Guillaume Tahon, Duygu Gök, Liesbeth Lebbe and Anne Willems. 2021. Description and functional testing of four species of the novel phototrophic genus Chioneia gen. nov., isolated from different East Antarctic environments. Syst. Appl. Microbiol. 44:126250.

Doi 10.1016/j.syapm.2021.126250

Description and functional testing of four species of the novel phototrophic genus *Chioneia* gen. nov., isolated from different East Antarctic environments

Guillaume Tahon[1,2], Duygu Gök[1], Liesbeth Lebbe[1] and Anne Willems[1]*

[1] Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium

[2] Current address: Laboratory of Microbiology, Agrotechnology and Food Sciences, Wageningen University, Wageningen, The Netherlands

Email addresses:

Guillaume Tahon: guillaume.tahon@wur.nl and Guillaume.Tahon@UGent.be

Duygu Gök: Duygu.Gok@UGent.be

Liesbeth Lebbe : Liesbeth.Lebbe@UGent.be

***Correspondence:**

Anne Willems

Laboratory of Microbiology, Department of Biochemistry and Microbiology

K.L. Ledeganckstraat 35

9000 Ghent, Belgium

Anne.Willems@UGent.be

+32 (0) 9 264 51 03

Abstract

Seven Gram-negative, aerobic, non-sporulating, motile strains were isolated from terrestrial (R-67880^T, R-67883, R-36501 and R-36677^T) and aquatic (R-39604, R-39161^T and R-39594^T) East Antarctic environments (i.e. soil and aquatic microbial mats), between 2007 and 2014. Analysis of near-complete 16S rRNA gene sequences revealed that the strains potentially form a novel genus in the family *Sphingomonadaceae* (Alphaproteobacteria). DNA-DNA reassociation and average nucleotide identity values indicated distinction from close neighbors in the family *Sphingomonadaceae* and showed that the seven isolates form four different species. The main central pathways present in the strains are the glycolysis, tricarboxylic acid cycle and pentose phosphate pathway. The strains can use only a limited number of carbon sources and mainly depend on ammonia and sulfate as a nitrogen and sulfur source, respectively. The novel strains showed the potential of aerobic anoxygenic phototrophy, based on the presence of bacteriochlorophyll *a* pigments, which was corroborated by the presence of genes for all building blocks for a type 2 photosynthetic reaction center in the annotated genomes. Based on the results of phenotypic, genomic, phylogenetic and chemotaxonomic analyses, the strains could be assigned four new species in the novel genus *Chioneia* gen. nov. in the family *Sphingomonadaceae*, for which the names C. *frigida* sp. nov. (R-67880^T, R-67883 and R-36501), *C. hiemis* sp. nov. (R-36677^T), *C. brumae* sp. nov. (R-39161^T and R-39604) and *C. algoris* sp. nov. (R-39594^T) are proposed.

Keywords

Aerobic anoxygenic phototrophy, Antarctica, ice-free soil, microbial mat, *pufLM*, *Sphingomonadaceae*

Introduction

Antarctica, Earth's southernmost continent, can be denominated a continent of extremes. Unlike any other place on our planet, Antarctica is constantly dominated by various extreme environmental conditions such as low temperatures, low humidity, limited precipitation and frequent freeze-thaw cycles [18, 73, 110]. Their constant influence, in combination with the absence of both vascular plants and large numbers of primary producers in West and East Antarctica, have resulted in an environment lacking nutrients [18, 49, 108]. To overcome these nutrient limitations, the microscopic life dominating these ecosystems may be expected to turn towards alternative energy sources. For phototrophic members of the microscopic communities inhabiting continental Antarctica, sunlight, which is abundantly present during the austral summer, may be an important alternative energy resource.

Two major types of phototrophy can be distinguished: rhodopsin-based and (bacterio)chlorophyll-based phototrophy. The first relies on a simple mechanism involving ion-pumping rhodopsin proteins [9, 107]. The latter, requiring a more complex mechanism , relies on (bacterio)chlorophyll-dependent photochemical reaction centers. Among bacteria, chlorophyll-dependent species have been found only in the Cyanobacteria, whereas bacteriochlorophyll-dependent species have been found in the Acidobacteria, Chlorobi, Chloroflexi, Firmicutes, Gemmatimonadetes and Proteobacteria [14, 52, 115, 116]. Bacteriochlorophyll-dependent species in the aforementioned phyla are referred to as anoxygenic phototrophic bacteria, most of which are known to be aerobic anoxygenic phototrophs [52, 116].

As part of several studies investigating the bacterial diversity in a number of Antarctic locations, a large number of strains were isolated between 2007 and 2014 [85, 86, 106]. Comparison of data obtained from these studies indicated that seven of them (i.e. strains R-67880^{T}, R-67883, R-36501, R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T}), isolated from soil or (aquatic) microbial mats in East Antarctica, had highly similar 16S rRNA gene sequences and potentially formed a novel genus in the family *Sphingomonadaceae*.

Taxa of the family *Sphingomonadaceae* are in general chemoorganotrophs, however, the presence of phototrophy within several genera of this family has been observed. For example, bacteriochlorophyll *a* pigments, the most common and abundant type of bacteriochlorophyll in aerobic anoxygenic phototrophs, have been detected in members of the multiple genera,

including *Blastomonas* [41], *Sandaracinobacter* [117], *Sandarakinorhabdus* [37] and *Polymorphobacter* [43].

Members of the *Sphingomonadaceae* are widespread on Earth and have been found, using culture-dependent and -independent approaches, in a variety of environments including marine [20, 44, 62, 63, 76], freshwater [16, 21, 37, 86] and wastewater habitats [82, 119], plants [28, 83, 96], clinical specimens [39], the gut of the house sparrow *Passer domesticus* [78], soil [2, 17, 112], and various extreme environments such as permafrost soil [43], Arctic and Antarctic habitats [15, 19, 36, 50, 65, 81, 85, 86, 94, 104, 106, 108] and hydrothermal vents [47].

In this study, the seven East Antarctic strains were further characterized using both laboratory experiments and genome-based analyses. Based on the results, we propose that they represent four new species in the novel photoheterotrophic genus *Chioneia* gen. nov. in the family *Sphingomonadaceae*, i.e. *C. frigida* sp. nov. (R-67880^{T}, R-67883 and R-36501), *C. hiemis* sp. nov. (R-36677^{T}), *C. brumae* sp. nov. (R-39161^{T} and R-39604) and *C. algoris* sp. nov. (R-39594^{T}).

Material and Methods

Isolation of bacterial strains

All strains in this study were obtained from samples originating from different East Antarctic sites (Table 1). Strains R-67880^T and R-67883 originate from KP43, a top surface soil sample taken on Utsteinen ridge, in the neighborhood of the Belgian Princess Elisabeth Station (PES) on 3 February 2009 [106]. Strains R-36501 and R-36677^T also originate from terrestrial samples (i.e. BB50 and BB115) taken in proximity of the PES, sampled in January 2007 [85]. Strains R-39604, R-39161^T and R-39594^T originate from SK5, an aquatic sample collected at Naka Tempyo in the Syowa Oasis region in January 2007 [86].

Strains R-67880^T and R-67883 were obtained from an isolation campaign focusing on the phototrophic bacterial diversity in ice-free top surface soil from Utsteinen, East Antarctica [106]. After a 21 week long non-shaking liquid enrichment in low-nutrient photoautotroph medium at 15 °C in an aerobic atmosphere, plates containing the same medium were inoculated with the enrichment and incubated under the same conditions. Isolates were picked up 7 months after plating [106].

Strain R-36501 was isolated from peptone-yeast-glucose-vitamin medium at 20 °C under aerobic conditions after 21 days. Strain R-36677^T was isolated from Reasoner's 2A (R2A) agar at 20 °C under aerobic conditions after 5 days**.** The two latter strains were isolated during a campaign which focused on the heterotrophic bacterial diversity in microbial mats in continental Antarctica at Utsteinen [85].

Finally, strains R-39604, R-39161^T and R-39594^T were all picked up from 1/10 R2A under an aerobic atmosphere, at 20 °C, 20 °C and 15 °C, after 9, 28 and 14 days, respectively. These three strains were isolated during a sampling campaign focusing on heterotrophic bacterial diversity in aquatic microbial mats from different Antarctic locations [86]. More details on environmental sampling and isolation of the abovementioned strains can be found in Tahon et al. [103] and Tahon and Willems [106] (i.e. strains R-67880^T and R-67883), Peeters, Ertz and Willems [85] (i.e. strains R-36501 and R-36677^T) and Peeters et al. [86] (i.e. strains R-39604, R-39161^T and R-39594^T). After isolation and purification, all strains were stored at -80 °C.

16S rRNA gene analysis

Partial 16S rRNA gene analysis (i.e. V1-V3 region) of strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T was previously performed in separate studies [85, 86, 106]. To improve their identification, the seven strains were revived and their DNA was

extracted by alkaline lysis [80]. Subsequently, the near-complete 16S rRNA gene was amplified using primers pA and pH, and sequenced using primers BKL1, *Gamma, Gamma, *O, *3, 3 and *R, as described by Tahon and Willems [106]. Afterwards, the amplified near-complete 16S rRNA sequences of strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T were compared to the EzBioCloud database to identify the closest cultivated and validly described neighbors based on the 16S rRNA gene [113].

MALDI-TOF MS profiling

The seven strains, including two closely related type strains (i.e. *Polymorphobacter multimanifer* DSM 102189^T and *Polymorphobacter fuscus* DSM 105347^T, as based on EzBioCloud results), were grown on R2A at their respective optimal growth temperatures and were subcultured weekly. Third generation pure cultures were used to obtain matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) fingerprints. Preparation of cell extracts and acquisition of MALDI-TOF MS fingerprints was performed as described by Dumolin et al. [31]. Briefly, a 1 µL loop of bacterial cells (Ansa/Microloop 1 µL) was suspended in 300 µL sterile MilliQ water and vortexed. Then, 900 µL absolute ethanol was added and samples were inverted to mix. Samples were centrifuged for 3 minutes (4 °C, 14000 rpm) and supernatants were discarded. After a short spin, residual ethanol was removed. Cell pellets were air-dried for 3 minutes at room temperature. The pellets were then suspended in 40 µL 70% formic acid in water and homogenized by pipetting up and down. Next, 40 µL acetonitrile was added and samples were vortexed and centrifuged for 2 minutes (4 °C, 14000 rpm). The supernatants were transferred to new tubes and were used for spotting for MALDI-TOF MS analysis. Raw mass spectra were extracted as t2d files. By using the Data Explorer 4.0 software (Applied Biosystems), these were converted to text files. Subsequently, these were imported in BioNumerics 7.5 (Applied Maths) and transformed into fingerprints. Pearson's product moment correlation was used to find similarities between fingerprints. Finally, spectra were clustered using unweighted pair group clustering with arithmetic means.

Genome sequencing and analyses

To allow genome-based analyses and comparison with publicly available genomes of closely related type strains of the family *Sphingomonadaceae*, the genomes of strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T were sequenced. DNA was extracted using an automated Maxwell® DNA preparation instrument (Promega), as previously described by Tahon et al. [102]. The DNA was sequenced at the Oxford Genomics Centre at

the Wellcome Centre for Human Genetics using the Illumina HiSeq 2500 platform (2x150 bp cycles). The genomes were assembled using Shovill 1.1.0 (https://github.com/tseemann/shovill) with the adaptor trimming command (--trim) enabled. Contigs smaller than 500 bp were discarded. The Quality Assessment Tool for Genome Analysis was used to obtain summary statistics of the different assemblies (e.g. number of contigs, N50, L50 and the G+C content) [40]. Genome completeness was determined using CheckM [84]. Final contigs of R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T were submitted for annotation using the Integrated Microbial Genomes-Expert Review (IMG-ER) platform [71] and the RAST server [5]. Genome assemblies and annotations of four closely related type strains based on 16S rRNA gene similarity were publicly available (Table 2). For *Sphingomonas spermidinifaciens* 9NM-10^T, a genome without annotation was available (accession number NWMW00000000). Therefore this was performed in this study using the RAST server [5].

AntiSMASH was used to identify secondary metabolite biosynthetic gene clusters in the genomes [12]. CRISPRCasFinder was used to find clustered regularly interspaced short palindromic repeats (CRISPRs) in the genomes of the strains [25]. PHASTER was used to identify the different prophage sequences in the genomes [4]. Glycosyl hydrolases (GH), glycosyl transferases (GT), polysaccharide lyases (PL), carboxyl esterases (CE), carbohydrate-binding modules (CBM) and auxiliary activities (AA) were identified through the dbCAN2 meta server [118], using the CAZy database [64]. Peptidases and sulfatases were identified using the MEROPS [91] and SulfAtlas [7] databases, respectively. Pairwise average nucleotide identity (ANI) was analyzed using the OrthoANIu tool of the EzBioCloud platform [114]. *In silico* DNA-DNA hybridization (*is*DDH) was performed using the DSMZ Genome Distance Calculator 2.1 [74].

Phylogenetic analyses

To confirm 16S rRNA gene data obtained by Sanger sequencing, full-length 16S rRNA genes of R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T were extracted from the genome assemblies using RNAmmer [58] and compared with the amplified sequences using MEGA X [55].

To support the results obtained using MALDI-TOF MS and EzBioCloud analyses, the 16S rRNA genes of the seven strains, together with those of closely related type strains of the family *Sphingomonadaceae* were included in a phylogenetic analysis. The 16S rRNA gene sequences of *Paracraurococcus ruber* NS89^T and *Craurococcus roseus* DSM 15488^T were used as an

outgroup. Using MEGA X, the sequences were aligned and a phylogenetic maximum likelihood tree (1000 bootstraps) was constructed [55]. Tree visualization was done using the iTOL software [60].

To detect presence of similar organisms in other environments, the 16S rRNA gene sequences of the novel strains were used in a BLAST search [45]. Highly similar sequences (>95%) were added to the above dataset, after which a phylogenetic maximum likelihood tree was constructed and visualized as described above. Additionally, the 16S rRNA gene sequences of the novel strains were subjected to an IMNGS search [59] using a 95% similarity threshold and a minimum size of 200. Datasets in which an abundance of more than 0.01% was detected were retained from the IMNGS results. The partial 16S rRNA gene sequences from these datasets and similar to the sequences of the novel strains were added to the alignment built for the 16S rRNA phylogenetic analysis using MAFFT (--add command) [48]. To confirm sequences and thus datasets in which partial 16S rRNA gene sequences similar to the novel strains were found, the resulting alignment was then used to create a maximum likelihood phylogenetic tree using IQ-TREE 2 [77].

The annotated genomes were also used to detect presence of the novel organisms in other environments. Using the Single Genomes vs. Metagenomes option in the IMG/MER database [70], each of the seven genomes was compared to the 41.056 available metagenomes. This analysis shows how genes in a genome are matched by BLAST to genes in different metagenomes or metagenomic categories.

To construct a phylogenetic *pufLM* tree, the type 2 photosynthetic reaction center L (light, *pufL*) and M (medium, *pufM*) subunit gene sequences of the seven new strains were extracted from the annotated genomes using the IMG-ER platform [71]. The same was performed for all genomes of the strains in the phylogenetic 16S rRNA tree, if the genomes showed the presence of these *pufLM* genes and if the genomes were publicly available. The *pufLM* sequences were aligned in MEGA X and a maximum likelihood phylogenetic tree (1000 bootstraps) was constructed [55]. Tree visualization was done using the iTOL software [60].

For a whole genome-based phylogenetic analysis, the genomes of all the strains that were included in the 16S rRNA phylogenetic analysis were taken into account, except if the genome of the strain was not publicly available. Using bcgTree, all genomes were screened for the presence of 107 single-copy core genes present in the majority of all currently known bacterial species, with default parameters, except that 1000 bootstraps were used [3]. The final tree was visualized using the iTOL software [60].

Fatty acid analysis

To analyze the cellular fatty acids, all strains, including closely related type strains, were grown on R2A at their respective optimal growth temperatures for a week. Harvesting, saponification, methylation and extraction and washing of the fatty acids were performed according to the Sherlock Microbial Identification System (MIDI) protocol, version 3.0 with a few modifications. Steps that were carried out differently were that cells were harvested from all quadrants, instead of only the third quadrant. In the washing step, a few drops of saturated NaCl were added before transferring the organic phase to a gas chromatography vial. The fatty acid extracts were analyzed using gas-liquid chromatography by the MIDI system (Microbial ID Inc.). After separation, these were identified by comparison to the MIDI Peak Library version 5.0. Results were corrected for the issue with the $C_{15:0}$ results in the MIDI database. Although there is a response for this fatty acid, it is not taken into account for the calculation of the percentages. Therefore, the MIDI results were recalculated by including the response for $C_{15:0}$.

Pigment extraction and spectral analysis

The seven new strains, including closely related type strains, were grown on R2A in light (i.e. on the lab bench near a window), at room temperature for a week. Pigment extraction in methanol was carried out according to Tahon et al. [101], except that the cells were washed in 1 mL MilliQ both times instead of in 2 mL MilliQ. For strain R-36677^T, the second washing step was not performed to improve precipitation of the cells. Additionally, for all strains, extracts were vacuum centrifuged at 30 °C for 10 minutes and then at 45 °C for 15 minutes to concentrate the pigment extracts. The extracts were characterized by spectrophotometry using the Spark multimode reader (Tecan). Absorbances were measured between 200 and 1000 nm. Presence of flexirubin-type pigments was determined by alkanalizing the samples with 0.1 M NaOH after which the spectra were examined for a bathychromatic shift characteristic of flexirubin-type pigments [98]. Additionally, pigment extracts were flooded with 20% KOH to examine presence of a color shift indicating presence of a flexirubin-type pigment [33].

Morphology and metabolic profile

For the observation of both colony and cell morphology, strains were grown on R2A for a week at their respective optimal temperatures. Colony morphology was observed using a Motic stereo microscope. Cell morphology was analyzed using phase-contrast microscopy (Olympus BX41). Gliding motility was determined by performing the hanging drop method [10]. Catalase activity was evaluated based upon gas production of bacterial cells in 10% (v/v) H_2O_2. Oxidase activity

of the strains was determined using 1% N, N, N', N'-tetramethyl-p-phenylenediamine. Gram staining was performed as described by MacFaddin [66].

Growth at different temperatures (i.e. 0, 4, 10, 15, 20, 25, 28, 30, 35, 37, 42, 45 and 52 °C) was evaluated on R2A in triplicate on a weekly basis for four weeks to identify the temperature growth range of the strains as well as the optimal growth temperature.

To determine whether the strains grew under microaerobic conditions, growth was performed in triplicate in microaerobic conditions (80% N_2, 15% CO_2 and 5% O_2) for two weeks on R2A at 20 °C for strains R-67880^T, R-36501, R-39604, R-39161^T and R-39594^T, *Sandarakinorhabdus limnophila* DSM 17366^T, and at 25 °C for strains R-67883, R-36677^T, *Polymorphobacter multimanifer* DSM 102189^T, *Sphingomonas spermidinifaciens* DSM 27571^T and *Polymorphobacter fuscus* DSM 105347^T. To assess growth in an anaerobic atmosphere (80% N_2, 10% CO_2 and 10% H_2, 10 ppm O_2), strains were grown in duplicate on R2A at 23 °C for two weeks.

To assess growth on different media, strains were grown on Tryptone Soya Agar (TSA), Nutrient Agar (NA), Marine Agar (MA), 10 and 100 times diluted R2A supplemented with Agar no. 2 (Neogen Culture Media) in triplicate at their optimal growth temperatures. Growth was evaluated on a weekly basis for four weeks. Growth in R2A Broth (R2B) was assessed in triplicate at 20 °C and evaluated by measuring optical densities at 600 nm daily for two weeks, using the Spark multimode reader (Tecan). Values >0.05 were considered as positive after correction of the values (i.e. subtraction of the blank).

Growth at different pH values (from pH 5 to 11, in steps of 0.5 pH units) was tested by growing the strains at 20 °C in triplicate in R2B, buffered with MES (pH 4.0-6.0), ACES (pH 6.1-7.5), TAPS (pH 7.7-9.1) or CAPS (pH 9.7-11.1) to a final concentration of 100 mM. Salt tolerance was evaluated by growing the strains in triplicate at 20 °C in R2B without salt (i.e. 0%) and R2B supplemented with NaCl to concentrations of 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.5, 3, 4 and 5%. To assess both pH and salt tolerance ranges and optima, optical densities were measured at 600 nm on a daily basis for two weeks using the Spark multimode reader (Tecan). Values >0.05 were considered as positive after correction of the values (i.e. subtraction of the blank).

Susceptibility of the strains to several antibiotics was tested by growing the strains at their optimal growth temperatures on R2A for two weeks and using antimicrobial susceptibility test discs (Oxoid). The tested antibiotics included ampicillin (10 µg), bacitracin (4 µg), chloramphenicol (30 µg), gentamicin (10 µg), tetracycline (30 µg) and vancomycin (30 µg).

Hydrolysis of Tweens 20, 40, 60 and 80 was tested by growing the strains for three weeks on R2A supplemented with 1% (v/v) Tween 20, 40, 60 and 80, respectively. Activity was present if a cloudy halo was formed around the colonies. Hydrolysis of starch was tested by growing the strains two weeks on R2A supplemented with 0.5% (w/v) starch and by subsequently flooding the cultures with 10% (v/v) Lugol's iodine solution. Activity was present if a transparent halo was formed around the colonies. Hydrolysis of casein was tested by evaluating growth of the strains for four weeks on R2A supplemented with 3% (w/v) casein. Hydrolysis of carboxymethylcellulose (CMC) was tested by growing the strains for four weeks on R2A supplemented with 0.5% (w/v) CMC and by flooding the cultures after growth with 0.2% aqueous Congo Red and a subsequent 1% (w/v) NaCl washing step. Activity was present if a transparent halo was formed around the colonies. DNA hydrolysis was tested by growing the strains for four weeks on DNase test agar and by flooding the cultures after growth with 1M HCl. Activity was present if a transparent halo was formed around the colonies. All hydrolysis tests were performed at the optimal growth temperatures of the strains.

API ZYM, API 50 CH, API 20 NE (bioMérieux) and GEN III MicroPlate (Biolog) tests were carried out to evaluate enzyme activity, the utilization of different carbon sources, the formation of acids and chemical sensitivity. API ZYM tests were performed as stated in the manufacturer's protocol. API 50 CH and API 20 NE tests were incubated at the respective optimal growth temperatures of the strains for a four-week period and were evaluated on a weekly basis. GEN III MicroPlate tests were incubated at the optimal growth temperatures of the strains for five to ten days, depending on the growth rate of the strains. Optical densities were measured at 590 nm using the Spark multimode reader (Tecan). Values >0.075 were considered as positive after correction of the values for the positive control (i.e. subtraction of the blank).

Availability of sequence data

The 16S rRNA sequences were deposited in the National Center for Biotechnology Information/GenBank database (NCBI) under the accession numbers MN601350 to MN601356 for strains R-36501, R-36677^T, R-39161^T, R-39594^T, R-39604, R-67880^T and R-67883, respectively. The annotated genome sequences of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T are publicly available on the IMG-ER platform under accession numbers Ga0401289, Ga0401290, Ga0401291, Ga0401292, Ga0401293, Ga0427977 and Ga0427589, respectively (Table 2). The raw sequence reads of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T are publicly

available in the NCBI database under accession numbers SRR12560621, SRR12560612, SRR12561322, SRR12564341, SRR12563995, SRR12563814 and SRR12564433, respectively.

Availability of biological material

Strains R-67880^T, R-36677^T, R-39161^T and R-39594^T have been deposited in the Belgian Co-ordinated Collections of Microorganisms (BCCM/LMG, Ghent, Belgium) under the accession numbers LMG 31952^T, LMG 31953^T, LMG 31951^T, LMG 31950^T, respectively. Additionally, strains R-67880^T, R-36677^T, R-39161^T and R-39594^T have been deposited in the Spanish Type Culture Collection under the accession numbers CECT 30379^T, CECT 30380^T, CECT 30378^T and CECT 30377^T, respectively. Strains R-67883, R-36501 and R-39604 have been deposited in the research collection of the Laboratory of Microbiology at Ghent University and are available for further research.

Results and Discussion

MALDI-TOF MS analysis

The clustering of the MALDI-TOF profiles shows two clear main clusters, one comprising the strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and one comprising the closely related strains *Polymorphobacter multimanifer* DSM 102189^T and *Polymorphobacter fuscus* DSM 105347^T (Figure S1). Within the cluster comprising the seven novel strains, three separate subclusters can be observed. The first, comprising strains R-67880^T, R-67883 and R-36501 shows a high similarity between the MALDI-TOF MS profiles of the respective strains, which could indicate that they belong to a single species. This is also the case for strains R-39604 and R-39161^T. The third cluster, comprising strains R-36677^T and R-39594^T, shows a lower similarity between the profiles in comparison to the former clusters. These strains could therefore represent two separate species.

Phylogenetic placement and phylogenomics

A recent analysis of bacteria isolated from Antarctic locations during various isolation campaigns showed that the partial 16S rRNA gene sequences of strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T were very similar (i.e. >99%). Furthermore, comparison of the partial sequences with the EzBioCloud database [113] revealed 16S rRNA gene similarities of ~95% with type strains present in the database, indicating the new strains potentially belong to a novel genus. To improve their placement, the near-complete 16S rRNA gene sequences of the seven strains was amplified and compared to the EzBioCloud database. Results indicated that the closest relatives of the novel strains were *Polymorphobacter multimanifer* 262-7^T, *Polymorphobacter arshaanensis* DJ1R-1^T, *Sphingomonas spermidinifaciens* 9NM-10^T, *Polymorphobacter fuscus* D40P^T and *Sandarakinorhabdus limnophila* DSM 17366^T, with similarities between 94.3 and 95.3%. *Polymorphobacter multimanifer* 262-7^T was the closest type strain of strains R-67880^T, R-67883 and R-36501 (sequence similarity of 94.88%), while *Polymorphobacter arshaanensis* DJ1R-1^T was the closest relative of strains R-39604 and R-39161^T (sequence similarity of 95.09%), strain R-36677^T (sequence similarity of 95.24%) and strain R-39594^T (sequence similarity of 95.02%). However, the description of *Polymorphobacter arshaanensis* DJ1R-1^T was only published when this study was already ongoing and could therefore not be included in the laboratory analyses anymore. As all these percentages are close to the cutoff to define a new genus [111], strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T may form

a new genus in the family *Sphingomonadaceae* (Alphaproteobacteria). Sequence similarities between the new strains are all >99%, as calculated by pairwise distances (number of differences) in MEGA X [55]. This would indicate that the seven new strains all belong to the same species, as the threshold of 98.65-98.7% is exceeded [51, 111].

To complement the results obtained using EzBioCloud, a 16S rRNA gene maximum-likelihood tree was generated (Figure 1). This tree shows that the novel strains form a clear and coherent separate cluster, which could indicate that they belong to a new genus. Similar to the MALDI-TOF MS dendrogram, subclusters can be observed within the main cluster of the novel strains, suggesting that strains R-67880^T, R-67883 and R-36501 group together, strains R-39604 and R-39161^T also group together, and strains R-36677^T and R-39594^T each may form a separate lineage. However, all appear highly related. As it is already well-known that analyses based on only 16S rRNA genes have a low resolving power to discriminate closely related species [1, 46], a phylogenomic tree based on whole-genome sequences (Figure 2) was also created using 107 single-copy bacterial core genes by bcgTree [3]. This led to a similar tree with more resolved clustering. However, interestingly, *Polymorphobacter arshaanensis* DJ1R-1^T did not cluster with other species of the genus *Polymorphobacter*. Its separate position suggests this species, classified solely using 16S rRNA gene information [88], might not be a member of the genus *Polymorphobacter* and should be reclassified in a novel genus.

Additionally, *is*DDH and ANI values could be calculated using the genome data (Table 3). The thresholds that are generally used for species delineation are 70% and 95%, for *is*DDH and ANI, respectively [51, 95, 109]. As the *is*DDH and ANI values of strains R-67880^T, R-67883 and R-36501 exceed the cutoffs, showing values between 99.70-99.90% and 99.97-99.98%, respectively (Table 3), these strains belong to the same species. Likewise, strains R-39604 and R-39161^T, which show *is*DDH and ANI values of 100.00 and 99.98%, respectively, belong to a single species. Other pairwise *is*DDH values between strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T range from 22.00 to 33.50%, while the ANI values vary between 78.85 and 87.53%, which are all below the recommended species thresholds (Table 3), indicating that strains R-36677^T and R-39594^T indeed each represent a separate species. However, based on 16S rRNA gene similarity between the new strains, it appeared that all strains belong to the same species, as described earlier. This confirms that analyses based on only 16S rRNA are not highly informative at species level [1, 46].

The *is*DDH and ANI values between strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and the most closely related reference type strains *Polymorphobacter multimanifer* DSM 102189^T, *Sphingomonas spermidinifaciens* 9NM-10^T, *Sandarakinorhabdus*

limnophila DSM 17366^T and *Polymorphobacter arshaanensis* DJ1R-1^T range from 18.20 to 19.70% and from 71.40 to 74.52%, respectively, which are also all below the threshold values (Table 3). Based on these analyses, strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T form four new species in a novel genus in the family *Sphingomonadaceae*, for which we propose the name *Chioneia* gen. nov. Unfortunately, for two of these four species there is currently only a single strain available despite our screening of a large collection of isolates from Antarctic samples. We therefore opt to propose new species for the separate strains R-39161^T and R-39594^T so that these organisms are named, documented and available for comparison in other studies.

General genome characteristics of *Chioneia* gen. nov. species

Several main genome characteristics of all strains investigated here can be found in Table 2. The genome sequence data all comply with the proposed minimal standards for taxonomic purposes [23]. Based on the bcgTree analysis [3], all genomes have at least 105 of the 107 major single-copy core genes. This, together with the CheckM results (data not shown) [84], indicates that the genomes are near-complete. Some differences in genome sizes could be observed. The genome sizes range from 3.04 to 4.36 Mb for the new strains (Table 2). Strains R-67883 (4.36 Mb), R-67880^T (4.36 Mb) and R-36501 (4.28 Mb), which form a new species based on the above analyses, show the largest genome sizes. These strains were isolated from inland terrestrial Antarctic samples, which are known to be low in nutrients and moisture [18, 49, 108]. As a result, the bacterial community members in these habitats may need adaptations to survive and a large genome size, which is likely linked to an increase in metabolic capacities, could thus allow these bacteria to produce a larger range of (new) enzymes [90]. This could be advantageous for the exploitation of the scarce nutrients [90]. Strains R-36677^T, R-39604, R-39161^T and R-39594^T, on the other hand, show smaller genome sizes (3.76, 3.05, 3.04 and 3.07 Mb, respectively), which might be due to their presence in the specific microniche of microbial mats from which they were isolated (Table1). It has been previously reported that small genomes are a consequence of genome streamlining in microniches, as a response to low nutrient availability in the environment, and that these indicate the absence of the required plasticity to survive in other environments [38, 87]. G+C contents of the new strains vary between 67.41 and 68.89% (Table 2).

All genomes contain the complete set of tRNAs, representing the full set of amino acids, and only one copy of the 16S rRNA gene (Table 2). However, multiple 16S rRNA copies are known to co-assembly into one 16S rRNA gene, due to the conserved nature of this gene [13].

Therefore, the presence of a single copy of the 16S rRNA gene was confirmed by both the absence of degenerations in the 16S rRNA sequence obtained by Sanger sequencing and by the ratio of the coverage of the contig including the 16S rRNA gene to the coverage of the other contigs for each strain [13].

Using CRISPRCasFinder [25], one CRISPR region could be detected in the genomes of strains R-36677^T, R-39604 and R-39161^T. Two CRISPR regions could be identified in the genome of strain R-39594^T, while CRISPR regions were absent in strains R-67883, R-67880^T and R-36501 (Table 2). These regions encode functions for the recognition and destruction of alien DNA and could be indicative of frequent contact with phages [72]. The genomes of strains R-36677^T, R-39604, R-39161^T and R-39594^T, which all contained CRISPR regions, also contained one or more prophage regions (Table 2). However, strains R-67883, R-67880^T and R-36501 also show the presence of multiple prophage regions, while no CRISPRs were observed. It should be noted, however, that all CRISPR regions observed in the genomes of the new strains were labeled as questionable, while the prophage regions were marked as incomplete.

Chioneia gen. nov. core metabolism

General carbon metabolism

The genome annotations of the novel strains allowed a first insight in the core metabolism of the four novel species as well as comparison with those of closely related type strains of other genera. Based on genome annotations of strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T, the core metabolism of all strains includes the glycolysis, the tricarboxylic acid cycle and the pentose phosphate pathway. The glycolysis can be started by using either α- or β-D-glucose, which is corroborated by the results of the API 20 NE test (Table S3). However, in the API 50 CH test, most strains showed a negative result for the use of different carbon sources, including D-glucose (Table S2). In general, the carbon sources that could be used by these strains were limited, based on the results of API 50 CH and API 20 NE tests (Tables S3 and S4). However, this is probably a consequence of the strains not properly growing in the provided standard medium for these tests. Therefore, these results should be interpreted with caution. Similarly, the Biolog GEN III assays showed no or poor reaction in all wells or even resulted in spontaneous reactions which could not be attributed to the strains. It should be noted that these tests are designed for fast-growing bacteria and normally should be evaluated after 24 to 48 hours of incubation, while incubations longer than 96 hours are not recommended. This complicates the growth and evaluation of slower growing bacteria in these

tests. Even on repeating the experiment, no improvements in the results could be obtained. Therefore, the Biolog GEN III results were not used in this study.

According to the genomes, pyruvate, the end product from the glycolysis, can further be transformed into acetyl-CoA, which can then enter the tricarboxylic acid cycle. The pentose phosphate pathway, which is present in all annotated genomes as well, can provide building blocks and NADPH for other biosynthetic processes. Key enzymes for carbon fixation are absent in genomes of the seven novel strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T. However, the annotated genome of one of the reference strains, *Polymorphobacter fuscus* DSM 105347^T, does show the presence of ribulose-1,5-bisphosphate carboxylase/oxygenase, the key enzyme in the Calvin-Benson-Bassham cycle for carbon fixation, thus hinting the autotrophic potential of this organism [6].

Respiration

The presence of an electron transport chain could be detected in the annotated genomes of the novel strains, as oxidative phosphorylation for the generation of ATP is present in all these genomes. The elements of this chain include NADH dehydrogenase, succinate dehydrogenase, cytochrome c oxidase and an F-type ATPase. The presence of the RegAB two-component system in the genome annotations on the other hand, could be indicative of the potential to increase the electron transport capacity. This system comprises RegA, a response regulator, and RegB, a histidine sensor kinase, which can be triggered by redox signals, resulting in the positive regulation of PetABC (cytochrome bc_1 complex), CydAB (quinol oxidase) and CycA (cytochrome c_2) [30]. In microoxic conditions, the same regulon is involved in triggering the CcoNOQP (cbb_3 cytochrome oxidase) operon, which allows respiration under low oxygen conditions due to the very high affinity of the cbb_3-type oxidase for oxygen [24]. This is corroborated by the ability of the strains to grow and survive under microaerobic conditions (Table 4).

Light-harvesting potential of *Chioneia* gen. nov.

Another interesting feature that could be detected in the genomes of the seven novel strains were *pufLM* genes. Additionally, these genes were found present in the genomes of several closely related strains. The *pufLM* genes are indicative for the presence of a type 2 photosynthetic reaction center in phototrophic bacteria and can thus be used as a marker to study the diversity of aerobic anoxygenic phototrophs [42, 52]. Therefore, a phylogenetic *pufLM* tree was constructed (Figure 3). This tree showed a similar clustering as the 16S rRNA

maximum-likelihood tree (Figure 1). Further evidence of phototrophy could be found by the presence of the bacteriochlorophyll biosynthesis pathway in the genome annotations. The seven novel genomes all have the necessary enzymes to produce both bacteriochlorophyll *a* and *b*. However, only BChl *a* was observed in the spectra of the pigment extracts (peak at ~780 nm) (Figure S2). Phototrophy has previously been detected in Antarctica in bacteria from soil [103, 105, 106], microbial mats [67], lakes [56, 57] and seawater and sea ice [53]. The presence of this phototrophic metabolism may give Antarctic phototrophic bacteria a competitive advantage over non-phototrophic bacteria in the extreme Antarctic conditions. A previous study has already demonstrated that phototrophic bacteria appear to be more abundant in oligotrophic ocean water in comparison to non-phototrophic bacteria which are more likely to occur in nutrient-rich environments, confirming the competitive advantages of phototrophic bacteria over non-phototrophic bacteria in nutrient-poor environments [54]. Nutritional status is an important regulatory factor of photosynthetic pigment synthesis, which has been proven in experiments with growth in conditions of low carbon or nutrients, resulting in transcriptional induction of bacteriochlorophyll synthesis [8, 100].

Based on the genome annotations, carotenoids can also be synthetized by the strains. This was confirmed by the spectra of the pigment extracts, which showed peaks at ~450 and ~475 nm and a shoulder peak at ~422 nm (Figure S2). The presence of carotenoids was also confirmed by the identification of two secondary metabolite gene clusters for the biosynthesis of carotenoids using antiSMASH [12]. While the pigment extracts of strains R-67883, R-67880^{T}, R-36501, R-36677^{T}, R-39604 and R-39161^{T} were yellow and only the R-39594^{T} extract showed a red-brown color, a highly similar spectrum could be observed for all strains, with near-identical peak maxima. Both BChl *a* and carotenoid pigments can have a role in the protection against UV radiation, while carotenoids in particular may contribute to membrane stability [27, 35, 92].

Environmental survival

Different strains (i.e. R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T}) originate from microbial mats, which are known to have a multi-layered structure in which different types of microorganisms live together and produce by-products that can serve as an energy source for other organisms in the microbial mat [93]. In the aquatic microbial mat that strains R-39604, R-39161^{T} and R-39594^{T} were isolated from, carbohydrates and other biopolymers produced by other organisms (e.g. cyanobacteria and microalgae) may be used as an energy source by these strains. In order to use these carbohydrates, these organisms should have the ability to degrade

these complex polymers. The genomes of strains R-39604, R-39161^T and R-39594^T hint at their ability to degrade such complex polymers by the presence of a large number of peptidases (~34, ~35 and ~35 per Mb, respectively) and glycosyl hydrolases (GH) (~6 per Mb for each strain). These strains all have multiple GH43 family enzymes (~1 per Mb), which are absent in the strains obtained from soil (i.e. R-67883, R-67880^T and R-36501). This could be explained by the activity of this family of enzymes with mainly α-L-arabinofuranosidases, endo-α-L-arabinanases, β-D-xylosidases and 1,3-β-galactosidases [75]. Their overall abundance in nature suggests an important role in accessing many complex glycans, which can be expected to be found in aquatic microbial mats [75]. Strain R-36677^T isolated from a microbial mat at the bank of a frozen lake, also shows the presence of GH43 enzymes (~1 per Mb). This could potentially be linked to the presence of a fungus, *Thelebolus microsporus*, in the black mat this strain was isolated from [85]. Fungi have the ability to produce complex carbohydrates, so in order for the bacteria to use these carbohydrates as an energy source, they will need the same enzymes to break these carbohydrates down [69].

Other compounds important in microbial mats are extracellular polymeric substances (EPS). These are rich in proteins and carbohydrates and have a function in supporting the structure of the mat, but also in substrate attachment [11]. They are likely produced by cyanobacteria and microalgae in aquatic microbial mats (i.e. isolation source of strains R-39604, R-39161^T and R-39594^T), but could also be produced by fungi (present in the same sample as strain R-36677^T) [11]. In case of strains R-67883 and R-67880^T, however, a single genus of microalgae, *Stichococcus*, has been detected in the soil sample the strains were isolated from [106]. To have the ability to attach themselves to the soil particles, these strains will probably need to produce EPS themselves as well, by using glycosyl transferases (GT). Especially glycosyl transferases of the GT2 and GT4 families are abundant in strains R-67883 and R-67880^T (~2 and ~3 per Mb for each strain, respectively), which might be linked to this EPS production.

Next to attachment, motility is an important bacterial feature and has been observed in the genome annotations of all strains, as all genes for flagellar assembly are present. However, not all strains showed motility during the laboratory experiments (Table 4).

In the dry and cold terrestrial Antarctic environment, different strategies are needed for bacteria to cope with the extreme environmental conditions. One strategy could be the expression of GH23 family enzymes, which are lytic transglycosylases, known to catalyze the non-hydrolytic cleavage of peptidoglycan in the bacterial cell wall [29]. These enzymes can have a role in many functions, e.g. synthesis, remodeling and degradation of the cell wall; for the detection of

cell wall-acting antibiotics; and for the insertion of secretion systems and flagellar assemblies into the cell wall [29]. This could explain the high number of GH23 family enzymes in strains R-67883, R-67880^T and R-36501 (~2 per Mb for each strain), which were isolated from exposed soil and gravel samples.

Besides GTs, GHs and peptidases, CEs, CBMs, AAs and PLs were detected in the annotated genomes as well (Table 2). Especially peptidases of the M20 and S9 families were more abundant in comparison to other peptidase families, as these were together responsible for one-fifth of all peptidases in the strains. Both peptidase families are ubiquitous and have a wide range of activities (i.e. metallopeptidases and aminohydrolases for M20 peptidases and serine peptidases for the S9 family of peptidases) [79].

Sulfatases were also present in the genome annotations, albeit in low amounts of ~1 per Mb in all strains (Table 2). These sulfatases were classified as either formylglycine-generating enzymes required for sulfatase activity (present in all strains except R-67883 and R-67880^T), or the ergothioneine biosynthesis protein EgtB, which was present in all strains. The former has a role in activating newly synthetized inactive sulfatases by converting a cysteine residue in the active site to formylglycine (oxoalanine) [79]. EgtB on the other hand, is a non-heme iron dependent sulfoxide synthase, involved in the biosynthesis of ergothioneine [99]. Ergothioneine is involved in protection against environmental stress by peroxides, superoxides and heavy metals, while it also plays a role in protection in low environmental nutrient conditions [26]. Other functions include maintaining microbial redox homeostasis and physiological adaptations in the secondary metabolism in response to environmental changes [26]. It seems that mainly peptidases, which have a function in protein degradation, are most abundant in these strains in comparison to other enzymes (Table 2).

Nitrogen, sulfur, phosphor and potassium metabolism

Based on the annotated genomes, the use of nitrogen sources is mainly restricted to ammonia, which can be converted to L-glutamine or L-glutamate in all strains by glutamine synthetase and glutamate dehydrogenase, respectively. Additionally, nitroalkanes may be used by converting them into nitrite using nitronate monooxygenase. Only strain R-39594^T shows the potential to use nitrile as a nitrogen source by converting it into ammonia using nitrilase. The presence of nitrilases in bacteria is rare, with only 6.8% of bacteria expressing nitrilases [22]. Little is known about these enzymes in bacteria, except for their role in supplying a nitrogen source for growth [22]. As predicted by the annotations, ammonia is the main nitrogen source

of the strains. When nitrogen availability is low, the bacteria appear to be capable of controlling the nitrogen levels by the GlnALG operon, consisting of GlnA, GlnL and GlnG (i.e. glutamine synthetase, a kinase and an effector protein, respectively) [68].

As hinted by the annotated genomes, the ability of the novel strains to use sulfur sources likely relies on the assimilatory sulfate to sulfide reduction pathway, for which they also have sulfate transporters. The bacteria are also able to transport phosphate, which is an essential nutrient for bacteria. To cope with phosphate limitation in the Antarctic environment, a Pho (phosphate) regulon is present in the bacterial genomes, which is controlled by a two-component system. This system consists of an inner-membrane histidine kinase sensor protein (PhoR) and a cytoplasmic transcriptional response regulator (PhoB) [97]. These can further lead to the transcription of phosphate scavengers, like phospholipases (PhoD), which are amongst the most common enzymes induced in response to phosphate limitation in bacteria [97]. Phosphonate can only be transported by R-67883 and R-67880^{T} and could serve as an extra source of phosphorus in the nutrient-scarce soil.

Potassium is another crucial component for bacteria and is for example used as an activator of intracellular enzymes and a regulator of internal pH [32]. The Kdp system, detected in the genome annotations of strains R-67883, R-67880^{T}, R-36501 and R-36677^{T}, could be important when potassium levels in the environment are low. Kdp is a P-type ATPase, consisting of the subunits KdpA, KdpB, KdpC and KdpF [32]. Only KdpF was absent from the genome annotations, which could be explained by this subunit not being essential in the functioning of the complex [32]. Other Kdp proteins present in the genomes are KdpD and KdpE, which are a sensor kinase and response regulator protein, respectively. These have a role in regulating the expression of the Kdp system on a transcriptional level [32].

The secretion system for proteins present in all strains, is the sec-SRP secretion system (SecA-signal recognition particle).

Environmental distribution of *Chioneia* gen.nov.

To determine the distribution of the new taxa on Earth, the 16S rRNA gene sequences of the seven novel strains were subjected to a blast search. A subsequent phylogenetic analysis including the most closely related sequences revealed that only five similar sequences have been deposited to date (Figure S3). Furthermore, these sequences (i.e. FJ946587, KC286808, HQ327177, MH041129 and MK346173) all originated from cold environments including glaciers and snow from China, and arctic snow, indicating these taxa are limited to cold extreme habitats.

Interestingly, one of the highly similar sequences (i.e. MK346173) originated from the genome sequence (i.e. VJWA00000000) of a recently cultured organism. Its 16S rRNA gene sequence clustered separately in between the sequences of our seven novel strains (Figure S3). Although deposited as *Sandarakinorhabdus* sp. LB1R16, this organism clearly belongs to the *Chioneia* gen. nov. which we propose here. To determine whether this strain grouped with one of the four novel species described here, ANI and *is*DDH values were determined between *Sandarakinorhabdus* sp. LB1R16 and R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T. These values all were far below the recommended thresholds of 70 and 95% for DDH and ANI, respectively [51, 95, 109]. An additional whole genome-based analysis using bcgTree confirmed the placement of *Sandarakinorhabdus* sp. LB1R16, as a separate branch in between our seven novel strains (Figure S4). Lastly, the genome of strain LB1R16 was screened for the presence of phototrophic genes, which were found in the genomes of R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T. The genome of *Sandarakinorhabdus* sp. LB1R16 also contained all necessary genes for bacteriochlorophyll-dependent phototrophy relying on a type II reaction center. A subsequent phylogenetic analysis of *pufLM* genes clearly showed that strain LB1R16 grouped as a separate branch in between the 7 new strains analyzed here (Figure S5). Taken together, these results indicate that *Sandarakinorhabdus* sp. LB1R16 represents a fifth novel species in *Chioneia* gen. nov.

Because a blast search will miss out on sequences from SRA datasets, the environmental distribution of the new taxa was also determined using an IMNGS search [59]. A total of 40.458 SRA datasets were found to contain reads more than 95% similar to those of the novel strains. Of these, 12.586 had an abundance of more than 0.01% of the new taxa. Subsequent phylogenetic placement of these short read sequences revealed that only 69 of the latter SRA datasets actually contained reads closely related to those of the seven novel strains. The other reads clustered together with other taxa in the family *Sphingomonadaceae*. These 69 datasets all originate from cold environments including (sub)alpine soil, Antarctic nunatak soil and glaciers. Whereas the abundance in the majority of datasets remained low (i.e. below 1%), interestingly, a very high abundance of reads was found in multiple datasets from rhizosphere of *Saxifraga oppositifolia* (up to 41%) , collected in Spitsbergen (Norway) and a dataset from roots and leaves of *Boechera stricta* (10%) collected in the Rocky Mountains (USA). Both are perennial herbs native to Arctic-Alpine regions (*Saxifraga oppositifolia*) [89] and montane regions of western North America (*Boechera stricta*) [61]. The high abundance of *Chioneia* species present in the proximity or on these plants may thus indicate a close cooperation between them.

Because the 16S rRNA gene is only a small fraction of an entire genome, the complete set of genes of the seven strains were also used to detect their presence in the environment. Using the Single Genome vs. Metagenomes search option of the IMG/MER database, their presence was screened in 41.056 metagenome datasets. However, no results were obtained. For other organisms, however, presence in the environment could be detected, indicating the analysis does work. Although it thus may seem that members of *Chioneia*. gen. nov. are absent in metagenomes, it should be noted that, to date, only very little metagenomes exist of environments where the taxa are found in, as observed with 16S rRNA gene analyses. Therefore it is likely that the taxa will be detected in the future, when more metagenomic sequencing is performed on samples from polar and alpine environments.

Based on 16S rRNA gene analysis, genome analyses and phenotypic and chemotaxonomic tests, the seven Antarctic isolates, R-67880^{T}, R-67883, R-36501, R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T}, form a new genus in the family *Sphingomonadaceae* for which the name *Chioneia* gen. nov. is proposed (Table 6), referring to the cold habitats these organisms are found in. Strains R-67880^{T}, R-67883 and R-36501 form the new species *C. frigida* (Table 6). Strains R-39604 and R-39161^{T} form the new species *C. brumae* (Table 7). Strain R-36677^{T} is currently the only representative of the new species *C. hiemis* and strain R-39594^{T} represents the new species *C. algorism* (Table 7).

Emended description of *Polymorphobacter multimanifer* Fukuda et al. 2014

The description is as given by Fukuda et al. [36] with the following amendments. Growth is observed at pH 7-10.5. Does not grow in a microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere. Cells grow on R2A, R2A/10, R2A/100, TSA, NA and weakly in R2B. Cells do not grow on MA. Hydrolysis of Tweens 40 and 60 is positive and hydrolysis of Tweens 20 and 80 is variable. Hydrolysis of DNA is positive. Does not hydrolyze starch, casein and CMC. Produces BChl *a*, but no flexirubin. *PufLM* genes are present. Susceptible to ampicillin and vancomycin, but resistant to bacitracin. The DNA G+C content of the type strain DSM 102189^T, as derived from the genome sequence, is 65.86%, its approximate genome size 3.92 Mb and its IMG-ER genome ID Ga0373304.

Emended description of *Polymorphobacter fuscus* Jia et al. 2015

The description is as given by Jia et al. [43] with the following amendments. Cells are autotrophic as the complete Calvin-Benson-Bassham pathway, including the key enzyme ribulose-1,5-biphosphate carboxylase/oxygenase, is present in the genome. Growth is observed at 0-37 °C and pH 5.5-11. Grows in a microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere. Cells grow on R2A/10, R2A/100 and NA. Cells do not grow on MA. Hydrolysis of Tweens 40 and 60 is positive and hydrolysis of Tweens 20 and 80 is variable. Hydrolysis of casein and DNA is positive. Does not hydrolyze starch and CMC. Does not produce flexirubin. Susceptible to tetracycline, but resistant to bacitracin. The DNA G+C content of the type strain DSM 105347^T, as derived from the genome sequence, is 66.82%, its approximate genome size 3.36 Mb and its IMG-ER genome ID Ga0373281.

Emended description of *Sandarakinorhabdus limnophila* Gich and Overmann 2006 emend. Kim et al. 2016

The description is as given by Gich and Overmann [37] emend. Kim et al. [50] with the following amendments. Colonies are entire, round, raised, opaque, round-edged and have a mat surface, when grown on R2A. Growth is observed at 0-37 °C and a pH between 6-10 with an optimum of pH 7.5. Does not grow in a microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere. Cells grow on R2A, R2A/10, R2A/100 and in R2B. Cells do not grow on TSA, NA and MA. Hydrolysis of Tweens 40, 60 and 80 is positive and hydrolysis of Tween 20 is variable. Does not hydrolyze casein, starch, CMC and DNA. Does not produce flexirubin. Susceptible to tetracycline, vancomycin, ampicillin, chloramphenicol and gentamicin, but resistant to bacitracin. The DNA G+C content of the type strain DSM 17366^T, as derived from

737 the genome sequence, is 64.49%, its approximate genome size 2.61 Mb and its IMG-ER
738 genome ID Ga0002310.
739
740 **Emended description of *Sphingomonas spermidinifaciens* Feng et al. 2017**
741 The description is as given by Feng et al. [34] with the following amendments. Growth occurs
742 at 0-45 °C and pH 5.5-11. Grows in a microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere.
743 Cells grow on R2A/10 and R2A/100. Hydrolysis of Tween 60 is positive and hydrolysis of
744 Tween 20 is variable. Hydrolysis of starch, casein and DNA is positive. Does not hydrolyze
745 CMC. Produces BChl *a* and carotenoids, but no flexirubin. *PufLM* genes are present.
746 Susceptible to tetracycline, vancomycin and gentamicin, intermediate sensitive to ampicillin,
747 but resistant to bacitracin and chloramphenicol. The DNA G+C content of the type strain 9NM-
748 10^T, as derived from the genome sequence, is 67.97%, its approximate genome size 3.64 Mb
749 and its NCBI genome ID NWMW00000000.
750

Abbreviations

AA, auxiliary activities

ANI, average nucleotide identity

BChl *a*, bacteriochlorophyll *a*

CBM, carbohydrate-binding module

CE, carboxyl esterase

CMC, carboxymethylcellulose

CRISPR, clustered regularly interspaced short palindromic repeat

EPS, extracellular polymeric substances

GH, glycosyl hydrolase

IMG-ER, Integrated Microbial Genomes-Expert Review

*is*DDH, *in silico* DNA-DNA hybridization

MA, marine agar

MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

NA, nutrient agar

NCBI, National Center for Biotechnology Information

PES, Princess Elisabeth Station

PL, polysaccharide lyase

pufLM, genes for the type 2 photosynthetic reaction center light and medium subunits

R2A, Reasoner's 2A agar

R2B, R2A broth

TSA, tryptone soya agar

Funding

This research made use of infrastructure supplied by EMBRC Belgium (FWO project GOH3817N and I001621N). The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the Flemish Supercomputer Centre (VSC) funded by Ghent University, the Hercules Foundation and the Flemish Government – department EWI.

Author contributions

Conceived and designed the experiments: GT, DG, LL, AW. Performed the experiments: GT, DG, LL. Analyzed the data: GT, DG, LL, AW. Wrote the paper: GT, DG, AW. All authors approved the final manuscript.

Acknowledgements

We thank the Oxford Genomics Centre at the Wellcome Centre for Human Genetics for the generation and initial processing of the sequencing data. We thank Peter Vandamme for advice on the etymology of the new names proposed.

Conflicts of interest statement

The authors declare that there are no conflicts of interest.

Ethical Committee approval for human or animal research

NA

Figure captions

Figure 1. Maximum-likelihood phylogenetic tree (1000 bootstrap replicates) of near-complete 16S rRNA gene sequences showing the relationship between strains R-67880^T, R-67883, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and closely related type strains of the family *Sphingomonadaceae*. *Paracraurococcus ruber* NS89^T and *Craurococcus roseus* DSM 15488^T were used as an outgroup. Only bootstrap values greater than 70% are shown. Scale bar indicates 0.05 substitutions per nucleotide position.

Figure 2. Maximum-likelihood phylogenetic tree (1000 bootstraps) based on 107 single-copy core genes of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and closely related type strains of the families *Sphingomonadaceae* and Erythrobacteraceae. Only bootstrap values higher than 70% are shown. The scale bar indicates 0.1 substitutions per site. *Paracraurococcus ruber* JCM 9931^T was used as an outgroup.

Figure 3. Maximum-likelihood phylogenetic tree (1000 bootstraps) of the *pufLM* gene sequences of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and closely related type strains of the families *Sphingomonadaceae* and Acetobacteraceae. Only bootstrap values higher than 70% are shown. The scale bar indicates 0.10 substitutions per site. *Erythrobacter longus* DSM 6997^T and *Erythromicrobium ramosum* DSM 8510^T were used as an outgroup.

List of Tables

Table 1. Overview of the samples of which the strains in this study were isolated from, including information about the sampling location and coordinates, sampling time and sample description.

Sample	Isolated strain(s)	Sampling location	Sampling site coordinates	Sampling time	Description of sample area	Reference
KP43	R-67880^T, R-67883	Utsteinen, Sør Rondane Mountains	71° 56′ 47.3″ S, 23° 20′ 44.6″ E	February 2009	Brown soil with dark green fragments	[106]
BB50	R-36501	Utsteinen, Sør Rondane Mountains	71° 57’ S, 23° 20’ E	January 2007	Gravel and green microbial/algal mat	[85]
BB115	R-36677^T	Utsteinen, Sør Rondane Mountains	71° 57’ S, 23° 20’ E	January 2007	Black mat on gravel and rock debris from the bank of a frozen lake	[85]
SK5	R-39604, R-39161^T, R-39594^T	Naka Tempyo, Syowa Oasis	69° 28' S, 39° 40' E	January 2007	Littoral epipsammic and interstitial microbial mat, brown or orange pigmented on top with a green surface layer, sampling depth 0.1 m	[86]

Table 2. Genome characteristics of the different new strains and type strains of closely related species of the family *Sphingomonadaceae*.

Strains: 1, R-67880^{T}; 2, R-67883; 3, R-36501; 4, R-36677^{T}; 5, R-39604; 6, R-39161^{T}; 7, R-39594^{T}; 8, *Polymorphobacter multimanifer* DSM 102189^{T}; 9, *Polymorphobacter fuscus* DSM 105347^{T}; 10, *Polymorphobacter arshaanensis* DJ1R-1^{T}; 11, *Sandarakinorhabdus limnophila* DSM 17366^{T}; 12, *Sphingomonas spermidinifaciens* 9NM-10^{T}. NA, no data available.

Characteristic/Strain	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Size (Mb)	4.36	4.36	4.28	3.76	3.05	3.04	3.07	3.92	3.36	3.26	2.61	3.64
Contigs	41	56	47	21	3	10	20	72	3	10	11	5
N50 (kb)	817	275	449	698	225	153	432	137	2786	1178	422	2102
L50	3	5	4	2	1	1	3	12	1	2	3	1
DNA G+C content (%)	67.41	67.41	67.47	67.76	68.89	68.89	68.19	65.86	66.82	64.98	64.49	67.97
Genes (total)	4320	4336	4241	3741	3068	3069	3104	3765	3232	3112	2586	3463
CDS (coding)	4258	4273	4179	3681	3006	3005	3043	3704	3172	3035	2534	3376
Genes (RNA)	56	57	56	54	57	56	53	53	55	54	52	52
5S rRNA	1	1	1	1	1	1	1	1	1	1	1	1
16S rRNA	1	1	1	1	1	1	1	1	1	1	1	1
23S rRNA	1	1	1	1	1	2	1	1	1	1	1	1
tRNA	46	47	46	46	47	48	46	47	46	48	44	46
Accession number	Ga0401290	Ga0401289	Ga0401291	Ga0401292	Ga0401293	Ga0427977	Ga0427589	Ga0373304	Ga0373281	SIHO00000000	Ga0002310	NWMW00000000
CRISPR (questionable)	0	0	0	1	1	1	2	2	2	1	2	0
CRISPR (high evidence)	0	0	0	0	0	0	0	0	0	0	1	0
Prophage region (incomplete)	4	3	4	1	2	2	3	5	1	3	0	2
Glycosyl hydrolases	30	29	29	25	17	17	18	25	36	15	9	64
Gycosyl transferases	26	26	26	13	14	14	12	21	24	13	8	32
Carboxyl esterases	5	5	5	3	2	2	3	2	4	2	0	3
Polysaccharide lyases	2	2	2	3	2	2	1	1	1	0	0	0
Carbohydrate-binding modules	0	0	0	2	1	1	0	0	0	0	0	0
Auxiliary activities	3	3	3	2	1	1	1	2	5	5	2	3
Peptidases	127	126	124	115	105	105	107	130	124	NA	120	NA
Sulfatases	1	1	1	2	2	2	4	10	6	NA	7	NA

Table 3. *In silico* DNA-DNA hybridization (*is*DDH) and pairwise average nucleotide identity (ANI) values (in %) between genomes of the different new strains and type strains of closely related species of the family *Sphingomonadaceae*.

Strains: 1, R-67883; 2, R-67880^T; 3, R-36501; 4, R-36677^T; 5, R-39604; 6, R-39161^T; 7, R-39594^T; 8, *Polymorphobacter multimanifer* DSM 102189^T; 9, *Polymorphobacter fuscus* DSM 105347^T; 10, *Polymorphobacter arshaanensis* DJ1R-1^T; 11, *Sandarakinorhabdus limnophila* DSM 17366^T; 12, *Sphingomonas spermidinifaciens* 9NM-10^T. Values in the upper and lower triangle (in bold) correspond to *is*DDH and ANI values, respectively.

Strain	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1		99.90	99.80	22.10	22.00	22.00	22.20	19.70	19.10	19.60	18.50	18.30
2	**99.98**		99.70	22.10	22.00	22.00	22.10	19.60	19.10	19.60	18.50	18.20
3	**99.98**	**99.97**		22.10	22.00	22.00	22.20	19.60	19.10	19.60	18.50	18.20
4	**78.85**	**78.96**	**78.87**		33.50	33.50	33.50	19.20	19.10	19.70	18.50	18.40
5	**78.87**	**79.05**	**78.90**	**87.39**		100.00	32.40	18.60	19.10	19.50	18.30	18.60
6	**78.97**	**79.02**	**78.87**	**87.38**	**99.98**		32.40	18.60	19.10	19.50	18.30	18.60
7	**79.04**	**78.88**	**78.98**	**87.53**	**86.96**	**86.82**		18.80	19.10	19.60	18.70	18.60
8	**72.85**	**72.92**	**72.95**	**72.77**	**72.47**	**72.38**	**72.50**		18.90	18.60	18.50	18.40
9	**73.81**	**73.72**	**73.86**	**74.07**	**73.89**	**74.07**	**73.56**	**74.62**		19.00	18.60	18.40
10	**74.39**	**74.29**	**74.36**	**74.49**	**74.40**	**74.49**	**74.52**	**72.26**	**73.70**		18.70	18.40
11	**71.65**	**71.96**	**71.61**	**71.46**	**71.40**	**71.46**	**71.42**	**72.84**	**73.55**	**71.82**		18.30
12	**71.58**	**71.77**	**71.60**	**71.78**	**71.92**	**71.78**	**71.76**	**70.12**	**70.63**	**71.28**	**69.61**	

Table 4. Phenotypic characteristics of the new strains and type strains of closely related species of the family *Sphingomonadaceae*

Strains: 1, R-67883; 2, R-67880^T; 3, R-36501; 4, R-36677^T; 5, R-39604; 6, R-39161^T; 7, R-39594^T; 8, *Polymorphobacter multimanifer* DSM 102189^T; 9, *Polymorphobacter fuscus* DSM 105347^T; 10, *Sandarakinorhabdus limnophila* DSM 17366^T; 11, *Sphingomonas spermidinifaciens* DSM 27571^T.

Results of API 20 NE, API 50 CH and API ZYM assays can be found in Tables S1 to S3. Strains have an entire colony and a rounded cell edge and are positive for catalase, hydrolysis of Tweens 40 and 60 and growth on 10 and 100 times diluted R2A. All strains are negative for Gram staining and the formation of spores. Strains are susceptible to tetracycline (30 µg) and vancomycin (30 µg). All strains have bacteriochlorophyll *a* and carotenoid pigments, but no flexirubin. Strains show no growth in an anaerobic atmosphere. +, positive; -, negative; w, weak; R, resistant; S, sensitive; I, intermediate sensitivity; v, variable; NG, no bacterial growth on medium.

Characteristic	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Colony morphology:											
Shape	round	round	round	round	round	round	round	round-irregular	round	round	curled
Elevation	convex	convex	convex	convex	convex	convex	convex	raised	raised	raised	convex
Transparancy	opaque	opaque	opaque	opaque	opaque	opaque	opaque	opaque	opaque	opaque	(semi-) transparent
Surface	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + mat	smooth + shiny
Size (mm)	~1	~1	~1	~1	≤1	≤1	<<1	~1	~1	<<1	~3-5
Color	orange-yellow	orange-yellow	orange-yellow	orange	orange-yellow	orange-yellow	red	red-brown	red-brown	red	yellow
Cell morphology:											
Shape	rod	rod	rod	rod-coccus	rod	long rod	rod	coccus + rod	rod-coccus	rod	rod
Size (µm)	2x1	2x1	2x1	2x1.5	2x1	2-3x1	2x1	1.5x1	1.5x1	2x1	2x1
Appearance	single-pair	single-pair	single-pair	single	single-pair-chain	single-pair-chain	pair-chain-heap	single-pair	single-pair	single-pair	single-pair
Motility	+	+	+	-	-	+	-	+	-	-	+
Oxidase	+	+	+	+	-	-	-	+	+	-	+
Growth on:											
TSA	w	w	w	-	w	-	-	+	+	-	+
NA	-	w	w	-	w	-	-	+	+	-	+
MA	-	-	-	-	-	-	-	-	-	-	+
R2B	+	+	+	+	+	+	+	w	+	+	+
Microaerobic	+	+	+	+	+	+	+	-	+	-	+
Growth conditions:											
Temperature range (°C)	0-30	0-30	0-30	4-30	0-35	0-30	4-28	0-30	0-37	0-37	0-45
Temperature optimum (°C)	28	25	25	28	25	25	20	28	30	30	30
Salinity range (% NaCl, w/v)	0-1	0-1	0-1	0-0.25	0-0.25	0-0.25	0	0-0.5	0-2	0-0.25	0-2.5
Salinity optimum (% NaCl, w/v)	0	0	0	0	0	0	0	0	0.25	0	0
pH range	5.5-10	5.5-10	5.5-10	6-9.5	6-8	6-8	6-10	7-10.5	5.5-11	6-10	5.5-11
pH optimum	7	7	7	8	6.5	6.5	6.5	7	7.5	7.5	7.5
Susceptible to:											

Characteristic	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Ampicillin	S	S	S	S	S	S	S	S	R	S	I
Bacitracin	R	R	R	S	S	R	R	R	R	R	R
Chloramphenicol	S	S	R	S	S	S	S	S	S	S	R
Gentamicin	S	S	I	S	S	S	S	S	S	S	S
Hydrolysis of:											
Tween20	+	+	+	v	v	v	v	v	v	v	v
Tween 80	+	+	+	v	v	+	+	v	v	+	+
Casein	-	+	+	-	-	-	-	-	+	-	+
Starch	+	+	+	-	-	-	+	-	-	-	+
Carboxymethylcellulose	-	-	-	NG	NG	-	-	-	-	-	-
DNA	-	-	-	NG	NG	NG	NG	+	+	-	+

841

842

Table 5. Fatty acid composition of strains R-67880^T, R-67883, R-36501, R-39594^T, R-36677^T, R-39161^T, R-39604 and type strains of closely related species of the family *Sphingomonadaceae*. Values shown are percentages of total fatty acids obtained in this study.

Strains: 1: R-67880 ^{T}, 2: R-67883, 3: R-36501, 4: R-39594^T, 5: R-36677^T, 6: R-39161 ^{T}, 7: R-39604, 8: *Polymorphobacter multimanifer* DSM 102189^T, 9: *Polymorphobacter fuscus* DSM 105347^T, 10: *Sphingomonas spermidinifaciens* DSM 27571^T, 11: *Sandarakinorhabdus limnophila* DSM 17366^T.TR: Trace amount (i.e. < 1%). [a]: Summed features represent groups of multiple fatty acids that cannot be separated by the Microbial Identification System. [b]: Unknown fatty acid with chain length (ECL) 18.814. -: Not detected

	1	2	3	4	5	6	7	8	9	10	11
$C_{12:00}$	6.0	3.2	8.9	21.4	5.7	29.8	4.1	11.7	4.2	2.3	7.1
$C_{12:0}$ 2-OH	9.2	8.7	9.9	12.4	5.8	11.4	7.4	2.1	TR	-	-
$C_{13:0}$ 2-OH	-	TR	TR	TR	-	TR	TR	1.3	TR	-	-
$C_{14:00}$	TR	TR	TR	TR	TR	-	TR	-	-	TR	TR
$C_{14:0}$ 2-OH	1.7	1.7	1.7	6.8	3.5	5.6	2.9	18.3	12.3	7.5	5.3
$C_{15:00}$	TR	TR	TR	TR	TR	-	TR	TR	TR	TR	TR
$C_{15:0}$ 2-OH	-	TR	TR	TR	-	1.1	-	5.7	1.6	TR	TR
$C_{15:1}$ ω6c	-	-	-	-	-	-	-	TR	1.5	-	-
$C_{16:00}$	3.8	3.3	2.8	4.4	4.9	3.6	7.1	4.5	2.1	10.4	9.2
$C_{16:0}$ 2-OH	-	-	-	-	-	-	TR	-	-	-	1.8
$C_{16:1}$ 2-OH	3.0	3.4	3.3	3.0	2.1	2.1	TR	-	-	-	-
$C_{16:1}$ ω5c	4.2	4.8	4.0	3.2	4.1	2.3	3.6	TR	1.0	1.5	1.6
iso-$C_{16:0}$ 3-OH	-	-	TR	-	-	TR	-	1.0	TR	-	TR
$C_{17:1}$ ω6c	6.3	6.5	4.7	2.8	3.2	3.5	5.1	9.7	18.8	3.0	6.9
$C_{17:1}$ ω8c	-	TR	TR	-	-	-	-	1.4	TR	-	-
10-methyl-$C_{17:0}$	-	-	-	TR	-	-	-	-	-	-	-
anteiso-$C_{17:0}$	-	-	TR	TR	-	-	-	1.1	-	-	-
$C_{18:00}$	-	TR	TR	TR	-	-	-	-	-	TR	2.1

	1	2	3	4	5	6	7	8	9	10	11
$C_{18:1}$ 2-OH	0.8	1.2	1.0	-	0.7	0.7	2.5	-	-	-	-
$C_{18:1}$ ω5c	-	-	-	-	-	-	-	-	-	1.6	-
$C_{18:1}$ ω7c	34.5	33.7	33.5	21.3	34.9	22.0	36.7	12.7	20.6	58.3	29.2
11-methyl-$C_{18:1}$ ω7c	-	-	-	-	-	-	-	-	-	6.3	-
$C_{19:00}$	-	-	-	-	-	-	-	-	-	-	-
10-methyl-$C_{19:0}$	-	-	TR	-	-	-	TR	-	-	-	-
cyclo-$C_{19:0}$ ω8c	-	-	-	-	-	-	-	-	-	TR	-
Summed Features[a]											
3 $C_{16:1}$ ω7c and/or iso-C_{15} 2-OH	29.4	31.1	25.0	19.6	34.1	16.5	27.2	34.5	25.3	6.3	34.0
5 $C_{18:2}$ ω6, 9c and/or anteiso-$C_{18:0}$	-	-	TR	-	-	-	-	-	-	-	-
Unknown ECL 18.814[b]	-	-	2.0	TR	-	-	-	3.1	-	-	-

849

850 **Table 6.** Description of *Chioneia* gen. nov. and the type species of the genus *Chioneia frigida* sp. nov.

Genus name	*Chioneia*	*Chioneia*
Species name		*Chioneia frigida*
Genus status	gen. nov.	-
Genus etymology	*Chioneia* (Chi.o'ni.a Gr. fem. n. *chion* snow or Gr. n. Chione, goddess of snow; N.L. fem. n. *Chioneia* referring to the cold environments where these bacteria have been isolated and are found in).	-
Type species of the genus	*Chioneia frigida*	-
Specific epithet	-	*frigida*
Species status	-	sp. nov.
Species etymology	-	*Chioneia frigida* (fri'gi.da. L. fem. adj. *frigida*, cold).
Description of the new taxon and diagnostic traits	Cells are strictly aerobic. Cells stain Gram-negative and divide by binary fission. Cells are motile, non-sporulating and catalase positive. Produce BChl *a* and carotenoid pigments, but no flexirubin. *PufLM* genes and other genes involved in bacteriochlorophyll-dependent phototrophy relying on a type II reaction center are present. Hence, members of *Chioneia* gen. nov. are aerobic anoxygenic photoheterotrophs. The core metabolism includes the glycolysis, the tricarboxylic acid cycle, the pentose phosphate and the assimilatory sulfate to sulfide reduction pathways. Genome sizes range from 3.04 Mb to 4.36 Mb. DNA G+C content varies between 67.41% and 68.89%. The genus is classified in the family Sphingomonadaceae based on 16S rRNA gene and whole genome analyses, and phenotypic and chemotaxonomic characteristics. The type species is *C. frigida*, represented by the type strain R-67880^T (= LMG 31952^T = CECT 30379^T).	**Description of *Chioneia frigida* sp. nov.** *Chioneia frigida* (fri.gi'da. L. fem. adj. *frigida*, cold). Cells are Gram-negative, catalase- and oxidase-positive, non-sporulating, rod-shaped (2x1 µm), appear single or as pairs and are motile by gliding. Colonies are orange-yellow, entire, round, convex, opaque, round-edged, have a smooth mat surface and are 1 mm in diameter, when grown on R2A. Growth is observed at 0-30 °C with an optimum at 25 °C, and a pH between 5.5-10 with an optimum of pH 7. NaCl is not required for growth, but cells can tolerate up to 1% (w/v) NaCl. Grows in an aerobic and microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere, but not in an anaerobic atmosphere (80% N_2, 10% CO_2 and 10% H_2, 10 ppm O_2). Cells grow on R2A/10, R2A/100 and in R2B. Cells grow weakly on TSA and NA and do not grow on MA. Hydrolysis of Tweens 20, 40, 60 and 80, casein and starch is positive. Does not hydrolyze carboxymethylcellulose and DNA. Produces BChl *a* and carotenoid pigments, but no flexirubin. *PufLM* genes are present. In the API ZYM gallery, the type strain R-67880^T is positive for alkaline phosphatase, esterase (C4), α-chymotrypsin and β-glucosidase activities. Lipase (C14), cystine arylamidase, β-galactosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, α-mannosidase, α-fucosidase, esterase lipase (C8), leucine arylamidase, valine arylamidase, trypsin, acid phosphatase, naphtol-AS-BI-phosphohydrolase, α-galactosidase and α-glucosidase activities are absent.

		According to the API 20 NE gallery, strain R-67880^T is positive for esculin, hydrolysis of gelatin and assimilation of D-glucose, N-acetyl-glucosamine, D-maltose, potassium gluconate and adipic acid. Weakly positive results have been obtained for urease, the PNPG test (β-galactosidase) and assimilation of L-arabinose, D-mannose, D-mannitol, malic acid, trisodium citrate and phenylacetic acid. Negative results have been obtained for nitrate reduction, indole production, glucose fermentation, arginine dihydrolase and the assimilation of capric acid. In the API 50 CH gallery, strain R-67880^T is positive for acid formation from D-adonitol, methyl-βD-xylopyranoside, esculin ferric citrate, D-cellobiose, D-maltose, amidon and gentiobiose. Weakly positive results have been obtained for acid formation from L-arabinose, D-ribose, D-xylose, L-xylose, D-glucose, salicin, D-trehalose, D-melezitose, xylitol and D-fucose. Susceptible to tetracycline, vancomycin, ampicillin, chloramphenicol and gentamicin, but resistant to bacitracin. The major cellular fatty acids are $C_{18:1}$ ω7c and summed feature 3 ($C_{16:1}$ ω7c and/or iso-C_{15} 2-OH) (Table 5). The DNA G+C content of the type strain R-67880^T, as derived from the genome sequence, is 67.41%, its approximate genome size 4.36 Mb and its IMG-ER genome ID Ga0401290. The description of strain R-67883 is identical to that of the type strain, with the following differences: The temperature optimum is at 28 °C. Cells do not grow on NA. Cells do not hydrolyze casein. In the API ZYM gallery, esterase lipase (C8), leucine arylamidase, trypsin, acid phosphatase and α-glucosidase activities are present. According to the API 20 NE gallery, strain R-67883 is weakly positive for the assimilation of potassium gluconate and adipic acid and negative for urease and the assimilation of phenylacetic acid. In the API 50 CH gallery, strain R-67883 is weakly positive for acid formation from D-arabinose, methyl-βD-xylopyranoside, D-galactose and methyl-αD-glucopyranoside and negative for acid formation from D-ribose, D-xylose, L-xylose, D-adonitol and salicin. The DNA G+C content of strain R-67883, as derived from the genome sequence, is 67.41%, its approximate genome size 4.36 Mb and its IMG-ER genome ID Ga0401289. The description of strain R-36501 is identical to that of the type strain, with the following differences: In the API ZYM gallery, leucine arylamidase activity is present. According to the API 20 NE gallery, strain R-36501 is positive for urease and weakly positive for the

		assimilation of potassium gluconate. In the API 50 CH gallery, strain R-36501 is positive for acid formation from D-lactose and potassium 5-ketogluconate, weakly positive for acid formation from D-arabinose, D-galactose, amygdalin, D-cellobiose and glycogen and negative for acid formation from L-xylose, D-glucose and amidon. Cells are intermediate sensitive to gentamicin. The DNA G+C content of strain R-36501, as derived from the genome sequence, is 67.47%, its approximate genome size 4.28 Mb and its IMG-ER genome ID Ga0401291. The type strain R-67880^{T} (= LMG 31952^{T} = CECT 30379^{T}) and strain R-67883 were isolated from an ice-free top surface soil sample taken on Utsteinen ridge, in the neighborhood of the Belgian Princess Elisabeth Station, Sør Rondane Mountains, East Antarctica. Strain R-36501 was isolated from a gravel and a green microbial/algal mat sample taken on Utsteinen, in the neighborhood of the Belgian Princess Elisabeth Station, Sør Rondane Mountains, East Antarctica.
Country of origin	-	Antarctica
Region of origin	-	Strains R-67880^{T} and R-67883 originate from Utsteinen ridge, in the neighborhood of the Belgian Princess Elisabeth Station, East Antarctica. Strain R-36501 also originates from Utsteinen,in proximity of the Princess Elisabeth Station, East Antarctica.
Date of isolation (dd/mm/yyyy)	-	R-67880^{T}: 01/05/2015 R-67883: 01/05/2015 R-36501: 01/06/2007
Source of isolation	-	R-67880^{T}: an ice-free top surface soil sample, KP43.
Sampling date (dd/mm/yyyy)	-	R-67880^{T} (KP43): February 2009 R-67883 (KP43): February 2009 R-36501(BB50): January 2007
Latitude (xx°xx′xx″N/S)	-	R-67880^{T} (KP43): 71° 56′ 47.3″ S R-67883 (KP43): 71° 56′ 47.3″ S R-36501(BB50): 71° 57’ S
Longitude (xx°xx′xx″E/W)	-	R-67880^{T} (KP43): 23° 20′ 44.6″ E R-67883 (KP43): 23° 20′ 44.6″ E R-36501(BB50): 23° 20’ E

Altitude (meters above sea level)	-	R-67880^{T} (KP43): 1362 R-67883 (KP43): 1362 R-36501(BB50): 1318
16S rRNA gene accession nr.	-	R-67880^{T}: MN601355 R-67883: MN601356 R-36501: MN601350
Genome accession number [RefSeq; EMBL; …]	-	R-67880^{T}: IMG-ER Ga0401290 R-67883: IMG-ER Ga0401289 R-36501: IMG-ER Ga0401291
Genome status	-	R-67880^{T}: Incomplete R-67883: Incomplete R-36501: Incomplete
Genome size	-	R-67880^{T}: 4.36 Mb R-67883: 4.36 Mb R-36501: 4.28 Mb
GC mol%	-	R-67880^{T}: 67.41 R-67883: 67.41 R-36501: 67.47
Number of strains in study	Seven	Three
Source of isolation of non-type strains		R-67883: an ice-free top surface soil sample, KP43. R-36501: a gravel and a green microbial/algal mat sample, BB50.
Information related to the Nagoya Protocol	-	-
Designation of the Type Strain	R-67880^{T}	R-67880^{T}
Strain Collection Numbers	LMG 31952^{T} = CECT 30379^{T}	LMG 31952^{T} = CECT 30379^{T}

851

852 **Table 7.** Description of *Chioneia hiemis* sp. nov., *Chioneia brumae* sp. nov. and *Chioneia algoris* sp. nov.

Genus name	*Chioneia*	*Chioneia*	*Chioneia*
Species name	*Chioneia hiemis*	*Chioneia brumae*	*Chioneia algoris*
Genus status	-	-	-
Genus etymology	-	-	-
Type species of the genus	-	-	
Specific epithet	*hiemis*	*brumae*	*algoris*
Species status	sp. nov.	sp. nov.	sp. nov.
Species etymology	*Chioneia hiemis* (hi'e.mis. L. gen. n. *hiemis*, from the cold, from winter).	*Chioneia brumae* (bru'mae. L. gen. n. *brumae*, from the winter, winter cold).	*Chioneia algoris* (al.go'ris. L. gen. n. *algoris*, from coldness).
Description of the new taxon and diagnostic traits	**Description of *Chioneia hiemis* sp. nov.** *Chioneia hiemis* (hi.e'mis. L. gen. n. *hiemis*, from the cold, from winter). Cells are Gram-negative, catalase- and oxidase-positive, non-sporulating, rod-coccoid-shaped (2x1.5 µm), appear as single cells and are motile by gliding. Colonies are orange, entire, round, convex, opaque, round-edged, have a smooth mat surface and are 1 mm in diameter, when grown on R2A. Growth is observed at 4-30 °C with an optimum at 28 °C, and a pH between 6-9.5 with an optimum of pH 8. NaCl is not required for growth, but cells can tolerate up to 0.25% (w/v) NaCl. Grows in an aerobic and microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere, but not in an anaerobic atmosphere (80% N_2, 10% CO_2 and 10% H_2, 10 ppm O_2). Cells grow on R2A/10, R2A/100 and in R2B. Cells do not grow on TSA, NA and MA. Hydrolysis of Tweens 40 and 60 is positive and hydrolysis of Tweens 20 and 80 is variable. Does not hydrolyze casein and starch. Produces BChl *a* and carotenoid pigments, but no flexirubin. *PufLM* genes are present.	**Description of *Chioneia brumae* sp. nov.** *Chioneia brumae* (bru'mae. L. gen. n. *brumae*, from the winter, winter cold). Cells are Gram-negative, catalase-positive and oxidase-negative, non-sporulating, rod-shaped (2-3x1 µm), appear as single cells, in pairs or in chains and are motile by gliding. Colonies are orange-yellow, entire, round, convex, opaque, round-edged, have a smooth mat surface and are ≤1 mm in diameter, when grown on R2A. Growth is observed at 0-30 °C with an optimum at 25 °C, and a pH between 6-8 with an optimum of pH 6.5. NaCl is not required for growth, but cells can tolerate up to 0.25% (w/v) NaCl. Grows in an aerobic and microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere, but not in an anaerobic atmosphere (80% N_2, 10% CO_2 and 10% H_2, 10 ppm O_2). Cells grow on R2A/10, R2A/100 and in R2B. Cells do not grow on TSA, NA and MA. Hydrolysis of Tweens 40, 60 and 80 is positive and hydrolysis of Tween 20 is variable. Does not hydrolyze casein, starch and CMC. Produces BChl *a* and carotenoid pigments, but no flexirubin. *PufLM* genes are present.	**Description of *Chioneia algoris* sp. nov.** *Chioneia algoris* (al.go'ris. L. gen. n. *algoris*, from coldness). Cells are Gram-negative, catalase-positive and oxidase-negative, non-sporulating, rod-shaped (2x1 µm), appear in pairs, chains and heaps and are motile by gliding. Colonies are red, entire, round, convex, opaque, round-edged, have a smooth mat surface and are <<1 mm in diameter, when grown on R2A. Growth is observed at 4-28 °C with an optimum at 20 °C, and a pH between 6-10 with an optimum of pH 6.5. Cells cannot tolerate NaCl. Grows in an aerobic and microaerobic (80% N_2, 15% CO_2 and 5% O_2) atmosphere, but not in an anaerobic atmosphere (80% N_2, 10% CO_2 and 10% H_2, 10 ppm O_2). Cells grow on R2A/10, R2A/100 and in R2B. Cells do not grow on TSA, NA and MA. Hydrolysis of Tweens 40, 60 and 80 is positive and hydrolysis of Tween 20 is variable. Hydrolysis of starch is positive. Does not hydrolyze casein and CMC. Produces BChl *a* and carotenoid pigments, but no flexirubin. *PufLM* genes are present. In the API ZYM gallery, alkaline phosphatase, esterase (C4), esterase lipase (C8) and α-chymotrypsin activities are present.

	In the API ZYM gallery, the type strain R-36677^T is positive for alkaline phosphatase and esterase (C4) activities. Lipase (C14), cystine arylamidase, β-galactosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, α-mannosidase, α-fucosidase, esterase lipase (C8), leucine arylamidase, valine arylamidase, trypsin, α-chymotrypsin, acid phosphatase, naphtol-AS-BI-phosphohydrolase, α-galactosidase, α-glucosidase and β-glucosidase activities are absent. According to the API 20 NE gallery, strain R-36677^T is positive for urease, esculin and the hydrolysis of gelatin. Weakly positive results have been obtained for the PNPG test (β-galactosidase) and the assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, adipic acid, malic acid, trisodium citrate and phenylacetic acid. Negative results have been obtained for nitrate reduction, indole production, glucose fermentation, arginine dihydrolase and the assimilation of capric acid. In the API 50 CH gallery, strain R-36677^T is positive for acid formation from esculin ferric citrate, D-maltose, D-lactose, gentiobiose and potassium 5-ketogluconate. Weakly positive results have been obtained for acid formation from amidon, D-lyxose and D-fucose. Susceptible to tetracycline, vancomycin, ampicillin, bacitracin, chloramphenicol and gentamicin. The major cellular fatty acids are $C_{18:1}$ ω7c and summed feature 3 ($C_{16:1}$ ω7c and/or iso-C_{15} 2-OH) (Table 5). The DNA G+C content of the type strain R-36677^T, as derived from the genome sequence, is 67.76%, its approximate genome size 3.76 Mb and its IMG-ER genome ID Ga0401292.	In the API ZYM gallery, the type strain R-39161^T is positive for alkaline phosphatase, esterase (C4), leucine arylamidase, α-chymotrypsin and acid phosphatase activities. Lipase (C14), cystine arylamidase, β-galactosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, α-mannosidase, α-fucosidase, esterase lipase (C8), valine arylamidase, trypsin, naphtol-AS-BI-phosphohydrolase, α-galactosidase, α-glucosidase and β-glucosidase activities are absent. According to the API 20 NE gallery, strain R-39161^T is positive for esculin and for the PNPG test (β-galactosidase). Weakly positive results have been obtained for urease and the assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, adipic acid, malic acid, trisodium citrate and phenylacetic acid. Negative results have been obtained for nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, hydrolysis of gelatin and the assimilation of capric acid. In the API 50 CH gallery, strain R-39161^T is positive for acid formation from D-adonitol, methyl-βD-xylopyranoside and esculin ferric citrate. Weakly positive results have been obtained for acid formation from D-saccharose and potassium 5-ketogluconate. Susceptible to tetracycline, vancomycin, ampicillin, chloramphenicol and gentamicin, but resistant to bacitracin. The major cellular fatty acids are $C_{12:00}$, $C_{18:1}$ ω7c and summed feature 3 ($C_{16:1}$ ω7c and/or iso-C_{15} 2-OH) (Table 5). The DNA G+C content of the type strain R-39161^T, as derived from the genome sequence, is 68.89%, its approximate genome size 3.04 Mb and its IMG-ER genome ID Ga0427977.	Lipase (C14), cystine arylamidase, β-galactosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, α-mannosidase, α-fucosidase, leucine arylamidase, valine arylamidase, trypsin, acid phosphatase, naphtol-AS-BI-phosphohydrolase, α-galactosidase, α-glucosidase and β-glucosidase activities are absent. According to the API 20 NE gallery, strain R-39594^T is positive for esculin, hydrolysis of gelatin and the PNPG test (β-galactosidase). Weakly positive results have been obtained for the assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, adipic acid, malic acid, trisodium citrate and phenylacetic acid. Negative results have been obtained for nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, urease and the assimilation of capric acid. In the API 50 CH gallery, strain R-39594^T is positive for acid formation from D-adonitol, methyl-βD-xylopyranoside, esculin ferric citrate, D-lactose, gentiobiose, D-turanose and potassium 5-ketogluconate. Weakly positive results have been obtained for acid formation from D-maltose and D-fucose. Susceptible to tetracycline, vancomycin, ampicillin, chloramphenicol and gentamicin, but resistant to bacitracin. The major cellular fatty acids are $C_{12:00}$, $C_{18:1}$ ω7c and summed feature 3 ($C_{16:1}$ ω7c and/or iso-C_{15} 2-OH) (Table 5). The DNA G+C content of the type strain R-39594^T, as derived from the genome sequence, is 68.19%, its approximate genome size 3.07 Mb and its IMG-ER genome ID Ga0427589. The type strain R-39594^T (= LMG 31950^T = CECT 30377^T) was isolated from a littoral epipsammic and interstitial

	The type strain R-36677^T (= LMG 31953^T = CECT 30380^T) was isolated from a black mat sample on gravel and rock debris from the bank of a frozen lake taken at Utsteinen, in the neighborhood of the Belgian Princess Elisabeth Station, Sør Rondane Mountains, East Antarctica.	The description of strain R-39604 is identical to that of the type strain, with the following differences: Growth is observed at 0-35 °C. Cells grow weakly on TSA and NA. Cells show variable hydrolysis of Tween 80. In the API ZYM gallery, no leucine arylamidase, α-chymotrypsin and acid phosphatase activities are present. According to the API 20 NE gallery, strain R-39604 is weakly positive for hydrolysis of gelatin. In the API 50 CH gallery, strain R-39604 is positive for acid formation from L-arabinose, D-fucose and potassium 5-ketogluconate, weakly positive for acid formation from amygdalin, salicin, D-cellobiose and D-raffinose and negative for acid formation from D-adonitol, methyl-βD-xylopyranoside and D-saccharose. Cells are susceptible to bacitracin. $C_{12:00}$ is not one of the major cellular fatty acids. The DNA G+C content of strain R-39604 is 68.89%, its approximate genome size 3.05 Mb and its IMG-ER genome ID Ga0401293. The type strain R-39161^T (= LMG 31951^T = CECT 30378^T) and strain R-39604 were isolated from a littoral epipsammic and interstitial microbial mat sample at 0.1 m sampling depth at Naka Tempyo in the Syowa Oasis region, East Antarctica.	microbial mat sample at 0.1 m sampling depth at Naka Tempyo in the Syowa Oasis region, East Antarctica.
Country of origin	Antarctica	Antarctica	Antarctica
Region of origin	Strain R-36677^T originates from Utsteinen, in the proximity of the Princess Elisabeth Station, East Antarctica.	Strains R-39604 and R-39161^T originate from Naka Tempyo in the Syowa Oasis, East Antarctica.	Strain R-39594^T originates from Naka Tempyo in the Syowa Oasis, East Antarctica.
Date of isolation (dd/mm/yyyy)	R-36677^T: 01/06/2007	R-39161^T: 01/03/2008 R-39604: 01/06/2008	R-39594^T: 01/06/2008
Source of isolation	R-36677^T: a black mat sample on gravel and rock debris from the bank of a frozen lake, BB115.	R-39161^T: a littoral epipsammic and interstitial microbial mat sample at 0.1 m sampling depth, SK5.	R-39594^T: a littoral epipsammic and interstitial microbial mat sample at 0.1 m sampling depth, SK5.
Sampling date (dd/mm/yyyy)	January 2007	January 2007	January 2007
Latitude (xx°xx′xx″N/S)	71° 57’ S	69° 28' S	69° 28' S
Longitude (xx°xx′xx″E/W)	23° 20’ E	39° 40' E	39° 40' E

Altitude (meters above sea level)	1318	288	288
16S rRNA gene accession nr.	R-36677^T: MN601351	R-39161^T: MN601352 R-39604: MN601354	R-39594^T: MN601353
Genome accession number [RefSeq; EMBL; …]	R-36677^T: IMG-ER Ga0401292	R-39161^T: IMG-ER Ga0427977 R-39604: IMG-ER Ga0401293	R-39594^T: IMG-ER Ga0427589
Genome status	R-36677^T: Incomplete	R-39161^T: Incomplete R-39604: Incomplete	R-39594^T: Incomplete
Genome size	R-36677^T: 3.76 Mb	R-39161^T: 3.04 Mb R-39604: 3.05 Mb	R-39594^T: 3.07 Mb
GC mol%	R-36677^T: 67.76	R-39161^T: 68.89 R-39604: 68.89	R-39594^T: 68.19
Number of strains in study	One	Two	One
Source of isolation of non-type strains		R-39604: a littoral epipsammic and interstitial microbial mat sample at 0.1 m sampling depth, SK5.	
Information related to the Nagoya Protocol	-	-	-
Designation of the Type Strain	R-36677^T	R-39161^T	R-39594^T
Strain Collection Numbers	LMG 31953^T = CECT 30380^T	LMG 31951^T = CECT 30378^T	LMG 31950^T = CECT 30377^T

853

854

References

[1] M. Achtman, M. Wagner (2008) Microbial diversity and the genetic nature of microbial species. Nat Rev Microbiol 6, 431-440.
[2] S. Akter, J. Du, K. Won, C.S. Yin, M. Kook, T.-H. Yi (2015) Sphingosinicella cucumeris sp. nov., isolated from soil of a cucumber garden. Antonie Van Leeuwenhoek 108, 1181-1188.
[3] M.J. Ankenbrand, A. Keller (2016) bcgTree: automatized phylogenetic tree building from bacterial core genomes. Genome 59, 783-791.
[4] D. Arndt, J.R. Grant, A. Marcu, T. Sajed, A. Pon, Y. Liang, D.S. Wishart (2016) PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16-W21.
[5] R. Aziz, D. Bartels, A. Best, M. DeJongh, T. Disz, R. Edwards, K. Formsma, S. Gerdes, E. Glass, M. Kubal, F. Meyer, G. Olsen, R. Olson, A. Osterman, R. Overbeek, L. McNeil, D. Paarmann, T. Paczian, B. Parrello, G. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko (2008) The RAST Server: Rapid Annotations using Subsystems Technology. BMC Genomics 9, 75.
[6] M.R. Badger, E.J. Bek (2008) Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO2 acquisition by the CBB cycle. J. Exp. Bot. 59, 1525-1541.
[7] T. Barbeyron, L. Brillet-Guéguen, W. Carré, C. Carrière, C. Caron, M. Czjzek, M. Hoebeke, G. Michel (2016) Matching the Diversity of Sulfated Biomolecules: Creation of a Classification Database for Sulfatases Reflecting Their Substrate Specificity. PLoS One 11, e0164846.
[8] J.T. Beatty (2002) On the natural selection and evolution of the aerobic phototrophic bacteria. Photosynthesis Res. 73, 109-114.
[9] O. Béjà, J.K. Lanyi (2014) Nature's toolkit for microbial rhodopsin ion pumps. Proc. Natl. Acad. Sci. U. S. A. 111, 6538-6539.
[10] J.-F. Bernardet, Y. Nakagawa, B. Holmes (2002) Proposed minimal standards for describing new taxa of the family Flavobacteriaceae and emended description of the family. Int. J. Syst. Evol. Microbiol. 52, 1049-1070.
[11] P.V. Bhaskar, N.B. Bhosle (2005) Microbial extracellular polymeric substances in marine biogeochemical processes. Curr. Sci. 88, 45-53.
[12] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S.Y. Lee, M.H. Medema, T. Weber (2019) antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 47, W81-W87.
[13] C.T. Brown, L.A. Hug, B.C. Thomas, I. Sharon, C.J. Castelle, A. Singh, M.J. Wilkins, K.C. Wrighton, K.H. Williams, J.F. Banfield (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523, 208-211.
[14] D.A. Bryant, N.-U. Frigaard (2006) Prokaryotic photosynthesis and phototrophy illuminated. Trends Microbiol. 14, 488-496.
[15] H.-J. Busse, E.B.M. Denner, S. Buczolits, M. Salkinoja-Salonen, A. Bennasar, P. Kämpfer (2003) *Sphingomonas aurantiaca* sp. nov., *Sphingomonas aerolata* sp. nov. and *Sphingomonas faeni* sp. nov., air- and dustborne and Antarctic, orange-pigmented, psychrotolerant bacteria, and emended description of the genus *Sphingomonas*. Int. J. Syst. Evol. Microbiol. 53, 1253-1260.
[16] H. Cai, H. Cui, Y. Zeng, M. An, H. Jiang (2018) Sandarakinorhabdus cyanobacteriorum sp. nov., a novel bacterium isolated from cyanobacterial aggregates in a eutrophic lake. Int. J. Syst. Evol. Microbiol. 68, 730-735.
[17] M. Cappelletti, D. Ghezzi, D. Zannoni, B. Capaccioni, S. Fedi (2016) Diversity of Methane-Oxidizing Bacteria in Soils from “Hot Lands of Medolla” (Italy) Featured by Anomalous High-Temperatures and Biogenic CO₂ Emission. Microbes Environ. 31, 369-377.
[18] S.C. Cary, I.R. McDonald, J.E. Barrett, D.A. Cowan (2010) On the rocks: the microbiology of Antarctic Dry Valley soils. Nat Rev Microbiol 8, 129-138.
[19] A. Chaya, N. Kurosawa, A. Kawamata, M. Kosugi, S. Imura (2019) Community Structures of Bacteria, Archaea, and Eukaryotic Microbes in the Freshwater Glacier Lake Yukidori-Ike in Langhovde, East Antarctica. Diversity 11, 105.

[20] C. Chen, Q. Zheng, Y.-N. Wang, X.-J. Yan, L.-K. Hao, X. Du, N. Jiao (2010) Stakelama pacifica gen. nov., sp. nov., a new member of the family Sphingomonadaceae isolated from the Pacific Ocean. Int. J. Syst. Evol. Microbiol. 60, 2857-2861.
[21] H. Chen, M. Jogler, M. Rohde, H.-P. Klenk, H.-J. Busse, B.J. Tindall, C. Spröer, J. Overmann (2013) Sphingobium limneticum sp. nov. and Sphingobium boeckii sp. nov., two freshwater planktonic members of the family Sphingomonadaceae, and reclassification of Sphingomonas suberifaciens as Sphingobium suberifaciens comb. nov. Int. J. Syst. Evol. Microbiol. 63, 735-743.
[22] V.P. Chhiba-Govindjee, C.W. van der Westhuyzen, M.L. Bode, D. Brady (2019) Bacterial nitrilases and their regulation. Appl. Microbiol. Biotechnol. 103, 4679-4692.
[23] J. Chun, A. Oren, A. Ventosa, H. Christensen, D.R. Arahal, M.S. da Costa, A.P. Rooney, H. Yi, X.-W. Xu, S. De Meyer, M.E. Trujillo (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68, 461-466.
[24] C. Cosseau, J. Batut (2004) Genomics of the ccoNOQP-encoded cbb3 oxidase complex in bacteria. Arch. Microbiol. 181, 89-96.
[25] D. Couvin, A. Bernheim, C. Toffano-Nioche, M. Touchon, J. Michalik, B. Néron, E.P. C. Rocha, G. Vergnaud, D. Gautheret, C. Pourcel (2018) CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. Nucleic Acids Res., gky425-gky425.
[26] B.M. Cumming, K.C. Chinta, V.P. Reddy, A.J.C. Steyn (2017) Role of Ergothioneine in Microbial Physiology and Pathogenesis. Antioxidants & Redox Signaling 28, 431-444.
[27] T.L. Dawson (2007) Light-harvesting and light-protecting pigments in simple life forms. Coloration Technology 123, 129-142.
[28] N. Delmotte, C. Knief, S. Chaffron, G. Innerebner, B. Roschitzki, R. Schlapbach, C. von Mering, J.A. Vorholt (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. U. S. A. 106, 16428-16433.
[29] D.A. Dik, D.R. Marous, J.F. Fisher, S. Mobashery (2017) Lytic transglycosylases: concinnity in concision of the bacterial cell wall. Crit. Rev. Biochem. Mol. Biol. 52, 503-542.
[30] J.M. Dubbs, F. Robert Tabita (2004) Regulators of nonsulfur purple phototrophic bacteria and the interactive control of CO2 assimilation, nitrogen fixation, hydrogen metabolism and energy generation. FEMS Microbiol. Rev. 28, 353-376.
[31] C. Dumolin, M. Aerts, B. Verheyde, S. Schellaert, T. Vandamme, F. Van der Jeugt, E. De Canck, M. Cnockaert, A.D. Wieme, I. Cleenwerck, J. Peiren, P. Dawyndt, P. Vandamme, A. Carlier (2019) Introducing SPeDE: High-Throughput Dereplication and Accurate Determination of Microbial Diversity from Matrix-Assisted Laser Desorption–Ionization Time of Flight Mass Spectrometry Data. mSystems 4, e00437-00419.
[32] W. Epstein 2003 The Roles and Regulation of Potassium in Bacteria in: Prog. Nucleic Acid Res. Mol. Biol., Academic Press, pp. 293-320.
[33] E. Fautz, H. Reichenbach (1980) A simple test for flexirubin-type pigments. FEMS Microbiol. Lett. 8, 87-91.
[34] G.-D. Feng, S.-Z. Yang, X. Xiong, H.-P. Li, H.-H. Zhu (2017) Sphingomonas spermidinifaciens sp. nov., a novel bacterium containing spermidine as the major polyamine, isolated from an abandoned lead–zinc mine and emended descriptions of the genus Sphingomonas and the species Sphingomonas yantingensis and Sphingomonas japonica. Int. J. Syst. Evol. Microbiol. 67, 2160-2165.
[35] N. Fong, M. Burgess, K. Barrow, D. Glenn (2001) Carotenoid accumulation in the psychrotrophic bacterium Arthrobacter agilis in response to thermal and salt stress. Appl. Microbiol. Biotechnol. 56, 750-756.
[36] W. Fukuda, Y. Chino, S. Araki, Y. Kondo, H. Imanaka, T. Kanai, H. Atomi, T. Imanaka (2014) Polymorphobacter multimanifer gen. nov., sp. nov., a polymorphic bacterium isolated from antarctic white rock. Int. J. Syst. Evol. Microbiol. 64, 2034-2040.
[37] F. Gich, J. Overmann (2006) Sandarakinorhabdus limnophila gen. nov., sp. nov., a novel bacteriochlorophyll a-containing, obligately aerobic bacterium isolated from freshwater lakes. Int. J. Syst. Evol. Microbiol. 56, 847-854.

[38] S.J. Giovannoni, J. Cameron Thrash, B. Temperton (2014) Implications of streamlining theory for microbial ecology. ISME J. 8, 1553-1565.
[39] M. Gomila, J. Gascó, A. Busquets, J. Gil, R. Bernabeu, J.M. Buades, J. Lalucat (2005) Identification of culturable bacteria present in haemodialysis water and fluid. FEMS Microbiol. Ecol. 52, 101-114.
[40] A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler (2013) QUAST: quality assessment tool for genome assemblies. Bioinformatics 29, 1072-1075.
[41] A. Hiraishi, H. Kuraishi, K. Kawahara (2000) Emendation of the description of *Blastomonas natatoria* (Sly 1985) Sly and Cahill 1997 as an aerobic photosynthetic bacterium and reclassification of *Erythromonas ursincola* Yurkov et al. 1997 as *Blastomonas ursincola* comb. nov. Int. J. Syst. Evol. Microbiol. 50, 1113-1118.
[42] J.F. Imhoff, T. Rahn, S. Künzel, S.C. Neulinger (2018) Photosynthesis Is Widely Distributed among Proteobacteria as Demonstrated by the Phylogeny of PufLM Reaction Center Proteins. Front. Microbiol. 8.
[43] L. Jia, X. Feng, Z. Zheng, L. Han, X. Hou, Z. Lu, J. Lv (2015) Polymorphobacter fuscus sp. nov., isolated from permafrost soil, and emended description of the genus Polymorphobacter. Int. J. Syst. Evol. Microbiol. 65, 3920-3925.
[44] R. Jiang, J.-X. Wang, B. Huang, K.-C. Yu, P. Zhang, J.-W. Zheng, X.-Z. Liu (2016) Phylogenetic analysis of bacterial community composition in sediments with organic contaminants from the Jiaojiang estuary in China. Mar. Pollut. Bull. 109, 558-565.
[45] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T.L. Madden (2008) NCBI BLAST: a better web interface. Nucleic Acids Res. 36, W5-W9.
[46] J. Jovel, J. Patterson, W. Wang, N. Hotte, S. O'Keefe, T. Mitchel, T. Perry, D. Kao, A.L. Mason, K.L. Madsen, G.K.S. Wong (2016) Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. Front. Microbiol. 7, 459.
[47] S. Kato, M. Ohkuma, A. Yamagishi 2015 Intra-Field Variation of Prokaryotic Communities On and Below the Seafloor in the Back-Arc Hydrothermal System of the Southern Mariana Trough in: J.-i. Ishibashi, K. Okino, M. Sunamura (Eds.) Subseafloor Biosphere Linked to Hydrothermal Systems: TAIGA Concept, Springer Japan, Tokyo, pp. 301-311.
[48] K. Katoh, M.C. Frith (2012) Adding unaligned sequences into an existing alignment using MAFFT and LAST. Bioinformatics 28, 3144-3146.
[49] A.D. Kennedy (1993) Water as a limiting factor in the Antarctic terrestrial environment - a biogeographical synthesis. Arct. Alp. Res. 25, 308-315.
[50] M. Kim, O. Kang, Y. Zhang, L. Ren, X. Chang, F. Jiang, C. Fang, C. Zheng, F. Peng (2016) Sphingoaurantiacus polygranulatus gen. nov., sp. nov., isolated from high-Arctic tundra soil, and emended descriptions of the genera Sandarakinorhabdus, Polymorphobacter and Rhizorhabdus and the species Sandarakinorhabdus limnophila, Rhizorhabdus argentea and Sphingomonas wittichii. Int. J. Syst. Evol. Microbiol. 66, 91-100.
[51] M. Kim, H.-S. Oh, S.-C. Park, J. Chun (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int. J. Syst. Evol. Microbiol. 64, 346-351.
[52] M. Koblížek (2015) Ecology of aerobic anoxygenic phototrophs in aquatic environments. FEMS Microbiol. Rev. 39, 854-870.
[53] E.Y. Koh, W. Phua, K.G. Ryan (2011) Aerobic anoxygenic phototrophic bacteria in Antarctic sea ice and seawater. Environ. Microbiol. Rep. 3, 710-716.
[54] Z.S. Kolber, C.L. Van Dover, R.A. Niederman, P.G. Falkowski (2000) Bacterial photosynthesis in surface waters of the open ocean. Nature 407, 177-179.
[55] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol. Evol. 35, 1547-1549.
[56] M. Labrenz, M.D. Collins, P.A. Lawson, B.J. Tindall, P. Schumann, P. Hirsch (1999) Roseovarius tolerans gen. nov., sp. nov., a budding bacterium with variable bacteriochlorophyll a production from hypersaline Ekho Lake. Int. J. Syst. Evol. Microbiol. 49, 137-147.

[57] M. Labrenz, B.J. Tindall, P.A. Lawson, M.D. Collins, P. Schumann, P. Hirsch (2000) Staleya guttiformis gen. nov., sp. nov. and Sulfitobacter brevis sp. nov., alpha-3-Proteobacteria from hypersaline, heliothermal and meromictic antarctic Ekho Lake. Int. J. Syst. Evol. Microbiol. 50, 303-313.
[58] K. Lagesen, P. Hallin, E.A. Rodland, H.H. Staerfeldt, T. Rognes, D.W. Ussery (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35.
[59] I. Lagkouvardos, D. Joseph, M. Kapfhammer, S. Giritli, M. Horn, D. Haller, T. Clavel (2016) IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. Sci. Rep. 6, 33721.
[60] I. Letunic, P. Bork (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47, W256-W259.
[61] J. Li, C. Bi, J. Tu, Z. Lu (2018) The complete mitochondrial genome sequence of Boechera stricta. Mitochondrial DNA B Resour 3, 896-897.
[62] S. Li, W. Zhou, D. Lin, K. Tang, N. Jiao (2016) Pacificimonas aurantium sp. nov., Isolated from the Seawater of the Pacific Ocean. Curr. Microbiol. 72, 752-757.
[63] K. Liu, S. Li, N. Jiao, K. Tang (2014) Pacificamonas flava gen. nov., sp. nov., a Novel Member of the Family Sphingomonadaceae Isolated from the Southeastern Pacific. Curr. Microbiol. 69, 96-101.
[64] V. Lombard, H. Golaconda Ramulu, E. Drula, P.M. Coutinho, B. Henrissat (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42, D490-D495.
[65] S. Lutz, L.A. Ziolkowski, L.G. Benning (2019) The Biodiversity and Geochemistry of Cryoconite Holes in Queen Maud Land, East Antarctica. Microorganisms 7, 160.
[66] J.F. MacFaddin 1980 Biochemical tests for identification of medical bacteria, 2nd ed. ed., Williams & Wilkins Co, Baltimore (Md.).
[67] M.T. Madigan, D.O. Jung, C.R. Woese, L.A. Achenbach (2000) *Rhodoferax antarcticus* sp. nov., a moderately psychrophilic purple nonsulfur bacterium isolated from an Antarctic microbial mat. Arch. Microbiol. 173, 269-277.
[68] B. Magasanik (1989) Regulation of transcription of the glnALG operon of Escherichia coli by protein phosphorylation. Biochimie 71, 1005-1012.
[69] S. Mahapatra, D. Banerjee (2013) Fungal exopolysaccharide: production, composition and applications. Microbiology insights 6, 1-16.
[70] V.M. Markowitz, I.M.A. Chen, K. Palaniappan, K. Chu, E. Szeto, Y. Grechkin, A. Ratner, B. Jacob, J. Huang, P. Williams, M. Huntemann, I. Anderson, K. Mavromatis, N.N. Ivanova, N.C. Kyrpides (2012) IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res. 40, D115-D122.
[71] V.M. Markowitz, K. Mavromatis, N.N. Ivanova, I.-M.A. Chen, K. Chu, N.C. Kyrpides (2009) IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics 25, 2271-2278.
[72] L.A. Marraffini, E.J. Sontheimer (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat. Rev. Genet. 11, 181-190.
[73] G.J. Marshall (2009) On the annual and semi-annual cycles of precipitation across Antarctica. IJCli 29, 2298-2308.
[74] J.P. Meier-Kolthoff, A.F. Auch, H.-P. Klenk, M. Göker (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14, 60.
[75] K. Mewis, N. Lenfant, V. Lombard, B. Henrissat (2016) Dividing the Large Glycoside Hydrolase Family 43 into Subfamilies: a Motivation for Detailed Enzyme Characterization. Appl. Environ. Microbiol. 82, 1686-1692.
[76] V.K. Michelou, J.G. Caporaso, R. Knight, S.R. Palumbi (2013) The Ecology of Microbial Communities Associated with Macrocystis pyrifera. PLoS One 8, e67480.
[77] B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R. Lanfear (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. 37, 1530-1534.
[78] L. Mirón, A. Mira, V. Rocha-Ramírez, P. Belda-Ferre, R.U. Cabrera-Rubio, J. Folch-Mallol, R. Cárdenas-Vázquez, A. DeLuna, A. Hernández, E. Maya-Elizarrarás, J. Schondube 2014 Gut Bacterial Diversity of the House Sparrow (Passer domesticus) Inferred by 16S rRNA Sequence Analysis, in.

[79] A.L. Mitchell, T.K. Attwood, P.C. Babbitt, M. Blum, P. Bork, A. Bridge, S.D. Brown, H.-Y. Chang, S. El-Gebali, M.I. Fraser, J. Gough, D.R. Haft, H. Huang, I. Letunic, R. Lopez, A. Luciani, F. Madeira, A. Marchler-Bauer, H. Mi, D.A. Natale, M. Necci, G. Nuka, C. Orengo, A.P. Pandurangan, T. Paysan-Lafosse, S. Pesseat, S.C. Potter, M.A. Qureshi, N.D. Rawlings, N. Redaschi, L.J. Richardson, C. Rivoire, G.A. Salazar, A. Sangrador-Vegas, C.J A. Sigrist, I. Sillitoe, G.G. Sutton, N. Thanki, P.D. Thomas, S.C E. Tosatto, S.-Y. Yong, R.D. Finn (2018) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. 47, D351-D360.
[80] S. Niemann, A. Puhler, H.V. Tichy, R. Simon, W. Selbitschka (1997) Evaluation of the resolving power of three different DNA fingerprinting methods to discriminate among isolates of a natural Rhizobium meliloti population. J. Appl. Microbiol. 82, 477-484.
[81] D. Obbels, E. Verleyen, M.-J. Mano, Z. Namsaraev, M. Sweetlove, B. Tytgat, R. Fernandez-Carazo, A. De Wever, S. D'hondt, D. Ertz, J. Elster, K. Sabbe, A. Willems, A. Wilmotte, W. Vyverman (2016) Bacterial and eukaryotic biodiversity patterns in terrestrial and aquatic habitats in the Sør Rondane Mountains, Dronning Maud Land, East Antarctica. FEMS Microbiol. Ecol. 92, fiw041.
[82] L. Pal, B. Kraigher, B. Brajer-Humar, M. Levstek, I. Mandic-Mulec (2012) Total bacterial and ammonia-oxidizer community structure in moving bed biofilm reactors treating municipal wastewater and inorganic synthetic wastewater. Bioresour. Technol. 110, 135-143.
[83] R. Pal, S. Bala, M. Dadhwal, M. Kumar, G. Dhingra, O. Prakash, S.R. Prabagaran, S. Shivaji, J. Cullum, C. Holliger, R. Lal (2005) Hexachlorocyclohexane-degrading bacterial strains Sphingomonas paucimobilis B90A, UT26 and Sp+, having similar lin genes, represent three distinct species, Sphingobium indicum sp. nov., Sphingobium japonicum sp. nov. and Sphingobium francense sp. nov., and reclassification of [Sphingomonas] chungbukensis as Sphingobium chungbukense comb. nov. Int. J. Syst. Evol. Microbiol. 55, 1965-1972.
[84] D.H. Parks, M. Imelfort, C.T. Skennerton, P. Hugenholtz, G.W. Tyson (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25, 1043-1055.
[85] K. Peeters, D. Ertz, A. Willems (2011) Culturable bacterial diversity at the Princess Elisabeth Station (Utsteinen, Sør Rondane Mountains, East Antarctica) harbours many new taxa. Syst. Appl. Microbiol. 34, 360-367.
[86] K. Peeters, E. Verleyen, D. Hodgson, P. Convey, D. Ertz, W. Vyverman, A. Willems (2012) Heterotrophic bacterial diversity in aquatic microbial mat communities from Antarctica. Polar Biol. 35, 543-554.
[87] M. Peimbert, L.D. Alcaraz, G. Bonilla-Rosso, G. Olmedo-Alvarez, F. García-Oliva, L. Segovia, L.E. Eguiarte, V. Souza (2012) Comparative Metagenomics of Two Microbial Mats at Cuatro Ciénegas Basin I: Ancient Lessons on How to Cope with an Environment Under Severe Nutrient Stress. Astrobiology 12, 648-658.
[88] D. Phurbu, Z.-X. Liu, H.-C. Liu, Y. Lhamo, P. Yangzom, A.-H. Li, Y.-G. Zhou (2020) Polymorphobacter arshaanensis sp. nov., containing the photosynthetic gene pufML, isolated from a volcanic lake. Int. J. Syst. Evol. Microbiol. 70, 1093-1098.
[89] M. Pietiläinen, H. Korpelainen (2013) Population genetics of purple saxifrage (Saxifraga oppositifolia) in the high Arctic archipelago of Svalbard. AoB PLANTS 5.
[90] J.A.G. Ranea, A. Grant, J.M. Thornton, C.A. Orengo (2005) Microeconomic principles explain an optimal genome size in bacteria. Trends Genet. 21, 21-25.
[91] N.D. Rawlings, M. Waller, A.J. Barrett, A. Bateman (2013) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 42, D503-D509.
[92] M.C.P.P. Reis-Mansur, J.S. Cardoso-Rurr, J.V.M.A. Silva, G.R. de Souza, V.d.S. Cardoso, F.R.P. Mansoldo, Y. Pinheiro, J. Schultz, L.B. Lopez Balottin, A.J.R. da Silva, C. Lage, E.P. dos Santos, A.S. Rosado, A.B. Vermelho (2019) Carotenoids from UV-resistant Antarctic Microbacterium sp. LEMMJ01. Sci. Rep. 9, 9554.
[93] J. Reitner 2011 Microbial Mats in: J. Reitner, V. Thiel (Eds.) Encyclopedia of Geobiology, Springer Netherlands, Dordrecht, pp. 606-608.

[94] L. Ren, X. Chang, F. Jiang, W. Kan, Z. Qu, X. Qiu, C. Fang, F. Peng (2015) Parablastomonas arctica gen. nov., sp. nov., isolated from high Arctic glacial till. Int. J. Syst. Evol. Microbiol. 65, 260-266.
[95] M. Richter, R. Rosselló-Móra (2009) Shifting the genomic gold standard for the prokaryotic species definition. PNAS 106, 19126-19131.
[96] R. Rivas, A. Abril, M.E. Trujillo, E. Velázquez (2004) Sphingomonas phyllosphaerae sp. nov., from the phyllosphere of Acacia caven in Argentina. Int. J. Syst. Evol. Microbiol. 54, 2147-2150.
[97] F. Santos-Beneit (2015) The Pho regulon: a huge regulatory network in bacteria. Front. Microbiol. 6, 402.
[98] S. Srinivasan, E.S. Joo, J.-J. Lee, M.K. Kim (2015) *Hymenobacter humi* sp. nov., a bacterium isolated from soil. Antonie Van Leeuwenhoek 107, 1411-1419.
[99] A.R. Stampfli, K.V. Goncharenko, M. Meury, B.N. Dubey, T. Schirmer, F.P. Seebeck (2019) An Alternative Active Site Architecture for O2 Activation in the Ergothioneine Biosynthetic EgtB from Chloracidobacterium thermophilum. J. Am. Chem. Soc. 141, 5275-5285.
[100] T. Suyama, T. Shigematsu, T. Suzuki, Y. Tokiwa, T. Kanagawa, K.V.P. Nagashima, S. Hanada (2002) Photosynthetic Apparatus in Roseateles depolymerans 61A Is Transcriptionally Induced by Carbon Limitation. Appl. Environ. Microbiol. 68, 1665-1673.
[101] G. Tahon, L. Lebbe, M. De Troch, K. Sabbe, A. Willems (2020) Leeuwenhoekiella aestuarii sp. nov., isolated from salt-water sediment and first insights in the genomes of Leeuwenhoekiella species. Int. J. Syst. Evol. Microbiol. 70, 1706-1719.
[102] G. Tahon, B. Tytgat, L. Lebbe, A. Carlier, A. Willems (2018) *Abditibacterium utsteinense* sp. nov., the first cultivated member of candidate phylum FBP, isolated from ice-free Antarctic soil samples. Syst. Appl. Microbiol. 41, 279-290.
[103] G. Tahon, B. Tytgat, P. Stragier, A. Willems (2016) Analysis of *cbbL*, *nifH*, and *pufLM* in Soils from the Sør Rondane Mountains, Antarctica, Reveals a Large Diversity of Autotrophic and Phototrophic Bacteria. Microb. Ecol. 71, 131-149.
[104] G. Tahon, B. Tytgat, A. Willems (2018) Diversity of key genes for carbon and nitrogen fixation in soils from the Sør Rondane Mountains, East Antarctica. Polar Biol. 41, 2181-2198.
[105] G. Tahon, B. Tytgat, A. Willems (2016) Diversity of Phototrophic Genes Suggests Multiple Bacteria May Be Able to Exploit Sunlight in Exposed Soils from the Sør Rondane Mountains, East Antarctica. Front. Microbiol. 7.
[106] G. Tahon, A. Willems (2017) Isolation and characterization of aerobic anoxygenic phototrophs from exposed soils from the Sør Rondane Mountains, East Antarctica. Syst. Appl. Microbiol. 40, 357-369.
[107] T. Tsukamoto, M. Demura, Y. Sudo (2014) Irreversible Trimer to Monomer Transition of Thermophilic Rhodopsin upon Thermal Stimulation. J. Phys. Chem. B 118, 12383-12394.
[108] B. Tytgat, E. Verleyen, M. Sweetlove, S. D'hondt, P. Clercx, E. Van Ranst, K. Peeters, S. Roberts, Z. Namsaraev, A. Wilmotte, W. Vyverman, A. Willems (2016) Bacterial community composition in relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica. FEMS Microbiol. Ecol. 92, fiw126.
[109] L.G. Wayne, D.J. Brenner, R.R. Colwell, P.A.D. Grimont, O. Kandler, M.I. Krichevsky, L.H. Moore, W.E.C. Moore, R.G.E. Murray, E. Stackebrandt, M.P. Starr, H.G. Truper (1987) Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Int. J. Syst. Evol. Microbiol. 37, 463-464.
[110] D.D. Wynn-Williams (1990) Microbial Colonization Processes In Antarctic Fellfield Soils : An Experimental Overview (Eleventh Symposium on Polar Biology). Proc. NIPR Symp. Polar Biol. 3, 164-178.
[111] P. Yarza, P. Yilmaz, E. Pruesse, F.O. Glockner, W. Ludwig, K.-H. Schleifer, W.B. Whitman, J. Euzeby, R. Amann, R. Rossello-Mora (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Micro 12, 635-645.
[112] J.-H. Yoon, S.-J. Kang, J.-S. Lee, S.-W. Nam, W. Kim, T.-K. Oh (2008) Sphingosinicella soli sp. nov., isolated from an alkaline soil in Korea. Int. J. Syst. Evol. Microbiol. 58, 173-177.

[113] S.-H. Yoon, S.-M. Ha, S. Kwon, J. Lim, Y. Kim, H. Seo, J. Chun (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int. J. Syst. Evol. Microbiol. 67, 1613-1617.
[114] S.-H. Yoon, S.-m. Ha, J. Lim, S. Kwon, J. Chun (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110, 1281-1286.
[115] V. Yurkov, J.T. Csotonyi 2009 New Light on Aerobic Anoxygenic Phototrophs in: C.N. Hunter, F. Daldal, M.C. Thurnauer, J.T. Beatty (Eds.) The Purple Phototrophic Bacteria, Springer Netherlands, Dordrecht, pp. 31-55.
[116] V. Yurkov, E. Hughes 2017 Aerobic Anoxygenic Phototrophs: Four Decades of Mystery in: P.C. Hallenbeck (Ed.) Modern Topics in the Phototrophic Prokaryotes: Environmental and Applied Aspects, Springer International Publishing, Cham, pp. 193-214.
[117] V. Yurkov, E. Stackebrandt, O. Buss, A. Vermeglio, V. Gorlenko, J.T. Beatty (1997) Reorganization of the Genus Erythromicrobium: Description of “Erythromicrobium sibiricum” as Sandaracinobacter sibiricus gen. nov., sp. nov., and of “Erythromicrobium ursincola” as Erythromonas ursincola gen. nov., sp. nov. Int. J. Syst. Evol. Microbiol. 47, 1172-1178.
[118] H. Zhang, T. Yohe, L. Huang, S. Entwistle, P. Wu, Z. Yang, P.K. Busk, Y. Xu, Y. Yin (2018) dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 46, W95-W101.
[119] J. Zhang, Z.-F. Lang, J.-W. Zheng, B.-J. Hang, X.-Q. Duan, J. He, S.-P. Li (2012) Sphingobium jiangsuense sp. nov., a 3-phenoxybenzoic acid-degrading bacterium isolated from a wastewater treatment system. Int. J. Syst. Evol. Microbiol. 62, 800-805.

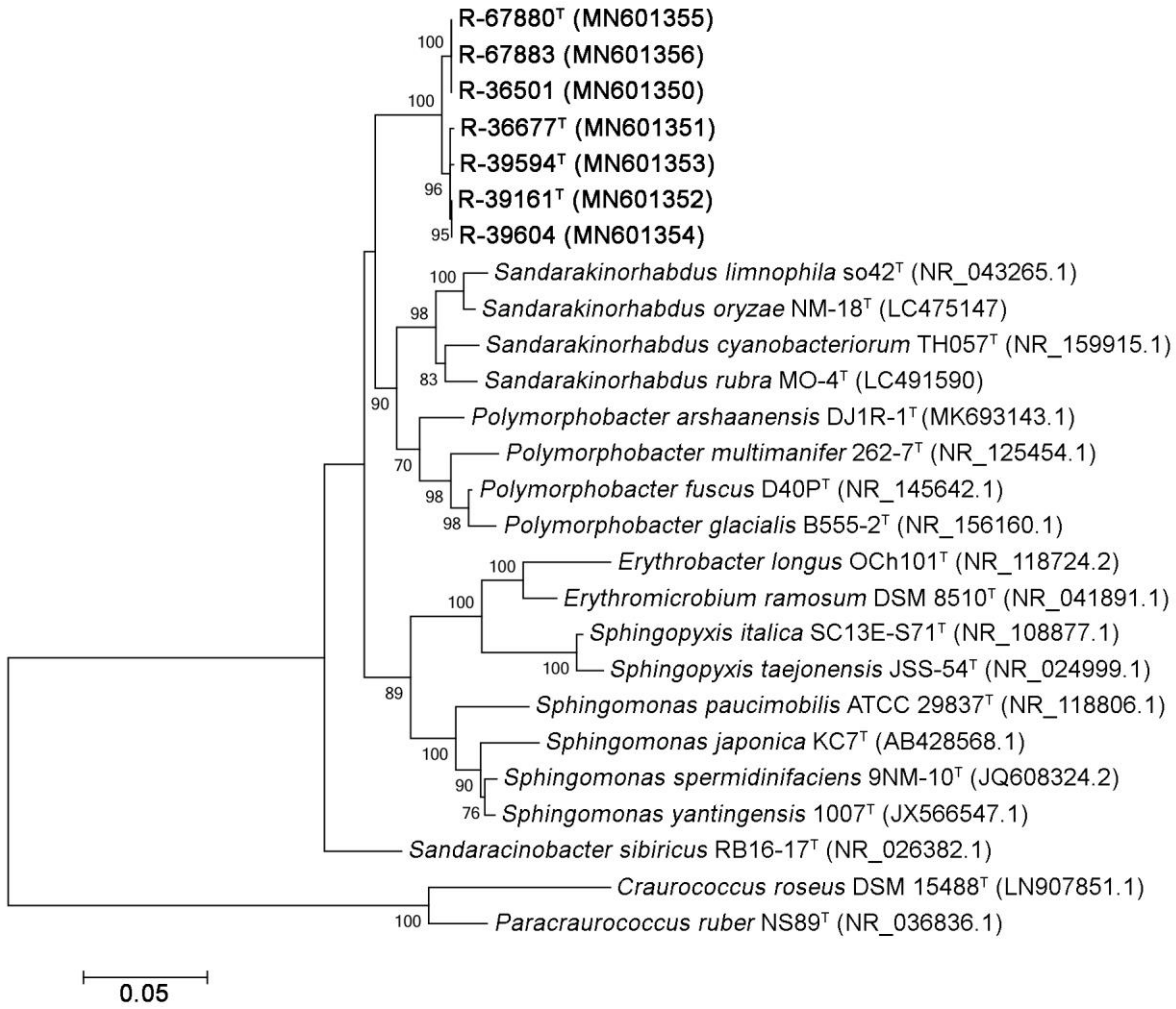

Figure 1. Maximum-likelihood phylogenetic tree (1000 bootstrap replicates) of near-complete 16S rRNA gene sequences showing the relationship between strains R-67880^{T}, R-67883, R-36501, R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T} and closely related type strains of the family *Sphingomonadaceae*. *Paracraurococcus ruber* NS89^{T} and *Craurococcus roseus* DSM 15488^{T} were used as an outgroup. Only bootstrap values greater than 70% are shown. Scale bar indicates 0.05 substitutions per nucleotide position.

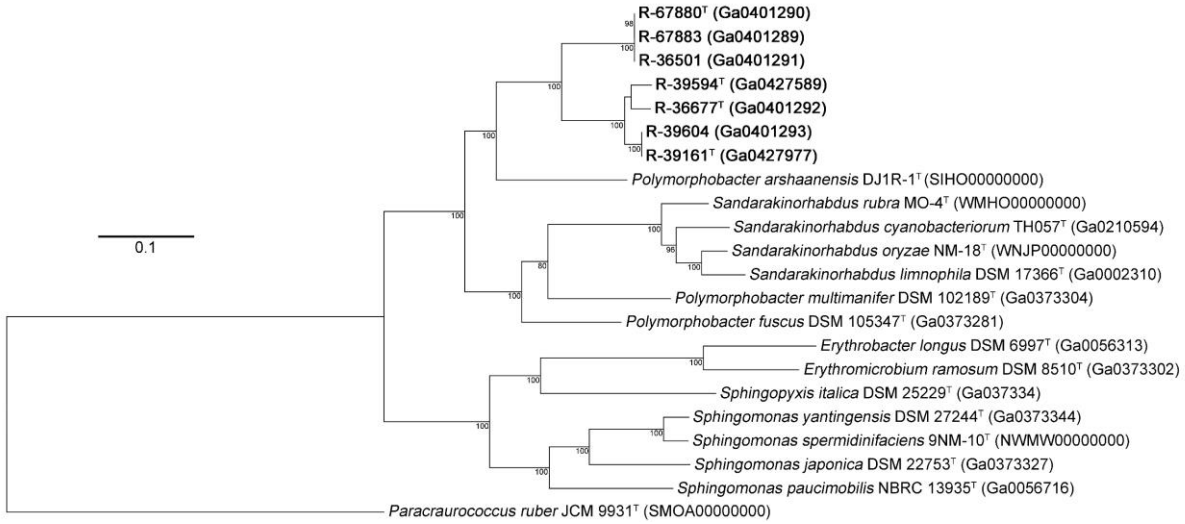

Figure 2. Maximum-likelihood phylogenetic tree (1000 bootstraps) based on 107 single-copy core genes of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and closely related type strains of the families *Sphingomonadaceae* and Erythrobacteraceae. Only bootstrap values higher than 70% are shown. The scale bar indicates 0.1 substitutions per site. *Paracraurococcus ruber* JCM 9931^T was used as an outgroup.

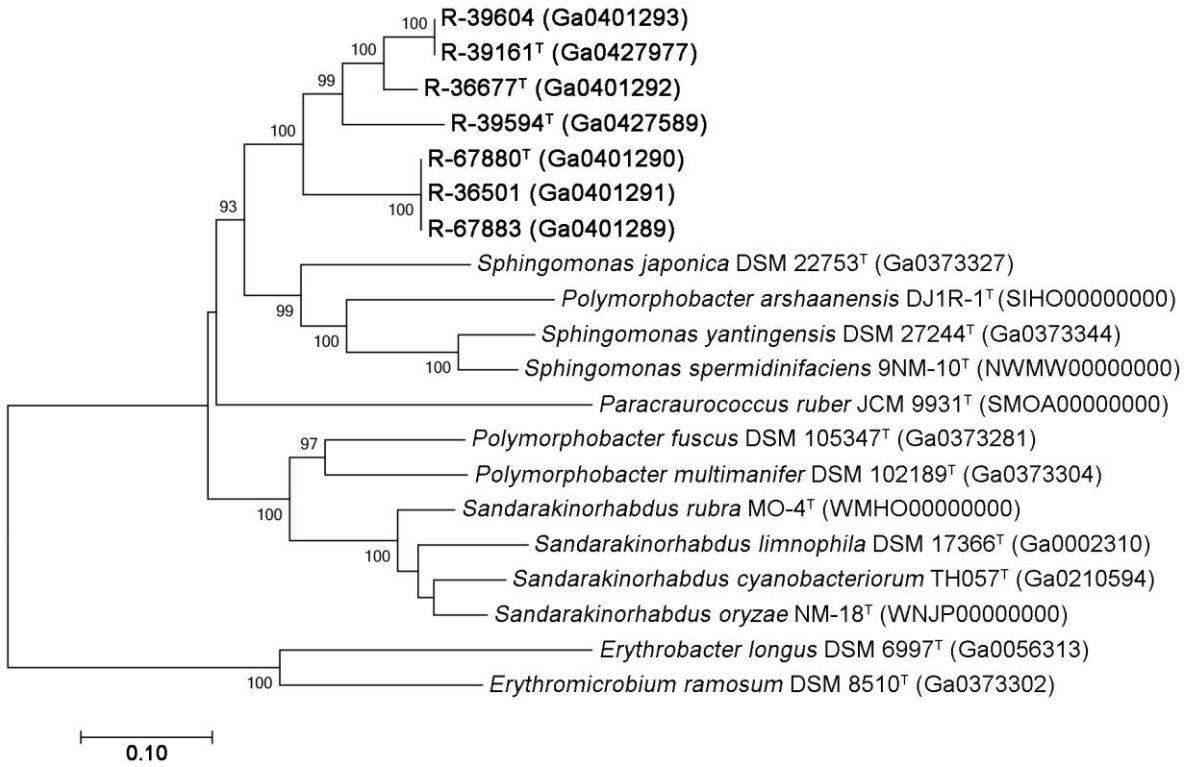

Figure 3. Maximum-likelihood phylogenetic tree (1000 bootstraps) of the *pufLM* gene sequences of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and closely related type strains of the families *Sphingomonadaceae* and Acetobacteraceae. Only bootstrap values higher than 70% are shown. The scale bar indicates 0.10 substitutions per site. *Erythrobacter longus* DSM 6997^T and *Erythromicrobium ramosum* DSM 8510^T were used as an outgroup.

Supplementary Material

Description and functional testing of four species of the novel phototrophic genus *Chioneia* gen. nov., isolated from different East Antarctic environments

Guillaume Tahon[1,2], Duygu Gök[1], Liesbeth Lebbe[1] and Anne Willems[1]*

[1] Laboratory of Microbiology, Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium

[2] Current address: Laboratory of Microbiology, Agrotechnology and Food Sciences, Wageningen University, Wageningen, The Netherlands

*** Correspondence:** Anne Willems: Anne.Willems@UGent.be

1. Supplementary Figures and Tables

1.1. Supplementary Figures

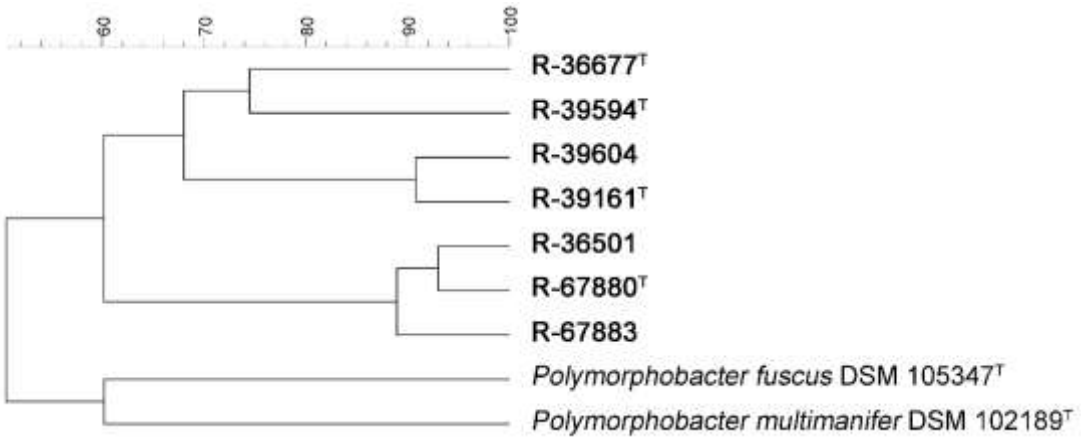

Fig. S1. Curve-based cluster analysis of mass spectra obtained from protein extracts using the Pearson product moment correlation coefficient and UPGMA cluster algorithm of strains R-67883, R-67880^{T}, R-36501, R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T}, and closely related strains *Polymorphobacter multimanifer* DSM 102189^{T} and *Polymorphobacter fuscus* DSM 105347^{T}.

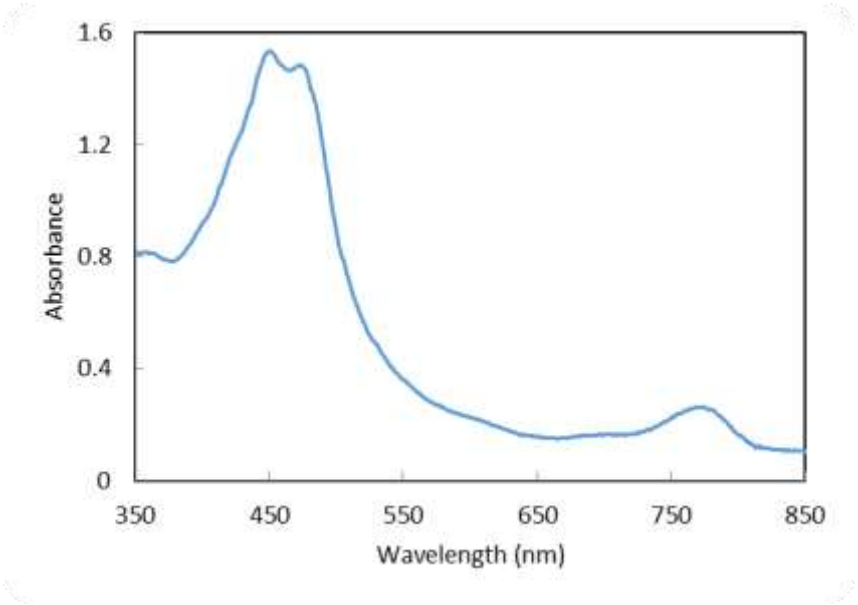

Fig. S2. Absorption spectrum of the pigment extract of strain R-39594^{T} in methanol.

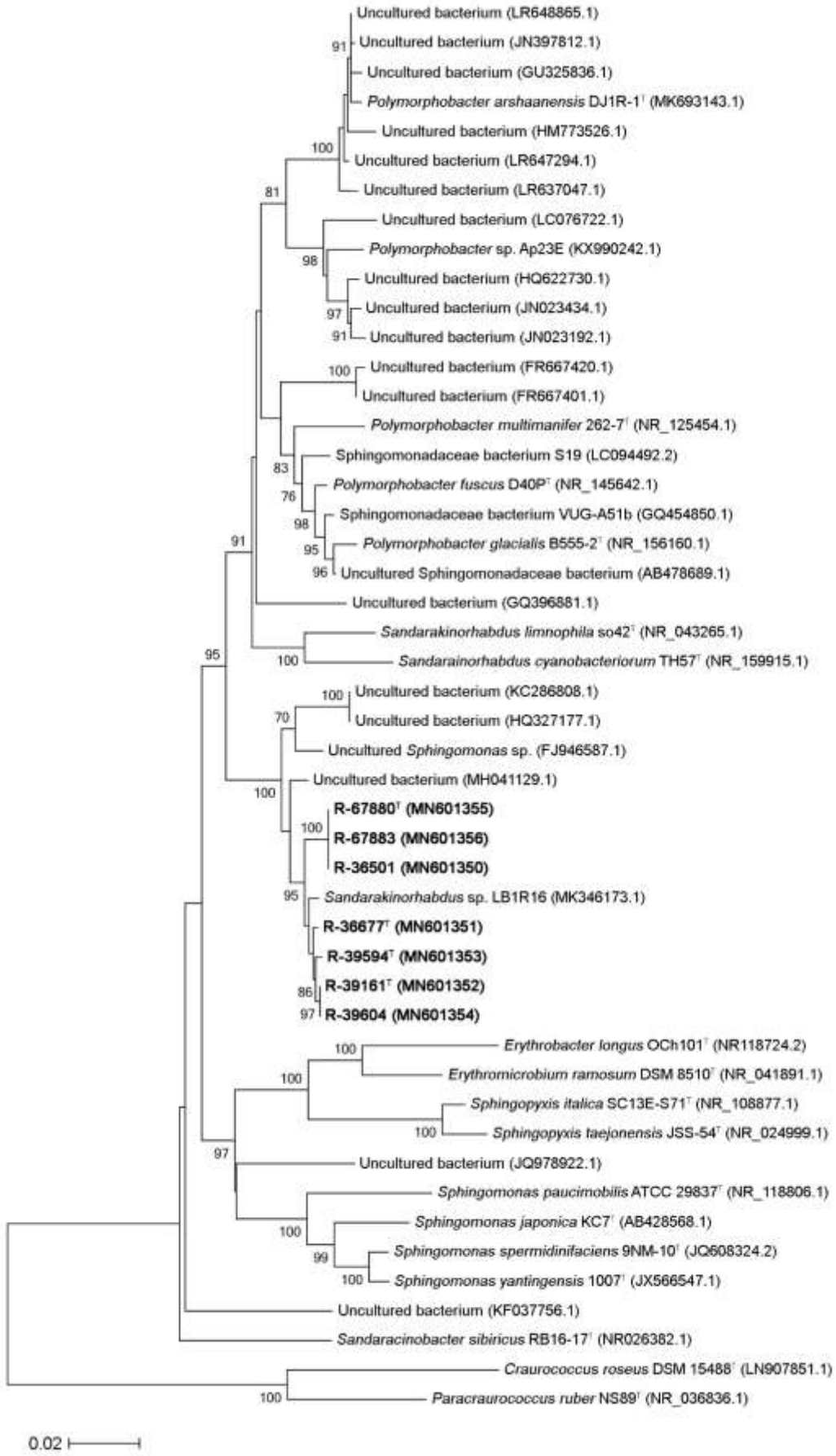
Uncultured bacterium (LR648865.1)
Uncultured bacterium (JN397812.1)
Uncultured bacterium (GU325836.1)
Polymorphobacter arshaanensis DJ1R-1T (MK693143.1)
Uncultured bacterium (HM773526.1)
Uncultured bacterium (LR647294.1)
Uncultured bacterium (LR637047.1)
Uncultured bacterium (LC076722.1)
Polymorphobacter sp. Ap23E (KX990242.1)
Uncultured bacterium (HQ622730.1)
Uncultured bacterium (JN023434.1)
Uncultured bacterium (JN023192.1)
Uncultured bacterium (FR667420.1)
Uncultured bacterium (FR667401.1)
Polymorphobacter multimanifer 262-7T (NR_125454.1)
Sphingomonadaceae bacterium S19 (LC094492.2)
Polymorphobacter fuscus D40PT (NR_145642.1)
Sphingomonadaceae bacterium VUG-A51b (GQ454850.1)
Polymorphobacter glacialis B555-2T (NR_156160.1)
Uncultured Sphingomonadaceae bacterium (AB478689.1)
Uncultured bacterium (GQ396881.1)
Sandarakinorhabdus limnophila so42T (NR_043265.1)
Sandarainorhabdus cyanobacteriorum TH57T (NR_159915.1)
Uncultured bacterium (KC286808.1)
Uncultured bacterium (HQ327177.1)
Uncultured Sphingomonas sp. (FJ946587.1)
Uncultured bacterium (MH041129.1)
R-67880T (MN601355)
R-67883 (MN601356)
R-36501 (MN601350)
Sandarakinorhabdus sp. LB1R16 (MK346173.1)
R-36677T (MN601351)
R-39594T (MN601353)
R-39161T (MN601352)
R-39604 (MN601354)
Erythrobacter longus OCh101T (NR118724.2)
Erythromicrobium ramosum DSM 8510T (NR_041891.1)
Sphingopyxis italica SC13E-S71T (NR_108877.1)
Sphingopyxis taejonensis JSS-54T (NR_024999.1)
Uncultured bacterium (JQ978922.1)
Sphingomonas paucimobilis ATCC 29837T (NR_118806.1)
Sphingomonas japonica KC7T (AB428568.1)
Sphingomonas spermidinifaciens 9NM-10T (JQ608324.2)
Sphingomonas yantingensis 1007T (JX566547.1)
Uncultured bacterium (KF037756.1)
Sandaracinobacter sibiricus RB16-17T (NR026382.1)
Craurococcus roseus DSM 15488T (LN907851.1)
Paracraurococcus ruber NS89T (NR_036836.1)
0.02

Fig. S3. Maximum-likelihood phylogenetic tree (1000 bootstrap replicates) of near-complete 16S rRNA gene sequences showing the relationship between strains R-67880^{T}, R-67883, R-36501, R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T} and closely related type strains and environmental sequences of the families *Sphingomonadaceae* and *Erythrobacteraceae*. *Paracraurococcus ruber* NS89^{T} and *Craurococcus roseus* DSM 15488^{T} were used as an outgroup. Only bootstrap values greater than 70% are shown. Scale bar indicates 0.02 substitutions per nucleotide position.

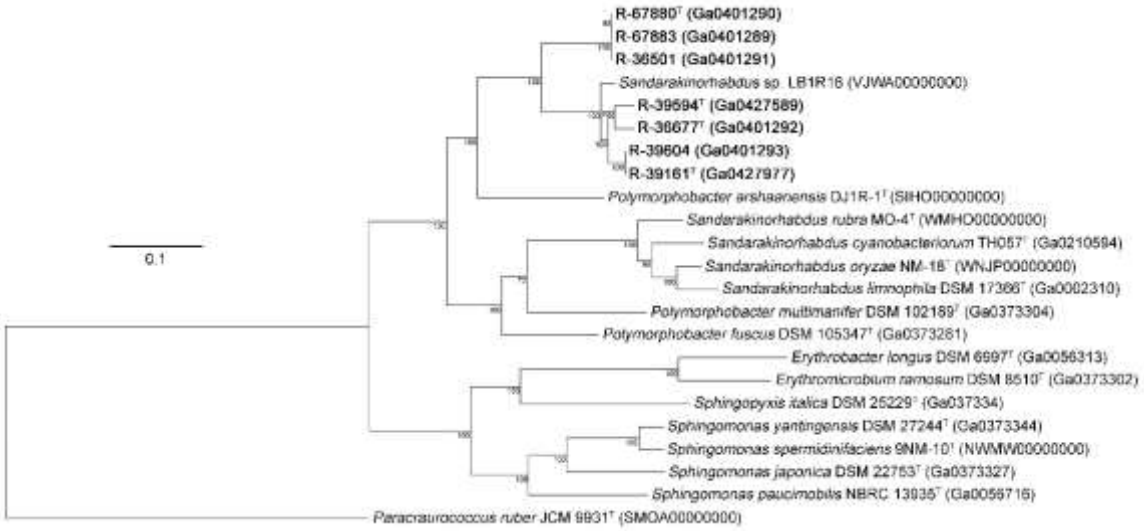

Fig. S4. Maximum-likelihood phylogenetic tree (1000 bootstraps) based on 107 single-copy core genes of strains R-67883, R-67880^{T}, R-36501, R-36677^{T}, R-39604, R-39161^{T} and R-39594^{T} and closely related type strains of the families *Sphingomonadaceae* and *Erythrobacteraceae*. Only bootstrap values higher than 70% are shown. The scale bar indicates 0.1 substitutions per site. *Paracraurococcus ruber* JCM 9931^{T} was used as an outgroup.

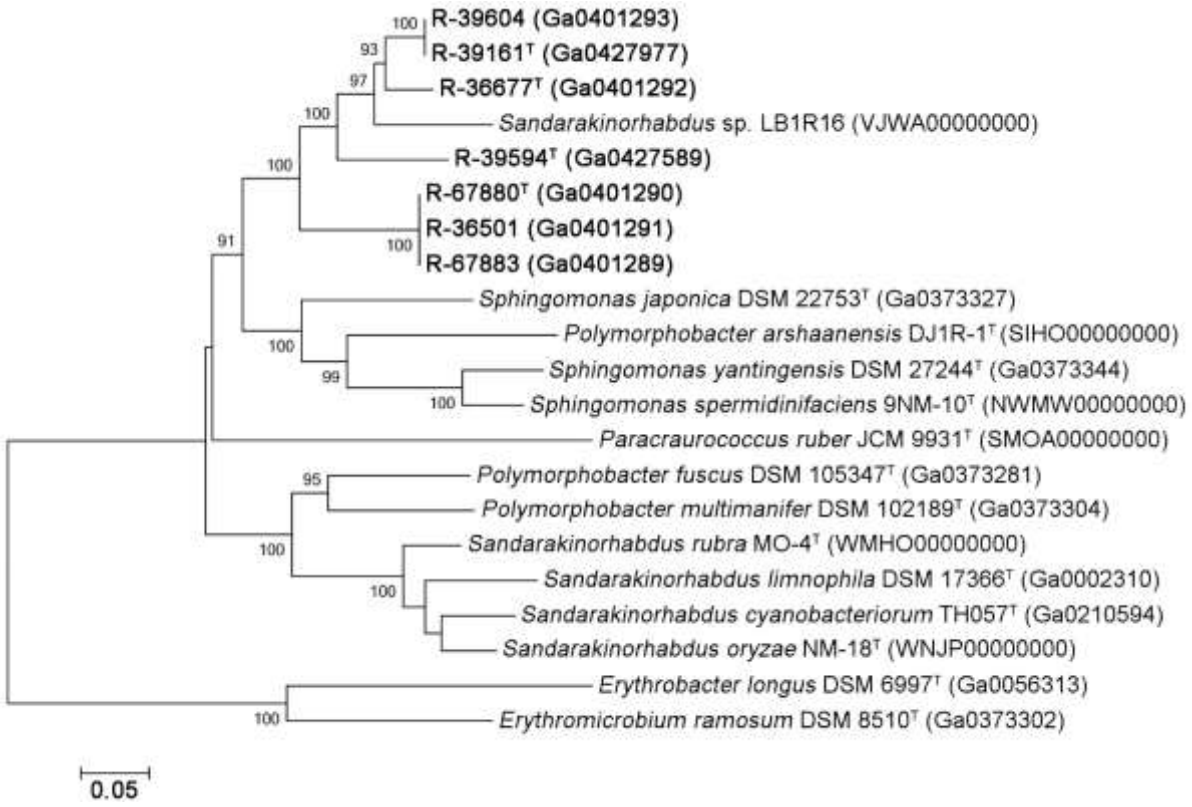

Fig. S5. Maximum-likelihood phylogenetic tree (1000 bootstraps) of the *pufLM* gene sequences of strains R-67883, R-67880^T, R-36501, R-36677^T, R-39604, R-39161^T and R-39594^T and closely related type strains of the families *Sphingomonadaceae* and *Acetobacteraceae*. Only bootstrap values higher than 70% are shown. The scale bar indicates 0.05 substitutions per site. *Erythrobacter longus* DSM 6997^T and *Erythromicrobium ramosum* DSM 8510^T were used as an outgroup.

1.2. Supplementary Tables

Table S1. Results based on API ZYM tests of the new strains and type strains of closely related species of the family *Sphingomonadaceae*.

Strains: 1, R-67883; 2, R-67880^T; 3, R-36501; 4, R-36677^T; 5, R-39604; 6, R-39161^T; 7, R-39594^T; 8, *Polymorphobacter multimanifer* DSM 102189^T; 9, *Polymorphobacter fuscus* DSM 105347^T; 10, *Sandarakinorhabdus limnophila* DSM 17366^T; 11, *Sphingomonas spermidinifaciens* DSM 27571^T. All strains are negative for Lipase (C14), Cystine arylamidase†-††, β-galactosidase††, β-glucuronidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase activities. +, positive; -, negative.

Test	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Alkaline phosphatase	+	+	+	+	+	+	+	+	+	-†	+
Esterase (C4)	+	+	+	+	+	+	+	+	+	-†	-††
Esterase lipase (C8)	+	-	-	-	-	-	+	+	-†	-†	+
Leucine arylamidase	+	-	+	-	-	+	-	+	+	+	+
Valine arylamidase	-	-	-	-	-	-	-	+	-†	-†	+
Trypsin	+	-	-	-	-	-	-	-†	-	-	-††
α-chymotrypsin	+	+	+	-	-	+	+	-	-†	-†	-††
Acid phosphatase	+	-	-	-	-	+	-	+	+	+	+
Naphthol-AS-BI-phosphohydrolase	-	-	-	-	-	-	-	+	+	-†	-††
α-galactosidase	-	-	-	-	-	-	-	-	+†	-	-
α-glucosidase	+	-	-	-	-	-	-	-	-	-	-††
β-glucosidase	+	+	+	-	-	-	-	-	-	-	-††

†Different results were reported by Fukuda *et al.* (2014), Jia *et al.* (2015) and Gich and Overmann (2006).

††Different results were reported by Fan *et al.* (2019).

Table S2. Results based on API 50 CH tests of the new strains and type strains of closely related species of the family *Sphingomonadaceae*.

Strains: 1, R-67883; 2, R-67880^T; 3, R-36501; 4, R-36677^T; 5, R-39604; 6, R-39161^T; 7, R-39594^T; 8, *Polymorphobacter multimanifer* DSM 102189^T; 9, *Polymorphobacter fuscus* DSM 105347^T; 10, *Sandarakinorhabdus limnophila* DSM 17366^T; 11, *Sphingomonas spermidinifaciens* DSM 27571^T. All strains are negative for acid formation from Glycerol†, Dulcitol†, Inositol†, D-Mannitol†, D-Sorbitol†, Inulin†, L-Fucose†, D-Arabitol†, L-Arabitol†, Potassium Gluconate† and Potassium 2-Ketogluconate†. +, positive; -, negative; w, weakly positive.

Test	1	2	3	4	5	6	7	8	9	10	11
Erythritol	-	-	-	-	-	-	-	-†	+	-†	-
D-Arabinose	w	-	w	-	-	-	-	-†	-†	w†	+††
L-Arabinose	w	w	w	-	+	-	-	+	+	w†	+
D-Ribose	-	w	w	-	-	-	-	w†	-†	w†	+††
D-Xylose	-	w	w	-	-	-	-	w†	+	-	+
L-Xylose	-	w	-	-	-	-	-	+	-†	-†	+††
D-Adonitol	-	+	+	-	-	+	+	+	-†	-†	+
Methyl-βD-Xylopyranoside	w	+	+	-	-	+	+	+	-†	+	+
D-Galactose	w	-	w	-	-	-	-	-	+	-	+
D-Glucose	w	w	-	-	-	-	-	-†	-†	-†	+
D-Fructose	-	-	-	-	-	-	-	-†	-†	-†	+††
D-Mannose	-	-	-	-	-	-	-	-†	-†	-†	+
L-Sorbose	-	-	-	-	-	-	-	-†	-†	-†	+††
L-Rhamnose	-	-	-	-	-	-	-	-†	+	-†	-
Methyl- αD-Mannopyranoside	-	-	-	-	-	-	-	-†	-†	-†	+
Methyl- αD-Glucopyranoside	w	-	-	-	-	-	-	w†	+	-†	+
N-acetylglucosamine	-	-	-	-	-	-	-	w†	-†	-†	-††
Amygdalin	-	-	w	-	w	-	-	w†	-†	-†	+
Arbutin	-	-	-	-	-	-	-	-†	-†	-†	+††
Esculin Ferric Citrate	+	+	+	+	+	+	+	+	+	+†	+
Salicin	-	w	w	-	w	-	-	w†	-†	-†	+††
D-Cellobiose	+	+	w	-	w	-	-	-†	-†	-†	+
D-Maltose	+	+	+	+	-	-	w	-†	-†	-†	+

Test	1	2	3	4	5	6	7	8	9	10	11
D-Lactose	-	-	+	+	-	-	+	-†	-†	-†	+
D-Melibiose	-	-	-	-	-	-	-	-†	-†	-†	+††
D-Saccharose	-	-	-	-	-	w	-	-†	-†	-†	+
D-Trehalose	w	w	w	-	-	-	-	-†	-†	-†	+††
D-Melezitose	w	w	w	-	-	-	-	-†	-†	-†	+
D-Raffinose	-	-	-	-	w	-	-	-†	-†	-†	+††
Amidon (starch)	+	+	-	w	-	-	-	-†	-†	-†	+††
Glycogen	-	-	w	-	-	-	-	-†	-†	-†	-
Xylitol	w	w	w	-	-	-	-	-†	-†	-†	-
Gentiobiose	+	+	+	+	-	-	+	-†	-†	-†	+
D-Turanose	-	-	-	-	-	-	+	-†	-†	-†	+
D-Lyxose	-	-	-	w	-	-	-	-†	+	-†	+††
D-Tagatose	-	-	-	-	-	-	-	w†	-†	-†	-
D-Fucose	w	w	w	w	+	-	w	+	+	w†	+
Potassium 5-Ketogluconate	-	-	+	+	+	w	+	+	+	w†	+††

†Different results were reported by Jia *et al.* (2015).

††Different results were reported by Fan *et al.* (2019).

Table S3. Results based on API 20 NE tests of the new strains and type strains of closely related species of the family *Sphingomonadaceae*.

Strains: 1, R-67883; 2, R-67880^T; 3, R-36501; 4, R-36677^T; 5, R-39604; 6, R-39161^T; 7, R-39594^T; 8, *Polymorphobacter multimanifer* DSM 102189^T; 9, *Polymorphobacter fuscus* DSM 105347^T; 10, *Sandarakinorhabdus limnophila* DSM 17366^T; 11, *Sphingomonas spermidinifaciens* DSM 27571^T. Strains are negative for the reduction of nitrates to nitrites and nitrates to nitrogen and indole production. +, positive; -, negative; w, weakly positive.

Test	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Glucose fermentation	-	-	-	-	-	-	-	-	-	-	+†
Arginine dihydrolase	-	-	-	-	-	-	-	+†	-	-†	-
Urease	-	w	+	+	w	w	-	w†	-	+	-
Esculin hydrolysis	+	+	+	+	+	+	+	+†	-†	+	+
Gelatin hydrolysis	+	+	+	+	w	-	+	+†	+†	-†	+
Para-nitrophenyl-βD-galactopyranosidase	w	w	w	w	+	+	+	w†	+†	+	+
Assimilation of:											
D-Glucose	+	+	+	w	w	w	w	+†	+†	+	+
L-Arabinose	w	w	w	w	w	w	w	+†	+†	+	+
D-Mannose	w	w	w	w	w	w	w	+†	-	+	+
D-Mannitol	w	w	w	w	w	w	w	+†	-	+†	w†
N-Acetyl-glucosamine	+	+	+	w	w	w	w	+†	-	+	+†
D-Maltose	+	+	+	w	w	w	w	+	-	+	+
Potassium gluconate	w	+	w	w	w	w	w	+†	+†	+†	+†
Capric acid	-	-	-	-	-	-	-	-	-	+†	-
Adipic acid	w	+	+	w	w	w	w	+†	+†	+†	+†
Malic acid	w	w	w	w	w	w	w	+	-	+†	+
Trisodium citrate	w	w	w	w	w	w	w	+†	-	-	w†
Phenylacetic acid	-	w	w	w	w	w	w	+†	-	+†	w†

†Different results were reported by Fukuda *et al.* (2014), Jia *et al.* (2015) and Feng *et al.* (2017).