

1 **Draft genome of *Metabacillus niabensis* strain 4T19^T isolated from cotton-waste**
2 **composts for mushroom cultivation**

3

4 Luis Johnson Kangale^{1,2}, Anthony Levasseur^{2,3}, Didier Raoult^{2,3,4}, Eric Ghigo^{2,5,*} and Pierre-
5 Edouard Fournier^{1,2,*}

6

7 ¹Aix-Marseille Univ, IRD, AP-HM, SSA, VITROME, Marseille, France

8 ²IHU-Méditerranée-Infection, Marseille, France

9 ³Aix-Marseille Univ, IRD, AP-HM, MEPHI, Marseille, France

10 ⁴Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz
11 University, Jeddah, Saudi Arabia

12 ⁵Techno-Jouvence, 19-21 Boulevard Jean Moulin 13385 Marseille cedex 05, France.

13

14 * **Corresponding authors:** Eric Ghigo and Pierre-Edouard Fournier

15

16 Aix-Marseille Université, Institut Hospitalo-Universitaire Méditerranée Infection,

17 19-21 Boulevard Jean Moulin 13385 Marseille cedex 05, France.

18 Phone: +33(0)413732401. Fax: +33(0)413732402

19 **E-mail:** eric.ghigo@gmail.com and pierre-edouard.fournier@univ-amu.fr

20

21 **Keywords:** *Metabacillus niabensis* strain 4T19^T, cotton-waste, genome

22

23 **Abbreviations :**

24 CSUR: Collection de Souches de l'Unité des Rickettsies.

25 DSMZ: German Collection of Microorganisms and Cell Cultures

26 **Abstract**

27 In this manuscript we present the draft genome sequence of *Metabacillus niabensis* strain
28 4T19^T (= CSUR Q2603^T = DSM 17723 = JCM 16399 = KACC 11279), that is a new
29 *Metabacillus* species isolated from cotton-waste composts. The genome sequence from
30 *Metabacillus niabensis* strain 4T19^T was assembled into 462 contigs for a total size of
31 4,987,608 bp with a G+C content of 35.5%.

32

33 **Announcement**

34 *Metabacillus fastidiosus* is considered as type species of genus *Metabacillus*¹. Strain 4T19^T
35 was isolated from cotton-waste composts for mushroom cultivation, and was described as
36 type strain of *Metabacillus niabensis*². Working on planarians microbiota, we identified a
37 bacterial strain, Marseille-P9898 (= CSUR P9898 = DSM 111480). We found that Marseille-
38 P9898 strain had similarity at the level of an 16S rRNA gene sequence with *Metabacillus*
39 *niabensis* strain 4T19^T (98.99%). This 16S rRNA gene similarity value does not discriminate
40 the strain Marseille-P9898 from the strain 4T19^T. However the characterization of the strain
41 Marseille-P9898 by a taxono-genomics approach require the genome from the strain 4T19^T,
42 but to date no genome sequence are available for the strain 4T19^T. Thus, here, we describe the
43 draft genome of *Metabacillus niabensis* strain 4T19^T. We purchased the strain 4T19^T (=DSM
44 17723) from the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
45 (DSMZ) and deposited in the Collection de Souches de l'Unité des Rickettsies (CSUR)
46 culture collection under the number Q2603. Strain 4T19^T was grown at 28°C for 24 hours on
47 Columbia agar supplemented with 5% sheep blood (bioMérieux, Marcy l'Etoile, France) in
48 aerobic atmosphere. To extract bacterial genomic DNA (gDNA) of the strain Q2603, a
49 mechanical treatment was first performed on a single colony, by glass beads acid washed
50 (Sigma Aldrich Chimie, Saint-Quentin-Fallavier, France) using a FastPrep-24TM 5G Grinder

51 (MP Biomedicals, Illkirch, France) at maximum speed (6.5) for 90s. Then after a 30 minutes
52 lysozyme incubation at 37°C. DNA (gDNA) was extracted using an EZ1 BioRobot and the
53 EZ1 DNA tissue kit (Cat No./ID: 953034, Qiagen, Hilden, Germany). The gDNA was
54 sequenced using MiSeq technology (Illumina Inc, San Diego, CA, USA)³ with the Nextera
55 XT DNA sample prep kit (Illumina), and Mate-Pair strategy. Foremost, gDNA was
56 fragmented, then amplified to 12 cycle of PCR, followed by the addition of the tag adapters,
57 and introduced dual-index barcodes. Libraries were purified on AMPure XP beads (Beckman
58 Coulter Inc, Fullerton, CA, USA), then normalized on specific beads following the Nextera
59 XT protocol, and pooled into a single library for sequencing. After 39 hours, automated
60 cluster generation and paired-end sequencing with dual index reads were performed in a
61 single run in 2x250-bp format. The 6,897,812 paired end reads of the Miseq run were checked
62 to according to quality using FastQC 0.11.8⁹ and trimmed using Trimmomatic 0.36.6⁸, with
63 default parameters. Trimmed reads were assembled using the Spades⁴ genome assembler
64 software (Galaxy 3.12.0+galaxy1) set with default parameters. Default parameters were
65 applied here and for all software (for k values, i.e., k-mer values of 127, 99, 77, 55, 33, and
66 21). Using default parameters of SSAKE-based Scaffolding of Pre-Assembled Contigs after
67 Extension (SSPACE 2.0)⁵ and GapFiller (GapFiller 1.10)⁶ allowed us to combine the
68 contigs. Briefly, we have obtained 792 scaffolds after assembly, then using SSAKE we
69 reduced to 520 scaffolds and the manual finishing allowed us to obtain 462 scaffolds
70 (singletons = 473, multi-contig scaffold = 15). Then manual finishing by using sequence
71 similarity searches and syntax block detection between closest species in the *Metabacillus*
72 genus (using type strain of the *Metabacillus halosaccharovorans*, *Metabacillus litoralis* and
73 *Metabacillus fastidiosus*). For contigs profiling, a Blast e-value threshold of 1e-10 was set to
74 perform similarity searches and contig length threshold was set to 500 pb and a contig
75 coverage threshold of 15% was applied. Finally, *Metabacillus niabensis* strain 4T19^T was

76 assembled into 462 contigs (N₅₀, 19,230 bp; L₅₀, 64; coverage, 5x) for a total size of 4,987,608
77 bp, with a G+C content of 35.5%. Genomic annotation was obtained by NCBI prokaryotic
78 genome annotation pipeline (pGAP)⁷. A total of 4,969 genes were identified, along with 11
79 rRNAs, 61 tRNAs, 6 ncRNA and 1tmRNA. A genomic comparison considering the size,
80 GC% and contig was performed (Table 1), evidencing that *Metabacillus niabensis* is different
81 of the other species of *Metabacillus*. To note, *Metabacillus niabensis* is a motile Gram
82 negative and aerobic bacteria, growing from 15 to 40°C. *Metabacillus niabensis* is catalase
83 and betagalactosidase positive, and oxydase negative, an the principal fatty acid find in this
84 bacteria is the 12-methyl-tetradecanoic acid¹⁰.

85

86 **Data availability**

87 The complete 16S rRNA gene sequence of *Metabacillus niabensis* have been deposited at
88 GenBank under the accession numbers AY998119. The genome sequence *Metabacillus*
89 *niabensis* have been deposited at GenBank under the accession number CADEPK010000000
90 (<https://www.ncbi.nlm.nih.gov/Traces/wgs/CADEPK01?display=contigs&page=1>). Illumina
91 MiSeq paired end sequencing raw data have been deposited under accession number
92 ERR4020000 (<https://www.ncbi.nlm.nih.gov/sra/ERR4020000>). The *Metabacillus niabensis*
93 strain 4T19^T is available at the Collection under the reference CSUR Q2603 = DSM 17723 =
94 JCM 16399 = KACC 11279.

95

96 **Author contribution statement**

97 LJK: conceived the experiments, realised the experiments, analysed the data, prepared figures,
98 wrote the manuscript. DR, EG, AL and PEF: designed the experiments, conceived the
99 experiments, analysed the data, wrote the manuscript.

100

101 **Acknowledgements**

102 The study was funded by the Méditerranée-Infection foundation, the National Research
103 Agency under the program “Investissements d’avenir”, reference ANR-10-IAHU-03 and by
104 Région Provence Alpes Côte d’Azur and European funding FEDER IHUBIOTK. LJK is
105 fellow of Méditerranée-Infection foundation. We also thank Aurelia Caputo for submitting the
106 genomic sequences to GenBank.

107

108 **Conflict of interest**

109 The authors have no conflicts of interest to declare. The funding sources had no role in the
110 study design, data collection and analysis, decision to publish, or manuscript preparation.

111 **References**

112

- 113 1. Patel S, Gupta RS. A phylogenomic and comparative genomic framework for resolving
114 the polyphyly of the genus *Bacillus*: Proposal for six new genera of *Bacillus* species,
115 *Peribacillus* gen. nov., *Cytobacillus* gen. nov., *Mesobacillus* gen. nov., *Neobacillus* gen.
116 nov., *Metabacillus* gen. nov. and *Alkalihalobacillus* gen. nov. *Int J Syst Evol Microbiol*,
117 70, 406-438 (2020).
- 118 2. Kwon, S.W. Lee, S. Y., Kim, B. Y., Weon, H. Y., Kim, J. B., Go, S. J., and Lee, G. B.
119 2007. *Bacillus niabensis* sp. nov., isolated from cotton-waste composts for mushroom
120 cultivation. *Int J Syst Evol Microbiol*, 57, 1909–1913.
- 121 3. Ravi, R. K., Walton, K. and Khosroheidari, M. 2018. MiSeq: A Next Generation
122 Sequencing Platform for Genomic Analysis. *Methods Mol. Biol.* 1706, 223–232.
- 123 4. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM,
124 Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G,
125 Alekseyev MA and Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and
126 its applications to single-cell sequencing. *J Comput Biol* 19, 455–477.
- 127 5. Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D. and Pirovano, W. 2011. Scaffolding
128 pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579.
- 129 6. Boetzer, M. and Pirovano, W. 2012. Toward almost closed genomes with GapFiller.
130 *Genome Biol* 13, R56.
- 131 7. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L,
132 Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome
133 annotation pipeline. *Nucleic Acids Res* 44, 6614–6624.
- 134 8. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina
135 sequence data. *Bioinformatics* 30, 2114 –2120.

- 136 9. Wingett, S. W.; Andrews, S. FastQ Screen: A Tool for Multi-Genome Mapping and
 137 Quality Control. *F1000Res* **2018**, 7. <https://doi.org/10.12688/f1000research.15931.2>.
- 138 10 Kangale, L. J.; Raoult, D. ; Ghigo, E. ; Fournier, P. E. *Metabacillus schmidtea* sp. nov.,
 139 Cultivated from Planarian *Schmidtea mediterranea* Microbiota. *Microbiol Res* 12(2) 299-316,
 140 <https://doi.org/10.3390/microbiolres12020021>.

141

142

143 Table 1: genomic comparison of *Metabacillus niabensis* with others genus *Bacillus*,
 144 *Mesobacillus*, *Metabacillus* and *Neobacillus*.

145 Genomic comparison considering the size, GC% and contigs. Refseq: numbers reference
 146 sequence in NCBI.

Name	Size (bp)	GC%	Contigs	Refseq
<i>Bacillus flexus</i>	3,906,163	37.6	259	BCVD01000001.1
<i>Bacillus acidicola</i>	5,137,992	39.4	10	LWJG01000001.1
<i>Mesobacillus foraminis</i>	5,730,823	43.0	35	SLVV01000001.1
<i>Metabacillus halosaccharovorans</i>	5,399,327	36.1	8	MTIR01000001.1
<i>Metabacillus litoralis</i>	5,230,624	35.9	1	NZ_CP033043.1
<i>Metabacillus niabensis</i>	4,987,608	35.5	462	NZ_CADEPK010000000
<i>Neobacillus niacini</i>	2,201,253	38.3	143	JRYQ01000001.1
<i>Metabacillus fastidiosus</i>	4,410,645	35.1	239	NZ_BCVG00000000.1
<i>Mesobacillus subterraneus</i>	4,571,170	43.9	42	RSFW01000001.1
<i>Bacillus cohnii</i>	4,899,141	36.1	1	NZ_CP018866.1

147