
This is the **accepted version** of the journal article:

Atanasov, K.E.; Galbis, D.M.; Cornadó, D.; [et al.]. «*Pseudomonas fitomaticsae* sp. nov., isolated at Marimurtra Botanical Garden in Blanes, Catalonia, Spain». *International Journal of Systematic and Evolutionary Microbiology*, Vol. 72 Núm. 10 (2022), p. 5557. DOI 10.1099/IJSEM.0.005557

This version is available at <https://ddd.uab.cat/record/273536>

under the terms of the  license

1 ***Pseudomonas fitomaticsae* sp. nov., isolated at Marimurtra Botanical Garden in Blanes,**
2 **Catalonia-Spain**

3
4 Kostadin Evgeniev Atanasov^{1,2}, David Miñana Galbis³, Deborah Cornadó⁴, Annabel Serpico⁴,
5 Guiomar Sánchez⁴, Montserrat Bosch⁴, Albert Ferrer^{1,5} and Teresa Altabella^{1,2}

6 ¹Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Bellaterra, Barcelona,
7 Spain.

8 ²Department of Biology, Healthcare and the Environment, Plant Physiology Section, Faculty
9 of Pharmacy and Food Sciences, Universitat de Barcelona, Spain.

10 ³Department of Biology, Healthcare and the Environment, Microbiology Section, Faculty of
11 Pharmacy and Food Sciences, Universitat de Barcelona, Spain.

12 ⁴Applied Microbiology and Biotechnology Unit, LEITAT Technological Center, Terrassa,
13 Spain.

14 ⁵Department of Biochemistry and Physiology, Faculty of Pharmacy and Food Sciences,
15 Universitat de Barcelona, Spain.

16 **Corresponding author:** Kostadin E. Atanasov

17 The email address for the corresponding author: evgenievatanassov@ub.edu

18 **Keywords:** *Pseudomonas fitomaticsae* FIT81^T, multilocus sequence analysis, genome-to-
19 genome comparison, genome assembly.

20 **Repositories:**

21 Bio Project: PRJNA705867; Bio Sample: SAMN19241968; SRR14663887; Assembly:
22 CP075567; 16s rRNA sequence ID: MZ773500.1.

23

24 **ABSTRACT**

25 In the framework of the research project called FITOMATICS, we have isolated and
26 characterized a bacterial plant-endophyte from the rhizomes of *Iris germanica*, hereafter
27 referred to as strain FIT81^T. The bacterium is Gram negative, rod-shaped with lophotrichous
28 flagella, and catalase and oxidase positive. The optimal growth temperature of strain FIT81^T is
29 28°C, although it can grow within a temperature range of 4°C to 32°C. The pH growth tolerance
30 ranges between 5 and 10, and it tolerates 4% (w/v) NaCl concentration. A 16S rRNA
31 phylogenetic analysis positioned strain FIT81^T within the genus *Pseudomonas*, and multilocus
32 sequence analysis (MLSA) revealed that *P. gozinkensis* IzPS32d^T, *P. glycinae* MS586^T, *P.*
33 *allokribbensis* IzPS23^T, '*P. kribbensis*' 46-2, and *P. koreensis* PS9-14^T are the top five most
34 closely related species, which were selected for further genome-to-genome comparisons, as
35 well as for physiological and chemotaxonomical characterization. The genome size of strain
36 FIT81^T is 6,492,796 base-pairs in length, with 60.6% of GC content. Average nucleotide
37 identity (ANI) and digital DNA-DNA hybridization (dDDH) yielded values of 93.6% and
38 56.1%, respectively, when FIT81^T genome was compared to that of the closest type-species *P.*
39 *gozinkensis* IzPS32d^T. Taken together, the obtained genomic, physiologic, and
40 chemotaxonomic data indicate that strain FIT81^T is different from its closest relative species,
41 which lead us to suggest that it is a novel species to be included in the list of type-strains with
42 the name *Pseudomonas fitomaticsae* FIT81^T (=CECT 30374^T =DSM 112699^T).

43

44 INTRODUCTION

45 Land plants are soil-dependent organisms that have the capacity to interact with a plethora of
46 multikingdom microorganisms. Among these are the so-called bacterial endophytes, which can
47 enter and reside within the plant tissues without causing disease symptoms but, on the contrary,
48 play a crucial role in promoting plant growth and defense by stimulating nutrient mobilization,
49 synthesis of hormones and nutrients, and protection against biotic and abiotic stress [1, 2]. In
50 addition to their recognized potential as plant-growth promoting organisms, endophytic
51 bacteria have gained special interest because of their capacity to produce a broad range of
52 bioactive metabolites, such as bactericides, antibiotics, fungicides, antivirals, and many others
53 [3, 4].

54 One of the most frequently genus of endophytic bacterial is *Pseudomonas*, which is a
55 ubiquitous Pseudomonadota, rod-shaped, aerobic, Gram negative and non-spore forming with
56 high capacity to utilize diverse compounds as carbon source, colonize a broad range of
57 eukaryotic host, and produce a wide array of specialized metabolites [5]. So far, a total of 299
58 *Pseudomonas* species with validly published and correct names are described for this genus
59 (<https://lpsn.dsmz.de/genus/pseudomonas>).

60 *Pseudomonas* species classification was first performed by using traditional phenotypical
61 characterization methods but, more recently, molecular techniques have emerged as key
62 analytical tools to aim species delineation. At present, multilocus sequence analysis (MLSA)
63 technique is used to study phylogenetic relationships among species. For instance, by using
64 MLSA, Peix *et al.* [6] suggested the division of *Pseudomonas* in three lineages: *P. fluorescens*,
65 *P. aeruginosa* and *P. pertucinogena*, from which *P. fluorescens* was also classified in seven
66 different groups. On the other hand, Hesse *et al.* [7] compared the genomes of type and none-
67 type strains, as well as of subspecies, and suggested the division of *Pseudomonas* in thirteen
68 groups of species, being *P. fluorescens* the group containing the highest number of species,

69 which was then divided in ten subgroups. More recently, Rudra and Gupta [8] used 118
70 concatenated conserved protein codifying genes from 255 genomes and showed a distinct
71 lineage for some halotolerant *Pseudomonas* species which were grouped in a Pertucinogena
72 clade. In further analysis, the authors identified 24 conserved signature indels in diverse cellular
73 functions proteins which supported the reclassification of *P. pertucinogena* to the new
74 proposed genus *Halopseudomonas* were the type species is *Halopseudomonas pertucinogena*.
75 In this work, we describe the characterization of a new *Pseudomonas* species candidate isolated
76 from the rhizomes of *Iris germanica*, *Pseudomonas fitomaticsae*, in the framework of a
77 research project (FITOMATICS) focused on the isolation and characterization of plant-
78 endophytic microorganisms with the potential to produce secondary specialized metabolites.

79 **ISOLATION AND ECOLOGY**

80 For bacteria isolation, rhizomes of *Iris germanica* from the Marimurtra Botanical Garden of
81 Blanes in Catalonia, Spain (41.67666° N; 2.80194 E) were used. Plants were examined for
82 pests, lesions or any disease symptoms, and only healthy plants were selected. Rhizomes were
83 collected *in situ*, maintaining plant integrity, and soil was detached manually from the
84 rhizomes. Samples were washed with sterile distilled water and phosphate buffered saline
85 solution (PBS) at pH 7.4, and surface sterilized essentially as reported by Kumar *et al.* [9]. The
86 rhizomes were first soaked in 70% ethanol for 5 min and then in 2% sodium hypochlorite
87 solution for 10 min, followed by 45 min in 70% ethanol. Finally, they were thoroughly washed
88 with sterile distilled water to remove any remaining ethanol. Endophytic bacteria were released
89 from the inner plant tissue by homogenizing the rhizomes in PBS. Then a two-fold serial
90 dilution of the crude-extract was prepared, and 50 µl of each sample were spread on Tryptic
91 Soy Agar medium (TSA, Sigma-Aldrich, St Louis, USA) plates, which were incubated at 30°C
92 for up to two weeks. As soon as colonies appeared, they were individually transferred to 3 ml
93 of Tryptic Soy Broth (TSB, Becton Dickinson, Franklin Lakes, USA) medium and cultured

94 over-night at 30°C with constant shaking at 190 rpm. After endophyte extraction, the rhizomes
95 were incubated in TSA medium at 30°C for up to three weeks, and no bacterial growth was
96 observed, thus demonstrating the effectiveness of the sterilization procedure. For long-term
97 strain preservation, 800 µl of an over-night culture were mixed with 200 µl of 80% sterile
98 glycerol and stored at -80°C.

99 **PHYLOGENETIC ANALYSIS**

100 For DNA extraction, 2 ml of strain FIT81^T over-night TSB culture was used. Cells were
101 recovered by centrifugation at 11,000 g for 3 min and the genomic DNA was extracted with
102 Wizard Genomic DNA Extraction Kit according to manufacturer's instructions (Promega,
103 Madison, USA). Thirty ng of the genomic DNA was used for the 16S rRNA gene PCR
104 amplification with the universal primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and
105 1492R (5'-GGTTACCTTGTTACGACTT-3') [10], and AccuPrime Pfx polymerase (2.5 units)
106 (Thermo Fisher Scientific, Waltham, USA). The PCR program was as follows: DNA
107 denaturalization at 95°C for 5 min, 25 cycles of amplification including denaturalization at
108 95°C for 15 s, annealing at 50°C for 30 s, 2 min extension at 72°C, and 10 min at 72°C for
109 complete synthesis of amplification product. The PCR product was analyzed by agarose gel
110 electrophoresis, purified using PCR clean-up columns (Macherey-Nagel GmbH & Co. KG,
111 Düren, Germany), and SANGER-sequenced using primers 27F or 1492R. The resulting partial
112 sequences were assembled in a unique consensus sequence with the DNA Dragon software
113 (SequentiX, Klein Raden, Germany), which was analyzed by nucleotide BLAST (blastn)
114 alignment on the NCBI database (<https://www.ncbi.nlm.nih.gov/>) under the parameters of
115 standard database, uncultured/environmental sample sequence exclusion, and search limited to
116 type material. Blast results positioned strain FIT81^T within the genus *Pseudomonas* with 100%
117 of sequence similarity to '*P. kribbensis*' 46-2, *P. glycinae* MS586^T and *P. saponiphila* DSM
118 9751^T. Next, the entire 16S rRNA sequence was extracted from the FIT81^T genome annotation,

119 and a species search on the EzBioCloud taxonomically united database [11] was performed. A
120 total of 33 species were selected from the EzTaxon and 13 from the NCBI nucleotide blast. *H.*
121 *pertucinogena* (homotypic synonym of *P. pertucinogena*) NBRC 14163^T (AB680571) was
122 used as a rooting group. The 16S rRNA sequences were aligned by the MUSCL algorithm [12]
123 and a phylogenetic tree was constructed by the Maximum Likelihood method based on the
124 Jukes-Cantor mode with MEGA7 [13, 14]. This analysis outputted an unprecise tree-branch
125 clustering of strain FIT81^T, showing 99.9% of sequence similarity to *P. gozinkensis* IzPS32d^T,
126 '*P. kribbensis*' 46-2, and *P. glycinae* MS4586^T (Fig. S1; Table S1). Hence, by using 16S rRNA
127 phylogeny, it was not possible to locate the newly isolated strain in any subgroup of the *P.*
128 *fluorescens* group [7]. Thus, a MLSA was conducted to improve resolution on the strain
129 phylogeny. Protein-coding housekeeping genes *rpoB*, *gyrB* and *rpoD* were extracted from the
130 NCBI genome annotation, whereas for some species genomes were annotated by Patric 3.6.12
131 with the Rapid Annotation using Subsystems Technology algorithm (RAST) [15]. Sequences
132 of the 16S rRNA, *rpoB*, *gyrB*, and *rpoD* genes were concatenated, aligned with MUSCL
133 algorithm, and a new MLSA-based Maximum Likelihood tree constructed (Fig. 1). The overall
134 sequences similarities of the FIT81^T concatenated housekeeping genes were as follows: 98.2%
135 to *P. gozinkensis* IzPS32d^T and *P. glycinae* MS4586^T; 97.9% to '*P. kribbensis*' 46-2; 97.4% to
136 *P. allokribbensis* IzPS23^T and 97.0% to the *P. koreensis* PS9-14^T (Table S2). These results
137 further supported the view that strain FIT81^T is closely related to *P. glycinae* MS4586^T and *P.*
138 *gozinkensis* IzPS32d^T and positioned strain FIT81^T as a new candidate of the *P. koreensis*
139 subgroup of the *P. fluorescens* group with a clear clustering differentiation to other *P. jessenii*
140 and *P. gessardii* subgroups. Thus, the selection of species for the whole genome-based species
141 comparisons and the physiological and chemical characterization was based on these MLSA-
142 based phylogenetic relationships.

143

144 GENOME FEATURES

145 Genomic DNA from strain FIT81^T was extracted as described above, treated with RNase A
146 DNase free (Panreac AppliChem, Darmstadt, Germany), and deproteinized by chloroform-
147 isoamyl alcohol treatment. The DNA was cleaned and concentrated by using Genomic DNA
148 Clean & Concentrator-10 columns (Zymo Research, Irvin, USA), and whole genome
149 sequencing was carried out by combining the Illumina Novaseq (2x150 bp) and PacBio RSII
150 platforms. The resulting sequence data were combined to generate a hybrid genome assembly
151 by using Unicycler software [16]. Genome size of FIT81^T is 6,492,796 base-pair length with
152 60.6% of GC content, which agrees with genome size and GC content of *Pseudomonas* species
153 isolated from environmental samples [17]. Gene annotation was performed by NCBI
154 Prokaryotic Genome Annotation Pipeline (PGAP) [18–20]. A total number of 5,742 protein-
155 coding genes, 19 rRNA coding genes (7 genes coding for 5S rRNAs, 6 genes coding for 16S
156 rRNAs and 6 genes coding for 23S rRNAs), and 72 tRNAs coding genes were predicted.
157 Genome circularization was achieved by visualization of FIT81^T genome with Bandage
158 software [21].

159 Furthermore, genome-to-genome computational comparisons were conducted between FIT81^T
160 and the top five related species. The average nucleotide identity (ANI) and orthologous ANI
161 (orthoANI) values obtained using the orthoANI Tool v0.93.1 [22] were, 93.6% and 94.2% to
162 the *P. gozinkensis* IzPS32d^T; 93.6% and 94.0% to *P. glycinae* MS4586^T; 91.0% and 91.5% to
163 *P. allokribbensis* IzPS23^T; 90.9% and 91.3% to '*P. kribbensis*' 46-2, 87.5% and 87.9% to the
164 *P. koreensis* PS9-14^T, respectively. Usearch ANI analysis (ANId) was performed with the ANI
165 calculator web-based tool from the EzBioCloud [23] and ANI maximal unique matches
166 (ANIm) estimated by JSpeciesWS [24]. The obtained ANId and ANIm values were very
167 similar to those from the orthoANI analysis. On the other hand, values obtained from the
168 BLAST-based ANI (ANId) calculated with JSpeciesWS were like the original ANI (Table 1).

169 Finally, we estimated the DNA-DNA hybridization (dDDH) values between species with the
170 Genome-to-Genome Distance Calculator (GGDC) from the DSMZ webtool ([https://ggdc-](https://ggdc-test.dsmz.de/ggdc.php#)
171 [test.dsmz.de/ggdc.php#](https://ggdc-test.dsmz.de/ggdc.php#)), with the BLAST+ local alignment algorithm using formula 2, which
172 consists of the sum of all identities within high-scoring segment pairs (HSPs) divided per
173 overall HSP length [25, 26]. Values from the dDDH were as follows: 56.1% [53.4 - 58.8%] to
174 *P. gozinkensis* IzPS32d^T; 54.9% [52.1 - 57.6%] to *P. glycinae* MS4586^T; 44.4% [41.9 - 47.0%]
175 to *P. allokribbensis* IzPS23^T; 43.9% [41.4 - 46.5%] to '*P. kribbensis*' 46-2; 34.7% [32.2 -
176 37.2%] to the *P. koreensis* PS9-14^T. All these genome-to-genome comparison percentages
177 were below the threshold of 95-96% (ANI) and 70% (dDDH) for species boundaries
178 delineation [27, 28], indicating there is enough nucleotide identity and dDDH differences to
179 propose that strain FIT81^T is a different species from its closest relatives, which is further
180 supported by the recently described methodology of bacteria and subspecies delineation
181 proposed by Meier-Kolthoff on the Type (Strain) Genome Server (TYGS; <https://tygs.dsmz.de>)
182 for prokaryote taxonomy [29]. FIT81^T genome was uploaded on the TYGS server and a total
183 number of 14 out of 15,682 type strains available in the TYGS database passed the analysis as
184 closest type strains. The position of FIT81^T in the phylogenetic tree constructed using the
185 Genome Blast Distance Phylogeny (GBDP) calculator (Fig. 2) agreed with the MLSA
186 phylogeny and the genome-to-genome comparison results.

187 **GENOME MINING AND IDENTIFICATION OF ENDOPHYTIC LIFE-STYLE** 188 **RELATED GENES**

189 Genome assemblies of strains FIT81^T, *P. gozinkensis* IzPS32d^T, and *P. glycinae* MS4586^T were
190 analyzed by using the Antibiotics and Secondary Metabolite Analysis Shell (antiSMASH) 6.0
191 webserver [30]. This computational analysis predicted the occurrence of 14 regions containing
192 10 putative biosynthetic gene clusters (BGCs) (Table S3). According to this, known cluster
193 similarities were found for the cyclic lipopeptide lokisin (71%; *Pseudomonas* sp. [31]), the

194 oxidative stress protectant aryl polyene (APE) Vf or methyl (2E,4E,6E,8E,10E,12E)-13-(4-
195 hydroxy-3,5-dimethylphenyl)trideca-2,4,6,8,10,12-hexaenoate (40%; *Aliivibrio fischeri*
196 ES114 [32]), the detoxing compound rimosamide (20%; *Streptomyces rimosus* subsp. *rimosus*
197 ATCC 10970 [33]), the chromophoric siderophore pyoverdine (21%, 20%, and 11%;
198 *Pseudomonas protegens* Pf-5[34, 35]), the bioactive cyclic lipopeptide fengycin (13%;
199 *Bacillus velezensis* FZB42 [36, 37]), and the antibiotic lankacidin C (13%; [38–42]).
200 Interestingly a similar cluster prediction was obtained by antiSMASH among the three bacterial
201 species. Nevertheless, differences were found regarding the BGC for the nonribosomal
202 bioactive cyclic lipopeptide, bacillomycin D, which was predicted into the genome of *P.*
203 *glycinae* MS4586^T (20%; *B. velezensis* FZB42 [43]), but not in FIT81^T nor in *P. gozinkensis*
204 IzPS32d^T genomes. In addition, subtle differences exist concerning the BGC for the antifungal
205 compound fragin, which was detected in *P. gozinkensis* IzPS32d^T and *P. glycinae* MS4586^T
206 (37%; *Burkholderia cenocepacia* H111 [44]), but not in FIT81^T (Table S3). Genomic regions
207 containing BGCs were also annotated by using the Minimum Information about a Biosynthetic
208 Gene cluster (MIBiG [45]) repository, and the top three similarity score-based compounds
209 annotated (Table S3). To summarize, the genome mining analysis suggested a high
210 conservation of the lokisin BGC and variation some among species in the case of the BGCs for
211 bacillomycin D and fragin. The usefulness of those clusters as species and subspecies
212 identification markers could be further investigated.

213 During last decade, several genes have been suggested to serve as markers for the identification
214 of bacterial endophytes [46–51]. Thus, besides the genome mining, we performed a
215 computational search to identify endophytic-capacity life-style genes. This analysis revealed
216 that several endophyte gene markers potentially involved in phosphate solubilization, hormone
217 synthesis and response, metabolite transport, secretion and delivery systems, cell wall
218 attachment, plant polymer/modification, transcriptional regulators, siderophores, and

219 detoxification are present in the genomes of FIT81^T, *P. gozinkensis* IzPS32d^T, and *P. glycinae*
220 MS586^T (Table S4). Interestingly, a detoxification-related gene coding for acetaldehyde
221 dehydrogenase was found only in the genome of FIT81^T. These results together with the way
222 in which FIT81^T was isolated would confirm the endophytic nature of FIT81^T.

223 **PHYSIOLOGY AND CHEMOTAXONOMY**

224 Growth of strain FIT81^T is vigorous and forms white-mucoid colonies. The bacterium can grow
225 within a temperature range from 4°C to 32°C in TSA medium, being 28°C the optimal growth
226 temperature. The pH growth tolerance in TSB medium was recorded within the range of pH5
227 to pH10, even though FIT81^T can resist alkaline conditions up to pH 12, at this pH growth is
228 seriously compromised. Limited growth capacity was observed at pH 4, and no growth was
229 observed at lower pH values after three weeks. Strain FIT81^T is a Gram-negative bacterium.
230 Uranyl acetate negative stain and transmission electron microscope (TEM) Bioscan Gatan,
231 JEM-1010 (JEOL) acquired image were used to visualize its lophotrichous flagella and to
232 measure the average cell size, which is 2±0.3 µm length and 0.8±0.2 µm width (N=10) (Fig.
233 3). Fluorescence production was observed when FIT81^T was cultured on solid King's B
234 medium (Duchefa Biochemie, Haarlem, TheNetherlands) at 28°C and irradiated by 365 nm UV
235 light [52].

236 Polar lipids and fatty acids methyl esters (FAME) were determined by the Identification
237 Service, Leibniz - Institute DSMZ – Deutsche Sammlung von Mikroorganismen und
238 Zellkulturen GmbH (Braunschweig, Germany) from an over-night culture of FIT81^T in TSB
239 medium incubated at 28°C with shaking at 190 rpm. Cells were pelleted by centrifugation at
240 4,000 g for 10 min at 4°C, mixed with cryoprotectant ATCC Reagent-18
241 (<https://www.atcc.org/resources/culture-guides/bacteriology-culture-guide>), and freeze-dried
242 until further analysis. Cellular fatty acids from 30 mg of cells biomass were analyzed after
243 conversion into fatty acid methyl esters (FAMES) by saponification, methylation and extraction

244 using minor modifications of the methods described by Miller [53] and Kuykendall *et al.* [54].
245 The FAME mixtures were then separated with gas chromatography and detected by a flame
246 ionization detector using Sherlock Microbial Identification System (MIS) (MIDI, Microbial
247 ID, Newark, DE 19711 U.S.A.). Peaks were automatically integrated, and fatty acid were
248 identified and quantified by the MIS Standard Software (Microbial ID). Summed feature 3 was
249 resolved by a GC-MS run on a Agilent GC-MS 7000D using an Agilent HP-5 ms UI 30 m x
250 250 μm x 0.25 μm column, with a helium flow of 1.2 ml, with an injection of 1 μl with split
251 ratio of 7.5:1. The oven program was as follows: initial temperature 170°C, ramp 3°C/min to
252 200°C, ramp 5°C/min to 270°C, ramp 120°C/min to 300°C and hold for 2 min. The inlet
253 temperature was set to 170°C and then linearly increased with 200°C/min up to 350°C and hold
254 for 5 min. The mass spectrometry parameters were set to aux temperature 230°C, source
255 temperature 230°C, and electron impact ionization at 70 eV with mass range of m/z 40-600 or
256 40-800, respectively. Peaks were identified based on retention time and mass spectra. The
257 position of single and double bounds was confirmed by derivatization to the corresponding
258 dimethyl disulfide adducts [55]. Branched-chain fatty acid positions, cyclo-positions and
259 multiple double bounds were determined by conversion to their 3-pyridylcarbinol and/or 4,4-
260 dimethyloxazoline (DMOX) derivatives [56–58].

261 FIT81^T showed a *Pseudomonas* species typical profile C_{10:0} 3-OH, C_{12:0}, and C_{12:0} 3-OH fatty
262 acids, which further supported its affiliation to genus *Pseudomonas* [59]. In addition, the major
263 FAME was the unsaturated sixteen carbon aliphatic chain (C_{16:0}) with an average value of 30.6,
264 which is slightly lower than those found in *P. gozinkensis* IzPS32d^T, *P. allokribbensis* IzPS23^T
265 [20], '*P. kribbensis*' 46-2 [60], and *P. koreensis* PS9-14^T, but higher than those described for *P.*
266 *glycinae* MS586^T [61]. Other fatty acids detected as FAMEs, were C_{18:1} ω 7c (14.4); C_{12:0} 3-
267 OH (4.4); C_{10:0} 3-OH (4.1); C_{12:0} 2-OH (3.9); C_{17:0} cyclo (3.8); C_{12:0} (3.7); C_{18:0} (1.48), and
268 traces of C_{10:0}, C_{12:1} 3-OH, and C_{14:0}. Moreover, the summed feature 3 of C_{16:1} ω 7c/C_{15:0} iso 2-

269 OH was clearly higher in FIT81^T than the levels found in *P. gozinkensis* IzPS32d^T [20], *P.*
270 *glycinae* MS586^T [61], *P. allokribbensis* IzPS23^T [20], and '*P. kribbensis*' 46-2 [60] (Table 2).
271 Polar lipids (PL) were extracted from 200 mg of freeze-dried cell material using a
272 chloroform:methanol:0.3% aqueous NaCl mixture and recovered into the chloroform phase
273 [62]. PL were then separated by two-dimensional silica gel thin layer chromatography (TLC)
274 by using chloroform:methanol:water and chloroform:methanol:acetic acid:water as mobile
275 phase for first and second direction, respectively. Total lipid material was detected by using
276 molybdato-phosphoric acid and specific functional groups were identified using spray reagents
277 specific for defined functional groups [63]. The detected major compounds species were
278 phosphatidylethanolamine (PE), diphosphatidylglycerol (DPG), phosphatidylglycerol (PG),
279 glycolipid (GL), phospholipid (PL), and aminolipid (AL) (Fig. 4).
280 Carbon source utilization and chemical sensitivity was assessed with the miniaturized
281 microplate assay of Biolog's GenIII (Biolog, Inc., Hayward, USA) following manufacturer's
282 instructions. The main differences between FIT81^T, *P. gozinkensis* IzPS32d^T, and *P. glycinae*
283 MS586^T were found in the utilization of D-arabitol, D-serine, D-fructose-6-phosphate, D-
284 galactose, pectin, D-fucose, inosine, D-galacturonic acid, L-galactonic acid lactone,
285 glucuronamide, D-malic acid, bromo-succinic acid, α -keto-butyric acid, acetoacetic acid, and
286 tween 40 (Table 3). FIT81^T exhibited chemical sensitivity to D-serine, guanidine HCl, lithium
287 chloride, sodium butyrate, and sodium bromate, but it was tolerant to sodium lactate, fusidic
288 acid, troleandomycin, rifamycin SV, minocycline, lincomycin, niaproof 4, vancomycin,
289 tetrazolium violet, tetrazolium blue, nalidixic acid, potassium tellurite, and the monobactam
290 aztreonam. FIT81^T shares high chemical compound tolerance with *P. gozinkensis* IzPS32d^T
291 [23], but differences were found in the case of fusidic acid and minocycline. Enzyme activities
292 of strain FIT81^T were analyzed on API 20 NE and API Zym strips (Biomerieux, Marcy-l'Étoile,
293 France). The results obtained revealed that FIT81^T exhibits arginine dihydrolase, esculin

294 hydrolysis (β -glucosidase), and gelatin hydrolysis (protease) activities. It can also grow with
295 N-acetyl-glucosamine while *P. gozinkensis* IzPS32d^T cannot. FIT81^T and *P. gozinkensis*
296 IzPS32d^T have both C4 and C8 esterase lipase activity. Moreover, FIT81^T exhibited alkaline
297 phosphatase, C14 lipase, trypsin, naphthol-AS-BI-phosphohydrolase, and acid phosphatase
298 activities, which have not been described in *P. gozinkensis* IzPS32d^T. It is worth noting that
299 acid phosphatase activity has been reported as a mechanism of phosphate solubilization from
300 organic compounds, which further supports our claim that FIT81^T is an endophytic bacterium.
301 Finally, the above biochemical comparisons among the five closest relative species to FIT81^T,
302 as well as additional biochemical traits screened in API 20E and API 50CH, are summarized
303 in more detail in Table 3.

304 In conclusion, the results from *de novo* genome sequencing, and the pangenome computational
305 comparative analyses based on the top five phylogenetically-related species to strain FIT81^T,
306 together with a thorough characterization of the bacterium chemo-physiological traits, leads to
307 the claim that strain FIT81^T represents a novel *Pseudomonas* species.

308 **DESCRIPTION OF *PSEUDOMONAS FITOMATICSAE* SP. NOV.**

309 *Pseudomonas fitomaticsae* (fi.to.ma'tics.ae. N.L. gen. n. *fitomaticsae*, referring to the
310 FITOMATICS research project), strain FIT81^T, isolated from the rhizomes of *Iris germanica*
311 accessed at the Marimurtra Botanical Garden, from Blanes, Catalonia, Spain (41.67666° N;
312 2.80194 E).

313 The bacterium is Gram negative, rod-shaped with lophotrichous flagellum, and its average cell
314 size is 2±0.3 μ m length and 0.8±0.2 μ m width. Its optimal growth temperature is 28°C, forming
315 white mucoid colonies in TSA medium, although it is also able to grow in a range of
316 temperatures comprised between 4°C and 32°C. Growth was observed at pH values ranging
317 from 5 to 10, with higher tolerance to alkaline pH than acidic one. The genome size is 6,492,796
318 base-pairs, with 60.6% GC content, and is predicted to harbor 5,742 protein-coding genes and

319 19 rRNA-coding genes. Among the latter, 7 genes are predicted to encode the 5S rRNA, 6
320 genes the 16S rRNA, and 6 genes the 23S rRNA. The type-strain of *P. fitomaticsae* FIT81^T
321 was deposited at international recognized cell-type culture collection (FIT81^T =CECT 30374^T
322 =DSM 112699^T) and is available for further biochemical, physiological, taxonomic, and
323 genetic studies. Sequence Read Archive (SRA) was deposited at the National Center for
324 Biotechnology Information (NCBI) database under the identification codes SRR14663887,
325 PRJNA705867 (Bio Project), and SAMN19241968 (Bio Sample). 16S rRNA sequence was
326 deposited in GenBank under the identification code MZ773500, and genome assembly was
327 assigned the code CP075567.

328 **Authors and contributors**

329 Kostadin E. Atanasov (KA) designed experiments with the contribution of David Miñana-
330 Galbis (DMG), Teresa Altabella (TA) and Albert Ferrer (AF). KA conducted experiments and
331 analyzed data with contribution of DMG, TA, and AF. Deborah Cornadó (DC) isolated strain
332 FIT81^T. Annabel Serpico (AS) participated in strain isolation and culturing. Guiomar Sánchez
333 (GS) and Montserrat Bosch (MB) designed and supervised DC and AS work. This article was
334 written by KA with contributions of all authors.

335 **Conflicts of interest**

336 Authors honestly declare that there are no economical, ethical, or publishing conflicts of
337 interest.

338 **Funding information**

339 This work was funded by grants RTC-2017-6431 from FEDER/Ministerio de Ciencia,
340 Innovación y Universidades-Agencia Estatal de Investigación (Spain), 2017SGR710 from the
341 Generalitat de Catalunya, and by the CERCA Programme of the Generalitat de Catalunya. We
342 also acknowledge financial support from the Spanish Ministerio de Economía y

343 Competitividad-Agencia Estatal de Investigación through the “Severo Ochoa Programme for
344 Centres of Excellence in R&D” SEV-2015- 0533 and CEX2019-000902-S.

345 **Acknowledgements**

346 We thank to Leigh A. Riley for help in genome annotation and submission at the NCBI
347 GenBank, Bethesda, Maryland USA. KA dedicates this manuscript in memory of Professor
348 Antonio Fernández Tiburcio from the Plant Physiology Section at the Faculty of Pharmacy and
349 Food Science of the Universitat de Barcelona.

350

351 **References**

- 352 1. **Kandel SL, Joubert PM, Doty SL.** Bacterial endophyte colonization and distribution
353 within plants. *Microorganisms* 2017;5:77.
- 354 2. **Berg G, Köberl M, Rybakova D, Müller H, Grosch R, et al.** Plant microbial diversity
355 is suggested as the key to future biocontrol and health trends. *FEMS Microbiol Ecol*
356 2017;93:10.1093/femsec/fix050.
- 357 3. **Gouda S, Das G, Sen SK, Shin HS, Patra JK.** Endophytes: A treasure house of
358 bioactive compounds of medicinal importance. *Front Microbiol* 2016;7:1538.
- 359 4. **Hagaggi NSA, Mohamed AAA.** Plant-bacterial endophyte secondary metabolite
360 matching: a case study. *Arch Microbiol* 2020;202:2679-2687.
- 361 5. **Rieusset L, Rey M, Muller D, Vacheron J, Gerin F, et al.** Secondary metabolites
362 from plant-associated *Pseudomonas* are overproduced in biofilm. *Microb Biotechnol*
363 2020;13(5):1562-1580.
- 364 6. **Peix A, Ramírez-Bahena MH, Velázquez E.** The current status on the taxonomy of
365 *Pseudomonas* revisited: An update. *Infect Genet Evol* 2018;57:106-116.
- 366 7. **Hesse C, Schulz F, Bull CT, Shaffer BT, Yan, et al.** Genome-based evolutionary
367 history of *Pseudomonas* spp. *Environ Microbiol* 2018;20:2142-2159.
- 368 8. **Rudra B, Gupta RS.** Phylogenomic and comparative genomic analyses of species of
369 the family *Pseudomonadaceae*: Proposals for the genera *Halopseudomonas* gen. nov. and
370 *Atopomonas* gen. nov., merger of the genus *Oblitimonas* with the genus *Thiopseudomonas*, and
371 transfer of some misclassified species of the genus *Pseudomonas* into other genera. *Int J Syst*
372 *Evol Microbiol.* 2021;71:10.1099/ijsem.0.005011.
- 373 9. **Kumar A, Singh R, Yadav A, Giri DD, Singh PK, et al.** Isolation and characterization
374 of bacterial endophytes of *Curcuma longa* L. *3 Biotech* 2016;6:60.

- 375 10. **Miller CS, Handley KM, Wrighton KC, Frischkorn KR, Thomas BC, et al.** Short-
376 read assembly of full-length 16S amplicons reveals bacterial diversity in subsurface sediments.
377 PLoS One 2013;8:e56018
- 378 11. **Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, et al.** Introducing EzBioCloud: A
379 taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies.
380 Int J Syst Evol Microbiol 2017;67:1613-1617.
- 381 12. **Edgar RC.** MUSCLE: Multiple sequence alignment with high accuracy and high
382 throughput. Nucleic Acids Res 2004;32:1792-1797.
- 383 13. **Kumar S, Stecher G, Tamura K.** MEGA7: Molecular evolutionary genetics analysis
384 version 7.0 for bigger datasets. Mol Biol Evol 2016;33:1870-1874.
- 385 14. **Matsubara H, Yamanaka T (1969).** Evolution of protein molecules. In: Mammalian
386 Protein Metabolism, pp. 21-132. Edited by Munro HN, Academic Press, New York.
- 387 15. **Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, et al.** RASTtk: A modular and
388 extensible implementation of the RAST algorithm for building custom annotation pipelines
389 and annotating batches of genomes. Sci Rep 2015;5:8365.
- 390 16. **Wick RR, Judd LM, Gorrie CL, Holt KE.** Unicycler: Resolving bacterial genome
391 assemblies from short and long sequencing reads. PLoS Comput Biol 2017;13:e1005595.
- 392 17. **Scales BS, Erb-Downward JR, Huffnagle IM, LiPuma JJ, Huffnagle GB.**
393 Comparative genomics of *Pseudomonas fluorescens* subclade III strains from human lungs.
394 BMC Genomics. 2015;16:1032.
- 395 18. **Haft DH, DiCuccio M, Badretdin A, et al.** RefSeq: An update on prokaryotic genome
396 annotation and curation. Nucleic Acids Res.2018;46:D851-D860.
- 397 19. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, et al.** NCBI
398 prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016;44:6614-6624.

- 399 20. **Morimoto Y, Lu YJ, Zuo H, Aibibula Z, Tohya M, et al.** *Pseudomonas*
400 *allokribbensis* sp. nov. and *Pseudomonas gozinkensis* sp. nov., Two new species isolated from
401 a volcanic island, Izu Oshima, Japan. *Curr Microbiol* 2021;78:1670-1677.
- 402 21. **Wick RR, Schultz MB, Zobel J, Holt KE.** Bandage: interactive visualization of de
403 novo genome assemblies. *Bioinformatics*. 2015;31:3350-3352.
- 404 22. **Lee I, Ouk Kim Y, Park SC, Chun J.** OrthoANI: An improved algorithm and software
405 for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 2016;66:1100-1103.
- 406 23. **Yoon SH, Ha SM, Lim J, Kwon S, Chun J.** A large-scale evaluation of algorithms to
407 calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281-1286.
- 408 24. **Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J.** JSpeciesWS: a web
409 server for prokaryotic species circumscription based on pairwise genome comparison.
410 *Bioinformatics* 2016;32:929-931.
- 411 25. **Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M.** Genome sequence-based species
412 delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*
413 2013;14:60.
- 414 26. **Auch AF, Klenk HP, Göker M.** Standard operating procedure for calculating genome-
415 to-genome distances based on high-scoring segment pairs. *Stand Genomic Sci* 2010;2:142-148.
- 416 27. **Chun J, Oren A, Ventosa A, et al.** Proposed minimal standards for the use of genome
417 data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461-466.
- 418 28. **Auch AF, von Jan M, Klenk HP, Göker M.** Digital DNA-DNA hybridization for
419 microbial species delineation by means of genome-to-genome sequence comparison. *Stand*
420 *Genomic Sci* 2010;2:117-134.
- 421 29. **Meier-Kolthoff JP, Göker M.** TYGS is an automated high-throughput platform for
422 state-of-the-art genome-based taxonomy. *Nat Commun* 2019;10:2182.

- 423 30. **Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, et al.**
424 antiSMASH 6.0: Improving cluster detection and comparison capabilities. *Nucleic Acids Res*
425 2021;49:W29-W35.
- 426 31. **Omoboye OO, Oni FE, Batool H, Yimer HZ, De Mot R, Höfte M.** *Pseudomonas*
427 cyclic lipopeptides suppress the rice blast fungus *Magnaporthe oryzae* by Induced resistance
428 and direct antagonism. *Front Plant Sci* 2019;10:901.
- 429 32. **Cimermancic P, Medema MH, Claesen J, Kurita K, Wieland Brown LC, et al.**
430 Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene
431 clusters. *Cell* 2014;158:412-421.
- 432 33. **McClure RA, Goering AW, Ju KS, Baccile JA, Schroeder FC, et al.** Elucidating the
433 rimosamide-detoxin natural product families and their biosynthesis using metabolite/gene
434 cluster correlations. *ACS Chem Biol* 2016;11:3452-3460.
- 435 34. **Stintzi A, Johnson Z, Stonehouse M, Ochsner U, Meyer JM, et al.** The pvc gene
436 cluster of *Pseudomonas aeruginosa*: role in synthesis of the pyoverdine chromophore and
437 regulation by PtxR and PvdS. *J Bacteriol* 1999;181:4118-4124.
- 438 35. **Stintzi A, Cornelis P, Hohnadel D, Meyer JM, Dean C, et al.** Novel pyoverdine
439 biosynthesis gene(s) of *Pseudomonas aeruginosa* PAO. *Microbiology* 1996;142:1181-1190.
- 440 36. **Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, et al.** Comparative
441 analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus*
442 *amyloliquefaciens* FZB42. *Nat Biotechnol* 2007;25:1007-1014.
- 443 37. **Koumoutsi A, Chen XH, Henne A, Liesegang H, Hitzeroth G, et al.** Structural and
444 functional characterization of gene clusters directing nonribosomal synthesis of bioactive
445 cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. *J Bacteriol* 2004;186:1084-
446 1096.

- 447 38. **Arakawa K, Sugino F, Kodama K, Ishii T, Kinashi H.** Cyclization mechanism for
448 the synthesis of macrocyclic antibiotic lankacidin in *Streptomyces rochei*. Chem Biol
449 2005;12:249-256.
- 450 39. **Kinashi H, Fujii S, Hatani A, Kurokawa T, Shinkawa H.** Physical mapping of the
451 linear plasmid pSLA2-L and localization of the eryAI and actI homologs. Biosci Biotechnol
452 Biochem 1998;62:1892-1897.
- 453 40. **Masanori Suwa, Hiroyuki Sugino, Akiko Sasaoka, Eijiro Mori, Shingo Fujii, et al.**
454 Identification of two polyketide synthase gene clusters on the linear plasmid pSLA2-L in
455 *Streptomyces rochei*. Gene 2000;246:123-131.
- 456 41. **Hiratsu K, Mochizuki S, Kinashi H.** Cloning and analysis of the replication origin
457 and the telomeres of the large linear plasmid pSLA2-L in *Streptomyces rochei*. Mol Gen Genet
458 2000;263:1015-1021.
- 459 42. **Mochizuki S, Hiratsu K, Suwa M, Ishii T, Sugino F, et al.** The large linear plasmid
460 pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary
461 metabolism. Mol Microbiol 2003;48:1501-1510.
- 462 43. **Koumoutsi A, Chen XH, Henne A, Liesegang H, Hitzeroth G, et al.** Structural and
463 functional characterization of gene clusters directing nonribosomal synthesis of bioactive
464 cyclic lipopeptides in *Bacillus amyloliquefaciens* strain FZB42. J Bacteriol 2004;186:1084-
465 1096.
- 466 44. **Jenul C, Sieber S, Daeppen C, Mathew A, Lardi M, et al.** Biosynthesis of fragin is
467 controlled by a novel quorum sensing signal. Nat Commun 2018;9:1297.
- 468 45. **Satria A Kautsar, Kai Blin, Simon Shaw, Jorge C Navarro-Muñoz, Barbara R**
469 **Terlouw, et al.** MIBiG 2.0: A repository for biosynthetic gene clusters of known function.
470 Nucleic Acids Res 2020;48:D454-D458.

- 471 46. **Ali S, Duan J, Charles TC, Glick BR.** A bioinformatics approach to the determination
472 of genes involved in endophytic behavior in *Burkholderia* spp. *J Theor Biol* 2014;343:193-
473 198.
- 474 47. **Jahn L, Hofmann U, Ludwig-Müller J.** Indole-3-acetic acid is synthesized by the
475 endophyte *Cyanodermella asteris* via a tryptophan-dependent and - independent way and
476 mediates the interaction with a non-host plant. *Int J Mol Sci* 2021;22:2651.
- 477 48. **Oteino N, Lally RD, Kiwanuka S, Lloyd A, Ryan D, et al.** Plant growth promotion
478 induced by phosphate solubilizing endophytic *Pseudomonas isolates*. *Front Microbiol*
479 2015;6:745.
- 480 49. **Taghavi S, Garafola C, Monchy S, Newman L, Hoffman A, et al.** Genome survey
481 and characterization of endophytic bacteria exhibiting a beneficial effect on growth and
482 development of poplar trees. *Appl Environ Microbiol* 2009;75:748-757
- 483 50. **Cueva-Yesquén LG, Goulart MC, Attili de Angelis D, Nopper Alves M, Fantinatti-
484 Garboggini F.** Multiple plant growth-promotion traits in endophytic bacteria retrieved in the
485 vegetative stage from passionflower. *Front Plant Sci* 2021;11:621740.
- 486 51. **Jiménez-Gómez A, García-Estévez I, Escribano-Bailón MT, García-Fraile P,
487 Rivas R.** Bacterial fertilizers based on *Rhizobium laguerreae* and *Bacillus halotolerans*
488 enhance cichorium endivia L. Phenolic compound and mineral contents and plant development.
489 *Foods* 2021;10:424.
- 490 52. **King E, Ward MK, Raney DE.** Two simple media for the demonstration of pyocyanin
491 and fluorescin. *J Lab Clin Med* 1954;44:301–307.
- 492 53. **Miller LT.** Single derivatization method for routine analysis of bacterial whole-cell
493 fatty acid methyl esters, including hydroxy acids. *J Clin Microbiol* 1982;16:584-586.

- 494 54. **Kuykendall L D, Roy M A, O'Neill J J, Devine T E.** Fatty acids, antibiotic resistance,
495 and deoxyribonucleic acid homology groups of *Bradyrhizobiurn japonicum*. Int J Syst
496 Bacteriol 1988;38:358–361.
- 497 55. **Moss CW, Lambert-Fair MA.** Location of double bonds in monounsaturated fatty
498 acids of *Campylobacter cryaerophila* with dimethyl disulfide derivatives and combined gas
499 chromatography-mass spectrometry. J Clin Microbiol 1989;27:1467-1470.
- 500 56. **Yu QT, Liu BN, Zhang JY, Huang ZH.** Location of methyl branchings in fatty acids:
501 fatty acids in uropygial secretion of Shanghai duck by GC-MS of 4,4-dimethyloxazoline
502 derivatives. Lipids 1988;23:804-10.
- 503 57. **Spitzer V.** Structure analysis of fatty acids by gas chromatography--low resolution
504 electron impact mass spectrometry of their 4,4-dimethyloxazoline derivatives--a review. Prog
505 Lipid Res 1996;35:387-408.
- 506 58. **Harvey DJ.** Picolinyl esters as derivatives for the structural determination of long chain
507 branched and unsaturated fatty acids. Biomedical Mass Spectrometry 1982;9:33–38.
- 508 59. **Palleroni NJ (2005).** Genus I. *Pseudomonas* Migula 1894, 237AL (nom. cons., Opin.
509 5 of the Jud. Comm. 1952, 121). In Bergey's Manual of Systematic Bacteriology, 2nd edn.,
510 vol. 2B, pp. 323–379. Edited by Boone DR, Brenner DJ, Castenholz RW, Garrity GM, Krieg
511 NR, Staley JT. New York: Springer
- 512 60. **Chang DH, Rhee MS, Kim JS, Lee Y, Park MY, Kim H, et al.** *Pseudomonas*
513 *kribbensis* sp. nov., isolated from garden soils in Daejeon, Korea. Antonie Van Leeuwenhoek
514 2016;109:1433-1446.
- 515 61. **Jia J, Wang X, Deng P, Ma L, Baird SM, et al.** *Pseudomonas glycinae* sp. nov.
516 isolated from the soybean rhizosphere. Microbiology open 2020;9:e1101.
- 517 62. **Bligh EG, Dyer WJ.** A rapid method of total lipid extraction and purification. Can J
518 Biochem Physiol 1959;37:911-917.

- 519 63. **Tindall BJ, Sikorski J, Smibert RA, Krieg NR (2017).** Phenotypic Characterization
520 and the Principles of Comparative Systematics. In: Methods for General and Molecular
521 Microbiology. John Wiley & Sons, Ltd. pp. 330–393.
- 522 64. **Kwon SW, Kim JS, Park IC, Yoon SH, Park DH, et al.** *Pseudomonas koreensis* sp.
523 nov., *Pseudomonas umsongensis* sp. nov. and *Pseudomonas jinjuensis* sp. nov., novel species
524 from farm soils in Korea. Int J Syst Evol Microbiol 2003;53:21-27.

525 **Figure legends**

526 **Fig. 1.** Multilocus sequence analysis (MLSA) phylogenetic tree based on the 16S rRNA, *rpoB*,
527 *gyrB*, and *rpoD* concatenated sequences with an average length of 9,791 nucleotides showing
528 the position of strain FIT81^T among related species. Sequences were aligned by MUSCL
529 algorithm, and the phylogenetic tree was constructed with MEGA 7 software by using
530 Maximum Likelihood statistical method, Jukes-Cantor substitution method, and pairwise
531 deletion for the missing data. Bootstrap values >50% based on 1000 replications are indicated
532 on branches. Bar, 1 nt substitution per 100 nt.

533 **Fig. 2.** Strain FIT81^T Genome Blast Distance Phylogeny (GBDP) phylogenetic tree based on
534 genome data retrieved from the TYGS server. Branch lengths were calculated with formula D5
535 and the algorithm of greedy with trimming. Pseudo-bootstrap values >60% based on 100
536 replications are indicated on branches. Bar, 1 nt substitution per 100 nt.

537 **Fig. 3.** Strain FIT81^T uranyl acetate negative stain and image captured by transmission electron
538 microscopy (TEM).

539 **Fig. 4.** Two-dimensional thin-layer chromatography (TLC) of the polar lipids fraction of strain
540 FIT81^T. The position of phosphatidylethanolamine (PE), diphosphatidylglycerol (DPG),
541 phosphatidylglycerol (PG), glycolipid (GL), phospholipid (PL), and aminolipid (AL) is
542 indicated by arrows.

543 **Table 1.** Genome-to-genome computational relatedness of the top-five strain FIT81^T closest related species.

Species	Genome size (bp)	GC (%)	original ANI (%)	ortho ANI (%)	ANI _u (%)	ANI _b and [% aligned nucleotides]	ANI _m and [% aligned nucleotides]	dDDH and [% model CI]
strain FIT81^T	6.492.796	60.6	-	-	-	-	-	-
<i>P. gozinkensis</i> IzPS32d ^T	6.563.527	60.3	93.6	94.3	94.0	93.7 [83.2]	94.5 [84.8]	56.1% [53.4 - 58.8%]
<i>P. glycinae</i> MS586 ^T	6.396.728	60.5	93.6	94.0	93.8	93.2 [82.5]	94.4 [82.3]	54.9% [52.1 - 57.6%]
<i>P. allokribbensis</i> IzPS23 ^T	6.565.027	60.3	91.0	91.5	91.5	90.9 [82.5]	91.9 [82.6]	44.4% [41.9 - 47.0%]
' <i>P. kribbensis</i> ' 46-2	6.324.282	60.5	90.9	91.3	91.3	90.7 [81.5]	91.8 [81.3]	43.9% [41.4 - 46.5%]
<i>P. koreensis</i> PS9-14 ^T	6.123.913	60.5	87.5	87.9	87.5	86.9 [72.3]	89.0 [71.7]	34.7% [32.2 - 37.2%]

544

545 **Table 2.** Main cellular fatty acids detected as FAMES in strain FIT81^T and the closest related
546 species. Strains: 1, strain FIT81^T; 2, *P. gozinkensis* IzPS32d^T [20]; 3, *P. glycinae* MS586^T [61];
547 4, *P. allokribbensis* IzPS23^T [20]; 5, '*P. kribbensis*' 46-2 [60], and 6, *P. koreensis* Ps9-14^T [64].
548 Data for FIT81^T were obtained in this study whereas for other species were extracted from

Fatty acid	1	2	3	4	5	6
C _{10:0}	TR	ND	TR	ND	ND	ND
C _{10:0} 3-OH	4.1	4.2	6.6	4.1	5.4	2.2
C _{12:0} 2-OH	3.9	4.9	5.5	5.1	6.8	5.0
C _{12:0} 3-OH	4.4	5.0	6.7	5.4	7.5	4.0
C _{12:1} 3-OH	TR	ND	ND	ND	ND	ND
C _{12:0}	3.7	2	2.9	1.8	ND	TR
C _{14:0}	TR	TR	TR	TR	1.2	ND
C _{16:0}	30.6	32.4	22.9	31.9	33.4	33
C _{17:0} cyclo	3.8	17.7	10.3	16.1	15.1	2.0
C _{18:1} ω7c	14.4	ND	ND	ND	ND	ND
C _{18:0}	1.5	TR	TR	TR	1.6	TR
Summed feature 3 (C _{16:1} ω7c/C _{15:0} iso 2-OH)	30.3	18.3	23.6	19.8	16.8	37

549 literature.

550

551 **Table 3.** Biochemical features of strain FIT81^T and its related species.
552 Strains: 1, strain FIT81^T; 2, *P. gozinkensis* IzPS32d^T; 3, *P. glycinae* MS586^T; 4, *P.*
553 *allokribbensis* IzPS23^T; 5, '*P. kribbensis*' 46-2, and 6, *P. koreensis* Ps9-14^T. Data for FIT81^T,
554 *P. glycinae* MS586^T (= LMG 30275^T), and '*P. kribbensis*' 46-2 (= DSM 100278^T) were
555 obtained in this study whereas those for *P. gozinkensis* IzPS32d^T, *P. allokribbensis* IzPS23^T,
556 and *P. koreensis* Ps9-14^T were extracted from literature [20, 64]. Symbols legend are positive
557 (+), negative (-), weak or uncertain (w), variability in literature (-/+), and not determined (ND).

Characteristics	1	2	3	4	5	6
Fluorescence (fluorescein pigment)	+	+	+	+	-	+
Growth at 37°C	-	-	-	-	-	-/+
Growth on 4% NaCl	+	+	+	+	+	+
<i>Carbon source utilization:</i>						
D-Arabitol	+	-	+	-	+	+
D-Aspartic acid	-	-	-	-	w	+
D-Serine	-	+	+	-	w	ND
α-D-Glucose	+	+	+	+	+	+
D-Fructose-6-phosphate	w	-	w	-	+	ND
N-Acetyl-D-glucosamine	+	-	+	-	+	+
D-Galactose	+	w	+	w	+	+
Pectin	w	-	w	-	w	ND
Citrate	+	+	+	+	+	+
D-Fucose	w	-	w	-	+	ND
D-Mannose	+	+	+	+	+	+
Inosine	+	-	+	-	+	+
D-Galacturonic acid	w	-	w	-	+	-/+
L-Arabinose	+	+	+	+	+	+
Glycerol	+	+	+	ND	+	+
L-Galactonic acid lactone	w	-	w	-	+	+
D-Glucuronic acid	w	-	w	-	+	-/+
Glucuronamide	+	-	+	-	+	+
D-Malic acid	+	-	w	-	+	ND
Mucic acid	+	+	+	w	+	ND
Bromo-succinic acid	+	-	+	-	+	ND
α-Keto-butyric acid	w	-	w	-	w	-/+
Acetoacetic acid	w	-	w	-	w	ND
Quinic acid	+	+	+	+	+	+

D-Saccharic acid	+	+	+	-	+	+
Tween 40	+	w	+	-	+	+
<i>Enzymatic activity</i>						
Alkaline phosphatase	+	+	+	+	+	ND
C4 esterase	+	+	+	+	+	ND
C8 esterase lipase	+	+	+	+	+	ND
C14 lipase	+	ND	+	ND	+	ND
Leucin arylamidase	w	+	+	+	w	ND
Valine arylamidase	w	+	+	+	w	ND
Cystine arylamidase	w	ND	w	ND	w	ND
Trypsin	+	ND	+	ND	+	ND
α -Chymotrypsin	w	ND	w	ND	w	ND
Acid phosphatase	+	+	+	+	+	ND
Naphthol-AS-BI-phosphohydrolase	+	+	+	+	+	ND
Lysine decarboxylase	-	ND	-	ND	+	ND
Arginine dihydrolase	+	ND	+	ND	+	+
Ornithine decarboxylase	-	ND	-	ND	-	ND
Gelatin hydrolysis (protease)	+	ND	+	ND	+	+
Urease	+	ND	+	ND	+	ND
Esculin hydrolysis (β -glucosidase)	+	ND	+	ND	+	ND
β -Galactosidase (p-nitrophenyl- β -D-galactopyranosidase)	-	ND	-	ND	-	ND

558

***Pseudomonas fitomaticsae* sp. nov., isolated at Marimurtra Botanical Garden in Blanes, Catalonia-Spain**

Kostadin Evgeniev Atanasov^{1,2}, David Miñana Galbis³, Deborah Cornadó⁴, Annabel Serpico⁴, Guiomar Sánchez⁴, Montserrat Bosch⁴, Albert Ferrer^{1,5} and Teresa Altabella^{1,2}

¹Center for Research in Agricultural Genomics (CSIC-IRTA-UAB-UB), Bellaterra, Barcelona, Spain.

²Department of Biology, Healthcare and the Environment, Plant Physiology Section, Faculty of Pharmacy and Food Sciences, Universitat de Barcelona, Spain.

³Department of Biology, Healthcare and the Environment, Microbiology Section, Faculty of Pharmacy and Food Sciences, Universitat de Barcelona, Spain.

⁴Applied Microbiology and Biotechnology Unit, LEITAT Technological Center, Terrassa, Spain.

⁵Department of Biochemistry and Physiology, Faculty of Pharmacy and Food Sciences, Universitat de Barcelona, Spain.

Corresponding author: Kostadin E. Atanasov

The email address for the corresponding author: evgenievatanassov@ub.edu

List of Supplemental Figures and Tables

Supplemental Figure 1 (Fig. S1): Phylogenetic tree constructed on the 1,453 average 16S rRNA sequences length of the strain FIT81^T relatives. Sequence alignment was performed by using MUSCL algorithm and tree constructed by the statistical method of Maximum Likelihood, Jukes-Cantor substitution method, and pairwise deletion of gaps or missing data. Bootstrap test of phylogeny was performed with 1000 replications and values >50% indicated on branches. Bar, 1 nt substitution per 100 nt.

Supplemental Table 1 (Table S1): Species, strain designation, accession ID, and sequence identity for the 16S rRNA nucleotide sequence blast analysis of strain FIT81^T and its closest relative species.

Supplemental Table 2 (Table S2): Species, strain designation, assembly ID, and sequence identity for the MLSA nucleotide blast of strain FIT81^T and its closest relative species.

Supplemental Table 3 (Table S3): Genome mining of strain FIT81^T, *P. gozinkensis* IzPS32d^T; 3, *P. glycinae* MS586^T.

Supplemental Table 4 (Table S4): Identification of endophytic behavior genes: 1, strain FIT81^T; 2, *P. gozinkensis* IzPS32d^T; 3, *P. glycinae* MS586^T.

Fig. S1

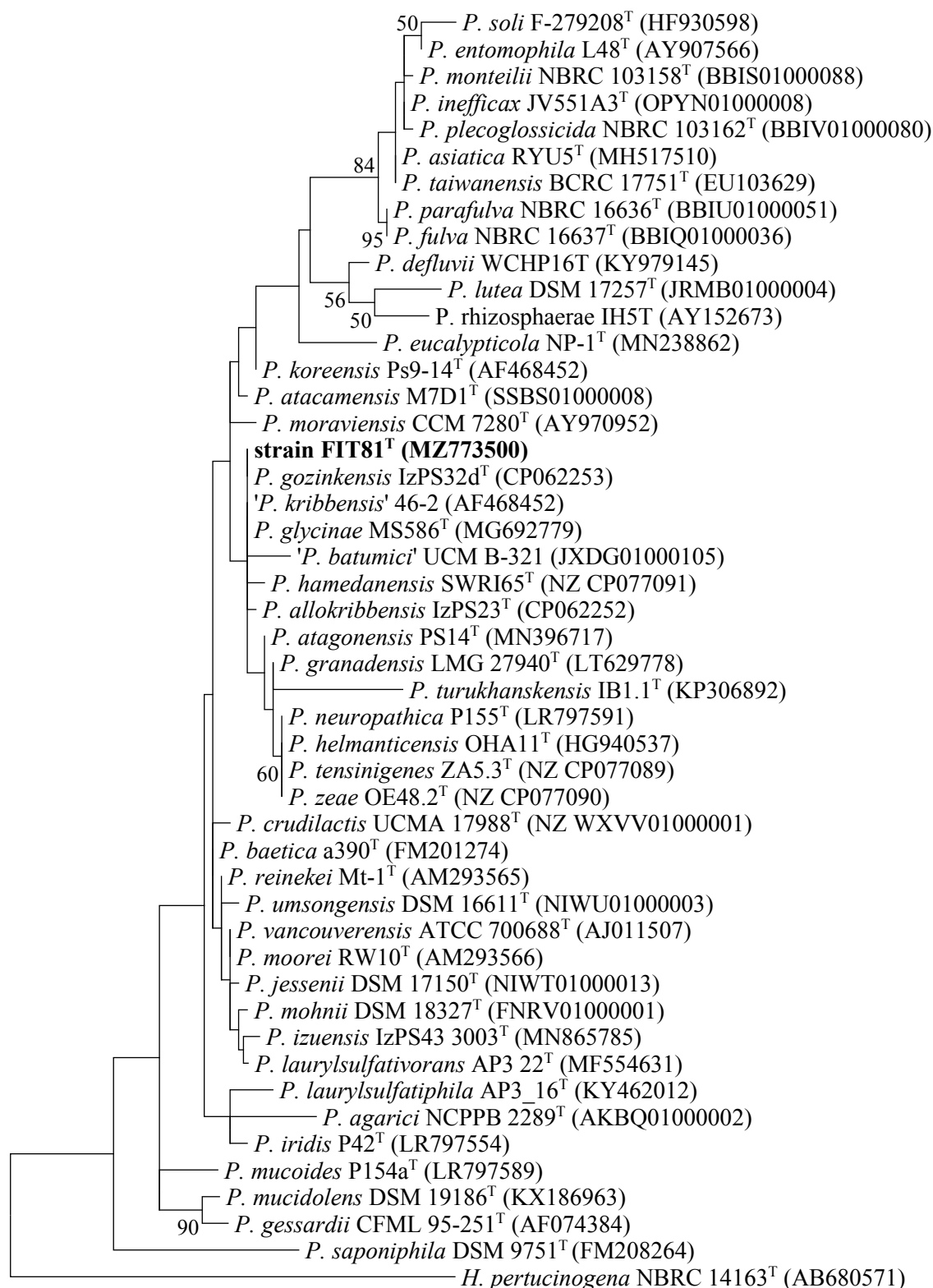


Table S1

Species	Strain	Accession ID	Identity	Cover	Length	E value
<i>'P. kribbensis'</i>	46-2	AF468452	99.9%	95%	1459	0.0
<i>P. gozinkensis</i>	IzPS32d ^T	CP062253	99.9%	100%	1533	0.0
<i>P. glycinae</i>	MS586 ^T	MG692779	99.9%	95%	1459	0.0
<i>P. allokribbensis</i>	IzPS23 ^T	CP062252	99.8%	100%	1533	0.0
<i>P. atacamensis</i>	M7D1 ^T	SSBS01000008	99.7%	95%	1459	0.0
<i>P. moraviensis</i>	CCM 7280 ^T	AY970952	99.6%	95%	1459	0.0
<i>P. granadensis</i>	LMG 27940 ^T	LT629778	99.6%	95%	1459	0.0
<i>P. koreensis</i>	Ps9-14 ^T	AF468452	99.6%	94%	1455	0.0
<i>P. reinekei</i>	Mt-1 ^T	AM293565	99.4%	95%	1459	0.0
<i>P. hamedanensis</i>	SWRI65 ^T	CP077091	99.4%	100%	1542	0.0
<i>P. vancouverensis</i>	ATCC 700688 ^T	AJ011507	99.3%	95%	1459	0.0
<i>P. jessenii</i>	DSM 17150 ^T	NIWT01000013	99.3%	95%	1459	0.0
<i>P. baetica</i>	a390 ^T	FM201274	99.2%	94%	1446	0.0
<i>P. neuropathica</i>	P155 ^T	LR797591	99.2%	80%	1232	0.0
<i>P. atagonensis</i>	PS14 ^T	MN396717	99.2%	95%	1460	0.0
<i>P. izuensis</i>	IzPS43_3003 ^T	MN865785	99.1%	95%	1459	0.0
<i>'P. batumici'</i>	UCM B-321	JXDG01000105	99.1%	95%	1459	0.0
<i>P. zeae</i>	OE48.2 ^T	CP077090	99.0%	100%	1542	0.0
<i>P. tensinigenes</i>	ZA5.3 ^T	CP077089	99.0%	100%	1542	0.0
<i>P. helmanticensis</i>	OHA11 ^T	HG940537	99.0%	95%	1459	0.0
<i>P. crudilactis</i>	UCMA 17988 ^T	WXVV01000001	99.0%	99%	1525	0.0
<i>P. laurylsulfativorans</i>	AP3_22 ^T	MF554631	98.8%	95%	1459	0.0

<i>P. umsongensis</i>	DSM 16611 ^T	NIWU01000003	98.8%	95%	1459	0.0
<i>P. mohnii</i>	DSM 18327 ^T	FNRV01000001	98.6%	95%	1459	0.0
<i>P. parafulva</i>	NBRC 16636 ^T	BBIU01000051	98.6%	95%	1459	0.0
<i>P. mucoides</i>	P154a ^T	LR797589	98.6%	80%	1232	0.0
<i>P. fulva</i>	NBRC 16637 ^T	BBIQ01000036	98.6%	95%	1459	0.0
<i>P. iridis</i>	P42 ^T	LR797554	98.5%	80%	1232	0.0
<i>P. laurylsulfatiphila</i>	AP3_16 ^T	KY462012	98.5%	95%	1460	0.0
<i>P. lutea</i>	DSM 17257 ^T	JRMB01000004	98.4%	95%	1459	0.0
<i>P. defluvii</i>	WCHP16 ^T	KY979145	98.3%	95%	1459	0.0
<i>P. asiatica</i>	RYU5 ^T	MH517510	98.3%	95%	1459	0.0
<i>P. taiwanensis</i>	BCRC 17751 ^T	EU103629	98.3%	94%	1451	0.0
<i>P. eucalypticola</i>	NP-1 ^T	MN238862	98.3%	92%	1423	0.0
<i>P. soli</i>	F-279,208 ^T	HF930598	98.2%	92%	1421	0.0
<i>P. turukhanskensis</i>	IB1.1 ^T	KP306892	98.2%	92%	1412	0.0
<i>P. gessardii</i>	CFML 95-251 ^T	AF074384	98.2%	95%	1459	0.0
<i>P. mucidolens</i>	DSM 19186 ^T	KX186963	98.2%	95%	1459	0.0
<i>P. inefficax</i>	JV551A3 ^T	OPYN01000008	98.2%	95%	1459	0.0
<i>P. agarici</i>	NCPPB 2289 ^T	AKBQ01000002	98.2%	95%	1459	0.0
<i>P. rhizosphaerae</i>	IH5 ^T	AY152673	98.2%	95%	1459	0.0
<i>P. plecoglossicida</i>	NBRC 103162 ^T	BBIV01000080	98.15%	95%	1459	0.0
<i>P. monteilii</i>	NBRC 103158 ^T	BBIS01000088	98.2%	95%	1459	0.0
<i>P. moorei</i>	RW10 ^T	AM293566	98.1%	95%	1459	0.0
<i>P. entomophila</i>	L48 ^T	AY907566	98.1%	95%	1459	0.0
<i>P. saponiphila</i>	DSM 9751 ^T	(FM208264)	96.5%	99%	1530	0.0

<i>H. pertucinogena</i>	NBRC 14163 ^T	(AB680571)	94.5%	95%	1459	0.0
-------------------------	-------------------------	------------	-------	-----	------	-----

Table S2

Species	Strain	Accession ID	Identity	Cover	Length	E value
<i>P. gozinkensis</i>	IzPS32d ^T	NZ_CP062253	98.2	99%	9873	0.0
<i>P. glycinae</i>	MS586 ^T	NZ_CP014205	98.2	100%	9799	0.0
' <i>P. kribbensis</i> '	46-2	NZ_CP029608	97.9	99%	9799	0.0
<i>P. allokribbensis</i>	IzPS23 ^T	NZ_CP062252	97.40	100%	9873	0.0
<i>P. koreensis</i>	Ps9-14 ^T	NZ_JAAQYM010000010	97.0	99%	9795	0.0
<i>P. atagonensis</i>	PS14 ^T	NZ_VXCA010000019	96.9	100%	9797	0.0
<i>P. atacamensis</i>	M7D1 ^T	NZ_SSBS010000010	96.5	100%	9799	0.0
<i>P. moraviensis</i>	CCM 7280 ^T	NZ_LT629788	96.4	100%	9799	0.0
<i>P. granadensis</i>	LMG 27940 ^T	NZ_LT629778	96.3	100%	9799	0.0
<i>P. baetica</i>	a390 ^T	NZ_PHHE010000001	96.1	99%	9786	0.0
<i>P. crudilactis</i>	UCMA 17988 ^T	NZ_WXVV010000001	96.1	100%	9865	0.0
<i>P. zeae</i>	OE48.2 ^T	NZ_CP077090	96.0	100%	9879	0.0
<i>P. tensinigenes</i>	ZA 5.3 ^T	NZ_CP077089	96.0	100%	9879	0.0
<i>P. neuropathica</i>	P155 ^T	NZ_JACOPX010000000	95.9	97%	9569	0.0
<i>P. helmanticensis</i>	OHA11 ^T	SAMN04488483	95.9	100%	9796	0.0
<i>P. hamedanensis</i>	SWRI65 ^T	NZ_CP012001	95.6	100%	9879	0.0
<i>P. izuensis</i>	IzPS43_3003 ^T	NZ_WTFT010000001	94.9	100%	9799	0.0
<i>P. vancouverensis</i>	ATCC 700688 ^T	NZ_RRZK010000005	94.9	99%	9799	0.0
<i>P. jessenii</i>	DSM 17150 ^T	NZ_FNTC010000002	94.8	100%	9799	0.0
<i>P. umsongensis</i>	DSM 16611 ^T	NZ_LT629767	94.7	100%	9799	0.0
<i>P. laurylsulfatiphila</i>	AP3_16 ^T	NZ_NIRS010000001	94.7	99%	9800	0.0
<i>P. laurylsulfativorans</i>	AP3_22 ^T	NZ_MUJK010000010	94.7	98%	9642	0.0

<i>P. iridis</i>	P42 ^T	NZ_JACOPU010000001	94.5	97%	9572	0.0
<i>P. reinekei</i>	Mt-1 ^T	NZ_LT629709	94.4	100%	9799	0.0
<i>P. moorei</i>	RW10 ^T	NZ_FNKJ01000003	94.2	100%	9799	0.0
<i>P. mohnii</i>	DSM 18327 ^T	NZ_FNRV01000002	94.2	100%	9799	0.0
<i>P. mucoides</i>	P154a ^T	NZ_JACOPW010000001	93.6	97%	9572	0.0
' <i>P. batumici</i> '	UCM B-321	NZ_JXDG01000001	93.6	100%	9799	0.0
<i>P. gessardii</i>	CFML 95-251 ^T	NZ_FNKR01000003	93.0	100%	9802	0.0
<i>P. saponiphila</i>	DSM 9751 ^T	NZ_FNTJ00000000	92.6	99%	9870	0.0
<i>P. mucidolens</i>	DSM 19186 ^T	NZ_LS483433	92.5	100%	9799	0.0
<i>P. agarici</i>	NCPPB 2289 ^T	NZ_FOAR01000077	91.4	99%	9799	0.0
<i>P. eucalypticola</i>	NP-1 ^T	NZ_CP056030	90.1	99%	9763	0.0
<i>P. rhizosphaerae</i>	IH5 ^T	NZ_CP009533	89.3	99%	9796	0.0
<i>P. plecoglossicida</i>	NBRC 103162 ^T	NZ_BBIV01000097	89.2	100%	9805	0.0
<i>P. entomophila</i>	L48 ^T	CT573326	89.1	99%	9805	0.0
<i>P. lutea</i>	DSM 17257 ^T	JRMB01000001	88.9	99%	9793	0.0
<i>P. inefficax</i>	JV551A3 ^T	NZ_OPYN01000001	88.9	100%	9805	0.0
<i>P. asiatica</i>	RYU5 ^T	NZ_BLJF01000001	88.9	100%	9805	0.0
<i>P. monteilii</i>	NBRC 103158 ^T	NZ_BBIS01000132	88.8	100%	9805	0.0
<i>P. soli</i>	F-279,208 ^T	NZ_CP009365	88.7	99%	9767	0.0
<i>P. taiwanensis</i>	BCRC 17751 ^T	NZ_KE384450	88.7	100%	9797	0.0
<i>P. parafulva</i>	NBRC 16636 ^T	NZ_BBIU01000055 NZ_BBIU01000000	- 88.3	100%	9802	0.0
<i>P. fulva</i>	NBRC 16637 ^T	LAWW01000001	88.0	100%	9802	0.0
<i>H. pertucinogena</i>	JCM 11590 ^T	NZ_BMNN00000000	84.1	86%	9814	0.0

Table S3

	Region	Type of metabolite	Position	Most similar known cluster	Cluster similarity	MIBiG comparison
strain FIT81 ^T	1	Arylpolyene	675,175 - 718,779	APE Vf	40%	Fragin (<i>Burkholderia cenocepacia</i> H111) Fellutamide B (<i>Aspergillus nidulans</i> FGSC A4) Showdomycin (<i>Streptomyces showdoensis</i>)
	2	RiPP-like	1,652,937 - 1,661,494	-	0	Duramycin (<i>Streptomyces cinnamoneus</i>) Butyriovibriocin AR10 (<i>Butyriovibrio fibrisolvens</i>) Citrulassin E (<i>Streptomyces glaucescens</i>)
	3	NRPS Terpene	2,318,719 - 2,394,454	Pyoverdin	20%	Sodorifen (<i>Serratia plymuthica</i> 4Rx13) 2-methylisoborneol (<i>Pseudanabaena</i> sp. dqh15) 2-methylisoborneol (<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350)
	4	RiPP-like	2,686,928 - 2,697,767	-	0	Lipopolysaccharide (<i>Xanthomonas campestris</i> pv. <i>campestris</i>) T3 toxin (<i>Cylindrospermopsis raciborskii</i> T3) Saxitoxin (<i>Cylindrospermopsis raciborskii</i> T3)

5	NRPS	2,811,220 - 2,886,660	Lokisin	71%	Gacamide A (<i>Pseudomonas fluorescens</i> Pf0-1) Bananamide (<i>Pseudomonas fluorescens</i>) Anikasin (<i>Pseudomonas fluorescens</i>)
6	NRPS	2,893,410 - 2,952,691	Rimosamide	28%	Massetolide A (<i>Pseudomonas fluorescens</i> SS101) Thalassospiramide A (<i>Tistrella bauzanensis</i>) Sevadecin (<i>Paenibacillus larvae</i>)
7	Siderophore	3,443,252 - 3,462,180	-	0%	Aerobactin (<i>Xenorhabdus szentirmaii</i> DSM 16338) Aerobactin (<i>Pantoea ananatis</i>) Ochrobactin (<i>Ochrobactrum anthropi</i>) Desferrioxamine (<i>Pantoea agglomerans</i>)
8	Betalactone	4,195,007 - 4,218,271	Fengycin	13%	Jadomycin (<i>Streptomyces venezuelae</i> ATCC 10712) Cepacin A (<i>Burkholderia ambifaria</i> IOP40-10) Bromopyrroles/bromophenols (<i>Marinomonas mediterranea</i> MMB-1)
9	Ranthipeptide	4,480,897 - 4,502,327	Pyoverdin	11%	Erythreapeptin (<i>Saccharopolyspora erythraea</i> NRRL 2338) Bicereucin (<i>Bacillus cereus</i> SJ1) FR-900098 (<i>Streptomyces rubellomurinus</i>)

<i>P. gozinkensis</i> IzPS32d ^T	10	NRPS	4,531,025 - 4,584,023	Pyoverdin	21%	Erythrochelin (<i>Saccharopolyspora erythraea</i> NRRL 2338) Anabaenopeptin NZ857/nostamide A (<i>Nostoc punctiforme</i> PCC 73102) monobactam (<i>Agrobacterium tumefaciens</i>)
	11	NAGGN	4,768,261 - 4,783,034	-	0%	Tabtoxin (<i>Pseudomonas syringae</i>) Bacilysin (<i>Bacillus</i> sp. CS93) 4-hydroxy-3-nitrosobenzamide (<i>Streptomyces murayamaensis</i>)
	12	Redox-cofactor	6,092,219 - 6,114,387	Lankacidin C	13%	Prodigiosin (<i>Serratia</i> sp.) Prodigiosin (<i>Serratia marcescens</i>) Prodigiosin (<i>Hahella chejuensis</i> KCTC 2396)
	1	NRPS-like	130.089 - 159.699	Fragin	37%	Fragin (<i>Burkholderia cenocepacia</i> H111) Livipeptin (<i>Streptomyces lividans</i> 1326) Aspergillic acid (<i>Aspergillus flavus</i> NRRL3357)
	2	Arylpolyene	488,336 - 531,964	APE Vf	40%	Aryl polyene (<i>Escherichia coli</i> CFT073) Aryl polyene (<i>Xenorhabdus doucetiae</i>) Caboxamycin (<i>Streptomyces</i> sp. NTK 937)
	3	RiPP-like	1,472,919 - 1,481,478	-	0%	Duramycin (<i>Streptomyces cinnamoneus</i>)

					Butyrivibriocin AR10 (<i>Butyrivibrio fibrisolvens</i>)
					Citrulassin E (<i>Streptomyces glaucescens</i>)
4	NRPS-like NRPS	2,103,351 - 2,187,903	Pyoverdin	20%	Livipeptin (<i>Streptomyces lividans</i> 1326)
					Fragin (<i>Burkholderia cenocepacia</i> H111)
					Fellutamide B (<i>Aspergillus nidulans</i> FGSC A4)
5	RiPP-like	2,415,999 - 2,426,838	-	0%	Lipopolysaccharide (<i>Xanthomonas campestris</i> pv. <i>campestris</i>)
					T3 toxin (<i>Cylindrospermopsis raciborskii</i> T3)
					Saxitoxin (<i>Cylindrospermopsis raciborskii</i> T3)
6	NRPS	2,664,065 - 2,733,716	Lokisin	71%	Gacamide A (<i>Pseudomonas fluorescens</i> Pf0-1)
					Bananamide (<i>Pseudomonas fluorescens</i>)
					Anikasin (<i>Pseudomonas fluorescens</i>)
7	NRPS	2,741,718 - 2,802,189	Rimosamide	21%	Massetolide A (<i>Pseudomonas fluorescens</i> SS101)
					Thalassospiramide A (<i>Tistrella bauzanensis</i>)
					Sevadycin (<i>Paenibacillus larvae</i>)
8	Siderophore	2,847,118 - 2,858,971	-	0%	Ethylenediaminesuccinic acid hydroxyarginine (EDHA) (<i>Streptomyces avermitilis</i> MA-4680 = NBRC 14893)

					Lipopolysaccharide (<i>Xanthomonas campestris</i> pv. <i>campestris</i>) streptonigrin (<i>Streptomyces flocculus</i>)
9	Thiopeptide	3,394,589 - 3,424,725	-	0%	Pseudopyronine (<i>Pseudomonas putida</i>) Pyrazomycin (<i>Streptomyces candidus</i>) Thaxteramide (<i>Jahnella</i> sp. MSr9139)
10	Betalactone	4,186,987 - 4,210,223	Fengycin	13%	Jadomycin (<i>Streptomyces venezuelae</i> ATCC 10712) Cepacin A (<i>Burkholderia ambifaria</i> IOP40-10) Bromopyrroles/bromophenols (<i>Marinomonas mediterranea</i> MMB-1)
11	Ranthipeptide	4,485,491 - 4,506,921	Pyoverdin	11%	FR-900098 (<i>Streptomyces rubellomurinus</i>) Succinoglycan (<i>Sinorhizobium meliloti</i>) Exopolysaccharides (<i>Burkholderia cenocepacia</i> J2315)
12	NRPS	4,539,277 - 4,592,278	Pyoverdin	20%	Erythrochelin (<i>Saccharopolyspora erythraea</i> NRRL 2338) Anabaenopeptin NZ857/nostamide A (<i>Nostoc punctiforme</i> PCC 73102) chloromyxamide (<i>Myxococcus</i> sp.)
13	NAGGN	4,747,766 - 4,762,638	-	0%	Tabtoxin (<i>Pseudomonas syringae</i>) Bacilysin (<i>Bacillus</i> sp. CS93)

<i>P. glyciniae</i> MS586 ^T					4-hydroxy-3-nitrosobenzamide (<i>Streptomyces murayamaensis</i>)	
	14	Redox-cofactor	5,936,419 - 5,958,575	Lankacidin C	13%	Prodigiosin (<i>Serratia sp.</i>) Prodigiosin (<i>Serratia marcescens</i>) Prodigiosin (<i>Hahella chejuensis</i> KCTC 2396)
	1	NAGGN	202,379 - 217,152	-	0%	Tabtoxin (<i>Pseudomonas syringae</i>) Bacilysin (<i>Bacillus sp.</i> CS93) 4-hydroxy-3-nitrosobenzamide (<i>Streptomyces murayamaensis</i>)
	2	NRPS	1,027,109 - 1,078,566	Pyoverdin	20%	Anabaenopeptin NZ857/nostamide A (<i>Nostoc punctiforme</i> PCC 73102) Monobactam (<i>Agrobacterium tumefaciens</i>) Aspirochlorine (<i>Aspergillus oryzae</i> RIB40)
	3	NRPS	2,076,149 - 2,134,645	Rimosamide	28%	Massetolide A (<i>Pseudomonas fluorescens</i> SS101) Sevadycin (<i>Paenibacillus larvae</i>) Curacomycin (<i>Streptomyces curacoi</i>)
	4	NRPS	2,291,136 - 2,358,877	Lokisin	71%	Bananamide (<i>Pseudomonas fluorescens</i>)

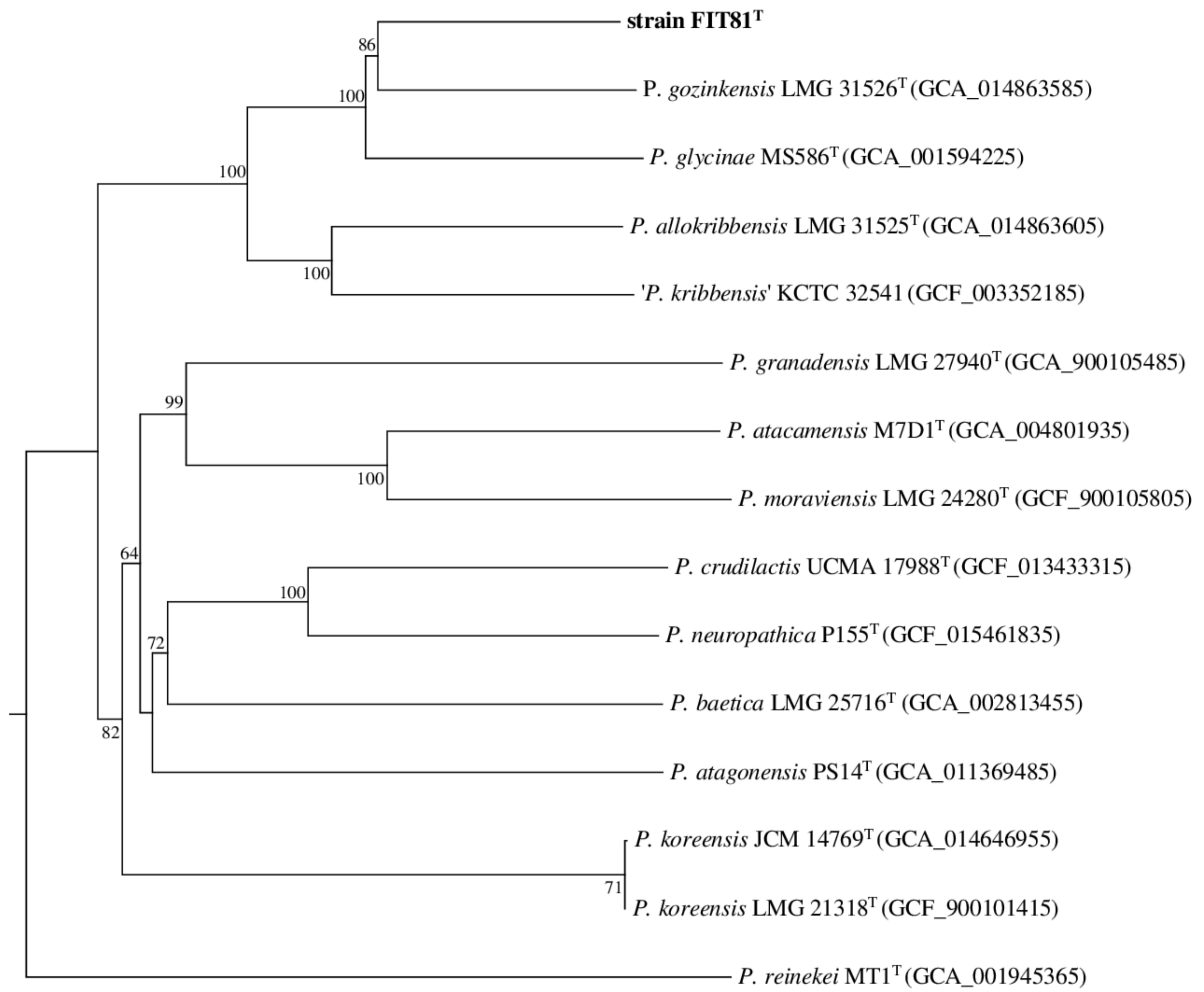
					Gacamide A (<i>Pseudomonas fluorescens</i> Pf0-1) Anikasin (<i>Pseudomonas fluorescens</i>)
5	RiPP-like	2,532,324 - 2,541,936	Bacillomycin D	20%	Lipopolysaccharide (<i>Xanthomonas campestris</i> pv. <i>campestris</i>) alkylresorcinol (<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350) Malleobactin (<i>Burkholderia thailandensis</i> E264)
6	Redox-cofactor	2,973,621 - 2,995,777	Lankacidin C	13%	Prodigiosin (<i>Serratia</i> sp.) Prodigiosin (<i>Serratia marcescens</i>) Prodigiosin (<i>Hahella chejuensis</i> KCTC 2396)
7	RiPP-like	4,555,244 - 4,566,131	-	0%	Duramycin (<i>Streptomyces cinnamoneus</i>) butyriovibriocin AR10 (<i>Butyriovibrio fibrisolvens</i>) citrulassin E (<i>Streptomyces Glaucescens</i>)
8	Betalactone	5,098,665 - 5,123,195	Fengycin	13%	Showdomycin (<i>Streptomyces showdoensis</i>) Jadomycin (<i>Streptomyces venezuelae</i> ATCC 10712) Cepacin A (<i>Burkholderia ambifaria</i> IOP40-10)
9	NRPS-like	5,205,531 - 5,235,101	Fragin	37%	Fragin (<i>Burkholderia cenocepacia</i> H111) Livipeptin (<i>Streptomyces lividans</i> 1326)

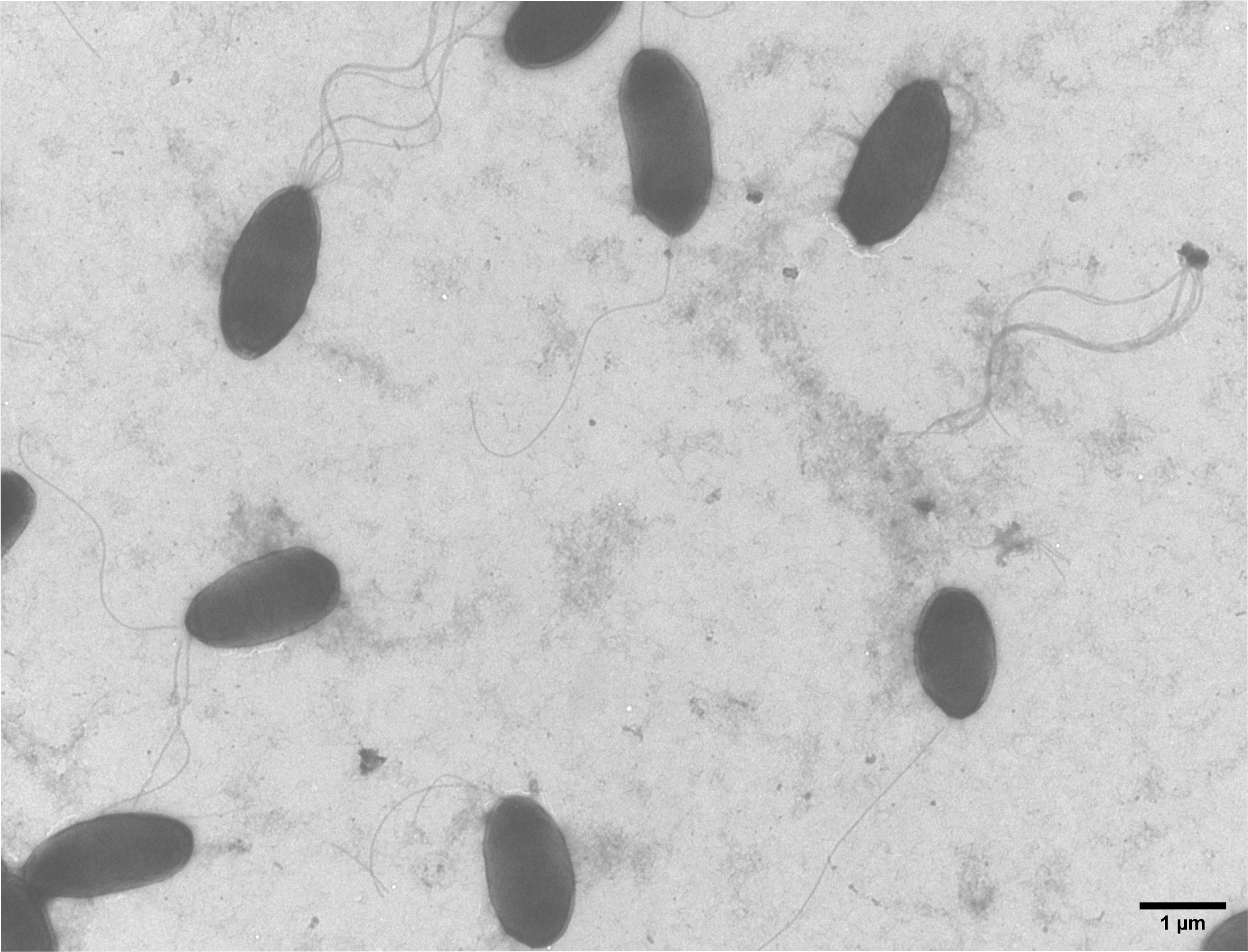
					Aspergillilic acid (<i>Aspergillus flavus</i> NRRL3357)
					Aryl polyene (<i>Escherichia coli</i> CFT073)
10	Arylpolyene	5,569,355 - 5,612,983	APE Vf	40%	Aryl polyene (<i>Xenorhabdus doucetiae</i>)
					Caboxamycin (<i>Streptomyces sp.</i> NTK 937)
					2-methylisoborneol (<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350)
11	NRPS-Terpene	6,231,816 -6,299,285	Pyoverdin	18%	Bacilysin (<i>Bacillus sp.</i> CS93)
					Pentostatine/vidarabine (<i>Streptomyces antibioticus</i>)

Table S4

		1	2	3
<i>Transport</i>	NAD(P)(+) transhydrogenase (AB-specific)	+	+	+
	Branched-chain amino acid ABC transporter	+	+	+
<i>Secretion and delivery system</i>	Type VI secretion ATPase, ClpV1	+	+	+
	Cupin	+	+	+
<i>Plant polymer degradation/modification</i>	Peptidase M48	+	+	+
	AsnC family transcriptional regulator	+	+	+
<i>Transcriptional regulator</i>	AraC family transcriptional regulator	+	+	+
	Transcriptional regulator, DeoR family	+	+	+
	Transcriptional regulator, LysR family	+	+	+
	Glutathione S-transferase	+	+	+
<i>Detoxification</i>	S-(hydroxymethyl) glutathione dehydrogenase/class III alcohol dehydrogenase	+	+	+
	2-dehydropantoate 2-reductase	+	+	+
	Acetaldehyde dehydrogenase	+	-	-
	Carbonate dehydratase	+	+	+
	Aldehyde dehydrogenase	+	+	+
<i>Redox potential maintenance</i>	L-lactate dehydrogenase	+	+	+
	Malate dehydrogenase	+	+	+
	3-hydroxyisobutyrate dehydrogenase	+	+	+
	ndvB	+	+	+
<i>Plant cell-wall attachment</i>	ndvC	+	+	+
	1-aminocyclopropane-1-carboxylate deaminase	+	+	+
<i>Hormones</i>	indole-3-glycerol phosphate synthase	+	+	+
	pqqA	+	+	+

	pqqB	+	+	+
	pqqC	+	+	+
	pqqD	+	+	+
	pqqE	+	+	+
	pqqF	+	+	+
<i>Other</i>	2-isopropylmalate synthase	+	+	+
	Diaminopimelate decarboxylase	+	+	+





1 μ m

