International Journal of Systematic and Evolutionary Microbiology

Ruthenibacterium lactatiformans gen. nov., sp. nov., a new anaerobic, lactate-producing member of the family Ruminococcaceae isolated from human feces

--Manuscript Draft--

Manuscript Number:	IJSEM-D-16-00053R2
Full Title:	Ruthenibacterium lactatiformans gen. nov., sp. nov., a new anaerobic, lactate-producing member of the family Ruminococcaceae isolated from human feces
Short Title:	Ruthenibacterium lactatiformans gen. nov., sp. nov.
Article Type:	Note
Section/Category:	New taxa - Firmicutes and related organisms
Keywords:	Ruthenibacterium; lactatiformans; Ruminococcaceae; human gut microbiota.
Corresponding Author:	Andrei N Shkoporov, M.D. Pirogov Russian National Research Medical University Moscow, RUSSIAN FEDERATION
First Author:	Andrei N Shkoporov, M.D.
Order of Authors:	Andrei N Shkoporov, M.D.
	Andrei V Chaplin
	Victoria A Shcherbakova, Ph.D.
	Natalia E Suzina, Ph.D.
	Lyudmila I Kafarskaia, Ph.D.
	Vladimir K Bozhenko, Ph.D.
	Boris A Efimov, Efimov
Manuscript Region of Origin:	RUSSIAN FEDERATION
Abstract:	Two novel strains of Gram-negative staining, rod-shaped, obligately anaerobic, non-sporeforming, non-motile bacteria were isolated from the feces of healthy human subjects. The strains designated as 585-1T and 668 are characterized by mesophilic fermentative metabolism, production of D-lactic, succinic, and acetic acids as end products of D-glucose fermentation, prevalence of C18:1ω9, C18:1ω9a, C16:0, and C16:1ω7-cis fatty acids, presence of glycine, glutamic acid, lysine, alanine, and aspartic acid in petidoglycan peptide moiety and lack of respiratory quinones. Whole genome sequencing revealed the DNA G+C content was 56.4-56.6 mol%. Complete 16S rRNA gene sequences shared 91.7/91.6% identity with Anaerofilum pentosovorans FaeT, 91.3/91.2% with Gemmiger formicilis ATCC 27749T, and 88.9/88.8% with Faecalibacterium prausnitzii ATCC 27768T. On the basis of chemotaxonomic and genomic properties it was concluded that the strains represent a new species in a new genus within the family Ruminococcaceae, for which the name Ruthenibacterium lactatiformans gen. nov., sp. nov. is proposed. The type strain is 585-1T (= DSM 100348T, = VKM B-2901T).

Ruthenibacterium lactatiformans gen. nov., sp. nov., a new anaerobic, lactate-producing member of the family *Ruminococcaceae* isolated from human feces

Andrei N. Shkoporov[1]*, Andrei V. Chaplin[1], Victoria A. Shcherbakova[2], Natalia E. Suzina[2], Lyudmila I. Kafarskaia[1], Vladimir K. Bozhenko[3], and Boris A. Efimov[1]

[1] Department of Microbiology and Virology, Pirogov Russian National Research Medical University, Moscow 117997, Russia

[2] Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Pushchino 142290, Russia

[3] Department of Molecular Biology and Experimental Tumor Therapies, Russian Scientific Center of Roentgenoradiology, Moscow 117997, Russia

* Corresponding author. Present address: APC Microbiome Institute, University College Cork, Cork, Ireland. Tel.: +353 21 490 1771. E-mail: andrey.shkoporov@ucc.ie.

Running title: *Ruthenibacterium lactatiformans* gen. nov., sp. nov.

Contents category: New taxa (*Firmicutes*)

The GenBank accession number for the 16S rRNA gene sequence of strains 585-1^T and 668 are KM098109 and KM098110, respectively.

Abstract

Two novel strains of Gram-negative staining, rod-shaped, obligately anaerobic, non-sporeforming, non-motile bacteria were isolated from the feces of healthy human subjects. The strains designated as 585-1^T and 668 are characterized by mesophilic fermentative metabolism, production of D-lactic, succinic, and acetic acids as end products of D-glucose fermentation, prevalence of $C_{18:1}\omega9$, $C_{18:1}\omega9a$, $C_{16:0}$, and $C_{16:1}\omega7$-cis fatty acids, presence of glycine, glutamic acid, lysine, alanine, and aspartic acid in petidoglycan peptide moiety and lack of respiratory quinones. Whole genome sequencing revealed the DNA G+C content was 56.4-56.6 mol%. Complete 16S rRNA gene sequences shared 91.7/91.6% identity with *Anaerofilum pentosovorans* Fae^T, 91.3/91.2% with *Gemmiger formicilis* ATCC 27749^T, and 88.9/88.8% with *Faecalibacterium prausnitzii* ATCC 27768^T. On the basis of chemotaxonomic and genomic properties it was concluded that the strains represent a new species in a new genus within the family *Ruminococcaceae*, for which the name *Ruthenibacterium lactatiformans* gen. nov., sp. nov. is proposed. The type strain is 585-1^T (= DSM 100348^T, = VKM B-2901^T).

The family *Ruminococcaceae* (Rainey, 2009) comprises a morphologically, physiologically, and ecologically divergent group of microorganisms within the order *Clostridiales*, the class *Clostridia* and the phylum *Firmicutes*, and was first described in the 2nd edition of Bergey's Manual of Systematic Bacteriology (De Vos *et al.*, 2009) on the basis of 16S rRNA gene sequence homology as opposed to the phenotypic classification used earlier. Historically, members of this group belonged to 'clostridial clusters' III and IV according to classification introduced by Collins *et al.* (1994).

Despite the recent advances in resolving the 'taxonomic conundrum' within the *Clostridia* class, significant discrepancies still exist between the major DNA databases (Yutin & Galperin, 2013; Lawson & Rainey, 2016). According to the List of Prokaryotic Names with Standing in Nomenclature (http://www.bacterio.net), the family *Ruminococcaceae* comprises 14 genera (*Acetanaerobacterium*, *Acetivibrio*, *Anaerofilum*, *Anaerotruncus*, *Ethanoligenens*, *Faecalibacterium*, *Fastidiosipila*, *Hydrogenoanaerobacterium*, *Oscillibacter*, *Oscillospira*, *Papillibacter*, *Ruminococcus*, *Sporobacter*, and *Subdoligranulum*) of which 12 were included into the original description of the family (Rainey, 2009). By contrast, the Ribosomal Database Project (RDP, http://rdp.cme.msu.edu) lists 21 genera within this family including *Butyricicoccus*, *Saccharofermentans*, *Flavonifractor*, *Pseudoflavonifractor*, *Cellulosibacter* and *Gemmiger*, as well as the provisional taxonomic groups '*Clostridium* cluster III' and '*Clostridium* cluster IV' (Collins et al., 1994), but excluding the genus *Oscillospira* which currently remains uncultured. The NCBI

Taxonomy database (http://www.ncbi.nlm.nih.gov/Taxonomy/) also proposes 21 genera within *Ruminococcaceae* partially overlapping with those in the RDP with the addition of *Caproiciproducens*, *Ercella*, *Mageeibacillus*, *Pseudobacteroides* and *Ruminiclostridium* (the latter taxon supersedes *Clostridium* clusters III and IV as suggested by Yutin & Galperin [2013]), to the exclusion of *Butyricicoccus*, *Cellulosibacter*, *Flavonifractor*, and *Pseudoflavonifractor*. These and other inconsistencies between different databases point out the necessity of further taxonomic improvements within the *Clostridia* class. It is pertinent to note that all bacterial names should be validly published in the International Journal of Systematic and Evolutionary Microbiology as required by the Bacteriological Code. It is only these names that have standing in the literature, which is often not reflected in the names used in DNA databases.

In spite of a high degree of phenotypic divergence, most members of the family *Ruminococcaceae* share a number of common features. The majority of genera within this family comprise strictly-anaerobic bacteria with a Gram-positive type of cell wall, albeit many species actually stain Gram-negative. Metabolism is chemoorganoheterotrophic fermentative with a variety of organic acids and H_2 produced as end products (Rainey, 2009). Some species are capable of anaerobic respiration by utilizing fumarate and sulphur as electron acceptors (van Gelder *et al*., 2014). Morphologically the family is very diverse and includes species with rod-shaped (Zellner *et al*., 1996; Duncan *et al*., 2002), coccoid (Sijpesteijn, 1948), and pleomorhic (Holmstrøm *et al*., 2004) cells, the most notable being a giant (10–40 × 3–6 μm) filamentous septate yet uncultured bacterium *Oscillospira guilliermondii* (Yanagita *et al*., 2003). Some species form spores whilst others are motile by peritrichous flagella (Grech-Mora *et al*., 1996; Zellner *et al*., 1996). A number of genera (e.g. *Ruminococcus*, *Faecalibacterium*, *Anaerotruncus*, *Fastidiosipila*, *Oscillospira*, and *Subdoligranulum*) are associated with human and animal hosts and have been isolated from feces, rumen and intestinal contents, and blood. Other representatives of the family have more diverse isolation sources including wastewater sludge, anaerobic digesters and bioreactors (Rainey, 2009).

One member of the *Ruminococcaceae* family, *Faecalibacterium prausnitzii*, has attracted special attention during the last decade due to its important role in the human gut. Placed in the genus *Fusobacterium* in 1973 (Cato *et al*., 1974), this acetate-consuming and butyrate-producing, extremely oxygen-sensitive, anaerobic organism was re-assigned two decades later to the *Clostridium leptum* group (clostridial cluster IV according to Collins taxonomy). Finally, it was renamed as *F. prausnitzii* in 2002 by Duncan *et al*. This bacterial species constitutes around 5% of total bacterial loads in the fecal samples from healthy adults as determined by metagenomic sequencing (Arumugam *et al*., 2011) and from 2% to 45% according to 16S library sequencing (our unpublished data). Depletion of this organism from the fecal microbiota has been implicated in the pathogenesis of inflammatory bowel disease (IBD; Sokol *et al*., 2009). Moreover, recent studies

have confirmed anti-inflammatory activity of *F. prausnitzii* live cultures, cell supernatant and certain purified components in animal models of IBD (Quévrain et al., 2015; Rossi et al., 2015). Interestingly, despite its extreme air sensitivity *F. prausnitzii* may actually benefit from low oxygen concentrations by using it for NADH regeneration through an extracellular electron shuttle (Khan *et al*., 2012).

During an ongoing culture-based study of human fecal microbiome, in healthy adults and children, two strains of strictly anaerobic Gram stain negative bacteria were isolated that presumably belonged to the family *Ruminococcaceae* but could not be classified to species level using routine identification approaches. Preliminary analysis has shown that the strains designated as 585-1^T and 668 were completely identical by their partial 16S rRNA gene sequences and were moderately related to *F. prausnitzii*, *Subdoligranulum variabile*, *Gemmiger formicilis*, and *Anaerofilum* species. The goal of the current study was to determine the taxonomic position of these strains using polyphasic approach.

Strain 585-1^T was isolated from a stool sample of a 31-year-old healthy Russian male where it was present at a concentration of ~ 1×10^8 c.f.u. g^{-1}. Strain 668 was isolated from the stool of a 5-year-old healthy Russian male child at a concentration of ~ 4×10^8 c.f.u. g^{-1}. Fecal samples were weighed, serially diluted with saline and spread over EG agar plates supplemented with 5% (v/v) defibrinated sheep blood. EG medium base consisted of (per 100 ml): 0.24 g Lab-Lemco powder (Oxoid), 1.0 g Proteose peptone No. 3 (BD-Difco), 0.5 g yeast extract (BD-Difco), 0.4 g Na_2HPO_4, 0.15 g glucose, 0.05 g soluble starch, 0.02 g L-cystine, 1.5 g agar, 0.05 g L-cysteine·HCl·H_2O. Plates were incubated in an atmosphere of 85% N_2, 10% H_2, 5% CO_2 at 37°C for 72 h in anaerobic jars (Schuett-Biotec). Well isolated colonies representative of each morphological type were selected and streaked out several times to obtain pure cultures on EG-blood agar. Upon isolation, strains 585-1^T and 668 were cultured anaerobically on EG broth, PYG broth (Thermo Fisher) or MRS broth (Himedia) supplemented with 5 mg l^{-1} haemin. Cultures were incubated at 37°C for 48-96 h. Strains were preserved by freeze-drying of bacterial suspensions frozen in 10% (w/v) sucrose, 1% (w/v) gelatin solution. Susceptibility of the strains to bile and NaCl was tested in EG broth supplemented with 0-3% (w/v) of Oxgall (Sigma-Aldrich) and 0-8% (w/v) of NaCl. Media were inoculated from fresh agar cultures and growth was examined visually after 48 hours. Physiological properties and enzyme profiles were determined using Vitek 2 ANC, Rapid ID 32A, and API 20A identification systems (bioMérieux) essentially according to manufacturer's instructions except for substitution of API 20A standard incubation medium for glucose-free MRS broth. Carbohydrate fermentation was studied by supplementing MRS broth with 2% (w/v) of every substrate tested instead of D-glucose as well as with 0.01% (w/v) of bromocresol purple. Cultures were incubated anaerobically for 72-96 h.

Analysis of short-chain fatty acids (SCFA) was performed in 168 h PYG broth culture supernatant using HPLC as described before (Shkoporov *et al*., 2015). Alcohols were analyzed using Pye-Unicam 304 gas chromatograph equipped with 2 m x 2 mm glass column packed with Porapak QS matrix (Fluka) in isocratic mode with column, injector, and detector temperatures of 100, 120, and 170°C, respectively. Cellular fatty acids and menaquinone profiles were analyzed in the late exponential phase cultures in MRS broth. Long-chain fatty acids were separated and detected using gas chromatography − mass spectrometry (GC-MS) according to Zhilina *et al*. (2012). Respiratory quinones were detected following the procedure of Collins (1985). For peptidoglycan amino acid analysis cell walls were prepared as described by Schleifer & Kandler, 1972. The cell wall preparations were hydrolysed with 6 M HCl at 105°C for 6 h (Schuman, 2011). Quantitative amino acid analysis was performed with an LC 600 amino acid analyser (Biotronic).

Ultrathin sections (50-60 nm thick) of culture pellets were prepared as described by Duda *et al*. (2009) and examined in a JEOL JEM-1200EX transmission electron microscope with accelerating voltage of 80 kV.

Genomic DNA was extracted and sequenced on Roche/454 GS Junior as described before (Shkoporov *et al*., 2015). Reads were assembled using Newbler v 2.7 into 108 contigs for strain 585-1^T (N50 = 101,736 bp) with a combined length of 4,111,078 bp and 20.0x coverage and 108 contigs for strain 668 (N50 = 114,065 bp) resulting in a combined length of 3,951,525 bp and 17.9x coverage. Draft genome sequences of strains 585-1^T and 668 were deposited in GenBank Nucleotide database under the accession numbers JXXK01000000 and LMUA01000000, respectively. Annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (Angiuoli *et al*., 2008) with metabolic pathways constructed using the KEGG automated annotation server (Moriya *et al*., 2007) followed by manual curation.

The complete 16S rRNA sequences of strains 585-1^T and 668 (KM098109 and KM098110, respectively) and of the type strains from the family *Ruminococcaceae* were aligned using MUSCLE (Edgar 2004).). All positions containing gaps and missing data were eliminated. Phylogenetic inference was performed using the neighbor-joining (NJ) approach in MEGA 6 (Tamura *et al*., 2013) with evolutionary distances calculated using Tamura-Nei substitution model (Tamura & Nei, 1993). The robustness of the tree topology was evaluated by a bootstrapping with 1000 re-samplings (Felsenstein, 1985). Furthermore, in order to check the validity of the NJ tree, maximum likelihood phylogeny was inferred by RAxML (Stamatakis, 2014) using the GTR model with gamma distributed rate heterogeneity and 1000 rapid bootstrap re-samplings. Additionally, phylogeny of strains 585-1^T and 668 was inferred by using core proteome sequences across 23 representative strains of the family *Ruminococcaceae.* To select conserved orthologous proteins encoded by publicly available *Ruminococcaceae* genomes we performed clustering of translated

genomic ORFs using OrthoMCL (Li *et al*. 2003) with an e-value cut-off 1E-5, percent identity cut-off 40% and MCL inflation index I = 1.1. As a result a core proteome of 204 conserved protein families with a single representative encoded by every *Ruminococcaceae* genome from the set was obtained. Amino acid sequences were concatenated and aligned using MUSCLE excluding all gaps. Phylogenetic inference was carried out using Neighbor-Joining in MEGA 6 with the JTT substitution model (Jones *et al*., 1992) and 1000 bootstrap re-samplings. For each reconstructed phylogeny *Clostridium perfringens* ATCC 13124^T was selected as an outgroup.

Average nucleotide identity (ANI) was calculated using 'Blast+'-based algorithm on JSpeciesWS server (Richter et al., 2015).

The novel strains described in this study were obligately anaerobic, non-sporeforming, non-motile, Gram-negative staining rods. Cells collected from 96 h EG blood agar plates were 1.6±0.3×0.4±0.1 μm in size and occurred singly and in pairs. Minute coccoid cells attached or budding from the poles of rods were seen in some light and electron micrographs. Transmission electron micrographs of ultrathin sections of strain 585-1^T revealed highly heterogenous cytoplasm with circular electron dense inclusion bodies (114±13 nm in diameter) and lamellar structures visible in some cells (Fig. 1). Cell envelope organization was elaborate with several electron dense and transparent layers visible, somewhat resembling a Gram-negative trilaminar cell envelope. However, KOH test was negative for both strains. Similar elements were previously seen in the cell envelope of *Gemmiger formicilis* (Gossling & Moore, 1975; Salanitro *et al*., 1976), a member of the family *Ruminococcaceae* moderately related to strains 585-1^T and 668 by 16S rRNA gene sequence. Other related members of the family, *A. agile* and *F. prausnitzii* were shown, however, to have a typical Gram-positive cell wall architecture with thin murein layer, despite their variable staining in Gram method (Zellner *et al*., 1996; Rossi *et al*., 2015). In addition, thin microcapsule was visible on the surface of strain 585-1^T. Flagella, pili, and other types of surface appendages were not detected in negatively stained preparations of strain 585-1^T.

After 96 h of anaerobic growth on EG blood agar colonies reached 0.15-0.4 mm in diameter and were non-haemolytic, colorless, circular, flat, dry, with entire margins and rough surface. Supplementation of EG, MRS, and PYG broth media with 0.5% (w/v) maltose and 5 mg l^{-1} haemin strongly stimulated growth of both strains which otherwise was poor even after prolonged incubation. The best overall growth support of the strains was obtained with MRS broth supplemented with maltose and haemin. The strains were not only resistant to 3% (w/v) of Oxgall in EG broth, but also demonstrated enhanced growth in the form of threadlike, ropy and mucous sediment. Both strains were resistant to up to 1% (w/v) NaCl in MRS broth. Using both MRS and EG broth growth was observed at 37°C but not at 32° or 42°C. Aesculine and starch were

hydrolized by both strains. Indole, hydrogen sulphide, catalase, urease, and gelatinase were not produced. Oxidase, nitrate reductase, and alkaline phosphatase reactions were also negative.

In Rapid ID 32A and Vitek 2 ANC identification panels, based on the use of chromogenic enzyme substrates, both strains demonstrated positive reactions for a number of glycosyl hydrolases, inlcuding α- and β-galactidases, α- and β-glucosidases, β-glucuronidase, α-mannosidase, β-N-acetyl-glucosaminidase, but not for α-arabinosidase, α-L-fucosidase, β-D-fucosidase, and β-mannosidase (Tables S1, S2). By contrast, all carbohydrate fermentation reactions included in these panels were negative. All chromogenic arylamidase (exopeptidase) tests included in the Rapid ID 32A and Vitek 2 ANC panels were negative for both strains, indicating that these microorganisms specialize in the utilization of carbohydrate rather than protein substrates as carbon source. Both strains were negative for most of the carbohydrate fermentation reactions in API 20A tests. Acid production was detected from maltose, salicin, and weakly from D-glucose and L-rhamnose (Table S3). In conventional carbohydrate fermentation tests acid production was detected from maltose, salicin, D-galactose, L-rhamnose, weakly from D-mannose, melibiose and D-sorbitol. Variable results were obtained with sucrose (Table S4). Growth on MRS without carbohydrates was very poor.

In disk-diffusion experiments strain 585-1^T was resistant to amikacin, ampicillin, azithromycin, cephalothin, clindamycin, levofloxacin, linezolid, and penicillin G, but sensitive to amoxyclav and vancomycin (Table S5).

When grown in PYG broth with 0.5% (w/v) glucose strains 585-1^T/668 produced 22.2/24.4 mM of D-lactate and 9.7/9.4 of mM succinate. In addition, strain 668 produced 7.2 mM of acetate. Supplementation of PYG with 0.5% (w/v) maltose led to increased production of succinic acid (17.2 mM) and formation of formic (11.0 mM) and acetic (6.9 mM) acids by strain 585-1^T. By contrast, growth of strains 585-1^T/668 on MRS broth supplemented with 0.5% (w/v) maltose resulted in a wide range fermentation end products including 8.5/8.1 mM of formate, 13.1/10.1 mM of acetate, 20.1/19.1 mM of D-lactate, 15.9/12.2 mM of succinate, 9.6/9.3 mM of propionate and 2.8/3.4 mM of butyrate. The overall composition of metabolic end-products in strains 585-1^T and 668 resembles those of other members of the family *Ruminococcaceae* isolated from human and animal feces. The striking predominance of lactic acid among the end-products relates this group of strains with the genus *Anaerofilum* (Zellner *et al*., 1996). However, unlike *Anaerofilum,* the novel strains produced D-lactate isomer.

Cellular fatty acids (CFA) in strain 585-1^T were mostly composed of monounsaturated species: $C_{18:1}\omega9$ (31.4-31.9%), $C_{18:1}\omega9$ aldehyde (20.6-21.0%), $C_{16:1}\omega7$-cis (5.3-6.0%). Palmitic acid (C16:0, 6.1-6.6%) and other saturated acids were present in minor amounts (Table S6). According to literature, cell membranes of the closest related genera are composed mostly of

saturated CFA, e.g.: $C_{14:0}$, $C_{16:0}$ and $C_{16:0}$, $C_{18:0}$, predominate in *Faecalibacterium* and *Subdoligranulum*, respectively (Jantzen & Hofstad, 1981; Holmstrøm *et al*., 2004), while iso-$C_{16:0}$, iso-$C_{12:0}$, anteiso-$C_{17:0}$ are found in *Ethanoligenens* (Xing *et al*., 2006). Species of *Ruminococcus* are quite diverse in their CFA composition, but $C_{14:0}$, $C_{15:0}$, $C_{16:0}$, iso-$C_{15:0}$, iso-$C_{16:0}$, anteiso-$C_{17:0}$, and anteiso-$C_{19:0}$ are the most common (Minato *et al*., 1988). Other genera of *Ruminococcaceae* that are closely related to strains 585-1^T and 668 have not been characterized yet in terms of CFA composition. Respiratory quinones were not detected in whole cell extracts from strains 585-1^T and 668. Amino acid composition of cell wall acid hydrolysates of the two strains was roughly identical and contained glycine (31.85- 33.99 nM), glutamic acid (31.46-33.75 nM), lysine (23.77-28.23 nM), alanine (23.36-28.87 nM), and aspartic acid (21.92-25.2 nM). The exact structure of of peptide moiety has to be determined. However, such amino acid composition is consistent with peptidoglycan type A4α (cross-linkage by L-Lys-D-Asp) according to Schleifer & Kandler (1972) nomenclature or type A11.31 according to Schumann (2011). Similarly to *Anaerofilum* the lactyl group of muramic acid is likely to be esterified with glycine.

The 4.11 Mbp draft genomic assembly of strain 585-1^T had an overall G+C content of 56.5 mol%, a total of 3,802 protein-coding genes, 51 tRNA, a set of rRNA genes and 1 CRISPR array. The combined length of strain 668 draft assembly was slightly smaller, 3.95 Mbp, with a G+C content of 56.6 mol% , encoding a total of 3,556 genes, 50 tRNA, a set of rRNA genes and 1 CRISPR array. Central carbon metabolism genes in both strains include a full complement for Embden–Meyerhof–Parnas and pentose phosphate pathways, an almost complete set of genes for Entner–Doudoroff pathway (excluding glucose-6-phosphate 1-dehydrogenase), a pyruvate ferredoxin oxidoreductase gene, and genes required for first carbon oxidation in citrate cycle. Therefore, the metabolic capabilities of the new strains presumably differ from those of *F. prausnitzii* and *Ruminococcus bromii* (which lack identifiable transaldolase gene) and *Ethanoligenens harbinense* (which lacks Entner–Doudoroff pathway genes, but possesses an almost complete TCA cycle). Unlike *F. prausnitzii* the novel strains do not have any recognizable genes for butyrate production from acetate including, acetyl-CoA acetyltransferase (thiolase), 3-hydroxyacyl-(β-hydroxybutyryl-) CoA dehydrogenase, 3-hydroxybutyryl-CoA dehydratase (crotonase), butyryl-CoA dehydrogenase, and butyryl-CoA:acetate CoA-transferase (Louis & Flint, 2009). No butyrate kinase genes were detected as well. Furthermore, unlike *F. prausnitzii* the new strains lack succinate dehydrogenase genes, but possess genes coding for lactate dehydrogenase, and aldehyde dehydrogenase. The presence of alcohol dehydrogenase genes can enable the new strains to produce ethanol, which, however, could not be detected in broth cultures under conditions used in this study.

The repertoire of glycan degradation enzymes encoded by genomes of strains 585-1^T and 668 include α- and β-galactidases, hexosaminidases, α- and β-mannosidases, β-glucuronidase

glucosylceramidase, α-L-fucosidase, α-amylase (only in 585-1^T) and sialidase (only in 668). Analysis of amino acid biosynthesis genes suggest that the strains are able to synthesize from common precursors most amino acids with the exception of tryptophan, tyrosine, alanine, arginine, and lysine. The strains appeared to be auxotrophic for most vitamins and enzyme cofactors, however an almost complete pathway of anaerobic transformation of uroporphyrinogen III into cobamide coenzyme (Roper *et al*., 2000) was found in both genomes.

A number of ABC-transporters were identified in the genomes of the new strains, which include predicted transport systems for spermidine/putrescine, raffinose/melibiose, methyl-galactoside, phosphate, phosphonate, branched-chain amino acid, oligopeptide, iron complex, zinc (only in 585-1^T), cobalt, nickel, and biotin. The genomes of strains 585-1^T and 668 encode proteins required for RecF homologous recombination pathway. In addition, strain 585-1^T, but not 668, possess several genes coding for β-lactamases, which correlates well with resistance of the former strain to penicillin G, ampicillin, and cephalothin.

The ecological distribution of the new bacterium across the human population and across various sites in human body has to be established. However, a brief search in NCBI 'nr' database revealed that numerous 16S rRNA gene sequences from uncultured bacteria with ≥97% identity to strain 585-1^T are present in 16S rRNA gene datasets from a number of studies, including a study of a Chinese family fecal microbiota (Li *et al*., 2008), study of ileal microbiota in patients with Crohn's disease (Li *et al*., 2012), and a study of gut microbiota in obese subjects (Ley *et al*., 2006). We also conducted a blastn search against an inhouse dataset of 16S rRNA gene sequences (V1-V3 region), which was generated using 454 platform from fecal samples of 19 healthy human subjects including the two subjects from whom strains 585-1^T and 668 were originally isolated. This search revealed the presence of sequences with ≥97% identity in 7 samples (36.8%) with relative abundance ranging from 0.01% to 0.4% of the total number of sequences.

Alignment of complete 16S rRNA genes from 585-1^T and 668 showed that they were 99.9% identical. The strains also had 91.7/91.6% identity to *Anaerofilum pentosovorans* Fae^T, 91.6/91.5% to *A. agile* F^T, 91.3/91.2% to *G. formicilis* ATCC 27749^T, 90.1% to *S. variabile* BI 114^T, and 88.9/88.8% identity to *F. prausnitzii* ATCC 27768^T. A search in NCBI 'wgs' database using 585-1^T 16S rRNA gene sequence revealed several highly identical (>99%) genes on gut shotgun metagenomic contigs (LBCJ01000006, LBCI01000027), as well as a 16S rRNA gene from unpublished draft assembly of human gut isolate '*Ruminococcaceae* bacterium cv2' (NZ_CYPT01000000). This 4.26 Mbp genomic sequence with a G+C content of 56.6 mol% has been included in the phylogenetic analysis described herein.

To establish the taxonomic positions of strains 585-1^T and 668 within the family *Ruminococcaceae* two different approaches of phylogenetic analyses were carried out. One was

based on 16S rRNA gene sequences from the type strains of family *Ruminococcaceae* and was conducted using both neighbor-joining and maximum likelihood inference methods (Fig. 2). The other was performed on a concatenated alignment of 204 conserved orthologous proteins (Table S7) encoded by the currently available complete and draft genomic sequences from the family *Ruminococcaceae* (Fig. 3) using the NJ approach. Both approaches reliably placed strains 585-1^T, 668 and cv2 inside the family *Ruminococcaceae* According to 16S rRNA phylogeny, the three strains formed a separate branch which was located as a sister clade to the *Gemmiger*/*Subdoligranulum*/*Faecalibacterium* clade. These two clades along with the genus *Anaerofilum* clade together were a part of a larger phylogenetic cluster, which also included *R. bromii*, [*Clostridium*] *leptum*, [*Clostridium*] *sporosphaeroides* and which corresponded to clostridial cluster IV in Collins's taxonomy. The conserved proteins tree had a similar topology and placed the three strains as a separate branch again as the sister clade to the *Subdoligranulum*/*Faecalibacterium* clade. As an additional phylogenetic measure ANI was calculated after pairwise blastn all-versus-all searches between 585-1^T, 668, cv2, *F. prausnitzii* A2-165^T and *S. variabile* DSM 15176^T genomes. The ANI between the strains 585-1^T, 668, cv2 ranged from 97.4 to 98.1%. *F. prausnitzii* A2-165^T had 69.3-69.8% ANI to the new strains and 72.6% to *S. variabile* DSM 15176^T. *S. variabile* DSM 15176^T in turn displayed 68.8-69.5% identity to strains 585-1^T, 668, cv2 and 72.8% to *F. prausnitzii* A2-165T.

Strains 585-1^T and 668 differ from phylogenetically related genera within the family *Ruminococcaceae* by cell morphology, physiological culture properties, enzymatic activity, spectrum of metabolic end-products, CFA composition, and genome characteristics. Based on the phenotypic and genotypic properties of strains 585-1^T and 668 it is concluded that they represent a new species in a new genus within the family *Ruminococcaceae*, for which the name *Ruthenibacterium lactatiformans* gen. nov., sp. nov is proposed. The main chemotaxonomic characteristics of the new taxon in comparison with some of the related genera are given in Table 1.

Description of *Ruthenibacterium* gen. nov.

Ruthenibacterium (Ru.the.ni.bac.te'ri.um. M.L. fem. n. *Ruthenia* medieval Latin name of Russia; Gr. dim. n. *bakterion* a small rod; N.L. neut. n. *Ruthenibacterium* a rod-shaped bacterium isolated in Russia).

Cells are Gram-negative, rod-shaped, obligately anaerobic, non-sporeforming, non-motile, 1.6±0.3×0.4±0.1 μm in size and occur singly and in pairs. Metabolism is chemoorganoheterotrophic fermentative. Optimal growth temperature is 37°C. D-lactate and succinate are the major end-products of fermentation. Predominant cellular fatty acids are $C1_{8:1}\omega9$, $C1_{8:1}\omega9a$, $C_{16:0}$, and $C_{16:1}\omega7$-cis. The peptidoglycan contains glycine, glutamic acid, lysine, alanine, and aspartic acid.

Menaquinones are not produced. Member of the family *Ruminococcaceae*. The type species is *Ruthenibacterium lactatiformans*.

Description of *Ruthenibacterium lactatiformans* sp. nov.

Ruthenibacterium lactatiformans (lac.ta.ti.for'mans. L. part. perf. pass. masc. *lactatus* fed with milk, lactate; L. part. adj. *formans* forming; N.L. part. adj. *lactatiformans* lactate forming).

Exhibits the following characteristics in addition to those given in the description of the genus. Growth on EG blood agar are visible after 72-96 h of anaerobic incubation at 37°C. Colonies are 0.15-0.4 mm in diameter, non-haemolytic, colorless, circular, entire, flat, dry, and with rough surface. Colonies are positive for aesculin and starch hydrolysis and tolerant to bile. In broth cultures growth is stimulated by 0.5% (w/v) maltose, 5 mg l^{-1} haemin, and 2-3% (w/v) of Oxgall. Indole, catalase, and urease are not produced. Gelatin is not digested. Acid is produced from D-glucose, D-galactose, maltose, salicin, L-rhamnose but not from L-arabinose, adonitol, lactose, D-mannitol, D-raffinose, and D-trehalose. In chromogenic substrates tests positive reactions are obtained for α- and β-galactidases, α- and β-glucosidases, β-glucuronidase, β-N-acetyl-glucosaminidase, but not for α-arabinosidase and α-fucosidase. The DNA G+C content is 56.4-56.6 mol%.

The type strain of the species, isolated from human feces, is 585-1^T (= DSM 100348^T, = VKM B-2901^T)

Acknowledgements

Authors thank Dr. Angela McCann, APC Microbiome Institute, University College Cork, for her kind help in editing the manuscript.

References

Angiuoli, S.V., Gussman, A., Klimke, W., Cochrane, G., Field, D., Garrity, G., Kodira, C.D., Kyrpides, N., Madupu, R. & other authors (2008). Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. *OMICS* **12**, 137-141.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T. & other authors (2011). Enterotypes of the human gut microbiome. *Nature* **473**, 174-180.

Cato, E.P., Salmon, C.W. & Moore, W.E.C. (1974). *Fusobacterium prausnitzii* (Hauduroy *et al*.) Moore and Holdeman: emended description and designation of neotype strain. *Int J Syst Bacteriol* **24,** 225-229.

Collins, M.D. (1985). Analysis of isoprenoid quinones. In *Methods in microbiology*, vol. 18, pp. 329–366. Edited by G. Gottschalk. New York: Academic Press.

Collins, M.D., Lawson, P.A., Willems, A., Cordoba, J.J., Fernandez-Garayzabal, J., Garcia, P., Cai, J., Hippe, H. & Farrow, J.A. (1994). The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol* **44,** 812-826.

De Vos, P., Garrity, G.M., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K-H. & Whitman, W.B. (2009). Bergey's Manual of Systematic Bacteriology, 2nd edn., Vol. 3: The Firmicutes. New York, NY: Springer-Verlag.

Duda, V.I., Suzina, N.E., Esikova, T.Z., Akimov, V.N., Oleinikov, R.R., Polivtseva, V.N., Abashina, T.N., Shorokhova, A.P. & Boronin A.M. (2009). A cytological characterization of the parasitic action of ultramicrobacteria NF1 and NF3 of the genus *Kaistia* on chemoorganotrophic and phototrophic bacteria. *FEMS Microbiol Ecol* **69,** 180-193.

Duncan, S.H., Hold, G.L., Harmsen, H.J., Stewart C.S. & Flint, H.J. (2002). Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* **52,** 2141–2146.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791.

Gossling, J. & Moore W.E.C. (1975). *Gemmiger formicilis* n. gen., n. sp., an anaerobic budding bacterium from intestines. *Int. J. Syst. Bacteriol.* **25,** 202-207.

Grech-Mora, I., Fardeau, M.L., Patel, B.K.C., Ollivier, B., Rimbault, A., Prensier, G., Garcia, J.L. & Garnier-Sillam, E. (1996). Isolation and characterization of *Sporobacter termitidis* gen nov sp. nov., from the digestive tract of the wood-feeding termite *Nasutitermes lujae*. *Int J Syst. Bacteriol* **46,** 512–518.

Holdeman, L.V., & Moore, W.E.C., (ed.). (1973). Anaerobe laboratory manual, 2nd edn. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Holmstrøm, K., Collins, M.D., Moller, T., Falsen E. & Lawson., P.A. (2004). *Subdoligranulum variabile* gen. nov., sp. nov. from human feces. *Anaerobe* **10,** 197–203.

Jantzen, E. & Hofstad, T. (1981). Fatty acids of *Fusobacterium* species: taxonomic implications. *J. Gen. Microbiol.* 123, 163-171.

Jones, D.T., Taylor, W.R. & Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8,** 275–282.

Khan, M.T., Duncan, S.H., Stams, A.J., van Dijl, J.M., Flint, H.J. & Harmsen, H.J., (2012). The gut anaerobe *Faecalibacterium prausnitzii* uses an extracellular electron shuttle to grow at oxic-anoxic interphases. *ISME J* **6,** 1578-1585.

Lawson, P. A. & Rainey, F. A. (2016). Proposal to restrict the genus *Clostridium* (Prazmowski) to *Clostridium butyricum* and related species. *Int J Syst Evol Microbiol* **66**, 1009–1016.

Ley, R.E., Turnbaugh, P.J., Klein, S. & Gordon, J.I. (2006). Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023.

Li, E., Hamm, C.M., Gulati, A.S., Sartor, R.B., Chen, H., Wu, X., Zhang, T., Rohlf, F.J., Zhu, W. & other authors (2012). Inflammatory bowel diseases phenotype, *C. difficile* and NOD2 genotype are associated with shifts in human ileum associated microbial composition. *PLoS One* **7**, e26284.

Li, L., Stoeckert, C.J. & Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research* **13,** 2178–2189.

Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., Zhang, Y., Shen, J., Pang, X. & other authors (2008). Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci USA* **105**, 2117-2122.

Louis, P. & Flint, H. J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* **294**, 1–8.

Minato, H., Ishibashi, S. & Hamaoka, T. (1988). Cellular fatty acid and sugar composition of representative strains of rumen bacteria. *J Gen Appl Microbiol* **34**, 303-319.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C. & Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35,** W182-5.

Quévrain, E., Maubert, M.A., Michon, C., Chain, F., Marquant, R., Tailhades, J., Miquel, S., Carlier, L., Bermúdez-Humarán, L.G., Pigneur, B. & other authors (2015). Identification of an anti-inflammatory protein from *Faecalibacterium prausnitzii*, a commensal bacterium deficient in Crohn's disease. *Gut*. In press.

Rainey, F.A. (2009). *Ruminococcaceae* fam. nov. In the *Bergey's Manual of Systematic Bacteriology*, 2nd edn., Vol. 3. Edited by P. Vos *et al.* New York, NY: Springer-Verlag.

Richter, M., Rosselló-Móra, R., Oliver Glöckner, F. & Peplies, J. (2015). JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics*. In press.

Roper, J.M., Raux, E., Brindley, A.A., Schubert, H.L., Gharbia, S.E., Shah, H.N. & Warren, M.J. (2000). The enigma of cobalamin (Vitamin B12) biosynthesis in *Porphyromonas gingivalis*. Identification and characterization of a functional corrin pathway. *J Biol Chem* **275**, 40316-40323.

Rossi, O., Khan, M.T., Schwarzer, M., Hudcovic, T., Srutkova, D., Duncan, S.H., Stolte, E.H., Kozakova, H., Flint, H.J., Samsom, J.N., Harmsen, H.J. & Wells, J.M. (2015). *Faecalibacterium prausnitzii* strain HTF-F and its extracellular polymeric matrix attenuate clinical parameters in DSS-induced colitis. *PLoS One* **10,** e0123013.

Salanitro, J.P., Muirhead, P.A. & Goodman, J.R. (1976). Morphological and physiological characteristics of *Gemmiger formicilis* isolated from chicken ceca. *Appl Environ Microbiol* **32**, 623-632.

Schleifer, K.H. & Kandler, O. (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* **36**, 407-477.

Schumann, P. (2011). Peptidoglycan structure. In *Taxonomy of Prokaryotes, Methods in Microbiology*, Vol. 38, pp. 101-129. Edited by F. Rainey & A. Oren. London: Academic Press.

Shkoporov, A.N., Chaplin, A.V., Khokhlova, E.V., Shcherbakova, V.A., Motuzova, O.V., Bozhenko, V.K., Kafarskaia, L.I. & Efimov, B.A. (2015). Description of *Alistipes inops* sp. nov. and *Coprobacter secundus* sp. nov. isolated from human feces. *Int J Syst Evol Microbiol* **65**, 4580-4588.

Sijpesteijn, A.K. (1948). Cellulose-decomposing bacteria from the rumen of cattle. Leiden University, Leiden: Eduard Ijdo N.V.

Sokol, H., Seksik, P., Furet, J.P., Firmesse, O., Nion-Larmurier, I., Beaugerie, L., Cosnes, J., Corthier, G., Marteau, P. & Doré, J. (2009). Low counts of Faecalibacterium prausnitzii in colitis microbiota. *Inflamm Bowel Dis* **15,** 1183-1189.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313.

Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**, 512-526.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* **30**, 2725-2729.

van Gelder, A.H., Sousa, D.Z., Rijpstra, W.I., Damsté, J.S., Stams, A.J. & Sánchez-Andrea, I. (2014). *Ercella succinigenes* gen. nov., sp. nov., an anaerobic succinate-producing bacterium. *Int J Syst Evol Microbiol* **64,** 2449-2454.

Wozny, M.A., Bryant M.P., Holdeman, L.V. and Moore, W.E.C. (1977). Urease assay and urease-producing species of anaerobes in the bovine rumen and human feces. *Appl Environ Microbiol* **33**, 1097-1104.

Yanagita, K., Manome, A., Meng, X.Y., Hanada, S., Kanagawa, T., Tsuchida, T., Mackie, R.I. & Kamagata, Y. (2003). Flow cytometric sorting, phylogenetic analysis and in situ detection of *Oscillospira guilliermondii*, a large, morphologically conspicuous but uncultured ruminal bacterium. *Int J Syst Evol Microbiol* **53,** 1609–1614.

Yutin, N. & Galperin, M.Y. (2013). A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* **15,** 2631-2641.

Zellner, G., Stackebrandt, E., Nagel, D., Messner, P., Weiss N. & Winter, J. (1996). *Anaerofilum pentosovorans* gen. nov., sp. nov., and *Anaerofilum agile* sp. nov., two new, strictly anaerobic, mesophilic, acidogenic bacteria from anaerobic bioreactors. *Int J Syst Bacteriol* **46,** 871–875.

Zhilina, T.N., Zavarzina, D.G., Panteleeva, A.N., Osipov, G.A., Kostrikina, N.A., Tourova, T.P. & Zavarzin, G.A. (2012). *Fuchsiella alkaliacetigena* gen. nov., sp. nov., an alkaliphilic, lithoautotrophic homoacetogen from a soda lake. *Int J Syst Evol Microbiol* **62**, 1666-1673.

Table 1. Differential characteristics of strain 585-1^T and 668 in comparison with some related genera. Data for other genera within the family *Ruminococcaceae* are taken from Cato *et al*. (1974), Duncan *et al*. (2002), Jantzen & Hofstad (1981), Holmstrøm *et al*. (2004), Gossling & Moore (1975), Zellner *et al*. (1996), Xing *et al*. (2006), Wozny *et al*. (1977), Minato *et al*. (1988) and Rainey (2009). NA, data not available; V, variable; W, weak; a/b/f/ib/iv/l/p/s/v/e/bdl;, fermentation end products (acetic, butyric, fumaric, isobutyric, isovaleric, lactic, propionic, succinic and valeric acids, ethanol, and 2,3-butanediol, respectively; capital and small letters indicate major and minor products, respectively).

	Strains 585-1 and 668	***Faecalibacterium***	***Subdoligranulum***	***Gemmiger***	***Anaerofilum***	***Ethanoligenens***	***Anaerotruncus***	***Ruminococcus***
Isolation source	Human feces	Human and animal feces	Human feces	Human feces and chicken ceca	Anaerobic sewage sludge	Anaerobic sewage sludge	Human feces and blood	Rumen, large bowel, or cecum of many animals and humans
Cell shape	Straight rods (occur singly and in pairs	Variable length straight rods (occur singly)	Coccoid and pleomorphic	Spherical to drop-like (often in pairs and short chains)	Thin straight rods (single and in pairs)	Rods and filaments	Thin rods	Cocci and coccobacilli (often in pairs and chains)
Cell size, µm	1.6±0.3×0.4±0.1	0.5–0.8 × 2.0–14.0	0.6–2.5	1–2.3 × 0.5–1.2	3.0–6.0 × 0.3–0.6	0.4–0.8 × 1.5–8.0	2–5 × 0.5	0.3–1.5 × 0.7–1.8
Gram stain result	-	-	-	-	+	+	+	V
Spores	-	-	-	-	-	-	+	-
Motility	-	-	-	-	+	+	-	-
Growth temperature, °C	37	37–45	37–45	37-45	18–44	20–44	36–40	37–42
Aesculin hydrolysis	+	+	+	+	V	+	-	V
Starch hydrolysis	+	V	-	V	-	W	-	V
Indole production	-	-	-	-	NA	+	+	-
Urease	-	-	-	-	NA	+	-	+
Optimal growth medium	EG, MRS	M2GSC, YCFA, Wilkins-Chalgrene broth (Oxoid)	M2GSC, Fastidious Anaerobe Broth (Oxoid)	E medium*	DSMZ medium 119	PYG broth	Brucella blood agar (Anaerobe Systems)	Rumen fluid agar
Major end products	L, S, a (PYG)	B, L, F, s†, p† (PYG-RF)	B, L, a, s	F, B, l, a (PYG)	L, A, E, F, bdl, CO_2	A, E, H_2, CO_2	A, B (PYG)	A, F, S, L, e, H_2 (PYG)
Major cellular fatty acids	$C_{18:1}\omega9$, $C_{18:1}\omega9a$, $C_{16:0}$, $C_{16:1}\omega7$-cis	$C_{14:0}$, $C_{16:0}$	$C_{16:0}$, $C_{18:0}$, $C_{18:1}\omega9$-cis	NA	NA	iso-$C_{16:0}$, iso-$C_{12:0}$, anteiso-$C_{17:0}$	NA	$C_{14:0}$, $C_{15:0}$, $C_{16:0}$, iso-$C_{15:0}$, iso-$C_{16:0}$, ai-$C_{17:0}$ ai-$C_{19:0}$
G+C content, mol%	56.4-56.6‡	47-67 (56.2-57.7‡)	52.2 (57.9‡)	59	54-55	47.8-49.0 (55.6‡)	53-54 (54.2‡)	37-47 (41.1-53.4‡)
Genome size, Mb	3.9-4.3§	2.9-3.3	3.25	NA	NA	3.0	3.7	2.2-4.6

* Holdeman & Moore, 1973
† trace amounts
‡ based on draft and complete genome assemblies
§ includes a closely related genome of strain '*Ruminococcaceae* bacterium cv2' (NZ_CYPT01000000)

Figure legends.

Figure 1. Transmission electron micrographs of ultrathin sections of strain 585-1^{T} cells showing overall cell morphology (a), spherical inclusion bodies (b), laminate structures (c), and cell wall organization (d). C, microcapsule; CW, multilayer cell wall; IB, inclusion bodies; IM, intracytoplasmic membrane structures; LB, lamellar bodies; MC, minute cells; N, nucleoid.

Figure 2. Neighbor-joining phylogenetic tree of 16S rRNA gene sequences. Evolutionary distances were computed using the Tamura-Nei substitution matrix. The scale bar represents 0.02 substitutions per nucleotide position. Accession number in Genbank database is given next to a strain name. Node labels represent bootstrap confidence levels obtained using Neighbor-joining/Maximum likelihood methods. Only the nodes with both bootstrap levels higher than 70% are labeled.

Figure 3. Neighbor-joining phylogenetic tree of concatenated sequences of 204 conserved proteins. Evolutionary distances were computed using the JTT matrix. The scale bar represents 0.1 substitutions per amino acid position. Accession numbers in Genbank database is given next to a strain name. Node labels represent bootstrap confidence levels obtained using neighbor-joining method.

Figure 1

Click here to download Figure Figure1.tif

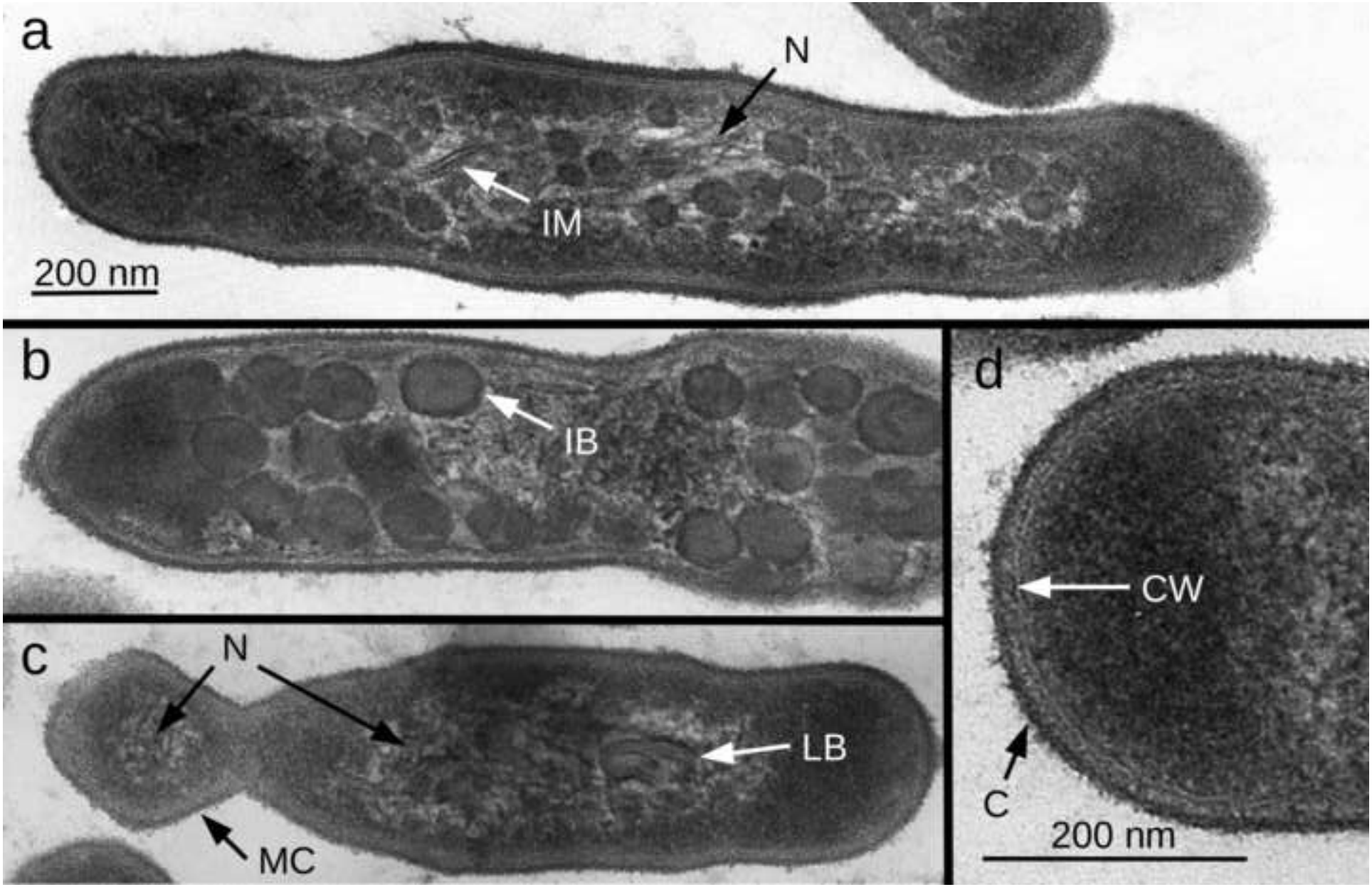


Figure 2

Click here to download Figure Figure2.ppt

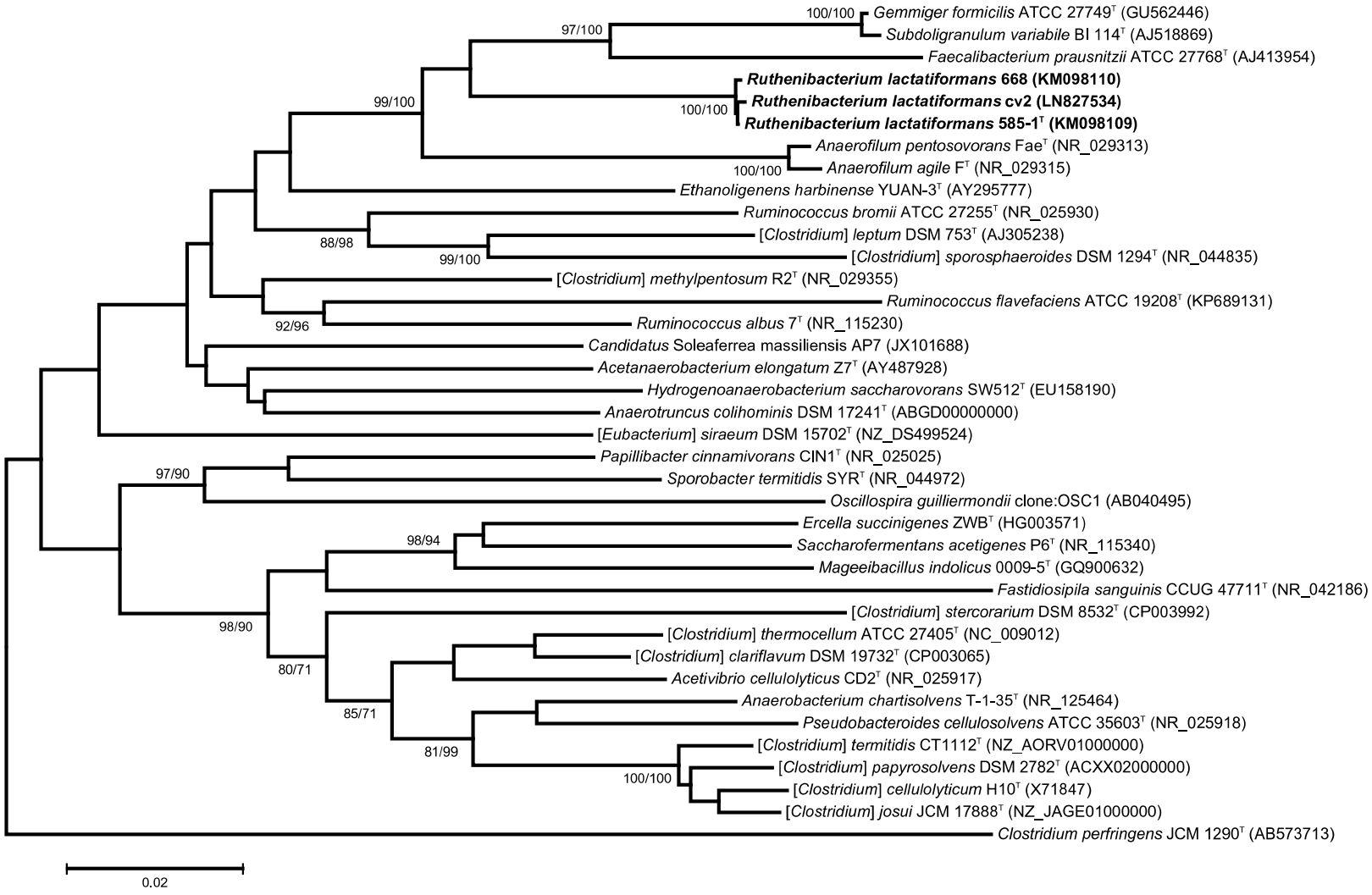
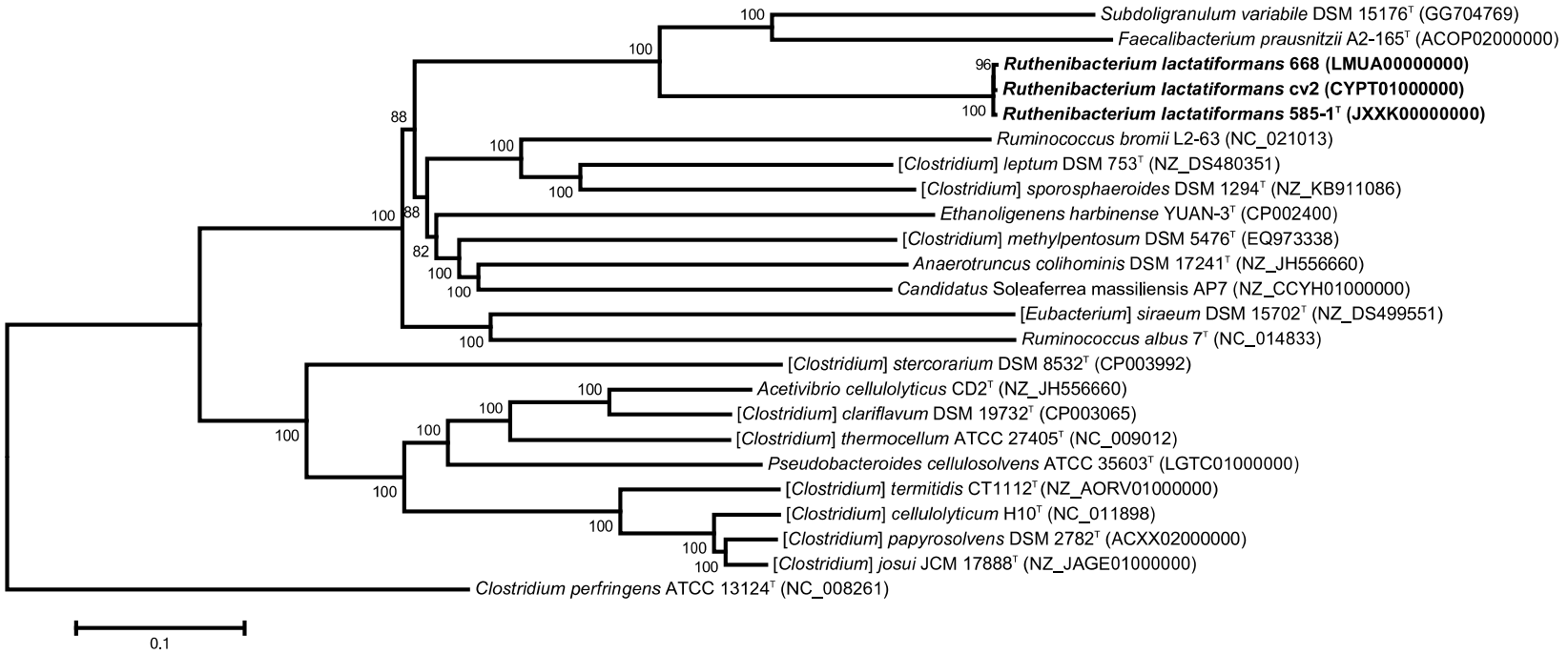


Figure 3

Click here to download Figure Figure3.ppt



Supplementary data

Ruthenibacterium lactatiformans gen. nov., sp. nov., a new anaerobic, lactate-producing member of the family _Ruminococcaceae_ isolated from human feces

Andrei N. Shkoporov, Andrei V. Chaplin, Victoria A. Shcherbakova, Natalia E. Suzina, Lyudmila I. Kafarskaia, Vladimir K. Bozhenko, and Boris A. Efimov

Table S1. Biochemical profiles of strains 585-1^T and 668 obtained using the Vitek 2 ANC identification panel.

	Biochemical test	Code	Strains	
			585-1	668
4	D-Galactose	dGal	-	-
5	Leucine arylamidase	LeuA	-	-
6	ELLMAN	ELLM	+	+
7	Phenylalanine Arylamidase	PheA	-	-
8	L-Proline arylamidase	ProA	-	-
10	L-Pyrrolydonyl arylamidase	PyrA	-	-
11	D-Cellobiose	dCEL	-	-
13	Tyrosine arylamidase	TyrA	-	-
15	Ala-Phe-Pro arylamidase	APPA	-	-
18	D-Glucose	dGLU	-	-
20	D-Mannose	dMNE	-	-
22	D-Maltose	dMAL	-	-
28	Sucrose	SAC	-	-
30	Arbutine	ARB	-	-
33	N-Acetyl-glucosamine	NAG	-	-
34	5-Bromo-4-chloro-3-indoxyl-β-glucoside	BGLUi	+	+
36	Urease	URE	-	-
37	5-Bromo-4-chloro-3-indoxyl-β-glucuronide	BGURi	+	+
39	5-Bromo-4-chloro-3-indoxyl-β-galactopyranoside	BGALi	+	+
41	α-Arabinosidase	AARA	+	+
42	5-Bromo-4-chloro-3-indoxyl-α-galactoside	AGALi	+	+
43	β-Mannosidase	BMAN	-	-
44	Arginine GP	ARG	-	-
45	Pyruvate	PVATE	-	-
51	Maltotriose	MTE	-	-
53	Aesculine, hydrolysis	ESC	+	+
54	β-D-Fucosidase	BdFUC	-	-
55	5-Bromo-4-chloro-3-indoxyl-β-N-acetyl-glucosamide	BNAGi	+	+
56	5-Bromo-4-chloro-3-indoxyl-α-mannoside	AMANi	+	+
57	α-L-Fucosidase	AIFUC	-	-
59	Phosphatase	PHOS	-	-
60	L-arabinose	IARA	-	-
61	D-Ribose 2	dRIB2	-	-
62	Phenylphosphonate	OPS	-	-
63	α-L-Arabinosidase	AARAF	-	-
64	L-Xylose	dXYL	-	-

Table S2. Biochemical profiles of strains 585-1^{T} and 668 obtained using the Rapid ID 32 A panel.

Biochemical test	Code	Substrate	Strains	
			585-1	668
Urease	URE	Urea	-	-
Arginine dihydrolase	ADH	L-arginine	-	-
α-galactosidase	αGAL	4-nitrophenyl-αD-galactopyranoside-	+	+
β–galactosidase	βGAL	4-nitrophenyl-βD-galactopyranoside	+	+
β–galacostidase 6 phophate	βGP	4-nitrophenyl-βD-galactopyranoside-6-phosphate-2CHA	-	-
α-glucosidase	αGLU	4-nitrophenyl-αD-glucopyranoside	+	+
β-glucosidase	βGLU	4-nitrophenyl-βD-glucopyranoside	+	+
α-arabinosidase	αARA	4-nitrophenyl- αL-arabinofuropyranoside	-	-
β-glucuronidase	βGUR	4-nitrophenyl-βD-glucuronide	+	+
β-N-acetyl glucosaminidase	βNAG	4-nitrophenyl-N-acetyl-β-D-glucosaminide	+	+
Mannose fermentation	MNE	D-mannose	-	-
Raffinose fermentation	RAF	D-raffinose	-	-
glutamate decarboxylase	GDC	Glutamic acid	-	-
α-fucosidase	αFUC	4-nitrophenyl- αL-fucopyranoside	-	-
nitrate reductase	NIT	Potassium nitrate	-	-
Indole production	IND	L-tryptophane	-	-
alkaline phosphatase	PAL	2-naphtyl-phosphate	-	-
Arginine arylamidase	ArgA	L-arginine- β-naphtylamide	-	-
Proline arylamidase	ProA	L-proline- β-naphtylamide	-	-
leucyl glycine arylamidase	LGA	L-leucyl-L-glycine - β-naphtylamide	-	-
phenylalanine arylamidase	PheA	L- phenylalanine - β-naphtylamide	-	-
leucine arylamidase	LeuA	L-leucine- β-naphtylamide	-	-
Pyroglutamic acid arylamidase	PyrA	Pyroglutamic acide - β-naphtylamide	-	-
Tyrosine arylamidase	TyrA	L-tyrosine- β-naphtylamide	-	-
Alanine arylamidase	AlaA	L-alanyl-L-alanine- β-naphtylamide	-	-
Glycine arylamidase	GlyA	L-glycine- β-naphtylamide	-	-
Histidine arylamidase	HysA	L-histidine- β-naphtylamide	-	-
Glutamyl Glutamic acid arylamidase	GGA	L-glutamyl-L-glutamic acide-- β-naphtylamide	-	-
Serine arylamidase	SerA	L-serine- β-naphtylamide	-	-

Table S3. Biochemical characteristics of strains 585-1^{T} and 668 determined using API 20 A identification system with MRS medium (readings after 72 h incubation at 37°C).

Characteristic	***585***	***668***
Indole production	-	-
Urease activity	-	-
Acid production from D-glucose	weak	weak
Acid production from D-mannitol	-	-
Acid production from lactose	-	-
Acid production from sucrose	-	-
Acid production from maltose	+	+
Acid production from salicin	+	+
Acid production from D-xylose	-	-
Acid production from L-arabinose	-	-
Gelatin digestion	-	-
Aesculin hydrolysis	+	+
Acid production from glycerol	-	-
Acid production from D-cellobiose	-	-
Acid production from D-mannose	-	-
Acid production from D-melezitose	-	-
Acid production from D-raffinose	-	-
Acid production from D-sorbitol	-	-
Acid production from L-rhamnose	weak	weak
Acid production from D-trehalose	-	-
Catalase production	-	-

Table S4. Carbohydrate fermentation profiles of strains 585-1^{T} and 668 determined in MRS medium supplemented with 0.01% (w/v) bromocresol purple indicator (readings after 96 h incubation at 37°C).

Characteristic	***585***	***668***
Acid production from D-glucose	+	+
Acid production from D-mannitol	-	-
Acid production from lactose	-	-
Acid production from sucrose	+	weak
Acid production from maltose	+	+
Acid production from salicin	+	+
Acid production from D-galactose	+	+
Acid production from L-arabinose	-	-
Acid production from adonitol	-	-
Acid production from α-methyl-D-glucoside	-	-
Acid production from D-mannose	weak	weak
Acid production from melibiose	weak	weak
Acid production from D-raffinose	-	-
Acid production from D-sorbitol	weak	weak
Acid production from L-rhamnose	+	+
Acid production from D-trehalose	-	-
Acid production from amygdalin	-	-

Table S5. Antibiotic susceptibility profile of strain 585-1^{T} determined using disk-diffusion method on (growth inhibition zones were measured 72 h after inoculation on EG agar).

Antibiotic		Amount of substance in disk, μg	Diameter of growth inhibition zone, mm 585-1	**Sensitivity, Yes/No**
Amikacin	**Ak**	30	5	No
Amoxyclav	**Ac**	30	50	Yes
Ampicillin	**A**	2	5	No
Azithromycin	**At**	15	5	No
Cephalothin	**Ch**	30	5	No
Clindamycin	**Cd**	2	12	No
Levofloxacin	**Le**	5	5	No
Linezolid	**Lz**	30	20	No
Penicillin G	**P**	10 U	5	No
Vancomycin	**Va**	30	18	Yes

Table S6. Cellular fatty acids (CFA) analysis from 2 mg of washed and dried cells of strain 585-1.

CFA species*	Peak area (percentage of total), Experiment 1	Peak area (percentage of total), Experiment 2
14:1ω3	0.3%	0.7%
14:0	0.9%	2.2%
ai15	0.1%	0.4%
15:0	0.2%	0.3%
15:1ω6	0.1%	0.2%
i16	0.1%	0.1%
16:1ω9	0.2%	0.2%
16:1ω7c	**5.3%**	**6.1%**
16:1ω7t	0.2%	0.2%
16:1ω5	0.2%	0.1%
16:0	**6.1%**	**6.6%**
16:1ω7a	1.6%	1.4%
16a	3.0%	2.1%
i17	0.9%	0.8%
ai17	0.7%	2.9%
17:1ω8	0.8%	1.4%
17:0	0.7%	0.9%
i17a	0.1%	0.1%
ai17a	1.0%	0.7%
17:1ω8a	0.6%	0.5%
17a	0.1%	0.1%
18:2	0.2%	0.3%
18:1ω9	**31.9%**	**31.4%**
18:1ω7c	5.3%	4.8%
18:1ω7t	3.5%	3.0%
18:1ω5	0.4%	0.4%
18:0	2.3%	1.6%
18:1ω9a	**21.1%**	**20.6%**
18:1ω7ca	5.2%	3.8%
18:1ω7ta	1.0%	1.0%
18:1ω5a	0.3%	0.5%
18a	1.2%	0.9%
i19	0.4%	0.0%
19:0	0.1%	0.6%
i19a	0.1%	0.1%
ai19a	0.2%	0.1%
20:1ω9	0.4%	0.5%
10h18	3.1%	2.3%
20:0	0.2%	0.1%
	100.0%	100.0%

* c, *cis*; t, *trans*; i, *iso*; ai, *anteiso*; a, aldehyde; h, hydroxy

Table S7. List of 204 core proteins used for phylogenetic inference.

Locus tag in str	Protein annotation
TQ39_RS06915	peptidyl-tRNA hydrolase
TQ39_RS00675	RNA methyltransferase
TQ39_RS11970	ATP-dependent DNA helicase RecG
TQ39_RS07070	gamma-glutamyl-phosphate reductase
TQ39_RS06745	Holliday junction resolvase
TQ39_RS15555	thiamine pyrophosphokinase
TQ39_RS02850	RNA polymerase sigma factor RpoD
TQ39_RS10525	hypothetical protein
TQ39_RS05190	50S ribosomal protein L18
TQ39_RS05235	50S ribosomal protein L16
TQ39_RS13860	hypothetical protein
TQ39_RS00710	transcription antitermination protein NusB
TQ39_RS05200	30S ribosomal protein S8
TQ39_RS05765	mRNA interferase PemK
TQ39_RS10565	aminopeptidase
TQ39_RS05275	30S ribosomal protein S10
TQ39_RS02285	16S rRNA maturation RNase YbeY
TQ39_RS09960	non-canonical purine NTP pyrophosphatase
TQ39_RS04565	hypothetical protein
TQ39_RS06225	cysteine--tRNA ligase
TQ39_RS13430	phosphoglucosamine mutase
TQ39_RS01830	excinuclease ABC subunit B
TQ39_RS00625	queuine tRNA-ribosyltransferase
TQ39_RS05145	30S ribosomal protein S13
TQ39_RS10570	ribosomal protein S12 methylthiotransferase
TQ39_RS05280	RNA-binding protein
TQ39_RS10515	riboflavin biosynthesis protein RibF
TQ39_RS16585	ATPase
TQ39_RS05260	50S ribosomal protein L23
TQ39_RS05130	DNA-directed RNA polymerase subunit alpha
TQ39_RS05175	50S ribosomal protein L15
TQ39_RS09950	RNA-binding protein
TQ39_RS05240	30S ribosomal protein S3
TQ39_RS08300	chorismate synthase
TQ39_RS10005	50S ribosomal protein L35
TQ39_RS14235	50S ribosomal protein L27
TQ39_RS14210	transcriptional regulator
TQ39_RS14240	50S ribosomal protein L21
TQ39_RS10530	ribosome-binding factor A
TQ39_RS15615	hypothetical protein
TQ39_RS16435	hypothetical protein
TQ39_RS05725	cytidylate kinase
TQ39_RS00085	alanine--tRNA ligase
TQ39_RS10075	ferredoxin-NADP+ reductase subunit alpha
TQ39_RS04365	tRNA sulfurtransferase ThiI
TQ39_RS00745	stage III sporulation protein AC
TQ39_RS01765	hypothetical protein

TQ39_RS12370	50S ribosomal protein L13
TQ39_RS16590	sigma-70 family RNA polymerase sigma factor
TQ39_RS14225	cysteine--tRNA ligase
TQ39_RS08165	hypothetical protein
TQ39_RS00670	hypothetical protein
TQ39_RS00090	histidine triad nucleotide-binding protein
TQ39_RS11410	ribosome biogenesis GTPase Der
TQ39_RS09185	membrane protein insertase
TQ39_RS10550	transcription termination factor NusA
TQ39_RS00595	DNA repair protein RadA
TQ39_RS07895	nucleoside triphosphate pyrophosphohydrolase
TQ39_RS15600	primosomal protein N'
TQ39_RS05620	RNA methyltransferase
TQ39_RS05165	adenylate kinase
TQ39_RS10805	adenylosuccinate lyase
TQ39_RS04610	DUF378 domain-containing protein
TQ39_RS10070	dihydropyrimidine dehydrogenase subunit A
TQ39_RS11560	argininosuccinate synthase
TQ39_RS13510	1-deoxy-D-xylulose-5-phosphate reductoisomerase
TQ39_RS10685	50S ribosomal protein L11
TQ39_RS09220	hypothetical protein
TQ39_RS02265	DNA mismatch repair protein MutS
TQ39_RS02365	50S rRNA methyltransferase
TQ39_RS02870	IMPACT family protein
TQ39_RS10830	phosphoribosylformylglycinamidine cyclo-ligase
TQ39_RS05790	phenylalanine--tRNA ligase subunit beta
TQ39_RS10510	30S ribosomal protein S15
TQ39_RS09275	16S rRNA methyltransferase
TQ39_RS05195	50S ribosomal protein L6
TQ39_RS05245	50S ribosomal protein L22
TQ39_RS07880	sporulation protein YabP
TQ39_RS13505	RIP metalloprotease RseP
TQ39_RS09270	kinase to dihydroxyacetone kinase
TQ39_RS09295	serine--tRNA ligase
TQ39_RS05215	50S ribosomal protein L24
TQ39_RS09350	ADP-ribose pyrophosphatase
TQ39_RS05700	tRNA-specific adenosine deaminase
TQ39_RS03410	triose-phosphate isomerase
TQ39_RS08545	RNA-binding protein
TQ39_RS10575	hypothetical protein
TQ39_RS15605	guanylate kinase
TQ39_RS10690	50S ribosomal protein L11
TQ39_RS10520	pseudouridine synthase
TQ39_RS08450	hypothetical protein
TQ39_RS08075	methionine--tRNA ligase
TQ39_RS13865	hypothetical protein
TQ39_RS14230	GTPase CgtA
TQ39_RS09250	tRNA modification GTPase
TQ39_RS01880	50S ribosomal protein L9
TQ39_RS04625	uracil phosphoribosyltransferase
TQ39_RS10375	nifR3 family TIM-barrel protein
TQ39_RS02210	elongation factor Ts
TQ39_RS10840	amidophosphoribosyltransferase

TQ39_RS02855	DNA primase
TQ39_RS06530	hypothetical protein
TQ39_RS09845	dimethyladenosine transferase
TQ39_RS00665	arginine repressor ArgR
TQ39_RS00740	stage III sporulation protein AD
TQ39_RS00630	aspartate-semialdehyde dehydrogenase
TQ39_RS02270	DNA repair protein RecO
TQ39_RS09190	ribonuclease P protein component
TQ39_RS15545	stage IV sporulation protein A
TQ39_RS02880	50S ribosomal protein L31
TQ39_RS16490	cell division protein FtsZ
TQ39_RS02295	hypothetical protein
TQ39_RS04700	hypothetical protein
TQ39_RS12375	30S ribosomal protein S9
TQ39_RS04400	formate--tetrahydrofolate ligase
TQ39_RS02275	GTPase Era
TQ39_RS09415	stage 0 sporulation protein
TQ39_RS09865	cell division protein FtsE
TQ39_RS08555	signal recognition particle protein
TQ39_RS02835	DNA-directed RNA polymerase subunit beta'
TQ39_RS13500	PolC-type DNA polymerase III
TQ39_RS13440	Holliday junction DNA helicase RuvB
TQ39_RS05230	50S ribosomal protein L29
TQ39_RS15665	hypothetical protein
TQ39_RS10410	N(6)-L-threonylcarbamoyladenine synthase TsaD
TQ39_RS05210	50S ribosomal protein L5
TQ39_RS10535	translation initiation factor IF-2
TQ39_RS00985	ATP-dependent DNA helicase PcrA
TQ39_RS11435	glucose-6-phosphate isomerase
TQ39_RS10715	hypothetical protein
TQ39_RS05270	50S ribosomal protein L3
TQ39_RS15925	YggS family pyridoxal phosphate enzyme
TQ39_RS02215	30S ribosomal protein S2
TQ39_RS00720	hypothetical protein
TQ39_RS05265	50S ribosomal protein L4
TQ39_RS13525	ribosome recycling factor
TQ39_RS00615	preprotein translocase subunit YajC
TQ39_RS06905	transcription-repair coupling factor
TQ39_RS08560	DNA-binding protein
TQ39_RS03480	adenylosuccinate synthetase
TQ39_RS01110	UDP-N-acetylenolpyruvoylglucosamine reductase
TQ39_RS12250	endonuclease III
TQ39_RS09900	elongation factor P
TQ39_RS01785	DNA mismatch repair protein MutL
TQ39_RS10670	50S ribosomal protein L7/L12
TQ39_RS10000	50S ribosomal protein L20
TQ39_RS10695	preprotein translocase subunit SecE
TQ39_RS10680	50S ribosomal protein L1
TQ39_RS09200	chromosomal replication initiator protein DnaA
TQ39_RS15645	sporulation protein YtfJ
TQ39_RS05180	50S ribosomal protein L30
TQ39_RS05155	translation initiation factor IF-1
TQ39_RS10845	N5-carboxyaminoimidazole ribonucleotide mutase

TQ39_RS00500	hypothetical protein
TQ39_RS04605	DUF378 domain-containing protein
TQ39_RS01165	NrdR family transcriptional regulator
TQ39_RS01190	isoleucine--tRNA ligase
TQ39_RS09995	hypothetical protein
TQ39_RS03760	glycine--tRNA ligase
TQ39_RS08550	30S ribosomal protein S16
TQ39_RS05140	30S ribosomal protein S11
TQ39_RS10555	ribosome assembly cofactor RimP
TQ39_RS08270	3-phosphoshikimate 1-carboxyvinyltransferase
TQ39_RS13520	UDP pyrophosphate synthase
TQ39_RS01770	hypothetical protein
TQ39_RS18975	GTP-binding protein YchF
TQ39_RS03400	phosphoglycerate kinase
TQ39_RS01125	HPr kinase
TQ39_RS13455	GTP-binding protein
TQ39_RS04585	AsnC family transcriptional regulator
TQ39_RS00700	exodeoxyribonuclease VII large subunit
TQ39_RS00620	queuine tRNA-ribosyltransferase
TQ39_RS09970	ribonuclease III
TQ39_RS06240	UDP-N-acetylmuramoylalanine--D-glutamate ligase
TQ39_RS08100	ribulose-phosphate 3-epimerase
TQ39_RS01725	hypothetical protein
TQ39_RS08005	branched chain amino acid aminotransferase
TQ39_RS08310	chorismate mutase
TQ39_RS05250	30S ribosomal protein S19
TQ39_RS01105	RNase adaptor protein RapZ
TQ39_RS05785	phenylalanine--tRNA ligase subunit alpha
TQ39_RS01870	hypothetical protein
TQ39_RS15670	peptidase M50
TQ39_RS00505	recombinase RecR
TQ39_RS09955	non-canonical purine NTP pyrophosphatase
TQ39_RS01865	hypoxanthine phosphoribosyltransferase
TQ39_RS02840	DNA-directed RNA polymerase subunit beta
TQ39_RS00230	NAD synthetase
TQ39_RS01650	GTP-binding protein
TQ39_RS00105	DNA polymerase III subunit delta
TQ39_RS15585	hypothetical protein
TQ39_RS10545	hypothetical protein
TQ39_RS02440	preprotein translocase subunit SecG
TQ39_RS05255	50S ribosomal protein L2
TQ39_RS01875	DNA helicase
TQ39_RS01100	hypothetical protein
TQ39_RS15560	ribosome small subunit-dependent GTPase A
TQ39_RS05125	50S ribosomal protein L17
TQ39_RS11175	RNA methyltransferase
TQ39_RS11220	30S ribosomal protein S20
TQ39_RS13530	UMP kinase
TQ39_RS01635	heat-shock protein Hsp33
TQ39_RS11190	50S ribosomal protein L11 methyltransferase
TQ39_RS15610	hypothetical protein