Description of Paenibacillus albicerus sp. nov. and Niallia alba sp. nov., isolated from digestive syrup.

Vipool Thorat
National Centre For Cell Science
Kiran Kirdat
NCCS: National Centre For Cell Science; Tujlaram chaturchand College, Baramati
Bhavesh Tiwarekar
NCCS: National Centre For Cell Science
Priyanka Dhanavade
NCCS: National Centre For Cell Science
Prachi Karodi
NCCS: National Centre For Cell Science
Yogesh Shouche
NCCS: National Centre For Cell Science
Shivaji Sathe
NCCS: National Centre For Cell Science
Tushar Lodha
NCCS: National Centre For Cell Science
Amit Yadav (amityadav@nccs.res.in)
National Centre For Cell Science https://orcid.org/0000-0002-4882-9075

Research Article

Keywords: Paenibacillus, Bacillus, Niallia, fungal diastase, α-amylase, pepsin, polyethylene glycol

Posted Date: July 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-758723/v1

Abstract

Two aerobic, Gram-stain variable, catalase and oxidase-positive, spore-forming, motile rods of strain UniB2^T and UniB3^T, were isolated from digestive syrup containing fungal diastase (10 mg/ml), pepsin (2 mg/ml) and sugar base containing polyethene glycol. Based on 16S rRNA gene sequence analysis, strain UniB2^T has the highest sequence similarity with *Paenibacillus humicus* NBRC 102415^T (98.30 %) and strain UniB3^T showed the highest sequence similarity with *Niallia circulans* DSM 11^T (98.95 %). The DNA G+C content of UniB2^T was 63.74 mol %. The dDDH and ANI values between the strain UniB2^T and its phylogenetically close relative were <38.3 % and <89.55 %, respectively. The major fatty acids of the strain UniB2^T were C16:00 (13.9 %), $C_{15:0}$ anteiso (39.7 %), $C_{17:0}$ anteiso (15.5 %). The DNA G+C content of UniB3^T was 35.6 mol%. The dDDH and ANI values between the strain UniB3^T and its close relatives were <29.1 % and 84.62%, respectively. The major fatty acids of strain UniB3^T were C16:0 (13.54 %), C15:0 anteiso (40.09 %) and C17:0 anteiso (16.03 %). Major polar lipids for both strains were Diphosphotidylglycerol and Phosphatidylethanolamine. Additionally, both the strains showed unique carbon utilization and assimilation pattern that differentiated them from their phylogenetically related neighbours. These phenotypic, genotypic and chemotaxonomic characters indicated the strains UniB2^T and UniB3^T represent two novel species for which the names *Paenibacillus albicerus* sp. nov. (Type strain UniB2^T =MCC 3997^T =KCTC 43095^T =JCM34513^T) and *Niallia alba* sp. nov. (Type strain UniB3^T = MCC 3998^T =KCTC 43235^T =JCM 34492^T) are proposed.

Introduction

Genus *Paenibacillus* was reclassified and separated from the *Bacillus* genus in 1993 (Ash et al. 1993). The name *Paenibacillus* came into existence from the Latin word *'Paene'* means 'almost', so *Paenibacillus* means 'almost Bacilli'. The genus *Paenibacillus* was initially placed in a former 'group 3' of genus *Bacillus*. Comparative 16S rRNA sequence analyses revealed that group 3 bacilli were phylogenetically distinct and sparsely related to type species *Bacillus subtilis*. Strains belonging to genus *Paenibacillus* have been reported from various environments like soil, rhizosphere, water, plant, animal, and human clinical samples (Yoon et al. 2005; Roux et al. 2008; Zhang et al. 2016; Hwang and Ghim 2017; Simon et al. 2017; Yun et al. 2017). The genus is composed of more than 320 validly published species based on the List of Prokaryotic Names with Standing in Nomenclature (Parte et al. 2020); some of these species exhibits various activities like nitrogen fixation (Heulin et al. 1994), phosphorous solubilization (Singh and Singh 1993), antibiotic production (Piuri et al. 1998) and production of chitinase (Mavingui and Heulin 1994).

Genus *Niallia* has been recently proposed and separated from the genus *Bacillus* (Gupta et al. 2020). Early classification of *Bacillus* species were based on two characteristics: aerobic growth and endospore formation (Logan and De Vos 2009; Logan et al. 2009). There were more than 570 species reported in this genus to date. *Bacillus* has long appeared as a heterogeneous phylogenetic group (Parte et al. 2020). Most of the species of this genus show minimal characters and exhibits substantial

polyphyly. Gupta et al. (2020) identified conserved signature indels (CSIs) specific for each identified clade and proposed reclassification of 148 *Bacillus* species into 17 novel genera and 11 existing genera. They identified 128 CSIs, which exhibit uniqueness to the specific clade and provide reliable means for their distinction (Gupta et al. 2020). This report describes two novel species belonging to the genus *Paenibacillus* and *Niallia* isolated from the digestive syrup, closely related to the *Paenibacillus humicus* (Vaz-Moreira et al. 2007) and *Niallia circulans* (Gupta et al. 2020).

Methods And Materials

Bacterial isolation and culture conditions

Strain UniB2^T and UniB3^T were isolated from the digestive syrup containing digestive enzymes [diastase; isolated from *Aspergillus niger* (10 mg/ml) and pepsin (2 mg/ml)] used in digestive troubles. The contaminated syrup was centrifuged, and the pellet was serially diluted and spread on a Nutrient agar (NA) medium (M001; Himedia, India), and cultures were grown at 30 ^{0} C for 48 hours.

Morphological, physiological and biochemical characteristics

The size and shape of the cells were determined by scanning electron microscopy (Zeiss, EVO 18, Version 6.02), as seen in supplementary Figure S1. The Gram character for both the strains was determined by Gram staining kit (K001-KT; Himedia, India) and observed by light microscopy (Model BX53; Olympus, USA). Motility was analyzed by the hanging drop method under a light microscope (Model BX53; Olympus, USA). Oxidase and catalase activity were investigated using oxidase discs (DD018; Himedia, India) and bubble production in 3 % (v/v) H2O2, respectively. The strains were optimally grown on NA at 30^{0}C (5 – 40^{0}C range) in 96 hours. The pH range for growth was determined at pH 5-12 with increments of 1 pH unit with different buffer systems like acetate (4 to 5 unit), phosphate (6 to 8 units) and glycine – sodium hydroxide (9 to 12 units). Growth of strains was detected using a Bioscreen C microbiology reader (Oy Growth Curves AB Inc., Finland). Salt tolerance for growth was examined in Nutrient broth (NB) with 0–4 % NaCl with an increment of 0.5 % at 30 °C for five days, and the growth was observed by a 'Bioscreen C Microbiology' reader. The pH, salinity, and temperature optimization were carried out on NB for five days. Biochemical characteristics like utilization, assimilation of different carbon sources and enzyme activity against different substrates were tested by analytical profile index, biomerieux (API 20NE, API ZYM) and Biolog GEN III Microplate assay according to the instructions of the manufacturers.

Chemotaxonomic characterization

The whole-cell fatty acids (Fatty Acid Methyl ester, FAME) and polar lipids were determined for chemotaxonomic analyses. Forty-eight hours old cultures, grown on tryptic soy agar at pH 7.3±0.2 (M1968; Himedia) of strain UniB2^T, NBRC 102415^T, NBRC 101214^T, UniB3^T, DSM 11^T and DSM 15077^T were harvested for cell mass and extracted total cell fatty acids were analyzed by RTSBA6 library on Sherlock Microbial Identification System (Version 6.1; MIDI) (Sasser 2001). For polar lipid extraction, cell

mass was harvested from cultures in the logarithmic phase (Bligh and Dyer 1959; Card 1973; Minnikin et al. 1984).

Phylogenetic and genotypic analysis

All obtained colonies were subjected to genomic DNA extraction (Ausubel et al. 1994), followed by their identification by 16S rRNA gene sequencing. The 16S rRNA gene PCR was done with universal 16S rRNA gene primers (Sambrook and Russell 2001), and the EzBioCloud database(Yoon et al. 2017) was used to identify the closest type relative. The genomic DNA of the strains UniB2^T and UniB3^T were sequenced on the Illumina MiSeq and Oxford Nanopore Technology (ONT) platforms. The reads were quality checked using FastQC version 0.10.1 (Brown et al. 2017). The ONT sequencing data were base-called with quality filtering (>Q7) using Guppy version 3.5.4. All QC passed Illumina reads (>Q30), and ONT reads were used to generate a hybrid assembly in Unicycler version 0.4.8 (Wick et al., 2017). The genome quality was assessed using QUAST (v5.0.2) (Gurevich et al. 2013) and CheckM (v1.1.3) (Parks et al. 2015). The genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). 16S rRNA gene phylogeny was constructed using the software packages MEGA version 7.0 (Kumar et al. 2016) with three algorithms, the neighbour-joining (Saitoh 1987), maximum-likelihood (Felsenstein 1981) and maximum-parsimony (Fitch 1971) methods. Further, to support the taxonomic position of strain UniB2^T and UniB3^T, a core-genome phylogenetic tree was constructed using BPGA, version 1.3.0 (Chaudhari et al. 2016) and UBCG, version 3.0 (Na et al. 2018) tools for both strains (Fig. 1 and 2). The BPGA tool searched for core genes from the genomes of 17 members phylogenetically close to strain UniB2^T and among 20 members related to UniB3^T. The BPGA pipeline generated orthologous amino acid sequences clusters using the integrated USEARCH algorithm (Edgar 2010). Among these, 20 random clusters were aligned and concatenated. Finally, the phylogenetic tree was reconstructed using the BPGA concatenated sequences by the neighbour-joining method in MEGA 7. A total of 12,595 and 14,771 (UniB2^T and UniB3^T) amino acid positions in BPGA final dataset were used for the tree generation. The length of the UBCG concatenated alignment was 86,964 bp containing 92 marker genes for UniB2^T and 85,746 bp containing 92 marker genes for UniB3^T identified using HMMER (Potter et al. 2018) and predicted using Prodigal (Hyatt et al. 2010) search.

The orthoANI values were calculated using the orthoANI calculator (Yoon et al. 2017). The digital DNA–DNA hybridization (dDDH) values were calculated using Genome-to-Genome Distance Calculator (GGDC) web server and tested using recommended formula 2 (Richter and Rosselló-Móra 2009; Auch et al. 2010; Meier-Kolthoff et al. 2014).

Results And Discussion

Morphological, physiological and biochemical characteristics

Morphologically, colonies of strain UniB2^T were mucoid, brownish-yellow in colour with an entire margin and convex elevation. In contrast, colonies of strain UniB3^T were off-white, circular, entire in the margin,

flat opaque. The strain UniB2^T and UniB3^T were found motile, rod in shape with 0.8 to 1.0 µM width; however, their length varied from 2 to 3.5 and 2 to 4.5 µM, respectively. Both the strains showed varied gram character. The strains were able to grow on pH 6- 8 (Optimum 7.0; 5.0 -12.0 pH range in increment of 1 pH unit), salinity 0.5- 2 % (Optimum 0.5 %; 0.5 – 4 % range in increment of 0.5 %).

According to API 20NE tests, strain UniB2^T showed positive results for hydrolysis of aesculin and P-nitrophenyl-β-D-galactopyranoside and assimilation of glucose, mannitol, N-acetyl-glucosamine, maltose, whereas it showed negative results for arabinose, gluconate and malate. Among three, only *Paenibacillus pasadenensis* NBRC 101214^T showed a positive result for assimilation of arabinose and malate. For gluconate assimilation, strain UniB2^T alone showed a negative result. By API ZYM tests, the strain UniB2^T was positive for ά-galactosidase, ß-galactosidase, ά glucosidase, ß glucosidase activity. The strain UniB2^T tested negative for esterase activity, while its phylogenetically closest neighbours (*Paenibacillus humicus* NBRC 102415^T and *Paenibacillus pasadenensis* NBRC 101214^T) tested positive. Sugar utilization patterns were checked by Biolog GEN III microbial identification system. Strain UniB2^T did not utilize N-acetyl-D-galactosamine, glycerol, methyl pyruvate, L-malic acid but NBRC 102415^T and NBRC 101214 ^{T} were tested positive for utilizing these sugars. The strain NBRC 102415^T and NBRC 101214^T were tested negative for utilization of 3-methyl glucose, D-fucose, L-Rhamnose, L-arginine, L-aspartic acid, whereas strain UniB2^T was able to utilize those sugars. The strain UniB2^T and NBRC 101214^T were positive for utilizing gentiobiose, α-D-glucose, L-fucose, D-fructose-6-Phosphate, D-galacturonic acid, β-hydroxy-D, L-butyric acid, acetic acid, whereas strain NBRC 102415^T was tested negative for these sugars. Strain UniB2^T and NBRC 102415^T were tested negative for L-lactic acid, D-malic acid, bromo-succinic acid, whereas strain *Paenibacillus pasadenensis* NBRC 101214^T showed positive results for utilization of these sugars. Strain UniB2^T was sensitive to rifamycin SV, lithium chloride, whereas NBRC 102415^T and NBRC 101214^T were insensitive to these chemicals. Both NBRC 102415^T and NBRC 101214^T were sensitive to D-serine, guanidine-HCl whereas strain UniB2^T was found resistant to these chemicals. The morphological, cultural, physiological and biochemical characteristics of strain UniB2^T are given in Table 1.

Strain UniB3^T showed positive results for hydrolysis of P-nitrophenyl-βD-galactopyranoside and assimilation of arabinose, mannose, N-acetyl- glucosamine, gluconate, malate and citrate. At the same time, its closest relatives *Niallia nealsonii* DSM 15077^T, showed a negative result for hydrolysis of P-nitrophenyl-β-D-galactopyranoside and assimilation of arabinose, gluconate, malate and citrate. In comparison, *Niallia circulans* DSM 11^T showed a negative result for assimilation of arabinose mannose, N- acetyl glucosamine and malate, according to API 20NE tests. According to the API ZYM test, Strain UniB3^T showed naphthol AS-BI-phosphohydrolase, ß-galactosidase, ά-glucosidase enzyme activity whereas both the closest neighbours were negative for ß-galactosidase, ά-glucosidase activity and showed positive activity for naphthol AS-BI-phosphohydrolase similar to strain UniB3^T. Sugar utilization pattern was tested by Biolog GEN III microbial identification system; strain UniB3^T and DSM 11^T showed a positive result for utilization of dextrin, D-maltose, D-cellobiose, D-turanose, stachyose, D-raffinose, α-D-

lactose, β-methyl-D-glucoside, D-salicin, N-acetyl-D-glucosamine, N-acetyl-D-mannosamine, inosine, D-sorbitol, D-mannitol, glycerol, D-gluconic acid, methyl pyruvate, tween 40, acetoacetic acid, acetic acid whereas the strain DSM 15077^T showed a negative result for these sugars. The strain DSM 15077^T showed a positive sugar utilization reaction for N-acetyl-D-galactosamine, 3-methyl glucose, D-fucose, D-galacturonic acid, Mucic acid, β-hydroxy-D, L-butyric acid and other both the strains were negative for these sugars. Strain UniB3^T and DSM 15077^T showed a positive result for the utilization of L-fucose, L-rhamnose, D-glucuronic acid, and strain DSM 11^T showed negative utilization results for these sugars. The strain DSM 11^T showed positive results for Myo-inositol, L-alanine, L-malic acid, whereas strain UniB3^T and DSM 15077^T showed negative results for these sugars. Strain UniB3^T was not observed utilizing bromo-succinic acid, D-glucose-6-Phosphate, while both the strains in the comparative study were observed to use these sugars. Strain UniB3^T was found sensitive to troleandomycin and rifamycin SV antibiotics, while both the other strains were not found susceptible to both antibiotics. Strain UniB3^T and DSM 11^T were found sensitive to tetrazolium blue, 1 % sodium lactate, nalidixic lithium chloride, Lithium chloride, aztreonam, sodium bromate, whereas strain DSM 15077^T was not found sensitive to these chemicals (Table 2).

Chemotaxonomic characterization

The major fatty acids in strain UniB2^T were C16:00 (13.9 %), C15:0 anteiso (39.7 %), C17:0 anteiso (15.5 %). In Strain UniB3^T, major fatty acids were C16:0 (13.54 %), C15:0 anteiso (40.09 %) and C17:0 anteiso (16.03 %). All data related to the fatty acid composition of UniB2^T and UniB3^T are presented in Tables 3 and 4, respectively.

Major Polar lipids of strain UniB2^T were Diphosphotidylglycerol and Phosphatidylethanolamine and possessed unknown phospholipid (PL1, PL2, and PL3) and unknown amino phospholipids (APL 1, APL2, and APL3). APL4 was not observed in Strain UniB2^T but detected in phylogenetically closest neighbours of strain UniB2^T, whereas PL2 and APL3 were not detected in strain NBRC 102415^T and PL1, PL2 and APL3 were not detected in strain NBRC 101214^T. Diphosphotidylglycerol and Phosphatidylethanolamine were the major polar lipids found in strain UniB3^T. This strain possessed unknown phospholipid PL1. Phospholipid PL1, PL2 and PL3 were not detected in strain DSM 15077^T, and PL4 was not detected in strain UniB3^T but detected in DSM 11^T and DSM 15077^T (Fig. S2).

Phylogenetic and genotypic analysis

The 16S rRNA gene sequence of strain UniB2^T (MK751590) and UniB3^T (MK751589) were generated and deposited at NCBI GenBank. The strain UniB2^T showed the closest match with *Paenibacillus humicus* NBRC 102415^T (Vaz-Moreira et al. 2007) (98.30 %, BIMD01000104) followed by *Paenibacillus pasadenensis* NBRC 101214T (Osman et al. 2006) (98.06 %, AY167820). The strain Unib3^T showed the closest match with *Niallia circulans* DSM 11^T (Jordan 1890; Gupta et al. 2020) (98.95 %, AY724690),

followed by *Niallia nealsonii* DSM 15077^T (Venkateswaran et al. 2003; Gupta et al. 2020) (98.47 %, EU656111).

Sequencing of UniB2^T and UniB3^T on Illumina MiSeq platform generated 4,373,992 reads and 3,315,487 reads (150×2 chemistry), respectively. ONT sequencing generated 159,965 reads and 165,209 reads respectively for UniB2^T and UniB3^T. The final UniB2^T assembly contained one scaffold corresponding to 5,168,992 bp. This genome was found to be 99.03 % complete with 248× depth coverage. The final UniB3^T assembly had two scaffolds corresponding to 5,241,236 bp. This genome was found to be 92.78 % complete with 184× depth coverage. The DNA G+C contents were 63.74 mol % and 35.6 mol % for UniB2^T and UniB3^T. The strain UniB2^T was predicted to have 4434 protein-coding genes, 82 tRNA and 24 rRNA genes, whereas strain UniB3^T was predicted to have 4574 protein-coding genes, 83 tRNA and 33 rRNA genes (Zhao et al. 2011).

The strain UniB2^T clustered in a monophyletic clade with *P. pasadenensis* NBRC 101214^T and *P. humicus* NBRC 102415^T in the phylogenetic tree drawn using the 16S rRNA gene sequences (Fig. S3). Similarly, the strain UniB3^T clustered in a monophyletic clade with *Niallia circulans* DSM 11^T (Fig. S4). This analysis was supported by the core-genome phylogenetic tree constructed using BPGA and UBCG

The orthoANI value for strain UniB2^T against NBRC 101214^T was 89.55 %, with NBRC 102415^T, it was 77.72 %, and with "*P. herberti*" R33, it was 76.2 %. The digital DNA–DNA hybridization (dDDH) values for strain UniB2^T against NBRC 101214^T, NBRC 102415^T and R33 were 38.3, 22.6 and 21.2 %, respectively. The orthoANI value for strain UniB3^T against DSM 11^T was 84.62 %, and with DSM 15077^T, it was 74.2 %. The dDDH values for strain UniB3^T against DSM 11^T and DSM 15077^T were 29.1 and 22.2 %, respectively. The orthoANI and dDDH values computed for both the strains strongly support them to be novel species. Therefore, based on physiological, chemotaxonomic characters, phylogenetic and genomics indices, we propose strains UniB2^T and UniB3^T represent two novel species named *Paenibacillus albicerus sp. nov.* and *Niallia alba sp. nov.* are proposed, respectively.

Description of *Paenibacillus albicerus* sp. nov.

Paenibacillus albicerus (albi'ce.rus L. masc. adj. *albicerus*, pale yellow, the colour of colonies.)

Aerobic, catalase and oxidase-positive and Gram-Stain variable in nature. Spore-forming, motile rod with 0.8 to 1.0 µM width and 2 to 3.5 um in length. Colonies are mucoid, brownish-yellow in colour with an entire margin and convex elevation. The colony size on solid media (Nutrient agar) is 2-3 mm. Cells can grow on 5 – 40 0C temperature range (Optimum 30 0 C) on NA medium and able to grow on pH ranges from 6.0-8.0 (Optimum 7.0) unit and salinity 0.5 - 2.0 % (Optimum 0.5 %). Cells show positive results for hydrolysis of aesculin and P-nitrophenyl-βD-galactopyranoside and assimilation of glucose, mannitol, N-acetyl-glucosamine, maltose and shows negative results for gluconate assimilation. Cells show esterase lipase, naphthol AS-BI-phosphohydrolase, ά-galactosidase, ß-galactosidase, ά-glucosidase and ß-glucosidase activity. It utilizes D-sorbitol, 3-methyl glucose, D-fucose, L-Rhamnose, L-arginine, L-aspartic

acid, gentiobiose, α-D-glucose, L-fucose, D-fructose-6-Phosphate, D-galacturonic acid, β-hydroxy-D, L-butyric acid, acetic acid. It shows sensitivity towards nalidixic lithium chloride, lithium chloride, rifamycin SV. C16:00 (13.9 %), C15:0 anteiso (39.7 %) and C17:0 anteiso (15.5 %) as major fatty acids and Diphosphotidylglycerol and Phosphatidylethanolamine as major polar lipids. The G+C content of type strain UniB2T is 63.74 mol %.

The type strain UniB2^T (= MCC 3997 = KCTC 43095=JCM 34513) was isolated from digestive syrup from Pune, India.

Description of *Niallia alba* sp. nov.

Niallia alba (al'ba L. fem. adj. *alba*, white, the colour of colonies).

The cell is rod-shaped motile, aerobic, catalase and oxidase-positive, Gram-Stain variable in nature. The colonies are off white, circular, entire in the margin, flat and non-transparent. Rods are 0.8 to 1 µM in width and 2 to 4.5 µM in length. The optimum temperature for growth is 30 ^{0}C and can grow in a temperature range of 5 – 40^{0}C on a Nutrient Agar medium. Able to grow on pH ranges from 6.0-8.0 (Optimum 7.0) unit and salinity 0.5 - 2.0 % (Optimum 0.5 %). The cell is negative for glucose fermentation. Cells can hydrolyze P-nitrophenyl-βD-galactopyranoside and assimilates arabinose, mannose, N-acetyl, glucosamine, gluconate, malate, citrate. It shows naphthol AS-BI-phosphohydrolase, ß-galactosidase, ά-glucosidase enzyme activities. Strain UniB3^T can utilize a variety of sugars and chemicals like dextrin, D-maltose, D-cellobiose, D-turanose, stachyose, D-raffinose, α-D-lactose, β-methyl-D-glucoside, D-salicin, N-acetyl-D-glucosamine, N-acetyl-D-mannosamine, Inosine, D-sorbitol, D-mannitol, glycerol, D-gluconic Acid, methyl pyruvate, tween 40, acetoacetic acid, acetic acid. Sensitive to antibiotics and chemicals like troleandomycin, rifamycin SV, tetrazolium Blue, 1 % sodium lactate, nalidixic, lithium chloride, lithium chloride, aztreonam, sodium bromate. Major fatty acids were C16:0 (13.54 %), C15:0 anteiso (40.09 %) and C17:0 anteiso (16.03 %). Diphosphotidylglycerol and Phosphatidylethanolamine as major polar lipids. The G+C content of the type strain UniB3^T is 35.6 mol %.

The type strain UniB3^T (=MCC 3998 =KCTC 43235 =JCM 34492) was isolated from digestive syrup from Pune, India.

Declarations

Funding information

Authors acknowledge the project funding and fellowships to BT by the National Medicinal Plant Board (NMPB), Ministry of Ayush, Government of India under grant number Z.18017/187/CSS/R&D/KR-03/2019-20-NMPB-IV A; authors also acknowledge the funding by the Department of Biotechnology (DBT), Government of India under grant number BT/COORD.II/01/03/2016 (NCMR) used for in-house laboratory facilities. Authors gratefully acknowledge the University Grant Commission (UGC) of the

Government of India for providing CSIR-UGC NET-JRF fellowship to KK (Ref. No. 857/CSIR-UGC NET JUNE 2017).

Author contributions

VT, KK, PD and PK carried out the polyphasic taxonomy experiments; VT, KK, and BT did the phenotypic and genome data analysis and wrote the first draft of the manuscript. YS and SS provided the project funding and in house facilities. TL and AY did the formal analysis, reviewed, edited and finalized the manuscript. AY conceptualized, lead the investigation and provided the funding for the study. All the authors reviewed and approved the final version of the paper.

Compliance with ethical standards

Conflicts of interest

The authors declare that there are no conflicts of interest.

Ethical statement

No human or animal subjects were recruited in this study.

Abbreviations

OrthoANI	Orthologous average nucleotide identity
dDDH	Digital DNA-DNA hybridization
API	Analytical Profile Index
OGRI	Overall Genome Relatedness Index
DPG	Diphosphotidylglycerol
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PL	Unknown phospholipid lipid
APL	Unknown amino phospholipid
L	Unknown lipid.

References

Ash C, Priest FG, Collins MD (1993) Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Antonie Van Leeuwenhoek 64:253–260

Auch AF, Klenk H-P, Göker M (2010) Standard operating procedure for calculating genome-to-genome distances based on high-scoring segment pairs. Stand Genomic Sci 2:142–148

Ausubel FM, Brent R, Kingston RE, et al (1994) Current Protocols in Molecular Biology. New York: John Willey & Sons

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification. Can J Biochem Physiol 37:911–917

Brown J, Pirrung M, McCue LA (2017) FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. Bioinformatics 33:3137–3139

Card GL (1973) Metabolism of phosphatidylglycerol, phosphatidylethanolamine, and cardiolipin of Bacillus stearothermophilus. J Bacteriol 114:1125–1137

Chaudhari NM, Gupta VK, Dutta C (2016) BPGA-an ultra-fast pan-genome analysis pipeline. Sci Rep 6:1–10

Edgar R (2010) Usearch. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States)

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17:368–376

Fitch WM (1971) Toward defining the course of evolution: minimum change for a specific tree topology. Syst Biol 20:406–416

Gupta RS, Patel S, Saini N, Chen S (2020) 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 o. Int J Syst Evol Microbiol 70:5753–5798

Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075

Heulin T, Berge O, Mavingui P, et al (1994) *Bacillus polymyxa* and *Rahnella aquatilis*, the dominant N2-fixing bacteria associated with wheat roots in French soils. Eur J Soil Biol 30:35–42

Hwang Y-J, Ghim S-Y (2017) *Paenibacillus aceris* sp. nov., isolated from the rhizosphere of *Acer okamotoanum*, a plant native to Ulleungdo Island, Republic of Korea. Int J Syst Evol Microbiol 67:1039–1045

Hyatt D, Chen G-L, LoCascio PF, et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119

Jordan EO (1890) A report on certain species of bacteria observed in sewage. A Rep Biol Work Lawrence Exp Station Incl an Acc methods Employ results obtained Microsc Bacteriol Investig Sew water Rep water supply sewerage, part 2

Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874

Logan NA, Berge O, Bishop AH, et al (2009) Proposed minimal standards for describing new taxa of aerobic, endospore-forming bacteria. Int J Syst Evol Microbiol 59:2114–2121

Logan NA, De Vos P (2009) Genus I. *bacillus*. Bergey's Man Syst Bacteriol 3:21–128

Mavingui P, Heulin T (1994) In vitro chitinase and antifungal activity of a soil, rhizosphere and rhizoplane population of *Bacillus polymyxa*. Soil Biol Biochem 26:801–803

Meier-Kolthoff JP, Klenk H-P, Göker M (2014) Taxonomic use of DNA G+ C content and DNA–DNA hybridization in the genomic age. Int J Syst Evol Microbiol 64:352–356

Minnikin DE, O'donnell AG, Goodfellow M, et al (1984) An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. J Microbiol Methods 2:233–241

Na S-I, Kim YO, Yoon S-H, et al (2018) UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J Microbiol 56:280–285

Osman S, Satomi M, Venkateswaran K (2006) *Paenibacillus pasadenensis* sp. nov. and *Paenibacillus barengoltzii* sp. nov., isolated from a spacecraft assembly facility. Int J Syst Evol Microbiol 56:1509–1514. https://doi.org/10.1099/ijs.0.64085-0

Parks DH, Imelfort M, Skennerton CT, et al (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055

Parte AC, Carbasse JS, Meier-Kolthoff JP, et al (2020) List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol 70:5607

Piuri M, Sanchez-Rivas C, Ruzal SM (1998) A novel antimicrobial activity of a *Paenibacillus polymyxa* strain isolated from regional fermented sausages. Lett Appl Microbiol 27:9–13

Potter SC, Luciani A, Eddy SR, et al (2018) HMMER web server: 2018 update. Nucleic Acids Res 46:W200–W204

Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131. https://doi.org/10.1073/pnas.0906412106

Roux V, Fenner L, Raoult D (2008) *Paenibacillus provencensis* sp. nov., isolated from human cerebrospinal fluid, and *Paenibacillus urinalis* sp. nov., isolated from human urine. Int J Syst Evol Microbiol 58:682–687

Saitoh N (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 10:471–483

Sambrook J, Russell DW (2001) Molecular Cloning-Sambrook & Russel-Vol. 1, 2, 3. Cold Springs Harb Lab Press

Sasser M (2001) "Tracking" a Strain Using the Sherlock Microbial Identification System (MIS). Tech Note 102:

Simon L, Škraban J, Kyrpides NC, et al (2017) *Paenibacillus aquistagni* sp. nov., isolated from an artificial lake accumulating industrial wastewater. Antonie Van Leeuwenhoek 110:1189–1197

Singh HP, Singh TA (1993) The interaction of rockphosphate, Bradyrhizobium, vesicular-arbuscular mycorrhizae and phosphate-solubilizing microbes on soybean grown in a sub-Himalayan mollisol. Mycorrhiza 4:37–43

Vaz-Moreira I, Faria C, Nobre MF, et al (2007) *Paenibacillus humicus* sp. nov., isolated from poultry litter compost. Int J Syst Evol Microbiol 57:2267–2271. https://doi.org/10.1099/ijs.0.65124-0

Venkateswaran K, Kempf M, Chen F, et al (2003) *Bacillus nealsonii* sp. nov., isolated from a spacecraft-assembly facility, whose spores are gamma-radiation resistant. Int J Syst Evol Microbiol 53:165–172. https://doi.org/10.1099/ijs.0.02311-0

Wick RR, Judd LM, Gorrie CL, Holt KE (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595

Yoon J-H, Kang S-J, Yeo S-H, Oh T-K (2005) *Paenibacillus alkaliterrae* sp. nov., isolated from an alkaline soil in Korea. Int J Syst Evol Microbiol 55:2339–2344

Yun J-H, Lee J-Y, Kim PS, et al (2017) *Paenibacillus apis* sp. nov. and *Paenibacillus intestini* sp. nov., isolated from the intestine of the honey bee *Apis mellifera*. Int J Syst Evol Microbiol 67:1918–1924

Yoon S-H, Ha S-M, Kwon S, et al (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613

Zhang J, Ma X-T, Gao J-S, et al (2016) *Paenibacillus oryzae* sp. nov., isolated from rice roots. Int J Syst Evol Microbiol 66:5000–5004

Zhao Y, Wu J, Yang J, et al (2011) PGAP: pan-genomes analysis pipeline. Bioinformatics 28:416–418

Tables

Table 1: Differential phenotypic characteristic between strain UniB2^T and its closest phylogenetic neighbours. *Paenibacillus humicus* (NBRC 102415^T), *Paenibacillus pasadenensis* (NBRC 101214^T), +, Positive; -, negative; S, sensitive; R, resistant; ND, not determined. All data were obtained from this study.

Characteristics	UniB2^T	NBRC 102415^T	NBRC 101214^T
Gram Staining	Gram Variable	Gram stain Positive	Gram Stain Negative
Spore Formation	Yes	Yes	Yes
Motility	Motile	Motile	Motile
Salinity tolerance range for growth (%)	0- 2 %	ND	ND
PH range for Growth	6-8	ND	ND
Maximum growth temperature	40 ^{0}C	ND	ND
API 20NE test			
Hydrolysis of			
Gelatin	-	-	+
Arabinose	-	-	+
Gluconate	-	+	+
Malate	-	-	+
API ZYM test			
Esterase	-	+	+
Esterase Lipase	+	-	+
ά-galactosidase	+	-	-
ß-galactosidase	+	-	-
ά-glucosidase	+	+	-
ß-glucosidase	+	+	-
Biolog GEN III			
D-Sorbitol	+	+	-
L-Histidine	-	-	+
D-Gluconic Acid	-	+	-
Tween 40	-	+	-
N-Acetyl-D-Gelactosamine	-	+	+
3-Methyl Glucose	+	-	-
D-Fucose	+	-	-

L-Rhanmnose	+	-	-
Glycerol	-	+	+
L-Arginine	+	-	-
L-Aspartic Acid	+	-	-
Methyl Pyruvate	-	+	+
L-Malic Acid	-	+	+
Gentibiose	+	-	+
α-D-Glucose	+	-	+
L-Fucose	+	-	+
D-Fructose-6-PO4	+	-	+
D-Galacturonic Acid	+	-	+
L-Lactic Acid	-	-	+
D-Malic -Acid	-	-	+
Bromo-Succinic Acid	-	-	+
β-Hydroxy-D, L-Butyric Acid	+	-	+
Acetic Acid	+	-	+
Chemical Sensitivity Assay			
D-Serine	R	S	S
Rifamycin SV	S	R	R
Guanidine-HCl	R	S	S
Nalidixic Lithium Chloride	S	R	S
Lithium Chloride	S	R	R
Sodium Bromate	R	R	S

Table 2. Differential phenotypic characteristic between strain UniB3^T and its closest phylogenetic neighbours. *Niallia circulans* (DSM 11^T) and *Niallia nealsonii* (DSM 15077^T). +, Positive; -, negative; ND, not determined. All data were obtained from this study.

Characteristics	UniB3^T	DSM 11^T	DSM 15077^T
Gram stain	Gram Variable	Gram Variable	Gram-Positive
Spore Formation	Yes	Yes	Yes
Motility	Motile	Motile	Motile
pH range for growth	6–8	ND	ND
Maximum growth temperature	40 ^{0}C	ND	ND
Salinity tolerance range for growth (%)	0–2	ND	ND
API ZYM Test			
ß-galactosidase	+	-	-
ά-glucosidase	+	-	-
API 20NE Test			
Malate Assimilation	+	-	-
P-nitrophenyl-βD-galactopyranoside Hydrolysis	+	+	-
Arabinose Assimilation	+	-	-
Mannose Assimilation	+	-	+
N-acetyl-glucosamine Assimilation	+	-	+
Gluconate Assimilation	+	+	-
Citrate Assimilation	+	+	-
Biolog GEN III Tests			
D-Glucose-6-PO4	-	+	+
Dextrin	+	+	-
D-Maltose	+	+	-
D-Cellobiose	+	+	-
D-Turanose	+	+	-
Stachyose	+	+	-
D-Raffinose	+	+	-
α-D-Lactose	+	+	-
β-Methyl-D-Glucoside	+	+	-
D-Salicin	+	+	-

N-Acetyl-D-Glucosamine	+	+	-
N-Acetyl-D-Mannosamine	+	+	-
N-Acetyl-D-Gelactosamina	-	-	+
3-Methyl Glucose	-	-	+
D-Fucose	-	-	+
L-Fucose	+	-	+
L-Rhanmnose	+	-	+
Inosine	+	+	-
D-Sorbitol	+	+	-
D-Mannitol	+	+	-
Myo-Inositol	-	+	-
Glycerol	+	+	-
L-Alanine	-	+	-
D-Gelacturonic Acid	-	-	+
D-Gluconic Acid	+	+	-
D-Glucuronic Acid	+	-	+
Mucic Acid	-	-	+
Phenylacetic Acid	-	-	-
Methyl Pyruvate	+	+	-
L-Malic Acid	-	+	-
Tween 40	+	+	-
β-Hydroxy-D, L-Butyric Acid	-	-	+
Acetoacetic Acid	+	+	-
Acetic Acid	+	+	-
Chemical Sensitivity Assay			
Troleandomycin	S	R	R
Rifamycin SV	S	R	R
1 % Sodium Lactate	S	S	R
Monocycline	R	R	R

Guanidine-HCl	R	S	R
Tetrazolium Blue	S	S	R
Nalidixic Lithium Chloride	S	S	R
Lithium Chloride	S	S	R
Aztreonam	S	S	R
Sodium Bromate	S	S	R

Table 3. Cellular fatty acid content (%) of strain and its closest phylogenetic neighbours. *Paenibacillus humicus* (NBRC 102415^T), *Paenibacillus pasadenensis* (NBRC 101214^T), -, not determined. All data were obtained from this study.

Fatty Acids	UniB2^T	NBRC 102415^T	NBRC 101214^T
Saturated			
$C_{12:00}$	0.2	-	-
$C_{14:00}$	1.01	1.39	1.1
$C_{16:00}$	13.97	9.4	11.19
$C_{18:00}$	0.35	-	-
Branched			
$C_{14:0}$ iso	0.61	-	1.5
$C_{15:0}$ iso	9.05	9.32	5.32
$C_{16:0}$ iso	8.57	2.96	14.56
$C_{17:0}$ iso	8.92	4.56	7.11
$C_{11:0}$ anteiso	0.28	-	-
$C_{13:0}$ anteiso	0.235	-	-
$C_{15:0}$ anteiso	39.77	57.47	40.26
$C_{17:0}$ anteiso	15.58	13.71	18.95
$C_{17:1}$ anteiso ω9c	0.21	-	-
Unsaturated			
$C_{16:1}$ ω11c	0.52	-	-
$C_{18:1}$ ω9c	0.23	-	-
$C_{20:1}$ ω7c	0.22	1.19	-
$C_{16:1}$ ω7c alcohol	0.2	-	-

Table 4. Cellular fatty acid content (%) of strain and its closest phylogenetic neighbours. *Niallia circulans* (DSM 11^T) and *Niallia nealsonii* (DSM 15077^T), -, not determined. All data were obtained from

this study.

Fatty acids	UniB3^T	DSM 11^T	DSM 15077^T
Saturated			
C12:0	0.17	_	
C14:0	0.98	8.42	9.3
C16:0	13.54	11.9	22.04
C18:0	0.35	0.35	0.48
Branched			
C13:0 iso	-	0.29	0.67
C14:0 iso	0.6	8.45	8.63
C15:0 iso	8.98	11.69	12.93
C16:0 iso	8.09	14.78	8.16
C17:0 iso	9.01	2.87	1.95
C18:0 iso	_	0.28	
C11:0 anteiso	0.3		
C13:0 anteiso	0.24	_	0.32
C15:0 anteiso	40.09	31.2	27.99
C17:0 anteiso	16.03	9.03	6.75
Unsaturated			
C16:1 w11c	0.51	_	
C17:1 w9c	0.24	_	
C16:1 w7c	0.18	_	_
C18:1 w9c	0.24	_	0.43
C20:1 w7c	0.44	0.21	
Sum In Feature 2*	_	0.31	_
Sum In Feature 3*	_	0.2	0.34
Summed Feature 2*	_	0.31	_
Summed Feature 3*	_	0.2	0.34

*Sum In Feature 2, unknown 10.9525; Sum In Feature 3, 16:1 w6c/16:1 w7c; Summed Feature 2, 14:0 3OH/16:1 iso I; Summed Feature 3, 16:1 w7c/16:1 w6c

Figures

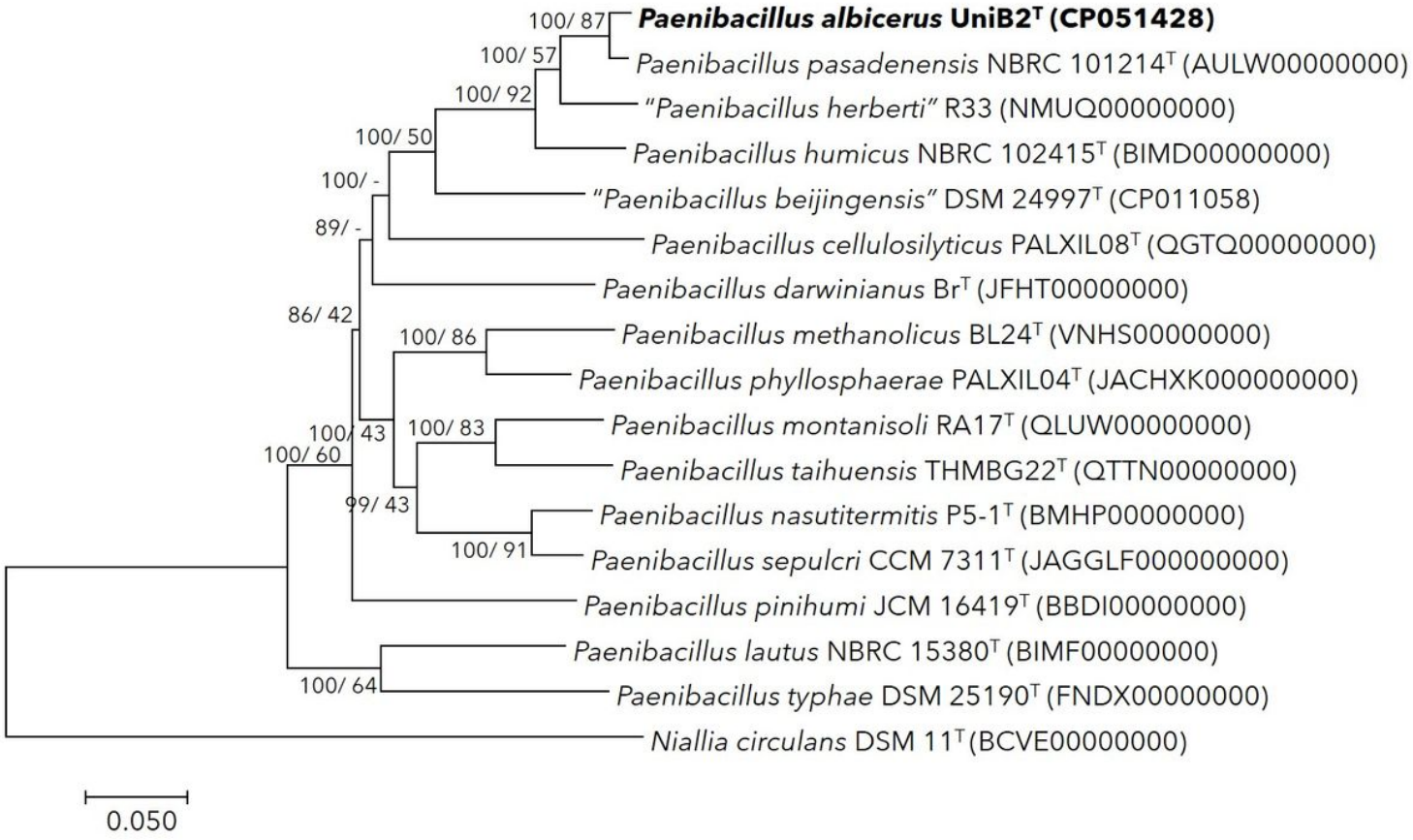

Figure 1

The combined pan-genome phylogenetic tree of strain UniB2T based on orthologous protein and gene sequences derived using BPGA and UBCG tools. The BPGA- and UBCG-based phylogenetic trees were reconstructed in MEGA 7 using the neighbour-joining method and maximum likelihood methods, respectively, at 1000 bootstrap replicates. Figures at branch points are bootstrap values seen in trees obtained using BPGA and UBCG, respectively. The respective sequences of Niallia circulans ATCC 4513 (BCVE00000000) were used as an outgroup. Bar indicates the number of substitutions per site.

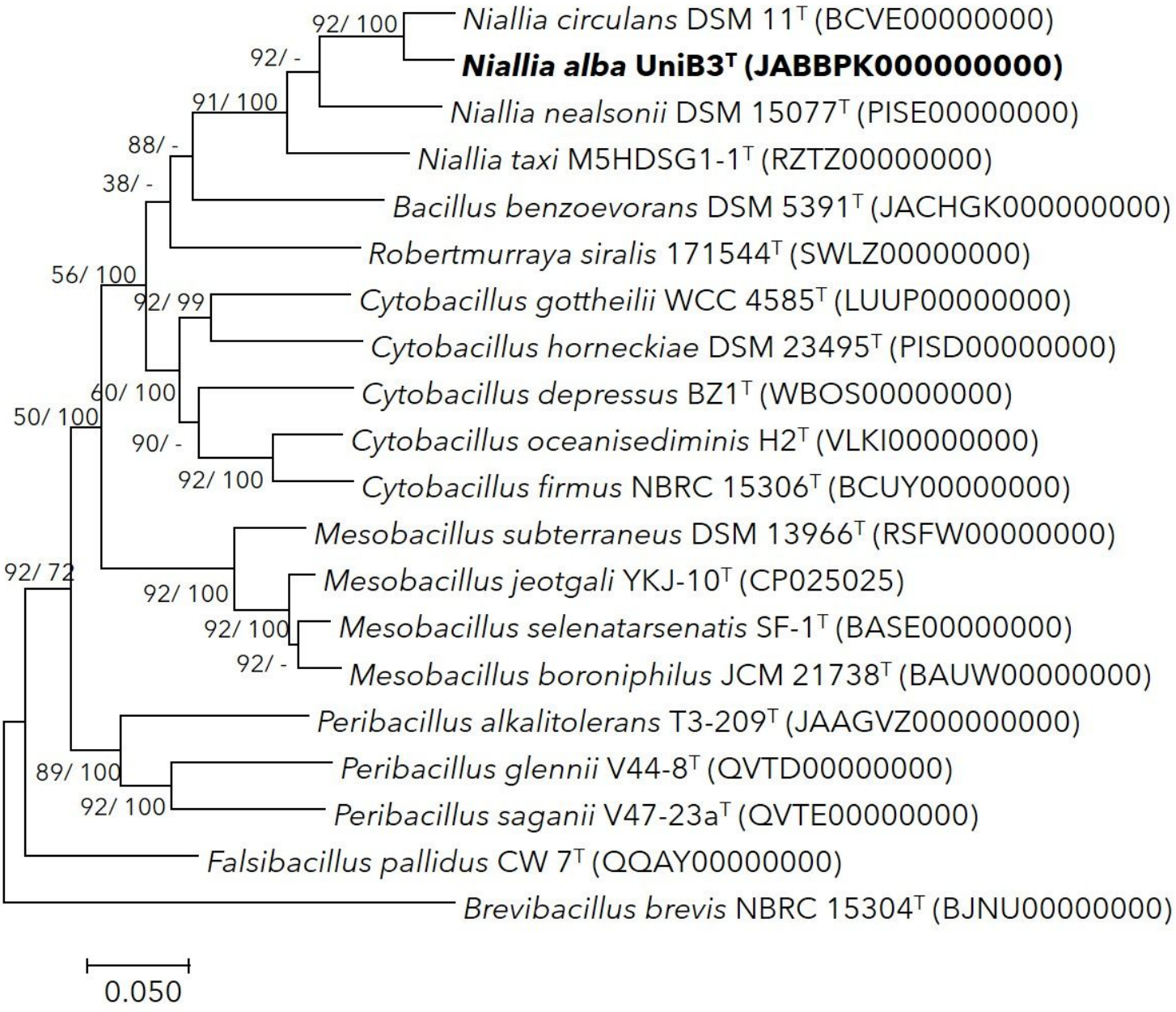

Figure 2

The combined pan-genome phylogenetic tree of strain UniB3T based on orthologous protein and gene sequences derived using BPGA and UBCG tools. The BPGA- and UBCG-based phylogenetic trees were reconstructed in MEGA 7 using the neighbour-joining method and maximum likelihood methods, respectively, at 1000 bootstrap replicates. Figures at branch points are bootstrap values seen in trees obtained using UBCG and BPGA, respectively. The respective sequences of Brevibacillus brevis NBRC 15304 (BJNU00000000) were used as an outgroup. Bar indicates the number of substitutions per site.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 1BacillusUniB3MCC3998.pdf
- 2BacillusUniB3KCTC43235.pdf
- 3BacillusUniB3JCM34492.pdf
- 4PaenibacillusUniB2MCC3997.pdf
- 5PaenibacillusUniB2KCTC43095.pdf
- 6PaenibacillusUniB2JCM34513.pdf
- SupplementaryFigures.pptx