

# Northumbria Research Link

Citation: Castro, Jean Franco, Nouioui, Imen, Sangal, Vartul, Choi, Seonbin, Yang, Seung-Jo, Kim, Byung-Yong, Trujillo, Martha, Riesco, Raul, Montero-Calasanz, Maria del Carmen, Rahmani, Tara, Bull, Alan, Sutcliffe, Iain, Asenjo, Juan, Andrews, Barbara and Goodfellow, Michael (2018) *Blastococcus atacamensis* sp. nov., a novel strain adapted to life in the Yungay core region of the Atacama Desert. *International Journal of Systematic and Evolutionary Microbiology*, 68 (9). ISSN 1466-5026

Published by: Society for General Microbiology

URL: <http://dx.doi.org/10.1099/ijsem.0.002828>  
<<http://dx.doi.org/10.1099/ijsem.0.002828>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/34970/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

1 ***Blastococcus atacamensis* sp. nov., a novel strain adapted to life in the Yungay core**  
2 **region of the Atacama Desert**

3

4 Jean Franco Castro<sup>1,2</sup>, Imen Nouioui<sup>1</sup>, Vartul Sangal<sup>3</sup>, Seonbin Choi<sup>4</sup>, Seung-Jo Yang<sup>4</sup>,  
5 Byung-Yong Kim<sup>4</sup>, Martha E. Trujillo<sup>5</sup>, Raul Riesco<sup>4</sup>, Maria del Carmen Montero-Calasanz<sup>1</sup>,  
6 Tara PD Rahmani<sup>1</sup>, Alan T. Bull<sup>6</sup>, Iain C. Sutcliffe<sup>3</sup>, Juan A. Asenjo<sup>2</sup>, Barbara Andrews<sup>2</sup>,  
7 Michael Goodfellow<sup>1\*</sup>

8

9 <sup>1</sup> School of Natural and Environmental Sciences, Ridley Building 2, Newcastle University,  
10 Newcastle upon Tyne NE1 7RU, United Kingdom

11 <sup>2</sup> Centre for Biotechnology and Bioengineering (CeBiB), Department of Chemical  
12 Engineering and Biotechnology, University of Chile, Beauchef 851, Santiago, Chile

13 <sup>3</sup> Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, NE1  
14 8ST, United Kingdom

15 <sup>4</sup> ChunLab, Inc., 1, Gwanaka-ro, Gwanak-gu, Seoul 151015, South Korea

16 <sup>5</sup> Departamento de Microbiología y Genética, Universidad de Salamanca, Campus Miguel de  
17 Unamuno, 37007 Salamanca, Spain

18 <sup>6</sup> School of Biosciences, University of Kent, Canterbury CT2 7NJ, Kent, United Kingdom

19

20

21 \*To whom correspondence should be addressed: phone +44-(0)191 2087706; email:  
22 [m.goodfellow@ncl.ac.uk](mailto:m.goodfellow@ncl.ac.uk).

23

24 Running title: *Blastococcus atacamensis* sp. nov.

25

26 Subject category: New taxa: Actinobacteria

27

28 The GenBank accession numbers for the 16S rRNA gene and genome sequences of strain  
29 P6<sup>T</sup> are KX926540 and POQU00000000, respectively. The genome accession number of  
30 *Blastococcus saxobsidens* DSM 44509<sup>T</sup> is POQT00000000.

31

32 Abbreviations: A<sub>2</sub>pm, diaminopimelic acid; ANI, average nucleotide identity; dDDH, digital  
33 DNA–DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; NJ, neighbour-  
34 joining; ML, maximum-likelihood; MP, maximum-parsimony; MUSCLE, Multiple  
35 Sequence Comparison by Log-Expectation; T3PKS, type III polyketide synthase.

36

37 Keywords: polyphasic taxonomy, stress and biosynthetic genes, whole-genome sequences

38

### 39 **Abstract**

40 A polyphasic study was undertaken to establish the taxonomic status of a *Blastococcus* strain  
41 isolated from an extreme hyper-arid Atacama Desert soil. The isolate, strain P6<sup>T</sup>, was found  
42 to have chemotaxonomic and morphological properties consistent with its classification in  
43 the genus *Blastococcus*. It was shown to form a well-supported branch in the *Blastococcus*  
44 16S rRNA gene tree together with the type strains of *Blastococcus capsensis* and  
45 *Blastococcus saxobsidens* and was distinguished from the latter, its closest phylogenetic  
46 neighbour, by a broad range of phenotypic properties. The draft whole genome sequence of  
47 isolate P6<sup>T</sup> showed 83.6% average nucleotide identity, 83.0% average amino acid identity  
48 and a digital DNA:DNA hybridisation value of 27.8% in comparison with the genome  
49 sequence of *B. saxobsidens* DSM 44509<sup>T</sup>, values consistent with its assignment to a separate  
50 species. Based on these data it is proposed that isolate P6<sup>T</sup> (NCIMB 15090<sup>T</sup> = NRRL B-  
51 65468<sup>T</sup>) be assigned to the genus *Blastococcus* as *Blastococcus atacamensis* sp. nov.  
52 Analysis of the whole genome sequence of *B. atacamensis* P6<sup>T</sup>, with 3,778 open reading  
53 frames and a genome size of 3.9 Mb showed the presence of genes and gene clusters that  
54 encode for properties that reflect its adaptation to the extreme environmental conditions that  
55 prevail in Atacama Desert soils.

56

57 The actinobacterial genus *Blastococcus* was proposed by Ahrens and Moll [1] and the  
58 description subsequently emended by Urzì *et al.* [2], Lee [3] and Hezbri *et al.* [4],  
59 respectively. The genus *Blastococcus* together with the genera *Cumulibacter* [5],  
60 *Geodermatophilus* [6], *Klenkia* [7] and *Modestobacter* [8] belong to the family  
61 *Geodermatophilaceae* [9, 10] of the order *Geodermatophilales* [11]. Members of all of these  
62 taxa share genomic features, as exemplified by multiple copies of the *trwC* gene (conjugative

63 relaxase) [12], have modest growth requirements [13], show unusual resistance to oxidative  
64 stress [14] and tend to be associated with arid biomes, such as desert and high altitude soils  
65 and with the surfaces of ancient monuments and natural stones [2, 15, 16].

66 Blastococci form a well-supported clade in the *Geodermatophilaceae* 16S rRNA gene tree  
67 [4, 17] and can be distinguished from members of other genera classified in this family using  
68 a combination of phenotypic features [13]. They are Gram-stain positive, coccoid-shaped  
69 bacteria that may be motile or non-motile and which may propagate by budding and multiple  
70 fission; they have *meso*-A<sub>2</sub>pm in the peptidoglycan, mainly unsaturated and iso-branched  
71 fatty acids; and complex phospholipid profiles which may include diphosphatidylglycerol,  
72 phosphatidylcholine, phosphatidylethanolamine and phosphatidylinositol [4, 18]. At the time  
73 of writing the genus encompasses 5 validly named species, namely *Blastococcus aggregatus*  
74 [1, 4, 19], the type species, *Blastococcus capsensis* [4], *Blastococcus endophyticus* [4, 17],  
75 *Blastococcus jejunensis* [3, 4], and *Blastococcus saxobsidens* [4, 19], and one validly named  
76 strain, *Blastococcus colisei* [20]. These bacteria were isolated from the Baltic Sea, an  
77 archaeological Roman pool in Tunisia, the leaves of the medicinal Chinese plant  
78 *Camptotheca acuminata*, beach sand off the coast of South Korea, a limestone sample from  
79 a church in Malta and an archaeological amphitheatre, respectively, and can be distinguished  
80 using a range of phenotypic properties [4]. In addition, “*Candidatus Blastococcus*  
81 *massiliensis*” was identified, from a stool sample of a patient with anorexia nervosa [21].

82 In a continuation of our studies on actinobacterial diversity in Atacama Desert habitats,  
83 several strains were recovered from an extreme hyper-arid soil that had colonial and  
84 morphological properties typical of blastococci. One of these isolates, strain P6<sup>T</sup>, was the  
85 subject of a polyphasic taxonomic study which showed that it represents a new *Blastococcus*  
86 species, for which the name *Blastococcus atacamensis* sp. nov. is proposed.

87

88 *Blastococcus* strains were isolated from an extreme hyper-arid soil sample in November 2011  
89 from the Yungay core region of the Atacama Desert on the eastern flank of the Cerro Aguas  
90 Blancas (24°06'18.6"S/70°01'55.6W) at 1,033 metres above sea level. One gram of the soil  
91 sample was suspended in 4.0 ml of ¼ strength Ringer's solution (Oxoid, product No.  
92 BO0332D), this suspension was shaken on a tumble shaker prior to heating at 55°C for six  
93 minutes. Aliquots (100µl) of the 10<sup>-1</sup> and 10<sup>-2</sup> dilutions were spread, in triplicate, over GYM

94 *Streptomyces* (DSMZ medium No. 65) and *Geodermatophilus obscurus* media [22]  
95 supplemented with nalidixic acid ( $10\mu\text{g}\cdot\text{ml}^{-1}$ ), cycloheximide and nystatin (each at  $25\mu\text{g}\cdot\text{ml}^{-1}$ ).  
96 The isolation plates were dried for 15 minutes at room temperature before incubation, as  
97 recommended by Vickers and Williams [23]. After incubation at  $28^{\circ}\text{C}$  for 2 weeks, the  
98 presumptive *Blastococcus* isolates were counted and expressed as the number of colony  
99 forming units (cfu) per gram dry weight soil.

100 Small numbers of strains growing on the isolation plates were assigned to the genus  
101 *Blastococcus* as they formed characteristic small, circular, reddish pink colonies with entire  
102 margins; the highest count,  $3.7\cdot 10^3$  cfu/g dry weight soil, was recorded from the *G. obscurus*  
103 medium plates and the corresponding number on the GYM *Streptomyces* medium plates was  
104  $2.7\cdot 10^3$  cfu/g dry weight soil. A representative *Blastococcus* strain, isolate P6<sup>T</sup>, was taken  
105 from one of the GYM *Streptomyces* plates and along with the type strains of *Blastococcus*  
106 species was maintained on GYM slopes at room temperature and as suspensions of cells in  
107 20%, v/v glycerol at  $-20^{\circ}\text{C}$  and  $-80^{\circ}\text{C}$ .

108 The isolate was examined for motility and morphological properties using procedures  
109 described by Trujillo *et al.* [24]. Cultural features were recorded on modified Bennett's  
110 (DSMZ medium No. 894), GPHF (DSMZ medium No. 553), GYM *Streptomyces* (DSMZ  
111 medium No. 65), Luedemann's [6], potato dextrose (DSMZ medium No. 129), Reasoner's  
112 2A (DSMZ medium No. 830) and from tryptone-yeast extract, yeast extract-malt extract,  
113 oatmeal, inorganic salts-starch, glycerol-asparagine, peptone-yeast extract-iron and tyrosine  
114 agar (International *Streptomyces* Project [ISP] media 1–7; [25]) plates following incubation  
115 at  $28^{\circ}\text{C}$  for 3 weeks. The isolate was found to be a Gram-stain positive, non-motile  
116 actinobacterium that formed rods and coccoid shaped cells with evidence of budding (Fig.  
117 1). Strain P6<sup>T</sup> was observed to grow well on GYM *Streptomyces*, GPHF, modified Bennett's,  
118 potato dextrose, Luedemann's, Reasoner's 2A and yeast extract-malt extract agar, as  
119 exemplified in Figure S1, but poorly on ISP media 1, 3 to 7, generally producing red-orange  
120 or yellowish pink pigments; diffusible pigments were not observed on any of these media.

121 Biomass for most of the chemotaxonomic analyses carried out on isolate P6<sup>T</sup> was harvested  
122 from 1,000 ml yeast extract-malt extract broth ISP medium 2 that had been shaken in 500 ml  
123 baffled Erlenmeyer flasks, each flask containing 200 ml of medium, at 180 revolutions per  
124 minute (rpm) at  $28^{\circ}\text{C}$  for 2 weeks; the biomass was washed twice in distilled water and

125 freeze-dried. Biomass for the fatty acid analysis was prepared on PYGV agar (DSMZ  
126 medium No. 621), modified by the inclusion of 2 g of peptone instead of casein, 2 g of yeast  
127 extract and 10 ml of a 20% w/v glucose, after incubation at 20°C for 16 days and washed  
128 twice in sterile distilled water.

129 Isolate P6<sup>T</sup> was examined for chemotaxonomic markers known to be of value in the  
130 systematics of microorganisms classified in the genus *Blastococcus* [4, 18]. Standard  
131 chromatographic procedures were used to determine the isomers of A<sub>2</sub>pm [26], isoprenoid  
132 quinones [27], whole-cell sugars [28] and polar lipids [29], as modified by Kroppenstedt and  
133 Goodfellow [30]. Isoprenoid quinones extracted from *Micromonospora luteifusca* GUI2<sup>T</sup>  
134 [31] were used as standards. In turn, fatty acids extracted from the isolate were methylated,  
135 analysed using the protocol of the Sherlock Microbial Identification (MIDI) system, version  
136 5 [32] and the resultant peaks identified using the ACTIN 6 database.

137 In general, the chemotaxonomic properties of isolate P6<sup>T</sup> are consistent with its classification  
138 in the genus *Blastococcus* [4, 18, 20]. The organism contains *meso*-A<sub>2</sub>pm as the diagnostic  
139 diamino acid (Fig. S2); MK-9(H<sub>2</sub>) and MK-9(H<sub>4</sub>) as predominant isoprenologues in a  
140 proportion of 3:2 (Fig. S3); iso-C<sub>16:0</sub> (38.9%), iso-C<sub>16:1</sub> H (17.7%), iso-C<sub>15:0</sub> (14.2%) and 9-  
141 methyl-C<sub>16:0</sub> (5.7%) as major fatty acids; a polar lipid profile that includes  
142 diphosphatidylglycerol, phosphatidylcholine, phosphatidylethanolamine,  
143 phosphatidylglycerol and phosphatidylinositol (Fig. S4) and galactose and glucose as major  
144 sugars with a lesser proportion of ribose and traces of arabinose and xylose (Fig. S5).

145 Genomic DNA for 16S rRNA gene sequencing was extracted from 25 ml ISP 2 broth  
146 incubated at 28°C, shaken at 180 rpm for 10 days; 5 ml of the fully grown culture were used  
147 for genomic DNA extraction, following the protocol used by Kieser *et al.* [33] though in this  
148 case incubation with protease K was conducted at 60°C until the solution became clear  
149 (~1.5h); the quality of the isolated genomic DNA was checked in a 1%, w/v agarose gel run  
150 at 70V for 1.5h. PCR-mediated amplification of the 16S rRNA gene was performed in a final  
151 volume of 25 µl using the standard primers 27F and 1525r [34], 100 ng of genomic DNA and  
152 MyFi™ Mix (Bioline, UK), following the manufacturer's instructions; the PCR conditions  
153 were 5min at 95°C followed by 30 cycles of 30sec at 95°C, 30sec at 55°C and 23sec at 72°C.  
154 Two µl of the resulting PCR mixture was passed through a 1%, w/v agarose gel from which  
155 a single band of the expected size (about 1,500 bp) was visualised. The rest of the PCR

156 mixture was cleaned with exonuclease I and shrimp alkaline phosphatase (NEB, UK;  
157 #E2622S) and sent to Geneius (Cramlington, UK) for sequencing, using a BigDye®  
158 terminator v3.1 cycle sequencing kit (Thermo Fisher Scientific Inc.) on an ABI sequencer  
159 model 3730xl; the sequence was assembled using Pregap4 and Gap4 from Staden Package  
160 version 2.0.0b9 [35].

161

162 The fragment of the 16S rRNA gene sequence of isolate P6<sup>T</sup> (1,340 bp) was compared with  
163 corresponding 16S rRNA gene sequences of the type strains of the *Blastococcus* species, the  
164 sequence of “*Candidatus B. massiliensis*” AP3 and those of *Cumulibacter manganitolerans*  
165 DSM 103787<sup>T</sup>, *G. obscurus* DSM 43160<sup>T</sup>, ‘*Klenkia marina*’ DSM 45722<sup>T</sup> and  
166 *Modestobacter multiseptatus* DSM 44406<sup>T</sup> all of which were retrieved from the EzBioCloud  
167 server (<http://www.ezbiocloud.net/>; [36]). Alignment of the nucleotide sequences was  
168 achieved with MUSCLE [37]. Phylogenetic trees were inferred with the MEGA7 suite of  
169 programs version 7.0 [38] using the NJ [39], ML [40] and MP algorithms [41] with 1,000  
170 bootstrap repetitions [42]; evolutionary distances were calculated with the Kimura 2-  
171 parameters model [43]. Sequence similarity values were calculated based on the alignment  
172 generated with MUSCLE [37], using PHYDIT software version 3.0. In addition, pairwise  
173 sequence similarities and phylogenetic reconstruction were performed using the method  
174 recommended by Meier-Kolthoff *et al.* [44] and the GGDC web server [45] available at  
175 <http://ggdc.dsmz.de/phylogeny-service.php#> and the Genome-to-Genome Distance  
176 Calculator (GGDC; [44]) to validate the results obtained using the MEGA7 software. The  
177 resultant trees were rooted with *G. obscurus* DSM 43160<sup>T</sup>, *K. marina* YIM M1315 and *M.*  
178 *multiseptatus* DSM 44406<sup>T</sup> using FigTree version 1.4.2  
179 (<http://tree.bio.ed.ac.uk/software/figtree/>). All of the Figures were edited in Inkscape version  
180 0.9 (<https://inkscape.org/en/download/>).

181 Strain P6<sup>T</sup> was found to form a well-supported subclade in the *Blastococcus* 16S rRNA gene  
182 tree together with the type strains of *B. capsensis* and *B. saxobsidens* (Fig. 2). It is most  
183 closely related to type strain of the *B. saxobsidens* sharing a 16S rRNA gene similarity of  
184 99.5%, a value that corresponds to 7 nucleotide (nt) differences at 1,339 locations. The  
185 corresponding 16S rRNA gene sequence similarity value between P6<sup>T</sup> and *B. capsensis* BMG  
186 804<sup>T</sup> was 99.1%, which equates to 10 nt differences at 1,241 sites. The 16S rRNA gene

187 sequence similarities between strain P6<sup>T</sup> and the other *Blastococcus* type strains fell within  
188 the range 97.7 and 98.1%, values corresponding to between 26 and 31 nt differences. The  
189 same pairwise similarity values were recorded between the isolate and the type strains of *B.*  
190 *capsensis* and *B. saxobsidens* using the GGDC server.

191 Isolate P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> were examined for a broad range of phenotypic  
192 properties. Enzyme profiles were determined using API ZYM strips (BioMérieux) by  
193 following the manufacturer's instructions, while GEN III microplates (Biolog Inc., Hayward,  
194 CA, USA) were used to test for the ability of the strains to oxidise carbon and nitrogen  
195 sources and to determine resistance to inhibitory compounds using inoculating fluid (IF-C,  
196 Biolog Inc.) and a cell density of 86% transmittance in an OmniLog instrument (Biolog Inc.)  
197 set at 28°C. Data from the triplicated cultures recorded in Phenotypic Mode from the GENIII  
198 microplates were analysed using opm package 1.0.6 [46] for R [47], using R studio [48].  
199 Many of the remaining tests were carried out using ISP 2 agar [25] as the basal medium. The  
200 strains were examined for their ability to grow at a range of pH values (pH 5–10 at single  
201 unit intervals; pH was adjusted by adding drops of either 1N NaOH or 1N HCl as described  
202 by Montero-Calasanz *et al.* [49]) and temperatures (4, 10, 20, 28, 37, 40, 45 and 50°C) and  
203 in the presence of various concentrations of sodium chloride (1.0, 1.5, 3.0, 5.0, 7.0, 15.0 and  
204 20%, w/v). Apart from the temperature tests these features were recorded after incubation at  
205 28°C. Results of all of these tests were recorded after incubation for 3 weeks. The ISP 2  
206 medium was also used to test the capacity of the strains to degrade casein (1%), cellulose  
207 (1%), elastin (0.3%), guanine (0.5%), hypoxanthine (0.4%), L-tyrosine (0.4%), uric acid  
208 (0.5%) and xanthine (0.4%), their ability to degrade tributyrin was determined using  
209 tributyrin agar (Sigma-Aldrich). Results of all of these tests were recorded after incubation  
210 at 28°C for 14 days. Aesculin (0.1%) and arbutin (0.1%) hydrolysis was established using the  
211 media and methods described by Williams *et al.* [50], the hydrolysis of urea (0.2%, w/v) after  
212 Christensen [51] and nitrate reduction following Schaal *et al.* [52]. Catalase production was  
213 detected by the formation of bubbles after mixing a fresh drop of 3% hydrogen peroxide to  
214 fresh growth of the cultures on glass slides. Oxidase activity was determined in a 1%, w/v  
215 solution of *N-N-N'-N'*-tetramethyl-1,4-phenyldiamine (Sigma-Aldrich) and the development  
216 of a blue purple colour was recorded as a positive result [53]. The degradation and tolerance



217 tests were carried out in triplicate using a cell suspension equivalent to 5.0 on the McFarland  
218 scale [54].

219 The triplicated tests on isolate P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> gave identical results for  
220 all of the phenotypic tests, apart from the ability of isolate P6<sup>T</sup> to degrade arbutin.

221 The phenotypic properties of isolate P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> strains were  
222 compared with those of the other type strains of *Blastococcus* species, which had mainly been  
223 examined using the same procedures, as exemplified by the API, GENIII microplate,  
224 tolerance and catalase tests [4, 20]. It can be seen from Table 1 that the isolate can be  
225 distinguished from all of the *Blastococcus* type strains, including *B. capsensis* DSM 46835<sup>T</sup>  
226 and *B. saxobsidens* DSM 44509<sup>T</sup>, its nearest phylogenetic neighbours, using a combination  
227 of phenotypic features. It can be distinguished from each of these organisms by its ability to  
228 grow at 10 and 45°C, to reduce nitrate to nitrite and use dextrin,  $\alpha$ -keto-butyric acid and D-  
229 malic acid, and by its inability to use glucuronamide and D-saccharic acid as sole carbon  
230 sources. It can also be separated from the *B. saxobsidens* type strain by its capacity to oxidise  
231 pectin and D-salicin. It is also evident from Table 1 that several chemotaxonomic features  
232 support the separation of isolate P6<sup>T</sup> from the type strains of *Blastococcus* species. Isolate  
233 P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> were also found to share many phenotypic features:  
234 they produce acid and alkaline phosphatases,  $\alpha$ -chymotrypsin, cysteine arylamidase, esterase  
235 (C4), esterase lipase (C8),  $\alpha$ -glucosidase, naphthol-AS-B1-phosphohydrolase and valine  
236 arylamidase; degraded aesculin, tributyrin and starch; but do not produce lipase (C14),  $\alpha$ -  
237 mannosidase, or  $\alpha$ -fucosidase or hydrolyse allantoin and urea, or degrade casein, cellulose,  
238 elastin, guanine, hypoxanthine, L-tyrosine, uric acid, or xanthine. In addition, they can utilise  
239  $\alpha$ - and  $\beta$ -hydroxy-butyric acid, D-cellobiose, D-fucose,  $\beta$ -gentiobiose, D-gluconic acid,  $\alpha$ -D-  
240 lactose, D-salicin, sucrose and D-turanose but not L-alanine, D-arabitol, D-aspartic acid,  
241 glycyl-proline, gelatin, mucic acid or *p*-hydroxy-phenylacetic acid.

242 Biomass for sequencing the whole-genome of isolate P6<sup>T</sup> was prepared in a 1.5 ml of brain-  
243 heart infusion broth at 28°C in a shaking incubator (180 rpm) for 2 days. Genomic DNA of  
244 strains was extracted using the QIAamp DNA extraction kit (Qiagen, USA) according to the  
245 manufacturer's instructions. The purity and concentration of the extracted genomic DNA  
246 were measured using the Nanodrop spectrophotometer (NanoDrop Technologies, UK).  
247 Genome sequencing was performed on an Illumina MiSeq instrument (Illumina); the reads

248 were assembled by using SPAdes 3.9.0 [55] and contigs smaller than 1,000 bp in size were  
249 discarded. The draft assemblies have been submitted to the GenBank (accession numbers:  
250 POQU000000000 and POQT000000000) and is publicly available.

251 The genomes were annotated using the RAST annotation pipeline [56] and a sequenced based  
252 comparison performed in the SEED Viewer [56, 57]. A digital DNA:DNA hybridisation  
253 (dDDH) value was calculated between the genomes of strain P6<sup>T</sup> and *B. saxobsidens* DSM  
254 44509<sup>T</sup> using the GGDC server [45]. BLAST-based ANI and AAI between the strains were  
255 calculated using the online resource from the K. Konstantinidis group (<http://enve-omics.ce.gatech.edu/>; [58]).

257 The draft genomes of strain P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup>, contained 3,778 and 4,348  
258 open reading frames, respectively, and were 3.9 Mb and 4.5 Mb in size with average *in silico*  
259 DNA G+C contents of 73.1 and 74.3 mol%. The dDDH value between the genome of the  
260 two strains was 27.8% (C. I. 25.4-30.3%), which is well below the 70% threshold used to  
261 confirm the species status of novel strains [59]. The corresponding ANI and AAI indices  
262 were 84.6± 4.5 and 83.0± 13.0, values below the threshold used for prokaryotic species  
263 delineation [58, 60, 61].

264 The draft genomes of isolates P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> were examined using the  
265 antiSMASH server [62] to detect putative biosynthetic gene clusters. The genome of isolate  
266 P6<sup>T</sup> was found to encode for a T3PKS and corresponding residues that make up the catalytic  
267 triad found in RppA, a T3PKS involved in the biosynthesis of pentaketide 1,3,6,8-  
268 tetrahydroxynaphthalene in *Streptomyces griseus* [63, 64]. The T3PKS of isolate P6<sup>T</sup> showed  
269 94% sequence identity with a corresponding sequence detected in the genome of the *B.*  
270 *saxobsidens* strain and 83% identity with a putative T3PKS encoded in the genome of *G.*  
271 *obscurus* DSM 43160<sup>T</sup> (Gobs\_4821; UniProt [65], accession number: D2S5V1). The gene  
272 that encodes for the T3PKS of isolate P6<sup>T</sup> was surrounded by other biosynthetic genes, such  
273 as one encoding for a methyltransferase and others encoding regulatory and transport  
274 proteins, thereby suggesting the presence of a biosynthetic gene cluster though the  
275 functionality and product generated by this putative biocluster has still to be established. The  
276 genomes of isolate P6<sup>T</sup> and the *B. saxobsidens* strain were also found to harbour genes  
277 encoding for polyprenyl synthetase and phytoene synthase, enzymes involved in the  
278 biosynthesis of terpenoid compounds [66, 67]. The genome of the *B. saxobsidens* strain

279 contains two genes that encode for proteins that contain the IucA/IucC domain (Pfam [68]  
280 accession: pfam04183) which is involved in the biosynthesis of siderophore compounds [69].  
281 The genome of the type strain of *Modestobacter caceserii* which, like P6<sup>T</sup>, was isolated from  
282 an extreme hyper-arid soil sample collected from the Yungay core region of the Atacama  
283 Desert, contained a siderophore gene cluster predicted to encode for deferoxamine; the  
284 genome of this organism also contained gene clusters encoding for type II and III polyketides  
285 and terpenes [70]. These preliminary datasets suggest that the genomes of  
286 *Geodermatophilaceae* strains have the capacity to produce specialised metabolites such as  
287 polyketides and siderophores. However, antiSMASH does not necessarily detect all of the  
288 gene clusters in genomes, as exemplified by the failure to identify the hygromycin A gene  
289 cluster in *Streptomyces leewenhoekii* C34<sup>T</sup> [71], moreover an improved genome assembly is  
290 required for a more precise interpretation of predicted biosynthesis gene clusters in the  
291 genome of isolate P6<sup>T</sup>.

292

293 A comparison of the genomes of isolate P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> showed that  
294 the genome of the former contains 474 genes that are absent from the genome of the latter,  
295 including those involved in stress responses (2 copies of *terA* and 3 copies of *terD* genes; P6-  
296 peg 850-P6-peg 855) [72]. In contrast, the genome of the *B. saxobsidens* strain harbours  
297 2,848 genes that are absent from isolate P6<sup>T</sup> (data not shown); these genes include multiple  
298 copies of the *tetA*, *tetB*, and *tetC* genes that are involved in tricarboxylate/citrate transport.  
299 However, most of the unique genes (54-66%) found in the genomes of these strains encode  
300 for hypothetical proteins.

301 The genomes of isolate P6<sup>T</sup> and the *B. saxobsidens* DSM 44509<sup>T</sup> contained 119 and 147  
302 genes, respectively that are associated with stress responses (Table S1). Each of the strains  
303 contained two genes involved in carbon starvation, one encoding for carbon starvation  
304 protein A and the other for a carbon storage regulatory protein indicating that they are adapted  
305 to life in low carbon environments [73-75]. Similarly, four genes belonging to the CspA  
306 family associated with responses to cold-shock [76], 13 genes of the *dnaK* gene cluster that  
307 respond to heat shock [77] and seven genes associated with the biosynthesis, uptake and  
308 utilisation of trehalose, which are considered to help in responses to heat and desiccation  
309 stress [78], are conserved in each of the strains. The genomes of the strains also contain genes

310 belonging to the *uvrABC* DNA repair system that assists in UV resistance [79], as well as  
311 Rec proteins (RecA, RecX and those involved in the RecBCD and RecFOR pathways) that  
312 are responsible for stabilising genomes [80]. The P6<sup>T</sup> genome contains a *coxGMLS* gene  
313 cluster and a *coxD* gene, whereas the *B. saxobsidens* strain has five copies of the *coxM* gene  
314 and two copies of the *coxS* gene, though the gene encoding for the *coxG* protein is absent;  
315 *cox* genes code for the utilisation of carbon monoxide, thereby indicating that these  
316 organisms may have a chemolithoautotrophic lifestyle [81].

317 The proteins involved in responses to oxidative stress are, with minor exceptions, conserved  
318 in the genomes of isolate P6<sup>T</sup> and *B. saxobsidens* DSM 44509<sup>T</sup> (Table S1). The P6<sup>T</sup> genome  
319 contains eight genes involved in responses to osmotic stress, notably choline dehydrogenase,  
320 and ABC transporter proteins for betaine, glycine and L-proline uptake and a high affinity  
321 choline uptake protein (*betT*) [82, 83]. In turn, only the *B. saxobsidens* strain has a *sox* gene  
322 cluster encoding the subunit of sarcosine oxidase, along with additional copies of *betT* and  
323 transport proteins. Sarcosine oxidase is associated with responses to osmotic stress [84, 85].

324

325 It is clear from the chemotaxonomic, genomic, morphological and phylogenetic data that  
326 isolate P6<sup>T</sup> is a *bona fide* member of the genus *Blastococcus*. It can be distinguished from the  
327 type strains of *Blastococcus* species using a broad range of phenotypic features and from *B.*  
328 *saxobsidens* BC448<sup>T</sup>, its close phylogenetic relative, by low ANI and AAI indices and by a  
329 low *in silico* DNA:DNA pairing value. It can be concluded that the isolate should be  
330 recognised as a new *Blastococcus* species, for which we propose the name *Blastococcus*  
331 *atacamensis* sp. nov.

332

333 This is the first description of a novel *Blastococcus* species from the Atacama Desert though  
334 there are grounds for believing that others will follow [86], especially since culture-  
335 independent studies show that blastococci are part of the core microbiome of hyper- and  
336 extreme hyper-arid soils of the desert landscape [87]. It is also interesting that the genomes  
337 of *B. atacamensis* P6<sup>T</sup> and *M. caceserii* KNN 45-2b [70], another isolate from the Yungay  
338 core region, contain genes or gene clusters associated with an ability to cope with low levels  
339 of carbon [73-75], osmotic stress [82], high UV radiation [79] and heat tolerance and  
340 desiccation (biosynthesis and uptake of trehalose; [78]). The genomes of these strains also

341 contain multiple *cox* genes suggesting that *Geodermatophilaceae* strains from the Atacama  
342 may be able to use carbon monoxide as a carbon and energy source, an observation in line  
343 with the suggestion that facultatively chemoautotrophic bacteria may sustain microbial  
344 communities in the nutrient impoverished high altitude Atacama Desert soils [88, 89].  
345 Biological adaptations such as these may account for the presence of blastococci in habitats  
346 characterised by scarcity of available water, low nutrient availability and extremes of  
347 temperature and UV radiation levels [2, 14].

348

349 **Description of *Blastococcus atacamensis* sp. nov.**

350 *Blastococcus atacamensis* (a.ta.cam.en'sis. N.L. masc. adj. *atacamensis*; belonging to the  
351 Atacama Desert, the source of the isolate).

352

353 Gram-stain positive, oxidase-negative actinobacterium which forms non-motile, rod-and  
354 coccoid-elements with evidence of budding. Round orange colonies with entire margins are  
355 formed on yeast extract-malt extract agar. Grows from 10–45°C, optimally ~35°C, from pH  
356 6–12, optimally ~pH 7.0 and in the presence of 3%, w/v sodium chloride. Degrades starch  
357 and tributyrin but not guanine. Arbutin is hydrolysed. Utilises L-glutamic acid (amino acid),  
358 D-glucose, glycerol,  $\alpha$ -methyl-D-glucoside, *N*-acetyl-D-glucosamine (sugars), acetic acid,  
359 acetoacetic acid,  $\gamma$ -amino-*n*-butyric acid,  $\alpha$ - and  $\beta$ -hydroxybutyric acid,  $\alpha$ -ketobutyric acid,  
360 D-gluconic acid,  $\alpha$ -ketoglutaric acid, D-malic acid and propionic acid (organic acids), but not  
361 D-mannose, D-melibiose, *N*-acetyl-neuraminic acid or L-rhamnose (sugars) or butyric acid,  
362 mucic acid, D-saccharic acid,  $\alpha$ -hydroxyphenylacetic acid or bromosuccinic acid (organic  
363 acids); is resistant to aztreonam (antibiotic), lithium chloride and potassium tellurite (heavy  
364 metals) and Tween 40 (surfactant) but is sensitive to fusidic acid, lincomycin, minocycline,  
365 rifamycin SV, troleandomycin and vancomycin (antibiotics), guanidine HCl (chaotropic  
366 agent), tetrazolium blue and tetrazolium violet (redox indicators), sodium bromide, sodium  
367 formate and sodium lactate (salts) and niaproof (surfactant) (GENIII microplates). Additional  
368 phenotypic features are cited either in the text or in Table 1. The predominant fatty acids are  
369 iso-C<sub>16:0</sub>, iso-C<sub>16:1</sub> H, iso-C<sub>15:0</sub> and 9-methyl-C<sub>16:0</sub>; the major sugars are galactose, glucose  
370 and ribose; the polar lipid profile contains diphosphatidylglycerol, phosphatidylcholine,  
371 phosphatidylethanolamine, phosphatidylglycerol and phosphatidylinositol, an unidentified

372 lipid and an unidentified phosphoglycolipid. Additional chemotaxonomic properties are  
373 typical of the genus. The *in silico* DNA G+C content of the type strain is 73.1 mol%. The  
374 type strain, P6<sup>T</sup> (=NCIMB 15090<sup>T</sup> = NRRL B-65468<sup>T</sup>) was isolated from an extreme hyper-  
375 arid soil sample from the Yungay core region of the Atacama Desert.

376

### 377 **Acknowledgements**

378 ATB and JAA thank the Royal Society for IJP grant 100654, BA, JAA, ATB, JFC and MG  
379 are grateful for support from the UK Newton Project for UK-Chile collaboration (grant JIC  
380 CA586), BA, JAA and JFC are indebted to the Programme of CONICYT (Chile) for funding  
381 of the Centre for Biotechnology and Bioengineering, CeBiB (project FB0001), IN for a  
382 postdoctoral fellowship from Newcastle University, TR for a scholarship from the Indonesian  
383 Government, MET for a grant from the Ministerio de Economía y Competitividad (CGL-  
384 52735-P) and ATB and MG for Emeritus Fellowships from the Leverhulme Trust.

385

### 386 **Funding information**

387 This research was supported by Newcastle University in collaboration with Centre for  
388 Biotechnology and Bioengineering, CeBiB, Santiago, Chile.

### 389 **Conflicts of interest**

390 The authors declare that they do not have any conflict of interest.

### 391 **Ethical statement**

392 The authors have not carried out any studies involving human participants or animals.

### 393 **References**

- 394 1. **Ahrens R, Moll G.** Ein neues knospendes Bakterium aus der Ostsee. *Archiv für*  
395 *Mikrobiologie* 1970;70(3):243-265.
- 396 2. **Urzi C, Brusetti L, Salamone P, Sorlini C, Stackebrandt E et al.** Biodiversity of  
397 *Geodermatophilaceae* isolated from altered stones and monuments in the Mediterranean  
398 basin. *Environ Microbiol* 2001;3(7):471-479.
- 399 3. **Lee SD.** *Blastococcus jejuensis* sp. nov., an actinomycete from beach sediment, and  
400 emended description of the genus *Blastococcus* Ahrens and Moll 1970. *Int J Syst Evol*  
401 *Microbiol* 2006;56(Pt 10):2391-2396.
- 402 4. **Hezbri K, Louati M, Nouioui I, Gtari M, Rohde M et al.** *Blastococcus capsensis*  
403 sp. nov., isolated from an archaeological Roman pool and emended description of the genus  
404 *Blastococcus*, *B. aggregatus*, *B. saxobsidens*, *B. jejuensis* and *B. endophyticus*. *Int J Syst Evol*  
405 *Microbiol* 2016;66(11):4864-4872.

- 406 5. **Huang J, Li J, Cao M, Liao S, Wang G.** *Cumulibacter manganitolerans* gen. nov.,  
407 sp. nov., isolated from sludge of a manganese mine. *Int J Syst Evol Microbiol*  
408 2017;67(8):2646-2652.
- 409 6. **Luedemann GM.** *Geodermatophilus*, a new genus of the *Dermatophilaceae*  
410 (*Actinomycetales*). *J Bacteriol* 1968;96(5):1848-1858.
- 411 7. **Montero-Calasanz MdC, Meier-Kolthoff JP, Zhang D-F, Yaramis A, Rohde M**  
412 **et al.** Genome-scale data call for a taxonomic rearrangement of *Geodermatophilaceae*.  
413 *Frontiers in Microbiology* 2017;8(2501).
- 414 8. **Mevs U, Stackebrandt E, Schumann P, Gallikowski CA, Hirsch P.** *Modestobacter*  
415 *multiseptatus* gen. nov., sp. nov., a budding actinomycete from soils of the Asgard Range  
416 (Transantarctic Mountains). *Int J Syst Evol Microbiol* 2000;50(1):337-346.
- 417 9. **Normand P.** *Geodermatophilaceae* fam. nov., a formal description. *Int J Syst Evol*  
418 *Microbiol* 2006;56(10):2277-2278.
- 419 10. **Normand P, Daffonchio D, Gtari M.** The family *Geodermatophilaceae*. In:  
420 Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (editors). *The Prokaryotes:*  
421 *Actinobacteria*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. pp. 361-379.
- 422 11. **Sen A, Daubin V, Abrouk D, Gifford I, Berry AM et al.** Phylogeny of the class  
423 *Actinobacteria* revisited in the light of complete genomes. The orders ‘*Frankiales*’ and  
424 *Micrococcales* should be split into coherent entities: proposal of *Frankiales* ord. nov.,  
425 *Geodermatophilales* ord. nov., *Acidothemales* ord. nov. and *Nakamurellales* ord. nov. *Int J*  
426 *Syst Evol Microbiol* 2014;64(11):3821-3832.
- 427 12. **Chouaia B, Crotti E, Brusetti L, Daffonchio D, Essoussi I et al.** Genome sequence  
428 of *Blastococcus saxobsidens* DD2, a stone-inhabiting bacterium. *J Bacteriol*  
429 2012;194(10):2752-2753.
- 430 13. **Normand P, Benson DR.** Family IV. *Geodermatophilaceae* Normand 2006, 2277<sup>VP</sup>  
431 (Effective publication: Normand, Orso, Cournoyer, Jeannin, Chapelon, Dawson, Evtushenko  
432 and Misra 1996, 8.). In: Goodfellow M, Kämpfer P, Busse H-J, Trujillo ME, Suzuki K-i et  
433 al. (editors). *Bergey’s manual of systematic bacteriology: The Actinobacteria, Part A*. New  
434 York, NY: Springer; 2012. p. 528.
- 435 14. **Gtari M, Essoussi I, Maaoui R, Sghaier H, Boujmil R et al.** Contrasted resistance  
436 of stone-dwelling *Geodermatophilaceae* species to stresses known to give rise to reactive  
437 oxygen species. *FEMS Microbiol Ecol* 2012;80(3):566-577.
- 438 15. **Eppard M, Krumbein WE, Koch C, Rhiel E, Staley JT et al.** Morphological,  
439 physiological, and molecular characterization of actinomycetes isolated from dry soil, rocks,  
440 and monument surfaces. *Arch Microbiol* 1996;166(1):12-22.
- 441 16. **Nie G-X, Ming H, Li S, Zhou E-M, Cheng J et al.** *Geodermatophilus nigrescens*  
442 sp. nov., isolated from a dry-hot valley. *Antonie van Leeuwenhoek* 2012;101(4):811-817.
- 443 17. **Zhu WY, Zhang JL, Qin YL, Xiong ZJ, Zhang DF et al.** *Blastococcus*  
444 *endophyticus* sp. nov., an actinobacterium isolated from *Camptotheca acuminata*. *Int J Syst*  
445 *Evol Microbiol* 2013;63(Pt 9):3269-3273.
- 446 18. **Stackebrandt E, Schumann P.** Genus II. *Blastococcus*. In: Goodfellow M, Kämpfer  
447 P, Busse H-J, Trujillo ME, Suzuki K-I et al. (editors). *Bergeys Manual of Systematic*  
448 *Bacteriology (The Actinobacteria, Part A)*. New York: Springer; 2012. pp. 531–536.
- 449 19. **Urzì C, Salamone P, Schumann P, Rohde M, Stackebrandt E.** *Blastococcus*  
450 *saxobsidens* sp. nov., and emended descriptions of the genus *Blastococcus* Ahrens and Moll  
451 1970 and *Blastococcus aggregatus* Ahrens and Moll 1970. *Int J Syst Evol Microbiol*  
452 2004;54(Pt 1):253-259.

- 453 20. **Hezbri K, Nouioui I, Rohde M, Schumann P, Gtari M et al.** *Blastococcus colisei*  
454 sp. nov, isolated from an archaeological amphitheatre. *Antonie van Leeuwenhoek*  
455 2017;110(3):339-346.
- 456 21. **Pfleiderer A, Lagier JC, Armougom F, Robert C, Vialettes B et al.** Culturomics  
457 identified 11 new bacterial species from a single anorexia nervosa stool sample. *Eur J Clin*  
458 *Microbiol Infect Dis* 2013;32(11):1471-1481.
- 459 22. **Atlas RM.** *Handbook of microbiological media*, 4<sup>th</sup> ed: CRC Press; 2010.
- 460 23. **Vickers JC, Williams ST.** An assessment of plate inoculation procedures for the  
461 enumeration and isolation of soil streptomycetes. *Microbios Letters* 1987;35:113-117.
- 462 24. **Trujillo ME, Goodfellow M, Busarakam K, Riesco R.** *Modestobacter lapidis* sp.  
463 nov. and *Modestobacter muralis* sp. nov., isolated from a deteriorated sandstone historic  
464 building in Salamanca, Spain. *Antonie van Leeuwenhoek* 2015;108(2):311-320.
- 465 25. **Shirling EB, Gottlieb D.** Methods for characterization of *Streptomyces* species. *Int*  
466 *J Syst Evol Microbiol* 1966;16(3):313-340.
- 467 26. **Staneck JL, Roberts GD.** Simplified approach to identification of aerobic  
468 Actinomycetes by thin-layer chromatography. *Applied Microbiology* 1974;28(2):226-231.
- 469 27. **Collins MD, Goodfellow M, Minnikin DE, Alderson G.** Menaquinone composition  
470 of mycolic acid-containing actinomycetes and some sporoactinomycetes. *Journal of Applied*  
471 *Bacteriology* 1985;58(1):77-86.
- 472 28. **Lechevalier MP, Lechevalier H.** Chemical composition as a criterion in the  
473 classification of aerobic actinomycetes. *Int J Syst Evol Microbiol* 1970;20(4):435-443.
- 474 29. **Minnikin DE, Patel PV, Alshamaony L, Goodfellow M.** Polar lipid composition in  
475 the classification of *Nocardia* and related bacteria. *Int J Syst Evol Microbiol* 1977;27(2):104-  
476 117.
- 477 30. **Kroppenstedt R, Goodfellow M.** The family *Thermomonosporaceae*:  
478 *Actinocorallia*, *Actinomadura*, *Spirillispora* and *Thermomonospora*. In: Dworkin M, Falkow  
479 S, Schleifer KH, E S (editors). *The Prokaryotes, Archaea and Bacteria Firmicutes,*  
480 *Actinomycetes.* New York: Springer; 2006. pp. 682-724.
- 481 31. **Carro L, Riesco R, Sproer C, Trujillo ME.** *Micromonospora luteifusca* sp. nov.  
482 isolated from cultivated *Pisum sativum*. *Syst Appl Microbiol* 2016;39(4):237-242.
- 483 32. Sasser M. Identification of bacteria by gas chromatography of cellular fatty acids.  
484 Newark, DE: MIDI Inc.1990.
- 485 33. **Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA.** *Practical*  
486 *Streptomyces genetics.* Norwich, UK: John Innes Foundation; 2000.
- 487 34. **Lane DJ.** 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (editors).  
488 *Nucleic acid techniques in bacterial systematics.* New York: John Wiley & Sons; 1991. pp.  
489 115-175.
- 490 35. **Staden R, Beal KF, Bonfield JK.** The Staden package, 1998. *Methods Mol Biol*  
491 2000;132:115-130.
- 492 36. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y et al.** Introducing EzBioCloud: A  
493 taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst Evol*  
494 *Microbiol* 2016;67(5):1613-1617.
- 495 37. **Edgar RC.** MUSCLE: multiple sequence alignment with high accuracy and high  
496 throughput. *Nucleic Acids Res* 2004;32(5):1792-1797.
- 497 38. **Kumar S, Stecher G, Tamura K.** MEGA7: molecular evolutionary genetics analysis  
498 version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33(7):1870-1874.



- 499 39. **Saitou N, Nei M.** The neighbor-joining method: a new method for reconstructing  
500 phylogenetic trees. *Mol Biol Evol* 1987;4(4):406-425.
- 501 40. **Felsenstein J.** Evolutionary trees from DNA sequences: a maximum likelihood  
502 approach. *J Mol Evol* 1981;17(6):368-376.
- 503 41. **Fitch WM.** Toward defining the course of evolution: minimum change for a specific  
504 tree topology. *Systematic Zoology* 1971;20(4):406-416.
- 505 42. **Felsenstein J.** Confidence limits on phylogenies: an approach using the bootstrap.  
506 *Evolution* 1985;39(4):783-791.
- 507 43. **Kimura M.** A simple method for estimating evolutionary rates of base substitutions  
508 through comparative studies of nucleotide sequences. *J Mol Evol* 1980;16(2):111-120.
- 509 44. **Meier-Kolthoff JP, Goker M, Sproer C, Klenk HP.** When should a DDH  
510 experiment be mandatory in microbial taxonomy? *Arch Microbiol* 2013;195(6):413-418.
- 511 45. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M.** Genome sequence-based  
512 species delimitation with confidence intervals and improved distance functions. *BMC*  
513 *Bioinformatics* 2013;14:60-60.
- 514 46. **Vaas LAI, Sikorski J, Hofner B, Fiebig A, Buddruhs N et al.** opm: an R package  
515 for analysing OmniLog® phenotype microarray data. *Bioinformatics* 2013;29(14):1823-  
516 1824.
- 517 47. **R Core Team.** R: a language and environment for statistical computing version 3.3.1.  
518 R Foundation for Statistical Computing, Vienna, Austria; 2016. <http://www.r-project.org/>.
- 519 48. **RStudio Team.** RStudio: integrated development for R version 0.99.903. RStudio,  
520 Inc., Boston, MA; 2015. <http://www.rstudio.com/>.
- 521 49. **Montero-Calasanz MC, Hofner B, Goker M, Rohde M, Sproer C et al.**  
522 *Geodermatophilus poikilotrophi* sp. nov.: a multitolerant actinomycete isolated from  
523 dolomitic marble. *BioMed research international* 2014;2014:914767.
- 524 50. **Williams ST, Goodfellow M, Alderson G, Wellington EMH, Sneath PHA et al.**  
525 Numerical classification of *Streptomyces* and related genera. *Microbiology*  
526 1983;129(6):1743-1813.
- 527 51. **Christensen WB.** Urea decomposition as a means of differentiating proteus and  
528 paracolon cultures from each other and from *Salmonella* and *Shigella* types. *J Bacteriol*  
529 1946;52(4):461-466.
- 530 52. **Schaal KP, Yassin AF, Stackebrandt E.** The family *Actinomycetaceae*: the genera  
531 *Actinomyces*, *Actinobaculum*, *Arcanobacterium*, *Varibaculum*, and *Mobiluncus*. In: Dworkin  
532 M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (editors). *The Prokaryotes – a*  
533 *handbook on the biology of bacteria: Archaea Bacteria: Firmicutes, Actinomycetes*. New  
534 York: Springer; 2006. pp. 430-537.
- 535 53. **Kovacs N.** Identification of *Pseudomonas pyocyanea* by the oxidase reaction. *Nature*  
536 1956;178(4535):703.
- 537 54. **Murray PR.** *Manual of clinical microbiology*. Washington, D.C.: ASM Press; 1999.
- 538 55. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M et al.** SPAdes: a new  
539 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*  
540 2012;19(5):455-477.
- 541 56. **Aziz RK, Bartels D, Best AA, DeJongh M, Disz T et al.** The RAST server: rapid  
542 annotations using subsystems technology. *BMC Genomics* 2008;9(1):75.
- 543 57. **Aziz RK, Devoid S, Disz T, Edwards RA, Henry CS et al.** SEED servers: high-  
544 performance access to the SEED genomes, annotations, and metabolic models. *PLOS ONE*  
545 2012;7(10):e48053.

- 546 58. **Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P et al.**  
547 DNA–DNA hybridization values and their relationship to whole-genome sequence  
548 similarities. *Int J Syst Evol Microbiol* 2007;57(1):81-91.
- 549 59. **Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O et al.** Report of  
550 the ad hoc committee on reconciliation of Approaches to bacterial systematics. *Int J Syst Evol*  
551 *Microbiol* 1987;37(4):463-464.
- 552 60. **Richter M, Rosselló-Móra R.** Shifting the genomic gold standard for the prokaryotic  
553 species definition. *Proc Natl Acad Sci U S A* 2009;106(45):19126-19131.
- 554 61. **Chun J, Rainey FA.** Integrating genomics into the taxonomy and systematics of the  
555 Bacteria and Archaea. *Int J Syst Evol Microbiol* 2014;64(Pt 2):316-324.
- 556 62. **Weber T, Blin K, Duddela S, Krug D, Kim HU et al.** antiSMASH 3.0—a  
557 comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids*  
558 *Res* 2015;43(W1):W237-W243.
- 559 63. **Funa N, Ohnishi Y, Fujii I, Shibuya M, Ebizuka Y et al.** A new pathway for  
560 polyketide synthesis in microorganisms. *Nature* 1999;400(6747):897-899.
- 561 64. **Yu D, Xu F, Zeng J, Zhan J.** Type III polyketide synthases in natural product  
562 biosynthesis. *IUBMB Life* 2012;64(4):285-295.
- 563 65. **The UniProt Consortium.** UniProt: the universal protein knowledgebase. *Nucleic*  
564 *Acids Res* 2017;45(D1):D158-D169.
- 565 66. **Sandmann G.** Carotenoid biosynthesis and biotechnological application. *Archives of*  
566 *Biochemistry and Biophysics* 2001;385(1):4-12.
- 567 67. **Klassen JL.** Phylogenetic and evolutionary patterns in microbial carotenoid  
568 biosynthesis are revealed by comparative genomics. *PLOS ONE* 2010;5(6):e11257.
- 569 68. **Finn RD, Cogill P, Eberhardt RY, Eddy SR, Mistry J et al.** The Pfam protein  
570 families database: towards a more sustainable future. *Nucleic Acids Res* 2016;44(D1):D279-  
571 D285.
- 572 69. **Barona-Gómez F, Wong U, Giannakopoulos AE, Derrick PJ, Challis GL.**  
573 Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces*  
574 *coelicolor* M145. *J Am Chem Soc* 2004;126(50):16282-16283.
- 575 70. **Busarakam K, Bull AT, Trujillo ME, Riesco R, Sangal V et al.** *Modestobacter*  
576 *caceresii* sp. nov., novel actinobacteria with an insight into their adaptive mechanisms for  
577 survival in extreme hyper-arid Atacama Desert soils. *Syst Appl Microbiol* 2016;39(4):243-  
578 251.
- 579 71. **Gomez-Escribano JP, Castro JF, Razmilic V, Chandra G, Andrews B et al.** The  
580 *Streptomyces leeuwenhoekii* genome: de novo sequencing and assembly in single contigs of  
581 the chromosome, circular plasmid pSLE1 and linear plasmid pSLE2. *BMC Genomics*  
582 2015;16:485.
- 583 72. **Anantharaman V, Iyer LM, Aravind L.** Ter-dependent stress response systems:  
584 novel pathways related to metal sensing, production of a nucleoside-like metabolite, and  
585 DNA-processing. *Molecular BioSystems* 2012;8(12):3142-3165.
- 586 73. **Schultz JE, Matin A.** Molecular and functional characterization of a carbon  
587 starvation gene of *Escherichia coli*. *J Mol Biol* 1991;218(1):129-140.
- 588 74. **Lucchetti-Miganeh C, Burrowes E, Baysse C, Ermel G.** The post-transcriptional  
589 regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages  
590 of infection in animal hosts. *Microbiology* 2008;154(1):16-29.
- 591 75. **Rasmussen JJ, Vegge CS, Frøkiær H, Howlett RM, Kroghfelt KA et al.**  
592 *Campylobacter jejuni* carbon starvation protein A (CstA) is involved in peptide utilization,

593 motility and agglutination, and has a role in stimulation of dendritic cells. *J Med Microbiol*  
594 2013;62(8):1135-1143.

595 76. **Essoussi I, Ghodhbane-Gtari F, Amairi H, Sghaier H, Jaouani A et al.** Esterase  
596 as an enzymatic signature of *Geodermatophilaceae* adaptability to Sahara desert stones and  
597 monuments. *J Appl Microbiol* 2010;108(5):1723-1732.

598 77. **Li J-s, Bi Y-t, Dong C, Yang J-f, Liang W-d.** Transcriptome analysis of adaptive  
599 heat shock response of *Streptococcus thermophilus*. *PLOS ONE* 2011;6(10):e25777.

600 78. **Reina-Bueno M, Argandoña M, Nieto JJ, Hidalgo-García A, Iglesias-Guerra F**  
601 **et al.** Role of trehalose in heat and desiccation tolerance in the soil bacterium *Rhizobium etli*.  
602 *BMC Microbiol* 2012;12(1):207.

603 79. **Normand P, Gury J, Pujic P, Chouaia B, Crotti E et al.** Genome sequence of  
604 radiation-resistant *Modestobacter marinus* strain BC501, a representative actinobacterium  
605 that thrives on calcareous stone surfaces. *J Bacteriol* 2012;194(17):4773-4774.

606 80. **Manthei KA, Hill MC, Burke JE, Butcher SE, Keck JL.** Structural mechanisms of  
607 DNA binding and unwinding in bacterial RecQ helicases. *Proceedings of the National*  
608 *Academy of Sciences* 2015;112(14):4292-4297.

609 81. **Lorite MJ, Tachil J, Sanjuán J, Meyer O, Bedmar EJ.** Carbon monoxide  
610 dehydrogenase activity in *Bradyrhizobium japonicum*. *Appl Environ Microbiol*  
611 2000;66(5):1871-1876.

612 82. **Boncompagni E, Dupont L, Mignot T, Osteras M, Lambert A et al.**  
613 Characterization of a *Snorhizobium meliloti* ATP-binding cassette histidine transporter also  
614 involved in betaine and proline uptake. *J Bacteriol* 2000;182(13):3717-3725.

615 83. **Nau-Wagner G, Opper D, Rolbetzki A, Boch J, Kempf B et al.** Genetic control of  
616 osmoadaptive glycine betaine synthesis in *Bacillus subtilis* through the choline-sensing and  
617 glycine betaine-responsive GbsR repressor. *J Bacteriol* 2012;194(10):2703-2714.

618 84. **Kappes RM, Kempf B, Kneip S, Boch J, Gade J et al.** Two evolutionarily closely  
619 related ABC transporters mediate the uptake of choline for synthesis of the osmoprotectant  
620 glycine betaine in *Bacillus subtilis*. *Mol Microbiol* 1999;32(1):203-216.

621 85. **Mandon K, Osteras M, Boncompagni E, Trinchant JC, Spennato G et al.** The  
622 *Sinorhizobium meliloti* glycine betaine biosynthetic genes (*betlCBA*) are induced by choline  
623 and highly expressed in bacteroids. *Mol Plant Microbe Interact* 2003;16(8):709-719.

624 86. **Bull AT, Asenjo JA, Goodfellow M, Gomez-Silva B.** The Atacama Desert:  
625 technical resources and the growing importance of novel microbial diversity. *Annu Rev*  
626 *Microbiol* 2016;70:215-234.

627 87. **Idris H, Goodfellow M, Sandenon R, Asenjo JA, Bull AT.** Actinobacterial rare  
628 biosphere and dark matter revealed in habitats of the Atacama Desert. *Scientific Reports*  
629 2017;7.

630 88. **Costello EK, Halloy SR, Reed SC, Sowell P, Schmidt SK.** Fumarole-supported  
631 islands of biodiversity within a hyperarid, high-elevation landscape on Socompa Volcano,  
632 Puna de Atacama, Andes. *Appl Environ Microbiol* 2009;75(3):735-747.

633 89. **Lynch RC, King AJ, Farías ME, Sowell P, Vitry C et al.** The potential for microbial  
634 life in the highest-elevation (>6000 m.a.s.l.) mineral soils of the Atacama region. *Journal of*  
635 *Geophysical Research: Biogeosciences* 2012;117(G2):G02028.

636

639 **Table 1:** Phenotypic properties that differentiate isolate P6<sup>T</sup> from the *Blastococcus* type strains.

<b>Characteristics</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Cell shape	Cocci	Cocci, rods, vibrios <sup>d</sup>	Cocci	Cocci, rods, vibrios	Cocci <sup>e</sup>	Cocci, rods <sup>c</sup>	Cocci <sup>d</sup>
Bud formation	+	+ <sup>d</sup>	–	+	– <sup>e</sup>	+ <sup>c</sup>	– <sup>d</sup>
Germ tube	+	+ <sup>d</sup>	–	+	– <sup>e</sup>	– <sup>c</sup>	– <sup>d</sup>
Motility	–	+ <sup>d</sup>	–	–	– <sup>e</sup>	+ <sup>c</sup>	+ <sup>d</sup>
Pigmentation	Light-pink to mild red-orange	Pink	Bright orange	Coral	White to pink to black <sup>e</sup>	Apricot	Pink to orange
Temperature growth range (°C)	10–45	10–40 <sup>d</sup>	20–30	10–40	10–45 <sup>e</sup>	10–37 <sup>c</sup>	20–37 <sup>d</sup>
<b>Biochemical tests:</b>							
Catalase	+	+	–	–	+ <sup>e</sup>	+ <sup>c</sup>	+ <sup>*</sup>
Nitrate to nitrite reduction	+	– <sup>e</sup>	– <sup>b</sup>	–	– <sup>e</sup>	+ <sup>e</sup>	– <sup>e</sup>
<b>API ZYM tests:</b>							
Acid phosphatase	+	– <sup>e</sup>	– <sup>b</sup>	–	+ <sup>e</sup>	– <sup>e</sup>	+ <sup>‡</sup>
Alkaline phosphatase	+	– <sup>e</sup>	+ <sup>b</sup>	+	+ <sup>e</sup>	– <sup>e</sup>	+ <sup>e</sup>
Esterase lipase (C 8)	+	+ <sup>e</sup>	+ <sup>b</sup>	–	+ <sup>e</sup>	– <sup>e</sup>	+ <sup>e</sup>
α-Glucosidase	+	– <sup>e</sup>	+ <sup>b</sup>	+	+ <sup>e</sup>	– <sup>e</sup>	+ <sup>‡</sup>
Naphthol-AS-BI-phosphohydrolase	+	– <sup>e</sup>	– <sup>b</sup>	–	+ <sup>e</sup>	– <sup>e</sup>	+ <sup>e</sup>
Valine arylamidase	+	– <sup>e</sup>	– <sup>b</sup>	+	+ <sup>e</sup>	+ <sup>e</sup>	+ <sup>‡</sup>
<b>GENIII Biolog microplates:</b>							
<b>Oxidation of Amino acids:</b>							
L-Alanine	–	+	–	+	v	+	–

Glycyl-L-proline(dipeptide)	–	+	v	–	–	–	–
<b>Monosaccharides:</b>							
<i>N</i> -Acetyl-neuraminic acid	–	+	–	–	–	–	v
Glucuronamide	–	v	+	–	+	v	+
<b>Disaccharide:</b>							
$\beta$ -Gentiobiose	+	+	–	+	+	+	+
<b>Sugar alcohol:</b>							
D-Salicin	+	–	–	v	v	–	+
<b>Polymers:</b>							
Dextrin	+	+	–	+	+	–	–
Gelatin	–	+	–	–	v	+	–
Pectin	+	–	+	v	–	–	–
Tween 40	+	+	–	+	+	+	v
<b>Organic acids:</b>							
D-Gluconic acid	+	–	+	+	+	v	+
$\beta$ -hydroxy-Butyric acid	+	–	+	+	v	+	+
$\alpha$ -keto-Butyric acid	+	v	–	+	–	+	–
D-Malic acid	+	+	+	v	+	+	–
Methyl pyruvate	+	–	–	v	–	–	v
Mucic acid	–	–	–	v	–	+	–
D-Saccharic acid	–	–	–	+	+	v	+
<b>Phospholipids</b>	DPG, L, PC PE, PG, PGL, PI	DPG, PI, GPL, PC	DPG, PE, PC, PI, 2PL	DPG, PC, PI, GPL, PE, OH-PE, 6PL	DPG, PI, PE, PC	DPG, PC, PME, PE, PI, GPL	DPG, PE, PC, PI, GLP, 3PL
<b>Diagnostic sugars</b>	Glu, Gal, Rib; traces of Ara, Xyl	Rib, Ara, Man, Glu	Glu, Rha, Rib	Glu, Gal, Rib	Glu, Gal; traces of	Rha, Rib, Xyl, Glu;	Glu, Gal; traces of Rib, Man

<b>Menaquinones (MK)</b>	MK-9(H <sub>4</sub> ), MK-9(H <sub>2</sub> )	MK-8(H <sub>4</sub> ), MK-9(H <sub>4</sub> )	MK-9(H <sub>4</sub> ), MK-9, MK- 9(H <sub>2</sub> )	MK-9(H <sub>4</sub> ), MK-9, MK- 8(H <sub>4</sub> )	Rha, Rib, Man, Ara MK-9(H <sub>4</sub> ), MK-9, MK- 8(H <sub>4</sub> )	traces of Man MK-9(H <sub>4</sub> ), MK-9	MK-9(H <sub>4</sub> ), MK-8(H <sub>4</sub> ), MK-9, MK- 9(H <sub>6</sub> ), MK- 9(H <sub>2</sub> )
<b>DNA G + C content (mole %)</b>	73.1	73.9 <sup>d</sup>	73.7	73.2	71.6 <sup>e</sup>	72.3 <sup>c</sup>	73.5

640

641

642 Strains: **1**, Isolate P6<sup>T</sup>; **2**, *B. aggregatus* DSM 4725<sup>T</sup>; **3**, *B. capsensis* DSM 46835<sup>T</sup>; **4**, *B. colisei* DSM 46837<sup>T</sup>; **5**, *B. endophyticus* DSM  
643 45413<sup>T</sup>; **6**, *B. jejuensis* DSM 19597<sup>T</sup>; **7**, *B. saxobsidens* DSM 44509<sup>T</sup>. Data for **1** and **7** are from this study; those for **2** to **6** are from  
644 Hezbri et al. [20] unless indicated. + positive; – negative; v variable. Data taken from <sup>a</sup> Ahrens and Moll [1]; <sup>b</sup> Hezbri et al. [4]; <sup>c</sup> Lee [3];  
645 <sup>d</sup> Urzì et al. [19] and <sup>e</sup> Zhu et al. [17]; \* recorded as negative by Hezbri et al. [4]; † recorded as negative by Zhu et al. [17]. Abbreviations:  
646 Ara, arabinose; Glu, glucose; Gal, galactose; Man, mannose; Rha, rhamnose; Rib, ribose; Xyl, xylose; DPG, diphosphatidylglycerol;  
647 GPI, glycoposphatidylinositol; PE, phosphatidylethanolamine; PME, phosphatidyl-*N*-methylethanolamine; PC, phosphatidylcholine;  
648 PI, phosphatidylinositol; unidentified: GPL, glycopospholipid, L, lipid, PL, phospholipid; MK, menaquinone.

649 **Legends for Figures**

650

651 **Figure 1.** Phase contrast image of isolate P6<sup>T</sup> following growth on ISP 2 at 28°C for 7 days  
652 showing the presence of coccoid and rod-shaped elements and evidence of budding. Scale  
653 bar: 5µm.

654 **Figure 2.** Neighbour-joining tree based on partial 16S rRNA gene sequences (1,239  
655 nucleotides) showing the relationships between isolate P6<sup>T</sup> and the type strains of  
656 *Blastococcus* species and the candidatus strain. Asterisks indicate branches of the tree that  
657 were also found using the maximum-likelihood (ML) and maximum-parsimony (MP) tree-  
658 making algorithms. Numbers at the nodes indicate levels of bootstrap support (%) above 50%  
659 based on a neighbour-joining analysis of 1,000 resampled datasets. Genbank accession  
660 numbers are indicated in parentheses. The scale bar indicates the number of substitutions per  
661 nucleotide position.

662

663 **Supplementary Figures**

664 Fig. S1. Light microscopy images of colonies of isolate P6<sup>T</sup> following growth on ISP 2 at  
665 28°C for 3 weeks.

666 Fig. S2. Thin-layer chromatographs showing (a) the presence of diaminopimelic acid (A<sub>2</sub>pm)  
667 isomers and (b) the presence of *meso*-A<sub>2</sub>pm in whole-cell hydrolysates of isolate P6<sup>T</sup>.

668 Fig. S3. Menaquinone profile of isolate P6<sup>T</sup>. Isoprenoid quinones extracted from  
669 *Micromonospora luteifusca* GUI2<sup>T</sup> [1] were used as standards.

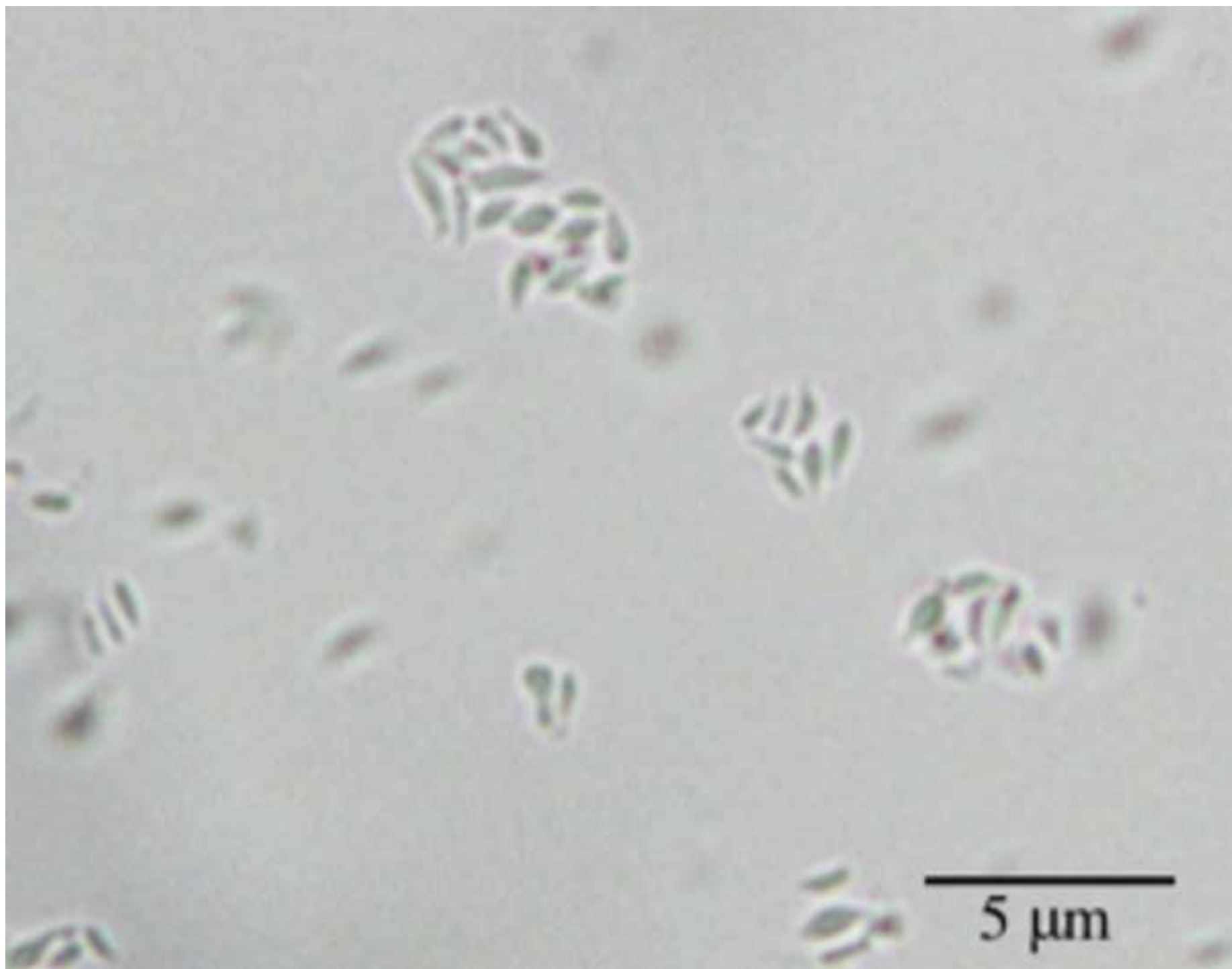
670 Fig. S4. Bi-dimensional thin-layer chromatography of polar lipids of isolate P6<sup>T</sup>.

671 Fig. S5. Thin-layer chromatographs showing the presence of (a) standard sugars and (b)  
672 sugars in whole-cell hydrolysates of isolate P6<sup>T</sup>.

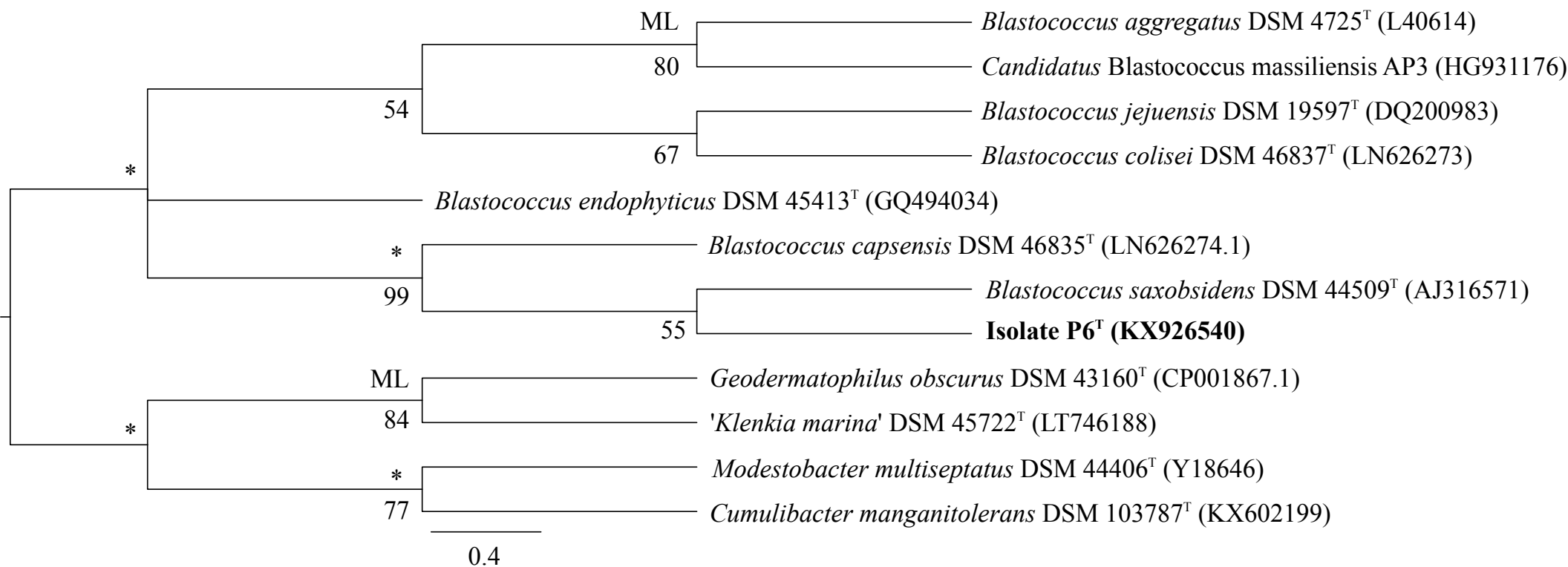
673

674

675







**International Journal of Systematic and Evolutionary Microbiology****Supplementary data*****Blastococcus atacamensis* sp. nov., a novel strain adapted to life in the Yungay core region of the Atacama Desert**

Jean Franco Castro<sup>1,2</sup>, Imen Nouioui<sup>1</sup>, Vartul Sangal<sup>3</sup>, Seonbin Choi<sup>4</sup>, Seung-Jo Yang<sup>4</sup>, Byung-Yong Kim<sup>4</sup>, Martha E. Trujillo<sup>5</sup>, Raul Riesco<sup>4</sup>, Maria del Carmen Montero-Calasanz<sup>1</sup>, Tara PD Rahmani<sup>1</sup>, Alan T. Bull<sup>6</sup>, Iain C. Sutcliffe<sup>3</sup>, Juan A. Asenjo<sup>2</sup>, Barbara Andrews<sup>2</sup>, Michael Goodfellow<sup>1\*</sup>

<sup>1</sup> School of Natural and Environmental Sciences, Ridley Building 2, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom

<sup>2</sup> Centre for Biotechnology and Bioengineering (CeBiB), Department of Chemical Engineering and Biotechnology, University of Chile, Beauchef 851, Santiago, Chile

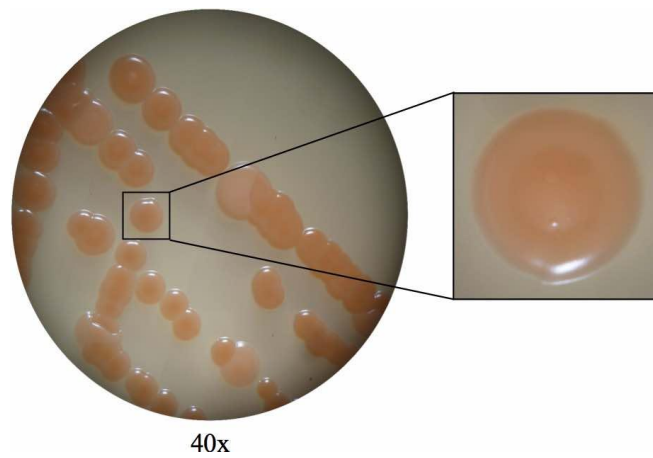
<sup>3</sup> Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, United Kingdom

<sup>4</sup> ChunLab, Inc., 1, Gwanaka-ro, Gwanak-gu, Seoul 151015, South Korea

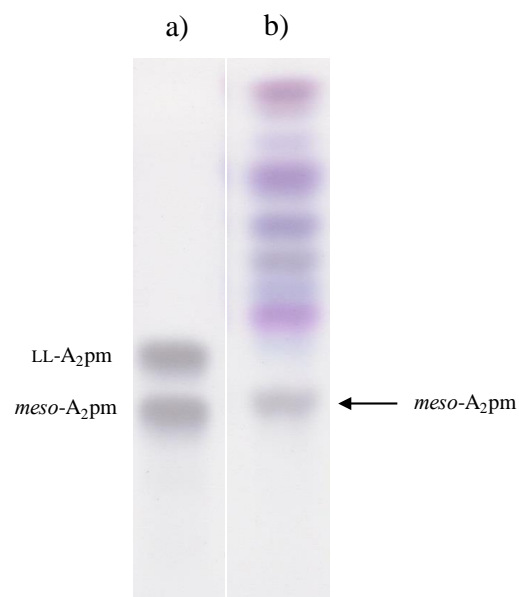
<sup>5</sup> Departamento de Microbiología y Genética, Universidad de Salamanca, Campus Miguel de Unamuno, 37007 Salamanca, Spain

<sup>6</sup> School of Biosciences, University of Kent, Canterbury CT2 7NJ, Kent, United Kingdom

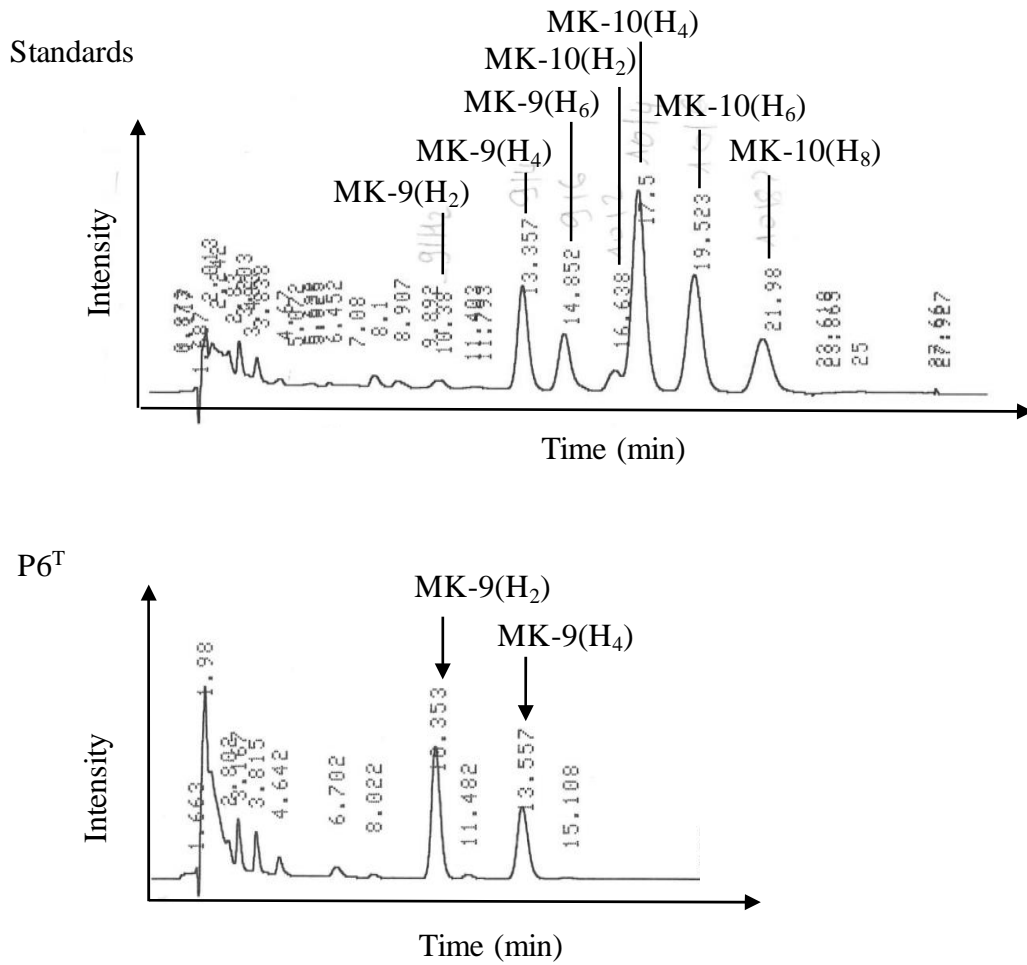
\*To whom correspondence should be addressed: phone +44-(0)191 2087706; email: [m.goodfellow@ncl.ac.uk](mailto:m.goodfellow@ncl.ac.uk).



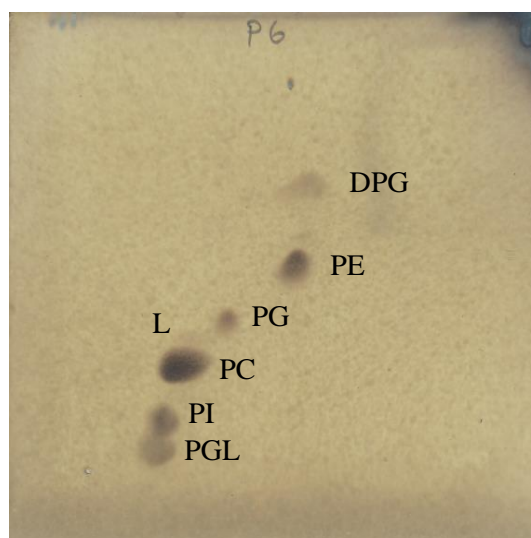
**Supplementary Figure S1:** Light microscopy images of colonies of isolate P6<sup>T</sup> following growth on ISP 2 at 28°C for 3 weeks.



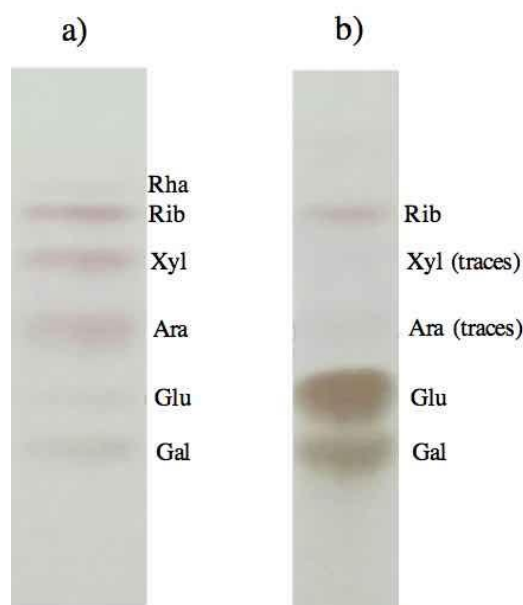
**Supplementary Fig. S2:** Thin-layer chromatographs showing (a) the presence of diaminopimelic acid (A<sub>2</sub>pm) isomers and (b) the presence of *meso*-A<sub>2</sub>pm in whole-cell hydrolysates of isolate P6<sup>T</sup>. a) Standards A<sub>2</sub>pm; b) isolate P6<sup>T</sup>.



**Supplementary Fig. S3:** Menaquinone profile of isolate P6<sup>T</sup>. Isoprenoid quinones extracted from *Micromonospora luteifusca* GUI2<sup>T</sup> [1] were used as standards.



**Supplementary Fig. S4:** Bi-dimensional thin-layer chromatography of polar lipids of isolate P6<sup>T</sup> using molybdato-phosphoric acid reagent (5 %). DPG: diphosphatidylglycerol; L: unidentified lipid; PC: phosphatidylcholine; PE: phosphatidylethanolamine; PG: phosphatidylglycerol; PGL: Unknown phosphoglycolipid; PI: phosphatidylinositol. Solvent 1: chloroform:methanol:distilled water (65:25:4); solvent 2: chloroform:glacial acetic acid:methanol:distilled water (80:12:15:4).



**Supplementary Fig. S5.** Thin-layer chromatographs showing the presence of (a) standard sugars and (b) sugars in whole-cell hydrolysates of isolate P6<sup>T</sup>. a) Standard sugars; b) sugars present in isolate P6<sup>T</sup>. Rha: rhamnose; Rib: ribose; Xyl: xylose; Ara: arabinose; Glu: glucose; Gal: galactose.

**Supplementary Table 1. A list of stress related genes present in strains P6<sup>T</sup> and DSM 44509<sup>T</sup>**

Presence	Category	Role	Organism A	Organism B
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Cold shock	Cold shock protein CspA	P6_peg1370, P6_peg2786	DSM44509_peg149, DSM44509_peg3572
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Cold shock	Cold shock protein CspC	P6_peg575	DSM44509_peg176
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Cold shock	Cold shock protein CspG	P6_peg2467	DSM44509_peg2911
P6 <sup>T</sup>	Detoxification	Various polyols ABC transporter, periplasmic substrate-binding protein	P6_peg1164, P6_peg1169	
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Chaperone protein DnaJ	P6_peg274, P6_peg1954	DSM44509_peg382, DSM44509_peg2184
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Chaperone protein DnaK	P6_peg226	DSM44509_peg384
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Heat shock protein GrpE	P6_peg227	DSM44509_peg383
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Heat-inducible transcription repressor HrcA	P6_peg1955	DSM44509_peg2183
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	HspR, transcriptional repressor of DnaK operon	P6_peg275	DSM44509_peg381
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Hypothetical radical SAM family enzyme in heat shock gene cluster, similarity with CPO of BS HemN-type	P6_peg1959	DSM44509_peg2178
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Nucleoside 5-triphosphatase RdgB (dHATP, dITP, XTP-specific) (EC 3.6.1.15)	P6_peg6	DSM44509_peg2317
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Ribonuclease PH (EC 2.7.7.56)	P6_peg5	DSM44509_peg2316
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-)	P6_peg1953	DSM44509_peg2185
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	Translation elongation factor LepA	P6_peg1981	DSM44509_peg2165
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	rRNA small subunit methyltransferase I	P6_peg2916	DSM44509_peg235
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Heat shock	tmRNA-binding protein SmpB	P6_peg3717	DSM44509_peg3854

P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Osmotic stress	Choline dehydrogenase (EC 1.1.99.1)	P6_peg827, P6_peg3514	DSM44509_peg321, DSM44509_peg1542
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Osmotic stress	Glycine betaine ABC transport system permease protein	P6_peg1488	DSM44509_peg112
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Osmotic stress	High-affinity choline uptake protein BetT	P6_peg2419	DSM44509_peg286, DSM44509_peg1388, DSM44509_peg1393, DSM44509_peg1579
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Osmotic stress	L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)	P6_peg944, P6_peg1490	DSM44509_peg110, DSM44509_peg1084
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Osmotic stress	L-proline glycine betaine ABC transport system permease protein ProW (TC 3.A.1.12.1)	P6_peg1489	DSM44509_peg111, DSM44509_peg1082, DSM44509_peg1083
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Osmotic stress	L-proline glycine betaine binding ABC transporter protein ProX (TC 3.A.1.12.1)	P6_peg1487	DSM44509_peg113, DSM44509_peg279
DSM 44509 <sup>T</sup>	Osmotic stress	Sarcosine oxidase alpha subunit (EC 1.5.3.1)		DSM44509_peg1385
DSM 44509 <sup>T</sup>	Osmotic stress	Sarcosine oxidase beta subunit (EC 1.5.3.1)		DSM44509_peg1383
DSM 44509 <sup>T</sup>	Osmotic stress	Sarcosine oxidase delta subunit (EC 1.5.3.1)		DSM44509_peg1384
DSM 44509 <sup>T</sup>	Osmotic stress	Sarcosine oxidase gamma subunit (EC 1.5.3.1)		DSM44509_peg1386
DSM 44509 <sup>T</sup>	Osmotic stress	Aquaporin Z		DSM44509_peg31
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Putative Holliday junction resolvase YggF	P6_peg2241	DSM44509_peg1712
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.-)	P6_peg1953	DSM44509_peg2185
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)	P6_peg3406	DSM44509_peg3484, DSM44509_peg3485
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	P6_peg65, P6_peg2230	DSM44509_peg1701, DSM44509_peg2761
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Lactoylglutathione lyase (EC 4.4.1.5)	P6_peg2435	DSM44509_peg1561
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Acetyl-CoA:Cys-GlcN-Ins acetyltransferase, mycothiol synthase MshD	P6_peg3623	DSM44509_peg3541
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Formaldehyde dehydrogenase MscR, NAD/mycothiol-dependent (EC 1.2.1.66)	P6_peg2339	DSM44509_peg334, DSM44509_peg1381
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Glycosyltransferase MshA involved in mycothiol biosynthesis (EC 2.4.1.-)	P6_peg3580	DSM44509_peg3513
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	L-cysteine:1D-myo-inositol 2-amino-2-deoxy-alpha-D-glucopyranoside ligase MshC	P6_peg3347	DSM44509_peg1070
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Mycothiol S-conjugate amidase Mca	P6_peg321	DSM44509_peg2370

P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase MshB	P6_peg1013	DSM44509_peg1794
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Putative hydrolase in cluster with formaldehyde/S-nitrosomycothiol reductase MscR	P6_peg2338	DSM44509_peg333
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	S-nitrosomycothiol reductase MscR	P6_peg2339	DSM44509_peg334, DSM44509_peg1381
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Uncharacterized protein Rv0487/MT0505 clustered with mycothiol biosynthesis gene	P6_peg3579	DSM44509_peg3512
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Alkyl hydroperoxide reductase subunit C-like protein	P6_peg3238	DSM44509_peg984
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Catalase (EC 1.11.1.6)	P6_peg613	DSM44509_peg1819, DSM44509_peg3715
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Organic hydroperoxide resistance protein	P6_peg279	DSM44509_peg379
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Organic hydroperoxide resistance transcriptional regulator	P6_peg280	DSM44509_peg378
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Phytochrome, two-component sensor histidine kinase (EC 2.7.3.-)	P6_peg28, P6_peg2209	DSM44509_peg2040, DSM44509_peg4012
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Redox-sensitive transcriptional regulator (AT-rich DNA-binding protein)	P6_peg3266, P6_peg3267	DSM44509_peg3930, DSM44509_peg3931
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Transcriptional regulator, FUR family	P6_peg612, P6_peg3592	DSM44509_peg3526, DSM44509_peg3714
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Zinc uptake regulation protein ZUR	P6_peg1798	DSM44509_peg407, DSM44509_peg1351, DSM44509_peg4073
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	bacteriophytochrome heme oxygenase BphO	P6_peg29	DSM44509_peg4013
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Catalase (EC 1.11.1.6)	P6_peg613	DSM44509_peg1819, DSM44509_peg3715
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	P6_peg874	DSM44509_peg3078
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	NAD-dependent protein deacetylase of SIR2 family	P6_peg1289, P6_peg2601	DSM44509_peg219, DSM44509_peg3132
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Nicotinamidase (EC 3.5.1.19)	P6_peg366	DSM44509_peg1160, DSM44509_peg2303
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	Oxidative stress	Nicotinate phosphoribosyltransferase (EC 2.4.2.11)	P6_peg365	DSM44509_peg2304
P6 <sup>T</sup>	Oxidative stress	CoA-disulfide reductase (EC 1.8.1.14)	P6_peg1645	
DSM 44509 <sup>T</sup>	Oxidative stress	Peroxidase (EC 1.11.1.7)		DSM44509_peg3715
DSM 44509 <sup>T</sup>	Oxidative stress	transcriptional regulator, Crp/Fnr family		DSM44509_peg2379



P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Flavoheomoprotein (Hemoglobin-like protein) (Flavoheomoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)	P6_peg3406	DSM44509_peg3484, DSM44509_peg3485
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Hemoglobin-like protein HbO	P6_peg1865	DSM44509_peg685
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	P6_peg301, P6_peg436, P6_peg521, P6_peg665, P6_peg699, P6_peg726, P6_peg1016, P6_peg1019, P6_peg1054, P6_peg1080, P6_peg1176, P6_peg1199, P6_peg1204, P6_peg1205, P6_peg1216, P6_peg1449, P6_peg1524, P6_peg1682, P6_peg1833, P6_peg1919, P6_peg1934, P6_peg1935, P6_peg2120, P6_peg2153, P6_peg3156, P6_peg3327, P6_peg3329, P6_peg3342, P6_peg3368, P6_peg3376, P6_peg3647, P6_peg3742	DSM44509_peg440, DSM44509_peg460, DSM44509_peg522, DSM44509_peg523, DSM44509_peg538, DSM44509_peg739, DSM44509_peg915, DSM44509_peg916, DSM44509_peg924, DSM44509_peg947, DSM44509_peg1062, DSM44509_peg1369, DSM44509_peg1372, DSM44509_peg1510, DSM44509_peg1787, DSM44509_peg1791, DSM44509_peg1814, DSM44509_peg1826, DSM44509_peg1922, DSM44509_peg2018, DSM44509_peg2150, DSM44509_peg2427, DSM44509_peg2461, DSM44509_peg2476, DSM44509_peg2477, DSM44509_peg2814, DSM44509_peg2873, DSM44509_peg3354, DSM44509_peg3770, DSM44509_peg3968, DSM44509_peg4177, DSM44509_peg4271
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Carbon starvation protein A	P6_peg1951	DSM44509_peg975
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Carbon storage regulator	P6_peg3731	DSM44509_peg4059
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	ABC-type Fe <sup>3+</sup> -siderophore transport system, permease 2 component	P6_peg780, P6_peg1125	DSM44509_peg709, DSM44509_peg3042
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Flavoheomoprotein (Hemoglobin-like protein) (Flavoheomoglobin) (Nitric oxide dioxygenase) (EC 1.14.12.17)	P6_peg3406	DSM44509_peg3484, DSM44509_peg3485
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	GTP-binding protein HflX	P6_peg208	DSM44509_peg1601
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Anti-sigma B factor antagonist RsbV	P6_peg1497	DSM44509_peg715, DSM44509_peg716, DSM44509_peg1537
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	RNA polymerase sigma factor SigB	P6_peg600, P6_peg2515	DSM44509_peg2957, DSM44509_peg3720
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Serine phosphatase RsbU, regulator of sigma subunit	P6_peg602, P6_peg1036, P6_peg1041, P6_peg1190, P6_peg1192, P6_peg1237, P6_peg1322, P6_peg2184, P6_peg2763, P6_peg2953, P6_peg2961, P6_peg3019	DSM44509_peg298, DSM44509_peg490, DSM44509_peg714, DSM44509_peg1766, DSM44509_peg2084, DSM44509_peg2260, DSM44509_peg2460, DSM44509_peg2490, DSM44509_peg2491, DSM44509_peg2564, DSM44509_peg2572, DSM44509_peg2959, DSM44509_peg3627, DSM44509_peg3647, DSM44509_peg4104, DSM44509_peg4267, DSM44509_peg4269
P6 <sup>T</sup> and DSM 44509 <sup>T</sup>	no subcategory	Serine-protein kinase RsbW (EC 2.7.11.1)	P6_peg601	DSM44509_peg2958
DSM 44509 <sup>T</sup>	no subcategory	Putative SigmaB associated two-component system sensor protein		DSM44509_peg2261