



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-03149954v1>

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.



HAL Authorization

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^{1,2}, Awa Diop³, Mamadou Lamine Tall^{1,2}, Matthieu Million^{1,2},
Saber Khelaifia^{1,2}, Marie Maraninchi⁴, Didier Raoult^{1,2}, Pierre-Edouard Fournier^{2,3},
Grégory Dubourg^{1,2*}

1. Aix Marseille Univ, IRD198, AP-HM, MEPHI, Marseille, France.
2. IHU Méditerranée Infection, Marseille, France
3. Aix Marseille Univ, IRD, SSA, VITROME, IHU-Méditerranée Infection, Marseille, France.
4. Aix Marseille Université, NORT “Nutrition, Obesity and Risk of Thrombosis”, INSERM1062, INRA1260, Marseille, France.

***Corresponding author:** Gregory Dubourg, ¹Aix Marseille Univ, IRD198, APMH, MEPHI, IHU-Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13385, Marseille cedex 05, France.

Tel: +33 413 732 401

Fax: +33 413 732 402

email: gregory.dubourg@univ-amu.fr

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

Abstract

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

Introduction

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

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

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

Material and methods

Sample collection, strain Marseille-P2082^T isolation and identification

A stool sample was collected in November 2011 from a 33 year-old obese French woman with a body mass index of 38.6 kg/m². Written and informed consent was obtained

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

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

Negativicoccus succinicivorans strain ADV 07/08/06-B-1388^T (Marchandin et al. 2010) was generously provided by Professor Hélène Marchandin (HydroSciences Montpellier, CNRS, IRD, University of Montpellier, Department of Microbiology, Nîmes University Hospital, France) and cultured under comparable conditions.

Phylogenetic analysis

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

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

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

Phenotypic and chemical characteristic analysis

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

Genome sequencing and assembly

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

Genome annotation and comparison

For the genome annotation of strain Marseille-P2082, Open Reading Frames (ORFs) were predicted using Prodigal (Hyatt et al. 2010) and the predicted protein sequences were searched against the GenBank and Clusters of Orthologous Groups (COGs) databases by BLASTP as previously described (Togo et al. 2017). The draft genome of *N. succinivorans* strain ADV 07/08/06-B-1388^T was obtained in this study. For comparison, genomes were automatically retrieved for closely related species in the 16S rRNA tree: *Dialister micraerophilus* strain DSM 19965 (NZ_AFB000000000), *Dialister pneumosintes* strain F0677 (NZ_CP017037), *Dialister succinatiphilus* strain YIT 11850 (NZ_ADLT000000000), *Dialister invisus* strain DSM 15470 (NZ_ACIM000000000) and *Selenomonas bovis* strain DSM 23594 (NZ_ARLB010000000) were selected. For each selected genome, complete genome sequence, proteomes sequences and Orfeomes sequences were retrieved from the FTP of NCBI. As a part of the taxonogenomics description, all proteomes were analysed with proteinOrtho (Lechner et al. 2011). Then for each pair of genomes, a similarity score was computed. This score is the mean value of nucleotide similarity between all pairs of orthologues between the two genomes studied (Ramasamy et al. 2014). An annotation of the entire proteome was performed to define the distribution of functional classes of predicted genes according to the clusters of orthologous groups of proteins. The genome of strain Marseille-P2082 was locally aligned 2-by-2 using BLSAT algorithm (Kent 2002; Auch et al. 2010) against each of the selected previously cited genomes and digital DNA-DNA hybridization (dDDH) values were estimated from a generalised model (Meier-Kolthoff et al.

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

Results and discussion

MALDI-TOF analysis

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

Phylogenetic analysis

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

Phenotypic, biochemical and chemotaxonomic characterisation

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

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

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

175 trypsin, α -chymotrypsin, α -galactosidase, β -galactosidase, β -glucuronidase, α -glucosidase, β -
176 glucosidase, α -mannosidase and α -fucosidase activities were negative.

177 Using API[®] 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- α D-glucopyranoside, methyl- α -D-mannopyranoside, Methyl- β 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^T compared with its close relatives are detailed in Table 1.

193 The most abundant cellular fatty acids of strain Marseille-P2082^T were identified as
194 C_{18:1n9} (37.2%), C_{18:0} (34.7%), C_{16:0} (21.3%), C_{18:2n6} (6.3%) and a minor amount of C_{14:0}
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,

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

Genome comparison

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

The dDDH values of strain Marseille-P2082 ranged from 23.2%±4.8% with *S. bovis* to 53.7±5.3 % with *N. succinicivorans* strain ADV (Table 3). The OrthoANI values ranged from 63.35% between strain Marseille-P2082 and *D. pneumosintes* to 93.78% with *N. succinicivorans*. These values are lower than the threshold value (95-96%) recommended to 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.

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

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

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

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

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

Data deposit. The whole-genome shotgun sequence of *N. succinivorans* strain ADV 07/08/06-B-1388^T has been deposited in NCBI under accession number PRJEB34910.

Acknowledgements

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

grateful to Professor Hélène Marchandin for providing the *Negativicoccus succinicivorans* strain ADV 07/08/06-B-1388^T.

Funding

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

Conflict of interest

The authors declare no conflict of interest.

Author’s contributions

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

References

Auch AF, von Jan M, Klenk H-P, Göker M (2010) Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. Stand Genomic Sci 2:117–134. <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>

- Togo AH, Diop A, Dubourg G, et al (2016a) *Butyricimonas phoceensis* sp. nov., a new anaerobic species isolated from the human gut microbiota of a French morbidly obese patient. *New Microbes New Infect* 14:38–48. <https://doi.org/10.1016/j.nmni.2016.07.010>
- Togo AH, Durand G, Khelaifia S, et al (2017) *Fournierella massiliensis*, gen. nov., sp. nov., a new human-associated member of the family *Ruminococcaceae*. *Int J Syst Evol Microbiol* 67:1393–1399. <https://doi.org/10.1099/ijsem.0.001826>
- Togo AH, Khelaifia S, Lagier J-C, et al (2016b) Noncontiguous finished genome sequence and description of *Paenibacillus ihumii* sp. nov. strain AT5. *New Microbes New Infect* 10:142–150. <https://doi.org/10.1016/j.nmni.2016.01.013>
- Togo AH, Khelaifia S, Valero R, et al (2016c) ‘*Negativicoccus massiliensis*’, a new species identified from human stool from an obese patient after bariatric surgery. *New Microbes New Infect* 13:43–44. <https://doi.org/10.1016/j.nmni.2016.05.018>
- Turnbaugh PJ, Ley RE, Mahowald MA, et al (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444:1027–1031. <https://doi.org/10.1038/nature05414>
- Zhang H, DiBaise JK, Zuccolo A, et al (2009) Human gut microbiota in obesity and after gastric bypass. *Proc Natl Acad Sci U S A* 106:2365–2370. <https://doi.org/10.1073/pnas.0812600106>

366 **Table 1:** Differential characteristics of strain Marseille-P2082 (1) compared with the closely related species *N. succinicivorans* 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)

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

Glycerol	+	—	—	+	—	—	—
Inositol	—	—	—	+	—	+	—
Lactose	v	—	—	+	—	+	+
Maltose	+	—	—	+	+	+	—
Mannitol	+	—	—	—	+	+	—
Mannose	+	—	—	+	—	+	+
Melibiose	—	—	—	+	—	+	+
Raffinose	v	—	—	—	w	+	+
Rhamnose	—	—	—	—	—	+	—
Ribose	—	—	—	+	—	+	—
Salicin	v	—	—	+	—	+	+
Sorbitol	—	—	—	+	—	—	—
Sucrose	+	—	—	+	+	+	+
Trehalose	+	—	—	+	—	+	+
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
G+C content (mol %)	51.1	48.3	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^T**

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

375

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

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
Strain Marseille-P2082	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*

	1	2	3	4	5	6	7
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

Figure legends

Figure 1: Reference mass spectrum from strain Marseille-P2082^T. Spectra from 12 individual colonies were compared and a reference spectrum was generated.

Figure 2: Gel view comparing strain Marseille-P2082^T to other closely related species.

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

Figure 3: Maximum Likelihood phylogenetic tree based on 16S rRNA sequence comparison highlighting the position of strain Marseille-P2082^T compared to other closely related species.

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

Figure 4: Graphical circular map of the genome of strain Marseille-P2082^T.

From outside to the centre: Contigs (red/grey), 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), GC content.

Figure 5: Distribution of functional classes of predicted genes according to the COGs of strain Marseille-P2082^T 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 μm .