

Characterization of the microaerophilic, bacteriochlorophyll *a*-containing bacterium *Gemmatimonas phototrophica* sp. nov., and emended descriptions of the genus *Gemmatimonas* and *Gemmatimonas aurantiaca*

Yonghui Zeng,¹ Vadim Selyanin,¹ Martin Lukeš,¹ Jason Dean,¹ David Kaftan,^{1,2} Fuying Feng³ and Michal Koblížek^{1,2}

¹Center Algatech, Institute of Microbiology CAS, Opatovický mlýn, 379 81 Třeboň, Czech Republic

²Faculty of Science, University of South Bohemia, 370 05 České Budějovice, Czech Republic

³Institute for Applied & Environmental Microbiology, Inner Mongolia Agricultural University, 306 Zhaowuda Road, 010 018 Huhhot, PR China

Correspondence

Yonghui Zeng
yonghui.sci@gmail.com
or
Michal Koblížek
koblizek@alga.cz

A red-pigmented, bacteriochlorophyll (BChl) *a*-producing strain, AP64^T, was isolated previously from the freshwater Swan Lake located in the western Gobi Desert. Based on its 16S rRNA gene sequence identity (96.1 %) to the type strain *Gemmatimonas aurantiaca* T-27^T, the new isolate was tentatively classified as a member of the bacterial phylum *Gemmatimonadetes*. Here, we report its formal description and polyphasic characterization. Strain AP64^T grew best on agar media under 9.8–15.2 % atmospheric oxygen. The cells were rods, dividing by symmetrical or asymmetrical binary fission. Budding structures were also observed. Its genomic DNA G + C content was 64.4 % (from the draft genome sequence). Phylogenetic analysis based on the 16S rRNA gene sequence clearly separated AP64^T from related species. Its genotypic differentiation from phylogenetically close relatives was further supported by performing *in silico* DNA–DNA hybridization and calculating average nucleotide identity, whereas the high percentage (67.3 %) of shared conserved proteins between strain AP64^T and *Gemmatimonas aurantiaca* T-27^T supports the classification of the two strains into the same genus. Strain AP64^T contained C_{16:1}, C_{14:1} and C_{18:1}ω9c as predominant fatty acids. The main respiratory quinone was menaquinone 8 (MK-8). The most distinctive feature of strain AP64^T was the presence of fully functional purple bacterial photosynthetic reaction centres. The main CO₂-fixation pathways were absent. Strain AP64^T was capable of growth and BChl production in constant darkness. Thus, strain AP64^T is a facultatively photoheterotrophic organism. It represents a novel species of the genus *Gemmatimonas*, for which the name *Gemmatimonas phototrophica* sp. nov. is proposed. The type strain is AP64^T (=DSM 29774^T=MCCC 1K00454^T). Emended descriptions of the genus *Gemmatimonas* and *Gemmatimonas aurantiaca* are also provided.

INTRODUCTION

Members of the phylum *Gemmatimonadetes* are distributed widely across various natural environments (DeBruyn

Abbreviations: AFM, atomic force microscopy; ANI, average nucleotide identity; BChl, bacteriochlorophyll; FTIR, Fourier-transform infrared; iDDH, *in silico* DNA–DNA hybridization; POCP, percentage of conserved proteins; WGA, whole-genome alignment.

The GenBank/EMBL/DDBJ accession number for the 16S rRNA gene sequence of strain AP64^T is KF481682. The accession number for the genome sequence of strain AP64^T is AUXF00000000.

et al., 2011; Hanada & Sekiguchi, 2014; Zhang *et al.*, 2003). Environmental 16S rRNA gene sequence surveys have revealed that they inhabit various terrestrial (e.g. soils, activated sludge and saline cave waters) as well as marine (e.g. deep-sea sediments, gas hydrates, arctic seawater, coastal mobile mud and marine sponge symbionts) habitats (Kamagata, 2010). High-throughput 16S rRNA gene sequencing indicated that members of the phylum *Gemmatimonadetes* represent a recognizable fraction of soil microbial communities, with relative abundances ranging from 0.2 to 6.5 % (DeBruyn *et al.*, 2011). The

phylogenetic divergence of the members of this phylum can be as large as that in the phyla *Actinobacteria* and *Proteobacteria* (Hanada & Sekiguchi, 2014). However, the physiology and environmental role of members of the phylum *Gemmatisationadetes* are largely unknown because of the very limited number of cultivated species. So far, two genera, *Gemmatisation* and '*Gemmatisation*', the name of which has not yet been validly published, have been proposed in this phylum, with only the type species and type strain reported in each genus, i.e. *Gemmatisation aurantiaca* T-27^T (Zhang *et al.*, 2003), isolated from a wastewater treatment reactor, and '*Gemmatisation kalamazonensis*' KBS708 (DeBruyn *et al.*, 2013), isolated from organically managed agricultural soil in the USA.

As reported previously (Zeng *et al.*, 2014), we isolated a bacteriochlorophyll (BChl) *a*-containing bacterium designated AP64^T from a freshwater desert lake located in Inner Mongolia, northern China. Its 16S rRNA gene shares 96.1 % sequence identity with that of *Gemmatisation aurantiaca* T-27^T, while it shows only 90.1 % identity to that of '*Gemmatisation kalamazonensis*' KBS708, suggesting that AP64^T and *Gemmatisation aurantiaca* T-27^T belong to the same genus, whereas '*Gemmatisation kalamazonensis*' KBS708 represents a different class of the phylum *Gemmatisationadetes*. Here, we attempt to employ a polyphasic approach to characterize the genotypic and phenotypic features of strain AP64^T further in comparison to the only known strain of the genus *Gemmatisation*, *Gemmatisation aurantiaca* T-27^T. Comparative genomics was also performed to corroborate the taxonomic classification. The name *Gemmatisation phototrophica* sp. nov. is proposed for this isolate.

METHODS

Isolation, enrichment and laboratory cultivation. Details on source environment, isolation strategy and preliminary molecular identification were reported by Zeng *et al.* (2014). Briefly, near-shore lake water was sampled at a depth of 0.5 m on 15 December 2011 from the freshwater Swan Lake (42.005° N 101.585° E), which is located at the northern margin of the western Gobi Desert. Lake water was diluted 100 times and then plated onto half-strength standard R2A agar plates (DSMZ medium 830). Colonies formed after 3–6 weeks of incubation at 25 °C under a 12 : 12 h light/dark regime. BChl *a*-containing colonies were identified based on their near-infrared fluorescence using a modified FluorCam 800MF fluorescence imaging system (PSI Ltd) (for the principles and detailed design of the system, see Zeng *et al.*, 2014). A red-pigmented strain, AP64^T, was selected for further characterization because of its high 16S rRNA gene sequence similarity to a member of the rare bacterial phylum *Gemmatisationadetes*.

The optimized growth medium R2A⁺ was established for strain AP64^T (Zeng *et al.*, 2014), containing (l⁻¹) 0.5 g yeast extract, 0.5 g proteose peptone, 0.5 g glucose, 0.5 g soluble starch, 0.3 g sodium pyruvate, 0.3 g K₂HPO₄, 0.05 g MgSO₄, 1.0 ml vitamin solution (refer to DSMZ medium 462), 0.05 ml trace element solution SL-6 (DSMZ medium 27) and 15 g agar. The pH was adjusted to 7.5–8.0 by addition of NaOH solution. Optimum growth was achieved after 10–14 days of incubation under microaerophilic conditions. Optimum oxygen concentration for growth was determined by inoculating cells into soft-agar R2A⁺ medium [0.6 % (w/v) agar]. The oxygen

profile of the deep agar was measured with a home-made Clark-type oxygen microelectrode at the Department of Biology of the University of Southern Denmark. The oxygen tension was originally reduced by evacuating the inoculated agar plates in a glass exsiccator [700 mbar (70 kPa)]. Later, the strain was grown in a Memmert INCO 108med incubator under an 88 % N₂, 10 % O₂ and 2 % CO₂ atmosphere at 28 ± 1 °C. Illumination (12 : 12 h light/dark regime) was provided by cool white LEDs. No or very slow growth (> 3 months) was observed in the corresponding liquid medium, as we reported before (Zeng *et al.*, 2014). To investigate the possible influence of the light source on growth, we compared the use of LED lights, fluorescence lamps and tungsten lamps, but we did not observe any difference in the growth of strain AP64^T on agar plates (see below for details on comparing the growth of colonies on agar plates). A 1-year experiment of cultivating AP64^T under continuous darkness was performed by transferring fresh colonies onto new agar plates every 2 months. Growth was monitored by measuring colony size and by observing their colour intensity.

Microscopy. Cells were scraped from 2-week-old R2A⁺ agar plates and resuspended into liquid R2A⁺ medium. The cells were pre-adapted in the liquid medium for a couple of hours, and the samples were then fixed with 2 % formaldehyde, collected onto 0.2 µm polycarbonate filters, dried and stained with 4',6-diamidino-2-phenylindole (DAPI) dissolved in a 3 : 1 mixture of Citifluor AF1 and Vectashield at a final concentration of 1 mg ml⁻¹. Infrared epifluorescence microscopy was performed using a Zeiss Axio Imager. D2 microscope equipped with a Hamamatsu EMCCD camera C9100-02. An infrared emission (> 850 nm) image was captured (200 ms–1 s exposure) to confirm the expression of BChl *a* in the cells. Cell morphology was examined with both DAPI and infrared images. Scanning and transmission electron microscopy images were obtained with a JEOL JSM-7401F scanning electron microscope and a JEOL JEM-1010 transmission electron microscope, respectively, using standard protocols, at the Laboratory of Electron Microscopy, Biology Centre of ASCR, České Budějovice, Czech Republic. For atomic force microscopy (AFM), cells of strain AP64^T were resuspended in a buffer containing 10 mM HEPES, pH 7.3, 150 mM NaCl and 5 mM MgCl₂. The cells were immobilized on a gelatin-coated muscovite mica by applying 50 µl cell suspension onto the dry gelatin surface. Following 30 min of immobilization, the excess liquid was removed and the cells were imaged immediately in contact mode in liquid (whole cells) or in air (high-resolution imaging of the cell surface) with a Pico SPM atomic force microscope (Agilent Technologies) equipped with a small 5.5 µm scan-size scanner using MSCT silicon nitride probes (Bruker; nominal spring constant 0.01 N m⁻¹) with a resolution of 1024 × 1024 pixels. The scan rate was 1–3 lines per second. All images were only first-order flattened.

Phylogenetic analyses. The 16S rRNA gene sequence of strain AP64^T was retrieved from its genome sequence (GenBank accession no. AUXF00000000). High-quality reference sequences (length > 1450 nt with no ambiguous positions and having been subjected to a chimera check) were retrieved from the Greengenes database (DeSantis *et al.*, 2006; <http://greengenes.lbl.gov/>). Highly similar sequences with ≤ 0.5 % divergence were removed prior to the following analyses. Finally, 284 high-quality reference sequences were obtained. Multiple alignment was performed using the SILVA Incremental Aligner tool available on the SILVA website (<http://www.arb-silva.de/>), which takes advantage of a high-quality pre-aligned rRNA database and is supposed to provide very high alignment accuracy (Pruesse *et al.*, 2012). Highly variable positions in the alignment (a nucleotide position shared in < 70 % sequences) were excluded for inference of phylogenetic trees. Finally, 1130 conserved positions were used to reconstruct phylogenetic trees using the neighbour-joining algorithm implemented in the MEGA 6.0 software (Tamura *et al.*, 2013). A maximum composite likelihood model was used for inferring the neighbour-joining tree. A uniform rate of

nucleotide substitution was used. To estimate the confidence of tree topologies, bootstrap resampling for 1000 replicates was performed.

DNA base composition, *in silico* DNA–DNA hybridization (iDDH) and average nucleotide identity (ANI). The genomic G + C content of strain AP64^T was calculated based on its draft genome sequence (4.7 Mbp). iDDH was performed between AP64^T and reference strains using the Genome-to-Genome Distance Calculator (GGDC 2.0) web server (Auch *et al.*, 2010; Meier-Kolthoff *et al.*, 2013). Genome sequences of *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazoonensis*' KBS708 were retrieved from GenBank with accession nos AP009153.1 and CP007128.1, respectively. To corroborate the iDDH results, ANI, which represents the overall level of similarity between two genome sequences (Goris *et al.*, 2007; Richter & Rosselló-Móra, 2009), was calculated using the online ANI calculation tool on the EzBiocloud web server (Kim *et al.*, 2012).

Physiological and biochemical characterization. Since strain AP64^T only grew well on agar plates, we standardized the procedure of strain inoculation in order to enable comparison of growth on different agar plates: (i) use an inoculation loop (1 mm in diameter) to harvest cells from a freshly grown plate (2 weeks old); (ii) suspend cells into 1 ml R2A⁺ liquid medium; (iii) perform serial dilution of the cell suspension by factors of 10⁰, 10² and 10⁴; (iv) for each agar plate, drip 10 µl of each cell dilution onto the agar surface and allow to air dry for 10–20 min to form a series of strain spots with a diameter of ~6 mm; (v) compare the size and colour intensity of strain spots on different plates after 2–4 weeks of incubation. All growth and substrate utilization tests were performed with a tungsten lamp as a light source. The comparison strain T-27^T was purchased from DSMZ, Germany (DSM 14586).

Growth at 0, 0.5, 1, 2, 4, 8, 12 and 15 g NaCl l⁻¹ was examined on R2A⁺ agar plates. The pH range for growth was investigated at pH 4–11 in increments of 1 pH unit, with an additional test at pH 7.5. The following buffer solutions were used: acetate buffer solutions (acetic acid/sodium acetate) for pH 4–6, NaH₂PO₄/Na₂HPO₄ for pH 6–8 and phosphate-base buffer solutions (Na₂HPO₄/NaOH) for pH 8–11. Cell motility was determined by inoculating cultures with a needle in a straight-line stab into semi-solid R2A⁺ medium (0.6 % agar, w/v) and then observing the medium turbidity after 4 weeks of incubation. Growth at 4, 16, 20, 25, 30 and 37 °C was examined on R2A⁺ agar plates with an incubation time of 4 weeks. Catalase activity was determined by observing bubble production in a 3 % H₂O₂ solution. Oxidase activity was determined by monitoring the oxidation of tetramethyl *p*-phenylenediamine dihydrochloride on filter paper. Fermentative growth was examined on R2A⁺ slants in cap-tightened bottles with the head atmosphere replaced with nitrogen gas. Anaerobic phototrophic growth was tested under the same conditions, but the R2A⁺ agar was supplemented with common electron donors for phototrophs: 10 mM Na₂S, 5 mM Na₂S₂O₃ and 5 mM Na₂SO₃. For tests of anaerobic growth based on anaerobic respiration, R2A⁺ agar was supplemented with 0.2 g KNO₃, 30 mM DMSO or 30 mM trimethylamine *N*-oxide (TMAO). Antibiotic susceptibility tests were performed by comparing growth at different regions of an agar plate with 20 µl solution of test antibiotic spread onto the centre of the agar surface (~1 cm in diameter). Carbon source utilization was assayed under aerobic and illuminated conditions using a basal medium supplemented with 0.5 g of each test substrate l⁻¹, as described previously by Zhang *et al.* (2003). The basal medium contained (l⁻¹) 0.44 g KH₂PO₄, 0.1 g (NH₄)₂SO₄, 0.1 g MgSO₄·7H₂O and 1 ml vitamin solution (refer to DSMZ medium 462), pH 7.5.

Other analyses. Cells were harvested from 2-week-old agar plates. For spectroscopic analyses, the cells were suspended in 70 % glycerol to reduce scattering. *In vivo* absorption spectra were recorded using a Shimadzu UV3000 spectrophotometer. For pigment analyses, the cells

were extracted with 1 ml 100 % methanol for 5 min. The extracts were analysed using an Agilent 1100 Series HPLC-MS/MS system, as described before (Koblížek *et al.*, 2010). Respiratory quinones were extracted using two-phase extraction with 3 ml methanol/hexane (1 : 2, v/v) followed by addition of 2 ml hexane and 2 ml NaCl solution in water (3 g l⁻¹). The upper, hexane layer was collected and evaporated with nitrogen. The collected fraction was dissolved in methanol/heptane (10 : 2, v/v) and analysed by isocratic reversed-phase chromatography using a C18 column with methanol/heptane (10 : 2, v/v) as the mobile phase. Chromatography was performed using a Shimadzu Prominence-i system equipped with a photodiode detector. Quinones were detected at 275 nm and identified based on retention times and UV absorption spectra.

For fatty acid analysis, cells were scraped carefully from 2-week-old agar plates to avoid agar contamination and were suspended in 800 µl distilled water. Two millilitres methanol and 1 ml dichloromethane were added, and the suspension was sonicated for 10 min in a cooled sonication bath. After sonication, 1 ml distilled water and 1 ml dichloromethane were added and the suspension was briefly vortexed. After centrifugation at 500 *g* for 10 min at 4 °C, the organic phases separated. The lower, dichloromethane phase was dried and the extract was diluted in dichloromethane/methanol (2 : 1, v/v) at a concentration of 1 mg lipid extract per 50 µl and stored at -70 °C until further analysis. Methyl esters of fatty acids were prepared using a method described previously (Kainz *et al.*, 2002). Quantitative and qualitative analysis was performed by means of GC-FID on TRACE 1300 (Thermo Finnigan) as described previously (Lukeš *et al.*, 2014). Amounts of individual fatty acids were calculated by the use of known quantities of internal standards and corrected by multiplication of integrated peaks with correction factors of FID response.

For Fourier-transform infrared (FTIR) spectroscopy, cells scraped from agar plates were suspended in 100 µl distilled water and 5 µl of the suspension was spread immediately on a single well of a Si 384-well plate. Pure distilled water was used as a blank for background subtraction. Samples were measured on a Nicolet IS10 spectrometer (Thermo Nicolet) equipped with a microarray reader compartment and a DTGS detector. The compartment with the detector was flushed continuously with dry nitrogen to avoid moisture. Absorbance spectra were collected between 400 and 4000 cm⁻¹ at a spectral resolution of 4 cm⁻¹ and 128 scans were co-added and averaged. A Blackman–Harris three-term apodization function was used with a zero-filling factor of 2. Spectra were baseline-corrected by the use of the rubber-band method.

Comparative genomics. To eliminate discrepancies in microbial genome annotation results and the associated problems in data comparison caused by the application of different annotation platforms, genome sequences of strain AP64^T, *Gemmatimonas aurantiaca* T-27^T (GenBank accession no. AP009153) and '*Gemmatirosa kalamazoonensis*' KBS708 (CP007128) were reannotated with the microbial genome annotation web server RAST (Aziz *et al.*, 2008). The percentage of conserved proteins (POCP) (Qin *et al.*, 2014) in each pair of genomes was calculated as [(C1 + C2)/(T1 + T2)] × 100 %, where C1 and C2 represent the numbers of conserved proteins in the two genomes when each serves as the query genome, and T1 and T2 represent the total numbers of predicted proteins in the two genomes. Conserved proteins were defined as having a BLASTP match with an E-value of less than 1e-5, sequence identity of more than 40 % and an alignment coverage of the query protein sequence of more than 50 %, as recommended by Qin *et al.* (2014). Whole genome alignments (WGA) were performed and visualized in the Mauve software (version 2.3.1) (Darling *et al.*, 2010). For WGA analysis, the draft genome sequence of strain AP64^T, which consisted of seven non-contiguous contigs, was reordered into a fake chromosome sequence

with reference to the chromosome sequence of *Gemmatimonas aurantiaca* T-27^T. Orthologues were calculated based on progressive Mauve alignments using the cut-off value >50 % for both sequence identity and alignment coverage. Both annotated and unannotated genes that were predicted to be positionally orthologous by whole-genome multiple alignment were counted.

RESULTS AND DISCUSSION

Cultivation and physiology

Strain AP64^T was isolated from shallow near-shore water of a freshwater desert lake, the Swan Lake, in northern China. It formed tiny (~0.3 mm), round, smooth and red-pigmented colonies on agar plates after incubation at 28 °C for 2 weeks. It grew very slowly (>3 months) in the corresponding liquid medium, and therefore all comparison experiments performed in this work were based on cells grown on agar plates.

Growth of strain AP64^T occurred at 16–30 °C, with an optimum temperature of 25–30 °C. The pH range for growth was 6.0–9.0, with an optimum at pH 7.5–8.0. The strain did not require NaCl for growth, but it tolerated up to 2 g NaCl l⁻¹. Deletion tests of components in the R2A⁺ medium that we used before (Zeng *et al.* 2014) demonstrated that soluble starch, MgSO₄ and ammonium acetate were not indispensable, since no noticeable growth delay was observed when they were absent from the medium. Therefore, we simplified the formula of the R2A⁺ medium further to (l⁻¹) 0.5 g yeast extract, 0.5 g proteose peptone, 0.2 g glucose, 0.3 g sodium pyruvate, 0.3 g K₂HPO₄, 1.0 ml vitamin solution and 0.05 ml trace element solution. Glucose seemed to have only a minimal effect, and therefore its amount was reduced. Neither AP64^T nor *Gemmatimonas aurantiaca* T-27^T yielded comparable biomass on nutrient-rich agar plates, e.g. Luria–Bertani agar, nutrient agar and double-nutrient-strength R2A⁺, to that on R2A⁺ agar plates, suggesting that copiotrophic growth is not preferred by these two members of the phylum *Gemmatimonadetes*.

We tested the oxygen preference of AP64^T using the deep-agar culture method, and we found that the strain grew optimally under reduced (9.8–15.2 %) atmospheric oxygen; at normal oxygen concentration (21 %), its growth was inhibited significantly. No growth was observed under any tested anaerobic conditions in a 4-week incubation. The highest yield of biomass for strain AP64^T was reached after 10–14 days of growth under micro-oxic conditions. In contrast, *Gemmatimonas aurantiaca* T-27^T preferred a fully oxic atmosphere, and reached maximum biomass on agar plates after 5–7 days.

Morphology

Cells of strain AP64^T were rod-shaped, 0.3–0.5 µm wide and most commonly 1–6 µm long (Fig. 1). Filaments up to 12 µm long were documented (Fig. 1a). Similar filamentous cells

were also observed for ‘*Gemmatirosa kalamazoonensis*’ KBS708 (DeBruyn *et al.*, 2013) but were not reported for *Gemmatimonas aurantiaca* T-27^T (Zhang *et al.* 2003). Structures of a Gram-negative cell envelope were observed with transmission electron microscopy (Fig. 1b). Electron-dense small bodies and transparent, large, vesicle-like structures were present inside cells (Fig. 1b). Cells reproduced by binary fission (Fig. 1b) and often showed budding morphology (Fig. 1c), similar to that reported in *Gemmatimonas aurantiaca* T-27^T (Zhang *et al.* 2003) and ‘*Gemmatirosa kalamazoonensis*’ KBS708 (DeBruyn *et al.*, 2013). An ongoing ternary fission through budding was observed (Fig. 1c), indicating the capability of strain AP64^T to perform various types of cell division. AFM images showed hexagonal structures (~30 nm in diameter) at the cell surface resembling an S-layer (Fig. 1d). Cells were motile, consistent with the fact that its genome contains flagellar biosynthesis genes, although flagella were not observed directly by electron microscopy.

Biochemical and chemotaxonomic characteristics

Strain AP64^T and *Gemmatimonas aurantiaca* T-27^T could be differentiated by colony size and colour, optimum pH, optimum salinity and oxygen concentrations for growth, capacity to grow in liquid medium, susceptibility to antibiotics and the species of pigments contained (Table 1). Strain AP64^T grew well with yeast extract as a sole carbon source. Weak growth with peptone was observed. During 30 days of incubation, no growth was observed with the following compounds as the sole carbon source: Casamino acids, sodium succinate, sodium acetate, sodium pyruvate, potato starch, sucrose, L-glutamic acid, L-leucine, L-arginine, L-alanine, L-isoleucine, L-arabinose, D-sorbitol and D-mannitol. Strain AP64^T exhibited natural resistance to ampicillin, penicillin, paramycin sulfate, polymyxin B sulfate and nystatin, but it was sensitive to neomycin, vancomycin, bacitracin and gentamicin. The major respiratory quinone was menaquinone 8 (MK-8). The dominant fatty acids were C_{16:1}, C_{14:1} and C_{18:1ω9c} (Table 2). The dominant fatty acid species observed in *Gemmatimonas aurantiaca* T-27^T (C_{14:1}, C_{16:1} and C_{16:0}) differ from those originally reported by Zhang *et al.* (2003), i.e. iso-C_{15:0}, C_{16:1} and C_{14:0}. The reason could be the application of different growth conditions [agar plates in this study vs liquid medium in Zhang *et al.* (2003)], since the fatty acid composition can be strongly influenced by the growth state and conditions of the bacterial cells.

The cells were positive for oxidase and catalase. The G + C content of strain AP64^T was calculated as 64.4 % based on the draft genome sequence. Cells of strain AP64^T and *Gemmatimonas aurantiaca* T-27^T had a very similar composition of basic macromolecules, as seen from FTIR spectra (Fig. 2). The two strains contained carotenoids of the oscillol series, with oscillol 2,2'-dirhamnosides as the major carotenoid (Takaichi *et al.*, 2010; Zeng *et al.*, 2014). In addition, strain AP64^T contained BChl *a* and spirilloxanthin, which were absent from *Gemmatimonas aurantiaca* T-27^T (Zeng *et al.*, 2014). Interestingly, cells that were grown continuously in

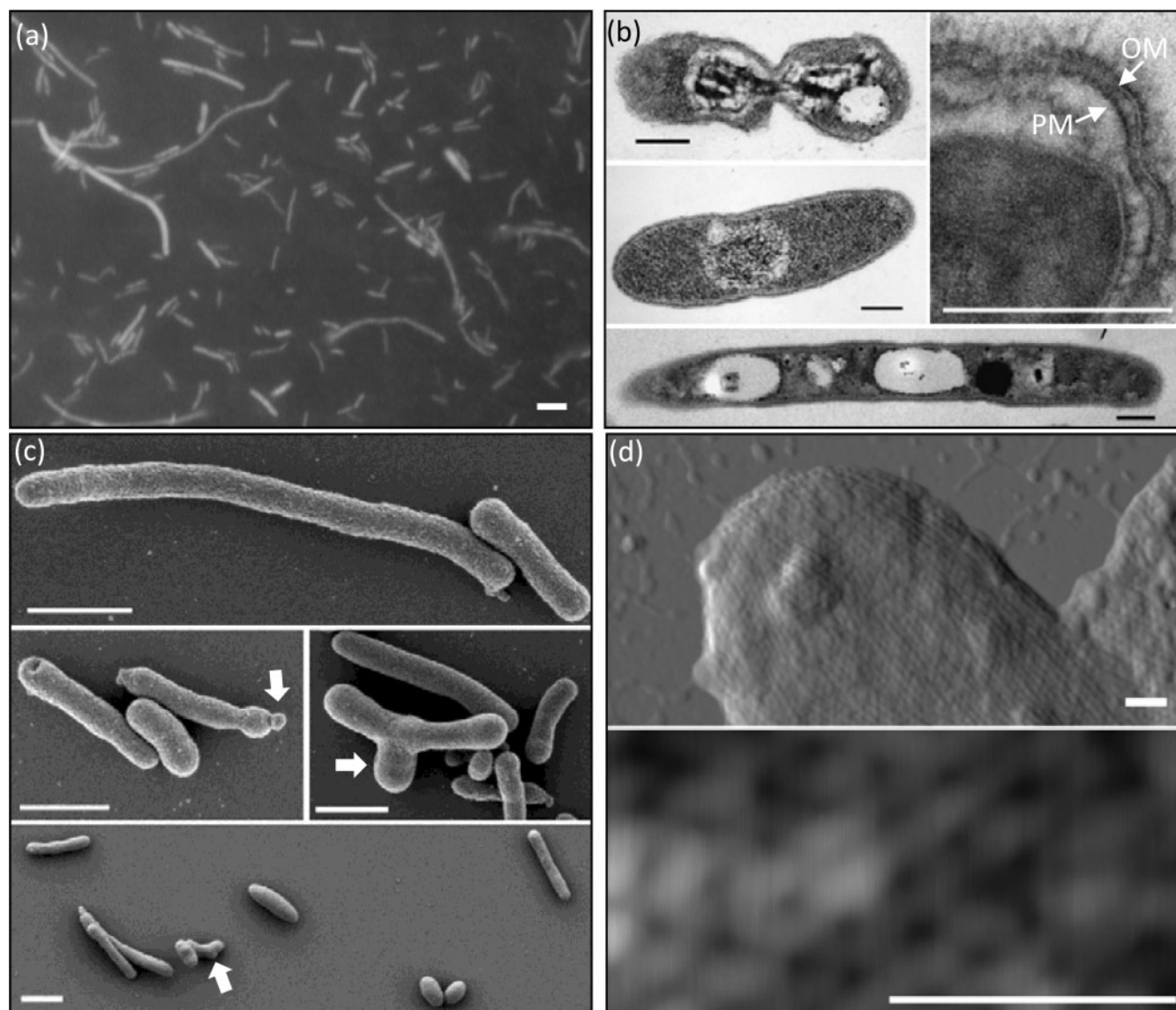


Fig. 1. Microscope images of cells of strain AP64^T. (a) Infrared fluorescence image showing rod- and filament-shaped cells. (b) Transmission electron micrographs. OM, Outer membrane; PM, plasma membrane. (c) Scanning electron micrographs. Budding structures are marked with arrows. (d) AFM deflection images showing porous surface layer. Bars, 1 μ m (a, c), 200 nm (b), 100 nm (d).

the dark and subcultured for 1 year still maintained BChl *a* synthesis, as seen from the absorption spectra (Fig. 3). Light did not inhibit BChl *a* synthesis. This shows that the photosynthetic apparatus was expressed constitutively in cells of strain AP64^T.

Phylogeny and genomic traits

Phylogenetic analysis of 16S rRNA gene sequences indicated that strain AP64^T was closely related to *Gemmatimonas aurantiaca* T-27^T (96.1 % sequence identity), forming a tight phylogenetic cluster (Fig. 4), but was distantly related to

'*Gemmatirosa kalamazonensis*' KBS708 (91.2 % sequence identity). Further evidence came from *in silico* comparison of genomic sequences of these three strains. ANI represents mean identity values between a given pair of genomes, and has been proposed as a robust parameter that can replace 16S rRNA gene sequence comparison for species delineation in the genomic era (Goris *et al.*, 2007; Konstantinidis & Tiedje, 2005). ANI between strain AP64^T and *Gemmatimonas aurantiaca* T-27^T and between strain AP64^T and '*Gemmatirosa kalamazonensis*' KBS708 was 73.51 ± 0.06 and 68.08 ± 0.02 %, respectively, much less than the proposed boundary of 95–96 % for defining a novel species (Goris

Table 1. Comparison of characteristics of members of the genus *Gemmatimonas*

Both strains form round, smooth colonies and motile cells. Both strains grow at 16–30 °C, with optimum growth at 25–30 °C. Both strains are catalase- and oxidase-positive and are susceptible to neomycin, bacitracin, gentamicin and vancomycin and resistant to ampicillin, nystatin, penicillin and paramycin sulfate. Neither strain requires Na⁺ ions for growth, and both strains show optimum growth in the absence of added NaCl. Neither strain shows anaerobic or autotrophic growth or grows on nutrient-rich media. Data are from this study unless indicated.

Characteristic	AP64 ^T	<i>Gemmatimonas aurantiaca</i> T-27 ^T
Colony properties (20 days)		
Diameter (mm)	~0.2–0.4	~1.0–2.0
Colour	Red	Light pink
Cell properties		
Shape	Short to long rod	Short rod*
Width (µm)	0.3–0.5	0.7*
Length (µm)	1–12	2.5–3.2*
Capsule-like cellular structure	Yes	No*
DNA G + C content (%)†	64.4	64.3
Genome size (Mbp)†	4.70 (draft)	4.64 (complete)
pH for growth		
Range	6.0–9.0	6.0–10.0
Optimum	7.5–8.0	7.0–7.5
NaCl concentration for growth (g l ⁻¹)	0–2.0	0–8.0
Oxygen requirement		
Preferred mode	Microaerophilic	Aerobic
Other tolerated mode	None	Microaerophilic
Grows well in liquid medium	No	Yes
Major quinone	MK-8	MK-9
Polymyxin B sulfate susceptibility	–	+
Pigment biosynthesis		
Oscillo 2,2'-dirhamnoside	Yes‡	Yes*‡
Spirilloxanthin	Yes	No
BChl <i>a</i>	Yes	No

*Data from Zhang *et al.* (2003).

†Based on published genome sequences.

‡Data from Zeng *et al.* (2014).

et al., 2007). Similarly, the iDDH values (19.30 ± 2.29 % between strain AP64^T and *Gemmatimonas aurantiaca* T-27^T and 18.70 ± 2.27 % between strain AP64^T and '*Gemmatirosa kalamazonensis*' KBS708) were significantly lower than the traditional cut-off value 70 % used for species delineation. Recently, a boundary value of 50 % POCP was proposed for genus definition (Qin *et al.*, 2014). We calculated POCP values between AP64^T and *Gemmatimonas aurantiaca* T-27^T (67.3 %), between AP64^T and '*Gemmatirosa kalamazonensis*' KBS708 (47.9 %) and between *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazonensis*' KBS708 (47.6 %). In summary, the evidence from analyses of phylogeny and genomic similarity implies that strain AP64^T and *Gemmatimonas aurantiaca* T-27^T belong to the same genus, whereas '*Gemmatirosa kalamazonensis*' KBS708 represents a different genus.

We compared genomic similarity and calculated the numbers of shared orthologues among strain AP64^T, *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazonensis*' KBS708 (Fig. 5). The genomes of AP64^T and

Table 2. Cellular fatty acids of strain AP64^T and *Gemmatimonas aurantiaca* T-27^T

Cells were harvested from 2-week-old agar plates. Values are mean percentages \pm SD from triplicate measurements, determined in this study. ND, Not detected; TR, trace amount.

Fatty acid	AP64 ^T	<i>Gemmatimonas aurantiaca</i> T-27 ^T
C _{12:1}	1.1 \pm 0.6	0.5 \pm 0.1
C _{14:0}	3.2 \pm 0.7	2.7 \pm 0.4
C _{14:1}	15.5 \pm 1.8	44.9 \pm 2.3
C _{15:1}	TR	ND
C _{16:0}	12.1 \pm 3.8	12.3 \pm 1.6
C _{16:1}	47.7 \pm 7.6	31.8 \pm 2.1
C _{16:2}	1.9 \pm 0.4	1.3 \pm 0.1
C _{18:1} ω _{9t}	3.5 \pm 1.1	1.7 \pm 0.7
C _{18:1} ω _{9c}	14.3 \pm 9.1	1.9 \pm 0.4
C _{18:1} ω ₇	0.4 \pm 0.1	2.3 \pm 1.7
C _{18:2} ω ₆	0.1 \pm 0.2	0.3 \pm 0.1

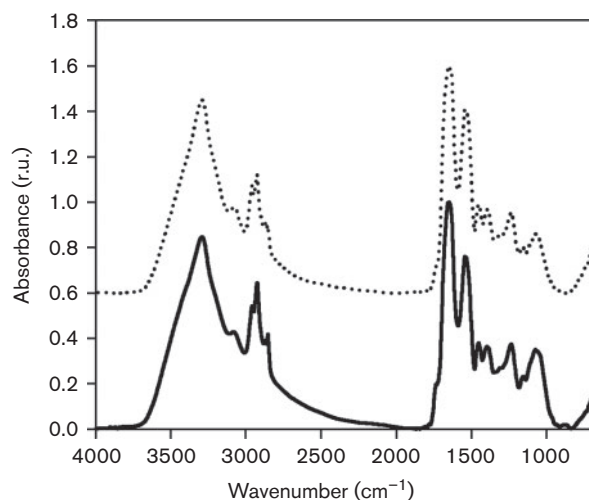


Fig. 2. FTIR spectra of cells of strain AP64^T (solid line) and *Gemmatimonas aurantiaca* T-27^T (dotted line), demonstrating the high similarity in their cellular chemical compositions.

Gemmatimonas aurantiaca T-27^T were highly conserved in terms of both genomic similarity and gene synteny (Fig. 5, left), with no large genomic rearrangement events detected. In contrast, a large portion of the genome in AP64^T and '*Gemmatirosa kalamazoonensis*' KBS708 has been subjected to rearrangement, and less than 50 % genomic regions were found to be conserved. A similar pattern was seen when comparing the numbers of shared orthologues in each pair of genomes (Fig. 5, right). The number of unique orthologues (1010) shared by AP64^T and *Gemmatimonas aurantiaca* T-27^T was much greater than the numbers shared by either AP64^T and '*Gemmatirosa kalamazoonensis*' KBS708 (92) or '*Gemmatirosa kalamazoonensis*' KBS708 and *Gemmatimonas aurantiaca* T-27^T (114), indicating a close relationship between AP64^T and *Gemmatimonas aurantiaca* T-27^T at the genomic level.

To probe the unique genes that occurred in the genome of strain AP64^T, we performed TBLASTN searches in the genomes of '*Gemmatirosa kalamazoonensis*' KBS708 and *Gemmatimonas aurantiaca* T-27^T for homologues of all predicted proteins in AP64^T using NCBI local BLAST tools. We used a less stringent set of parameters (>30 % sequence identity and >30 % sequence coverage in alignment) to define homologues in order to increase the chance of obtaining phototrophic *Gemmatimonas*-specific genes. A total of 419 ORFs were identified as unique genes in AP64^T (Table 3), among which a large number (358) were hypothetical genes with as-yet unidentified functions. The most distinctive characteristic of AP64^T was the presence of a photosynthesis gene cluster in the genome that contains 28 genes involved in BChl biosynthesis, expression and assembly of the photosystem, carotenoid biosynthesis and regulatory functions. The photosystem in AP64^T may provide an additional energy source, representing an advantage over purely heterotrophic members of the genus

Gemmatimonas when striving to live in harsh environments like the desert lake where AP64^T was isolated. To cope with such environments, AP64^T seems to have developed more strategies. For instance, the presence of genes encoding cyanophycin synthase and trehalase in the genome of strain AP64^T (Table 3) suggests that AP64^T may use cyanophycin and trehalose as energy-storage material. Cyanophycin is rich in nitrogen as well as in carbon, and its storage inside bacterial cells has a possible relation to enduring long starvation. In addition to serving as an energy source, trehalose is also known as a protective agent that enables bacterial cells to cope with cold and desiccation stresses (Potts, 1994). The presence of trehalose perhaps contributes to the capacities of the members of the phylum *Gemmatimonadetes* to resist environmental stresses, since 16S rRNA gene clone sequences associated with this phylum have often been found in arid environments.

Emended description of the genus *Gemmatimonas* Zhang *et al.* 2003

The genus is as described by Zhang *et al.* (2003) with the following emendations and modifications. Cells are pigmented, with oscillol 2,2'-dirhamnoside as the major carotenoid, grow under aerobic or micro-oxic conditions and do not grow under anaerobic conditions. The major type of respiratory quinones are menaquinones differing in the number of isoprenoid side-chain residues. The main fatty acids of cells grown on agar plates are C_{16:1}, C_{14:1} and C_{16:0}. The DNA G+C content of the genomes of the type strains is 64.3–64.4 %.

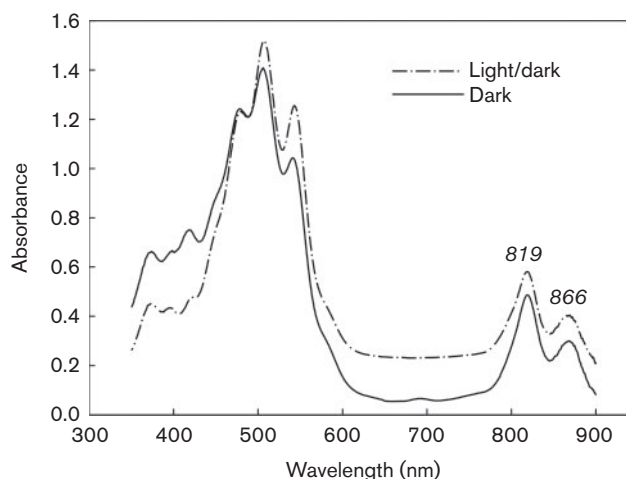


Fig. 3. *In vivo* absorption spectra of cells of strain AP64^T grown under a 12:12 h light/dark cycle and under continuous darkness. For cultures under a light/dark cycle, the strain was cultured for 2 weeks. For cultures under continuous darkness, colonies were transferred to fresh agar plates every 2 months during a 1-year experiment. Cells were harvested from the latest agar plates for analysis.

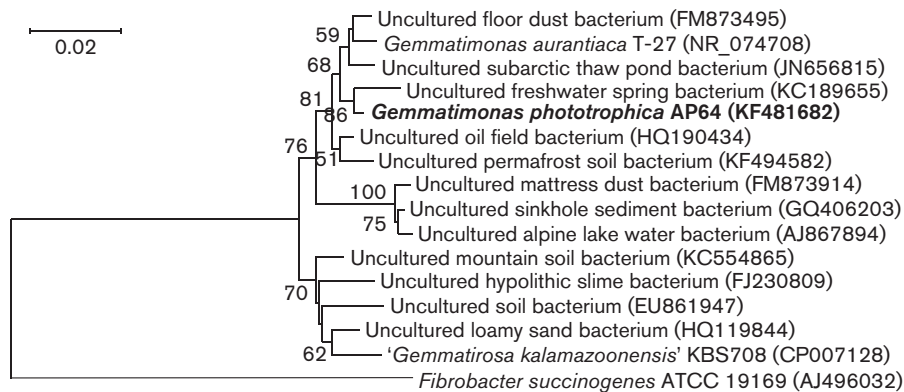


Fig. 4. Phylogenetic tree of 16S rRNA gene sequences of representative members of the *Gemmatimonadetes*. A total of 285 sequences (>1450 bp long) from environmental clones and pure cultures was multiply aligned, and 1130 gap-free conserved positions were used for inference of neighbour-joining trees. Bootstrap resampling (1000 times) was performed. The sequence from *Fibrobacter succinogenes* ATCC 19169^T was used as an outgroup. Only the branches containing strain AP64^T, *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazoonensis*' KBS708 are shown in the figure. Bar, 2% nucleotide substitution.

Emended description of *Gemmatimonas aurantiaca* Zhang *et al.* 2003

The species is as described by Zhang *et al.* (2003) with the following emendations and modifications. Cells grow under either aerobic or microaerobic conditions. Weak growth on agar plates occurs at 16 °C. Tolerates salinity up to 8 g NaCl l⁻¹. Resistant to ampicillin, penicillin, paramycin sulfate and nystatin, but susceptible to neomycin, bacitracin, gentamicin, vancomycin and polymyxin B sulfate.

Description of *Gemmatimonas phototrophica* sp. nov

Gemmatimonas phototrophica [pho.to.tro'phi.ca. Gr. n. *phôs* -otos light; N.L. adj. *trophicus* (from Gr. adj. *trophikos*) nursing, tending or feeding; N.L. fem. adj. *phototrophica* feeding on light, phototrophic].

Cells grow only on agar medium, and the colonies are tiny, round, smooth and red-pigmented. Cells grow best under a reduced oxygen atmosphere (9.8–15.2 % oxygen), but do

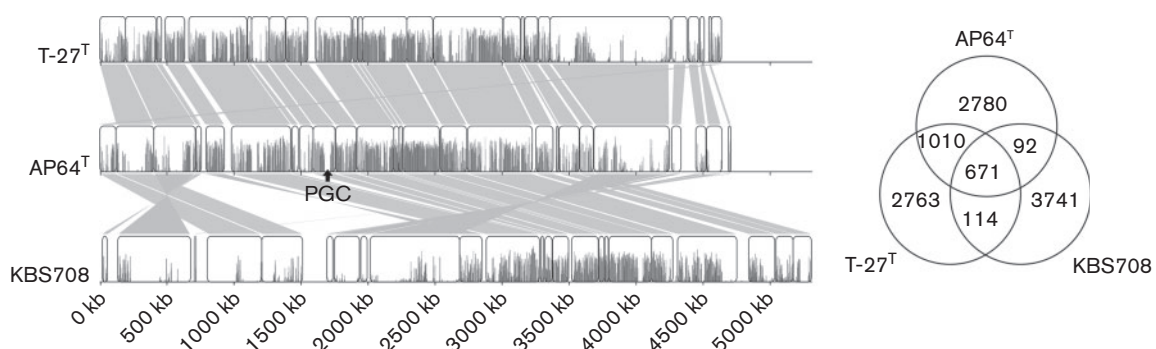


Fig. 5. WGA (left) and statistics of shared and distinct orthologues (right) of strain AP64^T, *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazoonensis*' KBS708. The complete chromosome sequences of *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazoonensis*' KBS708 were retrieved from GenBank as described in the text. The draft genome of strain AP64^T was reordered with reference to the genome of *Gemmatimonas aurantiaca* T-27^T. Each block represents a region that is presumably homologous and internally free of genomic rearrangement. The height of the similarity profile within blocks was calculated to be inversely proportional to the mean alignment column entropy over a region of the alignment, corresponding to the mean level of conservation in that region of the genomic sequence. Colours show the homologous regions shared by the genomes. Regions that are conserved among all three genomes are shown in red. Regions conserved only among subsets of the genomes have been differently colour-coded. The position of the photosynthesis gene cluster (PGC) is marked.

Table 3. Unique ORFs predicted in the genome of strain AP64^T

Genome sequences of *Gemmatimonas aurantiaca* T-27^T and '*Gemmatirosa kalamazoonensis*' KBS708 were used as references for function prediction. Thresholds of >30 % protein sequence identity and >30 % sequence coverage in alignment were used to define a unique ORF. Twenty-eight unique genes in AP64^T that are involved in the photosynthesis gene cluster (*bchP2G-ppsR-aerR-bchFNBHLM-acsF-lhaA-puhABC-pufCMLAB-bchZYXC-crtF-bchID*; Zeng *et al.*, 2014) are not shown. A further 358 ORFs are hypothetical or of unknown function, giving a total of 419 ORFs.

Predicted function	Contig	Gene position
(S)-2,3-di-O-geranylgeranyl glyceryl phosphate synthase	3	107152–108039
23S ribosome-associated endonuclease	1	44067–43063
Antitoxin protein HigA	2	42393–42022
Bacteriophage N4 adsorption protein B	4	213034–211490
Cation antiporter (Na ⁺ /Ca ²⁺)	1	923633–922611
Citrate transporter	1	280760–282079
Cyanophycin synthase	1	794245–797091
Cyanophycin synthase	1	797084–798934
Divinyl protochlorophyllide <i>a</i> 8-vinyl-reductase*	2	338640–339923
Flp pilus assembly protein RcpC	1	603701–604663
Flp pilus assembly protein TadB	1	608632–609594
Gentamicin 3'-N-acetyltransferase	2	272128–272571
Isoaspartyl dipeptidase	1	793345–792173
Lactone-specific esterase	5	308828–309778
Macrophage infectivity potentiator-related protein	1	1183716–1184261
Mg-protoporphyrin IX monomethyl ester cyclase BchE*	1	241690–243291
Phenylalanine 4-hydroxylase	7	136851–135904
Photosynthetic reaction centre cytochrome <i>c</i> subunit*	4	17445–18401
Poly(glycerol-phosphate) alpha-glucosyltransferase	2	689882–690982
Pyridine nucleotide-disulfide oxidoreductase	2	206696–205449
Sialidase	1	360693–361925
Sodium-calcium exchanger	7	146636–145683
Spore peptidoglycan hydrolase (<i>N</i> -acetylglucosaminidase)	4	50238–49183
Sugar transporter	2	783861–785156
Sulfate adenyltransferase subunit 1	1	226411–228303
Sulfate adenyltransferase subunit 2	1	225491–226411
Toxin protein HigB	2	219811–220008
Toxin protein HigB	2	42763–42443
Toxin protein VapC	4	328152–328397
Transcriptional regulator, DeoR family	7	264581–263589
Trehalase	2	785156–786784
Type II/IV secretion system ATPase TadZ/CpaE	1	605925–607136
Type II/IV secretion system ATP hydrolase TadA/CpaF	1	607139–608635

*Photosynthesis-related gene located outside the photosynthesis gene cluster.

not grow under anaerobic conditions. Cells are rod-shaped, 0.3–0.5 µm wide and most commonly 1–6 µm long, occasionally forming filaments up to 12 µm long. The temperature range for growth is 20–30 °C, and weak growth may occur at 16 °C. The pH range for growth is 6.0–9.0, with an optimum at pH 7.5–8.0. NaCl is not required, but concentrations up to 2 g l⁻¹ are tolerated. Cells reproduce by binary fission or budding. Produce pigments as oscillol 2,2'-dirhamnoside, BChl *a* and carotenoids of the spirilloxanthin series. Cells grow well with yeast extract as the sole carbon source. Weak growth with peptone is observed. No growth is observed with the following compounds as the sole carbon source: Casamino acids, sodium succinate, sodium acetate, sodium pyruvate, potato starch, sucrose,

L-glutamic acid, L-leucine, L-arginine, L-alanine, L-isoleucine, L-arabinose, D-sorbitol and D-mannitol. The major respiratory quinone is MK-8. Resistant to ampicillin, penicillin, paramycin sulfate, polymyxin B sulfate and nystatin, but susceptible to neomycin, vancomycin, bacitracin and gentamicin. Cells are motile and positive for oxidase and catalase activities. The dominant fatty acids are C_{16:1}, C_{14:1} and C_{18:1}ω9c. Its genome contains a complete photosynthesis gene cluster. Photosynthetic reaction centres appear to be expressed constitutively.

The type strain, AP64^T (=DSM 29774^T=MCCC 1K00454^T), was isolated at 0.5 m depth of near-shore water of Swan Lake, western Gobi Desert, Inner Mongolia, northern

China. Based on the draft genome sequence, the DNA G + C content of the type strain is 64.4 %.

ACKNOWLEDGEMENTS

We thank Anni Glud for providing the oxygen microelectrodes and for help with measurements. This work was supported financially by the projects GAČR P501/12/G055 and Algain CZ.1.07/2.3.00/30.0059.

REFERENCES

- Auch, A. F., von Jan, M., Klenk, H. P. & Göker, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2, 117–134.
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M. & other authors (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75.
- Darling, A. E., Mau, B. & Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5, e11147.
- DeBruyn, J. M., Nixon, L. T., Fawaz, M. N., Johnson, A. M. & Radosevich, M. (2011). Global biogeography and quantitative seasonal dynamics of Gemmatimonadetes in soil. *Appl Environ Microbiol* 77, 6295–6300.
- DeBruyn, J. M., Fawaz, M. N., Peacock, A. D., Dunlap, J. R., Nixon, L. T., Cooper, K. E. & Radosevich, M. (2013). *Gemmatirosa kalamazoonesis* gen. nov., sp. nov., a member of the rarely-cultivated bacterial phylum Gemmatimonadetes. *J Gen Appl Microbiol* 59, 305–312.
- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P. & Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72, 5069–5072.
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P. & Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57, 81–91.
- Hanada, S. & Sekiguchi, Y. (2014). The phylum Gemmatimonadetes. In *The Prokaryotes*, 4th edn, vol. 11, pp. 677–681. Edited by E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt & F. Thompson. Berlin, Heidelberg: Springer.
- Kainz, M., Lucotte, M. & Parrish, C. C. (2002). Methyl mercury in zooplankton – the role of size, habitat, and food quality. *Can J Fish Aquat Sci* 59, 1606–1615.
- Kamagata, Y. (2010). Phylum XXI. *Gemmatimonadetes* Zhang, Sekiguchi, Hanada, Hugenholtz, Kim, Kamagata and Nakamura 2003, 1161^{VP}. In *Bergey's Manual of Systematic Bacteriology*, 2nd edn, vol. 4, pp. 781–784. Edited by N. R. Krieg, J. T. Staley, D. R. Brown, B. P. Hedlund, B. J. Paster, N. L. Ward, W. Ludwig & W. B. Whitman. New York: Springer.
- Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H. & other authors (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* 62, 716–721.
- Koblížek, M., Mičoušková, J., Kolber, Z. & Kopecký, J. (2010). On the photosynthetic properties of marine bacterium COL2P belonging to *Roseobacter* clade. *Arch Microbiol* 192, 41–49.
- Konstantinidis, K. T. & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci U S A* 102, 2567–2572.
- Lukeš, M., Procházková, L., Schmidt, V., Nedbalová, L. & Kaftan, D. (2014). Temperature dependence of photosynthesis and thylakoid lipid composition in the red snow alga *Chlamydomonas* cf. *nivalis* (Chlorophyceae). *FEMS Microbiol Ecol* 89, 303–315.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P. & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14, 60.
- Potts, M. (1994). Desiccation tolerance of prokaryotes. *Microbiol Rev* 58, 755–805.
- Pruesse, E., Peplies, J. & Glöckner, F. O. (2012). SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28, 1823–1829.
- Qin, Q. L., Xie, B. B., Zhang, X. Y., Chen, X. L., Zhou, B. C., Zhou, J., Oren, A. & Zhang, Y. Z. (2014). A proposed genus boundary for the prokaryotes based on genomic insights. *J Bacteriol* 196, 2210–2215.
- Richter, M. & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 106, 19126–19131.
- Takaichi, S., Maoka, T., Takasaki, K. & Hanada, S. (2010). Carotenoids of *Gemmatimonas aurantiaca* (Gemmatimonadetes): identification of a novel carotenoid, deoxyoscollol 2-rhamnoside, and proposed biosynthetic pathway of oscollol 2,2'-dirhamnoside. *Microbiology* 156, 757–763.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30, 2725–2729.
- Zeng, Y., Feng, F., Medová, H., Dean, J. & Koblížek, M. (2014). Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc Natl Acad Sci U S A* 111, 7795–7800.
- Zhang, H., Sekiguchi, Y., Hanada, S., Hugenholtz, P., Kim, H., Kamagata, Y. & Nakamura, K. (2003). *Gemmatimonas aurantiaca* gen. nov., sp. nov., a Gram-negative, aerobic, polyphosphate-accumulating micro-organism, the first cultured representative of the new bacterial phylum Gemmatimonadetes phyl. nov. *Int J Syst Evol Microbiol* 53, 1155–1163.