



**HAL**  
open science

## Draft genome and description of *Negativicoccus massiliensis* strain Marseille-P2082, a new species isolated from the gut microbiota of an obese patient

Amadou Hamidou Togo, Awa Diop, Mamadou Lamine Tall, Matthieu Million, Saber Khelaifia, Marie Maraninchi, Didier Raoult, Pierre-Edouard Fournier, Gregory Dubourg

### ► To cite this version:

Amadou Hamidou Togo, Awa Diop, Mamadou Lamine Tall, Matthieu Million, Saber Khelaifia, et al.. Draft genome and description of *Negativicoccus massiliensis* strain Marseille-P2082, a new species isolated from the gut microbiota of an obese patient. *Antonie van Leeuwenhoek*, 2020, 113 (7), pp.997-1008. 10.1007/s10482-020-01414-5 . hal-03149954

**HAL Id: hal-03149954**

**<https://hal.science/hal-03149954>**

Submitted on 13 May 2021

**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.

1 **Draft genome and description of *Negativicoccus massiliensis* strain Marseille-P2082, a**  
2 **new species isolated from the gut microbiota of an obese patient**

3 Amadou Hamidou Togo<sup>1,2</sup>, Awa Diop<sup>3</sup>, Mamadou Lamine Tall<sup>1,2</sup>, Matthieu Million<sup>1,2</sup>,  
4 Saber Khelaifia<sup>1,2</sup>, Marie Maraninchi<sup>4</sup>, Didier Raoult<sup>1,2</sup>, Pierre-Edouard Fournier<sup>2,3</sup>,  
5 Grégory Dubourg<sup>1,2\*</sup>

6 1. Aix Marseille Univ, IRD198, AP-HM, MEPHI, Marseille, France.

7 2. IHU Méditerranée Infection, Marseille, France

8 3. Aix Marseille Univ, IRD, SSA, VITROME, IHU-Méditerranée Infection, Marseille,  
9 France.

10 4. Aix Marseille Université, NORT “Nutrition, Obesity and Risk of Thrombosis”,  
11 INSERM1062, INRA1260, Marseille, France.

12  
13 **\*Corresponding author:** Gregory Dubourg, <sup>1</sup>Aix Marseille Univ, IRD198, APHM, MEPHI,  
14 IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13385, Marseille cedex 05,  
15 France.

16 **Tel:** +33 413 732 401

17 **Fax:** +33 413 732 402

18 **email:** gregory.dubourg@univ-amu.fr

19 **Keywords:** *Culturomics; taxonomy; Negativicoccus massiliensis; human gut microbiota;*  
20 *obesity; bariatric surgery*

21 **Abstract**

22 Strain Marseille-P2082, an anaerobic, non-motile, asporogenous, Gram-negative, coccoid  
23 bacterium was isolated from the faeces of a 33 year-old obese French woman before bariatric  
24 surgery. The isolate exhibits 98.65% 16S rRNA gene nucleotide sequence similarity with  
25 *Negativicoccus succinicivorans* strain ADV 07/08/06-B-1388<sup>T</sup>, its current closest  
26 phylogenetic neighbour with standing in nomenclature. However, the dDDH relatedness  
27 between the new isolate and *N. succinicivorans* type strain ADV 07/08/06-B-1388<sup>T</sup> is  $52.5 \pm$   
28 2.7%. Strain Marseille-P2082 has a genome of 1,360,589 bp with a 51.1% G+C content. Its  
29 major fatty acids were identified as C<sub>18:1n9</sub>, C<sub>18:0</sub> and C<sub>16:0</sub>. Based on its phenotypic, genomic  
30 and phylogenetic characteristics, strain Marseille-P2082<sup>T</sup> (= CSURP2082 (Collection de  
31 Souches de l'Unité des Rickettsies) = DSM 100853) is proposed as the type strain of the  
32 novel species *Negativicoccus massiliensis* sp. nov. The 16S rRNA gene sequence and whole-  
33 genome shotgun sequence have been deposited in EMBL-EBI under accession numbers  
34 LN876651 and LT700188, respectively.

## 35 **Introduction**

36 Obesity is a major public health problem that increases the risk of several diseases  
37 such as metabolic diseases (type II diabetes) and cardiovascular diseases. According to Ng et  
38 al (2014), the world rate of obesity has increased between 1980 and 2013. Obesity results  
39 from an imbalance between feed quantity intake and energy expenditure. This process  
40 involves both genetic and environmental factors (Frayling et al. 2007; Cecil et al. 2008). More  
41 recently, obesity has been associated with an imbalance of the gut microbiota composition  
42 (Ley et al. 2006; Turnbaugh et al. 2006; Million et al. 2012).

43 Several methods of treating obesity have been proposed, mainly dietary and lifestyle  
44 measures, medical and surgical treatments. Bariatric surgery is one of the most effective  
45 treatments for obesity. It is also associated with an increase in microbial diversity (Zhang et  
46 al. 2009; Kong et al. 2013).

47 For the purpose of assessing gut microbiota dynamics from obese people before and after  
48 bariatric surgery, we studied the gut microbiota of stool samples from obese subjects before  
49 and after bariatric surgery by culturomics (Lagier et al. 2015). During this study, we isolated a  
50 new anaerobic, Gram-negative coccus strain Marseille-P2082<sup>T</sup>, from a stool sample of a 33  
51 year-old French woman living in Marseille after bariatric surgery. This new bacterium has  
52 been previously reported as a new species announcement without a thorough description and  
53 was provisionally named "*Negativicoccus massiliensis*" strain AT7 (Togo et al. 2016c).  
54 Here, we describe the phenotypic characteristics of strain Marseille-P2082<sup>T</sup> (formerly AT7)  
55 together with the description of its complete genome sequencing and annotation.

## 56 57 **Material and methods**

### 58 **Sample collection, strain Marseille-P2082<sup>T</sup> isolation and identification**

59 A stool sample was collected in November 2011 from a 33 year-old obese French  
60 woman with a body mass index of 38.6 kg/m<sup>2</sup>. Written and informed consent was obtained

61 from the patient at the Nutrition, Metabolic disease and Endocrinology service, in la Timone  
62 Hospital, Marseille, France. The study and the assent procedure were approved by the local  
63 ethics committee of IFR 48, under number 09 -022. The stool sample was stored at -80°C  
64 after collection and studied by a microbial culturomics approach.

65 The strain was isolated under anaerobic conditions. To isolate this new strain, 1g of stool  
66 sample was injected into an anaerobic blood culture vial (BACTEC Lytic/10 vials of  
67 anaerobic culture/F), enriched with 4ml (5%) of rumen juice sterilised by filtration and 4ml  
68 (5%) of sheep blood, then incubated at 37°C. A one-month follow-up was conducted as  
69 described elsewhere (Togo et al. 2017). Emerged colonies were cultivated under the same  
70 conditions for isolation and then identified by MALDI-TOF as described elsewhere (Seng et  
71 al. 2009).

72 *Negativicoccus succinicivorans* strain ADV 07/08/06-B-1388<sup>T</sup> (Marchandin et al. 2010) was  
73 generously provided by Professor H el ene Marchandin (HydroSciences Montpellier, CNRS,  
74 IRD, University of Montpellier, Department of Microbiology, N imes University Hospital,  
75 France) and cultured under comparable conditions.

## 76 **Phylogenetic analysis**

77 The 16S rRNA gene sequencing of the strain Marseille-P2082 was performed as  
78 previously reported (Drancourt et al. 2000) using the fD1-rP2 primers, a GeneAmp PCR  
79 System 2720 thermal cycler (Applied Bio systems, Bedford, MA, USA) and an ABI Prism  
80 3130-XL capillary sequencer (Applied Biosciences, Saint Aubin, France).

81 For taxonomic assignment, CodonCode Aligner (101 Victoria Street Centerville, MA  
82 02632, USA) was used to correct sequences and BLASTn searches was performed in the  
83 NCBI (National Centre for Biotechnology Information) web server at  
84 <http://blast.ncbi.nlm.nih.gov/gate1.inist.fr/Blast.cgi> using the 16S ribosomal RNA sequences  
85 (Bacteria and Archaea) database and excluding all sequences from uncultured species. Only

86 reference sequences from the type strains of the closely related validly named species were  
87 considered. Sequences alignment was performed by a multiple sequence alignment program,  
88 ClustalW (Thompson et al. 2002). The Maximum Likelihood method based on the Kimura 2-  
89 parameter model was used to infer the tree. Rapid bootstrapping in conjunction with the  
90 autoMRE bootstopping criterion (Pattengale et al. 2010) and subsequent search for the best  
91 tree was used; for Maximum Parsimony analysis, 1000 bootstrapping replicates were used in  
92 conjunction with tree-bisection-and-reconnection branch swapping and ten random sequence  
93 addition replicates. Reliability of the nodes was estimated as to the percentages of bootstrap  
94 values obtained by repeating the analysis 1000 times to generate a consensus tree. The  
95 sequences were checked for a compositional bias using the X<sup>2</sup> test as implemented in PAUP\*  
96 (Swofford 2002).

#### 97 **Phenotypic and chemical characteristic analysis**

98 Colony morphology and pigmentation were observed after cultivation of the strain on  
99 Columbia agar (bioMérieux, Marcy l'Etoile, France) at 37°C for 48 hours. Phenotypic  
100 characteristics of the strain was determined as previously described (Togo et al. 2016a, 2017).  
101 API strips: (API<sup>®</sup> 20A, API<sup>®</sup> Rapid ID 32A, API<sup>®</sup> ZYM, API<sup>®</sup> and API<sup>®</sup> 50 CH) were used  
102 according to the manufacturer's instruction (bioMérieux). Growth temperature range from 25,  
103 28, 37, 45 and 55°C was tested in aerobic, anaerobic and microaerophilic conditions. The  
104 growth of strain Marseille-P2082 was attempted at various pH values (6, 6.5, 7 and 8.5). The  
105 salt tolerance of the strain was also tested using various NaCl concentrations (5, 10, 50, 75  
106 and 100g/l) on Schaedler agar enriched with 5% Sheep Blood (bioMérieux) in an anaerobic  
107 atmosphere at 37°C. Anaerobic conditions were generated by incubating the culture in an  
108 anaerobic jar using the GENbag anaer system (bioMérieux). Cellular fatty acid methyl ester  
109 analysis was performed by Gas Chromatography/Mass Spectrometry (GC/MS) as described  
110 elsewhere (Togo et al. 2017).

111 **Genome sequencing and assembly**

112 Genomic DNA (gDNA) of strain Marseille-P2082 and *N. succinivorans* strain ADV  
113 07/08/06-B-1388<sup>T</sup> was sequenced using the MiSeq Technology (Illumina Inc, San Diego, CA,  
114 USA) with the Mate Pair strategy and assembled as previously described (Togo et al. 2016b).

115 **Genome annotation and comparison**

116 For the genome annotation of strain Marseille-P2082, Open Reading Frames (ORFs)  
117 were predicted using Prodigal (Hyatt et al. 2010) and the predicted protein sequences were  
118 searched against the GenBank and Clusters of Orthologous Groups (COGs) databases by  
119 BLASTP as previously described (Togo et al. 2017). The draft genome of *N. succinivorans*  
120 strain ADV 07/08/06-B-1388<sup>T</sup> was obtained in this study. For comparison, genomes were  
121 automatically retrieved for closely related species in the 16S rRNA tree: *Dialister*  
122 *micraerophilus* strain DSM 19965 (NZ\_AFBB000000000), *Dialister pneumosintes* strain  
123 F0677 (NZ\_CP017037), *Dialister succinatiphilus* strain YIT 11850 (NZ\_ADLT000000000),  
124 *Dialister invisus* strain DSM 15470 (NZ\_ACIM000000000) and *Selenomonas bovis* strain  
125 DSM 23594 (NZ\_ARLB010000000) were selected. For each selected genome, complete  
126 genome sequence, proteomes sequences and Orfeomes sequences were retrieved from the  
127 FTP of NCBI. As a part of the taxonogenomics description, all proteomes were analysed with  
128 proteinOrtho (Lechner et al. 2011). Then for each pair of genomes, a similarity score was  
129 computed. This score is the mean value of nucleotide similarity between all pairs of  
130 orthologues between the two genomes studied (Ramasamy et al. 2014). An annotation of the  
131 entire proteome was performed to define the distribution of functional classes of predicted  
132 genes according to the clusters of orthologous groups of proteins. The genome of strain  
133 Marseille-P2082 was locally aligned 2-by-2 using BLSAT algorithm (Kent 2002; Auch et al.  
134 2010) against each of the selected previously cited genomes and digital DNA-DNA  
135 hybridization (dDDH) values were estimated from a generalised model (Meier-Kolthoff et al.

136 2013), and the Orthologous Average Nucleotide Identity (OrthoANI) were calculated using  
137 Orthologous Average Nucleotide Identity Tool (OAT) software (Lee et al. 2016).

## 138 **Results and discussion**

### 139 **MALDI-TOF analysis**

140 The reference spectrum generated from isolated colonies was unable to be matched  
141 with any of those in the Bruker database (Fig. 1). A gel view was performed in order to detect  
142 the spectral differences of strain Marseille-P2082 and closely related species (Fig. 2).

### 143 **Phylogenetic analysis**

144 The 16S rRNA phylogenetic analysis showed that strain Marseille-P2082<sup>T</sup> exhibits a  
145 98.65% sequence similarity with *N. succinicivorans* (Marchandin et al. 2010), which is  
146 classified in the family *Veillonellaceae* (Rogosa 1971). A Maximum Likelihood phylogenetic  
147 tree (Fig. 3) based on 16S rRNA gene sequences showed that strain Marseille-P2082 is  
148 closely related to *N. succinicivorans*. The 16S rRNA sequence of strain Marseille-P2082 has  
149 been deposited in EMBL-EBI under accession number LN876651.

150 **Phenotypic, biochemical and chemotaxonomic characterisation**

151 Strain Marseille-P2082 was observed to be a Gram-stain negative, non-motile, coccoid  
152 bacterium. The strain exhibits no catalase nor oxidase activity. Growth was found to occur  
153 between 28 and 55°C, but optimal growth was observed at 37°C after 72 hours of incubation  
154 in both anaerobic and microaerophilic atmospheres on 5% sheep blood Columbia agar  
155 (bioMérieux). No growth of the strain was observed in an aerobic atmosphere. No growth of  
156 the bacterium was observed using NaCl concentration of 10g/l or more on 5% sheep blood  
157 Schaedler agar (bioMérieux). Strain Marseille-P2082 was found to grow at pH values ranging  
158 from 6 to 8.5 but the optimal pH was 7.5. Colonies were observed to be circular, translucent  
159 and very small, with a mean diameter less than 0.5 mm. No haemolysis was seen on  
160 Columbia agar (bioMérieux) after 72 hours of incubation in anaerobic condition at 37°C.  
161 Cells were found to have a diameter ranging from 0.2 to 0.5µm by electron microscopy  
162 (Supplementary Fig. 1).

163 Using rapid ID 32A strips, arginine dihydrolase, arginine arylamidase, histidine  
164 arylamidase, leucyl glycine arylamidase, phenylalanine arylamidase and tyrosine arylamidase  
165 activities were found to be positive. Urease and indole are not produced, and nitrate is not  
166 reduced. Mannose and raffinose are not fermented. Alkaline phosphatase, alanine  
167 arylamidase, glycine arylamidase, glutamyl glutamic acid arylamidase, leucine arylamidase,  
168 proline arylamidase, pyroglutamic acid arylamidase, serine arylamidase, glutamic acid  
169 decarboxylase, n-acetyl-β-glucosaminidase, α-arabinose, α-galactosidase, β-galactosidase, β-  
170 galactosidase 6 phosphate, α-glucosidase, β-glucosidase, β-glucuronidase, and α-fucosidase  
171 activities were found to be negative.

172 Using API® Zym strips, positive reactions were observed for esterase, esterase lipase, valine  
173 arylamidase, acid phosphatase, and naphthol-AS-BI-phosphohydrolase, but alkaline  
174 phosphatase, lipase, leucine arylamidase, cystine arylamidase, N-acetyl-β-glucosaminidase,

175 trypsin,  $\alpha$ -chymotrypsin,  $\alpha$ -galactosidase,  $\beta$ -galactosidase,  $\beta$ -glucuronidase,  $\alpha$ -glucosidase,  $\beta$ -  
176 glucosidase,  $\alpha$ -mannosidase and  $\alpha$ -fucosidase activities were negative.  
177 Using API<sup>®</sup> 20A strips, positive reactions were observed with D-glucose, D-lactose, D-  
178 maltose, D-mannose, D-melezitose, D-saccharose, and glycerol. Indole production, esculin  
179 hydrolyse, urease, D-cellulose, D-mannitol, D-raffinose, D-sorbitol D-trehalose, D-xylose,  
180 gelatine, salicin, L-arabinose and L-rhamnose were found to be negative.  
181 Using API 50 CH strips, positive fermentation reactions were observed with aesculin, D-  
182 galactose, D-glucose, D-fructose, D-lactose, D-maltose, D-mannose, D-melezitose, D-ribose,  
183 D-saccharose, D-tagatose, D-turanose, glycerol, N-acetylglucosamine, potassium 5-  
184 cetogluconate and starch. Negative reactions were observed for fermentation of adonitol,  
185 amygdaline, arbutine, D-arabinose, D-arabitol, D-cellobiose, D-fucose, D-lyxose, D-mannitol,  
186 D-melibiose, D-raffinose, D-sorbitol, D-trehalose, D-xylose, dulcitol, erythritol, gentiobiose,  
187 glycogen, inositol, inulin, L-arabinose, L-arabitol, L-fucose, L-rhamnose, L-Sorbose, L-  
188 xylose, methyl- $\alpha$ D-glucopyranoside, methyl- $\alpha$ -D-mannopyranoside, Methyl- $\beta$ D-  
189 xylopyranoside, potassium gluconate, potassium 2-ceto gluconate, salicin, and xylitol. The  
190 strain Marseille-P2082 differs from *N. succinicivorans* in its ability to ferment 14 sugars  
191 whereas *N. succinicivorans* is assaccharolytic. The differential characteristics of strain  
192 Marseille-P2082<sup>T</sup> compared with its close relatives are detailed in Table 1.

193 The most abundant cellular fatty acids of strain Marseille-P2082<sup>T</sup> were identified as  
194 C<sub>18:1n9</sub> (37.2%), C<sub>18:0</sub> (34.7%), C<sub>16:0</sub> (21.3%), C<sub>18:2n6</sub> (6.3%) and a minor amount of C<sub>14:0</sub>  
195 (<1%).

## 196 **Genome properties**

197 The draft genome from strain Marseille-P2082 has been deposited in EMBL-EBI  
198 under accession number LT700188. The draft genome is 1, 360, 589 bp long with a 51.1%  
199 G+C content (Fig. 4). It is composed of 1 scaffold (1 contig). Of the 1,353 predicted genes,

200 1,276 are protein-coding genes, 68 are RNAs genes (12 rRNA, 52 tRNA and 4 other RNA)  
201 and 9 were pseudogenes. A total of 995 genes (77.9 %) were assigned a putative function (by  
202 COGs or by NR blast), 23 genes (1.8%) were identified as ORFans, and 5 genes were  
203 associated with polyketide synthase (PKS) or non-ribosomal peptide synthetase  
204 (NRPS)(Conway and Boddy 2013). The remaining 253 protein coding genes (19.8%) were  
205 annotated as hypothetical proteins. Using ARG-ANNOT (Gupta et al. 2014), no antibiotic  
206 resistance genes were found. The distribution of genes into COGs functional categories is  
207 presented in Table 2.

### 208 **Genome comparison**

209 The draft genome (1.36Mb) of the strain Marseille-P2082 is larger than those of *D.*  
210 *succinatiphilus*, *D. pneumosintes* and *D. micraerophilus* (1.25, 1.27 and 1.28 Mb respectively)  
211 but smaller than those of *N. succinicivorans*, *D. invisus* and *S. bovis* (1.48, 1.89 and 2.69 Mb).  
212 The G+C content of strain Marseille-P2082<sup>T</sup> (51.1%) is higher than those of *D. pneumosintes*,  
213 *D. micraerophilus*, *D. invisus* and *D. succinatiphilus* (35.2%, 35.4%, 45.5% and 50.4%  
214 respectively), similar to that of *N. succinicivorans*, but smaller than that of *S. bovis* (59.2%).  
215 The gene content of strain Marseille-P2082 (1,353) is smaller than those of *N.*  
216 *succinicivorans*, *D. invisus*, *D. succinatiphilus* and *S. bovis* (1,454, 1,829, 2,280, and 2,569  
217 respectively) but higher than those of *D. micraerophilus* and *D. pneumosintes* (1,237 and  
218 1,263 respectively). However, the gene distribution into COGs categories was similar among  
219 all compared genomes (Table 2, Fig. 5).

220 The dDDH values of strain Marseille-P2082 ranged from 23.2%±4.8% with *S. bovis* to  
221 53.7±5.3 % with *N. succinicivorans* strain ADV (Table 3). The OrthoANI values ranged from  
222 63.35% between strain Marseille-P2082 and *D. pneumosintes* to 93.78% with *N.*  
223 *succinicivorans*. These values are lower than the threshold value (95-96%) recommended to  
224 delineate prokaryotic species (Konstantinidis and Tiedje 2005; Rodriguez-R and

225 Konstantinidis 2014). The OrthoANI values between strain Marseille-P2082 and closely  
226 related species are shown in Table 4 and Fig. 6.

227 Based on phenotypic, chemotaxonomic, genomic and phylogenetic characteristics, strain  
228 Marseille-P2082<sup>T</sup> (=CSUR P2082<sup>T</sup>, DSM 100853<sup>T</sup>) is designated as the type strain of a novel  
229 species, here named *Negativicoccus massiliensis*.

230 **Description of “*Negativicoccus massiliensis*” sp. nov.**

231 *Negativicoccus massiliensis* (*mas.si.li.en'sis*. L. masc. adj. *massiliensis*, of Massilia, the Latin  
232 name of Marseille, where the type strain was isolated).

233 Cells are Gram-negative, non-motile, asporogenous and coccoid, with a mean diameter of  
234 0.5µm. Columbia agar-grown colonies are circular, flat, translucent with a mean diameter  
235 smaller than 0.5 mm after 72 hours of incubation in anaerobic conditions. Catalase and  
236 oxidase activities are absent. The optimum growth temperature is 37°C. Arginine hydrolase  
237 activity is positive. Not able to reduce nitrate to nitrite. Indole and urease are not produced,  
238 gelatin is not liquefied and aesculin is not hydrolysed. The G+C content of the draft genome  
239 of the type strain is 51.1%.

240 The type strain Marseille-P2082<sup>T</sup> (=CSUR P2082<sup>T</sup>, =DSM 100853<sup>T</sup>) was isolated  
241 from a stool sample of a morbidly obese French woman. The habitat of this microorganism is  
242 the human gut. The 16S rRNA gene sequence and whole-genome shotgun sequence of strain  
243 Marseille-P2082<sup>T</sup> have been deposited in EMBL-EBI under accession numbers LN876651  
244 and LT700188, respectively.

245

246 **Data deposit.** The whole-genome shotgun sequence of *N. succinicivorans* strain ADV  
247 07/08/06-B-1388<sup>T</sup> has been deposited in NCBI under accession number PRJEB34910.

248

249 **Acknowledgements**

250 The authors thank the Xegen Company (www.xegen.fr) for automating the genomic  
251 annotation process. The authors also thank Noémie Labas for the genome sequencing. We are

252 grateful to Professor H el ene Marchandin for providing the *Negativicoccus succinicivorans*  
253 strain ADV 07/08/06-B-1388<sup>T</sup>.

#### 254 **Funding**

255 This work was supported by the French Government under the « Investissements  
256 d’avenir » (Investments for the Future) program managed by the Agence Nationale de la  
257 Recherche (ANR, fr: National Agency for Research) under the number IHU M editerran ee  
258 Infection 10-IAHU-03.

#### 259 **Conflict of interest**

260 The authors declare no conflict of interest.

#### 261 **Author’s contributions**

262 AHT isolated for the first time the strain Marseille-P2802, performed its phenotypic  
263 characterization and wrote the manuscript; AD and MLD actively participated to genome  
264 analysis; SK actively participated in the laboratory project in which the strain was isolated;  
265 MM and MM actively participated to the specimen collection and the study design; DR  
266 designed and directed the project; PEF corrected the manuscript, verified the accuracy of the  
267 Latin name of the strain; GD corrected the manuscript and acted as the corresponding author.

268

269

#### 270 **References**

271 Auch AF, von Jan M, Klenk H-P, G oker M (2010) Digital DNA-DNA hybridization for  
272 microbial species delineation by means of genome-to-genome sequence comparison.  
273 Stand Genomic Sci 2:117–134. <https://doi.org/10.4056/sigs.531120>

274 Cecil JE, Tavendale R, Watt P, et al (2008) An obesity-associated FTO gene variant and  
275 increased energy intake in children. *N Engl J Med* 359:2558–2566.  
276 <https://doi.org/10.1056/NEJMoa0803839>

277 Conway KR, Boddy CN (2013) ClusterMine360: a database of microbial PKS/NRPS  
278 biosynthesis. *Nucleic Acids Res* 41:D402-407

279 Drancourt M, Bollet C, Carlioz A, et al (2000) 16S ribosomal DNA sequence analysis of a  
280 large collection of environmental and clinical unidentifiable bacterial isolates. *J Clin*  
281 *Microbiol* 38:3623–3630

282 Frayling TM, Timpson NJ, Weedon MN, et al (2007) A common variant in the FTO gene is  
283 associated with body mass index and predisposes to childhood and adult obesity.  
284 *Science* 316:889–894. <https://doi.org/10.1126/science.1141634>

285 Gupta SK, Padmanabhan BR, Diene SM, et al (2014) ARG-ANNOT, a new bioinformatic  
286 tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents*  
287 *Chemother* 58:212–220

288 Hyatt D, Chen G-L, Locascio PF, et al (2010) Prodigal: prokaryotic gene recognition and  
289 translation initiation site identification. *BMC Bioinformatics* 11:119.  
290 <https://doi.org/10.1186/1471-2105-11-119>

291 Kent WJ (2002) BLAT--the BLAST-like alignment tool. *Genome Res* 12:656–664

292 Kong L-C, Tap J, Aron-Wisnewsky J, et al (2013) Gut microbiota after gastric bypass in  
293 human obesity: increased richness and associations of bacterial genera with adipose  
294 tissue genes. *Am J Clin Nutr* 98:16–24. <https://doi.org/10.3945/ajcn.113.058743>

295 Konstantinidis KT, Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. *J*  
296 *Bacteriol* 187:6258–6264. <https://doi.org/10.1128/JB.187.18.6258-6264.2005>

297 Lagier J-C, Hugon P, Khelaifia S, et al (2015) The rebirth of culture in microbiology through  
298 the example of culturomics to study human gut microbiota. *Clin Microbiol Rev*  
299 28:237–264. <https://doi.org/10.1128/CMR.00014-14>

300 Lechner M, Findeiss S, Steiner L, et al (2011) Proteinortho: detection of (co-)orthologs in  
301 large-scale analysis. *BMC Bioinformatics* 12:124. [https://doi.org/10.1186/1471-2105-](https://doi.org/10.1186/1471-2105-12-124)  
302 12-124

303 Lee I, Ouk Kim Y, Park S-C, Chun J (2016) OrthoANI: An improved algorithm and software  
304 for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103.  
305 <https://doi.org/10.1099/ijsem.0.000760>

306 Ley RE, Turnbaugh PJ, Klein S, Gordon JI (2006) Microbial ecology: human gut microbes  
307 associated with obesity. *Nature* 444:1022–1023. <https://doi.org/10.1038/4441022a>

308 Marchandin H, Teyssier C, Campos J, et al (2010) *Negativicoccus succinicivorans* gen. nov.,  
309 sp. nov., isolated from human clinical samples, emended description of the family  
310 *Veillonellaceae* and description of *Negativicutes* classis nov., *Selenomonadales* ord.  
311 nov. and *Acidaminococcaceae* fam. nov. in the bacterial phylum Firmicutes. *Int J Syst*  
312 *Evol Microbiol* 60:1271–1279

313 Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M (2013) Genome sequence-based species  
314 delimitation with confidence intervals and improved distance functions. *BMC*  
315 *Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>

316 Million M, Maraninchi M, Henry M, et al (2012) Obesity-associated gut microbiota is  
317 enriched in *Lactobacillus reuteri* and depleted in *Bifidobacterium animalis* and  
318 *Methanobrevibacter smithii*. *Int J Obes* 36:817–825.  
319 <https://doi.org/10.1038/ijo.2011.153>

320 Ng M, Fleming T, Robinson M, et al (2014) Global, regional, and national prevalence of  
321 overweight and obesity in children and adults during 1980–2013: a systematic analysis

322 for the Global Burden of Disease Study 2013. *The Lancet* 384:766–781.  
323 [https://doi.org/10.1016/S0140-6736\(14\)60460-8](https://doi.org/10.1016/S0140-6736(14)60460-8)

324 Pattengale ND, Alipour M, Bininda-Emonds ORP, et al (2010) How many bootstrap  
325 replicates are necessary? *J Comput Biol J Comput Mol Cell Biol* 17:337–354.  
326 <https://doi.org/10.1089/cmb.2009.0179>

327 Ramasamy D, Mishra AK, Lagier J-C, et al (2014) A polyphasic strategy incorporating  
328 genomic data for the taxonomic description of novel bacterial species. *Int J Syst Evol*  
329 *Microbiol* 64:384–391

330 Rodriguez-R LM, Konstantinidis KT (2014) Bypassing cultivation to identify bacterial  
331 species: Culture-independent genomic approaches identify credibly distinct clusters,  
332 avoid cultivation bias, and provide true insights into microbial species. *Microbe Mag*  
333 9:111–118. <https://doi.org/10.1128/microbe.9.111.1>

334 Rogosa M (1971) Transfer of *Veillonella Prevot* and *Acidaminococcus Rogosa* from  
335 *Neisseriaceae* to *Veillonellaceae* fam. nov., and the Inclusion of *Megasphaera Rogosa*  
336 in *Veillonellaceae*. *Int J Syst Bacteriol* 21:231–233. [https://doi.org/10.1099/00207713-](https://doi.org/10.1099/00207713-21-3-231)  
337 21-3-231

338 Seng P, Drancourt M, Gouriet F, et al (2009) Ongoing revolution in bacteriology: routine  
339 identification of bacteria by matrix-assisted laser desorption ionization time-of-flight  
340 mass spectrometry. *Clin Infect Dis* 49:543–551. <https://doi.org/10.1086/600885>

341 Swofford D (2002) PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods).  
342 Version 4.0b10. Sinauer Associates, Sunderland.

343 Thompson JD, Gibson TJ, Higgins DG (2002) Multiple sequence alignment using ClustalW  
344 and ClustalX. *Curr Protoc Bioinforma* Chapter 2:Unit 2.3.  
345 <https://doi.org/10.1002/0471250953.bi0203s00>

346 Togo AH, Diop A, Dubourg G, et al (2016a) *Butyricimonas phoceensis* sp. nov., a new  
347 anaerobic species isolated from the human gut microbiota of a French morbidly obese  
348 patient. *New Microbes New Infect* 14:38–48.  
349 <https://doi.org/10.1016/j.nmni.2016.07.010>

350 Togo AH, Durand G, Khelaifia S, et al (2017) *Fournierella massiliensis*, gen. nov., sp. nov., a  
351 new human-associated member of the family *Ruminococcaceae*. *Int J Syst Evol*  
352 *Microbiol* 67:1393–1399. <https://doi.org/10.1099/ijsem.0.001826>

353 Togo AH, Khelaifia S, Lagier J-C, et al (2016b) Noncontiguous finished genome sequence  
354 and description of *Paenibacillus ihumii* sp. nov. strain AT5. *New Microbes New*  
355 *Infect* 10:142–150. <https://doi.org/10.1016/j.nmni.2016.01.013>

356 Togo AH, Khelaifia S, Valero R, et al (2016c) ‘*Negativicoccus massiliensis*’, a new species  
357 identified from human stool from an obese patient after bariatric surgery. *New*  
358 *Microbes New Infect* 13:43–44. <https://doi.org/10.1016/j.nmni.2016.05.018>

359 Turnbaugh PJ, Ley RE, Mahowald MA, et al (2006) An obesity-associated gut microbiome  
360 with increased capacity for energy harvest. *Nature* 444:1027–1031.  
361 <https://doi.org/10.1038/nature05414>

362 Zhang H, DiBaise JK, Zuccolo A, et al (2009) Human gut microbiota in obesity and after  
363 gastric bypass. *Proc Natl Acad Sci U S A* 106:2365–2370.  
364 <https://doi.org/10.1073/pnas.0812600106>  
365

366 **Table 1:** Differential characteristics of strain Marseille-P2082 (1) compared with the closely related species *N. succinivorans* strain ADV (Data  
367 from Marchandin et al., 2010) (2); *Dialister micraerophilus* (3); *Dialister invisus* (4); *Dialister succinatiphilus* (5); *Dialister pneumosintes* (6);  
368 *Selenomonas bovis* (7)

<b>Properties</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Motility	–	–	–	–	+	–	+
Spore formation	–	–	–	–	–	–	–
Microaerophilic growth	+	+	+	–	NR	–	–
Indole production	–	–	–	–	NR	–	–
Nitrate reductase	–	–	+	+	+	+	–
Acid production from:							
<b>Aesculin</b>	+	–	–	+	–	+	+
Arabinose	–	–	–	v	–	+	+
Cellobiose	–	–	–	+	–	+	+
<b>Fructose</b>	+	–	–	+	+	+	–
<b>Galactose</b>	+	–	–	+	–	+	–
<b>Glucose</b>	+	–	–	+	+	+	+

---

<b>Glycerol</b>	+	-	-	+	-	-	-
Inositol	-	-	-	+	-	+	-
Lactose	v	-	-	+	-	+	+
<b>Maltose</b>	+	-	-	+	+	+	-
<b>Mannitol</b>	+	-	-	-	+	+	-
<b>Mannose</b>	+	-	-	+	-	+	+
Melibiose	-	-	-	+	-	+	+
Raffinose	v	-	-	-	w	+	+
Rhamnose	-	-	-	-	-	+	-
Ribose	-	-	-	+	-	+	-
Salicin	v	-	-	+	-	+	+
Sorbitol	-	-	-	+	-	-	-
<b>Sucrose</b>	+	-	-	+	+	+	+
<b>Trehalose</b>	+	-	-	+	-	+	+
Xylose	v	-	-	v	-	+	-

---

Xylitol	-	-	-	+	-	+	-
Isolation source	Human feces	human clinical samples	Human tongue biofilm	Cattle rumen	Human gingival crevice	Human and Pig feces	Yak rumen
<b>G+C content (mol %)</b>	<b>51.1</b>	<b>48.3</b>	38.1	58.1	57.3	58.1	59.2

369 + = Positive; - = Negative; w = weakly positive; v = Variable; NR = Data Not Reported, In bold difference between strain Marseille-P2082 and

370 *N. succinicivorans* strain ADV



372 **Table 2: Number of genes associated with the 26 general COG functional categories of strain Marseille-P2082<sup>T</sup>**

373 1) Strain Marseille-P2082; 2) *N. succinicivorans* strain ADV; 3) *Dialister micraerophilus*; 4) *Dialister invisus*; 5) *Dialister succinatiphilus*; 6)

374 *Dialister pneumosintes*; 7) *Selenomonas bovis*

Code	1	2	3	4	5	6	7	Description
<b>J</b>	170	172	172	182	178	172	190	Translation, ribosomal structure and biogenesis
<b>A</b>	0	0	0	0	0	0	0	RNA processing and modification
<b>K</b>	58	60	62	79	98	59	149	Transcription
<b>L</b>	79	76	77	80	104	76	118	Replication, recombination and repair
<b>B</b>	1	1	0	0	0	1	0	Chromatin structure and dynamics
<b>D</b>	20	22	25	21	24	18	40	Cell cycle control, cell division, chromosome partitioning
<b>Y</b>	0	0	0	0	0	0	0	Nuclear structure
<b>V</b>	31	32	37	50	47	46	58	Defence mechanisms
<b>T</b>	45	44	40	57	95	38	131	Signal transduction mechanisms
<b>M</b>	82	85	85	108	122	93	161	Cell wall/membrane/envelope biogenesis
<b>N</b>	8	8	8	10	12	11	80	Cell motility
<b>Z</b>	0	0	0	0	0	0	0	Cytoskeleton
<b>W</b>	4	4	9	6	9	7	12	Extracellular structures
<b>U</b>	22	22	27	25	24	23	40	Intracellular trafficking, secretion, and vesicular transport
<b>O</b>	47	49	47	61	64	43	89	Posttranslational modification, protein turnover, chaperones
<b>X</b>	1	9	7	30	31	2	37	Mobilome: prophages, transposons
<b>C</b>	46	46	45	67	78	42	115	Energy production and conversion
<b>G</b>	40	42	35	45	75	41	143	Carbohydrate transport and metabolism
<b>E</b>	91	96	87	142	195	76	196	Amino acid transport and metabolism
<b>F</b>	62	64	58	64	66	56	75	Nucleotide transport and metabolism
<b>H</b>	84	84	50	105	109	61	129	Coenzyme transport and metabolism

<b>I</b>	38	39	42	41	54	39	52	Lipid transport and metabolism
<b>P</b>	69	72	50	110	84	47	88	Inorganic ion transport and metabolism
<b>Q</b>	16	15	11	15	21	10	31	Secondary metabolites biosynthesis, transport and catabolism
<b>R</b>	90	87	81	125	151	81	181	General function prediction only
<b>S</b>	61	63	49	64	72	49	87	Function unknown
<b>-</b>	241	329	219	439	649	230	556	Not in cog

375

376 **Table 3: Pairwise comparison of strain Marseille-P2082 with other species using GGDC, formula 2 (dDDH estimates based on identities /**  
 377 **HSP length)\***

378 1) Strain Marseille-P2082; 2) *N. succinicivorans* strain ADV; 3) *Dialister micraerophilus*; 4) *Dialister invisus*; 5) *Dialister succinatiphilus*;  
 379 6)*Dialister pneumosintes*; 7) *Selenomonas bovis*

	1	2	3	4	5	6	7
<b>Strain Marseille-P2082</b>	100%	53.7%±5.3	33.2% ± 4.9	30.5% ± 4.9	30.1% ± 4.9	27.7% ± 4.8	23.2% ± 4.8
<i>N. succinicivorans</i>		100%	34.1% ± 4.9	23.9% ± 4.8	27.8% ± 4.9	27% ± 4.9	22.4% ± 4.8
<i>D. micraerophilus</i>			100%	20.9% ± 4.6	23% ± 4.7	19% ± 4.6	28.8% ± 4.9
<i>D. invisus</i>				100%	22.7% ± 4.8	22.9% ± 4.7	28.7% ± 4.8
<i>D. succinatiphilus</i>					100%	24.5% ± 4.8	25.6% ± 4.8
<i>D. pneumosintes</i>						100%	26.3% ± 4.9
<i>S. bovis</i>							100%

380

381 **Table 4: OrthoANI values calculated from the OAT software of strain Marseille-P2082 compared with those of its phylogenetically close**  
 382 **neighbours**

383 1) Strain Marseille-P2082; 2) *N. succinicivorans* strain ADV; 3) *Dialister micraerophilus*; 4) *Dialister invisus*; 5) *Dialister succinatiphilus*; 6)  
 384 *Dialister pneumosintes*; 7) *Selenomonas bovis*

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Strain Marseille-P2082	100%	93.78%	64.35%	64.67%	64.83%	63.86%	64.22%
<i>N. succinicivorans</i>		100%	64.52%	64.46%	64.87%	63.35%	63.95%
<i>D. micraerophilus</i>			100%	67.90%	67.05%	68.17%	62.13%
<i>D. invisus</i>				100%	70.81%	67.46%	62.88%
<i>D. succinatiphilus</i>					100%	66.62%	63.98%
<i>D. pneumosintes</i>						100%	61.51%
<i>S. bovis</i>							100%

385

386 **Figure legends**

387 **Figure 1: Reference mass spectrum from strain Marseille-P2082<sup>T</sup>.** Spectra from 12  
388 individual colonies were compared and a reference spectrum was generated.

389 **Figure 2: Gel view comparing strain Marseille-P2082<sup>T</sup> to other closely related species.**

390 The gel view displays the raw spectra of loaded spectrum files arranged in a pseudo-gel like  
391 look. The x-axis records the m/z value. The left y-axis displays the running spectrum number  
392 originating from subsequent spectra loading. The peak intensity is expressed by a Grey scale  
393 scheme code. The colour bar and the right y-axis indicate the relation between the colour a  
394 peak is displayed with and the peak intensity in arbitrary units. Displayed species are  
395 indicated on the left.

396 **Figure 3: Maximum Likelihood phylogenetic tree based on 16S rRNA sequence**  
397 **comparison highlighting the position of strain Marseille-P2082<sup>T</sup> compared to other**  
398 **closely related species.**

399 The corresponding GenBank accession numbers for 16S rRNA genes of each strain are  
400 indicated in brackets. Sequences were aligned using ClustalW with default parameters and  
401 Phylogenies were inferred by the GGDC web server available at <http://ggdc.dsmz.de/> using  
402 the DSMZ phylogenomics pipeline. The scale bar represents a 3% nucleotide sequence  
403 divergence.

404 **Figure 4: Graphical circular map of the genome of strain Marseille-P2082<sup>T</sup>.**

405 From outside to the centre: Contigs (red/grey), COG category of genes on the forward strand  
406 (three circles), genes on forward strand (blue circle), genes on the reverse strand (red circle),  
407 COG category on the reverse strand (three circles), GC content.

408 **Figure 5: Distribution of functional classes of predicted genes according to the COGs of**  
409 **strain Marseille-P2082<sup>T</sup> compared with closely related species.**

410 1) Strain Marseille-P2082; 2) *N. succinicivorans* strain ADV; 3) *Dialister micraerophilus*; 4)  
411 *Dialister invisus*; 5) *Dialister succinatiphilus*; 6) *Dialister pneumosintes*; 7) *Selenomonas*  
412 *bovis*.

413 **Figure 6:** Heatmap generated with OrthoANI values calculated using the OAT software  
414 between strain Marseille-P2082 and other closely related species with standing in  
415 nomenclature.

416 **Supplementary Figure 1:** Transmission electron microscopy of strain Marseille-P2082T  
417 using a Tecnai G20 (FEI company) at an operating voltage of 60kV. The scale bar represents  
418 200  $\mu\text{m}$ .