Systematic and Applied Microbiology 42 (2019) 488–494

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: http://www.elsevier.com/locate/syapm

Pantoea endophytica sp. nov., novel endophytic bacteria isolated from maize planting in different geographic regions of northern China

Jun-lian Gao [a], Jing Xue [a], Hui Yan [c], Shuai Tong [b], Mohammad Sayyar Khan [d], Li-wei Wang [a], Xiao-jie Mao [b], Xiuhai Zhang [a,*], Jian-guang Sun [b,*]

[a] *Beijing Agro-Biotechnology Research Center, Beijing Academy of Agriculture and Forestry/Beijing Key Laboratory of Agricultural Genetic Resources and Biotechnology, Beijing 100097, PR China*
[b] *Key Laboratory of Microbial Resources, Ministry of Agriculture and Rural Affairs/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, PR China*
[c] *College of Animal Science and Technology, Hebei Agricultural University, Baoding 071001, PR China*
[d] *Genomics and Bioinformatics Laboratory, Institute of Biotechnology and Genetic Engineering (IBGE), Faculty of Crop Production Sciences, The University of Agriculture, Peshawar 25000, Pakistan*

ARTICLE INFO

Article history:
Received 31 January 2019
Received in revised form 24 May 2019
Accepted 1 June 2019

Keywords:
Pantoea endophytica
Zea mays L.
Phylogeny
Genome
Polyphasic approach

ABSTRACT

Four endophytic bacterial strains were isolated from root, stem and leaf of maize planted in different regions of northern China. The four strains possessed almost identical 16S rRNA gene sequences. However, REP-PCR fingerprint patterns discriminated that they were not from one clonal origin. Furthermore, the average nucleotide identity (ANI) values among them were higher than 95%, suggesting they all belong to one species. Based on 16S rRNA gene phylogeny, the four strains were clustered together with *Pantoea rodasii* LMG 26273^T and *Pantoea rwandensis* LMG 26275^T, but on a separate branch. Multilocus sequence analysis (MLSA) indicated that the four strains form a novel *Pantoea* species. Authenticity of the novel species was confirmed by ANI comparisons between strain 596^T and its closest relatives, since obtained values were considerably below the proposed thresholds for the species delineation. The genome size of 596^T was 5.1Mbp, comprising 4896 predicted genes with DNA G + C content of 57.8 mol%. The respiratory quinone was ubiquinone-8 (Q-8) and the polar lipid profile consisted of phosphatidylethanolamin, diphosphatidylglycerol, phosphatidylglycerol, unidentified aminophospholipid and unidentified phospholipid. The major fatty acids of strain 596^T were $C_{16:0}$, summed feature 2 ($C_{12:0}$ aldehyde), summed feature 3 ($C_{16:1}\omega7c$ and/or $C_{16:1}\omega6c$) and summed feature 8 ($C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$). Based on phylogenetic, genomic, phenotypic and chemotaxonomic data, the four isolates are considered to represent a novel species of the genus *Pantoea*, for which the name *Pantoea endophytica* sp. nov., is proposed, with 596^T (= DSM 100,785^T = CGMCC 1.15280^T) as type strain.

The genus *Pantoea* belongs to the family *Erwiniaceae* [1] and was proposed by Gavini et al. [21] with *Pantoea agglomerans* as the type species and the genus has been subsequently emended over the years as more species have been classified [4,34]. The taxonomic status of members of genus *Pantoea* have undergone some changes, three *Pantoea* species: *Pantoea citrea*, *Pantoea punctata* and *Pantoea terrea* have been reclassified as belonging to the genus *Tatumella* [8]; four *Pantoea* species: *Pantoea calida*, *Pantoea gaviniae*, *Pantoea intestinalis* and *Pantoea theicola* have been moved to the genus *Mixta* [36]. Meanwhile *Pectobacterium cypripedii* (Hori 1911) was reclassified as *Pantoea cypripedii* [4]. At the time of writing, this genus comprises 19 recognized species (http://www.bacterio.net/Pantoea.html). Members of the genus *Pantoea* are isolated from a number of diverse habitats, such as plants, soil, water, animals [21], trees [5,7], plant seeds [6,23], black tea [40], human gut [37], the fruiting body of mushroom [31] and human clinical samples [4]. Some *Pantoea* species are known to interact with plants and may confer beneficial or deleterious effects to their hosts [11]. For example, *P. agglomerans* was initially identified as the causative agent of disease in a wide range of hosts, such as onion, gypsophila, beet and wisteria, beach pea, cotton, rice, maize [3,7,17,21,34], but many strains belonging to this species also occur as plant epiphytes or endophytes and exhibit plant growth promotion or biocontrol effects [18,29,35]. Several *Pantoea* species *P. vagans*, *P. agglomerans*, *P. ananatis*, *Pantoea stewartii* ssp. *stewartii* and ssp. *indologenes* were

---

[*] Corresponding authors.
E-mail addresses: zhangxiuhai@baafs.net.cn (X. Zhang), sunjianguang@caas.cn (J.-g. Sun).

https://doi.org/10.1016/j.syapm.2019.06.001

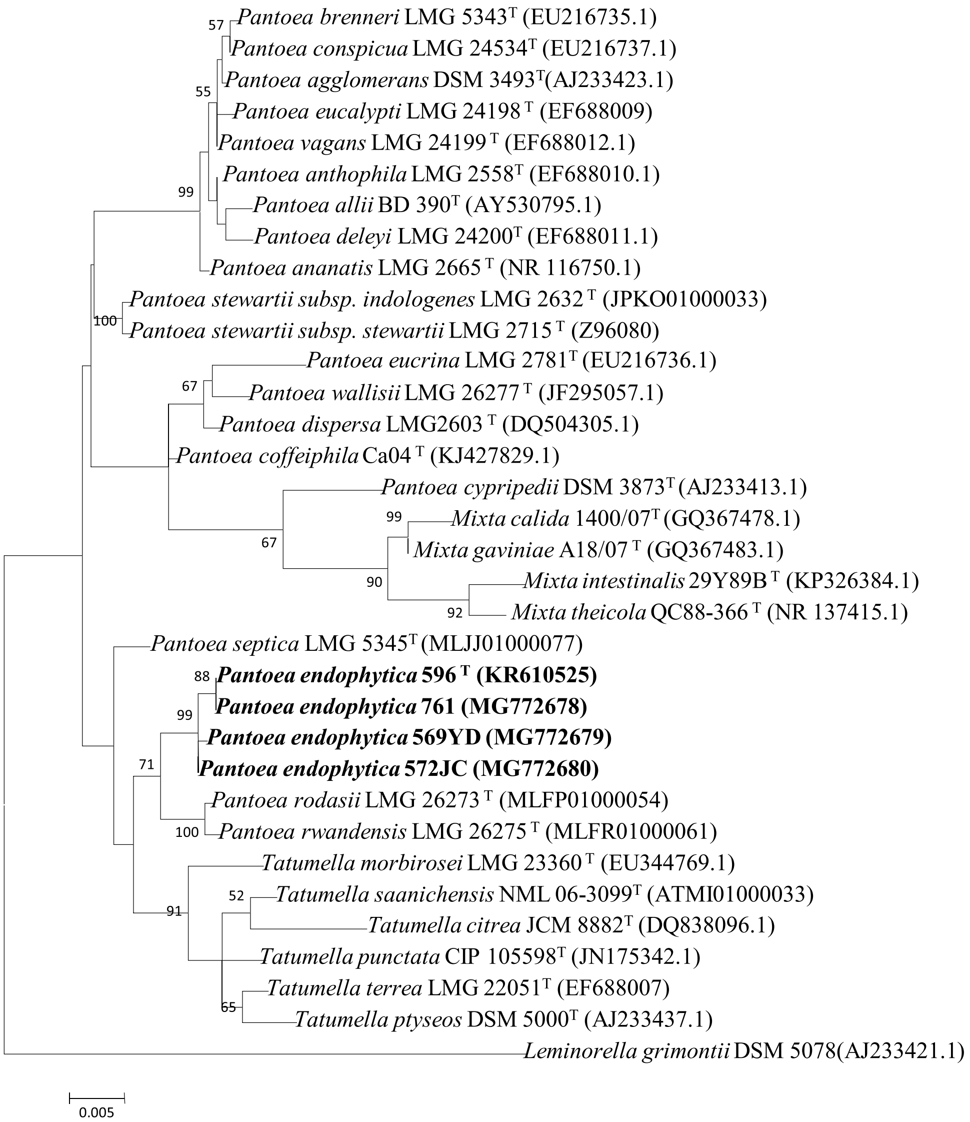

Fig. 1. Maximum-likelihood (ML) phylogenetic tree based on 16S rRNA gene sequences showing the relationship among strains of *Pantoea endophytica* sp. nov. (boldfaced) and related species within the genera *Pantoea, Mixta and Tatumella*. Bootstrap analyses were determined based on 1000 resamplings; values >50% were shown.The16S rRNA gene sequence of *Leminorella grimontii* DSM 5078 was used as an outgroup. Bar, 0.005 substitutions per nucleotide position.

isolated from maize [7,16,34]. In a diversity survey of maize endophytic bacteria in northern China, four endophytic bacterial strains were isolated from maize planted in different regions of northern China. In the present study, the four strains were taxonomically compared with the type strains of related *Pantoea* species by a polyphasic approach. As result, a novel species, *Pantoea endophytica* sp. nov., is proposed.

The four strains were isolated from internal plant tissue of maize in different geographic regions of northern China: strains 596^T and 761 from root of maize (*Zea mays* L., Wantian) in Beijing, strain 569YD from leaf of maize (*Zea mays* L., Jinqing 708) in Tongliao City of Inner Mongolia, strain 572JC from stem of maize (*Zea mays* L., Nendan No.1) in Suihua City of Heilongjiang Province. The isolation of endophytic bacterial strains was performed according to the procedure as described by Gao et al. [20].The four strains were originally obtained with different nitrogen-free mediums, strain 596^T and 569YD from Döbereiner nitrogen-free agar [30], meanwhile strain 572JC and 761 from CCM nitrogen-free agar [24]. Later their routine cultivation was performed on LB medium or trypticase soy agar (Becton Dickinson) at 30 °C, for 3 days. They were stored in 30% (w/v) glycerol at −80 °C for a longer storage. The closely related species to strain 596^T, namely *Pantoea rodasii* DSM 26611^T and *Pantoea rwandensis* DSM 105076^T determined by 16S rRNA gene sequence similarity were used as reference strains under the same laboratory conditions.

The 16S rRNA genes were amplified by colony PCR, for construction of the DNA template for 16S rRNA gene amplification, a loop-full of biomass was scraped off the R2A agar plate, suspended in 100 μl ddH$_2$O and lysed by 10 min boiling and afterwards 5 min freezing. After centrifugation, the supernatant was used as the template for PCR. The 16S rRNA gene was amplified using the universal primers 27 F and 1492R [28]. Purified PCR products were sequenced using an ABI 3730 sequencer following the manufacturer's protocols. The 16S rRNA gene sequences of the four strains were determined by direct sequencing and compared against related available 16S rRNA gene sequences in the EzBioCloud database (http://www.ezbiocloud.net/) [43]. In this study, the genome sequences of the four novel strains were determined following the method described below. The four house-keeping genes (*atpD*, *gyrB*, *infB* and *rpoB*) of the four strains were extracted by Rapid Annotation System Technology (RAST) from their genome sequences [9]. The four genes were aligned against related avail-

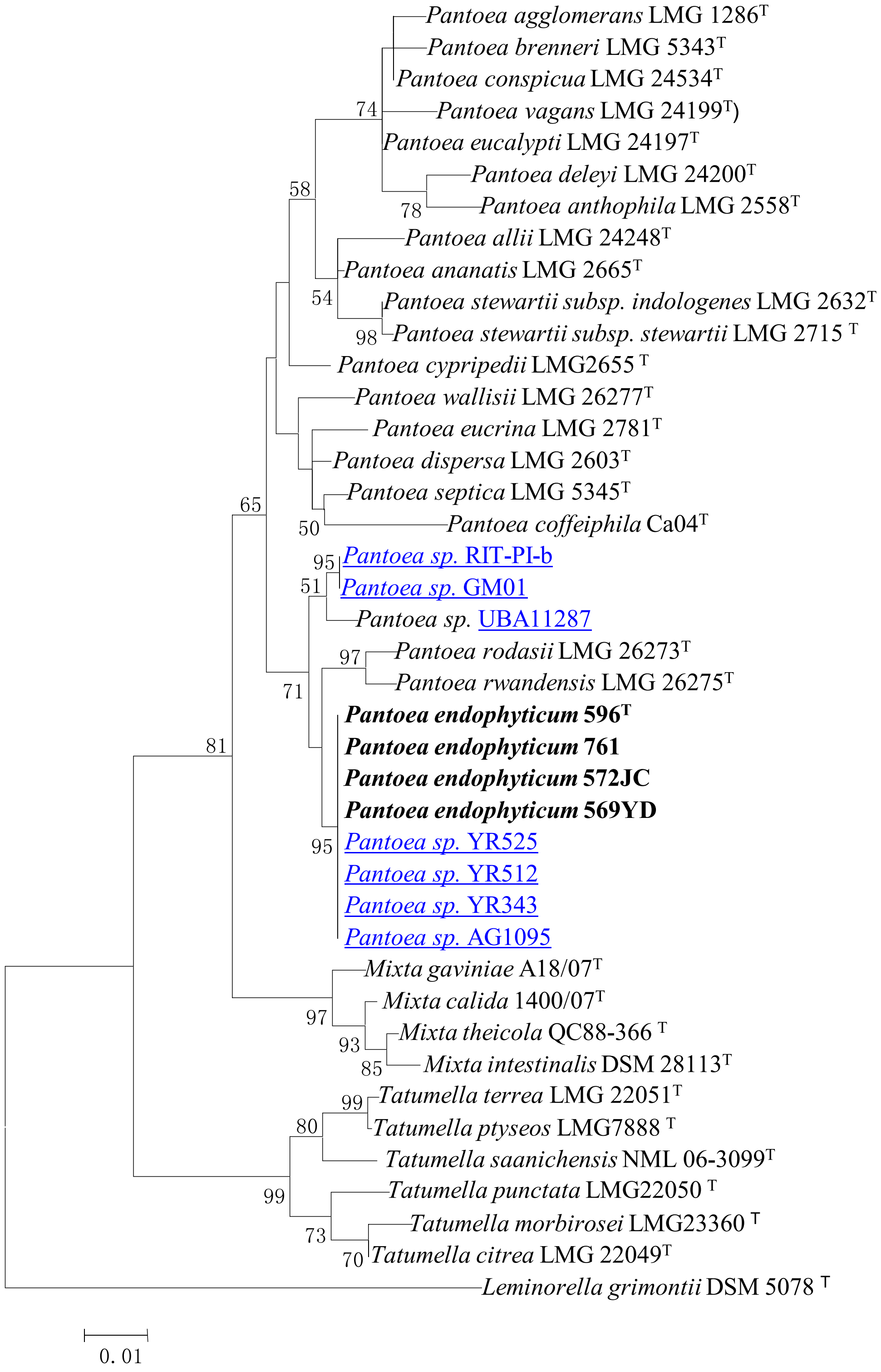

Fig. 2. Maximum likelihood (ML) phylogenetic tree based on concatenated amino acid sequences of AtpD, GyrB, InfB and RpoB showing the relationship among strains of *Pantoea endophytica* sp. nov. (boldfaced) and related species within the genera *Pantoea*, *Mixta* and *Tatumella*. Bootstrap analyses were determined based on 1000 resamplings; values >50% were shown. *Leminorella grimontii* DSM 5078 was used as an outgroup. Bar, 0.01 substitutions per nucleotide position.

able *atpD*, *gyrB*, *infB* and *rpoB* gene sequences in the NCBI GenBank database (http://www.ncbi.nlm.nic.gov/) and translated into protein sequences using the software package of MEGA version 7.0 [26]. The alignments were trimmed to the following protein lengths: AtpD – 219 aa, GyrB – 195 aa, InfB – 204 aa and RpoB – 212 aa which were subsequently concatenated for multilocus sequence analysis (MLSA). Phylogenetic trees based on individual and concatenated sequences were constructed using MEGA 7.0 [26] after multiple alignment of data by CLUSTAL X program [41]. The phylogenetic trees were reconstructed with the maximum-likelihood method [17]. In each case bootstrap values were calculated based on 1000 replications. Nearest-neighbor-interchange was applied to the maximum-likelihood analysis. The strains 596^{T}, 761, 569YD and 572JC possessed almost identical 16S

Table 1
Phenotypic characteristics that differentiate the four strains of *Pantoea endophytica* sp. nov and the closely related species *Pantoea rodasii* DSM 26611^{T} and *Pantoea rwandensis* DSM 105076^{T}.

Characteristics	Strain					
	1	2	3	4	5	6
Oxidation of (Biolog GENIII):						
Dextrin	–	–	–	–	+	+
D-Maltose	+	+	+	+	–	+
D-Trehalose	+	+	–	+	–	+
D-Cellobiose	–	–	–	–	–	+
Gentiobiose	+	+	–	+	–	+
Sucrose	–	–	–	–	–	+
ß-Methyl-D-glucoside	+	+	+	+	–	+
D-salicin	+	+	+	+	–	+
N-Acetyl-D-glucosamine	+	+	+	+	–	+
N-Acetyl-ß-D-mannosamine	+	+	+	+	–	+
N-Acetyl-D-galactosamine	–	–	–	–	–	+
α-D-Glucose	+	+	+	+	–	+
D-Mannose	–	+	–	+	–	+
D-Fructose	–	–	–	+	–	–
L-Rhamnose	+	+	+	+	–	+
Inosine	–	–	+	–	–	–
D-Sorbitol	–	–	–	–	–	+
D-Mannitol	+	+	–	+	–	+
D-Arabitol	+	+	+	+	–	+
myo-Inositol	+	+	+	+	+	–
D-Glucose-6-PO_4	–	+	+	+	–	+
Troleandomycin	+	+	+	+	–	+
L-Alanine	+	+	+	+	–	+
L-Arginine	+	+	+	+	–	–
L-Aspartic acid	+	+	+	+	–	+
L-Glutamic acid	+	+	+	+	–	+
L-Histidine	+	+	+	+	–	–
L-Serine	–	–	–	+	–	+
Lincomycin	+	+	+	+	–	+
Pectin	–	–	–	–	+	+
Glucuronamide	–	–	–	+	+	–
Mucic acid	+	+	+	+	–	+
Quinic acid	–	+	+	+	+	–
D-Saccharic acid	+	+	+	+	–	+
Methyl pyruvate	+	+	+	–	–	+
L-Lactic acid	+	+	–	–	+	+
Citric acid	+	+	+	–	–	+
Bromo-succinic acid	–	+	+	+	–	–
Lithium chloride	+	–	+	–	–	+
γ-Amino-butryric acid	+	+	+	+	+	–
Acetic acid	+	+	+	+	+	–
Aztreonam	–	+	+	+	–	+
Acid production from (API 50CH):						
Erythritol	+	+	+	+	(+)	–
D-Arabinose	–	+	(+)	+	(+)	–
L-Arabinose	+	+	+	+	+	(+)
L-Xylose	–	–	–	–	(+)	–
D-Adonitol	–	–	–	–	+	–
Dulcitol	+	–	–	–	(+)	+
Inositol	+	+	+	+	+	–
D-Sorbitol	–	–	–	–	–	+
Methyl-α-D-mannopyranosid	–	–	–	–	(+)	–
Methyl-α-D-glucopyranosid	–	–	–	–	(+)	–
N-Acetylglucosamine	–	–	–	–	+	(+)
Amygdalin	–	–	–	–	(+)	(+)
Arbutin	+	+	+	+	+	(+)
D-Lactose	(+)	–	–	–	(+)	+
D-Melibiose	+	+	+	+	+	–
D-Saccharose	+	+	–	–	+	(+)
D-Trehalose	+	+	+	+	+	(+)
Inulin	–	+	–	–	(+)	–
Gentiobiose	+	+	(+)	(+)	+	(+)
D-Turanose	+	+	–	–	–	–
D-Lyxose	+	+	+	+	(+)	+
D-Tagatose	(+)	–	–	–	–	–
D-Fucose	+	+	+	+	+	–
L-Fucose	(+)	–	–	–	(+)	–
D-Arabitol	+	+	+	+	+	–
Potassium 5-ketogluconate	–	–	–	–	+	–
Enzyme production (API ZYM):						
Esterase (C 4)	+	(+)	+	(+)	+	(+)
Esterase Lipase (C 8)	+	+	+	+	+	(+)
Lipase (C 14)	+	–	–	–	–	–
Valine arylamidase	–	–	–	–	(+)	+
Cystine arylamidase	–	–	–	–	–	+
Trypsin	–	–	–	–	–	+
ß-galactosidase	–	–	–	–	+	+
Reaction of (API 20E):						
Arginine dihydrolase	+	–	–	+	–	+
Tryptophane deaminase	+	(+)	(+)	(+)	+	(+)
Citrate utilization	+	+	+	+	–	+
Fermentation-oxidation of :						
Inositol	+	+	+	+	+	–
D-Sorbitol	+	+	+	+	–	+
D-Melibiose	+	+	+	+	+	(+)

Species: 1, Strain 596^{T}; 2, Strain 761; 3, Strain 569 YD; 4, Strain 572 JC; 5, *Pantoea rodasii* DSM 26611^{T}; 6, *Pantoea rwandensis* DSM 105076^{T}.
Note: All data are from this study. +, positive; (+), weak positive; –, negative.

Table 2
Cellular fatty acid composition (%) of the four strains of *Pantoea endophytica* sp. nov. and the closely related species *Pantoea rodasii* DSM 26611^{T} and *Pantoea rwandensis* DSM 105076^{T}.

Fatty acid	Strain					
	1	2	3	4	5	6
$C_{10:0}$	0.1	0.1	0.1	0.1	0.1	0.1
$C_{12:0}$	7.9	3.9	3.8	3.8	5.7	6.7
$C_{12:0}$ 3-OH	0.1	0.1	0.1	0.1	0.1	0.1
$C_{13:0}$	0.4	0.3	0.3	0.3	0.6	0.2
$C_{14:0}$	9.7	6.5	5.8	5.8	9.7	8.5
$C_{15:0}$ 3-OH	0.1	0.1	0.1	ND	0.2	ND
$C_{14:0}$2-OH	ND	ND	ND	ND	ND	0.4
$C_{16:0}$	24.4	31.7	32.7	32.1	26.0	25.5
$C_{16:1}\omega5c$	ND	0.1	0.1	0.1	0.1	0.1
$C_{17:0}$	0.5	1.0	0.9	1.1	0.8	0.5
$C_{17:0}$ cyclo	8.7	9.8	10.6	10.8	6.7	11.2
$C_{17:0}$ 3-OH	0.1	ND	0.1	0.2	ND	0.2
$C_{17:1}\omega8c$	0.1	0.2	0.2	0.2	0.2	ND
$C_{18:0}$	0.2	0.2	0.3	0.3	0.2	0.1
$C_{19:0}$ cyclo $\omega8c$	0.5	0.6	0.8	0.9	0.5	0.4
iso-$C_{19:0}$	ND	ND	0.1	0.1	ND	0.1
Summed feature 1[a]	0.3	ND	ND	ND	0.5	ND
Summed feature 2[a]	19.1	8.3	8.7	8.6	16.3	14.4
Summed feature 3[a]	15.8	20.6	19.2	18.9	23.0	16.2
Summed feature 8[a]	11.8	16.6	16.5	16.8	9.5	15.3

Strain number: 1, Strain 596^{T}; 2, Strain 761; 3, Strain 569 YD; 4, Strain 572 JC; 5, *Pantoea rodasii* DSM 26611^{T}; 6, *Pantoea rwandensis* DSM 105076^{T}. All data are from this study. Cells of all strains were harvested after growth on TSA medium at 30 for 48 h.Values are percentages of components of each fatty acid in total fatty acids; ND, not detected.

[a] Summed features are groups of two or three fatty acids that are treated together for the purpose of evaluation in the MIDI system and include both peaks with discrete ECLs as well as those where the ECLs are not reported separately. Summed feature 1 was listed as $C_{15:1}$ iso H and/or $C_{13:0}$ 3-OH; summed feature 2 was listed as $C_{12:0}$ aldehyde; summed feature 3 was listed as $C_{16:1}\omega6c$ and/or $C_{16:1}\omega7c$; summed feature 8 was listed as $C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$.

rRNA gene sequences (share 99.9–100% sequence similarities).The maximum-likelihood phylogenetic tree (Fig. 1) constructed from 16S rRNA gene sequences of the four new isolates and related reference species in the genera *Pantoea*, *Mixta and Tatumella* indicated that the four new isolates were clustered together and were closely related to *P. rodasii* LMG 26273^{T} and *P. rwandensis* LMG 26275^{T} with sequence similarities of (98.6–98.7%) and (98.2–98.3%), respectively. The novel isolates 596^{T}, 761, 569YD and 572JC possessed very similar sequences for each of the protein-encoding genes studied and formed a separate phylogenetic lineage in the MLSA phylogenetic tree (Fig. 2).The similarity between their concatenated amino acid sequences of AtpD, GyrB, InfB and RpoB and that

Table 3
Description of *Pantoea endophytica* sp. nov. according to Digital Protologue TA00773 assigned by the www.imedea.uib.es/dprotologue website.

Taxonumber	TA00773
Species name	*Pantoea endophytica*
Genus name	*Pantoea*
Specific epithet	*endophytica*
Species status	sp. nov.
Species etymology	en.do.phy'ti.ca. Gr. pref. endo-, within; Gr. n. phyton, plant; L. fem. suff. -ica, adjectival suffix used with the sense of belonging to; N.L. fem. adj. endophytica, within plant, endophytic
Designation of the type strain	596
Strain collection numbers	CGMCC 1.15280 = DSM 100785
16S rRNA gene accession number	KR610525
Genome accession number [RefSeq]	PJRT00000000
Genome status	Draft
Genome size	5165kbp
GC mol%	57.8
Country of origin	China
Region of origin	Beijing
Date of isolation	2013-07-26
Source of isolation	Root of Maize
Sampling date	2013-06-25
Geographic location	Fangshan District of Beijing
Number of strains in study	4
Source of isolation of non-type strains	Root of Maize
Growth medium, incubation conditions [temperature, pH, and further information] used for standard cultivation	Good growth occurs on Luria–Bertani (LB) agar, tryptic soy agar (TSA), nutrient agar (NA), R2A agar and Döbereiner nitrogen-free medium at 30 °C
Conditions of preservation	In 30% (v/v) glycerol and stored at −80 °C
Gram stain	Negative
Cell shape	Rod
Cell size	Cells are approximately 0.7–0.9 µm in width and 1.4–1.7 µm in length, after 25 h of incubation at 30 °C in LB broth.
Colony morphology	Colonies are circular with entire margins, yellow, convex, glossy and 2.0–3.0 mm in diameter on trypticase soy agar (TSA) plates after 3 days of incubation at 30 °C.
Temperature range for growth(°C)	Growth occurs at 5–45 °C, with optimal growth at 30 °C
pH range for growth	Growth occurs at pH 5.0–9.0, with optimal growth at pH 7.0
NaCl range for growth (%, w/v)	Growth occurs at a NaCl concentration of 0–7.0% (w/v)
Relationship to O_2	aerobic
Tests with API 50CH	Acid is produced from glycerol, erythritol, L-arabinose, D-ribose, D-galactose, D-xylose, D-glucose, D-fructose, D-mannose, L-rhamnose, dulcitol, inositol, D-mannitol, arbutin, esculin ferric citrate, salicin, D-cellobiose, D-maltose, D-melibiose, sucrose, D-trehalose, gentiobiose, D-lyxose, D-fucose, D-turanose and D-arabitol; and weakly produced from D-lactose, D-tagatose and L-fucose; but not produced from D-raffinose, starch, D-arabinose, glycogen, amygdalin, L-xylose, D-adonitol, methyl-β-D-xylopyranoside, L-sorbose, D-sorbitol, methyl-α-D-mannopyranoside, methyl-α-D-glucopyranoside, inulin, D-melezitose, xylitol, L-arabitol, potassium gluconate, potassium 2-ketogluconate, potassium 5-ketogluconate and N-acetylglucosamine.
Tests with API 20E	Positive for β-galactosidase, arginine dihydrolase, citrate utilization, tryptophane deaminase, Voges–Proskauer reactions/ acetoin production, fermentation - oxidation of D-glucose, D-mannitol, inositol, D-sorbitol, L-rhamnose, D-saccharose, D-melibiose, amygdalin and L-arabinose; but negative for lysine decarboxylase, ornithine decarboxylase, H_2S production, urease, indole production and gelatinase.
Tests with API ZYM	Positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine arylamidase, acid phosphatase, naphthol-AS-BI-phosphohydrolase, β-glucosidase and N-acetyl-β-glucosaminidase activities; but negative for valine arylamidase, cystine arylamidase, trypsin, β-galactosidase, β-glucuronidase, α-glucosidase, α-mannosidase, α-chymotrypsin, α-galactosidase and α-fucosidase activities.
Tests with Biolog GEN III	Positive for utilization of D-maltose, D-trehalose, gentiobiose, β-methyl-D-glucoside, D-salicin, N-Acetyl-D-glucosamine, N-acetyl-β-D -mannosamine, α-D-glucose, D-fructose, D-galactose, L-rhamnose, 1% sodium lactate, D-mannitol, D-arabitol, myo-inositol, glycerol, D-fructose-6-PO4,troleandomycin, rifamycin SV, glycyl-L-prolin, L-alanine, L-arginine, L-aspartic acid, L-glutamic acid, L-histidine, lincomycin, guanidine HCl, niaproof 4, D-galacturonic acid, L-galactonic acid lactone, D-gluconic acid, D-glucuronic acid, mucic acid, D-saccharic acid, vancomycin, tetrazolium violet, tetrazolium blue, methyl pyruvate, L-lactic acid, citric acid, L-malic acid, lithium chloride, γ-amino-butryric acid and acetic acid; but negative for utilization of dextrin, D-cellobiose, sucrose, D-turanose, stachyose, D-raffinose, α-D-lactos, D-melibiose, N-acetyl-D-galactosamine, N-acetyl -neuraminic acid, D-mannose, 3-methyl glucose, D-fucose, L-fucose, Inosine, fusidic acid, D-serine, D-sorbitol, D-glucose-6-PO4, D-aspartic acid, minocycline, gelatin, L-pyroglutamic acid, L-serine, pectin, glucuronamide, quinic acid, p-hydroxy-phenylacetic acid, D-lactic acid methyl ester, α-keto-glutaric acid, D-malic acid, bromo-succinic acid, nalidixic acid, potassium tellurite, tween 40,α-hydroxy-butyric acid, β-hydroxy-D, L-butyric acid, α-keto-butyric acid, acetoacetic acid, propionic acid, formic acid, aztreonam, sodium butyrate, sodium bromate
Oxidase	Negative
Catalase	Positive
Major fatty acids	$C_{16:0}$, summed feature 2 ($C_{12:0}$ aldehyde), summed feature 3 ($C_{16:1}\omega7c$ and/or $C_{16:1}\omega6c$) and summed feature 8 ($C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$).
Predominant isoprenoid quinone	Ubiquinone Q-8
Major polar lipids	Phosphatidylethanolamin (PE), diphosphatidylglycerol (DPG) and aminophospholipid (APL)
Known pathogenicity	None

of the closest relatives, *P. rodasii* LMG 26273^T and *P. rwandensis* LMG 26275^T all were 99.0%.

To further discriminate individual strains of *P. endophytica*, REP–PCR was applied according to the procedure described by Rodríguez et al. [38]. The REP–PCR patterns are similar but there are unique bands for the strains, which excluded the possibility of a clonal origin (Supplementary Fig. S1).

The whole genomic DNA of the four new isolates used for genome sequencing was extracted using the TIANamp Bacteria DNA Kit (Tiangen Biotech Beijing CO., Ltd.). Genome sequencing was carried out using Illumina next-generation sequencing technology. Genome assembly and gap filling were conducted on Software SOAPdenovo2.0 and GapCloser1.05 [33]. Average nucleotide identity (ANI) values were calculated using JSpecies [25]. Gene prediction was performed using Glimmer 3.0 [14]; core genes were analyzed by using PGAP [44]; tRNA genes were predicted by using tRNAscan-SE [32] and rRNA genes were detected by using RNAmmer [27]. The genomes of *P. endophytica* sp. nov. strains and one reference strain *P. rodasii* DSM 26611^T were acquired in the present study, mean while that of another type strain of *P. rwandensis* DSM 105076^T and additional seven *Pantoea* sp. strains (AG1095, YR525, GM01, RIT-PI-b, YR343, YR512 and UBA11287), which were included in the ANI and MLSA analyses according to the dendrogram based on genomic BLAST in NCBI website (https://www.ncbi.nlm.nih.gov/genome/?term=Pantoea+sp#), were extracted from the GenBank database. Characters of their genomes were listed in supplementary Table S1. The type strain 596^T had a genome of approximate 5.1Mbp, comprising 4896 predicted genes, 74 tRNA and 11 rRNA genes. The ANI values among the *Pantoea endophytica* sp. nov. strains were 98.2%–98.7% (Supplementary Table S2), meaning all the four strains belong to one species. But the ANI values between the strain 596^T and *P. rodasii* DSM 26611^T and *P. rwandensis* DSM 105076^T were 83.4% and 73.2%, respectively (Supplementary Table S2), which was much lower than the threshold of 95% or 95%–96% for differentiating bacterial species [12,22]. The strain AG1095, one of the seven *Pantoea* sp. strains, share 98.4%–98.6% ANI values with the four strains of *P. endophytica* sp. nov., meaning the strain AG1095 also belong to *P. endophytica* sp. nov. In addition, strain 596^T shared 1990 core genes with *P. rodasii* DSM 26611^T and *P. rwandensis* DSM 105076^T; contained 2855 accessory genes and 625 unique genes; and 2 genes were found in the other genomes, but not in the genome of 596^T (exclusively absent genes) (supplementary Table S1).The assembled genomes of strain 596^T and *P. rodasii* DSM26611^T have been deposited in GenBank database under the accession number PJRT00000000 and PIQI00000000, respectively. While the whole genome sequences data of strains 761, 569YD and 572JC have been deposited in the Genome Warehouse in BIG Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under the accession number GWHAABN00000000, GWHAABM00000000 and GWHAABO00000000 that are publicly accessible at http://bigd.big.ac.cn/gwh.

Colony properties of the type strain 596^T were observed on trypticase soy agar (Becton Dickinson). Cell morphology was examined using scanning electron microscopy (S-3400N, Hitachi) after 25-h-incubation of strain 596^T in LB broth at 30 °C. Cells are approximately 0.7–0.9 µm in width and 1.4–1.7 µm in length (Supplementary Fig. S2). Gram staining was performed by using the bioMérieux Gram Stain kit according to the manufacturer's instructions. Growth on nutrient agar (NA), Luria-Bertani (LB) agar, trypticase soy agar (TSA) and R2A agar was evaluated at 30 °C after incubation for 3 days.Growth was tested at 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50 °C on TSA agar medium. The pH range for growth was determined by measuring the optical densities (wavelength 600 nm) of TSB broth cultures after 3 days. pH was adjusted prior to sterilization to pH 4–10 (at intervals of 1.0 pH unit) using appropriate biological buffers [10,15]: citrate/Na_2HPO_4 buffer, pH 4.0–5.0; phosphate buffer, pH 6.0–7.0; Tris buffer, pH 8.0–9.0 and $NaHCO_3/Na_2CO_3$ buffer, pH 10.0. Verification of the pH after autoclaving revealed only minor changes. Tolerance of NaCl was determined using TSA agar medium adjusted to various NaCl concentrations in the range of 1.0–8.0% (w/v; intervals of 1%). Carbon source utilization, acid production and additional physiological and biochemical characterization were performed using the Biolog GEN III microtest system (Biolog, USA), API 20E, API 50CHE systems and API ZYM systems (bioMérieux, France) according to the manufacturer's instructions. Phenotypic characteristics of strain 596^T and the reference strains are presented in Table 1 and in the Protologue tables.

Cellular fatty acid profiles of the four *P. endophytica* sp. nov. strains were comparatively analyzed with the reference strains *P. rodasii* DSM 26611^T and *P. rwandensis* DSM 105076^T. All the mentioned strains were incubated for 2 days on TSA agar at 30 °C. The cellular fatty acids were saponified, methylated and extracted according to the protocol of the Sherlock Microbial Identification System (MIDI) [39]. The fatty acids were analyzed using a gas chromatography (Hewlett Packard 6890). The fatty acid methyl esters were identified and quantified by using the TSBA 6 database of the Sherlock Microbial Identification System. The major fatty acids of strain 596^T were $C_{16:0}$, summed feature 2 ($C_{12:0}$ aldehyde), summed feature 3 ($C_{16:1}\omega7c$ and/or $C_{16:1}\omega6c$) and summed feature 8 ($C_{18:1}\omega7c$ and/or $C_{18:1}\omega6c$), which was similar to the other three strains 761, 572JC, 569YD and the reference strains *P. rodasii* DSM 26611^T and *P. rwandensis* DSM 105076^T under the same laboratory conditions. The minor differences in the fatty acid profiles among the four strains are shown in Table 2. Polar lipids of the type strains 596^T and the reference strain *P. rodasii* DSM 26611^T were extracted according to the method of Bligh and Dyer [2]. The biomass used for lipids extraction was obtained from cultures growing in LB broth at 30 °C for 2 days. Separation of polar lipids and detection of individual lipids were performed just as described by Gao et al. [19]. Strain 596^T and its closely related species *P. rodasii* DSM 26611^T displayed similar polar lipids compositions (Supplementary Fig. S3). Their polar lipids consisted of phosphatidylethanolamin (PE), diphosphatidylglycerol (DPG), phosphatidylglycerol (PG), unidentified aminophospholipid (APL) and unidentified phospholipid (PL).Respiratory quinones of the type strain 596^T was extracted and purified as described by Collins [13] and was analysed by HPLC [42]. This step was carried out on an Agilent LC-5510 HPLC fitted with a reverse phase column (Diamonsil^{TM} C_{18} 250 × 4.6 mm, 5 µm id) using acetonitrile: isopropyl alcohol 5:3 (v/v) as the eluant, the flow rate was 1.25 mL/min, ubiquinone was detected at 275 nm. The respiratory quinone was ubiquinone-8 (Q-8).

Based upon the above results, the four bacterial strains isolated from maize planted in different geographic regions of northern China can be defined as representatives of a novel species in the genus *Pantoea*. Therefore, *P. endophytica* sp. nov. is proposed. Description of the new species is given in Table 3 with the taxonumber TA00773.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 31870003) and Science & Technology Innovation Project of Beijing Academy of Agricultural and Forestry Sciences (KJCX20170203).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.syapm.2019.06.001.

References

[1] Adeolu, M., Alnajar, S., Naushad, S., Gupta, R.S. (2016) Genome-based phylogeny and taxonomy of the *Enterobacteriales* proposal for *Enterobacterales* ord. nov. divided into the families *Enterobacteriaceae*, *Erwiniaceae* fam. nov. *Pectobacteriaceae* fam. nov. *Yersiniaceae* fam. nov. *Hafniaceae* fam. nov. *Morganellac* fam. nov. and *Budviciaceae* fam. nov. Int. J. Syst. Evol. Microbiol. 66, 5575–5599.

[2] Bligh, E.G., Dyer, J.W. (1959) A rapid method of total lipid extraction and purification. Can. J. Biochem. Physiol. 37, 911–917.

[3] Brady, C.L., Cleenwerck, I., Venter, S., Vancanneyt, M., Swings, J., Coutinho, T. (2008) Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA). Syst. Appl. Microbiol. 31, 447–460.

[4] Brady, C.L., Cleenwerck, I., Venter, S.N., Engelbeen, K., De Vos, P., Coutinho, T.A. (2010) Emended description of the genus *Pantoea*, description of four species from human clinical samples, *Pantoea septica* sp. nov., *Pantoea eucrina* sp. nov., *Pantoea brenneri* sp. nov. and *Pantoea conspicua* sp. nov., and transfer of *Pectobacterium cypripedii* (Hori 1911) Brenner et al. 1973 emend. Hauben et al. 1998 to the genus as *Pantoea cypripedii* comb. nov. Int. J. Syst. Evol. Microbiol. 60, 2430–2440.

[5] Brady, C.L., Cleenwerck, I., van der Westhuizen, L., Venter, S.N., Coutinho, T.A., De Vos, P. (2012) *Pantoea rodasii* sp. nov., *Pantoea rwandensis* sp. nov. and *Pantoea wallisii* sp. nov., isolated from Eucalyptus. Int. J. Syst. Evol. Microbiol. 62, 1457–1464.

[6] Brady, C.L., Goszczynska, T., Venter, S.N., Cleenwerck, I., De, V.P., Gitaitis, R.D., et al. (2011) *Pantoea allii* sp. nov. isolated from onion plants and seed. Int. J. Syst. Evol. Microbiol. 61, 932–937.

[7] Brady, C.L., Venter, S.N., Cleenwerck, I., Engelbeen, K., Vancanneyt, M., Swings, J., Coutinho, T.A. (2009) *Pantoea vagans* sp. nov., *Pantoea eucalypti* sp. nov., *Pantoea deleyi* sp. nov. and *Pantoea anthophila* sp. nov. Int. J. Syst. Evol. Microbiol. 59, 2339–2345.

[8] Brady, C.L., Venter, S.N., Cleenwerck, I., Vandemeulebroecke, K., De Vos, P., Coutinho, T.A. (2010) Transfer of *Pantoea citrea*, *Pantoea punctata* and *Pantoea terrea* to the genus *Tatumella* emend. as *Tatumella citrea* comb. nov., *Tatumella punctata* comb. nov. and *Tatumella terrea* comb. nov. and description of *Tatumella morbirosei* sp. nov. Int. J. Syst. Evol. Microbiol. 60, 484–494.

[9] Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., et al. (2015) RASTtk: a modular and extensible implementation of the rast algorithm for building custom annotation pipelines and annotating batches of genomes. Sci. Rep. 5 (8365), 8365.

[10] Breznak, J.A., Costilow, R.N. (2007) Physicochemical factors in growth. In: Beveridge, T.J., Breznak, J.A., Marzluf, G.A., Schmidt, T.M., Snyder, L.R. (Eds.), Methods for General and Molecular Bacteriology, Third Edition, American Society for Microbiology, Washington, DC, pp. 309–329.

[11] Chen, C., Xin, K., Liu, H., Cheng, J., Shen, X., Wang, Y., et al. (2017) *Pantoea alhagi*, a novel endophytic bacterium with ability to improve growth and drought tolerance in wheat. Sci. Rep. 7, 41564.

[12] Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., Costa, M.S., Rooney, A.P., Yi, H., Xu, X.W., Meyer, S., Trujillo, M.E. (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. Int. J. Syst. Evol. Microbiol. 68 (1), 461–466.

[13] Collins, M.D. (1985) Isoprenoid quinone analysis in classification and identification. In: Goodfellow, M., Minnikin, D.E. (Eds.), Chemical Methods in Bacterial Systematics, Academic Press, London, pp. 267–287.

[14] Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L. (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23, 673–679.

[15] Delory, G.E., King, E.J. (1945) A sodium carbonate-bicarbonate buffer for alkaline phosphatases. Biochem. J. 39, 245.

[16] Dutkiewicz, J., Mackiewicz, B., Kinga Lemieszek, M., Golec, M., Milanowski, J. (2016) *Pantoea agglomerans*: a mysterious bacterium of evil and good. Part III. Deleterious effects: infections of humans, animals and plants. Ann. Agric. Environ. Med. 23, 197–205.

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

[18] Feng, Y., Shen, D., Song, W. (2006) Rice endophyte pantoea agglomerans ys19 promotes host plant growth and affects allocations of host photosynthates. J. Appl. Microbiol. 100, 938–945.

[19] Gao, J.L., Sun, P., Wang, X.M., Lv, F.Y., Mao, X.J., Sun, J.G. (2017) *Rhizobium wenxiniae* sp. nov. an endophytic bacterium isolated from maize root. Int. J. Syst. Evol. Microbiol. 67, 2798–2803.

[20] Gao, J.L., Yuan, M., Wang, X.M., Qiu, T.L., Lv, F.Y., Yang, M.M., Sun, J.G. (2016) *Paenibacillus radicis* sp. nov., a endophytic bacterium isolated from maize root in China. Int. J. Syst. Evol. Microbiol. 66, 801–811.

[21] Gavini, F., Mergaert, J., Beji, A., Mielcarek, C., Izard, D., Kersters, K., De Ley, J. (1989) Transfer of *Enterobacter agglomerans* (Beijerinck 1988) Ewing and Fife 1972 to *Pantoea* gen. nov. as *Pantoea agglomerans* comb. nov. and description of *Pantoea dispersa* sp. nov. Int. J. Syst. Bacteriol. 39, 337–345.

[22] Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J.M. (2007) DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. Int. J. Syst. Evol. Microbiol. 57 (1), 81–91.

[23] Gueule, D., Fourny, G., Ageron, E., Le, F.M.A., Vandenbogaert, M., Grimont, P.A., et al. (2015) *Pantoea coffeiphila* sp. nov. cause of 'potato taste' of Arabica coffee from African great lakes region. Int. J. Syst. Evol. Microbiol. 65, 23–29.

[24] Hegazi, N.A., Hamza, A.M., Osman, A., Ali, S., Sedik, M.Z., Fayez, M. (1998) Modified combined carbon N-deficient medium for isolation, enumeration and biomass production of diazotrophs. In: Kauser, M.A., Sajjad, M.M. (Eds.), Nitrogen Fixation with Nonlegumes, Kluwer Academic Publishers, Dordrecht, pp. 247–253.

[25] Jiménez, G., Urdiain, M., Cifuentes, A., López-López, A., Blanch, A.R., Tamames, J., Kämpfer, P., Kolstø, A., Ramón, D., Martínez, J.F. (2013) Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. Syst. Appl. Microbiol. 36 (6), 383–391.

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

[27] Lagesen, K., Hallin, P., Rodland, E.A., Staerfeldt, H.H., Rognes, T., Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 35 (9), 3100–3108.

[28] Lane, D.J. (1991) 16S/23S rRNA sequencing. In: Stackerandt, E., Goodfellow, M. (Eds.), Nucleic Acid Techniques in Bacterial Systematic, Wiley, Chichester, pp. 115–175.

[29] Lim, J.A., Lee, D.H., Kim, B.Y., Heu, S. (2014) Draft genome sequence of *Pantoea agglomerans* R190, a producer of antibiotics against phytopathogens and foodborne pathogens. J. Biotechnol. 188, 7–8.

[30] Lima, E., Boddey, R.M., Döbereiner, J. (1987) Quantification of biological nitrogen fixation associated with sugar cane using a 15n aided nitrogen balance. Soil Biol. Biochem. 19 (2), 165–170.

[31] Liu, Y., Wang, S., Zhang, D., Wei, S., Zhao, S., Chen, S., Xu, F. (2013) *Pantoea beijingensis* sp. nov., isolated from the fruiting body of *Pleurotus eryngii*. Anton. Leeuw. 104, 1039–1047.

[32] Lowe, T.M., Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic. Acids. Res. 25 (5), 955–964.

[33] Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y. (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1 (1), 1–6.

[34] Mergaert, J., Verdonck, L., Kersters, K. (1993) Transfer of *Erwinia ananas* (synonym, Erwinia uredovora) and *Erwinia stewartii* to the genus *Pantoea* emend. as *Pantoea ananas* (Serrano 1928) comb. nov. and *Pantoea stewartii* (Smith 1898) comb. nov., respectively, and description of *Pantoea stewartii* subsp. *indologenes* subsp. nov. Int. J. Syst. Bacteriol. 43, 162–173.

[35] Mishra, A., Chauhan, P.S., Chaudhry, V., Tripathi, M., Nautiyal, C.S. (2011) Rhizosphere competent *Pantoea agglomerans* enhances maize (*Zea mays*) and chickpea (*Cicer arietinum* L.) growth, without altering the rhizosphere functional diversity. Anton. Leeuw. Int. J. Gen. 100, 405–413.

[36] Palmer, M., Steenkamp, E.T., Coetzee, M.P.A., Avontuur, J.R., Chan, W.Y., Zyl, E.V., Blom, J., Venter, S.N. (2018) *Mixta* gen. nov. a new genus in the *Erwiniaceae*. Int. J. Syst. Evol. Microbiol. 68 (4), 1396–1407.

[37] Prakash, O., Nimonkar, Y., Vaishampayan, A., Mishra, M., Kumbhare, S., Josef, N., et al. (2015) *Pantoea intestinalis* sp. nov. isolated from the human gut. Int. J. Syst. Evol. Microbiol. 65 (10), 3352–3358.

[38] Rodriguez, J.M., Lopez-Romalde, S., Beaz, R., Alonso, M.C., Castro, D., Romalde, J.L. (2006) Molecular fingerprinting of Vibrio tapetis strains using threePCR-based methods: ERIC–PCR, REP–PCR and RAPD. Dis. Aquat. Org. 69, 175–183.

[39] Sasser, M. 1990 Identification of Bacteria by Gas Chromatography of Cellular Fatty Acids, MIDI Technical Note 101, MIDI Inc., Newark, DE.

[40] Tanaka, Y.K., Horie, N., Mochida, K., Yoshida, Y., Okugawa, E., Nanjo, F. (2015) *Pantoea theicola* sp. nov. isolated from black tea. Int. J. Syst. Evol. Microbiol. 65 (10), 3313–3319.

[41] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. (1997) The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids. Res. 25, 4876–4882.

[42] Wu, C., Lu, X., Qin, M., Wang, Y., Ruan, J. (1989) The analysis of menaquinone compound in microbial cells by HPLC. Microbiology [English translation of Microbiology (Beijing)] 16, 176–178.

[43] Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J. (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. Int. J. Syst. Evol. Microbiol. 67 (5), 1613–1617.

[44] Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., Yu, J. (2012) PGAP: pan-genomes analysis pipeline. Bioinformatics 28 (3), 416–418.