

1 **Title: *Rhizobium desertarenae* sp. nov., isolated from the Saline Desert Soil from the Rann
2 of Kachchh, India.**

3 **Running Title: *Rhizobium desertarenae* sp. nov.,**

4 **Authors and affiliations:**

5 Mitesh Khairnar¹, Ashwini Hagar¹, Avinash Narayan², Kunal Jain², Datta Madamwar², Yogesh
6 Shouche¹, Praveen Rahi^{1*}

7 ¹National Centre for Microbial Resource, National Centre for Cell Science, Pune, Maharashtra,
8 India

9 ² Post-Graduate Department of Biosciences, UGC-Centre of Advanced Study, Satellite Campus,
10 Vadtal Road, Sardar Patel University, Bakrol, Anand, Gujarat, India

11 **Corresponding author details:**

12 Praveen Rahi

13 National Centre for Microbial Resource, National Centre for Cell Science, Pune, Maharashtra
14 411007, India

15 *Email ID: praveen_rahi22@yahoo.co.in; praveen@nccs.res.in

16 **Keywords:** Saline Desert; Zobell marine agar; Polyphasic taxonomy; MALDI-TOF MS
17 Biotyping

18

19 The GenBank accession numbers for the 16S rRNA gene sequences of strain ADMK78^T is
20 MK942856. The genome sequence has been deposited in GenBank under the accession number
21 CP058350-CP058352. The type strain is available with different culture collections under the
22 accession numbers MCC 3400^T; KACC 21383^T; and JCM 33657^T.

23

24

25 **Title: *Rhizobium desertarenae* sp. nov., isolated from the Saline Desert Soil from the Rann
26 of Kachchh, India.**

27 **Running Title: *Rhizobium desertarenae* sp. nov.,**

28 **Authors and affiliations:**

29 Mitesh Khairnar¹, Ashwini Hagar¹, Avinash Narayan², Kunal Jain², Datta Madamwar², Yogesh
30 Shouche¹, Praveen Rahi^{1*}

31 ¹National Centre for Microbial Resource, National Centre for Cell Science, Pune, Maharashtra,
32 India

33 ² Post-Graduate Department of Biosciences, UGC-Centre of Advanced Study, Satellite Campus,
34 Vadtal Road, Sardar Patel University, Bakrol, 388 315, Anand, Gujarat, India

35 **Abstract:**

36 A novel bacterial strain designated ADMK78^T was isolated from the saline desert soil. The cells
37 were rod-shaped, Gram-negative, and non-motile. The strain ADMK78^T grows best at 28°C and
38 pH 7.0 and can tolerate up to 2% (w/v) NaCl. Based on 16S rRNA gene phylogeny, the strain
39 ADMK78^T belongs to the genus *Rhizobium*, with the highest similarity to *Rhizobium*
40 *wuzhouense* W44^T (98.7%) and *Rhizobium ipomoeae* shin9-1^T (97.9%). Core-genes based
41 phylogenetic analysis revealed that the strain ADMK78^T forms a distinct branch in between
42 *Rhizobium ipomoeae* shin9-1^T and *Rhizobium selenitireducens* BAA-1503^T. The average
43 nucleotide identity of ADMK78^T was less than 82%, to members of the family *Rhizobiaceae*.
44 The genomic DNA G+C content of strain ADMK78^T is 58.6 mol%. The major fatty acids of

45 strain ADMK78^T were C_{18:0} and C_{18:1} ω7c. The strain ADMK78^T showed differences in
46 physiological, phenotypic, and protein profiles estimated by MALDI-TOF MS to its closest
47 relatives. Based on the phenotypic, chemotaxonomic properties, and phylogenetic analyses, the
48 strain ADMK78^T could be distinguished from the recognized species of the genus *Rhizobium*. It
49 is suggested to represent a novel species of this genus, for which the name *Rhizobium*
50 *desertarenae* sp. nov. is proposed. The type strain is ADMK78^T (=MCC 3400^T; KACC 21383^T;
51 JCM 33657^T).

52

53 **Introduction:**

54 The genus *Rhizobium* is a large group of bacteria, and most species are known for their
55 symbiotic fixation of nitrogen within the root nodules of leguminous plants. The family
56 *Rhizobiaceae* has now been classified into 16 genera [1,2]. *Rhizobium* is one of the main genera
57 in the family *Rhizobiaceae*, and it was first proposed in 1889 [3,4]. Presently, the genus
58 *Rhizobium* comprises 90 recognized species (<https://lpsn.dsmz.de/genus/rhizobium>). Members of
59 the genus *Rhizobium* are characterized as Gram-stain-negative, non-spore-forming, rod-shaped,
60 aerobic, chemo-organotrophic, and possess C_{18:1} and C_{18:1} ω7c as the predominant fatty acid and
61 have a DNA G+C content of between 55 and 66 mol% [5,6]. Non-symbiotic and free-living
62 members of *Rhizobium* have been found in various soils, including the rhizosphere [7,8],
63 bioreactor [9], and beach sand [10].

64 **Isolation and Ecology:**

65 During the investigations on bacterial diversity of the saline desert soil collected (23.7337° N,
66 69.8597° E) from the Rann of Kachchh, India, a bacterial strain ADMK78^T, was isolated on
67 Zobell Marine Agar. Rann of Kachchh is reputed to be the largest salt desert in the world and is a

68 transitional area between marine and terrestrial ecosystems [11]. The region experiences
69 diagonal fluctuations of average temperatures with a high temperature of 50 °C during summers
70 and drops below freezing during winters. Due to the hot and hypersaline environment, there is a
71 vast possibility of identifying novel microbes with high economic and industrial potential. The
72 newly isolated strain ADMK78^T was maintained on Zobell Marine Agar at 37 °C, and preserved
73 at -80 °C as a suspension in 20% (v/v) glycerol and by lyophilization with 20% (w/v) skimmed
74 milk. The pure culture was processed for MALDI-TOF MS-based identification, and a
75 comparison of the MALDI-TOF MS spectrum of ADMK78^T with the Biotype 3.0 database
76 resulted in no reliable identification.

77 **16S RNA phylogeny:**

78 High-quality genomic DNA was extracted from the strain following the JGI protocol version 3
79 for bacterial genomic DNA isolation using CTAB [12]. The 16S rRNA gene sequence was
80 amplified using universal primers (27f: 5'-AGAGTTGATCCTGGCTCAG-3' and 1492r: 5'-
81 TACGGCTACCTTGTACGACTT-3') according to the methods described by Gulati *et al.* [13],
82 and the amplified product was directly sequenced using the ABI PRISM Big Dye Terminator
83 v3.1 Cycle Sequencing kit on a 3730xl Genetic Analyzer (Applied BioSystems, Thermo
84 Scientific, USA). The similarity search for the 16S rRNA gene sequence of strain ADMK78^T
85 was performed against the type strains of prokaryotic species in the EzBioCloud's valid species
86 database [14]. The strain ADMK78^T showed the highest similarity to *Rhizobium wuzhouense*
87 W44^T (98.7%) and followed by *Rhizobium ipomoeae* shin9-1^T (97.9%). The 16S rRNA gene
88 sequence of strain ADMK78^T was used as queries to closely related gene sequences using the
89 NCBI BLASTn tool [15] and the non-redundant nucleotide database. Higher than 99% sequence
90 similarity was recorded for two sequences. The first one was of Alpha proteobacterium

91 (EU770254.1) associated with *Microcystis aeruginosa* culture, and the second one was of
92 *Ciceribacter* sp. strain AIY3W (MH463946.2) isolated from low salinity lakes on Tibetan
93 Plateau. Multiple alignments of sequences of strain ADMK78^T and its nearest neighbours
94 retrieved from EzBioCloud's server and NCBI GenBank, and phylogenetic analyses were
95 performed using MEGA software (version 7.0) [16]. Bootstrap values were determined based on
96 1000 replications. The 16S rRNA gene sequence of *Bradyrhizobium japonicum* USDA 6^T was
97 used as an outgroup. The strain ADMK78^T formed a separate branch along with the Alpha
98 proteobacterium and *Ciceribacter* sp. AIY3W (Fig. 1). The valid species of the genus
99 *Ciceribacter* were placed in a distinct group, which was far from the group formed by
100 ADMK78^T, Alpha proteobacterium, and *Ciceribacter* sp. AIY3W. The overall topologies were
101 similar for the phylogenetic trees obtained with the ML, MP and NJ methods. *Rhizobium*
102 *ipomoeae* shin9-1^T appears to be the closest phylogenetic neighbour to the group formed by
103 ADMK78^T, Alpha proteobacterium, and *Ciceribacter* sp. AIY3W in all three trees constructed
104 by different methods.

105 **Genome Features:**

106 Genome sequencing was performed using a hybrid approach of two platforms, first on an
107 Illumina MiSeq platform with 2 x 250 bp v2 chemistry, followed by sequencing with Oxford
108 Nanopore Technology (ONT) on a minION platform. The Nanopore reads were assembled using
109 Canu v. 2.0 [17] with default settings. The overlaps between the ends of circular contigs were
110 identified using NUCmer v. 3.1 [18] and removed using a custom Perl script. Two rounds of
111 polishing was performed using the paired-end Illumina reads. In each round the Illumina reads
112 were mapped to the genome assembly using bowtie2 v. 2.3.4.1 [19] with default parameters,
113 followed by polishing using Pilon v. 1.23 [20] with default settings. The genome sequence

114 quality of strain ADMK78^T was as per the genome standards proposed by Chun et al. [21], and
115 the detailed genome features are provided in Table 1. The genome of ADMK78^T had a size of
116 4,342,374 bp, which is smaller than the genome size of symbiotic members of *Rhizobium* and in
117 the range of the sizes of the non-symbiotic strains of *Rhizobium* (Table 1). It consisted of a
118 circular chromosome of 3,590,542 bp and two circular plasmids of 708,533 and 43,299 bp. The
119 overall genome sequencing coverage for the strain ADMK78^T was 147.5x, with an N50 value of
120 3,590,542 bp. Whole-genome sequences were annotated using the RAST [22] web server
121 (<http://rast.nmpdr.org/rast.cgi>). The genome of strain ADMK78^T contains 4377 protein-coding
122 sequences (CDS), of which 64 genes assigned to the stress response functions, like heat and cold
123 shock, hyperosmotic stress, and protection from reactive oxygen. We could not find any
124 nitrogen-fixation and nodulation genes in the genome of strain ADMK78^T.

125 The bacterial core gene-based phylogenetic analysis was carried out using the UBCG pipeline
126 [23] from the concatenated sequences of 92 core genes extracted by UBCG, and a maximum-
127 likelihood phylogenetic tree was inferred using RAxML version 8.2.8 [24] with the
128 GTRGAMMA model and 100 bootstrap replications. The Average Nucleotide Identity (ANI)
129 was determined between strain ADMK78^T and closely related strains of the *Rhizobiaceae* family
130 using FastANI [25], and a heatmap representation of the calculated ANI values was constructed
131 using DisplayR (<https://www.displayr.com/>). The genome-based phylogenetic analysis placed
132 strain ADMK78^T as an independent branch, with *Rhizobium ipomoeae* shin9-1^T as the closest
133 neighbour and followed by *Rhizobium rosettiformans* W3^T, *Rhizobium wuzhouense* W44^T and
134 *Rhizobium glycinendophyticum* CL12^T (Fig. 2). The highest ANI values of ADMK78^T was for
135 *Rhizobium rosettiformans* W3 (82%), followed by *Rhizobium wuzhouense* W44^T (81.3%) and
136 *Rhizobium ipomoeae* shin9-1^T (81%) (Fig. S1; Table 1). The dDDH relatedness values of

137 ADMK78^T with the reference strains were below 22% (Table 1), suggesting the strain
138 ADMK78^T is a novel species [19]. The genomic DNA G+C content of strain ADMK78^T was
139 58.6 mol%, which is well within the range (i.e., 55-66 mol %) of the genus *Rhizobium* [4].

140 **Physiology and Chemotaxonomy:**

141 The type strains of *Rhizobium ipomoeae* shin9-1^T and *Ciceribacter lividus* MSSRFBL1^T were
142 obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) and the
143 BCCM/LMG Bacteria Collection, Belgium (LMG), respectively. Both type strains were used as
144 reference strains and evaluated together under identical experimental conditions to those for
145 strain ADMK78^T.

146 For analysis of chemotaxonomic features, the strain was grown on Zobell Marine Agar at 28 °C,
147 and cell biomass was harvested after 24 h. Preparation and analysis of fatty acid methyl esters
148 were performed as described by Sasser [25] using the Microbial Identification System (MIDI)
149 and the Microbial Identification software package (Sherlock version 6.1; MIDI database,
150 TSBA6). The primary fatty acids detected in strain ADMK78^T were C_{18:1} ω 7c and C_{18:0} (Table
151 S1). Small proportions of C_{16:0}, C_{18:1} ω 7c 11-methyl, C_{18:0} 3OH, C_{20:1} ω 7c, and summed feature 2 (C_{12:0}
152 aldehyde) were also detected for strain ADMK78^T. The strain ADMK78^T exhibit higher proportions
153 of C_{18:0}, similar to *Ciceribacter lividus* MSSRFBL1^T (Table S1), which was missing in
154 *Rhizobium wuzhouense* W44^T [26] and present in relatively lower amounts in *Rhizobium*
155 *ipomoeae* shin9-1^T. A relatively higher proportion of C_{18:1} ω 7c 11-methyl and C_{20:1} ω 7c in the fatty
156 acids profile of strain ADMK78^T, differentiate it from the reference strains (Table S1).

157 Whole-cell proteins were extracted using ethanol/formic acid after 24 h growth on TSA, to
158 generate the Mean Spectral Profile (MSP). The proteins ranging from 2-20 KDa were analyzed
159 by matrix-assisted laser desorption/ionization time-of-flight mass spectrometer (MALDI-TOF

160 MS) autoflex speed (Bruker Daltonik GmbH, Germany) [27]. A total of 27 replicate spectra were
161 used to generate a new MSP of strain [28], which was compared with MSPs of the reference
162 strains *Rhizobium ipomoeae* shin9-1^T and *Ciceribacter lividus* MSSRFBL1^T generated during
163 this study following same procedure. The mean spectra profile of strain ADMK78^T has 33
164 unique peaks in comparison to the closely related taxa out of a total of 70 peaks (Table S2),
165 which attributed to the discrimination of ADMK78^T from the closely related species.
166 Morphological, physiological, and biochemical tests for strain ADMK78^T were performed on
167 Zobell Marine Agar plates incubated under aerobic conditions. The colonies of strain
168 ADMK78^T, were circular and translucent on Zobell Marine Agar (Fig. S3). Gram-staining (K001
169 and K004, Himedia, India) was used following the manufacturer's instructions. Hanging drop
170 technique was used to check the motility. Scanning electron microscopy was performed to
171 observe cell morphology as described in Rahi *et al.* [29]. These analyses revealed that strain
172 ADMK78^T is a rod-shaped bacterium with cell size ranging from 0.3-0.5×1.5-2 μ m, Gram stain
173 negative, (Fig. S2) and non-motile. Oxidase disc (DD018, Himedia, India) was used for testing
174 oxidase activity, and catalase activity was determined by bubble formation in a 3% (v/v) H₂O₂
175 solution. The strain was positive for both oxidase and catalase. Growth at different temperatures
176 (4, 10, 15, 20, 28, 37, 45 and 55 °C), NaCl concentrations [0-2% (w/v) at 0.5% intervals] and pH
177 values (4.0-11.0 at 1.0 pH unit intervals) was examined after incubation in Zobell Marine broth
178 for 7 days in automated microbial growth analyzer (Bioscreen C, OY Growth Curves, Finland).
179 The initial pH of the inoculation broth was adjusted using 1 M HCl and 1 M NaOH. The strain
180 ADMK78^T grows at 10-45 °C (optimum 28°C), pH ranging from 4-10 (optimum 7.0), and NaCl
181 concentration tolerance up to 2%.

182 Biochemical characteristics, enzyme activities, and oxidation/or reduction of carbon sources
183 were performed using the API 20E and API ZYM systems (07584D and 25200, bioMérieux,
184 France) and Biolog GN III system (OmniLog, Biolog, USA) following manufacturers'
185 instructions. The Biolog test showed that out of the substrates present in the GENIII BIOLOG
186 microplate, ADMK78^T showed activity for 71 substrates, of which a weak reaction was recorded
187 for four substrates (Table S3). The strain ADMK78^T can be distinguished from its closest
188 phylogenetic neighbours based on features listed in Table 2.
189 The genotypic and phenotypic data generated for the strain ADMK78^T revealed that the strain
190 represents a novel species in the genus *Rhizobium*, for which the name *Rhizobium desertarenae*
191 sp. nov. is proposed.

192 **Description of *Rhizobium desertarenae* sp. nov.**

193 ***Rhizobium desertarenae*** (de.sert.a.re'nae. L. neut. n. *desertum* desert; L. fem. n. *arena* sand;
194 N.L. gen. n. *desertarenae* of desert sand).

195 Cells are Gram-negative, straight rods with round ends (0.3-0.5×1.5-2 μ m), and non-motile.
196 Colonies grown on Zobell Marine Agar are 1-3 mm in diameter, circular, raised with an entire
197 margin, and translucent opacity. The optimal temperature for growth is 28 °C, and the optimal
198 pH is 7.0. Growth occurs in the absence of NaCl with up to 2% tolerance in Zobell Marine broth.
199 It is oxidase and catalase positive. The strain showed positive results in Biolog GN III analyses
200 for utilization of D-maltose, D-trehalose, D-cellobiose, D-gentiobiose, sucrose, D-turanose, α -
201 D-lactose, D-melibiose, β -methyl-d-glucoside, D-salicin, N-acetyl-D-glucosamine, N- acetyl- β -
202 D-mannosamine, N-acetyl-D-galactosamine, α -D-glucose, D-mannose, D-fructose, D-galactose,
203 D-fucose, L-fucose, L-rhamnose, inosine, D-sorbitol, D-mannitol, D-arabitol, myo-inositol,
204 glycerol, D-glucose-6-phosphate, D-fructose-6-phosphate, D-aspartic acid, Glycyl-L-proline,

205 glycyl-L-proline, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, L-
206 pyroglutamic acid, L-serine, pectin, D-galacturonic acid, D-gluconic acid, D-glucuronic acid,
207 glucuronamide, mucic acid, D-saccharic acid, p-hydroxy-phenylacetic acid, D-lactic acid methyl
208 ester, L-lactic acid, citric acid, α -keto-glutaric acid, D,L-malic acid, bromo-succinic acid,
209 Tween 40, γ -amino-butyric acid, α -hydroxy-butyric acid, α -hydroxy-D,L butyric acid, α -keto-
210 butyric acid, acetoacetic acid, propionic acid, acetic acid, formic acid, sodium lactate,
211 tetrazolium violet and blue, nalidixic acid, lithium chloride (Table S3). Positive results in API
212 ZYM strips for leucine arylamidase, trypsin, naphthol-AS-BI-phosphohydrolase, α -glucosidase,
213 N-acetyl- β -glucosaminidase activities (Table S4). C_{18:0} and C_{18:1 ω 7c} are the predominant cellular
214 fatty acids. The DNA G+C content of the type strain is 58.6 mol%.

215 The type strain ADMK78^T (=MCC 3400^T; KACC 21383^T; JCM 33657^T) was isolated from
216 saline desert sand collected from the Kutch District of Gujarat, India. The GenBank sequence
217 accession number of the genome sequence is CP058350- CP058352, and 16S rRNA gene
218 sequence of strain ADMK78^T is MK942856.

219 **Funding Information:**

220 Financial support by the Department of Biotechnology (BT/Coord. II/01/03/2016) and
221 (BT/PR8218/BCE/8/1044/2013).

222 **Acknowledgments:**

223 The authors thank Prof. Aharon Oren, Department of Plant and Environmental Sciences, The
224 Hebrew University of Jerusalem, Israel and Prof. Bernhard Schink, Department of Biology,
225 University of Konstanz, Germany, for their expert suggestions concerning the correct species
226 name, species epithet and Latin etymology.

227 **Conflicts of Interest:**

228 The authors declare that there are no conflicts of interest.

229 **Ethical Statement:**

230 The experiments reported in this manuscript did not involve human participants and/or animals.

231 **ABBREVIATIONS:**

232 CDS= coding sequence, CTAB= Cetyl trimethylammonium bromide; MALDI-TOF MS= Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometer; MSP= Mean Spectral Profile; UBCG= Up-to-date bacterial core genes.

235

236 **REFERENCES:**

- 237 1. **Lindström K, Amsalu AA, Mousavi SA.** Evolution and taxonomy of nitrogen-fixing
238 organisms with emphasis on rhizobia. In: de Bruijn FJ (editor). Biological Nitrogen
239 Fixation. Hoboken, NJ: John Wiley & Sons; 2015. pp. 21–38.
- 240 2. **Lindström K, Mousavi SA.** Rhizobium and other N-fixing symbioses. In: Encyclopedia
241 of Life Science (ELS). Chichester: John Wiley & Sons; 2010.
- 242 3. **Frank, B.** U" ber die Pilzsymbiose der Leguminosen. Ber Dtsch Bot Ges 1889;7, 332–
243 346 (in German).
- 244 4. **Kuykendall, L. D.** Family I. Rhizobiaceae. In Bergey's Manual of Systematic
245 Bacteriology, 2005; 2nd edn, vol. 2, Part C, pp. 324–361. Edited by D. J. Brenner, N. R.
246 Krieg, J. T. Staley & G. M. Garrity.
- 247 5. **Tighe, S. W., de Lajudie, P., Dipietro, K., Lindstro" m, K., Nick, G. & Jarvis, B. D.**
248 Analysis of cellular fatty acids and phenotypic relationships of Agrobacterium,
249 radyrhizobium, Mesorhizobium, Rhizobium and Sinorhizobium species using the
250 Sherlock Microbial Identification System. Int J Syst Evol Microbiol 2000;50, 787–801.

- 251 6. **Young, J. M., Kuykendall, L. D., Martínez-Romero, E., Kerr, A. & Sawada, H.** A
252 revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the
253 inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium* undicola de
254 Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*,
255 *R. undicola* and *R. vitis*. *Int J Syst Evol Microbiol* 2001;51, 89–103.
- 256 7. **Sheu SY, Chen ZH, Young CC, Chen WM.** *Rhizobium ipomoeae* sp. nov., isolated
257 from a water convolvulus field. *Int J Syst Evol Microbiol* 2016;66:1633–1640.
- 258 8. **Panday D, Schumann P, Das SK.** *Rhizobium pusense* sp. nov., isolated from the
259 rhizosphere of chickpea (*Cicer arietinum* L.). *Int J Syst Evol Microbiol* 2011;61:2632–
260 2639.
- 261 9. **Quan ZX, Bae HS, Baek JH, Chen WF, Im WT, et al.** *Rhizobium daejeonense* sp. nov.
262 isolated from a cyanide treatment bioreactor. *Int J Syst Evol Microbiol* 2005;55:2543–
263 2549.
- 264 10. **Ramana C V., Parag B, Girija KR, Raghu Ram B, Venkata Ramana V, et al.**
265 *Rhizobium subbaraonis* sp. nov., an endolithic bacterium isolated from beach sand. *Int J*
266 *Syst Evol Microbiol* 2013;63:581–585.
- 267 11. **Pandit AS, Joshi MN, Bhargava P, Shaikh I, Ayachit GN, et al.** A snapshot of
268 microbial communities from the Kutch: one of the largest salt deserts in the World.
269 *Extremophiles* 2015;19:973–987.
- 270 12. **Drive M, Creek W.** Bacterial DNA isolation CTAB protocol bacterial genomic DNA
271 isolation using CTAB. *Doe Jt Genome Inst* 2012;4.

- 272 13. **Gulati A, Rahi P, Vyas P.** Characterization of phosphate-solubilizing fluorescent
273 pseudomonads from the rhizosphere of seabuckthorn growing in the cold deserts of
274 Himalayas. *Curr Microbiol* 2008;56:73–79.
- 275 14. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y et al.** Introducing EzBioCloud: a
276 taxonomically United database of 16S rRNA gene sequences and whole-genome
277 assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
- 278 15. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** Basic local alignment search
279 tool. *J Mol Biol* 1990;215:403–410.
- 280 16. **Kumar S, Stecher G, Tamura K.** MEGA7: Molecular Evolutionary Genetics Analysis
281 Version 7.0 for Bigger Datasets. *Mol Biol Evol* 2016;33:1870–4.
- 282 17. **Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, et al.** Canu: scalable and
283 accurate long-read assembly via adaptive k-mer weighting and repeat separation.
284 *Genome Res* 2017;27:722–736.
- 285 18. **Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, et al.** Versatile and open
286 software for comparing large genomes. *Genome Biol*;5. Epub ahead of print 2004. DOI:
287 10.1186/gb-2004-5-2-r12.
- 288 19. **Langmead B, Salzberg SL.** Fast gapped-read alignment with Bowtie 2. *Nat Methods*
289 2012;9:357–359.
- 290 20. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, et al.** Pilon: An integrated tool for
291 comprehensive microbial variant detection and genome assembly improvement. *PLoS
292 One*;9. Epub ahead of print 2014. DOI: 10.1371/journal.pone.0112963.
- 293 21. **Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, et al.** Proposed minimal
294 standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol*

- 295 *Microbiol* 2018;68:461–466.
- 296 22. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T et al.** The RAST server: rapid
297 annotations using subsystems technology. *BMC Genomics* 2008;9:75–15.
- 298 23. **S-I N, Kim YO, Yoon S-H, Ha S, Baek I et al.** UBCG: up-to-date bacterial core gene
299 set and pipeline for phylogenomic tree reconstruction. *J Microbiol* 2018;56:280–285.
- 300 24. **Stamatakis A.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of
301 large phylogenies. *Bioinformatics* 2014;30:1312–1313.
- 302 25. **Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S.** High
303 throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries.
304 *Nat Commun* 2018;9:5114.
- 305 26. **Sasser M.** Technical note # 101 identification of bacteria by gas chromatography of
306 cellular fatty acids. *Stat* 2001:1–6.
- 307 27. **Tao Yuan, Lihui Liu, Shufen Huang, Ali Hussein Taher, Zhiyuan Tan et al.**
308 *Rhizobium wuzhouense* sp. nov., isolated from roots of *Oryza officinalis*. *Int J Syst Evol
309 Microbiol* 2018;68:2918–2923.
- 310 28. **Kurli R, Chaudhari D, Pansare AN, Khairnar M, Shouche YS, et al.** Cultivable
311 microbial diversity associated with cellular phones. *Front Microbiol*;9. Epub ahead of
312 print 7 June 2018. DOI: 10.3389/fmicb.2018.01229.
- 313 29. **Rahi P, Prakash O, Shouche YS.** Matrix-Assisted laser desorption/ionization time-of-
314 flight mass-spectrometry (MALDI-TOF MS) based microbial identifications: challenges
315 and scopes for microbial ecologists. *Front Microbiol* 2016;7:1359.

316 30. **Rahi P, Kurli R, Khairnar M, Jagtap S, Pansare AN et al.** Description of
317 Lysinibacillus telephonicus sp. nov., isolated from the screen of a cellular phone. *Int J*
318 *Syst Evol Microbiol* 2017;67:2289–2295.

319 **Figure legends:**

320 **Fig. 1:** Phylogenetic tree was inferred by using the Maximum Likelihood method and Tamura-
321 Nei model. The percentage of trees in which the associated taxa clustered together is shown next
322 to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying
323 Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the
324 Tamura-Nei model, and then selecting the topology with superior log likelihood value. Empty
325 circles indicate branches of the tree that were also recovered using the neighbour-joining method,
326 circles with dots indicate recovery with the maximum-parsimony method, and black filled circles
327 indicate that all three methods recovered the corresponding nodes. *Bradyrhizobium japonicum*
328 USDA 6^T (D11345) was used as an outgroup. Bar, 0.1 substitutions per nucleotide position.

329
330

331 **Fig. 2:** Phylogenetic tree inferred by the UBCG phylogenomics pipeline using the concatenated
332 alignment of 92 core genes, of strain ADMK78^T and its closely related taxa. Percentage of
333 bootstrap values are given at branching points. Bar, 0.05 substitution per position.

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

Table 1: Standard genome features of strain ADMK78^T and related type strains of the family *Rhizobiaceae*

Strain (GenBank accession numbers)	Genome size (Mbp)	Total contigs (nos.)	DNA G+C content (mol%)	N50 (kb)	Genome coverage	Genome ANI with ADMK78 ^T (%)	dDDH relatedness with ADMK78 ^T (%)	Difference in G+C content (%)
ADMK78 ^T (CP058350-CP058352)	4.31	03	58.6	3,590	147.5x	100.0	100	0
<i>Rhizobium wuzhouense</i> W44 (NZ_QJRY01000001)	4.93	32	61.6	493	277.4x	81.3	22.0	3.0
<i>Rhizobium rosettiformans</i> W3(NZ_STGU00000000)	4.98	86	61.7	288	200.0x	82.0	22.0	3.03
<i>Rhizobium glycinendophyticum</i> CL12(NZ_VFYP00000000)	4.84	16	61.1	2,995	206.0x	81.2	21.8	2.45
<i>Rhizobium ipomoeae</i> shin9-1(NZ_STGV00000000)	4.39	23	60.0	535	200.0x	81.0	21.6	1.78
<i>Rhizobium selenitireducens</i> ATCC BAA-1503 (JAEG00000000)	4.98	32	63.5	-	Unknown	79.9	20.7	4.88
<i>Rhizobium subbaraonis</i> JC85(NZ_OBQD00000000)	6.58	91	63.1	255	153x	78.8	20.7	4.44
<i>Ciceribacter lividus</i> DSM 25528 (NZ_QPIX00000000)	4.52	37	63.2	302	226.0x	79.2	20.2	4.58
<i>Rhizobium endolithicum</i> JC140 (NZ_CABFWF00000000)	4.18	62	62.7	206	64x	79.0	20.2	4.11
<i>Rhizobium sophoriradicis</i> CCBAU 03470(RQIH00000000)	6.67	-	61.3	-	170x	78.70	20.1	2.65
<i>Rhizobium leguminosarum</i> USDA 2370 (QBLB00000000)	7.66	108	60.8	-	50.0x	78.50	19.8	1.95

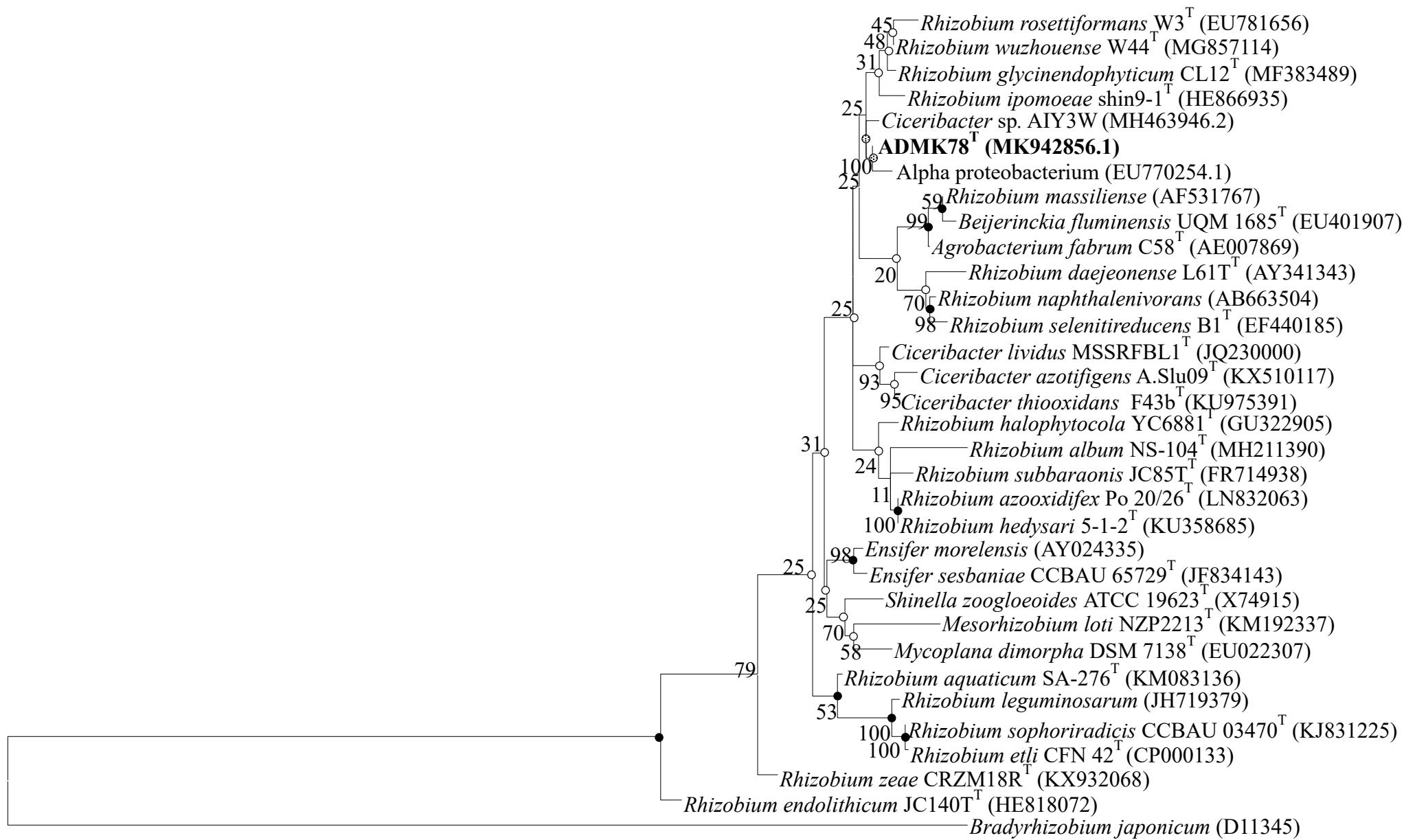
357 **Table 2.** Differentiating characteristics of strain ADMK78^T in comparison to its closest
358 phylogenetic neighbours.
359 strain 1, *Rhizobium* sp. ADMK78^T; 2, *Rhizobium wuzhouense* W44^T; 3, *Rhizobium ipomoeae* sp.
360 shin9-1^T; 4, *Ciceribacter lividus* MSSRFBL1^T

Characteristics	1	2*	3	4
Isolation source	Saline soil (Desert)	Roots of <i>Oryza</i> <i>officinalis</i>	Field	Rhizosphere soil of chickpea
Colony colour	Cream	Cream	Cream	Bluish black
pH range for growth	4.0–11(7)	5–8	7.0–9.0(7)	6.0–8.5(7)
Temperature range for growth (°C) (optimum)	10–45 (28)	15–40	10–45 (30)	10–45 (28)
NaCl range for growth (%) (optimum)	0–2.0 (1.5)	0–2.0	0–3.0 (1.5)	0.5–1.5 (1)
D-raffinose	-	+	-	+
rifamycin sv	-	-	+	+
Gelatin	-	+	-	-
L-arginine	+	-	+	+
β-galactosidase	-	+	+	-
glucosidase	-	ND	+	-
DNA G+C content (mol%)	58.6	61.6	60.0	63.2

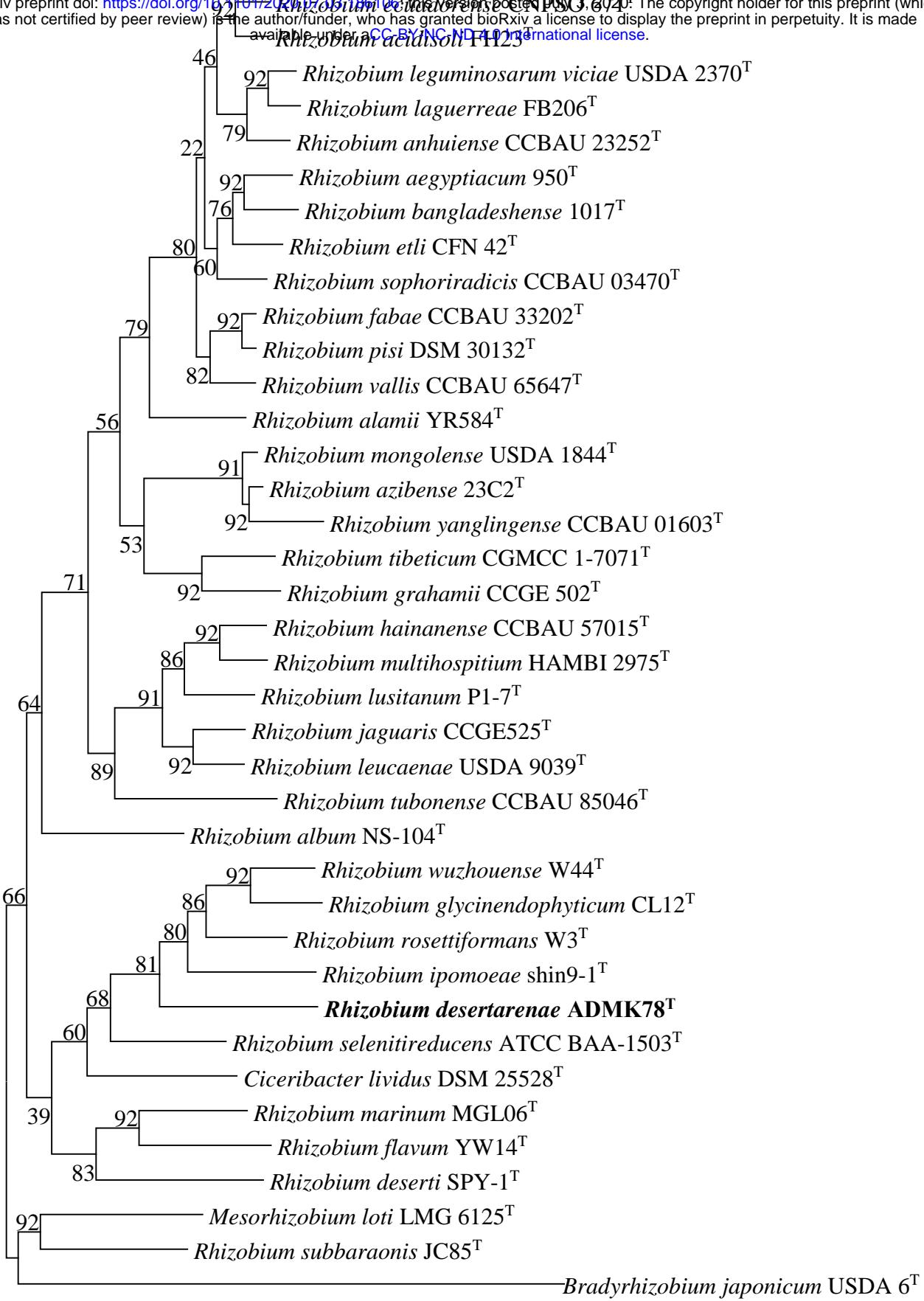
361 *Data from Tao Yuan et al., [27]

362
363
364
365

366



0.10



0.05