille-P1302<sup>T</sup> using a Tecnai G20  
450 transmission electron microscope. Left, the scale bar represents 10μm whereas it represents  
451 5μm in the right picture.

452

453



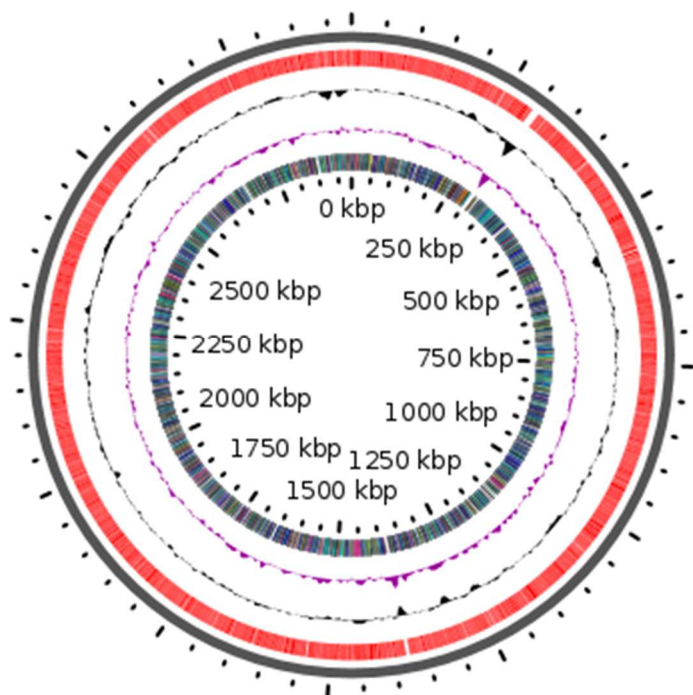
454

455 **Figure 3.** Gram staining image of strain Marseille-P1302<sup>T</sup>

456



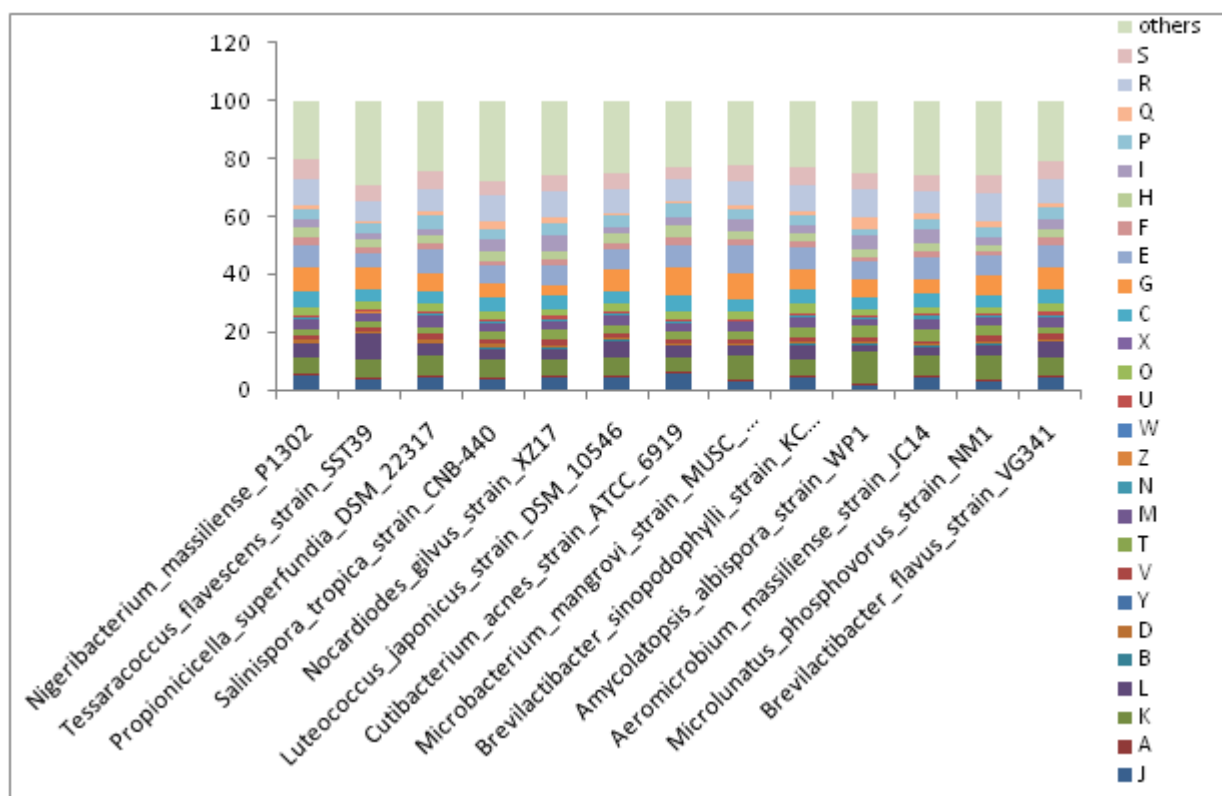
457



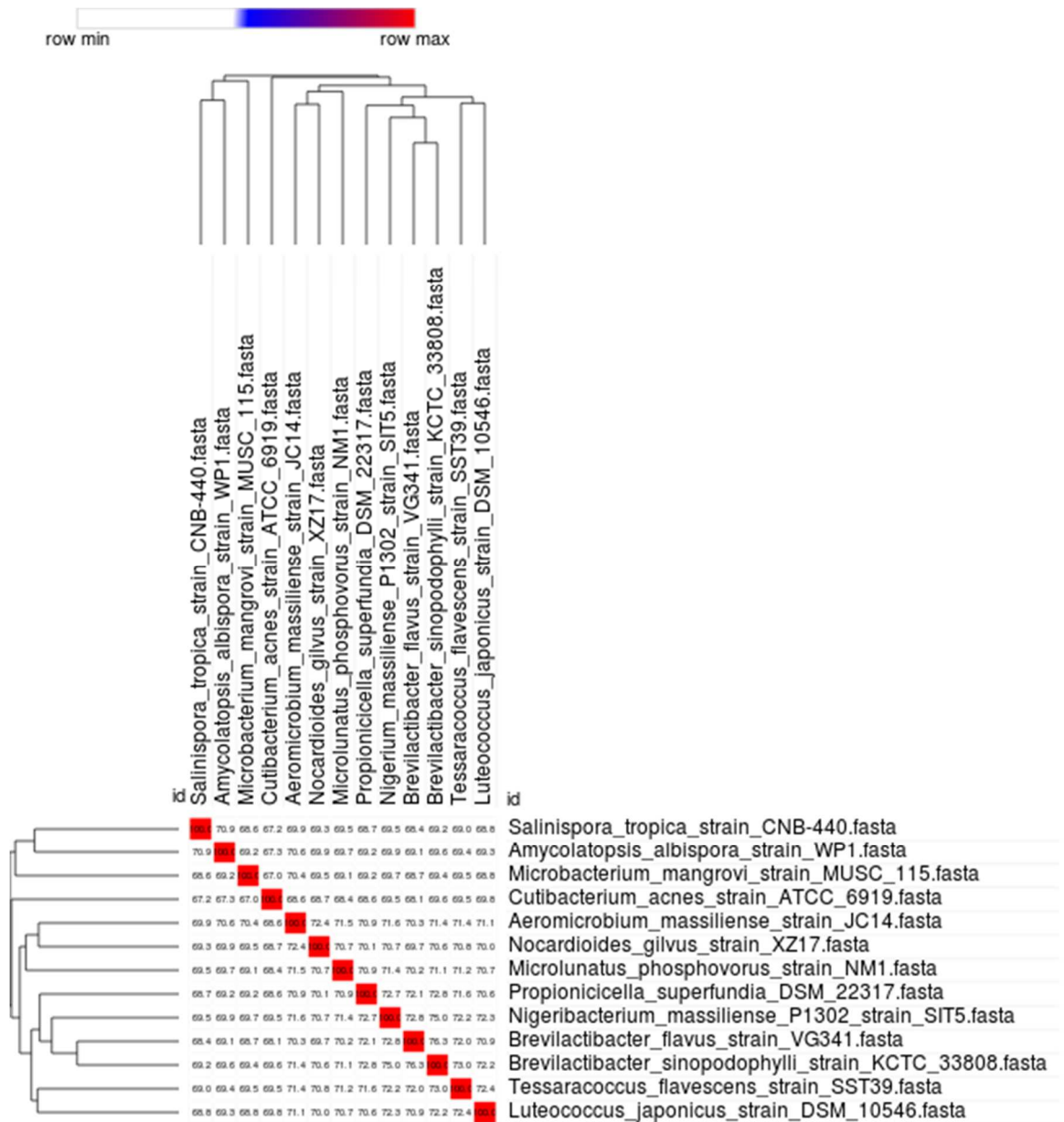
458

459 **Figure 4.** Graphical circular map of the chromosome from strain Marseille-P1302<sup>T</sup>. From the inside out  
460 the circle represents the COG categories, GC skew, G+C content and CDS.

461



462  
 463 **Figure 5.** Distribution of genes from strain Marseille-P1302<sup>T</sup> and closely related species into COGs  
 464 functional categories. Category abbreviations are as follows: C, energy production and conversion; E,  
 465 amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport  
 466 and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; P, inorganic  
 467 ion transport and metabolism; Q, secondary metabolites biosynthesis, transport and catabolism; X,  
 468 mobilome: prophages, transposons; A, RNA processing and modification; B, chromatin structure and  
 469 dynamics; J, translation, ribosomal structure and biogenesis; K, transcription; L, replication,  
 470 recombination and repair; D, cell cycle control, cell division, chromosomal partitioning; M, cell  
 471 wall/membrane/envelope biogenesis; N, cell motility; O, posttranslational modification, protein  
 472 turnover, chaperones; T, signal transduction mechanisms; U, intracellular trafficking, secretion, and  
 473 vesicular transport; V, defense mechanisms; W, extracellular structures; Z, cytoskeleton; R, general  
 474 function predicted only; S, function unknown.



476

477 **Figure 6.** OrthoANI values obtained when comparing strain Marseille-P1302<sup>T</sup> to phylogenetically  
 478 closely related strains.

479

480

481 **Table 1.** Differential characteristics of strain Marseille P1302 and related species

Characteristics	1	2	3	4	5	6	7	8	9	10	11	12	13
Gram stain	+	+	+	+	+	+	+	+	+	+	+	+	+
Production of													
Catalase	+	+	+	+	+	+	na	+	-	+	+	-	+
Oxidase	-	-	-	-	-	+	na	-	+	+	+	-	+
Utilization of													
L-Arabinose	-	na	na	na	-	+	na	na	na	na	-	na	na
D-Galactose	+	na	-	na	na	+	+	+	na	na	-	na	+
D-Glucose	+	+	na	+	+	+	na	na	na	na	-	na	+
D-Fructose	+	na	+	na	na	+	na	na	na	na	+	na	+
N-acetylglucosamine	+	na	+	na	-	na	na	na	na	na	-	na	+
Esculine	+	na	na	-	+	+	na	na	+	na	+	na	+
D-Maltose	+	na	na	-	na	+	na	na	na	na	-	na	+
D-Arabitol	+	na	+	na	na	na	na	na	-	na	-	na	+
Enzymatic activities													
Alkaline Phosphatase	-	+	na	-	-	+	+	+	-	+	-	-	na
Esterase (C4)	-	+	na	na	+	-	+	+	+(w)	+	+	na	na
Esterase Lipase (C8)	-	+	na	na	+	+	+	+	+	+	+	na	na
Lipase (C14)	-	-	na	+	-	-	-	-	-	-	-	na	na
Leucine arylamidase	+	+	na	-	+	+	+	+	+	+	+	na	na
Valine arylamidase	+	+	na	na	w	-	+	+	-	+	+/-	na	na

482

483

484

485

Cystine arylamidase	-	+	na	na	-	-	-	-	-	+	-	na	na
Trypsin	-	-	na	na	na	+	-	-	-	+	-	na	na
$\alpha$ -chymotrypsin	-	-	na	na	na	+	-	-	-	+	+/-	na	na
Acid phosphatase	-	+	na	+	-	+	-	+	-	+	+	na	na
Naphthol-AS-BI-	+	+	na	na	-	-	+	-	+(w)	+	+	na	
phosphohydrolase													
$\alpha$ -galactosidase	-	-	na	-	na	+	-	+	-	+	+/-	-	na
$\beta$ -galactosidase	+	+	na	+	na	+	-	+	-	+	+	-	na
$\beta$ -glucuronidase	-	-	na	-	w	-	-	+	-	-	-	-	na
$\alpha$ -glucosidase	+	+	na	-	+	+	+	+	-	+	+	+/-	na
$\beta$ -glucosidase	+	+	na	-	+	+	+	+	-	+	+	+	na
N-acetyl- $\beta$ -	-	+	na	+	na	+	+	+	-	+	-	-	na
glucosaminidase													
$\alpha$ -mannosidase	-	-	na	na	na	-	-	-	-	+	-	na	na
$\alpha$ -fucosidase	-	-	na	-	na	-	-	-	-	+	-	-	na
Motility	Non motile	Non motile	na	Non motile	motile	Non motile	Non motile	Non motile	Non motile	Non motile	Non motile	Non motile	Non motile
Spore formation													
Source	Human gut	Deep sea sediment	Mangrove soil	Lesion from human facial skin	Human gut	Soil	Marine sediment	Beach sediment	Namtso lake	Activated sludge	Sinopodophyllum laves	Contaminated Ground water	Raw milk

486 Taxa: 1, strain Marseille-P1302<sup>T</sup>; 2, *Amycolatopsis albispora* strain WP1<sup>T</sup>; 3, *Microbacterium mangrovi* strain MUSC 115<sup>T</sup>; 4, *Cutibacterium acnes* strain ATCC 6919<sup>T</sup>; 5,

487 *Aeromicrobium massiliense* strain JC14<sup>T</sup>; 6, *Luteococcus japonicus* strain IFO<sup>T</sup>; 7, *Salinispora tropica* strain CNB 440<sup>T</sup>; 8, *Tessaracoccus flavescens* strain SST 39<sup>T</sup>; 9,

488 *Nocardiodes gilvus* strain XZ17<sup>T</sup>; 10, *Microlunatus phosphovorus* strain NM-1<sup>T</sup>, 11, *Brevilactibacter sinopodophylli* strain KCTC 33808<sup>T</sup>, 12, *Propionicicella superfundia*  
489 strain DSM 22317<sup>T</sup>, 13, *Brevilactibacter flavus* strain VG341<sup>T</sup>, Abbreviations: +, positive; -, negative; na, not available; w, weak.

490  
491

**Table 2.** Cellular fatty acid composition of strain Marseille-P1302<sup>T</sup>

<b>Fatty acids</b>	<b>Name</b>	<b>Mean relative %<sup>a</sup></b>	492
15:0 anteiso	12-methyl-tetradecanoic acid	88.9 ± 0.5	493
14:0 iso	12-methyl-Tridecanoic acid	3.7 ± 0.1	494
16:0	Hexadecanoic acid	2.2 ± 0.2	495
14:0	Tetradecanoic acid	1.1 ± 0.1	496
16:0 iso	14-methyl-Pentadecanoic acid	TR	497
15:0	Pentadecanoic acid	TR	498
13:0 anteiso	10-methyl-Dodecanoic acid	TR	499
5:0 anteiso	2-methyl-Butanoic acid	TR	500
15:0 iso	13-methyl-tetradecanoic acid	TR	501
18:2n6	9,12-Octadecadienoic acid	TR	502
18:1n9	9-Octadecenoic acid	TR	503
17:0	Heptadecanoic acid	TR	504
17:0 anteiso	14-methyl-Hexadecanoic acid	TR	505
18:0	Octadecanoic acid	TR	506
			507
			508
			509
			510
			511
			512
			513
			514
			515
			516
			517
			518
			519

520 <sup>a</sup> Mean peak area percentage; TR, trace amount

521

522

523 **Table 3.** Number of genes in strain Marseille-P1302<sup>T</sup> associated with the 25 general COG functional  
 524 categories.  
 525

<b>Code</b>	<b>Value</b>	<b>% Value<sup>a</sup></b>	<b>Description</b>
J	149	5.51	Translation
A	1	0.03	RNA processing and modification
K	150	5.54	Transcription
L	141	5.22	Replication, recombination and repair
B	0	-	Chromatin structure and dynamics
D	26	0.96	Cell cycle control, mitosis and meiosis
Y	0	-	Nuclear structure
V	34	1.25	Defence mechanism
T	71	2.62	Signal transduction mechanism
M	98	3.62	Cell wall/membrane biogenesis
N	2	0.07	Cell motility
Z	0	-	Cytoskeleton
W	0	-	Extracellular structures
U	21	0.77	Intracellular trafficking and secretion
O	80	2.95	Posttranslational modification, protein turnover, chaperones
X	0	0	Mobilome: prophages, transposons
C	151	5.58	Energy production and conversion
G	213	7.87	Carbohydrate transport and metabolism
E	218	8.06	Amino acid transport and metabolism
F	73	2.69	Nucleotide transport and metabolism
H	85	3.14	Coenzyme transport and metabolism
I	79	2.92	Lipid transport and metabolism
P	95	3.51	Inorganic ion transport and metabolism
Q	33	1.22	Secondary metabolites biosynthesis, transport and catabolism
R	257	9.5	General function prediction only
S	190	7.02	Function unknown
others	537	19.85	Not in COGs

526 **a:** the value based on the total number of protein coding genes in the annotated genome



527 Category abbreviations are as follows: C, energy production and conversion; E, amino acid transport  
528 and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H,  
529 coenzyme transport and metabolism; I, lipid transport and metabolism; P, inorganic ion transport and  
530 metabolism; Q, secondary metabolites biosynthesis, transport and catabolism; X, mobilome: prophages,  
531 transposons; A, RNA processing and modification; B, chromatin structure and dynamics; J, translation,  
532 ribosomal structure and biogenesis; K, transcription; L, replication, recombination and repair; D, cell  
533 cycle control, cell division, chromosomal partitioning; M, cell wall/membrane/envelope biogenesis; N,  
534 cell motility; O, posttranslational modification, protein turnover, chaperones; T, signal transduction  
535 mechanisms; U, intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; W,  
536 extracellular structures; Z, cytoskeleton; R, general function predicted only; S, function unknown.

537

538 **Table 4.** Details of resistance genes, plasmids and virulence factors detected in strain Marseille-P1302<sup>T</sup>

Database	Genes	Coverage %	Identity %
CARD			
(comprehensive antibiotic resistance database)			
	Streptomyces_cinnamoneus_EF-Tu_mutants_conferring_resistance_to_elfamycin	99.58	82.09
	otr(A)	15.36	74.03
	novA	21.36	67.83
	Streptomyces_rishiriensis_parY_mutant_conferring_resistance_to_aminocoumarin	17.38	73.27
	Streptomyces_rishiriensis_parY_mutant_conferring_resistance_to_aminocoumarin	30.63	66.81
	mtrA	70.93	75.2
	Streptomyces_rishiriensis_parY_mutant_conferring_resistance_to_aminocoumarin	94.6	73.91
	novA	84.42	66.83
	novA	31.04	68.2
Resfinder			
	otr(A)_1	15.36	74.03
ARG-ANNOT			
	(Tet)otrA	15.36	74.03
Plasmidfinder			
	-	-	-

---

VFDB

---

tssH-5/clpV	12.18	72.43
clpV1	24.58	69.74
ideR	58.44	74.2
clpV1	14.91	67.24
tssH-5/clpV	21.06	67.86
relA	83.52	68.05
katA	61.91	67.63
phzC1	25.29	73.27

---

539

540

541 **Table 5.** Genomic characteristics of strain Marseille-P1302<sup>T</sup> and other compared species

Species	Strain	Size (Mb)	G+C (%mol)	Gene content	tRNA	rRNA
<i>Nigeribacterium massiliense</i>	Marseille-P1302 <sup>T</sup>	2.93	70.38	2704	46	9
<i>Amycolatopsis albispora</i>	WP1 <sup>T</sup>	9.12	72.2	8434	62	12
<i>Microbacterium mangrovi</i>	MUSC 115 <sup>T</sup>	4.41	70.9	4066	50	3
<i>Cutibacterium acnes</i>	ATCC 6919 <sup>T</sup>	2.49	60	2308	46	9
<i>Aeromicrobium massiliense</i>	JC14 <sup>T</sup>	3.31	77.32	3290	48	5
<i>Luteococcus japonicus</i>	DSM 10546 <sup>T</sup>	3.45	67	3270	55	9
<i>Salinispora tropica</i>	CNB 440 <sup>T</sup>	5.18	71.5	4662	61	9
<i>Tessaracoccus flavescens</i>	SST 39 <sup>T</sup>	3.60	68.4	3515	52	6
<i>Nocardiodes gilvus</i>	XZ17 <sup>T</sup>	3.33	68.9	3153	52	5
<i>Microbacterium phosphovorius</i>	NM-1 <sup>T</sup>	5.68	68	5190	51	3
<i>Brevilactibacter sinopodophyllii</i>	KCTC 33808 <sup>T</sup>	3.07	70.79	2917	50	5
<i>Propionicicella superfundia</i>	DSM 22317 <sup>T</sup>	3.61	69.99	3295	48	3
<i>Brevilactibacter flavus</i>	VG341 <sup>T</sup>	3.00	67.69	2771	49	3

542

543

544

545 **Table 6.** dDDH values obtained by comparing strain Marseille-P1302<sup>T</sup> with other closely related species.

	<i>Amycolatopsis albipora</i> strain WP1 <sup>T</sup> (NZ_CP 15163)	<i>Microbacterium mangrove</i> strain MUSC 115 <sup>T</sup> (NZ_JTDK 00000000.1)	<i>Cutibacterium acnes</i> strain ATCC 6919 <sup>T</sup> (NZ_CP 023676)	<i>Aeromicrobium massiliense</i> strain JC14 <sup>T</sup> (CAHG0 1000001.1)	<i>Luteococcus japonicus</i> strain DSM 10546 <sup>T</sup> (NZ_RKH G00000000. 1)	<i>Salinispor a tropica</i> strain CNB 440(NC_0 09380.1)	<i>Tessaracoccus flavescens</i> strain SST 39 <sup>T</sup> (NZ_CP01 9607.1)	<i>Nocardiodes gilvus</i> strain XZ17 <sup>T</sup> (NZ_ QHKY0000 0000.1)	<i>Microtholunatus phosphovoru</i> s strain NM- 1 <sup>T</sup> (NC_0156 35.1)	<i>Propionisicella superfundia</i> DSM 22317 <sup>T</sup> (AUIA0 1000001.1)	<i>Brevilactibacter sinopodophyllii</i> KCTC 33808 <sup>T</sup> (SDMQ01 000001.1)	<i>Brevilactibacter flavus</i> strain VG341 <sup>T</sup> (PPCV 01000001.1)
dDDH	18.7	17.9	22.8	18.4	20.3	20.2	19.8	18.6	19.8	19.1	19.8	19.4

546