International Journal of Systematic and Evolutionary Microbiology

Sphingobacterium hungaricum sp. nov. a novel species on the borderline of the genus Sphingobacterium

--Manuscript Draft--

Manuscript Number:	IJSEM-D-21-00404R1
Full Title:	Sphingobacterium hungaricum sp. nov. a novel species on the borderline of the genus Sphingobacterium
Article Type:	Taxonomic Description
Section/Category:	New taxa - Bacteroidetes
Keywords:	Sphingobacterium hungaricum; new taxon; Bacteroidetes; Sphingobacteriaceae; borderline of the genus
Corresponding Author:	Ákos Tóth National Agricultural Research and Innovation Center Budapest, HUNGARY
First Author:	Ákos Tóth
Order of Authors:	Ákos Tóth
	Rózsa Máté
	József Kutasi
	Erika Tóth
	Károly Bóka
	András Táncsics
	István Nagy
	Gábor Kovács
	Judit Kosztik
	Ildikó Bata-Vidács
	József Kukolya
Manuscript Region of Origin:	HUNGARY
Abstract:	A Gram-reaction-negative bacterial strain, designated Kb22 T , was isolated from agricultural soil and characterised using a polyphasic approach to determine its taxonomic position. On the basis of 16S rRNA gene sequence analysis, the strain shows highest similarity (94.39%) with Sphingobacterium nematocida M-SX103 T . The highest ANI (71.83%) value was found with Sphingobacterium composti T5-12 T , and the highest AAI (66.65%) value was found with Sphingobacterium olei HAL-9 T . Cells are aerobic, non-motile rods. The isolate was found to be positive for catalase and oxidase tests. The assembled genome of strain Kb22 T has a total length of 4,06 Mb, the DNA G+C content is 38.1 mol%. The only isoprenoid quinone is menaquinone 7 (MK-7). The major fatty acids are iso-C 15:0 (28.4%), summed feature 3 (C 16:1 ω 7 c and/or iso-C 15:0 2-OH) (25.7%) and iso-C 17:0 3-OH (19.7%). Based on phenotypic characteristics and phylogenetic analysis, it is concluded that strain Kb22 T is a member of the genus Sphingobacterium and represents a novel species, for which the name Sphingobacterium hungaricum sp. nov. is proposed. The type strain of the species is strain Kb22 T (=LMG 31574 = NCAIM B.02638).
Author Comments:	
Opposed Reviewers:	
Suggested Reviewers:	Peter Schumann DSMZ: Leibniz Institut - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH

	psc@dsmz.de
	Károly Márialigeti ELTE: Eotvos Lorand Tudomanyegyetem marialigeti.karoly@ttk.elte.hu
Additional Information:	
Question	**Response**
Does this article report on work with humans or animals?	No
Does this article include details (names, initials, hospital numbers), images, or videos relating to an individual person?	No

Sphingobacterium hungaricum sp. nov. a novel species on the borderline of the genus *Sphingobacterium*

Ákos Tóth[1,*] , Rózsa Máté[2], József Kutasi[2], Erika Tóth[3], Károly Bóka[4], András Táncsics[5], István Nagy[6,7], Gábor Kovács[6,8] Judit Kosztik[1], Ildikó Bata-Vidács[1], József Kukolya[1]

[1]Research Group for Food Biotechnology, Institute of Food Science and Technology, Hungarian University of Agriculture and Life Sciences, Budapest, Hungary; [2]BioFil Microbiological, Biotechnological and Biochemical Ltd., Budapest, Hungary; [3]Department of Microbiology, Eötvös Loránd University, Budapest, Hungary; [4]Department of Plant Anatomy, Eötvös Loránd University, Budapest, Hungary; [5]Department of Molecular Ecology, Institute of Aquaculture and Environmental Safety, Hungarian University of Agriculture and Life Sciences. Gödöllő, Hungary; [6]SeqOmics Biotechnology Ltd., Mórahalom, Hungary; [7]Institute of Biochemistry, Biological Research Centre, Eötvös Loránd Research Network, Szeged, Hungary; [8]University of Sopron, Sopron, Hungary

[*]Correspondence: Ákos Tóth, Tel.: +36 1 796 0400; E-mail address: toth.akosgergely@gmail.com

Keywords: *Sphingobacterium hungaricum*; new taxon; *Bacteroidetes, Sphingobacteriaceae;* borderline of the genus

Abbreviations: AAI, Amino Acid Identity; ANI, Average Nucleotide Identity; dDDH, digital DNA–DNA Hybridisation; DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen (German Collection of Microorganisms and Cell Cultures); GGDC, Genome-to-Genome Distance Calculator; GH, Glycoside Hydrolase; GNL, aminoglycolipid; LB agar, Luria-Bertani agar; L, uncharacterised lipid; MiGA, Microbial Genomes Atlas; TYGS, Type Strain Genome Server; MLSA, MultiLocus Sequence Analysis; RAST, Rapid Annotation using Subsystem Technology; PE, phosphatidylethanolamine; PGL, phosphoglycolipid; PL, phospholipid.

Author notes: The GenBank accession numbers for the 16S rRNA gene sequence and the whole genome of *Sphingobacterium hungaricum* strain Kb22^T are MF471353 and PRDK00000000, respectively.

A Gram-reaction-negative bacterial strain, designated Kb22^T, was isolated from agricultural soil and characterised using a polyphasic approach to determine its taxonomic position. On the basis of 16S rRNA gene sequence analysis, the strain shows highest similarity (94.39%) with *Sphingobacterium nematocida* M-SX103^T. The highest ANI (71.83%) value was found with *Sphingobacterium composti* T5-12^T, and the highest AAI (66.65%) value was found with *Sphingobacterium* olei HAL-9^T. Cells are aerobic, non-motile rods. The isolate was found to be positive for catalase and oxidase tests. The assembled genome of strain Kb22^T has a total length of 4,06 Mb, the DNA G+C content is 38.1 mol%. The only isoprenoid quinone is menaquinone 7 (MK-7). The major fatty acids are iso-$C_{15:0}$ (28.4%), summed feature 3 ($C_{16:1}$ *ω7c* and/or iso-$C_{15:0}$ 2-OH) (25.7%) and iso-$C_{17:0}$ 3-OH (19.7%).

Based on phenotypic characteristics and phylogenetic analysis, it is concluded that strain Kb22^T is a member of the genus *Sphingobacterium* and represents a novel species, for which the name *Sphingobacterium hungaricum* sp. nov. is proposed. The type strain of the species is strain Kb22^T (=LMG 31574 = NCAIM B.02638).

Introduction

The family *Sphingobacteriaceae* was proposed by Steyn *et al.* [1] and emended by García-López *et al.* [2]. *Sphingobacteriaceae* belong to the order *Sphingobacteriales*, class *Sphingobacteriia* and phylum '*Bacteroidetes*'. The type genus of the family is *Sphingobacterium*. At the time of writing, the family includes 15 validly published genera (https://lpsn.dsmz.de/ April, 2021 [3, 4]). According to Steyn *et al.*, members of the family are Gram-stain-negative, non-motile rods, having iso-$C_{15:0}$, iso-$C_{15:0}$ 2-OH, iso-$C_{15:0}$ 3-OH, $C_{16:0}$, $C_{16:1}$ *ω7c*, $C_{16:O}$ 3-OH and iso-$C_{17:O}$ 3-OH as predominant fatty acids; phosphatidylethanolamine as the major polar lipid and menaquinone-7 (MK-7) as the major respiratory quinone [1]. The members of *Sphingobacteriaceae* have been mostly isolated from soil and water. The genus *Sphingobacterium* was proposed by Yabuuchi *et al.* [5] and emended by Wauters *et al.* [6]. At the time of writing, the genus includes 58 taxa with validly published and correct names (https://lpsn.dsmz.de/ April, 2021 [3, 4]), the type species is *Sphingobacterium spiritivorum* [5]. The main characteristics of the genus are the negative result of Gram-staining, positive catalase and oxidase tests, rod shape, non-motility and MK-7

as predominant isoprenoid quinone. Most of the *Sphingobacterium* strains have been isolated from soil, rhizosphere or composts.

Isolation and ecology

Strain Kb22^T was isolated from an agricultural field in the Great Hungarian Plain, Hungary. The approximate geographical coordinates were 47° 11' 56" N and 19° 00' 46" E. Before sampling, maize was harvested from the field. The soil was fertilised, and its pH was moderately alkaline. After sampling, the soil particles were homogenised by vortexing and serially diluted with peptone water (9 g peptone, 1 g NaCl, in 1000 ml dH_2O). It was subsequently spread onto xylan containing agar (1 g $NaNO_3$; 1 g K_2HPO_4; 3 g NaCl; 0.5 g $MgCl_2$; 0.5 g yeast extract; 0.5 g peptone; 3 g xylan; 25 g agar; 1000 ml dH_2O) and incubated at 10 °C for 5 days. Single colonies on the plates were purified on the same medium. The isolate is routinely maintained on LB medium (DSM medium No. 381, www.dsmz.de) at 28 °C and pH 7.5.

16S phylogeny

Sphingobacterium is a relatively large genus, at the time of writing (June of 2021) the genus includes 58 validly published species with correct name. New species in the genus have already been described with 16S rRNA gene similarity of 90.0% [7] and 99.1% [8]. The golden standard for species delineation was DNA-DNA hybridisation, then Stackebrandt and Goebel proposed a boundary of 16S rRNA gene sequence similarity of 97% for species delineation based on a correlation analysis between DDH values and 16S rRNA gene sequence identities [9]. Thus, until recently, taxonomic frameworks primarily based on 16S rRNA gene sequences. However, this threshold value has been updated to 98.7% [10]. As for genera, the generally used genus threshold of 95% 16S rRNA gene identity has been recently revised to a minimal value of 94.5% [11].

DNA was extracted from Kb22^T liquid culture grown in LB medium. Genomic DNA isolation and 16S rRNA gene amplification were performed according to Tóth *et al.* [12]. The partial 16S rRNA gene sequence of strain Kb22^T (MF471353) was compared with the EzBioCloud Database (http://www.ezbiocloud.net/taxonomy) [13] for an approximate phylogenetic affiliation. After Sanger sequencing of the 16S rRNA gene, a genome sequencing project of Kb22^T was carried out. According to the comparisons with the complete 16S rRNA gene sequences in the EzBioCloud Database, the highest level of sequence similarity occurred with

Sphingobacterium nematocida M-SX103^T (94.39%) [14], followed by *Sphingobacterium composti* T5-12^T (94.31%) [15].

Phylogenetic tree based on 16S rRNA gene was inferred by using the neighbor-joining method [16] with Kimura's two-parameter calculation model [17]. Tree topologies and distances were evaluated by bootstrap analysis based on 1000 replicates. Evolutionary analyses were conducted in MEGA X [18]. The phylogenetic tree based on 16S rRNA gene sequences suggested that strain Kb22^T forms a distinct phyletic lineage in the *Sphingobacteriaceae* family.

Genome features

The genome of strain Kb22^T was sequenced with Illumina MiSeq sequencing technology as described previously [19]. Genome assembly was performed by SPAdes v. 3.9.1; CLC NGS Cell v. 11.0. Genome completeness and contamination values were examined by TypeMet tool of MiGA server (http://microbial-genomes.org/) [20]. Annotation of the genome was performed by NCBI Prokaryotic Genome Annotation Pipeline v4.4 with Best-placed reference protein set and GeneMarkS+ methods [21, 22] and Rapid Annotation using Subsystem Technology server v. 2.0 (RAST; https://rast.nmpdr.org) [23]. The completeness and contamination metrics of the genome were found to be 96.2% and 1.9%, respectively. Other quality metrics of genome sequencing and assembly were as follows: 227-fold genome coverage, contig N50=578,756, number of contigs was 10. The genome size and G+C content of Kb22^T were 4,056,205 bp and 38.1 mol%, respectively. According to the annotation, there were 3603 genes, 3548 CDSs and 55 RNA genes in the genome. The coding density was 89.51%.

The anti-SMASH server was used to identify the secondary metabolite biosynthesis gene clusters [24]. Four putative biosynthetic gene clusters (arylpolyene, resorcinol and two furans) were found in 3 genomic regions.

Strain Kb22^T was isolated on xylan containing minimal agar, and it was also able to grow on media containing mannan or cellulose as sole carbon source. The genome annotation revealed 55 glycoside hydrolases (GHs) in 21 different GH families (Table 1). These enzyme genes may play a role in the breakdown and modification of carbohydrates in soil. Some of these enzymes are active on plant cell wall polysaccharides and potentially have role in the breakdown of lignocelluloses. The main polysaccharide components in lignocellulose are cellulose and xylan. The key enzymes for the decomposition of these two main components

are a xylanases, β-xylosidases, cellulases and β-glucosidases. The presence of these genes was examined in *Sphingobacterium hungaricum* Kb22^T, *Sphingobacterium composti* KCTC 12578^T, *Sphingobacterium olei* HAL-9^T [25] and *Sphingobacterium nematocida* DSM_24091^T (Table 2). According to the annotation of genomes (GCA_015210005, GCA_009829075, GCA_005048855, GCA_900168125), we found the key genes required for xylan degradation in Kb22^T, but in the case of the endophytic *Sphingobacterium nematocida* DSM_24091^T no xylanase or cellulase genes were found. The results revealed differences in the gene set of the compared strains, and it is important to note that lignocellulose degradation in nature is performed by microbial communities. The glycoside hydrolase sequences were identified using the Carbohydrate Active Enzymes database (http://www.cazy.org/) and the Interpro web service (https://www.ebi.ac.uk/interpro/) [26, 27].

Genome based phylogeny

As the number of whole genome sequences increased, it became possible to introduce different overall genome related indexes (OGRI). However, according to Chun et al., OGRI does not have a taxonomic resolution above the species level, so a multigene-based phylogenomic treeing approach should be chosen for defining genera [28]. Others recommend the use of genus-specific ANI and AF (Aligment Fraction) values [29]. Recently, the genomic-scale phylogenetic, AAI and conserved signature indels (CSIs) based classification of broad genera seems to be suitable [30-32].

The 16S rRNA gene similarity found for Kb22^T is at the genus boundary, thus also MLSA and whole genome sequence based methods were used to determine the phylogenetic position of the strain.

Multilocus sequence analyses were performed by autoMLST webserver (https://automlst.ziemertlab.com/) [33]. The concatenated sequences were made from 78 genes (Table S1). Phylogenetic tree based on concatenated sequences was inferred using the neighbor-joining method [16] with Kimura's two-parameter calculation model [17] using MEGA version X [18]. Tree topologies and distances were evaluated by bootstrap analysis based on 1000 replicates. Based on the MLSA tree, Kb22^T was found to be a member of the genus *Sphingobacterium* as a novel species (Fig. 1).

Genome-based relatedness between Kb22^T and *S. spiritivorum* NCTC 11386^T, the type strain of the type species, was determined based on ANI using OrthoANI (https://www.ezbiocloud.net/tools/ani) algorithm [34] and dDDH (identities/HSP length) with

GGDC service of DSMZ (http://ggdc.dsmz.de/) [35]. The OrthoANIu and dDDH values between Kb22^T and *S. spiritivorum* NCTC 11386^T were 70.37 and 19.90%, much lower than the generally accepted species boundary of 95~96 and 70%, respectively [35-37]. The highest OrthoANIu value(71.83%) was found with *Sphingobacterium composti* T5-12^T.

As part of the phylogenomic studies, TYGS (https://tygs.dsmz.de/) [38] and MiGA (http://microbial-genomes.org/) [20] webservers were used. Phylogenomic treeing by TYGS confirmed the MLSA result that Kb22^T represents a new species within genus *Sphingobacterium* (Fig. 2). An AAI matrix with type strains in the *Sphingobacteriaceae* family was calculated by the latest extension of the Microbial Genome Atlas (https://xsede.microbial-genomes.org/). The lowest value was observed with *Anseongella ginsenosidimutans* Gsoil 524^T (52%), the highest with *Sphingobacterium spiritivorum* NCTC 11386^T (66%) (Table 3). Genbank assembly accessions: *Albibacterium bauzanense* DSM 22554^T - GCA_004339765.1 [39]; *Anseongella ginsenosidimutans* Gsoil 524^T - GCA_008033235.1 [40]; *Arcticibacter svalbardensis* MN12-7^T - GCA_000403135.1 [41]; *Mucilaginibacter paludis* DSM 18603^T - GCA_000166195.3 [42]; *Nubsella zeaxanthinifaciens* TDMA-5^T - GCA_003313335.1 [43]; *Olivibacter sitiensis* DSM 17696^T - GCA_000427965.1 [44]; *Parapedobacter koreensis* Jip14^T - GCA_900109365.1 [45]; *Pararcticibacter amylolyticus* FJ4-8^T - GCA_003130405.1 [46]; *Pedobacter heparinus* DSM 2366^T - GCA_000023825.1 [1]; *Pelobium manganitolerans* YS-25^T - GCA_003609575.1 [47]; *Pseudosphingobacterium domesticum* DSM 18733^T - GCA_900109575.1 [48]; *Solitalea koreensis* DSM 21342^T - GCA_900182575.1 [49]; *Sphingobacterium spiritivorum* NCTC 11386^T - GCA_900457435.1 [5].

According to MiGA distance analysis, the closest relatives are *Sphingobacterium olei* HAL-9^T (accession: GCA_005048855) (66.65% AAI) and *Sphingobacterium alkalisoli* Y3L14^T (accession: GCA_005049105**)** (66.39% AAI) [50]. The p-value of taxonomic novelty at genus level is 0.655 and at species level is 0.002. Summarising the phylogenetic assessments, the taxonomic position of Kb22^T is a borderline case. In accordance with the recommendation of Chun *et al*., the phylogenomic results were considered decisive for genus level rank [28].

Physiologic, morphologic and chemotaxonomic characterisation

Biomass for chemical and molecular studies was obtained by cultivation in shaker flasks (180 r.p.m.) using LB medium at 30 °C for 32 h. Colony morphology of strain Kb22^T was determined on LB agar medium by directly observing single colonies. Cell morphology of

strain Kb22^T was observed under electron microscope. Pigments were investigated as described previously by Daood and Biacs [51] and by Bernardet and Bowman [52]. The Gram reaction was determined with a non-staining method as described by Buck *et al*. [53]. Oxidase activity was studied with OXI oxidase test strip (Diagnostics s.r.o.). Catalase production was demonstrated by the method of Barrow and Feltham [54]. Growth at different temperatures (from 4 to 50 °C), NaCl tolerance (0–5% w/v) and pH tolerance (pH 4–10, using increments of 0.5 pH unit, pH values were adjusted with HCl or NaOH) were determined using LB medium. Growth at pH 4-10 was examined in flasks and 96-well plates with continuous monitoring of optical density. Acid production from different carbon sources, the assimilation of different substrates and the enzymatic activities of strain Kb22^T were investigated with API 50 CH, API 20 NE and API ZYM kits (BioMérieux) according to the manufacturer's instructions. The API 50CH and 20NE tests were read after 24–48 h incubation at 28 °C. Anaerobic and microaerophilic growth was checked on TSA medium using the Anaerocult A and C systems (Merck). The physiological characteristics were examined by side-by-side analysis with *Sphingobacterium composti* T5-12^T [15].

LB medium is used for general laboratory cultivation, but the novel strain also grew well on TSA, nutrient and R2A media. After 72 h growth on LB at 30 °C, colonies were found to be 1 mm in diameter, circular, non-mucoid, smooth and orange. Cells of Kb22^T produced flexirubin-type pigments. Strain Kb22^T was found to be Gram-reaction-negative, aerobic, positive for oxidase and catalase, rod-shaped bacterium (Fig. S1). Cells were observed to be non-motile, grow in 0.0–2.0% (w/v) NaCl, at a pH range 6.5 to 8.5 and at a temperature range between 10 and 35 °C. Optimal growth was observed at 30 °C, 1.0% (w/v) NaCl and pH 7.5. Mean cell size of Kb22^T was found to be 0.5 µm in diameter and 1.0-1.5 µm in length.

According to API 50 CH test, Kb22^T was positive for acid production from D-arabinose, D-glucose, D-fructose, D-mannose, L-rhamnose, methyl-α-D-mannopyranoside, methyl-α-D-xylopyranoside, N-acetyl-glucosamine, amygdaline, arbutine, esculine, salicine, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, D-melezitose, D-raffinose, starch, gentiobiose, D-turanose and L-fucose. β-galactosidase activity, hydrolysis of esculin and assimilation of glucose, mannose, N-acetyl-glucosamine and maltose were demonstrated by using the API 20 NE test. In the API ZYM test, strain Kb22^T was positive for alkaline phosphatase, esterase lipase (C8), leucine arylamidase, naphthol-AS-BI-phosphohydrolase, β-glucosidase, N-acetyl-β-glucosaminidase and trypsin. Differential phenotypic characteristics between strain Kb22^T and *Sphingobacterium composti* T5-12^T are given in Table 4. According to comparative genomics performed on MicroScope platform, the strain specific

CDSs account for 80.577% in Kb22^T genome and 83.065% in *Sphingobacterium composti* T5-12^T genome [55].

Analyses of chemotaxonomic traits were carried out by DSMZ Identification Service, Braunschweig, Germany. For polar lipid and respiratory quinone analyses, the strains were cultivated in LB liquid medium at 30 °C, 180 r.p.m. until exponential growth was reached. The fatty acid profiles of strain Kb22^T and *Sphingobacterium composti* T5-12^T were analysed using active growing cultures on LB at 30 °C.

According to the DSMZ Identification Service, fatty acid methyl esters (FAMEs) were obtained following the method of Miller [56] and Kuykendall *et al*. [57]. FAMEs were separated by gas chromatography, detected by a flame ionisation detector using Sherlock Microbial Identification System (MIS) (MIDI, Microbial ID, Newark, DE 19711 U.S.A.) and identified by using the TSBA40 4.10 database of the Microbial Identification System. Summed feature components were identified thereafter by GC/MS.

The predominant cellular fatty acids of strain Kb22^T were observed to be iso-$C_{15:0}$ (28.4%), summed feature 3 ($C_{16:1}\omega 7c$ and/or iso-$C_{15:0}$ 2-OH) (25.7%) and iso-$C_{17:0}$ 3-OH (19.7%). The fatty acid profile was found to be similar to that of related species, in accordance with the description of *Sphingobacterium* genus [5]. However, the ratios of the components were different. The complete fatty acid composition is shown in Table 5.

The respiratory quinones were extracted from freeze dried material and purified by a silica-based solid phase extraction. Purified samples were further analysed by HPLC and UHPLC-ESI-qTOF system [58, 59; dsmz.de]. The only respiratory quinone of Kb22^T was menaquinone-7 (MK-7).

Polar lipids were studied according to Tindall *et al.* [58-60, dsmz.de]. Strain Kb22^T exhibited a complex polar lipid profile consisting of phosphatidylethanolamine (PE) and phosphoglycolipid (PGL) as dominant elements and an uncharacterised aminoglycolipid (GNL), six uncharacterised phospholipids (PL) and five uncharacterised lipids (L) (Fig. S2). Though the domination of PE is characteristic of other species in the genus, the presence and ratio of other components were different [5].

In conclusion, phenotypic, biochemical, chemotaxonomic and phylogenetic information of strain Kb22^T support its classification as a novel species of *Sphingobacterium*, for which the name *Sphingobacterium hungaricum* sp. nov. is proposed. The GenBank accession numbers for the 16S rRNA gene sequence and the whole genome of *Sphingobacterium hungaricum* strain Kb22^T are MF471353 and PRDK00000000, respectively.

Description of *Sphingobacterium hungaricum* sp. nov.

Sphingobacterium hungaricum (hun.ga'ri.cum. M.L. neut. adj. *hungaricum* of or belonging to Hungary, where the type strain was isolated)

Cells are strictly aerobic, Gram-reaction-negative straight rods and non-motile. It grows well on LB, TSA, nutrient and R2A media. Colonies have orange pigmentation on TSA after 48 h of incubation. Cells are 0.5 µm in diameter and 1.0-1.5 µm in length. It grows at 10–35 °C (optimum, 30 °C) and at NaCl concentrations of 0–2 w/v % (optimum, 1 w/v %). It is positive for oxidase, catalase, β-galactosidase, alkaline phosphatase, esterase lipase (C8), leucine arylamidase, naphthol-AS-BI-phosphohydrolase, β-glucosidase, N-acetyl-β-glucosaminidase, trypsin, hydrolysis of esculin and assimilation of glucose, mannose, N-acetyl-glucosamine, and maltose. Strain Kb22^T was found to be positive for acid production from D-arabinose, D-glucose, D-fructose, D-mannose, L-rhamnose, methyl-α-D-mannopyranoside, methyl-α-D-xylopyranoside, N-acetyl-glucosamine, amygdaline, arbutine, esculine, salicine, D-cellobiose, D-maltose, D-lactose, D-melibiose, D-saccharose, D-trehalose, D-melezitose, D-raffinose, starch, gentiobiose, D-turanose, and L-fucose. The major fatty acids are iso-$C_{15:0}$ (28.4%), summed feature 3 ($C_{16:1}$ *ω7c* and/or iso-$C_{15:0}$ 2-OH) (25.7%) and iso-$C_{17:0}$ 3-OH (19.7%). The only respiratory quinone is MK-7. The major polar lipid is phosphatidylethanolamine. The type strain is Kb22^T (=LMG 31574 = NCAIM B.02638), isolated from agricultural field in the Great Hungarian Plain, Hungary. The DNA G+C content of the type strain is 38.1 mol%, the genome is 4.06 Mb. The GenBank accession number of the genome is PRDK00000000.

Funding Information

This work was supported by the NVKP_16-1-2016-0009 project. Ákos Tóth was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (Grant No. BO/00342/18).

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. Steyn PL, Segers P, Vancanneyt M, Sandra P, Kersters K, Joubert JJ. Classification of heparinolytic bacteria into a new genus, *Pedobacter*, comprising four species:

Pedobacter heparinus comb. nov., *Pedobacter piscium* comb. nov., *Pedobacter africanus* sp. nov. and *Pedobacter saltans* sp. nov. proposal of the family *Sphingobacteriaceae* fam. nov. *Int J Syst Bacteriol* 1998;48(1):165-177.

2. Garcia-Lopez M., Meier-Kolthoff JP, Tindall BJ, Gronow S, Woyke T, Kyrpides NC, Hahnke RL, Goker M. Analysis of 1000 type-strain genomes improves taxonomic classification of *Bacteroidetes*. *Front. Microbiol* 2019;10:2083. DOI: 10.3389/fmicb.2019.02083
3. Euzéby JP. List of bacterial names with standing in nomenclature: a folder available on the internet. *International Journal of Systematic Bacteriology* 1997;47:590-592. DOI: 10.1099/00207713-47-2-590
4. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC and Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *International Journal of Systematic and Evolutionary Microbiology* 2020;70:5607-5612. DOI: 10.1099/ijsem.0.004332
5. Yabuuchi E, Kaneko T, Yano I, Moss CW, Miyoshi N. *Sphingobacterium* gen. nov., *Sphingobacterium spiritivorum* comb. nov., *Sphingobacterium multivorum* comb. nov., *Sphingobacterium mizutae* sp. nov., and *Flavobacterium indologenes* sp. nov.: glucose-nonfermenting Gram-negative rods in CDC groups IIK-2 and IIb. *Int. J. Syst. Bacteriol.* 1983;33:580-598.
6. Wauters G, Janssens M, De Baere T, Vaneechoutte M, Deschaght P. Isolates belonging to CDC group II-i belong predominantly to *Sphingobacterium mizutaii* Yabuuchi *et al.* 1983: emended descriptions of *S. mizutaii* and of the genus *Sphingobacterium*. *Int J Syst Evol Microbiol* 2012;62:2598-2601.
7. Jiang S, Chen M, Su S, Yang M, Li A, Zhang C, Lin M, Zhang W, Luo X. *Sphingobacterium arenae* sp. nov., isolated from sandy soil. *Int J Syst Evol Microbiol* 2014;64:248-253.
8. Lee DH, Hur JS, Kahng HY. *Sphingobacterium cladoniae* sp. nov., isolated from lichen, *Cladonia* sp., and emended description of *Sphingobacterium siyangense*. *Int J Syst Evol Microbiol* 2013;63:755-760.
9. Stackebrandt E, Goebel BM. Taxonomic Note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* 1994;44:846–849.
10. Stackebrandt E, Ebers J. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006;33:152–155.

11. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K, Whitman WB, Euzéby J, Amann R, Móra RR. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;12:635-645. DOI:10.1038/ nrmicro3330
12. Tóth Á, Baka E, Bata-Vidács I, Luzics S, Kosztik J *et al*. *Micrococcoides hystricis* gen. nov., sp. nov., a novel member of the family *Micrococcaceae*, phylum *Actinobacteria*. *Int J Syst Evol Microbiol* 2017;67:2758-2765.
13. Kim OS, Cho YJ, Lee K, Yoon SH, Kim M *et al*. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* 2012;62:716–721.
14. Liu J, Yang LL, Xu CK, Xi JQ, Yang FX, Zhou F, Zhou Y, Mo MH, Li WJ. *Sphingobacterium nematocida* sp. nov., a nematicidal endophytic bacterium isolated from tobacco. *Int J Syst Evol Microbiol* 2012;62:1809-1813.
15. Ten LN, Liu Q-M, Im W-T, Aslam Z, Lee S-T. *Sphingobacterium composti* sp. nov., a novel DNase-producing bacterium isolated from compost. *J Microbiol Biotechnol* 2006;16:1728-1733.
16. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;4:406–425.
17. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 1980:16:111-120.
18. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 2018:35:1547-1549.
19. Szuroczki S, Khayer B, Sproer C, Toumi M, Szabo A, Felfoldi T, Schumann P, Toth E. *Arundinibacter roseus* gen. nov., sp. nov., a new member of the family *Cytophagaceae*. *Int J Syst Evol Microbiol* 2019;69:2076-2081.
20. Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, et al. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of *Archaea* and *Bacteria* at the whole genome level. Nucleic Acids Res. 2018;46:W282–W288. DOI:10.1093/nar/gky467
21. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP *et al*. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 2016;44(14):6614-6624.

22. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D *et al*. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 2016;4(44):D733-D745.
23. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, *et al*. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 2008;9:75.
24. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY *et al*. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Weber Nucleic Acids Research* 2019 DOI: 10.1093/nar/gkz310. 29. Buck JD.
25. Liu B, Yang X, Sheng M, Yang Z, Qiu J, Wang C, He J. *Sphingobacterium olei* sp. nov., isolated from oil-contaminated soil. *Int J Syst Evol Microbiol* 2020; 70:1931-1939.
26. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The Carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495.
27. Blum M, Chang H, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S, RichardsonL, Salazar GA, Williams L, Bork P, Bridge A, Gough J, Haft DH, Letunic I, Marchler-Bauer A, Mi H, Natale DA, Necci M, Orengo CA, Pandurangan AP, Rivoire C, Sigrist CJA, Sillitoe I, Thanki N, Thomas PD, Tosatto SCE, Wu CH, Bateman A and Finn RD The InterPro protein families and domains database: 20 years on. *Nucleic Acids Research*, Nov 2020, DOI: 10.1093/nar/gkaa977.
28. Chun J, Oren A, Ventosa A, Christensen H, Arahal DR, da Costa MS. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol.* 2018;68:461–466.
29. Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH, Emerson DA. Genus definition for *Bacteria* and *Archaea* based on a standard genome relatedness index. *mBio.* 2020;11, e02475–19. DOI: 10.1128/mBio.02475-19
30. Gupta RS, Patel S, Saini N, Chen S. Robust demarcation of 17 distinct *Bacillus* species clades, proposed as novel *Bacillaceae* genera, by phylogenomics and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. nov. and proposal for an emended genus *Bacillus* limiting it only to the members of the Subtilis and Cereus clades of species. *Int J Syst Evol Microbiol*. 2020 Nov;70(11):5753-5798. DOI: 10.1099/ijsem.0.004475. Epub 2020 Oct 27. Erratum in: *Int J Syst Evol Microbiol*. 2020 Dec;70(12):6531-6533. PMID: 33112222.

31. Lawson PA, Rainey FA. Proposal to restrict the genus *Clostridium* Prazmowski to *Clostridium butyricum* and related species. *Int J Syst Evol Microbiol.* 2016 Feb;66(2):1009-1016. DOI: 10.1099/ijsem.0.000824. Epub 2015 Dec 7. PMID: 26643615.
32. Nicholson AC, Gulvik CA, Whitney AM, Humrighouse BW, Bell ME, Holmes B, Steigerwalt AG, Villarma A, Sheth M, Batra D, Rowe LA, Burroughs M, Pryor JC, Bernardet JF, Hugo C, Kämpfer P, Newman JD, McQuiston JR. Division of the genus *Chryseobacterium:* Observation of discontinuities in amino acid identity values, a possible consequence of major extinction events, guides transfer of nine species to the genus *Epilithonimonas*, eleven species to the genus *Kaistella*, and three species to the genus *Halpernia* gen. nov., with description of *Kaistella daneshvariae* sp. nov. and *Epilithonimonas vandammei* sp. nov. derived from clinical specimens. *Int J Syst Evol Microbiol.* 2020 Aug;70(8):4432-4450. DOI: 10.1099/ijsem.0.003935. Epub 2020 Jan 2. PMID: 32735208; PMCID: PMC7660247.
33. Alanjary M, Steinke K, Ziemert N. AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential, *Nucleic Acids Research*, Volume 47, Issue W1, 02 July 2019, Pages W276–W282, https://doi.org/10.1093/nar/gkz282
34. Yoon SH, Ha SM, Lim JM, Kwon SJ, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek.* 2017;10:1281–1286.
35. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
36. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P *et al*. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
37. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 2009;106:19126–19131.
38. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun*. 2019;10:2182. DOI: 10.1038/s41467-019-10210-3

39. Garcia-Lopez M, Meier-Kolthoff JP, Tindall BJ, Gronow S, Woyke T, Kyrpides NC, Hahnke RL, Goker M. Analysis of 1,000 Type-Strain Genomes Improves Taxonomic Classification of *Bacteroidetes*. *Front Microbiol* 2019; 10:2083.
40. Siddiqi MZ, Liu Q, Kang MS, Kim MS, Im WT. *Anseongella ginsenosidimutans* gen. nov., sp. nov., isolated from soil cultivating ginseng. *Int J Syst Evol Microbiol* 2016;66:1125x-11130.
41. Prasad S, Manasa BP, Buddhi S, Pratibha MS, Begum Z, Bandi S, Tirunagari P, Shivaji S. *Arcticibacter svalbardensis* gen. nov., sp. nov., of the family *Sphingobacteriaceae* in the phylum *Bacteroidetes*, isolated from Arctic soil. *Int J Syst Evol Microbiol* 2013;63:1627-1632.
42. Pankratov TA, Tindall BJ, Liesack W, Dedysh SN. *Mucilaginibacter paludis* gen. nov., sp. nov. and *Mucilaginibacter gracilis* sp. nov., pectin-, xylan- and laminarin-degrading members of the family *Sphingobacteriaceae* from acidic Sphagnum peat bog. *Int J Syst Evol Microbiol* 2007;57:2349-2354.
43. Asker D, Beppu T, Ueda K. *Nubsella zeaxanthinifaciens* gen. nov., sp. nov., a zeaxanthin-producing bacterium of the family *Sphingobacteriaceae* isolated from freshwater. *Int J Syst Evol Microbiol* 2008;58:601-606.
44. Ntougias S, Fasseas C, Zervakis GI. *Olivibacter sitiensis* gen. nov., sp. nov., isolated from alkaline olive-oil mill wastes in the region of Sitia, Crete. *Int J Syst Evol Microbiol* 2007;57:398-404.
45. Kim MK, Na JR, Cho DH, Soung NK, Yang DC. *Parapedobacter koreensis* gen. nov., sp. nov. *Int J Syst Evol Microbiol* 2007;57:1336-1341.
46. Cai YY, Dong WW, Hu YL, Jiang X, Wang YW, Zhou YQ, Peng N, Liang YX, Zhao SM. *Pararcticibacter amylolyticus* gen. nov., sp. nov., Isolated from a Rotten Hemp Rope, and Reclassification of *Pedobacter tournemirensis* as *Pararcticibacter tournemirensis* comb. nov. *Curr Microbiol* 2020;77:320-326.
47. Xia X, Wu S, Han Y, Liao S, Wang G. *Pelobium manganitolerans* gen. nov., sp. nov., isolated from sludge of a manganese mine. *Int J Syst Evol Microbiol* 2016;66:4954-4959.
48. Vaz-Moreira I, Nobre MF, Nunes OC, Manaia CM. *Pseudosphingobacterium domesticum* gen. nov., sp. nov., isolated from home-made compost. *Int J Syst Evol Microbiol* 2007;57:1535-1538.

462 49. Weon HY, Kim BY, Lee CM, Hong SB, Jeon YA, Koo BS, Kwon SW. *Solitalea*
463 *koreensis* gen. nov., sp. nov. and the reclassification of [*Flexibacter*] canadensis as
464 *Solitalea canadensis* comb. nov. *Int J Syst Evol Microbiol* 2009;59:1969-1975.
465 50. Xu L, Sun JQ, Wang LJ, Gao ZW, Sun LZ, Wu XL. *Sphingobacterium alkalisoli* sp.
466 nov., isolated from a saline-alkaline soil. *Int J Syst Evol Microbiol* 2017;67:1943-
467 1948.
468 51. Daood HG, Biacs AP. Simultaneous determination of Sudan dyes and carotenoids in
469 red pepper and tomato products by HPLC. *J Chromatogr Sci* 2005: 43:461–465
470 52. Bernardet JF, Bowman PJ. (2006). The prokaryotes: a handbook on the biology of
471 bacteria. Proteobacteria: Delta and Epsilon Subclasses. Deeply Rooted Bacteria. 7.
472 481-532.
473 53. Buck JD. Nonstaining (KOH) method for determination of Gram reactions of marine
474 bacteria. *Appl Environ Microbiol* 1982;44:992–993.
475 54. Barrow GI, Feltham RKA. Cowan and Steel's Manual for the Identification of
476 Medical Bacteria, 3rd ed. Cambridge: Cambridge University Press; 2004
477 55. Vallenet, D., Engelen, S., Mornico, D., Cruveiller, S., Fleury, L., Lajus, A., Rouy, Z.,
478 Roche, D., Salvignol, G., Scarpelli, C., & Médigue, C. (2009). MicroScope: a platform
479 for microbial genome annotation and comparative genomics. *Database : the journal of*
480 *biological databases and curation, 2009*, bap021. DOI: 10.1093/database/bap021
481 56. Miller LT. A single derivatization method for bacterial fatty acid methyl esters
482 including hydroxy acids. *J Clin Microbiol* 1982;16:584-586.
483 57. Kuykendall LD, Roy MA, O'Neill JJ, Devine TE. Fatty acids, antibiotic resistance,
484 and deoxyribonucleic acid homology groups of *Bradyrhizobium japonicum*. *Int J Syst*
485 *Evol Microbiol* 1988;38:358-361.
486 58. Tindall BJ. A comparative study of the lipid composition of *Halobacterium*
487 *saccharovorum* from various sources. *Syst Appl Microbiol* 1990a;13:128-130.
488 59. Tindall BJ. Lipid composition of *Halobacterium lacusprofundi*. *FEMS Microbiol.*
489 *Letts* 1990b;66:199-202.
490 60. Tindall BJ, Sikorski J, Smibert RM, Kreig NR. *Phenotypic characterization and the*
491 *principles of comparative systematics.* In Methods for General and Molecular
492 Microbiology 3rd ed. 2007 pp. 330-393.
493 61. Lefort V, Desper R, Gascuel O. FastME 2.0: A comprehensive, accurate, and fast
494 distance-based phylogeny inference program. *Mol Biol Evol*. 2015;32:2798–2800.
495 DOI: 10.1093/molbev/msv150

62. Farris JS. Estimating phylogenetic trees from distance matrices. *Am Nat.* 1972;106: 645–667.

Figures and Tables

Table 1. Glycoside hydrolases coded in *Sphingobacterium hungaricum* Kb22^T genome

GH family	NCBI accession number	Activities in Family
GH2	MBE8713017; MBE8713638; MBE8712575; MBE8715037	β-galactosidase (EC 3.2.1.23) ; β-mannosidase (EC 3.2.1.25); β-glucuronidase (EC 3.2.1.31); α-L-arabinofuranosidase (EC 3.2.1.55); exo-β-glucosaminidase (EC 3.2.1.165); α-L-arabinopyranosidase (EC 3.2.1.-); β-galacturonidase (EC 3.2.1.-); β-xylosidase (EC 3.2.1.37); β-D-galactofuranosidase (EC 3.2.1.146); β-glucosidase (EC 3.2.1.21) and others
GH3	MBE8715078; MBE8713443; MBE8713973	β-glucosidase (EC 3.2.1.21); xylan 1,4-β-xylosidase (EC 3.2.1.37); α-L-arabinofuranosidase (EC 3.2.1.55); glucan 1,4-β-glucosidase (EC 3.2.1.74); β-1,2-glucosidase (EC 3.2.1.-); β-1,3-glucosidase (EC 3.2.1.-); xyloglucan-specific exo-β-1,4-glucanase / exo-xyloglucanase (EC 3.2.1.155) and others
GH5	MBE8712452	endo-β-1,4-glucanase / cellulase (EC 3.2.1.4); endo-β-1,4-xylanase (EC 3.2.1.8); β-glucosidase (EC 3.2.1.21); β-mannosidase (EC 3.2.1.25); glucan β-1,3-glucosidase (EC 3.2.1.58); exo-β-1,4-glucanase / cellodextrinase (EC 3.2.1.74); glucan endo-1,6-β-glucosidase (EC 3.2.1.75); mannan endo-β-1,4-mannosidase (EC 3.2.1.78); cellulose β-1,4-cellobiosidase (EC 3.2.1.91); steryl β-glucosidase (EC 3.2.1.104); endo-β-1,6-galactanase (EC 3.2.1.164); mannan transglycosylase (EC endo-β-1,3-glucanase / laminarinase (EC 3.2.1.39); β-N-acetylhexosaminidase (EC 3.2.1.52); chitosanase (EC 3.2.1.132); β-D-galactofuranosidase (EC 3.2.1.146); β-galactosylceramidase (EC 3.2.1.46); α-L-arabinofuranosidase (EC 3.2.1.55) and others
GH13	MBE8713329; MBE8712407	α-amylase (EC 3.2.1.1); pullulanase (EC 3.2.1.41); cyclomaltodextrinase (EC 3.2.1.54); trehalose-6-phosphate hydrolase (EC 3.2.1.93); α-glucosidase (EC 3.2.1.20); maltotetraose-forming α-amylase (EC 3.2.1.60); isoamylase (EC 3.2.1.68); maltotriose-forming α-amylase (EC 3.2.1.116); trehalose synthase (EC 5.4.99.16); amylo-α-1,6-glucosidase (EC 3.2.1.33); oligosaccharide α-4-glucosyltransferase (EC 2.4.1.161) and others
GH15	MBE8712115	glucoamylase (EC 3.2.1.3); glucodextranase (EC 3.2.1.70); α,α-trehalase (EC 3.2.1.28); dextran dextrinase (EC 2.4.1.2)
GH16	MBE8713572; MBE8714674	xyloglucan:xyloglucosyltransferase (EC 2.4.1.207); endo-1,3-β-glucanase / laminarinase (EC 3.2.1.39); endo-1,3(4)-β-glucanase (EC 3.2.1.6); licheninase (EC 3.2.1.73); β-agarase (EC 3.2.1.81); xyloglucanase (EC 3.2.1.151); endo-β-1,3-galactanase (EC 3.2.1.181); hyaluronidase (EC 3.2.1.35); endo-β-1,4-galactosidase (EC 3.2.1.-); chitin β-1,6-glucanosyltransferase (EC 2.4.1.-); β-glycosidase (EC 3.2.1.-); β-carrageenase (EC 3.2.1.-) and others
GH20	MBE8715038	β-hexosaminidase (EC 3.2.1.52); lacto-N-biosidase (EC 3.2.1.140); β-1,6-N-acetylglucosaminidase (EC 3.2.1.-); β-6-SO3-N-acetylglucosaminidase (EC 3.2.1.-)
GH25	MBE8715245	lysozyme (EC 3.2.1.17)
GH26	MBE8712451; MBE8712453	β-mannanase (EC 3.2.1.78); exo-β-1,4-mannobiohydrolase (EC 3.2.1.100); β-1,3-xylanase (EC 3.2.1.32); lichenase / endo-β-1,3-1,4-glucanase (EC 3.2.1.73); mannobiose-producing exo-β-mannanase (EC 3.2.1.-)
GH31	MBE8712408; MBE8713639	α-glucosidase (EC 3.2.1.20); α-galactosidase (EC 3.2.1.22); α-mannosidase (EC 3.2.1.24); α-1,3-glucosidase (EC 3.2.1.84); α-xylosidase (EC 3.2.1.177); α-glucan lyase (EC 4.2.2.13); isomaltosyltransferase (EC 2.4.1.-); α-N-acetylgalactosaminidase (EC 3.2.1.49) and others
GH32	MBE8713030	invertase (EC 3.2.1.26); endo-inulinase (EC 3.2.1.7); exo-inulinase (EC 3.2.1.80); fructan β-(2,6)-fructosidase/6-exohydrolase (EC 3.2.1.154); sucrose:fructan 6-fructosyltransferase (EC 2.4.1.10); fructan:fructan 6G-fructosyltransferase (EC 2.4.1.243); levan fructosyltransferase (EC 2.4.1.-); cycloinulo-oligosaccharide fructanotransferase (EC 2.4.1.-) and others
GH36	MBE8712331	α-galactosidase (EC 3.2.1.22); α-N-acetylgalactosaminidase (EC 3.2.1.49); stachyose synthase (EC 2.4.1.67); raffinose synthase (EC 2.4.1.82)

GH43	MBE8714090; MBE8715079; MBE8714498; MBE8712801; MBE8712802; MBE8714088; MBE8714089; MBE8715346	β-xylosidase (EC 3.2.1.37); α-L-arabinofuranosidase (EC 3.2.1.55); xylanase (EC 3.2.1.8); α-1,2-L-arabinofuranosidase (EC 3.2.1.-); β-1,3-xylosidase (EC 3.2.1.-); exo-β-1,3-galactanase (EC 3.2.1.145) and others
GH65	MBE8713327; MBE8713430	α,α-trehalase (EC 3.2.1.28); maltose phosphorylase (EC 2.4.1.8); trehalose phosphorylase (EC 2.4.1.64); trehalose-6-phosphate phosphorylase (EC 2.4.1.216) and others
GH76	MBE8713771; MBE8715179; MBE8713770	α-1,6-mannanase (EC 3.2.1.101); α-glucosidase (EC 3.2.1.20)
GH78	MBE8713718; MBE8713719; MBE8713720; MBE8713061; MBE8713026; MBE8715340	α-L-rhamnosidase (EC 3.2.1.40); rhamnogalacturonan α-L-rhamnohydrolase (EC 3.2.1.174); L-Rhap-α-1,3-D-Apif -specific α-1,3-L-rhamnosidase (EC 3.2.1.-)
GH88	MBE8713573	d-4,5-unsaturated β-glucuronyl hydrolase (EC 3.2.1.-)
GH92	MBE8715180; MBE8715337; MBE8713866; MBE8712429; MBE8713427; MBE8712806; MBE8714415; MBE8715336	mannosyl-oligosaccharide α-1,6-mannosidase (EC 3.2.1.-); α-mannosidase (EC 3.2.1.24); α-1,2-mannosidase (EC 3.2.1.-); α-1,3-mannosidase (EC 3.2.1.-); α-1,4-mannosidase (EC 3.2.1.-) and others
GH106	MBE8713721	α-L-rhamnosidase (EC 3.2.1.40); rhamnogalacturonan α-L-rhamnohydrolase (EC 3.2.1.174)
GH125	MBE8714531	exo-α-1,6-mannosidase (EC 3.2.1.-)
GH29/GH95	MBE8713785; MBE8713440; MBE8714269; MBE8715039	α-L-fucosidase (EC 3.2.1.51); α-1,3/1,4-L-fucosidase (EC 3.2.1.111); α-1,2-L-fucosidase (EC 3.2.1.63) α-L-galactosidase (EC 3.2.1.-)

503

Table 2. Genes annotated as xylanase, β-xylosidase, cellulase or β-glucosidase in the genome of Kb22^T and related strains. GenBank accession numbers are GCA_015210005 (*Sphingobacterium hungaricum* Kb22^T), GCA_009829075 (*Sphingobacterium composti* KCTC 12578^T), GCA_005048855 (*Sphingobacterium olei* HAL-9^T) and GCA_900168125 (*Sphingobacterium nematocida* DSM_24091^T). The GH family numbers are in parentheses.

	S. hungaricum Kb22^T	*S. composti* KCTC 12578^T	*S. olei* HAL-9^T	*S. nematocida* DSM_24091^T
isolation source of the strain	soil	compost	soil	leaf tissue
xylanase	MBE8715079(GH43) MBE8714090(GH43)	-	WP_136902756 (GH10) WP_136902760 (GH10)	-
β-xylosidase	MBE8712801(GH43)	-		-
cellulase	-	WP_159637017 (GH5) WP_159637239 (GH5)	WP_136901870 (GH5)	-
β-glucosidase	MBE8715078(GH3) MBE8713443(GH3)	-	WP_136900542 (GH1) WP_136902682 (GH3)	-

Table 3. AAI matrix with Kb22^T and type strains in the *Sphingobacteriaceae* family. Genbank assembly accessions: *Albibacterium bauzanense* DSM 22554^T - GCA_004339765.1; *Anseongella ginsenosidimutans* Gsoil 524^T - GCA_008033235.1; *Arcticibacter svalbardensis* MN12-7^T - GCA_000403135.1; *Mucilaginibacter paludis* DSM 18603^T - GCA_000166195.3; *Nubsella zeaxanthinifaciens* TDMA-5^T - GCA_003313335.1; *Olivibacter sitiensis* DSM 17696^T - GCA_000427965.1; *Parapedobacter koreensis* Jip14^T - GCA_900109365.1; *Pararcticibacter amylolyticus* FJ4-8^T - GCA_003130405.1; *Pedobacter heparinus* DSM 2366^T - GCA_000023825.1; *Pelobium manganitolerans* YS-25^T - GCA_003609575.1; *Pseudosphingobacterium domesticum* DSM 18733^T - GCA_900109575.1; *Solitalea koreensis* DSM 21342^T - GCA_900182575.1; *Sphingobacterium spiritivorum* NCTC 11386^T - GCA_900457435.1.

Sphingobacterium hungaricum Kb22^T	*Albibacterium bauzanense* DSM 22554^T	*Parapedobacter koreensis* Jip14^T	*Sphingobacterium spiritivorum* NCTC 11386^T	*Anseongella ginsenosidimutans* Gsoil 524^T	*Olivibacter sitiensis* DSM 17696^T	*Solitalea koreensis* DSM 21342^T	*Pararcticibacter amylolyticus* FJ4-8^T	*Pelobium manganitolerans* YS-25^T	*Pedobacter heparinus* DSM 2366^T	*Mucilaginibacter paludis* DSM 18603^T	*Arcticibacter svalbardensis* MN12-7^T	*Pseudosphingobacterium domesticum* DSM 18733^T	*Nubsella zeaxanthinifaciens* TDMA-5^T	
100	58	60	66	52	59	54	58	56	57	57	57	60	57	*Sphingobacterium hungaricum* Kb22^T
58	100	59	59	52	59	54	58	56	56	57	57	61	57	*Albibacterium bauzanense* DSM 22554^T
60	59	100	62	53	60	54	58	56	56	57	57	61	57	*Parapedobacter koreensis* Jip14^T
66	59	62	100	52	60	54	59	56	57	57	57	61	57	*Sphingobacterium spiritivorum* NCTC 11386^T
52	52	53	52	100	52	54	53	53	52	53	52	53	53	*Anseongella ginsenosidimutans* Gsoil 524^T
59	59	60	60	52	100	54	58	57	57	57	57	64	57	*Olivibacter sitiensis* DSM 17696^T
54	54	54	54	54	54	100	55	55	55	55	54	55	55	*Solitalea koreensis* DSM 21342^T

58	58	58	59	53	58	55	100	58	58	59	63	60	58	*Pararcticibacter amylolyticus* FJ4-8^{T}
56	56	56	56	53	57	55	58	100	58	57	57	57	58	*Pelobium manganitolerans* YS-25^{T}
57	56	56	57	52	57	55	58	58	100	58	58	57	65	*Pedobacter heparinus* DSM 2366^{T}
57	57	57	57	53	57	55	59	57	58	100	58	58	57	*Mucilaginibacter paludis* DSM 18603^{T}
57	57	57	57	52	57	54	63	57	58	58	100	58	57	*Arcticibacter svalbardensis* MN12-7^{T}
60	61	61	61	53	64	55	60	57	57	58	58	100	57	*Pseudosphingobacterium d omesticum* DSM 18733^{T}
57	57	57	57	53	57	55	58	58	65	57	57	57	100	*Nubsella zeaxanthinifaciens* TDMA-5^{T}

523

524 **Table 4.** Differential phenotypic characteristics between strain Kb22^T and *Sphingobacterium*
525 *composti* T5-12^T. Data are from this study, except G+C content of *Sphingobacterium*
526 *composti* T5-12^T [9]
527 Strains: 1, Kb22^T; 2, *Sphingobacterium composti* T5–12^T

	1	2
Isolation source	soil	compost
Temperature range for growth (°C) (optimum)	10-35 (30)	10-40 (35)
Growth with NaCl (%) (optimum)	0-2 (1)	0.5-3 (0.5)
pH range for growth (optimum)	6.0-8.5 (7.5)	5.5-8.5 (7.0)
G+C content (mol%)	38.1	35.2
API 50CH		
D-arabinose	+	-
D-galactose	-	+
rhamnose	+	-
mannitol	-	+
methyl-α-D-mannopyranoside	+	-
methyl-α-D-glucopyranoside	+	-
N-acetyl glucosamine	+	-
amygdalin	+	-
arbutin	+	-
salicin	+	-
D-cellobiose	+	-
D-maltose	+	-
D-lactose	+	-
D-melibiose	+	-
D-saccharose	+	-
D-trehalose	+	-
melesitose	+	-
D-raffinose	+	-
starch	+	-
gentiobiose	+	-
turanose	+	-
API 20 NE		
hydrolysis of esculin	+	-
assimilation of mannitol	-	+
assimilation of maltose	+	-
API ZYM		
esterase lipase (C8)	+	-
β-glucosidase	+	-
α-chymotrypsin	-	+

528

Table 5. Cellular fatty acid composition of *Sphingobacterium hungaricum* Kb22^T and *Sphingobacterium composti* T5-12^T. Strains: 1, Kb22^T; 2, *Sphingobacterium composti* T5-12^T; tr, trace amount (<1%). Summed Features are fatty acids that cannot be resolved reliably from another fatty acid using the chromatographic conditions chosen. The MIDI system groups these fatty acids together as one feature with a single percentage of the total. The unknown fatty acids have no name listed in the peak library file of the MIDI system. ECL, equivalent chain length.

	1	**2**
iso-$C_{15:0}$	28.4	39.7
iso-$C_{15:0}$ 3-OH	2.7	1.8
iso-$C_{15:1}$ F	-	2.0
iso-$C_{15:1}$ G	5.1	
iso-$C_{17:0}$ 3-OH	19.7	17.6
iso-$C_{17:1}$ ω9*c*	10.1	10.6
summed feature 1 (iso-$C_{15:1}$ H/$C_{13:0}$ 3-OH)	-	1.2
summed feature 3 ($C_{16:1}$ ω7*c*/iso-$C_{15:0}$ 2-OH)	25.7	20.6
summed feature 4 (iso-$C_{17:1}$ I/anteiso-$C_{17:1}$ B)	1.3	-
unknown (ECL 13.565)	1.1	1.1
unknown (ECL 16.582)	1.6	1.3

Fig. 1. MLSA tree showing the relationship of strain Kb22^T to closely related species. Genes for MLSA were selected by autoMLST webserver. The phylogenetic tree based on concatenated genes was inferred using the neighbor-joining method with Kimura's two-parameter calculation model.

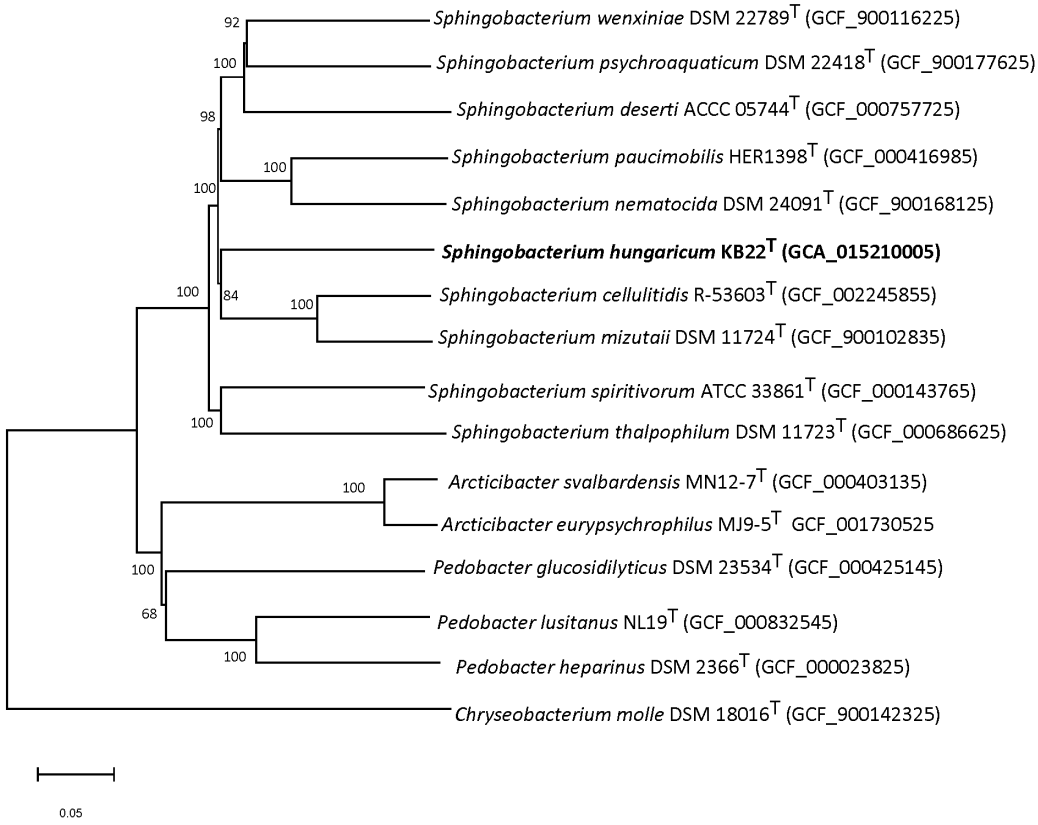
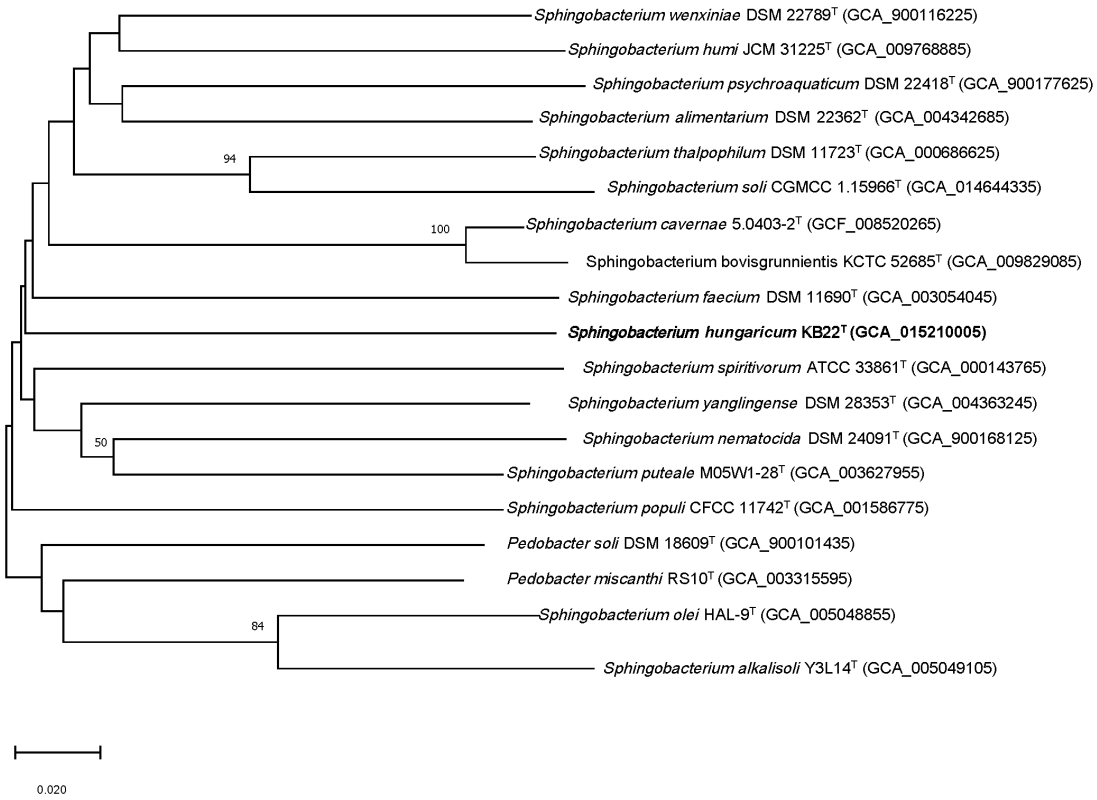


Fig. 2. Tree inferred with FastME 2.1.6.1 [61] from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula $_{\delta}5$. The numbers above branches are GBDP pseudo-bootstrap support values from 100 replications, with an average branch support of 39.9%. The tree was rooted at the midpoint [62]. Assembly accessions are in parentheses.



555 **Table S1.** Genes selected for MLSA by autoMLST

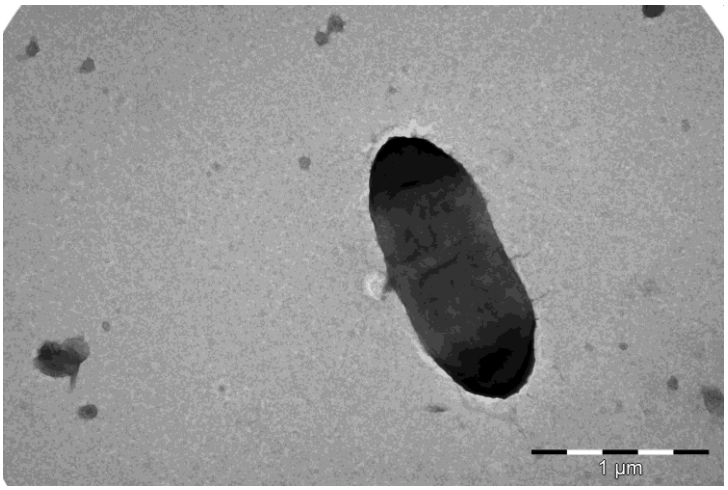
TIGFRAM entries	Gene abbreviation	Biological process	Gene name
TIGR01798	cit_synth_I	Energy metabolism	citrate (Si)-synthase
TIGR00055	uppS	Cell envelope	di-trans,poly-cis-decaprenylcistransferase
TIGR03953	rplD_bact	Protein synthesis	50S ribosomal protein uL4
TIGR00635	ruvB	DNA metabolism	Holliday junction DNA helicase RuvB
TIGR00447	pth	Protein synthesis	aminoacyl-tRNA hydrolase
TIGR00114	lumazine-synth	Biosynthesis of cofactors, prosthetic groups, and carriers	6,7-dimethyl-8-ribityllumazine synthase
TIGR00138	rsmG_gidB	Protein synthesis	16S rRNA (guanine(527)-N(7))-methyltransferase RsmG
TIGR01394	TypA_BipA	Regulatory functions	GTP-binding protein TypA/BipA
TIGR00059	L17	Protein synthesis	ribosomal protein bL17
TIGR00150	T6A_YjeE	Protein synthesis	tRNA threonylcarbamoyl adenosine modification protein YjeE
TIGR01520	FruBisAldo_II_A	Energy metabolism	fructose-bisphosphate aldolase, class II
TIGR00482	TIGR00482	Biosynthesis of cofactors, prosthetic groups, and carriers	nicotinate (nicotinamide) nucleotide adenylyltransferase
TIGR00242	TIGR00242	Regulatory functions	division/cell wall cluster transcriptional repressor MraZ
TIGR00420	trmU	Protein synthesis	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
TIGR00575	dnlj	DNA metabolism	DNA ligase, NAD-dependent
TIGR00422	valS	Protein synthesis	valine--tRNA ligase
TIGR00461	gcvP	Energy metabolism	glycine dehydrogenase
TIGR01169	rplA_bact	Protein synthesis	ribosomal protein uL1
TIGR00928	purB	Purines, pyrimidines, nucleosides, and nucleotides	adenylosuccinate lyase
TIGR00096	TIGR00096	Protein synthesis	16S rRNA (cytidine(1402)-2'-O)-methyltransferase
TIGR01163	rpe	Energy metabolism	ribulose-phosphate 3-epimerase
TIGR00510	lipA	Biosynthesis of cofactors, prosthetic groups, and carriers	lipoyl synthase
TIGR01071	rplO_bact	Protein synthesis	ribosomal protein uL15
TIGR00487	IF-2	Protein synthesis	translation initiation factor IF-2
TIGR03284	thym_sym	Purines, pyrimidines, nucleosides, and nucleotides	thymidylate synthase

TIGR00922	nusG	Transcription	transcription termination/antitermination factor NusG
TIGR00628	ung	DNA metabolism	uracil-DNA glycosylase
TIGR01066	rplM_bact	Protein synthesis	ribosomal protein uL13
TIGR01063	gyrA	DNA metabolism	DNA gyrase, A subunit
TIGR00967	3a0501s007	Protein fate	preprotein translocase, SecY subunit
TIGR01009	rpsC_bact	Protein synthesis	ribosomal protein uS3
TIGR01021	rpsE_bact	Protein synthesis	ribosomal protein uS5
TIGR00043	TIGR00043	Protein synthesis	rRNA maturation RNase YbeY
TIGR00042	TIGR00042	DNA metabolism	non-canonical purine NTP pyrophosphatase, RdgB/HAM1 family
TIGR00445	mraY	Cell envelope	phospho-N-acetylmuramoyl-pentapeptide-transferase
TIGR01051	topA_bact	DNA metabolism	DNA topoisomerase I
TIGR00382	clpX	Protein fate	ATP-dependent Clp protease, ATP-binding subunit ClpX
TIGR02013	rpoB	Transcription	DNA-directed RNA polymerase, beta subunit
TIGR00065	ftsZ	Cellular processes	cell division protein FtsZ
TIGR00064	ftsY	Protein fate	signal recognition particle-docking protein FtsY
TIGR00048	rRNA_mod_RlmN	Protein synthesis	23S rRNA (adenine(2503)-C(2))-methyltransferase
TIGR00165	S18	Protein synthesis	ribosomal protein bS18
TIGR00166	S6	Protein synthesis	ribosomal protein bS6
TIGR02729	Obg_CgtA	Protein synthesis	Obg family GTPase CgtA
TIGR00414	serS	Protein synthesis	serine--tRNA ligase
TIGR00431	TruB	Protein synthesis	tRNA pseudouridine(55) synthase
TIGR00416	sms	DNA metabolism	DNA repair protein RadA
TIGR00228	ruvC	DNA metabolism	crossover junction endodeoxyribonuclease RuvC
TIGR02386	rpoC_TIGR	Transcription	DNA-directed RNA polymerase, beta' subunit
TIGR00212	hemC	Biosynthesis of cofactors, prosthetic groups, and carriers	hydroxymethylbilane synthase
TIGR00755	ksgA	Protein synthesis	ribosomal RNA small subunit methyltransferase A
TIGR00952	S15_bact	Protein synthesis	ribosomal protein uS15
TIGR00436	era	Protein synthesis	GTP-binding protein Era
TIGR01302	IMP_dehydrog	Purines, pyrimidines, nucleosides, and nucleotides	inosine-5'-monophosphate dehydrogenase
TIGR00615	recR	DNA metabolism	recombination protein RecR

TIGR03263	guanyl_kin	Purines, pyrimidines, nucleosides, and nucleotides	guanylate kinase
TIGR00430	Q_tRNA_tgt	Protein synthesis	tRNA-guanine transglycosylase
TIGR01050	rpsS_bact	Protein synthesis	ribosomal protein uS19
TIGR02012	tigrfam_recA	DNA metabolism	protein RecA
TIGR00521	coaBC_dfp	Biosynthesis of cofactors, prosthetic groups, and carriers	phosphopantothenoylcysteine decarboxylase / phosphopantothenate--cysteine ligase
TIGR01059	gyrB	DNA metabolism	DNA gyrase, B subunit
TIGR00337	PyrG	Purines, pyrimidines, nucleosides, and nucleotides	CTP synthase
TIGR03594	GTPase_EngA	Protein synthesis	ribosome-associated GTPase EngA
TIGR00499	lysS_bact	Protein synthesis	lysine--tRNA ligase
TIGR00088	trmD	Protein synthesis	tRNA (guanine(37)-N(1))-methyltransferase
TIGR00086	smpB	Protein synthesis	SsrA-binding protein
TIGR00496	frr	Protein synthesis	ribosome recycling factor
TIGR00479	rumA	Protein synthesis	23S rRNA (uracil-5-)-methyltransferase RumA
TIGR01171	rplB_bact	Protein synthesis	ribosomal protein uL2
TIGR02348	GroEL	Protein fate	chaperonin GroL
TIGR00959	ffh	Protein fate	signal recognition particle protein
TIGR01379	thiL	Biosynthesis of cofactors, prosthetic groups, and carriers	thiamine-phosphate kinase
TIGR00033	aroC	Amino acid biosynthesis	chorismate synthase
TIGR01135	glmS	Central intermediary metabolism	glutamine-fructose-6-phosphate transaminase (isomerizing)
TIGR01357	aroB	Amino acid biosynthesis	3-dehydroquinate synthase
TIGR01032	rplT_bact	Protein synthesis	ribosomal protein bL20
TIGR03631	uS13_bact	Protein synthesis	ribosomal protein uS13
TIGR01011	rpsB_bact	Protein synthesis	ribosomal protein uS2

558 **Fig. S1.** Transmission electron micrograph of strain Kb22^{T}

559

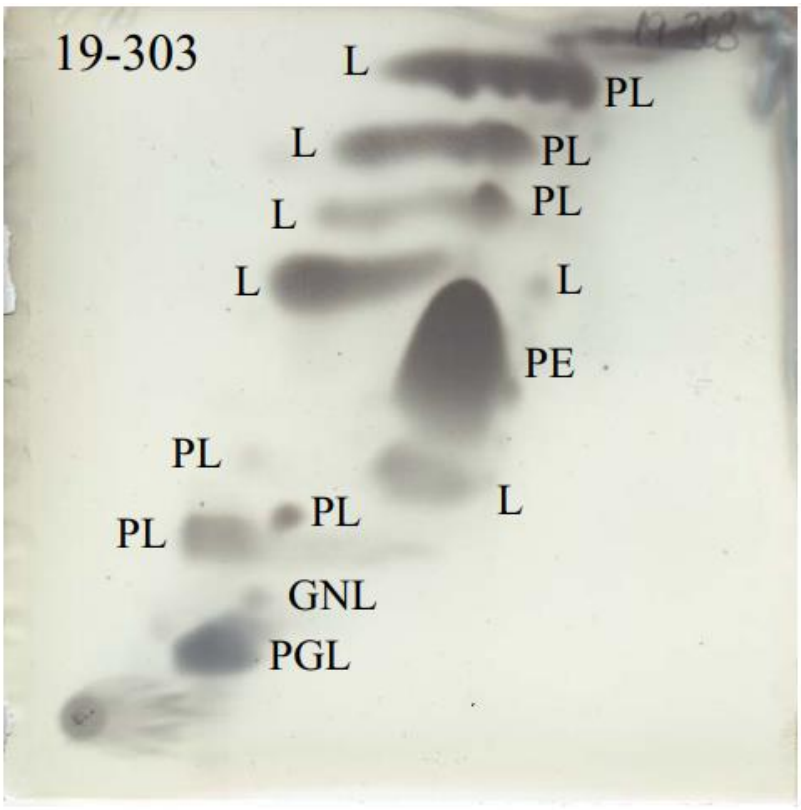


560

561

562 **Fig S2.** Polar lipids of strain Kb22^{T} after two-dimensional thin-layer chromatography (TLC)
563 PE, phosphatidylethanolamine; PGL, phosphoglycolipid; GNL, aminoglycolipid; PL,
564 phospholipid; L, uncharacterised lipid

565



566
567